

50376
1993
217

50376
1993
217

N° d'ordre : 1142

THÈSE

présentée à

L'UNIVERSITÉ DES SCIENCES ET TECHNOLOGIES DE LILLE

pour obtenir

LE TITRE DE DOCTEUR DE L'UNIVERSITÉ

SPÉCIALITÉ : MATHÉMATIQUES APPLIQUÉES



par

N'GUESSAN Assi



CONTRIBUTION A L'ANALYSE STATISTIQUE D'UNE MESURE DE SÉCURITÉ ROUTIÈRE

Soutenue le 24 Juin 1993 devant la Commission d'Examen :

Président : M. ROUSSIGNOL, Université de Lille

Rapporteurs : R. THEODORESCU, Université Laval - Canada

S. LASSARRE, I.N.RE.T.S - Arcueil

Examineurs : M. ROUSSIGNOL, Université de Lille

C. LANGRAND, Université de Lille



Ce travail est le fruit de plusieurs années de recherche et de nombreuses personnes physiques et morales ont contribué de loin ou de près à sa réalisation.

Je tiens à remercier Monsieur le Professeur Michel ROUSSIGNOL qui a bien voulu être le Président du jury de cette thèse, pour l'intérêt qu'il a porté à ce travail et pour ses remarques, notamment sur les aspects appliqués qu'il renferme.

Un grand merci aussi à tout le personnel de l'UFR de Mathématiques et à toute l'équipe du Laboratoire de Probabilités et Statistique qui m'ont donné le confort nécessaire à la réalisation de cette thèse, en particulier à Madame Nelly HANOUNE pour m'avoir initié aux traitements statistiques des données.

J'adresse mes sincères remerciements à Monsieur le Professeur Radu THEODORESCU pour avoir accepté d'être rapporteur de cette thèse. Ses remarques pertinentes 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 a bien voulu être rapporteur de cette recherche, pour la précision de ses remarques ainsi que pour ses idées et conseils concernant les développements futurs à envisager.

Cette thèse doit avant tout d'exister à Monsieur le Professeur Claude LANGRAND, qui a dirigé mes recherches et m'a toujours guidé et soutenu avec une extrême compétence et de grandes qualités, tant humaines que scientifiques. Il fait partie des rares personnes fantastiques qu'il m'a été donné de rencontrer.

Enfin, pour conclure mieux que je ne saurais le faire, je ne peux résister au plaisir de remercier:

ma famille, pour la nourriture, la chaleur, et l'affection qui m'ont permis de grandir et de me structurer;

mes maîtres, qui m'ont transmis les connaissances lentement accumulées par l'humanité depuis qu'elle interroge l'univers;

mes amis qui m'ont entouré d'un amour irremplaçable. Mais comme toute chose a une fin, c'est à moi d'achever l'ouvrage.

INTRODUCTION GENERALE

I. Approche chronologique	6
II. Approche "avant-après"	8

Chapitre 1 : METHODE STATISTIQUE DE TANNER

1.1 Introduction	15
1.2 Modèle d'analyse	16
1.3 Analyse sous l'hypothèse $H_0 : \theta_1 = \theta_2 = \dots = \theta_s = \theta$	18
1.3.1 Estimation de l'effet de la mesure	18
1.3.2 Distribution asymptotique de $\hat{\theta}$ et $Log(\hat{\theta})$	19
1.3.3 Test de l'hypothèse H_0	24
1.3.4 Approximation de la variabilité de l'effet aux différents sites	27
1.4 Analyse en cas d'hétérogénéité de l'effet	29
1.4.1 Estimation de l'effet de la mesure	29
1.4.2 Approximation de la variance asymptotique de $Log(\hat{\theta})$	29
1.4.3 Test de l'hypothèse d'efficacité nulle	32
1.4.4 Une application: transformation de carrefours en ronds-points	33

Chapitre 2 : UNE METHODE D'ANALYSE A L'AIDE D'UN TABLEAU "AVANT-APRES" $2 \times s \times r$

2.1 Introduction	39
2.2 Modèles d'analyse	40
2.2.1 Notations	40
2.2.2 Modèles	41
2.3 Analyse sous l'hypothèse $H_0 : \theta_1 = \theta_2 = \dots = \theta_s = \theta$	42
2.3.1 Résultats préliminaires	42
2.3.2 Estimation de l'effet	44
2.3.3 Approximation de la variance asymptotique de $Log(\hat{\theta})$	46
2.3.4 Test de l'hypothèse $H_0 : \theta_1 = \theta_2 = \dots = \theta_s = \theta$ et approximation de la variabilité de l'effet aux différents sites	49

2.4 Analyse en cas d'hétérogénéité de l'effet	54
2.4.1 Introduction	54
2.4.2 Estimation de l'effet	55
2.4.3 Une méthode d'approximation de la variance asymptotique de $\text{Log}(\hat{\theta})$	55
2.4.4 Approximation du coefficient d'hétérogénéité ϕ	59
2.4.5 Test de l'hypothèse d'efficacité nulle	63

Chapitre 3 : DISTRIBUTION ASYMPTOTIQUE DE $\hat{A}_{s,r}^2$

3.1 Introduction	67
3.2 Notations et lemmes préliminaires	68
3.2.1 Notations	68
3.2.2 Lemmes préliminaires	69
3.3 Distributions asymptotiques de $\hat{A}_{s,r}^2$ et $\tilde{A}_{s,r}^2$	71
3.3.1 Distribution asymptotique de $\tilde{A}_{s,r}^2$	71
3.3.2 Distribution asymptotique de $\hat{A}_{s,r}^2$	74
3.4 Calcul du rang de la matrice D	77
3.5 Utilisations pratiques de $\hat{A}_{s,r}^2$	79
3.5.1 Test de l'hypothèse H_0	79
3.5.2 Approximation de la variabilité de l'effet aux différents sites	79

Chapitre 4 : PROCEDURES NUMERIQUES

4.1 Structure théorique	83
4.1.1 Approche générale	83
4.1.2 Approche par la bibliothèque NAG	84
4.2 Structure pratique	85
4.2.1 Programme principal	85
4.2.2 Sous-programmes	86

4.3 Applications	89
4.3.1 Description générale des procédures numériques	89
4.3.2 Estimation de l'effet de la mesure sur données simulées	83
4.3.3 Comparaison des modèles	91

ANNEXES :

Annexe 2	99
A2.1 Approximation de la distribution asymptotique de $\hat{\theta}$ et $Log(\hat{\theta})$ sous $H_0 : \theta_1 = \theta_2 = \dots = \theta_s = \theta$	99
A2.2 Estimation des paramètres et information de Fisher dans le cadre du modèle 2	101
Annexe 3	103
A3.1 Méthode itérative de calcul du rang de B_o	103
A3.2 Calcul explicite de la matrice $\tilde{\Omega}$	109
Annexe 4	113
programmes FORTRAN	
Références bibliographiques	137

INTRODUCTION GENERALE

Dans l'analyse d'une mesure de sécurité routière, les méthodes statistiques représentent l'ultime étape du processus. Ces méthodes n'ont pas un statut méthodologique bien établi [voir OCDE (1970,1981)]. On peut subdiviser ces méthodes en deux grandes classes: une première dite "approche avant-après" basée sur l'analyse de tableaux de contingence [voir Fleischer (1981)] ou issus de plans quasi-expérimentaux [voir Lassarre (1981,1985); Hoc (1983)]; et une deuxième dite "approche chronologique ou longitudinale" basée sur les techniques d'analyse d'interventions [voir Box et Tiao (1975); Droesbeke et al. 1989]. Le choix de la méthode appropriée dépend fortement de la nature de l'étude menée et de la variable aléatoire (critère) utilisée pour apprécier l'efficacité ou l'inefficacité de la mesure. Cette variable aléatoire (v.a.) doit être adaptée à la nature de la mesure et des sites soumis à l'intervention. Quelle que soit la méthode et quelle que soit la v.a. utilisée, l'évaluation statistique d'une mesure de sécurité routière repose sur la prise en compte

- d'un ensemble de valeurs observées de la v.a. dans le passé (ou avant la prise de la mesure), à des dates généralement fixées et équidistantes
- d'un ensemble de valeurs observées de la v.a. après la mesure
- d'un ensemble d'hypothèses de travail permettant de justifier l'emploi de ces méthodes et d'établir l'efficacité ou l'inefficacité de la mesure.

La technique statistique la plus courante consiste à comparer un échantillon de nombres d'accident avant et après une intervention au moyen de tests statistiques sous diverses hypothèses et avec différents niveaux de signification. Cette comparaison tient souvent compte d'un échantillon témoin ou de contrôle. Mais pour de nombreuses raisons les sites traités (expérimentaux) peuvent être par nature particuliers, si bien qu'il est parfois difficile de trouver des sites non traités comparables; remarquons aussi que dans le cas des mesures imposées par législation ou réglementation tous les sites sont traités simultanément. Cette absence de témoin signifie que, pour juger si le remède (on considère parfois que l'insécurité routière constitue une maladie qu'il convient de guérir) a apporté un changement dans la

période après, il est nécessaire de supposer que les valeurs "après" auraient été du même type que celles de la période avant, si le traitement n'était pas intervenu. Cette hypothèse de base, peut être modifiée de deux manières:

- a) si l'on dispose d'une longue série de données retrospectives sur les v.a. concernées, on pourra s'en servir, pour élaborer une tendance susceptible d'être extrapolée par la suite pour établir une meilleure estimation de la valeur attendue;
- b) si l'on dispose également de données "avant" concernant d'autres v.a. (non touchées par le traitement) et s'il est possible de déterminer une relation entre celles-ci et les v.a. étudiées, on pourra alors procéder à des observations concernant ces autres variables après la mise en oeuvre du traitement afin d'estimer également de manière plus fine les valeurs attendues de ces variables. Les techniques statistiques employées dans ce contexte sont basées sur la modélisation des séries chronologiques.

Nous présentons de façon succincte un panorama des deux méthodes et utilisons dans ce travail la v.a. "nombre d'accidents" sans préjuger de l'utilité d'autres v.a. [voir Council et al. 1977; OCDE 1981].

I. Approche chronologique.

Soient $X_t, X_{t-1}, X_{t-2}, \dots$ les v.a. donnant l'effectif du nombre des accidents observés au temps (période) $t, t-1, t-2, \dots$ (on suppose pour simplifier que les durées entre deux observations consécutives sont égales) et B l'opérateur retard tel que

$$B X_t = X_{t-1}.$$

Box et Tiao (1975) ont proposé l'emploi de certaines formes fonctionnelles pour modéliser des interventions ou des ruptures survenues dans un processus stochastique. Leur idée est de présenter ces ruptures comme conséquences de chocs exogènes sur la série; ces chocs peuvent avoir selon le cas un effet permanent ou transitoire. Ils ont étudié en détail un modèle du type:

$$X_t = \sum_{j=1}^m \frac{\omega_j(B)}{\delta_j(B)} Y_{jt} + \varepsilon_t \quad t = 1, 2, \dots, T$$

où ω_j et δ_j ($j = 1, 2, \dots, m$) sont des polynômes, ε_t est un processus ARIMA [voir Box et Jenkins (1976); Gouriéroux et Monfort (1990); Dreesbeke et al. (1989); David et Michaud (1989)]. La v.a. Y_{jt} est soit une variable "de saut" définie par

$$S_{T_i}(t) = \mathbf{1}_{[T_i, +\infty)}(t)$$

(elle est alors destinée à rendre compte de l'influence d'un phénomène commençant à la date T_i , par exemple un changement de réglementation) , soit une variable "*impulsion*" définie par :

$$I_{T_i}(t) = \begin{cases} 0 & \text{si } t \neq T_i \\ 1 & \text{si } t = T_i \end{cases}$$

(elle est alors destinée à rendre compte de l'influence sur X_t d'un phénomène ayant eu lieu à la date T_i uniquement, par exemple une grève), soit une variable "*palier*" définie par:

$$P_{it} = 1_{[T_{1i}, T_{2i}]}(t)$$

(elle est alors destinée à rendre compte d'un phénomène transitoire ayant eu lieu entre les dates T_{1i} et T_{2i} , par exemple une modification provisoire de réglementation). Diverses applications de ces modèles ont été proposés dans le domaine de la sécurité routière [voir Wiorkowski et Heckard (1977); Bhattacharyya et Layton (1979); Lassarre et Tan (1982); Helfenstein (1990)]. Par exemple Lassarre et Tan analysent l'efficacité de trois mesures de sécurité routière (limitation de vitesse, ceinture de sécurité et loi sur l'alcool). Le modèle qu'ils proposent est

$$X_t = \sum_{j=1}^3 \omega_j Y_{jt} + \varepsilon_t ; \quad t = 1, 2, \dots, T$$

où Y_{1t}, Y_{2t} et Y_{3t} sont des v.a. "*palier*" et ε_t , un processus ARIMA dont la forme dépend de la série étudiée (accidents ou tués). Ils proposent alors une définition du gain relatif de chaque mesure et le modèle

$$X_t = \frac{\omega B}{1 - \delta B} I_t + \varepsilon_t$$

pour la loi sur l'alcool (juillet 1978), I_t étant une variable "*impulsion*". Avec ce dernier modèle, l'effet immédiat est plus important en rase campagne qu'en zone urbaine .

Harvey et Durbin(1986) ont proposé une deuxième classe de modèles de séries chronologiques applicables à l'analyse d'une intervention survenue dans un processus. Cette classe qu'ils ont baptisé " modèles structurels" a été appliquée à l'efficacité de la loi sur la ceinture de sécurité prise en Grande Bretagne en 1983. La forme générale de leur modèle est:

$$X_t = \mu_t + \gamma_t + \sum_{j=1}^m \alpha_j Y_{jt} + \theta W_t + \varepsilon_t \quad t = 1, 2, \dots, T$$

où X_t est l'observation au temps t ; μ_t , γ_t et ε_t représentent la tendance, la saisonnalité et l'irrégularité (bruit) de la série; Y_{jt} est la j^{eme} composante de la variable explicative au temps t (cette dernière variable peut être le trafic, la vitesse ... [voir Lassarre 1986]; α_j est un coefficient à estimer; W_t est la variable d'intervention et θ le coefficient représentant l'effet de la mesure. Ils ont proposé une forme trigonométrique à γ_t et pour μ_t la forme suivante:

$$\begin{cases} \mu_t &= \mu_{t-1} + \beta_t + \eta_t \\ \beta_t &= \beta_{t-1} + \zeta_t \end{cases}$$

où ζ_t est un bruit indépendant de ε_t et η_t . La forme explicite de W_t dépend de l'effet que l'intervention est supposée apporter à la série étudiée. Dans le cas simple où on suppose que l'effet de l'intervention à l'instant T_0 est immédiat et égal à une quantité θ , la variable W_t prend la forme d'une variable "de saut". On peut aussi imaginer qu'il y a un effet instantané au temps $t = T_0$ et que cet effet décroît ou croît par la suite ou que l'effet augmente petit à petit suivant l'intervention. Il existe différentes modifications ou extensions de ces modèles [voir Harvey 1985] et elles dépendent toutes de la forme de la tendance ou de la saisonnalité spécifiée et du bruit. Avec ce modèle, Harvey et Durbin ont pu estimer l'effet de la mesure en terme de nombre de tués, de tués et blessés graves et fournir des intervalles de confiance pour chaque catégorie d'utilisateurs de la route. Une analyse plus fine des accidents corporels permet de distinguer les catégories d'utilisateurs directement affectées par la loi (conducteurs, passagers avant) et celles qui ne le sont pas (passagers arrière, piétons et cyclistes).

II. Approche "avant-après" .

Afin d'analyser l'effet d'une mesure de sécurité routière en un site particulier, on compte les accidents avant cette mesure (période: t) et après celle-ci (période: $t + 1$) et il faut alors décider si la modification du nombre des accidents (passant par exemple de x_t à x_{t+1}) est plus importante que celle qu'on aurait observée si la mesure n'était pas efficace. Pour cela, on prend en compte des facteurs extérieurs (variations du nombre des accidents sur un site équivalent non soumis à la mesure de sécurité, nombre d'utilisateurs, ...). L'interprétation des observations d'accidents doit tenir compte de l'évolution du trafic et des conditions météorologiques; si en étudiant d'autres données, on peut s'attendre à ce que le nombre d'accidents augmente d'un facteur c (le rapport de contrôle, coefficient

d'évolution), cela même si la mesure n'est pas efficace, la statistique adéquate [voir Garwood et Newby ,1970] à utiliser est:

$$A^2 = \frac{(x_{t+1} - cx_t)^2}{(x_{t+1} + x_t) c}.$$

Ils comparent cette statistique à la valeur que dépasse une v.a. de loi de χ_1^2 (chi-deux à 1 degré de liberté) avec une probabilité α ($\alpha = 0,05$). Si la valeur de la statistique A^2 dépasse 3,841 ils concluent que l'hypothèse nulle (la mesure n'est pas efficace) est à rejeter et que probablement il y a efficacité réelle. Sinon, une telle conclusion n'est pas du tout évidente, bien qu'il puisse arriver que des observations supplémentaires relatives au même site puissent révéler une amélioration réelle. Comment donc estimer l'efficacité?

Si on note $\hat{X}_{t+1}$ la v.a. donnant l'effectif d'accidents "attendu" après la mesure dans le site traité, l'efficacité de la mesure de sécurité routière est estimée par

$$1 - \frac{E(X_{t+1})}{E(\hat{X}_{t+1})},$$

où $E(.)$ désigne l'espérance mathématique par rapport à une certaine distribution de probabilité. En pratique, on estime le rapport

$$\theta = \frac{E(X_{t+1})}{E(\hat{X}_{t+1})},$$

(indice d'efficacité que nous appellerons abusivement effet de la mesure) en supposant que les données d'accidents sont rangées dans des tableaux avant-après soit sous forme de tableaux de contingence [voir Fleischer, 1981] soit sous forme de strates [voir Jenicek et Cléroux p.180] et en faisant une hypothèse de distribution sur les v.a. X_t et X_{t+1} ; souvent on suppose que

$$E(\hat{X}_{t+1}) = \theta c \hat{E}(X_t)$$

c'est-à-dire que la mesure de sécurité routière a un effet multiplicateur sur la moyenne des accidents (on conçoit aisément dans le domaine de la sécurité routière [voir Oppe (1979)] qu'une mesure tend à faire diminuer la v.a. d'un certain pourcentage et non pas d'un certain incrément). Les caractéristiques de la distribution sur les v.a. X_t et X_{t+1} diffèrent d'une étude à une autre et doivent tenir compte de la structure temporelle (avant-après) et ponctuelle du nombre des accidents.

Elles peuvent être décrites soit par une distribution de Poisson [voir Gustavsson et Svensson(1976); Danielson (1986,1988); Lassarre (1977,1981)] soit par une distribution binomiale ou multinomiale [voir Tanner(1958); N'Guessan et Langrand 1992] soit par une distribution binomiale négative [voir Boyer et al.1990; Fridstrom 1991] soit par des modèles Log-linéaires [voir Brühning et al.1985].

Il est également conseillé dans certaines études d'utiliser la procédure bayésienne empirique [voir Gouriéroux et Monfort; Arbous et Kerrich (1951); Abbess et al. (1981); Hauer (1980)] . L'avantage de la méthode bayésienne empirique est d'éliminer le biais introduit par l'effet de regression vers la moyenne largement discuté dans la littérature [voir James (1971); Davis (1976); Hauer (1980); Chinn et Heller (1981); Wright et al. (1988); Schimittlen (1989)]. Lorsque l'on tient à analyser l'efficacité d'une mesure de sécurité routière sur plusieurs sites à l'aide de tableaux avant-après, il faut considérer le modèle de Tanner (chapitre 1) et ceux de N'Guessan et Langrand (1993).

Quelle direction de recherche faut-il choisir parmi ce flot de modèles et de méthodes d'analyse? Notre préférence pour l'approche de Tanner est due à trois grandes raisons. La première est qu'en pratique les données d'accidents issues d'une prise de mesure de sécurité routière sont généralement rangées dans des tableaux avant-après. La deuxième est que, d'après Lassarre (1985), *"seuls le modèle multi-sites de Tanner et le modèle Log-linéaire en intégrant le facteur aléatoire site fournissent un cadre complet pour inférer à partir de l'efficacité d'une mesure de sécurité estimée sur un échantillon de sites à la population de sites"* . La troisième est que du point de vue statistique, l'approche de Tanner reste la plus complète car elle propose un test sur la significativité de l'efficacité dans le cas d'homogénéité et même dans le cas d'hétérogénéité.

Nous nous sommes donc intéressés au modèle et à l'approche de Tanner au chapitre 1 de ce travail. Nous réorganisons et introduisons certaines conditions permettant de comprendre au mieux ses résultats. En particulier, nous déterminons la distribution asymptotique exacte de la statistique qu'il utilise pour éprouver l'hypothèse de similarité de l'effet aux différents sites.

Tout en gardant la philosophie d'approche de Tanner, nous généralisons au chapitre 2 ses résultats au cas où l'on est en présence d'un tableau "avant-après" comportant des sites et différents types d'accidents. Deux modèles généralisant celui de Tanner sont proposés ainsi qu'une méthode pour estimer les paramètres

de ces modèles. Nous suggérons ensuite une statistique pour tester l'homogénéité de l'effet aux différents sites. La distribution asymptotique de cette statistique [voir N'Guessan et Langrand (1993)] est effectuée au chapitre 3. Nous généralisons dans ce chapitre la statistique de Tanner. Notre travail se termine au chapitre 4 par des procédures numériques permettant de simuler des données d'accidents et de mettre en oeuvre la méthode proposée.

CHAPITRE 1:
MÉTHODE STATISTIQUE
DE TANNER

1.1.- INTRODUCTION

Tanner (1958) a mis au point un modèle et une méthode statistiques pour estimer et analyser l'efficacité d'une mesure de sécurité routière. Pour s sites supposés représentatifs de l'ensemble des sites, il dispose du nombre des accidents survenus avant et après la mise en place de la mesure. A chaque site est associée une zone de contrôle dans laquelle la mesure n'est pas directement appliquée et qui lui permet de prendre en compte l'évolution du nombre des accidents avant et après la mesure indépendamment de celle-ci ; il définit ainsi un rapport de contrôle pour chaque site à l'aide du quotient nombre des accidents "après" sur nombre des accidents "avant" dans la zone de contrôle.

Comme il est difficile d'estimer l'amélioration moyenne si ces rapports de contrôle ne sont pas les mêmes en tous les sites, Tanner propose d'analyser l'efficacité de la mesure de sécurité routière sous l'hypothèse de similarité en tous les sites de l'effet et suggère une statistique pour tester cette similarité. S'il n'y a pas de similarité, il faut examiner soigneusement la significativité du changement moyen. Il propose une méthode pour estimer et tester l'efficacité de la mesure imposée et l'applique à la transformation de carrefours en ronds-points.

Ce chapitre présente la méthode statistique de Tanner. Nous exposons à la section 1.2 le modèle de Tanner et son principe de modélisation. Nous présentons à la section 1.3 son analyse sous l'hypothèse de similarité de l'effet de la mesure de sécurité routière. La distribution asymptotique de l'estimateur de l'effet de la mesure et le détail des calculs sont donnés à la sous-section 1.3.2. Nous étudions à la sous-section 1.3.3 la distribution asymptotique de la statistique proposée par Tanner pour éprouver la similarité de l'effet. L'approximation de la variabilité de l'effet aux différents sites est effectuée à la sous-section 1.3.4. Sa méthode d'analyse lorsque la similarité n'est pas acquise est donnée à la section 1.4. La sous-section 1.4.2 discute en particulier de sa méthode d'approximation de la variance asymptotique de l'estimateur de l'effet. L'application fournie par Tanner fait l'objet de la sous-section 1.4.4.

1.2.- MODÈLE D'ANALYSE

On note

s le nombre de sites expérimentaux où la mesure est appliquée ;

X_{1k} [resp. X_{2k}] la variable aléatoire (v.a.) donnant l'effectif d'accidents sur le site k avant [resp. après] la mesure et dont une valeur observée est x_{1k} [resp. x_{2k}] ($k = 1, 2, \dots, s$) ;

$x_{.k} = x_{1k} + x_{2k}$ l'effectif total d'accidents survenus au site k au cours des deux périodes ($x_{.k}$ est déterminé) ;

$N = \sum_{k=1}^s x_{.k}$ le nombre total d'accidents enregistré sur l'ensemble des sites ;

c_k ($0 < c_k$) le rapport de contrôle spécifique au site k (c_k est supposé non aléatoire et parfaitement bien déterminé par les effectifs d'accidents observés, avant et après, dans le site de contrôle associé au site k) ;

$\theta_k^* = x_{2k}/c_k x_{1k}$ mesure l'effet apparent du changement au site expérimental k (c'est-à-dire le rapport du nombre d'accidents après au nombre d'accidents escompté au site k si la mesure n'avait pas d'effet et si le site de contrôle se comportait comme la zone de contrôle).

On suppose $k = 1, 2, \dots, s$ fixé et on note X_{2k}^* la v.a. donnant l'effectif d'accidents sur le site k après la mesure si celle-ci n'avait pas d'effet sur le nombre d'accidents et si la zone de contrôle se comportait comme le site k . Selon la modélisation de Tanner (1958), on a :

$$E(X_{2k}^*) = \theta_k^* c_k \hat{E}(X_{1k}) \quad (1.2.1.)$$

$k = 1, 2, \dots, s$. Une valeur observée de X_{2k}^* est donnée par :

$$x_{2k}^* = \theta_k^* c_k x_{1k} \quad (1.2.2.)$$

$k = 1, 2, \dots, s$. Le nombre total d'accidents espéré est $x_k^* = (1 + \theta_k^* c_k) x_{1k}$. Par conséquent la proportion d'accidents dans la période avant [resp. après] est donné par :

$$q_{1k}^* = \frac{1}{1 + \theta_k^* c_k}, \text{ [resp. } q_{2k}^* = \frac{\theta_k^* c_k}{1 + \theta_k^* c_k}] \quad (1.2.3.)$$

1.3.- ANALYSE SOUS L'HYPOTHÈSE $H_0: \theta_1 = \theta_2 = \dots \theta_s = \theta$

Tanner suggère d'utiliser cette méthode d'estimation et le test de signification associée seulement lorsque des considérations a priori sur les données du problème permettent de supposer qu'il n'y a aucune variation de l'effet de la mesure d'un site à un autre.

1.3.1.- Estimation de l'effet

Sous l'hypothèse $H_0: \theta_1 = \theta_2 = \dots = \theta_s = \theta$, la vraisemblance relativement à la suite de $\vec{v}.a.$ $(X_{11}, X_{21}), \dots, (X_{1s}, X_{2s})$ s'écrit :

$$\mathcal{L}(\theta) = \prod_{k=1}^s \left(\frac{x_{\cdot k}!}{x_{1k}! x_{2k}!} \right) \frac{(\theta c_k)^{x_{2k}}}{(1 + \theta c_k)^{x_{\cdot k}}} \quad (1.3.1.)$$

La Log-vraisemblance est donnée, à une constante additive près par :

$$L(\theta) = \sum_{k=1}^s \{x_{2k} \text{Log } \theta + x_{2k} \text{Log } c_k - x_{\cdot k} \text{Log}(1 + \theta c_k)\} \quad (1.3.2.)$$

On montre alors que l'estimateur du maximum de vraisemblance (EMV) $\hat{\theta}$ de θ satisfait à l'équation non linéaire

$$\begin{cases} \sum_{k=1}^s \left[\frac{x_{2k} - \hat{\theta} c_k x_{1k}}{1 + \hat{\theta} c_k} \right] = 0 \\ \hat{\theta} > 0 \end{cases} \quad (1.3.3.)$$

L'équation (1.3.3.) ne peut se résoudre que numériquement par approximation successives et on ne peut donc pas obtenir une forme explicite de $\hat{\theta}$ sauf dans le cas particulier où $c_1 = c_2 = \dots = c_s$. Pour faire face à ce problème numérique, Hauer (1992) a proposé deux estimateurs de l'effet de la mesure qui sont :

$$\tilde{\theta} = \frac{(\sum_{k=1}^s X_{2k}) / (\sum_{k=1}^s c_k X_{1k})}{1 + \frac{\sum_{k=1}^s c_k^2 X_{1k}}{(\sum_{k=1}^s c_k X_{1k})^2}}, \quad (1.3.4.)$$

$$\bar{\theta} = \exp \left[\frac{\sum_{k=1}^s \left(\frac{X_{2k} X_{1k}}{X_{2k} + X_{1k}} \right) \text{Log } \theta_k^*}{\sum_{k=1}^s \frac{X_{2k} X_{1k}}{X_{2k} + X_{1k}}} \right] \quad (1.3.5.)$$

$k = 1, 2, \dots, s$. Plus généralement, Tanner suppose que (X_{1k}, X_{2k}) , ($k = 1, 2, \dots, s$) est une suite de vecteurs aléatoires ($\vec{v}.a.$) indépendants suivant chacun une distribution multinomiale.

$$\mathcal{M}(x_{.k}; q_{1k}(\theta_k), q_{2k}(\theta_k))$$

$$q_{1k}(\theta) = \frac{1}{1 + \theta_k c_k}, \quad q_{2k}(\theta) = \frac{\theta_k c_k}{1 + \theta_k c_k} \quad (1.2.4)$$

où θ_k ($0 < \theta_k$) est le paramètre effet de la mesure au site expérimental k c'est-à-dire la valeur que prendrait, θ_k^* si les v.a. X_{1k} et X_{2k} prenaient elles-mêmes pour valeur leur espérance mathématique. L'indépendance des $\vec{v}.a.$ (X_{1k}, X_{2k}) est en partie justifiée par le fait que les accidents surviennent indépendamment les uns des autres. Selon certaines modélisations, ([voir par exemple Danielson (1988)]) il est souhaitable de supposer que X_{1k} et X_{2k} sont des v.a. indépendantes suivant une distribution de Poisson de paramètre respectif λ_k et $\theta_k c_k \lambda_k$. Cette optique ne sera pas abordée dans ce travail.

Les difficultés dans l'approche de modélisation de Tanner résident essentiellement dans le choix de la zone de contrôle et dans l'hypothèse "non aléatoire" effectuée sur le coefficient de contrôle c_k . Récemment Hauer (1991) 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 l'effet aléatoire inhérent au comptage des accidents sur ce site ne dégrade pas la précision d'estimation. L'utilisation d'un site de comparaison est beaucoup plus une façon de contrôler le biais dans l'estimation de l'effet de la mesure qu'une façon de réduire l'erreur de prédiction. Dans cette optique, l'utilisation de zone de contrôle introduite par Tanner en 1958 est toujours d'actualité et ce malgré quelques réserves de certains auteurs. Ce choix de zone de contrôle devrait être amélioré par prise en compte des récents résultats de Hauer (1991) et Hauer et al (1991).

Preuve : En réécrivant l'équation de vraisemblance sous la forme

$$\sum_{k=1}^s X_{1k} - \sum_{k=1}^s \frac{x_{.k}}{1 + \hat{\theta}_{c_k}} = 0 \quad (1.3.3.)$$

et en utilisant la définition de la v.a. z , on montre que

$$\sum_{k=1}^s X_{1k} - \sum_{k=1}^s \frac{x_{.k}}{1 + \theta_{c_k}} \left[\frac{1}{1 + \omega_k(e^z - 1)} \right] = 0 \quad (1.3.8.)$$

$\omega_k = \theta_{c_k}/(1 + \theta_{c_k})$. D'après les propriétés asymptotiques des EMV, on a $\forall \varepsilon > 0, \forall \eta > 0, \exists N(\varepsilon, \eta) \in \mathbb{N}^*$ tel que $\forall N \geq N(\varepsilon, \eta)$

$$\text{Proba}\{|z| < \varepsilon\} \geq 1 - \eta.$$

Il s'en suit alors que $\forall \varepsilon > 0, \forall \eta > 0, \exists N(\varepsilon, \eta)$ tel que $\forall N \geq N(\varepsilon, \eta)$

$$\text{Proba}\{|e^z - 1| < \varepsilon\} \geq 1 - \eta.$$

Par conséquent, on peut écrire $\forall N \geq N(\varepsilon, \eta)$

$$\frac{1}{1 + \omega_k(e^z - 1)} = 1 - \omega_k(e^z - 1) + \omega_k^2(e^z - 1)^2 - \omega_k^3(e^z - 1)^3 + \omega_k^4(e^z - 1)^4 + O_p(1).$$

En substituant cette dernière expression dans (1.3.8.) et en utilisant le fait que $\forall N \geq N(\varepsilon, \eta)$ $e^z - 1 = z + \frac{1}{2!}z^2 + \frac{1}{3!}z^3 + \frac{1}{4!}z^4 + O_p(1)$, on a $\forall \varepsilon > 0, \forall \eta > 0, \exists N(\varepsilon, \eta)$ tel que $\forall N \geq N(\varepsilon, \eta)$

$$\xi = a_1 z + \left(\frac{a_1}{2} - a_2\right)z^2 + \left(\frac{a_1}{6} - a_2 + a_3\right)z^3 + \left(\frac{a_1}{24} - \frac{7}{12}a_2 + \frac{3}{2}a_3 - a_4\right)z^4 + O_p(1)$$

Lemme 1.3.2.- Sous l'hypothèse $\theta_1 = \theta_2 = \dots = \theta_s = \theta$, on a

$$E(\xi) = 0, E(\xi^2) = a_1, E(\xi^3) = a_1 - 2a_2, E(\xi^4) = a_1 - 6a_2 + 6a_3 + 3a_1^2$$

Preuve : On réécrit ξ sous la forme $\xi = \sum_{k=1}^s \left[\frac{x_{.k}}{1 + \theta_{c_k}} - X_{1k} \right]$. Puisque $E(X_{1k}) = x_{.k}/(1 + \theta_{c_k})$, il s'en suit alors que

$$\xi = - \sum_{k=1}^s [X_{1k} - E(X_{1k})].$$

et pour lesquels il montre sur des résultats de simulations que $\tilde{\theta}$ et $\bar{\theta}$ sont moins biaisés que $\hat{\theta}$ et d'erreur quadratique moyenne minimale. Il suggère alors d'utiliser $\tilde{\theta}$ ou $\bar{\theta}$ comme estimateur de l'effet. Malgré ce choix pratique et malgré la forme explicite de $\tilde{\theta}$ et $\bar{\theta}$, il est extrêmement difficile de connaître théoriquement leur biais et leur variance asymptotique ($N \rightarrow \infty$). Ce qui n'est pas le cas de $\hat{\theta}$.

En effet, si on dérive $L(\theta)$ deux fois par rapport à $\text{Log } \hat{\theta}$, on montre que la variance asymptotique de $\text{Log } \hat{\theta}$ est donnée par

$$\text{var Log } \hat{\theta} = \frac{1}{\sum_{k=1}^s \frac{\theta c_k x_{.k}}{(1+\theta c_k)^2}} \quad (1.3.6)$$

Si on admet l'hypothèse $H_0: \theta_1 = \theta_2 = \dots = \theta_s = \theta$, alors Tanner préconise pour éprouver l'hypothèse supplémentaire $\theta = 1$ (la mesure n'est pas efficace) d'utiliser la v.a.

$$Z = \frac{\text{Log } \hat{\theta}}{[\text{var Log } \hat{\theta}]^{\frac{1}{2}}} \quad (1.3.7.)$$

qui, sous l'hypothèse $\theta = 1$, soit asymptotiquement une $\mathcal{N}(0; 1)$ d'après les propriétés des EMV.

1.3.2.- Distribution asymptotique de $\hat{\theta}$ et $\text{Log } \hat{\theta}$

D'après les propriétés des EMV [voir Monfort p. 105] $\hat{\theta}$ et $\text{Log } \hat{\theta}$ sont asymptotiquement normalement distribués. Pour décider entre $\hat{\theta}$ et $\text{Log } \hat{\theta}$ lors du test de l'hypothèse $\theta = 1$, le choix de Tanner repose sur une approximation asymptotique des coefficients d'asymétrie (skewness) et d'aplatissement (kurtosis) [voir Saporta, p. 29]. La connaissance de ces coefficients nécessite celle des quatre premiers moments de $\hat{\theta}$ et $\text{Log } \hat{\theta}$. Par conséquent, les développements de Taylor en probabilité utilisés seront limités à l'ordre quatre. On pose :

$$\begin{aligned} \xi &= \sum_{k=1}^s \frac{X_{2k} - \theta c_k X_{1k}}{1 + \theta c_k} ; \\ z &= \text{Log } \hat{\theta} - \text{Log } \theta ; \\ a_m &= \sum_{k=1}^s \frac{x_{.k} \theta^m c_k^m}{(1 + \theta c_k)^{m+1}} ; m = 1, 2, 3, 4 \end{aligned}$$

Lemme 1.3.1.- *Pour $N \rightarrow \infty$, on montre que*

$$\xi = a_1 z + \left(\frac{a_1}{2} - a_2\right) z^2 + \left(\frac{a_1}{6} - a_2 + a_3\right) z^3 + \left(\frac{a_1}{24} - \frac{7}{12} a_2 + \frac{3}{2} a_3 - a_4\right) z^4 + O_p(1)$$

Comme les v.a. $X_{1k} - E(X_{1k})$ ($k = 1, 2, \dots, s$) sont indépendantes et centrées, on a $E(\xi) = 0$;

$$\begin{aligned} E(\xi^2) &= \sum_{k=1}^s \text{var } X_{1k} \\ &= \sum_{k=1}^s \frac{x_{\cdot k} \theta_{c_k}}{(1 + \theta_{c_k})^2} = a_1 ; \end{aligned}$$

$$\begin{aligned} E(\xi^3) &= - \sum_{k=1}^s E(X_{1k} - EX_{1k})^3 \\ &= - \sum_{k=1}^s \frac{x_{\cdot k} \theta_{c_k}}{(1 + \theta_{c_k})^2} \left[\frac{\theta_{c_k}}{1 + \theta_{c_k}} - \frac{1}{1 + \theta_{c_k}} \right] \end{aligned}$$

$$E(\xi^3) = a_1 - 2a_2 ;$$

$$\begin{aligned} E(\xi^4) &= \sum_{k=1}^s E(X_{1k} - EX_{1k})^4 + 6 \sum_{k=1}^{s-1} \sum_{m>k} E(X_{1k} - EX_{1k})^2 E(X_{1m} - EX_{1m})^2 \\ &= \sum_{k=1}^s \left\{ 3 \left(\frac{\theta_{c_k} x_{\cdot k}}{1 + \theta_{c_k}} \right)^2 + \frac{\theta_{c_k} x_{\cdot k}}{(1 + \theta_{c_k})^2} \left[1 - 6 \frac{\theta_{c_k}}{(1 + \theta_{c_k})^2} \right] \right\} \\ &\quad + 6 \sum_{k=1}^{s-1} \sum_{m>k} \frac{\theta_{c_k} x_{\cdot k}}{(1 + \theta_{c_k})^2} \frac{\theta_{c_m} x_{\cdot m}}{(1 + \theta_{c_m})^2} \end{aligned}$$

$$E(\xi^4) = a_1 - 6a_2 + 6a_3 + 3a_1^2.$$

Lemme 1.3.3.- *Sous la condition $\hat{\theta} = \theta + o_p(1)(N \rightarrow \infty)$, on montre (voir Tanner, p. 340) que*

$$z = A_1 \xi + A_2 \xi^2 + A_3 \xi^3 + A_4 \xi^4 + O_p(1)$$

où

$$A_1 = 1/a_1; A_2 = (a_2 - \frac{a_1}{2})/a_1^3; A_3 = (6a_2^2 - 3a_1a_2 + a_1^2 - 3a_1a_3)/3a_1^5 ,$$

$$A_4 = (-\frac{a_1^3}{4} + a_1^2a_2 + a_1^2a_3 + a_1^2a_4 - 5a_1a_2a_3 - \frac{5}{2}a_1a_2^2 + 5a_2^2)/a_1^7 .$$

Proposition 1.3.1.- *En plus de la condition du lemme 1.3.3., on suppose que :*

$$\forall m > 4, \quad g_m E(\xi^m) = O(N^{-2})$$

où les coefficients g_m sont fonctions de a_1, a_2, a_3, a_4 . On montre alors qu'une approximation asymptotique des quatre premiers moments de z est :

$$\begin{aligned}
E(z) &= \frac{2a_2 - a_1}{2a_1^2} - \frac{2a_2 - a_1}{3a_1^5} [6a_2^2 - 3a_1a_2 + a_1^2 - 3a_1a_3] + \frac{3}{a_1^5} \\
&\quad \left[-\frac{a_1^3}{4} + a_1^2a_2 + a_1^2a_3 + a_1^2a_4 - 5a_1a_2a_3 - \frac{5}{2}a_1a_2^2 + 5a_2^3 \right] + O(N^{-2}) ; \\
E(z^2) &= \frac{1}{a_1} - \frac{(2a_2 - a_1)^2}{a_1^4} + \frac{3}{a_1^4} [5a_2^2 - 3a_1a_2 + \frac{11}{12}a_1^2 - 2a_1a_3] + O(N^{-2}) ; \\
E(z^3) &= \frac{7(2a_2 - a_1)}{2a_1^3} = O(N^{-2}) ; \\
E(z^4) &= \frac{3}{a_1^2} = O(N^{-2})
\end{aligned}$$

Preuve : *succinte.* La condition $g_m E(\xi^m) = O(N^{-2})$ permet de négliger tous les termes d'ordre N^{-p} ($p > 2$) dans l'approximation des quatre premiers moments de z . Par exemple :

$$\begin{aligned}
E(z^4) &= A_1^4 E(\xi^4) + O(N^{-2}) \\
&= \frac{1}{a_1^4} (a_1 - 6a_2 + 6a_3 + 3a_1^2) + O(N^{-2}) \\
E(z^4) &= \frac{3}{a_1^2} + O(N^{-2})
\end{aligned}$$

Proposition 1.3.2.- On pose $y = (\hat{\theta} - \theta)/\theta$. On montre, sous les mêmes conditions que la proposition 1.3.1, que

$$\begin{aligned}
E(y) &= \frac{1 + 8a_2}{8a_1^2} + \frac{7(2a_2 - a_1)}{12a_1^2} + \frac{1}{2a_1^4} [11a_2^2 - 5a_1a_2 + \frac{7}{4}a_1^2 - 6a_1a_3] \\
&\quad - \frac{2a_2 - a_1}{3a_1^5} [6a_2^2 - 3a_1a_2 + a_1^2 - 3a_1a_3] \\
&\quad + \frac{3}{a_1^5} \left[+\frac{a_1^3}{4} + a_1^2a_2 + a_1^2a_3 + a_1^2a_4 - 5a_1a_2a_3 - \frac{5}{2}a_1a_2^2 + 5a_2^3 \right] + O(N^{-2}) , \\
E(y^2) &= \frac{1}{a_1} + \frac{7}{4a_1^2} + \frac{7(2a_2 - a_1)}{2a_1^3} - \frac{(2a_2 - a_1)^2}{a_1^4} \\
&\quad + \frac{3}{a_1^4} [5a_2^2 - 3a_1a_3 + \frac{11}{12}a_1^2 - 2a_1a_3] + O(N^{-2}) , \\
E(y^3) &= \frac{7a_2 + a_1}{a_1^3} + O(N^{-2}) ; \\
E(y^4) &= \frac{3}{a_1^2} + O(N^{-2})
\end{aligned}$$

Preuve : Par définition de y et z , on a $N \rightarrow \infty$

$$y = e^z - 1 = z + \frac{1}{2}z^2 + \frac{1}{6}z^3 + \frac{1}{24}z^4 + O_p(1)$$

On utilise alors la proposition 1.3.2. pour déduire le résultat.

