

Partie 1

1) La densité d'une loi de Pareto de paramètres a et λ est définie par :

$$f(x) = \frac{dF(x)}{dx} = \begin{cases} \lambda \frac{a^\lambda}{x^{\lambda+1}} & \text{si } x > a \\ 0 & \text{sinon} \end{cases}$$

La vraisemblance des observations x_1, \dots, x_n s'en déduit aisément (on rappelle qu'un échantillon (X_1, \dots, X_n) est une suite de n variables aléatoires indépendantes et de même loi) :

$$\begin{aligned} f(x_1, \dots, x_n; \lambda) &= \prod_{i=1}^n f(x_i) \\ &= \prod_{i=1}^n \frac{\lambda a^\lambda}{x_i^{\lambda+1}} I_{]a, +\infty[}(x_i) \\ &= \frac{(\lambda a^\lambda)^n}{(\prod_{i=1}^n x_i)^{\lambda+1}} \prod_{i=1}^n I_{]a, +\infty[}(x_i) \end{aligned}$$

où $I_{]a, +\infty[}(x_i)$ est la fonction indicatrice définie sur $]a, +\infty[$ ($I_{]a, +\infty[}(x_i) = 1$ si $x > a$ et $I_{]a, +\infty[}(x_i) = 0$ si $x \leq a$). On a alors en supposant que toutes les observations vérifient $x_i > a$:

$$\ln f(x_1, \dots, x_n; \lambda) = n \ln (\lambda a^\lambda) - (\lambda + 1) \sum_{i=1}^n \ln x_i$$

En étudiant le signe de $\frac{\partial \ln f(x_1, \dots, x_n)}{\partial \lambda}$, on obtient :

$$\begin{aligned} \frac{\partial \ln f(x_1, \dots, x_n; \lambda)}{\partial \lambda} &\geq 0 \iff \frac{n}{\lambda} + n \ln a - \sum_{i=1}^n \ln x_i \geq 0 \\ \iff \frac{n}{\lambda} &\geq \sum_{i=1}^n (\ln x_i - \ln a) = \sum_{i=1}^n \ln \frac{x_i}{a} \\ \iff \lambda &\leq \left[\frac{1}{n} \sum_{i=1}^n \ln \left(\frac{x_i}{a} \right) \right]^{-1} \end{aligned}$$

La vraisemblance possède donc un maximum global unique obtenu pour $\lambda = \frac{1}{n} \sum_{i=1}^n \ln \frac{x_i}{a}$ d'où

$$\boxed{\hat{\lambda}_{MV} = \frac{n}{\sum_{i=1}^n \ln \left(\frac{x_i}{a} \right)}}$$

2) La borne de Cramer-Rao d'un estimateur non biaisé du paramètre λ est définie par

$$BCR(\lambda) = \frac{-1}{E \left[\frac{\partial^2 \ln f(x_1, \dots, x_n; \lambda)}{\partial \lambda^2} \right]} = \frac{-1}{E \left[-\frac{n}{\lambda^2} \right]}$$

d'où

$$\boxed{BCR(\lambda) = \frac{\lambda^2}{n}}$$

3)

- 3.1) Un changement de variables élémentaire permet d'obtenir la densité de la variable aléatoire V_i :

$$f(v_i) = \lambda e^{-\lambda v_i} \quad v_i \geq 0$$

qui suit une loi Gamma de paramètres λ et 1, i.e.

$$\boxed{V_i \sim \Gamma(\lambda, 1)}$$

- 3.2) En étudiant la fonction caractéristique de $X + Y$ et en utilisant l'indépendance entre les variables X et Y , on obtient :

$$\begin{aligned} E \left[e^{it(X+Y)} \right] &= E \left[e^{itX} \right] E \left[e^{itY} \right] \\ &= \frac{1}{\left(1 - \frac{it}{\alpha}\right)^{\beta_1}} \frac{1}{\left(1 - \frac{it}{\alpha}\right)^{\beta_2}} \\ &= \frac{1}{\left(1 - \frac{it}{\alpha}\right)^{\beta_1 + \beta_2}} \end{aligned}$$

qui est la fonction caractéristique d'une loi Gamma $\Gamma(\alpha, \beta_1 + \beta_2)$. On en déduit alors

$$V = \sum_{i=1}^n \ln V_i = \sum_{i=1}^n \ln \left(\frac{X_i}{a} \right) \sim \Gamma(\lambda, n)$$

