INSTITUT NATIONAL DE LA STATISTIQUE ET DES ÉTUDES ÉCONOMIQUES

ÉCOLE NATIONALE DE LA STATISTIQUE ET DE L'ANALYSE DE L'INFORMATION

Concours interne d'attaché statisticien de l'Insee

SESSION 2019

ÉPREUVE DE MATHÉMATIQUES ET STATISTIQUES

Durée 4 heures

Coefficient 3

Sans documents – L'usage de la calculatrice est interdit

Le sujet comprend 6 pages

Le sujet se compose de 4 exercices.

Exercice I .... 5 pts

Soit une variable X, absolument continue, suivant une loi de Rayleigh de paramètre $\sigma > 0$, c'est-à dire que sa densité est donné par :

$$\begin{cases} f(x) = \lambda x e^{-\frac{x^2}{2\sigma^2}} & \text{si } x \geq 0 \\ f(x) = 0 & \text{sinon} \end{cases}$$

où λ est définie pour que f soit une densité.

On admet que $\lim\limits_{x \to +\infty} \int_0^x e^{-\frac{x^2}{2\sigma^2}} dt = \sigma \sqrt{\frac{\pi}{2}}$.

Partie A

1. (a) Calculer $\int_0^x f(t)\,\mathrm{d}t$ pour $x \geqslant 0$.
   En déduire $\lim\limits_{x \to +\infty} \int_0^x f(t)\,\mathrm{d}t$.
   Déterminer enfin la valeur de λ.
   (b) Déterminer la fonction F de répartition de la loi de Rayleigh.
   (c) Montrer que $E(X) = \sigma\sqrt{\frac{\pi}{2}}$.
   (d) À l'aide d'une intégration par partie, calculer $E(X^2)$.
2. Soit U une variable aléatoire de loi uniforme sur $[0, 1]$.
   (a) Calculer la fonction de répartition de U.
   (b) On note $Z = \sigma\sqrt{-2\ln U}$. Calculer la fonction de répartition de Z.
   En déduire que Z suit une loi de Rayleigh de paramètre σ.
   (c) Soit $Y = X^2$. Montrer que Y suit une loi exponentielle dont on déterminera le paramètre.

Partie B

La hauteur maximale en mètres de la crue annuelle d'un fleuve est modélisée par une suite de variables aléatoires (H_i) indépendantes de loi de Rayleigh de paramètre σ.
On note h_i une réalisation de la variable aléatoire H_i.
Pendant dix ans les experts ont observé les hauteurs d'eau suivantes :

Hauteur h_i	6.43	4.38	1.93	5.12	2.46	3.30	5.98	2.56	2.19	7.29

On a calculé les grandeurs suivantes : $\sum_{i=1}^{10} h_i \simeq 41$, $\sum_{i=1}^{10} h_i^2 \simeq 207$ et $\sqrt{\frac{2}{\pi}} \simeq 0,8$.

On note $\hat{H} = \dfrac{1}{n}\sum\limits_{i=1}^{n} H_i$ et $\tilde{S} = \dfrac{1}{2n}\sum\limits_{i=1}^{n} H_i^2$ deux estimateurs.

On rappelle que le biais d'un estimateur $\hat{\theta}$ cherchant à estimer la valeur du paramètre θ est : Biais$(\hat{\theta}) = E(\hat{\theta}) - \theta$. On dit qu'un estimateur est sans biais si son biais est nul.

1. Calculer l'espérance de $\tilde{S}$. Est-ce un estimateur sans biais de σ^2 ?
   En déduire une estimation de la valeur numérique de σ^2.
2. (a) Calculer l'espérance de $\hat{H}$.
   (b) Est-ce un estimateur sans biais de σ ?
   (c) Proposer un estimateur $\tilde{H}$ sans biais de σ.
   (d) Calculer la variance de $\tilde{H}$.
   (e) Donner une estimation de la valeur numérique de σ.
3. Calculer la probabilité que la hauteur d'eau de la crue ne dépasse pas 6m en 100 ans (on ne demande pas de valeur numérique).

