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

Année Universitaire
2013-2014
Première Année

Méthodes d'estimation
Devoir surveillé : Mars 2014

Enseignantes : H.Mallek et K.Moalla

Durée : 1 heure 30

Exercice 1 *Soit* $(X_1, ..., X_n)$ *un échantillon associé à un modèle statistique* $\left(\mathbb{R}_+^n, \left(Q_\theta^{\otimes n}\right)_{\theta>0}\right)$ *tel que pour tout* $\theta > 0$, Q_θ *est la loi sur* $\mathbb{R}_+$ *de densité*

$$f_\theta(x) = \frac{1}{2\sqrt{x\theta}} \mathbf{1}_{]0,\theta]}(x).$$

1. *Vérifier que* $\widehat{\theta}_{1,n} = \max_{1\leq i\leq n} X_i$ *est l'estimateur du maximum de vraisemblance de* θ.
2. *Calculer la fonction de répartition et la densité de* $\widehat{\theta}_{1,n}$. *En déduire les quantités* $E_\theta\left(\widehat{\theta}_{1,n}\right)$ *et* $V_\theta\left(\widehat{\theta}_{1,n}\right)$.
3. *Calculer l'espérance de* X_1. *En déduire un estimateur* $\widehat{\theta}_{2,n}$ *de* θ *et calculer son espérance.*
4. *Calculer les quantités* $E_\theta\left(\widehat{\theta}_{.,n} - \theta\right)^2$ *pour chacun des estimateurs* $\widehat{\theta}_{1,n}$ *et* $\widehat{\theta}_{2,n}$.

Exercice 2 (Spécifique au groupe AB)

Pour $\theta \in]0,1[$, *on considère la loi* P_θ *de densité* f_θ *définie sur* $\mathbb{R}$ *par*

$$f_\theta(x) = \theta^2 \exp(-\theta x)\,\mathbf{1}_{(x\geq 0)} + \theta(1-\theta)\exp(\theta x)\,\mathbf{1}_{(x\leq 0)}.$$

On suppose que les conditions de dérivation sous le signe d'intégration sont remplies.

1. *Calculer l'information de Fisher* I *du modèle* $\left(\mathbb{R}, (P_\theta)_{\theta\in]0,1[}\right)$.
2. *On note* P'_θ *la loi suivie par la valeur absolue d'une variable de loi* P_θ.
   - a- *Vérifier que* P'_θ *est la loi exponentielle de paramètre* θ.
   - b- *Calculer l'information de Fisher* I' *du modèle* $\left(\mathbb{R}, (P'_\theta)_{\theta\in]0,1[}\right)$.
3. *Comparer* I' *et* I.

Exercice 3 (Spécifique au groupe CD)

Soient $X_1, \cdots, X_n$ les revenus annuels de n ménages choisis au hasard dans une certaine population. On suppose que X, la variable aléatoire "Revenu" suit la loi de Dagum de paramètre θ, $\theta \in \Theta = \mathbb{R}_+^*$, autrement dit, sa fonction de répartition associée s'écrit

$$F(x, \theta) = \left(1 + x^{-2}\right)^{-\theta} \mathbb{1}_{\mathbb{R}_+^*}(x)$$

1. *Vérifier que la densité de probabilité associée à X a pour expression*

$$f(x, \theta) = \frac{2\theta}{x^3} \frac{1}{(1 + x^{-2})^{\theta+1}} \mathbb{1}_{\mathbb{R}_+^*}(x)$$

2. *Vérifier que cette loi appartient à la famille exponentielle à un paramètre. (on précisera $T(\underline{X})$, la statistique exhaustive complète associée au modèle).*

3. *Ecrire la vraisemblance du modèle;*

4. *Calculer $E_\theta [T(\underline{X})]$.*

5. *Donner l'expression de $\widehat{\theta}_n$, l'estimateur du maximum de vraisemblance de θ.*

Corrigé Exercice 1 $f_\theta(x) = \dfrac{1}{2\sqrt{x\theta}}\mathbf{1}_{]0,\theta]}(x)$.

1. $\mathcal{L}(\underline{x},\theta) = \dfrac{1}{2^n\theta^{\frac{n}{2}}\prod\limits_{i=1}^{n} x_i^{\frac{1}{2}}}\mathbf{1}_{\{\max\ x_i \leq \theta\}}\mathbf{1}_{\{\min\ x_i > 0\}}..$

   Posons $g(\theta) = \dfrac{1}{2^n\theta^{\frac{n}{2}}\prod\limits_{i=1}^{n} x_i^{\frac{1}{2}}}$, *g est strictement décroissante sur l'intervalle* $[\max X_i, +\infty[$.