Proposition 1.3.3.- *Sous les mêmes conditions que la proposition 1.3.1., on montre qu'une approximation asymptotique ($N \rightarrow \infty$) des coefficients d'asymétrie γ_1 et d'aplatissement γ_2 est donnée par :*

$$(i) \quad \gamma_1(\text{Log } \hat{\theta}) = \frac{2(2a_2 - a_1)}{a_1^{3/2}} + O(N^{-3/2}), \gamma_2(\text{Log } \hat{\theta}) = 3 + O(N^{-1})$$

$$(ii) \quad \gamma_1(\hat{\theta}) = \frac{4a_2 + a_1}{a_1^{3/2}} + O(N^{-3/2}), \gamma_2(\hat{\theta}) = 3 + O(N^{-1})$$

Preuve :

i) En appliquant la proposition 1.3.1, on montre que

$$E(z - Ez)^2 = \frac{1}{a_1} + \frac{3}{a_1^4} [5a_2^2 - 3a_1a_2 + \frac{11}{12}a_1^2 - 2a_1a_3] - \frac{5(2a_2 - a_1)^2}{4a_1^4} + O(N^{-2})$$

$$E(z - Ez)^3 = \frac{2(2a_2 - a_1)}{a_1^3} + O(N^{-2}), E(z - Ez)^4 = \frac{3}{a_1^2} + O(N^{-2})$$

On en déduit alors que

$$\gamma_1(\text{Log } \hat{\theta}) = \frac{E(z - Ez)^3}{[E(z - Ez)^2]^{3/2}} = \frac{2(2a_2 - a_1)}{a_1^{3/2}} + O(N^{-3/2})$$

$$\gamma_2(\text{Log } \hat{\theta}) = \frac{E(z - Ez)^4}{[E(z - Ez)^2]^2} = 3 + O(N^{-1})$$

ii) De façon similaire, on déduit de la proposition 1.3.2 que

$$\gamma_1(\hat{\theta}) = \frac{4a_2 + a_1}{a_1^{3/2}} + O(N^{-3/2}), \gamma_2(\hat{\theta}) = 3 + O(N^{-1})$$

Remarque 1.3.1.- Les approximations de $E(z)$ et $E(y)$ montrent que $\text{Log } \hat{\theta}$ et $\hat{\theta}$ sont des estimateurs asymptotiquement sans biais de $\text{Log } \theta$ et θ . Ces moments donnent un ordre de grandeur du biais. D'autre part l'approximation de $E(z - Ez)^2$ montre, quand $N \rightarrow \infty$, que :

$$\text{var } z = \frac{1}{a_1} + \frac{3}{a_1^4} [5a_2^2 - 3a_1a_2 + \frac{11}{22}a_1^2 - 2a_1a_3] - \frac{5(2a_2 - a_1)^2}{4a_1^4} \quad (1.3.9).$$

Pour de faibles valeurs de N , le terme de droite peut être important. Par exemple si $c_1 = c_2 = \dots = c_s = 1$, on a :

$$\text{var } z = \frac{(1 + \theta)^2}{\theta N} \left[1 + \frac{3 - 2\theta + 3\theta^2}{2\theta} \times \frac{1}{N} \right] \quad (1.3.10.)$$

$\lim_{\theta \rightarrow 1} \left[\frac{3 - 2\theta + 3\theta^2}{2\theta} \right] = 2$, d'où le facteur $(1 + 2/N)$ préconisé par Tanner pour corriger $\text{var Log } \hat{\theta}$ donnée par l'expression (1.3.6.). De même, en étudiant le rapport

$$\frac{\gamma_1(\text{Log } \hat{\theta})}{\gamma_1(\hat{\theta})} \approx \frac{4a_2 - 2a_1}{4a_2 + a_1} \quad (1.3.11.)$$

et la distribution de $\text{Log } \hat{\theta}$ [Tanner p. 340-341], il suggère d'utiliser $\text{Log } \hat{\theta}$ pour la construction du test de l'hypothèse d'efficacité nulle ($\theta = 1$).

1.3.3.- Test de l'hypothèse $H_0: \theta_1 = \theta_2 = \dots = \theta_s$

Pour éprouver H_0 $\theta_1 = \theta_2 = \dots = \theta_s$, Tanner suggère de comparer la statistique

$$\hat{A}_s^2 = \sum_{k=1}^s \frac{(X_{2k} - \hat{\theta}_{c_k} X_{1k})^2}{\hat{\theta}_{c_k} x_{.k}} \quad (1.3.12.)$$

à une v.a. de loi de χ_{s-1}^2 (chi-deux à $s-1$ degrés de liberté), où $\hat{\theta}$ est l'EMV obtenu au paragraphe 1.3.1. Mais il ne donne pas la forme exacte de la distribution asymptotique ($N \rightarrow \infty$) de $\hat{A}_s^2$. Lassarre (1977) montre, sous l'hypothèse supplémentaire $\theta_1 = \theta_2 = \dots = \theta_s = \theta = 1$ et en utilisant un résultat de Chase (1972), que la statistique $\hat{A}_s^2$ est asymptotiquement de loi de χ_{s-1}^2 . On étudie dans ce paragraphe la distribution asymptotique de $\hat{A}_s^2$ sous la seule hypothèse H_0 d'égalité des effets. On note :

$$x^{1/2} = (x_{.1}^{1/2}, \dots, x_{.s}^{1/2})^T \in \mathbf{R}^s ;$$

$$\varepsilon = (\varepsilon_1, \dots, \varepsilon_s)^T \in \mathbf{R}^s \text{ le } \vec{v}.a. \text{ tel que } \varepsilon_k = [X_{1k} - x_{.k} q_{1k}(\theta)] / x_{.k}^{1/2} ;$$

$$\hat{\nu} = (\nu_1, \dots, \nu_s)^T \in \mathbf{R}^s \text{ le } \vec{v}.a. \text{ tel que } \hat{\nu}_k = x_{.k}^{1/2} [q_{1k}(\hat{\theta}) - q_{1k}(\theta)] ;$$

$$\hat{z} = (z_1, \dots, z_s)^T \in \mathbf{R}^s \text{ le } \vec{v}.a. \text{ tel que } \hat{z}_k = [\varepsilon_k - \hat{\nu}_k] / [q_{1k}(\theta) q_{2k}(\theta)] ;$$

$\mathbb{P}$ la matrice diagonale de taille s telle que

$$\mathbb{P} = \begin{bmatrix} q_{11}(\theta) q_{21}(\theta) & 0 & \dots & 0 \\ 0 & q_{12}(\theta) q_{22}(\theta) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & q_{1s}(\theta) q_{2s}(\theta) \end{bmatrix}$$

$$\mu = x^{1/2T} \mathbb{P} x^{1/2}$$

$$\Lambda = \frac{1}{\mu} \mathbb{P} x^{1/2} x^{1/2T}$$

Lemme 1.3.4.- *Le v.a. ε est asymptotiquement ($N \rightarrow \infty$) normalement distribué de moyenne nulle et de matrice de variance-covariance $\mathbb{P}$.*

Preuve : Vu l'indépendance des v.a. X_{1k} ($k = 1, 2, \dots, s$) et puisque sous l'hypothèse H_0 , X_{1k} suit une binomiale $x_{\cdot k}$ et $q_{1k}(\theta)$, on a :

$$E(\varepsilon_k) = x_{\cdot k}^{-1/2} E(X_{1k} - x_{\cdot k} q_{1k}(\theta)) = 0$$

et

$$E(\varepsilon_k \varepsilon_m) = \delta_{km} \cdot q_{1m}(\theta) q_{1k}(\theta)$$

où δ_{km} est le symbole de Kronecker.

Lemme 1.3.5.- *Soit A une matrice carrée de taille n telle que $A^2 = \alpha A$ pour $\alpha \neq 0$. Alors*

- i) *A possède une valeur propre non nulle égale à α et $n - 1$ valeurs propres identiques à zéro.*
- ii) *$I - A$ possède une valeur propre égale à $1 - \alpha$ et $n - 1$ valeurs propres identiques à 1.*

Preuve :

- i) Soit $\vec{u}$ un vecteur propre non nul de A associée à la valeur propre λ , alors $A\vec{u} = \lambda\vec{u}$. Soit $A^2\vec{u} = \alpha A\vec{u}$. Par suite

$$\lambda(\alpha - \lambda)\vec{u} = 0 \quad \text{c'est-à-dire} \quad \lambda = 0 \quad \text{ou} \quad \alpha = \lambda$$

- ii) On montre de même que si $\vec{v}$ est un vecteur propre non nul de $I - A$ associé à la valeur propre λ alors $\vec{v}$ est vecteur propre de A associé à la valeur propre $1 - \lambda$.

Théorème 1.3.1.- *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 $\hat{A}_s^2$ a même distribution asymptotique ($N \rightarrow \infty$) qu'une v.a. de loi de χ_{s-1}^2 .*

Preuve : En utilisant la relation

$$\frac{[X_{1k} - x_{\cdot k} q_{1k}(\hat{\theta})]^2}{x_{\cdot k} q_{1k}(\hat{\theta}) q_{2k}(\hat{\theta})} = \frac{1}{q_{1k}(\theta) q_{2k}(\theta)} \left[\frac{X_{1k} - x_{\cdot k} q_{1k}(\theta)}{x_{\cdot k}^{1/2}} - \frac{x_{\cdot k} [q_{1k}(\hat{\theta}) - q_{1k}(\theta)]}{x_{\cdot k}^{1/2}} \right]^2$$

on montre que

$$\frac{[X_{2k} - \hat{\theta} c_k X_{1k}]^2}{\hat{\theta} c_k x_{\cdot k}} = \frac{(\varepsilon_k - \hat{\nu}_k)^2}{q_{1k}(\theta) q_{2k}(\theta)} [1 + o_p(1)]$$

Par conséquent

$$\hat{A}_s^2 = \sum_{k=1}^s \frac{(\varepsilon_k - \hat{\nu}_k)^2}{q_{1k}(\theta) q_{2k}(\theta)} [1 + o_p(1)] \quad (1.3.13.)$$

La démonstration consiste à déterminer la distribution asymptotique de la forme quadratique

$$\hat{z}^T \hat{z} = \sum_{k=1}^s \frac{(\varepsilon_k - \hat{\nu}_k)^2}{q_{1k}(\theta) q_{2k}(\theta)} \quad (1.3.14.)$$

En utilisant un développement de Taylor en probabilité à l'ordre un de $q_{1k}(\hat{\theta})$ au voisinage de θ , on montre d'une part que

$$\hat{\nu}_k = x_{\cdot k}^{1/2} \left(\frac{\partial q_{1k}(u)}{\partial u} \right)_{u=\theta} (\hat{\theta} - \theta) + o_p(1),$$

et d'autre part que

$$\hat{\theta} - \theta = -\frac{\theta}{\mu} \sum_{m=1}^s x_{\cdot m}^{1/2} \varepsilon_m + o_p(1)$$

En combinant ces deux dernières égalités, on a

$$\hat{\nu}_k = \mu_k \sum_{m=1}^s x_{\cdot m}^{1/2} \varepsilon_m + o_p(1) \quad (1.3.15)$$

avec $\mu_k = \frac{x_{\cdot k}^{1/2}}{\mu} q_{1k}(\theta) q_{2k}(\theta)$ ($k = 1, 2, \dots, s$). La relation (1.3.15) permet d'écrire que :

$$\hat{z}_k = \left(\frac{1}{q_{1k}(\theta) q_{2k}(\theta)} \right)^{1/2} [(1 - \mu_k x_{\cdot k}^{1/2}) \varepsilon_k - \sum_{\substack{m=1 \\ m \neq k}}^s \mu_k x_{\cdot m}^{1/2} \varepsilon_m] + o_p(1)$$

$k = 1, 2, \dots, s$. En notation matricielle, on obtient

$$\hat{z} = \mathbb{P}^{-1/2} [I_S - \Lambda] \varepsilon + o_p(1) \quad (1.3.16)$$

La forme quadratique [voir par exemple Molinari] $\hat{z}^T \hat{z}$ est distribuée comme une v.a. dont la loi est de la forme $\sum_{m=1}^s \lambda_m z_m^2$, 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 (on utilise aussi le lemme 1.3.5)

$$\Omega = (I_s - \Lambda^T) \mathbb{P}^{-1} (I_s - \Lambda) \mathbb{P}$$

En développant $\mathbb{P}^{-1} (I_s - \Lambda) \mathbb{P}$ et en remarquant que $\mathbb{P}^{-1} \Lambda \mathbb{P} = \Lambda^T$ (ceci implique Λ et Λ^T sont des matrices semblables donc admettent les mêmes valeurs propres), on montre que

$$\Omega = I_s - \Lambda^T$$

$\Lambda^{T^2} = \Lambda^T$, par suite (lemme 1.3.6) $I_s - \Lambda^T$ admet une valeur propre nulle et $(s - 1)$ valeurs propres identiques à 1. Par conséquent la forme quadratique $\hat{z}^T \hat{z}$ est distribuée comme une v.a. de loi de χ_{s-1}^2 .

Sur le plan pratique, on compare la statistique $\hat{A}_s^2$ à la valeur α_{s-1} que dépasse une v.a. de loi de χ_{s-1}^2 avec une probabilité α (par exemple $\alpha = 0,05$. Tanner suggère d'utiliser $\alpha < 0,20$). Si la valeur calculée de $\hat{A}_s^2$ est raisonnablement plus grande que celle lue, on rejette l'hypothèse H_0 c'est-à-dire qu'il y a vraisemblablement une différence dans l'effet de la mesure aux différents sites. Comment estimer alors une telle variabilité? Il propose une méthode utilisant la statistique $\hat{A}_s^2$ pour donner une approximation de cette variabilité.

1.3.4.- Approximation de la variabilité de l'effet aux différents sites.

En plus de l'hypothèse de distribution faite sur les $\vec{v}.a.$ (X_{1k}, X_{2k}) ($k = 1, 2, \dots, s$), Tanner suppose que $\theta_1, \theta_2, \dots, \theta_s$ est une suite de v.a. identiquement distribuée de moyenne θ et de variance σ^2 . En supposant θ connue et donnée par (1.3.3), il propose une méthode d'approximation pour estimer σ^2 et l'interprète comme une mesure de la variabilité de l'effet aux différents sites. On note par la suite E^X [resp. $E^{\mathcal{D}}$] l'espérance mathématique par rapport à la distribution des $\vec{v}.a.$ (X_{1k}, X_{2k}) [resp. θ_k] ($k = 1, 2, \dots, s$).

Proposition 1.3.4.- *On suppose que*

(c₁) $\sum_{m=1}^L \omega_{k,m+2} E^{\mathcal{D}}[|\theta_k - \theta|^{m+2}] = O(1)$, pour tout entier $L \geq 2$ où $\omega_{k,m+2}$ sont des coefficients réels positifs fonction de θ .

(c₂) $\hat{\theta} - \theta = o_p(1) (N \rightarrow \infty)$

alors un estimateur asymptotique ($N \rightarrow \infty$) de σ^2 est donné par

$$\frac{\hat{A}_s^2 - (s-1)}{\sum_{k=1}^s \frac{c_k(x_{\cdot k} + \theta c_k - 2)}{\theta(1 + \theta c_k)^2}}$$

Preuve : On considère la fonction aléatoire

$$u \mapsto \hat{A}_s^2(u) = \sum_{k=1}^s \frac{(X_{2k} - u c_k X_{1k})^2}{u c_k x_{\cdot k}}$$

obtenue en remplaçant $\hat{\theta}$ par u dans l'expression de la statistique $\hat{A}_s^2$ et où X_{1k} [resp. X_{2k}] suit une distribution binomiale de taille $x_{\cdot k}$ et de paramètre $1/(1 + \theta_k c_k)$ [resp. $\theta_k c_k/(1 + \theta_k c_k)$]. La méthode consiste à évaluer $E(\hat{A}_s^2) = E^{\mathcal{D}} E^X[\hat{A}_s^2(u)]$. Par un calcul direct, Tanner montre que

$$E^X[\hat{A}_s^2(u)] = \sum_{k=1}^s \frac{x_{\cdot k} c_k (\theta_k - u)^2}{u(1 + \theta_k c_k)^2} + \sum_{k=1}^s \frac{\theta_k}{u} \left(\frac{1 + u c_k}{1 + \theta_k c_k} \right)^2 \quad (1.3.17.)$$

Au point $u = \theta$, on a

$$\begin{aligned} E(\hat{A}_s^2) &= E^{\mathcal{D}} E^X(\hat{A}_s^2(\theta)) = \sum_{k=1}^s \frac{x_{\cdot k} c_k}{\theta} E^{\mathcal{D}} \left[\frac{(\theta_k - \theta)^2}{(1 + \theta_k c_k)^2} \right] \\ &\quad + \sum_{k=1}^s \frac{(1 + \theta c_k)^2}{\theta} E^{\mathcal{D}} \left[\frac{\theta_k}{(1 + \theta_k c_k)^2} \right] \end{aligned}$$

En posant $z_k = \theta_k - \theta$ et en utilisant le développement de Taylor en probabilité à l'ordre 2 des fonctions aléatoires $(\theta_k - \theta)^2/[(1 + \theta_k c_k)^2]$ et $\theta_k/[(1 + \theta_k c_k)^2]$, on montre, sous la condition (c_1) , que

$$E^{\mathcal{D}} E^X[\hat{A}_s^2(\theta)] = s + \sigma^2 \sum_{k=1}^s \frac{c_k(x_{\cdot k} + \theta c_k - 2)}{\theta(1 + \theta c_k)^2} + O(N) \quad (1.3.18.)$$

Si σ^2 est suffisamment petite (si la variabilité de l'effet de la mesure aux différents sites est négligeable) l'inégalité de Tchebychev permet d'écrire

$$\forall \varepsilon > 0, \forall \eta > 0 \quad \text{Proba}\{|\theta_k - \theta| < \varepsilon\} < \eta, \quad \forall k = 1, 2, \dots, s$$

On a donc $\theta_1 = \theta_2 = \dots = \theta_s = \theta$ sauf pour un ensemble de probabilité nulle. Dans ces conditions le théorème 1.3.1. permet d'utiliser $s - 1$ comme degré de liberté. Par conséquent on a :

$$E^{\mathcal{D}} E^X[\hat{A}_s^2(\theta)] = s - 1 + \sigma^2 \sum_{k=1}^s \frac{c_k(x_{\cdot k} + \theta c_k - 2)}{\theta(1 + \theta c_k)^2} + O(N) \quad (1.3.19.)$$

En estimant le membre de droite par $\hat{A}_s^2$ et en utilisant la condition (c_2) , on obtient l'estimateur asymptotique ($N \rightarrow \infty$) de σ^2 désiré.

1.4.- ANALYSE EN CAS D'HÉTÉROGÉNÉITÉ DE L'EFFET

1.4.1.- Estimation de l'effet de la mesure

On suppose comme précédemment que $\theta_1, \theta_2, \dots, \theta_s$ est une suite de v.a. provenant d'une population d'espérance mathématique θ et de variance σ^2 . Il interprète alors θ comme l'effet de la mesure de sécurité routière et σ^2 comme une estimation de la variabilité de l'effet aux différents sites expérimentaux. En notant toujours $\hat{\theta}$ l'estimateur de l'effet, Tanner propose une méthode d'approximation pour estimer θ et évaluer $\text{var Log } \hat{\theta}$. Il utilise pour cela le résultat du paragraphe 1.3.4.

Il admet pour $N \rightarrow \infty$ que $\hat{\theta}$ vérifie toujours l'équation

$$\sum_{k=1}^s \left\{ \frac{x_{2k} - \hat{\theta} c_k x_{1k}}{1 + \hat{\theta} c_k} \right\} = 0 \quad (1.3.1.)$$

Cette hypothèse suppose donc l'utilisation de l'E.M.V comme estimateur de l'espérance mathématique de la distribution dont proviennent les v.a. θ_k . Cette approche semble raisonnable si la variabilité de l'effet de la mesure aux différents sites est suffisamment faible.

1.4.2.- Approximation de la variance asymptotique de $\text{Log } \hat{\theta}$

Pour $N \rightarrow \infty$, on considère la fonction aléatoire

$$\psi: u \mapsto \sum_{k=1}^s \left\{ \frac{X_{2k} - u c_k X_{1k}}{1 + u c_k} \right\}$$

et on pose

$$\phi = \frac{[\hat{A}_s^2 - (s-1)] \sum_{k=1}^s \left[\frac{c_k(x_{\cdot k} + \theta c_k - 2)}{\theta(1 + \theta c_k)^2} \cdot \frac{\theta c_k x_{\cdot k}}{(1 + \theta c_k)^2} \right]}{\sum_{k=1}^s \frac{c_k(x_{\cdot k} + \theta c_k - 2)}{\theta(1 + \theta c_k)^2} \cdot \sum_{k=1}^s \frac{\theta c_k x_{\cdot k}}{(1 + \theta c_k)^2}} \quad (1.4.1.)$$

Remarque 1.4.1.-

i) Si on suppose $\theta c_k - 2$ négligeable par rapport à $x_{\cdot k}$ ($k = 1, 2, \dots, s$) alors :

$$\phi = \frac{[\hat{A}_s^2 - (s-1)] \sum_{k=1}^s \left[\frac{\theta c_k x_{\cdot k}}{(1 + \theta c_k)^2} \right]^2}{\left[\sum_{k=1}^s \frac{\theta c_k x_{\cdot k}}{(1 + \theta c_k)^2} \right]^2} \quad (1.4.2.)$$

- ii) Si on tient maintenant compte de l'étude de la fonction qui à $x \mapsto x/(1+x^2)$ [voir chapitre 2. fig. 1] et en multipliant le membre de droite de (1.4.2.) par $s/(s-1)$, on obtient :

$$\phi = \left[\frac{\hat{A}_s^2}{s-1} - 1 \right] \frac{s \sum_{k=1}^s x_{\cdot k}^2}{N^2} \quad (1.4.3.)$$

- iii) Si de plus $x_{\cdot 1} = x_{\cdot 2} = \dots = x_{\cdot s}$ alors

$$1 + \phi = \frac{\hat{A}_s^2}{s-1} \quad (1.4.4.)$$

Lemme 1.4.1.- *Sous les mêmes conditions que la proposition 1.3.4, on a :*

$$\text{var } \psi(\theta) = \sum_{k=1}^s \frac{\theta c_k x_{\cdot k}}{(1 + \theta c_k)^2} + \sigma^2 \sum_{k=1}^s \frac{x_{\cdot k} c_k^2 (x_{\cdot k} + \theta c_k - 2)}{(1 + \theta c_k)^4} + O(N)$$

Preuve : En réécrivant ψ au point $u = \theta$ sous la forme

$$\psi(\theta) = \sum_{k=1}^s \frac{x_{\cdot k}}{1 + \theta c_k} - \sum_{k=1}^s X_{1k}$$

on a

$$\text{var } \psi(\theta) = \sum_{k=1}^s \text{var } X_{1k} \quad (1.4.5.)$$

car $x_{\cdot k}$ est supposé non aléatoire et les v.a. X_{1k} ($k = 1, 2, \dots, s$) sont indépendantes. Par définition, on a :

$$\text{var}(X_{1k}) = E^{\mathcal{D}} E^X [(X_{1k})^2] - [E^{\mathcal{D}} E^X (X_{1k})]^2 \quad (1.4.6.)$$

$$E^{\mathcal{D}} E^X (X_{1k}) = x_{\cdot k} E^{\mathcal{D}} \left[\frac{1}{1 + \theta_k c_k} \right] \quad (1.4.7.)$$

$$E^{\mathcal{D}} E^X [(X_{1k})^2] = x_{\cdot k} E^{\mathcal{D}} \left(\frac{1}{1 + \theta_k c_k} \right) + x_{\cdot k} (x_{\cdot k} - 1) E^{\mathcal{D}} \left[\frac{1}{(1 + \theta_k c_k)^2} \right] \quad (1.4.8.)$$

Sous la condition (c_1) de la proposition 1.3.4, on a :

$$E^{\mathcal{D}} \left[\frac{1}{1 + \theta_k c_k} \right] = \frac{1}{1 + \theta c_k} + \frac{c_k^2}{(1 + \theta c_k)^2} \sigma^2 + O(x_{\cdot k}) \quad (1.4.9)$$

$$E^{\mathcal{D}} \left[\frac{1}{(1 + \theta_k c_k)^2} \right] = \frac{1}{(1 + \theta c_k)^2} + \frac{3c_k^2}{(1 + \theta c_k)^4} \sigma^2 + O(x_{\cdot k}) \quad (1.4.10)$$

En substituant (1.4.9.) et (1.4.10.) dans (1.4.6.) et ensuite dans (1.4.5) on a :

$$\text{var } \psi(\theta) = \sum_{k=1}^s \frac{\theta c_k x_{\cdot k}}{(1 + \theta c_k)^2} + \sigma^2 \sum_{k=1}^s \frac{x_{\cdot k} c_k^2 (x_{\cdot k} + \theta c_k - 2)}{(1 + \theta c_k)^4} + O(N).$$

$\text{var } \psi(\theta)$ est interprétée comme étant les éventuelles variations pouvant provenir à la fois de l'hypothèse de distribution faite sur les $\vec{v}.a.$ (X_{1k}, X_{2k}) et de celle effectuée sur les v.a. θ_k .

Lemme 1.4.2.- *Si en plus des conditions du lemme 1.4.1, on suppose que*

$$(c_3) \text{ Log } \hat{\theta} = \text{Log } \theta + O_p(r_N), \text{ avec } r_N \text{ et } N^{-2} r_N^3 \rightarrow O \text{ quand } N \rightarrow \infty$$

$$(c_4) E^{\mathcal{D}} E^X \text{Log } \hat{\theta} = \text{Log } \theta$$

$$(c_5) E^{\mathcal{D}} E^X [\text{Log } \hat{\theta} - \text{Log } \theta]^2 < \infty, \text{ alors}$$

$$\text{var Log } \hat{\theta} = \frac{\sum_{k=1}^s \frac{\theta c_k x_{\cdot k}}{(1 + \theta c_k)^2} + \sigma^2 \sum_{k=1}^s \frac{x_{\cdot k} c_k^2 (x_{\cdot k} + \theta c_k - 2)}{(1 + \theta c_k)^4}}{[\sum_{k=1}^s \frac{\theta c_k x_{\cdot k}}{(1 + \theta c_k)^2}]^2}$$

Preuve : En utilisant le développement de Taylor en probabilité au voisinage de θ (voir Wolter p. 223), on peut écrire, sous la condition (c_3)

$$\psi(\hat{\theta}) - \psi(\theta) = (\text{Log } \hat{\theta} - \text{Log } \theta) \left(\frac{\partial \psi(u)}{\partial \text{Log } u} \right)_{u=\theta} + O_p(r_N^2)$$

Soit alors

$$(\text{Log } \hat{\theta} - \text{Log } \theta)^2 = \frac{[\psi(\hat{\theta}) - \psi(\theta)]^2}{\left(\frac{\partial \psi(u)}{\partial \text{Log } u} \right)_{u=\theta}^2} + O_p(N^{-2} r_N^3) \quad (1.4.11.)$$

En tenant compte des conditions (c_4) , (c_5) et du fait que $\psi(\hat{\theta}) = 0$, on a pour $N \rightarrow \infty$

$$\text{var Log } \hat{\theta} = \frac{E^{\mathcal{D}} E^X [(\psi(\theta))^2]}{\left(\frac{\partial \psi(u)}{\partial \text{Log } u} \right)_{u=\theta}^2} \quad (1.4.12.)$$

où $\left(\frac{\partial \psi(u)}{\partial \text{Log } u} \right)_{u=\theta} = - \sum_{k=1}^s \frac{\theta c_k x_{\cdot k}}{(1 + \theta c_k)^2}$. On montre d'autre part que

$$\begin{aligned} E^{\mathcal{D}} E^X [(\psi(\theta))^2] &= \text{var } \psi(\theta) + [E^{\mathcal{D}} E^X (\psi(\theta))]^2 \\ &= \text{var } \psi(\theta) + O(N^2) \end{aligned}$$

En substituant dans (1.4.12) on a

$$\text{var Log } \hat{\theta} = \frac{\text{var } \psi(\theta)}{[\sum_{k=1}^s \frac{\theta c_k x_{\cdot k}}{(1 + \theta c_k)^2}]^2} \quad N \rightarrow \infty$$

où l'expression de $\text{var } \psi(\theta)$ est donnée par le lemme 1.4.1.

Proposition 1.4.1.- *Sous les mêmes conditions que le lemme 1.4.2., on montre qu'une approximation asymptotique ($N \rightarrow \infty$) de la variance de $\text{Log } \hat{\theta}$ est*

$$\text{var Log } \hat{\theta} = \frac{1 + \phi}{\sum_{k=1}^s \frac{\theta c_k x_{\cdot k}}{(1 + \theta c_k)^2}}$$

où $\phi \simeq \left[\frac{\hat{A}_s^2}{s-1} - 1 \right] \frac{s \sum_{k=1}^s x_{\cdot k}^2}{N^2}$.

Preuve : D'après le lemme 1.4.2., on a :

$$\text{var Log } \hat{\theta} = \frac{1}{\sum_{k=1}^s \frac{\theta c_k x_{\cdot k}}{(1 + \theta c_k)^2}} \left[1 + \sigma^2 \frac{\sum_{k=1}^s \frac{x_{\cdot k} c_k^2 (x_{\cdot k} + \theta c_k - 2)}{(1 + \theta c_k)^4}}{\sum_{k=1}^s \frac{\theta c_k x_{\cdot k}}{(1 + \theta c_k)^2}} \right] + O\left(\frac{1}{N}\right)$$

En remplaçant σ^2 par son expression donnée à la proposition 1.3.4., on a :

$$\text{var Log } \hat{\theta} = \frac{1 + \phi}{\sum_{k=1}^s \frac{\theta c_k x_{\cdot k}}{(1 + \theta c_k)^2}}$$

où ϕ est donné par l'expression (1.4.1). On utilise alors (1.4.3) de la remarque 1.4.1 pour obtenir l'approximation de ϕ désirée. Tanner suggère d'utiliser le facteur $(1 + \phi)$ lorsque la statistique $\hat{A}_s^2$ est significative, soit

$$\text{Proba}\{\chi_{s-1}^2 > \alpha_{s-1}\} = \alpha < 0.20$$

Pour de grandes valeurs de α ($\alpha > 0.20$) l'erreur aléatoire due à ϕ peut l'emporter sur l'avantage d'avoir un estimateur non biaisé de $\text{var Log } \hat{\theta}$.

1.4.3.- Test de l'hypothèse d'efficacité nulle

Il s'agit de tester l'hypothèse $\theta = 1$ (l'efficacité de la mesure de sécurité routière est nulle) c'est-à-dire tester la moyenne d'une certaine distribution égale à une valeur donnée. Tanner suggère la statistique

$$Z = \frac{\text{Log } \hat{\theta}}{[\text{var Log } \hat{\theta}]^{1/2}} \quad (1.4.13.)$$

où $\text{var Log } \hat{\theta} = \frac{(1 + \phi)}{\sum_{k=1}^s \frac{\theta c_k x_{\cdot k}}{(1 + \theta c_k)^2}}$.

Cette statistique est construite comme celle donnée par l'expression (1.3.7.) mais où la variance asymptotique ($N \rightarrow \infty$) de $\text{Log } \hat{\theta}$ est corrigée par intégration

du facteur $(1 + \phi)$. Il suppose sous l'hypothèse $\theta = 1$ que la statistique Z admet pour distribution asymptotique une loi $\mathcal{N}(0; 1)$. On rejette $\theta = 1$ si la valeur de Z est en valeur absolue supérieure à 1.96.

1.4.4.- Une application : Transformation de carrefours en ronds-points

Les données proviennent d'une transformation de sept carrefours (sites expérimentaux) en ronds-points [Tanner p. 338]. Les observations obtenues sont regroupées dans le tableau ci-après :

Sites	Avant	Après	Total	Contrôle
n° 1	1	6	7	1,04
n° 2	6	3	9	1,25
n° 3	9	5	14	1,11
n° 4	16	5	21	2,36
n° 5	10	0	10	1,13
n° 6	2	2	4	1,69
n° 7	5	0	5	1,61

Pour le premier carrefour (site n° 1) le tableau montre qu'il n'y a eu 1 accident avant l'aménagement et 6 accidents après, alors que le rapport de contrôle $c_1 = 1,04$: c'est-à-dire que dans la zone de contrôle il y a eu, par rapport à la période avant, 4% d'accidents en plus dans la période après. D'autre part, tous les rapports de contrôle sont supérieurs à 1. Il y a donc une augmentation du nombre d'accidents dans les zones de contrôle après la transformation.

Pour estimer l'effet (recherche de $\hat{\theta}$), on résoud l'équation (1.3.3) en utilisant un algorithme de type gradient [voir chapitre 4]. Pour faire fonctionner ces algorithmes itératifs, il faut se donner un point de départ (valeur initiale). Afin de montrer l'importance d'un tel point, on a essayé différentes valeurs initiales. Les résultats sont consignés dans le tableau ci-après :

Valeur initiale	Nombre d'itérations	Estimation
0,022	8	0,283
0,050	6	0,283
0,100	4	0,283
0,280	0	0,283

On remarque que si l'on donne 0,022 comme valeur de départ de l'algorithme, on obtient 0,283 comme solution de (1.3.3.) après 8 itérations. Si le point de départ est mal spécifié, on peut ne pas avoir de solution ou on peut augmenter le nombre d'itérations. La valeur estimée de l'effet est $\hat{\theta} = 0,283$ soit :

$$\text{Log } \hat{\theta} = -1,27$$

La valeur calculée de la statique $\hat{A}_s^2$ avec $\hat{\theta} = 0,283$ est :

x_{2k}	$\hat{\theta}x_{1kc_k}$	$(x_{2k} - \hat{\theta}x_{1kc_k})^2$	$\hat{\theta}x_{.kc_k}$	$(x_{2k} - \hat{\theta}x_{1kc_k})^2 / \hat{\theta}x_{.kc_k}$
6	0,291	32,59	2,038	15,99
3	2,100	0,81	3,15	0,257
5	2,797	4,85	4,35	1,115
5	10,570	31,02	13,87	2,236
0	3,160	9,98	3,164	3,154
2	0,940	1,123	1,893	0,593
0	2,254	5,08	2,254	2,253

$\hat{A}_s^2 = 25,6$. Avec $\alpha = 5\%$, le test du χ_6^2 donne comme valeur 12,59. Cette dernière valeur est fortement significative. Ce qui veut dire que l'effet de l'installation d'un rond-point n'est pas le même à tous les carrefours. Il est donc nécessaire d'introduire le facteur $1 + \phi$ dans l'estimateur de $\text{var Log } \hat{\theta}$ si l'on souhaite des conclusions générales concernant l'effet de ce type de changement plutôt que des

conclusions dans le cas particulier des sept carrefours étudiés. Pour tester l'écart entre θ et 1, on calcule

$$\text{var Log} \hat{=} \frac{1 + \phi}{\sum_{k=1}^s \frac{\theta c_k x_{\cdot k}}{(1 + \theta c_k)^2}} = 0,3192$$

en faisant $\theta = 1$ au dénominateur. La statistique

$$Z = \frac{\text{Log } \hat{\theta}}{[\text{var Log } \hat{\theta}]^{1/2}} = \frac{-1,27}{0,565} = -2,25$$

Avec $\alpha = 5\%$, il conclut que les ronds-points ainsi étudiés tendent à faire diminuer le nombre des accidents puisque $\hat{\theta} < 1$. On peut, si les données l'exige, calculer $\text{var Log } \hat{\theta}$ en faisant $\theta = 0,283$ au dénominateur. Quelques simplifications sont aussi données par Tanner.

CHAPITRE 2:
ANALYSE DE L'EFFET
A L'AIDE D'UN TABLEAU
"AVANT-APRÈS"
 $2 \times s \times r$

2.1.- INTRODUCTION

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 (expérimentaux) et sur des sites sur lesquels la mesure n'est pas directement appliquée. On n'aborde pas ici le problème de la régression vers la moyenne largement discuté dans la littérature [voir, Hauer (1980 a); Wright et al. (1988)]. Pour tenir compte des difficultés liées au choix de la période d'observation, au choix des sites traités et des sites témoins ([38]), 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 l'instauration de la mesure. Il y a alors autant de rapports de contrôle pour chaque site que de types d'accidents retenus dans l'analyse. 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. Comment combiner ces données d'accidents et les rapports de contrôle pour juger de l'efficacité ou non de cette mesure ?

Différentes approches statistiques (voir [49], [64], [27], [23]) peuvent être proposées pour apporter une réponse à cette question. Tanner [voir chapitre 1] a proposé un modèle et une méthode pour estimer et tester l'efficacité d'une mesure de sécurité routière prise pour plusieurs sites dans le cas où n'est considéré qu'un seul type d'accident. En restant dans l'optique de Tanner, on construit des modèles (section 2.2) basés sur une distribution multinomiale et permettant de répartir les accidents de chaque site 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 analyse, dans une première étape (section 2.3), l'efficacité sous l'hypothèse d'égalité de l'effet aux différents sites. On estime (paragraphe 2.3.2) l'effet et certains paramètres auxiliaires ainsi que leur variance asymptotique (précision). On suggère, ensuite, une statistique (paragraphe 2.3.4) pour non seulement éprouver cette hypothèse d'égalité mais aussi pour donner une approximation de la variabilité de l'effet aux différents sites lorsque cette hypothèse d'égalité n'est pas retenue. Dans ces conditions, des complications tant théoriques que pratiques apparaissent dans l'analyse. On propose alors une méthode d'approximation (section 2.4) pour estimer l'efficacité et tester l'hypothèse d'efficacité nulle. On généralise ainsi les résultats de Tanner à des tableaux "avant-après" de

grandes dimensions.

2.2.- MODELES

2.2.1.- Notations

On suppose que les données dont on dispose se présentent sous la forme suivante :

	Avant	Après
	type 1 $\cdots$ type j $\cdots$ type r	type 1 $\cdots$ type j $\cdots$ type r
expérimental 1	$x_{111} \quad \cdots \quad x_{1j1} \quad \cdots \quad x_{1r1}$	$x_{211} \quad \cdots \quad x_{2j1} \quad \cdots \quad x_{2r1}$
$\vdots$	$\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots$	$\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots$
expérimental k	$x_{11k} \quad \cdots \quad x_{1jk} \quad \cdots \quad x_{1rk}$	$x_{21k} \quad \cdots \quad x_{2jk} \quad \cdots \quad x_{2rk}$
$\vdots$	$\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots$	$\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots$
expérimental s	$x_{11s} \quad \cdots \quad x_{1js} \quad \cdots \quad x_{1rs}$	$x_{21s} \quad \cdots \quad x_{2js} \quad \cdots \quad x_{2rs}$

On note

s le nombre de sites expérimentaux où la mesure est appliquée ;

r le nombre de types d'accidents retenus dans l'analyse ;

$X_{1k} = (X_{11k}, X_{12k}, \dots, X_{1rk})$ (resp. $X_{2k} = (X_{21k}, \dots, X_{2rk})$) le vecteur aléatoire donnant les effectifs des r types d'accidents sur le site expérimental $n^{\circ}k$ ($k = 1, 2, \dots, s$) 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 sur l'ensemble des sites 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} est supposé strictement positif, non aléatoire et parfaitement bien déterminé par les effectifs d'accidents observés, avant et après, dans le site de contrôle associé au site k) ;

$p_{jk} : (0 < p_{jk} < 1)$ la probabilité pour qu'un accident qui se produit dans le site expérimental k avant la mesure soit du type j , $\sum_{j=1}^r p_{jk} = 1$;

$P_k = (p_{1k}, p_{2k}, \dots, p_{rk})^T$ le vecteur à r dimensions de composantes les p_{jk} ;

$\theta_k^* = \sum_{j=1}^r x_{2jk} / \sum_{j=1}^r c_{jk} x_{1jk}$, l'effet apparent de la mesure au site k ; il représente le rapport du nombre total d'accidents observé après au nombre total d'accidents qu'on pourrait s'attendre à observer après si la mesure n'avait aucun effet au site k (comportement semblable à celui observé dans le site de contrôle).

2.2.2.- Modèles.

Les modèles statistiques généralement utilisés pour analyser l'effet d'une mesure de sécurité routière dépendent fortement de la nature de l'étude menée et des informations disponibles concernant les variables aléatoires observées.

On modélise l'effet d'une mesure à l'aide soit de distribution de Poisson [voir Hauer (1980a); Danielson (1986, 1988); Lassarre (1977, 1985)] soit de distribution binomiale négative [Tanner, 1958] ou encore de distribution binomiale négative [Boyer et al (1990)].

On suppose ici que la mesure a un effet multiplicateur sur le nombre des accidents indépendamment du type d'accidents. Soit donc x_{2jk}^* le nombre d'accidents du type j ($j = 1, 2, \dots, r$) qu'on pourrait s'attendre à observer après si la mesure n'avait aucun effet au site k et si ce site avait un comportement semblable à celui observé dans le site de contrôle associé, on pose alors:

$$\text{modèle 1 : } x_{2jk}^* = \theta_k^* c_{jk} x_{1jk} \quad j = 1, 2, \dots, r$$

$$\text{modèle 2 : } x_{2jk}^* = \theta_k^* c_k^* x_{1jk} \quad j = 1, 2, \dots, r;$$

avec $c_k^* = \sum_{j=1}^r c_{jk} p_{jk}^*$, $p_{jk}^* = x_{1jk} / (\sum_{m=1}^r x_{1mk})$.

Dans ces conditions, la proportion d'accidents du typ j ($j = 1, 2, \dots, r$) escompté au site k est donnée par :

$$\begin{aligned} \text{modèle 1 : } q_{1jk}(\theta_k^*, P_k^*) &= \frac{p_{jk}^*}{1 + \theta_k^* c_k^*}, \quad q_{2jk}(\theta_k^*, P_k^*) = \frac{\theta_k^* c_{jk} p_{jk}^*}{1 + \theta_k^* c_k^*} \\ \text{modèle 2 : } q_{1jk}(\theta_k^*, P_k^*) &= \frac{p_{jk}^*}{1 + \theta_k^* c_k^*}, \quad q_{2jk}(\theta_k^*, P_k^*) = \frac{\theta_k^* c_k^* p_{jk}^*}{1 + \theta_k^* c_k^*} \end{aligned}$$

On fait alors l'hypothèse que les vecteurs aléatoires (X_{1k}, X_{2k}) , ($k = 1, 2, \dots, s$) sont indépendants et suivent chacun une distribution multinomiale :

$$\begin{aligned} \mathcal{M}(x_{..k} ; q_{11k}(\theta_k, P_k), \dots, q_{1jk}(\theta_k, P_k), \dots, q_{1rk}(\theta_k, P_k), \\ q_{21k}(\theta_k, P_k), \dots, q_{2jk}(\theta_k, P_k), \dots, q_{2rk}(\theta_k, P_k)) \end{aligned} :$$

$$\text{modèle 1 : } q_{1jk}(\theta_k, P_k) = \frac{p_{jk}}{1 + \theta_k c_k}, \quad q_{2jk}(\theta_k, P_k) = \frac{\theta_k c_{jk} p_{jk}}{1 + \theta_k c_k} \quad (j = 1, 2, \dots, r)$$

$$\text{modèle 2 : } q_{1jk}(\theta_k, P_k) = \frac{p_{jk}}{1 + \theta_k c_k}, \quad q_{2jk}(\theta_k, P_k) = \frac{\theta_k c_k p_{jk}}{1 + \theta_k c_k} \quad (j = 1, 2, \dots, r),$$

avec $\theta_k (0 < \theta_k)$ le paramètre effet de la mesure au site expérimental k c'est-à-dire la valeur que prendrait θ_k^* si le vecteur aléatoire (X_{1k}, X_{2k}) prenait pour valeur son espérance mathématique,

$$c_k = \sum_{j=1}^r c_{jk} p_{jk} \quad k = 1, 2, \dots, s$$

le coefficient moyen de contrôle au site k . Les deux modèles ainsi construits tiennent compte à la fois de la structure du site expérimental avant la mesure (p_{jk}) et du comportement du site contrôle (c_{jk}) au cours des deux périodes. Ils représentent une généralisation du modèle de Tanner.

Etant donné les rapports de contrôle c_{jk} , on peut proposer d'estimer c_k en donnant des valeurs systématiques aux probabilités p_{jk} puisque ces paramètres interviennent avant l'imposition de la mesure. L'approche adoptée dans ce travail est un prolongement en dimension trois du travail de Tanner; elle s'effectue en deux étapes et permet d'éviter le choix systématique pour les probabilités p_{jk} . On propose, dans une première étape (paragraphe 2.3.2.), d'estimer simultanément les p_{jk} et l'effet de la mesure et de donner la précision de ces estimateurs sous l'hypothèse $H_0 : \theta_1 = \theta_2 = \dots = \theta_s$ (similarité de l'effet de la mesure aux différents sites). Une statistique est alors proposée pour éprouver cette hypothèse (chapitre 3). Lorsqu'elle est rejetée, on garde toujours les mêmes estimations mais on corrige la précision sur l'effet de la mesure en incorporant un coefficient d'hétérogénéité à sa variance asymptotique.

