



PRÉPASIMPLE

prepasimple.fr

ESCP — 2024

Sujet de mathématiques approfondies

Mathématiques approfondies — Prépa ECG

Annales de concours · Sujet

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

Pour tout réel x , $\lfloor x \rfloor$ désigne sa partie entière, c'est-à-dire l'entier relatif vérifiant $\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$.

Si A est un ensemble fini non vide, on note $\text{Card}(A)$ son nombre d'éléments. On convient que $\text{Card}(\emptyset) = 0$.

Pour les programmes Python, on suppose effectuées les importations suivantes :

Python

```
import numpy as np

import numpy.random as rd

import matplotlib.pyplot as plt
```

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 $(X_1(\omega), X_2(\omega), \dots, X_n(\omega))$ en :

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

On admet que les applications $X_{(n,k)}$ ainsi définies sont des variables aléatoires réelles sur $(\Omega, \mathcal{A}, P)$. Pour tous entiers k, n tels que $1 \leq k \leq n$, $X_{(n,k)}$ 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 $X_{(n,1)} = \min_{1 \leq i \leq n} X_i$ et $X_{(n,n)} = \max_{1 \leq i \leq n} X_i$.

Dans toutes les questions faisant intervenir $X_{(n, \lfloor \alpha n \rfloor)}$ ou sa fonction de répartition, avec $0 < \alpha < 1$, on considère les entiers n suffisamment grands pour que $1 \leq \lfloor \alpha n \rfloor \leq n$.

1. Soit $n \in \mathbb{N}^*$.

1. (a) Exprimer, à l'aide de F , la fonction de répartition $F_{(n,n)}$ de $X_{(n,n)}$. En déduire que $X_{(n,n)}$ admet comme densité la fonction $f_{(n,n)}$ définie par :

$$\forall x \in \mathbb{R}, \quad f_{(n,n)}(x) = n f(x) F(x)^{n-1}.$$

1. (b) Par une méthode similaire, montrer que $X_{(n,1)}$ admet comme densité $f_{(n,1)}$ définie par :

$$\forall x \in \mathbb{R}, \quad f_{(n,1)}(x) = n f(x) (1 - F(x))^{n-1}.$$

2. Soit $x \in \mathbb{R}$ et soient n, 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 :

$$Y_{i,x} = \begin{cases} 1 & \text{si } X_i \leq x, \\ 0 & \text{sinon.} \end{cases}$$

2. (a) Montrer que $Y_{i,x}$ suit une loi de Bernoulli dont on précisera le paramètre.

2. (b) Montrer l'égalité entre les événements :

$$[X_{(n,k)} \leq x] = \left[\sum_{i=1}^n Y_{i,x} \geq k \right].$$

2. (c) En déduire que la fonction de répartition $F_{(n,k)}$ de $X_{(n,k)}$ 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, pour tout $z \in \mathbb{R}$, par :

$$b_{n,\ell}(z) = \binom{n}{\ell} z^\ell (1-z)^{n-\ell}, \quad a_{n,k}(z) = \sum_{\ell=k}^n b_{n,\ell}(z).$$

On convient également que pour tout $n \in \mathbb{N}^*$, la fonction $b_{n-1,n}$ est identiquement nulle.

3. (a) Soient $n \in \mathbb{N}^*$ et $\ell \in \{1, \dots, n\}$. Montrer que pour tout $z \in \mathbb{R}$:

$$b'_{n,\ell}(z) = n(b_{n-1,\ell-1}(z) - b_{n-1,\ell}(z)).$$

3. (b) Soient $n \in \mathbb{N}^*$ et $k \in \{1, \dots, n\}$. Dédurre de la question précédente une expression de $a'_{n,k}$ en fonction d'une des fonctions $b_{m,\ell}$, avec m et ℓ à déterminer en fonction de n et k .

3. (c) Soient $n \in \mathbb{N}^*$ et $k \in \{1, \dots, n\}$. Montrer que $X_{(n,k)}$ 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 $0 < F(x) < 1$.

