



PRÉPASIMPLE

prepasimple.fr

HEC — 2021

Corrigé de mathématiques approfondies

Mathématiques approfondies — Prépa ECG

Annales de concours · Corrigé

Corrigé

Rappels et conventions utilisés dans tout le corrigé.

Polynômes : le degré d'un produit de deux fonctions polynomiales non nulles est la somme des degrés ; le coefficient dominant d'un produit est le produit des coefficients dominants ; deux fonctions polynomiales égales sur $\mathbb{R}$ ont les mêmes coefficients ; $(U_0, \dots, U_r)$ est une base de F_r , qui est donc de dimension $r + 1$. Les coordonnées d'un vecteur dans une base sont uniques.

Pour tout $r \in \mathbb{N}$, $V_{r+1} = V_r \times (U_1 - rU_0)$, c'est-à-dire $x^{r+1} = x^r(x - r)$ pour tout réel x ; on en déduit $U_1 V_r = V_{r+1} + rV_r$.

Loi de Poisson : si $X \hookrightarrow \mathcal{P}(\lambda)$, $P([X = k]) = e^{-\lambda} \frac{\lambda^k}{k!}$ pour $k \in \mathbb{N}$, $E(X) = \mathbf{V}(X) = \lambda$; une somme de variables de Poisson mutuellement indépendantes de paramètres $\lambda_1, \dots, \lambda_n$ suit la loi de Poisson de paramètre $\lambda_1 + \dots + \lambda_n$.

Moments : une variable discrète A admet un moment d'ordre m si $E(|A|^m)$ existe ; si A admet un moment d'ordre m , elle admet un moment de tout ordre $j \leq m$ (car $|A|^j \leq 1 + |A|^m$). Si A et B admettent une variance, $\mathbf{Cov}(A, B)$ existe et la covariance est bilinéaire et symétrique ; $\mathbf{Cov}(A, c) = 0$ pour une constante c ; $\mathbf{V}(A) = 0$ si et seulement si A est presque sûrement constante.

Inégalité de Cauchy-Schwarz pour la covariance : si A et B admettent une variance, $\mathbf{Cov}(A, B)^2 \leq \mathbf{V}(A)\mathbf{V}(B)$. Preuve, lorsque $\mathbf{V}(A) > 0$: pour tout réel λ , $0 \leq \mathbf{V}(\lambda A + B) = \lambda^2 \mathbf{V}(A) + 2\lambda \mathbf{Cov}(A, B) + \mathbf{V}(B)$; ce trinôme du second degré en λ , toujours positif, a un discriminant négatif ou nul : $4\mathbf{Cov}(A, B)^2 - 4\mathbf{V}(A)\mathbf{V}(B) \leq 0$.

Estimation : un estimateur T_n de $\varphi(\theta)$ est dit convergent s'il converge en probabilité vers $\varphi(\theta)$ pour toute valeur de θ . Condition suffisante : si $E(T_n) \rightarrow \varphi(\theta)$ et $\mathbf{V}(T_n) \rightarrow 0$, alors T_n est convergent (inégalité de Bienaymé-Tchebychev). Loi faible des grands nombres : si (A_i) est une suite de variables indépendantes de même loi admettant une espérance m et une variance,

$$\frac{1}{n} \sum_{i=1}^n A_i$$

converge en probabilité vers m . Si (Y_n) converge en probabilité vers y et (Z_n) vers z ,

alors $(Y_n + Z_n)$ converge en probabilité vers $y + z$ et, pour toute fonction h continue sur $\mathbb{R}$, $(h(Y_n))$ converge en probabilité vers $h(y)$.

Croissances comparées : pour tout réel a et tout réel $\lambda > 0$, $\frac{k^a \lambda^k}{k!} \xrightarrow{k \rightarrow +\infty} 0$.

Partie 1 – Polynômes factoriels

1.a) Méthode : raisonnement par l'absurde sur le terme de plus haut degré.

Supposons $\lambda_0 Q_0 + \dots + \lambda_r Q_r = 0$ avec des λ_k non tous nuls ; l'ensemble $\{k ; \lambda_k \neq 0\}$ est une partie non vide de $\llbracket 0, r \rrbracket$, donc m existe, et $\lambda_k = 0$ pour $k > m$. Ainsi

$$\lambda_m Q_m = - \sum_{k=0}^{m-1} \lambda_k Q_k.$$

Le membre de gauche est une fonction polynomiale de degré d_m (car $\lambda_m \neq 0$) ; le membre de droite est nul ou de degré au plus $\max(d_0, \dots, d_{m-1}) = d_{m-1} < d_m$ (si $m = 0$, il est nul). C'est absurde.

Conclusion

la seule combinaison linéaire nulle de $Q_0, \dots, Q_r$ est celle dont tous les coefficients sont nuls : une famille de fonctions polynomiales de degrés deux à deux distincts (ici strictement croissants) est libre.

1.b) Méthode : une famille libre de $r + 1$ vecteurs de F_r , espace de dimension $r + 1$, en est une base.

Comme $d_0, \dots, d_r$ sont des entiers naturels strictement croissants, une récurrence immédiate donne $d_k \geq k$ pour tout k , en particulier $d_r \geq r$.

Si $d_r \leq r$ (donc $d_r = r$) : tous les Q_k sont de degré $d_k \leq d_r = r$, donc appartiennent à F_r ; la famille est libre (question a)) et compte $r + 1 = \dim F_r$ vecteurs : c'est une base de F_r .

Si $d_r > r$: $Q_r \notin F_r$, donc la famille n'est pas une famille de vecteurs de F_r et ne peut pas en être une base.

Enfin $d_r = r$ équivaut à $d_k = k$ pour tout $k \in \llbracket 0, r \rrbracket$ (les entiers $d_0 < \dots < d_r$ sont alors $r + 1$ entiers distincts de $\llbracket 0, r \rrbracket$).

Conclusion

$(Q_0, \dots, Q_r)$ est une base de F_r si et seulement si $d_r = r$, c'est-à-dire si et seulement si $\deg Q_k = k$ pour tout $k \in \llbracket 0, r \rrbracket$; c'est une condition nécessaire et suffisante.

2.a) Méthode : un produit de réels est nul si et seulement si l'un des facteurs est nul.

Pour $r \geq 1$,

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

si et seulement si $x \in \{0, 1, \dots, r-1\}$. Pour $r = 0$, $V_0 = 1$ n'a pas de racine.

Conclusion

les racines de V_r sont $0, 1, \dots, r-1$ (toutes simples) si $r \geq 1$; V_0 n'a aucune racine.

2.b) Méthode : application de 1.b).

V_k est un produit de k facteurs de degré 1 et de coefficient dominant 1 : V_k est de degré k , de coefficient dominant 1. La famille $(V_0, \dots, V_r)$ vérifie donc $\deg V_k = k$ pour tout k , et 1.b) s'applique.

Conclusion

$(V_0, V_1, \dots, V_r)$ est une base de F_r .

2.c) Méthode 1 — Récurrence sur r (méthode suggérée). Pour $r = 2$, $x^2 = x^2 - x = x^2 - \frac{2 \times 1}{2}x$, ce qui est la formule (avec un reste nul). Supposons la formule vraie au rang $r \geq 2$. Alors, quand $x \rightarrow +\infty$:

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

Or $x^{r-1} = o(x^r)$ et $(x-r)o(x^{r-1}) = o(x^r)$. Comme $r + \frac{r(r-1)}{2} = \frac{(r+1)r}{2}$, on obtient $x^{r+1} = x^{r+1} - \frac{(r+1)r}{2}x^r + o(x^r)$, ce qui est la formule au rang $r+1$.

Méthode 2 — Développement du produit. En développant

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

le terme en x^r vaut x^r et le terme en x^{r-1} s'obtient en choisissant une fois $-k$ et $r-1$ fois x : son coefficient est

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

Les autres termes sont de degré au plus $r-2$, donc négligeables devant x^{r-1} en $+\infty$.

