République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITÉ MOHAMED KHIDER, BISKRA

FACULTÉ des SCIENCES EXACTES et des SCIENCES de la NATURE et de la VIE

DÉPARTEMENT DE MATHÉMATIQUES

Mémoire présenté en vue de l'obtention du Diplôme **:**

MASTER en Mathématiques

Option : **Statistique**

Par

NOM Prénom

Djebbari Ismahane

Titre :

Sur la distribution de Gumbel et ses appliquations

Membres du Comité d'Examen :

Dr.	BENATIA Fateh	UMKB	Président
Dr.	**BRAHIMI Brahime**	UMKB	Encadreur
Dr.	DHIABI Samra	UMKB	Examinateur

Juin 2018

Dédicace

Je dédie ce modeste travail

A mes très chers parents, qui m'ont tellement soutenu tout

Au long de ce travail, conseil et encouragements qu'ils trouvent la marque de tout mon amour pour eux

J'espère qui Dieu m'aide pour leur faire plaisir

Mes très chers frères : Ahmed,Iyad,Walid,Siraj

Mes très chères sœurs : Hadjer,Aycha,Rima

Mes très chers aimes et mes sœurs : hanan,afaf,souha,khadija,nisrin,

A tous mes collègues de promotion 2018

A tous les enseignants de ma période d'étude

A tous ceux que j'aime et m'aiment

REMERCIEMENTS

Au terme de ce travail nous remerciements Dieu qui a donné
patience, force et volonté et le courage pour réalisé ce mémoire je tene au terme de ce
travail à exprimer notre plus grand remerciement à members du jury.
Je tene particulierement à remercier notre directrice de travail
Mame "Abdeli Djihan" pour avoir accepté d'encadrer ce
travail et nous avoir dirigé guide, conseillé et encourage, aussi nos remercions l'encadreur
Ms"Brahimi Brahim" qui a terminer son travail avec nous,
ainsi que sa bonne volonté, sa patience et ses précicieux conseils, qu'il nous prodigué tout
au long de ce travail.
Ce travail a été realise au ministere de sciences exacte et
science de la nature et de vie de L'Université Mohamed
Khider Biskra. A ce titre, je remercier tous ceux qui nous
ont aides de prés ou de loin.Nous exprimons note profonde reconnaissance à nos parents.
Je remercier s'adressent en finalité à tous mes collegues de je
promotion pour leur soutien et encouragement.

Table des matières

Table des figures

Introduction

Au cours des dernières années entre (1920-1940) c'est le début de la théorème des valeurs extremes(TVE) quels sont considérons une branche des statistiques ,quand plusieurs des sientifiques travaillont par des bases théorique pour les valeurs extremes .Il y a deux statisticens britaniques (R.AFisher et L.H.C Tippett 1928),on déduit la forme limites de la fonction de répartitions des maximas de l'échantillon des variables aléatoires .En 1943 Gnedenko a apporté des conditions nécessaires et suffisantes pour la convergence faible dont la probabilité d'aparition

Les évènements extremes sont généralement des évènements rares .Le but de la théorème des valeurs extremes(TVE) est de modéliser et de décrire la servenue et l'intensité d'évènement dits rares ,elle se repose principalement sur des distributions limites des extremes. Pour une présentation assez complète du sujet,nous renvoyons à l'ouvrage de référence de Embrechts, Klüppelberg et Mikosch (1997) rappelant les principaux résultats théoriques sur la TVE et à Reiss et Thomas (2001) qui proposent un certain nombre d'exemples pratiques, en finance, en assurance et en sciences environnementales ,les applications ont commencé suite aux travaux de Gumbel en 1958.

La modélisation des évènements extrêmes (tremblement de terre, inondation, accidents nucléaires, crises financiéres, krachs boursiers, chocs pétroliers,...etc.) est aujourd'hui un champ de recherches particuliérement actif, notamment par l'importance de leurs impacts économiques et sociaux.On a retrouve deux modéles :("loi généralisées des extremes GEV"),("loi de paréto généralisées GPD" et leurs domaines d'attractions de Fré-

chet,Weibull et Gumbel ,cette dernière est un cas particulière de la distribution des valeurs extremes pour modéliser la distribution du maximum ,la distribution de Gumbel est plus utilisée en hydrologie et climatologie pour éstimer des valeurs extremes par plusieurs des méthodes d'éstimation nous les montionnons méthode des moments , L.moment,maximum de vraisemblance et des moments pondérers généralisées
Ce mémoire est organisée en deux chapitres :
chapitre 01 : est consacré à présenté des quelques idées théorique sur la théorié des valeurs extremes (TVE) nous montionnons la définition des statistiques d'ordre et les distributions limites et leurs domaine d'attraction nous parlons aussi sur l'éstimation d'indice des valeurs extremes et la méthode (POT).
chapitre 02 : dans ce chapitre on à présenté la distribution de Gumbel avec des propriétés et des relations de cette distribution on suite on discute l'estimation pour la loi de Gumbel avec des méthodes d'éstimations.

Chapitre 1

Quelques rappels sur la théorie des valeurs extrêmes

La théorie des valeurs extrêmes a été crée pour la modélisation des événements rares. Elle joue un rôle très important pour l'étude de la mésure et la gestion des risques ënancière.

Dans ce chapitre, nous rappelons quelques notions essentielles sur la théorie des valeurs extrêmes (TVE). Nous commençons d'abord par la définition des statistiques d'ordre, ensuite nous présentons les distributions limites des extrêmes et leurs domaines d'attraction et présentée l'éstimation d'indice des valeurs extrèmes et on à trouver la méthode peak over threshold (POT).

1.1 Présentation de la théorie des valeurs extrêmes

1.1.1 Statistique d'ordre

Soit $X_1, ..., X_n$, n variables aléatoires (iid) de distribution commune $F(x)$ est une fonction de répartition

$$F(x) = \Pr(X \leq x). \tag{1.1}$$

On concidére les variables aleatoires rangés par ordre croissant comme suit

$$X_{1,n} \leq .... \leq X_{n,n},$$

appeler statistique d'ordre.