4. (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 :

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

et en déduire que $\lim_{n \rightarrow +\infty} F_{(n, \lfloor \alpha n \rfloor)}(x) = 0$.

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

5. On suppose 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 $x \in \mathbb{R} \setminus I$.
5. (a) Montrer que F établit une bijection entre les intervalles I et $]0, 1[$.
5. (b) On note F^{-1} la bijection réciproque de F restreinte à I et on fixe $\alpha \in]0, 1[$. Montrer que $(X_{(n, \lfloor \alpha n \rfloor)})$ converge en loi vers une variable aléatoire certaine égale à $F^{-1}(\alpha)$.
5. (c) En déduire que $(X_{(n, \lfloor \alpha n \rfloor)})$ converge en probabilité vers une variable aléatoire certaine égale à $F^{-1}(\alpha)$.
6. 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 appelé tri par insertion, défini de la façon suivante :

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

Par exemple, en partant du tableau de contenu $[38, 28, 35, 24, 31, 15]$, les étapes donnent successivement :

$[38, 28, 35, 24, 31, 15]$,

$[28, 38, 35, 24, 31, 15]$,

$[28, 35, 38, 24, 31, 15]$,

$[24, 28, 35, 38, 31, 15]$,

$[24, 28, 31, 35, 38, 15]$,

$[15, 24, 28, 31, 35, 38]$.

Le tri doit être «en place» : 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 cet algorithme.

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

Python

```
def ech(n):  
    # A completer  
    pass  
  
def X(n, k):  
    res = ech(n)  
    # A completer  
  
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)))
```

