Examen Statistique
Mercredi 4 décembre 2013

Exercice 1

Enoncé

On considère n variables aléatoires $X_1, ..., X_n$ indépendantes suivant la même loi discrète à valeur dans l'ensemble des entiers non nuls $\mathbb{N}^* = \mathbb{N} \setminus \{0\}$ définie par

$$P\left[X_i = x_i; \theta\right] = \frac{\theta^{x_i - 1}}{(1+\theta)^{x_i}}, \qquad x_i = 1, 2, ...$$

avec $\theta > 0$. On admettra que la moyenne et la variance d'une telle loi sont définies par $E\left[X_i\right] = \theta + 1$ et $\mathrm{var}(X_i) = \theta\left(\theta + 1\right)$.

1) Montrer que la vraisemblance de $(x_1, ..., x_n)$ admet un unique maximum global pour une valeur de θ que l'on déterminera. En déduire l'estimateur du maximum de vraisemblance du paramètre θ noté $\widehat{\theta}_{\mathrm{MV}}$.

2) L'estimateur $\widehat{\theta}_{\mathrm{MV}}$ est-il un estimateur sans biais et convergent du paramètre θ ?

3) Déterminer la borne de Cramer-Rao pour un estimateur non biaisé du paramètre θ. L'estimateur $\widehat{\theta}_{\mathrm{MV}}$ est-il l'estimateur efficace du paramètre θ ?

5) On désire désormais estimer le paramètre $a = \frac{1}{1+\theta} \in \left]0, 1\right[$.

- Quel est l'estimateur du maximum de vraisemblance du paramètre a noté $\widehat{a}_{\mathrm{MV}}$?
- On suppose que le paramètre a est muni d'une loi a priori (appelée loi beta de paramètres α et β) de densité
$$p(a) = \frac{a^{\alpha-1}(1-a)^{\beta-1}}{B\left(\alpha, \beta\right)} I_{]0,1[}(a)$$
où $I_{]0,1[}(a)$ est la fonction indicatrice sur l'intervalle $]0, 1[$ ($I_{]0,1[}(a) = 1$ si $a \in]0, 1[$ et $I_{]0,1[}(a) = 1$ si $a \notin]0, 1[$) et où $B(\alpha, \beta)$ est la fonction beta dont l'expression n'est pas importante dans cet exercice. Montrer que la loi a posteriori de $a|\, x_1, ..., x_n$ est aussi une loi beta dont on précisera les paramètres.
- Déterminer l'estimateur du maximum a posteriori du paramètre a noté $\widehat{a}_{\mathrm{MAP}}$ et étudier son comportement lorsque $n \to \infty$.

Réponses

1) (2pts) La vraisemblance de $(x_1, ..., x_n)$ est définie par

$$\begin{aligned} L(x_1, ..., x_n; \theta) &= \prod_{i=1}^{n} P\left[X_i = x_i; \theta\right] \\ &= \prod_{i=1}^{n} \left[\frac{\theta^{x_i - 1}}{(1+\theta)^{x_i}}\right] \\ &= \frac{\theta^{n\overline{x} - n}}{(1+\theta)^{n\overline{x}}}. \end{aligned}$$

avec $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$. La log-vraisemblance s'écrit donc

$$\ln L(x_1, ..., x_n; \theta) = n\left(\overline{x} - 1\right)\ln\theta - n\overline{x}\ln\left(1+\theta\right).$$

En étudiant le signe de $\frac{\partial \ln L(x_1,...,x_n;\theta)}{\partial\theta}$, on obtient :

$$\begin{aligned} \frac{\partial \ln L(x_1, ..., x_n; \theta)}{\partial\theta} &\geq 0 \Longleftrightarrow \frac{\overline{x} - 1}{\theta} - \frac{\overline{x}}{1+\theta} \geq 0, \\ &\Longleftrightarrow \theta \leq \overline{x} - 1. \end{aligned}$$

On remarque donc (il suffit de faire un tableau de variation) que la vraisemblance possède un maximum global unique défini par

$$\theta = \overline{x} - 1.$$

d'où

$$\widehat{\theta}_{\mathrm{MV}} = \overline{X} - 1 \text{ avec } \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