Proposition 1 (Fonction de répartition de la $k^{éme}$ statistique d'ordre) *Soit $X_{k,n}$ est la $k^{éme}$ statistique d'ordre alors :*

$$F_{X_{k,n}}(x) = \sum_{i=k}^{n} \binom{n}{i} [F(x)]^i [1 - F(x)]^{n-i}, \ 1 \leq \ k \leq n$$

Définition 1 (Fonction de répartition empirique) *La fonction F_n est donnée par*

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} 1_{]-\infty,x[}(X_i), x \in \mathbb{R} \tag{1.2}$$

la fonction inverse généralisée de F notée par Q, telle que

$$Q(t) = F^{-1}(t) = \inf\{s, F(s) \geq t\}, 0 < t < 1.$$

Proposition 2 *Les termes du maximum et du minimum est notée par :*

$$X_{1,n} = \min(X_1, ..., X_n) \qquad (X_{1,n} \textit{ le statistique du minimum}) \tag{1.3}$$

$$X_{n,n} = max(X_1, ..., X_n) \quad (X_{n,n} \textit{ le statistique du maximum}) \tag{1.4}$$

1.1.2 Distribution du maximum et du minimum

La formule du distribution exact est

$$F_{X_{n,n}}(x) = [F(x)]^n \quad, \ -\infty \leq x \leq +\infty \tag{1.5}$$

$$F_{X_{1,n}}(x) = 1 - [1 - F(x)]^n, \quad -\infty \leq x \leq +\infty \tag{1.6}$$

Par la dirivation des deux expressions 1.6,1.5alors :

$$\begin{cases} f_{X_{1,n}}(x) & = & n[1 - F(x)]^{n-1} f(x) & -\infty \leq x \leq +\infty \\ f_{X_{n,n}}(x) & = & n[F(x)]^{n-1} f(x) & -\infty \leq x \leq +\infty \end{cases}$$

1.1.3 Théoréme de Fisher et Tippet 1928 Gnedenko 1943

Considérons n variable aléatoires $X_1......X_n$ (iid) de mème loi et notons $S_n = X_1 + ... + X_n$ leur somme.Si l'espérance et la variance de ces variables aléatoires sont finies,le théoréme central limite(TCL) permet d'affirmer que :

$$\frac{S_n - nE(X)}{\sqrt{n}\sigma} \rightarrow^{loi} N(0.1)$$

Ce théoreme permet de connètre le comportement de $\frac{S_n - a_n}{b_n}$ pour des suites a_n, b_n ,de cas symétrique le résultat de EVT est donné les lois limite du maximum corrèctement normaliser par des suites normalisantes a_n, b_n .

Soit $(X_n)_{n\geq1}$ une suite de variable aléatoire (i.i.d) avec 1.1.S'il existe deux suites normlisantes réelles $\left((a_n)_{n\geq1} > 0\ , (b_n)_{n\geq1} \in \mathbb{R}\right)$ et une loi non-dégénérée H telle que :

$$\lim_{n\to\infty} \Pr\left(\frac{X_{n,n} - a_n}{b_n} \leq x\right) = \lim_{n\to\infty} F^n(a_n + b_n x) = H_\alpha(x), \forall x \in \mathbb{R}$$

alors Λ, Ψ_α et $\phi_\alpha$sont les distribution limites possibles pour le maximum.

$$\Lambda(x) = \exp(-\exp(-x)) \quad , x \in \mathbb{R} \qquad \text{(distribution de Gumbel)} \tag{1.7}$$

$$\Psi_\alpha(x) = \begin{cases} 1 & , \quad x>0 \\ \exp\left(-(-x)^\alpha\right) & , \quad x \leq 0 \end{cases} \quad \text{(distribution de Weibul } (\alpha<0)\text{)}$$

$$\phi_\alpha(x) = \begin{cases} 0 \ , & x<0 \\ \exp\left(-x^{-\alpha}\right) \ , & x \ \geq 0 \end{cases} \quad \text{(distribution de Fréchet } (\alpha>0)\text{)}$$

Exemple 1 (Cas de la loi exponentielle) *Supposons F suit une loi exponentielle de paramètre* 1 *donc* $F(x) = 1 - \exp(-x)$ *.En posant* $a_n = \log(n), b_n = 1$

$$\begin{aligned} \Pr\left(\frac{X_{n,n} - a_n}{b_n} \leq x\right) &= F^n(b_n x + a_n) \\ &= \Pr\left(X_{n,n} \leq x + \log(n)\right) \\ &= \Pr\left[X \leq x + \log(n)\right]^n \\ &= \left(1 - \exp(-x - \log(n))\right)^n \\ &= \left(1 + \frac{(-1)}{n}\exp(-x)\right)^n . \\ &= \exp(-\exp(-x)) \to^{n\to\infty} \Lambda(x) \end{aligned}$$

1.2 Distribution généralisée des valeurs extrêmes

La distribution généralisée des valeurs extrêmes (GEV) notée par :

$$H_{\mu,\sigma,\xi}(y) = \begin{cases} \exp\left\{-\left[1+\xi\left(\frac{y-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\} & \text{pour} \quad 1 + \frac{\xi}{\sigma}(y-\mu) \ >0 \\ \exp\left(-\exp\left(-\frac{y-\mu}{\sigma}\right)\right) & \text{pour} \quad \xi = 0 \quad ,y \in \mathbb{R} \end{cases} \tag{1.8}$$

avec $\mu \in \mathbb{R}, \sigma>0$.

En remplacent $\frac{y-\mu}{\sigma}$ par x on à trouvée la formule suivant :

$$H_\xi(x) = \begin{cases} \exp\left[-(1+\xi x)^{-\frac{1}{\xi}}\right] & \text{si} \quad \xi \neq 0 \ , \ 1 + \xi x>0 \\ \exp(-\exp(-x)) & \text{si} \quad \xi = 0, x \in \mathbb{R} \end{cases}$$

La fonction de densité est donnée par :

$$h_\xi(x) = \begin{cases} \frac{1}{\sigma}\left(1+\xi\left(\frac{y-\mu}{\sigma}\right)\right)^{-\frac{1}{\xi}-1}\exp\left(-\left(1+\xi\frac{y-\mu}{\sigma}\right)^{-\frac{1}{\xi}}\right) & \text{si} \quad \xi \neq 0 \\ \frac{1}{\sigma}\left(\exp\left(-\exp\left(-\frac{y-\mu}{\sigma}\right)-\frac{y-\mu}{\sigma}\right)\right) & \text{si} \quad \xi = 0 \end{cases}$$

1. Si $\xi = \frac{1}{\alpha} > 0$ alors $\phi_\alpha(x) = H\left(\frac{1}{\alpha}\right)(\alpha(x-1))$.
2. Si $\xi = -\frac{1}{\alpha} < 0$ alors $\Psi_\alpha(x) = H\left(-\frac{1}{\alpha}\right)(\alpha(x+1))$.
3. Si $\xi = 0$ alors $\Lambda(x) = H(0)$

1.2.1 Fonctions à variation régulière

Une fonction mésurable $g : \mathbb{R}^+ \to \mathbb{R}^+$ est à variation régulière $(g \in RV(\alpha))$. à l'infini si et seulement si il existe un réel α tel que :

$$lim_{t\to\infty}\frac{g(tx)}{g(t)} = x^\alpha, (x>0) \tag{1.9}$$

Les fonctions à variation lente sont notées $l(x)$ on a :

$$lim_{t\to\infty}\frac{l(tx)}{l(t)} = 1, (x>0)$$

Exemple 2 *Soit $l(x)$ une fonction logarithmique telque : $l(t) = \log(t)$ d aprés la formule de l'Opitale on a :*

$$\lim_{t\to\infty}\frac{\log(tx)}{\log(t)} = \lim_{t\to\infty}\frac{\frac{x}{tx}}{\frac{1}{t}} = 1$$

Alors $l(x)$ est a variation lente.

