M1 Actuariat & Econométrie et Statistiques, année 2021–2022.

Numéro copie (obligatoire) :

Statistique Inférentielle
Contrôle terminal

Jeudi 13 janvier

Durée 2h, documents, téléphone, calculatrice interdits

Veuillez soigner la présentation 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.

En page 17, vous trouverez un aide-mémoire qui reprend toutes les principales distributions ainsi que quelques définitions générales sur les tests.

Enfin, comme dirait Marilyn Monroe, actrice nord-américaine et philosophe, “N'arrête pas quand tu es fatigué, arrête quand tu as fini”. Traduction : même si à un moment vous vous dites que vous ne savez plus rien faire et bien on essaye encore jusqu'à la fin.

Le barême (indicatif) prévu est le suivant : 9-6-7

Rappel :

- Une loi exponentielle de paramètre $\lambda >$ est une loi Gamma$(1, \lambda)$
- Une loi du χ^2_n à n degrés de liberté est une loi Gamma $(n/2, 1/2)$
- Si Z suit une loi Gamma (r, λ) alors $Y = 2\lambda Z$ suit une loi du χ^2_{2r}.

Exercice 1 : Booba : "Tout le monde peut s'en sortir, aucune cité n'a de barreaux."

1. Pour $p > 0$, on considère une variable aléatoire de fonction de répartition :

$$F_p(x) = \begin{cases} x^p & , 0 < x < 1, \\ 1 & , x \geq 1 \\ 0 & , x \leq 0 \end{cases}$$

Est-ce que F_p est une fonction de répartition ? Si oui, c'est très bien, si non, rajouter la constante pour que ce soit une fonction de répartition.

Si $F_p(x)$ fdr $\Rightarrow \lim_{x \to \infty} F_p(x) = 1$

De plus on prolonge par continuité F_p tq $F_p(0) = 0$ et $F_p(1) = 1$

$\Rightarrow$ OK

2. Soit X une variable aléatoire de fonction de répartition F_p, calculer $\mathbb{E}(X)$.

$$\mathbb{E}[X] = \int_0^1 x\, p x^{p-1} dx = \int_0^1 p x^{p} = \frac{p}{p+1}$$

$$p = 1 \implies f_p(x) = 1$$

$$\mathbb{E}[X]$$

3. Calculer la loi de $-\ln X$.

$$f_p(x) = \begin{cases} p x^{p-1} & \text{si } 0 < x < 1 \\ 0 & \text{sinon} \end{cases}$$

$x \in]0,1[$
$\implies e^{-x} \in]e^{-1}, 1[$

$$F_{-\ln(X)}(x) = \mathbb{P}(-\ln X \leq x) = \mathbb{P}(\ln X \geq -x)$$
$$= \mathbb{P}(X \geq e^{-x}) = 1 - \mathbb{P}(X \leq e^{-x})$$
$$= 1 - e^{-px} \quad \text{car } 0 < e^{-x} < 1$$

$$\implies f_{-\ln X}(x) = p e^{-px} \mathbb{1}_{\mathbb{R}_+}(x)$$

$$\implies -\ln X \sim \mathcal{E}(p)$$

4. X est la variable aléatoire qui désigne le taux de défaut des cleints d'une entreprise de crédit sur une durée d'un mois, on suppose que sa fonction de répartition est F_p. Comment peut-on interpréter le paramètre p en terme de risque de défaut ? Comment interpréter le fait que $p = 1$?

$p \Rightarrow$ rapidité d'apparit° de défaut

si $p = 1$ linéaire

5. Déterminer l'estimateur $\hat{p}_n$ du maximum de vraisemblance du paramètre p.

$L = p^n \prod_{i=1}^{n} x_i^{p-1} \, \mathbb{1}_{\min x_i \geq 0} \, \mathbb{1}_{\max x_i \leq 1}$ Les hyp sont ok

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

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

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

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

$\Rightarrow \hat{p}_n = \frac{-n}{\sum \ln(x_i)}$ est l'EMV de p

6. Calculer le biais de $\hat{p}_n$. $\hat{p}_n$ est-il efficace ? Exhaustif ?

$$B_p(\hat{p}_n) = \mathbb{E}[\hat{p}_n] - p = \mathbb{E}\left[\frac{-n}{\sum \ln(X_i)}\right] - p$$

$$= n\,\mathbb{E}\left[\frac{1}{\sum -\ln(X_i)}\right] - p = n\,\mathbb{E}\left[\frac{1}{Y}\right] - p = n \times \frac{p}{n-1} - p$$

$$Y = \sum -\ln(X_i) \sim \Gamma(n, p) \qquad = \frac{p}{n-1} \xrightarrow[\infty]{} 0$$

