MATHEMATIQUES
Variables alétoires à densité

Le sujet se compose d'un ***exercice*** *et de* ***deux problèmes****. On prendra soin de lire l'ensemble des énoncés avant de commencer à composer.*
Il sera tenu compte de la présentation et en particulier de $\boxed{\text{l'encadrement}}$ *des résultats.*
*L'usage de la calculatrice n'****est pas*** *autorisé au cours de l'épreuve.*

Exercice (oral agro-véto 2022) :

Dans tout l'exercice, les variables aléatoires sont définies sur le même espace probabilisé $(\Omega, \tau, \mathbb{P})$
Si X *est une variable aléatoire, on notera* $\mathbb{E}(X)$ *son espérance et* $\mathbb{V}(X)$ *sa variance.*
Soient a et b deux réels strictement positifs et f la fonction définie sur $\mathbb{R}$ par :

$$f(x) = \begin{cases} 0 & \text{si } x < b \\ a\dfrac{b^a}{x^{a+1}} & \text{si } x \geq b \end{cases}$$

① a) Soit $\alpha \in \mathbb{R}$. Montrer que $\displaystyle\int_b^{+\infty} \frac{dt}{t^\alpha}$ converge si et seulement si $\alpha > 1$ et que :

$$\forall \alpha > 1, \int_b^{+\infty} \frac{dt}{t^\alpha} = \frac{1}{(\alpha-1)b^{\alpha-1}}$$

b) Montrer que f est une densité de probabilité.
On dit qu'une variable aléatoire suit la loi de Pareto de paramètres a et b si elle admet f pour densité.

② Soit X une variable aléatoire suivant la loi de Pareto de paramètres a et b.
Déterminer la fonction de répartition de X.

③ a) Soit U une variable aléatoire suivant la loi uniforme sur $]0, 1[$.
Montrer que $bU^{-\frac{1}{a}}$ suit la loi de Pareto de paramètres a et b.

b) Écrire une fonction Python `pareto` prenant en argument deux réels strictement positifs a et b et simulant X.

c) Comment pourriez-vous estimer l'espérance de X à l'aide de Python ?

④ a) Montrer que X admet une espérance si, et seulement si, $a > 1$ et que, le cas échéant :

$$\mathbb{E}(X) = \frac{ab}{a-1}$$

b) Montrer que X admet une variance si, et seulement si, $a > 2$ et que, le cas échéant :

$$\mathbb{V}(X) = \frac{ab^2}{(a-1)^2(a-2)}$$

⑤ On suppose désormais que $b = 1$ et pour tout $n \in \mathbb{N}^*$, on pose $W_n = \ln(X_n)$.
Montrer que W_n suit une loi exponentielle dont on précisera le paramètre. En déduire l'espérance et la variance de W_n.

Problème 1 (sujet Agro-véto - épreuve B 2024) :

I. Résultats préliminaires

1. Soit $p \in \mathbb{N}$.

   a) Montrer que $\forall j \in [\![0, p]\!], \dbinom{p}{j} \dfrac{1}{1+j} = \dbinom{p+1}{j+1} \dfrac{1}{p+1}$.

   b) Montrer que $\forall j \in [\![0, p-1]\!] \dbinom{p+1}{j+1} = \dbinom{p}{j+1} + \dbinom{p}{j}$.

2. Soit X une variable aléatoire suivant une loi exponentielle de paramètre $\lambda > 0$. Donner, sans justification, une fonction de densité, la fonction de répartition, l'espérance et la variance de X en fonction de λ.

3. a) Démontrer que, pour tout entier $i > 1$, on a :

$$\int_i^{i+1} \frac{dt}{t} \leqslant \frac{1}{i} \leqslant \int_{i-1}^{i} \frac{dt}{t}.$$

   ✐ *On pourra faire un dessin avec des surfaces qui donnent du sens à cette double inégalité.*

   b) Justifier l'équivalent : $\ln(n+1) \underset{n\to+\infty}{\sim} \ln(n)$.

   c) En déduire un équivalent de $\displaystyle\sum_{k=1}^{n} \frac{1}{k}$ lorsque n tend vers $+\infty$.

II. Quelques résultats autour de la loi exponentielle

On se donne un entier naturel $n > 0$, un réel strictement positif λ et une famille de n variables aléatoires notées $X_1, \ldots, X_n$, indépendantes et identiquement distribuées selon la loi exponentielle de paramètre λ.
On définit $X_{(1)} = \min(X_1, \ldots, X_n)$ et $X_{(n)} = \max(X_1, \ldots, X_n)$.

