REVUE DE STATISTIQUE APPLIQUÉE

A. N'GUESSAN
A. ESSAI
C. LANGRAND

Estimation multidimensionnelle des contrôles et de l'effet moyen d'une mesure de sécurité routière

Revue de statistique appliquée, tome 49, nº 2 (2001), p. 85-102
<http://www.numdam.org/item?id=RSA_2001__49_2_85_0>

© Société française de statistique, 2001, tous droits réservés.

L'accès aux archives de la revue « Revue de statistique appliquée » (http://www.sfds.asso.fr/publicat/rsa.htm) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM
Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Rev. Statistique Appliquée, 2001, **XLIX** (2), 85-102

ESTIMATION MULTIDIMENSIONNELLE DES CONTRÔLES ET DE L'EFFET MOYEN D'UNE MESURE DE SÉCURITÉ ROUTIÈRE

A. N'Guessan*, A. Essai**, C. Langrand*

**Laboratoire de Statistique et Probabilités : CNRS, FRE 2222.*
Université de Lille 1, 59655 Villeneuve d'Ascq Cedex, France.
E-mail : Assi.Nguessan@eudil.fr, Claude.Langrand@univ-lille1.fr
***Laboratoire d'Analyse Numérique et d'Optimisation,*
Université de Lille 1, 59655 Villeneuve d'Ascq Cedex, France.
E-mail : essai@ano.univ-lille1.fr

RÉSUMÉ

On s'intéresse à l'estimation de l'effet d'une mesure de sécurité routière appliquée à S sites expérimentaux, une zone de contrôle étant associée à chacun d'eux. On suppose que chaque site comporte R ($R > 1$) types d'accidents mutuellement exclusifs. On fait l'hypothèse que les données d'accidents relatifs à cette mesure sont mises sous la forme de tableaux avant-après. On construit alors, pour chaque site expérimental, un modèle multinomial prenant en compte l'effet moyen à tous les sites et la tendance moyenne des variations en nombre d'accidents dans la zone de contrôle de chaque site. On propose une méthode pour estimer le vecteur paramètre lié au modèle. Des tableaux de données d'accidents simulés nous permettent de résoudre numériquement les équations du maximum de vraisemblance et d'étudier l'erreur quadratique moyenne.

Mots-clés : *Mesure de sécurité routière, Accident de la route, Méthode avant-après, Modèle multinomial, Maximum de vraisemblance sous contraintes, Erreur quadratique moyenne.*

ABSTRACT

We deal with the estimation of the effect of a road safety measure applied on S target sites with a control area for each site. We suppose that the accidents data recorded at each site are classified in R ($R > 1$) mutually exclusive types. We adopt the before-after technique and assume that at any one target site the total number of accidents recorded is multinomially distributed between the periods and the types of accidents. An estimation method of the unknown vector parameter is suggested. Some simulated accidents data allow us to solve numerically the maximum likelihood's equations and to study the consistence of the estimator via the mean square error.

Keywords : *Road Safety measure, Accident data, Before-after method, Multinomial pattern, Restraints Maximum likelihood, Mean Square Error.*

1. Introduction

Les méthodes statistiques utilisées dans l'estimation avant-après de l'efficacité d'une mesure de sécurité routière dépendent fortement des données disponibles et des objectifs fixés lors de la mise en place de cette mesure (*voir* Lassarre (1985) [8], Hauer (1997) [7]). Quelle que soit la méthode, le principe de base de l'estimation consiste à prédire la valeur moyenne de la variable aléatoire (ou critère d'appréciation) utilisée si la mesure n'avait pas été appliquée. En pratique, on utilise souvent le nombre d'accidents comme variable aléatoire. Pour un seul site par exemple, on compte les nombres d'accidents survenus avant l'application de cette mesure (période t_1) et après celle-ci (période t_2) et il faut pouvoir estimer si la modification du nombre des accidents (passant par exemple de n_{t_1} à n_{t_2}) est significativement différente de celle qu'on aurait observée si la mesure n'avait pas été appliquée ou n'était pas efficace. L'interprétation des résultats doit tenir compte d'un certain nombre de facteurs (évolution du trafic, conditions météorologiques...); si en étudiant d'autres données (par exemple les variations du nombre des accidents sur un site équivalent non soumis à la mesure) on peut voir que le nombre d'accidents est multiplié par un facteur c (coefficient de contrôle, ou d'évolution), l'appréciation de l'efficacité de la mesure doit tenir compte de ce facteur c. Lorsqu'une mesure est appliquée à plusieurs sites en même temps, il n'est pas en général réaliste de considérer, pour des raisons de localisation des sites par exemple, que le même facteur c puisse être utilisé quel que soit le site. C'est pourquoi Tanner (1958) [15] a introduit la notion de zone de contrôle et défini un coefficient de contrôle pour chaque site à l'aide du quotient nombre des accidents « après » sur nombre des accidents « avant » dans la zone de contrôle. Supposons maintenant que, la mesure ayant été appliquée à plusieurs sites, on dispose du nombre total d'accidents pour chaque site expérimental au cours des deux périodes. Puis, de façon rétrospective, on s'intéresse au niveau de chaque site à l'effet de la mesure sur un certain nombre de types d'accidents (accidents mortels, blessés graves, blessés légers, dégâts matériels, etc). Une telle démarche induit une troisième dimension (typologie d'accidents) au modèle et à la statistique de Tanner et impose d'estimer l'effet de la mesure par rapport à des tableaux avant-après de données d'accidents. C'est dans cette optique que N'guessan et Langrand (1993) [13] ont généralisé l'approche de Tanner à des tableaux de données d'accidents avant-après avec contrôle. Ils supposent que la même mesure de sécurité routière est appliquée à plusieurs sites chacun comportant différents types d'accidents mutuellement exclusifs et étant associé à une zone de contrôle spécifique. Ils suggèrent des tests statistiques pour juger si l'effet de la mesure est le même quel que soit le site et quel que soit le type d'accident. On s'intéresse essentiellement dans ce travail à la recherche par simulation d'une solution aux équations du maximum de vraisemblance et à la convergence numérique de cet estimateur. La section 2, est consacrée aux données et à la formulation du problème. On expose dans la section 3 le mécanisme de construction du modèle statistique utilisé pour répartir l'effectif total d'accidents de chaque site à la fois entre les deux périodes et entre les différents types d'accidents. La méthode d'estimation du vecteur paramètre lié à ce modèle et incorporant l'effet moyen de la mesure est exposée à la section 4. La section 5 est exclusivement dédiée aux résultats numériques sur données simulées. On étudie en particulier l'ordre de grandeur de l'erreur quadratique moyenne pour des tableaux avant-après de faibles et grands effectifs d'accidents.

