M1 Actuariat, économétrie et statistiques, année 2021–2022.

NOM, Prénom :

Statistique Inférentielle
Contrôle continu

Lundi 17 Octobre
Durée 1h30, documents, téléphone, calculatrice interdits

Le barême (indicatif) prévu est le suivant : 6–7–8 (on tiendra compte (grave) de la présentation et de la clareté des explications):

On rappelle que la densité d'une loi $\Gamma(\alpha, \beta)$ vaut :

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

Exercice 1

Soit $X_1, X_2, ..., X_n$ un échantillon de variables aléatoires de même loi X, où X admet pour densité

$$f_\theta(x) = \frac{2x}{\theta^2} \mathbb{1}_{[0,\theta]}$$

1. Calculer la fonction de répartition de X, son espérance et sa variance.

$F_\theta(x) = \mathbb{P}(X \leq x) = \int_0^x \frac{2t}{\theta^2} dt = \left[\frac{t^2}{\theta^2} \right]_0^x = \frac{x^2}{\theta^2}$ si $x \in [0,\theta]$

$\mathbb{E}[X] = \int_0^\theta \frac{2x^2}{\theta^2} dx = \left[\frac{2}{\theta^2} \frac{x^3}{3} \right]_0^\theta = \frac{2}{3}\theta$

$\mathbb{E}[X^2] = \int_0^\theta \frac{2x^3}{\theta^2} dx = \frac{2}{\theta^2} \left[\frac{x^4}{4} \right]_0^\theta = \frac{\theta^2}{2}$

$V(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{\theta^2}{2} - \frac{4\theta^2}{9} = \frac{\theta^2}{18}$

2. Ah mais vous y avez cru ?? RRroooooo lalala, vous auriez vu vos têtes ! Mais non, bien évidemment, c'est pas un CC, je suis pas encore à ce niveau de fdp. C'est juste un entrainement. On va le faire pendant 1h15/30 et on corrige dans la seconde partie du partie. Du coup, question suivante : Calculer l'estimateur du maximum de vraisemblance $\hat{\theta}_n$ de θ ?

Ici $\mathcal{X} = [0;\theta]$ donc on n'a pas (H1) $\longrightarrow$ pas de $\ln L$ et $\frac{\partial}{\partial\theta}\ln L$ ...

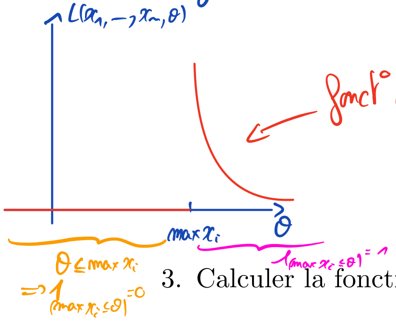
$$L(x_1,\dots,x_n,\theta) = \prod_{i=1}^n f_\theta(x_i) = \prod_{i=1}^n \frac{2x_i}{\theta^2} \mathbb{1}_{[0;\theta]}(x_i)$$

$$= \frac{2^n}{\theta^{2n}} \prod_{i=1}^n x_i \underbrace{\prod_{i=1}^n \mathbb{1}_{[0;\theta]}(x_i)}_{= \begin{cases} 1 & \text{si tous les } x_i \in [0;\theta] \\ 0 & \text{sinon.} \end{cases}}$$

i.e. $\min x_i \geq 0$, $\max x_i \leq \theta$.

$$= \frac{2^n}{\theta^{2n}} \prod_{i=1}^n x_i \,\mathbb{1}_{(\min x_i \geq 0)} \mathbb{1}_{(\max x_i \leq \theta)}$$

A $x_1,\dots,x_n$ fixé et $\min x_i > 0$



fonct° décroissante en θ car $\propto \frac{1}{\theta^{2n}}$

$$\Longrightarrow \hat{\theta}_n = \max_{i\in[1,n]} X_i$$

3. Calculer la fonction de répartition et la densité de $\hat{\theta}_n$.

$$\mathbb{P}(\hat{\theta}_n \leq x) = \mathbb{P}(\max X_i \leq x) = \mathbb{P}(X_1 \leq x, \dots, X_n \leq x) = \mathbb{P}(X_i \leq x)^n$$

$$= \left(\frac{x^2}{\theta^2}\right)^n = \left(\frac{x}{\theta}\right)^{2n} \text{ si } x \in [0;\theta]$$

$$\Longrightarrow \text{densité } f_{\hat{\theta}_n}(x) = \frac{2n\, x^{2n-1}}{\theta^{2n}} \mathbb{1}_{[0;\theta]}(x)$$