Comparaison : la méthode 1 est celle suggérée et se prête à la récurrence ; la méthode 2 donne directement le coefficient exact.

Conclusion

pour tout $r \geq 2$, $x^r \underset{x \rightarrow +\infty}{=} x^r - \frac{r(r-1)}{2}x^{r-1} + o(x^{r-1})$. Plus précisément, le coefficient de x^{r-1} dans V_r est $-\frac{r(r-1)}{2}$ (y compris pour $r=1$, où il vaut 0).

3.a) Méthode : coordonnées dans une base.

$U_r \in F_r$ et $(V_0, \dots, V_r)$ est une base de F_r (2.b)) : U_r s'écrit de manière unique comme combinaison linéaire de $V_0, \dots, V_r$.

Conclusion

il existe une unique famille $(\sigma(r, k))_{0 \leq k \leq r}$ telle que

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

3.b) Méthode : évaluations en des racines des V_k et identification de coefficients.

On utilise l'égalité

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

valable pour tout réel x , et le fait que $j^k = 0$ dès que $k > j$ (j est alors racine de V_k), tandis que $j^j = j!$.

(1) En $x = 0$: $0^r = 0$ (car $r \geq 1$), et $0^k = 0$ pour $k \geq 1$, $0^0 = 1$. Donc $\sigma(r, 0) = 0$.

(2) En $x = 1$: $1 = \sigma(r, 0) + \sigma(r, 1) \times 1^1 = 0 + \sigma(r, 1)$, donc $\sigma(r, 1) = 1$. Coefficient de x^r : V_r est de coefficient dominant 1 et les V_k , $k < r$, sont de degré $< r$; donc le coefficient de x^r dans

$$\sum \sigma(r, k) V_k$$

est $\sigma(r, r)$, qui vaut 1.

(3) Coefficient de x^{r-1} : seuls V_r et V_{r-1} contribuent; d'après 2.c) (méthode 2), il vaut

$\sigma(r, r) \times \left(-\frac{r(r-1)}{2}\right) + \sigma(r, r-1) \times 1$. Dans U_r , ce coefficient est nul. Donc

$\sigma(r, r-1) = \frac{r(r-1)}{2}$ (pour $r = 1$, on retrouve $\sigma(1, 0) = 0$).

(4) Pour $r \geq 2$, en $x = 2$:

$$2^r = \sigma(r, 0) + \sigma(r, 1) \times 2 + \sigma(r, 2) \times 2^2 = 0 + 2 + 2\sigma(r, 2),$$

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

Conclusion

les relations (1), (2), (3) et (4) sont établies.

3.c) Méthode : multiplier la décomposition de U_r par U_1 et utiliser $U_1 V_k = V_{k+1} + kV_k$, puis l'unicité des coordonnées dans la base $(V_0, \dots, V_{r+1})$ de F_{r+1} .

$$\begin{aligned} U_{r+1} &= U_1 U_r = \sum_{k=0}^r \sigma(r, k) U_1 V_k = \sum_{k=0}^r \sigma(r, k) V_{k+1} + \sum_{k=0}^r k \sigma(r, k) V_k = \sum_{j=1}^{r+1} \sigma(r, j-1) V_j \\ &+ \sum_{k=0}^r k \sigma(r, k) V_k. \end{aligned}$$

Pour $k \in \llbracket 1, r \rrbracket$, le coefficient de V_k dans cette expression est $\sigma(r, k-1) + k \sigma(r, k)$. Par unicité de la décomposition de U_{r+1} dans la base $(V_0, \dots, V_{r+1})$ (question 3.a) au rang $r+1$, ce coefficient est $\sigma(r+1, k)$.

Conclusion

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) Méthode : récurrence sur r .

Soit $\mathcal{P}(r)$: « pour tout $k \in \llbracket 1, r \rrbracket$, $\sigma(r, k) \in \mathbb{N}^*$ ». $\mathcal{P}(1)$ est vraie car $\sigma(1, 1) = 1$. Supposons $\mathcal{P}(r)$. Pour $k \in \llbracket 1, r \rrbracket$, $\sigma(r+1, k) = \sigma(r, k-1) + k \sigma(r, k)$, où $k \sigma(r, k) \in \mathbb{N}^*$ et $\sigma(r, k-1) \in \mathbb{N}$ (il vaut 0 si $k = 1$ d'après (1), et appartient à $\mathbb{N}^*$ si $k \geq 2$). Donc $\sigma(r+1, k) \in \mathbb{N}^*$. Enfin $\sigma(r+1, r+1) = 1$ d'après (2). Ainsi $\mathcal{P}(r+1)$ est vraie.

Conclusion

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

3.e) Code adapté en Python.

Méthode : calculer de proche en proche la liste $(\sigma(r+1, k))_{1 \leq k \leq r+1}$ à partir de la liste $(\sigma(r, k))_{1 \leq k \leq r}$ grâce à 3.c), avec $\sigma(r+1, 1) = \sigma(r+1, r+1) = 1$.

Dans le code, la liste s contient $\sigma(r, 1), \dots, \sigma(r, r)$: $s[k-1]$ vaut $\sigma(r, k)$, les indices commençant à 0. La nouvelle liste t , initialisée avec des 1, reçoit pour $2 \leq k \leq r$ la valeur $t[k-1] = s[k-2] + k * s[k-1]$, c'est-à-dire $\sigma(r, k-1) + k \sigma(r, k)$; ses premier et dernier termes restent égaux à 1. La boucle sur r de 1 à 4 produit et affiche les listes pour $r+1$ variant de 2 à 5.

Python

```
s = [1]                                # liste des sigma(1, k), 1 <= k <= 1
for r in range(1, 5):
    t = [1] * (r + 1)                  # sigma(r+1, 1) = sigma(r+1, r+1) = 1
    for k in range(2, r + 1):
        t[k - 1] = s[k - 2] + k * s[k - 1]
    s = t
    print(s)
```

Ce code a été exécuté; il affiche $[1, 1]$, puis $[1, 3, 1]$, puis $[1, 7, 6, 1]$, puis $[1, 15, 25, 10, 1]$. Ces valeurs sont cohérentes avec 3.b) : par exemple pour $r = 5$, $\sigma(5, 2) = 2^4 - 1 = 15$ et $\sigma(5, 4) = \frac{5 \times 4}{2} = 10$.

Variante utilisant `np.ones`, comme le suggère l'énoncé : on stocke les $\sigma(r, k)$ dans un tableau carré dont la case $[r-1, k-1]$ contient $\sigma(r, k)$.

Python

```
import numpy as np

S = np.ones((5, 5))

for r in range(1, 5):
    for k in range(2, r + 1):
        S[r, k - 1] = S[r - 1, k - 2] + k * S[r - 1, k - 1]

for r in range(2, 6):
    print(S[r - 1, 0:r])
```

Dans ce tableau, les cases utiles sont celles d'indices $[r-1, k-1]$ avec $1 \leq k \leq r$; la case $[r-1, 0]$ vaut $\sigma(r, 1) = 1$ et la case $[r-1, r-1]$ vaut $\sigma(r, r) = 1$ grâce à l'initialisation par des 1; $S[r - 1, 0 : r]$ extrait les r premiers coefficients de la ligne d'indice $r-1$. Ce code a été exécuté et affiche les mêmes listes (sous forme de réels).

Conclusion

les listes affichées sont $(1, 1)$, $(1, 3, 1)$, $(1, 7, 6, 1)$ et $(1, 15, 25, 10, 1)$.

4.a) Méthode : coordonnées de V_r dans la base canonique.

$V_r \in F_r$ et $(U_0, \dots, U_r)$ est une base de F_r : il existe une unique famille $(s(r, k))_{0 \leq k \leq r}$ telle que

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

c'est-à-dire

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

pour tout réel x (l'égalité de fonctions polynomiales équivaut à l'égalité des coefficients).

Conclusion

la famille $(s(r, k))_{0 \leq k \leq r}$ existe et est unique.

4.b) Méthode : $V_{r+1} = V_r \times (U_1 - rU_0)$ et identification des coefficients.