Théorème 1 (Représentation de Karamata) *l est une fonction à variation lente si et seulement si*

$$l(x) = a(x)\left\{-\int_1^x \frac{g(t)}{t}dt\right\}, (x>0)$$

où :

$$a(x) \to_{x\to\infty} a, a \in]0,\infty[\quad et \quad g(t) \to_{t\to\infty} 0 \ .$$

Proposition 3 (Condition àvariation réguliere du premier ordre) *si viréfie les équivalents suivants :*

$$\begin{aligned} F \in DA\left(\phi_{-\frac{1}{\alpha}}\right) &\Leftrightarrow 1 - F(x) \text{ est àvariation réguliere d'indice } \left(-\tfrac{1}{\alpha}\right) \\ &\Leftrightarrow \lim_{t\to\infty} \tfrac{1-F(tx)}{1-F(t)} = x^{-\frac{1}{\alpha}}, \quad (x>0) \Leftrightarrow \lim_{t\to\infty} \tfrac{U(tx)}{U(t)} = x^{\alpha} \end{aligned}$$

avec la fonction de valeurs extreme de quantile U notée :

$$U(t) = F^{-1}(1 - 1/t) = (1/\overline{F})^{-1}(t), t \geq 1.$$

1.2.2 Caractirisation des domaines d'attractions

Dans cette partie nous allons trouver trois domaine d'attraction différent. Soit $F(x)$ une fonction de répartition avec α un indice de valeur extrême et l une fonction à variation lente ou a_n, b_n les constantes du normalisation.

Domaine d'attraction de ϕ_α

Théorème 2 *$F(x)$ appartient au domaine d'attraction de Fréchet avec $(\alpha>0)$ ssi $x_F = +\infty$ et $1 - F(x) = x^{-\alpha} l(x)$*

Si $F \in DA(\phi_\alpha)$ alors :

$$\Pr\left(\frac{X_{n,n} - a_n}{b_n} \leq x\right) \to_{n\to\infty} \phi_\alpha(x)$$

$$a_n = F^{-1}\left(1 - \frac{1}{n}\right), b_n = 0.$$

F une fonction de distribution absolument continu avec la densité f si :

$$lim_{x\to x_F}\frac{xf(x)}{1-F(x)}=\alpha>0$$

Alors $F\in DA(\phi_\alpha)$.

Exemple 3 (Loi de Pareto) *Soit* $F(x)$ *une fonction de répartition de loi de Pareto telque* $(x,\alpha\rangle 0)$:

$$F(x)=1-x^{-\alpha}$$
$$f(x)=\alpha x^{-\alpha-1}$$

on a : $lim_{x\to x_F}\frac{xf(x)}{F(x)}=lim_{x\to x_F}\alpha\frac{xx^{-\alpha-1}}{x^{-\alpha}}=\alpha\rangle 0$
Alors $F\in DA(\phi_\alpha)$ *et* $a_n=F^{-1}\left(1-\frac{1}{n}\right), b_n=0$
D'ou :$lim_{n\to\infty}\Pr\left(\frac{X_{n,n}-a_n}{b_n}\leq x\right)$
$=lim_{n\to\infty}\Pr\left(X_{n,n}\leq n^{\frac{1}{\alpha}}x\right)$
$=\lim_{n\to\infty}F^n\left(n^{\frac{1}{\alpha}}x\right)$
$=\lim_{n\to\infty}\left(1-\frac{x^{-\alpha}}{n}\right)^n=\exp\left(-x^{-\alpha}\right)$

Domaine d'attraction de Ψ_α

Théorème 3 $F(x)$ *appartien au domaine d'attraction de Weibull avec* $(\alpha<0)$ *ssi* $x_F<+\infty$ *et* $[1-F(x)]\left(x_F-\frac{1}{x}\right)=x^{-\alpha}l(x)$. *Si* $F\in DA(\Psi_\alpha)$ *alors :*

$$\Pr\left(\frac{X_{n,n}-a_n}{b_n}\leq x\right)\to_{n\to\infty}\Psi_\alpha(x)$$

$$a_n=x_F-F^{-1}\left(1-\frac{1}{n}\right)\ ,\quad b_n=x_F$$

F une fonction de distribution absolument continu avec la densité f qui est positive sur un certain intervalle fini (t, xF). Si

$$\lim_{x \to x_F} \frac{(x_F - x) f(x)}{1 - F(x)} = \alpha > 0$$

Alors $\quad F \in DA(\Psi_\alpha)$.

Domaine d'attraction de Λ_α

Définition 2 *(Fonction de Von Mises) F une fonction de distribution est dite fonction de Von Mises avec la fonction auxiliaire b tel que :*

$$1 - F(x) = a \exp\left\{-\int_t^x \frac{1}{b(s)}\right\}, \ t \quad \leq \ x \leq x_F, (a > 0)$$

et b est une fonction positive absolument continue (par rapport la mésure de Lebesgue) avec la densité b' ayant :

$$\lim_{x \to x_F} b'(x) = 0. \tag{1.10}$$

Théorème 4 *$F(x)$ appartient au domaine d'attraction de Gumbel ssi ($t < x_F$) tels que*

$$1 - F(x) = c(x) \exp\left\{-\int_t^x \frac{h(s)}{b(s)}\right\} \quad , \ t \ \leq \ x \leq \ x_F \tag{1.11}$$

et c et h sont des fonctions mésurables tels que $c(x) \to c > 0$, $h(x) \to 1$ quand $x \to x_F$. Avec 1.10. Si $F \in DA(\Lambda_\alpha)$ alors :

$$\Pr\left(\frac{X_{n,n} - a_n}{b_n} \leq x\right) \to_{n \to \infty} \Lambda_\alpha$$

$$a_n = c(b_n), b_n = F^{-1}\left(1 - \frac{1}{n}\right).$$

soit F une fonction de Von Mises absolument continue sur un certain intervalle fini (t, x_F) avec La fonction auxiliaire notée :

$$a(x) = \frac{1 - F(x)}{f\,(x)} \tag{1.12}$$

La figure suivant expliquer les densités de loi GEV :

