STATISTIQUE MATHÉMATIQUE - MS ENSAI 2026-2027
EXERCICE 2.22. CORRIGÉ

On considère un n-échantillon $X = (X_1, ..., X_n)$ d'une variable aléatoire de loi de Bernoulli de paramètre $\theta \in]0,1[$. On note $S_n = \sum_{i=1}^n X_i$ et $\overline{X} = S_n/n$ la moyenne empirique de l'échantillon.

1°. Déterminer la vraisemblance du n-échantillon $X_1, ..., X_n$. Le modèle est-il exponentiel ?

2°. Déterminer un estimateur $\widehat{\theta}_n$ de θ par la méthode du maximum de vraisemblance, puis par la méthode des moments. Que constatez-vous ?

3°. Montrer que cet estimateur est fortement consistant et converge en moyenne quadratique. Démontrer que $\widehat{\theta}_n$ est asymptotiquement normal et déterminer la variance de la loi limite.

4°. Déterminer un intervalle de confiance asymptotique pour θ, de niveau $1-\alpha$.

5°. $\widehat{\theta}_n$ est-il exhaustif ? Minimal ? Complet ?

6°. Le modèle est-il régulier ?

7°. Démontrer que $\widehat{\theta}_n$ est un estimateur VUMSB (sans biais de variance minimale) de θ.

8°. Calculer l'information de Fisher relative à θ et montrer que $\widehat{\theta}_n$ est un estimateur efficace de θ.

On considère que la loi *a priori* de θ est une loi bêta $\mathscr{B}(a,b)$, avec $a, b > 0$, dont la densité est donnée par

$$f(\theta) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \theta^{a-1}(1-\theta)^{b-1} \mathbb{1}_{[0,1]}(\theta) \tag{1}$$

où Γ est la fonction Gamma d'Euler donnée par

$$\Gamma(a) = \int_0^{+\infty} x^{a-1} e^{-x} dx \tag{2}$$

On rappelle que l'espérance d'une variable aléatoire H de loi bêta est donnée par

$$\mathbb{E}[H] = \frac{a}{a+b} \tag{3}$$

9°. Déterminer la loi *a posteriori* de θ sachant $[X = x]$. En déduire que l'estimateur de la moyenne *a posteriori* peut s'exprimer au travers de la statistique S_n. Montrer que cet estimateur est biaisé, mais asymptotiquement sans biais.

10°. On suppose $n > 2$. On souhaite estimer la valeur de θ^2. Démontrer que $Y = X_1 X_2$ est un estimateur sans biais de θ^2. Démontrer que $T_n = \mathbb{E}[Y|S_n]$ est un estimateur sans biais de variance minimale pour θ^2.

11°. On se propose de calculer explicitement l'expression de T_n. Calculer $\mathbb{P}[Y = 1|S_n = s]$ (on distinguera les cas $s < 2$ et $s \geq 2$). En déduire que

$$T_n = \frac{S_n(S_n-1)}{n(n-1)} \tag{4}$$

Correction

1°. La vraisemblance de l'échantillon s'écrit

$$L(x,\theta) = \prod_{i=1}^n \mathbb{P}[X_i = x_i] = \theta^s (1-\theta)^{n-s} \tag{5}$$

où $s = \sum_{i=1}^n x_i$.

On voit facilement que

$$L(x,\theta) = \exp\left(n\overline{x} \ln\left(\frac{\theta}{1-\theta}\right) - n\ln(1-\theta) \right) \tag{6}$$

ce qui montre que le modèle est exponentiel et que $\overline{X}$ est la statistique naturelle.

2°. La log-vraisemblance l est bien définie et un calcul immédiat déjà fait en cours et dans plusieurs exercices donne

$$l(x,\theta) = s\ln\theta + (n-s)\ln(1-\theta) \tag{7}$$

Cette fonction est de classe C^2 (et même C^∞) sur $]0,\infty[$ et par suite

$$\frac{\partial l}{\partial \theta}(x,\theta) = \frac{s}{\theta} - \frac{n-s}{1-\theta} = \frac{s}{\theta(1-\theta)} - \frac{n}{1-\theta} \tag{8}$$

$$= \frac{s - n\theta}{\theta(1-\theta)} \tag{9}$$

qui s'annule en $\theta = s/n$ et cette valeur est bien un maximum car $l(x,\theta)$ change de signe en s/n (ou bien on peut également constaté que la dérivée seconde de l est négative).

