Master 1SPA
Faculté de maths

Examen Statistique inférentielle

Exercise 1 *On dit que X suit une loi Gamma de paramètres p et θ ($p > 0$,$\theta > 0$), notée $\gamma(p, \theta)$, si sa densité (par rapport à la mesure de Lebesgue) est :*

$$f(x) = \frac{\theta^p}{\Gamma(p)} \exp(-\theta x) x^{p-1} 1_{[0,+\infty[}(x)$$

ou de façon équivalente, si sa fonction caractéristique vaut $\varphi_X(t) = \frac{1}{\left(1-\frac{it}{\theta}\right)^p}$ pour tout réel t. On rappelle les propriétés suivantes de la fonction Gamma :

$$\forall \alpha > 0,\ \Gamma(\alpha) = \int_0^\infty x^{\alpha-1} \exp(-x) dx;\ \Gamma(\alpha+1) = \alpha\Gamma(\alpha);\ \Gamma(1/2) = \sqrt{\pi}$$

1. Vérifier que la loi $\gamma(p, \theta)$ est bien une loi de probabilité.

2. Calculer $E(X^k)$ pour $k \geq 1$. En déduire que $E(X) = \frac{p}{\theta}$ et $Var(X) = \frac{p}{\theta^2}$.

3. Si $a > 0$, montrer que $\frac{X}{a} \sim \gamma(p, a\theta)$.

4. Soient X et Y deux v.a. indépendantes de lois respectives $\gamma(p_1, \theta)$ et $\gamma(p_2, \theta)$. Montrer que $X + Y \sim \gamma(p_1 + p_2, \theta)$.

5. Si $X_1, ..., X_n$ sont n variables aléatoires indépendantes de même loi $\gamma(1, \theta)$ (dite loi exponentielle de paramètre θ), donner la loi de la somme $S_n = X_1 + ... + X_n$ et calculer $E(S_n)$ et $Var(S_n)$.

Exercise 2 *On observe $X_1, ..., X_n$ indépendantes et de même loi. On suppose qu'il existe $\theta > 0$ tel que cette loi admet la densité*

$$f(x) = \frac{1}{\sqrt{2\pi\theta^2}} \exp\left(\frac{-x^2}{2\theta^2}\right)$$

1. On veut estimer $\tau = \theta^2$. Proposer un estimateur $\widehat{\tau}$ de τ et étudier sa loi.

2. Construire un intervalle de confiance au niveau $1 - \alpha$ de la forme $[S_1, S_2]$ tel que $P(\tau < S_1) = P(\tau > S_2) = \alpha/2$. Lorsque $n = 10$, $\widehat{\tau} = 2$ et $\alpha = 0.05$, calculer l'intervalle de confiance obtenu.

3. Constuire un test de niveau de $H_0 : \tau = 3$ contre $H_1 : \tau > 3$. Pour un seuil de 5%, lorsque $n = 10$ et qu'on observe $\widehat{\tau} = 4$, rejette-t-on l'hypothèse nulle ?

Indication : $F^{-1}\chi^2_{10}(0.025) \simeq 3.25$, $F^{-1}\chi^2_{10}(0.975) \simeq 20.48$, et $\Phi^{-1}(0.975) \simeq 2$.

Exercise 3 *Soient X une v.a. admettant pour densité*

$$f(x) = \frac{1}{\theta} \exp\left(-\frac{(x-\theta)}{\theta}\right) 1_{[\theta,+\infty[}(x)$$

où $\theta > 0$ est inconnu, et $(X_1, ..., X_n)$ un échantillon de même loi que X.

1. Estimer θ par la méthode des moments.

2. Quelle est la loi limite de l'estimateur $\widehat{\theta}$ ainsi obtenu ? Etudier également sa convergence en moyenne quadratique.

