Dans ce sujet, par convention un vecteur x est un vecteur colonne. On notera $x^\top$ la transposée du vecteur (ou de la matrice) x. Pour $x \in \mathbb{R}^n$, $\|x\|$ désigne la norme euclidienne : $\|x\|^2 := x^\top x$.
$\mathbb{N}^*$ désigne l'ensemble des entiers naturels strictement positifs ; $\mathbb{R}_+$ est l'ensemble des réels positifs ou nuls et $\mathbb{R}_+^*$ est l'ensemble des réels strictement positifs.
Pour $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, on pose $x_{n:n} := \max(x_1, \ldots, x_n)$.
Une loi Gamma de paramètres (p, λ) (définition IV.3-12 du polycopié) a pour espérance p/λ et pour variance p/λ^2.

Exercice 1. Soient $n \in \mathbb{N}$, $n \geq 3$ et $0 \leq t_1 < \cdots < t_n$ des instants connus. Nous définissons les vecteurs $\mathbf{1}$, $\mathbf{x}$ et la matrice $\mathbf{X}$ de taille $n \times 2$ par

$$\mathbf{1} := \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{R}^n, \qquad \mathbf{x} := \begin{bmatrix} t_1 \\ \vdots \\ t_n \end{bmatrix} \in \mathbb{R}^n, \qquad \mathbf{X} := [\mathbf{1}, \mathbf{x}] \in \mathbb{R}^{n \times 2} .$$

Les hypothèses sur les composantes de $\mathbf{x}$ entraînent que $\mathbf{X}$ est de rang 2 et par suite, $\mathbf{X}^\top \mathbf{X}$ est inversible. Soit $\mathbf{Y} \in \mathbb{R}^n$ une observation du modèle statistique

$$(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \{\mathrm{N}(\mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathrm{I}_n), (\boldsymbol{\beta}, \sigma^2) \in \Theta := \mathbb{R}^2 \times \mathbb{R}_+^*\}) , \qquad \mathbf{Y} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} .$$

Pour $\theta = (\boldsymbol{\beta}, \sigma^2) \in \Theta$, nous notons $\mathrm{L}_n(\theta, \mathbf{Y})$ la vraisemblance de l'observation $\mathbf{Y}$, donnée par

$$\mathrm{L}_n(\theta, \mathbf{Y}) = (2\pi\sigma^2)^{-n/2} \exp(-\|\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}\|^2 / 2\sigma^2) .$$

Soit $\Theta_0 := \{(\beta_1, \beta_2, \sigma^2) \in \Theta, \beta_2 = 0\}$.

1. Montrer [en justifiant soigneusement le résultat] que l'estimateur du maximum de vraisemblance $(\hat{\boldsymbol{\beta}}, \widehat{\sigma}^2)$ de $(\boldsymbol{\beta}, \sigma^2)$ est donné par

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{Y} , \qquad \hat{\sigma}^2 = n^{-1} \|\mathbf{Y} - \Pi \mathbf{Y}\|^2 , \tag{1}$$

   où $\Pi := \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top$ est le projecteur orthogonal sur l'espace vectoriel engendré par les colonnes de $\mathbf{X}$.

2. En déduire que

$$\sup_{\theta \in \Theta} \mathrm{L}_n(\theta, \mathbf{Y}) = (2\pi\widehat{\sigma}^2)^{-n/2} \mathrm{e}^{-n/2} .$$

3. Montrer que $\sup_{\theta \in \Theta_0} \mathrm{L}_n(\theta, \mathbf{Y}) = \left(2\pi\tilde{\sigma}^2\right)^{-n/2} \mathrm{e}^{-n/2}$ où $\tilde{\sigma}^2 := n^{-1} \|\mathbf{Y} - \Pi_0 \mathbf{Y}\|^2$ avec $\Pi_0 = n^{-1} \mathbf{1}\mathbf{1}^\top$.

4. Montrer que

$$\tilde{\sigma}^2 = \left(\widehat{\sigma}^2 + n^{-1} \mathbf{Y}^\top \{\Pi - \Pi_0\} \mathbf{Y}\right)$$

   où $\hat{\sigma}^2$ est l'estimateur défini par (1).

Nous considérons le test

$$H_0 : \ \theta \in \Theta_0, \quad \text{contre} \quad H_1 : \ \theta \notin \Theta_0. \tag{2}$$

Nous étudions le test du Rapport de Vraisemblance Généralisée (RVG) dont la statistique est donnée par

$$\Lambda(\mathbf{Y}) := \frac{\sup_{\theta\in\Theta_0} \mathrm{L}_n(\theta, \mathbf{Y})}{\sup_{\theta\in\Theta} \mathrm{L}_n(\theta, \mathbf{Y})} \ .$$

Pour $\alpha \in]0,1[$, nous considérons le test de fonction critique

$$\phi_\alpha(\mathbf{Y}) := \mathbb{1}_{[0,c_\alpha]}(\Lambda(\mathbf{Y})) \ ,$$

où $c_\alpha \geq 0$ vérifie : $\sup_{\theta\in\Theta_0} \mathbb{P}_\theta\left(\phi_\alpha(\mathbf{Y}) = 1\right) \leq \alpha$.

5. Montrer qu'il existe $k_\alpha \geq 0$ tel que le test RVG ϕ_α est équivalent au test de fonction critique $\psi_\alpha(\mathbf{Y}) := \mathbb{1}_{[k_\alpha,\infty[}(\mathrm{F}(\mathbf{Y}))$ où

$$\mathrm{F}(\mathbf{Y}) := (n-2)\frac{\mathbf{Y}^\top\{\Pi - n^{-1}\mathbf{1}\mathbf{1}^\top\}\mathbf{Y}}{n\hat{\sigma}^2} \ .$$

6. Montrer que, pour tout $\theta \in \Theta_0$, sous $\mathbb{P}_\theta$, la statistique $\mathrm{F}(\mathbf{Y})$ est distribuée suivant une loi de Fisher (voir polycopié, Section IV-3.6) dont on précisera les paramètres.
7. En déduire, pour $\alpha \in]0,1[$, un test de niveau α pour le problème (2).