2. Données et formulation du problème

On considère dans ce travail qu'une mesure de sécurité routière (aménagement d'un carrefour, revètement d'un tronçon d'autoroute,...) a été appliquée à S sites. On suppose disposer, sur chacun des S sites, du nombre total d'accidents avant et après l'imposition de la mesure pour des périodes de longueurs convenables (généralement égales et souvent d'au moins un an). La question qui est posée ici est la suivante : Comment estimer l'effet moyen de la mesure relative à l'ensemble des S sites, dans le cas où chacun d'eux comporte $R(R > 1)$ types d'accidents? Les données dont on dispose se trouvent dans le tableau 1 suivant :

TABLEAU 1

	Avant					Après					
	type 1	$\cdots$	type r	$\cdots$	type R	type 1	$\cdots$	type r	$\cdots$	type R	total
site 1	n_{111}	$\cdots$	n_{1r1}		n_{1R1}	n_{211}	$\cdots$	n_{2r1}	$\cdots$	n_{2R1}	n_1
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
site s	n_{11s}	$\cdots$	n_{1rs}		n_{1Rs}	n_{21s}	$\cdots$	n_{2rs}	$\cdots$	n_{2Rs}	n_s
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$		$\vdots$			$\vdots$
site S	n_{11S}	$\cdots$	n_{1rS}		n_{1RS}	n_{21S}	$\cdots$	n_{2rS}	$\cdots$	n_{2RS}	n_S
											N

où n_{1rs} (resp. n_{2rs}) représente le nombre d'accidents du type r du site s avant (resp. après) la mise en place de la mesure, n_s le nombre total d'accidents du site s et $N = \sum_{s=1}^{S} n_s$. Afin de tenir compte de certains facteurs extérieurs (tendances générales des débits de circulation, variations des limitations de vitesse, conditions météorologiques...) on suppose également disposer pendant les mêmes périodes de données sur des zones de contrôle non soumises à la mesure et associées aux sites expérimentaux qui permettent de construire le tableau de coefficients de contrôle suivant :

TABLEAU 2

	type 1	$\cdots$	type r	$\cdots$	type R
site 1	c_{11}	$\cdots$	c_{r1}		c_{R1}
$\vdots$	$\vdots$		$\vdots$		$\vdots$
site s	c_{1s}	$\cdots$	c_{rs}		c_{Rs}
$\vdots$	$\vdots$		$\vdots$		$\vdots$
site S	c_{1S}	$\cdots$	c_{rS}		c_{RS}

c_{rs} étant le coefficient multiplicatif qu'il faut employer dans la zone de contrôle associée au site s et pour le type r d'accidents pour passer du nombre d'accidents observés dans la période « avant » au nombre observé dans la période « après ». Il s'agit alors de combiner les données des tableaux 1 et 2 pour estimer convenablement l'effet moyen de la mesure relativement à tous les sites et types d'accidents et de proposer ainsi une réponse à la question précédente.

3. Modèle statistique

Considérons la période avant (resp. après) l'implémentation de la mesure et notons X_{1rs} (resp. X_{2rs}) la variable aléatoire (v.a.) donnant le nombre d'accidents du type r sur le site expérimental s. Pour modéliser l'effet moyen de la mesure de sécurité, on suppose que le nombre moyen $E(X_{2rs})$ d'accidents du type r au site s qu'on pourrait s'attendre à observer dans la période « après » si ce site avait un comportement semblable à celui de la zone de contrôle associée est :

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

où $E(X_{1rs})$ est l'espérance mathématique de la v.a. X_{1rs} et θ^* le rapport des espérances mathématiques du nombre total d'accidents observés dans la période « après », et du nombre d'accidents qu'on pourrait s'attendre à observer dans cette période si chaque site se comportait comme sa zone de contrôle :

$$\theta^* = \frac{\displaystyle\sum_{s=1}^{S}\sum_{r=1}^{R} E(X_{2rs})}{\displaystyle\sum_{s=1}^{S}\sum_{r=1}^{R} c_{rs} E(X_{1rs})}$$

Si le site s a un comportement semblable à celui de la zone de contrôle associée et si l'on estime $E(X_{1rs})$ par n_{1rs}, valeur observée de X_{1rs} lue dans le tableau 1, alors l'expression (1) devient

$$\theta^* c_{rs} n_{1rs}. \tag{2}$$

Sous ces mêmes hypothèses et en considèrant les périodes « avant » et « après », le nombre total d'accidents espérés au site s est estimé par :

$$n_s^* = n_{11s} + n_{12s} + \cdots + n_{1Rs} + \theta^* c_{1s} n_{11s} + \theta^* c_{2s} n_{12s} + \cdots + \theta^* c_{Rs} n_{1Rs}$$

Cette dernière égalité implique que les proportions d'accidents espérés par type au site s au cours des deux périodes (avant et après) sont estimées par

$$\frac{n_{11s}}{n_s^*}, \frac{n_{12s}}{n_s^*}, \ldots, \frac{n_{1Rs}}{n_s^*}, \frac{\theta^* c_{1s} n_{11s}}{n_s^*}, \frac{\theta^* c_{2s} n_{12s}}{n_s^*}, \ldots, \frac{\theta^* c_{Rs} n_{1Rs}}{n_s^*} \tag{3}$$

soit encore par

$$\frac{p_{1s}^*}{1+\theta^* c_s^*}, \frac{p_{2s}^*}{1+\theta^* c_s^*}, \cdots, \frac{p_{Rs}^*}{1+\theta^* c_s^*}, \frac{\theta^* c_{1s} p_{1s}^*}{1+\theta^* c_s^*}, \frac{\theta^* c_{2s} p_{2s}^*}{1+\theta^* c_s^*}, \cdots, \frac{\theta^* c_{Rs} p_{Rs}^*}{1+\theta^* c_s^*}$$

en posant

$$p_{rs}^* = n_{1rs}/(\sum_{j=1}^{R} n_{1js}), \quad c_s^* = \sum_{r=1}^{R} c_{rs} p_{rs}^*.$$

p_{rs}^* est une estimation de la probabilité pour qu'un accident qui se produit au site s avant la mise en place de la mesure soit du type r, avec $\sum_{r=1}^{R} p_{rs}^* = 1, s = 1, 2, \cdots, S$. Pour s fixé, c_s^* est la somme des valeurs c_{rs} du tableau 2 pondérée par les p_{rs}^* $(r = 1, 2, \cdots, R)$. c_s^* représente une estimation du coefficient de contrôle moyen associé au site s si ce site se comportait comme sa zone de contrôle associée.

On pose alors

$$q_{1rs}^* = \frac{p_{rs}^*}{1+\theta^* c_s^*}, \quad q_{2rs}^* = \frac{\theta^* c_{rs} p_{rs}^*}{1+\theta^* c_s^*}, \quad r = 1, 2, \cdots, R. \tag{4}$$

Par construction, les q_{trs}^* ($t = 1$ ou 2) représentent la répartition des n_s^* accidents en $2R$ catégories mutuellement exclusives entre les périodes « avant » et « après » si le site s se comporte comme sa zone de contrôle et si θ^* est l'effet moyen de la mesure; on a

$$\sum_{r=1}^{R} \sum_{t=1}^{2} q_{trs}^* = 1. \tag{5}$$

Les considérations précédentes nous conduisent à proposer la modélisation suivante : on considère le vecteur aléatoire ($\overrightarrow{v}.a.$)

$$X^{(s)} = (X_{11s}, X_{12s}, \cdots, X_{1Rs}, X_{21s}, X_{22s}, \cdots, X_{2Rs}) \qquad s = 1, 2, \ldots, S$$

prenant simultanément en compte les v.a. $X_{11s}, X_{12s}, \cdots, X_{1Rs}, X_{21s}, X_{22s}, \cdots, X_{2Rs}$ nombres d'accidents des R types avant et après la mesure au site expérimental s, de valeur observée $n^{(s)} = (n_{11s}, n_{12s}, \cdots, n_{1Rs}, n_{21s}, n_{22s}, \cdots, n_{2Rs})$ (ligne s du tableau 1). Pour s fixé, on s'intéresse à la loi de probabilité du $\overrightarrow{v}.a.$ $X^{(s)}$ étant donnés les tableaux 1 et 2. En s'inspirant des relations (4) et (5), on fait alors l'hypothèse que pour s fixé le $\overrightarrow{v}.a.$ $X^{(s)}$ est de loi multinomiale :

$$Proba[X^{(s)} = n^{(s)}] = \left(\frac{n_s!}{\prod_{t=1}^{2} \prod_{r=1}^{R} n_{trs}!} \right) \prod_{t=1}^{2} \prod_{r=1}^{R} q_{trs}^{n_{trs}} \tag{6}$$

où

$$q_{1rs} = \frac{p_{rs}}{1+\theta c_s}, \quad q_{2rs} = \frac{\theta c_{rs} p_{rs}}{1+\theta c_s}, \quad r = 1, 2, \cdots, R \tag{7}$$

avec θ ($\theta > 0$) paramètre effet moyen de la mesure de sécurité routière relative à l'ensemble des sites et des types d'accidents, les p_{rs} ($0 < p_{rs} < 1$) étant des paramètres auxiliaires liés aux zones de contrôle et aux sites expérimentaux tels que

$$\sum_{r=1}^{R} p_{rs} = 1, \qquad c_s = \sum_{r=1}^{R} c_{rs} p_{rs}. \tag{8}$$

