



PRÉPASIMPLE

prepasimple.fr

HEC — 2020

Corrigé de mathématiques approfondies

Mathématiques approfondies — Prépa ECG

Annales de concours · Corrigé

Corrigé

Partie I : développements en série

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, la série converge.

Conclusion

la série

$$\sum_{k \geq 1} \frac{x^k}{k}$$

converge.

b) Soit $m \in \mathbb{N}^*$ et $t \in]0, 1[$. Comme $t \neq 1$, la somme géométrique donne

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

donc :

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

Conclusion

l'égalité est vérifiée (elle l'est aussi pour $t = 0$).

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

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

Le majorant tend vers 0 lorsque $m \rightarrow +\infty$.

Conclusion

par encadrement, $\int_0^x \frac{t^m}{1-t} dt \rightarrow 0$ quand $m \rightarrow +\infty$.

- d) Les fonctions de b) sont continues sur $[0, x]$; en intégrant l'égalité de b) sur $[0, x]$:

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

Or $\int_0^x \frac{dt}{1-t} = [-\ln(1-t)]_0^x = -\ln(1-x)$, et

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

Donc

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

qui tend vers $-\ln(1-x)$ d'après c).

Conclusion

pour tout $x \in]0, 1[$,

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

- 2) a) Soit $x \in [-c, c]$. Pour tout k , $|a_k x^k| \leq |a_k| c^k$, terme général d'une série convergente.
Par comparaison, la série

$$\sum a_k x^k$$

est absolument convergente, donc convergente.

Conclusion

$$f(x) = a_0 + \sum_{k=1}^{+\infty} a_k x^k$$

est bien défini pour tout $x \in [-c, c]$.

- b) La série définissant M_m converge, car $|a_k| c^{k-m-1} = c^{-m-1} |a_k| c^k$. Soit $x \in [-c, c]$.
Pour $k \geq m+1$, $|x|^{k-m-1} \leq c^{k-m-1}$, donc
 $|a_k x^k| = |a_k| |x|^{m+1} |x|^{k-m-1} \leq |x|^{m+1} |a_k| c^{k-m-1}$. Pour tout entier $K \geq m+1$, par l'inégalité triangulaire :

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

La série

$$\sum a_k x^k$$

converge d'après a) ; on fait tendre K vers $+\infty$ (passage à la limite dans une inégalité large).

Conclusion

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

pour tout $x \in [-c, c]$.

c) Soit $m \in \mathbb{N}^*$. Pour $x \in [-c, c]$,

$$f(x) - a_0 - \sum_{k=1}^m a_k x^k = \sum_{k=m+1}^{+\infty} a_k x^k.$$

D'après b), pour $x \in [-c, c]$ non nul :

$$\frac{1}{|x|^m} \left| f(x) - a_0 - \sum_{k=1}^m a_k x^k \right| \leq M_m |x| \rightarrow 0 \quad \text{lorsque } x \rightarrow 0$$

Conclusion

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

au voisinage de 0.

d) Raisonnons par l'absurde : supposons la suite (a_k) non nulle et notons m le plus petit indice tel que $a_m \neq 0$. D'après c) à l'ordre $m+1 \geq 1$, et puisque $a_k = 0$ pour $k < m$:

$$f(x) = a_m x^m + a_{m+1} x^{m+1} + o(x^{m+1}) = a_m x^m + o(x^m)$$

Pour $x \in]0, c]$, $f(x) = 0$, donc $\frac{f(x)}{x^m} = 0$; mais $\frac{f(x)}{x^m} = a_m + o(1) \rightarrow a_m$ lorsque $x \rightarrow 0^+$. Ainsi $a_m = 0$, ce qui est absurde.

Conclusion

si f est nulle sur $]0, c]$, la suite $(a_k)_{k \in \mathbb{N}}$ est nulle.

Partie II : estimateur du maximum de vraisemblance, un exemple

3) a) Pour tout $x \in \mathbb{N}^*$, $P^\theta([X = x]) = (1 - \theta)^{x-1}\theta$ avec $\theta \in]0, 1[$.

Conclusion

sous P^θ , X suit la loi géométrique de paramètre θ ; en particulier $E^\theta(X) = \frac{1}{\theta}$ et $V^\theta(X) = \frac{1 - \theta}{\theta^2}$.

b) $\bar{X}$ est une fonction de $(X_1, \dots, X_n)$ ne dépendant pas de θ : c'est un estimateur. Par (2), chaque X_i suit la loi géométrique de paramètre θ , donc par linéarité de l'espérance :

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

Conclusion

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

c) Question adaptée au programme actuel : le risque quadratique est hors programme ; la question demande la variance de l'estimateur, qui mesure ici l'écart quadratique moyen à $\frac{1}{\theta}$ puisque l'estimateur est sans biais.

