Gulf Journal of Mathematics
Vol 14, Issue 2 (2023) 40-53

ISSN: 2309-4966
https://doi.org/10.56947/gjom.v14i2.1177

EXISTENCE, UNIQUENESS AND STRONG CONSISTENCY OF THE MAXIMUM LIKELIHOOD ESTIMATOR FOR A MODEL OF ACCIDENTS FREQUENCIES

ISSA CHERIF GERALDO*

Abstract. The aim of this paper is to prove the existence, uniqueness and strong consistency (i.e. almost sure convergence to the true unknown value) of the maximum likelihood estimator (MLE) of the vector parameter for a statistical model used in statistics applied to road safety. In the general case, the strong consistency of the MLE may be established by using the well-known result by Abraham Wald (in 1949) or its variants under a set of conditions. However, for the model considered in this paper, all these conditions are very difficult to verify because of the great dimension of the parameter space and the rather complex expression of the log-likelihood function. To circumvent these difficulties, we first demonstrate that the MLE exists and is unique afterwards we demonstrate the strong consistency of the MLE using the properties of the model and some theorems of mathematical analysis.

1. Introduction

In parametric statistics, most of the problems require the estimation of a parameter vector $\boldsymbol{\theta} \in \Theta$, where $\Theta \subset \mathbb{R}^d$ and d is the number of components of $\boldsymbol{\theta}$. The parameter vector may be unconstrained ($\Theta = \mathbb{R}^d$) or constrained ($\Theta \subsetneq \mathbb{R}^d$) and the maximum likelihood estimator $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$, if it exists, is obtained as

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta} \in \Theta}{\operatorname{argmax}}\, \ell(\boldsymbol{\theta}), \tag{1.1}$$

where $\ell(\boldsymbol{\theta})$ is the log-likelihood function. Although the maximum likelihood method is generally efficient in most cases, some statistical models may suffer from the non-existence of the MLE [4, 8, 9, 18] or its non-uniqueness [1, 15, 16]. When existence and uniqueness of the MLE are guaranteed, another desirable property of the MLE is its strong consistency that is its almost sure (a.s.) convergence (convergence with probability equal to one when the sample size tends to $+\infty$) to the true unknown parameter vector $\boldsymbol{\theta}^0$ of the model from which the observed data come.

Date: Received: May 19, 2022; Accepted: Oct 20, 2022.

* Corresponding author.

2010 *Mathematics Subject Classification.* Primary 62F10, 62H12; Secondary 62F30.

Key words and phrases. Maximum likelihood, Multinomial distribution, Strong consistency, Road safety measure.

To our knowledge, the strong consistency of the MLE was established by Wald [20] under some regularity conditions. Although the latter claimed that the underlying assumptions are easy to check in practice, a few of them involve the calculation of upper bounds which are not obvious to obtain. Although Wald's proof of consistency has opened the way for more studies on the strong consistency of the MLE (see for example [3, 5, 7]), it remains a central result in Statistics. When the MLE is available in closed-form, the study of its strong consistency could certainly be done directly by using its closed-form expression and the properties of the underlying statistical distribution.

In this paper, we are interested in a statistical model comprising a $s-$tuple ($s > 0$) of independent discrete random vectors $(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_s)$ following multinomial distributions whose parameters are linked. This model was proposed by [12] for estimating the average effect of a road safety measure implemented on s experimental sites. It has a parameter vector $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}^\mathsf{T})^\mathsf{T}$ where $\alpha > 0$ is the average effect (in the multiplicative sense) of the measure, $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\mathsf{T}, \ldots, \boldsymbol{\beta}_s^\mathsf{T})^\mathsf{T}$ where for each $k = 1, \ldots, s$, $\boldsymbol{\beta}_k$ is a probability vector associated with the site k.

The aim of this paper is to prove the existence, uniqueness and strong consistency of the MLE $\hat{\boldsymbol{\theta}} = (\hat{\alpha}, \hat{\boldsymbol{\beta}}^\mathsf{T})^\mathsf{T}$ of the parameter vector of the model. In the case $s = 1$, the existence and uniqueness of the MLE have been proved by [11] using the likelihood equations to obtain closed-form expressions of $\hat{\alpha}$ and $\hat{\boldsymbol{\beta}}$ and the strong consistency has been established by [6] by exploiting these closed-form expressions. In this paper, we extend these results to the general case $s \geqslant 1$. We get closed-form expressions of the components of $\hat{\boldsymbol{\beta}}$ as functions of $\hat{\alpha}$ but, unlike the case $s = 1$, it seems impossible to get closed-form expression of $\hat{\alpha}$ (this latter being the solution to a non-linear implicit equation). To circumvent this difficulty, we use the properties of the model and some theorems of mathematical analysis.

The remainder of this paper is organized as follows. Section 2 gives more details on the statistical model, the assumptions and notations used in this paper. The first main contribution of this paper (the proof of the existence and the uniqueness of the MLE) is given in Section 3 afterwards the second main contribution (the proof of strong consistency) is given in Section 4. The last section of the paper is dedicated to the concluding remarks.

2. Statistical model, assumptions and notations

2.1. Statistical model and assumptions. Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_s$ be s vectors such that for all $k = 1, \ldots, s$, $\boldsymbol{X}_k = (X_{11k}, \ldots, X_{1rk}, X_{21k}, \ldots, X_{2rk})^\mathsf{T}$ where r is the number of accident severity levels, X_{1jk} (respectively X_{2jk}) is the random number of accidents of severity level j recorded on site k before (respectively after) the implementation of the measure. The model of [12] is built under the following assumptions:

(A1): For all $k = 1, \ldots, s$, the vector $\boldsymbol{X}_k$ has the multinomial distribution

$$\boldsymbol{X}_k \rightsquigarrow \mathscr{M}(n_k, \boldsymbol{\pi}_k(\boldsymbol{\theta})), \tag{2.1}$$

where

(A2): $\boldsymbol{\pi}_k(\boldsymbol{\theta}) = (\pi_{11k}(\boldsymbol{\theta}), \ldots, \pi_{1rk}(\boldsymbol{\theta}), \pi_{21k}(\boldsymbol{\theta}), \ldots, \pi_{2rk}(\boldsymbol{\theta}))^\mathsf{T}$,

$$\pi_{1jk}(\boldsymbol{\theta}) = \frac{\beta_{jk}}{1+\alpha\langle \boldsymbol{z}_k, \boldsymbol{\beta}_k\rangle}, \quad \pi_{2jk}(\boldsymbol{\theta}) = \frac{\alpha z_{jk}\beta_{jk}}{1+\alpha\langle \boldsymbol{z}_k, \boldsymbol{\beta}_k\rangle}, \quad j = 1, \ldots, r, \tag{2.2}$$

$\boldsymbol{z}_k = (z_{1k}, \ldots, z_{rk})^\mathsf{T}$ is a vector of positive coefficients (see [12] for more details) and $\langle \boldsymbol{z}_k, \boldsymbol{\beta}_k\rangle = \sum_{j'=1}^r z_{j'k}\beta_{j'k}$.

