

HEC Paris – ESCP BS

Mathématiques approfondies – Mathématiques II

Corrigé du sujet adapté – 2020

Les questions et sous-questions conservent la numérotation de l'énoncé adapté. Pour alléger, lorsque le paramètre θ est fixé, on écrit parfois P , E , V au lieu de P^θ , E^θ , V^θ .

Partie I – Développement en série

Partie I – Deux résultats sur les séries entières

Calcul de $\sum_{k \geq 1} x^k/k$ et unicité des coefficients d'une somme de série $\sum a_k x^k$ nulle sur $]0, c]$.

Corrigé – Partie I.

1. La série $\sum x^k/k$, pour $x \in]0, 1[$.

a) Pour $k \geq 1$, $0 \leq \frac{x^k}{k} \leq x^k$, et la série géométrique $\sum x^k$ converge car $0 < x < 1$. Par comparaison de séries à termes positifs,

$$\sum_{k \geq 1} \frac{x^k}{k} \text{ converge.}$$

b) Pour $t \in]0, 1[$, $t \neq 1$ donc $\sum_{k=0}^{m-1} t^k = \frac{1-t^m}{1-t}$, d'où

$$\frac{t^m}{1-t} + \sum_{k=0}^{m-1} t^k = \frac{t^m + 1 - t^m}{1-t} = \frac{1}{1-t}.$$

c) Pour $t \in [0, x]$, $0 \leq \frac{t^m}{1-t} \leq \frac{t^m}{1-x}$. Par croissance de l'intégrale ($0 < x$),

$$0 \leq \int_0^x \frac{t^m}{1-t} dt \leq \frac{1}{1-x} \int_0^x t^m dt = \frac{x^{m+1}}{(m+1)(1-x)} \xrightarrow{m \rightarrow +\infty} 0.$$

Par encadrement,

$$\lim_{m \rightarrow +\infty} \int_0^x \frac{t^m}{1-t} dt = 0.$$

d) Les fonctions de b) sont continues sur $[0, x]$; on intègre l'égalité de b) (valable aussi en $t = 0$) sur $[0, x]$:

$$-\ln(1-x) = \int_0^x \frac{t^m}{1-t} dt + \sum_{k=0}^{m-1} \frac{x^{k+1}}{k+1} = \int_0^x \frac{t^m}{1-t} dt + \sum_{k=1}^m \frac{x^k}{k}.$$

En faisant tendre m vers $+\infty$ grâce à c) :

$$\sum_{k=1}^{+\infty} \frac{x^k}{k} = -\ln(1-x).$$

2. Unicité du développement.

- a) Pour $x \in [-c, c]$ et $k \in \mathbb{N}$, $|a_k x^k| \leq |a_k| c^k$. La série $\sum |a_k| c^k$ converge, donc par comparaison $\sum a_k x^k$ converge absolument, donc converge.

$$f \text{ est bien définie sur } [-c, c].$$

- b) La série définissant M_m converge car $|a_k| c^{k-m-1} = c^{-(m+1)} |a_k| c^k$. Pour $x \in [-c, c]$ et $N \geq m+1$, par l'inégalité triangulaire,

$$\left| \sum_{k=m+1}^N a_k x^k \right| \leq \sum_{k=m+1}^N |a_k| |x|^{m+1} |x|^{k-m-1} \leq |x|^{m+1} \sum_{k=m+1}^N |a_k| c^{k-m-1},$$

$$\left| \sum_{k=m+1}^N a_k x^k \right| \leq M_m |x|^{m+1},$$

puisque $|x| \leq c$ et que les termes de M_m sont positifs. En passant à la limite quand $N \rightarrow +\infty$:

$$\left| \sum_{k=m+1}^{+\infty} a_k x^k \right| \leq M_m |x|^{m+1}.$$

- c) Pour $x \in [-c, c] \setminus \{0\}$, d'après b),

$$\left| \frac{f(x) - a_0 - \sum_{k=1}^m a_k x^k}{x^m} \right| = \frac{\left| \sum_{k=m+1}^{+\infty} a_k x^k \right|}{|x|^m} \leq M_m |x| \xrightarrow{x \rightarrow 0} 0.$$

$$f(x) = a_0 + \sum_{k=1}^m a_k x^k + o(x^m) \quad (x \rightarrow 0).$$

- d) Supposons f nulle sur $]0, c]$ et (a_k) non nulle ; soit p le plus petit indice tel que $a_p \neq 0$.

