



PRÉPASIMPLE

prepasimple.fr

HEC

Mathématiques — Option économique

2015

Mathématiques appliquées

Mathématiques appliquées — Prépa ECG

Annales de concours · Sujet

Exercice

Soit n un entier supérieur ou égal à 2 et $\mathcal{B} = (e_1, e_2, \dots, e_n)$ la base canonique de $\mathbb{R}^n$.

Soit v un vecteur donné de $\mathbb{R}^n$, de coordonnées $v_1, v_2, \dots, v_n$ dans la base $\mathcal{B}$, qui vérifie $\sum_{i=1}^n v_i = 1$.

Soit f l'application définie sur $\mathbb{R}^n$ qui, à tout vecteur $x = (x_1, x_2, \dots, x_n)$, associe le vecteur

$$f(x) = x - \left(\sum_{i=1}^n x_i\right)v.$$

1.a) Montrer que f est un endomorphisme de $\mathbb{R}^n$.

1.b) Montrer que $f \circ f = f$.

2. Déterminer le spectre de f .

3.a) Montrer que le vecteur y appartient à l'image de f , notée $\text{Im } f$, si et seulement si $f(y) = y$.

3.b) Montrer que la dimension de $\text{Im } f$ est inférieure ou égale à $n - 1$.

3.c) Montrer que, pour tout $i \in \{1, \dots, n - 1\}$, on a $(e_i - e_{i+1}) \in \text{Im } f$.

3.d) En déduire une base et la dimension de $\text{Im } f$. Quel est le rang de f ?

4.a) Déterminer une base du noyau de f .

4.b) Quels sont les sous-espaces propres de f ?

4.c) L'endomorphisme f est-il diagonalisable?

5. Écrire la matrice M de l'endomorphisme f dans la base canonique de $\mathbb{R}^n$ et la matrice M' de f dans une base de vecteurs propres.

Problème

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

On considère une variable aléatoire X à valeurs strictement positives suivant la loi exponentielle de paramètre $\lambda > 0$, de densité

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x > 0, \\ 0 & \text{si } x \leq 0. \end{cases}$$

Pour un paramètre réel θ , un estimateur S_n est dit sans biais si son espérance existe et vaut θ . On le dit convergent si, pour tout $\varepsilon > 0$, $P(|S_n - \theta| \geq \varepsilon) \rightarrow 0$ lorsque $n \rightarrow +\infty$.

Partie I. Comparaison de deux estimateurs de $\frac{1}{\lambda}$

L'objectif de cette partie est de comparer deux estimateurs sans biais et convergents du paramètre inconnu $\frac{1}{\lambda}$.

Pour tout $n \in \mathbb{N}^*$, soit $(X_1, X_2, \dots, X_n)$ un n -échantillon de variables aléatoires à valeurs strictement positives, indépendantes et de même loi que X .

On pose, pour tout $n \in \mathbb{N}^*$,

$$Y_n = \sum_{i=1}^n X_i, \quad \bar{X}_n = \frac{1}{n} Y_n, \quad M_n = \max(X_1, X_2, \dots, X_n).$$

1. Rappeler sans démonstration les valeurs de l'espérance $E(X)$, de la variance $V(X)$ ainsi que l'expression de la fonction de répartition F_X de la variable aléatoire X .
2. Calculer $E(\bar{X}_n)$ et $V(\bar{X}_n)$. En déduire que $\bar{X}_n$ est un estimateur sans biais et convergent du paramètre $\frac{1}{\lambda}$.
- 3.a) Montrer qu'une densité f_{M_n} de M_n est donnée par

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

3.b) Établir, pour tout $n \in \mathbb{N}^*$, l'existence de l'espérance $E(M_n)$ de la variable aléatoire M_n .

3.c) En posant $z = 1 - e^{-\lambda x}$, justifier, pour tout $a > 0$, l'égalité

$$\int_0^a x e^{-\lambda x} (1 - e^{-\lambda x})^{n-1} dx = -\frac{1}{\lambda^2} \int_0^{1-e^{-\lambda a}} z^{n-1} \ln(1-z) dz.$$

3.d) En déduire que l'on a

$$E(M_n) = -\frac{n}{\lambda} \int_0^1 z^{n-1} \ln(1-z) dz.$$

L'intégrale à droite désigne la limite lorsque $b \rightarrow 1^-$ des intégrales sur $[0, b]$; on justifiera son existence à l'aide de la question précédente et de l'existence de $E(M_n)$.

3.e) Montrer que la fonction $z \mapsto (1-z)(1 - \ln(1-z))$, définie sur $[0, 1[$, est une primitive de $z \mapsto \ln(1-z)$.

À l'aide d'une intégration par parties, en déduire, pour tout $n \in \mathbb{N}^*$, une relation entre $E(M_{n+1})$ et $E(M_n)$. On effectuera d'abord l'intégration par parties sur $[0, b]$, avec $b < 1$, puis on justifiera le passage à la limite.