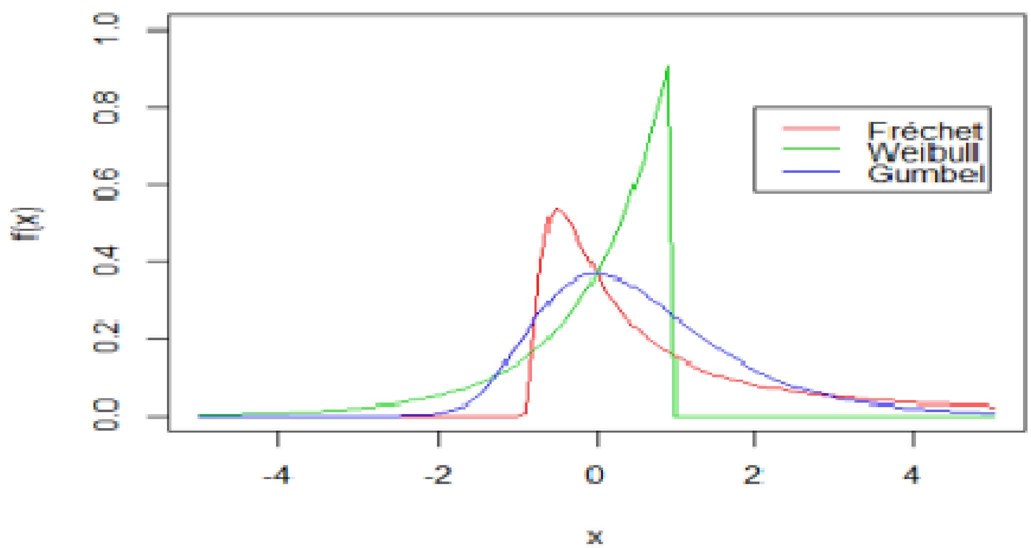

FIG. 1.1 – La densité de loi GEV,les courbes rouge, verte et bleu correspondent respectivement aux d'indices ξ= 1 etξ = -1,ξ = 0

Dans ce tableau on a expliquée quelques domaine d'attraction du certaine lois :

Domaine d'attraction	Gumbel ($\alpha = 0$)	Weibull (α<0)	Fréchet(α>0)
Lois	Normale Exponentielle Gamma Weibull Lognormale	Beta Uniforme	Cauchy. Parito .Student Cauchy loggamma

TAB. 1.1 – Les domaines d'attractions des quelques exemple de lois classés

1.3 Estimation de l'indice de queue

Les deux éstimateurs les plus populaires sont les éstimateurs de Hill (1975) et de Pickands (1975), il y'a aussi des méthodes d'éstimation comme la Méthode du maximum de vraisemblance, méthode des moments pondérés et la méthode des moments pondérés généralisés . On note $X_{1,n}, ..., X_{n,n}$ les statistiques d'ordre associées à $X_1, ..., X_n$ .n variable aléatoire (iid) de fonction de répartition F avec ξ l'indice de queue réel et k le nombre des excès $(1<k<n)$.

1.3.1 Estimateur de Hill

L'estimateur de Hill utilise plus pour domaine de Fréchet $(\alpha>0)$ et notée par la formule suivant :

$$\hat{\xi}_H^{(k,n)} = \frac{1}{k}\sum_{i=1}^{k} \log\left(X_{n-i+1,n}\right) - \log\left(X_{n-k,n}\right) \tag{1.13}$$

1.3.2 Estimateur de Pickands

Sa formule de l'éstimateur de Pickands est donnée par :

$$\hat{\xi}_P^{(k,n)} = \frac{1}{\ln 2} \ln \frac{X_{n-k+1,n} - X_{n-2k+1,n}}{X_{n-2k+1,n} - X_{n-4k+1,n}} \tag{1.14}$$

1.3.3 L'estimateur des moments

C'est un estimateur proposé par Dekkers *et al* [4]. Comme pour l'estimateur de Hill, il n'est utilisable que pour le domaine d'attraction de Fréchet $(\alpha>0)$ .Cet estimateur est défini par :

$$\hat{\xi}_m^{(k,n)} = 1 + \hat{\xi}_m^{(1)} - \frac{1}{2}\left(1 - \frac{\left(\hat{\xi}_m^{(1)}\right)^2}{\hat{\xi}_m^{(2)}}\right)^{-1} \tag{1.15}$$

avec $\hat{\xi}_m^{(i)} = \frac{1}{k}\sum_{i=1}^{k}\left(\log\left(X_{n-i+1,n}\right) - \log\left(X_{n-k,n}\right)\right)^i$, pour $(i=1)$ on trouve l'éstimateur de Hill $\hat{\xi}_H^{(k,n)}$.

1.3.4 Estimateur du maximum de vraisemblance

Soit l'échantillons $X = (X_1, X_2, ..., X_n)$ iid de densité $p_{\xi,\mu,\sigma}(X)$ avec $\theta = (\xi, \mu, \sigma)$ la formule de la fonction de vraisemblance est :

$$L(\theta, X) = \prod_{i=1}^{n} p_\theta(X_i)$$

L'éstimateur $\hat{\theta}$ est donné par la résolution suivant :

$$\begin{cases} \frac{\partial \log l(\theta)}{\partial \theta} = 0 \\ \frac{\partial^2 \log l(\theta)}{\partial^2 \theta} < 0. \end{cases}$$

D'autre part si $\xi = 0$ (loi de Gumbel) la fonction logarithmique de vraisemblance notée :

$$\log L(0, \mu, \sigma, X) = -n \log \sigma - \sum_{i=1}^{n} \exp\left(-\frac{x_i - \mu}{\sigma}\right) - \sum_{i=1}^{n} \frac{x_i - \mu}{\sigma},$$

par la dirivation de cette fonction de deux paramètre , nous trouvons la forme suivant pour résoudre :

$$\begin{cases} \frac{\partial \log l}{\partial \sigma} = 0 \Leftrightarrow n + \sum_{i=1}^{n} \frac{x_i - \mu}{\sigma}\left[\exp\left(-\frac{x_i - \mu}{\sigma}\right) - 1\right] & = 0 \\ \frac{\partial \log l}{\partial \mu} = 0 \Leftrightarrow n - \sum_{i=1}^{n} \exp\left(-\frac{x_i - \mu}{\sigma}\right) & = 0 \end{cases}$$

1.4 La méthode Peaks Over Theshold

1.4.1 Définitions

La fonction de répartition

La méthode des exés ou le dépassement d'un seuil sont formé à observer toutes les valeurs d'axcédent de certain seuil élevé, le but est d'analyser leurs comportement asymptotique.