4. Calculer le biais et la variance $\hat{\theta}_n$. $\hat{\theta}_n$ est-il convergent ?

$$\mathbb{E}[\hat{\theta}_n] = \int_0^\theta x \frac{2n x^{2n-1}}{\theta^{2n}} dx = \int_0^\theta \frac{2n x^{2n}}{\theta^{2n}} dx = \left[\frac{2n}{\theta^{2n}} \frac{x^{2n+1}}{2n+1}\right]_0^\theta = \frac{2n}{2n+1}\theta \longrightarrow \theta.$$

$\Rightarrow \hat{\theta}_n$ est asympt. sans biais pour θ.

$$\mathbb{E}[\hat{\theta}_n^2] = \int_0^\theta x^2 \frac{2n x^{2n-1}}{\theta^{2n}} dx = \left[\frac{2n}{\theta^{2n}} \frac{x^{2n+2}}{2n+2}\right]_0^\theta = \frac{2n}{2n+2}\theta^2$$

$$V(\hat{\theta}_n) = \mathbb{E}[\hat{\theta}_n^2] - \mathbb{E}[\hat{\theta}_n]^2 = \frac{2n\theta^2}{2n+2} - \frac{(2n\theta)^2}{(2n+1)^2} = \frac{2n\theta^2}{(2n+2)(2n+1)^2} \longrightarrow 0.$$

$\oplus \Rightarrow \hat{\theta}_n$ convergent pour θ.

5. Montrer que

$$2n\left(1 - \frac{\hat{\theta}_n}{\theta}\right) \xrightarrow{\mathcal{L}} \mathcal{E}(1)$$

On veut $2n\left(1 - \frac{\hat{\theta}_n}{\theta}\right) \xrightarrow{\mathcal{L}} \mathcal{E}(1)$

$$\mathbb{P}\left(2n\left(1-\frac{\hat{\theta}_n}{\theta}\right) \leq x\right) = \mathbb{P}\left(1-\frac{\hat{\theta}_n}{\theta} \leq \frac{x}{2n}\right) = \mathbb{P}\left(\hat{\theta}_n \geq \theta\left(1-\frac{x}{2n}\right)\right)$$
$$= 1 - \mathbb{P}\left(\hat{\theta}_n \leq \theta\left(1-\frac{x}{2n}\right)\right)$$
$$= 1 - \left(\frac{\theta\left(1-\frac{x}{2n}\right)}{\theta}\right)^{2n} = 1 - \left(1-\frac{x}{2n}\right)^{2n}$$
$$= 1 - \exp\left(2n \ln\left(1-\frac{x}{2n}\right)\right)$$

Donc $$\mathbb{P}\left(2n\left(1-\frac{\hat{\theta}_n}{\theta}\right) \leq x\right) = 1 - \exp\Big(2n \underbrace{\ln\left(1-\frac{x}{2n}\right)}_{\sim -\frac{x}{2n}}\Big) \sim \underbrace{1-\exp(-x)}_{\text{f.d.r d'une } \mathcal{E}(1)}$$

$$\longrightarrow 2n\left(1-\frac{\hat{\theta}_n}{\theta}\right) \xrightarrow{\mathcal{L}} \mathcal{E}(1)$$

6. Calculer l'estimateur $\hat{\theta}_M$ par la méthode des moments de θ.

Rappel : Nb de moments $\mathbb{E}[X^p]$ = Dimens° du paramètre

Ici, $\theta \in \mathbb{R} \longrightarrow$ On a besoin d'un moment i.e $\mathbb{E}[X]$

On a $\mathbb{E}[X] = M_1(\theta)$ et $\hat{\theta}_M$ est tq $M_1(\hat{\theta}_M) = \bar{X}_n$

Ici, on a $\mathbb{E}[X] = \frac{2}{3}\theta$

$\hat{\theta}_M$ est tq $\frac{2}{3}\hat{\theta}_M = \bar{X}_n \Longrightarrow \hat{\theta}_M = \frac{3}{2}\bar{X}_n$

7. Calculer le biais et la variance de $\hat{\theta}_M$. Est-il convergent ?

$\mathbb{E}[\hat{\theta}_M] = \mathbb{E}[\frac{3}{2}\bar{X}_n] = \frac{3}{2}\mathbb{E}[\bar{X}_n] = \frac{3}{2}\mathbb{E}[X_i] = \theta.$

$V(\hat{\theta}_M) = V(\frac{3}{2}\bar{X}_n) = \frac{9}{4}V(\bar{X}_n) = \frac{9}{4}\frac{V(X_i)}{n} = \frac{\theta^2}{8n} \longrightarrow 0$

$\hat{\theta}_M$ est convergent pour θ.

