



PRÉPASIMPLE

prepasimple.fr

HEC / ESCP

Mathématiques 2

2021

Corrigé de mathématiques approfondies

Mathématiques approfondies — Prépa ECG

Annales de concours · Corrigé

Partie 1 – Polynômes factoriels

On utilise la notation

$$x^r = \prod_{j=0}^{r-1} (x - j),$$

avec $x^0 = 1$. Les fonctions U_r et V_r sont respectivement $x \mapsto x^r$ et $x \mapsto x^r$.

1.a)

On raisonne par l'absurde en examinant le coefficient du terme de plus haut degré.

Supposons qu'une relation

$$\sum_{k=0}^r \lambda_k Q_k = 0$$

existe avec au moins un coefficient non nul. L'ensemble des indices concernés est fini et non vide; son maximum m est donc défini.

Si $a_m \neq 0$ est le coefficient dominant de Q_m , alors $\lambda_m Q_m$ a pour degré d_m et pour coefficient dominant $\lambda_m a_m \neq 0$. Pour $k < m$, on a $d_k < d_m$, donc les autres termes ne contribuent pas au coefficient de degré d_m .

Le coefficient de degré d_m de

$$\sum_{k=0}^m \lambda_k Q_k$$

vaut donc $\lambda_m a_m \neq 0$. Cette somme ne peut pas être le polynôme nul, ce qui contredit l'hypothèse.

La famille $(Q_0, \dots, Q_r)$ est ainsi libre dans F .

1.b)

Une condition nécessaire et suffisante est $d_r \leq r$. Avec la croissance stricte des degrés, elle équivaut à $d_k = k$ pour tout $k \in \llbracket 0, r \rrbracket$, et également à $d_r = r$.

En effet, si la famille est une base de F_r , tous ses éléments appartiennent à F_r , donc $d_r \leq r$.

Réciproquement, sous cette condition, les $r + 1$ polynômes appartiennent à F_r . Ils sont libres par 1.a), tandis que $\dim F_r = r + 1$. Ils forment donc une base de F_r .

Enfin, les degrés sont des entiers naturels strictement croissants : $d_k \geq k$ et $d_k \leq d_r - (r - k)$. Si $d_r \leq r$, ces deux inégalités imposent $d_k = k$.

2.a)

Pour $r = 0$, $V_0 = 1$ n'a aucune racine.

Pour $r \geq 1$, la factorisation

$$V_r(x) = \prod_{k=0}^{r-1} (x - k)$$

montre que ses racines sont exactement $0, 1, \dots, r - 1$. Chacune est simple, car les facteurs sont distincts et tous les autres facteurs sont non nuls au point considéré.

2.b)

Le polynôme V_k est de degré k et de coefficient dominant 1, y compris $V_0 = 1$, de degré 0.

La famille $(V_0, \dots, V_r)$ est donc libre par 1.a). Ses $r + 1$ éléments appartiennent à F_r , de dimension $r + 1$: elle est une base de F_r .

2.c)

Méthode 1 — Récurrence sur le nombre de facteurs

Pour $r = 2$, $x^2 = x^2 - x$, ce qui donne exactement la formule demandée.

Supposons, pour un entier $r \geq 2$, que

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

L'identité $x^{r+1} = (x-r)x^r$ donne alors

$$x^{r+1} = x^{r+1} - \left(r + \frac{r(r-1)}{2}\right)x^r + o(x^r).$$

Pour justifier le reste, si $h(x) = o(x^{r-1})$, alors

$$\frac{(x-r)h(x)}{x^r} = \left(1 - \frac{r}{x}\right) \frac{h(x)}{x^{r-1}} \rightarrow 0.$$

Les autres termes non écrits sont de degré au plus $r-1$ et sont donc également négligeables devant x^r . Or $r + r(r-1)/2 = r(r+1)/2$. La formule est ainsi établie au rang $r+1$, puis pour tout $r \geq 2$ par récurrence.

Méthode 2 — Lecture des deux premiers coefficients

Dans

$$\prod_{k=0}^{r-1} (x-k),$$

le terme de degré r provient du choix de x dans tous les facteurs et a pour coefficient 1.

Pour obtenir un terme de degré $r-1$, on choisit le terme constant dans exactement un facteur. Le coefficient correspondant est

$$-\sum_{k=0}^{r-1} k = -\frac{r(r-1)}{2}.$$

Le reste est un polynôme R de degré au plus $r-2$. Chaque monôme de $R(x)/x^{r-1}$ tend vers 0, donc $R(x) = o(x^{r-1})$.

Cette lecture des coefficients évite la récurrence ; elle est particulièrement utile pour les produits de facteurs du premier degré.

3.a)

Le polynôme U_r appartient à F_r , et $(V_0, \dots, V_r)$ en est une base par 2.b). Il possède donc une unique décomposition dans cette base :

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

Les coefficients $\sigma(r, k)$ sont ses coordonnées dans la base des polynômes factoriels.

3.b)

On part de l'identité, valable pour tout $x \in \mathbb{R}$,

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

Pour $r \geq 1$, l'évaluation en $x = 0$ donne $0 = \sigma(r, 0)$, car $0^0 = 1$ et $0^k = 0$ pour $k \geq 1$.

L'évaluation en $x = 1$ donne $1 = \sigma(r, 0) + \sigma(r, 1)$, tous les termes d'indice $k \geq 2$ étant nuls. Ainsi $\sigma(r, 1) = 1$.

La comparaison des coefficients de degré r donne $\sigma(r, r) = 1$, car seul V_r a ce degré et son coefficient dominant vaut 1.

Pour $r \geq 2$, le coefficient de degré $r-1$ du membre de gauche est nul. Grâce à 2.c), celui du membre de droite vaut

$$-\frac{r(r-1)}{2}\sigma(r, r) + \sigma(r, r-1).$$

Il en résulte $\sigma(r, r-1) = r(r-1)/2$. Pour $r = 1$, la même formule signifie $\sigma(1, 0) = 0$, déjà établi.

