CC3 2024/2025 – Durée 3h

Les documents et appareils électroniques (calculatrice, téléphone, ordinateur, ...) sont interdits. **Toutes les réponses doivent être justifiées.**

Exercice 1 - Chaleur vague ou vague de chaleur ?

Un climatologue veut estimer le temps de retour noté θ d'une vague de chaleur en Europe à partir d'observations météorologiques ces Δt dernières années et ainsi confirmer l'hypothèse selon laquelle ce temps de retour est inférieur à θ_0 , c'est à dire celui déterminé pour les décennies précédentes. Il s'intéresse pour ce faire à la variable aléatoire X modélisant le temps d'arrivée d'une nouvelle vague de chaleur. Cependant il n'a accès qu'à l'observation de la variable aléatoire N décrivant le nombre d'occurrence de vague de chaleur dans un interval de temps Δt .

1. Sous l'hypothèse que N suit la loi de Poisson $\mathcal{P}(\theta^{-1}\Delta t)$, montrer que X suit la loi exponentielle $\mathcal{E}(\theta^{-1})$. On pourra remarquer que $\mathbb{P}(X > \Delta t) = \mathbb{P}_\theta(N = 0)$.

Pour les 2 questions suivantes, on suppose que le climatologue a accès aux $X_1, \dots, X_n$.

2. Quel est le modèle ? Montrer que son information de Fisher vaut $I_1(\theta) = \theta^{-2}$.
3. Calculer un estimateur $\hat{\theta}_1$ par méthode des moments, et déterminer son comportement asymptotique. $\hat{\theta}_1$ est-il asymptotiquement efficace ?

À partir de maintenant on suppose que le climatologue est en situation réelle, et n'a accès qu'à $Y_1, \dots, Y_n$, où $Y_i = \mathbb{1}_{N_i > 0}$ et les $N_1, \dots, N_n$ sont i.i.d. de loi celle de N .

4. Donner le modèle basé sur les observations des Y_i , et l'estimateur du maximum de vraisemblance de θ noté $\hat{\theta}_2$ dans ce modèle.
5. Montrer que

$$\sqrt{n}(\hat{\theta}_2 - \theta) \rightsquigarrow \mathcal{N}(0, (e^{\theta^{-1}\Delta t} - 1)\theta^4(\Delta t)^{-2}).$$

6. (*) En déduire que la variance asymptotique est toujours supérieure à celle de $\hat{\theta}_1$. *Indication : On pourra utiliser le fait que la fonction $f(z) = (e^z - 1)z^{-2}$ est convexe et bornée inférieurement par 1 sur $\mathbb{R}_+$. Interpréter ce résultat et déduire la fenêtre de temps d'observation Δt qui minimise la variance asymptotique de $\hat{\theta}_2$.*
7. Afin de confirmer l'hypothèse du climatologue, donner les hypothèses de test sur le temps de retour θ . Réécrivez ces hypothèse en fonction de $p(\theta) = 1 - e^{-\theta^{-1}\Delta t}$, où l'on fixera $\Delta t = \theta_0 \log(2)$. Montrer que le test du rapport de vraisemblance est alors de la forme $\mathbb{1}_{\bar{Y} \geq u}$ pour un certain $u > 0$ à calibrer.

Solution 1 -

1. On a

$$\begin{aligned}\mathbb{P}(X \leq \Delta t) &= 1 - \mathbb{P}(N = 0) \\ &= 1 - \frac{(\theta^{-1}\Delta t)^0}{0!} e^{-\theta^{-1}\Delta t} \\ &= 1 - e^{-\theta^{-1}\Delta t}.\end{aligned}$$

On en déduit que $X \sim \mathcal{E}(\theta^{-1})$.