- Si $p \geq 1$, c) avec $m = p$ donne $f(x) = a_p x^p + o(x^p)$, donc $\frac{f(x)}{x^p} \rightarrow a_p$ quand $x \rightarrow 0^+$.
- Si $p = 0$, c) avec $m = 1$ donne $f(x) = a_0 + a_1 x + o(x)$, donc $f(x) \rightarrow a_0$ quand $x \rightarrow 0^+$.

Dans les deux cas, la fonction considérée est nulle sur $]0, c]$, donc sa limite a_p est nulle : contradiction.

$$\text{Si } f \text{ est nulle sur }]0, c], \text{ alors } a_k = 0 \text{ pour tout } k \in \mathbb{N}.$$

Partie II – Estimateur du maximum de vraisemblance

Partie II – Loi géométrique

$X \hookrightarrow \mathcal{G}(\theta)$; estimateurs $\bar{X}$, $T = \frac{1}{n} \sum \frac{1}{X_i}$ et $U = \sum \frac{n}{X_i}$.

Corrigé – Partie II.

3. L'estimateur $\bar{X}$.

a) $X(\Omega) = \mathbb{N}^*$ et $P^\theta([X = x]) = (1 - \theta)^{x-1}\theta$:

$$X \hookrightarrow \mathcal{G}(\theta) \text{ (loi géométrique de paramètre } \theta).$$

b) Chaque X_i suit la loi $\mathcal{G}(\theta)$, donc admet une espérance égale à $1/\theta$. Par linéarité,

$$E^\theta(\bar{X}) = \frac{1}{n} \sum_{i=1}^n \frac{1}{\theta} = \frac{1}{\theta}.$$

$\bar{X}$ est une statistique (fonction de $X_1, \dots, X_n$ ne dépendant pas de θ) :

$$\bar{X} \text{ est un estimateur sans biais de } 1/\theta.$$

c) L'estimateur étant sans biais, son risque quadratique est sa variance. Les X_i sont indépendantes et admettent une variance $\frac{1-\theta}{\theta^2}$, donc

$$R_{1/\theta}(\bar{X}) = V^\theta(\bar{X}) = \frac{1}{n^2} \sum_{i=1}^n V^\theta(X_i) = \frac{1-\theta}{n\theta^2}.$$

4. L'estimateur T .

a) Par le théorème de transfert, $\frac{1}{X}$ admet une espérance si et seulement si la série à termes positifs $\sum_{x \geq 1} \frac{1}{x} (1 - \theta)^{x-1} \theta$ converge. Avec $1 - \theta \in]0, 1[$ et la question 1.d :

$$\sum_{x=1}^{+\infty} \frac{1}{x} (1 - \theta)^{x-1} \theta = \frac{\theta}{1 - \theta} \sum_{x=1}^{+\infty} \frac{(1 - \theta)^x}{x} = \frac{\theta}{1 - \theta} (-\ln(1 - (1 - \theta))) = \frac{\theta \ln \theta}{\theta - 1}.$$

D'après (2), chaque $\frac{1}{X_i}$ a même loi que $\frac{1}{X}$; par linéarité,

$$E^\theta(T) = \frac{1}{n} \sum_{i=1}^n E^\theta\left(\frac{1}{X_i}\right) = \frac{\theta \ln \theta}{\theta - 1}.$$

b) T est une statistique, donc un estimateur de θ , de biais

$$b_\theta(T) = E^\theta(T) - \theta = \frac{\theta \ln \theta - \theta(\theta - 1)}{\theta - 1} = \theta \frac{\ln \theta - (\theta - 1)}{\theta - 1}.$$

Par stricte concavité de $\ln$, $\ln \theta < \theta - 1$ pour $\theta \neq 1$; comme $\theta - 1 < 0$ et $\theta > 0$,

$$b_\theta(T) > 0.$$

5. Maximum de vraisemblance.

a) Pour $(x_1, \dots, x_n) \in (\mathbb{N}^*)^n$,

$$L(x_1, \dots, x_n, \theta) = \prod_{i=1}^n (1 - \theta)^{x_i - 1} \theta = \theta^n (1 - \theta)^{\sum_i x_i - n} > 0,$$

donc

$$\ln L(x_1, \dots, x_n, \theta) = n \ln \theta - \left(n - \sum_{i=1}^n x_i \right) \ln(1 - \theta).$$

b) Notons $s = \sum_i x_i$. Comme les x_i sont des entiers ≥ 1 non tous égaux à 1, $s > n$. La fonction $\ell : \theta \mapsto n \ln \theta + (s - n) \ln(1 - \theta)$ est dérivable sur $]0, 1[$ et

$$\ell'(\theta) = \frac{n}{\theta} - \frac{s - n}{1 - \theta} = \frac{n(1 - \theta) - (s - n)\theta}{\theta(1 - \theta)} = \frac{n - s\theta}{\theta(1 - \theta)}.$$