Exercice 2. Soit $(X_1, \ldots, X_n)$ un n-échantillon du modèle statistique

$$(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+), \left\{p_\theta \cdot \mathrm{Leb} \ : \ \theta = (a,\beta) \in \Theta := \mathbb{R}_+^* \times \mathbb{R}_+^*\right\}) \ ,$$

où, pour tout $x \in \mathbb{R}_+$ et $\theta = (a,\beta) \in \Theta$, la densité de probabilité p_θ par rapport à la mesure de Lebesque Leb sur $\mathbb{R}_+$, est donnée par

$$p_\theta(x) := \frac{a}{\beta^a} x^{a-1} \mathbb{1}_{[0,\beta]}(x).$$

1. Montrer [en justifiant soigneusement le résultat] que l'estimateur du maximum de vraisemblance de (a,β) est donné par

$$\hat{\beta}_n := X_{n:n} \quad \text{et} \quad \hat{a}_n := \left\{\frac{1}{n}\sum_{i=1}^n \log(\hat{\beta}_n/X_i)\right\}^{-1} \ .$$

2. Montrer que pour tout $\theta \in \Theta$,

$$\mathbb{P}_{n,\theta}(\hat{\beta}_n \leq x) = \begin{cases} 0 & \text{si} \quad x \leq 0 \\ (x/\beta)^{an} & \text{si} \quad 0 \leq x \leq \beta \\ 1 & \text{si} \quad x \geq \beta \ . \end{cases}$$

3. En déduire que pour tout $\theta = (a,\beta) \in \Theta$, $\sqrt{n}\log(\hat{\beta}_n/\beta) \stackrel{\mathbb{P}_{n,\theta}-\text{prob}}{\longrightarrow} 0$.
4. Montrer que pour $\theta = (a,\beta) \in \Theta$, sous p_θ.Leb, la loi de la variable $Y_1 := \log(\beta/X_1)$ est Gamma$(1,a)$ (voir Définition IV-3.12 dans le polycopié).
5. Montrer que pour tout $\theta = (a,\beta) \in \Theta$, $n^{-1/2}\sum_{i=1}^n\{\log(\beta/X_i) - 1/a\}$ converge en loi vers une limite que l'on identifiera.
6. En déduire que la suite d'estimateurs $\{\hat{a}_n, \ n \in \mathbb{N}\}$ est asymptotiquement normale (on précisera la variance asymptotique).

7. Pour $\alpha \in]0,1[$, déterminer un intervalle de confiance de a de probabilité de couverture asymptotique $1-\alpha$.
8. Montrer que la suite de fonctions $G_n(X_1,\ldots,X_n;a,\beta) := na(1-\hat{\beta}_n/\beta)$ est asymptotiquement pivotale. En déduire que la suite de fonctions $\tilde{G}_n(X_1,\ldots,X_n;\beta) := n\hat{a}_n(1-\hat{\beta}_n/\beta)$ est asymptotiquement pivotale.
9. Pour $\alpha \in]0,1[$, déterminer un intervalle de confiance de β de probabilité de couverture asymptotique $1-\alpha$.

Exercice 3. Nous considérons un problème de classification binaire avec abstention : la règle d'apprentissage a une possibilité de s'abstenir de classifier.
Nous utilisons les notations du polycopié, Chapitre III-1 : nous supposons que l'observation (X,Y) (où X est le vecteur de attributs et Y est l'étiquette) est un vecteur aléatoire défini sur un espace de probabilité $(\Omega,\mathcal{F},\mathbb{P})$ à valeurs dans $\mathbb{R}^d \times \{0,1\}$; et nous notons η la fonction de régression :

$$x \in \mathbb{R}^d \mapsto \eta(x) := \mathbb{P}\left(Y=1 \,|\, X=x\right) \ .$$

Nous définissons deux fonctions : une règle de classification $g : \mathbb{R}^d \to \{0,1\}$ et une règle d'abstention $r : \mathbb{R}^d \to \{0,1\}$. Lorsque $r(x)=0$, la règle d'apprentissage se prononce et décide $g(x)$; lorsque $r(x)=1$, la règle d'apprentissage s'abstient. Nous considérons la perte ℓ_ρ définie par

$$\ell_\rho(g(x),r(x),y) := \mathbb{1}_{\{g(x)\neq y\}}\mathbb{1}_{\{0\}}(r(x)) + \rho\,\mathbb{1}_{\{1\}}(r(x))$$

où $\rho \in [0,+\infty[$. Cette perte exprime que lorsque la règle d'apprentissage s'abstient, elle subit une perte ρ et lorsqu'elle se prononce, elle subit une perte de type $0-1$.
Dans la suite, les fonctions g et r sont toujours supposées mesurables ; le risque associé à la perte ℓ_ρ est noté

$$R_\rho(g,r) := \mathbb{E}[\ell_\rho(g(X),r(X),Y)] \ .$$

1. Montrer que pour toute paire de fonctions (g,r), nous avons

$$R_\rho(g,r) = \mathbb{E}[\mathbb{1}_{\{0\}}(r(X))\{\eta(X)\mathbb{1}_{\{0\}}(g(X)) + (1-\eta(X))\mathbb{1}_{\{1\}}(g(X))\}] + \rho\,\mathbb{P}(r(X)=1) \ . \tag{3}$$

Nous posons $h(x) := \min(\eta(x), 1-\eta(x))$.

2. Donner l'expression d'une règle de classification g_* telle que pour toute paire de fonctions (g,r) et tout $\rho \geq 0$,

$$R_\rho(g,r) \geq R_\rho(g_*,r) \ , \qquad \text{avec } R_\rho(g_*,r) = \mathbb{E}[\mathbb{1}_{\{0\}}(r(X))\ h(X)] + \rho\ \mathbb{P}(r(X)=1) \ .$$

3. Montrer que pour tout $\rho \geq 0$ et pour toute règle d'abstention r, on a

$$R_\rho(g_*,r) = \rho + \mathbb{E}[\mathbb{1}_{\{0\}}(r(X))\{h(X)-\rho\}]$$

4. Comment choisir la règle d'abstention r pour $\rho = 0$?
5. Comment choisir la règle d'abstention r pour $\rho \geq 1/2$?
6. Lorsque $\rho \in]0,1/2[$, quelle est la règle d'abstention optimale r_* ?

Exercice 4. Soit $Z = (X_1, \ldots, X_n)$ un n-échantillon du modèle $(\mathbb{N}, \mathcal{P}(\mathbb{N}), (\mathbb{P}_\theta)_{\theta \in \Theta})$, où $\Theta := \mathbb{N}^*$ et, pour tout $\theta \in \Theta$, $\mathbb{P}_\theta$ est la loi de densité

$$p_\theta(x) := \theta^{-1} \mathbb{1}_{\{1,\ldots,\theta\}}(x) = \begin{cases} \theta^{-1} & \text{si } x \in \{1, \ldots, \theta\} \\ 0 & \text{sinon} \end{cases} , \quad x \in \mathbb{N} ,$$

par rapport à la mesure de comptage sur $\mathbb{N}$.

Soit $\theta_0 \in \Theta$ et $\alpha \in]0, 1[$, tels que $\theta_0 \alpha^{1/n} \in \mathbb{N}^*$.

1. Montrer que pour toute statistique S définie sur $\mathbb{N}^n$ et pour tout $\theta > \theta_0$,

$$\mathbb{E}_\theta[S(Z) \mathbb{1}_{\{X_{n:n} \le \theta_0\}}] = \mathbb{E}_{\theta_0}[S(Z)](\theta_0/\theta)^n .$$

On considère le test

$$H_0 : \theta \le \theta_0, \quad \text{contre} \quad H_1 : \theta > \theta_0. \tag{4}$$

Nous allons montrer que le test randomisé de fonction critique

$$\phi_{\mathrm{m}}(Z) := \begin{cases} 1 & \text{si } X_{n:n} > \theta_0 \\ \alpha & \text{si } X_{n:n} \le \theta_0 \end{cases}$$

est U. P. P.(α) pour (4).

