Code sujet : 283

BCE
BANQUE COMMUNE D'ÉPREUVES

Conception : ESCP BS – HEC Paris

MATHÉMATIQUES 2 APPROFONDIES

FILIÈRE ÉCONOMIQUE ET COMMERCIALE

VOIE GÉNÉRALE

Jeudi 25 avril 2024, de 14 h. à 18 h.

La présentation, la lisibilité, l'orthographe, la qualité de la rédaction, la clarté et la précision des raisonnements entreront pour une part importante dans l'appréciation des copies.
Les candidats sont invités à ***encadrer*** *dans la mesure du possible les résultats de leurs calculs.*
Aucun document n'est autorisé. ***L'utilisation de toute calculatrice et de tout matériel électronique est interdite.*** *Seule l'utilisation d'une règle graduée est autorisée.*
Si au cours de l'épreuve, un candidat repère ce qui lui semble être une erreur d'énoncé, il la signalera sur sa copie et poursuivra sa composition en expliquant les raisons des initiatives qu'il sera amené à prendre.

Toutes les variables aléatoires de ce problème sont supposées définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.
On rappelle que pour tout x réel, $\lfloor x \rfloor$ désigne la partie entière de x, c'est à dire le nombre entier relatif $\lfloor x \rfloor$ vérifiant $\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$.
Si A est un ensemble fini non vide, on note $\mathrm{Card}(A)$ le nombre d'éléments de A. Si $A = \emptyset$, on convient que $\mathrm{Card}(A) = 0$.
Pour les programmes en Python, on suppose importées :

- la bibliothèque `numpy` sous l'alias `np`,
- la bibliothèque `numpy.random` sous l'alias `rd`,
- la bibliothèque `matplotlib.pyplot` sous l'alias `plt`.

Le mot **FIN** marque la fin de l'énoncé.

Première partie

Soit f une densité de probabilité sur $\mathbb{R}$, X une variable aléatoire réelle ayant f pour densité et F sa fonction de répartition.

Soit $(X_i)_{i\in\mathbb{N}^*}$ un échantillon de X, c'est-à-dire une suite de variables aléatoires réelles mutuellement indépendantes, de même loi que X.

Pour tout entier $n \geq 1$, on ordonne, pour chaque $\omega \in \Omega$, le n-uplet $\big(X_1(\omega), X_2(\omega), \ldots, X_n(\omega)\big)$ en :

$$X_{(n,1)}(\omega) \leq X_{(n,2)}(\omega) \leq \cdots \leq X_{(n,n)}(\omega),$$

Tournez la page S.V.P.

et **on admet** que les applications $X_{(n,k)}$ ainsi définies sont des variables aléatoires réelles sur $(\Omega, \mathcal{A}, \mathbb{P})$.

Pour tous k,n entiers tels que $1 \leq k \leq n$, la variable aléatoire $X_{(n,k)}$ donne la k-ième valeur obtenue en classant dans l'ordre croissant les valeurs prises par les n premières X_i.

On remarque en particulier que $X_{(n,1)} = \min\limits_{1\leq i\leq n}(X_i)$ et $X_{(n,n)} = \max\limits_{1\leq i\leq n}(X_i)$.

1. Soit $n \in \mathbb{N}^*$.
   (a) Exprimer, à l'aide de F, la fonction de répartition $F_{(n,n)}$ de la variable aléatoire $X_{(n,n)}$. En déduire que $X_{(n,n)}$ admet comme densité la fonction $f_{(n,n)}$ définie par :
   $$\forall x \in \mathbb{R}, \quad f_{(n,n)}(x) = nf(x)F(x)^{n-1}.$$
   (b) Par une méthode similaire, montrer que $X_{(n,1)}$ admet comme densité $f_{(n,1)}$ définie par :
   $$\forall x \in \mathbb{R}, \quad f_{(n,1)}(x) = nf(x)(1-F(x))^{n-1}.$$