2.3.- ANALYSE DE L'EFFET DE LA MESURE SOUS L'HYPOTHÈSE H_0

2.3.1.- Résultats préliminaires

On considère dans ce paragraphe l'analyse de l'effet de la mesure sous l'hypothèse $H_0 : \theta_1 = \theta_2 = \dots = \theta_s$ (égalité de l'effet aux différents sites). On note $\theta (0 < \theta)$ cet effet commun et $P^T = (P_1^T, P_2^T, \dots, P_s^T)$ le vecteur à sr dimensions de composantes les $P_k (k = 1, 2, \dots, s)$. Soit $\Theta = (\theta, P)^T$ l'élément de $\mathbb{R}^{1+sr}$ ainsi défini. On pose alors :

$$L(\Theta) = \sum_{k=1}^s f_k((X_{1k}, X_{2k}); \Theta)$$

la Log-vraisemblance relative à la suite de vecteurs aléatoires (X_{1k}, X_{2k}) , $(k = 1, 2, \dots, s)$ sous l'hypothèse H_0 . Θ est supposé de plus être un point de $\mathbb{R}^{1+sr}$ dont toutes les composantes sont strictement positives ($\Theta > 0$) vérifiant des contraintes du type

$$h_k(\Theta) = 0, \quad (k = 1, 2, \dots, s).$$

On note également pour $k = 1, \dots, s$, $\mathcal{O}_k = \{P_k \in \mathbb{R}^r, 0 < p_{jk}, \sum_{j=1}^r p_{jk} = 1\}$; $\mathcal{O} = \mathcal{O}_1 \times \mathcal{O}_2 \times \dots \times \mathcal{O}_s$ et $\mathcal{V} = \mathbb{R}^{*+} \times \mathcal{O}$ le sous-ensemble de $\mathbb{R}^{1+sr}$ ainsi défini supposé contenir Θ_0 la vraie valeur inconnue du paramètre. Le premier problème à résoudre est donc l'estimation de Θ .

La recherche de l'estimateur du maximum de vraisemblance de Θ est équivalente à la recherche d'un estimateur du maximum de vraisemblance contraint (EMVC) $\hat{\Theta}$ de Θ c'est-à-dire d'un élément de $\mathcal{V}$. On donne ci-après quelques résultats relatifs à l'existence et à la distribution asymptotique d'un EMVC. Gouriéroux et Monfort (1989) présentent diverses méthodes d'estimation sous contraintes. Aitchison et Silvey (1958), Silvey (1959) démontrent sous certaines conditions l'existence et la normalité asymptotique de $\hat{\Theta}$ dans le cas d'observation i.i.d. (indépendantes identiquement distribuées). Lee (1979b) applique les résultats de Silvey à l'analyse de matrice de variances-covariances sous contraintes. Crowder (1984) discute du cas des observations non i.i.d. avec contraintes.

On suppose que $f_k((X_{1k}, X_{2k}); \Theta)$ et $h_k, (k = 1, 2, \dots, s)$ sont continûment différentiables dans un voisinage $\mathcal{V}_0$ de Θ_0 .

Si on note λ le vecteur des s multiplicateurs de Lagrange et $\mathcal{A} = ((\frac{\partial h_1}{\partial \Theta}, \frac{\partial h_2}{\partial \Theta}, \dots, \frac{\partial h_s}{\partial \Theta}))^T$ la matrice jacobienne de taille $(s, 1 + rs)$, associée aux contraintes, évaluée au point $\Theta = \Theta_0$ et supposé de rang plein, on sait alors que l'EMVC $\hat{\Theta}$ s'il existe vérifie les conditions du premier ordre suivantes :

$$(2.3.1) \quad \ell(\hat{\Theta}) + \mathcal{A}^T(\hat{\Theta}) \cdot \lambda = 0$$

$$(2.3.2) \quad h_k(\hat{\Theta}) = 0 \quad ; \quad k = 1, 2, \dots, s$$

avec $\hat{\Theta} \in \mathcal{V}$ et $\ell(\hat{\Theta}) = (\partial L / \partial \Theta)_{(\Theta=\hat{\Theta})}$. On pose alors

$$J = -E \left[\left(\frac{\partial^2 L}{\partial \Theta \partial \Theta} \right)_{(\Theta=\Theta_0)} \right],$$

la matrice d'information de Fisher associée à la suite de vecteurs aléatoires $(X_{11}, X_{21}), (X_{12}, X_{22}), \dots, (X_{1s}, X_{2s})$.

Théorème 2.3.1 (Aitchison et Silvey).- *Sous certaines conditions de régularité, il existe des variables aléatoires $\hat{\Theta}$ et $\hat{\lambda}$ telles que :*

- a) $\hat{\Theta} - \Theta_0 = o_p(1)$
- b) *avec une probabilité tendant vers 1 quand $N \rightarrow \infty$, $\hat{\Theta}$ et $\hat{\lambda}$ sont solutions des équations (2.3.1) et (2.3.2) telles que $\hat{\Theta}$ appartiennent à un certain voisinage $\mathcal{V}_N$ de Θ_0 (avec $\mathcal{V}_N \rightarrow \{\Theta_0\}$ quand $N \rightarrow \infty$).*
- c) *avec une probabilité tendant vers 1 quand $N \rightarrow \infty$, $\hat{\Theta}$ maximise $L(\Theta)$ dans $\mathcal{V}_N$.*
- d) *les vecteurs $N^{1/2}(\hat{\Theta} - \Theta_0)$ et $N^{1/2}\hat{\lambda}$ sont asymptotiquement normalement distribués de moyenne nulle et de variances-covariances*

$$\begin{pmatrix} Q(\Theta_0) & O \\ O & -R(\Theta_0) \end{pmatrix} \text{ où } \begin{pmatrix} Q(\Theta_0) & S^T(\Theta_0) \\ S(\Theta_0) & R(\Theta_0) \end{pmatrix} = \begin{pmatrix} J & -\mathcal{A}^T \\ -\mathcal{A} & O \end{pmatrix}^{-1}$$

Remarque 2.3.1.- Si la matrice $J(\Theta_0)$ n'est pas définie positive, Silvey (section 6) propose une correction de la matrice d'information. Le point (d) du théorème 2.3.1. est alors légèrement modifié. Si de plus les $x_{..k}$ sont distincts, il faut conditionner la matrice d'information utilisée (Silvey, section 7) . Le théorème 2.3.1. est obtenu sous l'hypothèse que les vecteurs aléatoires $X_1, X_2, \dots, X_s$ sont i.i.d. Lorsque les vecteurs aléatoires sont indépendants mais non identiquement distribués, Crowder propose une correction de la matrice $J(\Theta_0)$ en faisant des hypothèses sur $\mathcal{A}$. Il prolonge alors le théorème 2.3.1.

2.3.2.- Estimation simultanée de l'effet et des p_{jk}

La fonction de vraisemblance de la suite de vecteurs aléatoires $(X_{11}, X_{22}), \dots, (X_{1s}, X_{2s})$ sous l'hypothèse H_0 relativement au modèle 1 s'écrit :

$$\mathcal{L}(\Theta) = \prod_{k=1}^s \left(\frac{x_{..k}!}{\prod_{j=1}^r x_{1jk}! x_{2jk}!} \right) \left\{ \prod_{j=1}^r p_{jk}^{x_{.jk}} \frac{(\theta c_{jk})^{x_{2jk}}}{(1 + \theta c_k)^{x_{.jk}}} \right\}$$

avec $x_{.jk} = x_{1jk} + x_{2jk}$.

La log-vraisemblance est donnée à une constante additive près par :

$$L(\Theta) = \sum_{k=1}^s \sum_{j=1}^r \{x_{.jk} \log p_{jk} + x_{2jk} \log c_{jk} + x_{2jk} \log \theta - x_{.jk} \log(1 + \theta c_k)\}.$$

En dérivant la fonction $F(\Theta) = -L(\Theta)$ par rapport à Θ et en annulant les dérivées partielles, on montre que l'EMVC de Θ vérifie le système d'équations non linéaires :

$$\left\{ \begin{array}{l} (2.3.1) \quad \sum_{k=1}^s \sum_{j=1}^r \left\{ \frac{x_{2jk} - \hat{\theta} \hat{c}_k x_{1jk}}{1 + \hat{\theta} \hat{c}_k} \right\} = 0 \\ (2.3.2) \quad \frac{x_{\cdot jk}}{\hat{p}_{jk}} - \frac{\hat{\theta} c_{jk} x_{\cdot \cdot k}}{1 + \hat{\theta} \hat{c}_k} = 0 ; (k = 1, 2, \dots, s; j = 1, 2, \dots, r) \\ \hat{\Theta} \in \mathcal{V} \end{array} \right.$$

avec $\hat{c}_k = \sum_{j=1}^r c_{jk} \hat{p}_{jk}$. La détermination de la fonction de vraisemblance et des équations de vraisemblance dans le cadre du modèle 2 est donnée en annexe. L'équation (2.3.1) est obtenue en dérivant F par rapport au paramètre commun θ . Elle est semblable à celle proposée par Tanner. Le système (2.3.2) est obtenu en dérivant F par rapport aux p_{jk} et est utilisé pour estimer convenablement les coefficients de contrôle c_k via les estimations des p_{jk} .

Différentes approches de résolution du système ci-dessus sont possibles [Gouriéroux et Monfort (1989), Lee (1979a)]. Les algorithmes que nous proposons sont basés sur la méthode du gradient projeté [voir Minoux p. 197]. L'un des avantages de ces algorithmes est qu'ils évitent l'inversion de la matrice d'information à chaque étape. De surcroît, la forme de $\mathcal{A}$ permet l'existence d'un point $\hat{\Theta} \in \mathcal{V}$ vérifiant les équations de vraisemblance. On a le résultat suivant:

Théorème 2.3.2. - *On suppose la donnée d'un point $\hat{\Theta}_1 \in \mathcal{V}$ solution des systèmes (2.3.1) et (2.3.2). Alors la solution optimale du problème*

$$\left\{ \begin{array}{l} \text{Minimiser } \left(\frac{\partial F}{\partial \Theta}(\hat{\Theta}_1) \right)^T \cdot \vec{d} \\ \text{avec :} \\ \mathcal{A} \cdot \vec{d} = 0 \\ \text{et } \vec{d} \text{ vecteur unitaire de } \mathbb{R}^{1+sr} \end{array} \right.$$

est $\vec{d} = \vec{v} / \|\vec{v}\|$ où $\vec{v}$ est la projection de $-\left(\frac{\partial F}{\partial \Theta}(\hat{\Theta}_1) \right)$ sur $\mathcal{V}$.

Démonstration : En calculant les vecteurs $\left(\frac{\partial h_k}{\partial \Theta} \right)$ ($k = 1, 2, \dots, s$), on montre que la matrice $\mathcal{A}$ de taille $(s, 1 + sr)$ est donnée par

$$\mathcal{A} = \begin{pmatrix} 0 & \mathbf{1}_r^T & O^T & \dots & O^T \\ 0 & O^T & \mathbf{1}_r^T & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & O^T \\ 0 & O^T & \dots & O^T & \mathbf{1}_r^T \end{pmatrix},$$

$\mathbf{1}_r = (1, \dots, 1)^T \in \mathbb{R}^r$. On montre alors que la matrice $\mathcal{A}\mathcal{A}^T$ de taille (s, s) est

$$\mathcal{A}\mathcal{A}^T = \begin{pmatrix} \mathbf{1}_r^T \mathbf{1}_r & O & \dots & O \\ O & \mathbf{1}_r^T \mathbf{1}_r & \ddots & \vdots \\ \vdots & \ddots & \ddots & O \\ O & \dots & O & \mathbf{1}_r^T \mathbf{1}_r \end{pmatrix} = rI_s.$$

Par conséquent, la matrice $\mathcal{A}$ est de rang plein. On applique alors le résultat donné dans (Minoux, p. 199) pour achever la démonstration.

Une fois obtenue la direction de déplacement $\vec{d} = \vec{d}/\|\vec{v}\|$, on détermine le déplacement maximal autorisé par les contraintes c'est-à-dire $\alpha_{\max}$ tel que

$$\alpha_{\max} = \text{Max}\{\alpha/\alpha \geq 0, \hat{\Theta}_1 + \alpha\vec{v} \in \mathcal{V}\}.$$

Le point $\hat{\Theta}_2$ suivant est alors sélectionné comme un point minimisant $F(\hat{\Theta}_1 + \alpha\vec{v})$ sur le segment $[0, \alpha_{\max}]$. Au point $\hat{\Theta}_2$, on devra déterminer la nouvelle direction de déplacement, et ainsi de suite. L'algorithme se poursuit donc de cette façon tant que $\vec{v} \neq \vec{O}$. Cette méthode peut être considérée comme un outil intéressant au moins pour la résolution de problèmes d'optimisation non linéaires avec contraintes linéaires. L'aspect pratique et sa mise en oeuvre sont discutés au chapitre 4.

2.3.3.- Estimation de la variance asymptotique de $\log \hat{\theta}$.

$\hat{\theta}$ étant l'estimateur du maximum de vraisemblance de θ^0 , on sait d'après les propriétés du maximum de vraisemblance que $\log \hat{\theta}$ est un estimateur du maximum de vraisemblance de $\log \theta^0$. Il suffit d'appliquer le théorème d'invariance fonctionnelle [voir Monfort p. 105] à la fonction :

$$\begin{aligned} \mathcal{V} &\longrightarrow \mathbb{R} \times S \\ g : \Theta &\longrightarrow (\log \theta, P)^T = g(\Theta) \end{aligned}$$

laissant P invariant et transformant θ en $\log \theta$ var $\log \hat{\theta}$ ($N \rightarrow \infty$) est obtenue en inversant [voir théorème 2.3.1, paragraphe 2.3.1] la matrice

$$\begin{pmatrix} \mathbb{J} & -\mathcal{A}^T \\ -\mathcal{A} & O \end{pmatrix}$$

où $\mathbb{J} = -E \left[\frac{\partial^2 L}{\partial g(\Theta) \partial g(\Theta)} \right]$ est la matrice carrée de taille $1 + sr$ évaluée au point $g(\Theta^0) = (\log \theta^0, P^0)^T$.

Le lemme suivant [voir Aitchison et Silvey, lemme 3] permet d'obtenir l'inverse de la matrice partitionnée ci-dessus :

Lemme 2.3.1. *La matrice partitionnée*

$$\begin{pmatrix} \mathbb{J} & -\mathcal{A}^T \\ -\mathcal{A} & O \end{pmatrix}$$

est inversible et son inverse

$$\begin{pmatrix} \mathbb{Q} & \mathbb{S}^T \\ \mathbb{S} & \mathbb{R} \end{pmatrix}$$

est donnée par

$$\begin{aligned} \mathbb{Q} &= \mathbb{J}^{-1} [I_{1+sr} - \mathcal{A}^T (\mathcal{A} \mathbb{J}^{-1} \mathcal{A}^T)^{-1} \mathcal{A} \mathbb{J}^{-1}] \\ \mathbb{S}^T &= -\mathbb{J} \mathcal{A}^T (\mathcal{A} \mathbb{J}^{-1} \mathcal{A}^T)^{-1} \\ \mathbb{R} &= -(\mathcal{A} \mathbb{J}^{-1} \mathcal{A}^T)^{-1}. \end{aligned}$$

Preuve : La démonstration repose sur le fait que $\mathbb{J}$ est définie positive et $\mathcal{A}$ de rang plein.

La variance asymptotique de $\log \hat{\theta}$ est alors donnée par l'élément de la première ligne et première colonne de la matrice

$$\mathbb{Q} = \mathbb{J}^{-1} [I_{1+sr} - \mathcal{A}^T (\mathcal{A} \mathbb{J}^{-1} \mathcal{A}^T)^{-1} \mathcal{A} \mathbb{J}^{-1}].$$

En calculant explicitement la matrice $\mathbb{J}$, on montre que

$$\mathbb{J}^{-1} = \begin{pmatrix} (\mathbb{J}_{\theta\theta} - \mathbb{J}_{P\theta}^T \mathbb{J}_{PP}^{-1} \mathbb{J}_{P\theta})^{-1} & -\mathbb{J}_{\theta\theta}^{-1} \mathbb{J}_{P\theta}^T (\mathbb{J}_{PP} - \mathbb{J}_{P\theta} \mathbb{J}_{\theta\theta}^{-1} \mathbb{J}_{P\theta}^T)^{-1} \\ -\mathbb{J}_{PP}^{-1} \mathbb{J}_{P\theta} (\mathbb{J}_{\theta\theta} - \mathbb{J}_{P\theta}^T \mathbb{J}_{PP}^{-1} \mathbb{J}_{P\theta})^{-1} & (\mathbb{J}_{PP} - \mathbb{J}_{P\theta} \mathbb{J}_{\theta\theta}^{-1} \mathbb{J}_{P\theta}^T)^{-1} \end{pmatrix}$$

où $\mathbb{J}_{\theta\theta} = \sum_{k=1}^s \frac{\theta c_k x_{..k}}{(1+\theta c_k)^2}$, $\mathbb{J}_{P\theta}$ est l'élément de $\mathbb{R}^{sr}$ donné par

$$\mathbb{J}_{P\theta} = \begin{pmatrix} \mathbb{J}_{P_1\theta} \\ \mathbb{J}_{P_2\theta} \\ \vdots \\ \mathbb{J}_{P_s\theta} \end{pmatrix}$$

avec $\mathbb{J}_{P_k\theta} = \frac{\theta x_{..k}}{(1+\theta c_k)^2} (c_{1k}, c_{2k}, \dots, c_{rk})^T \in \mathbb{R}^r$ et où $\mathbb{J}_{PP}$ est la matrice carrée de taille sr définie par

$$\mathbb{J}_{PP} = \begin{pmatrix} \mathbb{J}_{P_1P_1} & O & \dots & O \\ O & \mathbb{J}_{P_2P_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & O \\ O & \dots & O & \mathbb{J}_{P_sP_s} \end{pmatrix}, \quad \text{avec}$$

$$\mathbb{J}_{P_k P_k} = \frac{x_{..k}}{1 + \theta c_k} \begin{pmatrix} \frac{1 + \theta c_{1k}}{p_{1k}} - \frac{\theta^2 c_{1k}^2}{1 + \theta c_k} & -\frac{\theta^2 c_{1k} c_{2k}}{1 + \theta c_k} & \cdots & -\frac{\theta^2 c_{1k} c_{rk}}{1 + \theta c_k} \\ -\frac{\theta^2 c_{2k} c_{1k}}{1 + \theta c_k} & \frac{1 + \theta c_{2k}}{p_{2k}} - \frac{\theta^2 c_{2k}^2}{1 + \theta c_k} & \ddots & \vdots \\ \vdots & \ddots & \ddots & -\frac{\theta^2 c_{r-1k} c_{rk}}{1 + \theta c_k} \\ -\frac{\theta^2 c_{rk} c_{1k}}{1 + \theta c_k} & \cdots & -\frac{\theta^2 c_{rk} c_{r-1k}}{1 + \theta c_k} & \frac{1 + \theta c_{rk}}{p_{rk}} - \frac{\theta^2 c_{rk}^2}{1 + \theta c_k} \end{pmatrix}$$

($k = 1, 2, \dots, s$).

Lemme 2.3.2. *On montre que*

$$\text{var log } \hat{\theta}_{(N \rightarrow \infty)} = \frac{1 - \beta \left[\sum_{k=1}^s \left(\frac{\theta c_k x_{..k}}{(1 + \theta c_k)^2} - \mathbb{J}_{P_k \theta}^T \mathbb{J}_{P_k P_k}^{-1} \mathbb{J}_{P_k \theta} \right) \right]}{\sum_{k=1}^s \left\{ \frac{\theta c_k x_{..k}}{(1 + \theta c_k)^2} - \mathbb{J}_{P_k \theta}^T \mathbb{J}_{P_k P_k}^{-1} \mathbb{J}_{P_k \theta} \right\}}$$

où β est l'élément de la première ligne et première colonne de la matrice :

$$\mathbb{J}^{-1} \mathcal{A}^T [\mathcal{A} \mathbb{J}^{-1} \mathcal{A}^T]^{-1} \mathcal{A} \mathbb{J}^{-1}.$$

Preuve : D'après le lemme 2.3.1, $\text{var log } \hat{\theta}$ est l'élément de la première ligne et première colonne ($N \rightarrow \infty$) de la matrice

$$\mathbb{Q} = \mathbb{J}^{-1} - \mathbb{J}^{-1} \mathcal{A}^T [\mathcal{A} \mathbb{J}^{-1} \mathcal{A}]^{-1} \mathcal{A} \mathbb{J}^{-1}.$$

Quelques manipulations usuelles sur l'inversion de matrices partitionnées montre que l'élément de la première ligne et de la première colonne de $\mathbb{J}^{-1}$ est donné par

$$\alpha_{\theta\theta} = \frac{1}{\mathbb{J}_{\theta\theta} - \sum_{k=1}^s \mathbb{J}_{P_k \theta}^T \mathbb{J}_{P_k P_k}^{-1} \mathbb{J}_{P_k \theta}}.$$

En notant β l'élément de la première ligne et première colonne de la matrice

$$\mathbb{J}^{-1} \mathcal{A}^T [\mathcal{A} \mathbb{J}^{-1} \mathcal{A}]^{-1} \mathcal{A} \mathbb{J}^{-1}$$

et en remplaçant $\mathbb{J}_{\theta\theta}$ par son expression, on montre que

$$\text{var log } \hat{\theta}_{(N \rightarrow \infty)} = \alpha_{\theta\theta} - \beta,$$

d'où le résultat du lemme 2.3.2. ■

Remarque 2.3.2.- La variance asymptotique de $\log \hat{\theta}$ ne peut donc être obtenue que de façon numérique. Lorsque $r = 1$, les paramètres p_{jk} sont identiques

à 1. Les modèles 1 et 2 coïncident avec celui de Tanner et la matrice $\mathcal{A}$ est identiquement nulle. La matrice $\mathbb{J}$ est alors identique au réel

$$\mathbb{J}_{\theta\theta} = \sum_{k=1}^s \frac{\theta c_k x_{..k}}{(1 + \theta c_k)^2}.$$

Le résultat du lemme 2.3.2 est identique à celui de Tanner : c'est-à-dire

$$\text{var} \log \hat{\theta}_{(N \rightarrow \infty)} = \frac{1}{\sum_{k=1}^s \frac{\theta c_k x_{..k}}{(1 + \theta c_k)^2}}.$$

On note dans la suite

$$\rho = \text{var} \log \hat{\theta}_{(N \rightarrow \infty)}$$

la variance obtenue au lemme 2.3.2.

2.3.4.- Test de l'hypothèse $H_0 : \theta_1 = \theta_2 = \dots = \theta_S$ et approximation de la variabilité de l'effet aux différents sites.

La suite $\theta_1^*, \theta_2^*, \dots, \theta_S^*$ des effets apparents est une estimation de l'effet aux différents sites dont on dispose en pratique. Si l'on admet l'hypothèse $H_0 : \theta_1 = \theta_2 = \dots = \theta_S$ (égalité des effets) alors la variation pouvant existée entre ces effets apparents ne peut parvenir, vue la définition des θ_k^* , que de l'hypothèse de distribution multinomiale effectuée sur les vecteurs aléatoires (X_{1k}, X_{2k}) . On propose alors, pour éprouver H_0 , de comparer la statistique

$$\hat{A}_{s,r}^2 = \sum_{k=1}^s \sum_{j=1}^r \sum_{t=1}^2 \frac{(X_{tjk} - x_{..k} q_{tjk}(\hat{\theta}, \hat{P}_k))^2}{x_{..k} q_{tjk}(\hat{\theta}, \hat{P}_k)}$$

à la valeur α_ν que dépasse une variable aléatoire de loi de χ_ν^2 avec une probabilité α , où $\hat{\theta}$ et $\hat{P}_k$ représentent, sous H_0 , les estimateurs de θ et de P_k déduits de l'EMV de θ basé sur l'ensemble de toutes les données.

Si la valeur calculée $\hat{A}_{s,r}^2$ n'est pas suffisamment grande, on peut penser à une non-variation dans l'effet de la mesure aux différents sites. Dans ces conditions, on établit l'inefficacité ($\theta = 1$) de la mesure en utilisant la variable aléatoire (v.a.)

$$(2.3.3) \quad z = \frac{\log \hat{\theta}}{\rho^{1/2}}$$

dont la loi est sous l'hypothèse $\theta = 1$, asymptotiquement $\mathcal{N}(0;1)$ d'après les propriétés des EMV (théorème 2.3.1). On rejette, l'hypothèse $\theta = 1$ au niveau de signification 5%, si la valeur observée est en valeur absolue supérieure à 1,96.

Si par contre la valeur calculée $\hat{A}_{s,r}^2$ est trop grande, on peut conclure qu'il y a vraisemblablement une différence dans l'effet de la mesure aux différents sites, on rejette alors H_0 . Une telle prise de décision exige la connaissance exacte de ν (chapitre 3). Notons que dans le cas du rejet de cette hypothèse H_0 , on désirera souvent apprécier la variabilité entre les θ_k ($k = 1, 2, \dots, s$) qui ne sont pas directement observables. On propose ci-après une méthode d'approximation basée sur la statistique $\hat{A}_{s,r}^2$ pour estimer la variabilité de l'effet aux différents sites (proposition 2.3.1). On utilise ensuite cette approximation (voir section 2.4) pour corriger la variance asymptotique ($N \rightarrow \infty$) de $\log \hat{\theta}$ lorsque H_0 est rejetée.

En plus de l'hypothèse de distribution faite sur les vecteurs aléatoires (X_{1k}, X_{2k}) , on suppose dans la suite que $\theta_1, \theta_2, \dots, \theta_s$ est une suite de v.a. provenant d'une distribution dont la loi a pour espérance mathématique θ et pour variance σ^2 . On suppose de plus que les paramètres θ et p_{jk} sont connus et donnés par les équations (2.3.1) et (2.3.2). On se propose alors d'estimer σ^2 . On note E^X [resp. $E^{\mathcal{P}}$] l'espérance mathématique par rapport à la distribution des $\vec{v}.a. (X_{1k}, X_{2k})$ [resp. θ_k] $k = 1, 2, \dots, s$.

Proposition 2.3.1. *On suppose que N et $x_{..k} = x_{..k}(N)$ tendent vers ∞ de telle sorte que $\lim_{N \rightarrow \infty} \frac{x_{..k}(N)}{N} = \tau_k < \infty$ et que*

(c1) $\hat{p}_{jk} - p_{jk} = o_p(1)$ ($N \rightarrow \infty$) $\forall k, j$ fixé où $\hat{p}_{jk}$ est l'estimateur donné par l'équation (2.3.1).

(c2) $\sum_{m=1}^L \omega_{k,m+2} E^{\mathcal{P}}(|\theta_k - \theta|^{m+2}) = O(1)$, ($N \rightarrow \infty$) pour tout entier $L \geq 2$ où $\omega_{k,m+2}$ sont des coefficients réels fonction de θ et p_{jk}

(c3) $\hat{\theta} - \theta = o_p(1)$ ($N \rightarrow \infty$)

alors un estimateur asymptotique de σ^2 est :

$$\frac{A_{s,r}^2 - (sr - 1)}{\sum_{k=1}^s \left[\frac{c_k(x_{..k} + \theta c_k - 2)}{\theta(1 + \theta c_k)^2} \right] + (r - 1) \sum_{k=1}^s \left[\frac{c_k}{\theta} \frac{\theta c_k - 1}{(1 + \theta c_k)^2} \right]}.$$

Preuve : On suppose que $\forall \varepsilon > 0$ donné, $\exists N_\varepsilon \in \mathbb{N}^*$ tel que $\forall N \geq N_\varepsilon$ tel que $\forall N \geq N_\varepsilon$ $(\hat{\theta}, \hat{P}^T)^T \in \mathcal{V}_N$, où $\mathcal{V}_N$ est un voisinage ouvert de l'espérance mathématique θ et des probabilités p_{jk} .

On pose $U_{N_\varepsilon} = \mathcal{V}_{N_\varepsilon} \cap \mathbb{R}_+^*$ (la restriction de $\mathcal{V}_{N_\varepsilon}$ à $\mathbb{R}_+^*$) voisinage ouvert de

l'espace mathématiques θ . On considère alors la fonction aléatoire définie par :

$$U_{N_\varepsilon} \rightarrow \mathbb{R}_+^*$$

$$z \rightarrow \hat{A}_{s,r}^2(z) = \sum_{k=1}^r \sum_{j=1}^r \sum_{t=1}^2 \frac{(X_{tjk} - \hat{X}_{tjk}(z))^2}{\hat{X}_{tjk}(z)}$$

où $\hat{X}_{tjk}(z) = x_{..k} q_{tjk}(z, \hat{P}_k)$ et où X_{tjk} suit une distribution binomiale de paramètres $x_{..k}$ et $q_{tjk}(\theta_k, P_k)$ (voir modèles), θ_k étant par hypothèse une v.a. d'espérance mathématique θ et de variance σ^2 . Il faut alors évaluer l'espérance mathématique de $\hat{A}_{s,r}^2(z)$ en utilisant l'approche proposée par Tanner (paragraphe 5). Par définition, on a :

$$E^X(\hat{A}_{s,r}^2(z)) = \sum_{k=1}^s \sum_{j=1}^r \sum_{t=1}^2 E^X \left[\frac{(X_{tjk} - \hat{X}_{tjk}(z))^2}{\hat{X}_{tjk}(z)} \right].$$

En développant les termes entre crochets et en utilisant le modèle 1, on montre que

$$\begin{aligned} \sum_{j=1}^r E^X \left[\frac{(X_{1jk} - \hat{X}_{1jk}(z))^2}{\hat{X}_{1jk}(z)} \right] &= \frac{x_{..k}(1 + z\hat{c}_k)}{(1 + \theta_k c_k)^2} \left(\sum_{j=1}^r \frac{p_{jk}^2}{\hat{p}_{jk}} \right) \\ &\quad - \frac{1 + z\hat{c}_k}{(1 + \theta_k c_k)^2} \left(\sum_{j=1}^r \frac{p_{jk}^2}{\hat{p}_{jk}} \right) + \frac{1 + z\hat{c}_k}{1 + \theta_k c_k} \left(\sum_{j=1}^r \frac{p_{jk}^2}{\hat{p}_{jk}} \right) - 2 \frac{x_{..k}}{1 + \theta_k c_k} + \frac{x_{..k}}{1 + z\hat{c}_k}, \\ \sum_{j=1}^r E^X \left[\frac{(X_{2jk} - \hat{X}_{2jk}(z))^2}{\hat{X}_{2jk}(z)} \right] &= \frac{x_{..k}(1 + z\hat{c}_k)\theta_k^2}{z(1 + \theta_k c_k)^2} \left(\sum_{j=1}^r \frac{c_{jk}p_{jk}^2}{\hat{p}_{jk}} \right) \\ &\quad + \frac{(1 + z\hat{c}_k)\theta_k}{z(1 + \theta_k c_k)} \left(\sum_{j=1}^r \frac{p_{jk}}{\hat{p}_{jk}} \right) - \frac{1 + z\hat{c}_k}{(1 + \theta_k c_k)^2} \frac{\theta_k^2}{z} \left(\sum_{j=1}^r \frac{c_{jk}p_{jk}^2}{\hat{p}_{jk}} \right) - 2 \frac{x_{..k}\theta_k c_k}{1 + \theta_k c_k} \\ &\quad + \frac{x_{..k}z}{1 + z\hat{c}_k} \left(\sum_{j=1}^r c_{jk}\hat{p}_{jk} \right). \end{aligned}$$

La condition (c1) permet d'écrire que $\hat{c}_k = c_k + o_p(1)$ et que :

$$\begin{aligned} \sum_{j=1}^r \frac{p_{jk}^2}{\hat{p}_{jk}} &= \sum_{j=1}^r p_{jk}[1 + o_p(1)], \\ \sum_{j=1}^r \frac{p_{jk}}{\hat{p}_{jk}} &= \sum_{j=1}^r [1 + o_p(1)], \\ \sum_{j=1}^r \frac{c_{jk}p_{jk}^2}{\hat{p}_{jk}} &= \sum_{j=1}^r c_{jk}p_{jk}[1 + o_p(1)]. \end{aligned}$$

Par substitution, on obtient :

$$\begin{aligned}
\sum_{j=1}^r E^X \left[\frac{(X_{1jk} - \hat{X}_{1jk}(z))^2}{\hat{X}_{1jk}(z)} \right] &= \frac{x_{..k}(1 + zc_k)}{(1 + \theta_k c_k)^2} \\
&\quad + r \frac{(1 + zc_k)c_k}{(1 + \theta_k c_k)} - \frac{(1 + zc_k)}{(1 + \theta_k c_k)^2} - \frac{2x_{..k}}{1 + \theta_k c_k} \\
&\quad + \frac{x_{..k}}{1 + zc_k} + o_p(1), \\
\\
\sum_{j=1}^r E^X \left[\frac{(X_{2jk} - \hat{X}_{2jk}(z))^2}{\hat{X}_{2jk}(z)} \right] &= \frac{x_{..k}(1 + zc_k)\theta_k^2 c_k}{z(1 + \theta_k c_k)^2} \\
&\quad + r \frac{(1 + zc_k)\theta_k c_k}{z(1 + \theta_k c_k)} - \frac{(1 + zc_k)\theta_k^2 c_k}{z(1 + \theta_k c_k)^2} - \frac{2x_{..k}\theta_k c_k}{1 + \theta_k c_k} \\
&\quad + \frac{x_{..k} z c_k}{1 + zc_k} + o_p(1),
\end{aligned}$$

En sommant ces deux dernières égalités et en réduisant au même dénominateur, on montre que

$$\begin{aligned}
\sum_{j=1}^r \sum_{t=1}^2 E^X \left[\frac{(X_{tjk} - \hat{X}_{tjk}(z))^2}{\hat{X}_{tjk}(z)} \right] &= r - 1 + r \frac{\theta_k(1 + zc_k)^2}{z(1 + \theta_k c_k)^2} \\
&\quad + \frac{c_k(x_{..k} + r - 1)(\theta_k - z)^2}{z(1 + \theta_k c_k)^2} + o_p(1).
\end{aligned}$$

On en déduit alors que

$$E^X(\hat{A}_{s,r}^2(z)) = s(r-1) + r \sum_{k=1}^s \frac{\theta_k(1 + zc_k)^2}{z(1 + \theta_k c_k)^2} + \sum_{k=1}^s \frac{c_k(x_{..k} + r - 1)(\theta_k - z)^2}{z(1 + \theta_k c_k)^2} + o_p(1).$$

$E^X(\hat{A}_{s,r}^2(z))$ est donc la somme de deux variables aléatoires Y_1 et Y_2 telles que

$$E^X(Y_1) = s(r-1) + r \sum_{k=1}^s \frac{\theta_k(1 + zc_k)^2}{z(1 + \theta_k c_k)^2} + \sum_{k=1}^s \frac{c_k(x_{..k} + r - 1)(\theta_k - z)^2}{z(1 + \theta_k c_k)^2}$$

et $Y_2 = o_p(1)$. Pour la suite de la démonstration, on assimile $E^X(\hat{A}_{s,r}^2(z))$ à sa partie principale Y_1 et on admet que $E^{\mathcal{D}}(Y_2) \rightarrow 0$ quand $N \rightarrow \infty$. En posant $u_k =$

$\theta_k - z$ et en développant les expressions $\theta_k/[(1 + \theta_k c_k)^2]$ et $(\theta_k - z)^2/[(1 + \theta_k c_k)^2]$ en puissance de u_k , on a :

$$\frac{(1 + z c_k)^2 \theta_k}{z(1 + \theta_k c_k)^2} = 1 + \frac{1 - z c_k}{z(1 + z c_k)} u_k + \frac{c_k^2 z - 2 c_k}{z(1 + z c_k)^2} u_k^2$$

$$+ \sum_{m \geq 1} \left\{ \frac{(-1)^{m+1} (m+2) c_k^{m+2}}{z(1 + z c_k)^{m+1}} + \frac{(-1)^{m+2} (m+3) c_k^{m+2}}{(1 + z c_k)^{m+2}} \right\} u_k^{m+2},$$

$$\frac{c_k(x_{..k} + r - 1)(\theta_k - z)^2}{z(1 + \theta_k c_k)^2} = \frac{c_k(x_{..k} + r - 1)}{z(1 + z c_k)^2} u_k^2$$

$$+ (x_{..k} + r - 1) \sum_{m \geq 1} \frac{(-1)^{m+1} (m+1) c_k^{m+1}}{z(1 + z c_k)^{m+2}} u_k^{m+2}.$$

En utilisant la condtion (c2), on a

$$E^{\mathcal{D}} \left[\frac{c_k(x_{..k} + r - 1)(\theta_k - z)^2}{z(1 + \theta_k c_k)^2} \right] = \frac{c_k(x_{..k} + r - 1)}{z(1 + c_k)^2} E^{\mathcal{D}}(u_k^2) + O(x_{..k}),$$

$$E^{\mathcal{D}} \left[r \frac{(1 + z c_k)^2 \theta_k}{z(1 + \theta_k c_k)^2} \right] = r + r \frac{(1 - z c_k)}{z(1 + z c_k)} E^{\mathcal{D}}(u_k)$$

$$+ r \left(\frac{z c_k^2 - 2 c_k}{z(1 + z c_k)^2} \right) E^{\mathcal{D}}(u_k^2) + O(x_{..k}).$$

$$E[\hat{A}_{s,r}^2(z)] = E^{\mathcal{D}} E^X[\hat{A}_{s,r}^2(z)]$$

$$= s(r - 1) + r \sum_{k=1}^s E^{\mathcal{D}} \left[\frac{\theta_k(1 + z c_k)^2}{z(1 + \theta_k c_k)^2} \right]$$

$$+ \sum_{k=1}^s E^{\mathcal{D}} \left[\frac{c_k(x_{..k} + r - 1)(\theta_k - z)^2}{z(1 + \theta_k c_k)^2} \right] + O(N)$$

$$E[\hat{A}_{s,r}^2(z)] = s(2r - 1) + r \sum_{k=1}^s \frac{1 - z c_k}{z(1 + z c_k)} E^{\mathcal{D}}(u_k)$$

$$+ \sum_{k=1}^s \frac{c_k[x_{..k} + z c_k - 2 + (r - 1)(z c_k - 1)]}{z(1 + z c_k)^2} E^{\mathcal{D}}(u_k^2) + O(N).$$

Pour $N \rightarrow \infty$ et $z = \theta$, on obtient sous la condition (c3) que

$$E[\hat{A}_{s,r}^2] = E^{\mathcal{D}} E^X E[\hat{A}_{s,r}^2(z)]$$

$$= s(2r - 1) + \sigma^2 \left[\sum_{k=1}^s \left\{ \frac{c_k(x_{..k} + \theta c_k - 2)}{\theta(1 + \theta c_k)^2} \right\} + (r - 1) \sum_{k=1}^s \frac{c_k(\theta c_k - 1)}{\theta(1 + \theta c_k)^2} \right]$$

$$+ O(N)$$

car pour $z = \theta$ on a $E^{\mathcal{D}}(u_k) = 0$ et $E^{\mathcal{D}}(u_k^2) = \sigma^2$. Si l'on pense que $\hat{A}_{s,r}^2$ peut donner une estimation de $E(\hat{A}_{s,r}^2)$ alors σ^2 peut être approchée par

$$\frac{\hat{A}_{s,r}^2 - s(2r - 1)}{\sum_{j=1}^s \frac{c_k(x_{..k} + \theta c_k - 2)}{\theta(1 + \theta c_k)^2} + (r - 1) \sum_{k=1}^s \frac{c_k(\theta c_k - 1)}{\theta(1 + \theta c_k k)^2}}.$$

En fait si σ^2 n'est pas grande, on pense que les θ_k ne sont pas très loin de θ . On montre alors au chapitre 3 que

$$sr - 1 \leq E(\hat{A}_{s,r}^2) \leq s(2r - 1).$$

Ce qui fait que l'estimation de σ^2 ci-dessus est majorée par

$$\frac{\hat{A}_{s,r}^2 - (sr - 1)}{\sum_{k=1}^s \frac{c_k(x_{..k} + \theta c_k - 2)}{\theta(1 + \theta c_k)^2} + (r - 1) \sum_{k=1}^s \frac{c_k(\theta c_k - 1)}{\theta(1 + \theta c_k)^2}}.$$

Comme on veut évaluer la dispersion des θ_k , σ^2 ne vaut pas plus que

$$\frac{\hat{A}_{s,r}^2 - (sr - 1)}{\sum_{k=1}^s \frac{c_k(x_{..k} + \theta c_k - 2)}{\theta(1 + \theta c_k)^2} + (r - 1) \sum_{k=1}^s \frac{c_k(\theta c_k - 1)}{\theta(1 + \theta c_k)^2}},$$

d'où le résultat proposé.

2.4.- ANALYSE EN CAS D'HETEGROGENEITE DE L'EFFET

2.4.1. Introduction.

Si la similarité de l'effet, en tous les sites, de la mesure de sécurité routière n'est pas établie de façon significative, l'analyse statistique de l'efficacité à l'aide de tableaux "avant-après" $2 \times s \times r$ devient extrêmement complexe. Cette complexité provient non seulement de la structure tridimensionnelle du tableau d'analyse mais surtout du fait que l'effet d'une mesure en un site donné n'est pas directement observable. De surcroît, les seuls estimateurs de l'effet aux différents sites dont on dispose réellement (l'effet apprend θ_k^* en est une estimation) sont généralement fonctions aléatoires non linéaires des $\vec{v}.a.$ (X_{1k}, X_{2k}) . On proposera une méthode d'analyse consistant à construire une statistique du type (2.3.3) [voir paragraphe 2.3.4] en intégrant un coefficient (qualifié d'hétérogénéité) à ρ .

On suppose, en plus de l'hypothèse de distribution faite sur les $\vec{v}.a.$ (X_{1k}, X_{2k}) , que $\theta_1, \theta_2, \dots, \theta_s$ est une suite de v.a. provenant d'une distribution dont la loi a pour espérance mathématique θ inconnue et pour variance σ^2 inconnue. La procédure statistique la plus satisfaisante consiste à admettre que la loi

de la distribution des $\theta_1, \theta_2, \dots, \theta_s$ appartient à une certaine famille et utiliser les méthodes statistiques usuelles d'estimation pour estimer θ et σ^2 . Malheureusement, l'on ne dispose d'aucune observation de cet échantillon en pratique. Par conséquent, une méthode d'approximation est proposée pour obtenir un estimateur asymptotique $(N \rightarrow \infty)\hat{\theta}$ de θ et une évaluation de $\text{var log } \hat{\theta}$. Cette méthode d'approximation utilise la distribution multinomiale des *v.a.* (X_{1k}, X_{2k}) dont les probabilités de classe $q_{tjk}(\theta_k, P_k)$ sont cette fois-ci des *v.a.* De telles distributions sont généralement qualifiées [voir Morton, 1991] d'extra-multinomiales. La méthode d'approximation est faite en deux étapes. On estime premièrement l'effet de la mesure θ et les p_{jk} . En supposant cet effet connu, on propose une approximation asymptotique de $\text{var log } \hat{\theta}$ via l'estimation de σ^2 donnée par la proposition 2.3.1.

2.4.2. Estimation de l'effet de la mesure.

On suppose comme au paragraphe 2.3 que θ et p_{jk} satisfont, dans le cadre du modèle 1, aux équations :

$$\left\{ \begin{array}{l} (2.3.1) \quad \sum_{k=1}^s \sum_{j=1}^r \left\{ \frac{x_{2jk} - \hat{\theta} \hat{c}_k x_{1jk}}{1 + \hat{\theta} \hat{c}_k} \right\} = 0 \\ (2.3.2) \quad \frac{x_{\cdot jk}}{\hat{p}_{jk}} - \frac{\hat{\theta} c_{jk} x_{\cdot \cdot k}}{1 + \hat{\theta} \hat{c}_k} = 0 ; (k = 1, 2, \dots, s; j = 1, 2, \dots, r) \\ \hat{\theta} \in \mathcal{V} \end{array} \right.$$

avec $\hat{c}_k = \sum_{j=1}^r c_{jk} \hat{p}_{jk}$. Une telle approche est raisonnable si toute fois la variation entre les θ_k est relativement faible. Une estimation de cette variation est donnée par la proposition 2.3.1.