$\Rightarrow$ asympt SB

$$V(\hat{p}_n) = V\left(\frac{n}{\sum -\ln(X_i)}\right) = n^2\, V\left(\frac{1}{\Gamma(n,p)}\right) = n^2 \times \frac{p^2}{(n-1)^2(n-2)}$$

$$g : x \mapsto \frac{nx}{n-1} \Rightarrow g'(x) = \frac{n}{n-1} \qquad \Rightarrow \frac{(g'(p))^2}{I_n(p)} = \frac{n^2}{(n-1)^2} \times \frac{p^2}{n} = \frac{np^2}{(n-1)^2}$$

$$I_n(p) = -\mathbb{E}\left[\frac{\partial^2}{\partial p^2} \ell L\right] = \frac{n}{p^2}$$

Or $V(\hat{p}_n) = \dfrac{n^2 p^2}{(n-1)^2(n-2)} \neq \dfrac{np^2}{(n-1)^2}$ $\Rightarrow$ non efficace.

$$L = p^n \prod x_i^{p-1} \quad \text{et } \hat{p}_n = \frac{-n}{\sum \ln X_i} \Rightarrow \prod x_i = \exp\left(-\frac{n}{\hat{p}_n}\right)$$

$$\Rightarrow L(x_1, \dots, x_n, p) = \underbrace{p^n \exp\left(-\frac{n}{\hat{p}_n} \times (p-1)\right)}_{\varphi(p, \hat{p}_n)} \quad \underbrace{1}_{\psi(x_1, \dots, x_n)}$$

$\Rightarrow$ exhaustif.

7. Construire un intervalle de confiance bilatéral de niveau $1-\alpha$ pour p à l'aide d'un quantile d'une loi du χ^2.

$\hat{p}_n = \frac{-n}{\sum \ln(X_i)}$

On a mq $-\ln(X) \sim \mathcal{E}(p) \Rightarrow \sum -\ln(X_i) \sim \Gamma(n,p)$

$\Rightarrow \frac{1}{\sum -\ln(X_i)}$ suit une loi inverse gamma de paramètre n et p.

Donc $2p \underbrace{\sum -\ln(X_i)}_{\Gamma(n,\theta)} \sim \chi^2_{2n}$

$\Rightarrow \frac{2np}{\hat{p}_n} = 2p\sum -\ln(X_i)$ Donc $\frac{2np}{\hat{p}_n} \sim \chi^2_{2n}$ On note $Z \sim \chi^2_{2n}$

$\Rightarrow \mathbb{P}\left(-Z_{1-\alpha/2} \leq \frac{2np}{\hat{p}_n} \leq Z_{1-\alpha/2}\right) = \alpha$

$\Rightarrow \mathbb{P}\left(\underbrace{-\frac{\hat{p}_n Z_{1-\alpha/2}}{2n}}_{A_n} \leq p \leq \underbrace{\frac{\hat{p}_n Z_{1-\alpha/2}}{2n}}_{B_n}\right) = \alpha$

Donc l'IC est $\left[-\frac{\hat{p}_n Z_{1-\alpha/2}}{2n} ; \frac{\hat{p}_n Z_{1-\alpha/2}}{2n}\right]$

8. Montrer la normalité de $\hat{p}_n$.

Les hypothèses H1 à H7 sont vérifiées

Donc on a $\sqrt{n}(\hat{p}_n - p) \longrightarrow \mathcal{N}(0, I^{-1}(p))$

Où $I(p) = \frac{1}{n} I_n(p) = \frac{1}{n} - \mathbb{E}\left[\frac{\partial^2}{\partial p^2} \ln L\right] = \frac{1}{n} \times \frac{n}{p^2} \Rightarrow I^{-1}(p) = p^2$

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

9. Déterminer une fonction g telle que ? pas de sens car $\sqrt{n}(\hat{p}_n - p) \xrightarrow{\mathcal{L}} \mathcal{N}(0, p^2)$

$$\sqrt{n}(\hat{p}_n - p) \to \mathcal{N}(0, 1)$$

10. Construire deux intervalles de confiance asymptotique en se basant sur les 2 dernières questions.

$$L = p^n \prod_{i=1}^n x_i^{p-1} \mathbb{1}_{\min X_i \geq 0} \mathbb{1}_{\max X_i \leq 1}$$

11. Soit $p_0 < p_1$. Déterminer un test plus puissant pour

$$H_0 \ : \ p = p_0, \qquad \text{contre} \qquad H_1 \ : \ p = p_1.$$

$$V_{p_1,p_0} = \frac{L(x_1,\dots,x_n,p_1)}{L(x_1,\dots,x_n,p_0)} = \frac{p_1^n \prod_{i=1}^n x_i^{p_1-1}}{p_0^n \prod_{i=1}^n x_i^{p_0-1}} = \left(\frac{p_1}{p_0}\right)^n \prod_{i=1}^n x_i^{p_1-p_0} = \left(\frac{p_1}{p_0}\right)^n e^{(p_0-p_1)\sum_{i=1}^n -\ln(X_i)}$$

Or On sait que $2pZ \sim \chi^2_{2n}$ avec $Z \sim \Gamma(n,p)$

et on a mq $-\ln X \sim \mathcal{E}(p) \Rightarrow \sum -\ln X_i \sim \Gamma(n,p)$