2. Soit $x \in \mathbb{R}$ et soient n et k deux entiers tels que $1 \leq k \leq n$.
   Pour tout $i \in \mathbb{N}^*$, on note $Y_{i,x}$ la variable aléatoire indicatrice de l'événement $[X_i \leq x]$. C'est à dire
   $$Y_{i,x} = \begin{cases} 1 & \text{si } X_i \leqslant x \\ 0 & \text{sinon} \end{cases}$$
   (a) Montrer que $Y_{i,x}$ suit une loi de Bernoulli dont on précisera le paramètre.
   (b) Montrer l'égalité entre les événements :
   $$[X_{(n,k)} \leq x] = \left[\sum_{i=1}^{n} Y_{i,x} \geq k\right].$$
   (c) En déduire que la fonction de répartition $F_{(n,k)}$ de la variable aléatoire $X_{(n,k)}$ est donnée par :
   $$\forall x \in \mathbb{R}, \quad F_{(n,k)}(x) = \sum_{\ell=k}^{n} \binom{n}{\ell} F(x)^{\ell}(1-F(x))^{n-\ell}.$$

3. Pour tout triplet d'entiers (n,k,ℓ) tel que $0 \leq k \leq n$ et $0 \leq \ell \leq n$, on introduit les fonctions polynomiales $b_{n,\ell}$ et $a_{n,k}$ définies respectivement par, pour tout $z \in \mathbb{R}$:
   $$b_{n,\ell}(z) = \binom{n}{\ell} z^{\ell}(1-z)^{n-\ell} \qquad \text{et} \qquad a_{n,k}(z) = \sum_{\ell=k}^{n} b_{n,\ell}(z).$$
   On convient également que pour tout $n \in \mathbb{N}^*$, la fonction $b_{n-1,n}$ est identiquement nulle.
   (a) Soient $n \in \mathbb{N}^*$ et $\ell \in [\![1,n]\!]$. Montrer que pour tout $z \in \mathbb{R}$:
   $$b'_{n,\ell}(z) = n\big(b_{n-1,\ell-1}(z) - b_{n-1,\ell}(z)\big).$$
   (b) Soient $n \in \mathbb{N}^*$ et $k \in [\![1,n]\!]$. Déduire de la question précédente une expression de $a'_{n,k}$ en fonction de l'une des fonctions $b_{m,\ell}$ avec m et ℓ à déterminer en fonction de n et k.
   (c) Soient $n \in \mathbb{N}^*$ et $k \in [\![1,n]\!]$. Montrer que la variable aléatoire $X_{(n,k)}$ admet comme densité $f_{(n,k)}$ définie par :
   $$\forall x \in \mathbb{R}, \quad f_{(n,k)}(x) = n\binom{n-1}{k-1} f(x)F(x)^{k-1}(1-F(x))^{n-k}.$$

4. Soit x un réel tel que $0 < F(x) < 1$.

(a) Soient $\alpha \in]F(x), 1[$ et $\varepsilon > 0$ tel que $\alpha - F(x) > \varepsilon$.
Montrer que pour tout entier n suffisamment grand, on a l'inclusion d'événements :

$$\left[\sum_{i=1}^{n} Y_{i,x} \geq \lfloor \alpha n \rfloor\right] \subset \left[\left|\frac{1}{n}\sum_{i=1}^{n} Y_{i,x} - F(x)\right| \geq \varepsilon\right],$$

et en déduire que :

$$\lim_{n\to\infty} F_{(n,\lfloor \alpha n \rfloor)}(x) = 0.$$

(b) Soient $\alpha \in]0, F(x)[$ et $\varepsilon > 0$ tel que $F(x) - \alpha > \varepsilon$.
Montrer que pour tout $n \in \mathbb{N}^*$:

$$\left[\sum_{i=1}^{n} Y_{i,x} \geq \lfloor \alpha n \rfloor\right] \supset \left[\left|\frac{1}{n}\sum_{i=1}^{n} Y_{i,x} - F(x)\right| < \varepsilon\right],$$

et en déduire que :

$$\lim_{n\to\infty} F_{(n,\lfloor \alpha n \rfloor)}(x) = 1.$$

5. On suppose pour cette question qu'il existe un intervalle ouvert $I =]a, b[$, avec éventuellement $a = -\infty$ ou $b = +\infty$, tel que f est continue sur I, $f(x) > 0$ pour tout $x \in I$ et $f(x) = 0$ pour tout $x \in \mathbb{R} \setminus I$.