2) (1,5pt) On a de manière très classique

$$\begin{aligned} E\left[\widehat{\theta}_{\mathrm{MV}}\right] &= E\left[\overline{X} - 1\right] = \theta \\ \mathrm{var}\left[\widehat{\theta}_{\mathrm{MV}}\right] &= \mathrm{var}\left[\overline{X} - 1\right] = \frac{\mathrm{var}\left[X_1\right]}{n} = \frac{\theta\left(\theta+1\right)}{n}. \end{aligned}$$

L'estimateur $\widehat{\theta}_{\mathrm{MV}}$ est donc un estimateur sans biais et convergent du paramètre θ.

3) (1,5pt) La borne de Cramer-Rao pour un estimateur non biaisé du paramètre θ est définie par

$$\mathrm{BCR}\left(\theta\right) = \frac{-1}{E\left[\frac{\partial^2 \ln L(X_1,...,X_n;\theta)}{\partial\theta^2}\right]}$$

Mais

$$\frac{\partial \ln L(x_1, ..., x_n; \theta)}{\partial\theta} = \frac{n\left(\overline{x} - 1\right)}{\theta} - \frac{n\overline{x}}{1+\theta}$$

et donc

$$\frac{\partial^2 \ln L(x_1, ..., x_n; \theta)}{\partial\theta^2} = \frac{-n\left(\overline{x} - 1\right)}{\theta^2} + \frac{n\overline{x}}{\left(1+\theta\right)^2}.$$

On en déduit

$$E\left[\frac{\partial^2 \ln L(X_1,...,X_n;\theta)}{\partial\theta^2}\right] = \frac{-n}{\theta} + \frac{n}{(1+\theta)} = \frac{-n}{\theta\,(1+\theta)}$$

et par conséquent

$$\mathrm{BCR}\,(\theta) = \frac{\theta\,(1+\theta)}{n}.$$

Puisque var$\left[\widehat{\theta}_{\mathrm{MV}}\right]$ =BCR(θ), l'estimateur $\widehat{\theta}_{\mathrm{MV}}$ est l'estimateur efficace du paramètre θ.
4)

- (1pt) En utilisant la propriété d'invariance fonctionnelle, on a directement

$$\begin{aligned}\widehat{a}_{\mathrm{MV}} &= \frac{1}{1+\widehat{\theta}_{\mathrm{MV}}}\\ &= \frac{1}{\overline{X}}\end{aligned}$$

- (2pts) La loi a posteriori de $a|\,x_1,...,x_n$ s'écrit

$$\begin{aligned}f\,(a|\,x_1,...,x_n) &\propto L(x_1,...,x_n;a)p(a)\\ &\propto L\left(x_1,...,x_n;\theta=\frac{1-a}{a}\right)a^{\alpha-1}(1-a)^{\beta-1}I_{]0,1[}\,(a)\\ &\propto a^{n+\alpha-1}\,(1-a)^{n\overline{x}-n+\beta-1}\,I_{]0,1[}\,(a)\end{aligned}$$

  Cette densité est la densité d'une loi beta

$$a|\,x_1,...,x_n \sim B\,(n+\alpha, n\overline{x}-n+\beta)$$

- (1pt) L'estimateur du maximum a posteriori du paramètre a noté $\widehat{a}_{\mathrm{MAP}}$ est obtenu en maximisant le logarithme de la loi a posteriori $f\,(a|\,x_1,...,x_n)$. Mais

$$\ln\left[f\,(a|\,x_1,...,x_n)\right] = (n+\alpha-1)\ln a + (n\overline{x}-n+\beta-1)\ln(1-a)$$

  d'où

$$\begin{aligned}\frac{\partial\ln\left[f\,(a|\,x_1,...,x_n)\right]}{\partial a} &\geq 0 \Longleftrightarrow \frac{n+\alpha-1}{a} - \frac{n\overline{x}-n+\beta-1}{1-a} \geq 0\\ &\Longleftrightarrow a \leq \frac{n+\alpha-1}{n\overline{x}+\alpha+\beta-2}\end{aligned}$$

