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

Année Universitaire
2012-2013
1ère Année

Méthodes d'estimation : Examen Final

mai 2013

Enseignants : Mme Mallek et M. Rammeh — Durée : 1 heure 30mn

Exercice 1 *Soit* $(X_1, ..., X_n)$ *un n-échantillon de loi* $\mathcal{N}(\theta, \theta)$, $\theta > 0$.

1. *Ecrire la fonction de vraisemblance associée à l'échantillon.*
2. *Déterminer* $\widehat{\theta}_n^{M.M}$, *l'estimateur de* θ *par la méthode des moments déterminé à partir de* $E_\theta\left(X^2\right)$.
3. *Donner sa distribution asymptotique (c-à-d: Etudier le comportement asymptotique de l'estimateur).*
   (Indication : $Var\left(X^2\right) = \theta^4 + 6\theta^3 + 4\theta^2 + \theta$*).*
4. *Vérifier que le modèle appartient à la famille exponentielle.*
5. *Proposer une statistique exhaustive pour* θ.
6. *Démontrer que* $\widehat{\theta_n}^{MV}$, *l'estimateur du maximum de vraisemblance de* θ, *est identique à* $\widehat{\theta_n}^{M.M}$.
7. *Calculer l'information de Fisher du modèle.*

Exercice 2 *On dispose d'un n-échantillon issu d'une loi uniforme* $\mathcal{U}_{[0,e^\theta]}$. *On cherche à estimer* $\theta \in \mathbb{R}$.

1. *Proposer une statistique exhaustive pour le modèle. (Pour toute la suite, on supposera cette statistique complète).*
2. *Prouver que l'estimateur du maximum de vraisemblance de* θ *a pour expression* $\widehat{\theta_n}^{MV} = \ln \max X_i$.
3. *Déterminer sa densité et vérifier qu'il est biaisé.*
4. *Construire à partir de* $\widehat{\theta_n}^{MV}$ *un nouvel estimateur sans biais* T_n *dont on déterminera la variance.*
5. *On se propose maintenant de construire un intervalle de confiance pour* θ. *Montrer que* $Z = n(\theta - \ln \max X_i)$ *est une statistique pivotale (on utilisera la fonction de répartition de* $\widehat{\theta_n}^{MV}$*). En déduire un intervalle de confiance pour* θ *de niveau de confiance* $1 - \alpha$.
6. **Facultatif***: Verifier que* T_n *est une statistique exhaustive pour* θ.
7. **Facultatif***: Montrer que* T_n *est un estimateur sans biais de variance minimale. Est-il efficace? Justifier.*

Correction Exercice 1 *:*

1. $\mathcal{L}\left(\underline{x},\theta\right)=\exp\left[-\frac{1}{2\theta}\sum_{i=1}^{n}\left(x_i-\theta\right)^2-\frac{n}{2}\ln 2\pi\theta\right] \qquad \theta>0.$

2. $E_\theta\left(X^2\right)=\theta+\theta^2 \Longleftrightarrow \theta^2+\theta-\mu_2=0.$

   $\Delta=1+4\mu_2 \qquad \theta_1=\frac{-1-\sqrt{1+4\mu_2}}{2}<0$ *rejetée*

   $\theta_2=\frac{-1+\sqrt{1+4\mu_2}}{2}=-\frac{1}{2}+\sqrt{\frac{1}{4}+\mu_2}.$

   $\overbrace{E_\theta\left(X^2\right)}^{E_\theta[g(X)]}=\overbrace{\theta+\theta^2}^{h(\theta)} \Longleftrightarrow \theta=h^{-1}\left(E_\theta\left(X^2\right)\right).$

   Donc $\widehat{\theta}^M=h^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}X_i^2\right)=-\frac{1}{2}+\sqrt{\frac{1}{4}+\frac{1}{n}\sum_{i=1}^{n}X_i^2}.$

3. $h:\theta\longmapsto\theta^2+\theta,$ *h est de classe* C^1 *et* $\frac{\partial h}{\partial\theta}=2\theta+1>0.$

   $Var\left[g\left(X\right)\right]=Var\left(X^2\right)=\theta^4+6\theta^3+4\theta^2+\theta.$

   Donc $\sqrt{n}\left(\widehat{\theta}^M-\theta\right)\xrightarrow{Loi}\mathcal{N}\left(0,\left(\frac{1}{h'(\theta)}\right)^2Var(g(X))\right)$

