

Sujet type HEC – ESSEC 2.

La partie I peut être posée à tous les concours.

Toutes les variables aléatoires qui interviennent dans ce problème sont considérées comme définies sur des espaces probabilisés non nécessairement identiques, mais qui, par souci de simplification, seront tous notés $(\Omega, \mathcal{A}, \mathbf{P})$

Partie I : loi de Laplace.

Soit $\alpha \in \mathbf{R}$ et $\beta > 0$. On dit qu'une variable aléatoire réelle à densité suit une loi de Laplace de paramètre (α, β) , notée $\mathcal{L}(\alpha, \beta)$, si elle admet comme densité la fonction $h_{\alpha, \beta}$ donnée par :

$$\forall t \in \mathbf{R}, \quad h_{\alpha, \beta}(t) = \frac{1}{2\beta} \exp\left(-\frac{|t - \alpha|}{\beta}\right)$$

- 1) Vérifier que $h_{\alpha, \beta}$ est bien une densité de probabilité d'une variable aléatoire réelle.
- 2) Déterminer la fonction de répartition, notée G , de la loi $\mathcal{L}(0, 1)$. On notera $g = h_{0, 1}$ une densité pour cette loi.
- 3) On suppose que Y suit la loi $\mathcal{L}(0, 1)$.
 - a) Montrer que $\beta Y + \alpha$ suit la loi $\mathcal{L}(\alpha, \beta)$.
 - b) En déduire la fonction de répartition $H_{\alpha, \beta}$ de la loi $\mathcal{L}(\alpha, \beta)$ en fonction de G .
 - c)
 - (i) Montrer que Y admet un moment à tout ordre $r \in \mathbf{N}^*$, noté $m_r(Y)$.
 - (ii) Vérifier que $\mathbf{E}(Y)$ et $\mathbf{V}(Y)$ valent respectivement 0 et 2.
 - (iii) Pour $r \in \mathbf{N}^*$, calculer $m_r(Y)$ en fonction de r .
 - d) Déduire de ce qui précède l'existence et les valeurs de l'espérance et de la variance d'une variable aléatoire réelle qui suit la loi $\mathcal{L}(\alpha, \beta)$.
- 4) a) Établir que G est une bijection de $\mathbf{R}$ sur $]0, 1[$.
 - b) Montrer que l'équation $G(x) = \frac{1}{2}$ admet une unique solution que l'on déterminera.
 - c) Établir que la fonction qui, à tout réel x associe $G(x)(1 - G(x))$, est paire.
 - d) Montrer que l'application réciproque G^{-1} de G est définie par :

$$G^{-1}(x) = \begin{cases} \ln(2x) & \text{si } 0 < x \leq 1/2 \\ -\ln(2(1-x)) & \text{si } 1/2 \leq x < 1 \end{cases}$$

- e) Soit U une variable suivant la loi uniforme sur $]0, 1[$.
Montrer que $Y = G^{-1}(U)$ suit la loi de Laplace $\mathcal{L}(0, 1)$.
 - f)
 - (i) Définir sous Python une fonction `InvG(x)` renvoyant l'expression de $G^{-1}(x)$ pour x donné en entrée.
 - (ii) Définir sous Python une fonction `Laplace(alpha, beta)` simulant la loi $\mathcal{L}(\alpha, \beta)$. On utilisera la fonction `InvG`.
- 5) Pour tout entier n de $\mathbf{N}^*$, on considère la fonction g_n définie par :

$$\forall x \in \mathbf{R}, \quad g_n(x) = g(x) (1 + xe^{-n|x|})$$

a) Montrer que g_n définit une densité de probabilité.

Pour tout n de $\mathbf{N}^*$, on désigne par Y_n une variable aléatoire de densité g_n , et on note G_n la fonction de répartition de Y_n .

b) Établir pour tout réel x , la majoration suivante : $|G_n(x) - G(x)| \leq \frac{1}{ne} \times G(x)$

c) En déduire que la suite de variables aléatoires $(Y_n)_{n \geq 1}$ converge en loi vers une variable aléatoire suivant la loi $\mathcal{L}(0, 1)$

Partie II : étude de deux estimateurs de β pour la loi $\mathcal{L}(0, \beta)$.