Si on note C_s la v.a. prenant les valeurs $c_{1s}, c_{2s}, \cdots, c_{Rs}$ de la ligne s du tableau 2 avec les probabilités $p_{1s}, p_{2s}, \cdots, p_{Rs}$ alors $c_s = \sum_{r=1}^{R} c_{rs} p_{rs}$ est l'espérance mathématique de la v.a. C_s et est appelé coefficient moyen de contrôle associé au site expérimental s. Les paramètres auxiliaires $(p_{rs})_{1 \leqslant r \leqslant R}$ sont interprétés comme loi de probabilité de C_s. Les $c_1, c_2, \cdots, c_S$ seront entièrement déterminés une fois les lois de probabilité $((p_{rs})_{1 \leqslant r \leqslant R})_{1 \leqslant s \leqslant S}$ connues.

4. Estimation simultanée des coefficients moyens de contrôle et de l'effet moyen de la mesure

Par construction du modèle, on peut distinguer deux types de paramètres parmi ceux associés aux S sites expérimentaux et R types d'accidents : le paramètre d'intérêt θ et l'ensemble des paramètres auxiliaires $P_s = (p_{1s}, p_{2s}, \cdots, p_{Rs})^T$, $s = 1, 2, \cdots, S$. Dans la suite de ce travail, on s'intéresse à l'estimation simultanée de ces paramètres et on pose alors $\Theta = (\theta, P_1^T, P_2^T, \ldots, P_S^T)^T$ l'élément de $\mathbb{R}^{1+SR}$ ainsi défini (en réalité on est dans $\mathbb{R}^{1+S(R-1)}$ puisque $\sum_{r=1}^{R} p_{rs} = 1$). On fait alors l'hypothèse que les $\overrightarrow{v}.a.$ $X^{(1)}, X^{(2)}, \ldots, X^{(S)}$ sont indépendants et suivent chacun une distribution multinomiale dont la forme est donnée par (6) et (7). Etant donné les tableaux 1 et 2, la fonction de vraisemblance associée à la suite des $\overrightarrow{v}.a.$ $X^{(1)}, X^{(2)}, \ldots, X^{(S)}$ est :

$$L(X^{(1)}, X^{(2)}, \ldots, X^{(S)}; \Theta) = \prod_{s=1}^{S} l_s(X^{(s)}; \Theta) \tag{9}$$

où

$$l_s(X^{(s)}; \Theta) = \frac{n_s!}{\prod_{t=1}^{2} \prod_{m=1}^{R} n_{tms}!} \prod_{r=1}^{R} \frac{p_{rs}^{n_{1rs}} (\theta c_{rs})^{n_{2rs}}}{(1 + \theta \sum_{j=1}^{R} c_{js} p_{js})^{n_{1rs}}} \tag{10}$$

avec $n_{.rs} = n_{1rs} + n_{2rs}$. La log-vraisemblance est donnée à une constante additive près par :

$$\mathcal{L}(X^{(1)}, X^{(2)}, \ldots, X^{(S)}; \Theta)$$
$$= \sum_{s=1}^{S} \sum_{r=1}^{R} \{n_{.rs} log_e(p_{rs}) + n_{2rs} log_e(\theta) - n_{.rs} \log_e(1 + \theta \sum_{m=1}^{R} c_{ms} p_{ms})\}. \quad (11)$$

En dérivant la fonction $\Theta \longmapsto \mathcal{L}(X^{(1)}, X^{(2)}, \ldots, X^{(S)}; \Theta)$ par rapport aux composantes de Θ et en annulant les dérivées partielles (tout en tenant compte de ce que $\sum_{r=1}^{R} p_{rs} = 1$, $s = 1, 2, \ldots, S$) on montre que l'estimateur, s'il existe, $\hat{\Theta}$ de Θ est solution du système d'équations non linéaires sous contraintes :

$$\begin{cases} \sum_{s=1}^{S} \sum_{r=1}^{R} \dfrac{n_{2rs} - \hat{\theta}\hat{c}_s n_{1rs}}{1 + \hat{\theta}\hat{c}_s} = 0; \\ n_{.rs} - \dfrac{n_s \hat{p}_{rs}(\hat{\theta} c_{rs} + 1)}{1 + \hat{\theta}\hat{c}_s} = 0, \\ (s = 1, 2, \cdots, S; \quad r = 1, 2, \cdots, R); \\ \hat{\theta} > 0; \quad 0 < \hat{p}_{rs} < 1; \\ \hat{c}_s = \sum_{r=1}^{R} c_{rs} \hat{p}_{rs}, \quad s = 1, 2, \cdots, S; \\ \sum_{r=1}^{R} \hat{p}_{rs} = 1, \quad s = 1, 2, \cdots, S. \end{cases} \quad (12)$$

$\hat{\Theta}$ est l'estimation du maximum de vraisemblance sous contraintes linéaires et de bornes (*voir* par exemple, [1], [2], [5], [11]). On aborde ici uniquement les aspects numériques de l'existence de $\hat{\Theta}$ par résolution et simulation du système (12).

5. Résultats numériques de simulation

5.1. Principe de simulation des données avant-après avec contrôle

L'étude numérique des équations de vraisemblance (12) et donc de l'existence numérique de $\hat{\Theta}$ nécessite la connaissance des tableaux de données (tableaux 1 et 2). Cette résolution numérique comporte deux grandes phases. Une première phase de simulation du tableau des coefficients de contrôle c_{rs} (tableau 2) et du tableau des données avant-après (tableau 1) par programmation en FORTRAN du modèle multinomial construit à la section 3. Puis suit la phase de résolution proprement dite (recherche numérique de $\hat{\Theta}$) par utilisation des algorithmes (E04ZCF et E04UCF) d'optimisation du logiciel NAG [11] faisant appel aux tableaux 1 et 2 précédemment simulés.