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

Pour $k \in \llbracket 1, r \rrbracket$, le coefficient de x^k est $s(r, k-1) - r s(r, k)$, et par unicité (4.a) au rang $r+1$ il vaut $s(r+1, k)$.

Valeur de $s(r, 1)$. Pour $r \geq 1$, $s(r, 0) = V_r(0) = 0$ (0 est racine de V_r). Pour $k = 1$, la relation donne $s(r+1, 1) = -r s(r, 1)$, avec $s(1, 1) = 1$ (car $V_1 = U_1$). Par récurrence, $s(r, 1) = (-1)^{r-1}(r-1)!$.

Conclusion

$s(r+1, k) = s(r, k-1) - r s(r, k)$ pour $k \in \llbracket 1, r \rrbracket$, et $s(r, 1) = (-1)^{r-1}(r-1)!$ pour tout $r \in \mathbb{N}^*$.

4.c) Méthode : récurrence sur r .

Soit $\mathcal{Q}(r)$: « pour tout $k \in \llbracket 1, r \rrbracket$, $(-1)^{r+k}s(r, k) > 0$ ». $\mathcal{Q}(1)$ est vraie car $s(1, 1) = 1$.
Supposons $\mathcal{Q}(r)$. Pour $k \in \llbracket 1, r \rrbracket$, d'après 4.b) :

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

Le second terme est strictement positif ($r \geq 1$ et $\mathcal{Q}(r)$) ; le premier est nul si $k = 1$ (car $s(r, 0) = 0$) et strictement positif si $k \geq 2$ ($\mathcal{Q}(r)$). La somme est strictement positive. Pour $k = r+1$, $s(r+1, r+1) = 1$ (V_{r+1} est de coefficient dominant 1) et $(-1)^{2r+2} = 1$. Donc $\mathcal{Q}(r+1)$ est vraie.

Conclusion

pour tout $r \in \mathbb{N}^*$ et tout $k \in \llbracket 1, r \rrbracket$, $s(r, k)$ est non nul et du signe de $(-1)^{r+k}$.

4.d) Méthode : substituer les décompositions l'une dans l'autre, puis identifier dans la base canonique.

$$U_r = \sum_{k=0}^r \sigma(r, k) V_k = \sum_{k=0}^r \sigma(r, k) \sum_{\ell=0}^k s(k, \ell) U_\ell = \sum_{\ell=0}^r \left(\sum_{k=\ell}^r \sigma(r, k) s(k, \ell) \right) U_\ell$$

(échange de deux sommes finies sur l'ensemble des couples (k, ℓ) tels que $0 \leq \ell \leq k \leq r$). Par unicité des coordonnées de U_r dans la base $(U_0, \dots, U_r)$, le coefficient de U_r vaut 1 et celui de U_ℓ , pour $\ell < r$, vaut 0. Pour $\ell = r$, la somme se réduit à $\sigma(r, r)s(r, r)$.

Conclusion

$\sigma(r, r)s(r, r) = 1$ et, pour tout $\ell \in \llbracket 0, r-1 \rrbracket$,

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

4.e) Méthode : utiliser 4.d) avec $\sigma(r, r) = 1$ et $\sigma(r, r-1) = \frac{r(r-1)}{2}$.

D'après 4.d) et (2), $s(r, r) = \frac{1}{\sigma(r, r)} = 1$ (ce que l'on savait déjà : V_r est de coefficient dominant 1).

Pour $r \geq 2$, 4.d) avec $\ell = r-1$ donne $\sigma(r, r-1)s(r-1, r-1) + \sigma(r, r)s(r, r-1) = 0$, d'où $s(r, r-1) = -\sigma(r, r-1) = -\frac{r(r-1)}{2}$, ce que confirme 2.c).

Conclusion

$$s(r, r) = 1 \text{ pour tout } r \in \mathbb{N}^* \text{ et } s(r, r-1) = -\frac{r(r-1)}{2} \text{ pour tout } r \geq 2.$$

Partie 2 – Quelques propriétés de la loi de Poisson

5.a) Méthode : cours sur la loi de Poisson et formule de Kœnig-Huygens.

Pour $X \hookrightarrow \mathcal{P}(\theta)$, on a $E(X) = \theta$ et $V(X) = \theta$, d'où $E(X^2) = V(X) + E(X)^2 = \theta + \theta^2$.

Conclusion

$$E(X) = V(X) = \theta \text{ et } E(X^2) = \theta + \theta^2.$$

5.b) Méthode : utiliser la relation admise

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

avec $\sigma(r, 0) = 0$ et les valeurs de $\sigma(r, k)$ affichées en 3.e) (et $\sigma(1, 1) = 1$).

Les valeurs $(\sigma(2, 1), \sigma(2, 2)) = (1, 1)$, $(\sigma(3, k))_{1 \leq k \leq 3} = (1, 3, 1)$ et $(\sigma(4, k))_{1 \leq k \leq 4} = (1, 7, 6, 1)$ donnent directement les écritures demandées. Vérification pour $X^2 : X^2 + X^1 = X(X-1) + X = X^2$.

Conclusion

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

5.c) Méthode 1 — Critère de comparaison pour une série à termes positifs. Soit $r \in \mathbb{N}$.

Comme X est à valeurs dans $\mathbb{N}$, X^r admet une espérance si et seulement si la série

$$\sum_{k \geq 0} k^r e^{-\theta} \frac{\theta^k}{k!},$$

à termes positifs, converge (formule de transfert). Par croissances comparées (avec

$a = r + 2$), $k^2 \times k^r \frac{\theta^k}{k!} \xrightarrow{k \rightarrow +\infty} 0$, donc il existe un rang k_0 à partir duquel

$0 \leq k^r e^{-\theta} \frac{\theta^k}{k!} \leq \frac{1}{k^2}$. La série de Riemann

$$\sum \frac{1}{k^2}$$

converge, donc la série étudiée converge.

Méthode 2 — Utiliser les puissances descendantes. Pour tout entier $j \geq 1$ et tout $k \in \mathbb{N}$, $k^{\underline{j}}$ est nul si $k < j$ (l'un des facteurs est nul) et vaut $\frac{k!}{(k-j)!} > 0$ si $k \geq j$. La série

$$\sum_{k \geq j} \frac{k!}{(k-j)!} e^{-\theta} \frac{\theta^k}{k!} = e^{-\theta} \theta^j \sum_{i \geq 0} \frac{\theta^i}{i!}$$

(on pose $i = k - j$) est à termes positifs et convergente (série exponentielle). Donc $X^{\underline{j}}$ admet une espérance (c'est le calcul de 6.b) ci-dessous, qui n'utilise pas 5.c)). Par la relation admise,

$$X^r = \sum_{k=0}^r \sigma(r, k) X^{\underline{k}}$$

est une combinaison linéaire de variables admettant une espérance : elle admet une espérance.

Comparaison : la méthode 1 est la plus directe ; la méthode 2 réutilise les puissances descendantes et prépare 6.b).

Conclusion

pour tout $r \in \mathbb{N}$, $E(X^r)$ existe : X admet des moments de tous ordres.

6.a) Méthode : $(X^r)^m$ est un polynôme en X .

Soit $m \in \mathbb{N}^*$. La fonction V_r^m est polynomiale, de la forme

$$\sum_{j=0}^{rm} a_j U_j;$$

donc

$$(X^r)^m = \sum_{j=0}^{rm} a_j X^j$$

est une combinaison linéaire de variables admettant une espérance (5.c)) : elle admet une espérance. De plus $X^r \geq 0$ (vu en 5.c), méthode 2, puisque X est à valeurs dans $\mathbb{N}$, donc $|X^r|^m = (X^r)^m$.

Conclusion

pour tout $m \in \mathbb{N}^*$, $E(|X^r|^m)$ existe : X^r admet des moments de tous ordres.

6.b) Méthode : formule de transfert et série exponentielle.

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

(les termes d'indice $k < r$ sont nuls, puis on a posé $i = k - r$; la série est à termes positifs et convergente).