2. Montrer que le test randomisé de fonction critique ϕ_{m} est de niveau α pour (4).
3. Montrer que pour tout test randomisé de niveau α de fonction critique $\phi : Z \mapsto \phi(Z) \in [0, 1]$, et pour tout $\theta > \theta_0$,

$$\mathbb{E}_\theta[\phi_{\mathrm{m}}(Z)] - \mathbb{E}_\theta[\phi(Z)] \ge \mathbb{E}_\theta \left[(\alpha - \phi(Z)) \mathbb{1}_{\{X_{n:n} \le \theta_0\}}\right] \ge 0 .$$

   Conclure.

On considère maintenant

$$H_0 : \theta = \theta_0, \quad \text{contre} \quad H_1 : \theta \ne \theta_0. \tag{5}$$

Nous considérons le test bilatéral de fonction critique

$$\phi_{\mathrm{b}}(Z) := \begin{cases} 1 & \text{si } X_{n:n} > \theta_0 \text{ ou si } X_{n:n} \le \theta_0 \alpha^{1/n} \\ 0 & \text{autrement} \end{cases}$$

Nous allons montrer que ϕ_{b} est U. P. P.(α) pour (5). Dans la suite, ϕ est la fonction critique d'un test randomisé de niveau α de (5).

4. Montrer que le test randomisé de fonction critique ϕ_{b} est de niveau α pour le test (5).
5. Montrer que pour tout $\theta > \theta_0$, nous avons

$$\mathbb{E}_\theta[\phi_{\mathrm{b}}(Z)] - \mathbb{E}_\theta[\phi(Z)] \ge \mathbb{E}_\theta \left[(\phi_{\mathrm{b}}(Z) - \phi(Z)) \mathbb{1}_{\{X_{n:n} \le \theta_0\}}\right] \ge 0 .$$

6. Montrer que pour tout $\theta < \theta_0$, nous avons

$$\mathbb{E}_\theta[\phi_{\mathrm{b}}(Z)] - \mathbb{E}_\theta[\phi(Z)] = \mathbb{P}_\theta \left(X_{n:n} \le \theta_0 \alpha^{1/n}\right) - \mathbb{E}_\theta[\phi(Z)] .$$

7. En déduire que pour $\theta < \theta_0$, nous avons $\mathbb{E}_\theta[\phi_{\mathrm{b}}(Z)] - \mathbb{E}_\theta[\phi(Z)] \ge 0$; on distinguera les cas $\theta \le \theta_0 \alpha^{1/n}$ et $\theta_0 \alpha^{1/n} \le \theta < \theta_0$.
   Conclure.

1 Solutions

Solution . 1. La log-vraisemblance est définie sur Θ par (à une constante additive près)

$$\theta \mapsto -\frac{n}{2}\ln(\sigma^2) - \frac{\|\mathbf{Y}-\mathbf{X}\boldsymbol{\beta}\|^2}{2\sigma^2}.$$

Maximiser en $\boldsymbol{\beta}$ cette expression revient à minimiser en $\boldsymbol{\beta}$ la forme quadratique

$$\boldsymbol{\beta} \mapsto \|\mathbf{Y}-\mathbf{X}\boldsymbol{\beta}\|^2 = \boldsymbol{\beta}^\top\mathbf{X}^\top\mathbf{X}\boldsymbol{\beta} - 2\boldsymbol{\beta}^\top\mathbf{X}^\top\mathbf{Y} + \mathbf{Y}^\top\mathbf{Y}\ ;$$

Le minimum est atteint au point $\hat{\boldsymbol{\beta}}$ qui est tel que $\mathbf{X}\hat{\boldsymbol{\beta}}$ est la projection orthogonale de $\mathbf{Y}$ sur l'espace engendré par $(\mathbf{1},\mathbf{x})$. En utilisant le théorème de Pythagore,

$$\mathbf{X}^T(\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}}) = \mathbf{0}_{2\times 1},$$

ce qui implique : $\hat{\boldsymbol{\beta}} := (\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\mathbf{Y}$. Nous avons donc, pour tout $\sigma^2 > 0$

$$\mathrm{L}_n(\boldsymbol{\beta},\sigma^2,\mathbf{Y}) \le \mathrm{L}_n(\hat{\boldsymbol{\beta}},\sigma^2,\mathbf{Y})\ ;$$

il reste à maximiser le terme de droite par rapport à σ^2 :

$$\sigma^2 \mapsto -\frac{n}{2}\ln(\sigma^2) - \frac{\|\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}}\|^2}{2\sigma^2}.$$

Cette fonction tend vers $-\infty$ quand $\sigma^2 \to 0^+$ et $\sigma^2 \to +\infty$; elle est continue sur $\mathbb{R}_+^*$ et sa dérivée possède un unique zero donné par

$$\hat{\sigma}^2 := \frac{1}{n}\|\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}}\|^2\ .$$

Ainsi, nous avons établi $\mathrm{L}_n(\boldsymbol{\beta},\sigma^2,\mathbf{Y}) \le \mathrm{L}_n(\hat{\boldsymbol{\beta}},\sigma^2,\mathbf{Y}) \le \mathrm{L}_n(\hat{\boldsymbol{\beta}},\hat{\sigma}^2,\mathbf{Y})$ pour tout $\theta \in \Theta$: $(\hat{\boldsymbol{\beta}},\hat{\sigma}^2)$ est bien le maximum de vraisemblance.

2. Nous avons $\sup_{\theta\in\Theta}\mathrm{L}_n(\theta,\mathbf{Y}) = \mathrm{L}_n(\hat{\boldsymbol{\beta}},\hat{\sigma}^2,\mathbf{Y})$ ce qui entraine

$$\sup_{\theta\in\Theta}\mathrm{L}_n(\theta,\mathbf{Y}) = (2\pi\hat{\sigma}^2)^{-n/2}\exp(-\|\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}}\|^2/(2\hat{\sigma}^2))\ .$$

En utilisant $\hat{\sigma}^2 = n^{-1}\|\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}}\|^2$ dans l'argument de la fonction exponentielle, nous obtenons que

$$\exp(-\|\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}}\|^2/(2\hat{\sigma}^2)) = \exp(-n/2)$$

ce qui conclut la démonstration.