- 3.3) Des calculs simples permettent d'obtenir la moyenne et la variance de l'estimateur du maximum de vraisemblance de λ :

$$\begin{aligned} E \left[\hat{\lambda}_{MV} \right] &= E \left[\frac{n}{V} \right] \\ &= \int_0^{+\infty} \frac{n}{v} \frac{\lambda^n}{\Gamma(n)} v^{n-1} e^{-\lambda v} dv \\ &= \frac{n\lambda^n}{\Gamma(n)} \int_0^{+\infty} v^{n-2} e^{-\lambda v} dv \\ &= \frac{n\lambda}{\Gamma(n)} \int_0^{+\infty} x^{n-2} e^{-x} dx \\ &= \frac{n\lambda\Gamma(n-1)}{\Gamma(n)} = \frac{n}{n-1}\lambda \end{aligned}$$

On en déduit que $\lambda^* = \frac{n-1}{n} \hat{\lambda}_{MV} = \frac{n-1}{\sum_{i=1}^n \ln \left(\frac{X_i}{a} \right)}$ est un estimateur non biaisé de λ .

De la même façon, on a

$$\begin{aligned} E \left[\hat{\lambda}_{MV}^2 \right] &= E \left[\frac{n^2}{V^2} \right] \\ &= \int_0^{+\infty} \frac{n^2}{v^2} \frac{\lambda^n}{\Gamma(n)} v^{n-1} e^{-\lambda v} dv \end{aligned}$$

$$\begin{aligned}
&= \frac{n^2 \lambda^n}{\Gamma(n)} \int_0^{+\infty} v^{n-3} e^{-\lambda v} dv \\
&\stackrel{x=\lambda v}{=} \frac{n^2 \lambda^n}{\Gamma(n) \lambda^{n-2}} \int_0^{+\infty} x^{n-3} e^{-x} dx \\
&= \frac{n^2 \lambda^2 \Gamma(n-2)}{\Gamma(n)} \\
&= \frac{n^2 \lambda^2}{(n-1)(n-2)}
\end{aligned}$$

d'où

$$\begin{aligned}
\text{var} \hat{\lambda}_{MV} &= \frac{n^2 \lambda^2}{(n-1)(n-2)} - \frac{n^2}{(n-1)^2} \lambda^2 \\
&= \frac{n^2 \lambda^2}{(n-1)} \left[\frac{1}{n-2} - \frac{1}{n-1} \right] \\
&= \frac{n^2 \lambda^2}{(n-1)^2 (n-2)}
\end{aligned}$$

et

$$\boxed{\text{var} \lambda^* = \frac{\lambda^2}{n-2}}$$

L'estimateur λ^* est donc non biaisé, convergent mais n'est il n'est pas efficace.

4) La loi a posteriori du paramètre λ vérifie

$$f(\lambda | x_1, \dots, x_n) \propto e^{-c\lambda} \lambda^{d-1} \frac{(\lambda a^\lambda)^n}{(\prod_{i=1}^n x_i)^\lambda}$$

c'est-à-dire

$$\begin{aligned}
f(\lambda | x_1, \dots, x_n) &\propto \lambda^{n+d-1} e^{-\lambda [c - n \ln a + \sum_{i=1}^n \ln x_i]} \\
&\propto \lambda^{n+d-1} e^{-\lambda [c + \sum_{i=1}^n \ln \frac{x_i}{a}]}
\end{aligned}$$

On voit donc que la loi a posteriori de λ est une loi Gamma de paramètres $n+d$ et $c - \sum_{i=1}^n \ln \frac{x_i}{a}$:

$$\lambda | x_1, \dots, x_n \sim \Gamma \left(n+d, c + \sum_{i=1}^n \ln \frac{x_i}{a} \right)$$

La maximisation de $\ln f(\lambda | x_1, \dots, x_n)$ conduit à

$$\begin{aligned}
\frac{\partial \ln f(\lambda | x_1, \dots, x_n)}{\partial \lambda} &\geq 0 \iff \frac{n+d-1}{\lambda} - \left[c + \sum_{i=1}^n \ln \frac{x_i}{a} \right] \geq 0 \\
&\iff \lambda \leq \frac{n+d-1}{c + \sum_{i=1}^n \ln \left(\frac{x_i}{a} \right)}
\end{aligned}$$

On en déduit que l'estimateur MAP de est :

$$\boxed{\hat{\lambda}_{MAP} = \frac{n+d-1}{c + \sum_{i=1}^n \ln \left(\frac{x_i}{a} \right)}}$$

On remarque que

$$\hat{\lambda}_{MAP} = \hat{\lambda}_{MV} \frac{1 + \frac{d-1}{n}}{\frac{c}{\sum_{i=1}^n \ln\left(\frac{X_i}{a}\right)} + 1}$$

et donc asymptotiquement (n grand), l'estimateur $\hat{\lambda}_{MAP}$ se comporte comme $\hat{\lambda}_{MV}$. Pour n "petit", on a

$$\hat{\lambda}_{MAP} \approx \frac{d-1}{c},$$

qui est le maximum de la loi a posteriori de λ . Lorsque n grand, l'estimateur a tendance à oublier l'information a priori. Pour n petit, l'estimateur MAP a tendance à faire confiance aux informations a priori et à oublier le terme lié aux données.

