PARTIEL
MARS 2013

La durée de l'examen est de 3 heures. Vous me rendrez :
- sur une copie, les réponses aux questions théoriques,
- le listing *commenté* des fonctions ou des scripts dont vous aurez eu besoin,
- les graphiques (n'oubliez pas qu'au lieu de différencier les courbes par leur couleur, vous pouvez changer la matière du trait).

Exercice 1. *Valeurs extrêmes*

Soit $x_0 > 0$ et $k > 0$. La loi de Pareto de paramètres x_0 et k est définie comme la loi de fonction de répartition $F(x) = (1 - (x/x_0)^{-k})\mathbb{1}_{x \geq x_0}$.

(1) Ecrire une fonction qui prend en argument n et N et renvoit une matrice de taille $n \times N$ de variables aléatoires indépendantes de loi de Pareto de paramètres $x_0 = 2$ et $k = 20$.

(2) Calculer théoriquement l'espérance d'une loi de Pareto de paramètres x_0 et $k > 1$.

(3) Illustrer graphiquement la loi des grands nombres pour un échantillon de Pareto.

(4) On cherche maintenant à calculer l'intégrale suivante
$$I = \int_1^\infty \arctan(2x) x^{-21} dx.$$
Expliquer quelle méthode numérique il est possible d'utiliser ici, et donner une valeur approchée de I.

(5) Dans le point précédent, déterminer théoriquement la taille de l'échantillon nécessaire pour obtenir, avec une confiance de 95%, une estimation affectée d'une erreur au plus de 10^{-2}.

Le but de la suite de cet exercice est d'illustrer le résultat suivant.

Théorème. *(Gnedenko, 1943)*
Soit $X_1, \ldots, X_n$ des variables indépendantes de même fonction de répartition F. On note $M_n = \max(X_1, \ldots, X_n)$. Sous des conditions générales sur F, il existe trois paramètres a_n, b_n et γ tels que :
$$\forall x \in \mathbb{R}, \quad \lim_{n\to\infty} \mathbb{P}\left(\frac{M_n - a_n}{b_n} \leq x\right) = H_\gamma(x),$$
avec
$$H_\gamma(x) = \begin{cases} \exp(-x_+^{-1/\gamma}) & \text{si } \gamma > 0 \\ \exp(-e^{-x}) & \text{si } \gamma = 0 \\ \exp(-(-x)_+^{-1/\gamma}) & \text{si } \gamma < 0 \end{cases}$$
où $y_+ = \max(0, y)$.

(6) Donner la fonction de répartition de $(M_n - a_n)/b_n$ en fonction de F.

(7) On suppose que F est la fonction de répartition d'une loi de Pareto de paramètres x_0 et k. On choisit $a_n = 0$, $b_n = x_0.n^{1/k}$. Montrer théoriquement le théorème dans ce cas (avec un γ à déterminer).

(8) Simuler un N-échantillon de même loi que M_n dans le cas où F est la distribution de Pareto de paramètres 2 et 20.

(9) Illustrer graphiquement le théorème dans ce cas, en comparant la fonction de répartition empirique de votre échantillon à la fonction $H_{1/k}$ pour n grand.

(10) On suppose maintenant que F est la distribution uniforme sur $[0,1]$. On prend $a_n = 1$, $b_n = 1/n$. Montrer théoriquement que le théorème est aussi vrai dans ce cas (avec un γ à déterminer).

(11) Illustrer graphiquement le théorème dans ce cas, en comparant un histogramme de votre échantillon à la densité limite pour n grand.

Exercice 2. *Processus de Poisson*

Soit (T_n) une suite de variables aléatoires indépendantes de loi exponentielle $\mathcal{E}(\lambda)$, $S_0 = 0$ et pour tout $k \geq 1$, $S_k = \sum_{i=1}^{k} T_i$. On définit également, pour tout réel positif t, la variable aléatoire

$$X_t = \sup\{k \in \mathbb{N}, \quad S_k \leq t\}.$$

Le processus aléatoire $(X_t)_{t\geq 0}$ est appelé processus de Poisson d'intensité λ.

(1) Simuler une matrice $n \times N$ de variables aléatoires indépendantes de loi exponentielle $\mathcal{E}(\lambda)$. En déduire une matrice $(n+1) \times N$ comportant en colonnes N vecteurs indépendants de même loi que $S_0, \ldots, S_n$. On pourra prendre $\lambda = 2$, $n = 50$ et $N = 1000$.

(2) Soit l un entier quelconque. Que vaut X_t pour $t = S_l$?

(3) Dessiner sur un même graphique trois trajectoires du processus (X_t).

(4) Soit t un réel positif et l un entier. Montrer que $\mathbb{P}(X_t = l) = \mathbb{E}(\phi(t, S_l, T_{l+1}))$ avec ϕ une fonction à déterminer, et en déduire que X_t suit une loi de Poisson $\mathcal{P}(\lambda t)$.

(5) Soit $t = 2$. Afin de vérifier le résultat précédent, comparer graphiquement les fréquences empiriques pour un N-échantillon de même loi que X_t et les fréquences théoriques d'une loi $\mathcal{P}(\lambda t)$.

(6) Essayer d'illustrer avec Matlab la propriété suivante :

$$\forall t \geq 0 \qquad \mathbb{P}(X_{t+h} - X_t > 1) = o(h) \text{ lorsque } h \to 0+$$

Exercice 3. *Calcul de quantiles par l'algorithme de Robbins-Monroe*

Soit U une fonction continue strictement croissante et a un réel. On cherche à résoudre l'équation $U(x) = a$. Pour cela on construit des suites de variables aléatoires (Y_n) et (X_n) telles que

$$\begin{cases} X_{n+1} = X_n + \dfrac{10}{n+1}(a - Y_{n+1}) \\ \mathbb{E}(Y_{n+1}|\mathcal{F}_n) = U(X_n) \end{cases}$$

où $\mathcal{F}_n$ est la tribu engendrée par $X_0, \ldots, X_n, Y_1, \ldots, Y_n$. On initialise la suite X_n par une valeur déterministe quelconque : $X_0 = x$. On peut montrer que, si la suite Y_n est bornée presque sûrement, X_n converge presque sûrement vers le réel x_a qui réalise $U(x_a) = a$. Ce résultat est admis.

On suppose maintenant que U est la fonction de répartition d'une variable aléatoire Z. Soit (Z_n) des variables indépendantes de même loi que Z. On pose pour tout $n \geq 0$,

$$Y_{n+1} = \mathbb{1}_{Z_{n+1} \leq X_n}.$$

(1) Montrer que $\mathbb{E}(Y_{n+1}|\mathcal{F}_n) = U(X_n)$.

(2) Soit q le quantile d'ordre α de la loi U, c'est-à-dire le réel q tel que $U(q) = 1 - \alpha$. Expliquer comment obtenir une valeur approchée de q à partir des variables Z_n.

On définit la loi de Student à d degrés de liberté, notée $\mathcal{T}(d)$, comme la loi de

$$Z_{d+1} / \sqrt{(Z_1^2 + \cdots + Z_{d+1}^2)/d}$$

où $Z_1, \ldots, Z_{d+1}$ sont des variables indépendantes de loi $\mathcal{N}(0,1)$.

(3) Ecrire une fonction qui prend n et d en argument et renvoit un n-échantillon de loi de Student $\mathcal{T}(d)$.

(4) Mettre en oeuvre la méthode de la question 2 pour calculer le quantile d'ordre $\alpha = 0.15$ de la loi de Student $\mathcal{T}(5)$. Comparer avec `qt(1-alpha,5)`.