C. R. Acad. Sci. Paris, Ser. I 353 (2015) 1147–1152

Contents lists available at ScienceDirect

C. R. Acad. Sci. Paris, Ser. I

www.sciencedirect.com

Probability theory/Statistics

A note on the strong consistency of a constrained maximum likelihood estimator used in crash data modeling

À propos de la consistance forte d'un estimateur du maximum de vraisemblance sous contraintes utilisé dans la modélisation des données d'accidents

Issa Cherif Geraldo [a,b], Assi N'Guessan [b], Kossi Essona Gneyou [a,c]

[a] *Département de mathématiques et informatique, Université catholique de l'Afrique de l'Ouest, Unité universitaire du Togo (UCAO–UUT), 01 B.P. 1502 Lomé 01, Lomé, Togo*
[b] *Laboratoire Paul-Painlevé, UMR CNRS 8524, Université de Lille-1, 59655 Villeneuve d'Ascq cedex, France*
[c] *Département de mathématiques, Faculté des sciences, Université de Lomé, B.P. 1515 Lomé, Togo*

ARTICLE INFO

Article history:
Received 1 May 2015
Accepted after revision 28 September 2015
Available online 30 October 2015

Presented by the Editorial Board

ABSTRACT

In this note, we consider the Maximum Likelihood Estimator (MLE) of the vector parameter $\Phi = (\theta, \phi^{\mathrm{T}})^{\mathrm{T}}$ of dimension R $(R > 1)$ used in crash-data modeling where $\theta > 0$ and ϕ belongs to the simplex of order $R - 1$. We prove the strong consistency of this constrained estimator making capital out of the cyclic form between the components of the MLE.
© 2015 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

RÉSUMÉ

Dans cette note, nous considérons l'estimateur du maximum de vraisemblance (EMV) du vecteur paramètre $\Phi = (\theta, \phi^{\mathrm{T}})^{\mathrm{T}}$ de dimension R $(R > 1)$ utilisé dans la modélisation des données d'accidents où $\theta > 0$ et ϕ appartient au simplexe d'ordre $R - 1$. Nous démontrons la consistance forte de cet estimateur sous contraintes en exploitant la forme cyclique entre les composantes de cet estimateur.
© 2015 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

1. Introduction

Let X be a $\mathbb{R}^d$ valued random variable defined on a probability space $(\Omega, \mathcal{A}, \mathrm{P})$ with a probability density function depending on a vector parameter Φ. The Maximum Likelihood Estimator (MLE) $\hat{\Phi}_n$ of Φ can be obtained by solving the optimization problem

$$\hat{\Phi}_n = \arg\max_{\Phi \in S} L(\Phi)$$

where L is the log-likelihood function calculated on a sample of n i.i.d. observations of X and S is the parameter space (the set of all possible values of Φ).

E-mail addresses: cherifgera@gmail.com (I.C. Geraldo), assi.nguessan@polytech-lille.fr (A. N'Guessan), kossi_gneyou@yahoo.fr (K.E. Gneyou).

http://dx.doi.org/10.1016/j.crma.2015.09.025
1631-073X/© 2015 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

One of the most desired properties of the estimator $\hat{\Phi}_n$ is its consistency, i.e. its asymptotic convergence to the true value Φ^0. This property was studied in the literature by many authors (see, e.g., [2,6,7,14,15]). The main result on the strong consistency was established by Wald [15], who gave regularity conditions under which the MLE is strongly consistent. However, all these conditions may be hard to check in practice if the dimension of the parameter space is large and the probability density function (or the likelihood) takes some complex forms. Nevertheless, introducing modifications in Wald's work, some authors, among which [5,6,12,14], obtained useful results on the consistency under less restrictive conditions. Van der Vaart [14] established general consistency properties of M-estimators presenting the MLE as a special case of M-estimators. But it is still possible that the MLE is not consistent even when it exists, as shown by the examples given in [4].

