



PRÉPASIMPLE

prepasimple.fr

HEC — 2025

Sujet de mathématiques approfondies

Mathématiques approfondies — Prépa ECG

Annales de concours · Sujet

Notations et rappels

Toutes les variables aléatoires sont définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$. Si Z est une variable aléatoire réelle, on note $E(Z)$ son espérance et $V(Z)$ sa variance, si elles existent.

On note φ et Φ les fonctions définies, pour tout $x \in \mathbb{R}$, par :

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \quad \Phi(x) = \int_{-\infty}^x \varphi(t) dt.$$

La fonction Gamma est définie, pour tout réel $x > 0$, par :

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt.$$

Elle vérifie, pour tout $x > 0$, $\Gamma(x+1) = x\Gamma(x)$.

Pour les programmes Python, on dispose du formulaire à la fin du sujet. On importe les bibliothèques suivantes :

Python

```
import numpy as np
import matplotlib.pyplot as plt
import numpy.random as rd
```

Première partie

Soit $(X_i)_{i \in \mathbb{N}^*}$ une famille de variables aléatoires mutuellement indépendantes, chacune de loi normale $\mathcal{N}(0, 1)$. On suppose que $X_i(\omega) \neq 0$ pour tout $i \in \mathbb{N}^*$ et tout $\omega \in \Omega$. On définit, pour tout entier $n \in \mathbb{N}^*$:

$$S_n = \sum_{i=1}^n X_i^2.$$

La loi de S_n est appelée loi du khi-deux à n degrés de liberté et on notera $S_n \hookrightarrow \chi^2(n)$. Cette notation désigne ici la loi définie par cette somme de carrés ; aucun résultat supplémentaire sur cette loi n'est supposé connu.

On pose aussi, pour tout $n \in \mathbb{N}^*$, $W_n = \frac{1}{2}S_n$.

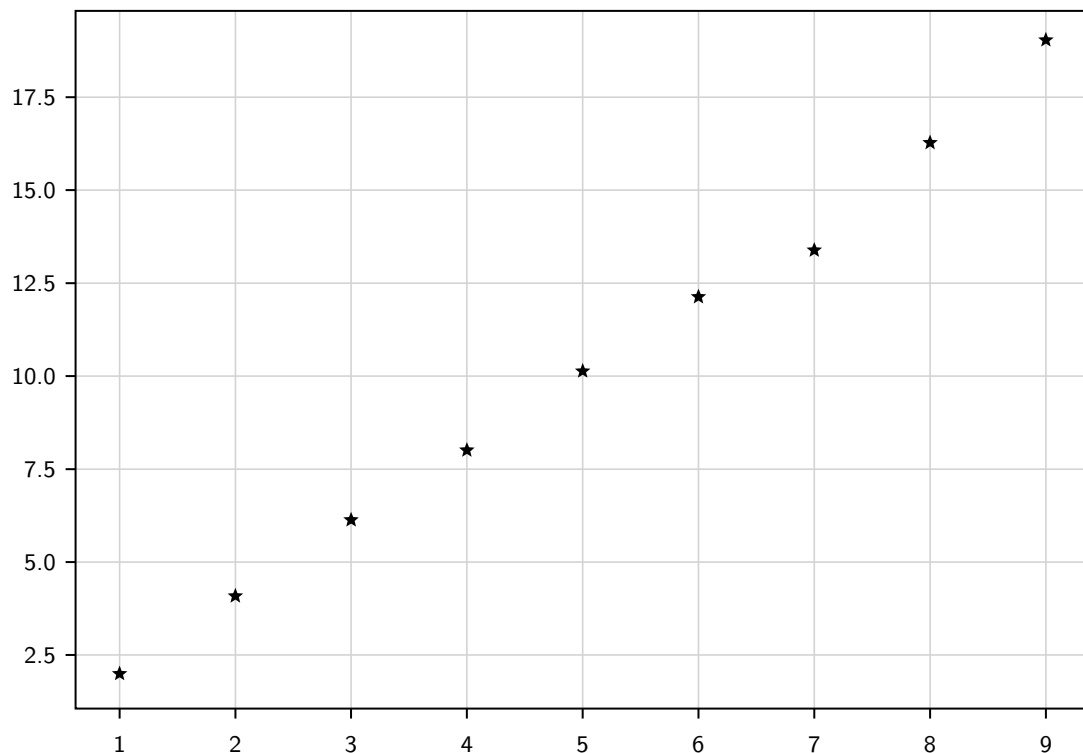
1. (a) Montrer que, pour tout $n \in \mathbb{N}^*$, S_n admet une espérance et calculer $E(S_n)$. On admet que $E(S_n^2)$ existe pour tout $n \in \mathbb{N}^*$.
1. (b) Écrire une fonction Python `simul(n)` qui renvoie une réalisation de S_n lorsqu'on lui fournit un entier naturel non nul n en entrée.
1. (c) On exécute ensuite le programme suivant :

