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

Numéro copie (obligatoire) :

# Statistique Inférentielle
# Contrôle terminal

Vendredi 10 janvier

**Durée 2h, documents, téléphone, calculatrice interdits**

**Veuillez soigner la présentation (Aithem...) et bien justifier les résultats**

**Vous rédigez directement sur le sujet, vous notez le numéro de la copie sur le sujet (si il n'y a pas de numéro vous en inventez un que vous notez sur la copie et sur le sujet) et à la fin de l'épreuve, vous glissez le sujet dans la copie et vous rendez le tout.**

Le barême (indicatif) prévu est le suivant : 16-8

**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 $\mathbb{R}^k$, $\Sigma$ 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 $\theta$.

**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

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

1. Donner $\hat{\theta}_M$ l'estimateur de $\theta$ par la méthode des moments.

Dimension 1, $\mathbb{E}[X] = \int_0^1 x\,\theta(1-x)^{\theta-1}dx = [-x(1-x)^\theta]_0^1 + \int_0^1 (1-x)^\theta dx = [-\frac{1}{\theta+1}(1-x)^{\theta+1}]_0^1 = \frac{1}{\theta+1}$

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

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

2. Donner l'estimateur du maximum de vraisemblance $\hat{\theta}_n$ de $\theta$.

$f_\theta(x) = \theta(1-x)^{\theta-1}\mathbb{1}_{[0,1]}(x) \Rightarrow$ Indep du support et les autres hypothèses OK

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

On considère les $x_i \in [0,1]$

$\ln L = n\ln\theta + \sum_1^n (\theta-1)\ln(1-x_i)$

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

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

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

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

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

3. Cet estimateur est-il exhaustif ?

$$L = \theta^n \left(\prod_{1}^{n} (1-x_i)^{\theta-1}\right) \mathbb{1}_{\min x_i \geq 0} \mathbb{1}_{\max x_i \leq 1}$$

et $\hat{\theta}_n = \dfrac{-n}{\sum \ln(1-x_i)} \Rightarrow \prod(1-x_i) = \exp\left(-\dfrac{n}{\hat{\theta}_n}\right)$

$$\Rightarrow L = \underbrace{\theta^n \times \exp\left(-\frac{n}{\hat{\theta}_n} \times (\theta-1)\right)}_{\varphi(\theta,\hat{\theta}_n)} \underbrace{\mathbb{1}_{\min x_i \geq 0} \mathbb{1}_{\max x_i \leq 1}}_{\psi(x_1,\dots,x_n)}$$

Donc exhaustif

4. calculer la loi de la v.a. $Y = -\log(1-X)$.

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

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

$u = \ln(1-x)$

$du = \dfrac{-1}{1-x} dx$

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

$\rightarrow Y \sim \mathcal{E}(\theta)$

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

$u = -\log(1-x)$

$du = (1-x)^{-1} dx$

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

$\mathbb{1}_{[0,1[}(x) = \mathbb{1}_{[0,+\infty[}(u)$

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

D'où $Y \sim \mathcal{E}(\theta)$

5. L'estimateur $\hat{\theta}_n$ est-il sans biais ? Donner un estimateur sans biais de $\theta$.

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

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

$= n\,\mathbb{E}\left[\frac{1}{\sum y_i}\right]$

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

et $\frac{1}{\sum Y_i}$ suit une loi inverse gamma de paramètre $n, \theta$ $\Rightarrow n\,\mathbb{E}\left[\frac{1}{\sum Y_i}\right] = n \frac{\theta}{n-1}$

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

$T_n = \frac{n-1}{n}\hat{\theta}_n = \frac{-(n-1)}{\sum \ln(1-x_i)}$ est un estimateur SB de $\theta$

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

6. L'estimateur $\hat{\theta}_n$ est-il convergent ?

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

Variance d'une inverse gamma $n, \theta$

$\Rightarrow \begin{cases} B_\theta(\hat{\theta}_n) \xrightarrow[\infty]{} 0 \\ \oplus \\ V(\hat{\theta}_n) \xrightarrow[\infty]{} 0 \end{cases}$ Convergent.