4. Calculer $\mathbb{P}(X_{(1)} > x)$ pour tout réel x strictement positif et en déduire que la variable $X_{(1)}$ suit une loi exponentielle dont on précisera le paramètre.

5. Prouver que $X_{(n)}$ est une variable aléatoire de densité :

$$x \mapsto n\lambda e^{-\lambda x}(1 - e^{-\lambda x})^{n-1} 1_{\mathbb{R}}(x).$$

6. En déduire que :

$$\mathbb{E}(X_{(n)}) = n\lambda \sum_{j=0}^{n-1} \binom{n-1}{j} (-1)^j \int_0^{+\infty} x e^{-(1+j)\lambda x} dx$$

   puis, en utilisant les résultats de la partie I, que :

$$\mathbb{E}(X_{(n)}) = \frac{1}{\lambda} \sum_{j=0}^{n-1} \binom{n}{j+1} \frac{(-1)^j}{j+1}.$$

7. Dans cette question on veut prouver que :

$$\forall p \in \mathbb{N}^*, \qquad \sum_{j=0}^{p-1} \binom{p}{j+1} \frac{(-1)^j}{j+1} = \sum_{k=1}^{p} \frac{1}{k}.$$

   a) Soit $p \in \mathbb{N}$. Déterminer la valeur de $\displaystyle\sum_{k=0}^{p+1} \binom{p+1}{k} (-1)^k$.

   b) En utilisant les résultats de la partie I, prouver le résultat souhaité par récurrence.

8. En déduire un équivalent de $\mathbb{E}(X_{(n)})$.

On définit la suite $u = \left(\lambda \mathbb{E}(X_{(n)}) - \ln(n)\right)_{n \in \mathbb{N}^*}$.

9. Déterminer un équivalent de $u_{n+1} - u_n$ et en déduire que la série $\sum (u_{n+1} - u_n)$ est convergente.

10. En déduire que la suite u est convergente. On ne cherchera pas à calculer sa limite.

III. Loi de Gumbel

On reprend les notations de la partie précédente et on rappelle que la variable X_1 suit une loi exponentielle de paramètre λ. On définit la fonction f par :

$$f : x \mapsto \lambda \exp(-x) \exp\left[-\lambda \exp(-x)\right].$$

11. Justifier que la fonction f est bien une densité de probabilité.
    On appellera cette loi *la loi de Gumbel de paramètre* λ.
    Indication : on pourra utiliser le changement de variable $y = \exp(-x)$.

12. Montrer que la variable aléatoire $Y = -\ln X_1$ suit la loi de Gumbel.

13. Soit $x \in \mathbb{R}$. Justifier la limite suivante :

$$\lim_{n \to +\infty} \left(1 + \frac{x}{n}\right)^n = \exp(x).$$

14. Montrer que si $\lambda = 1$, alors la suite $(X_{(n)} - \ln(n))_{n \in \mathbb{N}^*}$ converge en loi vers Y.

Problème 2 (sujet Agro-véto - épreuve B 2015) :

Dans ce problème, $\mathbb{N}$, $\mathbb{N}^*$, $\mathbb{R}$ et $\mathbb{R}_+^*$ désignent respectivement les ensembles des entiers naturels, des entiers naturels non nuls, des nombres réels et des nombres réels strictement positifs.

Pour $\alpha \in \mathbb{R}$, et $r \in \mathbb{R}_+^*$, on définit la fonction $\ell_{\alpha,r}$ sur $\mathbb{R}$ par :

$$\ell_{\alpha,r}(t) = \begin{cases} \alpha \dfrac{t^{\alpha-1}}{r^\alpha} & \text{si } t \in]0, r[\\ 0 & \text{sinon} \end{cases}$$

① Montrer que $\ell_{\alpha,r}$ est une densité de probabilité sur $\mathbb{R}$ si, et seulement si, $\alpha > 0$.

On dit qu'une variable aléatoire réelle X suit la loi $\mathcal{L}$ de paramètres α et r si X admet $\ell_{\alpha,r}$ pour densité. On notera : $X \hookrightarrow \mathcal{L}(\alpha, r)$

✐ *Dans toute la suite de ce problème, on note X une variable aléatoire de loi $\mathcal{L}(\alpha, r)$.*

② On s'intéresse tout d'abord aux propriétés élémentaires de cette loi.

a) Déterminer la fonction de répartition F de X.

b) Que vaut $\mathbb{P}(X \leq 0)$?

c) Reconnaître la loi de la variable aléatoire $W = -\ln\left(\dfrac{X}{r}\right)$. Préciser son espérance et sa variance.