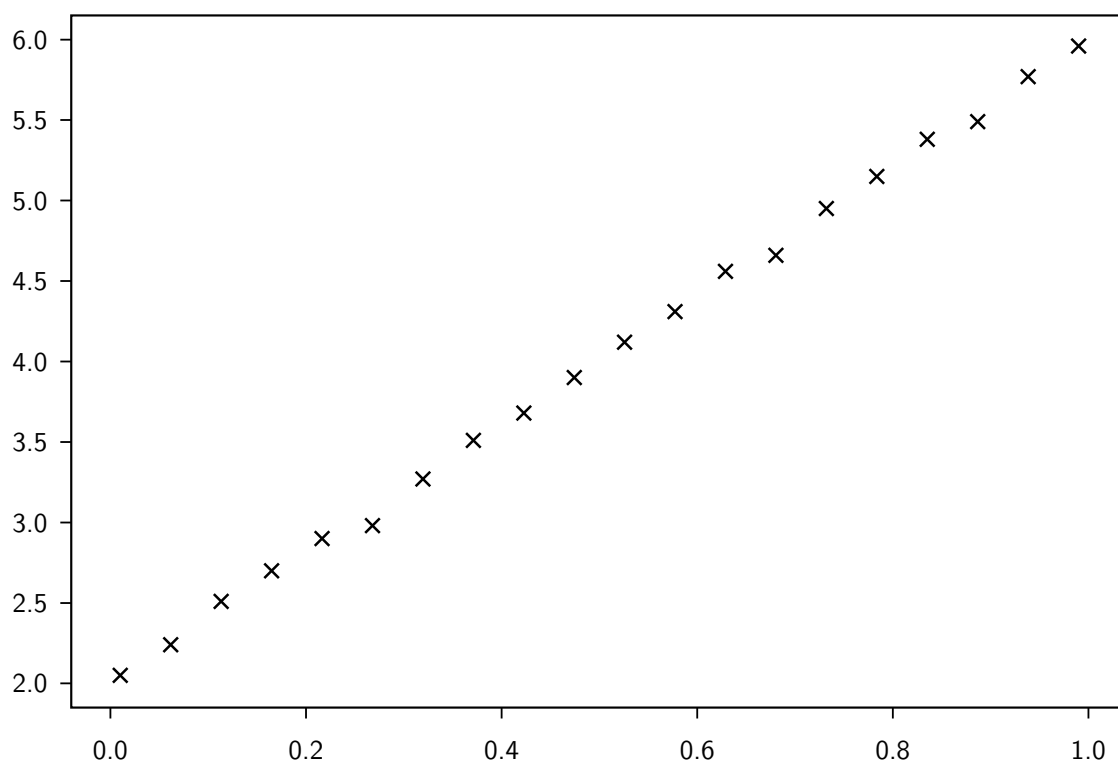
```
plt.plot(A, B, "xk")
plt.show()
```

Les commentaires « À compléter » repèrent les zones à renseigner. Dans `ech`, `pass` est une instruction vide à remplacer par le corps demandé.

La fonction `ech`, supposée 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 $(X_1, \dots, X_n)$.

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

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



Les positions des croix sont reproduites à la précision du graphique imprimé.

À quoi correspondent les points affichés sur cette représentation graphique ?

7. (c) Utiliser la représentation graphique pour proposer une loi usuelle compatible avec la loi commune des X_i simulées par ech. Expliquer pourquoi une représentation de valeurs simulées ne permet pas de déterminer cette loi avec certitude.
7. (d) Compléter le corps de ech par une ou plusieurs lignes de code pour que le script produise une représentation de valeurs simulées proche de celle ci-dessus.

Deuxième partie

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

8. Soit $n \in \mathbb{N}^*$ fixé.

8. (a) On note $(T_1, \dots, T_n)$ une famille de variables aléatoires mutuellement indépendantes telle que, pour tout $i \in \{1, \dots, n\}$, T_i suit la loi exponentielle $\mathcal{E}((n - i + 1)\lambda)$. Pour tout $k \in \{1, \dots, n\}$, on note $S_k = T_1 + \dots + T_k$.

Montrer par récurrence sur $k \in \{1, \dots, n\}$ que S_k suit la loi de densité $g_{n,k}$ définie par :

$$g_{n,k}(x) = \begin{cases} n \binom{n-1}{k-1} \lambda e^{-\lambda(n-k+1)x} (1 - e^{-\lambda x})^{k-1} & \text{si } x > 0, \\ 0 & \text{si } x \leq 0. \end{cases}$$

8. (b) Soit $k \in \{1, \dots, n\}$. Vérifier que S_k suit la même loi que $X_{(n,k)}$ lorsque les X_i suivent toutes $\mathcal{E}(\lambda)$.
9. Soit (n, k) un couple d'entiers tel que $1 \leq k \leq n$. Montrer que $X_{(n,k)}$ admet une espérance et une variance données par :

$$E(X_{(n,k)}) = \frac{1}{\lambda} \sum_{i=1}^k \frac{1}{n - i + 1}, \quad V(X_{(n,k)}) = \frac{1}{\lambda^2} \sum_{i=1}^k \frac{1}{(n - i + 1)^2}.$$

10. (a) Montrer que pour tout entier $j \geq 1$ et tout $x \in [j, j+1]$, on a :

$$0 \leq \frac{1}{j} - \frac{1}{x} \leq \frac{2}{x^2}.$$

10. (b) En déduire l'encadrement, pour $1 \leq k \leq n$:

$$\frac{1}{\lambda} \ln \left(\frac{n+1}{n-k+1} \right) \leq E(X_{(n,k)}) \leq \frac{1}{\lambda} \ln \left(\frac{n+1}{n-k+1} \right) + \frac{2}{\lambda} \left(\frac{1}{n-k+1} - \frac{1}{n+1} \right).$$

10. (c) Montrer que pour tout $\alpha \in]0, 1[$:

$$\lim_{n \rightarrow +\infty} E(X_{(n, \lfloor \alpha n \rfloor)}) = -\frac{1}{\lambda} \ln(1 - \alpha).$$