Let $\mathbb{S}_{r-1} = \left\{(p_1, \ldots, p_r)^\mathsf{T} \in [0,1]^r, \ \sum_{j=1}^r p_j = 1\right\}$. The statistical model thus defined has a parameter vector satisfying the following assumption:

(A3): $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}^\mathsf{T})^\mathsf{T}$ where $\alpha > 0$, $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\mathsf{T}, \ldots, \boldsymbol{\beta}_s^\mathsf{T})^\mathsf{T}$ and for all $k = 1, \ldots, s$, $\boldsymbol{\beta}_k = (\beta_{1k}, \ldots, \beta_{rk})^\mathsf{T} \in \mathbb{S}_{r-1}$.

The parameter of interest α represents the average effect (in the multiplicative sense) of the measure. If $\alpha < 1$, then it could be concluded that the measure enabled to reduce the number of accidents on the different experimental sites. The sr secondary parameters β_{jk} ($k = 1, \ldots, s$, $j = 1, \ldots, r$) respectively represent the probabilities that an accident occurring on site k before the implementation of the measure is of severity level j.

2.2. **Notations.** Let $k = 1, \ldots, s$ and $j = 1, \ldots, r$. In the remainder of the paper, we adopt the following notations: $X_{\bullet jk} = X_{1jk} + X_{2jk}$, $X_{1\bullet k} = \sum_{j=1}^r X_{1jk}$, $X_{2\bullet k} = \sum_{j=1}^r X_{2jk}$, $X_{1\bullet\bullet} = \sum_{k=1}^s \sum_{j=1}^r X_{1jk}$ and $X_{2\bullet\bullet} = \sum_{k=1}^s \sum_{j=1}^r X_{2jk}$. Almost sure (a.s.) convergence is represented by $\xrightarrow{a.s.}$.

3. Existence and uniqueness of the maximum likelihood estimator

As is customary in the search for the expression of the maximum likelihood estimator (MLE), we will deliberately confuse in this section, the random variables and their realizations. Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_s$ be realizations of (2.1). It is proved in [12] that the log-likelihood is defined up to an additive constant by

$$\ell(\boldsymbol{\theta}) = \sum_{k=1}^s \sum_{j=1}^r \left\{ X_{\bullet jk} \log \beta_{jk} + X_{2jk} \log \alpha - X_{\bullet jk} \log\Big(1 + \alpha\langle \boldsymbol{z}_k, \boldsymbol{\beta}_k\rangle\Big)\right\}. \tag{3.1}$$

The MLE $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$, if it exists, is the solution to the following constrained optimization problem:

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta} \in \mathbb{R}_+^* \times (\mathbb{S}_{r-1})^s}{\operatorname{argmax}} \ \ell(\boldsymbol{\theta}). \tag{3.2}$$

It is proved (see [12]) that, if it exists, $\hat{\boldsymbol{\theta}}$ is also the solution of the following system of non-linear equations:

$$\begin{cases} \displaystyle\sum_{k=1}^s \sum_{j=1}^r \frac{X_{2jk} - \hat{\alpha} X_{1jk}\langle \boldsymbol{z}_k, \hat{\boldsymbol{\beta}}_k\rangle}{1 + \hat{\alpha}\langle \boldsymbol{z}_k, \hat{\boldsymbol{\beta}}_k\rangle} = 0 & \text{(3.3a)} \\[2ex] \displaystyle X_{\bullet jk} - \frac{n_k \hat{\beta}_{jk}(\hat{\alpha} z_{jk} + 1)}{1 + \hat{\alpha}\langle \boldsymbol{z}_k, \hat{\boldsymbol{\beta}}_k\rangle} = 0, \quad k = 1, \ldots, s, \quad j = 1, \ldots, r, & \text{(3.3b)} \end{cases}$$

where $\hat{\boldsymbol{\theta}} = (\hat{\alpha}, \hat{\boldsymbol{\beta}}^\mathsf{T})^\mathsf{T}$, $\hat{\boldsymbol{\beta}} = (\hat{\boldsymbol{\beta}}_1^\mathsf{T}, \ldots, \hat{\boldsymbol{\beta}}_s^\mathsf{T})^\mathsf{T}$ and for all $k = 1, \ldots, s$, $\hat{\boldsymbol{\beta}}_k = (\hat{\beta}_{1k}, \ldots, \hat{\beta}_{rk})^\mathsf{T}$.

In the case $s = 1$, [11] demonstrated the existence and uniqueness of the solution to the problem (3.3). In this paper, we extend this result to the general case $s \geqslant 1$. To do this, we will use Equation (3.3b) to get the $\hat{\beta}_{jk}$'s as functions of $\hat{\alpha}$ afterwards we will use Equation (3.3a) to get $\hat{\alpha}$ in an implicit form.

The following lemma shows that the subsystem (3.3b) can be decomposed into s subsystems of r linear equations with r unknowns.

Lemma 3.1. *Given $\hat{\alpha} > 0$, subsystem (3.3b) is equivalent to the following s linear systems of equations*

$$\boldsymbol{\Omega}_{\hat{\alpha},k}\hat{\boldsymbol{\beta}}_k = \frac{1}{n_k}\boldsymbol{X}_{\bullet k}, \quad k = 1, \ldots, s, \tag{3.4}$$

where for all $k = 1, \ldots, s$, $\boldsymbol{X}_{\bullet k} = (X_{\bullet 1k}, \ldots, X_{\bullet rk})^\mathsf{T}$ and

$$\boldsymbol{\Omega}_{\hat{\alpha},k} = \begin{bmatrix} 1 + \left(1 - \frac{X_{\bullet 1k}}{n_k}\right)\hat{\alpha}z_{1k} & -\frac{X_{\bullet 1k}}{n_k}\hat{\alpha}z_{2k} & \ldots & -\frac{X_{\bullet 1k}}{n_k}\hat{\alpha}z_{rk} \\ -\frac{X_{\bullet 2k}}{n_k}\hat{\alpha}z_{1k} & 1 + \left(1 - \frac{X_{\bullet 2k}}{n_k}\right)\hat{\alpha}z_{2k} & \ldots & -\frac{X_{\bullet 2k}}{n_k}\hat{\alpha}z_{rk} \\ \vdots & \ddots & \ddots & \vdots \\ -\frac{X_{\bullet rk}}{n_k}\hat{\alpha}z_{1k} & -\frac{X_{\bullet rk}}{n_k}\hat{\alpha}z_{2k} & \ldots & 1 + \left(1 - \frac{X_{\bullet rk}}{n_k}\right)\hat{\alpha}z_{rk} \end{bmatrix}.$$

