

Lille
n° d'ordre
H529

UNIVERSITÉ DE LILLE I

THÈSE

présentée pour obtenir

L'habilitation à diriger des recherches

Spécialité Mathématiques

Par

Assi N'GUESSAN

**Approches statistiques de l'évaluation d'une
mesure : cas de la sécurité routière**

Soutenue le 23 octobre 2006 devant la commission d'examen :

M.	François Bellavance	Rapporteur
M.	Jean-Louis Bon	Examineur
M.	Claude Langrand	Directeur
M.	Sylvain Lassarre	Rapporteur
M.	Michel Roussignol	Rapporteur



Remerciements

Ce mémoire synthétise mon travail de recherche de ces dernières années et de nombreuses personnes physiques et morales ont contribué de loin ou de près à sa réalisation. Je profite donc de ce court espace pour leur exprimer ma gratitude et ma reconnaissance.

Si aujourd'hui je peux présenter ce mémoire c'est grâce au soutien constant de Monsieur le Professeur Claude LANGRAND, qui a dirigé mes recherches et m'a toujours guidé avec une extrême compétence et de grandes qualités, tant humaines que scientifiques, depuis mes premiers balbutiements de chercheur. Il m'a apporté le goût de la recherche et je le remercie très vivement.

Je tiens à remercier Monsieur le Professeur Michel ROUSSIGNOL qui, après ma thèse de doctorat, a encore bien voulu être rapporteur de ce mémoire, pour l'intérêt qu'il continue de porter à mon travail de recherche et pour l'ensemble de ses remarques.

J'adresse mes sincères remerciements à Monsieur le Professeur François BELLAVANCE pour avoir accepté d'être rapporteur de ce mémoire de recherche sur la sécurité routière. Ses remarques et ses données d'accidents m'ont permis d'améliorer certains points de ce travail.

Je tiens à exprimer ma reconnaissance à Monsieur Sylvain LASSARRE, Directeur de recherches à l'INRETS (Institut National de Recherche sur les Transports et leur Sécurité), qui m'a accompagné tout au long de mes travaux de recherche et qui a bien voulu être rapporteur de ce mémoire; qu'il soit remercié pour ses remarques et les idées qu'il m'a données pour les développements futurs de ce travail.

Malgré ses préoccupations scientifiques et ses lourdes charges administratives, Monsieur le professeur Jean-Louis BON a accepté d'être membre du jury de cette soutenance. Je le remercie sincèrement d'avoir accepté de juger ce manuscrit.

Un grand merci aussi à tout le personnel de l'École Polytechnique Universitaire de Lille et à celui du laboratoire de Mathématiques Paul Painlevé qui m'ont donné le confort nécessaire à la réalisation de cette synthèse scientifique.

Messieurs Patrick LE BRETON du SETRA (Service d'Études Techniques des Routes et Autoroutes) et Marcel TRUFFIER du CETE (Centre d'Études Techniques de l'Équipement) Nord-Picardie m'ont constamment soutenu dans la recherche de données d'application et m'ont toujours associé aux différents projets de recherche en sécurité routière. Je tiens à leur dire merci.

Enfin, je ne peux résister au plaisir de remercier mon épouse Joëlle, mes enfants Cécilia, Antoine et Louis, pour la chaleur et l'affection qu'ils n'ont cessé de m'apporter durant tout ce temps consacré à ma recherche.



Résumé

Le but de ce travail de recherche est l'évaluation statistique d'une mesure de sécurité routière par prise en compte de plusieurs sites expérimentaux ; chacun comporte différents types d'accidents liés à la gravité et est associé à une zone de contrôle. Nous construisons, sous certaines hypothèses, des distributions multinomiales permettant de répartir les données d'accidents de chaque site expérimental entre les différents types d'accidents et les périodes avant et après la mise en place de la mesure. Les probabilités de classe de chacune de nos distributions multinomiales combinent les données des sites expérimentaux et des zones de contrôle ; elles prennent en compte le paramètre effet moyen de la mesure, les poids aléatoires utilisés pour pondérer les coefficients de contrôle ainsi que la tendance moyenne des variations dans la zone de contrôle de chaque site. L'estimation des paramètres liés à nos modèles statistiques est établie en utilisant les propriétés du maximum de vraisemblance sous contraintes. Par la suite, nous étudions l'existence et la convergence théoriques ainsi que numériques des estimateurs du maximum de vraisemblance sous contraintes de l'effet moyen et des risques d'accidents. L'étude de l'existence, de la convergence ainsi que de la distribution asymptotique des estimateurs passe par l'inversion des matrices d'informations associées à nos modèles d'analyse et par l'obtention des éléments diagonaux de leurs inverses. Or, ces matrices d'information sont généralement très creuses et leurs dimensions sont des fonctions linéaires croissantes du nombre des sites expérimentaux et des types d'accidents. Leur inversion directe par les procédures classiques pose souvent des problèmes d'instabilité numérique et on n'est pas certain de la qualité des inverses numériques obtenus. Nous proposons une approche par complément de Schur pour contourner ces difficultés et estimer par la même occasion la variabilité des paramètres des modèles en calculant directement certains éléments d'une matrice de variances-covariances. L'utilisation de la méthode du complément de Schur permet également de vérifier de façon formelle la condition d'inversibilité des matrices d'information et d'obtenir une expression formelle de la variance des estimateurs ; ceci permet l'étude de leur comportement asymptotique ainsi que la construction d'intervalles de confiance et de tests d'hypothèse relatifs à la nullité de l'efficacité moyenne de la mesure. Nous appliquons ces résultats théoriques à des aménagements routiers effectués sur certaines routes nationales de la région Nord Pas de Calais ainsi qu'à l'évaluation de l'impact sur les nombres d'accidents de l'installation de panneaux-réclame le long de certaines autoroutes de Montréal. Enfin, nous étudions et obtenons la loi asymptotique de certaines statistiques utilisées pour éprouver l'hypothèse d'égalité des effets moyens par site ou par type d'accidents. Les conclusions et perspectives sont structurées autour de l'utilisation de nouvelles méthodes d'estimation, de l'introduction de nouvelles familles de modèles multivariés et de la recherche de solutions initiales adaptées aux données de nos approches statistiques de l'évaluation d'une mesure.

Table des matières

Remerciements

Résumé

1 Introduction générale	5
2 Modèles et estimation sous contraintes	9
2.1 Introduction	9
2.2 Données et Objectifs	11
2.2.1 Structure des données	11
2.2.2 Objectifs principaux et formulation du problème	12
2.3 Modèles statistiques	13
2.3.1 Modèle 1	13
2.3.2 Modèle 2	14
2.3.3 Modèle 3	15
2.4 Estimation multivariée sous contraintes	16
2.4.1 Introduction	16
2.4.2 Résultats asymptotiques généraux	17
3 Complément de Schur et approximations d'erreurs standards	19
3.1 Introduction	19
3.2 Complément de Schur	20
3.2.1 Définition et premières propriétés	20
3.2.2 Quelques applications à l'inversion formelle de matrices	20
3.3 Étude des matrices d'information liées aux modèles d'évaluation	21
3.3.1 Structure générale des matrices d'information	21
3.3.2 Étude de l'inversibilité des matrices d'information	22
3.3.3 Application à l'estimation des écart-types de l'effet moyen et des différents types de risque d'accidents	25
3.3.4 Construction d'intervalles de confiance et de tests d'hypothèse	28
4 Quelques exemples d'application	29
4.1 Introduction	29
4.2 Données simulées	30
4.2.1 Principe de simulation des données d'accident avant-après avec contrôle	30
4.2.2 Estimation et convergence numériques de solutions	30
4.2.3 Analyse des lois de probabilité liées aux différents risques d'accidents	32

4.3 Quelques études de cas réels	33
4.3.1 Notation et modèle	33
4.3.2 Estimation de l'effet moyen et des différentes lois de probabilité	34
4.3.3 Données de Tanner 1958	35
4.3.4 Évaluation d'aménagements de sécurité sur certaines routes nationales	36
4.4.5 Evaluation de l'effet de panneaux-réclame le long de certaines autoroutes	40
5 Conclusions et perspectives	43
5.1 tests d'hypothèse d'égalité des effets.....	43
5.1.1 Introduction	43
5.1.2 Modèles statistiques	44
5.1.3 Statistiques de test et loi limite	45
5.2 Estimation à l'aide de la méthode du chi-deux minimum modifié	47
5.2.1 Principe de la méthode	47
5.2.2 Résultats asymptotiques	48
5.3 Quelques perspectives	49
Annexe A01 Statistics Vol. 39, No. 4, (2005) 303-314	51
Annexe A02 Computational Statistics and Data Analysis, 00, 000-000 (2006)	65
Annexe A03 J. of Comput. and Applied Mathematics, 177, 331-345 (2005)	85
Annexe A04 C. R. Math. Rep. Acad. Sci. Canada vol. 27, (3), 2005 pp. 84-91	101
Annexe A05 Rev. Statistique Appliquée XLIX (2), 83-100 (2001)	110
Annexe A06 Proceedings of the fourth International Conference, Eldon Y. Li Series Editor, Cal. Poly. State University, Vol. 4, 243-252, (2005)	130
Annexe A07 Comptes Rendus Acad. Sci. Paris, Série I, 337, 219-222, (2003)	142
Annexe A08 Comptes Rendus Acad. Sci. Paris, Série I, 317, 401-404,(1993)	148
Annexe A09 Technical Report, Center for Research on Transportation, Université de Montréal, CRT-2003-12, (2003)	154
Annexe A10 Technical Report, IRMA, Vol. 66, No. III (2006)	174
Annexe A11 Technical Report, IRMA, Vol. 39, No. XI (1996)	195
Annexe A12 Technical Report, IRMA, Vol. 47, No. VII (1998)	208
Références bibliographiques	216

Chapitre 1

Introduction générale

Le but de ce travail de recherche est l'évaluation statistique d'une mesure de sécurité routière par prise en compte de différents types d'accidents liés ou non à la gravité. Dans ce domaine de l'évaluation d'une mesure de sécurité routière où les approches statistiques n'ont pas un statut méthodologique bien établi [8, 10, 11, 16, 27, 30, 34, 35, 43, 49, 50, 62, 79, 84, 103, 106], une méthode et trois questions principales ont motivé ce travail : la généralisation du modèle statistique de Tanner[103] à des tableaux de données intégrant la dimension type d'accidents ; l'estimation de l'effet moyen de la mesure, de différents types de risque d'accidents et de leur variabilité ; la comparaison des effets moyens dans les sites expérimentaux et pour les différents types de risque d'accidents.

La méthode statistique mise en place par Tanner est la plus ancienne, la plus couramment citée et utilisée. Elle consiste, en considérant un seul type d'accidents, à combiner les effectifs observés avant et après la mise en place de la mesure de sécurité, relevés sur plusieurs sites expérimentaux à ceux provenant de plusieurs zones associées à ces sites expérimentaux. Cette combinaison se fait au moyen d'un modèle et de tests paramétriques sous diverses hypothèses de travail. En particulier, une des hypothèses de base effectuée par cet auteur consiste à supposer que les coefficients utilisés dans son modèle et obtenus à partir de rapport de nombres d'accidents observés dans les zones de contrôle sont fixes et sans erreur. Or indépendamment de la mesure, des facteurs extérieurs (conditions météorologiques, trafic moyen, vitesse moyenne, comportement des usagers de la route, etc) concourent à la réalisation et à l'aggravation des accidents de la route. L'hypothèse qui consiste alors à supposer que les coefficients de contrôle sont fixes et sans erreur semble difficile à tenir car ces coefficients sont calculés à l'aide de nombres d'accidents soumis aussi à des influences aléatoires et que même si la mesure n'est pas directement appliquée dans les zones de contrôle, elle peut conduire à des modifications de comportement : la vitesse moyenne baisse localement aux alentours des radars fixes alors qu'elle augmente sensiblement quelques kilomètres plus loin. Une meilleure connaissance de la structure (distribution de probabilité, tendance moyenne) des coefficients des zones de contrôle permettrait de mieux modéliser les données d'accidents des sites expérimentaux, de mieux évaluer l'effet de la mesure et devrait permettre une bonne appréciation des risques d'accidents dans les zones de contrôle ainsi que dans les sites expérimentaux.

Nous avons donc cherché à la fois à étendre le modèle de Tanner à plusieurs sites expérimentaux chacun comportant plusieurs types d'accidents et à lever l'hypothèse effectuée sur les coefficients de contrôle. Ceci implique donc de généraliser le modèle de Tanner à de grands tableaux (section 2.2) dans lesquels on peut lire, pour chacun des sites expérimentaux et pour chaque type d'accidents, les effectifs observés avant et après la mise en place de la mesure et de supposer que les coefficients de contrôle sont l'objet d'une certaine influence aléatoire dont il faut tenir compte. Pour cela, nous avons utilisé le modèle multinomial, généralisation naturelle du modèle binomial de Tanner, et considéré que la mesure avait un effet multiplicateur sur la moyenne des accidents (on conçoit aisément dans le domaine de la sécurité routière qu'une mesure tend à faire baisser la variable aléatoire d'un certain pourcentage et non pas d'un certain incrément), ainsi nous avons obtenu trois généralisations (section 2.3, annexes A02 p.65, A04 p.101 et A05 p. 110) du modèle de base de Tanner.

Les deux premières distributions multinomiales ainsi construites permettent de répartir les données d'accidents de chaque site expérimental entre les différents types d'accidents et les périodes avant et après. L'hypothèse concernant les coefficients de contrôle a été levée en pondérant ces coefficients, pour une zone de contrôle fixée, par une loi de probabilité discrète inconnue dont les caractéristiques sont utilisées pour construire les probabilités de classe de nos distributions multinomiales. Ces probabilités de classe combinent les données des sites expérimentaux et des zones de contrôle et prennent en compte le paramètre effet moyen de la mesure, la distribution de probabilité associée aux coefficients de contrôle ainsi que la tendance moyenne des variations dans les zones de contrôle de chaque site. Ces résultats font l'objet des annexes A02 p. 65 et A05 p.110 (voir aussi les paragraphes 2.3.1 et 2.3.2 de la section 2.3). Une approche classique de la modélisation d'une mesure de sécurité routière est de considérer le logarithme de l'effet moyen plutôt que l'effet moyen. En utilisant ce principe, nous avons construit une fonction bijective qui transforme l'effet en son logarithme et laisse invariant les autres paramètres de nos modèles. Ainsi, nous avons obtenu un troisième modèle (annexe A04 p. 101 et paragraphe 2.3.3) dont la structure formelle s'apparente aux modèles multinomiaux à lien logistique.

Le problème de l'estimation des paramètres liés aux modèles reste la question centrale des approches que nous proposons. Or, il s'avère que les modèles statistiques obtenus comportent, par construction, deux ensembles de paramètres. Un ensemble de dimension un relatif à l'effet moyen de la mesure et un autre ensemble, de dimension strictement supérieure à un, décrivant les différentes lois de probabilité utilisées pour pondérer les coefficients de contrôle de chaque zone. Par conséquent le vecteur paramètre associé à chaque modèle est soumis à des contraintes. De ce fait, la question cruciale de l'estimation de l'effet moyen de la mesure et des risques d'accidents revient à un problème d'optimisation sous contraintes. Nous avons utilisé les propriétés (voir paragraphe 2.4.2) de l'estimateur du maximum de vraisemblance sous contraintes pour répondre à cette question. Nous avons ainsi adapté les résultats généraux proposés par Crowder ([20]) au cas de la sécurité routière.

L'existence et la convergence théoriques des estimateurs du maximum de vraisemblance sous contraintes de l'effet moyen et des risques d'accidents sont ainsi étudiées (voir annexe A11 p. 195).

Pour étudier la distribution asymptotique des estimateurs, il est nécessaire d'estimer la variabilité de l'effet moyen et des risques d'accidents. Cela passe par l'obtention des éléments diagonaux de l'inverse de la matrice d'information de Fisher. On note que les matrices d'information de Fisher associées à nos modèles d'analyse sont généralement très creuses et leurs dimensions sont des fonctions linéaires croissantes du nombre de sites expérimentaux et de types d'accidents. Leur inversion directe par les procédures classiques pose souvent des problèmes d'instabilité numérique et on n'est pas certain de la qualité des inverses numériques obtenus.

Nous proposons une approche par complément de Schur (voir par exemple [28], [80] [95],[100] [110]) pour contourner ces difficultés et estimer par la même occasion la variabilité de l'effet moyen et des risques d'accident en calculant directement certains éléments d'une matrice de variances covariances. Il y a trois avantages à l'utilisation de la méthode du complément de Schur, le premier est qu'elle permet de vérifier de façon formelle la condition d'inversibilité des matrices d'information, essentielle pour établir l'existence du maximum de vraisemblance sous contraintes, le deuxième est d'éviter numériquement l'inversion de la matrice d'information de Fisher, le troisième est d'obtenir une expression formelle de la variance des estimateurs permettant l'étude de leurs comportements ainsi que la construction d'intervalles de confiance. Nous utilisons pour cela certaines propriétés (paragraphe 3.2.2) du complément de Schur d'une matrice et celles du produit de Kronecker de deux matrices.

Les résultats algébriques obtenus sont exposés au paragraphe 3.3.2 (voir aussi annexe A03 p. 85). Ces résultats sont ensuite utilisés pour déduire la structure générale de la distribution asymptotique (annexes A07 p.142 et A09 p. 154) de l'estimateur du maximum de vraisemblance sous contraintes de l'effet moyen et des risques d'accident. En utilisant toujours le complément de Schur, nous exploitons en profondeur cette structure générale pour obtenir l'expression formelle explicite des variances et des covariances asymptotiques (paragraphe 3.3.3 et annexe A01 p. 51) des composantes de l'estimateur sous contraintes. Ces expressions analytiques sont alors utilisées pour construire (paragraphe 3.3.4 et annexe A04 p. 101) des intervalles de confiance et des tests d'hypothèse relatifs à la nullité de l'efficacité moyenne de la mesure.

Quelques exemples d'application de nos travaux de recherche sur l'évaluation font l'objet des annexes A02 p. 65, A05 p. 110, A06 p. 130 et A10 p. 174. Les annexes A02 et A05 traitent des aspects pratiques de l'existence et de la convergence des estimateurs en utilisant des données simulées. L'objectif visé dans ces deux premiers travaux est de soumettre notre approche d'estimation à différents tableaux de données d'accidents et d'étudier le comportement de la solution des équations de vraisemblance sous contraintes. Dans un premier temps, nous nous sommes intéressés à l'estimation de l'effet moyen et des tendances moyennes des variations en nombre d'accidents dans les zones de contrôle (paragraphe 4.2.2 et annexe A05). Dans un deuxième temps, nous avons étudié numériquement l'existence et la convergence des lois de probabilité utilisées pour pondérer les coefficients de contrôle de chaque zone. Pour ce faire, nous avons utilisé une approche graphique comparant les histogrammes des différentes lois de probabilité en nous servant de plusieurs tableaux de nombres d'accidents simulés (voir paragraphe 4.2.3 et annexe A02). L'annexe A06 représente la première grande application de nos approches à des données réelles. En effet, nous montrons comment analyser les données liées à la transformation de sept carrefours en ronds-points traitée par Tanner. Nous appliquons notre principe de

pondération aléatoire des coefficients de contrôle aux données de ce dernier auteur et obtenons les mêmes résultats sans être obligé de recourir à une procédure de test pour estimer l'écart-type de l'effet moyen. Dans l'annexe A10 (voir paragraphes 4.3.4 et 4.3.5), nous présentons trois études de cas réels et récents. Les deux premières applications proviennent d'une série d'opérations d'aménagements de sécurité effectués sur les routes nationales du département Pas de Calais (France). Ces opérations visent à résoudre les problèmes de sécurité sur le réseau existant, en s'attachant à traiter en priorité les zones repérées comme les plus dangereuses. La dernière application est relative à l'évaluation de l'effet de panneaux publicitaires installés le long de certains échangeurs et de certaines autoroutes de la ville de Montréal (Québec, Canada). L'ensemble des exemples d'application montre bien que nos approches statistiques de l'évaluation s'appliquent à la fois à des tableaux de données d'accidents de petits effectifs et de grands effectifs.

Nous avons structuré les conclusions et perspectives autour de trois grands points. Dans un premier point, nous abordons la question récurrente en évaluation statistique d'une mesure appliquée à plusieurs sites comportant plusieurs types : a-t-on le même effet pour tous les sites ou tous les types d'accident ?

Cette question revient à tester l'hypothèse d'égalité des effets moyens par site ou par type d'accident. Nous proposons des modèles et des statistiques (paragraphe 5.1) pour éprouver cette hypothèse. La loi asymptotique de ces statistiques est obtenue (voir annexe A08 p. 148 et paragraphe 5.1.3) sous certaines conditions de régularité grâce à l'étude de certaines formes quadratiques aléatoires. Différentes statistiques du type chi-deux sont ainsi déduites et utilisées pour éprouver l'homogénéité de l'efficacité moyenne de la mesure. Les distributions asymptotiques obtenues montrent clairement la place importante occupée par la statistique du chi-deux dans les approches d'évaluation. Nous proposons, dans un deuxième point, une deuxième approche d'estimation (paragraphe 5.2.2 et annexe A12 p. 208) basée sur la méthode du chi-deux minimum modifié (voir par exemple [17] p. 425-426, [36]). Dans un troisième point (voir paragraphe 5.1.3), nous présentons quelques perspectives d'amélioration de nos résultats actuels. Ces perspectives vont de la méthode d'estimation à l'introduction de nouvelles familles de modèles multivariés en passant par la recherche de solution initiale adaptée aux données de nos approches statistiques de l'évaluation d'une mesure.

Chapitre 2

Modèles et estimation sous contraintes

2.1 Introduction

Les modèles statistiques utilisés dans l'évaluation de l'efficacité d'une mesure de sécurité routière dépendent fortement des données disponibles et des objectifs fixés lors de la mise en place de cette mesure [2, 4, 5, 6, 7, 8, 16, 22, 23, 37, 40-48, 51, 57-62]. Quel que soit le modèle, le principe de base de l'évaluation consiste à prédire la valeur moyenne de la variable aléatoire utilisée (ou critère d'appréciation) si la mesure n'avait pas été appliquée. En pratique, on utilise souvent le nombre d'accidents comme variable aléatoire.

Pour un seul site par exemple, on compte les nombres d'accidents survenus avant l'application de cette mesure (période t_1) et après celle-ci (période t_2) et il faut pouvoir juger si la modification du nombre des accidents (passant par exemple de n_{t_1} à n_{t_2}) est significativement différente de celle qu'on aurait observée si la mesure n'avait pas été appliquée ou n'était pas efficace. L'interprétation des résultats doit tenir compte d'un certain nombre de facteurs (évolution du trafic, conditions météorologiques, comportement des usagers, ...); si en étudiant d'autres données (par exemple les variations du nombre des accidents sur un site équivalent non soumis à la mesure) on s'aperçoit que le nombre d'accidents est multiplié par un facteur c (coefficient de contrôle, ou d'évolution), la modélisation de l'efficacité de la mesure doit tenir compte de ce facteur c .

Lorsqu'une mesure est appliquée simultanément à plusieurs sites (supposés être un échantillon représentatif de l'ensemble des sites), il n'est pas en général réaliste de considérer, pour des raisons de localisation, ou de similarité des sites ([44]), ou de régression vers la moyenne [1, 15, 23, 24, 55, 101, 107, 108] par exemple, que le même facteur c puisse être utilisé quel que soit le site. C'est pourquoi Tanner (1958) a introduit la notion de zone de contrôle et défini un coefficient de contrôle pour chaque site expérimental à l'aide du quotient nombre des accidents "après" sur nombre des accidents "avant" dans la zone de contrôle. Par suite, il a proposé un modèle binomial pour combiner les données d'accidents de chaque site expérimental avec celles de la zone de contrôle qui lui est associée. Son modèle dépend non seulement du paramètre décrivant l'effet de la mesure au site expérimental considéré mais aussi du coefficient de contrôle spécifique à ce site. Or une hypothèse fondamentale de travail effectuée par Tanner est le fait

qu'il suppose que ces dits coefficients de contrôle, obtenus à l'aide de rapports de nombres d'accident, sont sans erreur et soumis à aucune variation aléatoire.

Dans cette partie de notre travail de recherche, nous continuons d'utiliser cette approche avant-après avec zone de contrôle mais supposons que les coefficients de contrôle sont soumis à des aléas. Nous proposons une approche probabiliste pour prendre en compte ces aléas en construisant des distributions de probabilité utilisées pour pondérer les différents coefficients de contrôle. Les poids ainsi attribués permettent de moduler la variabilité entre les coefficients de contrôle et de construire des modèles statistiques multivariés applicables à des tableaux avant-après sans agréger les données qui correspondent à différents types d'accidents. On considère dans ce mémoire qu'une mesure de sécurité routière est appliquée à plusieurs sites et que l'on dispose du nombre total d'accidents pour chaque type et pour chaque site expérimental au cours des deux périodes d'analyse (période précédant et période suivant l'instauration de la mesure). Une telle démarche induit une troisième dimension (typologie d'accidents) dans l'approche avant-après d'une mesure de sécurité routière et impose de modéliser l'effet moyen de cette mesure par rapport à des tableaux de données d'accidents en incorporant la modélisation de différents types de risque d'accidents. Les différents modèles obtenus (annexes A02, p. 65 A04 p. 101, A05 p. 110) sont bâtis sur le principe du modèle multinomial et permettent d'estimer à la fois l'effet moyen de la mesure et des indices de risques associés aux différents types d'accident de l'étude considérée.

Ainsi, nous construisons plusieurs modèles statistiques multivariés (voir [9], [70], [79-89]) permettant de combiner des tableaux de données d'accidents avant-après avec contrôle et proposons différentes généralisations du modèle d'évaluation de Tanner. Au paragraphe 2.3.1 (voir annexe A05 p. 110), nous exposons le premier modèle multinomial dont la dimension est égale à deux (les deux périodes d'analyse de la mesure) fois le nombre de types d'accidents retenus. Ce modèle prend en compte le paramètre effet moyen de la mesure à tous les sites, le coefficient spécifique à chaque type d'accidents de la zone de contrôle, le poids aléatoire associé à chaque coefficient de contrôle ainsi que la tendance moyenne des variations en nombre d'accidents dans la zone de contrôle associée à chaque site expérimental. Ce premier modèle utilise directement le coefficient de contrôle spécifique à chaque type d'accident dans la construction de la loi multinomiale permettant de répartir le nombre total d'accidents de chaque site expérimental. Or, ces coefficients individuels peuvent être souvent instables et soumis à d'éventuelles variabilités. De ce fait, nous construisons un deuxième modèle (paragraphe 2.3.2). Ce deuxième modèle évite l'utilisation directe des coefficients individuels et prend plutôt en compte la tendance moyenne des variations en nombre d'accidents des zones de contrôle. Ces tendances moyennes ont l'avantage d'être plus stables.

La construction de ces deux modèles multinomiaux est basée sur une approche de répartition aléatoire du nombre total d'accidents de chaque site expérimental entre d'une part les deux périodes d'analyse et d'autre part les différents types d'accidents. Les lois de probabilité des deux modèles statistiques dépendent, par construction, d'un vecteur paramètre constitué de l'effet moyen de la mesure et des poids associés aux différents coefficients de chaque zone de contrôle. De ce fait, l'espace des paramètres est constitué du produit de la demi droite réelle positive avec des simplexes. L'estimation de ce vecteur paramètre est une estimation sous

contraintes faite à l'aide de la méthode du maximum de vraisemblance. Outre les propriétés asymptotiques bien connues du maximum de vraisemblance, on montre en sécurité routière (voir par exemple [103]) qu'il est préférable d'estimer le logarithme de l'effet moyen plutôt que l'effet moyen lui même. En utilisant ces résultats asymptotiques, on modifie le vecteur paramètre à l'aide de l'application qui laisse invariant les poids et transforme l'effet moyen en son logarithme. Cette transformation bijective permet d'obtenir un troisième modèle exposé dans le paragraphe 2.3.3 (annexe A04 p. 85) dont la structure s'apparente à celle des modèles à lien logistique ([52], [68]).

2.2 Données et objectifs

On considère donc qu'une mesure de sécurité routière (aménagement d'un carrefour, revêtement d'un tronçon d'autoroute, ...) a été appliquée à $S(S > 0)$ sites. On suppose disposer, sur chacun des S sites expérimentaux, du nombre total d'accidents avant et après l'imposition de la mesure pour des périodes de longueurs convenables. La question qui est posée ici est la suivante : comment modéliser et estimer l'effet moyen de la mesure relative à l'ensemble des S sites, dans le cas où chacun d'eux comporte $R(R > 1)$ types d'accident ?

2.2.1 Structure des données

Les données dont on dispose se trouvent dans le tableau 1 suivant :

	type 1	...	type j	...	type R	type 1	...	type j	...	type R	total
site 1	x_{111}	...	x_{1j1}	...	x_{1R1}	x_{211}	...	x_{2j1}	...	x_{2R1}	n_1
⋮	⋮		⋮		⋮	⋮		⋮		⋮	⋮
site k	x_{11k}	...	x_{1jk}	...	x_{1Rk}	x_{21k}	...	x_{2jk}	...	x_{2Rk}	n_k
⋮	⋮		⋮		⋮	⋮		⋮		⋮	⋮
site S	x_{11S}	...	x_{1jS}	...	x_{1RS}	x_{21S}	...	x_{2jS}	...	x_{2RS}	n_S

table 1 : Sites expérimentaux

où x_{1jk} (resp. x_{2jk}) représente le nombre d'accidents du type j enregistré au site expérimental k avant (resp. après) la mise en place de la mesure, $n_k = \sum_{j=1}^R (x_{1jk} + x_{2jk})$ le nombre total d'accidents observés au site $k(k = 1, 2, \dots, S)$.

Des facteurs (comme par exemple la tendance générale du trafic routier, les variations des limitations de vitesse, les conditions météorologiques, ...) ainsi que le comportement des usagers (en particulier des conducteurs vis à vis de la mesure : changement d'itinéraire de conduite par exemple) peuvent avoir une influence directe ou indirecte sur l'effet visible de la mesure prise. Pour prendre donc en compte ces facteurs extérieurs sur les données d'accident, on suppose qu'à chaque site expérimental est associée une zone (ou site) de contrôle où la mesure n'est pas directement appliquée. On suppose alors que les nombres d'accidents pour les sites de contrôle et pour les mêmes périodes avant-après sont connus, disponibles pour l'ensemble des types d'accident sous forme de coefficients de contrôle fournis par la table 2 suivante :

	type 1	...	type j	...	type R
contrôle 1	c_{11}	...	c_{j1}	...	c_{R1}
⋮	⋮		⋮		⋮
contrôle k	c_{1k}	...	c_{jk}	...	c_{Rk}
⋮	⋮		⋮		⋮
contrôle S	c_{1S}	...	c_{jS}	...	c_{RS}

table 2 : zones ou sites de contrôle

où pour une zone de contrôle fixée, c_{jk} est le rapport du nombre d'accidents du type j de la période après sur celui de la période avant.

2.2.2 Objectifs principaux et formulation du problème

Comment modéliser et estimer l'effet moyen de la mesure de sécurité routière sur différents types d'accidents, et prendre formellement en considération les informations (nombres d'accidents) dans les zones de contrôle sans agréger systématiquement les données d'accidents disponibles ?

L'approche que nous proposons consiste à combiner les données des sites expérimentaux et des zones de contrôle en utilisant le principe des modèles multiplicatifs [48, 90-92, 93] souvent adopté en sécurité routière. Cette approche permet la prise en compte du paramètre effet moyen de la mesure de sécurité routière ainsi que la construction, pour chaque site expérimental k , d'un vecteur paramètre de dimension R . Chaque vecteur paramètre de dimension R ainsi construit est alors utilisé pour pondérer les coefficients de la zone de contrôle associée au site expérimental k . De ce fait, au problème de modélisation et d'estimation de l'effet moyen s'adjoint celui de S vecteurs de probabilités (ou de poids) chacun de dimension R .

Le premier objectif est donc de construire des modèles statistiques paramétriques qui combinent simultanément les données d'accidents des tables 1 et 2, l'effet moyen de la mesure et les différents vecteurs de poids. Le deuxième objectif est d'estimer l'ensemble des paramètres liés aux modèles afin de quantifier l'ampleur de l'effet de la mesure et de son incidence sur les différents types de risques d'accident présents dans l'étude.

Par la suite, on note X_{1jk} (resp. X_{2jk}) la variable aléatoire donnant l'effectif du nombre d'accidents du type j sur le site expérimental k avant (resp. après) l'implémentation de la mesure. On s'intéresse alors, pour un site expérimental k fixé, à la loi de probabilités du vecteur aléatoire discret $(X_{1k}, X_{2k})^T$ de dimension $2R$, étant donné n_k le nombre total d'accidents enregistrés à ce site, avec

$$X_{1k} = (X_{11k}, X_{12k}, \dots, X_{1Rk})^T \text{ (resp. } X_{2k} = (X_{21k}, X_{22k}, \dots, X_{2Rk})^T)$$

le vecteur aléatoire de dimension R donnant les effectifs d'accidents, avant (resp. après) la mesure.

2.3 Modèles statistiques

Nous exposons partiellement, dans cette section, le mécanisme de construction des modèles statistiques utilisés pour répartir l'effectif total d'accidents de chaque site expérimental à la fois entre les deux périodes et entre les différents types d'accidents. le principe de base de cette construction est présentée dans l'annexe A05 p 110.

2.3.1 Modèle 1 (annexe A05)

Pour modéliser l'effet moyen de la mesure de sécurité, on suppose que le nombre moyen $E(X_{2jk})$ d'accidents du type j au site k qu'on pourrait s'attendre à observer dans la période "après" si ce site expérimental avait un comportement semblable à celui de la zone de contrôle associée est :

$$\theta^* c_{jk} E(X_{1jk}) \tag{1}$$

où $E(X_{1jk})$ est l'espérance mathématique de la variable aléatoire X_{1jk} et

$$\theta^* = \left[\sum_{k=1}^S \sum_{j=1}^R E(X_{2jk}) \right] / \left[\sum_{k=1}^S \sum_{j=1}^R c_{jk} E(X_{1jk}) \right].$$

Sous ces mêmes hypothèses et en utilisant les données d'accidents des tables 1 et 2, et en estimant $E(X_{1jk})$ par x_{1jk} , on déduit que les proportions d'accidents espérés par type au site expérimental k au cours des deux périodes (avant et après) sont estimées par :

$$\frac{p_{1k}^*}{1 + \theta^* c_k^*}, \frac{p_{2k}^*}{1 + \theta^* c_k^*}, \dots, \frac{p_{Rk}^*}{1 + \theta^* c_k^*}, \frac{\theta^* c_{1k} p_{1k}^*}{1 + \theta^* c_k^*}, \frac{\theta^* c_{2k} p_{2k}^*}{1 + \theta^* c_k^*}, \dots, \frac{\theta^* c_{Rk} p_{Rk}^*}{1 + \theta^* c_k^*} \tag{2}$$

en posant

$$p_{mk}^* = x_{1mk} / \left(\sum_{j=1}^R x_{1jk} \right), \quad c_k^* = \sum_{m=1}^R c_{mk} p_{mk}^*.$$

p_{jk}^* est une estimation de la probabilité pour qu'un accident qui se produit au site expérimental k avant la mise en place de la mesure soit du type j , avec

$$p_{jk}^* > 0, \quad \sum_{j=1}^R p_{jk}^* = 1, \quad (k = 1, 2, \dots, S),$$

et c_k^* , pour k fixé, est la somme des valeurs c_{jk} de la table 2 pondérées par les p_{jk}^* . c_k^* représente une estimation du coefficient de contrôle moyen associé au site k si ce site se comportait comme sa zone de contrôle associée.

On pose alors

$$\pi_{1jk}^* = \frac{p_{jk}^*}{1 + \theta^* c_k^*}, \quad \pi_{2jk}^* = \frac{\theta^* c_{jk} p_{jk}^*}{1 + \theta^* c_k^*}, \quad j = 1, 2, \dots, R. \tag{3}$$

Par construction, les π_{tjk}^* ($t = 1, 2$) représentent la répartition du nombre total d'accidents espérés au site expérimental k en $2R$ catégories mutuellement exclusives entre les périodes "avant" et "après" si le site k se comporte comme sa zone de contrôle et si θ^* est l'effet moyen de la mesure. On a

$$\sum_{j=1}^R \sum_{t=1}^2 \pi_{tjk}^* = 1. \quad (4)$$

En s'inspirant de ces deux dernières relations, on fait l'hypothèse que, pour k fixé, le vecteur $(X_{1k}, X_{2k})^T$ suit une loi multinomiale notée $\mathcal{M}(n_k; \Pi_{1k}(\Theta), \Pi_{2k}(\Theta))$, où

$$\Pi_{tk}(\Theta) = (\pi_{t1k}(\Theta), \pi_{t2k}(\Theta), \dots, \pi_{tRk}(\Theta))$$

est le vecteur de dimension R de composantes $\pi_{tjk} : \mathbb{R}^{1+sr} \mapsto]0; 1[$ données par :

$$\pi_{1jk}(\Theta) = \frac{p_{jk}}{1 + \theta \sum_{m=1}^R c_{mk} p_{mk}} ; \quad \pi_{2jk}(\Theta) = \frac{\theta c_{jk} p_{jk}}{1 + \theta \sum_{m=1}^R c_{mk} p_{mk}}, \quad (j = 1, 2, \dots, R), \quad (5)$$

où $\Theta = (\theta, P^T)^T$ de dimension $(1 + SR)$ est le vecteur des paramètres, θ ($\theta > 0$) le paramètre effet moyen de la mesure et où $P = (P_1^T, P_2^T, \dots, P_S^T)^T$ avec $P_k = (p_{1k}, p_{2k}, \dots, p_{Rk})^T$, vecteur de dimension R est tel que :

$$P_k \in \mathcal{S}_k^{(R-1)} = \{(p_{1k}, p_{2k}, \dots, p_{Rk})^T \in \mathbb{R}^R, p_{jk} > 0, \sum_{j=1}^R p_{jk} = 1\}. \quad (6)$$

2.3.2 Modèle 2 (annexe A02)

Bien que le modèle 1 soit une extension considérable de celui de Tanner (1958) à des types d'accidents, il comporte, à travers la structure du paramètre π_{2jk} , les coefficients individuels c_{jk} supposés fixes et connus. Mais, les données d'accidents dans les zones de contrôle sont également aléatoires. Et comme précédemment notifié, il peut exister une variabilité entre les valeurs c_{jk} d'un même site expérimental et le fait de ne pas corriger cette variation dans la modélisation peut avoir une influence sur l'estimation des paramètres. Pour améliorer la conception du modèle 1 et proposer une alternative, on modifie le paramètre π_{2jk} en remplaçant le coefficient individuel c_{jk} par la tendance moyenne c_k . On en déduit un deuxième modèle multinomial en supposant, pour k fixé, que le vecteur $(X_{1k}, X_{2k})^T$ suit une loi multinomiale notée $\mathcal{M}(n_k; \Pi_{1k}(\Theta), \Pi_{2k}(\Theta))$, où

$$\Pi_{tk}(\Theta) = (\pi_{t1k}(\Theta), \pi_{t2k}(\Theta), \dots, \pi_{tRk}(\Theta))$$

est le vecteur de dimension R de composantes $\pi_{tjk} : \mathbb{R}^{1+sr} \mapsto]0; 1[$ données par :

$$\pi_{1jk}(\Theta) = \frac{p_{jk}}{1 + \theta \sum_{m=1}^R c_{mk} p_{mk}} ; \quad \pi_{2jk}(\Theta) = \frac{\theta p_{jk} \sum_{m=1}^R c_{mk} p_{mk}}{1 + \theta \sum_{m=1}^R c_{mk} p_{mk}}, \quad (j = 1, 2, \dots, R), \quad (7)$$

où $\Theta = (\theta, P^T)^T$ de dimension $(1 + SR)$ est le vecteur des paramètres, $\theta (\theta > 0)$ le paramètre effet moyen de la mesure et où $P = (P_1^T, P_2^T, \dots, P_S^T)^T$ avec $P_k = (p_{1k}, p_{2k}, \dots, p_{Rk})^T$, vecteur de dimension R est tel que :

$$P_k \in \mathcal{S}_k^{(R-1)} = \{(p_{1k}, p_{2k}, \dots, p_{Rk})^T \in \mathbb{R}^R, p_{jk} > 0, \sum_{j=1}^R p_{jk} = 1\}. \quad (8)$$

Le vecteur Θ comporte deux ensembles de paramètres ; θ et les vecteurs P_k . Or, par construction, chaque sous-vecteur P_k appartient au simplexe d'ordre R . De ce fait, le vecteur paramètre Θ est soumis à des contraintes de bornes ($\theta > 0$), $0 < p_{jk} < 1$) et à des contraintes linéaires :

$$h_k(\Theta) = \langle \mathbf{1}_r, P_k \rangle - 1 = 0, \quad (k = 1, 2, \dots, S), \quad (9)$$

où \langle, \rangle est le produit scalaire classique et $\mathbf{1}_r = (1, \dots, 1)^T \in \mathbb{R}^R$ le vecteur des unités.

2.3.3 Modèle 3 (annexe A04)

Dans la pratique de l'évaluation d'une mesure de sécurité routière, il est souvent indiqué de considérer le logarithme de l'effet plutôt que l'effet lui-même. Une des raisons de cette pratique se trouve dans le résultat asymptotique obtenu par Tanner. En partant de ce résultat, nous avons construit la transformation bijective g_1

$$g_1 : \mathbb{R}^{1+SR} \mapsto \mathbb{R}^{1+SR}; \Theta \mapsto g_1(\Theta) = \Theta_1$$

qui laisse le vecteur P invariant et transforme $\theta \mapsto \beta = \log_e(\theta)$. Une conséquence intéressante de la bijection g_1 est qu'elle permet de passer outre la contrainte sur le paramètre θ en transférant la recherche numérique de solution pour l'effet moyen de $]0, +\infty[$ dans $] - \infty, +\infty[$. Un tel transfert donne plus de souplesse à l'estimation des paramètres et en particulier à celle liée au paramètre θ . En utilisant cette fonction bijective, on obtient un troisième modèle multinomial. Ainsi, pour k fixé, on suppose que le vecteur $(X_{1k}, X_{2k})^T$ suit une loi multinomiale notée $\mathcal{M}(n_k; \Pi_{1k}(\Theta_1), \Pi_{2k}(\Theta_1))$, où

$$\Pi_{tk}(\Theta_1) = (\pi_{t1k}(\Theta_1), \pi_{t2k}(\Theta_1), \dots, \pi_{tRk}(\Theta_1))$$

est le vecteur de dimension R de composantes $\pi_{tjk} : \mathbb{R}^{1+SR} \mapsto]0; 1[$ données par :

$$\pi_{1jk}(\Theta_1) = \frac{p_{jk}}{1 + e^\beta \sum_{m=1}^R c_{mk} p_{mk}} ; \quad \pi_{2jk}(\Theta_1) = \frac{e^\beta c_{jk} p_{jk}}{1 + e^\beta \sum_{m=1}^R c_{mk} p_{mk}}, \quad (j = 1, 2, \dots, R), \quad (10)$$

$\Theta_1 = (\beta, P^T)^T$ de dimension $(1 + SR)$ est le vecteur des paramètres où $\beta (\beta \in] - \infty; +\infty[)$ représente le logarithme de l'effet moyen et où $P = (P_1^T, P_2^T, \dots, P_S^T)^T$ avec $P_k = (p_{1k}, p_{2k}, \dots, p_{Rk})^T$, vecteur de dimension R est tel que :

$$P_k \in \mathcal{S}_k^{(R-1)} = \{(p_{1k}, p_{2k}, \dots, p_{Rk})^T \in \mathbb{R}^R, p_{jk} > 0, \sum_{j=1}^R p_{jk} = 1\}. \quad (11)$$

Par construction des modèles, chaque vecteur auxiliaire P_k appartient au simplexe d'ordre R . Par conséquent, chacun de ces vecteurs peut être associé à une loi de probabilité qui faut estimer (voir section 2.4) au même titre que le paramètre lié à l'effet moyen de la mesure. Selon la façon dont on estime le vecteur P_k (voir par exemple l'expression (2) p. 13), chaque composante p_{jk} , pour k fixé, peut être interprétée comme un risque d'accidents. En pratique, on aménage un site parce que les risques d'accidents avant l'aménagement sont anormalement élevés par rapport à d'autres entités du réseau. L'effet espéré de l'aménagement routier (mesure de sécurité routière) est donc de faire baisser les risques d'accidents. De ce fait, pour mieux apprécier l'impact de la mesure de sécurité routière sur les différents types de risque d'accidents, il convient de distinguer la période avant de celle d'après et d'inclure les données de contrôle dans la définition de ces risques. A cette fin, nous utilisons, pour le site expérimental k fixé, le paramètre $\pi_{1jk}(\Theta)$ (resp. $\pi_{2jk}(\Theta)$) comme indicateur du risque d'accident du type j au cours de la période avant (resp. après). En effet, la probabilité conditionnelle (conditionnellement au nombre total d'accidents observés au site expérimental k et aux données de la zone de contrôle associée) $\pi_{1jk}(\Theta)$ (resp. $\pi_{2jk}(\Theta)$), représente, par construction, la probabilité qu'un accident survenu au niveau du site expérimental k soit du type j avant (resp. après). La connaissance de ces différents risques d'accidents avant et après et donc de la loi de probabilité des vecteurs aléatoires $(X_{1k}, X_{2k})^T$ passe par l'estimation du vecteur Θ .

2.4 Estimation multivariée sous contraintes

2.4.1 Introduction

Après la modélisation, nous nous sommes intéressés dans cette partie de nos travaux, à l'estimation des paramètres liés à nos approches statistiques en utilisant la méthode du maximum de vraisemblance sous contraintes. Il est hors de nos propos de faire une revue complète de cette approche d'estimation car différents auteurs [3, 13, 20, 26, 31, 35, 38, 63, 64, 65, 66, 77, 79, 84, 88, 96, 98, 99, 104, 105] ont déjà abordé la question et proposé des résultats asymptotiques sous des conditions de régularité diverses et variées.

La méthode d'estimation que nous proposons correspond au cas particulier où les variables de l'échantillon sont supposées indépendantes non identiquement distribuées et dont les lois sont soumises à des contraintes linéaires et de bornes. Dans une première série de travaux (paragraphe 2.4.2 et annexe A11 p. 190), nous avons étudié l'existence théorique de solutions en adaptant les conditions de régularité proposées par Crowder [20] et que nous rappelons ci-dessous. Ensuite, nous exploitons les conditions et les résultats de ce dernier auteur pour analyser la distribution asymptotique (annexes A07 p. 142 et A09 p. 154) de l'estimateur du maximum de vraisemblance sous contraintes relatif à l'effet moyen et aux risques globaux d'accident.

On suppose que $\{(X_{1k}, X_{2k})^T, 1 \leq k \leq S\}$ est un échantillon de S couples de vecteurs aléatoires indépendants tels que la loi de chaque couple soit une loi multinomiale donnée par l'un des modèles (par exemple le modèle 1) de la section 2.3. On note $\mathcal{L}(\Theta)$ la log-vraisemblance

associée à cette suite et supposée continûment différentiable. La recherche de l'existence de l'estimateur $\hat{\Theta}$ du maximum de vraisemblance de Θ^0 (la vraie valeur) revient au problème :

$$\text{Maximiser } \mathcal{L}(\Theta) \quad \text{sous les contraintes } h_k(\Theta) = \mathbf{0}; \quad k = 1, 2, \dots, S. \quad (12)$$

où $h_k(\Theta)$ sont les fonctions de contraintes supposées également continûment différentiables. Cette recherche de solution au problème de l'évaluation multidimensionnelle d'une mesure de sécurité routière est donc un cas particulier du problème plus général de maximum de vraisemblance sous contraintes. Cette recherche de solution est équivalente à la résolution du système

$$\left(\frac{\partial \mathcal{L}}{\partial \Theta}\right)_{\hat{\Theta}} + H_{\hat{\Theta}}^T \hat{\lambda} = \mathbf{0}, \quad h(\hat{\Theta}) = \mathbf{0} \quad (13)$$

où $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_S)^T$ est le vecteur de dimension S des multiplicateurs de Lagrange, $h = (h_1, h_2, \dots, h_S)$ et où $H_{\Theta} = (H_1, H_2)$ est la matrice de dimension $S \times (1 + SR)$ avec

$$H_1 = \left(\frac{\partial h_1}{\partial \theta}, \dots, \frac{\partial h_S}{\partial \theta}\right)^T, \quad H_2 = \left(\frac{\partial h_1}{\partial P}, \dots, \frac{\partial h_S}{\partial P}\right)^T$$

respectivement des matrices de dimension $S \times 1$ et $S \times (SR)$ des dérivées partielles des fonctions de contraintes par rapport aux composantes du vecteur paramètre Θ . On pose par la suite

$$J_{\Theta} = E\left(-\frac{\partial^2 \mathcal{L}}{\partial \Theta \partial \Theta^T}\right)$$

la matrice d'information de Fisher. La recherche d'une solution aux équations de vraisemblance sous contraintes est intimement liée au fait que la matrice d'information de Fisher et la matrice des contraintes soient de plein rang. Afin d'étudier l'existence d'une telle solution, on admettra dans un premier temps que ces matrices sont définies positives. Les résultats obtenus sur la question d'inversibilité formelle des matrices d'information sont exposés au chapitre 3.

2.4.2 Résultats asymptotiques généraux

Nous donnons dans ce paragraphe les conditions générales utilisées pour établir l'existence et la convergence en probabilité et en loi de l'estimateur du maximum de vraisemblance sous contraintes. Ces conditions ne sont pas uniques et dépendent souvent de la structure du modèle et des contraintes. Les conditions utilisées dans ce travail sont similaires à celles utilisées par Crowder [20] .

On admet les conditions de régularité suivantes :

- (C1) J_{Θ} et $H_{\Theta} J_{\Theta}^{-1} H_{\Theta}^T$ sont définies positives ;
- (C2) $E\left[\frac{\partial \mathcal{L}}{\partial \Theta}\right] = \mathbf{0}$ et $Var\left[\left(\frac{\partial \mathcal{L}}{\partial \Theta}\right)_{\Theta^0}\right] = J_{\Theta^0}$;
- (C3) $\forall \epsilon > 0$ il existe $\delta \in (0, \epsilon]$ tel que

$$\text{Proba} \left[\sup_{\|\Theta - \Theta^0\| = \delta} (\Theta - \Theta^0)^T F(X; \Theta) < \delta^2 \right] \longrightarrow 1, \quad n_k \text{ et } N = \sum_{k=1}^S n_k \longrightarrow + \infty$$

où $F(X; \Theta) = (I_{1+SR} - \Phi_{\Theta} H_{\Theta}) \{ J_{\Theta}^{-1} (\frac{\partial \mathcal{L}(\Theta)}{\partial \Theta})_{\Theta^0} + M_{\Theta} (\Theta - \Theta^0) \}$ avec $M_{\Theta} = J_{\Theta}^{-1} \left(J_{\Theta} + (\frac{\partial^2 \mathcal{L}(\Theta)}{\partial \Theta \partial \Theta^T}) \right)$ est une matrice $(1+SR) \times (1+SR)$, $\Phi_{\Theta} = J_{\Theta}^{-1} H_{\Theta}^T (H_{\Theta} J_{\Theta}^{-1} H_{\Theta}^T)^{-1}$ est une matrice $(1+SR) \times S$ et X le vecteur aléatoire constitué des S couples (X_{1k}, X_{2k}) ;

(C4) $M_{\hat{\Theta}} \xrightarrow{Proba} 0$, $\Phi_{\hat{\Theta}} - \Phi_{\Theta^0} \xrightarrow{Proba} 0$, $|\Phi_{\Theta^0}| < +\infty$ quand n_k et $N = \sum_{k=1}^S n_k \rightarrow +\infty$

(C5) $(\frac{\partial \mathcal{L}}{\partial \Theta})_{\Theta^0} \stackrel{Distr.}{\sim} \mathcal{N}(\mathbf{0}, J_{\Theta^0})$.

Les conditions (C1) et (C2) sont des conditions classiques relatives à la fonction score (le gradient de la log-vraisemblance). La condition (C1), par exemple, est essentielle car elle sert à construire la fonction $F(X; \Theta)$ utilisée pour établir l'existence théorique de $\hat{\Theta}$. Si cette condition (C1) n'est pas vérifiée, certains auteurs (par exemple [97, 99, 102]) proposent de modifier la matrice d'information en utilisant convenablement la structure du modèle et des contraintes. La condition (C3) permet ensuite d'utiliser une version probabiliste du théorème du point fixe évoqué par exemple dans Aitchison et Silvey [3]. Les conditions (C4) et (C5) servent à établir la convergence en probabilité et en loi de l'estimateur du maximum de vraisemblance sous contraintes.

Théorème 2.4.1 (Annexe A11)

Sous les conditions de régularité (C1) à (C4), il existe $(\hat{\Theta}, \hat{\lambda})$ une solution convergente aux équations du maximum de vraisemblance sous contraintes.

Globalement, la démonstration consiste à appliquer le théorème de la moyenne (c'est-à-dire un développement limité d'ordre 1 des équations de vraisemblance) dans un voisinage compact (au sens de la topologie euclidienne) de Θ^0 . Ensuite, on montre, dans ce voisinage compact, que les équations de vraisemblance sous contraintes sont équivalentes à

$$F(X, \Theta) - (\Theta - \Theta^0) = \mathbf{0}.$$

En appliquant, enfin, un équivalent du théorème du point fixe, une solution $\hat{\Theta}$ existe, avec $\|\hat{\Theta} - \Theta^0\| < \delta$, si avec une probabilité tendant vers 1, $(\Theta - \Theta^0)^T \{F(X, \Theta) - (\Theta - \Theta^0)\} < 0$ pour $\|\hat{\Theta} - \Theta^0\| = \delta$

Théorème 2.4.2 (Annexe A07 et A09)

Sous les conditions (C1) à (C5), il existe une solution convergente aux équations de vraisemblance sous contraintes et on a asymptotiquement (n_k et $N = \sum_{k=1}^S n_k \rightarrow +\infty$)

$$(\hat{\Theta} - \Theta^0) \stackrel{Distr.}{\sim} \mathcal{N}(\mathbf{0}; J_{\Theta}^{-1} - J_{\Theta}^{-1} H_{\Theta}^T (H_{\Theta} J_{\Theta}^{-1} H_{\Theta}^T)^{-1} H_{\Theta} J_{\Theta}^{-1}) \quad (14)$$

On utilise le même principe du théorème de la moyenne dans le même voisinage compact pour montrer que le vecteur aléatoire constitué de l'estimateur du maximum de vraisemblance sous contraintes et une transformation des multiplicateurs de Lagrange est asymptotiquement normalement distribué.

Chapitre 3

Complément de Schur et approximation d'erreurs standards

3.1 Introduction

La plupart des méthodes multivariées en statistique appliquée exige l'inversion de matrices d'information. En effet, les algorithmes itératifs du type Newton-Raphson utilisés pour estimer les paramètres nécessitent à chaque itération l'inversion d'une matrice. D'autre part, on a besoin de connaître, en fin d'itération, une approximation de la précision relative de chaque paramètre du modèle. Ceci va nous conduire à étudier les éléments de l'inverse d'une matrice d'information. Dans le cas des approches statistiques de l'évaluation proposées, nous connaissons la structure formelle des matrices d'information et de la matrice de contraintes. Il faut donc en extraire les écart-types en utilisant les résultats du théorème 2.4.2. Or ces matrices sont souvent creuses et de tailles croissantes par rapport aux dimensions des tableaux de données d'accident. Il faut alors pouvoir vérifier la condition fondamentale ($C1$) relative à la non singularité de nos matrices d'information et surtout pouvoir disposer d'une expression analytique de la variance de chaque paramètre afin de bâtir des outils simples d'aide à la décision (intervalles de confiance et tests d'hypothèse). Cela passe non seulement par l'étude de l'inversibilité de la matrice d'information considérée mais aussi par l'analyse détaillée des éléments de son inverse. A cet effet, nous utilisons la technique du complément de Schur [28, 80, 95, 100, 110] brièvement rappelé à la section 3.2. Ainsi, nous démontrons l'inversibilité de nos matrices d'information et obtenons de surcroît l'expression analytique des composantes de chaque inverse en fonction des données du problème et des paramètres du modèle utilisé. Au paragraphe 3.3.2 (voir annexes A03 et A07), nous présentons les résultats obtenus sur l'inversibilité des matrices d'information liées à notre problématique. Nous exploitons alors ces résultats pour estimer les variances et covariances (paragraphe 3.3.3 et annexe A01) de l'estimateur sous contraintes de l'effet moyen et des risques d'accident. Ces variances sont utilisées pour construire (paragraphe 3.3.4 et annexe A04) des intervalles de confiance et pour bâtir des tests d'hypothèse relatifs à la nullité de l'effet moyen de la mesure et des risques d'accident.

3.2 Complément de Schur

3.2.1 Définition et premières propriétés

Considérons la matrice partitionnée suivante

$$\mathbf{M} = \begin{pmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{R} & \mathbf{S} \end{pmatrix}$$

où la matrice carrée \mathbf{P} est supposée non singulière (invertible). Alors le complément de Schur de \mathbf{P} dans \mathbf{M} est défini par

$$(\mathbf{M}/\mathbf{P}) = \mathbf{S} - \mathbf{R}\mathbf{P}^{-1}\mathbf{Q}.$$

En utilisant cette définition, on a la formule de diagonalisation par blocs de Aitken cité par Zhang et al [110] :

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{R}\mathbf{P}^{-1} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{R} & \mathbf{S} \end{pmatrix} \begin{pmatrix} \mathbf{I} & -\mathbf{P}^{-1}\mathbf{Q} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} = \begin{pmatrix} \mathbf{P} & \mathbf{0} \\ \mathbf{0} & (\mathbf{M}/\mathbf{P}) \end{pmatrix} \quad (1)$$

Il s'en suit alors deux grands résultats algébriques concernant les calculs de rang et de déterminant de matrices. On montre en effet que si le complément de Schur de \mathbf{P} dans \mathbf{M} existe alors

$$\begin{aligned} \text{rang}(\mathbf{M}) &= \text{rang}(\mathbf{P}) + \text{rang}((\mathbf{M}/\mathbf{P})) \\ &= \text{rang}(\mathbf{P}) + \text{rang}(\mathbf{S} - \mathbf{R}\mathbf{P}^{-1}\mathbf{Q}) \end{aligned}$$

et

$$\begin{aligned} \det(\mathbf{M}) &= \det(\mathbf{P}) \times \det((\mathbf{M}/\mathbf{P})) \\ &= \det(\mathbf{P}) \times \det(\mathbf{S} - \mathbf{R}\mathbf{P}^{-1}\mathbf{Q}) \end{aligned}$$

Le résultat concernant le calcul de rang est souvent cité dans la littérature comme la formule d'additivité de rang de Guttman alors que celui relatif au déterminant est appelé formule du déterminant de Schur. Comme la matrice \mathbf{P} est supposée non singulière, l'invertibilité de \mathbf{M} est équivalente à celle de (\mathbf{M}/\mathbf{P}) . Cette dernière propriété a fait l'objet de notre attention dans nos travaux. Nous rappelons brièvement quelques résultats principaux concernant l'inversion de matrices.

3.2.2 Quelques applications à l'inversion formelle de matrices

Théorème 3.2.1 *On suppose que la matrice carrée*

$$\mathbf{M} = \begin{pmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{R} & \mathbf{S} \end{pmatrix}$$

est non singulière (invertible) avec \mathbf{P} et \mathbf{S} deux matrices carrées. On a les résultats suivants :

(i) *si \mathbf{P} est non singulière alors le complément de Schur $(\mathbf{M}/\mathbf{P}) = \mathbf{S} - \mathbf{R}\mathbf{P}^{-1}\mathbf{Q}$ est également non singulière et de plus*

$$\mathbf{M}^{-1} = \begin{pmatrix} \mathbf{P}^{-1} + \mathbf{P}^{-1}\mathbf{Q}(\mathbf{M}/\mathbf{P})^{-1}\mathbf{R}\mathbf{P}^{-1} & -\mathbf{P}^{-1}\mathbf{Q}(\mathbf{M}/\mathbf{P})^{-1} \\ -(\mathbf{M}/\mathbf{P})^{-1}\mathbf{R}\mathbf{P}^{-1} & (\mathbf{M}/\mathbf{P})^{-1} \end{pmatrix},$$

(ii) si \mathbf{S} est non singulière et \mathbf{P} non nécessairement inversible alors le complément de Schur $(\mathbf{M}/\mathbf{S}) = \mathbf{P} - \mathbf{Q}\mathbf{S}^{-1}\mathbf{R}$ est également non singulière et de plus

$$\mathbf{M}^{-1} = \begin{pmatrix} (\mathbf{M}/\mathbf{S})^{-1}\mathbf{R}\mathbf{P}^{-1} & -(\mathbf{M}/\mathbf{S})^{-1}\mathbf{Q}\mathbf{S}^{-1} \\ -\mathbf{S}^{-1}\mathbf{R}(\mathbf{M}/\mathbf{S})^{-1} & \mathbf{S}^{-1} + \mathbf{S}^{-1}\mathbf{R}(\mathbf{M}/\mathbf{S})^{-1}\mathbf{Q}\mathbf{S}^{-1} \end{pmatrix},$$

(iii) si maintenant \mathbf{P} et \mathbf{S} sont toutes deux non singulières alors les compléments de Schur $(\mathbf{M}/\mathbf{P}) = \mathbf{S} - \mathbf{R}\mathbf{P}^{-1}\mathbf{Q}$ et $(\mathbf{M}/\mathbf{S}) = \mathbf{P} - \mathbf{Q}\mathbf{S}^{-1}\mathbf{R}$ sont également non singulières et de plus

$$(\mathbf{P} - \mathbf{Q}\mathbf{S}^{-1}\mathbf{R})^{-1} = \mathbf{P}^{-1} + \mathbf{P}^{-1}\mathbf{Q}(\mathbf{S} - \mathbf{R}\mathbf{P}^{-1}\mathbf{Q})^{-1}\mathbf{R}\mathbf{P}^{-1}.$$

Corollaire 3.2.2 On suppose que la matrice carrée

$$\mathbf{M} = \begin{pmatrix} \mathbf{P} & \mathbf{q} \\ \mathbf{r}^T & s \end{pmatrix}$$

est non singulière (inversible) avec \mathbf{P} matrice carrée inversible, \mathbf{q} et \mathbf{r} deux vecteurs et s un scalaire non nul. Alors le complément de Schur (\mathbf{M}/s) s'écrit

$$(\mathbf{M}/s) = \mathbf{P} - \frac{1}{s}\mathbf{q}\mathbf{r}^T$$

et on a

$$\left(\mathbf{P} - \frac{1}{s}\mathbf{q}\mathbf{r}^T\right)^{-1} = \mathbf{P}^{-1} + \frac{1}{s - \mathbf{r}^T\mathbf{P}^{-1}\mathbf{q}}\mathbf{P}^{-1}\mathbf{q}\mathbf{r}^T\mathbf{P}^{-1}$$

où le complément de Schur $(\mathbf{M}/\mathbf{P}) = s - \mathbf{r}^T\mathbf{P}^{-1}\mathbf{q}$ est un scalaire supposé non nul.

D'autres résultats algébriques ainsi que des cas particuliers d'inversion de matrices utilisant le complément de Schur sont disponibles dans la littérature (voir par exemple Zhang et al 2006, Ouellette 1981). Ces derniers auteurs donnent également une revue importante des applications du complément de Schur à la statistique. Dans le cadre de nos travaux de recherche sur l'évaluation, nous combinons cette technique du complément de Schur et la structure particulière des modèles statistiques pour démontrer l'inversibilité formelle des matrices d'information et déduire l'expression analytique des écart-types des paramètres. Les résultats obtenus sont exposés dans les annexes A01, A03, A04, A07 et A09. Nous donnons, ci-après, un extrait de ces résultats.

3.3 Étude des matrices d'information liées aux modèles d'évaluation

3.3.1 Structure générale des matrices d'information

Les conditions générales et théoriques utilisées pour établir l'existence, la convergence en probabilité et en loi de l'estimateur du maximum de vraisemblance sous contraintes reposent essentiellement sur l'inversibilité de matrices d'information partitionnées de la forme :

$$\Gamma_{\Theta} = \begin{bmatrix} J_{\Theta} & H_{\Theta}^T \\ H_{\Theta} & \mathcal{O}_{S,S} \end{bmatrix}; \quad J_{\Theta} = \begin{bmatrix} \tau_{\Theta} & U_{\Theta}^T \\ U_{\Theta} & B_{\Theta} \end{bmatrix} \quad (2)$$

où $\tau_{\Theta} = E(-\frac{\partial^2 \mathcal{L}}{\partial \theta \partial \theta})$ est un scalaire strictement positif, $U_{\Theta} = E(-\frac{\partial^2 \mathcal{L}}{\partial \theta \partial P})$ est un vecteur de dimension (SR) , $B_{\Theta} = E(-\frac{\partial^2 \mathcal{L}}{\partial P \partial P})$ est une matrice de dimension $(SR) \times (SR)$ et $\mathcal{O}_{S,S}$ la matrice $S \times S$ constituée de zéros. D'après le point (i) du théorème 3.2.1, l'inversibilité de la matrice d'information Γ_{Θ} est équivalente à celle des matrices J_{Θ} et $H_{\Theta} J_{\Theta}^{-1} H_{\Theta}^T$. La forme explicite des composantes de chacun des blocs (par exemple pour le modèle de la proposition 2.3.1 du chapitre 2) est donnée dans les annexes A03 et A04.

Différentes méthodes (voir par exemple [39, 63-65, 97, 102, 107]) peuvent être utilisées pour étudier la non-singularité de la matrice Γ_{Θ} . Nous démontrons l'inversibilité formelle de cette matrice et obtenons l'expression analytique explicite des éléments de son inverse en utilisant les propriétés algébriques du complément de Schur et du produit de Kronecker de deux matrices. Les résultats exposés au paragraphe 3.3.2 sont relatifs au modèle de la proposition 2.3.1 (modèle 1) et s'appliquent parfaitement aux autres modèles (propositions 2.3.2 et 2.3.3) du chapitre 2 en utilisant la décomposition proposée dans l'annexe A04.

3.3.2 Étude de l'inversibilité des matrices d'information

Pour $k = 1, 2, \dots, S$ fixé, on note $Z_k = (c_{1k}, c_{2k}, \dots, c_{Rk})^T$, le vecteur de dimension R des coefficients de contrôle associés au site expérimental k et on définit le scalaire strictement positif $\gamma_k = n_k / (1 + \theta \langle Z_k, P_k \rangle)$,

$$V_{\Theta,k} = \theta \frac{\gamma_k^{1/2}}{n_k^{1/2}} Z_k \quad \text{et} \quad \Omega_{\Theta,k} = \text{diag} \left(\frac{1 + \theta c_{1k}}{p_{1k}}, \dots, \frac{1 + \theta c_{Rk}}{p_{Rk}} \right),$$

respectivement vecteur de dimension R et matrice diagonale de dimension $R \times R$. En utilisant le modèle 1, on montre que

$$\tau_{\Theta} = \sum_{k=1}^S \gamma_k (\langle Z_k, P_k \rangle) / (\theta (1 + \theta \langle Z_k, P_k \rangle)), \quad U_{\Theta} = (U_{\Theta,1}^T, \dots, U_{\Theta,S}^T)^T \quad \text{avec} \\ U_{\Theta,k} = \frac{\gamma_k^{3/2}}{\theta n_k^{1/2}} V_{\Theta,k} \quad \text{et} \quad B_{\Theta} = \text{Bloc-diag}(B_{\Theta,1}, \dots, B_{\Theta,S}) \quad \text{avec} \quad B_{\Theta,k} = \gamma_k (\Omega_{\Theta,k} - V_{\Theta,k} V_{\Theta,k}^T).$$

Théorème 3.3.1 (Annexe A07)

En utilisant par exemple le modèle 1, on montre que

$$(i) \quad \| Z_k \|_{\Omega_{\Theta,k}^{-1}}^2 < \frac{\langle Z_k, P_k \rangle}{\theta};$$

$$(ii) \quad \| V_{\Theta,k} \|_{\Omega_{\Theta,k}^{-1}}^2 < 1.$$

Théorème 3.3.2 (Annexe A07)

On pose $t_k = (1 - \| V_{\Theta,k} \|_{\Omega_{\Theta,k}^{-1}}^2)^{-1}$ pour $(k = 1, 2, \dots, s)$. On montre alors que B_{Θ}^{-1} existe et que

$$B_{\Theta}^{-1} = \text{Bloc-diag}(B_{\Theta,1}^{-1}, \dots, B_{\Theta,S}^{-1})$$

où

$$B_{\Theta,k}^{-1} = \gamma_k^{-1} (\Omega_{\Theta,k}^{-1} + t_k \Omega_{\Theta,k}^{-1} V_{\Theta,k} V_{\Theta,k}^T \Omega_{\Theta,k}^{-1}).$$

Pour démontrer l'existence de B_{Θ}^{-1} , on démontre d'abord l'inversibilité de la matrice $B_{\Theta,k} = E(-\frac{\partial^2 \mathcal{L}}{\partial P_k \partial P_k})$ en construisant une nouvelle matrice de dimension $(R+1) \times (R+1)$ de telle sorte que l'un des compléments de Schur de cette nouvelle matrice soit égale à $B_{\Theta,k}$. Ensuite, on montre que les deux compléments de Schur de cette nouvelle matrice existent et sont inversibles. On en déduit alors l'inverse de $B_{\Theta,k}$ en utilisant le point (iii) du théorème 3.2.1. Les deux résultats suivants permettent d'obtenir l'inverse des matrices d'information désirées.

Théorème 3.3.3 (Annexe A03)

On montre que

i) $(J_{\Theta}/B_{\Theta}) > 0$, ii) $(J_{\Theta}/\tau_{\Theta})^{-1} = B_{\Theta}^{-1} + (J_{\Theta}/B_{\Theta})^{-1} B_{\Theta}^{-1} U_{\Theta} U_{\Theta}^T B_{\Theta}^{-1}$
 où (J_{Θ}/B_{Θ}) (resp. $(J_{\Theta}/\tau_{\Theta})$) est le complément de Schur de B_{Θ} (resp. de τ_{Θ}) dans J_{Θ} et défini par $(J_{\Theta}/B_{\Theta}) = \tau_{\Theta} - U_{\Theta}^T B_{\Theta}^{-1} U_{\Theta}$ (resp. $(J_{\Theta}/\tau_{\Theta}) = B_{\Theta} - U_{\Theta} \tau_{\Theta}^{-1} U_{\Theta}^T$).

Connaissant l'inverse des deux compléments de Schur associés à la matrice d'information de Fisher J_{Θ} , on peut alors en déduire l'inversibilité de cette matrice d'information. Le résultat ci-dessous finalise l'objectif fixé à savoir la vérification de la condition (C1) du paragraphe 2.4.2.

Théorème 3.3.4 (Annexe A03)

On définit

$$B_{\Theta,k}^{(0)} = \begin{bmatrix} B_{\Theta,k} & -1_R \\ 1_R^T & 0 \end{bmatrix}, \quad \Lambda_{\Theta}^{(0)} = \begin{bmatrix} \Lambda_{\Theta} & \xi_{\Theta} \\ \xi_{\Theta}^T & 0 \end{bmatrix}, \quad \Lambda_{\Theta}^{(1)} = \begin{bmatrix} \Lambda_{\Theta} & -\xi_{\Theta} \\ \xi_{\Theta}^T & (J_{\Theta}/B_{\Theta}) \end{bmatrix},$$

respectivement une matrice de dimension $(1+R) \times (1+R)$ et deux matrices de dimension $(1+S) \times (1+S)$ avec

$$\Lambda_{\Theta} = \text{diag} \left[(B_{\Theta,1}^{(0)}/B_{\Theta,1}), \dots, (B_{\Theta,S}^{(0)}/B_{\Theta,S}) \right] \quad (3)$$

une matrice de dimension $S \times S$ et

$$\xi_{\Theta} = H_2 B_{\Theta}^{-1} U_{\Theta} \quad (4)$$

un vecteur de dimension S . On montre que

- (i) $(B_{\Theta,k}^{(0)}/B_{\Theta,k}) > 0$ pour $k = 1, 2, \dots, S$;
 (ii) $H_{\Theta} J_{\Theta}^{-1} H_{\Theta}^T$ est inversible et on a

$$\begin{aligned} (H_{\Theta} J_{\Theta}^{-1} H_{\Theta}^T)^{-1} &= (\Lambda_{\Theta}^{(1)}/(J_{\Theta}/B_{\Theta}))^{-1} \\ &= \Lambda_{\Theta}^{-1} - \Lambda_{\Theta}^{-1} \xi_{\Theta} ((J_{\Theta}/B_{\Theta}) + (\Lambda_{\Theta}^{(0)}/\Lambda_{\Theta}))^{-1} \xi_{\Theta}^T \Lambda_{\Theta}^{-1} \end{aligned} \quad (5)$$

où $(\Lambda_{\Theta}^{(1)}/(J_{\Theta}/B_{\Theta}))$ est le complément de Schur de (J_{Θ}/B_{Θ}) dans $\Lambda_{\Theta}^{(1)}$.

Corollaire 3.3.5 (annexe A03)

(i)

$$H_2^T (\Lambda_{\Theta}^{(1)}/(J_{\Theta}/B_{\Theta}))^{-1} H_2 = A_{\Theta} \otimes \mathbf{1}_{R,R},$$

une matrice $(SR) \times (SR)$ où \otimes est le produit de Kronecker, $\mathbf{1}_{R,R}$ est la matrice $R \times R$ des unités;

$A_{\Theta} = ((A_{k,m}))$ est la matrice $S \times S$ dont le $(k,m)^e$ élément est donné par

$$A_{k,m} = \frac{\delta_{k,m}}{\|\mathbf{1}_R\|_{B_{\Theta,k}^{-1}}^2} - \frac{[(\Lambda_{\Theta}^{(1)}/\Lambda_{\Theta})]^{-1} t_k t_m \gamma_k \gamma_m \Delta_{\Theta,k} \Delta_{\Theta,m}}{n_k n_m \|\mathbf{1}_R\|_{B_{\Theta,k}^{-1}}^2 \|\mathbf{1}_R\|_{B_{\Theta,m}^{-1}}^2}. \quad (6)$$

$\delta_{k,m}$ est le symbole de Kronecker, $\Delta_{\Theta,k} = Z_k^T \Omega_{\Theta,k}^{-1} \mathbf{1}_R$, $k = 1, 2, \dots, S$;

(ii)

$$(J_{\Theta}/\tau_{\Theta})^{-1} H_2^T (\Lambda_{\Theta}^{(1)}/(J_{\Theta}/B_{\Theta}))^{-1} H_2 (J_{\Theta}/\tau_{\Theta})^{-1} = G_{\Theta}$$

où G_{Θ} est la matrice $(SR) \times (SR)$ dont l'élément $G_{k,m}$ est également une matrice $R \times R$ donnée par

$$G_{k,m} = \sum_{j=1}^S \left[\left(\sum_{i=1}^S A_{i,j} J_{k,i} \right) \mathbf{1}_{R,R} \right] J_{j,m} \quad (7)$$

avec $k, m = 1, 2, \dots, S$.

Corollaire 3.3.6 (annexe A03)

Γ_{Θ} est inversible

$$\Gamma_{\Theta}^{-1} = \begin{bmatrix} W_{\Theta} & \Phi_{\Theta} \\ \Phi_{\Theta}^T & -(H_{\Theta} J_{\Theta}^{-1} H_{\Theta}^T)^{-1} \end{bmatrix}, \quad W_{\Theta} = \begin{bmatrix} W_{\Theta}(1,1) & W_{\Theta}^T(2,1) \\ W_{\Theta}(2,1) & W_{\Theta}(2,2) \end{bmatrix} \quad (8)$$

où $W_{\Theta}(1,1)$ est un scalaire strictement positif, $W_{\Theta}(2,2)$ et $W_{\Theta}(2,1)$ respectivement une matrice $(SR) \times (SR)$ et un vecteur de dimension (SR) donnés par :

$$\begin{aligned} W_{\Theta}(1,1) &= (\Lambda_{\Theta}^{(1)}/\Lambda_{\Theta})^{-1} \\ W_{\Theta}(2,2) &= (J_{\Theta}/\tau_{\Theta})^{-1} - (J_{\Theta}/\tau_{\Theta})^{-1} H_2^T R_{\Theta}^{-1} H_2 (J_{\Theta}/\tau_{\Theta})^{-1} \\ W_{\Theta}(2,1) &= -(J_{\Theta}/B_{\Theta})^{-1} [B_{\Theta}^{-1} U_{\Theta} - (J_{\Theta}/\tau_{\Theta})^{-1} H_2^T R_{\Theta}^{-1} \xi_{\Theta}]. \end{aligned} \quad (9)$$

où $R_{\Theta} = -(\Gamma_{\Theta}/J_{\Theta})$ est l'opposé du complément de Schur de J_{Θ} dans Γ_{Θ} .

La matrice W_{Θ} est le développement formel de la matrice de variances covariances de l'estimateur du maximum de vraisemblance sous contraintes (voir théorème 2.4.2) relativement aux modèles et aux contraintes exposés au chapitre 2. Les résultats des paragraphes suivants permettent non seulement d'expliciter ces composantes mais aussi de les interpréter. En particulier, on montre que la composante $W_{\Theta}(1,1)$ est une estimation de la variance de la première composante de $\hat{\Theta}$ c'est-à-dire du paramètre associé à l'efficacité moyenne de la mesure. On en déduit alors des intervalles de confiance pour l'effet moyen de la mesure et les risques d'accident ainsi qu'une procédure de test d'hypothèse.

3.3.3 Application à l'estimation des écart-types de l'effet moyen et des différents types de risque d'accidents

Les résultats de ce paragraphe donnent les détails techniques des composantes de la matrice de variances et covariances de l'estimateur du maximum de vraisemblance sous contraintes. On utilise pour cela le modèle 1 ainsi que les contraintes linéaires associées. On rappelle que le but poursuivi est de fournir la structure explicite des variances et de covariances des composantes du vecteur des paramètres afin de bâtir des intervalles de confiance et construire des statistiques pour évaluer l'hypothèse d'effet moyen nul c'est-à-dire comparer θ à l'unité. Les deux théorèmes ci-dessous et leur corollaire résultent de manipulations algébriques combinant les caractéristiques du modèle 1, les contraintes, le produit de Kronecker et les propriétés du complément de Schur. Les détails algébriques sont exposés dans les annexes A03 et A07. Les résultats statistiques font l'objet de l'annexe A01.

Théorème 3.3.7 (annexe A01)

Sous les conditions (C1) à (C5) et en utilisant le modèle 1, on démontre que le vecteur de dimension R des covariances asymptotiques entre $\hat{\theta}$ et le vecteur \hat{P}_k est donnée par :

$$\text{cov}(\hat{P}_k; \hat{\theta}) = \mathcal{S}_{\Theta,k} \times Z_k; \quad k = 1, 2, \dots, S \quad (10)$$

où $Z_k = (c_{1k}, c_{2k}, \dots, c_{Rk})^T$ est le vecteur de dimension R des coefficients de contrôle, $\mathcal{S}_{\Theta,k}$ la matrice $R \times R$ dont les éléments $\mathcal{S}_{\Theta,k}(i, j)$ sont donnés par

$$\mathcal{S}_{\Theta,k}(i, j) = -\frac{t_k}{n_k(\Lambda_{\Theta}^{(1)}/\Lambda_{\Theta})} \begin{cases} \frac{\gamma_k p_{ik}}{1+\theta c_{ik}} - \frac{(1+\rho_{\Theta,k} c_{ik}) p_{ik}^2}{\|\mathbf{1}_R\|_{B_{\Theta,k}^{-1}}^2 (1+\theta c_{ik})^2} & i = j \\ -\frac{(1+\rho_{\Theta,k} c_{ik}) p_{ik} p_{jk}}{\|\mathbf{1}_R\|_{B_{\Theta,k}^{-1}}^2 (1+\theta c_{ik})(1+\theta c_{jk})} & i \neq j \end{cases} \quad (11)$$

avec $(\Lambda_{\Theta}^{(1)}/\Lambda_{\Theta})$ le complément de Schur de Λ_{Θ} dans $\Lambda_{\Theta}^{(1)}$, $\mathbf{1}_{R,R} = \mathbf{1}_R \mathbf{1}_R^T$ la matrice $R \times R$ des unités et

$$\rho_{\Theta,k} = \frac{t_k \Delta_{\Theta,k} \theta^2 \gamma_k}{n_k}, \quad \Delta_{\Theta,k} = Z_k^T \Omega_{\Theta,k}^{-1} \mathbf{1}_R, \quad t_k = (1 - \|V_{\Theta,k}\|_{\Omega_{\Theta,k}^{-1}}^2)^{-1} \quad (12)$$

Sous les conditions de régularités déjà énoncées, on montre les égalités suivantes :

$$\text{cov}(\hat{P}_k; \hat{\theta}) = -\frac{1}{(\Lambda_{\Theta}^{(1)}/\Lambda_{\Theta})} \left[I_R - B_{\Theta,k}^{-1} \frac{\mathbf{1}_{R,R}}{\|\mathbf{1}_R\|_{B_{\Theta,k}^{-1}}^2} \right] B_{\Theta,k}^{-1} U_{\Theta,k}$$

et

$$B_{\Theta,k}^{-1} U_{\Theta,k} = \frac{t_k \gamma_k}{n_k} \Omega_{\Theta,k}^{-1} Z_k, \quad B_{\Theta,k}^{-1} \mathbf{1}_{R,R} \Omega_{\Theta,k}^{-1} = \gamma_k^{-1} [\Omega_{\Theta,k}^{-1} (\mathbf{1}_R \mathbf{1}_R^T + \rho_{\Theta,k} Z_k \mathbf{1}_R^T) \Omega_{\Theta,k}^{-1}]$$

On déduit par substitution la matrice de covariances suivante :

$$\text{cov}(\hat{P}_k; \hat{\theta}) = \mathcal{S}_{\Theta,k} \times Z_k$$

où $S_{\Theta,k}$ est la matrice $R \times R$ ci-dessous.

$$S_{\Theta,k} = -\frac{t_k}{n_k(\Lambda_{\Theta}^{(1)}/\Lambda_{\Theta})} \left[\gamma_k \Omega_{\Theta,k}^{-1} - \frac{1}{\|\mathbf{1}_R\|_{B_{\Theta,k}^{-1}}^2} \Omega_{\Theta,k}^{-1} (\mathbf{1}_{R,R} + \rho_{\Theta,k} Z_k \mathbf{1}_R^T) \Omega_{\Theta,k}^{-1} \right].$$

Corollaire 3.3.8 (annexe A01)

L'expression explicite de la covariance asymptotique entre $\hat{\theta}$ et \hat{p}_{mk} est donnée par

$$\begin{aligned} \frac{n_k(\Lambda_{\Theta}^{(1)}/\Lambda_{\Theta})}{t_k} [\text{cov}(\hat{p}_{mk}, \hat{\theta})] &= \left(-\frac{\gamma_k p_{mk}}{1 + \theta c_{mk}} + \frac{(1 + \rho_{\Theta,k} c_{mk}) p_{mk}^2}{\|\mathbf{1}_R\|_{B_{\Theta,k}^{-1}}^2 (1 + \theta c_{mk})^2} \right) c_{mk} \\ &+ \frac{(1 + \rho_{\Theta,k} c_{mk}) p_{mk}}{\|\mathbf{1}_R\|_{B_{\Theta,k}^{-1}}^2 (1 + \theta c_{mk})} \left(\Delta_{\Theta,k} - \frac{p_{mk} c_{mk}}{1 + \theta c_{mk}} \right) \end{aligned} \quad (13)$$

pour $m = 1, 2, \dots, R$.

Théorème 3.3.9 (annexe A01)

Sous les conditions (C1) à (C5) et en utilisant le modèle 1, on démontre que la matrice $R \times R$ des variances et covariances asymptotiques entre les vecteurs \hat{P}_k et \hat{P}_m est donnée par :

$$\text{cov}(\hat{P}_k; \hat{P}_m) = J_{k,m} - \sum_{j=1}^S \left[\left(\sum_{i=1}^S A_{i,j} J_{k,i} \right) \mathbf{1}_{R,R} \right] J_{j,m} \quad (14)$$

$k, m = 1, 2, \dots, S$ avec $J_{k,m}$ la matrice $R \times R$ donnée par le complément de Schur de (J_{Θ}/B_{Θ}) dans la matrice $(1 + R) \times (1 + R)$ définie par

$$\begin{bmatrix} (J_{\Theta}/B_{\Theta}) & U_{\Theta,m}^T B_{\Theta,m}^{-1} \\ B_{\Theta,k}^{-1} U_{\Theta,k} & \delta_{k,m} B_{\Theta,k}^{-1} \end{bmatrix},$$

$\delta_{k,m}$ le symbole de Kronecker et $A_{k,m}$ des coefficients réels définis au corollaire 3.2.7.

Dans le cas d'une mesure de sécurité routière appliquée à un seul site expérimental ($S = 1$) comportant plusieurs types d'accident ($R > 1$), les résultats sont plus affinés et permettent d'obtenir l'expression analytique de la covariance entre les différents risques d'accident ainsi que celle de leur variance. Les corollaires ci-après exposent le contenu de ces résultats.

Corollaire 3.3.10 (annexe A01)

En supposant ($S = 1$) et ($R > 1$), la covariance asymptotique entre les différentes composantes du vecteur \hat{P}_k est donnée par

$$\text{cov}(\hat{p}_{jk}, \hat{p}_{mk}) = \gamma_k^{-1} w_{jk} w_{mk} [\varepsilon_k \gamma_k c_{jk} c_{mk} - A_{k,k} \gamma_k^{-1} x_{jk} x_{mk}], \quad j \neq m = 1, 2, \dots, R \quad (15)$$

où $w_{jk} = \frac{p_{jk}}{1 + \theta c_{jk}}$ les éléments de la matrice diagonale $\Omega_{\Theta,k}^{-1}$, $\varepsilon_k = \frac{t_k}{n_k} \left[\theta^2 - (J_{\Theta}/B_{\Theta})^{-1} \frac{t_k \gamma_k^2}{n_k} \right]$, $x_{jk} = 1 + \varepsilon_k \Delta_{\Theta,k} \gamma_k c_{jk}$; ($j = 1, 2, \dots, R$).

Corollaire 3.3.11 (annexe A01)

En supposant ($S = 1$) et ($R > 1$), la variance asymptotique de chaque composante du vecteur \hat{P}_k est donnée par

$$\sigma^2(\hat{p}_{jk}) = \gamma_k^{-1} [w_{jk} + w_{jk}^2 (\varepsilon_k \gamma_k c_{jk}^2 - A_{k,k} \gamma_k^{-1} x_{jk}^2)] \quad (16)$$

où ($j = 1, 2, \dots, R$).

On montre dans ce cas particulier que

$$\text{cov}(\hat{P}_k, \hat{P}_k) = J_{k,k} - A_{k,k} J_{k,k} \mathbf{1}_{R,R} J_{k,k}$$

où la $R \times R$ matrice $J_{k,k}$ est donnée par

$$J_{k,k} = \gamma_k^{-1} [\Omega_{\Theta,k}^{-1} + \varepsilon_k \gamma_k \Omega_{\Theta,k}^{-1} Z_k Z_k^T \Omega_{\Theta,k}^{-1}].$$

D'où le résultat

$$\text{cov}(\hat{P}_k, \hat{P}_k) = \gamma_k^{-1} [\Omega_{\Theta,k}^{-1} + \Omega_{\Theta,k}^{-1} (\varepsilon_k \gamma_k Z_k Z_k^T - A_{k,k} \gamma_k^{-1} X_k X_k^T) \Omega_{\Theta,k}^{-1}]$$

où X_k le vecteur de dimension R et de coordonnées $x_{mk} = 1 + \varepsilon_k \Delta_{\Theta,k} \gamma_k c_{mk}$.

Corollaire 3.3.12 (annexe A01)

En utilisant le modèle 1, on a les résultats suivants :

- (i) $\Delta_{\Theta,k} = \sum_{m=1}^R \frac{p_{mk} c_{mk}}{1 + \theta c_{mk}}$,
- (ii) $\| \mathbf{1}_R \|_{B_{\Theta,k}^{-1}}^2 = \gamma_k^{-1} [\text{trace}(\Omega_{\Theta,k}^{-1}) + t_k \frac{\theta^2 \gamma_k \Delta_{\Theta,k}^2}{n_k}]$,
- (iii) $(\Lambda_{\Theta}^{(1)} / \Lambda_{\Theta}) = \sum_{k=1}^S \frac{\gamma_k^2}{n_k} \left[\left(\frac{\langle Z_k, P_k \rangle}{\theta} - \frac{\gamma_k t_k}{n_k} \| Z_k \|_{\Omega_{\Theta,k}^{-1}}^2 \right) + \frac{t_k^2 \Delta_{\Theta,k}^2}{n_k \| \mathbf{1}_R \|_{B_{\Theta,k}^{-1}}^2} \right]$.

