

IMPACT D'UN AMÉNAGEMENT DE SÉCURITÉ ROUTIÈRE SUR LA GRAVITÉ DES ACCIDENTS DE LA ROUTE

Assi N'GUESSAN*, Marcel TRUFFIER**

RÉSUMÉ

Nous proposons l'évaluation d'une mesure de sécurité routière à partir des données d'accidents survenus en un site expérimental et dans une zone de contrôle qui lui est associée. Les données d'accidents du site et de la zone sont réparties en plusieurs types mutuellement exclusifs et en deux périodes d'observation (avant et après). Nous utilisons un modèle multinomial particulier pour modéliser le total des accidents observés au niveau du site expérimental en tenant compte du nombre des accidents de la zone de contrôle. Ensuite, nous estimons les paramètres du modèle ainsi que leurs erreurs standards en utilisant une approche par complément de Schur. Nous appliquons cette approche à des aménagements routiers effectués sur le réseau des routes nationales du Pas de Calais.

Mots-clés : Mesure de sécurité routière, données d'accidents de la route, Modèle de Poisson conditionnel, Modèle multinomial, Maximum de vraisemblance sous contraintes, Complément de Schur.

ABSTRACT

We propose the evaluation of a road safety layout using crash data that occurred on an experimental site and a matched control site. The crash data in both sites are divided in several mutually exclusive types and in two observation periods (before and after). We use a specific multinomial model to model the total number of crashes in the experimental site while taking into consideration the number of crashes in the control site. We estimate the model parameters and their standard errors using a Schur complement approach. We applied this approach on new road safety layouts done on the Pas de Calais national road network.

Keywords : Road Safety layouts, Accident data, Conditional Poisson Model, Multinomial model, Constrained Maximum Likelihood, Schur complement.

* Ecole Polytechnique Universitaire de Lille et Laboratoire Paul Painlevé UMR CNRS 8542 – Université de Lille 1, 59655 Villeneuve d'Ascq Cedex, France – Assi.Nguessan@polytech-lille.fr.

** Centre d'Etudes Techniques de l'Équipement Nord Picardie, 2 rue de Bruxelles, BP 275 59019 Lille Cedex, France – Marcel.Truffier@gouv-equipement.fr.

1. Introduction

Les travaux d'aménagements routiers (installation de giratoires, construction d'ilôts en dur, modification du marquage au sol, etc), effectués sur des zones reconnues dangereuses de certaines routes nationales, font partie des mesures de sécurité routière prises pour essayer de résoudre les problèmes d'insécurité sur ce type de réseau routier. Une fois les travaux terminés et la circulation rétablie, se pose la question de l'évaluation de la mesure prise, l'aménagement a-t-il fait baisser de façon significative la dangerosité de la zone concernée? Autrement dit, comment évaluer l'efficacité de tels aménagements routiers?

A cette fin, les autorités de tutelle relèvent les nombres d'accidents mortels, graves et légers de la zone aménagée (site expérimental) et ceux d'une zone de contrôle de la route nationale à laquelle appartient la dite zone aménagée. Les données ainsi obtenues sont relatives à deux périodes d'observation : une période précédant le début des travaux et une autre suivant la fin de ceux-ci. Il s'agit donc d'utiliser l'ensemble de ces données pour non seulement répondre à la question posée mais aussi analyser l'impact de l'aménagement sur les différents types de gravité d'accident.

La méthode d'évaluation que nous proposons est basée sur la technique du complément de Schur (voir par exemple Ouellette (1981), Zhang (2005)) et utilise une version particulière du modèle multinomial construit par N'Guessan *et al.* (2001) lorsqu'une mesure de sécurité routière est appliquée à plusieurs sites expérimentaux et plusieurs types d'accidents, chaque site étant associé à une zone de contrôle dont le but est de prendre en compte l'effet de certains facteurs indépendamment de la mesure (intensité du trafic, conditions météorologiques, ...). Cette approche par complément de Schur nous permet non seulement d'estimer formellement les paramètres du modèle statistique utilisé mais aussi de proposer une approximation de leur erreur standard. Pour ce faire, nous transformons d'abord les équations du maximum de vraisemblance sous contraintes en une équation non linéaire et en un sous système d'équations linéaires auquel nous appliquons un principe de résolution par complément de Schur. Nous construisons ainsi un algorithme itératif d'estimation des paramètres adapté aux données d'accidents de l'étude et ne faisant plus appel à des procédures de type Newton-Raphson utilisées par N'Guessan *et al.* (2001). Ensuite, nous complétons l'estimation des paramètres par celle de leur erreur standard en appliquant les récents résultats obtenus par ces mêmes auteurs à la matrice d'information associée au modèle. Des intervalles de confiance sont alors proposés pour chaque paramètre du modèle d'analyse afin de juger au mieux de la pertinence de l'efficacité moyenne de l'aménagement routier et de son impact sur les différentes gravités d'accidents. Enfin, nous appliquons nos résultats à une étude de cas concernant certaines opérations d'aménagements routiers effectués dans le département du Pas De Calais.

La suite de notre article est organisée de la façon suivante. Dans la section 2, nous exposons le modèle statistique utilisé pour combiner les données des deux entités d'observation, et donnons une interprétation pratique des paramètres du modèle statistique en terme d'effet de l'aménagement et en terme de risques

d'accidents. Dans la section 3, nous présentons la méthode d'estimation de l'effet moyen de l'aménagement et de son impact sur les différents types de risque d'accidents. Cette section 3 s'achève sur un bref rappel de résultats récents relatifs à l'estimation d'écart-types et d'intervalles de confiance par complément de Schur. La section 4 traite d'une étude de cas. En effet, nous analysons une des opérations du plan régional d'aménagements de sécurité routière effectuée entre 1997 et 1999 dans le département du Pas de Calais. Cette étude traite de la modification du marquage au sol sur un tronçon de la route nationale 17 afin d'empêcher les dépassements anarchiques de certains usagers. Cet aménagement routier a eu lieu en 1999 et se situe au niveau des communes de Vimy et d'Avion. Nous discutons, dans la section 5, de l'approche de modélisation proposée, des résultats obtenus ainsi que des possibilités d'extension du modèle statistique et de la méthode d'analyse. Les annexes A et B reprennent quelques détails techniques concernant le système d'estimation des paramètres, l'algorithme et le calcul direct de la matrice d'information utilisée pour estimer les erreurs standards.

2. Modélisation de l'impact d'un aménagement routier en présence de plusieurs types de gravité d'accidents

Pour modéliser et estimer l'impact d'un aménagement routier sur différents types de gravité d'accidents, et prendre en considération les informations (nombres d'accidents) disponibles dans une zone de contrôle, nous utilisons le modèle statistique construit par N'Guessan *et al.* (2001). Ce modèle est une extension de celui décrit par Tanner (1958) pour évaluer l'effet moyen d'une modification apportée au réseau routier, au cas où plusieurs types d'accidents sont considérés dans le même modèle, comme par exemple dans les présentes études, les accidents mortels, avec blessés graves ou avec blessés légers.