Etant donné donc S (le nombre de sites expérimentaux) et R (le nombre de types d'accidents), on suppose que les coefficients c_{rs} $(s = 1, 2, \cdots, S;\ r = 1, 2, \cdots, R)$ proviennent d'une $v.a.$ de loi $U_{]\frac{1}{2},\frac{5}{2}[}$ (loi uniforme entre $\frac{1}{2}$ et $\frac{5}{2}$). Pour l'obtention du tableau 1, on pose $\theta^0 = \frac{1}{2}$ et on fait l'hypothèse que les composantes du vecteur $P_s^0 = (p_{1s}^0, p_{2s}^0, \cdots, p_{Rs}^0)^T$ vraie valeur du vecteur P_s proviennent d'une $v.a.$ de loi $U_{[\alpha,1-\alpha]}$ (avec $\alpha = 10^{-5}$ de manière à permettre l'écriture de la log-vraisemblance) et sont normalisées de manière que $\sum_{r=1}^{R} p_{rs}^0 = 1;\ s = 1, 2, \cdots, S$. En utilisant ces valeurs, on détermine alors les vrais coefficients moyens de contrôle

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

puis ensuite les vraies probabilités

$$q_{1rs}^0 = \frac{p_{rs}^0}{1 + \theta^0 c_s^0}, \quad q_{2rs}^0 = \frac{c_{rs}\theta^0 p_{rs}^0}{1 + \theta^0 c_s^0}, \quad r = 1, 2, \cdots, R$$

associée à la distribution multinomiale relative au $\overrightarrow{v}.a.$ $X^{(s)}$. Enfin pour s $(s = 1, 2, \cdots, S)$ fixé, on génère l'effectif total n_s du site s et on répartit de façon aléatoire n_s entre les périodes « avant » et « après » par utilisation des probabilités q_{1rs}^0 et q_{2rs}^0. On obtient ainsi les valeurs observées n_{trs} du tableau 1 avec $\sum_{r=1}^{R}\sum_{t=1}^{2} n_{trs} = n_s$, $s = 1, 2, \cdots, S$. Plusieurs tableaux de différentes tailles et de différents effectifs ont été générés en fixant les valeurs de S et R et en choisissant n_s dans différentes tranches d'accidents. Les résultats numériques (*voir* paragraphe 5.3) sont obtenus pour n_s appartenant aux tranches d'accidents $T_1, \ldots, T_{10}$ suivantes :

$$\{]0, 30[,\ ]30, 50[,\ ]50, 300[,\ ]300, 1000[,\ ]1000, 3000[,\ ]3000, 5000[,\ ]5000, 7000[,\ ]7000, 9000[,\ ]9000, 11000[,\ ]11000, 13000[\}\ .$$

5.2. Estimation et convergence de $\hat{\Theta}$

Le vecteur paramètre à estimer $\Theta \in \mathbb{R}^{1+SR}$ a pour vraie valeur

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

avec θ^0 fixé à 0.5 et $P_1^0, P_2^0, \cdots, P_S^0$ des vecteurs de dimension R. L'objectif principal de la simulation est de rechercher d'abord une solution $\hat{\Theta}$ aux équations de vraisemblance et d'étudier la convergence de $\hat{\Theta}$ vers Θ^0 par l'intermédiaire de

l'analyse de l'erreur quadratique moyenne (EQM) qui mesure l'adéquation entre $\hat{\Theta}$ et Θ^o et est définie par

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

On utilise ensuite les composantes $\hat{P}_1, \hat{P}_2, \cdots, \hat{P}_S$ de $\hat{\Theta}$ pour estimer (*voir* paragraphe 5.3) et étudier la convergence des $\hat{c_s}$ vers c_s^0 en analysant l'EMC (Ecart Moyen de Contrôle) et l'EQMC (Ecart Quadratique Moyen de Contrôle) définis ci-après :

$$EMC(\hat{c}_s, c_s^0) = \frac{1}{S} \sum_{s=1}^{S} |\hat{c}_s - c_s^0|; \quad EQMC(\hat{c}_s, c_s^0) = \frac{1}{S} \sum_{s=1}^{S} (\hat{c}_s - c_s^0)^2.$$

Enfin on se focalise sur le comportement de la première composante (paragraphe 5.4) $\hat{\theta}$ de $\hat{\Theta}$ qui est l'estimateur de l'effet moyen de la mesure de sécurité dont la vraie valeur est $\theta^0 = 0.5$.

TABLEAU 3
Valeurs de $\hat{\Theta}$ pour $S = 4$ et $R = 2$.

Θ^0	$\hat{\Theta}$						
	$T=1$	$T=2$	$T=4$	$T=6$	$T=8$	$T=9$	$T=10$
5.0000e-1	5.3094e-1	5.0439e-1	4.6923e-1	4.7080e-1	4.7807e-1	4.8208e-1	4.8236e-1
9.3099e-1	8.0479e-1	8.6918e-1	9.2232e-1	9.2797e-1	9.3141e-1	9.3126e-1	9.3174e-1
6.9010e-2	1.9521e-1	1.3082e-1	7.7680e-2	7.2030e-2	6.8590e-2	6.8740e-2	6.8260e-2
6.8019e-1	9.9999e-1	5.0774e-1	6.8164e-1	6.9053e-1	6.9104e-1	6.9561e-1	6.9461e-1
3.1981e-1	1.0000e-5	4.9226e-1	3.1836e-1	3.0947e-1	3.0896e-1	3.0439e-1	3.0539e-1
4.1127e-1	4.6512e-1	3.8411e-1	3.9899e-1	4.1712e-1	4.1677e-1	4.0974e-1	4.1172e-1
5.8873e-1	5.3488e-1	6.1589e-1	6.0101e-1	5.8288e-1	5.8323e-1	5.9026e-1	5.8828e-1
6.6746e-1	8.8382e-1	6.6582e-1	6.6267e-1	6.6807e-1	6.6211e-1	6.7331e-1	6.6320e-1
3.3254e-1	1.1618e-1	3.3418e-1	3.3733e-1	3.3193e-1	3.3789e-1	3.2669e-1	3.3680e-1
EQM	3.7420e-2	7.6237e-3	1.6095e-4	1.2821e-4	9.2758e-5	9.6645e-5	8.5027e-5

TABLEAU 4
Valeurs de $\hat{\Theta}$ pour $S = 4$ et $R = 3$.

Θ^0	$\hat{\Theta}$						
	$T = 1$	$T = 2$	$T = 4$	$T = 6$	$T = 8$	$T = 9$	$T = 10$
5.0000e-1	4.1035e-1	4.5889e-1	5.0496e-1	4.8721e-1	4.9125e-1	4.8978e-1	4.9024e-1
2.6775e-1	1.5461e-1	2.7960e-1	2.7525e-1	2.7457e-1	2.7090e-1	2.7097e-1	2.7167e-1
1.9850e-2	1.0000e-5	1.0000e-5	1.6330e-2	1.7380e-2	1.9180e-2	1.8730e-2	1.8800e-2
7.1241e-1	8.4538e-1	7.2039e-1	7.0843e-1	7.0805e-1	7.0991e-1	7.1030e-1	7.0953e-1
2.1941e-1	3.9929e-1	1.9002e-1	2.2650e-1	2.2371e-1	2.2338e-1	2.2374e-1	2.1945e-1
3.2103e-1	6.0070e-1	5.1472e-1	2.9518e-1	3.0859e-1	3.2144e-1	3.2393e-1	3.2357e-1
4.5956e-1	1.0000e-5	2.9526e-1	4.7832e-1	4.6770e-1	4.5518e-1	4.5233e-1	4.5698e-1
3.1570e-1	6.0772e-1	3.5021e-1	3.3186e-1	3.1235e-1	3.1378e-1	3.1286e-1	3.1667e-1
1.5729e-1	1.0000e-5	1.5312e-1	1.5034e-1	1.6176e-1	1.5443e-1	1.5581e-1	1.5422e-1
5.2702e-1	3.9227e-1	4.9667e-1	5.1780e-1	5.2589e-1	5.3180e-1	5.3133e-1	5.2911e-1
3.9505e-1	4.8206e-1	4.3417e-1	3.7513e-1	3.8995e-1	3.9367e-1	3.8916e-1	3.9118e-1
4.8130e-1	4.2259e-1	4.2579e-1	4.9792e-1	4.9061e-1	4.8124e-1	4.9242e-1	4.8983e-1
1.2365e-1	9.5350e-2	1.4004e-1	1.2695e-1	1.1945e-1	1.2509e-1	1.1843e-1	1.1900e-1
EQM	3.8512e-2	5.7439e-3	1.7372e-4	4.9044e-5	1.2835e-5	3.1863e-5	1.9789e-5

TABLEAU 5
Valeurs de EQM($\hat{\Theta}, \Theta^0$) pour les 10 tranches d'accidents.