Dans le cas particulier où la mesure est appliquée à un site expérimental avec plusieurs types d'accident, on obtient la variance asymptotique (voir annexe A06 p. 130) de l'effet moyen à l'aide de l'expression :

$$\sigma^2(\hat{\theta}) = \frac{1}{\frac{\gamma_k^2}{n_k} \left[\left(\frac{\langle Z_k, P_k \rangle}{\theta} - \frac{\gamma_k t_k}{n_k} \| Z_k \|_{\Omega_{\Theta,k}^{-1}}^2 \right) + \frac{t_k^2 \Delta_{\Theta,k}^2}{n_k \| \mathbf{1}_R \|_{B_{\Theta,k}^{-1}}^2} \right]}.$$

Ces derniers résultats permettent d'étendre le problème d'estimation ponctuelle à une approche d'estimation ensembliste en construisant des intervalles de confiance non seulement pour l'effet moyen de la mesure mais aussi pour les différents risques d'accident associés à cette mesure. Par ailleurs, ces résultats ont l'avantage d'offrir des outils simples et accessibles à tous les praticiens.

3.3.4 Construction d'intervalles de confiance et de tests d'hypothèse

Théorème 3.3.13 (annexe A04)

Sous les conditions (C1) à (C5) et en utilisant le modèle 1, on montre que les intervalles de confiance asymptotiques de niveau $(1 - \alpha)100\%$ associés aux paramètres $\beta = \log_e(\theta)$ et p_{jk} sont donnés par

$$\begin{aligned}
 IC_{1-\alpha}(\beta) &= \left[\hat{\beta} - \frac{z_{\alpha/2}}{(\Lambda_{\Theta_1}^{(1)}/\Lambda_{\Theta_1})^{1/2}}; \hat{\beta} + \frac{z_{\alpha/2}}{(\Lambda_{\Theta_1}^{(1)}/\Lambda_{\Theta_1})^{1/2}} \right] \\
 IC_{1-\alpha}(p_{jk}) &= \left[\max(0; \hat{p}_{jk} - z_{\alpha/2}\hat{\sigma}_{jk}); \min(1; \hat{p}_{jk} + z_{\alpha/2}\hat{\sigma}_{jk}) \right] \\
 &\quad (k = 1, 2, \dots, S; j = 1, 2, \dots, R)
 \end{aligned} \tag{17}$$

où $z_{\alpha/2}$ et $\hat{\sigma}_{jk}^2$ sont respectivement obtenus à partir de la table de la loi normale standard et des éléments diagonaux de la matrice $(SR) \times (SR)$ donnée par

$$(J_{\Theta_1}/\tau_{\Theta_1})^{-1} - (J_{\Theta_1}/\tau_{\Theta_1})^{-1} H_2^T R_{\Theta_1}^{-1} H_2 (J_{\Theta_1}/\tau_{\Theta_1})^{-1}$$

Corollaire 3.3.14 (annexe A04)

Sous les conditions (C1) à (C5) et en utilisant le modèle 1, on montre que l'intervalle de confiance asymptotique de niveau $(1 - \alpha)100\%$ associé au paramètre θ est donné par

$$IC_{1-\alpha}(\theta) = \left[EXP\left(\hat{\beta} - \frac{z_{\alpha/2}}{(\Lambda_{\Theta_1}^{(1)}/\Lambda_{\Theta_1})^{1/2}}\right); EXP\left(\hat{\beta} + \frac{z_{\alpha/2}}{(\Lambda_{\Theta_1}^{(1)}/\Lambda_{\Theta_1})^{1/2}}\right) \right] \tag{18}$$

Corollaire 3.3.15 (annexe A04)

Sous les conditions (C1) à (C5), on évalue l'efficacité moyenne de la mesure relativement au modèle 1 en testant

$$H_0 : \beta = 0 \quad \text{contre} \quad H_1 : \beta \neq 0$$

et en utilisant le fait qu'asymptotiquement on a

$$\frac{\hat{\beta}}{(\Lambda_{\Theta_1}^{(1)}/\Lambda_{\Theta_1})^{-1/2}} \stackrel{\text{Distr.}}{\sim} \mathcal{N}(0; 1) \tag{19}$$

Les résultats d'extraction de variances et de covariances ainsi que ceux relatifs à la construction des intervalles de confiance peuvent être étendus au modèle 2. Il suffit, pour cela, de définir et de décomposer convenablement la matrice d'information de Fisher associée à ce modèle. L'hypothèse $H_0 : \beta = 0$ est équivalente à comparer l'effet moyen de la mesure à l'unité c'est-à-dire à l'hypothèse $H_0 : \theta = 1$ consistant à dire qu'en moyenne l'efficacité de la mesure est nulle : il n'y a aucun impact statistiquement significatif. Si H_0 est rejetée et que $\theta < 1$ ou $\beta < 0$ alors la mesure tend à faire baisser en proportion le nombre moyen des accidents pour l'ensemble des sites et des types d'accidents analysés. Sinon, cette mesure semble a priori inefficace par rapport aux nombres d'accident observés et aux coefficients de contrôle utilisés.

Chapitre 4

Quelques exemples d'application

4.1 Introduction

Nous proposons, dans ce chapitre, les résultats (annexes *A02*, *A05*, *A06* et *A10*) concernant la mise en œuvre pratique de nos travaux de recherche théorique. Dans les annexes *A02* et *A05*, nous utilisons des données simulées. Notre ambition est d'analyser, par simulation, le comportement (existence et convergence) de la solution des équations de vraisemblance. Nous avons donc soumis l'approche d'estimation à des tableaux de données simulées en faisant varier, le nombre de sites expérimentaux, le nombre de types d'accident et les effectifs marginaux des tableaux. L'étude des erreurs quadratiques moyennes et des histogrammes des lois de probabilités permettant de pondérer les données de contrôle montrent de très bonnes approximations. Dans l'annexe *A06* (voir paragraphe 4.3.3), nous appliquons notre approche méthodologique à la transformation de sept carrefours en ronds-points traitée par Tanner. L'objectif est double. Il s'agit de montrer que les approches statistiques mises en place sont applicables moyennant une réorganisation des données de Tanner. Ensuite et surtout, il s'agit de pondérer aléatoirement les coefficients de contrôle en utilisant la distribution de probabilités d'une variable aléatoire dont la loi est à estimer au même titre que l'effet moyen de l'aménagement des sept carrefours. Dans les paragraphes 4.3.4 et 4.3.5 (voir annexe *A10*), nous présentons trois exemples d'étude de cas réels et récents. Les deux premiers exemples (installation de giratoire et modification d'une signalisation au sol) proviennent d'une série d'opérations d'aménagements routiers effectués dans le département du Pas de Calais (France). Ces opérations d'aménagement de sécurité s'insèrent dans la démarche sécurité des usagers sur les routes nationales existantes mises en place par les autorités de tutelle. Cette démarche vise à résoudre les problèmes de sécurité sur le réseau existant, en s'attachant à traiter en priorité les zones repérées - selon des critères pré-définis - comme les plus dangereuses du réseau. Le dernier exemple (paragraphe 4.3.5) concerne l'évaluation de l'effet de panneaux-réclame le long de certains échangeurs et de certaines autoroutes de Montréal de la province du Québec (Canada). Ces exemples montrent bien que nos approches statistiques s'appliquent à la fois à des tableaux d'accidents de petits effectifs et à des tableaux d'accidents de grands effectifs.

4.2 Données simulées

4.2.1 Principe de simulation des données d'accident avant-après avec contrôle

En pratique, l'on dispose de tableaux de données avant-après sur les sites expérimentaux et sur les zones de contrôle. Le vecteur Θ (voir section 2.3) qui définit l'effet moyen de la mesure et les différentes lois de probabilité associées aux contrôles individuels n'est pas connu et doit être estimé à l'aide de ces données. Dans cette partie nous exposons les travaux de recherche sur l'estimation numérique (annexes A02 et A05) de ces paramètres.

Notre intention n'est pas de faire une étude de cas mais de soumettre notre méthodologie à plusieurs configurations de données ; allant de tableaux dont les effectifs marginaux sont faibles à des tableaux dont les effectifs marginaux sont élevés. Pour ce faire, nous avons simulé, pour chaque modèle, plusieurs tableaux de données d'accident en faisant varier à la fois le nombre de sites expérimentaux, le nombre de types d'accident et l'effectif total par site. La vraie valeur Θ^0 du vecteur Θ est fixée quelle que soit la configuration.

Des programmes FORTRAN permettant de simuler des lois multinomiales ont été mis en place à cet effet. Une fois les tableaux de données obtenus, nous faisons appel à l'algorithme de Newton-Raphson via la bibliothèque numérique NAG [76] pour résoudre les équations du maximum de vraisemblance sous contraintes. Nous comparons ensuite la solution numérique de chaque configuration à la vraie valeur en utilisant différents critères statistiques.

4.2.2 Estimation et convergence numériques de solutions

Nous étudions l'existence d'une solution aux équations du maximum de vraisemblance sous contraintes ; pour le modèle 1 par exemple (voir annexe A05), les équations de vraisemblance sous contraintes sont :

$$\left\{ \begin{array}{l} \sum_{k=1}^S \sum_{j=1}^R \frac{n_{2jk} - \hat{\theta} \hat{c}_k n_{1jk}}{1 + \hat{\theta} \hat{c}_k} = 0; \\ n_{.jk} - \frac{n_k \hat{p}_{jk} (\hat{\theta} c_{jk} + 1)}{1 + \hat{\theta} \hat{c}_k} = 0, \\ (k = 1, 2, \dots, S; \quad j = 1, 2, \dots, R); \\ \hat{c}_k = \sum_{j=1}^R c_{jk} \hat{p}_{jk}, \quad (k = 1, 2, \dots, S); \end{array} \right. \quad (1)$$

sous les contraintes

$$\left\{ \begin{array}{l} \hat{\theta} > 0; \quad 0 < \hat{p}_{jk} < 1; \\ \sum_{j=1}^R \hat{p}_{jk} = 1, \quad (k = 1, 2, \dots, S). \end{array} \right.$$

Les équations de vraisemblance relatives au modèle 2 (voir annexe A02) sont identiques à un vecteur additif près. Ensuite, nous montrons la convergence de cette solution vers la vraie valeur du paramètre en utilisant l'erreur quadratique moyenne définie par :

$$MSE(\hat{\Theta} - \Theta^0) = \frac{1}{1 + SR} \sum_{m=1}^{1+SR} (\hat{\Theta}_m - \Theta_m^0)^2.$$

Bien que la dimension des espaces de solution soit une fonction croissante du nombre de sites expérimentaux et du nombre de types d'accident, les résultats (voir annexes A02, A05) obtenus montrent que $MSE \sim O(10^{-d})$ avec $d > 1$. Un exemple de ces résultats est présenté dans la figure 4.1. Dans ce graphe, on a fixé le nombre de types d'accidents à 10 et fait varier le nombre de sites expérimentaux de 5 à 20. Pour un couple (S, R) fixé, on estime $\hat{\Theta}$ pour différentes tranches ("Accident slices") de valeurs de nombres d'accidents par site expérimental : $n_k \in \{]0; 30[,]30; 50[, \dots,]11000; 13000[\}$. On calcule ensuite l'erreur quadratique moyenne en utilisant les équations de vraisemblance liées aux modèles. La figure 4.1 (voir annexe A06 p. 130) utilise le modèle 1 et permet de montrer que pour la première tranche d'accidents (faibles effectifs d'accidents : $n_k \in \{]0; 30[$), la MSE est de l'ordre de 10^{-2} alors qu'elle est de l'ordre de 10^{-5} pour les tranches plus élevées (grands effectifs d'accidents : $n_k > 8000$).

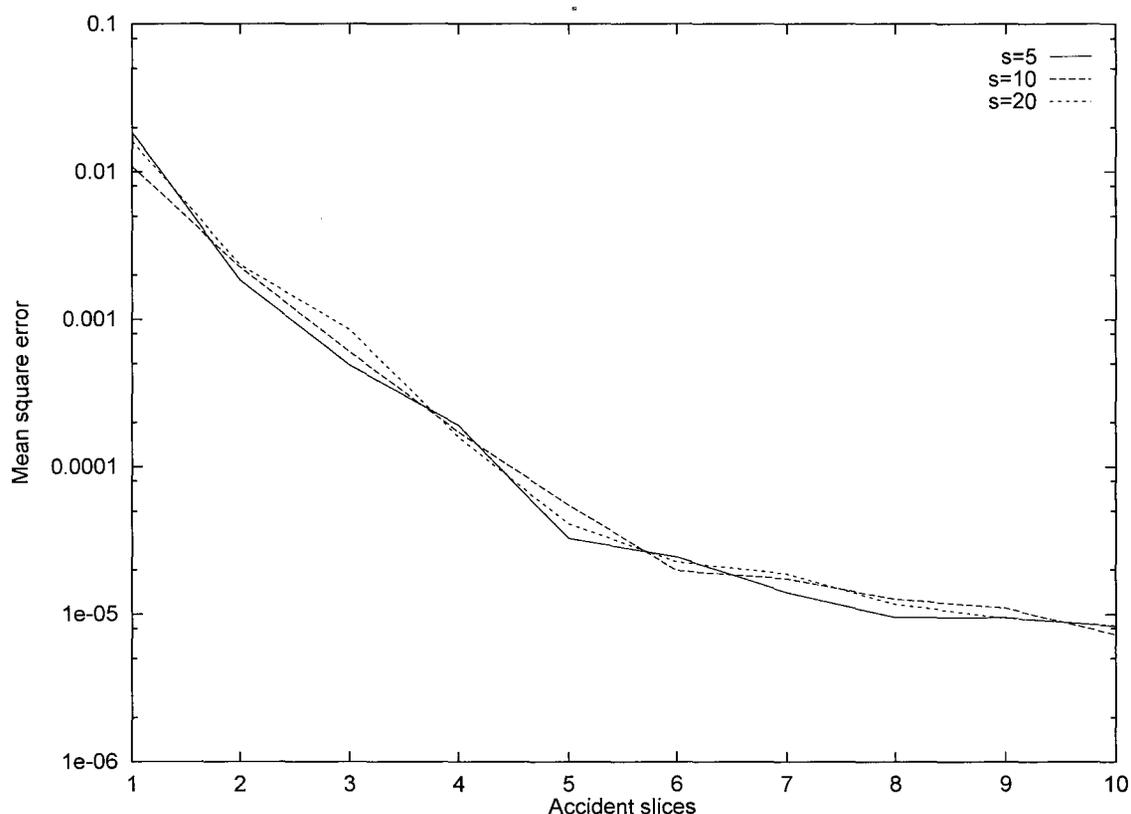


FIG. 4.1 – $R=10$.

En plus de l'étude de l'existence et de la convergence de l'estimateur du maximum de vraisemblance, nous avons examiné la convergence du paramètre effet moyen de la mesure ainsi que celle des coefficients de contrôle moyens associés aux zones de contrôle. L'étude consacrée aux contrôles moyens (annexe A05 p. 110) est motivée par le rôle important joué par ces contrôles dans la conception des modèles proposés, dans l'interprétation de l'effet moyen et donc de l'évaluation de cette mesure. La définition de ces contrôles moyens passe par la pondération

des coefficients de contrôle individuels. De ce fait, pour chaque site, nous avons analysé les lois de probabilité servant de poids aux coefficients de contrôle individuels.

4.2.3 Analyse des lois de probabilité liées aux différents risques d'accidents

Les résultats sont exposés dans les annexes (A02 p. 65 et A05 p. 110). Dans ces deux travaux, nous avons mis l'accent sur l'estimation des lois de probabilité liées aux différents risques d'accident et utilisées pour pondérer les coefficients individuels de chaque zone de contrôle. En effet, les poids p_{jk} et les coefficients moyens $c_k = \sum_{k=1}^R c_{jk} p_{jk}$, par construction et par définition, représentent la loi de probabilité et l'espérance mathématique des variables aléatoires associées aux coefficients de contrôle de chaque zone k . Il est donc important de bien estimer la loi de ces variables. C'est cette question cruciale qui est examinée dans les annexes A02 (pour le modèle 2) et A05 (pour le modèle 1). On compare par simulation l'histogramme et l'espérance mathématique des vraies distributions à ceux obtenus par notre approche. On montre que lorsque les effectifs marginaux d'accident augmentent, on obtient une très bonne approximation des lois de probabilité associées aux coefficients de contrôle individuels de chaque zone k . Les graphiques 4.2 et 4.3 montrent que, pour S et R fixés respectivement à 2 et à 5, il y a une bonne adéquation entre la loi de probabilité théorique et celle estimée par notre approche lorsque l'on passe de petits effectifs à de grands effectifs d'accidents.

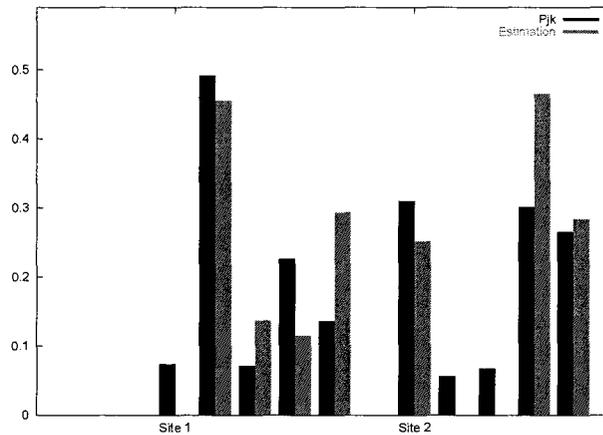
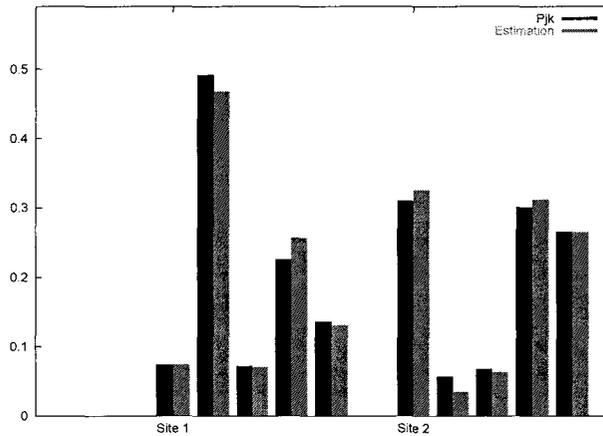


FIG. 4.2 – $S= 2.$; $R=5.$; $n_k \in]0; 30[$


 FIG. 4.3 – S=2.;R=5.; $n_k \in]1000; 3000[$

4.3 Quelques études de cas réels

Nous exposons (voir annexe A10) les premières applications de nos approches théoriques à des études de cas réels. Ces mesures de sécurité routière concernent un seul site expérimental associé à une zone de contrôle et en présence de différents types de risque d'accidents. Les quatre exemples d'application sont extraits des annexes A06 et A10 et utilisent le modèle de l'annexe A05. Le premier exemple (paragraphe 4.3.3 et annexe A06) est une application de nos approches aux données d'accident de Tanner. Les deux exemples suivants proviennent d'une série d'opérations d'aménagements routiers effectués dans le département du Pas de Calais (France). Le dernier exemple d'application concerne l'installation de panneaux publicitaires le long de certains échangeurs et de certaines autoroutes de Montréal (Québec, Canada).

4.3.1 Notation et modèle

En utilisant l'annexe A05, on note R ($R > 1$) le nombre de types d'accidents, x_{1j} (resp. x_{2j}) le nombre d'accidents du type j ($j = 1, 2, \dots, R$) observés avant (resp. après) la mise en place de l'aménagement routier, $n = \sum_{j=1}^R (x_{1j} + x_{2j})$ le nombre total d'accidents enregistrés dans la zone aménagée durant les deux périodes d'observation, c_1, c_2, \dots, c_R les coefficients de contrôle, X_{1j} (resp. X_{2j}) la variable aléatoire donnant le nombre d'accidents du type j de la période avant (resp. après) l'aménagement, dont une valeur observée est x_{1j} (resp. x_{2j}). D'après le modèle 1 (voir section 2.3.1), la loi de probabilité conditionnelle (conditionnellement au nombre total d'accidents observés dans le site expérimental et aux coefficients de contrôle) de la paire (X_1, X_2) où $X_1 = (X_{11}, X_{12}, \dots, X_{1R})^T$ et $X_2 = (X_{21}, X_{22}, \dots, X_{2R})^T$ est donnée par la loi multinomiale, de dimension $2R$, de paramètres $n, \pi_{1j}(\Theta), \pi_{2j}(\Theta)$ avec

$$\pi_{1j}(\Theta) = \frac{p_j}{1 + \theta \sum_{m=1}^R p_m c_m}, \quad \pi_{2j}(\Theta) = \frac{\theta c_j p_j}{1 + \theta \sum_{m=1}^R p_m c_m}, \quad j = 1, 2, \dots, R, \quad (2)$$

sous les contraintes

$$p_j > 0, \quad \text{et} \quad \sum_{j=1}^R p_j = 1; \quad (3)$$

$\Theta = (\theta, P^T)^T$, de dimension $(1 + R)$, est le vecteur des paramètres inconnus du modèle où θ ($\theta > 0$) décrit l'effet moyen (ou indice d'efficacité moyenne) de l'aménagement et où $P = (p_1, p_2, \dots, p_R)^T$, de dimension R , est un vecteur de probabilités. Il s'agit d'estimer Θ .

4.3.2 Estimation de l'effet moyen et des différentes lois de probabilité

Cette estimation est obtenue à l'aide de la méthode du maximum de vraisemblance. Il s'agit de maximiser la fonction log-vraisemblance donnée par

$$\begin{aligned} \mathcal{L}(X_1, X_2; \Theta) = & \text{constante} \\ & + \sum_{j=1}^R \left\{ X_{\bullet j} \log_e(p_j) + X_{2j} \log_e(\theta) - X_{\bullet j} \log_e \left(1 + \theta \sum_{m=1}^R c_m p_m \right) \right\} \end{aligned} \quad (4)$$

avec $X_{\bullet j} = X_{1j} + X_{2j}$. En prenant les dérivées partielles premières de $\mathcal{L}(X_1, X_2; \Theta)$ par rapport à $\theta, p_1, p_2, \dots, p_R$ et en égalant à zéro, on montre que l'estimateur du maximum de vraisemblance $\hat{\Theta}$ de Θ , s'il existe, est solution du système d'équations non linéaires :

$$\begin{cases} \sum_{j=1}^R \{ X_{2j} - X_{\bullet j} \frac{\hat{\theta} \sum_{m=1}^R c_m \hat{p}_m}{1 + \hat{\theta} \sum_{m=1}^R c_m \hat{p}_m} \} = 0 \\ X_{\bullet j} - n \frac{\hat{p}_j (1 + c_j \hat{\theta})}{1 + \hat{\theta} \sum_{m=1}^R c_m \hat{p}_m} = 0, \quad (j = 1, 2, \dots, R); \\ 0 < \hat{p}_j \text{ avec } \sum_{j=1}^R \hat{p}_j = 1. \end{cases} \quad (5)$$

La mise en œuvre de la méthode du maximum de vraisemblance sous contraintes passe par la résolution numérique ce système. Cette résolution fait appel à une procédure itérative du type Newton-Raphson. Le point de convergence de cette procédure est une estimation ponctuelle de $\hat{\theta}, \hat{p}_1, \hat{p}_2, \dots, \hat{p}_R$. Pour apprécier l'impact de l'aménagement routier sur les nombres d'accidents, on compare la valeur estimée du paramètre θ à 1. Une valeur de θ significativement supérieure à 1 indique que l'aménagement routier a eu un effet non bénéfique sur le nombre moyen des accidents de la zone aménagée par rapport aux données de la zone de contrôle. Au contraire, une valeur de θ significativement inférieure à 1 témoigne d'une baisse en proportion du nombre des accidents de la zone aménagée par rapport à la zone de contrôle. Enfin, une valeur de θ non significativement différente de 1 indique que l'aménagement routier n'a pas eu d'effet par rapport à l'évolution moyenne du nombre des accidents observés dans la zone de contrôle : on parle d'efficacité moyenne nulle. Pour mieux estimer les paramètres des modèles et évaluer l'impact de la mesure, on construit des intervalles de confiance en utilisant les résultats de l'annexe A04. On obtient les intervalles de confiance à $(1 - \alpha)100\%$ suivants :

$$\begin{aligned}
IC_{1-\alpha}(\theta) &= \left[\max(0; \hat{\theta} - \epsilon_{\alpha/2}\sigma(\hat{\theta})); \hat{\theta} + \epsilon_{\alpha/2}\sigma(\hat{\theta}) \right]; \\
IC_{1-\alpha}(\hat{p}_j) &= \left[\max(0; \hat{p}_j - \epsilon_{\alpha/2}\sigma(\hat{p}_j)); \min(1; \hat{p}_j + \epsilon_{\alpha/2}\sigma(\hat{p}_j)) \right], \\
(j = 1, 2, \dots, R),
\end{aligned} \tag{6}$$

où $\epsilon_{\alpha/2}$ est obtenu de façon usuelle à l'aide de la table de la loi normale centrée et réduite, $\sigma(\bullet)$ l'erreur standard (E.S) c'est-à-dire l'écart-type estimé de chaque composante du paramètre Θ . Les valeurs ponctuelles de $\hat{\theta}$, \hat{p}_1 , \hat{p}_2 , \dots , \hat{p}_R sont ensuite utilisées pour obtenir une estimation des probabilités conditionnelles $\pi_{1j}(\Theta)$ et $\pi_{2j}(\Theta)$ et apprécier l'incidence de l'aménagement routier sur les risques d'accidents avant et après la mise en place de celui-ci.

4.3.3 Données de Tanner 1958 (annexe A06)

Dans l'annexe A06 p. 130, nous réexaminons les données fournies par Tanner (1958). Il s'agit, dans son travail, de données provenant de la transformation de sept carrefours en ronds-points. Dans son exemple, chaque rond-point est un site expérimental auquel il associe une zone de contrôle. Il propose une méthode pour évaluer l'effet moyen de cette transformation en supposant que chaque coefficient de contrôle est sans erreur. L'ensemble de ces sept sites appartenant à une même entité géographique, nous proposons donc d'appliquer notre approche multivariée à ces données de la façon suivante : chaque carrefour transformé (chaque site expérimental) représente un type d'accident. L'entité géographique comporte donc sept types de risque d'accidents. Les sept coefficients individuels ne sont plus sans erreur mais sont pondérés par des poids aléatoires qu'il faut estimer au même titre que l'effet moyen de la transformation. Les résultats obtenus à l'aide de notre approche sont les suivants :

Tableau 4.1 : Estimation des paramètres du modèle

paramètres	estimation (E.S)	B_{inf} à 95%	B_{sup} à 95%
$\hat{\theta}$	0.2833 (0.0800)	0.1265	0.4401
\hat{p}_1	0.1103 (0.0389)	0.0340	0.1865
\hat{p}_2	0.1356 (0.0417)	0.0538	0.2173
\hat{p}_3	0.2173 (0.0524)	0.1145	0.3200
\hat{p}_4	0.2568 (0.0465)	0.1656	0.3479
\hat{p}_5	0.1545 (0.0449)	0.0664	0.2425
\hat{p}_6	0.0552 (0.0267)	0.0028	0.1075
\hat{p}_7	0.0703 (0.0295)	0.0124	0.1281

On obtient la même estimation (0,2833) que Tanner c'est-à-dire une efficacité moyenne de 71,67%. L'écart-type estimé est 0,08 obtenu en utilisant l'approche par complément de Schur. Cela évite le calcul d'un coefficient complexe proposé par Tanner pour estimer la variance de l'effet moyen. L'intervalle [0,1265; 0,4401] contient la vraie valeur de l'effet moyen réel du changement avec une probabilité de 0,95. En utilisant cet intervalle de confiance au niveau 95%

associé au paramètre θ , on teste la significativité de l'effet moyen de la transformation de sept carrefours en ronds-points. En effet, puisque $1 \notin [0, 1265; 0, 4401]$, on rejette l'hypothèse nulle $H_0 : \theta = 1$ (ou de façon équivalente $\log(\theta) = 0$) contre $H_1 : \theta \neq 1$ avec un risque 5%, concluant ainsi à une efficacité moyenne significative du changement apporté.

4.3.4 Évaluation d'aménagements de sécurité sur certaines routes nationales (annexe A10)

Les deux exemples d'application de cette sous-section concernent certaines opérations d'aménagements routiers effectuées dans le département du Pas de Calais (France). La première application est relative à l'aménagement, de 1997 à 1998, d'un giratoire sur la route nationale 43 au niveau de la commune de Vermelles. La deuxième étude traite de la modification du marquage au sol sur un tronçon de la route nationale 17 afin d'empêcher les dépassements anarchiques de certains usagers. Ce deuxième aménagement routier a eu lieu en 1999 et se situe au niveau des communes de Vimy et d'Avion.

Installation d'un giratoire sur route nationale

Un giratoire (rond-point) a été aménagé au niveau de la commune de Vermelles (département du Pas de Calais) au croisement de la route départementale 75 (RD75) et de la route nationale 43 (RN43). Les travaux ont duré deux ans (de 1997 à 1998). Afin d'évaluer cet aménagement de sécurité, on a relevé le nombre des accidents mortels (M), des accidents graves (G), et des accidents légers (L) au niveau de la zone aménagée (site expérimental : Vermelles) au cours des quatre années précédant le début des travaux et des quatre années suivant la fin de ceux-ci. Les données d'accident cumulées sur chaque période de 4 ans sont consignées dans le tableau 4.2 suivant :

Tableau 4.2 : Données d'accidents du site expérimental

	période avant (4 ans)			période après (4 ans)			total
	M	G	L	M	G	L	
zone aménagée	3	5	6	0	0	3	17

Pour des périodes d'observation de longueur 4 ans, les accidents mortels sont passés de 3 avant à 0 après. Il y a eu donc 3 accidents avec au moins un tué au cours des 4 années précédant les travaux et 0 accident avec au moins un tué au cours des 4 années suivant les travaux. De même, les accidents graves (accidents avec au moins un blessé grave et aucun tué) sont passés de 5 avant à 0 après. Quant aux accidents légers (accidents sans tué ni blessé grave) ils ont baissé de moitié. Pour analyser et apprécier convenablement l'effet moyen de l'aménagement ainsi que son impact probable sur les risques de gravité, on combine, pour des périodes de même longueur, les données d'accident du site expérimental à celles relevées sur une portion de la RN43 servant de zone de contrôle. Les données de cette zone de contrôle sont fournies sous forme de rapports de nombres d'accidents (tableau 4.3) :

Tableau 4.3 : Valeurs des coefficients de contrôle

	c_1	c_2	c_3
zone de contrôle	0.6666	0.3930	0.5000

Par exemple la valeur 0.6666 est le rapport du nombre total des accidents mortels enregistrés sur la portion de la RN43 servant de zone de contrôle au cours des quatre années suivant la fin des travaux sur le nombre des accidents mortels durant les quatre années précédant le début des travaux. On note, de façon générale, une baisse du nombre de chaque type d'accident après les travaux au niveau du site aménagé et une baisse en proportion du nombre de ceux-ci sur la portion de la route nationale servant de contrôle. Par exemple pour les accidents graves, on est passé au niveau du site expérimental de 5 accidents graves au cours des quatre années précédant l'aménagement à 0 au cours des quatre années le suivant, pendant que sur toute la portion de route nationale servant de contrôle, on a observé un rapport de contrôle de 0.3930 soit une baisse en proportion d'environ 60, 70% des accidents graves.

Les estimations de l'effet moyen de l'aménagement et de la loi de probabilités servant de poids aux coefficients de contrôle sont disponibles dans le tableau 4.4. Les valeurs entre parenthèses représentent l'erreur standard (E.S) c'est-à-dire l'écart-type estimé de chaque paramètre du modèle. La valeur estimée de l'effet moyen de cet aménagement est de 0,4328 soit une baisse moyenne en proportion de $(1 - 0,4328) \times 100 = 56,72\%$ des accidents de la zone aménagée par rapport à la tendance moyenne de l'évolution des accidents survenus dans la zone de contrôle au cours de la période de quatre ans suivant l'installation du giratoire. L'intervalle de confiance au niveau 95% associé à l'effet moyen est $[0; 0,9738]$ avec une borne supérieure plus petite que 1. Ainsi, avec 95 chances sur 100, un tel aménagement a apporté, dans la période de quatre ans après sa mise en place, une baisse moyenne en proportion statistiquement significative du nombre des accidents de la zone aménagée comparativement à la tendance moyenne du nombre des accidents de la zone de contrôle.

Les valeurs estimées des paramètres p_1 , p_2 et p_3 à l'aide de la méthode du maximum de vraisemblance sont respectivement 0,1663, 0,3052 et 0,5285. Ces valeurs donnent un ordre de grandeur du niveau de risque de chaque type d'accidents de la zone aménagée au cours des quatre années précédant l'aménagement si l'on tient compte des données de la zone de contrôle. Par exemple, après ajustement à l'aide des coefficients de contrôle, on estime en moyenne que 16,63% des accidents survenus dans la zone aménagée au cours des quatre années précédant l'installation du giratoire ont été mortels. L'intervalle de confiance au niveau 95% $([0; 0,3419])$ associé au paramètre p_1 montre qu'il est fort probable que le risque d'accidents mortels de la zone aménagée ne soit pas statistiquement significativement différent de 0 au cours des quatre années précédant l'installation du giratoire.

Tableau 4.4 : Estimation des paramètres du modèle

paramètres	estimation (E.S)	B_{inf} à 95%	B_{sup} à 95%
$\hat{\theta}$	0.4328 (0.2760)	0	0.9738
\hat{p}_1	0.1663 (0.086)	0	0.3419
\hat{p}_2	0.3052 (0.1175)	0.0749	0.5355
\hat{p}_3	0.5285 (0.1329)	0.2680	0.7889

En revanche, il est fort probable que le risque d'accidents graves de la zone aménagée soit statistiquement significativement différent de 0 au cours des quatre années précédant l'aménagement du giratoire car l'intervalle de confiance au niveau 95% associé à p_2 ne contient pas 0.

Pour apprécier, au niveau de la zone aménagée, l'incidence de l'installation du giratoire sur les différents risques d'accidents, on utilise la valeur des probabilités conditionnelles $\pi_{2j}(\hat{\Theta})$ et $\pi_{1j}(\hat{\Theta})$. Le calcul de ces risques d'accidents par période de temps utilise les valeurs estimées des paramètres et celles des coefficients de contrôle. Les valeurs obtenues sont disponibles dans le tableau 4.5 :

Tableau 4.5 : Estimation des risques d'accidents par période de temps

différents types de risque	avant : $\pi_{1j}(\hat{\Theta})$	après : $\pi_{2j}(\hat{\Theta})$
accidents mortels	0.1370	0.0395
accidents graves	0.2513	0.0426
accidents légers	0.4352	0.0942

Par exemple, si l'on estime l'effet moyen de l'aménagement à 0,4328 et si les coefficients de contrôle sont ceux donnés par le tableau 4.3 alors au cours de la période des huit années d'analyse, le risque d'avoir un accident mortel avant l'aménagement de la zone de Vermelles est de l'ordre de $0,1370 = 0,1663/(1 + 0,4328 \times 0,4950)$ où $0,4950 = 0,1663 \times 0,6666 + 0,3052 \times 0,3930 + 0,5285 \times 0,5000$. De même, conditionnellement aux données de contrôle et aux valeurs estimées des paramètres du modèle, on a 3,95% de chances, soit $0,0395 = 0,1663 \times 0,4328 \times 0,6666/(1 + 0,4328 \times 0,4950)$, qu'un accident qui s'est produit pendant les huit années d'analyse soit un accident mortel survenu au cours de la période des quatre années suivant l'aménagement. Dans les mêmes conditions, on a 25,13% (resp. 4,26%) de chances qu'un accident qui s'est produit pendant les huit années d'analyse soit un accident grave survenu au cours de la période des quatre années précédant (resp. suivant) l'aménagement du giratoire. Les valeurs du risque d'accidents légers avant et après sont obtenus de façon similaire en utilisant 0,5285 et 0,5000. Globalement, selon les données disponibles, la longueur des périodes utilisées et les valeurs estimées des paramètres, on pense vraisemblablement que cet aménagement a eu localement (au niveau de la zone aménagée) un impact significatif sur les différents types de risque d'accidents de l'étude.

Analyse de la modification d'une signalisation au sol

Un autre aménagement routier a été effectué en 1999 sur la route nationale 17 (RN17) au niveau des communes de Vimy et d'Avion du département Pas de Calais. En effet, la route, entre ces deux communes, était une trois voies classique, à deux sens. L'aménagement a consisté à ce que l'on appelle un "marquage à l'Italienne". On modifie le marquage au sol pour qu'il soit impossible de doubler en même temps dans chaque sens. Sur une courte distance, on donne deux voies pour un sens (on peut alors doubler dans ce sens) et une seule voie sur l'autre sens séparées par une ligne continue. Plus loin, on inverse pour que l'autre sens puisse doubler et ainsi de suite. Les données relevées au niveau de la zone aménagée (Vimy-Avion) sont les suivantes (tableau 4.6) :

Tableau 4.6 : Données d'accidents du site expérimental

	période avant (4 ans)			période après (4 ans)			total
	<i>M</i>	<i>G</i>	<i>L</i>	<i>M</i>	<i>G</i>	<i>L</i>	
zone aménagée	4	4	16	1	1	7	33

Pour des périodes d'observation de longueur 4 ans, les accidents mortels sont passés de 4 avant à 1 après. Il y a eu donc 4 accidents avec au moins un tué au cours de la période précédant les travaux et 1 accident avec au moins un tué au cours de la période suivant ceux-ci. Pendant ce temps les accidents légers (accidents sans tué ni blessé grave) répertoriés ont baissé plus que de moitié. Pour les mêmes longueurs de périodes, on observe, sur la portion de la RN17 servant de zone de contrôle, les rapports de nombres d'accidents suivants (tableau 4.7) :

Tableau 4.7 : Valeurs des coefficients de contrôle

	c_1	c_2	c_3
zone de contrôle	0.5190	0.4220	0.5600

Globalement on note une baisse en proportion du nombre des accidents de la zone de contrôle, entre la période de 4 ans suivant les travaux et la période de 4 ans précédant ceux-ci. Les résultats numériques relatifs à ces données d'accidents sont les suivants (tableau 4.8) :

Tableau 4.8 : Estimation des paramètres du modèle

paramètres	estimation (E.S)	B_{inf} à 95%	B_{sup} à 95%
$\hat{\theta}$	0.7054 (0.2759)	0.1646	1.2462
\hat{p}_1	0.1525 (0.0643)	0.0264	0.2785
\hat{p}_2	0.1605 (0.0672)	0.0287	0.2922
\hat{p}_3	0.6870 (0.1023)	0.4864	0.8875

L'indice d'efficacité moyenne estimé à l'aide des données des tableaux 4.6 et 4.7 est de 0,7054, ce qui correspond à une baisse moyenne en proportion de $29,46\% = (1 - 0,7054) \times 100\%$ du nombre d'accidents de la zone aménagée (Vimy-Avion) par rapport à la tendance moyenne du nombre des accidents de la zone de contrôle. Les valeurs estimées des trois paramètres \hat{p}_1 , \hat{p}_2 et \hat{p}_3 sont respectivement 0,1525, 0,1605 et 0,6870. Ces valeurs permettent d'apprécier le niveau probable des trois types de risque d'accidents du site expérimental au cours des quatre années précédant la mise en place des travaux de marquage au sol. Les intervalles de confiance au niveau 95% associés à ces trois paramètres montrent qu'au cours des quatre années précédant les travaux de marquage au sol, on estime, avec une probabilité de 0,95, que le risque d'observer un accident mortel est compris entre 2,64 et 27,85%, celui d'observer un accident grave entre 2,87% et 29,22%, celui d'observer un accident léger entre 48,64% et 88,75%.

On teste la significativité de l'effet moyen de ce type d'aménagement en utilisant l'intervalle de confiance au niveau 95% associé au paramètre θ . Cet intervalle de confiance montre qu'on a une fourchette de valeurs dont la borne inférieure (resp. supérieure) est strictement plus grande que

0 (resp. 1). En effet, comme $1 \in [0.1646; 1.2462]$, on ne peut rejeter l'hypothèse $H_0 : \theta = 1$ contre $H_1 : \theta \neq 1$ avec une erreur de première espèce de 5%. La valeur estimée de θ pouvant excéder 1, il est fort probable que ce type d'aménagement routier n'apporte aucun impact significatif sur le nombre moyen des accidents au niveau de la zone aménagée par rapport à l'évolution du nombre moyen des accidents de la zone de contrôle. Même si dans le cas présent, on note une baisse en proportion de 29,46% du nombre moyen des accidents de la zone aménagée (Vimy-Avion), le test ci-dessus montre que la valeur de l'indice d'efficacité moyenne n'est pas statistiquement significativement différente de 1 pour conclure de façon systématique à une efficacité avérée de ce type de marquage au sol.

Cependant, on peut apprécier l'impact de ce type de marquage sur les différents types de risque d'accidents par période de temps au niveau du site expérimental si l'on admet que l'effet moyen de cet aménagement sur les nombres d'accidents est de l'ordre de 0,7054 et que les rapports de nombres d'accidents sont ceux donnés dans le tableau 4.7. On utilise les valeurs estimées des probabilités conditionnelles $\pi_{1j}(\hat{\Theta})$ et $\pi_{2j}(\hat{\Theta})$ du modèle multinomial. Ces valeurs estimées sont présentées dans le tableau 4.9 :

Tableau 4.9 : Estimation des risques d'accidents par période de temps

différents types de risque	avant : $\pi_{1j}(\hat{\Theta})$	après : $\pi_{2j}(\hat{\Theta})$
accidents mortels	0.1109	0.0400
accidents graves	0.1167	0.0347
accidents légers	0.4996	0.1973

L'interprétation de ces probabilités est identique à celle de l'installation du giratoire. Si on estime à 0,7054 l'effet moyen de ce type de marquage sur les nombres d'accidents de la zone aménagée et après ajustement avec les coefficients de contrôle alors on a 11,09% (resp. 4%) de chances qu'un accident qui s'est produit pendant les huit années d'analyse soit un accident mortel survenu au cours de la période des quatre années précédant (resp. suivant) l'aménagement. Dans les mêmes conditions, on a 11,67% (resp. 3,47%) de chances qu'un accident qui s'est produit pendant les huit années d'analyse soit un accident grave survenu au cours de la période des quatre années précédant (resp. suivant) l'aménagement du giratoire. Les valeurs du risque d'accidents légers avant et après sont obtenus de façon similaire en utilisant 0,6870 et 0,5600. Ces valeurs estimées des risques d'accidents sur les huit années d'analyse du marquage au sol donnent deux enseignements : le premier enseignement est qu'on a, par période, sensiblement la même probabilité d'avoir un accident mortel ou grave ; le deuxième enseignement est que ce type de marquage au sol semble avoir localement un impact probable significatif sur les différents types de risque d'accidents de la zone aménagée.

4.3.5 Évaluation de l'effet de panneaux-réclame le long de certaines autoroutes (annexe A10)

Des panneaux-réclame ont été installés en octobre 1993 au niveau de l'échangeur Décarie (jonction des autoroutes 15 et 40) de l'île de Montréal. Dans le tableau 4.10, nous présentons

les nombres d'accidents survenus dans cet échangeur pour les 12 mois avant et les 12 mois après l'apparition des premiers panneaux-réclame, soit de novembre 1992 à octobre 1993 et de décembre 1993 à novembre 1994.

Tableau 4.10 : **Données d'accidents de l'échangeur Décarie**

	période avant (12 mois)			période après (12 mois)			total
	<i>MG</i>	<i>L</i>	<i>DMS</i>	<i>MG</i>	<i>L</i>	<i>DMS</i>	
échangeur Décarie	11	50	325	12	43	300	741

Les accidents mortels ou graves (*MG*) sont passés de 11 avant l'installation des panneaux-réclame à 12 après. Il y a eu donc 11 accidents avec au moins un tué ou un blessé grave au cours des 12 mois précédant l'installation des panneaux-réclame au niveau de l'échangeur contre 12 accidents avec au moins un tué ou blessé grave au cours de la période des 12 mois suivant celle-ci. De même, les accidents légers (*L*) sont passés de 50 avant à 43 après et les accidents avec dommages matériels seulement (*DMS*) sont passés de 325 avant à 300 après. Afin d'évaluer l'effet des panneaux-réclame sur les accidents dans cet échangeur et apprécier son incidence sur leur gravité, on a relevé les mêmes types d'accidents sur une zone de contrôle constituée de l'échangeur Anjou et d'un tronçon de l'autoroute 40 secteur Est. Les données d'accident ainsi obtenues ont permis de dresser les coefficients de contrôle du tableau 4.11.

Tableau 4.11 : **Valeurs des coefficients de contrôle**

	c_1	c_2	c_3
zone de contrôle	1	0.9428	1.0424

Ces coefficients de contrôle font apparaître une stabilité des collisions mortelles ou graves ($c_1 = 1$) au niveau de la zone de contrôle, une petite baisse en proportion de 5,72% des accidents légers ($c_2 = 0,9428$) de la période après par rapport à la période avant, et une légère augmentation en proportion de 4,24% des accidents avec dommages matériels seulement ($c_3 = 1,0424$). Les paramètres estimés à partir des tableaux 4.10 et 4.11 sont présentés dans le tableau 4.12.

Tableau 4.12 : **Estimation des paramètres du modèle**

paramètres	estimation (E.S)	B_{inf} à 95%	B_{sup} à 95%
$\hat{\theta}$	0.8946 (0.0657)	0.7656	1.0235
\hat{p}_1	0.0314 (0.0065)	0.0187	0.0442
\hat{p}_2	0.1307 (0.0131)	0.1051	0.1563
\hat{p}_3	0.8379 (0.0252)	0.7884	0.8875

En moyenne, après ajustement à l'aide des coefficients de contrôle, on estime que le nombre d'accidents a diminué de 10,54%, soit $(1 - 0,8946) \times 100\%$, sur la période de 12 mois après par rapport à la période de 12 mois avant l'apparition des premiers panneaux-réclame. Cependant, cette diminution n'est pas statistiquement significativement différente de 0. En effet, l'intervalle de confiance au niveau 95% associé au paramètre θ contient la valeur 1.

Tableau 4.13 : Estimation des risques d'accidents par période de temps

différents types de risque	avant : $\pi_{1j}(\hat{\Theta})$	après : $\pi_{2j}(\hat{\Theta})$
accidents MG	0.0163	0.0146
accidents L	0.0681	0.0575
accidents DMS	0.4365	0.4070

On peut conclure qu'il n'y a, en général, aucun impact statistiquement significatif des panneaux-réclame sur les nombres d'accidents répertoriés dans l'échangeur Décarie au cours de la période de 12 mois suivant leur installation comparativement à l'évolution moyenne des accidents de la zone de contrôle. En ce qui concerne la gravité des accidents durant les 12 mois précédant l'installation des panneaux, on note que le risque d'accidents mortels ou graves est estimé à 3,14%, celui des accidents légers à 13,07%, et le risque d'accidents avec dommages matériels seulement s'élève à 83,79%. Le tableau 4.13 montre qu'il n'y a pas d'incidence importante (statistiquement significative) sur les différents risques d'accidents de la période de 12 mois après par rapport à la période de 12 mois avant l'installation des panneaux-réclame dans l'échangeur Décarie.

Chapitre 5

Conclusions et perspectives

5.1 Tests d'hypothèse d'égalité des effets

5.1.1 Introduction

En pratique, on cherche souvent à comparer l'effet moyen de la mesure au niveau des différents sites expérimentaux ou des différents types d'accidents. Par exemple, si une mesure est prise pour différentes villes expérimentales (disons dix villes pour fixer les idées) avec pour chacune différents types de gravité d'accident (blessés graves, blessés légers, dégâts matériels uniquement), on cherche à comparer l'effet moyen dans ces villes ou pour différents types de gravité. La mesure a-t-elle fait baisser le nombre d'accidents de façon identique dans chacune des dix villes ? Peut-on affirmer que la mesure a eu, en proportion, le même effet sur chacune des gravités d'accident ?

La question fréquemment rencontrée dans la pratique est donc de savoir si la mesure (ou l'intervention) imposée a eu en moyenne le même effet au niveau de chaque site expérimental. D'un point de vue statistique, cette question revient à tester l'hypothèse d'égalité des effets moyens relativement aux sites expérimentaux ou aux différents types d'accidents. L'hypothèse de base (ou hypothèse nulle) à éprouver est donc l'égalité des effets moyens. En effet, il s'agit de comparer, sous cette hypothèse nulle, un ensemble de valeurs espérées (théoriques) de nombre d'accidents issues d'un modèle spécifique à un ensemble de valeurs observées après de la mise en place de cette mesure.

Diverses statistiques du type chi-deux [9, 12, 14, 18, 19, 21, 29, 32, 36, 61, 69, 71, 72, 73, 75, 78] peuvent être utilisées pour ajuster un tel modèle théorique à des données. L'utilisation pratique de ces statistiques exige la connaissance exacte de leur distribution asymptotique. Cette dernière est elle même liée à la structure des données et à la méthode d'estimation des paramètres du modèle d'analyse.

Dans le domaine de la sécurité routière et lorsque les données sont mises sous forme d'un tableau avant-après avec contrôle pour chaque site expérimental, nos statistiques représentent une extension immédiate de celle proposée par Tanner pour éprouver l'hypothèse d'égalité de l'effet dû à la transformation de carrefours en ronds-points. Nous avons établi la distribution asymptotique des statistiques utilisées en étudiant la loi de certaines formes quadratiques aléatoires.

Nos résultats (annexes A08p.148 et A12p.208) s'appliquent également aux situations particulières correspondant à un seul site expérimental et plusieurs types d'accidents ou à plusieurs sites expérimentaux et un seul type d'accidents. Pour ce faire, nous construisons des extensions des deux modèles de N'Guessan et al (2001, 2006) (voir annexes A02p.65 et A05p.110). Les méthodes utilisées pour estimer les paramètres sont le maximum de vraisemblance et le χ^2 -minimum modifié [17, 36]. Il existe d'autres méthodes d'estimation (voir par exemple [25, 53, 54, 56, 67, 94]) non abordées dans ce travail de recherche.

5.1.2 Modèles statistiques (annexe A08)

Les deux modèles d'analyse de cette section sont extraits de l'annexe A08. Le principe de leur construction, exposé dans [79], est similaire à celui des modèles de la section 2.3.1. On note θ_k ($\theta_k > 0$) le paramètre effet moyen de la mesure au site expérimental k , $k = 1, 2, \dots, S$. En s'inspirant du modèle 1 de la section 2.3.1, on fait l'hypothèse que, le vecteur $(X_{1k}, X_{2k})^T$ suit une loi multinomiale notée $\mathcal{M}(n_k; \Pi_{1k}(\Psi), \Pi_{2k}(\Psi))$, où

$$\Pi_{tk}(\Psi) = (\pi_{t1k}(\Psi), \pi_{t2k}(\Psi), \dots, \pi_{tRk}(\Psi))$$

est le vecteur de dimension R et de composantes $\pi_{tjk} : \mathbb{R}^{S+SR} \mapsto]0; 1[$ données par :

$$\pi_{1jk}(\Psi) = \frac{p_{jk}}{1 + \theta_k \sum_{m=1}^R c_{mk} p_{mk}} ; \quad \pi_{2jk}(\Psi) = \frac{\theta_k c_{jk} p_{jk}}{1 + \theta_k \sum_{m=1}^R c_{mk} p_{mk}}, \quad (j = 1, 2, \dots, R), \quad (1)$$

$\Psi = (\theta_1, \theta_2, \dots, \theta_S, P^T)^T$ de dimension $(S + SR)$ est le vecteur des paramètres et où $P = (P_1^T, P_2^T, \dots, P_S^T)^T$ avec $P_k = (p_{1k}, p_{2k}, \dots, p_{Rk})^T$, vecteur de dimension R est tel que :

$$P_k \in \mathcal{S}_k^{(R-1)} = \{(p_{1k}, p_{2k}, \dots, p_{Rk})^T \in \mathbb{R}^R, p_{jk} > 0, \sum_{j=1}^R p_{jk} = 1\}. \quad (2)$$

De même, en s'inspirant du modèle 2 de la section 2.3.2, on déduit un deuxième modèle multinomial en supposant, pour k fixé, que le vecteur $(X_{1k}, X_{2k})^T$ suit une loi multinomiale notée $\mathcal{M}(n_k; \Pi_{1k}(\Psi), \Pi_{2k}(\Psi))$, où

$$\Pi_{tk}(\Psi) = (\pi_{t1k}(\Psi), \pi_{t2k}(\Psi), \dots, \pi_{tRk}(\Psi))$$

est le vecteur de dimension R et de composantes $\pi_{tjk} : \mathbb{R}^{S+SR} \mapsto]0; 1[$ données par :

$$\pi_{1jk}(\Psi) = \frac{p_{jk}}{1 + \theta_k \sum_{m=1}^R c_{mk} p_{mk}} ; \quad \pi_{2jk}(\Psi) = \frac{\theta_k p_{jk} \sum_{m=1}^R c_{mk} p_{mk}}{1 + \theta_k \sum_{m=1}^R c_{mk} p_{mk}}, \quad (j = 1, 2, \dots, R), \quad (3)$$

$\Psi = (\theta_1, \theta_2, \dots, \theta_S, P^T)^T$ de dimension $(S + SR)$ est le vecteur des paramètres et où $P = (P_1^T, P_2^T, \dots, P_S^T)^T$ avec $P_k = (p_{1k}, p_{2k}, \dots, p_{Rk})^T$, vecteur de dimension R est tel que :

$$P_k \in \mathcal{S}_k^{(R-1)} = \{(p_{1k}, p_{2k}, \dots, p_{Rk})^T \in \mathbb{R}^R, p_{jk} > 0, \sum_{j=1}^R p_{jk} = 1\}. \quad (4)$$

Ces deux modèles ainsi que ceux de la section 2.3 peuvent aussi s'obtenir, sous certaines conditions, à l'aide de variables de Poisson conditionnellement à leur totaux [74].

5.1.3 Statistiques de test et loi limite

On propose des statistiques pour décider de l'existence ou non de variation entre les différents effets moyens et on étudie leur loi asymptotique. Il s'agit d'éprouver l'hypothèse :

$$H_0 : \theta_1 = \theta_2 = \dots = \theta_S = \theta.$$

On considère pour cela :

soit la statistique

$$\hat{A}_{S,R}^2 = \sum_{k=1}^S \sum_{j=1}^R \sum_{t=1}^2 \frac{(X_{tjk} - n_k \pi_{tjk}(\hat{\Theta}))^2}{n_k \pi_{tjk}(\hat{\Theta})}$$

où $\hat{\Theta}$ représente, sous l'hypothèse H_0 , l'estimateur du maximum de vraisemblance sous contraintes de Ψ déduit de celui de Θ (annexe A05 pour le modèle 1 et annexe A02 pour le modèle 2) obtenu au chapitre 2.

soit la statistique

$$\tilde{A}_{S,R}^2 = \sum_{k=1}^S \sum_{j=1}^R \sum_{t=1}^2 \frac{(X_{tjk} - n_k \pi_{tjk}(\tilde{\Theta}))^2}{n_k \pi_{tjk}(\tilde{\Theta})}$$

où $\tilde{\Theta}$ représente, sous l'hypothèse H_0 , l'estimateur du χ^2 minimum modifié ([17] p. 424-425) sous contraintes de Ψ basé sur l'ensemble des données. On étudie, dans l'annexe A08, la distribution asymptotique de ces statistiques. En plus des conditions de régularité du paragraphe 2.4.2, on suppose

(C6) $n_k = n_k(N)$ et $N = \sum_{k=1}^S n_k \rightarrow +\infty$ tels que

$$\lim_{N \rightarrow +\infty} \frac{n_k(N)}{N} = \zeta_k < +\infty.$$

Cette condition supplémentaire décrit le comportement de l'effectif total d'accidents par site expérimental au cours des deux périodes quand le nombre total d'accidents de l'ensemble des sites tend vers l'infini. Les deux théorèmes ci-dessus et leurs corollaires établissent les lois asymptotiques des statistiques utilisées.

Théorème 5.1.1 (annexe A08)

Sous l'hypothèse H_0 et les conditions (C1) à (C6) la statistique $\hat{A}_{S,R}^2$ est asymptotiquement distribuée selon χ_{SR-1}^2 (une loi du chi-deux à $SR - 1$ degrés de liberté).

Théorème 5.1.2 (annexe A08)

Sous l'hypothèse H_0 et les conditions (C1) à (C6) la statistique $\tilde{A}_{S,R}^2$ est asymptotiquement distribuée selon χ_{SR-1}^2 (une loi du chi-deux à $SR - 1$ degrés de liberté).

On déduit alors les corollaires ci-dessous. Ces corollaires permettent d'appliquer nos résultats à des situations particulières comme par exemple lorsque $S = 1$ et $R > 1$ (la mesure est appliquée à un site expérimental comportant plusieurs types d'accidents) ou lorsque $S > 1$ et $R = 1$ (on est ici dans le cas particulier des résultats de Tanner). Les premiers corollaires sont associés à l'estimation du maximum de vraisemblance sous contraintes et les deux suivants à la méthode du chi-deux minimum modifié sous contraintes.

Corollaire 5.1.3 (annexe A08)

Sous l'hypothèse H_0 , les conditions (C1) à (C6) et si $S = 1$ et $R > 1$ la statistique $\hat{A}_{1,R}^2$ est asymptotiquement distribuée selon χ_{R-1}^2 (une loi du chi-deux à $R - 1$ degrés de liberté).

Corollaire 5.1.4 (annexe A08)

Sous l'hypothèse H_0 , les conditions (C1) à (C6) et $S > 1$ et $R = 1$ la statistique $\hat{A}_{S,1}^2$ est asymptotiquement distribuée selon χ_{S-1}^2 (une loi du chi-deux à $S - 1$ degrés de liberté).

Corollaire 5.1.5 (annexe A08)

Sous l'hypothèse H_0 , les conditions (C1) à (C6) et si $S = 1$ et $R > 1$ la statistique $\tilde{A}_{1,R}^2$ est asymptotiquement distribuée selon χ_{R-1}^2 (une loi du chi-deux à $R - 1$ degrés de liberté).

Corollaire 5.1.6 (annexe A08)

Sous l'hypothèse H_0 , les conditions (C1) à (C6) et $S > 1$ et $R = 1$ la statistique $\tilde{A}_{S,1}^2$ est asymptotiquement distribuée selon χ_{S-1}^2 (une loi du chi-deux à $S - 1$ degrés de liberté).

Lemme 5.1.7 (annexe A08)

Sous les mêmes conditions que les théorèmes 5.1.1 et 5.1.2, on montre que le vecteur aléatoire Y de dimension $(2SR) \times 1$ et de composantes

$$Y_{tjk} = \frac{X_{tjk} - n_k \pi_{tjk}(\Theta^0)}{(n_k \pi_{tjk}(\Theta^0))^{1/2}}$$

est asymptotiquement normalement distribué de moyenne nulle, de matrice de variances et covariances $\Omega^* = I_{2SR} - \Pi^{1/2}(\Pi^{1/2})^T$, où $\Pi^{1/2}$ est le vecteur de dimension $2SR$ et de composantes $[\pi_{tjk}(\Theta^0)]^{1/2}$, et I_{2SR} la matrice identité de dimension $2SR$.

Pour obtenir la distribution asymptotique de ces statistiques, il est nécessaire d'étudier la loi (voir par exemple [69, 98 p.57-69]) des formes quadratiques aléatoires du type $Y^T A Y$ où Y est un vecteur aléatoire multinormal de moyenne nulle, de matrice de variances covariances Ω^* et A une matrice carrée symétrique. On sait alors que $Y^T A Y$ est distribuée comme un $\chi_{tr(A\Omega^*)}^2$ si et seulement si la matrice $A\Omega^*$ est idempotente.

Dans le cas des approches statistiques de l'évaluation d'une mesure que nous proposons, on montre que $A = I_{2SR} - D(D^T D)^{-1} D^T$ où la matrice D est telle $D^T D = J_{\Theta_0}$ (la matrice d'information de Fisher) sous H_0 et Ω^* donnée par le lemme 5.1.7. Différents résultats intermédiaires permettent alors de déduire que $\text{tr}(A\Omega^*) = SR - 1$. Il s'agit pour l'essentiel de considérer, sous H_0 , la fonctionnelle qui laisse θ invariant et transforme le vecteur $\beta_k = (p_{1k}, p_{2k}, \dots, p_{R-1k})^T$ en $P_k = (p_{1k}, p_{2k}, \dots, p_{R-1k}, p_{Rk} = 1 - \sum_{j=1}^{R-1} p_{jk})^T$. La détermination de la trace de la matrice $A\Omega^*$ est alors obtenue à l'aide de l'algorithme proposé par Gerstein [33] et utilisé dans [79].

Ces résultats permettent de disposer d'un outil d'aide à la décision relativement à l'homogénéité de l'efficacité moyenne de la mesure de sécurité routière au niveau de l'ensemble des sites expérimentaux. Cette décision est conditionnée par l'ensemble des données dont on dispose et par l'erreur de première espèce choisie.

Pour éprouver pratiquement l'homogénéité de l'efficacité moyenne de la mesure de sécurité routière, on compare généralement la statistique $\hat{A}_{S,R}^2$ ou $\tilde{A}_{S,R}^2$ à la valeur $\chi_{SR-1,\alpha}^2$ que dépasse une variable aléatoire de loi de chi-deux à $SR - 1$ degrés de liberté avec une probabilité α (0,01 ou 0,05). Si la valeur calculée de $\hat{A}_{S,R}^2$ (par exemple) est plus grande, on rejette au niveau α l'hypothèse H_0 (il y a vraisemblablement une différence significative dans l'effet moyen de la mesure aux différents sites expérimentaux). Sinon, il n'y pas de raison apparente de rejeter H_0 . Pour éprouver H_0 , on peut aussi envisager d'estimer une probabilité de signification adaptée à la statistique utilisée et de s'en servir pour prendre une décision appropriée.

5.2 Estimation à l'aide de la méthode du χ^2 minimum modifié

5.2.1 Principe de la méthode

Classiquement, on a recours à la méthode du maximum de vraisemblance avec ou sans contraintes pour estimer l'effet d'une mesure de sécurité routière. Les raisons proviennent des propriétés bien connues (convergence et normalité asymptotique) de cette famille d'estimateurs qui permettent d'obtenir des intervalles de confiance. Les raisons sont aussi numériques car sa mise en oeuvre pratique est facilitée par l'algorithme de Newton-Raphson [31, 94] qui fournit par la même occasion une approximation de la précision des estimations. Dans certaines applications, comme par exemple la modélisation dans les tableaux de contingence à l'aide de distributions multinomiales, il est connu que l'estimation du maximum de vraisemblance et la méthode du χ^2 minimum modifié coïncident ([17] p. 425-426). De plus les résultats, que nous obtenons, montrent clairement la place importante occupée par la statistique du χ^2 dans nos approches d'évaluation.

L'estimateur $\tilde{\phi}$ du χ^2 minimum modifié de l'effet moyen et des risques d'accident, s'il existe, est solution du système d'équations

$$\sum_{k=1}^S \sum_{j=1}^R \sum_{t=1}^2 \frac{X_{tjk} - n_k \pi_{tjk}}{\pi_{tjk}} \times \frac{\partial \pi_{tjk}}{\partial \phi_m} = \mathbf{0}; \quad m = 1, 2, \dots, 1 + S(R - 1), \quad (5)$$

où ϕ_m représente la m^e composante du vecteur $\phi = (\theta, \beta_1^T, \beta_2^T, \dots, \beta_S^T)^T$, avec $\beta_k = (p_{1k}, p_{2k}, \dots, p_{R-1k})^T$ et π_{tjk} obtenu en utilisant les modèles de la section 2.3. Nous

construisons une suite $(\phi_l, l \geq 1)$ convergente de vecteurs dont la limite $\tilde{\phi}$ est solution du système d'équations ci-dessus. En combinant la solution $\tilde{\phi}$ et la fonctionnelle qui laisse θ invariant et transforme le vecteur β_k en $P_k = (p_{1k}, p_{2k}, \dots, p_{R-1k}, p_{Rk})^T$ avec $p_{Rk} = 1 - \sum_{j=1}^{R-1} p_{jk}$, ($k = 1, 2, \dots, S$), nous montrons l'existence de $\tilde{\phi}$.

5.2.2 Résultats asymptotiques

Théorème 5.2.1 (annexe A12)

On montre qu'il existe un estimateur du χ^2 minimum modifié $\tilde{\phi}$ de ϕ^0 (la vraie valeur) tel que pour $\epsilon_N > 0$ donné

$$\|\tilde{\phi} - \phi^0\| \longrightarrow 0 \text{ quand } N = \sum_{k=1}^S n_k \longrightarrow +\infty$$

avec une probabilité plus grande que $1 - \epsilon_N^{-2}$, où $\epsilon_N \longrightarrow +\infty$ quand $N = \sum_{k=1}^S n_k \longrightarrow +\infty$ tel que $N^{-1/2} \epsilon_N^2 \longrightarrow 0$.

Pour obtenir les résultats asymptotiques (existence et convergence en probabilité) de $\tilde{\phi}$, on construit donc une suite de vecteurs de Cauchy dans un voisinage $\mathcal{V}(\phi^0)$ compact non vide de ϕ^0 . Dans le cas de nos modèles, on note :

D_0 la matrice de taille $(2SR) \times (1 + S(R-1))$ dont chaque élément est donné par

$$\left[\frac{1}{(\pi_{tjk})^{1/2}} \frac{\partial \pi_{tjk}}{\partial \phi_m} \right]_{\phi=\phi^0};$$

$$\Gamma_N^{1/2} = \text{Diag} \left(\frac{n_1^{1/2}}{N^{1/2}} I_{2R}, \frac{n_2^{1/2}}{N^{1/2}} I_{2R}, \dots, \frac{n_S^{1/2}}{N^{1/2}} I_{2R} \right);$$

la matrice $(2SR) \times (2SR)$ où I_{2R} est la matrice identité d'ordre $2R$,

et on montre que l'estimateur du χ^2 -minimum modifié est obtenu de façon équivalente par la forme matricielle

$$(\tilde{\phi} - \phi^0) = N^{-1/2} (D_0^T \Gamma_N D_0)^{-1} (\Gamma_N^{1/2} D_0)^T Y + (D_0^T \Gamma_N D_0)^{-1} \bar{\xi}(\tilde{\phi}) \quad (6)$$

avec

$$\bar{\xi}(\phi) = (\bar{\xi}_1(\phi), \bar{\xi}_2(\phi), \dots, \bar{\xi}_{1+S(R-1)}(\phi))^T$$

vecteur de dimension $1 + S(R-1)$ et de composantes

$$\bar{\xi}_m = \frac{1}{N} \sum_{k=1}^S n_k \bar{\xi}_{k,m}(\phi)$$

où $\bar{\xi}_{k,m}(\phi)$ est une fonction réelle continûment différentiable et définie en utilisant $\pi_{tjk}(\phi)$ et ses dérivées partielles premières par rapport à ϕ . Réciproquement, il s'agit de montrer qu'il existe asymptotiquement un point de $\mathcal{V}(\phi^0)$ vérifiant la forme matricielle ci-dessus et donc solution de l'équation (5) du paragraphe 5.2.1. Les résultats ci-dessous permettent de montrer cette réciproque. Les lemmes 5.2.2 et 5.2.3 permettent de construire une suite de points de $\mathcal{V}(\phi^0)$ et d'étudier sa convergence en probabilité. Le théorème 5.2.1 permet de montrer l'unicité asymptotique de la solution.

Lemme 5.2.2 (*annexe A12*)

Avec une probabilité plus grande que $1 - \epsilon_N^{-2}$, on a

$$|\bar{\xi}_m(\phi^{(2)}) - \bar{\xi}_m(\phi^{(1)})| \leq c_1 \|(\phi^{(2)}) - (\phi^{(1)})\| \left(\|\phi^{(1)} - \phi^0\| + \|\phi^{(2)} - \phi^0\| + \zeta_0^{1/2} \frac{\epsilon_N}{N^{1/2}} \right)$$

où c_1 est une constante positive indépendante de $\phi^{(1)}$, $\phi^{(2)}$, m , N et $\epsilon_N \rightarrow +\infty$ quand $N = \sum_{k=1}^S n_k \rightarrow +\infty$ tel que $N^{-1/2} \epsilon_N^2 \rightarrow 0$, $\zeta_0^{1/2} = \lim_{N \rightarrow +\infty} [\sum_{k=1}^S (\frac{n_k}{N})^{1/2}]$, $\phi^{(1)}$ et $\phi^{(2)}$ deux points de $\mathcal{V}(\phi^0)$ un voisinage non vide de ϕ^0 .

Lemme 5.2.3 (*annexe A12*)

On définit $(\phi_l, l \geq 1)$ une suite de points de $\mathcal{V}(\phi^0)$ telle que $\phi_0 = \phi^0$ et

$$\tilde{\phi}_l = \phi^0 + N^{-1/2} (D_0^T \Gamma_N D_0)^{-1} (\Gamma_N^{1/2} D_0)^T Y + (D_0^T \Gamma_N D_0)^{-1} \tilde{\xi}(\tilde{\phi}_{l-1}).$$

Alors, avec une probabilité plus grande que $1 - \epsilon_N^{-2}$, on a $\forall l > 0$

$$\|\phi_{l+1} - \phi_l\| \leq c_2 [(2c_2 + \zeta_0^{1/2})c_3]^l \left(\frac{\epsilon_N}{N^{1/2}} \right)^{l+1}$$

où c_2 et c_3 sont des constantes positives indépendantes de N et l , $\epsilon_N \rightarrow +\infty$ quand $N = \sum_{k=1}^S n_k \rightarrow +\infty$ tel que $N^{-1/2} \epsilon_N^2 \rightarrow 0$, $\zeta_0^{1/2} = \lim_{N \rightarrow +\infty} [\sum_{k=1}^S (\frac{n_k}{N})^{1/2}]$.

5.3 Quelques perspectives

La méthode d'estimation des paramètres (effet moyen et différents risques d'accident) reste un point crucial dans les approches d'évaluation traitées dans nos travaux de recherche. En général, les paramètres des modèles sont estimés à l'aide de l'algorithme de Newton-Raphson appliqué à la log-vraisemblance sous contraintes : l'espace des paramètres est égal à la demi-droite réelle positive croisée à des produits de simplexes. Dans le cas de l'évaluation d'une mesure à l'aide de tableaux de données d'accident, l'application de cet algorithme à nos modèles donne des résultats très satisfaisants.

Néanmoins, ces tableaux de données d'accident peuvent être, en pratique, très creux. Par exemple, les périodes d'observations peuvent être très courtes ou le nombre de sites expérimentaux et de types d'accident élevé ou la zone aménagée est de nature particulière de telle sorte que certains effectifs d'accident soient très faibles voire nuls. A ce problème d'effectifs d'accident,

il faut ajouter celui de la dimension de l'espace des paramètres qui est, par construction, une fonction croissante du nombre de sites expérimentaux et du nombre de types d'accident, et le fait que la majeure partie des paramètres est à rechercher dans des produits de simplexes. On est donc face à une problématique d'estimation où certains paramètres sont dans des intervalles $]0, 1[$ et leur nombre peut, très rapidement, devenir gigantesque. De telles situations, inhérentes à nos approches statistiques, peuvent rendre difficile le calcul de la log-vraisemblance et compliquent l'obtention d'une solution initiale adaptée à notre méthode d'estimation.

Pour remédier à ces différents problèmes spécifiques aux approches statistiques proposées, et mettre des outils simples à disposition du praticien, nous comptons explorer plusieurs pistes de recherche. Pour la recherche d'une solution initiale adaptée à nos procédures d'estimation, nous comptons décomposer les équations de vraisemblance de l'annexe 05 (voir paragraphe 2.4.3) en plusieurs sous-systèmes linéaires et adapter certaines propriétés du complément de Schur liées aux systèmes d'équations linéaires. Ainsi, on pourra contourner le problème de la dimensionnalité, fournir et automatiser une solution initiale.

Une deuxième piste de recherche serait l'adaptation de certaines méthodes d'estimation comme par exemple l'EM algorithme et ses extensions [25, 67, 109] ou le MM algorithme [31, 53, 54, 56, 94 p. 253-255]. Enfin, une autre voie de recherche semble être une reparamétrisation des modèles d'évaluation en s'inspirant de la transformation logarithmique proposée dans l'annexe A04. Ceci permettrait alors de résoudre en partie la sensibilité des sous espaces de solutions, introduirait une structure logistique (voir par exemple [52, 68]) dans la loi de nos modèles multinomiaux, et établirait un pont entre les modèles de regression logistique et nos approches statistiques d'évaluation d'une mesure de sécurité routière.

Annexe A01 :

Statistics, **Vol. 39, No. 4, 303-314 (2005)**

A covariance components estimation procedure when modelling a road safety measure in terms of linear constraints

A covariance components estimation procedure when modelling a road safety measure in terms of linear constraints

ASSI N'GUESSAN*† and CLAUDE LANGRAND‡

†Ecole Polytechnique Universitaire de Lille et, Laboratoire Paul Painlevé CNRS, UMR 8542, Université de Lille 1, 59655 Villeneuve d'Ascq Cedex, France

‡Laboratoire Paul Painlevé CNRS, UMR 8542, U.F.R de Mathématiques pures et appliquées, Université de Lille 1, 59655 Villeneuve d'Ascq Cedex, France

(Received 15 December 2003; revised 7 December 2004; in final form 9 March 2005)

This paper deals with the asymptotic estimation of the covariance between the average effect of a road safety measure and the accident risks when the multidimensional parameter of interest is constrained. The explicit asymptotic covariance is achieved through Schur complements technical method. This approach also provides a powerful tool to formally compute the explicit covariance between two different accident risks whatever the experimental sites. The easiness of the obtention of the asymptotic covariances without matrix inversion is an attractive aspect of this new approach. Some examples of formal estimation are discussed to back up the method.

Keywords: Road safety measure; Accident data; Multinomial model; Constrained maximum likelihood; Fisher information matrix; Schur complements; Asymptotic covariance matrix; Formal estimation

AMS Classification: 62F10; 62F12; 62F25; 62F30

1. Introduction

The purpose of this paper is to give a formal and analytical expression of the asymptotic covariance between the components of interest of a constrained vector parameter arising in a linear problem of restricted maximum likelihood estimation (RMLE) of road safety measure modelling.

The general framework for RMLE is well-known and is not discussed here. Aitchison and Silvey [1] consider the case of i.i.d. observations along the lines due to Cramér [2] and prove, under certain conditions, asymptotic existence, consistency and normality of the RMLE. Silvey [3] considers the i.i.d. case, but allowing the Fisher information matrix to be singular. He uses the alternative approach to prove consistency due to Wald [4]. Magnus [5] considers an alternative way of dealing with RMLE asymptotic results in the case of non i.i.d. observations using linear structures properties. Crowder [6] considers the case of non i.i.d.

*Corresponding author. Email: assi.nguessan@polytech-lille.fr; Fax: +33 3 28 76 73 01.

observations when the parameters are constrained. His asymptotic results and formal methods stem from those of Aitchison and Silvey [1] and Silvey [3]. Also, more emphasis is placed in Crowder's [6] paper on the precise manner in which the constraints may be used to construct a suitable nonsingular version of Fisher information matrix and the connection with parameter identifiability. Don [7] generalizes the well-known results of Aitchison and Silvey [1] using g-inverses approach [8]. He deals properly with the loss of rank in the bordered information matrix, while the assumption of i.i.d. is replaced by some much more general assumptions. Neuenschwander and Flury [9] give an alternative proof of a theorem by Silvey [3] for the asymptotic covariance matrix of the maximum likelihood estimation (m.l.e.) under different types of constraints in the parameter space. They show how the different types of constraints can be treated in a unified way. Some related studies are also connected to RMLE [10–14].

Whatever the case considered and the method used in RMLE, it is clear that a central role is played by the inverse, if it exists, of the bordered information matrix which has the form

$$\Gamma_\alpha = \begin{pmatrix} J_\alpha & H_\alpha^T \\ H_\alpha & \mathbf{O} \end{pmatrix}, \tag{1}$$

where parameter α , Fisher information matrix J_α and constraint matrix H_α are best defined subsequently. The general version of matrix (1) was extensively studied by Rao [15] and Rao and Yanai [8]. All the results on the inverse of matrix (1) are characterized under some rank conditions on the block matrices in the partitions. In the present paper, we consider a version of matrix (1) connected to multidimensional road safety measure modelling and thoroughly study the analytical expression of the block elements apart from the diagonal of its inverse using the Schur complements approach [16–18]. One particularly focuses attention on the explicit formula of the covariance between any two components of the unknown vector parameter.

Section 2 is devoted to the bordered multinomial model and the information matrix of interest. Section 3 is exclusively centred on the main results. We particularly use the model and constraints mentioned subsequently to clarify the covariances. Some formal examples are discussed in section 4 to back up the method. More particularly, we have matrix (1) dimensions vary in order to fully prove what our method offers. Some technical lemmas are derived in Appendix A in order to support the foregoing main results.

2. Bordered multinomial model and information matrix

2.1 Bordered multinomial model

Let us consider a collection $(Y_{11}, Y_{21}), (Y_{12}, Y_{22}), \dots, (Y_{1s}, Y_{2s})$ of s independent couples of $2r -$ dimensional random vectors with $r > 1$ such that each couple has a multinomial distribution whose cell probabilities are linked to an unknown vector parameter α subjected to a set of restraints $h(\alpha) = 0$ where $h = (h_1, h_2, \dots, h_s)$.

Following N'Guessan *et al.* [19] and N'Guessan [20–22], we assume, for fixed k , that

$$(Y_{1k}, Y_{2k}) \sim \mathcal{M} \left(n_k; \prod_{1k}(\alpha), \prod_{2k}(\alpha) \right)$$

where

$$\prod_{tk}(\alpha) = (\pi_{t1k}(\alpha), \pi_{t2k}(\alpha), \dots, \pi_{trk}(\alpha)); \quad t = 1, 2$$

and $n_k (n_k > 0)$ is a given integer, $\alpha = (\beta_0, \beta^T)^T \in \mathbb{R}^{1+sr}$ is the unknown vector parameter, $\beta_0 (\beta_0 > 0)$ a scalar, $\beta = (\beta_1^T, \dots, \beta_s^T)^T \in \mathbb{R}^{sr}$ with $\beta_k = (\beta_{1k}, \beta_{2k}, \dots, \beta_{rk})^T \in \mathbb{R}^r$ and $\beta_{jk} > 0$. The cell probabilities $\pi_{tjk}(\alpha)$ of the $r \times 1$ vector probability $\Pi_{tk}(\alpha)$ are given by a continuous differentiable functional link $g: \mathbb{R}^{1+sr} \mapsto]0, 1[$ [so that $\pi_{tjk}(\alpha) = g(z_k, \alpha)$

$$g(Z_k, \alpha) = \begin{cases} \frac{\beta_{jk}}{1 + \beta_0 \langle Z_k, \beta_k \rangle} & t = 1; j = 1, 2, \dots, r \\ \frac{z_{jk} \beta_0 \beta_{jk}}{1 + \beta_0 \langle Z_k, \beta_k \rangle} & t = 2; j = 1, 2, \dots, r \end{cases} \quad (2)$$

with

$$h_k(\alpha) = \langle \mathbf{1}_r, \beta_k \rangle - 1; \quad k = 1, 2, \dots, s \quad (3)$$

where $\langle \cdot, \cdot \rangle$ is the usual inner product, $Z_k = (z_{1k}, z_{2k}, \dots, z_{rk})^T$ a given extra set of observations connected to each couple (Y_{1k}, Y_{2k}) and $\mathbf{1}_r = (1, \dots, 1)^T \in \mathbb{R}^r$, the vector of unity. The construction mechanism of the statistical pattern mentioned earlier is displayed in N'Guessan *et al.* [19]. Let us indicate that in road safety modelling β_0 represents the average effect when combining all data and vector β_k , for fixed k , denotes the probability distribution (or accident risk) connected to extra data Z_k such that $\sum_{j=1}^r \beta_{jk} = 1$. This is the natural and practical interpretation of constraints (3).

2.2 Bordered information matrix

Let $\mathcal{L}(\alpha)$ denote the logarithm of the likelihood function for a given s independent couples of $2r$ -dimensional random vectors (Y_{1k}, Y_{2k}) from a true model has an unknown parameter value α^0 . We know that α^0 satisfies a set of restrictions $h_k(\alpha^0) = 0, k = 1, 2, \dots, s$. Then the well-known RMLE problem is

$$\text{Max } \mathcal{L}(\alpha) \text{ subject to } h_k(\alpha) = 0; \quad k = 1, 2, \dots, s \quad (4)$$

where $\mathcal{L}(\alpha)$ is given to one additive constant

$$\mathcal{L}(\alpha) = \sum_{k=1}^s \sum_{j=1}^r \{y_{.jk} \log_e(\beta_{jk}) + y_{2jk} \log_e(\beta_0) - y_{.jk} \log_e(1 + \beta_0 \langle Z_k, \beta_k \rangle)\} \quad (5)$$

with $y_{.jk} = y_{1jk} + y_{2jk}$ and $h_k(\alpha)$ given by equation (3).

This is not the place to study sufficient conditions for the existence and consistency [5, 7] of a solution $\hat{\alpha}$. N'Guessan *et al.* [19] study numerically the existence and consistency of $\hat{\alpha}$ via wide simulation results and mean square error. We will follow assumptions and lines due to Crowder [6] and used in N'Guessan [20–22].

Let us assume, then, that a restricted ML estimator $\hat{\alpha}$ exists asymptotically almost surely and is strongly consistent. Furthermore, the asymptotic Fisher information matrix J_α is then given by the following $(1 + sr) \times (1 + sr)$ matrix:

$$J_\alpha = \mathbb{E} \left(- \frac{\partial^2 \mathcal{L}}{\partial \alpha \partial \alpha^T} \right)$$

evaluate at different points on the line segment joining α^0 and $\hat{\alpha}$, the restraint ML, and the $s \times (1 + sr)$ constraint matrix $H_\alpha = (H_1, H_2)$ with $H_1 = (\partial h / \partial \beta_0)$ an $s \times 1$

306

A. N'Guessan and C. Langrand

matrix, $H_2 = (\partial h/\partial \beta)$ an $s \times (sr)$ matrix. Consequently, the $(1 + s(r + 1)) \times (1 + s(r + 1))$ bordered information matrix is given by

$$\Gamma_\alpha = \begin{bmatrix} J_\alpha & H_\alpha^\top \\ H_\alpha & \circ_{s,s} \end{bmatrix}, \quad J_\alpha = \begin{bmatrix} \tau_\alpha & U_\alpha^\top \\ U_\alpha & B_\alpha \end{bmatrix}, \tag{6}$$

where J_α is a $(1 + sr) \times (1 + sr)$ matrix and H_α is an $s \times (1 + sr)$ matrix as follows

$$U_\alpha = \begin{bmatrix} U_{\alpha,1}^\top \\ U_{\alpha,2}^\top \\ \vdots \\ U_{\alpha,s}^\top \end{bmatrix}; \quad B_\alpha = \begin{bmatrix} B_{\alpha,1} & \circ & \cdots & \circ \\ \circ & B_{\alpha,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \circ \\ \circ & \cdots & \circ & B_{\alpha,s} \end{bmatrix}; \quad H_\alpha = \begin{bmatrix} 0 & 1_r^\top & 0_r^\top & \cdots & 0_r^\top \\ 0 & 0_r^\top & 1_r^\top & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0_r^\top \\ 0 & 0_r^\top & \cdots & 0_r^\top & 1_r^\top \end{bmatrix}, \tag{7}$$

with $\tau_\alpha = \mathbb{E}(-\partial^2 \mathcal{L}/\partial \beta_0 \partial \beta_0)$ a scalar, $U_{\alpha,k}$ an $r \times 1$ vector whose components are $\mathbb{E}(-\partial^2 \mathcal{L}/\partial \beta_0 \partial \beta_{jk})$, ($j = 1, 2, \dots, r$), $B_{\alpha,k}$ an $r \times r$ matrix whose entries are $\mathbb{E}(-\partial^2 \mathcal{L}/\partial \beta_{jk} \partial \beta_{mk})$, ($j, m = 1, 2, \dots, r$), and $0_r^\top = (0, \dots, 0)$ an $r \times 1$ vector of zero. So taking the second derivative of the negative of $\mathcal{L}(\alpha)$ with respect to the components of α , one shows that

$$U_{\alpha,k} = \frac{\gamma_k^{3/2}}{\beta_0 n_k^{1/2}} V_{\alpha,k}, \quad B_{\alpha,k} = \gamma_k (\Omega_{\alpha,k} - V_{\alpha,k} V_{\alpha,k}^\top) \tag{8}$$

with

$$V_{\alpha,k} = \beta_0 \frac{\gamma_k^{1/2}}{n_k^{1/2}} (z_{1k}, z_{2k}, \dots, z_{rk})^\top; \quad \Omega_{\alpha,k} = \text{diag} \left(\frac{1 + \beta_0 z_{1k}}{\beta_{1k}}, \dots, \frac{1 + \beta_0 z_{rk}}{\beta_{rk}} \right),$$

where

$$\gamma_k = \frac{n_k}{1 + \beta_0 \langle Z_k, \beta_k \rangle}, \quad \text{for } k = 1, 2, \dots, s \quad \text{and} \quad \tau_\alpha = \sum_{k=1}^s \frac{\gamma_k \langle Z_k, \beta_k \rangle}{\beta_0 (1 + \beta_0 \langle Z_k, \beta_k \rangle)}.$$

The rest of the paper looks at the analytical expression of the covariance between any two components of the constrained ML estimator $\hat{\alpha}$. The following theorems and corollaries concern the main results and give the formal expressions of the asymptotic covariances using Schur complement. More precisely, we are interested in the asymptotic covariance between the parameter of interest β_0 (the average effect of a road safety measure) and any element of the subparameter $\beta \in \mathbb{R}^{sr}$ which denotes the accident risk in control areas and in the asymptotic covariance between two different accident risks.

3. Main results

This section sketches the demonstration of the main results concerning the asymptotic covariance between the different components of vector $\hat{\alpha}$. Thus, for any number of experimental sites, Theorem 3.1 and Corollary 3.2 deal with the explicit formal expression of the covariance between the average effect of a road safety measure and the different types of risk whereas Theorem 3.3 discusses the structure of the covariances between the different accident types, whatever the accident type and the experimental site considered. In particular, Corollary 3.4

sets the explicit covariance matrix between the different types of accident risks when the measure is applied to a single site only.

The restraints m.l.e. (RMLE) of α_0 , provided it exists, is vector $\hat{\alpha}$ which satisfies

$$\left(\frac{\partial L}{\partial \alpha}\right)_{\hat{\alpha}} + H_{\hat{\alpha}}^T \hat{\lambda} = 0, \quad h(\hat{\alpha}) = 0 \quad (9)$$

where λ is a vector of s Lagrange multipliers. Let $M_{\alpha} = J_{\alpha}^{-1}(J_{\alpha} + (\partial^2 L / \partial \alpha \partial \alpha))$ a $(1 + sr) \times (1 + sr)$ matrix, $R_{\alpha} = H_2(J_{\alpha} / \tau_{\alpha})^{-1} H_2^T$ an $s \times s$ matrix, $\Phi_{\alpha} = J_{\alpha}^{-1} H_{\alpha}^T R_{\alpha}^{-1}$ a $(1 + sr) \times s$ matrix, $\xi_{\alpha} = H_2 B_{\alpha}^{-1} U_{\alpha}$ an $s \times 1$ vector, $\Lambda_{\alpha} = \text{diag}(\|\mathbf{1}_r\|_{B_{\alpha,1}^{-1}}^2, \|\mathbf{1}_r\|_{B_{\alpha,2}^{-1}}^2, \dots, \|\mathbf{1}_r\|_{B_{\alpha,s}^{-1}}^2)$ an $s \times s$ matrix with $\|\mathbf{1}_r\|_{B_{\alpha,k}^{-1}}^2 = \mathbf{1}_r^T B_{\alpha,k}^{-1} \mathbf{1}_r$, and

$$\Lambda_{\alpha}^{(1)} = \begin{bmatrix} \Lambda_{\alpha} & -\xi_{\alpha} \\ \xi_{\alpha}^T & \left(\frac{J_{\alpha}}{B_{\alpha}}\right) \end{bmatrix} \quad (10)$$

an $(s + 1) \times (s + 1)$ matrix. We assume for the rest of the work that the following conditions are verified:

- (A1) $(\hat{\alpha}, \hat{\lambda})$ is a consistent solution of the latter equation.
 (A2) $M_{\hat{\alpha}} \xrightarrow{\text{Proba.}} 0$, $\Phi_{\hat{\alpha}} - \Phi_{\alpha^0} \xrightarrow{\text{Proba.}} 0$, $|\Phi_{\alpha^0}| < \infty$, when n_k and $N = \sum_{k=1}^s n_k \rightarrow \infty$
 (A3) $(\partial L / \partial \alpha)_{\alpha^0} \stackrel{\text{Distr.}}{\sim} \mathcal{N}(0, J_{\alpha^0})$.