En utilisant les notations données dans N'Guessan *et al.* (2001), on pose R ($R > 1$) le nombre de types d'accidents, x_{1j} (resp. x_{2j}) le nombre d'accidents du type j ($j = 1, 2, \dots, R$) observés avant (resp. après) la mise en place de l'aménagement routier, $n = \sum_{j=1}^R (x_{1j} + x_{2j})$ le nombre total d'accidents enregistrés sur la zone aménagée durant les deux périodes d'observation. On suppose également disposer, pendant ces deux périodes, des données d'accidents relatives à la portion de la route nationale utilisée comme contrôle et on note c_1, c_2, \dots, c_R les coefficients de contrôle où c_j est le quotient du nombre d'accidents du type j enregistrés au cours de la période suivant les travaux d'aménagement sur le nombre d'accidents du type j enregistrés au cours de la période précédant ceux-ci. Par la suite, on note pour la zone aménagée X_{1j} (resp. X_{2j}) la variable aléatoire donnant le nombre d'accidents du type j de la période avant (resp. après) l'aménagement, dont une valeur observée est x_{1j} (resp. x_{2j}). En utilisant un principe d'effet multiplicateur de l'aménagement sur le nombre moyen d'accidents par type, on montre (voir N'Guessan *et al.* 2001) que la loi de probabilité conditionnelle (conditionnellement au nombre total d'accidents observés avant et après au niveau du site expérimental) de la paire (X_1, X_2) où $X_1 = (X_{11}, X_{12}, \dots, X_{1R})^T$ et $X_2 = (X_{21}, X_{22}, \dots, X_{2R})^T$

est une loi multinomiale de dimension $2R$ dont les probabilités de classe sont données par $\pi_{1j}(\Theta)$ et $\pi_{2j}(\Theta)$ avec

$$\pi_{1j}(\Theta) = \frac{p_j}{1 + \beta_0 \sum_{m=1}^R p_m c_m}, \quad \pi_{2j}(\Theta) = \frac{\beta_0 c_j p_j}{1 + \beta_0 \sum_{m=1}^R p_m c_m}, \quad j = 1, 2, \dots, R, \quad (1)$$

sous les contraintes

$$p_j > 0, \quad \text{et} \quad \sum_{j=1}^R p_j = 1; \quad (2)$$

$\Theta = (\beta_0, P^T)^T$, de dimension $(1 + R)$, est le vecteur des paramètres inconnus du modèle où β_0 ($\beta_0 > 0$) décrit l'impact moyen (ou indice d'efficacité moyenne) de l'aménagement sur les nombres d'accidents, $P = (p_1, p_2, \dots, p_R)^T$, de dimension R , est un vecteur de probabilités dont chaque composante p_j permet d'apprécier le risque global d'accident de type j au niveau du site aménagé pour l'ensemble des deux périodes d'étude. Comme, $\pi_{1j}(\Theta)$ (resp. $\pi_{2j}(\Theta)$) représente, par construction du modèle statistique, la probabilité qu'un accident survenu au niveau de la zone aménagée soit du type j avant (resp. après), on propose d'utiliser ces probabilités comme des indicateurs du risque d'accidents du type j avant (resp. après) au niveau de la zone aménagée conditionnellement aux données de la zone de contrôle associée.

Pour interpréter l'impact global de l'aménagement routier sur l'ensemble des effectifs d'accidents, on compare le coefficient β_0 à 1. Une valeur de β_0 supérieure à 1 indique que l'aménagement routier n'a pas fait baisser, de façon significative, les effectifs de la zone aménagée par rapport aux données d'accidents de la zone de contrôle. Au contraire, une valeur de β_0 significativement inférieure à 1 témoigne d'une baisse importante en proportion du nombre des accidents de la zone aménagée par rapport à la zone de contrôle. Enfin, une valeur de β_0 non significativement différente de 1 indique que l'aménagement routier n'a pas eu d'effet sur l'ensemble des données d'accidents de la zone aménagée comparativement aux données de la zone de contrôle. On parle alors d'efficacité moyenne nulle.

3. Estimation sous contraintes des paramètres du modèle et de leur erreur standard

3.1. Estimation de l'effet moyen de l'aménagement et de son impact sur les différents risques d'accidents de l'étude

L'évaluation de l'effet moyen de l'aménagement et de son impact sur les différents types de risque d'accidents exige l'estimation préalable des paramètres $\beta_0, p_1, p_2, \dots, p_R$. Cette estimation est obtenue à l'aide de la méthode du maximum de vraisemblance. En effet, étant donné x_{1j} (resp. x_{2j}) le nombre d'accidents du type j ($j = 1, 2, \dots, R$) observés avant (resp. après)

et en utilisant la loi multinomiale (1), la fonction de vraisemblance associée à la paire (X_1, X_2) est (voir N'Guessan *et al.* (2001)) :

$$l(X_1, X_2; \Theta) = \frac{n!}{2 \prod_{t=1}^R \prod_{m=1}^R x_{tm}!} \left[\prod_{j=1}^R \frac{p_j^{x_{\bullet j}} (\beta_0 c_j)^{x_{2j}}}{(1 + \beta_0 \sum_{m=1}^R c_m p_m)^{x_{\bullet j}}} \right]$$

où $x_{\bullet j} = x_{1j} + x_{2j}$. Il s'agit de maximiser la fonction $l(X_1, X_2; \Theta)$ par rapport aux composantes du vecteur Θ ou de façon équivalente de maximiser la fonction log-vraisemblance donnée par

$$\mathcal{L}(X_1, X_2; \Theta) = \text{constante} + \sum_{j=1}^R \left\{ X_{\bullet j} \log_e(p_j) + X_{2j} \log_e(\beta_0) - X_{\bullet j} \log_e(1 + \beta_0 \sum_{m=1}^R c_m p_m) \right\}. \quad (3)$$

Puisque, par construction, le vecteur P appartient au simplexe d'ordre R , l'estimation des paramètres se fait dans le cadre de la méthode du maximum de vraisemblance sous contraintes (voir par exemple Aitchison et Silvey 1958, Crowder 1984, Haber et Brown 1986, Matthews et Crowther 1995). En prenant les dérivées partielles premières de la log-vraisemblance (voir annexe B) par rapport à $\beta_0, p_1, p_2, \dots, p_R$ et au multiplicateur de Lagrange associé à la contrainte $\sum_{j=1}^R p_j = 1$, et en égalant à zéro, on montre que l'estimateur du maximum de vraisemblance $\hat{\Theta}$ de Θ , s'il existe, est solution du système d'équations non linéaires :

$$\left\{ \begin{array}{l} \sum_{j=1}^R \left\{ X_{2j} - X_{\bullet j} \frac{\hat{\beta}_0 \sum_{m=1}^R c_m \hat{p}_m}{1 + \hat{\beta}_0 \sum_{m=1}^R c_m \hat{p}_m} \right\} = 0 \\ X_{\bullet j} - n \frac{\hat{p}_j (1 + c_j \hat{\beta}_0)}{1 + \hat{\beta}_0 \sum_{m=1}^R c_m \hat{p}_m} = 0, \quad (j = 1, 2, \dots, R); \\ 0 < \hat{p}_j \text{ avec } \sum_{j=1}^R \hat{p}_j = 1. \end{array} \right. \quad (4)$$

Ce système d'équations non linéaires est un cas particulier de celui de N'Guessan *et al.* (2001) où l'aménagement routier concerne un seul site expérimental et plusieurs types d'accident. La mise en oeuvre proposée par ces derniers auteurs passe par la résolution simultanée des deux équations du système et nécessite un vecteur solution initiale et l'inversion d'une matrice à chaque itération. Pour ce faire, les auteurs donnent manuellement une valeur initiale à β_0 proche de zéro en prenant le rapport $1/R$ comme solution initiale

pour chaque paramètre p_j . Le point de convergence numérique de l'algorithme itératif, de type Newton-Raphson proposé par ces auteurs, est alors utilisé comme solution des équations du maximum de vraisemblance. Or, il est bien connu qu'une solution optimale d'un algorithme itératif dépend fortement des données du problème mais aussi de la solution initiale utilisée.

