



PRÉPASIMPLE

prepasimple.fr

HEC / ESCP

Mathématiques 2

2021

Sujet de mathématiques approfondies

Mathématiques approfondies — Prépa ECG

Annales de concours · Sujet

Partie 1 – Polynômes factoriels

On note F l'espace vectoriel des fonctions polynomiales à coefficients réels et, pour tout entier naturel r , on note F_r le sous-espace 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 $x^0 = 1$, de telle sorte que 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 $d_0 < d_1 < \dots < d_r$.

1.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 $m = \max\{k \in \llbracket 0, r \rrbracket : \lambda_k \neq 0\}$, démontrer que l'hypothèse précédente est absurde. Qu'a-t-on ainsi démontré?

1.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^r et défini par

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

avec la convention $x^0 = 1$. On pose alors

$$\forall r \in \mathbb{N}, \quad V_r : x \mapsto x^r.$$

La fonction V_r appartient à F .

2.a) Quelles sont les racines de V_r ?

2.b) Démontrer que la famille $(V_0, V_1, \dots, V_r)$ est une base de F_r .

2.c) Démontrer que, pour tout entier $r \geq 2$,

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

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 l'espace F_r .

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

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

3.b) Établir les relations suivantes :

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

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

$$\forall r \in \mathbb{N}^*, \quad \sigma(r, r-1) = \frac{r(r-1)}{2}. \quad (3)$$

$$\forall r \in \llbracket 2, +\infty \rrbracket, \quad \sigma(r, 2) = 2^{r-1} - 1. \quad (4)$$

3.c) Démontrer que, pour tout $r \in \mathbb{N}^*$ et tout $k \in \llbracket 1, r \rrbracket$,

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

3.d) En déduire que $\sigma(r, k)$ est un entier naturel non nul pour tout $r \in \mathbb{N}^*$ et tout $k \in \llbracket 1, r \rrbracket$.

3.e) Écrire un code Python qui affiche les listes $(\sigma(r, k))_{1 \leq k \leq r}$ pour r variant de 2 à 5.

On pourra utiliser `numpy.ones((n, p))`, qui renvoie un tableau de taille (n, p) dont tous les coefficients valent 1. La commande `print(a)` affiche le tableau ou la liste `a`.

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

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

4.b) Soit $r \in \mathbb{N}^*$. Démontrer que

$$\forall k \in \llbracket 1, r \rrbracket, \quad s(r+1, k) = s(r, k-1) - rs(r, k).$$

En déduire la valeur de $s(r, 1)$.

4.c) Déduire de 4.b) que, pour tout $r \in \mathbb{N}^*$ et tout $k \in \llbracket 1, r \rrbracket$, le signe de $s(r, k)$ est celui de $(-1)^{r+k}$.

4.d) Démontrer que $\sigma(r, r)s(r, r) = 1$ et que

$$\forall \ell \in \llbracket 0, r-1 \rrbracket, \quad \sum_{k=\ell}^r \sigma(r, k)s(k, \ell) = 0.$$

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

Partie 2 – Quelques propriétés de la loi de Poisson

Sous réserve d'existence, on note respectivement $E(A)$ et $V(A)$ 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 et suivant la 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 coefficients définis aux questions 3.a) et 4.a), on a

$$\forall r \geq 1, \quad X^r = \sum_{k=0}^r \sigma(r, k) X^k \quad \text{et} \quad X^r = \sum_{k=0}^r s(r, k) X^k.$$

On admet ces deux résultats sans démonstration.

5.a) Rappeler sans démonstration les valeurs respectives de $E(X)$, $V(X)$ et $E(X^2)$.

5.b) Exprimer X , X^2 , X^3 et X^4 en fonction des variables aléatoires X^1 , X^2 , X^3 et X^4 .

5.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.

6.b) Pour tout entier $r \geq 1$, exprimer $E(X^r)$ en fonction de r et de θ .

6.c) Calculer $E(X^3)$ et $E(X^4)$ en fonction de θ .

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

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

7.a) Pour tout entier $k \geq 0$, calculer 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 .

7.b) Vérifier que $XX^r = X^{r+1} + rX^r$ pour tout entier $r \geq 1$. En déduire que $\text{Cov}(X, X^r) = r\theta^r$.

7.c) Calculer $\text{Cov}(\partial_1 g(\theta, X), X^r)$ et en déduire l'inégalité

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

Partie 3 – Estimation ponctuelle de fonctions du paramètre θ

Le contexte et les notations sont ceux de la partie 2. On suppose que le paramètre $\theta > 0$ est inconnu et on cherche à estimer $\varphi(\theta)$, où φ est une fonction dérivable sur $]0, +\infty[$.