8. Calculer la normalité asymptotique de $\hat{\theta}_M$. ⊛ en dimens° 1

Rappel: Si on a l'estimateur des moments ⊛ $\hat{\theta}_M = g(\bar{X}_n)$

On prend le TCL

$$\sqrt{n}\,(\bar{X}_n - \mathbb{E}[X_1]) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2(\theta))$$

$V(X_1)$

On applique la méthode Delta au TCL et à g

$$\sqrt{n}\,(\underbrace{g(\bar{X}_n)}_{=\hat{\theta}_M} - \underbrace{g(\mathbb{E}[X_1])}_{=\theta}) \xrightarrow{\mathcal{L}} \mathcal{N}(0, (g'(\mathbb{E}[X_1]))^2\sigma^2(\theta))$$

Ici, $\hat{\theta}_M = \frac{3}{2}\bar{X}_n$ i.e $g(x) = \frac{3}{2}x \longrightarrow g'(x) = \frac{3}{2}$

On prend le TCL appliqué à $X_1, \ldots, X_n$

$$\sqrt{n}\,(\bar{X}_n - \underbrace{\tfrac{3}{2}\theta}_{\mathbb{E}[X]}) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \underbrace{\tfrac{\theta^2}{18}}_{V(X)})$$

On applique la méthode Delta au TCL et à g

$$\sqrt{n}\,(\underbrace{g(\bar{X}_n)}_{\hat{\theta}_M} - \underbrace{g(\tfrac{2}{3}\theta)}_{\theta}) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \underbrace{\underbrace{(g'(\tfrac{2}{3}\theta))^2}_{(\frac{3}{2})^2} \times \tfrac{\theta^2}{18}}_{=\frac{\theta^2}{8}})$$

$$\Rightarrow \sqrt{n}\,(\hat{\theta}_M - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \tfrac{\theta^2}{8})$$

9. Quel estimateur entre $\hat{\theta}_n$ et $\hat{\theta}_M$ est préférable au sens du risque quadratique.

$$\mathcal{R}(\hat{\theta}_M, \theta) = B_\theta(\hat{\theta}_M) + V(\hat{\theta}_M) = 0 + \frac{\theta^2}{8n} = \frac{\theta^2}{8n}$$

$$\mathcal{R}(\hat{\theta}_n, \theta) = B_\theta(\hat{\theta}_n) + V(\hat{\theta}_n) \sim \frac{1}{n^2}$$

$$\Rightarrow \boxed{\mathcal{R}(\hat{\theta}_n, \theta) \leq \mathcal{R}(\hat{\theta}_M, \theta)}$$

Exercice 2

On dispose d'un échantillon $X_1, ..., X_n$ i.i.d. issu d'une loi de Poisson de paramètre θ.

1. Ecrire le modèle statistique et montrer qu'il est dominé.

Le modèle est $(\mathbb{N}, \mathcal{P}_{\text{art de } \mathbb{N}}, \mathcal{P}(\lambda), \lambda \in \mathbb{R}_+^* \subset \mathbb{R})^n$

La loi de Poisson est dominée par la mesure de comptage.

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

Ici on a (H1) à (H4) vérifiées.

$$L(x_1, \dots, x_n, \theta) = \prod_{i=1}^n \mathbb{P}(X_i = x_i) = \prod_{i=1}^n \frac{\theta^{x_i}}{x_i!} e^{-\theta}$$

$$= \frac{\theta^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} e^{-n\theta}. \Rightarrow \ln(L) = \sum_{i=1}^n x_i \ln(\theta) - \ln \prod x_i! - n\theta.$$

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

$\hat{\theta}_n$ est tq $\dfrac{\sum_{i=1}^n x_i}{\theta} - n = 0 \Rightarrow \boxed{\hat{\theta}_n = \dfrac{\sum_{i=1}^n x_i}{n}}$

$$\frac{\partial^2}{\partial \theta^2} \ln L = -\frac{\sum_{i=1}^n x_i}{\theta^2} < 0$$

$\Rightarrow \hat{\theta}_n$ est l'EMV de θ.

3. Montrer que l'estimateur est exhaustif.

$\hat{\theta}_n$ est exhaustif car $L(x_1, \ldots, x_n, \theta) = \frac{\theta^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!} e^{-n\theta} = \underbrace{\theta^{\overbrace{n\bar{X}_n}^{n\hat{\theta}_n}} e^{-n\theta}}_{\varphi(\hat{\theta}_n, \theta)} \times \underbrace{\left(\prod_{i=1}^{n} x_i!\right)}_{\psi(x_1, \ldots, x_n)}$

$\Rightarrow \hat{\theta}_n$ exhaustif pour θ.

4. Cet estimateur est-il sans biais ? Convergent ? Efficace ?

$\mathbb{E}[\hat{\theta}_n] = \mathbb{E}[\bar{X}_n] = \theta \Rightarrow \hat{\theta}_n$ sans biais pour θ

$V(\hat{\theta}_n) = V(\bar{X}_n) = \frac{V(X_i)}{n} = \frac{\theta}{n} \longrightarrow 0$

$\Rightarrow \hat{\theta}_n$ Convergent pour θ.