Enfin, pour $r \geq 2$, l'évaluation en $x = 2$ donne

$$2^r = 2\sigma(r, 1) + 2\sigma(r, 2),$$

car $2^k = 0$ pour $k \geq 3$. On obtient

$$\sigma(r, 2) = 2^{r-1} - 1.$$

Les quatre relations sont ainsi démontrées, avec leurs cas limites.

3.c)

Pour $k \geq 0$, on a $V_{k+1}(x) = (x - k)V_k(x)$, donc

$$xV_k(x) = V_{k+1}(x) + kV_k(x).$$

En multipliant la décomposition de U_r par x , il vient

$$U_{r+1} = \sum_{j=0}^r \sigma(r, j)V_{j+1} + \sum_{j=0}^r j\sigma(r, j)V_j.$$

Pour $1 \leq k \leq r$, le coefficient de V_k est $\sigma(r, k-1) + k\sigma(r, k)$. Par unicité de la décomposition dans la base $(V_0, \dots, V_{r+1})$,

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

Le coefficient de V_{r+1} est $\sigma(r, r) = 1$, en accord avec $\sigma(r+1, r+1) = 1$.

3.d)

On raisonne par récurrence sur r . Pour $r = 1$, $\sigma(1, 1) = 1$ est un entier strictement positif.

Supposons que tous les $\sigma(r, k)$, pour $1 \leq k \leq r$, soient des entiers strictement positifs.

On sait que $\sigma(r, 0) = 0$.

Pour $1 \leq k \leq r$, la relation de 3.c) exprime $\sigma(r+1, k)$ comme la somme de deux entiers naturels, dont $k\sigma(r, k)$ est strictement positif. Ce coefficient est donc un entier strictement positif. Le dernier coefficient vaut $\sigma(r+1, r+1) = 1$.

La propriété est établie à tous les rangs.

3.e)

Code adapté en Python. Le tableau `ligne` contient, au rang courant r , les coefficients d'indices 0 à r . On construit un nouveau tableau pour ne pas écraser les coefficients encore nécessaires.

Python

```
import numpy as np

ligne = np.array([0, 1], dtype=int)

for r in range(2, 6):
    nouvelle = np.zeros(r + 1, dtype=int)
    nouvelle[r] = 1
    for k in range(1, r):
        nouvelle[k] = ligne[k - 1] + k * ligne[k]
    ligne = nouvelle
    print(ligne[1:])
```

La boucle extérieure parcourt exactement 2, 3, 4, 5, car la borne supérieure de range est exclue. La boucle intérieure calcule les coefficients d'indices 1 à $r - 1$, puis le coefficient d'indice r est fixé à 1.

Ce code a été exécuté; ses sorties sont :

$$r = 2 : \quad (1, 1).$$

$$r = 3 : \quad (1, 3, 1).$$

$$r = 4 : \quad (1, 7, 6, 1).$$

$$r = 5 : \quad (1, 15, 25, 10, 1).$$

L'option `dtype=int` demande un tableau d'entiers; `ligne[1:]` extrait les coefficients à partir de l'indice 1.

4.a)

Le polynôme V_r appartient à F_r , dont $(U_0, \dots, U_r)$ est la base canonique. Il existe donc une unique famille de coordonnées $(s(r, k))_{0 \leq k \leq r}$ telle que

$$V_r = \sum_{k=0}^r s(r, k) U_k.$$

Autrement dit, pour tout réel x ,

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

En particulier, $s(0, 0) = 1$. Pour $r \geq 1$, l'évaluation en 0 donne $s(r, 0) = 0$, et le coefficient dominant donne $s(r, r) = 1$.

4.b)

L'identité $V_{r+1}(x) = (x - r)V_r(x)$ donne

$$V_{r+1}(x) = \sum_{j=0}^r s(r, j)x^{j+1} - r \sum_{j=0}^r s(r, j)x^j.$$

Pour $1 \leq k \leq r$, la comparaison des coefficients de x^k conduit à

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

Avec $k = 1$ et $s(r, 0) = 0$, on obtient $s(r+1, 1) = -rs(r, 1)$. Comme $s(1, 1) = 1$, une récurrence donne

$$s(r, 1) = (-1)^{r-1}(r-1)! \quad (r \geq 1).$$

4.c)

On démontre par récurrence que les coefficients concernés sont non nuls et ont le signe annoncé.

Pour $r = 1$, $s(1, 1) = 1$ a le signe de $(-1)^2$.

Supposons la propriété vraie au rang r . Pour $2 \leq k \leq r$, $s(r, k-1)$ a le signe de $(-1)^{r+k-1}$, tandis que $-rs(r, k)$ a le signe de $(-1)^{r+k+1}$. Ces deux signes sont identiques, car les exposants diffèrent de 2.

La relation de 4.b) additionne donc deux nombres non nuls de même signe. Leur somme est non nulle et a le signe de $(-1)^{r+1+k}$.

Pour $k = 1$, la formule de 4.b) donne directement le signe de $s(r+1, 1)$. Pour $k = r+1$, on a $s(r+1, r+1) = 1$, dont le signe est celui de $(-1)^{2r+2}$.

Tous les indices sont traités : $s(r, k)$ a le signe de $(-1)^{r+k}$ pour $1 \leq k \leq r$.

4.d)

On substitue la décomposition de V_k dans celle de U_r :

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

Toutes les sommes étant finies, leur regroupement ne pose aucune difficulté :

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

Par unicité des coefficients d'un polynôme, le coefficient de degré r est 1, et tous ceux de degrés inférieurs sont nuls. Ainsi

$$\sigma(r, r) s(r, r) = 1$$

et

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

4.e)

Méthode 1 — Coefficients du produit

Le coefficient dominant de V_r est 1, donc $s(r, r) = 1$ pour tout $r \geq 1$.