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 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

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

et

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

8.a) 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 $]0, +\infty[$ et calculer les dérivées partielles $\partial_1 F(\theta, k_1, \dots, k_n)$ et $\partial_1 G(\theta, k_1, \dots, k_n)$.

8.b) 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.

Conventions pour les sommes utilisées ci-dessous : la notation

$$\sum_{(k_1, \dots, k_n) \in \mathbb{N}^n} u_{k_1, \dots, k_n}$$

désigne la limite, lorsqu'elle existe, des sommes finies portant sur $0 \leq k_1, \dots, k_n \leq m$, lorsque m tend vers $+\infty$.

On rappelle que, si les n séries

$$\sum_{k_i \geq 0} |v_{i, k_i}|$$

convergent et si

$$\forall (k_1, \dots, k_n) \in \mathbb{N}^n, \quad |u_{k_1, \dots, k_n}| \leq \prod_{i=1}^n |v_{i, k_i}|,$$

alors les sommes finies des $|u_{k_1, \dots, k_n}|$ sont majorées uniformément en m . On admet que les sommes finies des $u_{k_1, \dots, k_n}$ ont alors une limite, indépendante de l'ordre de sommation.

On rappelle la formule de transfert dans ce cadre : si la somme

$$\sum_{(k_1, \dots, k_n) \in \mathbb{N}^n} |u_{\theta}(k_1, \dots, k_n)| F(\theta, k_1, \dots, k_n)$$

est finie, alors la variable aléatoire $U_{\theta} = u_{\theta}(X_1, \dots, X_n)$ admet une espérance finie et

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

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 dit que T est un estimateur régulier de $\varphi(\theta)$ lorsque les trois conditions suivantes sont satisfaites, pour tout $\theta > 0$:

$$E(T) = \varphi(\theta). \quad (R_1)$$

$$V(T) \text{ existe.} \quad (R_2)$$

$$\varphi'(\theta) = \sum_{(k_1, \dots, k_n) \in \mathbb{N}^n} t(k_1, \dots, k_n) \partial_1 F(\theta, k_1, \dots, k_n). \quad (R_3)$$

La condition (R_3) suppose également que la somme des valeurs absolues des termes du second membre est finie.

9) Dans cette question, on suppose que T est un estimateur régulier de $\varphi(\theta)$.

9.a) La variable aléatoire T est-elle un estimateur sans biais de $\varphi(\theta)$?

9.b) Pourquoi la condition (R_3) n'est-elle pas une conséquence directe de la condition (R_1) ?

10) Soit T un estimateur régulier de $\varphi(\theta)$.

10.a) Établir les égalités suivantes :

$$\forall \theta > 0, \quad \varphi'(\theta) = E(TZ_\theta) = \text{Cov}(T, Z_\theta).$$

10.b) En déduire l'inégalité

$$\forall \theta > 0, \quad V(T) \geq \frac{(\varphi'(\theta))^2}{I(\theta)},$$

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) .

11.a) Compléter le code Python suivant en justifiant la réponse.