2.4.3. Une méthode d'approximation de la variance de $\log \hat{\theta}$.

La méthode d'approximation proposée consiste à corriger ρ ($\text{var log } \hat{\theta}$ sous l'hypothèse H_0 : voir paragraphe 2.3.3). On suppose $N \rightarrow \infty$ et on considère la fonction aléatoire

$$\begin{aligned} \Psi : U_{N_\epsilon} &\longrightarrow \mathbb{R} \\ u &\longrightarrow \sum_{k=1}^s \sum_{j=1}^r \left\{ \frac{X_{2jk} - u c_k X_{1jk}}{1 + u c_k} \right\}. \end{aligned}$$

Lemme 2.1.1. *Sous les mêmes conditions que celle de la proposition 2.3.1, on a :*

$$(2.4.1) \quad \text{var } \Psi(\theta) = \sum_{k=1}^s \frac{\theta c_k x_{\cdot \cdot k}}{(1 + \theta c_k)^2} + \sigma^2 \sum_{k=1}^s \frac{x_{\cdot \cdot k} c_k^2 (x_{\cdot \cdot k} + \theta c_k - 2)}{(1 + \theta c_k)^4} + O(N)$$

Preuve : Il s'agit d'évaluer la variance de la fonction aléatoire Ψ qui dépend en fait de deux familles de distribution : celles des $\vec{v}.a.(X_{1k}, X_{2k})$ et celle des v.a. θ_k . On réécrit, au point $u = \theta$, Ψ sous la forme

$$\Psi(\theta) = \sum_{k=1}^r \frac{x_{..k}}{1 + \theta c_k} - \sum_{k=1}^s \sum_{j=1}^r X_{1jk}.$$

Sous cette forme, il est facile de montrer que

$$(2.4.2) \quad \text{var } \Psi(\theta) = \sum_{k=1}^s \text{var} \left[\sum_{j=1}^r X_{1jk} \right]$$

car $x_{..k}$ est non aléatoire et (X_{1k}, X_{2k}) sont des $\vec{v}.a.$ indépendants.

$$(2.4.3) \quad \text{var} \left[\sum_{j=1}^r X_{1jk} \right] = \sum_{j=1}^r E^{\mathcal{D}} E^X [(X_{1jk})^2] \\ + 2 \sum_{j=1}^{r-1} \sum_{m>j} E^{\mathcal{D}} E^X [X_{1jk} X_{1mk}] - \left[\sum_{j=1}^r E^{\mathcal{D}} E^X (X_{1jk}) \right]^2.$$

En utilisant les expressions relatives aux moments factoriels d'une distribution multinomiale (voir Bishop et al., p. 442), on montre que

$$(2.4.4) \quad E^{\mathcal{D}} E^X [X_{1jk}] = x_{..k} p_{jk} E^{\mathcal{D}} \left[\frac{1}{1 + \theta_k c_k} \right],$$

$$(2.4.5) \quad E^{\mathcal{D}} E^X [(X_{1jk})^2] = x_{..k} p_{jk} E^{\mathcal{D}} \left(\frac{1}{1 + \theta_k c_k} \right) + x_{..k} (x_{..k} - 1) p_{jk}^2 E^{\mathcal{D}} \left[\frac{1}{(1 + \theta_k c_k)^2} \right],$$

$$(2.4.6) \quad E^{\mathcal{D}} E^X [X_{1jk} X_{1mk}] = x_{..k} (x_{..k} - 1) p_{jk} p_{mk} E^{\mathcal{D}} \left[\frac{1}{(1 + \theta_k c_k)^2} \right], m > j.$$

En remplaçant ces trois dernières expressions dans (2.4.3), on a

$$(2.4.7) \quad \text{var} \left[\sum_{j=1}^r X_{1jk} \right] = x_{..k} E^{\mathcal{D}} \left[\frac{1}{1 + \theta_k c_k} \right] + x_{..k} (x_{..k} - 1) E^{\mathcal{D}} \left[\frac{1}{(1 + \theta_k c_k)^2} \right] \\ - x_{..k}^2 \left[E^{\mathcal{D}} \left(\frac{1}{1 + \theta_k c_k} \right) \right]^2.$$

Pour évaluer le membre de droite de l'expression (2.4.7), on développe en probabilité au voisinage de θ les v.a. $1/[1 + \theta_k c_k]$ et $1/[(1 + \theta_k c_k)^2]$. On utilise pour cela le principe de calcul introduit à la proposition 2.3.1.

On a sous les conditions (c1) à (c3) (voir proposition 2.3.1),

$$\begin{aligned} E^{\mathcal{D}} \left[\frac{1}{1 + \theta_k c_k} \right] &= E^{\mathcal{D}} \left[\sum_{n=0}^{\infty} \frac{(-1)^n c_k^n}{(1 + \theta_k c_k)^{n+1}} (\theta_k - \theta)^n \right] \\ &= \frac{1}{1 + \theta_k c_k} + \frac{c_k^2}{(1 + \theta_k c_k)^3} \sigma^2 + O(1), \\ E^{\mathcal{D}} \left[\frac{1}{(1 + \theta_k c_k)^2} \right] &= E^{\mathcal{D}} \left[\sum_{n=0}^{\infty} \frac{(-1)^n (n+1) c_k^n}{(1 + \theta_k c_k)^{n+2}} (\theta_k - \theta)^n \right] \\ &= \frac{1}{(1 + \theta_k c_k)^2} + \frac{3c_k^2}{(1 + \theta_k c_k)^4} \sigma^2 + O(1). \end{aligned}$$

On obtient donc

$$(2.4.8) \quad \text{var} \left(\sum_{j=1}^r X_{1jk} \right) = \frac{\theta_k c_k x_{..k}}{(1 + \theta_k c_k)^2} + \frac{x_{..k} c_k^2 (x_{..k} + \theta_k c_k - 2)}{(1 + \theta_k c_k)^4} \sigma^2 + O(x_{..k}).$$

En portant cette dernière expression dans (2.4.2), on a

$$\text{var} \Psi(\theta) = \sum_{k=1}^s \frac{\theta_k c_k x_{..k}}{(1 + \theta_k c_k)^2} + \sigma^2 \sum_{k=1}^s \frac{x_{..k} c_k^2 + \theta_k c_k - 2}{(1 + \theta_k c_k)^4} + O(N).$$

Cette variance est interprétée comme étant les éventuelles variations pouvant provenir non seulement de l'hypothèse de distribution multinomiale faite sur les $\vec{v}.a.(X_{1k}, X_{2k})$ mais aussi des variations dues à la distribution des θ_k .

Lemme 2.4.2. *Si en plus des conditions du lemme 2.4.1, on suppose que*

(c4) $\log \hat{\theta} = \log \theta + O_p(r_N)$, avec $r_N \rightarrow 0$ et $N^{-2} r_N^3 \rightarrow 0$ quand $N \rightarrow \infty$.

(c5) $E^{\mathcal{D}} E^X(\log \hat{\theta}) = \log \theta$.

(c6) $E^{\mathcal{D}} E^X \{(\log \hat{\theta} - \log \theta)^2\} < \infty$, alors

$$\text{var} \log \hat{\theta} = \frac{\sum_{k=1}^r \frac{\theta_k c_k x_{..k}}{(1 + \theta_k c_k)^2} + \sigma^2 \sum_{k=1}^s \frac{x_{..k} c_k^2 (x_{..k} + \theta_k c_k - 2)}{(1 + \theta_k c_k)^4}}{\left(\sum_{k=1}^s \frac{\theta_k c_k x_{..k}}{(1 + \theta_k c_k)^2} \right)^2}$$

Preuve : Elle est basée sur le développement de Taylor en probabilité. On ne retiendra que les termes linéaires car seuls ces termes interviennent dans

les formes d'estimation de la variance d'un estimateur ou d'une fonction d'un estimateur. Néanmoins sous certaines conditions, le développement peut s'étendre à des polynômes de degré supérieur. En utilisant donc le développement de Taylor en probabilité à l'ordre un de Ψ au voisinage de l'espérance mathématique θ , on peut écrire sous la condition (c4) (voir Chernoff, corollaire 1 ; Wolter p. 223) $\Psi(\hat{\theta}) - \Psi(\theta) = (\log \hat{\theta} - \log \theta) \left(\frac{\partial \Psi(u)}{\partial \log u} \right)_{u=\theta} + O_p(r_N^2)$. Soit alors

$$(2.4.9) \quad (\log \hat{\theta} - \log \theta)^2 = \frac{(\Psi(\hat{\theta}) - \Psi(\theta))^2}{\left(\frac{\partial \Psi(u)}{\partial \log u} \right)_{u=\theta}^2} + O_p(N^{-2}r_N^3).$$

En tenant compte des conditions (c5) à (c6) et du fait que $\Psi(\hat{\theta}) = 0$ (par construction de Ψ), on a

$$(2.4.10) \quad \text{var} \log \hat{\theta} = \frac{E^{\mathcal{D}} E^X[(\Psi(\theta))^2]}{\left(\frac{\partial \Psi(u)}{\partial \log u} \right)_{u=\theta}^2}; \quad (N \rightarrow \infty)$$

où

$$\left(\frac{\partial \Psi(u)}{\partial \log u} \right)_{u=\theta}^2 = - \sum_{k=1}^s \frac{\theta c_k x_{..k}}{(1 + \theta c_k)^2}.$$

En utilisant les conditions du lemme 2.4.1, on montre que

$$E^{\mathcal{D}} E^X(\Psi(\theta)) = -\sigma^2 \sum_{k=1}^s \frac{x_{..k} c_k^2}{(1 + \theta c_k)^3} + O(N) = O(N).$$

Par conséquent

$$\begin{aligned} E^{\mathcal{D}} E^X[\Psi(\theta)]^2 &= \text{var} \Psi(\theta) + [E^{\mathcal{D}} E^X(\Psi(\theta))]^2 \\ &= \text{var} \Psi(\theta) + O(N^2). \end{aligned}$$

D'après le lemme 2.4.1, on a

$$E^{\mathcal{D}} E^X[\Psi(\theta)]^2 = \sum_{k=1}^s \frac{\theta c_k x_{..k}}{(1 + \theta c_k)^2} + \sigma^2 \sum_{k=1}^s \frac{x_{..k} c_k^2 (x_{..k} + \theta c_k - 2)}{(1 + \theta c_k)^4} + O(N).$$

Par substitution dans (2.4.10), on a le résultat du lemme 2.4.2.

Proposition 2.4.1. *Sou les mêmes conditions que celles des lemmes 2.4.1 et 2.4.2, on montre qu'une approximation de la variance asymptotique de $\log \hat{\theta}$ est :*

$$(2.4.11) \quad \text{var} \log \hat{\theta} = (1 + \phi) \rho$$

où une approximation de ϕ est donnée par :

$$\phi \simeq \left[\frac{\hat{A}_{s,r}^2}{sr - 1} - 1 \right] \left[\frac{sr \left(\sum_{k=1}^s x_{..k}^2 \right)}{N^2 \left[1 + \frac{s(r-1)}{N} \right]} \right].$$

Preuve : D'après les lemmes 2.4.1 et 2.4.2, on a

$$\text{var} \log \hat{\theta} = \frac{1}{\sum_{k=1}^s \frac{\theta_{c_k} x_{..k}}{(1+\theta_{c_k})^2}} \left[1 + \sigma^2 \frac{\sum_{k=1}^s \frac{x_{..k} c_k^2 (x_{..k} + \theta_{c_k} - 2)}{(1+\theta_{c_k})^4}}{\sum_{k=1}^s \frac{\theta_{c_k} x_{..k}}{(1+\theta_{c_k})^2}} \right].$$

En utilisant l'estimation de σ^2 donnée par la proposition 2.3.1, on a :

$$\text{var} \log \hat{\theta} = \frac{1 + \phi}{\sum_{k=1}^s \frac{\theta_{c_k} x_{..k}}{(1+\theta_{c_k})^2}}$$

avec

$$(2.4.12) \quad \phi = \frac{[\hat{A}_{s,r}^2 - (sr - 1)] \sum_{k=1}^s \left\{ \frac{\theta_{c_k} (x_{..k} + \theta_{c_k} - 2)}{(1+\theta_{c_k})^2} \cdot \frac{\theta_{c_k} x_{..k}}{(1+\theta_{c_k})^2} \right\}}{\left(\sum_{k=1}^s \frac{\theta_{c_k} x_{..k}}{(1+\theta_{c_k})^2} \right) \left[\sum_{k=1}^s \left\{ \frac{\theta_{c_k} (x_{..k} + \theta_{c_k} - 2)}{(1+\theta_{c_k})^2} \right\} + (r - 1) \sum_{k=1}^s \frac{\theta_{c_k} (\theta_{c_k} - 1)}{(1+\theta_{c_k})^2} \right]}.$$

Pour simplifier ϕ et retrouver le résultat de la proposition 2.4.1, nous allons procéder à des approximations de (2.4.12).

2.4.4. Approximation du coefficient d'hétérogénéité ϕ .

Si le terme $\theta_{c_k} - 2$ est négligeable par rapport à $x_{..k}$ dans l'expression $x_{..k} + \theta_{c_k} - 2$ alors ϕ devient :

$$(2.4.13) \quad \phi = \frac{[\hat{A}_{s,r}^2 - (sr - 1)] \sum_{k=1}^s \left\{ \frac{\theta_{c_k} x_{..k}}{(1+\theta_{c_k})^2} \cdot \frac{\theta_{c_k} x_{..k}}{(1+\theta_{c_k})^2} \right\}}{\left[\sum_{k=1}^s \frac{\theta_{c_k} x_{..k}}{(1+\theta_{c_k})^2} \right] \left[\sum_{k=1}^s \frac{\theta_{c_k} x_{..k}}{(1+\theta_{c_k})^2} + (r - 1) \sum_{k=1}^s \frac{\theta_{c_k} (\theta_{c_k} - 1)}{(1+\theta_{c_k})^2} \right]}.$$

Posons $n_k = \frac{\theta_{c_k} x_{..k}}{(1+\theta_{c_k})^2}$; ($k = 1, 2, \dots, s$). On suppose aussi que

$$\frac{1}{s} \sum_{k=1}^s n_k \neq \frac{r-1}{8}.$$

En tenant compte du comportement de la fonction (voir fig. 2) $x \mapsto x(x-1)/(1+x)^2$ pour les $x > 0$, on montre que :

$$\frac{[\hat{A}_{s,r}^2 - (sr - 1)] (\sum_{k=1}^s n_k^2)}{(\sum_{k=1}^s n_k) [s(r-1) + \sum_{k=1}^s n_k]} \leq \phi \leq \frac{[\hat{A}_{s,r}^2 - (sr - 1)] (\sum_{k=1}^s n_k^2)}{(\sum_{k=1}^s n_k) \left(\sum_{k=1}^s n_k - \frac{s(r-1)}{8} \right)}.$$

On peut majorer la borne supérieure de ϕ en minimisant son dénominateur c'est-à-dire en remplaçant $\sum_{k=1}^s n_k - \frac{s(r-1)}{8}$ par $\sum_{k=1}^s n_k - s(r-1)$. Cependant, pour obtenir une bonne précision c'est-à-dire une variance asymptotique de $\log \hat{\theta}$ minimale, il faut minimiser ϕ . On suggère donc de prendre

$$\phi = \frac{[\hat{A}_{s,r}^2 - (sr - 1)] (\sum_{k=1}^s n_k^2)}{(\sum_{k=1}^s n_k) [\sum_{k=1}^s n_k + s(r-1)]}.$$

Si on tient compte, maintenant, du comportement de la fonction (fig. 1) $x \mapsto x/(1+x)^2$ pour les $x > 0$ alors on a

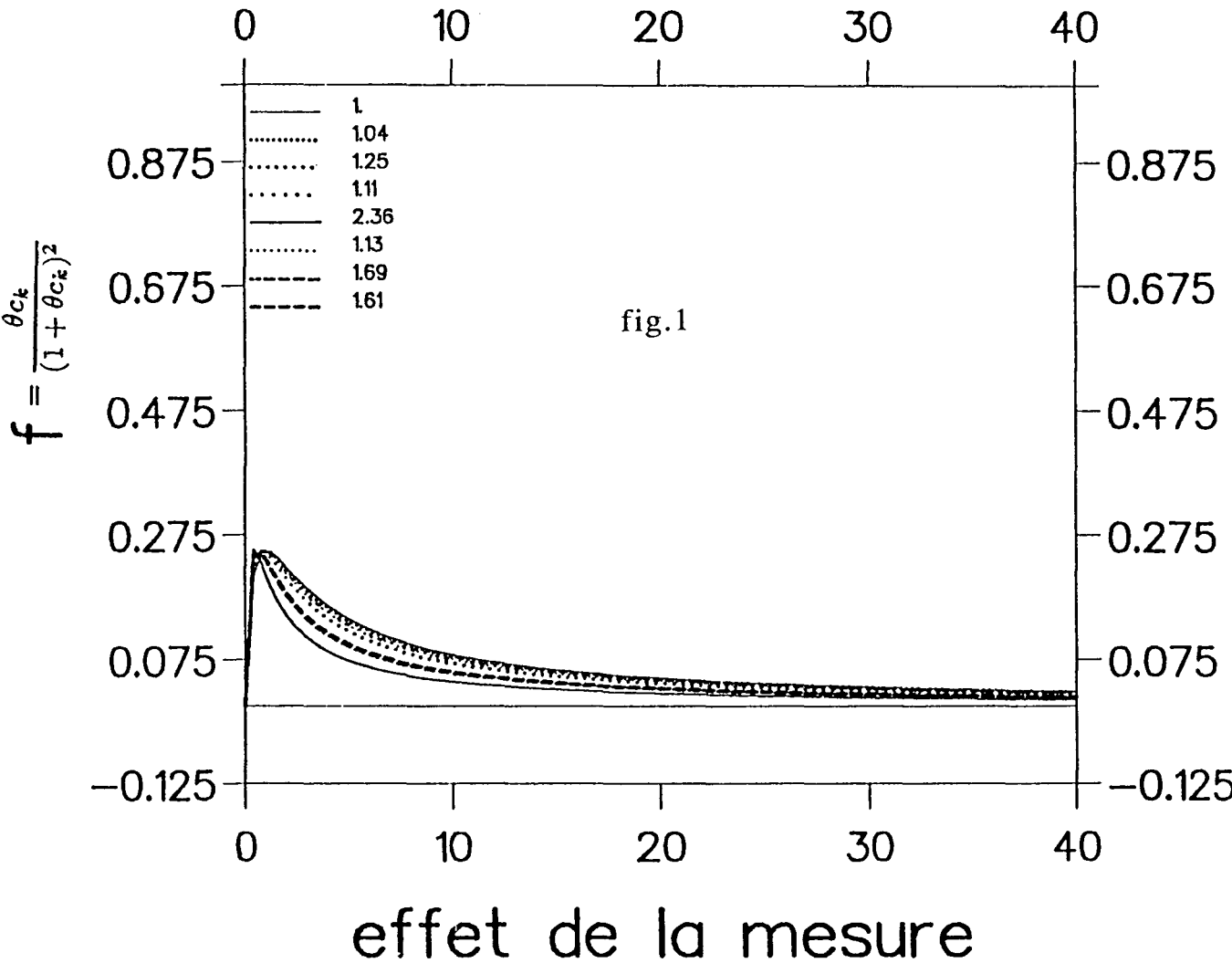
$$0 \leq \sum_{k=1}^s n_k \leq \frac{N}{4} < N.$$

Une approximation de l'expression (2.4.14) est donnée par

$$(2.4.15) \quad \phi \simeq \left[\frac{\hat{A}_{sr}}{sr-1} - 1 \right] \frac{(sr-1) (\sum_{k=1}^s x_{..k}^2)}{N^2 \left[1 + \frac{s(r-1)}{N} \right]}.$$

En multipliant cette dernière expression par $sr/(sr-1)$, on obtient le coefficient d'hétérogénéité ϕ donné par la proposition 2.4.1. Pour différentes valeurs c on représente la fonction $f = \theta c/(1+\theta c)^2$ [resp. $\theta c(\theta c - 1)/(1+\theta c)^2$] en fonction de θ . Les valeurs de c sont celles données dans [Tanner].

Remarque 2.4.1. Pour $r = 1$, l'approximation de ϕ donnée dans la proposition 2.4.1 est identique à celle de Tanner. Les résultats du chapitre 3 montre $\hat{A}_{s,r}^2(r=1)$ a asymptotiquement ($N \rightarrow \infty$) la même distribution qu'une variable aléatoire de loi χ_{s-1}^2 . Ceci permet de montrer que ϕ est une généralisation du coefficient d'hétérogénéité de Tanner. On peut aussi simplifier ϕ en comparant $s(r-1)$ et N . Il faut en pratique multiplier l'expression (2.4.11) donnant $\text{var} \log \hat{\theta}$ par un facteur correcteur obtenu en inversant la matrice d'information de Fisher (voir paragraphe 2.3.3). Ces considérations d'ordre pratiques ne seront pas abordées ici.



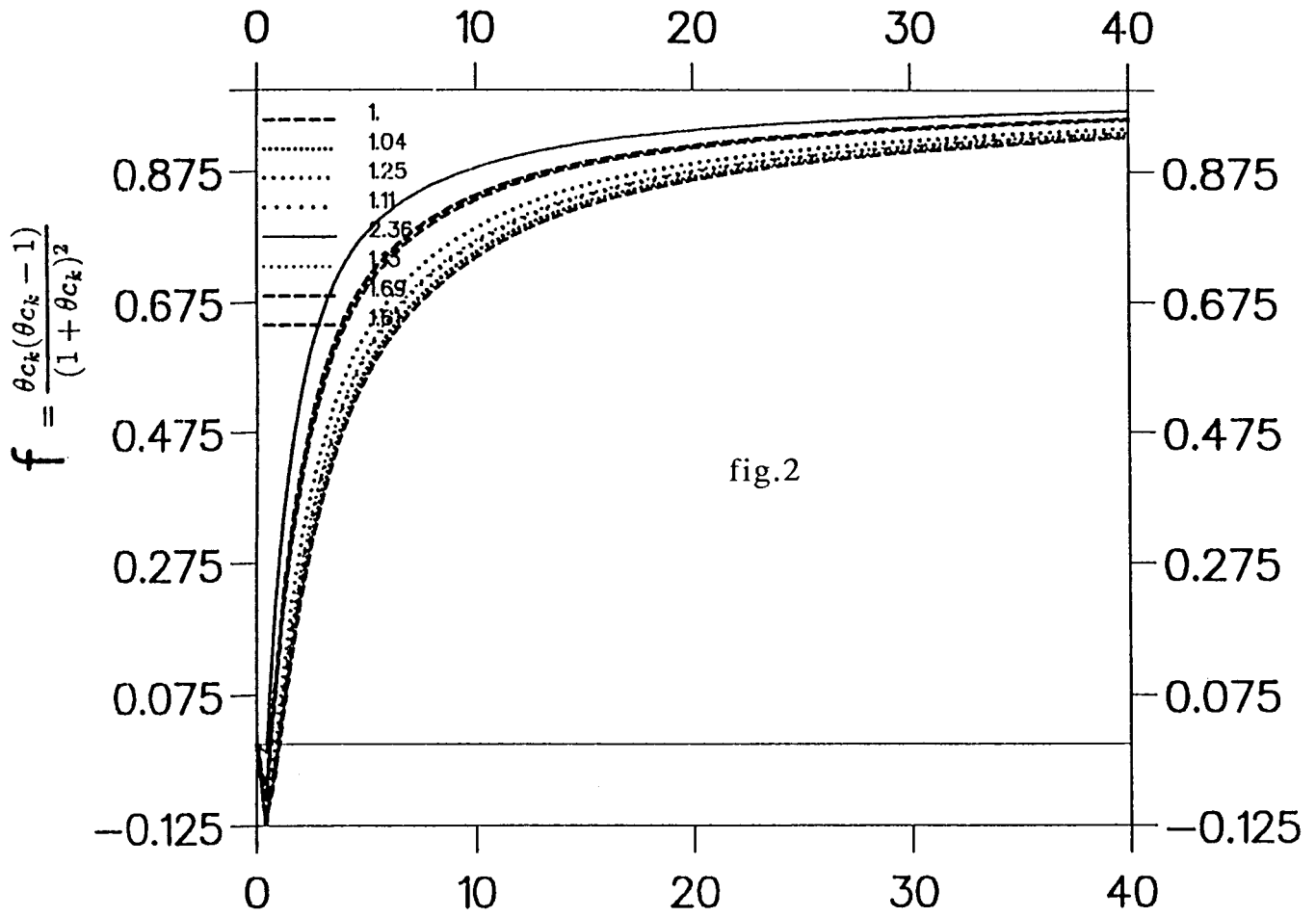


fig.2

effet de la mesure

2.4.5. Test de l'hypothèse d'efficacité nulle.

Il s'agit de tester si $\theta = 1$ (l'efficacité de la mesure est nulle : c'est-à-dire tester la moyenne d'une certaine distribution égale à une valeur donnée). On propose d'utiliser la v.a.

$$(2.4.16) \quad Z = \frac{\log \hat{\theta}}{[\text{var} \log \hat{\theta}]^{1/2}}$$

où $\hat{\theta}$ est obtenue en résolvant les systèmes (2.3.1) et (2.3.2) et

$$(2.4.17) \quad \text{var} \log \hat{\theta} \simeq \rho(1 + \phi) \quad (N \rightarrow \infty).$$

Cette approche consiste donc à procéder comme sous l'hypothèse H_0 d'égalité des effets mais en corrigeant la variance asymptotique ρ par intégration du coefficient $(1 + \phi)$ donné par la proposition 2.4.1. En admettant que la variable aléatoire donnée par l'expression (2.4.17) a pour loi asymptotique $\mathcal{N}(0, 1)$, sous l'hypothèse $\theta = 1$, on rejette l'hypothèse $\theta = 1$ si la valeur de Z est en valeur absolue supérieure à 1,96.

Remarque 2.4.2 Il serait intéressant d'étudier sous certaines conditions la normalité asymptotique de la statistique

$$Z = \frac{\log \hat{\theta}}{\rho^{1/2}(1 + \phi)^{1/2}}.$$

Cet aspect de la recherche ne sera pas abordé dans ce travail.

CHAPITRE 3:
DISTRIBUTION
ASYMPTOTIQUE
DE $\hat{A}_{s,r}^2$

3.1. INTRODUCTION

Diverses statistiques du type chi-deux peuvent être utilisées (*voir* [5], [19]) pour ajuster un modèle théorique à des données. Ces statistiques comparent un ensemble de valeurs espérées (théoriques) issues d'un modèle spécifique à un ensemble de valeurs observées provenant d'une expérience. L'utilisation pratique de ces statistiques ([26]) 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 (*voir* par exemple [10],[18],[21],[54],[56],[57],[59]). Dans le domaine de la sécurité routière, on fait souvent appel à la statistique introduite par Tanner ([70]) et discutée au chapitre 1 pour éprouver l'hypothèse d'égalité de l'effet de la mesure (*voir* aussi [46]). Lorsque les données sont mises sous forme d'un tableau avant-après de grandes dimensions, on propose [*voir* N'Guessan et Langrand 1993] un test d'adéquation pour décider de l'existence ou non de variations entre les différents effets. On considère pour cela:

- soit la statistique

$$\hat{A}_{s,r}^2 = \sum_{k=1}^s \sum_{j=1}^r \sum_{t=1}^2 \frac{(X_{tjk} - x_{..k}q_{tjk}(\hat{\theta}, \hat{P}_k))^2}{x_{..k}q_{tjk}(\hat{\theta}, \hat{P}_k)}$$

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 de $\Theta = (\theta, P_1^T, \dots, P_s^T)^T$ basé sur l'ensemble de toutes les données (*voir* N'guessan et Langrand [15]).

- soit la statistique

$$\tilde{A}_{s,r}^2 = \sum_{k=1}^s \sum_{j=1}^r \sum_{t=1}^2 \frac{(X_{tjk} - x_{..k}q_{tjk}(\tilde{\theta}, \tilde{P}_k))^2}{x_{..k}q_{tjk}(\tilde{\theta}, \tilde{P}_k)}$$

où $\tilde{\theta}$ et $\tilde{P}_k$ 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. Les résultats sont obtenus dans le cadre du modèle 1 et utilisent certaines techniques de Cramer, de Chernoff et Lehmann. Des calculs similaires (*voir* annexe A4) montrent que le modèle 2 donne les mêmes conclusions. Après quelques

notations et lemmes préliminaires (paragraphe 3.2), on donne les principaux résultats (théorèmes 1 et 2) et leurs démonstrations au paragraphe 3.3. On discute, au paragraphe 3.4, d'un nouvel algorithme de calcul de rang de matrice. Ce chapitre se termine (paragraphe 3.5) par une discussion sur les utilisations pratiques de la statistique $\hat{A}_{s,r}^2$.

3.2. NOTATIONS et LEMMES PRELIMINAIRES

3.2.1 Notations

On pose:

$Q^T = (Q_1^T, \dots, Q_k^T, \dots, Q_s^T)$ le vecteur à $2sr$ dimensions où

$Q_k^T = (q_{11k}(\theta, P_k), \dots, q_{1rk}(\theta, P_k), q_{21k}(\theta, P_k), \dots, q_{2rk}(\theta, P_k))$ est le vecteur à $2r$ dimensions de composantes les $q_{tjk}(\theta, P_k)$;

$\beta_k^T = (p_{1k}, \dots, p_{r-1k})$, ($k = 1, 2, \dots, s$);

$\Phi = (\theta, \beta_1^T, \dots, \beta_k^T, \dots, \beta_s^T)^T$ et g l'application de $\mathbb{R}^{1+s(r-1)}$ dans $\mathbb{R}^{1+sr}$ qui à Φ associe $g(\Phi) = \Theta$ dont la matrice jacobienne notée M_o est de taille $(1 + sr, 1 + s(r - 1))$;

$\Pi^T = (\Pi_1^T, \dots, \Pi_k^T, \dots, \Pi_s^T) \in \mathbb{R}^{2sr}$ où

$\Pi_k^T = (\pi_{11k}(\theta, \beta_k), \dots, \pi_{1rk}(\theta, \beta_k), \pi_{2rk}(\theta, \beta_k), \dots, \pi_{2rk}(\theta, \beta_k))$ est le vecteur à $2r$ dimensions de composantes les $\pi_{tjk}(\theta, \beta_k) = q_{tjk}(\theta, P_k)$;

D_o la matrice de taille $(2sr, 1 + s(r - 1))$ dont chaque élément est donné par l'expression $(\pi_{tjk}^{-1/2} \frac{\partial \pi_{tjk}}{\partial \Phi_m})$ évaluée au point $\Phi = \Phi^o$, où Φ^o est la vraie valeur du paramètre Φ ;

B_o la matrice de taille $(2sr, 1 + sr)$ dont chaque élément est donné par l'expression $(q_{tjk}^{-1/2} \frac{\partial q_{tjk}}{\partial \Theta_m})$ évaluée au point $\Theta^o = g(\Phi^o)$;

Y le vecteur aléatoire à $2sr$ dimensions de composantes les Y_{tjk} telles que $Y_{tjk} = [X_{tjk} - x_{..k}\pi_{tjk}^o] / [x_{..k}\pi_{tjk}^o]^{1/2}$ avec $\pi_{tjk}^o = \pi_{tjk}(\Phi^o)$.

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 $q_{tjk}(\Theta)$ pour $q_{tjk}(\theta, P_k)$ et $\pi_{tjk}(\Phi)$ pour $\pi_{tjk}(\theta, \beta_k)$.

$\Lambda^{1/2}$ la matrice bloc-diagonale de taille $2sr$

$$\Lambda^{1/2} = \begin{pmatrix} \Lambda_1^{1/2} & O & \dots & O \\ O & \Lambda_2^{1/2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & O \\ O & \dots & O & \Lambda_s^{1/2} \end{pmatrix}$$

où $\Lambda_k^{1/2} = (x_{..k})^{1/2} I_{2r}$ (I_{2r} matrice identité de $\mathbb{R}^{2r}$) ($k = 1, 2, \dots, s$);

$V_k^T = (x_{..k})^{1/2} (\pi_{11k}^{-1/2} \frac{\partial \pi_{11k}}{\partial \theta}, \dots, \pi_{1rk}^{-1/2} \frac{\partial \pi_{1rk}}{\partial \theta}, \pi_{21k}^{-1/2} \frac{\partial \pi_{21k}}{\partial \theta}, \dots, \pi_{2rk}^{-1/2} \frac{\partial \pi_{2rk}}{\partial \theta})$ le vecteur à $2r$ dimensions évaluée au point $\Phi = \Phi^o$;

W_k la matrice de taille $(2r, r-1)$ évaluée au point $\Phi = \Phi^o$

$$W_k = (x_{..k})^{1/2} \begin{pmatrix} \pi_{11k}^{-1/2} \frac{\partial \pi_{11k}}{\partial p_{1k}} & \dots & \pi_{11k}^{-1/2} \frac{\partial \pi_{11k}}{\partial p_{r-1k}} \\ \vdots & & \vdots \\ \pi_{1rk}^{-1/2} \frac{\partial \pi_{1rk}}{\partial p_{1k}} & \dots & \pi_{1rk}^{-1/2} \frac{\partial \pi_{1rk}}{\partial p_{r-1k}} \\ \pi_{21k}^{-1/2} \frac{\partial \pi_{21k}}{\partial p_{1k}} & \dots & \pi_{21k}^{-1/2} \frac{\partial \pi_{21k}}{\partial p_{r-1k}} \\ \vdots & & \vdots \\ \pi_{2rk}^{-1/2} \frac{\partial \pi_{2rk}}{\partial p_{1k}} & \dots & \pi_{2rk}^{-1/2} \frac{\partial \pi_{2rk}}{\partial p_{r-1k}} \end{pmatrix};$$

$D_k = I_{2r} - W_k(W_k^T W_k)^{-1} W_k^T$, la matrice carrée de taille $2r$, [on montre (voir annexe A3) que $\text{rang}(W_k) = r-1$] ($k = 1, 2, \dots, s$);

$\alpha = \sum_{k=1}^s V_k^T D_k V_k$ réel strictement positif;

$V^T = (V_1^T D_1, \dots, V_s^T D_s)$ vecteur à $2sr$ dimensions.

3.2.2 Lemmes préliminaires

Lemme 3.2.1.- *Le vecteur aléatoire Y est asymptotiquement ($N \rightarrow \infty$) normalement distribué de moyenne nulle et de matrice de variances-covariances $I_{2sr} - \Pi^{1/2} \Pi^{1/2^T}$, où $\Pi^{1/2}$ est le vecteur à $2sr$ dimensions de composantes les $(\pi_{ijk}^o)^{1/2}$ et I_{2sr} matrice identité de $\mathbb{R}^{2sr}$.*

Preuve: par définition des π_{tjk} et par construction des variables aléatoires X_{tjk} , il est clair sous H_0 , que $E(Y)$ est nulle. La matrice de variances-covariances est donc égale à $E(Y Y^T)$. On montre que:

$$E(Y_{tjk} Y_{mlp}) = \frac{E(X_{tjk} X_{mlp}) - x_{..k} x_{..p} q_{tjk}^o q_{mlp}^o}{[x_{..k} x_{..p} q_{tjk}^o q_{mlp}^o]^{1/2}}$$

où $q_{tjk}^o = q_{tjk}(\Theta^o)$. En utilisant l'indépendance des vecteurs aléatoires X_k , ($k = 1, 2, \dots, s$) et les moments factoriels d'une loi multinomiale (voir [5] p.442), on a: $E(X_{tjk}X_{mlp}) = x_{..k}x_{..p}q_{tjk}^o q_{mlp}^o$ si $k \neq p$; $E(X_{tjk}X_{mlp}) = x_{..k}(x_{..k} - 1)q_{tjk}^o q_{mlk}^o$ si $k = p$ et $t \neq m$ ou $j \neq l$; $E(X_{tjk}^2) = x_{..k}q_{tjk}^o(1 - q_{tjk}^o) + x_{..k}^2(q_{tjk}^o)^2$ si $k = p$ et $t = m$ et $j = l$. On déduit alors que $E(Y_{tjk}Y_{mlp}) = 0$ si $k \neq p$; $E(Y_{tjk}Y_{mlp}) = -(q_{tjk}^o q_{mlp}^o)^{1/2}$ si $k = p$ et $t \neq m$ ou $j \neq l$; $E(Y_{tjk}^2) = 1 - q_{tjk}^o$ si $k = p$ et $t = m$ et $j = l$.

Lemme 3.2.2.- *La matrice carrée $\frac{1}{\alpha}VV^T$ de taille $2sr$ possède une valeur propre identique à 1 et $2sr - 1$ valeurs propres nulles.*

Preuve: $(\frac{1}{\alpha}VV^T)(\frac{1}{\alpha}VV^T) = \frac{1}{\alpha}VV^T$ (car $VV^T = \alpha$). La matrice $\frac{1}{\alpha}VV^T$ est donc symétrique et idempotente.

Lemme 3.2.3.- *La matrice carrée $D_k - \Pi_k^{1/2}\Pi_k^{1/2^T}$ de taille $2r$ possède r valeurs propres identique à 1 et r valeurs propres nulles.*

Preuve : on note d'une part que $D_k^T = D_k$, $D_k^2 = D_k$ pour tout $k = 1, 2, \dots, s$ et d'autre part que $D_k\Pi_k^{1/2} = \Pi_k^{1/2}$ car $W_k^T\Pi_k^{1/2} = O$ (d'après les relations $\sum_{j=1}^r \sum_{t=1}^2 \pi_{tjk}(\Phi) = 1$). En utilisant cette remarque, on montre que: $(D_k - \Pi_k^{1/2}\Pi_k^{1/2^T})^2 = D_k - \Pi_k^{1/2}\Pi_k^{1/2^T}$ et $(D_k - \Pi_k^{1/2}\Pi_k^{1/2^T})^T = D_k - \Pi_k^{1/2}\Pi_k^{1/2^T}$. Par suite les valeurs propres de $D_k - \Pi_k^{1/2}\Pi_k^{1/2^T}$ sont soit 1 soit 0. Le nombre de valeurs propres égales à 1 est: $tr(D_k - \Pi_k^{1/2}\Pi_k^{1/2^T}) = 2r - (r - 1) - 1 = r$.

Lemme 3.2.4.- *Soient X un vecteur aléatoire de dimension n normalement distribué de moyenne nulle et de matrice de variances-covariances Ω et A une matrice carrée symétrique de taille n . Alors la forme quadratique X^TAX est distribuée comme $\sum_{j=1}^n \lambda_j Z_j^2$ où les variables aléatoires Z_j sont i.i.d. de loi $\mathcal{N}(0; 1)$ et où les λ_j sont les valeurs propres non nécessairement distinctes de la matrice $A\Omega$.*

Preuve: [voir [54] p.116; [5] p.473].

Lemme 3.2.5.- *Soit $f : E \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ une application définie sur un ouvert E de $\mathbb{R}^n$. Pour tout i , $1 \leq i \leq n$ et tout j , $1 \leq j \leq n$, si les fonctions $(\frac{\partial^2 f}{\partial x_i \partial x_j})$ et $(\frac{\partial^2 f}{\partial x_j \partial x_i})$ existent dans un voisinage du point $a_o \in E$ et sont continues en a_o alors*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a_o) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a_o).$$

3.3. DISTRIBUTIONS ASYMPTOTIQUES DE $\hat{A}_{s,r}^2$ et $\tilde{A}_{s,r}^2$.

3.3.1 Distribution asymptotique de $\tilde{A}_{s,r}^2$

Théorème 1.- *Sous la condition (c) imposée pour $k = 1, 2, \dots, s$:*

(c) N et $x_{..k} = x_{..k}(N)$ tendent vers ∞ de sorte que $\lim_{N \rightarrow \infty} \frac{x_{..k}(N)}{N} = \tau_k$;

la statistique $\tilde{A}_{s,r}^2$ a pour distribution asymptotique ($N \rightarrow \infty$) une loi de χ_{sr-1}^2 .

La condition (c) (voir [59]) 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.