2. Posons $\lambda = \theta^{-1}$. Le modèle est $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+), (\mathcal{E}(\lambda)^{\otimes n})_{\theta > 0})$ et est dominé par la mesure de Lebesgue sur $\mathbb{R}_+$. Pour $x > 0$ et $\theta > 0$, la densité pour une observation est donnée par

$$p_\theta(x) = \theta^{-1} e^{-x/\theta}.$$

Comme $p_\theta > 0$ sur $]0; +\infty[$, $\ell_\theta(x) = \log(p_\theta(x)) = -\log(\theta) - x/\theta$ est bien définie. Cette fonction étant dérivable en θ pour tout $x > 0$, on a

$$I_1(\theta) = E_\theta \left(\frac{-1}{\theta} + \frac{X}{\theta^2} \right)^2 = \theta^{-4} E_\theta (X - \theta)^2 = \theta^{-2}.$$

3. Comme $X_1 \in L_1(P_\theta)$, on a $E_\theta(X_1) = \theta$, un estimateur par méthode des moments est donc

$$\hat{\theta}_1 = \bar{X}_n.$$

Comme $E_\theta[X_1^2] < +\infty$, on a

$$\sqrt{n}(\bar{X} - \theta) \rightsquigarrow \mathcal{N}(0, \theta^2) = \mathcal{N}(0, I(\theta)^{-1}),$$

et on en déduit que $\hat{\theta}_1$ est asymptotiquement efficace.

4. Le modèle pour les Y_i est

$$(\{0, 1\}^n, \mathcal{P}(\{0, 1\}^n), (\mathcal{B}(p(\theta))^{\otimes n})_{\theta > 0}),$$

avec $p(\theta) = 1 - e^{-\theta^{-1}\Delta t}$. Il est dominé par la mesure $(\delta_0 + \delta_1)^{\otimes n}$ et la vraisemblance s'écrit

$$V_{y_{1:n}}(\theta) = \binom{n}{n\bar{y}} p(\theta)^{n\bar{y}} (1 - p(\theta))^{n - n\bar{y}} \mathbb{1}_{p(\theta) \in]0, 1[},$$

qui atteint son maximum en $p(\theta) = \bar{y}$. Par continuité de la fonction $p(\theta)$ bijective sur $[0, 1]$, on en déduit que

$$\hat{\theta}_2 = -\frac{\Delta t}{\log(1 - \bar{Y})}$$

5. Comme $\mathbb{E}[Y_1^2] < +\infty$, on a

$$\sqrt{n}(\bar{Y} - p(\theta)) \rightsquigarrow \mathcal{N}(0, p(\theta)(1 - p(\theta)))$$

ou encore

$$\sqrt{n}(1 - \bar{Y} - e^{-\lambda\Delta t}) \rightsquigarrow \mathcal{N}(0, (1 - e^{-\lambda\Delta t})e^{-\lambda\Delta t})$$

et la delta-méthode donne pour les fonctions différentiables $y \mapsto \log(y)$ et $y \mapsto 1/y$ sur $\mathbb{R}_+$,

$$\begin{aligned} \sqrt{n}(\log(1 - \bar{Y}) + \lambda\Delta t) &\rightsquigarrow \mathcal{N}(0, e^{\lambda\Delta t} - 1), \\ \sqrt{n}\left(\frac{\Delta t}{\log(1 - \bar{Y})} + \frac{1}{\lambda}\right) &\rightsquigarrow \frac{1}{\lambda^2} \mathcal{N}\left(0, \frac{e^{\lambda\Delta t} - 1}{(\Delta t)^2}\right), \end{aligned}$$

c'est à dire

$$\sqrt{n}(\hat{\theta}_2 - \theta) \rightsquigarrow \mathcal{N}(0, (e^{\theta^{-1}\Delta t} - 1)\theta^4(\Delta t)^{-2}).$$

6. En posant $\Delta t = a\theta$ avec $a > 0$, on obtient

$$(e^{\theta^{-1}\Delta t} - 1)\theta^4(\Delta t)^{-2} = (e^a - 1)a^{-2}\theta^2.$$

Or $f(a) = (e^a - 1)a^{-2} > 1$ pour tout $a > 0$, ce qui montre que la variance asymptotique de $\hat{\theta}_2$ est supérieure à celle de $\hat{\theta}_1$. Plus précisément $f'(a)$ est négative sur $[0, a_0]$ et positive sur $[a_0, +\infty[$ et s'annule en a_0 solution de $(a - 2)e^a + 2 = 0$. On trouve $a_0 \approx 1.6$ et $f(a_0) \approx 1.5 > 1$.

