Université de Carthage
Ecole Supérieure de la Statistique
et de l'Analyse de l'Information

Année Universitaire
2010-2011
1ère Année
Durée : 1h30.

Devoir Surveillé
Méthodes d'estimation

Exercice 1 *Soit X une v.a. de loi Bernoulli de paramètre $\theta \in]0,1[$. On pose $\eta = \ln \frac{\theta}{1-\theta}$. Montrer, en reparamétrisant la loi de X par η, que cette loi appartient à une famille exponentielle canonique à un paramètre.*

Exercice 2 *Soit $\theta \in \mathbb{R}_+^*$ et une fonction $\phi :]\theta, \theta+1[\to \mathbb{R}_+^*$ telle que $\int_\theta^{\theta+1} \phi(x)dx < \infty$. On considère un échantillon $\underline{X} = (X_1, \ldots, X_n)$ issu d'une loi de probabilité de densité*

$$f(x\,,\,\theta) = c(\theta)\phi(x)\mathbb{1}_{]\theta,\theta+1[}(x)$$

où $c(\theta) = \left(\int_\theta^{\theta+1} \phi(x)dx\right)^{-1}$.

Partie I
- *(a) Calculer la fonction vraisemblance.*
- *(b) En déduire l'expression d'une statistique exhaustive.*

Partie II Dans cette partie, $\phi(x) = 1$ pour tout $x \in]\theta, \theta+1[$.
- *(a) Donner dans ce cas l'expression d'une statistique exhaustive qu'on désignera par T.*
- *(b) Montrer qu'il existe une application ψ non nulle telle que $E_\theta(\psi(T)) = 0$ pour tout $\theta > 0$.*
- *(c) T est-elle complète ?*

Exercice 3 *Soit $\underline{X} = (X_1, \ldots, X_n)$ un $n-$échantillon de loi de Poisson de paramètre $\theta > 0$. On considère $\overline{X}_n$ la moyenne empirique associée à $\underline{X} = (X_1, \ldots, X_n)$.*

1. *Montrer que la loi du modèle appartient à une famille exponentielle.*
2. *Calculer la fonction vraisemblance.*
3. *Calculer $E(n\overline{X}_n)$[1].*
4. *En déduire l'expression de l'EMV.*
5. *Retrouver le résultat de la question précédente en utilisant la fonction vraisemblance.*

Exercice 4 *Soit $\underline{X} = (X_1, \ldots, X_n)$ un n-échantillon d'un variable aléatoire X de loi exponentielle de paramètre $1+\theta$ où $\theta > 0$.*

1. *Calculer[2] $E(X^2)$.*
2. *En déduire, en utilisant la méthode des moments et la question précédente, un estimateur par la méthode des moments de θ.*
3. *(Facultatif) Quelle est la loi asymptotique de la loi de l'estimateur calculé à la question précédente.*

---

1. On admet que si X et Y sont deux v.a. indépendantes de lois de Poisson respectivement de paramètre μ et λ, alors $X+Y$ est aussi de loi de Poisson $\mu+\lambda$
2. Rappelons que $\int_0^\infty x^n e^{-x}dx = n!$, $\forall n \in \mathbb{N}$.

Corrigé de l'exercice 1 $X \rightsquigarrow b(\theta) \qquad \theta \in]0,1[$.

$\eta = \ln \dfrac{\theta}{1-\theta}$.

$P[X=x] = \theta^x (1-\theta)^{1-x}\, \mathbb{1}_{\{0,1\}}(x)\, \theta^x (1-\theta)^{1-x}$

$$\begin{aligned}
\mathcal{L}(\underline{x},\theta) &= \theta^{\sum_{i=1}^n x_i} (1-\theta)^{n-\sum_{i=1}^n x_i}\, \mathbb{1}_{\{0,1\}^n}(\underline{x}) \\
&= \exp\left(\ln\theta \cdot \sum_{i=1}^n x_i + \ln(1-\theta)\cdot\left(n - \sum_{i=1}^n x_i\right)\right) \mathbb{1}_{\{0,1\}^n}(\underline{x}) \\
&= \exp\left(\ln \tfrac{\theta}{1-\theta} \sum_{i=1}^n x_i + n\ln(1-\theta)\right) \mathbb{1}_{\{0,1\}^n}(\underline{x}) \\
&= \exp\left(\eta \sum_{i=1}^n x_i + n\ln\left(\frac{\exp\eta}{1+\exp\eta}\right)\right) \mathbb{1}_{\{0,1\}^n}(\underline{x})
\end{aligned}$$

Il s'agit d'une famille exponentielle à un paramètre sous forme canonique avec $T(\underline{X}) = \sum_{i=1}^n X_i \qquad d(\eta) = n\ln\left(\dfrac{\exp\eta}{1+\exp\eta}\right) \qquad S(\underline{X}) = 0$ *et* $A = \{0,1\}^n$ *indépendant de* η.