7. Calculer l'information de Fisher du modèle.

$$I_n(\theta) = -\mathbb{E}\left[\frac{\partial^2 \ln L}{\partial \theta^2}\right] = \frac{n}{\theta^2} \quad \text{et } I(\theta) = \frac{1}{\theta^2}$$

8. L'estimateur $\hat{\theta}_n$ est-il efficace ?

Efficace si $V(\hat{\theta}_n) = \dfrac{(g'(\theta))^2}{I_n(\theta)}$ $\qquad g : x \mapsto \dfrac{nx}{n-1} \qquad g'(x) = \dfrac{n}{n-1}$

$$\frac{(g'(\theta))^2}{I_n(\theta)} = \frac{n^2}{(n-1)^2} \times \frac{1}{n/\theta^2} = \frac{n\theta^2}{(n-1)^2} \neq V(\hat{\theta}_n) = \frac{n^2\theta^2}{(n-1)^2(n-2)}$$

Donc non efficace

9. Donner la normalité asymptotique de $\hat{\theta}_n$.

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

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

10. Donner la normalité asymptotique de $\hat{\theta}_M$ (estimateur des moments).

$\hat{\theta}_M = \frac{1}{\bar{X}_n} - 1$

On applique le TCL à $\bar{X}_n$ $\Rightarrow \sqrt{n}\left(\bar{X}_n - \underbrace{\frac{1}{\theta+1}}_{\mathbb{E}[X_n]}\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \underbrace{V(\bar{X}_n)}_{= V(X) = \frac{\theta}{(\theta+1)^2(\theta+2)}}\right)$

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

$g(\bar{X}_n) = \hat{\theta}_M$

$g\left(\frac{1}{\theta+1}\right) = \theta \quad g'(x) = -\frac{1}{x^2}$

$g'\left(\frac{1}{\theta+1}\right) = -(\theta+1)^2$

$\Rightarrow \sqrt{n}\left(g(\bar{X}_n) - g\left(\frac{1}{\theta+1}\right)\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, V(\bar{X}_n) \times g'\left(\frac{1}{\theta+1}\right)^2\right)$

$\Rightarrow \sqrt{n}(\hat{\theta}_M - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{\theta}{(\theta+1)^2(\theta+2)} \times (\theta+1)^4\right)$

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

11. Donner un intervalle de confiance de niveau asymptotique $1-\alpha$ pour $\theta$.

On a $\sqrt{n}\,\frac{(\hat{\theta}_n - \theta)}{\theta} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$. On remplace le $\theta$ du bas par $\hat{\theta}_n$ car $\frac{\theta}{\hat{\theta}_n} \xrightarrow{\mathbb{P}} 1$ ($\hat{\theta}_n \xrightarrow{\mathbb{P}} \theta$ car EMV)

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

$$\Rightarrow \mathbb{P}\left(-z_{1-\alpha/2} \leq \sqrt{n}\left(\frac{\hat{\theta}_n - \theta}{\hat{\theta}_n}\right) \leq z_{1-\alpha/2}\right) \longrightarrow 1-\alpha$$

$$\Rightarrow \mathbb{P}\left(-\frac{\hat{\theta}_n z_{1-\alpha/2}}{\sqrt{n}} \leq \hat{\theta}_n - \theta \leq \frac{\hat{\theta}_n z_{1-\alpha/2}}{\sqrt{n}}\right) \longrightarrow 1-\alpha$$

$$\Rightarrow \mathbb{P}\left(\hat{\theta}_n - \frac{\hat{\theta}_n z_{1-\alpha/2}}{\sqrt{n}} \leq \theta \leq \hat{\theta}_n + \frac{\hat{\theta}_n z_{1-\alpha/2}}{\sqrt{n}}\right) \longrightarrow 1-\alpha$$

Donc l'IC de niveau asympt $1-\alpha$ pour $\theta$ est $\hat{\theta}_n\left(1 - \frac{z_{1-\alpha/2}}{\sqrt{n}}\right)$ et $\hat{\theta}_n\left(1 + \frac{z_{1-\alpha/2}}{\sqrt{n}}\right)$

12. Construire un text de niveau asymptotique $\alpha$ pour tester $H_0 : \theta = \theta_0$ contre $H_1 : \theta = \theta_1 > \theta_0$.

