Examen Théorie des Valeurs Extrêmes 2016-2017

ENSAE

Sans document - Sans calculatrice - 2h00

Questions de cours et exercices d'applications du cours (6 points):

1. Que signifie F appartient au domaine d'attraction d'une distribution max-stable G ($F \in D(G)$)?

On rappelle que la distribution de Pareto généralisée GPD(β, ξ) est définie par

$$G^{p}_{\xi,\beta}(x) = \begin{cases} 1 - [1 + \xi(x/\beta)]_+^{-1/\xi} & \text{si } \xi \neq 0, \\ 1 - e^{-x/\beta} & \text{si } \xi = 0, \end{cases}$$

Expliquer à quel domaine d'attraction appartient la distribution de Pareto généralisée GPD(β, ξ).

2. Soit B une variable aléatoire de distribution Binomiale(n, p)

$$P(B = k) = C_n^k p^k (1-p)^{n-k}$$

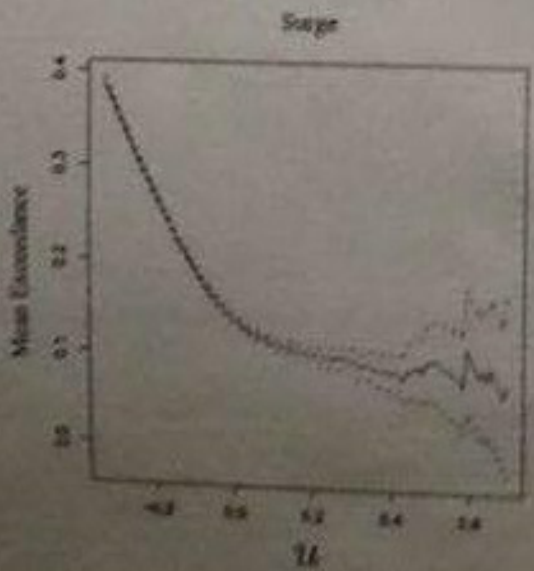
et N une variable aléatoire de distribution Poisson(λ). On pose $\lambda = np$. Montrer que

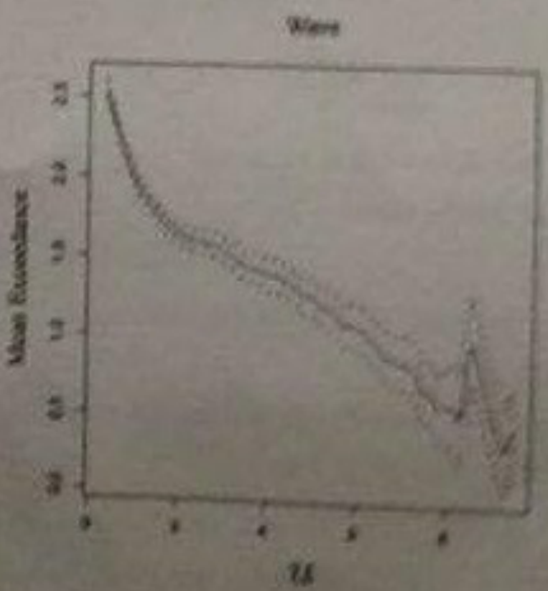
$$P(B = k) = P(N = k)\left(\frac{\left(1 - \frac{\lambda}{n}\right)^n}{e^{-\lambda}} \frac{1}{\left(1 - \frac{\lambda}{n}\right)^k} \prod_{i=1}^{k}\left(1 - \frac{i-1}{n}\right)\right)$$

En déduire que si $n \to \infty$, $p \to 0$ tel que $\lambda = np$, alors B converge en distribution vers N.

Expliquer comment ce résultat peut être utilisé pour caractériser la loi du maximum ou des statistiques d'ordre élevé.

3. 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 (6 points):

A. On considère la distribution $F(x) = 1 - e^{1/x}$ pour $x < 0$, et $F(x) = 1$ pour $x \geq 0$.

1. Montrer que la suite

$$u_n = \frac{1}{\ln \tau - \ln n}$$

assure que la condition $n\bar{F}(u_n) \to \tau > 0$ est satisfaite. En déduire la limite de $P(M_n \leq u_n)$ si $M_n = \max(X_1, ..., X_n)$ avec (X_i) i.i.d. de loi F quand $n \to \infty$.

