

**Partie 1**

1) La moyenne d'une loi exponentielle est donnée dans la table de la loi Gamma :

$$E[X_i] = \frac{1}{\lambda} \implies \lambda = \frac{1}{E[X_i]}$$

On obtient l'estimateur des moments en remplaçant  $E[X_i]$  par son estimateur habituel

$$\hat{\lambda}_{MO} = \left( \frac{1}{n} \sum_{i=1}^n X_i \right)^{-1} = \frac{1}{\bar{X}},$$

avec la notation habituelle  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ .

2) La vraisemblance de l'échantillon  $(X_1, \dots, X_n)$  s'écrit :

$$\begin{aligned} L(x_1, \dots, x_n; \lambda) &= \prod_{i=1}^n f(x_i; \lambda) \\ &= \prod_{i=1}^n \lambda e^{-\lambda x_i} \\ &= \lambda^n \exp\left(-\lambda \sum_{i=1}^n x_i\right) \end{aligned}$$

On a alors :

$$\ln L(x_1, \dots, x_n; \lambda) = n \ln(\lambda) - \lambda \left( \sum_{i=1}^n x_i \right)$$

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

$$\begin{aligned} \frac{\partial \ln L(x_1, \dots, x_n; \lambda)}{\partial \lambda} &\geq 0 \iff \frac{n}{\lambda} - \sum_{i=1}^n x_i \geq 0 \\ &\iff \lambda \leq \frac{n}{\sum_{i=1}^n x_i} = \frac{1}{\bar{x}}, \end{aligned}$$

La vraisemblance possède donc un maximum global unique obtenu pour  $\lambda = \frac{1}{\bar{x}}$  d'où

$$\hat{\lambda}_{MV} = \left( \frac{1}{n} \sum_{i=1}^n X_i \right)^{-1} = \frac{1}{\bar{X}}$$

L'estimateur des moments du paramètre  $\lambda$  déterminé à la première question coïncide donc avec l'estimateur du maximum de vraisemblance du paramètre  $\lambda$ .

3)

- La fonction caractéristique de  $U = \frac{1}{n} \sum_{k=1}^n X_k$  est définie par

$$\begin{aligned} E [e^{itU}] &= E \left[ \exp \left( \frac{it}{n} \sum_{k=1}^n X_k \right) \right] \\ &= \prod_{k=1}^n E \left[ \exp \left( \frac{it}{n} X_k \right) \right] \\ &= \prod_{k=1}^n \frac{1}{1 - \frac{it}{n\lambda}} = \frac{1}{\left(1 - \frac{it}{n\lambda}\right)^n} \end{aligned}$$

On voit donc que la fonction caractéristique de  $U$  est celle d'une loi Gamma  $\Gamma(n\lambda, n)$  d'où

$$U \sim \Gamma(n\lambda, n)$$

- Le biais et la variance de l'estimateur  $\widehat{\lambda}_{MV}$  sont définis par

$$\begin{aligned} \text{Biais}(\widehat{\lambda}_{MV}) &= E [\widehat{\lambda}_{MV}] - \lambda \\ \text{Var} [\widehat{\lambda}_{MV}] &= E [\widehat{\lambda}_{MV}^2] - E [\widehat{\lambda}_{MV}]^2 \end{aligned}$$

avec

$$\begin{aligned} E [\widehat{\lambda}_{MV}] &= E \left[ \frac{1}{U} \right] \\ &= \int_0^\infty \frac{1}{u} g(u) du \\ &= \int_0^\infty \frac{1}{u} \frac{(n\lambda)^n}{\Gamma(n)} e^{-n\lambda u} u^{n-1} du \\ &= \frac{(n\lambda)^n}{\Gamma(n)} \int_0^\infty \left( \frac{x}{n\lambda} \right)^{n-2} e^{-x} \frac{dx}{n\lambda} \\ &= \frac{n\lambda}{\Gamma(n)} \int_0^\infty x^{n-2} e^{-x} dx \\ &= n\lambda \frac{\Gamma(n-1)}{\Gamma(n)} = \frac{n}{n-1} \lambda \end{aligned}$$

On en déduit un estimateur non biaisé de  $\lambda$  :

$$\lambda^* = \frac{n-1}{n} \widehat{\lambda}_{MV} = \frac{n-1}{\sum_{i=1}^n X_i}$$