$H_0 : \theta = \theta_0$ vs $H_1 : \theta = \theta_1 > \theta_0$

$$V_{\theta_1,\theta_0} = \frac{\theta_1^n \prod (1-x_i)^{\theta_1}}{\theta_0^n \prod (1-x_i)^{\theta_0}} = \left(\frac{\theta_1}{\theta_0}\right)^n \exp\left(-\frac{n}{\hat{\theta}_n}\right)^{\theta_1-\theta_0} = \left(\frac{\theta_1}{\theta_0}\right)^n \exp\left(-\frac{n(\theta_1-\theta_0)}{\hat{\theta}_n}\right)$$

$\searrow$ en $\hat{\theta}_n$

$\hat{\theta}_n$ est une stat exhaustive

On rejette $H_0$ si $\hat{\theta}_n < k_\alpha$ où $k_\alpha$ tq $\mathbb{P}_{H_0}(\hat{\theta}_n \leq k_\alpha) = \alpha$

Or sous $H_0$, $\sqrt{n}\,\frac{(\hat{\theta}_n - \theta_0)}{\theta_0} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$

$Z \sim \mathcal{N}(0,1)$

$$\Rightarrow \mathbb{P}_{H_0}\left(\sqrt{n}\,\frac{(\hat{\theta}_n - \theta_0)}{\theta_0} < \underbrace{\frac{\sqrt{n}(k_\alpha - \theta_0)}{\theta_0}}\right) = \alpha$$

$z_\alpha$ = quantile d'ordre $\alpha$ d'une $\mathcal{N}(0,1)$

Donc on rejette $H_0$ si $\hat{\theta}_n < \theta_0\left(1 - \frac{z_\alpha}{\sqrt{n}}\right)$

$$\Rightarrow \sqrt{n}\,\frac{(k_\alpha - \theta_0)}{\theta_0} = z_\alpha$$

$$\Rightarrow k_\alpha = \theta_0\left(1 - \frac{z_\alpha}{\sqrt{n}}\right)$$

13. Construire un text de niveau asymptotique $\alpha$ pour tester $H_0 : \theta = \theta_0$ contre $H_1 : \theta < \theta_0$.

$H_0 : \theta = \theta_0$ vs $H_1 : \theta < \theta_0$

$$V_{\theta,\theta_0} = \left(\frac{\theta}{\theta_0}\right)^n \exp\left(\frac{n\overbrace{(\theta_0 - \theta)}^{>0}}{\hat{\theta}_n}\right) \nearrow \text{ en } \hat{\theta}_n$$

$\hat{\theta}_n$ stat exhaustive

On rejette $H_0$ si $\hat{\theta}_n \geq k'_\alpha$ où $k'_\alpha$ est tq $\mathbb{P}_{H_0}(\hat{\theta}_n \geq k'_\alpha) = \alpha$

Or sous $H_0$, $\sqrt{n}\,\dfrac{(\hat{\theta}_n - \theta_0)}{\theta_0} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$

$Z \sim \mathcal{N}(0,1)$

$$\Rightarrow \mathbb{P}_{H_0}\left(\sqrt{n}\,\frac{(\hat{\theta}_n - \theta_0)}{\theta_0} \geq \underbrace{\sqrt{n}\,\frac{(k'_\alpha - \theta_0)}{\theta_0}}_{Z_{1-\alpha}}\right) = \alpha$$

$Z_{1-\alpha}$ = quantile d'ordre $\alpha$ d'une $\mathcal{N}(0,1)$ $\Rightarrow k'_\alpha = \theta_0\left(1 - \dfrac{Z_{1-\alpha}}{\sqrt{n}}\right)$

Donc on rejette $H_0$ si $\hat{\theta}_n \geq \theta_0\left(1 - \dfrac{Z_{1-\alpha}}{\sqrt{n}}\right)$

**Exercice 2 :**

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 $k = 1, ..., 4$.