Ainsi $\ell' > 0$ sur $]0, n/s[$ et $\ell' < 0$ sur $]n/s, 1[$, avec $n/s \in]0, 1[$: ℓ admet un unique point de maximum, n/s . Comme $L = \exp \circ \ell$ avec $\exp$ strictement croissante,

$$\theta = \frac{n}{\sum_{i=1}^n x_i} \text{ est l'unique valeur maximisant } L(x_1, \dots, x_n, \theta).$$

6. L'estimateur U .

a) La fonction $g : t \mapsto \frac{1}{t}$ est de classe $\mathcal{C}^2$ sur $\mathbb{R}_+^*$, avec $g'(t) = -\frac{1}{t^2}$ et $g''(t) = \frac{2}{t^3}$. La formule de Taylor avec reste intégral à l'ordre 1 entre $a = \frac{1}{\theta}$ et $b = \frac{k}{n}$ (deux réels de $\mathbb{R}_+^*$) donne

$$g(b) = g(a) + g'(a)(b - a) + \int_a^b (b - t)g''(t) dt,$$

soit, avec $g(a) = \theta$ et $g'(a) = -\theta^2$:

$$\frac{n}{k} = \theta - \theta^2 \left(\frac{k}{n} - \frac{1}{\theta} \right) + \int_{1/\theta}^{k/n} \left(\frac{k}{n} - t \right) \frac{2}{t^3} dt.$$

b) Notons $\Sigma = \sum_{i=1}^n X_i$ et I_k l'intégrale de a). Comme chaque $X_i \geq 1$, $\Sigma(\Omega) \subset \llbracket n, +\infty \rrbracket$, et $U = \frac{n}{\Sigma} \in]0, 1]$ est bornée, donc admet une espérance. Par le théorème de transfert, la série $\sum_{k \geq n} \frac{n}{k} P([\Sigma = k])$ converge et vaut $E^\theta(U)$. D'après a),

$$\frac{n}{k} P([\Sigma = k]) = \theta P([\Sigma = k]) - \theta^2 \left(\frac{k}{n} - \frac{1}{\theta} \right) P([\Sigma = k]) + P([\Sigma = k]) I_k.$$

Les événements $([\Sigma = k])_{k \geq n}$ forment un système complet, donc $\sum_{k \geq n} P([\Sigma = k]) = 1$; Σ admet une espérance $E(\Sigma) = \frac{n}{\theta}$, donc $\sum_{k \geq n} \frac{k}{n} P([\Sigma = k]) = \frac{1}{\theta}$. Les deux premières séries convergent, donc la troisième aussi, et

$$E^\theta(U) = \theta - \theta^2 \left(\frac{1}{\theta} - \frac{1}{\theta} \right) + \sum_{k=n}^{+\infty} P([\Sigma = k]) I_k = \theta + \sum_{k=n}^{+\infty} P([\Sigma = k]) I_k.$$

U est une statistique, donc un estimateur de θ , et

$$b_\theta(U) = \sum_{k=n}^{+\infty} P\left(\left[\sum_{i=1}^n X_i = k\right]\right) \int_{1/\theta}^{k/n} \left(\frac{k}{n} - t\right) \frac{2}{t^3} dt.$$

- c) Si $\frac{k}{n} \geq \frac{1}{\theta}$, l'intégrande de I_k est positif sur $[\frac{1}{\theta}, \frac{k}{n}]$ et les bornes sont dans l'ordre croissant ; si $\frac{k}{n} < \frac{1}{\theta}$, $I_k = \int_{k/n}^{1/\theta} (t - \frac{k}{n}) \frac{2}{t^3} dt$ avec un intégrande positif. Donc $I_k \geq 0$; de plus, si $\frac{k}{n} \neq \frac{1}{\theta}$, l'intégrande est continu, positif et non identiquement nul sur un segment non réduit à un point, donc $I_k > 0$.