The present work is motivated by our need to provide a proof of the strong convergence property of the MLE of Φ proposed in [10,11] for statistical analysis of accident data on an experimental site where observed accidents can be classified into R mutually exclusive categories, $R \in \mathbf{N}^*$. In their before–after study in order to assess the impact of a measure on the occurrence of accidents, N'Guessan et al. [10,11] considered a random vector $X = (X_{11}, \dots, X_{1R}, X_{21}, \dots, X_{2R})^{\mathrm{T}}$ whose components are positive non-zero discrete random variables such that X_{1r} (resp. X_{2r}), $r = 1, \dots, R$, represents the number of crashes of type r that have occurred in the "before" (resp. "after") period. This model also integrates a vector of known non-random components denoted by $C = (c_1, \dots, c_R)^{\mathrm{T}}$. It is assumed that X follows the multinomial distribution $X \sim \mathcal{M}(n; \pi_1(\Phi), \pi_2(\Phi))$ where n denotes a positive integer representing the total number of independent accidents in both before and after periods, that is $\sum_{t=1}^{2}\sum_{r=1}^{R} X_{\mathrm{tr}} = n$. Here $\pi_1(\Phi) = (\pi_{11}(\Phi), \dots, \pi_{1R}(\Phi))^{\mathrm{T}}$, $\pi_2(\Phi) = (\pi_{21}(\Phi), \dots, \pi_{2R}(\Phi))^{\mathrm{T}}$ with

$$\pi_{1r}(\Phi) = \frac{\phi_r}{1 + \theta \sum_{j=1}^{R} c_j \phi_j}, \quad \pi_{2r}(\Phi) = \frac{\theta\, c_r \phi_r}{1 + \theta \sum_{j=1}^{R} c_j \phi_j}, \qquad \forall r = 1, \dots, R \tag{1}$$

and $\sum_{t=1}^{2}\sum_{r=1}^{R} \pi_{\mathrm{tr}}(\Phi) = 1$. The random vector X has a probability density function depending on a multidimensional parameter $\Phi = (\theta, \phi^{\mathrm{T}})^{\mathrm{T}}$, where $\theta \in \mathbb{R}_+^*$ and $\phi = (\phi_1, \dots, \phi_R)^{\mathrm{T}}$ satisfies $\sum_{r=1}^{R} \phi_r = 1$ and belongs to the simplex of dimension $R-1$. The scalar θ represents the unknown average effect of the road safety measure, while each ϕ_r $(r = 1, 2, \dots, R)$ denotes the global accident risk of type r before and after the application of the road safety measure. The coefficients $c_1, \dots, c_R$ are given positive real numbers.

The existence of the constrained MLE $\hat{\Phi}_n$ of model (1) has been studied by [8] and an application is given in [11]. The numerical convergence properties of $\hat{\Phi}_n$ to the true values were recently studied by N'Guessan and Geraldo [10] using intensive simulation studies. They found that the MLE $\hat{\Phi}_n$ given by (2) converges numerically to the true value of the parameter whenever n tends to $+\infty$. We then make up their results by showing the strong convergence of $\hat{\Phi}_n$ given by (2) below.

So the aim of this note is to give a theoretical proof of the strong convergence of the estimator $\hat{\Phi}_n$ when n tends to infinity. The remainder of the note is organized as follows. In Section 2, we give some preliminary results. The main results on the strong convergence of the MLE in the crash control model are presented in Section 3. These main results are divided into three theorems. In Section 4, we present some concluding remarks.

2. Preliminary results

Throughout the paper, the subscript n is used to indicate that the estimators depend on the sample size n. It is proven in [9] that the log-likelihood associated with an observed data $x = (x_{11}, \dots, x_{1r}, x_{21}, \dots, x_{2r})$ satisfying $\sum_{t=1}^{2}\sum_{r=1}^{R} x_{\mathrm{tr}} = n$, is defined up to an additive constant by $\ell(\Phi) = \sum_{r=1}^{R}[\underline{x}_r \log(\phi_r) + x_{2r}\log(\theta) - \underline{x}_r \log(1 + \theta\sum_{m=1}^{R} c_m\phi_m)]$ where $\underline{x}_r = x_{1r} + x_{2r}$, $r = 1, \dots, R$. The MLE $\hat{\Phi}_n$ of Φ is then given by the following lemma.

Lemma 2.1. *(See [8].) The components $\hat{\theta}_n$ and $\hat{\phi}_n$ of the MLE $\hat{\Phi}_n$ satisfy*

$$\begin{cases} \hat{\theta}_n = \dfrac{\sum_{m=1}^{R} X_{2m}}{\left(\sum_{m=1}^{R} c_m \hat{\phi}_{n,m}\right) \times \left(\sum_{m=1}^{R} X_{1m}\right)} \\ \hat{\phi}_{n,r} = \dfrac{1}{1 - \frac{1}{n}\sum_{m=1}^{R} \frac{\hat{\theta}_n c_m \underline{X}_m}{1 + \hat{\theta}_n c_m}} \times \dfrac{\underline{X}_r}{n(1 + \hat{\theta}_n c_r)}, \qquad r = 1, 2, \dots, R \end{cases} \tag{2}$$

with $\underline{X}_r = X_{1r} + X_{2r}$, $r = 1, \dots, R$.