3.f) On pose, pour tout $n \in \mathbb{N}^*$,

$$u_n = \sum_{j=1}^n \frac{1}{j}.$$

Déduire de la question précédente que $E(M_n) = \frac{1}{\lambda} u_n$.

4. On pose, pour tout $n \in \mathbb{N}^*$, $M'_n = \frac{M_n}{u_n}$ et

$$v_n = \sum_{j=1}^n \frac{1}{j^2}.$$

On admet que $V(M_n) = \frac{1}{\lambda^2} v_n$.

4.a) Calculer $E(M'_n)$ et $V(M'_n)$.

4.b) Justifier la convergence de la suite $(v_n)_{n \in \mathbb{N}^*}$. Déterminer $\lim_{n \rightarrow +\infty} u_n$.

4.c) Dédire des questions 4.a) et 4.b) que M'_n est un estimateur sans biais et convergent du paramètre $\frac{1}{\lambda}$.

4.d) Soit Q la fonction polynomiale définie sur $\mathbb{R}$ par

$$Q(x) = \sum_{j=1}^n \left(x - \frac{1}{j} \right)^2.$$

À l'aide de l'étude de Q , établir l'inégalité $u_n^2 \leq n v_n$.

4.e) Comparer alors $V(M'_n)$ et $V(\bar{X}_n)$. Conclure.

Partie II. Un exemple

Les notations et le contexte sont ceux de la partie I.

Dans cette partie, on suppose que la durée de vie, en heures, d'un composant électronique est une variable aléatoire X à valeurs strictement positives suivant la loi exponentielle de paramètre $\lambda > 0$.

On suppose qu'en cas de panne, le composant électronique est immédiatement remplacé par un composant neuf dont la durée de vie est indépendante et de même loi que celle des composants précédents.

Pour $j \in \mathbb{N}^*$, on note X_j la variable aléatoire égale à la durée de vie du j -ième composant. Pour tout $n \in \mathbb{N}^*$, on constitue ainsi un échantillon $(X_1, X_2, \dots, X_n)$ de

variables aléatoires à valeurs strictement positives, indépendantes et de même loi que X .

On note $(x_1, x_2, \dots, x_n)$ la réalisation du n -échantillon $(X_1, X_2, \dots, X_n)$ et l'on pose

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i.$$

5.a) Donner une interprétation de $\frac{1}{\lambda}$. Dans quelle unité s'exprime $\frac{1}{\lambda}$?

5.b) Soit $p \in]0, 1[$. Déterminer, en fonction de p et de λ , l'unique réel h_p pour lequel on a $P(X > h_p) = p$.

5.c) En construire un estimateur H_n sans biais et convergent pour le paramètre h_p .

5.d) On suppose que, sur un échantillon de 100 composants, on a obtenu

$$\sum_{i=1}^{100} x_i = 10^5$$

heures. Donner une estimation de $h_{1/2}$; on donne $\ln 2 \simeq 0,7$.

6. On admet sans démonstration que, pour tout $n \in \mathbb{N}^*$, la fonction de répartition F_{Y_n} de la variable aléatoire Y_n est donnée par

$$F_{Y_n}(x) = \begin{cases} 1 - e^{-\lambda x} \sum_{k=0}^{n-1} \frac{(\lambda x)^k}{k!} & \text{si } x > 0, \\ 0 & \text{sinon.} \end{cases}$$

Soit $t > 0$ un réel fixé et $N(t)$ la variable aléatoire égale au nombre de pannes dans l'intervalle $[0, t[$.

6.a) Déterminer la loi de la variable aléatoire $N(t)$.

6.b) Donner une estimation « naturelle » du nombre de pannes dans l'intervalle $[0, t[$.

7. L'objectif de cette question est de déterminer un intervalle de confiance asymptotique du paramètre inconnu $\frac{1}{\lambda}$, au niveau de confiance $1 - \alpha$, avec $0 < \alpha < 1$.

On entend par là une suite d'intervalles aléatoires I_n , calculables à partir de l'échantillon et indépendants de la valeur inconnue de λ , telle que $P(\frac{1}{\lambda} \in I_n) \rightarrow 1 - \alpha$.

7.a) Soit Φ la fonction de répartition de la loi normale centrée réduite, et soient t_1 et t_2 les réels qui vérifient $\Phi(t_1) = 1 - \frac{\alpha}{2}$ et $\Phi(t_2) = \frac{\alpha}{2}$. Montrer que $t_2 = -t_1$.

7.b) On pose $R_n = \sqrt{n}(\lambda \bar{X}_n - 1)$. Justifier que la suite de variables aléatoires $(R_n)_{n \in \mathbb{N}^*}$ converge en loi vers une variable aléatoire suivant la loi normale centrée réduite.

7.c) En déduire $\lim_{n \rightarrow +\infty} P(-t_1 \leq R_n \leq t_1) = 1 - \alpha$.

7.d) Déterminer un intervalle de confiance asymptotique de $\frac{1}{\lambda}$ au niveau de confiance $1 - \alpha$, en inversant les inégalités de la question précédente. On précisera les bornes lorsque $n > t_1^2$ et la forme de l'intervalle lorsque $n \leq t_1^2$.

