Statistique

Examen du 12 juin 2002 : Correction

EXERCICE 1

1. Comme $S_n = \min(T_1, \ldots, T_n)$,

   $$P(S_n \geq s) = P(T_i \geq s\,, i = 1, \ldots, n) = P(T_1 \geq s)^n = \exp\{-n\ell s\}$$

   donc

   $$S_n \sim \mathcal{E}\mathrm{xp}(n\ell)$$

   De même,

   $$P(R_n \leq r) = P(T_i \leq r, i = 1, \ldots, n) = P(T_1 \leq r)^n = (1 - \exp\{-\ell r\})^n \ ,$$

   qui ne correspond pas à une loi standard.

2. La variable nS_n a une loi $\mathcal{E}\mathrm{xp}(\ell)$, quelque soit n, donc elle converge en loi vers $\mathcal{E}\mathrm{xp}(\ell)$.

3. Comme

   $$P(R_n/n \leq r) = P(R_n \leq nr) = (1 - \exp\{-n\ell r\})^n \ ,$$

   le terme de gauche tend vers 1, car, par passage au log, $n\log(1 - \exp\{-n\ell r\})$ est équivalent à $-n\exp\{-n\ell r\}$ qui tend vers 0 pour tout $r > 0$, et la suite $(R_n/n)_{n\geq 1}$ converge **en loi** vers 0.

4. La suite $(R_n/n)_{n\geq 1}$ converge **en probabilité** vers 0 pour la même raison.

5. $M_n = (T_1 + \ldots + T_n)/n$ est la moyenne empirique des T_i, donc elle converge en probabilité vers la moyenne des T_i, $1/\ell$, par la loi des grands nombres *[cours]* ou par Cebycev *[re-cours]*.

Exercice 2

1. On a

$$\begin{aligned}\int_0^1\int_0^1 x_1 e^{-x_1x_2}dx_1dx_2 &= \int_0^1 \left[-e^{-x_1x_2}\right]_{x_2=0}^{x_2=1} dx_1\\ &= \int_0^1 \left(1-e^{-x_1}\right)dx_1 = \left[x_1+e^{-x_1}\right]_0^1\\ &= 1+e^{-1}-1 = e^{-1} = c^{-1}\,.\end{aligned}$$

2. Par le théorème de factorisation, les composantes X_1 et X_2 sont dépendantes.

3. Comme $Z = \max(X_1, X_2)$,

$$\begin{aligned}F_Z(z) &= P(Z\le z) = P(X_1\le z, X_2\le z)\\ &= e\int_0^z\int_0^z x_1e^{-x_1x_2}dx_1dx_2\\ &= e\int_0^z \left[-e^{-x_1x_2}\right]_{x_2=0}^{x_2=z} dx_1\\ &= e\int_0^z \left(1-e^{-zx_1}\right)dx_1\\ &= \left[x_1+e^{-zx_1}/z\right]_0^z\\ &= e\left(z+\frac{e^{-z^2}}{z}-\frac{1}{z}\right)\end{aligned}$$

   et $P(Z<0.762)$ vaut $\varrho = 0.5000709$ par résolution numérique.

4. La loi de N, nombre de Z_i indépendants dépassant 0.762, est une loi binomiale de coefficients n et 0.5, $\mathcal{B}(n, 0.5)$.

5. Pour $n = 100$, comme $\text{var}(N) = n\,0.5\,0.5 = 0.25\,n$, on peut faire l'approximation

$$\sqrt{n}\,\frac{N-0.5\,n}{n} \approx \mathcal{N}(0, 0.25)$$

   en vertu du théorème central limit. Donc on a approximativement

$$\frac{N}{n} \sim \mathcal{N}\left(0.5, \frac{0.25}{n}\right)\,.$$

Comme

$$
\begin{aligned}
0.95 &= P\left(\frac{N - 0.5\,n}{\sqrt{0.25n}} \le 1.68\right) \\
&= P\left(N - 0.5\,n \le 0.84\sqrt{n}\right) \\
&= P\left(N \le 0.84\sqrt{n} + 0.5\,n\right)\,,
\end{aligned}
$$