Proof. Subsystem (3.3b) can be rewritten as

$$X_{\bullet jk}\Big(1 + \hat{\alpha}\sum_{j'=1}^{r} z_{j'k}\hat{\beta}_{j'k}\Big) - n_k\hat{\beta}_{jk}(1 + \hat{\alpha}z_{jk}) = 0, \quad j = 1, \ldots, r, \quad k = 1, \ldots, s.$$

For all $k = 1, \ldots, s$, we have

$$\sum_{\substack{j'=1 \\ j' \neq j}}^{r} \Big(X_{\bullet jk}\hat{\alpha}z_{j'k}\hat{\beta}_{j'k}\Big) - \Big(n_k + (n_k - X_{\bullet jk})\hat{\alpha}z_{jk}\Big)\hat{\beta}_{jk} = -X_{\bullet jk}, \quad j = 1, \ldots, r,$$

and then (after division by $-n_k$),

$$\left(1 + \left(1 - \frac{X_{\bullet jk}}{n_k}\right)\hat{\alpha}z_{jk}\right)\hat{\beta}_{jk} - \sum_{\substack{j'=1 \\ j' \neq j}}^{r}\left(\frac{X_{\bullet jk}}{n_k}\hat{\alpha}z_{j'k}\right)\hat{\beta}_{j'k} = \frac{X_{\bullet jk}}{n_k}, \quad j = 1, \ldots, r,$$

which completes the proof. ☐

The following lemma gives the closed-form expression of the solution to Equation (3.4).

Lemma 3.2. *For all $k = 1, \dots, s$, the matrix $\mathbf{\Omega}_{\hat{\alpha},k}$ is invertible and, therefore, Equation (3.4) has a unique vector solution $\hat{\boldsymbol{\beta}}_k$ given by*

$$\hat{\boldsymbol{\beta}}_k = \frac{1}{\displaystyle\sum_{m=1}^{r} \frac{X_{\bullet mk}}{1+\hat{\alpha} z_{mk}}} \left(\frac{X_{\bullet 1k}}{1+\hat{\alpha} z_{1k}}, \dots, \frac{X_{\bullet rk}}{1+\hat{\alpha} z_{rk}} \right)^{\mathsf{T}}. \tag{3.5}$$

Proof. The proof uses the notion of Schur complement applied to the inversion of block-defined matrices (see for example [13, p. 34]). For all $k = 1, \dots, s$, the matrix $\mathbf{\Omega}_{\hat{\alpha},k}$ of Equation (3.4) has the same structure as its homonym defined in Lemma 3.4 of [11]. The arguments which prove that $\mathbf{\Omega}_{\hat{\alpha},k}$ is non-singular are similar to those used in the first lines of the proof of Lemma 3.4 of [11]. Let

$$\mathbf{\Delta}_{\hat{\alpha},k} = \begin{bmatrix} 1+\hat{\alpha} z_{1k} & & \bigcirc \\ & \ddots & \\ \bigcirc & & 1+\hat{\alpha} z_{rk} \end{bmatrix}$$

be a diagonal matrix of order $r \times r$ and

$$\boldsymbol{M}_{\hat{\alpha},k} = \begin{bmatrix} \mathbf{\Delta}_{\hat{\alpha},k} & \hat{\alpha} \boldsymbol{X}_{\bullet k} \\ \boldsymbol{z}_k^{\mathsf{T}} & 1 \end{bmatrix}.$$

By Lemmas 3.1 and 3.3 of the same paper [11], the Schur complement of $\mathbf{\Delta}_{\hat{\alpha},k}$ in $\boldsymbol{M}_{\hat{\alpha},k}$, denoted $(\boldsymbol{M}_{\hat{\alpha},k}/\mathbf{\Delta}_{\hat{\alpha},k})$, is

$$(\boldsymbol{M}_{\hat{\alpha},k}/\mathbf{\Delta}_{\hat{\alpha},k}) = \frac{1}{n_k} \sum_{m=1}^{r} \frac{X_{\bullet mk}}{1+\hat{\alpha} z_{mk}} > 0$$

and the inverse of $\mathbf{\Omega}_{\hat{\alpha},k}$ is given by

$$\mathbf{\Omega}_{\hat{\alpha},k}^{-1} = \mathbf{\Delta}_{\hat{\alpha},k}^{-1} + \frac{\hat{\alpha}}{n_k} \mathbf{\Delta}_{\hat{\alpha},k}^{-1} \boldsymbol{X}_{\bullet k} (\boldsymbol{M}_{\hat{\alpha},k}/\mathbf{\Delta}_{\hat{\alpha},k})^{-1} \boldsymbol{z}_k^{\mathsf{T}} \mathbf{\Delta}_{\hat{\alpha},k}^{-1}.$$

By applying Theorem 3.5 of the same paper, we get

$$\hat{\boldsymbol{\beta}}_k = \frac{1}{n_k} (\boldsymbol{M}_{\hat{\alpha},k}/\mathbf{\Delta}_{\hat{\alpha},k})^{-1} \mathbf{\Delta}_{\hat{\alpha},k}^{-1} \boldsymbol{X}_{\bullet k}$$

and the proof is completed. □

The result concerning the existence and uniqueness of the MLE $\hat{\boldsymbol{\theta}}$ is given by the following theorem.

Theorem 3.3. *The MLE $\hat{\boldsymbol{\theta}} = (\hat{\alpha}, \hat{\beta}_{11}, \dots, \hat{\beta}_{r1}, \dots, \hat{\beta}_{1s}, \dots, \hat{\beta}_{rs})^{\mathsf{T}}$ is given by:*

$$\left\{ \begin{aligned} & \sum_{k=1}^{s} \sum_{j=1}^{r} \frac{X_{\bullet jk}}{1+\hat{\alpha} z_{jk}} = X_{1\bullet\bullet} && (3.6a) \\ & \hat{\beta}_{jk} = \frac{1}{\displaystyle\sum_{j'=1}^{r} \frac{X_{\bullet j'k}}{1+\hat{\alpha} z_{j'k}}} \times \frac{X_{\bullet jk}}{1+\hat{\alpha} z_{jk}}, \quad k = 1, \dots, s, \quad j = 1, \dots, r. && (3.6b) \end{aligned} \right.$$

Moreover, Equation (3.6a) has a unique solution $\hat\alpha$ and, therefore, the MLE $\hat{\boldsymbol\theta}$ is unique.