  La loi a posteriori $f\,(a|\,x_1,...,x_n)$ possède donc un unique maximum global, d'où

$$\widehat{a}_{\mathrm{MAP}} = \frac{n+\alpha-1}{n\overline{X}+\alpha+\beta-2}$$

- (1pt) On peut exprimer $\widehat{\theta}_{\text{MAP}}$ en fonction de $\widehat{\theta}_{\text{MV}}$ puisque

$$\begin{aligned}\widehat{\theta}_{\text{MAP}} &= \frac{1+\frac{\alpha-1}{n}}{\overline{X}+\frac{\alpha+\beta-2}{n}} \\ &= \frac{1+\frac{\alpha-1}{n}}{\frac{1}{\widehat{\theta}_{\text{MV}}}+\frac{\alpha+\beta-2}{n}}\end{aligned}$$

L'estimateur $\widehat{\theta}_{\text{MAP}}$ se comporte donc comme $\widehat{\theta}_{\text{MV}}$ lorsque $n \to \infty$. Lorsqu'on a beaucoup d'observations, on fait confiance à ces observations et donc l'effet de la loi a priori est négligeable.

Exercice 2

Enoncé

On considère n vecteurs aléatoires $X_1, ..., X_n$ appartenant à $\mathbb{R}^2$ définis par

$$X_i = M\theta + e_i$$

où M est une matrice connue de taille 2×2, θ est un vecteur inconnu de $\mathbb{R}^2$ et

$$e_i \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \sigma^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right).$$

Pour simplifier, on notera $e_i \sim \mathcal{N}(0, \sigma^2 I_2)$ (où 0 est le vecteur nul de $\mathbb{R}^2$ et I_2 est la matrice identité de taille 2×2) et on suppose que les vecteurs $e_1, ..., e_n$ sont indépendants. On suppose également (au moins dans un premier temps) que le paramètre σ^2 est connu. On désire effectuer le test d'hypothèses

$$\begin{aligned} H_0 &: \theta = \theta_0 = (0,0)^T \\ H_1 &: \theta = \theta_1 \neq \theta_0 \end{aligned}$$

avec $M\theta_1 \neq 0$.

1) Montrer que la statistique du test de Neyman-Pearson peut s'écrire

$$T = \theta_1^T M^T \sum_{i=1}^{n} X_i$$

et donner la région critique de ce test.

2) On rappelle que si X est un vecteur gaussien de $\mathbb{R}^2$ de moyenne μ et de matrice de covariance Σ et si a est un vecteur non nul de $\mathbb{R}^2$, alors $a^T X$ est une variable aléatoire gaussienne de moyenne $E\left[a^T X\right] = a^T \mu$ et de variance $a^T \Sigma a$. En déduire la loi de T sous les deux hypothèses H_0 et H_1.

3) En utilisant les résultats de la question précédente, exprimer les risques de première et de seconde espèces α et β en fonction du seuil du test de Neyman Pearson noté S_α, du paramètre θ_1 et de la fonction $\phi(x) = 1 - F(x)$ où F est la fonction de répartition de la loi normale $\mathcal{N}(0,1)$. En déduire l'expression analytique des courbes COR de ce détecteur et montrer qu'elles dépendent de n, σ et de $\nu = \|M\theta_1\| = \sqrt{\theta_1^T M^T M \theta_1}$. Représenter graphiquement la forme approximative de ces courbes COR pour différentes valeurs des paramètres n et de σ.

4) On suppose désormais que le paramètre σ^2 est inconnu. Déterminer l'estimateur du maximum de vraisemblance de ce paramètre. En déduire la règle de décision liée au test du rapport des vraisemblances maximales.