Conclusion

pour tout entier $r \geq 1$, $E(X^r) = \theta^r$ (formule encore vraie pour $r = 0$).

6.c) Méthode : linéarité de l'espérance appliquée aux écritures de 5.b), puis 6.b).

$$E(X^3) = E(X^3) + 3E(X^2) + E(X^1) = \theta^3 + 3\theta^2 + \theta;$$

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

On retrouve au passage $E(X^2) = \theta^2 + \theta$.

Conclusion

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

7) Question adaptée au programme actuel : l'énoncé utilise la notation ∂_1 pour une fonction dont la seconde variable est entière ; on a précisé dans le sujet que $\partial_1(g)(\theta, k)$ désigne la dérivée de $\theta \mapsto g(\theta, k)$ à k fixé, et que la même convention vaut dans la partie 3. Il s'agit donc uniquement de dérivées de fonctions d'une variable réelle.

7.a) Méthode : simplifier le logarithme puis dériver par rapport à θ .

Pour $\theta > 0$ et $k \in \mathbb{N}$, $f(\theta, k) > 0$ et $g(\theta, k) = k \ln \theta - \theta - \ln(k!)$. La fonction $\theta \mapsto g(\theta, k)$ est dérivable sur $\mathbb{R}_+^*$ et

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

Conclusion

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

7.b) Méthode : la relation $U_1 V_r = V_{r+1} + r V_r$ (rappels), puis la formule

$$\text{Cov}(A, B) = E(AB) - E(A)E(B).$$

Pour tout réel x , $x \times x^r = x^r((x-r) + r) = x^{r+1} + r x^r$; appliqué à $x = X(\omega)$, cela donne $XX^r = X^{r+1} + r X^r$.

Les variables X et X^r admettent une variance (5.c) et 6.a)), donc leur covariance existe, et par linéarité et 6.b) :

$$\mathbf{Cov}(X, X^r) = E(X^{r+1}) + rE(X^r) - E(X)E(X^r) = \theta^{r+1} + r\theta^r - \theta \times \theta^r = r\theta^r.$$

Conclusion

$$XX^r = X^{r+1} + rX^r \text{ et } \mathbf{Cov}(X, X^r) = r\theta^r.$$

7.c) Méthode : bilinéarité de la covariance, puis inégalité de Cauchy-Schwarz (démontrée dans les rappels).

Par bilinéarité, et puisque la covariance avec une constante est nulle :

$$\mathbf{Cov}(\partial_1(g)(\theta, X), X^r) = \mathbf{Cov}\left(\frac{X}{\theta} - 1, X^r\right) = \frac{1}{\theta} \mathbf{Cov}(X, X^r) = r\theta^{r-1}.$$

Par ailleurs $\mathbf{V}(\partial_1(g)(\theta, X)) = \frac{1}{\theta^2} \mathbf{V}(X) = \frac{1}{\theta} > 0$. L'inégalité de Cauchy-Schwarz donne

$$r^2 \theta^{2r-2} \leq \frac{1}{\theta} \mathbf{V}(X^r), \text{ soit } \mathbf{V}(X^r) \geq r^2 \theta^{2r-1}.$$

Conclusion

$$\mathbf{Cov}(\partial_1(g)(\theta, X), X^r) = r\theta^{r-1} \text{ et, pour tout } \theta > 0 \text{ et tout } r \in \mathbb{N}^*, \mathbf{V}(X^r) \geq r^2 \theta^{2r-1}.$$

Partie 3 – Estimation ponctuelle de fonctions du paramètre θ

8.a) Méthode : écrire F sous forme explicite.

Fixons $(k_1, \dots, k_n) \in \mathbb{N}^n$ et posons $s = k_1 + \dots + k_n$. Alors

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

La fonction $\theta \mapsto \theta^s e^{-n\theta}$ est dérivable sur $\mathbb{R}_+^*$ (produit de fonctions dérivables, $s \in \mathbb{N}$), de dérivée $(s\theta^{s-1} - n\theta^s)e^{-n\theta}$ (formule valable aussi pour $s = 0$, le premier terme étant alors nul). Donc $\partial_1(F)(\theta, k_1, \dots, k_n) = \left(\frac{s}{\theta} - n\right) F(\theta, k_1, \dots, k_n)$.

Comme $F > 0$, $G = \ln F = s \ln \theta - n\theta - \ln(k_1! \times \dots \times k_n!)$ est dérivable sur $\mathbb{R}_+^*$, de dérivée $\frac{s}{\theta} - n$.

Conclusion

les deux fonctions sont dérivables sur $\mathbb{R}_+^*$ et, avec

$$s = \sum_{i=1}^n k_i,$$

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

8.b) Méthode : exprimer Z_θ à l'aide de $S_n = X_1 + \dots + X_n$, qui suit la loi de Poisson de paramètre $n\theta$ (stabilité de la loi de Poisson par somme de variables mutuellement indépendantes).

D'après 8.a), $Z_\theta = \frac{S_n}{\theta} - n$. Donc $E(Z_\theta) = \frac{n\theta}{\theta} - n = 0$ et

$$\mathbf{V}(Z_\theta) = \frac{1}{\theta^2} \mathbf{V}(S_n) = \frac{n\theta}{\theta^2} = \frac{n}{\theta} > 0.$$

Conclusion

$$Z_\theta \text{ est centrée et } I(\theta) = \mathbf{V}(Z_\theta) = \frac{n}{\theta} > 0.$$

9.a) Méthode : définition d'un estimateur sans biais.

Le paramètre θ étant inconnu, la condition (R_1) s'entend pour tout $\theta > 0$: $E(T) = \varphi(\theta)$ pour tout $\theta > 0$. C'est exactement la définition d'un estimateur sans biais de $\varphi(\theta)$.

Conclusion

oui, un estimateur régulier de $\varphi(\theta)$ est sans biais.

9.b) Méthode : écrire (R_1) comme une égalité entre fonctions de θ et examiner ce que demanderait sa dérivation.

D'après la formule de transfert, (R_1) s'écrit : pour tout $\theta > 0$,

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

Le membre de droite est une somme infinie de fonctions dérivables de θ . Pour obtenir (R_3) , il faudrait dériver cette somme terme à terme, c'est-à-dire échanger la dérivation et la sommation infinie. Or la dérivée d'une somme infinie de fonctions n'est pas, en général, la somme des dérivées ; aucun résultat du cours ne permet de faire cet échange ici. De plus, la condition (R_3) exige l'absolue convergence de la série

$$\sum t \partial_1(F) = \sum t \left(\frac{s}{\theta} - n \right) F,$$

qui ne découle pas de celle de

$$\sum tF :$$

le facteur $\frac{s}{\theta} - n$ n'est pas borné.

Conclusion

(R_3) revient à dériver terme à terme l'égalité (R_1) , ce qui n'est pas automatique : il faut à la fois justifier l'échange de la dérivée et de la somme infinie et l'absolue convergence de la série dérivée. C'est pourquoi (R_3) est imposée comme une condition supplémentaire.

10.a) Méthode : écrire $\partial_1(F) = \partial_1(G) \times F$ et appliquer la formule de transfert rappelée.

Pour tout $(k_1, \dots, k_n)$, comme $G = \ln F$ avec $F > 0$, $\partial_1(G) = \frac{\partial_1(F)}{F}$, donc $t \partial_1(F) = u_\theta F$ où $u_\theta(k_1, \dots, k_n) = t(k_1, \dots, k_n) \partial_1(G)(\theta, k_1, \dots, k_n)$. D'après (R_3) , la série

$$\sum u_\theta F$$

est absolument convergente; la formule de transfert montre alors que

$U_\theta = u_\theta(X_1, \dots, X_n) = T \times Z_\theta$ est d'espérance finie et que

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

Enfin T (condition (R_2)) et Z_θ (8.b)) admettent une variance, donc $\mathbf{Cov}(T, Z_\theta)$ existe et vaut $E(T Z_\theta) - E(T)E(Z_\theta) = E(T Z_\theta)$ puisque Z_θ est centrée.