Proof. Introducing a Lagrange multiplier in order to cope with the linear constraint $\sum_{r=1}^{R} \phi_r = 1$, $\hat{\Phi}_n$ is obtained as a solution of the non-linear system:

$$\begin{cases} \displaystyle\sum_{r=1}^{R} \left[x_{2r} - \underline{x}_r \frac{\hat{\theta}_n \sum_{m=1}^{R} c_m \hat{\phi}_{n,m}}{1 + \hat{\theta}_n \sum_{m=1}^{R} c_m \hat{\phi}_{n,m}} \right] = 0 \\ \underline{x}_r - n \dfrac{\hat{\phi}_{n,r}(1 + c_r \hat{\theta}_n)}{1 + \hat{\theta}_n \sum_{m=1}^{R} c_m \hat{\phi}_{n,m}} = 0, \quad r = 1, 2, \dots, R. \end{cases} \tag{3}$$

The first line of System (2) follows from the first line of (3). The expression of $\hat{\phi}_n$ is obtained by transforming the second line of (3) into a linear system whose unique vector solution is $\hat{\phi}_n$. For further details, refer to [8,11] and the references therein. □

Let us recall some important lemmas that will be the key for establishing our strong convergence results. The first lemma is provided by the continuous mapping theorem of [14, p. 7]. The second is due to the strong law of large numbers. The third lemma states conditions under which the convergence of a sequence of injective functions (f_n) implies that of their inverses (f_n^{-1}) [1, Theorem 2].

Lemma 2.2. *(See [14].) Let $g:\mathbb{R}^k \to \mathbb{R}^m$ be continuous at every point of a Borel set A such that $\mathrm{P}(X \in A) = 1$. If X_n converges almost surely (a.s.) to X, then $g(X_n)$ converges almost surely to $g(X)$.*

Lemma 2.3. *If the random vector $X = (X_{11}, \dots, X_{1R}, X_{21}, \dots, X_{2R})$ has the multinomial distribution $\mathcal{M}(n;\pi)$ with $\pi = (\pi_{11}, \dots, \pi_{1R}, \pi_{21}, \dots, \pi_{2R})$ then, as $n \to +\infty$: $\frac{1}{n}X \xrightarrow{a.s.} (\pi_{11}, \dots, \pi_{1R}, \pi_{21}, \dots, \pi_{2R})$.*

Proof. The vector X can be thought of as the sum of n independent multinomial vectors $Y_1, \dots, Y_n$ with parameters 1 and π. Then, the mathematical expectation of Y_i is $E(Y_i) = \pi$. By the strong law of large numbers, the random sequence $n^{-1}X$ converges almost surely to π. □

The following lemmas are related to the uniform convergence of sequences of functions on a metric space.

Lemma 2.4. *(See [1].) If (f_n) is a sequence of injection mappings on a metric space E, taking values in a locally compact metric space G and converging uniformly to f on E and if f^{-1} is a continuous mapping on $G_1 \subset G$, then f_n^{-1} converges uniformly to f^{-1} on every compact set contained in $\mathrm{int}(G_1) \cap (\cap_n f_n(E))$ where $\mathrm{int}(G_1)$ denotes the interior of G_1.*

Lemma 2.5. *(See [13].) Let f_n be a sequence of continuous functions on a set D. If f_n converges uniformly to f, then $f_n(u_n)$ converges to $f(u)$ for all sequences u_n in D convergent to $u \in D$.*

3. Main results

Theorem 3.1. *As n tends to $+\infty$, the random variable $\hat{\theta}_n$ converges a.s. to θ^0 if and only if the random vector $\hat{\phi}_n$ converges a.s. to ϕ^0.*