3. Nous cherchons le maximum de

$$(\beta_0,\sigma_0^2) \mapsto -\frac{n}{2}\ln(\sigma_0^2) - \frac{\|\mathbf{Y}-\mathbf{1}\beta_0\|^2}{2\sigma_0^2}.$$

Nous utilisons un raisonnement similaire à la question 1, en remplaçant $\mathbf{X}$ par $\mathbf{1}$. Nous obtenons donc

$$\widehat{\beta}_0 := n^{-1}\mathbf{1}^\top\mathbf{Y} = \frac{1}{n}\sum_{i=1}^n Y_i\ , \qquad \widehat{\sigma}_0^2 := \frac{1}{n}\|\mathbf{Y}-\widehat{\beta}_0\mathbf{1}\|^2\ .$$

Nous avons $\sup_{\theta \in \Theta_0} \mathrm{L}_n(\theta, \mathbf{Y}) = \mathrm{L}_n(\widehat{\beta}_0, \widehat{\sigma}_0^2, \mathbf{Y})$ et donc

$$\sup_{\theta \in \Theta_0} \mathrm{L}_n(\theta, \mathbf{Y}) = (2\pi\hat{\sigma}_0^2)^{-n/2} \exp(-n/2) \; ;$$

il reste à relier $\hat{\sigma}_0^2$ et $\hat{\sigma}^2$. On note $\Pi_0 := n^{-1}\mathbf{1}\mathbf{1}^\top$, le projecteur orthogonal sur l'espace vectoriel engendré par $\mathbf{1}$. Nous avons $\Pi\Pi_0 = \Pi_0\Pi = \Pi_0$ et donc $(\mathrm{I}_n - \Pi)(\Pi - \Pi_0) = 0$, $\Pi^2 = \Pi$ et $\Pi^\top = \Pi$. Par conséquent

$$\begin{aligned} n\hat{\sigma}_0^2 &= \|(\mathrm{I}_n - \Pi_0)\mathbf{Y}\|^2 = \|(\mathrm{I}_n - \Pi)\mathbf{Y} + (\Pi - \Pi_0)\mathbf{Y}\|^2 \\ &= \|(\mathrm{I}_n - \Pi)\mathbf{Y}\|^2 + \|(\Pi - \Pi_0)\mathbf{Y}\|^2 = n\hat{\sigma}^2 + \|(\Pi - \Pi_0)\mathbf{Y}\|^2 \; , \end{aligned}$$

ce qui conclut la preuve.

4. Nous écrivons

$$\begin{aligned} \Lambda \leq c_\alpha \iff \Lambda^{-2/n} \geq c_\alpha^{-2/n} &\iff 1 + \frac{\mathbf{Y}^T \left(\Pi - n^{-1}\mathbf{1}\mathbf{1}^\top\right) \mathbf{Y}}{n\hat{\sigma}^2} \geq c_\alpha^{-2/n} \\ &\iff \frac{\mathbf{Y}^T \left(\Pi - n^{-1}\mathbf{1}\mathbf{1}^\top\right) \mathbf{Y}}{n\hat{\sigma}^2} \geq c_\alpha^{-2/n} - 1 \\ &\iff F \geq (n-2)\left(c_\alpha^{-2/n} - 1\right) \; . \end{aligned}$$

Noter que puisque $\Theta_0 \subset \Theta$, nous avons $\Lambda \leq 1$ et donc $c_\alpha \leq 1$. Par suite, $c_\alpha^{-n/2} \geq 1$ et le terme $(n-2)\left(c_\alpha^{-2/n} - 1\right)$ est bien positif ou nul. Cela établit l'existence de k_α.

5. Soit $\theta = (\beta_1, 0, \sigma^2) \in \Theta_0$. Sous $\mathbb{P}_\theta$

$$\mathbf{Y} \sim \mathrm{N}(\beta_1 \mathbf{1}, \sigma^2 \mathrm{I}_n) \; . \tag{6}$$

Nous écrivons

$$\begin{aligned} F &= \left(\frac{\mathbf{Y}^T \left(\Pi - n^{-1}\mathbf{1}\mathbf{1}^\top\right) \mathbf{Y}}{\sigma^2}\right) \times \left(\frac{n\hat{\sigma}^2}{(n-2)\sigma^2}\right)^{-1} \\ &= \left(\frac{\mathbf{Y}^T \left(\Pi - n^{-1}\mathbf{1}\mathbf{1}^\top\right) \mathbf{Y}}{\sigma^2}\right) \times \left(\frac{\|\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}\|^2}{(n-2)\sigma^2}\right)^{-1} \end{aligned}$$

Nous allons démontrer successivement que (i) le premier terme du produit suit une loi du χ^2 à 1 degré de liberté ; (ii) le second terme est l'inverse d'une loi de χ^2 à $(n-2)$ degrés de liberté ; (iii) les deux termes du produit sont indépendants. Par application de la définition IV.3-26, nous en déduisons que sous $\mathbb{P}_\theta$ pour $\theta \in \Theta_0$: F suit une loi de Fisher à $(1, n-2)$ degrés de liberté.

Etape 1 En utilisant (6), nous voyons que sous $\mathbb{P}_\theta$, $\mathbf{Y}/\sigma \sim \mathrm{N}(\beta_1\mathbf{1}, \mathrm{I}_n)$. Or, Π est un projecteur orthogonal sur le sous-espace de dimension 2, engendré par les vecteurs $(\mathbf{1}, \mathbf{x})$; $\Pi_0 := n^{-1}\mathbf{1}\mathbf{1}^\top$ est le projecteur sur l'espace engendré par le vecteur $\mathbf{1}$; donc $\Pi - n^{-1}\mathbf{1}\mathbf{1}^\top$ est le projecteur orthogonal sur l'espace de dimension un engendré par $\mathbf{x} - \mathbf{1}\mathbf{1}^\top\mathbf{x}/n$. Ce vecteur étant orthogonal à $\mathbf{1}$, nous en déduisons que

$$\frac{\mathbf{Y}^T \left(\Pi - \Pi_0\right) \mathbf{Y}}{\sigma^2} = \| \left(\Pi - \Pi_0\right) \frac{\mathbf{Y}}{\sigma} \|^2 = \| \left(\Pi - \Pi_0\right) \left(\frac{\mathbf{Y} - \beta_1\mathbf{1}}{\sigma}\right) \|^2 \; .$$

En conséquence, ce terme est la norme au carré du projeté orthogonal sur un espace vectoriel de dimension un, d'un vecteur gaussien centré, de matrice de covariance l'identité. Sa loi est un χ^2 de paramètre 1 ; en tant que v.a., il appartient à la tribu engendrée par $\mathbf{X}^\top \mathbf{Y}$.