Exercice II .... 5 pts

1. Soit f une application linéaire de $\mathbb{R}^3$ dans $\mathbb{R}^3$ dont la matrice associée dans la base canonique est notée M. On suppose que f est diagonalisable. On rappelle que $f^2 = f \circ f$.
   (a) Montrer que f^2 est diagonalisable.
   (b) Déterminer les valeurs propres de f^2 en fonction de celles de f.
2. Soit g l'application linéaire de $\mathbb{R}^3$ dans $\mathbb{R}^3$ dont la matrice dans la base canonique de $\mathbb{R}^3$ est :
$$A = \begin{pmatrix} 2 & 1 & 0 \\ -3 & -2 & 2 \\ -1 & -1 & 1 \end{pmatrix}.$$
   On note I la matrice identité de $\mathcal{M}_3(\mathbb{R})$.
   (a) Déterminer la matrice A^2 puis montrer que $A^4 = I$. En déduire les valeurs propres possibles de la matrice A.
   (b) Analyse du spectre de g
      i. Donner une base (u) de $\ker(g - Id)$.
      ii. Déterminer $\ker(g + Id)$. Donner la dimension de cet espace. La valeur -1 est-elle valeur propre de A ?
      iii. En déduire que g n'est pas diagonalisable.
   (c) Analyse du spectre de g^2
      i. Résoudre $A^2X = -X$ où X est un vecteur de $\mathbb{R}^3$. En déduire une base (v, w) de $\ker(g^2 + Id)$.
      ii. Montrer que (u, v, w) est une base de $\mathbb{R}^3$.
      iii. Écrire la matrice de g^2 dans la base (u, v, w).
3. Soit f une application linéaire de $\mathbb{R}^3$ dans $\mathbb{R}^3$ telle que f^2 soit diagonalisable. A-t-on nécessairement f diagonalisable ?

Exercice III .... 5 pts

On s'intéresse à un bandit manchot (machine à sous) qui fonctionne de la manière suivante : il présente 3 écrans A, B et C. Lorsqu'on lance une partie, chacun des 3 écrans fait apparaître un symbole ($\star$,$\circ$, $\square$), au hasard. Les résultats des 3 écrans sont indépendants. La partie est gagnée si les 3 symboles obtenus sont les mêmes. On donne deux exemples de parties ci-dessous :

écran	A	B	C
	$\star$	$\circ$	$\circ$

Une partie perdue

écran	A	B	C
	$\square$	$\square$	$\square$

Une partie gagnée

On définit les évènements G : "la partie est gagnée" et E : "la partie est perdue".

Partie A

1. Calculer la probabilité des événements G et E.
2. Gain du casino
   On suppose qu'un joueur doit payer 3 euros à chaque fois qu'il veut jouer. En cas de partie gagnée il remporte 15 euros.
   On suppose que le bandit manchot est utilisé 5000 fois dans une journée.
   (a) On note M_i la variable aléatoire qui représente le gain du casino à la i-ième partie, pour i entre 1 et 5000. Donner la loi des variables aléatoires M_i ainsi que leur espérance.
   (b) Quel est le gain moyen du casino sur une journée ?
3. Première partie gagnée
   Un joueur décide de miser jusqu'à gagner une partie. On note X la variable aléatoire égale au nombre de parties jouées pour gagner la première fois (y compris la première partie gagnée).
   (a) Donner la loi de la variable aléatoire X.
   (b) Pour tout $i \in \mathbb{N}$, montrer que la probabilité que le joueur joue au plus i parties avant de gagner pour la première fois (toujours y compris la partie gagnée) est $1 - \left(\frac{8}{9}\right)^i$.
   (c) Donner l'espérance et la variance de X (on ne demande pas de démonstration).

Partie B

On souhaite dorénavant prendre en compte dans le calcul des probabilités la possibilité que le bandit dysfonctionne. Lorsque c'est le cas, les écrans A et C affichent toujours le symbole $\star$, l'écran B continuant à afficher des résultats au hasard. On note D : "la machine dysfonctionne", et on pose $p = P(D)$. Le cas où le bandit manchot fonctionne normalement est donc décrit dans la partie A.

1. Calculer la probabilité conditionnelle des événements E et G sachant D : $P_D(E)$ et $P_D(G)$.
2. En déduire que la probabilité de E est $\frac{8-2p}{9}$.
3. Soit R la variable aléatoire égale au gain du casino lors d'une partie jouée. Déterminer la valeur maximale de p pour que l'espérance de gain soit positive.
4. Un joueur joue une partie, on suppose qu'il gagne. Quelle est la probabilité, en fonction de p, que le bandit manchot ait dysfonctionné lors de la partie ?