Python

```
import numpy as np

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

Pour un tableau M de taille (n, p) , `np.sum(M, axis=0)` renvoie le tableau de taille $(p,)$ contenant les sommes des colonnes, et `np.sum(M, axis=1)` renvoie le tableau de taille $(n,)$ contenant les sommes des lignes.

11.b) À l'aide de la commande `matplotlib.pyplot.hist`, avec `density=True`, on a tracé les histogrammes des 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 $I(\theta) = n/\theta$.

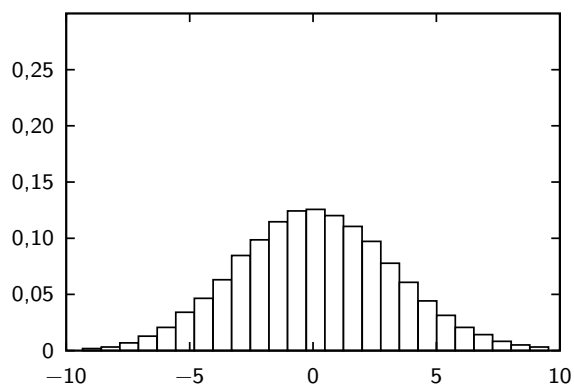


Figure A

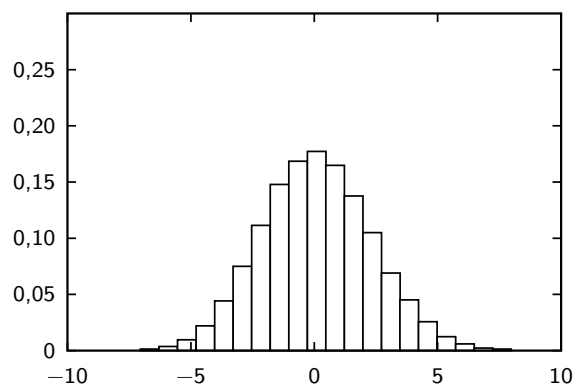


Figure B

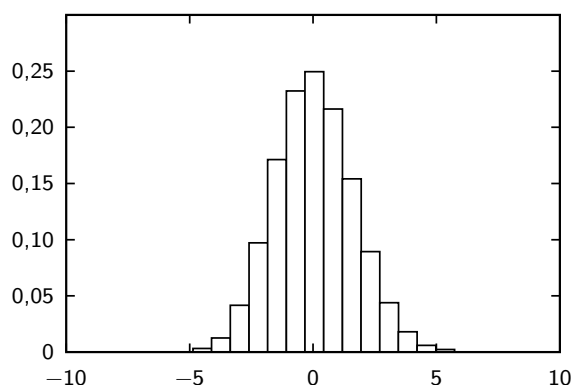


Figure C

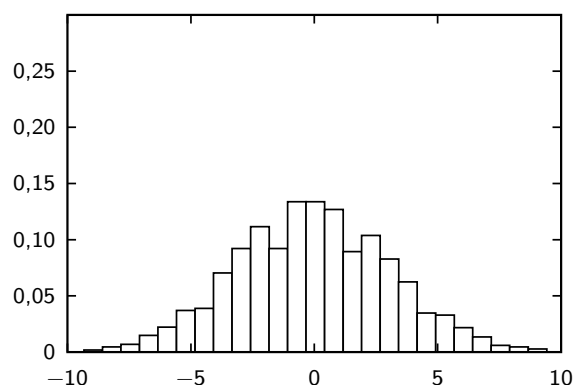


Figure D

12) Soit un entier $r \geq 1$. On suppose ici que $\varphi(\theta) = \theta^r$ et, pour tout entier $n \geq 1$, on pose

$$S_n = X_1 + \dots + X_n \quad \text{et} \quad M_{r,n} = \frac{S_n(S_n - 1) \dots (S_n - r + 1)}{n^r}.$$

12.a) Rappeler sans démonstration la loi de S_n ainsi que son espérance et sa variance.

12.b) Démontrer que $M_{r,n}$ est un estimateur régulier de θ^r .

Pour établir (R_3) , on admettra que la somme est absolument convergente.

En déduire que

$$\forall \theta > 0, \quad V(M_{r,n}) \geq \frac{r^2 \theta^{2r-1}}{n}.$$

12.c) Dans cette question, on suppose que $r = 2$. Calculer la variance de $M_{2,n}$ et

démontrer que la suite $(M_{2,n})_{n \geq 1}$ d'estimateurs de θ^2 est convergente.

12.d) Pour un entier $r \geq 1$ quelconque, la suite $(M_{r,n})_{n \geq 1}$ d'estimateurs de θ^r est-elle convergente ? On pourra commencer par calculer, en fonction de l'entier $k \in \mathbb{N}^*$, un équivalent de $E(S_n^k)$ lorsque n tend vers $+\infty$.

