Correction du Devoir surveillé

Exercice 1 : (Questions du cours)

Les trois questions de cet exercice sont indépendantes :

1. Soit la fonction f définie sur $\mathbb{R}$ par :

$$f(t) = \begin{cases} 2t & \text{si} \quad 0 < t < 1 \\ 0 & \text{sinon.} \end{cases}$$

(a) Vérifier que f définit une densité de probabilité.

. $f(t) \geq 0 \ \forall t$.

. $\int_{-\infty}^{+\infty} f(t)dt = \int_0^1 2tdt = [t^2]_0^1 = 1$.

Donc f est bien une densité de probabilité.

(b) Soit T une variable aléatoire de densité f. Calculer son espérance $E(T)$ et sa variance $V(T)$.

$$E[T] = \int_{-\infty}^{-\infty} t.f(t)dx = \int_0^1 2t^2 dt = [\frac{2}{3}t^3]_0^1 = \frac{2}{3}$$

$$V[T] = E[T^2] - E[T]^2 = \int_0^1 2t^3 - (\frac{2}{3})^2 = [\frac{2}{4}t^4]_0^1 - (\frac{2}{3})^2 = \frac{2}{4} - \frac{4}{9} = \frac{1}{18}$$

2. Soit X une variable aléatoire qui suit une loi exponentielle de paramétre $\lambda > 0$. Donner sans justification l'espérance de la variable aléatoire X ainsi que sa variance.

$X \sim \mathcal{E}(\lambda)$, sa fonction de densité est défine par

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si} \quad x > 0 \\ 0 & \text{sinon.} \end{cases}$$

$$E[X] = \frac{1}{\lambda} \quad , \quad V[X] = \frac{1}{\lambda^2}$$

3. Soient X et Y deux variables aléatoires qui suivent respectivement une loi normale $\mathcal{N}(1,1)$ et une loi normale $\mathcal{N}(2,1)$. On suppose que X et Y sont indépendantes. Quelle est la loi suivie par la variable aléatoire $X - Y$ en justifiant votre réponse.

$$X \sim \mathcal{N}(1,1) \ , \ Y \sim \mathcal{N}(2,1)$$

X et Y sont indépendantes alors : $X - Y \sim \mathcal{N}(1-2, \sqrt{1+1}) = \mathcal{N}(-1, \sqrt{2})$.
$\Rightarrow$ Stabilité de la loi Normale.

Exercice 2

Un ordinateur est programmé pour effectuer des calculs de comptabilité à partir des bases de données afin d'obtenir des bilans (un par base de donnée).
On dit qu'un bilan est "éronné" s'il y a au moins une erreur de comptabilité pour le bilan. On définit la variable aléatoire X qui représente le nombre de bilans érronés. Lorsque le nombre de bilans traités devient assez grand, X peut être approchée par une loi normale de moyenne égale à 20 bilans et de variance $19,2$.

1. Calculer la probabilité pour que le nombre de bilans érronés ne dépasse pas 15.
X : ' Nombre des bilans érronés '

$$X \sim \mathcal{N}(20, \sqrt{19,2})$$

$$P(X < 15) = P(\frac{X-20}{\sqrt{19,2}} < \frac{15-20}{\sqrt{19,2}}) = P(Z < \frac{15-20}{\sqrt{19,2}})$$

$$= P(Z < \frac{-5}{\sqrt{19,2}}) = P(Z > \frac{5}{\sqrt{19,2}} = P(Z > 1,141) = 0,12714$$

où $Z \sim \mathcal{N}(0,1)$.

2. Calculer la probabilité tel que le nombre de bilans érronés soit compris entre 17 et 22.

$$P(17 < X < 22) = P(\frac{17-20}{\sqrt{19,2}} < Z < \frac{22-20}{\sqrt{19,2}}) = P(Z < \frac{2}{\sqrt{19,2}}) - P(Z < \frac{-3}{\sqrt{19,2}})$$

$$= 1 - P(Z > \frac{2}{\sqrt{19,2}}) - P(Z > \frac{3}{\sqrt{19,2}})$$

$$= 1 - P(Z > 0,45) - P(Z > 0,68)$$

$$= 1 - 0,32636 - 0,24825 = 0,42539$$

3. Trouver la valeur entière de x_0 telle que $P(20 - x_0 \leq X \leq 20 + x_0) = 0.95$.

$$P(20 - x_0 < X < 20 + x_0) = 0,95$$

$$P(-x_0 < X - 20 < x_0) = 0,95$$

$$P(\frac{-x_0}{\sqrt{19,2}} \leq \frac{X-20}{\sqrt{19,2}} \leq \frac{x_0}{\sqrt{19,2}}) = 0,95$$

$$P(Z < \frac{x_0}{\sqrt{19,2}}) - P(Z < \frac{-x_0}{\sqrt{19,2}}) = 0,95$$

$$1 - P(Z \geq \frac{x_0}{\sqrt{19,2}}) - P(Z \geq \frac{x_0}{\sqrt{19,2}}) = 0,95$$

$$2P(Z \geq \frac{x_0}{\sqrt{19,2}}) = 0,05$$

$$P(Z \geq \frac{x_0}{\sqrt{19,2}}) = 0,025$$

Par la suite $\frac{x_0}{\sqrt{19,2}} = 1,96$, d'où $x_0 = 1,96 \times \sqrt{19,2} = 8,101$.