THEOREM 3.1 Under conditions (A1)–(A3) the asymptotic $r \times 1$ subcovariance matrix between the $r \times 1$ vector $\hat{\beta}_k$ and the scalar $\hat{\beta}_0$ of the RMLE $\hat{\alpha}$ is

$$\text{cov}(\hat{\beta}_k; \hat{\beta}_0) = S_{\alpha,k} \times Z_k; \quad k = 1, 2, \dots, s \quad (11)$$

where Z_k is an $r \times 1$ given data vector and $S_{\alpha,k}$ is the $r \times r$ matrix with entries $S_{\alpha,k}(i, j)$ as follows

$$S_{\alpha,k}(i, j) = -\frac{t_k}{n_k(\Lambda_{\alpha}^{(1)} / \Lambda_{\alpha})} \begin{cases} \frac{\gamma_k \beta_{ik}}{1 + \beta_0 z_{ik}} - \frac{(1 + \rho_{\alpha,k} z_{ik}) \beta_{ik}^2}{\|\mathbf{1}_r\|_{B_{\alpha,k}^{-1}}^2 (1 + \beta_0 z_{ik})^2} & i = j \\ -\frac{(1 + \rho_{\alpha,k} z_{ik}) \beta_{ik} \beta_{jk}}{\|\mathbf{1}_r\|_{B_{\alpha,k}^{-1}}^2 (1 + \beta_0 z_{ik})(1 + \beta_0 z_{jk})} & i \neq j \end{cases} \quad (12)$$

with $(\Lambda_{\alpha}^{(1)} / \Lambda_{\alpha})$ the Schur complement of Λ_{α} in $\Lambda_{\alpha}^{(1)}$, $\mathbf{1}_{r,r} = \mathbf{1}_r \mathbf{1}_r^T$ the $r \times r$ matrix with unity and

$$\rho_{\alpha,k} = \frac{t_k \Delta_{\alpha,k} \beta_0^2 \gamma_k}{n_k}, \quad \Delta_{\alpha,k} = Z_k^T \Omega_{\alpha,k}^{-1} \mathbf{1}_r, \quad t_k = (1 - \|V_{\alpha,k}\|_{\Omega_{\alpha,k}^{-1}}^2)^{-1} \quad (13)$$

Proof (Outline) Under conditions (A1)–(A3) and combining model (2), constraints (3) and Lemmas 1–7 (Appendix A) we show that the $r \times 1$ asymptotic covariance matrix between

308

A. N'Guessan and C. Langrand

vector $\hat{\beta}_k$ and scalar $\hat{\beta}_0$ is given by

$$\text{cov}(\hat{\beta}_k; \hat{\beta}_0) = -\frac{1}{(\Lambda_\alpha^{(1)}/\Lambda_\alpha)} \left[I_r - B_{\alpha,k}^{-1} \frac{\mathbf{1}_{r,r}}{\|\mathbf{1}_r\|_{B_{\alpha,k}^{-1}}^2} \right] B_{\alpha,k}^{-1} U_{\alpha,k}$$

for $k = 1, 2, \dots, s$. Now using expression of $B_{\alpha,k}^{-1}$ (Lemma 1, part iii) and the relation between vectors $U_{\alpha,k}$ and Z_k , we get the relations

$$B_{\alpha,k}^{-1} U_{\alpha,k} = \frac{t_k \gamma_k}{n_k} \Omega_{\alpha,k}^{-1} Z_k \quad \text{and} \quad B_{\alpha,k}^{-1} \mathbf{1}_{r,r} \Omega_{\alpha,k}^{-1} = \gamma_k^{-1} [\Omega_{\alpha,k}^{-1} (\mathbf{1}_r, \mathbf{1}_r^T + \rho_{\alpha,k} Z_k \mathbf{1}_r^T) \Omega_{\alpha,k}^{-1}]$$

So after matrices manipulations, we obtain

$$\text{cov}(\hat{\beta}_k; \hat{\beta}_0) = \mathcal{S}_{\alpha,k} \times Z_k$$

where the $r \times r$ matrix $\mathcal{S}_{\alpha,k}$ is as follows

$$\mathcal{S}_{\alpha,k} = -\frac{t_k}{n_k (\Lambda_\alpha^{(1)}/\Lambda_\alpha)} \left[\gamma_k \Omega_{\alpha,k}^{-1} - \frac{1}{\|\mathbf{1}_r\|_{B_{\alpha,k}^{-1}}^2} \Omega_{\alpha,k}^{-1} (\mathbf{1}_{r,r} + \rho_{\alpha,k} Z_k \mathbf{1}_r^T) \Omega_{\alpha,k}^{-1} \right]$$

and the entries $\mathcal{S}_{\alpha,k}(i, j)$ follow from a straightforward calculation. ■

COROLLARY 3.2 *Under the same assumptions as Theorem 3.1, the explicit expression of the asymptotic covariance between $\hat{\beta}_0$ and any other estimator $\hat{\beta}_{mk}$ is given by*

$$\begin{aligned} \frac{n_k (\Lambda_\alpha^{(1)}/\Lambda_\alpha)}{t_k} [\text{cov}(\hat{\beta}_{mk}, \hat{\beta}_0)] &= \left(-\frac{\gamma_k \beta_{mk}}{1 + \beta_0 z_{mk}} + \frac{(1 + \rho_{\alpha,k} z_{mk}) \beta_{mk}^2}{\|\mathbf{1}_r\|_{B_{\alpha,k}^{-1}}^2 (1 + \beta_0 z_{mk})^2} \right) z_{mk} \\ &+ \frac{(1 + \rho_{\alpha,k} z_{mk}) \beta_{mk}}{\|\mathbf{1}_r\|_{B_{\alpha,k}^{-1}}^2 (1 + \beta_0 z_{mk})} \left(\Delta_{\alpha,k} - \frac{\beta_{mk} z_{mk}}{1 + \beta_0 z_{mk}} \right) \end{aligned} \quad (14)$$

for $m = 1, 2, \dots, r$.

Remark 1 The proof of this corollary stems from a straightforward development of the result given in Theorem 3.1 and uses the expression of $\Delta_{\alpha,k}$ displayed in Lemma 5 (Appendix A). So we leave out the details.

THEOREM 3.3 *Under the same assumptions as Theorem 3.1, the $r \times r$ between covariance matrix of $\hat{\beta}_k$ and $\hat{\beta}_m$ is*

$$\text{cov}(\hat{\beta}_k; \hat{\beta}_m) = J_{k,m} - \sum_{j=1}^s \left[\left(\sum_{i=1}^s A_{i,j} J_{k,i} \right) \mathbf{1}_{r,r} \right] J_{j,m} \quad (15)$$

$k, m = 1, 2, \dots, s$, where $J_{k,m}$ is an $r \times r$ matrix given by the Schur complements of (J_α/B_α) in the $(1+r) \times (1+r)$ matrix

$$\begin{bmatrix} (J_\alpha/B_\alpha) & U_{\alpha,m}^T B_{\alpha,m}^{-1} \\ B_{\alpha,k}^{-1} U_{\alpha,k} & \delta_{k,m} B_{\alpha,k}^{-1} \end{bmatrix},$$

and

$$A_{k,m} = \frac{\delta_{k,m}}{\|\mathbf{1}_r\|_{B_{\alpha,k}^{-1}}^2} - \frac{(\Lambda_\alpha^{(1)}/\Lambda_\alpha)^{-1} t_k t_m \gamma_k \gamma_m \Delta_{\alpha,k} \Delta_{\alpha,m}}{n_k n_m \|\mathbf{1}_r\|_{B_{\alpha,k}^{-1}}^2 \|\mathbf{1}_r\|_{B_{\alpha,m}^{-1}}^2}, \quad (16)$$

with $\delta_{k,m}$ the usual Kronecker symbol.

Remark 2 The proof stems from an algebraic manipulation of the $(sr) \times (sr)$ matrix $W_\alpha(2, 2)$ given in Lemma 6 of the Appendix A. So we just give the outline. Using the Kronecker product, we deduce that $H_2^T R_\alpha^{-1} H_2 = A_\alpha \otimes \mathbf{1}_{r,r}$, where A_α is the $s \times s$ matrix with (k, m) th element given by $A_{k,m}$. So we get the result, using Lemma 3, the postmultiplication of $H_2^T R_\alpha^{-1} H_2$ by $(J_\alpha / \tau_\alpha)^{-1}$ and finally the premultiplication of the same matrix.

COROLLARY 3.4 *Under the same assumptions as Theorem 3.1 and when the measure is applied to a single site ($s = 1$) with different accident types ($r > 1$), the covariances between the different components of the $r \times 1$ vector $\hat{\beta}_k$ are as follows*

$$\text{cov}(\hat{\beta}_{jk}, \hat{\beta}_{mk}) = \gamma_k^{-1} w_{jk} w_{mk} [\theta_k \gamma_k z_{jk} z_{mk} - A_{k,k} \gamma_k^{-1} x_{jk} x_{mk}], \quad j \neq m = 1, 2, \dots, r \quad (17)$$

where $w_{jk} = (\beta_{jk} / (1 + \beta_0 z_{jk}))$ the diagonal elements of matrix $\Omega_{\alpha,k}^{-1}$, $\theta_k = (t_k / n_k) [\beta_0^2 - (J_\alpha / B_\alpha)^{-1} (t_k \gamma_k^2 / n_k)]$, $x_{jk} = 1 + \theta_k \Delta_{\alpha,k} \gamma_k z_{jk}$. ($j = 1, 2, \dots, r$).

Proof (Outline) For $s = 1$ and $r > 1$, and in view of Theorem 3.3 and of expression of matrix $B_{\alpha,k}^{-1}$ (see Lemma 1 below), the covariance matrix of the $r \times 1$ vector $\hat{\beta}_k$ is given by

$$\text{cov}(\hat{\beta}_k, \hat{\beta}_k) = J_{k,k} - A_{k,k} J_{k,k} \mathbf{1}_{r,r} J_{k,k}$$

where the $r \times r$ matrix $J_{k,k}$ is as follows

$$J_{k,k} = \gamma_k^{-1} [\Omega_{\alpha,k}^{-1} + \theta_k \gamma_k \Omega_{\alpha,k}^{-1} Z_k Z_k^T \Omega_{\alpha,k}^{-1}].$$

So after a manipulation of matrices, one obtains

$$\text{cov}(\hat{\beta}_k, \hat{\beta}_k) = \gamma_k^{-1} [\Omega_{\alpha,k}^{-1} + \Omega_{\alpha,k}^{-1} (\theta_k \gamma_k Z_k Z_k^T - A_{k,k} \gamma_k^{-1} X_k X_k^T) \Omega_{\alpha,k}^{-1}]$$

where X_k is the $r \times 1$ vector of elements $x_{mk} = 1 + \theta_k \Delta_{\alpha,k} \gamma_k z_{mk}$. The result follows from the elements beside the diagonal of the latter matrix. ■

4. Examples

The main aim of the examples of formal estimation provided subsequently is to back up the covariance estimation approach proposed in this work. These examples also enable us to realize our method's importance and usefulness when the dimensions of the considered matrices are great.

4.1 Example 1: $s = 1$ and $r = 2$

We suppose here that the number of experimental sites is equal to one and that there are two accident types with n_k , the total number of accidents for the site, and $Z_k = (z_{1k}, z_{2k})^T$ the data from the control area. The parameter is defined by $\alpha = (\beta_0, \beta_k)^T \in \mathbb{R}^3$ and $\beta_k = (\beta_{1k}, \beta_{2k})^T \in \mathbb{R}^2$ with $\beta_0 > 0$, $\beta_{jk} > 0$ and $\beta_{1k} + \beta_{2k} = 1$. So the dimension of the parameter space is three

310

A. N'Guessan and C. Langrand

and we have one constraint. We show that the 4×4 constrained Fisher information matrix is

$$\Gamma_\alpha = \begin{pmatrix} \tau_\alpha & \frac{\gamma_k^{3/2}}{\beta_0 n_k^{1/2}} v_{1k} & \frac{\gamma_k^{3/2}}{\beta_0 n_k^{1/2}} v_{2k} & 0 \\ \frac{\gamma_k^{3/2}}{\beta_0 n_k^{1/2}} v_{1k} & b_{11}^{(k)} & b_{12}^{(k)} & 1 \\ \frac{\gamma_k^{3/2}}{\beta_0 n_k^{1/2}} v_{2k} & b_{21}^{(k)} & b_{22}^{(k)} & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

where $\tau_\alpha = (\gamma_k^2 \langle Z_k, \beta_k \rangle) / \beta_0 n_k$, $v_{mk} = \beta_0 (\gamma_k^{1/2} / n_k^{1/2}) z_{mk}$, $\gamma_k = n_k / (1 + \beta_0 \langle Z_k, \beta_k \rangle)$, and

$$\begin{pmatrix} b_{11}^{(k)} & b_{12}^{(k)} \\ b_{21}^{(k)} & b_{22}^{(k)} \end{pmatrix} = \gamma_k \begin{pmatrix} \frac{1 + \beta_0 z_{1k}}{\beta_{1k}} - \frac{\beta_0^2 z_{1k}^2 \gamma_k}{n_k} & -\frac{\beta_0^2 z_{1k} z_{2k} \gamma_k}{n_k} \\ -\frac{\beta_0^2 z_{2k} z_{1k} \gamma_k}{n_k} & \frac{1 + \beta_0 z_{2k}}{\beta_{2k}} - \frac{\beta_0^2 z_{2k}^2 \gamma_k}{n_k} \end{pmatrix}.$$

In order to get the asymptotic covariance matrix of the constrained m.l.e. $\hat{\alpha}$, the classical approach consists of inverting matrix Γ_α and take the leading 3×3 submatrix. The method proposed here enables us to avoid doing this inversion by using main results as follows: setting

$$J_\alpha = \begin{pmatrix} \tau_\alpha & \frac{\gamma_k^{3/2}}{\beta_0 n_k^{1/2}} V_{\alpha,k}^\top \\ \frac{\gamma_k^{3/2}}{\beta_0 n_k^{1/2}} V_{\alpha,k} & B_{\alpha,k} \end{pmatrix}; \quad \Omega_{\alpha,k} = \begin{pmatrix} \frac{1 + \beta_0 z_{1k}}{\beta_{1k}} & 0 \\ 0 & \frac{1 + \beta_0 z_{2k}}{\beta_{2k}} \end{pmatrix}$$

with $B_{\alpha,k} = \gamma_k [\Omega_{\alpha,k} - V_{\alpha,k} V_{\alpha,k}^\top]$ a 2×2 matrix, $V_{\alpha,k} = \beta_0 (\gamma_k^{1/2} / n_k^{1/2}) (z_{1k}, z_{2k})^\top$; $H_2 = (1, 1)$, and using Corollaries 3.2 and 3.4 we obtain both the covariance between the 2×1 vector $\hat{\beta}_k$ and the scalar $\hat{\beta}_0$ and the covariance within the vector $\hat{\beta}_k$ i.e. covariance between the elements $\hat{\beta}_{1k}$ and $\hat{\beta}_{2k}$ as follows:

$$\text{cov}(\hat{\beta}_{1k}, \hat{\beta}_0) = \frac{n_k^{-1} t_k \beta_{1k}}{[(\Lambda_\alpha^{(1)} / \Lambda_\alpha)](1 + \beta_0 z_{1k})} \left[-\gamma_k z_{1k} + \frac{(1 + \rho_{\alpha,k} z_{1k}) \Delta_{\alpha,k}}{\|\mathbf{1}_r\|_{B_{\alpha,k}^{-1}}^2} \right],$$

$$\text{cov}(\hat{\beta}_{2k}, \hat{\beta}_0) = \frac{n_k^{-1} t_k \beta_{2k}}{[(\Lambda_\alpha^{(1)} / \Lambda_\alpha)](1 + \beta_0 z_{2k})} \left[-\gamma_k z_{2k} + \frac{(1 + \rho_{\alpha,k} z_{2k}) \Delta_{\alpha,k}}{\|\mathbf{1}_r\|_{B_{\alpha,k}^{-1}}^2} \right]$$

and

$$\text{cov}(\hat{\beta}_{1k}, \hat{\beta}_{2k}) = \gamma_k^{-1} w_{1k} w_{2k} [\theta_k \gamma_k z_{1k} z_{2k} - A_{k,k} \gamma_k^{-1} x_{1k} x_{2k}].$$

So we obtain the analytical expressions for covariance between any two components of the 3×3 vector $\hat{\alpha}$.

4.2 Example 2: $s = 2$ and $r = 2$

We have two experimental sites and two accident types with n_k, n_m the total number of accidents in each site and $Z_k = (z_{1k}, z_{2k})^T$, $Z_m = (z_{1m}, z_{2m})^T$ the data from each associated control area. The parameter and the constraints, defined by $\alpha = (\beta_0, \beta_k, \beta_m)^T \in \mathbb{R}^5$ where $\beta_k = (\beta_{1k}, \beta_{2k})^T \in \mathbb{R}^2$, $\beta_m = (\beta_{1m}, \beta_{2m})^T \in \mathbb{R}^2$ with $\beta_0 > 0$, $\beta_{ij} > 0$, $\beta_{1k} + \beta_{2k} = 1$ and $\beta_{1m} + \beta_{2m} = 1$. So the dimension of the parameter space is five and we have two constraints. One notes that the additional experimental site to Example 1 has extended the dimension of the parameter space by two and constrained information matrix is a 7×7 matrix given by

$$\Gamma_\alpha = \begin{pmatrix} \tau_\alpha & \frac{\gamma_k^{3/2}}{\beta_0 n_k^{1/2}} v_{1k} & \frac{\gamma_k^{3/2}}{\beta_0 n_k^{1/2}} v_{2k} & \frac{\gamma_m^{3/2}}{\beta_0 n_m^{1/2}} v_{1m} & \frac{\gamma_m^{3/2}}{\beta_0 n_m^{1/2}} v_{2m} & 0 & 0 \\ \frac{\gamma_k^{3/2}}{\beta_0 n_k^{1/2}} v_{1k} & b_{11}^{(k)} & b_{12}^{(k)} & 0 & 0 & 1 & 0 \\ \frac{\gamma_k^{3/2}}{\beta_0 n_k^{1/2}} v_{2k} & b_{21}^{(k)} & b_{22}^{(k)} & 0 & 0 & 1 & 0 \\ \frac{\gamma_m^{3/2}}{\beta_0 n_m^{1/2}} v_{1m} & 0 & 0 & b_{11}^{(m)} & b_{12}^{(m)} & 0 & 1 \\ \frac{\gamma_m^{3/2}}{\beta_0 n_m^{1/2}} v_{2m} & 0 & 0 & b_{21}^{(m)} & b_{22}^{(m)} & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{pmatrix}$$

where $\tau_\alpha = (\gamma_k^2 \langle Z_k, \beta_k \rangle / \beta_0 n_k) + (\gamma_m^2 \langle Z_m, \beta_m \rangle / \beta_0 n_m)$ and the other terms are defined in the same way as in Example 1.

Let us suppose for example that we wish to estimate $\text{cov}(\hat{\beta}_0, \hat{\beta}_{2m})$, *i.e.* the covariance between the measure's average effect and the type 2 accident risk in the second experimental site. Applying Corollary 3.2 results, we obtain the desired covariance as follows:

$$\begin{aligned} \frac{n_m (\Lambda_\alpha^{(1)} / \Lambda_\alpha)}{t_m} [\text{cov}(\hat{\beta}_{2m}, \hat{\beta}_0)] &= \left(-\frac{\gamma_m \beta_{2m}}{1 + \beta_0 z_{2m}} + \frac{(1 + \rho_{\alpha,m} z_{2m}) \beta_{2m}^2}{\|\mathbf{1}_r\|_{B_{\alpha,m}^{-1}}^2 (1 + \beta_0 z_{2m})^2} \right) z_{2m} \\ &+ \frac{(1 + \rho_{\alpha,m} z_{2m}) \beta_{2m}}{\|\mathbf{1}_r\|_{B_{\alpha,m}^{-1}}^2 (1 + \beta_0 z_{2m})} \left(\Delta_{\alpha,m} - \frac{\beta_{2m} z_{2m}}{1 + \beta_0 z_{2m}} \right) \end{aligned}$$

with

$$\left(\frac{\Lambda_\alpha^{(1)}}{\Lambda_\alpha} \right) = \sum_{i \in \{k, m\}} \frac{\gamma_i^2}{n_i} \left[\left(\frac{\langle Z_i, \beta_i \rangle}{\beta_0} - \frac{\gamma_i t_i}{n_i} \|Z_i\|_{\Omega_{\alpha,i}^{-1}}^2 \right) + \frac{t_i^2 \Delta_{\alpha,i}^2}{n_i \|\mathbf{1}_r\|_{B_{\alpha,i}^{-1}}^2} \right].$$

For the covariance between the different accident risks, we use Theorem 3.3 calculating the following 2×2 matrices

$$\text{cov}(\hat{\beta}_k; \hat{\beta}_m) = J_{k,m} - \sum_{j=1}^2 \left[\left(\sum_{i=1}^2 A_{i,j} J_{k,i} \right) \mathbf{1}_{r,r} \right] J_{j,m}$$

$k, m = 1, 2$ and taking the elements outside the main diagonal. The explicit form of the latter matrix is obtained after some matrix developments. The problem of the 7×7 matrix inversion is thus replaced by a 2×2 matrix addition problem.

4.3 Example 3: $s = 20$ and $r = 10$

In this example, the dimension of the matrix of interest Γ_α is 221×221 . The latter matrix's analytical expression is not given here for a question of space. To estimate the covariances between the different components of estimator $\hat{\alpha}$, our method consists of using Corollary 3.2 and Theorem 3.3 formulas. This approach consists of making 10×10 matrix additions rather than inverting a 221×221 matrix. On the whole, we note that the size of the constrained information matrix is an increasing function of the experimental site number and the number of accident types linked to the road safety measure. To obtain the desired covariances, sophisticated numerical and statistical tools are needed to invert matrices whose dimensions grow rather rapidly. The approach we suggest in this paper replaces the direct inversion of such a matrix with products/multiplications and additions of smaller-size matrices. It also enables us to have the explicit analytical expression of the desired covariances.

Acknowledgements

We wish to thank Professor Claude Brezinski, who drew our attention to the paper about Schur complements. We are also deeply grateful to the anonymous referee for judicious notes and comments which have enabled a complete remodelling of our paper. Part of this work was carried out while the first author was an invited researcher of CRT (Centre for Research on Transportation) of Université de Montréal, whose partial support is gratefully acknowledged. This research was partially supported by France-Maroc action intégrée programme no. MA01/02.

Appendix A

Technical lemmas used in the main results of section 3 are given subsequently, without demonstration. These intermediate technical results stem from those of N'Guessan [20–22] and use some Schur complement properties [23]. So we leave out the proof.

LEMMA 1 For $k = 1, 2, \dots, s$

- (i) $\|Z_k\|_{\Omega_{\alpha,k}^{-1}}^2 < \frac{\langle Z_k, \beta_k \rangle}{\beta_0}$;
- (ii) $\|V_{\alpha,k}\|_{\Omega_{\alpha,k}^{-1}}^2 < \frac{\beta_0 \langle Z_k, \beta_k \rangle}{1 + \beta_0 \langle Z_k, \beta_k \rangle} < 1$;
- (iii) matrix $B_{\alpha,k}$ is nonsingular and

$$B_{\alpha,k}^{-1} = \gamma_k^{-1} (\Omega_{\alpha,k}^{-1} + t_k \Omega_{\alpha,k}^{-1} V_{\alpha,k} V_{\alpha,k}^T \Omega_{\alpha,k}^{-1}),$$

with $t_k = (1 - \|V_{\alpha,k}\|_{\Omega_{\alpha,k}^{-1}}^2)^{-1}$.

LEMMA 2 (i) $\|U_\alpha\|_{B_\alpha^{-1}}^2 < \tau_\alpha$, (ii) $(J_\alpha/B_\alpha) > 0$, where (J_α/B_α) is the Schur complement of B_α in J_α .

LEMMA 3 (J_α/τ_α) the Schur complement of τ_α in J_α is a $(sr) \times (sr)$ nonsingular matrix with

$$(J_\alpha/\tau_\alpha)^{-1} = (J_{k,m})_{(1 \leq k \leq s; 1 \leq m \leq s)}$$

where each block $J_{k,m}$ is an $r \times r$ matrix given by the Schur complements of (J_α/B_α) in the $(1+r) \times (1+r)$ matrix

$$\begin{bmatrix} (J_\alpha/B_\alpha) & U_{\alpha,m}^\top B_{\alpha,m}^{-1} \\ B_{\alpha,k}^{-1} U_{\alpha,k} & \delta_{k,m} B_{\alpha,k}^{-1} \end{bmatrix}.$$

LEMMA 4 The $s \times s$ matrix $R_\alpha = H_2(J_\alpha/\tau_\alpha)^{-1} H_2^\top$ is nonsingular and

$$R_\alpha^{-1} = \Lambda_\alpha^{-1} - [(\Lambda_\alpha^{(1)}/\Lambda_\alpha)]^{-1} \Lambda_\alpha^{-1} \xi_\alpha \xi_\alpha^\top \Lambda_\alpha^{-1}.$$

LEMMA 5 We get the following results

$$\begin{aligned} \text{(i)} \quad \Delta_{\alpha,k} &= \sum_{m=1}^r \frac{\beta_{mk} z_{mk}}{1 + \beta_0 z_{mk}}, \\ \text{(ii)} \quad \|\mathbf{1}_r\|_{B_{\alpha,k}^{-1}}^2 &= \gamma_k^{-1} \left[\text{trace}(\Omega_{\alpha,k}^{-1}) + t_k \frac{\beta_0^2 \gamma_k \Delta_{\alpha,k}^2}{n_k} \right], \\ \text{(iii)} \quad (\Lambda_\alpha^{(1)}/\Lambda_\alpha) &= \sum_{k=1}^s \frac{\gamma_k^2}{n_k} \left[\left(\frac{\langle Z_k, \beta_k \rangle}{\beta_0} - \frac{\gamma_k t_k}{n_k} \|Z_k\|_{\Omega_{\alpha,k}^{-1}}^2 \right) + \frac{t_k^2 \Delta_{\alpha,k}^2}{n_k \|\mathbf{1}_r\|_{B_{\alpha,k}^{-1}}^2} \right]. \end{aligned}$$

LEMMA 6 Matrix Γ_α is nonsingular and

$$\Gamma_\alpha^{-1} = \begin{bmatrix} W_\alpha & J_\alpha^{-1} H_\alpha^\top R_\alpha^{-1} \\ R_\alpha^{-1} H_\alpha J_\alpha^{-1} & -R_\alpha^{-1} \end{bmatrix}$$

where W_α , the leading $(1+rs) \times (1+sr)$ matrix, is given by

$$W_\alpha = \begin{bmatrix} W_\alpha(1,1) & W_\alpha^\top(2,1) \\ W_\alpha(2,1) & W_\alpha(2,2) \end{bmatrix}$$

with $W_\alpha(1,1)$, $W_\alpha(2,1)$, and $W_\alpha(2,2)$, respectively, a scalar, an $(sr) \times 1$ vector and an $(sr) \times (sr)$ matrix as follows:

$$\begin{aligned} W_\alpha(1,1) &= (\Lambda_\alpha^{(1)}/\Lambda_\alpha)^{-1} \\ W_\alpha(2,1) &= -(J_\alpha/B_\alpha)^{-1} [B_\alpha^{-1} U_\alpha - (J_\alpha/\tau_\alpha)^{-1} H_2^\top R_\alpha^{-1} \xi_\alpha] \\ W_\alpha(2,2) &= (J_\alpha/\tau_\alpha)^{-1} - (J_\alpha/\tau_\alpha)^{-1} H_2^\top R_\alpha^{-1} H_2 (J_\alpha/\tau_\alpha)^{-1}. \end{aligned}$$

LEMMA 7 Under conditions (A1)–(A3) the asymptotic distribution of RMLE $\hat{\alpha}$ is:

$$\hat{\theta} \stackrel{\text{Distr.}}{\sim} \mathcal{N}(\theta^0; W_{\theta^0})$$

where $W_\alpha = J_\alpha^{-1} - J_\alpha^{-1} H_\alpha^\top R_\alpha^{-1} H_\alpha J_\alpha^{-1}$ is a $(1+sr) \times (1+sr)$ partitioned matrix given by the latter lemma.

References

- [1] Aitchison, J. and Silvey, S.D., 1958, Maximum likelihood estimation of parameters subject to restraints. *Annals of Mathematical Statistics*, **29**, 813–829.
- [2] Cramer, H., 1946, *Mathematical Methods of Statistics* (Princeton University Press).
- [3] Silvey, S.D., 1959, The Lagrangian multiplier test. *Annals of Mathematical Statistics*, **30**, 389–407.
- [4] Wald, A., 1949, Note on the consistency of the maximum likelihood estimate. *Annals of Mathematical Statistics*, **20**, 595–601.
- [5] Magnus, J.R., 1988, *Linear Structures* (London: Charles Griffin and Company Ltd).
- [6] Crowder, M., 1984, On the constrained maximum likelihood estimation with non i.i.d. observations. *Annals of Institute of Statistics Mathematics*, **36A**, 239–249.
- [7] Don, F.J.H., 1985, The use of generalized inverses in restricted maximum likelihood. *Linear Algebra and Its Applications*, **70**, 225–240.
- [8] Rao, C.R. and Yanai, H., 1985, Generalized inverses of partitioned matrices useful in statistical applications. *Linear Algebra and its Applications*, **70**, 105–115.
- [9] Neuenschwander, B.E. and Flury, B.D., 1997, A note on Silvey's (1959) theorem. *Statistics and Probability Letter*, **36**, 307–317.
- [10] Harville, D.A., 1977, Maximum likelihood approaches to variance component estimation and to related problems. *Journal of American Society Association*, **72**(358), 320–340.
- [11] Lee, S.-Y., 1979, Constrained estimation in covariance structure analysis. *Biometrika*, **66**(3), 539–545.
- [12] Haber, M. and Brown, M.B., 1986, Maximum likelihood methods for log-linear models when expected frequencies are subject to linear constraints. *Journal of American Society Association*, **81**(394), 477–482.
- [13] Matthews, G.B. and Crowther, N.A.S., 1995, A maximum likelihood estimation procedure when modelling in terms of constraints. *South African Statistics Journal*, **29**, 29–50.
- [14] Van Eeden, C., 1996, Estimation in restricted parameter spaces—some history and some recent developments. *CWI Quarterly*, **9**(1 and 2), 69–76.
- [15] Rao, C.R., 1973, *Linear Statistical Inference and Its Applications* (2nd edn) (New York: John Wiley).
- [16] Ouellette, D.V., 1981, Schur complements and statistics. *Linear Algebra and Its Applications*, **36**, 187–295.
- [17] Fan, Y., 2002, Schur complements and its applications to symmetric nonnegative and Z -matrices. *Linear Algebra and Its Applications*, **353**, 289–307.
- [18] Zhang, F., 2005, *Schur Complement and Its Applications* (Springer Verlag).
- [19] N'Guessan, A., Essai, A. and Langrand, C., 2001, Estimation multidimensionnelle des contrôles et de l'effet moyen d'une mesure de sécurité routière. *Revue de Statistique Appliquée*, **XLIX**(2), 83–100.
- [20] N'Guessan, A., 2003, On a use of Schur complements for an inverse of a constrained accident data information matrix. *Technical Report, Publications de l'Institut de Recherche en Mathématiques Avancées (IRMA)*, **60**(IV), 1–15.
- [21] N'Guessan, A., 2003, Constrained covariance matrix estimation in road accident modelling with Schur complements. *Comptes rendus de l'Académie des Sciences, Paris, Série I*, **337**, 219–222.
- [22] N'Guessan, A., 2003, Constrained estimation of a road safety covariance matrix using Schur complements. Technical Report, Centre for Research on Transportation, CRT-2003-12, Université de Montréal.
- [23] N'Guessan, A. and Langrand, C., 2005, A Schur complement approach for computing subcovariance matrices arising in road safety measure modelling. *Journal of Computational and Applied Mathematics*, **177**, 331–345.

Annexe A02 :

Computational Statistics & Data Analysis, Vol. 00, No. 00, 000-000 (2006)

An estimation method of the average effect and the different accident risks when modelling a road safety measure : A simulation study

ARTICLE IN PRESS

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Computational Statistics & Data Analysis III (III) III-III

**COMPUTATIONAL
STATISTICS
& DATA ANALYSIS**
www.elsevier.com/locate/csda

An estimation method of the average effect and the different accident risks when modelling a road safety measure: A simulation study

Assi N'Guessan^{a,*}, Azeddine Essai^b, Modeste N'Zi^c

^a *Ecole Polytechnique Universitaire de Lille et Laboratoire Paul Painlevé, UMR CNRS 8542 Université de Lille 1, 59655 Villeneuve d'Ascq Cedex, France*

^b *Laboratoire d'Analyse Numérique et d'Optimisation, U.F.R de Mathématiques pures et appliquées Université de Lille 1, 59655 Villeneuve d'Ascq Cedex, France*

^c *Laboratoire de Probabilités et Statistique, U.F.R de Mathématiques et d'Informatique Université de Cocody, Abidjan Côte d'Ivoire*

Received 13 July 2004; received in revised form 19 January 2005; accepted 5 September 2005

Abstract

The combination of road accident frequencies before and after a similar change at a given number of sites are considered. Each target site includes different accident types and is linked to a specific control area. At any one target site it is assumed that the total number of accidents recorded is multinomially distributed between the before period and the after period and also between several mutually exclusive types. The parameter of the distribution depends on the different accident risks in the control area linked to each site as well as on the average effect of the change. A method of estimating simultaneously the average effect and the accident risks in control areas is suggested. Some simulated accidents data allow us to study the existence and consistence of the linear constrained estimator of the unknown vector parameter.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Road safety measure; Accident data; Multinomial model; Constrained maximum likelihood; Accident risk; Mean square error; Before–after estimation; Simulation

1. Introduction

In this paper, we consider that a road safety measure (crossroad lay-out, surface of a motorway section, etc.) is simultaneously applied to several sites (experimental sites), each site presenting several mutually exclusive types of accidents (fatal accidents, seriously injured people, slightly injured people, material damage, etc.). We also consider that to each experimental site a control area is associated, with the same accident types, but where the measure is not directly applied. The control areas are used not only as comparison sites but also and mainly enable us to take into account the impact of some factors (as traffic-flow and speed variations before and after the applied measure, weather conditions, experimental sites location, etc.) on the applied measure effect. These factors have a direct or indirect influence on the observed data in the control areas and act upon the accident risks in these areas. We therefore have to

* Corresponding author. Tel.: +33 3 28 76 74 57; fax: +33 3 28 76 73 01.

E-mail addresses: assi.nguessan@polytech-lille.fr (A. N'Guessan), essai@ano.univ-lille1.fr (A. Essai), nzim@ucocody.ci (M. N'Zi).

ARTICLE IN PRESS

2

A. N'Guessan et al. / Computational Statistics & Data Analysis ■■■ (■■■■) ■■■–■■■

include the impact of these factors into the statistical models used to analyse the measure mean effect if we want to interpret correctly the scope of this measure at the experimental sites.

The statistical models used in the analysis of a road safety measure efficiency heavily depend on the data and on the target set along with the measure (see for instance Hauer, 1997; Barry et al., 1999; Beenstock et al., 2001; Abdel-Aty and Radwan, 2000; Gaudry and Lassarre, 2000; Lassarre, 2001). According to some points of view and some practices in the field of road safety, it is advisable to model the accident data according to Poisson distribution (see for example Lassarre, 1977; Hauer, 1997), conditional or truncated Poisson distribution (for example, Danielson, 1986, Danielson, 1988; Hauer, 1997), negative binomial distribution (Hauer, 2001; Allain and Brenac, 2001). Even if Poisson model or conditional Poisson model is a common and convenient assumption in accident count analysis, are accidents Poisson distributed? It is very important to evaluate this Poisson assumption in practice using statistical tests (see for example Nicholson, 1985, 1986; Nicholson and Wong, 1993). The question is which approach and assumption give the better results and under what conditions.

In this paper we have adopted Tanner (1958)'s before–after philosophy. In fact, the latter author introduced the notion of control area and defined a control coefficient for each experimental site using the ratio of the accident number after to before in the control area. Thereafter, he proposed a binomial model to combine accident frequencies. The difficulties in that modelling approach mainly lie in the choice of the control area and in the 'free from error' assumption he makes about control coefficients. Hauer (1991) suggested a notion of similarity to choose a comparison site. However, it seems rather difficult to ground the choice of a comparison site only on this notion of similarity. It is vital that the accident number on the supposed comparison site or control site should be large enough whatever the accident types for the random phenomenon inherent to accident counting on this site not to impair the accuracy of the estimation of the safety measure's average effect. Our approach includes that random phenomenon on the control site data by generalising Tanner's model through the random distribution of the control coefficients. The model's construction and the estimation method stem from those of N'Guessan (1993), N'Guessan et al. (2001, 2005).

This material is organised as follows. Section 2 presents the available before–after accident data tables and the main purpose of the paper. Section 3 deals with construction mechanism of the parametric statistical pattern used to combine simultaneously the accident data tables of experimental sites and control areas. We build a multinomial model for each experimental site taking into account the periods, the different accident types and the parameters to be estimated. We also give some remarks to justify the choice of our model choice. Section 4 presents the estimation method for the unknown vector parameter (i.e. the measure's average effect and accident risks in control areas) based on the constrained maximum likelihood estimation. Section 5 is exclusively centred on the numerical simulated results. We first deal with the way we simulate the available Tables 1 and 2. Then we prove the numerical existence and convergence of a solution for the constrained maximum likelihood equations using different simulated accident data tables. Thereafter, we study the existence and convergence of both the measure's average effect and the accident risks in each control area. Section 6 puts forward the general analysis of the simulated results and some concluding remarks. Section 7 presents some discussions about the statistical pattern building strategy used to combine data from experimental sites and from control areas. Further simulated results about existence and convergence are given in Appendix A. The constrained log-likelihood function and its first derivatives are displayed in Appendix B.

2. Data and main purpose

This paper considers that a road safety measure (crossroad lay-out, surface of a motorway section, etc.) has been applied to S sites. Each target site counts R ($R > 1$) mutually exclusive accident types and is linked to a specific control area where the measure is not directly applied. Let us suppose that the accident numbers are known for each site and each type before and after the measure for significant periods (often equal to and generally longer than one year). Then the available data are found in Table 1, where n_{1rs} (resp. n_{2rs}) stands for the number of type r accidents on site s before (resp. after) the setting up of the measure, and $n_s = \sum_{t=1}^2 \sum_{r=1}^R n_{trs}$ for the total accident number recorded at each site s ($s = 1, 2, \dots, S$). Let us also suppose that, for the same periods, we know the accident data for control areas which are linked to the target sites. The available control coefficients data are found in Table 2, where c_{rs} is the multiplying coefficient which has to be used in the control area linked to site s and for type r accidents to convert the accident number in the before period into the accident number in the after period. Due to external factors (such as traffic flow, speed limit variation, weather conditions, etc.) and random phenomena inherent to accident counting, we also suppose

ARTICLE IN PRESS

Table 1

	Before					After					Total
	Type 1	...	Type <i>r</i>	...	Type <i>R</i>	Type 1	...	Type <i>r</i>	...	Type <i>R</i>	
Site 1	n_{111}	...	n_{1r1}	...	n_{1R1}	n_{211}	...	n_{2r1}	...	n_{2R1}	n_1
⋮	⋮		⋮		⋮	⋮		⋮		⋮	⋮
Site <i>s</i>	n_{11s}	...	n_{1rs}	...	n_{1Rs}	n_{21s}	...	n_{2rs}	...	n_{2Rs}	n_s
⋮	⋮		⋮		⋮	⋮		⋮		⋮	⋮
Site <i>S</i>	n_{11S}	...	n_{1rS}	...	n_{1RS}	n_{21S}	...	n_{2rS}	...	n_{2RS}	n_S
											<i>N</i>

Table 2

	Type 1	...	Type <i>r</i>	...	Type <i>R</i>
Control area 1	c_{11}	...	c_{r1}	...	c_{R1}
⋮	⋮		⋮		⋮
Control area <i>s</i>	c_{1s}	...	c_{rs}	...	c_{Rs}
⋮	⋮		⋮		⋮
Control area <i>S</i>	c_{1S}	...	c_{rS}	...	c_{RS}

that the control coefficients are not free from error i.e. we make the assumption that the control coefficients linked to each site come from an unknown probability distribution.

So the main purpose of this paper is to estimate the measure's average effect for the whole set of sites and for the whole set of types including the estimation of the random distribution of the control coefficients among the accident types. One then has to combine Tables 1 and 2 data to efficiently estimate both the average effect of the measure and the random distribution (interpreted as accident risks) connected to each control area.

3. Statistical pattern building strategy

In the following we denote by X_{1rs} (resp. X_{2rs}) the random variable giving the type *r* accident number on experimental site *s* in the period before (resp. after), one observed value of which is n_{1rs} (resp. n_{2rs}). We are here interested mainly in the joint probability distribution of the 2*R*- dimensional random vector $X^{(s)} = (X_1^{(s)}, X_2^{(s)})$ for a fixed experimental site *s* i.e. the joint distribution of the sequence

$$(X_{11s}, X_{12s}, \dots, X_{1Rs}, X_{21s}, X_{22s}, \dots, X_{2Rs})$$

given n_s the total number of accidents recorded at site *s*. The retrospective repartition of n_s accidents accounts into 2*R*-types of accident is analogue to the random placement of n_s indistinguishable objects into 2*R*-containers. It is similar to the partitions of combinatorial analysis. So we make the assumption that the sequence

$$(X_{11s}, X_{12s}, \dots, X_{1Rs}, X_{21s}, X_{22s}, \dots, X_{2Rs})$$

is the result of n_s independent trials with 2*R*-possible outcomes, each trial involving the random allocation of one of the n_s observations to one of the 2*R*-types of accident. Hence the natural joint probability of the $(X_{11s}, X_{12s}, \dots, X_{1Rs}, X_{21s}, X_{22s}, \dots, X_{2Rs})$ is given by a multinomial distribution denoted

$$\mathcal{M}(n_s; \pi_{11s}, \pi_{12s}, \dots, \pi_{1Rs}, \pi_{21s}, \pi_{22s}, \dots, \pi_{2Rs}), \tag{1}$$

ARTICLE IN PRESS

4

A. N'Guessan et al. / Computational Statistics & Data Analysis ■■■ (■■■) ■■■-■■■

where

$$\sum_{r=1}^R \sum_{t=1}^2 \pi_{trs} = 1,$$

with π_{trs} ($0 < \pi_{trs} < 1$) the theoretical cell probability associated to the random component X_{trs} .

So we have to construct a link function (for a fixed s) between π_{trs} and the parameters describing both the mean effect of the measure linked to Table 1 and the probability distribution linked to each data line in Table 2. The construction of this link function is based on a forecast principle used in the modelling of a road safety measure's effect. Actually, studies about the road safety effect of interventions are usually retrospective quasi-experiments. Therefore, one key task is to predict what the safety of the experimental site would have been without the intervention i.e. if that site had a similar performance to that of the linked comparison site. Such predictions can be made by several methods, one of them is to combine fittingly control area data and the hypothesis according to which the measure has a multiplicative effect (see for instance Oppe, 1979; Hauer et al., 1991) on expected numbers of accidents in the after period whatever the site and the type of accident may be. We will use this latter idea in the statistical pattern building strategy.

So given data from Tables 1 and 2, and if experimental site s has a similar performance to that of the linked control area, then we assume that the mean number of type r accidents that could be expected in the after period is given by

$$\theta^* c_s^* n_{1rs} \quad (r = 1, 2, \dots, R), \tag{2}$$

where

$$\theta^* = \left(\sum_{s=1}^S \sum_{j=1}^R n_{2js} \right) / \left(\sum_{s=1}^S c_s^* \sum_{j=1}^R n_{1js} \right), \quad c_s^* = \sum_{r=1}^R c_{rs} p_{rs}^*,$$

$$p_{rs}^* = n_{1rs} / \left(\sum_{j=1}^R n_{1js} \right), \tag{3}$$

are respectively the apparent average effect of the safety measure (i.e. the ratio of the total number of accidents recorded in the after period to the total number of accidents which could be expected in that period if each site performed as its own control area), the apparent mean control coefficient of site s (i.e. the sum of c_{rs} values in Table 2 weighted by values of p_{rs}^*), and p_{rs}^* the apparent risk of type r accidents of site s with $\sum_{r=1}^R p_{rs}^* = 1$, for a fixed site s . So under the same assumptions and considering the before and after periods, the total number of expected accidents in site s is estimated thanks to

$$n_s^* = (1 + \theta^* c_s^*) \sum_{r=1}^R n_{1rs}. \tag{4}$$

Then if site s performs as its control area and if the measure's average effect is estimated by the apparent average effect θ^* , we deduce an estimate π_{trs}^* of the theoretical probability π_{trs} as follows:

$$\pi_{1rs}^* = \frac{p_{rs}^*}{1 + \theta^* c_s^*}, \quad \pi_{2rs}^* = \frac{\theta^* c_s^* p_{rs}^*}{1 + \theta^* c_s^*} \quad (r = 1, 2, \dots, R). \tag{5}$$

Through construction, the values of π_{trs}^* represent the apparent repartitions of n_s^* accidents into $2R$ mutually exclusive categories between the before and the after periods if site s performs as its control area and if the measure's average effect is estimated by θ^* .

Now, we know that, in practice, an experimental site will never exactly behave as its associated control area where the measure is not directly enforced. It is therefore important to randomly combine Tables 1 and 2 data if we want to model correctly probabilities π_{trs} according to the measure's mean effect and the accident risks in the

ARTICLE IN PRESS

A. N'Guessan et al. / Computational Statistics & Data Analysis ■■■ (■■■) ■■■–■■■

5

control areas. So let us denote by $\theta(\theta > 0)$ the unknown parameter average effect of the road safety measure for all sites and accident types and let us define Z_s the random variable taking values $c_{1s}, c_{2s}, \dots, c_{Rs}$ of Table 2 line s with

$$0 < p_{js} = \text{Proba}[Z_s = c_{js}] \quad \text{and} \quad \sum_{j=1}^R p_{js} = 1 \quad (s = 1, 2, \dots, S). \quad (6)$$

Drawing our inspiration from the structure of the proportions $\pi_{tr_s}^*$ above, we then assume that for a fixed site s the joint probability of the sequence

$$(X_{11s}, X_{12s}, \dots, X_{1Rs}, X_{21s}, X_{22s}, \dots, X_{2Rs})$$

is given by the multinomial distribution

$$\mathcal{M}(n_s; \pi_{11s}, \pi_{12s}, \dots, \pi_{1Rs}, \pi_{21s}, \pi_{22s}, \dots, \pi_{2Rs})$$

with

$$\pi_{tr_s} = g(\theta, P_s) = \begin{cases} \frac{p_{rs}}{1 + \theta c_s}, & t = 1, \quad r = 1, 2, \dots, R, \\ \frac{\theta p_{rs} c_s}{1 + \theta c_s}, & t = 2, \quad r = 1, 2, \dots, R, \end{cases} \quad (7)$$

where $P_s = (p_{1s}, p_{2s}, \dots, p_{Rs})^T$ is the unknown $R \times 1$ vector of accident risks connected to the control area s , $c_s = \sum_{j=1}^R c_{js} p_{js}$ is the unknown mean control coefficient of control area s i.e. the mean of random variable Z_s .

Remark 1. Another way of building the joint distribution above is to use the well-known conditional Poisson model property. In fact, assuming that $X_{11s}, X_{12s}, \dots, X_{1Rs}, X_{21s}, X_{22s}, \dots, X_{2Rs}$ is a sequence of independent random variables where X_{1rs} (resp. X_{2rs}) is a Poisson distribution with mean λ_{1rs} (resp. $\theta c_s \lambda_{1rs}$), then their conditional joint distribution (conditional to the total number of accidents in both the before and the after periods) is a multinomial distribution similar to model (7) above. We will use this common Poisson model assumption in the simulated data below.

Remark 2. From some points of view, an alternative model choice (see for instance N'Guessan et al., 2001, 2005a,b; N'Guessan, 2003a–c) of the probabilities in the after period is $\theta c_{rs} p_{rs} / (1 + \theta c_s)$. Our choice here i.e. the use of c_s instead of c_{rs} in the after period is also governed by some further motivations. The first one is that the individual control coefficients are usually unstable and not free from error because calculated from ratio of accidents in control areas whereas mean values c_s are more stable, free from error and constant. The second one is that our model represents a natural generalisation of Tanner (1958)'s binomial model.

4. Constrained maximum likelihood estimation of average effect and accident risks

Through construction of the pattern, two types of parameters can be spotted out among those linked to S target sites and R accident types: θ , the interest parameter, and the set of auxiliary parameters $P_s = (p_{1s}, p_{2s}, \dots, p_{Rs})^T$, $s = 1, 2, \dots, S$. As this paper unfolds, one focuses on the simultaneous estimation of those parameters and one then poses $\Theta = (\theta, P_1^T, P_2^T, \dots, P_S^T)^T$, the element from \mathbb{R}^{1+SR} thus defined (or actually from $\mathbb{R}^{1+S(R-1)}$ since $\sum_{r=1}^R p_{rs} = 1$). We then assume that the sequence of random vectors $X^{(1)}, X^{(2)}, \dots, X^{(S)}$ are independent and follow a multinomial distribution whose form is given by (7). So according to Tables 1 and 2 and the log-likelihood function linked displayed in the appendix, we show that the constrained maximum likelihood estimator $\hat{\Theta}$ of Θ , if existing, is the solution

ARTICLE IN PRESS

6

A. N'Guessan et al. / Computational Statistics & Data Analysis ■■■ (■■■■) ■■■–■■■

to the system of non-linear equations:

$$\begin{cases} \sum_{s=1}^S \sum_{r=1}^R \frac{n_{2rs} - \hat{\theta} \hat{c}_s n_{1rs}}{1 + \hat{\theta} \hat{c}_s} = 0, \\ n_{.rs} - \frac{n_s \hat{p}_{rs} (\hat{\theta} c_{rs} + 1)}{1 + \hat{\theta} \hat{c}_s} - \frac{n_{2,s} \hat{p}_{rs} (\hat{c}_s - c_{rs})}{\hat{c}_s} = 0, \\ (s = 1, 2, \dots, S; \quad r = 1, 2, \dots, R), \\ \hat{c}_s = \sum_{r=1}^R c_{rs} \hat{p}_{rs} \quad (s = 1, 2, \dots, S), \end{cases} \tag{8}$$

under the constraints

$$\begin{cases} \hat{\theta} > 0, \quad 0 < \hat{p}_{rs} < 1, \\ \text{and} \\ \sum_{r=1}^R \hat{p}_{rs} = 1 \quad (s = 1, 2, \dots, S). \end{cases}$$

Let us note that $\hat{\theta}$, the solution to the system above, is clearly a maximum likelihood estimation method under restraints (RMLE) which is an approach to estimation that maximizes the likelihood over a restricted space. This is not the place to study theoretical sufficient conditions for existence and consistency of $\hat{\theta}$. The general framework for RMLE is well-known and is not discussed here (see for instance Aitchison and Silvey, 1958; Silvey, 1959; Crowder, 1984; Neuenschwander and Flury, 1997; Haber and Brown, 1986; Matthews and Crowther, 1995). We here numerically study the existence and consistency of $\hat{\theta}$ and thereafter the probability distribution of each random variable Z_s ($s=1, 2, \dots, S$) via wide simulation results and graphic (curves and histograms) comparisons.

5. Numerical simulated results

The main object of the simulation is first to look for a solution (i.e. existence) to the constrained equations above and secondly to focus on the convergence of components $\hat{P}_1, \hat{P}_2, \dots, \hat{P}_S$ of $\hat{\theta}$ in order to estimate the probability distribution of control ratios connected to each site. The numerical study of the constrained likelihood equations above and therefore of the numerical existence and convergence of $\hat{\theta}$ imply the knowledge of data Tables 1 and 2. So numerical simulated results count two main phases. The first phase deals with the simulation of control coefficients c_{rs} (Table 2) and of the before–after data table (Table 1) by a FORTRAN programming of the multinomial pattern built in Section 3. Then comes the resolution phase proper (the numerical research of $\hat{\theta}$) with the use of NAG (2003) software of optimisation algorithms referring to the above-mentioned tables.

5.1. Simulation principle of before–after data with control area

Given S (the number of target sites) and R (the number of accident types), one supposes that coefficients c_{rs} ($s = 1, 2, \dots, S; r = 1, 2, \dots, R$) come from a random variable of law $U_{|1/2, 5/2|}$ (a uniform law between $\frac{1}{2}$ and $\frac{5}{2}$). To fill in Table 1, one poses $\theta^0 = \frac{1}{2}$ and assumes that the components of vector $P_s^0 = (p_{1s}^0, p_{2s}^0, \dots, p_{Rs}^0)^T$, the true value of vector P_s , come from a random variable of law $U_{[\alpha, 1-\alpha]}$ (with $\alpha = 10^{-5}$ so that the log-likelihood can be written) and are normalised such as $\sum_{r=1}^R p_{rs}^0 = 1; s = 1, 2, \dots, S$. Using those values, one then defines the true mean control ratio for each site as follows:

$$c_s^0 = \sum_{r=1}^R c_{rs} p_{rs}^0 \quad (s = 1, 2, \dots, S)$$

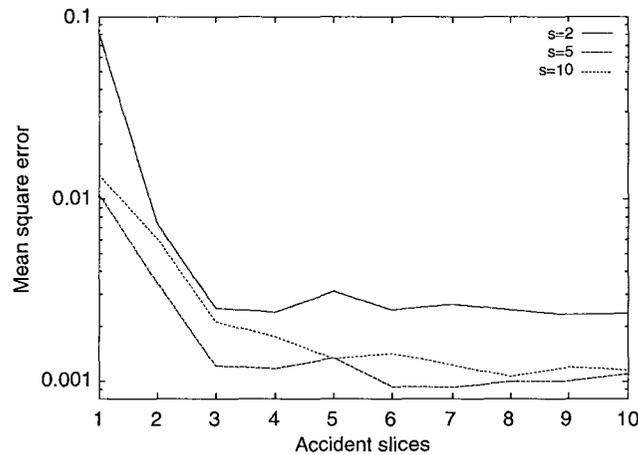
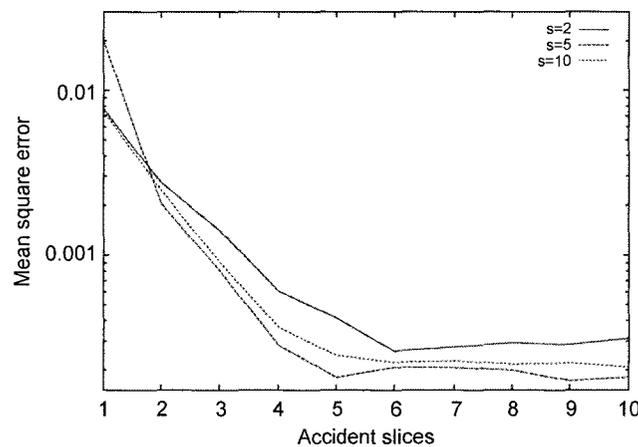
and then the true probabilities

$$\pi_{1rs}^0 = \frac{P_{rs}^0}{1 + \theta^0 c_s^0} \quad \pi_{2rs}^0 = \frac{P_{rs}^0 \theta^0 c_s^0}{1 + \theta^0 c_s^0} \quad r = 1, 2, \dots, R$$

ARTICLE IN PRESS

A. N'Guessan et al. / Computational Statistics & Data Analysis ■■■ (■■■■) ■■■-■■■

7

Fig. 1. $R = 4$.Fig. 2. $R = 10$.

linked to the multinomial distribution of $X^{(s)}$. Finally, for a set value of s ($s = 1, 2, \dots, S$) we generate the total number n_s of accidents on site s from a Poisson distribution. We then retrospectively distribute n_s among the $2R$ -types of accidents using both the probabilities π_{1rs}^0 and π_{2rs}^0 and a FORTRAN program. The observed values of n_{trs} of Table 1 with $\sum_{r=1}^R \sum_{t=1}^2 n_{trs} = n_s$, $s = 1, 2, \dots, S$ are then found. Several tables of different sizes and with different numbers have been generated by setting the values of S and R and by choosing n_s in different accident intervals. The numerical results are found for n_s belonging to the following accident intervals T_1, \dots, T_{10} :

$$\{]0, 30[,]30, 50[,]50, 300[,]300, 1000[,]1000, 3000[,]3000, 5000[,]5000, 7000[,]7000, 9000[,]9000, 11\,000[,]11\,000, 13\,000[\}.$$

5.2. Numerical existence and convergence of $\hat{\Theta}$

The true value of parameter vector $\Theta \in \mathbb{R}^{1+SR}$ which has to be estimated is

$$\Theta^0 = \left(\theta^0, P_1^{0T}, P_2^{0T}, \dots, P_S^{0T} \right)^T$$

with θ^0 set to 0.5 and $P_1^0, P_2^0, \dots, P_S^0$ being vectors of dimension R . We look for a solution $\hat{\Theta}$ to the likelihood equations and then study the convergence of $\hat{\Theta}$ to Θ^0 using the analysis of the mean square error (MSE) which measures the

ARTICLE IN PRESS

8

A. N'Guessan et al. / Computational Statistics & Data Analysis ■■■ (■■■) ■■■-■■■

adequacy between $\hat{\theta}$ and θ^0 and which is defined by (Figs. 1 and 2)

$$\text{MSE}(\hat{\theta}, \theta^0) = \frac{1}{1 + RS} \sum_{m=1}^{1+RS} (\hat{\theta}_m - \theta_m^0)^2.$$

5.3. Estimation of accident risks in control areas

We compare $\hat{P}_1, \hat{P}_2, \dots, \hat{P}_S$ the RMLE to the true R -dimensional probability vectors $P_1^0, P_2^0, \dots, P_S^0$ by analysing their histogram for different values of S and R . Part of the results are displayed below.

6. General analysis of the results and concluding remarks

Table 3 gives the different values of vector $\hat{\theta}$, the solution (i.e. existence) to likelihood equations for fixed $S = 4$ and $R = 2$. The first column of Table 3 gives the true value of parameter θ^0 . In the next columns we find the solution $\hat{\theta}$ for n_s belonging to some of the accident intervals (column T). Those solutions belong to \mathbb{R}^9 for Table 3. We notice a decreasing order of magnitude of $\text{MSE}(\hat{\theta}, \theta^0)$ when numbers n_s go from]0, 30[to]11 000, 13 000[i.e. the convergence of $\hat{\theta}$ to θ^0 . Table 4 completes this first study of $\text{MSE}(\hat{\theta}, \theta^0)$ for the whole set of the accident intervals (column T) and for relatively high values of S and R . The order of magnitude reaches 10^{-4} for the four columns of Table 4 where the space of the parameters is included in \mathbb{R}^{51} in the first case and in \mathbb{R}^{101} in the second case.

A second set of studies was carried out to evaluate the performance of $\text{MSE}(\hat{\theta}, \theta^0)$ for small and important numbers of accidents. The value of n_s was then set in a given interval and S , the site number, and R , the type of accidents, were made to vary. That enables us to note an important improvement in the estimation and in the convergence when the numbers go from]0, 30[to]1000; 3000[, i.e. when the accident numbers per site grow. That convergence of $\text{MSE}(\hat{\theta}, \theta^0)$ to 0 with the increase of accident numbers can be seen on flow charts 1 and 2 for different set values of S and R . The charts obtained for the first values of S and R clearly show that $\text{MSE}(\hat{\theta}, \theta^0)$ is approximately 10^{-2} for values of $n_s \in]0, 30[$ and 10^{-4} for values of $n_s \in]11 000, 13 000[$. One notes a certain stability of the MSE from charts 1 to 2, i.e. when number R of the accident types grows. Likewise, graphs 1 and 2 as well as Table 4 results show a very good stability of the MSE when the value of the R accident type number is set and when the number of experimental sites S goes from 2 to 5 or from 5 to 10.

Along with those studies, we have analysed the results obtained for the probability distribution connected to each set of control area ratios. Part of the results are recorded in Tables 5 and 6. Thus, the values given in those tables are obtained by using elements of θ^0 and $\hat{\theta}$ of Table 3 except the first component. So we have four experimental sites and two types of accident. One notices (in Tables 5 and 6) that the higher the numbers of accidents ($n_s \in]0, 30[$ and $n_s \in]7000, 9000[$)

Table 3
Existence of $\hat{\theta}$ for $S = 4$ and $R = 2$

$\theta^0 (\times 10^{-1})$	$\hat{\theta}$						
	$T = 1$	$T = 2$	$T = 4 (\times 10^{-1})$	$T = 6 (\times 10^{-1})$	$T = 8 (\times 10^{-1})$	$T = 9 (\times 10^{-1})$	$T = 10 (\times 10^{-1})$
5.0000	5.504×10^{-1}	5.208×10^{-1}	4.690	4.803	4.899	4.490	4.895
0.8430	1.000×10^{-4}	0.6390×10^{-1}	0.6220	0.6390	0.8140	0.888	0.8160
9.156	9.999×10^{-1}	9.361×10^{-1}	9.3780	9.3610	9.1860	9.112	9.184
0.2710	0.223×10^{-1}	1.000×10^{-4}	0.5730	0.2980	0.2700	0.280	0.250
9.728	9.777×10^{-1}	9.999×10^{-1}	9.4260	9.7020	9.7300	9.720	9.750
4.059	5.917×10^{-1}	5.2160×10^{-1}	4.370	4.1880	4.120	4.098	4.1770
5.940	4.082×10^{-1}	4.7830×10^{-1}	5.6290	5.8120	5.880	5.902	5.823
7.163	8.817×10^{-1}	7.1320×10^{-1}	7.445	7.1320	7.166	7.102	7.109
2.836	1.182×10^{-1}	2.8670×10^{-1}	2.5540	2.8670	2.834	2.898	2.891

ARTICLE IN PRESS

A. N'Guessan et al. / Computational Statistics & Data Analysis 111 (2010) 111–111

9

Table 4
Convergence of $\hat{\theta}$ for the 10 accident intervals

T	$R = 2$	$R = 10$
$S = 5$		
1	3.8451×10^{-2}	2.0127×10^{-2}
2	2.0946×10^{-2}	2.1294×10^{-3}
3	1.0205×10^{-2}	9.6242×10^{-4}
4	8.9454×10^{-2}	4.3707×10^{-4}
5	5.0994×10^{-3}	3.3561×10^{-4}
6	3.8658×10^{-3}	2.6104×10^{-4}
7	4.5289×10^{-3}	2.3812×10^{-4}
8	4.8712×10^{-3}	2.3464×10^{-4}
9	4.8080×10^{-3}	1.8590×10^{-4}
10	4.8539×10^{-3}	1.9941×10^{-4}
$S = 10$		
1	2.9566×10^{-2}	7.4789×10^{-3}
2	8.4961×10^{-3}	2.4548×10^{-3}
3	4.2690×10^{-3}	9.2732×10^{-4}
4	4.4180×10^{-3}	3.6441×10^{-4}
5	3.6573×10^{-3}	2.4489×10^{-4}
6	3.7994×10^{-3}	2.2200×10^{-4}
7	3.8811×10^{-3}	2.2637×10^{-4}
8	3.8863×10^{-3}	2.2379×10^{-4}
9	3.9175×10^{-3}	2.2019×10^{-4}
10	3.7677×10^{-3}	2.2000×10^{-4}

Table 5
RMLE of the probability distributions

$n_s \in T_1$	P_1^0	\hat{P}_1
<i>Site 1</i>		
Type 1	0.0843	0.0001
Type 2	0.9156	0.9999
$n_s \in T_1$	P_1^0	\hat{P}_1
<i>Site 2</i>		
Type 1	0.0271	0.0223
Type 2	0.9728	0.9777
$n_s \in T_{12}$	P_1^0	\hat{P}_1
<i>Site 1</i>		
Type 1	0.0843	0.0814
Type 2	0.9156	0.9186
$n_s \in T_{12}$	P_2^0	\hat{P}_2
<i>Site 2</i>		
Type 1	0.0271	0.0270
Type 2	0.9728	0.9730

are, the closer to the true values of P_s^0 the values of \hat{P}_s are. That convergence of the RMLE probability distribution to the true distribution is confirmed by the histograms displayed in Figs. 3–10 and in the appendix (Figs. 11–16). Each block of three figures represents the same number of S and R . We obtain good approximations of the true probability distribution of the control area ratios when n_s shifts from]0, 30[to]11 000, 13 000[. For example Figs. 3–5 concern the

ARTICLE IN PRESS

10

A. N'Guessan et al. / Computational Statistics & Data Analysis ■■■ (■■■■) ■■■-■■■

Table 6
RMLE of the probability distributions

$n_s \in T_i$	P_3^0	\hat{P}_3
Site 3		
Type 1	0.4059	0.5917
Type 2	0.5940	0.4082
$n_s \in T_1$	P_4^0	\hat{P}_4
Site 4		
Type 1	0.7163	0.8817
Type 2	0.2836	0.1182
$n_s \in T_{12}$	P_3^0	\hat{P}_3
Site 3		
Type 1	0.4059	0.4120
Type 2	0.5940	0.5880
$n_s \in T_{12}$	P_4^0	\hat{P}_4
Site 4		
Type 1	0.7163	0.7166
Type 2	0.2836	0.2834

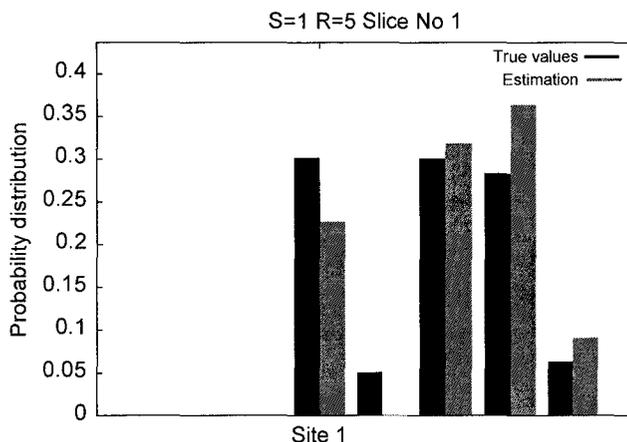


Fig. 3. $R = 5, S = 1$, slice No. 1.

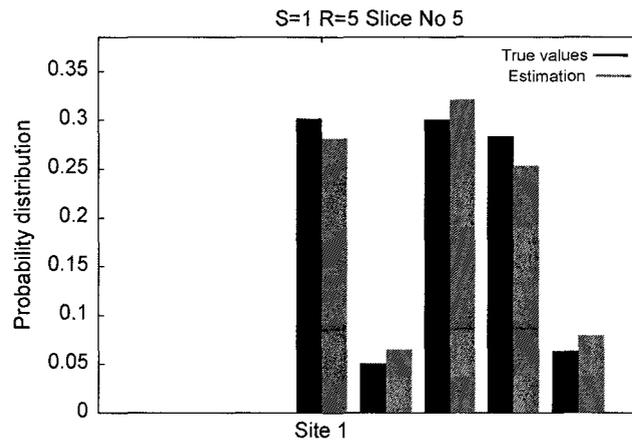
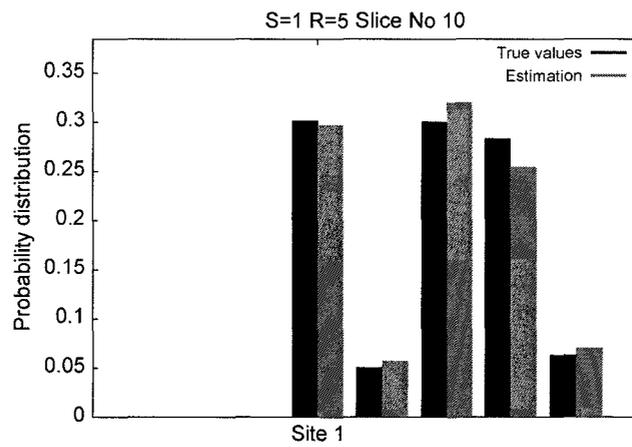
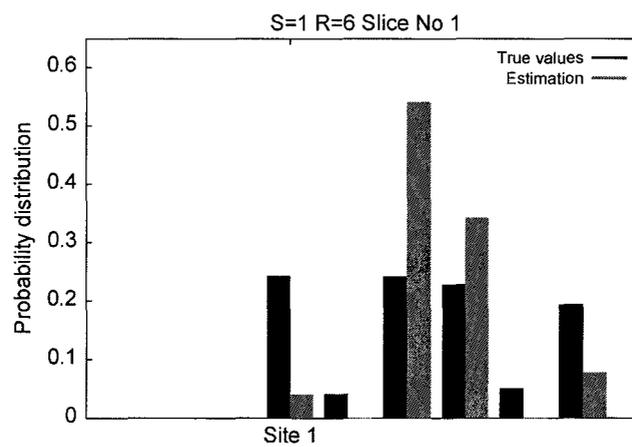
estimation of the probability distribution for one experimental site and five types of accident and three slices of accidents ($[0, 30[,]1000, 3000[,]11\ 000, 13\ 000[$). The black histograms of each figure gives the true probability distribution and the other one the RMLE probability distribution. Figs. 3 and 5 show that the higher the accident number grows, the closer the RMLE probability distribution \hat{P}_s tends to the true distribution P_s^0 . The conclusion is the same for Figs. 15 and 16.

The method's drawback is that the dimension of the parameters' space increases with the number of sites and accident types. Yet the results show that the MSE is a $O(10^{-\alpha})$ with $\alpha > 1$. We also have good fittings between true and estimated values of accident risks. We now are thinking of using the Schur complement approach (see N'Guessan and Langrand, 2005a,b) results about covariance matrices computing in order to get the precision of both the average effect and the accident risks.

ARTICLE IN PRESS

A. N'Guessan et al. / Computational Statistics & Data Analysis ■■■ (■■■■) ■■■-■■■

11

Fig. 4. $R = 5$, $S = 1$, slice No. 5.Fig. 5. $R = 5$, $S = 1$, slice No. 10.Fig. 6. $R = 6$, $S = 1$, slice No. 1.

ARTICLE IN PRESS

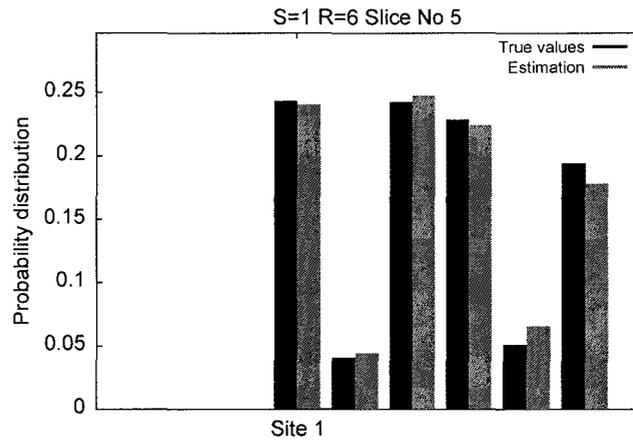


Fig. 7. $R = 6, S = 1$, slice No. 5.

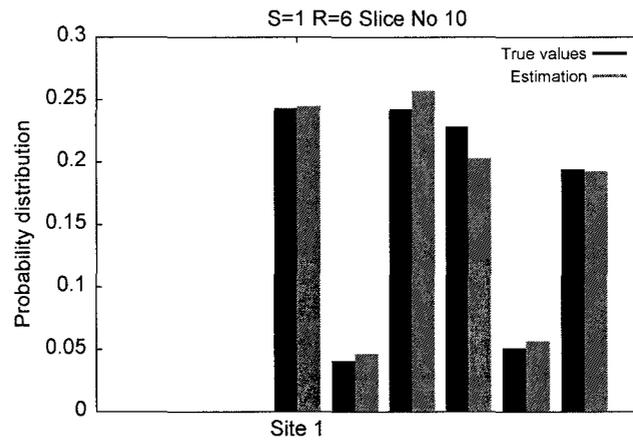


Fig. 8. $R = 6, S = 1$, slice No. 10.

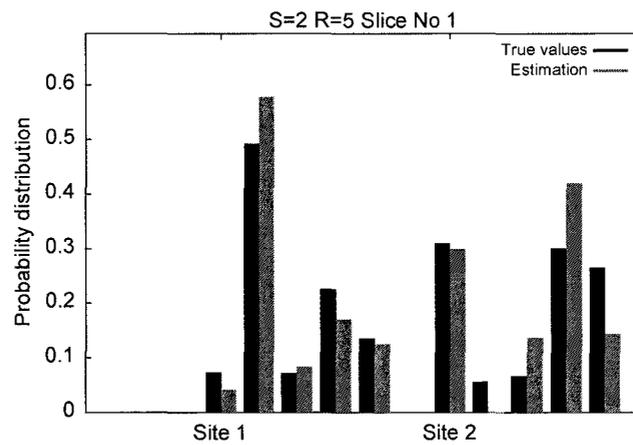
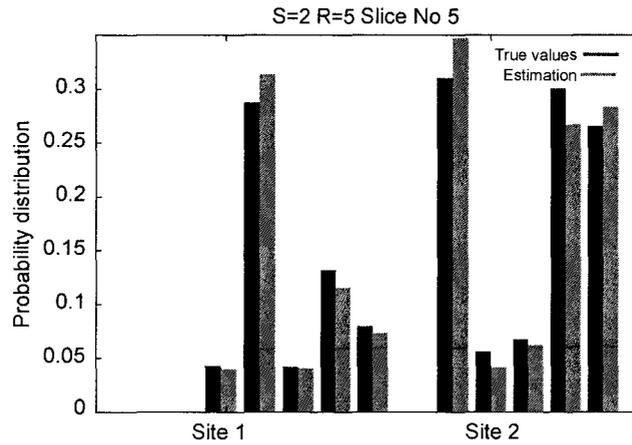


Fig. 9. $R = 5, S = 2$, slice No. 1.

Fig. 10. $R = 5$, $S = 2$, slice No. 5.

7. Discussion

Our main aim is to estimate a road safety measure (intervention)'s mean effect in the presence of several control coefficients per control site. Because of variability and instability in the definition of the control coefficient ratios, we begin with modelling the different control coefficients of each control site. Then we combine this information with the data from the different experimental sites using the multiplicative pattern approach.

7.1. Modelling ratios of two random counts

Since each of the control coefficients $c_{r,s}$ is ratio of two accident counts of the same type in the same control area, these coefficients are therefore automatically random because they are ratios of two random variables. So the assumption that the control coefficients are free from error (see Tanner, 1958) is unrealistic. If the two counts are large, the variance of the estimate of control coefficients will probably be small. But if the two accident counts in the control areas are small then the control coefficients are often not stable and of potentially huge variability. A basic question is then raised: how to model the observations of a random variable coming from the ratio of two random variables? Two answer strategies are possible. The first one consists in considering that the two accident counts enabling to calculate the ratios stem from a certain known or assumed distribution (generally poisson law) and to deduce the ratio's distribution when possible (see for instance Pham-Gia and Turkkan, 2005). The second strategy consists in modelling directly the ratios without making a prior assumption of distributions on the two accident numbers used for the calculation. This is what we present in this paper. Consequently, one defines a probability distribution linked to the $c_{r,s}$ coefficients of each control area. This is the direction we give to the random variables Z_s . The distribution estimation and the use of the average of each Z_s enable to take the variability and the inconstancy problems into account.

7.2. The multiplicative effect assumption in road safety measure modelling

In view of the definitions in (3), the assumption in (2) seems difficult to accept and unrealistic for any real intervention. Really, expression (2) comes from the use of the assumption of the multiplicative effect generally developed in an intervention analysis (Oppe, 1979). In fact, if after the measure, the experimental sites behaved as the linked control sites, we suppose, for a fixed experimental site, that the mean type r accident number one could expect in the period after the intervention is given by

$$\tilde{\mathbb{E}}(X_{2rs}) = \theta^* c_s^* \mathbb{E}(X_{1rs}) \quad (r = 1, 2, \dots, R), \quad (9)$$

where

$$\theta^* = \frac{\left[\sum_{s=1}^S \sum_{r=1}^R \mathbb{E}(X_{2rs}) \right]}{\left[\sum_{s=1}^S \sum_{r=1}^R c_s^* \mathbb{E}(X_{1rs}) \right]},$$

ARTICLE IN PRESS

14

A. N'Guessan et al. / Computational Statistics & Data Analysis ■■■ (■■■■) ■■■–■■■

$c^* = \sum_{r=1}^R c_{rs} p_{rs}^*$ with $p_{rs}^* = \mathbb{E}(X_{1rs}) / \left[\sum_{m=1}^R \mathbb{E}(X_{1ms}) \right]$ are respectively an estimation of the average and of the distribution probability of each random variable Z_s , $\mathbb{E}(X_{1rs})$ (resp. $\mathbb{E}(X_{2rs})$) denotes the unknown mean type r accident number on site s for the before (resp. after) period. Therefore, if under the assumption of a similar behaviour between the experimental sites and their linked control areas, we estimate $\mathbb{E}(X_{1rs})$ by n_{1rs} and $\mathbb{E}(X_{2rs})$ by n_{2rs} the available data in Table 1, then we deduce expressions (2) and (3) from (9).

Acknowledgements

The first author wishes to thank Professor Claude Langrand of Université de Lille 1 for his helpful comments and suggestions. We are also deeply grateful to the two anonymous referees and to the Associate Editor for judicious notes and comments which have enabled a complete remodelling of our paper. Part of this work was carried out while Professor N'Zi was an invited researcher of Université de Lille 1, whose partial support is gratefully acknowledged.

Appendix A

Further numerical results about existence and convergence are given in Figs. 11–16.

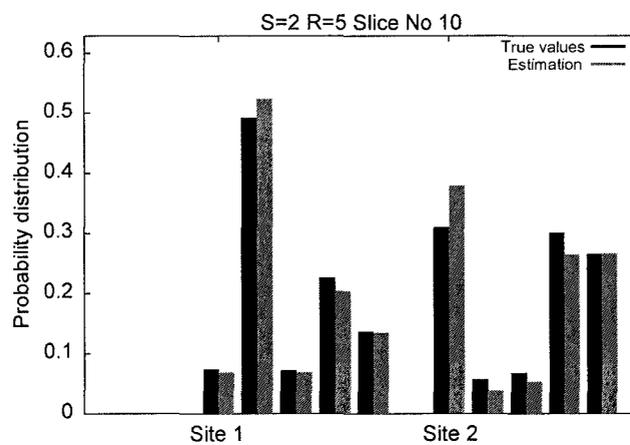


Fig. 11. $R = 5, S = 2$, slice No. 10.

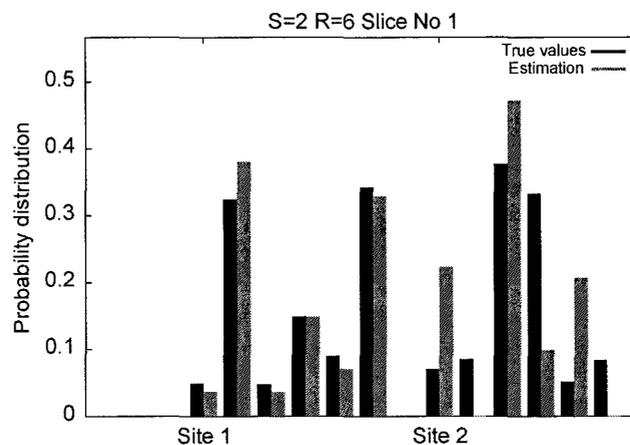
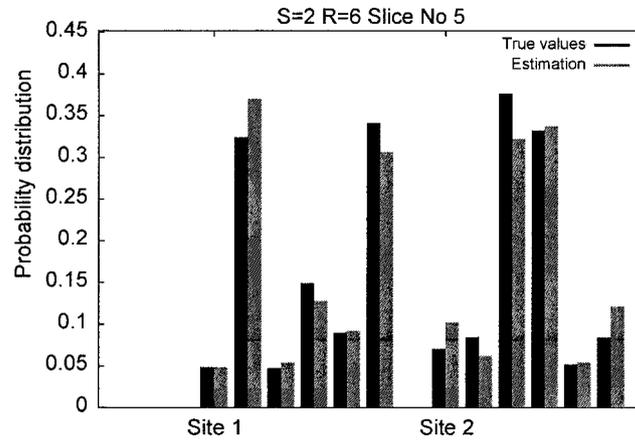
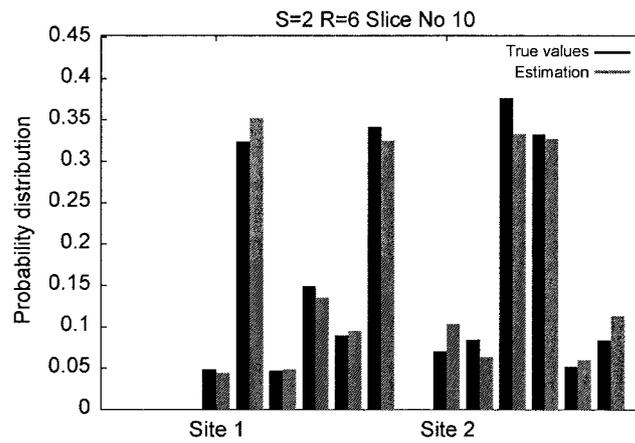
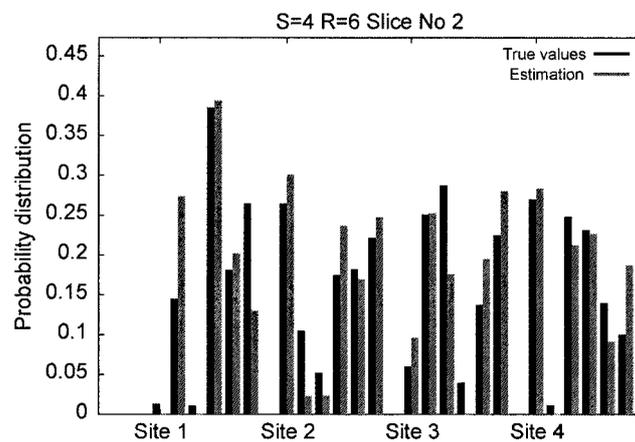


Fig. 12. $R = 6, S = 2$, slice No. 1.

ARTICLE IN PRESS

A. N'Guessan et al. / Computational Statistics & Data Analysis ■■■ (■■■■) ■■■-■■■

15

Fig. 13. $R = 6$, $S = 2$, slice No. 5.Fig. 14. $R = 6$, $S = 2$, slice No. 10.Fig. 15. $R = 6$, $S = 4$, slice No. 2.

ARTICLE IN PRESS

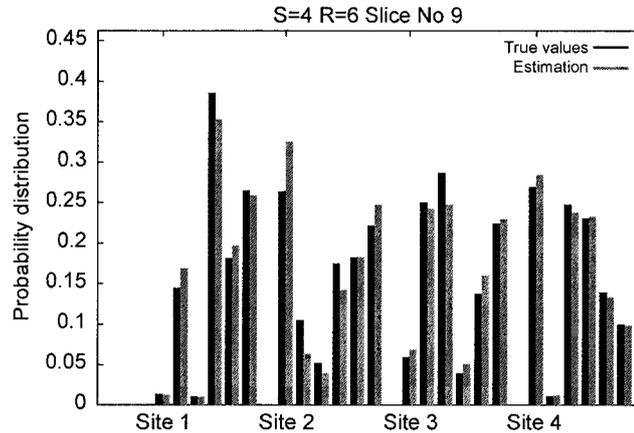


Fig. 16. $R = 6, S = 4$, slice No. 9.

Appendix B. Constrained maximum likelihood equations

Through construction of the pattern, two types of parameters can be spotted out among those linked to S target sites and R accident types: θ , the interest parameter, and the set of auxiliary parameters $P_s = (p_{1s}, p_{2s}, \dots, p_{Rs})^T$, $s = 1, 2, \dots, S$. As this paper unfolds, one focuses on the simultaneous estimation of those parameters and one then poses $\Theta = (\theta, P_1^T, P_2^T, \dots, P_S^T)^T$, the element from \mathbb{R}^{1+SR} thus defined (or actually from $\mathbb{R}^{1+S(R-1)}$ since $\sum_{r=1}^R p_{rs} = 1$). We then assume that the sequence of random vectors $X^{(1)}, X^{(2)}, \dots, X^{(S)}$ is independent and follows a multinomial distribution whose form is given by (7). According to Tables 1 and 2, the likelihood function linked to the sequence of $X^{(1)}, X^{(2)}, \dots, X^{(S)}$ is

$$L(X^{(1)}, X^{(2)}, \dots, X^{(S)}; \Theta) = \prod_{s=1}^S l_s(X^{(s)}; \Theta), \tag{10}$$

$$l_s(X^{(s)}; \Theta) = \frac{n_s!}{\prod_{l=1}^2 \prod_{m=1}^R n_{lms}!} \prod_{r=1}^R \frac{p_{rs}^{n_{rs}} (\theta \sum_{j=1}^R c_{js} p_{js})^{n_{2rs}}}{(1 + \theta \sum_{j=1}^R c_{js} p_{js})^{n_{rs}}}, \tag{11}$$

with $n_{rs} = n_{1rs} + n_{2rs}$. The log-likelihood is given to one constant additive

$$\begin{aligned} \mathcal{L}(X^{(1)}, X^{(2)}, \dots, X^{(S)}; \Theta) = & \sum_{s=1}^S \sum_{r=1}^R \left\{ n_{rs} \log_e(p_{rs}) + n_{2rs} \log_e(\theta) \right. \\ & - n_{rs} \log_e \left(1 + \theta \sum_{m=1}^R c_{ms} p_{ms} \right) \\ & \left. + n_{2rs} \log_e \left(\sum_{m=1}^R c_{ms} p_{ms} \right) \right\}. \end{aligned}$$

It is sufficient to derive the function

$$h(\lambda_1, \lambda_2, \dots, \lambda_S; \Theta) = \sum_{s=1}^S \sum_{r=1}^R \left\{ n_{rs} \log_e(p_{rs}) + n_{2rs} \log_e(\theta) - n_{rs} \log_e \left(1 + \theta \sum_{m=1}^R c_{ms} p_{ms} \right) + n_{2rs} \log_e \left(\sum_{m=1}^R c_{ms} p_{ms} \right) \right\} + \sum_{s=1}^S \lambda_s \left(1 - \sum_{r=1}^R p_{rs} \right)$$

according to parameters Θ where $\lambda_1, \lambda_2, \dots, \lambda_S$ are Lagrange's multipliers. We obtain:

$$\frac{\partial h}{\partial \theta} = \sum_{s=1}^S \sum_{r=1}^R \left(\frac{n_{2rs} - n_{1rs} \theta c_s}{\theta (1 + \theta c_s)} \right),$$

and for $k = 1, \dots, S; j = 1, \dots, R$

$$\frac{\partial h}{\partial p_{jk}} = \frac{n_{jk}}{p_{jk}} - \frac{n_k \theta c_{jk}}{1 + \theta c_k} + \frac{n_{2,k} c_{jk}}{c_k} - \lambda_k,$$

where $n_{2,k} = \sum_{m=1}^R n_{2mk}$. Multiplying this last equation by p_{jk} , summing on j and setting to zero, we deduce that

$$\hat{\lambda}_k = n_{2,k} + \frac{n_k}{1 + \theta c_k}, \quad k = 1, \dots, S.$$

Setting the partial derivatives of h to zero and replacing $\hat{\lambda}_k$ by its expression, we obtain the maximum likelihood equations under the linear restraints of paragraph 4.

References

- Abdel-Aty, M.A., Radwan, A.E., 2000. Modeling traffic accident occurrence and involvement. *Accid. Anal. Prev.* 32, 633–642.
- Aitchison, J., Silvey, S.D., 1958. Maximum likelihood estimation of parameters subject to restraints. *Ann. Math. Statist.* 29, 813–829.
- Allain, E., Brenac, T., 2001. The application of generalized linear models to the investigation of accident frequencies at road sites. The Poisson model and its extensions. *Recherche Transports Sécurité* 72, 3–18.
- Barry, S., Ginpil, S., O'Neill, T.J., 1999. The effectiveness of air bags. *Accid. Anal. Prev.* 31, 781–787.
- Beenstock, M., Gafni, D., Goldin, E., 2001. The effect of traffic policing on road safety in Israel. *Accid. Anal. Prev.* 33, 73–80.
- Crowder, M., 1984. On the constrained maximum likelihood estimation with non i.i.d. observations. *Ann. Inst. Statist. Math.* 36 (A), 239–249.
- Danielson, S., 1986. A comparison of two methods for estimating the effect of a countermeasure in the presence of regression effects. *Accid. Anal. Prev.* 18 (1), 13–23.
- Danielson, S., 1988. Estimation of the effects of countermeasures on different types of accidents in the presence of regression effects. *Accid. Anal. Prev.* 20 (4), 289–298.
- Gaudry, M., Lassarre, S., 2000. Structural Road Accidents Models. The International DRAG Family. Pergamon, Oxford.
- Haber, M., Brown, M.B., 1986. Maximum likelihood methods for log-linear models when expected frequencies are subject to linear constraints. *JASA* 81 (394), 477–482.
- Hauer, E., 1991. Comparison groups in road safety studies: an analysis. *Accid. Anal. Prev.* 23 (26), 609–622.
- Hauer, E., 1997. Observational Before–After Studies in Road Safety Estimating the Effect of Highway and Traffic Engineering Measures on Road Safety. Pergamon, Oxford.
- Hauer, E., 2001. Overdispersion in modelling accidents on road sections and in empirical Bayes estimation. *Accid. Anal. Prev.* 33, 799–808.
- Hauer, E., Ng, J.C.N., Papaioannou, P., 1991. Prediction in road safety studies: an empirical inquiry. *Accid. Anal. Prev.* 23 (6), 595–607.