On rejette Ho si $\left(\frac{p_1}{p_0}\right)^n e^{(p_0-p_1)\sum -\ln(X_i)} \geq k_\alpha$ avec k_α tq $\mathbb{P}_{H_0}\left(\left(\frac{p_1}{p_0}\right)^n e^{(p_0-p_1)\sum -\ln X_i} \geq k_\alpha\right) = \alpha$

Sous Ho, $2p_0 \sum -\ln X_i \sim \chi^2_{2n}$

$$\Rightarrow \mathbb{P}_{H_0}\left(\left(\frac{p_1}{p_0}\right)^n e^{\frac{p_0-p_1}{2p_0} 2p_0 \sum -\ln X_i} \geq k_\alpha\right) = \alpha$$

$$\left(\frac{p_1}{p_0}\right)^n e^{\frac{1}{2}(1-\frac{p_1}{p_0})Y} \geq k_\alpha$$

$$Y \leq \frac{\ln\left(\left(\frac{p_1}{p_0}\right)^{-n} k_\alpha\right)}{\frac{1}{2}(1-\frac{p_1}{p_0})}$$

$$\Rightarrow \mathbb{P}_{H_0}\left(\underbrace{2p_0 \sum -\ln X_i}_{\chi^2_{2n}} \leq \underbrace{\frac{2\ln\left(\left(\frac{p_1}{p_0}\right)^{-n} k_\alpha\right)}{\frac{1}{2}(1-\frac{p_1}{p_0})}}_{\chi^2_{2n,1-\alpha}}\right) = \alpha$$

Donc rejette Ho si $2p_0 \sum_{i=1}^n -\ln(X_i) \leq \chi^2_{2n,1-\alpha}$

12. Déterminer un test uniformément plus puissant pour les hypothèses

$$H_0 \ : \ p = 1, \qquad \text{contre} \qquad H_1 \ : \ p > 1.$$

$$V_{p,1} = p^n \prod_{i=1}^n x_i^{p-1} = p^n e^{(1-p)\sum_{i=1}^n -\ln(X_i)} = p^n e^{\frac{(1-p)t}{2}}$$

$\searrow$ en t car $p>1$

Soit $T_n = 2\sum_{i=1}^n -\ln(X_i)$ qui est une stat exhaustive

Donc on rejette Ho si $T_n \leq k_\alpha$ avec k_α tq $\mathbb{P}_{H_0}(T_n \leq k_\alpha) = \alpha$

$\Leftrightarrow V_{p,1} \geq \tilde{k}_\alpha$

Or Sous Ho, $T_n \sim \chi^2_{2n}$

$$\Rightarrow \mathbb{P}_{H_0}(T_n \leq \underbrace{k_\alpha}_{= \chi^2_{2n,\alpha}}) = \alpha$$

Donc On rejette Ho si $2\sum_{i=1}^n -\ln(X_i) \leq \chi^2_{2n,\alpha}$

Exercice 2 : Marcel Proust "*Il n'y a pas de réussites faciles ni d'échecs définitifs.*"

On considère $X_1, ..., X_n$ des v.a.i.i.d. telles que $X_i \sim \mathcal{N}(m, \sigma^2)$. On se propose de tester $H_0 : m = m_0$ contre $H_1 : m \neq m_0$. Dans ce cas là, on ne peut pas utiliser un test de Neyman-Pearson ou se variante pour les tests composites. Spoiler : On va donc faire un test du rapport du maximum de vraisemblance (cf rappel sur les tests 18).

1. En se basant sur l'écriture du test page 18, déterminer les espaces Θ_0 et Θ_1.

$\Theta_0 = \{m_0\} \times \mathbb{R}_+^*$ et $\Theta_1 = \mathbb{R} \setminus \{m_0\} \times \mathbb{R}_+^*$

$= (m_0, \sigma^2)$ $\sigma^2 \in \mathbb{R}_+^*$ $\quad = (m, \sigma^2)$ $m \in \mathbb{R} \setminus \{m_0\}$ $\sigma^2 \in \mathbb{R}_+^*$

2. Déterminer l'estimateur du maximum de vraisemblance $(\hat{m}, \hat{\sigma^2})$ sur Θ.

$X_i \sim \mathcal{N}(m, \sigma^2)$ Les hypothèses sont ok

$\Rightarrow L = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(x_i - m)^2\right) = (2\pi)^{-n/2} (\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - m)^2\right)$

$\Rightarrow \ln L = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - m)^2$

$\frac{\partial \ln L}{\partial m} = \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - m)$ $\Rightarrow \hat{m}$ est tq $\frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \hat{m}) = 0 \Rightarrow \hat{m} = \overline{X_n}$

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

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

$\Rightarrow \hat{\sigma}^2 = \frac{1}{n} \sum_{1}^{n} (x_i - m)^2$

$\frac{\partial \ln L}{\partial m \partial \sigma^2} = -\frac{\sum_{i=1}^{n} (x_i - m)}{(\sigma^2)^2}$ en $\hat{m}, \hat{\sigma}^2 = 0$