Dans cette partie, on suppose que β est un réel strictement positif que l'on cherche à estimer.

On considère une variable aléatoire R suivant la loi $\mathcal{L}(0, \beta)$ et une suite de variables aléatoires $(R_n)_{n \in \mathbf{N}^*}$ mutuellement indépendantes et de même loi que R (loi de Laplace $\mathcal{L}(0, \beta)$).

Pour $n \in \mathbf{N}^*$, on pose enfin $T_n = \frac{1}{n} \sum_{k=1}^n |R_k|$ et $M_n = \max(|R_1|, \dots, |R_n|)$.

6) Montrer que $|R|$ suit la loi exponentielle de paramètre $\frac{1}{\beta}$.

7) Montrer que $\mathbf{E}(T_n) = \beta$, et que :

$$\forall \varepsilon > 0, \quad \lim_{n \rightarrow +\infty} \mathbf{P}(|T_n - \beta| > \varepsilon) = 0.$$

T_n sera donc considéré comme un estimateur convergent de β .

8) Montrer que M_n est une variable à densité et qu'on peut choisir pour densité la fonction μ_n définie par :

$$\forall t \in \mathbf{R}, \quad \mu_n(t) = \begin{cases} 0 & \text{si } t < 0 \\ \frac{ne^{-t/\beta}}{\beta} (1 - e^{-t/\beta})^{n-1} & \text{si } t \geq 0. \end{cases}$$

9) a) Soit A un réel strictement positif et $n \in \mathbf{N}^*$. Par le changement de variables $x = e^{-t/\beta}$ montrer que :

$$\int_0^A te^{-t/\beta} (1 - e^{-t/\beta})^{n-1} dt = -\beta^2 \int_{e^{-A/\beta}}^1 (1-x)^{n-1} \ln(x) dx.$$

b) Pour $\varepsilon \in]0, 1]$ et $n \in \mathbf{N}^*$, on pose : $I_n(\varepsilon) = \int_{\varepsilon}^1 -n(1-x)^{n-1} \ln(x) dx$.

(i) Calculer $I_1(\varepsilon)$ pour $\varepsilon \in]0, 1]$ et montrer que : $\lim_{\varepsilon \rightarrow 0+} I_1(\varepsilon) = 1$.

(ii) A l'aide d'une intégration par parties, montrer que : pour $\varepsilon > 0$ et $n \in \mathbf{N}^*$

$$I_{n+1}(\varepsilon) - I_n(\varepsilon) = (1-\varepsilon)^n \varepsilon \ln(\varepsilon) - \int_{\varepsilon}^1 (1-x)^n dx$$

(iii) En déduire que : pour tout $n \in \mathbf{N}^*$, $I_n(\varepsilon)$ admet une limite finie, notée u_n , lorsque ε tend vers $0+$ et que :

$$\forall n \in \mathbf{N}^*, \quad u_{n+1} - u_n = \frac{1}{n+1}.$$

(iv) Montrer que : pour tout $n \in \mathbf{N}^*$, $u_n = \sum_{k=1}^n \frac{1}{k}$.

(v) Dédurre des questions précédentes que M_n admet une espérance et que :

$$\forall n \in \mathbf{N}^*, \quad \mathbf{E}(M_n) = \beta u_n$$

On admettra que M_n admet une variance et que $\mathbf{V}(M_n) = \beta^2 v_n$ avec $v_n = \sum_{k=1}^n \frac{1}{k^2}$.

10) On pose, pour $n \in \mathbf{N}^*$, $\widetilde{M}_n = \frac{1}{u_n} M_n$.

a) Calculer $\mathbf{E}(\widetilde{M}_n)$ et $\mathbf{V}(\widetilde{M}_n)$.

b) Déterminer la limite de u_n quand n tend vers $+\infty$ et justifier que $(v_n)_{n \in \mathbf{N}^*}$ est une suite convergente.

c) Dédurre des deux questions précédentes que :

$$\forall \varepsilon > 0, \quad \lim_{n \rightarrow +\infty} \mathbf{P} \left(\left| \widetilde{M}_n - \beta \right| > \varepsilon \right) = 0.$$

Ce résultat permet de considérer que $\widetilde{M}_n$ est un estimateur convergent de β .

d) Montrer que, pour tout entier $k \in \mathbf{N}^*$, $\frac{1}{k+1} \leq \ln(k+1) - \ln(k) \leq \frac{1}{k}$.