Proof. We know that $\hat{\phi}_n = (\hat{\phi}_{n,1}, \dots, \hat{\phi}_{n,R}) \in \mathbb{R}^R$ converges a.s. to $\phi^0 = (\phi_1^0, \dots, \phi_R^0) \in \mathbb{R}^R$ if and only if, for all $r = 1, \dots, R$, $\hat{\phi}_{n,r} \to \phi_r^0$ a.s. Thus, it is sufficient to prove that for all $r = 1, \dots, R$, $\hat{\theta}_n \longrightarrow \theta^0$ *a.s.* implies that $\hat{\phi}_{n,r} \longrightarrow \phi_r^0$ *a.s.* Now let us suppose that $\hat{\theta}_n \to \theta^0$ a.s. Observing that $\sum_{m=1}^R \underline{X}_m = n$, we get

$$\hat{\phi}_{n,r} = \frac{(X_{1r} + X_{2r})/(1 + \hat{\theta}_n c_r)}{\sum_{m=1}^R (X_{1m} + X_{2m})/(1 + \hat{\theta}_n c_m)}. \tag{4}$$

Moreover, we can write $\hat{\phi}_{n,r} = g_r(\frac{X_{11}}{n}, \dots, \frac{X_{1R}}{n}, \frac{X_{21}}{n}, \dots, \frac{X_{2R}}{n}, \hat{\theta}_n)$ where g_r is the continuous function from $\mathbb{R}^{2R+1}$ to $\mathbb{R}$ defined by $(b_1, \dots, b_R, a_1, \dots, a_R, \theta) \mapsto \frac{(b_r + a_r)/(1+\theta c_r)}{\sum_{m=1}^R (b_m + a_m)/(1+\theta c_m)}$. Using Lemma 2.3, we have almost surely

$$\left(\frac{X_{11}}{n}, \dots, \frac{X_{1R}}{n}, \frac{X_{21}}{n}, \dots, \frac{X_{2R}}{n}, \hat{\theta}_n\right) \to (\pi_{11}^0, \dots, \pi_{1R}^0, \pi_{21}^0, \dots, \pi_{2R}^0, \theta^0)$$

as $n \to \infty$. Applying the continuous mapping theorem (Lemma 2.2) and relations (1), we get as $n \to +\infty$,

$$\hat{\phi}_{n,r} \quad \to \quad \frac{\left(\pi_{1r}^0 + \pi_{2r}^0\right)/(1 + \theta^0 c_r)}{\sum_{m=1}^R \left(\pi_{1m}^0 + \pi_{2m}^0\right)/(1 + \theta^0 c_m)} = \phi_r^0 \quad \text{a.s.}$$

This proves that $\hat{\phi}_{n,r}$ converges to ϕ_r^0 a.s.

Now let us assume that $\hat{\phi}_n \longrightarrow \phi^0$ a.s. or equivalently $\hat{\phi}_{n,r} \longrightarrow \phi_r^0$ a.s. for all $r = 1, \dots, R$. From Lemma 2.1, we get

$$\hat{\theta}_n = \frac{\sum_{m=1}^R (X_{2m}/n)}{\sum_{m=1}^R (X_{1m}/n)} \times \frac{1}{\sum_{m=1}^R c_m \hat{\phi}_{n,m}} = g\left(\frac{X_{11}}{n}, \dots, \frac{X_{1R}}{n}, \frac{X_{21}}{n}, \dots, \frac{X_{2R}}{n}, \hat{\phi}_{n,1}, \dots, \hat{\phi}_{n,R}\right)$$

where g is the continuous function defined from $\mathbb{R}^{3R}$ to $\mathbb{R}$ by

$$(b_1, \dots, b_R, a_1, \dots, a_R, \phi_1, \dots, \phi_R) \quad \mapsto \quad \frac{\sum_{m=1}^R a_m}{\sum_{m=1}^R b_m} \times \frac{1}{\sum_{m=1}^R c_m \phi_m}.$$

We again apply Lemmas 2.2 and 2.3 and get

$$\hat{\theta}_n \xrightarrow[n\to+\infty]{a.s.} g(\pi_{11}^0,\dots\pi_{1R}^0,\pi_{21}^0,\dots,\pi_{2R}^0,\phi_1^0,\dots,\phi_R^0)=\theta^0. \qquad \square$$

Theorem 3.1 shows that the almost sure convergence of $\hat{\phi}_n$ to ϕ^0 is equivalent to the almost sure convergence of $\hat{\theta}_n$ to θ^0. To prove that the MLE $\hat{\Phi}_n=(\hat{\theta}_n,\hat{\phi}_n^{\mathrm{T}})^{\mathrm{T}}$ converges almost surely, it is then sufficient by Theorem 3.1 to prove for example that $\hat{\theta}_n$ converges almost surely to θ^0. With that in mind, we first prove that the a.s. limit of $\hat{\theta}_n$ exists and then show that this a.s. limit is equal to θ^0.