(a) Montrer que la fonction de répartition associée F établit une bijection entre les intervalles I et $]0, 1[$.

(b) On note F^{-1} la bijection réciproque de la fonction F restreint à I et on fixe $\alpha \in]0, 1[$.
Montrer que la suite de variables aléatoires $(X_{(n,\lfloor \alpha n \rfloor)})$ converge en loi vers une variable aléatoire certaine égale à $F^{-1}(\alpha)$.

(c) En déduire que la suite de variables aléatoires $(X_{(n,\lfloor \alpha n \rfloor)})$ converge en probabilité vers une variable aléatoire certaine égale à de $F^{-1}(\alpha)$.

6. Dans cette question, on souhaite écrire une fonction Python d'en-tête `def tri(T):` permettant de trier dans l'ordre croissant les valeurs d'un tableau unidimensionnel `T` (de type `numpy.array`) sans utiliser de fonction de tri prédéfinie dans Python.
Pour cela, on utilise un algorithme de tri appelé *tri par insertion* et défini de la façon suivante :

— On agit sur les coefficients du tableau `T`, que l'on suppose numérotés de 0 à $s - 1$.

— Pour chaque k compris entre 1 et $s - 1$, avant l'étape k les coefficients numérotés 0 à $k - 1$ de `T` ont déjà été classés dans l'ordre croissant, les coefficients suivants étant inchangés. L'étape k consiste alors à insérer la valeur x du coefficient numéroté k de `T` en bonne position parmi les $k + 1$ premiers coefficients de `T` afin qu'à l'issue de l'étape k les coefficients numérotés 0 à k de `T` soient classés dans l'ordre croissant, les coefficients suivants étant inchangés.

Par exemple en partant du tableau `T` dont le contenu est `[38 28 35 24 31 15]`, les étapes de l'algorithme de tri par insertion donnent successivement :

```
[38 28 35 24 31 15]
[28 38 35 24 31 15]
[28 35 38 24 31 15]
[24 28 35 38 31 15]
[24 28 31 35 38 15]
[15 24 28 31 35 38]
```


Tournez la page S.V.P.

Le tri doit être ≪en place ≫, c'est à dire que le tableau se retrouve trié en fin de fonction au lieu de retourner un autre tableau, version triée du premier.

Écrire la fonction `tri` demandée en mettant en œuvre l'algorithme de tri par insertion décrit ci-dessus.

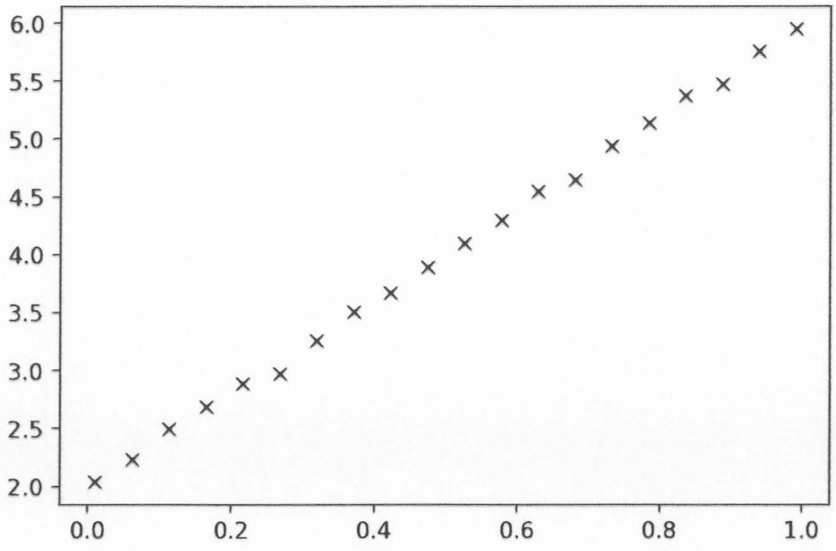
7. On écrit le programme Python suivant à la suite de la fonction `tri` :

```
def ech(n):
    *****

def X(n,k):
    res = ech(n)
    *****

n = 2000 ; na = 20
A = np.linspace(0.01, 0.99, na)
B = np.zeros(na)
for k in range(na):
    a = A[k]
    B[k] = X(n, int(np.floor(a*n)))
plt.plot(A, B, "xk")
plt.show()
```

La fonction `ech`, que l'on suppose déjà écrite pour les trois questions suivantes (a, b et c), renvoie, sous la forme d'un tableau unidimensionnel de longueur n, une réalisation du n-uplet de variables aléatoires $(X_1, X_2, \ldots, X_n)$.