$I_n(\theta) = \mathbb{E}\left[-\frac{\partial^2}{\partial\theta^2} \ln L(X_1, \ldots, X_n, \theta)\right] = \mathbb{E}\left[\frac{\sum_{i=1}^{n} X_i}{\theta^2}\right] = \frac{n}{\theta}$

Ici, $\hat{\theta}_n$ est sans biais donc $K_\theta(\hat{\theta}_n) = I(\theta)^{-1} = \frac{\theta}{n}$

$V(\hat{\theta}_n) = \frac{\theta}{n}$ $\Rightarrow \hat{\theta}_n$ efficace pour θ

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

Comme $\hat{\theta}_n$ est l'EMV de θ et (H1) à (H7) vérifiées, on a

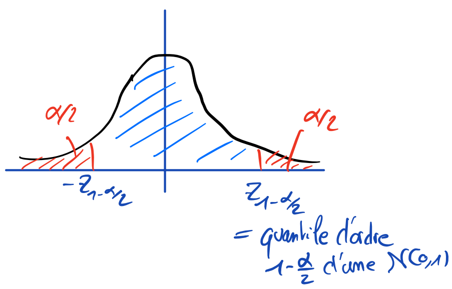
$$\sqrt{n}\,(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, I^{-1}(\theta))$$

Ici, $I_n(\theta) = \frac{n}{\theta} \implies I(\theta) = \frac{1}{\theta}$

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

6. (Bonus : hors programme) Donner un intervalle de confiance asymptotique de niveau $1 - \alpha$ de θ.

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



On a $\mathbb{P}(-z_{1-\frac{\alpha}{2}} \leq Z \leq z_{1-\frac{\alpha}{2}}) = 1-\alpha$

donc

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

Rappel : $\left.\begin{array}{l} Y_n \xrightarrow{\mathcal{L}} Y \\ X_n \xrightarrow{\mathbb{P}} 1 \end{array}\right) Y_n X_n \xrightarrow{\mathcal{L}} Y$

Or $\hat{\theta}_n \xrightarrow{\mathbb{P}} \theta$ $\qquad \sqrt{\frac{\theta}{\hat{\theta}_n}} \xrightarrow{\mathbb{P}} 1$

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

Du coup, $\mathbb{P}\left(-z_{1-\frac{\alpha}{2}} \leq \sqrt{n}\,\frac{\hat{\theta}_n - \theta}{\sqrt{\hat{\theta}_n}} \leq z_{1-\frac{\alpha}{2}}\right) \longrightarrow 1-\alpha.$

et

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

7. On pose

$$\hat{\theta}^* = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$$

Cet estimateur est-il biaisé ? Quel estimateur entre $\hat{\theta}_n$ et $\hat{\theta}^*$ est préférable au sens du risque quadratique ?

Biais de $\hat{\theta}_m$: $\xrightarrow{\text{Psychopathes}}$ $\mathbb{E}[\hat{\theta}_m^*] = \frac{1}{m-1}\sum_{i=1}^{m}\left(\mathbb{E}[X_i^2] - \frac{2}{n}\sum_{j=1}^{n}\mathbb{E}[X_iX_j] + \frac{1}{m^2}\mathbb{E}\left(\left(\sum_{j=1}^{m}X_j\right)^2\right)\right)$

$\longrightarrow$ Rappel : Si $X_1, \ldots, X_m$ iid tq $X_i \sim \mathbb{P}$ et $\mathbb{V}(X_i) = \sigma^2 < \infty$

alors $\frac{1}{m-1}\sum_{i=1}^{m}(X_i - \overline{X}_m)^2$ est un ESB convergent de θ.

Ici, $\mathbb{V}(X_i) = \mathbb{E}[X_i] = \theta$ donc $\mathbb{E}[\hat{\theta}_m^*] = \theta$.

On a $\hat{\theta}_m$ et $\hat{\theta}_m^*$ qui sont sans biais

$\mathcal{R}(\hat{\theta}_m, \theta) = \mathbb{V}(\hat{\theta}_m)$

$\mathcal{R}(\hat{\theta}_m^*, \theta) = \mathbb{V}(\hat{\theta}_m^*)$ $\Longrightarrow \mathcal{R}(\hat{\theta}_n, \theta) < \mathcal{R}(\hat{\theta}_m^*, \theta)$

mais $\hat{\theta}_m$ est efficace pour θ i.e. $\forall\, T_m$ un estimateur sans biais de θ

$\mathbb{V}(T_m) > \mathbb{V}(\hat{\theta}_m)$

8. On veut maintenant estimer $\lambda = e^{-\theta} = \mathbb{P}(X = 0)$. On pose

$$Y_i = \mathbb{1}_{X_i=0} = \begin{cases} 1 \text{ si } X_i = 0 \text{ avec proba } \mathbb{P}(X_i=0) = \lambda \\ 0 \text{ sinon avec proba } 1-\lambda. \end{cases}$$

Donner la loi de Y_1, son espérance et sa variance.