Pour $r \geq 2$, le coefficient de x^{r-1} est la somme des contributions obtenues en choisissant un seul terme constant dans le produit. Comme en 2.c),

$$s(r, r-1) = - \sum_{j=0}^{r-1} j = -\frac{r(r-1)}{2}.$$

Méthode 2 — Relations entre les deux changements de base

Par 3.b), $\sigma(r, r) = 1$. La première relation de 4.d) impose donc $s(r, r) = 1$.

Pour $r \geq 2$, prenons $\ell = r - 1$ dans la seconde relation de 4.d). Elle s'écrit

$$\sigma(r, r-1)s(r-1, r-1) + \sigma(r, r)s(r, r-1) = 0.$$

Les deux coefficients diagonaux valent 1 et $\sigma(r, r-1) = r(r-1)/2$. On retrouve $s(r, r-1) = -r(r-1)/2$.

La première approche lit directement le produit ; la seconde montre comment exploiter des relations de changement de base.

Partie 2 – Quelques propriétés de la loi de Poisson

Dans toute cette partie, $\theta > 0$ et $P(X = m) = e^{-\theta}\theta^m/m!$ pour $m \in \mathbb{N}$.

5.a)

Les moments usuels de la loi de Poisson donnent

$$E(X) = \theta, \quad V(X) = \theta.$$

La formule de Huygens donne alors

$$E(X^2) = V(X) + (E(X))^2 = \theta + \theta^2.$$

5.b)

En utilisant les lignes de coefficients calculées en 3.e), on obtient les identités polynomiales, donc les égalités entre variables aléatoires suivantes :

$$X = X^1.$$

$$X^2 = X^1 + X^2.$$

$$X^3 = X^1 + 3X^2 + X^3.$$

$$X^4 = X^1 + 7X^2 + 6X^3 + X^4.$$

5.c)

Comme $X \geq 0$, l'existence du moment d'ordre $p \geq 1$ revient à la convergence de

$$\sum_{m \geq 0} m^p P(X = m).$$

Méthode 1 — Domination de la fin de la série par une série géométrique

Pour $m \geq 1$, posons $a_m = m^p \theta^m / m! > 0$. On calcule

$$\frac{a_{m+1}}{a_m} = \frac{\theta}{m+1} \left(1 + \frac{1}{m}\right)^p \rightarrow 0.$$

Il existe donc un entier $m_0 \geq 1$ tel que, pour tout $m \geq m_0$, $a_{m+1} \leq a_m/2$. Par récurrence,

$$\forall j \geq 0, \quad a_{m_0+j} \leq a_{m_0} 2^{-j}.$$

La série de droite converge, donc la fin de

$$\sum a_m$$

converge par comparaison. Ajouter les premiers termes et multiplier par $e^{-\theta}$ préserve la convergence. Ainsi $E(X^p) < +\infty$.

On n'utilise ici aucun critère général de convergence fondé sur un quotient : la majoration géométrique est explicitement établie.

Méthode 2 — Domination par une exponentielle

Pour $x \geq 0$, la série exponentielle, dont tous les termes sont positifs, donne $e^x \geq x^p/p!$. D'autre part,

$$E(e^X) = e^{-\theta} \sum_{m=0}^{+\infty} \frac{(\theta e)^m}{m!} = \exp(\theta(e-1)) < +\infty.$$

Par domination,

$$E(X^p) \leq p! \exp(\theta(e-1)) < +\infty.$$

Cette seconde méthode fournit d'un seul coup une majoration pour tous les moments. Le moment d'ordre 0 vaut 1 : X admet donc des moments de tous ordres.

6.a)

Fixons $r \geq 1$ et un ordre $p \geq 1$. Pour $m \in \mathbb{N}$, $m^r = 0$ si $m < r$, tandis que, si $m \geq r$,

$$0 \leq m^r = m(m-1) \cdots (m-r+1) \leq m^r.$$

Par conséquent,

$$0 \leq (X^r)^p \leq X^{rp}.$$

Le majorant admet une espérance finie par 5.c). Par domination, X^r admet un moment d'ordre p . Comme p est quelconque, elle admet des moments de tous ordres.

6.b)

L'existence de l'espérance est établie en 6.a). La formule de transfert et l'annulation du produit pour $m < r$ donnent

$$E(X^r) = \sum_{m=r}^{+\infty} \frac{m!}{(m-r)!} e^{-\theta} \frac{\theta^m}{m!}.$$

Après simplification et changement d'indice $j = m - r$,

$$E(X^r) = \theta^r e^{-\theta} \sum_{j=0}^{+\infty} \frac{\theta^j}{j!} = \theta^r.$$

La formule vaut aussi pour $r = 0$, puisque $X^0 = 1$.

6.c)

Les variables intervenant ont une espérance finie par 5.c) et 6.a). En appliquant la linéarité aux identités de 5.b),

$$E(X^3) = \theta + 3\theta^2 + \theta^3$$

et

$$E(X^4) = \theta + 7\theta^2 + 6\theta^3 + \theta^4.$$

7.a)

Pour $k \in \mathbb{N}$ et $\theta > 0$, $f(\theta, k) > 0$, donc son logarithme est défini. On a

$$g(\theta, k) = k \ln \theta - \ln(k!) - \theta.$$

Cette fonction est dérivable par rapport à θ sur $]0, +\infty[$, et

$$\partial_1 g(\theta, k) = \frac{k}{\theta} - 1.$$

Ainsi

$$\partial_1 g(\theta, X) = \frac{X}{\theta} - 1 = \frac{X - \theta}{\theta}.$$

Cette variable admet une variance, vaut 0 en espérance et vérifie

$$V(\partial_1 g(\theta, X)) = \frac{V(X)}{\theta^2} = \frac{1}{\theta}.$$

7.b)