Les X_i sont indépendantes (hypothèse (1)) et admettent une variance, donc :

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

Conclusion

$$V^\theta(\bar{X}) = \frac{1-\theta}{n\theta^2}.$$

- 4) a) La variable $\frac{1}{X}$ est à valeurs dans $]0, 1]$, donc bornée : elle admet une espérance.
Par le théorème de transfert, puis par 1) d) appliqué à $x = 1 - \theta \in]0, 1[$:

$$E^\theta\left(\frac{1}{X}\right) = \sum_{x=1}^{+\infty} \frac{1}{x} (1-\theta)^{x-1} \theta = \frac{\theta}{1-\theta} \sum_{x=1}^{+\infty} \frac{(1-\theta)^x}{x} = \frac{\theta}{1-\theta} (-\ln(1 - (1-\theta))) = \frac{-\theta \ln(\theta)}{1-\theta}$$

Par (2) et la linéarité de l'espérance,

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

Conclusion

$$\text{pour tout } \theta \in \Theta, \quad E^\theta(T) = \frac{\theta \ln(\theta)}{\theta - 1}.$$

- b) T est une fonction de $(X_1, \dots, X_n)$ ne dépendant pas de θ : c'est un estimateur de θ .
Son biais vaut :

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

Posons $k(t) = \ln(t) - t + 1$ sur $]0, 1]$. On a $k'(t) = \frac{1}{t} - 1 > 0$ sur $]0, 1[$, donc k est strictement croissante sur $]0, 1]$ et $k(1) = 0$: $k(\theta) < 0$ pour $\theta \in]0, 1[$. Comme $\theta > 0$ et

$\theta - 1 < 0$, le quotient est strictement positif.

Conclusion

$$T \text{ est un estimateur de } \theta \text{ de biais } b_{\theta}(T) = \frac{\theta(\ln(\theta) - \theta + 1)}{\theta - 1} > 0.$$

5) a) Par (4) et (2),

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

En prenant le logarithme :

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

Conclusion

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

b) Notons

$$\sigma = \sum_{i=1}^n x_i.$$

Les x_i sont dans $\mathbb{N}^*$ et ne sont pas tous égaux à 1, donc $\sigma > n$ et $\frac{n}{\sigma} \in]0, 1[$. La fonction $\psi : \theta \mapsto n \ln(\theta) + (\sigma - n) \ln(1 - \theta)$ est dérivable sur $]0, 1[$ et :

$$\psi'(\theta) = \frac{n}{\theta} - \frac{\sigma - n}{1 - \theta} = \frac{n(1 - \theta) - (\sigma - n)\theta}{\theta(1 - \theta)} = \frac{n - \sigma\theta}{\theta(1 - \theta)}$$

Le dénominateur est strictement positif ; ψ' est strictement positive sur $]0, \frac{n}{\sigma} [$ et strictement négative sur $]\frac{n}{\sigma}, 1 [$. Donc ψ est strictement croissante puis strictement décroissante : elle atteint son maximum en $\frac{n}{\sigma}$ et seulement en ce point. Comme $\exp$ est strictement croissante, il en est de même de $L = \exp \circ \psi$.

(Si tous les x_i valent 1, $L = \theta^n$ est strictement croissante sur $]0, 1 [$ et n'admet pas de maximum.)

Conclusion

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

est l'unique valeur de $\theta \in]0, 1 [$ qui maximise la vraisemblance.

6) a) Soit $\theta \in \Theta$, $k \geq n$, $a = \frac{1}{\theta}$ et $b = \frac{k}{n}$, deux réels strictement positifs.

Méthode 1 — Par la formule de Taylor avec reste intégral.

La fonction $g : t \mapsto \frac{1}{t}$ est de classe C^2 sur $]0, +\infty [$, intervalle contenant a et b , 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 donne :

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

avec $g(b) = \frac{n}{k}$, $g(a) = \theta$ et $g'(a) = -\theta^2$.

Méthode 2 — Par un calcul direct de l'intégrale.

Une primitive de $t \mapsto (b-t)\frac{2}{t^3} = \frac{2b}{t^3} - \frac{2}{t^2}$ sur $]0, +\infty[$ est $t \mapsto -\frac{b}{t^2} + \frac{2}{t}$. Donc :

$$\int_a^b (b-t)\frac{2}{t^3} dt = \left(-\frac{1}{b} + \frac{2}{b}\right) - \left(-\frac{b}{a^2} + \frac{2}{a}\right) = \frac{n}{k} + \theta^2 \frac{k}{n} - 2\theta$$

Par ailleurs $\theta - \theta^2 \left(\frac{k}{n} - \frac{1}{\theta}\right) = 2\theta - \theta^2 \frac{k}{n}$, et la somme des deux expressions vaut $\frac{n}{k}$.

La méthode 1 explique d'où vient la formule (développement de $\frac{1}{t}$ autour de l'espérance $\frac{1}{\theta}$ de $\bar{X}$) ; la méthode 2 la vérifie sans aucun théorème.

