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

Année Universitaire
2019-2020
1ère Année

Examen Final : Statistique inférentielle 1

Mars 2020

Aucun document autorisé. Toutes les réponses doivent être justifiées

(2 pages)

Cours de Mme Héla Ouaili-Mallek

Durée : 1h 30mn

Exercice 1 *On donne la densité de la loi* $\beta eta\ (\alpha, \theta)$, *où* α *et* θ *sont deux réels inconnus et strictement positifs.*

$$f(x\,,\,\alpha\,\,,\,\,\theta) = \frac{\Gamma(\alpha+\theta)}{\Gamma(\alpha)\,\Gamma(\theta)} x^{\alpha-1}(1-x)^{\theta-1}\,\mathbb{1}_{]0,1[}(x)$$

Partie 1: *le couple* (α, θ) *est inconnu*
On dispose d'un n-échantillon issu de ce modèle

1. *Ecrire la vraisemblance du n-échantillon.*
2. *Vérifier si le modèle appartient à la famille exponentielle.*
3. *Proposer une statistique* $T_1(\underline{X})$ *exhaustive pour le modèle et vérifier si elle est complète.*
4. *Donner l'expression de* $E_\theta[T_1(\underline{X})]$ *en fonction de la fonction* Γ *et de sa dérivée.*

Partie 2: *on fixe* $\alpha = 1$
On dispose d'un n-échantillon issu d'une variable aléatoire X *de loi* $\beta eta\ (1, \theta)$.

5. *Ecrire la vraisemblance du n-échantillon.*
6. *Calculer* $E_\theta[X]$. *En déduire un estimateur par la méthode des moments,* $\widehat{\theta}_1$.
7. *Vérifier si le nouveau modèle appartient à la famille exponentielle.*
8. *Proposer une statistique exhaustive,* $T_2(\underline{X})$. *Calculer son espérance et sa variance.*
9. *Déterminer* $\widehat{\theta}_2$, *l'estimateur du maximum de vraisemblance.*
10. *Calculer* $I_n(\theta)$, *l'information de Fisher du n-échantillon.*

Partie 3:

On considère la variable aléatoire $Y = -\ln(1-X)$

11. *Vérifier que* $Y \rightsquigarrow \mathcal{E}(\theta)$

12. *En déduire* $E_\theta\left[\widehat{\theta}_2\right]$ *et vérifier que* $\widehat{\theta}_2$ *n'est pas efficace pour* θ.

13. *Proposer, alors, un estimateur* T^*, *uniformément de variance minimale pour* θ.

$BCR = \frac{\theta^2}{n}$ $\quad$ $var(\hat{\theta}_2) = n\theta^2$

Corrigé de l'exercice 1 :

Partie 1:

1. $f(x,\alpha,\theta)=\frac{\Gamma(\alpha+\theta)}{\Gamma(\alpha)\Gamma(\theta)}x^{\alpha-1}(1-x)^{\theta-1}\,\mathbb{1}_{]0,1[}(x)$

   $\mathcal{L}(\underline{x},\alpha,\theta)=\left(\frac{\Gamma(\alpha+\theta)}{\Gamma(\alpha)\Gamma(\theta)}\right)^n\prod_{i=1}^n x_i^{\alpha-1}\prod_{i=1}^n(1-x_i)^{\theta-1}\,\mathbb{1}_{]0,1[^n}(\underline{x})$

2. $$l(\underline{x},\alpha,\theta)=\exp\begin{pmatrix}(\alpha-1)\sum\limits_{i=1}^n\ln x_i+(\theta-1)\sum\limits_{i=1}^n\ln(1-x_i)\\ +n\ln(\Gamma(\alpha+\theta))-n\ln(\Gamma(\alpha))-n\ln(\Gamma(\theta))\end{pmatrix}\mathbb{1}_{]0,1[^n}(\underline{x})$$

   $$l(\underline{x},\alpha,\theta)=\exp\begin{pmatrix}\alpha\sum\limits_{i=1}^n\ln x_i+\theta\sum\limits_{i=1}^n\ln(1-x_i)+n\ln(\Gamma(\alpha+\theta))\\ -n\ln(\Gamma(\alpha))-n\ln(\Gamma(\theta))-\sum\limits_{i=1}^n\ln x_i-\sum\limits_{i=1}^n\ln(1-x_i)\end{pmatrix}\mathbb{1}_{]0,1[^n}(\underline{x})$$

