M2 Statistique 2014-2015
Parcours « Biostatistique et Statistiques Industrielles » et « Actuariat »

Statistique des valeurs extrêmes *(1ère Session)*

(*Durée : 2h00. Seule la calculatrice est autorisée.*)

Notations utilisées dans le sujet :

- Pour tout échantillon $W_1, \ldots, W_N$, on note $W_{1,N} \leq \ldots \leq W_{N,N}$ l'échantillon ordonné associé.
- Pour toute fonction $\Phi(.)$ croissante, $\Phi^{\leftarrow}(y) := \inf\{x | \Phi(x) \geq y\}$.
- $X \stackrel{\mathcal{L}}{=} Y$ signifie que X et Y ont même loi.

Questions de cours [4 pts] – Dans ces questions, $F(.)$ désigne une fonction de répartition de point terminal x_F (éventuellement infini).

1) Rappeler la définition du point terminal x_F.

2) Rappeler la définition du quantile $q(\alpha)$ d'ordre $\alpha \in]0, 1[$ associé à $F(.)$.

3) On suppose que $F(.)$ appartient au domaine d'attraction de Weibull avec un indice des valeurs extrêmes γ. Quel est le signe de γ ? Le point terminal peut-il être infini ? Donner en fonction de x_F, de γ et d'une fonction à variations lentes $L(.)$ l'expression de $F(.)$.

4) Soit X une variable aléatoire de fonction de répartition $F(.)$. Donner la définition de la fonction de répartition des excès de X au dessus d'un seuil $u < x_F$ ainsi que de la fonction *moyenne des excès*. Enoncer le théorème de Pickands.

Exercice 1 [5 pts] – On considère la fonction $F : \mathbb{R} \mapsto \mathbb{R}$ définie pour $\theta > 0$ et $\lambda > 0$ par

$$F(x) = \left(1 - (1 + x^{\theta})^{-\lambda}\right) \mathbb{I}_{\{x>0\}}.$$

1) Montrer que $F(.)$ est une fonction de répartition et donner son point terminal.

2) Montrer que $F(.)$ appartient au domaine d'attraction de Fréchet et donner, en fonction de θ et λ , l'indice des valeurs extrêmes γ associé.

3) On pose $L : [0, \infty[\mapsto [0, \infty[$ la fonction définie par $L(x) = x^{1/\gamma}(1 - F(x))$ pour tout $x > 0$. Quelle est la principale caractéristique de la fonction $L(.)$?

4) On pose $\Delta : [0, \infty[\mapsto [0, \infty[$ la fonction définie par $\Delta(x) = xL'(x)/L(x)$ pour tout $x > 0$ où $L'(.)$ désigne la dérivée de la fonction $L(.)$. Montrer que $\Delta(.)$ est une fonction à variations régulières dont vous préciserez l'indice.

On considère à présent n variables aléatoires indépendantes $X_1, \ldots, X_n$ de même fonction de répartition $F(.)$ et on pose $a_n = F^{\leftarrow}(1 - 1/n)$.

5) Déterminer la loi limite de la variable aléatoire

$$Y_n := \max \left\{ \frac{X_1}{a_n}; \ldots; \frac{X_n}{a_n} \right\}.$$

Expliquer pourquoi on retrouve ainsi le résultat de la question 2) ?

Exercice 2 [11 pts] – Soit $F : \mathbb{R} \mapsto [0, 1]$ une fonction de répartition de point terminal $x_F < +\infty$ et telle que pour tout $x < x_F$, $F(x) = 1 - (x_F - x)^{-1/\gamma} L\left((x_F - x)^{-1}\right)$, où $\gamma < 0$ et $L : [0, \infty[\frac{1}{2} \mapsto [0, \infty[$ est une fonction à variations lentes.

1) Montrer que la fonction $\bar{F}^* : [0, \infty[\frac{1}{2} \mapsto [0, 1]$ telle que $\bar{F}^*(x) = 1 - F(x_F - 1/x)$ pour tout $x > 0$ est une fonction décroissante et à variations régulières d'indice $1/\gamma$.

On rappelle que si $U : [0, \infty[\frac{1}{2} \mapsto [0, \infty[$ est une fonction décroissante et à variations régulières d'indice $\alpha < 0$ alors la fonction $U^{\leftarrow}(1/.)$ est à variations régulières d'indice $-1/\alpha$.

2) Après avoir remarqué que $(\bar{F}^*)^{\leftarrow}(x) = (x_F - F^{\leftarrow}(1 - x))^{-1}$ pour tout $x \in [0, 1]$, montrer qu'il existe une fonction à variations lentes $\ell(.)$ telle que pour tout $\alpha \in [0, 1]$

$$F^{\leftarrow}(1 - \alpha) = x_F - \alpha^{-\gamma} \ell(\alpha^{-1}).$$

On considère à présent n variables aléatoires $X_1, \dots, X_n$ indépendantes et de même fonction de répartition $F(.)$. On introduit également une suite $(k_n)_{n\in\mathbb{N}}$ telle que $k_n \to \infty$ et $k_n/n \to 0$ lorsque $n \to \infty$.

3) Montrer que $\{\frac{1}{2}X_{n-i+1,n},\ i = 1,\dots,k_n\} \stackrel{\mathcal{L}}{=} \{x_F - U_{i,n}^{-\gamma}\ell(U_{i,n}^{-1}),\ i = 1,\dots,k_n\}$, où $U_1,\dots,U_n$ sont des variables aléatoires indépendantes de loi uniforme sur $[0,1]$.

4) En déduire que

$$\frac{1}{k_n}\sum_{i=1}^{k_n}\log\left(\frac{x_F - X_{n-i+1,n}}{x_F - X_{n-k_n,n}}\right) \stackrel{\mathcal{L}}{=} \gamma\frac{1}{k_n}\sum_{i=1}^{k_n}\log\frac{U_{i,n}^{-1}}{U_{k_n+1,n}^{-1}} + \frac{1}{k_n}\sum_{i=1}^{k_n}\log\frac{\ell(U_{i,n}^{-1})}{\ell(U_{k_n+1,n}^{-1})}.$$

5) En utilisant la représentation de Rényi, montrer que

$$\sum_{i=1}^{k_n}\log\frac{U_{i,n}^{-1}}{U_{k_n+1,n}^{-1}} \stackrel{\mathcal{L}}{=} \sum_{i=1}^{k_n}E_i,$$

où $E_1,\dots,E_{k_n}$ sont des variables aléatoires indépendantes de loi exponentielle de paramètre 1.

On admettra pour la suite que

$$\frac{1}{k_n}\sum_{i=1}^{k_n}\log\frac{\ell(U_{i,n}^{-1})}{\ell(U_{k_n+1,n}^{-1})} \stackrel{\mathrm{P}}{\to} 0.$$

6) Montrer que

$$\frac{1}{k_n}\sum_{i=1}^{k_n}\log\left(\frac{x_F - X_{n-i+1,n}}{x_F - X_{n-k_n,n}}\right) \stackrel{\mathrm{P}}{\to} \gamma.$$

7) En supposant que la fonction à variations lentes $L(.)$ converge à l'infini vers une constante $c > 0$, montrer que $n^{-\gamma}(x_F - X_{n,n})$ converge en loi (vous préciserez la loi limite). Qu'en déduisez-vous sur la convergence en probabilité de $X_{n,n}$?

8) Proposer un estimateur de γ (on ne cherchera pas à montrer que cet estimateur converge en probabilité vers γ).