Gulf Journal of Mathematics ISSN: 2309-4966
Vol 12, Issue 2 (2022) 49-56 https://doi.org/10.56947/gjom.v12i2.733

MAXIMUM LIKELIHOOD ESTIMATION IN THE GENERALIZED EXTREME VALUE REGRESSION MODEL FOR BINARY DATA

LO FATIMATA[1], BA DEMBA BOCAR[1*] AND DIOP ABA[2]

ABSTRACT. Logistic regression model is obviously used to study a binary response variable Y (presence or absence of an event) in practice but may have limitations in the case of rare event. Thus, Generalized extreme value regression model is proposed when the dependent variable Y represents a rare event. Therefore, the quantile function of the GEV distribution is used as link function to investigate the relationship between the binary outcome Y and a set of potential predictors $\mathbf{X}$. In this article we develop a maximum likelihood estimation procedure int he generalized extreme value regression model. We establish the asymptotic properties (existence, consistency and asymptotic normality) of the proposed maximum likelihood estimator.

1. ITRODUCTION

Generalized extreme value (GEV) regression model has become a standard tool to investigate the relationship between a binary response Y and a set of potential predictors [10]. The binary response may represent, for example, the occurrence of some outcome of interest ($Y = 1$ if the outcome occurred and $Y = 0$ otherwise). When the dependent variable Y represents a rare event, the logistic regression model ([1], [4]) shows relevant drawbacks.

When the logistic regression model is employed, it is assumed that the response curve between the covariates and the probability is symmetric. This assumption may not always be true, and it may be severely violated when the number of observations in the two response categories are significantly different from each other. This unbalance is not uncommon when we consider binary rare events data (i.e. binary dependent variables with a very small number of ones) which happens with only a small probability. [8] used GEV regression model, that uses the quantile function of the GEV distribution as link function, for analyzing default probabilities and find that its predictive performance is better than that of the logistic regression predictive model. They point out that logistic regression could underestimate the probability of default of rare events since the link function is a symmetric function. [10] used this model for an application to B2B electronic

Date: Received: Nov 6, 2021; Accepted: Apr 20, 2022.

* Corresponding author.

2010 *Mathematics Subject Classification.* Primary 62F12; Secondary 62F03.

Key words and phrases. reneralized extreme value, regression model, maximum likelihood estimation, consistency, asymptotic normality.

payments system adoption.

The main aim of this paper is to develop a maximum likelihood procedure in the GEV regression model by taking the quantile function of the GEV distribution as link function. On the contrary and to the best of our knowledge, maximum likelihood estimator (MLE) of the unknown parameter β in the GEV regression model have never been investigated. The present article aims at filling this gap. We construct the MLE and we establish its asymptotic properties.

The rest of the paper is organized as follows. In Section 2, we describe the GEV regression model and the proposed estimation procedure. In Section 3, we rigorously establish the asymptotic properties of the resulting estimator. A discussion and some perspectives for future work are given in Section 4.

2. GEV REGRESSION MODEL

Let $(Y_1, \mathbf{X}_1), \ldots, (Y_n, \mathbf{X}_n)$ be independent and identically distributed copies of the random vector $(Y, \mathbf{X})$ defined on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$. For every individual $i = 1, \ldots, n$, Y_i is a binary response variable indicating say, the occurrence of some outcome of interest ($Y_i = 1$ if the outcome occurred and $Y_i = 0$ otherwise). Let $\mathbf{X}_i = (1, X_{i2}, \ldots, X_{ip})'$ be random vectors of predictors or covariates. Let $\pi(x_i) = \mathbb{P}(Y_i = 1|\mathbf{X}_i = x_i)$ the conditional probability. Since we consider the class of Generalized Linear Models, we suggest the GEV cumulative distribution function proposed by [8] as the response curve given by

$$\begin{aligned} \pi(x_i) &= 1 - \exp\left([-(1 - \tau(\beta_1 + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}))_+]^{-1/\tau}\right) \\ &= 1 - GEV(-\beta' x_i; \tau) \end{aligned} \tag{2.1}$$

where $\beta = (\beta_1, \ldots, \beta_p)' \in \mathbb{R}^p$ is an unknown regression parameter measuring the association between potential predictors and the binary response Y and $GEV(x; \tau)$ represents the cumulative probability at x for the GEV distribution with a location parameter $\mu = 0$, a scale parameter $\sigma = 1$, an unknown shape parameter τ.