$N^\star$ vaut $0.84\sqrt{n} + 0.5\,n = 58.4\,.$

6. Par la formule du changement de variable, le jacobien vaut σ^2 et donc

$$f_\sigma(y) = \frac{y_1 \exp\{-y_1 y_2/\sigma^2\}}{\sigma^3} \mathbb{I}_{0 \le y_1, y_2 \le \sigma}$$

7. Pour un échantillon iid $Y_1, \ldots, Y_n$ de loi f_σ, la vraisemblance vaut

$$\sigma^{-3n} \left(\prod_{i=1}^n y_{i1}\right) \exp\left\{-\sum y_{i1} y_{i2}/\sigma^2\right\} \mathbb{I}_{\sigma > \max_i \max(y_{i1}, y_{i2})}$$

et une statistique exhaustive est

$$T(Y_1, \ldots, Y_n) = \left(\sum Y_{i1} Y_{i2}, \max_i \max(Y_{i1}, Y_{i2})\right)$$

8. Comme

$$
\begin{aligned}
\mathbb{E}[X_1 X_2] &= e \int_0^1 \int_0^1 x_1^2 x_2 e^{-x_1 x_2} dx_1 dx_2 \\
&= e \int_0^1 (1 - e^{-x_1} - x_1 e^{-x_1}) dx_1 \\
&= 3 - e = 0.2817\,,
\end{aligned}
$$

$\mathbb{E}[Y_1 Y_2] = 0.2817\sigma^2$ et

$$\frac{\sum Y_{i1} Y_{i2}}{0.2817\,n}$$

est un estimateur sans biais de σ^2. (Ce n'est pas le seul!)

QCM

1. Si (X_1, X_2) est un vecteur normal de moyenne (μ_1, μ_2) et de matrice de variance-covariance $\begin{pmatrix} 1 & \varrho \\ \varrho & 1 \end{pmatrix}$, la loi conditionnelle de X_2 sachant $X_1 = x_1$ est

   ○ $\mathcal{N}_1(\mu_2 + x_1, \varrho^2)$ ⊗ $\mathcal{N}_1(\mu_2 + \varrho(x_1 - \mu_1), 1 - \varrho^2)$
   ○ $\mathcal{N}_1(\varrho(x_1 - \mu_1), 1 + \varrho^2)$ ○ $\mathcal{N}_1(\mu_2 + (x_1 - \mu_1)/\varrho, 1/\varrho^2)$

2. Si X_1 et X_2 sont des variables aléatoires indépendantes de lois gamma $\mathcal{G}a(\alpha_1, 1)$ et $\mathcal{G}a(\alpha_2, 1)$, la variable aléatoire $X_1/(X_1 + X_2)$ a comme loi

   ○ $\mathcal{G}a(\alpha_1/(\alpha_1 + \alpha_2), 1)$ ○ $\mathcal{G}a(\alpha_1/\alpha_2, \alpha_1^2/(\alpha_1 + \alpha_2))$
   ○ $\mathcal{N}(\alpha_1/(\alpha_1 + \alpha_2), \alpha_1\alpha_2/(\alpha_1 + \alpha_2)^2)$ ⊗ $\mathcal{B}e(\alpha_1, \alpha_2)$

3. Soient $X_1, \ldots, X_n$ n variables indépendantes de fonctions génératrices respectives $F_1, \ldots, F_n$. Alors $\bar{X}_n = (X_1 + \ldots + X_n)/n$ a comme fonction génératrice

   ○ $G(t) = \sqrt[n]{F_1(t) \ldots F_n(t)}$ ○ $G(t) = F_1(t/n) + \ldots + F_n(t/n)$
   ○ $G(t) = (F_1(t) + \ldots + F_n(t))/n$ ⊗ $G(t) = F_1(t/n) \ldots F_n(t/n)$

4. Si X est une variable aléatoire de fonction de répartition F_X et U une variable uniforme,

   ⊗ X a même loi que $F_X^{-1}(U)$ ○ U a même loi que $F_X^{-1}(X)$
   ○ U a même loi que $F_X(X)$ ○ $F_X(U)$ a même loi que X

5. Si $X_1, \ldots, X_n$ est un échantillon iid de loi $\mathcal{U}(-\theta, \theta)$, une statistique exhaustive minimale est

   ○ $X_{(n)} = \max X_i$ ○ $X_{(1)} = \min X_i$
   ⊗ $\max |X_i|$ ○ $(X_{(1)}, X_{(n)})$