

Rappels de statistique mathématique  
*Corrigé des travaux dirigés n° 5*

Guillaume Lacôte  
Bureau E03

✉ [Guillaume.Lacote@ensae.fr](mailto:Guillaume.Lacote@ensae.fr)

☞ <http://ensae.no-ip.com/SE222/>

Corrigé de l'exercice 1

☞ Q1 Par définition on a pour  $y \in \mathbb{R}^+$

$$\begin{aligned} f(y; a, b) &= \frac{\partial F}{\partial \cdot}(y, a, b) \\ &= aby^{b-1}e^{-ay^b} \end{aligned}$$

Il en découle que

$$h(y; a, b) = aby^{b-1}\mathbf{1}_{y \geq 0}$$

☞ Q2 La fonction de hasard  $h(y)$  s'écrit comme le rapport  $\frac{f(y)}{S(y)}$ . Or

$$\begin{aligned} f(y) &= \frac{\partial F(\cdot; a, b)}{\partial y}(y) \\ &= \lim_{dy \rightarrow 0^+} \frac{\mathbb{P}(Y \in [y, y + dy])}{dy} \end{aligned}$$

et par ailleurs  $S(y) = 1 - F(y) = \mathbb{P}(Y \geq y)$ .

Autrement dit

$$\begin{aligned} h(y) &= \lim_{dy \rightarrow 0^+} \frac{1}{dy} \frac{\mathbb{P}(Y \in [y, y + dy])}{\mathbb{P}(Y \geq y)} \\ &= \lim_{dy \rightarrow 0^+} \frac{1}{dy} \frac{\mathbb{P}(Y \in [y, y + dy] \wedge Y \geq y)}{\mathbb{P}(Y \geq y)} \\ &= \lim_{dy \rightarrow 0^+} \frac{1}{dy} \mathbb{P}(\text{NON}(Y \geq y + dy) \mid Y \geq y) \end{aligned}$$

En d'autres termes  $h(y)dy$  mesure la probabilité, pour un individu encore au chômage à la date  $y$ , de **sortir du chômage** peu après  $y$  (au sens "après  $y + dy$ ", ceci pour  $dy$  "voisin" de 0). Supposer que  $h$  est constante revient à supposer que la probabilité de sortir d'une période chômage de durée  $dy$  ne dépend pas de sa date  $y$ , c'est-à-dire qu'il n'y a pas d'*effet mémoire*. Sachant que

$$\begin{aligned} h(y) &= \frac{aby^{b-1}e^{-ay^b}}{e^{-ay^b}} \\ &= aby^{b-1} \end{aligned}$$

cette condition est vérifiée **si, et seulement si**  $b = 1$ .

Lorsque cette condition est vérifiée, on constate que la loi de durée du chômage s'écrit simplement pour  $y \geq 0$

$$f(y; a, 1) = ae^{-ay}$$

dont on remarque qu'il s'agit d'une loi exponentielle de paramètre  $a$ .

☞ Q3 Etudions les variations de  $h(y; \cdot, \cdot)$  en fonction de  $a$  et  $b$  à  $y \geq 0$  donné. Il vient

– Selon  $a$  :

$$\begin{aligned} \frac{\partial h}{\partial a}(y, a, b) &= by^{b-1} \\ &> 0 \quad \text{car } b > 0 \end{aligned}$$

Ainsi la situation du chômage s'améliore (au sens où la probabilité de sortie s'accroît) lorsque  $a$  s'accroît.

– Selon  $b$  :

$$\begin{aligned} \frac{\partial h}{\partial b}(y, a, b) &= ay^{b-1} + ab \ln(y)y^{b-1} \\ &= ay^{b-1} (1 + b \ln(y)) \end{aligned}$$

En conséquence

– Si  $y \in ]0, e^{-\frac{1}{b}}[$ , soit si  $b > -\frac{1}{\ln(y)}$ , la situation du chômage *se détériore* lorsque  $b$  croît.

– Si  $y \in ]e^{-\frac{1}{b}}, +\infty[$ , soit si  $b < -\frac{1}{\ln(y)}$ , la situation du chômage *s'améliore* lorsque  $b$  croît.

On suppose dans cette partie  $b = 1$ . Le modèle est alors uniquement paramétré par  $a$ .