For $\tau \to 0$, the previous model (2.1) becomes the response curve of the log-log model, for $\tau > 0$ and $\tau < 0$ it becomes the Frechet and Weibull response curve respectively, a particular case of the GEV one.
The link function of the GEV model is given by

$$\frac{1 - [\log(1 - \pi(x_i))]^{-\tau}}{\tau} = \beta' x_i \tag{2.2}$$

that represents a noncanonical link function.
Note the parameters μ and σ are set to fixed constants for model identifiability ([10]).

The likelihood function for the unknown p-dimensional parameter β from the independent sample $(y_1, \mathbf{x}_1), \ldots, (y_n, x_n)$ is as follows:

$$L_n(\beta) = \prod_{i=1}^{n} [1 - \mathrm{GEV}(-\beta' x_i; \tau)]^{y_i} \times [\mathrm{GEV}(-\beta' x_i; \tau)]^{1-y_i}. \tag{2.3}$$

We define the maximum likelihood estimator $\widehat{\beta}_n$ as the solution of the p-dimensional score equation

$$\dot{l}_n(\beta) := \frac{\partial \log L_n(\beta)}{\partial \beta} = 0, \tag{2.4}$$

In the following, we shall be interested in the asymptotic properties of the maximum likelihood estimator $\hat{\beta}_n$ of β. Before proceeding, we need to set some further notations.

Define first the $(p \times n)$ matrice

$$\mathbb{X} = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ X_{12} & X_{22} & \ldots & X_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ X_{1p} & X_{2p} & \ldots & X_{np} \end{pmatrix}$$

Let also $C(\beta) = (C_i(\beta))_{1 \leq i \leq n}$ be the n-dimensional column vector defined as

$$C_i(\beta) = -x_{ij} \frac{ln\pi(x_i) \times y_i - \pi(x_i)}{(1 + \tau\beta' x_i)(1 - \pi(x_i))}, \quad i = 1, \ldots, n$$

Then, simple algebra shows that the score equation can be rewritten as

$$\dot{l}_n(\beta) = XC(\beta) = 0.$$

If $H = (H_{ij})_{1 \leq i \leq a, 1 \leq j \leq b}$ denotes some $(a \times b)$ matrix, we will denote by $H_{\bullet j}$ it's j-th column $(j = 1, \ldots, b)$ that is, $H_{\bullet j} = (H_{1j}, \ldots, H_{aj})'$. Then, it will be useful to rewrite the score vector as

$$\dot{l}_n(\beta) = \sum_{i=1}^{n} \mathbb{X} \bullet j C_j(\beta).$$

We shall further note $\ddot{l}_n(\theta)$ the $(k \times k)$ matrix of second derivatives of $l_n(\beta)$, that is $\ddot{l}_n(\beta) = \frac{\partial^2 l_n(\beta)}{\partial\beta\partial\beta'}$. Let $\mathbf{D}(\beta) = (\mathbf{D}_{ij}(\beta))_{1 \leq i,j \leq n}$ be the $(n \times n)$ diagonal matrix with i-th diagonal elements $(i = 1, \ldots, n)$ given by

$$D_{ii}(\beta) = \frac{x_i^2 \frac{ln\pi(x_i)}{1-\pi(x_i)}(1 - \pi(x_i))}{1 + \tau\beta' x_i} - \frac{x_i \frac{ln\pi(x_i)}{1-\pi(x_i)}(y_i - \pi(x_i))}{(1 + \tau\beta x_i)^2} \left[\frac{\tau x_i - \tau x_i \pi(x_i) + \pi(x_i) \log \pi(x_i)}{1 - \pi(x_i)} \right]$$

Then, some algebra shows that $\ddot{l}_n$ can be expressed as

$$\ddot{l}_n(\beta) = -\mathbb{X}\mathbb{D}(\beta)\mathbb{X}'.$$

In the next section, we turn to the asymptotic theory for the proposed estimator $\hat{\beta}_n$ of β.