$\Longrightarrow Y_i \sim \mathcal{B}ern(\lambda)$ $\longrightarrow \mathbb{E}[X_i] = \lambda$

$\longrightarrow \mathbb{V}(Y_i) = \lambda(1-\lambda)$

9. On se propose d'estimer λ par

$$\hat{\lambda}^* = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

Cet estimateur est-il convergent ?

LGN : $X_1, \dots, X_n$ iid tq $\mathbb{E}[X_i] = \mu$ $\quad \frac{1}{n}\sum X_i \xrightarrow{\mathbb{P}} \mu$.

Ici, on a $\hat{\lambda}_n^* = \frac{1}{n}\sum_{i=1}^{n} Y_i \xrightarrow{\mathbb{P}} \lambda$ donc $\hat{\lambda}_n^*$ est convergent

10. Ecrire la vraisemblance de l'échantillon relativement au paramètre λ et calculer l'estimateur du maximum de vraisemblance $\hat{\lambda}_n$ de λ.

On a

$$L(x_1, \dots, x_n, \theta) = \frac{\theta^{\sum x_i}}{\prod x_i!} e^{-n\theta}$$

On pose $\lambda = e^{-\theta} \Rightarrow \theta = -\ln \lambda$.

On a $L(x_1, \dots, x_n, \lambda) = \frac{(-\ln \lambda)^{\sum x_i}}{\prod x_i!} \lambda^n$

$$\ln L(x_1, \dots, x_n, \lambda) = \left(\sum x_i\right) \ln(-\ln \lambda) - \ln\left(\prod x_i!\right) + n \ln \lambda.$$

$$\frac{\partial}{\partial \lambda} \ln L(x_1, \dots, x_n, \lambda) = \frac{\sum x_i}{\lambda \ln(\lambda)} + \frac{n}{\lambda}$$

$\hat{\lambda}_n$ est tq $\frac{\sum x_i}{\hat{\lambda}_n \ln \hat{\lambda}_n} + \frac{n}{\hat{\lambda}_n} = 0 \Rightarrow \hat{\lambda}_n = e^{-\frac{\sum_{i=1}^{n} x_i}{n}} = e^{-\bar{X}_n}$

$$\frac{\partial^2}{\partial \lambda^2} \ln L = -\sum_{i=1}^{n} x_i \frac{(\lambda \ln \lambda)'}{(\lambda \ln \lambda)^2} - \frac{n}{\lambda^2} < 0$$

$\Rightarrow \hat{\lambda}_n$ est l'EMV de λ.

11. Donner la normalité asymptotique de $\hat{\lambda}_n$.

$$I_n(\lambda) = V\left(\frac{\sum X_i}{\lambda \ln \lambda} + \frac{n}{\lambda}\right) = \frac{V(\sum X_i)}{\lambda^2 \ln(\lambda)^2} = \frac{-n}{\lambda^2 \ln(\lambda)}$$

$$I(\lambda) = -\frac{1}{\lambda^2 \ln(\lambda)}$$

$\hat{\lambda}_n$ est l'EMV de λ, on a $\sqrt{n}\,(\hat{\lambda}_n - \lambda) \xrightarrow{\mathcal{L}} \mathcal{N}(0, I^{-1}(\lambda))$

donc $\sqrt{n}\,(\hat{\lambda}_n - \lambda) \xrightarrow{\mathcal{L}} \mathcal{N}(0, -\lambda^2 \ln \lambda)$

Exercice 3

On suppose qu'au cours de l'année écoulée, le chiffre d'affaire des entreprises ayant un chiffre d'affaire inférieur à $c > 0$ (fixé) peut être représenté par une variable aléatoire X absolument continue de densité :

$$f(x) = \frac{x^{\frac{1}{\theta}-1}}{\theta c^{\frac{1}{\theta}}} \mathbb{1}_{\{0 \leq x < c\}}$$

où $\theta > 0$ est un paramètre, c est supposé connu. Soient $X_1, .., X_n$ n variables aléatoires i. i. d. de même loi que X.