Conclusion

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

b) Notons

$$\Sigma_n = \sum_{i=1}^n X_i,$$

$p_k = P^\theta([\Sigma_n = k])$ et J_k l'intégrale de a). Comme les X_i sont à valeurs dans $\mathbb{N}^*$, Σ_n est à valeurs dans $\llbracket n, +\infty \rrbracket$, et $U = \frac{n}{\Sigma_n}$ est une fonction de l'échantillon ne dépendant pas de θ : c'est un estimateur de θ , à valeurs dans $]0, 1]$, donc admettant une espérance. Par le théorème de transfert et a) :

$$\frac{n}{k} p_k = \theta p_k - \theta^2 \left(\frac{k}{n} - \frac{1}{\theta}\right) p_k + J_k p_k$$

La série

$$\sum p_k$$

converge vers 1, et

$$\sum k p_k$$

converge vers $E^\theta(\Sigma_n) = \frac{n}{\theta}$ (linéarité), donc

$$\sum \left(\frac{k}{n} - \frac{1}{\theta} \right) p_k$$

converge vers $\frac{1}{n} \cdot \frac{n}{\theta} - \frac{1}{\theta} = 0$. La série

$$\sum \frac{n}{k} p_k$$

converge vers $E^\theta(U)$. Par différence, la série

$$\sum J_k p_k$$

converge et :

$$E^\theta(U) = \theta - 0 + \sum_{k=n}^{+\infty} J_k p_k$$

Conclusion

U est un estimateur de θ de biais

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

- c) Signe de J_k . Si $\frac{k}{n} > \frac{1}{\theta}$, la fonction $t \mapsto \left(\frac{k}{n} - t \right) \frac{2}{t^3}$ est continue, positive sur $\left[\frac{1}{\theta}, \frac{k}{n} \right]$ et non identiquement nulle, donc $J_k > 0$. Si $\frac{k}{n} < \frac{1}{\theta}$, $J_k = \int_{k/n}^{1/\theta} \left(t - \frac{k}{n} \right) \frac{2}{t^3} dt > 0$ pour la même raison. Si $\frac{k}{n} = \frac{1}{\theta}$, $J_k = 0$. Dans tous les cas $J_k \geq 0$, et $p_k \geq 0$.

Pour $k = n$: $\frac{k}{n} = 1 < \frac{1}{\theta}$, donc $J_n > 0$; et $p_n = P^{\theta}(\cap_i [X_i = 1]) = \theta^n > 0$ par (1). Tous les termes de la série étant positifs :

$$b_{\theta}(U) \geq p_n J_n > 0$$

Conclusion

$b_{\theta}(U) > 0$ pour tout $\theta \in \Theta$.

7) Notons

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

et fixons $\theta \in \Theta$.

Suite (T_n) . Les variables $\frac{1}{X_i}$ sont indépendantes, de même loi, bornées (donc admettent une variance), d'espérance $m(\theta) = \frac{\theta \ln(\theta)}{\theta - 1}$ (4) a)). Par la loi faible des grands nombres, T_n converge en probabilité vers $m(\theta)$. Or $m(\theta) - \theta = b_{\theta}(T) = \beta > 0$ d'après 4) b) (ce biais ne dépend pas de n). Avec $\delta = \frac{\beta}{2}$, si $|T_n - m(\theta)| < \delta$ alors

$|T_n - \theta| \geq \beta - |T_n - m(\theta)| > \delta$, donc :

$$P^\theta (|T_n - \theta| > \delta) \geq 1 - P^\theta (|T_n - m(\theta)| \geq \delta) \rightarrow 1$$

La suite (T_n) ne converge donc pas en probabilité vers θ .

Suite (U_n) . Les X_i sont indépendantes, de même loi géométrique, d'espérance $\frac{1}{\theta}$ et de variance finie ; par la loi faible des grands nombres, $\bar{X}_n$ converge en probabilité vers $\frac{1}{\theta}$.

On a $U_n = \frac{1}{\bar{X}_n}$ avec $\bar{X}_n \geq 1$.

Méthode 1 — Par une majoration directe.

Comme $\bar{X}_n \geq 1$:

$$|U_n - \theta| = \frac{\theta \left| \frac{1}{\theta} - \bar{X}_n \right|}{\bar{X}_n} \leq \theta \left| \bar{X}_n - \frac{1}{\theta} \right|.$$

Pour $\varepsilon > 0$:

$$P^\theta (|U_n - \theta| > \varepsilon) \leq P^\theta \left(\left| \bar{X}_n - \frac{1}{\theta} \right| > \frac{\varepsilon}{\theta} \right) \rightarrow 0$$

Méthode 2 — Par composition avec une fonction continue.

La fonction $\gamma : y \mapsto \frac{1}{\max(y, 1)}$ est continue sur $\mathbb{R}$ (elle vaut 1 sur $] -\infty, 1]$ et $\frac{1}{y}$ sur $[1, +\infty[$, les deux expressions coïncidant en 1). Comme $\bar{X}_n \geq 1$, $U_n = \gamma(\bar{X}_n)$. Par composition de la convergence en probabilité par une fonction continue sur $\mathbb{R}$, U_n converge en probabilité vers $\gamma\left(\frac{1}{\theta}\right) = \theta$, car $\frac{1}{\theta} > 1$.