Proof. Expression (3.6b) is a simple consequence of (3.5). Thus, the main part of the proof consists in proving that Equation (3.6a) admits the MLE $\hat\alpha$ as unique solution. We can write Equation (3.3a) under the following equivalent forms:

$$\begin{aligned}\sum_{k=1}^{s}\sum_{j=1}^{r}\frac{X_{2jk}-\hat\alpha X_{1jk}\langle\boldsymbol z_k,\hat{\boldsymbol\beta}_k\rangle}{1+\hat\alpha\langle\boldsymbol z_k,\hat{\boldsymbol\beta}_k\rangle}=0 &\iff \sum_{k=1}^{s}\frac{\sum_{j=1}^{r}\Big(X_{2jk}-\hat\alpha X_{1jk}\langle\boldsymbol z_k,\hat{\boldsymbol\beta}_k\rangle\Big)}{1+\hat\alpha\langle\boldsymbol z_k,\hat{\boldsymbol\beta}_k\rangle}=0\\ &\iff \sum_{k=1}^{s}\frac{X_{2\bullet k}-\hat\alpha X_{1\bullet k}\langle\boldsymbol z_k,\hat{\boldsymbol\beta}_k\rangle}{1+\hat\alpha\langle\boldsymbol z_k,\hat{\boldsymbol\beta}_k\rangle}=0\\ &\iff \sum_{k=1}^{s}\frac{n_k-X_{1\bullet k}-\hat\alpha X_{1\bullet k}\langle\boldsymbol z_k,\hat{\boldsymbol\beta}_k\rangle}{1+\hat\alpha\langle\boldsymbol z_k,\hat{\boldsymbol\beta}_k\rangle}=0\end{aligned}$$

since $X_{1\bullet k}+X_{2\bullet k}=n_k$. Thus,

$$\begin{aligned}\sum_{k=1}^{s}\sum_{j=1}^{r}\frac{X_{2jk}-\hat\alpha X_{1jk}\langle\boldsymbol z_k,\hat{\boldsymbol\beta}_k\rangle}{1+\hat\alpha\langle\boldsymbol z_k,\hat{\boldsymbol\beta}_k\rangle}=0 &\iff \sum_{k=1}^{s}\frac{n_k-X_{1\bullet k}\Big(1+\hat\alpha\langle\boldsymbol z_k,\hat{\boldsymbol\beta}_k\rangle\Big)}{1+\hat\alpha\langle\boldsymbol z_k,\hat{\boldsymbol\beta}_k\rangle}=0\\ &\iff \sum_{k=1}^{s}\left(\frac{n_k}{1+\hat\alpha\langle\boldsymbol z_k,\hat{\boldsymbol\beta}_k\rangle}-X_{1\bullet k}\right)=0\end{aligned}$$

hence

$$\sum_{k=1}^{s}\frac{n_k}{1+\hat\alpha\langle\boldsymbol z_k,\hat{\boldsymbol\beta}_k\rangle}=X_{1\bullet\bullet} \tag{3.7}$$

From (3.6b), for all $k=1,\dots,s$, we have:

$$\begin{aligned}1+\hat\alpha\langle\boldsymbol z_k,\hat{\boldsymbol\beta}_k\rangle &= 1+\hat\alpha\sum_{j=1}^{r}z_{jk}\hat\beta_{jk}\\ &=1+\left(\sum_{j=1}^{r}\frac{\hat\alpha z_{jk}X_{\bullet jk}}{1+\hat\alpha z_{jk}}\right)\Bigg/\left(\sum_{j'=1}^{r}\frac{X_{\bullet j'k}}{1+\hat\alpha z_{j'k}}\right)\\ &=\left(\sum_{j=1}^{r}\frac{X_{\bullet jk}}{1+\hat\alpha z_{jk}}+\sum_{j=1}^{r}\frac{\hat\alpha z_{jk}X_{\bullet jk}}{1+\hat\alpha z_{jk}}\right)\Bigg/\left(\sum_{j'=1}^{r}\frac{X_{\bullet j'k}}{1+\hat\alpha z_{j'k}}\right)\\ &=\left(\sum_{j=1}^{r}\frac{(1+\hat\alpha z_{jk})X_{\bullet jk}}{1+\hat\alpha z_{jk}}\right)\Bigg/\left(\sum_{j'=1}^{r}\frac{X_{\bullet j'k}}{1+\hat\alpha z_{j'k}}\right)\\ &=\left(\sum_{j=1}^{r}X_{\bullet jk}\right)\Bigg/\left(\sum_{j'=1}^{r}\frac{X_{\bullet j'k}}{1+\hat\alpha z_{j'k}}\right)\\ &=n_k\Bigg/\left(\sum_{j'=1}^{r}\frac{X_{\bullet j'k}}{1+\hat\alpha z_{j'k}}\right),\end{aligned}$$

hence

$$\frac{n_k}{1+\hat{\alpha}\langle \boldsymbol{z}_k, \hat{\boldsymbol{\beta}}_k\rangle} = \sum_{j'=1}^{r} \frac{X_{\bullet j'k}}{1+\hat{\alpha} z_{j'k}}. \tag{3.8}$$

After replacing (3.8) in (3.7), we get Equation (3.6a).

To complete the proof, we need to show that the solution $\hat{\alpha}$ to Equation (3.6a) exists and is unique. Let Ψ be the function from $[0, +\infty[$ to $]0, \sum_{k=1}^{s} n_k]$ defined by:

$$\Psi(u) = \sum_{k=1}^{s}\sum_{j=1}^{r} \frac{X_{\bullet jk}}{1+u z_{jk}}. \tag{3.9}$$