ARTICLE IN PRESS

18

A. N'Guessan et al. / Computational Statistics & Data Analysis ■■■ (■■■■) ■■■–■■■

- Lassarre, S., 1977. A propos des tests statistiques, sur variables poissonniennes, utilisées dans le domaine de la sécurité routière, *Revue de Statistique Appliquée* XXV (3).
- Lassarre, S., 2001. Analysis of progress in road safety in ten European countries. *Accid. Anal. Prev.* 33, 743–751.
- Mathews, G.B., Crowther, N.A.S., 1995. A maximum likelihood estimation procedure when modelling in terms of constraints. *South African Statist. J.* 29, 29–50.
- NAG (2003). Numerical Algorithms Groups Limited, 2000. Ecole Polytechnique Universitaire de Lille, Université de Lille I.
- Neuenschwander, B.E., Flury, B.D., 1997. A note on Silvey's (1959) theorem. *Statist. & Probab. Lett.* 36, 307–317.
- N'Guessan, A., 1993. Contribution à l'analyse statistique d'une mesure de sécurité routière. Ph.D. Thesis, Université de Lille 1, France.
- N'Guessan, A., 2003a. On a use of Schur complements for an inverse of a constrained accident data information matrix. PUB. IRMA, LILLE 60 (VI).
- N'Guessan, A., 2003b. Constrained covariance matrix estimation in road accident modelling with Shur complements. *C.R. Acad. Sci. Paris Ser. I* 337, 219–222.
- N'Guessan, A., 2003c. Constrained estimation of a road safety covariance matrix using Shur complements. Technical Report, Centre for Research on Transportation, No. CRT-2003-12, Université de Montréal, 18pp.
- N'Guessan, A., Langrand, C., 2005a. A Schur complement approach for computing subcovariance matrices arising in a road safety measure modelling. *J. Comput. Appl. Math.* 177, 331–345.
- N'Guessan, A., Langrand, C., 2005b. A covariance components estimation procedure when modelling a road safety measure in terms of linear constraints. *Statistics* 39 (4), 303–314.
- N'Guessan, A., Essai, A., Langrand, C., 2001. Estimation multidimensionnelle des contrôles et de l'effet moyen d'une mesure de sécurité routière. *Rev. Statistique Appliquée* XLIX (2), 83–100.
- Nicholson, A.J., 1985. The variability of accident counts. *Accid. Anal. Prev.* 17, 47–56.
- Nicholson, A.J., 1986. The randomness of accident counts. *Accid. Anal. Prev.* 18, 193–198.
- Nicholson, A., Wong, Y.-D., 1993. Are accidents Poisson distributed? A statistical test. *Accid. Anal. Prev.* 23 (26), 609–622.
- Oppe, S., 1979. The use of multiplicative models for analysis of road safety data. *Accid. Anal. Prev.* 11, 101–115.
- Pham-Gia, T., Turkkan, N., 2005. Distributions of ratios of random variables from the power-quadratic exponential family and applications. *Statistics* 39 (4), 355–372.
- Silvey, S.D., 1959. The Lagrangian multiplier test. *Ann. Math. Statist.* 30, 389–407.
- Tanner, J.C., 1958. A problem in the combination of accident frequencies. *Biometrika* 45, 331–342.

Annexe A03 :

Journal of Computational and Applied Mathematics, Vol. 177, 331-345 (2005)

A Schur complement approach for computing subcovariance matrices arising in a road safety measure modelling



ELSEVIER

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Journal of Computational and Applied Mathematics 177 (2005) 331–345

JOURNAL OF
COMPUTATIONAL AND
APPLIED MATHEMATICSwww.elsevier.com/locate/cam

A Schur complement approach for computing subcovariance matrices arising in a road safety measure modelling

Assi N'Guessan^{a,*}, Claude Langrand^b^a*Ecole Polytechnique Universitaire de Lille et Laboratoire de Mathématiques Appliquées CNRS, FRE 2222, Université de Lille 1, 59655 Villeneuve d'Ascq, Cedex, France*^b*Laboratoire de Mathématiques Appliquées CNRS, FRE 2222.- U.F.R. de Mathématiques pures et appliquées, Université de Lille 1, 59655 Villeneuve d'Ascq, Cedex, France*

Received 9 March 2004; received in revised form 20 July 2004

Abstract

This paper deals with the determination of the analytical expression of a real partitioned matrix inverse. A formal inversion method is suggested through Schur complements and the elements of any inverse matrix block are explained, irrespective of the size of the considered matrix. The suggested approach does not need any matrix inversion numerical program. This methodology is applied to a Fisher information matrix relating to a multidimensional modelling of road accident data when a road safety measure is applied on different experiment sites. An example of formal calculation and interpretation is given to support our approach.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Matrix approximation; Schur complement; Fisher information matrix; Asymptotic covariance matrix; Formal estimation; Road safety measure; Accident data; Multinomial model; Restricted maximum likelihood

1. Introduction

Most statistical studies involving data collecting (random phenomenon modelling, experiment planning, opinion poll, etc.) not only bring out the problems of parameter estimation (looking for optimal solutions) but also those related to the evaluation of the accuracy of those estimations. In many of the

* Corresponding author. Department of Statistics and Computer Sciences, Bât. Polytech'Lille, Université de Lille 1, 59655 Villeneuve d'Ascq, France. Tel.: +33 3 28 76 74 57; fax: +33 3 28 76 73 01.

E-mail addresses: assi.nguessan@polytech-lille.fr (A. N'Guessan), claudelangrand@univ-lille1.fr (C. Langrand).

statistical evaluation problems, particularly in multivariate statistics, interest parameters are not functionally independent, which means there are relations—constraints—between them. These constraints add further difficulties in bringing out the solutions (estimations) and their accuracy. One of the most popular—as well as the most frequently used—statistical tools to overcome those difficulties still used is the inversion of the so-called Fisher information matrix. This matrix definition uses both the notion of partial derivative of the logarithm of the likelihood function and the linear operator called mathematical expectancy, a notion to which we will come back later in our paper. The concept of the Fisher information matrix is well known and of capital importance in the theory of the parameter statistical estimation (see for example [4,7,9,10,18,22]) and even in the theory of linear systems (see for example [17,20]). Indeed this matrix intervenes in many estimation iterative algorithms, with or without constraints, and its inverse, called asymptotic variance and covariance matrix, generates a measure of the estimation accuracy and of the degree of the linear relations which may exist among the different parameters of the considered model.

In this paper, we focus on the analytical expression of the elements of some blocks of the inverse of a particular Fisher information matrix arising from a statistical model analysis of a road safety measure set up by N'Guessan et al. (see [13,16]). There are certainly other methods (for example [6,23,24,27]) to reach this expression but the one we suggest is based on the Schur complement approach [2,8,19,26] and enables us to get the formal expression of the desired elements. This work stems from N'Guessan's one [14]. Our results go further than those of the latter author and explain the analytical structure of the elements of any block of the inverse of the information matrix considered, irrespective of the size of this matrix. The paper is thus organised. Section 2 states all the work notations and hypotheses used thereafter. Section 3 presents the main technical results and sketches their demonstrations. Section 4 offers an illustrative example coming from the multidimensional modelling of a safety road and accident risk measure. We conclude with an appendix describing the mechanism to obtain the analysed Fisher information matrix structure.

2. Notations and assumptions

Let us consider $s (s > 0)$ and $r (r > 1)$ two given integers; $n_k (n_k > 0)$ a given integer, $k = 1, 2, \dots, s$; $z_k = (z_{1k}, z_{2k}, \dots, z_{rk})^T$ a $r \times 1$ vector of given real data with $z_{jk} > 0, k = 1, 2, \dots, s; j = 1, 2, \dots, r$; and note, $\alpha = (\beta_0, \beta^T)^T$ a $(1 + sr) \times 1$ unknown vector parameter, where $\beta_0 (\beta_0 > 0)$ is a real number, $\beta = (\beta_1^T, \beta_2^T, \dots, \beta_s^T)^T \in \mathbb{R}^{sr}$, with $\beta_k = (\beta_{1k}, \beta_{2k}, \dots, \beta_{rk})^T$ a $r \times 1$ vector and $\beta_{jk} > 0$; for $k = 1, 2, \dots, s$ $\gamma_k = n_k / (1 + \beta_0 \langle z_k, \beta_k \rangle)$ a real number, where $\langle \cdot, \cdot \rangle$ is the usual inner product,

$$V_{\alpha,k} = \frac{\beta_0 \gamma_k}{n_k} (z_{1k}, z_{2k}, \dots, z_{rk})^T,$$

a $r \times 1$ vector,

$$\Omega_{\alpha,k} = \text{diag} \left(\frac{1 + \beta_0 z_{1k}}{\beta_{1k}}, \frac{1 + \beta_0 z_{2k}}{\beta_{2k}}, \dots, \frac{1 + \beta_0 z_{rk}}{\beta_{rk}} \right),$$

a $r \times r$ matrix, $B_{\alpha,k} = \gamma_k (\Omega_{\alpha,k} - V_{\alpha,k} V_{\alpha,k}^T)$, a $r \times r$ matrix.

Then, we assume that the unknown vector parameter α is subjected to s set of restraints

(A1) $h_k(\alpha) = 0$, $k = 1, 2, \dots, s$, where functions $h_k : \mathbb{R}^{1+sr} \mapsto \mathbb{R}$ are given by

(A2) $h_k(\alpha) = \langle \mathbf{1}_r, \beta_k \rangle - 1$ with $\mathbf{1}_r = (1, \dots, 1)^T \in \mathbb{R}^r$, the vector of unity. Therefore we note that

$$J_\alpha = \begin{bmatrix} \tau_\alpha & U_\alpha^T \\ U_\alpha & B_\alpha \end{bmatrix}, \quad \Gamma_\alpha = \begin{bmatrix} J_\alpha & H_\alpha^T \\ H_\alpha & 0_{s,s} \end{bmatrix}, \quad (1)$$

where τ_α , U_α , B_α are, respectively, a positive real number, a $(sr) \times 1$ vector and a $(sr) \times (sr)$ matrix given by

$$\tau_\alpha = \sum_{k=1}^s \frac{\gamma_k \langle z_k, \beta_k \rangle}{\beta_0 (1 + \beta_0 \langle z_k, \beta_k \rangle)}, \quad U_\alpha = (U_{\alpha,1}^T, U_{\alpha,2}^T, \dots, U_{\alpha,s}^T)^T, \\ B_\alpha = \text{bloc-diag}(B_{\alpha,1}, B_{\alpha,2}, \dots, B_{\alpha,s})$$

with $U_{\alpha,k} = (\gamma_k / \beta_0) V_{\alpha,k}$ and $H_\alpha = (H_1, H_2)$ a $s \times (1 + sr)$ matrix, where

$$H_1 = \left(\frac{\partial h_1}{\partial \beta_0}, \frac{\partial h_2}{\partial \beta_0}, \dots, \frac{\partial h_s}{\partial \beta_0} \right)^T, \quad H_2 = \left(\frac{\partial h_1}{\partial \beta}, \frac{\partial h_2}{\partial \beta}, \dots, \frac{\partial h_s}{\partial \beta} \right)^T$$

are, respectively, $s \times 1$ and $s \times (sr)$ matrices. In the present paper, we investigate the inversion of matrices J_α and Γ_α under assumptions (A1) and (A2) and using the Schur complement approach (see for instance [19]).

3. Main results

Theorem 3.1. For $k = 1, 2, \dots, s$

- (i) $\|z_k\|_{\Omega_{\alpha,k}^{-1}}^2 < \langle z_k, \beta_k \rangle / \beta_0$;
- (ii) $\|V_{\alpha,k}\|_{\Omega_{\alpha,k}^{-1}}^2 < \beta_0 \langle z_k, \beta_k \rangle / (1 + \beta_0 \langle z_k, \beta_k \rangle)^2 < 1$.

Proof. (i) Using the fact that $x/(1+x) < 1$ for all $x > 0$, one obtains after matrix manipulations, $\|z_k\|_{\Omega_{\alpha,k}^{-1}}^2 < (1/\beta_0) (\sum_{m=1}^r \beta_{mk} z_{mk}) = \langle z_k, \beta_k \rangle / \beta_0$.

(ii) This part follows from (i) and the relation between vectors V_k and z_k . \square

Theorem 3.2. (i) For $k=1, 2, \dots, s$, matrix $B_{\alpha,k}$ is nonsingular and $B_{\alpha,k}^{-1} = \gamma_k^{-1} (\Omega_{\alpha,k}^{-1} + t_k \Omega_{\alpha,k}^{-1} V_{\alpha,k} V_{\alpha,k}^T \Omega_{\alpha,k}^{-1})$ with $t_k = (1 - \|V_{\alpha,k}\|_{\Omega_{\alpha,k}^{-1}}^2)^{-1}$.

- (ii) $\|U_\alpha\|_{B_\alpha^{-1}}^2 < \tau_\alpha$.

Proof. (i) Let us define

$$\Omega_{\alpha,k}^{(1)} = \begin{bmatrix} \Omega_{\alpha,k} & V_{\alpha,k} \\ V_{\alpha,k}^T & 1 \end{bmatrix},$$

334 A. N'Guessan, C. Langrand / Journal of Computational and Applied Mathematics 177 (2005) 331–345

a $(1 + r) \times (1 + r)$ matrix and $(\Omega_{\alpha,k}^{(1)}/\Omega_{\alpha,k})$ (resp. $(\Omega_{\alpha,k}^{(1)}/1)$) the Schur complement of $\Omega_{\alpha,k}$ (resp. of 1) in $\Omega_{\alpha,k}^{(1)}$. Theorem 3.1 implies that

$$(\Omega_{\alpha,k}^{(1)}/\Omega_{\alpha,k}) = 1 - \|V_{\alpha,k}\|_{\Omega_{\alpha,k}^{-1}}^2 > 0.$$

So using Woodburry [29] formula (see also [21, pp. 73–78; 19, pp. 201–204]), we can deduce that $(\Omega_{\alpha,k}^{(1)}/1)$ is nonsingular and

$$(\Omega_{\alpha,k}^{(1)}/1)^{-1} = \Omega_{\alpha,k}^{-1} + \Omega_{\alpha,k}^{-1} V_{\alpha,k} (\Omega_{\alpha,k}^{(1)}/\Omega_{\alpha,k})^{-1} V_{\alpha,k}^T \Omega_{\alpha,k}^{-1}.$$

The result follows from the relation $B_{\alpha,k} = \gamma_k (\Omega_{\alpha,k}^{(1)}/1)$.

(ii) Matrix manipulations and relations on one hand between vectors $U_{\alpha,k}$ and $V_{\alpha,k}$ and on the other hand between matrices $B_{\alpha,k}$ and $\Omega_{\alpha,k}$ imply that

$$\|U_{\alpha}\|_{B_{\alpha}^{-1}}^2 = \frac{1}{\beta_0^2} \sum_{k=1}^s \frac{\gamma_k \|V_{\alpha,k}\|_{\Omega_{\alpha,k}^{-1}}^2}{1 - \|V_{\alpha,k}\|_{\Omega_{\alpha,k}^{-1}}^2}. \tag{2}$$

Using Theorem 3.1, one obtains

$$\|U_{\alpha}\|_{B_{\alpha}^{-1}}^2 < \sum_{k=1}^s \frac{\gamma_k (1 + \beta_0 \langle z_k, \beta_k \rangle)}{\beta_0^2} \|V_{\alpha,k}\|_{\Omega_{\alpha,k}^{-1}}^2 < \tau_{\alpha}. \quad \square \tag{3}$$

Theorem 3.3. *Let us define*

$$B_{\alpha}^{(0)} = \begin{bmatrix} B_{\alpha} & U_{\alpha} \\ U_{\alpha}^T & 0 \end{bmatrix}, \quad \Gamma_{k,m} = \begin{bmatrix} (J_{\alpha}/B_{\alpha}) & U_{\alpha,m}^T B_{\alpha,m}^{-1} \\ B_{\alpha,k}^{-1} U_{\alpha,k} & \delta_{k,m} B_{\alpha,k}^{-1} \end{bmatrix},$$

respectively, a $(1 + sr) \times (1 + sr)$ matrix and a $(1 + r) \times (1 + r)$ matrix with $k, m = 1, 2, \dots, s$ and $\delta_{k,m}$ the Kronecker symbol. Then

- (i) $(J_{\alpha}/B_{\alpha}) = \tau_{\alpha} + (B_{\alpha}^{(0)}/B_{\alpha}) > 0$;
- (ii) $(J_{\alpha}/\tau_{\alpha})$ is a $(sr) \times (sr)$ nonsingular matrix and

$$(J_{\alpha}/\tau_{\alpha})^{-1} = \begin{bmatrix} J_{1,1} & J_{1,2} & \dots & J_{1,s} \\ J_{2,1} & J_{2,2} & \dots & J_{2,s} \\ \vdots & \vdots & \ddots & \vdots \\ J_{s,1} & J_{s,2} & \dots & J_{s,s} \end{bmatrix}, \tag{4}$$

where each block $J_{k,m}$ is a $r \times r$ matrix given by $(\Gamma_{k,m}/(J_{\alpha}/B_{\alpha}))$, the Schur complement of (J_{α}/B_{α}) in $\Gamma_{k,m}$.

Proof. (i) Matrix manipulations and relations on the one hand between vectors $U_{\alpha,k}$ and $V_{\alpha,k}$ and on the other hand between matrices $B_{\alpha,k}^{-1}$ and $\Omega_{\alpha,k}^{-1}$ imply that

$$-(B_{\alpha}^{(0)}/B_{\alpha}) = \frac{1}{\beta_0^2} \sum_{k=1}^s \frac{\gamma_k \|V_{\alpha,k}\|_{\Omega_{\alpha,k}^{-1}}^2}{1 - \|V_{\alpha,k}\|_{\Omega_{\alpha,k}^{-1}}^2}.$$

So doubly using part (ii) of Theorem 3.1, we obtain

$$-(B_{\alpha}^{(0)}/B_{\alpha}) < \sum_{k=1}^s \frac{\gamma_k \langle z_k, \beta_k \rangle}{\beta_0(1 + \beta_0 \langle z_k, \beta_k \rangle)} = \tau_{\alpha}.$$

Therefore, we can deduce that (J_{α}/B_{α}) is nonsingular and

$$(J_{\alpha}/B_{\alpha}) = \tau_{\alpha} + (B_{\alpha}^{(0)}/B_{\alpha}) > 0. \quad (5)$$

(ii) Since B_{α}^{-1} exists and $(J_{\alpha}/B_{\alpha}) > 0$ then, using the same argument as in part (i) of Theorem 3.2, the $(sr) \times (sr)$ matrix $(J_{\alpha}/\tau_{\alpha})$ is nonsingular and

$$(J_{\alpha}/\tau_{\alpha})^{-1} = B_{\alpha}^{-1} + (J_{\alpha}/B_{\alpha})^{-1} B_{\alpha}^{-1} U_{\alpha} U_{\alpha}^T B_{\alpha}^{-1}. \quad (6)$$

The result follows from matrix manipulations. \square

Theorem 3.4. Let us define

$$B_{\alpha,k}^{(0)} = \begin{bmatrix} B_{\alpha,k} & 1_r \\ 1_r^T & 0 \end{bmatrix}, \quad A_{\alpha}^{(0)} = \begin{bmatrix} A_{\alpha} & \xi_{\alpha} \\ \xi_{\alpha}^T & 0 \end{bmatrix}, \quad A_{\alpha}^{(1)} = \begin{bmatrix} A_{\alpha} & -\xi_{\alpha} \\ \xi_{\alpha}^T & (J_{\alpha}/B_{\alpha}) \end{bmatrix},$$

respectively, a $(1+r) \times (1+r)$ matrix and two $(1+s) \times (1+s)$ matrices, where

$$A_{\alpha} = \text{diag}[-(B_{\alpha,1}^{(0)}/B_{\alpha,1}), \dots, -(B_{\alpha,s}^{(0)}/B_{\alpha,s})] \quad (7)$$

is a $s \times s$ matrix and

$$\xi_{\alpha} = H_2 B_{\alpha}^{-1} U_{\alpha} \quad (8)$$

is a $s \times 1$ vector. Then

- (i) $-(B_{\alpha,k}^{(0)}/B_{\alpha,k}) > 0$ for $k = 1, 2, \dots, s$;
(ii) $(A_{\alpha}^{(1)}/(J_{\alpha}/B_{\alpha}))$, the Schur complement of (J_{α}/B_{α}) in $A_{\alpha}^{(1)}$ is a $s \times s$ nonsingular matrix and

$$(A_{\alpha}^{(1)}/(J_{\alpha}/B_{\alpha}))^{-1} = A_{\alpha}^{-1} - A_{\alpha}^{-1} \xi_{\alpha} ((J_{\alpha}/B_{\alpha}) + (A_{\alpha}^{(0)}/A_{\alpha}))^{-1} \xi_{\alpha}^T A_{\alpha}^{-1}. \quad (9)$$

Proof. (i) $B_{\alpha,k}$ is nonsingular so $(B_{\alpha,k}^{(0)}/B_{\alpha,k})$ exists and

$$-(B_{\alpha,k}^{(0)}/B_{\alpha,k}) = 1_r^T B_{\alpha,k}^{-1} 1_r > 0$$

for $k = 1, 2, \dots, s$.

(ii) Part (i) implies that A_x is nonsingular. Since $(J_x/B_x) > 0$, then $(A_x^{(1)}/(J_x/B_x))$ is nonsingular and

$$\begin{aligned} (A_x^{(1)}/(J_x/B_x))^{-1} &= A_x^{-1} - A_x^{-1} \xi_x (A_x^{(1)}/A_x)^{-1} \xi_x^T A_x^{-1} \\ &= A_x^{-1} - A_x^{-1} \xi_x ((J_x/B_x) + (A_x^{(0)}/A_x))^{-1} \xi_x^T A_x^{-1}. \quad \square \end{aligned} \tag{10}$$

Theorem 3.5. Let us note $R_x = -(\Gamma_x/J_x)$, the $s \times s$ matrix given by the negative of the Schur complement of J_x in Γ_x . R_x is nonsingular and

$$R_x^{-1} = A_x^{-1} - ((J_x/B_x) + (A_x^{(0)}/A_x))^{-1} A_x^{-1} \xi_x \xi_x^T A_x^{-1}. \tag{11}$$

Proof. Using Theorems 3.2 and 3.3 and the well-known results of Schur complements about the partitioned matrix inverse (see for instance [19, pp. 201–204]), matrix J_x is nonsingular and

$$J_x^{-1} = \begin{bmatrix} (J_x/B_x)^{-1} & -(J_x/B_x)^{-1} U_x^T B_x^{-1} \\ -B_x^{-1} U_x (J_x/B_x)^{-1} & (J_x/\tau_x)^{-1} \end{bmatrix}. \tag{12}$$

Now using assumptions (A1) and (A2) above and the expression of $(J_x/\tau_x)^{-1}$ (see the proof of Theorem 3.3), we obtain

$$\begin{aligned} R_x &= H_2 (J_x/\tau_x)^{-1} H_2^T \\ &= A_x + (J_x/B_x)^{-1} \xi_x \xi_x^T \\ &= (A_x^{(1)}/(J_x/B_x)). \end{aligned} \tag{13}$$

The result follows from Theorem 3.4. \square

Theorem 3.6. Let us note $\sigma_x^2 = [(A_x^{(1)}/A_x)]^{-1}$, where $(A_x^{(1)}/A_x)$ is the Shur complement of A_x in $A_x^{(1)}$; $\mathbf{1}_{r,r}$ the $r \times r$ matrix of unity; $\Delta_{x,k} = z_k^T \Omega_{x,k}^{-1} \mathbf{1}_r$, $k = 1, 2, \dots, s$; $A_x = ((A_{k,m}))$, a matrix $s \times s$ with (k, m) th element

$$A_{k,m} = \frac{\delta_{k,m}}{\|\mathbf{1}_r\|_{B_{x,k}}^{-1}} - \frac{\sigma_x^2 t_k t_m \gamma_k \gamma_m \Delta_{x,k} \Delta_{x,m}}{n_k n_m \|\mathbf{1}_r\|_{B_{x,k}}^2 \|\mathbf{1}_r\|_{B_{x,m}}^2}. \tag{14}$$

Then

(i)

$$H_2^T R_x^{-1} H_2 = A_x \otimes \mathbf{1}_{r,r}$$

is a $(sr) \times (sr)$ matrix, where \otimes is the Kronecker product;

(ii)

$$(J_x/\tau_x)^{-1} H_2^T R_x^{-1} H_2 (J_x/\tau_x)^{-1} = G_x,$$

where G_x is a $(sr) \times (sr)$ matrix, where each $G_{k,m}$ block is also a $r \times r$ matrix with

$$G_{k,m} = \sum_{j=1}^s \left[\left(\sum_{i=1}^s A_{i,j} J_{k,i} \right) \mathbf{1}_{r,r} \right] J_{j,m}, \tag{15}$$

where $k, m = 1, 2, \dots, s$.

Proof. This proof stems from an algebraic manipulation of matrices. So we give the outline.

(i) Using Theorems 3.4 and 3.5 and the relations

$$\mathbf{1}_{r,r} B_{\alpha,k}^{-1} U_{\alpha,k} U_{\alpha,m}^T B_{\alpha,m}^{-1} \mathbf{1}_{r,r} = \frac{t_k t_m A_{\alpha,k} A_{\alpha,m}}{n_k n_m} \mathbf{1}_{r,r}, \tag{16}$$

we can deduce that

$$H_2^T R_\alpha^{-1} H_2 = A_\alpha \otimes \mathbf{1}_{r,r}. \tag{17}$$

(ii) In view of Theorem 3.3, the $r \times r$ elements $J_{k,m}$ of $(J_\alpha/\tau_\alpha)^{-1}$ are

$$J_{k,m} = \delta_{k,m} B_{\alpha,k}^{-1} - (J_\alpha/B_\alpha)^{-1} B_{\alpha,k}^{-1} U_{\alpha,k} U_{\alpha,m}^T B_{\alpha,m}^{-1}.$$

So using part (ii) of Theorem 3.3, the postmultiplication of $H_2^T R_\alpha^{-1} H_2$ by $(J_\alpha/\tau_\alpha)^{-1}$ and finally the premultiplication of the same matrix, we get

$$\begin{aligned} & (J_\alpha/\tau_\alpha)^{-1} H_2^T R_\alpha^{-1} H_2 (J_\alpha/\tau_\alpha)^{-1} \\ &= \begin{bmatrix} \sum_{j=1}^s \left[\left(\sum_{i=1}^s J_{1,i} A_{i,j} \right) \mathbf{1}_{r,r} \right] J_{j,1} & \sum_{j=1}^s \left[\left(\sum_{i=1}^s J_{1,i} A_{i,j} \right) \mathbf{1}_{r,r} \right] J_{j,2} & \dots & \sum_{j=1}^s \left[\left(\sum_{i=1}^s J_{1,i} A_{i,j} \right) \mathbf{1}_{r,r} \right] J_{j,s} \\ \sum_{j=1}^s \left[\left(\sum_{i=1}^s J_{2,i} A_{i,j} \right) \mathbf{1}_{r,r} \right] J_{j,1} & \sum_{j=1}^s \left[\left(\sum_{i=1}^s J_{2,i} A_{i,j} \right) \mathbf{1}_{r,r} \right] J_{j,2} & \dots & \sum_{j=1}^s \left[\left(\sum_{i=1}^s J_{2,i} A_{i,j} \right) \mathbf{1}_{r,r} \right] J_{j,s} \\ \dots & \dots & \dots & \dots \\ \sum_{j=1}^s \left[\left(\sum_{i=1}^s J_{s,i} A_{i,j} \right) \mathbf{1}_{r,r} \right] J_{j,1} & \sum_{j=1}^s \left[\left(\sum_{i=1}^s J_{s,i} A_{i,j} \right) \mathbf{1}_{r,r} \right] J_{j,2} & \dots & \sum_{j=1}^s \left[\left(\sum_{i=1}^s J_{s,i} A_{i,j} \right) \mathbf{1}_{r,r} \right] J_{j,s} \end{bmatrix} \end{aligned} \tag{18}$$

and thereafter we obtain matrix G_α . \square

Theorem 3.7. Matrix Γ_α is nonsingular and

$$\Gamma_\alpha^{-1} = \begin{bmatrix} W_\alpha & J_\alpha^{-1} H_\alpha^T R_\alpha^{-1} \\ R_\alpha^{-1} H_\alpha J_\alpha^{-1} & -R_\alpha^{-1} \end{bmatrix}, \tag{19}$$

where W_α , the leading $(1 + rs) \times (1 + sr)$ matrix, is defined by

$$W_\alpha = \begin{bmatrix} W_\alpha(1, 1) & W_\alpha^T(2, 1) \\ W_\alpha(2, 1) & W_\alpha(2, 2) \end{bmatrix} \tag{20}$$

with $W_\alpha(1, 1)$, $W_\alpha(2, 1)$ and $W_\alpha(2, 2)$, respectively, a scalar, a $(sr) \times 1$ vector and a $(sr) \times (sr)$ matrix as follows:

$$\begin{aligned} W_\alpha(1, 1) &= (A_\alpha^{(1)}/A_\alpha)^{-1}, \\ W_\alpha(2, 1) &= -(J_\alpha/B_\alpha)^{-1} [B_\alpha^{-1} U_\alpha - (J_\alpha/\tau_\alpha)^{-1} H_2^T R_\alpha^{-1} \xi_\alpha], \\ W_\alpha(2, 2) &= (J_\alpha/\tau_\alpha)^{-1} - (J_\alpha/\tau_\alpha)^{-1} H_2^T R_\alpha^{-1} H_2 (J_\alpha/\tau_\alpha)^{-1}. \end{aligned} \tag{21}$$

Proof (Outline). Combining well-known results about the inverse of partitioned matrix (see for instance [1, Lemma 3]) and the theorems above, we get the expression of Γ_α^{-1} above, where

$$W_\alpha = J_\alpha^{-1} - J_\alpha^{-1} H_2^T R_\alpha^{-1} H_2 J_\alpha^{-1}. \tag{22}$$

So using (A1), (A2) and the expression of J_α^{-1} (see proof of Theorem 3.5), we obtain the block-components of matrix W_α as follows:

$$\begin{aligned} W_\alpha(1, 1) &= (J_\alpha/B_\alpha)^{-1} - (J_\alpha/B_\alpha)^{-1} \xi_\alpha^T R_\alpha^{-1} \xi_\alpha (J_\alpha/B_\alpha)^{-1}, \\ W_\alpha(2, 1) &= -B_\alpha^{-1} U_\alpha (J_\alpha/B_\alpha)^{-1} + (J_\alpha/\tau_\alpha)^{-1} H_2^T R_\alpha^{-1} \xi_\alpha (J_\alpha/B_\alpha)^{-1}, \\ W_\alpha(2, 2) &= (J_\alpha/\tau_\alpha)^{-1} - (J_\alpha/\tau_\alpha)^{-1} H_2^T R_\alpha^{-1} H_2 (J_\alpha/\tau_\alpha)^{-1}. \end{aligned} \tag{23}$$

The result follows from matrices manipulations. \square

Lemma 3.8. *We get the following results:*

(i)

$$\begin{aligned} -(B_{\alpha,k}^{(0)}/B_{\alpha,k}) &= \|\mathbf{1}_r\|_{B_{\alpha,k}^{-1}}^2 \\ &= \gamma_k^{-1} \left[\text{trace}(\Omega_{\alpha,k}^{-1}) + t_k \frac{\beta_0^2 \gamma_k^2 A_{\alpha,k}^2}{n_k^2} \right], \end{aligned} \tag{24}$$

(ii)

$$(J_\alpha/B_\alpha) = \sum_{i=1}^s \frac{\hat{\gamma}_i^2}{n_i} \left(\frac{\langle z_i, \beta_i \rangle}{\beta_0} - \frac{\gamma_i t_i}{n_i} \|z_i\|_{\Omega_{\alpha,i}^{-1}}^2 \right), \tag{25}$$

(iii)

$$\begin{aligned} -(A_\alpha^{(0)}/A_\alpha) &= \|\xi_\alpha\|_{A_\alpha^{-1}}^2 \\ &= \sum_{i=1}^s \frac{\gamma_i^2 t_i^2 A_{\alpha,i}^2}{n_i^2 \|\mathbf{1}_r\|_{B_{\alpha,i}^{-1}}^2}, \end{aligned} \tag{26}$$

(iv)

$$(A_\alpha^{(1)}/A_\alpha) = (J_\alpha/B_\alpha) - (A_\alpha^{(0)}/A_\alpha). \tag{27}$$

Remark 1. Lemma 3.8 provides from matrix manipulations. So we leave out the proof. Details are displayed in a recent technical report [15].

4. Application to road safety measure modelling

4.1. Problem formulation and data

We apply the above results to a statistical road accident data modelling when a road safety measure is applied on different target sites. Our model and estimation method stem from those of N'Guessan et al. [13,16]. The latter authors deal with a multidimensional combination of road accident frequencies before

and after a similar change (crossroad lay-out, surface of a motorway section, etc.) at s ($s > 0$) target sites. Each target site counts r ($r > 1$) accident types (fatal accidents, seriously injured people, slightly injured people, material damage, etc.) and is linked to a specific control area. In order to take some external factors into account (such as traffic flow, speed limit variation, weather conditions, etc.) they suppose that they know the data set $(z_{1k}, z_{2k}, \dots, z_{rk})^T$, ($k = 1, 2, \dots, s$) for the specific control areas, where the measure (or change) is not directly applied, but which are linked to the target sites. They built a statistical model to share the total accident number n_k ($k = 1, 2, \dots, s$) in each site. The unknown vector parameter α of their model depends on the global average effect β_0 of the change as well as on the vector of accident risks $(\beta_{1k}, \beta_{2k}, \dots, \beta_{rk})^T$ of each control area k such that

$$\sum_{j=1}^r \beta_{jk} = 1, \quad (28)$$

where $k = 1, 2, \dots, s$. Their method of estimating simultaneously the global effect and the accident risks in control areas, i.e., α , is based on the maximization of function

$$\mathcal{L}(\alpha) = \text{Cte} + \sum_{k=1}^s \sum_{j=1}^r \{y_{.jk} \log_e(\beta_{jk}) + y_{2jk} \log_e(\beta_0) - y_{.jk} \log_e(1 + \beta_0 \langle z_k, \beta_k \rangle)\} \quad (29)$$

under assumptions (A1) and (A2), where $y_{.jk} = y_{1jk} + y_{2jk}$ and y_{1jk} (resp. y_{2jk}) stands for the number of type j of accidents on site k before (resp. after) the setting up of the change and $n_k = \sum_{j=1}^r y_{.jk}$. Function $\mathcal{L}(\alpha)$ is called log-likelihood and the method proposed is clearly the restricted maximum likelihood estimation (RMLE) method which is an approach to estimation that maximizes the logarithm of a likelihood over a restricted space. This approach of estimation is the well known applied mathematical method that maximizes a function subjected to restraints. The general statistical framework is well known and is not discussed here (see for instance [1, 3, 5, 6, 10, p. 172; 11, 12, 25, 28]).

4.2. Computing information matrix using second partial derivatives

Information matrix Γ_α defined below and related to function $\mathcal{L}(\alpha)$ and assumptions (A1) and (A2) plays an important role in RMLE method. Not only does it provide the solution for unknown vector parameter α but its inverse, if it exists, is also used to compute the covariance matrix of RMLE estimates. The general statistical framework shows that a central role is played by the inverse, if it exists, of the bordered information matrix

$$\Gamma_\alpha = \begin{pmatrix} J_\alpha & H_\alpha^T \\ H_\alpha & 0 \end{pmatrix}, \quad (30)$$

where the $(1 + sr) \times (1 + sr)$ Fisher information matrix J_α uses the second partial derivatives of the negative of function $\mathcal{L}(\alpha)$ with respect to the elements of α and defined as follows:

$$J_\alpha = \mathbb{E} \left(- \frac{\partial^2 \mathcal{L}}{\partial \alpha \partial \alpha^T} \right)$$

with $\mathbb{E}(\cdot)$ the usual statistical expectation operator (see the appendix for details). Taking the second partial derivatives of the negative of function $\mathcal{L}(\alpha)$ and setting

$$\mathbb{E}(y_{1jk}) = \frac{n_k \beta_{jk}}{1 + \beta_0 \langle z_k, \beta_k \rangle}, \quad \mathbb{E}(y_{2jk}) = \frac{n_k \beta_{jk} z_{jk} \beta_0}{1 + \beta_0 \langle z_k, \beta_k \rangle}, \tag{31}$$

we obtain the general form of the elements of matrix J_α as follows:

$$J_\alpha = \begin{pmatrix} \tau & U_{\alpha,1}^\top & U_{\alpha,2}^\top & \dots & U_{\alpha,s}^\top \\ U_{\alpha,1} & B_{\alpha,1} & 0 & \dots & 0 \\ U_{\alpha,2} & 0 & B_{\alpha,2} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ U_{\alpha,s} & 0 & \dots & 0 & B_{\alpha,s} \end{pmatrix}. \tag{32}$$

4.3. Computing the inverse of the information matrix and interpretation

Without loss of generality, we suppose that $s = 1$ and $r = 2$ i.e., the number of experimental site is equal to one and there are two accident types with n_k the given total number of accidents on the site and $z_k = (z_{1k}, z_{2k})^\top$ the given data from the control area. Parameter α is defined by $\alpha = (\beta_0, \beta_k)^\top \in \mathbb{R}^3$ and $\beta_k = (\beta_{1k}, \beta_{2k})^\top \in \mathbb{R}^2$ with $\beta_0 > 0$, $\beta_{jk} > 0$, and $\beta_{1k} + \beta_{2k} = 1$. So the dimension of the parameter space is three and we have one constraint. We show that the 4×4 constrained or bordered Fisher information matrix is

$$\Gamma_\alpha = \begin{pmatrix} \tau_\alpha & \frac{\gamma_k}{\beta_0} v_{1k} & \frac{\gamma_k}{\beta_0} v_{2k} & 0 \\ \frac{\gamma_k}{\beta_0} v_{1k} & b_{11}^{(k)} & b_{12}^{(k)} & 1 \\ \frac{\gamma_k}{\beta_0} v_{2k} & b_{21}^{(k)} & b_{22}^{(k)} & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix},$$

where $\tau_\alpha = \gamma_k^2 \langle z_k, \beta_k \rangle / \beta_0 n_k$, $v_{mk} = \gamma_k \beta_0 z_{mk} / n_k$, $\gamma_k = n_k / (1 + \beta_0 \langle z_k, \beta_k \rangle)$ and

$$\begin{aligned} B_\alpha &= B_{\alpha,k} \\ &= \begin{pmatrix} b_{11}^{(k)} & b_{12}^{(k)} \\ b_{21}^{(k)} & b_{22}^{(k)} \end{pmatrix} \\ &= \gamma_k \begin{pmatrix} 1 + \beta_0 z_{1k} - \frac{\beta_0^2 z_{1k}^2 \gamma_k^2}{n_k^2} & -\frac{\beta_0^2 z_{1k}^2 z_{2k}^2 \gamma_k^2}{n_k^2} \\ -\frac{\beta_0^2 z_{2k}^2 z_{1k}^2 \gamma_k^2}{n_k^2} & 1 + \beta_0 z_{2k} - \frac{\beta_0^2 z_{2k}^2 \gamma_k^2}{n_k^2} \end{pmatrix}. \end{aligned} \tag{33}$$

In order to reach the variances (accuracy) and covariances (linear relations) of the components of sub-parameter

$$\beta_k = \begin{pmatrix} \beta_{1k} \\ \beta_{2k} \end{pmatrix} \in \mathbb{R}^2, \quad \beta_{1k} + \beta_{2k} = 1,$$

we must get the elements of the 4×4 symmetric matrix

$$\Gamma_x^{-1} = \begin{pmatrix} \Gamma_x^{-1}(1, 1) & \Gamma_x^{-1}(1, 2) & \Gamma_x^{-1}(1, 3) & \Gamma_x^{-1}(1, 4) \\ & \Gamma_x^{-1}(2, 2) & \Gamma_x^{-1}(2, 3) & \Gamma_x^{-1}(2, 4) \\ & & \Gamma_x^{-1}(3, 3) & \Gamma_x^{-1}(3, 4) \\ & & & \Gamma_x^{-1}(4, 4) \end{pmatrix}$$

and thereafter we take the 2×2 element-block

$$\text{cov}(\beta_k, \beta_k) = \begin{pmatrix} \Gamma_x^{-1}(2, 2) & \Gamma_x^{-1}(2, 3) \\ \Gamma_x^{-1}(2, 3) & \Gamma_x^{-1}(3, 3) \end{pmatrix},$$

where $\Gamma_x^{-1}(2, 2)$, $\Gamma_x^{-1}(3, 3)$ and $\Gamma_x^{-1}(2, 3)$ are, respectively, interpreted as the precisions related to the estimation of parameters β_{1k} , β_{2k} and the degree of linear relation between β_{1k} and β_{2k} .

The method proposed here allows us to avoid doing this inversion and gives the analytical expressions of the elements of matrix $\text{cov}(\beta_k, \beta_k)$ by using Theorem 3.7 as follows:

$$J_x = \begin{pmatrix} \tau_x & \frac{\gamma_k}{\beta_0} V_{x,k}^T \\ \frac{\gamma_k}{\beta_0} V_{x,k} & B_{x,k} \end{pmatrix}, \quad \Omega_{x,k} = \begin{pmatrix} \frac{1 + \beta_0 z_{1k}}{\beta_{1k}} & 0 \\ 0 & \frac{1 + \beta_0 z_{2k}}{\beta_{2k}} \end{pmatrix}$$

and $B_{x,k} = \gamma_k [\Omega_{x,k} - V_{x,k} V_{x,k}^T]$ a 2×2 matrix and

$$V_{x,k} = \frac{\beta_0 \gamma_k}{n_k} \begin{pmatrix} z_{1k} \\ z_{2k} \end{pmatrix}, \quad H_2^T = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

We show that the Schur complement of B_x in J_x is given by

$$(J_x/B_x) = \frac{\gamma_k^2}{n_k} \left(\frac{\langle z_k, \beta_k \rangle}{\beta_0} - \frac{\gamma_k t_k}{n_k} \|z_k\|_{\Omega_{x,k}^{-1}}^2 \right), \quad \|z_k\|_{\Omega_{x,k}^{-1}}^2 = \omega_{1k} z_{1k}^2 + \omega_{2k} z_{2k}^2$$

with

$$\omega_{jk} = \frac{\beta_{jk}}{1 + \beta_0 z_{jk}}, \quad j = 1, 2$$

the elements of the 2×2 diagonal matrix $\Omega_{x,k}^{-1}$ and

$$\|\xi_x\|_{A_x^{-1}}^2 = \frac{\gamma_k^2 t_k^2 \Delta_{x,k}^2}{n_k^2 \|\mathbf{1}_r\|_{B_{x,k}^{-1}}^2} \quad \text{with } \Delta_{x,k} = \omega_{1k} z_{1k} + \omega_{2k} z_{2k}.$$

So using Theorems 3.6 and 3.7 we get

$$\text{cov}(\beta_k, \beta_k) = J_{k,k} - A_{k,k} J_{k,k} \mathbf{1}_{2,2} J_{k,k}, \tag{34}$$

342 *A. N'Guessan, C. Langrand / Journal of Computational and Applied Mathematics 177 (2005) 331–345*

a 2×2 matrix, where

$$I_{2,2} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad J_{k,k} = \gamma_k^{-1} \begin{bmatrix} \omega_{1k} + \theta_k \omega_{1k}^2 z_{1k}^2 & \theta_k \omega_{1k} \omega_{2k} \\ \theta_k \omega_{1k} \omega_{2k} & \omega_{2k} + \theta_k \omega_{2k}^2 z_{2k}^2 \end{bmatrix}$$

with

$$\theta_k = \frac{t_k \gamma_k^2}{n_k^2} [\beta_0^2 - (J_x/B_x)^{-1} t_k \gamma_k] \quad \text{and} \quad A_{k,k} = \frac{1}{\|\mathbf{1}_r\|_{B_{x,k}}^{-1}} - \frac{\sigma_x^2 t_k^2 \gamma_k^2 \Delta_{x,k}^2}{n_k^2 \|\mathbf{1}_r\|_{B_{x,k}}^{-4}}$$

with

$$\sigma_x^2 = [(J_x/B_x) + \|\xi_x\|_{\Lambda_x^{-1}}^2]^{-1}.$$

Thereafter, we deduce the explicit expressions of the variances and the covariances of subparameter β_k as follows:

$$\Gamma_x^{-1}(j+1, j+1) = \sigma^2(\beta_{jk}) = \gamma_k^{-1} [\omega_{jk} + \omega_{jk}^2 (\theta_k z_{jk}^2 - A_{k,k} \gamma_k^{-1} x_{jk}^2)], \quad j = 1, 2,$$

$$\Gamma_x^{-1}(2, 3) = \text{cov}(\beta_{1k}; \beta_{2k}) = \gamma_k^{-1} \omega_{1k} \omega_{2k} [\theta_k z_{1k} z_{2k} - A_{k,k} \gamma_k^{-1} x_{1k} x_{2k}],$$

where $x_{jk} = 1 + \theta_k \Delta_{x,k} z_{jk}$. The main strength of our approach is its applicability to any dimension of matrix Γ_x without having to invert it. It saves considerable time and enables great precision in calculations and is accessible to everyone. We do not need any numerical inversion program to succeed in doing it, only matrices manipulations (product and addition) are called for.

Acknowledgements

We would like to thank Prof. Claude Brezinski for turning our attention to the paper on Schur complement and we are also deeply grateful to the anonymous referee for his judicious notes and comments which have enabled a complete remodelling of our paper. Part of this work was set up while the first author was a guest researcher at Montreal University CRT (Research Center on Transportation), the technical and material helps of which we would also like to thank.

Appendix A. Computing of the Fisher information matrix

Differentiating $\mathcal{L}(\alpha)$ with respect to each of the $(1 + sr)$ components of α , we obtain the first partial derivatives as follows:

$$\frac{\partial \mathcal{L}(\alpha)}{\partial \beta_0} = \sum_{k=1}^s \sum_{m=1}^r \left\{ \frac{y_{2mk} - y_{1mk} \beta_0 \langle z_k, \beta_k \rangle}{\beta_0 (1 + \beta_0 \langle z_k, \beta_k \rangle)} \right\},$$

$$\frac{\partial \mathcal{L}(\alpha)}{\partial \beta_{mk}} = \frac{y_{mk}}{\beta_{mk}} - \frac{y_{mk} \beta_0 z_{mk}}{1 + \beta_0 \langle z_k, \beta_k \rangle} \quad (k = 1, 2, \dots, s; \quad m = 1, 2, \dots, r),$$

where $y_{.mk} = y_{1mk} + y_{2mk}$ and $y_{.k} = \sum_{j=1}^r y_{.jk}$. So we deduce the second partial derivatives as follows:

$$\frac{\partial^2 \mathcal{L}(\alpha)}{\partial \beta_0 \partial \beta_0} = \sum_{k=1}^s \left\{ -\frac{y_{2.k}}{\beta_0^2} + \frac{y_{.k} (\langle z_k, \beta_k \rangle)^2}{(1 + \beta_0 \langle z_k, \beta_k \rangle)^2} \right\},$$

$$\frac{\partial^2 \mathcal{L}(\alpha)}{\partial \beta_0 \partial \beta_{mk}} = -\frac{y_{.k} z_{mk}}{(1 + \beta_0 \langle z_k, \beta_k \rangle)^2} \quad k = 1, 2, \dots, s, \quad m = 1, 2, \dots, r,$$

$$\frac{\partial^2 \mathcal{L}(\alpha)}{\partial \beta_{jp} \partial \beta_{mk}} = \begin{cases} 0 & \text{if } k \neq p, \\ \frac{y_{.k} z_{mk} z_{jk} \beta_0^2}{(1 + \beta_0 \langle z_k, \beta_k \rangle)^2} & \text{if } k = p \text{ and } j \neq m, \\ -\frac{y_{mk}}{\beta_{mk}^2} + \frac{y_{.k} \beta_0^2 z_{mk}^2}{(1 + \beta_0 \langle z_k, \beta_k \rangle)^2} & \text{if } k = p \text{ and } j = m \end{cases}$$

$k, p = 1, 2, \dots, s; m, j = 1, 2, \dots, r$

and thereafter the general structure of the $(1 + sr) \times (1 + sr)$ matrix J_α is given by

$$J_\alpha = \begin{pmatrix} J_{\beta_0, \beta_0} & J_{\beta_0, \beta_1}^T & J_{\beta_0, \beta_2}^T & \cdots & J_{\beta_0, \beta_s}^T \\ J_{\beta_0, \beta_1} & J_{\beta_1, \beta_1} & \mathbf{0} & \cdots & \mathbf{0} \\ J_{\beta_0, \beta_2} & \mathbf{0} & J_{\beta_2, \beta_2} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \mathbf{0} \\ J_{\beta_0, \beta_s} & \mathbf{0} & \cdots & \mathbf{0} & J_{\beta_s, \beta_s} \end{pmatrix}.$$

where J_{β_0, β_0} is a scalar given by

$$J_{\beta_0, \beta_0} = \mathbb{E} \left(-\frac{\partial^2 \mathcal{L}(\alpha)}{\partial \beta_0 \partial \beta_0} \right) = \sum_{k=1}^s \left\{ \frac{1}{\beta_0^2} \mathbb{E}(y_{2.k}) - \frac{(\langle z_k, \beta_k \rangle)^2}{(1 + \beta_0 \langle z_k, \beta_k \rangle)^2} \mathbb{E}(y_{.k}) \right\}.$$

J_{β_0, β_k} is a $r \times 1$ vector given by

$$J_{\beta_0, \beta_k} = \left[\mathbb{E} \left(-\frac{\partial^2 \mathcal{L}(\alpha)}{\partial \beta_0 \partial \beta_{1k}} \right), \mathbb{E} \left(-\frac{\partial^2 \mathcal{L}(\alpha)}{\partial \beta_0 \partial \beta_{2k}} \right), \dots, \mathbb{E} \left(-\frac{\partial^2 \mathcal{L}(\alpha)}{\partial \beta_0 \partial \beta_{rk}} \right) \right]^T$$

with the m th element given by

$$\mathbb{E} \left(-\frac{\partial^2 \mathcal{L}(\alpha)}{\partial \beta_0 \partial \beta_{mk}} \right) = \frac{z_{mk}}{(1 + \beta_0 \langle z_k, \beta_k \rangle)^2} \mathbb{E}(y_{.k})$$

and J_{β_k, β_k} a $r \times r$ matrix with (j, m) th elements

$$\mathbb{E} \left(-\frac{\partial^2 \mathcal{L}(\alpha)}{\partial \beta_{jk} \partial \beta_{mk}} \right) = \begin{cases} -\frac{z_{mk} z_{jk} \beta_0^2}{(1 + \beta_0 \langle z_k, \beta_k \rangle)^2} \mathbb{E}(y_{.k}) & \text{if } j \neq m, \\ \frac{1}{\beta_{mk}^2} \mathbb{E}(y_{.mk}) - \frac{\beta_0^2 z_{mk}^2}{(1 + \beta_0 \langle z_k, \beta_k \rangle)^2} \mathbb{E}(y_{.k}) & \text{if } j = m. \end{cases}$$

344 A. N'Guessan, C. Langrand / *Journal of Computational and Applied Mathematics* 177 (2005) 331–345

Now setting

$$\mathbb{E}(y_{1jk}) = \frac{n_k \beta_{jk}}{1 + \beta_0 \langle z_k, \beta_k \rangle}, \quad \mathbb{E}(y_{2jk}) = \frac{n_k \beta_{jk} z_{jk} \beta_0}{1 + \beta_0 \langle z_k, \beta_k \rangle}$$

for $k = 1, 2, \dots, s$ and $m = 1, 2, \dots, r$, we get $J_{\beta_0, \beta_0} = \tau$, $J_{\beta_0, \beta_k} = U_{x,k}$ and $J_{\beta_k, \beta_k} = B_{x,k}$.

References

- [1] J. Aitchison, S.D. Silvey, Maximum likelihood estimation of parameters subject to restraints, *Ann. Math. Statist.* 29 (1958) 813–829.
- [2] C. Brezinski, M.R. Zaglia, A Schur complement approach to a general extrapolation algorithm, *Linear Algebra Appl.* 368 (2003) 279–301.
- [3] N.R. Cook, Restricted maximum likelihood, in: P. Armitage, T. Colton (Eds.), *Encyclopedia of Biometrics*, vol. 5, Wiley, 1998, pp. 3827–3830.
- [4] H. Cramer, *Mathematical Methods of Statistics*, Princeton University Press, Princeton, 1946.
- [5] M. Crowder, On the constrained maximum likelihood estimation with non i.i.d. observations, *Ann. Inst. Statist. Math.* 36A (1984) 239–249.
- [6] F.J.H. Don, The use of generalized inverses in restricted maximum likelihood, *Linear Algebra Appl.* 70 (1985) 225–240.
- [7] L.A. Escobar, W.Q. Meeker, Fisher information matrices with censoring, truncation, and explanatory variables, *Statist. Sinica* 8 (1998) 221–237.
- [8] Y. Fan, Schur complements and its applications to symmetric nonnegative and Z-matrices, *Linear Algebra Appl.* 353 (2002) 289–307.
- [9] A. Klein, G. Mélard, Computation of the Fisher information matrix for times series models, *J. Comput. Appl. Math.* 64 (1995) 57–68.
- [10] J.R. Magnus, *Linear Structures*, Charles Griffin and Company Ltd., London, 1988.
- [11] G.B. Matthews, N.A.S. Crowther, A maximum likelihood estimation procedure when modelling in terms of constraints, *South African Statist. J.* 29 (1995) 29–50.
- [12] B.E. Neuenschwander, B.D. Flury, A note on Silvey's (1959) Theorem, *Statist. Probab. Lett.* 36 (1997) 307–317.
- [13] A. N'Guessan, On a use of Schur complements for an inverse of a constrained accident data information matrix, vol. 60, no. VI, Pub. IRMA, Lille, 2003.
- [14] A. N'Guessan, Constrained covariance matrix estimation in road accident modelling with Shur complements, *C.R. Acad. Sci. Paris Ser. I* 337 (2003) 219–222.
- [15] A. N'Guessan, Constrained Estimation of a Road Safety covariance matrix using Shur complements, Technical Report—Centre for Research on Transportation, no. CRT-2003-12, Université de Montréal, 2003, 18pp.
- [16] A. N'Guessan, A. Essai, C. Langrand, Estimation multidimensionnelle des contrôles et de l'effet moyen d'une mesure de sécurité routière, *Rev. Statist. Appl.* XLIX (2) (2001) 83–100.
- [17] R.J. Ober, The Fisher information matrix for linear systems, *Systems Control Lett.* 47 (2002) 221–226.
- [18] R.J. Ober, Q. Zou, Z. Lin, Calculation of the Fisher information matrix for multidimensional data sets, *IEEE Trans. Signal Process.* 51 (10) (2003) 2679–2691.
- [19] D.V. Ouellette, Schur complements and statistics, *Linear Algebra Appl.* 36 (1981) 187–295.
- [20] R.L.M. Peeters, B. Hanzon, Symbolic Computation of Fisher information matrices for parametrized state-space systems, *Automatica* 35 (1999) 1059–1071.
- [21] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.F. Flannery, *Numerical Recipes in C The Art of Scientific Computing*, second ed., Cambridge University Press, Cambridge, 1992.
- [22] C.R. Rao, *Linear Statistical Inference and its Applications*, second ed., Wiley, New York, 1973.
- [23] C.R. Rao, H. Yanai, Generalized inverses of partitioned matrices useful in statistical applications, *Linear Algebra Appl.* 70 (1985) 105–115.
- [24] M. Sagae, K. Tanabe, Symbolic Cholesky decomposition of the variance–covariance matrix of the negative multinomial distribution, *Statist. Probab. Lett.* 15 (1992) 103–108.
- [25] S.D. Silvey, The Lagrangian multiplier test, *Ann. Math. Statist.* 30 (1959) 389–407.

A. N'Guessan, C. Langrand / Journal of Computational and Applied Mathematics 177 (2005) 331–345 345

- [26] G.P.H. Styan, in: T. Pukkila, S. Puntanen (Eds.), *Schur Complements and Linear Statistical Models*, 1985, pp. 37–75.
- [27] K. Tanabe, M. Soga, An exact Cholesky decomposition and the generalized inverse of the variance–covariance matrix of the multinomial distribution, with applications, *J. Roy. Statist. Soc. B* 54 (1) (1992) 211–219.
- [28] C. Van Eeden, Estimation in restricted parameter spaces—some history and some recent developments, *CWI Quart.* 9 (1,2) (1996) 69–76.
- [29] M.A. Woodbury, *Inverting modified matrices*, Memorandum Report 42, Statistical Research Group, Princeton, NJ, 1950.

Annexe A04 :

C. R. Math. Rep. Acad. Sci. Canada, Vol. 27, (3), 84-91 (2005)

A Confidence Interval Estimation problem using the Schur complement approach, with application

C. R. Math. Rep. Acad. Sci. Canada Vol. 27, (3), 2005 pp. 84–91

A CONFIDENCE INTERVAL ESTIMATION PROBLEM USING THE SCHUR COMPLEMENT APPROACH, WITH APPLICATION

ASSI N'GUESSAN AND FRANÇOIS BELLAVANCE

Presented by Vlastimil Dlab, FRSC

ABSTRACT. We present a method based on the Schur complement approach to build asymptotic confidence intervals linked to the maximum likelihood estimator of a vector of parameters under constraints. This approach makes it possible to obtain the formal expression of the standard error of each component of the vector without direct inversion of the Fisher information matrix. We then give an application of this method to the modelling and the confidence interval estimation of the average effect of a road safety measure and the accident risks of different types.

RÉSUMÉ. Nous proposons une méthode basée sur la technique du complément de Schur pour construire des intervalles de confiance asymptotiques relatifs à l'estimateur du maximum de vraisemblance d'un vecteur paramètre soumis à des contraintes. Cette méthode permet d'obtenir l'expression formelle de l'écart-type de chaque composante du vecteur sans inverser directement la matrice d'information de Fisher. Nous indiquons ensuite une application de cette méthodologie à la modélisation et à l'estimation par intervalle de confiance de l'effet moyen d'une mesure de sécurité routière et de différents risques d'accident.

1. Introduction and assumptions. Let's consider $\{(X_{1k}, X_{2k}), 1 \leq k \leq s\}$, a collection of s pairs of discreet independent random vectors where $X_{1k} = (X_{11k}, X_{12k}, \dots, X_{1rk})^T$ and $X_{2k} = (X_{21k}, X_{22k}, \dots, X_{2rk})^T$ are $r \times 1$ random vectors with $r > 1$. We assume that, for fixed k , each pair (X_{1k}, X_{2k}) has a multinomial distribution denoted by $\mathcal{M}(n_k; \Pi_{1k}(\theta), \Pi_{2k}(\theta))$, where n_k ($n_k > 0$) is a given integer, $\Pi_{tk}(\theta) = (\pi_{t1k}(\theta), \pi_{t2k}(\theta), \dots, \pi_{trk}(\theta))$ $t = 1, 2$, is a $r \times 1$ vector of cell probabilities, $\theta = (\gamma, \beta^T)^T$ an unknown vector of parameters whose dimension is $(1 + sr) \times 1$, γ being a real number, $\beta = (\beta_1^T, \beta_2^T, \dots, \beta_s^T)^T$ being a vector whose dimension is $(sr) \times 1$, and $\beta_k = (\beta_{1k}, \beta_{2k}, \dots, \beta_{rk})^T \in \mathbb{R}^r$ with $0 < \beta_{jk} < 1$. We propose an estimation method for the asymptotic confidence interval of each element of θ . This estimation method is based on the formal inversion of the Fisher information matrix using Schur complement (see Ouellette [11], Zhang [16]). There are certainly other methods (see for instance Don [3], Rao and Yanai [12], Tanabe and Sagae [14]) to invert the Fisher information matrix but the one we suggest is based on the Schur complement approach and

Received by the editors on March 14, 2005.

AMS subject classification: 15A09, 62F12, 62F25, 62F30, 62H10, 62H12, 62P30.

Keywords: logistic multinomial model, constrained maximum likelihood, confidence interval, Fisher information matrix, formal asymptotic variance, Schur complement, road safety measure, accident risk, accident data.

© Royal Society of Canada 2005.

A CONFIDENCE INTERVAL ESTIMATION PROBLEM

85

generalizes N'Guessan's results [7].

Let $\mathcal{L}(\theta) = \sum_{k=1}^s \mathcal{L}_k(\theta)$ the log-likelihood linked to the s pairs with $\mathcal{L}_k(\theta)$, the one associated to pair k , and we assume for the rest of the work that the following conditions are verified:

(A1) $\forall k, \mathcal{L}_k(\theta)$ is continuously differentiable such that

$$\frac{\partial \mathcal{L}_k(\theta)}{\partial \beta_m} \equiv 0_r, \quad k \neq m, \quad k, m = 1, 2, \dots, s$$

where 0_r is a $r \times 1$ vector of zeros;

(A2) Fisher information matrix $J_\theta = \mathbb{E}(-\frac{\partial^2 \mathcal{L}(\theta)}{\partial \theta \partial \theta^T})$ is thus partitioned:

$$J_\theta = \begin{bmatrix} \tau_\theta & U_\theta^T \\ U_\theta & B_\theta \end{bmatrix}$$

where τ_θ is a real number, U_θ is a $(sr) \times 1$ vector and B_θ a $(sr) \times (sr)$ matrix, respectively defined by

$$\tau_\theta = \mathbb{E}\left(-\frac{\partial^2 \mathcal{L}(\theta)}{\partial \gamma \partial \gamma}\right), \quad U_\theta = (U_{\theta,1}^T, U_{\theta,2}^T, \dots, U_{\theta,s}^T)^T, \\ B_\theta = \text{block-diag}(B_{\theta,1}, B_{\theta,2}, \dots, B_{\theta,s})$$

with, for a fixed value of k , $U_{\theta,k}$ a $r \times 1$ vector of components $\mathbb{E}(-\frac{\partial^2 \mathcal{L}(\theta)}{\partial \gamma \partial \beta_{j,k}})$ and $B_{\theta,k}$, an $r \times r$ matrix with elements $\mathbb{E}(-\frac{\partial^2 \mathcal{L}(\theta)}{\partial \beta_{j,k} \partial \beta_{m,k}})$, ($j = 1, 2, \dots, r$; $m = 1, 2, \dots, r$);

(A3) $\forall k$, there exists $\Omega_{\theta,k}$, a nonsingular $r \times r$ diagonal matrix, $V_{\theta,k}$, a $r \times 1$ vector, a_k and b_k two strictly positive real numbers such that

(i)

$$B_{\theta,k} = a_k(\Omega_{\theta,k} - V_{\theta,k}V_{\theta,k}^T),$$

(ii)

$$U_{\theta,k} = b_k V_{\theta,k};$$

(A4) $\forall k (k = 1, 2, \dots, s)$, we set

$$\Omega_{\theta,k}^{(0)} = \begin{bmatrix} \Omega_{\theta,k} & -V_{\theta,k} \\ V_{\theta,k}^T & 0 \end{bmatrix}, \quad B_{\theta,k}^{(0)} = \begin{bmatrix} B_{\theta,k} & -1_r \\ 1_r^T & 0 \end{bmatrix}$$

two $(r + 1) \times (r + 1)$ matrices and we assume that

(i)

$$0 < (\Omega_{\theta,k}^{(0)} / \Omega_{\theta,k}) < 1,$$

86

ASSI N'GUESSAN AND FRANÇOIS BELLAVANCE

(ii)

$$0 < \sum_{k=1}^s \frac{b_k^2}{a_k} \frac{(\Omega_{\theta,k}^{(0)}/\Omega_{\theta,k})}{1 - (\Omega_{\theta,k}^{(0)}/\Omega_{\theta,k})} < \tau_{\theta},$$

(iii)

$$0 < (B_{\theta,k}^{(0)}/B_{\theta,k}),$$

where $(\Omega_{\theta,k}^{(0)}/\Omega_{\theta,k})$ (resp. $(B_{\theta,k}^{(0)}/B_{\theta,k})$) is the Schur complement of $\Omega_{\theta,k}$ in $\Omega_{\theta,k}^{(0)}$ (resp. the Schur complement of $B_{\theta,k}$ in $B_{\theta,k}^{(0)}$);

(A5) $\forall k$, there exists $h_k: \mathbb{R}^{1+sr} \mapsto \mathbb{R}$, $\theta \mapsto h_k(\theta)$ a continuously differentiable function such that

(i)

$$h_k(\theta) = 0,$$

(ii)

$$\frac{\partial h_k}{\partial \gamma} \equiv \mathbf{0}_s,$$

(iii)

$$\frac{\partial h_k}{\partial \beta_m} \equiv \delta_{km} \mathbf{1}_r,$$

$k, m = 1, 2, \dots, s$ with δ_{km} being the Kronecker symbol, $\mathbf{0}_s = (0, \dots, 0)^T \in \mathbb{R}^s$, and $\mathbf{1}_r = (1, \dots, 1)^T \in \mathbb{R}^r$.

2. Preliminary results using Schur complement.

THEOREM 2.1. Under (A1) to (A4), we show that

(i) $(J_{\theta}/B_{\theta}) > 0$,(ii) $(J_{\theta}/\tau_{\theta})^{-1} = B_{\theta}^{-1} + (J_{\theta}/B_{\theta})^{-1} B_{\theta}^{-1} U_{\theta} U_{\theta}^T B_{\theta}^{-1}$,

where (J_{θ}/B_{θ}) (resp. $(J_{\theta}/\tau_{\theta})$) is the Schur complement of B_{θ} (resp. of τ_{θ}) in J_{θ} and is defined by $(J_{\theta}/B_{\theta}) = \tau_{\theta} - U_{\theta}^T B_{\theta}^{-1} U_{\theta}$ (resp. $(J_{\theta}/\tau_{\theta}) = B_{\theta} - U_{\theta} \tau_{\theta}^{-1} U_{\theta}^T$).

THEOREM 2.2. Let $H_{\theta} = (H_{\gamma}, H_{\beta})$ the $s \times (1 + sr)$ matrix with $H_{\gamma} = (\frac{\partial h_1}{\partial \gamma}, \frac{\partial h_2}{\partial \gamma}, \dots, \frac{\partial h_s}{\partial \gamma})^T$, $H_{\beta} = (\frac{\partial h_1}{\partial \beta}, \frac{\partial h_2}{\partial \beta}, \dots, \frac{\partial h_s}{\partial \beta})^T$, being respectively $s \times 1$ and $s \times (sr)$ matrices, $\xi_{\theta} = H_{\beta} B_{\theta}^{-1} U_{\theta}$ being an $s \times 1$ vector,

$$\Lambda_{\theta} = \text{diag}((B_{\theta,1}^{(0)}/B_{\theta,1}), \dots, (B_{\theta,s}^{(0)}/B_{\theta,s}))$$

an $s \times s$ matrix,

$$\Lambda_{\theta}^{(1)} = \begin{bmatrix} \Lambda_{\theta} & -\xi_{\theta} \\ \xi_{\theta}^T & (J_{\theta}/B_{\theta}) \end{bmatrix}, \quad \Gamma_{\theta} = \begin{bmatrix} J_{\theta} & H_{\theta}^T \\ H_{\theta} & \mathbf{O}_{s,s} \end{bmatrix}$$

A CONFIDENCE INTERVAL ESTIMATION PROBLEM

87

being respectively $(s + 1) \times (s + 1)$ and $(1 + sr + s) \times (1 + sr + s)$ matrices with $\bigcirc_{s,s}$ the $s \times s$ matrix with all entries equal to zero. Then, under conditions (A1) to (A5), we show that

- (i) $R_\theta = -(\Gamma_\theta/J_\theta)$, the negative of the Schur complement of J_θ in Γ_θ , is a nonsingular $s \times s$ matrix and $R_\theta^{-1} = \Lambda_\theta^{-1} - \Lambda_\theta^{-1}\xi_\theta(\Lambda_\theta^{(1)}/\Lambda_\theta)^{-1}\xi_\theta^T\Lambda_\theta^{-1}$.
- (ii) Matrix Γ_θ is nonsingular and

$$(1) \quad \Gamma_\theta^{-1} = \begin{bmatrix} W_\theta & J_\theta^{-1}H_\theta^T R_\theta^{-1} \\ R_\theta^{-1}H_\theta J_\theta^{-1} & -R_\theta^{-1} \end{bmatrix}, \quad W_\theta = \begin{bmatrix} W_\theta(1,1) & W_\theta^T(2,1) \\ & W_\theta(2,2) \end{bmatrix}$$

where the diagonal components $W_\theta(1,1)$ and $W_\theta(2,2)$ of matrix W_θ are respectively a scalar and an $(sr) \times (sr)$ matrix defined as follows:

$$(2) \quad \begin{aligned} W_\theta(1,1) &= (\Lambda_\theta^{(1)}/\Lambda_\theta)^{-1} \\ W_\theta(2,2) &= (J_\theta/\tau_\theta)^{-1} - (J_\theta/\tau_\theta)^{-1}H_\beta^T R_\beta^{-1}H_\beta(J_\theta/\tau_\theta)^{-1} \\ W_\theta(2,1) &= -(J_\theta/B_\theta)^{-1}[B_\theta^{-1}U_\theta - (J_\theta/\tau_\theta)^{-1}H_\beta^T R_\beta^{-1}\xi_\theta]. \end{aligned}$$

REMARK 1. We give an outline of the proofs of Theorems 2.1 and 2.2.

From matrix manipulations and (A3) (i) and (A4) (i) we obtain

$$(3) \quad (J_\theta/B_\theta) = \tau_\theta - \sum_{k=1}^s \frac{b_k^2}{a_k} \frac{(\Omega_{\theta,k}^{(0)}/\Omega_{\theta,k})}{1 - (\Omega_{\theta,k}^{(0)}/\Omega_{\theta,k})}$$

Now using (A4) (i) and (A3) (ii), we get part (i) of Theorem 2.1. The rest of the proof uses theorems in N'Guessan [7, Section 2] and N'Guessan and Langrand [10, Section 3] and general applications of well-known results of Schur complement about the inverse of a partitioned matrix.

3. Asymptotic confidence interval.

THEOREM 3.1. Let $M_\theta = J_\theta^{-1}(J_\theta + (\frac{\partial^2 \mathcal{L}(\theta)}{\partial \theta \partial \theta^T}))$ a $(1 + sr) \times (1 + sr)$ matrix, $\Phi_\theta = J_\theta^{-1}H_\theta^T R_\theta^{-1}$ a $(1 + sr) \times s$ matrix, λ the $s \times 1$ vector of the Lagrange multipliers, θ^0 the true value of the vector of parameters θ , and we suppose that, in addition to conditions (A1) to (A5),

- (A6) $(\hat{\theta}, \hat{\lambda})$, the maximum likelihood estimators of θ and λ , exist and converge.
- (A7) $M_{\hat{\theta}} \xrightarrow{\text{Proba}} 0$, $\Phi_{\hat{\theta}} - \Phi_{\theta^0} \xrightarrow{\text{Proba}} 0$, $|\Phi_{\theta^0}| < \infty$ when n_k and $N = \sum_{k=1}^s n_k \rightarrow +\infty$.
- (A8) $(\frac{\partial \mathcal{L}}{\partial \theta})_{\theta^0} \stackrel{\mathcal{D}}{\sim} \mathcal{N}(0, J_{\theta^0})$.

Then the asymptotic confidence intervals, at the $(1 - \alpha)$ confidence level, for the θ parameter components are given by the following formal expressions:

$$(4) \quad \begin{aligned} IC_{1-\alpha}(\gamma) &= \left[\hat{\gamma} - \frac{z_{\alpha/2}}{(\Lambda_{\theta}^{(1)}/\Lambda_{\theta})^{1/2}} ; \hat{\gamma} + \frac{z_{\alpha/2}}{(\Lambda_{\theta}^{(1)}/\Lambda_{\theta})^{1/2}} \right] \\ IC_{1-\alpha}(\beta_{jk}) &= [\max(0; \hat{\beta}_{jk} - z_{\alpha/2}\hat{\sigma}_{jk}) ; \min(1; \hat{\beta}_{jk} + z_{\alpha/2}\hat{\sigma}_{jk})] \\ &\quad (k = 1, 2, \dots, s; j = 1, 2, \dots, r) \end{aligned}$$

where $z_{\alpha/2}$ and $\hat{\sigma}_{jk}^2$ are respectively obtained through the standard normal law table and the elements of the main diagonal of the $(sr) \times (sr)$ matrix $(J_{\theta}/\tau_{\theta})^{-1} - (J_{\theta}/\tau_{\theta})^{-1}H_{\beta}^T R_{\theta}^{-1}H_{\beta}(J_{\theta}/\tau_{\theta})^{-1}$.

REMARK 2. The main point of the proof consists in applying the mean value theorem, in the neighbourhood of θ^0 , to the constrained maximum likelihood equations and uses a version of a theorem in Crowder [2, Section 3] to show that asymptotically

$$(5) \quad \begin{pmatrix} \hat{\theta} - \theta^0 \\ \eta_{\hat{\theta}} \end{pmatrix} \underset{D)}{\underset{\sim}{\mathcal{N}}} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} ; \begin{pmatrix} W_{\theta^0} & 0 \\ 0 & R_{\theta^0} \end{pmatrix} \right)$$

with $\eta_{\hat{\theta}} = R_{\theta}\hat{\lambda}$ being a $s \times 1$ vector and $W_{\theta} = J_{\theta}^{-1} - J_{\theta}^{-1}H_{\theta}^T R_{\theta}^{-1}H_{\theta}J_{\theta}^{-1}$ a $(1 + sr) \times (1 + sr)$ matrix obtained by inverting matrix Γ_{θ} with the Schur complement method. We then use conditions (A1) to (A5) along with Theorems 2.1 and 2.2 to deduce the asymptotic variance structure of the components of $\hat{\theta}$. Part of the method proposed here is clearly related to the constrained maximum likelihood estimation (RMLE) method which is an approach to estimation that maximizes the logarithm of a likelihood over a restricted space. This approach of estimation is the well-known applied mathematical method that maximizes a function subjected to restraints. The general statistical framework is well known and is not discussed here (see for instance Aitchison and Sylvey [1], Sylvey [13], Don [3], Magnus [5, p. 172]). However one recalls that the size of matrix Γ_{θ} increases very quickly with r and s . The main strength of the Schur complement approach is its applicability to any values of s and r , *i.e.*, to any dimension of matrix Γ_{θ} , without being obliged to invert it.

4. Application: Evaluation of a road safety measure.

4.1. *A logistic multinomial model.* We apply the above results to a statistical road accident data modelling problem when a road safety measure is applied on different sites. We consider a multidimensional combination of road accident frequencies before and after the introduction of a road safety measure (crossroad lay-out, surface of a motorway section, *etc.*) at s ($s > 0$) experimental sites. Each site counts r ($r > 1$) accident types (fatal accidents, seriously injured

A CONFIDENCE INTERVAL ESTIMATION PROBLEM

89

people, slightly injured people, material damage only) over a fixed period of time. To model the average effect of the safety measure and the accident risks of different types, let's consider the period before (resp. after) the introduction of the safety measure (or change) and note $X_{1k} = (X_{11k}, X_{12k}, \dots, X_{1rk})^T$ (resp. $X_{2k} = (X_{21k}, X_{22k}, \dots, X_{2rk})^T$) the random vector giving the r accident types number on experimental site k . In order to take some external factors into account (such as traffic flow, speed limit variation, weather conditions, ...), each experimental site is matched to a control site where the safety measure was not applied. We further assume that the vector of accident data for control site k ($k = 1, 2, \dots, s$), over the same time period, is fixed and known and is given by $y_k = (y_{1k}, y_{2k}, \dots, y_{rk})^T$ where y_{jk} denotes the ratio of the number of accidents of type j for the period after to the period before in control site k and uses a variation of a control ratio defined by Tanner [15].

We then assume that $\{(X_{1k}, X_{2k}), 1 \leq k \leq s\}$ is a collection of s independent $2r$ -dimensional random vector and that, for fixed k , each couple (X_{1k}, X_{2k}) has a logistic multinomial distribution (see Hosmer and Lemeshow [4], Mehta and Patel [6]) denoted by $\mathcal{LM}(n_k; \Pi_{1k}(\theta), \Pi_{2k}(\theta))$, where n_k ($n_k > 0$) is the observed total accident number in experimental site k ,

$$\Pi_{tk}(\theta) = (\pi_{t1k}(\theta), \pi_{t2k}(\theta), \dots, \pi_{trk}(\theta))$$

is the $r \times 1$ vector of cell probabilities with component $\pi_{tjk} : \mathbb{R}^{1+sr} \mapsto]0; 1[$ given by a logistic link as follows:

$$(6) \quad \begin{aligned} \pi_{1jk}(\theta) &= \frac{\beta_{jk}}{1 + e^\gamma \sum_{m=1}^r y_{mk} \beta_{mk}}; \\ \pi_{2jk}(\theta) &= \frac{y_{jk} \beta_{jk} e^\gamma}{1 + e^\gamma \sum_{m=1}^r y_{mk} \beta_{mk}}, \quad (j = 1, 2, \dots, r), \end{aligned}$$

and where $\theta = (\gamma, \beta^T)^T$ is the $(1 + sr) \times 1$ unknown vector of parameters with γ the logarithm of the average effect, $\beta = (\beta_1^T, \beta_2^T, \dots, \beta_s^T)^T$ with $\beta_k = (\beta_{1k}, \beta_{2k}, \dots, \beta_{rk})^T$ a $r \times 1$ vector of accident risks connected to the control site k such that

$$(7) \quad \beta_k \in \mathcal{S}_k^{(r-1)} = \left\{ (\beta_{1k}, \beta_{2k}, \dots, \beta_{rk})^T \in \mathbb{R}^r, \beta_{jk} > 0, \sum_{j=1}^r \beta_{jk} = 1 \right\}.$$

The logistic link between $\pi_{tjk}(\theta)$ and vector θ comes from the combination of the multinomial model built by N'Guessan *et al.* [9] and used by N'Guessan [8] and the functional $g : \mathbb{R}^{1+sr} \mapsto \mathbb{R}^{1+sr}$, $\theta \mapsto g(\theta) = (e^\gamma, \beta^T)^T$ which leaves the vector β invariant and transforms γ in e^γ . Likewise, the fact that the sub-vector β_k belongs to the simplex $\mathcal{S}_k^{(r-1)}$ means that the vector of parameters θ has the following restraints ($0 < \beta_{jk} < 1$) and the linear constraints $h_k(\theta) = 0$ with

$$(8) \quad h_k(\theta) = \langle \mathbf{1}_r, \beta_k \rangle - 1, \quad (k = 1, 2, \dots, s)$$

where $\langle \cdot, \cdot \rangle$ is the classic inner product. The latter functions show that assumptions (A5) is satisfied and

$$H_\gamma = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^s; \quad H_\beta = \begin{pmatrix} \mathbf{1}_r^T & \mathbf{0}_r^T & \dots & \mathbf{0}_r^T \\ \mathbf{0}_r^T & \mathbf{1}_r^T & \ddots & \mathbf{0}_r^T \\ \vdots & \ddots & \ddots & \mathbf{0}_r^T \\ \mathbf{0}_r^T & \dots & \mathbf{0}_r^T & \mathbf{1}_r^T \end{pmatrix}_{s \times (sr)}$$

Using model (6), we show that assumptions (A1) to (A4) are also satisfied with (9)

$$\mathcal{L}(\theta) = \text{constant} + \sum_{k=1}^s \sum_{j=1}^r \left\{ x_{jk} \log_e(\beta_{jk}) + x_{2jk} \gamma - x_{jk} \log_e(1 + e^\gamma \sum_{m=1}^r y_{mk} \beta_{mk}) \right\}$$

where $x_{jk} = x_{1jk} + x_{2jk}$ and x_{1jk} (resp. x_{2jk}) stands for the number of accidents of type j on site k before (resp. after) the introduction of the safety measure.

4.2. *Asymptotic confidence interval for the average effect.* Taking the second partial derivatives of the negative of $\mathcal{L}(\theta)$ and using the usual statistical expectation operator, we obtain

$$(10) \quad \begin{aligned} a_k &= \frac{n_k}{1 + e^\gamma \langle y_k, \beta_k \rangle}, \quad b_k = \frac{a_k}{(1 + e^\gamma \langle y_k, \beta_k \rangle)^{1/2}}, \\ \Omega_{\theta,k} &= \text{Diag} \left(\frac{1 + y_{1k} e^\gamma}{\beta_{1k}}, \dots, \frac{1 + y_{rk} e^\gamma}{\beta_{rk}} \right), \\ V_{\theta,k} &= \frac{e^\gamma}{(1 + \langle y_k, \beta_k \rangle e^\gamma)^{1/2}} (y_{1k}, \dots, y_{rk})^T. \end{aligned}$$

Using the latter expressions, we obtain the explicit expression of the asymptotic variance of $\hat{\gamma}$ as follows

$$(11) \quad \sigma^2(\hat{\gamma}) = \frac{1}{(\Lambda_\theta^{(1)}/\Lambda_\theta)}, \quad (\Lambda_\theta^{(1)}/\Lambda_\theta) = (J_\theta/B_\theta) + (\Lambda_\theta^{(0)}/\Lambda_\theta),$$

with

$$\begin{aligned} (J_\theta/B_\theta) &= \sum_{k=1}^s a_k \left\{ \frac{e^\gamma \langle y_k, \beta_k \rangle}{1 + e^\gamma \langle y_k, \beta_k \rangle} - \frac{(\Omega_{\theta,k}^{(0)}/\Omega_{\theta,k})}{1 - (\Omega_{\theta,k}^{(0)}/\Omega_{\theta,k})} \right\}, \\ (\Lambda_\theta^{(0)}/\Lambda_\theta) &= \sum_{k=1}^s \frac{a_k}{n_k} \left\{ \frac{1}{(B_{\theta,k}^{(0)}/B_{\theta,k})} \frac{\Delta_{\theta,k}^2}{[1 - (\Omega_{\theta,k}^{(0)}/\Omega_{\theta,k})]^2} \right\} \end{aligned}$$

and

$$\Lambda_\theta^{(0)} = \begin{bmatrix} \Lambda_\theta & -\xi_\theta \\ \xi_\theta^T & 0 \end{bmatrix}$$

A CONFIDENCE INTERVAL ESTIMATION PROBLEM

91

an $(s + 1) \times (s + 1)$ matrix, $\Delta_{\theta,k} = \mathbf{1}_r^T \Omega_{\theta,k}^{-1} V_{\theta,k}$ a scalar. Thereafter, we deduce the asymptotic confidence interval. We can also get the confidence interval for the other components of the vector of parameters θ . We only need matrix manipulations (see N'Guessan [8], N'Guessan and Langrand [10] for useful technical details).

ACKNOWLEDGEMENT. Part of this work was set up while the first author was a guest researcher at CRT (Centre for Research on Transportation), University of Montreal, whom we would like to thank for their technical and material help.

REFERENCES

1. J. Aitchison and S. D. Silvey, *Maximum likelihood estimation of parameters subject to restraints*. Ann. Math. Statist. **29** (1958), 813–829.
2. M. Crowder, *On the constrained maximum likelihood estimation with non i.i.d. observations*. Ann. Inst. Statist. Math. **36** (1984), 239–249.
3. F. J. H. Don, *The use of generalized inverses in restricted maximum likelihood*. Linear Algebra Appl. **70** (1985), 225–240.
4. D. W. Hosmer and S. Lemeshow, *Applied logistic regression*. 2nd edition, John Wiley & Sons, 2000.
5. J. R. Magnus, *Linear Structures*. Charles Griffin & Co., London, 1988.
6. C. R. Mehta and N. R. Patel, *Exact logistic regression: theory and examples*. Statistics in Medicine **14** (1995), 2143–2160.
7. A. N'Guessan, *Constrained covariance matrix estimation in road accident modelling with Schur complements*. C. R. Acad. Sci. Paris Ser. I **337** (2003), 219–222.
8. ———, *Constrained estimation of a road safety covariance matrix using Schur complements*. Technical Report, Centre for Research on Transportation, Université de Montréal, CRT-2003-12.
9. A. N'Guessan, A. Essai and C. Langrand, *Estimation multidimensionnelle des contrôles et de l'effet moyen d'une mesure de sécurité routière*. Rev. Statist. Appl. (2) **XLIX** (2001), 83–100.
10. A. N'Guessan and C. Langrand, *A Schur complement approach for computing sub-covariance matrices arising in a road safety measure modelling*. J. Comput. Appl. Math. **177** (2005), 331–345.
11. D. V. Ouellette, *Schur complements and statistics*. Linear Algebra Appl. **36** (1981), 187–295.
12. C. R. Rao and H. Yanai, *Generalized inverses of partitioned matrices useful in statistical applications*. Linear Algebra Appl. **70** (1985), 105–115.
13. S. D. Silvey, *The Lagrangian multiplier test*. Ann. Math. Statist. **30** (1959), 389–407.
14. K. Tanabe and M. Sagae, *An Exact Cholesky Decomposition and the Generalized Inverse of the Variance-Covariance Matrix of the Multinomial Distribution, with Applications*. J. Roy. Statist. Soc. Ser. B (1) **54** (1992), 211–219.
15. J. C. Tanner, *A problem in the combination of accident frequencies*. Biometrika **45** (1958), 331–342.
16. F. Zhang (ed.), *Schur complement and its applications*. Springer-Verlag, 2005.

Polytech'Lille and Laboratoire
Paul Painlevé CNRS, UMR 8542
Université de Lille 1
Bât. Polytech'Lille
59655 Villeneuve d'Ascq Cedex
France
email: assi.nguessan@polytech-lille.fr

HEC Montréal and Laboratory on
Transportation Safety
Centre for Research on Transportation
Université de Montréal,
P.O. Box 6128
Succ. Centre-ville
Montréal, Québec
H3C 3J7
email: francois.bellavance@hec.ca

Annexe A05 :

Rev. Statistique Appliquée, **XLIX (2)**, 83-100 (2001)

Estimation multidimensionnelle des contrôles et de l'effet moyen d'une mesure de sécurité routière

Rev. Statistique Appliquée, 2001, XLIX (2), 85-102

ESTIMATION MULTIDIMENSIONNELLE DES CONTRÔLES ET DE L'EFFET MOYEN D'UNE MESURE DE SÉCURITÉ ROUTIÈRE

A. N'Guessan*, A. Essai**, C. Langrand*

**Laboratoire de Statistique et Probabilités : CNRS, FRE 2222.*

Université de Lille 1, 59655 Villeneuve d'Ascq Cedex, France.

E-mail : Assi.Nguessan@eudil.fr, Claude.Langrand@univ-lille1.fr

***Laboratoire d'Analyse Numérique et d'Optimisation,*

Université de Lille 1, 59655 Villeneuve d'Ascq Cedex, France.

E-mail : essai@ano.univ-lille1.fr

RÉSUMÉ

On s'intéresse à l'estimation de l'effet d'une mesure de sécurité routière appliquée à S sites expérimentaux, une zone de contrôle étant associée à chacun d'eux. On suppose que chaque site comporte R ($R > 1$) types d'accidents mutuellement exclusifs. On fait l'hypothèse que les données d'accidents relatifs à cette mesure sont mises sous la forme de tableaux avant-après. On construit alors, pour chaque site expérimental, un modèle multinomial prenant en compte l'effet moyen à tous les sites et la tendance moyenne des variations en nombre d'accidents dans la zone de contrôle de chaque site. On propose une méthode pour estimer le vecteur paramètre lié au modèle; Des tableaux de données d'accidents simulés nous permettent de résoudre numériquement les équations du maximum de vraisemblance et d'étudier l'erreur quadratique moyenne.

Mots-clés : Mesure de sécurité routière, Accident de la route, Méthode avant-après, Modèle multinomial, Maximum de vraisemblance sous contraintes, Erreur quadratique moyenne.

ABSTRACT

We deal with the estimation of the effect of a road safety measure applied on S target sites with a control area for each site. We suppose that the accidents data recorded at each site are classified in R ($R > 1$) mutually exclusive types. We adopt the before-after technique and assume that at any one target site the total number of accidents recorded is multinomially distributed between the periods and the types of accidents. An estimation method of the unknown vector parameter is suggested. Some simulated accidents data allow us to solve numerically the maximum likelihood's equations and to study the consistence of the estimator via the mean square error.

Keywords : Road Safety measure, Accident data, Before-after method, Multinomial pattern, Restraints Maximum likelihood, Mean Square Error.