   Finalement, $\sqrt{n}\left(\widehat{\theta}^M-\theta\right)\xrightarrow{Loi}\mathcal{N}\left(0,\frac{\theta^4+6\theta^3+4\theta^2+\theta}{\left(2\theta+1\right)^2}\right)$

4. $\mathcal{L}\left(\underline{x},\theta\right)=\exp\left[-\frac{1}{2\theta}\sum_{i=1}^{n}x_i^2-\frac{n}{2}\left(\theta+\ln\theta\right)+\sum_{i=1}^{n}x_i-\frac{n}{2}\ln 2\pi\right]$

   Le modèle appartient à la famille exponentielle à un paramètre avec $c\left(\theta\right)=-\frac{1}{2\theta},$

   $d\left(\theta\right)=-\frac{n}{2}\left(\theta+\ln\theta\right),T\left(\underline{x}\right)=\sum_{i=1}^{n}x_i^2$ *et* $S\left(\underline{x}\right)=\sum_{i=1}^{n}x_i-\frac{n}{2}\ln 2\pi.$ $A^n=\mathbb{R}^{\ltimes}$ *indépendant de* θ.

5. $T\left(\underline{X}\right)=\sum_{i=1}^{n}X_i^2$ *est la statistique exhaustive complète associée à la famille exponentielle.*$\mathcal{L}$

6. $\mathcal{L}\left(\underline{x},\theta\right)=\exp\left[-\frac{1}{2\theta}\sum_{i=1}^{n}\left(x_i-\theta\right)^2-\frac{n}{2}\ln 2\pi\theta\right]$

   $\mathcal{L}\left(\underline{x},\theta\right)=\exp\left[-\frac{1}{2\theta}\sum_{i=1}^{n}x_i^2+\sum_{i=1}^{n}x_i-\frac{n\theta}{2}-\frac{n}{2}\ln\theta-\frac{n}{2}\ln 2\pi\right] \qquad \lambda>0.$

   Il s'agit d'une famille exponentielle où $\Theta=\mathbb{R}_+^*$ *est un ouvert. De plus,* $c:\theta\longmapsto-\frac{1}{2\theta}$ *est injective et de classe* C^2 *et*

   $d:\theta\longmapsto-\frac{n}{2}\left(\theta+\ln\theta\right)$ *est de classe* C^2*. Aussi, l'emv, s'il existe, est solution de l'équation* $E_\theta\left(T\left(\underline{X}\right)\right)=T\left(\underline{x}\right).$

   $E_\theta\left(T\left(\underline{X}\right)\right)=E_\theta\left(\sum_{i=1}^{n}X_i^2\right)=nE_\theta\left(X^2\right)=n\left(\theta+\theta^2\right)=n\theta\left(1+\theta\right).$

   $E_\theta\left(T\left(\underline{X}\right)\right)=T\left(\underline{x}\right)\Longleftrightarrow\sum_{i=1}^{n}x_i^2=n\theta\left(1+\theta\right).$

$n\theta^2+n\theta-\sum\limits_{i=1}^{n}x_i^2=0$

$\Delta=n^2+4n\sum\limits_{i=1}^{n}x_i^2=\left[2n\sqrt{\frac{1}{4}+\frac{1}{n}\sum\limits_{i=1}^{n}x_i^2}\right]^2>0$

$\theta_1=\dfrac{-n-2n\sqrt{\frac{1}{4}+\frac{1}{n}\sum\limits_{i=1}^{n}x_i^2}}{2n}<0 \qquad \textit{rejetée}$

$\theta_2=\dfrac{-n+2n\sqrt{\frac{1}{4}+\frac{1}{n}\sum\limits_{i=1}^{n}x_i^2}}{2n}=-\frac{1}{2}+\sqrt{\frac{1}{4}+\frac{1}{n}\sum\limits_{i=1}^{n}x_i^2}$

Donc $\widehat{\theta}^{MV}=\widehat{\theta}^{M}$

7. *Comme le modèle appartient à la famille exponentielle, les 3 hypothèses sont valides.*

   $I_n(\theta)=-E\left[\frac{\partial^2}{\partial\theta^2}\ln\mathcal{L}(\underline{x},\theta)\right]$