En déduire que : pour tout entier $n \in \mathbf{N}^*$, $\ln(n) + \frac{1}{n} \leq u_n \leq 1 + \ln(n)$, puis que $u_n \underset{n \rightarrow +\infty}{\sim} \ln(n)$.

e) Montrer que : $\mathbf{V}(\widetilde{M}_n) \underset{n \rightarrow +\infty}{=} o(\mathbf{V}(T_n))$.

En quoi cela suggère que $\widetilde{M}_n$ est un meilleur estimateur que T_n ?

Partie III : estimateurs de θ pour la loi $\mathcal{L}(\theta, 1)$.

Soit θ un paramètre réel inconnu et X une variable aléatoire à densité suivant la loi $\mathcal{L}(\theta, 1)$.

Soit n un entier naturel . On considère un $(2n+1)$ -échantillon $(X_1, X_2, \dots, X_{2n+1})$ de variables aléatoires réelles indépendantes et de même loi $\mathcal{L}(\theta, 1)$

Partie III - A.

11) Déterminer la fonction de répartition F de la variable aléatoire X .

12) Soit x un réel fixé. Pour tout i de $\llbracket 1, 2n+1 \rrbracket$, on note Z_i la variable aléatoire de Bernoulli telle que $\mathbf{P}(Z_i = 1) = \mathbf{P}(X_i \leq x)$

a) Établir l'indépendance des variables aléatoires $Z_1, Z_2, \dots, Z_{2n+1}$

b) Soit S_{2n+1} la variable aléatoire définie par : $S_{2n+1} = \sum_{i=1}^{2n+1} Z_i$.

Quelle est la loi de probabilité de S_{2n+1} ?

Préciser l'espérance et la variance de S_{2n+1} .

13) On pose $\overline{X}_{2n+1} = \frac{1}{2n+1} \sum_{i=1}^{2n+1} X_i$.

a) Montrer que $\mathbf{E}(\overline{X}_{2n+1}) = \theta$ et calculer $\mathbf{V}(\overline{X}_{2n+1})$.

b) Montrer que :

$$\forall \varepsilon > 0, \quad \lim_{n \rightarrow +\infty} \mathbf{P} \left(\left| \overline{X}_{2n+1} - \theta \right| > \varepsilon \right) = 0$$

$\overline{X}_{2n+1}$ est alors considéré comme un estimateur convergent de θ

Partie III - B.

Le contexte de cette partie est identique à celui de la partie **III - A** précédente.

Pour tout ω de Ω , on réordonne par ordre croissant les réels $X_1(\omega), X_2(\omega), \dots, X_{2n+1}(\omega)$, et on note $\hat{X}_1(\omega), \hat{X}_2(\omega), \dots, \hat{X}_{2n+1}(\omega)$ les nombres ainsi rangés, c'est-à-dire que

$$\hat{X}_1(\omega) \leq \hat{X}_2(\omega) \leq \dots \leq \hat{X}_{2n+1}(\omega).$$

On définit ainsi $(2n+1)$ variables aléatoires $\hat{X}_1, \hat{X}_2, \dots, \hat{X}_{2n+1}$, telles que

$$\hat{X}_1 \leq \hat{X}_2 \leq \dots \leq \hat{X}_{2n+1}$$

qui constituent un réarrangement par ordre croissant des variables aléatoires $X_1, X_2, \dots, X_{2n+1}$. On admet que $\mathbf{P}\left(\left[\hat{X}_1 < \hat{X}_2 < \dots < \hat{X}_{2n+1}\right]\right) = 1$.

On s'intéresse dans cette partie à la variable aléatoire $\hat{X}_{n+1}$.