1. Introduction

Les méthodes statistiques utilisées dans l'estimation avant-après de l'efficacité d'une mesure de sécurité routière dépendent fortement des données disponibles et des objectifs fixés lors de la mise en place de cette mesure (*voir* Lassarre (1985) [8], Hauer (1997) [7]). Quelle que soit la méthode, le principe de base de l'estimation consiste à prédire la valeur moyenne de la variable aléatoire (ou critère d'appréciation) utilisée si la mesure n'avait pas été appliquée. En pratique, on utilise souvent le nombre d'accidents comme variable aléatoire. Pour un seul site par exemple, on compte les nombres d'accidents survenus avant l'application de cette mesure (période t_1) et après celle-ci (période t_2) et il faut pouvoir estimer si la modification du nombre des accidents (passant par exemple de n_{t_1} à n_{t_2}) est significativement différente de celle qu'on aurait observée si la mesure n'avait pas été appliquée ou n'était pas efficace. L'interprétation des résultats doit tenir compte d'un certain nombre de facteurs (évolution du trafic, conditions météorologiques...); si en étudiant d'autres données (par exemple les variations du nombre des accidents sur un site équivalent non soumis à la mesure) on peut voir que le nombre d'accidents est multiplié par un facteur c (coefficient de contrôle, ou d'évolution), l'appréciation de l'efficacité de la mesure doit tenir compte de ce facteur c . Lorsqu'une mesure est appliquée à plusieurs sites en même temps, il n'est pas en général réaliste de considérer, pour des raisons de localisation des sites par exemple, que le même facteur c puisse être utilisé quel que soit le site. C'est pourquoi Tanner (1958) [15] a introduit la notion de zone de contrôle et défini un coefficient de contrôle pour chaque site à l'aide du quotient nombre des accidents « après » sur nombre des accidents « avant » dans la zone de contrôle. Supposons maintenant que, la mesure ayant été appliquée à plusieurs sites, on dispose du nombre total d'accidents pour chaque site expérimental au cours des deux périodes. Puis, de façon rétrospective, on s'intéresse au niveau de chaque site à l'effet de la mesure sur un certain nombre de types d'accidents (accidents mortels, blessés graves, blessés légers, dégâts matériels, etc). Une telle démarche induit une troisième dimension (typologie d'accidents) au modèle et à la statistique de Tanner et impose d'estimer l'effet de la mesure par rapport à des tableaux avant-après de données d'accidents. C'est dans cette optique que N'guessan et Langrand (1993) [13] ont généralisé l'approche de Tanner à des tableaux de données d'accidents avant-après avec contrôle. Ils supposent que la même mesure de sécurité routière est appliquée à plusieurs sites chacun comportant différents types d'accidents mutuellement exclusifs et étant associé à une zone de contrôle spécifique. Ils suggèrent des tests statistiques pour juger si l'effet de la mesure est le même quel que soit le site et quel que soit le type d'accident. On s'intéresse essentiellement dans ce travail à la recherche par simulation d'une solution aux équations du maximum de vraisemblance et à la convergence numérique de cet estimateur. La section 2, est consacrée aux données et à la formulation du problème. On expose dans la section 3 le mécanisme de construction du modèle statistique utilisé pour répartir l'effectif total d'accidents de chaque site à la fois entre les deux périodes et entre les différents types d'accidents. La méthode d'estimation du vecteur paramètre lié à ce modèle et incorporant l'effet moyen de la mesure est exposée à la section 4. La section 5 est exclusivement dédiée aux résultats numériques sur données simulées. On étudie en particulier l'ordre de grandeur de l'erreur quadratique moyenne pour des tableaux avant-après de faibles et grands effectifs d'accidents.

2. Données et formulation du problème

On considère dans ce travail qu'une mesure de sécurité routière (aménagement d'un carrefour, revêtement d'un tronçon d'autoroute,...) a été appliquée à S sites. On suppose disposer, sur chacun des S sites, du nombre total d'accidents avant et après l'imposition de la mesure pour des périodes de longueurs convenables (généralement égales et souvent d'au moins un an). La question qui est posée ici est la suivante : Comment estimer l'effet moyen de la mesure relative à l'ensemble des S sites, dans le cas où chacun d'eux comporte $R (R > 1)$ types d'accidents? Les données dont on dispose se trouvent dans le tableau 1 suivant :

TABLEAU 1

	Avant			Après			total
	type 1	... type r	... type R	type 1	... type r	... type R	
site 1	n_{111}	... n_{1r1}	n_{1R1}	n_{211}	... n_{2r1}	... n_{2R1}	n_1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
site s	n_{11s}	... n_{1rs}	n_{1Rs}	n_{21s}	... n_{2rs}	... n_{2Rs}	n_s
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
site S	n_{11S}	... n_{1rS}	n_{1RS}	n_{21S}	... n_{2rS}	... n_{2RS}	n_S
							N

où n_{1rs} (resp. n_{2rs}) représente le nombre d'accidents du type r du site s avant (resp. après) la mise en place de la mesure, n_s le nombre total d'accidents du site

s et $N = \sum_{s=1}^S n_s$. Afin de tenir compte de certains facteurs extérieurs (tendances

générales des débits de circulation, variations des limitations de vitesse, conditions météorologiques...) on suppose également disposer pendant les mêmes périodes de données sur des zones de contrôle non soumises à la mesure et associées aux sites expérimentaux qui permettent de construire le tableau de coefficients de contrôle suivant :

TABLEAU 2

	type 1	... type r	... type R
site 1	c_{11}	... c_{r1}	c_{R1}
⋮	⋮	⋮	⋮
site s	c_{1s}	... c_{rs}	c_{Rs}
⋮	⋮	⋮	⋮
site S	c_{1S}	... c_{rS}	c_{RS}

c_{rs} étant le coefficient multiplicatif qu'il faut employer dans la zone de contrôle associée au site s et pour le type r d'accidents pour passer du nombre d'accidents observés dans la période « avant » au nombre observé dans la période « après ». Il s'agit alors de combiner les données des tableaux 1 et 2 pour estimer convenablement l'effet moyen de la mesure relativement à tous les sites et types d'accidents et de proposer ainsi une réponse à la question précédente.

3. Modèle statistique

Considérons la période avant (resp. après) l'implémentation de la mesure et notons X_{1rs} (resp. X_{2rs}) la variable aléatoire (v.a.) donnant le nombre d'accidents du type r sur le site expérimental s . Pour modéliser l'effet moyen de la mesure de sécurité, on suppose que le nombre moyen $E(X_{2rs})$ d'accidents du type r au site s qu'on pourrait s'attendre à observer dans la période « après » si ce site avait un comportement semblable à celui de la zone de contrôle associée est :

$$\theta^* c_{rs} E(X_{1rs}) \quad (1)$$

où $E(X_{1rs})$ est l'espérance mathématique de la v.a. X_{1rs} et θ^* le rapport des espérances mathématiques du nombre total d'accidents observés dans la période « après », et du nombre d'accidents qu'on pourrait s'attendre à observer dans cette période si chaque site se comportait comme sa zone de contrôle :

$$\theta^* = \frac{\sum_{s=1}^S \sum_{r=1}^R E(X_{2rs})}{\sum_{s=1}^S \sum_{r=1}^R c_{rs} E(X_{1rs})}$$

Si le site s a un comportement semblable à celui de la zone de contrôle associée et si l'on estime $E(X_{1rs})$ par n_{1rs} , valeur observée de X_{1rs} lue dans le tableau 1, alors l'expression (1) devient

$$\theta^* c_{rs} n_{1rs}. \quad (2)$$

Sous ces mêmes hypothèses et en considérant les périodes « avant » et « après », le nombre total d'accidents espérés au site s est estimé par :

$$n_s^* = n_{11s} + n_{12s} + \dots + n_{1Rs} + \theta^* c_{1s} n_{11s} + \theta^* c_{2s} n_{12s} + \dots + \theta^* c_{Rs} n_{1Rs}$$

Cette dernière égalité implique que les proportions d'accidents espérés par type au site s au cours des deux périodes (avant et après) sont estimées par

$$\frac{n_{11s}}{n_s^*}, \frac{n_{12s}}{n_s^*}, \dots, \frac{n_{1Rs}}{n_s^*}, \frac{\theta^* c_{1s} n_{11s}}{n_s^*}, \frac{\theta^* c_{2s} n_{12s}}{n_s^*}, \dots, \frac{\theta^* c_{Rs} n_{1Rs}}{n_s^*} \quad (3)$$

soit encore par

$$\frac{p_{1s}^*}{1 + \theta^* c_s^*}, \frac{p_{2s}^*}{1 + \theta^* c_s^*}, \dots, \frac{p_{Rs}^*}{1 + \theta^* c_s^*}, \frac{\theta^* c_{1s} p_{1s}^*}{1 + \theta^* c_s^*}, \frac{\theta^* c_{2s} p_{2s}^*}{1 + \theta^* c_s^*}, \dots, \frac{\theta^* c_{Rs} p_{Rs}^*}{1 + \theta^* c_s^*}$$

en posant

$$p_{rs}^* = n_{1rs} / \left(\sum_{j=1}^R n_{1js} \right), \quad c_s^* = \sum_{r=1}^R c_{rs} p_{rs}^*$$

p_{rs}^* est une estimation de la probabilité pour qu'un accident qui se produit au site s avant la mise en place de la mesure soit du type r , avec $\sum_{r=1}^R p_{rs}^* = 1, s = 1, 2, \dots, S$. Pour s fixé, c_s^* est la somme des valeurs c_{rs} du tableau 2 pondérée par les p_{rs}^* ($r = 1, 2, \dots, R$). c_s^* représente une estimation du coefficient de contrôle moyen associé au site s si ce site se comportait comme sa zone de contrôle associée.

On pose alors

$$q_{1rs}^* = \frac{p_{rs}^*}{1 + \theta^* c_s^*}, \quad q_{2rs}^* = \frac{\theta^* c_{rs} p_{rs}^*}{1 + \theta^* c_s^*}, \quad r = 1, 2, \dots, R. \tag{4}$$

Par construction, les q_{trs}^* ($t = 1$ ou 2) représentent la répartition des n_s^* accidents en $2R$ catégories mutuellement exclusives entre les périodes « avant » et « après » si le site s se comporte comme sa zone de contrôle et si θ^* est l'effet moyen de la mesure; on a

$$\sum_{r=1}^R \sum_{t=1}^2 q_{trs}^* = 1. \tag{5}$$

Les considérations précédentes nous conduisent à proposer la modélisation suivante : on considère le vecteur aléatoire ($\vec{v}.a.$)

$$X^{(s)} = (X_{11s}, X_{12s}, \dots, X_{1Rs}, X_{21s}, X_{22s}, \dots, X_{2Rs}) \quad s = 1, 2, \dots, S$$

prenant simultanément en compte les v.a. $X_{11s}, X_{12s}, \dots, X_{1Rs}, X_{21s}, X_{22s}, \dots, X_{2Rs}$ nombres d'accidents des R types avant et après la mesure au site expérimental s , de valeur observée $n^{(s)} = (n_{11s}, n_{12s}, \dots, n_{1Rs}, n_{21s}, n_{22s}, \dots, n_{2Rs})$ (ligne s du tableau 1). Pour s fixé, on s'intéresse à la loi de probabilité du $\vec{v}.a. X^{(s)}$ étant donnés les tableaux 1 et 2. En s'inspirant des relations (4) et (5), on fait alors l'hypothèse que pour s fixé le $\vec{v}.a. X^{(s)}$ est de loi multinomiale :

$$Proba[X^{(s)} = n^{(s)}] = \left(\frac{n_s!}{\prod_{t=1}^2 \prod_{r=1}^R n_{trs}!} \right) \prod_{t=1}^2 \prod_{r=1}^R q_{trs}^{n_{trs}} \tag{6}$$

90

A. N'GUESSAN, A. ESSAI, C. LANGRAND

où

$$q_{1rs} = \frac{p_{rs}}{1 + \theta c_s}, \quad q_{2rs} = \frac{\theta c_{rs} p_{rs}}{1 + \theta c_s}, \quad r = 1, 2, \dots, R \quad (7)$$

avec θ ($\theta > 0$) paramètre effet moyen de la mesure de sécurité routière relative à l'ensemble des sites et des types d'accidents, les p_{rs} ($0 < p_{rs} < 1$) étant des paramètres auxiliaires liés aux zones de contrôle et aux sites expérimentaux tels que

$$\sum_{r=1}^R p_{rs} = 1, \quad c_s = \sum_{r=1}^R c_{rs} p_{rs}. \quad (8)$$

Si on note C_s la v.a. prenant les valeurs $c_{1s}, c_{2s}, \dots, c_{Rs}$ de la ligne s du tableau 2 avec les probabilités $p_{1s}, p_{2s}, \dots, p_{Rs}$ alors $c_s = \sum_{r=1}^R c_{rs} p_{rs}$ est l'espérance mathématique de la v.a. C_s et est appelé coefficient moyen de contrôle associé au site expérimental s . Les paramètres auxiliaires $(p_{rs})_{1 \leq r \leq R}$ sont interprétés comme loi de probabilité de C_s . Les c_1, c_2, \dots, c_S seront entièrement déterminés une fois les lois de probabilité $((p_{rs})_{1 \leq r \leq R})_{1 \leq s \leq S}$ connues.

4. Estimation simultanée des coefficients moyens de contrôle et de l'effet moyen de la mesure

Par construction du modèle, on peut distinguer deux types de paramètres parmi ceux associés aux S sites expérimentaux et R types d'accidents : le paramètre d'intérêt θ et l'ensemble des paramètres auxiliaires $P_s = (p_{1s}, p_{2s}, \dots, p_{Rs})^T$, $s = 1, 2, \dots, S$. Dans la suite de ce travail, on s'intéresse à l'estimation simultanée de ces paramètres et on pose alors $\Theta = (\theta, P_1^T, P_2^T, \dots, P_S^T)^T$ l'élément de \mathbb{R}^{1+SR}

ainsi défini (en réalité on est dans $\mathbb{R}^{1+S(R-1)}$ puisque $\sum_{r=1}^R p_{rs} = 1$). On fait alors

l'hypothèse que les $\vec{v}.a. X^{(1)}, X^{(2)}, \dots, X^{(S)}$ sont indépendants et suivent chacun une distribution multinomiale dont la forme est donnée par (6) et (7). Etant donné les tableaux 1 et 2, la fonction de vraisemblance associée à la suite des $\vec{v}.a. X^{(1)}, X^{(2)}, \dots, X^{(S)}$ est :

$$L(X^{(1)}, X^{(2)}, \dots, X^{(S)}; \Theta) = \prod_{s=1}^S l_x(X^{(s)}; \Theta) \quad (9)$$

où

$$l_x(X^{(s)}; \Theta) = \frac{n_s!}{\prod_{t=1}^2 \prod_{m=1}^R n_{tms}!} \prod_{r=1}^R \frac{p_{rs}^{n_{rs}} (\theta c_{rs})^{n_{2rs}}}{(1 + \theta \sum_{j=1}^R c_{js} p_{js})^{n_{rs}}} \quad (10)$$

avec $n_{.rs} = n_{1rs} + n_{2rs}$. La log-vraisemblance est donnée à une constante additive près par :

$$\mathcal{L}(X^{(1)}, X^{(2)}, \dots, X^{(S)}; \Theta) = \sum_{s=1}^S \sum_{r=1}^R \{n_{.rs} \log_e(p_{rs}) + n_{2rs} \log_e(\theta) - n_{.rs} \log_e(1 + \theta \sum_{m=1}^R c_{ms} p_{ms})\}. \quad (11)$$

En dérivant la fonction $\Theta \mapsto \mathcal{L}(X^{(1)}, X^{(2)}, \dots, X^{(S)}; \Theta)$ par rapport aux composantes de Θ et en annulant les dérivées partielles (tout en tenant compte de ce que $\sum_{r=1}^R p_{rs} = 1, s = 1, 2, \dots, S$) on montre que l'estimateur, s'il existe, $\hat{\Theta}$ de Θ est solution du système d'équations non linéaires sous contraintes :

$$\left\{ \begin{array}{l} \sum_{s=1}^S \sum_{r=1}^R \frac{n_{2rs} - \hat{\theta} \hat{c}_s n_{1rs}}{1 + \hat{\theta} \hat{c}_s} = 0; \\ n_{.rs} - \frac{n_s \hat{p}_{rs} (\hat{\theta} c_{rs} + 1)}{1 + \hat{\theta} \hat{c}_s} = 0, \\ (s = 1, 2, \dots, S; \quad r = 1, 2, \dots, R); \\ \hat{\theta} > 0; \quad 0 < \hat{p}_{rs} < 1; \\ \hat{c}_s = \sum_{r=1}^R c_{rs} \hat{p}_{rs}, \quad s = 1, 2, \dots, S; \\ \sum_{r=1}^R \hat{p}_{rs} = 1, \quad s = 1, 2, \dots, S. \end{array} \right. \quad (12)$$

$\hat{\Theta}$ est l'estimation du maximum de vraisemblance sous contraintes linéaires et de bornes (voir par exemple, [1], [2], [5], [11]). On aborde ici uniquement les aspects numériques de l'existence de $\hat{\Theta}$ par résolution et simulation du système (12).

5. Résultats numériques de simulation

5.1. Principe de simulation des données avant-après avec contrôle

L'étude numérique des équations de vraisemblance (12) et donc de l'existence numérique de $\hat{\Theta}$ nécessite la connaissance des tableaux de données (tableaux 1 et 2). Cette résolution numérique comporte deux grandes phases. Une première phase de simulation du tableau des coefficients de contrôle c_{rs} (tableau 2) et du tableau des données avant-après (tableau 1) par programmation en FORTRAN du modèle multinomial construit à la section 3. Puis suit la phase de résolution proprement dite (recherche numérique de $\hat{\Theta}$) par utilisation des algorithmes (E04ZCF et E04UCF) d'optimisation du logiciel NAG [11] faisant appel aux tableaux 1 et 2 précédemment simulés.

Etant donné donc S (le nombre de sites expérimentaux) et R (le nombre de types d'accidents), on suppose que les coefficients c_{rs} ($s = 1, 2, \dots, S$; $r = 1, 2, \dots, R$) proviennent d'une *v.a.* de loi $U_{] \frac{1}{2}, \frac{5}{2} [}$ (loi uniforme entre $\frac{1}{2}$ et $\frac{5}{2}$). Pour l'obtention du tableau 1, on pose $\theta^0 = \frac{1}{2}$ et on fait l'hypothèse que les composantes du vecteur $P_s^0 = (p_{1s}^0, p_{2s}^0, \dots, p_{Rs}^0)^T$ vraie valeur du vecteur P_s proviennent d'une *v.a.* de loi $U_{[\alpha, 1-\alpha]}$ (avec $\alpha = 10^{-5}$ de manière à permettre l'écriture de la log-vraisemblance) et sont normalisées de manière que $\sum_{r=1}^R p_{rs}^0 = 1$; $s = 1, 2, \dots, S$. En utilisant ces valeurs, on détermine alors les vrais coefficients moyens de contrôle

$$c_s^0 = \sum_{r=1}^R c_{rs} p_{rs}^0, \quad (s = 1, 2, \dots, S)$$

puis ensuite les vraies probabilités

$$q_{1rs}^0 = \frac{p_{rs}^0}{1 + \theta^0 c_s^0}, \quad q_{2rs}^0 = \frac{c_{rs} \theta^0 p_{rs}^0}{1 + \theta^0 c_s^0}, \quad r = 1, 2, \dots, R$$

associée à la distribution multinomiale relative au \vec{v} . *a.* $X^{(s)}$. Enfin pour s ($s = 1, 2, \dots, S$) fixé, on génère l'effectif total n_s du site s et on répartit de façon aléatoire n_s entre les périodes « avant » et « après » par utilisation des probabilités q_{1rs}^0 et q_{2rs}^0 . On obtient ainsi les valeurs observées n_{trs} du tableau 1 avec $\sum_{r=1}^R \sum_{t=1}^2 n_{trs} = n_s$, $s = 1, 2, \dots, S$. Plusieurs tableaux de différentes tailles et de différents effectifs ont été générés en fixant les valeurs de S et R et en choisissant n_s dans différentes tranches d'accidents. Les résultats numériques (voir paragraphe 5.3) sont obtenus pour n_s appartenant aux tranches d'accidents T_1, \dots, T_{10} suivantes :

$$\{]0, 30[,]30, 50[,]50, 300[,]300, 1000[,]1000, 3000[,]3000, 5000[,]5000, 7000[,]7000, 9000[,]9000, 11000[,]11000, 13000\}.$$

5.2. Estimation et convergence de $\hat{\Theta}$

Le vecteur paramètre à estimer $\Theta \in \mathbb{R}^{1+SR}$ a pour vraie valeur

$$\Theta^0 = (\theta^0, P_1^{0T}, P_2^{0T}, \dots, P_S^{0T})^T$$

avec θ^0 fixé à 0.5 et $P_1^0, P_2^0, \dots, P_S^0$ des vecteurs de dimension R . L'objectif principal de la simulation est de rechercher d'abord une solution $\hat{\Theta}$ aux équations de vraisemblance et d'étudier la convergence de $\hat{\Theta}$ vers Θ^0 par l'intermédiaire de

l'analyse de l'erreur quadratique moyenne (EQM) qui mesure l'adéquation entre $\hat{\Theta}$ et Θ^0 et est définie par

$$EQM(\hat{\Theta}, \Theta^0) = \frac{1}{1 + RS} \sum_{m=1}^{1+RS} (\hat{\Theta}_m - \Theta_m^0)^2.$$

On utilise ensuite les composantes $\hat{P}_1, \hat{P}_2, \dots, \hat{P}_S$ de $\hat{\Theta}$ pour estimer (voir paragraphe 5.3) et étudier la convergence des \hat{c}_s vers c_s^0 en analysant l'EMC (Ecart Moyen de Contrôle) et l'EQMC (Ecart Quadratique Moyen de Contrôle) définis ci-après :

$$EMC(\hat{c}_s, c_s^0) = \frac{1}{S} \sum_{s=1}^S |\hat{c}_s - c_s^0|; \quad EQMC(\hat{c}_s, c_s^0) = \frac{1}{S} \sum_{s=1}^S (\hat{c}_s - c_s^0)^2.$$

Enfin on se focalise sur le comportement de la première composante (paragraphe 5.4) $\hat{\theta}$ de $\hat{\Theta}$ qui est l'estimateur de l'effet moyen de la mesure de sécurité dont la vraie valeur est $\theta^0 = 0.5$.

TABLEAU 3
Valeurs de $\hat{\Theta}$ pour $S = 4$ et $R = 2$.

Θ^0	$\hat{\Theta}$						
	$T = 1$	$T = 2$	$T = 4$	$T = 6$	$T = 8$	$T = 9$	$T = 10$
5.0000e-1	5.3094e-1	5.0439e-1	4.6923e-1	4.7080e-1	4.7807e-1	4.8208e-1	4.8236e-1
9.3099e-1	8.0479e-1	8.6918e-1	9.2232e-1	9.2797e-1	9.3141e-1	9.3126e-1	9.3174e-1
6.9010e-2	1.9521e-1	1.3082e-1	7.7680e-2	7.2030e-2	6.8590e-2	6.8740e-2	6.8260e-2
6.8019e-1	9.9999e-1	5.0774e-1	6.8164e-1	6.9053e-1	6.9104e-1	6.9561e-1	6.9461e-1
3.1981e-1	1.0000e-5	4.9226e-1	3.1836e-1	3.0947e-1	3.0896e-1	3.0439e-1	3.0539e-1
4.1127e-1	4.6512e-1	3.8411e-1	3.9899e-1	4.1712e-1	4.1677e-1	4.0974e-1	4.1172e-1
5.8873e-1	5.3488e-1	6.1589e-1	6.0101e-1	5.8288e-1	5.8323e-1	5.9026e-1	5.8828e-1
6.6746e-1	8.8382e-1	6.6582e-1	6.6267e-1	6.6807e-1	6.6211e-1	6.7331e-1	6.6320e-1
3.3254e-1	1.1618e-1	3.3418e-1	3.3733e-1	3.3193e-1	3.3789e-1	3.2669e-1	3.3680e-1
EQM	3.7420e-2	7.6237e-3	1.6095e-4	1.2821e-4	9.2758e-5	9.6645e-5	8.5027e-5

TABLEAU 4
Valeurs de $\hat{\Theta}$ pour $S = 4$ et $R = 3$.

Θ^0	$\hat{\Theta}$						
	$T = 1$	$T = 2$	$T = 4$	$T = 6$	$T = 8$	$T = 9$	$T = 10$
5.0000e-1	4.1035e-1	4.5889e-1	5.0496e-1	4.8721e-1	4.9125e-1	4.8978e-1	4.9024e-1
2.6775e-1	1.5461e-1	2.7960e-1	2.7525e-1	2.7457e-1	2.7090e-1	2.7097e-1	2.7167e-1
1.9850e-2	1.0000e-5	1.0000e-5	1.6330e-2	1.7380e-2	1.9180e-2	1.8730e-2	1.8800e-2
7.1241e-1	8.4538e-1	7.2039e-1	7.0843e-1	7.0805e-1	7.0991e-1	7.1030e-1	7.0953e-1
2.1941e-1	3.9929e-1	1.9002e-1	2.2650e-1	2.2371e-1	2.2338e-1	2.2374e-1	2.1945e-1
3.2103e-1	6.0070e-1	5.1472e-1	2.9518e-1	3.0859e-1	3.2144e-1	3.2393e-1	3.2357e-1
4.5956e-1	1.0000e-5	2.9526e-1	4.7832e-1	4.6770e-1	4.5518e-1	4.5233e-1	4.5698e-1
3.1570e-1	6.0772e-1	3.5021e-1	3.3186e-1	3.1235e-1	3.1378e-1	3.1286e-1	3.1667e-1
1.5729e-1	1.0000e-5	1.5312e-1	1.5034e-1	1.6176e-1	1.5443e-1	1.5581e-1	1.5422e-1
5.2702e-1	3.9227e-1	4.9667e-1	5.1780e-1	5.2589e-1	5.3180e-1	5.3133e-1	5.2911e-1
3.9505e-1	4.8206e-1	4.3417e-1	3.7513e-1	3.8995e-1	3.9367e-1	3.8916e-1	3.9118e-1
4.8130e-1	4.2259e-1	4.2579e-1	4.9792e-1	4.9061e-1	4.8124e-1	4.9242e-1	4.8983e-1
1.2365e-1	9.5350e-2	1.4004e-1	1.2695e-1	1.1945e-1	1.2509e-1	1.1843e-1	1.1900e-1
EQM	3.8512e-2	5.7439e-3	1.7372e-4	4.9044e-5	1.2835e-5	3.1863e-5	1.9789e-5

TABLEAU 5
Valeurs de $EQM(\hat{\Theta}, \Theta^0)$ pour les 10 tranches d'accidents.

T	S = 5		T	S = 20	
	R = 2	R = 10		R = 2	R = 10
1	2.6850E-2	1.8480E-2	1	3.8793E-2	1.6054E-2
2	5.3214E-3	1.8482E-3	2	6.0657E-3	2.3454E-3
3	8.9237E-4	4.8812E-4	3	2.4164E-3	8.5019E-4
4	7.8695E-4	1.9116E-4	4	4.5135E-4	1.5882E-4
5	2.6904E-4	3.2655E-5	5	7.2394E-5	4.0908E-5
6	5.8576E-5	2.4412E-5	6	3.2943E-5	2.2612E-5
7	3.9579E-5	1.4034E-5	7	3.0454E-5	1.8778E-5
8	1.5870E-5	9.5464E-6	8	4.8501E-5	1.1726E-5
9	1.6832E-5	9.4963E-6	9	2.3679E-5	9.3814E-6
10	8.2177E-6	8.2518E-6	10	2.1514E-5	8.3590E-6

TABLEAU 6
Valeurs de $EQM(\hat{\Theta}, \Theta^0)$ pour $n_s \in]0, 30[$.

S/R	2	3	4	5	6	7	8	9	10
1	2.347e-2	2.726e-2	1.270e-1	1.930e-3	1.342e-2	5.890e-3	1.772e-2	4.711e-2	6.632e-3
2	1.282e-2	9.293e-3	5.170e-2	2.160e-3	9.044e-3	1.474e-2	2.196e-3	4.206e-2	2.440e-2
3	1.145e-2	3.002e-2	2.444e-2	5.957e-3	2.209e-2	1.155e-2	8.153e-3	7.182e-3	1.917e-2
4	3.742e-2	3.851e-2	4.879e-2	2.322e-2	8.693e-3	3.092e-2	1.228e-2	1.290e-2	6.966e-3
5	2.685e-2	4.082e-2	1.217e-2	1.421e-2	1.409e-2	1.994e-2	1.094e-2	6.691e-3	1.848e-2
6	6.811e-2	5.576e-2	1.089e-2	2.455e-2	1.816e-2	1.230e-2	1.825e-2	2.182e-2	1.055e-2
7	5.306e-2	1.740e-2	1.683e-2	3.150e-2	1.133e-2	2.406e-2	2.084e-2	1.100e-2	1.710e-2
8	4.241e-2	1.865e-2	1.564e-2	6.814e-3	2.715e-2	1.805e-2	7.295e-3	1.451e-2	1.692e-2
9	5.680e-2	4.341e-2	3.781e-2	2.049e-2	1.697e-2	8.426e-3	2.703e-2	1.005e-2	7.622e-3
10	1.613e-2	2.726e-2	1.377e-2	3.513e-2	1.389e-2	1.589e-2	1.548e-2	4.664e-3	1.084e-2

TABLEAU 7
Valeurs de $EQM(\hat{\Theta}, \Theta^0)$ pour $n_s \in]1000, 3000[$.

S/R	2	3	4	5	6	7	8	9	10
1	9.171e-4	1.600e-3	2.722e-4	4.290e-4	4.496e-4	3.142e-4	1.083e-4	9.915e-5	2.029e-4
2	6.568e-5	1.309e-4	6.783e-5	1.333e-4	5.466e-5	5.733e-5	1.101e-4	6.047e-5	3.740e-5
3	1.778e-5	2.159e-4	6.407e-5	8.938e-5	1.053e-4	7.605e-5	7.212e-5	6.265e-5	7.335e-5
4	3.430e-4	6.474e-5	1.279e-4	8.235e-5	6.360e-5	8.585e-5	1.220e-4	7.270e-5	4.409e-5
5	2.690e-4	7.882e-5	9.696e-5	1.288e-4	5.424e-5	5.920e-5	4.764e-5	5.065e-5	3.266e-5
6	3.238e-5	1.689e-4	4.080e-5	9.685e-5	4.810e-5	3.265e-5	7.108e-5	4.138e-5	6.603e-5
7	1.575e-4	2.163e-4	9.277e-5	1.103e-4	8.133e-5	1.039e-4	4.449e-5	5.136e-5	5.366e-5
8	2.852e-4	1.147e-4	1.109e-4	7.999e-5	1.462e-4	5.997e-5	5.060e-5	4.875e-5	5.415e-5
9	1.538e-4	9.020e-5	7.464e-5	1.086e-4	9.135e-5	4.888e-5	4.992e-5	4.379e-5	4.923e-5
10	9.104e-5	1.236e-4	9.078e-5	8.324e-5	6.447e-5	6.406e-5	6.046e-5	7.232e-5	5.495e-5

5.3. Estimation et convergence des contrôles moyens

TABLEAU 8
Estimation des contrôles moyens pour $S = 4$ et $R = 2$.

c_s^0	\hat{c}_s						
	$T = 1$	$T = 2$	$T = 4$	$T = 6$	$T = 8$	$T = 9$	$T = 10$
1.3522	1.5012	1.4252	1.3624	1.3558	1.3517	1.3519	1.3513
1.9271	2.4512	1.6444	1.9294	1.9440	1.9449	1.9523	1.9507
1.3618	1.3865	1.3493	1.3562	1.3645	1.3643	1.3611	1.3620
1.0523	8.9612e-1	1.0535	1.0558	1.0519	1.0562	1.0481	1.0554
EMC	1.3540e-1	5.5232e-2	2.6107e-3	5.6889e-3	5.9208e-3	5.0079e-3	6.5096e-3
EQMC	8.0487e-2	.1341e-2	3.8520e-5	7.6836e-5	8.4519e-5	1.6431e-4	1.4231e-4

TABLEAU 9
Estimation des contrôles moyens pour $S = 4$ et $R = 3$.

c_s^0	\hat{c}_s						
	$T = 1$	$T = 2$	$T = 4$	$T = 6$	$T = 8$	$T = 9$	$T = 10$
1.5513	1.5759	1.5308	1.5458	1.5469	1.5497	1.5493	1.5491
1.3377	1.4667	1.2303	1.3561	1.3477	1.3427	1.3422	1.3369
1.0763	1.2720	1.1061	1.0880	1.0753	1.0733	1.0731	1.0758
1.3383	1.2550	1.3219	1.3550	1.3389	1.3403	1.3388	1.3377
EMC	6.6486e-2	2.8671e-2	1.0301e-2	1.2852e-3	5.5014e-4	9.7155e-5	1.0570e-3
EQMC	1.5611e-2	3.2824e-3	1.9652e-4	3.0317e-5	1.0155e-5	8.8405e-6	1.6337e-6

TABLEAU 10
Valeurs de EMC pour $n_s \in]0, 30[$.

S/R	2	3	4	5	6	7	8	9	10
1	1.147e-2	2.550e-2	4.328e-1	8.259e-2	1.531e-2	1.073e-1	3.668e-2	7.383e-1	9.564e-2
2	1.346e-1	1.103e-1	2.688e-1	3.929e-2	4.097e-2	8.928e-2	5.126e-2	3.249e-1	1.708e-1
3	3.564e-2	1.556e-1	3.277e-2	5.591e-2	1.907e-1	4.254e-2	4.983e-3	1.361e-3	2.470e-1
4	1.354e-1	6.649e-2	1.018e-1	1.736e-1	4.203e-2	9.166e-3	4.907e-2	3.129e-2	1.278e-2
5	4.316e-3	2.296e-1	3.215e-2	9.013e-2	6.403e-2	7.545e-2	8.186e-3	1.508e-2	5.714e-2
6	1.862e-2	5.818e-2	1.594e-3	1.504e-1	5.499e-2	7.606e-2	1.397e-1	4.484e-2	3.997e-2
7	1.122e-1	1.173e-2	6.156e-2	7.971e-2	7.701e-3	7.682e-2	9.809e-2	1.751e-2	8.651e-2
8	1.012e-1	2.059e-2	9.182e-2	3.437e-2	5.015e-2	1.209e-2	7.337e-2	7.725e-2	7.207e-2
9	1.034e-1	6.315e-3	1.821e-1	1.196e-1	2.248e-2	4.765e-2	1.439e-1	1.504e-1	2.233e-2
10	1.863e-2	4.172e-2	5.736e-2	1.137e-2	2.724e-3	4.354e-2	1.066e-1	1.338e-2	3.226e-2

TABLEAU 11
Valeurs de EMC pour $n_s \in]1000, 3000[$.

S/R	2	3	4	5	6	7	8	9	10
1	3.643e-4	2.856e-2	2.818e-2	1.784e-2	2.720e-2	9.270e-3	2.117e-3	2.466e-2	6.808e-3
2	5.777e-3	1.629e-2	3.887e-3	8.571e-3	3.637e-3	1.989e-2	7.520e-4	1.614e-2	2.839e-3
3	2.108e-3	2.901e-3	2.383e-3	4.995e-4	6.700e-3	6.573e-4	1.101e-2	6.698e-3	3.923e-3
4	8.725e-3	2.323e-3	1.723e-3	9.243e-4	6.826e-3	4.683e-4	4.377e-3	4.849e-4	9.939e-3
5	6.571e-3	5.984e-3	4.482e-3	4.655e-3	1.173e-3	2.954e-3	3.996e-3	4.507e-4	2.726e-3
6	2.167e-3	7.422e-3	1.556e-3	8.796e-3	3.866e-4	6.370e-3	2.710e-3	5.346e-3	5.013e-3
7	1.674e-3	5.131e-4	2.960e-3	1.236e-3	1.381e-2	9.190e-3	2.399e-3	6.367e-4	8.007e-3
8	1.950e-3	9.858e-3	5.267e-4	3.522e-3	2.361e-3	1.026e-3	4.723e-3	4.985e-3	3.846e-3
9	9.869e-4	5.584e-3	8.248e-3	9.781e-4	3.626e-3	5.889e-3	2.667e-3	8.648e-4	2.151e-3
10	1.920e-3	2.858e-3	7.850e-3	4.333e-3	6.714e-5	9.285e-4	3.228e-4	1.814e-3	1.598e-3

TABLEAU 12
Valeurs de EQMC pour $n_s \in]0, 30[$.

S/R	2	3	4	5	6	7	8	9	10
1	1.315e-4	6.501e-4	1.873e-1	6.821e-3	2.343e-4	1.151e-2	1.345e-3	5.450e-1	9.147e-3
2	2.031e-2	1.282e-2	1.550e-1	7.661e-3	1.705e-3	1.129e-2	1.102e-2	2.846e-1	3.104e-2
3	2.871e-2	6.841e-2	3.149e-2	1.109e-2	8.077e-2	2.057e-2	3.156e-2	8.861e-3	1.565e-1
4	8.049e-2	1.561e-2	4.747e-2	6.469e-2	1.908e-2	1.753e-2	3.084e-2	7.883e-3	4.494e-3
5	2.505e-3	1.147e-1	1.525e-2	2.960e-2	9.892e-3	1.207e-1	1.860e-3	1.281e-2	4.845e-2
6	1.869e-2	4.478e-2	1.805e-3	7.734e-2	5.667e-2	7.960e-2	3.936e-2	2.602e-2	8.418e-2
7	6.582e-2	1.827e-2	2.495e-2	8.567e-2	3.131e-2	3.218e-2	3.634e-2	7.098e-3	4.656e-2
8	3.227e-2	2.201e-3	3.174e-2	1.206e-2	7.312e-2	5.255e-3	1.165e-2	8.066e-2	1.221e-2
9	1.066e-1	5.492e-2	1.472e-1	6.402e-2	1.482e-2	2.739e-2	1.250e-1	4.800e-2	2.331e-2
10	2.209e-2	1.475e-2	1.819e-2	9.020e-3	2.016e-2	3.946e-2	6.558e-2	7.523e-3	2.303e-2

TABLEAU 13
Valeurs de EQMC pour $n_s \in]1000, 3000[$.

S/R	2	3	4	5	6	7	8	9	10
1	1.327e-7	8.156e-4	7.941e-4	3.182e-4	7.397e-4	8.593e-5	4.482e-6	6.079e-4	4.635e-5
2	4.801e-5	4.124e-4	5.512e-5	4.185e-4	1.727e-5	4.051e-4	3.988e-6	2.767e-4	1.792e-4
3	1.348e-5	3.752e-4	3.936e-5	5.443e-5	1.173e-4	1.691e-4	3.889e-4	1.165e-4	2.089e-4
4	6.198e-4	2.019e-5	9.667e-5	6.508e-5	5.416e-5	2.711e-5	6.954e-5	6.220e-5	1.009e-4
5	1.043e-4	7.047e-5	8.137e-5	2.580e-4	1.192e-4	1.006e-4	1.674e-4	1.426e-4	8.175e-5
6	3.132e-5	2.407e-4	6.797e-5	2.159e-4	1.741e-4	1.196e-4	1.280e-4	2.991e-4	2.411e-4
7	2.096e-4	2.914e-4	1.466e-4	1.531e-4	3.350e-4	4.119e-4	1.263e-4	6.231e-5	1.679e-4
8	4.351e-4	2.342e-4	2.074e-4	2.066e-4	3.829e-4	1.925e-4	1.972e-4	2.418e-4	9.696e-5
9	3.502e-4	2.687e-4	2.367e-4	2.632e-4	3.385e-4	1.640e-4	9.895e-5	6.178e-5	7.156e-5
10	8.210e-5	2.082e-4	1.573e-4	1.601e-4	1.123e-4	1.018e-4	8.255e-5	2.191e-4	7.166e-5

5.4. Estimation et convergence de l'effet commun de la mesure

TABLEAU 14
Valeur de $|\hat{\theta} - \frac{1}{2}|$ pour $S = 4$.

R	$T = 1$	$T = 2$	$T = 4$	$T = 6$	$T = 8$	$T = 9$	$T = 10$
2	3.094e-2	4.394e-3	3.077e-2	2.920e-2	2.193e-2	1.792e-2	1.764e-2
3	8.965e-2	4.111e-2	4.961e-3	1.279e-2	8.748e-3	1.022e-2	9.765e-3
4	1.440e-1	8.236e-4	2.659e-3	2.079e-2	1.829e-2	2.048e-2	1.860e-2
5	1.990e-1	4.665e-2	3.570e-2	7.751e-3	2.235e-5	2.907e-3	5.794e-3
6	2.237e-1	3.798e-2	9.156e-3	9.118e-3	1.311e-2	1.763e-2	1.724e-2
7	2.466e-1	7.232e-2	2.417e-2	1.413e-2	1.754e-2	1.130e-2	1.049e-2
8	3.006e-2	4.634e-2	1.461e-2	2.392e-2	1.638e-2	1.991e-2	1.717e-2
9	2.562e-2	1.688e-1	4.343e-4	4.114e-3	1.388e-2	1.140e-2	1.118e-2
10	1.771e-1	5.026e-2	3.456e-2	1.711e-2	1.761e-2	2.150e-2	1.648e-2

TABLEAU 15
Valeur de $|\hat{\theta} - \frac{1}{2}|$ pour $R = 4$.

S	T = 1	T = 2	T = 4	T = 6	T = 8	T = 9	T = 10
1	5.000e-1	5.686e-2	1.236e-2	1.439e-2	2.832e-2	3.703e-2	3.562e-2
2	3.388e-1	7.728e-2	2.211e-4	8.970e-3	3.814e-3	8.832e-3	9.407e-3
3	8.473e-2	3.753e-2	3.694e-2	2.320e-2	2.955e-2	2.993e-2	2.886e-2
4	1.440e-1	8.236e-4	2.659e-3	2.079e-2	1.829e-2	2.048e-2	1.860e-2
5	9.641e-2	4.448e-2	2.952e-2	1.520e-3	2.829e-3	4.496e-3	3.655e-3
6	1.861e-2	1.474e-2	1.251e-2	1.785e-3	4.581e-3	4.886e-3	4.831e-3
7	1.935e-1	2.695e-2	5.266e-2	2.582e-2	1.693e-2	1.406e-2	1.251e-2
8	1.744e-1	8.770e-2	1.661e-3	4.951e-3	1.905e-3	1.468e-3	6.916e-4
9	4.464e-2	3.109e-2	4.067e-3	2.074e-3	8.238e-4	5.313e-3	2.534e-3
10	7.171e-2	2.162e-2	9.629e-3	9.504e-4	1.859e-3	4.166e-3	5.709e-3

5.5. Analyse générale des résultats

Le tableau 3 (resp. 4) donne les différentes valeurs du vecteur $\hat{\Theta}$ solution des équations de vraisemblance pour $S = 4$ et $R = 2$ (resp. $R = 3$) fixés. La première colonne de chacun des deux tableaux donne la vraie valeur Θ^o du paramètre. Les colonnes suivantes sont les solutions obtenues pour n_s appartenant à 7 tranches (colonne T : $T = 1, \dots, T = 10$) d'accidents. Ces solutions appartiennent à \mathbb{R}^9 pour le tableau 3 (resp. \mathbb{R}^{13} pour le tableau 4). On constate une décroissance de l'ordre de grandeur de $EQM(\hat{\Theta}, \Theta^o)$ lorsque les effectifs n_s passent de $]0, 30[$ à $]11000, 13000[$. Pour les tableaux 3 et 4 $EQM(\hat{\Theta}, \Theta^o)$ passe de $O(10^{-2})$ à $O(10^{-5})$. Le tableau 5 complète cette première étude de $EQM(\hat{\Theta}, \Theta^o)$ pour l'ensemble des 10 tranches (colonne : T) d'accidents et pour des valeurs de S et de R relativement élevées. L'ordre de grandeur atteint 10^{-6} pour la deuxième ($S = 5; R = 10$) et quatrième colonne ($S = 20; R = 10$) du tableau 5 où l'espace des paramètres est inclus dans \mathbb{R}^{51} dans le premier cas et dans \mathbb{R}^{201} dans le second cas.

Une deuxième étude a été mise en œuvre afin d'apprécier le comportement de $EQM(\hat{\Theta}, \Theta^o)$ pour des effectifs d'accidents faibles et élevés. On a alors fixé la valeur de n_s dans une tranche donnée et fait varier le nombre de sites S et de types R d'accidents. Les tableaux 6 (resp. 7) donnent des valeurs de $EQM(\hat{\Theta}, \Theta^o)$ de l'ordre de 10^{-2} (resp. 10^{-5}). Ils permettent d'observer une amélioration notable de l'estimation et de la convergence lorsque les effectifs passent de $]0, 30[$ à $]1000; 3000[$ c'est-à-dire lorsque les effectifs d'accidents par site augmentent. Cette convergence de $EQM(\hat{\Theta}, \Theta^o)$ vers 0 quand augmentent les effectifs d'accidents est visible sur les figures 1 à 3 obtenues pour différentes valeurs fixées de S et R . Les graphes obtenus pour les premières valeurs de S et de R montrent bien que $EQM(\hat{\Theta}, \Theta^o)$ est approximativement de 10^{-2} pour les $n_s \in]0, 30[$ et de 10^{-5} pour les $n_s \in]11000, 13000[$.

ESTIMATION MULTIDIMENSIONNELLE DES CONTRÔLES MOYEN ET...

99

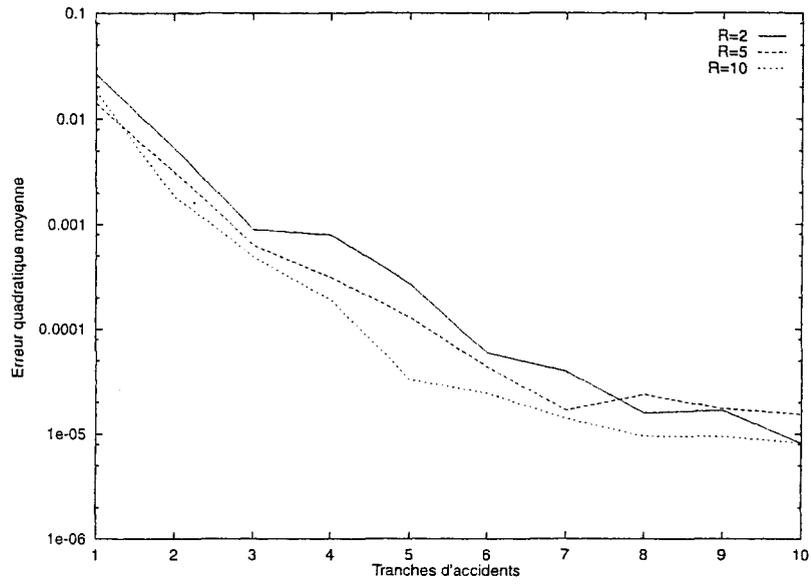


FIGURE 1
S = 5

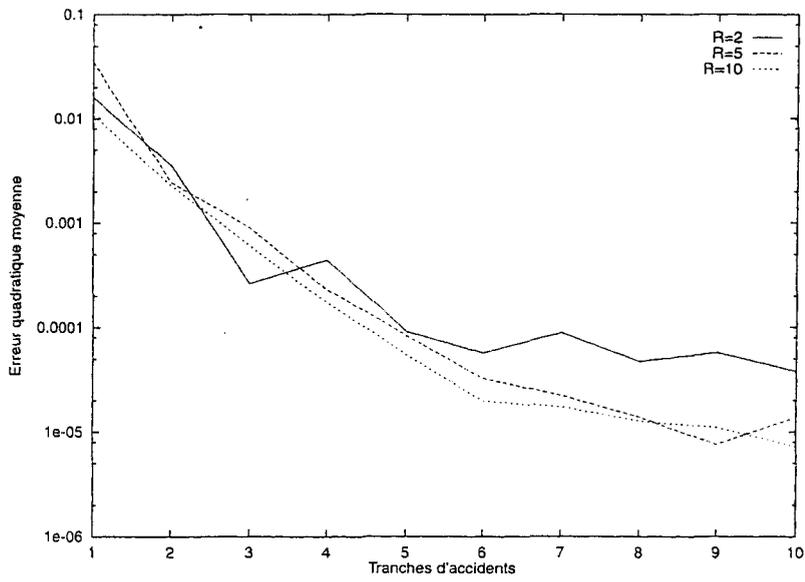


FIGURE 2
S = 10

100

A. N'GUESSAN, A. ESSAI, C. LANGRAND

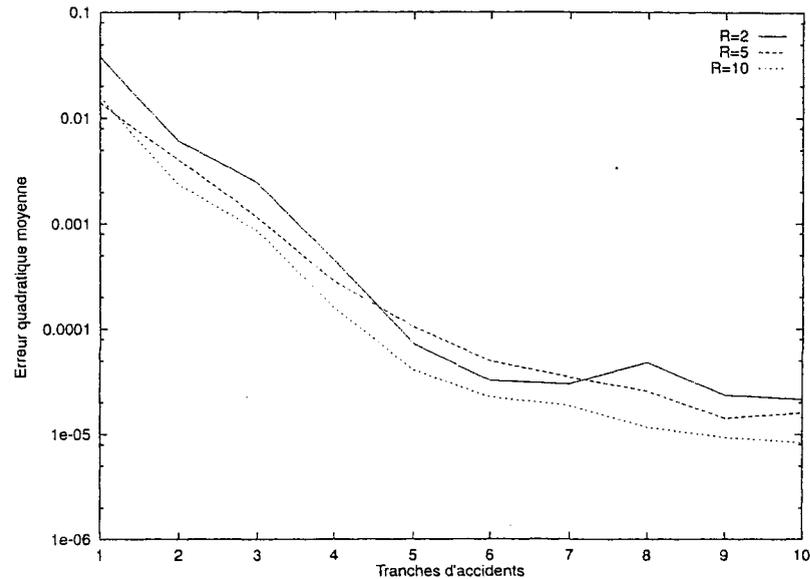


FIGURE 3
 $S = 20$

En parallèle à cet ensemble d'études, nous avons analysé les résultats obtenus pour les coefficients moyens de contrôle (espérance mathématique des v.a. C_s) et pour l'effet commun de la mesure. Les résultats sont consignés dans les tableaux 8 à 15. Le tableau 8 (resp.9) est lié au tableau 3 (resp. 4). Ainsi les valeurs de c_s^o et \hat{c}_s données dans ces tableaux sont obtenues en utilisant la formule (8) et les éléments de Θ^o et de $\hat{\Theta}$ du tableau 3 (resp. 4) sauf la première composante. On remarque, par exemple pour le tableau 8, que les valeurs de \hat{c}_s sont d'autant plus proches des vraies valeurs c_s^o que les effectifs sont élevés (colonne $T = 4, \dots, T = 10$). Cette convergence est mesurée par les critères EMC (Erreur Moyenne de Contrôle) et EQMC (Erreur Quadratique Moyenne de Contrôle) dont l'ordre de grandeur respectif varie de 10^{-1} (pour $n_s \in]0, 30[$) à 10^{-3} (pour $n_s \in]11000, 13000[$) et de 10^{-2} (pour $n_s \in]0, 30[$) à 10^{-6} (pour $n_s \in]11000, 13000[$). On obtient même des approximations de l'ordre de 10^{-5} et 10^{-6} lorsque que R passe de 2 à 3 (tableau 9). Quant aux tableaux 10 et 12 (resp. 11 et 13), ils vont de paire avec le tableau 6 (resp. 7). Ces quatre tableaux permettent d'apprécier une assez bonne stabilité de l'ordre de grandeur des deux critères EMC et EQMC. Enfin les tableaux 14 et 15 donnent l'estimation de l'effet commun de la mesure c'est-à-dire la première composante de $\hat{\Theta}$ et complètent les résultats des tableaux 3 et 4. On note dans les tableaux 3 et 4 que plus les effectifs d'accidents augmentent plus $\hat{\theta}$ tend vers 0.5 la vraie valeur de l'effet commun. Cette convergence est confirmée par la stabilité de l'ordre de grandeur des valeurs des tableaux 14 et 15, la valeur absolue de l'écart entre l'estimation et la vraie valeur atteint quelques fois 10^{-4} .

6. Remarques et conclusion

Selon certains points de vue et certaines pratiques dans le domaine de la sécurité routière (Lassarre 1977 [9], Danielson 1986 [3] et 1988 [4], Hauer 1997 [7]), il est souhaitable de modéliser les données d'accidents à l'aide de distribution de Poisson. Nous avons adopté dans ce travail la philosophie de Tanner. Les difficultés dans l'approche de modélisation de ce dernier résident essentiellement dans le choix de la zone de contrôle et dans l'hypothèse « non aléatoire » qu'il fait sur les coefficients de contrôle. Hauer (1991) [6] a proposé une notion de similarité pour choisir un site de comparaison. Néanmoins, il semble assez difficile de baser le choix d'un site de comparaison sur cette seule notion de similarité. Il est essentiel que le nombre d'accidents du site supposé de comparaison ou de contrôle soit suffisamment élevé afin que le phénomène aléatoire inhérent au comptage des accidents sur ce site ne dégrade pas la précision de l'estimation de l'effet de la mesure de sécurité. Notre approche intègre ce phénomène aléatoire sur le site de contrôle en imposant que les coefficients associés à ce site proviennent d'une distribution de probabilité estimée simultanément avec l'effet commun de la mesure.

Nos résultats généralisent l'approche de Tanner par le fait que les coefficients de contrôle sont distribués de façon aléatoire entre les types d'accidents et complètent ceux de N'Guessan (1993) [12], N'Guessan et Langrand (1992) [14]. L'inconvénient de la méthode est que la dimension de l'espace des paramètres croît avec le nombre de sites et de types d'accidents. Cependant, les résultats montrent que l'erreur quadratique moyenne est un $O(10^{-\beta})$ avec $\beta > 1$.

Bibliographie

- [1] AITCHISON J., SILVEY S.D.(1958), *Maximum likelihood estimation of parameters subject to restraints*, Annals of mathematical statistics, 29, 813–829.
- [2] CROWDER M.(1984), *On the constrained maximum likelihood estimation with non i.i.d. observations*, Ann. Inst. Statist. Math.36, A, 239–249.
- [3] DANIELSON S. (1986), *A comparison of two methods for estimating of the effect of a countermeasure in the presence of regression effects*, Accid. Anal. and Prev. 18, n°1, 13–23.
- [4] DANIELSON S. (1988), *Estimation of the effects of countermeasures on differents types of accidents in the presence of regression effects*, Accid. Anal. and Prev. 20, n°4, 289–298.
- [5] GOURIÉROUX CH., MONFORT A. (1989), *Statistique et modèles économétriques : vol n°1*, Edition Economica.
- [6] HAUER E. (1991), *Comparison Groups in road safety studies : An analysis*, Accid. Anal. and Prev. 23, n°26, 609–622.
- [7] HAUER E. (1997), *Observational before-after Studies in Road Safety : Estimating the effect of highway and traffic engineering measures on road safety*, Pergamon Edition.

- [8] LASSARRE S.(1985), *Panorama des méthodes associées à l'évaluation de type avant-après*, Colloque international sur l'évaluation des mesures locales de sécurité routière. Paris.
- [9] LASSARRE S.(1977), *A propos des tests statistiques, sur variables poissonniennes, utilisées dans le domaine de la sécurité routière*, Revue de Statistique Appliquée XXV, n°3.
- [10] NAG, (2000) *Numerical Algorithms Groups*, C.R.I. Université de Lille 1.
- [11] NEUENSCHWANDER B.E., FLURY B.D. (1997), *A note on Silvey's (1959) theorem*, Statistics and Probability Letters 36, 307–317.
- [12] N'GUESSAN A. (1993), *Contribution à l'analyse statistique d'une mesure de sécurité routière*, Thèse de Doctorat Université de Lille 1. France.
- [13] N'GUESSAN A., LANGRAND C. (1993), *Sur la distribution asymptotique de certaines statistiques utilisées dans le domaine de la sécurité routière*, C.R. Acad. Sci, t.317, Série I, 401–404.
- [14] N'GUESSAN A., LANGRAND C. (1992), *Une méthode d'estimation de l'effet moyen d'une mesure de sécurité routière*, XXIV^{èmes} Journées de Statistique de l'ASU, Bruxelles.
- [15] TANNER J.C. (1958), *A problem in the combination of accident frequencies*, Biometrika, 45, 331–342.

Annexe A06 :

Proceedings of the Conference on Information and Management Sciences, Eldon Y. Li Series Editor, California Polytechnic State University, Vol. 4, 243-252 (2005)

A Problem in the Multiple combination of Accident Frequencies Informations with Applications

Assi N'Guessan , Claude Langrand

**A Problem in the Multiple Combination of Accident
Frequencies Informations with Applications**

*Proceedings of the Fourth International Conference on Information and Management
Sciences*, Series Editor: Eldon Y. Li, California Polytechnic State University,
Vol. 4 (2005) pp. 243 – 252.

A Problem in the Multiple Combination of Accident Frequencies Informations with Application

Assi N'Guessan^a, Claude Langrand^{a,b}

(a) Laboratoire Paul Painlevé CNRS, UMR 8542 - Université de Lille 1 59655 Villeneuve d'Ascq Cédex, France

(b) UFR de mathématiques pures et appliquées - Université de Lille 1 59655 Villeneuve d'Ascq Cédex, France

assi.nguessan@polytech-lille.fr, claude.langrand@univ-lille1.fr

Abstract : We consider a multidimensional combination of road accident frequencies before and after the introduction of a road safety measure at different experimental sites with a control site for each of them. Each experimental site counts several mutually exclusive types of accidents over a two fixed periods (before and after) of time. A multinomial distribution is proposed to model the total accident number in each experimental site. An estimation method of the parameters and their asymptotic standard errors is suggested. Some simulated accidents data allow us to solve numerically the constrained maximum likelihood equations and to study the existence and the convergence of the estimator via the mean square error. We then reanalyse a set data concerning the effects of installing approximatively circular roundabouts at cross-roads.

Keywords: Road Safety Measure, Accident data, Poisson model, Conditional Poisson model, Multinomial model, Constrained Maximum Likelihood, Risk accident, Mean square error, Before-after Estimation, Simulation.

1 Introduction

We consider a statistical accident frequencies modelling problem when a road safety measure (crossroad lay-out, surface of a motorway section, etc) is simultaneously applied on different experimental sites. Each experimental site counts several mutually exclusive types of accidents (fatal accidents, seriously injured people, slightly injured people, material damage only) over two fixed periods (before and after) of time. Then we are interested in the modelling and the estimation of the average effect of the measure for the whole set of experimental sites and for the whole set of accident types. However one knows that some external factors (as traffic-flow and speed variations before and after the applied measure, weather conditions, experimental sites location, etc) have a direct or indirect influence on the road accident frequencies before and after the

introduction of a road safety measure (or change). We therefore have to include the impact of these factors into the statistical models used to analyse the average effect of the measure if we want to interpret correctly the scope of this road safety measure at all experimental sites and all types of accident. The question is which approach and model we must use if we want to take into account the impact of these external factors on the applied measure average effect.

The statistical models and methods used to overcome this problem heavily depend on the data and on the experimental set along with the measure (*see for instance* Hauer 1991, 1997, Barry et al 2000, Beenstock et al 2001, Abdel-Aty et al 2000, Lassarre 2001). According to some points of view and some practices in the field of road safety, it is advisable to model the accident data according to Poisson distribution (*see for example* Lassarre 1977, Hauer 1997), conditional or truncated Poisson distribution (for example Danielson 1986 and 1988, Hauer 1997), negative binomial distribution (Hauer 2001, Allain and Brenac 2001). Even if Poisson model or conditional Poisson model is a common and convenient assumption in accident count analysis, are accidents Poisson distributed? It is very important to evaluate this Poisson assumption in practice using statistical tests (*see for example* Nicholson 1985 and 1986, Nicholson and Wong 1993).

In this paper we have adopted Tanner(1958)'s before-after philosophy. In fact, this author introduced the notion of control area and defined a control coefficient for each experimental site using the ratio of the accident number after to before in the control area. Thereafter, he proposed a binomial model to combine accident frequencies. The difficulties in that modelling approach mainly lie in the choice of the control area and in the 'free from error' assumption he makes about control coefficients. Hauer (1991) suggested a notion of similarity to choose a comparison site. However, it seems rather difficult to ground the choice of a comparison site only on this notion of similarity. It is vital that the

accident number on the supposed comparison site or control site should be important enough whatever the accident types for the random phenomenon inherent to accident counting on this site not to impair the accuracy of the estimation of the safety measure's average effect.

With that in mind, N'Guessan et al (2001) and N'Guessan (2003b) generalised Tanner's approach to before-after tables of accident with different control area ratios for each site. In their papers they supposed that the same road safety measure had been applied to several sites, each one including mutually exclusive different accident types and linked to a specific control area. N'Guessan et al (2001) are interested in the multidimensional estimation of the mean control and the mean effect of a measure on different sites and different types of accidents. The latter authors deal with the maximum likelihood estimation of both the average effect of the road safety measure and the mean control ratios of each control area via simulation studies and mean square error for wide before-after data tables. But they don't include the estimation of the probability distributions connected to their model and used to define each mean control ratio. Our approach completes N'Guessan et al's (2001) model and method through the use of the conditional Poisson model and through the estimation of the probability distributions (interpreted as accident risks) used to define the mean control ratio connected to each control site.

This material is organised as follows. Section 2 is devoted to the observed accident data frequencies and the main purpose of the paper. Section 3 deals with the statistical pattern used to combine simultaneously the accident data tables of experimental sites and control sites. We build a multinomial model for each experimental site taking into account the periods, the different accident types, accident data from experimental sites and from control sites, and the parameters to be estimated. The model obtained is based on the conditional Poisson model. Section 4 presents the estimation method for the unknown vector of parameters (i.e. the measure's average effect and accident risks in control sites) based on the constrained maximum likelihood estimation. We also propose a method based on Schur complement approach in order to reach the asymptotic standard errors of the components of the vector of parameters. Section 5 is exclusively centred on the numerical results. We first study the numerical existence and convergence of a solution for the constrained maximum likelihood equations using different simulated accident data tables. Thereafter we apply our method to a set of real data provided from Tanner's (1958) results. The constrained log-likelihood function and its first derivatives, the constrained Fisher

information matrix and the Schur complement estimation of the standard errors are displayed in the appendix A and B.

2 Data and Notations

We consider a multidimensional combination of road accident frequencies before and after the introduction of a road safety measure (crossroad lay-out, surface of a motorway section, etc) at $S(S > 0)$ experimental sites. Each site counts $R(R > 1)$ accident types (fatal accidents, seriously injured people, slightly injured people, material damage only) over a fixed period of time. We suppose that the total accident number is known for each experimental site and the available data are given as follows :

	type 1	...	type j	...	type R
site 1	x_{111}	...	x_{1j1}	...	x_{1R1}
⋮	⋮		⋮		⋮
site k	x_{11k}	...	x_{1jk}	...	x_{1Rk}
⋮	⋮		⋮		⋮
site S	x_{11S}	...	x_{1jS}	...	x_{1RS}

table 1a : before

where x_{1jk} (resp. x_{2jk}) stands for the number of type j of accidents on site k before (resp. after) the setting up of the measure.

	type 1	...	type j	...	type R
site 1	x_{211}	...	x_{2j1}	...	x_{2R1}
⋮	⋮		⋮		⋮
site k	x_{21k}	...	x_{2jk}	...	x_{2Rk}
⋮	⋮		⋮		⋮
site S	x_{21S}	...	x_{2jS}	...	x_{2RS}

table 1b : after

In order to take some external factors into account (such as traffic flow, speed limit variation, weather conditions,...), each experimental site is matched to a control site where the safety measure was not applied. We further assume that the accident data for control sites, over the same time period, are fixed and known and are given by the following control coefficients table :

	type 1	...	type j	...	type R
control 1	z_{11}	...	z_{j1}	...	z_{R1}
⋮	⋮		⋮		⋮
control k	z_{1k}	...	z_{jk}	...	z_{Rk}
⋮	⋮		⋮		⋮
control S	z_{1S}	...	z_{jS}	...	z_{RS}

table 2 : Control sites

where z_{jk} denotes the ratio of the number of accidents of type j for the period after to the period before in the control site k and uses a version of a control ratio defined by Tanner (1958). One then has to combine accident frequencies from all the whole set of experimental sites and from the whole set of types including the ratios of the number of accidents of the whole set of control sites. In the following we note X_{1jk} (resp. X_{2jk}) the random variable giving the type j accident number on experimental site k in the period before (resp. after).

3 Statistical pattern

For a fixed experimental site k , we then are interested mainly in the joint probability distribution of the $2R$ -dimensional random vector (X_{1k}, X_{2k}) , given the observed total number of accidents recorded at experimental site k , where $X_{1k} = (X_{11k}, X_{12k}, \dots, X_{1Rk})^T$ (resp. $X_{2k} = (X_{21k}, X_{22k}, \dots, X_{2Rk})^T$) is the random vector giving the R accident types number on experimental site k before (resp. after) the introduction of the safety measure (or change). The construction of this joint probability distribution is based on a forecast principle used in the modelling of a road safety measure's average effect. Actually, when a road safety measure is applied, we suppose that it has a multiplicative effect (see for instance Oppe 1979, Hauer et al. 1991) on the expected numbers of accidents whatever the site and the type of accident may be. So, for fixed experimental k , we assume that $(X_{11k}, \dots, X_{1Rk}, X_{21k}, \dots, X_{2Rk})$ is a sequence of independent random variables, where X_{1jk} (resp. X_{2jk}) is Poisson distributed with mean λ_{1jk} (resp. λ_{2jk}), with

$$\lambda_{2jk} = \beta_0 z_{jk} \lambda_{1jk}, \quad (j = 1, 2, \dots, R) \quad (1)$$

β_0 ($\beta_0 > 0$) is the unknown parameter average effect of the road safety for all experimental sites and all accident types. Then, for fixed experimental site k and using the conditional Poisson distribution (conditional to the observed total number of accidents in both the before and the after periods) property, the joint probability distribution of the $2R$ -dimensional random vector (X_{1k}, X_{2k}) is given by the multinomial distribution denoted $\mathcal{M}(n_k; \Pi_{1k}(\Theta), \Pi_{2k}(\Theta))$, where n_k ($0 < n_k = \sum_{t=1}^2 \sum_{m=1}^R x_{tmk}$) is the observed total accident number in experimental site k ,

$$\Pi_{tk}(\Theta) = (\pi_{t1k}(\Theta), \pi_{t2k}(\Theta), \dots, \pi_{tRk}(\Theta)) \quad (2)$$

is the $R \times 1$ vector of cell probabilities with component $\pi_{tjk} : \mathbb{R}^{1+SR} \rightarrow]0, 1[$ given as follows :

$$\begin{aligned} \pi_{1jk}(\Theta) &= \frac{p_{jk}}{1 + \beta_0 \sum_{m=1}^R z_{mk} p_{mk}}; \\ \pi_{2jk}(\Theta) &= \frac{\beta_0 z_{jk} p_{jk}}{1 + \beta_0 \sum_{m=1}^R z_{mk} p_{mk}}. \end{aligned} \quad (3)$$

($j = 1, 2, \dots, R$), and where

$\Theta = (\beta_0, P^T)^T$ is the $(1 + SR) \times 1$ unknown vector of parameters with β_0 the average effect, $P = (P_1^T, P_2^T, \dots, P_S^T)^T$ with $P_k = (p_{1k}, p_{2k}, \dots, p_{Rk})^T$ a $R \times 1$ vector of probabilities connected to the control site k such that $P_k \in S_k^{(R-1)}$ with

$$S_k^{(R-1)} = \{P_k \in \mathbb{R}^R, p_{jk} > 0, \sum_{j=1}^R p_{jk} = 1\}. \quad (4)$$

Likewise, the fact that the sub-vector P_k belongs to the simplex of order R means that the vector of parameters Θ has the following restraints ($0 < p_{jk} < 1$) and the linear constraints

$$h_k(\Theta) = 0; h_k(\Theta) = 1 - \sum_{j=1}^R p_{jk}, \quad (5)$$

with ($k = 1, 2, \dots, S$). Drawing our inspiration from the constraints and the structure of cell probabilities $\pi_{tjk}(\Theta)$, we then define ζ_k the random variable taking values $z_{1k}, z_{2k}, \dots, z_{Rk}$ such that, for fixed control site k ,

$$p_{jk} = \text{Proba}[\zeta_k = z_{jk}]; \sum_{j=1}^R p_{jk} = 1. \quad (6)$$

So the main purpose of this paper is to estimate both the safety measure average effect (β_0) for the whole set of experimental sites and for the whole set of accident types and the accident risks ($p_{jk}, j = 1, 2, \dots, R; k = 1, 2, \dots, S$) including the estimation of their asymptotic standard errors. One then has to combine accident data from experimental sites (tables 1a, 1b) and from control sites (table 2) to efficiently estimate both parameters and their asymptotic standard errors.

4 Constrained Maximum likelihood estimation

4.1 Joint estimation of the average effect and the accident risks

Through construction of the pattern, two types of parameters can be spotted out among those linked to S target sites and R accident types : β_0 , the interest parameter, and the set of auxiliary parameters $P_k = (p_{1k}, p_{2k}, \dots, p_{Rk})^T, k = 1, 2, \dots, S$. As this paper unfolds, one focuses on the simultaneous estimation of those parameters and one then poses $\Theta = (\beta_0, P_1^T, P_2^T, \dots, P_S^T)^T$, the element from \mathbb{R}^{1+SR} thus defined (or actually from $\mathbb{R}^{1+S(R-1)}$ since $\sum_{j=1}^R p_{jk} = 1$.) We then assume that the sequence of random vectors $(X_{11}, X_{21}), (X_{12}, X_{22}), \dots, (X_{1S}, X_{2S})$ are independent

and follow a multinomial distribution whose form is given above. So according to tables 1a, 1b and 2 and the log-likelihood function linked displayed in the appendix, we show that the constrained maximum likelihood estimator $\hat{\Theta}$ of Θ , if existing, is the solution to the system of non-linear equations :

$$\begin{cases} \sum_{k=1}^S \sum_{j=1}^R \frac{x_{2jk} - \hat{\beta}_0 \hat{c}_k x_{1jk}}{1 + \hat{\beta}_0 \hat{c}_k} = 0; \\ x_{jk} - \frac{n_k \hat{p}_{jk} (\hat{\beta}_0 z_{jk} + 1)}{1 + \hat{\beta}_0 \hat{c}_k} = 0, \\ (k = 1, 2, \dots, S; \quad j = 1, 2, \dots, R); \\ \hat{c}_k = \sum_{j=1}^R z_{jk} \hat{p}_{jk}, \quad (k = 1, 2, \dots, S); \end{cases} \quad (7)$$

under the constraints

$$\begin{cases} \hat{\beta}_0 > 0; \quad 0 < \hat{p}_{jk} < 1; \\ \sum_{j=1}^R \hat{p}_{jk} = 1, \quad (k = 1, 2, \dots, S). \end{cases} \quad (8)$$

Let's note that $\hat{\Theta}$, the solution to the system above, is clearly a maximum likelihood estimation method under restraints (RMLE) which is an approach to estimation that maximizes the likelihood over a restricted space. This is not the place to study theoretical sufficient conditions for existence and convergence of $\hat{\Theta}$. The general framework for RMLE is well-known and is not discussed here (see for instance Aitchison and Silvey 1958, Silvey 1959, Crowder 1984). N'Guessan et al (2001) numerically study the existence and convergence of $\hat{\Theta}$ via wide simulation results and graphic curves comparisons. A few part of their simulated results are displayed below. We add to those simulated results a real application provides from Tanner's (1958) data concerning the effects of installing circular roundabouts at cross-roads.

4.2 Asymptotic standard errors estimation using Schur complement

In order to reach the asymptotic standard errors of the components of the vector of parameters, we need the diagonal elements of the inverse of the matrix (see appendix for details)

$$\Gamma_{\Theta} = \begin{pmatrix} J_{\Theta} & H_{\Theta}^T \\ H_{\Theta} & O_{S,S} \end{pmatrix}, \quad (9)$$

where J_{Θ} , H_{Θ} , $O_{S,S}$ are, respectively, the $(1 + SR) \times (1 + SR)$ Fisher information matrix, the $S \times (1 + SR)$ matrix of constraints, the $S \times S$ matrix with all entries equal to zero. The general version of the latter matrix was extensively studied by Rao (1973, section 4i.1) and Rao and Yanai (1985). All the results on the inverse of this matrix are characterized under some rank conditions on the block matrices in the partitions. However one recalls that the size of matrix Γ_{Θ} increases very quickly with the number S of experimental sites and the number R of accident types. So to obtain the desired variances

and therefore the desired asymptotic standard errors, sophisticated numerical and statistical tools are needed to invert matrices whose dimensions grow rather rapidly. We propose another way to overcome this problem.

The approach we use in this paper is based on the notion of Schur complement (see Ouellette 1981, Zhang 2005) and provides from the recent results obtained by N'Guessan (2003a,b) and by N'Guessan and Langrand (2005a,b). This approach replaces the direct inversion of such a matrix with products/multiplications and additions of smaller-sizes matrices. It also enables to have the explicit analytical expression of the desired asymptotic standard errors. The main strength of the Schur complement approach proposed is its applicability to any values of number of experimental sites and of number of different types without being obliged to invert matrix Γ_{Θ} .

For example when the measure is applied to a single experimental site ($S = 1$) with different accident types $R(R > 1)$, then the asymptotic variances are given by (see appendix for details)

$$\sigma^2(\hat{\beta}_0) = [(J_{\Theta}/B_{\Theta}) + \frac{\gamma_k^2 t_k^2 \Delta_{\Theta,k}^2}{n_k^2 \|1_R\|_{B_{\Theta,k}^{-1}}^2}]^{-1} \quad (10)$$

and

$$\sigma^2(\hat{p}_{jk}) = \gamma_k^{-1} [\omega_{jk} + \omega_{jk}^2 (\theta_k \gamma_k z_{jk}^2 - A_{k,k} \gamma_k^{-1} y_{jk}^2)]; \quad (11)$$

$$j = 1, 2, \dots, R,$$

where

$$(J_{\Theta}/B_{\Theta}) = \frac{\gamma_k^2}{n_k} \left[\left(\frac{\sum_{m=1}^R z_{mk} p_{mk}}{\beta_0} - \frac{\gamma_k t_k}{n_k} \|Z_k\|_{\Omega_{\Theta,k}^{-1}}^2 \right) \right] \quad (12)$$

with $y_{jk} = 1 + \theta_k \gamma_k \Delta_{\Theta,k} z_{jk}$ and t_k , $\|Z_k\|_{\Omega_{\Theta,k}^{-1}}^2$, $\|1_R\|_{B_{\Theta,k}^{-1}}^2$, ω_{jk} , $A_{k,k}$, $\Delta_{\Theta,k}$, and θ_k best defined in appendix below. The explicit analytical expressions of the asymptotic standard errors of the measure's average effect and the accident risks are obtained via the square root of $\sigma^2(\hat{\beta}_0)$ and $\sigma^2(\hat{p}_{jk})$. Thereafter the asymptotic confidence intervals, at the $(1 - \alpha)$ confidence level, for the measure's effect and the accident risks are given by the following formal expressions

$$\begin{aligned} IC_{1-\alpha}(\beta_0) &= [LCI(\beta_0); UCI(\beta_0)]; \\ IC_{1-\alpha}(p_{jk}) &= [LCI(p_{jk}); UCI(p_{jk})] \quad (13) \\ &(j = 1, 2, \dots, R), \end{aligned}$$

where

$$LCI(\beta_0) = \max(0; \hat{\beta}_0 - \epsilon_{\alpha/2} \sigma(\hat{\beta}_0))$$

$$UCI(\beta_0) = \hat{\beta}_0 + \epsilon_{\alpha/2}\sigma(\hat{\beta}_0)$$

$$LCI(p_{jk}) = \max(0; \hat{p}_{jk} - \epsilon_{\alpha/2}\sigma(\hat{p}_{jk}))$$

$$UCI(p_{jk}) = \min(1; \hat{p}_{jk} + \epsilon_{\alpha/2}\sigma(\hat{p}_{jk}))$$

with $\epsilon_{\alpha/2}$ obtained through the standard normal law table.

5 Numerical results

The numerical results count two parts. The first part presents some few numerical simulated studies of the constrained likelihood equations above and therefore of the numerical existence and convergence of $\hat{\Theta}$ (see for example N'Guessan et al 2001 for wide simulated results). The second part concerns a numerical example provided from Tanner's (1958). We then re-analyze the data concerning the effects of installing approximately circular roundabouts at cross-roads.

5.1 Numerical simulated study : Existence and convergence

We look for a solution $\hat{\Theta}$ to the likelihood equations and then study the convergence of $\hat{\Theta}$ to Θ^0 using the analysis of the mean square error (MSE) which measures the adequacy between $\hat{\Theta}$ and Θ^0 and which is defined by :

$$MSE(\hat{\Theta}, \Theta^0) = \frac{1}{1 + RS} \sum_{m=1}^{1+RS} (\hat{\Theta}_m - \Theta_m^0)^2.$$

The numerical results are found for n_k belonging to the following accident intervals : $T_1 =]0, 30[$, $T_2 =]30, 50[$, $T_3 =]50, 300[$, $T_4 =]300, 1000[$, $T_5 =]1000, 3000[$, $T_6 =]3000, 5000[$, $T_7 =]5000, 7000[$, $T_8 =]7000, 9000[$, $T_9 =]9000, 11000[$, $T_{10} =]11000, 13000[$:

Table 3 gives the different values of $MSE(\hat{\Theta}, \Theta^0)$ for n_s belonging to the different accident intervals .

	$R = 2$	$R = 10$
T_1	$2.6850E - 2$	$1.8480E - 2$
T_2	$5.3214E - 3$	$1.8482E - 3$
T_3	$8.9237E - 4$	$4.8812E - 4$
T_4	$7.8695E - 4$	$1.9116E - 4$
T_5	$2.6904E - 4$	$3.2655E - 5$
T_6	$5.8576E - 5$	$2.4412E - 5$
T_7	$3.9579E - 5$	$1.4034E - 5$
T_8	$1.5870E - 5$	$9.5464E - 6$
T_9	$1.6832E - 5$	$9.4963E - 6$
T_{10}	$8.2177E - 6$	$8.2518E - 6$

table 3 : $S = 5$

	$R = 2$	$R = 10$
T_1	$3.8793E - 2$	$1.6054E - 2$
T_2	$6.0657E - 3$	$2.3454E - 3$
T_3	$2.4164E - 3$	$8.5019E - 4$
T_4	$4.5135E - 4$	$1.5882E - 4$
T_5	$7.2394E - 5$	$4.0908E - 5$
T_6	$3.2943E - 5$	$2.2612E - 5$
T_7	$3.0454E - 5$	$1.8778E - 5$
T_8	$4.8501E - 5$	$1.1726E - 5$
T_9	$2.3679E - 5$	$9.3814E - 6$
T_{10}	$2.1514E - 5$	$8.3590E - 6$

table 4 : $S = 20$

Table 4 completes this first study of $MSE(\hat{\Theta}, \Theta^0)$ for the whole set of the accident intervals and for relatively high values of S and R . The order of magnitude reaches 10^{-6} for table 4 where the space of the parameters is included in \mathbb{R}^{41} in the first case and in \mathbb{R}^{201} in the second case. Further studies was carried out to evaluate the performance of $MSE(\hat{\Theta}, \Theta^0)$ for small and important numbers of accidents. The value of n_s was then set in a given interval and S , the experimental site number, and R , the type of accidents, were made to vary. That enables us to note an important improvement in the estimation and in the convergence when accident numbers go from $]0, 30[$ to $]1000, 3000[$, i.e. when the observed accident numbers per site grow. Figure 1 below completes the convergence of the solution and shows a certain stability of the MSE for different values of S .

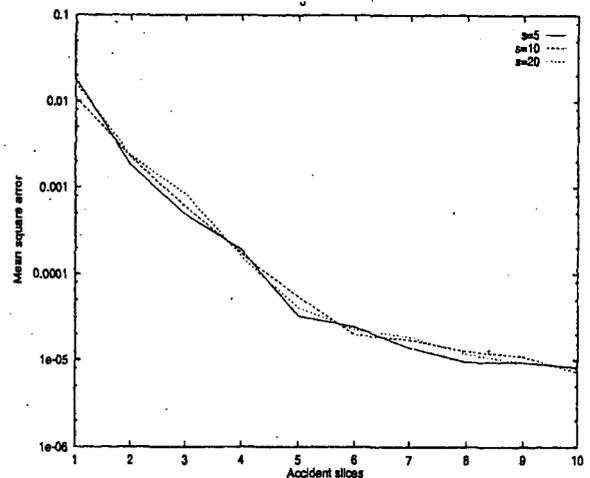


FIG. 1 - $R=10$.

5.2 Application : Tanner's (1958) data

Here we revisit Tanner's data. His data concern the effects of installing approximately circular roundabouts at seven sites of the same country. We transform the seven sites in seven types (column t_1 to t_7) of accident and the country in one site. So the data are as follows :

These solutions belong to \mathbb{R}^{11} and \mathbb{R}^{51} for table 3. We notice a decreasing order of magnitude of $MSE(\hat{\Theta}, \Theta^0)$ when numbers n_s (observed total accident numbers per experimental site) go from $]0, 30[$ to $]11000, 13000[$ i.e. the convergence of $\hat{\Theta}$ to Θ^0 .

Accident data : Before period

t_1	t_2	t_3	t_4	t_5	t_6	t_7
1	6	9	16	10	2	5

Accident data : After period

t_1	t_2	t_3	t_4	t_5	t_6	t_7
6	3	5	5	0	2	0

control data

t_1	t_2	t_3	t_4	t_5	t_6	t_7
1.04	1.25	1.11	2.36	1.13	1.69	1.61

Thus for type 1, there was 1 accident in the before period, 6 in the after period, while in the whole of the control site concerned, there were 4% more casualties in the after period than in the before period.

The estimation of the average effect and of the risks accident are given by (7). To solve this system, an initial value of Θ was inserted in a fortran program of NAG (2003) software of optimisation algorithms. So we get the value of parameters (column $\hat{\Theta}$) below. The standard errors (column S.E.) are obtained using Schur complement, expressions (10), (11) and Microsoft Excel. The average effect of changing crossroads to roundabouts is estimated at 0.2833, that is to say an efficiency of 71.67%. We obtain the same results than Tanner.

Estimation of parameters

parameter	Θ	$\hat{\Theta}$	S.E.
effect	β_0	0.2833	0.0800
risk 1	p_{1k}	0.1103	0.0389
risk 2	p_{2k}	0.1356	0.0417
risk 3	p_{3k}	0.2173	0.0524
risk 4	p_{4k}	0.2568	0.0465
risk 5	p_{5k}	0.1545	0.0449
risk 6	p_{6k}	0.0552	0.0267
risk 7	p_{7k}	0.0703	0.0295

With regard to the risks of accident in the control sites, we observe a high value (25.68%) for the control site number 4, fitting well with an increase over 136% of accidents in this site from the before period to the after period. On the other hand, the risks are less high in the sites 6 and 7, even though the variations of the accidents numbers for the periods before and after remain large. Using the standard errors (S.E.) and the expressions 12 and 13, we get 95% confidence intervals for the set of parameters of the model; these confidence intervals put forward more general conclusions about the effect and the efficiency of the change in consideration (crossroads to roundabouts). Thus, the interval [0.1265; 0.4401] contains, with a probability of 0.95, the true value of the average effect of the change. So, the real average efficiency vary from 55.99% to 87.35%, with a probability of 95%.

95% Confidence Intervals

parameter	Θ	LCI	UCI
effect	β_0	0.1265	0.4401
risk 1	p_{1k}	0.0340	0.1865
risk 2	p_{2k}	0.0538	0.2173
risk 3	p_{3k}	0.1145	0.3200
risk 4	p_{4k}	0.1656	0.3479
risk 5	p_{5k}	0.0664	0.2425
risk 6	p_{6k}	0.0028	0.1075
risk 7	p_{7k}	0.0124	0.1281

It's too possible to test the significance of the average effect of the change, using the confidence interval of the β_0 parameter : since $1 \notin [0.1265; 0.4401]$, we reject the hypotesis $H_0 : \beta_0 = 1$ against $H_1 : \beta_0 \neq 1$ with a risk of 5%, concluding therefore to the average efficiency of the change. Likewise, using the confidence intervals for p_{jk} , we can judge the risks of the 7 types of accident. We prove $H_0 : p_{jk} = 0$ against $H_1 : p_{jk} \neq 0$ for ($j = 1, 2, \dots, 7$). With a risk of 5%, we conclude that theses risks are not null, except may be, for the type number 6, where $LCI(p_{6k}) = 0.0028$.

Appendix A : Computing of the Constrained Maximum Likelihood Equations and of the information matrix

This appendix gives the recent main results concerning the Schur complement approach for computing subcovariance matrices (see N'Guessan and Langrand 2005a). Through construction of the pattern, two types of parameters can be spotted out among those linked to S experimental sites and R accident types : β_0 , the interest parameter, and the set of auxiliary parameters $P_k = (p_{1k}, p_{2k}, \dots, p_{Rk})^T, k = 1, 2, \dots, S$. As this paper unfolds, one focuses on the joint estimation of those parameters and one then poses $\Theta = (\beta_0, P_1^T, P_2^T, \dots, P_S^T)^T$, the element from \mathbb{R}^{1+SR} thus defined (or actually from $\mathbb{R}^{1+S(R-1)}$ since $\sum_{j=1}^R p_{jk} = 1$). We then assume that random vectors $(X_{11}, X_{21}), (X_{12}, X_{22}), \dots, (X_{1S}, X_{2S})$, are independent and follow a multinomial distribution whose form is given by (4). So according to tables 1a, 1b and 2, the likelihood function linked to the sequence of $(X_{11}, X_{21}), (X_{12}, X_{22}), \dots, (X_{1S}, X_{2S})$ is :

$$L(\Theta) = \prod_{k=1}^S l_k(\Theta) \tag{14}$$

where

$$l_k(\Theta) = C(k) \prod_{r=1}^R \frac{p_{rk}^{x_{rk}} (\beta_0 z_{rk})^{x_{rk}}}{(1 + \beta_0 \sum_{j=1}^R z_{jk} p_{jk})^{x_{rk}}} \tag{15}$$

with $x_{.rk} = x_{1rk} + x_{2rk}$ and

$$C(k) = \frac{n_k!}{2^R \prod_{t=1}^2 \prod_{m=1}^R x_{tmk}!}$$

The log-likelihood is given by :

$$\begin{aligned} \mathcal{L}(\Theta) = & \text{constante} + \sum_{k=1}^S \sum_{r=1}^R \{x_{.rk} \log_e(p_{rk}) \\ & + x_{2rk} \log_e(\beta_0) \\ & - x_{.rk} \log_e(1 + \beta_0 \sum_{m=1}^R z_{mk} p_{mk})\} \end{aligned} \quad (16)$$

It is sufficient to derive the function

$$\begin{aligned} g(\lambda_1, \lambda_2, \dots, \lambda_S; \Theta) = & \mathcal{L}(\Theta) \\ & + \sum_{k=1}^S \lambda_k (1 - \sum_{r=1}^R p_{rk}) \end{aligned}$$

according to parameters Θ where $\lambda_1, \lambda_2, \dots, \lambda_S$ are the Lagrange's multipliers. We obtain :

$$\frac{\partial g}{\partial \beta_0} = \sum_{s=1}^S \sum_{r=1}^R \left(\frac{x_{2rs} - x_{1rs} \beta_0 c_s}{\beta_0 (1 + \beta_0 c_s)} \right)$$

and for $k = 1, \dots, S; j = 1, \dots, R$

$$\frac{\partial g}{\partial p_{jk}} = \frac{x_{.jk}}{p_{jk}} - \frac{n_k \beta_0 z_{jk}}{1 + \beta_0 c_k} - \lambda_k$$

Multiplying this last equation par p_{jk} , summing on j and setting to zero, we deduce that

$$\hat{\lambda}_k = \frac{n_k}{1 + \beta_0 \hat{c}_k}, \quad k = 1, \dots, S.$$

Setting the partial derivatives of g to zero and replacing $\hat{\lambda}_k$ by its expression, we obtain the maximum likelihood equations under linear restraints.