$i \in [\![1,n]\!]$

$$\mathbb{E}[X_i] = \int_0^\infty x \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} dx = \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)} x^\alpha e^{-\beta x} dx = \frac{\alpha}{\beta} \int_0^\infty \underbrace{\frac{\beta^{\alpha+1}}{\Gamma(\alpha+1)} x^{\alpha+1-1} e^{-\beta x}}_{f_{\Gamma(\alpha+1,\beta)}} dx = \frac{\alpha}{\beta}$$

$$\mathbb{E}[X_i^2] = \int_0^\infty x^2 \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} dx = \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha+1} e^{-\beta x} dx$$

$$= \frac{\alpha(\alpha+1)}{\beta^2} \int_0^\infty f_{\Gamma(\alpha+2,\beta)}(x) dx = \frac{\alpha(\alpha+1)}{\beta^2}$$

$$\mathbb{E}[X_i^3] = \frac{\alpha(\alpha+1)(\alpha+2)}{\beta^3}$$

$$\mathbb{E}[X_i^4] = \frac{\alpha(\alpha+1)(\alpha+2)(\alpha+3)}{\beta^4}$$

2. Calculer $(\hat{\alpha}, \hat{\beta})$ l'estimateur des moments de $(\alpha, \beta)$.

Dimens° 2 $\quad \mathbb{E}[X_i] = \frac{\alpha}{\beta}$ et $\mathbb{E}[X_i^2] = \frac{\alpha(\alpha+1)}{\beta^2}$

$$\Rightarrow \begin{cases} \dfrac{\hat{\alpha}_M}{\hat{\beta}_M} = \overline{X_n} \\ \dfrac{\hat{\alpha}_M(\hat{\alpha}_M+1)}{\hat{\beta}_M^2} = \overline{X_n^2} \end{cases} \Longrightarrow \begin{cases} \hat{\alpha}_M = \dfrac{\overline{X_n}^2}{\overline{X_n^2} - \overline{X_n}^2} \\ \hat{\beta}_M = \dfrac{\overline{X_n}}{\overline{X_n^2} - \overline{X_n}^2} \end{cases}$$

3. Ecrire le Théorème Centrale Limite pour $(\overline{X}_n, \overline{X^2}_n)$.

$$\sqrt{n}\begin{pmatrix} \overline{X}_n - \mathbb{E}[\overline{X}_n] \\ \overline{X^2_n} - \mathbb{E}[\overline{X^2_n}] \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \begin{pmatrix} V(\overline{X}_n) & Cov(\overline{X}_n, \overline{X^2_n}) \\ Cov(\overline{X^2_n}, \overline{X}_n) & V(\overline{X^2_n}) \end{pmatrix}\right)$$

$$\mathbb{E}[\overline{X}_n] = \mathbb{E}[X_i] = \frac{\alpha}{\beta} \qquad \mathbb{E}[\overline{X^2_n}] = \frac{\alpha(\alpha+1)}{\beta^2}$$

$$Cov(\overline{X}_n, \overline{X^2_n}) = Cov(X_1, X_1^2) = \mathbb{E}\left[\left(X_1 - \frac{\alpha}{\beta}\right)\left(X_1^2 - \frac{\alpha(\alpha+1)}{\beta^2}\right)\right]$$

$$= \mathbb{E}[X_1^3] - \frac{\alpha(\alpha+1)}{\beta^2} \times \frac{\alpha}{\beta} - \frac{\alpha}{\beta}\frac{\alpha(\alpha+1)}{\beta^2} + \frac{\alpha\alpha(\alpha+1)}{\beta^3}$$

$$= \frac{\alpha(\alpha+1)(\alpha+2)}{\beta^3} - \frac{\alpha^2(\alpha+1)}{\beta^3} = \frac{\alpha^2(\alpha+1) + 2\alpha(\alpha+1) - \alpha^2(\alpha+1)}{\beta^3}$$

$$= \frac{2\alpha(\alpha+1)}{\beta^3}$$

$$V(\overline{X^2_n}) = V(X_1^2) = \mathbb{E}[X_1^4] - \mathbb{E}[X_1^2]^2$$

$$= \frac{\alpha(\alpha+1)(\alpha+2)(\alpha+3)}{\beta^4} - \frac{\alpha^2(\alpha+1)^2}{\beta^4}$$

$$= \frac{\alpha(\alpha+1)(\alpha(\alpha+1) + 4\alpha + 6) - \alpha^2(\alpha+1)^2}{\beta^4}$$

$$= \frac{\alpha(\alpha+1)(4\alpha+6)}{\beta^4}$$

$$\Rightarrow \sqrt{n}\begin{pmatrix} \overline{X}_n - \frac{\alpha}{\beta} \\ \overline{X^2_n} - \frac{\alpha(\alpha+1)}{\beta^2} \end{pmatrix} \longrightarrow \mathcal{N}\left(0, \underbrace{\begin{pmatrix} \alpha/\beta^2 & 2\alpha(\alpha+1)/\beta^3 \\ 2\alpha(\alpha+1)/\beta^3 & \frac{\alpha(\alpha+1)(4\alpha+6)}{\beta^4} \end{pmatrix}}_{\Sigma}\right)$$

$$V(\overline{X}_n) = \frac{\alpha(\alpha+1)}{\beta^2} - \frac{\alpha^2}{\beta^2} = \frac{\alpha}{\beta^2}$$