Conclusion

pour tout $\theta > 0$, $\varphi'(\theta) = E(T \times Z_\theta) = \mathbf{Cov}(T, Z_\theta)$.

10.b) Méthode : inégalité de Cauchy-Schwarz (rappels).

T et Z_θ admettent une variance, donc

$(\varphi'(\theta))^2 = \mathbf{Cov}(T, Z_\theta)^2 \leq \mathbf{V}(T)\mathbf{V}(Z_\theta) = \mathbf{V}(T)I(\theta)$. Comme $I(\theta) > 0$, on peut diviser.

Conclusion

pour tout $\theta > 0$, $\mathbf{V}(T) \geq \frac{(\varphi'(\theta))^2}{I(\theta)}$.

11.a) Code adapté en Python.

Méthode : identifier, dans le tableau X , l'emplacement d'un n -échantillon.

Le tableau X a n lignes et N colonnes; chaque colonne j contient n réalisations

indépendantes de la loi $\mathcal{P}(\theta)$, c'est-à-dire une réalisation $(X_{1,j}, \dots, X_{n,j})$ de l'échantillon $(X_1, \dots, X_n)$. D'après 8.b), $Z_\theta = \frac{S_n}{\theta} - n = \frac{S_n - n\theta}{\theta}$: il faut donc, pour chaque colonne, sommer les n coefficients, ce qui correspond à une somme sur l'indice de ligne i . C'est l'instruction `np.sum(X, axis=0)`, qui renvoie un tableau de N valeurs ; la fonction renvoie alors bien N réalisations de Z_θ .

Python

```
import numpy as np

import numpy.random as rd

def Z_th(n, theta):
    X = rd.poisson(theta, (n, N))

    ech = (np.sum(X, axis=0) - n*theta) / theta

    return ech
```

Conclusion

il faut compléter par `axis=0`.

11.b) Méthode : comparer les variances $I(\theta) = \frac{n}{\theta}$, puis, à variance égale, la structure des valeurs prises par Z_θ .

Les quatre variables Z_θ sont centrées, de variances respectives $\frac{10}{4} = 2,5$ pour $(10, 4)$, $\frac{20}{4} = 5$ pour $(20, 4)$, $\frac{40}{4} = 10$ pour $(40, 4)$ et $\frac{50}{5} = 10$ pour $(50, 5)$. Plus la variance est grande, plus l'histogramme est étalé et, son aire valant 1, moins il est haut. De plus, $Z_\theta = \sqrt{n/\theta} \times \frac{S_n - n\theta}{\sqrt{n\theta}}$ et, d'après le théorème limite central, la loi de $\frac{S_n - n\theta}{\sqrt{n\theta}}$ (somme centrée réduite de n variables indépendantes de même loi) est proche de la loi normale centrée réduite : l'histogramme de Z_θ est proche de la densité de la loi normale de variance $\frac{n}{\theta}$, dont le maximum vaut $\frac{1}{\sqrt{2\pi n/\theta}}$, soit environ 0,25 ; 0,18 ; 0,13 et 0,13.

On en déduit que la figure C (la plus haute, hauteur environ 0,25) correspond à $(10, 4)$ et la figure B (hauteur environ 0,18) à $(20, 4)$. Les figures A et D, de même étalement et de hauteur environ 0,13, correspondent aux deux couples de variance 10.

Pour les distinguer, on observe que $Z_\theta = \frac{S_n - n\theta}{\theta}$ ne prend que des valeurs de la forme

$\frac{j - n\theta}{\theta}$ avec $j \in \mathbb{N}$, régulièrement espacées de $\frac{1}{\theta}$: de 0,25 pour $\theta = 4$ (multiples de 0,25 puisque $n\theta = 160$) et de 0,2 pour $\theta = 5$ (multiples de 0,2 puisque $n\theta = 250$). Les classes des histogrammes ont une largeur voisine de 0,75 (valeur lue sur la figure). Pour $\theta = 4$, $0,75 = 3 \times 0,25$: si les bornes des classes sont des multiples de 0,25 (par exemple $-10 + 0,75m$), chaque classe contient exactement trois valeurs possibles, voisines, et l'histogramme est régulier. Pour $\theta = 5$, $\frac{0,75}{0,2} = 3,75$: selon leur position, les classes contiennent trois ou quatre valeurs possibles, et leurs hauteurs varient irrégulièrement d'une classe à sa voisine : l'histogramme est irrégulier.

Vérification par simulation : le code suivant a été exécuté avec des classes de largeur 0,75 ; la commande `np.var` renvoie la variance empirique de l'échantillon et `plt.hist` renvoie en premier lieu le tableau des hauteurs des classes, dont on affiche cinq valeurs centrales.

Python

```
import numpy as np

import numpy.random as rd

import matplotlib.pyplot as plt

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

N = 100000

classes = np.arange(-10, 10.01, 0.75)

for (n, theta) in [(10, 4), (20, 4), (40, 4), (50, 5)]:
    ech = Z_th(n, theta)
    hauteurs, bornes, rectangles = plt.hist(ech, bins=classes, density=True)
    plt.show()
    print((n, theta), np.var(ech), hauteurs[11:16])
```

Les variances empiriques obtenues sont environ 2,49 ; 4,99 ; 10,03 et 9,91. Les histogrammes pour (10, 4), (20, 4) et (40, 4) sont réguliers (pour (40, 4) : 0,113 ; 0,123 ; 0,128 ; 0,122 ; 0,112) ; celui pour (50, 5) présente des sauts de hauteur entre classes

voisines (par exemple 0,092; 0,132; 0,135; 0,129; 0,088 autour de 0), comme la figure D.

Conclusion

figure A pour (40, 4), figure B pour (20, 4), figure C pour (10, 4) et figure D pour (50, 5).

12.a) Méthode : stabilité de la loi de Poisson par somme de variables mutuellement indépendantes.

Conclusion

$$S_n \hookrightarrow \mathcal{P}(n\theta), E(S_n) = V(S_n) = n\theta.$$

12.b) Méthode : appliquer à $S_n \hookrightarrow \mathcal{P}(n\theta)$ les résultats de la partie 2 (valables pour tout paramètre strictement positif), puis 10.b).

On a $M_{r,n} = \frac{S_n^r}{n^r} = t(X_1, \dots, X_n)$ avec $t(k_1, \dots, k_n) = \frac{(k_1 + \dots + k_n)^r}{n^r}$, application de $\mathbb{N}^n$ dans $\mathbb{R}$ indépendante de θ .

(R_1) : d'après 6.b) appliqué à S_n , $E(S_n^r) = (n\theta)^r$, donc $E(M_{r,n}) = \theta^r$ pour tout $\theta > 0$.

(R_2) : d'après 6.a) appliqué à S_n , S_n^r admet un moment d'ordre 2, donc $M_{r,n}$ admet une variance.

(R_3) : la série

$$\sum t \partial_1(F)$$

est absolument convergente (admis, NB). Comme en 10.a), $t \partial_1(F) = t \partial_1(G) F$, et la formule de transfert donne

$$\begin{aligned} \sum_{(k_1, \dots, k_n) \in \mathbb{N}^n} t(k_1, \dots, k_n) \partial_1(F)(\theta, k_1, \dots, k_n) &= E(M_{r,n} Z_\theta) = E\left(\frac{S_n^r}{n^r} \left(\frac{S_n}{\theta} - n\right)\right) \\ &= \frac{1}{\theta n^r} E(S_n S_n^r) - n E(M_{r,n}). \end{aligned}$$

D'après 7.b) appliqué à S_n , $S_n S_n^r = S_n^{r+1} + r S_n^r$, d'espérance $(n\theta)^{r+1} + r(n\theta)^r$. Donc la somme vaut $n\theta^r + r\theta^{r-1} - n\theta^r = r\theta^{r-1} = \varphi'(\theta)$.