The following result shows the almost sure convergence of $\hat{\theta}_n$.

Theorem 3.2. *There exists a constant $\mu>0$ and a subset $N\subset\Omega$ such that $\mathrm{P}(N)=0$ and*

$$\forall\omega\in\Omega\setminus N,\quad \lim_{n\to\infty}\hat{\theta}_n(\omega)=\mu. \tag{5}$$

Proof. Set $\varphi_n(u)=\sum_{m=1}^R \frac{\underline{X}_m/n}{1+u\,c_m}$, $u\in]0,+\infty[$ and for all $t=1,2$, denote $X_{t+}=\sum_{m=1}^R X_{tm}$.

We first show that for all $\omega\in\Omega$, the real valued function $\varphi_{\omega,n}(u)=\varphi_n(u)(\omega)=\sum_{m=1}^{R}\dfrac{\underline{X}_m(\omega)/n}{1+uc_m}$ is a continuous bijective mapping from $]0,+\infty[$ to $]0,1[$ and that $\hat{\theta}_n(\omega)=\varphi_{\omega,n}^{-1}(X_{1+}(\omega)/n)$.

By Equation (4), we have the relationship $\hat{\phi}_{n,r}=\frac{\underline{X}_r/(1+\hat{\theta}_n c_r)}{\sum_{m=1}^R \underline{X}_m/(1+\hat{\theta}_n c_m)}$ which enables to write

$$\sum_{r=1}^R c_r\hat{\phi}_{n,r}=\frac{\sum_{r=1}^R\left(c_r\underline{X}_r/(1+\hat{\theta}_n c_r)\right)}{\sum_{m=1}^R\left(\underline{X}_m/(1+\hat{\theta}_n c_m)\right)}.$$

By the first line of System (2) in Lemma 2.1, we have

$$\hat{\theta}_n=\frac{X_{2+}}{X_{1+}}\frac{1}{\sum_{m=1}^R c_m\hat{\phi}_{n,m}}=\frac{X_{2+}}{X_{1+}}\frac{\sum_{m=1}^R\left(\underline{X}_m/(1+\hat{\theta}_n c_m)\right)}{\sum_{r=1}^R\left(c_r\underline{X}_r/(1+\hat{\theta}_n c_r)\right)}.$$

This is equivalent to

$$\frac{X_{2+}}{X_{1+}}\sum_{m=1}^R\frac{\underline{X}_m}{1+\hat{\theta}_n c_m}=\sum_{m=1}^R\frac{\hat{\theta}_n c_m\underline{X}_m}{1+\hat{\theta}_n c_m}.$$

We then deduce that $\sum_{m=1}^R\frac{\underline{X}_m}{1+\hat{\theta}_n c_m}=X_{1+}$. Divide the last equality by the sample size n and get

$$\sum_{m=1}^R\frac{\underline{X}_m(\omega)/n}{1+\hat{\theta}_n(\omega)\,c_m}=\frac{X_{1+}(\omega)}{n},\quad \forall\omega\in\Omega. \tag{6}$$

It is obvious that the random real function $\varphi_{\omega,n}(u)$ has a strictly negative derivative with respect to u and satisfies for all $\omega\in\Omega$:

$$1=\lim_{u\to0}\varphi_{\omega,n}(u)=\sum_{m=1}^R\underline{X}_m(\omega)/n \qquad 0=\lim_{u\to+\infty}\varphi_{\omega,n}(u).$$

Hence $\varphi_{\omega,n}(u)$ is a continuous and bijective mapping from $]0,+\infty[$ to $]0,1[$ and since $X_{1+}(w)/n\in\,]0,1[$, Equation (6) yields $\hat{\theta}_n(\omega)=\varphi_{\omega,n}^{-1}(X_{1+}(\omega)/n)$.

Let us now prove that there exists a subset $N\subset\Omega$ with $\mathrm{P}(N)=0$ such that for all $\omega\in\Omega\setminus N$, the sequence of real functions $\varphi_{\omega,n}(u)$ converges uniformly to some function $\varphi(u)$ on $]0,+\infty[$. The almost sure convergence of the statistic $\varphi_n(u)=\sum_{m=1}^R\dfrac{\underline{X}_m/n}{1+uc_m}$ to $\varphi(u)$ will then follow.