   Le modèle appartient à la famille exponentielle à 2 paramètres avec

   $c_1(\alpha,\theta)=\alpha \qquad c_2(\alpha,\theta)=\theta$

   $T_1(\underline{x})=\sum\limits_{i=1}^n\ln x_i \qquad T_2(\underline{x})=\sum\limits_{i=1}^n\ln(1-x_i)$

   $d(\alpha,\theta)=n\ln(\Gamma(\alpha+\theta))-n\ln(\Gamma(\alpha))-n\ln(\Gamma(\theta))$

   $S(\underline{x})=-\sum\limits_{i=1}^n\ln x_i-\sum\limits_{i=1}^n\ln(1-x_i) \qquad A=]0,1[^n$ *indépendant de* θ.

3. *On a* $T(\underline{X})=\begin{pmatrix}\sum\limits_{i=1}^n\ln X_i\\ \sum\limits_{i=1}^n\ln(1-X_i)\end{pmatrix}$ *est une statistique exhaustive. Elle est nécessairement complète puisque le modèle appartient à la famille exponentielle.*

4. $$l(\underline{x},\alpha,\theta)=\exp\begin{pmatrix}\alpha\sum\limits_{i=1}^n\ln x_i+\theta\sum\limits_{i=1}^n\ln(1-x_i)+n\ln(\Gamma(\alpha+\theta))\\ -n\ln(\Gamma(\alpha))-n\ln(\Gamma(\theta))-\sum\limits_{i=1}^n\ln x_i-\sum\limits_{i=1}^n\ln(1-x_i)\end{pmatrix}\mathbb{1}_{]0,1[^n}(\underline{x})$$

   La famille exponentielle est sous forme canonique. De plus $\Theta=\{(\alpha,\theta)\}=\mathbb{R}_+^*\times\mathbb{R}_+^*$ *est un ouvert.*

   $E_\theta[T(\underline{X})]=-\begin{pmatrix}\frac{\partial}{\partial\alpha}d(\alpha,\theta)\\ \frac{\partial}{\partial\theta}d(\alpha,\theta)\end{pmatrix}$

   $\frac{\partial}{\partial\alpha}d(\alpha,\theta)=n\frac{\frac{\partial}{\partial\alpha}\Gamma(\alpha+\theta)}{\Gamma(\alpha+\theta)}-n\frac{\frac{\partial}{\partial\alpha}\Gamma(\alpha)}{\Gamma(\alpha)}$

   $\frac{\partial}{\partial\theta}d(\alpha,\theta)=n\frac{\frac{\partial}{\partial\theta}\Gamma(\alpha+\theta)}{\Gamma(\alpha+\theta)}-n\frac{\frac{\partial}{\partial\theta}\Gamma(\theta)}{\Gamma(\theta)}$

Partie 2:

5. $f(x,\theta)=\frac{\Gamma(1+\theta)}{\Gamma(\theta)}(1-x)^{\theta-1}\,\mathbb{1}_{]0,1[}(x)=\theta(1-x)^{\theta-1}\,\mathbb{1}_{]0,1[}(x)$

   $\mathcal{L}(\underline{x},\theta)=\theta^n\prod_{i=1}^n(1-x_i)^{\theta-1}\,\mathbb{1}_{]0,1[^n}(\underline{x})$

6. $E_\theta[X] = \int_0^1 x\theta(1-x)^{\theta-1}\,dx$

$u = x \qquad u' = 1 \qquad v' = \theta(1-x)^{\theta-1} \qquad v = -(1-x)^\theta$

$E_\theta[X] = \left[-x(1-x)^\theta\right]_0^1 + \int_0^1 (1-x)^\theta\,dx = \left[\frac{-1}{\theta+1}(1-x)^{\theta+1}\right]_0^1 = \frac{1}{\theta+1}$

$q(\theta) = E_\theta[g(X)] \qquad q(\theta) = \frac{1}{\theta+1} \qquad g(x) = x$

$\theta = \frac{1}{E_\theta[X]} - 1 \Longrightarrow \hat{\theta}_1 = \frac{1}{\overline{X}_n} - 1$ *est un estimateur par la méthode des moments*

7. $\mathcal{L}(\underline{x}, \theta) = \theta^n \prod_{i=1}^n (1-x_i)^{\theta-1}\, \mathbb{1}_{]0,1[^n}(\underline{x})$

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

Le modèle appartient à la famille exponentielle à un paramètre avec

$c(\theta) = \theta \qquad T(\underline{x}) = \sum_{i=1}^n \ln(1-x_i) \qquad d(\theta) = n\ln\theta \qquad S(\underline{x}) = -\sum_{i=1}^n \ln(1-x_i)$

$A =]0,1[^n$ *indépendant de* θ.