- **Démonstration.** 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:

$$\sum_{k=1}^s \sum_{j=1}^r \sum_{t=1}^2 \frac{(X_{tjk} - x_{..k} \pi_{tjk})}{\pi_{tjk}} \frac{\partial \pi_{tjk}}{\partial \Phi_m} = 0; \quad m = 1, 2, \dots, 1 + s(r-1).$$

Soit $F_m : \vartheta_N(\Phi^o) \subseteq \mathbb{R}^{1+s(r-1)} \rightarrow \mathbb{R}^{1+s(r-1)}$ l'application définie dans un voisinage ouvert $\vartheta_N(\Phi^o)$ de Φ^o par

$$F_m(\Phi) = \sum_{k=1}^s \sum_{j=1}^r \sum_{t=1}^2 \frac{(X_{tjk} - x_{..k} \pi_{tjk}(\Phi))}{\pi_{tjk}(\Phi)} \left(\frac{\partial \pi_{tjk}}{\partial \Phi_m} \right)_{(\Phi)}; \quad m = 1, 2, \dots, 1 + s(r-1).$$

Par différenciation de F_m au voisinage de Φ^o , l'équation ci-dessus devient:

$$F_m(\Phi^o) + \sum_{l=1}^{1+s(r-1)} (\tilde{\Phi}_l - \Phi_l^o) \left(\frac{\partial F_m}{\partial \Phi_l} \right)_{(\Phi=\Phi^o)} + R_m(\tilde{\Phi}, \Phi^o) = 0; \quad m = 1, 2, \dots, 1 + s(r-1)$$

où

$$R_m(\tilde{\Phi}, \Phi^o) = \sum_{l=1}^{1+s(r-1)} \sum_{n=1}^{1+s(r-1)} \frac{1}{2} \left(\frac{\partial^2 F_m}{\partial \Phi_l \partial \Phi_n} \right)_{(\Phi=\Phi^*)} (\tilde{\Phi}_l - \Phi_l^o) (\tilde{\Phi}_n - \Phi_n^o)$$

et où Φ^* est un point du segment joignant Φ^o et $\tilde{\Phi}$. On peut montrer (voir [72] p.223) que $R_m(\tilde{\Phi}, \Phi^o) = O_p(\epsilon_N^2)$ où $\epsilon_N \rightarrow 0$ quand $N \rightarrow \infty$. Un calcul direct de dérivation montre que:

$$\left(\frac{\partial F_m}{\partial \Phi_l} \right)_{(\Phi=\Phi^o)} = - \sum_{k=1}^s \sum_{j=1}^r \sum_{t=1}^2 \frac{x_{..k}}{\pi_{tjk}^o} \left(\frac{\partial \pi_{tjk}}{\partial \Phi_m} \right)_o \left(\frac{\partial \pi_{tjk}}{\partial \Phi_l} \right)_o + \omega_{ml}^{(1)}; \quad m = 1, 2, \dots, 1 + s(r-1)$$

où

$$\omega_{ml}^{(1)} = \sum_{k=1}^s \sum_{j=1}^r \sum_{t=1}^2 \left(\frac{X_{tjk} - x_{..k} \pi_{tjk}^o}{\pi_{tjk}^o} \right) \left[\left(\frac{\partial^2 \pi_{tjk}}{\partial \Phi_l \partial \Phi_m} \right)_o - \frac{1}{\pi_{tjk}^o} \left(\frac{\partial \pi_{tjk}}{\partial \Phi_m} \right)_o \left(\frac{\partial \pi_{tjk}}{\partial \Phi_l} \right)_o \right].$$

Par substitution de $(\frac{\partial F_m}{\partial \Phi_l})_{(\Phi=\Phi^o)}$ l'équation ci-dessus devient:

$$\begin{aligned} \sum_{l=1}^{1+s(r-1)} (\tilde{\Phi}_l - \Phi_l^o) \left[\sum_{k=1}^s \sum_{j=1}^r \sum_{t=1}^2 \frac{x_{..k}}{\pi_{tjk}^o} \left(\frac{\partial \pi_{tjk}}{\partial \Phi_m} \right)_o \left(\frac{\partial \pi_{tjk}}{\partial \Phi_l} \right)_o \right] = \\ \sum_{k=1}^s \sum_{j=1}^r \sum_{t=1}^2 \left(\frac{X_{tjk} - x_{..k} \pi_{tjk}^o}{\pi_{tjk}^o} \right) \left(\frac{\partial \pi_{tjk}}{\partial \Phi_l} \right)_o + \omega_m(\tilde{\Phi}, \Phi^o); \\ \omega_m(\tilde{\Phi}, \Phi^o) = R_m(\tilde{\Phi}, \Phi^o) + \sum_{l=1}^{1+s(r-1)} (\tilde{\Phi}_l - \Phi_l^o) \omega_{ml}^{(1)}; \quad m = 1, 2, \dots, 1 + s(r-1) \end{aligned}$$

En tenant compte des relations $(\partial \pi_{tjk} / \partial p_{ml}) = 0$ pour tout $l \neq k$ (par construction des modèles 1 et 2), on obtient en notation matricielle:

$$(D^T D)(\tilde{\Phi} - \Phi^o) = D^T Y + \omega$$

avec $D = \Lambda^{1/2} D_o$ et ω le vecteur aléatoire de composantes les ω_m . Puisque que $\omega = o_p(1)$ (voir [17] p.431) et $\text{rang}(D) = 1 + s(r-1)$ [voir paragraphe 3.4], on peut écrire que:

$$(\tilde{\Phi} - \Phi^o) = (D^T D)^{-1} D^T Y + o_p(1).$$

Soit maintenant $\tilde{A}$ le vecteur à $2sr$ composantes définies par

$$\tilde{A}_{tjk} = [X_{tjk} - x_{..k} \pi_{tjk}(\tilde{\Phi})] / [x_{..k} \pi_{tjk}(\tilde{\Phi})]^{1/2}.$$

$\tilde{A}_{tjk}$ s'écrit aussi:

$$\begin{aligned} \tilde{A}_{tjk} = \frac{X_{tjk} - x_{..k} \pi_{tjk}(\Phi^o)}{[x_{..k} \pi_{tjk}(\Phi^o)]^{1/2}} - (x_{..k})^{1/2} \frac{\pi_{tjk}(\tilde{\Phi}) - \pi_{tjk}(\Phi^o)}{[\pi_{tjk}(\Phi^o)]^{1/2}} \\ + \frac{X_{tjk} - x_{..k} \pi_{tjk}(\tilde{\Phi})}{(x_{..k})^{1/2}} \left(\frac{1}{[\pi_{tjk}(\tilde{\Phi})]^{1/2}} - \frac{1}{[\pi_{tjk}(\Phi^o)]^{1/2}} \right) \end{aligned}$$

En appliquant à $\pi_{tjk}(\tilde{\Phi}) - \pi_{tjk}(\Phi^o)$ le développement de Taylor en probabilité à l'ordre un au voisinage de Φ^o (on suppose que $\tilde{\Phi} \in \mathcal{V}_N(\Phi^o)$ pour $N \rightarrow \infty$) et l'inégalité de Tchebycheff au dernier terme de droite [voir Murthy et Gafarian p.191], on montre que:

$$\tilde{A}_{tjk} = Y_{tjk} - \frac{(x_{..k})^{1/2}}{[\pi_{tjk}(\Phi^o)]^{1/2}} \sum_{l=1}^{1+s(r-1)} (\tilde{\Phi}_l - \Phi_l^o) \left(\frac{\partial \pi_{tjk}}{\partial \Phi_l} \right)_o + o_p(1) :$$

c'est-à-dire $\tilde{A} = Y - D(\tilde{\Phi} - \Phi^o) + o_p(1)$. En substituant dans cette dernière égalité l'expression de $\tilde{\Phi} - \Phi^o$ ci-dessus, on obtient finalement

$$\tilde{A} = (I_{2sr} - D(D^T D)^{-1} D^T)Y + o_p(1).$$

Il nous faut donc connaître la distribution asymptotique de

$$A^T A = Y^T [I_{2sr} - D(D^T D)^{-1} D^T]Y$$

où $A = (I_{2sr} - D(D^T D)^{-1} D^T)Y$. En vertu du lemme 3.2.4 et des relations $(\partial \pi_{tjk} / \partial p_{ml}) = 0$ pour tout $l \neq k$ on montre que $A^T A$ est distribuée comme $\sum_{m=1}^{2sr} \lambda_m Z_m^2$ où les variables aléatoires Z_m sont i.i.d. de loi $\mathcal{N}(0; 1)$ et où les λ_m sont les valeurs propres non nécessairement distinctes de la matrice

$$\tilde{\Omega} = I_{2sr} - \Pi^{1/2} \Pi^{1/2^T} - D(D^T D)^{-1} D^T.$$

On montre (voir annexe A3) que $\tilde{\Omega} = D_{\Pi^{1/2}} - \frac{1}{\alpha} V V^T$, avec

$$D_{\Pi^{1/2}} = \begin{pmatrix} D_{\Pi_1^{1/2}} & \bigcirc & \dots & \bigcirc \\ \bigcirc & D_{\Pi_2^{1/2}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \bigcirc \\ \bigcirc & \dots & \bigcirc & D_{\Pi_s^{1/2}} \end{pmatrix}$$

où $D_{\Pi_k^{1/2}} = D_k - \Pi_k^{1/2} \Pi_k^{1/2^T}$. En utilisant d'une part les lemmes 3.2.2 et 3.2.3 et d'autre part les égalités $W_k^T \Pi_k^{1/2} = O$ et $V_k^T \Pi_k^{1/2} = 0$ (d'après les relations $\sum_{j=1}^r \sum_{t=1}^2 \pi_{tjk}(\Phi) = 1$), on montre que $\tilde{\Omega}^T = \tilde{\Omega}$ et :

$$\begin{aligned} \tilde{\Omega}^2 &= (D_{\Pi^{1/2}} - \frac{1}{\alpha} V V^T)^2 \\ &= D_{\Pi^{1/2}} - \frac{1}{\alpha} D_{\Pi^{1/2}} V V^T - \frac{1}{\alpha} (D_{\Pi^{1/2}} V V^T)^T + \frac{1}{\alpha} V V^T \\ &= D_{\Pi^{1/2}} - \frac{1}{\alpha} V V^T. \end{aligned}$$

On déduit alors que $\tilde{\Omega}$ est une matrice symétrique et idempotente. Par conséquent ses valeurs propres sont soit 1 soit 0. Le nombre de valeurs propres égales à 1 est:

$$tr(\tilde{\Omega}) = \sum_{k=1}^s tr(D_{\Pi_k^{1/2}}) - tr(\frac{1}{\alpha} V V^T) = sr - 1.$$

3.3.2 Distribution asymptotique de $\hat{A}_{s,r}^2$

Théorème 2.- *Sous la condition (c) imposée pour $k = 1, 2, \dots, s$:*

(c) N et $x_{..k} = x_{..k}(N)$ tendent vers ∞ de sorte que $\lim_{N \rightarrow \infty} \frac{x_{..k}(N)}{N} = \tau_k$;

la statistique $\hat{A}_{s,r}^2$ a même distribution asymptotique ($N \rightarrow \infty$) que la variable aléatoire :

$$\kappa_{s,r}^2 = Z + \sum_{m=1}^{1+s(r-1)} \lambda_m Z_m^2$$

où Z est une variable aléatoire de loi de χ_{sr-1}^2 , où les variables aléatoires Z_m sont i.i.d. de loi $\mathcal{N}(0; 1)$, indépendantes de Z , et où les λ_m ($0 < \lambda_m < 1$) peuvent dépendre de θ et des p_{jk} ($k = 1, 2, \dots, s; j = 1, 2, \dots, r$).

- **Démonstration.** On note à une constante additive près

$$L(\Phi) = \sum_{k=1}^s \sum_{j=1}^r \sum_{t=1}^2 x_{tjk} \log(\pi_{tjk}(\theta, \beta_k))$$

la log- vraisemblance relativement au paramètre Φ et J la matrice d'information de Fisher associée et définie par

$$J_{ml} = -E\left[\left(\frac{\partial^2 L(\Phi)}{\partial \Phi_m \partial \Phi_l}\right)_{(\Phi=\Phi^o)}\right].$$

En calculant explicitement le terme entre crochets et puisque $E(X_{tjk}) = x_{..k} \pi_{tjk}(\Phi^o)$ (par définition de π_{tjk}), on montre que:

$$J_{ml} = J_{ml}^* + \sum_{k=1}^s \sum_{j=1}^r \sum_{t=1}^2 \frac{x_{..k}}{\pi_{tjk}(\Phi^o)} \left(\frac{\partial \pi_{tjk}}{\partial \Phi_l}\right)_{(\Phi=\Phi^o)} \left(\frac{\partial \pi_{tjk}}{\partial \Phi_m}\right)_{(\Phi=\Phi^o)}$$

où

$$J_{ml}^* = -E\left[\sum_{k=1}^s \sum_{j=1}^r \sum_{t=1}^2 \frac{X_{tjk}}{\pi_{tjk}(\Phi^o)} \left(\frac{\partial^2 \pi_{tjk}}{\partial \Phi_l \partial \Phi_m}\right)_{(\Phi=\Phi^o)}\right].$$

On remarque que le terme

$$\sum_{k=1}^s \sum_{j=1}^r \sum_{t=1}^2 \frac{x_{..k}}{\pi_{tjk}(\Phi^o)} \left(\frac{\partial \pi_{tjk}}{\partial \Phi_l}\right)_{(\Phi=\Phi^o)} \left(\frac{\partial \pi_{tjk}}{\partial \Phi_m}\right)_{(\Phi=\Phi^o)}$$

n'est rien d'autre que l'élément de la m^e ligne et l^e colonne de la matrice $D^T D = D_o^T \Lambda D_o$. On décompose ainsi la matrice d'information de Fisher sous la forme

$$J = D^T D + J^*.$$

D'après le lemme 3.2.5,

$$\left(\frac{\partial^2 \pi_{tjk}}{\partial \Phi_l \partial \Phi_m}\right)_{(\Phi=\Phi^o)} = \left(\frac{\partial^2 \pi_{tjk}}{\partial \Phi_m \partial \Phi_l}\right)_{(\Phi=\Phi^o)}.$$

Par suite la matrice J^* est symétrique. Ce qui entraîne la symétrie de J . L'estimateur du maximum de vraisemblance $\hat{\Phi}$ de Φ est obtenu en résolvant le système d'équations:

$$\left(\frac{\partial L}{\partial \Phi_m}\right)(\hat{\Phi}) = 0 \quad m = 1, 2, \dots, 1 + s(r-1)$$

avec

$$\frac{\partial L}{\partial \Phi_m}(\Phi) = \sum_{k=1}^s \sum_{j=1}^r \sum_{t=1}^2 \frac{x_{tjk}}{\pi_{tjk}(\Phi)} \frac{\partial \pi_{tjk}(\Phi)}{\partial \Phi_m}.$$

D'après les propriétés du maximum de vraisemblance et en utilisant la décomposition ci-dessus de la matrice J , on montre que

$$(\hat{\Phi} - \Phi^o) = (D^T D + J^*)^{-1} \left(\frac{\partial L}{\partial \Phi}\right)_{(\Phi=\Phi^o)} + o_p(1),$$

où $\frac{\partial L}{\partial \Phi}$ est le vecteur à $1 + s(r-1)$ dimensions de composantes les

$$\frac{\partial L}{\partial \Phi_m} = \sum_{k=1}^s \sum_{j=1}^r \sum_{t=1}^2 \frac{X_{tjkk}}{\pi_{tjk}(\Phi^o)} \left(\frac{\partial \pi_{tjk}}{\partial \Phi_m}\right)_{(\Phi=\Phi^o)}.$$

D'autre part la m^e composante du vecteur aléatoire $D^T Y$ s'écrit:

$$\sum_{k=1}^s \sum_{j=1}^r \sum_{t=1}^2 \frac{(X_{tjk} - x_{..k} \pi_{tjk}(\Phi^o))}{\pi_{tjk}(\Phi^o)} \left(\frac{\partial \pi_{tjk}}{\partial \Phi_m}\right)_{(\Phi=\Phi^o)} ; \quad m = 1, 2, \dots, 1 + s(r-1).$$

En développant le terme entre crochets, cette composante est identique à

$$\sum_{k=1}^s \sum_{j=1}^r \sum_{t=1}^2 \frac{X_{tjk}}{\pi_{tjk}(\Phi^o)} \left(\frac{\partial \pi_{tjk}}{\partial \Phi_m}\right)_{(\Phi=\Phi^o)},$$

car

$$\sum_{k=1}^s \sum_{j=1}^r \sum_{t=1}^2 x_{..k} \left(\frac{\partial \pi_{tjk}}{\partial \Phi_m}\right)_{(\Phi=\Phi^o)} = 0,$$

(d'après les relations $\sum_{j=1}^r \sum_{t=1}^2 \pi_{tjk}(\Phi) = 1$). Par conséquent $\left(\frac{\partial L}{\partial \Phi}\right)_{(\Phi=\Phi^o)} = D^T Y$. On obtient par substitution

$$(\hat{\Phi} - \Phi^o) = (D^T D + J^*)^{-1} D^T Y + o_p(1).$$

Si on désigne maintenant par $\hat{A}$ le vecteur aléatoire à $2sr$ dimensions obtenu en remplaçant $\tilde{\Phi}$ par $\hat{\Phi}$ dans l'expression du vecteur $\tilde{A}$, on montre de façon similaire que: $\hat{A} = Y - D(\hat{\Phi} - \Phi^o) + o_p(1)$. En utilisant l'expression de $(\hat{\Phi} - \Phi^o)$ donnée ci-dessus, on obtient finalement

$$\hat{A} = [I_{2sr} - D(D^T D + J^*)^{-1} D^T] Y + o_p(1).$$

Comme dans le cas de $\tilde{A}_{s,r}^2$, et puisque $\hat{A}_{s,r}^2 = \hat{A}^T \hat{A}$, il nous faut donc déterminer la distribution asymptotique ($N \rightarrow \infty$) de la variable aléatoire

$$A^T A = Y^T [I_{2sr} - D(D^T D + J^*)^{-1} D^T]^2 Y$$

où $A = [I_{2sr} - D(D^T D + J^*)^{-1} D^T] Y$. En combinant les lemmes 3.2.1 et 3.2.4, on montre que $A^T A$ a même distribution asymptotique que $\sum_{m=1}^{2sr} \lambda_m Z_m^2$ où les variables aléatoires Z_m sont i.i.d. de loi $\mathcal{N}(0; 1)$ et où les λ_m sont les valeurs propres non nécessairement distinctes de la matrice

$$\hat{\Omega} = [I_{2sr} - D(D^T D + J^*)^{-1} D^T]^2 [I_{2sr} - \Pi^{1/2} \Pi^{1/2^T}].$$

En remarquant que $D^T \Pi^{1/2} = O$, on a finalement

$$\hat{\Omega} = I_{2sr} - \Pi^{1/2} \Pi^{1/2^T} - D(D^T D + J^*)^{-1} D^T - D(D^T D + J^*)^{-1} J^* (D^T D + J^*)^{-1} D^T.$$

Etant donné J et $D^T D$, il existe [voir Chernoff et Lehmann p.584] une matrice S carrée inversible de taille $1 + s(r-1)$ et une matrice diagonale

$$M = \begin{pmatrix} \mu_1 & O & \dots & O \\ O & \mu_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & O \\ O & \dots & O & \mu_{1+s(r-1)} \end{pmatrix}$$

avec $0 < \mu_i < 1$ telles que

$$(D^T D)^{-1} = S S^T, \quad (D^T D + J^*)^{-1} = S M S^T.$$

Soient $U_1, U_2, \dots, U_{1+s(r-1)}$ les colonnes de la matrice DS avec $U_k \in \mathbb{R}^{2sr}$ et $\Pi_1^*, \Pi_2^*, \dots, \Pi_s^*$ les vecteurs de $\mathbb{R}^{2sr}$ définis par

$$\Pi_m^{*T} = [\delta_{m1} \Pi_1^T, \dots, \delta_{mk} \Pi_k^T, \dots, \delta_{ms} \Pi_s^T] \quad m = 1, 2, \dots, s;$$

où δ_{mj} est le symbole de Kronecker. On a d'une part

$$D(D^T D)^{-1} D^T = (DS)(DS)^T = \sum_{m=1}^{1+s(r-1)} U_m U_m^T$$

et d'autre part

$$(DS)(DS)^T = \begin{pmatrix} U_1^T U_1 & \cdots & U_1^T U_{1+s(r-1)} \\ \vdots & \ddots & \vdots \\ U_{1+s(r-1)}^T U_1 & \cdots & U_{1+s(r-1)}^T U_{1+s(r-1)} \end{pmatrix}.$$

Par suite $(DS)^T(DS) = S^T D^T D S = I_{1+s(r-1)}$. On en déduit alors que $U_1, U_2, \dots, U_{1+s(r-1)}$ est une suite de famille de vecteurs unitaires deux à deux orthogonaux de $\mathbb{R}^{2sr}$. Par construction des vecteurs $\Pi_1^*, \Pi_2^*, \dots, \Pi_s^*$, on a

$$D^T \Pi_m^{*1/2} = O, \quad \Pi_k^{*1/2T} \Pi_m^{*1/2} = \delta_{km} \quad (m = 1, 2, \dots, s; k = 1, 2, \dots, s)$$

et

$$U_j^T \Pi_m^{*1/2} = 0 \quad m = 1, 2, \dots, s; j = 1, 2, \dots, 1 + s(r-1).$$

Par conséquent $\Pi_1^*, \Pi_2^*, \dots, \Pi_s^*, U_1, U_2, \dots, U_{1+s(r-1)}$ forment une famille de vecteurs unitaires deux à deux orthogonaux de $\mathbb{R}^{2sr}$. Si $V_1^*, V_2^*, \dots, V_{sr-1}^*$ complète cette famille en une base de $\mathbb{R}^{2sr}$, on a

$$\begin{aligned} \hat{\Omega} &= I_{2sr} - \Pi^{1/2} \Pi^{1/2T} - 2DSM(DS)^T + DSM^2(DS)^T \\ &= I_{2sr} - \sum_{m=1}^s \Pi_m^{*1/2} \Pi_m^{*1/2T} - \sum_{m=1}^{1+s(r-1)} U_m U_m^T + \sum_{m=1}^{1+s(r-1)} (1 - \mu_m)^2 U_m U_m^T \\ &= \sum_{m=1}^{sr-1} V_m^* V_m^{*T} + \sum_{m=1}^{1+s(r-1)} (1 - \mu_m)^2 U_m U_m^T. \end{aligned}$$

On en déduit alors que la matrice $\hat{\Omega}$ possède $sr - 1$ valeurs propres identiques à 1, s valeurs propres égales à 0 et $\lambda_m = (1 - \mu_m)^2$ ($m = 1, 2, \dots, 1 + s(r-1)$) valeurs propres comprises entre 0 et 1.

3.4 CALCUL DU RANG DE LA MATRICE D.

Lemme 3.4.1 .- ([66]p.27) Soient A et B deux matrices réelles de tailles respectives (m, n) et (r, m) telles que $\text{rang}(A) = m$ et $\text{rang}(B) = r$. Alors $\text{rang}(BA) = r$.

Lemme 3.4.2. - ([31]) Soit $A = (a_{ij})$ une matrice réelle de taille (m, n) . Pour tout couple d'indices (i, j) tels que $2 \leq i \leq m$ et $2 \leq j \leq n$, on définit le sous-déterminant $(2, 2)$:

$$d_{ij} = \det \begin{pmatrix} a_{11} & a_{1j} \\ a_{i1} & a_{ij} \end{pmatrix} = a_{11}a_{ij} - a_{i1}a_{1j}.$$

Si $a_{11} \neq 0$ alors $\text{rang}(A) = 1 + \text{rang} \begin{pmatrix} d_{22} & \cdots & d_{2n} \\ \vdots & & \vdots \\ d_{m2} & \cdots & d_{mn} \end{pmatrix}.$

Lemme 3.4.3. - Soient a un réel non nul, V un vecteur de $\mathbb{R}^l$ et O le vecteur nul de $\mathbb{R}^m$. Si B est une matrice réelle de taille (l, m) et de rang n , alors

$$\text{rang} \begin{pmatrix} a & O^T \\ V & B \end{pmatrix} = 1 + n.$$

Lemme 3.4.4. - Soit $B^{(k)} = ((\frac{\partial Q_k}{\partial p_k}))$ la matrice de taille $(2r, r)$ donnant les dérivées partielles du vecteur Q_k par rapport à $p_{1k}, \dots, p_{rk}$. On a $\text{rang}(B^{(k)}) = r$, pour tout $k = 1, 2, \dots, s$.

Preuve. Un calcul explicite de dérivation dans le cas du modèle 1 montre que

$$\text{rang}(B^{(k)}) = \text{rang} \begin{pmatrix} \Delta_k^{(1)} & -\theta c_{2k}p_{1k} & -\theta c_{3k}p_{1k} & \cdots & -\theta c_{rk}p_{1k} \\ -\theta c_{1k}p_{2k} & \Delta_k^{(2)} & -\theta c_{3k}p_{2k} & \cdots & -\theta c_{rk}p_{2k} \\ -\theta c_{1k}p_{3k} & -\theta c_{2k}p_{3k} & \Delta_k^{(3)} & \cdots & -\theta c_{rk}p_{3k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\theta c_{1k}p_{rk} & -\theta c_{2k}p_{rk} & -\theta c_{3k}p_{rk} & \cdots & \Delta_k^{(r)} \end{pmatrix}$$

où $\Delta_k^{(m)} = (1 + \theta c_k) - \theta c_{mk}p_{mk}$, $m = 1, 2, \dots, r$. Le calcul du rang de la matrice de droite se fait de façon itérative. On calcule à chaque itération les coefficients d_{ij} (voir lemme 3.4.2) associés à la matrice de l'étape suivante. On montre ainsi, après $r - 1$ itérations, que $\text{rang}(B^{(k)}) = r - 1 + \text{rang}[\Delta_k^{(1)} - \theta \sum_{j=2}^{r-2} c_{jk}p_{jk}] = r$.

Lemme 3.4.5 - $\text{Rang}(B_o) = 1 + sr$

Preuve. Après dérivation et simplifications (voir annexe A3), on montre que

$$\text{rang}(B_o) = \text{rang} \begin{pmatrix} U^{(1)} & B^{(1)} & O & \cdots & O \\ U^{(2)} & O & B^{(2)} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & O \\ U^{(s)} & O & \cdots & O & B^{(s)} \end{pmatrix}$$

où $U^{(k)} (k = 1, 2, \dots, s)$ est le vecteur à $2r$ dimensions $((\frac{\partial Q_k}{\partial \theta}))$. En utilisant le lemme 3.4.2, on a après r itérations $Rang(B_o) = 1 + r + \sum_{k=2}^s rang(B^{(k)})$. Le lemme 3.4.4 permet alors d'obtenir le résultat.

Lemme 3.4.6.- $Rang(D_o) = 1 + s(r - 1)$

Preuve. On montre que $D_o = B_o M_o$, où $rang(M_o) = 1 + s(r - 1)$. On utilise alors le lemme 3.4.1.

Remarque $rang(D) = rang(\Lambda^{1/2} D_o) = 1 + s(r - 1)$.

3.5. UTILISATIONS PRATIQUES DE $\hat{A}_{s,r}^2$.

3.5.1 Test de l'hypothèse H_0

Pour éprouver H_0 , on compare la statistique $\hat{A}_{s,r}^2$ à la valeur α_ν que dépasse une variable aléatoire de loi de χ_ν^2 avec une probabilité α . Si, au niveau α , la valeur calculée de $\hat{A}_{s,r}^2$ est trop grande, on peut conclure qu'il y a vraisemblablement une différence dans l'effet de la mesure aux différents sites, on rejette l'hypothèse H_0 . Une telle prise de décision est conditionnée par la connaissance exacte de ν . En utilisant le théorème 2, on peut envisager une approximation du degré de liberté ν . On montre (voir [59], expression(63)) que:

$$Proba[\chi_{s(2r-1)}^2 \leq x] \leq \lim_{N \rightarrow \infty} Proba[\hat{A}_{s,r}^2 \leq x] \leq Proba[\chi_{sr-1}^2 \leq x]$$

où x est un point de continuité de la fonction de répartition de $\kappa_{s,r}^2$. Pour éprouver pratiquement l'homogénéité de l'efficacité de la mesure de sécurité routière, on compare généralement la statistique $\hat{A}_{s,r}^2$ à la valeur α_{sr-1} . Si, la valeur calculée de $\hat{A}_{s,r}^2$ est plus grande on rejette, au niveau α , l'hypothèse d'homogénéité de l'effet dans les différents sites. D'après les inégalités ci-dessus et le théorème 2, cette méthode conduit, en fait, à rejeter à tort l'hypothèse avec une probabilité supérieure à α .

3.5.2 Approximation de la variabilité de l'effet aux différents sites

On a montré [voir proposition 2.3.1, chapitre 2] que

$$E(\hat{A}_{s,r}^2) = s(2r - 1) + \sigma^2 \sum_{k=1}^s \frac{c_k (x_{..k} + \theta c_k - 2)}{\theta (1 + \theta c_k)^2} + (r - 1) \sum_{k=1}^s \frac{c_k (\theta c_k - 1)}{\theta (1 + \theta c_k)^2} + O(N).$$

Si $\sigma^2 = 0$ alors pour tout k ($k = 1, 2, \dots, s$) $\theta_k = \theta$ presque sûrement. En utilisant l'inégalité ci-dessus, on montre que l'estimateur asymptotique ($N \rightarrow \infty$) $\hat{\sigma}^2$ de σ^2 vérifie les inégalités:

$$\frac{\hat{A}_{s,r}^2 - s(2r - 1)}{\sum_{k=1}^s \frac{c_k (x_{..k} + \theta c_k - 2)}{\theta (1 + \theta c_k)^2} + (r - 1) \sum_{k=1}^s \frac{c_k (\theta c_k - 1)}{\theta (1 + \theta c_k)^2}} \leq \hat{\sigma}^2$$

et

$$\hat{\sigma}^2 \leq \frac{\hat{A}_{s,r}^2 - (sr - 1)}{\sum_{k=1}^s \frac{c_k (x_{..k} + \theta c_k - 2)}{\theta (1 + \theta c_k)^2} + (r - 1) \sum_{k=1}^s \frac{c_k (\theta c_k - 1)}{\theta (1 + \theta c_k)^2}}$$

d'où le choix (pessimiste) préconisé pour $\hat{\sigma}^2$. Un bon choix serait de comparer pratiquement ces deux bornes et de ne retenir que celle qui donne l'estimateur le moins biaisé. Quelques résultats de simulations sont fournis dans N'Guessan et Langrand(1992).

CHAPITRE 4:

PROCÉDURES NUMÉRIQUES

La détermination d'estimations obtenues par minimisation ou maximisation de critères (maximum de vraisemblance, moindres carrés, méthodes des moments ...) est une question difficile à résoudre sous forme analytique surtout lorsque les paramètres à estimer sont soumis à des contraintes. Elle peut cependant être traitée en utilisant des procédures numériques. Différentes approches sont discutées dans la littérature [voir par exemple Minoux,1987] et sont basées sur la méthode du gradient. Ces algorithmes numériques employés de manière classique en statistique ont leurs avantages et leurs inconvénients [voir par exemple Gouriéroux et Monfort,p.481]. Ceux que nous proposons dans ce travail sont des adaptations de la méthode du gradient projeté[voir Minoux p.197]. Après une présentation générale de la structure théorique de ces algorithmes, nous donnons en deuxième partie les programmes FORTRAN et leurs descriptions.

4.1. Structure théorique

4.1.1. Approche générale

On s'intéresse, pour fixer les idées, au problème de minimisation d'une fonction numérique $F(\Theta)$ supposée continûment différentiable et définie sur un ensemble Ω de $\mathbb{R}^n$ (n est entier dépendant du nombre de paramètres à estimer). On suppose aussi que le vecteur Θ satisfait aux relations $h_i(\Theta) \leq 0$, ($i=1,2,\dots,m$), où les fonctions h_i sont supposées continûment différentiables et appelées fonctions contraintes. Ce problème peut aussi se formuler de la façon suivante:

$$(B.1) \quad \left\{ \begin{array}{l} \text{Minimiser } F(\Theta) \\ \text{sous les contraintes} \\ h_i(\Theta) \leq 0, (i=1,2,\dots,m) \\ \Theta \in \Omega \end{array} \right.$$

On note dans la suite:

$g(\Theta) = \partial F(\Theta)/\partial \Theta$, le gradient de la fonction F au point Θ et

$G(\Theta) = \partial^2 F(\Theta)/\partial \Theta \partial \Theta$, le hessien.

Un algorithme de type gradient se présente comme un processus itératif permettant d'engendrer une suite d'approximation $\Theta^{(0)}, \Theta^{(1)}, \dots, \Theta^{(q)}$ du minimum (local) recherché. Ce type d'algorithme se présente sous la forme suivante:

$$(B.2) \quad \Theta^{(q+1)} = \Theta^{(q)} + \mu^{(q)} d^{(q)}, \mu^{(q)} \in \mathbb{R}^+, d^{(q)} \in \mathbb{R}^n,$$

et est construit en deux temps. On se donne un point initial $\Theta^{(0)}$ et on détermine d'abord une direction de déplacement $d^{(q)}$ en fonction de $g(\Theta^{(q)})$ telle que:

$$(B.3) \quad d^{(q)} = - G(\Theta^{(q)})^{-1} g(\Theta^{(q)}),$$

puis on choisit le pas $\mu^{(q)}$ de manière à ce que la valeur de la fonction F soit plus petite que celle de l'étape d'avant. Toutes les variantes de cet algorithme consistent en le choix du pas $\mu^{(q)}$ et du vecteur $d^{(q)}$. Dans le cas de l'algorithme du gradient projeté, il faut en plus projeter à chaque étape le vecteur déplacement sur la frontière du domaine, afin de s'assurer que le nouveau point obtenu appartienne à Ω . Minoux [p.199, théorème8] donne sous certaines conditions un résultat théorique général qui assure l'existence d'une solution au problème (B.1).

Les principes et propriétés statistiques de cette famille d'algorithmes sont largement discutés dans la littérature [voir par exemple Gouriéroux et Monfort, ch 13]. On se restreint ici, pour simplifier la présentation, à l'estimation par la méthode du maximum de vraisemblance, mais les résultats se transposent aux autres méthodes d'estimation. Du point de vue pratique, on ne s'intéresse pas uniquement à l'estimation de l'EMVC, mais aussi à la variance asymptotique de cet estimateur. Il serait donc souhaitable que les algorithmes fournissent également la valeur de cette variance. On remplace alors dans l'expression (B.2) $G(\Theta(q))$ par son espérance mathématique:

$$(B.4) \quad G(\Theta(q)) = E[\partial^2 F(\Theta)/\partial \Theta \partial \Theta](\Theta(q)),$$

on obtient alors l'algorithme du score de Fisher [Lee et Jennrich, 1979]. Il en résulte donc que la précision estimée de l'estimateur est bien déterminée à chaque étape de l'algorithme. D'autres modifications de la matrice $G(\Theta)$ sont discutées dans la littérature [Minoux, p.114] ainsi que d'autres procédures numériques d'estimation comme l'E.M. algorithme [Chauveau (1991, 1992); Wu, 1983].

4.1.2. Approche par la bibliothèque NAG

La bibliothèque NAG (Numerical Algorithms Group) propose plusieurs sous-programmes pour minimiser une fonction numérique $F(\Theta)$ sous des contraintes d'ordre général. Ces contraintes peuvent être des contraintes de bornes (contraintes sur les composantes de Θ), des contraintes linéaires ou/et non linéaires. L'utilisateur doit fournir des programmes qui définissent la fonction à minimiser, le vecteur gradient g et les fonctions h_j . On utilise dans ce travail le sous-programme E04VCF pour rechercher les EMVC de Θ . E04VCF utilise un algorithme de Programmation Quadratique Séquentielle (SQP) dans lequel la direction de déplacement d est solution d'un problème quadratique (QP). Ce dernier problème est résolu par le sous-programme E04NAF. Ces algorithmes traitent séparément les bornes, les contraintes linéaires et non linéaires. On suppose que le problème à résoudre est de la forme:

$$(B.5) \quad \left\{ \begin{array}{l} \text{Minimiser } F(\Theta) \\ \Theta \in \mathbb{R}^n \\ \text{sous les contraintes} \\ l \leq \begin{bmatrix} \Theta \\ A_L \Theta \\ c(\Theta) \end{bmatrix} \leq u \end{array} \right. ,$$

où A_L est une matrice réelle constante, $c(\Theta)$ est le vecteur contenant les contraintes non linéaires. La matrice A_L et/ou le vecteur $c(\Theta)$ peuvent ne pas exister. Les vecteurs u et l du problème (B.5) contiennent à la fois les bornes supérieures et inférieures de toutes les contraintes du problème. Le sous-programme de recherche de la direction $d(q)$ est obtenu en résolvant un problème de la forme:

$$(B.6) \quad \left\{ \begin{array}{l} \text{Minimiser } g^T d + 1/2 d^T G d \\ d \in \mathbb{R}^n \\ \text{sous les contraintes} \\ l_1 \leq \begin{bmatrix} d \\ A_1 d \end{bmatrix} \leq u_1 \end{array} \right. .$$

Le vecteur g et la matrice G sont respectivement le gradient et le hessien de F évalués tous deux au point $\Theta(q)$. Les vecteurs l_1, u_1 et la matrice A_1 sont définis en fonction de (B.5).

La résolution de ce sous problème quadratique est obtenue en utilisant la procédure E04NAF. On pourra consulter la documentation de NAG ainsi que [Fletcher,1981; Gill et al,1981] pour une étude plus approfondie de ces algorithmes. La bibliothèque NAG propose également d'autres algorithmes plus élaborés pour rechercher à chaque itération un point $\Theta^{(q)}$ vérifiant les conditions de Kuhn-Tucker [Minoux, 1987].

4.2. Structure pratique.

4.2.1. Programme principal

Algorithme Principal

```

| BLOC.A
|   Donner les dimensions du tableau d'analyse: s,r
|   Initialisation
| BLOC.B1
|   Donner les valeurs de  $\theta$  et  $\sigma^2$ 
|   Appel du sous-programme GENENOR pour simuler les effets  $q_k$ 
|       ( $k=1,2,\dots,s$ )
|   Appel du sous-programme G05CAF de la bibliothèque NAG pour simuler
|       les probabilités  $P_{jk}$  ( $k=1,2,\dots,s;j=1,2,\dots,r$ )
|   Appel des sous-programmes G05CBF et G05DAF de NAG pour simuler
|       les coefficients de contrôle  $c_{jk}$  ( $k=1,2,\dots,s;j=1,2,\dots,r$ )
|   Calcul des vrais coefficients moyens  $c_k$  ( $k=1,2,\dots,s$ ).
| BLOC.C
|   Si MODELE = 1 alors
|       Appel du sous-programme GMULT1 pour simuler un tableau
|       2xsxr avec le modèle1.
|       Appel du sous programme E04VCF ou E04NAF de NAG pour
|       rechercher les estimateurs du maximum de vraisemblance.
|   Sinon
|       Appel du sous-programme GMULT2 pour simuler un tableau
|       2xsxr avec le modèle2.
|       Appel du sous programme E04VCF ou E04NAF de NAG pour
|       rechercher les estimateurs du maximum de vraisemblance.
|   Fin de Si.
```

¹En pratique, le BLOC.B est remplacé par un tableau de données d'accidents.

| BLOC.D

| Appel du sous-programme INFO pour estimer la variance
 | asymptotique de l'effet moyen sous l'hypothèse H_0 par inversion
 | de l'information de Fisher.
 | Appel du sous-programme ESTIMA pour calculer A^2 , discuter de son
 | comportement asymptotique et estimée σ^2 .
 | Appel du sous-programme NORMAL pour estimer le coefficient d'hété-
 | rogénéité phi et calculer la statistique de comparaison.
 FIN *Algorithme Principal*

4.2.2. Sous programmes

Sous-programme: GENENOR

a. Spécification

SUBROUTINE GENENOR (IS,IR,TETA,SIGMA,TETAK)

integer IS,IR

real TETA,SIGMA,TETAK(IS)

b. Paramètres

IS - représente le nombre s de sites expérimentaux

IR - représente le nombre r de types d'accidents

TETA - réel représentant l'effet moyen ou l'effet commun q

SIGMA - réel représentant la racine carrée de σ^2 TETAK- vecteur réel de dimension IS donnant à la sortie de GENENOR
l'effet global simulé de la mesure au site n°k: q_k ($k=1,2,\dots,IS$),
à partir d'une loi normale $N(\theta, \sigma^2)$

c. Description

Entrer

Appel G05CBF pour initialiser le générateur
 $k=1,2,\dots,IS$ $TETAK(k) \leftarrow G05DDF(TETA, SIGMA)$

fin boucle sur k

Retour.

Sous-programme: GMULT

a. Spécification

SUBROUTINE GMULT (IS,IR,TETAK,PJK,COEC,COCM,QTJK1,
NE,IXE)

integer NE(IS),IXE(IS,2*IR)

real PJK(IS,IR),COEC(IS,IR),COCM(IS,IR),QTJK1(IS,2*IR)

b. Paramètres

QTJK1 - matrice réelle de taille (IS,2*IR) contenant les vraies proba-
bilités q_{tjk} ($t=1,2,k=1,2,\dots,IS;j=1,2,\dots,IR$) de la dis-
tribution multinomiale relativement au modèle1.PJK - matrice réelle de taille (IS,IR) contenant les vraies proba-
bilités p_{jk} ($k=1,2,\dots,IS;j=1,2,\dots,IR$) avant la mesure.

- COEC - matrice réelle de taille (IS,IR) contenant les coefficients de contrôle $c_{jk}(k=1,2,\dots,IS;j=1,2,\dots,IR)$.
- COCM - vecteur réel de dimension IS contenant les vrais coefficients moyen de contrôle: $c_k(k=1,2,\dots,IS)$.
- NE - vecteur de nombres entiers de taille IS contenant, les effectifs totaux $x_{..k}^{(1)}$, à la sortie de GMULT.
- IXE - matrice de nombres entiers et de taille (IS,2*IR) contenant, à la sortie de GMULT, les effectifs $x_{tjk}^{(1)}$ d'accidents de chaque site et de chaque type ($t=1,2;j=1,2,\dots,IR;k=1,2,\dots,IS$).

c. Description

Entrer

Calcul des probabilités q_{tjk} du modèle 1

$k=1,2,\dots,IS$

$j=1,2,\dots,IR$

$QTJK1(k,j) \leftarrow PJK(k,j)/(1+TETAK(k)*COCM(k))$

$QTJK1(k,j+IR) \leftarrow TETAK(k)*COEC(k,j)*QTJK1(k,j)$

fin boucle sur j

fin boucle sur k

Choix du nombre total NE(k) d'accidents de chaque site expérimental par appel du sous-programme G05EYF de NAG.

$k=1,2,\dots,IS$

$NE(k) \leftarrow G05EYF(R1, NR1)$

fin boucle sur k

Simulation site par site du nombre d'accidents de chaque catégorie: répartition de NE(k) dans 2r classes selon une distribution multinomiale.

On découpe [0,1] en 2r classes et on note SOM(k,m) ($k=1,2,\dots,IS;m=1,2,\dots,2*IR$) les points de découpage: avec SOM(k,1)=0 et SOM(k,2*IR)=1. On tire ensuite une probabilité p avec le sous-programme G05CAF de NAG et on teste si p appartient à]SOM(k,j);,SOM(k,j+1)[.

$k=1,2,\dots,IS$

$SOM(k,1) \leftarrow 0$

$m=2,3,\dots,2*IR+1$

$SOM(k,m) \leftarrow SOM(m-1;k)+QTJK1(k,m-1)$

fin boucle sur m

fin boucle sur k

$k=1,2,\dots,IS$

(ETQ) $p \leftarrow G05CAF(X)$

$m=1,2,\dots,2*IR$

si $p \in]SOM(k,m);SOM(k,m+1)[$ alors

$IXE(k,m) \leftarrow IXE(k,m)+1$

$IXE(k,m) \leftarrow IXE(k,m)-1$

sinon

$IXE(k,m) \leftarrow IXE(k,m)$

$NE(k) \leftarrow NE(k)$

fin de si

fin boucle sur m

si ($NE(k) > 0$) aller à ETQ

fin boucle sur k
Retour.

Sous-programme: E04VCF

a. Spécification

SUBROUTINE E04VCF (itmax, msglvl, n, nclin, ncnln, nctotl, nrowa, nrowj, nrowr, bigbnd, epsaf, eta, ftol, A, BL, BU, featol, confun, OBJFUN, cold, fealin, orthog, X, istate, R, iter, C, cjac, OBJF, OBJGRD, clamda, iwork, liwork, work, lwork, ifail)

Remarque: ce sous-programme de NAG permet d'optimiser un critère (minimiser ou maximiser une fonction réelle) sous des contraintes d'ordre général. Il est utilisé dans ce travail pour rechercher les estimateurs du maximum de vraisemblance sous contraintes de Θ par résolution d'un problème de type (B.5). Il existe d'autres procédures numériques permettant de traiter le même problème. Nous donnons plus loin une deuxième procédure. Toutes ces procédures de NAG peuvent être remplacées par des sous-programmes de SAS ou GLIM permettant d'optimiser un critère.

b. Paramètres

b.1. Paramètres d'entrée

ITMAX	- Nombre maximal d'itérations nécessaire pour la recherche de l'EMVC.
MSGLVL	- Permet de filtrer les sorties; par exemple MSGLVL=1 ne fournira que la solution optimale.
N	- Nombre de paramètres à estimer.
NCLIN	- Nombre de contraintes linéaires associées au problème.
NCNLN	- Nombre de contraintes non linéaires.
NCTOTL	- est la somme N+NCLIN+NCNLN.
NROWA	- Nombre de lignes de la matrice A [voir problème (B.5)]; $NROWA \geq \max(1, NCLIN)$.
NROWJ	- Nombre de lignes de la matrice Jacobienne associée aux contraintes non linéaires; $NROWJ \geq \max(1, NCNLN)$.
NROWR	- Nombre de ligne du facteur R de Cholesky associé au hessien de la fonction à minimiser; $NROWR \geq N$.
BIGBND	- Cette valeur est associée à l'infini.
EPSAF	- L'erreur absolue commise sur le calcul de la fonction F au point initial.
ETA	- La valeur initiale de $\mu^{(q)}$ permettant de minimiser au mieux la fonction F le long du vecteur $g^{(q)}$. $ETA \approx 0,9$ pour des problèmes avec contraintes non linéaires.
FTOL	- seuil de tolérance indiquant la précision exigée sur le calcul de la fonction F au point optimal.
A	- Matrice des contraintes linéaires de taille NROWA et N représentant la matrice A_L du problème (B.5). Si NCLIN est nul la matrice A n'existe pas.
BL	- Vecteur réel de dimension NCTOTL contenant toutes les bornes inférieures du problème à traiter.
BU	- Vecteur réel de dimension NCTOTL contenant toutes les

FEATOL	bornes supérieures du problème à traiter. - Vecteur réel dimension NCTOTL contenant tous les seuils de tolérance qui définissent la violation maximale permise sur chaque contrainte afin qu'un point soit considéré comme réalisable.
CONFUN	- Sous-programme fourni par l'utilisateur et permettant de calculer les contraintes non linéaires et le jacobien associé. SUBROUTINE CONFUN(mode, ncnln, n, nrowj, x,c, cjac, nstate)
MODE	- Entier naturel indiquant une quelconque défaillance dans l'évaluation des contraintes non linéaires.
X	- Réel de dimension N représentant le point en lequel la vraisemblance (le critère) est évaluée. Ce point donne une estimation de la solution au problème. Sa valeur initiale est très importante.
C	- Réel de dimension NROWJ contenant les valeurs des contraintes non linéaires c_j ($j=1,2,\dots,NCNLN$).
CJAC	- Matrice réelle de taille (NROWJ,N) représentant la matrice jacobienne associée aux contraintes non-linéaires.
OBJFUN	- Sous-programme fourni par l'utilisateur et permettant de définir la fonction à minimiser et le gradient associé. SUBROUTINE OBJFUN(mode, n, x, objf, objgrd, nstate)
OBJF	- Réel contenant la valeur de $F(\Theta)$.
OBJGRD	- réel de dimension N contenant le vecteur $g(\Theta)$.

b.2. Paramètres de sortie

ITER	- Entier naturel désignant le nombre d'itérations nécessaire à l'obtention du minimum local.
CLAMDA	- Vecteur réel de dimension NCTOTL contenant les multiplicateurs de Lagrange associés à chaque contraintes.
IWORK, LIWORK, WORK, LWORK:	sont des paramètres de travail spécifiques à E04VCF. Une description complète est

dispo-

IFAIL	nible dans la documentation NAG. - Entier indiquant la nature de l'erreur pouvant arriver au cours de la procédure de minimisation.
-------	--

Remarque: Une description plus complète des paramètres associés à la procédure E04VCF est disponible dans la documentation NAG.

4.4. APPLICATIONS

4.4.1 Description générale des procédures numériques

On considère l'estimation de l'effet de la mesure sous l'aspect pratique. Comme nous ne disposons pas de données réelles, nous présenterons quelques exemples avec des données simulées.

Nous avons simulé des tableaux d'accidents correspondant à un nombre de sites s et pour différentes valeurs de r qui correspondent au nombre de types d'accidents étudiés. Afin de n'être pas obligé de rentrer ces différentes données "à la main", nous avons utilisé un sous-programme qui engendre les effectifs totaux d'accidents par site expérimental suivant une loi de Poisson.

En ce qui concerne les sites de contrôle, seuls les coefficients de contrôle c_{jk} - qui représentent le quotient du nombre d'accidents après par le nombre d'accidents avant dans la zone de contrôle k et pour le type j - ont été introduits (sous-programmes utilisés

G05CBF, G05DAF); les valeurs prises par ces coefficients ont été limitées à l'intervalle $[0,5 ; 2,5]$.

Le programme G05CAF génère les probabilités p_{jk} qui représentent, pour un accident qui se produit avant l'instauration de la mesure dans le site k , la probabilité pour qu'il soit de type j .