Etape 2 Puisque $\mathrm{I}_n - \Pi$ est un projecteur orthogonal sur un espace de dimension $(n-2)$ et orthogonal au vecteur $\mathbf{1}$, nous avons

$$\sigma^{-2}\|\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}\|^2 = \|(\mathrm{I}_n - \Pi)\frac{\mathbf{Y}}{\sigma}\|^2 = \|(\mathrm{I}_n - \Pi)\frac{\mathbf{Y} - \beta_1 \mathbf{1}}{\sigma}\|^2$$

En conséquence, ce terme est la norme au carré du projeté orthogonal sur un espace vectoriel de dimension $(n-2)$, d'un vecteur gaussien centré, de matrice de covariance l'identité. Sa loi est un χ^2 de paramètre $(n-2)$; en tant que v.a., il appartient à la tribu engendrée par $(\mathrm{I}_n - \Pi)\mathbf{Y}$.

Etape 3 Les vecteurs $(\mathrm{I}_n - \Pi)\mathbf{Y} \in \mathbb{R}^n$ et $\mathbf{X}^\top\mathbf{Y} \in \mathbb{R}^2$ sont deux vecteurs gaussiens ; ils sont de plus décorrélés puisque

$$\mathbb{E}_\theta\left[(\mathrm{I}_n - \Pi)(\mathbf{Y} - \beta_1\mathbf{1})(\mathbf{Y}^\top - \beta_1\mathbf{1}^\top)\mathbf{X}\right] = \sigma^2(\mathrm{I}_n - \Pi)\mathbf{X} = 0_{n\times 2} \ .$$

Par suite, ils sont indépendants et les deux termes du produit dans l'expression de F le sont aussi.

6. Pour garantir le niveau du test, nous cherchons k_α tel que

$$\sup_{\theta\in\Theta_0} \mathbb{P}_\theta(F \geq k_\alpha) \leq \alpha \ .$$

Il suffit de définir k_α comme le quantile d'ordre $1-\alpha$ d'une loi de Fisher à $(1, n-2)$ degrés de liberté. Toute valeur k'_α **supérieure** à k_α convient aussi (même si une étude de la puissance du test devrait révéler que le choix du quantile est le meilleur).

Solution . 1. Sous $\mathbb{P}_{n,\theta}$, les v.a. $\{X_i, i \leq n\}$ sont i.i.d. donc pour tout $\theta \in \Theta$, la vraisemblance est donnée par

$$\theta \mapsto \prod_{i=1}^n p_\theta(X_i) = \frac{a^n}{\beta^{an}} \prod_{i=1}^n \left(X_i^{a-1}\mathbb{1}_{[0,\beta]}(X_i)\right) = \frac{a^n}{\beta^{an}}\left(\prod_{i=1}^n X_i^{a-1}\right)\mathbb{1}_{[0,\beta]}(X_{n:n}) \ .$$

De façon équivalente, la log-vraisemblance vaut

$$\theta \mapsto \ell_n(\theta) := \begin{cases} n\ln a - na\ln\beta + (a-1)\sum_{i=1}^n \ln X_i & \text{si } \beta \geq X_{n:n} \\ -\infty & \text{sinon.} \end{cases}$$

Vue comme une fonction de la variable β, nous avons $\ell_n(a,\beta) \leq \ell_n(a, X_{n:n})$ pour tout $a > 0$, en observant que $\beta \mapsto -na\ln\beta$ est décroissante sur $[X_{n:n}, +\infty[$ pour tout $a > 0$. Il vient que $\hat{\beta}_n = X_{n:n}$. La relation $\ell_n(a,\beta) \leq \ell_n(a, X_{n:n})$ pour tout $a > 0$, montre qu'il suffit de maximiser en a la fonction $a \mapsto \ell_n(a, X_{n:n})$. Un calcul de dérivée première donne

$$a \mapsto \frac{n}{a} - n\ln X_{n:n} + \sum_{i=1}^n \ln X_i \ .$$

La dérivée seconde est négative sur $\mathbb{R}_+^*$ et il est aisé de vérifier que $\lim_{a\to 0^+} \ell_n(a, X_{n:n}) = \lim_{a\to+\infty} \ell_n(a, X_{n:n}) = -\infty$. Par suite, le zero de la dérivée première est le maximum de $a \mapsto \ell_n(a, X_{n:n})$ et on trouve

$$\hat{a}_n := \frac{1}{n^{-1}\sum_{i=1}^n \ln(X_{n:n}/X_i)}.$$

2. $\hat{\beta}_n \geq 0$, donc pour tout $t \geq 0$ nous calculons

$$\begin{aligned}\mathbb{P}_{n,\theta}(\hat{\beta}_n \leq t) &= \mathbb{P}_{n,\theta}(X_1 \leq t, \ldots, X_n \leq t) = (\mathbb{P}_{n,\theta}(X_1 \leq t))^n \\ &= \left(\frac{a}{\beta^a}\int_0^{\beta\wedge t} x^{a-1}\mathrm{d}x\right)^n = \frac{(t\wedge\beta)^{an}}{\beta^{an}} \;.\end{aligned}$$

On en déduit que

$$\mathbb{P}_{n,\theta}(\hat{\beta}_n \leq t) = \begin{cases} 0 & \text{si } t \leq 0 \\ \frac{t^{an}}{\beta^{an}} & \text{si } t \in [0,\beta] \\ 1 & \text{sinon.}\end{cases}$$

3. Nous avons $\mathbb{P}_{n,\theta}(\hat{\beta}_n \leq \beta) = 1$ donc $\mathbb{P}_{n,\theta}(\ln(\beta/\hat{\beta}_n) \geq 0) = 1$. Soit $\varepsilon > 0$; nous avons

$$\mathbb{P}_{n,\theta}\left(\sqrt{n}\left|\ln(\hat{\beta}_n/\beta)\right| \geq \varepsilon\right) = \mathbb{P}_{n,\theta}\left(-\sqrt{n}\ln(\hat{\beta}_n/\beta) \geq \varepsilon\right) = \mathbb{P}_{n,\theta}\left(\hat{\beta}_n \leq \beta\exp(-\varepsilon/\sqrt{n})\right) \;.$$

Puisque $-\varepsilon/\sqrt{n} < 0$, nous avons $0 \leq \beta\exp(-\varepsilon/\sqrt{n}) < \beta$; nous déduisons de la question (2) que

$$\mathbb{P}_{n,\theta}\left(\sqrt{n}\left|\ln(\hat{\beta}_n/\beta)\right| \geq \varepsilon\right) = \exp(-\varepsilon an/\sqrt{n}) = \exp(-\varepsilon a\sqrt{n}) \;.$$

Le terme de droite tend vers zero quand $n \to +\infty$. Ainsi, pour tout $\varepsilon > 0$,

$$\lim_n \mathbb{P}_{n,\theta}\left(\sqrt{n}\left|\ln(\hat{\beta}_n/\beta)\right| \geq \varepsilon\right) = 0$$

ce qui établit la convergence en probabilités.