(a) Compléter la fonction `X` afin qu'elle renvoie une réalisation de la variable aléatoire $X_{(n,k)}$ lorsqu'on lui fournit en entrée deux entiers de type Python `int` n et k tels que $1 \leq k \leq n$.
On pourra utiliser la fonction `tri`.

(b) En exécutant le programme, on obtient l'affichage suivant :



À quoi correspondent les points affichés sur cette représentation graphique ?

(c) Utiliser la représentation graphique pour déterminer la loi commune des X_i simulées par la fonction `ech`.

(d) Compléter le corps de la fonction `ech` par une ou plusieurs lignes de code pour que le script produise une représentation graphique de valeurs simulées proche de la représentation graphique ci-dessus.

Deuxième partie

On s'intéresse dans cette partie au cas où $(X_i)_{i\in\mathbb{N}^*}$ est une famille de variables aléatoires mutuellement indépendantes suivant toutes la même loi exponentielle $\mathcal{E}(\lambda)$, avec $\lambda \in]0,+\infty[$.

8. Soit $n \in \mathbb{N}^*$ fixé.

   (a) On note $(T_1,\dots,T_n)$ une famille de variables aléatoires mutuellement indépendantes telles que :
   $$\text{pour tout } i \in [\![1,n]\!],\ T_i \text{ suit la loi exponentielle } \mathcal{E}\big((n-i+1)\lambda\big).$$
   Pour tout $k \in [\![1,n]\!]$, on note de plus $S_k = T_1+\cdots+T_k$.
   Montrer par récurrence sur $k \in [\![1,n]\!]$ que la variable aléatoire S_k suit la loi de densité $g_{n,k}$ définie par :
   $$\forall x \in \mathbb{R},\quad g_{n,k}(x) = \begin{cases} n\dbinom{n-1}{k-1}\lambda e^{-\lambda(n-k+1)x}(1-e^{-\lambda x})^{k-1} & \text{si } x>0,\\[2ex] 0 & \text{si } x \leqslant 0.\end{cases}$$

   (b) Soit $k \in [\![1,n]\!]$.
   Vérifier que S_k suit la même loi que $X_{(n,k)}$ lorsque les X_i suivent toutes la loi $\mathcal{E}(\lambda)$.

9. Soit (n,k) un couple d'entiers tel que $1 \le k \le n$.
   Montrer que $X_{(n,k)}$ admet une espérance et une variance données par :
   $$\mathbb{E}(X_{(n,k)}) = \frac{1}{\lambda}\sum_{i=1}^{k}\frac{1}{n-i+1} \qquad \text{et} \qquad \mathbb{V}(X_{(n,k)}) = \frac{1}{\lambda^2}\sum_{i=1}^{k}\frac{1}{(n-i+1)^2}.$$

10. (a) Montrer que pour tout entier $j \ge 1$ et tout $x \in [j,j+1]$, on a :
    $$0 \le \frac{1}{j} - \frac{1}{x} \le \frac{2}{x^2}.$$

    (b) En déduire l'encadrement, pour $(n,k)\in\mathbb{N}^2$ tel que $1 \le k \le n$:
    $$\frac{1}{\lambda}\ln\left(\frac{n+1}{n-k+1}\right) \le \mathbb{E}(X_{(n,k)}) \le \frac{1}{\lambda}\ln\left(\frac{n+1}{n-k+1}\right) + \frac{2}{\lambda}\left(\frac{1}{n-k+1} - \frac{1}{n+1}\right).$$

    (c) Montrer que pour tout $\alpha \in]0,1[$:
    $$\lim_{n\to\infty}\mathbb{E}\big(X_{(n,\lfloor \alpha n\rfloor)}\big) = -\frac{1}{\lambda}\ln(1-\alpha).$$

11. Montrer que pour tout $\alpha \in]0,1[$:
    $$\lim_{n\to\infty}\mathbb{V}\big(X_{(n,\lfloor \alpha n\rfloor)}\big) = 0.$$