L'identité $X^{r+1} = (X - r)X^r$ donne, sans aucune division,

$$XX^r = X^{r+1} + rX^r.$$

Les moments nécessaires existent par 5.c) et 6.a), de même que l'espérance du produit, qui est dominé par X^{r+1} . On peut donc écrire

$$\text{Cov}(X, X^r) = E(XX^r) - E(X)E(X^r).$$

Par 6.b),

$$\text{Cov}(X, X^r) = \theta^{r+1} + r\theta^r - \theta\theta^r = r\theta^r.$$

7.c)

Par bilinéarité de la covariance et 7.b),

$$\text{Cov}\left(\frac{X}{\theta} - 1, X^r\right) = \frac{1}{\theta} \text{Cov}(X, X^r) = r\theta^{r-1}.$$

Les deux variables admettent des moments d'ordre 2. L'inégalité de Cauchy-Schwarz appliquée à leurs versions centrées donne

$$\text{Cov}(\partial_1 g(\theta, X), X^r)^2 \leq V(\partial_1 g(\theta, X)) V(X^r).$$

Ainsi

$$r^2 \theta^{2r-2} \leq \frac{1}{\theta} V(X^r).$$

Comme $\theta > 0$, on conclut

$$V(X^r) \geq r^2 \theta^{2r-1}.$$

Partie 3 – Estimation ponctuelle de fonctions du paramètre θ

On fixe $n \geq 1$ et on note

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

Les variables X_i sont mutuellement indépendantes et de loi $\mathcal{P}(\theta)$.

8.a)

Pour un n -uplet fixé $(k_1, \dots, k_n)$, posons

$$K = \sum_{i=1}^n k_i.$$

L'indépendance donne

$$F(\theta, k_1, \dots, k_n) = e^{-n\theta} \frac{\theta^K}{\prod_{i=1}^n k_i!} > 0.$$

Sur $]0, +\infty[$, cette fonction est dérivable et son logarithme l'est aussi. Plus précisément,

$$G(\theta, k_1, \dots, k_n) = -n\theta + K \ln \theta - \sum_{i=1}^n \ln(k_i!).$$

Il en résulte

$$\partial_1 G(\theta, k_1, \dots, k_n) = \frac{K}{\theta} - n$$

et

$$\partial_1 F(\theta, k_1, \dots, k_n) = F(\theta, k_1, \dots, k_n) \left(\frac{K}{\theta} - n \right).$$

On a donc l'identité utile $\partial_1 F = F \partial_1 G$.

8.b)

D'après 8.a),

$$Z_\theta = \frac{S_n - n\theta}{\theta}.$$

Les X_i possèdent des moments d'ordre 2, donc S_n et Z_θ aussi. Par linéarité,

$$E(S_n) = n\theta, \quad E(Z_\theta) = 0.$$

L'indépendance entraîne l'annulation des covariances entre X_i et X_j pour $i \neq j$. Ainsi

$$V(S_n) = \sum_{i=1}^n V(X_i) = n\theta.$$

On obtient

$$I(\theta) = V(Z_\theta) = \frac{n\theta}{\theta^2} = \frac{n}{\theta} > 0.$$

Question adaptée au programme actuel — Le rappel qui suit cette question dans le sujet définit les sommes par des limites de sommes finies sur des rectangles d'indices. La

domination par un produit de séries convergentes et la formule de transfert sont conservées; aucune théorie générale des familles sommables n'est nécessaire.

La majoration du rappel se vérifie déjà au niveau fini :

$$\sum_{k_1=0}^m \cdots \sum_{k_n=0}^m |u_{k_1, \dots, k_n}| \leq \prod_{i=1}^n \left(\sum_{k_i=0}^m |v_{i, k_i}| \right).$$

Le membre de droite est majoré par le produit des n sommes des séries de valeurs absolues. Les règles de passage à une somme indépendante de l'ordre et le transfert sont rappelés dans l'énoncé, et non déduits d'un échange automatique de limites.

9.a)

La condition (R_1) est exactement la propriété $E(T) = \varphi(\theta)$, pour tout $\theta > 0$. Le biais vaut donc

$$E(T) - \varphi(\theta) = 0.$$

Tout estimateur régulier est ainsi sans biais.

9.b)

La condition (R_1) et la formule de transfert donnent

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

Pour obtenir (R_3) en dérivant, il faudrait échanger une dérivée avec une somme infinie. La seule existence de cette somme pour chaque valeur de θ ne justifie pas cet échange.

La condition (R_3) impose donc une identité supplémentaire, ainsi que la convergence absolue de la somme des termes dérivés. Elle n'est pas une conséquence directe de (R_1) .

10.a)

Puisque T admet une variance par (R_2) et que Z_θ admet une variance par 8.b), Cauchy-Schwarz donne

$$E(|TZ_\theta|) \leq \sqrt{E(T^2)E(Z_\theta^2)} < +\infty.$$

La formule de transfert s'applique donc au produit. Avec $\partial_1 F = F\partial_1 G$ et (R_3) ,

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

Enfin, $E(Z_\theta) = 0$, donc

$$\text{Cov}(T, Z_\theta) = E(TZ_\theta) - E(T)E(Z_\theta) = \varphi'(\theta).$$

10.b)

Méthode 1 — Cauchy-Schwarz pour les variables centrées

Les moments d'ordre 2 existent, et 10.a) donne

$$(\varphi'(\theta))^2 = \text{Cov}(T, Z_\theta)^2 \leq V(T) V(Z_\theta) = V(T) I(\theta).$$

Comme $I(\theta) > 0$, on peut diviser et obtenir

$$V(T) \geq \frac{(\varphi'(\theta))^2}{I(\theta)}.$$

Méthode 2 — Non-négativité d'une variance

Pour tout réel c , la variable $T - cZ_\theta$ admet une variance et

$$0 \leq V(T - cZ_\theta) = V(T) - 2c\varphi'(\theta) + c^2I(\theta).$$

Prenons $c = \varphi'(\theta)/I(\theta)$, ce qui est permis par la stricte positivité de $I(\theta)$. Il vient

$$0 \leq V(T) - \frac{(\varphi'(\theta))^2}{I(\theta)},$$

soit la même borne. Cette seconde méthode indique aussi quelle combinaison linéaire intervient lorsque l'égalité est atteinte.

