Université de Strasbourg
Statistique Mathématique

Département de mathématique
L3 S6 année 2016-2017

Épreuve de rattrapage : UE Statistique Mathématique

Solution

Exercice 2 :

1. (-3-4+0+1+5+5+2-1-5+2+6+7)/12=5/4.
2. -5 ; -4 ; -3 ; -1 ; 0 ; 1 ; 2 ; 2 ; 5 ; 5 ; 6 ; 7.
3. $n = 12$ est pair. Donc la médiane vaut $\frac{X_{6,n}+X_{7,n}}{2} = \frac{1+2}{2} = 1.5$ où $X_{i,n}$ est la ième statistique d'ordre, ie $X_{1,n} \leq ... \leq X_{n,n}$.
4. Si $12p \in \mathbb{N}$ alors $q_{12,p} = (X_{12p,n} + X_{12p+1,n})/2$. Donc
   1er quartile=$q_{12,1/4} = (-3-1)/2 = -2$.
   2ème quartile=médiane=$q_{12,1/2} = (1+2)/2 = 3/2$.
   3ème quartile=$q_{12,3/4} = (5+5)/2 = 5$.
5. Étendue=$X_{n,n} - X_{1,n} = 7 - (-5) = 12$.

Exercice 2 :

1. On écrit la vraisemblance
$$L(x_1, ..., x_n; \theta) = 1\!\!1_{x_{1,n}>\theta} 1\!\!1_{x_{n,n}<\theta+1} = g((x_{1,n}, x_{n,n}); \theta).$$
   Donc $(X_{1,n}, X_{n,n})$ est une statistique exhaustive. Pour montrer qu'elle est minimale, il suffit de voir que
$$\frac{L(X_1, ..., X_n; \theta)}{L(Y_1, ..., Y_n; \theta)} = \frac{1\!\!1_{X_{1,n}>\theta} 1\!\!1_{X_{n,n}<\theta+1}}{1\!\!1_{Y_{1,n}>\theta} 1\!\!1_{Y_{n,n}<\theta+1}}$$
   est indépendant de θ ssi $X_{1,n} = Y_{1,n}$ et $X_{n,n} = Y_{n,n}$.

2. Il s'agit de faire un changement de variable. Les anciennes coordonnées sont $(X_{1,n}, X_{n,n})$ et les nouvelles $(R = X_{n,n} - X_{1,n}, U = X_{1,n})$. Le jacobien vaut -1. On a donc
$$\begin{aligned} f_{(R,U)}(r,u) &= f_{(X_{1,n},X_{n,n})}(u, r+u)|1| \\ &= n(n-1)r^{n-2} 1\!\!1_{\theta<u<r+u<\theta+1}. \end{aligned}$$
   Donc
$$\begin{aligned} f_R(r) &= \int f_{(R,U)}(r,u)du \\ &= n(n-1)r^{n-2}(1-r) 1\!\!1_{0<r<1} \end{aligned}$$
   qui est indépendant de θ. Donc R est libre.

Exercice 3 :

1. Clairement $f_\theta(x) > 0$ et $\int_{\mathbb{R}} f_\theta(x)dx = \int_0^\infty 2\theta x\, e^{-\theta x^2} dx = [-e^{-\theta x^2}]_0^\infty = 1$. Donc f_θ est bien une densité.

2. En utilisant l'indépendance des X_i, on forme la vraisemblance

$$\begin{aligned} L(x_1, ..., x_n; \theta) &= \prod_{i=1}^{n} \left\{ 2\theta x_i \, e^{-\theta x_i^2} \mathbb{1}_{\{x_i \geq 0\}} \right\} \\ &= \left(2^n \prod_{i=1}^{n} x_i \, \mathbb{1}_{\{\min(x_i) \geq 0\}} \right) \left(\theta^n \, e^{-\theta \sum_{i=1}^{n} x_i^2} \right) \\ &= h(x_1, ..., x_n) g\left(\sum_{i=1}^{n} x_i^2; \theta \right). \end{aligned}$$

Donc $\sum_{i=1}^{n} X_i^2$ est une statistique exhaustive.

3. La fonction $x \mapsto \frac{n}{x}$ est bijective sur $\mathbb{R}_*^+$. Donc $\widehat{\theta} = \frac{n}{\sum_{i=1}^{n} X_i^2}$ est une statistique exhaustive pour θ.

4. $\mathbb{P}(Y \leq t) = \mathbb{P}\left(X^2 < \frac{t}{\theta}\right) = \mathbb{P}\left(X \leq \sqrt{\frac{t}{\theta}}\right)$ si $t > 0$, sinon cette probabilité vaut 0. Par conséquent $F_Y(t) = F_X\left(\sqrt{\frac{t}{\theta}}\right)$. En dérivant, on obtient la densité de Y : $f_Y(t) = \frac{1}{2\sqrt{t\theta}} f_X\left(\sqrt{\frac{t}{\theta}}\right) = e^{-t}\mathbb{1}_{\{t \geq 0\}}$. De ce fait, Y est de loi exponentielle de paramètre 1, i.e. une Gamma$(1, 1)$.

5. En tant que somme de variables aléatoires indépendantes de loi Gamma$(1, 1)$, on a $\sum_{i=1}^{n} Y_i$ qui suit une Gamma$(n, 1)$.

