

HEC Paris – ESCP Business School

Mathématiques II

Mathématiques approfondies

Sujet adapté – 2021

Les parties du sujet contenant des éléments n'étant plus au programme depuis 2022 ont été adaptées pour convenir aux nouveaux programmes de la filière Mathématiques approfondies en vigueur depuis 2023.

Table des matières

1	Partie I – Polynômes factoriels	2
	<i>Thèmes abordés : espaces vectoriels ; applications linéaires ; Python.</i>	2
2	Partie II – Quelques propriétés de la loi de Poisson	3
	<i>Thèmes abordés : probabilités discrètes ; estimations.</i>	3
3	Partie III – Estimation ponctuelle de fonctions du paramètre θ	4
	<i>Thèmes abordés : probabilités discrètes ; estimations ; Python.</i>	4
4	Partie IV – Le cas $\varphi(\theta) = \theta$	6
	<i>Thèmes abordés : probabilités discrètes ; convergences de variables ; estimations.</i>	6

Problème – Polynômes factoriels et estimation pour la loi de Poisson

Partie I – Polynômes factoriels

Thèmes abordés : espaces vectoriels ; applications linéaires ; Python.

On note F , l'espace vectoriel des fonctions polynomiales à coefficients réels et, pour tout entier naturel r , F_r le sous-espace vectoriel des fonctions polynomiales à coefficients réels de degré inférieur ou égal à r . On note U_k la fonction $x \mapsto x^k$, avec la convention $x^0 = 1$; la base canonique de F_r est notée $(U_0, U_1, \dots, U_r)$.

1. Soit r un entier naturel. On considère une famille $(Q_0, Q_1, \dots, Q_r)$ de fonctions polynomiales de degrés respectifs $d_0, d_1, \dots, d_r$ avec $d_0 < d_1 < \dots < d_r$.
 - a) On suppose qu'il existe des réels $\lambda_0, \lambda_1, \dots, \lambda_r$ non tous nuls tels que $\lambda_0 Q_0 + \lambda_1 Q_1 + \dots + \lambda_r Q_r = 0$. En considérant $m = \max\{k \in \{0, \dots, r\} : \lambda_k \neq 0\}$, démontrer que l'hypothèse précédente est absurde. Qu'a-t-on ainsi démontré ?
 - b) À quelle condition la famille $(Q_0, Q_1, \dots, Q_r)$ est-elle une base de F_r ? On précisera s'il s'agit d'une condition nécessaire, suffisante ou nécessaire et suffisante.
2. Pour tout entier naturel r , le réel « x puissance r descendante » est noté $x^{\underline{r}}$ et défini par

$$\forall x \in \mathbb{R}, \quad x^{\underline{r}} = x(x-1) \cdots (x-r+1) = \prod_{k=0}^{r-1} (x-k),$$

avec la convention $x^{\underline{0}} = 1$. On pose donc, pour tout $r \in \mathbb{N}$, $V_r : x \mapsto x^{\underline{r}}$.

- a) Quelles sont les racines de V_r ?
- b) Démontrer que la famille $(V_0, V_1, \dots, V_r)$ est une base de F_r .
- c) Démontrer que, pour tout entier $r \geq 2$,

$$x^{\underline{r}} \underset{x \rightarrow +\infty}{=} x^r - \frac{r(r-1)}{2} x^{r-1} + o(x^{r-1}).$$

On pourra raisonner par récurrence sur r .

3. Ici, l'entier $r \geq 1$ est fixé et on compare la famille $(V_k)_{0 \leq k \leq r}$ à la base canonique $(U_k)_{0 \leq k \leq r}$ de F_r .

- a) Démontrer qu'il existe une unique famille $(\sigma(r, k))_{0 \leq k \leq r}$ de nombres réels telle que

$$U_r = \sum_{k=0}^r \sigma(r, k) V_k.$$

- b) Établir les relations suivantes :

$$\forall r \in \mathbb{N}^*, \quad \sigma(r, 0) = 0, \quad \sigma(r, 1) = \sigma(r, r) = 1, \quad \sigma(r, r-1) = \frac{r(r-1)}{2},$$

et, pour tout $r \in \{2, 3, \dots\}$, $\sigma(r, 2) = 2^{r-1} - 1$.