Partie 4 – Le cas $\varphi(\theta) = \theta$

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

13) Pour tout entier $n \geq 1$, on pose

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i.$$

13.a) Démontrer que $\bar{X}_n$ est un estimateur régulier du paramètre θ .

13.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 .

14) Soit Y un estimateur régulier de θ . Pour tout réel α , on pose

$$\psi(\alpha) = \bar{X}_n + \alpha(Y - \bar{X}_n).$$

14.a) Vérifier que $\psi(\alpha)$ est un estimateur régulier de θ pour tout $\alpha \in \mathbb{R}$.

14.b) En déduire que

$$\text{Cov}(\bar{X}_n, Y) = \frac{\theta}{n}.$$

14.c) Exprimer $V(Y - \bar{X}_n)$ en fonction de $V(Y)$ et de $V(\bar{X}_n)$. En déduire qu'un estimateur efficace de θ est presque sûrement unique.

15) Pour tout entier $n \geq 2$, on pose

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

15.a) Exprimer W_n en fonction de

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

et de $\bar{X}_n^2$.

15.b) Démontrer que W_n est un estimateur sans biais de θ .

15.c) Démontrer que W_n admet une variance, qu'on ne cherchera pas à calculer.

15.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 le résultat suivant : si une suite réelle (a_n) converge vers $a \in \mathbb{R}$ et si deux suites de variables aléatoires (Y_n) et (Z_n) convergent en probabilité vers les réels y et z respectivement, alors $(a_n(Y_n - Z_n))$ converge en probabilité vers $a(y - z)$.

16) 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.

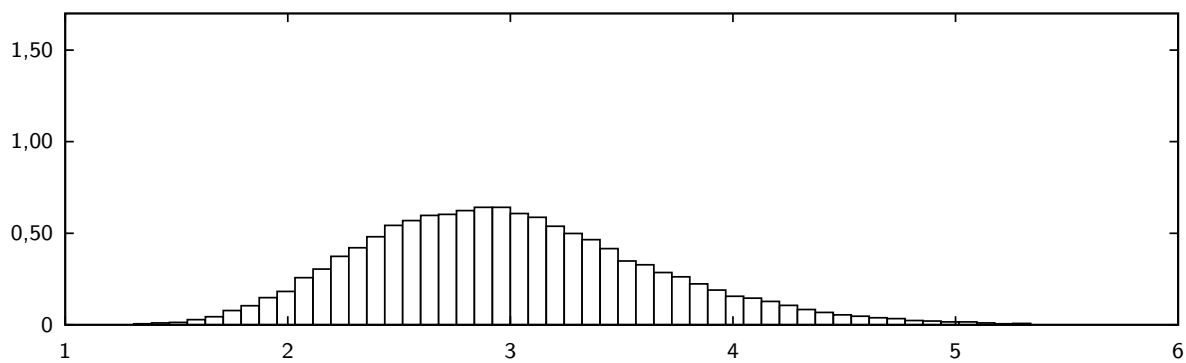


Figure E

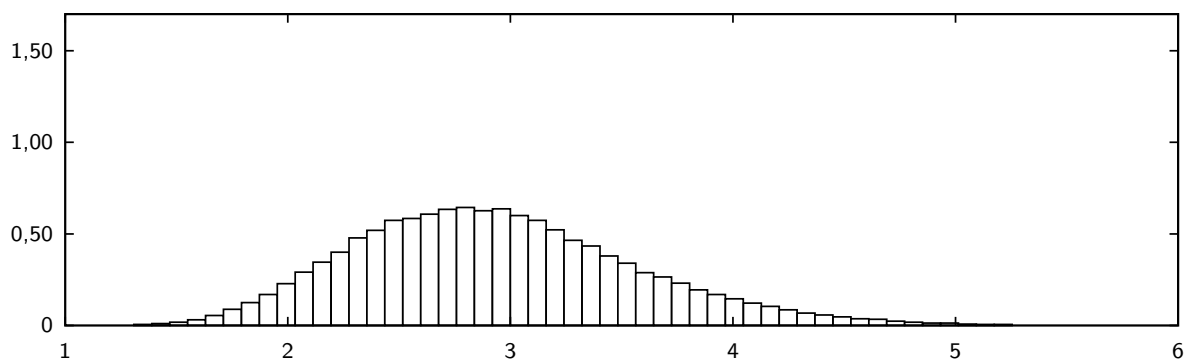


Figure F

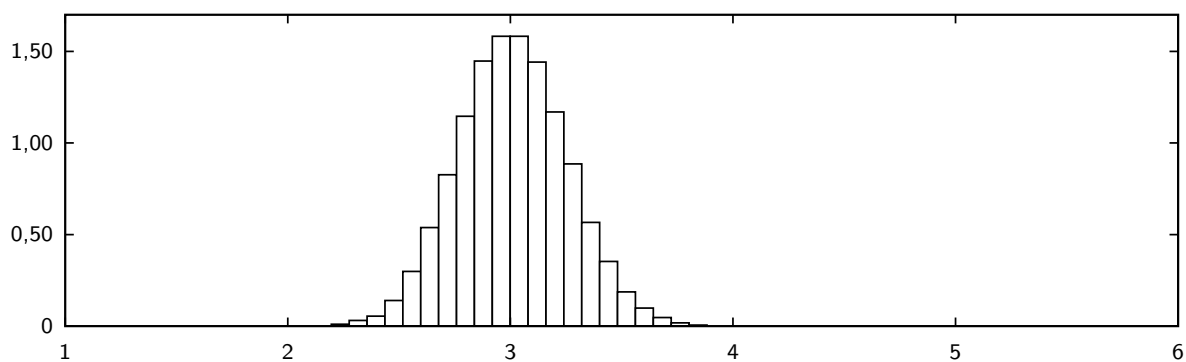


Figure G