Ce résultat montre que, dans le régime asymptotique, 1) les occurrences sont moins informatives pour estimer θ que les temps d'apparition des vagues de chaleur, 2) qu'il existe une fenêtre d'observation optimale $\Delta t \approx 1.6\theta$ au sens de la variance asymptotique.

7. On pose $H_0 : \theta = \theta_0$ et $H_1 : 0 < \theta < \theta_0$. En prenant $\Delta t = \log(2)\theta_0$, les hypothèses se réécrivent $H_0 : p = 1 - e^{-\theta_0^{-1}\Delta t} = 1/2$ et $H_1 : p = 1 - e^{-\theta^{-1}\Delta t} \in]1/2, 1[$ avec $n\bar{Y} \sim \mathcal{B}(n, p)$. Le rapport de vraisemblance s'écrit alors

$$RV_{H_1, H_0} = \frac{\sup_{p > 1/2} p^{n\bar{y}}(1 - p)^{n(1 - \bar{y})}}{(1/2)^n}$$

La dérivée de la fonction à maximiser est positive

$$\partial_p(p^{n\bar{y}}(1 - p)^{n(1 - \bar{y})}) = p^{n\bar{y} - 1}(1 - p)^{n - n\bar{y} - 1}(n\bar{y} - pn) \geq 0$$

pour tout $p \leq \bar{y}$ et cette dérivée s'annule en $\hat{p} = \bar{y}$. Le maximiseur sur $]1/2, 1[$ est donc $\hat{p} = \mathbb{1}_{\bar{y} > 1/2} \bar{y} + \mathbb{1}_{\bar{y} \leq 1/2} \frac{1}{2}$, et

$$\begin{aligned} RV_{H_1, H_0} &= \mathbb{1}_{\bar{y} > 1/2} \bar{y}^{n\bar{y}}(1 - \bar{y})^{n(1 - \bar{y})} 2^n + \mathbb{1}_{\bar{y} \leq 1/2} (1/2)^{n\bar{y}}(1 - 1/2)^{n(1 - \bar{y})} 2^n \\ &= \mathbb{1}_{\bar{y} > 1/2} (\bar{y}^{\bar{y}}(1 - \bar{y})^{(1 - \bar{y})})^n 2^n + \mathbb{1}_{\bar{y} \leq 1/2}. \end{aligned}$$

Dans le cas où $\bar{y} \leq 1/2$, $RV_{H_1, H_0} = 1$. Dans le cas où $\bar{y} > 1/2$, $RV_{H_1, H_0} = 2^n [\bar{y}^{\bar{y}}(1-\bar{y})^{(1-\bar{y})}]^n = 2^n \exp(n g(\bar{y}))$, avec $g(x) = x \log(x) + (1-x) \log(1-x)$, pour $x \in [0, 1]$. Comme $g'(x) = \log(x/(1-x))$, g est croissante sur $[1/2, 1]$. Par ailleurs $2^n e^{ng(1/2)} = 1$. On en déduit alors que RV_{H_1, H_0} est une fonction **croissante** de $\bar{y}$, et donc que, pour tout $u \in \mathbb{R}$,

$$RV_{H_1, H_0} \geq u \Leftrightarrow \bar{y} \geq s_u,$$

pour $s_u \in [0, 1]$. On en déduit que le test du rapport de vraisemblance de niveau α est de la forme $\mathbb{T} = \mathbb{1}_{1/2 - \bar{Y} \leq u_\alpha}$, ou u_α est un paramètre à calibrer.

Exercice 2 - Analyse de discours politique

Une équipe de chercheurs Brestois cherche à analyser et comparer des discours de politiques sur la base du champ lexical employé. À ces fins, sur un corpus de taille moyenne n (connue), ils comptent les occurrences de mots faisant référence à des thèmes choisis (chômage, détournement de fond public, etc.). On suppose qu'il y a d thèmes.