- c) Démontrer que, pour tout $r \in \mathbb{N}^*$ et tout $k \in \{1, \dots, r\}$,

$$\sigma(r+1, k) = \sigma(r, k-1) + k\sigma(r, k).$$

- d) En déduire que $\sigma(r, k)$ est un entier naturel non nul pour tout $r \in \mathbb{N}^*$ et tout $k \in \{1, \dots, r\}$.
- e) Écrire un code Python qui affiche les listes $(\sigma(r, k))_{1 \leq k \leq r}$ pour r variant de 2 à 5.
4. a) Démontrer que, pour tout entier naturel r , il existe une unique famille $(s(r, k))_{0 \leq k \leq r}$ de nombres réels telle que

$$\forall x \in \mathbb{R}, \quad x^r = \sum_{k=0}^r s(r, k) x^k.$$

- b) Soit $r \in \mathbb{N}^*$. Démontrer que, pour tout $k \in \{1, \dots, r\}$,

$$s(r+1, k) = s(r, k-1) - r s(r, k).$$

En déduire la valeur de $s(r, 1)$.

- c) Déduire de 4.b) que, pour tout $r \in \mathbb{N}^*$ et tout $k \in \{1, \dots, r\}$, le signe de $s(r, k)$ est celui de $(-1)^{r+k}$.
- d) Démontrer que $\sigma(r, r) s(r, r) = 1$ et que, pour tout $\ell \in \{0, \dots, r-1\}$,

$$\sum_{k=\ell}^r \sigma(r, k) s(k, \ell) = 0.$$

- e) Calculer $s(r, r)$ pour tout $r \in \mathbb{N}^*$ et $s(r, r-1)$ pour tout entier $r \geq 2$.

Partie II – Quelques propriétés de la loi de Poisson

Thèmes abordés : probabilités discrètes ; estimations.

Sous réserve d'existence, on note respectivement $\mathbb{E}(A)$ et $\mathbb{V}(A)$ l'espérance et la variance d'une variable aléatoire A , et $\text{Cov}(A, B)$ la covariance de deux variables aléatoires discrètes A et B .

Dans cette partie, X est une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, qui suit une loi de Poisson de paramètre $\theta > 0$. Pour tout entier $r \geq 1$, on pose $X^r = X(X-1) \cdots (X-r+1)$, avec la convention $X^0 = X^0 = 1$. Avec les suites définies en 3.a) et 4.a), on a, pour tout entier $r \geq 1$,

$$X^r = \sum_{k=0}^r \sigma(r, k) X^k \quad \text{et} \quad X^r = \sum_{k=0}^r s(r, k) X^k.$$

On admet ces deux résultats sans démonstration.

5. a) Rappeler sans démonstration les valeurs respectives de $\mathbb{E}(X)$, $\mathbb{V}(X)$ et $\mathbb{E}(X^2)$.
- b) Exprimer X , X^2 , X^3 et X^4 en fonction des variables aléatoires X^1 , X^2 , X^3 et X^4 .
- c) Démontrer que la variable aléatoire X admet des moments de tous ordres.
6. a) Justifier que, pour tout entier $r \geq 1$, la variable aléatoire X^r admet des moments de tous ordres.
- b) Pour tout entier $r \geq 1$, exprimer $\mathbb{E}(X^r)$ en fonction de r et de θ .
- c) Calculer $\mathbb{E}(X^3)$ et $\mathbb{E}(X^4)$ en fonction de θ .

7. Pour tout entier naturel k et pour tout réel $\theta > 0$, on pose

$$f(\theta, k) = \frac{\theta^k}{k!} e^{-\theta} \quad \text{et} \quad g(\theta, k) = \ln(f(\theta, k)).$$

- Pour tout entier $k \geq 0$, calculer l'expression de la dérivée partielle $\partial_1 g(\theta, k)$ et exprimer la variable aléatoire $\partial_1 g(\theta, X)$ en fonction de X et de θ .
- Vérifier que $XX^r = X^{r+1} + rX^r$ pour tout entier $r \geq 1$. En déduire que $\text{Cov}(X, X^r) = r\theta^r$.
- Calculer $\text{Cov}(\partial_1 g(\theta, X), X^r)$ et en déduire l'inégalité

$$\forall \theta > 0, \forall r \in \mathbb{N}^*, \quad \mathbb{V}(X^r) \geq r^2 \theta^{2r-1}.$$

Partie III – Estimation ponctuelle de fonctions du paramètre θ

Thèmes abordés : probabilités discrètes ; estimations ; Python.

