Final SQ20, Printemps 2010

La précision et la clarté de la rédaction seront prises en compte dans l'évaluation de la copie.
Les tables des lois statistiques, et une calculatrice sont les seuls documents autorisés.
Le barême, donné à titre indicatif, est susceptible de modification.
Chaque exercice doit être rédigé sur une feuille différente.

Exercice 1 Pareto (14 points)

La loi de Pareto de paramètre $k \in \mathbb{R}_+^\times$, notée $Par(k)$, est une loi continue à valeur dans $[1, +\infty[$ de densité

$$f(x) = \frac{k}{x^{k+1}} \quad \text{pour } x \in [1, +\infty[.$$

Dans tout l'exercice, X désignera une variable aléatoire de loi de Pareto $Par(k)$.

Partie A : Distribution de Pareto

1. Calculer la fonction de répartition F_X de X.
2. Calculer l'espérance de X (on supposera $k > 1$).
3. Soit X la variable aléatoire qui représente le salaire mensuel d'un salarié (en milliers d'euros). Dans cette question, on suppose que $X \sim Par(2)$.
   (a) Calculer α tel que $\mathbb{P}(X \geq \alpha) = 0,2$. Que représente α ?
   (b) On s'intéresse maintenant à la richesse cumulée des 20% des salariés les plus riches. Pour la valeur de α obtenue à la question précédente, calculer
   $$\int_\alpha^{+\infty} x f_X(x)\,\mathrm{d}x.$$
   Remarque : si vous n'avez pas calculé α, prenez une valeur quelconque qui vous semble plausible (le salaire mensuel médian en France est environ de 1,6k€).
   (c) Hors barême : Interpréter le rapport $\dfrac{\int_\alpha^{+\infty} x f_X(x)\,\mathrm{d}x}{\int_1^{+\infty} x f_X(x)\,\mathrm{d}x}$ en termes de répartition des revenus.

Partie B : Estimation de k

Soit $(X_1, X_2, \cdots, X_n)$ une suite de variables aléatoires indépendantes, et identiquement distribuées selon la loi $Par(k)$. On cherche à estimer la valeur du paramètre k.

1. Pour un échantillon $e_n = (x_1, \ldots, x_n)$, donner l'expression de $\ln(L(x_1, \ldots, x_n, k))$, où L est la fonction de vraissemblance :
$$L(x_1, \ldots, x_n, k) = \mathbb{P}(X_1 = x_1, \cdots, X_n = x_n) = \prod_{i=1}^n f_{X_i}(x_i).$$
2. En déduire que l'estimateur de maximum de vraissemblance du paramètre k est donné par
$$T = \frac{n}{\sum_{i=1}^n \ln(X_i)}.$$

3. On peut montrer que $E[T] = \frac{n}{n-1}k$, et donc que T est un estimateur biaisé de k. Application : on s'intéresse au salaire mensuel en milliers d'euros de vingt salariés :

1,7	1,3	1,3	1,1	1,2	2,6	1,1	1,0	1,0	1,0
1,1	2,1	2,3	1,9	1,0	2,1	1,3	1,2	9,2	1,2

   En déduire une estimation non biaisée de k.

Partie C : Étude de la variable $\ln(X)$

Soit $X \sim Par(k)$, et soit $Y = \ln(X)$. La variable Y est donc à valeurs dans $[0, +\infty[$.

1. Soit $t \geq 0$. Montrer que $F_Y(t) = F_X(e^t)$.
2. En déduire que Y suit une loi exponentielle de paramètre k : $Y \sim \mathcal{E}(k)$.
3. Énoncer le théorème central limite pour la variable $\overline{Y}$.
4. En considérant l'estimateur $T = \frac{1}{\overline{Y}}$ de la partie précédente, montrer que pour n assez grand, on peut considérer que
$$\sqrt{n}(k \times \overline{Y} - 1) \approx \mathcal{N}(0, 1).$$
5. Utiliser la variable de test de la question précédente pour calculer un intervalle de confiance à 95% de k pour les salaires de nos vingt salariés :

1,7	1,3	1,3	1,1	1,2	2,6	1,1	1,0	1,0	1,0
1,1	2,1	2,3	1,9	1,0	2,1	1,3	1,2	9,2	1,2

Partie D : Test d'adéquation à la loi de Pareto

Dans les parties précédentes, nous avons supposé que la distribution des richesses dans une population donnée suit une loi de Pareto (aussi appelée loi des 80/20 : 20% de la population concentre 80% des richesses). Nous allons maintenant tester cette hypothèse. Pour cela, on considère les revenus mensuels d'un échantillon de 100 personnes choisies de manière aléatoire. Les résultats sont regroupés dans le tableau suivant :

Salaire (en k€)	[1 ;1,5[	[1,5 ;2[	[2 ;2,5[	[2,5 ;3[	[3,+∞[
Effectif	50	28	10	5	7

La première valeur du tableau signifie par exemple que sur les 100 personnes interrogées, il y en a 50 qui gagnent entre 1000€ inclus et 1500€ exclus par mois.
Tester $H_0 : X \sim Par(2)$ contre $H_1 : X \nsim Par(2)$ au risque $\alpha = 0,05$ (où X désigne la variable aléatoire qui à une personne donnée associe son salaire mensuel en k€).

Exercice 2 Couple de va (6 points)

Soient X et Y deux variables indépendantes de même loi uniforme sur $[0,1]$. Soit $U = \frac{Y}{X}$.

1. Donner la densité du couple (X, Y).
2. Soit $t > 0$. Représenter l'ensemble $D = \{0 \leq x, y \leq 1 \mid y \geq tx\}$.
3. Montrer que la fonction de répartition de U est donnée par

$$F_U(t) = \begin{cases} \frac{t}{2} & \text{si } 0 < t \leq 1 \\ 1 - \frac{1}{2t} & \text{si } t > 1 \end{cases} .$$

4. Donner la densité de U et la représenter graphiquement.
5. Calculer $\mathbb{P}(U > 1)$. Montrer que U a une espérance infinie.