③ Soit $\lambda, \mu \in \mathbb{R}_+^*$. Soient U_1 et U_2 deux variables aléatoires réelles indépendantes qui suivent des lois exponentielles de paramètre λ et μ respectivement.

a) Donner une densité de $Y = -U_2$.

b) On rappelle que si V_1 et V_2 sont deux variables aléatoires indépendantes de densités respectives v_1 et v_2, alors $Z = V_1 + V_2$ est encore une variable aléatoire réelle à densité, dont une densité est données par $z \longmapsto \displaystyle\int_{-\infty}^{\infty} v_1(u)v_2(t-u)du$.
Montrer qu'une densité de $U_1 - U_2$ est donnée par :

$$g_{\lambda,\mu} : t \longmapsto \begin{cases} \dfrac{\lambda\mu}{\lambda+\mu} e^{-\lambda t} & \text{si } t > 0 \\ \dfrac{\alpha\mu}{\lambda+\mu} e^{\mu t} & \text{si } t \leq 0 \end{cases}$$

c) Soient X et Y deux variables aléatoires réelles indépendantes, respectivement de lois $\mathcal{L}(\alpha, r)$ et $\mathcal{L}(\beta, s)$ avec $(\alpha, \beta, r, s) \in (\mathbb{R}_+^*)^4$. A l'aide de la question qui précède et de la question 2.c), donner la loi de $Z = -\ln\left(\dfrac{sX}{rY}\right)$ et en déduire que :

$$\mathbb{P}(X < Y) = \begin{cases} \dfrac{\beta}{\alpha+\beta}\left(\dfrac{s}{r}\right)^\alpha & \text{si } s < r \\ 1 - \dfrac{\alpha}{\alpha+\beta}\left(\dfrac{r}{s}\right)^\beta & \text{si } r \leq s \end{cases}$$

④ On applique maintenant le résultat de la question précédente à un modèle de test de dépistage d'une maladie canine H au sein d'une population $\mathcal{P}$. La présence de cette maladie chez un individu de $\mathcal{P}$ se note par la modification de la loi de concentration de deux bactéries A et B présentes dans l'estomac. Plus précisément, ces concentrations respectives X et Y, en UFCml^{-1} (Unité Formant Colonie par millilitre), sont deux variables aléatoires réelles indépendantes de lois respectives $\mathcal{L}(\lambda, r)$ et $\mathcal{L}(\beta, s)$, avec $(\alpha, \beta, r, s) \in (\mathbb{R}_+^*)^4$.

On a statistiquement les valeurs suivantes :

— Pour les sujets *non atteints* de la maladie H :
$$\alpha = 2\,;\ \beta = 3\,;\ r = 100 \text{ UFC } ml^{-1}\,;\ s = 50 \text{ UFC } ml^{-1}$$
— Pour les sujets *atteints* de la maladie H :
$$\alpha = 4\,;\ \beta = 2\,;\ r = 400 \text{ UFC } ml^{-1}\,;\ s = 800 \text{ UFC } ml^{-1}$$

Le test T consiste à effectuer un prélèvement sanguin d'un individu $\mathcal{C}$ et $\mathcal{P}$. Une fois ce prélèvement mis en culture dans des conditions adéquates, les deux bactéries A et B entrent en concurrence. Au bout de quelques heures, seule celle dont la concentration était la plus forte subsiste, l'autre ayant totalement disparu. Cette procédure permet donc de savoir lequel des événements $(X < Y)$ ou $(X > Y)$ est réalisé pour $\mathcal{C}$.
On note $R = (X < Y)$. Le test est positif si R est réalisé, négatif sinon.
On note M l'événement « le sujet $\mathcal{C} \in \mathcal{P}$ est atteint de la maladie H ». L'événement contraire d'un événement E sera systématiquement noté $\overline{E}$ dans la suite.

a) Donner les probabilités conditionnelles $\mathbb{P}(R|M) = \mathbb{P}_M(R)$ et $\mathbb{P}(R|\overline{M}) = \mathbb{P}_{\overline{M}}(R)$.

b) Un sondage a permis d'estimer $\mathbb{P}(M) = \dfrac{1}{40}$. Donner la probabilité qu'un sujet testé soit atteint de la maladie H sachant que son test T est positif. Qu'en pensez-vous ?

- FIN -