Remarque : l'absolue convergence admise peut se démontrer. Avec

$$s = \sum k_j,$$

on a $0 \leq s^r \leq s^r$ et

$$s \leq \prod_{j=1}^n (1 + k_j)$$

(le développement du produit contient $1 + k_1 + \dots + k_n$ et des termes positifs), d'où, puisque $n^r \geq 1$,

$$|t \partial_1(F)| \leq s^r \left(\frac{s}{\theta} + n \right) F \leq \left(n + \frac{1}{\theta} \right) \prod_{j=1}^n (1 + k_j)^{r+1} f(\theta, k_j),$$

et chaque série

$$\sum_k (1 + k)^{r+1} f(\theta, k)$$

converge, car X admet des moments de tous ordres (5.c)).

Ainsi $M_{r,n}$ est un estimateur régulier de θ^r , et 10.b) avec $\varphi'(\theta) = r\theta^{r-1}$ et $I(\theta) = \frac{n}{\theta}$

donne $\mathbf{V}(M_{r,n}) \geq \frac{r^2 \theta^{2r-2} \theta}{n}$.

Conclusion

$M_{r,n}$ est un estimateur régulier de θ^r et, pour tout $\theta > 0$, $\mathbf{V}(M_{r,n}) \geq \frac{r^2 \theta^{2r-1}}{n}$.

12.c) Méthode : écrire $(x^2)^2$ dans la base des puissances descendantes, puis utiliser 6.b) pour $S_n \hookrightarrow \mathcal{P}(\lambda)$ avec $\lambda = n\theta$.

Pour tout réel x ,

$$x^4 + 4x^3 + 2x^2 = x(x-1)[(x-2)(x-3) + 4(x-2) + 2] = x(x-1)(x^2 - x) = (x^2)^2.$$

Donc

$$E((S_n^2)^2) = \lambda^4 + 4\lambda^3 + 2\lambda^2 \text{ et } \mathbf{V}(S_n^2) = \lambda^4 + 4\lambda^3 + 2\lambda^2 - (\lambda^2)^2 = 4\lambda^3 + 2\lambda^2.$$

Par suite $\mathbf{V}(M_{2,n}) = \frac{4n^3\theta^3 + 2n^2\theta^2}{n^4} = \frac{4\theta^3}{n} + \frac{2\theta^2}{n^2}$ (on retrouve bien $\mathbf{V}(M_{2,n}) \geq \frac{4\theta^3}{n}$, conformément à 12.b)).

Convergence : $M_{2,n}$ est sans biais ($E(M_{2,n}) = \theta^2$) et $\mathbf{V}(M_{2,n}) \xrightarrow{n \rightarrow +\infty} 0$. Par l'inégalité de

Bienaymé-Tchebychev, pour tout $\varepsilon > 0$, $P(|M_{2,n} - \theta^2| \geq \varepsilon) \leq \frac{\mathbf{V}(M_{2,n})}{\varepsilon^2} \xrightarrow{n \rightarrow +\infty} 0$.

Conclusion

$\mathbf{V}(M_{2,n}) = \frac{4\theta^3}{n} + \frac{2\theta^2}{n^2}$ et la suite $(M_{2,n})_{n \geq 1}$ est une suite convergente d'estimateurs de θ^2 .

12.d) Méthode 1 — Calcul asymptotique de la variance (méthode suggérée). Pour $k \in \mathbb{N}^*$, la relation admise appliquée à S_n et 6.b) donnent

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

qui est un polynôme en n de degré k et de coefficient dominant $\sigma(k, k)\theta^k = \theta^k \neq 0$. Donc $E(S_n^k) \underset{n \rightarrow +\infty}{\sim} (n\theta)^k$, et en particulier $E(S_n^k) = o(n^{2r})$ pour $k < 2r$.

La fonction V_r^2 est polynomiale, de degré $2r$ et de coefficient dominant 1 :

$$V_r^2 = U_{2r} + \sum_{j=0}^{2r-1} c_j U_j.$$

Donc

$$E((S_n^r)^2) = E(S_n^{2r}) + \sum_{j=0}^{2r-1} c_j E(S_n^j) = (n\theta)^{2r} + o(n^{2r}),$$

d'où $E(M_{r,n}^2) = \frac{E((S_n^r)^2)}{n^{2r}} \xrightarrow{n \rightarrow +\infty} \theta^{2r}$ et $V(M_{r,n}) = E(M_{r,n}^2) - \theta^{2r} \xrightarrow{n \rightarrow +\infty} 0$. Comme $M_{r,n}$ est sans biais, l'inégalité de Bienaymé-Tchebychev conclut comme en 12.c).

Méthode 2 — Loi faible des grands nombres. On a

$$M_{r,n} = \prod_{j=0}^{r-1} A_{j,n}$$

avec $A_{j,n} = \frac{S_n}{n} - \frac{j}{n}$. Par la loi faible des grands nombres ($X_1, \dots, X_n$ indépendantes, de même loi, d'espérance θ et admettant une variance), $\frac{S_n}{n}$ converge en probabilité vers θ . La suite constante (certaine) $-\frac{j}{n}$ converge en probabilité vers 0 (pour $\varepsilon > 0$, $P\left(\frac{j}{n} \geq \varepsilon\right) = 0$ dès que $n > \frac{j}{\varepsilon}$). Par somme, $A_{j,n}$ converge en probabilité vers θ . Pour un produit : si A_n converge en probabilité vers a et B_n vers b , alors $A_n B_n = \frac{1}{4}((A_n + B_n)^2 - (A_n - B_n)^2)$; $A_n + B_n$ converge vers $a + b$ et $A_n - B_n = A_n + (-B_n)$ vers $a - b$ (somme, et composition par $x \mapsto -x$), puis leurs carrés vers $(a + b)^2$ et $(a - b)^2$ (composition par $x \mapsto x^2$, continue), et finalement $A_n B_n$ converge vers $\frac{1}{4}((a + b)^2 - (a - b)^2) = ab$ (somme, et compositions par $x \mapsto -x$ et $x \mapsto \frac{x}{4}$). Par récurrence sur le nombre de facteurs, $M_{r,n}$ converge en probabilité vers θ^r .

Comparaison : la méthode 1 suit l'indication et donne en plus $\mathbf{V}(M_{r,n}) \rightarrow 0$; la méthode 2 est plus courte et n'utilise que la loi faible des grands nombres et les opérations sur les convergences en probabilité.

Conclusion

oui, pour tout $r \geq 1$, la suite $(M_{r,n})_{n \geq 1}$ est une suite convergente d'estimateurs de θ^r .

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

13.a) Méthode : reconnaître un cas particulier de 12.b).

$\bar{X}_n = \frac{S_n}{n} = M_{1,n}$ (pour $r = 1$, $S_n^1 = S_n$). D'après 12.b) avec $r = 1$, $M_{1,n}$ est un estimateur régulier de $\theta^1 = \theta$.

Conclusion

$\bar{X}_n$ est un estimateur régulier de θ .

13.b) Méthode : appliquer 10.b) avec $\varphi(\theta) = \theta$, puis calculer $\mathbf{V}(\bar{X}_n)$.

Avec $\varphi'(\theta) = 1$ et $I(\theta) = \frac{n}{\theta}$, l'inégalité de 10.b) devient : pour tout estimateur régulier T de θ et tout $\theta > 0$, $\mathbf{V}(T) \geq \frac{\theta}{n}$. Or $\mathbf{V}(\bar{X}_n) = \frac{1}{n^2} \mathbf{V}(S_n) = \frac{n\theta}{n^2} = \frac{\theta}{n}$.

Conclusion

tout estimateur régulier T de θ vérifie $\mathbf{V}(T) \geq \frac{\theta}{n}$, et $\bar{X}_n$ réalise l'égalité : sa variance est minimale parmi les estimateurs réguliers de θ , c'est-à-dire que $\bar{X}_n$ est efficace (au sens défini juste après).

14.a) Méthode : vérifier les trois conditions ; pour (R_3) , établir l'absolue convergence puis utiliser la formule de transfert et 10.a).