8. $T_2(\underline{X}) = \sum_{i=1}^n \ln(1-X_i)$ *est donc une statistique exhaustive.*

On a une famille exponentielle à un paramètre sous forme canonique et Θ *est un ouvert. Donc*

$E_\theta[T_2(\underline{X})] = -d'(\theta) = -\frac{n}{\theta} \qquad Var[T_2(\underline{X})] = -d''(\theta) = \frac{n}{\theta^2}$

9. *On a une famille exponentielle à un paramètre sous forme canonique où* Θ *est un ouvert, c est injective et c et d sont de classe* C^2. *Donc l'estimateur du maximum de vraisemblance, s'il existe, est solution de l'équation*

$$E_\theta[T_2(\underline{X})] = T_2(\underline{x})$$

$E_\theta[T_2(\underline{X})] = T_2(\underline{x}) \Longleftrightarrow -\frac{n}{\theta} = \sum_{i=1}^n \ln(1-x_i) \Longleftrightarrow \theta = \frac{-n}{\sum_{i=1}^n \ln(1-x_i)}$

$\Longrightarrow \hat{\theta}_2 = \frac{-n}{\sum_{i=1}^n \ln(1-X_i)}$ *est l'emv.*

10. **Le modèle appartient à la famille exponentielle. Les hypothèses H_1, H_2 et H_3 sont donc valides.**

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

$$\frac{\partial^2}{\partial\theta^2} l(x,\theta) = \frac{\partial}{\partial\theta}\left(\sum_{i=1}^{n} \ln(1-x_i) + \frac{n}{\theta}\right) = -\frac{n}{\theta^2} \Longrightarrow I_n(\theta) = \frac{n}{\theta^2}$$

Partie 3:

11. $Y = -\ln(1-X) = g(X)$

$$E_\theta[Y] = \int_0^1 g(x) f_X(x,\theta)\,dx = \int_0^1 g(x)\,\theta(1-x)^{\theta-1}\,dx$$

$$y = -\ln(1-x) \Longleftrightarrow x = 1-\exp(-y) \qquad dx = \exp(-y)\ dy$$

$$E_\theta[Y] = \int_0^{+\infty} y\theta(\exp(-y))^{\theta-1}\exp(-y)\ dy = \int_0^{+\infty} y\theta\exp(-\theta y)\ dy$$

$$f_Y(y,\theta) = \theta\exp(-\theta y)\,\mathbb{1}_{]0,+\infty[}(y)$$

$$Y \rightsquigarrow \mathcal{E}(\theta)$$

12. $Y \rightsquigarrow \mathcal{E}(\theta) \Longrightarrow \sum_{i=1}^{n} Y_i = -\sum_{i=1}^{n}\ln(1-X_i) \rightsquigarrow \gamma(n,\theta)$

$$\widehat{\theta}_2 = \frac{-n}{\sum_{i=1}^{n}\ln(1-X_i)} = \frac{n}{\sum_{i=1}^{n} Y_i} \Longrightarrow E_\theta\left[\widehat{\theta}_2\right] = \int_0^{+\infty} \frac{n}{t}\frac{\theta^n}{\Gamma(n)} t^{n-1}\exp(-\theta t)\,dt$$

$$E_\theta\left[\widehat{\theta}_2\right] = \int_0^{+\infty} n\frac{\theta^n}{\Gamma(n)} t^{(n-1)-1}\exp(-\theta t)\,dt = \frac{n}{n-1}\theta\int_0^{+\infty}\frac{\theta^{n-1}}{\Gamma(n-1)} t^{(n-1)-1}\exp(-\theta t)\,dt$$

$$E_\theta\left[\widehat{\theta}_2\right] = \frac{n}{n-1}\theta$$

13. $S^* = \frac{n-1}{n}\widehat{\theta}_2 \Longrightarrow E_\theta[S^*] = \theta < +\infty$

De plus, $S^* = -\frac{n-1}{\sum_{i=1}^{n}\ln(1-X_i)}$ *est fonction d'une statistique exhaustive complète,* S^* *est dnc uvmb.*