Préparation au Capes de Mathématiques

Probabilités - Problèmes des médianes

(Concernant l'indépendance des différentes parties : sauf la question II.C.5, la partie II est indépendante de la partie I ; sauf la question III.2, la partie III est indépendante des parties I et II.)

Les variables aléatoires réelles (v.a.r.) X que nous considérons dans ce problème sont de deux types :
1/ X est à valeurs dans $\mathbb{Z}$
2/ la loi de probabilité de X admet une densité f, strictement positive sur un intervalle fermé I et nulle en dehors de I, continue, sauf éventuellement aux bornes de l'intervalle I.

Préliminaires (non notés)
Soit X une v.a.r. et F sa fonction de répartition (f.d.r.), c'est-à-dire, $\forall x \in \mathbb{R}, F(x) = P(X \leq x)$. Selon que la v.a.r. X est de type 1 ou de type 2, donner une expression de la f.d.r. F et rappeler ses principales propriétés.

Partie I

On définit l'ensemble des *médianes* de X par :

$$M(X) = \{m \in \mathbb{R}\,; P(X < m) \leq \frac{1}{2} \leq P(X \leq m)\} \quad (= \{m \in \mathbb{R}\,; \lim_{x \to m_-} F(x) \leq \frac{1}{2} \leq F(m)\}).$$

I.1. Montrer que, lorsque X est une v.a.r. de type 2, $M(X)$ est un singleton.

I.2. Soit X une v.a.r. de type 1, F sa f.d.r. et A le sous-ensemble de $\mathbb{Z}$ défini par :

$$A = \{m \in \mathbb{Z}\,; P(X \leq m) \geq 1/2\}.$$

Montrer que A admet un plus petit élément m_0.
En déduire que, si $F(m_0) > \frac{1}{2}$, alors on a $M(X) = \{m_0\}$ et si $F(m_0) = \frac{1}{2}$, alors il existe un entier naturel non nul n tel que l'on ait $M(X) = [m_0, m_0 + n]$.

I.3. Déterminer $M(X)$ dans les cas suivants :
a. X suit la loi de Bernoulli de paramètre p (avec $p \in]0, 1[$);
b. X suit la loi binomiale $B(n, 1/2)$ (avec $n \in \mathbb{N}^*$);
c. X suit la loi normale centrée réduite $N(0, 1)$;
d. X suit la loi exponentielle de paramètre λ (avec $\lambda \in \mathbb{R}_+^*$).

I.4. Soit a un nombre réel. Montrer les égalités :

$$M(aX) = \{am\,; m \in M(X)\} \text{ et } M(a + X) = \{a + m\,; m \in M(X)\}.$$

I.5 L'assertion suivante est-elle exacte : pour tout couple (X, Y) de v.a.r. de type 1, $M(X + Y) = \{m + m'\,; m \in M(X), m' \in M(Y)\}$?

I.6. Soit X une v.a.r. admettant un moment d'ordre 1.
L'assertion suivante est-elle exacte : $E(X) \in M(X)$?

I.7. Soit X une v.a.r. admettant un moment d'ordre 2 et m une médiane de X.
Montrer l'inégalité ($\mathbb{I}_A$ désigne l'indicatrice de l'événement A et $\sigma^2(X)$ la variance de X) :

$$\sigma^2(X) \geq \max(E((X - E(X))^2 \mathbb{I}_{[X \geq m]}), E((X - E(X))^2 \mathbb{I}_{[X \leq m]})).$$

1.8. En distinguant les cas $m \geq E(X)$ et $m \leq E(X)$, déduire de l'inégalité précédente l'inégalité : $\mid E(X) - m \mid \leq \sqrt{2\sigma^2(X)}$.

Partie II

On pose : $\forall n \in \mathbb{N}^*, \forall \lambda \in \mathbb{R}_+^*,\ P_n(\lambda) = e^{-\lambda} \sum_{k=0}^{n} \frac{\lambda^k}{k!}$ et $I_n = \int_0^1 ((1-x)e^x)^n dx$.

II.A.1. Montrer : $\forall n \in \mathbb{N}^*,\ \forall \lambda \in \mathbb{R}_+^*,\ 1 - P_n(\lambda) = \frac{1}{n!} \int_0^{\lambda} e^{-x} x^n dx$.

II.A.2. En déduire : $\forall n \in \mathbb{N}^*,\ 1 - P_n(n) = e^{-n} \frac{n^{n+1}}{n!} I_n$ et $\forall \lambda \in \mathbb{R}_+^*,\ P_n^{'}(\lambda) = -e^{-\lambda} \frac{\lambda^n}{n!}$

II.A.3. On pose : $\Psi(0) = \frac{1}{2}$ et $\forall x \in]0,1[,\ \Psi(x) = \frac{-x-\ln(1-x)}{x^2}$. Montrer que Ψ est une bijection croissante de [0,1[sur $[\frac{1}{2}, +\infty[$.