Cette borne a été démontrée ici avec les outils du programme ; aucun résultat général sur les bornes de variance en estimation n'est supposé connu.

11.a)

Code adapté en Python. Chaque colonne du tableau X représente un échantillon de n variables indépendantes de loi $\mathcal{P}(\theta)$. Il faut sommer les colonnes, donc utiliser `axis=0`.

Python

```
import numpy as np

def Z_th(n, theta, N):
    X = np.random.poisson(theta, size=(n, N))
    ech = (np.sum(X, axis=0) - n * theta) / theta
    return ech
```

Le tableau renvoyé est de taille $(N,)$: il contient N réalisations de $Z_\theta = (S_n - n\theta)/\theta$. Les paramètres doivent vérifier $n \geq 1$, $N \geq 1$ et $\theta > 0$.

Ce code a été exécuté avec chacun des quatre couples proposés et $N = 100\,000$. Les tailles des sorties ont été contrôlées, ainsi que leurs moyennes proches de 0 et leurs variances proches de n/θ . Ces observations numériques vérifient le comportement du programme ; elles ne remplacent pas la démonstration de 8.b).

11.b)

Question adaptée au programme actuel — Le nom de la commande de tracé est remplacé par la commande Python `hist`, avec `density=True` pour représenter une densité de fréquences. Les quatre histogrammes sont reproduits dans l'énoncé en SVG à partir de mesures sur l'original.

Toutes les distributions représentées sont centrées. Leurs variances théoriques sont :

Couple (n, θ)	Variance n/θ	Écart-type
(10, 4)	2,5	$\sqrt{2,5}$
(20, 4)	5	$\sqrt{5}$
(40, 4)	10	$\sqrt{10}$
(50, 5)	10	$\sqrt{10}$

La figure C est la plus concentrée autour de 0 : elle correspond à (10, 4). La figure B a une dispersion intermédiaire : elle correspond à (20, 4).

Les figures A et D sont les plus étalées et correspondent aux deux couples restants, (40, 4) et (50, 5).

L'association suggérée par l'aspect des barres est

A : (40, 4), B : (20, 4), C : (10, 4), D : (50, 5).

Il faut cependant distinguer cette lecture graphique d'une identification démontrée. Les couples (40, 4) et (50, 5) ont exactement la même moyenne et la même variance pour Z_θ . La variance seule ne permet donc pas de départager A et D.

Les supports diffèrent : les valeurs successives de Z_θ sont espacées de $1/4$ pour $(40, 4)$ et de $1/5$ pour $(50, 5)$. Cette différence peut contribuer aux irrégularités des barres, mais les observations simulées et les bornes exactes des classes ne sont pas fournies.

L'association A–D indiquée est plausible ; sa confirmation certaine nécessiterait ces données ou le code de tracé original. La conclusion certaine à partir des dispersions est l'attribution de B et C et l'identification de la paire de couples associée à A et D.

12.a)

Par stabilité de la loi de Poisson par somme de variables indépendantes,

$$S_n \sim \mathcal{P}(n\theta).$$

Ses moments usuels sont

$$E(S_n) = n\theta, \quad V(S_n) = n\theta.$$

12.b)

On vérifie successivement les trois conditions de régularité, sans dériver de somme infinie.

L'application

$$t(k_1, \dots, k_n) = \frac{K^r}{n^r}, \quad K = \sum_{i=1}^n k_i,$$

ne dépend pas de θ : $M_{r,n}$ est bien un estimateur.

Pour (R_1) , on applique 6.b) à $S_n \sim \mathcal{P}(n\theta)$:

$$E(M_{r,n}) = \frac{E(S_n^r)}{n^r} = \frac{(n\theta)^r}{n^r} = \theta^r.$$

Pour (R_2) , S_n^r admet des moments de tous ordres par 6.a), appliquée au paramètre $n\theta > 0$. En particulier, $M_{r,n}$ admet un moment d'ordre 2 et donc une variance.

Pour (R_3) , l'énoncé admet la convergence absolue de la somme concernée. Avec 8.a) et le transfert, son second membre est $E(M_{r,n}Z_\theta)$. Or $E(Z_\theta) = 0$, donc

$$E(M_{r,n}Z_\theta) = \frac{1}{n^r\theta} \text{Cov}(S_n^r, S_n).$$

La formule de 7.b), appliquée à la loi $\mathcal{P}(n\theta)$, donne

$$\text{Cov}(S_n^r, S_n) = r(n\theta)^r.$$

Ainsi

$$E(M_{r,n}Z_\theta) = r\theta^{r-1} = (\theta^r)'$$

Les trois conditions sont remplies : $M_{r,n}$ est régulier. D'après 10.b),

$$V(M_{r,n}) \geq \frac{(r\theta^{r-1})^2}{n/\theta} = \frac{r^2\theta^{2r-1}}{n}.$$

On admet uniquement la convergence absolue explicitement prévue par la question ; l'identité de (R_3) vient d'un calcul de covariance.

12.c)

Posons $\lambda = n\theta$. Comme $S_n \sim \mathcal{P}(\lambda)$, tous les moments utilisés ci-dessous existent. On a $M_{2,n} = (S_n^2 - S_n)/n^2$.