   Donc l'estimateur dumaximum de vraisemblance de θ *est* $\widehat{\theta}_{1,n} = \max\limits_{1\leq i\leq n} X_i$

2. $P_\theta\left[\widehat{\theta}_{1,n} \leq x\right] = P_\theta\left[\max\limits_{1\leq i\leq n} X_i \leq x\right] = P_\theta[X_1 \leq x, ..., X_n \leq x] = (F_X(x))^n$.

   On a pour tout $x \in]0,\theta]$, $F_X(x) = \int\limits_0^x \dfrac{1}{2\sqrt{t\theta}}dt = \dfrac{1}{\sqrt{\theta}}\int\limits_0^x \dfrac{1}{2\sqrt{t}}dx = \dfrac{1}{\sqrt{\theta}}\left[\sqrt{t}\right]_0^x = \dfrac{\sqrt{x}}{\sqrt{\theta}} = \sqrt{\dfrac{x}{\theta}}$.

   D'où $F_X(x) = \sqrt{\dfrac{x}{\theta}}\mathbf{1}_{]0,\theta]}(x) + \mathbf{1}_{]\theta,+\infty]}(x)$.

   $P_\theta\left[\widehat{\theta}_{1,n} \leq x\right] = \left(\dfrac{x}{\theta}\right)^{\frac{n}{2}}\mathbf{1}_{]0,\theta]}(x) + \mathbf{1}_{]\theta,+\infty]}(x)$.

   $f_{\widehat{\theta}_{1,n}}(x) = \dfrac{n}{2}\dfrac{x^{\frac{n}{2}-1}}{\theta^{\frac{n}{2}}}\mathbf{1}_{]0,\theta]}(x)$.

   $E_\theta\left(\widehat{\theta}_{1,n}\right) = \int\limits_0^\theta \dfrac{n}{2}x\dfrac{x^{\frac{n}{2}-1}}{\theta^{\frac{n}{2}}}dx = \dfrac{n}{2}\left(\dfrac{1}{\theta}\right)^{\frac{n}{2}}\int\limits_0^\theta x^{\frac{n}{2}}\ dx = \dfrac{n}{2}\left(\dfrac{1}{\theta}\right)^{\frac{n}{2}}\dfrac{1}{\frac{n}{2}+1}\theta^{\frac{n}{2}+1} = \dfrac{n\theta}{n+2}$.

   $E_\theta\left(\widehat{\theta}_{1,n}\right)^2 = \int\limits_0^\theta \dfrac{n}{2}x^2\dfrac{x^{\frac{n}{2}-1}}{\theta^{\frac{n}{2}}}dx = \dfrac{n}{2}\left(\dfrac{1}{\theta}\right)^{\frac{n}{2}}\int\limits_0^\theta x^{\frac{n}{2}+1}\ dx = \dfrac{n}{2}\left(\dfrac{1}{\theta}\right)^{\frac{n}{2}}\dfrac{1}{\frac{n}{2}+2}\theta^{\frac{n}{2}+2} = \dfrac{n\theta^2}{n+4}$.

   $V_\theta\left(\widehat{\theta}_{1,n}\right) = \dfrac{n}{n+4}\theta^2 - \left(\dfrac{n}{n+2}\theta\right)^2 = n\theta^2\left(\dfrac{1}{n+4} - \dfrac{n}{(n+2)^2}\right) = n\theta^2\dfrac{(n+2)^2 - n(n+4)}{(n+4)(n+2)^2}$

   $V_\theta\left(\widehat{\theta}_{1,n}\right) = \dfrac{4n\theta^2}{(n+4)(n+2)^2}$

3. $E_\theta(X_1) = \int\limits_0^\theta x\dfrac{1}{2\sqrt{x\theta}}dx = \int\limits_0^\theta \dfrac{\sqrt{x}}{2\sqrt{\theta}}dx = \dfrac{1}{2\sqrt{\theta}}\int\limits_0^\theta \sqrt{x}dx = \dfrac{1}{2\sqrt{\theta}}\dfrac{\theta^{\frac{3}{2}}}{\frac{3}{2}} = \dfrac{\theta}{3}$.

   On a $h(\theta) = E_\theta(X) \Longleftrightarrow \theta = h^{-1}(E_\theta(X)) \Longleftrightarrow \theta = 3E_\theta(X)$

   Donc $\widehat{\theta}_{2,n} = 3\overline{X}_n$ *est un estimateur de* θ *par la méthode des moments.*

   $E_\theta\left(\widehat{\theta}_{2,n}\right) = 3E_\theta(X) = \theta$.