Notons t_1 et t_2 les applications de $\mathbb{N}^n$ dans $\mathbb{R}$, indépendantes de θ , telles que

$\bar{X}_n = t_1(X_1, \dots, X_n)$ et $Y = t_2(X_1, \dots, X_n)$. Alors

$\psi(\alpha) = (1 - \alpha)\bar{X}_n + \alpha Y = t_\alpha(X_1, \dots, X_n)$ avec $t_\alpha = (1 - \alpha)t_1 + \alpha t_2$, indépendante de θ .

(R_1) : par linéarité de l'espérance, $E(\psi(\alpha)) = (1 - \alpha)\theta + \alpha\theta = \theta$.

(R_2) : $\bar{X}_n$ et Y admettent un moment d'ordre 2 et $\psi(\alpha)^2 \leq 2(1 - \alpha)^2\bar{X}_n^2 + 2\alpha^2 Y^2$ (car $2ab \leq a^2 + b^2$), donc $\psi(\alpha)$ admet un moment d'ordre 2, donc une variance.

(R_3) : d'après (R_3) pour $\bar{X}_n$ et pour Y , il existe des séries absolument convergentes

$$\sum_k v_{j,k}$$

et

$$\sum_k w_{j,k}$$

$(1 \leq j \leq n)$ telles que

$$|t_1 \partial_1(F)| \leq \prod_{j=1}^n |v_{j,k_j}|$$

et

$$|t_2 \partial_1(F)| \leq \prod_{j=1}^n |w_{j,k_j}|.$$

Alors

$$|t_\alpha \partial_1(F)| \leq |1 - \alpha| \prod_{j=1}^n |v_{j,k_j}| + |\alpha| \prod_{j=1}^n |w_{j,k_j}| \leq \prod_{j=1}^n \left((1 + |\alpha|)|v_{j,k_j}| + (1 + |\alpha|)|w_{j,k_j}| \right),$$

la dernière inégalité venant de ce que le développement du produit de droite, dont tous

les termes sont positifs, contient

$$(1 + |\alpha|)^n \prod_j |v_{j,k_j}|$$

et

$$(1 + |\alpha|)^n \prod_j |w_{j,k_j}|,$$

avec $(1 + |\alpha|)^n \geq \max(|1 - \alpha|, |\alpha|)$. Chaque série

$$\sum_k (1 + |\alpha|) (|v_{j,k}| + |w_{j,k}|)$$

converge : la série

$$\sum t_\alpha \partial_1(F)$$

est absolument convergente. Comme en 10.a), $t_\alpha \partial_1(F) = t_\alpha \partial_1(G) F$ et la formule de transfert donne, avec la linéarité de l'espérance et 10.a) appliqué à $\bar{X}_n$ et à Y (estimateurs réguliers de θ , donc $E(\bar{X}_n Z_\theta) = E(Y Z_\theta) = 1$) :

$$\begin{aligned} \sum_{(k_1, \dots, k_n) \in \mathbb{N}^n} t_\alpha \partial_1(F) &= E(\psi(\alpha) Z_\theta) = (1 - \alpha) E(\bar{X}_n Z_\theta) + \alpha E(Y Z_\theta) = (1 - \alpha) + \alpha = 1 \\ &= \varphi'(\theta). \end{aligned}$$

Conclusion

pour tout réel α , $\psi(\alpha)$ est un estimateur régulier de θ .

14.b) Méthode 1 — Minimalité de la variance de $\bar{X}_n$ (déduction demandée). Posons $D = Y - \bar{X}_n$, $c = \text{Cov}(\bar{X}_n, D)$ et $v = \mathbf{V}(D)$. D'après 14.a) et 13.b), pour tout réel α , $\mathbf{V}(\psi(\alpha)) \geq \frac{\theta}{n} = \mathbf{V}(\bar{X}_n)$, soit

$$\mathbf{V}(\bar{X}_n) + 2\alpha c + \alpha^2 v \geq \mathbf{V}(\bar{X}_n), \text{ c'est-à-dire } 2\alpha c + \alpha^2 v \geq 0.$$

Pour $\alpha > 0$, on divise par α : $2c + \alpha v \geq 0$, et quand $\alpha \rightarrow 0^+$, $c \geq 0$. Pour $\alpha < 0$, on obtient $2c + \alpha v \leq 0$, et quand $\alpha \rightarrow 0^-$, $c \leq 0$. Donc $c = 0$, c'est-à-dire $\text{Cov}(\bar{X}_n, Y) - \mathbf{V}(\bar{X}_n) = 0$.

Méthode 2 — Calcul direct par 10.a). D'après 8.b), $Z_\theta = \frac{S_n}{\theta} - n = \frac{n}{\theta} \bar{X}_n - n$. **Par 10.a) appliqué à Y (avec $\varphi'(\theta) = 1$),** $1 = \text{Cov}(Y, Z_\theta) = \frac{n}{\theta} \text{Cov}(Y, \bar{X}_n)$.

Comparaison : la méthode 1 suit la déduction suggérée et n'utilise que la minimalité de la variance de $\bar{X}_n$; la méthode 2 est plus rapide mais exploite le lien particulier entre Z_θ et $\bar{X}_n$.

Conclusion

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

14.c) Méthode : développer la variance d'une différence, puis utiliser la caractérisation des variables de variance nulle.

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

Soit Y un estimateur efficace de θ . Sa variance est minimale parmi les estimateurs réguliers, donc $\mathbf{V}(Y) \leq \mathbf{V}(\bar{X}_n) = \frac{\theta}{n}$; d'après 13.b), $\mathbf{V}(Y) \geq \frac{\theta}{n}$. Donc $\mathbf{V}(Y) = \mathbf{V}(\bar{X}_n)$ et $\mathbf{V}(Y - \bar{X}_n) = 0$: la variable $Y - \bar{X}_n$ est presque sûrement égale à la constante

$E(Y - \bar{X}_n) = \theta - \theta = 0$. Ainsi $Y = \bar{X}_n$ presque sûrement. Si Y_1 et Y_2 sont deux estimateurs efficaces, on a donc $Y_1 = \bar{X}_n = Y_2$ presque sûrement.

Conclusion

$V(Y - \bar{X}_n) = V(Y) - V(\bar{X}_n)$, et un estimateur efficace de θ est presque sûrement égal à $\bar{X}_n$: il est presque sûrement unique.

15.a) Méthode : développer le carré.

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

Conclusion

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

15.b) Méthode : linéarité de l'espérance et formule de Kœnig-Huygens.

$$E(X_i^2) = \theta + \theta^2 \text{ (5.a)} \text{ et } E(\bar{X}_n^2) = V(\bar{X}_n) + E(\bar{X}_n)^2 = \frac{\theta}{n} + \theta^2. \text{ Donc}$$

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

Conclusion

W_n est un estimateur sans biais de θ .

15.c) Méthode : majorer W_n^2 par une variable admettant une espérance.

D'après 15.a),

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

puisque $n\overline{X}_n^2 \geq 0$. Par l'inégalité

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

(inégalité de Cauchy-Schwarz dans $\mathbb{R}^n$), on obtient

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

Le membre de droite admet une espérance, car chaque X_i admet un moment d'ordre 4 (5.c)). Par comparaison de variables aléatoires positives, W_n^2 admet une espérance.

Conclusion

W_n admet un moment d'ordre 2, donc une variance.

15.d) Méthode : loi faible des grands nombres et opérations sur les convergences en probabilité; on démontre d'abord le résultat proposé.