Le contexte et les notations sont ceux de la partie 2. On suppose que le paramètre inconnu $\theta \in \mathbb{R}_+^*$ et on cherche à estimer $\varphi(\theta)$, où φ est une fonction dérivable sur $\mathbb{R}_+^*$. Pour $n \in \mathbb{N}^*$, on considère dans toute cette partie un n -échantillon $(X_1, \dots, X_n)$ de variables aléatoires mutuellement indépendantes qui, comme X , suivent toutes la loi de Poisson de paramètre θ . Pour tout $\theta > 0$ et tout $(k_1, \dots, k_n) \in \mathbb{N}^n$, on pose

$$F(\theta, k_1, \dots, k_n) = \mathbb{P} \left(\bigcap_{i=1}^n [X_i = k_i] \right) = \prod_{i=1}^n f(\theta, k_i),$$

$$G(\theta, k_1, \dots, k_n) = \ln(F(\theta, k_1, \dots, k_n)).$$

- Démontrer que les fonctions $\theta \mapsto F(\theta, k_1, \dots, k_n)$ et $\theta \mapsto G(\theta, k_1, \dots, k_n)$ sont dérivables sur $\mathbb{R}_+^*$ et calculer les dérivées partielles $\partial_1 F$ et $\partial_1 G$.
 - Pour tout $\theta > 0$, on pose $Z_\theta = \partial_1 G(\theta, X_1, \dots, X_n)$. Démontrer que Z_θ est centrée et admet une variance strictement positive, notée $I(\theta)$, que l'on calculera.

On rappelle que s'il existe n séries absolument convergentes $\sum_{k_i \in \mathbb{N}} v_{i,k_i}$ telles que, pour tout $(k_1, \dots, k_n) \in \mathbb{N}^n$, $|u_{k_1, \dots, k_n}| \leq |v_{1,k_1}| \times \dots \times |v_{n,k_n}|$, alors la série $\sum_{(k_1, \dots, k_n) \in \mathbb{N}^n} u_{k_1, \dots, k_n}$ est dite absolument convergente et sa somme est bien définie. On rappelle l'énoncé de la formule de transfert : si la série $\sum_{(k_1, \dots, k_n) \in \mathbb{N}^n} u_\theta(k_1, \dots, k_n) F(\theta, k_1, \dots, k_n)$ est absolument convergente, alors la variable aléatoire discrète $U_\theta = u_\theta(X_1, \dots, X_n)$ est d'espérance finie et

$$\mathbb{E}(U_\theta) = \sum_{(k_1, \dots, k_n) \in \mathbb{N}^n} u_\theta(k_1, \dots, k_n) F(\theta, k_1, \dots, k_n).$$

Soit $t : \mathbb{N}^n \rightarrow \mathbb{R}$ une application indépendante de θ . On peut considérer $T = t(X_1, \dots, X_n)$ comme un estimateur de $\varphi(\theta)$. On dira que T est un *estimateur régulier* de $\varphi(\theta)$ lorsque les trois conditions suivantes sont satisfaites :

$$\begin{aligned} \mathbb{E}(T) &= \varphi(\theta) \quad (R_1), & \mathbb{V}(T) &\text{existe} \quad (R_2), \\ \forall \theta > 0, \quad \varphi'(\theta) &= \sum_{(k_1, \dots, k_n) \in \mathbb{N}^n} t(k_1, \dots, k_n) \partial_1 F(\theta, k_1, \dots, k_n) \quad (R_3). \end{aligned}$$

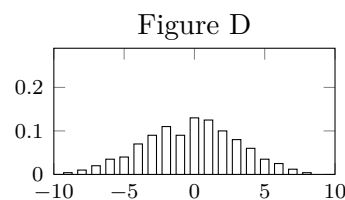
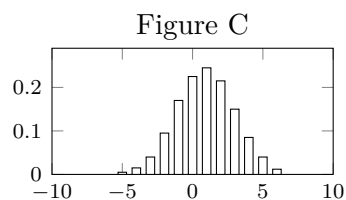
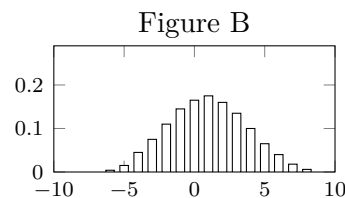
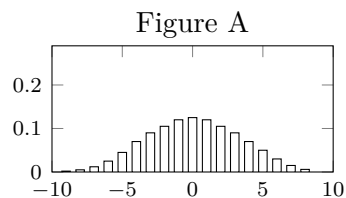
La condition (R_3) sous-entend que le second membre est la somme d'une série absolument convergente.