7.e) Que se passe-t-il lorsque α est proche de 0 ou lorsque α est proche de 1 ?

7.f) Pour $\alpha = 0,05$, on donne $t_1 \simeq 2$. Montrer qu'avec l'échantillon de la question 5.d), la réalisation de l'intervalle de confiance asymptotique de $\frac{1}{\lambda}$ au niveau de confiance 0,95 est, après arrondi des bornes, [833; 1250].

Partie III. Un résultat asymptotique

Les notations et le contexte sont ceux des parties précédentes.

Pour tout $n \in \mathbb{N}^*$, on pose $T_n = M_n - \frac{1}{\lambda} \ln n$, et l'on note F_{T_n} la fonction de répartition de T_n .

8.a) Montrer que, pour tout réel x , on a $F_{T_n}(x) = \left(F_X\left(x + \frac{1}{\lambda} \ln n\right)\right)^n$.

8.b) Pour chaque réel x fixé, déterminer un entier naturel N_x , qui dépend de x , tel que, pour tout entier $n \geq N_x$, on ait $x + \frac{1}{\lambda} \ln n > 0$.

En déduire que, pour tout entier $n \geq N_x$, on a $F_X\left(x + \frac{1}{\lambda} \ln n\right) = 1 - \frac{e^{-\lambda x}}{n}$.

8.c) Montrer que, pour tout réel x , on a $\lim_{n \rightarrow +\infty} F_{T_n}(x) = \exp(-e^{-\lambda x})$.

9. Soit F la fonction définie sur $\mathbb{R}$ par $F(x) = \exp(-e^{-\lambda x})$.

9.a) Justifier que F est de classe C^∞ sur $\mathbb{R}$ et montrer que F réalise une bijection de $\mathbb{R}$ sur l'intervalle $]0, 1[$.

9.b) En déduire que F est la fonction de répartition d'une variable aléatoire T admettant une densité f_T continue sur $\mathbb{R}$, que l'on déterminera ; on dit que T suit la loi de Gumbel de paramètre λ .

9.c) Justifier la convergence en loi de la suite $(T_n)_{n \in \mathbb{N}^*}$ vers la variable aléatoire T .

10.a) Établir l'existence de l'espérance $E(T)$ de la variable aléatoire T . Si l'on utilise le changement de variable $t = e^{-\lambda x}$, on l'effectuera d'abord sur des segments, puis on traitera les limites en 0 et en $+\infty$.

10.b) On pose $Z = -\frac{1}{\lambda} \ln(\lambda X)$. Montrer que les variables aléatoires Z et T sont de même loi.

10.c) Justifier l'égalité $E(T) = -\frac{1}{\lambda} \int_0^{+\infty} e^{-t} \ln t \, dt$. Pour la borne 0, on pourra utiliser $0 \leq -e^{-t} \ln t \leq -\ln t$ sur $]0, 1]$ et une primitive de $\ln t$.

10.d) À l'aide de la concavité de la fonction $\ln$ sur $\mathbb{R}_+^*$, établir l'inégalité $E(T) \geq 0$.

11. On suppose dans cette question que $\lambda = 1$.

11.a) Expliciter la bijection réciproque G de la fonction F .

11.b) On considère le programme Python suivant :

Python

```
import numpy as np

import matplotlib.pyplot as plt

x = np.linspace(-2, 2, 400)
y = np.exp(-np.exp(-x))
plt.plot(x, y)
```

```
plt.plot(y, x)
plt.show()
```

La commande `np.linspace(a, b, m)` renvoie les m réels régulièrement espacés de a à b , bornes comprises. Les fonctions `np.exp` et `np.log` agissent terme à terme sur les tableaux. La commande `plt.plot(a, b)` relie les points dont les abscisses et les ordonnées figurent respectivement dans les tableaux a et b ; les deux tracés sont ici affichés sur les mêmes axes par `plt.show()`.

11.b)(i) Le réel 0 fait-il partie des nombres renvoyés par la commande `x = np.linspace(-2, 2, 400)` ?

11.b)(ii) Quel sera le résultat de l'exécution de ce programme ?

11.c) Soit U une variable aléatoire suivant la loi uniforme sur l'intervalle $]0, 1[$. Quelle est la loi de la variable aléatoire $G(U)$?

11.d) Établir l'inégalité $E(T) \leq 1$.

11.e) Par une méthode de votre choix, écrire les commandes Python qui permettent de simuler la loi de T .

On pourra utiliser `np.random.random()`, qui renvoie un nombre pseudo-aléatoire uniforme dans $[0, 1[$, ou `np.random.exponential(scale=1.0)`, qui simule une loi exponentielle de moyenne 1. Si une valeur nulle est obtenue, on la retirera avant de prendre un logarithme.

11.f) Écrire les commandes Python qui permettent de renvoyer une valeur numérique approchée de $E(T)$ en utilisant la méthode de Monte-Carlo.