Python

```
def f(n, N):
    t = 0
    for k in range(N):
        t = t + simul(n)**2
    return t/N - n**2

A = np.arange(1, 10)
B = np.zeros(9)
for n in A:
    B[n-1] = f(n, 50000)
plt.plot(A, B, "*k")
plt.grid()
plt.show()
```

On obtient la figure ci-dessous. Que peut-on conjecturer sur la variable aléatoire S_n ? Expliquer ce qui motive votre conjecture.



2. (a) Montrer que W_1 est une variable aléatoire à densité et déterminer une densité de W_1 .
2. (b) En déduire la valeur de $\Gamma\left(\frac{1}{2}\right)$.
2. (c) En déduire que, pour tout $n \in \mathbb{N}^*$, W_n suit la loi gamma $\gamma\left(\frac{n}{2}\right)$.
2. (d) Retrouver alors la valeur de $E(S_n)$ et déterminer la valeur de $\mathbb{V}(S_n)$.
3. (a) Montrer que, pour tout entier $n \geq 3$, $\frac{1}{W_n}$ admet une espérance et que :

$$E\left(\frac{1}{W_n}\right) = \frac{\Gamma\left(\frac{n}{2} - 1\right)}{\Gamma\left(\frac{n}{2}\right)}.$$

En déduire que $E\left(\frac{1}{S_n}\right) = \frac{1}{n-2}$.

3. (b) Dédurre de la question précédente que, pour tout entier $n \geq 3$, $\frac{1}{\sqrt{S_n}}$ admet une espérance.

Soit Y une variable aléatoire indépendante des $(X_k)_{k \geq 1}$ et de loi $\mathcal{N}(0, 1)$. Pour tout $n \geq 1$, on pose :

$$T_n = \frac{Y}{\sqrt{S_n/n}}.$$

La loi de T_n est appelée loi de Student de paramètre n et on note $T_n \hookrightarrow \mathcal{T}(n)$. Cette notation désigne ici la loi définie par ce quotient ; aucun résultat supplémentaire sur cette loi n'est supposé connu.

On admet que T_n est une variable aléatoire à densité possédant une densité strictement positive et continue sur $\mathbb{R}$.

4. Soit $\alpha \in]0, 1[$. Montrer l'existence et l'unicité d'un réel $t_{n,\alpha}$ tel que $P(|T_n| \leq t_{n,\alpha}) = 1 - \alpha$.

5. (a) Montrer que, pour tout entier $n \geq 3$, $E(T_n)$ existe et vaut 0.

5. (b) Montrer que, pour tout entier $n \geq 3$, $V(T_n)$ existe et vaut $\frac{n}{n-2}$.

5. (c) En déduire que, pour tout entier $n \geq 3$:

$$E((T_n - Y)^2) = \frac{2n-2}{n-2} - \sqrt{2n} E\left(\frac{1}{\sqrt{W_n}}\right).$$