Partie 2

1) Le test de Neyman-Pearson est défini par

$$\text{Rejet de } H_0 \text{ si } \frac{f(x_1, \dots, x_n; \lambda_1)}{f(x_1, \dots, x_n; \lambda_0)} > S_\alpha$$

Mais

$$\begin{aligned} \frac{f(x_1, \dots, x_n; \lambda_1)}{f(x_1, \dots, x_n; \lambda_0)} > S_\alpha &\Leftrightarrow \frac{\lambda_1^n a^{n\lambda_1} (\prod_{i=1}^n x_i)^{\lambda_0+1}}{\lambda_0^n a^{n\lambda_0} (\prod_{i=1}^n x_i)^{\lambda_1+1}} > S_\alpha \\ &\Leftrightarrow (\lambda_0 - \lambda_1) \sum_{i=1}^n \ln x_i > \nu_\alpha \end{aligned}$$

Puisque $\lambda_0 - \lambda_1 < 0$, un test équivalent est

$$\boxed{\text{Rejet de } H_0 \text{ si } V = \sum_{i=1}^n \ln\left(\frac{X_i}{a}\right) < K_\alpha}$$

La moyenne de $V_i = \ln\left(\frac{X_i}{a}\right)$ est la moyenne d'une loi Gamma $\Gamma(\lambda, 1)$, c'est-à-dire

$$E[V_i] = \frac{1}{\lambda}$$

Donc, puisque le paramètre λ est "petit" sous l'hypothèse H et "grand" sous l'hypothèse H_1 , $E[V_i]$ est "petit" lorsque l'hypothèse H_1 est vérifiée. En vertu de la loi des grands nombres on sait que

$$\frac{1}{n} \sum_{i=1}^n \ln\left(\frac{X_i}{a}\right) \xrightarrow{n \rightarrow \infty} E[V_i]$$

Donc lorsque $\sum_{i=1}^n \ln\left(\frac{X_i}{a}\right)$ est "petit", il est logique d'accepter H_1 et donc de rejeter H_0 .

2) A partir de la définition de α , on obtient

$$\begin{aligned} \alpha &= P[\text{Rejeter } H_0 \mid H_0 \text{ vraie}] \\ &= P[V < K_\alpha \mid \lambda = \lambda_0] \\ &= P[V < K_\alpha \mid V \sim \Gamma(\lambda_0, n)] \\ &= \int_0^{K_\alpha} \frac{\lambda_0^n}{\Gamma(n)} v^{n-1} e^{-\lambda_0 v} dv \\ &= \int_{x=\lambda_0 v}^{\lambda_0 K_\alpha} \frac{1}{\Gamma(n)} x^{n-1} e^{-x} dx \\ &= Q_n(\lambda_0 K_\alpha) \end{aligned}$$

On en déduit

$$\boxed{K_\alpha = \frac{1}{\lambda_0} Q_n^{-1}(\alpha)}$$

3) La puissance du test de Neyman-Pearson est définie par

$$\begin{aligned} \pi &= 1 - \beta = P[\text{Rejeter } H_0 \mid H_1 \text{ vraie}] \\ &= P[V < K_\alpha \mid \lambda = \lambda_1] \\ &= P[V < K_\alpha \mid V \sim \Gamma(\lambda_1, n)] \\ &= \int_0^{K_\alpha} \frac{\lambda_1^n}{\Gamma(n)} v^{n-1} e^{-\lambda_1 v} dv \\ &= \int_{x=\lambda_1 v}^{\lambda_1 K_\alpha} \frac{1}{\Gamma(n)} x^{n-1} e^{-x} dx \\ &= Q_n(\lambda_1 K_\alpha) \end{aligned}$$

d'où

$$\boxed{\pi = Q_n\left(\frac{\lambda_1}{\lambda_0} Q_n^{-1}(\alpha)\right)}$$

Le rapport signal sur bruit du test est donc

$$SNR = \frac{\lambda_1}{\lambda_0}$$

et les deux tests proposés ont la même puissance (et donc la même performance).