4. Soit h une fonction mesurable positive. Nous avons

$$\begin{aligned}\mathbb{E}\left[h(\log(\beta/X_1))\right] &= \int_0^\beta \frac{a}{\beta^a} h(\log(\beta/x))\, x^{a-1}\mathrm{d}x \\ &= \int_0^{+\infty} \frac{a}{\beta^a} h(y)\, \beta^{a-1}\exp(-(a-1)y)\,\beta\exp(-y)\mathrm{d}y \\ &= a\int_0^{+\infty} h(y)\exp(-ay)\mathrm{d}y.\end{aligned}$$

en ayant fait le changement de variable $y \leftarrow \ln(\beta/x)$ soit aussi $x = \beta\exp(-y)$. On reconnaît une loi Gamma$(1,a)$ ou de façon équivalente, une loi exponentielle de paramètre a.

5. Sous $\mathbb{P}_{n,\theta}$, les v.a. $\{\ln(\beta/X_i), i \leq n\}$ sont i.i.d. Gamma(1,a). Leur espérance est $1/a$ et leur variance est $1/a^2$. Par suite, le TCL pour des v.a. i.i.d. possédant un moment d'ordre 2 entraine

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n \ln(\beta/X_i) - \frac{1}{a}\right) \overset{\mathbb{P}_{n,\theta}}{\Longrightarrow} \mathrm{N}(0, 1/a^2).$$

6. Nous écrivons

$$\frac{1}{\hat{a}_n} = \frac{1}{n}\sum_{i=1}^{n} \ln(\hat{\beta}_n/X_i) = \ln(\hat{\beta}_n/\beta) + \frac{1}{n}\sum_{i=1}^{n} \ln(\beta/X_i)$$

puis

$$\sqrt{n}\left(\frac{1}{\hat{a}_n} - \frac{1}{a}\right) = \sqrt{n}\ln(\hat{\beta}_n/\beta) + \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}\ln(\beta/X_i) - \frac{1}{a}\right) .$$

Le second terme converge en loi vers $\mathrm{N}(0, 1/a^2)$ et le premier terme converge en probabilités vers zero. On en déduit par le lemme de Slutsky que

$$\sqrt{n}\left(\frac{1}{\hat{a}_n} - \frac{1}{a}\right) \overset{\mathbb{P}_{n,\theta}}{\Longrightarrow} \mathrm{N}(0, 1/a^2).$$

Enfin, on applique la méthode delta avec la fonction $\phi : u \mapsto 1/u$, qui est de classe C^1 sur $\mathbb{R}_+$. On en déduit que

$$\sqrt{n}\,(\hat{a}_n - a) \overset{\mathbb{P}_{n,\theta}}{\Longrightarrow} \phi'(1/a)\ \mathrm{N}(0, 1/a^2) \equiv -a^2\ \mathrm{N}(0, 1/a^2) \equiv \mathrm{N}(0, a^2).$$

7. Notons z_r le quantile d'ordre r d'une loi $\mathrm{N}(0, 1)$. Nous venons de voir que

$$\mathbb{P}_{n,\theta}\left(-z_{1-\alpha/2} \leq a^{-1}\sqrt{n}\,(\hat{a}_n - a) \leq z_{1-\alpha/2}\right) \to 1 - \alpha$$

dont nous déduisons que

$$\mathbb{P}_{n,\theta}\left(a \in \left[\frac{\hat{a}_n}{1 + \frac{z_{1-\alpha/2}}{\sqrt{n}}}, \frac{\hat{a}_n}{1 - \frac{z_{1-\alpha/2}}{\sqrt{n}}}\right]\right) \to 1 - \alpha .$$

8. Soit $\theta \in \Theta$. Calculons la loi de $an(1 - \hat{\beta}_n/\beta)$ sous $\mathbb{P}_{n,\theta}$. Observons que $\mathbb{P}_{n,\theta}(0 \leq \hat{\beta}_n/\beta \leq 1) = 1$. Par suite, nous considérons $t \geq 0$:

$$\begin{aligned}\mathbb{P}_{n,\theta}\left(an(1 - \hat{\beta}_n/\beta) \leq t\right) &= \mathbb{P}_{n,\theta}\left(1 - \frac{t}{an} \leq \hat{\beta}_n/\beta\right) = \mathbb{P}_{n,\theta}\left(\beta - \frac{\beta t}{an} \leq \hat{\beta}_n\right) \\ &= 1 - \mathbb{P}_{n,\theta}\left(\hat{\beta}_n < \beta - \frac{\beta t}{an}\right) = \begin{cases} 1 - \left(1 - \frac{t}{an}\right)^{an} & \text{si } t \leq na \\ 1 & \text{sinon.}\end{cases}\end{aligned}$$

Nous avons utilisé la question (2) dans la dernière égalité. Déterminons la loi asymptotique : soit $t \geq 0$. Lorsque $n \to \infty$, la condition $t \leq an$ est vérifiée pour tout n assez grand, et nous avons d'autre part

$$\lim_n 1 - \left(1 - \frac{t}{an}\right)^{an} = 1 - \exp(-t).$$

Nous en déduisons que $an(1 - \hat{\beta}_n/\beta) \overset{\mathbb{P}_{n,\theta}}{\Longrightarrow} \mathcal{E}(1)$. La loi limite est indépendante de θ : G_n est asymptotiquement pivotale.

Nous avons établi que $\{\hat{a}_n, n \geq 1\}$ est asymptotiquement normale (voir question (6)), ce qui entraine (par Slutsky) qu'elle est consistante : $\hat{a}_n/a \overset{\mathbb{P}_{n,\theta}-\mathrm{prob}}{\longrightarrow} 1$. Par suite, le lemme de Slutsky entraine que

$$n\hat{a}_n(1 - \hat{\beta}_n/\beta) = \frac{\hat{a}_n}{a}\ an(1 - \hat{\beta}_n/\beta) \overset{\mathbb{P}_{n,\theta}}{\Longrightarrow} \mathcal{E}(1).$$

9. Notons q_r le quantile d'ordre r d'une loi exponentielle de paramètre 1. Nous venons d'établir que
$$\lim_n \mathbb{P}_{n,\theta}\left(n\hat{a}_n(1-\hat{\beta}_n/\beta) \leq q_{1-\alpha}\right) = 1-\alpha$$
dont nous déduisons que
$$\lim_n \mathbb{P}_{n,\theta}\left(\beta \leq \frac{\hat{\beta}_n}{1-\frac{q_{1-\alpha}}{n\hat{a}_n}}\right) = 1-\alpha\ .$$