Nous proposons dans ce travail une procédure itérative (voir annexe A) pour automatiser la recherche de $\hat{\Theta}$, prendre en compte les données d'accident, construire une solution initiale adaptée à chaque étude, et éviter l'inversion d'une matrice à chaque itération. Notre approche est fondée sur le résultat suivant.

THÉOREME 3.1. — *On montre que les composantes du vecteur $\hat{\Theta}$, solution du système d'équations (4), sont données par*

$$\left\{ \begin{array}{l} \hat{\beta}_0 = \frac{\sum_{m=1}^R X_{2m}}{\left(\sum_{m=1}^R c_m \hat{p}_m\right) \left(\sum_{m=1}^R X_{1m}\right)}; \\ \hat{p}_j = \frac{1}{1 - \frac{1}{n} \sum_{m=1}^R \frac{X_{\bullet m} \hat{\beta}_0 c_m}{1 + \hat{\beta}_0 c_m}} \times \frac{X_{\bullet j}}{n(1 + \hat{\beta}_0 c_j)}, \quad (j = 1, 2, \dots, R); \\ 0 < \hat{p}_j \text{ avec } \sum_{j=1}^R \hat{p}_j = 1. \end{array} \right. \quad (5)$$

La preuve (voir annexe A pour les grandes lignes) consiste à montrer, dans un premier temps, que pour le paramètre β_0 fixé, le deuxième sous-système du système (4) des équations de vraisemblance est équivalent à un système linéaire par rapport aux paramètres p_1, p_2, \dots, p_R . La matrice et le second membre de ce système linéaire sont fonction de β_0 et des données d'accident de l'étude considérée. Dans un second temps, on démontre que la matrice associée au système linéaire est formellement (c'est-à-dire sans utilisation de valeurs numériques) inversible. On en déduit alors l'expression des \hat{p}_j et par substitution celle de $\hat{\beta}_0$.

Du fait de la dépendance fonctionnelle entre les composantes du vecteur $\hat{\Theta}$, une solution optimale ne peut être obtenue que de façon itérative. A cette fin, on initialise β_0 à 0. En utilisant l'algorithme décrit en annexe A, on estime à l'étape $k \geq 1$, les $\hat{p}_j^{(k)}$ à l'aide de la deuxième équation du système (5), puis $\hat{\beta}_0^{(k)}$ et ainsi de suite.

3.2. Approximation des erreurs standards via le complément de Schur

Le processus d'évaluation de l'aménagement routier passe également par la construction d'intervalles de confiance et par la mise en place de tests

d'hypothèses comme, par exemple, tester l'hypothèse d'efficacité moyenne nulle. Pour ce faire, on a besoin d'inverser la matrice d'information associée aux paramètres et aux contraintes et d'en extraire les éléments diagonaux. Dans le cas particulier d'un seul site aménagé et de plusieurs types de gravité d'accidents, cette matrice prend la forme particulière suivante :

$$\Gamma_{\Theta} = \begin{bmatrix} \frac{\gamma^2 \sum_{m=1}^R c_m p_m}{n \beta_0} & \frac{\gamma^2 c_1}{n} & \dots & \frac{\gamma^2 c_R}{n} & 0 \\ \frac{\gamma^2 c_1}{n} & \gamma \left(\frac{1 + \beta_0 c_1}{p_1} - \frac{\gamma \beta_0^2 c_1^2}{n} \right) & \dots & -\frac{\gamma^2 \beta_0^2 c_1 c_R}{n} & 1 \\ \frac{\gamma^2 c_2}{n} & -\frac{\gamma^2 \beta_0^2 c_2 c_1}{n} & \ddots & -\frac{\gamma^2 \beta_0^2 c_2 c_R}{n} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\gamma^2 c_R}{n} & -\frac{\gamma^2 \beta_0^2 c_R c_1}{n} & \dots & \gamma \left(\frac{1 + \beta_0 c_R}{p_R} - \frac{\gamma \beta_0^2 c_R^2}{n} \right) & 1 \\ 0 & 1 & \dots & 1 & 0 \end{bmatrix}$$

où $\gamma = n / (1 + \beta_0 \sum_{m=1}^R c_m p_m)$. Cette matrice d'information sous contraintes linéaires est de dimension $(2+R) \times (2+R)$ et est obtenue à l'aide de l'espérance des dérivées partielles secondes de la fonction

$$\mathcal{L}_1(X_1, X_2; \Theta, \lambda) = \mathcal{L}(X_1, X_2; \Theta) - \lambda \left(\sum_{j=1}^R p_j - 1 \right) \quad (6)$$

par rapport aux composantes du vecteur paramètre Θ et au multiplicateur de Lagrange λ (voir les détails des calculs en annexe B). Une fois les données d'accident connues et les paramètres estimés, l'inversion de Γ_{Θ} est formellement obtenue à l'aide des résultats donnés dans N'Guessan et Langrand (2005a, 2005b). Ces résultats, basés sur la technique du complément de Schur, ont l'avantage de remplacer l'inversion directe de Γ_{Θ} par des produits et des sommes de matrices de dimensions plus petites que celles de Γ_{Θ} . De plus, ces résultats s'appliquent à un nombre quelconque de sites expérimentaux et de types d'accidents et permettent d'obtenir l'expression analytique de la variance de chaque composante. On peut alors construire des intervalles de confiance (N'Guessan et Bellavance (2005)) et bâtir des tests d'hypothèses relatifs à l'effet moyen de l'aménagement et aux différents types de risque de gravité d'accident. En utilisant la technique du complément de Schur, on obtient les intervalles de confiances à $(1 - \alpha)100\%$ suivants :

$$\begin{aligned} IC_{1-\alpha}(\beta_0) &= \left[\max(0; \hat{\beta}_0 - \epsilon_{\alpha/2} \sigma(\hat{\beta}_0)); \hat{\beta}_0 + \epsilon_{\alpha/2} \sigma(\hat{\beta}_0) \right]; \\ IC_{1-\alpha}(p_j) &= \left[\max(0; \hat{p}_j - \epsilon_{\alpha/2} \sigma(\hat{p}_j)); \min(1; \hat{p}_j + \epsilon_{\alpha/2} \sigma(\hat{p}_j)) \right], \quad (7) \\ &\quad (j = 1, 2, \dots, R), \end{aligned}$$

où $\epsilon_{\alpha/2}$ est obtenu de façon usuelle à l'aide de la table de la loi normale centrée et réduite, $\sigma(\bullet)$ est l'écart-type de chaque composante du paramètre Θ (voir par exemple N'Guessan et Langrand (2005a)).