3. ASYMPTOTIC THEORY

We first state some regularity conditions that will be needed to study the asymptotic theory.

C1: The covariates are bounded that is, there exist compact sets $G \subset \mathbb{R}^p$ such that $\mathbf{X}_i \in G$ for every $i = 1, 2, \ldots$ For every $i = 1, 2, \ldots,$ $j = 2, \ldots, p$, $\text{var}[X_{ij}] > 0$. For every $i = 1, 2, \ldots,$ the X_{ij} $(j = 1, \ldots, p)$ are linearly independent.

C2: The true parameter value β_0 lies in the interior of known compact sets $\mathcal{G} \subset \mathbb{R}^p$.

C3: The Hessian matrix $\ddot{l}_n(\theta)$ is negative definite and of full rank, for every $n = 1, 2, \ldots$ Let λ_n and Λ_n be respectively the smallest and largest eigenvalues of $\mathbb{X}\mathbf{D}(\beta_0)\mathbb{X}'$. There exists a finite positive constant c such that $\Lambda_n/\lambda_n < c$ for every $n = 1, 2, \ldots$

C4: Let A denote a $p \times p$ matrix such that $\|A\| = \max_{1 \le i \le p} |\lambda_i|$.

Now we establish rigorously the existence, consistency and asymptotic normality of the maximum likelihood estimator $\hat{\beta}_n$ in model (2.1). We prove the following results:

Theorem 3.1 (Existence and consistency)**.** *Under the conditions* ***C1-C4****, The maximum likelihood estimator $\hat{\beta}_n$ exists almost surely as $n \to \infty$ and converges almost surely to β_0, if and only if λ_n tends to infinity as $n \to \infty$.*

Proof. Let γ_n be defined as $\gamma_n(\beta) = -(\mathbb{X}\mathbf{D}(\beta)\mathbb{X}')\dot{l}_n(\beta)$. Then $\gamma_n(\beta_0)$ converges almost surely to 0 as $n \to \infty$. Note that

$$\gamma_n(\beta_0) = \left(\frac{1}{n}\ddot{l}_n(\beta_0)\right)^{-1} \times \left(\frac{1}{n}\dot{l}_n(\beta_0)\right).$$

The condition C3 implies that, $(\frac{1}{n}\ddot{l}_n(\beta_0))^{-1}$ converges to a matrix Σ. Then,

$$\frac{1}{n}\dot{l}_n(\beta_0) = \frac{1}{n}\mathbb{X}C(\beta_0) = \begin{pmatrix} \frac{1}{n}\sum_{i=1}^n X_{i1}C_i(\beta_0) \\ \vdots \\ \frac{1}{n}\sum_{i=1}^n X_{ip}C_i(\beta_0) \end{pmatrix}$$

converges almost surely to 0 as $n \to \infty$.
Indeed for all $i = 1, \ldots, n$ and $j = 1, \ldots, p$ we have

$$\begin{aligned} \mathbb{E}[X_{ij}C_i(\beta_0)] &= \mathbb{E}[\mathbb{E}(X_{ij}C_i(\beta_0)|\mathbf{X}_i = x_i)] \\ &= \mathbb{E}[X_{ij}\mathbb{E}(C_i(\beta_0)|\mathbf{X}_i = x_i] \end{aligned}$$

where

$$\begin{aligned} \mathbb{E}(C_i(\beta_0)|\mathbf{X}_i = x_i) &= \mathbb{E}\left(y_i\frac{x_i \log \pi(x_i)}{1 + \tau\beta' x_i} - (1 - y_i)\frac{x_i \log \pi(x_i)\pi(x_i)}{(1 + \tau\beta' x_i)(1 - \pi(x_i))}|X_i = x_i\right) \\ &= \frac{x_i \log \pi(x_i)}{1 + \tau\beta' x_i}\mathbb{E}(y_i|X_i = x_i) - \frac{x_i \log \pi(x_i)\pi(x_i)}{(1 + \tau\beta' x_i)(1 - \pi(x_i))}\mathbb{E}(1 - y_i|X_i = x_i) \\ &= \frac{x_i \log \pi(x_i)}{1 + \tau\beta' x_i}\pi(x_i) - \frac{x_i \log \pi(x_i)\pi(x_i)}{(1 + \tau\beta' x_i)(1 - \pi(x_i))}(1 - \pi(x_i)) \\ &= 0 \end{aligned}$$