Pour $k=1,2,\dots,s$ l'effet réel θ_k de la mesure au site k est supposé de loi Normale:

$$N(\theta, \sigma^2)$$

avec $\theta = 0,5$ et où σ^2 a été choisie petite de manière à ce que les θ_k soient strictement positifs (sous-programme GENENOR).

Le programme introduit ensuite le calcul des c_k :

$$c_k = \sum_{j=1}^r c_{jk} p_{jk}$$

et celui des probabilités q_{tjk} définies par:

$$(1) \text{ Modèle1: } q_{1jk}(\theta_k, P_k) = \frac{p_{jk}}{1 + \theta_k c_k}, \quad q_{12k}(\theta_k, P_k) = \frac{\theta_k c_k p_{jk}}{1 + \theta_k c_k} ; (j=1,2,\dots,r)$$

$$(2) \text{ Modèle2: } q_{1jk}(\theta_k, P_k) = \frac{p_{jk}}{1 + \theta_k c_k}, \quad q_{12k}(\theta_k, P_k) = \frac{\theta_k c_k p_{jk}}{1 + \theta_k c_k} ; (j=1,2,\dots,r)$$

(Sous-programme GMULT).

L'estimation de l'effet par la résolution des équations de vraisemblance, s'effectue à l'aide de E04VCF ou E04NAF et utilise la méthode du gradient projeté.

L'approximation de la variance asymptotique de l'estimateur de l'effet par inversion de l'information de Fisher est faite à l'aide du sous-programme INFO.

Le programme introduit l'étude du comportement asymptotique de la statistique du chapitre 3 : sous-programme KHIDEUX.

L'analyse de l'effet de la mesure en cas de non similarité s'effectue à l'aide du sous-programme NORMAL.

Le but principal de cette simulation est d'estimer (résolution des équations de vraisemblance: chapitre 2) les paramètres des deux modèles et de comparer ces modèles en terme d'estimations et de nombre d'itérations. On compare ensuite l'estimation de l'effet à celui donné par l'estimateur empirique (EE) :

$$\theta^* = \frac{1}{s} \sum_{k=1}^s \theta_k .$$

Les résultats concernant les sous-programmes INFO, KHIDEUX et NORMAL feront l'objet de nos prochaines investigations. Nous avons simulé deux ensembles de données. Une première étude (paragraphe 4.4.2) a été faite avec

$$\theta = 0,5 \text{ et } \sigma = 0,001$$

afin d'obtenir des valeurs de θ_k strictement positifs et de travailler sous H_0 . Une deuxième étude (paragraphe 4.4.3) a été effectuée avec

$$\theta = 0,5 \text{ et } \sigma = 0,160.$$

Ces premières simulations montrent que les deux modèles donnent sensiblement les mêmes valeurs du point de vue de l'estimation de l'effet.

4.4.2 Estimation de l'effet de la mesure sur données simulées

Le tableau suivant rassemble les estimations obtenues pour les différents cas étudiés.

Tableau 1 (Estimations de l'effet)						
	r=2		r=3		r=4	
	Modèle1	Modèle2	Modèle1	Modèle2	Modèle1	Modèle2
s=1	0,376	0,507	0,427	0,429	0,625	0,633
s=2	0,507	0,515	0,469	0,469	0,499	0,503
s=3	0,488	0,488	0,492	0,483	0,529	0,524
s=4	0,468	0,472	0,494	0,491	0,529	0,526
s=5	0,506	0,511	0,538	0,540	0,530	0,530
s=6	0,482	0,483	0,502	0,500	0,488	0,489
s=7	0,505	0,505	0,518	0,519	0,493	0,497
s=8	0,488	0,485	0,503	0,501	0,504	0,505
s=9	0,516	0,516	0,523	0,520	0,505	0,507
s=10	0,488	0,493	0,493	0,503	0,524	0,533
s=11	0,493	0,493	0,478	0,480	0,489	0,492
s=12	0,507	0,508	0,519	0,525	0,508	0,517
s=13	0,513	0,516	0,504	0,506	0,516	0,511
s=14	0,490	0,492	0,472	0,472	0,513	0,515
s=15	0,486	0,486	0,501	0,501	0,486	0,487
s=16	0,510	0,511	0,512	0,516	0,507	0,512
s=17	0,494	0,497	0,499	0,500	0,487	0,496
s=18	0,503	0,505	0,503	0,502	0,499	0,495
s=19	0,495	0,498	0,504	0,504	0,495	0,491
s=20	0,525	0,531	0,512	0,508	0,527	0,530

On note alors que les deux modèles donnent sensiblement les mêmes valeurs du point de vue de l'estimation de l'effet .

L'exemple que nous présentons ci-après a été effectué avec le modèle 1. Le nombre de sites expérimentaux s est égal à 4 et le nombre r de type d'accidents à 2 (accidents à tués et accidents à blessés). La vraie valeur de l'effet est

$\theta = 0,5.$

Les vraies valeurs des risques p_{jk} , des coefficients c_{jk} et des coefficients moyens c_k sont rassemblées dans le tableau 2:

	r = 2				
	p_{jk}		c_{jk}		c_k
	tués	blessés	tués	blessés	
site n°1	0,604	0,395	1,270	2,451	1,737
site n°2	0,482	0,519	2,451	0,812	1,601
site n°3	0,186	0,814	1,631	1,173	1,258
site n°4	0,413	0,587	0,812	1,534	1,236

Par exemple pour le site n°1, on a 6 chances sur 10 (0,604) d'être tué et 4 chances sur 10 (0,395) d'être blessé dans un accident survenu avant la prise de la mesure. Après l'instauration de la mesure, on enregistre 27% d'augmentation des accidents à tués

($c_{jk} = 1,27$) et plus 145% ($c_{jk} = 2,451$) des accidents à blessés dans la zone de contrôle. Le coefficient de contrôle moyen

$$c_1 = 0,604 \times 1,27 + 0,395 \times 2,451 = 1,737$$

montre une augmentation moyenne de 73% des accidents mortels dans la zone de contrôle n°1.

A partir du tableau 2 et de l'effet réel, on calcule les probabilités q_{ijk} . Ensuite, on simule le tableau 3 ci-dessous comme l'indique la description de l'algorithme principal:

Tableau 3
(accidents graves)

Avant		Après		
Tués	Blessés	Tués	Blessés	Total
81	38	34	64	217
37	28	35	15	115
28	130	24	67	249
32	58	12	45	147

On note que le site n°1 enregistre 81 accidents à tués avant la mesure et 34 après soit plus de 50% de réduction alors que les accidents à blessés sont passés de 38 avant à 64 après la mesure. On est passé de 119 accidents au total avant la mesure (site n°1) à 98 après. Indépendamment donc du fait que les données soient simulées, on peut émettre quelques interprétations. En effet, l'évolution des accidents mortels laisse penser que la mesure a eu un effet favorable pour les accidents à tués et défavorable pour les accidents à blessés. Ce gain de vie humaine peut s'expliquer par un changement (dû peut-être à la mesure) du comportement des usagers de la route ou par une variation du trafic entre les périodes avant et après. On peut également avancer l'hypothèse d'un phénomène de migration d'accidents à la fois de tués vers blessés et de site traité vers site non traité appartenant à la zone de contrôle n°1. Il est donc important que toute conclusion éventuelle tienne compte de facteurs locaux.

Dans la pratique, on ne dispose que du tableau 3 et des c_{jk} du tableau 2. A partir donc des valeurs des c_{jk} et du tableau 3, on estime les paramètres θ et p_{jk} en se donnant une valeur initiale de ces paramètres. On obtient, après 11 itérations, une valeur de 0,47 comme estimation de l'effet moyen et le tableau 4 ci-dessous comme estimations des p_{jk} et des c_k .

Tableau 4 (Estimations des p_{jk} et des c_k)

p_{jk}		c_k
0,603	0,397	1,739
0,518	0,482	1,661
0,188	0,812	1,259
0,347	0,653	1,283

On obtient une bonne estimation de l'effet ainsi que des p_{jk} des trois premiers sites.

4.4.3 Comparaison des modèles

Les tableaux 5 à 10 représentent les estimations obtenues pour différentes valeurs de s et de r . Le tableau 5, par exemple, donne le cas $s = 3$ et $r = 2$ pour les deux modèles. On a 7 paramètres

$$(\theta, p_{11}, p_{21}, p_{12}, p_{22}, p_{13}, p_{23})^T$$

à estimer. La colonne "Point de Départ: (Dép)" est la valeur initiale fournie au sous-programme E04VCF. Après 9 itérations [voir Iter (tableau 11)] le modèle 1 donne

$$(0,4869; 0,5085; 0,4915; 0,2642; 0,7360; 0,9064; 0,0935)^T$$

comme estimation des paramètres. Alors que le modèle 2, avec le même point de départ, donne après 8 itérations [voir Iter (tableau 12)]

$$(0,4827; 0,5204; 0,4796; 0,2756; 0,7244; 0,9113; 0,0887)^T$$

comme estimation des paramètres. La colonne "Vraie valeur" représente la vraie valeur des paramètres pour chacun des deux modèles.

Le tableau 1-5 résume les vraies valeurs des paramètres p_{jk} , des coefficients c_{jk} et c_k utilisés pour simuler les tableaux d'accidents donnés par les tableaux 2-5 et 3-5. On remarque que les nombres d'accidents avant l'instauration de la mesure sont identiques pour les deux modèles alors que les accidents après diffèrent. Cela est dû à la construction des modèles. Ces données des tableaux 2-5 et 3-5 sont utilisées pour rechercher une estimation de l'effet de la mesure et des probabilités p_{jk} .

D'une façon générale, les deux modèles donnent sensiblement les mêmes résultats avec un nombre d'itérations très peu différents. On a comparé aussi, sur quelques exemples, l'estimateur empirique (EE) de l'effet de la mesure à l'estimateur donné par la méthode proposée (MV) [voir tableaux 11 et 12]. L'EE semble beaucoup plus stable que le nôtre. Malheureusement l'on ne dispose pas des valeurs de l'effet en pratique. Par conséquent, il est impossible d'utiliser l'EE comme estimateur de l'effet de la mesure de sécurité routière.

Tableau 1-5 (vraies valeurs des p_{jk} , c_{jk} et c_k pour $s = 3$ et $r = 2$)

	P _{jk}		C _{jk}		C _k
	j = 1	j = 2	j = 1	j = 2	
k = 1	0,500	0,490	1,270	2,451	1,849
k = 2	0,313	0,687	2,451	0,812	1,324
k = 3	0,900	0,100	1,631	1,173	1,585

Tableau 2-5 (Accidents simulés avec le modèle 1)

s = 3	r = 2				Total
	Avant		Après		
	j = 1	j = 2	j = 1	j = 2	
k = 1	48	48	33	58	187
k = 2	43	125	58	54	280
k = 3	116	15	83	3	217

Tableau 3-5 (Accidents simulés avec le modèle 2)

s = 3	r = 2				Total
	Avant		Après		
	j = 1	j = 2	j = 1	j = 2	
k = 1	48	48	50	41	187
k = 2	43	125	33	79	280
k = 3	116	15	82	4	217

Tableau 5

r = 2						
Modèle 1				Modèle 2		
s = 3	Point de départ	Estimateur	Vraie valeur	Point de départ	Estimateur	Vraie valeur
	0,2000	0,4869	0,5000	0,2000	0,4827	0,5000
	0,5000	0,5085	0,5100	0,5000	0,5204	0,5101
	0,5000	0,4915	0,4900	0,5000	0,4796	0,4899
	0,5000	0,2642	0,3127	0,5000	0,2756	0,3127
	0,5000	0,7360	0,6872	0,5000	0,7244	0,6873
	0,5000	0,9064	0,8999	0,5000	0,9113	0,8999
	0,5000	0,0935	0,1000	0,5000	0,0887	0,1000

Tableau 6

r = 2						
Modèle 1				Modèle 2		
s = 5	Point de départ	Estimateur	Vraie valeur	Point de départ	Estimateur	Vraie valeur
	0,2800	0,5091	0,5000	0,2800	0,5073	0,5000
	0,5000	0,1726	0,1503	0,5000	0,1704	0,1503
	0,5000	0,8274	0,8497	0,5000	0,8296	0,8497
	0,5000	0,5646	0,5606	0,5000	0,5634	0,5606
	0,5000	0,4354	0,4394	0,5000	0,4366	0,4394
	0,5000	0,5916	0,5646	0,5000	0,5973	0,5646
	0,5000	0,4084	0,4354	0,5000	0,4026	0,4353
	0,5000	0,6238	0,5699	0,5000	0,5982	0,5699
	0,5000	0,3761	0,4301	0,5000	0,4017	0,4301
	0,5000	0,3660	0,3666	0,5000	0,3632	0,3666
	0,5000	0,6340	0,6334	0,5000	0,6368	0,6334

Tableau 7

r = 2						
	Modèle 1			Modèle 2		
	Point de départ	Estimateur	Vraie valeur	Point de départ	Estimateur	Vraie valeur
s = 8	0,2800	0,4948	0,5000	0,2800	0,4907	0,5000
	0,5000	0,3000	0,2741	0,5000	0,3130	0,2741
	0,5000	0,6998	0,7258	0,5000	0,6870	0,7258
	0,5000	0,5614	0,5377	0,5000	0,5946	0,5377
	0,5000	0,4385	0,4623	0,5000	0,4053	0,4623
	0,5000	0,1650	0,1633	0,5000	0,1626	0,1633
	0,5000	0,8350	0,8366	0,5000	0,8374	0,8366
	0,5000	0,4886	0,4625	0,5000	0,5077	0,4625
	0,5000	0,5113	0,5375	0,5000	0,4923	0,5375
	0,5000	0,2620	0,2725	0,5000	0,2621	0,2725
	0,5000	0,7380	0,7275	0,5000	0,7379	0,7275
	0,5000	0,5858	0,5531	0,5000	0,5494	0,5531
	0,5000	0,4141	0,4468	0,5000	0,4506	0,4468
	0,5000	0,3454	0,3011	0,5000	0,3537	0,3011
	0,5000	0,6545	0,6988	0,5000	0,6463	0,6988
	0,5000	0,7198	0,6972	0,5000	0,7177	0,6972
	0,5000	0,2801	0,3027	0,5000	0,2823	0,3027

Tableau 8

r = 3						
	Modèle 1			Modèle 2		
	Point de départ	Estimateur	Vraie valeur	Point de départ	Estimateur	Vraie valeur
s = 3	0,2800	0,4360	0,5000	0,2800	0,4447	0,5000
	0,3333	0,4274	0,4519	0,3333	0,4183	0,4519
	0,3333	0,3736	0,4340	0,3333	0,3727	0,4340
	0,3333	0,1990	0,1141	0,3333	0,2089	0,1141
	0,3333	0,5420	0,5311	0,3333	0,5074	0,5311
	0,3333	0,4079	0,4219	0,3333	0,4425	0,4220
	0,3333	0,0500	0,0469	0,3333	0,5000	0,0469
	0,3333	0,2581	0,2800	0,3333	0,2484	0,2800
	0,3333	0,5000	0,0266	0,3333	0,0500	0,0266
	0,3333	0,6920	0,6933	0,3333	0,7016	0,6933

Tableau 9

r = 3						
	Modèle 1			Modèle 2		
	Point de départ	Estimateur	Vraie valeur	Point de départ	Estimateur	Vraie valeur
s = 5	0,2800	0,5134	0,5000	0,2800	0,5134	0,5000
	0,3333	0,0500	0,0686	0,3333	0,0592	0,0686
	0,3333	0,4155	0,3878	0,3333	0,4033	0,3879
	0,3333	0,5345	0,5435	0,3333	0,5374	0,5435
	0,3333	0,2009	0,2301	0,3333	0,2165	0,2301
	0,3333	0,4165	0,4347	0,3333	0,4022	0,4347
	0,3333	0,3825	0,3351	0,3333	0,3813	0,3351
	0,3333	0,3424	0,3435	0,3333	0,3297	0,3435
	0,3333	0,2414	0,2593	0,3333	0,2470	0,2593
	0,3333	0,4162	0,3972	0,3333	0,4232	0,3972
	0,3333	0,5814	0,5908	0,3333	0,5734	0,5908
	0,3333	0,0500	0,0203	0,3333	0,0500	0,0203
	0,3333	0,3686	0,3888	0,3333	0,3765	0,3888
	0,3333	0,0864	0,1236	0,3333	0,0811	0,1236
	0,3333	0,4556	0,3998	0,3333	0,4595	0,3998
	0,3333	0,4580	0,4766	0,3333	0,4593	0,4766

Tableau 10

r = 4						
	Modèle 1			Modèle 2		
	Point de départ	Estimateur	Vraie valeur	Point de départ	Estimateur	Vraie valeur
s = 3	0,2000	0,4972	0,5000	0,2000	0,4967	0,5000
	0,5000	0,3695	0,3613	0,5000	0,3647	0,3613
	0,5000	0,3801	0,3470	0,5000	0,3906	0,3470
	0,5000	0,0788	0,0912	0,5000	0,0674	0,0912
	0,5000	0,1715	0,2004	0,5000	0,1771	0,2004
	0,5000	0,4140	0,4436	0,5000	0,4211	0,4436
	0,5000	0,0500	0,0493	0,5000	0,0500	0,0493
	0,5000	0,4860	0,4630	0,5000	0,4788	0,4630
	0,5000	0,0500	0,0440	0,5000	0,0500	0,0440
	0,5000	0,3170	0,2914	0,5000	0,3270	0,2914
	0,5000	0,1375	0,1575	0,5000	0,1344	0,1575
	0,5000	0,2526	0,2022	0,5000	0,2280	0,2022
	0,5000	0,2929	0,3488	0,5000	0,3106	0,3489

Tableau 11: Modèle 1

	Dép			Iter			MV			EE		
	r = 2	r = 3	r = 4	r = 2	r = 3	r = 4	r = 2	r = 3	r = 4	r = 2	r = 3	r = 4
s = 3	0,20	0,28	0,20	9	11	16	0,49	0,43	0,49	0,47	0,47	0,47
s = 5	0,28	0,28	0,28	11	18	26	0,51	0,51	0,52	0,50	0,50	0,50
s = 8	0,28	0,20	0,28	14	28	40	0,49	0,49	0,48	0,50	0,50	0,50

Tableau 12: Modèle 2

	Dép			Iter			MV			EE		
	r = 2	r = 3	r = 4	r = 2	r = 3	r = 4	r = 2	r = 3	r = 4	r = 2	r = 3	r = 4
s = 3	0,20	0,28	0,20	9	11	16	0,49	0,43	0,49	0,47	0,47	0,47
s = 5	0,28	0,28	0,28	11	18	26	0,51	0,51	0,52	0,50	0,50	0,50
s = 8	0,28	0,20	0,28	14	28	40	0,49	0,49	0,48	0,50	0,50	0,50

ANNEXES

ANNEXE 2

A.2.1. Approximation de la distribution asymptotique de $\hat{\theta}$ et $\log \hat{\theta}$ sous l'hypothèse $H_0 : \theta_1 = \theta_2 = \dots = \theta_s = \theta$.

On s'intéresse à la distribution asymptotique ($N \rightarrow \infty$) de $\hat{\theta}$ et de $\log \hat{\theta}$ où $\hat{\theta}$ est la première composante de Θ et où les p_{jk} ($k = 1, 2, \dots, r; j = 1, 2, \dots, r$) sont supposées connues. On sait d'après les propriétés des EMV (voir Monfort, p. 105) que $\log \hat{\theta}$ est un estimateur du maximum de vraisemblance de $\log \theta$. De plus, $\hat{\theta}$ et $\log \hat{\theta}$ sont asymptotiquement normalement distribués. Par conséquent, pour construire la statistique permettant d'éprouver l'hypothèse $\theta = 1$, il faut faire un choix entre $\hat{\theta}$ et $\log \hat{\theta}$. Le choix que nous faisons ci-dessous est basé sur la comparaison des coefficients d'asymétrie (skewness) et d'aplatissement (kurtosis) (voir Saporta, p. 29). La connaissance de ces coefficients nécessite celle des quatre premiers moments de $\hat{\theta}$ et $\log \hat{\theta}$. On propose alors une méthode d'approximation pour y parvenir. On pose pour cela $z = \log \hat{\theta} - \log \theta$ et on suppose que :

$$\forall \varepsilon > 0, \exists N_\varepsilon \in \mathbb{N}^* \text{ tel que } \forall N \geq N_\varepsilon \text{ Proba}[e^z - 1| \geq \varepsilon] \leq \varepsilon.$$

En remplaçant $\hat{\theta}$ par θe^z dans l'équation (2.3.1), on a

$$(2.3.1.) \quad 0 = \sum_{k=1}^s \sum_{j=1}^r X_{1jk} - \sum_{k=1}^s \frac{x_{..k}}{1 + \theta c_k e^z}.$$

On fait un développement en probabilité en puissance de z jusqu'à l'ordre 4 du deuxième membre de droite. $\forall N \geq N_\varepsilon$, on a

$$0 = \sum_{k=1}^r \sum_{j=1}^r X_{1jk} - \sum_{k=1}^s \frac{x_{..k}}{1 + \theta c_k} \left[1 - \frac{\theta c_k (e^z - 1)}{1 + \theta c_k} + \left(\frac{\theta c_k (e^z - 1)}{1 + \theta c_k} \right)^2 - \left(\frac{\theta c_k (e^z - 1)}{1 + \theta c_k} \right)^3 + \left(\frac{\theta c_k (e^z - 1)}{1 + \theta c_k} \right)^4 + R(z) \right]$$

$$\text{où } R(z) = \sum_{m \geq 5} (-1)^m \left(\frac{\theta c_k}{1 + \theta c_k} \right)^m (e^z - 1)^m.$$

$\forall N \geq N_\varepsilon$, $\text{Proba}[|e^z - 1| \geq \varepsilon]$. Ce qui implique que

$$\forall N \geq N_\varepsilon, |R(z)| \leq \sum_{m \geq 5} \left(\frac{\theta c_k}{1 + \theta c_k} \right)^m \leq \sum_{m=0}^{\infty} \left(\frac{\theta c_k}{1 + \theta c_k} \right)^m \leq M_\varepsilon$$

où $M_\varepsilon = 1 + \theta \sup_{k=1,2,\dots,s} c_k$. On en déduit que $R(z) = O_p(1)$. Après quelques simplifications, on montre que

$$\xi = a_1 z + \left(\frac{a_1}{2} - a_2 \right) z^2 + \left(\frac{a_1}{6} - a_2 + a_3 \right) z^3 + \left(\frac{a_1}{24} - \frac{7}{12} a_2 + \frac{3}{2} a_3 - a_4 \right) z^4 + O_p(1),$$

avec

$$\begin{aligned}\xi &= \sum_{k=1}^s \sum_{j=1}^r \left\{ \frac{X_{2jk} - \theta c_k X_{1jk}}{1 + \theta c_k} \right\}, a_1 = \sum_{k=1}^s \frac{\theta c_k x_{..k}}{(1 + \theta c_k)^2}, a_2 = \sum_{k=1}^s \frac{\theta^2 c_k^2 x_{..k}}{(1 + \theta c_k)^3} \\ a_3 &= \sum_{k=1}^s \frac{\theta^3 c_k^3 x_{..k}}{(1 + \theta c_k)^4}, a_4 = \sum_{k=1}^s \frac{\theta^4 c_k^4 x_{..k}}{(1 + \theta c_k)^5}. \\ \forall N \geq N_\varepsilon \quad z &= A_1 \xi + A_2 \xi^2 + A_3 \xi^3 + A_4 \xi^4 + O_p(1),\end{aligned}$$

où (voir Tanner, p. 340)

$$\begin{aligned}A_1 &= 1/a_1 \\ A_2 &= (a_2 - \frac{1}{2}a_1)/a_1^3 \\ A_3 &= (6a_1 - 3a_1a_2 + a_1^2 - 3a_1a_3)/3a_1^5 \\ A_4 &= (-\frac{1}{4}a_1^3 + a_1^2a_2 + a_1^2a_3 + a_1^2a_4 - 5a_1a_2a_3 - \frac{5}{2}a_1a_2^2 + 5a_2^3)/a_1^7.\end{aligned}$$

Pour obtenir une approximation des quatre premiers moments de z , il faut connaître ceux de ξ . On réécrit pour cela ξ sous la forme $\xi = -\sum_{k=1}^s [U_k - E(U_k)]$, où $U_k = \sum_{j=1}^r X_{1jk}$.

D'après les propriétés de la distribution multinomiale (Bishop et al, p. 442-443) et en utilisant les relations $\sum_{j=1}^r p_{jk} = 1$ et $c_k = \sum_{j=1}^r c_{jk} p_{jk}$ (par construction des modèles 1 et 2), on montre que

$$\begin{aligned}E(U_k) &= \frac{x_{..k}}{1 + \theta c_k}, E[U_k - E(U_k)]^2 = \frac{\theta c_k x_{..k}}{(1 + \theta c_k)^2}, \\ E(U_k - EU_k)^3 &= -\frac{\theta c_k x_{..k}}{(1 + \theta c_k)^2} + 2\frac{\theta^2 c_k^2 x_{..k}}{(1 + \theta c_k)^3} \\ E(U_k - EU_k)^4 &= 3 \left[\frac{\theta c_k x_{..k}}{(1 + \theta c_k)^2} \right]^2 + \frac{\theta c_k x_{..k}}{(1 + \theta c_k)^2} \left[1 - 6\frac{\theta c_k}{(1 + \theta c_k)^2} \right].\end{aligned}$$

En utilisant l'indépendance des U_k et le fait que $E(U_k - EU_k) = 0$, ($k = 1, 2, \dots, s$), on obtient :

$$\begin{aligned}E(\xi) &= 0, E(\xi^2) = a_1, E(\xi^3) = a_1 - 2a_2, \\ E(\xi^4) &= a_1 - 6a_2 + 6a_3 + 3a_1^2.\end{aligned}$$

On en déduit alors qu'une approximation des coefficients d'asymétrie γ_1 et d'aplatissement γ_2 de $\log \hat{\theta}$ est donnée par :

$$\gamma_1(\log \hat{\theta}) = \frac{2(2a_2 - a_1)}{a_1^{3/2}} + O(N^{-3/2}), \gamma_2(\log \hat{\theta}) = 3 + O(N^{-1}).$$

Pour estimer γ_1 et γ_2 de $\hat{\theta}$, on pose $y = \frac{\hat{\theta} - \theta}{\theta}$ c'est-à-dire $y = e^z - 1$. En utilisant l'approximation des quatre premiers moments de z , on montre que :

$$\gamma_1(\hat{\theta}) = \frac{4a_2 + a_1}{a_1^{3/2}} + O(N^{-3/2}), \gamma_2(\hat{\theta}) = 3 + O(N^{-1}).$$

Ces approximations montrent, lorsque $N \rightarrow \infty$, que les coefficients d'asymétrie et d'aplatissement de $\log \hat{\theta}$ et $\hat{\theta}$ tendent vers ceux de la loi normale (voir Saporta, p. 29) c'est-à-dire vers 0 et 3. On note aussi que $\gamma_2(\hat{\theta})$ et $\gamma_2(\log \hat{\theta})$ sont du même ordre de grandeur alors que $\gamma_1(\log \hat{\theta})$ est numériquement plus petit que $\gamma_1(\hat{\theta})$ (voir Tanner, p. 341-342 pour le détail). D'autre part, en calculant

$$E(z) = A_1 E(\xi) + A_2 E(\xi)^2 + A_3 E(\xi^3) + A_4 E(\xi^4),$$

on montre que

$$E(\log \hat{\theta} - \log \theta) = O(N^{-1})$$

($\log \hat{\theta}$ est un estimateur asymptotiquement sans biais de $\log \theta$). On choisit, par conséquent d'utiliser $\log \hat{\theta}$ pour la construction du test de l'hypothèse d'efficacité nulle.

A.2.2. Estimation des paramètres et information de Fisher dans le cadre du modèle 2.

La log-vraisemblance relativement au modèle 2 est donnée à une constante additive près par :

$$L(\Theta) = \sum_{k=1}^s \sum_{j=1}^r \{x_{\cdot jk} \log p_{jk} + x_{2jk} \log c_k + x_{2jk} \log \theta - x_{\cdot jk} \log(1 + \theta c_k)\}.$$

En dérivant cette fonction aléatoire par rapport aux éléments du point Θ et en annulant les dérivées partielles, on montre que l'EMVC de Θ vérifie le système d'équation non linéaires :

$$\left\{ \begin{array}{l} \sum_{k=1}^s \sum_{j=1}^r \left\{ \frac{x_{2jk}}{\hat{\theta}} - \frac{x_{\cdot jk} \hat{c}_k}{1 + \hat{\theta} \hat{c}_k} \right\} = 0 \\ \frac{x_{\cdot jk}}{\hat{p}_{jk}} - \frac{\hat{\theta} c_{jk} x_{\cdot \cdot k}}{1 + \hat{\theta} \hat{c}_k} + \frac{x_{2 \cdot k} c_{jk}}{\hat{c}_k} = 0 \quad (k = 1, 2, \dots, s; j = 1, 2, \dots, r) \\ \hat{\theta} \in \mathcal{V} \end{array} \right.$$

avec $\hat{c}_k = \sum_{\theta=1}^r c_{jk} \hat{p}_{jk}$. La matrice de variances- covariances est obtenue comme au paragraphe 2.3.3 où la matrice $\mathbb{J}$ est identique à celle du modèle 1 sauf pour les sous-matrices $\mathbb{J}_{P_k P_k}$ ($k = 1, 2, \dots, s$) dont les éléments sont donnés par :

$$E \left[\left(\frac{\partial^2 L}{\partial p_{jk} \partial P_{mk}} \right) \right] = \begin{cases} x_{..k} \left[\frac{1}{p_{mk}} + \frac{\theta^2 c_{mk}^2}{c_k (1 + \theta c_k)^2} \right], & \text{si } j = m \\ x_{..k} \left(\frac{\theta c_{jk} c_{mk}}{c_k (1 + \theta c_k)^2} \right), & \text{si } j \neq m \end{cases}$$

($j = 1, 2, \dots, r; m = 1, 2, \dots, r$).

ANNEXE - A3

A3.1 - MÉTHODE ITÉRATIVE DE CALCUL DU RANG DE B_0 A3.1.a - Rang de $B^{(k)}$

Par définition $B^{(k)} = ((\frac{\partial Q_k}{\partial p_{1k}}, \frac{\partial Q_k}{\partial p_{2k}}, \dots, \frac{\partial Q_k}{\partial p_{rk}}))$ la matrice de taille $(2r, r)$ donnant les dérivées partielles du vecteur Q_k par rapport à $p_{1k}, p_{2k}, \dots, p_{rk}$. De façon plus détaillée, on a :

$$B^{(k)} = \begin{bmatrix} \frac{\partial q_{11k}}{\partial p_{1k}} & \dots & \frac{\partial q_{11k}}{\partial p_{rk}} \\ \vdots & & \vdots \\ \frac{\partial q_{1rk}}{\partial p_{1k}} & \dots & \frac{\partial q_{1rk}}{\partial p_{rk}} \\ \frac{\partial q_{21k}}{\partial p_{1k}} & \dots & \frac{\partial q_{21k}}{\partial p_{rk}} \\ \vdots & & \vdots \\ \frac{\partial q_{2rk}}{\partial p_{1k}} & \dots & \frac{\partial q_{2rk}}{\partial p_{rk}} \end{bmatrix}$$

MODÈLE 1

Sous l'hypothèse $\theta_1 = \theta_2 = \dots = \theta_s$, on a pour $j = 1, 2, \dots, r$ et $m = 1, 2, \dots, r$ ($\partial q_{2jk} / \partial p_{mk} = \theta c_{jk} (\partial q_{1jk} / \partial p_{mk})$ (car $q_{2jk} = \theta c_{jk} q_{1jk}$). Par suite on a la relation :

$$\begin{bmatrix} \frac{\partial q_{21k}}{\partial p_{1k}} & \dots & \frac{\partial q_{21k}}{\partial p_{rk}} \\ \vdots & \ddots & \vdots \\ \frac{\partial q_{2rk}}{\partial p_{1k}} & \dots & \frac{\partial q_{2rk}}{\partial p_{rk}} \end{bmatrix} = \begin{bmatrix} \theta c_{1k} & & O \\ & \theta c_{2k} & \\ & & \ddots \\ O & & & \theta c_{rk} \end{bmatrix} \begin{bmatrix} \frac{\partial q_{11k}}{\partial p_{1k}} & \dots & \frac{\partial q_{11k}}{\partial p_{rk}} \\ \vdots & \ddots & \vdots \\ \frac{\partial q_{1rk}}{\partial p_{1k}} & \dots & \frac{\partial q_{1rk}}{\partial p_{rk}} \end{bmatrix}$$

D'après les propriétés du calcul du rang de matrice, on a

$$\text{rang}(B^{(k)}) = \text{rang} \begin{bmatrix} \frac{\partial q_{11k}}{\partial p_{1k}} & \dots & \frac{\partial q_{11k}}{\partial p_{rk}} \\ \vdots & \ddots & \vdots \\ \frac{\partial q_{1rk}}{\partial p_{1k}} & \dots & \frac{\partial q_{1rk}}{\partial p_{rk}} \end{bmatrix}$$

Un calcul explicite de dérivation montre que pour $j = 1, 2, \dots, r$; $m = 1, 2, \dots, r$

$$\frac{\partial q_{1jk}}{\partial p_{mk}} = \begin{cases} \frac{(1+\theta c_k) - \theta c_{mk} p_{mk}}{(1+\theta c_k)^2} & \text{si } m = j \\ -\frac{\theta c_{mk} p_{jk}}{(1+\theta c_k)^2} & \text{si } m \neq j \end{cases}$$

Cette dernière relation permet d'écrire :

$$\text{rang}(B^{(k)}) = \text{rang} \begin{bmatrix} \Delta_k^{(1)} & -\theta c_{2k} p_{1k} & -\theta c_{3k} p_{1k} & \cdots & -\theta c_{rk} p_{1k} \\ -\theta c_{1k} p_{2k} & \Delta_k^{(2)} & -\theta c_{3k} p_{2k} & \cdots & -\theta c_{rk} p_{2k} \\ -\theta c_{1k} p_{3k} & -\theta c_{2k} p_{3k} & \Delta_k^{(3)} & \cdots & -\theta c_{rk} p_{3k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\theta c_{1k} p_{rk} & -\theta c_{2k} p_{rk} & \cdots & -\theta c_{3k} p_{rk} & \Delta_k^{(r)} \end{bmatrix}$$

où $\Delta_k^{(m)} = (1 + \theta c_k) - \theta c_{mk} p_{mk}$; $m = 1, 2, \dots, r$. Le calcul du rang de la matrice de droite se fait de façon itérative. On calcule à chaque itération les coefficients d_{ij} (lemme 3.4.1., paragraphe 3.4) associés à la matrice de l'étape suivante. On remarque que sur chaque ligne le paramètre p_{jk} ($j = 1, 2, \dots, r$) reste invariant alors que sur chaque colonne le coefficient c_{jk} reste invariant.

A l'itération 1, on a les valeurs suivantes pour d_{ij} :

$$\begin{aligned} d_{22} &= \begin{vmatrix} \Delta_k^{(1)} & -\theta c_{2k} p_{1k} \\ -\theta c_{2k} p_{1k} & \Delta_k^{(2)} \end{vmatrix} = [1 + \theta c_k - \theta c_{1k} p_{1k}] \\ &\quad [1 + \theta c_k - \theta c_{2k} p_{2k}] - \theta^2 c_{1k} p_{2k} c_{2k} p_{1k} \\ d_{22} &= (1 + \theta c_k)(\Delta_k^{(1)} - \theta c_{2k} p_{2k}) \\ d_{2r} &= \begin{vmatrix} \Delta_k^{(1)} & -\theta c_{rk} p_{1k} \\ -\theta c_{1k} p_{2k} & -\theta c_{rk} p_{2k} \end{vmatrix} = -\theta c_{rk} p_{2k} [1 + \theta c_k - \theta c_{1k} p_{1k}] - \theta^2 c_{1k} c_{rk} p_{1k} p_{2k} \\ d_{2r} &= -(1 + \theta c_k) \theta c_{rk} p_{2k} \\ d_{r2} &= \begin{vmatrix} \Delta_k^{(1)} & -\theta c_{2k} p_{1k} \\ -\theta c_{1k} p_{rk} & -\theta c_{2k} p_{rk} \end{vmatrix} = -(1 + \theta c_k) \theta c_{2k} p_{rk} \\ d_{rr} &= \begin{vmatrix} \Delta_k^{(1)} & -\theta c_{rk} p_{1k} \\ -\theta c_{1k} p_{rk} & -\Delta_k^{(r)} \end{vmatrix} = (1 + \theta c_k)(\Delta_k^{(1)} - \theta c_{rk} p_{rk}) \end{aligned}$$

Puisque $\Delta_k^{(1)} \neq 0$, on a (par application du lemme 3.4.1.) à l'itération 1

$$\begin{aligned} \text{rang}(B^k) &= 1 + \text{rang} \begin{pmatrix} (1 + \theta c_k)(\Delta_k^{(1)} - \theta c_{2k} p_{2k}) & \cdots & -(1 + \theta c_k) \theta c_{rk} p_{rk} \\ \vdots & \ddots & \vdots \\ -(1 + \theta c_k) \theta c_{2k} p_{rk} & \cdots & (1 + \theta c_k)(\Delta_k^{(1)} - \theta c_{rk} p_{rk}) \end{pmatrix} \\ &= 1 + \text{rang} \begin{pmatrix} \Delta_k^{(1)} - \theta c_{2k} p_{2k} & -\theta c_{3k} p_{2k} & \cdots & -\theta c_{rk} p_{2k} \\ -\theta c_{2k} p_{3k} & \Delta_k^{(1)} - \theta c_{3k} p_{3k} & \cdots & -\theta c_{rk} p_{3k} \\ \vdots & \vdots & \ddots & \vdots \\ -\theta c_{2k} p_{rk} & -\theta c_{3k} p_{rk} & \cdots & \Delta_k^{(1)} - \theta c_{rk} p_{rk} \end{pmatrix} \end{aligned}$$

La matrice de départ (avant l'itération 1) est de taille r . Celle résultant de la première itération est de taille $r - 1$. Cette dernière est obtenue en supprimant la première ligne et la première colonne de la matrice de départ et remplaçant la diagonale principale de la matrice de départ par $\Delta_k^{(1)} - \theta c_{mk} p_{mk}$, $m = 2, 3, \dots, r$. La quantité $\Delta_k^{(1)} - \theta c_{2k} p_{2k}$ joue le même rôle que $\Delta_k^{(1)}$. On pose alors $\Delta_{2k}^{(1)} = \Delta_k^{(1)} - \theta c_{2k} p_{2k}$.

Itération 2 Puisque $\Delta_{2k}^{(1)} \neq 0$, on a (on calcule comme précédemment les coefficients d_{ij} issus de la dernière matrice).

$$\text{rang}(B^{(k)}) = 2$$

$$+\text{rang} \begin{pmatrix} \Delta_{2k}^{(1)} - \theta c_{3k} p_{3k} & -\theta c_{4k} p_{3k} & \cdots & -\theta c_{rk} p_{3k} \\ -\theta c_{3k} p_{4k} & \Delta_{2k}^{(1)} - \theta c_{4k} p_{4k} & \cdots & -\theta c_{rk} p_{4k} \\ \vdots & \vdots & \ddots & \vdots \\ -\theta c_{3k} p_{rk} & -\theta c_{4k} p_{rk} & \cdots & \Delta_{2k}^{(1)} - \theta c_{rk} p_{rk} \end{pmatrix}$$

Itération m ($m < r$) On montre (puisque $\Delta_{mk}^{(1)} = \Delta_k^{(1)} - \theta \sum_{j=2}^m c_{jk} p_{jk} \neq 0$),

$$\text{rang}(B^{(k)}) = m$$

$$+\text{rang} \begin{pmatrix} \Delta_{mk}^{(1)} - \theta c_{m+1k} p_{m+1k} & -\theta c_{m+2k} p_{m+1k} & \cdots & -\theta c_{rk} p_{m+1k} \\ -\theta c_{m+1k} p_{m+2k} & \Delta_{mk}^{(1)} - \theta c_{m+2k} p_{m+2k} & & \\ \vdots & \vdots & \ddots & \vdots \\ -\theta c_{m+1k} p_{rk} & -\theta c_{m+2k} p_{rk} & \cdots & \Delta_{mk}^{(1)} - \theta c_{rk} p_{rk} \end{pmatrix}$$

Ainsi à l'itération $r - 2 > 2$ ($\Delta_k^{(1)} - \theta \sum_{j=2}^{r-2} c_{jk} p_{jk} = 1 + \theta c_{r-1k} p_{r-1k} + \theta c_{rk} p_{rk} \neq 0$) on obtient :

$$\text{rang}(B^{(k)}) = r - 2$$

$$+\text{rang} \begin{pmatrix} \Delta_k^{(1)} - \theta \sum_{j=2}^{r-1} c_{jk} p_{jk} & -\theta c_{rk} p_{r-1k} \\ -\theta c_{rk} p_{r-1k} & \Delta_k^{(1)} - \theta \sum_{j=2}^{r-2} c_{jk} p_{jk} - \theta c_{rk} p_{rk} \end{pmatrix}$$

$$\text{rang}(B^{(k)}) = r - 2 + 1 + \text{rang}[\Delta_k^{(1)} - \theta \sum_{j=2}^{r-2} c_{jk} p_{jk}] = r - 1 + 1 = r$$

MODÈLE 2

Calcul de $B^{(k)}$ et $U^{(k)}$. On montre par dérivation que :

$$B^{(k)} = \frac{1}{(1 + \theta_{c_k})^2} \begin{bmatrix} \Delta_k^{(1)} & -\theta_{c_{2k}} p_{1k} & \cdots & -\theta_{c_{rk}} p_{1k} \\ -\theta_{c_{1k}} p_{2k} & \Delta_k^{(2)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & -\theta_{c_{rk}} p_{r-1k} \\ -\theta_{c_{1k}} p_{rk} & \cdots & -\theta_{c_{r-1k}} p_{rk} & \Delta_k^{(r)} \\ \rho_k^{(1)} & \theta_{c_{2k}} p_{1k} & \cdots & \theta_{c_{rk}} p_{1k} \\ -\theta_{c_{1k}} p_{2k} & \rho_k^{(2)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \theta_{c_{rk}} p_{r-1k} \\ \theta_{c_{1k}} p_{rk} & \cdots & \theta_{c_{r-1k}} p_{rk} & \rho_k^{(r)} \end{bmatrix}$$

avec $\rho_k^{(m)} = \theta_{c_k}(1 + \theta_{c_k}) + \theta_{c_{mk}} p_{mk}$; $m = 1, 2, \dots, r$ et

$$U^{(k)} = \frac{c_k}{(1 + \theta_{c_k})^2} (-p_{1k}, -p_{2k}, \dots, -p_{rk}, p_{1k}, p_{2k}, \dots, p_{rk})^T$$

A3.1.b - Rang de B_0

Posons pour $k = 1, 2, \dots, s$

$$\begin{bmatrix} \sqrt{q_{11k}} & & & & \\ & \ddots & & & \\ & & \sqrt{q_{1rk}} & & \bigcirc \\ & & & \sqrt{q_{21k}} & \\ & & & & \ddots \\ \bigcirc & & & & & \sqrt{q_{2rk}} \end{bmatrix} = E_k$$

Par définition de B_0 , on a

$$B_0 = \begin{bmatrix} E_1 & \bigcirc & \cdots & \bigcirc \\ \bigcirc & E_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \bigcirc \\ \bigcirc & \cdots & \bigcirc & E_s \end{bmatrix} \begin{bmatrix} U^{(1)} & B^{(1)} & \bigcirc & \cdots & \bigcirc \\ U^{(2)} & \bigcirc & B^{(2)} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \bigcirc \\ U^{(s)} & \bigcirc & \cdots & \bigcirc & B^{(s)} \end{bmatrix}$$

Par suite

$$\text{rang}(B_0) = \text{rang} \begin{bmatrix} U^{(1)} & B^{(1)} & \bigcirc & \cdots & \bigcirc \\ U^{(2)} & \bigcirc & B^{(2)} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \bigcirc \\ U^{(s)} & \bigcirc & \cdots & \bigcirc & B^{(s)} \end{bmatrix}$$

MODÈLE 1

On obtient par dérivation

$$U^{(k)} = \frac{1}{(1 + \theta c_k)^2} [-c_k p_{1k}, \dots, -c_k p_{rk}, c_{1k} p_{1k}, \dots, c_{rk} p_{rk}]^T \in \mathbf{R}^{2r}.$$

Pour $k = 1, 2, \dots, s$ la matrice $((U^{(k)} B^{(k)}))$ de taille $(2r, 1 + r)$ est donnée par

$$((U^{(k)} B^{(k)})) = \frac{1}{(1 + \theta c_k)^2} \begin{bmatrix} -c_k p_{1k} & \Delta_k^{(1)} & -\theta c_{2k} p_{1k} & \cdots & -\theta c_{rk} p_{1k} \\ -c_k p_{2k} & -\theta c_{1k} p_{2k} & \Delta_k^{(2)} & \cdots & -\theta c_{rk} p_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -c_k p_{rk} & -\theta c_{1k} p_{rk} & -\theta c_{2k} p_{rk} & \cdots & \Delta_k^{(r)} \\ c_{1k} p_{1k} & \theta c_{1k} \Delta_k^{(1)} & -\theta^2 c_{1k} c_{2k} p_{1k} & \cdots & -\theta^2 c_{1k} c_{rk} p_{1k} \\ c_{2k} p_{2k} & -\theta^2 c_{2k} c_{1k} p_{2k} & \theta c_{2k} \Delta_k^{(2)} & \cdots & -\theta^2 c_{2k} c_{rk} p_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{rk} p_{rk} & -\theta^2 c_{rk} c_{1k} p_{rk} & -\theta^2 c_{rk} c_{2k} p_{rk} & \cdots & \theta c_{rk} \Delta_k^{(r)} \end{bmatrix}$$

En appliquant comme précédemment le lemme 3.2.2. et en remarquant pour tout $k = 1, 2, \dots, s$ que $\Delta_k^{(1)} - \theta \sum_{j=2}^r c_{jk} p_{jk} = 1$, on montre après r itérations que

$$\begin{aligned} \text{rang}(B_0) &= 1 + r + \text{rang} \begin{bmatrix} B^{(2)} & \bigcirc & \cdots & \bigcirc \\ \bigcirc & B^{(3)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \bigcirc \\ \bigcirc & \cdots & \bigcirc & B^{(s)} \end{bmatrix} \\ &= 1 + r + \sum_{k=2}^s \text{rang}(B^{(k)}) \\ \text{rang}(B_0) &= 1 + r + (s - 1)r = 1 + sr \end{aligned}$$

Exemple : On suppose $s = 2$ et $r = 2$. On note k et m le numéro de ces deux sites. B_0 est une matrice de taille $(8, 5)$

$$\text{rang}(B_0) = \text{rang} \begin{bmatrix} U^{(k)} & B^{(k)} & \bigcirc \\ U^{(m)} & \bigcirc & B^{(m)} \end{bmatrix}$$

Après dérivation et simplifications, on montre que

$$\text{rang}(B_0) =$$

$$\text{rang} \begin{bmatrix} -c_k p_{1k} & \Delta_k^{(1)} & -\theta c_{2k} p_{1k} & 0 & 0 \\ -c_k p_{2k} & -\theta c_{1k} p_{2k} & \Delta_k^{(2)} & 0 & 0 \\ p_{1k} & \theta \Delta_k^{(1)} & -\theta^2 c_{2k} p_{1k} & 0 & 0 \\ p_{2k} & -\theta^2 c_{1k} p_{2k} & \theta \Delta_k^{(2)} & 0 & 0 \\ -c_m p_{1m} & 0 & 0 & \Delta_m^{(1)} & -\theta c_{2m} p_{1m} \\ -c_m p_{2m} & 0 & 0 & -\theta c_{1m} p_{2m} & \Delta_m^{(2)} \\ p_{1m} & 0 & 0 & +\theta \Delta_m^{(1)} & -\theta^2 c_{2m} p_{1m} \\ p_{2m} & 0 & 0 & -\theta^2 c_{1m} p_{2m} & \theta \Delta_m^{(2)} \end{bmatrix}$$

Itération 1 ($c_k p_{1k} \neq 0$) : on calcule les coefficients d_{ij} associés et on montre après simplifications que

$$\text{rang}(B_0) = 1$$

$$+\text{rang} \begin{bmatrix} p_{2k} & -1 & 0 & 0 \\ -\Delta_k^{(1)} & \theta c_{2k} & 0 & 0 \\ p_{2k}(\theta c_{1k} p_{1k} - 1) & \theta[c_{2k} p_{2k} - c_k] & 0 & 0 \\ c_m p_{1m} \Delta_k^{(1)} & -c_m p_{1m} \theta c_{2k} & -\Delta_m^{(1)} & \theta c_{2m} p_{1m} \\ c_m p_{2m} \Delta_k^{(1)} & -c_m p_{2m} \theta c_{2k} & \theta c_{1m} p_{2m} & \Delta_m^{(2)} \\ -p_{1m} \Delta_k^{(1)} & p_{1m} \theta c_{2k} & -\theta \Delta_m^{(1)} & \theta^2 c_{2m} p_{1m} \\ -p_{2m} \Delta_k^{(1)} & p_{2m} \theta c_{2k} & \theta^2 c_{1m} p_{2m} & -\theta \Delta_m^{(2)} \end{bmatrix}$$

Après cette itération, on remarque que la matrice résultante est de taille (7,4). Le bloc correspondant au site n° k s'est transformé en une matrice de taille (3,4). Alors que le bloc correspondant au site n° m est de taille (4,4).