For all $m=1,\dots,R$, write $\frac{\underline{X}_m}{n}=g_m(\frac{X_{11}}{n},\dots,\frac{X_{1R}}{n},\frac{X_{21}}{n},\dots,\frac{X_{2R}}{n})$ where g_m is the continuous mapping defined from $\mathbb{R}^{2R}$ to $\mathbb{R}$ by $g_m(b_1,\dots,b_R,a_1,\dots,a_R)=b_m+a_m$. Applying Lemmas 2.2 and 2.3, we get

$$\frac{\underline{X}_m}{n}\xrightarrow{a.s.}\alpha_m^0=g_m(\pi_{11}^0,\dots,\pi_{1R}^0,\pi_{21}^0,\dots,\pi_{2R}^0)=\frac{(1+\theta^0c_m)\phi_m^0}{1+\theta^0\sum_{k=1}^R c_k\phi_k^0}.$$

Equivalently [3, p. 68], there exists a null set N_m such that

$$\forall \omega \in \Omega \setminus N_m, \quad \lim_{n\to\infty} \frac{\underline{X}_m(\omega)}{n} = \alpha_m^0. \tag{7}$$

The set $E_1 = \cup_{m=1}^R N_m$ satisfies $\mathrm{P}(E_1) = 0$ and

$$\forall \omega \in \Omega \setminus E_1, \quad \lim_{n\to\infty} \varphi_{\omega,n}(u) = \lim_{n\to\infty} \sum_{m=1}^R \frac{\underline{X}_m(\omega)/n}{1+u\,c_m} = \sum_{m=1}^R \frac{\alpha_m^0}{1+u\,c_m} = \varphi(u). \tag{8}$$

Thus we have proved that for all $\omega \in \Omega \setminus E_1$, the sequence of functions $\varphi_{\omega,n}$ converges simply to φ on $]0,+\infty[$. Moreover,

$$\begin{aligned}\sup_{u\in]0,+\infty[} |\varphi_{\omega,n}(u) - \varphi(u)| &= \sup_{u\in]0,+\infty[} \left| \sum_{m=1}^R \frac{\underline{X}_m(\omega)/n}{1+u\,c_m} - \sum_{m=1}^R \frac{\alpha_m^0}{1+u\,c_m} \right| \leqslant \sup_{u\in]0;+\infty[} \sum_{m=1}^R \left| \frac{\underline{X}_m(\omega)/n - \alpha_m^0}{1+u\,c_m} \right| \\ &= \sup_{u\in]0,+\infty[} \sum_{m=1}^R \frac{|\underline{X}_m(\omega)/n - \alpha_m^0|}{1+u\,c_m} \leqslant \sum_{m=1}^R |\underline{X}_m(\omega)/n - \alpha_m^0|\end{aligned}$$

because $\forall u \in]0,+\infty[$, $\forall c_m > 0$, $\frac{1}{1+u\,c_m} \leqslant 1$. It follows by (7) that

$$\sup_{u\in]0,+\infty[} |\varphi_{\omega,n}(u) - \varphi(u)| \leqslant \sum_{m=1}^R |\underline{X}_m(\omega)/n - \alpha_m^0| \longrightarrow 0 \text{ as } n \to +\infty.$$

This proves the uniform convergence of the sequence $\varphi_{\omega,n}$ to φ on $]0,+\infty[$. □

Remark 1. In summary, we have proved that $\varphi_{\omega,n}$ is a sequence of bijective functions taking values in $]0,1[$ that is locally compact. Moreover, $\varphi_{\omega,n}$ converges uniformly to φ and φ^{-1} is continuous (as the inverse of a non-zero continuous function). Thus the conditions of Lemma 2.4 are satisfied and hence the sequence $\varphi_{\omega,n}^{-1}$ converges uniformly to φ^{-1}. Moreover, the sequence X_{1+}/n satisfies

$$\frac{X_{1+}}{n} = \tilde{g}\left(\frac{X_{11}}{n}, \ldots, \frac{X_{1R}}{n}, \frac{X_{21}}{n}, \ldots, \frac{X_{2R}}{n}\right)$$

where $\tilde{g}$ is the continuous mapping defined from $\mathbb{R}^{2R}$ to $\mathbb{R}$ by $\tilde{g}(b_1, \ldots, b_R, a_1, \ldots, a_R) = \sum_{m=1}^R b_m$. Apply again Lemmas 2.2 and 2.3 and get

$$\frac{X_{1+}}{n} \xrightarrow{a.s.} \tilde{g}(\pi_{11}^0, \ldots, \pi_{1R}^0, \pi_{21}^0, \ldots, \pi_{2R}^0) = \frac{1}{1+\theta^0 \sum_{k=1}^R c_k \phi_k^0}.$$