Soit $X_1, X_{2,}..X_n$, n variable aléatoire (iid) de fonction de répartition F,on définit les excés par $y_j = x_j - u$ (u appelé le seuil). La formule de la fonction de répartition est,

$$\begin{aligned} F_u(y) &= \Pr(X - u \leq y / X > u) \\ &= \frac{F(u+y) - F(u)}{1 - F(u)} \end{aligned}$$

pour $\quad 0 \leq y \leq x_F - u$, ce équivalent avec :

$$F_u(x) = \Pr(X \leq x / X > u) = \frac{F(x) - F(u)}{1 - F(u)}, \text{ pour } \; x \geq u$$

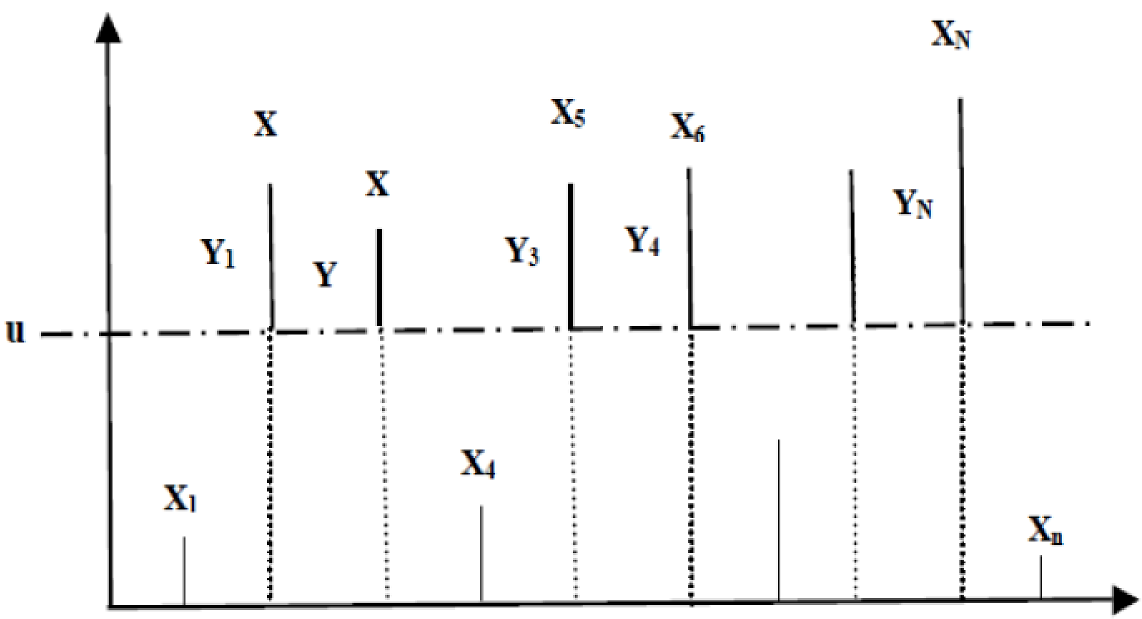

FIG. 1.2 – Méthode des excés

1.4.2 Théoréme de pickands-Belkema-de Haan

L'objectif de ce théoréme est pour la modélisation des excés ,elle est lié entre la loi GPD et le comportement asymptotique de la distribution des excés ,soit l'échantillons : $X_1, ..., X_n$

n variable aléatoire (iid) comune de distribution de F

Théorème 5 *Soit F appartient au domaine d'attraction de la distribution limite des queues associé avec la fonction des exés $F_u(y)$ ssi s'il exist un fonction positive $\varphi(u)$ tell que :*

$$\lim_{u \to y_F} \sup_{0<y_F<y_F-u} \left| F_u(y) - H^{GPD}_{\xi,\varphi(u)}(y) \right| = 0$$

et $H^{GPD}_{\xi,\varphi(u)}$ la distriution de Paréto généralisée (GPD) avec $(y \geq 0)$ notée par :

$$H^{GPD}_{\xi,\varphi(u)} = \begin{cases} 1 - \left(1 - \xi \frac{y}{\varphi}\right)^{-\frac{1}{\xi}} & si \quad \xi \neq 0 \\ 1 - \exp\left(-\frac{y}{\varphi}\right) & si \quad \xi = 0. \end{cases}$$

Chapitre 2

La distribution de Gumbel

2.1 La présentation de distribution de Gumbel

Dans la théorie des probabilités et les statistiques, la distribution de Gumbel (distribution généralisée des valeurs extrêmes de type I) est utilisée pour modéliser la distribution du maximum (ou du minimum) d'un certain nombre d'échantillons de diverses distributions. Cette distribution peuvent décrire des phénomènes éxtremes de durée de vie,de couts de sinistres en assurance,de crues aussi utilisée pour représenter la distribution du niveau maximum d'une rivière ,etc .D'une inondation ou d'une autre catastrophe naturelle. L'applicabilité potentielle de la distribution de Gumbel pour représenter la distribution des maxima est liée à la théorie des valeurs extrêmes, ce qui indique qu'elle sera probablement utile si la distribution des données d'échantillon sous-jacentes est de type normal ou exponentiel. Le reste de cet partie se réfère à Gumbel pour modéliser la distribution de la valeur maximale.La distribution de Gumbel (loi de Fisher -Typpet de type I ou loi double exponentiel) est un cas particuliere de la distribution des valeurs extremes géneralisées (GEV) avec l'alternativity parfois utilise pour se référer la distribution de Laplace il est lié à la loi de Gompertz et nommé d'aprés Emil Julius Gumbel (1891-1966) d'après ses articles originaux décrivant la distribution.Il est à remarquer que plus le nombre de paramètres d'une loi est grand, plus l'incertitude dans l'estimation est importante. Pratiquement il est par conséquent préférable d'éviter l'utilisation de lois à trois paramètres ou plus. Il est à remarquer que plus le nombre de paramètres d'une loi est grand, plus l'incertitude dans l'estimation est importante. Pratiquement il est par conséquent préférable d'éviter l'utilisation de lois à trois paramètres ou plus. Est la forme limite de la distribution de la valeur maximale d'un échantillon de n valeurs. Le maximum annuel d'une variable étant considéré comme le maximum de 365 valeurs journalières, cette loi doit ainsi être capable de décrire les séries de maxima annuels.La loi de Gumbel est trés utilisée en hydrologie et en climatologie pour éstimer les valeurs extrémes des phénomènes.