Méthode 1 — Moments ordinaires

Le carré du numérateur vaut $S_n^4 - 2S_n^3 + S_n^2$. D'après 5.a) et 6.c),

$$E(S_n^2) = \lambda^2 + \lambda,$$

$$E(S_n^3) = \lambda^3 + 3\lambda^2 + \lambda,$$

$$E(S_n^4) = \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda.$$

En combinant ces expressions,

$$E((S_n^2 - S_n)^2) = \lambda^4 + 4\lambda^3 + 2\lambda^2.$$

Puisque $E(M_{2,n}) = \theta^2$, la formule de Huygens donne

$$V(M_{2,n}) = \frac{\lambda^4 + 4\lambda^3 + 2\lambda^2}{n^4} - \theta^4 = \frac{4\theta^3}{n} + \frac{2\theta^2}{n^2}.$$

Méthode 2 — Produit de polynômes factoriels

Pour tout réel x , une expansion directe donne

$$(x^2)^2 = x^4 + 4x^3 + 2x^2.$$

En effet, le membre de droite vaut

$$(x^4 - 6x^3 + 11x^2 - 6x) + 4(x^3 - 3x^2 + 2x) + 2(x^2 - x) = x^4 - 2x^3 + x^2.$$

En prenant l'espérance et en utilisant les moments factoriels de paramètre λ ,

$$E((S_n^2)^2) = \lambda^4 + 4\lambda^3 + 2\lambda^2.$$

Le calcul de variance est alors identique et donne

$$V(M_{2,n}) = \frac{4\theta^3}{n} + \frac{2\theta^2}{n^2}.$$

La première méthode utilise les moments ordinaires disponibles ; la seconde exploite directement la structure factorielle de l'estimateur.

Pour la convergence, le biais est nul et la variance tend vers 0. Plus explicitement, pour tout $\varepsilon > 0$, Bienaymé-Tchebychev donne

$$P(|M_{2,n} - \theta^2| \geq \varepsilon) \leq \frac{4\theta^3/n + 2\theta^2/n^2}{\varepsilon^2} \rightarrow 0.$$

La suite $(M_{2,n})$ converge en probabilité vers θ^2 .

12.d)

On fixe $r \geq 1$ et $\theta > 0$.

Méthode 1 — Moments et variance

Pour $k \geq 1$, la décomposition de 3.a) et les moments factoriels donnent

$$E(S_n^k) = \sum_{j=1}^k \sigma(k, j)(n\theta)^j.$$

Le terme de degré k a pour coefficient $\sigma(k, k) = 1$. Comme la somme est finie et $\theta > 0$,

$$E(S_n^k) \sim (n\theta)^k \quad (n \rightarrow +\infty).$$

Le polynôme $(x^r)^2$ est unitaire de degré $2r$. Il existe donc des réels $a_0, \dots, a_{2r-1}$ tels que

$$(x^r)^2 = x^{2r} + \sum_{k=0}^{2r-1} a_k x^k.$$

Par conséquent,

$$E(M_{r,n}^2) = \frac{E(S_n^{2r})}{n^{2r}} + \sum_{k=0}^{2r-1} a_k \frac{E(S_n^k)}{n^{2r}}.$$

Le premier terme tend vers θ^{2r} . Pour $1 \leq k \leq 2r-1$, l'équivalent précédent entraîne $E(S_n^k)/n^{2r} \rightarrow 0$; pour $k=0$, ce quotient est $1/n^{2r} \rightarrow 0$. Ainsi

$$E(M_{r,n}^2) \rightarrow \theta^{2r}.$$

Comme $E(M_{r,n}) = \theta^r$, on a $V(M_{r,n}) \rightarrow 0$. Bienaymé-Tchebychev montre donc la convergence en probabilité vers θ^r .

Méthode 2 — Moyenne empirique et majoration directe du produit

Posons $Q_n = S_n/n$. Son espérance vaut θ et sa variance θ/n , donc $Q_n \rightarrow \theta$ en probabilité. De plus,

$$M_{r,n} = \prod_{j=0}^{r-1} \left(Q_n - \frac{j}{n} \right).$$

On utilise l'identité finie, obtenue en remplaçant les facteurs un à un,

$$\prod_{j=0}^{r-1} a_j - \prod_{j=0}^{r-1} b_j = \sum_{j=0}^{r-1} (a_j - b_j) \left(\prod_{i<j} a_i \right) \left(\prod_{i>j} b_i \right).$$

Sur l'événement $|Q_n - \theta| \leq 1$, tous les facteurs $Q_n - j/n$ et θ sont en valeur absolue majorés par $C = \theta + r + 1$. L'identité donne alors

$$|M_{r,n} - \theta^r| \leq rC^{r-1} \left(|Q_n - \theta| + \frac{r-1}{n} \right).$$

Pour $\varepsilon > 0$, choisissons $0 < \delta \leq 1$ tel que $rC^{r-1}\delta \leq \varepsilon/2$. Pour n assez grand, $rC^{r-1}(r-1)/n \leq \varepsilon/2$. Ainsi

$$P(|M_{r,n} - \theta^r| > \varepsilon) \leq P(|Q_n - \theta| > \delta) \rightarrow 0.$$

Cette preuve ne requiert aucun calcul de moment d'ordre élevé ni résultat général sur des opérations en convergence en loi. Dans les deux méthodes, $(M_{r,n})$ est convergente pour tout entier fixé $r \geq 1$.

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

13.a)

Méthode 1 — Cas particulier de l'estimateur factoriel

Pour $r = 1$, on a $M_{1,n} = S_n/n = \bar{X}_n$. La question 12.b) a démontré que $M_{r,n}$ est régulier pour tout $r \geq 1$. Il en résulte directement que $\bar{X}_n$ est un estimateur régulier de θ .

Méthode 2 — Vérification directe des trois conditions

L'application $t(k_1, \dots, k_n) = (k_1 + \dots + k_n)/n$ est indépendante de θ . Par les moments de S_n ,

$$E(\bar{X}_n) = \theta, \quad V(\bar{X}_n) = \frac{\theta}{n}.$$

Les conditions (R_1) et (R_2) sont donc satisfaites. Pour (R_3) , l'identité $\partial_1 F = F \partial_1 G$ conduit à la somme de transfert de $\bar{X}_n Z_\theta$. Cette somme est absolument convergente par Cauchy-Schwarz, puisque les deux variables admettent un moment d'ordre 2.