Remarque : les écart-types utilisés dans la construction de ces intervalles sont obtenus en supposant n grand (généralement une valeur de n supérieure à 30 est utilisée comme critère asymptotique). Les intervalles de confiance ainsi construits complètent les valeurs ponctuelles des paramètres et permettent de juger convenablement de l'impact (négatif ou positif) de l'aménagement routier sur l'incidence des accidents de la route surtout lorsque l'effectif n est proche de la limite de validité comme dans l'étude de cas ci-dessous.

4. Application à l'évaluation de la modification d'une signalisation au sol

Un aménagement routier a été effectué en 1999 sur la route nationale 17 (RN17) au niveau des communes de Vimy et d'Avion du département Pas de Calais. En effet, la route, entre ces deux communes, était une trois voies classique, à deux sens. L'aménagement a consisté à ce que l'on appelle un «marquage à l'Italienne». On modifie le marquage au sol pour qu'il soit impossible de doubler en même temps dans chaque sens. Sur une courte distance, on donne deux voies pour un sens (on peut alors doubler dans ce sens) et une seule voie sur l'autre sens séparées par une ligne continue. Plus loin, on inverse pour que l'autre sens puisse doubler et ainsi de suite. Les données relevées au niveau de la zone aménagée (Vimy-Avion) sont les suivantes (tableau 1) :

TABLEAU 1.

	période avant (4 ans)			période après (4 ans)			total
	M	G	L	M	G	L	
zone aménagée	4	4	16	1	1	7	33
zone de contrôle	27	64	182	14	27	102	416

Pour des périodes d'observation de longueur 4 ans, les accidents mortels sont passés de 4 avant à 1 après. Il y a eu donc 4 accidents avec au moins un tué au cours de la période précédant les travaux et 1 accident avec au moins un tué au cours de la période suivant ceux-ci. Pendant ce temps les accidents légers (accidents sans tué ni blessé grave) répertoriés ont baissé plus que de moitié. Pour les mêmes longueurs de périodes, on observe, sur la portion de la RN17 servant de zone de contrôle, les rapports suivants (tableau 2) :

TABLEAU 2.

	c_1	c_2	c_3
zone de contrôle	0.5190	0.4220	0.5600

Les valeurs du tableau 2 sont obtenues à l'aide des données d'accidents de la zone de contrôle du tableau 1 : $c_1 = 14/27$, $c_2 = 27/64$, $c_3 = 102/182$.

Globalement on note une baisse en proportion du nombre des accidents de la zone de contrôle, entre la période de 4 ans suivant les travaux et la période de 4 ans précédant ceux-ci. Les résultats numériques relatifs à ces données d'accidents sont les suivants (tableau 3) :

TABLEAU 3. — Estimation des paramètres du modèle

paramètres	estimation (E.S)	B_{inf} à 95%	B_{sup} à 95%
$\hat{\beta}_0$	0.7054 (0.2759)	0.1646	1.2462
\hat{p}_1	0.1525 (0.0643)	0.0264	0.2785
\hat{p}_2	0.1605 (0.0672)	0.0287	0.2922
\hat{p}_3	0.6870 (0.1023)	0.4864	0.8875

L'indice d'efficacité moyenne estimé à l'aide des données des tableaux 1 et 2 et du système d'équations (4) est de 0.7054, ce qui correspond à une baisse moyenne en proportion de $29,46\% = (1 - 0,7054) \times 100\%$ de l'ensemble des accidents de la zone aménagée (Vimy-Avion) par rapport à la tendance moyenne de la zone de contrôle. Les valeurs estimées des trois paramètres \hat{p}_1 , \hat{p}_2 et \hat{p}_3 à l'aide du système d'équations (4) sont respectivement 0,1525, 0,1605 et 0,6870. Ces valeurs permettent d'apprécier le niveau des trois types de risque d'accidents du site expérimental au cours des huit années utilisées pour analyser les travaux de marquage au sol. Cette répartition montre qu'au cours des huit années d'analyse des travaux de marquage au sol, 15,25% des accidents répertoriés au niveau de la zone aménagée ont été mortels, 16,05% ont été graves et 68,70% ont été légers.

On peut également tester la significativité de l'effet moyen de ce type d'aménagement en utilisant l'intervalle de confiance au niveau 95% associé au paramètre β_0 . Cet intervalle de confiance montre qu'on a une fourchette de valeurs dont la borne inférieure (resp. supérieure) est strictement plus grande que 0 (resp. 1). En effet, comme $1 \in [0.1646; 1.2462]$, on ne peut rejeter l'hypothèse $H_0 : \beta_0 = 1$ contre $H_1 : \beta_0 \neq 1$ avec une erreur de première espèce de 5%. La valeur estimée de β_0 pouvant même excéder 1, ce type d'aménagement routier est susceptible d'induire un effet pervers sur le nombre moyen des accidents au niveau de la zone aménagée. Même si dans le cas présent, on note une baisse en proportion de 29,46% du nombre moyen des accidents de la zone aménagée (Vimy-Avion), le test ci-dessus montre que la valeur de l'indice d'efficacité moyenne n'est pas statistiquement significativement différente de 1 pour conclure de façon systématique à une efficacité avérée de ce type de marquage au sol. En pratique, il convient de procéder à une analyse en fonction des périodes, des données observées et de la zone de contrôle utilisée afin d'obtenir des conclusions adéquates.

Pour apprécier l'impact de ce type de marquage sur les différents types de risque d'accidents par période de temps au niveau du site expérimental, on utilise les probabilités conditionnelles $\pi_{1j}(\hat{\Theta})$ et $\pi_{2j}(\hat{\Theta})$. La valeur estimée de ces probabilités est donnée dans le tableau 4 :

TABLEAU 4. — Estimation des risques d'accidents par période de temps

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

Les données d'accidents répertoriés et les périodes d'analyse utilisées permettent d'observer une baisse perceptible des trois types de risque d'accidents de la zone aménagée. On estime qu'au cours des quatre années précédant le marquage au sol à Vimy-Avion, le risque d'accidents mortels de la zone aménagée était de l'ordre de 11,09%, celui des accidents graves de 11,67% et celui des accidents légers de 49,96%. Au cours de la période des quatre années suivant la mise en service de la zone, ces différents risques d'accidents sont descendus respectivement à 4%, 3,47% et 19,73%. Selon les données d'accidents répertoriés et la longueur des périodes utilisées, ce type de marquage semble avoir localement un impact significatif sur les différents types de risque d'accidents de la zone aménagée.

5. Discussion

5.1. Principe de modélisation

Par construction du modèle multinomial, on constate que le paramètre β_0 intervient dans la probabilité $\pi_{1j}(\Theta)$ avant aménagement. Il en est de même dans le cas de la loi binomiale proposée par Tanner (1958) et dont nous nous sommes inspirés. Il est donc légitime de se poser la question suivante : pourquoi un paramètre présenté comme un indice d'efficacité d'un aménagement intervient-il dans une probabilité d'accident avant ? Il faut noter que cette présence de β_0 se fait à travers l'expression $1 + \beta_0 \sum_{m=1}^R c_m p_m$ qui n'est rien d'autre qu'un facteur de normalisation à 1 de la loi multinomiale utilisée pour répartir l'effectif total d'accidents de la zone aménagée. En terme de nombre d'accidents et selon le principe de construction du modèle d'analyse (voir par exemple N'Guessan *et al.* (2001)), le paramètre β_0 n'intervient que dans les nombres moyens d'accidents après aménagement.