La méthode 1 donne une majoration quantitative ; la méthode 2 est plus rapide mais demande de prolonger $y \mapsto \frac{1}{y}$ en une fonction continue sur $\mathbb{R}$ tout entier, car le théorème de composition porte sur des fonctions continues sur $\mathbb{R}$.

Conclusion

(U_n) est une suite convergente d'estimateurs de θ (bien que chaque U_n soit biaisé), alors que (T_n) converge en probabilité vers $\frac{\theta \ln(\theta)}{\theta - 1} \neq \theta$ et n'est donc pas une suite convergente d'estimateurs de θ .

Partie III : statistique exhaustive, un exemple

- 8) a) Pour $x \in \{0, 1\}$, $P^\theta([X = x]) = \theta^x(1 - \theta)^{1-x}$ (cela vaut $1 - \theta$ si $x = 0$ et θ si $x = 1$).
Par (4) et (2) :

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

Conclusion

$$L(x_1, \dots, x_n, \theta) = \theta^{\sum_i x_i} (1 - \theta)^{\sum_i (1-x_i)}.$$

- b) Ici

$$s(x_1, \dots, x_n) = \sum_i x_i$$

et $s(B^n) = \llbracket 0, n \rrbracket$. Comme

$$\sum_i (1 - x_i) = n - \sum_i x_i,$$

on pose $g(u, \theta) = \theta^u(1 - \theta)^{n-u} \geq 0$ pour $u \in \llbracket 0, n \rrbracket$ et $h = 1$. D'après a),
 $L(x_1, \dots, x_n, \theta) = g(s(x_1, \dots, x_n), \theta) h(x_1, \dots, x_n)$.

Conclusion

S est exhaustive.

- 9) a) Sous P^θ , S suit la loi binomiale $\mathcal{B}(n, \theta)$ (somme de n variables de Bernoulli indépendantes), donc $P^\theta([S = k]) = \binom{n}{k} \theta^k (1 - \theta)^{n-k} > 0$. Notons
 $A_x = [X_1 = x_1] \cap \dots \cap [X_n = x_n]$.

Si

$$\sum_i x_i \neq k,$$

$A_x \cap [S = k] = \emptyset$ et la probabilité conditionnelle est nulle.

Si

$$\sum_i x_i = k,$$

$A_x \subset [S = k]$ et, par (1) et 8) a), $P^\theta(A_x) = \theta^k (1 - \theta)^{n-k}$, donc :

$$P_{[S=k]}^\theta(A_x) = \frac{\theta^k (1 - \theta)^{n-k}}{\binom{n}{k} \theta^k (1 - \theta)^{n-k}} = \frac{1}{\binom{n}{k}}$$

Conclusion

$$P_{[S=k]}^\theta(A_x) \text{ vaut } \frac{1}{\binom{n}{k}} \text{ si}$$

$$\sum_i x_i = k$$

et 0 sinon : sachant $[S = k]$, le vecteur $(X_1, \dots, X_n)$ suit la loi uniforme sur les $\binom{n}{k}$ suites de $\{0, 1\}^n$ comportant exactement k coefficients égaux à 1, loi qui ne dépend pas de θ .

b) Méthode 1 — Par dénombrement.

L'événement $[X_1 = 1]$ est la réunion disjointe des A_x pour $x \in \{0, 1\}^n$ tel que $x_1 = 1$.
D'après a), seuls ceux vérifiant

$$\sum_i x_i = k$$

contribuent, chacun pour $\frac{1}{\binom{n}{k}}$. Si $k \geq 1$, ces suites sont déterminées par le choix des

$k - 1$ positions des autres 1 parmi les $n - 1$ positions $2, \dots, n$: il y en a $\binom{n-1}{k-1}$. Donc

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

Si $k = 0$, aucune suite ne convient et la probabilité vaut $0 = \frac{0}{n}$.

Méthode 2 — Par symétrie et linéarité.

La loi conditionnelle de a) est invariante par permutation des coordonnées, donc

$P_{[S=k]}^\theta([X_i = 1])$ ne dépend pas de i ; notons-la q . Sous $P_{[S=k]}^\theta$, $S = k$ presque sûrement et, par linéarité de l'espérance (chaque X_i étant une variable de Bernoulli) :

$$k = E_{[S=k]}^\theta(S) = \sum_{i=1}^n P_{[S=k]}^\theta([X_i = 1]) = nq.$$

La méthode 1 est un calcul direct ; la méthode 2, sans dénombrement, montre en plus que le résultat vaut pour chaque X_i , ce qui servira en 10) c).

Conclusion

$$P_{[S=k]}^\theta([X_1 = 1]) = \frac{k}{n} \text{ pour tout } \theta \in \Theta.$$

10) Code adapté en Python. Le script Scilab de l'énoncé d'origine est remplacé par un code Python de même comportement : `rd.binomial(1, theta, (n, N))` remplace `grand`, `np.sum(U, 0)` remplace la somme par colonnes, le tableau de booléens `K` et l'extraction `U[:, K]` remplacent `find` et `U(1:n,K)`, `np.mean(Y, 1)` remplace la moyenne par lignes. Les valeurs affichées par l'énoncé d'origine sont conservées. Le code a été exécuté ; une exécution a donné des valeurs comprises entre 0,394 et 0,406, avec 19 958 colonnes retenues.

a) `U` est un tableau à $n = 10$ lignes et $N = 100\,000$ colonnes de simulations indépendantes de la loi de Bernoulli de paramètre $\theta = 0,3$. Une colonne contient donc 10 simulations indépendantes de cette loi.