Itération 2 ($p_{2k} \neq 0$). On applique l'algorithme à la matrice de droite. On obtient

$$\text{rang}(B_0) = 1 + 1 + \text{rang} \begin{bmatrix} -1 & 0 & 0 \\ -p_{2k} & 0 & 0 \\ c_m p_{1m} & -\Delta_m^{(1)} p_{2k} & \theta c_{2m} p_{1m} p_{2k} \\ c_m p_{2m} & \theta c_{1m} p_{2m} p_{2k} & \Delta_m^{(2)} p_{2k} \\ -p_{1m} & -\theta \Delta_m^{(1)} p_{2k} & \theta^2 c_{2m} p_{1m} p_{2k} \\ -p_{2m} & \theta^2 c_{1m} p_{2m} p_{2k} & -\theta \Delta_m^{(2)} p_{2k} \end{bmatrix}$$

En simplifiant, on obtient finalement

$$\text{rang}(B_0) = 2 + \text{rang} \begin{bmatrix} -1 & 0 & 0 \\ -c_m p_{1m} & \Delta_m^{(1)} & -\theta c_{2m} \\ -c_m p_{2m} & -\theta c_{1m} p_{2m} & \Delta_m^{(2)} \\ p_{1m} & \theta \Delta_m^{(1)} & -\theta^2 c_{2m} p_{1m} \\ p_{2m} & -\theta^2 c_{1m} p_{2m} & \theta \Delta_m^{(2)} \end{bmatrix}$$

à $1 + s(r - 1)$ équations et $1 + s(r - 1)$ inconnus. Les $s(r - 1)$ dernières équations permettent d'écrire, pour $k = 1, 2, \dots, s$

$$(\tilde{\beta}_k - \beta_k^o) = (W_k^T W_k)^{-1} W_k^T Y_k - (W_k^T W_k)^{-1} W_k^T V_k (\tilde{\theta} - \theta^o) + o_p(1).$$

Par substitution dans la première équation, on a

$$(\tilde{\theta} - \theta^o) = \frac{1}{\alpha} \sum_{m=1}^s V_m^T D_m Y_m + o_p(1).$$

En remplaçant $(\tilde{\theta} - \theta^o)$ par cette dernière expression, on obtient finalement

$$\begin{aligned} (\tilde{\beta}_k - \beta_k^o) &= (W_k^T W_k)^{-1} W_k^T [I_{2r} - \frac{1}{\alpha} V_k V_k^T D_k] Y_k \\ &\quad - \frac{1}{\alpha} \sum_{\substack{m=1 \\ m \neq k}}^s (W_k^T W_k)^{-1} W_k^T V_k V_m^T D_m Y_m + o_p(1) \end{aligned}$$

($k = 1, 2, \dots, s$). Ces deux dernières égalités permettent de mettre l'égalité

$$\tilde{\Phi} - \Phi^o = (D^T D)^{-1} D^T Y + o_p(1)$$

sous la forme :

$$\begin{bmatrix} \tilde{\theta} - \theta^o \\ \tilde{\beta}_1 - \beta_1^o \\ \tilde{\beta}_2 - \beta_2^o \\ \vdots \\ \tilde{\beta}_s - \beta_s^o \end{bmatrix} = \begin{bmatrix} \frac{1}{\alpha} V_1^T D_1 & \frac{1}{\alpha} V_2^T D_2 & \dots & \frac{1}{\alpha} V_s^T D_s \\ \tilde{B}_1 [I_{2r} - \frac{1}{\alpha} V_1 V_1^T D_1] & -\tilde{B}_1 \frac{V_1 V_2^T}{\alpha} D_2 & \dots & -\tilde{B}_1 \frac{V_1 V_s^T}{\alpha} D_s \\ -\tilde{B}_2 [\frac{1}{\alpha} V_2 V_1^T D_1] & \tilde{B}_2 [I_{2r} - \frac{1}{\alpha} V_2 V_2^T D_2] & \ddots & \vdots \\ \vdots & \ddots & \ddots & -\tilde{B}_{s-1} \frac{V_{s-1} V_s^T}{\alpha} D_s \\ -\tilde{B}_s \frac{1}{\alpha} V_s V_1^T D_1 \dots & -\tilde{B}_s \frac{V_s V_{s-1}^T}{\alpha} D_{s-1} & \dots & \tilde{B}_s [I_{2r} - \frac{V_s V_s^T}{\alpha} D_s] \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_s \end{bmatrix} + o_p(1)$$

où pour $k = 1, 2, \dots, s$ les matrices $\tilde{B}_k$ de taille $(r - 1, 2r)$ sont définies par $\tilde{B}_k = (W_k^T W_k)^{-1} W_k^T$. En combinant cette dernière égalité matricielle et la relation $\tilde{A} = Y - D(\tilde{\Phi} - \Phi^o) + o_p(1)$, on obtient par identification :

$$D(D^T D)^{-1} D^T =$$

$$\begin{bmatrix} \frac{1}{\alpha} V_1 V_1^T D_1 + W_1 \tilde{B}_1 [I_{2r} - \frac{V_1 V_1^T}{\alpha} D_1] & \cdots & \frac{1}{\alpha} V_1 V_s^T D_s - W_1 \tilde{B}_1 \frac{V_1 V_s^T}{\alpha} D_s \\ \frac{1}{\alpha} V_2 V_1^T D_1 - W_2 \tilde{B}_2 \frac{V_2 V_1^T}{\alpha} D_1 & \cdots & \frac{1}{\alpha} V_2 V_s^T D_s - W_2 \tilde{B}_2 \frac{V_2 V_s^T}{\alpha} D_s \\ \vdots & \ddots & \vdots \\ \frac{1}{\alpha} V_s V_1^T D_1 - W_s \tilde{B}_s \frac{V_s V_1^T}{\alpha} D_1 & \cdots & \frac{1}{\alpha} V_s V_s^T D_s + W_s \tilde{B}_s [I_{2r} - \frac{V_s V_s^T}{\alpha} D_s] \end{bmatrix}$$

On note que $W_k \tilde{B}_k = I_{2r} - D_k$, par suite on a :

$$\begin{cases} \frac{1}{\alpha} V_k V_k^T D_k + W_k \tilde{B}_k [I_{2r} - \frac{V_k V_k^T}{\alpha} D_k] & = I_{2r} - D_k + \frac{1}{\alpha} D_k V_k V_k^T D_k \\ \frac{1}{\alpha} V_m V_k^T D_k - W_m \tilde{B}_m \frac{1}{\alpha} V_m V_k^T D_k & = \frac{1}{\alpha} D_m V_m (D_k V_k)^T, m \neq k \end{cases}.$$

On peut donc mettre la matrice $D(D^T D)^{-1} D^T$ sous la forme

$$D(D^T D)^{-1} D^T = \begin{bmatrix} I_{2r} - D_1 & \bigcirc & \cdots & \bigcirc \\ \bigcirc & I_{2r} - D_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \bigcirc \\ \bigcirc & \cdots & \bigcirc & I_{2r} - D_s \end{bmatrix} + \frac{1}{\alpha} \begin{bmatrix} D_1 V_1 V_1^T D_1 & D_1 V_1 V_2^T D_2 & \cdots & D_1 V_1 V_s^T D_s \\ D_2 V_2 V_1^T D_1 & D_2 V_2 V_2^T D_2 & \cdots & D_2 V_2 V_s^T D_s \\ \vdots & \vdots & \ddots & \vdots \\ D_s V_s V_1^T D_1 & D_s V_s V_2^T D_2 & \cdots & D_s V_s V_s^T D_s \end{bmatrix}.$$

D'après le lemme 3.2.1., on a :

$$I_{2sr} - \Pi^{1/2} \Pi^{1/2T} = \begin{bmatrix} I_{2sr} - \Pi_1^{1/2} \Pi_1^{1/2T} & \bigcirc & \cdots & \bigcirc \\ \bigcirc & I_{2sr} - \Pi_s^{1/2} \Pi_s^{1/2T} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \bigcirc \\ \bigcirc & \cdots & \bigcirc & I_{2sr} - \Pi_s^{1/2} \Pi_s^{1/2T} \end{bmatrix}$$

Par conséquent on obtient, sous forme matricielle,

$$\tilde{\Omega} = \begin{bmatrix} D_{\Pi_1^{1/2}} & \bigcirc & \cdots & \bigcirc \\ \bigcirc & D_{\Pi_2^{1/2}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \bigcirc \\ \bigcirc & \cdots & \bigcirc & D_{\Pi_s^{1/2}} \end{bmatrix} - \frac{1}{\alpha} \begin{bmatrix} D_1 V_1 V_1^T D_1 & \cdots & D_1 V_1 V_s^T D_s \\ \vdots & \ddots & \vdots \\ D_s V_s V_1^T D_1 & \cdots & D_s V_s V_s^T D_s \end{bmatrix}$$

où $D_{\Pi_k^{1/2}} = D_k - \Pi_k^{1/2} \Pi_k^{1/2T}$; $k = 1, 2, \dots, s$. On montre ainsi que

$$\tilde{\Omega} = D_{\Pi^{1/2}} - \frac{1}{\alpha} V V^T .$$

En utilisant le lemme 3.2.3., on montre que la matrice $D_{\Pi^{1/2}}$ est symétrique et idempotente c'est-à-dire $D_{\Pi^{1/2}}^T = D_{\Pi^{1/2}}$ et $D_{\Pi^{1/2}}^2 = D_{\Pi^{1/2}}$. Par conséquent, $\text{tr}(D_{\Pi^{1/2}}) = \sum_{k=1}^s \text{tr}(D_k - \Pi_k^{1/2} \Pi_k^{1/2T}) = sr$. On note aussi que

$$D_{\Pi^{1/2}} V V^T = V V^T$$

(on utilise les relations $W_k^T \Pi_k^{1/2} = 0$ et $V_k^T \Pi_k^{1/2} = 0$). On en déduit alors que $\tilde{\Omega}^T = \tilde{\Omega}$ et $\tilde{\Omega}^2 = \tilde{\Omega}$. $\tilde{\Omega}$ est donc une matrice symétrique et idempotente, par suite $\text{tr}(\tilde{\Omega}) = \text{tr}(D_{\Pi^{1/2}} - \text{tr}(\frac{1}{\alpha} V V^T)) = sr - 1$.

ANNEXE 4 - Programmes informatiques

Modèle 1

```

C .... Recherche de solutions aux equations de vraisemblance
C par utilisation de la Subroutine E04VCF de la Bibliotheque NAG
C ( C.I.T.I Lille). Ce programme a ete teste sur les donnees
C de Tanner avant d'etre applique a des tableaux 2xSxR....
C .... MODELE 1.
C
C ..... Parametres .....
C Integer IS, IR
C parameter (IS=2,IR=2)
C
C Integer n, nclin, ncnln, nctotl, nrowa, nrowj, liwork,
C * lwork, nrowr
C parameter (n=1+IS*IR,nclin=IS,ncnln=0,nctotl=n+nclin+ncnln,
C * nrowa=IS,nrowj=1,nrowr=n,liwork=3*n+nclin+ncnln,
C * lwork=3*n*(n+2)*(nclin+ncnln)+n*nrowj
C * +12*n+5*(nclin+ncnln)+nrowa+5*n+nrowj+nclin)
C
C integer nin, nout
C parameter (nin=5,nout=6)
C real zero, one, two, bigbnd
C parameter (zero=0.0e0,one=1.0e0 ,two=2.0e0)
C
C ..... Local Scalars .....
C
C real ftol, OBJF,EPSTAF, EPSMCH,ETA,RTEPS
C INTEGER i,ifail,itmax,j,msglvl,iter,mode,nstate
C LOGICAL COLD, FEALIN, ORTHOG
C
C ..... Locals Arrays
C real A(nrowa,n),bl(nctotl), bu(nctotl),c(nrowj),
C * cjac(nrowj,n),clambda(nctotl), OBJGRD(n),
C * work(lwork), x(n), featol(nctotl), r(nrowr,n),
C * HESS(n+is,n+is),VARCO(n+is,n+is)
C integer istate(nctotl), iwork(liwork)
C
C real X02AAF
C EXTERNAL X02AAF
C ..... Arrays in common
C
C Integer IXE1(IS,IR),IXE2(IS,IR),IXE(IS,2*IR),NE(IS),NE1(IS)
C Integer IXE3(IS,IR),NE2(IS),k
C Integer G05ECF,G05EYF
C Real Pjk(is,ir),se(is),cocm(is),teta,tetak(is)
C Real Qtjk1(is,2*ir),sigma,sigma2(3),TCAR,
C * DQtjk1(2*ir,ir+1),SQtjk1(1+ir,1+ir),RJQtjk1(2*ir),
C * VJQtjk1(2*ir,2*ir),EJQtjk1(2*ir),QQtjk1(2*ir,2*ir),
C * FQtjk1(1+ir,1+ir),FQINV(1+ir,1+ir),
C * GSIG(2*ir,1+ir),FSIG(2*ir,2*ir),SQINV(1+ir,1+ir),
C * RQQtjk1(2*ir,2*ir),GRAO(2*ir,1+ir),FRAO(2*ir,2*ir),
C * RAO(2*ir,2*ir),FICHR(2*ir,2*ir),RRAO(2*ir),ERAO(2*ir),
C * VRAO(2*ir,2*ir),
C * sigmak1(2*is),sigmak2(3),tetabar
C Real XT0,XT1,XT2,XT3,XC,G05CBF,CMC(IS),class(is,2*ir)
C Real COEC(IS,IR),PROB(IS,IR),EFFAP(IS),G05DAF
C
C COMMON/ACC1/IXE1, IXE2,IXE3,IXE, NE, NE1,NE2
C COMMON/FACC1/ COEC
C EXTERNAL CONFUN,E04VCF,E04ZCF,OBJFUN,X04ABF,CMOY,ESTIMA,F02ABF
C EXTERNAL G05DAF,G05CBF,G05ECF,G05EYF,G05CAF,GMULT,GENENOR,F01AAF

```

```

INTRINSIC ABS,SQRT
c ..... Executable Statements .....
c
c .... Choix de la vraie valeur de teta: effet moyen ....
teta = .50
c
c   write(nout,fmt=88996) teta
c .... Choix de la variance sigma-carre ....
sigma = 10e-6
sigma = sqrt(sigma)
c
c .... Generation de l'effet de la mesure de chaque site ....
CALL GENENOR(is,ir,teta,sigma,tetak)
write(nout,*) (tetak(i),i=1,is)
c
c   tetabar = 0.
c   do 41 k=1,is
c     tetabar = tetabar+tetak(k)
41  continue
c   tetabar = tetabar/is
c   write(nout,*) tetabar
c
c .... Choix des vraies probabilites Pjk avant la mesure ....
do 15 k=1,is
se(k) = 0.
do 13 j=1,ir
Pjk(k,j) = G05CAF(X)
se(k) = se(k)+Pjk(k,j)
13  continue
do 14 j=1,ir
14  Pjk(k,j) = Pjk(k,j)/se(k)
15  continue
c
c .... Choix des coefficients de controle COEC(k,j) ....
do 32 k=1,is
do 30 j=1,ir
CALL G05CBF(k*j)
COEC(k,j) = G05DAF(.5,2.5)
30  continue
32  continue
c .... Calcul des vrais coefficients moyens de controle "c(k)" ....
do 42 k=1,is
cocm(k) = 0.
do 40 j=1,ir
cocm(k) = cocm(k)+COEC(k,j)*Pjk(k,j)
40  continue
42  continue
c
c ..... Simulation de tableaux Avant-After 2xSxR .....
c
CALL GMULT(IS,IR,tetak,Pjk,COEC,cocm,Qtjk1,NE,IXE)
c
c   do 7 k=1,IS
c     do 5 j=1,IR
c       IXE1(k,j) = IXE(k,j)
c       IXE2(k,j) = IXE(k,j)+IR)
5    continue
7    continue
c   DO 24 k=1,IS
c     NE1(k) = 0
c     NE2(k) = 0
c     do 22 m=1,IR
c       NE1(k) = NE1(k)+ixe1(k,m)
c       NE2(k) = NE2(k)+ixe2(k,m)
22  continue

```

```

      do 23 j=1,IR
c      PROB(k,j) = float(ixe1(k,j))/NE1(k)
      PROB(k,j) = 1./IR
23      continue
24 CONTINUE
c
      do 8 k=1,is
        write(6,88997) (Pjk(k,m),m=1,ir),(COEC(k,l),l=1,ir),cocm(k)
8      continue
        write(6,88998)
        do 9 k=1,is
          write(6,89002) (Qtjk1(k,l),l=1,2*ir)
9      continue
        write(6,88998)
        write(6,88999)
        write(6,89000) is,ir
        do 12 k=1,is
          write(6,89001) (ixe(k,m),m=1,2*ir),ne(k)
12     continue
c
      DO 80 k=1,IS
        do 78 j=1,IR
          ix3(k,j) = ix1(k,j)+ix2(k,j)
78      continue
80 CONTINUE
c
c      .... Appel de la subroutine E04VCF pour la recherche des EMVC ....
      write(nout,fmt=99999)
c
      call X04ABF(1,nout)
      itmax=100
      msglvl=5
      eta = 0.2e0
      epsmch = X02AAF(zero)
c      ftol = 10.0*epsmch
      ftol = 1.0e-5
      rsteps = sqrt(epsmch)
      do 2 j=1,nctotl
        featol(j) = rsteps
2      continue
      bigbnd=1.0e10
c
c      .... A = Matrice des contraintes lineaires .....
      do 11 j=1,n
        do 10 i=1,nclin
          A(i,j)= 0.
10      continue
11      continue
        do 17 i=1,nclin
          do 16 j=(i-1)*IR+2,i*IR+1
            A(i,j) = 1.
16      continue
17      continue
c      .... Ensemble de toutes les bornes (inf et sup) du probleme ....
c      On commence par le vecteur BL(j) de toutes les bornes inf.
c      Ensuite, on remplit le vecteur BU(j) de toutes les bornes sup.
c      Il faut entrer d'abord les bornes sur les variables du probleme,
c      ensuite les bornes sur les contraintes lineaires generales, et
c      .... enfin les bornes sur les contraintes non lineaires generales....
      BL(1) = 0.
      do 18 j=2,n
        BL(j) = .05
18      continue
      do 19 j=n+1,nctotl
        BL(j) = 1.

```

```

19 continue
   BU(1) = bigbnd
   do 20 j=2,n
     BU(j) = .998
20 continue
   do 90 j=n+1,nctotl
     BU(j) = 1.
90 continue
c
c   .... X = point de Depart de l'optimisation ....
c   do 25 k=1,IS
     i=0
     do 21 j=(k-1)*IR+2,k*IR+1
       i=i+1
       X(j)=PROB(k,i)
21 continue
25 continue
   X(1) = .0385
c
   do 26 k=1,IS
     CMC(k) = CMOY(k,IS,IR,n,X,COEC)
26 continue
   write(nout,*) (CMC(k),k=1,IS)
   write(nout,fmt=99996) (X(j),j=1,n)
c
nstate = 1
mode = 1
c
c
c   .... Definition de la precision absolue sur la OBJF ....
c   CALL OBJFUN(mode,n,x,OBJF,OBJGRD,nstate)
c   epsaf = epsmch*ABS(OBJF)
c   COLD = .TRUE.
c   FEALIN = .TRUE.
c   ORTHOG = .TRUE.
c
c   Check the gradients using routine E04ZCF
c
c   ifail=-1
c   call E04ZCF(n,ncnln,nrowj,CONFUN,OBJFUN,c,cjac,objf,objgrd,
1     x,work,lwork,ifail)
c   if(ifail.eq.0) then
c     write(nout,fmt=99998)
c
c   Solve the problem from a cold start.
c
c   ifail=-1
c   call E04VCF(itmax,msglvl,n,ncnln,ncnln,nctotl,nrowa,nrowj,nrowr,
1     bigbnd,epsaf,eta,ftol,A,BL,BU,featol,CONFUN,
2     OBJFUN,cold,fealin,orthog,X,istate,R,iter,c,cjac,
3     OBJF,OBJGRD,clamda,iwork,liwork,work,lwork,ifail)
c   if(ifail.eq.0) then
c
c     do 140 j=1,n
c       X(j)=X(j)
140 continue
c   ... Reset the absolute precision of the objective function ....
c   epsaf = epsmch*ABS(OBJF)
c   COLD = .false.
c   write(nout,fmt=99997)
c   write(nout,fmt=99996) (X(j),j=1,n)
c
c   ifail = -1
c   call E04VCF(itmax,msglvl,n,ncnln,ncnln,nctotl,nrowa,nrowj,nrowr,
1     bigbnd,epsaf,eta,ftol,A,BL,BU,featol,CONFUN,

```

```

2      OBJFUN,cold,fealin,orthog,X,istate,R,iter,c,cjac,
3      OBJF,OBJGRD,clamda,iwork,liwork,work,lwork,ifail)
END IF
  if (ifail.ne.0) write(nout,fmt=99995) ifail
  else
    write(nout,fmt=99994) ifail
  end if
c
  do 28 k=1,IS
    CMC(k) = CMOY(k,IS,IR,n,X,COEC)
28  continue
    write(nout,*) (CMC(k),k=1,IS)
c
c      *****
c
c      ***** Definition et inversion de la matrice de Fisher *****
c      augmentee : HESS
c      *****
c
  call INFO(is,ir,n,X,coec,cmc,PROB,NE,HESS,VARCO)
c
  write(nout,2912) ((HESS(i,j),j=1,n+is),i=1,n+is)
  if (ifail.ne.0) write(nout,2996) ifail
c
  write(nout,2995) ((VARCO(i,j),j=1,n+is),i=1,n+is)
c
c      *****
c      Etude dela distribution asymptotique de la Statistique
c      T-deux
c      *****
c
  CALL ESTIMA(is,ir,n,X,COEC,CMC,PROB,Qtjk1,IXE,NE,class,DQtjk1,
*      SQtjk1,RJQtjk1,VJQtjk1,EJQtjk1,FQtjk1,FQINV,QQtjk1,
*      GSIG,FSIG,RQQtjk1,SQINV,GRAO,FRAO,RAO,FICHR,RAAO,
*      VRAO,ERAO,
*      TCAR,sigma1,sigma2,sigma2)
c
  write (nout,fmt=79995) ifail
  write (nout,fmt=79994) (RJQtjk1(i),i=1,2*ir)
  write (nout,fmt=79998)
  do 128 i=1,2*ir
    write (nout,79997) (VJQtjk1(i,j),j=1,2*ir)
128  continue
  write (nout,fmt=79994) (RAAO(i),i=1,2*ir)
  write (nout,fmt=79998)
  do 130 i=1,2*ir
    write (nout,79997) (VRAO(i,j),j=1,2*ir)
130  continue

  stop
c
99999 format(/ ' E04VCF Example Program Results: Modele1 ',/1x)
99998 format(/ ' Successful exit from E04ZCF.')
99997 format(/ ' A run of the same exemple with a warm start ...')
99996 format(/ ' Initial  X.',/ 5(1x,f10.5))
99995 format(/ ' E04VCF terminated with ifail = ' ,I3)
99994 format(/ ' Incorrect gradients. ifail = ' , I3)
88996 format(/ ' La vraie valeur de teta est =', f6.4,/)
88997 format(4(1x,f10.7),',',4(1x,f10.7),',',f10.7)
88998 format( /)
88999 format(/, ' *****   Tableau Avant-Apres 2xSxR   *****',/)
89000 format(' Nombre de Sites S = ' ,i3,','Nombre de Types R = ' ,i3, /)
89001 format(9(5x,i5))
89002 format(4(1x,f10.7),',',4(1x,f10.7))
79998 format (/ ' vecteurs propres ')
79997 format ( ' ',16f9.4)

```

```

79995 format (/ ' erreur dans F02ABF ifail = ',i2)
79994 format (' valeurs propres',/( ' ',16f9.4))
2912 format(/ ' Info Fisher est: '// (1H , 7f10.4))
2996 format(25Hoerror in F01AAF ifail = ,i2)
2995 format(/ ' l"inverse est '// (1H , 7f10.7))
c
  end

c
  .... Definition du coefficient de controle moyen ....
  FUNCTION CMOY(k,IS,IR,n,X,COEC)
    integer IS,IR,i,j,k,n
    real X(n),COEC(IS,IR)
    real CMOY
    CMOY = 0.
    i=0
    do 90 j=(k-1)*IR+2,k*IR+1
      i=i+1
    CMOY = CMOY+COEC(k,i)*X(j)
90 continue
    return
  END

c
  Subroutine OBJFUN(MODE,N,X,OBJF,OBJGRD,NSTATE)
C
  .... Scalars Arguments ....
  Integer IS, IR
  Parameter(IS=2,IR=2)
  integer ixel(is,ir),ixe2(IS,IR),ne(IS),ne1(IS),ne2(IS)
  integer mode,n,nstate,i,j,k,ixe3(IS,IR),ixe(IS,2*IR)
  real OBJF,COEC(IS,IR),CMC(IS),OBJGRD(n),
  * X(n)
  COMMON/ACC1/IXE1, IXE2,IXE3, IXE, NE,NE1,NE2
  COMMON/FACC1/ COEC
c
  ..... Executable statements .....
  EXTERNAL CMOY
c
  ..... Fonction de Vraisemblance .....
  OBJF = 0.
  DO 95 k=1,IS
    OBJF = OBJF+NE2(k)*LOG(X(1))-
1    NE(k)*LOG(1.+X(1)*CMOY(k,IS,IR,n,X,COEC))
    i=0
    do 93 j=(k-1)*IR+2,k*IR+1
      i=i+1
    OBJF = OBJF+IXE3(k,i)*LOG(X(j))+IXE2(k,i)*LOG(COEC(k,i))
93 continue
95 continue
  OBJF = - OBJF

c
c
  .... Vecteur gradient ....
  OBJGRD(1) = 0.
  do 105 k=1,IS
    OBJGRD(1) = OBJGRD(1)+(NE2(k)/X(1))-
1    (NE(k)*CMOY(k,IS,IR,n,X,COEC))/
2    (1.+X(1)*CMOY(k,IS,IR,n,X,COEC))
105 continue
c
  do 110 k=1,IS
    m=0
    do 108 j=(k-1)*IR+2,k*IR+1
      m=m+1
    OBJGRD(j) = (IXE3(k,m)/X(j))-
1    (X(1)*COEC(k,m)*NE(k))/(1.+X(1)*CMOY(k,IS,IR,n,X,COEC))
108 continue
110 continue
c
  do 112 m = 1,n

```

```

      OBJGRD(m) = - OBJGRD(m)
112 CONTINUE
      return
      end
c
      SUBROUTINE CONFUN(mode,ncnln,n,nrowj,x,c,cjac,nstate)
c   The zero elements of jacobian matrix are set only once. This
c   occurs during the first call to confun (nstate = 1).
c
c   ..... Parameters .....
      Integer IS,IR
      Parameter(IS=1,IR=2)
      real      zero, two, one
      parameter (zero=0.0e0,one=1.0e0,two=2.0e0)
      Integer ixel(IS,IR),ixe2(IS,IR),ne(IS),ixe(IS,2*IR),ne1(IS)
      integer    mode, n, ncnln, nrowj, nstate,ixe3(IS,IR),ne2(IS)
      Real      X(n), c(nrowj),cjac(nrowj,n),coec(IS,IR)
c   ..... Scalars Arguments .....
      n = 1+IS*IR
c   ..... Local Scalars .....
      if (nstate.eq.1) then
      do 116 j=1,n
      do 114 i=1,nrowj
      cjac(i,j) = 0.
114  continue
116  continue
c
      end if
      return
      end
c
      SUBROUTINE GMULT(IS,IR,tetak,Pjk,COEC,cocm,Qtjk1,NE,IXE)
c   .... Simulation d'une loi multinomiale par utilisation de la
c   .... bibliotheque NAG .....
c   On tire de facon aleatoire le nombre total d'accidents de
c   chaque site experimental avec une distribution de Poisson.
c   On distribue de facon multinomiale Avant et Apres ce nombre
c   dans les deux fois ir categories d'accidents.
c
      Integer is,ir
      Integer ne(is),ntot(1000),ifail,ixe(is,2*ir),i,j,k
      Integer G05ECF,G05EYF,NR1
      Real    Pjk(is,ir),som(1000,1000),cocm(is),G05DAF
      Real    Qtjk1(is,2*ir),tetak(is),COEC(is,ir),G05CAF
      Real    XT0,XT1,XT2,XT3,XC,R1(10000)
      External G05ECF,G05EYF,G05CAF,G05DAF
c
c   .... Calcul des Probabilites Qtjk de la loi multinomiale ....
      do 55 k=1,is
      do 51 j=1,ir
      Qtjk1(k,j) = Pjk(k,j)/(1.+tetak(k)*cocm(k))
      Qtjk1(k,j+ir) = tetak(k)*COEC(k,j)*Pjk(k,j)/
      * (1.+tetak(k)*cocm(k))
51  continue
55  continue
c
c   .... Choix aleatoire du nombre ne(k)=x1..k d'accidents du site k
c   G05CBF permet d'initialiser le tirage de ne(k); Choisir
c   le parametre de G05CBF constant si on veut les ne(k) identiques.
      do 2 k=1,is
      CALL G05CBF(is*ir+(k-1))
      ifail = 0
      XC = G05DAF(100.,300.)
      XT0 = sqrt(XC)
      XT1 = 20.+20.*XT0

```

```

      XT2 = XC+7.15*XT0
      XT3 = XC-7.15*XT0
      NR1 = MAX(int(XT1),int(XT2+8.5))-MAX(0,int(XT3))+4)
      CALL G05ECF(XC,R1,NR1,IFAIL)
      if(ifail.ne.0) stop
      ne(k) = G05EYF(R1,NR1)
2. continue
c .... Initialisation de l'effectif ix de chaque classe
do 11 m=1,is
  do 11 j=1,2*ir
11   ix(m,j) = 0
c .... Decoupage de l'intervalle (0,1) en 2*ir classes par
c utilisation des Qtjk1(k,j). Les som(k,j) representent les
c points de decoupage.
do 18 k=1,is
  som(k,1) = 0.
  do 17 m=2,2*ir+1
    som(k,m) = som(k,m-1)+Qtjk1(k,m-1)
17 continue
18 continue
c
c .... Remplissage des elements ix par tirage d'une proba p
c et comparaison de p aux som(k,j) ....
do 21 k=1,is
  ntot(k) = ne(k)
1005  p = G05CAF(X)
  do 20 j=1,2*ir
    if (som(k,j).lt.p.and.p.lt.som(k,j+1)) then
      ix(k,j)=ix(k,j)+1
      ntot(k)=ntot(k)-1
    else
      ix(k,j)=ix(k,j)
      ntot(k) = ntot(k)
    end if
20 continue
  if (ntot(k).gt.0) go to 1005
21 continue
c
  return
end
c
  subroutine GENENOR(is,ir,n,teta,sigma,tetak)
c .... Programme de generation de nombres pseudo-aleatoires
c provenant d'une distribution normale ....
c
  integer is,ir
  real G05DDF,G05CBF,tetak(is),teta,sigma
  external G05DDF,G05CBF
c
  call G05CBF(0)
  do 20 i=1,is
    tetak(i) = G05DDF(teta,sigma)
20 continue
  return
end
c
  SUBROUTINE ESTIMA(is,ir,n,X,COEC,CMC,PROB,Qtjk1,IXE,NE,class,
* DQtjk1,SQtjk1,RJQtjk1,VJQtjk1,EJQtjk1,FQtjk1,
* FQINV,QQtjk1,GSIG,FSIG,RQQtjk1,SQINV,GRAO,FRAO,
* RAO,FICHR,RRAO,VRAO,ERAO,
* TCAR,sigmak1,sigmak2,sigma2)
  Integer is,ir,n,i,j,k,m,jk,ia,IXE(is,2*ir),NE(is)
  Real X(n),CMC(is),Qtjk1(is,2*ir),TCAR,sigma2(3),COEC(is,ir),
* DQtjk1(2*ir,ir+1),SQtjk1(1+ir,1+ir),RJQtjk1(2*ir),
* VJQtjk1(2*ir,2*ir),EJQtjk1(2*ir),QQtjk1(2*ir,2*ir),

```

```

*   FQtjk1(1+ir,1+ir),FQINV(1+ir,1+ir),WKSPACE(10000),
*   GSIG(2*ir,1+ir),FSIG(2*ir,2*ir),SQINV(1+ir,1+ir),
*   RQQtjk1(2*ir,2*ir).GRAO(2*ir,1+ir),FRAO(2*ir,2*ir),
*   RAO(2*ir,2*ir),FICHR(2*ir,2*ir),RRAO(2*ir),ERAO(2*ir),
*   VRAO(2*ir,2*ir),
*   PROB(is,ir),class(is,2*ir),sigmak1(2*is),sigmak2(3)
External CMOY,F02ABF,F01AAF
c
do 25 k=1,is
  i=0
  do 21 j=(k-1)*ir+2,k*ir+1
    i=i+1
    PROB(k,i) = X(j)
21  continue
25  continue
c
do 55 k=1,is
  do 51 j=1,ir
    Qtjk1(k,j) = PROB(k,j)/(1.+X(1)*CMC(k))
    Qtjk1(k,j+ir) = PROB(k,j)*COEC(k,j)*X(1)/(1.+X(1)*CMC(k))
51  continue
  do 53 m=1,ir
    class(k,m) = NE(k)*Qtjk1(k,m)
    class(k,m+ir) = NE(k)*Qtjk1(k,m+ir)
53  continue
55  continue
c
c .... Calcul de la Statistique T-Carre ....
TCAR = 0.
do 62 k=1,is
  do 61 j=1,2*ir
    TCAR = TCAR+((IXE(k,j)-class(k,j))**2)/class(k,j)
61  continue
62  continue
c
c ***** Comportement asymptotique de T-carre *****
c
do 64 k=1,is
  jk = 0
  do 63 j=(k-1)*ir+2,k*ir+1
    jk=jk+1
    DQtjk1(jk,1) = -cmc(k)*X(jk+1)
    DQtjk1(jk+ir,1) = -coec(k,jk)*X(jk+1)
63  continue
64  continue
c
do 67 k=1,is
  do 66 j=(k-1)*ir+2,k*ir+1
    do 65 jk=1,ir
      if (jk.eq.j-1) then
        DQtjk1(jk,j)=1.+X(1)*cmc(k)-X(1)*coec(k,j-1)*X(jk+1)
        DQtjk1(jk+ir,j)=X(1)*coec(k,jk)*DQtjk1(jk,j)
      else
        DQtjk1(jk,j)=-X(1)*coec(k,j-1)*X(jk+1)
        DQtjk1(jk+ir,j)=X(1)*coec(k,jk)*DQtjk1(jk,j)
      end if
65  continue
66  continue
67  continue
c
do 70 k=1,is
  do 69 j=1,ir+1
    do 68 m=1,2*ir
      DQtjk1(m,j)= DQtjk1(m,j)/((1.+X(1)*cmc(k))**2)
      DQtjk1(m,j)= DQtjk1(m,j)/sqrt(Qtjk1(k,m))

```

```

68  continue
69  continue
70  continue
c
  do 73 j=1,1+ir
    do 72 m=1,1+ir
      SQtjk1(j,m) = 0.
      do 71 i=1,2*ir
        SQtjk1(j,m)=SQtjk1(j,m)+DQtjk1(i,j)*DQtjk1(i,m)
71    continue
72  continue
73  continue
c
  do 78 k=1,is
    do 76 m=1,2*ir
      do 74 j=1,2*ir
        QQtjk1(m,j) = sqrt(Qtjk1(k,m)*Qtjk1(k,j))
74    continue
76  continue
78  continue
c
  do 82 k=1,is
    FQtjk1(1,1) = -cmc(k)*ne(k)/X(1)*((1.+X(1)*cmc(k))**2)
    do 80 j=1,ir
      FQtjk1(1,j+1)= - ne(k)*coec(k,j)/((1.+X(1)*cmc(k))**2)
      FQtjk1(j+1,1)= FQtjk1(1,j+1)
80  continue
82  continue
    do 88 k=1,is
      jk=0
      do 86 j=(k-1)*ir+2,k*ir+1
        jk=jk+1
        do 84 m=1,ir
          if(jk.eq.m) then
            FQtjk1(jk+1,m+1) = -(1.+X(1)*coec(k,jk))/X(j)
            FQtjk1(jk+1,m+1) = FQtjk1(jk+1,m+1)+((X(1)*coec(k,jk))
*              **2)/(1.+X(1)*cmc(k))
            FQtjk1(jk+1,m+1) = FQtjk1(jk+1,m+1)*ne(k)/
*              (1.+X(1)*cmc(k))
          else
            FQtjk1(jk+1,m+1) = ne(k)*(X(1)**2)*coec(k,jk)*
*              coec(k,m)/((1.+X(1)*cmc(k))**2)
          end if
84    continue
86    continue
88    continue
      do 92 j=1,ir+1
        do 90 m=1,ir+1
          FQtjk1(j,m) = FQtjk1(j,m)+SQtjk1(j,m)
90    continue
92    continue
c
  ia = 1+ir
  ifail=1
c  call F01AAF(FQtjk1,ia,ia,FQINV,ia,WKSPCE,ifail)
c  if(ifail.ne.0) stop
c
  ifail=1
c  call F01AAF(SQtjk1,ia,ia,SQINV,ia,WKSPCE,ifail)
c  if(ifail.ne.0) stop
  do 95 j=1,2*ir
    do 94 m=1,1+ir
      GSIG(j,m) = 0.
      GRAO(j,m) = 0.
    do 93 i=1,1+ir

```

```

        GSIG(j,m) = GSIG(j,m)+DQtjk1(j,i)*FQINV(i,m)
        GRAO(j,m) = GRAO(j,m)+DQtjk1(j,i)*SQINV(i,m)
93    continue
94    continue
95    continue
    do 98 j=1,2*ir
        do 97 m=1,2*ir
            FSIG(j,m) = 0.
            FRAO(j,m) = 0.
            do 96 i=1,1+ir
                FSIG(j,m) = FSIG(j,m)+GSIG(j,i)*DQtjk1(m,i)
                FRAO(j,m) = FRAO(j,m)+GRAO(j,i)*DQtjk1(m,i)
96            continue
97            continue
98        continue
c
    do 102 j=1,2*ir
        do 100 m=1,2*ir
            if(j.eq.m) then
                FSIG(j,m) = -FSIG(j,m)-QQtjk1(j,m)+1.
                FRAO(j,m) = -FRAO(j,m)+1.
                RQQtjk1(j,m) = -QQtjk1(j,m)+1.
            else
                FSIG(j,m) = -FSIG(j,m)-QQtjk1(j,m)
                FRAO(j,m) = -FRAO(j,m)
                RQQtjk1(j,m) = -QQtjk1(j,m)
            end if
100        continue
102    continue
c
    do 108 j=1,2*ir
        do 106 m=1,2*ir
            RAO(j,m) = 0.
            do 104 i=1,2*ir
                RAO(j,m) = RAO(j,m)+FRAO(j,i)*RQQtjk1(i,m)
104            continue
106        continue
108    continue
        do 114 j=1,2*ir
            do 112 m=1,2*ir
                FICHR(j,m) = 0.
                do 110 i=1,2*ir
                    FICHR(j,m) = FICHR(j,m)+RAO(j,i)*FRAO(i,m)
110                continue
112            continue
114        continue
c
    ia = 2*ir
    ifail = 1
c    call F02ABF(FSIG,ia,ia,RJQtjk1,VJQtjk1,ia,EJQtjk1,ifail)
c    if (ifail.ne.0) stop
c
    ifail = 1
c    call F02ABF(FICHR,ia,ia,RRAO,VRAO,ia,ERA0,ifail)
c    if (ifail.ne.0) stop
c
c    **** Estimation de la variance sigma-carre ****
c
    do 215 k=1,is
        sigmak1(k) = cmc(k)*(NE(k)+X(1)*cmc(k)-2)/
*            X(1)*((1.+X(1)*cmc(k))**2)
        sigmak1(k+is) = cmc(k)*(X(1)*cmc(k)-1)/
*            X(1)*((1.+X(1)*cmc(k))**2)
215    continue
    sigmak2(1) = 0.