1. Déterminer la loi de $\ln c - \ln X$.

Soit $x \in \mathbb{R}$, $\mathbb{P}(\ln c - \ln X \leq x) = \mathbb{P}(\ln \frac{c}{X} \leq x)$

$= \mathbb{P}(\frac{c}{X} \leq e^{x}) = \mathbb{P}(X \geq ce^{-x})$

$= 1 - \mathbb{P}(X \leq ce^{-x})$

Si $x < 0$, $ce^{-x} \geq c \Rightarrow \mathbb{P}(X \leq ce^{-x}) = 1 \Rightarrow \mathbb{P}(\ln c - \ln(X) \leq x) = 0$

Si $x > 0$, $\mathbb{P}(\ln c - \ln X \leq x) = 1 - \int_0^{ce^{-x}} \frac{r^{1/\theta - 1}}{\theta c^{1/\theta}} dr = 1 - \frac{\theta}{\theta c^{1/\theta}} [r^{1/\theta}]_0^{ce^{-x}}$

$= 1 - e^{-\frac{x}{\theta}}$

Donc $f_{\ln c - \ln X}(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}} \mathbb{1}_{\mathbb{R}_+}(x)$.

On reconnait donc que $\ln c - \ln X \sim \mathcal{E}(\frac{1}{\theta})$

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

$L(x_1, —, x_n, \theta) = \frac{1}{\theta^n c^{n/\theta}} \prod_{i=1}^{n} x_i^{1/\theta - 1} = \theta^{-n} \left(\prod_{i=1}^{n} \frac{x_i}{c}\right)^{1/\theta} \prod_{i=1}^{n} x_i^{-1}$

Les hypothèses 1 à 8 sont vérifiées.

$\ln L = -n \ln \theta - \frac{1}{\theta} \sum_{i=1}^{n} \ln c - \ln x_i - \sum_{i=1}^{n} \ln x_i$

$\frac{\partial}{\partial \theta} \ln L = -\frac{n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^{n} \ln c - \ln x_i$ $\rightarrow$ $\hat{\theta}_n$ est tq $-\frac{n}{\hat{\theta}_n} + \frac{1}{\hat{\theta}_n^2} \sum \ln c - \ln x_i = 0 \Rightarrow \hat{\theta}_n = \frac{1}{n} \sum_{i=1}^{n} (\ln c - \ln x_i)$

$\frac{\partial^2}{\partial \theta^2} \ln L = + \frac{n}{\theta^2} - \frac{2}{\theta^3} \sum_{i=1}^{n} \ln c - \ln x_i$

3. L'estimateur $\hat{\theta}_n$ est-il sans biais ? Asympt. sans biais ?

$$\mathbb{E}[\hat{\theta}_n] = \frac{1}{n}\mathbb{E}\left[\sum_{i=1}^{n}(\ln c - \ln X_i)\right] = \mathbb{E}[\ln c - \ln X] = \theta \quad \text{car} \quad \ln c - \ln X \sim \mathcal{E}\left(\frac{1}{\theta}\right)$$

Donc l'estimateur $\hat{\theta}_n$ est sans biais.

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

$$V(\hat{\theta}_n) = \frac{1}{n^2} V\left(\sum_{i=1}^{n}(\ln c - \ln X_i)\right) = \frac{1}{n} V(\ln c - \ln X) = \frac{1}{n}\theta^2 \xrightarrow[n\to\infty]{} 0$$

Donc $\hat{\theta}_n$ est convergent.

5. L'estimateur $\hat{\theta}_n$ est-il exhaustif ?

$$L(x_1, \ldots, x_n, \theta) = \frac{1}{\theta^n c^{n/\theta}} \prod_{i=1}^{n} x_i^{1/\theta - 1}$$

$$= \frac{1}{\theta^n} \prod_{i=1}^{n} x_i^{-1} \left(\prod_{i=1}^{n} \frac{x_i}{c} \right)^{1/\theta}$$

or $\exp(\hat{\theta}_n) = \left(\prod_{i=1}^{n} \frac{c}{x_i} \right)^n$

donc $L(x_1, \ldots, x_n, \theta) = \underbrace{\frac{1}{\theta^n} \exp\left(-\frac{\hat{\theta}_n}{n\theta}\right)}_{\varphi(\hat{\theta}_n, \theta)} \underbrace{\prod_{i=1}^{n} x_i^{-1}}_{\psi(x_1, \ldots, x_n)}$

D'où $\hat{\theta}_n$ est exhaustif.

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

* $\mathbb{E}[\hat{\theta}_n] = \theta$

$$-\frac{\partial^2}{\partial \theta^2} \ln(L) = -\frac{n}{\theta^2} + \frac{2}{\theta^3} \left(\sum_{i=1}^{n} \ln c - \ln(x_i) \right)$$

$$I_n(\theta) = -\frac{n}{\theta^2} + \frac{2n}{\theta^2} = \frac{n}{\theta^2}$$

$$V(T_n) = \frac{\theta^2}{n} = \frac{1}{I_n(\theta)}$$

Donc l'estimateur $\hat{\theta}_n$ est efficace

Rappel :

* On dit que l'estimateur T_n est efficace s'il vérifie :
$$V(T_n) = K_{T_n}(\theta) = \frac{(g'(\theta))^2}{I_n(\theta)} \quad \text{où } g(\theta) = \mathbb{E}_\theta[T_n]$$

* Si T_n n'est pas efficace mais que $\frac{K_{T_n}(\theta)}{V(T_n)} \xrightarrow[\infty]{} 1$
T_n est asympt. efficace.