4. Donner la normalité asymptotique de $(\hat{\alpha}, \hat{\beta})$.

On applique la méthode Delta du TCL bidimensionnel appliqué à $(\overline{X}_n, \overline{X^2_n})$ avec

$$g : (x,y) \longmapsto \left(\frac{x}{y}, \frac{x(x+1)}{y^2}\right) \qquad \frac{\partial g}{\partial x} = \left(\frac{1}{y}, \frac{2x+1}{y^2}\right) \qquad \frac{\partial g}{\partial y} = \left(-\frac{x}{y^2}, \frac{-2x(x+1)}{y^3}\right)$$

$$\Rightarrow \sqrt{n}\begin{pmatrix} \hat{\alpha}_n - \alpha \\ \hat{\beta}_n - \beta \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}(0, D_g \Sigma D_g')$$

$$D_g = \begin{pmatrix} \frac{1}{\beta} & -\frac{\alpha}{\beta^2} \\ \frac{2\alpha+1}{\beta^2} & \frac{-2\alpha(\alpha+1)}{\beta^3} \end{pmatrix}$$

matrice Jacobienne de g calculé en $(\alpha, \beta)$

**Exercice 3 :**

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

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

$\mathbb{P}(X_i > 0) = \frac{1+\theta}{2}$

1. Montrer que $K$ est une statistique exhaustive.

Calculons la densité de $K$ notée $f_K$

Soit $k \in [\![0, n]\!]$.

$f_K(k) = \mathbb{P}(K=k) = \binom{n}{k} \mathbb{P}(X_1>0)^k (1-\mathbb{P}(X_1>0))^{n-k}$

$= \binom{n}{k} \left(\int_0^1 \frac{1+\theta}{2} dx\right)^k \left(1 - \int_0^1 \frac{1+\theta}{2} dx\right)^{n-k} = \binom{n}{k} \left(\frac{1+\theta}{2}\right)^k \left(\frac{1-\theta}{2}\right)^{n-k}$

$= \binom{n}{k} \left(\frac{1-\theta}{2}\right)^n \left(\frac{1+\theta}{1-\theta}\right)^k$

Donc $K$ est bien exhaustive

2. Calculer $\hat{\theta}_n$ l'estimateur du maximum de vraisemblance de $\theta$. Où $n_1 = n - K$ et $n_2 = K$

$L = \prod_{i=1}^{n} f(x_i) = \prod_{i=1}^{n=n_1+n_2} \left(\frac{1-\theta}{2}\right)^{n_1} \times \left(\frac{1+\theta}{2}\right)^{n_2} = \prod_{i=1}^{n} \left(\frac{1-\theta}{2}\right)^{n-K} \left(\frac{1+\theta}{2}\right)^{K}$

$\Rightarrow L = \frac{1}{2^n} (1-\theta)^{2n-K} (1+\theta)^{n+K}$

$\ln L = -n \ln(2) + (2n-K)\ln(1-\theta) + (n+K)\ln(1+\theta)$

$\frac{\partial \ln L}{\partial \theta} = -\frac{(2n-K)}{1-\theta} + \frac{(n+K)}{1+\theta} \Rightarrow \hat{\theta}_n$ est tq $-\frac{(2n-K)}{1-\hat{\theta}_n} + \frac{(n+K)}{1+\hat{\theta}_n} = 0 \Rightarrow -(2n-K)(1+\hat{\theta}_n) + (n+K)(1-\hat{\theta}_n) = 0$

$\Rightarrow \hat{\theta}_n = \frac{2K-n}{3n}$

$\frac{\partial^2 \ln L}{\partial \theta^2} = \frac{(2n-K)}{(1-\theta)^2} - \frac{(n+K)}{(1+\theta)^2} = \frac{-3n - 3n\theta^2 - 2n\theta + 4K\theta}{(1-\theta^2)^2} < 0$