4. $E_\theta\left(\widehat{\theta}_{1,n} - \theta\right)^2 = E_\theta\left(\widehat{\theta}_{1,n} - E_\theta\left(\widehat{\theta}_{1,n}\right)\right)^2 + \left(E_\theta\left(\widehat{\theta}_{1,n}\right) - \theta\right)^2$ *(König)*

   $E_\theta\left(\widehat{\theta}_{1,n} - \theta\right)^2 = Var\left(\widehat{\theta}_{1,n}\right) + \left(\dfrac{n\theta}{n+2} - \theta\right)^2$

$$E_\theta\left(\widehat{\theta}_{1,n}-\theta\right)^2=\frac{4n\theta^2}{(n+4)(n+2)^2}+\frac{4\theta^2}{(n+2)^2}$$

$$E_\theta\left(\widehat{\theta}_{1,n}-\theta\right)^2=\frac{2n+4}{n+4}\frac{4\theta^2}{(n+2)^2}.$$

$$E_\theta\left(\widehat{\theta}_{2,n}-\theta\right)^2=Var\left(\widehat{\theta}_{2,n}\right)=Var\left(3\overline{X}_n\right)=\frac{9}{n}Var\left(X\right).$$

$$E_\theta\left(X^2\right)=\int_0^\theta x^2\frac{1}{2\sqrt{x\theta}}dx=\int_0^\theta\frac{\sqrt{x}}{2\sqrt{\theta}}dx=\frac{1}{2\sqrt{\theta}}\int_0^\theta x\sqrt{x}dx=\frac{1}{2\sqrt{\theta}}\frac{\theta^{\frac{5}{2}}}{\frac{5}{2}}=\frac{\theta^2}{5}.$$

$$Var\left(X\right)=\frac{\theta^2}{5}-\frac{\theta^2}{9}=\frac{4}{45}\theta^2.$$

D'où $E_\theta\left(\widehat{\theta}_{2,n}-\theta\right)^2=\dfrac{4}{5n}\theta^2.$

$\widehat{\theta}_{1,n}$ *est donc préférable à* $\widehat{\theta}_{2,n}$ *selon l'erreur quadratique moyenne.*

Corrigé Exercice 2 $f_\theta(x)=\theta^2\exp(-\theta x)\mathbf{1}_{(x\geq 0)}+\theta(1-\theta)\exp(\theta x)\mathbf{1}_{(x\leq 0)}$

1. $I(\theta)=-E_\theta\left(\dfrac{\partial^2}{\partial\theta^2}\ln f_\theta(X)\right)$

   $$\ln f_\theta(x)=\left[2\ln\theta-\theta x\right]\mathbf{1}_{(x\geq 0)}+\left[\ln\theta+\ln(1-\theta)+\theta x\right]\mathbf{1}_{(x\leq 0)}$$

   $$\frac{\partial}{\partial\theta}\ln f_\theta(x)=\left(\frac{2}{\theta}-x\right)\mathbf{1}_{(x\geq 0)}+\left(\frac{1}{\theta}-\frac{1}{1-\theta}+x\right)\mathbf{1}_{(x\leq 0)}$$

   $$\frac{\partial^2}{\partial\theta^2}\ln f_\theta(x)=-\frac{2}{\theta^2}\mathbf{1}_{(x\geq 0)}+\left(-\frac{1}{\theta^2}-\frac{1}{(1-\theta)^2}\right)\mathbf{1}_{(x\leq 0)}$$

   $$E_\theta\left(\frac{\partial^2}{\partial\theta^2}\ln f_\theta(X)\right)=\int_{\mathbb{R}^+}-2\exp(-\theta x)\ dx-\int_{\mathbb{R}^-}\left(\frac{1-\theta}{\theta}+\frac{\theta}{1-\theta}\right)\exp(\theta x)\ dx$$

   $$E_\theta\left(\frac{\partial^2}{\partial\theta^2}\ln f_\theta(X)\right)=-\frac{2}{\theta}-\left(\frac{1-\theta}{\theta}+\frac{\theta}{1-\theta}\right)\left[\frac{1}{\theta}\exp(\theta x)\right]_{-\infty}^0$$

   $$E_\theta\left(\frac{\partial^2}{\partial\theta^2}\ln f_\theta(X)\right)=-\frac{2}{\theta}-\frac{1}{\theta}\left(\frac{1-\theta}{\theta}+\frac{\theta}{1-\theta}\right)=-\frac{1}{\theta^2(1-\theta)}$$

   Finalement, $I(\theta)=\dfrac{1}{\theta^2(1-\theta)}$

2. *Soit* $Y=|X|$.