Comme $E(Z_\theta) = 0$,

$$E(\bar{X}_n Z_\theta) = \text{Cov} \left(\frac{S_n}{n}, \frac{S_n - n\theta}{\theta} \right) = \frac{V(S_n)}{n\theta} = 1.$$

Ce résultat est la dérivée de $\varphi(\theta) = \theta$, donc (R_3) est vérifiée. Cette vérification directe peut être réutilisée sans la construction des moments factoriels.

13.b)

Pour $\varphi(\theta) = \theta$, on a $\varphi'(\theta) = 1$ et $I(\theta) = n/\theta$. La borne de 10.b) devient

$$V(T) \geq \frac{\theta}{n}$$

pour tout estimateur régulier T de θ .

Or $V(\bar{X}_n) = \theta/n$. La moyenne empirique atteint donc cette borne : elle a la variance minimale parmi les estimateurs réguliers et est efficace au sens défini par l'énoncé.

14.a)

Pour un réel fixé α , on a $\psi(\alpha) = (1 - \alpha)\bar{X}_n + \alpha Y$. Les fonctions définissant Y et $\bar{X}_n$ ne dépendent pas de θ , et α non plus : cette combinaison est bien un estimateur.

Par linéarité de l'espérance,

$$E(\psi(\alpha)) = (1 - \alpha)\theta + \alpha\theta = \theta.$$

Les deux variables admettent un moment d'ordre 2. L'inégalité $(u + v)^2 \leq 2u^2 + 2v^2$ montre que leur combinaison en admet également un : (R_2) est vérifiée.

Enfin, les deux sommes de (R_3) sont absolument convergentes. Leur combinaison linéaire l'est aussi, car la valeur absolue d'un terme combiné est majorée par la somme des valeurs absolues correspondantes, multipliées par $|1 - \alpha|$ et $|\alpha|$. La linéarité donne pour cette somme

$$(1 - \alpha) \times 1 + \alpha \times 1 = 1.$$

C'est la dérivée de θ . Les trois conditions sont remplies : $\psi(\alpha)$ est régulier pour tout réel fixé α .

14.b)

Méthode 1 — Minimum de la variance de la combinaison

Posons $D = Y - \bar{X}_n$. La variable D admet une variance. Pour tout $\alpha \in \mathbb{R}$,

$$V(\psi(\alpha)) = V(\bar{X}_n) + 2\alpha \operatorname{Cov}(\bar{X}_n, D) + \alpha^2 V(D).$$

Comme $\psi(\alpha)$ est régulier, sa variance est au moins $\theta/n = V(\bar{X}_n)$. Le polynôme

$$q(\alpha) = 2\alpha \operatorname{Cov}(\bar{X}_n, D) + \alpha^2 V(D)$$

est donc positif ou nul pour tout α et vaut 0 en 0. La fonction polynomiale q possède un minimum en 0 ; sa dérivée doit y être nulle. Ainsi

$$0 = q'(0) = 2 \operatorname{Cov}(\bar{X}_n, D).$$

Par bilinéarité,

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

Méthode 2 — Relation directe avec Z_θ

L'identité de 10.a), appliquée à Y , donne $\operatorname{Cov}(Y, Z_\theta) = 1$. Or

$$Z_\theta = \frac{n}{\theta}(\bar{X}_n - \theta).$$

Par bilinéarité et symétrie de la covariance,

$$1 = \frac{n}{\theta} \text{Cov}(Y, \bar{X}_n).$$

Puisque $n > 0$ et $\theta > 0$, on retrouve $\text{Cov}(\bar{X}_n, Y) = \theta/n$.

La première méthode exploite l'efficacité et suit la construction de la question précédente; la seconde utilise directement l'identité de covariance établie en partie 3.

14.c)

La formule de variance d'une différence et 14.b) donnent

$$V(Y - \bar{X}_n) = V(Y) + V(\bar{X}_n) - 2 \text{Cov}(Y, \bar{X}_n) = V(Y) - V(\bar{X}_n).$$

Si Y est efficace, sa variance vaut nécessairement θ/n , puisque la moyenne empirique atteint la variance minimale. On a alors $V(Y - \bar{X}_n) = 0$.

Par ailleurs, les deux estimateurs sont sans biais, donc $E(Y - \bar{X}_n) = 0$. La formule de Huygens donne $E((Y - \bar{X}_n)^2) = 0$. Une variable aléatoire positive d'espérance nulle étant nulle presque sûrement, on conclut

$$Y = \bar{X}_n \quad \text{presque sûrement.}$$

L'estimateur efficace est donc unique à égalité presque sûre près, pour chaque $\theta > 0$.

15.a)

On développe la somme des carrés en utilisant

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

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

Pour $n \geq 2$, le dénominateur est strictement positif, et

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

En posant

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

on peut également écrire

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

15.b)

L'expression de 15.a) montre que W_n est intégrable : chaque X_i^2 et $\bar{X}_n^2$ admet une espérance finie. La moyenne empirique a pour espérance θ et pour variance θ/n , donc

$$E(\bar{X}_n^2) = \theta^2 + \frac{\theta}{n}.$$

Par linéarité,

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

La fonction définissant W_n ne dépend pas de θ , et son biais est nul : W_n est un estimateur sans biais de θ .

15.c)

Il suffit de montrer que W_n admet un moment d'ordre 2, sans calculer sa variance.

La définition de W_n et l'expression de 15.a) donnent

$$0 \leq W_n \leq \frac{1}{n-1} \sum_{i=1}^n X_i^2.$$

L'inégalité de Cauchy-Schwarz pour les n réels $X_1^2, \dots, X_n^2$ donne

$$\left(\sum_{i=1}^n X_i^2 \right)^2 \leq n \sum_{i=1}^n X_i^4.$$

Ainsi

$$0 \leq W_n^2 \leq \frac{n}{(n-1)^2} \sum_{i=1}^n X_i^4.$$

Chaque X_i possède un moment d'ordre 4 par 5.c), et la somme est finie. Le majorant admet donc une espérance finie. Par domination, $E(W_n^2) < +\infty$: W_n admet une variance.