Une autre question récurrente dans les modèles avant-après avec contrôle est le fait de considérer que les coefficients de contrôle sont fixes et connus alors qu'ils sont définis à l'aide de nombres d'accidents. Si ces nombres sont importants, cette hypothèse paraît concevable. Par contre, si ces nombres d'accidents sont faibles, les coefficients de contrôle peuvent être très variables d'un type de gravité à un autre et le fait de ne pas tenir compte de cette variation supplémentaire dans le modèle peut induire une sous-estimation ou une sur-estimation des intervalles de confiance. Pour améliorer cette façon d'utiliser les données de la zone de contrôle, il conviendrait de faire une analyse de sensibilité aux différents coefficients de contrôle en les supposant

tous égaux à 1, c'est-à-dire faire une évaluation sans la présence de zone de contrôle. Il faudrait alors comparer la longueur des intervalles de confiance pour mieux apprécier l'impact de l'aménagement. On peut également, tout en gardant cette définition des coefficients de contrôle, adopter une approche de modélisation proposée par N'Guessan et al. (2006) pour pondérer les éventuelles variations entre les différents coefficients de contrôle d'une même zone. Il faudrait alors pouvoir adapter notre méthode d'estimation à leurs équations de vraisemblance.

Par ailleurs et selon certains points de vue, on peut se poser la question de l'aspect aléatoire des nombres totaux d'accidents ainsi que celle de l'utilisation d'autres méthodes comme le bootstrap (Efron et Tibshirani, 1993, p.45) ou les méthodes bayésiennes (voir par exemple Hauer 1997, p. 175). C'est cette dernière méthode que nous introduisons dans ce travail mais en supposant l'effectif total n non aléatoire et en utilisant un principe d'effet multiplicateur (voir par exemple Oppe, 1979). Certes, il est possible de supposer que les effectifs totaux sont distribués selon une loi de Poisson (Voir par exemple Nicholson 1986, Nicholson et Wong 1993). Mais en pratique, l'on ne dispose que d'un seul effectif total par site expérimental. Il faut alors s'intéresser, de façon rétrospective, à l'impact de l'aménagement sur les différents types de gravité d'accident : d'où la répartition multinomiale proposée.

5.2. Principe d'estimation des paramètres

L'estimation des paramètres par la méthode classique du maximum de vraisemblance sous contraintes utilisée dans N'Guessan *et al.* (2001) est fonction de la donnée d'un vecteur solution initiale et de l'inversion d'une matrice d'information à chaque itération. Les contraintes (en particulier les contraintes de bornes) liées au vecteur Θ font que certaines de ces composantes peuvent tendre vers 0 lorsqu'on utilise directement la méthode classique avec une solution initiale mal adaptée. Il est alors impossible de définir la log-vraisemblance. Outre sa simplicité et son accessibilité pour les utilisateurs, l'algorithme d'estimation proposé dans ce travail permet d'adapter la solution initiale aux données d'accident de l'aménagement. De surcroît, on n'est plus obligé de fournir manuellement une solution initiale et les solutions obtenues sont très proches de celles fournies par le maximum de vraisemblance sous contraintes. De plus, notre approche permet d'éviter l'inversion de la matrice d'information à chaque itération. Du coup, on gagne en nombre d'itérations par rapport à l'algorithme classique de Newton-Raphson. Même si la dimension de la matrice d'information relative à l'exemple de la section 4 est faible, cette dimension peut, selon la nature de l'aménagement, être très élevée (voir par exemple N'Guessan et Langrand (2005a)) et la matrice d'information associée creuse.

Annexe A : Recherche de solutions aux équations de vraisemblance sous contraintes

Nous présentons ici les détails techniques de la preuve du théorème 3.1 ainsi que la procédure itérative par complément de Schur permettant la recherche

de $\hat{\Theta}$, le choix d'une solution initiale adaptée aux données d'accident de l'étude. Le résultat principal est donné par le théorème 5.1 ci-dessous. Sa démonstration passe par quelques résultats intermédiaires structurés en deux lemmes et par un bref rappel de l'utilisation du complément de Schur pour l'inversion d'une matrice partitionnée en blocs (voir par exemple Zhang (2005) p. 163-167).

Considérons la matrice partitionnée suivante

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

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

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

On suppose maintenant que la matrice carrée

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

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

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

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

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

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

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

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

THÉORÈME 5.1. — *Pour tout β_0 ($\beta_0 > 0$) fixé , on montre que le vecteur $P(\beta_0)$ solution des équations*

$$X_{\bullet j} - n \frac{\hat{p}_j(1 + c_j \hat{\beta}_0)}{R} = 0, \quad (j = 1, 2, \dots, R);$$

$$1 + \hat{\beta}_0 \sum_{m=1}^R c_m \hat{p}_m$$

$$0 < \hat{p}_j \text{ avec } \sum_{j=1}^R \hat{p}_j = 1. \quad (8)$$

est donné par :

$$P(\beta_0) =$$

$$\frac{1}{1 - \frac{1}{n} \sum_{m=1}^R \left\{ \frac{\beta_0 c_m}{1 + \beta_0 c_m} X_{\bullet m} \right\}} \left(\frac{X_{\bullet 1}}{n(1 + \beta_0 c_1)}, \frac{X_{\bullet 2}}{n(1 + \beta_0 c_2)}, \dots, \frac{X_{\bullet R}}{n(1 + \beta_0 c_R)} \right)^T \quad (9)$$

On se donne une valeur initiale $\hat{\beta}_0^{(0)} = 0$ et on calcule le vecteur initial $\hat{P}^{(0)} = x_{\bullet}/n$ obtenu en remplaçant β_0 par zéro dans l'expression (9). Le vecteur initial

$$\hat{\Theta}^{(0)} = \left(0, \frac{x_{\bullet 1}}{n}, \frac{x_{\bullet 2}}{n}, \dots, \frac{x_{\bullet R}}{n} \right)^T$$

est ensuite utilisé pour amorcer l'algorithme d'approximation de $\hat{\Theta}$.