	$S = 5$	
T	$R = 2$	$R = 10$
1	2.6850E-2	1.8480E-2
2	5.3214E-3	1.8482E-3
3	8.9237E-4	4.8812E-4
4	7.8695E-4	1.9116E-4
5	2.6904E-4	3.2655E-5
6	5.8576E-5	2.4412E-5
7	3.9579E-5	1.4034E-5
8	1.5870E-5	9.5464E-6
9	1.6832E-5	9.4963E-6
10	8.2177E-6	8.2518E-6

	$S = 20$	
T	$R = 2$	$R = 10$
1	3.8793E-2	1.6054E-2
2	6.0657E-3	2.3454E-3
3	2.4164E-3	8.5019E-4
4	4.5135E-4	1.5882E-4
5	7.2394E-5	4.0908E-5
6	3.2943E-5	2.2612E-5
7	3.0454E-5	1.8778E-5
8	4.8501E-5	1.1726E-5
9	2.3679E-5	9.3814E-6
10	2.1514E-5	8.3590E-6

TABLEAU 6
Valeurs de EQM($\hat{\Theta}, \Theta^0$) pour $n_s \in]0, 30[$.

S/R	2	3	4	5	6	7	8	9	10
1	2.347e-2	2.726e-2	1.270e-1	1.930e-3	1.342e-2	5.890e-3	1.772e-2	4.711e-2	6.632e-3
2	1.282e-2	9.293e-3	5.170e-2	2.160e-3	9.044e-3	1.474e-2	2.196e-3	4.206e-2	2.440e-2
3	1.145e-2	3.002e-2	2.444e-2	5.957e-3	2.209e-2	1.155e-2	8.153e-3	7.182e-3	1.917e-2
4	3.742e-2	3.851e-2	4.879e-2	2.322e-2	8.693e-3	3.092e-2	1.228e-2	1.290e-2	6.966e-3
5	2.685e-2	4.082e-2	1.217e-2	1.421e-2	1.409e-2	1.994e-2	1.094e-2	6.691e-3	1.848e-2
6	6.811e-2	5.576e-2	1.089e-2	2.455e-2	1.816e-2	1.230e-2	1.825e-2	2.182e-2	1.055e-2
7	5.306e-2	1.740e-2	1.683e-2	3.150e-2	1.133e-2	2.406e-2	2.084e-2	1.100e-2	1.710e-2
8	4.241e-2	1.865e-2	1.564e-2	6.814e-3	2.715e-2	1.805e-2	7.295e-3	1.451e-2	1.692e-2
9	5.680e-2	4.341e-2	3.781e-2	2.049e-2	1.697e-2	8.426e-3	2.703e-2	1.005e-2	7.622e-3
10	1.613e-2	2.726e-2	1.377e-2	3.513e-2	1.389e-2	1.589e-2	1.548e-2	4.664e-3	1.084e-2

TABLEAU 7
Valeurs de EQM($\hat{\Theta}, \Theta^0$) pour $n_s \in]1000, 3000[$.

S/R	2	3	4	5	6	7	8	9	10
1	9.171e-4	1.600e-3	2.722e-4	4.290e-4	4.496e-4	3.142e-4	1.083e-4	9.915e-5	2.029e-4
2	6.568e-5	1.309e-4	6.783e-5	1.333e-4	5.466e-5	5.733e-5	1.101e-4	6.047e-5	3.740e-5
3	1.778e-5	2.159e-4	6.407e-5	8.938e-5	1.053e-4	7.605e-5	7.212e-5	6.265e-5	7.335e-5
4	3.430e-4	6.474e-5	1.279e-4	8.235e-5	6.360e-5	8.585e-5	1.220e-4	7.270e-5	4.409e-5
5	2.690e-4	7.882e-5	9.696e-5	1.288e-4	5.424e-5	5.920e-5	4.764e-5	5.065e-5	3.266e-5
6	3.238e-5	1.689e-4	4.080e-5	9.685e-5	4.810e-5	3.265e-5	7.108e-5	4.138e-5	6.603e-5
7	1.575e-4	2.163e-4	9.277e-5	1.103e-4	8.133e-5	1.039e-4	4.449e-5	5.136e-5	5.366e-5
8	2.852e-4	1.147e-4	1.109e-4	7.999e-5	1.462e-4	5.997e-5	5.060e-5	4.875e-5	5.415e-5
9	1.538e-4	9.020e-5	7.464e-5	1.086e-4	9.135e-5	4.888e-5	4.992e-5	4.379e-5	4.923e-5
10	9.104e-5	1.236e-4	9.078e-5	8.324e-5	6.447e-5	6.406e-5	6.046e-5	7.232e-5	5.495e-5

5.3. Estimation et convergence des contrôles moyens

TABLEAU 8
Estimation des contrôles moyens pour $S = 4$ et $R = 2$.

c_s^0	$\hat{c}_s$						
	$T = 1$	$T = 2$	$T = 4$	$T = 6$	$T = 8$	$T = 9$	$T = 10$
1.3522	1.5012	1.4252	1.3624	1.3558	1.3517	1.3519	1.3513
1.9271	2.4512	1.6444	1.9294	1.9440	1.9449	1.9523	1.9507
1.3618	1.3865	1.3493	1.3562	1.3645	1.3643	1.3611	1.3620
1.0523	8.9612e-1	1.0535	1.0558	1.0519	1.0562	1.0481	1.0554
EMC	1.3540e-1	5.5232e-2	2.6107e-3	5.6889e-3	5.9208e-3	5.0079e-3	6.5096e-3
EQMC	8.0487e-2	.1341e-2	3.8520e-5	7.6836e-5	8.4519e-5	1.6431e-4	1.4231e-4

TABLEAU 9
Estimation des contrôles moyens pour $S = 4$ et $R = 3$.

c_s^0	$\hat{c}_s$						
	$T = 1$	$T = 2$	$T = 4$	$T = 6$	$T = 8$	$T = 9$	$T = 10$
1.5513	1.5759	1.5308	1.5458	1.5469	1.5497	1.5493	1.5491
1.3377	1.4667	1.2303	1.3561	1.3477	1.3427	1.3422	1.3369
1.0763	1.2720	1.1061	1.0880	1.0753	1.0733	1.0731	1.0758
1.3383	1.2550	1.3219	1.3550	1.3389	1.3403	1.3388	1.3377
EMC	6.6486e-2	2.8671e-2	1.0301e-2	1.2852e-3	5.5014e-4	9.7155e-5	1.0570e-3
EQMC	1.5611e-2	3.2824e-3	1.9652e-4	3.0317e-5	1.0155e-5	8.8405e-6	1.6337e-6

TABLEAU 10
Valeurs de EMC pour $n_s \in]0, 30[$.

S/R	2	3	4	5	6	7	8	9	10
1	1.147e-2	2.550e-2	4.328e-1	8.259e-2	1.531e-2	1.073e-1	3.668e-2	7.383e-1	9.564e-2
2	1.346e-1	1.103e-1	2.688e-1	3.929e-2	4.097e-2	8.928e-2	5.126e-2	3.249e-1	1.708e-1
3	3.564e-2	1.556e-1	3.277e-2	5.591e-2	1.907e-1	4.254e-2	4.983e-3	1.361e-3	2.470e-1
4	1.354e-1	6.649e-2	1.018e-1	1.736e-1	4.203e-2	9.166e-3	4.907e-2	3.129e-2	1.278e-2
5	4.316e-3	2.296e-1	3.215e-2	9.013e-2	6.403e-2	7.545e-2	8.186e-3	1.508e-2	5.714e-2
6	1.862e-2	5.818e-2	1.594e-3	1.504e-1	5.499e-2	7.606e-2	1.397e-1	4.484e-2	3.997e-2
7	1.122e-1	1.173e-2	6.156e-2	7.971e-2	7.701e-3	7.682e-2	9.809e-2	1.751e-2	8.651e-2
8	1.012e-1	2.059e-2	9.182e-2	3.437e-2	5.015e-2	1.209e-2	7.337e-2	7.725e-2	7.207e-2
9	1.034e-1	6.315e-3	1.821e-1	1.196e-1	2.248e-2	4.765e-2	1.439e-1	1.504e-1	2.233e-2
10	1.863e-2	4.172e-2	5.736e-2	1.137e-2	2.724e-3	4.354e-2	1.066e-1	1.338e-2	3.226e-2

TABLEAU 11
Valeurs de EMC pour $n_s \in]1000, 3000[$.

