Banque « Agro »
A et AE - 0502

MATHÉMATIQUES
Épreuve B

Durée : 3 heures 30 minutes

L'usage de la calculatrice est interdit pour cette épreuve.

Les parties I et II sont indépendantes.

Préliminaire.

La plupart des expériences aléatoires conduisent à l'étude de variables aléatoires réelles obéissant à des lois dont le type est connu, mais qui dépendent d'un paramètre réel lié à l'expérience. Ce problème a pour objectif de donner des méthodes afin d'estimer la valeur numérique de ce paramètre, que nous désignerons dans la suite par la lettre θ.

Notations.

Si X est une variable aléatoire, on notera, sous réserve d'existence, $E(X)$ l'espérance de X et $V(X)$ sa variance ; F_X désignera la fonction de répartition de X et f_X une densité éventuelle. Soit n un entier naturel non nul, on appellera $n-$échantillon de X toute suite $(X_1, \ldots, X_n)$ de variables aléatoires mutuellement indépendantes et suivant la même loi que X.

Si X est une variable aléatoire dont la loi dépend d'un paramètre θ, et $(X_1, \ldots, X_n)$ un n-échantillon de X, une variable aléatoire T_n, fonction de $(X_1, \ldots, X_n)$, est un estimateur sans biais et convergent de θ si $E(T_n) = \theta$ et $\lim\limits_{n \to +\infty} V(T_n) = 0$. Le terme estimateur utilisé dans ce problème signifie par abus de langage que ces deux conditions sont vérifiées.

Enfin, on dira que l'estimateur T_n est meilleur que l'estimateur T'_n si pour tout n, entier naturel assez grand, $V(T_n) \leq V(T'_n)$.

0. Exemple introductif.

Une population donnée contient une proportion inconnue, θ, d'individus possédant un certain caractère. On considère une variable aléatoire X suivant la loi de Bernoulli de paramètre θ. On prélève, avec remise, n individus de cette population, et on désigne par X_i la variable aléatoire associée au i-ème tirage. Enfin, pour tout n, entier naturel non nul, on note $T_n = \frac{X_1 + \ldots + X_n}{n}$.

0.1. Rappelez les valeurs de $E(X)$ et de $V(X)$.

0.2. Montrer que T_n est un estimateur de θ.

I. Première partie

Soit θ un réel strictement positif ; soient X une variable aléatoire de loi uniforme sur l'intervalle $]0, \theta[$, et $(X_1, \ldots, X_n)$ un $n-$échantillon de X.

I.1.a. Déterminer une densité f_X de la variable aléatoire X, ainsi que sa fonction de répartition F_X, puis exprimer, en fonction de θ, les valeurs de $E(X)$ et de $V(X)$.

I.1.b. Pour tout n, entier naturel non nul, on pose $T_n = \dfrac{2}{n}(X_1 + \ldots + X_n)$; montrer que T_n est un estimateur de θ.

On se propose, dans la suite de cette première partie, de construire d'autres estimateurs de θ liés à cette dernière variable aléatoire, T_n.

I.2. Soit n un entier naturel non nul, on considère les deux variables aléatoires

$$Y = \max(X_1, \ldots, X_n) \text{ et } Z = \min(X_1, \ldots, X_n).$$

I.2.a. Déterminer en fonction de θ les expressions de F_Y, f_Y, $E(Y)$, et $V(Y)$.

I.2.b. Déterminer en fonction de θ les expressions de F_Z, f_Z, $E(Z)$, et $V(Z)$.
Indication : Pour déterminer F_Z, on pourra calculer $P[Z > z]$.

I.2.c. On remarque que $V(Z) = V(Y)$. Justifier cette égalité sans recourir au calcul.

I.3. Pour $n \in \mathbb{N}^*$, on considère la variable aléatoire $T'_n = \frac{n+1}{n} Y$.

I.3.a. Montrer que T'_n est un estimateur de θ.

I.3.b. T'_n est-il un meilleur estimateur que l'estimateur T_n défini à la question I.1.b ?

I.4. Pour $n \in \mathbb{N}^*$, on considère la variable aléatoire $T''_n = Y + Z$.

I.4.a. On rappelle que la covariance de Y et de Z, notée $Cov(Y, Z)$ vérifie :
$|Cov(Y, Z)| \leq \sqrt{V(Y)V(Z)}$. En déduire que $V(T''_n) \leq 4V(Y)$.