The function Ψ is differentiable and its derivative is such that

$$\Psi'(u) = -\sum_{k=1}^{s}\sum_{j=1}^{r} \frac{X_{\bullet jk} z_{jk}}{(1+u z_{jk})^2} < 0,$$

hence function Ψ is strictly decreasing. We also have

$$\Psi(0) = \sum_{k=1}^{s}\sum_{j=1}^{r} X_{\bullet jk} = \sum_{k=1}^{s} n_k \quad \text{and} \quad \lim_{u\to+\infty} \Psi(u) = 0,$$

hence $\Psi([0, +\infty[) =]0, \sum_{k=1}^{s} n_k]$ and Ψ is bijective. Moreover, for all $k = 1,\dots,s$, we have $0 < X_{1jk} < X_{\bullet jk}$ because $X_{\bullet jk} = X_{1jk} + X_{2jk}$. So,

$$0 < \sum_{k=1}^{s}\sum_{j=1}^{r} X_{1jk} < \sum_{k=1}^{s}\sum_{j=1}^{r} X_{\bullet jk}$$

which means

$$0 < X_{1\bullet\bullet} < \sum_{k=1}^{s} n_k.$$

As Ψ is bijective and $X_{1\bullet\bullet} \in \Psi([0, +\infty[) =]0, \sum_{k=1}^{s} n_k]$, the solution $\hat{\alpha}$ to Equation $\Psi(u) = X_{1\bullet\bullet}$ (equivalent to Equation (3.6a)) exists and is unique. As $\hat{\alpha}$ is unique and each $\hat{\beta}_{jk}$ is obtained as a function of $\hat{\alpha}$, we conclude that the MLE $\hat{\boldsymbol{\theta}}$ is also unique. □

4. Strong consistency of the MLE

Let $(\Omega, \mathcal{A}, \mathrm{P})$ be the probability space linked to random vectors $\boldsymbol{X}_1, \dots, \boldsymbol{X}_s$ where for all $k = 1, \dots, s$, $\boldsymbol{X}_k = (X_{11k}, \dots, X_{1rk}, X_{21k}, \dots, X_{2rk})^\top$ satisfies Assumptions (A1), (A2), (A3) and the true unknown value of $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}^\top)^\top$ is denoted $\boldsymbol{\theta}^0 = (\alpha^0, (\boldsymbol{\beta}^0)^\top)^\top$ where $\boldsymbol{\beta}^0 = ((\boldsymbol{\beta}^0_1)^\top, \dots, (\boldsymbol{\beta}^0_s)^\top)^\top$ and for all $k = 1, \dots, s$, $\boldsymbol{\beta}^0_k = (\beta^0_{1k}, \dots, \beta^0_{rk})^\top$. The strong consistency (the almost sure convergence of the estimator $\hat{\boldsymbol{\theta}}$ to the true unknown value $\boldsymbol{\theta}^0$) has been proved in the case $s = 1$ by [6]. In this section, we extend their results by proving the strong consistency in the general case $s \geqslant 1$.

The study of almost sure (a.s.) convergence requires the precise definition of the sample size n which will tend to $+\infty$. In our context, this common sample size must be set such that $n \to +\infty$ if and only if for all $k = 1, \dots, s$, $n_k \to +\infty$.

One can argue that there are several ways to tend to infinity. As we are studying convergence, we assume (without loss of generality) that:

(A4): $n_1 = n_2 = \cdots = n_s = n$.

Assumption (A4) makes sense because of the strong law of large numbers (SLLN) and the properties of the multinomial distribution. For example, if we set $n = \min(n_1, \ldots, n_s)$ and if n is great enough, it is possible by the SLLN to rearrange the components of $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_s$ such that for all $k = 1, \ldots, s$, $\sum_{i=1}^{2} \sum_{j=1}^{r} X_{ijk} = n$.

Let us give the following two lemmas.

Lemma 4.1. *Let $\boldsymbol{\theta}^0$ be the true value of parameter vector $\boldsymbol{\theta}$ and for all $k = 1, \ldots, s$, $\boldsymbol{\pi}_k^0 = (\pi_{11k}^0, \ldots, \pi_{1rk}^0, \pi_{21k}^0, \ldots, \pi_{2rk}^0)^\intercal$ be a vector of class probabilities where, for all $i = 1, 2$ and $j = 1, \ldots, r$, $\pi_{ijk}^0 = \pi_{ijk}(\boldsymbol{\theta}^0)$ (see Formula (2.2) for the definition of $\pi_{ijk}(\boldsymbol{\theta})$). If the random vector $\boldsymbol{X}_k = (X_{11k}, \ldots, X_{1rk}, X_{21k}, \ldots, X_{2rk})^\intercal$ follows the multinomial distribution $\mathscr{M}(n, \boldsymbol{\pi}_k^0)$, then, when $n \to +\infty$,*

(i) $X_{ijk}/n \xrightarrow{a.s.} \pi_{ijk}^0, \quad i = 1, 2, \quad j = 1, \ldots, r.$

(ii) $X_{\bullet jk}/n \xrightarrow{a.s.} \pi_{\bullet jk}^0, \quad j = 1, \ldots, r.$

(iii) $X_{1\bullet\bullet} \xrightarrow{a.s.} \gamma^0$

where for all $k = 1, \ldots, s$ and $j = 1, \ldots, r$,

$$\pi_{\bullet jk}^0 = \pi_{1jk}^0 + \pi_{2jk}^0 = \frac{\beta_{jk}^0(1 + \alpha^0 z_{jk})}{1 + \alpha^0 \sum_{j'=1}^{r} z_{j'k}\beta_{j'k}^0} \tag{4.1}$$

and

$$\gamma^0 = \sum_{k=1}^{s} \frac{1}{1 + \alpha^0 \sum_{j'=1}^{r} z_{j'k}\beta_{j'k}^0}. \tag{4.2}$$

Proof. (i) It is proven (see for example [21, Lemma 2.42]) that the marginal distribution of the random variable X_{ijk} is the binomial distribution $\mathscr{B}(n, \pi_{ijk}^0)$. Hence, by application of the strong law of large numbers, X_{ijk}/n converges a.s. to π_{ijk}^0.

(ii) By item (i), $X_{1jk}/n \xrightarrow{a.s.} \pi_{1jk}^0$ and $X_{2jk}/n \xrightarrow{a.s.} \pi_{2jk}^0$. Thus $X_{\bullet jk}/n = (X_{1jk}/n + X_{2jk}/n) \xrightarrow{a.s.} (\pi_{1jk}^0 + \pi_{2jk}^0) = \pi_{\bullet jk}^0$.

