

Exercice 1 1) a) Soit $a > 0$. Déterminer l'espérance, la variance et la fonction caractéristique de la loi uniforme sur l'intervalle $[0, a]$.

b) Soient X et Y deux variables aléatoires indépendantes de loi uniforme sur $[0, a]$. Calculer la densité de la variable $Z = X + Y$.

2) Pour tout réel $r > 0$ notons :

$$\Gamma(r) = \int_0^{+\infty} x^{r-1} e^{-x} dx.$$

a) Montrer que $\Gamma(r+1) = r\Gamma(r)$. Quel est l'expression de $\Gamma(n)$ pour $n \in \mathbb{N}^*, n \geq 2$?

b) On dit qu'une variable aléatoire X suit une loi de *gamma de paramètre* r (notée γ_r) si sa densité est

$$f(x) = \frac{1}{\Gamma(r)} x^{r-1} e^{-x} \mathbf{1}_{[0, +\infty]}(x).$$

Vérifier qu'il s'agit bien d'une densité. Calculer son espérance et sa variance. Calculer la fonction caractéristique de la loi γ_1 .

c) Soit X une variable aléatoire suivant une loi normale centrée réduite. Montrer que $\frac{X^2}{2}$ suit une loi $\gamma_{1/2}$.

d) Soient X et Y deux variables aléatoires indépendantes de loi respective γ_r et γ_s . Calculer la loi de la variable $Z = X + Y$, et en déduire que

$$\frac{\Gamma(r)\Gamma(s)}{\Gamma(r+s)} = \int_0^1 t^{r-1} (1-t)^{s-1} dt.$$

En déduire la fonction caractéristique d'une loi γ_n avec $n \in \mathbb{N}^*$.

Donner la loi fonction caractéristique d'une somme de n carrés de v.a. i.i.d. suivant une loi normale centrée réduite.

Exercice 2 Montrer que pour tout $(x_1, \dots, x_n) \in \mathbb{R}^n$, on a

$$\min_{a \in \mathbb{R}} \sum_{i=1}^n (x_i - a)^2 = \sum_{i=1}^n (x_i - \bar{x}_n)^2.$$

Exercice 3 On considère les modèles statistiques suivants :

1) $(\mathbb{R}, (\mathbb{P}_\theta)_{\theta \in \Theta})$ où $\mathbb{P}_\theta = \delta_\theta$ pour $\theta \in \Theta \subset \mathbb{R}$. Montrer que ce modèle est dominé si et seulement si, Θ est dénombrable.

2) $(\mathbb{R}, (\mathbb{P}_\theta)_{\theta \in \Theta})$ où $\mathbb{P}_\theta$ est la loi de Poisson de paramètre $\theta \in \Theta \subset \mathbb{R}_+^*$. Est ce un modèle dominé ? Si oui, quelle est sa vraisemblance ?

Exercice 4 On considère le modèle $(\mathbb{R}^n, (\mathbb{P}_\theta)_{\theta \in \Theta})$ de Kotz où la vraisemblance L est définie, pour tout $x \in \mathbb{R}^n$ et pour tout $\theta \in \mathbb{R}^n$, par :

$$L(x, \theta) = \frac{1}{n(2\pi)^{n/2}} \|x - \theta\|^2 \exp\left(-\frac{1}{2}\|x - \theta\|^2\right),$$

où $\|x\|$ désigne la norme euclidienne standard de x .

1) Vérifiez qu'il s'agit bien d'une vraisemblance.

2) Montrer que, quel que soit $\theta \in \mathbb{R}^n$, les projections canoniques X_i ne sont pas indépendantes.

3) Néanmoins, mettez en évidence le fait qu'elles ne sont pas corrélées en montrant que la matrice de covariance (qui ne dépend pas de θ) vaut $\frac{n+2}{n} I_n$.

Exercice 5 1) Soient $X_1, \dots, X_n$ des variables aléatoires réelles i.i.d. de loi $\mathbb{P}_\theta$ paramétrée par $\theta \in \Theta$. Notons $X_{(1)} = \min_{1 \leq i \leq n} X_i$ et $X_{(n)} = \max_{1 \leq i \leq n} X_i$.