7. L'estimateur $\hat{\theta}_n$ est-il asymptotiquement normal ? Si oui, préciser sa variance asymptotique ?

Les hypothèses 7 et 8 sont vérifiées

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

Or, $I(\theta) = \frac{1}{\theta^2}$

donc $\sqrt{n}(\hat{\theta}_n - \theta) \longrightarrow \mathcal{N}(0\,;\,\theta^2)$

8. On suppose que θ est connu et que c est le paramètre inconnu. Donner l'estimateur $\hat{c}_n$ du maximum de vraisemblance de c.

Le support n'est pas homogène

$$L(x_1, \ldots, x_n, c) = \frac{1}{\theta^n c^{n/\theta}} \prod_{i=1}^{n} x_i^{\frac{1}{\theta}-1} \prod_{i=1}^{n} \mathbb{1}_{\{0 \leq x_i < c\}} = \underbrace{\frac{1}{\theta^n c^{n/\theta}}}_{\text{fonct° décroissante par rapport à } c} \prod_{i=1}^{n} x_i^{\frac{1}{\theta}-1} \mathbb{1}_{\{0 \leq \prod_{i=1}^{n} x_i < c\}}$$

Ainsi, $\hat{c}_n = \max\limits_{i \in [\![0\,;\,n]\!]} X_i$

9. L'estimateur $\hat{c}_n$ est-il exhaustif ?

$$L(x_1, -, x_n, c) = \underbrace{\frac{1}{c^{n/\theta}} 1_{\{\hat{c}_n \leq c\}}}_{\varphi(\hat{c}_n ; c)} \times \underbrace{\frac{1}{\theta^n} \prod_{i=1}^{n} x_i^{1/\theta - 1}}_{\psi(x_1, -, x_n)}$$

Donc $\hat{c}_n$ est exhaustif.

10. L'estimateur c_n est-il sans biais ?

Soit $x \in [0, c[$,

$$\mathbb{P}\left(\max_{i \in [\![1,n]\!]} X_i \leq x\right) = \prod_{i=1}^{n} \mathbb{P}(X_i \leq x) \overset{\text{Va iid}}{=} \mathbb{P}(X_1 \leq x)^n$$

$$\mathbb{P}(X_1 \leq x) = \int_0^x \frac{t^{1/\theta - 1}}{\theta c^{1/\theta}} dt = \frac{1}{c^{1/\theta}} \left[t^{1/\theta} \right]_0^x = \left(\frac{x}{c}\right)^{1/\theta}$$

$$\Rightarrow \underbrace{\mathbb{P}\left(\max_{i \in [\![1,n]\!]} X_i \leq x\right)}_{= F_{\max_{i \in [\![1,n]\!]} X_i}(x)} = \left(\frac{x}{c}\right)^{n/\theta} \Longrightarrow \boxed{f_{\max_{i \in [\![1,n]\!]} X_i}(x) = \frac{n}{\theta c^{n/\theta}} x^{\frac{n}{\theta} - 1}}$$

$$\Rightarrow \mathbb{E}[\hat{c}_n] = \int x f_{\hat{c}_n}(x) dx = \frac{n}{\theta c^{n/\theta}} \int_0^c x^{n/\theta} dx = \frac{n}{\theta c^{n/\theta}} \frac{1}{\frac{1}{\theta} + 1} c^{\frac{n}{\theta} + 1}$$

$$= c \frac{n/\theta}{1 + n/\theta} \xrightarrow[\infty]{} c$$

Donc $\hat{c}_n$ est asympt sans biais.

