XVI CT 2019-2020

M1 Actuariat & Econométrie et Statistiques, année 2019–2020.

Statistique Inférentielle
Contrôle terminal

Vendredi 10 janvier

Durée 2h, documents, téléphone, calculatrice interdits

Rappel 1 : densité d'une loi $\Gamma(\alpha,\beta)$

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \mathbb{1}_{\mathbb{R}^+}(x).$$

Rappel 2 : densité d'une loi $Beta(\alpha,\beta)$ vaut

$$\begin{aligned} f(x) &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \mathbb{1}[0,1](x) \\ &= \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} \mathbb{1}[0,1](x) \end{aligned}$$

et

$$\Gamma(x) = (x-1)\Gamma(x)$$

Rappel 3 : Méthode Delta

On considère T_n un vecteur aléatoire de $\mathbb{R}^k$, Σ une matrice de covariance. On suppose que

$$a_n(T_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0,\Sigma)$$

avec $a_n \to \infty$. Alors, pour toute fonction g de classe C^1, $g(T_n)$ converge en probabilité vers $g(\theta)$ et

$$a_n(g(T_n) - g(\theta)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, D_g \Sigma D_g^t)$$

où D_g est la matrice Jacobienne de g calculée en θ.

Exercice 1 :

On considère un échantillon $X_1, ..., X_n$ de v.a.i.i.d. de loi $Beta(1,\theta)$ i.e. de densité

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

1. Donner $\hat{\theta}_M$ l'estimateur de θ par la méthode des moments.
2. Donner l'estimateur du maximum de vraisemblance $\hat{\theta}_n$ de θ.
3. Cet estimateur est-il exhaustif ?
4. calculer la loi de la v.a. $Y = -\log(1-X)$.
5. L'estimateur $\hat{\theta}_n$ est-il sans biais ? Donner un estimateur sans biais de θ.
6. L'estimateur $\hat{\theta}_n$ est-il convergent ?
7. Calculer l'information de Fisher du modèle.
8. L'estimateur $\hat{\theta}_n$ est-il efficace ?
9. Donner la normalité asymptotique de $\hat{\theta}_n$.
10. Donner la normalité asymptotique de $\hat{\theta}_M$ (estimateur des moments).
11. Donner un intervalle de confiance de niveau asymptotique $1-\alpha$ pour θ.
12. Construire un text de niveau asymptotique α pour tester $H_0 \ : \ \theta = \theta_0$ contre $H_1 \ : \ \theta = \theta_1 > \theta_0$.
13. Construire un text de niveau asymptotique α pour tester $H_0 \ : \ \theta = \theta_0$ contre $H_1 \ : \ \theta < \theta_0$.

<u>Exercice 2 :</u>

On considère un échantillon $X_1, ..., X_n$ de v.a.i.i.d. de loi $\Gamma(\alpha, \beta)$ (cf page 1).

1. Calculer $\mathbb{E}(X_i^k)$ pour $k = 1, ..., 4$.
2. Calculer $(\hat{\alpha}, \hat{\beta})$ l'estimateur des moments de (α, β).
3. Ecrire le Théorème Centrale Limite pour $(\overline{X}_n, \overline{X^2}_n)$.
4. Donner la normalité asymptotique de $(\hat{\alpha}, \hat{\beta})$.

<u>Exercice 3 :</u>

On considère un échantillon $X_1, ..., X_n$ de v.a.i.i.d. de densité

$$f(x) = \begin{cases} \frac{1-\theta}{2} & \text{si } x \ \in]-1, 0] \\ \\ \frac{1+\theta}{2} & \text{si } x \ \in]0, 1] \\ \\ 0 & \text{sinon} \end{cases}$$

On pose

$$K = \sum_{i=1}^{n} \mathbb{1}\{X_i > 0\}$$

1. Montrer que K est une statistique exhaustive.
2. Calculer $\hat{\theta}_n$ l'estimateur du maximum de vraisemblance de θ.
3. Donner la loi de K.
4. Calculer le biais de $\hat{\theta}_n$.
5. $\hat{\theta}_n$ est-il convergent ?
6. Donner la normalité asymptotique de $\hat{\theta}_n$.
7. Déterminer un intervalle de confiance de θ de niveau asymptotique $1 - \alpha$.
8. Construire un test de niveau asymptotique α pour tester $H_0 \ : \ \theta = \theta_0$ contre $H_1 \ : \ \theta = \theta_1 > \theta_0$.