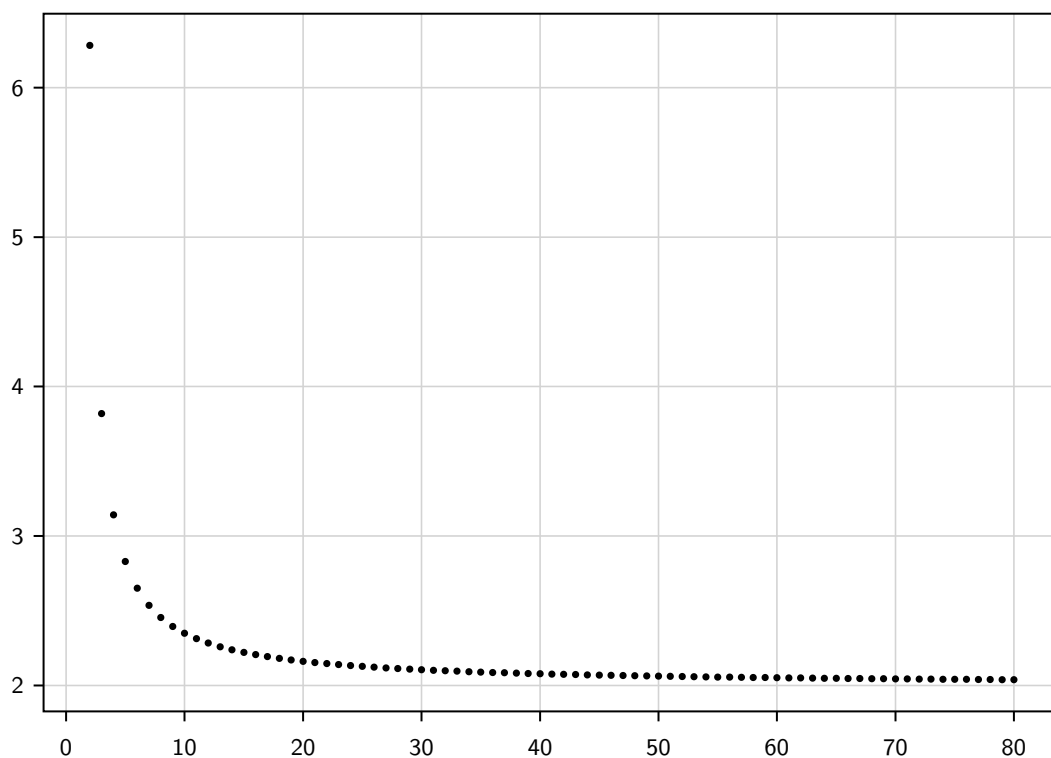
6. Pour tout entier $n \geq 2$, on pose $u_n = E\left(\frac{1}{\sqrt{W_n}}\right)$.

6. (a) Montrer que la suite $(u_n)_{n \geq 2}$ est décroissante et déterminer la valeur de u_2 .
6. (b) Montrer que, pour tout entier $n \geq 2$, on a $u_{n+1}u_n = \frac{2}{n-1}$.
6. (c) Écrire une fonction Python d'en-tête `suite_u(n)` qui renvoie, sous la forme d'un tableau NumPy, les $n-1$ valeurs $u_2, u_3, \dots, u_n$.
6. (d) On exécute ensuite le programme suivant :

Python

```
U = suite_u(80)
V = np.arange(2, 81)
plt.plot(V, V*U**2, ".k")
plt.grid()
plt.show()
```

On obtient la figure suivante. Que peut-on conjecturer concernant la suite $(u_n)_{n \geq 2}$?



6. (e) Montrer que $u_n \underset{n \rightarrow +\infty}{\sim} \sqrt{\frac{2}{n}}$.

7. Montrer que la suite $(T_n)_{n \geq 2}$ converge en probabilité vers Y .

Deuxième partie

Dans cette partie, on démontre le résultat suivant : si X et Y sont deux variables aléatoires indépendantes, chacune de loi normale $\mathcal{N}(0, 1)$, les variables aléatoires $\frac{X+Y}{\sqrt{2}}$ et $\frac{X-Y}{\sqrt{2}}$ sont indépendantes et de loi normale $\mathcal{N}(0, 1)$.

Soient $a, b \in \mathbb{R}$ fixés jusqu'à la fin de cette partie. On considère la partie $\mathcal{A} \subset \mathbb{R}^2$ définie par :

$$\mathcal{A} = \{(x, y) \in \mathbb{R}^2 \mid x + y \leq a \text{ et } x - y \leq b\}.$$

On pose $c = \frac{a+b}{2}$ et $d = \frac{a-b}{2}$.

Soient X et Y deux variables aléatoires indépendantes, chacune de loi normale $\mathcal{N}(0, 1)$. On pose $W = \frac{X+Y}{\sqrt{2}}$ et $Z = \frac{X-Y}{\sqrt{2}}$.

Pour la preuve, on admet uniquement l'identité particulière suivante :

$$P((X, Y) \in \mathcal{A}) = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} \mathbf{1}_{\mathcal{A}}(x, y) \varphi(x) \, dx \right) \varphi(y) \, dy,$$

où $\mathbf{1}_{\mathcal{A}}(x, y)$ vaut 1 si $(x, y) \in \mathcal{A}$ et 0 sinon. Les intégrales ci-dessus sont interprétées comme des intégrales successives d'une variable réelle.

On rappelle qu'un demi-plan défini par une inégalité affine large est une partie fermée de $\mathbb{R}^2$, et qu'une intersection finie de parties fermées est fermée.

8. Dessiner la partie $\mathcal{A}$ dans le cas où $a = 2$ et $b = -1$.

9. Soit y un réel tel que $y \geq d$.

9. (a) Montrer que $(x, y) \in \mathcal{A}$ si et seulement si $x \in]-\infty, a - y]$.