(iii) By item (i), $X_{1jk}/n \xrightarrow{a.s.} \pi_{1jk}^0$. Thus

$$\frac{X_{1\bullet\bullet}}{n} = \sum_{k=1}^{s}\sum_{j=1}^{r} \frac{X_{1jk}}{n} \quad \xrightarrow{a.s.} \quad \sum_{k=1}^{s}\sum_{j=1}^{r} \pi_{1jk}^0 = \sum_{k=1}^{s}\sum_{j=1}^{r} \frac{\beta_{jk}^0}{1 + \alpha^0 \sum_{j'=1}^{r} z_{j'k}\beta_{j'k}^0}$$
$$= \sum_{k=1}^{s} \frac{1}{1 + \alpha^0 \sum_{j'=1}^{r} z_{j'k}\beta_{j'k}^0}.$$

□

Lemma 4.2. *For all $\omega \in \Omega$, let $\varphi_{\omega,n}$ and φ be the functions from $[0,+\infty[$ to $]0,s]$ respectively defined by*

$$\varphi_{\omega,n}(u) = \sum_{k=1}^{s}\sum_{j=1}^{r} \frac{X_{\bullet jk}(\omega)/n}{1+uz_{jk}} \quad \textit{and} \quad \varphi(u) = \sum_{k=1}^{s}\sum_{j=1}^{r} \frac{\pi^0_{\bullet jk}}{1+uz_{jk}},$$

where for all $k = 1,\dots,s$ and $j = 1,\dots,r$, $\pi^0_{\bullet jk}$ is defined by Formula (4.1). The functions $\varphi_{\omega,n}$ and φ are bijective with inverses respectively denoted $\varphi^{-1}_{\omega,n}$ and φ^{-1}. Moreover, there exists a set $E_1 \subset \Omega$ such that $\mathrm{P}(E_1) = 0$ and for all $\omega \in \Omega \setminus E_1$, the sequence of inverse functions $\varphi^{-1}_{\omega,n}$ converges uniformly to φ^{-1}.

Proof. The functions $\varphi_{\omega,n}$ and φ are both strictly decreasing (their respective derivatives are strictly negative). Moreover,

$$\lim_{u\to+\infty} \varphi_{\omega,n}(u) = 0, \quad \varphi_{\omega,n}(0) = \frac{1}{n}\sum_{k=1}^{s}\sum_{j=1}^{r} X_{\bullet jk}(\omega) = \frac{1}{n}\sum_{k=1}^{s} n = s,$$

$$\lim_{u\to+\infty} \varphi(u) = 0, \quad \text{and} \quad \varphi(0) = \sum_{k=1}^{s}\left(\sum_{j=1}^{r} \pi^0_{\bullet jk}\right) = \sum_{k=1}^{s} 1 = s.$$

Thus $\varphi_{\omega,n}([0,+\infty[) = \varphi([0,+\infty[) =]0,s]$ and we can conclude that both functions are bijective. For all $k = 1,\dots,s$ and $j = 1,\dots,r$, $X_{\bullet jk}/n \xrightarrow{a.s.} \pi^0_{\bullet jk}$ (Lemma 4.1) and, by definition of the a.s. convergence, there exists a set $N_{jk} \subset \Omega$ such that $\mathrm{P}(N_{jk}) = 0$ and

$$\forall \omega \in \Omega \setminus N_{jk}, \quad \lim_{n\to+\infty} \frac{X_{\bullet jk}(\omega)}{n} = \pi^0_{\bullet jk}. \tag{4.3}$$

Thus, the set $E_1 = \bigcup_{k=1}^{s}\bigcup_{j=1}^{r} N_{jk}$ is such that $\mathrm{P}(E_1) = 0$ and for all $\omega \in \Omega \setminus E_1$,

$$\lim_{n\to+\infty} \varphi_{\omega,n}(u) = \lim_{n\to+\infty} \sum_{k=1}^{s}\sum_{j=1}^{r} \frac{X_{\bullet jk}(\omega)/n}{1+uz_{jk}} = \sum_{k=1}^{s}\sum_{j=1}^{r} \frac{\pi^0_{\bullet jk}}{1+uz_{jk}} = \varphi(u).$$

Thus, for all $\omega \in \Omega \setminus E_1$, the sequence of functions $\varphi_{\omega,n}$ converges simply to φ. Moreover,

$$\begin{aligned}
\sup_{u\in[0,+\infty[} |\varphi_{\omega,n}(u) - \varphi(u)| &= \sup_{u\in[0,+\infty[} \left|\sum_{k=1}^{s}\sum_{j=1}^{r} \frac{X_{\bullet jk}(\omega)/n - \pi^0_{\bullet jk}}{1+uz_{jk}}\right| \\
&\leqslant \sup_{u\in[0,+\infty[} \sum_{k=1}^{s}\sum_{j=1}^{r} \left|\frac{X_{\bullet jk}(\omega)/n - \pi^0_{\bullet jk}}{1+uz_{jk}}\right| \\
&\leqslant \sup_{u\in[0,+\infty[} \sum_{k=1}^{s}\sum_{j=1}^{r} \frac{\left|X_{\bullet jk}(\omega)/n - \pi^0_{\bullet jk}\right|}{1+uz_{jk}} \\
&\leqslant \sum_{k=1}^{s}\sum_{j=1}^{r} \left|\frac{X_{\bullet jk}(\omega)}{n} - \pi^0_{\bullet jk}\right|
\end{aligned}$$

since for all $u \in [0, +\infty[$, $j = 1, \dots, r$ and $k = 1, \dots, s$, $1/(1 + uz_{jk}) \leqslant 1$. So we have

$$\sup_{u \in [0,+\infty[} |\varphi_{\omega,n}(u) - \varphi(u)| \leqslant \sum_{k=1}^{s} \sum_{j=1}^{r} \left| \frac{X_{\bullet jk}(\omega)}{n} - \pi^0_{\bullet jk} \right| \xrightarrow[n\to+\infty]{} 0$$

which proves the uniform convergence of the sequence of bijective functions $\varphi_{\omega,n}$ to φ. By Theorem 2 of [2], we conclude that the sequence of functions $\varphi^{-1}_{\omega,n}$ converges uniformly to φ^{-1}. □

It is well known that the almost sure (a.s.) convergence of a random vector is equivalent to the a.s. convergence of each of its components [14, Proposition 6.55]. To prove that $\hat{\boldsymbol{\theta}}$ converges almost surely (a.s.) to $\boldsymbol{\theta}^0$, it is then sufficient to prove that $\hat{\alpha}$ converges a.s. to α^0 and each $\hat{\beta}_{jk}$ converges a.s. to β^0_{jk}.

We start with the following theorem on the a.s. convergence of $\hat{\alpha}$ and which is a generalization of Theorem 3.2 of [6].