Pour $k \geq n$, par indépendance et (3),

$$\begin{aligned} P([\Sigma = k]) &\geq P([X_1 = k - n + 1] \cap [X_2 = 1] \cap \dots \cap [X_n = 1]) \\ &= P([X = k - n + 1])P([X = 1])^{n-1} > 0. \end{aligned}$$

Parmi $k = n$ et $k = n + 1$, l'un au moins vérifie $\frac{k}{n} \neq \frac{1}{\theta}$, donc la somme de termes positifs de b) contient un terme strictement positif :

$$b_\theta(U) > 0 \text{ pour tout } \theta \in]0, 1[.$$

7. Convergence des suites d'estimateurs.

- Les variables $\frac{1}{X_i}$ sont indépendantes, de même loi, bornées (à valeurs dans $]0, 1[$) donc admettent une variance. Par la loi faible des grands nombres,

$$T_n \xrightarrow[n \rightarrow +\infty]{P} E\left(\frac{1}{X}\right) = \frac{\theta \ln \theta}{\theta - 1} \neq \theta \quad (\text{d'après 4.b}).$$

- Les X_i sont indépendantes, de même loi et admettent une variance ; par la loi faible des grands nombres, $\bar{X}_n \xrightarrow{P} \frac{1}{\theta}$. La fonction h définie par $h(x) = \frac{1}{x}$ si $x \geq 1$ et $h(x) = 1$ si $x < 1$ est continue sur $\mathbb{R}$, et $U_n = h(\bar{X}_n)$ car $\bar{X}_n \geq 1$. Par composition par une fonction continue, $U_n \xrightarrow{P} h(1/\theta) = \theta$ (car $1/\theta > 1$).

(U_n) converge en probabilité vers θ : c'est une suite d'estimateurs convergente ;
 (T_n) converge en probabilité vers $\frac{\theta \ln \theta}{\theta - 1} \neq \theta$: elle n'est pas convergente.

Partie III – Statistique exhaustive

Partie III – Échantillon de Bernoulli

$X_i \hookrightarrow \mathcal{B}(\theta)$, $S = \sum X_i$; loi conditionnelle sachant $[S = k]$ et simulations.

Corrigé – Partie III.

8. Exhaustivité de S .

- a) Pour $x_i \in \{0, 1\}$, $P^\theta([X_i = x_i]) = \theta^{x_i}(1 - \theta)^{1-x_i}$ (les deux cas $x_i = 1$ et $x_i = 0$ se vérifient directement). Donc

$$L(x_1, \dots, x_n, \theta) = \prod_{i=1}^n \theta^{x_i}(1 - \theta)^{1-x_i} = \theta^{\sum_{i=1}^n x_i} (1 - \theta)^{\sum_{i=1}^n (1-x_i)}.$$

- b) Avec $s(x_1, \dots, x_n) = \sum_i x_i$, on a $\sum_i (1 - x_i) = n - s(x_1, \dots, x_n)$. En posant $g(u, \theta) = \theta^u (1 - \theta)^{n-u} \geq 0$ pour $u \in s(B^n) = \llbracket 0, n \rrbracket$ et $h = 1$ sur B^n , on obtient $L(x_1, \dots, x_n, \theta) = g(s(x_1, \dots, x_n), \theta)h(x_1, \dots, x_n)$:

S est exhaustive.

9. Loi conditionnelle sachant $[S = k]$.

- a) Les X_i étant mutuellement indépendantes de loi $\mathcal{B}(\theta)$, $S \hookrightarrow \mathcal{B}(n, \theta)$, donc $P([S = k]) = \binom{n}{k} \theta^k (1 - \theta)^{n-k} > 0$. Notons $A = [X_1 = x_1] \cap \dots \cap [X_n = x_n]$; on a $A \subset [S = \sum_i x_i]$.
- Si $\sum_i x_i \neq k$, $A \cap [S = k] = \emptyset$ et $P_{[S=k]}(A) = 0$.
 - Si $\sum_i x_i = k$, $A \cap [S = k] = A$ et

$$P_{[S=k]}(A) = \frac{L(x_1, \dots, x_n, \theta)}{P([S = k])} = \frac{\theta^k (1 - \theta)^{n-k}}{\binom{n}{k} \theta^k (1 - \theta)^{n-k}} = \frac{1}{\binom{n}{k}}.$$

$$P_{[S=k]}^\theta([X_1 = x_1] \cap \dots \cap [X_n = x_n]) = \begin{cases} 1/\binom{n}{k} & \text{si } \sum_i x_i = k, \\ 0 & \text{sinon,} \end{cases}$$

qui ne dépend pas de θ : sachant $[S = k]$, $(X_1, \dots, X_n)$ suit la loi uniforme sur les n -uplets de $\{0, 1\}^n$ comportant exactement k coordonnées égales à 1.