Exercice IV ....5 pts

Partie A

On souhaite dans cette partie montrer la convexité de la fonction

$$b : \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto x^2 + y^2,$$

c'est-à-dire montrer que

$$\forall (x, y), (x', y') \in \mathbb{R}^2, \forall t \in [0, 1], b\left(t(x, y) + (1 - t)(x', y')\right) \leqslant t.b(x, y) + (1 - t).b(x', y').$$

On fixe deux points $(x, y), (x', y')$ de $\mathbb{R}^2$, distincts.

1. Soit la fonction $\varphi : \mathbb{R} \to \mathbb{R}, t \mapsto b(t(x, y)+(1-t)(x', y'))-[t.b(x, y) + (1 - t).b(x', y')]$. Calculer une expression de $\varphi(t)$ pour $t \in \mathbb{R}$.
2. Calculer φ' et φ''.
3. Montrer que φ' est strictement croissante sur $\mathbb{R}$.
4. (a) Montrer qu'il existe un unique point $t_0 \in]0, 1[$ tel que $\varphi'(t_0) = 0$.
   (b) En déduire le tableau de signe de φ' et le tableau de variations de φ sur $[0, 1]$.
5. Conclure.

Partie B

Un boulanger souhaite optimiser sa production de viennoiseries. Pour simplifier on considère la situation suivante : on suppose que le boulanger ne produit que des pains au chocolat standard (en nombre x, en centaines) et des pains au chocolat maxi (en nombre y, en centaines également). Il souhaite choisir les nombres de pains x et y à produire pour optimiser sa recette $\frac{1}{3}x + y$ (un pain au chocolat standard est vendu 3 fois moins cher qu'un pain au chocolat maxi). On suppose que le seul ingrédient utilisé est de la farine, et qu'un pain standard nécessite 50g de farine, un pain maxi 100g. Il y a 10kg de farine disponible. Enfin une commande de 100 viennoisieries (pains standards ou maxis) a été passée et doit être assurée.

On traduit le problème à l'aide des fonctions suivantes : soient $f_1, f_2, f_3 : \mathbb{R}^2 \to \mathbb{R}$ données par $f_1(x, y) = y,\ f_2(x, y) = x + y,\ f_3(x, y) = -\frac{1}{2}x - y$.
On pose

$$\begin{aligned}
\mathcal{P}_1 &= (f_1)^{-1}([0, +\infty[) = \{(x, y) \in \mathbb{R}^2; f_1(x, y) \geqslant 0\} \\
\mathcal{P}_2 &= (f_2)^{-1}([1, +\infty[) = \{(x, y) \in \mathbb{R}^2; f_2(x, y) \geqslant 1\} \\
\mathcal{P}_3 &= (f_3)^{-1}([-1, +\infty[) = \{(x, y) \in \mathbb{R}^2; f_3(x, y) \geqslant -1\}
\end{aligned}$$

et $\mathcal{P} = \mathcal{P}_1 \cap \mathcal{P}_2 \cap \mathcal{P}_3$. On note $A_1 = (0, 1),\ A_2 = (2, 0),\ A_3 = (1, 0)$.

L'objectif principal de cette partie est d'établir que tout point M de $\mathcal{P}$ s'écrit comme barycentre des points A_1, A_2, A_3, c'est à dire que

$$\boxed{\forall M \in \mathcal{P},\ \text{il existe } \alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}_+ \text{ tels que } \alpha_1 + \alpha_2 + \alpha_3 = 1 \text{ et } M = \alpha_1 A_1 + \alpha_2 A_2 + \alpha_3 A_3.}$$

On pourra si besoin admettre ce résultat dans la partie III.

1. Expliquer en quoi les contraintes $\{f_2 \geqslant 1\}$ et $\{f_3 \geqslant -1\}$ modélisent certains impératifs du boulanger.