S/R	2	3	4	5	6	7	8	9	10
1	3.643e-4	2.856e-2	2.818e-2	1.784e-2	2.720e-2	9.270e-3	2.117e-3	2.466e-2	6.808e-3
2	5.777e-3	1.629e-2	3.887e-3	8.571e-3	3.637e-3	1.989e-2	7.520e-4	1.614e-2	2.839e-3
3	2.108e-3	2.901e-3	2.383e-3	4.995e-4	6.700e-3	6.573e-4	1.101e-2	6.698e-3	3.923e-3
4	8.725e-3	2.323e-3	1.723e-3	9.243e-4	6.826e-3	4.683e-4	4.377e-3	4.849e-4	9.939e-3
5	6.571e-3	5.984e-3	4.482e-3	4.655e-3	1.173e-3	2.954e-3	3.996e-3	4.507e-4	2.726e-3
6	2.167e-3	7.422e-3	1.556e-3	8.796e-3	3.866e-4	6.370e-3	2.710e-3	5.346e-3	5.013e-3
7	1.674e-3	5.131e-4	2.960e-3	1.236e-3	1.381e-2	9.190e-3	2.399e-3	6.367e-4	8.007e-3
8	1.950e-3	9.858e-3	5.267e-4	3.522e-3	2.361e-3	1.026e-3	4.723e-3	4.985e-3	3.846e-3
9	9.869e-4	5.584e-3	8.248e-3	9.781e-4	3.626e-3	5.889e-3	2.667e-3	8.648e-4	2.151e-3
10	1.920e-3	2.858e-3	7.850e-3	4.333e-3	6.714e-5	9.285e-4	3.228e-4	1.814e-3	1.598e-3

TABLEAU 12
Valeurs de EQMC pour $n_s \in]0, 30[$.

S/R	2	3	4	5	6	7	8	9	10
1	1.315e-4	6.501e-4	1.873e-1	6.821e-3	2.343e-4	1.151e-2	1.345e-3	5.450e-1	9.147e-3
2	2.031e-2	1.282e-2	1.550e-1	7.661e-3	1.705e-3	1.129e-2	1.102e-2	2.846e-1	3.104e-2
3	2.871e-2	6.841e-2	3.149e-2	1.109e-2	8.077e-2	2.057e-2	3.156e-2	8.861e-3	1.565e-1
4	8.049e-2	1.561e-2	4.747e-2	6.469e-2	1.908e-2	1.753e-2	3.084e-2	7.883e-3	4.494e-3
5	2.505e-3	1.147e-1	1.525e-2	2.960e-2	9.892e-3	1.207e-1	1.860e-3	1.281e-2	4.845e-2
6	1.869e-2	4.478e-2	1.805e-3	7.734e-2	5.667e-2	7.960e-2	3.936e-2	2.602e-2	8.418e-2
7	6.582e-2	1.827e-2	2.495e-2	8.567e-2	3.131e-2	3.218e-2	3.634e-2	7.098e-3	4.656e-2
8	3.227e-2	2.201e-3	3.174e-2	1.206e-2	7.312e-2	5.255e-3	1.165e-2	8.066e-2	1.221e-2
9	1.066e-1	5.492e-2	1.472e-1	6.402e-2	1.482e-2	2.739e-2	1.250e-1	4.800e-2	2.331e-2
10	2.209e-2	1.475e-2	1.819e-2	9.020e-3	2.016e-2	3.946e-2	6.558e-2	7.523e-3	2.303e-2

TABLEAU 13
Valeurs de EQMC pour $n_s \in]1000, 3000[$.

S/R	2	3	4	5	6	7	8	9	10
1	1.327e-7	8.156e-4	7.941e-4	3.182e-4	7.397e-4	8.593e-5	4.482e-6	6.079e-4	4.635e-5
2	4.801e-5	4.124e-4	5.512e-5	4.185e-4	1.727e-5	4.051e-4	3.988e-6	2.767e-4	1.792e-4
3	1.348e-5	3.752e-4	3.936e-5	5.443e-5	1.173e-4	1.691e-4	3.889e-4	1.165e-4	2.089e-4
4	6.198e-4	2.019e-5	9.667e-5	6.508e-5	5.416e-5	2.711e-5	6.954e-5	6.220e-5	1.009e-4
5	1.043e-4	7.047e-5	8.137e-5	2.580e-4	1.192e-4	1.006e-4	1.674e-4	1.426e-4	8.175e-5
6	3.132e-5	2.407e-4	6.797e-5	2.159e-4	1.741e-4	1.196e-4	1.280e-4	2.991e-4	2.411e-4
7	2.096e-4	2.914e-4	1.466e-4	1.531e-4	3.350e-4	4.119e-4	1.263e-4	6.231e-5	1.679e-4
8	4.351e-4	2.342e-4	2.074e-4	2.066e-4	3.829e-4	1.925e-4	1.972e-4	2.418e-4	9.696e-5
9	3.502e-4	2.687e-4	2.367e-4	2.632e-4	3.385e-4	1.640e-4	9.895e-5	6.178e-5	7.156e-5
10	8.210e-5	2.082e-4	1.573e-4	1.601e-4	1.123e-4	1.018e-4	8.255e-5	2.191e-4	7.166e-5

5.4. Estimation et convergence de l'effet commun de la mesure

TABLEAU 14
Valeur de $|\hat{\theta} - \frac{1}{2}|$ *pour* $S = 4$.

R	$T = 1$	$T = 2$	$T = 4$	$T = 6$	$T = 8$	$T = 9$	$T = 10$
2	3.094e-2	4.394e-3	3.077e-2	2.920e-2	2.193e-2	1.792e-2	1.764e-2
3	8.965e-2	4.111e-2	4.961e-3	1.279e-2	8.748e-3	1.022e-2	9.765e-3
4	1.440e-1	8.236e-4	2.659e-3	2.079e-2	1.829e-2	2.048e-2	1.860e-2
5	1.990e-1	4.665e-2	3.570e-2	7.751e-3	2.235e-5	2.907e-3	5.794e-3
6	2.237e-1	3.798e-2	9.156e-3	9.118e-3	1.311e-2	1.763e-2	1.724e-2
7	2.466e-1	7.232e-2	2.417e-2	1.413e-2	1.754e-2	1.130e-2	1.049e-2
8	3.006e-2	4.634e-2	1.461e-2	2.392e-2	1.638e-2	1.991e-2	1.717e-2
9	2.562e-2	1.688e-1	4.343e-4	4.114e-3	1.388e-2	1.140e-2	1.118e-2
10	1.771e-1	5.026e-2	3.456e-2	1.711e-2	1.761e-2	2.150e-2	1.648e-2

TABLEAU 15
Valeur de $|\hat{\theta} - \frac{1}{2}|$ *pour* $R = 4$.