Corrigé de l'exercice 2 $\theta > 0 \qquad \phi :]\theta,\theta+1[\to \mathbb{R}_+^* \qquad \int_\theta^{\theta+1} \phi(x)dx < \infty$

$f(x\,,\,\theta) = c(\theta)\phi(x)\mathbb{1}_{]\theta,\theta+1[}(x) \qquad$ *où* $c(\theta) = \left(\int_\theta^{\theta+1}\phi(x)dx\right)^{-1}$

Partie I (a) $\mathcal{L}(\underline{x},\theta) = (c(\theta))^n \prod_{i=1}^n \phi(x_i)\mathbb{1}_{\{\min x_i > \theta\}}\mathbb{1}_{\{\max x_i < \theta+1\}}$

(b)

$$\mathcal{L}(\underline{x},\theta) = (c(\theta))^n \prod_{i=1}^n \phi(x_i)\mathbb{1}_{\{\min x_i > \theta\}}\mathbb{1}_{\{\max x_i < \theta+1\}} = g(T(\underline{x}),\theta)\cdot 1$$

Ainsi, d'après le théorème de factorisation, la statistique $T(\underline{X}) = (\min X_i, \max X_i)$ *est une statistique exhaustive.*

Partie II $\phi(x) = 1$ *pour tout* $x \in]\theta,\theta+1[$.

On a alors $f(x\,,\,\theta) = c(\theta)\mathbb{1}_{]\theta,\theta+1[}(x) = \left(\int_\theta^{\theta+1} dx\right)^{-1} \mathbb{1}_{]\theta,\theta+1[}(x) = \mathbb{1}_{]\theta,\theta+1[}(x)$

$X \rightsquigarrow \mathcal{U}_{]\theta,\theta+1[}$

(a) $\mathcal{L}(\underline{x},\theta) = \mathbb{1}_{\{\min x_i > \theta\}}\mathbb{1}_{\{\max x_i < \theta+1\}}$
$T = T(\underline{X}) = (\min X_i, \max X_i)$ *est une statistique exhaustive pour le modèle.*

(b) *Considérons l'application* $\psi : (a,b) \longmapsto b-a$
Soit la variable aléatoire Z *définie par* $Z = X - \theta$
$Z \rightsquigarrow \mathcal{U}_{]0,0+1[}$. *La loi de* Z *est indépendante de* θ.
Considérons maintenant l'application $\psi : (a,b) \longmapsto b-a$
$\psi(T) = \max X_i - \min X_i = \max Z_i - \min Z_i$.

(c) *On a alors* $E_\theta(\psi(T)) = 0$ *pour tout* $\theta > 0$
T *n'est donc pas complète puisque* $E_\theta(\psi(T)) = 0$ *pour tout* $\theta > 0$ *et* ψ *n'est pas l'application nulle.*

Corrigé de l'exercice 3 $X \rightsquigarrow \mathcal{P}(\theta) \qquad \theta >$

1. $f(x,\theta) = P[X = x] = \dfrac{\theta^{x_i}}{x_i!} \exp -\theta \quad \mathbb{1}_{\mathbb{N}}(x)$

   $f(x,\theta) = \exp(\ln\theta \cdot x - \theta - \ln(x!))\, \mathbb{1}_{\mathbb{N}}(x)$

   Il s'agit bien d'une famille exponentielle à un paramètre avec $t(x) = x \qquad c(\theta) = \ln\theta \qquad \delta(\theta) = -\theta \qquad S(\underline{X}) = 0$ *et* $A = \mathbb{N}$ *indépendant de* θ.