2.1.1 Expréssion mathématique

Fonction de répartition

Une variable aléatoire continue x suit une loi de Gumbel si sa fonction de répartition $F(x)$ ou représente la cdf (cumulative distribution function)) lorsque x varie de $-\infty$ à $+\infty$ (l'étendu) nous l'écrivant :

$$F(x) = \exp\left(-\exp\left(-\frac{x-\alpha}{\beta}\right)\right) \tag{2.1}$$

– Le pdf de la distribution de Gumbel est donné par :

$$F(x) = \frac{1}{\beta}\exp\left(-\frac{x-\alpha}{\beta} - \exp\left(-\frac{x-\alpha}{\beta}\right)\right)$$

avec $\beta\rangle 0$

Fonction de densité

Donnés par la formule suivant :

$$f(x) = \frac{1}{\beta} \exp\left(-\frac{x-\alpha}{\beta}\right) \exp\left(-\exp\left(-\frac{x-\alpha}{\beta}\right)\right) \tag{2.2}$$

avec α le parametre de position ou le mode

β le parametre d'échélle ou de dipersion

2.2 Les proprietés de la distribution de Gumbel

Cette distribution est caractérisée par des types suivant :

2.2.1 La moyenne :

On à notée la moyenne avec $\gamma = 0.5772...$(constante d'Euler-Mascheroni) par :

$$\mu = E(x) = \alpha + \beta\gamma \tag{2.3}$$

2.2.2 La variance :

Soit $V(x)$ désigne la variable aléatoire de la variance

$$\sigma^2 = V(x) = \frac{\pi^2}{6}\beta^2, \tag{2.4}$$

et l'étendu : $(-\infty \leq x \leq +\infty)$

2.2.3 Le médiane :

notée par :

$$Me = \alpha + \beta \ln(\ln 2)$$

2.3 Caractiristique de la distribution de Gumbel

Certaines des caractéristiques spécifiques de la distribution de Gumbel sont :

- La forme de la distribution de Gumbel est biaisée vers la gauche. Le pdf Gumbel n'a pas de paramètre de forme. Cela signifie que la pdf de Gumbel n'a qu'une seule forme, qui ne change pas.
- Le pdf de Gumbel a un paramètre de localisation α qui est égal au mode mais diffère de la médiane et de la moyenne. C'est parce que la distribution de Gumbel n'est pas symétrique par rapport à son α.
- Lorsque α diminue, le pdf est décalé vers la gauche. Lorsque α augmente, le pdf est décalé vers la droite.
- Si x a une distribution de Gumbel, alors la distribution conditionnelle de $Y = -x$ étant donné que Y est positif, ou de manière équivalente étant donné que x est négatif, a une distribution de Gompertz. Le cdf G de Y est lié à F, la cdf de x, par la formule

$$G\left(y\right) = P\left(Y \leq y\right) = P\left(X \geq -y \backslash x \leq 0\right) = \left(F\left(0\right) - F\left(-y\right)\right)/F\left(0\right)$$

  ,pour $y > 0$. Par conséquent, les densités sont liées par $g\left(y\right) = f\left(-y\right)/F\left(0\right)$,la densité de Gompertz est proportionnelle à une densité de Gumbel réfléchie, limitée à la demi-droite positive.
- Si x une exponentielle est avec la moyenne 1, alors $-ln\left(x\right)$ a une distribution Gumbel standard.

Remarque 1 *Si le cas de la loi standard de Gumbel est* $\{\alpha = 0, \beta = 1\}$ *Alors :*

$$F\left(x\right) = \exp\left(-\exp\left(-x\right)\right) \tag{2.5}$$

$$f\left(x\right) = \exp\left(-\left(x + \exp\left(-x\right)\right)\right)$$

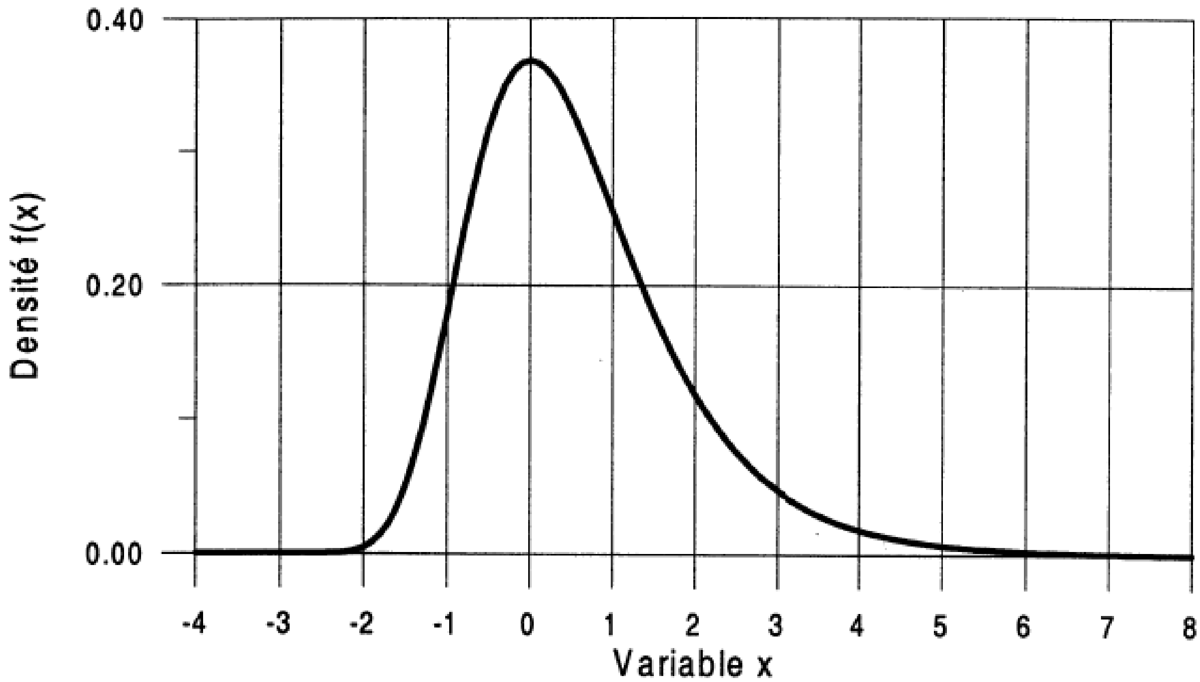

FIG. 2.1 – Fonction de densité f(x) de la loi de gumbel de paramètre $\alpha = 0, \beta = 1$

Dans ce cas

le mode : 0

le médiane est : $\quad -\ln(\ln 2) = 0.3665,$

la moyenne est : $\gamma = 0.5772..$*et l'écart type est :* $\frac{\pi}{\sqrt{6}}$.

