Maxime TCHIBOZO

Devoir Maison de Statistiques.

Exercice 1 : Exploration des données, recherche de la loi.

2. a) Soit $p \in]0,1[$.

Posons $y = F^{-1}(p,0,1)$ et $z = F^{-1}(p,\mu,\sigma^2)$

D'une part, $p = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$. D'autre part, $p = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$

$$\int_{-\infty}^{z} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} = \int_{-\infty}^{\frac{z-\mu}{\sqrt{\sigma^2}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{\tilde{t}^2}{2}} d\tilde{t}$$ par le changement de variable C^1, bijectif $\tilde{t} \leftarrow \frac{t-\mu}{\sqrt{\sigma^2}}$

$$= \int_{-\infty}^{\frac{z-\mu}{\sqrt{\sigma^2}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$ par le changement de variable C^1, bijectif $t \leftarrow \tilde{t}$

On a montré : $F(y) = F\left(\frac{z-\mu}{\sqrt{\sigma^2}}\right)$. Or F est injective. Donc

$y = \frac{z-\mu}{\sqrt{\sigma^2}}$, i.e. $F^{-1}(p,\mu,\sigma^2) = \mu + F^{-1}(p,0,1)$.

3. b) Posons $y = F^{-1}(p,\lambda)$ et $z = F^{-1}(p,1)$.

$p = \int_{-\infty}^{y} \lambda e^{-\lambda x} dx$ et $p = \int_{-\infty}^{z} e^{-\lambda x} dx$

$$\int_{-\infty}^{y} \lambda e^{-\lambda x} dx = \int_{-\infty}^{\lambda y} e^{-\hat{x}} d\hat{x}$$ par le changement de variable $\hat{x} \leftarrow \lambda x$

Par injectivité de F, on a : $\lambda y = z$, soit $F^{-1}(p,\lambda) = \frac{1}{\lambda} F^{-1}(p,1)$

Exercice 2 : Estimation ponctuelle des paramètres de loi exponentielle.

1. Supposons le modèle $\{P_\lambda, \lambda > 0\}$ régulier au sens du théorème de Cramer-Rao.

$\mathbb{E}[T_1] = \frac{1}{n} \times n\, \mathbb{E}[X_1] = \frac{1}{\lambda}$ donc T_1 est un estimateur sans biais.

D'après l'inégalité de Cauchy-Schwarz, $\quad Var(T_1(x)) \geq \dfrac{(g_1'(\lambda))^2}{I(\lambda)}$.

D'une part : $\quad Var(T_1(x)) = \frac{1}{n^2} \cdot n \cdot Var(X_1)$ car les X_i sont iid.

$$= \frac{1}{n\lambda^2}$$

D'autre part : $\quad \dfrac{(g_1'(\lambda))^2}{I(\lambda)} = \dfrac{\left(-\frac{1}{\lambda^2}\right)^2}{n \cdot Var(I_1(\lambda))}$ car les X_i sont iid.

$$I_1(\lambda) = Var_\lambda\left[\frac{\partial}{\partial\lambda}[\log(\lambda e^{-\lambda x})]\right]$$

$$= Var_\lambda\left[\frac{\partial}{\partial\lambda}[\log(\lambda) - \lambda x]\right]$$

$$= Var_\lambda\left[\frac{1}{\lambda} - x\right] = Var(-x) = \frac{1}{\lambda^2}$$

$$\frac{(g_1'(\lambda))^2}{I(\lambda)} = \frac{\frac{1}{\lambda^4}}{\frac{n}{\lambda^2}} = \frac{1}{n\lambda^2}$$

On a donc $\quad Var(T_1(x)) = \dfrac{(g_1'(\lambda))^2}{I(\theta)}$

Ce qui correspond au cas d'égalité de Cramer-Rao.

T_1 est donc un estimateur uniformément de variance minimale parmi les estimateurs sans biais de g_1.

2.2. Soit $\alpha > 0$. $\quad R(\lambda, \tilde{T}_{1,\alpha}) = Var(\tilde{T}_{1,\alpha}) + b(\lambda, \tilde{T}_{1,\alpha})^2$

$$= \alpha^2 Var(T_1) + \left(\mathbb{E}_\lambda[g_1(\lambda) - \tilde{T}_{1,\alpha}(x)]\right)^2$$

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

On a donc $R(\lambda, \tilde{T}_{1,\alpha}) < R(\lambda, T_1)$ quand

$$\alpha^2\left(\frac{1}{n\lambda^2}+\frac{1}{\lambda^2}\right) - \frac{2}{\lambda^2}\alpha + \frac{1}{\lambda^2} - \frac{1}{n\lambda^2} < 0$$

$$\Leftrightarrow \quad a\alpha^2 + b\alpha + c < 0 \quad \text{où} \quad \begin{cases} a = \frac{1}{\lambda^2}\left(1+\frac{1}{n}\right) \\ b = -\frac{2}{\lambda^2} \\ c = \frac{1}{\lambda^2}\left(1-\frac{1}{n}\right) \end{cases}$$

$$\Delta = b^2 - 4ac$$

$$= \frac{4}{\lambda^4} - \frac{4}{\lambda^4}\left(1+\frac{1}{n}\right)\left(1-\frac{1}{n}\right)$$

$$= \frac{4}{\lambda^2}\left[\frac{1}{n^2}\right] > 0.$$

$$\alpha_1 = \frac{\frac{2}{\lambda^2} - \frac{2}{n\lambda}}{\frac{2}{\lambda^2}\left(1+\frac{1}{n}\right)} \qquad \alpha_2 = \frac{\frac{2}{\lambda^2} + \frac{2}{n\lambda}}{\frac{2}{\lambda^2}\left(1+\frac{1}{n}\right)}$$