Moreover, we have

$$\begin{aligned}\mathbb{V}ar(X_{ip}C_i(\beta_0)) &= \mathbb{E}(\mathbb{V}ar(X_{ip}C_i(\beta_0))|\mathbf{X}_i = x_i] + \mathbb{V}ar[\mathbb{E}(X_{ip}C_i(\beta_0))|\mathbf{X}_i = x_i] \\ &= \mathbb{E}[X_{ip}^2 \mathbb{V}ar(C_i(\beta_0)|\mathbf{X}_i = x_i] \\ &= \mathbb{E}\left[X_{ip}^2\left(\frac{x_i \log \pi(x_i)}{1+\tau\beta' x_i} + \frac{x_i \log \pi(x_i)}{(1+\tau\beta' x_i)(1-\pi(x_i))}\right)^2 \mathbb{V}ar(y_i|\mathbf{X}_i = x_i)\right] \\ &\leq \mathbb{E}\left[X_{ip}^2\left(\frac{x_i ln\pi(x_i)}{1+\tau\beta' x_i} + \frac{x_i ln\pi(x_i)}{(1+\tau\beta' x_i)(1-\pi(x_i))}\right)^2\right] := c_3\end{aligned}$$

The conditions C1, C2 and C3 implie that $c_3 < \infty$, then

$$\sum_{i=1}^{n} \frac{\mathbb{V}ar(C_i(\beta_0)}{i^2} \leq c_3 \sum_{i=1}^{n} \frac{1}{i^2} \leq \infty.$$

Kolmogorov's strong law of large numbers ensures that, as $n \to \infty$

$$\frac{1}{n}\sum_{i=1}^{n}\{X_{ij}C_i(\beta_0) - E(X_{ij}C_i(\beta_0))\} = \frac{1}{n}\sum_{i=1}^{n} C_i(\beta_0)X_{ij} \ \to \ 0 \text{ as } n \to \infty.$$

Finaly, as $n \to \infty$ $\frac{1}{n}\dot{l}_n(\beta_0)$ and $\gamma_n(\beta_0)$ converge almost surely to 0. Let $\epsilon > 0$, the almost surely convergence of $\gamma_n(\beta_0)$ implies then for almost every $\omega \in \Omega$, there exists an integer value $n(\epsilon,\omega)$ such that fo any $n \geq n(\epsilon,\omega)$, $\|\gamma_n(\beta_0)\| \leq \epsilon$ or equivalently, $0 \in B(\gamma_n(\beta_0), \epsilon)$. In particulary, let $\epsilon = (1-c)s$ with $0 < c < 1$ such as in Lemma 4.2 of [2]. Since γ_n satisfies the Lipschitz condition, the Lemma Lemma 4.2 of [2] ensures that there exists an element of $B(\beta_0, s)$ (let denote this element by $\hat{\beta}_n$) such that $\gamma_n(\hat{\beta}_n) = -(\mathbb{X}\mathbf{D}(\hat{\beta}_n)\mathbb{X}')\dot{l}_n(\hat{\beta}_n) = 0$. The condition C3 implies that $\dot{l}_n(\hat{\beta}_n) = 0$ and that $\hat{\beta}_n$ is the unique maximizer of l_n. To summarize, we have shown that for almost every $\omega \in \Omega$ and for every $s > 0$, there exists an integer value $n(s,\omega)$ such that if $n \geq n(s,\omega)$ then the maximum likelihood estimator $\hat{\beta}_n$ exists, and $\|\hat{beta}_n - \beta_0\| \leq s$ (that is, $\hat{\beta}_n$ converges almost surely to β_0).
which concludes the proof. □

We now turn to the convergence in distribution of the proposed estimator $\hat{\beta}_n$ of β, which is stated by the following theorem:

Theorem 3.2 (Asymptotic normality). *Assume that the conditions **C1-C4** hold and that $\hat{\beta}_n$ converges almost surely to β_0. Let $\widehat{\Sigma} = \mathbb{X}\mathbf{D}(\hat{\beta}_n)\mathbb{X}'$ and I_p denote the identity matrix of order p. Then $\widehat{\Sigma}_n^{\frac{1}{2}}(\hat{\beta}_n - \beta_0)$ converges in distribution to the Gaussian vector $\mathcal{N}(0, I_p)$.*

Proof. A Taylor expansion of the score function is as

$$0 = \dot{l}_n(\hat{\beta}_n) = \dot{l}_n(\beta_0) + \ddot{l}_n(\tilde{\beta}_n)(\hat{\beta}_n - \beta_0)$$

where $\tilde{\beta}_n$ lies between $\hat{\theta}_n$ and β_0, and thus $\dot{l}_n(\beta_0) = -\ddot{l}_n(\tilde{\beta}_n)(\hat{\beta}_n - \beta_0)$. Let $\tilde{\Sigma}_n = -\ddot{l}_n(\tilde{\beta}_n) = \mathbb{X}\mathbf{D}(\tilde{\beta}_n)\mathbb{X}'$ and $\Sigma_{n,0} = \mathbb{X}\mathbf{D}(\beta_0)\mathbb{X}'$. Now

$$\widehat{\Sigma}_n^{\frac{1}{2}}(\hat{\beta}_n - \beta_0) = \left[\widehat{\Sigma}_n^{\frac{1}{2}}\tilde{\Sigma}_n^{\frac{1}{2}}\right]\left[\tilde{\Sigma}_n^{\frac{1}{2}}\Sigma_{n,0}^{\frac{1}{2}}\right]\Sigma_{n,0}^{\frac{1}{2}}\left[\tilde{\Sigma}_n(\widehat{\theta}_n - \theta_0)\right]. \tag{3.1}$$

The two terms in brackets in (3.1) converge almost surely to I_p. To see this, we show for example that $\|\tilde{\Sigma}_n^{-\frac{1}{2}}\Sigma_{n,0}^{\frac{1}{2}} - I_p\| \longrightarrow 0$ almost surely as $n \to \infty$. First, note that

$$\begin{aligned} \|\tilde{\Sigma}_n^{-\frac{1}{2}}\Sigma_{n,0}^{\frac{1}{2}} - I_p\| &= \|\tilde{\Sigma}_n^{-\frac{1}{2}}\left(\Sigma_{n,0}^{\frac{1}{2}} - \tilde{\Sigma}_n^{\frac{1}{2}}\right)\| \\ &\leq \|\tilde{\Sigma}_n^{\frac{1}{2}}\|\|\Sigma_{n,0}^{\frac{1}{2}} - \tilde{\Sigma}_n^{\frac{1}{2}}\| \\ &\leq \Lambda_n^{\frac{1}{2}}\|\tilde{\Sigma}_n^{\frac{1}{2}}\|\|\Lambda_n^{-\frac{1}{2}}\left(\Sigma_{n,0}^{\frac{1}{2}} - \tilde{\Sigma}_n^{\frac{1}{2}}\right)\| \end{aligned} \tag{3.2}$$

and

$$\Lambda_n^{-1}\|\Sigma_{n,0} - \tilde{\Sigma}_n\| = \Lambda_n^{-1}\|\mathbb{X}(\mathbf{D}(\beta_0) - \mathbf{D}(\tilde{\beta}_n))\mathbb{X}'\|.$$

