Tasdawit n'Bgayet
Université de Béjaïa

Université A. Mira Bejaia
Faculté des Sciences Exactes
Département de Mathématiques
Master 2 (A.M)

Corrigé de l'Examen de Statistique Inférentielle

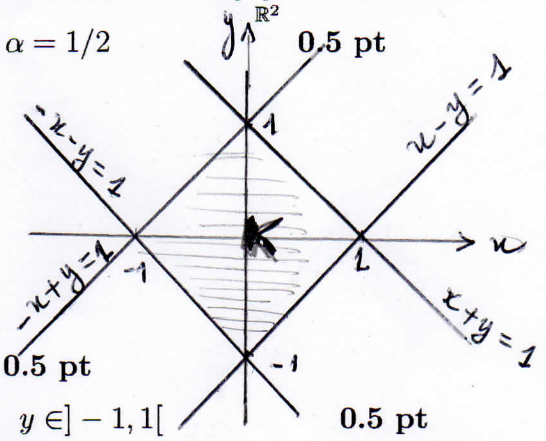
Exercice 1. (03 points)

$\mathbf{K} = \{(x, y) \in \mathbb{R}^2 : |x| + |y| < 1\}$ (carré de coté $\sqrt{2}$), $f(x, y) = \alpha \mathbf{1}_{\boldsymbol{K}}(x, y)$.

1. f est une densité de probabilité si $f(x, y) \geq 0$, $\forall (x, y) \in \mathbb{R}^2$ et $\int\int_{\mathbb{R}^2} f(x, y)dxdy = 1$

On a $\int\int_{\mathbb{R}^2} f(x, y)dxdy = \alpha \int\int_{\mathbb{K}} dxdy = 2\alpha = 1 \Rightarrow \alpha = 1/2$ **0.5 pt**

D'où $f(x, y) = \frac{1}{2}\mathbf{1}_K(x, y)$.



2. Déterminons les lois marginales de X et Y.

$f_X(x) = \int_{\mathbb{R}} f(x, y)dy$

Si $x \in]-1, 0]$: $f_X(x) = \frac{1}{2}\int_{-x-1}^{1+x} dy = x + 1$

Si $x \in [0, 1[$: $f_X(x) = \frac{1}{2}\int_{x-1}^{1-x} dy = 1 - x$

D'où $f_X(x) = 1 - |x|, \quad x \in]-1, 1[$ **0.5 pt**

De la même manière, on trouve $f_Y(y) = 1 - |y|, \quad y \in]-1, 1[$ **0.5 pt**

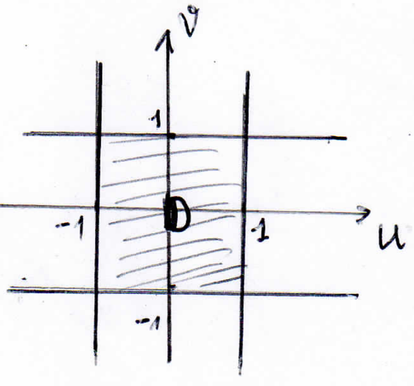
3. Les variables aléatoires X et Y ne sont pas indépendantes car $f(x, y) \neq f_X(x)f_Y(y)$ **0.25 pt**

4. Densité de la variable aléatoire $X + Y$: Utilisons la méthode du changement de variable.

On pose $\begin{cases} U = X + Y \\ V = X - Y \end{cases}$

$\begin{cases} u = x + y \\ v = x - y \end{cases} \Longleftrightarrow \begin{cases} x = \frac{u+v}{2} \\ y = \frac{v-u}{2} \end{cases}$

$J(x, y) = -2 \neq 0$. D'où $f_{((U,V)}(u, v) = \frac{1}{4}\mathbf{1}_{\mathbf{D}}(u, v)$ **0.5 pt**



avec $\mathbf{D} = \{(u, v) \in \mathbb{R}^2 : |u| < 1, |v| < 1\}$ (carré de coté 2)

$f_{X+Y}(u) = f_U(u) = \int_{\mathbb{R}} f_{(U,V)}(u, v)dv = \int_{-1}^{1} \frac{1}{4}dv = \frac{1}{2}$

Donc $f_{X+Y}(u) = \begin{cases} \frac{1}{2}, & u \in]-1, 1[; \\ 0, & \text{sinon.} \end{cases}$

Donc $U = X + Y$ suit la loi Uniforme sur $]-1, 1[$ **0.25 pt**

Par symétrie, $V = X - Y$ suit la loi Uniforme sur $]-1, 1[$ **0.25 pt**

Comme $f_{(U,V)}(u, v) = f_U(u)f_V(v)$, les variables aléatoires U et V sont indépendantes. **0.25 pt**