Pour un politique fixé, on modélise le nombre d'occurrence de mots du j -ème thème, noté E_j , par une loi $\mathcal{P}(np_j)$ (loi de Poisson), pour $1 \leq j \leq d$, avec $p_j \in]0, 1[$ et $\sum_{j=1}^d p_j \leq 1$. On suppose de plus que $(E_1, \dots, E_d)$ sont mutuellement indépendantes. Le style discursif du politique en question est alors caractérisé par le vecteur $\mathbf{p} = (p_1, \dots, p_d)^T$.

Estimation du style discursif

Dans cette partie l'homme politique est fixé, et on observe $\mathbf{E} = (E_1, \dots, E_d)^T$.

1. Donner un modèle pour cette expérience. Est-il dominé ?
2. Construire un estimateur $\hat{\mathbf{p}}$ de $\mathbf{p}$ par la méthode de votre choix (moments ou maximum de vraisemblance).
3. Calculer la matrice de covariance de $\hat{\mathbf{p}}$. On admet que l'information de Fisher vaut $I(\mathbf{p}) = nD(\mathbf{p})^{-1}$, où $D(\mathbf{p})$ est la matrice diagonale de coefficients $\mathbf{p}$. $\hat{\mathbf{p}}$ est-il efficace ?
4. On admet que le modèle est régulier. En déduire que, si $\hat{\mathbf{p}}_2$ est un autre estimateur non-biaisé de $\mathbf{p}$, on a

$$E_{\mathbf{p}} \|\hat{\mathbf{p}}_2 - \mathbf{p}\|^2 \geq E_{\mathbf{p}} \|\hat{\mathbf{p}} - \mathbf{p}\|^2.$$

5. On suppose que $n \rightarrow +\infty$. Montrer que

$$\sqrt{n}D(\mathbf{p})^{-1/2}(\hat{\mathbf{p}} - \mathbf{p}) \rightsquigarrow \mathcal{N}(0, I_d).$$

(On rappelle que si P_1 et P_2 sont deux lois de Poisson indépendantes de paramètres λ_1 et λ_2 , alors $P_1 + P_2 \sim \mathcal{P}(\lambda_1 + \lambda_2)$.)

6. Donner la loi limite de $\|\sqrt{n}D(\mathbf{p})^{-1/2}(\hat{\mathbf{p}} - \mathbf{p})\|^2$, et en déduire une ellipse de niveau de confiance asymptotique 98% sur $\mathbf{p}$.

7. (*) Montrer que $I(\mathbf{p}) = nD(\mathbf{p})^{-1}$.
8. (*) Justification du modèle : on suppose qu'un corpus est une suite indépendante de thèmes $(M_i)_{i=1,\dots,N}$, à valeurs dans $\llbracket 1, d \rrbracket$, et de longueur N ou $N \sim \mathcal{P}(n)$ et N est indépendante de $(M_i)_{i \geq 1}$. La loi de M_1 est caractérisée par $\mathbf{p} = (\mathbb{P}(M_1 = 1), \dots, \mathbb{P}(M_1 = d))^T$.
 Pour $i \geq 1$, on note $\mathbf{E}_i = (\mathbb{1}_{M_i=1}, \dots, \mathbb{1}_{M_i=d})^T$, et on note $\mathbf{E} = \sum_{i=1}^N \mathbf{E}_i$. Montrer que $(E_1, \dots, E_d)$ (coordonnées de $\mathbf{E}$) sont indépendantes, et que, pour tout $j \in \llbracket 1, d \rrbracket$, $E_j \sim \mathcal{P}(np_j)$.

Solution 2 -

1. Modèle : $\left(\mathbb{N}^d, \mathcal{P}(\mathbb{N}^d), \left(\bigotimes_{j=1}^d \mathcal{P}(np_j) \right)_{\mathbf{p} \in B_d} \right)$, où $B_d = \{\mathbf{u} \in]0, 1[^d \mid \sum_{j=1}^d u_j \leq 1\}$. Il est dominé par $\#^d$ (mesure de comptage sur $\mathbb{N}^d$).
2. Comme $E_{\mathbf{p}}(\|\mathbf{E}\|) < +\infty$, on peut regarder les moments d'ordre 1. On a

$$E_{\mathbf{p}}(\mathbf{E}) = n\mathbf{p},$$

l'estimateur par moments correspondant est alors $\hat{\mathbf{p}} = \mathbf{E}/n$.