LEMME 5.2. — *On note*

$$Z = (c_1, c_2, \dots, c_R)^T, \quad X_{\bullet} = (X_{\bullet 1}, X_{\bullet 2}, \dots, X_{\bullet R})^T, \quad 1_R = (1, \dots, 1)^T$$

les vecteurs de dimension $1 \times R$ ainsi définis avec $X_{\bullet j} = X_{1j} + X_{2j}$ et

$$M_{\beta_0} = \text{diag}(1 + \beta_0 c_1, \dots, 1 + \beta_0 c_R),$$

la matrice diagonale $R \times R$ ainsi définie. Alors, on a

$$(i) \quad \frac{\beta_0}{n} Z^T M_{\beta_0}^{-1} X_{\bullet} < 1; \quad (ii) \quad \frac{\beta_0}{n} Z^T M_{\beta_0}^{-1} X_{\bullet} + \frac{1}{n} 1_R^T M_{\beta_0}^{-1} X_{\bullet} = 1.$$

Démonstration. — Les deux points de ce lemme proviennent de manipulations matricielles. On montre, par exemple, que

$$\frac{\beta_0}{n} Z^T M_{\beta_0}^{-1} X_{\bullet} = \frac{1}{n} \sum_{j=1}^R \frac{X_{\bullet j} \beta_0 c_j}{1 + \beta_0 c_j} < 1$$

■

LEMME 5.3. — *Soit*

$$M_{\beta_0}^{(1)} = \begin{bmatrix} M_{\beta_0} & (\frac{\beta_0}{n})^{1/2} X_{\bullet} \\ (\frac{\beta_0}{n})^{1/2} Z^T & 1 \end{bmatrix} \quad (10)$$

la matrice $(1 + R) \times (1 + R)$ ainsi définie. Alors $(M_{\beta_0}^{(1)}/M_{\beta_0})$ et $(M_{\beta_0}^{(1)}/1)$ respectivement les compléments de Schur de M_{β_0} dans $M_{\beta_0}^{(1)}$ et de 1 dans $M_{\beta_0}^{(1)}$ existent et sont inversibles avec :

$$(i) \quad (M_{\beta_0}^{(1)}/M_{\beta_0}) > 0;$$

$$(ii) \quad (M_{\beta_0}^{(1)}/1)^{-1} = (M_{\beta_0} - \frac{\beta_0}{n} X_{\bullet} Z^T)^{-1} = M_{\beta_0}^{-1} + \frac{\beta_0}{n} M_{\beta_0}^{-1} X_{\bullet} (M_{\beta_0}^{(1)}/M_{\beta_0})^{-1} Z^T M_{\beta_0}^{-1};$$

$$(iii) \quad \frac{1}{n} 1_R^T (M_{\beta_0}^{(1)}/1)^{-1} X_{\bullet} - 1 = 0.$$

Démonstration. — (i) Par définition, $(M_{\beta_0}^{(1)}/M_{\beta_0}) = 1 - \frac{\beta_0}{n} Z^T M_{\beta_0}^{-1} X_{\bullet}$. D'où le résultat d'après le point (i) du lemme 5.2. Le point (ii) est une conséquence

immédiate des propriétés du complément de Schur (voir par exemple Ouellette 1981, Zhang ed. 2005). En utilisant l'expression de l'inverse de $(M_{\beta_0}^{(1)}/1)^{-1}$ et le point (ii) du lemme 5.2 on a

$$\begin{aligned} \frac{1}{n} \mathbf{1}_R^T (M_{\beta_0}^{(1)}/1)^{-1} X_{\bullet} &= \frac{1}{n} \mathbf{1}_R^T M_{\beta_0}^{-1} X_{\bullet} (M_{\beta_0}^{(1)}/M_{\beta_0})^{-1} \\ &= \frac{(M_{\beta_0}^{(1)}/M_{\beta_0})}{(M_{\beta_0}^{(1)}/M_{\beta_0})} \\ &= 1 \end{aligned}$$

■

On peut maintenant démontrer le résultat du théorème 5.1. En effet, on montre, après quelques transformations, que le système

$$\begin{aligned} X_{\bullet j} - n \frac{\hat{p}_j (1 + c_j \hat{\beta}_0)}{R} &= 0, \quad (j = 1, 2, \dots, R); \\ 1 + \hat{\beta}_0 \sum_{m=1}^R c_m \hat{p}_m & \\ 0 < \hat{p}_j \text{ avec } \sum_{j=1}^R \hat{p}_j &= 1, \end{aligned}$$

est équivalent à

$$\left\{ \begin{array}{l} \hat{p}_1 (1 + c_1 \hat{\beta}_0) - \hat{\beta}_0 \frac{X_{\bullet 1}}{n} \left(\sum_{m=1}^R c_m \hat{p}_m \right) = \frac{X_{\bullet 1}}{n} \\ \hat{p}_2 (1 + c_2 \hat{\beta}_0) - \hat{\beta}_0 \frac{X_{\bullet 2}}{n} \left(\sum_{m=1}^R c_m \hat{p}_m \right) = \frac{X_{\bullet 2}}{n} \\ \dots \\ \hat{p}_R (1 + c_R \hat{\beta}_0) - \hat{\beta}_0 \frac{X_{\bullet R}}{n} \left(\sum_{m=1}^R c_m \hat{p}_m \right) = \frac{X_{\bullet R}}{n} \\ \hat{p}_1 + \hat{p}_2 + \dots + \hat{p}_R = 1, \\ 0 < \hat{p}_j \quad (j = 1, 2, \dots, R) \end{array} \right. \quad (11)$$

On en déduit alors le système linéaire suivant :

$$\begin{bmatrix} (M_{\beta_0}^{(1)}/1) & \mathbf{1}_R \\ \mathbf{1}_R^T & 0 \end{bmatrix} \times \begin{bmatrix} P \\ 0 \end{bmatrix} = \begin{bmatrix} n^{-1} X_{\bullet} \\ 1 \end{bmatrix}, \quad (12)$$

où

$$(M_{\beta_0}^{(1)}/1) = \begin{bmatrix} 1 + (1 - \frac{X_{\bullet 1}}{n})\beta_0 c_1 & -\frac{X_{\bullet 1}}{n}\beta_0 c_2 & \dots & -\frac{X_{\bullet 1}}{n}\beta_0 c_R \\ -\frac{X_{\bullet 2}}{n}\beta_0 c_1 & 1 + (1 - \frac{X_{\bullet 2}}{n})\beta_0 c_2 & \dots & -\frac{X_{\bullet 2}}{n}\beta_0 c_R \\ \vdots & \ddots & \ddots & \vdots \\ -\frac{X_{\bullet R}}{n}\beta_0 c_1 & -\frac{X_{\bullet R}}{n}\beta_0 c_2 & \dots & 1 + (1 - \frac{X_{\bullet R}}{n})\beta_0 c_R \end{bmatrix} \quad (13)$$

En utilisant les points (i) et (ii) du lemme 5.3, on démontre que la matrice

$$\begin{bmatrix} (M_{\beta_0}^{(1)}/1) & \mathbf{1}_R \\ \mathbf{1}_R^T & 0 \end{bmatrix}$$

est formellement inversible. On en déduit alors que

$$P = (M_{\beta_0}^{(1)}/1)^{-1} \frac{X_{\bullet}}{n} - \|\mathbf{1}_R^T\|_{(M_{\beta_0}^{(1)}/1)^{-1}}^{-2} \left[\frac{1}{n} \mathbf{1}_R^T (M_{\beta_0}^{(1)}/1)^{-1} X_{\bullet} - 1 \right] (M_{\beta_0}^{(1)}/1)^{-1} \mathbf{1}_R.$$

Les points (iii) puis (ii) du même lemme 5.3 permettent alors de déduire que

$$\begin{aligned} P &= (M_{\beta_0}^{(1)}/1)^{-1} \frac{X_{\bullet}}{n} \\ &= M_{\beta_0}^{-1} \frac{X_{\bullet}}{n} + (M_{\beta_0}^{(1)}/M_{\beta_0})^{-1} \times \left(\frac{\beta_0}{n} Z^T M_{\beta_0}^{-1} X_{\bullet} \right) \times M_{\beta_0}^{-1} \frac{X_{\bullet}}{n} \\ &= \left[1 + (M_{\beta_0}^{(1)}/M_{\beta_0})^{-1} (1 - (M_{\beta_0}^{(1)}/M_{\beta_0})) \right] M_{\beta_0}^{-1} \frac{X_{\bullet}}{n} \\ &= (M_{\beta_0}^{(1)}/M_{\beta_0})^{-1} M_{\beta_0}^{-1} \frac{X_{\bullet}}{n} \end{aligned}$$

D'où la solution $P(\beta_0)$ du théorème 5.1. Le paramètre β_0 est strictement positif et d'après le point (i) du lemme 5.1, $(M_{\beta_0}^{(1)}/M_{\beta_0})^{-1}$ est strictement positif, par conséquent chaque composante p_j du vecteur $P(\beta_0)$ est donc positive.