9. (b) En déduire que :

$$\int_{-\infty}^{+\infty} \mathbf{1}_{\mathcal{A}}(x, y) \varphi(x) \, dx = \Phi(a - y).$$

10. Montrer de la même façon que, pour tout $y \leq d$, on a :

$$\int_{-\infty}^{+\infty} \mathbf{1}_{\mathcal{A}}(x, y) \varphi(x) \, dx = \Phi(b + y).$$

11. (a) Montrer que $\mathcal{A}$ est une partie fermée de $\mathbb{R}^2$.

11. (b) Montrer que :

$$P((X, Y) \in \mathcal{A}) = \int_0^{+\infty} (\varphi(d + z) + \varphi(d - z)) \Phi(c - z) \, dz.$$

12. Montrer que :

$$P((X, Y) \in \mathcal{A}) = \int_0^{+\infty} \left(\int_{-\infty}^c (\varphi(d + z) + \varphi(d - z)) \varphi(t - z) \, dt \right) dz.$$

On admet pour la suite l'égalité particulière suivante :

$$P((X, Y) \in \mathcal{A}) = \int_{-\infty}^c \left(\int_0^{+\infty} (\varphi(d + z) + \varphi(d - z)) \varphi(t - z) \, dz \right) dt.$$

13. Montrer que, pour tous $u, v \in \mathbb{R}$:

$$\varphi(u)\varphi(v) = \varphi\left(\frac{u+v}{\sqrt{2}}\right)\varphi\left(\frac{u-v}{\sqrt{2}}\right).$$

14. Montrer que :

$$P((X, Y) \in \mathcal{A}) = \frac{1}{\sqrt{2}} \int_{-\infty}^c \left(\varphi\left(\frac{t+d}{\sqrt{2}}\right) \Phi\left(\frac{t-d}{\sqrt{2}}\right) + \varphi\left(\frac{t-d}{\sqrt{2}}\right) \Phi\left(\frac{t+d}{\sqrt{2}}\right) \right) dt.$$

15. Montrer que :

$$P((X, Y) \in \mathcal{A}) = \Phi\left(\frac{c+d}{\sqrt{2}}\right) \Phi\left(\frac{c-d}{\sqrt{2}}\right).$$

16. Conclure quant à l'objectif de cette partie.

Troisième partie

Soit $n \in \mathbb{N}$ tel que $n \geq 2$. On munit l'espace vectoriel $\mathbb{R}^n$ du produit scalaire euclidien standard, noté $\langle \cdot, \cdot \rangle$. On considère une base orthonormée $(a_1, \dots, a_n)$ de $\mathbb{R}^n$ telle que $a_1 = \frac{1}{\sqrt{n}}(1, \dots, 1)$.

17. Soit $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. On pose $\bar{x} = \frac{x_1 + \dots + x_n}{n}$.

17. (a) Montrer que :

$$\sum_{k=2}^n \langle x, a_k \rangle a_k = (x_1 - \bar{x}, \dots, x_n - \bar{x}).$$

17. (b) En déduire que :

$$\sum_{k=2}^n \langle x, a_k \rangle^2 = \sum_{i=1}^n (x_i - \bar{x})^2.$$

Soient $X_1, \dots, X_n$ des variables aléatoires mutuellement indépendantes, chacune de loi normale $\mathcal{N}(0, 1)$. On pose $X = (X_1, \dots, X_n)$. Pour $k \in \{1, \dots, n\}$, on note $\langle X, a_k \rangle$ la variable aléatoire Y_k définie par $Y_k(\omega) = \langle X(\omega), a_k \rangle$. Quitte à modifier les Y_k sur un ensemble de probabilité nulle, on peut supposer $Y_k(\omega) \neq 0$ pour tout $\omega \in \Omega$ et tout $k \in \{1, \dots, n\}$.

Dans le reste du problème, on admet le résultat particulier suivant : les variables aléatoires $Y_1, \dots, Y_n$ sont mutuellement indépendantes, chacune de loi normale $\mathcal{N}(0, 1)$.

18. (a) Retrouver le résultat de la deuxième partie à l'aide du résultat admis ci-dessus.

18. (b) Soit $(\beta_1, \beta_2) \in \mathbb{R}^2 \setminus \{(0, 0)\}$. Montrer que les variables aléatoires $R_1 = X_1 + X_2$ et $R_2 = \beta_1 X_1 + \beta_2 X_2$ sont indépendantes si et seulement si $\text{Cov}(R_1, R_2) = 0$.