$\Rightarrow (\hat{m}, \hat{\sigma}^2)$ est bien l'EMV sur Θ

3. Déterminer l'estimateur du maximum de vraisemblance $\hat{\sigma}_0^2$ sur Θ_0.

On se place sur $\Theta_0 = (m_0, \sigma^2)$ $\sigma^2 \in \mathbb{R}_+^*$ Donc m_0 est fixé et connu

On cherche l'EMV $\hat{\sigma}_0^2$ de σ^2 sur Θ_0

$$L = (2\pi)^{-n/2} (\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - m_0)^2\right)$$

$$\Rightarrow \ln L = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum(x_i - m_0)^2$$

$$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_1^n (x_i - m_0)^2$$

$\Rightarrow \hat{\sigma}_0^2$ est tq $\dfrac{-n}{2\hat{\sigma}_0^2} + \dfrac{1}{2(\hat{\sigma}_0^2)^2}\sum(x_i - m_0)^2 = 0$

$$\Rightarrow \hat{\sigma}_0^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - m_0)^2$$

$$= \overline{X_n^2} - 2m_0\overline{X}_n + m_0^2$$

$$\frac{\partial^2 \ln L}{\partial(\sigma^2)^2} = \frac{n}{2(\sigma^2)^2} - \frac{1}{(\sigma^2)^3}\sum_1^n (x_i - m_0)^2$$

en $\hat{\sigma}_0^2$, $\dfrac{\partial^2 \ln L}{\partial(\sigma^2)^2} = \dfrac{n}{2(\sigma^2)^2} - \dfrac{n}{(\sigma^2)^2} = \dfrac{-n}{2(\sigma^2)^2} < 0$ donc OK

$$\hat{\sigma}_0^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - m_0)^2$$

4. Calculer $\lambda(x_1, ..., x_n)$.

$$\lambda(x_1, \dots, x_n) = \frac{\sup_{\theta \in \Theta} L(x,\theta)}{\sup_{\theta \in \Theta_0} L(x,\theta)} = \frac{L(x_1, \dots, x_n, \hat{m}, \hat{\sigma}^2)}{L(x_1, \dots, x_n, m_0, \hat{\sigma}_0^2)} = \frac{(2\pi)^{-n/2}(\hat{\sigma}^2)^{-n/2}\exp\left(-\frac{1}{2\hat{\sigma}^2} \times n\hat{\sigma}^2\right)}{(2\pi)^{-n/2}(\hat{\sigma}_0^2)^{-n/2}\exp\left(-\frac{1}{2\hat{\sigma}_0^2} \times n\hat{\sigma}_0^2\right)}$$

$$= \left(\frac{\hat{\sigma}^2}{\hat{\sigma}_0^2}\right)^{-n/2}$$

Or $\hat{\sigma}^2 = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(x_i - \bar{x})^2$

et $\hat{\sigma}_0^2 = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(x_i - m_0)^2$

$$\hat{\sigma}_0^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x} + \bar{x} - m_0)^2 = \frac{1}{n}\sum_{i=1}^{n}\left((x_i - \bar{x})^2 + 2(x_i - \bar{x})(\bar{x} - m_0) + (\bar{x} - m_0)^2\right)$$

$$= \underbrace{\frac{1}{n}\sum(x_i - \bar{x})^2}_{\hat{\sigma}^2} + \frac{1}{n}(\bar{x} - m_0)\underbrace{\sum_1^n(x_i - \bar{x})}_{=0} + (\bar{x} - m_0)^2$$

$$\Rightarrow \hat{\sigma}_0^2 = \hat{\sigma}^2 + (\bar{x} - m_0)^2$$

$$\Rightarrow \lambda(x_1, \dots, x_n) = \left(\frac{\hat{\sigma}^2}{\hat{\sigma}_0^2}\right)^{-n/2}$$

$$= \left(1 + \left(\frac{\bar{x} - m_0}{\hat{\sigma}}\right)^2\right)^{n/2}$$

5. En posant, $T_n = \sqrt{n-1}\frac{\bar{X}-m_0}{\hat{\sigma}}$, montrer que $\lambda(x_1, ..., x_n)$ est une fonction croissante en $\|T_n\|$.

On pose $T_n = \sqrt{n-1}\,\frac{\bar{X}-m_0}{\hat{\sigma}}$

$\lambda(x_1, \dots, x_n) = \left(1+\left(\frac{\bar{x}-m_0}{\hat{\sigma}}\right)^2\right)^{n/2} \Rightarrow \lambda(x_1, \dots, x_n) = \left(1+\frac{1}{n-1}T_n^2\right)^{n/2}$

Donc on a bien que $\lambda(x_1, \dots, x_n)$ est une fonct° $\nearrow$ en $\|T_n\|$

6. Quelle est la loi de T_n (avec un minimum d'explication) ?

$T_n = \sqrt{n-1}\,\frac{\bar{X}-m_0}{\hat{\sigma}}$

On a $\hat{\sigma}^2 \xrightarrow{P} \sigma^2$

Donc on peut confondre σ^2 et $\hat{\sigma}^2$ pour exprimer la loi de T_n

On a $\sqrt{n}\,\frac{\bar{X}_n - m_0}{\hat{\sigma}} \sim \mathcal{N}(0,1)$

On remplace $\hat{\sigma}^2$ par $S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$