3. Comme $E_{\mathbf{p}}\|\mathbf{E}\|^2 < +\infty$, $\text{Cov}_{\mathbf{p}}(\hat{\mathbf{p}})$ est bien définie. Comme $j \neq k$ implique $E_j \perp\!\!\!\perp E_k$, $\text{Cov}_{\mathbf{p}}(\hat{\mathbf{p}})$ est diagonale, de coefficients

$$\text{Cov}_{\mathbf{p}}(\hat{\mathbf{p}})_{j,j} = \text{Var}_{\mathbf{p}}(E_j/n) = \frac{p_j}{n},$$

pour $j = 1, \dots, d$. On en déduit que

$$\text{Cov}_{\mathbf{p}}(\hat{\mathbf{p}}) = \frac{1}{n}D(\mathbf{p}) = I(\mathbf{p})^{-1},$$

et donc que $\hat{\mathbf{p}}$ est efficace.

4. Le modèle étant régulier, si $\hat{\mathbf{p}}_2$ est un estimateur sans biais de $\mathbf{p}$, la borne de Cramer-Rao stipule que $\text{Cov}_{\mathbf{p}}(\hat{\mathbf{p}}_2) \succcurlyeq I(\mathbf{p})^{-1}$ (pour l'ordre sur les matrices symétriques). En particulier, on a

$$E_{\mathbf{p}}\|\hat{\mathbf{p}}_2 - \mathbf{p}\|^2 = \text{Tr}(\text{Cov}_{\mathbf{p}}(\hat{\mathbf{p}}_2)) \geq \text{Tr}(I(\mathbf{p})^{-1}) = \text{Tr}(\text{Cov}_{\mathbf{p}}(\hat{\mathbf{p}})) = E_{\mathbf{p}}\|\hat{\mathbf{p}} - \mathbf{p}\|^2.$$

5. On peut écrire $\mathbf{E} = \sum_{i=1}^n \mathbf{E}_i$, où les $\mathbf{E}_i$ sont i.i.d. de loi $\bigotimes_{j=1}^d \mathcal{P}(p_j)$. Comme $E_{\mathbf{p}}\|\mathbf{E}_1\|^2 < +\infty$, le Théorème central limite donne

$$\sqrt{n}(\hat{\mathbf{p}} - \mathbf{p}) \sim \sqrt{n} \left(\frac{\sum_{i=1}^n \mathbf{E}_i}{n} - \mathbf{p} \right) \rightsquigarrow \mathcal{N}(0, \text{Cov}_{\mathbf{p}}(\mathbf{E}_1)).$$

Or $\text{Cov}_{\mathbf{p}}(\mathbf{E}_1) = D(\mathbf{p})$. On en déduit donc que

$$\sqrt{n}D(\mathbf{p})^{-1/2}(\hat{\mathbf{p}} - \mathbf{p}) \rightsquigarrow \mathcal{N}(0, I_d).$$

6. Comme $u \mapsto \|u\|^2$ est continue, on a

$$\|\sqrt{n}D(\mathbf{p})^{-1/2}(\hat{\mathbf{p}} - \mathbf{p})\|^2 \rightsquigarrow \|\mathcal{N}(0, I_d)\|^2 \sim \chi^2(d).$$

Comme $\hat{\mathbf{p}}$ est consistant (via loi des grands nombres ou question 5-), on en déduit que $D(\hat{\mathbf{p}}) \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} D(\mathbf{p})$. Le lemme de Slutsky donne alors

$$\|\sqrt{n}D(\hat{\mathbf{p}})^{-1/2}(\hat{\mathbf{p}} - \mathbf{p})\|^2 \rightsquigarrow \sim \chi^2(d).$$