De la même façon, on a

$$\begin{aligned}
 E \left[ \widehat{\lambda}_{MV}^2 \right] &= E \left[ \frac{1}{U^2} \right] \\
 &= \int_0^\infty \frac{1}{u^2} g(u) du \\
 &= \int_0^\infty \frac{1}{u^2} \frac{(n\lambda)^n}{\Gamma(n)} e^{-n\lambda u} u^{n-1} du \\
 &= \frac{(n\lambda)^n}{\Gamma(n)} \int_0^\infty \left( \frac{x}{n\lambda} \right)^{n-3} e^{-x} \frac{dx}{n\lambda} \\
 &= \frac{(n\lambda)^2}{\Gamma(n)} \int_0^\infty x^{n-3} e^{-x} dx \\
 &= (n\lambda)^2 \frac{\Gamma(n-2)}{\Gamma(n)} = \frac{(n\lambda)^2}{(n-1)(n-2)}
 \end{aligned}$$

La variance de  $\widehat{\lambda}_{MV}$  est donc

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

On en déduit la variance de  $\lambda^*$  :

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

- La borne de Cramer-Rao d'un estimateur non biaisé du paramètre  $\lambda$  est définie par

$$\begin{aligned}
 BCR(\lambda) &= \frac{-1}{E \left[ \frac{\partial^2 \ln L(X_1, \dots, X_n; \lambda)}{\partial \lambda^2} \right]} \\
 &= \frac{\lambda^2}{n} < \frac{\lambda^2}{n-2}
 \end{aligned}$$

L'estimateur  $\lambda^*$  n'est donc pas l'estimateur efficace de  $\lambda$ .

- 4) La loi a posteriori du paramètre  $\lambda$  vérifie

$$f(\lambda | x_1, \dots, x_n) \propto (e^{-\alpha \lambda} \lambda^{\beta-1}) \lambda^n \exp \left( -\lambda \sum_{i=1}^n x_i \right)$$

c'est-à-dire

$$f(\lambda|x_1, \dots, x_n) \propto \lambda^{\beta+n-1} \exp\left(-\lambda \left[\sum_{i=1}^n x_i + \alpha\right]\right)$$

On voit donc que la loi a posteriori de  $\lambda$  est une loi Gamma de paramètres  $\alpha + \sum_{i=1}^n x_i$  et  $\beta + n$  :

$$\lambda|x_1, \dots, x_n \sim \Gamma\left(\alpha + \sum_{i=1}^n x_i, \beta + n\right)$$

L'estimateur MMSE est la moyenne de cette la loi a posteriori. D'après la table, il s'écrit donc

$$\hat{\lambda}_{MMSE} = \frac{\beta + n}{\alpha + \sum_{i=1}^n X_i} = \frac{1 + \frac{\beta}{n}}{\bar{X} + \frac{\alpha}{n}}$$

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

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

On en déduit que l'estimateur MAP de  $\lambda$  est :

$$\hat{\lambda}_{MAP} = \frac{\beta + n - 1}{\sum_{i=1}^n X_i + \alpha}$$

On remarque que

$$\hat{\lambda}_{MAP} = \frac{\frac{\beta-1}{n} + 1}{\bar{X} + \frac{\alpha}{n}}$$

et donc asymptotiquement ( $n$  grand), les estimateurs  $\hat{\lambda}_{MMSE}$  et  $\hat{\lambda}_{MAP}$  se comportent comme  $\hat{\lambda}_{MV}$ . Pour  $n$  "petit", on a

$$\hat{\lambda}_{MAP} \approx \frac{\beta - 1}{\alpha} \text{ et } \hat{\lambda}_{MMSE} \approx \frac{\beta}{\alpha}$$

qui sont respectivement l'argument du maximum et la moyenne de la loi a priori de  $\lambda$ . Lorsque  $n$  grand, les estimateurs Bayésiens ont tendance à oublier l'information a priori. Pour  $n$  petit, les estimateurs Bayésiens font confiance aux informations a priori et oublient le terme lié aux données.