2. $\mathcal{L}(\underline{x},\theta) = \prod_{i=1}^{n} f(x_i,\theta) = \exp\left(\ln\theta \cdot \sum_{i=1}^{n} x_i - n\theta - \sum_{i=1}^{n} \ln(x_i!)\right) \mathbb{1}_{\mathbb{N}^n}(\underline{x})$

3. $E(n\overline{X}_n) = E\left(\sum_{i=1}^{n} X_i\right) = n\theta$

4. *Le modèle appartient à la famille exponentielle.* $\Theta = \mathbb{R}_+^*$ *est un ouvert.* $T(\underline{X}) = \sum_{i=1}^{n} X_i$ *est la statistique exhaustive.* $c : \theta \longmapsto \ln\theta$ *est injective et de classe* C^2 *et* $d : \theta \longmapsto -n\theta$ *est aussi de classe* C^2.

   $E(T(\underline{X})) = T(\underline{x}) \Longleftrightarrow n\theta = \sum_{i=1}^{n} x_i \Longrightarrow \widehat{\theta} = \overline{X}_n$ *est l'estimateur du maximum de vraisemblance.*

5. $l(\underline{x},\theta) = \left(\ln\theta \cdot \sum_{i=1}^{n} x_i - n\theta - \sum_{i=1}^{n} \ln(x_i!)\right) \quad \mathbb{1}_{\mathbb{N}^n}(\underline{x})$

   $$\begin{cases} \dfrac{\partial l(\underline{x},\theta)}{\partial\theta} = 0 \\ \dfrac{\partial^2 l(\underline{x},\theta)}{\partial\theta^2} < 0 \end{cases} \Longleftrightarrow \begin{cases} \dfrac{1}{\theta}\sum_{i=1}^{n} x_i - n = 0 \\ -\dfrac{1}{\theta^2} < 0 \end{cases} \Longleftrightarrow \begin{cases} \theta = \dfrac{1}{n}\sum_{i=1}^{n} x_i \\ -\dfrac{1}{\theta^2} < 0 \end{cases}$$

   On retrouve le même résultat.

Corrigé de l'exercice 4 $X \rightsquigarrow \mathcal{E}(1+\theta) \qquad \theta > 0.$

1. $f(x,\theta) = \frac{1}{1+\theta} \exp -(1+\theta)x \quad \mathbb{1}_{\{x>0\}}$

   $E(X^2) = \int_0^{+\infty} \frac{x^2}{1+\theta} \exp -(1+\theta)x \; dx$

   Posons $u = (1+\theta)x \Longrightarrow du = (1+\theta)dx$

   $E(X^2) = \int_0^{+\infty} \dfrac{u^2}{(1+\theta)^4} \exp -u \; du = \dfrac{2!}{(1+\theta)^4} = \dfrac{2}{(1+\theta)^4}$

2. $\phi(\theta) = \int_0^{+\infty} g(x)\, P(dx) \qquad$ *avec* $\phi : \theta \longmapsto \dfrac{2}{(1+\theta)^4}$ *et* $g : x \longmapsto x^2$

   ϕ *est injective sur son ensemble image.*

   $\theta = \phi^{-1}\left(\int_0^{+\infty} x^2 P(dx)\right)$

   $\phi^{-1}(y) = \dfrac{2}{\sqrt[4]{y}} - 1$

   $\theta = \frac{2}{\sqrt[4]{\int_0^{+\infty} x^2 P(dx)}} - 1 = \dfrac{2}{\sqrt[4]{E(X^2)}} - 1 = \dfrac{2}{\sqrt[4]{m_2}} - 1$

   Conclusion $\widehat{\theta} = \dfrac{2}{\sqrt[4]{\frac{1}{n}\sum_{i=1}^{n} X_i^2}} - 1$

3. $\widehat{\theta}$ *est un estimateur pnar la méthode des moments et* $\theta = h(m_2) = 2m_2^{-\frac{1}{4}} - 1$

$h\,:\,m_2 \longmapsto \dfrac{2}{\sqrt[4]{m_2}} - 1$ *h est au moins de classe* $\mathcal{C}^1$. $\widehat{\theta}$ *est donc asymptotiquement normal et converge en loi vers* $\mathcal{N}\left(\theta, \sigma_h^2\right)$ *avec*

$$\sigma_h^2 = Var\left(\frac{\partial h\left(\theta\right)}{\partial m_2} X_1^2\right) = Var\left(-\tfrac{1}{2} m_2^{-\frac{5}{4}} X_1^2\right) = \frac{1}{4} m_2^{-\frac{5}{2}} Var\left(X^2\right)$$