Theorem 4.3. *The MLE $\hat{\alpha}$ of α converges a.s. to the true value α^0 i.e. there exists a subset $N \subset \Omega$ such that $\mathrm{P}(N) = 0$ and*

$$\forall \omega \in \Omega \setminus N, \quad \lim_{n\to+\infty} \hat{\alpha}(\omega) = \alpha^0. \tag{4.4}$$

Proof. By dividing Equation (3.6a) by n, we get

$$\forall \omega \in \Omega, \quad \sum_{k=1}^{s} \sum_{j=1}^{r} \frac{X_{\bullet jk}(\omega)/n}{1 + \hat{\alpha}(\omega) z_{jk}} = \frac{X_{1\bullet\bullet}(\omega)}{n}$$

which is equivalent to

$$\forall \omega \in \Omega, \quad \hat{\alpha}(\omega) = \varphi^{-1}_{\omega,n}\left(X_{1\bullet\bullet}(\omega)/n\right), \tag{4.5}$$

where $\varphi_{\omega,n}$ is defined by Lemma 4.2. By Lemma 4.1, $X_{1\bullet\bullet}/n \xrightarrow{a.s.} \gamma^0$ and therefore, there exists a null probability set $E_2 \subset \Omega$ such that

$$\forall \omega \in \Omega \setminus E_2, \quad \lim_{n\to+\infty} \frac{X_{1\bullet\bullet}(\omega)}{n} = \gamma^0.$$

Let $N = E_1 \cup E_2$ where E_1 is defined by Lemma 4.2. The set N is such that $\mathrm{P}(N) = 0$. For all $\omega \in \Omega \setminus N$, the sequence of inverse functions $\varphi^{-1}_{\omega,n}$ converges uniformly to φ^{-1} and the real sequence $X_{1\bullet\bullet}(\omega)/n$ converges to γ^0. Therefore, by Theorem 7.3.5 of [17], Equation (4.5) implies that

$$\forall \omega \in \Omega \setminus N, \quad \hat{\alpha}(\omega) \xrightarrow[n\to+\infty]{} \varphi^{-1}(\gamma^0).$$

Since φ is bijective (Lemma 4.2) and

$$\begin{aligned}\varphi(\alpha^0) &= \sum_{k=1}^{s}\sum_{j=1}^{r} \frac{\pi^0_{\bullet jk}}{1+\alpha^0 z_{jk}} \\ &= \sum_{k=1}^{s}\sum_{j=1}^{r} \frac{\beta^0_{jk}}{1+\alpha^0 \sum_{j'=1}^{r} z_{j'k}\beta^0_{j'k}} \\ &= \sum_{k=1}^{s} \frac{\sum_{j=1}^{r}\beta^0_{jk}}{1+\alpha^0 \sum_{j'=1}^{r} z_{j'k}\beta^0_{j'k}} \\ &= \sum_{k=1}^{s} \frac{1}{1+\alpha^0 \sum_{j'=1}^{r} z_{j'k}\beta^0_{j'k}} \\ &= \gamma^0,\end{aligned}$$

we have $\varphi^{-1}(\gamma^0) = \alpha^0$ and we conclude that

$$\forall \omega \in \Omega \setminus N, \quad \hat{\alpha}(\omega) \xrightarrow[n\to+\infty]{} \varphi^{-1}(\gamma^0) = \alpha^0.$$

The proof is thus completed. □

The result concerning the strong consistency of the MLE $\hat{\boldsymbol{\theta}}$ is given by the following theorem.

Theorem 4.4. *The MLE $\hat{\boldsymbol{\theta}} = (\hat{\alpha}, \hat{\boldsymbol{\beta}}^\top)^\top$ of $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}^\top)^\top$ is strongly consistent i.e. $\hat{\boldsymbol{\theta}} = (\hat{\alpha}, \hat{\boldsymbol{\beta}}^\top)^\top$ converges a.s. to $\boldsymbol{\theta}^0 = (\alpha^0, (\boldsymbol{\beta}^0)^\top)^\top$.*

Proof. By Theorem 4.3, $\hat{\alpha}$ converges a.s. to α^0. Dividing the numerator and denominator of Equation (3.6b) by n, we get for all $k = 1,\dots,s$ and $j = 1,\dots,r$,

$$\hat{\beta}_{jk} = \frac{1}{\displaystyle\sum_{j'=1}^{r} \frac{X_{\bullet j'k}/n}{1+\hat{\alpha} z_{j'k}}} \times \frac{X_{\bullet jk}/n}{1+\hat{\alpha} z_{jk}} = g_j\Big(\frac{X_{\bullet 1k}}{n}, \dots, \frac{X_{\bullet rk}}{n}, \hat{\alpha}\Big)$$

where g_j is the continuous function from $(\mathbb{R}_+^*)^{r+1}$ to $\mathbb{R}_+^*$ defined by:

$$g_j(y_1, \dots, y_r, u) = \frac{1}{\displaystyle\sum_{m=1}^{r} \frac{y_m}{1+u z_{mk}}} \times \frac{y_j}{1+u z_{jk}}$$

By Lemma 4.1 and Theorem 4.3, we have

$$\Big(\frac{X_{\bullet 1k}}{n}, \dots, \frac{X_{\bullet rk}}{n}, \hat{\alpha}\Big) \xrightarrow{a.s.} (\pi^0_{\bullet 1k}, \dots, \pi^0_{\bullet rk}, \alpha^0).$$

By applying the continuous-mapping theorem [19, Theorem 2.3], we get

$$\hat{\beta}_{jk} \quad \xrightarrow{a.s.} \quad g_j(\pi^0_{\bullet 1k}, \dots, \pi^0_{\bullet rk}, \alpha^0) = \frac{1}{\displaystyle\sum_{m=1}^{r} \frac{\pi^0_{\bullet mk}}{1+\alpha^0 z_{mk}}} \times \frac{\pi^0_{\bullet jk}}{1+\alpha^0 z_{jk}}$$

where

$$\frac{\pi^0_{\bullet jk}}{1+\alpha^0 z_{jk}} = \frac{\beta^0_{jk}(1+\alpha^0 z_{jk})}{(1+\alpha^0 z_{jk})\Big(1+\alpha^0\sum_{j'=1}^r z_{j'k}\beta^0_{j'k}\Big)} = \frac{\beta^0_{jk}}{1+\alpha^0\sum_{j'=1}^r z_{j'k}\beta^0_{j'k}}$$

and

$$\sum_{m=1}^r \frac{\pi^0_{\bullet mk}}{1+\alpha^0 z_{mk}} = \sum_{m=1}^r \frac{\beta^0_{mk}}{1+\alpha^0\sum_{j'=1}^r z_{j'k}\beta^0_{j'k}} = \frac{\sum_{m=1}^r \beta^0_{mk}}{1+\alpha^0\sum_{j'=1}^r z_{j'k}\beta^0_{j'k}} = \frac{1}{1+\alpha^0\sum_{j'=1}^r z_{j'k}\beta^0_{j'k}}.$$

Thus,

$$\hat{\beta}_{jk} \quad \xrightarrow{a.s.} \quad g_j(\pi^0_{\bullet 1k},\ldots,\pi^0_{\bullet rk},\alpha^0) = \beta^0_{jk}.$$

Since $\hat{\alpha} \xrightarrow{a.s.} \alpha^0$ and $\hat{\beta}_{jk} \xrightarrow{a.s.} \beta^0_{jk}$ for all $k = 1, \ldots, s$ and $j = 1, \ldots, r$, we can conclude that $\hat{\boldsymbol{\theta}} = (\hat{\alpha}, \hat{\boldsymbol{\beta}}^\mathsf{T})^\mathsf{T}$ converges a.s. to $\boldsymbol{\theta}^0 = (\alpha^0, (\boldsymbol{\beta}^0)^\mathsf{T})^\mathsf{T}$. The proof is thus completed. □

5. Concluding remarks

In this paper, we have considered a statistical model for accident frequencies combining a given number of multinomial distributions whose parameters are dependent. We proved the existence, the uniqueness and the strong consistency of the maximum likelihood estimator (MLE) $\hat{\boldsymbol{\theta}}$ of the parameter vector $\boldsymbol{\theta}$. We thus generalize the results of [11] and [6].