S	$T=1$	$T=2$	$T=4$	$T=6$	$T=8$	$T=9$	$T=10$
1	5.000e-1	5.686e-2	1.236e-2	1.439e-2	2.832e-2	3.703e-2	3.562e-2
2	3.388e-1	7.728e-2	2.211e-4	8.970e-3	3.814e-3	8.832e-3	9.407e-3
3	8.473e-2	3.753e-2	3.694e-2	2.320e-2	2.955e-2	2.993e-2	2.886e-2
4	1.440e-1	8.236e-4	2.659e-3	2.079e-2	1.829e-2	2.048e-2	1.860e-2
5	9.641e-2	4.448e-2	2.952e-2	1.520e-3	2.829e-3	4.496e-3	3.655e-3
6	1.861e-2	1.474e-2	1.251e-2	1.785e-3	4.581e-3	4.886e-3	4.831e-3
7	1.935e-1	2.695e-2	5.266e-2	2.582e-2	1.693e-2	1.406e-2	1.251e-2
8	1.744e-1	8.770e-2	1.661e-3	4.951e-3	1.905e-3	1.468e-3	6.916e-4
9	4.464e-2	3.109e-2	4.067e-3	2.074e-3	8.238e-4	5.313e-3	2.534e-3
10	7.171e-2	2.162e-2	9.629e-3	9.504e-4	1.859e-3	4.166e-3	5.709e-3

5.5. Analyse générale des résultats

Le tableau 3 (resp. 4) donne les différentes valeurs du vecteur $\hat{\Theta}$ solution des équations de vraisemblance pour $S = 4$ et $R = 2$ (resp. $R = 3$) fixés. La première colonne de chacun des deux tableaux donne la vraie valeur Θ^o du paramètre. Les colonnes suivantes sont les solutions obtenues pour n_s appartenant à 7 tranches (colonne T : $T = 1, \ldots, T = 10$) d'accidents. Ces solutions appartiennent à $\mathbb{R}^9$ pour le tableau 3 (resp. $\mathbb{R}^{13}$ pour le tableau 4). On constate une décroissance de l'ordre de grandeur de $EQM(\widehat{\Theta}, \Theta^o)$ lorsque les effectifs n_s passent de $]0, 30[$ à $]11000, 13000[$. Pour les tableaux 3 et 4 $EQM(\widehat{\Theta}, \Theta^o)$ passe de $O(10^{-2})$ à $O(10^{-5})$. Le tableau 5 complète cette première étude de $EQM(\hat{\Theta}, \Theta^o)$ pour l'ensemble des 10 tranches (colonne : T) d'accidents et pour des valeurs de S et de R relativement élevées. L'ordre de grandeur atteint 10^{-6} pour la deuxième ($S = 5; R = 10$) et quatrième colonne ($S = 20; R = 10$) du tableau 5 où l'espace des paramètres est inclus dans $\mathbb{R}^{51}$ dans le premier cas et dans $\mathbb{R}^{201}$ dans le second cas.

Une deuxième étude a été mise en œuvre afin d'apprécier le comportement de $EQM(\hat{\Theta}, \Theta^o)$ pour des effectifs d'accidents faibles et élevés. On a alors fixé la valeur de n_s dans une tranche donnée et fait varier le nombre de sites S et de types R d'accidents. Les tableaux 6 (resp. 7) donnent des valeurs de $EQM(\hat{\Theta}, \Theta^o)$ de l'ordre de 10^{-2} (resp. 10^{-5}). Ils permettent d'observer une amélioration notable de l'estimation et de la convergence lorsque les effectifs passent de $]0, 30[$ à $]1000; 3000[$ c'est-à-dire lorsque les effectifs d'accidents par site augmentent. Cette convergence de $EQM(\hat{\Theta}, \Theta^o)$ vers 0 quand augmentent les effectifs d'accidents est visible sur les figures 1 à 3 obtenues pour différentes valeurs fixées de S et R. Les graphes obtenus pour les premières valeurs de S et de R montrent bien que $EQM(\hat{\Theta}, \Theta^o)$ est approximativement de 10^{-2} pour les $n_s \in]0, 30[$ et de 10^{-5} pour les $n_s \in]11000, 13000[$.