Algorithme de recherche de solutions aux équations de vraisemblance sous contraintes

Etape 0 : Initialisation

$\hat{\beta}_0^{(0)} = 0$ et on calcule le vecteur initial $\hat{P}^{(0)} = x_{\bullet}/n$;

Ratio = $[\sum_{j=1}^R x_{2j}] / [\sum_{j=1}^R (x_{1j} + x_{2j})]$;

Etape k ($k \geq 1$) :

Calculer :

$(M_{\hat{\beta}_0^{(k-1)}}^{(1)}/1)^{-1}$: (inverse du complément de Schur de 1 dans $M_{\hat{\beta}_0^{(k-1)}}^{(1)}$);

$\hat{P}^{(k)} = (M_{\hat{\beta}_0^{(k-1)}}^{(1)}/1)^{-1} \frac{X_{\bullet}}{n}$; $\hat{c}^{(k)} = \sum_{j=1}^R c_j \hat{p}_j^{(k)}$;

$\hat{\beta}_0^{(k)} = [\sum_{j=1}^R x_{2j}] / [\hat{c}^{(k)} \times \sum_{j=1}^R x_{1j}]$;

$\hat{RT}^{(k)} = [\hat{\beta}_0^{(k)} \hat{c}^{(k)}] / [1 + \hat{\beta}_0^{(k)} \hat{c}^{(k)}]$;

Si $\hat{RT}^{(k)} < Ratio$, alors remplacer $\hat{\beta}_0^{(k-1)}$ par $\hat{\beta}_0^{(k)}$ et aller à **Etape k**;

Sinon $\hat{\beta}_0 \in [\hat{\beta}_0^{(k-1)}, \hat{\beta}_0^{(k)}]$;

Calculer $\hat{\beta}_0$ par interpolation linéaire en utilisant les couples de valeurs

$(\hat{\beta}_0^{(k-1)}; \hat{RT}^{(k-1)})$, $(\hat{\beta}_0; Ratio)$, $(\hat{\beta}_0^{(k)}; \hat{RT}^{(k)})$,

et déterminer $\hat{P} = (M_{\hat{\beta}_0}^{(1)}/1)^{-1} \frac{x_{\bullet}}{n}$;

FIN.

Le calcul de $\hat{\beta}_0^{(k)}$ provient de la transformation de la première équation du système (4). D'autre part, on montre que cette première équation est équivalente à

$$\frac{\hat{\beta}_0 \sum_{m=1}^R c_m \hat{p}_m}{1 + \hat{\beta}_0 \sum_{m=1}^R c_m \hat{p}_m} = \frac{\sum_{m=1}^R X_{2m}}{\sum_{m=1}^R (X_{1m} + X_{2m})}.$$

On en déduit alors que

$$\hat{\beta}_0 = \frac{\sum_{m=1}^R X_{2m}}{\left[\sum_{m=1}^R c_m \hat{p}_m \right] \times \left[\sum_{m=1}^R X_{1m} \right]},$$

et par suite les expressions de $\hat{RT}^{(k)}$ et de *Ratio*. On compare alors les deux membres de cette dernière équation. Si le membre de gauche est inférieur à celui de droite alors on passe à l'étape $k+1$, sinon on conclut que $\hat{\beta}_0$ appartient à l'intervalle $[\hat{\beta}_0^{(k-1)}; \hat{\beta}_0^{(k)}]$. On estime alors $\hat{\beta}_0$ par interpolation linéaire et ensuite \hat{p}_j par utilisation de la deuxième équation de (5). En pratique, l'algorithme utilise les deux valeurs tests $\{0, 1\}$ pour les deux premières étapes d'estimation de β_0 . Si la valeur solution $\hat{\beta}_0 \in [0, 1]$ alors on partitionne cet intervalle en deux et on recherche (en utilisant l'algorithme) lequel des deux intervalles est susceptible de contenir la solution $\hat{\beta}_0$. Sinon, on initialise $\hat{\beta}_0$ à 1 et on réitère le processus avec un intervalle de borne inférieure 1.

Annexe B : Détermination de la matrice d'information sous contraintes

On considère la fonction

$$\mathcal{L}_1(X_1, X_2; \Theta, \lambda) = \mathcal{L}(X_1, X_2; \Theta) - \lambda \left(\sum_{j=1}^R p_j - 1 \right)$$

où λ , le multiplicateur de Lagrange associé à la contrainte $\sum_{j=1}^R p_j = 1$, est égal à $n / (1 + \beta_0 \sum_{j=1}^R c_j p_j)$. La matrice d'information relative à l'ensemble des paramètres et du multiplicateur de Lagrange est définie par :

$$\Gamma_{\Theta} = \begin{bmatrix} E\left(-\frac{\partial^2 \mathcal{L}_1}{\partial \Theta \partial \Theta}\right) & E\left(-\frac{\partial^2 \mathcal{L}_1}{\partial \Theta \partial \lambda}\right) \\ E\left(-\frac{\partial^2 \mathcal{L}_1}{\partial \Theta \partial \lambda}\right)^T & E\left(-\frac{\partial^2 \mathcal{L}_1}{\partial \lambda \partial \lambda}\right) \end{bmatrix}$$

avec $E\left(-\frac{\partial^2 \mathcal{L}_1}{\partial \Theta \partial \Theta}\right)$ est la matrice carrée de dimension $(1+R) \times (1+R)$ constituée de l'espérance mathématique des dérivées partielles secondes de \mathcal{L}_1 par rapport aux composantes du vecteur Θ , $E\left(-\frac{\partial^2 \mathcal{L}_1}{\partial \Theta \partial \lambda}\right)$ est le vecteur de dimension $1+R$ constitué de l'espérance mathématique des dérivées partielles secondes de \mathcal{L}_1 par rapport aux composantes du vecteur Θ et du multiplicateur de

Lagrange, et $E(-\frac{\partial^2 \mathcal{L}_1}{\partial \lambda \partial \lambda})$ est le scalaire réel obtenu à l'aide de l'espérance mathématique de la dérivée partielle seconde de \mathcal{L}_1 par rapport à λ . On montre que les dérivées partielles premières de \mathcal{L}_1 sont données par :

$$\begin{aligned} \frac{\partial \mathcal{L}_1}{\partial \beta_0} &= \frac{X_{2\bullet}}{\beta_0} - \frac{n \sum_{j=1}^R c_j p_j}{1 + \beta_0 \sum_{j=1}^R c_j p_j}; \\ \frac{\partial \mathcal{L}_1}{\partial p_j} &= \frac{X_{\bullet j}}{p_j} - \frac{n c_j \beta_0}{1 + \beta_0 \sum_{m=1}^R c_m p_m} - \lambda, \quad j = 1, 2, \dots, R; \\ \frac{\partial \mathcal{L}_1}{\partial \lambda} &= -(\sum_{j=1}^R p_j - 1). \end{aligned}$$