2.3.1 La relation entre loi de Gumbel et période de retour

Soit T la période de retour d'une valeurs x et $F(x)$ désigne la fonction de répartition de Gumbel alors cette relation notée par :

$$T = \frac{1}{1 - F(x)}$$

$$\Pr(X \geq x_T) = 1 - F(x_T) = \frac{1}{T} \tag{2.6}$$

avec la valeur x_T est appelée le quantile de période de retour T et est une fonction de T

Exemple 4 *Soit x une variable aléatoire distribuée selon une loi Gumbel de paramètres α et β. Selon la fonction de répartition correspondante est donnée par :*

$$F(x_T) = \exp\left(-\exp\left(-\frac{x_T - \alpha}{\beta}\right)\right)$$

Le quantile x_T de période de retour$T$ de cette variable aléatoire est tel que :

$$1 - \exp\left(-\exp\left(-\frac{x_T - \alpha}{\beta}\right)\right) = \frac{1}{T}$$

En appliquant la transformé logarithmique deux fois de chaque côté de l'équation et en isolant x_T on a que :

$$x_T = \alpha - \beta \ln\left[-\ln\left(1 - \frac{1}{T}\right)\right] \tag{2.7}$$

2.3.2 La variable réduite u

On a donnée la distribution et la fréquence cumulée : $F(x) = \exp(-\exp(-u))$ telque u la variable réduite notée :

$$u = \frac{x - \alpha}{\beta}$$

alors on à donnée

$$u = -\ln\left[-\ln F(x)\right]$$

2.3.3 Expréssion d'un quantile

Pour trouver la valeur x_q (quantile) correspondant à la fréquence cumulée $F(x_q) = q$,en fonctions des deux paramètres α, β :

$$x_q = \alpha + \beta u_q \tag{2.8}$$

il suffit d'inverser la variable réduite puis en remplacer u_q par :$- \ln[- \ln F(x)]$

2.4 Estimation pour la loi de Gumbel

2.4.1 L'ajustement de la loi de Gumbel

Est une méthode statistique de prédiction consistant à étudier les événements passés,propriètè d'un processus donné(hydrologique ou autre) afin d'en définir les probabilités d'apparition future,cette prédiction repose sur la définition et la mise en oeuvre d'un modéle fréquentiel .qui est une équation modélisant le comportement statistique d'un processus .

La dernière étape de la modélisation est la spécification ou l'ajustement du modéle c'est -à-dire la détermination de ses paramètre . On explique donc à ce stade que la validité de l'analyse fréquentielle repose bien d'avantange sur la pertinence du modèle adapté que sur la technique numérique d'ajustement utiliser deux méthode d'ajustement

– Méthode analytique

– Méthode graphique

l'ajustement repose sur le fait que l'expréssion d'un quantille correspond à l'équation ,d'une droite .En conséquence les points de la série à ajuster peuvent étre reporté dans un systeme d'axés :il est alors possible de tracer la droite qui passe le mieux par ces points et en déduire les deux paramètre et définissant la loi .Il s'agit donc essentiellement d'estimer la probabililé de non -dépassement $F(x)$ qu'il convient d'attribuer à chaque valeurs x_i .

Elles reposent toutes sur un tri de la série par valeurs croissantes permettant d'associer à chaque valeur son rang r .

En analyse fréquentielle plusieurs méthodes ont été développées pour l'estimation des paramètres des distributions d'intérêt c'est :

- Le maximum de vraisemblance ;
- La méthode des moments ;
- La méthode des moments de probabilités pondérées (PWM) ;
- La méthode des L-moments (LM)

On a fait l'éstimation pour la distribution de Gumbel par les méthode suivant :

2.4.2 Méthode des moments

Cette méthode consiste à égaler les moments échantillonnaux et les moments théoriques de la loi choisie ,soit $x_1, x_2, ..., x_n$ l'échantillons de donnée à disposition on à l'éstimateur standard de la moyenne et de la variance :

$$\begin{aligned} \hat{\mu} &= \bar{x} = \tfrac{1}{n}\sum_{i=1}^{n} x_i \\ \hat{\sigma} &= s^2 = \tfrac{1}{n-1}\sum_{i=1}^{n} (x_i - \bar{x})^2 \end{aligned}$$

Pour les deux premiers moments théoriques de la loi de Gumbel,s'utilise à partir des paramètres de position et d'échèlle de la manière suivant :

$$\begin{aligned} \mu &= \alpha + \beta\gamma \\ \sigma^2 &= \tfrac{\pi^2}{6}\beta^2 \end{aligned}$$

par la méthode des moments on a utilise les formules suivant pour l'éstimation des paramètre $\hat{\alpha}, \hat{\beta}$:

$$\begin{aligned} \hat{\alpha} &= \hat{\mu} - \hat{\beta}\gamma \\ \hat{\beta} &= \tfrac{\sqrt{6}}{\pi}\hat{\sigma} \end{aligned}$$

2.4.3 Méthode du maximum de vraisemblance

Cette méthode consiste à considérer que les paramétres inconnus de la loi à ajuster est notée par :

$$L = \pi_{i=1}^{n} f\left(x_i, \alpha, \beta\right)$$

ou f la fonction de densité de la loi .

il s'agit de trouver les valeurs α, β :

soit :

$$\frac{\delta \ln L}{\delta \alpha} = 0$$

$$\frac{\delta \ln L}{\delta \beta} = 0$$

Pour l'éstimation des paramétres $\hat{\alpha}, \hat{\beta}$

$$\hat{\alpha} = \hat{\beta} \ln \left(\frac{n}{\sum_{i=1}^{n} \exp\left(-\frac{x_i}{\hat{\beta}}\right)} \right)$$

$$\hat{\beta} = \bar{x} - \frac{\sum_{i=1}^{n} x_i \exp\left(-\frac{x_i}{\hat{\beta}}\right)}{\sum_{i=1}^{n} \exp\left(-\frac{x_i}{\hat{\beta}}\right)}$$

2.4.4 Méthode des moindres réctangles

Cette méthode revient donc à minimiser la distance du point à sa projection orthogonale sur la droite de régression $Y = \alpha X + \beta$ avec $X = -\ln\left(-\ln\left(F\left(Y\right)\right)\right)$

alors dans le cas de la loi de Gumbel l'axe est remplacé par l'axe de la variable réduite de Gumbel et l'axe par celui de la variable hydrologique on notée par cette méthode les éstimateurs suivant :

$$\hat{\beta} = \frac{S_X}{S_Y}$$

$$\hat{\alpha} = \bar{x} - \beta \bar{Y}$$

2.4.5 Méthode des L.moment

Le but de cette méthode est de réaliser un ajustement lorsque les moments classiques ne conviennent pas. Les deux paramètres et sont obtenus très simplement à partir des valeurs des deux premiers L-moments de la loi de Gumbel et des estimations calculées sur l'échantillon :

$$\begin{cases} \hat{\beta} = \frac{\lambda 2}{\ln 2} \\ \hat{\alpha} = \hat{\lambda}_1 - \hat{\beta}\gamma \end{cases}$$