☞ Q1 La vraisemblance du modèle est donnée pour  $(y_1, \dots, y_n) \in \mathbb{R}^n$  par

$$\begin{aligned} L(y; a, 1) &= \prod_{i=1}^n (ae^{-ay_i} \mathbf{1}_{y_i \geq 0}) \\ &= (a^n \mathbf{1}_{\min_i y_i \geq 0}) \cdot e^{-a(\sum_{i=1}^n y_i)} \end{aligned}$$

On constate donc que le modèle est exponentiel, de statistique exhaustive  $T(Y) = \sum_{i=1}^n Y_i$ .

☞ Q2 La log-vraisemblance est alors pour  $(y_1, \dots, y_n) \in \mathbb{R}^{+n}$

$$\ln L(y; a, 1) = n \ln a - a \sum_{i=1}^n y_i$$

Le score s'écrit donc

$$\begin{aligned} S_a(y; a, 1) &= \frac{\partial \ln L}{\partial a}(y; a, 1) \\ &= \frac{n}{a} - \sum_{i=1}^n y_i \end{aligned}$$

Or

$$\begin{aligned} \mathbb{E}(y_i) &= \int_{\mathbb{R}} y_i \cdot ae^{-ay_i} \mathbf{1}_{y_i \geq 0} dy_i \\ &= - \int_0^{+\infty} (-au) e^{-au} du \\ &= + \frac{1}{a} [e^{-au}]_0^{+\infty} \\ &= \frac{1}{a} \end{aligned}$$

Il s'ensuit que

$$\mathbb{E}(S_a(y; a, 1)) = \frac{n}{a} - \sum_{i=1}^n \mathbb{E}(y_i) = 0$$

de sorte que le vecteur score est bien centré (ce qui est une propriété générale, voir TD exercice 2 )

☞ Q3 De la condition nécessaire  $\frac{\partial \ln L}{\partial a}(\hat{a}_0) = 0$  on tire  $\hat{a}_0 = \frac{1}{\bar{y}}$ ; réciproquement  $\hat{a}_0$  maximise bien log-vraisemblance car celle-ci est concave.

Par ailleurs, on a

$$\begin{aligned} \mathbb{E}(\hat{a}_0) &= \mathbb{E}\left(\frac{1}{\bar{y}}\right) \\ &> \frac{1}{\mathbb{E}(\bar{y})} \quad (\text{d'après l'inégalité de Jensen}) \\ &= \frac{1}{a} \end{aligned}$$

Autrement dit,  $\hat{a}_0$  est biaisé et surestime systématiquement la vraie valeur  $a$ .

On pourrait même calculer le biais exact, en s'appuyant sur le fait que  $(Y_1 + \dots + Y_n) \sim \Gamma(n, a)$  en calculant explicitement  $\int_{\mathbb{R}} \frac{n}{s} f_{\Gamma(a)}(s) ds = \frac{n-1}{n-1} a$ ; on en déduirait que  $\frac{n-1}{n} \hat{a}_0$  est un estimateur sans biais de  $a$ .

☞ Q4 On a

$$I_n(a) = -\mathbb{E}\left(\frac{\partial^2 \ln L}{\partial a^2} = -\frac{n}{a^2}\right)$$

et donc

$$\mathbb{V}_{as}(\hat{a}_0) = \frac{a^2}{n}$$

(Attention, c'est bien  $I_n(a)^{-1}$  et non pas  $I_1(a)^{-1}$  car on étudie la loi asymptotique de  $(\hat{a}_0 - \mathbb{E}(\hat{a}_0))$  et non pas celle de  $\sqrt{n}(\hat{a}_0 - \mathbb{E}(\hat{a}_0))$ )

☞ Q1 La vraisemblance du modèle est donnée pour  $(y_1, \dots, y_n) \in \mathbb{R}^n$  par

$$\begin{aligned} L(y; a, b) &= \prod_{i=1}^n (aby_i^{b-1} e^{-ay_i^b} \mathbf{1}_{y_i \geq 0}) \\ &= ((ab)^n \mathbf{1}_{\min_i y_i \geq 0}) \cdot \left(\prod_{i=1}^n y_i^{b-1}\right) \cdot e^{-\sum_{i=1}^n (a \cdot y_i^b)} \end{aligned}$$

On constate donc que le modèle n'est **pas** exponentiel, car on ne peut dissocier les termes  $y_i$  du paramètre  $(a, b)$  là où ils interviennent (à savoir  $\sum_{i=1}^n y_i^b$ ). Donc en vertu du théorème de factorisation  $(Y_1, \dots, Y_n)$  est exhaustive et minimale, et il n'existe donc pas de statistique exhaustive comportant moins de  $n$  composantes.