Exercice 2. (05 points)

1. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires réelles indépendantes telles que $\forall n \geq 1$, $X_n \hookrightarrow \mathcal{E}xp(\lambda_n = n)$.

a) Fonction de répartition de $Y_n = \min\limits_{1\leq i\leq n} X_i$ pour tout $n \geq 1$.

$$\begin{aligned}
F_{Y_n}(y) &= \mathbb{P}(\min_{1\leq i\leq n} X_i \leq y) \\
&= 1 - \mathbb{P}(\min_{1\leq i\leq n} X_i > y) \\
&= 1 - \prod_{i=1}^{n}[\mathbb{P}(X_i > y)] \\
&= 1 - \prod_{i=1}^{n}[1 - \mathbb{P}(X_i \leq y)] \\
&= 1 - \prod_{i=1}^{n}(1 - F_{X_i}(y)) \\
&= 1 - \prod_{i=1}^{n} e^{-iy} = 1 - e^{-y\sum\limits_{i=1}^{n} i} \\
&= 1 - e^{-\frac{n(n+1)}{2}y}, \quad y \geq 0 \qquad \textbf{3pts}
\end{aligned}$$

Donc Y_n suit $\exp(\frac{n(n+1)}{2})$.

b) On a $\forall \varepsilon > 0 : \mathbb{P}(|Yn - 0| > \varepsilon) = \mathbb{P}(Yn > \varepsilon) = e^{-\frac{n(n+1)}{2}y}$

On a $\lim\limits_{n\to\infty} e^{-\frac{n(n+1)}{2}y} = 0$

D'où Y_n converge en probabilité vers 0. **0.5pt**

2. Soit $(T_n)_{n\geq 1}$ une suite de v.a.r. indépendantes de même loi $\mathcal{N}(0,1)$ et $S_n = \frac{1}{n}\sum\limits_{k=1}^{n}\sqrt{k}\, T_k$.
On ne peut pas appliquer le théorème central limite pour l'étude de la convergence en loi de la suite (S_n) car les variables aléatoires $\sqrt{k}\, T_k$ ne sont pas identiquement distribuées. **0.25pt**

3. On a $\mathbb{E}(S_n) = 0$ et $V(S_n) = \frac{n+1}{2n}$

Donc $S_n \rightsquigarrow \mathcal{N}(0, \frac{n+1}{2n})$. **0.5pt**

La fonction caractéristique de S_n est donnée par: $\varphi_{S_n}(t) = e^{-\frac{t^2(n+1)}{4n}}$, $\forall t \in \mathbb{R}$. **0.5pt**

$\lim\limits_{n\to\infty} \varphi_{S_n}(t) = e^{-\frac{t^2}{4}}$, (fonction caractéristique de $\mathcal{N}(0, 1/2)$ et est continue en 0)

D'où S_n converge en loi vers une variable aléatoire de loi $\mathcal{N}(0, 1/2)$. **0.25pt**

Exercice 3. (12 points)

Soit $(X_1, X_2, \ldots, X_n)$ un échantillon aléatoire simple issu d'une variable aléatoire X de densité de probabilité

$$f(x;\theta) = \begin{cases} \frac{1}{x\sqrt{2\pi\theta}} e^{-\frac{1}{2}\frac{(\ln x)^2}{\theta}}, & x > 0; \\ 0, & \text{sinon.} \end{cases}$$

Où θ est un paramètre réel inconnu strictement positif.

1. Déterminons l'estimateur du maximum de vraisemblance T_n de θ.

La vraisemblance de l'échantillon est: $L(\underline{x}, \theta) = \prod_{i=1}^{n} f(x_i, \theta) = \frac{1}{(2\pi\theta)^{n/2} \prod_{i=1}^{n} x_i} \exp\{-\frac{1}{2} \sum_{i=1}^{n} \frac{(\ln x_i)^2}{\theta}\}$

$\ln L(\underline{x}, \theta) = \frac{-n}{2} \ln(2\pi\theta) + \sum_{i=1}^{n} \ln x_i - \frac{1}{2\theta} \sum_{i=1}^{n} (\ln x_i)^2$

$\frac{\partial \ln L(\underline{x},\theta)}{\partial\theta} = -\frac{n}{2\theta} + \frac{\sum_{i=1}^{n}(\ln x_i)^2}{2\theta^2}$

$\frac{\partial \ln L(\underline{x},\theta)}{\partial\theta} = 0 \Longrightarrow \hat{\theta} = \frac{1}{n} \sum_{i=1}^{n} (\ln x_i)^2$