Démonstration du résultat proposé. Posons $D_n = Y_n - Z_n$ et $d = y - z$. Comme $-Z_n$ converge en probabilité vers $-z$ (composition par $x \mapsto -x$, continue), $D_n = Y_n + (-Z_n)$ converge en probabilité vers d (somme). Écrivons $a_n D_n = a D_n + (a_n - a) D_n$. D'une part, $a D_n$ converge en probabilité vers ad (composition par $x \mapsto ax$). D'autre part, soit $\varepsilon > 0$. Comme $|D_n| \leq |D_n - d| + |d|$, si $|D_n - d| < 1$, alors $|(a_n - a) D_n| \leq |a_n - a|(1 + |d|)$. Or $a_n \rightarrow a$, donc il existe un rang n_0 tel que $|a_n - a|(1 + |d|) < \varepsilon$ pour $n \geq n_0$. Pour $n \geq n_0$, on a donc l'inclusion $\{|(a_n - a) D_n| \geq \varepsilon\} \subset \{|D_n - d| \geq 1\}$, d'où $P(|(a_n - a) D_n| \geq \varepsilon) \leq P(|D_n - d| \geq 1) \xrightarrow{n \rightarrow +\infty} 0$. Ainsi $(a_n - a) D_n$ converge en probabilité vers 0, et par somme $a_n D_n$ converge en probabilité vers $ad = a(y - z)$.

Suite $(\overline{X}_n)$: les X_i sont indépendantes, de même loi, d'espérance θ et admettent une variance; par la loi faible des grands nombres, $\overline{X}_n$ converge en probabilité vers θ , pour tout $\theta > 0$. (On peut aussi invoquer 12.c) : $\overline{X}_n$ est sans biais et $\mathbf{V}(\overline{X}_n) = \frac{\theta}{n} \rightarrow 0$.)

Suite (W_n) : d'après 15.a), $W_n = a_n(Y_n - Z_n)$ avec $a_n = \frac{n}{n-1} \xrightarrow{n \rightarrow +\infty} 1$,

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

et $Z_n = \overline{X}_n^2$. Les variables X_i^2 sont indépendantes (fonctions de variables indépendantes), de même loi, d'espérance $\theta + \theta^2$, et admettent une variance puisque X admet un moment d'ordre 4 : par la loi faible des grands nombres, Y_n converge en probabilité vers $\theta + \theta^2$. Par composition par $x \mapsto x^2$, continue, Z_n converge en probabilité vers θ^2 . Le résultat démontré donne la convergence en probabilité de W_n vers $1 \times (\theta + \theta^2 - \theta^2) = \theta$.

Conclusion

les suites $(\overline{X}_n)_{n \geq 1}$ et $(W_n)_{n \geq 2}$ sont deux suites convergentes d'estimateurs de θ .

16) Méthode : comparer les dispersions et les centres des trois estimateurs, en s'appuyant sur les questions précédentes.

Valeur de θ : les trois estimateurs ont des distributions centrées près de la même valeur, et l'histogramme le plus étroit (figure G) est centré en 3 ; on lit donc $\theta \approx 3$ (valeur lue sur la figure).

Dispersions : $V(\overline{X}_{50}) = \frac{\theta}{50}$, soit un écart-type $\sqrt{3/50} \approx 0,245$. D'après le théorème limite central, la loi de $\overline{X}_{50}$ est proche d'une loi normale de cet écart-type, dont la densité maximale vaut $\frac{1}{0,245\sqrt{2\pi}} \approx 1,63$: c'est la hauteur de la figure G. Les figures E et F, nettement plus étalées et environ deux fois et demie moins hautes, correspondent à W_n et W'_n . C'est cohérent avec 13.b) et 14.c) : $\overline{X}_n$ est l'estimateur efficace, et un estimateur régulier qui n'est pas presque sûrement égal à $\overline{X}_n$ a une variance strictement plus grande que $\frac{\theta}{n}$.

Distinction entre E et F : $W'_n = \frac{n-1}{n}W_n = 0,98 W_n$, avec $W_n \geq 0$, donc $W'_n \leq W_n$: l'histogramme de W'_n est celui de W_n légèrement contracté vers 0, donc légèrement décalé vers la gauche, avec une hauteur presque identique (multipliée par $\frac{1}{0,98}$). De plus $E(W_n) = \theta = 3$ (15.b)) alors que $E(W'_n) = 0,98 \theta = 2,94$. La figure F, décalée vers la

gauche par rapport à E, correspond donc à W'_n .

Vérification par simulation : le code suivant a été exécuté. La commande `np.mean(X, axis=0)` calcule la moyenne de chaque colonne, c'est-à-dire de chacun des N échantillons ; dans $X - \text{Xbar}$, le tableau `Xbar` de taille N est soustrait à chaque ligne de `X`, de sorte que la colonne j de $X - \text{Xbar}$ contient les écarts $X_{i,j} - \bar{X}_{n,j}$.

Python

```
import numpy as np

import numpy.random as rd

N = 100000

n = 50

theta = 3

X = rd.poisson(theta, (n, N))

Xbar = np.mean(X, axis=0)

W = np.sum((X - Xbar)**2, axis=0) / (n - 1)

Wp = np.sum((X - Xbar)**2, axis=0) / n

print(np.mean(Xbar), np.std(Xbar), np.mean(W), np.std(W), np.mean(Wp), np.std(Wp))
```

Il affiche des moyennes empiriques voisines de 3,00 pour `Xbar` et `W` et de 2,94 pour `Wp`, avec des écarts-types empiriques voisins de 0,245 pour `Xbar`, 0,654 pour `W` et 0,641 pour `Wp`. Une loi normale d'écart-type 0,65 a une densité maximale d'environ 0,61, ce qui est cohérent avec la hauteur lue sur E et F (environ 0,65, la distribution de W_n étant un peu dissymétrique).

Conclusion

la figure G correspond à $\bar{X}_n$, la figure E à W_n et la figure F à W'_n .

Relevé des adaptations

3.e) Code adapté en Python. L'original demande un code Scilab utilisant la commande `disp` et suggère la commande `ones(n,p)`. L'énoncé adapté demande un code Python affichant les listes au moyen de `print` et présente `np.ones((n, p))` de la bibliothèque `numpy`, avec son instruction d'importation et son rôle. Justification : seul Python est au programme.

7) Précision de notation, sans changement du contenu : l'énoncé adapté explique que $\partial_1(g)(\theta, k)$ désigne la dérivée au point θ de la fonction $\theta \mapsto g(\theta, k)$, l'entier k étant fixé, et que la même convention vaut pour F et G dans la partie 3. Justification : la seconde variable étant entière, il s'agit de dérivées de fonctions d'une variable réelle ; la notation de l'original est conservée.

11.a) Code adapté en Python. La fonction Scilab `function ech=Z_th(n, theta) ... endfunction`, la commande `grand(n,N,'poi',theta)` et les instructions `sum(M,'r')` et `sum(M,'c')` sont remplacées par une fonction Python `def Z_th(n, theta)`, par `rd.poisson(theta, (n, N))` (bibliothèque `numpy.random` importée sous le nom `rd`) et par `np.sum(M, axis=0)` et `np.sum(M, axis=1)`, présentées dans l'énoncé avec leur rôle et la taille des tableaux renvoyés. L'énoncé précise que N est une variable préalablement définie. La partie à compléter est conservée ; la réponse attendue passe de `'r'` à `axis=0`. Justification : seul Python est au programme.

11.b) Code adapté en Python. La commande Scilab `histplot` est remplacée par `plt.hist` de la bibliothèque `matplotlib.pyplot`, avec la précision que les histogrammes sont normalisés (aire totale égale à 1), ce que faisait `histplot`. La figure des quatre histogrammes est reproduite à partir de l'original dans le fichier compagnon `sujethec_escp_maths2_2021_figures_A_D.png`, accompagnée d'une description des valeurs lues.

16) La figure des trois histogrammes est reproduite à partir de l'original dans le fichier compagnon `sujethec_escp_maths2_2021_figures_E_G.png`, accompagnée d'une description des valeurs lues.

Choix de méthode sans modification de l'énoncé : l'inégalité de Cauchy-Schwarz pour la covariance est démontrée dans les rappels avant d'être utilisée ; l'inégalité de 10.b)

(inégalité de Cramér-Rao) n'est jamais supposée connue, elle est établie dans le sujet lui-même ; le théorème limite central n'est utilisé qu'à titre d'interprétation des histogrammes en 11.b) et 16).