11. Calculer l'estimateur $\hat{c}_M$ par la méthode des moments.

- Méthode des moments :

Si θ est de dimens° p, $\mathbb{E}[X_i^k] = M_k(\theta)$

et $\frac{1}{m}\sum_{i=1}^{m} x_i^k = m_k$

$\hat{\theta}_{MM}$ est tq $\begin{cases} M_1(\hat{\theta}_{MM}) = m_1 \\ \vdots \\ M_p(\hat{\theta}_{MM}) = m_p \end{cases}$

On a besoin d'un seul moment

Méthode des moments : $\mathbb{E}[X] = M_1(c)$ et $\bar{X}_m = m_1$

$\hat{c}_M$ est tq $M_1(\hat{c}_M) = m_1$

$$\Rightarrow \mathbb{E}[X] = \int_0^c x \frac{x^{1/\theta - 1}}{\theta c^{1/\theta}} dx = \int_0^c \frac{x^{1/\theta}}{\theta c^{1/\theta}} dx$$

$$= \frac{1}{\theta c^{1/\theta}} \times \frac{1}{\frac{1}{\theta}+1} c^{1/\theta + 1} = \frac{c}{1+\theta}$$

$$\underset{\text{Méthode des moments}}{\Longrightarrow} \frac{\hat{c}_M}{1+\theta} = \bar{X}_m \Longrightarrow \boxed{\hat{c}_M = (1+\theta)\bar{X}_m}$$

12. L'estimateur $\hat{c}_M$ est-il sans biais ? Convergent ?

$$\mathbb{E}[\hat{c}_M] = (1+\theta)\,\mathbb{E}[\bar{X}_m] = c$$

$$\mathbb{E}[\bar{X}_m] = \mathbb{E}\left[\frac{1}{m}\sum X_i\right] = \frac{1}{m}\sum \mathbb{E}[X_i] = \frac{1}{m} \times m\,\mathbb{E}[X_i] = \mathbb{E}[X_i] = \frac{c}{1+\theta}.$$

Donc l'estimateur $\hat{c}_M$ est sans biais

$$V(\hat{c}_M) = V((1+\theta)\bar{X}_m) = (1+\theta)^2 \times V(\bar{X}_m)$$

$$= (1+\theta)^2 \times \frac{1}{m^2}\sum_{i=1}^{m} V(X_i)$$

$$= (1+\theta)^2 \times \frac{1}{m} V(X)$$

$$= (1+\theta)^2 \times \frac{1}{m} \times \frac{\theta^2 c^2}{(1+2\theta)(1+\theta)^2}$$

$$= \frac{\theta^2 c^2}{(1+2\theta)m} \xrightarrow[m\to\infty]{} 0$$

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

$$\mathbb{E}[X^2] = \int_0^c x^2 \frac{x^{1/\theta - 1}}{\theta c^{1/\theta}} dx = \int_0^c \frac{x^{1/\theta + 1}}{\theta c^{1/\theta}} dx$$

$$= \frac{1}{\theta c^{1/\theta}} \frac{1}{\frac{1}{\theta}+2} c^{1/\theta + 2} = \frac{c^2}{1+2\theta}$$

$$\Rightarrow V(X) = \frac{c^2}{1+2\theta} - \frac{c^2}{(1+\theta)^2}$$

$$= \frac{c^2(1+\theta)^2 - c^2(1+2\theta)}{(1+2\theta)(1+\theta)^2} = \frac{\theta^2 c^2}{(1+2\theta)(1+\theta)^2}$$

Donc $\begin{cases} \mathbb{E}[\hat{c}_M] = c \\ V(\hat{c}_M) \xrightarrow[m\to\infty]{} 0 \end{cases} \Rightarrow$ L'estimateur $\hat{c}_M$ est convergent.

13. Montrer la normalité asymptotique de $\hat{c}_M$ et préciser sa variance asymptotique.

On a $\hat{C}_M = g(\bar{X}_m)$ et $c = g(\mathbb{E}[X])$ avec $g : x \mapsto (1+\theta)x$

Donc grâce à la méthode Delta,

$$\sqrt{m}\,(\hat{C}_M - c) \longrightarrow \mathcal{N}\left(0, \frac{\theta^2 c^2}{1+2\theta}\right)$$

14. On suppose maintenant que c et θ sont inconnus. Calculer l'estimateur du maximum de vraisemblance de (c, θ).

$\frac{x^{1/\theta - 1}}{\theta c^{1/\theta}} \mathbb{1}_{\{0 \leq x < c\}}$

Soit $\beta = (c, \theta)$

$$L(x_1, \ldots, x_m, \beta) = \prod_{i=1}^{m} \frac{x_i^{1/\theta - 1}}{\theta c^{1/\theta}} \mathbb{1}_{\{0 \leq x < c\}} = \frac{1}{\theta^m c^{m/\theta}} \prod_{i=1}^{m} x_i^{1/\theta - 1} \mathbb{1}_{\min X_i \geq 0} \mathbb{1}_{\max X_i < c}$$

β^* est EMV de (c, θ) si L est maximal en β^*

c.à.d lorsque $\beta^* = \left(\min_{i \in [\![1,m]\!]} X_i, \max_{i \in [\![1,m]\!]} X_i\right)$

Donc $\beta^* = (c^*, \theta^*) = \left(\min_{i \in [\![1,m]\!]} X_i, \max_{i \in [\![1,m]\!]} X_i\right).$