Réponses

1) (2pts) Le test de Neyman-Pearson est défini par

$$\text{Rejet de } H_0 \text{ si} \quad \frac{L(x_1, ..., x_n; \theta_1)}{L(x_1, ..., x_n; \theta_0)} > k_\alpha.$$

Mais la vraisemblance de X_i s'écrit (rappelons nous le cours sur les vecteurs gaussiens)

$$\begin{aligned} f(x_i; \theta) &= \frac{1}{2\pi\sigma^2} \exp\left[-\frac{1}{2\sigma^2}(x_i - M\theta)^T (x_i - M\theta)\right] \\ &= \frac{1}{2\pi\sigma^2} \exp\left[-\frac{1}{2\sigma^2}\|x_i - M\theta\|^2\right] \end{aligned}$$

et en conséquence la vraisemblance est

$$L(x_1, ..., x_n; \theta) = \frac{1}{(2\pi\sigma^2)^n} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^n \|x_i - M\theta\|^2\right]$$

d'où

$$\begin{aligned} \frac{L(x_1, ..., x_n; \theta_1)}{L(x_1, ..., x_n; \theta_0)} &> k_\alpha \Leftrightarrow \sum_{i=1}^n \left[\|x_i\|^2 - \|x_i - M\theta_1\|^2\right] > K_\alpha \\ &\Leftrightarrow \theta_1^T M^T \sum_{i=1}^n x_i > S_\alpha \end{aligned}$$

On en déduit le test équivalent

$$\text{Rejet de } H_0 \text{ si} \quad \theta_1^T M^T \sum_{i=1}^n x_i > S_\alpha.$$

La statistique du test de Neyman-Pearson est donc

$$T = \theta_1^T M^T \sum_{i=1}^n X_i.$$

La région critique de ce test est définie par

$$\left\{ x_1, ..., x_n \middle| \theta_1^T M^T \sum_{i=1}^n x_i > S_\alpha \right\}.$$

2) (2pts) Sous l'hypothèse H_0, on a

$$X_i \,|H_0 \sim \mathcal{N}\left(0, \sigma^2 I_2\right)$$

tandis que sous l'hypothèse H_1

$$X_i \,|H_0 \sim \mathcal{N}\left(M\theta_1, \sigma^2 I_2\right).$$

En utilisant le rappel de l'énoncé, on obtient

$$\begin{aligned}
\theta_1^T M^T X_i \,|H_0 &\sim \mathcal{N}\left(0, \sigma^2 \theta_1^T M^T M \theta_1\right) = \mathcal{N}\left(0, \sigma^2 \left\|M\theta_1\right\|^2\right) \\
\theta_1^T M^T X_i \,|H_1 &\sim \mathcal{N}\left(\theta_1^T M^T M \theta_1, \sigma^2 \theta_1^T M^T M \theta_1\right) = \mathcal{N}\left(\left\|M\theta_1\right\|^2, \sigma^2 \left\|M\theta_1\right\|^2\right)
\end{aligned}$$

d'où, puisque les vecteurs X_i sont indépendants

$$\begin{aligned}
T\,|H_0 &\sim \mathcal{N}\left(0, n\sigma^2 \left\|M\theta_1\right\|^2\right) \\
T\,|H_1 &\sim \mathcal{N}\left(n\left\|M\theta_1\right\|^2, n\sigma^2 \left\|M\theta_1\right\|^2\right)
\end{aligned}$$

3) (4pts) Le risque de première espèce α est défini par

$$\begin{aligned}
\alpha &= P\left[\text{Rejeter } H_0 \mid H_0 \text{ vraie}\right] \\
&= P\left[T > S_\alpha \middle| \theta = 0\right] \\
&= P\left[\frac{T}{\sqrt{n\sigma^2 \left\|M\theta_1\right\|^2}} > \frac{S_\alpha}{\sqrt{n\sigma^2 \left\|M\theta_1\right\|^2}} \middle| \theta = 0\right] \\
&= \phi\left[\frac{S_\alpha}{\sqrt{n\sigma^2 \left\|M\theta_1\right\|^2}}\right].
\end{aligned}$$

De la même façon

$$\begin{aligned}
\beta &= P\left[\text{Rejeter } H_1 \mid H_1 \text{ vraie}\right] \\
&= P\left[T < S_\alpha \middle| \theta = \theta_1\right] \\
&= P\left[\frac{T - n\left\|M\theta_1\right\|^2}{\sqrt{n\sigma^2 \left\|M\theta_1\right\|^2}} < \frac{S_\alpha - n\left\|M\theta_1\right\|^2}{\sqrt{n\sigma^2 \left\|M\theta_1\right\|^2}} \middle| \theta = \theta_1\right] \\
&= 1 - \phi\left(\frac{S_\alpha - n\left\|M\theta_1\right\|^2}{\sqrt{n\sigma^2 \left\|M\theta_1\right\|^2}}\right).
\end{aligned}$$