Corrigé

Exo1

1. Comme f est positive, il suffit de vérifier que $\int_{\mathbb{R}} f(x)dx = 1$. Ceci s'obtient:

$$\begin{aligned}
\int_{\mathbb{R}} f(x)dx &= \int_0^\infty \frac{\theta^p}{\Gamma(p)} \exp(-\theta x) x^{p-1} dx \\
&= \int_0^\infty \frac{(\theta x)^{p-1}}{\Gamma(p)} \exp(-\theta x)\theta dx \qquad \text{on pose } y = \theta x \text{ et donc } dy = \theta dx \\
&= \frac{1}{\Gamma(p)} \int_0^\infty y^{p-1} \exp(-y) dy = 1
\end{aligned}$$

2. De la même façon, on a

$$\begin{aligned}
E(X^k) &= \int_{\mathbb{R}} x^k f(x)dx = \int_0^\infty x^k \frac{\theta^p}{\Gamma(p)} \exp(-\theta x) x^{p-1} dx \\
&= \int_0^\infty \frac{(\theta x)^{k+p-1}}{\theta^k \Gamma(p)} \exp(-\theta x)\theta dx = \frac{1}{\theta^k \Gamma(p} \int_0^\infty y^{k+p-1} \exp(-y) dy \\
&= \frac{\Gamma(k+p)}{\theta^k \Gamma(p)}
\end{aligned}$$

Ainsi, on a

$$E(X) = \tfrac{\Gamma(1+p)}{\theta\Gamma(p)} = \tfrac{p}{\theta} \quad \text{et} \quad E(X^2) = \tfrac{\Gamma(2+p)}{\theta^2\Gamma(p)} = \tfrac{p(p+1)}{\theta^2}$$
$$Var(X) = E(X^2) - (EX)^2 = \tfrac{p(p+1)}{\theta^2} - \left(\tfrac{p}{\theta}\right)^2 = \tfrac{p}{\theta^2}$$

3. pour $a > 0$

$$P(\tfrac{X}{a} < y) = P(X < ay) = F_X(ay)$$

$\Longrightarrow$

$$\begin{aligned}
f_{\frac{X}{a}}(y) &= a f_X(ay) = a\frac{\theta^p}{\Gamma(p)} \exp(-\theta ay)(ay)^{p-1} 1_{[0,+\infty[}(y) \\
&= \frac{(a\theta)^p}{\Gamma(p)} \exp(-\theta ay) y^{p-1} 1_{[0,+\infty[}(y)
\end{aligned}$$

$\Longrightarrow$

$$\frac{X}{a} \sim \gamma(p, a\theta)$$

4. Pour déterminer la loi de $X+Y$, on calcule sa fonction caractéristique $\varphi_{X+Y}(t) = E[e^{it(X+Y)}]$. Grâce à l'indépendance, cette fonction est égale au produit des fonctions caractéristiques de X et de Y :

$$\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t) = \frac{1}{(1-\frac{it}{\theta})^{p_1}} \times \frac{1}{(1-\frac{it}{\theta})^{p_2}} = \frac{1}{(1-\frac{it}{\theta})^{p_1+p_2}}$$

Ainsi, $X+Y$ est la fonction caractéristique de la loi $\gamma(p_1+p_2,\theta)$.

6. D'après la question précédente (et par une récurrence triviale),

$$S_n \sim \gamma(n,\theta).$$

En utilisant la question 2, nous avons donc $E(S_n) = \frac{n}{\theta}$, $Var(S_n) = \frac{n}{\theta^2}$.

Exo2

1. Les X_i suivent une loi normale $N(0,\theta^2)$. On sait donc que $Var(X_i) = E(X_i^2) = \theta^2 = \tau$. Pour estimer τ, on propose l'estimateur suivant : $\widehat{\tau} = \frac{1}{n}\sum_{i=1}^{n} X_i^2$. On a alors $\frac{n\widehat{\tau}}{\tau} = \sum_{i=1}^{n}\left(\frac{X_i}{\theta}\right)^2$. Or comme pour tout i, $\frac{X_i}{\theta} \sim N(0,1)$ et comme les v.a. X_i sont i.i.d., on en déduit que $\frac{n\widehat{\tau}}{\tau} \sim \chi^2(n)$. Cette caractérisation de la loi de $\widehat{\tau}$ suffit, cependant on peut remarquer qu'une loi du $\chi^2(n)$ est aussi une loi $\gamma\left(\frac{n}{2},\frac{1}{2}\right)$. Ainsi,