11. Montrer que pour tout $\alpha \in]0, 1[$:

$$\lim_{n \rightarrow +\infty} V(X_{(n, \lfloor \alpha n \rfloor)}) = 0.$$

12. On fixe $\alpha \in]0, 1[$.

12. (a) Soit $\varepsilon > 0$. Montrer que :

$$P \left(\left| X_{(n, \lfloor \alpha n \rfloor)} + \frac{1}{\lambda} \ln(1 - \alpha) \right| \geq \varepsilon \right) \leq \frac{1}{\varepsilon^2} E \left(\left(X_{(n, \lfloor \alpha n \rfloor)} + \frac{1}{\lambda} \ln(1 - \alpha) \right)^2 \right).$$

12. (b) En déduire que $(X_{(n, \lfloor \alpha n \rfloor)})$ converge en probabilité vers une variable aléatoire certaine égale à $-\frac{1}{\lambda} \ln(1 - \alpha)$.

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

14. On souhaite estimer λ à l'aide de l'observation de l'échantillon $(X_i)_{i \in \mathbb{N}^*}$. Pour tout $m \in \mathbb{N}^*$ et $t > 0$, on introduit :

$$V_{m,t} = \frac{\text{Card}(\{i \in \{1, \dots, m\} : X_i > t\}) + 1}{m}.$$

14. (a) Montrer que pour tout $t > 0$, $(V_{m,t})_{m \in \mathbb{N}^*}$ converge en probabilité vers une constante que l'on précisera.

14. (b) Soient $\varepsilon > 0$, $m \in \mathbb{N}^*$ et $t > 0$. Montrer que :

$$P(|-\ln(V_{m,t}) - \lambda t| \leq \varepsilon) \geq P(|V_{m,t} - e^{-\lambda t}| \leq e^{-\lambda t}(1 - e^{-\varepsilon})).$$

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

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

$$W_m = \frac{1}{t_2 - t_1} \ln \left(\frac{\text{Card}(\{i \in \{1, \dots, m\} : X_i > t_1\}) + 1}{\text{Card}(\{i \in \{1, \dots, m\} : X_i > t_2\}) + 1} \right).$$

Montrer que W_m est un estimateur convergent de λ .

Troisième partie

15. On admet pour cette question le théorème suivant.

Soit $n \in \mathbb{N}^*$ fixé. Il existe une famille $(T_{n,1}, \dots, T_{n,n})$ de variables aléatoires mutuellement indépendantes telle que :

- Pour tout $i \in \{1, \dots, n\}$, $T_{n,i}$ suit $\mathcal{E}((n - i + 1)\lambda)$.
- Pour tout $k \in \{1, \dots, n\}$, $X_{(n,k)} = T_{n,1} + \dots + T_{n,k}$.

En quoi ce théorème donne-t-il un résultat plus fort que celui prouvé en question 8. (b) ?

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

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

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

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

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

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

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

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

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

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

19. (a) Soit $m \in \mathbb{N}^*$. On pose $Z_1 = \lfloor mX_1 \rfloor$ et $Z_2 = \lfloor mX_2 \rfloor$. Montrer que Z_1, Z_2 suivent toutes deux $\mathcal{L}(q)$, où q est à préciser.

19. (b) On conserve les notations précédentes et on pose $k_1 = \lfloor mx \rfloor, k_2 = \lfloor my \rfloor$. Montrer les inclusions :

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

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

19. (c) En déduire l'encadrement :

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

19. (d) Montrer que :

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

20. Soient X_1, X_2 deux variables aléatoires indépendantes suivant $\mathcal{E}(\lambda)$. On note $X_{(1)}$ et $X_{(2)}$ les variables précédemment notées $X_{(2,1)}$ et $X_{(2,2)}$.

Pour tout $\omega \in \Omega$, si $X_1(\omega) \leq X_2(\omega)$, alors $X_{(1)}(\omega) = X_1(\omega)$ et $X_{(2)}(\omega) = X_2(\omega)$; dans le cas contraire, $X_{(1)}(\omega) = X_2(\omega)$ et $X_{(2)}(\omega) = X_1(\omega)$.

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

20. (b) Montrer que pour tous réels positifs x, y :

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

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