Exemple 5 *Le tableau suivant présenté les pluis journalières maximum annuelles à la station de Lutry*

chronologiquement		Par valeurs croissantes				
Année (1)	Pj_{max} (2)	Année (3)	Pj_{max} (4)	Rang r (5)	$\hat{F}$ (6)	u (7)
69	46.8	80	35.2	1	0.033	-1.22
70	57.8	72	36.2	2	0.100	-0.83
71	51.5	83	37.6	3	0.167	-0.58
72	36.2	69	46.8	4	0.233	-0.38
73	53.4	82	49.5	5	0.300	-0.19
74	62.0	71	51.5	6	0.367	0.00
75	62.6	73	53.4	7	0.433	0.18
76	61	79	54.9	8	0.500	0.37
77	58.2	70	57.8	9	0.567	0.57
78	85.7	77	58.2	10	0.633	0.78
79	54.9	76	61	11	0.700	1.03
80	35.2	74	62	12	0.767	1.33
81	64.6	75	62.6	13	0.833	1.70
82	49.5	81	64.6	14	0.900	2.25
83	37.6	78	85.7	15	0.967	3.38

TAB. 2.1 – Les pluis journalières maximum annuelles à la station de Lutry

par la méthode des moments on a utilise les formules suivant pour l'éstimation des paramètre $\hat{\alpha}, \hat{\beta}$: Pour l'exemple du tableau 2.1 l'échantillons présente les caractiristiques suivants :

$$\begin{cases} \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = 54.4667 \\ s = 12.5214 \end{cases}$$

les éstimateurs pour la population sont alors :

$$\begin{cases} \bar{x} = 54.4667 \\ \hat{\sigma} = s\sqrt{\frac{n}{n-1}} = 12.5214\sqrt{\frac{15}{15-1}} = 12.96 \end{cases}$$

La méthode des moments consiste à égaler les moments des échantillons avec les moments théoriques de la loi de gumbel. Par la méthode des moments les paramètres $\hat{\alpha}$ *et* $\hat{\beta}$ *sont calculés d'après les formules :*

$$\begin{cases} \hat{\beta} \; = \frac{\sqrt{6}}{\pi}12.96 = 0.7797 * 12.96 = 10.1 \\ \hat{\alpha} = \hat{\mu} - \hat{\beta}\gamma = 54.47 - 10.1 * 0.5773 = 48.6 \end{cases}$$

Par la Méthode des moindres réctangles On va continue avec le tableau 2.1 nous obtenons : puis

$$\begin{cases} \hat{\beta} = \frac{S_X}{Su} = \frac{12.5214}{1.1933} = 10.5 \\ \hat{\alpha} \; = \bar{x} - \; \hat{\beta}\bar{u} = 54.4667 - 10.5 * 0.558 = 48.6 \end{cases}$$

Conclusion

La théorie des valeurs extrêmes au cours des dernières années a fait l'objet d'une grande attention théorique et pratique à l'étude des comportements asymptotique des grands observation des échantillons des variables aléatoires iid et comme la théorème de Fisher et Tippet donnent un cas c'est la loi de Gumbel cette dernière est plus utilisée et appliquer en hydrologie et climatologie pour éstimer des valeurs extremes et décrire des phénomènes éxtremes de durée de vie,de couts de sinistres en assurance,de crues .D'une inondation ou d'une autre catastrophe naturelle

Bibliographie

[1] Beirlant . J.Goegebeur,Y,Segers,J et Teugels,J (2004).Statistic of Etremes Théory and Applications.Wiley

[2] Coles, S. (2001). An Introduction to Statistical Modeling of Extremes Values : Springer, Londres.

[3] De Haan L.Ferreira.A(2006) Extreme Value Theory.Springer -Verlag

[4] Embrechts P.Kluppelberg C,Mikosch.T(1997).Modeling Extremal Event for Insurance and Finance .Berlin-Spinger

[5] GianFausto.S.and .R.ross(2007).Extremes in nature.university delSalento.Italy

[6] Gumbel,E.J.(1958).Statistics of Extremes.Colombia.université,New york

[7] J.Galambos,J.Lechner and ,E.Simiu(1994).Extreme Value Theory and Application ,Boston , London

[8] Jonathan El Methni(2013).Contribution à l'estimaion de quantiles extremes Application à des donners environnnemontales ,Université de Grenoble

[9] Julie carreau(cours 3-2014).Théorie des valeurs .Hydro-siences.Montpellier

[10] Louiza Soltane (2017) .Analyse des valeurs extremes en presences de censure ,mémoire de doctorat ,université de Biskra

[11] Necir.A (2017-2018).Cours de 2ème année master

[12] P.Dubreuil (1974) Initiation à l'analyse hydrologique .université de Paris

[13] Pascal Lezaud (2005).Introduction à la théorie des valeurs extremes,mémoire de doctorat ,université de Nice

[14] Pathé.Ndao().Modélisation des valeurs extremes conditionnelles en présence de censure,mémoire de doctorat .Université Gaston .Berger de Saint -Louis

[15] Paul .Meylan(1999).Hydrologie fréquentielle .Ecole Polytechnique Fédérale de Lausanna-suisse

[16] Pikands.J.(1975).Statistical inférence using extreme order statistics.Ann-statist3

[17] R.D.Reiss.M.Thomas(1997).Statistical Analysis of Extreme Values .Boston.Berlin

[18] Samuel.Kotz (2000) .Extreme value distribution(theory and application).The university of Nottingham

[19] Y.Brunet-Moret(1969).Etude de quelques lois statistiques utilisées en hydrologie O.R.S.T.O.M

Annexe B : Abréviations et Notations

Les différentes abréVariations et notations utilisées tout au long de ce mémoire sont expliquées ci-dessous.

TVE	Théorie des valeurs extrêmes
GEV	Distribution des valeurs extrêmes généralisée
GPD	Distribution de Paréto généralisée
POT	Peaks-Over-Threshold
iid	Indépendantes et identiquement distribuées.
x_F	Point terminal.
F_n	Fonction de répartition empirique.
$\rightarrow^{loi}$	Converge en loi
x_T	Le qantil de périod de retour
cdf	Cumulative distribution function
pdf	prababilité distribution function
TCL	Téorème central limite