Ainsi, $\widehat{\theta} = \overline{X} = S/n$ est l'estimateur du maximum de vraisemblance.

Pour déterminer l'estimateur des moments, on égale le moment théorique d'ordre un $\mathbb{E}[X] = \theta$ avec le moment empirique $\overline{X}$. L'estimateur des moments est la valeur de θ qui égale ces deux quantités. On voit donc que $\overline{X} = \widehat{\theta}_n$ est également l'estimateur des moments.

3°. Appliquons la loi forte des grands nombres à la suite $(X_n)_n$. Ceci est licite car les X_i sont des v.a.i.i.d. intégrables ($\mathbb{E}[|X_1|] = \theta < \infty$). On a alors :

$$\widehat{\theta}_n = \overline{X} \underset{N\to+\infty}{\longrightarrow} \mathbb{E}[X_1] = \theta \text{ p.s.} \tag{10}$$

Ceci prouve que $\widehat{\theta}_n$ est fortement consistant.

$\widehat{\theta}_n$ est également sans biais car $\mathbb{E}[\overline{X}] = n \times \mathbb{E}[X_1]/n = \theta$.

Nous pouvons également appliquer le théorème de la limite centrale car la suite des $(X_n)_n$ est de carré intégrable. On a alors

$$\sqrt{n}\left(\widehat{\theta}_n - \theta\right) \rightsquigarrow \mathscr{N}(0, \sigma^2) \tag{11}$$

où $\sigma^2 = \theta(1-\theta)$ et $\rightsquigarrow$ indique la convergence en loi. Ainsi, $\widehat{\theta}_n$ est asymptotiquement normal.

4°. De la question précédente et des propriétés de la convergence en loi, on déduit que pour tout a, b dans $\mathbb{R}$,

$$\lim_{n\to+\infty} \mathbb{P}\left[a \le \sqrt{n}\left(\frac{\widehat{\theta}_n - \theta}{\sigma}\right) \le b\right] = \Pi(b) - \Pi(a) \quad (12)$$

où Π est la fonction de répartition d'une loi normale standard et $\sigma = \sqrt{\theta(1-\theta)}$. Mais θ est inconnu donc σ aussi. Nous appliquons donc la méthode de Wald en estimant θ par $(\overline{X}(1-\overline{X}))^{1/2}$. Le lemme de Slutsky assure que la limite précédente est toujours valable et nous donne un IC asymptotique au niveau $1-\alpha$:

$$IC_{1-\alpha}(\theta) = \left[\overline{X} - \frac{q}{\sqrt{n}}\sqrt{\overline{X}(1-\overline{X})}\ ;\ \overline{X} + \frac{q}{\sqrt{n}}\sqrt{\overline{X}(1-\overline{X})}\right]$$

où $q = q_{1-\alpha/2}$ est le quantile d'ordre $1-\alpha/2$ d'une loi normale standard.

5°. L'espace canonique des paramètres (c'est l'image de la fonction $\Lambda(\theta) = \ln(\theta/(1-\theta))$) est égal à $\mathbb{R}$ (quitte à faire une petite étude de fonction) ; il contient donc un ouvert (lui-même). Par suite, la statistique est minimale et également complète.

6°. Oui ! Le modèle est régulier : il est dominé par la mesure de comptage, homogène car à support sur $\{0,1\}^n$, la vraisemblance est strictement positive et de classe C^2 en tant que fonction de θ. Enfin,la dérivée de la log-vraisemblance est de carré intégrable.

7°. On a :

$$\mathbb{E}_\theta[\overline{X}] = \mathbb{E}_\theta[X] = \theta \quad (13)$$

$\overline{X}$ est donc un estimateur sans biais de θ et est également complet. D'après le théorème de Lehmann-Scheffé, on en déduit que

$$T = \mathbb{E}_\theta[\overline{X}|\overline{X}] = \overline{X} \quad (14)$$

est VUMSB.