14) a) Pour tout réel x , justifier l'égalité entre événements suivante : $\left[\hat{X}_{n+1} \leq x\right] = [S_{2n+1} \geq n+1]$.

b) En déduire la fonction de répartition $\hat{F}_{n+1}$ de $\hat{X}_{n+1}$ en fonction de F (on exprimera cette fonction sous forme d'une somme que l'on ne cherchera pas à calculer).

15) On note $\hat{f}_{n+1}$ une densité de $\hat{X}_{n+1}$, et $\hat{g}_{n+1}$ une densité de $(\hat{X}_{n+1} - \theta)$.

a) Établir pour tout j de $\llbracket 0; 2n \rrbracket$, l'égalité suivante : $(j+1) \binom{2n+1}{j+1} = (2n-j+1) \binom{2n+1}{j}$.

b) En déduire, pour tout x réel, l'égalité :

$$\hat{f}_{n+1}(x) = \frac{(2n+1)!}{(n!)^2} (F(x))^n (1-F(x))^n f(x)$$

c) Établir, pour tout x réel, l'égalité suivante :

$$\hat{g}_{n+1}(x) = \frac{(2n+1)!}{(n!)^2} (G(x))^n (1-G(x))^n g(x)$$

où g et G ont été définies dans la partie I.

d) En utilisant la question 4, montrer que $\mathbf{E}(\hat{X}_{n+1}) = \theta$.

16) Dans cette question, on étudie le comportement de la suite $(\hat{X}_{n+1})_{n \in \mathbf{N}}$, lorsque n tend vers $+\infty$. On désigne par $\hat{h}_{n+1}$ une densité de la variable aléatoire $\sqrt{2n+1}(\hat{X}_{n+1} - \theta)$.

a) Montrer que pour tout réel x , on a :

$$\hat{h}_{n+1}(x) = \frac{1}{\sqrt{2n+1}} \times \frac{(2n+1)!}{(n!)^2} \left(G\left(\frac{x}{\sqrt{2n+1}}\right)\right)^n \left(1 - G\left(\frac{x}{\sqrt{2n+1}}\right)\right)^n g\left(\frac{x}{\sqrt{2n+1}}\right)$$

b) Écrire, lorsque u tend vers 0, le développement limité à l'ordre 2 de e^{-u} , et le développement limité à l'ordre 1 de $\ln(1-u)$.

c) Soit x un réel fixé. Montrer que lorsque n tend vers $+\infty$, on a :

$$G\left(\frac{x}{\sqrt{2n+1}}\right) \left(1 - G\left(\frac{x}{\sqrt{2n+1}}\right)\right) = \frac{1}{4} \left(1 - \frac{x^2}{2n+1} + o\left(\frac{1}{2n+1}\right)\right)$$

d) En déduire que :

$$\lim_{n \rightarrow +\infty} 4^n \left[G\left(\frac{x}{\sqrt{2n+1}}\right) \left(1 - G\left(\frac{x}{\sqrt{2n+1}}\right)\right) \right]^n = e^{-x^2/2}$$

e) On admet que :

$$\binom{2n}{n} \times \frac{1}{4^n} \underset{n \rightarrow +\infty}{\sim} \frac{1}{\sqrt{\pi n}}$$

Montrer alors que pour tout réel x , on a :

$$\lim_{n \rightarrow +\infty} \hat{h}_{n+1}(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

17) On admet que le résultat de la question précédente entraîne la convergence en loi de la suite de variables aléatoires $\sqrt{2n+1} \left(\hat{X}_{n+1} - \theta \right)$ vers une variable aléatoire qui suit la loi normale centrée réduite.

On rappelle que si Φ désigne la fonction de répartition de la loi normale centrée réduite, on a

$$\Phi(1,96) \simeq 0,975.$$

Montrer qu'un intervalle de confiance $[I_n; J_n]$ pour le paramètre θ au risque $\alpha = 0,05$, est donné par :

$$I_n = \hat{X}_{n+1} - \frac{1,96}{\sqrt{2n+1}} \quad \text{et} \quad J_n = \hat{X}_{n+1} + \frac{1,96}{\sqrt{2n+1}}$$