   $\frac{\partial}{\partial\theta}\ln\mathcal{L}(\underline{x},\theta)=\frac{1}{2\theta^2}\sum\limits_{i=1}^{n}x_i^2-\frac{n}{2}\left(1+\frac{1}{\theta}\right)$

   $\frac{\partial^2}{\partial\theta^2}\ln\mathcal{L}(\underline{x},\theta)=-\frac{1}{\theta^3}\sum\limits_{i=1}^{n}x_i^2+\frac{n}{2\theta^2}$

   $I_n(\theta)=-E\left[-\frac{1}{\theta^3}\sum\limits_{i=1}^{n}X_i^2+\frac{n}{2\theta^2}\right]=E\left[\frac{1}{\theta^3}\sum\limits_{i=1}^{n}X_i^2-\frac{n}{2\theta^2}\right]=\frac{n}{\theta^3}E\left[X^2\right]-\frac{n}{2\theta^2}$

   $I_n(\theta)=\frac{n}{\theta^3}\left(\theta+\theta^2\right)-\frac{n}{2\theta^2}=\frac{n}{2\theta^2}+\frac{n}{\theta}=\frac{n(1+2\theta)}{2\theta^2}.$

   $BCR\left(\theta^2\right)=\dfrac{\left(\psi'(\theta)\right)^2}{I_n(\theta)}=\dfrac{\left(\frac{1}{2\sqrt{\theta}}\right)^2}{n(1+2\theta)}(2\theta)^2=\dfrac{\theta}{n(1+2\theta)}.$

Correction Exercice 2 $X\rightsquigarrow\mathcal{U}_{[0,e^\theta]} \qquad f_X(x,\theta)=\frac{1}{e^\theta}\mathbb{1}_{[0,e^\theta]}(x)=e^{-\theta}\mathbb{1}_{[0,e^\theta]}(x).$

1. $\mathcal{L}(\underline{x},\theta)=e^{-n\theta}\mathbb{1}_{\left\{\max x_i<e^\theta\right\}}\mathbb{1}_{\{\min x_i>0\}}=e^{-n\theta}\mathbb{1}_{\{\ln\max x_i<\theta\}}\mathbb{1}_{\{\min x_i>0\}}.$

   D'après le théorème de factorisation, $T(\underline{X})=MaxX_i$ *est exhaustive.*

2. $\mathcal{L}(\underline{x},\theta)$ *est une fonction strictement décroissante en* θ*. Ainsi, sur l'intervalle* $[\ln\max x_i,+\infty[$, *elle atteint son maximum en* $\widehat{\theta_n}^{MV}=\ln\max X_i$*.*

3. $F_{\widehat{\theta_n}^{MV}}(x)=P_\theta\left[\widehat{\theta_n}^{MV}<x\right]=P_\theta\left[\max X_i<\exp x\right]=\left[F_X(\exp x)\right]^n$

   $F_X(x)=xe^{-\theta}\mathbb{1}_{[0,e^\theta]}(x)+\mathbb{1}_{]e^\theta+\infty[}(x)$

   $F_{\widehat{\theta_n}^{MV}}(x)=\left[\exp xe^{-\theta}\right]^n\mathbb{1}_{[0,e^\theta]}(\exp x)+\mathbb{1}_{]e^\theta+\infty[}(\exp x)$

   $F_{\widehat{\theta_n}^{MV}}(x)=\exp n(x-\theta)\,\mathbb{1}_{]-\infty,\theta]}(x)+\mathbb{1}_{]\theta+\infty[}(x)$

   $f_{\widehat{\theta_n}^{MV}}(x)=n\exp n(x-\theta)\,\mathbb{1}_{]-\infty,\theta]}(x).$

$E_\theta\left[\widehat{\theta_n}^{MV}\right] = \int_{-\infty}^{\theta} nx \exp n\left(x-\theta\right)dx = e^{-n\theta}\left[x\ e^{nx}\right]_{-\infty}^{\theta} - e^{-n\theta}\int_{-\infty}^{\theta}\exp nx\ dx$

$E_\theta\left[\widehat{\theta_n}^{MV}\right] = e^{-n\theta}\left(\theta e^{n\theta}\right) - e^{-n\theta}\frac{1}{n}e^{n\theta} = \theta - \frac{1}{n} \neq \theta$