Déterminer la fonction de répartition $X_{(n)}$, puis celle de $X_{(1)}$.

Lorsque $\mathbb{P}_\theta$ est absolument continue par rapport à la mesure de Lebesgue, déterminer leurs fonctions de densité.

Déterminer dans ce cas la fonction de densité du couple $(X_{(1)}, X_{(n)})$.

Déterminer la fonction de densité de la variable $\frac{X_{(1)} + X_{(n)}}{2}$.

2) On considère la *statistique d'ordre* notée φ définie par :

$$\begin{aligned} \varphi : \quad \mathbb{R}^n &\rightarrow \mathbb{R}^n \\ (x_1, \dots, x_n) &\mapsto (x_{(1)}, \dots, x_{(n)}) \end{aligned}$$

où $(x_{(1)}, \dots, x_{(n)})$ désigne l'ordonnement par ordre croissant des $\{x_1, \dots, x_n\}$, i.e. $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$.

On s'intéresse à la variable aléatoire $\varphi(X_1, \dots, X_n)$. On suppose que $\mathbb{P}_\theta$ est absolument continue par rapport à la mesure de Lebesgue λ .

a) Montrer que φ prend ses valeurs λ -presque sûrement dans :

$$\Omega = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid x_1 < \dots < x_n\}.$$

b) En remarquant que pour tout $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ il existe une permutation s de l'ensemble $\{1, \dots, n\}$ t.q. $\varphi(x) = (x_{s(1)}, \dots, x_{s(n)})$, exprimer la vraisemblance du modèle image par φ du modèle $(\mathbb{R}^n, (\mathbb{P}_\theta^{\otimes n})_{\theta \in \Theta})$ en fonction de la vraisemblance L de ce dernier modèle.

Exprimer ensuite la vraisemblance en fonction de la densité $\frac{d\mathbb{P}_\theta}{d\lambda}$.

En déduire la densité de $X_{(i)}$.

Exercice 6 On considère un échantillon i.i.d. $(X_1, \dots, X_n)$ de taille $n \geq 4$ d'un modèle réel $(\mathbb{R}, (\mathbb{P}_\theta)_{\theta \in \mathbb{R}^4})$ où pour tout $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$, θ_1 est la moyenne de la loi $\mathbb{P}_\theta$ et où pour $2 \leq i \leq 4$, θ_i est le moment centré d'ordre i de $\mathbb{P}_\theta$. En d'autres termes, pour tout $i \in \{1, \dots, n\}$, on a

$$\theta_1 = E(X_i) \quad \text{et} \quad \theta_j = E((X_i - \theta_1)^j) \quad \text{pour } j = 2, 3, 4.$$

On considère la moyenne empirique de $\overline{X_n}$ et la variance empirique S_n^2 :

$$\overline{X_n} = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{et} \quad S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \overline{X_n})^2.$$

1) Montrer que $V_\theta(S_n^2)$ ne change pas si l'on suppose $\theta_1 = 0$.

2) Montrer que

$$S_n^2 = \frac{1}{2n(n-1)} \sum_{i=1}^n \sum_{j=1}^n (X_i - X_j)^2.$$

En déduire que

$$V_\theta(S_4^2) = \frac{1}{4}(\theta_4 - \frac{1}{3}\theta_2^2).$$

3) Démontrez alors, par récurrence sur n que

$$V_\theta(S_n^2) = \frac{1}{n}(\theta_4 - \frac{n-3}{n-1}\theta_2^2).$$

4) Déterminez $Cov(\overline{X_n}, S_n^2)$ en fonction de $\theta_1, \theta_2, \theta_3$ et θ_4 . Sous quelle condition a-t-on $Cov(\overline{X_n}, S_n^2) = 0$?

5) Pour quelles valeurs de θ les variables $\overline{X_n}$ et S_n^2 ne sont pas indépendantes ?