19. On introduit les variables aléatoires :

$$\bar{X} = \frac{X_1 + \dots + X_n}{n}, \quad U = \sum_{i=1}^n (X_i - \bar{X})^2.$$

19. (a) Donner la loi de $\bar{X}$.

19. (b) Exprimer $\bar{X}$ et U à l'aide des variables aléatoires $Y_k, k \in \{1, \dots, n\}$.

19. (c) Montrer que $\bar{X}$ et U sont indépendantes et que U suit la loi $\chi^2(n-1)$.

On suppose pour la fin du problème que $Z = (Z_1, \dots, Z_n)$ est un n -uplet de variables aléatoires mutuellement indépendantes, chacune de loi normale $\mathcal{N}(\mu, \sigma^2)$, avec $\sigma > 0$, où (μ, σ^2) est inconnu. On pose :

$$\bar{Z} = \frac{Z_1 + \dots + Z_n}{n}, \quad V = \sum_{i=1}^n (Z_i - \bar{Z})^2.$$

Comme précédemment, on pourra supposer $V(\omega) \neq 0$ pour tout $\omega \in \Omega$.

20. (a) Pour $i \in \{1, \dots, n\}$, on pose $X_i = \frac{Z_i - \mu}{\sigma}$. On pose de plus $\bar{X} = \frac{X_1 + \dots + X_n}{n}$ et

$$U = \sum_{i=1}^n (X_i - \bar{X})^2.$$

Exprimer V en fonction de U et σ .

20. (b) En déduire que $\frac{1}{n-1}V$ est un estimateur sans biais de σ^2 .

20. (c) Montrer que $\frac{\bar{Z} - \mu}{\sqrt{V/(n(n-1))}}$ suit la loi de Student $\mathcal{T}(n-1)$, telle qu'elle a été définie dans la première partie.

21. Soit $\alpha \in]0, 1[$ et $t_{n-1, \alpha}$ le réel tel que $P(|T| \leq t_{n-1, \alpha}) = 1 - \alpha$ lorsque $T \hookrightarrow \mathcal{T}(n-1)$. Construire, en fonction de $\bar{Z}$, V , n et $t_{n-1, \alpha}$, un intervalle de confiance de μ au niveau de confiance $1 - \alpha$.

Dans la pratique, les valeurs $t_{n-1, \alpha}$ sont tabulées et permettent ainsi la construction d'intervalles de confiance non asymptotiques pour des modèles gaussiens.

Aide-mémoire Python

I. Mathématiques générales

Python

```
import numpy as np
```

Commande	Rôle
<code>np.pi</code>	Valeur approchée de π .
<code>np.linspace(a, b, n)</code>	Tableau de n valeurs uniformément réparties de a à b , bornes incluses.
<code>np.zeros([n, m])</code>	Tableau nul à n lignes et m colonnes.
<code>np.zeros(n)</code>	Tableau nul à une dimension, de longueur n .
<code>np.arange(a, b, eps)</code>	Tableau de valeurs de a à b exclu, avec un pas eps .
<code>np.shape(M)</code>	Dimensions de M sous forme d'un tuple.

II. Simulations probabilistes

Python

```
import numpy.random as rd
```

Les coefficients des tableaux simulés ci-dessous sont indépendants. L'argument `[q, r]` produit un tableau à q lignes et r colonnes ; l'argument `n` produit un tableau à une dimension de longueur n .

Commandes	Loi de chaque coefficient
<code>rd.exponential(a, [q, r]),</code> <code>rd.exponential(a, n)</code>	Loi exponentielle de taux $1/a$, pour $a > 0$.
<code>rd.normal(m, d, [q, r]),</code> <code>rd.normal(m, d, n)</code>	Loi normale de moyenne m et de variance d^2 , pour $d > 0$.
<code>rd.gamma(m, a, [q, r]),</code> <code>rd.gamma(m, a, n)</code>	Loi gamma de forme $m > 0$ et d'échelle $a > 0$.

La dernière loi est la loi de aG , où $G \hookrightarrow \gamma(m)$; son taux est $1/a$.

III. Graphiques

Python

```
import matplotlib.pyplot as plt
```

Commande	Rôle
<code>plt.plot(X, Y, options)</code>	Trace les points de coordonnées données par X et Y, selon options.
<code>plt.grid()</code>	Affiche le quadrillage.
<code>plt.show()</code>	Affiche le graphique.

L'option "`*k`" représente des étoiles noires et l'option "`.k`" des points noirs.