CB2 - sujet 2 type Maths II - Vendredi 14 mars 2025

Toutes les variables aléatoires qui interviennent dans ce problème sont supposées définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$

Pour toute variable aléatoire réelle X admettant une densité f_X sur $\mathbb{R}$, on note $\mathcal{D}_X$ l'ensemble des réels x tels que la variable aléatoire e^{xX} admette une espérance.

On note G_X la fonction définie sur $\mathcal{D}_X$ par : $G_X(x) = E\left(e^{xX}\right)$

On dit que G_X est la fonction génératrice de X. On admet que si X et Y sont telles que G_X et G_Y coïncident sur un intervalle ouvert non vide alors X et Y ont même loi.

Dans tous les programmes Python ci-dessous, on suppose que l'on a déjà fait les imports `import numpy as np` et `import numpy.random as rd`

Partie I : Etude de fonctions génératrices

1. Généralités
   Soit X une variable aléatoire de densité f_X.

   (a) Vérifier que $0 \in \mathcal{D}_X$ et préciser la valeur de $G_X(0)$.

   (b) Etant donnés a et b réels tels que $a \neq 0$, montrer que $x \in \mathcal{D}_{aX+b}$ si et seulement si $ax \in \mathcal{D}_X$ et montrer que dans ce cas, $G_{aX+b}(x) = e^{xb}.G_X(ax)$.

   (c) Soit $n \in \mathbb{N}^*$ et $X_1, \dots, X_n$ des variables aléatoires mutuellement indépendantes et de même loi que X, on note $S_n = \sum_{i=1}^{n} X_i$.
   Montrer que $\forall x \in \mathcal{D}_X, G_{S_n}(x) = (G_X(x))^n$

   (d) Justifier que pour tout $x \in \mathcal{D}_X$, $G_X(x) = \int_{-\infty}^{+\infty} e^{xt} f_X(t)dt$.

2. Cas de la loi normale centrée-réduite
   On suppose que $U \hookrightarrow \mathcal{N}(0,1)$

   (a) Montrer que G_U est définie sur $\mathbb{R}$ et que $\forall x \in \mathbb{R}, G_U(x) = e^{\frac{x^2}{2}}$

   (b) Montrer que, pour tout $n \in \mathbb{N}$, U admet un moment d'ordre n, c'est-à-dire que $E(U^n)$ existe.

   (c) Montrer que si n est impair, alors $E(U^n) = 0$.

   (d) Montrer que : $\forall p \geq 1, E\left(U^{2p}\right) = (2p-1)E\left(U^{2p-2}\right)$

   (e) Pour $p \in \mathbb{N}$, donner l'expression de $E\left(U^{2p}\right)$ en fonction de p à l'aide de factorielles.

   (f) Montrer que : $\forall x \in \mathbb{R}, G_U(x) = \sum_{n=0}^{+\infty} \frac{E\left(U^n\right)}{n!} x^n$

3. Cas de la loi exponentielle
   Soient λ un réel strictement positif, on suppose que $Y \hookrightarrow \mathcal{E}(\lambda)$

   (a) Déterminer $\mathcal{D}_Y$ et montrer que pour tout $x \in \mathcal{D}_Y$, $G_Y(x) = \dfrac{\lambda}{\lambda - x}$.

   (b) Vérifier que $\forall x \in]-\lambda, \lambda[, G_Y(x) = \displaystyle\sum_{n=0}^{+\infty} \frac{E(Y^n)}{n!} x^n$

   (c) De quel réel (exprimé sous forme d'une fraction) le résultat affiché par le script Python ci-dessous devrait-il être proche pour n suffisamment grand ?
   Justifier et donner le nom de la méthode utilisée.

```
n=int(input('n='))
Y=rd.exponential(1/2,n)
M=np.mean(np.exp(-Y))
print(M)
```

4. Convergence
   Pour $n \in \mathbb{N}^*$, on suppose $\lambda > 0$ et $Y_1, \ldots, Y_n$ des variables indépendantes, identiquement distribuées de loi $\mathcal{E}(\lambda)$ et on note $Z_n = \sum_{i=1}^{n} Y_i$.