$$\widehat{\tau} \sim \gamma\left(\frac{n}{2},\frac{1}{2\tau}\right)$$

2. On note $F_{\chi^2(n)}$ la f.d.r. de la loi du $\chi^2(n)$ et $F^{-1}_{\chi^2(n)}$la fonction quantile (qui est ici l'inverse de $F_{\chi^2(n)}$ car cette dernière est une bijection). On peut alors écrire :

$$P\left(F^{-1}_{\chi^2(n)}(\alpha/2) \le \frac{n\widehat{\tau}}{\tau} \le F^{-1}_{\chi^2(n)}(1-\alpha/2)\right) = 1-\alpha$$

donc

$$P\left(\frac{n\widehat{\tau}}{F^{-1}_{\chi^2(n)}(1-\alpha/2)} \le \tau \le \frac{n\widehat{\tau}}{F^{-1}_{\chi^2(n)}(\alpha/2)}\right) = 1-\alpha$$

En posant $S_1 = \frac{n\widehat{\tau}}{F^{-1}_{\chi^2(n)}(1-\alpha/2)}$ et $S_2 = \frac{n\widehat{\tau}}{F^{-1}_{\chi^2(n)}(\alpha/2)}$, on en déduit que

$$I_\alpha = [0.97, 6.16]$$

3. Le test rejette H_0 si $\widehat{\tau} > 3n^{-1}F^{-1}_{\chi^2(n)}(1-\alpha/2)$ au niveau α. Pour $n=10$, et $\alpha = 5\%$, nous avons $F^{-1}_{\chi^2(n)}(1-\alpha/2) \simeq 20.48$. Ainsi H_0 est acceptée lorsque $\widehat{\tau} = 4$.

Exo3

1-

$$EX = \int_{\mathbb{R}} xf(x)dx = \int_{\theta}^{\infty} x\frac{1}{\theta}\exp(-\frac{x-\theta}{\theta})dx = \theta\int_{0}^{\infty}(y+1)\exp(-y)dy = 2\theta$$

en posant

$$x-\theta=\theta y \Longrightarrow dx=\theta dy \text{ et } x=\theta(y+1) \text{ et } \int\limits_0^\infty (y+1)\exp(-y)dy=2$$

Ceci nous permet de proposer l'estimateur de θ suivant:

$$\widehat{\theta}=\frac{\overline{X}}{2}$$

2-

$var X=\int\limits_\theta^\infty x^2 \frac{1}{\theta}\exp(-\frac{x-\theta}{\theta})dx-4\theta^2=\theta^2$ car $\int\limits_0^\infty (y+1)^2\exp(-y)dy=5$

D'après le TLC, on a

$$\sqrt{n}\left(\frac{\overline{X}-2\theta}{\theta}\right)\rightarrow\mathcal{N}(0,1)$$

$\Longrightarrow \sqrt{n}\left(\frac{\overline{X}-2\theta}{\theta}\right)=\sqrt{n}\left(\frac{2\widehat{\theta}-2\theta}{\theta}\right)=\frac{2}{\theta}\sqrt{n}(\widehat{\theta}-\theta)\rightarrow\mathcal{N}(0,1)$ on en déduit que

$$\sqrt{n}(\widehat{\theta}-\theta)\rightarrow\mathcal{N}(0,\frac{\theta^2}{4})$$

Le risque quadratique

on a $E\widehat{\theta}=\theta \Longrightarrow$

$$E(\widehat{\theta}-\theta)^2=var\widehat{\theta}=var\frac{\overline{X}}{2}$$

or $var\frac{\overline{X}}{2}=\frac{1}{4n}var X=\frac{\theta^2}{4n}$ d'où le risque quadratique tend vers 0 quand $n\rightarrow\infty$.