$\widehat{\theta_n}^{MV}$ *est donc biaisé.*

4. *Soit* $T_n = \widehat{\theta_n}^{MV} + \frac{1}{n} = \ln\max X_i + \frac{1}{n} \qquad E_\theta\left[T_n\right] = \theta$

   $Var\left[T_n\right] = Var\left[\widehat{\theta_n}^{MV}\right] = \int_{-\infty}^{\theta} nx^2\exp n\left(x-\theta\right)dx - \left(\theta - \frac{1}{n}\right)^2$

   $Var\left[T_n\right] = e^{-n\theta}\left[x^2e^{nx}\right]_{-\infty}^{\theta} - e^{-n\theta}\int_{-\infty}^{\theta} 2x\exp nx\ dx - \left(\theta-\frac{1}{n}\right)^2$

   $Var\left[T_n\right] = e^{-n\theta}\left(\theta^2e^{n\theta}\right) - 2e^{-n\theta}\left(\left[\frac{1}{n}x\ e^{nx}\right]_{-\infty}^{\theta} - \int_{-\infty}^{\theta}\frac{1}{n}\exp nx\ dx\right) - \left(\theta-\frac{1}{n}\right)^2$

   $Var\left[T_n\right] = \theta^2 - 2\left(\frac{1}{n}\theta - \frac{1}{n^2}\right) - \left(\theta - \frac{1}{n}\right)^2 = \frac{1}{n^2}$

5. $Z = n(\theta - \ln\max X_i)$

   $F_Z\left(z\right) = P_\theta\left[n(\theta-\ln\max X_i) < z\right] = P_\theta\left[\ln\max X_i > \theta - \frac{1}{n}z\right]$

   $F_Z\left(z\right) = 1 - F_{\widehat{\theta_n}^{MV}}\left(\theta - \frac{1}{n}z\right)$

   $F_Z\left(z\right) = 1 - \exp\left(-z\right)\mathbb{1}_{\{z>0\}} \qquad Z \rightsquigarrow \mathcal{E}\left(1\right).$

   La loi de Z *est indépendante de* θ *alors que cette statistique s'exprime en fonction de* θ*; elle est donc pivotale.*

   La loi exponentielle étant unimodale, on suppose $0.05 < 2\min\left(F_Z\left(x^*\right), 1 - F_Z\left(x^*\right)\right).$ *Ainsi,* $\beta = \frac{\alpha}{2}$ *correspond à l'intervalle de dispersion optimal.*

   $P\left[F_Z^{-1}\left(0.025\right) < n(\theta - \ln\max X_i) < F_Z^{-1}\left(0.975\right)\right] = 0.95$

   $P\left[\ln\max X_i + \frac{1}{n}F_Z^{-1}\left(0.025\right) < \theta < \ln\max X_i + \frac{1}{n}F_Z^{-1}\left(0.975\right)\right] = 0.95$

   $IC_{0.95}\left(\theta\right) = \left[\ln\max X_i + \frac{1}{n}F_Z^{-1}\left(0.025\right)\ ;\ \ln\max X_i + \frac{1}{n}F_Z^{-1}\left(0.975\right)\right]$

   $F_Z\left(z\right) = 1 - \exp\left(-z\right) \qquad F_Z^{-1}\left(u\right) = -\ln\left(1-u\right)$

   $IC_{0.95}\left(\theta\right) = \left[\ln\max X_i - \frac{1}{n}\ln\left(0.975\right)\ ;\ \ln\max X_i - \frac{1}{n}\ln\left(0.025\right)\right]$

6. $T_n = \ln\max X_i + \frac{1}{n}$

   Il s'agit d'une transformée strictement monotone d'une statistique exhaustive, elle est donc exhaustive.

7. $T_n = \ln\max X_i + \frac{1}{n}$; T_n *est sans biais de* θ *avec* θ *finie.* $T\left(\underline{X}\right) = \max X_i$ *est exhaustive complète. D'après Lehman Scheffe,* $E_\theta\left[T_n \backslash T\left(\underline{X}\right)\right]$ *est uvmb. Or* $\widehat{\theta}_2 = \ln\max X_i + \frac{1}{n} = \ln T\left(\underline{X}\right) + \frac{1}{n}$.

   Donc $T_n = E_\theta\left[T_n\backslash T\left(\underline{X}\right)\right]$ *est uvmb de* θ.

   Cet estimateur ne peut être efficace car il n'existe pas d'estimateur efficace, l'emv étant biaisé.