8°. Nous avons vu que le modèle est régulier et que la log-vraisemblance l est bien définie. L'information de Fisher existe donc et l'on a :

$$l(x,\theta) = s\ln\theta + (n-s)\ln(1-\theta) \quad (15)$$

Par suite

$$\frac{\partial l}{\partial\theta}(x,\theta) = \frac{s}{\theta} - \frac{n-s}{\theta} = \frac{s}{\theta(1-\theta)} - \frac{n}{1-\theta} \quad (16)$$

Ainsi,

$$\mathbb{I}_X(\theta) = \mathbb{V}_\theta\left(\frac{S}{\theta(1-\theta)}\right) \quad (17)$$

$$= \frac{1}{\theta^2(1-\theta)^2}\mathbb{V}_\theta(S) = \frac{n}{\theta(1-\theta)} \quad (18)$$

Un calcul rapide montre que $\mathbb{V}_\theta(\overline{X}) = \theta(1-\theta)/n = \mathbb{I}_X(\theta)^{-1}$. $\overline{X}$ atteint donc la borne de Cramer Rao. Par suite il est efficace pour θ.

9°. La loi *a posteriori* est proportionnelle à

$$L(\theta|s) = \theta^s(1-\theta)^{n-s}\theta^{a-1}(1-\theta)^{b-1}\mathbb{1}_{[0,1]}(\theta) \quad (19)$$

$$= \theta^{s+a-1}(1-\theta)^{n-s+b-1}\mathbb{1}_{[0,1]}(\theta) \quad (20)$$

et suit donc une loi bêta de paramètres $\mathscr{B}(s+a, n-s+b)$, dont l'espérance (*a posteriori*) est

$$\Delta(x) = \mathbb{E}[H|X = x] = \frac{s+a}{a+b+n} \quad (21)$$

et donc

$$\Delta = \mathbb{E}[H|X] = \frac{S_n + a}{a+b+n} \quad (22)$$

$$\mathbb{E}[\Delta] = \frac{n\theta + a}{a+b+n} \overset{n\to+\infty}{\to} \theta \quad (23)$$

10°. Les X_i son i.i.d. donc:

$$\mathbb{E}_\theta[Y] = \mathbb{E}_\theta[X_1]\mathbb{E}_\theta[X_2] = \mathbb{E}_\theta[X_i]^2 = \theta^2. \quad (24)$$

Ainsi, Y est un estimateur sans biais de θ^2. On considère maintenant la statistique $T_n = \mathbb{E}_\theta[Y \mid S_n]$. Montrons que T_n est un estimateur sans biais de variance minimale de θ^2. S_n est une statistique complète pour θ et pour θ^2. En effet, la complétude est une propriété de la famille de lois du modèle et non de la fonction du paramètre. Quand on change le paramétrage de θ en θ^2, on ne change pas la famille de lois et donc S_n reste exhaustive et complète pour θ^2. On peut alors appliquer directement le théorème de Lehmann-Scheffé, et l'on déduit que

$$T_n = \mathbb{E}[Y|S_n] \quad (25)$$

est l'estimateur VUMSB de θ^2.

11°. Si $s = 0, 1$, alors $S_n = s = \sum_{i=1}^n x_i$ implique qu'au plus l'un des x_i est non nul. En ce cas le produit est nul et

$$T_n(x) = \mathbb{E}[Y|S_n = s] = 0 \quad (26)$$

Si $s \ge 2$, Y est non nul seulement si les deux premiers X_i sont non nuls. Les X_i valent 0 ou 1 donc leur produit également: $Y = 0$ ou 1 et son espérance conditionnelle sachant $[S = s]$ est donc égale à $\mathbb{P}[Y = 1|S_n = s]$.

$$\mathbb{P}[Y = 1|S_n = s] = \frac{\mathbb{P}\left([X_1 = 1] \cap [X_2 = 1] \cap [\sum_{i=3}^n X_i = s-2]\right)}{\mathbb{P}[S = s]},$$

et par indépendance et identique distribution des X_i,

$$\mathbb{P}[Y = 1|S_n = s] = \frac{\theta^2\begin{pmatrix} n-2 \\ s-2 \end{pmatrix}\theta^{s-2}(1-\theta)^{n-s}}{\begin{pmatrix} n \\ s \end{pmatrix}\theta^s(1-\theta)^{n-s}} \quad (27)$$

$$= \begin{pmatrix} n-2 \\ s-2 \end{pmatrix} \Big/ \begin{pmatrix} n \\ s \end{pmatrix} = \frac{s(s-1)}{n(n-1)} \quad (28)$$

Ainsi,

$$T_n = \mathbb{E}[Y|S_n] = \frac{S_n(S_n-1)}{n(n-1)}. \quad (29)$$

Autres méthodes pour les questions 10° et 11°.