12. (a) Soit $\varepsilon > 0$. Montrer que :
    $$\mathbb{P}\left(\left|X_{(n,\lfloor \alpha n\rfloor)} + \frac{1}{\lambda}\ln(1-\alpha)\right| \ge \varepsilon\right) \leqslant \frac{1}{\varepsilon^2}\,\mathbb{E}\left(\left(X_{(n,\lfloor \alpha n\rfloor)} + \frac{1}{\lambda}\ln(1-\alpha)\right)^2\right).$$

(b) En déduire que la suite de variables aléatoires $\left(X_{(n,\lfloor \alpha n\rfloor)}\right)$ converge en probabilité vers une variable aléatoire certaine égale à $-\dfrac{1}{\lambda}\ln(1-\alpha)$.

13. Quelle question de la partie 1 permet de retrouver le résultat précédent ?

14. On souhaite estimer la valeur du paramètre λ à l'aide de l'observation de l'échantillon $(X_i)_{i\in\mathbb{N}^*}$. Pour cela on introduit pour tout $m \in \mathbb{N}^*$ et $t \in]0,+\infty[$, la variable aléatoire :

$$V_{m,t} = \frac{\mathrm{Card}\big(\{i \in [\![1,m]\!],\ X_i > t\}\big)+1}{m}$$

(a) Montrer que pour tout $t \in]0,+\infty[$, la suite de variables aléatoires $(V_{m,t})_{m\in\mathbb{N}^*}$ converge en probabilité vers une constante que l'on précisera.

(b) Soient $\varepsilon \in \mathbb{R}_+^*$, $m \in \mathbb{N}^*$ et $t \in]0,+\infty[$. Montrer que :

$$\mathbb{P}\left(|-\ln(V_{m,t}) - \lambda t| \leqslant \varepsilon\right) \geqslant \mathbb{P}\left(\left|V_{m,t} - e^{-\lambda t}\right| \leqslant e^{-\lambda t}(1-e^{-\varepsilon})\right)$$

En déduire que la suite de variables aléatoires $(-\ln(V_{m,t}))_{m\in\mathbb{N}^*}$ converge en probabilité vers une constante que l'on précisera.

(c) On fixe deux réels t_1 et t_2 tels que $0 < t_1 < t_2$.
Pour tout $m \in \mathbb{N}^*$, on pose :

$$W_m = \frac{1}{t_2-t_1}\ln\left(\frac{\mathrm{Card}\big(\{i \in [\![1,m]\!],\ X_i > t_1\}\big)+1}{\mathrm{Card}\big(\{i \in [\![1,m]\!],\ X_i > t_2\}\big)+1}\right)$$

Montrer que W_m est un estimateur convergent de λ.

Troisième partie

15. On admet **pour cette question** le théorème suivant :
Théorème : *Soit* $n \in \mathbb{N}^*$ *fixé.*
Il existe une famille $(T_{n,1},\ldots,T_{n,n})$ *de variables aléatoires mutuellement indépendantes telles que :*
— *Pour tout* $i \in [\![1,n]\!]$*, la variable aléatoire* $T_{n,i}$ *suit la loi exponentielle* $\mathcal{E}\big((n-i+1)\lambda\big)$;
— *Pour tout* $k \in [\![1,n]\!]$*, on a* $X_{(n,k)} = T_{n,1}+\cdots+T_{n,k}$.

En quoi ce théorème donne-t-il un résultat plus fort que ce qui a été prouvé en question 8b ?

****

Le but de cette partie est de démontrer ce théorème dans le cas où $n=2$.

Soit $q \in]0,1[$ et soit $p = 1-q$.
On dit qu'une variable aléatoire Z à valeurs dans $\mathbb{N}$ suit la loi $\mathcal{L}(q)$ si :

$$\forall k \in \mathbb{N},\quad \mathbb{P}(Z=k) = pq^k.$$

16. On suppose que Z suit la loi $\mathcal{L}(q)$. Donner la valeur de l'espérance et de la variance de Z. Déterminer $\mathbb{P}(Z \geq \ell)$ pour tout $\ell \in \mathbb{N}$.