Donc $\hat{\theta}_n = \frac{2K-n}{3n}$ est l'EMV de $\theta$

3. Donner la loi de $K$.

Calculons la densité de K notée $f_K$

Soit $k \in [\![0, n]\!]$.

$$f_K(k) = \mathbb{P}(K=k) = \binom{n}{k} \mathbb{P}(X_1>0)^k (1-\mathbb{P}(X_1>0))^{n-k}$$

$$= \binom{n}{k} \left(\int_0^1 \frac{1+\theta}{2} dx\right)^k \left(1-\int_0^1 \frac{1+\theta}{2} dx\right)^{n-k} = \binom{n}{k} \left(\frac{1+\theta}{2}\right)^k \left(\frac{1-\theta}{2}\right)^{n-k}$$

$$= \binom{n}{k} \left(\frac{1-\theta}{2}\right)^n \left(\frac{1+\theta}{1-\theta}\right)^k$$

$$K \sim \mathcal{B}\left(n, \frac{1+\theta}{2}\right)$$

4. Calculer le biais de $\hat{\theta}_n$.

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

$$\mathbb{E}[\hat{\theta}_n] = \mathbb{E}\left[\frac{2K-n}{3n}\right] = \frac{2}{3n}\mathbb{E}[K] - \frac{1}{3}$$

$$\mathbb{E}[K] = \sum_{k=0}^{n} k\, \mathbb{P}(K=k) = \sum_{k=1}^{n} k \binom{n}{k} \left(\frac{1+\theta}{2}\right)^k \left(\frac{1-\theta}{2}\right)^{n-k}$$

On sait que $k\binom{n}{k} = n\binom{n-1}{k-1}$

$$\Rightarrow \mathbb{E}[K] = \sum_{k=1}^{n} n\binom{n-1}{k-1} \left(\frac{1+\theta}{2}\right)^k \left(\frac{1-\theta}{2}\right)^{n-k} = \sum_{k=0}^{n-1} n\binom{n}{k} \left(\frac{1+\theta}{2}\right)^{k+1} \left(\frac{1-\theta}{2}\right)^{n-1-k}$$

$$\Rightarrow \mathbb{E}[K] = n\left(\frac{1+\theta}{2}\right) \sum_{k=0}^{n-1} \binom{n}{k} \left(\frac{1+\theta}{2}\right)^k \left(\frac{1-\theta}{2}\right)^{n-1-k}$$

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

Binôme de newton



$$\Rightarrow \mathbb{E}[\hat{\theta}_n] = \frac{2}{3n} \times n\left(\frac{1+\theta}{2}\right) - \frac{1}{3} = \frac{\theta}{3}$$

$$\Rightarrow B_\theta(\hat{\theta}_n) = \frac{\theta}{3} - \theta = -\frac{2\theta}{3}$$

5. $\hat{\theta}_n$ est-il convergent ?

Non car $B_\theta(\hat{\theta}_n) = \frac{-2\theta}{3}$ Donc $\hat{\theta}_n$ n'est ni SB ni asympt SB

Donc non CV

6. Donner la normalité asymptotique de $\hat{\theta}_n$.

Hypothèses OK

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

$I_n(\theta) = -\mathbb{E}\left[\frac{\partial^2 \ell L}{\partial \theta^2}\right] = -\mathbb{E}\left[\frac{-3n-3n\theta^2-2n\theta+4k\theta}{(1-\theta^2)^2}\right]$

$= \frac{1}{(1-\theta^2)^2}\left(3n+3n\theta^2+2n\theta-4\theta\,\mathbb{E}[k]\right)$

Or $\mathbb{E}[k] = \frac{n(1+\theta)}{2}$

$\Rightarrow 3n+3n\theta^2+2n\theta+4\theta\,\mathbb{E}[k] = 3n+3n\theta^2+2n\theta-4\theta n\left(\frac{1+\theta}{2}\right)$

$= 3n+3n\theta^2+2n\theta-2n\theta-2n\theta^2$

$= 3n+n\theta^2$

$\Rightarrow I_n(\theta) = \frac{3n+n\theta^2}{(1-\theta^2)^2}$ $\Rightarrow I(\theta) = \frac{3+\theta^2}{(1-\theta^2)^2}$