## Partie 2

1) Le test de Neyman-Pearson est défini par

$$\text{Rejet de } H_0 \text{ si } \frac{L(x_1, \dots, x_n; \theta_1)}{L(x_1, \dots, x_n; \theta_0)} > S_\alpha$$

Mais

$$\begin{aligned}
\frac{L(x_1, \dots, x_n; \theta_1)}{L(x_1, \dots, x_n; \theta_0)} &> S_\alpha \Leftrightarrow \frac{\prod_{i=1}^n \frac{3}{\theta_1} x_i^2 \exp\left(-\frac{x_i^3}{\theta_1}\right)}{\prod_{i=1}^n \frac{3}{\theta_0} x_i^2 \exp\left(-\frac{x_i^3}{\theta_0}\right)} > S_\alpha \\
&\Leftrightarrow \left(\frac{\theta_0}{\theta_1}\right)^n \exp\left(\left(\frac{1}{\theta_0} - \frac{1}{\theta_1}\right) \sum_{i=1}^n x_i^3\right) \\
&\Leftrightarrow n \ln\left(\frac{\theta_0}{\theta_1}\right) + \left(\frac{1}{\theta_0} - \frac{1}{\theta_1}\right) \sum_{i=1}^n x_i^3 > \nu_\alpha
\end{aligned}$$

Pour  $\theta_0 < \theta_1$ , on a  $\frac{1}{\theta_0} - \frac{1}{\theta_1} > 0$ . Un test équivalent est donc

$$\text{Rejet de } H_0 \text{ si } T_n = \sum_{i=1}^n X_i^3 > K_\alpha$$

2) On effectue le changement de variables  $Y = \frac{2}{\theta} X^3$  qui est bijectif de  $\mathbb{R}^+$  dans  $\mathbb{R}^+$ . Le jacobien de ce changement de variables est

$$J = \left| \frac{dX}{dY} \right| = \frac{1}{3} \left( \frac{\theta}{2} Y \right)^{-2/3} \frac{\theta}{2}$$

La densité de  $Y$  s'écrit donc

$$\begin{aligned}
g(y) &= \frac{1}{3} \left( \frac{\theta}{2} y \right)^{-2/3} \frac{\theta}{2} \frac{3}{\theta} \left( \frac{\theta}{2} y \right)^{2/3} \exp\left(-\frac{y}{2}\right) \mathbb{I}_{\mathbb{R}^+}(y) \\
&= \frac{1}{2} \exp\left(-\frac{y}{2}\right) \mathbb{I}_{\mathbb{R}^+}(y)
\end{aligned}$$

Cette dernière expression est bien la densité d'une loi du chi-deux à deux degrés de liberté, i.e.  $Y \sim \chi_2^2$ . On en conclut à l'aide du rappel

$$\frac{2}{\theta} T_n = \frac{2}{\theta} \sum_{i=1}^n X_i^3 \sim \chi_{2n}^2$$

À partir de la définition de  $\alpha$ , on obtient

$$\begin{aligned}
\alpha &= P[\text{Rejeter } H_0 \mid H_0 \text{ vraie}] \\
&= P[T_n > K_\alpha \mid \theta = \theta_0] \\
&= P\left[\frac{2}{\theta_0} T_n > \frac{2}{\theta_0} K_\alpha \mid \frac{2}{\theta_0} T_n \sim \chi_{2n}^2\right]
\end{aligned}$$

Si  $g_{2n}(u)$  est la densité d'une loi du  $\chi_{2n}^2$ , on pose

$$\int_x^\infty g_{2n}(u) du = \Phi_{2n}(x)$$

et on obtient

$$\alpha = \Phi_{2n} \left( \frac{2}{\theta_0} K_\alpha \right) \Leftrightarrow K_\alpha = \frac{\theta_0}{2} \Phi_{2n}^{-1}(\alpha)$$

3) On donne  $\theta_0 = 1/2, \theta_1 = 1, n = 10, \alpha = 0.05$  et  $\sum_{i=1}^n x_i^3 = 18$ . Les tables donnent

$$\Phi_{20}^{-1}(0.05) = 31.410 \implies K_\alpha = \frac{31.410}{4} \simeq 7.85$$

Puisque  $\sum_{i=1}^n x_i^3 = 18 > K_\alpha$ , on rejette l'hypothèse  $H_0$  avec le risque  $\alpha = 0.05$ . La puissance du test est définie par

$$\begin{aligned} \pi &= P[\text{Rejeter } H_0 \mid H_1 \text{ vraie}] \\ &= P[T_n > K_\alpha \mid \theta = \theta_1] \\ &= P \left[ \frac{2}{\theta_1} T_n > \frac{2}{\theta_1} K_\alpha \mid \frac{2}{\theta_1} T_n \sim \chi_{2n}^2 \right] \\ &= \Phi_{2n} \left( \frac{2}{\theta_1} K_\alpha \right) \end{aligned}$$