☞ Q2 La log-vraisemblance s'écrit pour  $(y_1, \dots, y_n) \in \mathbb{R}^{+n}$

$$\ln L(y; a, b) = n \ln a + n \ln b + (b-1) \sum_{i=1}^n \ln y_i - a \sum_{i=1}^n y_i^b$$

On a

$$\begin{cases} \frac{\partial \ln L}{\partial a}(y; a, b) &= \frac{n}{a} - \sum_{i=1}^n y_i^b \\ \frac{\partial \ln L}{\partial b}(y; a, b) &= \frac{n}{b} + \sum_{i=1}^n \ln y_i - a \sum_{i=1}^n \ln y_i y_i^b \end{cases}$$

et donc l'estimateur  $(\hat{a}, \hat{b})$  de  $(a, b)$  vérifie

$$\begin{cases} \hat{a} &= \frac{n}{\sum_{i=1}^n y_i^{\hat{b}}} \\ \hat{a} \sum_{i=1}^n \ln y_i \hat{y}_i^{\hat{b}} - \frac{n}{\hat{b}} &= \sum_{i=1}^n \ln y_i \end{cases}$$

On constate que  $\hat{b}$ , et par suite  $\hat{a}$ , n'est pas exprimable sous forme analytique. Pour autant l'étude de la fonction  $\left( \begin{matrix} \mathbb{R}^+ & \rightarrow & \mathbb{R} \\ b & \mapsto & \frac{1}{b} + \alpha - \frac{\sum_{i=1}^n \beta_i y_i^b}{\sum_{i=1}^n y_i^b} \end{matrix} \right)$  assure de l'existence et de l'unicité de  $\hat{b}$ , et donc de  $\hat{a}$ .

☞ Q3 On a successivement

$$\begin{aligned} \frac{\partial^2 \ln L}{\partial a^2}(y; a, b) &= -\frac{n}{a^2} \\ \frac{\partial^2 \ln L}{\partial a \partial b}(y; a, b) &= -\sum_{i=1}^n (\ln y_i y_i^b) \\ \frac{\partial^2 \ln L}{\partial b^2}(y; a, b) &= -\frac{n}{b^2} - a \sum_{i=1}^n ((\ln y_i)^2 y_i^b) \end{aligned}$$

En conséquence

$$\mathbb{V}_{as} \left( \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} \right) = \begin{pmatrix} n \mathbb{E}(y_1^b \ln y_1) & n \mathbb{E}(y_1^b \ln y_1) \\ n \mathbb{E}(y_1^b \ln y_1) & \frac{n}{b^2} + n a \mathbb{E}((\ln y_1)^2 y_1^b) \end{pmatrix}^{-1}$$

Il reste alors à exprimer  $\mathbb{E}(y_1^b \ln(y_1)^i)$  pour  $i \in \{1, 2\}$ ; ce calcul peut être fait au moyen des intégrales eulériennes.

Remarquons pour ce faire que si  $Z$  suit une loi de Weibull de paramètre  $(\alpha, \beta)$ , de densité  $f_{\alpha, \beta}(x) = \alpha \theta x^{\alpha-1} e^{-\theta x^\alpha} \mathbf{1}_{x>0}$ , alors pour  $p \in \mathbb{N}^*$  on a

$$\begin{aligned} \mathbb{E}(Z^\alpha \ln(Z)^p) &= \int_0^{+\infty} \theta z^\alpha \ln^p(z) \alpha z^{\alpha-1} e^{-\theta z^\alpha} dz \\ &= \frac{1}{\alpha^p \theta} \int_0^{+\infty} \ln^p\left(\frac{u}{\theta}\right) u e^{-u} du \quad \text{en posant } u = \theta z^\alpha \end{aligned}$$

En particulier pour  $p \in \{1, 2\}$  on a

$$\begin{aligned} \mathbb{E}(Z^\alpha \ln(Z)) &= \frac{1}{\alpha \theta} \int_0^{+\infty} \ln\left(\frac{u}{\theta}\right) u e^{-u} du \\ \mathbb{E}(Z^\alpha (\ln Z)^2) &= \frac{1}{\alpha^2 \theta} \int_0^{+\infty} \ln^2\left(\frac{u}{\theta}\right) u e^{-u} du \end{aligned}$$

Or ( TD 2, exercice 2 )

$$\begin{aligned} \mathbb{E}(X^\alpha \ln(X)) &= \frac{1 - \gamma - \ln \theta}{\theta \alpha} \\ \mathbb{E}(X^\alpha (\ln X)^2) &= \frac{1}{\alpha^2 \theta} \left( \frac{\pi^2}{6} - 2\gamma + \gamma^2 - 2(1 - \gamma) \ln(\theta) + \ln(\theta)^2 \right) \end{aligned}$$