On a donc $R(\lambda, \tilde{T}_{1,\alpha}) < R(\lambda, T_1)$ pour $\alpha \in \left]\frac{\lambda(n-1)}{n+1}, \frac{\lambda(n+1)}{n+1}\right[$

Ce résultat n'est pas en contradiction avec 2.1. car $\tilde{T}_{1,\alpha}$ n'est pas un estimateur sans biais

3) Soit Φ telle que $\Phi(\underbrace{q_2(\lambda)}_{= \frac{\ln(2)}{\lambda}}) = E[X] = \frac{1}{\lambda}$.

$$\Phi^{-1}\left(\Phi(q_2(\lambda))\right) = \Phi^{-1}\left(\frac{1}{\lambda}\right)$$

$$\Leftrightarrow \quad \frac{\ln(2)}{\lambda} = \Phi^{-1}\left(\frac{1}{\lambda}\right)$$

D'où, $\Phi^{-1}(t) = \ln(2)\cdot t \quad \forall t > 0.$

On en déduit l'estimateur par la méthode des moments :

$$\Phi^{-1}\left(\frac{1}{n}\sum \Phi(x_i)\right) = \frac{\ln(2)}{n}\sum x_i$$

$$= \ln(2)\cdot T_1.$$

$T_2(X) = \ln(2) * T_1(X).$

Notre échantillon comporte $\widehat{55}$ valeurs.

D'après la question précédente, le risque de T_2 est inférieur à celui de T_1 si $\alpha \in \left]\lambda\frac{(n-1)}{(n+1)}, \lambda\right[$

$$\Leftrightarrow \alpha = \ln(2) \in \left]\lambda \cdot \frac{54}{56} ; \lambda\right[$$

$$\Leftrightarrow \ln(2) \in \left]\underset{\approx 0,03}{0,96} \cdot \lambda ; \lambda\right[.$$

Nos approximation expérimentale de ~~λ~~ λ sont 5 à 10 fois inférieurs à $\ln(2)$. On peut être confiant que le risque de T_2 n'est pas inférieur à celui de T_1.

Exercice 3 : Test sur le paramètre d'une loi.

1) Soit $T(X) = X_1 + \dots + X_n$

$$\Phi_T(t) = \Phi_{X_1+\dots+X_n}(t) = \prod_{i=1}^{n} \Phi_{X_i}(t) \quad \text{par indépendance des } X_i$$

$$= \frac{1}{\left(1 - it\frac{1}{\lambda}\right)^n}$$

Par identification, on en déduit $T \sim \text{Gamma}\left(n, \frac{1}{\lambda}\right)$

Soit $\mathcal{X}_1$ notre région de rejet.

Par définition du test au niveau α, on cherche δ tel que $\mathbb{P}_{\lambda_0}[T(X) \in \mathcal{X}_1] \leq \alpha$

i.e. $\mathbb{P}_{\lambda_0}[\text{le coût moyen d'un accident} > \$1\text{ Md}] \leq \alpha$.

$$\Leftrightarrow \mathbb{P}_{\lambda_0}[\delta(X) = 1] \leq \alpha$$

On en déduit la règle de décision $\delta(X) = \mathbb{1}_{\mathcal{X}_1}(T(X))$

Notre échantillon comporte $\underset{\sim}{\overset{n}{55}}$ valeurs.

D'après la question précédente. le risque de T_2 est inférieur à celui de T_1 si $\alpha \in \left]\lambda\frac{(n-1)}{(n+1)}, \lambda\right[$

$$\Leftrightarrow \alpha = \ln(2) \in \left]\lambda \cdot \frac{54}{56} ; \lambda\right[$$

$$\Leftrightarrow \underset{\approx 0{,}05}{\ln(2)} \in \left]0{,}96 \cdot \lambda ; \lambda\right[.$$

Nos ~~[illegible]~~ approximations expérimentales de ~~[illegible]~~ λ sont 5 à 10 fois inférieures à $\ln(2)$. On peut être confiant que le risque de T_2 n'est pas inférieur à celui de T_1.

Exercice 3 : Test sur le paramètre d'une loi.

1) Soit $T(X) = X_1 + \dots + X_n$

$$\Phi_T(t) = \Phi_{X_1+\dots+X_n}(t) = \prod_{i=1}^{n} \Phi_{X_i}(t) \quad \text{par indépendance des } X_i$$

$$= \frac{1}{(1 - it\frac{1}{\lambda})^n}$$

Par identification, on en déduit $T \sim \text{Gamma}(n, \frac{1}{\lambda})$

Soit $\mathcal{X}_1$ notre région de rejet.

Par définition du test au niveau α, on cherche δ tel que : $\mathbb{P}_{\lambda_0}[T(X) \in \mathcal{X}_1] \leq \alpha$

i.e. $\mathbb{P}_{\lambda_0}[$ le coût moyen d'un accident $> \$1\ Md] \leq \alpha$.

$$\Leftrightarrow \mathbb{P}_{\lambda_0}[\delta(X) = 1] \leq \alpha$$

On en déduit la règle de décision $\delta(X) = \mathbb{1}_{\mathcal{X}_1}(T(x))$