6. En utilisant les questions précédentes

$$\begin{aligned} \mathbb{E}(\widehat{\theta}) &= \mathbb{E}\left(\frac{n\theta}{\sum_{i=1}^{n} Y_i} \right) = n\theta \int_0^\infty \frac{1}{x} \frac{1}{\Gamma(n)} x^{n-1} e^{-x} dx \\ &= n\theta \frac{\Gamma(n-1)}{\Gamma(n)} = \frac{n}{n-1}\theta. \end{aligned}$$

Par conséquent $\widehat{\theta}$ n'est pas un estimateur sans biais, mais c'est un estimateur asymptotiquement sans biais.
$\widetilde{\theta} = \frac{n-1}{n}\widehat{\theta} = \frac{n-1}{\sum_{i=1}^{n} X_i^2}$ est donc un estimateur sans biais de θ.

7. On va utiliser la loi faible des grands nombres. On a

$$\widetilde{\theta} = \frac{n-1}{n} \frac{\theta}{\frac{1}{n}\sum_{i=1}^{n} Y_i} \xrightarrow{\mathbb{P}} \frac{\theta}{\mathbb{E}Y} = \theta.$$

Toutes les conditions sont requises pour utiliser cette loi faible des grands nombres car les Y_i sont indépendantes et de loi Gamma$(1, 1)$.

8. L'estimateur $\widetilde{\theta}$ étant sans biais, il suffit de calculer sa variance. Pour cela on procède comme pour l'espérance :

$$\begin{aligned} \mathbb{E}(\widetilde{\theta}^2) &= \left(\frac{n-1}{n} \right)^2 \mathbb{E}\left(\frac{n^2\theta^2}{(\sum_{i=1}^{n} Y_i)^2} \right) = (n-1)^2\theta^2 \int_0^\infty \frac{1}{x^2} \frac{1}{\Gamma(n)} x^{n-1} e^{-x} dx \\ &= (n-1)^2\theta^2 \frac{\Gamma(n-2)}{\Gamma(n)} = \frac{n-1}{n-2}\theta^2. \end{aligned}$$

Par conséquent $\mathrm{Var}(\widetilde{\theta}) = \frac{\theta^2}{n-2}$ qui converge vers 0, donc $\mathbb{E}((\widetilde{\theta} - \theta)^2) \to 0$ et l'estimateur $\widetilde{\theta}$ est convergeant en moyenne quadratique.

Exercice 4 :

1. $X + 2Y - Z$ est une combinaison linéaire des composantes de $\aleph$ qui est un vecteur gaussien. Donc $X + 2Y - Z$ est une variable aléatoire gaussienne, centrée et de variance 18.

2. $\begin{pmatrix} X+Y \\ Y-Z \end{pmatrix}$ est un vecteur gaussien car toute combinaison linéaire de ses composantes s'exprime comme une combinaison linéaire de X, Y et Z, donc des composantes du vecteur $\aleph$ qui est gaussien. Le vecteur est centré et sa matrice de variance-covariance est $\begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}$.

3. $\aleph$ est un vecteur gaussien, donc indépendance équivaut à covariance nulle. D'après la matrice de variance-covariance Y et Z sont donc indépendantes.

4. Deux variables sont orthogonales si leur produit scalaire est nul. Par définition, le produit scalaire entre deux variables est l'espérance du produit de ces deux variables. Si une de ces variables au moins est centrée, cela revient donc à dire qu'elles sont orthogonales si leur covariance est nulle. Par conséquent Y et Z sont orthogonales.

5. D'après la question 3. Y et Z sont indépendantes, donc la densité du couple $\begin{pmatrix} Y \\ Z \end{pmatrix}$ est le produit des densités : $f_{(Y,Z)}(y,z) = f_Y(y) f_Z(z) = \frac{1}{\sqrt{2}\sqrt{2\pi}} e^{-\frac{1}{4}y^2} \times \frac{1}{\sqrt{3}\sqrt{2\pi}} e^{-\frac{1}{6}z^2} = \frac{1}{2\sqrt{6}\pi} e^{-\frac{y^2}{4} - \frac{z^2}{6}}$.

6. $\frac{Y}{\sqrt{2}}$ est de loi $\mathcal{N}(0,1)$ et $\frac{Z}{\sqrt{3}}$ est de loi $\mathcal{N}(0,1)$. Elles sont par ailleurs indépendantes. Donc $\frac{Y^2}{2} + \frac{Z^2}{3}$ est de loi chi-deux à deux degrés de liberté.

7. $\mathbb{E}[X|Y]$ est le projeté orthogonal de X sur $L^2(Y)$ qui coincide avec le projeté orthogonal de X sur $\mathrm{Vect}(1, Y)$. On a
$$\mathbb{E}[X|Y] = \mathbb{E}(X) + \frac{Cov(X,Y)}{Var(Y)}(Y - \mathbb{E}(Y)) = \frac{Y}{2}.$$

8. $$\mathbb{E}[X|(Y,Z)] = \mathbb{E}(X) + \Gamma_{X,(Y,Z)} \Gamma_{(Y,Z)}^{-1} \begin{pmatrix} Y - \mathbb{E}(Y) \\ Z - \mathbb{E}(Z) \end{pmatrix} = \frac{Y}{2} - \frac{Z}{3}.$$

9. $$\mathbb{E}[(X - \mathbb{E}[X|(Y,Z)])^2] = \Gamma_X - \Gamma_{X,(Y,Z)} \Gamma_{(Y,Z)}^{-1} (\Gamma_{X,(Y,Z)})' = \frac{1}{6}.$$