Dautre part $\frac{\partial^2 \ln L(\underline{x},\theta)}{\partial\theta^2} = \frac{1}{\theta^2}(\frac{n}{2} - \frac{\sum_{i=1}^{n}(\ln x_i)^2}{\theta})$

au point $\theta = \hat{\theta}$, $\frac{\partial^2 \ln L(\underline{x},\hat{\theta})}{\partial\theta^2} = -\frac{n}{2\hat{\theta}^2} < 0$

D'où l'E.M.V de θ est $T_n = \frac{1}{n} \sum_{i=1}^{n} (\ln X_i)^2$. **3pts**

2. Loi de probabilité de $Y = \ln X$:

$\mathbb{P}(Y \leq y) = \mathbb{P}(\ln X \leq y) = \mathbb{P}(X \leq e^y) = F_X(e^y)$

D'où la densité de Y est $f_Y(y) = e^y f_{X(e^y)} = \frac{1}{\sqrt{2\pi\theta}} \exp\{-\frac{y^2}{2\theta}\}$, $y \in \mathbb{R}$. **2pts**

Donc $Y \rightsquigarrow \mathcal{N}(0, \theta)$. **1 pt**

On en déduit $\mathbb{E}(Y) = 0$ **0.5 pt** et $V(Y) = \theta$. **0.5 pt**

Par conséquent $\frac{Y}{\sqrt{\theta}} \rightsquigarrow \mathcal{N}(0, 1)$. **1 pt**

3. On a $\mathbb{E}(T_n) = \mathbb{E}(Y^2) = V(Y) + \mathbb{E}^2(Y) = \theta$. Donc T_n est un estimateur sans biais. **0.5 pt**

$V(T_n) = \frac{1}{n} V(Y^2)$. Or comme $\frac{Y}{\sqrt{\theta}} \rightsquigarrow \mathcal{N}(0, 1) \Longrightarrow \frac{Y^2}{\theta} \rightsquigarrow \chi_1^2$ D'où $Y^2 \rightsquigarrow \theta\chi_1^2$ par conséquent $V(Y^2) = 2\theta^2$ et $V(T_n) = \frac{2\theta^2}{n}$

$\lim_{n\to\infty} V(T_n) = 0$. Donc T_n est convergent de θ. **1 pt**

Pour l'efficacitén, on calcule l'information de Fisher apporté par l'échantillon sur θ:

$$\begin{aligned} I_n(\theta) &= -\mathbb{E}(\frac{\partial^2 \ln L(\underline{x}, \theta)}{\partial\theta^2}) \\ &= -\mathbb{E}(\frac{n}{2\theta} - \frac{\sum_{i=1}^{n}(\ln x_i)^2}{\theta^3} \\ &= -\frac{n}{2\theta} + \frac{n}{\theta^3}\mathbb{E}(Y^2) \\ &= -\frac{n}{2\theta} + \frac{n\theta}{\theta^3} = \frac{n}{2\theta^2} \end{aligned}$$ **0.5 pt**

On a la borne de Frechet $B_F = \frac{1}{I_n(\theta)} = \frac{2\theta^2}{n} = V(T_n)$. D'où T_n est un estimateur efficace de θ. **0.5 pt**

4. $IC_{1-\alpha}(\theta)$?

On a T_n est le meilleur estimateur possible de θ et $\frac{nT_n}{\theta} = \sum_{i=1}^{n} \frac{\ln X_i}{\sqrt{\theta}}$

Or $\frac{\ln X_i}{\sqrt{\theta}} \rightsquigarrow \mathcal{N}(0,1) \Longrightarrow \frac{nT_n}{\theta} \rightsquigarrow \chi^2_n$

L'intervalle de probabilité est

$\mathbb{P}(k_1 \leq \frac{nT_n}{\theta} \leq k_2) = 1-\alpha$ avec $k_1 = \phi^{-1}_{\chi^2_n}(\frac{\alpha}{2})$ et $k_2 = \phi^{-1}_{\chi^2_n}(\frac{1-\alpha}{2})$. D'où l'intervalle de confiance pour θ est

$$IC_{1-\alpha}(\theta) = [\frac{nt_n}{k_2}, \frac{nt_n}{k_1}] \quad \text{avec} \quad t_n = \frac{1}{n}\sum_{i=1}^{n}(\ln x_i)^2. \qquad \textbf{1 pt}$$

5. Application numérique:

Estimations ponctuelle: $\hat{\theta} = t_n = \frac{1}{10}\sum_{i=1}^{10} i^2 = \frac{n(n+1)(2n+1)}{6} = \frac{77}{2} = 38.5$ **0.25 pt**

Estimation par intervalle de confiance: Sur la table de χ^2_{10}, on lit $k_1 = 3.25$ et $k_2 = 20.48$

Donc

$$IC_{0.95}(\theta) = [18.8,\ 118.46] \qquad \textbf{0.25 pt}$$