Note also that $\hat{\beta}_n$ converges almost surely to β_0 (that is, for every $w \in \tilde{\Omega}$, where $\tilde{\Omega} \subset \Omega$ and $\mathbb{P}(\tilde{\Omega}) = 1$). Let $w \in \tilde{\Omega}$. By the same arguments as in the proof of Lemma 4.2 of [2], for every $\epsilon > 0$, there exists a positive $n(\epsilon, w) \in \mathbb{N}$ such that if $n \geq n(\epsilon, w)$, then $\Lambda_n^{-1}\|\mathbb{X}(\mathbf{D}(\beta_0) - \mathbf{D}(\tilde{\beta}_n))\mathbb{X}'\| < \epsilon$. Hence $\Lambda_n^{-1}\|\mathbb{X}(\mathbf{D}(\beta_0) - \mathbf{D}(\tilde{\beta}_n))\mathbb{X}'\|$ converges almost surely to 0. By continuity of the map $x \to x^{\frac{1}{2}}$, $\|\Lambda_n^{-\frac{1}{2}}\left(\Sigma_{n,0}^{\frac{1}{2}} - \tilde{\Sigma}_n^{\frac{1}{2}}\right)\|$ converges also almost surely to 0. Moreover, for n sufficiently large, there exists a positive constant $A < \infty$, such that almost surely, $\|\tilde{\Sigma}_n^{-\frac{1}{2}}\| \leq A\frac{\Lambda_n^{\frac{1}{2}}}{\lambda_n^{1/2}} < Ac_2^{\frac{1}{2}}$. It follows from (3.2) and the condition C3 that $\|(\Sigma_{n,0}^{\frac{1}{2}}\tilde{\Sigma}_n^{-\frac{1}{2}}) - I_p\|$ converges almost surely to 0.

It remains for us to show that $\Sigma_{n,0}^{-\frac{1}{2}}(\tilde{\Sigma}_n(\widehat{\theta}_n - \theta_0))$ converges in distribution to $\mathcal{N}(0, I_p)$. Note that $\Sigma_{n,0}^{-\frac{1}{2}}(\tilde{\Sigma}_n(\hat{\beta}_n - \beta_0)) = \Sigma_{n,0}^{-\frac{1}{2}}\sum_{j=1}^{2n}\mathbb{X}_{\bullet j}C_j(\beta_0)$. This convergence holds if we can check the following conditions:

i) $\max_{1\leq j\leq n}\mathbb{X}'_{\bullet j}(\mathbb{X}\mathbb{X}')^{-1}\mathbb{X}_{\bullet j} \to 0$ if $n \to \infty$. Condition i) follows by noting that

$$0 \leq \max_{1\leq j\leq n}\mathbb{X}'_{\bullet j}(\mathbb{X}\mathbb{X}')^{-1}\mathbb{X}_{\bullet j} \leq \max_{1\leq j\leq n}\|\mathbb{X}_{\bullet j}\|^2\|(\mathbb{X}\mathbb{X}')^{-1}\| = \max_{1\leq j\leq n}\frac{\|\mathbb{X}_{\bullet j}\|^2}{\tilde{\lambda}_n}.$$

and that $\|\mathbb{X}_{\bullet j}\|^2$ is bounded above, by C1 and C2. Moreover, $\frac{1}{\tilde{\lambda}_n} \to 0$ as $n \to \infty$.

ii) $\sup_{1\leq j\leq n}\mathbb{E}\left[C_j^2(\beta_0)\mathbf{1}_{\{|C_j(\theta_0)|>c\}}\right] \to 0$ as $n \to \infty$. Condition ii) follows by noting that the components $C_j(\beta_0)$ of C are bounded under C1, C2 and C3. Finally, for every $i = 1, \ldots, n$, $\mathbb{E}(C_j^2(\beta_0)) = Var(C_j(\beta_0))$ since $\mathbb{E}(C_j(\beta_0)) = 0$. Therefore, for all $j \in \{1, \ldots, n\}$ under conditions C1, C2 and C3, $\mathbb{V}ar(C_j(\beta_0)) > 0$.

iii) $\inf_{1\leq j\leq n}\mathbb{E}(C_j^2(\beta_0)) > 0$. Condition iii) follows by using similar arguments, that is $\mathbb{V}ar(C_j(\beta_0)) > 0$ for all $j = 1, \ldots, n$.