L'application numérique donne  $\pi = \Phi_{2n}(2K_\alpha) = \Phi_{2n}(15.61)$ . Les tables donnent donc

$$0.70 < \pi < 0.80$$

4) Les courbes COR sont définies à partir des relations

$$\begin{aligned} \alpha &= \Phi_{2n} \left( \frac{2}{\theta_0} K_\alpha \right) \\ \pi &= \Phi_{2n} \left( \frac{2}{\theta_1} K_\alpha \right) \end{aligned}$$

En éliminant  $K_\alpha$ , on obtient

$$K_\alpha = \frac{\theta_0}{2} \Phi_{2n}^{-1}(\alpha) \implies \pi = \Phi_{2n} \left( \frac{\theta_0}{\theta_1} G_{2n}^{-1}(\alpha) \right)$$

Cette dernière relation liant  $\pi$  à  $\alpha$  définit les courbes COR. Le paramètre  $\theta_1$  est soumis à la contrainte  $\theta_1 > \theta_0 \iff \frac{\theta_0}{\theta_1} < 1$ . Quand on augmente le rapport  $\frac{\theta_0}{\theta_1}$  (où quand on diminue  $\frac{\theta_0}{\theta_1}$ ), la puissance augmente. Les courbes COR montrent que le paramètre important réglant la performance du test est le rapport  $\frac{\theta_1}{\theta_0}$ .

### Partie 3

Le test de Kolmogorov est défini par

$$\text{Rejet de } H_0 \text{ si } \sup_{y \in \mathbb{R}} |F_n(y) - F(y)| > K_\alpha,$$

c'est-à-dire si l'écart entre la fonction de répartition théorique de la loi du  $\chi_2^2$  et la fonction de répartition empirique de  $(y_1, \dots, y_5)$  est trop grande. Nous avons vu en cours que cet écart pouvait se déterminer comme suit

$$\sup_{y \in \mathbb{R}} |F_n(y) - F(y)| = \sup_{i=1, \dots, n} \left\{ \max \left\{ \left| F(y_i) - \frac{i-1}{n} \right|, \left| F(y_i) - \frac{i}{n} \right| \right\} \right\}$$

En utilisant le tableau :

$F(y_1) = 0.42$	$F(y_2) = 0.68$	$F(y_3) = 0.93$	$F(y_4) = 0.98$	$F(y_5) = 0.99$
-----------------	-----------------	-----------------	-----------------	-----------------

on obtient ( $\frac{1}{n} = \frac{1}{5} = 0.2$ )

$$\begin{aligned} \max \left\{ \left| F(y_1) - \frac{0}{n} \right|, \left| F(y_1) - \frac{1}{n} \right| \right\} &= \max \{0.42, 0.22\} = 0.42 \\ \max \left\{ \left| F(y_2) - \frac{1}{n} \right|, \left| F(y_2) - \frac{2}{n} \right| \right\} &= \max \{0.48, 0.28\} = 0.48 \\ \max \left\{ \left| F(y_3) - \frac{2}{n} \right|, \left| F(y_3) - \frac{3}{n} \right| \right\} &= \max \{0.53, 0.33\} = 0.53 \\ \max \left\{ \left| F(y_4) - \frac{3}{n} \right|, \left| F(y_4) - \frac{4}{n} \right| \right\} &= \max \{0.38, 0.18\} = 0.38 \\ \max \left\{ \left| F(y_5) - \frac{4}{n} \right|, \left| F(y_5) - \frac{5}{n} \right| \right\} &= \max \{0.19, 0.01\} = 0.19 \end{aligned}$$

d'où

$$\sup_{y \in \mathbb{R}} |F_n(y) - F(y)| = 0.53$$

- L'écart entre les deux fonctions de répartition est supérieur aux deux seuils  $S_{0.01} = 0.352$  et  $S_{0.05} = 0.294$ . Donc, on rejette  $H_0$  avec  $\alpha = 0.01$  et  $\alpha = 0.05$ .

- On a

$$\alpha = P[\text{Rejeter } H_0 \mid H_0 \text{ vraie}]$$

donc plus  $\alpha$  est grand, plus on rejette  $H_0$  et donc plus le seuil est faible.

- On ne peut calculer la puissance du test car la loi de  $\sup_{y \in \mathbb{R}} |F_n(y) - F(y)|$  n'est pas connue sous  $H_0$ .