Ceci permet alors de conclure en posant  $\theta = a$  et  $\alpha = b$  que<sup>1</sup>

$$\mathbb{V}_{as} \left( \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} \right) = \begin{pmatrix} \frac{n}{a^2} & \frac{n^{1-\gamma-\ln a}}{ab} \\ \frac{n^{1-\gamma-\ln a}}{ab} & \frac{n}{b^2} \left( 1 + \frac{\pi^2}{6} - 2\gamma + \gamma^2 - 2(1 - \gamma) \ln a + (\ln a)^2 \right) \end{pmatrix}^{-1}$$

☞ Q4 Remarquons que pour toute matrice  $A$ ,  $A^{-1} = \frac{1}{|A|} \text{Com}(A)'$  où  $\text{Com}(A)$  désigne la comatrice  $A$ ; par suite

$$I_n(a, 1)^{-1} = \frac{1}{|I_n(a, 1)|} \begin{pmatrix} n + na \mathbb{E}((\ln y_1)^2 y_1) & -n \mathbb{E}(y_1 \ln y_1) \\ -n \mathbb{E}(y_1 \ln y_1) & \frac{n}{a^2} \end{pmatrix}$$

Par ailleurs,

$$|I_n(a, 1)| = \frac{n}{a^2} (n + na \mathbb{E}((\ln y_1)^2 y_1)) - (n \mathbb{E}(y_1 \ln y_1))^2$$

En conséquence,

$$\begin{aligned} \mathbb{V}_{as}(\hat{a}) &= \frac{n + na \mathbb{E}((\ln y_1)^2 y_1)}{|I_n(a, 1)|} \\ &= \frac{n + na \mathbb{E}((\ln y_1)^2 y_1)}{\frac{n}{a^2} (n + na \mathbb{E}((\ln y_1)^2 y_1)) - (n \mathbb{E}(y_1 \ln y_1))^2} \\ &= \frac{a^2}{n} \left( 1 + \frac{a^2 (n \mathbb{E}(y_1 \ln y_1))^2}{(n + na \mathbb{E}((\ln y_1)^2 y_1)) - \frac{a^2}{n} (n \mathbb{E}(y_1 \ln y_1))^2} \right) \\ &\geq \frac{a^2}{n} \\ &= \mathbb{V}_{as}(\hat{a}_0) \end{aligned}$$

Ainsi estimer sous contrainte permet d'obtenir un estimateur *plus efficace* qu'estimer sans contrainte.

☞ Q5 Lorsque l'échantillon est de petit taille, une simulation permettrait d'étudier la distribution

$\left( \begin{matrix} \hat{a} \\ \hat{b} \end{matrix} \right)$ , de la façon suivante :

- Observant  $(y_1, \dots, y_n)$ , on en tire  $\left( \begin{matrix} \hat{a}^0 \\ \hat{b}^0 \end{matrix} \right)$  estimateur du maximum de vraisemblance  $\left( \begin{matrix} a^0 \\ b^0 \end{matrix} \right)$ .

<sup>1</sup>Bien que la tentation soit forte, nous résisterons ici au plaisir d'inverser cette matrice.

- Soit alors  $\mathcal{W}^0$  la loi de Weibull de paramètre  $\hat{a}^0, \hat{b}^0$ ; on tire alors un échantillon indépendant de taille  $N$  selon  $\mathcal{W}^0$ , soit  $(y^0_1, \dots, y^0_N)$ .
- de cet échantillon on tire alors  $\begin{pmatrix} \hat{a}^1 \\ \hat{b}^1 \end{pmatrix}$ , et ainsi de suite.

Bien-entendu, à la limite lorsque  $N \rightarrow +\infty$  on a  $\begin{pmatrix} \hat{a}^1 \\ \hat{b}^1 \end{pmatrix} = \begin{pmatrix} \hat{a}^0 \\ \hat{b}^0 \end{pmatrix}$ ; il en va de même lorsque le nombre d'itération devient arbitrairement grand.

L'idée serait donc plutôt de se limiter à une ou deux itérations, mais de **répéter le tirage** pour obtenir la distribution empirique de  $\begin{pmatrix} \hat{a}^1 \\ \hat{b}^1 \end{pmatrix}$ .