9. Dans cette question, on suppose que T est un estimateur régulier de $\varphi(\theta)$.
- La variable aléatoire T est-elle un estimateur sans biais de $\varphi(\theta)$?
 - Pourquoi la condition (R_3) n'est-elle pas une conséquence directe de la condition (R_1) ?
10. Soit T un estimateur régulier du paramètre $\varphi(\theta)$.
- Établir les égalités suivantes : $\forall \theta > 0, \varphi'(\theta) = \mathbb{E}(TZ_\theta) = \text{Cov}(T, Z_\theta)$.
 - En déduire l'inégalité

$$\forall \theta > 0, \quad \mathbb{V}(T) \geq \frac{(\varphi'(\theta))^2}{I(\theta)},$$
 où $I(\theta)$ a été définie à la question 8.b).
11. On cherche à simuler un échantillon de N réalisations de Z_θ pour différents couples (n, θ) .
- Compléter le code Python suivant en justifiant votre réponse. Ici, `numpy` est importé sous le nom `np` et N est fixé.

```
def z_th(n, theta):
    X = np.random.poisson(theta, size=(n,N))
    ech = (np.sum(X, axis=.....) - n*theta)/theta
    return ech
```

- À l'aide d'un histogramme, on a représenté les échantillons obtenus pour les couples $(n, \theta) = (10, 4), (20, 4), (40, 4)$ et $(50, 5)$. À quels couples correspondent les figures suivantes ? On pourra admettre que $I(\theta) = n/\theta$.



Partie IV – Le cas $\varphi(\theta) = \theta$

Thèmes abordés : probabilités discrètes ; convergences de variables ; estimations.

Le contexte et les notations sont ceux des parties 2 et 3. Dans cette partie, on compare deux estimateurs du paramètre inconnu θ .

12. Pour tout entier $n \geq 1$, on pose $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$.

- a) Démontrer que $\bar{X}_n$ est un estimateur régulier du paramètre θ .
- b) Que devient l'inégalité de 10.b) ?

On dit qu'un estimateur régulier de θ est *efficace* lorsque sa variance est minimale parmi les estimateurs réguliers de θ .

13. Soit Y un estimateur régulier de θ . Pour tout réel α , on pose $\psi(\alpha) = \bar{X}_n + \alpha(Y - \bar{X}_n)$.

- a) Vérifier que $\psi(\alpha)$ est un estimateur régulier de θ pour tout $\alpha \in \mathbb{R}$.
- b) En déduire que $\text{Cov}(\bar{X}_n, Y) = \frac{\theta}{n}$.
- c) Exprimer $\mathbb{V}(Y - \bar{X}_n)$ en fonction de $\mathbb{V}(Y)$ et de $\mathbb{V}(\bar{X}_n)$. En déduire qu'un estimateur efficace de θ est presque sûrement unique.

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

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

- a) Exprimer W_n en fonction de $\sum_{i=1}^n X_i^2$ et de $\bar{X}_n^2$.
- b) Démontrer que W_n est un estimateur sans biais de θ .
- c) Démontrer que W_n admet une variance (qu'on ne cherchera pas à calculer).
- d) Étudier la convergence des deux suites d'estimateurs $(\bar{X}_n)_{n \geq 1}$ et $(W_n)_{n \geq 2}$ du paramètre inconnu θ . On pourra démontrer que si une suite réelle (a_n) converge vers $a \in \mathbb{R}$ et si deux suites (Y_n) et (Z_n) de variables aléatoires convergent en probabilité vers les réels y et z , alors $(a_n(Y_n - Z_n))$ converge en probabilité vers $a(y - z)$.

15. On simule des échantillons de N réalisations des estimateurs $\bar{X}_n$, W_n et $W'_n = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2$, avec $n = 50$. En comparant les figures suivantes, relier chaque histogramme à l'estimateur qui lui correspond.

Figure E

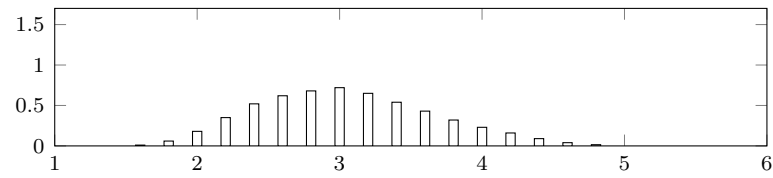


Figure F

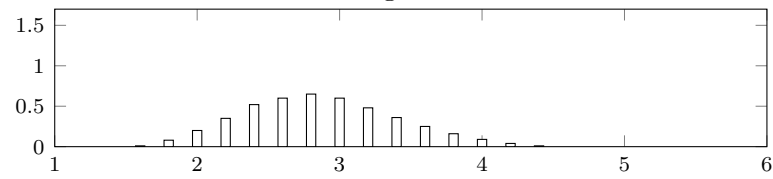


Figure G