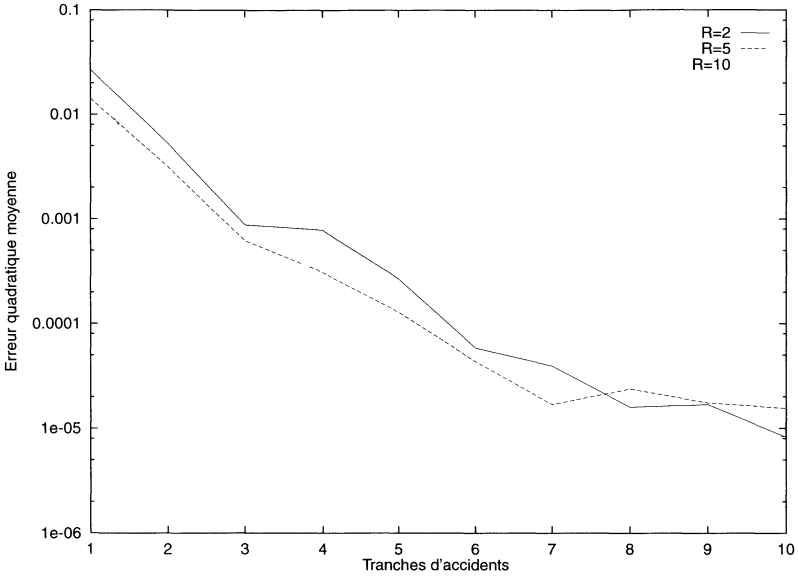

FIGURE 1
$S = 5$

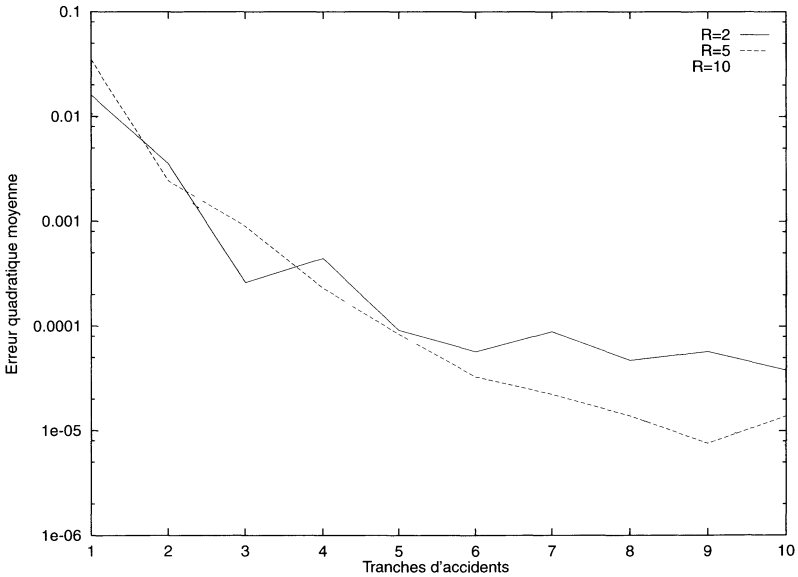

FIGURE 2
$S = 10$

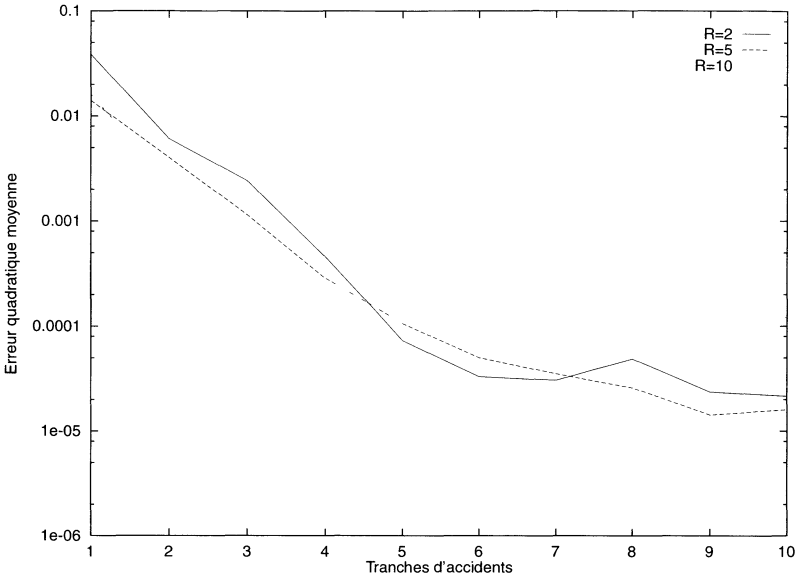

FIGURE 3
$S = 20$

En parallèle à cet ensemble d'études, nous avons analysé les résultats obtenus pour les coefficients moyens de contrôle (espérance mathématique des v.a. C_s) et pour l'effet commun de la mesure. Les résultats sont consignés dans les tableaux 8 à 15. Le tableau 8 (resp.9) est lié au tableau 3 (resp. 4). Ainsi les valeurs de c_s^o et $\hat{c_s}$ données dans ces tableaux sont obtenues en utilisant la formule (8) et les éléments de Θ^o et de $\hat{\Theta}$ du tableau 3 (resp. 4) sauf la première composante. On remarque, par exemple pour le tableau 8, que les valeurs de $\hat{c_s}$ sont d'autant plus proches des vraies valeurs c_s^o que les effectifs sont élevés (colonne $T = 4, \ldots, T = 10$). Cette convergence est mesurée par les critères EMC (Erreur Moyenne de Contrôle) et EQMC (Erreur Quadratique Moyenne de Contrôle) dont l'ordre de grandeur respectif varie de 10^{-1} (pour $n_s \in]0, 30[$) à 10^{-3} (pour $n_s \in]11000, 13000[$) et de 10^{-2} (pour $n_s \in]0, 30[$) à 10^{-6} (pour $n_s \in]11000, 13000[$). On obtient même des approximations de l'ordre de 10^{-5} et 10^{-6} lorsque que R passe de 2 à 3 (tableau 9). Quant aux tableaux 10 et 12 (resp. 11 et 13), ils vont de paire avec le tableau 6 (resp. 7). Ces quatre tableaux permettent d'apprécier une assez bonne stabilité de l'ordre de grandeur des deux critères EMC et EQMC. Enfin les tableaux 14 et 15 donnent l'estimation de l'effet commun de la mesure c'est-à-dire la première composante de $\hat{\Theta}$ et complètent les résultats des tableaux 3 et 4. On note dans les tableaux 3 et 4 que plus les effectifs d'accidents augmentent plus $\hat{\theta}$ tend vers 0.5 la vraie valeur de l'effet commun. Cette convergence est confirmée par la stabilité de l'ordre de grandeur des valeurs des tableaux 14 et 15, la valeur absolue de l'écart entre l'estimation et la vraie valeur atteint quelques fois 10^{-4}.

6. Remarques et conclusion

Selon certains points de vue et certaines pratiques dans le domaine de la sécurité routière (Lassarre 1977 [9], Danielson 1986 [3] et 1988 [4], Hauer 1997 [7]), il est souhaitable de modéliser les données d'accidents à l'aide de distribution de Poisson. Nous avons adopté dans ce travail la philosophie de Tanner. Les difficultés dans l'approche de modélisation de ce dernier résident essentiellement dans le choix de la zone de contrôle et dans l'hypothèse « non aléatoire » qu'il fait sur les coefficients de contrôle. Hauer (1991) [6] a proposé une notion de similarité pour choisir un site de comparaison. Néanmoins, il semble assez difficile de baser le choix d'un site de comparaison sur cette seule notion de similarité. Il est essentiel que le nombre d'accidents du site supposé de comparaison ou de contrôle soit suffisammment élevé afin que le phénomène aléatoire inhérent au comptage des accidents sur ce site ne dégrade pas la précision de l'estimation de l'effet de la mesure de sécurité. Notre approche intègre ce phénomène aléatoire sur le site de contrôle en imposant que les coefficients associés à ce site proviennent d'une distribution de probabilité estimée simultanément avec l'effet commun de la mesure.

Nos résultats généralisent l'approche de Tanner par le fait que les coefficients de contrôle sont distribués de façon aléatoire entre les types d'accidents et complètent ceux de N'Guessan (1993) [12], N'Guessan et Langrand (1992) [14]. L'inconvénient de la méthode est que la dimension de l'espace des paramètres croît avec le nombre de sites et de types d'accidents. Cependant, les résultats montrent que l'erreur quadratique moyenne est un $O(10^{-\beta})$ avec $\beta > 1$.

Bibliographie

[1] AITCHISON J., SILVEY S.D.(1958), *Maximum likelihood estimation of parameters subjet to restraints*, Annals of mathematical statistics, 29, 813–829.

[2] CROWDER M.(1984), *On the constrained maximum likelihood estimation with non i.i.d. observations*, Ann. Inst. Statist. Math.36, A, 239–249.

[3] DANIELSON S. (1986), *A comparison of two methods for estimating of the effect of a countermeasure in the presence of regression effects*, Accid. Anal. and Prev. 18, n^o1, 13–23.

[4] DANIELSON S. (1988), *Estimation of the effects of countermeasures on differents types of accidents in the presence of regression effects*, Accid. Anal. and Prev. 20, n^o4, 289–298.

[5] GOURIÉROUX CH., MONFORT A. (1989), *Statistique et modèles économétriques : vol* n^o1, Edition Economica.

[6] HAUER E. (1991), *Comparison Groups in road safety studies : An analysis*, Accid. Anal. and Prev. 23, n^o26, 609–622.

[7] HAUER E. (1997), *Observasional before-after Studies in Road Safety : Estimating the effect of highway and traffic engineering measures on road safety*, Pergamon Edition.

[8] LASSARRE S.(1985), *Panorama des méthodes associées à l'évaluation de type avant-après*, Colloque international sur l'évaluation des mesures locales de sécurité routière. Paris.

[9] LASSARRE S.(1977), *A propos des tests statistiques, sur variables poissonniennes, utilisées dans le domaine de la sécurité routière*, Revue de Statistique Appliquée XXV, $n^o 3$.

[10] NAG, *(2000) Numerical Algorithms Groups*, C.R.I. Université de Lille1.

[11] NEUENSCHWANDER B.E., FLURY B.D. (1997), *A note on Silvey's (1959) theorem*, Statistics and Probability Letters 36, 307–317.

[12] N'GUESSAN A. (1993), *Contribution à l'analyse statistique d'une mesure de sécurité routière*, Thèse de Doctorat Université de Lille 1. France.

[13] N'GUESSAN A., LANGRAND C. (1993), *Sur la distribution asymptotique de certaines statistiques utilisées dans le domaine de la sécurité routière*, C.R. Acad. Sci, t.317, Série I, 401–404.

[14] N'GUESSAN A., LANGRAND C. (1992), *Une méthode d'estimation de l'effet moyen d'une mesure de sécurité routière*, XXIVèmes Journées de Statistique de l'ASU, Bruxelles.

[15] TANNER J.C. (1958), *A problem in the combination of accident frequencies*, Biometrika, 45, 331–342.