Conclusion

chaque colonne de `U` est une simulation du vecteur $(X_1, \dots, X_{10})$ sous P^θ avec $\theta = 0,3$, et les N colonnes sont des simulations indépendantes de ce vecteur.

b) `S` contient, pour chaque colonne de `U`, la somme de ses coefficients : c'est une

simulation de

$$S = \sum_i X_i$$

pour chacune des N simulations du vecteur. K repère les colonnes pour lesquelles cette somme vaut $k = 4$, et Y est formé de ces seules colonnes. On ne conserve donc que les simulations pour lesquelles l'événement $[S = 4]$ est réalisé (méthode de rejet).

Pour une colonne donnée, qui simule $(X_1, \dots, X_n)$, la loi de cette colonne conditionnellement au fait qu'elle soit retenue est la loi conditionnelle de $(X_1, \dots, X_n)$ sachant $[S = 4]$; les colonnes étant indépendantes, les colonnes retenues sont des simulations indépendantes de cette loi conditionnelle. Leur nombre est aléatoire, d'espérance $N P^\theta([S = 4]) = N \binom{10}{4} 0,3^4 0,7^6 \approx 20\,012$.

Conclusion

les colonnes de Y simulent la loi conditionnelle de $(X_1, \dots, X_n)$ sachant $[S = k]$.

- c) Le coefficient d'indice i de M est la proportion de 1 dans la ligne i de Y , c'est-à-dire la fréquence de l'événement $[X_i = 1]$ parmi les simulations conditionnées par $[S = 4]$. Par la loi faible des grands nombres, il approche $P_{[S=4]}^\theta([X_i = 1]) = \frac{4}{10} = 0,4$ (question 9) b) et sa méthode 2, valable pour tout i).

Les dix valeurs obtenues sont toutes proches de 0,4 (entre 0,394 et 0,406), l'ordre de grandeur des écarts étant celui de l'écart-type $\sqrt{0,4 \times 0,6/20\,000} \approx 0,0035$. On note que 0,4 diffère de $\theta = 0,3$: sachant $[S = 4]$, la loi du vecteur ne dépend plus de θ , ce qu'illustre la simulation. Enfin, chaque colonne de Y ayant exactement quatre coefficients égaux à 1, la somme des coefficients de M vaut exactement 4 (aux arrondis d'affichage près : la somme des valeurs affichées vaut 4,0000001).

Conclusion

les coefficients de M sont des approximations de $\frac{k}{n} = 0,4$, valeur indépendante de θ , conformément à l'exhaustivité de S .

- 11)** Code adapté en Python. L'instruction Scilab $C=Y*Y'/\text{length}(K)$ devient $C = \text{np.dot}(Y, \text{np.transpose}(Y)) / \text{np.shape}(Y)[1]$; les coefficients $C(1,1)$ et $C(1,2)$ deviennent $C[0, 0]$ et $C[0, 1]$, les indices Python commençant à 0.

Notons m le nombre de colonnes de Y et $Y_{i,j} \in \{0, 1\}$ ses coefficients. Pour tous i, l :

$$C_{i,l} = \frac{1}{m} \sum_{j=1}^m Y_{i,j} Y_{l,j}$$

- a)** Y a n lignes et m colonnes, donc $Y^t Y$ est carrée d'ordre n . Comme $Y_{0,j}^2 = Y_{0,j}$ (valeurs 0 ou 1),

$$C_{0,0} = \frac{1}{m} \sum_j Y_{0,j},$$

qui est le premier coefficient de M .

Conclusion

C est une matrice carrée d'ordre $n = 10$, et $C[0, 0] = 0,4019917$, premier coefficient de M .

- b)** Pour $i \neq l$, $C_{i,l}$ est la fréquence, parmi les simulations retenues, de l'événement $[X_i = 1] \cap [X_l = 1]$. Elle approche donc $P_{[S=k]}^\theta([X_i = 1] \cap [X_l = 1])$. D'après 9) a), cette probabilité vaut $\frac{1}{\binom{n}{k}}$ fois le nombre de suites ayant k coefficients égaux à 1 dont ceux d'indices i et l , soit $\binom{n-2}{k-2}$ (pour $k \geq 2$) :

$$\frac{\binom{n-2}{k-2}}{\binom{n}{k}} = \frac{k(k-1)}{n(n-1)} = \frac{4 \times 3}{10 \times 9} = \frac{2}{15}$$

Conclusion

on s'attend à $C[0, 1] \approx \frac{2}{15} \approx 0,133$, et de même pour tous les coefficients non diagonaux. (Le code exécuté en Python a donné des coefficients non diagonaux compris entre 0,128 et 0,139.)

c) La somme de tous les coefficients vaut :

$$\sum_{i,l} C_{i,l} = \frac{1}{m} \sum_{j=1}^m \left(\sum_i Y_{i,j} \right) \left(\sum_l Y_{l,j} \right) = \frac{1}{m} \sum_{j=1}^m k^2 = k^2$$

car chaque colonne retenue a exactement k coefficients égaux à 1. On retrouve la cohérence avec a) et b) : $10 \times 0,4 + 90 \times \frac{2}{15} = 4 + 12 = 16$.