2. En prenant $\tau = e^{-x}$ et en utilisant un développement limité, en déduire qu'il existe $(a_n) > 0$ et (b_n) tel que

$$\lim_{n\to\infty} P(M_n \leq a_n x + b_n) = \exp(-\exp(-x)).$$

B. On considère la distribution $F(x) = K(1 - e^{-x})$ pour $0 < x < x_F$.

1. Donner la valeur de K en fonction de x_F.

2. Trouver $(a_n) > 0$ tel que, pour $x < 0$,

$$\lim_{n\to\infty} P(M_n \leq a_n x + x_F) = \exp(x).$$

Que concluez-vous de cet exercice?

Exercice 2 (5 points):

Soit (X_i) des variables aléatoires i.i.d. de loi F. On suppose qu'il existe deux suites $(u_n^{(1)})$ et $(u_n^{(2)})$ telles que pour $i = 1, 2$

$$n\bar{F}(u_n^{(i)}) \to \tau_i$$

avec $0 < \tau_1 \leq \tau_2$.

1. On définit $N_n^{(i)} = \sum_{i=1}^n \mathrm{I}_{\{X_i \leq u_n^{(i)}\}}$. Il est possible alors de montrer que

$$\lim_{n\to\infty} P\left(N_n^{(1)} = k_1, N_n^{(2)} = k_1 + k_2\right) = \frac{\tau_1^{k_1}}{k_1!}\frac{(\tau_2 - \tau_1)^{k_2}}{k_2!}e^{-\tau_2}.$$

Quel résultat retrouve-t-on si $\tau_1 = \tau_2$.

2. Exprimer l'évènement

$$\{X_{(1)} \leq u_n^{(1)}, X_{(2)} \leq u_n^{(2)}\}$$

en fonction de $N_n^{(1)}$ et $N_n^{(2)}$. On rappelle que $X_{(1)}$ et $X_{(2)}$ sont les deux plus grandes statistiques d'ordre de $(X_i)_{i=1,...,n}$.

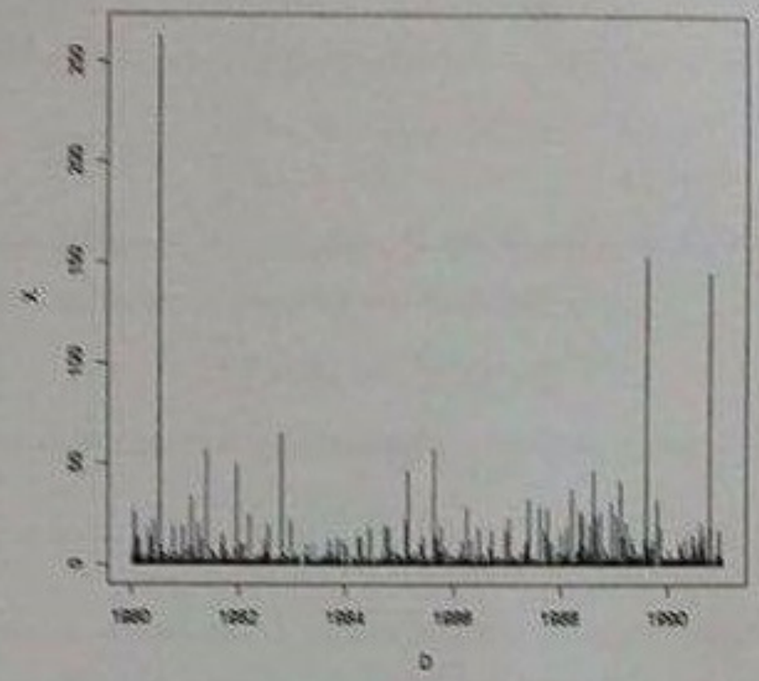
3. En supposant que F appartient au domaine d'attraction de G pour deux suites $(a_n) > 0$ et (b_n), montrer que pour $x_1 > x_2$

$$\begin{aligned}
&\lim_{n\to\infty} P(X_{(1)} \leq a_n x_1 + b_n, X_{(2)} \leq a_n x_2 + b_n) \\
= \;& G(x_2)\{\ln G(x_1) - \ln G(x_2) + 1\}.
\end{aligned}$$

Quel est l'intérêt de cette distribution?

Question méthodologique (3 points):

Un réassureur dispose des données de sinistres suivantes pour tarifier un excédent de sinistres de priorité 100 et de portée illimitée.



On rappelle qu'en cas de sinsitre X_i, le réassureur rembourse à l'assureur le montant $(X_i - 100)_+$. Quelle méthodologie lui proposez-vous? Donner une réponse littérale.