On en déduit alors les dérivées partielles secondes suivantes :

$$\frac{\partial^2 \mathcal{L}_1}{\partial \lambda \partial \lambda} = 0, \quad \frac{\partial^2 \mathcal{L}_1}{\partial \lambda \partial \beta_0} = 0, \quad \frac{\partial^2 \mathcal{L}_1}{\partial \lambda \partial p_j} = -1, \quad j = 1, 2, \dots, R;$$

$$\begin{aligned} \frac{\partial^2 \mathcal{L}_1}{\partial \beta_0 \partial p_j} &= -\frac{n c_j}{(1 + \beta_0 \sum_{m=1}^R c_m p_m)^2}, \quad j = 1, 2, \dots, R; \\ \frac{\partial^2 \mathcal{L}_1}{\partial \beta_0 \partial \beta_0} &= -\frac{X_{2\bullet}}{\beta_0^2} + \frac{n (\sum_{m=1}^R c_m p_m)^2}{(1 + \beta_0 \sum_{m=1}^R c_m p_m)^2}; \end{aligned}$$

et

$$\frac{\partial^2 \mathcal{L}_1}{\partial p_j \partial p_k} = \begin{cases} -\frac{X_{\bullet j}}{p_j^2} + \frac{n \beta_0^2 c_j^2}{(1 + \beta_0 \sum_{m=1}^R c_m p_m)^2}; & j = k; \\ \frac{n \beta_0^2 c_j c_k}{(1 + \beta_0 \sum_{m=1}^R c_m p_m)^2} & (j \neq k). \end{cases}$$

En utilisant la loi multinomiale de la paire (X_1, X_2) , on montre que $E(X_{\bullet j}) = [n p_j (1 + \beta_0 c_j)] / [1 + \beta_0 \sum_{m=1}^R c_m p_m]$, et $E(X_{2\bullet}) = [n \beta_0 (\sum_{m=1}^R c_m p_m)] / [1 + \beta_0 \sum_{m=1}^R c_m p_m]$. En substituant ces deux dernières expressions dans l'espérance mathématique de l'opposé des dérivées partielles secondes de \mathcal{L}_1 , on a les

résultats ci-après :

$$E\left(-\frac{\partial^2 \mathcal{L}_1}{\partial \beta_0 \partial p_j}\right) = \frac{n c_j}{(1 + \beta_0 \sum_{m=1}^R c_m p_m)^2}, j = 1, 2, \dots, R;$$

$$E\left(-\frac{\partial^2 \mathcal{L}_1}{\partial \beta_0 \partial \beta_0}\right) = \frac{n \left(\sum_{m=1}^R c_m p_m\right)}{\beta_0 (1 + \beta_0 \sum_{m=1}^R c_m p_m)^2};$$

et

$$E\left(-\frac{\partial^2 \mathcal{L}_1}{\partial p_j \partial p_k}\right) = \begin{cases} \frac{n}{1 + \beta_0 \sum_{m=1}^R c_m p_m} \left[\frac{1 + \beta_0 c_j}{p_j} - \frac{\beta_0^2 c_j^2}{1 + \beta_0 \sum_{m=1}^R c_m p_m} \right], & j = k; \\ -\frac{n \beta_0^2 c_j c_k}{(1 + \beta_0 \sum_{m=1}^R c_m p_m)^2} & (j \neq k). \end{cases}$$

En posant $\gamma = n / (1 + \beta_0 \sum_{m=1}^R c_m p_m)$, on obtient l'expression de Γ_Θ donnée au paragraphe 3.2.

Remerciements. — Nous remercions la direction du Centre d'Etude Technique de l'Équipement (CETE) Nord Picardie pour nous avoir autorisés à utiliser les données d'accident provenant de différents aménagements routiers. Nous tenons également à apporter notre profonde gratitude au Professeur Claude Langrand et aux deux arbitres anonymes pour leurs remarques et suggestions. La pertinence de leurs commentaires nous ont permis d'améliorer le fond et la forme de notre manuscrit.

Références

- AITCHISON J., SILVEY S.D. (1958). *Maximum Likelihood Estimation of parameters subject to restraints*, Annals of mathematical statistics, 29, 813-829.
- CROWDER M. (1984). *On the constrained maximum likelihood estimation with non i.i.d. observations*, Ann. Inst. Statist. Math. 36, A, 239-249.
- EFRON B., TIBSHIRANI R.J. (1993). *An introduction to the bootstrap* Chapman and Hall.
- HABER M., BROWN M.B. (1986). *Maximum Likelihood methods for log-linear models when expected frequencies are subject to linear constraints*, J.A.S.A. vol. 81, n°394, 477-482
- HAUER E. (1997). *Observational Before-After studies in road safety*, Pergamon
- MATTHEWS G.B., CROWTHER N.A.S. (1995). *A maximum likelihood estimation procedure when modelling in terms of constraints*, South African Statist. J. n°29, 29-50
- N'GUESSAN A. (2003). *Constrained covariance matrix estimation in road accident modelling with Schur complements*, C.R. Acad. Sci. Paris, Ser. I 337, 219-222.

- N'GUESSAN A., ESSAI A., LANGRAND C. (2001). *Estimation multidimensionnelle des contrôles et de l'effet moyen d'une mesure de sécurité routière*, Rev. Statistique Appliquée. XLIX(2), 83-100.
- N'GUESSAN A., LANGRAND C. (2005a). *A covariance components estimation procedure when modelling a road safety measure in terms of linear constraints*, Statistics, vol 39, n° 4, 303-314.
- N'GUESSAN A., LANGRAND C. (2005b). *A Schur complement approach for computing subcovariance matrices arising in a road safety measure modelling*, Journal of Computational and Applied Mathematics, 177, 331-345.
- N'GUESSAN A., BELLAVANCE F. (2005). *A confidence interval estimation problem using the Schur complement approach, with application*, C.R. Math. Rep. Acad. Sci. Canada Vol. 27, (3), pp. 84-91.
- N'GUESSAN A., ESSAI A., N'ZI M. (2006). *An estimation method of the average effect and the different accident risks when modelling a road safety measure : A simulation study*, Computational Statistics & Data Analysis, 51, 1260-1277.
- NICHOLSON A.J. (1986). *The randomness of accident count*, Accid. Anal. and Prev. vol n°18, pp. 193-198.
- NICHOLSON A.J., WONG Y. D. (1993). *Are accidents Poisson distributed? A statistical test*, Accid. Anal. and Prev. 23, vol n°26, pp. 609-622.
- OPPE S. (1979). *The randomness of accident count*, Accid. Anal. and Prev. vol n°11, pp. 101-115.
- OUELLETTE D.V., (1981). *Schur complements and statistics*, Linear Algebra an its applications, 36, 187-295.
- TANNER J.C, (1958). *A problem in the combination of accident frequencies*, Biometrika, 45, 331-342.
- ZHANG F. ed, (2005). *Schur complement and Its applications*, (Springer Verlag).