Soit q le quantile d'ordre 98% d'une loi $\chi^2(d)$. On en déduit alors que

$$\mathcal{E} = \left\{ \mathbf{u} \mid \|\sqrt{n}D(\hat{\mathbf{p}})^{-1/2}(\hat{\mathbf{p}} - \mathbf{u})\|^2 \leq q \right\} = \left\{ \mathbf{u} \mid \sum_{j=1}^d n \frac{(u_j - \hat{p}_j)^2}{\hat{p}_j} \leq q \right\}$$

est une ellipse de niveau de confiance asymptotique 98% (on a, pour tout $\mathbf{p}$, $P_{\mathbf{p}}(\mathbf{p} \in \mathcal{E}) \xrightarrow[n \rightarrow +\infty]{} 98\%$).

7. Pour $\mathbf{p} \in B_d$, la densité de $P_{\mathbf{p}}$ est donnée par

$$f_{\mathbf{p}}(x_{1:d}) = \prod_{j=1}^d e^{-np_j} \frac{(np_j)^{x_j}}{x_j!},$$

pour $x_{1:d} \in \mathbb{N}^d$. Comme, pour tout $x_{1:d} \in \mathbb{N}^d$ et $\mathbf{p} \in B_d$, $f_{\mathbf{p}}(x_{1:d}) > 0$, la log-vraisemblance est bien définie par

$$\ell_{\mathbf{p}}(x_{1:d}) = \sum_{j=1}^d -np_j + x_j \log(np_j) - \log(x_j!).$$

Pour tout $x_{1:d} \in \mathbb{N}^d$, $\ell_{\mathbf{p}}(x_{1:d})$ est différentiable en $\mathbf{p}$, de gradient

$$\dot{\ell}_{\mathbf{p}}(x_{1:d}) = \nabla_{\mathbf{p}} \ell_{\mathbf{p}}(x_{1:d}) = \left(-n + \frac{x_j}{p_j} \right)_{j=1, \dots, d}^T.$$

Comme $E_{\mathbf{p}} \|\dot{\ell}_{\mathbf{p}}(\mathbf{E})\|^2 < +\infty$, l'information de Fisher est bien définie et s'écrit

$$\begin{aligned} I(\mathbf{p}) &= E_{\mathbf{p}} \left(\dot{\ell}_{\mathbf{p}}(\mathbf{E}) \dot{\ell}_{\mathbf{p}}(\mathbf{E})^T \right) \\ &= \left(\left[\left(\frac{E_j}{p_j} - n \right) \left(\frac{E_k}{p_k} - n \right) \right]_{j,k \in \llbracket 1, d \rrbracket} \right). \end{aligned}$$

Or, $E_{\mathbf{p}}(E_j/p_j) = n$, et si $j \neq k$, $E_j \perp E_k$. On en déduit que $I(\mathbf{p})$ est diagonale, avec, pour $j \in \llbracket 1, d \rrbracket$,

$$I(\mathbf{p})_{j,j} = E_{\mathbf{p}} (E_j/p_j - n)^2 = \frac{\text{Var}_{\mathbf{p}}(E_j)}{p_j^2} = \frac{n}{p_j}.$$

On en déduit que

$$I(\mathbf{p}) = nD(\mathbf{p})^{-1},$$

où $D(\mathbf{p})$ est la matrice diagonale de coefficients $\mathbf{p}$.

8. Pour $s \geq 1$, et $k_{1:d}$, on a

$$\mathbb{P}(\mathbf{E} = (k_1, \dots, k_d)^T \mid N = s) = \frac{s!}{k_1! \dots k_d!} \prod_{j=1}^d p_j^{k_j} \mathbb{1}_{N=s}.$$

Comme $(M_i)_{i \geq 1}$ et N sont indépendantes, on a alors

$$\begin{aligned} \mathbb{P}(\mathbf{E} = (k_1, \dots, k_d)^T) &= \sum_{s=1}^{+\infty} e^{-ns} \frac{n^s}{s!} \frac{s!}{k_1! \dots k_d!} \prod_{j=1}^d p_j^{k_j} \mathbb{1}_{s=\sum_{j=1}^d k_j} \\ &= \prod_{j=1}^d e^{-np_j} \prod_{j=1}^d \frac{n^{k_j}}{k_j!} p_j^{k_j}. \end{aligned}$$

On en déduit que $\mathbf{E} \sim \bigotimes_{j=1}^d \mathcal{P}(np_j)$.