Furthermore, the asymptotic Fisher information matrix J_Θ is then given by the following $(1 + SR) \times (1 + SR)$ matrix :

$$J_\Theta = \mathbb{E} \left(- \frac{\partial^2 \mathcal{L}}{\partial \Theta \partial \Theta^T} \right)$$

evaluated at different points on the line segment joining Θ^0 and $\hat{\Theta}$ the constrained ML, and the $S \times (1 + SR)$ constraint matrix $H_\Theta = (H_1, H_2)$ with $H_1 = \frac{\partial h}{\partial \beta_0}$ a $S \times 1$ matrix, $H_2 = \frac{\partial h}{\partial P}$ a $S \times (SR)$ matrix. Consequently, the $(1 + S(R + 1)) \times (1 + S(R + 1))$ bordered information matrix is given by

$$\Gamma_\Theta = \begin{bmatrix} J_\Theta & H_\Theta^T \\ H_\Theta & \mathbf{0}_{S,S} \end{bmatrix}, \quad J_\Theta = \begin{bmatrix} \tau_\Theta & U_\Theta^T \\ U_\Theta & B_\Theta \end{bmatrix}, \quad (17)$$

where U_Θ is a $(SR) \times 1$ vector as follows

$$U_\Theta = [U_{\Theta,1}^T, U_{\Theta,2}^T, \dots, U_{\Theta,S}^T]^T, \quad (18)$$

and J_Θ and H_Θ are matrices as follows

$$B_\Theta = \begin{bmatrix} B_{\Theta,1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & B_{\Theta,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & B_{\Theta,S} \end{bmatrix}, \quad (19)$$

$$H_\Theta = \begin{bmatrix} 0 & 1_R^T & 0_R^T & \dots & 0_R^T \\ 0 & 0_R^T & 1_R^T & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0_R^T \\ 0 & 0_R^T & \dots & 0_R^T & 1_R^T \end{bmatrix}, \quad (20)$$

with $\tau_\Theta = \mathbb{E} \left(- \frac{\partial^2 \mathcal{L}}{\partial \beta_0 \partial \beta_0} \right)$ a scalar, $U_{\Theta,k}$ a $R \times 1$ vector whose components are $\mathbb{E} \left(- \frac{\partial^2 \mathcal{L}}{\partial \beta_0 \partial p_{jk}} \right)$, ($j = 1, 2, \dots, R$), $B_{\Theta,k}$ a $R \times R$ matrix which entries are $\mathbb{E} \left(- \frac{\partial^2 \mathcal{L}}{\partial p_{jk} \partial p_{mk}} \right)$, ($j, m = 1, 2, \dots, R$), and $0_R^T = (0, \dots, 0)$ a $R \times 1$ vector of zero. So taking the second derivative of the negative of $\mathcal{L}(\Theta)$ with respect to the components of Θ , one shows (see N'Guessan and Langrand 2005a for details) that

$$\tau_\Theta = \sum_{k=1}^S \frac{\gamma_k \langle Z_k, P_k \rangle}{\beta_0 (1 + \beta_0 \langle Z_k, P_k \rangle)}, \quad (21)$$

where

$$\langle Z_k, P_k \rangle = \sum_{m=1}^R z_{mk} p_{mk}$$

with

$$Z_k = (z_{1k}, z_{2k}, \dots, z_{Rk})^T,$$

and, for $k = 1, 2, \dots, S$,

$$\gamma_k = \frac{n_k}{1 + \beta_0 \langle Z_k, P_k \rangle},$$

$$U_{\Theta,k} = \frac{\gamma_k^{3/2}}{\beta_0 n_k^{1/2}} V_{\Theta,k} \quad (22)$$

and

$$B_{\Theta,k} = \gamma_k (\Omega_{\Theta,k} - V_{\Theta,k} V_{\Theta,k}^T) \quad (23)$$

with

$$V_{\Theta,k} = \beta_0 \frac{\gamma_k^{1/2}}{n_k^{1/2}} Z_k,$$

$$\Omega_{\Theta,k} = \text{diag} \left(\frac{1 + \beta_0 z_{1k}}{p_{1k}}, \dots, \frac{1 + \beta_0 z_{Rk}}{p_{Rk}} \right).$$

lemma 1 For $k = 1, 2, \dots, s$

- (i) $\| Z_k \|^2_{\Omega_{\Theta,k}^{-1}} < \frac{\langle Z_k, P_k \rangle}{\beta_0}$;
- (ii) $\| V_{\Theta,k} \|^2_{\Omega_{\Theta,k}^{-1}} < \frac{\beta_0 \langle Z_k, P_k \rangle}{1 + \beta_0 \langle Z_k, P_k \rangle} < 1$;
- (iii) matrix $B_{\Theta,k}$ is nonsingular and

$$B_{\Theta,k}^{-1} = \gamma_k^{-1} (\Omega_{\Theta,k}^{-1} + t_k \Omega_{\Theta,k}^{-1} V_{\Theta,k} V_{\Theta,k}^T \Omega_{\Theta,k}^{-1}),$$

with $t_k = (1 - \| V_{\Theta,k} \|^2_{\Omega_{\Theta,k}^{-1}})^{-1}$

lemma 2

- (i) $\| U_{\Theta} \|^2_{B_{\Theta}^{-1}} < \tau_{\Theta}$,
- (ii) $(J_{\Theta}/B_{\Theta}) > 0$, where (J_{Θ}/B_{Θ}) is the Schur complement of B_{Θ} in J_{Θ} .

lemma 3

$(J_{\Theta}/\tau_{\Theta})$ the Schur complement of τ_{Θ} in J_{Θ} is a (SR) \times (SR) nonsingular matrix with

$$(J_{\Theta}/\tau_{\Theta})^{-1} = (J_{k,m})_{(1 \leq k \leq S; 1 \leq m \leq S)}$$

where each block $J_{k,m}$ is a $R \times R$ matrix given by the Schur complement of (J_{Θ}/B_{Θ}) in the $(1 + R) \times (1 + R)$ matrix

$$\begin{bmatrix} (J_{\Theta}/B_{\Theta}) & U_{\Theta,m}^T B_{\Theta,m}^{-1} \\ B_{\Theta,k}^{-1} U_{\Theta,k} & \delta_{k,m} B_{\Theta,k}^{-1} \end{bmatrix}$$

with $\delta_{k,m}$ is the Kronecker symbol.

lemma 4

The $S \times S$ matrix $Q_{\Theta} = H_2(J_{\Theta}/\tau_{\Theta})^{-1} H_2^T$ is nonsingular and

$$Q_{\alpha}^{-1} = \Lambda_{\Theta}^{-1} - [(\Lambda_{\Theta}^{(1)}/\Lambda_{\Theta})]^{-1} \Lambda_{\Theta}^{-1} \xi_{\Theta} \xi_{\Theta}^T \Lambda_{\Theta}^{-1}.$$

lemma 5 We get the following results

$$(i) \Delta_{\Theta,k} = \sum_{m=1}^R \frac{z_{mk} p_{mk}}{1 + \beta_0 z_{mk}},$$

$$(ii) \| \mathbf{1}_R \|^2_{B_{\Theta,k}^{-1}} = \gamma_k^{-1} [\text{trace}(\Omega_{\Theta,k}^{-1}) + t_k \frac{\beta_0^2 \gamma_k \Delta_{\Theta,k}^2}{n_k}],$$

$$(iii) (\Lambda_{\Theta}^{(1)}/\Lambda_{\Theta}) = \sum_{k=1}^S \frac{\gamma_k^2}{n_k} \left[\left(\frac{\langle Z_k, P_k \rangle}{\beta_0} - \frac{\gamma_k t_k}{n_k} \| Z_k \|^2_{\Omega_{\Theta,k}^{-1}} \right) + \frac{t_k^2 \Delta_{\Theta,k}^2}{n_k \| \mathbf{1}_R \|^2_{B_{\Theta,k}^{-1}}} \right]$$

$$(iv) \| Z_k \|^2_{\Omega_{\Theta,k}^{-1}} = \sum_{m=1}^R \frac{p_{mk} z_{mk}^2}{1 + \beta_0 z_{mk}}$$

Appendix B : Computing of the asymptotic standard errors

This appendix gives the recent main results concerning the asymptotic covariance estimation using Schur complement (see N'Guessan and Langrand 2005b). The restraints m.l.e (RMLE) of Θ_0 , provided it exists, is vector $\hat{\Theta}$ which satisfies

$$\left(\frac{\partial \mathcal{L}}{\partial \Theta} \right)_{\hat{\Theta}} + H_{\hat{\Theta}}^T \hat{\lambda} = 0, \quad h(\hat{\Theta}) = 0 \quad (24)$$

where λ is a vector of S Lagrange multipliers. Let $M_{\Theta} = J_{\Theta}^{-1} \left(J_{\Theta} + \left(\frac{\partial^2 \mathcal{L}}{\partial \Theta \partial \Theta} \right) \right)$ a $(1 + SR) \times (1 + SR)$ matrix, $Q_{\Theta} = H_2(J_{\Theta}/\tau_{\Theta})^{-1} H_2^T$ a $S \times S$ matrix, $\Phi_{\Theta} = J_{\Theta}^{-1} H_{\hat{\Theta}}^T Q_{\Theta}^{-1}$ a $(1 + SR) \times S$ matrix, $\xi_{\Theta} = H_2 B_{\Theta}^{-1} U_{\Theta}$ a $S \times 1$ vector, $\Lambda_{\Theta} = \text{diag}(\| \mathbf{1}_R \|^2_{B_{\Theta,1}^{-1}}, \| \mathbf{1}_R \|^2_{B_{\Theta,2}^{-1}}, \dots, \| \mathbf{1}_R \|^2_{B_{\Theta,S}^{-1}})$ a $S \times S$ matrix with $\| \mathbf{1}_R \|^2_{B_{\Theta,k}^{-1}} = \mathbf{1}_R^T B_{\Theta,k}^{-1} \mathbf{1}_R$, and

$$\Lambda_{\Theta}^{(1)} = \begin{bmatrix} \Lambda_{\Theta} & -\xi_{\Theta} \\ \xi_{\Theta}^T & (J_{\Theta}/B_{\Theta}) \end{bmatrix} \quad (25)$$

a $(S + 1) \times (S + 1)$ matrix. We assume for the rest of the work that the following conditions are verified :

- (A1) $(\hat{\Theta}, \hat{\lambda})$ is a consistent solution of the latter equation.

$$(A2) M_{\hat{\Theta}} \xrightarrow{\text{Proba}} 0, \quad \Phi_{\hat{\Theta}} - \Phi_{\Theta_0} \xrightarrow{\text{Proba}} 0,$$

$|\Phi_{\Theta_0}| < \infty$, when n_k and $N = \sum_{k=1}^S n_k \rightarrow \infty$

$$(A3) \left(\frac{\partial \mathcal{L}}{\partial \Theta} \right)_{\Theta_0} \stackrel{\text{Distr.}}{\sim} \mathcal{N}(0, J_{\Theta_0}).$$

Theorem 1 Matrix Γ_{Θ} is nonsingular and

$$\Gamma_{\Theta}^{-1} = \begin{bmatrix} W_{\Theta} & J_{\Theta}^{-1} H_{\hat{\Theta}}^T Q_{\Theta}^{-1} \\ Q_{\Theta}^{-1} H_{\Theta} J_{\Theta}^{-1} & -Q_{\Theta}^{-1} \end{bmatrix}$$

where W_{Θ} , the leading $(1 + SR) \times (1 + SR)$ matrix, is given by

$$W_{\Theta} = \begin{bmatrix} W_{\Theta}(1,1) & W_{\Theta}^T(2,1) \\ W_{\Theta}(2,1) & W_{\Theta}(2,2) \end{bmatrix}$$

with $W_{\Theta}(1,1)$, $W_{\Theta}(2,1)$ and $W_{\Theta}(2,2)$ respectively a scalar, a $(SR) \times 1$ vector and a $(SR) \times (SR)$ matrix as follows :

$$W_{\Theta}(1,1) = (\Lambda_{\Theta}^{(1)}/\Lambda_{\Theta})^{-1} = \sigma^2(\hat{\beta}_0) \quad (26)$$

$$\begin{aligned}
 W_{\Theta}(2,1) &= -(J_{\Theta}/B_{\Theta})^{-1}[B_{\Theta}^{-1}U_{\Theta} \\
 &\quad - (J_{\Theta}/\tau_{\Theta})^{-1}H_2^T Q_{\Theta}^{-1}\xi_{\Theta}] \quad (27) \\
 &= COV(\hat{\beta}_0; \hat{P})
 \end{aligned}$$

$$\begin{aligned}
 W_{\Theta}(2,2) &= (J_{\Theta}/\tau_{\Theta})^{-1} \\
 &\quad - (J_{\Theta}/\tau_{\Theta})^{-1}H_2^T Q_{\Theta}^{-1}H_2(J_{\Theta}/\tau_{\Theta})^{-1} \quad (28) \\
 &= COV(\hat{P}; \hat{P})
 \end{aligned}$$

Theorem 2 Under conditions (A1) - (A3) the asymptotic distribution of $\hat{\Theta}$ is :

$$\hat{\Theta} \stackrel{Distr.}{\sim} \mathcal{N}(\Theta^0; W_{\Theta^0})$$

where $W_{\Theta} = J_{\Theta}^{-1} - J_{\Theta}^{-1}H_{\Theta}^T Q_{\Theta}^{-1}H_{\Theta}J_{\Theta}^{-1}$ is the $(1+SR) \times (1+SR)$ partitioned matrix given above.

Theorem 3 Under the same assumptions as theorem above, the $R \times R$ between covariance matrix of \hat{P}_k and \hat{P}_m is

$$\begin{aligned}
 COV(\hat{P}_k; \hat{P}_m) &= J_{k,m} \\
 &\quad - \sum_{j=1}^S \left[\left(\sum_{i=1}^S A_{i,j} J_{k,i} \right) \mathbf{1}_{R,R} \right] J_{j,m} \quad (29)
 \end{aligned}$$

$k, m = 1, 2, \dots, S$. Where $J_{k,m}$ is a $R \times R$ matrix given by the Schur complement of (J_{Θ}/B_{Θ}) in the $(1+R) \times (1+R)$ matrix

$$\begin{bmatrix} (J_{\Theta}/B_{\Theta}) & U_{\Theta,m}^T B_{\Theta,m}^{-1} \\ B_{\Theta,k}^{-1} U_{\Theta,k} & \delta_{k,m} B_{\Theta,k}^{-1} \end{bmatrix},$$

and

$$\begin{aligned}
 A_{k,m} &= \frac{\delta_{k,m}}{\|\mathbf{1}_R\|_{B_{\Theta,k}^{-1}}^2} \\
 &\quad - \frac{(\Lambda_{\Theta}^{(1)}/\Lambda_{\Theta})^{-1} t_k t_m \gamma_k \gamma_m \Delta_{\Theta,k} \Delta_{\Theta,m}}{n_k n_m \|\mathbf{1}_R\|_{B_{\Theta,k}^{-1}}^2 \|\mathbf{1}_R\|_{B_{\Theta,m}^{-1}}^2}, \quad (30)
 \end{aligned}$$

with $\delta_{k,m}$ the usual kronecker symbol and $\mathbf{1}_{R,R} = \mathbf{1}_R \mathbf{1}_R^T$ the $R \times R$ matrix with all entries equal to one.

Corollary 4 Under the same assumptions as theorem above and when the measure is applied to a single site ($S = 1$) with different accident types ($R > 1$), the covariances between the different components of the $R \times 1$ vector \hat{P}_k are as follows

$$\begin{aligned}
 cov(\hat{p}_{jk}, \hat{p}_{mk}) &= \gamma_k^{-1} w_{jk} w_{mk} [\theta_k \gamma_k z_{jk} z_{mk} \\
 &\quad - A_{k,k} \gamma_k^{-1} y_{jk} y_{mk}], \quad (31) \\
 (j \neq m = 1, 2, \dots, R)
 \end{aligned}$$

$$\begin{aligned}
 \sigma^2(\hat{p}_{jk}) &= \gamma_k^{-1} \{w_{jk} \\
 &\quad + w_{jk}^2 (\theta_k \gamma_k z_{jk}^2 - A_{k,k} \gamma_k^{-1} y_{jk}^2)\}, \quad (32) \\
 (j = 1, 2, \dots, R)
 \end{aligned}$$

where $w_{jk} = \frac{p_{jk}}{1+\beta_0 z_{jk}}$
 the diagonal elements of matrix $\Omega_{\Theta,k}^{-1}$,
 $\theta_k = \frac{t_k}{n_k} \left[\beta_0^2 - (J_{\Theta}/B_{\Theta})^{-1} \frac{t_k \gamma_k^2}{n_k} \right]$,
 $y_{jk} = 1 + \theta_k \Delta_{\Theta,k} \gamma_k z_{jk}$; ($j = 1, 2, \dots, R$).

Références

- [1] Abdel-Aty M.A., Radwan A.E (2000), *Modeling traffic accident occurrence and involvement*, *Accid. Anal. and Prev.* 32, 633-642.
- [2] Aitchison J., Silvey S.D, *Maximum Likelihood Estimation of parameters subject to restraints* *Annals of mathematical statistics*, 29,813-829, 1958.
- [3] Allain E and Brenac T., *The application of generalized linear models to the investigation of accident frequencies at road sites. The Poisson model and its extensions* *Recherche Transports Sécurité* n°72, 3-18, 2001.
- [4] Barry S., Ginpil S, O'Neill, T.J, *The effectiveness of air bags*, *Accid. Anal. and Prev.* 31, 781-787, 1999.
- [5] Beenstock M., Gafni D., Goldin E, *The effect of traffic policing on road safety in Israel*, *Accid. Anal. and Prev.* 33, 73-80, 2001
- [6] Crowder M. *On the constrained maximum likelihood estimation with non i.i.d. observations* *Ann. Inst. Statist. Math.* 36, A, 239-249, 1984.
- [7] Danielson S., *Estimation of the effects of countermeasures on different types of accidents in the presence of regression effects*, *Accid. Anal. and Prev.* 20, n°4 , 289-298, 1988.
- [8] Danielson S., *A comparison of two methods for estimating of the effect of a countermeasure in the presence of regression effects*, *Accid. Anal. and Prev.* 18, n°1 , 13-23, 1986.
- [9] Hauer E., *Observational before-after Studies in Road Safety Estimating the effect of highway and traffic engineering measures on road safety*, Pergamon Edition, 1997.
- [10] HAUER E., *Comparison Groups in road safety studies : An analysis*, *Accid. Anal. and Prev.* 23, n°26, 609-622, 1991.
- [11] HAUER E., *Overdispersion in modelling accidents on road sections and in Empirical Bayes estimation*, *Accid. Anal. and Prev.* 33, 799-808,2001.
- [12] HAUER E., NG J. C. N., PAPAIOANNOU P., *Prediction in road safety studies : An empirical inquiry*, *Accid. Anal. and Prev.* 23, n°6, 595-607, 1991.
- [13] LASSARRE S., *A propos des tests statistiques, sur variables poissonniennes, utilisées dans le domaine de la sécurité routière*, *Revue de Statistique Appliquée* XXV, n°3, 1977.
- [14] LASSARRE S., *Analysis of progress in road safety in ten European countries*, *Accident Analysis and Prevention* 33 (2001). 743-751, 2001.
- [15] NAG, *Numerical Algorithms Groups Limited, 2000*, Ecole Polytechnique Universitaire de Lille, Université de Lille1, 2003.
- [16] N'Guessan A. , *On a use of Schur complements for an inverse of a constrained accident data information matrix* *PUB. IRMA, LILLE* Vol. 60, N°VI, 2003a.
- [17] N'Guessan A. , *Constrained covariance matrix estimation in road accident modelling with Schur complements* *C.R. Acad. Sci. Paris, Ser. I* 337 (2003) 219-222, 2003b.
- [18] N'Guessan A., Essai A., Langrand C., *Estimation multidimensionnelle des contrôles et de l'effet moyen d'une mesure de sécurité routière* *Rev. Statistique Appliquée.* XLIX(2), 83-100, 2001.
- [19] N'Guessan A., Langrand C., *A Schur complement approach for computing subcovariance matrices arising in a road safety measure modelling* *Journal of Computational and Applied Mathematics*, 177, 331-345, 2005a.

- [20] N'Guessan A., Langrand C., *A covariance components estimation procedure when modelling a road safety measure in terms of linear constraints*. Statistics 2005b.
- [21] NICHOLSON A. J., *The variability of accident counts*, *Accid. Anal. and Prev.* 17, 47-56, 1985.
- [22] NICHOLSON A. J., *The randomness of accident counts*, *Accid. Anal. and Prev.* 18, 193-198, 1986.
- [23] NICHOLSON A., WONG Y-D., *Are accidents Poisson distributed? A statistical test*, *Accid. Anal. and Prev.* 23, n°26, 609-622, 1993.
- [24] Oppe S., *The use of multiplicative models for analysis of road safety data*. *Accid. Anal. and Prev.* vol. 11, 101-115, 1979.
- [25] Ouellette D.V., *Schur complements and statistics* *Linear Algebra and its applications*, 36 187-295, 1981.
- [26] Rao C.R., *Linear Statistical Inference and Its Applications*, (2nd ed.), John Wiley, New York, 1973.
- [27] Rao C.R. and Yanai H., *Generalized Inverses of Partitioned Matrices Useful in Statistical Applications*, *Linear Algebra and its Applications*, 70 : 105-115, 1985.
- [28] Silvey S.D., *The Lagrangian multiplier test* *Ann. Math. Statist.* 30, 389-407, 1959.
- [29] Tanner J.C. (1958), *A problem in the combination of accident frequencies*, *Biometrika*, 45, 331-342.
- [30] Zhang F. ed., *Schur complement and Its applications* (Springer Verlag), 2005.

Dr. Assi N'Guessan is Assistant Professor at the Department of Statistics and Computer Sciences, Polytech'Lille, University of Lille 1, France.

Dr. Claude Langrand is Professor at the Department of Mathematics, University of Lille 1 - France

Annexe A07 :

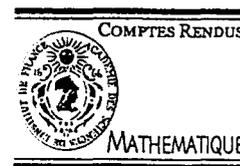
C. R. Acad. Sci. Paris, **Série I**, **337**, **219-222 (2003)**

Constrained estimation of a road safety covariance matrix using Schur complements

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

C. R. Acad. Sci. Paris, Ser. I 337 (2003) 219–222



Statistics/Probability Theory

Constrained covariance matrix estimation in road accident modelling with Schur complements

Assi N'Guessan

École polytechnique universitaire de Lille et laboratoire de mathématiques appliquées: CNRS, FRE 2222, Université de Lille 1, 59655 Villeneuve d'Ascq cedex, France

Received 31 March 2003; accepted after revision 17 June 2003

Presented by Paul Deheuvels

Abstract

We consider a collection $(Y_{11}, Y_{21}), \dots, (Y_{1s}, Y_{2s})$ of s independent couples of $2r$ -dimensional random vector with $r > 1$. We assume that each couple has a multinomial distribution linked to an unknown vector parameter and an extra set data. We deal with a formal inversion of a Fisher's information matrix connected to those couples using Schur complements approach. *To cite this article: A. N'Guessan, C. R. Acad. Sci. Paris, Ser. I 337 (2003).*

© 2003 Académie des sciences. Published by Éditions scientifiques et médicales Elsevier SAS. All rights reserved.

Résumé

Estimation sous contraintes d'une matrice de covariances dans la modélisation d'accidents de la route par compléments de Schur. On considère $(Y_{11}, Y_{21}), \dots, (Y_{1s}, Y_{2s})$ une collection de s couples indépendants de vecteurs aléatoires de dimension $2r$ avec $(r > 1)$. On suppose que chaque couple est distribué selon une loi multinomiale dépendant à la fois d'un vecteur paramètre inconnu et d'un ensemble de données supplémentaire. On étudie l'inversion formelle de l'information de Fisher issue de ces s couples par le biais de la technique des compléments de Schur. *Pour citer cet article: A. N'Guessan, C. R. Acad. Sci. Paris, Ser. I 337 (2003).*

© 2003 Académie des sciences. Published by Éditions scientifiques et médicales Elsevier SAS. All rights reserved.

1. Introduction and motivation

N'Guessan, Essai and Langrand [6] deal with the multidimensional estimation of the effect of road safety measure applied to s ($s > 0$) entities (sites), each one including r ($r > 1$) mutually exclusive different accident types and linked to a specific control area. They consider a collection $(Y_{11}, Y_{21}), (Y_{12}, Y_{22}), \dots, (Y_{1s}, Y_{2s})$ of s independent couples of $2r$ -dimensional random vector and assume that, for fixed k , each couple (Y_{1k}, Y_{2k}) has multinomial distribution denoted by $\mathcal{M}(n_k; \Pi_{1k}(\alpha), \Pi_{2k}(\alpha))$, where

$$\Pi_{1k}(\alpha) = (\pi_{11k}(\alpha), \pi_{12k}(\alpha), \dots, \pi_{1rk}(\alpha)),$$

E-mail address: Assi.Nguessan@polytech-lille.fr (A. N'Guessan).

1631-073X/\$ – see front matter © 2003 Académie des sciences. Published by Éditions scientifiques et médicales Elsevier SAS. All rights reserved.

doi:10.1016/S1631-073X(03)00309-1

220

A. N'Guessan / C. R. Acad. Sci. Paris, Ser. I 337 (2003) 219–222

n_k ($n_k > 0$) is a given integer, $\alpha = (\beta_0, \beta^T)^T \in \mathbb{R}^{1+sr}$ a unknown vector parameter, β_0 ($\beta_0 > 0$) a scalar, $\beta = (\beta_1^T, \dots, \beta_s^T)^T \in \mathbb{R}^{sr}$ with $\beta_k = (\beta_{1k}, \beta_{2k}, \dots, \beta_{rk})^T \in \mathbb{R}^r$ and $\beta_{jk} > 0$. The components $\pi_{irk}(\alpha)$ of the $r \times 1$ vector probability $\Pi_{rk}(\alpha)$ are given by a continuous differentiable functional link $g: \mathbb{R}^{1+sr} \mapsto]0, 1[$ so that $\pi_{irk}(\alpha) = g(z_k, \alpha)$ and

$$g(z_k, \alpha) = \begin{cases} \frac{\beta_{jk}}{1 + \beta_0 \langle z_k, \beta_k \rangle}, & t = 1; j = 1, 2, \dots, r, \\ \frac{z_{jk} \beta_0 \beta_{jk}}{1 + \beta_0 \langle z_k, \beta_k \rangle}, & t = 2; j = 1, 2, \dots, r, \end{cases} \tag{1}$$

where $\langle \cdot, \cdot \rangle$ is the usual inner product, $z_k = (z_{1k}, z_{2k}, \dots, z_{rk})^T$ a given extra set data connected to the control area of each couple k . The latter authors are interested in the m.l.e. $\hat{\alpha}$ of α and its numerical consistency under the restraints $h_k(\alpha) = 0$ where

$$h_k(\alpha) = \langle \mathbf{1}_r, \beta_k \rangle - 1; \quad k = 1, 2, \dots, s, \tag{2}$$

with $\mathbf{1}_r = (1, \dots, 1)^T \in \mathbb{R}^r$, the vector of unity. The main motivation of this Note is to make their paper complete by investigating the asymptotic precision of $\hat{\alpha}$, i.e., the asymptotic variances–covariances matrix.

2. Technical intermediate results

In the following we note $H_\alpha = (H_1, H_2)$ the constraints matrix with $H_1 = \frac{\partial h}{\partial \beta_0}$ a $s \times 1$ matrix, $H_2 = \frac{\partial h}{\partial \beta}$ a $s \times (sr)$ matrix, $h = (h_1, h_2, \dots, h_s)$, $L(\alpha)$ the log-likelihood function based on the s couples, $J_\alpha = E(-\frac{\partial^2 L}{\partial \alpha \partial \alpha})$, the $(1 + sr) \times (1 + sr)$ information matrix evaluated at different points on the line segment joining α^0 (the true vector parameter) and $\hat{\alpha}$. Statistical aspects of asymptotic results of restraints m.l.e. $\hat{\alpha}$ and related problems have been extensively discussed in Aitchison and Silvey [1], Crowder [2,3], Harville [4], Lee [5], and Van Eeden [9]. It can be shown under regularity conditions, that $\hat{\alpha}$ is asymptotically normally distributed and its covariance matrix strongly depends on the fact that both J_α and $H_\alpha J_\alpha^{-1} H_\alpha^T$ are nonsingular. Technical lemmas below elaborate these points. The technical demonstrations and further straightforward matrix manipulations are displayed in N'Guessan [7].

Lemma 2.1. Let us note $V_{\alpha,k} = \frac{\beta_0}{1 + \beta_0 \langle z_k, \beta_k \rangle} (z_{1k}, z_{2k}, \dots, z_{rk})^T$ a $r \times 1$ vector and $\Omega_{\alpha,k} = \text{diag}(\frac{1 + \beta_0 z_{1k}}{\beta_{1k}}, \frac{1 + \beta_0 z_{2k}}{\beta_{2k}}, \dots, \frac{1 + \beta_0 z_{rk}}{\beta_{rk}})$ a $r \times r$ matrix. Then for $k = 1, 2, \dots, s$: (i) $\|z_k\|_{\Omega_{\alpha,k}^{-1}}^2 < \frac{\langle z_k, \beta_k \rangle}{\beta_0}$; (ii) $\|V_k\|_{\Omega_{\alpha,k}^{-1}}^2 < \frac{\beta_0 \langle z_k, \beta_k \rangle}{(1 + \beta_0 \langle z_k, \beta_k \rangle)^2} < 1$.

Lemma 2.2. Let us note $B_{\alpha,k} = \gamma_k (\Omega_{\alpha,k} - V_{\alpha,k} V_{\alpha,k}^T)$ with $\gamma_k = \frac{n_k}{1 + \beta_0 \langle z_k, \beta_k \rangle}$. Then (i) for $k = 1, 2, \dots, s$; $B_{\alpha,k}^{-1} = \gamma_k^{-1} (\Omega_{\alpha,k}^{-1} + t_k \Omega_{\alpha,k}^{-1} V_{\alpha,k} V_{\alpha,k}^T \Omega_{\alpha,k}^{-1})$; $t_k = (1 - \|V_k\|_{\Omega_{\alpha,k}^{-1}}^2)^{-1}$; (ii) $\|U_\alpha\|_{B_\alpha^{-1}}^2 < \tau_\alpha$; where $B_\alpha = \text{Bloc} - \text{diag}(B_{\alpha,1}, \dots, B_{\alpha,s})$, $U_\alpha = (U_{\alpha,1}^T, \dots, U_{\alpha,s}^T)^T$ a $(sr) \times 1$ vector with $U_{\alpha,k} = \frac{\gamma_k}{\beta_0} V_{\alpha,k}$ and $\tau_\alpha = \sum_{k=1}^s \frac{\gamma_k \langle z_k, \beta_k \rangle}{\beta_0 (1 + \beta_0 \langle z_k, \beta_k \rangle)}$.

Using model (1) above and after some straightforward derivations and matrix manipulations one obtains:

$$J_\alpha = \begin{pmatrix} \tau_\alpha & U_\alpha^T \\ U_\alpha & B_\alpha \end{pmatrix}.$$

Following Ouellette [8] we note (J_α / B_α) (resp. (J_α / τ_α)) the Schur [8] complement of B_α (resp. τ_α) in J_α and defined by $(J_\alpha / B_\alpha) = \tau_\alpha - U_\alpha^T B_\alpha^{-1} U_\alpha$ (resp. $(J_\alpha / \tau_\alpha) = B_\alpha - U_\alpha \tau_\alpha^{-1} U_\alpha^T$).

Lemma 2.3. (i) $(J_\alpha / B_\alpha) > 0$; (ii) (J_α / τ_α) is nonsingular and

$$(J_\alpha / \tau_\alpha)^{-1} = B_\alpha^{-1} + (J_\alpha / B_\alpha)^{-1} B_\alpha^{-1} U_\alpha U_\alpha^T B_\alpha^{-1}. \tag{3}$$

Lemma 2.4. (i) J_α^{-1} exists and

$$J_\alpha^{-1} = \begin{pmatrix} (J_\alpha/B_\alpha)^{-1} & -(J_\alpha/B_\alpha)^{-1}U_\alpha^T B_\alpha^{-1} \\ -B_\alpha^{-1}U_\alpha(J_\alpha/B_\alpha)^{-1} & (J_\alpha/\tau_\alpha)^{-1} \end{pmatrix}; \tag{4}$$

(ii)

$$(H_\alpha J_\alpha^{-1} H_\alpha^T)^{-1} = \Lambda_\alpha^{-1} - ((J_\alpha/B_\alpha) + \|\xi_\alpha\|_{\Lambda_\alpha^{-1}}^2)^{-1} \Lambda_\alpha^{-1} \xi_\alpha \xi_\alpha^T \Lambda_\alpha^{-1}, \tag{5}$$

where $\xi_\alpha = H_2 B_\alpha^{-1} U_\alpha$ a $s \times 1$ vector and $\Lambda_\alpha = \text{diag}(\|\mathbf{1}_r\|_{B_{\alpha,1}}^2, \|\mathbf{1}_r\|_{B_{\alpha,2}}^2, \dots, \|\mathbf{1}_r\|_{B_{\alpha,s}}^2)$ a $s \times s$ matrix.

Remark 1. Lemmas 2.3 and 2.4 provide from general applications of well-known results of Schur complement about the inverse of partitioned matrix (see Ouellette [8], pp. 201–204) and combine model (1) and constraints (2).

3. Main result

The restraints m.l.e. of α_0 , provided it exists, is the vector $\hat{\alpha}$ which satisfies

$$\left(\frac{\partial L}{\partial \alpha}\right)_{\hat{\alpha}} + H_{\hat{\alpha}}^T \hat{\lambda} = 0, \quad h(\hat{\alpha}) = 0, \tag{6}$$

where λ is a vector of s Lagrange multipliers.

Let $M_\alpha = J_\alpha^{-1}(J_\alpha + (\frac{\partial^2 L}{\partial \alpha^2}))$ a $(1 + sr) \times (1 + sr)$ matrix, $R_\alpha = H_2(J_\alpha/\tau_\alpha)^{-1}H_2^T$ a $s \times s$ matrix, $\Phi_\alpha = J_\alpha^{-1}H_\alpha^T R_\alpha^{-1}$ a $(1 + sr) \times s$ matrix.

Theorem 3.1. Suppose that

- (A1) $(\hat{\alpha}, \hat{\lambda})$ is a consistent solution of (6).
- (A2) $M_{\hat{\alpha}} \xrightarrow{\text{Proba}} 0, \Phi_{\hat{\alpha}} - \Phi_{\alpha^0} \xrightarrow{\text{Proba}} 0, |\Phi_{\alpha^0}| < \infty.$
- (A3) $(\frac{\partial L}{\partial \alpha})_{\alpha^0} \overset{\text{Distr.}}{\sim} \mathcal{N}(0, J_{\alpha^0}).$

Then the asymptotic components of the covariance matrix of $\hat{\alpha}$ are:

$$\begin{aligned} \sigma^2(\hat{\beta}_0) &= ((J_\alpha/B_\alpha) + \|\xi_\alpha\|_{\Lambda_\alpha^{-1}}^2)^{-1}, \\ \text{Var}(\hat{\beta}) &= (J_\alpha/\tau_\alpha)^{-1} - (J_\alpha/\tau_\alpha)^{-1}H_2^T R_\alpha^{-1}H_2(J_\alpha/\tau_\alpha)^{-1}, \\ \text{COV}(\hat{\beta}, \hat{\beta}_0) &= -(J_\alpha/B_\alpha)^{-1}(B_\alpha^{-1}U_\alpha - (J_\alpha/\tau_\alpha)^{-1}H_2^T R_\alpha^{-1}\xi_\alpha). \end{aligned}$$

Proof (outline). Application of the mean value theorem at α^0 to (6) and Lemma 2.4 give

$$\begin{aligned} \left(\frac{\partial L}{\partial \alpha}\right)_{\alpha^0} + J_\alpha(M_\alpha - I_{1+sr})(\hat{\alpha} - \alpha^0) + H_\alpha^T \hat{\lambda} &= 0, \\ H_\alpha(\hat{\alpha} - \alpha^0) &= 0. \end{aligned} \tag{7}$$

Under constraints (2) and (A1) to (A3) and using a version of a theorem of Crowder ([3], Section 2), then

$$\begin{pmatrix} \hat{\alpha} - \alpha^0 \\ \eta_{\hat{\alpha}} \end{pmatrix} \overset{\text{Distr.}}{\sim} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \begin{pmatrix} W_{\alpha^0} & 0 \\ 0 & R_{\alpha^0} \end{pmatrix}\right), \tag{8}$$

where $W_\alpha = J_\alpha^{-1} - J_\alpha^{-1}H_\alpha^T R_\alpha^{-1}H_\alpha J_\alpha^{-1}$; $\eta_{\hat{\alpha}} = R_\alpha \hat{\lambda}$. Now using formula (4) of Lemma 2.4 and after some matrix manipulation (N'Guessan [7], Appendix A.2) the components of W_{α^0} above follow. We will make capital out

of those results at subsequent papers in order to bring forward asymptotic confidence intervals and correlation coefficients (N'Guessan [7], Section 5). \square

Acknowledgements

I am especially grateful to Prof. Claude Langrand for his helpful comments and suggestions, which have led to an improvement in this article and my Ph.D. thesis in road safety measure analysis. I wish also thank Prof. Claude Brezinsky who drew my attention to the paper about Schur Complements. I thank the referees and Prof. Paul Deheveuls for constructive suggestions. This research was partially supported by *France–Maroc action intégrée programme N⁰MA01/02*.

References

- [1] J. Aitchison, S.D. Silvey, Maximum likelihood estimation of parameters subject to restraints, *Ann. Math. Statist.* 29 (1958) 813–829.
- [2] M. Crowder, Maximum likelihood estimation for dependent observations, *J. Roy. Statist. Soc.* 38 (1975) 45–53.
- [3] M. Crowder, On the constrained maximum likelihood estimation with non i.i.d. observations, *Ann. Inst. Statist. Math. Part A* 36 (1984) 239–249.
- [4] D.A. Harville, Maximum likelihood approaches to variance component estimation and to related problems, *J. Amer. Statist. Assoc.* 72 (358) (1977) 320–340.
- [5] S.-Y. Lee, Constrained estimation in covariance structure analysis, *Biometrika* 66 (3) (1979) 539–545.
- [6] A. N'Guessan, A. Essai, C. Langrand, Estimation multidimensionnelle des contrôles et de l'effet moyen d'une mesure de sécurité routière, *Rev. Statist. Appl.* XLIX (2) (2001) 83–100.
- [7] A. N'Guessan, On a use of Schur complements for an inverse of a constrained accident data information matrix, *Pub. IRMA Lille* 60 (VI) (2003).
- [8] D.V. Ouellette, Schur complements and statistics, *Linear Algebra Appl.* 36 (1981) 187–295.
- [9] C. Van Eeden, Estimation in restricted parameter spaces – some history and some recent developments, *CWI Quarterly* 9 (1 and 2) (1996) 69–76.

Annexe A08 :

C. R. Acad. Sci. Paris , **Série I**, 317, 401-404 (1993)

Sur la distribution asymptotique de certaines statistiques utilisées dans le domaine de la sécurité routière

Sur la distribution asymptotique de certaines statistiques utilisées dans le domaine de la sécurité routière

Assi N'GUESSAN et Claude LANGRAND

Résumé — L'étude porte sur l'effet d'une mesure de sécurité routière; on dispose, avant et après cette mesure, des effectifs d'accidents enregistrés dans s sites comportant chacun r types d'accidents. On note θ_k le paramètre effet de la mesure au site k ($k=1, 2, \dots, s$). On propose une statistique pour éprouver l'hypothèse $H_0: \theta_1 = \theta_2 = \dots = \theta_s = \theta$ et on étudie sous certaines conditions la loi asymptotique de cette statistique.

On the limiting Distribution of some statistics used in road safety studies

Abstract — To evaluate the effectiveness of a road safety measure, the number of accidents before and after the measure, classified in r categories, are observed on s sites, with θ_k the effect of the safety measure for the site k ($k=1, 2, \dots, s$). We propose a statistic to test the hypothesis $H_0: \theta_1 = \theta_2 = \dots = \theta_s = \theta$. Its asymptotic distribution is given under certain conditions.

1. INTRODUCTION ET MODÈLES. — On considère le problème de l'évaluation statistique de l'efficacité d'une mesure de sécurité routière dans le cas où les observations portent sur des sites traités et sur des sites de contrôle sur lesquels la mesure n'est pas directement appliquée. Pour tenir compte des difficultés liées au choix de la période d'observation, au choix des sites traités et des sites témoins (voir [1], [2]), on a utilisé des rapports de contrôle calculés pendant la même période sur des zones étendues entourant le site expérimental pour « corriger » le nombre d'accidents escompté après la mesure. On suppose ainsi disposer de l'observation avant et après cette mesure des effectifs d'accidents enregistrés dans différents sites expérimentaux comportant chacun plusieurs types d'accidents. Des modèles basés sur une distribution multinomiale permettent de répartir les accidents de chaque site expérimental entre les différents types en tenant compte de l'évolution de l'effectif de ceux-ci observés indépendamment de la mesure imposée (notion de sites de contrôle). On propose une statistique pour éprouver l'hypothèse H_0 et on étudie sous certaines conditions la loi limite de cette statistique. On note :

s le nombre de sites expérimentaux où la mesure est appliquée;

r le nombre de types (catégories) d'accidents retenus dans l'analyse;

$X_{1k} = (X_{11k}, X_{12k}, \dots, X_{1rk})$ [resp. $X_{2k} = (X_{21k}, X_{22k}, \dots, X_{2rk})$] le vecteur aléatoire (v. a.) donnant les effectifs des r types d'accidents sur le site expérimental k ($k=1, 2, \dots, s$) avant (resp. après) la mesure et dont une valeur observée est $(x_{11k}, x_{12k}, \dots, x_{1rk})$ [resp. $(x_{21k}, x_{22k}, \dots, x_{2rk})$];

$x_{..k}$ l'effectif total des accidents survenus dans le site k au cours des deux périodes, ($x_{..k}$ est déterminé);

$N = \sum_{k=1}^s x_{..k}$ le nombre total d'accidents enregistré au cours des deux périodes;

c_{jk} , ($k=1, 2, \dots, s$; $j=1, 2, \dots, r$) le rapport de contrôle spécifique à la catégorie j du site k ($c_{jk} > 0$, non aléatoire et parfaitement bien déterminé par les effectifs d'accidents observés, avant et après, dans le site de contrôle associé au site k);

Note présentée par Robert FORTET.

p_{jk} ($0 < p_{jk} < 1$) la probabilité pour qu'un accident qui se produit dans le site expérimental k avant la mesure soit du type j , $\sum_{j=1}^r p_{jk} = 1$;

$$c_k = \sum_{j=1}^r c_{jk} p_{jk}, \text{ le coefficient moyen de contrôle au site } k.$$

On suppose que les v.a. (X_{1k}, X_{2k}) , ($k=1, 2, \dots, s$) sont indépendants et suivent chacun une distribution multinomiale :

$$\mathcal{M}(x_{\cdot k}; q_{11k}(\theta_k, P_k), \dots, q_{1rk}(\theta_k, P_k), q_{21k}(\theta_k, P_k), \dots, q_{2rk}(\theta_k, P_k))$$

modèle 1 :

$$q_{1jk}(\theta_k, P_k) = \frac{p_{jk}}{1 + \theta_k c_k}, \quad q_{2jk}(\theta_k, P_k) = \frac{\theta_k c_{jk} p_{jk}}{1 + \theta_k c_k}, \quad (j=1, 2, \dots, r)$$

modèle 2 :

$$q_{1jk}(\theta_k, P_k) = \frac{p_{jk}}{1 + \theta_k c_k}, \quad q_{2jk}(\theta_k, P_k) = \frac{\theta_k c_k p_{jk}}{1 + \theta_k c_k}, \quad (j=1, 2, \dots, r)$$

avec $P_k = (p_{1k}, p_{2k}, \dots, p_{rk})^T \in \mathbb{R}^r$ et θ_k ($0 < \theta_k$) le paramètre effet de la mesure au site expérimental k . On propose ci-dessous un test d'adéquation pour décider de l'existence ou non de variations entre les différents effets θ_k . On considère pour cela :

- soit la statistique

$$\hat{A}_{s,r}^2 = \sum_{k=1}^s \sum_{j=1}^r \sum_{t=1}^2 \frac{(X_{tjk} - x_{\cdot k} q_{tjk}(\hat{\theta}, \hat{P}_k))^2}{x_{\cdot k} q_{tjk}(\hat{\theta}, \hat{P}_k)}$$

où $\hat{\theta}$ et \hat{P}_k représentent, sous l'hypothèse H_0 , les estimateurs de θ et P_k déduits de l'estimateur du maximum de vraisemblance (EMV) de $\Theta = (\theta, P_1^T, \dots, P_s^T)^T$ basé sur l'ensemble de toutes les données (voir [3]);

- soit la statistique

$$\bar{A}_{s,r}^2 = \sum_{k=1}^s \sum_{j=1}^r \sum_{t=1}^2 \frac{(X_{tjk} - x_{\cdot k} q_{tjk}(\bar{\theta}, \bar{P}_k))^2}{x_{\cdot k} q_{tjk}(\bar{\theta}, \bar{P}_k)}$$

où $\bar{\theta}$ et \bar{P}_k représentent, sous l'hypothèse H_0 , les estimateurs de θ et P_k déduits de l'estimateur du χ^2 minimum modifié de $\Theta = (\theta, P_1^T, \dots, P_s^T)^T$ basé sur l'ensemble de toutes les données. On se propose d'étudier le comportement asymptotique ($N \rightarrow \infty$) de ces statistiques.

2. DISTRIBUTIONS ASYMPTOTIQUES DE $\hat{A}_{s,r}^2$ ET $\bar{A}_{s,r}^2$.

THÉORÈME 1. - Sous la condition (c) imposée pour $k=1, 2, \dots, s$:

(c) N et $x_{\cdot k} = x_{\cdot k}(N)$ tendent vers ∞ de sorte que $\lim_{N \rightarrow \infty} (x_{\cdot k}(N)/N) = \tau_k$;

la statistique $\hat{A}_{s,r}^2$ a même distribution asymptotique ($N \rightarrow \infty$) que la variable aléatoire

(v.a.) : $\kappa_{s,r}^2 = Z + \sum_{m=1}^{\infty} \lambda_m Z_m^2$, où Z est une v.a. de loi de χ_{sr-1}^2 , où les v.a. Z_m sont i.i.d. de loi $\mathcal{N}(0; 1)$, indépendantes de Z , et où les λ_m ($0 \leq \lambda_m \leq 1$) peuvent dépendre de θ et des p_{jk} ($k=1, 2, \dots, s; j=1, 2, \dots, r$).

THÉORÈME 2. - Sous la condition (c) du théorème 1, la statistique $\bar{A}_{s,r}^2$ a pour distribution asymptotique ($N \rightarrow \infty$) une loi de χ_{sr-1}^2 .

Distribution asymptotique de certaines statistiques

403

La condition (c) traduit le comportement du nombre $x_{.k}$ d'accidents observés au cours des deux périodes sur le site k quand le nombre total N d'accidents enregistrés sur l'ensemble des sites tend vers l'infini. Les résultats sont obtenus dans le cadre du modèle 1 et utilisent certaines techniques de [4]. Des calculs similaires montrent que le modèle 2 donne les mêmes conclusions. On pose :

$$Q^T = (Q_1^T, \dots, Q_k^T, \dots, Q_s^T) \in \mathbb{R}^{2sr},$$

$$Q_k^T = (q_{11k}(\theta, P_k), \dots, q_{1rk}(\theta, P_k), q_{21k}(\theta, P_k), \dots, q_{2rk}(\theta, P_k)) \in \mathbb{R}^{2r},$$

$$\beta_k^T = (p_{1k}, \dots, p_{r-1k}), \quad (k=1, 2, \dots, s);$$

$$\Phi = (\theta, \beta_1^T, \dots, \beta_k^T, \dots, \beta_s^T)^T$$

g l'application de $\mathbb{R}^{1+s(r-1)}$ dans \mathbb{R}^{1+sr} qui à Φ associe $g(\Phi) = \Theta$

$$\Pi^T = (\Pi_1^T, \dots, \Pi_k^T, \dots, \Pi_s^T) \in \mathbb{R}^{2sr}$$

où

$$\Pi_k^T = (\pi_{11k}(\theta, \beta_k), \dots, \pi_{1rk}(\theta, \beta_k), \pi_{2rk}(\theta, \beta_k), \dots, \pi_{2rk}(\theta, \beta_k)) \in \mathbb{R}^{2r}$$

et $\pi_{ijk}(\theta, \beta_k) = q_{ijk}(\theta, P_k)$;

D_0 la matrice de taille $(2sr, 1+s(r-1))$ dont chaque élément est donné par l'expression $\pi_{ijk}^{-1/2} (\partial \pi_{ijk} / \partial \Phi_m)$ évaluée au point $\Phi = \Phi^0$, où Φ^0 est la vraie valeur du paramètre Φ ;

$Y \in \mathbb{R}^{2sr}$ le v. a. tel que $Y_{ijk} = [X_{ijk} - x_{.k} \pi_{ijk}^0] / [x_{.k} \pi_{ijk}^0]^{1/2}$ avec $\pi_{ijk}^0 = \pi_{ijk}(\Phi^0)$.

Le point (θ, P_k) [resp. (θ, β_k)] est déduit de Θ (resp. Φ) par l'intermédiaire d'une application laissant invariant θ et transformant $(P_1^T, \dots, P_k^T, \dots, P_s^T)^T$ [resp. $(\beta_1^T, \dots, \beta_k^T, \dots, \beta_s^T)^T$] en P_k^T (resp. β_k^T). Afin d'alléger les notations, on utilise dans la suite $q_{ijk}(\Theta)$ pour $q_{ijk}(\theta, P_k)$ et $\pi_{ijk}(\Phi)$ pour $\pi_{ijk}(\theta, \beta_k)$.

LEMME. — Le v. a. Y est asymptotiquement ($N \rightarrow \infty$) normalement distribué de moyenne nulle et de matrice de variance-covariance $I_{2sr} - \Pi^{1/2} \Pi^{1/2T}$, où $\Pi^{1/2}$ est le vecteur à $2sr$ dimensions de composantes les $(\pi_{ijk}^0)^{1/2}$ et I_{2sr} la matrice identité de \mathbb{R}^{2sr} .

2.1. Démonstration du théorème 2. — A l'aide de l'application g , on ramène la démonstration au point $\Phi \in \mathbb{R}^{1+s(r-1)}$. L'estimateur $\tilde{\Phi}$ de Φ obtenu avec la méthode du χ^2 minimum modifié est solution du système d'équations :

$$\sum_{k=1}^s \sum_{j=1}^r \sum_{i=1}^2 \frac{(X_{ijk} - x_{.k} \pi_{ijk})}{\pi_{ijk}} \frac{\partial \pi_{ijk}}{\partial \Phi_m} = 0; \quad m=1, 2, \dots, 1+s(r-1).$$

Par différentiation du membre de gauche de l'égalité ci-dessus au voisinage de Φ^0 et puisque la matrice $D = \Lambda^{1/2} D_0$ [où $\Lambda^{1/2} = \text{diag} (\Lambda_1^{1/2}, \Lambda_2^{1/2}, \dots, \Lambda_s^{1/2})$ est une matrice bloc-diagonale de taille $2sr$ avec $\Lambda_k^{1/2} = (x_{.k})^{1/2} I_{2r}$] est de plein rang, on a

$$(\tilde{\Phi} - \Phi^0) = (D^T D)^{-1} D^T Y + o_p(1).$$

Soit $\tilde{A} \in \mathbb{R}^{2rs}$ le v. a. défini par $\tilde{A}_{ijk} = [X_{ijk} - x_{.k} \pi_{ijk}(\tilde{\Phi})] / [x_{.k} \pi_{ijk}(\tilde{\Phi})]^{1/2}$. En utilisant le développement de Taylor en probabilité de $\pi_{ijk}(\tilde{\Phi})$ à l'ordre un au voisinage de Φ^0 , l'inégalité de Tchebycheff et l'expression de $(\tilde{\Phi} - \Phi^0)$ ci-dessus, on montre que :

$$\tilde{A} = (I_{2sr} - D(D^T D)^{-1} D^T) Y + o_p(1).$$

Il faut donc connaître la distribution asymptotique de $A^T A$ où

$$A = (I_{2sr} - D(D^T D)^{-1} D^T) Y.$$

En vertu du lemme on montre que $A^T A$ est distribuée comme $\sum_{m=1}^{2sr} \lambda_m Z_m^2$ où les v. a. Z_m sont i.i.d. de loi $\mathcal{N}(0; 1)$ et où les λ_m sont les valeurs propres de la matrice

$$\tilde{\Omega} = I_{2sr} - \Pi^{1/2} \Pi^{1/2T} - D(D^T D)^{-1} D^T.$$

Or $\tilde{\Omega}$ admet $(sr-1)$ valeurs propres égales à 1 et $(sr+1)$ valeurs propres égales à 0.

2.2. *Démonstration du théorème 1.* — On note à une constante additive près $L(\Phi) = \sum_{k=1}^s \sum_{j=1}^r \sum_{t=1}^2 x_{ijk} \log(\pi_{ijk}(\theta, \beta_k))$ la log-vraisemblance relativement au paramètre Φ et J la matrice d'information de Fisher associée et définie par

$$J_{ml} = -E[(\partial^2 L(\Phi)/\partial\Phi_m \partial\Phi_l)_{(\Phi=\Phi^0)}].$$

En calculant le terme entre crochets, on montre que :

$$J = D^T D + J^*$$

avec

$$J_{ml}^* = -E \left[\sum_{k=1}^s \sum_{j=1}^r \sum_{t=1}^2 \frac{X_{tjk}}{\pi_{ijk}(\Phi^0)} \left(\frac{\partial^2 \pi_{ijk}}{\partial\Phi_l \partial\Phi_m} \right)_{(\Phi=\Phi^0)} \right].$$

D'autre part, l'EMV $\hat{\Phi}$ de Φ est obtenu en résolvant le système d'équations :

$$\left(\frac{\partial L}{\partial\Phi_m} \right) (\hat{\Phi}) = 0, \quad m = 1, 2, \dots, 1+s(r-1).$$

En utilisant ces équations et la décomposition ci-dessus de la matrice J , on montre que

$$(\hat{\Phi} - \Phi^0) = (D^T D + J^*)^{-1} D^T Y + o_p(1).$$

Si on note maintenant $\hat{A} \in \mathbb{R}^{2sr}$ le v. a. obtenu en remplaçant $\tilde{\Phi}$ par $\hat{\Phi}$ dans l'expression de \tilde{A} , on montre de façon similaire que :

$$\hat{A} = [I_{2sr} - D(D^T D + J^*)^{-1} D^T] Y + o_p(1).$$

La statistique $\hat{A}_{s,r}^2$, a même distribution asymptotique que $\sum_{m=1}^{2sr} \lambda_m Z_m^2$ où les v. a. Z_m sont i.i.d. de loi $\mathcal{N}(0; 1)$ et où les λ_m sont les valeurs propres de la matrice

$$\tilde{\Omega} = I_{2sr} - \Pi^{1/2} \Pi^{1/2T} - D(D^T D + J^*)^{-1} D^T - D(D^T D + J^*)^{-1} J^* (D^T D + J^*)^{-1} D^T.$$

On démontre que $\tilde{\Omega}$ possède $(sr-1)$ valeurs propres égales à 1, s valeurs propres nulles et $(1+s(r-1))$ valeurs propres comprises entre 0 et 1.

Note remise le 16 mars 1993, acceptée le 22 avril 1993.

RÉFÉRENCES BIBLIOGRAPHIQUES

- [1] J.C. TANNER, A problem in the combination of accident frequencies. *Biometrika*, n° 45, 1958, p. 331-342.
- [2] E. HAUER, Comparison group in road safety studies: An analysis. *Accid. Anal. and Prev.*, 23, n° 6, 1991, p. 609-622.
- [3] A. N'GUESSAN et C. LANGRAND, Une méthode d'estimation de l'effet moyen d'une mesure de sécurité routière, *XXIV^{es} Journées de Statistique*, Bruxelles, 1992.
- [4] H. CHERNOFF et E. F. LEHMANN, The use of Maximum likelihood estimates in χ^2 -tests for goodness of fit, *Ann. Math. Statistics*, 25, 1954, p. 579-586.

Laboratoire de Statistique et Probabilités, UFR de Mathématiques Pures et Appliquées, Université des Sciences et Technologies de Lille, 59655 Villeneuve-d'Ascq Cedex, France.

Annexe A09 :

*Technical Report Center for Research on Transportation, Université de Montréal, **CRT-2003-12, 1-18 (2003)***

Constrained estimation of a road safety covariance matrix using Schur complements

**PUBLICATION
CRT-2003-12**

**CONSTRAINED ESTIMATION OF A
ROAD SAFETY COVARIANCE MATRIX
USING SCHUR COMPLEMENTS**

by

Assi N'Guessan¹

This research was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC).

¹. École Polytechnique Universitaire de Lille et Laboratoire de Mathématiques Appliquées, CNRS, FRE 2222. Université de Lille 1, 59655 Villeneuve d'Ascq Cedex, France (Assi.Nguessan@polytech-lille.fr)

Abstract

One is interested in the asymptotic matrix of the constrained m.l.e. (maximum likelihood estimator) of a multidimensional parameter connected to a road safety measure. He achieves the explicit asymptotic covariance matrix of the constrained m.l.e. and the explicit form of certain estimated standard errors through Schur complements technical inversion method. This approach also provides a powerful tool to compute formally the explicit estimated between and within variances covariance's matrices of block-components of the restraint m.l.e. ease of implementation without matrix program inversion is an attractive aspect of this new approach. Some examples of formal estimation are discussed to back up the method.

Keywords: Road safety measure, multinomial model, constrained maximum likelihood, Schur complements, asymptotic covariance's matrix, formal estimation.

Résumé

On s'intéresse à la matrice de covariances asymptotique du m.l.e. (estimateur du maximum de vraisemblance) sous contraintes d'un paramètre multidimensionnel lié à une mesure de sécurité routière. Il fait appel à une technique formelle d'inversion matricielle appelée compléments de Schur pour obtenir explicitement la matrice de covariances asymptotique du m.l.e. sous contraintes ainsi que l'estimation formelle de certains écart-types. Cette approche est un outil puissant qui permet également d'estimer la structure analytique des matrices de variances covariances inter et intra de bloc d'éléments du m.l.e. sous contraintes. La mise en œuvre simple et sans obligation de programmes informatiques d'inversion matricielle de cette nouvelle approche le rend d'autant plus intéressante. Quelques exemples d'estimation formelle sont fournis pour étayer la méthode.

Mots-clés: mesure de sécurité routière, modèle multinomial, maximum de vraisemblance sous contraintes, compléments de Schur, matrice de covariances asymptotique, estimation formelle.

1. Introduction and review

In many estimation problems, particularly in multivariate statistics, the parameters in a model are not functionally independent, that is, there exist certain relations connecting them. Nevertheless, it may be undesirable to eliminate redundancies for reasons of symmetry, or because the natural interpretation of the parameters would be lost, or because of mathematical untractability. Statistical aspects of estimation in restricted parameter spaces, asymptotic results for constrained estimation and related problems have been extensively discussed (*see for example* Harville 1997, Lee 1979, Van Eeden 1996).

More precisely, Aitchison and Silvey (1958) consider the case of i.i.d. observations. They prove, under certain conditions, asymptotic existence, consistency and asymptotic normality of a solution of the likelihood equations. Silvey (1959) deals with the i.i.d. case, but allowing the information matrix to be singular. He uses the alternative approach to proof of consistency due to Wald (1949). Crowder (1984) considers the case of non i.i.d. observations when the parameters are constrained and uses inequality constraints as an application. His asymptotic results and formal methods stem from those of Aitchison and Silvey (1958) and Silvey (1959). Also, more emphasis is placed in Crowder's (1984) paper on the precise manner in which the constraints may be used to construct a suitable nonsingular version of the information matrix and the connection with parameter identifiability. Neuenschwander and Flury (1997) give an alternative proof of a theorem by Silvey (1959) for the asymptotic covariance matrix of the maximum likelihood estimation (m.l.e.) under different types of constraints in the parameter space. They show how the different types of constraints can be treated in a unified way and applied to a variety of statistical estimation problems.

N'Guessan et al (2001) deal with the constrained estimation of the average effect of a road safety measure using before-after accident data tables. The latter authors consider independent non i.d. multinomial observations under linear constraints and are interested in the existence and convergence of the restraint m.l.e. connected to their model via simulation studies and mean square error. But they don't include the accuracy of the constrained m.l.e. i.e. the asymptotic covariances matrix. Also, their method's drawback is that both parameters's space dimension and constraints's number increase with the size of before-after data tables used and thereafter the dimension of Fisher's information matrix to invert. One knows that this situation sometimes affects the numerical inversion results and impair the accuracy of standard errors.

The main motivation here is to make their paper complete by investigating the asymptotic variances covariances matrix of their m.l.e. in order to obtain an explicit form of estimated standard errors without calling numerical matrix inversion programm. The formal methods combine Fisher's information matrix related to model below, the notion of Schur complements for the inversion of partitioned matrix (Ouellette 1981; N'Guessan 2003a) and the results of Crowder (1984).

The constrained multinomial model and notation in section 2, used to share the total accident number in each experimental site at the same time between both periods and among the different accident types, stem from those of N'Guessan et al (2001). In section 3 some technical intermediate results are derived to support the foregoing theorem and its corollaries. We particularly use the model and constraints below to derive that the Fisher's information matrix is nonsingular. Section 4 is exclusively centred on the main result. its proof is given in the appendix. In section 5 some formal estimation examples are discussed to back up the method. Some concluding remarks are displayed in section 6.

2. Statistical Model and notation

The statistical models used in the before-after estimation of a road safety measure efficiency heavily depend on the available data and on the targets set along with the measure (see for example Danielson (1986, 1988), Hauer (1997), Barry et al (1999), Abdel-Aty et al (2000), Gaudry and Lassarre (2000), Beenstock et al (2001)). Some factors such as traffic evolution or weather conditions, etc... must be considered when interpreting the estimation results; if when studying other data (such as the variations of the accident number on a similar site where the measure is not applied) one can notice that the accident number is multiplied by a c factor (control or evolution coefficient), that factor must be taken into account when appreciating the measure's efficiency. When a measure is simultaneously applied to several sites, considering the use of that very factor whatever the site is generally not realistic because of site location for example. That is why Tanner (1958) introduced the notion of control area and defined a control coefficient for each site using the ratio of the accident number after to before in the control area. Let's now suppose that, the measure having been applied to several sites, the total number of accidents for both periods in each target site is available. Then let's retrospectively focus on the measure's average effect in each site on a certain number of accidents (fatal accidents, seriously injured people, slightly injured people, material damage, etc...). Such an approach introduces a third dimension (accident typology) to Tanner's pattern and statistics and imposes the estimation of the measure's average effect through before-after tables of accident data. With that in mind, N'Guessan (1993), N'Guessan and Langrand (1993) generalised Tanner's approach to before-after tables of accident with control area. In their papers they supposed that the same road safety measure had been applied to several sites, each one including mutually exclusive different accident types and linked to a specific control area. Recently, N'Guessan et al (2001) are interested in the constrained m.l.e. of the average effect of a road safety measure in presence of random control area data and focus their attention on its numerical convergence via the value of the mean square error for before-after tables. They deal with the multidimensional estimation of the average effect of road safety measure applied to s ($s > 0$) entities (sites), each one including r ($r > 1$) mutually exclusive different accident types and linked to a specific control area. They consider a collection

$$(Y_{11}, Y_{21}), (Y_{12}, Y_{22}), \dots, (Y_{1s}, Y_{2s})$$

of s independent couples of $2r$ - dimensional random vector and assume that, for fixed k , each couple (Y_{1k}, Y_{2k}) has multinomial distribution denoted by $\mathcal{M}(n_k; \Pi_{1k}(\alpha), \Pi_{2k}(\alpha))$, where

$$\Pi_{tk}(\alpha) = (\pi_{t1k}(\alpha), \pi_{t2k}(\alpha), \dots, \pi_{trk}(\alpha)),$$

n_k ($n_k > 0$) is a given integer, $\alpha = (\beta_0, \beta^T)^T \in \mathbb{R}^{1+sr}$ a unknown vector parameter, β_0 ($\beta_0 > 0$) a scalar, $\beta = (\beta_1^T, \dots, \beta_s^T)^T \in \mathbb{R}^{sr}$ with $\beta_k = (\beta_{1k}, \beta_{2k}, \dots, \beta_{rk})^T \in \mathbb{R}^r$ and $\beta_{jk} > 0$. The components $\pi_{trk}(\alpha)$ of the $r \times 1$ vector probability $\Pi_{tk}(\alpha)$ are given by a continuous differentiable functional link $g : \mathbb{R}^{1+sr} \mapsto]0, 1[$ so that $\pi_{trk}(\alpha) = g(z_k, \alpha)$ and

$$g(z_k, \alpha) = \begin{cases} \frac{\beta_{jk}}{1 + \beta_0 \langle z_k, \beta_k \rangle} & t=1; j=1, 2, \dots, r \\ \frac{z_{jk} \beta_0 \beta_{jk}}{1 + \beta_0 \langle z_k, \beta_k \rangle} & t=2; j=1, 2, \dots, r \end{cases} \quad (1)$$

where \langle, \rangle is the usual inner product, $z_k = (z_{1k}, z_{2k}, \dots, z_{rk})^T$ a given extra set data connected to the control area of each couple k . The latter authors are interested in the m.l.e. $\hat{\alpha}$ of α and its numerical consistency under the restraints $h_k(\alpha) = 0$ where

$$h_k(\alpha) = \langle \mathbf{1}_r, \beta_k \rangle - 1; k = 1, 2, \dots, s \quad (2)$$

with $\mathbf{1}_r = (1, \dots, 1)^T \in \mathbb{R}^r$, the vector of unity. The construction mechanism of the statistical patten above is displayed in N'Guessan et al (2001). Let's indicate that in road safety modeling β_0 represents the average effect when combining all data tables and vector β_k , for fixed k , denotes the probability distribution (or risk accident) connected to extra data z_k of control area k such that $\sum_{j=1}^r \beta_{jk} = 1$. This is the natural and practical interpretation of the constraints (2).

In the following we note $H_\alpha = (H_1, H_2)$ the constraints matrix with $H_1 = \frac{\partial h}{\partial \beta_0}$ a $s \times 1$ matrix, $H_2 = \frac{\partial h}{\partial \beta}$ a $s \times (sr)$ matrix, $h = (h_1, h_2, \dots, h_s)$, $L(\alpha)$ the log-likelihood function based on the s couples, $J_\alpha = E(-\frac{\partial^2 L}{\partial \alpha \partial \alpha})$, the $(1 + sr) \times (1 + sr)$ information matrix evaluated at different points on the line segment joining α^0 (the true vector parameter) and $\hat{\alpha}$. It can be shown under regularity conditions, that $\hat{\alpha}$ is asymptotically normally distributed and its covariance matrix strongly depends on the fact that both J_α and $H_\alpha J_\alpha^{-1} H_\alpha^T$ are nonsingular. Technical lemmas below elaborate these points. The foregoing theorem concerns the main result and gives the formal expressions of the components of the asymptotic covariance matrix connected to model (1) and constraints (2). We achieve those explicit forms through the notion of Schur complements well presented by Ouellette (1981). She also presents an unified traitment of standard results on inverses, and g-inverses of partitioned matrices using Schur complements and discribes some areas of statistics in which this notion arises. It is the same for spatial prediction and temporal backcasting (Le et al. 2001) or log-linear model estimators (Rivest and Lévesque, 2001) or longitudinal data models (Frees, 2001) and so on.

3. Technical intermediate results

lemma 1 Let's note $V_{\alpha,k} = \frac{\beta_0}{1 + \beta_0 \langle z_k, \beta_k \rangle} (z_{1k}, z_{2k}, \dots, z_{rk})^T$ a $r \times 1$ vector and $\Omega_{\alpha,k} = \text{diag}(\frac{1 + \beta_0 z_{1k}}{\beta_{1k}}, \frac{1 + \beta_0 z_{2k}}{\beta_{2k}}, \dots, \frac{1 + \beta_0 z_{rk}}{\beta_{rk}})$ a $r \times r$ matrix. Then for $k = 1, 2, \dots, s$:

i) $\|z_k\|_{\Omega_{\alpha,k}^{-1}}^2 < \frac{\langle z_k, \beta_k \rangle}{\beta_0}$; ii) $\|V_{\alpha,k}\|_{\Omega_{\alpha,k}^{-1}}^2 < \frac{\beta_0 \langle z_k, \beta_k \rangle}{(1 + \beta_0 \langle z_k, \beta_k \rangle)^2} < 1$.

Proof. i) Using the fact that $\frac{x}{1+x} < 1$ for all $x > 0$, one obtains after matrix manipulations,

$$\|z_k\|_{\Omega_{\alpha,k}^{-1}}^2 < \frac{1}{\beta_0} (\sum_{m=1}^r \beta_{mk} z_{mk}) = \frac{\langle z_k, \beta_k \rangle}{\beta_0}.$$

ii) This part follows from i) and the relation between the vectors V_k and z_k ■

lemma 2 Let's note $B_{\alpha,k} = \gamma_k (\Omega_{\alpha,k} - V_{\alpha,k} V_{\alpha,k}^T)$ with $\gamma_k = \frac{n_k}{1 + \beta_0 \langle z_k, \beta_k \rangle}$, Then

i) For $k = 1, 2, \dots, s$;

$$B_{\alpha,k}^{-1} = \gamma_k^{-1} (\Omega_{\alpha,k}^{-1} + t_k \Omega_{\alpha,k}^{-1} V_{\alpha,k} V_{\alpha,k}^T \Omega_{\alpha,k}^{-1}); \quad t_k = (1 - \|V_{\alpha,k}\|_{\Omega_{\alpha,k}^{-1}}^2)^{-1}$$

ii) $\|U_\alpha\|_{B_\alpha^{-1}}^2 < \tau_\alpha$; where $B_\alpha = \text{Bloc-diag}(B_{\alpha,1}, \dots, B_{\alpha,s})$, $U_\alpha = (U_{\alpha,1}^T, \dots, U_{\alpha,s}^T)^T$ a $(sr) \times 1$ vector with $U_{\alpha,k} = \frac{\gamma_k}{\beta_0} V_{\alpha,k}$ and $\tau_\alpha = \sum_{k=1}^s \frac{\gamma_k \langle z_k, \beta_k \rangle}{\beta_0 (1 + \beta_0 \langle z_k, \beta_k \rangle)}$.

Proof. i) Lemma 1 implies that $t_k > 0$ for all k , so $B_{\alpha,k}^{-1}$ follows from Woodburry (1950) formula (see also Press et al (1992) pp. 73-78, Ouellette 1981) or Sherman-Morrison formula (1950).

ii) Matrix manipulations and relations on the one hand between vectors $U_{\alpha,k}$ and $V_{\alpha,k}$ and on the other hand matrices $B_{\alpha,k}$ and $\Omega_{\alpha,k}$ imply that

$$\|U_\alpha\|_{B_\alpha^{-1}}^2 = \frac{1}{\beta_0^2} \sum_{k=1}^s \frac{\gamma_k \|V_{\alpha,k}\|_{\Omega_{\alpha,k}^{-1}}^2}{1 - \|V_{\alpha,k}\|_{\Omega_{\alpha,k}^{-1}}^2} \quad (3)$$

Using lemma 1, one obtains

$$\|U_\alpha\|_{B_\alpha^{-1}}^2 < \sum_{k=1}^s \frac{\gamma_k(1 + \beta_0 \langle z_k, \beta_k \rangle)}{\beta_0^2} \|V_{\alpha,k}\|_{\Omega_{\alpha,k}^{-1}}^2 < \tau_\alpha \quad (4)$$

■ Now we are ready to deal with the inversion of J_α and $H_\alpha J_\alpha^{-1} H_\alpha^T$. Using model (1) above and after some derivations and matrix manipulations one obtains:

$$J_\alpha = \begin{pmatrix} \tau_\alpha & U_\alpha^T \\ U_\alpha & B_\alpha \end{pmatrix}.$$

Following Ouellette (1981), and N'Guessan (2003a, appendix B) we note (J_α/B_α) (resp. (J_α/τ_α)) the Schur complement of B_α (resp. τ_α) in J_α and defined by $(J_\alpha/B_\alpha) = \tau_\alpha - U_\alpha^T B_\alpha^{-1} U_\alpha$ (resp. $(J_\alpha/\tau_\alpha) = B_\alpha - U_\alpha \tau_\alpha^{-1} U_\alpha^T$).

lemma 3

- (i) $(J_\alpha/B_\alpha) > 0$
(ii) (J_α/τ_α) is nonsingular and

$$(J_\alpha/\tau_\alpha)^{-1} = B_\alpha^{-1} + (J_\alpha/B_\alpha)^{-1} B_\alpha^{-1} U_\alpha U_\alpha^T B_\alpha^{-1} \quad (5)$$

Proof. i) is a consequence of lemma 2 part ii)

ii) This result is derived from the combination of lemma 2 and Woodburry formula a special case of Schur complements inversion. ■

lemma 4

- i) J_α^{-1} exists and

$$J_\alpha^{-1} = \begin{pmatrix} (J_\alpha/B_\alpha)^{-1} & -(J_\alpha/B_\alpha)^{-1} U_\alpha^T B_\alpha^{-1} \\ -B_\alpha^{-1} U_\alpha (J_\alpha/B_\alpha)^{-1} & (J_\alpha/\tau_\alpha)^{-1} \end{pmatrix} \quad (6)$$

- ii)

$$(H_\alpha J_\alpha^{-1} H_\alpha^T)^{-1} = \Lambda_\alpha^{-1} - ((J_\alpha/B_\alpha) + \|\xi_\alpha\|_{\Lambda_\alpha^{-1}}^2)^{-1} \Lambda_\alpha^{-1} \xi_\alpha \xi_\alpha^T \Lambda_\alpha^{-1} \quad (7)$$

Proof. i) uses lemma 2 and the formal inversion of partitioned matrix via Shur complements.

ii) combines the constraints functions (2) and i). More precisely, straightforward derivations of constraints (2) show that

$$\frac{\partial h_k}{\partial \beta_m} \equiv \mathbf{1}_r^T \delta_{mk}; \quad k = 1, 2, \dots, s; m = 0, 1, 2, \dots, s \quad (8)$$

where δ_{mk} is the Kronecker symbol. So using the latter formula, part i) of lemma 4 and ii) of lemma 3 we can write:

$$\begin{aligned} H_\alpha J_\alpha^{-1} H_\alpha^T &= H_2 (J_\alpha/\tau_\alpha)^{-1} H_2^T - H_1 (J_\alpha/B_\alpha)^{-1} \xi_\alpha^T \\ &= H_2 (J_\alpha/\tau_\alpha)^{-1} H_2^T \\ &= \Lambda_\alpha + (J_\alpha/B_\alpha)^{-1} \xi_\alpha \xi_\alpha^T \end{aligned} \quad (9)$$

The same arguments used for part i) of lemma 2 are applied. ■

4. Main result

The restraints m.l.e of α_0 , provided it exists, is the vector $\hat{\alpha}$ which satisfies

$$\left(\frac{\partial L}{\partial \alpha}\right)_{\hat{\alpha}} + H_{\hat{\alpha}}^T \hat{\lambda} = 0, \quad h(\hat{\alpha}) = 0 \quad (10)$$

where λ is a vector of s Lagrange multipliers. Let $M_{\alpha} = J_{\alpha}^{-1} \left(J_{\alpha} + \left(\frac{\partial^2 L}{\partial \alpha \partial \alpha} \right) \right)$ a $(1 + sr) \times (1 + sr)$ matrix, $R_{\alpha} = H_2(J_{\alpha}/\tau_{\alpha})^{-1} H_2^T$ a $s \times s$ matrix, $\Phi_{\alpha} = J_{\alpha}^{-1} H_{\alpha}^T R_{\alpha}^{-1}$ a $(1 + sr) \times s$ matrix, $\xi_{\alpha} = H_2 B_{\alpha}^{-1} U_{\alpha}$ a $s \times 1$ vector, $\Lambda_{\alpha} = \text{diag}(\| \mathbf{1}_r \|_{B_{\alpha,1}^{-1}}^2, \| \mathbf{1}_r \|_{B_{\alpha,2}^{-1}}^2, \dots, \| \mathbf{1}_r \|_{B_{\alpha,s}^{-1}}^2)$ a $s \times s$ matrix with $\| \mathbf{1}_r \|_{B_{\alpha,k}^{-1}}^2 = \mathbf{1}_r^T B_{\alpha,k}^{-1} \mathbf{1}_r$.

Theorem 4.1 *Suppose that*

Assumption A1 : $(\hat{\alpha}, \hat{\lambda})$ is a consistent solution of the latter equation

Assumption A2 : $M_{\hat{\alpha}} \xrightarrow{\text{Proba}} 0$, $\Phi_{\hat{\alpha}} - \Phi_{\alpha_0} \xrightarrow{\text{Proba}} 0$, $|\Phi_{\alpha_0}| < \infty$.

Assumption A3 : $\left(\frac{\partial L}{\partial \alpha}\right)_{\alpha_0} \stackrel{\text{Distr.}}{\sim} \mathcal{N}(0, J_{\alpha_0})$

Then the asymptotic components of the covariance matrix of $\hat{\alpha}$ are:

$$\sigma^2(\hat{\beta}_0) = (J_{\alpha}/B_{\alpha})^{-1} - (J_{\alpha}/B_{\alpha})^{-2} \| \xi_{\alpha} \|_{R_{\alpha}^{-1}}^2$$

$$\text{Cov}(\hat{\beta}, \hat{\beta}_0) = -(J_{\alpha}/B_{\alpha})^{-1} [B_{\alpha}^{-1} U_{\alpha} - (J_{\alpha}/\tau_{\alpha})^{-1} H_2^T R_{\alpha}^{-1} \xi_{\alpha}]$$

$$\text{Var}(\hat{\beta}) = (J_{\alpha}/\tau_{\alpha})^{-1} - (J_{\alpha}/\tau_{\alpha})^{-1} H_2^T R_{\alpha}^{-1} H_2 (J_{\alpha}/\tau_{\alpha})^{-1}$$

Remark

The proof of the theorem and of the corollaries below are given in appendix. Let's note that assumption (A1) seems natural and the numerical behaviour is discussed in N'Guessan et al (2001). Theoretical behaviour are displayed for example in Crowder (1984). Assumption (A2) is designed to support $D_{\alpha_0}^{-1} D_{\hat{\alpha}} \xrightarrow{\text{Proba}} I_{1+s(r+1)}$ in the proof where alternatives are possible. Matrices $D_{\alpha_0}^{-1}$ and $D_{\hat{\alpha}}$ are also defined in appendix. The behaviour in (A3) is for instance discussed in Crowder (1975) and is used to compute matrix Σ_{α} . One also notes that $H_{\alpha} J_{\alpha}^{-1} H_{\alpha}^T = R_{\alpha}$. So using lemma 4, R_{α} is nonsingular and

$$R_{\alpha}^{-1} = \Lambda_{\alpha}^{-1} - \sigma^2(\hat{\beta}_0) \Lambda_{\alpha}^{-1} \xi_{\alpha} \xi_{\alpha}^T \Lambda_{\alpha}^{-1} \quad (11)$$

Corollary 4.2

$$\sigma^2(\hat{\beta}_0) = ((J_{\alpha}/B_{\alpha}) + \| \xi_{\alpha} \|_{\Lambda_{\alpha}^{-1}}^2)^{-1} \quad (12)$$

where

$$\| \xi_{\alpha} \|_{\Lambda_{\alpha}^{-1}}^2 = \sum_{k=1}^s \frac{t_k^2 \gamma_k^2}{n_k^2} \frac{\Delta_{\alpha,k}^2}{\| \mathbf{1}_r \|_{B_{\alpha,k}^{-1}}^2} \quad (13)$$

with

$$\Delta_{\alpha,k} = z_k^T \Omega_{\alpha,k}^{-1} \mathbf{1}_r \quad (14)$$

and

$$\| \mathbf{1}_r \|_{B_{\alpha,k}^{-1}}^2 = \gamma_k^{-1} \left[\text{trace}(\Omega_{\alpha,k}^{-1}) + t_k \left(\frac{\beta_0 \gamma_k \Delta_{\alpha,k}}{n_k} \right)^2 \right]. \quad (15)$$

Corollary 4.3 Under the same assumptions than theorem above, the $r \times 1$ covariance vector between the restraints m.l.e. of β_k and β_0 is

$$\text{Cov}(\hat{\beta}_k; \hat{\beta}_0) = S_{\alpha,k} z_k \quad (16)$$

where $S_{\alpha,k}$ is the $r \times r$ matrix as follows

$$S_{\alpha,k} = -\frac{t_k \sigma^2(\hat{\beta}_0)}{n_k} \left[\gamma_k \Omega_{\alpha,k}^{-1} - \frac{1}{\|\mathbf{1}_r\|_{B_{\alpha,k}^{-1}}^2} \Omega_{\alpha,k}^{-1} (\mathbf{1}_{r,r} + \rho_{\alpha,k} z_k \mathbf{1}_r^T) \Omega_{\alpha,k}^{-1} \right] \quad (17)$$

with

$$\rho_{\alpha,k} = \frac{t_k \Delta_{\alpha,k} \beta_0^2 \gamma_k^2}{n_k^2} \quad (18)$$

for $k = 1, 2, \dots, s$ and $\mathbf{1}_{r,r} = \mathbf{1}_r \mathbf{1}_r^T$ is the $r \times r$ matrix with unity.

Corollary 4.4 Under the same assumptions than theorem above, the $r \times r$ between covariance matrix of $\hat{\beta}_k$ and $\hat{\beta}_m$ is

$$\text{Cov}(\hat{\beta}_k; \hat{\beta}_m) = J_{k,m} - \sum_{j=1}^s \left[\left(\sum_{i=1}^s A_{i,j} J_{k,i} \right) \mathbf{1}_{r,r} \right] J_{j,m} \quad (19)$$

$k, m = 1, 2, \dots, s$. Where $J_{k,k}$ (resp. $J_{k,m}$ for $m \neq k$) is the Schur complement of (J_α/B_α) in the $(1+r) \times (1+r)$ matrix

$$\begin{bmatrix} (J_\alpha/B_\alpha) & U_{\alpha,k}^T B_{\alpha,k}^{-1} \\ B_{\alpha,k}^{-1} U_{\alpha,k} & B_{\alpha,k}^{-1} \end{bmatrix}, \quad (\text{resp. } \begin{bmatrix} (J_\alpha/B_\alpha) & U_{\alpha,m}^T B_{\alpha,m}^{-1} \\ B_{\alpha,k}^{-1} U_{\alpha,k} & 0 \end{bmatrix}),$$

and

$$A_{k,m} = \frac{\delta_{k,m}}{\|\mathbf{1}_r\|_{B_{\alpha,k}^{-1}}^2} - \frac{\sigma^2(\hat{\beta}_0) t_k t_m \gamma_k \gamma_m \Delta_{\alpha,k} \Delta_{\alpha,m}}{n_k n_m \|\mathbf{1}_r\|_{B_{\alpha,k}^{-1}}^2 \|\mathbf{1}_r\|_{B_{\alpha,m}^{-1}}^2} \quad (20)$$

If $k = m$ then we obtain the within $r \times r$ covariance of $\hat{\beta}_k$ as follows

$$\text{Cov}(\hat{\beta}_k; \hat{\beta}_k) = J_{k,k} - \sum_{j=1}^s \left[\left(\sum_{i=1}^s A_{i,j} J_{k,i} \right) \mathbf{1}_{r,r} \right] J_{j,k} \quad (21)$$

$k = 1, 2, \dots, s$.