Car comme les $X_i \sim \mathcal{N}(m, \sigma^2)$ on a $\hat{\sigma}^{-2}\sum(X_i - \bar{X}_n)^2 \sim \chi^2_{n-1}$

et $\bar{X}_n$ indep de $\sum(X_i - \bar{X}_n)^2$

Donc $\frac{\sqrt{n}\,\frac{\bar{X}_n - m_0}{\hat{\sigma}}}{\sqrt{\frac{\hat{\sigma}^{-2}}{n-1}\sum(X_i-\bar{X}_n)^2}} \sim t_{n-1}$

$\Rightarrow \sqrt{n}\,\frac{\bar{X}_n - m_0}{\hat{\sigma}} \sim t_{n-1}$

Rappel : Si
$Z \sim \mathcal{N}(0,1)$
$W \sim \chi^2_a$
$\Rightarrow \frac{Z}{\sqrt{W/a}} \sim t_a$

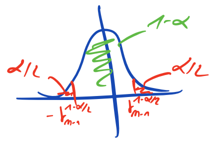
7. Donner la région de rejet du test de sensibilité α.

On a mq $\lambda(x_1, \dots, x_n)$ est une fonct° ↗ en $\|T_n\|$

Donc on rejette H_0 si $\{\lambda(x_1, \dots, x_n) > k_\alpha\} \Leftrightarrow \{\|T_n\| > \tilde{k}_\alpha\}$

Sous H_0 $(m = m_0)$, $T_n = \sqrt{n-1}\,\dfrac{\bar{X}_n - m_0}{\hat{\sigma}} \sim t_{n-1}$

Donc $\tilde{k}_\alpha$ est tq $\mathbb{P}_{H_0}(\|T_n\| > \tilde{k}_\alpha) = \alpha$ On note $T \sim t_{n-1}$



Donc on rejette H_0 si $\left|\sqrt{n-1}\,\dfrac{\bar{x}_n - m_0}{\hat{\sigma}}\right| > t_{n-1}^{1-\alpha/2}$

Donc $W = \left\{(x_1, \dots, x_n) \in \mathbb{R}^n,\ \left|\sqrt{n-1}\,\dfrac{\bar{x}_n - m_0}{\hat{\sigma}}\right| > t_{n-1}^{1-\alpha/2}\right\}$

région de rejet du test de sensibilité α

Exercice 3 : Winston Churchill : “Agissez toujours comme s’il était impossible d’échouer.”

On considère le modèle exponentiel translaté

$$\left(\mathbb{R}_+^*, \mathcal{B}_{\mathbb{R}_+^*}, \mathcal{P}_\theta, \theta = (\lambda, \beta) \in \mathbb{R}_+^* \times \mathbb{R}\right)$$

de densité

$$f(x, \theta) = \lambda \exp\left(-\lambda(x - \beta)\right) \mathbb{1}_{[\beta, \infty[}(x).$$

1. Calculer les moments d’ordre 1 et 2.

$$\mathbb{E}[X] = \int_\beta^\infty x \lambda \exp(-\lambda(x-\beta))\,dx = \left[-x\exp(-\lambda(x-\beta))\right]_\beta^\infty + \int_\beta^\infty \exp(-\lambda(x-\beta))\,dx$$

$$= \beta + \left[-\frac{1}{\lambda}\exp(-\lambda(x-\beta))\right]_\beta^\infty = \beta + \frac{1}{\lambda}$$

IPP
$u = x$ $u' = 1$
$v' = \lambda\exp(-\lambda(x-\beta))$ $v = -\exp(-\lambda(x-\beta))$

$$\mathbb{E}[X^2] = \int_\beta^\infty x^2 \lambda \exp(-\lambda(x-\beta))\,dx = \left[-x^2\exp(-\lambda(x-\beta))\right]_\beta^\infty + \int_\beta^\infty 2x\exp(-\lambda(x-\beta))\,dx$$

IPP
$u = x^2$ $u' = 2x$
$v' = \lambda\exp\ldots$ $v = -\exp\ldots$

$$= \beta^2 + \frac{2}{\lambda}\left(\beta + \frac{1}{\lambda}\right)$$

2. Déterminer l'estimateur $\hat{\theta}_n^{MM}$ par la méthode des moments.

Par la méthode des moments

$$\begin{cases} \hat{\beta}_{MM} + \dfrac{1}{\hat{\lambda}_{MM}} = \overline{X_n} \\ \hat{\beta}_{MM}^2 + \dfrac{2}{\hat{\lambda}_{MM}}\left(\hat{\beta}_{MM} + \dfrac{1}{\hat{\lambda}_{MM}}\right) = \overline{X_n^2} \end{cases} \Rightarrow \begin{cases} \hat{\lambda}_{MM} = \dfrac{1}{\sqrt{\overline{X_n^2} - \overline{X_n}^2}} \\ \hat{\beta}_{MM} = \overline{X_n} - \sqrt{\overline{X_n^2} - \overline{X_n}^2} \end{cases}$$

