STATISTIQUE MATHÉMATIQUE - CORRIGÉ DE L'EXAMEN.

ENSAI - Mastères spécialisés - 2024/2025 -

ENSAI École nationale de la statistique et de l'analyse de l'information

I. 14 points

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

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

où $s = \sum_{i=1}^n x_i$. $\boxed{\text{0.5 point}}$

On voit facilement que

$$L(x,\theta) = \exp\left(n\overline{x}\ln\left(\frac{\theta}{1-\theta}\right) - n\ln(1-\theta)\right) \tag{2}$$

ce qui montre que le modèle est exponentiel et que $\overline{X}$ est la statistique naturelle $\boxed{\text{0.5 point}}$

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

$$l(x,\theta) = s\ln\theta + (n-s)\ln(1-\theta) \tag{3}$$

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

$$\frac{\partial l}{\partial\theta}(x,\theta) = \frac{s}{\theta} - \frac{n-s}{1-\theta} = \frac{s}{\theta(1-\theta)} - \frac{n}{1-\theta} \tag{4}$$

$$= \frac{s-n\theta}{\theta(1-\theta)} \tag{5}$$

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. $\boxed{\text{1 point}}$

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. $\boxed{\text{0.5 point}}$

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 :

$$\widehat{\theta}_n = \overline{X} \underset{N\to+\infty}{\longrightarrow} \mathbb{E}[X_1] = \theta \text{ p.s.} \tag{6}$$

Ceci prouve que $\widehat{\theta}_n$ est fortement consistant $\boxed{\text{1 point}}$

$\widehat{\theta}_n$ est également sans biais car $\mathbb{E}[\overline{X}] = n \times \mathbb{E}[X_1]/n = \theta$. $\boxed{\text{0.5 point}}$

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

$$\sqrt{n}\left(\widehat{\theta}_n - \theta\right) \rightsquigarrow \mathcal{N}(0,\sigma^2) \tag{7}$$

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

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

$$\lim_{n\to+\infty} \mathbb{P}\left[a \le \sqrt{n}\left(\frac{\widehat{\theta}_n - \theta}{\sigma}\right) \le b\right] = \Pi(b) - \Pi(a) \tag{8}$$

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. $\boxed{\text{1.5 points}}$.

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. $\boxed{\text{1 point}}$

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. $\boxed{\text{0.5 point}}$

7°. On a :

$$\mathbb{E}_\theta[\overline{X}] = \mathbb{E}_\theta[X] = \theta \tag{9}$$

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

$$T = \mathbb{E}_\theta[\overline{X}|\overline{X}] = \overline{X} \tag{10}$$

est VUMSB. $\boxed{\text{1 point}}$

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 :

$$l(x,\theta) = s\ln\theta + (n-s)\ln(1-\theta) \tag{11}$$

Par suite

$$\frac{\partial l}{\partial\theta}(x,\theta) = \frac{s}{\theta} - \frac{n-s}{\theta} = \frac{s}{\theta(1-\theta)} - \frac{n}{1-\theta} \tag{12}$$

Ainsi,

$$\mathbb{I}_X(\theta) = \mathbb{V}_\theta\left(\frac{S}{\theta(1-\theta)}\right) \tag{13}$$

$$= \frac{1}{\theta^2(1-\theta)^2}\mathbb{V}_\theta(S) = \frac{n}{\theta(1-\theta)} \tag{14}$$

$\boxed{\text{1 point}}$

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. $\boxed{\text{1 point}}$

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

$$L(\theta|s) = \theta^s(1-\theta)^{n-s}\theta^{a-1}(1-\theta)^{b-1}\mathbb{1}_{[0,1]}(\theta) \tag{15}$$
$$= \theta^{s+a-1}(1-\theta)^{n-s+b-1}\mathbb{1}_{[0,1]}(\theta) \tag{16}$$

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

$$\Delta(x) = \mathbb{E}[H|X=x] = \frac{s+a}{a+b+n} \tag{17}$$

et donc

$$\Delta = \mathbb{E}[H|X] = \frac{S_n+a}{a+b+n} \tag{18}$$

$\boxed{\text{1 point}}$

$$\mathbb{E}[\Delta] = \frac{n\theta+a}{a+b+n} \overset{n\to+\infty}{\longrightarrow} \theta \tag{19}$$