- b) L'événement $[X_1 = 1]$ est la réunion disjointe des A associés aux $(1, x_2, \dots, x_n)$. Ceux de probabilité conditionnelle non nulle vérifient $x_2 + \dots + x_n = k - 1$: il y en a $\binom{n-1}{k-1}$ (aucun si $k = 0$). Donc

$$P_{[S=k]}^\theta([X_1 = 1]) = \frac{\binom{n-1}{k-1}}{\binom{n}{k}} = \frac{(n-1)! k! (n-k)!}{(k-1)! (n-k)! n!}, \quad P_{[S=k]}^\theta([X_1 = 1]) = \frac{k}{n}.$$

(Le cas $k = 0$ est immédiat : la probabilité vaut $0 = \frac{0}{n}$.) Le même calcul donne $P_{[S=k]}([X_i = 1]) = \frac{k}{n}$ pour tout i .

10. Simulation.

- a) $\mathbf{U}$ est une matrice $n \times N$ dont les coefficients sont des simulations indépendantes de la loi $\mathcal{B}(\theta)$, $\theta = 0,3$.

Une colonne de $\mathbf{U}$ simule le vecteur $(X_1, \dots, X_n)$ sous P^θ , $\theta = 0,3$, $n = 10$.

- b) $\mathbf{S}[j]$ est la somme de la colonne j de $\mathbf{U}$, c'est-à-dire une simulation de S . Les colonnes de $\mathbf{Y}$ sont les colonnes de $\mathbf{U}$ pour lesquelles cette somme vaut $k = 4$: ce sont les simulations de $(X_1, \dots, X_n)$ pour lesquelles l'événement $[S = k]$ est réalisé. Parmi des simulations indépendantes, celles qui réalisent $[S = k]$ sont des simulations indépendantes de la loi conditionnelle sachant $[S = k]$ (méthode de rejet).

Les colonnes de $\mathbf{Y}$ simulent la loi conditionnelle de $(X_1, \dots, X_n)$ sachant $[S = 4]$.

- c) $\mathbf{M}[i]$ est la moyenne des coefficients de la ligne i de $\mathbf{Y}$, qui valent 0 ou 1 : c'est la fréquence de réalisation de $[X_{i+1} = 1]$ parmi les simulations conditionnées par $[S = 4]$. D'après la loi faible des grands nombres, elle est proche de $P_{[S=4]}([X_{i+1} = 1]) = \frac{4}{10} = 0,4$ (question 9.b).

Tous les coefficients de $\mathbf{M}$ sont proches de $k/n = 0,4$, et non de $\theta = 0,3$.

Cela illustre que la loi conditionnelle sachant $[S = k]$ ne dépend pas de θ . On remarque aussi que la somme des coefficients de $\mathbf{M}$ vaut exactement 4 (aux arrondis près, 4,0000001) : chaque colonne de $\mathbf{Y}$ a pour somme 4.

11. La matrice $\mathbf{C}$.

Notons nb le nombre de colonnes de $\mathbf{Y}$, $Y = (y_{i,j})$, et $\text{nb} \cdot C = Y^t Y$.

- a) $Y \in \mathcal{M}_{n,\text{nb}}(\mathbb{R})$ donc $C = \frac{1}{\text{nb}} Y^t Y \in \mathcal{M}_n(\mathbb{R})$: c'est une matrice carrée 10×10 . Comme $y_{0,j} \in \{0, 1\}$, $y_{0,j}^2 = y_{0,j}$ et

$$c[0,0] = \frac{1}{\text{nb}} \sum_j y_{0,j}^2 = \frac{1}{\text{nb}} \sum_j y_{0,j} = \mathbf{M}[0], \quad \boxed{c[0,0] = 0,4019917.}$$

Plus généralement, les coefficients diagonaux de $\mathbf{C}$ sont ceux de $\mathbf{M}$.