15.d)

Convergence de la moyenne empirique

Méthode 1 — Bienaymé-Tchebychev

On a $E(\bar{X}_n) = \theta$ et $V(\bar{X}_n) = \theta/n$. Pour tout $\varepsilon > 0$,

$$P(|\bar{X}_n - \theta| \geq \varepsilon) \leq \frac{\theta}{n\varepsilon^2} \rightarrow 0.$$

Ainsi $\bar{X}_n \rightarrow \theta$ en probabilité.

Méthode 2 — Loi faible des grands nombres

Les X_i sont indépendantes, de même loi, d'espérance θ et de variance finie θ . La loi faible des grands nombres s'applique donc à leur moyenne et donne la même convergence en probabilité.

Le calcul de variance fournit une majoration explicite de l'erreur ; la loi des grands nombres permet de reconnaître directement le résultat général applicable.

Convergence de la variance empirique corrigée

Les variables X_i^2 sont indépendantes, car elles sont des fonctions de variables indépendantes, et de même loi. Elles ont pour espérance $\theta^2 + \theta$ et admettent une variance, puisque les X_i admettent un moment d'ordre 4.

La loi faible des grands nombres donne

$$A_n = \frac{1}{n} \sum_{i=1}^n X_i^2 \rightarrow \theta^2 + \theta \quad \text{en probabilité.}$$

La fonction $x \mapsto x^2$ est continue sur $\mathbb{R}$. La composition de la convergence en probabilité par cette fonction donne

$$\bar{X}_n^2 \rightarrow \theta^2 \quad \text{en probabilité.}$$

Par différence, $B_n = A_n - \bar{X}_n^2 \rightarrow \theta$ en probabilité. Enfin, $a_n = n/(n-1) \rightarrow 1$, et $W_n = a_n B_n$.

Pour justifier cette dernière étape avec une majoration explicite, la suite réelle $(a_n)_{n \geq 2}$ est bornée, par exemple par 2, et

$$|a_n B_n - \theta| \leq 2|B_n - \theta| + |a_n - 1|\theta.$$

Pour n assez grand, le second terme est inférieur à $\varepsilon/2$. Il vient

$$P(|W_n - \theta| \geq \varepsilon) \leq P(|B_n - \theta| \geq \varepsilon/4) \rightarrow 0.$$

Ainsi $W_n \rightarrow \theta$ en probabilité. Les deux suites d'estimateurs sont convergentes.

Justification de l'indication de l'énoncé

Supposons $a_n \rightarrow a$, $Y_n \rightarrow y$ et $Z_n \rightarrow z$ en probabilité. Choisissons une constante $C > 0$ telle que $|a_n| \leq C$ pour tout n assez grand. Alors

$$|a_n(Y_n - Z_n) - a(y - z)| \leq C|Y_n - y| + C|Z_n - z| + |a_n - a||y - z|.$$

Pour n assez grand, le dernier terme est au plus $\varepsilon/3$. Par inclusion d'événements et inégalité de l'union,

$$P(|a_n(Y_n - Z_n) - a(y - z)| > \varepsilon) \leq P\left(|Y_n - y| > \frac{\varepsilon}{3C}\right) + P\left(|Z_n - z| > \frac{\varepsilon}{3C}\right) \rightarrow 0.$$

Cela prouve exactement le résultat proposé, uniquement dans le cadre de la convergence en probabilité.

16)

Les trois estimateurs sont proches de θ pour $n = 50$, mais leurs centrages et leurs dispersions diffèrent.

La moyenne empirique est sans biais et sa variance vaut $\theta/50$. L'histogramme G est nettement plus concentré que les deux autres et centré près de 3 : il correspond à $\bar{X}_{50}$.

Pour distinguer les deux histogrammes étalés, on utilise une relation exacte, valable pour chaque échantillon :

$$W'_n = \frac{n-1}{n} W_n.$$

Pour $n = 50$,

$$W'_{50} = \frac{49}{50} W_{50}, \quad E(W_{50}) = \theta, \quad E(W'_{50}) = \frac{49}{50} \theta.$$

La variable W'_{50} est donc légèrement décalée vers la gauche et légèrement contractée par rapport à W_{50} . Son biais vaut $-\theta/50$. Sur les figures originales, F est légèrement décalée à gauche de E.

L'association attendue est donc :

Figure	Estimateur
E	W_{50}
F	W'_{50}
G	$\bar{X}_{50}$

La valeur du paramètre utilisée pour la simulation n'est pas explicitement donnée ; le centrage de G suggère $\theta \simeq 3$. Cette lecture approximative n'est pas une donnée supplémentaire du sujet. L'association repose sur la concentration de G et la contraction exacte reliant les deux variances empiriques, et non sur une identité générale entre un histogramme simulé et une loi théorique.

Relevé des adaptations

- 3.e) : code adapté en Python ; les commandes de création et d’affichage des tableaux sont remplacées par leurs équivalents Python. Le calcul et les listes demandées sont conservés.
- Après 8.b), rappels de sommation et de transfert : définition des sommes au moyen de limites de sommes finies sur des rectangles d’indices et maintien du rappel sur la convergence absolue. Cette formulation évite une théorie générale des familles sommables, tout en conservant les hypothèses nécessaires à (R_3) et à la suite du problème.
- 11.a) : code adapté en Python ; la taille de simulation N est rendue explicite dans les paramètres, et le blanc porte sur l’axe de sommation. Chaque colonne conserve le rôle d’un échantillon de taille n .
- 11.b) : commande de tracé remplacée par `matplotlib.pyplot.hist`, avec `density=True`. La figure est retranscrite en SVG dans l’énoncé à partir de l’original, avec les valeurs mesurées.
- 16) : la figure est retranscrite en SVG dans l’énoncé à partir de l’original, avec les valeurs mesurées. Les positions, hauteurs et graduations restent des lectures de la figure source, sans création de nouvelles données simulées.