Solution . 1. Nous avons
$$R_\rho(g,r) = \mathbb{E}\left[\mathbb{1}_{\{g(X)\neq Y\}}\mathbb{1}_{\{0\}}(r(X))\right] + \rho\mathbb{E}\left[\mathbb{1}_{\{1\}}(r(X))\right]\ . \tag{7}$$
Le second terme est égal à $\rho\mathbb{P}(r(X)=1)$. Travaillons le premier terme, que nous reécrivons
$$\mathbb{E}\left[\mathbb{1}_{\{0\}}(g(X))\mathbb{1}_{\{1\}}(Y)\mathbb{1}_{\{0\}}(r(X)) + \mathbb{1}_{\{1\}}(g(X))\mathbb{1}_{\{0\}}(Y)\mathbb{1}_{\{0\}}(r(X))\right]\ .$$
Par définition de l'espérance conditionnelle, nous savons que pour toutes fonctions mesurables positives h_1, h_2
$$\mathbb{E}\left[h_1(X)h_2(Y)\right] = \mathbb{E}\left[\mathbb{E}\left[h_2(Y)\,|\,X\right]h_1(X)\right].$$
En appliquant cette relation avec $h_2(Y) \leftarrow \mathbb{1}_{\{1\}}(Y)$ et $h_1(X) \leftarrow \mathbb{1}_{\{0\}}(g(X))\mathbb{1}_{\{0\}}(r(X))$, nous obtenons
$$\mathbb{E}\left[\mathbb{1}_{\{1\}}(Y)\mathbb{1}_{\{0\}}(g(X))\mathbb{1}_{\{0\}}(r(X))\right] = \mathbb{E}\left[\eta(X)\mathbb{1}_{\{0\}}(g(X))\mathbb{1}_{\{0\}}(r(X))\right]\ .$$
De même, nous démontrons que
$$\mathbb{E}\left[\mathbb{1}_{\{1\}}(g(X))\mathbb{1}_{\{0\}}(Y)\mathbb{1}_{\{0\}}(r(X))\right] = \mathbb{E}\left[(1-\eta(X))\mathbb{1}_{\{1\}}(g(X))\mathbb{1}_{\{0\}}(r(X))\right]\ .$$
Par suite le premier terme de (7) est égal à
$$\mathbb{E}\left[\left\{\eta(X)\mathbb{1}_{\{0\}}(g(X)) + (1-\eta(X))\mathbb{1}_{\{1\}}(g(X))\right\}\mathbb{1}_{\{0\}}(r(X))\right]\ .$$

2. Notons tout d'abord que pour tout $x\in\mathbb{R}^d$,
$$\eta(x)\mathbb{1}_{\{0\}}(g(x)) + (1-\eta(x))\mathbb{1}_{\{1\}}(g(X)) \geq h(x)\ . \tag{8}$$
Nous savons que $\eta\wedge(1-\eta) = \eta$ ssi $\eta \leq 1/2$ et que $\eta\wedge(1-\eta) = 1-\eta$ ssi $\eta\geq 1/2$. Par conséquent,
$$\eta(x)\mathbb{1}_{\{0\}}(g_*(x)) + (1-\eta(x))\mathbb{1}_{\{1\}}(g_*(x)) = h(x)$$
en définissant la règle g_* par
$$g_*(x) := \mathbb{1}_{\{[1/2,1]\}}(\eta(x))\ .$$
Dans ce cas, nous avons
$$R_\rho(g_*,r) = \mathbb{E}[\mathbb{1}_{\{0\}}(r(X))\ h(X)] + \rho\ \mathbb{P}(r(X)=1)\ .$$
En utilisant (3) et (8) nous avons, pour tout $\rho\in[0,1]$ et toute fonction r,
$$R_\rho(g,r) \geq \mathbb{E}[\mathbb{1}_{\{0\}}(r(X))\ h(X)] + \rho\ \mathbb{P}(r(X)=1) = R_\rho(g_*,r)\ .$$

3. On écrit $\mathbb{P}(r(X)=1) = 1 - \mathbb{P}(r(X)=0) = 1 - \mathbb{E}\left[\mathbb{1}_{\{0\}}(r(X))\right]$, ce qui entraine
$$R_\rho(g_*, r) = \mathbb{E}[\mathbb{1}_{\{0\}}(r(X))\, h(X)] + \rho\, \left(1 - \mathbb{E}\left[\mathbb{1}_{\{0\}}(r(X))\right]\right) = \rho + \mathbb{E}[\mathbb{1}_{\{0\}}(r(X))\ (h(X) - \rho)]\ .$$
4. Lorsque $\rho = 0$, nous avons
$$R_\rho(g_*, r) = \mathbb{E}[\mathbb{1}_{\{0\}}(r(X))\ h(X)].$$
Cette quantité peut être égale à zero en prenant $r_*(x) = 1$ pour tout $x \in \mathbb{R}^d$: la règle de classification s'abstient toujours.
5. Lorsque $\rho \geq 1/2$, nous avons
$$h(X) - \rho \leq 0$$
en observant que $h(x) \leq 1/2$. Par suite, nous pouvons rendre le risque minimal en prenant $r(x) = 0$ pour tout $x \in \mathbb{R}^d$: la règle de classification se prononce toujours.
6. Soit $\rho \in \,]0, 1/2[$. Dans ce cas $h(x) - \rho$ peut être positif ou négatif selon la valeur de $\eta(x)$. Par suite, nous proposons la stratégie
$$r_*(x) = 0 \quad \text{ssi} \quad h(x) - \rho \leq 0\ .$$
Vérifions que cette stratégie est optimale :
$$R_\rho(g_*, r) - R_\rho(g_*, r_*) = \mathbb{E}[\left(\mathbb{1}_{\{0\}}(r(X)) - \mathbb{1}_{\{h(X)\leq\rho\}}\right)\ \{h(X) - \rho\}]$$
Sur $\{h(X) > \rho\}$, il vient
$$\left(\mathbb{1}_{\{0\}}(r(X)) - \mathbb{1}_{\{h(X)\leq\rho\}}\right)\ \{h(X) - \rho\} = \mathbb{1}_{\{0\}}(r(X))\ \{h(X) - \rho\} \geq 0$$
et sur $\{h(X) \leq \rho\}$, il vient
$$\left(\mathbb{1}_{\{0\}}(r(X)) - \mathbb{1}_{\{h(X)\leq\rho\}}\right)\ \{h(X) - \rho\} = \left(\mathbb{1}_{\{0\}}(r(X)) - 1\right)\ \{h(X) - \rho\} \geq 0$$
dont on déduit que $R_\rho(g_*, r) - R_\rho(g_*, r_*) \geq 0$.

Solution . Dans la correction, nous omettons Z quand il n'y a pas d'ambiguïté, i.e. nous notons $\mathbb{E}_\theta[\phi]$ pour $\mathbb{E}_\theta[\phi(Z)]$.