$\boxed{\text{0.5 point}}$

10°. On a :

$$\mathbb{E}_\theta[Y] = \mathbb{E}_\theta[X_i]^2 = \theta^2. \tag{20}$$

$\boxed{\text{0.5 point}}$

S_n est une statistique complète de θ (mais pas forcément de θ^2 !) d'après les questions précédentes et Y un estimateur sans biais de θ^2. En appliquant à nouveau le théorème de Lehmann-Scheffé, on en déduit que

$$T_n = \mathbb{E}[Y|S_n] \tag{21}$$

est l'estimateur VUMSB de θ^2. $\boxed{\text{1 point}}$

En effet, l'estimateur VUMSB est une fonction $h(S)$ de S (car S est complete pour θ) qui doit vérifier

$$\mathbb{E}[h(S)] = \theta^2 = \mathbb{E}\left[\mathbb{E}\left[Y|\psi(S)\right]\right] \tag{22}$$

Il s'avère que $\mathbb{E}[Y|S_n]$ convient (et donc finalement $\psi = \mathrm{Id}$) d'après le calcul qui suit.

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

$$T_n(x) = \mathbb{E}[Y|S_n = s] = 0 \tag{23}$$

Si $s \geq 2$, 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,

$$\mathbb{P}[Y=1|S_n=s] = \frac{\theta^2\begin{pmatrix} n-2 \\ s-2 \end{pmatrix}\theta^{s-2}(1-\theta)^{n-s}}{\begin{pmatrix} n \\ s \end{pmatrix}\theta^s(1-\theta)^{n-s}} \tag{24}$$
$$= \begin{pmatrix} n-2 \\ s-2 \end{pmatrix} / \begin{pmatrix} n \\ s \end{pmatrix} = \frac{s(s-1)}{n(n-1)} \tag{25}$$

Ainsi,

$$T_n = \mathbb{E}[Y|S_n] = \frac{S_n(S_n-1)}{n(n-1)}\mathbb{1}_{[S_n\geq 2]} \tag{26}$$

$\boxed{\text{2 points}}$

II. 8 points

1°. Y_1 est une variable aléatoire de Bernoulli de paramètre $\mathbb{P}[X_1 > 2]$. $n\overline{Y}_n$ est la somme de n variables de Bernoulli indépendantes et identiquement distribuées, elle suit donc une loi binomiale de paramètres n et $p = \mathbb{P}[X_1 > 2]$ $\boxed{\text{1 point}}$. On a alors immédiatement

$$\mathbb{E}[Y_1] = \mathbb{P}[X_1 > 2] = \int_2^{+\infty} \theta e^{-\theta x}dx = e^{-2\theta} = \lambda. \tag{27}$$

$\boxed{\text{1 point}}$

2°.

$$\mathbb{P}[\overline{Y}_n \neq 0] = 1 - \mathbb{P}[\overline{Y}_n = 0] = 1-(1-\lambda)^n. \tag{28}$$

Comme $\lambda \in]0,1[$, $(1-\lambda)^n$ tend vers 0 avec n et la probabilité recherchée tend vers 1 $\boxed{\text{1 point}}$.

3°. Nous appliquons la loi forte des grands nombres aux Y_i, qui sont des v.a.i.i.d. ayant un moment d'ordre 1 ($\mathbb{E}[|Y_1|] = \mathbb{E}[Y_1] < \infty$) : $\overline{Y}_n$ converge presque sûrement vers $\mathbb{E}[Y_1] = \lambda$ lorsque n tend vers l'infini $\boxed{\text{1 point}}$. En appliquant le théorème de la limite centrale (les Y_i ont une variance finie), il vient également

$$\sqrt{n}\left(\overline{Y}_n - \lambda\right) \rightsquigarrow \mathcal{N}(0,\sigma^2) \tag{29}$$

avec $\sigma^2 = \lambda(\lambda-1) = e^{-2\theta}(1-e^{-2\theta})$ $\boxed{\text{1 point}}$.

4°. $\overline{Y}_n$ converge presque sûrement vers un nombre strictement positif. Par définition de la limite, cette quantité est donc aussi proche que l'on veut de $\lambda > 0$ pour n assez grand, elle est donc strictement positive pour n assez grand $\boxed{\text{1 point}}$.

5°. Il suffit d'appliquer aux résultats de la question 3° le théorème de l'application continue et la méthode delta à la fonction $\psi(x) = -\ln x/2$ qui est continue et dérivable pour $x > 0$. La loi forte des grands nombres assure alors que $\widehat{\theta}_n$ converge presque sûrement vers θ et

$$\sqrt{n}\left(\widehat{\theta}_n - \theta\right) \rightsquigarrow \mathcal{N}(0,\tilde{\sigma}^2) \tag{30}$$

avec $\tilde{\sigma}^2 = \psi'^2(\theta)\sigma^2 = e^{-2\theta}(1-e^{-2\theta})/(4\theta^2)$ $\boxed{\text{2 points}}$.