- b) $c[0,1] = \frac{1}{\text{nb}} \sum_j y_{0,j} y_{1,j}$ est la fréquence, parmi les simulations conditionnées, de l'événement $[X_1 = 1] \cap [X_2 = 1]$. Par la loi faible des grands nombres, elle approche

$$P_{[S=k]}([X_1 = 1] \cap [X_2 = 1]) = \frac{\binom{n-2}{k-2}}{\binom{n}{k}} = \frac{k(k-1)}{n(n-1)} = \frac{4 \times 3}{10 \times 9}.$$

(même dénombrement qu'en 9.b). Il en va de même pour tout couple (i, l) avec $i \neq l$:

$$c[i,1] \approx \frac{2}{15} \approx 0,133 \quad (i \neq l).$$

- c) La somme de tous les coefficients de C vaut

$$\sum_{i,l} c[i,1] = \frac{1}{\text{nb}} \sum_j \left(\sum_i y_{i,j} \right) \left(\sum_l y_{l,j} \right) = \frac{1}{\text{nb}} \sum_j k^2 = k^2,$$

car chaque colonne de $\mathbf{Y}$ a pour somme k .

La somme des coefficients de $\mathbf{C}$ vaut exactement $k^2 = 16$.

(Cohérence : $10 \times 0,4 + 90 \times \frac{2}{15} = 4 + 12 = 16$.)

Partie IV – Inégalité de Rao–Blackwell

Partie IV – Amélioration d'un estimateur sans biais

S exhaustive ; $R = E_{[S=.]}(T)$ est sans biais et de variance inférieure à celle de T .

Corrigé – Partie IV.

Préliminaires. B^n est dénombrable, donc $T = t(X_1, \dots, X_n)$ et S sont des variables discrètes. Pour $u \in s(B^n)$, écrivons $u = s(y_1, \dots, y_n)$; alors

$$P^\theta([S = u]) \geq P^\theta\left(\bigcap_i [X_i = y_i]\right) = \prod_i P^\theta([X = y_i]) > 0$$

d'après (1), (2) et (3). Pour $x = (x_1, \dots, x_n) \in B^n$ on note $A_x = \bigcap_i [X_i = x_i]$.

12. Espérance conditionnelle.

- a) Les événements A_x , $x \in B^n$, forment un système complet d'événements et $T = t(x)$ sur A_x . T admet une espérance, donc la famille $(t(x)|P(A_x))_{x \in B^n}$ est sommable.

Comme $P_{[S=u]}(A_x) \leq \frac{P(A_x)}{P([S=u])}$, la famille $(t(x)|P_{[S=u]}(A_x))_x$ est aussi sommable, et T admet une espérance pour $P_{[S=u]}^\theta$:

$$E_{[S=u]}^\theta(T) = \sum_{x \in B^n} t(x) P_{[S=u]}^\theta(A_x).$$

Chaque $P_{[S=u]}^\theta(A_x)$ est indépendant de θ (résultat admis), donc

$$E_{[S=u]}^\theta(T) \text{ existe et ne dépend pas de } \theta.$$

- b) Les événements $[S = u]$, $u \in s(B^n)$, sont deux à deux disjoints (une variable ne prend qu'une valeur en chaque ω), et pour tout ω , $S(\omega) = s(X_1(\omega), \dots, X_n(\omega)) \in s(B^n)$, donc leur réunion est Ω . L'ensemble $s(B^n)$ est dénombrable comme image d'un ensemble dénombrable.

$$([S = u])_{u \in s(B^n)} \text{ est un système complet d'événements.}$$

13. L'estimateur R .

Notons $\rho(u) = E_{[S=u]}(T)$ pour $u \in s(B^n)$; ainsi $r(x_1, \dots, x_n) = \rho(s(x_1, \dots, x_n))$ et $R = \rho(S)$.

- a) r ne dépend pas de θ (question 12.a), donc R est une statistique : c'est un estimateur de θ . On a $|\rho(u)| \leq E_{[S=u]}(|T|)$ et, par la formule de l'espérance totale appliquée à $|T|$ avec le système complet $([S = u])_{u \in s(B^n)}$, la série $\sum_u E_{[S=u]}(|T|)P([S = u])$ converge et vaut $E(|T|)$. Donc $\sum_u \rho(u)P([S = u])$ converge absolument : $R = \rho(S)$ admet une espérance (théorème de transfert) et, par la formule de l'espérance totale,

$$E^\theta(R) = \sum_{u \in s(B^n)} E_{[S=u]}(T) P^\theta([S = u]) = E^\theta(T) = \theta.$$