I.4.b. Déduire du 1.4.a. et du 1.1.b que T''_n est un estimateur de θ, meilleur que T_n.

II. Deuxième Partie. Loi de Pareto.

II.1 La loi $\gamma(p, \lambda)$.
p et λ désignent dans cette partie deux réels strictement positifs.
On désigne par $I(p, \lambda)$ l'intégrale :

$$I(p, \lambda) = \int_0^{+\infty} \lambda^p x^{p-1} e^{-\lambda x} dx$$

et on note $\Gamma(p) = I(p, 1)$.

II.1.a. Montrer que pour tout x, réel strictement positif, $x^{p-1}e^{-\lambda x} \leq x^{p-1}$, puis montrer que pour tout a réel, strictement positif, $\int_0^a x^{p-1}dx$ converge. En déduire la convergence de l'intégrale $I(p,\lambda)$ à la borne zéro.

II.1.b. Montrer que pour x suffisamment grand, $0 \leq x^{p-1}e^{-\lambda x} \leq \frac{1}{x^2}$; puis montrer que pour tout a réel, strictement positif, $\int_a^{+\infty} \frac{dx}{x^2}$ converge. En déduire la convergence de l'intégrale $I(p,\lambda)$ à la borne $+\infty$.

II.1.c. En effectuant le changement de variable $x = \frac{u}{\lambda}$, montrer que $I(p,\lambda) = \Gamma(p)$.

II.1.d. Calculer $\Gamma(1)$ et montrer, à l'aide d'une intégration par parties, que $\Gamma(p+1) = p\Gamma(p)$. En déduire que pour tout n, entier naturel non nul, $\Gamma(n) = (n-1)!$

II.1.e. Montrer que la fonction g définie par $g(u) = 0$ si $u \leq 0$ et $g(u) = \frac{1}{\Gamma(p)}\lambda^p u^{p-1}e^{-\lambda u}$ sinon, est la densité de probabilité d'une variable aléatoire U ; puis vérifier que $E(U) = \frac{p}{\lambda}$ et $V(U) = \frac{p}{\lambda^2}$.

Notation : On dira dans la suite du problème qu'une telle variable aléatoire U suit la loi $\gamma(p,\lambda)$.

II.2. Soient θ un réel strictement positif et X la variable aléatoire de densité f_X définie comme suit :

$$f_X(x) = 0 \text{ si } x \leq 1 \quad et \quad f_X(x) = \frac{1}{\theta}x^{-\frac{\theta+1}{\theta}} \text{ sinon.} \qquad (\textit{loi de Pareto})$$

II.2.a. Vérifier que f_X est bien une densité de probabilité et déterminer F_X.

II.2.b. Pour quelles valeurs de θ la variable aléatoire X admet-elle une espérance? Une variance? Calculer $E(X)$ et $V(X)$ lorsque ces valeurs existent.

II.2.c. Soit $Y = ln(X)$. On admettra que Y est une variable aléatoire ; déterminer en fonction de θ les expressions de F_Y, f_Y, $E(Y)$ et $V(Y)$.

II.2.d. On considère un $n-$échantillon $(Y_1, \ldots Y_n)$ de la variable aléatoire Y définie ci-dessus et on pose $S_n = Y_1 + \ldots + Y_n$ et $T_n = \frac{S_n}{n}$; montrer que T_n est un estimateur de θ.

II.2.e. On rappelle que pour tout couple (U,V) de variables aléatoires indépendantes de densités respectives f_U et f_V la variable $W = U + V$ admet une densité f_W définie par :

$$\forall w \in \mathbb{R},\ f_W(w) = \int_{-\infty}^{+\infty} f_U(w-t)f_V(t)dt.$$

Vérifier que, si U et V sont à valeurs positives, alors $\forall w \in \mathbb{R}, f_W(w) = \int_0^w f_U(w-t)f_V(t)dt$.

T.S.V.P.

II.2.f. On considère deux variables aléatoires indépendantes suivant la même loi exponentielle de paramètre λ, notée $\mathcal{E}(\lambda)$; déterminer une densité de leur somme, puis, plus généralement, montrer qu'une densité de la variable aléatoire S_n définie au II.2.d. est :

$$f_n(s) = \frac{s^{n-1}}{\theta^n (n-1)!} e^{-(s/\theta)} \quad \textit{pour} \quad s > 0 \quad \textit{et} \quad 0 \quad \textit{sinon.}$$

Quelle est la loi suivie par S_n?

FIN