On considère maintenant le cas de  $T$  observations  $Y_1, \dots, Y_T$  indépendantes, de lois respectives :

$$F(y; e^{\alpha t}, 1), t \in [1, T], \alpha \in \mathbb{R}$$

☞ Q1 La vraisemblance du modèle est donnée pour  $(y_1, \dots, y_n) \in \mathbb{R}^n$  par

$$\begin{aligned} L(y; \alpha) &= \prod_{t=1}^T e^{\alpha t} e^{-e^{\alpha t} y_t} (\mathbf{1}_{y_t \geq 0}) \\ &= \mathbf{1}_{\inf_{t \in [1, T]} y_t \geq 0} e^{\alpha \sum_{t=1}^T t} e^{-\sum_{t=1}^T e^{\alpha t} y_t} \end{aligned}$$

Une condition nécessaire sur l'estimateur  $\hat{\alpha}$  de  $\alpha$  est donc

$$\sum_{t=1}^T (t y_t e^{t \hat{\alpha}}) = \frac{T(T+1)}{2}$$

condition dont on vérifie qu'elle est suffisante puisque la log-vraisemblance est strictement concave car de dérivée seconde

$$\frac{\partial^2 \ln L}{\partial \alpha^2}(y; \alpha) = - \sum_{t=1}^T t^2 y_t e^{\alpha t} < 0 \quad \forall \alpha \in \mathbb{R}^+$$

☞ Q2 On a

$$\begin{aligned} u_t &= y_t - e^{-\hat{\alpha} t} \\ &= y_t - \mathbb{E}_{|\alpha=\hat{\alpha}}(y_t) \end{aligned}$$

Ainsi  $u_t$  représente l'erreur de la prévision empirique de  $y_t$ , différence entre la réalisation réelle de  $y_t$  et la meilleure prévision compte-tenu du modèle estimé.

☞ Q3 On a

$$\begin{aligned} \frac{\partial \ln L}{\partial \alpha}(y, \alpha) &= \frac{T(T+1)}{2} - \sum_{t=1}^T t y_t e^{\alpha t} \\ &= \sum_{t=1}^T (t - t y_t e^{\alpha t}) \\ &= \sum_{t=1}^T t (1 - y_t e^{\alpha t}) \\ &= \sum_{t=1}^T t e^{\alpha t} (e^{-\alpha t} - y_t) \end{aligned}$$

de sorte que

$$\frac{\partial \ln L}{\partial \alpha}(y, \hat{\alpha}) = \sum_{t=1}^T e^{\hat{\alpha} t} t u_t$$

Posons donc

$$\langle \cdot | \cdot \rangle_{\rho} : \begin{pmatrix} (\mathbb{R}^T)^2 & \rightarrow & \mathbb{R}^+ \\ ((x_t)_t, (y_t)_t) & \mapsto & \sum_{t=1}^T \rho^t x_t y_t \end{pmatrix}$$

Alors pour tout  $\rho \in \mathbb{R}^{+*}$ ,  $\langle \cdot | \cdot \rangle_{\rho}$  est un produit scalaire, et en outre

$$\frac{\partial \ln L}{\partial \alpha}(y, \hat{\alpha}) = 0 \Leftrightarrow \langle (1, \dots, T) | u \rangle_{e^{\hat{\alpha}}} = 0$$

\*  
\* \*

Corrigé de l'exercice 2

☞ Q1 On suppose pour simplifier que jusqu'à la question 4,  $T = +\infty$ . Sinon, il faudrait alourdir expressions des densités qui suivent en adjoignant l'indicatrice  $\mathbf{1}_{t \leq T}$  adéquate.

