

HEC Paris – ESCP Europe

Mathématiques II

Mathématiques approfondies

Sujet adapté – 2025

Le sujet 2025 est présenté sans modification de contenu, selon les programmes en vigueur de la filière Mathématiques approfondies.

Table des matières

1	Première partie – Lois du khi-deux et convergence	2
	<i>Thèmes abordés : variables à densité ; intégrales généralisées ; suites ; convergences de variables ; estimations ; Python.</i>	2
2	Deuxième partie – Transformation de deux lois normales	4
	<i>Thèmes abordés : fonctions de plusieurs variables ; variables à densité ; intégrales généralisées.</i>	4
3	Troisième partie – Théorème de Cochran et estimation gaussienne	6
	<i>Thèmes abordés : espaces vectoriels ; algèbre bilinéaire ; variables à densité ; estimations.</i>	6
	Formulaire Python	8

1 Première partie – Lois du khi-deux et convergence

Thèmes abordés : variables à densité ; intégrales généralisées ; suites ; convergences de variables ; estimations ; Python.

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

On note, pour $x \in \mathbb{R}$,

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

On rappelle que la fonction Gamma est définie, pour $x > 0$, par

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

La loi gamma à un paramètre de forme $a > 0$ sera notée $\gamma(a)$; elle admet la densité $x \mapsto x^{a-1}e^{-x}/\Gamma(a)$ sur $]0, +\infty[$.

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 $\omega \in \Omega$. Pour $n \in \mathbb{N}^*$, on définit

$$S_n = \sum_{i=1}^n X_i^2, \quad W_n = \frac{1}{2} S_n.$$

La loi de S_n est appelée loi du khi-deux à n degrés de liberté et est notée $\chi^2(n)$.

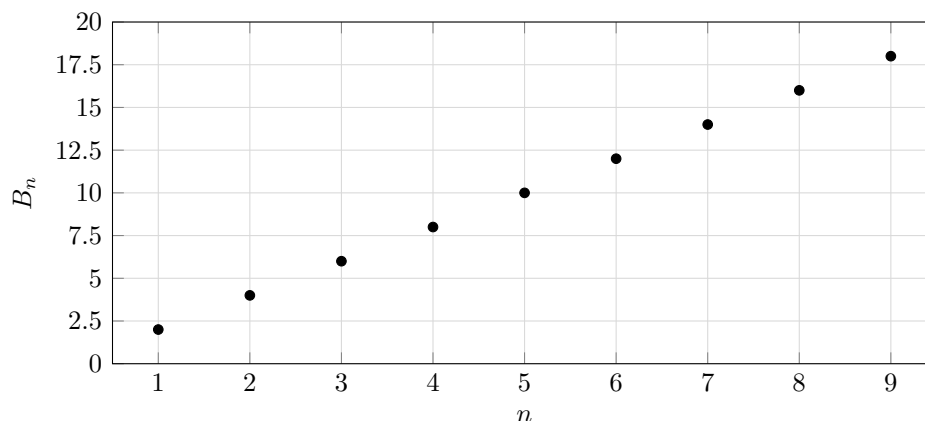
1. a) Montrer que, pour tout $n \in \mathbb{N}^*$, S_n admet une espérance et calculer $\mathbb{E}(S_n)$. On admet que $\mathbb{E}(S_n^2)$ existe pour tout $n \in \mathbb{N}^*$.
- b) Écrire une fonction Python `simul(n)` qui renvoie une réalisation de S_n lorsqu'on lui fournit un entier naturel non nul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

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

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

En déduire que $\mathbb{E}(1/S_n) = 1/(n-2)$.