10°. Estimation de θ^2 par reparamétrisation. On souhaite estimer θ^2 à partir d'un échantillon $X_1, \ldots, X_n$ de variables indépendantes de loi de Bernoulli de paramètre $\theta \in]0,1[$.

Posons $\lambda = \theta^2$. Puisque $\theta \in]0,1[$, on a $\lambda \in]0,1[$ et $\theta = \sqrt{\lambda}$. La loi de chaque observation X_i est maintenant une loi de Bernoulli de paramètre $\sqrt{\lambda}$:

$$\mathbb{P}_\lambda[X_i = x_i] = (\sqrt{\lambda})^{x_i}(1-\sqrt{\lambda})^{1-x_i}.$$

La vraisemblance de l'échantillon devient

$$L(x,\lambda) = \lambda^{S_n/2}(1-\sqrt{\lambda})^{n-S_n}.$$

D'après le théorème de factorisation, S_n est exhaustive pour le paramètre λ. Démontrons maintenant sa complétude dans cette nouvelle paramétrisation. Soit h une fonction telle que

$$\mathbb{E}_\lambda[h(S_n)] = 0, \ \forall \lambda \in]0,1[.$$

Comme S_n suit une loi binomiale de paramètres n et $\sqrt{\lambda}$, cette condition s'écrit

$$\sum_{s=0}^{n} h(s)\binom{n}{s}(\sqrt{\lambda})^s(1-\sqrt{\lambda})^{n-s} = 0.$$

En divisant par $(1-\sqrt{\lambda})^n > 0$,

$$\sum_{s=0}^{n} h(s)\binom{n}{s}\left(\frac{\sqrt{\lambda}}{1-\sqrt{\lambda}}\right)^s = 0.$$

Effectuons le changement de variable

$$t = \frac{\sqrt{\lambda}}{1-\sqrt{\lambda}}.$$

Lorsque λ parcourt $]0,1[$, t parcourt $]0,+\infty[$.
On obtient donc

$$\sum_{s=0}^{n} h(s)\binom{n}{s}t^s = 0, \qquad \forall\, t > 0.$$

Le membre de gauche est un polynôme de degré inférieur ou égal à n, nul sur un intervalle ouvert. Il est donc identiquement nul. Tous ses coefficients sont nuls. On en déduit

$$h(s) = 0, \ s = 0, \dots, n.$$

Ainsi, S_n est également complète dans le modèle paramétré par λ. Enfin, $Y = X_1X_2$ vérifie

$$\mathbb{E}_\lambda[Y] = (\sqrt{\lambda})^2 = \lambda.$$

L'estimateur Y est donc sans biais pour λ.
Puisque S_n est exhaustive et complète pour λ, le théorème de Lehmann–Scheffé montre que $T_n = \mathbb{E}_\lambda[Y \mid S_n]$ est l'unique estimateur sans biais de variance minimale de $\lambda = \theta^2$.

11°. Vérification directe de l'absence de biais. En utilisant l'expression obtenue précédemment,

$$\mathbb{E}_\theta[T_n] = \frac{\mathbb{E}_\theta[S_n^2] - \mathbb{E}_\theta[S_n]}{n(n-1)}.$$

Puisque $S_n \sim \mathscr{B}(n,\theta)$, on a

$$\mathbb{E}_\theta[S_n] = n\theta \text{ et } \mathbb{V}_\theta(S_n) = n\theta(1-\theta).$$

D'où

$$\mathbb{E}_\theta[S_n^2] = \mathbb{V}_\theta(S_n) + \mathbb{E}_\theta[S_n]^2 = n\theta(1-\theta) + n^2\theta^2.$$

Il vient alors

$$\mathbb{E}_\theta[T_n] = \frac{n\theta(1-\theta) + n^2\theta^2 - n\theta}{n(n-1)} = \frac{n(n-1)\theta^2}{n(n-1)} = \theta^2.$$

Clairement, il est plus facile de le prouver en utilisant les propriétés de l'espérance conditionnelle:

$$\mathbb{E}_\theta[T_n] = \mathbb{E}[\mathbb{E}[Y|S_n]] = \mathbb{E}[Y] = \theta^2.$$

L'estimateur est donc bien sans biais, comme annoncé à la question 10°.