Conclusion

la somme des coefficients de C vaut exactement $k^2 = 16$ (le code exécuté affiche 16,0).

Partie IV : inégalité de Rao-Blackwell

12) a) Soit $\theta \in \Theta$ et $u \in s(B^n)$. Il existe $x = (x_1, \dots, x_n) \in B^n$ tel que $s(x) = u$. Alors $\bigcap_i [X_i = x_i] \subset [S = u]$, et par (1), (2), (3) :

$$P^\theta([S = u]) \geq \prod_{i=1}^n P^\theta([X = x_i]) > 0$$

La probabilité conditionnelle $P_{[S=u]}^\theta$ est donc bien définie.

Écrivons $T = t(X_1, \dots, X_n)$. Comme B^n est dénombrable, T prend ses valeurs dans l'ensemble dénombrable $t(B^n)$. Pour $y \in t(B^n)$, l'événement $[T = y]$ est la réunion disjointe dénombrable des événements $A_x = \bigcap_i [X_i = x_i]$ pour $x \in B^n$ tel que $t(x) = y$, donc :

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

Chaque terme ne dépend pas de θ (résultat admis), donc $P_{[S=u]}^\theta([T = y])$ non plus : la loi de T sachant $[S = u]$ ne dépend pas de θ .

Existence : T est sans biais, donc admet une espérance sous P^θ , et

$$\sum_y |y| P^\theta([T = y])$$

converge. Comme $P_{[S=u]}^\theta([T = y]) \leq \frac{P^\theta([T = y])}{P^\theta([S = u])}$, la série

$$\sum_y |y| P_{[S=u]}^\theta([T = y])$$

converge par comparaison.

Conclusion

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

existe, et c'est la somme d'une série dont les termes ne dépendent pas de θ : sa valeur ne dépend pas de θ .

- b) Pour tout ω , $X_i(\omega) \in B$ pour tout i , donc $S(\omega) = s(X_1(\omega), \dots, X_n(\omega)) \in s(B^n)$: les événements $[S = u]$, $u \in s(B^n)$, recouvrent Ω . Ils sont deux à deux disjoints (S prend une seule valeur en chaque ω), et indexés par l'ensemble $s(B^n)$, image de l'ensemble dénombrable B^n , donc fini ou dénombrable.

Conclusion

$([S = u])_{u \in s(B^n)}$ est un système complet d'événements (chacun de probabilité strictement positive d'après a)).

- 13) Notons $\phi(u) = E_{[S=u]}(T)$ pour $u \in s(B^n)$, de sorte que $r(x) = \phi(s(x))$ et $R = \phi(S)$.

- a) R est une fonction de $(X_1, \dots, X_n)$ qui ne dépend pas de θ (d'après 12) a)) : c'est un estimateur. Soit $\theta \in \Theta$.

Existence de $E^\theta(R)$. Par l'inégalité triangulaire sur les séries, $|\phi(u)| \leq E_{[S=u]}^\theta(|T|)$ (cette espérance existe par le même argument qu'en 12) a), $|T|$ admettant une espérance). Par la formule de l'espérance totale appliquée à $|T|$ et au système complet de 12) b) :

$$\sum_{u \in s(B^n)} E_{[S=u]}^\theta(|T|) P^\theta([S = u]) = E^\theta(|T|) < +\infty$$

Donc

$$\sum_u |\phi(u)| P^\theta([S = u])$$

converge, et par le théorème de transfert, $R = \phi(S)$ admet une espérance égale à

$$\sum_u \phi(u) P^\theta([S = u]).$$

Valeur. Par la formule de l'espérance totale appliquée à T :

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

Conclusion

R est un estimateur sans biais de θ .

b) Fixons $\theta \in \Theta$ et notons $Q = P_{[S=u]}^\theta$ pour $u \in s(B^n)$.

Existence des espérances conditionnelles. T admet un moment d'ordre 2, donc $(T - \theta)^2$ admet une espérance ; par l'argument de 12) a), $E_{[S=u]}((T - \theta)^2)$ existe, ainsi que $E_{[S=u]}(T - \theta) = \phi(u) - \theta$ par linéarité.