- b) Déduire de la question précédente que, pour tout entier $n \geq 3$, $1/\sqrt{S_n}$ 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 est notée $\mathcal{T}(n)$. 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

$$\mathbb{P}(|T_n| \leq t_{n,\alpha}) = 1 - \alpha.$$

5. a) Montrer que, pour tout entier $n \geq 3$, $\mathbb{E}(T_n)$ existe et vaut 0.
 b) Montrer que, pour tout entier $n \geq 3$, $\mathbb{V}(T_n)$ existe et vaut $n/(n-2)$.
 c) En déduire que, pour tout entier $n \geq 3$,

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

6. Pour tout entier $n \geq 2$, on pose

$$u_n = \mathbb{E}\left(\frac{1}{\sqrt{W_n}}\right).$$

- a) Montrer que la suite $(u_n)_{n \geq 2}$ est décroissante et déterminer la valeur de u_2 .

b) Montrer que, pour tout entier $n \geq 2$,

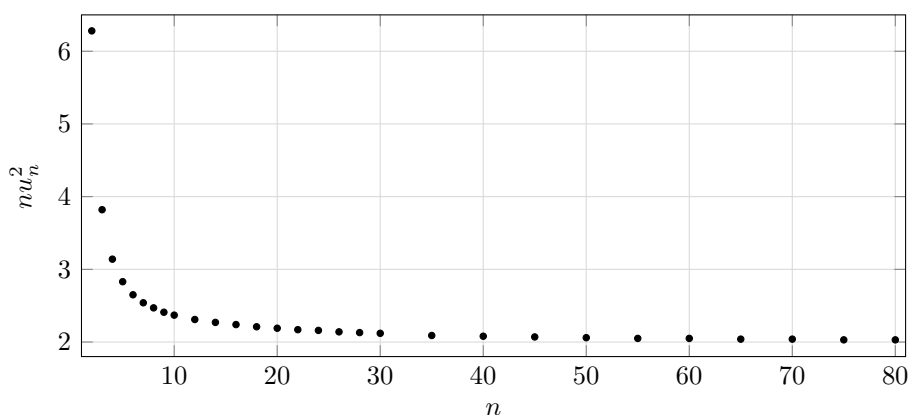
$$u_{n+1}u_n = \frac{2}{n-1}.$$

c) Écrire une fonction Python d'en-tête `suite_u(n)` qui renvoie sous la forme d'un tableau NumPy les $n-1$ valeurs $u_2, u_3, \dots, u_n$.