Les courbes COR du détecteur expriment la puissance $\pi = 1 - \beta$ en fonction de α. On a donc

$$\begin{aligned}
\pi &= 1 - \beta \\
&= \phi\left(\frac{S_\alpha - n\left\|M\theta_1\right\|^2}{\sqrt{n\sigma^2 \left\|M\theta_1\right\|^2}}\right) \\
&= \phi\left(\frac{\sqrt{n\sigma^2 \left\|M\theta_1\right\|^2}\phi^{-1}(\alpha) - n\left\|M\theta_1\right\|^2}{\sqrt{n\sigma^2 \left\|M\theta_1\right\|^2}}\right) \\
&= \phi\left(\phi^{-1}(\alpha) - \frac{\sqrt{n}}{\sigma}\left\|M\theta_1\right\|\right)
\end{aligned}$$

Puisque ϕ est une fonction décroissante, le test de Neyman Pearson est meilleur lorsque n est "grand" ou lorsque σ est "petit", ce qui est naturel.

4) (3pts) La log-vraisemblance sous l'hypothèse H_1 s'écrit

$$\ln L(x_1, ..., x_n; \theta, \sigma^2) = -n \ln\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} \|x_i - M\theta\|^2 .$$

En étudiant le signe de $\frac{\partial \ln L(x_1,...,x_n;\theta,\sigma^2)}{\partial \sigma^2}$, on obtient :

$$\begin{aligned} \frac{\partial \ln L(x_1, ..., x_n; \theta, \sigma^2)}{\partial \sigma^2} &\geq 0 \Longleftrightarrow \frac{-n}{\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^{n} \|x_i - M\theta\|^2 \geq 0, \\ &\Longleftrightarrow \sigma^2 \leq \frac{1}{2n} \sum_{i=1}^{n} \|x_i - M\theta\|^2 . \end{aligned}$$

On remarque donc (il suffit de faire un tableau de variation) que la vraisemblance possède un maximum global unique défini par

$$\sigma^2 = \frac{1}{2n} \sum_{i=1}^{n} \|x_i - M\theta\|^2$$

d'où

$$\widehat{\sigma^2}_{\mathrm{MV}} = \frac{1}{2n} \sum_{i=1}^{n} \|X_i - M\theta\|^2 .$$

Le test du rapport des vraisemblances maximales est défini par

$$\text{Rejet de } H_0 \text{ si} \qquad \frac{\sup\limits_{\sigma^2, \theta=\theta_1} L(x_1, ..., x_n; \theta, \sigma^2)}{\sup\limits_{\sigma^2, \theta=0} L(x_1, ..., x_n; \theta, \sigma^2)} > k_\alpha.$$

Des calculs élémentaires conduisent aux résultats suivants

$$\begin{aligned} \sup_{\sigma^2, \theta=\theta_1} L(x_1, ..., x_n; \theta, \sigma^2) &= L\left(x_1, ..., x_n; \theta_1, \widehat{\sigma^2}_{\mathrm{MV}}\right) \\ &= \left(\frac{n}{\pi}\right)^n \left(\sum_{i=1}^{n} \|x_i - M\theta_1\|^2\right)^{-n} \exp(-n) \end{aligned}$$

et

$$\sup_{\sigma^2, \theta=0} L(x_1, ..., x_n; \theta, \sigma^2) = \left(\frac{n}{\pi}\right)^n \left(\sum_{i=1}^{n} \|x_i\|^2\right)^{-n} \exp(-n).$$

Le test du rapport des vraisemblances maximales s'écrit donc

$$\text{Rejet de } H_0 \text{ si} \qquad \frac{\sum_{i=1}^{n} \|x_i\|^2}{\sum_{i=1}^{n} \|x_i - M\theta_1\|^2} > S_\alpha$$

où S_α est un seuil dépendant du risque de première espèce α.