Exercice 1 : $X_1, \dots, X_n$ iid tq $X_i \sim \mathcal{B}(1,\theta)$

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

1) Méthode des moments :

Dimens° 1 $\Rightarrow$ $\mathbb{E}[X_i] = \overline{X}_n$

$$\Rightarrow \frac{1}{1+\hat{\theta}_M} = \overline{X}_n \Rightarrow \hat{\theta}_M = \frac{1}{\overline{X}_n} - 1$$

2) $$L(x_1, \dots, x_n, \theta) = \prod_{i=1}^{n} \theta(1-x_i)^{\theta-1} \mathbb{1}_{[0,1]}(x_i)$$

$$= \theta^n \left(\prod_{i=1}^{n} (1-x_i)^{\theta-1}\right) \underbrace{\prod_{i=1}^{n} \mathbb{1}_{[0,1]}(x_i)}_{\min\limits_{i\in[1,n]} x_i \geq 0 \text{ et } \max\limits_{i\in[1,n]} x_i \leq 1}$$

Indep du Support

$$\Rightarrow \ln L = n \ln\theta + \sum_{i=1}^{n} (\theta-1)\ln(1-x_i) \qquad x_i \in [0,1]$$

$$\Rightarrow \frac{\partial \ln L}{\partial\theta} = \frac{n}{\theta} + \sum_{i=1}^{n} \ln(1-x_i)$$

$\hat{\theta}_n$ EMV $\Rightarrow \frac{n}{\hat{\theta}_n} + \sum \ln(1-x_i) = 0$

$$\Rightarrow \hat{\theta}_n = \frac{-n}{\sum_{i=1}^{n} \ln(1-X_i)}$$

$$\frac{\partial^2 \ln L}{\partial\theta^2} = -\frac{n}{\theta^2} < 0$$

Donc $\hat{\theta}_n = \frac{-n}{\sum_{i=1}^{n} \ln(1-x_i)}$ est bien l'EMV de θ.

3) $$L = \theta^n \left(\prod_{i=1}^{n} (1-x_i)^{\theta-1}\right) \prod_{i=1}^{n} \mathbb{1}_{[0,1]}(x_i)$$

$$\hat{\theta}_n = \frac{-n}{\sum_{i=1}^{n} \ln(1-X_i)} \Rightarrow \sum_{1}^{n} \ln(1-X_i) = -\frac{n}{\hat{\theta}_n} \Leftrightarrow \ln\left(\prod_{1}^{n} \ln(1-X_i)\right) = -\frac{n}{\hat{\theta}_n}$$

$$\Rightarrow \prod_{1}^{n} (1-x_i)^{\theta-1} = \exp\left(-\frac{n}{\hat{\theta}_n}\right)^{\theta-1} = \exp\left(-\frac{n(\theta-1)}{\hat{\theta}_n}\right)$$

$$\Rightarrow L = \underbrace{\theta^n \exp\left(-\frac{n(\theta-1)}{\hat{\theta}_n}\right)}_{\varphi(\theta,\hat{\theta}_n)} \underbrace{\prod_{1}^{n} \mathbb{1}_{[0,1]}(x_i)}_{\psi(x_1,\dots,x_n)}$$

Donc exhaustif.

4) Soit Y la VA tq $Y = -\log(1-X)$

$$\mathbb{E}[\varphi(Y)] = \int -\log(1-x)\,\theta(1-x)^{\theta-1}\,\mathbb{1}_{[0,1]}(x)\,dx$$

Changmt de var : $u = -\log(1-x)$

$$\frac{du}{dx} = (1-x)^{-1} \Rightarrow dx = (1-x)\,du$$

$$1-x = e^{-u}$$

$$= \int \theta u\, e^{-\theta u}\,\mathbb{1}_{[0,\infty[}(u)\,du$$

$$\Rightarrow Y = -\log(1-X) \sim \mathcal{E}(\theta)$$

5) $B_\theta(\hat{\theta}_n) = \mathbb{E}[\hat{\theta}_n] - \theta$

$$\mathbb{E}[\hat{\theta}_n] = \mathbb{E}\left[\frac{-n}{\sum \ln(1-X_i)}\right] = n\,\mathbb{E}\left[\frac{1}{\sum(-\ln(1-X_i))}\right] = n\,\mathbb{E}\left[\frac{1}{\sum Y_i}\right]$$