d) On exécute ensuite le programme suivant :

```
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 $(u_n)_{n \geq 2}$?

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 .

2 Deuxième partie – Transformation de deux lois normales

Thèmes abordés : fonctions de plusieurs variables ; variables à densité ; intégrales généralisées.

Dans cette partie, on démontre le résultat suivant.

Théorème. Soient X, Y deux variables aléatoires indépendantes, chacune de loi normale $\mathcal{N}(0, 1)$. Les variables aléatoires

$$\frac{X+Y}{\sqrt{2}} \quad \text{et} \quad \frac{X-Y}{\sqrt{2}}$$

sont alors indépendantes et de loi normale $\mathcal{N}(0, 1)$.

Pour la preuve, on utilisera sans démonstration le résultat suivant : si X, Y sont indépendantes, de densités respectives f, g , et si $A \subset \mathbb{R}^2$ est une partie fermée, alors

$$\mathbb{P}((X, Y) \in A) = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} \mathbf{1}_A(x, y) f(x) dx \right) g(y) dy,$$

où $\mathbf{1}_A(x, y)$ vaut 1 si $(x, y) \in A$ et 0 sinon.

Soient $a, b \in \mathbb{R}$ fixés jusqu'à la fin de cette partie. On considère

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

Soient X et Y deux variables aléatoires indépendantes, chacune de loi normale $\mathcal{N}(0, 1)$, et posons

$$W = \frac{X+Y}{\sqrt{2}}, \quad Z = \frac{X-Y}{\sqrt{2}}.$$

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

9. Soit y un réel tel que $y \geq d$.

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

b) En déduire que

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

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

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

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

b) Montrer que

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

12. Montrer que

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

On admet pour la suite que l'on peut changer l'ordre d'intégration dans cette formule, c'est-à-dire que

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

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

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

15. Montrer que

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

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

3 Troisième partie – Théorème de Cochran et estimation gaussienne

Thèmes abordés : espaces vectoriels ; algèbre bilinéaire ; variables à densité ; estimations.

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 $\bar{x} = (x_1 + \dots + x_n)/n$.

a) Montrer que

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

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 $\mathcal{N}(0, 1)$, et posons $X = (X_1, \dots, X_n)$. 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 que $Y_k(\omega) \neq 0$ pour tout $\omega \in \Omega$ et tout k .

Dans la suite, on utilisera sans preuve le théorème de Cochran : 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 théorème de Cochran.

b) Soit $(\beta_1, \beta_2) \in \mathbb{R}^2 \setminus \{(0, 0)\}$. Montrer que les variables aléatoires

$$R_1 = X_1 + X_2, \quad R_2 = \beta_1 X_1 + \beta_2 X_2$$

sont indépendantes si et seulement si $\text{Cov}(R_1, R_2) = 0$.

19. On introduit les variables aléatoires

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

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

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

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

On suppose pour la fin de ce 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 que $V(\omega) \neq 0$ pour tout $\omega \in \Omega$.

20. a) Pour $i \in \{1, \dots, n\}$, on pose $X_i = (Z_i - \mu)/\sigma$. On pose de plus

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

Exprimer V en fonction de U et σ .

- b) En déduire que $V/(n-1)$ est un estimateur sans biais de σ^2 .

- c) Montrer que

$$\frac{\bar{Z} - \mu}{\sqrt{V/(n(n-1))}}$$

suit la loi de Student $\mathcal{T}(n-1)$.

21. Soit $\alpha \in]0, 1[$ et soit $t_{n-1, \alpha}$ le réel tel que

$$\mathbb{P}(|T| \leq t_{n-1, \alpha}) = 1 - \alpha \quad \text{lorsque } T \sim \mathcal{T}(n-1).$$

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.

Formulaire Python

`numpy` est importé sous le nom `np`, `numpy.random` sous le nom `rd`, et `matplotlib.pyplot` sous le nom `plt`.

Instruction	Rôle
<code>np.pi</code>	Renvoie une valeur approchée de π .
<code>np.linspace(a,b,n)</code>	Crée une matrice ligne de n valeurs uniformément réparties entre a et b (inclus).
<code>np.zeros([n,m])</code>	Crée la matrice nulle de taille $n \times m$.
<code>np.zeros(n)</code>	Crée la matrice ligne nulle de taille n .
<code>np.arange(a,b,eps)</code>	Renvoie la liste des flottants de a à b (non compris), de pas constant eps .
<code>np.shape(M)</code>	Donne la taille de la matrice M sous forme d'un tuple.

Instruction	Rôle
<code>rd.exponential(a,[q,r])</code> <code>rd.exponential(a,n)</code>	Simule une réalisation d'une matrice (respectivement d'un vecteur) aléatoire dont les coefficients sont indépendants et suivent la loi exponentielle de paramètre $1/a$.
<code>rd.normal(m,d,[q,r])</code> <code>rd.normal(m,d,n)</code>	Simule une réalisation d'une matrice (respectivement d'un vecteur) aléatoire dont les coefficients sont indépendants et suivent la loi $\mathcal{N}(m, d^2)$.
<code>rd.gamma(m,a,[q,r])</code> <code>rd.gamma(m,a,n)</code>	Simule une réalisation d'une matrice (respectivement d'un vecteur) aléatoire dont les coefficients sont indépendants et suivent la loi $\Gamma(m, a)$.

Instruction	Rôle
<code>plt.plot(X,Y,options)</code>	Génère la courbe des points définis par les listes X et Y suivant les options graphiques indiquées.
<code>plt.grid()</code>	Affiche le quadrillage.
<code>plt.show()</code>	Affiche le graphique.