That is, there exists a null set E_2 such that

$$\forall \omega \in \Omega \setminus E_2, \lim_{n\to\infty} \frac{X_{1+}(\omega)}{n} = \gamma^0 = \frac{1}{1+\theta^0 \sum_{k=1}^R c_k \phi_k^0}.$$

The set $N = E_1 \cup E_2$ satisfies $\mathrm{P}(N) = 0$ and for all $\omega \in \Omega \setminus N$, the sequence $X_{1+}(\omega)/n$ is convergent and hence is bounded. That is, there exists a compact set $D \subset]0,1[$ such that $X_{1+}(\omega)/n \in D, \forall n > 0$. Moreover, since

$$\hat{\theta}_n(\omega) = \varphi_{\omega,n}^{-1}\left(\frac{X_{1+}(\omega)}{n}\right)$$

and $\varphi_{\omega,n}^{-1}$ converges uniformly to φ, we apply Lemma 2.5 to conclude that

$$\forall \omega \in \Omega \setminus N, \; \hat{\theta}_n(\omega) \longrightarrow \mu = \varphi^{-1}(\gamma^0) \text{ as } n \to \infty \tag{9}$$

where γ^0 is given above. This ends the proof of Theorem 3.2.

Theorem 3.3. *Set $\beta^0 = \sum_{r=1}^R c_r \phi_r^0$. Let F_{θ^0} be the function from $\mathbb{R}_+$ to $\mathbb{R}$ defined by:*

$$F_{\theta^0}(u) = u\left(\sum_{m=1}^R \frac{c_m(1+\theta^0 c_m)\phi_m^0}{1+uc_m}\right) - \theta^0\beta^0\left(\sum_{m=1}^R \frac{(1+\theta^0 c_m)\phi_m^0}{1+uc_m}\right).$$

Then,

i) the function F_{θ^0} has θ^0 as unique root on $\mathbb{R}_+$.
ii) the almost sure limit μ of $\hat{\theta}_n$ is equal to the unique root θ^0 of F_{θ^0} on $\mathbb{R}_+$.

Proof. i) We have $F_{\theta^0}(\theta^0) = 0$ and $F'_{\theta^0}(u) > 0$. So F_{θ^0} is strictly monotone and satisfies

$$\lim_{u\to 0} F_{\theta^0}(u) < 0 \quad \text{and} \quad \lim_{u\to +\infty} F_{\theta^0}(u) > 0.$$

So we get the first assertion of the theorem.

ii) Let us assume that as $n \to +\infty$, $\hat{\theta}_n \to \mu$ a.s. Let us then divide the numerator and the denominator of $\hat{\theta}_n$ given by Lemma 2.1 by n^2 and get

$$\hat{\theta}_n = \frac{\sum_{m=1}^{R}(X_{2m}/n) \times \sum_{m=1}^{R} \frac{X_m/n}{1+\hat{\theta}_n c_m}}{\sum_{m=1}^{R}(X_{1m}/n) \times \sum_{m=1}^{R} \frac{c_m X_m/n}{1+\hat{\theta}_n c_m}}. \tag{10}$$

By Lemma 2.3,

$$\frac{\sum_{m=1}^{R} X_{2m}/n}{\sum_{m=1}^{R} X_{1m}/n} \xrightarrow[n\to+\infty]{a.s.} \frac{\sum_{m=1}^{R} \pi_{2m}^0}{\sum_{m=1}^{R} \pi_{1m}^0} = \theta^0\beta^0$$

and

$$\frac{X_r/n}{1+\hat{\theta}_n c_r} = \frac{X_{1r}/n + X_{2r}/n}{1+\hat{\theta}_n c_r} \xrightarrow[n\to+\infty]{a.s.} \frac{\pi_{1r}^0 + \pi_{2r}^0}{1+\mu c_r} = \frac{(1+\theta^0 c_r)\phi_r^0}{(1+\theta^0\sum_{m=1}^{R} c_m\phi_m^0)(1+\mu c_r)}.$$

Thus, as $n \to +\infty$, the first and the second hands of equation (10) yield, almost surely,

$$\mu = \theta^0\beta^0 \times \left(\sum_{r=1}^{R} \frac{(1+\theta^0 c_r)\phi_r^0}{(1+\mu c_r)}\right) \Bigg/ \left(\sum_{r=1}^{R} \frac{c_r(1+\theta^0 c_r)\phi_r^0}{(1+\mu c_r)}\right).$$