$$\hat{\beta}^2 = \left(\overline{X_n} - \frac{1}{\hat{\lambda}}\right)^2 = \overline{X_n}^2 - \frac{2}{\hat{\lambda}}\overline{X_n} + \frac{1}{\hat{\lambda}}^2$$

$$\overline{X_n}^2 - \frac{2}{\hat{\lambda}}\overline{X_n} + \frac{1}{\hat{\lambda}^2} + \frac{2\hat{\beta}}{\hat{\lambda}} + \frac{2}{\hat{\lambda}^2} = \overline{X_n^2}$$

$$-\cancel{\frac{2}{\lambda}}(\cancel{\beta} + \cancel{\tfrac{1}{\lambda}}) + \frac{1}{\lambda^2} + \cancel{\frac{2\beta}{\lambda}} + \cancel{\frac{2}{\lambda^2}} = \overline{X_n^2} - \overline{X_n}^2 \Rightarrow \hat{\lambda}_{MM} = \frac{1}{\sqrt{\overline{X_n^2} - \overline{X_n}^2}}$$

$$\hat{\beta} = \overline{X_n} - \sqrt{\overline{X_n^2} - \overline{X_n}^2}$$

3. Déterminer l'estimateur du maximum de vraisemblance $\hat{\theta}_n^{MV}$.

$$L = \prod_{i=1}^{n} \lambda e^{-\lambda(x_i - \beta)} \mathbb{1}_{[\beta, +\infty[}(x_i) = \lambda^n \exp\left(-\lambda \sum_{i=1}^{n}(x_i - \beta)\right) \mathbb{1}_{\min X_i \geq \beta}$$

On considère λ fixé et on obtient l'EMV $\hat{\beta}_{MV}$ en maximisant L

$$\Rightarrow \hat{\beta}_{MV} = \min_{i \in [\![1,n]\!]} X_i$$

Ensuite on considère $\hat{\beta}_{MV}$ fixé $\Rightarrow$ hypothèses vérifiées (Indep du support...)

$$\Rightarrow \ln L = n \ln \lambda - \lambda \sum_{1}^{n}(x_i - \beta)$$

$$\Rightarrow \frac{\partial \ln L}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n}(x_i - \beta) \Rightarrow \hat{\lambda}_{MV} \text{ est la } \frac{n}{\hat{\lambda}_{MV}} - \sum(x_i - \beta) = 0 \Rightarrow \hat{\lambda}_{MV} = \frac{n}{\sum(x_i - \beta)}$$

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

Donc $\hat{\theta}_{MV} = (\hat{\lambda}_{MV}, \hat{\beta}_{MV})$ est l'EMV de θ.

4. On note $\hat{\theta}_n^{MV} = (\hat{\lambda}_n, \hat{\beta}_n)$, montrer que $\hat{\beta} - \beta \sim \mathcal{E}(n\lambda)$

$$\hat{\beta}_{MV} = \min_{i \in [\![1,n]\!]} X_i \implies \hat{\beta}_{MV} - \beta = \min_{i \in [\![1,n]\!]} (X_i - \beta)$$

$$\mathbb{P}\left(\min_{i \in [\![1,n]\!]} (X_i - \beta) \leq x\right) = 1 - \mathbb{P}\left(\min (X_i - \beta) > x\right)$$
$$= 1 - \mathbb{P}(X_1 - \beta > x)^n = 1 - (1 - \mathbb{P}(X_1 - \beta \leq x))^n$$
$$= 1 - (1 - \mathbb{P}(X_1 \leq x + \beta))^n$$
$$= 1 - \left(1 - \int_\beta^{x+\beta} \lambda e^{-\lambda(x-\beta)} dx\right)^n$$
$$= 1 - (1 - (-e^{-\lambda x} + 1))^n$$
$$= 1 - (e^{-n\lambda x})$$

Donc la fdr de $\hat{\beta} - \beta$ est la fdr d'une $\mathcal{E}(n\lambda)$

Donc $\hat{\beta} - \beta \sim \mathcal{E}(n\lambda)$

5. Calculer le risque quadratique de $\hat{\beta}_n$ et montrer que $\sqrt{n}(\hat{\beta}_n - \beta) \xrightarrow{\mathbb{P}} 0$.

$$\mathcal{R}(\beta, \hat{\beta}_n) = B_\beta(\hat{\beta}_n)^2 + V(\hat{\beta}_n)$$

$$B_\beta(\hat{\beta}_n) = \mathbb{E}[\hat{\beta}_n] - \beta = \mathbb{E}[\hat{\beta}_n - \beta] = \frac{1}{n\lambda}$$

$$V(\hat{\beta}_n) = V(\hat{\beta}_n - \beta) = \frac{1}{(n\lambda)^2} \implies \mathcal{R}(\beta, \hat{\beta}_n) = \left(\frac{1}{n\lambda}\right)^2 + \frac{1}{(n\lambda)^2} = \frac{2}{(n\lambda)^2}$$