$$R \text{ est un estimateur sans biais de } \theta.$$

b) Fixons θ . Pour une variable Z admettant un moment d'ordre 2 pour $P_{[S=u]}$, $E_{[S=u]}(Z^2) - (E_{[S=u]}(Z))^2 = E_{[S=u]}((Z - E_{[S=u]}(Z))^2) \geq 0$, ce qui justifie les inégalités de l'énoncé. Par linéarité, $E_{[S=u]}(T - \theta) = \rho(u) - \theta$, donc

$$(\rho(u) - \theta)^2 \leq E_{[S=u]}((T - \theta)^2).$$

T admet une variance, donc $(T - \theta)^2$ admet une espérance et, par la formule de l'espérance totale, $\sum_u E_{[S=u]}((T - \theta)^2)P([S = u])$ converge et vaut $E((T - \theta)^2) = V(T)$ (car $E(T) = \theta$). Par comparaison de séries à termes positifs, $\sum_u (\rho(u) - \theta)^2 P([S = u])$ converge : $(R - \theta)^2$ admet une espérance, donc $R - \theta$, puis R , admettent un moment d'ordre 2 : R admet une variance. Enfin, puisque $E(R) = \theta$,

$$\begin{aligned} V(R) &= E((R - \theta)^2) = \sum_u (\rho(u) - \theta)^2 P([S = u]) \\ &\leq \sum_u E_{[S=u]}((T - \theta)^2) P([S = u]) = V(T). \end{aligned}$$

$\forall \theta \in \Theta, \quad V^\theta(R) \leq V^\theta(T).$

Partie V – Un exemple d'estimateur sans biais optimal

Partie V – Échantillon de Poisson

$X_i \hookrightarrow \mathcal{P}(\theta)$: exhaustivité de $S = \sum X_i$, lois conditionnelles, optimalité de $\bar{X}$.

Corrigé – Partie V.

14. Le cas poissonnien.

a) Pour $(x_1, \dots, x_n) \in \mathbb{N}^n$,

$$L(x_1, \dots, x_n, \theta) = \prod_{i=1}^n e^{-\theta} \frac{\theta^{x_i}}{x_i!} = e^{-n\theta} \theta^{\sum_i x_i} \cdot \prod_{i=1}^n \frac{1}{x_i!}.$$

Avec $s(x_1, \dots, x_n) = \sum_i x_i$, $g(u, \theta) = e^{-n\theta} \theta^u \geq 0$ et $h(x_1, \dots, x_n) = \prod_i \frac{1}{x_i!} \geq 0$, on a (5) :

$$S = \sum_{i=1}^n X_i \text{ est exhaustive.}$$

b) Les X_i étant indépendantes de loi $\mathcal{P}(\theta)$, $S \hookrightarrow \mathcal{P}(n\theta)$ (stabilité des lois de Poisson). Comme en 9.a, si $\sum_i x_i \neq u$, la probabilité conditionnelle est nulle ; si $\sum_i x_i = u$,

$$P_{[S=u]}^\theta \left(\bigcap_i [X_i = x_i] \right) = \frac{e^{-n\theta} \theta^u \prod_i \frac{1}{x_i!}}{e^{-n\theta} \frac{(n\theta)^u}{u!}} = \frac{u!}{n^u x_1! \cdots x_n!}.$$

$$P_{[S=u]}^\theta \left(\bigcap_{i=1}^n [X_i = x_i] \right) = \begin{cases} \frac{u!}{n^u x_1! \cdots x_n!} & \text{si } \sum_i x_i = u, \\ 0 & \text{sinon,} \end{cases} \quad \text{indépendant de } \theta.$$

c) Soit $i \in \llbracket 1, n \rrbracket$ et $j \in \mathbb{N}$. Notons $S' = \sum_{l \neq i} X_l$: d'après le lemme des coalitions, S' est indépendante de X_i , et $S' \hookrightarrow \mathcal{P}((n-1)\theta)$. Si $j > u$, $[X_i = j] \cap [S = u] = \emptyset$. Si $0 \leq j \leq u$,

$$\begin{aligned} P_{[S=u]}([X_i = j]) &= \frac{P([X_i = j])P([S' = u-j])}{P([S = u])} = \frac{e^{-\theta} \frac{\theta^j}{j!} e^{-(n-1)\theta} \frac{((n-1)\theta)^{u-j}}{(u-j)!}}{e^{-n\theta} \frac{(n\theta)^u}{u!}} \\ &= \binom{u}{j} \left(\frac{1}{n}\right)^j \left(1 - \frac{1}{n}\right)^{u-j}. \end{aligned}$$

$$\text{Pour } P_{[S=u]}^\theta, \text{ chaque } X_i \text{ suit la loi binomiale } \mathcal{B}(u, \frac{1}{n}).$$