Donc $\sqrt{n}\,(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{(1-\theta^2)^2}{3+\theta^2}\right)$

7. Déterminer un intervalle de confiance de $\theta$ de niveau asymptotique $1-\alpha$.

On a $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{(1-\theta^2)^2}{3+\theta^2}\right) \Longrightarrow \sqrt{n}\frac{(\hat{\theta}_n - \theta)}{\sqrt{\frac{(1-\theta^2)^2}{3+\theta^2}}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$

$Z \sim \mathcal{N}(0,1)$

$$\Longrightarrow \mathbb{P}\left(-Z_{1-\frac{\alpha}{2}} \leq \sqrt{n}\frac{(\hat{\theta}_n - \theta)}{\sqrt{\frac{(1-\theta^2)^2}{3+\theta^2}}} \leq Z_{1-\frac{\alpha}{2}}\right) = 1-\alpha$$

Comme $\hat{\theta}_n$ EMV de $\theta$, $\hat{\theta}_n \xrightarrow{\mathbb{P}} \theta$ Donc on remplace les $\theta$ du bas par $\hat{\theta}_n$

$$\Longrightarrow \mathbb{P}\left(-Z_{1-\frac{\alpha}{2}} \leq \sqrt{n}\frac{(\hat{\theta}_n - \theta)}{\sqrt{\frac{(1-\hat{\theta}_n^2)^2}{3+\hat{\theta}_n^2}}} \leq Z_{1-\alpha/2}\right) = 1-\alpha$$

$$\Longrightarrow \mathbb{P}\left(\underbrace{\hat{\theta}_n - \frac{(1-\hat{\theta}_n^2) Z_{1-\alpha/2}}{\sqrt{n(3+\hat{\theta}_n^2)}}}_{A_n} \leq \theta \leq \underbrace{\hat{\theta}_n + \frac{(1-\hat{\theta}_n^2) Z_{1-\alpha/2}}{\sqrt{n(3+\hat{\theta}_n^2)}}}_{B_n}\right) = 1-\alpha$$

8. Construire un test de niveau asymptotique $\alpha$ pour tester $H_0 : \theta = \theta_0$ contre $H_1 : \theta = \theta_1 > \theta_0$.

$$V_{\theta_1,\theta_0} = \frac{L(\_,\theta_1)}{L(\_,\theta_0)} = \frac{2^{-n}(1-\theta_1)^{2n-K}(1+\theta_1)^{n+K}}{2^{-n}(1-\theta_0)^{2n-K}(1+\theta_0)^{n+K}} = \left(\frac{1-\theta_1}{1-\theta_0}\right)^{2n-K}\left(\frac{1+\theta_1}{1+\theta_0}\right)^{n+K}$$

$$= \left(\frac{1-\theta_1}{1-\theta_0}\right)^{2n}\left(\frac{1+\theta_1}{1+\theta_0}\right)^{n}\left(\frac{(1-\theta_0)(1+\theta_1)}{(1-\theta_1)(1+\theta_0)}\right)^{K}$$

On considère $\theta_1 > \theta_0 \geq 1$ Sinon même résultat

$K$ est exhaustive et $K \sim \mathcal{B}(n, \frac{1+\theta}{2})$

Donc $V_{\theta_1,\theta_0}$ est $\nearrow$ en fonct° de $K$

Donc on rejette $H_0$ si $K \geq k_\alpha$ avec $k_\alpha$ tq $\mathbb{P}_{H_0}(K \geq k_\alpha) = \alpha$

Sous $H_0$, $K \sim \mathcal{B}(n, \frac{1+\theta_0}{2})$

On note $B \sim \mathcal{B}(n, \frac{1+\theta_0}{2})$

$\Longrightarrow k_\alpha = B_{1-\alpha}^{n,\frac{1+\theta_0}{2}}$ ← quantile d'ordre $1-\alpha$ d'une $\mathcal{B}(n, \frac{1+\theta_0}{2})$

Donc on rejette $H_0$ si $\sum_{i=1}^{n} \mathbb{1}_{(X_i > 0)} \geq B_{1-\alpha}^{n,\frac{1+\theta_0}{2}}$