Exercice 3

Au second tour de la présidentielle, les électeurs de Syldavie doivent choisir entre deux condidats c et C_1 (On suppose que tous les électeurs se prononcent). Depuis peu, les habitants votent sur les bornes électroniques. On suppose que la durée de vie d'une borne (exprimée en années) est modélisée par une variable aléatoire continue notée X de fonction de répartition :

$$F_\theta(x) = \begin{cases} 1 - e^{\frac{-1}{2}\frac{x^2}{\theta^2}} & \text{si } \; x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$$

où $\theta > 0$ est un paramètre inconnu à estimer. Pour étudier le paramètre θ, on a effectué une suite de n expériences indépendantes qui ont donné les réalisations $x_1, , x_n$ de n variables aléatoires $X_1, , X_n$ i.i.d. de même loi que X.

1. Déterminer la fonction de densité f_θ de X.

$$f_\theta(x) = F_\theta^{'}(x) = \begin{cases} \dfrac{x}{\theta^2} e^{\frac{-x^2}{2\theta^2}} & \text{si } \; x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$$

2. (a) Calculer E(X). (On admettra que $\int_0^{+\infty} x^2 e^{-\frac{x^2}{2\theta^2}} dx = \theta^3 \sqrt{\frac{\pi}{2}}$).

$$E[X] = \int_0^{+\infty} \frac{x^2}{\theta^2} e^{\frac{-x^2}{2\theta^2}} dx$$

$$= \frac{1}{\theta^2}\theta^3\sqrt{\frac{\pi}{2}} dx = \theta\sqrt{\frac{\pi}{2}}$$

   (b) En déduire un estimateur T_1 de θ par la méthode des moments.

$$E[X] = \overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i = \theta\sqrt{\frac{\pi}{2}}$$

D'où un estimateur de θ par la méthode des moments est donné par

$$\tilde{\theta}_n = \sqrt{\frac{2}{\pi}}\frac{1}{n}\sum_{i=1}^{n} X_i = \sqrt{\frac{2}{\pi}}\overline{X_n}$$

3. (a) Déterminer la fonction de vraisemblance $L(x_1, x_n, , x_n, \theta)$.

$$L(x_1, x_n, , x_n, \theta) = \prod_{i=1}^{n} f_\theta(x_i)$$

$$= \prod_{i=1}^{n} \frac{x_i}{\theta^2} e^{\frac{-x_i^2}{2\theta^2}}$$

$$= \frac{1}{\theta^{2n}} (\prod_{i=1}^{n} x_i) e^{\frac{-1}{2\theta^2} \sum_{i=1}^{n} x_i^2}$$

(b) Montrer que la log-vraisemblance associée aux observations $x_1, , x_n$ s'écrit :

$$lnL(x_1, x_n, , x_n, \theta) = \sum_{i=1}^{n} ln(x_i) - 2nln(\theta) - \frac{1}{2\theta^2} \sum_{i=1}^{n} x_i^2$$

$$lnL(x_1, x_n, , x_n, \theta) = ln(\frac{1}{\theta^{2n}} (\prod_{i=1}^{n} x_i) e^{\frac{-1}{2\theta^2} \sum_{i=1}^{n} x_i^2})$$

$$= -2nln(\theta) + \sum_{i=1}^{n} ln(x_i) - \frac{1}{2\theta^2} \sum_{i=1}^{n} x_i^2$$

c) En déduire un estimateur T_2 de θ par la méthode du maximum de vraisemblance.

$$\frac{\partial}{\partial\theta} (lnL(x_1, x_n, , x_n, \theta) = -\frac{2n}{\theta} + \frac{1}{\theta^3} \sum_{i=1}^{n} x_i^2$$

$$\frac{\partial}{\partial\theta} (lnL(x_1, x_n, , x_n, \theta) = 0 \iff \frac{1}{\theta} [-2n + \frac{1}{\theta^2} \sum_{i=1}^{n} x_i^2] = 0$$

$$\iff -2n + \frac{1}{\theta^2} \sum_{i=1}^{n} x_i^2 = 0$$

$$\iff \frac{1}{\theta^2} \sum_{i=1}^{n} x_i^2 = 2n$$

$$\iff \theta^2 = \frac{1}{2n} \sum_{i=1}^{n} x_i^2$$

Donc un estimateur par la méthode de vraisemblance est donné par :

$$\widehat{\theta}_n = \sqrt{\frac{1}{2n}\sum_{i=1}^{n} X_i^2}$$

On vérifie de plus que :

$$\frac{\partial^2}{\partial^2\theta}(lnL(x_1, x_n, , x_n, \theta)_{\diagup\theta=\widehat{\theta}} \leq 0.$$

(d) Si on considère un échantillon de taille 100, calculer une estimation de θ par la méthode des moments sachant que $\sum_{i=1}^{100} x_i = 276$

$$\tilde{\theta}_n = \sqrt{\frac{2}{\pi}}.\frac{1}{100}\sum_{i=1}^{100} x_i = \sqrt{\frac{2}{\pi}}.\frac{1}{100}.276 = ...$$