   (a) Soit $n \in \mathbb{N}^*$
   - Exprimer la variable aléatoire centrée-réduite Z_n^* comme fonction affine de Z_n.
   - En déduire que $\mathcal{D}_{Z_n^*} = \left]-\infty, \sqrt{n}\right[$ et que $\forall x < \sqrt{n}$, $G_{Z_n^*}(x) = e^{-\sqrt{n}x}\left(1 - \dfrac{x}{\sqrt{n}}\right)^{-n}$

   (b) Montrer que $\forall x \in \mathbb{R}, \lim_{n\to+\infty} G_{Z_n^*}(x) = G_U(x)$ où $U \hookrightarrow \mathcal{N}(0,1)$.
   Ce résultat est-il étonnant?

Partie II : Loi du χ^2 centrée *(se lit "loi du khi-deux")*

On suppose que $n \in \mathbb{N}^*$ et que $U_1, \ldots, U_n$ sont des variables aléatoires mutuellement indépendantes et suivant toutes la loi $\mathcal{N}(0,1)$, on note $W_n = \sum_{i=1}^{n} U_i^2$.

1. Soit X une variable aléatoire suivant la loi $\gamma(\nu)$ où $\nu > 0$ (loi petit gamma de paramètre ν).
   Montrer que $\mathcal{D}_X =]-\infty, 1[$ et que $\forall x \in \mathcal{D}_X$, $G_X(x) = \dfrac{1}{(1-x)^{\nu}}$

2. Ecrire en Python un script simulant et affichant W_n, n étant entré par l'utilisateur.

3. (a) Déterminer la fonction génératrice de $V_1 = \dfrac{U_1^2}{2}$.

   (b) En déduire que $V_1 \hookrightarrow \gamma\left(\frac{1}{2}\right)$.

   (c) Donner la loi de $\frac{W_n}{2}$ pour tout $n \in \mathbb{N}^*$.

4. (a) Préciser l'espérance et la variance de W_n et montrer qu'une densité de W_n est la fonction f_n définie par $f_n(t) = \begin{cases} 0 & \text{si } t \leq 0 \\ \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} t^{\frac{n}{2}-1} e^{-\frac{t}{2}} & \text{si } t > 0 \end{cases}$.

   On dit que W_n suit la loi du χ^2 centré à n degrés de liberté notée $\chi^2(n)$

(b) Donner l'allure de la courbe représentative de f_4 en précisant la convexité et la demi-tangente à droite au point d'abscisse 0 .

5. Justifier que $\mathcal{D}_{W_n} =]-\infty, \frac{1}{2}[$ et que $\forall x < \frac{1}{2}$, $G_{W_n}(x) = (1-2x)^{-\frac{n}{2}}$

6. Pour $n \in \mathbb{N}^*$, on note F_n la fonction de répartition de W_n.

   Justifier que F_n réalise une bijection de $]0, +\infty[$ sur $]0, 1[$.

Partie III : Loi du χ^2 décentrée

On considère une suite $(M_i)_{i \in \mathbb{N}^*}$ de variables aléatoires mutuellement indépendantes telles que, pour $i \in \mathbb{N}^*$, $M_i \hookrightarrow \mathcal{N}(m_i, 1)$ où $m_i \in \mathbb{R}$.

Pour tout $n \in \mathbb{N}^*$, on note $S_n = \sum_{i=1}^{n} M_i^2$ et $\lambda_n = \sum_{i=1}^{n} m_i^2$

On dit que S_n suit la loi du χ^2 décentrée à n degrés de liberté, notée $\chi^2(n, \lambda_n)$

1. Pour $n \in \mathbb{N}^*$, exprimer $E(S_n)$ en fonction de n et λ_n

2. (a) Montrer que $\forall x < \frac{1}{2}$, $G_{M_1^2}(x) = (1-2x)^{-\frac{1}{2}} \times \exp\left(\dfrac{m_1^2 x}{1-2x}\right)$

   (b) En déduire, pour $x < \frac{1}{2}$, l'expression de $G_{S_n}(x)$ en fonction de x, n et λ_n

3. Nombre aléatoire de degrés de liberté

   Soit $\lambda > 0$, on note K une variable aléatoire suivant la loi de Poisson de paramètre λ et indépendante de la suite $(U_i)_{i \in \mathbb{N}^*}$ définie dans la partie II..
   Pour $n \in \mathbb{N}^*$, on note $H_n = U_1^2 + U_2^2 + \ldots + U_{n+2K}^2$ et on admet que H_n est une variable aléatoire à densité.