Indépendance. Si $u = 0$, toutes les X_i sont presque sûrement nulles pour $P_{[S=0]}$: elles sont (trivialement) indépendantes. Si $u \geq 1$ (et $n \geq 2$) :

$$P_{[S=u]}([X_1 = u] \cap [X_2 = u]) = 0 \quad (\text{car } X_1 + X_2 \leq S = u < 2u),$$

alors que

$$P_{[S=u]}([X_1 = u]) P_{[S=u]}([X_2 = u]) = \frac{1}{n^{2u}} > 0.$$

$$\text{Pour } u \geq 1, X_1, \dots, X_n \text{ ne sont pas indépendantes pour } P_{[S=u]}^\theta.$$

d) *Existence.* Pour $\theta > 0$,

$$\theta e^{n\theta} = \sum_{j=0}^{+\infty} \frac{n^j \theta^{j+1}}{j!} = \sum_{k=1}^{+\infty} \frac{n^{k-1} \theta^k}{(k-1)!} = \sum_{k=1}^{+\infty} \frac{k}{n} \cdot \frac{(n\theta)^k}{k!}.$$

$$\boxed{\varphi_k = \frac{k}{n} \text{ pour tout } k \in \mathbb{N} \text{ convient.}}$$

Unicité. Soit (ψ_k) une autre suite convenable. La série $\sum \psi_k \frac{(2n)^k}{k!}$ converge (cas $\theta = 2$), donc son terme général est borné par un réel M , et

$$\left| \psi_k \frac{n^k}{k!} \right| \cdot 1^k = \left| \psi_k \frac{(2n)^k}{k!} \right| \cdot \frac{1}{2^k} \leq \frac{M}{2^k}.$$

Ainsi $\sum \psi_k \frac{n^k}{k!}$ converge absolument ; de même pour (φ_k) . Posons $a_k = (\varphi_k - \psi_k) \frac{n^k}{k!}$: la série $\sum a_k 1^k$ est absolument convergente et, pour tout $x \in]0, 1]$,

$$f(x) = \sum_{k=0}^{+\infty} a_k x^k = \sum_k \varphi_k \frac{(nx)^k}{k!} - \sum_k \psi_k \frac{(nx)^k}{k!} = x e^{nx} - x e^{nx} = 0.$$

D'après la question 2.d (avec $c = 1$), $a_k = 0$ pour tout k , donc $\psi_k = \varphi_k$.

$$\boxed{\text{La suite } (\varphi_k) = \left(\frac{k}{n}\right)_{k \in \mathbb{N}} \text{ est l'unique suite convenable.}}$$

e) $\bar{X} = \frac{S}{n}$ est un estimateur sans biais de θ ($E(\bar{X}) = \theta$) admettant une variance, de risque quadratique $V(\bar{X}) = \frac{\theta}{n}$.

Soit T un estimateur sans biais de θ . Si T n'admet pas de variance pour une valeur de θ , son risque quadratique n'est pas fini et il n'y a rien à démontrer ; on suppose donc que T admet une variance pour tout θ . D'après 14.a et 14.b, les hypothèses de la partie IV sont satisfaites, avec $s(B^n) = \mathbb{N}$. Soit $R = \rho(S)$ l'estimateur de la question 13. Il est sans biais : pour tout $\theta > 0$,

$$\theta = E^\theta(R) = \sum_{k=0}^{+\infty} \rho(k) e^{-n\theta} \frac{(n\theta)^k}{k!}, \quad \text{soit} \quad \sum_{k=0}^{+\infty} \rho(k) \frac{(n\theta)^k}{k!} = \theta e^{n\theta}.$$

Par l'unicité de 14.d, $\rho(k) = \frac{k}{n}$ pour tout k , c'est-à-dire $R = \frac{S}{n} = \bar{X}$. La question 13.b donne alors, pour tout θ ,

$$R_\theta(\bar{X}) = V^\theta(\bar{X}) = V^\theta(R) \leq V^\theta(T) = R_\theta(T).$$

$$\boxed{\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \text{ est optimal parmi les estimateurs sans biais de } \theta.}$$