1. Pour $\theta > \theta_0$,
$$\begin{aligned}\mathbb{E}_\theta[S(Z)\mathbb{1}_{\{X_{n:n}\leq\theta_0\}}] &= \sum_{x_1,\ldots,x_n\in\{1,\ldots,\theta_0\}^n} S(x_1,\ldots,x_n)\theta^{-n} \\ &= \sum_{x_1,\ldots,x_n\in\{1,\ldots,\theta_0\}^n} S(x_1,\ldots,x_n)\theta_0^{-n}(\theta_0/\theta)^n\ .\end{aligned}$$
2. Soit $\theta \leq \theta_0$. Nous avons $\mathbb{E}_\theta[\phi_{\mathrm{m}}] = \alpha\mathbb{P}_\theta(X_{n:n} \leq \theta_0) + \mathbb{P}_\theta(X_{n:n} > \theta_0)$. Puisque $\theta \leq \theta_0$, nous avons $\mathbb{P}_\theta(X_{n:n} \leq \theta_0) = 1$ et $\mathbb{P}_\theta(X_{n:n} > \theta_0) = 0$. Par suite
$$\sup_{\theta\leq\theta_0} \mathbb{E}_\theta[\phi_{\mathrm{m}}] = \alpha.$$

3. Soit ϕ un test de niveau α.

$$\begin{aligned}\mathbb{E}_\theta[\phi_{\mathrm{m}}] - \mathbb{E}_\theta[\phi] &= \mathbb{E}_\theta\left[(\phi_{\mathrm{m}} - \phi)\, \mathbb{1}_{\{\{X_{n:n} > \theta_0\}\}}\right] + \mathbb{E}_\theta\left[(\phi_{\mathrm{m}} - \phi)\, \mathbb{1}_{\{X_{n:n} \leq \theta_0\}}\right] \\ &= \mathbb{E}_\theta\left[(1-\phi)\mathbb{1}_{\{X_{n:n} > \theta_0\}}\right] + \mathbb{E}_\theta\left[(\alpha - \phi)\mathbb{1}_{\{X_{n:n} \leq \theta_0\}}\right] \\ &\geq \mathbb{E}_\theta\left[(\alpha - \phi)\mathbb{1}_{\{X_{n:n} \leq \theta_0\}}\right] \\ &= \mathbb{E}_{\theta_0}\left[(\alpha - \phi)\mathbb{1}_{\{X_{n:n} \leq \theta_0\}}\right](\theta_0/\theta)^n \\ &= \mathbb{E}_{\theta_0}[\alpha - \phi]\,(\theta_0/\theta)^n \geq 0\end{aligned}$$

Nous avons utilisé la question 1, et utilisé que ϕ est de niveau α (ce qui entraine $\mathbb{E}_{\theta_0}[\phi(Z)] \leq \alpha$).

4. Calculons le niveau du test. En utilisant que $\theta_0\alpha^{1/n} \in \mathbb{N}^*$, nous avons

$$\mathbb{E}_{\theta_0}[\phi_{\mathrm{b}}] = \mathbb{P}_{\theta_0}(X_{n:n} \leq \theta_0\alpha^{1/n}) = \left(\frac{\theta_0\,\alpha^{1/n}}{\theta_0}\right)^n = \alpha.$$

5.

$$\begin{aligned}\mathbb{E}_\theta[\phi_{\mathrm{b}}] - \mathbb{E}_\theta(\phi) &= \mathbb{E}_\theta[(1-\phi)\mathbb{1}_{\{X_{n:n} > \theta_0\}}] + \mathbb{E}_\theta\left[(\phi_{\mathrm{b}} - \phi)\, \mathbb{1}_{\{X_{n:n} \leq \theta_0\}}\right] \\ &\geq \mathbb{E}_\theta\left[(\phi_{\mathrm{b}} - \phi)\, \mathbb{1}_{\{X_{n:n} \leq \theta_0\}}\right] = \mathbb{E}_\theta\left[\phi_{\mathrm{b}}\mathbb{1}_{\{X_{n:n} \leq \theta_0\,\alpha^{1/n}\}} - \phi\mathbb{1}_{\{X_{n:n} \leq \theta_0\}}\right] \\ &= \mathbb{P}_\theta\left(X_{n:n} \leq \theta_0\alpha^{1/n}\right) - \mathbb{E}_\theta[\phi\mathbb{1}_{\{X_{n:n} \leq \theta_0\}}] \\ &= \alpha\,(\theta_0/\theta)^n - \mathbb{E}_{\theta_0}[\phi\mathbb{1}_{\{X_{n:n} \leq \theta_0\}}]\,(\theta_0/\theta)^n = \alpha\,(\theta_0/\theta)^n - \mathbb{E}_{\theta_0}[\phi]\,(\theta_0/\theta)^n \geq 0\end{aligned}$$

Nous avons utilisé dans la dernière ligne : la question 1 et le fait que ϕ est de niveau α (donc $\mathbb{E}_{\theta_0}[\phi] \leq \alpha$).

6. Pour $\theta < \theta_0$, nous avons

$$\mathbb{E}_\theta[\phi_{\mathrm{b}}] - \mathbb{E}_\theta[\phi] = \mathbb{P}_\theta\left(X_{n:n} \leq \theta_0\alpha^{1/n}\right) + \mathbb{P}_\theta\left(X_{n:n} > \theta_0\right) - \mathbb{E}_\theta[\phi]\,.$$

et nous concluons en remarquant que pour $\theta < \theta_0$, $\mathbb{P}_\theta\left(X_{n:n} > \theta_0\right) = 0$.

7. Si $\theta \leq \theta_0\alpha^{1/n}$ alors $\mathbb{P}_\theta\left(X_{n:n} \leq \theta_0\alpha^{1/n}\right) - \mathbb{E}_\theta[\phi] = 1 - \mathbb{E}_\theta[\phi] \geq 0$.
Si $\theta_0\alpha^{1/n} \leq \theta \leq \theta_0$, alors

$$\mathbb{P}_\theta\left(X_{n:n} \leq \theta_0\alpha^{1/n}\right) = \left(\frac{\theta_0\,\alpha^{1/n}}{\theta}\right)^n\,.$$

D'autre part, puisque $\phi \geq 0$ et que le test est de niveau α

$$\begin{aligned}\mathbb{E}_\theta\left[\phi\right] &= \frac{1}{\theta_0^n}\left(\frac{\theta_0}{\theta}\right)^n \sum_{(x_1,\ldots,x_n)\in\{1,\ldots,\theta\}^n} \phi(x_1,\ldots,x_n) \\ &\leq \frac{1}{\theta_0^n}\left(\frac{\theta_0}{\theta}\right)^n \sum_{(x_1,\ldots,x_n)\in\{1,\ldots,\theta_0\}^n} \phi(x_1,\ldots,x_n) \\ &\leq \left(\frac{\theta_0}{\theta}\right)^n \mathbb{E}_{\theta_0}\left[\phi\right] \leq \alpha\left(\frac{\theta_0}{\theta}\right)^n\end{aligned}$$

Par conséquent, $\mathbb{E}_\theta[\phi_{\mathrm{b}}] - \mathbb{E}_\theta[\phi] \geq 0$.
En conclusion, nous avons montré que pour tout $\theta \neq \theta_0$, $\mathbb{E}_\theta[\phi_{\mathrm{b}}] - \mathbb{E}_\theta[\phi] \geq 0$; et ϕ_{b} est de niveau α. Il est donc UPP$[\alpha]$.