17. Soit $n \in \mathbb{N}^*$ et soit $(Z_1,\ldots,Z_n)$ une famille de variables aléatoires mutuellement indépendantes de loi $\mathcal{L}(q)$.
On pose $U_n = \min(Z_1,\ldots,Z_n)$.

(a) Calculer $\mathbb{P}(U_n \geq \ell)$ pour tout $\ell \in \mathbb{N}$.

(b) En déduire que U_n suit une loi $\mathcal{L}(r)$ pour un certain paramètre r à déterminer.

18. Soient Z_1, Z_2 deux variables aléatoires indépendantes et de même loi $\mathcal{L}(q)$. Soient k_1, k_2 deux entiers naturels. Montrer que :

$$\mathbb{P}\big([Z_1 \geq k_1] \cap [Z_2 - Z_1 \geq k_2]\big) = \frac{1}{1+q} q^{2k_1} q^{k_2}.$$

19. Soient X_1, X_2 deux variables aléatoires indépendantes suivant la loi exponentielle $\mathcal{E}(\lambda)$.
Soient x, y deux réels positifs ou nuls.

(a) Soit m un entier naturel non nul. On pose $Z_1 = \lfloor mX_1 \rfloor$ et $Z_2 = \lfloor mX_2 \rfloor$.
Montrer que les variables aléatoires Z_1 et Z_2 suivent toutes les deux la loi $\mathcal{L}(q)$, où q est un paramètre que l'on précisera.

(b) On conserve les notations de la question précédente et on pose de plus
$k_1 = \lfloor mx \rfloor$ et $k_2 = \lfloor my \rfloor$.
Montrer les inclusions entre événements :

$$[Z_1 \geq k_1 + 1] \subset [X_1 > x] \subset [Z_1 \geq k_1],$$

$$[Z_2 - Z_1 \geq k_2 + 2] \subset [X_2 - X_1 > y] \subset [Z_2 - Z_1 \geq k_2].$$

(c) En déduire l'encadrement :

$$\frac{1}{1+e^{-\lambda/m}} e^{-2\lambda\frac{\lfloor mx \rfloor+1}{m}} e^{-\lambda\frac{\lfloor my \rfloor+2}{m}} \leq \mathbb{P}\big([X_1 > x] \cap [X_2 - X_1 > y]\big) \leq \frac{1}{1+e^{-\lambda/m}} e^{-2\lambda\frac{\lfloor mx \rfloor}{m}} e^{-\lambda\frac{\lfloor my \rfloor}{m}}.$$

(d) Montrer que :

$$\mathbb{P}\big([X_1 > x] \cap [X_2 - X_1 > y]\big) = \frac{1}{2} e^{-2\lambda x} e^{-\lambda y}.$$

20. Soient X_1, X_2 deux variables aléatoires indépendantes suivant la loi exponentielle $\mathcal{E}(\lambda)$.
On note $X_{(1)}$ et $X_{(2)}$ les variables aléatoires notées précédemment $X_{(2,1)}$ et $X_{(2,2)}$.
Autrement dit pour tout $\omega \in \Omega$, si $X_1(\omega) \leq X_2(\omega)$, alors $X_{(1)}(\omega) = X_1(\omega)$ et $X_{(2)}(\omega) = X_2(\omega)$, et dans le cas contraire $X_{(1)}(\omega) = X_2(\omega)$ et $X_{(2)}(\omega) = X_1(\omega)$.

(a) Montrer que $\mathbb{P}(X_1 = X_2) = 0$. En déduire que les variables aléatoires $X_{(1)}$ et $X_{(2)} - X_{(1)}$ sont presque-sûrement à valeurs strictement positives.

(b) Montrer que pour tous réels positifs x et y :

$$\mathbb{P}\big([X_{(1)} > x] \cap [X_{(2)} - X_{(1)} > y]\big) = 2\,\mathbb{P}\big([X_1 > x] \cap [X_2 - X_1 > y]\big).$$

(c) Montrer que les variables aléatoires $T_1 = X_{(1)}$ et $T_2 = X_{(2)} - X_{(1)}$ sont indépendantes, de loi respective $\mathcal{E}(2\lambda)$ et $\mathcal{E}(\lambda)$.
Conclure quant à l'objectif de cette partie.

FIN