5. Examples

5.1. example 1: $s = 1$ and $r = 2$

We suppose here that the number of experimental site is equal to one and there are two accident types with n_k the total number of accident of the site and $z_k = (z_{1k}, z_{2k})^T$ the data from the control area. The parameter is defined by $\alpha = (\beta_0, \beta_k)^T \in \mathbb{R}^3$ and $\beta_k = (\beta_{1k}, \beta_{2k})^T \in \mathbb{R}^2$ with $\beta_0 > 0$, $\beta_{jk} > 0$, and $\beta_{1k} + \beta_{2k} = 1$. So the dimension of the parameter space is three and we have one constraint. Following asymptotic results (Crowder 1984 or Neuenschwander & Flury, 1997) and using model (1) and constraints

(2), and after some labouring derivations of the log-likelihood (N'Guessan et al. 2001), we show that the 4×4 constrained Fisher's information matrix is :

$$\Gamma_{\alpha} = \begin{pmatrix} \tau_{\alpha} & \frac{\gamma_k}{\beta_0} v_{1k} & \frac{\gamma_k}{\beta_0} v_{2k} & 0 \\ \frac{\gamma_k}{\beta_0} v_{1k} & b_{11}^{(k)} & b_{12}^{(k)} & 1 \\ \frac{\gamma_k}{\beta_0} v_{2k} & b_{21}^{(k)} & b_{22}^{(k)} & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

where $\tau_{\alpha} = \frac{\gamma_k^2 \langle z_k, \beta_k \rangle}{\beta_0 n_k}$, $v_{mk} = \frac{\gamma_k \beta_0 z_{mk}}{n_k}$, $\gamma_k = \frac{n_k}{1 + \beta_0 \langle z_k, \beta_k \rangle}$, and

$$\begin{pmatrix} b_{11}^{(k)} & b_{12}^{(k)} \\ b_{21}^{(k)} & b_{22}^{(k)} \end{pmatrix} = \gamma_k \begin{pmatrix} \frac{1 + \beta_0 z_{1k}}{\beta_{1k}} - \frac{\beta_0^2 z_{1k}^2 \gamma_k^2}{n_k^2} & -\frac{\beta_0^2 z_{1k} z_{2k} \gamma_k^2}{n_k^2} \\ -\frac{\beta_0^2 z_{2k} z_{1k} \gamma_k^2}{n_k^2} & \frac{1 + \beta_0 z_{2k}}{\beta_{2k}} - \frac{\beta_0^2 z_{2k}^2 \gamma_k^2}{n_k^2} \end{pmatrix}.$$

In order to get at the asymptotic covariance matrix of the constrained m.l.e $\hat{\alpha}$, one must obtain Γ_{α}^{-1} and take the leading 3×3 submatrix. The method proposed here allows everybody to avoid doing this inversion by using theorem 4.1 as follows;

setting

$$J_{\alpha} = \begin{pmatrix} \tau_{\alpha} & \frac{\gamma_k}{\beta_0} V_{\alpha,k}^T \\ \frac{\gamma_k}{\beta_0} V_{\alpha,k} & B_{\alpha,k} \end{pmatrix}; \quad \Omega_{\alpha,k} = \begin{pmatrix} \frac{1 + \beta_0 z_{1k}}{\beta_{1k}} & 0 \\ 0 & \frac{1 + \beta_0 z_{2k}}{\beta_{2k}} \end{pmatrix}$$

and $B_{\alpha,k} = \gamma_k [\Omega_{\alpha,k} - V_{\alpha,k} V_{\alpha,k}^T]$ a 2×2 matrix and

$$V_{\alpha,k} = \frac{\beta_0 \gamma_k}{n_k} (z_{1k}, z_{2k})^T; \quad H_2 = (1, 1),$$

and using corollary 4.2, the estimated standard error of $\hat{\beta}_0$ is

$$SE(\hat{\beta}_0) = \left[(J_{\hat{\alpha}}/B_{\hat{\alpha}}) + \|\xi_{\hat{\alpha}}\|_{\Lambda_{\hat{\alpha}}^{-1}}^2 \right]^{-1/2}$$

with

$$(J_{\hat{\alpha}}/B_{\hat{\alpha}}) = \frac{\hat{\gamma}_k^2}{n_k} \left(\frac{\langle z_k, \hat{\beta}_k \rangle}{\hat{\beta}_0} - \frac{\hat{\gamma}_k \hat{t}_k}{n_k} \|z_k\|_{\Omega_{\hat{\alpha},k}^{-1}}^2 \right); \quad \|\xi_{\hat{\alpha}}\|_{\Lambda_{\hat{\alpha}}^{-1}}^2 = \frac{\hat{\gamma}_k^2 \hat{t}_k^2 \Delta_{\hat{\alpha},k}^2}{n_k^2 \|\mathbf{1}_r\|_{B_{\hat{\alpha},k}^{-1}}^2}$$

and $\Delta_{\hat{\alpha},k} = \frac{\hat{\beta}_{1k} z_{1k}}{1 + \hat{\beta}_0 z_{1k}} + \frac{\hat{\beta}_{2k} z_{2k}}{1 + \hat{\beta}_0 z_{2k}}$. Now using the corollary 4.3, one obtains the covariances between the subvector parameter $\hat{\beta}_k$ and the first component $\hat{\beta}_0$ as follows (appendix B)

$$cov(\hat{\beta}_{mk}, \hat{\beta}_0) = \frac{t_k \sigma^2(\hat{\beta}_0)}{n_k} \left(-\frac{\gamma_k \beta_{mk} z_{mk}}{1 + \beta_0 z_{mk}} + \frac{(1 + \rho_{\alpha,k} z_{mk}) \beta_{mk} \Delta_{\alpha,k}}{(1 + \beta_0 z_{mk}) \|\mathbf{1}_r\|_{B_{\alpha,k}^{-1}}^2} \right) \quad (22)$$

for $m = 1, 2$. Now in view of the proof of corollary 4.4 (see appendix B for details) we obtain

$$Var(\hat{\beta}_k) = J_{k,k} - A_{k,k} J_{k,k} \mathbf{1}_{r,r} J_{k,k}$$

a 2×2 matrix and thereafter we deduce

$$\begin{aligned} \sigma^2(\hat{\beta}_{jk}) &= \gamma_k^{-1} [\omega_{jk} + \omega_{jk}^2 (\theta_k z_{jk}^2 - A_{k,k} \gamma_k^{-1} X_{jk}^2)]; j = 1, 2 \\ cov(\hat{\beta}_{1k}; \hat{\beta}_{2k}) &= \gamma_k^{-1} \omega_{1k} \omega_{2k} [\theta_k z_{1k} z_{2k} - A_{k,k} \gamma_k^{-1} X_{1k} X_{2k}] \end{aligned} \quad (23)$$

ω_{jk} is the diagonal elements of $\Omega_{\alpha,k}^{-1}$ and

$$\theta_k = \frac{t_k \gamma_k^2}{n_k^2} (\beta_0^2 - (J_\alpha/B_\alpha)^{-1} t_k \gamma_k); \quad X_{jk} = 1 + \theta_k \Delta_{\alpha,k} z_{jk}.$$

5.2. example 2: $s = 2$ and $r = 2$

We have two experimental sites and two accident types with n_k, n_m the total number of accident in each site and $z_k = (z_{1k}, z_{2k})^T, z_m = (z_{1m}, z_{2m})^T$ the data from each associated control area. The parameter and the constraints are defined by $\alpha = (\beta_0, \beta_k, \beta_m)^T \in \mathbb{R}^5$ where $\beta_k = (\beta_{1k}, \beta_{2k})^T \in \mathbb{R}^2, \beta_m = (\beta_{1m}, \beta_{2m})^T \in \mathbb{R}^2$ with $\beta_0 > 0, \beta_{ij} > 0, \beta_{1k} + \beta_{2k} = 1$ and $\beta_{1m} + \beta_{2m} = 1$. So the dimension of the parameter space is five and we have two constraints. One notes that the additional experimental site to example 1 has extended the dimension of the parameter space by two and the constrained information matrix is a 7×7 matrix given by

$$\Gamma_\alpha = \begin{pmatrix} \tau_\alpha & \frac{\gamma_k}{\beta_0} v_{1k} & \frac{\gamma_k}{\beta_0} v_{2k} & \frac{\gamma_m}{\beta_0} v_{1m} & \frac{\gamma_m}{\beta_0} v_{2m} & 0 & 0 \\ \frac{\gamma_k}{\beta_0} v_{1k} & b_{11}^{(k)} & b_{12}^{(k)} & 0 & 0 & 1 & 0 \\ \frac{\gamma_k}{\beta_0} v_{2k} & b_{21}^{(k)} & b_{22}^{(k)} & 0 & 0 & 1 & 0 \\ \frac{\gamma_m}{\beta_0} v_{1m} & 0 & 0 & b_{11}^{(m)} & b_{12}^{(m)} & 0 & 1 \\ \frac{\gamma_m}{\beta_0} v_{1m} & 0 & 0 & b_{21}^{(m)} & b_{22}^{(m)} & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{pmatrix}$$

where $\tau_\alpha = \frac{\gamma_k^2 \langle z_k, \beta_k \rangle}{\beta_0 n_k} + \frac{\gamma_m^2 \langle z_m, \beta_m \rangle}{\beta_0 n_m}$ and the other terms are defined in the same way as in the example 1. It is the same for vectors $V_{\alpha,k}, V_{\alpha,m}$ and matrices $\Omega_{\alpha,k}^{-1}$, and $\Omega_{\alpha,m}^{-1}$. Setting

$$J_\alpha = \begin{pmatrix} \tau_\alpha & \frac{\gamma_k}{\beta_0} V_{\alpha,k}^T & \frac{\gamma_m}{\beta_0} V_{\alpha,m}^T \\ \frac{\gamma_k}{\beta_0} V_{\alpha,k} & B_{\alpha,k} & \circ \\ \frac{\gamma_m}{\beta_0} V_{\alpha,m} & \circ & B_{\alpha,m} \end{pmatrix}, \quad B_\alpha = \begin{pmatrix} B_{\alpha,k} & \circ \\ \circ & B_{\alpha,m} \end{pmatrix}, \quad H_2 = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

and in view of corollary 4.2, the estimated standard error of the first component is

$$\hat{SE}(\hat{\beta}_0) = \left[(J_{\hat{\alpha}}/B_{\hat{\alpha}}) + \|\xi_{\hat{\alpha}}\|_{\Lambda_{\hat{\alpha}}^{-1}}^2 \right]^{-1/2}$$

with

$$(J_{\hat{\alpha}}/B_{\hat{\alpha}}) = \sum_{i \in \{k,m\}} \frac{\hat{\gamma}_i^2}{n_i} \left(\frac{\langle z_i, \hat{\beta}_i \rangle}{\hat{\beta}_0} - \frac{\hat{\gamma}_i \hat{t}_i}{n_i} \|z_i\|_{\Omega_{\hat{\alpha},i}^{-1}}^2 \right); \quad \|\xi_{\hat{\alpha}}\|_{\Lambda_{\hat{\alpha}}^{-1}}^2 = \sum_{i \in \{k,m\}} \frac{\hat{\gamma}_i^2 \hat{t}_i^2 \Delta_{\hat{\alpha},i}^2}{n_i^2 \|\mathbf{1}_r\|_{B_{\hat{\alpha},i}^{-1}}^2}.$$

The covariances between $\hat{\beta}_0$ and the two other estimators $\hat{\beta}_k$ and $\hat{\beta}_m$ are obtained in the same way as in example 1. We have here (see appendix B) one between covariances matrix $Cov(\hat{\beta}_k, \hat{\beta}_m)$ and two within covariances matrices $Var(\hat{\beta}_k)$ and $Var(\hat{\beta}_m)$. So the standard errors of $\hat{\beta}_{jk}$ and $\hat{\beta}_{jm}$ are obtained with the square root of the diagonal elements of each within covariances matrix. We don't need any numerical

inversion programm to succeed in doing it, only matrices manipulations (product and addition) are called for. For example (see for furthers technical details in appendix B) the 2×2 within matrix $Var(\hat{\beta}_k)$ is obtained as folows

$$Var(\hat{\beta}_k) = \gamma_k^{-1} \Omega_{\alpha,k}^{-1} + \Omega_{\alpha,k}^{-1} [\phi_{k,m} z_k z_k^T + a_{k,m} (z_k \mathbf{1}_r^T + \mathbf{1}_r z_k^T) - A_{k,k} \gamma_k^{-2} X_k X_k^T] \Omega_{\alpha,k}^{-1}$$

where

$$\phi_{k,m} = \gamma_k^{-1} \theta_k + 2A_{k,m} v_{k,m} \gamma_k^{-1} \theta_k - A_{m,m} v_{k,m}^2 \Delta_{\alpha,k}^{-2}, \quad a_{k,m} = A_{k,m} v_{k,m} \gamma_k^{-1} \Delta_{\alpha,k}^{-1}$$

and

$$v_{k,m} = (J_\alpha / B_\alpha)^{-1} \frac{t_k \gamma_k \Delta_{\alpha,k}}{n_k} \frac{t_m \gamma_m \Delta_{\alpha,m}}{n_m}$$

6. Concluding remarks

We can use any inversion method (Guttman 1946, ouellette 1981 pp. 255 - 266) or any computer programm (Press et al 1992, N'Guessan 1993) to calculate numerically and directly the asymptotic covariances matrix W_α (see appendix) by inverting the matrix

$$\Gamma_\alpha = \begin{pmatrix} J_\alpha & H_\alpha^T \\ H_\alpha & 0 \end{pmatrix}$$

and thereafter standard errors are obtained by taking the square root of the diagonal elements of the covariance matrix equal to the leading $(1 + sr) \times (1 + sr)$ submatrix in Γ_α^{-1} i.e. W_α . However one recalls that the dimension of the parameter space and the number of constrained related to the model above are very closed to s and r . So the size of matrix Γ_α increases with them i.e. with the dimension of accident data tables which are to be combined. Of course this latter relation may seriously affect the numerical values of Γ_α^{-1} when s and r are large, and sometimes leads to instability in the coefficients of Γ_α^{-1} and therefore in those of W_α . The main strength of our approach is its applicability to any dimension of matrix Γ_α without being obliged to invert it. It saves a considerable time and precision in calculations and is accessible to everyone. Theorem 4.1 provides a powerful tool to compute asymptotic variances and covariances based on relatively straightforward calculations and a reduction of dimensionality. Once data z_k, n_k have been given and scalars γ_k, τ_α and vectors $V_{\alpha,k}$, and matrices $\Omega_{\alpha,k}$ have been found, theorem 4.1 gives a general formula for the asymptotic covariance matrix W_α . Moreover, if \hat{W}_α denotes the matrix W_α evaluated at the restraint maximum likelihood estimate $\hat{\alpha}$, then estimated variances and covariances of the parameter estimates can be easily computed, estimated standard errors are obtained by taking the square root of the diagonal elements of the same matrix, and asymptotic normal theory confidence regions can be constructed as well. For example, the standard error of $\hat{\beta}_0$ the element of interest of $\hat{\alpha}$ is obtained as follows

$$SE(\hat{\beta}_0) = \left[(J_{\hat{\alpha}} / B_{\hat{\alpha}}) + \|\xi_{\hat{\alpha}}\|_{\Lambda_{\hat{\alpha}}^{-1}}^2 \right]^{-1/2}$$

and without matrix inversion. We can also get asymptotic correlation coefficients by capitalizing on the elements apart from the diagonal. Furthers numerical developments will make capital out of those results and remarks at subsequent papers in order to bring forward asymptotic confidence intervals and correlation coefficients for all components of the restraint m.l.e. $\hat{\alpha}$.

Appendix A : Proof of the main result

Proof. of the theorem 4.1

Application of the mean value theorem at α^0 to (10) gives

$$\left(\frac{\partial L}{\partial \alpha}\right)_{\alpha^0} + \left(\frac{\partial^2 L}{\partial \alpha \partial \alpha}\right)_{\alpha}(\hat{\alpha} - \alpha^0) + H_{\alpha}^T \hat{\lambda} = 0 \quad (24)$$

$$H_{\alpha}(\hat{\alpha} - \alpha^0) = 0.$$

Now combining lemma 4 and the constraints (2), then likelihood equations can be written as

$$\left(\frac{\partial L}{\partial \alpha}\right)_{\alpha^0} + J_{\alpha}(M_{\alpha} - I_{1+sr})(\hat{\alpha} - \alpha^0) + H_{\alpha}^T \hat{\lambda} = 0 \quad (25)$$

$$H_{\alpha}(\hat{\alpha} - \alpha^0) = 0.$$

A matrix version of the latter is as follows

$$D_{\alpha} \begin{pmatrix} \hat{\alpha} - \alpha^0 \\ \eta_{\hat{\alpha}} \end{pmatrix} = \begin{pmatrix} J_{\alpha}^{-1} \left(\frac{\partial L}{\partial \alpha}\right)_{\alpha^0} \\ 0 \end{pmatrix} \quad (26)$$

where

$$D_{\alpha} = \begin{pmatrix} I_{1+sr} - M_{\alpha} & -\Phi_{\alpha} \\ H_{\alpha} & 0_s \end{pmatrix}; \quad \eta_{\hat{\alpha}} = H_2(J_{\alpha}/\tau_{\alpha})^{-1} H_2^T \hat{\lambda}.$$

The matrix J_{α}^{-1} and D_{α} are evaluated at any point of the line segment joining α^0 and $\hat{\alpha}$. Under (A1) and (A2) and using a version of a theorem of Crowder (1984, section 2), then the latter system of equations can be written as

$$\begin{pmatrix} \hat{\alpha} - \alpha^0 \\ \eta_{\hat{\alpha}} \end{pmatrix} \underset{\text{Distr.}}{\sim} D_{\alpha^0}^{-1} \begin{pmatrix} J_{\alpha}^{-1} \left(\frac{\partial L}{\partial \alpha}\right)_{\alpha^0} \\ 0 \end{pmatrix} \quad (27)$$

where

$$D_{\alpha^0} = \begin{pmatrix} I_{1+sr} & -\Phi_{\alpha^0} \\ H_{\alpha^0} & 0_s \end{pmatrix}; \quad D_{\alpha^0}^{-1} = \begin{pmatrix} I_{1+sr} - \Phi_{\alpha^0} H_{\alpha^0}^{-1} & \Phi_{\alpha^0} \\ -H_{\alpha^0} & I_s \end{pmatrix}.$$

By (A3) the asymptotic distribution of the right-hand side of (27) is $\mathcal{N}(0, \Sigma_{\alpha^0})$ where

$$\Sigma_{\alpha^0} = D_{\alpha^0}^{-1} \begin{pmatrix} J_{\alpha^0}^{-1} & 0 \\ 0 & 0 \end{pmatrix} (D_{\alpha^0}^{-1})^T.$$

After some matrix manipulation, we obtain the formal expression of Σ_{α} as follows

$$\Sigma_{\alpha} = \begin{pmatrix} W_{\alpha} & 0 \\ 0 & R_{\alpha} \end{pmatrix}; \quad W_{\alpha} = J_{\alpha}^{-1} - J_{\alpha}^{-1} H_{\alpha}^T R_{\alpha}^{-1} H_{\alpha} J_{\alpha}^{-1}.$$

Now using part *i*) of lemma 4 and reminding that $H_1 = O_s$ we get:

$$J_{\alpha}^{-1} H_{\alpha}^T = \begin{pmatrix} -(J_{\alpha}/B_{\alpha})^{-1} \xi_{\alpha}^T \\ (J_{\alpha}/\tau_{\alpha})^{-1} H_2^T \end{pmatrix}$$

a $(1 + sr) \times s$ matrix and

$$H_\alpha^T J_\alpha^{-1} = \left(-(J_\alpha/B_\alpha)^{-1} \xi_\alpha, H_2(J_\alpha/\tau_\alpha)^{-1} \right)$$

a $s \times (1 + sr)$ matrix. One deduces the formal components of W_α as follows

$$W_\alpha = \begin{bmatrix} (J_\alpha/B_\alpha)^{-1} & -(J_\alpha/B_\alpha)^{-1} U_\alpha^T B_\alpha^{-1} \\ -B_\alpha^{-1} U_\alpha (J_\alpha/B_\alpha)^{-1} & (J_\alpha/\tau_\alpha)^{-1} \end{bmatrix} - \begin{bmatrix} -(J_\alpha/B_\alpha)^{-1} \xi_\alpha^T \\ (J_\alpha/\tau_\alpha)^{-1} H_2^T \end{bmatrix} \times R_\alpha^{-1} \begin{bmatrix} -(J_\alpha/B_\alpha)^{-1} \xi_\alpha, H_2(J_\alpha/\tau_\alpha)^{-1} \end{bmatrix}. \quad (28)$$

After some manipulations, the formal expressions for the variances covariances of $\hat{\alpha}$ are

$$\begin{aligned} W_\alpha^{(11)} &= (J_\alpha/B_\alpha)^{-1} - (J_\alpha/B_\alpha)^{-1} \xi_\alpha^T R_\alpha^{-1} \xi_\alpha (J_\alpha/B_\alpha)^{-1} \\ W_\alpha^{(12)} &= -(J_\alpha/B_\alpha)^{-1} U_\alpha^T B_\alpha^{-1} + (J_\alpha/B_\alpha)^{-1} \xi_\alpha^T R_\alpha^{-1} H_2(J_\alpha/\tau_\alpha)^{-1} \\ W_\alpha^{(21)} &= -B_\alpha^{-1} U_\alpha (J_\alpha/B_\alpha)^{-1} + (J_\alpha/\tau_\alpha)^{-1} H_2^T R_\alpha^{-1} \xi_\alpha (J_\alpha/B_\alpha)^{-1} \\ W_\alpha^{(22)} &= (J_\alpha/\tau_\alpha)^{-1} - (J_\alpha/\tau_\alpha)^{-1} H_2^T R_\alpha^{-1} H_2(J_\alpha/\tau_\alpha)^{-1}. \end{aligned} \quad (29)$$

■

Proof. of Corollary 4.2

Using lemma 4 part ii), one obtains

$$\begin{aligned} R_\alpha^{-1} \xi_\alpha &= \left(\Lambda_\alpha^{-1} - ((J_\alpha/B_\alpha) + \|\xi_\alpha\|_{\Lambda_\alpha^{-1}}^2)^{-1} \Lambda_\alpha^{-1} \xi_\alpha \xi_\alpha^T \Lambda_\alpha^{-1} \right) \xi_\alpha \\ &= \left(1 - ((J_\alpha/B_\alpha) + \|\xi_\alpha\|_{\Lambda_\alpha^{-1}}^2)^{-1} \|\xi_\alpha\|_{\Lambda_\alpha^{-1}}^2 \right) \Lambda_\alpha^{-1} \xi_\alpha \\ &= \frac{(J_\alpha/B_\alpha)}{(J_\alpha/B_\alpha) + \|\xi_\alpha\|_{\Lambda_\alpha^{-1}}^2} \Lambda_\alpha^{-1} \xi_\alpha \end{aligned} \quad (30)$$

then

$$\begin{aligned} \|\xi_\alpha\|_{R_\alpha^{-1}}^2 &= \xi_\alpha^T R_\alpha^{-1} \xi_\alpha \\ &= \frac{(J_\alpha/B_\alpha)}{(J_\alpha/B_\alpha) + \|\xi_\alpha\|_{\Lambda_\alpha^{-1}}^2} \|\xi_\alpha\|_{\Lambda_\alpha^{-1}}^2. \end{aligned} \quad (31)$$

We deduce

$$\begin{aligned} \sigma^2(\hat{\beta}_0) &= (J_\alpha/B_\alpha)^{-1} - (J_\alpha/B_\alpha)^{-2} \|\xi_\alpha\|_{R_\alpha^{-1}}^2 \\ &= \frac{1}{(J_\alpha/B_\alpha) + \|\xi_\alpha\|_{\Lambda_\alpha^{-1}}^2}. \end{aligned} \quad (32)$$

The explicit expressions of $\|\xi_\alpha\|_{\Lambda_\alpha^{-1}}^2$ and $\|\mathbf{1}_r\|_{B_{\alpha,k}^{-1}}^2$ provide from a formal application of lemma 2 as follows

$$\begin{aligned}
\| \xi_\alpha \|_{\Lambda_\alpha^{-1}}^2 &= \xi_\alpha^T \Lambda_\alpha^{-1} \xi_\alpha \\
&= U_\alpha^T B_\alpha^{-1} H_2^T \Lambda_\alpha^{-1} H_2 B_\alpha^{-1} U_\alpha \\
&= (U_{\alpha,1}^T B_{\alpha,1}^{-1}, \dots, U_{\alpha,s}^T B_{\alpha,s}^{-1}) H_2^T \Lambda_\alpha^{-1} H_2 (B_{\alpha,1}^{-1} U_{\alpha,1}, \dots, B_{\alpha,s}^{-1} U_{\alpha,s}) \\
&= (U_{\alpha,1}^T B_{\alpha,1}^{-1}, \dots, U_{\alpha,s}^T B_{\alpha,s}^{-1}) \begin{bmatrix} \frac{\mathbf{1}_r \mathbf{1}_r^T}{\| \mathbf{1}_r \|_{B_{\alpha,1}^{-1}}^2} & \circ & \dots & \circ \\ \circ & \frac{\mathbf{1}_r \mathbf{1}_r^T}{\| \mathbf{1}_r \|_{B_{\alpha,2}^{-1}}^2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \circ \\ \circ & \dots & \circ & \frac{\mathbf{1}_r \mathbf{1}_r^T}{\| \mathbf{1}_r \|_{B_{\alpha,s}^{-1}}^2} \end{bmatrix} (B_{\alpha,1}^{-1} U_{\alpha,1}, \dots, B_{\alpha,s}^{-1} U_{\alpha,s})^T \\
&= \sum_{k=1}^s U_{\alpha,k}^T B_{\alpha,k}^{-1} \frac{\mathbf{1}_r \mathbf{1}_r^T}{\| \mathbf{1}_r \|_{B_{\alpha,k}^{-1}}^2} B_{\alpha,k}^{-1} U_{\alpha,k} \\
&= \sum_{k=1}^s \frac{t_k^2 \gamma_k^2}{n_k^2} z_k^T \Omega_{\alpha,k}^{-1} \frac{\mathbf{1}_r \mathbf{1}_r^T}{\| \mathbf{1}_r \|_{B_{\alpha,k}^{-1}}^2} \Omega_{\alpha,k}^{-1} z_k \\
\| \xi_\alpha \|_{\Lambda_\alpha^{-1}}^2 &= \sum_{k=1}^s \frac{t_k^2 \gamma_k^2}{n_k^2} \frac{\Delta_{\alpha,k}^2}{\| \mathbf{1}_r \|_{B_{\alpha,k}^{-1}}^2},
\end{aligned} \tag{33}$$

and

$$\begin{aligned}
\| \mathbf{1}_r \|_{B_{\alpha,k}^{-1}}^2 &= \mathbf{1}_r^T B_{\alpha,k}^{-1} \mathbf{1}_r \\
&= \mathbf{1}_r^T \gamma_k^{-1} (\Omega_{\alpha,k}^{-1} + t_k \Omega_{\alpha,k}^{-1} V_{\alpha,k} V_{\alpha,k}^T \Omega_{\alpha,k}^{-1}) \mathbf{1}_r \\
&= \gamma_k^{-1} [\text{trace}(\Omega_{\alpha,k}^{-1}) + t_k \frac{\beta_0^2 \gamma_k^2}{n_k^2} \mathbf{1}_r^T \Omega_{\alpha,k}^{-1} z_k z_k^T \Omega_{\alpha,k}^{-1} \mathbf{1}_r] \\
&= \gamma_k^{-1} [\text{trace}(\Omega_{\alpha,k}^{-1}) + t_k \frac{\beta_0^2 \gamma_k^2 \Delta_{\alpha,k}^2}{n_k^2}]
\end{aligned} \tag{34}$$

■ **Proof.** of Corollary 4.3

Notice that

$$\text{Cov}(\hat{\beta}; \hat{\beta}_0) = W_\alpha^{(21)}$$

so using lemma 3 and 4 we deduce

$$\begin{aligned}
\text{Cov}(\hat{\beta}; \hat{\beta}_0) &= -\frac{1}{(J_\alpha/B_\alpha) + \| \xi_\alpha \|_{\Lambda_\alpha^{-1}}^2} [B_\alpha^{-1} U_\alpha - B_\alpha^{-1} H_2^T \Lambda_\alpha^{-1} \xi_\alpha] \\
&= -\frac{1}{(J_\alpha/B_\alpha) + \| \xi_\alpha \|_{\Lambda_\alpha^{-1}}^2} [I_{sr} - B_\alpha^{-1} H_2^T \Lambda_\alpha^{-1} H_2] B_\alpha^{-1} U_\alpha
\end{aligned} \tag{35}$$

with

$$H_2^T \Lambda_\alpha^{-1} H_2 = \text{Bloc-diag} \left(\frac{\mathbf{1}_r \mathbf{1}_r^T}{\| \mathbf{1}_r \|_{B_{\alpha,1}^{-1}}^2}, \dots, \frac{\mathbf{1}_r \mathbf{1}_r^T}{\| \mathbf{1}_r \|_{B_{\alpha,s}^{-1}}^2} \right)$$

a $(sr) \times (sr)$ matrix where each bloc k ($k = 1, 2, \dots, s$) is a $r \times r$ matrix given by the expression

$$\frac{\mathbf{1}_r \mathbf{1}_r^T}{\|\mathbf{1}_r\|_{B_{\alpha,k}^{-1}}^2} = \frac{1}{\|\mathbf{1}_r\|_{B_{\alpha,k}^{-1}}^2} \begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix}. \quad (36)$$

Hence

$$\text{Cov}(\hat{\beta}_k; \hat{\beta}_0) = -\frac{1}{(J_\alpha/B_\alpha) + \|\xi_\alpha\|_{\Lambda_\alpha^{-1}}^2} (I_r - B_{\alpha,k}^{-1} \frac{\mathbf{1}_{r,r}}{\|\mathbf{1}_r\|_{B_{\alpha,k}^{-1}}^2}) B_{\alpha,k}^{-1} U_{\alpha,k} \quad (37)$$

for $k = 1, 2, \dots, s$. Now using part i) of lemma 2 we show that

$$B_{\alpha,k}^{-1} U_{\alpha,k} = \frac{t_k \gamma_k}{n_k} \Omega_{\alpha,k}^{-1} z_k \quad (38)$$

and

$$B_{\alpha,k}^{-1} \mathbf{1}_{r,r} \Omega_{\alpha,k}^{-1} = \gamma_k^{-1} \left[\Omega_{\alpha,k}^{-1} (\mathbf{1}_r \mathbf{1}_r^T + \rho_{\alpha,k} z_k \mathbf{1}_r^T) \Omega_{\alpha,k}^{-1} \right] \quad (39)$$

So after some algebraic manipulations, one obtains the expression of matrix $\mathcal{S}_{\alpha,k}$ and the result of the corollary. ■

Proof. of Corollary 4.4

This proof is an algebraic manipulations of matrices. So we just give the outline. From theorem 4.1 $\text{Var}(\hat{\beta})$ is the $(sr) \times (sr)$ matrix

$$\text{Var}(\hat{\beta}) = (J_\alpha/\tau_\alpha)^{-1} - (J_\alpha/\tau_\alpha)^{-1} H_2^T R_\alpha^{-1} H_2 (J_\alpha/\tau_\alpha)^{-1}.$$

Now using lemma 3 and matrix $B_\alpha^{-1} = \text{Bloc-diag}(B_{\alpha,1}^{-1}, \dots, B_{\alpha,s}^{-1})$, one shows that

$$(J_\alpha/\tau_\alpha)^{-1} = \begin{bmatrix} B_{\alpha,1}^{-1} & \circ & \dots & \circ \\ \circ & B_{\alpha,2}^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \circ \\ \circ & \dots & \circ & B_{\alpha,s}^{-1} \end{bmatrix} - (J_\alpha/B_\alpha)^{-1} \begin{bmatrix} B_{\alpha,1}^{-1} U_{\alpha,1} \\ B_{\alpha,2}^{-1} U_{\alpha,2} \\ \vdots \\ B_{\alpha,s}^{-1} U_{\alpha,s} \end{bmatrix} \left[U_{\alpha,1}^T B_{\alpha,1}^{-1}, U_{\alpha,2}^T B_{\alpha,2}^{-1}, \dots, U_{\alpha,s}^T B_{\alpha,s}^{-1} \right]$$

So $(J_\alpha/\tau_\alpha)^{-1} = ((J_{k,m}))$; $k, m = 1, 2, \dots, s$ a $(sr) \times (sr)$ matrix whose entry is also a $r \times r$ matrix such that

$$J_{k,m} = \delta_{k,m} B_{\alpha,k}^{-1} - (J_\alpha/B_\alpha)^{-1} B_{\alpha,k}^{-1} U_{\alpha,k} U_{\alpha,m}^T B_{\alpha,m}^{-1}. \quad (40)$$

In view of (11), and the relations

$$\begin{aligned} \mathbf{1}_{r,r} B_{\alpha,k}^{-1} U_{\alpha,k} U_{\alpha,m}^T B_{\alpha,m}^{-1} \mathbf{1}_{r,r} &= \frac{t_k t_m \Delta_{\alpha,k} \Delta_{\alpha,m}}{n_k n_m} \mathbf{1}_{r,r} \\ \Delta_{\alpha,k} &= z_k^T \Omega_{\alpha,k}^{-1} \mathbf{1}_r = \mathbf{1}_r^T \Omega_{\alpha,k}^{-1} z_k \end{aligned} \quad (41)$$

we deduce

$$H_2^T R_\alpha^{-1} H_2 = A_\alpha \otimes \mathbf{1}_{r,r} \quad (42)$$

where \otimes is the Kronecker product, $A_\alpha = ((A_{k,m}))$ a $s \times s$ matrix whose entry is given by (20). So postmultiplication and premultiplication of $H_2^T R_\alpha^{-1} H_2$ by $(J_\alpha/\tau_\alpha)^{-1}$, we get

$$\begin{aligned}
& (J_\alpha/\tau_\alpha)^{-1} H_2^T R_\alpha^{-1} H_2 (J_\alpha/\tau_\alpha)^{-1} = \\
& = \begin{bmatrix} \sum_{j=1}^s [(\sum_{i=1}^s J_{1,i} A_{i,j}) \mathbf{1}_{r,r}] J_{j,1} & \sum_{j=1}^s [(\sum_{i=1}^s J_{1,i} A_{i,j}) \mathbf{1}_{r,r}] J_{j,2} & \dots & \sum_{j=1}^s [(\sum_{i=1}^s J_{1,i} A_{i,j}) \mathbf{1}_{r,r}] J_{j,s} \\ \sum_{j=1}^s [(\sum_{i=1}^s J_{2,i} A_{i,j}) \mathbf{1}_{r,r}] J_{j,1} & \sum_{j=1}^s [(\sum_{i=1}^s J_{2,i} A_{i,j}) \mathbf{1}_{r,r}] J_{j,2} & \dots & \sum_{j=1}^s [(\sum_{i=1}^s J_{2,i} A_{i,j}) \mathbf{1}_{r,r}] J_{j,s} \\ \dots & \dots & \dots & \dots \\ \sum_{j=1}^s [(\sum_{i=1}^s J_{s,i} A_{i,j}) \mathbf{1}_{r,r}] J_{j,1} & \sum_{j=1}^s [(\sum_{i=1}^s J_{s,i} A_{i,j}) \mathbf{1}_{r,r}] J_{j,2} & \dots & \sum_{j=1}^s [(\sum_{i=1}^s J_{s,i} A_{i,j}) \mathbf{1}_{r,r}] J_{j,s} \end{bmatrix} \quad (43)
\end{aligned}$$

So, we obtain the second term of the right hand of (19). The result follows from matrices addition. ■

Appendix B: Some technical details for examples 1 and 2

Example 1

For $s = 1$ and $r > 1$, and in view of lemma 2 part i), the $r \times r$ Schur complement matrix $J_{k,k}$ is

$$J_{k,k} = \gamma_k^{-1} \left[\Omega_{\alpha,k}^{-1} + \theta_k \Omega_{\alpha,k}^{-1} z_k z_k^T \Omega_{\alpha,k}^{-1} \right]. \quad (44)$$

So using the latter expression of $J_{k,k}$ and after some matrices manipulations, one obtains

$$\begin{aligned}
\text{Var}(\hat{\beta}_k) &= J_{k,k} - A_{k,k} J_{k,k} \mathbf{1}_{r,r} J_{k,k} \\
&= \gamma_k^{-1} \left[\Omega_{\alpha,k}^{-1} + \Omega_{\alpha,k}^{-1} (\theta_k z_k z_k^T - A_{k,k} \gamma_k^{-1} X_k X_k^T) \Omega_{\alpha,k}^{-1} \right] \quad (45)
\end{aligned}$$

where $X_k = (X_{1k}, \dots, X_{rk})^T$ a $r \times 1$ vector of elements $X_{jk} = 1 + \theta_k \Delta_{\alpha,k} z_{jk}$. In this way, we deduce

$$\begin{aligned}
\sigma^2(\hat{\beta}_{jk}) &= \gamma_k^{-1} [\omega_{jk} + \omega_{jk}^2 (\theta_k z_{jk}^2 - A_{k,k} \gamma_k^{-1} X_{jk}^2)]; j = 1, 2, \dots, r \\
\text{cov}(\hat{\beta}_{jk}; \hat{\beta}_{mk}) &= \gamma_k^{-1} \omega_{jk} \omega_{mk} [\theta_k z_{jk} z_{mk} - A_{k,k} \gamma_k^{-1} X_{jk} X_{mk}]; j \neq m = 1, 2, \dots, r. \quad (46)
\end{aligned}$$

A straightforward development of the $r \times r$ matrix $S_{\alpha,k}$ of corollary 4.3 implies the covariance between $\hat{\beta}_0$ and any other estimator $\hat{\beta}_{mk}$ as follows

$$\begin{aligned}
\frac{n_k}{t_k \sigma^2(\hat{\beta}_0)} \text{cov}(\hat{\beta}_{mk}, \hat{\beta}_0) &= - \left(\frac{\gamma_k \beta_{mk}}{1 + \beta_0 z_{mk}} - \frac{(1 + \rho_{\alpha,k} z_{mk}) \beta_{mk}^2}{\| \mathbf{1}_r \|_{B_{\alpha,k}^{-1}}^2 (1 + \beta_0 z_{mk})^2} \right) z_{mk} + \\
&+ \frac{(1 + \rho_{\alpha,k} z_{mk}) \beta_{mk}}{\| \mathbf{1}_r \|_{B_{\alpha,k}^{-1}}^2 (1 + \beta_0 z_{mk})} \left(\Delta_{\alpha,k} - \frac{\beta_{mk} z_{mk}}{1 + \beta_0 z_{mk}} \right) \quad (47)
\end{aligned}$$

for $m = 1, 2, \dots, r$.

Now for $r = 2$,

$$S_{\alpha,k} = - \frac{t_k \sigma^2(\hat{\beta}_0)}{n_k} \begin{pmatrix} \frac{\gamma_k \beta_{1k}}{1 + \beta_0 z_{1k}} - \frac{(1 + \rho_{\alpha,k} z_{1k}) \beta_{1k}^2}{\| \mathbf{1}_r \|_{B_{\alpha,k}^{-1}}^2 (1 + \beta_0 z_{1k})^2} & - \frac{(1 + \rho_{\alpha,k} z_{1k}) \beta_{1k} \beta_{2k}}{\| \mathbf{1}_r \|_{B_{\alpha,k}^{-1}}^2 (1 + \beta_0 z_{1k})(1 + \beta_0 z_{2k})} \\ - \frac{(1 + \rho_{\alpha,k} z_{2k}) \beta_{2k} \beta_{1k}}{\| \mathbf{1}_r \|_{B_{\alpha,k}^{-1}}^2 (1 + \beta_0 z_{2k})(1 + \beta_0 z_{1k})} & \frac{\gamma_k \beta_{2k}}{1 + \beta_0 z_{2k}} - \frac{(1 + \rho_{\alpha,k} z_{2k}) \beta_{2k}^2}{\| \mathbf{1}_r \|_{B_{\alpha,k}^{-1}}^2 (1 + \beta_0 z_{2k})^2} \end{pmatrix}$$

we get more details for the covariances between the 2×1 vector $\hat{\beta}_k$ and β_0 of example 1 as follows

$$\begin{aligned} \frac{n_k}{t_k \sigma^2(\hat{\beta}_0)} \text{cov}(\hat{\beta}_{1k}, \hat{\beta}_0) = & - \left(\frac{\gamma_k \beta_{1k}}{1 + \beta_0 z_{1k}} - \frac{(1 + \rho_{\alpha,k} z_{1k}) \beta_{1k}^2}{\| \mathbf{1}_r \|^2_{B_{\hat{\alpha},k}^{-1}} (1 + \beta_0 z_{1k})^2} \right) z_{1k} + \\ & + \left(\frac{(1 + \rho_{\alpha,k} z_{1k}) \beta_{1k} \beta_{2k}}{\| \mathbf{1}_r \|^2_{B_{\hat{\alpha},k}^{-1}} (1 + \beta_0 z_{1k})(1 + \beta_0 z_{2k})} \right) z_{2k} \end{aligned} \quad (48)$$

and

$$\begin{aligned} \frac{n_k}{t_k \sigma^2(\hat{\beta}_0)} \text{cov}(\hat{\beta}_{2k}, \hat{\beta}_0) = & \left(\frac{(1 + \rho_{\alpha,k} z_{2k}) \beta_{1k} \beta_{2k}}{\| \mathbf{1}_r \|^2_{B_{\hat{\alpha},k}^{-1}} (1 + \beta_0 z_{1k})(1 + \beta_0 z_{2k})} \right) z_{1k} - \\ & - \left(\frac{\gamma_k \beta_{2k}}{1 + \beta_0 z_{2k}} - \frac{(1 + \rho_{\alpha,k} z_{2k}) \beta_{2k}^2}{\| \mathbf{1}_r \|^2_{B_{\hat{\alpha},k}^{-1}} (1 + \beta_0 z_{2k})^2} \right) z_{2k}. \end{aligned} \quad (49)$$

Example 2

For $s = 2$ (two experimental sites noted k and m) and $r > 1$, A_α (resp. $H_2^T R_\alpha^{-1} H_2$) is a 2×2 (resp. $2r \times 2r$) matrix given by

$$A_\alpha = \begin{bmatrix} A_{k,k} & A_{k,m} \\ A_{k,m} & A_{m,m} \end{bmatrix}, \quad (\text{resp. } H_2^T R_\alpha^{-1} H_2 = \begin{bmatrix} A_{k,k} \mathbf{1}_{r,r} & A_{k,m} \mathbf{1}_{r,r} \\ A_{k,m} \mathbf{1}_{r,r} & A_{m,m} \mathbf{1}_{r,r} \end{bmatrix}).$$

Therefore we have

$$(J_\alpha / \tau_\alpha)^{-1} = \begin{bmatrix} J_{k,k} & J_{k,m} \\ J_{m,k} & J_{m,m} \end{bmatrix} \quad (50)$$

a $2r \times 2r$ matrix. So,

$$(J_\alpha / \tau_\alpha)^{-1} H_2^T R_\alpha^{-1} H_2 (J_\alpha / \tau_\alpha)^{-1} = \begin{bmatrix} \sum_{(j,i) \in \{k,m\}} J_{k,i} A_{i,j} \mathbf{1}_{r,r} J_{j,k} & \sum_{(j,i) \in \{k,m\}} J_{k,i} A_{i,j} \mathbf{1}_{r,r} J_{j,m} \\ \sum_{(j,i) \in \{k,m\}} J_{m,i} A_{i,j} \mathbf{1}_{r,r} J_{j,k} & \sum_{(j,i) \in \{k,m\}} J_{m,i} A_{i,j} \mathbf{1}_{r,r} J_{j,m} \end{bmatrix}$$

and we deduce two $r \times r$ within covariances matrix and one $r \times r$ between covariance matrix as follows

$$\text{Var}(\hat{\beta}_k) = J_{k,k} - \sum_{(j,i) \in \{k,m\}} J_{k,i} A_{i,j} \mathbf{1}_{r,r} J_{j,k} \quad (51)$$

$$\text{Var}(\hat{\beta}_m) = J_{m,m} - \sum_{(j,i) \in \{k,m\}} J_{m,i} A_{i,j} \mathbf{1}_{r,r} J_{j,m} \quad (52)$$

and

$$\text{Cov}(\hat{\beta}_k; \hat{\beta}_m) = J_{k,m} - \sum_{(j,i) \in \{k,m\}} J_{k,i} A_{i,j} \mathbf{1}_{r,r} J_{j,m}, \quad k \neq m. \quad (53)$$

For example

$$\begin{aligned} \text{Var}(\hat{\beta}_k) = & J_{k,k} - \sum_{(j,i) \in \{k,m\}} J_{k,i} A_{i,j} \mathbf{1}_{r,r} J_{j,k} \\ = & J_{k,k} - A_{k,k} J_{k,k} \mathbf{1}_{r,r} J_{k,k} - \\ & - A_{k,m} J_{k,m} \mathbf{1}_{r,r} J_{k,k} - A_{k,m} J_{k,k} \mathbf{1}_{r,r} J_{m,k} - A_{m,m} J_{k,m} \mathbf{1}_{r,r} J_{m,k}. \end{aligned} \quad (54)$$

Setting

$$v_{k,m} = (J_\alpha/B_\alpha)^{-1} \frac{t_k \gamma_k \Delta_{\alpha,k}}{n_k} \frac{t_m \gamma_m \Delta_{\alpha,m}}{n_m}$$

then

$$\begin{aligned} J_{k,m} \mathbf{1}_{r,r} J_{k,k} &= -v_{k,m} \gamma_k^{-1} \Delta_{\alpha,k}^{-1} \Omega_{\alpha,k}^{-1} (z_k \mathbf{1}_r^T + \theta_k \Delta_{\alpha,k} z_k z_k^T) \Omega_{\alpha,k}^{-1} \\ J_{k,k} \mathbf{1}_{r,r} J_{m,k} &= -v_{k,m} \gamma_k^{-1} \Delta_{\alpha,k}^{-1} \Omega_{\alpha,k}^{-1} (\mathbf{1}_r z_k^T + \theta_k \Delta_{\alpha,k} z_k z_k^T) \Omega_{\alpha,k}^{-1} \\ J_{k,m} \mathbf{1}_{r,r} J_{m,k} &= v_{k,m}^2 \Delta_{\alpha,k}^{-2} \Omega_{\alpha,k}^{-1} z_k z_k^T \Omega_{\alpha,k}^{-1}. \end{aligned} \quad (55)$$

So we deduce

$$\begin{aligned} \text{Var}(\hat{\beta}_k) &= \gamma_k^{-1} \left[\Omega_{\alpha,k}^{-1} + \Omega_{\alpha,k}^{-1} (\theta_k z_k z_k^T - A_{k,k} \gamma_k^{-1} X_k X_k^T) \Omega_{\alpha,k}^{-1} \right] + \\ &+ A_{k,m} v_{k,m} \gamma_k^{-1} \Delta_{\alpha,k}^{-1} \Omega_{\alpha,k}^{-1} [z_k \mathbf{1}_r^T + \mathbf{1}_r z_k^T + 2\theta_k \Delta_{\alpha,k} z_k z_k^T] \Omega_{\alpha,k}^{-1} - \\ &- A_{m,m} v_{k,m}^2 \Delta_{\alpha,k}^{-2} \Omega_{\alpha,k}^{-1} z_k z_k^T \Omega_{\alpha,k}^{-1} \end{aligned} \quad (56)$$

where

$$z_k \mathbf{1}_r^T + \mathbf{1}_r z_k^T = \begin{bmatrix} 2z_{1k} & z_{1k} + z_{2k} & \dots & z_{1k} + z_{rk} \\ z_{2k} + z_{1k} & 2z_{2k} & \dots & z_{2k} + z_{rk} \\ \vdots & \ddots & \ddots & \vdots \\ z_{rk} + z_{1k} & \dots & \dots & 2z_{rk} \end{bmatrix}; \quad z_k z_k^T = \begin{bmatrix} z_{1k}^2 & z_{1k} z_{2k} & \dots & z_{1k} z_{rk} \\ z_{2k} z_{1k} & z_{2k}^2 & \dots & z_{2k} z_{rk} \\ \vdots & \ddots & \ddots & \vdots \\ z_{rk} z_{1k} & \dots & \dots & z_{rk}^2 \end{bmatrix}$$

and the $r \times r$ matrix $X_k X_k^T$ is given as follows

$$\begin{bmatrix} (1 + \theta_k \Delta_{\alpha,k} z_{1k})^2 & (1 + \theta_k \Delta_{\alpha,k} z_{1k})(1 + \theta_k \Delta_{\alpha,k} z_{2k}) & \dots & (1 + \theta_k \Delta_{\alpha,k} z_{1k})(1 + \theta_k \Delta_{\alpha,k} z_{rk}) \\ (1 + \theta_k \Delta_{\alpha,k} z_{1k})(1 + \theta_k \Delta_{\alpha,k} z_{2k}) & (1 + \theta_k \Delta_{\alpha,k} z_{2k})^2 & \dots & (1 + \theta_k \Delta_{\alpha,k} z_{2k})(1 + \theta_k \Delta_{\alpha,k} z_{rk}) \\ \vdots & \ddots & \ddots & \vdots \\ (1 + \theta_k \Delta_{\alpha,k} z_{1k})(1 + \theta_k \Delta_{\alpha,k} z_{rk}) & \dots & \dots & (1 + \theta_k \Delta_{\alpha,k} z_{rk})^2 \end{bmatrix}$$

So

$$\text{Var}(\hat{\beta}_k) = \gamma_k^{-1} \Omega_{\alpha,k}^{-1} + \Omega_{\alpha,k}^{-1} [\phi_{k,m} z_k z_k^T + a_{k,m} (z_k \mathbf{1}_r^T + \mathbf{1}_r z_k^T) - A_{k,k} \gamma_k^{-2} X_k X_k^T] \Omega_{\alpha,k}^{-1}$$

where

$$\phi_{k,m} = \gamma_k^{-1} \theta_k + 2A_{k,m} v_{k,m} \gamma_k^{-1} \theta_k - A_{m,m} v_{k,m}^2 \Delta_{\alpha,k}^{-2}, \quad a_{k,m} = A_{k,m} v_{k,m} \gamma_k^{-1} \Delta_{\alpha,k}^{-1}.$$

acknowledgement

I am especially grateful to Prof. Claude Langrand for his helpful comments and suggestions, which have led to an improvement in this article which follows from my Ph.D thesis in road safety measure analysis. I wish also thank Professor Francois Bellavance of HEC Montreal for helpful discussions. Part of this work was carried out while the author was an invited researcher of CRT (Centre for Research on Transportation) of Université de Montréal, whose partial support is gratefully acknowledged.

References

- [1] Abdel-Aty M.A., Radwan A.E (2000), *Modeling traffic accident occurrence and involvement*, *Accid. Anal. and Prev.* 32, 633-642.

- [2] Aitchison J., Silvey S.D. (1958) *Maximum Likelihood Estimation of parameters subject to restraints* Annals of mathematical statistics, 29,813-829.
- [3] Barry S., Ginpil S, O'Neill, T.J (1999), *The effectiveness of air bags*, Accid. Anal. and Prev. 31, 781-787.
- [4] Beenstock M., Gafni D., Goldin E (2001), *The effect of traffic policing on road safety in Israel*, Accid. Anal. and Prev. 33, 73-80.
- [5] Crowder M. (1975) *Maximum likelihood estimation for dependent observations* J. R. Statist. Soc., 38, 45-53
- [6] Crowder M. (1984) *On the constrained maximum likelihood estimation with non i.i.d. observations* Ann. Inst. Statist. Math. 36, A, 239-249.
- [7] Danielson S. (1988), *Estimation of the effects of countermeasures on different types of accidents in the presence of regression effects*, Accid. Anal. and Prev. 20, n°4 , 289-298.
- [8] Danielson S. (1986), *A comparison of two methods for estimating of the effect of a countermeasure in the presence of regression effects*, Accid. Anal. and Prev. 18, n°1 , 13-23.
- [9] Frees E.W (2001) *Omitted variables in longitudinal data models* The Canadian Journal of Statistics vol. 29, No. 4, pp 573-595.
- [10] Gaudry M., Lassarre S.(2000), *Structural Road Accidents Models: The international DRAG Family*, Pergamon Ed.
- [11] Guttman L (1946) *Enlargement methods for computing the inverse matrix* Ann. Math. Statist. 17:336-343.
- [12] Harville D.A. (1977) *Maximum Likelihood Approaches to Variance Component Estimation and to related Problems* J.A.S.A. vol. 72, n°358, 320-340
- [13] Hauer E. (1997), *Observational before-after Studies in Road Safety: Estimating the effect of highway and traffic engineering measures on road safety*, Pergamon Edition.
- [14] Le Nhu D., Sun L., Zidek J.V. (2001) *Spatial prediction and temporal backcasting for environmental fields having monotone data patterns* The Canadian Journal of Statistics vol. 29, No. 4, pages 529-554.
- [15] Lee S-Y (1979) *Constrained estimation in covariance structure analysis* Biometrika 66, 3, 539-545.
- [16] Neuenschwander B.E., Flury B.D. (1997) *A note on Silvey's (1959) theorem* Statistics & Probability Letter 36, 307-317.
- [17] N'Guessan A. (1993), *Contribution à l'analyse statistique d'une mesure de sécurité routière*, Ph. D Thesis, Université de Lille 1. France.
- [18] N'Guessan A. (2003a) *On a use of Schur complements for an inverse of a constrained accident data information matrix* PUB. IRMA, LILLE Vol. 60, N°VI.
- [19] N'Guessan A., Langrand C. (1993), *Sur la distribution asymptotique de certaines statistiques utilisées dans le domaine de la sécurité routière*, C.R. Acad. Sci, t.317, Série I, 401-404.
- [20] N'Guessan A., Essai A., Langrand C. (2001) *Estimation multidimensionnelle des contrôles et de l'effet moyen d'une mesure de sécurité routière* Rev. Statistique Appliquée. XLIX(2), 83-100.
- [21] Ouellette D.V. (1981) *Schur complements and statistics* Linear Algebra and its applications 36: 187-295.
- [22] Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. F. (1992) *Numerical Recipes in C: The Art of Scientific Computing* Second Ed. Cambridge University Press.
- [23] Rivest L-P., Lévesque T. (2001) *Improved log-linear model estimators of abundance in capture-recapture experiments* The Canadian Journal of Statistics vol. 29, No. 4, pp 555-572.
- [24] Sherman J., Morrison W. J. (1950) *Adjustement of an inverse matrix corresponding a change in one element of a given matrix* Ann. Math. Statist. 21:124-127.
- [25] Silvey S.D. (1959) *The Lagrangian multiplier test* Ann. Math. Statist. 30, 389-407.
- [26] Tanner J.C. (1958), *A problem in the combination of accident frequencies*, Biometrika, 45, 331-342.
- [27] Van Eeden C. (1996) *Estimation in Restricted Parameter Spaces - Some history and Some Recent developments* CWI Quarterly Vol. 9 (1 and 2) , 69 - 76
- [28] Wald A. (1949) *Note on the consistency of the maximum likelihood estimate* Ann. Math. Statist. 20, 595-601.
- [29] Woodburry M. A. (1950) *Inverting modified matrices* Memorandum Report 42, Statistical Research Group, Princeton, N.J.

Annexe A10 :

*Technical Report, PUB IRMA, Lille, Vol. 66,
No. III, 1 - 20 (2006)*

Application d'une méthode statistique pour évaluer l'impact de certaines modifications des conditions de la route : étude de cas

PUB. IRMA, LILLE 2006
Vol. 66, N° III

Application d'une méthode statistique pour
évaluer l'impact de certaines modifications des
conditions de la route : étude de cas*

Assi N'GUESSAN
Laboratoire Paul Painlevé UMR CNRS 8524, Polytech'Lille
Université de Lille I 59655 Villeneuve d'Ascq, France
assi.nguessan.univ-lille1.fr

Marcel TRUFFIER
Centre d'Etudes Techniques de l'Équipement Nord-Picardie,
2, rue de Bruxelles, BP 275
59019 Lille Cedex, France
Marcel.Truffier@gouv-equipement.fr

Claude LANGRAND,
Laboratoire Paul Painlevé UMR CNRS 8524, UFR de Mathématiques
Université de Lille I 59655 Villeneuve d'Ascq, France
Claude.Langrand@univ-lille1.fr

François BELLAVANCE
HEC et Centre Interuniversitaire de Recherche
sur les réseaux d'entreprise, la logistique et le transport
Montréal, P.O. Box 6128, Succ. centre-ville Montréal, Qc, Canada H3C 3J7
francois.bellavance@hec.ca

*Preprint

III – 2

Abstract

We propose the evaluation of a road safety layout using crash data that occurred on an experimental site and a matched control site. The crash data in both sites are divided in several mutually exclusive types and in two observation periods (before and after). We use a specific multinomial pattern to model the total number of crashes in the experimental site while taking into consideration the number of crashes in the control site. We estimate the model parameters and their standard errors. We applied this approach on new road safety layouts done on the Nord Pas de Calais national road network and also on the impact of roadside billboards on accident rates in their immediate vicinity.

Résumé

Nous proposons l'évaluation d'une mesure de sécurité routière à partir des données d'accidents survenus en un site expérimental et dans une zone de contrôle qui lui est associée. Les données d'accidents du site et de la zone sont réparties en plusieurs types mutuellement exclusifs et en deux périodes d'observation (avant et après). Nous utilisons un modèle multinomial particulier pour modéliser le total des accidents observés au niveau du site expérimental en tenant compte du nombre des accidents de la zone de contrôle. Ensuite, nous estimons les paramètres du modèle ainsi que leurs erreurs standards. Nous appliquons cette approche à des aménagements routiers effectués sur le réseau des routes nationales du Nord Pas de Calais ainsi qu'à l'évaluation de l'impact sur les nombres d'accident de l'installation de panneaux-réclame le long des autoroutes.

Keys Words - Road Safety Measure, Accident data, Multinomial model, Constrained Maximum Likelihood, Schur complement.

AMS Classification - 62F10, 62F12, 62F25, 62F30.

1 Introduction

Les modifications des conditions de la route (installation de giratoires, construction d'ilôts en dur, modification du marquage au sol, etc), effectuées sur des zones reconnues dangereuses ou pas de certaines routes ou autoroutes, font partie des mesures de sécurité routière prises pour essayer de résoudre les problèmes d'insécurité sur ce type de réseau routier. Une fois les modifications terminées et la circulation rétablie, se pose la question de l'évaluation de la mesure prise, le changement des conditions de la route a-t-il fait baisser ou non de façon significative la dangerosité de la zone concernée? Autrement dit, comment évaluer l'efficacité de telles modifications routières? Ces mêmes questions se posent aussi lorsque l'on installe des panneaux-réclame le long de certaines routes en milieu urbain.

A cette fin, on relève les nombres d'accidents mortels, graves et légers de la zone aménagée (site expérimental) et ceux d'une zone de contrôle. Les données ainsi obtenues sont relatives à deux périodes d'observation : une période précédant le début des travaux et une autre suivant la fin de ceux-ci. Il s'agit donc d'utiliser l'ensemble de ces données pour non seulement répondre à la question posée mais aussi analyser l'impact de l'aménagement sur les différents types de gravité d'accident. La méthode d'évaluation que nous préconisons consiste à utiliser les données d'accidents de la zone de contrôle sous forme de rapports d'accidents et à appliquer le principe méthodologique mis au point par N'Guessan et al (2001). En effet, ces auteurs ont proposé un modèle et une méthode pour évaluer l'effet d'une mesure de sécurité routière appliquée à plusieurs sites expérimentaux et plusieurs types d'accidents, chaque site étant associé à une zone de contrôle dont le but est de prendre en compte l'effet de certains facteurs indépendamment de la mesure (intensité du trafic, conditions météorologiques, ...). Dans [6], ils ont étudié la convergence numérique de solutions à leur problème à l'aide de données simulées. Ensuite, ils ont utilisé la technique du complément de Schur (N'Guessan et Langrand 2005) pour obtenir l'expression formelle des variances et des covariances relatives aux estimateurs du maximum de vraisemblance des paramètres de leur modèle.

Notre objectif est donc d'appliquer leurs résultats à des études de cas. Dans la section 2, nous exposons le modèle statistique utilisé pour combiner les données des deux entités d'observation. Ensuite, nous donnons une interprétation concrète des paramètres du modèle statistique en terme d'effet de l'aménagement et en terme de risques probables d'accidents. Dans la section 3, nous présentons la méthode d'estimation de l'effet moyen de la modification des conditions de la route et de son incidence sur les différents types de risque d'accidents. Cette section 3 s'achève sur un rappel de résultats récents relatifs à l'estimation d'écart-types et d'intervalles de confiance. La section 4 traite des études de cas. En effet, nous analysons deux opérations du plan régional d'aménagements de sécurité routière effectuées entre 1997 et 1999 dans le département du Pas de Calais. La première application concerne l'aménagement de 1997 à 1998 d'un giratoire sur la route nationale 43 au niveau de la commune de Vermelles. La deuxième étude traite de la modification du marquage au sol sur un tronçon de

III – 4

la route nationale 17 afin d'empêcher les dépassements anarchiques de certains usagers. Ce deuxième aménagement routier a eu lieu en 1999 et se situe au niveau des communes de Vimy et d'Avion. Les deux derniers exemples concernent l'évaluation de l'effet de l'installation de panneaux-réclame le long de certaines autoroutes de la ville de Montréal et de son incidence sur différents risques d'accidents. Nous discutons, dans la section 5, de l'incidence de la structure des données utilisées sur les résultats obtenus ainsi que des limites et des possibilités d'extension du modèle statistique. L'annexe reprend quelques uns des récents résultats théoriques de N'Guessan et Langrand (2005) concernant l'estimation théorique de variances et de covariances par complément de Schur ([10], [12]).

2 Modélisation de l'effet moyen d'un aménagement routier en présence de plusieurs types d'accidents

Pour modéliser et estimer l'effet moyen d'un aménagement routier sur différents types d'accidents, et prendre formellement en considération les informations (nombres d'accidents) dans les zones de contrôle, nous utilisons la méthode proposée par N'Guessan et al. (2001). Cette méthode est une extension, du test décrit par Tanner (1958) pour évaluer l'effet moyen d'une modification apportée au réseau routier, au cas où plusieurs types d'accidents sont considérées dans le même modèle, comme par exemple dans la présente étude, les accidents mortelles, avec blessés graves, avec blessés légers ou avec dégâts matériels seulement.

En utilisant les notations données dans [6], on pose R ($R > 1$) le nombre de types d'accidents, x_{1j} (resp. x_{2j}) le nombre d'accidents du type j ($j = 1, 2, \dots, R$) observés avant (resp. après) la mise en place de l'aménagement routier, $n = \sum_{j=1}^R (x_{1j} + x_{2j})$ le nombre total d'accidents enregistrés sur la zone aménagée durant les deux périodes d'observation. On suppose également disposer, pendant ces deux périodes, des données d'accidents relatives à une zone de contrôle et on note c_1, c_2, \dots, c_R les coefficients de contrôle où c_j est le quotient du nombre d'accidents du type j enregistrés au cours de la période suivant les travaux d'aménagement sur le nombre d'accidents du type j enregistrés au cours de la période précédant ceux-ci. Par la suite, on note pour la zone aménagée X_{1j} (resp. X_{2j}) la variable aléatoire donnant le nombre d'accidents du type j de la période avant (resp. après) l'aménagement, dont une valeur observée est x_{1j} (resp. x_{2j}). On montre (voir N'Guessan et al 2001) que la loi de probabilité conditionnelle (conditionnellement au nombre total d'accidents observés avant et après au niveau du site expérimental) de la paire (X_1, X_2) où $X_1 = (X_{11}, X_{12}, \dots, X_{1R})^T$ et $X_2 = (X_{21}, X_{22}, \dots, X_{2R})^T$ est une loi multinomiale de dimension $2R$ dont les probabilités de classe sont données par $\pi_{1j}(\Theta)$

III - 5

et $\pi_{2j}(\Theta)$ avec

$$\pi_{1j}(\Theta) = \frac{p_j}{1 + \beta_0 \sum_{m=1}^R p_m c_m}, \quad \pi_{2j}(\Theta) = \frac{\beta_0 c_j p_j}{1 + \beta_0 \sum_{m=1}^R p_m c_m}, \quad j = 1, 2, \dots, R, \quad (1)$$

sous les contraintes

$$p_j > 0, \quad \text{et} \quad \sum_{j=1}^R p_j = 1; \quad (2)$$

$\Theta = (\beta_0, P^T)^T$, de dimension $(1 + R)$, est le vecteur des paramètres inconnus du modèle où β_0 ($\beta_0 > 0$) décrit l'effet moyen (ou indice d'efficacité moyenne) de l'aménagement, $P = (p_1, p_2, \dots, p_R)^T$, de dimension R , est un vecteur de paramètres auxiliaires dont les composantes servent de poids aux coefficients individuels c_j et permettent de définir le coefficient de contrôle moyen : $\sum_{m=1}^R p_m c_m$ (évolution moyenne du nombre d'accidents de la zone de contrôle associée à la zone aménagée). Si l'on se réfère au principe de base de la construction du modèle multinomial utilisé c'est-à-dire à la loi du couple (X_1, X_2) (voir [6]), le poids aléatoire p_j est un paramètre auxiliaire dont une valeur estimée, sous certaines conditions, est donnée par le rapport du nombre d'accidents du type j avant sur le nombre total d'accidents avant au site expérimental. Dans ces conditions, p_j représenterait la probabilité qu'un accident survenu dans la zone aménagée avant le début des travaux soit de type j si l'on n'utilise que les données d'accidents avant. Néanmoins, les contraintes sur les composantes du vecteur P montrent clairement qu'elles peuvent, certe, être associées à une loi de probabilité mais que leur estimation n'est pas unique. On propose, par la suite, d'estimer les composantes du vecteur P en même temps que l'estimation du paramètre β_0 en combinant les données d'accidents de la zone aménagée aux valeurs des coefficients de contrôle (voir section 3.1). La valeur estimée de p_j ainsi obtenue représente un ordre de grandeur probable du risque d'accidents de type j au niveau de la zone aménagée avant le début des travaux. Cette valeur est ensuite utilisée pour construire des fourchettes de valeurs (voir section 3.2) probables pour la vraie valeur du risque d'accidents de type j avant l'aménagement.

En pratique, on aménage un site parce que les nombres ou les risques d'accidents sont anormalement élevés. L'effet escompté de l'aménagement routier est donc de faire baisser les risques d'accidents du site expérimental. Néanmoins on constate, sur le terrain, que d'autres facteurs peuvent, dans le temps, avoir un impact sur le nombre des accidents (exemple : conditions météorologiques, densité du trafic, déplacement des accidents de la zone aménagée vers une zone environnante ...). De ce fait, il est important, dans la définition et le calcul des risques d'accidents, de distinguer la période avant de la période après et d'intégrer les données d'accidents de la zone de contrôle associée si l'on veut apprécier convenablement l'impact de cet aménagement sur les différents risques d'accidents du site expérimental. On propose d'utiliser la probabilité conditionnelle $\pi_{1j}(\Theta)$ (resp. $\pi_{2j}(\Theta)$) comme indicateur du risque d'accident du type j au cours de la période précédant (resp. suivant) l'aménagement du site expérimental. En effet, si l'effet escompté de l'aménagement est β_0 et si l'on pondère

III - 6

chaque coefficient individuel c_j par p_j alors $\pi_{1j}(\Theta)$ (resp. $\pi_{2j}(\Theta)$) représente, par construction du modèle statistique, la probabilité qu'un accident survenu au niveau de la zone aménagée au cours des deux périodes d'analyse soit du type j avant (resp. après) l'aménagement. Par conséquent, ces probabilités peuvent s'interpréter comme des indicateurs du risque d'accidents du type j avant (resp. après) au niveau de la zone aménagée conditionnellement aux données de la zone de contrôle associée. La comparaison de leur valeur peut ainsi permettre d'apprécier l'impact probable de l'aménagement de sécurité sur les différents types de risque d'accidents au niveau du site expérimental.

3 Estimation des paramètres du modèle et de leur erreur standard

3.1 Estimation ponctuelle de l'effet moyen et des différents types de risques d'accidents

L'évaluation de l'effet moyen de l'aménagement et de son impact sur les différents types de risque d'accidents passe par l'estimation des paramètres $\beta_0, p_1, p_2, \dots, p_R$. Cette estimation est obtenue à l'aide de la méthode du maximum de vraisemblance. Il s'agit de maximiser la fonction log-vraisemblance donnée par

$$\mathcal{L}(X_1, X_2; \Theta) = \text{constante} + \sum_{j=1}^R \left\{ X_{\bullet j} \log_e(p_j) + X_{2j} \log_e(\beta_0) - X_{\bullet j} \log_e \left(1 + \beta_0 \sum_{m=1}^R c_m p_m \right) \right\} \quad (3)$$

avec $X_{\bullet j} = X_{1j} + X_{2j}$. Du fait de (2), l'estimation des paramètres se fait dans le cadre de la méthode du maximum de vraisemblance sous contraintes (voir par exemple Aitchison et Silvey 1958, Crowder 1984, Haber and Brown 1986, Matthews and Crowther 1995). En prenant les dérivées partielles premières de $\mathcal{L}(X_1, X_2; \Theta)$ par rapport à $\beta_0, p_1, p_2, \dots, p_R$ et en égalant à zéro, on montre que l'estimateur du maximum de vraisemblance $\hat{\Theta}$ de Θ , s'il existe, est solution du système d'équations non linéaires :

$$\begin{cases} \sum_{j=1}^R \{ X_{2j} - X_{\bullet j} \frac{\hat{\beta}_0 \sum_{m=1}^R c_m \hat{p}_m}{1 + \hat{\beta}_0 \sum_{m=1}^R c_m \hat{p}_m} \} = 0 \\ X_{\bullet j} - n \frac{\hat{p}_j (1 + c_j \hat{\beta}_0)}{1 + \hat{\beta}_0 \sum_{m=1}^R c_m \hat{p}_m} = 0, \quad (j = 1, 2, \dots, R); \\ 0 < \hat{p}_j \text{ avec } \sum_{j=1}^R \hat{p}_j = 1. \end{cases} \quad (4)$$

La mise en oeuvre de la méthode du maximum de vraisemblance sous contraintes passe par la résolution numérique du système (4). Cette résolution fait appel à une procédure itérative du type Newton-Raphson. Le point de convergence de cette procédure est une estimation ponctuelle de $\hat{\beta}_0, \hat{p}_1, \hat{p}_2, \dots, \hat{p}_R$. Pour

III – 7

apprécier l'impact global de l'aménagement routier sur les nombres d'accidents, on compare le coefficient β_0 à 1. Une valeur de β_0 significativement supérieur à 1 indique que l'aménagement routier a eu un effet non bénéfique sur le nombre des accidents de la zone aménagée par rapport aux données de la zone de contrôle. Au contraire, une valeur de β_0 significativement inférieur à 1 témoigne d'une baisse moyenne importante en proportion du nombre des accidents de la zone aménagée par rapport à la zone de contrôle. Enfin, une valeur de β_0 non significativement différente de 1 indique que l'aménagement routier n'a pas eu d'effet.

3.2 Approximation des erreurs standards par complément de Schur

Le processus d'évaluation de l'aménagement routier passe par la construction d'intervalles de confiance et par la mise en place de tests d'hypothèse comme, par exemple, tester l'hypothèse d'efficacité moyenne nulle. Pour ce faire, on a besoin d'inverser la matrice d'information associée aux paramètres et aux contraintes et d'en extraire les éléments diagonaux. Dans le cas particulier d'un seul site aménagé et de plusieurs types de gravité d'accidents, cette matrice prend la forme particulière suivante :

$$\Gamma_{\Theta} = \begin{bmatrix} \frac{\gamma^2 \sum_m^R c_m p_m}{n \beta_0} & \frac{\gamma^2 c_1}{n} & \dots & \frac{\gamma^2 c_R}{n} & 0 \\ \frac{\gamma^2 c_1}{n} & \gamma \left(\frac{1 + \beta_0 c_1}{p_1} - \frac{\gamma \beta_0^2 c_1^2}{n} \right) & \dots & -\frac{\gamma^2 \beta_0^2 c_1 c_R}{n} & 1 \\ \frac{\gamma^2 c_2}{n} & -\frac{\gamma^2 \beta_0^2 c_2 c_1}{n} & \ddots & -\frac{\gamma^2 \beta_0^2 c_2 c_R}{n} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\gamma^2 c_R}{n} & -\frac{\gamma^2 \beta_0^2 c_R c_1}{n} & \dots & \gamma \left(\frac{1 + \beta_0 c_R}{p_R} - \frac{\gamma \beta_0^2 c_R^2}{n} \right) & 1 \\ 0 & 1 & \dots & 1 & 0 \end{bmatrix} \tag{5}$$

où $\gamma = n / (1 + \beta_0 \sum_m^R c_m p_m)$. Cette matrice d'information sous contraintes linéaires est de dimension $(2 + R) \times (2 + R)$ et est obtenue à l'aide de l'espérance des dérivées partielles secondes de $\mathcal{L}(X_1, X_2; \Theta)$ par rapport aux composantes du vecteur paramètre. Une fois les données d'accident connues et les paramètres estimés, l'inversion de Γ_{Θ} peut alors être effectuée à l'aide des résultats de N'Guessan et Langrand (2005a,b). Ces résultats, basés sur la technique du complément de Schur, ont l'avantage de remplacer l'inversion directe de Γ_{Θ} par des produits et des sommes de matrices de dimensions plus petites que celles de Γ_{Θ} . De plus, ces résultats s'appliquent à un nombre quelconque de sites expérimentaux et de types d'accidents et permettent d'obtenir l'expression analytique de la variance de chaque composante. On peut alors construire des intervalles de confiance (N'Guessan et Bellavance 2005) et bâtir des tests d'hypothèse relatifs à l'effet moyen de l'aménagement et aux différents types de risque de gravité d'accident. En utilisant la technique du complément de Schur (voir [5],[7], [8] et [9]), on obtient les intervalles de confiances à $(1 - \alpha)100\%$ suivants :

III – 8

$$\begin{aligned}
 IC_{1-\alpha}(\beta_0) &= \left[\max(0; \hat{\beta}_0 - \epsilon_{\alpha/2}\sigma(\hat{\beta}_0)); \hat{\beta}_0 + \epsilon_{\alpha/2}\sigma(\hat{\beta}_0) \right]; \\
 IC_{1-\alpha}(p_j) &= \left[\max(0; \hat{p}_j - \epsilon_{\alpha/2}\sigma(\hat{p}_j)); \min(1; \hat{p}_j + \epsilon_{\alpha/2}\sigma(\hat{p}_j)) \right], \quad (6) \\
 (j &= 1, 2, \dots, R),
 \end{aligned}$$

où $\epsilon_{\alpha/2}$ est obtenu de façon usuelle à l'aide de la table de la loi normale centrée et réduite, $\sigma(\bullet)$ est l'écart-type de chaque composante du paramètre Θ . L'obtention de ces écart-types est partiellement rappelée en annexe. Les intervalles de confiance complètent les valeurs ponctuelles des paramètres et permettent d'obtenir des fourchettes de valeurs probables pour l'ensemble des paramètres du modèle et de pouvoir, ainsi, juger de l'impact (négatif ou positif) de l'aménagement routier sur le nombre des accidents et sur les risques d'accidents.

4 Applications

4.1 Évaluation d'aménagements de sécurité sur certaines routes nationales

Les deux exemples d'application de cette sous-section concernent certaines opérations d'aménagements routiers effectuées dans le département du Pas de Calais (France). La première application est relative à l'aménagement, de 1997 à 1998, d'un giratoire sur la route nationale 43 au niveau de la commune de Vermelles. La deuxième étude traite de la modification du marquage au sol sur un tronçon de la route nationale 17 afin d'empêcher les dépassements anarchiques de certains usagers. Ce deuxième aménagement routier a eu lieu en 1999 et se situe au niveau des communes de Vimy et d'Avion.

4.1.1 Installation d'un giratoire sur route nationale

Un giratoire (rond-point) a été aménagé au niveau de la commune de Vermelles (département du Pas de Calais) au croisement de la route départementale 75 (RD75) et de la route nationale 43 (RN43). Les travaux ont duré deux ans (de 1997 à 1998). Afin d'évaluer cet aménagement de sécurité, on a relevé le nombre des accidents mortels (M), des accidents graves (G), et des accidents légers (L) au niveau de la zone aménagée (site expérimental : Vermelles) au cours des quatre années précédant le début des travaux et des quatre années suivant la fin de ceux-ci. Les données d'accident cumulées sur chaque période de 4 ans sont consignées dans le tableau 4.1 suivant :

Tableau 4.1 : Données d'accidents du site expérimental

	période avant (4 ans)			période après (4 ans)			total
	M	G	L	M	G	L	
zone aménagée	3	5	6	0	0	3	17

III – 9

Pour des périodes d'observation de longueur 4 ans, les accidents mortels sont passés de 3 avant à 0 après. Il y a eu donc 3 accidents avec au moins un tué au cours des 4 années précédant les travaux et 0 accident avec au moins un tué au cours des 4 années suivant les travaux. De même, les accidents graves (accidents avec au moins un blessé grave et aucun tué) sont passés de 5 avant à 0 après. Quant aux accidents légers (accidents sans tué ni blessé grave) ils ont baissé de moitié. Pour analyser et apprécier convenablement l'effet moyen de l'aménagement ainsi que son impact probable sur les risques de gravité, on combine, pour des périodes de même longueur, les données d'accident du site expérimental à celles relevées sur une portion de la RN43 servant de zone de contrôle. Les données de cette zone de contrôle sont fournies sous forme de rapports de nombres d'accidents (tableau 4.2) :

Tableau 4.2 : Valeurs des coefficients de contrôle

	c_1	c_2	c_3
zone de contrôle	0.6666	0.3930	0.5000

Par exemple la valeur 0.6666 est le rapport du nombre total des accidents mortels enregistrés sur la portion de la RN43 servant de zone de contrôle au cours des quatre années suivant la fin des travaux sur le nombre des accidents mortels durant les quatre années précédant le début des travaux. On note, de façon générale, une baisse du nombre de chaque type d'accidents après les travaux au niveau du site aménagé et une baisse en proportion du nombre de ceux-ci sur la portion de la route nationale servant de contrôle. Par exemple pour les accidents graves, on est passé au niveau du site expérimental de 5 accidents graves au cours des quatre années précédant l'aménagement à 0 au cours des quatre années le suivant, pendant que sur toute la portion de route nationale servant de contrôle, on a observé un rapport de contrôle de 0.3930 soit une baisse en proportion d'environ 60, 70% des accidents graves.

Les estimations de l'effet moyen de l'aménagement et de la loi de probabilité servant de poids aux coefficients de contrôle sont disponibles dans le tableau 4.3. Les valeurs entre parenthèses représentent l'erreur standard (E.S) c'est-à-dire l'écart-type estimé de chaque paramètre du modèle. La valeur estimée de l'effet moyen de cet aménagement est de 0,4328 soit une baisse moyenne en proportion de $(1 - 0,4328) \times 100 = 56,72\%$ des accidents de la zone aménagée par rapport à la tendance moyenne de l'évolution des accidents survenus dans la zone de contrôle au cours de la période de quatre ans suivant l'installation du giratoire. L'intervalle de confiance au niveau 95% associé à l'effet moyen est $[0; 0,9738]$ avec une borne supérieure plus petite que 1. Ainsi, avec 95 chances sur 100, un tel aménagement a apporté, dans la période de quatre ans après sa mise en place, une baisse moyenne en proportion statistiquement significative du nombre des accidents de la zone aménagée comparativement à la tendance moyenne du nombre des accidents de la zone de contrôle.

Les valeurs estimées des paramètres p_1 , p_2 et p_3 à l'aide de la méthode du maximum de vraisemblance sont respectivement 0,1663, 0,3052 et 0,5285. Ces

III – 10

valeurs donnent un ordre de grandeur du niveau de risque de chaque type d'accidents de la zone aménagée au cours des quatre années précédant l'aménagement si l'on tient compte des données de la zone de contrôle. Par exemple, après ajustement à l'aide des coefficients de contrôle, on estime en moyenne que 16,63% des accidents survenus dans la zone aménagée au cours des quatre années précédant l'installation du giratoire ont été mortels. L'intervalle de confiance au niveau 95% ($[0; 0,3419]$) associé au paramètre p_1 montre qu'il est fort probable que le risque d'accidents mortels de la zone aménagée ne soit pas statistiquement significativement différent de 0 au cours des quatre années précédant l'installation du giratoire.

Tableau 4.3 : Estimation des paramètres du modèle

paramètres	estimation (E.S)	B_{inf} à 95%	B_{sup} à 95%
$\hat{\theta}$	0.4328 (0.2760)	0	0.9738
\hat{p}_1	0.1663 (0.086)	0	0.3419
\hat{p}_2	0.3052 (0.1175)	0.0749	0.5355
\hat{p}_3	0.5285 (0.1329)	0.2680	0.7889

En revanche, il est fort probable que le risque d'accidents graves de la zone aménagée soit statistiquement significativement différent de 0 au cours des quatre années précédant l'aménagement du giratoire car l'intervalle de confiance au niveau 95% associé à p_2 ne contient pas 0.

Pour apprécier, au niveau de la zone aménagée, l'incidence de l'installation du giratoire sur les différents risques d'accidents, on utilise la valeur des probabilités conditionnelles $\pi_{2j}(\hat{\Theta})$ et $\pi_{2j}(\check{\Theta})$. Le calcul de ces risques d'accidents par période de temps utilise les valeurs estimées des paramètres et celles des coefficients de contrôle. Les valeurs obtenues sont disponibles dans le tableau 4.4 :

Tableau 4.4 : Estimation des risques d'accidents par période de temps

différents types de risque	avant : $\pi_{1j}(\hat{\Theta})$	après : $\pi_{2j}(\check{\Theta})$
accidents mortels	0.1370	0.0395
accidents graves	0.2513	0.0426
accidents légers	0.4352	0.0942

Par exemple, si l'on estime l'effet moyen de l'aménagement à 0,4328 et si les coefficients de contrôle sont ceux donnés par le tableau 4.2 alors au cours de la période des huit années d'analyse, le risque d'avoir un accident mortel avant l'aménagement de la zone de Vermelles est de l'ordre de $0,1370 = 0,1663/(1 + 0,4328 \times 0,4950)$ où $0,4950 = 0,1663 \times 0,6666 + 0,3052 \times 0,3930 + 0,5285 \times 0,5000$. De même, conditionnellement aux données de contrôle et aux valeurs estimées des paramètres du modèle, on a 3,95% de chances, soit $0,0395 = 0,1663 \times 0,4328 \times 0,6666/(1 + 0,4328 \times 0,4950)$, qu'un accident qui s'est produit pendant les huit années d'analyse soit un accident mortel survenu au cours de la

III - 11

période des quatre années suivant l'aménagement. Dans les mêmes conditions, on a 25, 13% (resp. 4, 26%) de chances qu'un accident qui s'est produit pendant les huit années d'analyse soit un accident grave survenu au cours de la période des quatre années précédant (resp. suivant) l'aménagement du giratoire. Les valeurs du risque d'accidents légers avant et après sont obtenus de façon similaire en utilisant 0,5285 et 0,5000. Globalement, selon les données disponibles, la longueur des périodes utilisées et les valeurs estimées des paramètres, on pense vraisemblablement que cet aménagement a eu localement (au niveau de la zone aménagée) un impact significatif sur les différents types de risque d'accidents de l'étude.

4.1.2 Analyse de la modification d'une signalisation au sol

Un autre aménagement routier a été effectué en 1999 sur la route nationale 17 (RN17) au niveau des communes de Vimy et d'Avion du département Pas de Calais. En effet, la route, entre ces deux communes, était une trois voies classique, à deux sens. L'aménagement a consisté à ce que l'on appelle un "marquage à l'Italienne". On modifie le marquage au sol pour qu'il soit impossible de doubler en même temps dans chaque sens. Sur une courte distance, on donne deux voies pour un sens (on peut alors doubler dans ce sens) et une seule voie sur l'autre sens séparées par une ligne continue. Plus loin, on inverse pour que l'autre sens puisse doubler et ainsi de suite. Les données relevées au niveau de la zone aménagée (Vimy-Avion) sont les suivantes (tableau 4.5) :

Tableau 4.5 : **Données d'accidents du site expérimental**

	période avant (4 ans)			période après (4 ans)			total
	<i>M</i>	<i>G</i>	<i>L</i>	<i>M</i>	<i>G</i>	<i>L</i>	
zone aménagée	4	4	16	1	1	7	33

Pour des périodes d'observation de longueur 4 ans, les accidents mortels sont passés de 4 avant à 1 après. Il y a eu donc 4 accidents avec au moins un tué au cours de la période précédant les travaux et 1 accident avec au moins un tué au cours de la période suivant ceux-ci. Pendant ce temps les accidents légers (accidents sans tué ni blessé grave) répertoriés ont baissé plus que de moitié. Pour les mêmes longueurs de périodes, on observe, sur la portion de la RN17 servant de zone de contrôle, les rapports de nombres d'accidents suivants (tableau 4.6) :

Tableau 4.6 : **Valeurs des coefficients de contrôle**

	<i>c</i> ₁	<i>c</i> ₂	<i>c</i> ₃
zone de contrôle	0.5190	0.4220	0.5600

Globalement on note une baisse en proportion du nombre des accidents de la zone de contrôle, entre la période de 4 ans suivant les travaux et la période de 4 ans précédant ceux-ci. Les résultats numériques relatifs à ces données d'accidents sont les suivants (tableau 4.7) :

III – 12

Tableau 4.7 : Estimation des paramètres du modèle

paramètres	estimation (E.S)	B_{inf} à 95%	B_{sup} à 95%
$\hat{\theta}$	0.7054 (0.2759)	0.1646	1.2462
\hat{p}_1	0.1525 (0.0643)	0.0264	0.2785
\hat{p}_2	0.1605 (0.0672)	0.0287	0.2922
\hat{p}_3	0.6870 (0.1023)	0.4864	0.8875

L'indice d'efficacité moyenne estimé à l'aide des données des tableaux 4.5 et 4.6 est de 0,7054, ce qui correspond à une baisse moyenne en proportion de 29,46% = $(1 - 0,7054) \times 100\%$ du nombre d'accidents de la zone aménagée (Vimy-Avion) par rapport à la tendance moyenne du nombre des accidents de la zone de contrôle. Les valeurs estimées des trois paramètres p_1 , p_2 et p_3 sont respectivement 0,1525, 0,1605 et 0,6870. Ces valeurs permettent d'apprécier le niveau probable des trois types de risque d'accidents du site expérimental au cours des quatre années précédant la mise en place des travaux de marquage au sol. Les intervalles de confiance au niveau 95% associés à ces trois paramètres montrent qu'au cours des quatre années précédant les travaux de marquage au sol, on estime, avec une probabilité de 0,95, que le risque d'observer un accident mortel est compris entre 2,64% et 27,85%, celui d'observer un accident grave entre 2,87% et 29,22%, celui d'observer un accident léger entre 48,64% et 88,75%.

On teste la significativité de l'effet moyen de ce type d'aménagement en utilisant l'intervalle de confiance au niveau 95% associé au paramètre θ . Cet intervalle de confiance montre qu'on a une fourchette de valeurs dont la borne inférieure (resp. supérieure) est strictement plus grande que 0 (resp. 1). En effet, comme $1 \in [0.1646; 1.2462]$, on ne peut rejeter l'hypothèse $H_0 : \theta = 1$ contre $H_1 : \theta \neq 1$ avec une erreur de première espèce de 5%. La valeur estimée de θ pouvant excéder 1, il est fort probable que ce type d'aménagement routier n'apporte aucun impact significatif sur le nombre moyen des accidents au niveau de la zone aménagée par rapport à l'évolution du nombre moyen des accidents de la zone de contrôle. Même si dans le cas présent, on note une baisse en proportion de 29,46% du nombre moyen des accidents de la zone aménagée (Vimy-Avion), le test ci-dessus montre que la valeur de l'indice d'efficacité moyenne n'est pas statistiquement significativement différente de 1 pour conclure de façon systématique à une efficacité avérée de ce type de marquage au sol.

Néanmoins, on peut apprécier l'impact de ce type de marquage sur les différents types de risque d'accidents par période de temps au niveau du site expérimental si l'on admet que l'effet moyen de cet aménagement sur les nombres d'accidents est de l'ordre de 0,7054 et que les rapports de nombres d'accidents sont ceux donnés dans le tableau 4.6. On utilise les valeurs estimées des probabilités conditionnelles $\pi_{1j}(\hat{\Theta})$ et $\pi_{2j}(\hat{\Theta})$ du modèle multinomial. Ces valeurs estimées sont présentées dans le tableau 4.8 :

Tableau 4.8 : Estimation des risques d'accidents par période de temps

III - 13

différents types de risque	avant : $\pi_{1j}(\hat{\Theta})$	après : $\pi_{2j}(\hat{\Theta})$
accidents mortels	0.1109	0.0400
accidents graves	0.1167	0.0347
accidents légers	0.4996	0.1973

L'interprétation de ces probabilités est identique à celle de l'installation du giratoire. Si on estime à 0,7054 l'effet moyen de cet type de marquage sur les nombres d'accidents de la zone aménagée et après ajustement avec les coefficients de contrôle alors on a 11,09% (resp. 4%) de chances qu'un accident qui s'est produit pendant les huit années d'analyse soit un accident mortel survenu au cours de la période des quatre années précédant (resp. suivant) l'aménagement. Dans les mêmes conditions, on a 11,67% (resp. 3,47%) de chances qu'un accident qui s'est produit pendant les huit années d'analyse soit un accident grave survenu au cours de la période des quatre années précédant (resp. suivant) l'aménagement du giratoire. Les valeurs du risque d'accidents légers avant et après sont obtenus de façon similaire en utilisant 0,6870 et 0,5600.

Ces valeurs estimées des risques d'accidents sur les huit années d'analyse du marquage au sol donnent deux enseignements : le premier enseignement est qu'on a, par période, sensiblement la même probabilité d'avoir un accident mortel ou grave ; le deuxième enseignement est que ce type de marquage au sol semble avoir localement un impact probable significatif sur les différents types de risque d'accidents de la zone aménagée.

4.2 Évaluation de l'effet de panneaux-réclame le long de certaines autoroutes de la ville de Montréal

4.2.1 Echangeur Décarie

Des panneaux-réclame ont été installés en octobre 1993 au niveau de l'échangeur Décarie (jonction des autoroutes 15 et 40) de l'île de Montréal. Dans le tableau 4.9, nous présentons les nombres d'accidents survenus dans cet échangeur pour les 12 mois avant et les 12 mois après l'apparition des premiers panneaux-réclame, soit de novembre 1992 à octobre 1993 et de décembre 1993 à novembre 1994.

Tableau 4.9 : Données d'accidents de l'échangeur Décarie

	période avant (12 mois)			période après (12 mois)			total
	<i>MG</i>	<i>L</i>	<i>DMS</i>	<i>MG</i>	<i>L</i>	<i>DMS</i>	
Décarie	11	50	325	12	43	300	741

Les accidents mortels ou graves (*MG*) sont passés de 11 avant l'installation des panneaux-réclame à 12 après. Il y a eu donc 11 accidents avec au moins un tué ou un blessé grave au cours des 12 mois précédant l'installation des panneaux-réclame au niveau de l'échangeur contre 12 accidents avec au moins un tué ou blessé grave au cours de la période des 12 mois suivant celle-ci. De même, les accidents légers (*L*) sont passés de 50 avant à 43 après et les accidents

III – 14

avec dommages matériels seulement (*DMS*) sont passés de 325 avant à 300 après. Afin d'évaluer l'effet des panneaux-réclame sur les accidents dans cet échangeur et apprécier son incidence sur leur gravité, on a relevé les mêmes types d'accidents sur une zone de contrôle constituée de l'échangeur Anjou et d'un tronçon de l'autoroute 40 secteur Est. Les données d'accident ainsi obtenues ont permis de dresser les coefficients de contrôle du tableau 4.10.

Tableau 4.10 : Valeurs des coefficients de contrôle

	c_1	c_2	c_3
zone de contrôle	1	0.9428	1.0424

Ces coefficients de contrôle font apparaître une stabilité du nombre des accidents mortelles ou graves ($c_1 = 1$) au niveau de la zone de contrôle, une petite baisse en proportion de 5,72% des accidents légers ($c_2 = 0,9428$) de la période après par rapport à la période avant, et une légère augmentation en proportion de 4,24% des accidents avec dommages matériels seulement ($c_3 = 1,0424$). Les paramètres estimés à partir des tableaux 4.9 et 4.10 sont présentés dans le tableau 4.11.

Tableau 4.11 : Estimation des paramètres du modèle

paramètres	cstimation (E.S)	B_{inf} à 95%	B_{sup} à 95%
$\hat{\theta}$	0.8946 (0.0657)	0.7656	1.0235
\hat{p}_1	0.0314 (0.0065)	0.0187	0.0442
\hat{p}_2	0.1307 (0.0131)	0.1051	0.1563
\hat{p}_3	0.8379 (0.0252)	0.7884	0.8875

En moyenne, après ajustement à l'aide des coefficients de contrôle, on estime que le nombre d'accidents a diminué de 10,54%, soit $(1 - 0,8946) \times 100\%$, sur la période de 12 mois après par rapport à la période de 12 mois avant l'apparition des premiers panneaux-réclame. Cependant, cette diminution n'est pas statistiquement significativement différente de 0. En effet, l'intervalle de confiance au niveau 95% associé au paramètre θ contient la valeur 1.

Tableau 4.12 : Estimation des risques d'accidents par période de temps

différents types de risque	avant : $\pi_{1j}(\hat{\Theta})$	après : $\pi_{2j}(\hat{\Theta})$
accidents MG	0.0163	0.0146
accidents L	0.0681	0.0575
accidents DMS	0.4365	0.4070

On peut conclure qu'il n'y a, en général, aucun impact statistiquement significatif des panneaux-réclame sur les nombres d'accidents répertoriés dans l'échangeur Décarie au cours de la période de 12 mois suivant leur installation comparativement à l'évolution moyenne des accidents de la zone de contrôle. En

III - 15

ce qui concerne la gravité des accidents durant les 12 mois précédant l'installation des panneaux, on note que le risque d'accidents mortels ou graves est estimé à 3,14%, celui des accidents légers à 13,07%, et le risque d'accidents avec dommages matériels seulement s'élève à 83,79%. Le tableau 4.12 montre qu'il n'y a pas d'incidence importante sur les différents risques d'accidents de la période de 12 mois après par rapport à la période de 12 mois avant l'installation des panneaux-réclame dans l'échangeur Décarie par rapport aux évolutions des nombres d'accidents de la zone de contrôle.

4.2.2 Echangeur Turcot

En juillet 1995, de nouveaux panneaux-réclame ont été installés au niveau de l'échangeur Turcot (jonction des autoroutes 15, 20 et 720) de l'île de Montréal. Pour évaluer l'effet de ces nouveaux panneaux, on combine les nombres d'accidents de cet échangeur à ceux d'une zone de contrôle obtenue par fusion de l'échangeur d'Anjou et de la jonction des autoroutes 13 et 40. Le tableau 4.13 présente les nombres d'accidents survenus dans l'échangeur Turcot pour les 12 mois précédant (juillet 1994 à juillet 1995) l'apparition des panneaux et pour les 12 mois suivant (août 1995 à juillet 1996) celle-ci.

Tableau 4.13 : Données d'accidents de l'échangeur Turcot

	période avant (12 mois)			période après (12 mois)			total
	<i>MG</i>	<i>L</i>	<i>DMS</i>	<i>MG</i>	<i>L</i>	<i>DMS</i>	
Turcot	4	20	133	3	29	143	332

Pour les mêmes périodes, des coefficients de contrôle ont été calculés pour chaque type d'accidents en utilisant les données de la zone de contrôle précédemment définie. Le tableau 4.14 ci-dessous présente la valeur de ces coefficients.