$\forall \varepsilon > 0,$

$$\mathbb{P}(|\sqrt{n}(\hat{\beta}_n - \beta)| \geq \varepsilon) = \mathbb{P}\left(|\hat{\beta} - \beta| \geq \frac{\varepsilon}{\sqrt{n}}\right) = \int_{\frac{\varepsilon}{\sqrt{n}}}^{\infty} n\lambda e^{-n\lambda x} dx = \left[-e^{-n\lambda x}\right]_{\frac{\varepsilon}{\sqrt{n}}}^{\infty}$$
$$= e^{-n\lambda \times \frac{\varepsilon}{\sqrt{n}}}$$

Car $\hat{\beta} - \beta \sim \mathcal{E}(n\lambda)$

$$\implies \mathbb{P}(|\sqrt{n}(\hat{\beta}_n - \beta)| \geq \varepsilon) = e^{-\sqrt{n}\lambda\varepsilon} \qquad \lambda > 0 \text{ et } \varepsilon \in \mathbb{R}_+^*$$

$$\implies \lim_{n \to \infty} \mathbb{P}(|\sqrt{n}(\hat{\beta}_n - \beta)| \geq \varepsilon) = 0$$

Donc $\sqrt{n}(\hat{\beta}_n - \beta) \xrightarrow{\mathbb{P}} 0$

6. En utilisant le Théorème Centrale Limite sur $(X_i - \beta)_{i\in\mathbb{N}}$, déterminer la loi asymptotique de $\overline{X} - \hat{\beta}_n$.

On applique le TCL à $(X_i - \beta)_{i\in\mathbb{N}}$

$$\Rightarrow \sqrt{n}\left(\overline{X}_n - \beta - \mathbb{E}[X_i - \beta]\right) \longrightarrow \mathcal{N}(0, \underbrace{V(X-\beta)}_{\lambda^{-2}})$$

$$\Rightarrow \sqrt{n}\left(\overline{X}_n - \beta - \frac{1}{\lambda}\right) \longrightarrow \mathcal{N}(0, \lambda^{-2})$$

Or $\sqrt{n}(\hat{\beta}_n - \beta) \xrightarrow{\mathbb{P}} 0$

$$\Rightarrow \sqrt{n}\left(\overline{X}_n - \hat{\beta}_n - \frac{1}{\lambda}\right) \longrightarrow \mathcal{N}(0, \lambda^{-2})$$

7. En déduire que $\sqrt{n}(\hat{\lambda} - \lambda) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \lambda^2)$.

$$\hat{\lambda} = \frac{n}{\sum(x_i - \beta)} = \frac{1}{\overline{X}_n - \beta}$$

Donc comme $\sqrt{n}\left(\overline{X}_n - \beta - \frac{1}{\lambda}\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \lambda^{-2})$

On applique la méthode Delta au TCL appliqué à $\overline{X} - \beta$ avec $g : x \mapsto \frac{1}{x}$

$g'(x) = -\frac{1}{x^2}$ $\quad g(\overline{X} - \beta) = \hat{\lambda}$ $\quad g\left(\frac{1}{\lambda}\right) = \lambda$

$$\Rightarrow \sqrt{n}\left(\underbrace{g(\overline{X}_n - \beta)}_{\hat{\lambda}_n} - \underbrace{g\left(\tfrac{1}{\lambda}\right)}_{\lambda}\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \lambda^{-2} \times \underbrace{\left(g'\left(\tfrac{1}{\lambda}\right)\right)^2}_{\lambda^4}\right)$$

$$\Rightarrow \sqrt{n}\left(\hat{\lambda} - \lambda\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \lambda^2)$$

8. Déterminer $(a, b) \in \mathbb{R}^2$ tels que $[\hat{\lambda}_n/a, \hat{\lambda}_n/b]$ soit un intervalle de confiance de niveau asymptotique $1-\alpha$ pour λ.

On a $\sqrt{n}(\hat{\lambda}-\lambda) \longrightarrow \mathcal{N}(0,\lambda^2) \Rightarrow \sqrt{n}\,\dfrac{(\hat{\lambda}-\lambda)}{\lambda} \longrightarrow \mathcal{N}(0,1)$

Or comme $\hat{\lambda} \xrightarrow{\mathbb{P}} \lambda$ car EMV on remplace le λ du bas par $\hat{\lambda}_n$

$$\Rightarrow \sqrt{n}\,\frac{(\hat{\lambda}_n-\lambda)}{\hat{\lambda}_n} \longrightarrow \mathcal{N}(0,1) \qquad Z \sim \mathcal{N}(0,1)$$

$$\Rightarrow \mathbb{P}\left(-Z_{1-\alpha/2} \leq \sqrt{n}\,\frac{(\hat{\lambda}_n-\lambda)}{\hat{\lambda}_n} \leq \underbrace{Z_{1-\frac{\alpha}{2}}}_{\text{quantile d'ordre } 1-\alpha/2 \text{ d'une } \mathcal{N}(0,1)}\right) \longrightarrow 1-\alpha$$