We have not considered numerical computation of the MLE $\hat{\boldsymbol{\theta}}$ in this paper (for example, Equation (3.6a) can only be solved using a numerical algorithm). We rather refer the reader to [10, 12] where three different algorithms (minorization-maximization (MM), gradient projection - expectation maximization (GP-EM) and Newton-Raphson algorithms) have been used for efficient computation of the MLE $\hat{\boldsymbol{\theta}}$. The performance of these algorithms has been examined through simulation studies and it was found that the mean square error (MSE)

$$\mathrm{MSE}(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}^0) = \frac{1}{1+sr}\left((\hat{\alpha}-\alpha^0)^2 + \sum_{k=1}^s\sum_{j=1}^r(\hat{\beta}_{jk}-\beta^0_{jk})^2\right)$$

decreases and tends to zero when the sample size increases. Thus, the numerical results obtained in [10, 12] provide good illustrations for the results demonstrated in our paper and, in return, our results provide a theoretical justification for the numerical convergence observed in these papers.

References

1. F. Balabdaoui, C. Durot, and H. Jankowski, *Behaviour of the monotone single index model under repeated measurements*, Sankhya A (2021), 1–27. https://doi.org/10.1007/s13171-021-00250-7

2. E. Barvínek, I. Daler, and J. Francu, *Convergence of sequences of inverse functions*, Arch. Math. (Brno) **27** (1991), no. 2, 201–204.
3. D. Chafaı and D. Concordet, *On the strong consistency of asymptotic M-estimators*, J. Statist. Plann. Inference **137** (2007), no. 9, 2774–2783. https://doi.org/10.1016/j.jspi.2006.09.027
4. D. J. Eck and C. J. Geyer, *Computationally efficient likelihood inference in exponential families when the maximum likelihood estimator does not exist*, Electron. J. Stat. **15** (2021), no. 1, 2105–2156. https://doi.org/10.1214/21-EJS1815
5. S. Fiorin, *The strong consistency for maximum likelihood estimates: a proof not based on the likelihood ratio*, C. R. Math. Acad. Sci. Paris **331** (2000), no. 9, 721–726. https://doi.org/10.1016/S0764-4442(00)01718-3
6. I. C. Geraldo, A. N'Guessan, and K. E. Gneyou, *A note on the strong consistency of a constrained maximum likelihood estimator used in crash data modeling*, C. R. Math. Acad. Sci. Paris **353** (2015), no. 12, 1147–1152. https://doi.org/10.1016/j.crma.2015.09.025
7. S. Kourouklis, *On the strong consistency of a solution to the likelihood equation*, Statist. Probab. Lett. **5** (1987), no. 1, 23–27.
8. M. M. Mittal and R. C. Dahiya, *Estimating the parameters of a doubly truncated normal distribution: Estimating the parameters*, Comm. Statist. Simulation Comput. **16** (1987), no. 1, 141–159. https://doi.org/10.1080/03610918708812582
9. M. M. Mittal and R. C. Dahiya, *Estimating the parameters of a truncated Weibull distribution*, Comm. Statist. Theory Methods **18** (1989), no. 6, 2027–2042. https://doi.org/10.1080/03610928908830020
10. A. Mkhadri, A. N'Guessan, and B. Hafidi, *An MM algorithm for constrained estimation in a road safety measure modeling*, Comm. Statist. Simulation Comput. **39** (2010), no. 5, 1057–1071.
11. A. N'Guessan, *Analytical existence of solutions to a system of nonlinear equations with application*, J. Comput. Appl. Math. **234** (2010), no. 1, 297–304. https://doi.org/10.1016/j.cam.2009.12.026
12. A. N'Guessan, A. Essai, and C. Langrand, *Estimation multidimensionnelle des contrôles et de l'effet moyen d'une mesure de sécurité routière*, Revue de Statistique Appliquée **49** (2001), no. 2, 85–102.
13. A. N'Guessan and M. Truffier, *Impact d'un aménagement de sécurité routière sur la gravité des accidents de la route*, J. SFdS **49** (2008), 23–41.
14. M. A. Proschan and P. A. Shaw, *Essentials of Probability Theory for Statisticians*, Chapman and Hall/CRC Texts in Statistical Science, CRC Press, New York, 2016. https://doi.org/10.1201/9781315370576
15. S. S. Shapiro and H. Zahedi, *Bernoulli trials and discrete distributions*, Journal of Quality Technology **22** (1990), no. 3, 193–205. https://doi.org/10.1080/00224065.1990.11979239
16. M. Sobel, *Comparisons of estimators of a slippage probability*, Comm. Statist. Theory Methods **24** (1995), no. 4, 1039–1055. https://doi.org/10.1080/03610929508831539
17. R. S. Strichartz, *The Way of Analysis*, Jones and Bartlett Publishers, Boston, 2000.
18. G. Tzavelas and D. Panagiotakos, *Statistical inference for the size-biased Weibull distribution*, J. Stat. Comput. Simul. **83** (2013), no. 7, 1252–1265. https://doi.org/10.1080/00949655.2012.657197
19. A. W. Van der Vaart, *Asymptotic Statistics*, Cambridge University Press, Cambridge (United Kingdom), 1998. https://doi.org/10.1017/CBO9780511802256
20. A. Wald, *Note on the consistency of the maximum likelihood estimate*, Ann. Math. Statistics **20** (1949), no. 4, 595–601. https://doi.org/10.1214/aoms/1177729952
21. L. Wasserman, *All of Statistics: a concise course in statistical inference*, Springer, New York, 2004.

Département de Mathématiques, Faculté des Sciences, Université de Lomé, 1 B.P. 1515 Lomé 1, Togo.
Email address: cherifgera@gmail.com