Or $Y_i \sim \mathcal{E}(\theta) \Rightarrow \sum Y_i \sim \Gamma(n,\theta)$

$\Rightarrow \frac{1}{\sum Y_i}$ suit une loi inverse gamma de paramètre n et θ

$$\Rightarrow \mathbb{E}[\hat{\theta}_n] = n \times \frac{\theta}{n-1}$$

$$\Rightarrow B_\theta(\hat{\theta}_n) = \frac{n\theta}{n-1} - \theta = \frac{\theta}{n-1} \xrightarrow[n\to\infty]{} 0$$

Donc $\hat{\theta}_n$ asympt sans biais

$\hat{\theta}'_n = \frac{n-1}{n}\hat{\theta}_n = \frac{-(n-1)}{\sum \ln(1-X_i)}$ est un estimateur sans biais de θ.

Car $\mathbb{E}[\hat{\theta}'_n] = \frac{n-1}{n}\mathbb{E}[\hat{\theta}_n] = \frac{n-1}{n} \times \frac{n}{n-1}\theta = \theta$

$$\Rightarrow B_\theta(\hat{\theta}'_n) = 0$$

6) On a déjà $B_\theta(\hat{\theta}_n) \xrightarrow[\infty]{} 0$

$$V(\hat{\theta}_n) = V\left(\frac{n}{\sum -\ln(1-X_i)}\right) = n^2\,V\left(\frac{1}{\sum Y_i}\right) \underset{\uparrow}{=} n^2\frac{\theta^2}{(n-1)^2(n-2)} \underset{\infty}{\sim} \frac{\theta^2}{n-2} \xrightarrow[n\to\infty]{} 0$$

Variance d'une gamma inverse de paramètre n,θ

Donc estimateur convergent.

7) $I_n(\theta) = -\mathbb{E}\left[\frac{\partial^2}{\partial\theta^2} \ln L\right]$ H4 Ok

$$= -\mathbb{E}\left[-\frac{n}{\theta^2}\right] = \frac{n}{\theta^2}$$

8) $\mathbb{E}[\hat{\theta}_n] = \frac{n}{n-1}\theta \implies g : x \mapsto \frac{n}{n-1}x \implies g'(x) = \frac{n}{n-1}$

$$\frac{(g'(\theta))^2}{I_n(\theta)} = \frac{\frac{n^2}{(n-1)^2}}{\frac{n}{\theta^2}} = \frac{n^2}{(n-1)^2} \times \frac{\theta^2}{n} = \frac{n\theta^2}{(n-1)^2}$$

Or $I_n(\theta) = \frac{n}{\theta^2} \neq \frac{n\theta^2}{(n-1)^2}$

Donc $\hat{\theta}_n$ n'est pas efficace

9) $\hat{\theta}_n = \frac{n}{\sum -\ln(1-X_i)}$

Toutes les hypothèses sont vérifiées $\implies$

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, I^{-1}(\theta))$$

Donc $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \theta^2)$

10) $\hat{\theta}_M = \frac{1}{\overline{X}_n} - 1 = g(\overline{X}_n)$ avec $g : x \mapsto \frac{1}{x} - 1$, $g'(x) = -\frac{1}{x^2}$

On applique le TCL à $X_1, \dots, X_n$ i.e $\sqrt{n}\,\frac{\overline{X}_n - \mathbb{E}[X_1]}{\sqrt{V(X_1)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$

donc $\sqrt{n}\,\frac{\overline{X}_n - \frac{1}{1+\theta}}{\sqrt{\frac{\theta}{(1+\theta)^2(2+\theta)}}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$

$$\implies \sqrt{n}\left(\overline{X}_n - \frac{1}{1+\theta}\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{\theta}{(1+\theta)^2(2+\theta)}\right)$$

On applique la Méthode Delta au TCL avec $g : x \mapsto \frac{1}{x} - 1$

$g(\overline{X}_n) = \hat{\theta}_M$, $g\left(\frac{1}{1+\theta}\right) = \theta$, $g'(x) = \frac{-1}{x^2}$, $\left(g'\left(\frac{1}{1+\theta}\right)\right)^2 = (1+\theta)^4$

$$\implies \sqrt{n}\left(g(\overline{X}_n) - g\left(\frac{1}{1+\theta}\right)\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(\frac{\theta}{(1+\theta)^2(2+\theta)} \times \left(g'\left(\frac{1}{1+\theta}\right)\right)^2\right)$$

$$\implies \sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{\theta(1+\theta)^2}{2+\theta}\right)$$