Justification des inégalités. Sous la probabilité Q , la variable $T - \theta$ admet un moment d'ordre 2, donc une variance positive ; par la formule de Koenig-Huygens :

$$E_{[S=u]}((T - \theta)^2) - \left(E_{[S=u]}(T - \theta)\right)^2 \geq 0$$

Variance de R . On a $(R - \theta)^2 = \psi(S)$ avec $\psi(u) = (\phi(u) - \theta)^2 = \left(E_{[S=u]}(T - \theta)\right)^2$, et $0 \leq \psi(u) \leq E_{[S=u]}((T - \theta)^2)$. Par la formule de l'espérance totale appliquée à $(T - \theta)^2$, et puisque $E^\theta(T) = \theta$:

$$\sum_{u \in s(B^n)} E_{[S=u]}((T - \theta)^2) P^\theta([S = u]) = E^\theta((T - \theta)^2) = V^\theta(T)$$

Par comparaison de séries à termes positifs,

$$\sum_u \psi(u) P^\theta([S = u])$$

converge et sa somme est au plus $V^\theta(T)$. Par le théorème de transfert, $(R - \theta)^2$ admet une espérance, donc $R = (R - \theta) + \theta$ admet un moment d'ordre 2 et une variance.

Comme $E^\theta(R) = \theta$ (question a)) :

$$V^\theta(R) = E^\theta((R - \theta)^2) = \sum_u \psi(u) P^\theta([S = u]) \leq V^\theta(T)$$

Conclusion

R admet une variance et $V^\theta(R) \leq V^\theta(T)$ pour tout $\theta \in \Theta$.

- 14) a) Pour $x \in \mathbb{N}$, $P^\theta([X = x]) = e^{-\theta} \frac{\theta^x}{x!} > 0$, donc l'hypothèse (3) est satisfaite. Par (4) et (2) :

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

Avec $s(B^n) = \mathbb{N}$, $g(u, \theta) = e^{-n\theta} \theta^u \geq 0$ et

$$h(x_1, \dots, x_n) = \frac{1}{\prod_i x_i!} \geq 0,$$

on a $L = g(s(x), \theta)h(x)$.

Conclusion

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

est exhaustive.

- b) Par stabilité de la loi de Poisson pour la somme de variables indépendantes (récurrence sur n), S suit la loi de Poisson de paramètre $n\theta$, donc

$$P^\theta([S = u]) = e^{-n\theta} \frac{(n\theta)^u}{u!} > 0. \text{ Notons } A_x = \bigcap_i [X_i = x_i].$$

Si

$$\sum_i x_i \neq u,$$

$A_x \cap [S = u] = \emptyset$ et la probabilité conditionnelle est nulle. Si

$$\sum_i x_i = u,$$

$A_x \subset [S = u]$ et, par (1) :

$$P_{[S=u]}^\theta(A_x) = \frac{e^{-n\theta}\theta^u / \prod_i x_i!}{e^{-n\theta}(n\theta)^u / u!} = \frac{u!}{n^u \prod_{i=1}^n x_i!}$$

Conclusion

$P_{[S=u]}^\theta(A_x)$ vaut

$$\frac{u!}{n^u \prod_i x_i!}$$

si

$$\sum_i x_i = u$$

et 0 sinon : elle ne dépend pas de θ .

- c) Supposons $n \geq 2$ (si $n = 1$, $X_1 = S = u$ presque sûrement sous $P_{[S=u]}^\theta$, loi binomiale $\mathcal{B}(u, 1)$). Soit $i \in \llbracket 1, n \rrbracket$ et $S'_i = S - X_i$, somme des $n - 1$ autres variables. Par le lemme des coalitions, X_i et S'_i sont indépendantes, et S'_i suit la loi de Poisson de

paramètre $(n-1)\theta$.

Soit $j \in \mathbb{N}$. Si $j > u$, $[X_i = j] \cap [S = u] = \emptyset$ car les X_i sont positives. Si $0 \leq j \leq u$, $[X_i = j] \cap [S = u] = [X_i = j] \cap [S'_i = u - j]$, donc :

$$P_{[S=u]}^\theta([X_i = j]) = \frac{e^{-\theta} \frac{\theta^j}{j!} e^{-(n-1)\theta} \frac{((n-1)\theta)^{u-j}}{(u-j)!}}{e^{-n\theta} \frac{(n\theta)^u}{u!}} = \binom{u}{j} \left(\frac{1}{n}\right)^j \left(1 - \frac{1}{n}\right)^{u-j}$$

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

Indépendance. Si $u \geq 1$ (et $n \geq 2$), $[X_1 = u] \cap [X_2 = u] \cap [S = u] = \emptyset$ car on aurait $S \geq 2u > u$; donc $P_{[S=u]}^\theta([X_1 = u] \cap [X_2 = u]) = 0$, alors que

$P_{[S=u]}^\theta([X_1 = u]) P_{[S=u]}^\theta([X_2 = u]) = \frac{1}{n^{2u}} > 0$. Si $u = 0$, toutes les X_i sont nulles presque sûrement sous $P_{[S=0]}^\theta$: des variables presque sûrement constantes sont indépendantes.