2. (a) Montrer que $\mathcal{P}_1 = \mathbb{R} \times \mathbb{R}_+$, puis dessiner $\mathcal{P}_1$ dans un repère, en mettant les x en abscisses et les y en ordonnées.
   (b) Dans le même repère, dessiner $\mathcal{P}_2$, $\mathcal{P}_3$, puis $\mathcal{P}$.
3. (a) Soient $t \in [0,1]$, $M = (x,y) \in \mathcal{P}_1, M' = (x',y') \in \mathcal{P}_1$.
   Montrer que $f_1(tM + (1-t)M') \geqslant 0$.
   (b) En déduire que $\mathcal{P}_1$ est un sous-ensemble convexe de $\mathbb{R}^2$, c'est à dire que
   $$\forall t \in [0,1], \forall M, M' \in \mathcal{P}_1, \text{ on a } tM + (1-t)M' \in \mathcal{P}_1.$$
   (c) Que peut-on dire de $\mathcal{P}_2$, $\mathcal{P}_3$?
   (d) Montrer que $\mathcal{P}$ est un sous-ensemble convexe de $\mathbb{R}^2$.
4. Dans cette question on souhaite montrer que tout point M de $\mathcal{P}$ s'écrit comme barycentre des points A_1, A_2, A_3, i.e. que pour tout point M de $\mathcal{P}$ il existe des réels positifs $\alpha_1, \alpha_2, \alpha_3$ vérifiant $\alpha_1 + \alpha_2 + \alpha_3 = 1$ et $M = \alpha_1 A_1 + \alpha_2 A_2 + \alpha_3 A_3$.
   (a) Vérifier que A_1 est bien un barycentre des points A_1, A_2, A_3.
   (b) On suppose dans cette question que $M \in P \backslash \{A_1\}$.
      i. Montrer qu'il existe un unique $\lambda \in \mathbb{R}$ tel que $f_3(M + \lambda(1,0)) = -1$, et que $\lambda \geqslant 0$.
      Placer le point $M + \lambda(1,0)$ sur le dessin et donner une interprétation en terme de barycentre avec les points A_1, A_2, A_3.
      ii. Montrer qu'il existe un unique $\mu \in \mathbb{R}$ tel que $f_2(M - \mu(1,0)) = 1$, et que $\mu \geqslant 0$.
      iii. Montrer qu'il existe $t \in [0,1]$ tel que
      $$M = t(M - \mu(1,0)) + (1-t)(M + \lambda(1,0)).$$
      Il est conseillé de faire un dessin !
      iv. Montrer que si M' et M'' sont des points s'écrivant comme barycentres de A_1, A_2, A_3, et si $s \in [0,1]$ alors $sM' + (1-s)M''$ s'écrit aussi comme barycentre de A_1, A_2, A_3.
      v. Utiliser les questions précédentes pour conclure que M s'écrit comme barycentre de A_1, A_2, A_3.

Partie C

Dans cette partie on souhaite optimiser la recette du boulanger, donc maximiser la fonction $r(x,y) = y + \frac{1}{3}x$ sur $\mathcal{P}$. Dans un deuxième temps on cherchera à optimiser le bénéfice donné par la fonction $b(x,y) = x^2 + y^2$ de la partie A.

1. (a) Soit $g : \mathbb{R}^2 \to \mathbb{R}$ linéaire. Montrer que pour tout $M \in \mathcal{P}$ on a :
   $$g(M) \leqslant \max(g(A_1), g(A_2), g(A_3))$$
   On utilisera le fait que M s'écrit comme barycentre des points A_1, A_2, A_3.
   (b) Calculer le maximum de r sur $\mathcal{P}$, et dire en quel(s) point(s) de $\mathcal{P}$ il est atteint.
2. (a) Soit $h : \mathbb{R}^2 \to \mathbb{R}$ convexe.
      i. Montrer que pour tout $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}_+$ tels que $\alpha_1 + \alpha_2 + \alpha_3 = 1$ on a
      $$h(\alpha_1 A_1 + \alpha_2 A_2 + \alpha_3 A_3) \leqslant \alpha_1 h(A_1) + \alpha_2 h(A_2) + \alpha_3 h(A_3).$$
      ii. Établir que h admet un maximum sur $\mathcal{P}$ et qu'il est atteint en A_1, A_2 ou A_3.
   (b) Calculer le maximum de b sur $\mathcal{P}$, et dire en quel(s) point(s) de $\mathcal{P}$ il est atteint.