```

```

    sigmak2(2) = 0.
    do 216 k=1,is
        sigmak2(1) = sigmak2(1)+sigmak1(k)
        sigmak2(2) = sigmak2(2)+sigmak1(k+is)
216 continue
    sigmak2(3) = sigmak2(1)+(ir-1)*sigmak2(2)
c
    sigma2(1) = (TCAR-(is*ir)+is+1)/sigmak2(3)
    sigma2(3) = 0.
    do 218 k=1,is
        sigma2(3) = sigma2(3)+NE(k)
218 continue
    sigma2(3) = is*(ir-1)+sigma2(3)/4
    sigma2(2) = (X(1)**2)*(TCAR-is*(ir-1))/sigma2(3)
    sigma2(3) = (X(1)**2)*(TCAR-(is-1)*(ir-1))/sigma2(3)
c
    return
end
c
SUBROUTINE INFO(is,ir,n,X,coec,cmc,PROB,NE,HESS,VARCO)
    integer is,ir,n,i,j,k,NE(is),ifail,ia,inut
    real HESS(n+is,n+is),VARCO(n+is,n+is),WKSPCE(10000),
    * X(n),coec(is,ir),cmc(is),PROB(is,ir)
    external F01AAF,CMOY
c
c .... Stockage de la matrice Hessienne. Elle est remplacee par
c l'information de Fisher.
    do 46 j=1,n+is
        do 44 i=1,n+is
            HESS(i,j) = 0.
44 continue
46 continue
c
    do 48 k=1,is
        HESS(1,1) = HESS(1,1)+(X(1)*ne(k)*CMOY(k,is,ir,n,X,coec))/
        * ((1.+X(1)*CMOY(k,is,ir,n,X,coec))**2)
48 continue
    do 52 k=1,is
        i=0
        do 50 j = (k-1)*ir+2,k*ir+1
            i=i+1
            HESS(1,j) = X(1)*coec(k,i)*ne(k)/
            * ((1.+X(1)*CMOY(k,is,ir,n,X,coec))**2)
            HESS(j,1) = HESS(1,j)
50 continue
52 continue
c
    do 58 k=1,is
        mk = 0
        do 56 m = (k-1)*ir+2,k*ir+1
            mk = mk+1
            jk = 0
            do 54 j = (k-1)*ir+2,k*ir+1
                jk = jk+1
                if (m.ne.j) then
                    HESS(m,j) = - ne(k)*(X(1)**2)*coec(k,mk)*coec(k,jk)/
                    * ((1.+X(1)*CMOY(k,is,ir,n,X,coec))**2)
                else
                    HESS(j,j) = ((1.+X(1)*coec(k,jk))/X(j))-
                    * ((X(1)*coec(k,jk))**2)/(1.+X(1)*CMOY(k,is,ir,n,X,coec))
                    HESS(j,j)=ne(k)*HESS(j,j)/(1.+X(1)*CMOY(k,is,ir,n,X,coec))
                end if
54 continue
56 continue
58 continue

```

```

c
  do 66 i=n+1,n+is
    do 64 j=(i-n-1)*ir+2,(i-n)*ir+1
      HESS(j,i) = -1.
64  continue
66  continue
    do 76 k=n+1,n+is
      do 74 j=1,n
        HESS(k,j) = HESS(j,k)

c    do 76 k=n+1,n+is
c      do 74 j=(k-n-1)*ir+2,(k-n)*ir+1
c        HESS(k,j) = -1.
74  continue
76  continue
    write(6 ,2912) ((HESS(i,j),j=1,n+is),i=1,n+is)
2912 format (7f10.4)

c ..... Inversion de HESS .....
  ia=n+is
  inut =n+is
  ifail=1
  call F01AAF(HESS,ia,ia,VARCO,inut,WKSPCE,ifail)
  write(6 ,2912) ((HESS(i,j),j=1,n+is),i=1,n+is)
  write(6 ,2912) ((VARCO(i,j),j=1,n+is),i=1,n+is)
  return
end

```

Modèle.2

```

C .... Recherche de solutions aux equations de vraisemblance
c par utilisation de la Subroutine E04VCF de la Bibliotheque NAG
C ( C.I.T.I Lille). Ce programme a ete teste sur les donnees
C de Tanner avant d'etre applique a des tableaux 2xSxR.....
C .... MODELE 2.
C
C ..... Parametres .....
Integer IS, IR
parameter (IS=1,IR=2)

c
  Integer  n, nclin, ncnln, nctotl, nrowa, nrowj, liwork,
*          lwork, nrowr
  parameter (n=1+IS*IR,nclin=IS,ncnln=0,nctotl=n+nclin+ncnln,
*          nrowa=IS,nrowj=1,nrowr=n,liwork=3*n+nclin+ncnln,
*          lwork=3*n*n+(n+2)*(nclin+ncnln)+n*nrowj
*          +12*n+5*(nclin+ncnln)+nrowa+5*n+nrowj+nclin)
  integer  nin, nout
  parameter (nin=5,nout=6)
  real     zero, one, two, bigbnd
  parameter (zero=0.0e0,one=1.0e0 ,two=2.0e0)

c
c ..... Local Scalars .....
c
  real  ftol, OBJF,EPSAF, EPSMCH,ETA,RTEPS
  INTEGER i,ifail,itmax,j,msglvl,iter,mode,nstate
  LOGICAL COLD, FEALIN, ORTHOG

c
c ..... Locals Arrays
c
  real  A(nrowa,n),bl(nctotl), bu(nctotl),c(nrowj),
1      cjac(nrowj,n),clamda(nctotl), OBJGRD(n),
2      work(lwork), x(n), featol(nctotl), r(nrowr,n)
  integer  istate(nctotl), iwork(liwork)

c
  real  X02AAF

```

```

EXTERNAL X02AAF
c ..... Arrays in common
c
Integer IXE1(IS,IR),IXE2(IS,IR),IXE(IS,2*IR),NE(IS),NE1(IS)
Integer IXE3(IS,IR),NE2(IS),ntot(is),k
Integer G05ECF,G05EYF
Real Pjk(is,ir),se(is),cocm(is),teta,tetak(is)
Real Qtjk2(is,2*ir),som(is,2*ir+1),sigma,sigma2(3),TCAR,
*   sigmak1(2*is),sigmak2(3),tetabar
Real XT0,XT1,XT2,XT3,XC,G05CBF,CMC(IS),class(is,2*ir)
Real COEC(IS,IR),PROB(IS,IR),EFFAP(IS),G05DAF

c
COMMON/ACC1/IXE1, IXE2,IXE3,IXE, NE, NE1,NE2
COMMON/FACC1/ COEC
EXTERNAL CONFUN, E04VCF, E04ZCF, OBJFUN, X04ABF,CMOY,ESTIMA
EXTERNAL G05DAF,G05CBF,G05ECF,G05EYF,G05CAF,GMULT,GENENOR
INTRINSIC ABS,SQRT
c ..... Executable Statements .....
c
c .... Choix de la vraie valeur de teta: effet moyen ....
teta = .80
c
write(nout,fmt=88996) teta
c .... Choix de la variance sigma-carre ....
sigma = .050000
sigma = sqrt(sigma)
c
c .... Generation de l'effet de la mesure de chaque site ....
CALL GENENOR(is,ir,teta,sigma,tetak)
write(nout,*) (tetak(i),i=1,is)
c
tetabar = 0.
do 41 k=1,is
tetabar = tetabar+tetak(k)
41 continue
tetabar = tetabar/is
write(nout,*) tetabar
c
c .... Choix des vraies probabilites Pjk avant la mesure ....
do 15 k=1,is
se(k) = 0.
do 13 j=1,ir
c CALL G05CBF((is*ir)-(k+j))
c Pjk(k,j) = G05DAF(0.,1.)
Pjk(k,j) = G05CAF(X)
se(k) = se(k)+Pjk(k,j)
13 continue
do 14 j=1,ir
14 Pjk(k,j) = Pjk(k,j)/se(k)
15 continue
c
c .... Choix des coefficients de controle COEC(k,j) ....
do 32 k=1,is
do 30 j=1,ir
CALL G05CBF(k*j)
COEC(k,j) = G05DAF(.5,2.5)
30 continue
32 continue
c .... Calcul des vrais coefficients moyens de controle "c(k)" ....
do 42 k=1,is
cocm(k) = 0.
do 40 j=1,ir
cocm(k) = cocm(k)+COEC(k,j)*Pjk(k,j)
40 continue
42 continue

```

```

c
c ..... Simulation de tableaux Avant-After 2xSxR .....
c
CALL GMULT(IS,IR,tetak,Pjk,COEC,cocm,Qtjk2,som,NE,ntot,IXE)
c
do 7 k=1,IS
do 5 j=1,IR
IXE1(k,j) = IXE(k,j)
IXE2(k,j) = IXE(k,j+IR)
5 continue
7 continue
DO 24 k=1,IS
NE1(k) = 0
NE2(k) = 0
do 22 m=1,IR
NE1(k) = NE1(k)+ixe1(k,m)
NE2(k) = NE2(k)+ixe2(k,m)
22 continue
do 23 j=1,IR
c PROB(k,j) = float(ixe1(k,j))/NE1(k)
PROB(k,j) = 1./IR
23 continue
24 CONTINUE
c
do 8 k=1,is
write(6,88997) (Pjk(k,m),m=1,ir),(COEC(k,l),l=1,ir),cocm(k)
8 continue
write(6,88998)
do 9 k=1,is
write(6,89002) (Qtjk2(k,l),l=1,2*ir),som(k,2*ir+1)
9 continue
write(6,88998)
write(6,88999)
write(6,89000) is,ir
do 12 k=1,is
write(6,89001) (ixe(k,m),m=1,2*ir),ne(k)
12 continue
c
DO 80 k=1,IS
do 78 j=1,IR
ixe3(k,j) = ixel(k,j)+ixe2(k,j)
78 continue
80 CONTINUE
c
c ..... Appel de la subroutine E04VCF pour la recherche des EMVC .....
write(nout,fmt=99999)
c
call X04ABF(1,nout)
itmax=100
msglvl=5
eta = 0.2e0
epsmch = X02AAF(zero)
c ftol = 10.0*epsmch
ftol = 1.0e-5
rsteps = sqrt(epsmch)
do 2 j=1,nctotl
featol(j) = rsteps
2 continue
bigbnd=1.0e10
c
c ..... A = Matrice des contraintes lineaires .....
do 11 j=1,n
do 10 i=1,n
A(i,j)= 0.
10 continue
11 continue

```

```

11 continue
  do 17 i=1,ncnln
    do 16 j=(i-1)*IR+2,i*IR+1
      A(i,j) = 1.
16 continue
17 continue
c .... Ensemble de toutes les bornes (inf et sup) du probleme ....
c   On commence par le vecteur BL(j) de toutes les bornes inf.
c   Ensuite, on remplit le vecteur BU(j) de toutes les bornes sup.
c   Il faut entrer d'abord les bornes sur les variables du probleme,
c   ensuite les bornes sur les contraintes lineaires generales, et
c .... enfin les bornes sur les contraintes non lineaires generales....
  BL(1) = 0.
  do 18 j=2,n
    BL(j) = .05
18 continue
  do 19 j=n+1,nctotl
    BL(j) = 1.
19 continue
  BU(1) = bigbnd
  do 20 j=2,n
    BU(j) = .998
20 continue
  do 90 j=n+1,nctotl
    BU(j) = 1.
90 continue
c
c .... X = point de Depart de l'optimisation ....
do 25 k=1,IS
  i=0
  do 21 j=(k-1)*IR+2,k*IR+1
    i=i+1
    X(j)=PROB(k,i)
21 continue
25 continue
  X(1)=.0200
c
  do 26 k=1,IS
    CMC(k) = CMOY(k,IS,IR,n,X,COEC)
26 continue
  write(nout,*) (CMC(k),k=1,IS)
  write(nout,fmt=99996) (X(j),j=1,n)
c
  nstate = 1
  mode = 1
c
c
c .... Definition de la precision absolue sur la OBJF .....
CALL OBJFUN(mode,n,x,OBJF,OBJGRD,nstate)
epsaf = epsmch*ABS(OBJF)
COLD = .TRUE.
FEALIN = .TRUE.
ORTHOG = .TRUE.
c
c Check the gradients using routine E04ZCF
c
  ifail=-1
  call E04ZCF(n,ncnln,nrowj,CONFUN,OBJFUN,c,cjac,objf,objgrd,
1    x,work,lwork,ifail)
  if(ifail.eq.0) then
    write(nout,fmt=99998)
c
c Solve the problem from a cold start.
c
  ifail=-1

```

```

      call E04VCF(itmax,msglvl,n,nclin,ncnln,nctotl,nrowa,nrowj,nrowr,
1         bigbnd,epsaf,eta,ftol,A,BL,BU,featol,CONFUN,
2         OBJFUN,cold,fealin,orthog,X,istate,R,iter,c,cjac,
3         OBJF,OBJGRD,clamda,iwork,liwork,work,lwork,ifail)
      if(ifail.eq.0) then
c
      do 140 j=1,n
        X(j)=X(j)
140 continue
c      ... Reset the absolute precision of the objective function ....
      epsaf = epsmch*ABS(OBJF)
      COLD = .false.
      write(nout,fmt=99997)
      write(nout,fmt=99996) (X(j),j=1,n)
c
      ifail = -1
      call E04VCF(itmax,msglvl,n,nclin,ncnln,nctotl,nrowa,nrowj,nrowr,
1         bigbnd,epsaf,eta,ftol,A,BL,BU,featol,CONFUN,
2         OBJFUN,cold,fealin,orthog,X,istate,R,iter,c,cjac,
3         OBJF,OBJGRD,clamda,iwork,liwork,work,lwork,ifail)
      END IF
      if (ifail.ne.0) write(nout,fmt=99995) ifail
      else
        write(nout,fmt=99994) ifail
      end if
c
      do 28 k=1,IS
        CMC(k) = CMOY(k,IS,IR,n,X,COEC)
28 continue
        write(nout,*) (CMC(k),k=1,IS)
c
      CALL ESTIMA(is,ir,n,X,COEC,CMC,PROB,Qtk2,IXE,NE,class,TCAR,
*         sigmak1,sigmak2,sigma2)
c
      stop
c
99999 format(/ ' E04VCF Example Program Results: Modele2 ',/1x)
99998 format(/ ' Successful exit from E04ZCF.')
99997 format(/ ' A run of the same exemple with a warm start ...')
99996 format(/ ' Initial  X.',/ 5(1x,f10.5))
99995 format(/ ' E04VCF terminated with ifail = ' ,I3)
99994 format(/ ' Incorrect gradients. ifail = ' , I3)
88996 format(' La vraie valeur de teta est =', f6.4,/ )
88997 format(2(1x,f10.7),' ',2(1x,f10.7),' ',f10.7)
88998 format( / )
88999 format(/, ' *****      Tableau Avant-Apres 2xSxR      *****',/)
89000 format(' Nombre de Sites S = ' ,i3,';Nombre de Types R = ' ,i3, /)
89001 format(5(5x,i5))
89002 format(2(1x,f10.7),' ',2(1x,f10.7),8x,f2.0)
c
      end

c      .... Definition du coefficient de controle moyen ....
      FUNCTION CMOY(k,IS,IR,n,X,COEC)
        integer IS,IR,i,j,k,n
        real X(n),COEC(IS,IR)
        real CMOY
        CMOY = 0.
        i=0
        do 90 j=(k-1)*IR+2,k*IR+1
          i=i+1
          CMOY = CMOY+COEC(k,i)*X(j)
90 continue
        return
      END

```

```

C
  Subroutine OBJFUN(MODE,N,X,OBJF,OBJGRD,NSTATE)
C ..... Scalars Arguments .....
  Integer IS, IR
  Parameter(IS=1,IR=2)
  Integer ixe1(IS,IR),ixe2(IS,IR),ne(IS),ne1(IS),ne2(IS)
  integer mode,n,nstate,i,j,k,ixe3(IS,IR),ixe(IS,2*IR)
  real OBJF,COEC(IS,IR),CMC(IS)
  real OBJGRD(n), X(n)
  COMMON/ACC1/IXE1, IXE2,IXE3, IXE, NE,NE1,NE2
  COMMON/FACC1/ COEC
c ..... Executable statements .....
EXTERNAL CMOY
c ..... Fonction de Vraisemblance .....
  OBJF = 0.
  DO 95 k=1,IS
  OBJF=OBJF+NE2(k)*LOG(X(1))+NE2(k)*LOG(CMOY(k,IS,IR,n,X,COEC))-
1  NE(k)*LOG(1.+X(1)*CMOY(k,IS,IR,n,X,COEC))
    i=0
    do 93 j=(k-1)*IR+2,k*IR+1
      i=i+1
      OBJF = OBJF+IXE3(k,i)*LOG(X(j))
93  continue
95  continue
  OBJF = - OBJF
c
c ..... Vecteur gradient .....
  OBJGRD(1) = 0.
  do 105 k=1,IS
    OBJGRD(1) = OBJGRD(1)+(NE2(k)/X(1))-
1    (NE(k)*CMOY(k,IS,IR,n,X,COEC))/
2    (1.+X(1)*CMOY(k,IS,IR,n,X,COEC))
105  continue
c
  do 110 k=1,IS
    m=0
    do 108 j=(k-1)*IR+2,k*IR+1
      m=m+1
      OBJGRD(j) = (IXE3(k,m)/X(j))-
1      (X(1)*COEC(k,m)*NE(k))/(1.+X(1)*CMOY(k,IS,IR,n,X,COEC))+
2      (COEC(k,m)*NE2(k))/CMOY(k,IS,IR,n,X,COEC)
108  continue
110  continue
c
  do 112 m = 1,n
    OBJGRD(m) = - OBJGRD(m)
112 CONTINUE
  return
end
c
  SUBROUTINE CONFUN(mode,ncnln,n,nrowj,x,c,cjac,nstate)
c The zero elements of jacobian matrix are set only once. This
c occurs during the first call to confun (nstate = 1).
c
c ..... Parameters .....
  Integer IS,IR
  Parameter(IS=2,IR=2)
  real zero, two, one
  parameter (zero=0.0e0,one=1.0e0,two=2.0e0)
  Integer ixe1(IS,IR),ixe2(IS,IR),ne(IS),ixe(IS,2*IR),ne1(IS)
  integer mode, n, ncnln, nrowj, nstate,ixe3(IS,IR),ne2(IS)
  Real X(n), c(nrowj),cjac(nrowj,n),coec(IS,IR)
c ..... Scalars Arguments .....
  n = 1+IS*IR

```

```

c ..... Local Scalars .....
c   if (nstate.eq.1) then
c     do 116 j=1,n
c       do 114 i=1,nrowj
c         cjac(i,j) = 0.
114    continue
116    continue
c   - - -
c   end if
c   return
c   end

c
c   SUBROUTINE GMULT(IS,IR,tetak,Pjk,COEC,cocm,Qtjk2,som,NE,ntot,IXE)
c   .... Simulation d'une loi multinomiale par utilisation de la
c   .... bibliotheque NAG .....
c   On tire de facon aleatoire le nombre total d'accidents de
c   chaque site experimental avec une distribution de Poisson.
c   On distribue de facon multinomiale Avant et Apres ce nombre
c   dans les deux fois ir categories d'accidents.
c
c   Integer is,ir
c   Integer ne(is),ntot(is),ifail,ixe(is,2*ir),i,j,k
c   Integer G05ECF,G05EYF,NR1
c   Real   Pjk(is,ir),som(is,2*ir+1),cocm(is),G05DAF
c   Real   Qtjk2(is,2*ir),tetak(is),COEC(is,ir),G05CAF
c   Real   XT0,XT1,XT2,XT3,XC,R1(10000)
c   External G05ECF,G05EYF,G05CAF,G05DAF
c
c   .... Calcul des Probabilites Qtjk de la loi multinomiale ....
c   do 55 k=1,is
c     do 51 j=1,ir
c       Qtjk2(k,j) = Pjk(k,j)/(1.+tetak(k)*cocm(k))
c       Qtjk2(k,j+ir) = tetak(k)*cocm(k)*Pjk(k,j)/
c       * (1.+tetak(k)*cocm(k))
51    continue
55    continue
c
c   .... Choix aleatoire du nombre ne(k)=x1..k d'accidents du site k
c   G05CBF permet d'initialiser le tirage de ne(k); Choisir
c   le parametre de G05CBF constant si on veut les ne(k) identiques.
c   do 2 k=1,is
c     CALL G05CBF(is*ir+(k-1))
c     ifail = 0
c     XC = G05DAF(100.,300.)
c     XT0 = sqrt(XC)
c     XT1 = 20.+20.*XT0
c     XT2 = XC+7.15*XT0
c     XT3 = XC-7.15*XT0
c     NR1 = MAX(int(XT1),int(XT2+8.5))-MAX(0,int(XT3))+4)
c     CALL G05ECF(XC,R1,NR1,IFAIL)
c     if (ifail.ne.0) stop
c     ne(k) = G05EYF(R1,NR1)
2    continue
c   .... Initialisation de l'effectif ix de chaque classe
c   do 11 m=1,is
c     do 11 j=1,2*ir
11    ix(m,j) = 0
c   .... Decoupage de l'intervalle (0,1) en 2*ir classes par
c   utilisation des Qtjk2(k,j). Les som(k,j) representent les
c   points de decoupage.
c   do 18 k=1,is
c     som(k,1) = 0.
c     do 17 m=2,2*ir+1
c       som(k,m) = som(k,m-1)+Qtjk2(k,m-1)
17    continue

```

```

18 continue
c
c .... Remplissage des elements ixé par tirage d'une proba p
c et comparaison de p aux som(k,j) ....
do 21 k=1,is
    ntot(k) = ne(k)
1005    p = G05CAF(X)
        do 20 j=1,2*ir
            if (som(k,j).lt.p.and.p.lt.som(k,j+1)) then
                ixé(k,j)=ixé(k,j)+1
                ntot(k)=ntot(k)-1
            else
                ixé(k,j)=ixé(k,j)
                ntot(k) = ntot(k)
            end if
        20 continue
        if (ntot(k).gt.0) go to 1005
21 continue
c
    return
end
c
    subroutine GENENOR(is,ir,teta,sigma,tetak)
c .... Programme de generation de nombres pseudo-aleatoires
c provenant d'une distribution normale ....
c
    integer is,ir
    real G05DDF,G05CBF,tetak(is),teta,sigma
    external G05DDF,G05CBF
c
    call G05CBF(0)
    do 20 i=1,is
        tetak(i) = G05DDF(teta,sigma)
20 continue
    return
end
c
SUBROUTINE ESTIMA(is,ir,n,X,COEC,CMC,PROB,Qtjk2,IXE,NE,class,
* TCAR,sigmak1,sigmak2,sigma2)
    Integer is,ir,n,i,j,k,m,IXE(is,2*ir),NE(is)
    Real X(n),CMC(is),Qtjk2(is,2*ir),TCAR,sigma2(3),COEC(is,ir),
    * PROB(is,ir),class(is,2*ir),sigmak1(2*is),sigmak2(3)
    External CMOY
c
    do 25 k=1,is
        i=0
        do 21 j=(k-1)*ir+2,k*ir+1
            i=i+1
            PROB(k,i) = X(j)
21 continue
25 continue
c
    do 55 k=1,is
        do 51 j=1,ir
            Qtjk2(k,j) = PROB(k,j)/(1.+X(1)*CMC(k))
            Qtjk2(k,j+ir) = PROB(k,j)*COEC(k,j)*X(1)/(1.+X(1)*CMC(k))
            class(k,j) = NE(k)*Qtjk2(k,j)
            class(k,j+ir) = NE(k)*Qtjk2(k,j+ir)
51 continue
55 continue
c
c .... Calcul de la Statistique T-Carre ....
TCAR = 0.
do 62 k=1,is
    do 61 j=1,2*ir

```

```

        TCAR = TCAR+((LXE(k,j)-class(k,j))**2)/class(k,j)
61  continue
62  continue
c
c  .... Estimation de la variance sigma-carre ....
do 65 k=1,is
    sigmak1(k) = cmc(k)*(NE(k)+X(1)*cmc(k)-2)/
*      X(1)*((1.+X(1)*cmc(k))**2)
    sigmak1(k+is) = cmc(k)*(X(1)*cmc(k)-1)/
*      X(1)*((1.+X(1)*cmc(k))**2)
65  continue
    sigmak2(1) = 0.
    sigmak2(2) = 0.
    do 70 k=1,is
        sigmak2(1) = sigmak2(1)+sigmak1(k)
        sigmak2(2) = sigmak2(2)+sigmak1(k+is)
70  continue
    sigmak2(3) = sigmak2(1)+(ir-1)*sigmak2(2)
c
    sigma2(1) = (TCAR-(is-1)*(2*ir-1))/sigmak2(3)
    sigma2(3) = 0.
    do 80 k=1,is
        sigma2(3) = sigma2(3)+NE(k)
80  continue
    sigma2(3) = is*(ir-1)+sigma2(3)/4
    sigma2(2) = (X(1)**2)*(TCAR-is*(ir-1)+2)/sigma2(3)
    sigma2(3) = (X(1)**2)*(TCAR-(is-1)*(2*ir-1))/sigma2(3)
c
    return
end

```

Modèle de Tanner

```

C  .... Recherche de solutions aux equations de vraisemblance
C  par utilisation de la Subroutine E04VDF de la Bibliotheque NAG
C  ( C.I.T.I Lille). Ce programme a ete teste sur un exemple
C  donne par NAG avant d'etre applique au Modele.Poisson (Modele2)
C  .... MARK ....
C
C  ..... Parametres .....
Integer IS
parameter (IS=7)
c
Integer n, nclin, ncnln, nctotl, nrowa, nrowj, liwork,
*      lwork
parameter (n=1,nclin=0,ncnln=0,nctotl=n+nclin+ncnln,
*      nrowa=1,nrowj=1,liwork=3*n+nclin+ncnln,
*      lwork=3*n*n+(n+2)*(nclin+ncnln)+n*nrowj
*      +12*n+5*(nclin+ncnln)+nrowa+5*n+nrowj+nclin)
integer nin, nout
parameter (nin=5,nout=6)
real zero, one, two, bigbnd
parameter (zero=0.0e0,one=1.0e0,two=2.0e0)
c
c  .... Local Scalars ....
c
real ctol, ftol, OBJF
INTEGER i,ifail,itmax,j,msglvl
c
c  .... Locals Arrays
real A(nrowa,n),bl(nctotl),bu(nctotl),c(nrowj),
1  cjac(nrowj,n),clamda(nctotl), OBJGRD(n),
2  work(lwork), x(n)
integer istate(nctotl), iwork(liwork)

```

```

c ..... Arrays in common
c
  Integer IXE1(IS),IXE2(IS),NE(IS)
  Real COEC(IS)
  Real FONCT,DPART
c ..... External Subroutines .....
COMMON/ACC1/IXE1, IXE2, NE
COMMON/FACC1/COEC
EXTERNAL CONFUN, E04VDF, E04ZCF, OBJFUN, X04ABF
c ..... Executable Statements .....
c read (nin,fmt=')')
write(nout,fmt=99999)
c
  DATA COEC/1.04,1.25,1.11,2.36,1.13,1.69,1.61/
  DATA IXE1/1,6,9,16,10,2,5/
  DATA IXE2/6,3,5,5,0,2,0/
  DATA NE/7,9,14,21,10,4,5/

  call X04ABF(1,nout)
  itmax=50
  msglvl=1
  ftol=1.0e-5
  ctol=1.0e-4
  bigbnd=1.0e21
c
c ..... A = Matrice des contraintes lineaires .....
do 15 j=1,n
  do 10 i=1,n
    A(i,j)=zero
  10 continue
  15 continue
c .... Ensemble de toutes les bornes (inf et sup) du probleme ....
c .... On commence par le vecteur BL(j) de toutes les bornes inf.
c .... Ensuite, on remplit le vecteur BU(j) de toutes les bornes sup.
c .... Il faut entrer d'abord les bornes sur les variables du probleme,
c .... ensuite les bornes sur les contraintes lineaires generales, et
c .... enfin les bornes sur les contraintes non lineaires generales....
do 30 j=1,nctotl
  BL(j) = zero
  BU(j) = bigbnd
30 continue
c .... X = point de Depart de l'optimisation ....
c DO 31 k=1,IS
c POIDS(k) = 1./(IS)
c EFFAP(k) = float(IXE2(k))/(IXE1(k)*COEC(k))
c 31 CONTINUE
  X(1) = 0.240
c DO 32 k=1,IS
c X(1) = X(1)+POIDS(k)*EFFAP(k)
c 32 CONTINUE
c
c DO 60 k=1,IS
c FIXE1(k) = one
c DO 50 i=2,IXE1(k)
c 50 FIXE1(k) = FIXE1(k)*i
c FIXE1(k) = log(FIXE1(k))
c 60 CONTINUE
c DO 75 k=1,IS
c FIXE2(k) = one
c DO 70 i=2,IXE2(k)
c 70 FIXE2(k) = FIXE2(k)*i
c FIXE2(k) = log(FIXE2(k))
c 75 CONTINUE
c DO 85 k=1,IS

```

```

c   FNE(k) = one
c   DO 80 i=2,NE(k)
c 80   FNE(k) = FNE(k)*i
c   FNE(k) = log(FNE(k))
c 85 CONTINUE
c   DO 90 k=1,IS
c   CONST(k) = NE(k)*LOG(NE(k))+EXP(-NE(k))-(FIXE1(k)+FIXE2(k))
c 90 CONTINUE

```

```

c   Check the gradients using routine E04ZCF

```

```

      ifail=1
      call E04ZCF(n,ncnln,nrowj,CONFUN,OBJFUN,c,cjac,objf,objgrd,
1      x,work,lwork,ifail)
      if(ifail.eq.0) then
        write(nout,fmt=99998)

```

```

c
c .....   Resolution du Probleme   .....
c

```

```

      ifail=1
      call E04VDF(itmax,msglvl,n,ncnln,ncnln,nctotl,nrowa,nrowj,ctol,
1      ftol,A,BL,BU,CONFUN,OBJFUN,x,istate,c,cjac,OBJF,
2      objgrd,clamda,iwork,liwork,work,lwork,ifail)
      else

```

```

c
      write(nout,fmt=99997) ifail
      end if
      write(*,*) x
      WRITE(*,*) OBJF
      WRITE(*,*) itmax,istate
      stop

```

```

c
99999 format(' E04VDF Example Program Results',/1x)
99998 format('/ Successful exit from E04ZCF.')
99997 format('/ Incorrect Gradients. ifail =',i3)
      end

```

```

FUNCTION FONCT(X)
PARAMETER (IS=7)
real coec(IS)
integer ixel(IS),ixe2(IS),NE(IS)
Real FONCT
COMMON/ACC1/IXE1, IXE2, NE
COMMON/FACC1/ COEC
FONCT = 0.
DO 92 k=1,IS
  FONCT = FONCT+IXE2(k)*log(X)-NE(K)*log(1.+X*COEC(k))
92 CONTINUE
END

```

```

FUNCTION DPART(X)
PARAMETER (IS=7)
real coec(IS)
integer ixel(IS),ixe2(IS),NE(IS)
COMMON/ACC1/IXE1, IXE2, NE
COMMON/FACC1/ COEC
Real DPART
DPART = 0.
DO 96 k=1,IS
  DPART = DPART+(float(IXE2(k))/X)-
1      (COEC(k)*NE(k))/(1.+X*COEC(k))
96 CONTINUE
END

```

```

c      Subroutine OBJFUN(MODE,N,X,OBJF,OBJGRD,NSTATE)

C ..... Scalars Arguments .....
real   OBJF, FONCT, DPART
integer mode,n,nstate,i
c ..... Array Arguments .....
real   OBJGRD(n), X(n)
c ..... Definition de la fonction de vraisemblance et du vecteur
c ..... gradient associe .....
c ..... Executable Statements .....
OBJF = - FONCT(X)
OBJGRD(1) = - DPART(X)
return
end

c      SUBROUTINE CONFUN(mode,ncnln,n,nrowj,x,c,cjac,nstate)
c      The zero elements of jacobian matrix are set only once. This
c      occurs during the first call to confun (nstate = 1).
c
c ..... Parameters .....
real   zero, two
parameter (zero=0.0e0,two=2.0e0)
c ..... Scalars Arguments .....
integer mode, n, ncnln, nrowj, nstate
c ..... Array Arguments .....
real   c(nrowj), cjac(nrowj,n), x
c ..... Local Scalars .....
integer i,j
c ..... Executable Statements .....
if (nstate.eq.1) then
  do 40 j=1,n
    c(j) = zero
    do 20 i=1,n
      cjac(i,j) = zero
20    continue
40  continue
  end if
  return
end

```

Références bibliographiques.-

- [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.(1981) 22(10) p.535-542.
- [2] Aitchison J., Silvey S.D. - *Maximum likelihood estimation of parameters subject to restraints.* Annals of mathematical statistics, (1958), 29, 813-829.
- [3] Arbous A.G., Kerrich J.E - *Accidents Statistics and the concept of accident-proneness.* Biometrics (1951) vol.7 p.340-432.
- [4] 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. (1979), n° 74, p.596-603.
- [5] 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).
- [6] Box G.E.P., Jenkins G.M. - *Time series analysis: Forecasting and control.* 1976 San-Francisco: Holden Day.
- [7] Box G.E.P., Tiao G.C. - *Intervention analysis with applications to economic and environmental problems.* J.A.S.A.(1975) 70 p.70-79
- [8] Boyer, M. et al. - *Econometric models of accident distributions.* Rapport C.R.T. 672 (1990), Université de Montréal - Canada.
- [9] Brünhing E., Ernt - *Log-linear models in effectiveness studies: an application to simultaneous before-after comparison with control group.* International meeting on the evaluation of local traffic safety measure (1985). Paris.
- [10] Chase G.R. - *On the Chi-square Test when the Parameters are estimated Independently of the Sample.* J.A.S.A. (1972) vol.67, n°339. p.609-611.
- [11] Chauveau D. - *Algorithmes EM et SEM pour un mélange censuré de distributions de défaillances application à la fiabilité d'équipements électroniques en période de jeunesse.* Rev. Statistique Appliquée, (1992). XXXX (2), p.67-76.
- [12] Chauveau D. - *Extensions des algorithmes EM et SEM à la reconnaissance de mélanges censurées de distributions de défaillances.* Thèse de Doctorat en maths.(1991) Université Paris sud - Centre d'Orsay. France.
- [13] Chernoff H. - *Large sample theory : parametric case.* Ann. Math. Statist. (1956) 27, 1-27.
- [14] Chernoff H., Lehmann E.F. - *The use of Maximum likelihood estimates in χ^2 -tests for goodness of fit.* Ann. Math. Statistics.(1954) 25, p.579-586 .
- [15] Chinn S., Heller R.F. - *Some further results concerning regression to the mean.* American Journal of Epidemiology (1981) vol.114 n°6 p.902-905.
- [16] Council F.M., Reinfurt D.W., Campbell B.J., Koediger F.L., Carroll C.L., Dutt A.K., Dunham J.R. - *Accident Research manual.* (1980) University of North Carolina, Chapell Hill.
- [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,(1984) 46,n°3, p.440-464.
- [19] Cressie N., Read T.R.C. - *Pearson's χ^2 and the Loglikelihood Ratio Statistic G^2 : A comparative Review.* Inter. Statistical Review. (1989) n°57, 1, p.19-43.

- [20] Crowder M. - *On constrained maximum likelihood estimation with non i.i.d. observations.* Ann. Inst. Statist. Math.(1984) 36, A, 239-249 .
- [21] Dale J.R. - *Asymptotic Normality of Goodness-of-fit-Statistics for Sparse Product Multinomials.* J.R.Statist. Soc. B, (1986) 48,n°1, p.48-59.
- [22] 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.(1986) vol 18,n° 1, p.13-23.
- [23] Danielson S. - *Estimation of the effects of countermeasures on different types of accidents in the presence of regression effects.* Accid.Anal.and Prev.(1988) vol.20,n° 4, p.289-298.
- [24] Davis C.E. - *The effect of regression to the mean in epidemiologic and clinical studies.* American Journal of Epidemiology (1976) vol.104 n°5 p.493-498.
- [25] Droesbeke J.J., Fichet B., Tassi Ph. - *Séries chronologiques: théorie et pratique des modèles ARIMA.* (1989) Economica.
- [26] Feinberg S.E. - *The Use of Chi-squared Statistics for Categorical Data problems.* J.R.Statist. Soc. B,(1979) 41,n°1, p.54-64.
- [27] Fleischer G.A. - *Contingency table analysis for road safety studies.* NATO Advanced study institute, series E (1981) n°42.
- [28] Fletcher R. - *Practical methods of optimization: (vol.2) constrained optimization* (1981) John W. Chichester.
- [29] Fridstrom L. - *In favor of aggregate econometric accident models.* 6th Inter. Conference on travel behaviour(1991). Quebec.
- [30] 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 (1970) Paris.
- [31] Gerstein L.J. - *A new algorithm for computing the rank of a matrix.* The American Math. Monthly.(1988) vol 95 p.950-952.
- [32] Gill P.E., al. - *Practical optimization.* (1981) Academic Press. London
- [33] Gouriéroux C., Monfort A. - *Séries chronologiques et modèles dynamiques.* Economica (1990).
- [34] Gouriéroux C., Monfort A. - *Statistique et modèles économétriques.*(1989) Economica.
- [35] Gustavsson J., Svensson A. - *A poisson regression model applied to class of road accidents with small frequencies.* Scand. Journal of Statist.(1976) vol.3 p.49-60.
- [36] Hauer E. - *Selection for treatment as a source of bias in before-after studies.* Traff.Eng.Control(1980b) 21(8/9), p.419-421.
- [37] Hauer E. - *Bias-by-selection: overestimation of the effectiveness of safety countermeasures caused by the process of selection for treatment.* Accid. Anal and Prev.(1980a) vol.12, n° 26, p.113-117.
- [38] Hauer E. - *Comparison groups in road safety studies: An analysis.* Accid.Anal. and Prev.(1991) vol.23 n° 6, p.609-622.
- [39] Hauer E., NG J.C.C., Papaloannou P. - *Prediction in road safety studies: an empirical inquiry.* Accid.Anal. and Prev.(1991) vol.23 n° 6, p.595-607.

- [40] Hauer E. - *A note on three estimators of safety effect*. Traffic Eng. and Control.(1992) p.388-393.
- [41] Harvey A.C. - *Trends and cycles in macroeconomic time series*. J. Bus. Econ. Statist (1985) n° 3, p.216-227.
- [42] Harvey A.C., Durbin J. - *The effects of seat-belt legislation on British road causalities: a case study in structural time series modelling*. J. R. Statist. Soc. A (1986), 149 part 3, p.187-227.
- [43] Helfenstein U. - *When did a reduced speed limit shows an effect? Explonatory identification of an intervention time*. Accid.Anal. and Prev.(1990) vol.22 n° 1, p.79-87.
- [44] Hoc J. M. - *L'analyse planifiée des données en psychologie*. P.U.F (1983) Paris.
- [45] James K.E. - *Regression toward the mean in uncontrolled clinical studies*. Biometrics (1973), 29 p.121-130.
- [46] Lassarre S. - *A propos des tests statistiques, sur variables poissonniennes, utilisés dans le domaine de la sécurité routière*. Revue de Statistique Appliquée.(1977) vol, XXV n° 3.
- [47] Lassarre S. - *Three statistical treatments of contingency tables in road safety studies*. NATO Advanced study institute (1981) serie E n° 42.
- [48] 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 (1982) 1 p.297-306.
- [49] Lassarre S. - *Panorama des méthodes associées à l'évaluation de type avant-après*. Colloque international sur l'évaluation des mesures locales de sécurité routière.(1985) Paris.
- [50] 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. (1986) vol.18, n°2 p.129-134.
- [51] Lee S.Y., Jenrich R.I. - *A study of algorithms for covariance structure analysis with specific comparisons using factors analysis*. Psychometrika (1979) vol. 44, n° 1
- [52] Lee S.Y. - *Constrained estimation in covariance structure analysis*. Biometrika (1979) 66, 3, p. 539-45.
- [53] Minoux M. *Programmation mathématique : théorie et algorithmes*. Dunod (1987).
- [54] Molinari L. - *Distribution of the Chi-squared test in nonstandard situations*. Biometrika (1977) n° 64,1, p.115-121.
- [55] Monfort A. - *Cours de statistique mathématique*. Economica (1982).
- [56] Moore D.S., Spruill M.C. - *Unified Large-Sample Theory of general Chi-squared Statistics for tests of fit*. Annals of Math. Statistics. (1975) 3, p.599-616 .
- [57] Moore D.S. - *Generalized Inverses, Wald's Method, and the Construction of Chi-squared Tests of Fit*. J.A.S.A.(1977) vol.72, n°357, p.131-137.
- [58] Morton R. - *Analysis of extra-multinomial data derived from extra-Poisson variables conditional on their total*. Biometrika (1991) 78 n° 1, p.1-6.

- [59] Murthy V.K., Gafarian A.V. - *The limiting distributions of some variations of the Chi-square Statistic*. The Annals of Math. Statistics. (1970) vol 41, n° 1, p.188-202.
- [60] N.A.G - *Numerical Algorithms Group*. (1992) C.I.T.I. Lille
- [61] N'Guessan A., Langrand C. - *Une méthode d'estimation de l'effet moyen d'une mesure de sécurité routière*. XXIV^{es} Journées de Statistique (1992), Bruxelles.
- [62] N'Guessan A., Langrand C. *Sur la distribution asymptotique de certaines statistiques utilisées dans le domaine de la sécurité routière*. C. R. Acad. Sci.(1993) Paris t.316.
- [63] O.C.D.E., groupe de recherche routière. - *Symposium sur l'utilisation des méthodes statistiques dans l'analyse des accidents de la route*. (1970) France.
- [64] O.C.D.E., groupe de recherche routière. - *Méthodes d'évaluation de mesures correctives de sécurité routière*. (1981) Paris.
- [65] Oppe S. - *The use of multiplicative models for analysis of road safety data*. Accid.Anal.and Prev.(1979) vol 11, p.101-115.
- [66] Rao C.R. - *Linear Statistical Inference and its Applications*. J.Wiley (1965).
- [67] Saporta G. - *Probabilités, Analyse des données et Statistique*. Technip (1990). Paris.
- [68] Schmittlein D.C. - *Surprising Inferences from Unsurprising observations: do conditional Expectations really regress to the mean?*. The American Statistician (1989) vol.43 n° 3 p.176-183.
- [69] Silvey S.D. *The Langrangan multiplier test* : Annals of Math. Statistics (1959) 30 p.389-407.
- [70] Tanner J.C. - *A problem in the combination of accident frequencies*. Biometrika (1958) n° 45, p.331-342.
- [71] 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.
- [72] Wolter K.M. - *Introduction to Variance Estimation*. Springer-Verlag.(1985)
- [73] 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.(1988) vol 20,n° 3, p.199-214.
- [74] Wu C.F. - *On the convergence properties of the E.M. algorithm*. Ann. Statist. (1983) 11, p.95-103

RÉSUMÉ

Nous considérons le problème de l'évaluation statistique de l'efficacité d'une mesure de sécurité routière à partir de l'observation avant et après cette mesure des effectifs d'accidents enregistrés dans s sites comportant chacun r types d'accidents.

Des modèles basés sur une distribution multinomiale permettent alors de répartir le nombre total des 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). Nous généralisons ainsi le modèle de Tanner.

Nous estimons et analysons l'efficacité de la mesure de sécurité routière sous l'hypothèse d'égalité de l'effet aux différents sites expérimentaux. Nous suggérons ensuite une statistique pour éprouver l'égalité et étudions sous certaines conditions la distribution asymptotique de cette statistique.

Lorsque l'égalité de l'effet n'est pas admise, nous faisons l'hypothèse supplémentaire que les différents effets de la mesure forment un échantillon provenant d'une certaine loi. Nous proposons dans ces conditions une méthode d'approximation pour estimer l'espérance mathématique et la variance de cette loi et nous fournissons des programmes informatiques qui permettent de mettre en oeuvre la méthode.

MOTS CLÉS

MESURE DE SECURITE ROUTIERE - ACCIDENTS DE LA ROUTE - MODELE MULTINOMIAL - MAXIMUM DE VRAISEMBLANCE CONTRAINT - INFORMATION DE FISHER - TEST DU χ^2 .