   (a) Ecrire en Python une fonction intitulée `def H(n,lambda):` simulant la variable H_n.

   (b) Soit $k \in \mathbb{N}$.
   - Quelle est la loi de H_n conditionnelle à $(K = k)$?
   - Préciser $E(H_n \mid (K = k))$

   (c) Montrer que H_n a une espérance et la calculer en fonction de n et λ.

   (d) Pour $k \in \mathbb{N}$ et $x < \frac{1}{2}$, calculer $E\left(e^{xH_n} \mid (K = k)\right)$.

   (e) En déduire l'expression de $G_{H_n}(x)$ pour tout $x < \frac{1}{2}$ puis la loi de H_n

Partie IV : Estimations de la variance d'une loi normale

Pour cette partie, m est un réel, σ un réel strictement positif.
On considère une suite $(X_i)_{i\in\mathbb{N}^*}$ de variables aléatoires mutuellement indépendantes et suivant toutes la loi $\mathcal{N}(m,\sigma^2)$. On cherche à estimer σ^2 par intervalle de confiance.

On suppose pour les questions de 1. à 3. que m est connue.

Pour tout $n \geq 2$, on note $T_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - m)^2$.

1. Justifier que $\frac{n}{\sigma^2}T_n$ suit la loi du χ^2 centré à n degrés de liberté.
2. Montrer que T_n est un estimateur sans biais et convergent de σ^2.
3. Soit $\alpha \in]0,1[$
   (a) On considère la fonction F_n définie au II. 6.
   Justifier que $0 < F_n^{-1}\left(\frac{\alpha}{2}\right) < F_n^{-1}\left(1-\frac{\alpha}{2}\right)$ et calculer $P\left(F_n^{-1}\left(\frac{\alpha}{2}\right) \leq \frac{n}{\sigma^2}T_n \leq F_n^{-1}\left(1-\frac{\alpha}{2}\right)\right)$
   (b) En déduire que $\left[\dfrac{nT_n}{F_n^{-1}\left(1-\frac{\alpha}{2}\right)};\dfrac{nT_n}{F_n^{-1}\left(\frac{\alpha}{2}\right)}\right]$ est un intervalle de confiance pour σ^2 au niveau de confiance $1-\alpha$.

Pour la suite de cette partie m est inconnue

4. n est un entier supérieur ou égal à 2 , on munit $\mathcal{M}_{n,1}(\mathbb{R})$ du produit scalaire canonique et on note $M_n = \frac{1}{n}\sum_{i=1}^{n}X_i$ et $V_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - M_n)^2$
   (a) Montrer que M_n est un estimateur sans biais et convergent de m.
   (b) Montrer que $\sum_{i=1}^{n}(X_i - M_n)^2 = \sum_{i=1}^{n}X_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n}X_i\right)^2$
5. Pour $x = (x_1,\ldots,x_n) \in \mathbb{R}^n$, on note $Q_n(x) = \sum_{i=1}^{n}x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n}x_i\right)^2$
   (a) Montrer que la matrice symétrique A de $\mathcal{M}_n(\mathbb{R})$ associée à la forme quadratique Q_n est donnée en Python par `A=np.eye(n)-1/n*np.ones([n,n])`
   (b) Justifier que A est la matrice dans la base canonique de $\mathbb{R}^n$ d'un projecteur orthogonal.
   (c) Préciser le rang de A.
   (d) En déduire qu'il existe une matrice orthogonale P et une matrice diagonale D, telles que $A = P.D.{}^tP$, où $D = Diag(1,...,1,0)$.
6. On note $X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$, et on pose $Y = {}^tP.X$. On note $Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}$.
   Montrer que $V_n = Q_n(X) = Y_1^2 + \cdots + Y_{n-1}^2$.
7. **On admet que les variables aléatoires $Y_1,\ldots,Y_{n-1},Y_n$ sont mutuellement indépendantes et qu'elles suivent toutes la loi $\mathcal{N}(0,\sigma^2)$**
   (a) Quelle est la loi de $\frac{1}{\sigma^2}Q_n(X)$?
   (b) En déduire, pour $\alpha \in]0,1[$, un intervalle de confiance au risque α pour σ^2