Exercice 3 - Cristallographie

Soit $(X_j)_{1 \leq j \leq n}$ des loi exponentielles indépendantes de paramètres $(s + t\kappa_j)_{1 \leq j \leq n}$. On suppose $\kappa_j > 0$ pour tout $j \geq 1$, $s \geq 0$ et $t > 0$. Les $(\kappa_j)_{1 \leq j \leq n}$ sont supposés connus, et sauf mention contrainte s et t sont supposés inconnus.

1. Quel est le modèle ? Est-il dominé ?
2. On suppose $s = 0$. Calculer l'estimateur du maximum de vraisemblance $\hat{t}$ de t .
3. Si $n \geq 3$, calculer l'espérance et la variance de $\hat{t}$.
4. En déduire un intervalle de niveau de confiance 95% sur t .

Pour la suite de l'exercice on suppose $t = 1$. On souhaite tester $H_0 : s = 0$ contre l'alternative $H_1 : s = \nu$ où $\nu > 0$ est fixé.

5. Sous H_0 , quelle est la loi de $\min_{j=1, \dots, n} X_j$?
6. Déduire de la question précédente un test de niveau α .
7. On suppose que pour tout $j = 1, \dots, n$, $\kappa_j = \kappa_j(n)$ dépend de n et que $\frac{1}{n} \sum_{j=1}^n \kappa_j(n) \xrightarrow{n \rightarrow +\infty}$
0. Montrer que la puissance du test précédent converge vers 1.
8. Écrire le test du rapport de vraisemblance. En déduire un test de niveau exactement α , que l'on exprimera en fonction des quantiles de $\sum_{j=1}^n \beta_j E_j$ avec $(E_j)_{1 \leq j \leq n}$ i.i.d. de loi exponentielles de paramètre à déterminer et $(\beta_j)_{1 \leq j \leq n}$ également à déterminer.
9. Montrer que la puissance du test du rapport de vraisemblance converge aussi vers 1 lorsque $\sum_{j=1}^n \kappa_j(n)/n \rightarrow 0$.

Solution 3 -

1. Modèle : $((0, \infty)^n, \mathcal{B}((0, \infty)^n), (\mathcal{E}(s + t\kappa_1) \otimes \cdots \otimes \mathcal{E}(s + t\kappa_n)_{s,t>0})$. Il est dominé par la mesure de Lebesgue sur $(0, \infty)$.
2. La log-vraisemblance est

$$\log V(X_1, \dots, X_n; 0, t) = \sum_{j=1}^n \log(t\kappa_j) - t\kappa_j X_j.$$

On a :

$$\partial_t \log V(X_1, \dots, X_n; 0, t) = \frac{n}{t} - \sum_{j=1}^n \kappa_j X_j,$$

$$\partial_t^2 \log V(X_1, \dots, X_n; 0, t) = -n/t^2 < 0.$$

Donc, l'EMV est :

$$\hat{t} = \left(\frac{1}{n} \sum_{j=1}^n \kappa_j X_j \right)^{-1}.$$

3. Remarquons que $\hat{t} = t/\Gamma(n, n)$. Donc :

$$\begin{aligned} \mathbb{E}[\hat{t}] &= t \int_0^\infty x^{-1} x^{n-1} e^{-nx} \frac{n^n}{\Gamma(n)} dx \\ &= t \frac{n^n}{n^{n-1} \Gamma(n)} \int_0^\infty x^{n-2} e^{-x} dx = t \frac{n^n \Gamma(n-1)}{n^{n-1} \Gamma(n)} = t \frac{n}{n-1}. \end{aligned}$$

$$\begin{aligned} \mathbb{E}[\hat{t}^2] &= t^2 \int_0^\infty x^{-2} x^{n-1} e^{-nx} \frac{n^n}{\Gamma(n)} dx \\ &= t^2 \frac{n^n}{n^{n-2} \Gamma(n)} \int_0^\infty x^{n-3} e^{-x} dx = t^2 \frac{n^n \Gamma(n-2)}{n^{n-2} \Gamma(n)} \\ &= t \frac{n^2}{(n-1)(n-2)}. \end{aligned}$$

