Examen Théorie des Valeurs Extrêmes 2013-2014

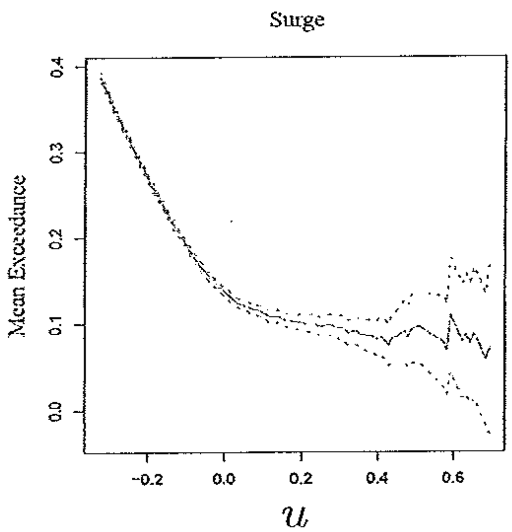
Master 2 SAF Pro

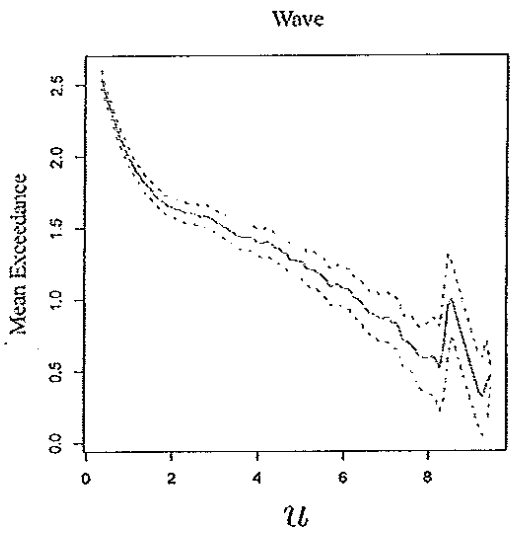
Sans document - Sans calculatrice - 1h30

Question de cours (8 points):

Répondre aux questions suivantes de manière la plus littérale possible, c'est-à-dire avec le minimum de mathématique.

1. Qu'appelle-t-on une distribution à queue épaisse et comment l'identifie-t-on dans la pratique?
2. Quel lien relie les distributions GEV (Generalised Extreme Value distibutions) et les distributions GPD (Generalised Pareto distributions)?
3. Quelles sont les procédures statistiques à utiliser pour estimer les paramètres des distributions GEV (discuter suivant la nature des observations)?
4. Quelle utilisation fait-on de la fonction de dépassement moyen empirique? Vous trouverez ci-dessous deux représentations graphiques de la fonction de dépassement moyen empirique pour deux grandeurs physiques (Surge et Wave) en fonction d'un seuil u. Quel comportement de la fonction de dépassement est attendu pour les grandes valeurs de u?





Exercice 1 (8 points):

On considère la distribution donnée par

$$F(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1 - \exp\left(-\frac{x}{1-x}\right) & \text{si } 0 \leq x \leq 1, \\ 1 & \text{si } x > 1. \end{cases}$$

Soient $X_1, ..., X_n$ une suite de variables aléatoires IID de distribution F et $M_n = \max(X_1, ..., X_n)$.

0. Est-ce que M_n peut avoir une distribution non dégénérée avec une suite de seuils non-linéaires?

1. En choisissant

$$a_n = \frac{1}{(1+\ln n)^2} \quad \text{et} \quad b_n = \frac{\ln n}{1+\ln n},$$

trouver la loi limite de $a_n^{-1}(M_n - b_n)$.

Soient $E_1, ..., E_n$ une suite de variables aléatoires IID de distribution exponentielle de paramètre 1 et $M_n^E = \max(E_1, ..., E_n)$.

2. Donner les constantes de normalisation a_n^E et b_n^E telles que $(a_n^E)^{-1}\left(M_n - b_n^E\right)$ converge en distribution vers la loi de Gumbel.

3. On considère la fonction

$$g(x) = \frac{x}{1+x}, \quad x > 0.$$

a. Montrer que $X_1 \stackrel{d}{=} g(E_1)$. En déduire que

$$\frac{M_n - b_n}{a_n} \stackrel{d}{=} \frac{g\left(M_n^E\right) - g(\ln n)}{g'(\ln n)}.$$

b. Montrer qu'il existe ζ_n tel que

$$g\left(M_n^E\right) - g(\ln n) = \left(M_n^E - \ln n\right) g'\left(\zeta_n\right)$$

avec $\ln n \leq \zeta_n \leq M_n^E$ si $\ln n \leq M_n^E$ et inversement.

c. Montrer que

$$\frac{M_n^E}{\ln n} \xrightarrow{P} 1,$$

et en déduire que

$$\frac{g'\left(\zeta_n\right)}{g'(\ln n)} \xrightarrow{P} 1.$$

Retrouver le résultat de la question 1.

Exercice 2 (4 points):

Soient $X_1, ..., X_n$ une suite de variables aléatoires indépendantes de loi Gaussienne centrée et réduite, et $M_n = \max(|X_1|, ..., |X_n|)$. On a dans ce cas là:

$$\frac{M_n - b_n}{a_n} \xrightarrow{d} \Lambda$$

où Λ a une distribution de Gumbel et

$$\begin{aligned} a_n &= (2\ln(2n))^{-1/2} \\ b_n &= (2\ln(2n))^{1/2} - \frac{1}{2}(2\ln(2n))^{-1/2}[\ln(\ln(2n)) + \log(4\pi)]. \end{aligned}$$

1. Montrer que

$$\frac{M_n}{b_n} \xrightarrow{P} 1.$$

On cherche à caractériser la loi asympotique du maximum de variables aléatoires indépendantes avec une distribution du Khi-deux à un degrès de liberté.

2. En utilisant une identité remarquable, montrer que si

$$\frac{\max(|X_1|, ..., |X_n|) - b_n}{a_n} \xrightarrow{d} \Lambda,$$

alors

$$\frac{\max(X_1^2, ..., X_n^2) - b_n^2}{2a_nb_n} \xrightarrow{d} \Lambda.$$

En déduire les constantes de normalisation pour la convergence du maximum de variables aléatoires indépendantes avec une distribution du Khi-deux à un degrès de liberté.