   (a) *pour* $x\geq 0$, $F_Y(x)=P\left[-x\leq X\leq x\right]\mathbf{1}_{(x\geq 0)}=F_X(x)-F_X(-x)$.
   $f_Y(x)=f_X(x)+f_X(-x)=\left(\theta^2\exp(-\theta x)+\theta(1-\theta)\exp(-\theta x)\right)\mathbf{1}_{(x\geq 0)}$
   $f_Y(x)=\theta\exp(-\theta x)\mathbf{1}_{(x\geq 0)}.$

   (b) $I'=-E_\theta\left(\dfrac{\partial^2}{\partial\theta^2}\ln f_Y(X)\right)=-E_\theta\left(\dfrac{\partial^2}{\partial\theta^2}(\ln\theta-\theta x)\right)$

   $I'=-E_\theta\left(-\dfrac{1}{\theta^2}\right)=\dfrac{1}{\theta^2}.$

3. $I'=\dfrac{1}{\theta^2}\leq\dfrac{1}{\theta^2(1-\theta)}$

Corrigé Exercice 3 $F(x,\theta) = \left(1+x^{-2}\right)^{-\theta} \mathbb{1}_{\mathbb{R}_+^*}(x)$

1. $f(x,\theta) = \frac{\partial}{\partial x} F(x,\theta) = -\theta \left(1+x^{-2}\right)^{-\theta-1} (-2)\, x^{-3} \mathbb{1}_{\mathbb{R}_+^*}(x)$
   $f(x,\theta) = \dfrac{2\theta}{x^3} \dfrac{1}{\left(1+x^{-2}\right)^{\theta+1}} \mathbb{1}_{\mathbb{R}_+^*}(x)$

2. $f(x,\theta) = \exp\left[\ln 2 + \ln\theta - 3\ln x - (\theta+1)\ln\left(1+x^{-2}\right)\right] \mathbb{1}_{\mathbb{R}_+^*}(x)$
   $$\mathcal{L}(\underline{x},\theta) = \exp\left[n\ln 2 + n\ln\theta - 3\sum_{i=1}^{n} \ln x_i - (\theta+1)\sum_{i=1}^{n} \ln\left(1+x_i^{-2}\right)\right] \mathbb{1}_{\left(\mathbb{R}_+^*\right)^n}(\underline{x})$$
   $$\mathcal{L}(\underline{x},\theta) = \exp\left[-\theta\sum_{i=1}^{n} \ln\left(1+x^{-2}\right) + n\ln\theta - 3\sum_{i=1}^{n} \ln x_i - \sum_{i=1}^{n} \ln\left(1+x_i^{-2}\right) + n\ln 2\right] \mathbb{1}_{\left(\mathbb{R}_+^*\right)^n}(\underline{x}).$$

3. *Le modèle appartient donc à la famille exponentielle,*
   $T(\underline{X}) = -\sum_{i=1}^{n} \ln\left(1+x^{-2}\right)$, *la statistique exhaustive complète,* $c(\theta) = \theta$, $d(\theta) = n\ln\theta$, $S(\underline{X}) = -3\sum_{i=1}^{n} \ln x_i - \sum_{i=1}^{n} \ln\left(1+x_i^{-2}\right)$ *et* $A = \left(\mathbb{R}_+^*\right)^n$ *indépendant de* θ.

4. *Il s'agit d'une famille exponentielle à un paramètre, sous forme canonique avec* $\Theta = \mathbb{R}_+^*$, *ouvert. Donc*
   $$E_\theta\left[T(\underline{X})\right] = -d'(\theta)$$
   $E_\theta\left[T(\underline{X})\right] = -d'(\theta) = -(n\ln\theta)' = -\dfrac{n}{\theta}$.

5. *Nous avons une famille exponentielle où* c *est injective et* c *et* d *sont de classe* $\mathcal{C}^2$, *de plus,* Θ *est un ouvert. Par conséquent, si l'équation de vraisemblance admet une solution* $\widehat{\theta}(\underline{x})$ *alors celle-ci est l'unique estimation du maximum de vraisemblance.*
   $$E_\theta\left[T(\underline{X})\right] = T(\underline{x}) \Longleftrightarrow -\frac{n}{\theta} = -\sum_{i=1}^{n} \ln\left(1+x_i^{-2}\right) \Longleftrightarrow \theta = \frac{n}{\sum_{i=1}^{n} \ln\left(1+x_i^{-2}\right)}$$
   Donc $\widehat{\theta} = \dfrac{n}{\sum_{i=1}^{n} \ln\left(1+X_i^{-2}\right)}$ *est l'estimateur du maximum de vraisemblance.*