Conclusion

sous $P_{[S=u]}^\theta$, chaque X_i suit la loi $\mathcal{B}\left(u, \frac{1}{n}\right)$; pour $n \geq 2$, les X_i ne sont pas indépendantes si $u \geq 1$, et elles le sont (trivialement) si $u = 0$.

d) Existence. Pour $\theta > 0$, par la série exponentielle :

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

(en posant $k = j + 1$ et en utilisant $\frac{1}{(k-1)!} = \frac{k}{k!}$). La suite $\varphi_k = \frac{k}{n}$ convient (avec $\varphi_0 = 0$).

Unicité. Soit (ψ_k) une autre suite telle que

$$\sum_k \psi_k \frac{(n\theta)^k}{k!}$$

converge et vaut $\theta e^{n\theta}$ pour tout $\theta > 0$. La série converge pour $\theta = 2$, donc son terme général tend vers 0 et est borné : il existe K tel que $|\psi_k| \frac{(2n)^k}{k!} \leq K$ pour tout k . Alors

$|\psi_k| \frac{n^k}{k!} \leq \frac{K}{2^k}$, terme d'une série convergente ; il en est de même pour (φ_k) . Posons $a_k = (\varphi_k - \psi_k) \frac{n^k}{k!}$: la série

$$\sum |a_k| 1^k$$

converge, ce qui est l'hypothèse de la question 2) avec $c = 1$. Pour $x \in]0, 1]$:

$$a_0 + \sum_{k=1}^{+\infty} a_k x^k = \sum_{k=0}^{+\infty} \varphi_k \frac{(nx)^k}{k!} - \sum_{k=0}^{+\infty} \psi_k \frac{(nx)^k}{k!} = xe^{nx} - xe^{nx} = 0$$

D'après 2) d), $a_k = 0$ pour tout k , donc $\psi_k = \varphi_k$.

Conclusion

l'unique suite convenable est $\varphi_k = \frac{k}{n}$ pour tout $k \in \mathbb{N}$.

- e) Question adaptée au programme actuel : « risque quadratique » est remplacé par « variance », et la comparaison porte sur les estimateurs sans biais admettant une variance ; pour un estimateur sans biais, il s'agit de la même quantité.

L'estimateur $\bar{X} = \frac{S}{n}$ est sans biais ($E^\theta(\bar{X}) = \theta$) et admet une variance, $V^\theta(\bar{X}) = \frac{\theta}{n}$.

Soit T un estimateur sans biais de θ admettant une variance pour tout θ . La statistique S est exhaustive (a)) et la probabilité conditionnelle sachant $[S = u]$ ne dépend pas de θ (b)), donc les questions 12) et 13) s'appliquent. Avec $\phi(u) = E_{[S=u]}(T)$, ne dépendant

pas de θ , l'estimateur $R = \phi(S)$ est sans biais (13) a)), la série étant absolument convergente :

$$\forall \theta > 0, \quad \sum_{u=0}^{+\infty} \phi(u) e^{-n\theta} \frac{(n\theta)^u}{u!} = \theta, \quad \text{donc} \quad \sum_{u=0}^{+\infty} \phi(u) \frac{(n\theta)^u}{u!} = \theta e^{n\theta}$$

Par l'unicité de d), $\phi(u) = \frac{u}{n}$ pour tout $u \in \mathbb{N}$, c'est-à-dire $R = \frac{S}{n} = \bar{X}$. L'inégalité de Rao-Blackwell (13) b)) donne alors, pour tout $\theta > 0$:

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

Conclusion

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

est un estimateur sans biais de θ de variance minimale parmi les estimateurs sans biais admettant une variance : il est optimal.

Relevé des adaptations

- 3) c) : Question adaptée au programme actuel. « Quel est le risque quadratique de cet estimateur ? » devient « Quelle est la variance de cet estimateur ? ». Justification : le risque quadratique est hors programme ; pour un estimateur sans biais, il coïncide avec la variance.
- 6) b) : précision de notation. La probabilité notée P dans le document fourni est notée P^θ . Justification : cohérence avec les notations du préambule.
- 10) : Code adapté en Python. Le script Scilab devient un code Python (`rd.binomial`, `np.sum(U, 0)`, sélection par un tableau de booléens, `np.mean(Y, 1)`), ces commandes

étant expliquées dans l'énoncé ; les valeurs affichées dans le document fourni sont conservées. Justification : seul Python est autorisé, et les commandes hors liste doivent être expliquées.

- 11) : Code adapté en Python. L'instruction $C=Y*Y'/\text{length}(K)$ devient $C = \text{np.dot}(Y, \text{np.transpose}(Y)) / \text{np.shape}(Y)[1]$; les coefficients $C(1,1)$ et $C(1,2)$ deviennent $C[0, 0]$ et $C[0, 1]$. Justification : seul Python est autorisé, avec une indexation commençant à 0.
- 14) e) : Question adaptée au programme actuel. « Son risque quadratique est inférieur ou égal à celui de tout autre estimateur sans biais » devient « sa variance est inférieure ou égale à celle de tout autre estimateur sans biais de θ admettant une variance pour toute valeur de θ ». Justification : le risque quadratique est hors programme ; l'inégalité de Rao-Blackwell (13) b)) suppose l'existence de la variance.