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

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

$$\Rightarrow \mathbb{P}\left(\underbrace{\hat{\lambda}_n\left(1-\frac{Z_{1-\alpha/2}}{\sqrt{n}}\right)}_{\frac{\hat{\lambda}_n}{a}} \leq \lambda \leq \underbrace{\hat{\lambda}_n\left(1+\frac{Z_{1-\alpha/2}}{\sqrt{n}}\right)}_{\frac{\hat{\lambda}_n}{b}}\right) \longrightarrow 1-\alpha$$

avec $\begin{cases} a = \dfrac{1}{1-\frac{Z_{1-\alpha/2}}{\sqrt{n}}} = \dfrac{\sqrt{n}}{\sqrt{n}-Z_{1-\alpha/2}} \\[2ex] b = \dfrac{1}{1+\frac{Z_{1-\alpha/2}}{\sqrt{n}}} = \dfrac{\sqrt{n}}{\sqrt{n}+Z_{1-\alpha/2}} \end{cases}$

Aide Mémoire

Principales lois de probabilité

- Loi Bernoulli, $Bern(p)$, $p \in [0,1]$

$$\mathbb{P}(X=1)=p, \quad \mathbb{P}(X=0)=1-p$$

- Loi Binomiale, $\mathcal{B}(n,p)$, $n \in \mathbb{N}^*$ et $p \in [0,1]$

$$\mathbb{P}(X=x)=\left(\begin{array}{c} x \\ n \end{array}\right) p^x(1-p)^{n-x}, \quad x \in \{0,...,n\}$$

- Loi Poisson, $\mathcal{P}(\lambda)$, $\lambda \in \mathbb{R}^{+*}$

$$\mathbb{P}(X=x)=\frac{\lambda^x}{x!}e^{-\lambda}, \quad x \in \mathbb{N}$$

- Loi Géométrique, $\mathcal{G}(p)$, $p \in [0,1]$

$$\mathbb{P}(X=x)=(1-p)^{x-1}p, \quad x \in \mathbb{N}^*$$

- Loi Normale, $\mathcal{N}(\mu,\sigma^2)$, $\mu \in \mathbb{R}$ et $\sigma^2 \in \mathbb{R}^{+*}$

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right), \quad x \in \mathbb{R}$$

- Loi Exponentielle, $\mathcal{E}(\lambda)$, $\lambda \in \mathbb{R}^{+*}$

$$f(x)=\lambda e^{-\lambda x}, \quad x \in \mathbb{R}^{+*}$$

- Loi Beta, $\mathcal{B}(\alpha,\beta)$, $\alpha>0$ et $\beta>0$

$$f(x)=\frac{\Gamma(\alpha+\beta}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}, \quad x \in [0,1]$$

- Loi Gamma, $\Gamma(\alpha,\theta)$, $\alpha>0$ et $\theta>0$

$$f(x)=\frac{x^{\alpha-1}\theta^\alpha}{\Gamma(\alpha)}e^{-\theta x}, \quad x>0$$

- Loi Student, $t(k)$, $k>0$

$$f(x)=\frac{1}{\sqrt{\pi k}}\frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)}\left(1-\frac{t^2}{k}\right)^{-(k+1)/2}, \quad x \in \mathbb{R}$$

- Loi du Khi-2, χ_n, $n>0$

$$f(x)=\frac{(1/2)^{n/2}}{\Gamma(n/2)}x^{k/2-1}e^{-x/2}, \quad x>0$$

Rappel sur les tests

- Test de Neyman-Pearson : $\Theta_0 = \{\theta_0\}$ contre $\Theta_1 = \{\theta_1\}$. Soit

  $$V_{\theta_0,\theta_1}(x) = \frac{L(x,\theta_1)}{L(x,\theta_0)},$$

  le test est de région de rejet :

  $$V_{\theta_0,\theta_1}(x) > k.$$

- Test d'une hypothèse composite : $\theta = \theta_0$ contre $\theta > \theta_0$.

  Si T est une statistique exhautive dont la fonction de vraissemblance $g(t,\theta)$ vérifie : pour $\theta > \theta_0$,

  $$\frac{g(t,\theta)}{g(t,\theta_0)} \text{ est une fonction croissante de } t.$$

  le test de zone de rejet $\widetilde{R_k} = \{T > k\}$ est u.p.p.

- Tests du rapport du maximum de vraisemblance : $H_0 \ : \ \theta \in \Theta_0$ contre $H_1 \ : \ \theta \in \Theta_1$.

  Région critique :

  $$W = \left\{ x \in \mathbb{R}^n \ / \ \frac{\sup_{\theta\in\Theta_1} L(x,\theta)}{\sup_{\theta\in\Theta_0} L(x,\theta)} > k_\alpha \right\},$$

  où k_α est choisit tel que $\sup\limits_{\theta\in\Theta_0} \mathbb{P}_\theta(W) = \alpha$.

  Si $\Theta_1 = \Theta_0^c$ alors

  $$W = \left\{ x \in \mathbb{R}^n \ / \ \lambda(x) = \frac{\sup_{\theta\in\Theta} L(x,\theta)}{\sup_{\theta\in\Theta_0} L(x,\theta)} > k_\alpha \right\}.$$