To summarize, we have proved that $\Sigma_{n,0}^{-\frac{1}{2}}(\tilde{\Sigma}_n(\hat{\beta}_n - \beta_0))$ converges in distribution to $\mathcal{N}(0, I_p)$. This result, combined with Slutsky's theorem and equation (3.1), implies that $\widehat{\Sigma}_n^{\frac{1}{2}}(\hat{\beta}_n - \beta_0)$ converges in distribution to $\mathcal{N}(0, I_p)$.
which concludes the proof. □

4. DISCUSSION AND PERSPECTIVES

In this paper, we have proposed a procedure of maximum likelihood estimation in the Generalized extreme value regression model by using the quantile function of the GEV distribution as link function which is not a cononical lik function. We have established the asymptotic properties (existence, consistency and asymptotic normality) of the proposed maximum likelihood estimator $\hat{\beta}_n$ of the parameter β. The maximum likelihood estimator of the parameter β will allow the calculation of odds ratios which are of major interest in epidemiology. It is also possible, using this method, to establish model validation tests (deviance tests, likelihood ratio tests).

It is interesting to note that it is also possible to estimate the shape parameter τ with this maximum likelihood method, but in this work we are particularly interested in the parameter of interest β.

Another issue of interest deals with: i) the inference on the shape parameter τ in the GEV regression model; ii) the inference in a high-dimensional setting, when the predictor dimension is much larger than the sample size (this problem arises, for example, in genetic studies where high-dimensional data are generated using microarray technologies. Some articles ([5], [6]) have addressed the estimation in the logistic model for binary data); iii) the inference in the GEV regression model with presence of zero inflation; iv) estimate the predictive probability and define the simultaneous confidence bands for this probability. Extending these methods to the GEV regression model constitutes another nontrivial topic for future work. It will also be interesting to study the statistical convergence of sequence $(\widehat{\beta}_n)_{n\geq 0}$ in subspaces (see [3] and [7] for more details in statistical convergence of sequences of subspaces).

ACKNOWLEDGMENTS

The authors are grateful to the associate editor and one referee for their detailed comments that helped us to improve a previous version of this article.

References

1. A. Agresti, *Categorical Data Analysis*. Wiley, New York, (2002).
2. A. Diop, A. Diop and J.F. Dupuy, *Maximum likelihood estimation in the logistic regression model with a cure fraction*, Electronic Journal of Statistics. Vol. **5**, (2011) 460–483.
3. F. Nuray, *Statistical convergence of sequences of subspaces*, Gulf J. Math., Vol. **10**, Issue **2** (2021), 31-37.
4. G. Celeux and J.P. Nakache, *Analyse Discriminante sur Variables Qualitatives*, Polytechnica, (1994).
5. J. Huan, S. Ma and C.H. Zhang, *The Iterated Lasso forHigh-Dimensional Logistic Regression*. Technical Report No. **392**, Iowa City, Iowa: The University of Iowa, (2008).
6. L. Meier, S. Van de Geer and P. Buhlmann, *The group Lasso for logistic regression*, Journal of the Royal Statistical Society, Series B **70**, (2008), 53–71.
7. M. Candan and E. E. Kara, *A study on topolojical and geometrical characteristics of new Banach sequence spaces*, Gulf J. Math., Vol **3**, Issue **4**, (2015), 67-84.
8. R. Calabrese and S. Osmetti, *Modelling SME Loan Defaults asRare Events: an Application to Credit Defaults*, Journal of Applied Statistics **40** (**6**), (2013), 1172-1188.

9. S.G. Coles, *An Introduction to Statistical Modeling of Extreme Values*, Springer, New York, (2000).
10. X. Wang and D.K. Dey, *Generalized extreme value regression for binary response data: An application to B2B electronic payments system adoption*, Ann. Appl. Stat. **4**, (2010), 2000-2023.

[1] DÉPARTEMENT DE MATHÉMATIQUES, UNIVERSITÉ IBA DER THIAM (UIDT), THIÈS, SÉNÉGAL.
Email address: fatymalo@hotmail.fr

[1] DÉPARTEMENT DE MATHÉMATIQUES, UNIVERSITÉ IBA DER THIAM (UIDT), THIÈS, SÉNÉGAL.
Email address: dbba@univ-thies.sn

[2] DÉPARTEMENT DE MATHÉMATIQUES, UNIVERSITÉ ALIOUNE DIOP DE BAMBEY (UADB), BAMBEY, SÉNÉGAL.
Email address: aba.diop@uadb.edu.sn