D'où

$$\text{Var} [\hat{t}] = t^2 \frac{n^2}{n-1} \left(\frac{1}{n-2} - \frac{1}{n-1} \right) = t^2 \frac{n^2}{(n-1)^2 (n-2)}.$$

4. Soit $x > 0$, en utilisant l'inégalité de Bienaymé-Cebicev on a, pour tout $t > 0$,

$$P_t \left(\left| \hat{t} - \frac{n}{n-1}t \right| > x \right) \leq \frac{t^2 n^2}{(n-1)^2 (n-2) x^2}.$$

En posant $x_0 = \sqrt{\frac{n^2}{(n-1)^2 (n-2) 5\%}}$, on a

$$P_t \left(\left| \hat{t} - \frac{n}{n-1}t \right| > tx_0 \right) \leq 5\%.$$

Or, pour tout $t > 0$,

$$\left| \hat{t} - \frac{n}{n-1}t \right| \leq tx_0 \Leftrightarrow \left(\frac{n}{n-1} - x_0 \vee 0 \right) \leq \frac{\hat{t}}{t} \leq \frac{n}{n-1} + x_0.$$

On en déduit que

$$\left[\frac{\hat{t}}{\frac{n}{n-1} + x_0}; \frac{\hat{t}}{\left(\frac{n}{n-1} - x_0 \vee 0 \right)} \right]$$

est un intervalle de niveau de confiance 95% (avec la convention que le terme de droite vaut $+\infty$ lorsque $\frac{n}{n-1} - x_0 \leq 0$).

5.

$$\mathbb{P}[\min_{j=1, \dots, n} X_j \geq x] = \prod_{i=1}^n \mathbb{P}[X_j \geq x] = e^{-x \sum_{j=1}^n \kappa_j}.$$

Donc $\min_{j=1, \dots, n} X_j$ suit une loi exponentielle de paramètre $\sum_{j=1}^n \kappa_j$.

6. On prend

$$T = \mathbf{1}_{\min_{j=1, \dots, n} X_j \leq q},$$

avec $q = \log(1/(1 - \alpha)) \left(\sum_{j=1}^n \kappa_j \right)^{-1}$.

7. On a

$$\begin{aligned} \mathbb{P}_{H_1}(\min_{j=1, \dots, n} X_j \leq q) &= 1 - e^{-q \sum_{j=1}^n (\nu + \kappa_j)} \\ &= 1 - \exp \left(\log(1 - \alpha) \frac{n\nu + \sum_{j=1}^n \kappa_j}{\sum_{j=1}^n \kappa_j} \right) \\ &= 1 - \exp \left(\log(1 - \alpha) \frac{\nu + \frac{1}{n} \sum_{j=1}^n \kappa_j}{\frac{1}{n} \sum_{j=1}^n \kappa_j} \right) \rightarrow 1. \end{aligned}$$

8. On a :

$$R = \prod_{i=1}^n \frac{(\nu + \kappa_j)}{\kappa_j} e^{-(\nu + \kappa_j)X_j + \kappa_j X_j}.$$

Ainsi

$$R \geq q \iff \sum_{i=1}^n X_j \leq q'.$$

En prenant q' le quantile d'ordre α de $\sum_{j=1}^n (\kappa_j)^{-1} \mathcal{E}(1)$ (qui est une loi diffuse), on a un test de niveau exact α .

9. D'après Neyman Pearson,

$$\mathbb{P}_{H_1} \left(\min_{j=1, \dots, n} X_j \leq q \right) \leq \mathbb{P}_{H_1} [R \geq q].$$