4) Pour n "grand", le théorème de la limite centrale permet d'approcher la loi de V par une loi normale de moyenne $m = nE[V_i]$ et de variance $\sigma^2 = n \text{var}[V_i]$. Puisque $V_i \sim \Gamma(\lambda, 1)$, on a $E[V_i] = \frac{1}{\lambda}$ et $\text{var}[V_i] = \frac{1}{\lambda^2}$ d'où

$$V \approx \mathcal{N}\left(\frac{n}{\lambda}, \frac{n}{\lambda^2}\right)$$

On a alors

$$\begin{aligned} \alpha &= P[\text{Rejeter } H_0 \mid H_0 \text{ vraie}] \\ &= P[V < K_\alpha \mid \lambda = \lambda_0] \\ &= P\left[V < K_\alpha \mid V \sim \mathcal{N}\left(\frac{n}{\lambda_0}, \frac{n}{\lambda_0^2}\right)\right] \\ &= P\left[W = \frac{V - \frac{n}{\lambda_0}}{\frac{\sqrt{n}}{\lambda_0}} < \frac{K_\alpha - \frac{n}{\lambda_0}}{\frac{\sqrt{n}}{\lambda_0}} \mid W \sim \mathcal{N}(0, 1)\right] \end{aligned}$$

Les tables de la loi normale $\mathcal{N}(0, 1)$ donnent alors

$$-\frac{\lambda_0}{\sqrt{n}} \left(K_\alpha - \frac{n}{\lambda_0}\right) \simeq 2.32 \text{ d'où } K_\alpha \simeq 25.33.$$

La puissance de ce test est

$$\begin{aligned} \pi &= 1 - \beta = P[\text{Rejeter } H_0 \mid H_1 \text{ vraie}] \\ &= P[V < K_\alpha \mid \lambda = \lambda_1] \\ &= P\left[V < K_\alpha \mid \mathcal{N}\left(\frac{n}{\lambda_1}, \frac{n}{\lambda_1^2}\right)\right] \\ &= P\left[W = \frac{V - \frac{n}{\lambda_1}}{\frac{\sqrt{n}}{\lambda_1}} < \frac{K_\alpha - \frac{n}{\lambda_1}}{\frac{\sqrt{n}}{\lambda_1}} \mid W \sim \mathcal{N}(0, 1)\right] \end{aligned}$$

Les tables de loi normale $\mathcal{N}(0, 1)$ donnent alors

$$\pi \simeq 0.95$$

5) La recherche des classes équiprobables se fait par intégration de la densité

$$\begin{aligned}\int_1^{c_1} \frac{1}{x^2} dx &= 1 - \frac{1}{c_1} = \frac{1}{4} \text{ d'où } c_1 = \frac{4}{3} \\ \int_{c_1}^{c_2} \frac{1}{x^2} dx &= \frac{1}{c_1} - \frac{1}{c_2} = \frac{1}{4} \text{ d'où } c_2 = 2 \\ \int_{c_2}^{c_3} \frac{1}{x^2} dx &= \frac{1}{c_2} - \frac{1}{c_3} = \frac{1}{4} \text{ d'où } c_3 = 4\end{aligned}$$

Les quatre classes sont donc définies par

$$\begin{aligned}\text{Classe } C_1 &: \left[1, \frac{4}{3}\right[\\ \text{Classe } C_2 &: \left[\frac{4}{3}, 2\right[\\ \text{Classe } C_3 &: [2, 4[\\ \text{Classe } C_4 &: [4, +\infty[\end{aligned}$$

Les nombres d'observations de chaque classe sont

$$n_1 = 9, n_2 = 7, n_3 = 3, n_4 = 1$$

et donc la statistique de test du χ^2 peut se calculer comme suit

$$\begin{aligned}\phi &= \sum_{i=1}^K \frac{(n_i - np_i)^2}{np_i} \\ &= \frac{1}{5} (16 + 4 + 4 + 16) = 8\end{aligned}$$

Le seuil se calcule à partir du risque α :

$$\begin{aligned}\alpha &= P[\text{Rejeter } H_0 \mid H_0 \text{ vraie}] \\ &= P[\phi > S_\alpha \mid \phi \sim \chi_3^2] = 0.01\end{aligned}$$

Les tables donnent

$$S_\alpha = 11.34$$

et donc on accepte l'hypothèse H_0 (les observations sont distribuées suivant la loi de Pareto $P_{1,1}$) avec le risque $\alpha = 0.01$. Par contre, pour $\alpha = 0.05$, on a $S_\alpha = 7.81 < \phi$. Donc on rejette l'hypothèse H_0 avec le risque $\alpha = 0.05$.

Barème

1ère partie : 10 pts

1) 2 pts

2) 1 pt

3) 3.1) 1pt, 3.2) 1pt, 3.3) calcul de moyenne 1pt + calcul de variance 1pt + λ^* sans biais, cv, non efficace : 1pt

4) 1 + 1 (MAP + analyse)

2ème partie :

1) Test de NP : 1pt + explications moyenne : 1pt

2) K_α : 1pt

3) π : 1pt, SNR + analyse : 1pt

4) Seuil + puissance 1pt + 1pt

5) Test du chi2 : classes 1pt, calcul de ϕ 1pt, calcul du seuil 1pt, conclusions 1pt.