II.A.4. Montrer : $\forall n \in \mathbb{N}^*,\ I_n = \int_0^1 e^{-nx^2\Psi(x)} dx \leq \sqrt{\frac{\pi}{2n}}$

II.A.5. Soit $\sigma \in]0,1[$ et $\delta = \Psi^{-1}(\frac{1}{2\sigma^2})$. Montrer : $\forall n \in \mathbb{N}^*,\ I_n \geq \frac{\sigma}{\sqrt{n}} \int_0^{\frac{\delta\sqrt{n}}{\sigma}} e^{-\frac{x^2}{2}} dx \geq \frac{\sigma}{\sqrt{n}} \int_0^{\delta\sqrt{n}} e^{-\frac{x^2}{2}} dx$

II.A.6. En déduire les équivalents au voisinage de l'infini :

$$I_n \sim \sqrt{\frac{\pi}{2n}} \text{ et } 1 - P_n(n) \sim \frac{n^n}{n!} e^{-n} \sqrt{\frac{\pi n}{2}}.$$

II.B. Soit $n \in \mathbb{N}^*$, $X_1, ..., X_n, n$ variables aléatoires indépendantes et de même loi de Poisson de paramètre 1 et $S_n = X_1 + ... + X_n$.

II.B.1. Montrer, par récurrence sur $\mathbb{N}^*$, que S_n suit une loi de Poisson de paramètre n.

II.B.2. En déduire que $P_n(n)$ tend vers $1/2$ lorsque n tend vers l'infini (utiliser le théorème de la limite centrale).

II.B.3. En déduire un équivalent pour $n!$ lorsque n tend vers l'infini (formule de Stirling).

II.C.1. Montrer : $\forall n \in \mathbb{N}^*,\ n \geq 2,\ \forall \lambda \in \mathbb{R}_+^*,\ P_n(\lambda) + P_n^{'}(\lambda) = P_{n-1}(\lambda)$.

II.C.2. En appliquant une formule de Taylor à l'ordre 2 entre $n-1$ et n, montrer que la suite $(P_n(n))$ est strictement décroissante.

II.C.3. Procéder de manière analogue pour montrer que la suite $(P_{n-1}(n))$ est strictement croissante.

II.C.4. Montrer que les suites $(P_{n-1}(n))$ et $(P_n(n))$ sont adjacentes.

II.C.5. En déduire que si X est une variable aléatoire de loi de Poisson de paramètre n ($n \in \mathbb{N}^*$), on a $M(X) = \{n\}$.

Partie III

Soit X une v.a.r. vérifiant, dans cette troisième partie l'une des trois hypothèses suivantes :
a. X suit une loi normale centrée et réduite ;
b. X suit une loi de Poisson de paramètre 1 ;
c. $X/2$ suit une loi de Bernoulli de paramètre $1/2$.

III.1. On pose $Y = X - E(X)$. Vérifier que Y est une variable aléatoire centrée et réduite.

Soit n un entier naturel non nul et $X_1, ..., X_n$, n v.a.r. indépendantes en probabilités et de même loi que X. On pose : $\forall k \in [1, n],\ Y_k = X_k - E(X_k)$ et $T_k = \sum_{j \leq k} Y_j$.

III.2. Montrer que, pour tout entier k de $[1, n]$, la variable aléatoire $T_n - T_k$ est bien du type 1 ou du type 2 et que, de plus, 0 est une médiane de $T_n - T_k$.

Soit $x \in \mathbb{R}_+$. On pose :

$$\Omega_1 = [T_1 > x] \text{ et } \forall k \in [2, n],\ \Omega_k = [\max\{T_j; j < k\} \leq x] \cap [T_k > x].$$

III.3. Montrer que les événements $(\Omega_k)_{k=1,\ldots,n}$ sont 2 à 2 incompatibles et de réunion $[\max\{T_j; j \leq n\} > x]$.

III.4. Montrer, pour tout entier k de $[1, n]$, et tout ε de $\mathbb{R}_+$ les inclusions suivantes :
(i) $[T_n - T_k \geq 0] \cap \Omega_k \subset [T_n > x] \cap \Omega_k$
(ii) $[T_n - T_k > 0] \cap \Omega_k \supset [T_n > x + \varepsilon] \cap \Omega_k \cap [Y_k < \varepsilon]$.

III.5. En déduire l'encadrement :

$$P([T_n > x + \varepsilon]) - \sum_{k=1}^{n} P(Y_k \geq \varepsilon) \leq \frac{1}{2} P([\max\{T_j; j \leq n\} > x]) \leq P([T_n > x]).$$

(On pourra écrire : $[T_n > x] = \cup_{k=1}^{n}([T_n > x] \cap \Omega_k)$ et utiliser (i), puis
$[T_n > x + \varepsilon] = \cup_{k=1}^{n}([T_n > x + \varepsilon] \cap \Omega_k \cap [Y_k < \varepsilon]) \cup \cup_{k=1}^{n}([T_n > x + \varepsilon] \cap \Omega_k \cap [Y_k \geq \varepsilon])$
et utiliser (ii).)

III.6. Soit θ un réel strictement positif.
Montrer que la suite de terme général $\sum_{j=1}^{n} P([Y_j \geq \theta\sqrt{n}])$ converge vers 0 .
(On pourra appliquer l'inégalité de Markov à Y_j^4.)

III.7. Montrer :

$$\forall \lambda \in \mathbb{R}_+, \lim_{n \to \infty} P([\max\{T_j; j \leq n\} > \lambda\sqrt{n}]) = \sqrt{\frac{2}{\pi}} \int_{\lambda}^{+\infty} e^{-\frac{x^2}{2}}\, dx.$$

III.8. Que se passe-t-il lorsque λ est négatif ?