Tableau 4.14 : Valeurs des coefficients de contrôle

	c_1	c_2	c_3
zone de contrôle	4.5	1.4230	1.5519

On note que tous les coefficients de contrôle sont largement supérieurs à 1. Par conséquent, il y a eu, dans la zone utilisée comme contrôle, de très importantes augmentations en proportion de tous les types d'accidents de la période des 12 mois après l'apparition des panneaux par rapport à la période des 12 mois avant. Il y a eu 4,5 fois plus d'accidents mortels ou graves (plus de 300% d'augmentation) dans la zone de contrôle au cours de la période suivant l'installation des panneaux par rapport à la période précédant celle-ci. Pendant les mêmes périodes de temps, on dénombre 1 seul accident mortel ou grave dans l'échangeur Turcot. Les valeurs estimées des paramètres du modèle à l'aide des tableaux 4.13 et 4.14 sont :

III – 16

Tableau 4.15 : Estimation des paramètres du modèle

paramètres	estimation (E.S)	B_{inf} à 95%	B_{sup} à 95%
$\hat{\theta}$	0.7131 (0.0785)	0.5590	0.8671
\hat{p}_1	0.0106 (0.0039)	0.0027	0.0184
\hat{p}_2	0.1549 (0.0212)	0.1132	0.1966
\hat{p}_3	0.8345 (0.0390)	0.7579	0.9110

L'estimation de l'effet moyen de l'apparition des panneaux-réclame dans l'échangeur Turcot sur les nombres d'accidents est de 0,7131. Par conséquent, on estime, en moyenne, que le nombre d'accidents a baissé de 28,69%, soit $(1 - 0,7131) \times 100\%$, sur la période de 12 mois après par rapport à la période de 12 mois avant et à l'évolution moyenne des accidents de la zone de contrôle. Il y a une probabilité de 0,95 que l'effet moyen réel de l'apparition de ces panneaux dans l'échangeur Turcot soit compris entre $[0,5590; 0,8671]$. Cet intervalle de confiance ne contient pas la valeur 1, nous pouvons conclure, par rapport à la tendance moyenne des accidents de la zone de contrôle, qu'il y a un impact bénéfique probable des panneaux-réclame sur les nombres d'accidents survenus dans l'échangeur Turcot au cours de la période de 12 mois suivant leur installation. Cependant, cet impact bénéfique (baisse de 28,69% du nombre des accidents) de l'installation des panneaux-réclame dans ce cas précis est en bonne partie expliqué par l'augmentation en proportion beaucoup plus importante (coefficients de contrôle largement plus grands que 1) dans la zone de contrôle que dans l'échangeur. L'intervalle de confiance au niveau 95% associé, par exemple, au paramètre p_1 montre qu'on a une probabilité de 0,95 que la vraie valeur du risque d'accident mortel ou grave avant l'apparition des panneaux-réclame dans l'échangeur Turcot soit comprise entre 0,27% et 1,84%. Les valeurs estimées données dans le tableau 4.16 montrent une faible (on passe de 7,32% à 7,43%) augmentation du risque d'accidents légers ainsi que du risque d'accidents avec dégâts matériels seulement (on passe de 39,44% à 43,68%) de la période après installation des panneaux à celle avant. La hausse relative du risque d'accidents mortels ou graves proviendrait de l'utilisation du coefficient de contrôle élevé des accidents mortels ou graves de la zone de contrôle. On note que les deux risques du tableau 4.16 liés aux accidents mortels ou graves appartiennent à l'intervalle de confiance au niveau 95% associé au paramètre p_1 . Par conséquent, avec 95 chances sur 100, l'installation des panneaux n'a pas eu d'incidence sur le risque d'accidents mortels ou graves statistiquement significative.

Tableau 4.16 : Estimation des risques d'accidents par période de temps

différents types de risque	avant : $\pi_{1j}(\hat{\Theta})$	après : $\pi_{2j}(\hat{\Theta})$
accidents MG	0.0050	0.0161
accidents L	0.0732	0.0743
accidents DMS	0.3944	0.4368

5 Discussion

Ces différents exemples illustrent parfaitement le fait qu'en pratique, les données d'accidents du site aménagé peuvent être faibles ou importantes, les coefficients de contrôle faibles ou élevés quelle que soit la longueur des périodes d'étude. Les interprétations doivent alors tenir compte de ces valeurs : en particulier celles des coefficients de contrôle car ils ont un impact important dans les estimations.

Une limite des deux premières études est le nombre peu élevé d'accidents répertoriés malgré la longueur importante des périodes utilisées. Dans le premier cas, un total de 17 accidents a été observé sur une période de huit années d'observation dont seulement 3 accidents dans la période suivant l'aménagement du giratoire. Dans le cas du marquage à l'italienne sur une portion de la route nationale n°17, 33 accidents ont été répertoriés dont 9 au cours de la période suivant la mise en service de la zone aménagée. Ces petites fréquences d'accidents influent sur le calcul des écart-types et donc sur la longueur des intervalles de confiance pour l'estimation de l'effet moyen des aménagements de sécurité routière.

Dans le premier cas, l'effet de l'installation du giratoire est tellement important qu'il est statistiquement significatif au niveau de confiance de 95% car la borne supérieure de l'intervalle de confiance est légèrement inférieure à 1. Par contre dans le deuxième cas, malgré un effet qui semble appréciable avec une diminution en proportion du nombre des accidents de 29,46% ($\hat{\beta}_0 = 0,7054$), le petit nombre d'observations et la faible valeur des coefficients de contrôle font en sorte que cette baisse en proportion n'est pas statistiquement significative, la borne supérieure de l'intervalle de confiance étant beaucoup plus grande que 1. Dans ces conditions, il n'est pas possible de porter un jugement définitif sur l'efficacité avérée des aménagements routiers du type "marquage à l'italienne" sur les nombres d'accidents.

Les deux exemples (en particulier celui de l'échangeur Turcot) concernant l'évaluation de l'effet de panneaux-réclame montrent également que la valeur des coefficients de contrôle influencent l'estimation des paramètres ainsi que des risques de gravité malgré les effectifs d'accidents élevés. A cause de l'impact que peut avoir la valeur des coefficients de contrôle dans l'estimation de l'effet moyen des panneaux-réclame, et qu'à priori on pourrait penser que l'installation de ces derniers le long des routes en milieu urbain augmenterait les probabilités d'avoir un accident, il serait intéressant de refaire les analyses (en particulier l'exemple de l'échangeur Décarie) avec des coefficients de contrôle égaux à 1 : c'est-à-dire évaluer l'effet des panneaux en utilisant que les données d'accidents de la zone où ils sont installés.

Enfin, le modèle permet d'estimer un effet moyen d'un aménagement routier sur la sécurité des usagers. Il serait intéressant de modifier le modèle de manière à obtenir une estimation de l'effet pour chacun des types d'accidents. Il faudrait alors remplacer le β_0 par un β_j . Dans ce contexte, il faut alors voir s'il est possible d'obtenir sans trop de difficultés les estimateurs du maximum de vraisemblance sous contraintes de ces paramètres.

III – 18

Annexe : Approximation d'erreurs standards par complément de Schur

La matrice d'information Γ_Θ peut se mettre sous la forme partitionnée

$$\Gamma_\Theta = \begin{bmatrix} \tau_\Theta & U_\Theta^T & 0 \\ U_\Theta & B_\Theta & \mathbf{1}_R \\ 0 & \mathbf{1}_R^T & 0 \end{bmatrix}$$

où $\tau_\Theta = \frac{\gamma^2 \sum_m^R c_m p_m}{n \beta_0}$ est un réel strictement positif, $U_\Theta = (\frac{\gamma^2 c_1}{n}, \frac{\gamma^2 c_2}{n}, \dots, \frac{\gamma^2 c_R}{n})^T$ est un vecteur de dimension R et

$$B_\Theta = \begin{bmatrix} \gamma(\frac{1+\beta_0 c_1}{p_1} - \frac{\gamma \beta_0^2 c_1^2}{n}) & -\frac{\gamma^2 \beta_0^2 c_1 c_2}{n} & \dots & -\frac{\gamma^2 \beta_0^2 c_1 c_R}{n} \\ -\frac{\gamma^2 \beta_0^2 c_2 c_1}{n} & \gamma(\frac{1+\beta_0 c_2}{p_2} - \frac{\gamma \beta_0^2 c_2^2}{n}) & \ddots & -\frac{\gamma^2 \beta_0^2 c_2 c_R}{n} \\ \vdots & \ddots & \ddots & \vdots \\ -\frac{\gamma^2 \beta_0^2 c_R c_1}{n} & -\frac{\gamma^2 \beta_0^2 c_R c_2}{n} & \dots & \gamma(\frac{1+\beta_0 c_R}{p_R} - \frac{\gamma \beta_0^2 c_R^2}{n}) \end{bmatrix}$$

une matrice de dimension $R \times R$. On pose alors

$$J_\Theta = \begin{bmatrix} \tau_\Theta & U_\Theta^T \\ U_\Theta & B_\Theta \end{bmatrix}, \quad \Omega_\Theta = \text{diag} \left(\frac{1 + \beta_0 c_1}{p_1}, \dots, \frac{1 + \beta_0 c_R}{p_R} \right), \quad V_\Theta = \beta_0 \frac{\gamma^{1/2}}{n^{1/2}} Z$$

respectivement deux matrices de dimension $(1 + R) \times (1 + R)$ et $R \times R$ et un vecteur de dimension R . On montre alors (voir N'Guessan et Langrand 2005a,b) les résultats suivants :

- (i) $\|Z\|_{\Omega_\Theta^{-1}}^2 < \frac{\langle Z, P \rangle}{\beta_0}$; (ii) $\|V_\Theta\|_{\Omega_\Theta^{-1}}^2 < \frac{\beta_0 \langle Z, P \rangle}{1 + \beta_0 \langle Z, P \rangle} < 1$;
 (iii) B_Θ est formellement inversible et

$$B_\Theta^{-1} = \gamma^{-1}(\Omega_\Theta^{-1} + t \Omega_\Theta^{-1} V_\Theta V_\Theta^T \Omega_\Theta^{-1}),$$

avec $t = (1 - \|V_\Theta\|_{\Omega_\Theta^{-1}}^2)^{-1}$, $\langle Z, P \rangle = \sum_{m=1}^R c_m p_m$, $U_\Theta = \frac{\gamma^{3/2}}{\beta_0 n^{1/2}} V_\Theta$,
 $B_\Theta = \gamma(\Omega_\Theta - V_\Theta V_\Theta^T)$

$$(i) \quad \Delta_\Theta = \sum_{m=1}^R \frac{c_m p_m}{1 + \beta_0 c_m}, \quad (ii) \quad \|\mathbf{1}_R\|_{B_\Theta^{-1}}^2 = \gamma^{-1} [\text{trace}(\Omega_\Theta^{-1}) + t \frac{\beta_0^2 \gamma \Delta_\Theta^2}{n}]$$

$$(iii) \quad (J_\Theta / B_\Theta) = \frac{\gamma^2}{n} \left[\frac{\langle Z, P \rangle}{\beta_0} - \frac{\gamma t}{n} \|Z\|_{\Omega_\Theta^{-1}}^2 \right],$$

$$(iv) \quad \|Z\|_{\Omega_\Theta^{-1}}^2 = \sum_{m=1}^R \frac{p_m c_m^2}{1 + \beta_0 c_m}$$

(J_Θ / B_Θ) est le complément de Schur de B_Θ dans J_Θ .

III - 19

On montre que

$$\text{cov}(\hat{p}_j, \hat{p}_m) = \gamma^{-1} w_j w_m [\theta \gamma c_j c_m - A \gamma^{-1} y_j y_m],$$

$$(j \neq m = 1, 2, \dots, R)$$

$$\sigma^2(\hat{p}_j) = \gamma^{-1} [w_j + w_j^2 (\theta \gamma c_j^2 - A \gamma^{-1} y_j^2)],$$

$$(j = 1, 2, \dots, R)$$

où $w_j = \frac{p_j}{1 + \beta_0 c_j}$, $\theta = \frac{t}{n} \left[\beta_0^2 - (J_\Theta / B_\Theta)^{-1} \frac{t \gamma^2}{n} \right]$, $y_j = 1 + \theta \Delta_\Theta \gamma c_j$, ($j = 1, 2, \dots, R$).

et

$$A = \frac{1}{\|\mathbf{1}_R\|_{B_\Theta^{-1}}^2} - \frac{\sigma^2(\beta_0) t^2 \gamma^2 \Delta_\Theta^2}{n^2 \|\mathbf{1}_R\|_{B_\Theta^{-1}}^4}.$$

Références

- [1] Aitchison J., Silvey S.D. (1958) *Maximum Likelihood Estimation of parameters subject to restraints* Annals of mathematical statistics, 29,813-829.
- [2] Crowder M. (1984) *On the constrained maximum likelihood estimation with non i.i.d. observations* Ann. Inst. Statist. Math. 36, A, 239-249.
- [3] Haber M., Brown M.B. (1986) *Maximum Likelihood methods for log-linear models when expected frequencies are subject to linear constraints* J.A.S.A. vol. 81, n°394, 477-482
- [4] Matthews G.B., Crowther N.A.S. (1995) *A maximum likelihood estimation procedure when modelling in terms of constraints* South African Statist. J. n°29, 29-50
- [5] N'Guessan A. (2003) *Constrained covariance matrix estimation in road accident modelling with Schur complements* C.R. Acad. Sci. Paris, Ser. I 337 (2003) 219-222.
- [6] N'Guessan A., Essai A., Langrand C. (2001) *Estimation multidimensionnelle des contrôles et de l'effet moyen d'une mesure de sécurité routière* Rev. Statistique Appliquée. XLIX(2), 83-100.
- [7] N'Guessan A., Langrand C., (2005a) *A covariance components estimation procedure when modelling a road safety measure in terms of linear constraints*. Statistics, vol 39, No 4 (2005) 303-314.
- [8] N'Guessan A., Langrand C., (2005b) *A Schur complement approach for computing subcovariance matrices arising in a road safety measure modelling* Journal of Computational and Applied Mathematics, 177, 331-345, 2005.
- [9] N'Guessan A., Bellavance F., (2005) *A confidence interval estimation problem using the Schur complement approach, with application* C.R. Math. Rep. Acad. Sci. Canada Vol. 27, (3), 2005 pp. 84-91.

III – 20

- [10] Ouellette D.V., (1981) *Schur complements and statistics* Linear Algebra and its applications, 36 187-295,1981.
- [11] Tanner J.C, (1958) *A problem in the combination of accident frequencies* Biometrika, 45, 331-342.
- [12] Zhang F. ed, (2005) *Schur complement and Its applications* (Springer Verlag), 2005.

acknowledgement - Nous souhaitons remercier Monsieur Patrick Le Breton du SETRA ainsi que la Direction du Centre d'Etude Technique de l'Equipement (CETE) Nord Picardie pour nous avoir autorisés à utiliser les données d'accident provenant de différents aménagements routiers.

Annexe A11 :

*Technical Report, PUB IRMA, Lille, Vol. 39,
No. XI, 1-13 (1996)*

**Existence theorem for the m.l.e. of the common road
safety effect**

PUB. IRMA, LILLE 1996
Vol. 39, N° XI

Existence theorem for the m.l.e. of the common road safety effect. *

Assi N'GUESSAN

Ecole Universitaire d'Ingénieurs de Lille
e-mail : nguessan@alea.univ-lille1.fr

Claude LANGRAND

Laboratoire de Statistique et Probabilité
Université des Sciences et Technologies de Lille
F-59655 Villeneuve d'Ascq Cedex (France)

Résumé

Soit (X_{1k}, X_{2k}) , $k = 1, 2, \dots, s$, une suite de vecteurs aléatoires indépendantes suivant chacun une distribution multinomiale $\mathcal{M}(x_{\cdot,k}; 1/(1+\theta c_k), \theta c_k/(1+\theta c_k))$ où les $c_k (c_k > 0)$ sont des réels donnés et θ le paramètre à estimer. Nous discutons dans ce travail de l'existence asymptotique de l'estimateur du maximum de vraisemblance de θ .

Abstract

Let (X_{1k}, X_{2k}) , $k = 1, 2, \dots, s$, be independent random vector with multinomial distribution $\mathcal{M}(x_{\cdot,k}; 1/(1+\theta c_k), \theta c_k/(1+\theta c_k))$ where the $c_k (c_k > 0)$ are given positive real numbers and θ the parameter to be estimated. This paper deals with the asymptotic existence of the maximum likelihood estimate of θ .

AMS classifications: 62A10 - 62F10 - 62F12 - 62F30.

Key Words: road safety measure, multinomial distribution, asymptotic existence, constrained maximum likelihood, Fisher's information, Chebychev inequality.

*Preprint. Duplication of this preprint for publication or sale is strictly prohibited without prior written permission of the authors

1 Introduction

The whole problem of existence of the maximum likelihood estimate (m.l.e) of a common road safety effect is closely connected to the estimation of a finite dimensional parameter subject to restraints. Naturally this problem has received considerable attention (Aitchison et Silvey (1958), Silvey (1959), Crowder (1984), Tadashi (1984), Sutradhar (1996)). This existence involves the introduction of various assumptions.

The problem we are confronted with is that described for example in Silvey 1959 i.e. we are given $x_{..1}$ observations on a random vector X_1 whose probability density function depends on parameters $\theta, p_{11}, p_{12}, \dots, p_{r1}$, $x_{..2}$ observations on a random vector X_2 whose probability density function depends on parameters $\theta, p_{12}, p_{22}, \dots, p_{r2}, \dots, x_{..s}$ observations on a random vector X_s whose probability function depends on parameters $\theta, p_{1s}, p_{2s}, \dots, p_{rs}$. And we are presented with the problem of estimation the true parameter

$$\Theta_0 = (\theta^0, p_{11}^0, p_{21}^0, \dots, p_{r1}^0, \dots, p_{1s}^0, p_{2s}^0, \dots, p_{rs}^0)^T \in \mathbb{R}^{1+sr}$$

belongs to a set $\nu = \{\Theta \in \mathbb{R}^{1+sr}, \theta > 0; h(\Theta) = 0\}$ where $h = (h_1, h_2, \dots, h_s)$, $h_k(\Theta) = \sum_{j=1}^r p_{jk} - 1$. We will denote by x a given set of $N = x_{..1} + x_{..2} + \dots + x_{..s}$ observations on the random independents vectors (X_1, X_2, \dots, X_s) , and $l(X; \Theta) = \log L(X; \Theta)$ will denote the value of the log-likelihood function at the point Θ , where

$$(1.1) \quad L(x; \Theta) = \prod_{k=1}^s \left(\frac{x_{..k}!}{\prod_{m=1}^r x_{1mk}! x_{2mk}!} \right) \prod_{m=1}^r \frac{p_{mk}^{x_{..k}} (\theta c_{mk})^{x_{2mk}}}{(1 + \theta \sum_{j=1}^r c_{jk} p_{jk})^{x_{..k}}}$$

where $x_{..k} = \sum_{m=1}^r x_{mk}$, $x_{2.k} = \sum_{m=1}^r x_{2mk}$ and $x_{.mk} = x_{1mk} + x_{2mk}$.

So, apart from an additive constant,

$$l(X; \Theta) = \sum_{k=1}^s f_k(X_k; \Theta)$$

where $f_k(X_k; \Theta) = x_{2.k} \log(\theta) - x_{..k} \log(1 + \theta \sum_{j=1}^r c_{jk} p_{jk}) + \sum_{j=1}^r x_{.jk} \log(p_{jk})$. If H_Θ denotes the $(1+sr) \times s$ matrix $(\partial h_k / \partial \Theta_m)$, if λ is a lagrangian multiplier in \mathbb{R}^s , then we propose to estimate the unknown parameter Θ_0 by a solution ; if such exists, of the equations

$$(1.2) \quad l'(X; \Theta) + H_\Theta \lambda = 0.$$

XI - 4

$$(1.3) \quad h(\Theta) = 0;$$

where $l'(X; \Theta)$ is the point in \mathbb{R}^{1+s} whose m^{th} component is $\partial l(X; \Theta) / \partial \theta_m$. We now introduce matrix $B_\Theta = -E(\partial^2 l / \partial \theta_m \partial \theta_l)$ the Fisher's information matrix which is supposed to be positive definitive and E denotes expected value when Θ is take as the true parameter. We also assume throughout the "usual regularity conditions" : $E[l'(X; \Theta_0)] = 0$ and $\text{var}[l'(X; \Theta_0)] = B_\Theta$. If B_Θ is not positive definitive, then the model or the constraints can be used to construct a suitable non-singular version of B_Θ (see Silvey 1959, Crowder 1984). Also, conditions imposed on h are to get $H_\Theta^T B_\Theta^{-1} H_\Theta$ non singular (since H_Θ is supposed to have full rank at Θ_0) and a closed neighbourhood of Θ_0 so that $\det(H_\Theta^T B_\Theta^{-1} H_\Theta)$ is bounded away from zero. Now, if $\Theta_0 \in \nu$ then under certain conditions, asymptotics results for the constrained m.l.e. stem from those of Aitchison et Silvey (1958), Silvey (1959) and Crowder (1984) can be used to show that the m.l.e. $\hat{\Theta}$ of the common road safety effect exists, is near Θ_0 when $x_{..1}, x_{..2}, \dots, x_{..s}$ and N are all large. The general formal method will be set out in a subsequent paper. We study here the special case of Tanner's model (Tanner, 1958) i.e. when $s > 2$ and r equal one.

2 General existence theorems

We present in this paragraph existence theorems of m.l.e. subject to restraints. In order to prove that the constrained m.l.e. exists in U_δ (some neighbourhood of Θ_0), if δ is sufficiently small, we will require the following lemma (see Aitchison et Silvey, 1958).

Lemma 2.1 *If g is a continuous function mapping \mathbb{R}^n into itself with the property that, for every Θ such that $\|\Theta\| = 1$, $\Theta^T g(\Theta) < 0$, then there exists a point $\hat{\Theta}$ such that $\|\hat{\Theta}\| < 1$ and $g(\hat{\Theta}) = 0$.*

Proof : The proof of this result is indebted to Aichison and Silvey (1958, lemma 2). We suppose that $g(\Theta) \neq 0$ for any Θ such that $\|\Theta\| \leq 1$. Then the function g_1 , defined on the unit sphere in \mathbb{R}^n by

$$g_1(\Theta) = \frac{g(\Theta)}{\|g(\Theta)\|},$$

XI - 5

is a continuous function mapping of this unit sphere into itself. Hence by the fixed point theorem there is a point Θ^* in the unit sphere such that $\Theta^* = g_1(\Theta^*)$. Also since $\|g_1(\Theta^*)\| = 1$ for every Θ in the unit sphere, it follows that $\|\Theta^*\| = 1$ and $\Theta^{*T}g_1(\Theta^*) = \Theta^{*T}\Theta^* = 1 > 0$. But this contradicts the fact that $\Theta^Tg(\Theta) < 0$ (and consequently that $\Theta^Tg_1(\Theta) < 0$) for every Θ such that $\|\Theta\| = 1$. Hence there is a point $\hat{\Theta}$ in the unit sphere such that $g(\hat{\Theta}) = 0$. It is obvious that $\|\hat{\Theta}\| \neq 1$. Hence $\|\hat{\Theta}\| < 1$.

Now, let us give the general existence theorems of $\hat{\Theta}$. Application of the mean value theorem at Θ_0 to the equations (1.3) and (1.4) gives :

$$(2.1) \quad l'(X; \Theta_0) + l''(\Theta_0, \Theta)(\Theta - \Theta_0) + H(\Theta, \Theta)\lambda = 0,$$

$$(2.2) \quad H^T(\Theta_0, \Theta)(\Theta - \Theta_0) = 0 ;$$

The general notation $H(\Theta_1, \Theta_2)$ is used to convey that the rows of the matrix are evaluated at possibly different points on the line segment joining Θ_1 and Θ_2 . Let

$$R(X; \Theta) = B_{\Theta}^{-1}\{B_{\Theta} + l''(\Theta_0, \Theta)\}$$

$$Q(\Theta) = B_{\Theta}^{-1}H(\Theta_0, \Theta)\{H^T(\Theta_0, \Theta)B_{\Theta}^{-1}H(\Theta, \Theta)\}^{-1}$$

$$W(X; \Theta) = [I_{1+sr} - Q(\Theta)H^T(\Theta_0, \Theta)] \cdot \{B_{\Theta}^{-1}l'(X; \Theta_0) + R(X; \Theta)(\Theta - \Theta_0)\}.$$

Theorem 2.2 (Crowder, 1984) *There exists a consistent solution of the likelihood equations if for every $\epsilon > 0 \exists \delta \in (0, \epsilon]$ s.t.*

$$\text{Proba}[\sup_{\|\Theta - \Theta_0\| = \delta} (\Theta - \Theta_0)^T W(X; \Theta) < \delta^2] \rightarrow 1 \text{ as } N \rightarrow \infty$$

Proof : Equation (2.1) is equivalent to

$$(2.3) \quad l'(X; \Theta_0) + B_{\Theta}\{R(X; \Theta) - I_{1+sr}\}(\Theta - \Theta_0) + H(\Theta, \Theta) \cdot \lambda = 0,$$

where

$$(2.4) \quad \lambda = -\{H^T(\Theta_0, \Theta)B_{\Theta}^{-1}H(\Theta, \Theta)\}^{-1}H^T(\Theta_0, \Theta) \cdot Z(\Theta)$$

with

$$Z(\Theta) = \{B_{\Theta}^{-1}l'(X; \Theta_0) + R(X; \Theta)(\Theta - \Theta_0)\}.$$

On multiplying (2.3) by B_{Θ}^{-1} and substituting for λ one obtains

$$(2.5) \quad W(X; \Theta) - (\Theta - \Theta_0) = 0.$$

XI - 6

The pair (2.4), (2.5) are equivalent to the original equations (2.1), (2.1) : if (2.5) has a solution, say $\hat{\Theta}$, the $\hat{\lambda}$ is defined by (2.4) and these are solutions of (2.1), (2.2). Now apply the equivalent of Brouwer's fixed point Theorem (lemma 2.1) : a solution $\hat{\Theta}$ exists, with $\|\hat{\Theta} - \Theta_0\| < \delta$, if $(\Theta - \Theta_0)^T(W(X; \Theta) - (\Theta - \Theta_0))$ is strictly negative for $\|\Theta - \Theta_0\| = \delta$.

Theorem 2.3 (Crowder, 1984) *There exists a consistent solution of the likelihood equations if for every $\varepsilon > 0$, $\exists M < \infty$, $d > 0$, $\delta > 0$, $p \in [0, 1]$ s.t. $\delta < \min(d, \varepsilon)$ and*

- (i) $\|Q(\Theta)H^T(\Theta_0, \Theta)\| \leq M$ for $\|\Theta - \Theta_0\| \leq d$ and all N ,
- (ii) $\text{Prob}[\|B_{\Theta}^{-1}l'(X; \Theta_0)\| < p\delta(M + \sqrt{1 + sr})^{-1}] \rightarrow 1$ as $N \rightarrow \infty$ when $\|\Theta - \Theta_0\| = \delta$,
- (iii) $\text{Prob}[\|R(X; \Theta)\| < (1 - p)(M + \sqrt{1 + sr})^{-1}] \rightarrow 1$ as $N \rightarrow \infty$ when $\|\Theta - \Theta_0\| = \delta$.

Proof : This is based on the following, which holds for $\|\Theta - \Theta_0\| = \delta$:

$$\|(\Theta - \Theta_0)^T W(X; \Theta)\| \leq \delta \|W(X; \Theta)\| \leq \delta \|I_{1+sr} - Q(\Theta)H^T(\Theta_0; \Theta)\| \cdot \Delta(\Theta)$$

where

$$\Delta(\Theta) = \{\|B_{\Theta}^{-1}l'(X; \Theta_0)\| + \delta\|R(X; \Theta)\|\}.$$

From (i), $\sup_{\|\Theta - \Theta_0\| = \delta} \|I_{1+sr} - Q(\Theta)H^T(\Theta_0, \Theta)\| \leq M + \sqrt{1 + sr}$. Hence

$$\text{Prob}\left[\sup_{\|\Theta - \Theta_0\| = \delta} (\Theta - \Theta_0)^T W(X; \Theta) < \delta^2\right] \geq \text{Prob}\left[\sup_{\|\Theta - \Theta_0\| = \delta} \|W(X; \Theta)\| < \delta\right]$$

$$\text{Prob}\left[\sup_{\|\Theta - \Theta_0\| = \delta} \|W(X; \Theta)\| < \delta\right] \geq \text{Prob}[A_1 \cap A_2]$$

where

$$A_1 = \{\sup_{\|\Theta - \Theta_0\| = \delta} \|B_{\Theta}^{-1}l'(X; \Theta_0)\| < p\delta(M + \sqrt{1 + sr})^{-1}\}$$

and

$$A_2 = \{\sup_{\|\Theta - \Theta_0\| = \delta} \|R(X; \Theta)\| < (1 - p)(M + \sqrt{1 + sr})^{-1}\}.$$

XI - 7

$R(X, \Theta)$ is Θ -continuous and $\{\Theta; \|\Theta - \Theta_0\| = \delta\}$ is compact, so (iii) implies $\text{Proba}[\sup_{\|\Theta - \Theta_0\| = \delta} \|R(X, \Theta)\| < (1 - p)(M + \sqrt{1 + sr})^{-1}] \rightarrow 1$; similarly for $B_{\Theta}^{-1}l(X; \Theta_0)$. Hence (2.6) $\rightarrow 1$ and the criterion of Theorem 2.1 is fulfilled.

Now, for $\|\Theta - \Theta_0\| = \delta$, $\{\Theta; (\Theta - \Theta_0)^T(W(X; \Theta) - (\Theta - \Theta_0)) < 0\}$ is equal to $\{\Theta; (\Theta - \Theta_0)^T W(X; \Theta) < \delta^2\}$. Hence with probability one when $N \rightarrow \infty$ equation $W(X; \Theta) - (\Theta - \Theta_0) = 0$ has a solution $\hat{\Theta}$ such that $\|\hat{\Theta} - \Theta_0\| < \delta$ and $W(X; \hat{\Theta}) - (\hat{\Theta} - \Theta_0) = 0$ Lemma 2.1.

3 Analysis of $R(X; \Theta)$ and $B_{\Theta}^{-1}l'(X; \Theta_0)$.

Without any loss of generality we assume that $r = 1$ and $s > 2$. This is the case of the so popular Tanner's model used in road safety measure (Tanner, 1958). In this case equations (1.3) and (1.4) reduce to the unique equation

$$(3.1) \quad l'(X; \theta) = 0, \theta > 0;$$

where $l'(X; \theta) = \sum_{k=1}^s \left\{ \frac{X_{2k} - \theta c_k X_{1k}}{\theta(1 + \theta c_k)} \right\}$, X_{1k} and X_{2k} are respectively, the random "before" and "after" counts of accidents at target site k . Observations of these random variables are denoted, respectively, x_{1k} and x_{2k} and $x_{..k}$ denotes the overall total count of accidents at site k during both periods. c_k is an adjustment factor which is positive, non-random and defined by the number of accidents observed "before" and "after" at the control site k and θ is the mean policy effect parameter ($\theta > 0$). When $c_1 = c_2 = \dots = c_s = c$, equation (3.1) is immediately solved and the m.l.e. of θ is given by:

$$\hat{\theta} = \frac{\sum_{k=1}^s X_{2k}}{c \sum_{k=1}^s X_{1k}}.$$

Since the denominator is a random variable, this is a biased estimator. An approximate correction for the bias and other estimators of safety effect are proposed in Hauer (1992). We do not compare in this paper the performance of those estimators. When, the adjustment factors c_k are different it is not immediately obvious how the average effect (θ) of the road safety measure should be estimated. We focus one's attention on the asymptotic ($N \rightarrow \infty$) existence solution of equation (3.1).

XI - 8

We suppose that (X_{1k}, X_{2k}) , $(k = 1, 2, \dots, s)$ are independent random vectors. Using (3.1), equations (2.1) and (2.2) become

$$(3.2) \quad l'(X; \theta_0) + B_\theta \{R(X; \theta) - 1\}(\theta - \theta_0) = 0, \quad \theta > 0;$$

where $B_\theta = \rho(\theta)/\theta$, $\rho(\theta) = \sum_{k=1}^s \frac{x_{..k}c_k}{(1+\theta c_k)^2}$; ;

$$R(X; \theta) = \sum_{k=1}^s \frac{1}{\rho(\theta)} \left\{ \frac{x_{..k}c_k}{(1+\theta c_k)^2} - \frac{X_{2k}}{\theta(1+\theta c_k)} \right\} - \sum_{k=1}^s \frac{c_k}{\rho(\theta)} \left\{ \frac{X_{2k} - \theta c_k X_{1k}}{(1+\theta c_k)^2} \right\};$$

$$B_\theta^{-1} l'(X; \theta_0) = \frac{\theta}{\rho(\theta)} \sum_{k=1}^s \left\{ \frac{X_{2k} - \theta_0 c_k X_{1k}}{\theta_0(1+\theta_0 c_k)} \right\}.$$

Theorem 3.1 For fixed s , let $c = c_{\min} | c_{\max}$, where $c_{\min} = \min(c_k)$, $c_{\max} = \max(c_k)$, $k \in \{1, 2, \dots, s\}$. For every $\varepsilon > 0$, there exists $q_\varepsilon(N) > 0$ s.t. $q_\varepsilon(N) \geq 2(c\varepsilon N)^{-1/2}$

$$\text{Proba}[|R(X; \theta)| < q_\varepsilon(N)] \rightarrow 1 \text{ as } N \rightarrow \infty.$$

Proof : The sequence of random vectors (X_{1k}, X_{2k}) are independent and has multinomial distribution with parameter $x_{..k}$, $\frac{1}{1+\theta c_k}$ and $\frac{\theta c_k}{1+\theta c_k}$. Hence, for $k = 1, 2, \dots, s$,

$$E(X_{1k}) = \frac{x_{..k}}{(1+\theta c_k)}, \quad E(X_{2k}) = \frac{x_{..k}\theta c_k}{(1+\theta c_k)};$$

$$\text{var}(X_{1k}) = \text{var}(X_{2k}) = \frac{x_{..k}\theta c_k}{(1+\theta c_k)^2}, \quad \text{cov}(X_{1k}, X_{2k}) = -\frac{x_{..k}\theta c_k}{(1+\theta c_k)^2}.$$

So $E(R(X; \theta)) = 0$ and

$$(3.3) \quad \text{var}(R(X; \theta)) = \frac{1}{\theta \rho(\theta)} - \frac{3}{\rho^2(\theta)} \sum_{k=1}^s \frac{x_{..k}c_k^2}{(1+\theta c_k)^4} \leq \frac{1}{\theta \rho(\theta)}$$

Application of Tchebychev inequality gives, for every given $\varepsilon > 0$,

$$\text{Proba}[|R(X; \theta)| \geq \varepsilon] \leq \frac{\text{var}(R(X; \theta))}{\varepsilon^2}.$$

XI - 9

Using formula (3.3), one obtains

$$(3.4) \quad \text{Proba}[|R(X; \theta)| \geq \varepsilon] \leq \frac{1}{\theta \rho(\theta) \varepsilon^2}.$$

Now it is easy to show that

$$(3.5) \quad \frac{\theta \cdot N c_{\min}}{(1 + \theta c_{\max})^2} \leq \theta \rho(\theta) \leq \frac{\theta N c_{\max}}{(1 + \theta c_{\min})^2}.$$

The function $\theta \rightarrow \theta c_{\min} / (1 + \theta c_{\max})^2$ is bounded by $\frac{\varepsilon}{4}$ and similarly the function $\theta \rightarrow \theta c_{\max} / (1 + \theta c_{\min})^2$ is bounded by $\frac{1}{4c}$. Hence we get

$$(3.6) \quad \frac{c \cdot N}{4} \leq \theta \rho(\theta) \leq \frac{N}{4c}.$$

Using (3.6) and (3.4), one obtains

$$(3.7) \quad \text{Proba}[|R(X; \theta)| \geq \varepsilon] \leq \frac{4}{\varepsilon^2 c N}.$$

Let $q_\varepsilon(N) > 0$ so that $q_\varepsilon(N) \geq 2c^{-1/2} \varepsilon^{-1/2} N^{-1/2}$, then

$$(3.8) \quad \text{Proba}\{|R(X; \theta)| \geq q_\varepsilon(N)\} \leq \text{Proba}\{|R(X; \theta)| \geq 2c^{-1/2} \varepsilon^{-1/2} N^{-1/2}\}.$$

Application of (3.7) gives

$$\text{Proba}\left\{|R(X; \theta)| \geq \frac{2}{c^{1/2} \varepsilon^{1/2} N^{1/2}}\right\} \leq \frac{4}{\left(\frac{2}{c^{1/2} \varepsilon^{1/2} N^{1/2}}\right)^2 c N} = \varepsilon,$$

so

$$1 - \varepsilon \leq \text{Proba}\{|R(X; \theta)| \leq q_\varepsilon(N)\} \leq 1$$

for every given $\varepsilon > 0$ and $N \rightarrow \infty$ (N large).

Theorem 3.2 Let $\mu(\theta_0) = (1 + \theta_0 c_{\max})^2 / c_{\min}$, where c_{\max} and c_{\min} are defined as in theorem 3.1 and θ_0 the true value of parameter θ . For every $\varepsilon > 0$, there exists $p_\varepsilon(N) > 0$ s.t. $p_\varepsilon(N) \geq \frac{\mu(\theta_0)}{\varepsilon^{1/2} N^{1/2}}$ $\text{Proba}[|B_\theta^{-1} l'(X; \theta_0)| < p_\varepsilon(N)] \rightarrow 1$ as $N \rightarrow \infty$ and θ near θ_0 .

XI - 10

Proof : Using Tchebychev inequality, one obtains

$$(3.9) \quad \text{Proba}[|B_{\theta}^{-1}l'(X; \theta_0)| \geq \varepsilon] \leq \frac{\text{var}(B_{\theta}^{-1}l'(X; \theta_0))}{\varepsilon^2}$$

for every given $\varepsilon > 0$. Now using expression of $B_{\theta}^{-1}l'(X; \theta_0)$, we get

$$\text{var}(B_{\theta}^{-1}l'(X; \theta_0)) = \frac{\theta^2}{\rho^2(\theta)} \cdot \frac{1}{\theta_0^2} \sum_{k=1}^s \frac{x_k \theta_0 c_k}{(1 + \theta_0 c_k)^2},$$

so

$$(3.10) \quad \text{var}(B_{\theta}^{-1}l'(X; \theta_0)) \leq \frac{N \cdot \theta^2}{\rho^2(\theta) \theta_0^2}.$$

For θ near θ_0 (3.10) becomes

$$(3.11) \quad \text{var}(B_{\theta}^{-1}l'(X; \theta_0)) \leq \frac{N}{\rho^2(\theta_0)} \leq \frac{(1 + \theta_0 c_{\max})^4}{N \cdot c_{\min}^2}.$$

Hence

$$(3.12) \quad \text{Proba}[|B_{\theta}^{-1}l'(X; \theta_0)| \geq \varepsilon] \leq \frac{(1 + \theta_0 c_{\max})^4}{N \varepsilon^2 c_{\min}^2}.$$

Now let $p_{\varepsilon}(N) > 0$ so that $p_{\varepsilon}(N) \geq \frac{\mu(\theta_0)}{\varepsilon^{1/2} N^{1/2}}$ and using (3.12), one obtains

$$\text{Proba}[|B_{\theta}^{-1}l'(X; \theta_0)| \geq p_{\varepsilon}(N)] \leq \text{Proba}[|B_{\theta}^{-1}l'(X; \theta_0)| \geq \frac{\mu(\theta_0)}{\varepsilon^{1/2} N^{1/2}}] \leq \varepsilon.$$

Hence $1 - \varepsilon \leq \text{Proba}[|B_{\theta}^{-1}l'(X; \theta_0)| < p_{\varepsilon}(N)] \leq 1$ for $N \rightarrow \infty$ and θ near θ_0 .

4 Asymptotic existence of the m.l.e.

Function $W(X; \theta)$ is defined by :

$$(4.1) \quad W(X; \theta) = B_{\theta}^{-1}l'(X; \theta_0) + R(X; \theta)(\theta - \theta_0).$$

A solution $\hat{\theta}$ exists, with $|\hat{\theta} - \theta_0| < \delta$, if $(\theta - \theta_0)^T \{W(X; \theta) - (\theta - \theta_0)\}$ is strictly negative for $|\theta - \theta_0| = \delta$.

Theorem 4.1 *There exists a consistent solution of the likelihood equation (3.1) if for every $\varepsilon > 0$, $\exists q_{\varepsilon}(N) \in [\frac{2}{\varepsilon^{1/2} \varepsilon^{1/2} N^{1/2}}, 1[$, $\delta > 0$ s.t. $\delta > \frac{\mu(\theta_0)}{(1 - q_{\varepsilon}(N)) \varepsilon^{1/2} N^{1/2}}$ and*

XI - 11

- (i) $\text{Proba}[|R(X; \theta)| < q_\epsilon(N)] \rightarrow 1$ as $N \rightarrow \infty$ when $|\theta - \theta_0| = \delta$.
- (ii) $\text{Proba}[|B_\theta^{-1}l'(X; \theta_0)| < (1 - q_\epsilon(N))\delta] \rightarrow 1$ as $N \rightarrow \infty$ when $|\theta - \theta_0| = \delta$.

Proof : For $|\theta - \theta_0| = \delta$.

$$(4.2) \quad |(\theta - \theta_0)^T W(X; \theta)| \leq \delta |W(X; \theta)| \leq \delta \{|B_\theta^{-1}l'(X; \theta_0)| + \delta |R(X; \theta)|\}.$$

Using theorem (3.1) there exists $q_\epsilon(N) \geq \frac{2}{\epsilon^{1/2} \sigma^{1/2} N^{1/2}}$ so that

$$(4.3) \quad \text{Proba}[|R(X; \theta)| < q_\epsilon(N)] \rightarrow 1 \text{ as } N \rightarrow \infty.$$

Similarly using theorem (3.2), we get $(p_\epsilon(N) \geq \mu(\theta_0)\epsilon^{-1/2}N^{-1/2})$

$\text{Proba}[|B_\theta^{-1}l'(X; \theta_0)| < p_\epsilon(N)] \rightarrow 1$ as $N \rightarrow \infty$ and θ near θ_0 (ie. δ sufficiently small). So with probability one and θ near θ_0 ,

$$(4.4) \quad |B_\theta^{-1}l'(X; \theta_0)| + \delta |R(X; \theta)| < p_\epsilon(N) + \delta q_\epsilon(N),$$

and $p_\epsilon(N) + \delta q_\epsilon(N) < \delta$ i.e. $p_\epsilon(N) < (1 - q_\epsilon(N))\delta$ if $1 - q_\epsilon(N) > 0$ and $\delta > \frac{p_\epsilon(N)}{1 - q_\epsilon(N)}$.

This implies that

$$(4.5) \quad \text{proba}[|B_\theta^{-1}l'(X; \theta_0)| < p_\epsilon(N)] \leq \text{Proba}[|B_\theta^{-1}l'(X; \theta_0)| < (1 - q_\epsilon(N))\delta].$$

Condition (ii) implies that $\text{Proba}[|B_\theta^{-1}l'(X; \theta_0)| < p_\epsilon(N)] \rightarrow 1$ as $N \rightarrow \infty$ when $|\theta - \theta_0| = \delta$. Using expressions (4.5) and (4.3), one obtains

$$|B_\theta^{-1}l'(X; \theta_0)| + \delta |R(X; \theta)| < (1 - q_\epsilon(N))\delta + \delta q_\epsilon(N) = \delta.$$

Hence, for $|\theta - \theta_0| = \delta$

$$|(\theta - \theta_0)^T W(X; \theta)| \leq \delta |W(X; \theta)| \leq \delta \{(1 - q_\epsilon(N))\delta + \delta q_\epsilon(N)\} = \delta^2,$$

with probability one as $N \rightarrow \infty$ i.e. condition of theorem 2.1 is then fulfilled.

Remark. Useful discussions about theorem 2.1 and theorem 2.2 are given in Crowder (1984). Theorem 2.2 gives a breakdown of the single condition of theorem 2.1 into more manageable bits. The reduction depends critically upon (i) of theorem 2.2, that $Q(\Theta)H^T(\Theta_0, \Theta)$ should be bounded uniformly

XI - 12

in both N and Θ . However with regard to the special case $r = 1$ and $s > 2$ there is no constraints. Hence $Q(\Theta)H^T(\Theta_0, \Theta)$ is equal to nul function and any positive real number can be used as M . On the other hand, one can take $p_\varepsilon(N)$ equal to $\mu(\theta_0)(\varepsilon N)^{-\frac{1}{2}}$ and $q_\varepsilon(N)$ equal to $2(\varepsilon cN)^{-\frac{1}{2}}$ in such a way that

$$(4.6) \quad \varepsilon > \delta > \frac{c^{\frac{1}{2}}\mu(\theta_0)}{(\varepsilon cN)^{\frac{1}{2}} - 2},$$

for a given $\varepsilon > 0$ and N large.

References

- [1] AITCHISON J., SILVEY S.D. (1958) Maximum likelihood estimation of parameters subject to restraints, *Ann. Math. Statist.*, 29, 813-828.
- [2] CRESSIE N., LAHIRI S.N. (1993) The asymptotic distribution of REML estimators, *Journal of Multivariate Analysis*, 45, 217-233.
- [3] CROWDER M. (1984) On constrained maximum likelihood estimation with non i.i.d. observations, , *Ann. Inst. Statist. Mth.*, 36, part A, 239-249.
- [4] HAUER E. (1992) A note on three estimators of safety effect. *Traffic Engineering and Control*, 388-3936
- [5] SILVEY S.D. (1959) The Langrangian multiplier test, *Ann. Math. Statist.*, 30, 389-407.
- [6] SUTRADHAR B.C. (1996) On joint estimation of regression and overdispersion paramaters in generalized linear models for longitudinal data, *J. of Multivariate Analysis*, 56, 90-119.

XI - 13

- [7] TADASHI N. (1984) Existence theorems of a MLE from a generalized censored data sample *Ann. Inst. Statist. Math.*, 36, part A, 375-393.

Annexe A12 :

*Technical Report, PUB IRMA, Lille, Vol. 47,
No. VII, 1-8 (1998)*

Etude asymptotique de l'existence d'un estimateur lié à
des données d'accidents de la route

PUB. IRMA, LILLE 1998
Vol. 47, N° VII

Etude asymptotique de l'existence d'un estimateur lié à des données d'accidents de la route

Assi N'GUESSAN

Laboratoire de Statistique et Probabilités EP 1765
U.F.R. de Mathématiques Pures et Appliquées, Bât. M2
Université des Sciences et Technologies de Lille
F-59665 Villeneuve d'Ascq Cedex, France

Résumé

On considère (Y_{1k}, Y_{2k}) , $(k = 1, 2, \dots, s)$, $s \in \mathbb{N}^*$ une suite de couples de vecteurs aléatoires (v.a.) indépendants suivant chacun une distribution multinomiale dont les probabilités de classe dépendent d'un vecteur paramètre inconnu $\Phi \in \mathbb{R}^d$, $d \in \mathbb{N}^*$. On s'intéresse à l'estimation de Φ^0 (la vraie valeur du paramètre) lorsqu'on utilise la méthode du chi-deux minimum modifié. On montre, sous certaines conditions, qu'il existe asymptotiquement un estimateur $\tilde{\Phi}$ convergeant en probabilité vers Φ^0 .

Abstract

Let (Y_{1k}, Y_{2k}) , $(k = 1, 2, \dots, s)$, $s \in \mathbb{N}^*$, be independent random vectors with multinomial distribution where the cell probabilities depend on an unknown parameter $\Phi \in \mathbb{R}^d$, where $d \in \mathbb{N}^*$. One shows, under certain conditions, that there exists asymptotically an estimator derived from the modified chi-square method which converges in probability to Φ^0 (the true value of the parameter).

Key words : Road safety measure, road accident, multinomial distribution, Estimation, chi-deux minimum method, asymptotic existence, convergence in Probability.

Mots clés : Accident de la route, Mesure de sécurité routière, distribution multinomiale, Estimation, chi-deux minimum, Existence asymptotique.

A.M.S. Subject classification : 62 F 10, 62 F 12, 62 H 12.

VII - 3

1 Introduction

On considère $k = 1, 2, \dots, s$ ($s \in \mathbb{N}^*$) et $Y_{1k} = (Y_{11k}, Y_{12k}, \dots, Y_{1rk})$ (resp. $Y_{2k} = (Y_{21k}, Y_{22k}, \dots, Y_{2rk})$) le vecteur aléatoire (v.a.) donnant les effectifs de r ($r \in \mathbb{N}^*$) types d'accidents sur le site traité k avant (resp. après) la mise en place d'une mesure de sécurité routière. Afin de tester l'hypothèse selon laquelle l'indice d'efficacité de la mesure est le même quelque soit le site traité et quelque soit le type d'accidents considéré, N'guessan et Langrand (1993) ont mis au point certaines statistiques et étudié leur distribution asymptotique. Ces statistiques ont été élaborées en supposant que les couples de v.a. (Y_{1k}, Y_{2k}) , ($k = 1, 2, \dots, s$) sont indépendants et suivent chacun une distribution multinomiale (voir N'guessan 1993 pour la construction du modèle multinomial).

$$\mathcal{M}(n_k; \pi_{11k}(\Phi), \dots, \pi_{1jk}(\Phi), \dots, \pi_{1rk}(\Phi), \\ \pi_{21k}(\Phi), \dots, \pi_{2jk}(\Phi), \dots, \pi_{2rk}(\Phi))$$

où n_k est l'effectif total des accidents survenus dans le site k au cours des deux périodes (n_k est déterminé), $\Phi \in \mathbb{R}^{1+s(r-1)}$ un vecteur paramètre de vraie valeur Φ° inconnue. Pour construire et établir la loi asymptotique d'une de leur statistique, ces auteurs ont utilisé $\tilde{\Phi}$, l'estimateur du χ^2 minimum modifié (voir Cràmer p. 425-426) de Φ° c'est-à-dire une solution, si elle existe, du système d'équations :

$$(1) \quad \sum_{k=1}^s \sum_{j=1}^r \sum_{t=1}^2 \frac{Y_{tjk} - n_k \pi_{tjk}}{\pi_{tjk}} \cdot \frac{\partial \pi_{tjk}}{\partial \Phi_m} = 0; \quad m = 1, 2, \dots, 1 + s(r-1).$$

Cependant les résultats asymptotiques ainsi obtenus n'abordent pas le problème primordial de l'existence de $\tilde{\Phi}$. On se propose donc d'étudier, dans ce travail, l'existence asymptotique ($N = \sum_{k=1}^s n_k \rightarrow \infty$) de $\tilde{\Phi}$.

2 Principaux résultats

On note :

D_0 la matrice de taille $(2sr, 1 + s(r-1))$ dont chaque élément est donné par l'expression $\pi_{tjk}^{-1/2} (\partial \pi_{tjk} / \partial \Phi_m)$ évalué au point $\Phi = \Phi^\circ$;

$D = \Lambda^{1/2} D_0$ [où $\Lambda^{1/2} = \text{diag}(\Lambda_1^{1/2}, \Lambda_2^{1/2}, \dots, \Lambda_s^{1/2})$ est une matrice bloc-diagonale de taille $2sr$ avec $\Lambda_k^{1/2} = n_k^{1/2} I_{2r}$];

$Z \in \mathbb{R}^{2sr}$ le v.a. tel que $Z_{tjk} = [Y_{tjk} - n_k \pi_{tjk}^\circ] / [n_k \pi_{tjk}^\circ]^{1/2}$ avec $\pi_{tjk}^\circ = \pi_{tjk}(\Phi^\circ)$;

VII - 4

$$\begin{aligned}
 R_m^{(k)}(\Phi) &= \sum_{j=1}^r \sum_{t=1}^2 \frac{Y_{tjk} - n_k \pi_{tjk}^\circ}{n_k} \left[\frac{1}{\pi_{tjk}} \frac{\partial \pi_{tjk}}{\partial \Phi_m} - \frac{1}{\pi_{tjk}^\circ} \left(\frac{\partial \pi_{tjk}}{\partial \Phi_m} \right)_\circ \right] \\
 &- \sum_{j=1}^r \sum_{t=1}^2 (\pi_{tjk} - \pi_{tjk}^\circ) \left[\frac{1}{\pi_{tjk}} \frac{\partial \pi_{tjk}}{\partial \Phi_m} - \frac{1}{\pi_{tjk}^\circ} \left(\frac{\partial \pi_{tjk}}{\partial \Phi_m} \right)_\circ \right] \\
 &- \sum_{j=1}^r \sum_{t=1}^2 \frac{1}{\pi_{tjk}^\circ} \left(\frac{\partial \pi_{tjk}}{\partial \Phi_m} \right)_\circ \left[(\pi_{tjk} - \pi_{tjk}^\circ) - \sum_{\ell=1}^{1+s(r-1)} \left(\frac{\partial \pi_{tjk}}{\partial \Phi_\ell} \right)_\circ (\Phi_\ell - \Phi_\ell^\circ) \right]
 \end{aligned}$$

$k = 1, 2, \dots, s; m = 1, 2, \dots, 1 + s(r - 1)$ fixés.

On admettra dans la suite les conditions ci-dessous ;

- (H1) $\sum_{j=1}^r \sum_{t=1}^2 \pi_{tjk}(\Phi) = 1$ pour $k = 1, 2, \dots, s$
- (H2) $\partial \pi_{tjk} \mid \partial \Phi_m$ et $\partial^2 \pi_{tjk} \mid \partial \Phi_m \partial \Phi_\ell$ sont continues
- (H3) la matrice de taille $(2sr, 1 + s(r - 1))$ dont les éléments sont $\partial \pi_{tjk} \mid \partial \Phi_m$ est de rang $1 + s(r - 1)$;
- (H4) $\pi_{tjk}(\Phi) > \rho^2 > 0$ pour tout k, j et t ;
- (H5) $n_k = n_k(N)$ pour $k = 1, 2, \dots, s$ et N et $n_k(N)$ tendent vers ∞ de telle sorte que

$$\lim_{N \rightarrow \infty} \frac{n_k(N)}{N} = \tau_k < \infty .$$

Soit $\delta_N > 0$ donné, pour k, j et t fixés et d'après l'inégalité de Tchebycheff, on a

$$(2) \quad \text{Proba} \left[|Y_{tjk} - n_k \pi_{tjk}^\circ| \geq \delta_N n_k^{1/2} \right] \leq \frac{\pi_{tjk}^\circ (1 - \pi_{tjk}^\circ)}{\delta_N^2} < \frac{\pi_{tjk}^\circ}{\delta_N^2}$$

Par conséquent

$$(3) \quad |Y_{tjk} - n_k \pi_{tjk}^\circ| < \delta_N n_k^{1/2} \quad \forall k, j, t$$

avec une probabilité plus grande que $1 - \delta_N^{-2}$. On supposera dans la suite de ce travail que les Y_{tjk} satisfont à la relation (3) et que $\delta_N \rightarrow \infty$ quand $N \rightarrow \infty$ de telle sorte que $N^{-1/2} \delta_N^2 \rightarrow 0$. On pourra prendre $\delta_N = N^a$ ($0 < a < \frac{1}{4}$).

Lemme 1 Soient $v_N(\Phi^\circ)$ un voisinage ouvert non vide de Φ° et $\bar{R}(\Phi) \in \mathbb{R}^{1+s(r-1)}$ le vecteur aléatoire (v.a.) de composantes $\bar{R}_m(\Phi) = \frac{1}{N} \sum_{k=1}^s n_k R_m^{(k)}(\Phi)$, $\Phi \in v_N(\Phi^\circ)$. Sous les hypothèses (H1) à (H5) et avec une probabilité plus grande que $1 - \delta_N^{-2}$, on a pour $\Phi^{(2)}$ et $\Phi^{(1)}$ de $v_N(\Phi^\circ)$

$$|\bar{R}_m(\Phi^{(2)}) - \bar{R}_m(\Phi^{(1)})| \leq C_0 \|\Phi^{(2)} - \Phi^{(1)}\| \left(\|\Phi^{(1)} - \Phi^\circ\| + \|\Phi^{(2)} - \Phi^\circ\| + \tau_0^{1/2} \frac{\delta_N}{N^{1/2}} \right)$$

où C_0 est une constante positive indépendante de $\Phi^{(2)}, \Phi^{(1)}, m, N$ et $\delta_N \rightarrow \infty$ quand $N \rightarrow \infty$, avec $\delta_N^2 N^{-1/2} \rightarrow 0$, et $\tau_0^{1/2} = \lim_{N \rightarrow \infty} \sum_{k=1}^s \left(\frac{n_k}{N} \right)^{1/2}$

VII - 5

Lemme 2 *Sous les hypothèses (H1) à (H5), on définit Φ_p une suite de points de $v_N(\Phi^\circ)$ telle que $\Phi_0 = \Phi^\circ$ et*

$$\Phi_p = \Phi^\circ + N^{-1/2}(D_0^T \Gamma D_0)^{-1}(\Gamma^{1/2} D_0)^T Z + (D_0^T \Gamma D_0)^{-1} \bar{R}(\Phi_{p-1})$$

où $p = 1, 2, \dots$, $\Gamma^{1/2} = \frac{1}{\sqrt{N}} \Lambda^{-1/2}$. Alors, avec une probabilité plus grande que $1 - \delta_N^{-2}$, on a pour $p > 0$

$$\|\Phi_{p+1} - \Phi_p\| \leq C_2 [(2C_2 + \tau_0^{1/2}) C_3]^p \left(\frac{\delta_N}{\sqrt{N}}\right)^{p+1}$$

où C_2 et C_3 sont des constantes indépendantes de N et p et où $\tau_0^{1/2}$ et $\bar{R}(\cdot)$ sont définies comme au lemme 1.

Théorème 1 *Sous les conditions (H1) à (H5), il existe asymptotiquement ($N \rightarrow \infty$) un estimateur du χ^2 minimum modifié $\tilde{\Phi}$ de Φ° tel que $\forall \delta_N > 0$ donné*

$$\|\tilde{\Phi} - \Phi^\circ\| \rightarrow 0 \quad \text{quand } N \rightarrow \infty$$

avec une probabilité plus grande que $1 - \delta_N^{-2}$, où $\delta_N \rightarrow \infty$ quand $N \rightarrow \infty$ tel que $N^{-1/2} \delta_N^2 \rightarrow 0$.

3 Ligne de démonstration

3.1 Démonstration du théorème

Soit $\tilde{\Phi}$ l'estimateur du χ^2 minimum modifié, s'il existe, de Φ° , alors sous les hypothèses (H1) à (H5), le système d'équations (1) est équivalent à la forme matricielle

$$(4) \quad \tilde{\Phi} = \Phi^\circ + N^{-1/2}(D_0^T \Gamma D_0)^{-1}(\Gamma^{1/2} D_0)^T Z + (D_0^T \Gamma D_0)^{-1} \bar{R}(\tilde{\Phi})$$

où $\tilde{\Phi} \in v_N(\Phi^\circ)$ ($N \rightarrow \infty$), $\Gamma^{1/2} = N^{-1/2} \Lambda^{1/2}$ et $\bar{R}(\tilde{\Phi})$ donné par le lemme 1. Par suite, tout point $\tilde{\Phi} \in v_N(\Phi^\circ)$ solution de (1) se met sous la forme (4).

Réciproquement, il s'agit de montrer qu'il existe asymptotiquement un point de $v_N(\Phi^\circ)$, vérifiant l'expression (4) et donc solution de (1). On considère donc Φ_ℓ la suite de points de $v_N(\Phi^\circ)$ définie par $\Phi_0 = \Phi^\circ$ et

$$(5) \quad \Phi_\ell = \Phi^\circ + N^{-1/2}(D_0^T \Gamma D_0)^{-1}(\Gamma^{1/2} D_0)^T Z + (D_0^T \Gamma D_0)^{-1} \bar{R}(\Phi_{\ell-1}), \ell > 0.$$

D'après le lemme 2, il existe des constantes K_2 et K_3 indépendantes de N et ℓ telles que

$$(6) \quad \|\Phi_{\ell+1} - \Phi_\ell\| \leq K_2 [(2K_2 + \tau_0^{1/2}) K_3]^\ell \left(\frac{\delta_N}{\sqrt{N}}\right)^{\ell+1}$$

avec une probabilité plus grande que $1 - \delta_N^{-2}$ quand $N \rightarrow \infty$ tel que $N^{-1/2} \delta_N^2 \rightarrow 0$. Par conséquent pour $N \rightarrow \infty$ la suite Φ_1, Φ_2, \dots de $v_N(\Phi^\circ)$ converge vers un point $\Phi^* \in v_N(\Phi^\circ)$ avec

$$(7) \quad \Phi^* - \Phi^\circ = (\Phi_1 - \Phi^\circ) + (\Phi_2 - \Phi_1) + (\Phi_3 - \Phi_2) + \dots$$

VII - 6

Par construction de la suite $(\Phi_\ell, \ell \geq 1)$, Φ^* satisfait à l'équation matricielle (4) et donc est solution de (1) avec une probabilité plus grande que $1 - \delta_N^{-2}$. D'après le lemme 2 et en utilisant la décomposition (7), on a que $\|\Phi^* - \Phi^\circ\| \rightarrow 0$ quand $N \rightarrow \infty$ avec une probabilité plus grande que $1 - \delta_N^{-2}$ c'est-à-dire tendant vers 1 quand $N \rightarrow \infty$. Il existe alors, avec une probabilité tendant vers 1, une solution du chi-deux minimum modifié au système (1).

Soit maintenant $\tilde{\Phi} \in v_N(\Phi^\circ)$ une autre solution de (1) tendant vers Φ° quand $N \rightarrow \infty$ alors d'après le lemme 1 on a

$$(8) \quad \begin{aligned} \|\tilde{\Phi} - \Phi^*\| &= \|(D_0^T \Gamma D_0)^{-1} (\bar{R}(\Phi^*) - \bar{R}(\tilde{\Phi}))\| \\ &\leq C_3 \|\Phi^* - \tilde{\Phi}\| \left(\|\tilde{\Phi} - \Phi^\circ\| + \|\Phi^* - \Phi^\circ\| + \tau_0^{1/2} \frac{\delta_N}{\sqrt{N}} \right) \end{aligned}$$

Puisque $\|\Phi^* - \Phi^\circ\|$ et $\|\tilde{\Phi} - \Phi^\circ\|$ tendent vers 0 quand $N \rightarrow \infty$ tel que $N^{-1/2} \delta_N^2 \rightarrow 0$ alors $\tilde{\Phi} = \Phi^*$ avec une probabilité plus grande que $1 - \delta_N^{-2}$ quand $N \rightarrow \infty$. Par conséquent, il existe asymptotiquement une solution unique $\tilde{\Phi}$ au système (4) et donc à (1) convergeant en probabilité vers Φ° .

3.2 Démonstration du lemme 1

On considère $\Phi^{(2)}$ et $\Phi^{(1)}$ deux points quelconques de $v_N(\Phi^\circ)$. Soient k et m fixés alors on obtient la relation ci-dessous

$$\begin{aligned} R_m^{(k)}(\Phi^{(2)}) - R_m^{(k)}(\Phi^{(1)}) &= \\ &\sum_{j=1}^r \sum_{t=1}^2 \frac{Y_{tjk} - n_k \pi_{tjk}^\circ}{n_k} \left[\frac{1}{\pi_{tjk}(\Phi^{(1)})} \left(\frac{\partial \pi_{tjk}}{\partial \Phi_m} \right)_{\Phi^{(1)}} - \frac{1}{\pi_{tjk}(\Phi^{(2)})} \left(\frac{\partial \pi_{tjk}}{\partial \Phi_m} \right)_{\Phi^{(2)}} \right] \\ &+ \sum_{j=1}^r \sum_{t=1}^2 (\pi_{tjk}(\Phi^{(1)}) - \pi_{tjk}^\circ) \left[\frac{1}{\pi_{tjk}(\Phi^{(2)})} \left(\frac{\partial \pi_{tjk}}{\partial \Phi_m} \right)_{\Phi^{(2)}} - \frac{1}{\pi_{tjk}(\Phi^{(1)})} \left(\frac{\partial \pi_{tjk}}{\partial \Phi_m} \right)_{\Phi^{(1)}} \right] \\ &+ \sum_{j=1}^r \sum_{t=1}^2 (\pi_{tjk}(\Phi^{(2)}) - \pi_{tjk}(\Phi^{(1)})) \left[\frac{1}{\pi_{tjk}(\Phi^{(2)})} \left(\frac{\partial \pi_{tjk}}{\partial \Phi_m} \right)_{\Phi^{(2)}} - \frac{1}{\pi_{tjk}^\circ} \left(\frac{\partial \pi_{tjk}}{\partial \Phi_m} \right)_{\Phi^0} \right] \\ &+ \sum_{j=1}^r \sum_{t=1}^2 \frac{1}{\pi_{tjk}^\circ} \left(\frac{\partial \pi_{tjk}}{\partial \Phi_m} \right)_0 \left[\pi_{tjk}(\Phi^{(2)}) - \pi_{tjk}(\Phi^{(1)}) \right] \\ &- \sum_{\ell=1}^{1+s(r-1)} \left(\frac{\partial \pi_{tjk}}{\partial \Phi_\ell} \right)_0 (\Phi_\ell^{(2)} - \Phi_\ell^{(1)}) \end{aligned}$$

La preuve du lemme consiste à majorer en probabilité chacun des termes du second membre de cette égalité. En utilisant la relation (3), les hypothèses (H1) et (H5) et le développement de Taylor à l'ordre 1 au voisinage de Φ° , le premier terme du second membre de l'égalité est majorée par l'expression $K_1 \|\Phi^{(1)} - \Phi^{(2)}\| \frac{\delta_N}{n_k^{1/2}}$, avec une probabilité plus grande que $1 - \delta_N^{-2}$ où K_1

VII - 7

est une constante indépendante de k, m et n_k ($k = 1, 2, \dots, s$). En appliquant le même raisonnement aux trois autres termes, on montre que :

$$(9) \quad \begin{aligned} |R_m^{(k)}(\Phi^{(1)}) - R_m^{(k)}(\Phi^{(2)})| &\leq K_1 \|\Phi^{(2)} - \Phi^{(1)}\| \frac{\delta_N}{n_k^{1/2}} \\ &+ K_2 \|\Phi^{(1)} - \Phi^{(0)}\| \|\Phi^{(1)} - \Phi^{(2)}\| \\ &+ K_3 \|\Phi^{(2)} - \Phi^{(0)}\| \|\Phi^{(1)} - \Phi^{(2)}\| \\ &+ K_4 \|\Phi^{(2)} - \Phi^{(1)}\|^2. \end{aligned}$$

En posant $C_0 = \max(K_1, K_2 + K_4, K_3 + K_4)$, il s'en suit, avec une probabilité plus grande que $1 - \delta_N^{-2}$, que

$$(10) \quad |R_m^{(k)}(\Phi^{(1)}) - R_m^{(k)}(\Phi^{(2)})| \leq C_0 \|\Phi^{(2)} - \Phi^{(1)}\| \left(\|\Phi^{(1)} - \Phi^{(0)}\| + \|\Phi^{(2)} - \Phi^{(0)}\| + \frac{\delta_N}{n_k^{1/2}} \right)$$

où C_0 est une constante indépendante de m, k, n_k ($k = 1, 2, \dots, s$), $\Phi^{(1)}$ et $\Phi^{(2)}$. On déduit de cette dernière inégalité que pour $N \rightarrow \infty$

$$(11) \quad \begin{aligned} |\bar{R}_m(\Phi^{(2)}) - \bar{R}_m(\Phi^{(1)})| &\leq \frac{1}{N} \sum_{k=1}^s n_k |R_m^{(k)}(\Phi^{(2)}) - R_m^{(k)}(\Phi^{(1)})| \\ &\leq C_0 \|\Phi^{(2)} - \Phi^{(1)}\| \left(\|\Phi^{(2)} - \Phi^{(0)}\| + \|\Phi^{(1)} - \Phi^{(0)}\| \right. \\ &\quad \left. + \tau_0^{1/2} \cdot \frac{\delta_N}{\sqrt{N}} \right) \end{aligned}$$

où $\tau_0^{1/2} = \lim_{N \rightarrow \infty} \sum_{k=1}^s \left(\frac{n_k}{N}\right)^{1/2}$, $\delta_N \rightarrow \infty$ avec N de telle sorte que $\delta_N^2 N^{-1/2} \rightarrow 0$.

3.3 Démonstration du lemme 2

Par construction des $R_m^{(k)}(\Phi)$, on a que $R_m^{(k)}(\Phi^0) = 0 \forall m$ et k . Par suite on a

$$(12) \quad \Phi_1 - \Phi^0 = N^{-1/2} (D_0^T \Gamma D_0)^{-1} (\Gamma^{1/2} D_0)^T Z$$

et pour $p > 0$

$$(13) \quad \Phi_{p+1} - \Phi_p = (D_0^T \Gamma D_0)^{-1} (\bar{R}(\Phi_p) - \bar{R}(\Phi_{p-1})).$$

D'après (H5), les matrices $(D_0^T \Gamma D_0)^{-1} (\Gamma^{1/2} D_0)^T$ et $(D_0^T \Gamma D_0)^{-1}$ sont asymptotiquement ($N \rightarrow \infty$) indépendantes de N . Par conséquent en combinant les relations (3), (12) et l'hypothèse (H4), on obtient

$$(14) \quad \|\Phi_1 - \Phi^0\| < C_2 \frac{\delta_N}{\sqrt{N}}$$

VII - 8

où C_2 est une constante indépendante de N . En utilisant de la même façon la relation (13) et le lemme 1, on a pour $p > 0$

(15)

$$\|\Phi_{p+1} - \Phi_p\| \leq C_3 \|\Phi_p - \Phi_{p-1}\| \left(\|\Phi_p - \Phi^o\| + \|\Phi_{p-1} - \Phi^o\| + \tau_0^{1/2} \cdot \frac{\delta_N}{\sqrt{N}} \right)$$

où C_3 est une constante indépendante de p et N . En combinant maintenant les inégalités (14) et (15), on montre par un mécanisme similaire à celui de Cràmer (1946, p.430-431) que

$$\|\Phi_{p+1} - \Phi_p\| \leq C_2 [(2C_2 + \tau_0^{1/2})C_3]^p \left(\frac{\delta_N}{\sqrt{N}}\right)^{p+1}, \quad p \geq 0$$

avec une probabilité plus grande que $1 - \delta_N^{-2}$ et $\delta_N \rightarrow \infty$ avec N tel que $\delta_N^{+2} N^{-1/2} \rightarrow 0$.

Références

N'GUESSAN A., LANGRAND C.(1993) "Sur la distribution asymptotique de certaines statistiques utilisées dans le domaine de la sécurité routière" C.R. Acad. Sci. Paris t.317, série I, n° 4, p.401-404.

N'GUESSAN A.(1993) "Contribution à l'analyse statistique d'une mesure de sécurité routière" Thèse de doctorat - Université de Lille 1.

CRAMER H. (1946) "Mathematical methods of statistics" Princeton University Press

Bibliographie

- [1] ABBESS C., JARRET D., WRIGHT C.C , *Accidents and black spot : estimating the effectiveness of remedial treatments, with special reference to the "regression-to- mean" effect.*. Traffic Eng. Control. 22(10) p.535-542, 1981.
- [2] ABDEL-ATY M.A., RADWAN A.E , *Modeling traffic accident occurrence and involvement*, Accid. Anal. and Prev. 32, 633-642, 2000.
- [3] AITCHISON J., SILVEY S.D, *Maximum Likelihood Estimation of parameters subject to restraints* Annals of mathematical statistics, 29,813-829, 1958.
- [4] ALLAIN E., BRENAC T., *The application of generalized linear models to the investigation of accident frequencies at road sites. The Poisson model and its extensions* Recherche Transports Sécurité n°72, 3-18, 2001.
- [5] ARBOUS A.G., KERRICH J.E, *Accidents Statistics and the concept of accident-proneness*, Biometrics, vol.7 p.340-432, 1951.
- [6] BARRY S., GINPIL S, O'NEILL, T.J, *The effectiveness of air bags*, Accid. Anal. and Prev. 31, 781-787, 1999.
- [7] BEENSTOCK M., GAFNI D., GOLDIN E, *The effect of traffic policing on road safety in Israel*, Accid. Anal. and Prev. 33, 73-80, 2001
- [8] BHATTACHARYYA M.N., LAYTON A.P., *Effectiveness of seat belt legislation on the Queensland road toll : an Australian case study in intervention analysis.* J.A.S.A., n° 74, p.596-603, 1979.
- [9] BISHOP Y.M.M., FEINBERG S.E., HOLLAND P.W., *Discrete Multivariate Analysis. Theory and Practice.* The M.I.T Press 9^e Ed., 1988.
- [10] BOX G.E.P., JENKINS G.M., *Time series analysis : Forecasting and control*, San-Francisco : Holden Day, 1976.
- [11] BOX G.E.P., TIAO G.C., *Intervention analysis with applications to economic and environmental problems.* J.A.S.A. 70 p.70-79, 1975.
- [12] CHASE G.R., *On the Chi-square Test when the Parameters are estimated Independently of the Sample.* J.A.S.A., vol.67, n°339. p.609-611, 1972.
- [13] CHERNOFF H., *Large sample theory : parametric case.* Ann. Math. Statist. 27, 1-27, 1956.
- [14] CHERNOFF H., LEHMANN E.F., *The use of Maximum likelihood estimates in χ^2 -tests for goodness of fit.* Ann. Math. Statistics. 25, p.579-586, 1954.
- [15] CHINN S., HELLER R.F., *Some further results concerning regression to the mean.* American Journal of Epidemiology, vol.114 n°6 p.902-905, 1981.
- [16] COUNCIL F.M., REINFURT D.W., CAMPBELL B.J., KOEDIGER F.L., CAROLL C.L., DUTT A.K., DUNHAM J.R., *Accident Research manual*, University of North Carolina, Chapell Hill. 1980.
- [17] CRAMER H., *Mathematical Methods of Statistics*, Princeton University Press, 1946.
- [18] CRESSIE N., READ T.R.C., *Multinomial Goodness-of-fit Tests*, J.R.Statist. Soc. B, 46, No. 3, p.440-464, 1984.
- [19] CRESSIE N., READ T.R.C., *Pearson's χ^2 and the Loglikelihood Ratio Statistic G^2 : A comparative Review.* Inter. Statistical Review, No 57, 1, p.19-43, 1989.
- [20] CROWDER M., *On the constrained maximum likelihood estimation with non i.i.d. observations* Ann. Inst.

- Statist. Math. 36, A, 239-249, 1984.
- [21] DALE J.R., *Asymptotic Normality of Goodness-of-fit-Statistics for Sparse Product Multinomials*. J.R.Statist. Soc. B, 48,n°1, p.48-59, 1986.
- [22] DANIELSON S., *Estimation of the effects of countermeasures on differents types of accidents in the presence of regression effects*, Accid. Anal. and Prev. 20, n°4 , 289-298, 1988.
- [23] DANIELSON S., *A comparison of two methods for estimating of the effect of a countermeasure in the presence of regression effects*, Accid. Anal. and Prev. 18, n°1 , 13-23, 1986.
- [24] DAVIS C.E., *The effect of regression to the mean in epidemiologic and clinical studies*. American Journal of Epidemiology, vol.104 n°5 p.493-498, 1976.
- [25] DEMPSTER A. P., LAIRD N. M., RUBIN D. B., *Maximum likelihood from incomplete data via the EM algorithm*, J. Roy. Stat. Soc. B, 39, 1-38, 1977.
- [26] DON F.J.H., *The Use of generalized inverses in restricted Maximum Likelihood*, Linear Algebra and its Applications, 70,225-240, 1985.
- [27] DROESBEKE J.J., FICHET B., TASSI PH., *Séries chronologiques : théorie et pratique des modèles ARIMA.*, Economica, 1989.
- [28] FAN Y., *Schur complements and its applications to symmetric nonnegative and Z-matrices*, Linear Algebra Appl., 353, 289-307, 2002.
- [29] FEINBERG S.E., *The Use of Chi-squared Statistics for Categorical Data problems*. J.R.Statist. Soc. B, 41,n°1, p.54-64, 1979.
- [30] FLEISCHER G.A., *Contingency table analysis for road safety studies*. NATO Advanced study institute, series E, n°42, 1981.
- [31] FLETCHER R., *Practical methods of optimization : (vol.2) constrained optimization*, John W. Chichester, 1981.
- [32] GARWOOD F., NEWBY R.F., *Utilisation du test du χ^2 pour la comparaison des fréquences d'accidents*, Symposium sur l'utilisation des méthodes statistiques dans l'analyse des accidents de la route, Paris, 1970.
- [33] GERSTEIN L.J., *A new algorithm for computing the rank of a matrix*. The American Math. Monthly, vol 95 p.950-952, 1988.
- [34] GOURIÉROUX C., MONFORT A., *Séries chronologiques et modèles dynamiques*. Economica, 1990.
- [35] GOURIÉROUX C., MONFORT A., *Statistique et modèles économétriques*, Economica, 1989.
- [36] GREENWOOD P. E., NICULIN M. S., *A guide to Chi-square testing*, John Wiley, New York, 1996.
- [37] GUSTAVSSON J., SVENSSON A., *A poisson regression model applied to class of road accidents with small frequencies*. Scand. Journal of Statist. Vol.3 p.49-60, 1976.
- [38] HABER M., BROWN M.B., *Maximum Likelihood methods for log-linear models when expected frequencies are subject to linear constraints*, Journal of American Society Association, vol. 81, n°394, 477-482, 1986.
- [39] HARVILLE D.A., *Maximum Likelihood Approaches to Variance Component Estimation and to related Problems*, Journal of American Society Association, vol. 72, n°358, 320-340, 1977.
- [40] HAUER E., *Selection for treatment as a source of bias in before-after studies*, Traff.Eng.Control, 21(8/9), p.419-421, 1980.
- [41] HAUER E., *Bias-by-selection : overestimation of the effectiveness of safety countermeasures caused by the process of selection for treatment*. Accid. Anal and Prev. Vol.12, n° 26, p.113-117, 1980.
- [42] HAUER E., *Comparison groups in road safety studies : An analysis*. Accid.Anal. and Prev. Vol.23 n° 6, p.609-622, 1991.
- [43] HAUER E., *Observational before-after Studies in Road Safety Estimating the effect of highway and traffic engineering measures on road safety*, Pergamon Edition, 1997.
- [44] HAUER E., *Comparison Groups in road safety studies : An analysis*, Accid. Anal. and Prev. 23, n°26, 609-622, 1991.
- [45] HAUER E., *Overdispersion in modelling accidents on road sections and in Empirical Bayes estimation*, Accid. Anal. and Prev. 33, 799-808, 2001.
- [46] HAUER E., *A note on three estimators of safety effect*. Traffic Eng. and Control. p.388-393, 1992.
- [47] HAUER E., NG J. C. N., PAPAIOANNOU P., *Prediction in road safety studies : An empirical inquiry*,

- Accid. Anal. and Prev. 23, n°6, 595-607, 1991.
- [48] HAUER E., NG J.C.C., PAPALOANNOU P., *Prediction in road safety studies : an empirical inquiry*. Accid. Anal. and Prev. Vol.23 n° 6, p.595-607, 1991.
- [49] HARVEY A.C., *Trends and cycles in macroeconomic time series*. J. Bus. Econ. Statist. n° 3, p.216-227, 1985.
- [50] HARVEY A.C., DURBIN J., *The effects of seat-belt legislation on British road casualties : a case study in structural time series modelling*. J. R. Statist. Soc. A, 149 part 3, p.187-227, 1986.
- [51] HELFENSTEIN U., *When did a reduced speed limit shows an effect ? Explonatory identification of an intervention time*. Accid. Anal. and Prev. Vol.22 n° 1, p.79-87, 1990.
- [52] HOSMER D.W., LEMESHOW S., *Applied logistic regression*, J.W. and Sons, 2nd Ed. 2000.
- [53] HUNTER D. R., *MM algorithms for generalized Bradley-Terry models*, Annals Stat., 2004.
- [54] HUNTER D. R., LANGE K., *Rejoinder to discussion of "Optimization transfer using surrogate objective functions"*, J. Comput. Graphical Stat. 9, 52-59, 2000.
- [55] JAMES K.E., *Regression toward the mean in uncontrolled clinical studies*. Biometrics, 29 p.121-130, 1973.
- [56] LANGE K., HUNTER D. R., YANG I., *Optimization transfer using surrogate objective functions (with discussion)*, J. Comput. Graphical Stat. 9, 1-20, 2000.
- [57] LASSARRE S., *Three statistical treatments of contingency tables in road safety studies*. NATO Advanced study institute, serie E n° 42, 1981.
- [58] LASSARRE S., TAN S.H., *Evaluation of safety measures on the frequency and gravity of traffic accidents in France by means of intervention analysis*, In Anderson, O.D. ed. Times series analysis, theory and practice, 1 p.297-306, 1982.
- [59] LASSARRE S., *Panorama des méthodes associées à l'évaluation de type avant-après*. Colloque international sur l'évaluation des mesures locales de sécurité routière, Paris, 1985.
- [60] LASSARRE S., *The introduction of the variables traffic-volume, speed and belt-warning into a predictive model of severity of accidents*. Accid. Anal. and Prev. Vol.18, n°2 p.129-134, 1986.
- [61] LASSARRE S., *A propos des tests statistiques, sur variables poissonniennes, utilisées dans le domaine de la sécurité routière*, Revue de Statistique Appliquée XXV, n°3, 1977.
- [62] LASSARRE S., *Analysis of progress in road safety in ten European countries*, Accident Analysis and Prevention 33 (2001) 743-751, 2001.
- [63] LEE S.Y., JENRICH R.I., *A study of algorithms for covariance structure analysis with specific comparisons using factors analysis*. Psychometrika, Vol. 44, n° 1, 1979.
- [64] LEE S.Y., *Constrained estimation in covariance structure analysis*. Biometrika, 66, 3, p. 539-45, 1979.
- [65] MAGNUS J.R., *Linear Structures*, Charles Griffin and Compagny LTD, London, 1988.
- [66] MATTHEWS G.B., CROWTHER N.A.S., *A maximum likelihood estimation procedure when modelling in terms of constraints*, South African Statistics Journal, n°29, 29-50, 1995.
- [67] MCLACHLAN G. J., KRISHNAN T., *The EM Algorithm and Extensions*, Wiley, New York, 1997.
- [68] METHA C.R., PATEL N.R., *Exact logistic regression : theory and examples*, Statistics in Medecine, vol. 14, 2143-2160, 1995.
- [69] MOLINARI L., *Distribution of the Chi-squared test in nonstandard situations*. Biometrika, No 64,1, p.115-121, 1977.
- [70] MONFORT A., *Cours de statistique mathématique*. Economica, 1982.
- [71] MOORE D.S., SPRUILL M.C., *Unified Large-Sample Theory of general Chi- squared Statistics for tests of fit*, Annals of Math. Statistics. 3, p.599-616. 1975.
- [72] MOORE D.S., *Generalized Inverses, Wald's Method, and the Construction of Chi-squared Tests of Fit*, J.A.S.A. Vol.72, n°357, p.131-137, 1977.
- [73] MORALES D., PARDO L., VAJDA I., *Some new Statistics of Testing Hypothesis in Parametric Models*, Journal of Multivariate Analysis, Vol. 62, p.137-168, 1997.
- [74] MORTON R., *Analysis of extra-multinomial data derived from extra-Poisson variables conditional on their total*, Biometrika, 78 n° 1, p.1-6, 1991.
- [75] MURTHY V.K., GAFARIAN A.V., *The limiting distributions of some variations of the Chi-square Statistic*,

- The Annals of Math. Statistics. Vol 41, n° 1, p.188-202, 1970.
- [76] NAG, *Numerical Algorithms Groups Limited, 2000*, Ecole Polytechnique Universitaire de Lille, Université de Lille1, 2003.
- [77] NEUENSCHWANDER B. E., FLURY B. D., *A note on Silvey's (1959) theorem*, Statistics & Probability Letter, 36, 307-317, 1997.
- [78] NEYMAN J., *Contribution to the theory of the Chi-square Test*, Proc. Berkeley Symp. Math. Statistics and Probability ,pp. 239-273, 1949.
- [79] N'GUESSAN A., *Contribution à l'analyse statistique d'une mesure de sécurité routière*, Thèse de Doctorat, Université des Sciences et Technologies de LILLE , 1993.
- [80] N'GUESSAN A. , *On a use of Schur complements for an inverse of a constrained accident data information matrix* PUB. IRMA, LILLE Vol. 60, N°VI, 2003.
- [81] N'GUESSAN A. , *Constrained covariance matrix estimation in road accident modelling with Schur complements* C.R. Acad. Sci. Paris, Ser. I 337, 219-222, 2003.
- [82] N'GUESSAN A., *Constrained Estimation of a Road Safety covariance matrix using Shur complements*, Technical Report - Centre for Research on Transportation ,Université de Montréal, N° CRT-2003-12,18 pages, 2003.
- [83] N'GUESSAN A., LANGRAND C., *A Schur complement approach for computing subcovariance matrices arising in road safety measure modelling*, Journal of Computational and Applied Mathematics 177, 331-345, 2005.
- [84] N'GUESSAN A., LANGRAND C. - *Une méthode d'estimation de l'effet moyen d'une mesure de sécurité routière*. XXIV^e Journées de Statistique (1992), Bruxelles.
- [85] N'GUESSAN A., LANGRAND C. *Sur la distribution asymptotique de certaines statistiques utilisées dans le domaine de la sécurité routière*. C. R. Acad. Sci. Paris, Série I, 217, 401-404,1993.
- [86] N'GUESSAN A., LANGRAND C., *A covariance components estimation procedure when modelling a road safety measure in terms of linear constraints* Statistics, Vol. 39, No.4, 303-314, 2005.
- [87] N'GUESSAN A., BELLAVANCE F., *A confidence interval estimation problem using the Schur complement approach, with application* C. R. Math. Rep. Acad. Sci. Canada Vol.27,(3), pp. 84-91, 2005.
- [88] N'GUESSAN A., ESSAI A., LANGRAND C., *Estimation multidimensionnelle des contrôles et de l'effet moyen d'une mesure de sécurité routière* Rev. Statistique Appliquée. XLIX(2), 83-100, 2001.
- [89] N'GUESSAN A., TRUFFIER M., LANGRAND C., BELLAVANCE F., *Application d'une méthode statistique pour évaluer l'impact de certaines modifications des conditions de la route : étude de cas*, PUB. IRMA, LILLE Vol. 66, N°III, 2006.
- [90] NICHOLSON A. J., *The variability of accident counts*, Accid. Anal. and Prev. 17, 47-56, 1985.
- [91] NICHOLSON A. J., *The randomness of accident counts*, Accid. Anal. and Prev. 18, 193-198, 1986.
- [92] NICHOLSON A., WONG Y-D., *Are accidents Poisson distributed? A statistical test*, Accid. Anal. and Prev. 23, n°26, 609-622, 1993.
- [93] OPPE S., *The use of multiplicative models for analysis of road safety data*. Accid. Anal. and Prev. vol. 11, 101-115, 1979.
- [94] ORTEGA J. M., RHEINBOLDT W. C., *Iterative Solutions of Nonlinear Equations in Several Variables*, Academic, New York, 1970.
- [95] OUELLETTE D.V., *Schur complements and statistics*, Linear Algebra and its applications, 36 187-295,1981.
- [96] RAO C.R., *Linear Statistical Inference and Its Applications*, 2nd Ed., John Wiley, New York, 1973.
- [97] RAO C.R.AND YANAI H., *Generalized Inverses of Partitioned Matrices Useful in Statistical Applications* Linear Algebra and its Applications, 70 : 105-115, 1985.
- [98] SEARLE S.R., *Linear models*, John Wiley, New York, 1971.
- [99] SILVEY S.D., *The Langrangian multiplier test*, Ann. Math. Statist. 30, 389-407,1959.
- [100] STYAN G. P. H., *Schur complements and linear statistical models*, In Pukkila & Puntanen eds. pp 37-75, 1985.
- [101] SCHMITTLEIN D.C., *Surprising Inferences from Unsurprising observations : do conditional Expectations really regress to the mean ?*. The American Statistician, vol.43 n° 3 p. 176-183, 1989.

-
- [102] TANABE K., SAGAE M., *An Exact Cholesky Decomposition and the Generalized Inverse of the Variance-Covariance Matrix of the Multinomial Distribution, with Applications*, J. R. Statist. Soc., B, 54, n^o1 pp. 211-219, 1992.
- [103] TANNER J.C., *A problem in the combination of accident frequencies*, Biometrika, 45, 331-342, 1958.
- [104] VAN EEDEN C., *Estimation in Restricted Parameter Spaces - Some history and Some Recent developments*, CWI Quarterly Vol. 9 (1and 2), 69-76, 1996.
- [105] WALD A., *Note on the consistency of the maximum likelihood estimate*, Annals of Mathematical Statistics, 20, 595-601, 1949.
- [106] WIORKOWSKI J.J., HECKARD R.F., *The use of time series analysis and intervention analysis to assess the effects of external factors on traffic indices : a case study of the effects of the speed limit reduction and energy crisis in the state of Texas*. Accid.Anal.and Prev.(1977) vol 19, p.229-247.
- [107] WOLTER K.M., *Introduction to Variance Estimation*, Springer-Verlag, 1985.
- [108] WRIGHT C.C., ABBESS C.R., JARRETT D.F., *Estimating the regression to mean effect associated with road accident black spot treatment a more realistic approach*, Accid.Anal.and Prev. Vol 20,n^o 3, p.199-214, 1988.
- [109] WU C.F., *On the convergence properties of the E.M. algorithm*. Ann. Statist. 11, p.95-103, 1983.
- [110] ZHANG F. ED, *Schur complement and Its applications*, Springer Verlag, 2005.