That is,

$$\mu \sum_{r=1}^{R} \frac{c_r(1+\theta^0 c_r)\phi_r^0}{(1+\mu c_r)} = \theta^0\beta^0 \times \sum_{r=1}^{R} \frac{(1+\theta^0 c_r)\phi_r^0}{(1+\mu c_r)},$$

which means that $F_{\theta^0}(\mu) = 0$ and hence $\mu = \theta^0$ by i). This completes the proof of Theorem 3.3. □

Theorem 3.4. *The MLE $\hat{\Phi}_n = (\hat{\theta}_n, \hat{\phi}_n)^{\mathrm{T}}$ converges a.s. to $\Phi^0 = (\theta^0, \phi^0)^{\mathrm{T}}$ with $\phi^0 = (\phi_1^0, \ldots, \phi_R^0)^{\mathrm{T}}$.*

Proof. This is a consequence of Theorem 3.1 and Theorem 3.3. Indeed, $\hat{\theta}_n$ converges a.s. to θ^0 and since by Theorem 3.1, the consistency of $\hat{\theta}_n$ is equivalent to that of $\hat{\phi}_n$, then $\hat{\phi}_n$ converges also almost surely to ϕ^0. Thus the vector $\hat{\Phi}_n = (\hat{\theta}_n, \hat{\phi}_n)^{\mathrm{T}}$ converges a.s. to the vector $\Phi^0 = (\theta^0, \phi^0)^{\mathrm{T}}$. □

4. Concluding remarks

We study the asymptotic strong consistency of a constrained maximum likelihood estimator of a vector parameter when a road safety measure has been applied to a target site. We intend to generalize our results to the multidimensional estimator proposed in [9] when we deal with the estimation of the effect of a road-safety measure applied on different target sites. Each target site counts R ($R > 1$) mutually exclusive accidents types and is linked to a specific control area where the measure is not directly applied.

References

[1] E. Barvínek, I. Daler, J. Francu, Convergence of sequences of inverse functions, Arch. Math. 27 (3–4) (1991) 201–204.
[2] D. Chafai, D. Concordet, On the strong consistency of asymptotic *M*-estimators, J. Stat. Plan. Inference 137 (2007) 2774–2783.
[3] K.L. Chung, A Course in Probability Theory, Academic Press, 2001.
[4] T.S. Ferguson, An inconsistent maximum likelihood estimate, J. Am. Stat. Assoc. 77 (380) (1982) 831–834.
[5] S. Fiorin, The strong consistency for maximum likelihood estimates: a proof not based on the likelihood ratio, C. R. Acad. Sci. Paris, Ser. I 331 (2000) 721–726.
[6] S. Kourouklis, On the strong consistency of a solution to the likelihood equation, Stat. Probab. Lett. 5 (1987) 23–27.
[7] J.F. Monahan, Numerical Methods of Statistics, 2nd edition, Cambridge University Press, 2011.
[8] A. N'Guessan, Analytical existence of solutions to a system of non-linear equations with application, J. Comput. Appl. Math. 234 (2010) 297–304.
[9] A. N'Guessan, A. Essai, C. Langrand, Estimation multidimensionnelle des contrôles et de l'effet moyen d'une mesure de sécurité routière, Rev. Stat. Appl. 49 (2) (2001) 85–102.
[10] A. N'Guessan, I.C. Geraldo, A cyclic algorithm for maximum likelihood estimation using Schur complement, Numer. Linear Algebra Appl. (2015), http://dx.doi.org/10.1002/nla.1999.
[11] A. N'Guessan, M. Truffier, Impact d'un aménagement de sécurité routière sur la gravité des accidents de la route, J. Soc. Fr. Stat. 149 (3) (2008) 23–41.
[12] B. Seo, B.G. Lindsay, Nearly universal consistency of maximum likelihood in discrete models, Stat. Probab. Lett. 83 (2013) 1699–1702.
[13] R.S. Strichartz, The Way of Analysis, Jones and Bartlett Books in Mathematics, Jones and Bartlett Publishers, 2000.
[14] A.W. Van der Vaart, Asymptotic Statistics, Cambridge University Press, 1998.
[15] A. Wald, Note on the consistency of the maximum likelihood estimate, Ann. Math. Stat. 20 (4) (1949) 595–601.