On a pour  $t \in \llbracket 1, T-1 \rrbracket$

$$\begin{aligned}
 \mathbb{P}(T_i^* = t+1 | T_i > t) &= \mathbb{P}(\text{L'individu } i \text{ reçoit au moins une offre intéressante entre } t \text{ et } t+1) \\
 &= \mathbb{P}(N_i^t \geq 1 \wedge \text{Une des } N_i^t \text{ offres propose un salaire } \geq \xi_i) \\
 &= \sum_{k=1}^{+\infty} \mathbb{P}(N_i^t = k \wedge \text{Une des } k \text{ offres propose un salaire } \geq \xi_i) \\
 &= \sum_{k=1}^{+\infty} \mathbb{P}(N_i^t = k) \cdot \mathbb{P}(\text{Une des } k \text{ offres propose un salaire } \geq \xi_i) \\
 &= \sum_{k=1}^{+\infty} \mathbb{P}(N_i^t = k) \cdot (1 - \mathbb{P}(\text{Toutes les } k \text{ offres proposent un salaire } < \xi_i)) \\
 &= \sum_{k=1}^{+\infty} \left( \frac{\lambda^k}{k!} e^{-\lambda} \right) \cdot (1 - F(\xi_i)^k) \\
 &= \sum_{k=1}^{+\infty} \frac{\lambda^k}{k!} e^{-\lambda} - e^{-\lambda} \sum_{k=1}^{+\infty} \frac{(\lambda F(\xi_i))^k}{k!} \\
 &= e^{-\lambda} (e^\lambda - 1) - e^{-\lambda} (e^{\lambda F(\xi_i)} - 1) \\
 &= 1 - e^{\lambda(F(\xi_i)-1)}
 \end{aligned}$$

Q2 On a pour  $t \in \llbracket 1, T-1 \rrbracket$

$$\begin{aligned}
 \mathbb{P}(T_i^* = t+1) &= \mathbb{P}(T_i^* = t+1 \wedge T_i > t) \quad \text{car } T_i^* - 1 < T_i \leq T_i^* \\
 &= \mathbb{P}(T_i^* = t+1 | T_i > t) \cdot \mathbb{P}(T_i > t) \\
 &= \mathbb{P}(T_i^* = t+1 | T_i > t) \cdot (1 - \mathbb{P}(T_i \leq t)) \\
 &= \mathbb{P}(T_i^* = t+1 | T_i > t) \cdot \left( 1 - \sum_{k=1}^t \mathbb{P}(T_i^* = k) \right)
 \end{aligned}$$

ce qui s'écrit encore, en notant  $u_t = \mathbb{P}(T_i^* = t)$  et  $\rho = 1 - e^{\lambda(F(\xi_i)-1)}$

$$u_{t+1} = \rho \left( 1 - \sum_{k=1}^t u_k \right)$$

avec  $u_1 = \mathbb{P}(T_i^* = 1) = \mathbb{P}(T_i^* = 1 | T_i > 0) = \rho$ .

Alors,  $u_2 = \rho(1 - \rho)$ , puis  $u_3 = \rho(1 - \rho - \rho(1 - \rho)) = \rho(1 - \rho)^2$ .

Soit donc  $t \geq 3$  tel que  $\forall t' \leq t, u_{t'} = \rho(1 - \rho)^{t'-1}$ ; alors

$$\begin{aligned}
 u_{t+1} &= \rho \left( 1 - \sum_{t'=1}^t u_{t'} \right) \\
 &= \rho \left( 1 - \sum_{t'=1}^t \rho(1 - \rho)^{t'-1} \right) \\
 &= \rho \left( 1 - \rho \sum_{t'=0}^{t-1} (1 - \rho)^{t'} \right) \\
 &= \rho \left( 1 - \rho \frac{1 - (1 - \rho)^t}{1 - (1 - \rho)} \right) \\
 &= \rho(1 - \rho)^t
 \end{aligned}$$

Ainsi  $\forall t \in \llbracket 1, T-1 \rrbracket$ ,  $u_{t+1} = \rho(1 - \rho)^t$ , ce qui s'écrit encore

$$\forall t \in \llbracket 1, T-1 \rrbracket, \mathbb{P}(T_i^* = t+1) = (1 - e^{\lambda(F(\xi_i)-1)}) e^{\lambda t(F(\xi_i)-1)}$$

Autre méthode : On a directement pour  $t \in \llbracket 1, T-1 \rrbracket$

$$\begin{aligned}
 \mathbb{P}(T_i^* = t+1) &= \mathbb{P}(T_i^* = t+1 \wedge T_i > t) \quad \text{car } T_i^* - 1 < T_i \leq T_i^* \\
 &= \mathbb{P}(T_i^* = t+1 | T_i > t) \cdot \mathbb{P}(T_i > t) \\
 &= \mathbb{P}(T_i^* = t+1 | T_i > t) \cdot \left( \prod_{j=0}^{t-1} \mathbb{P}(T_i > j+1 | T_i > j) \right) \\
 &= \mathbb{P}(T_i^* = t+1 | T_i > t) \cdot \left( \prod_{j=0}^{t-1} (1 - \mathbb{P}(T_i \leq j+1 | T_i > j)) \right) \\
 &= \mathbb{P}(T_i^* = t+1 | T_i > t) \cdot \left( \prod_{j=0}^{t-1} (1 - \mathbb{P}(T_i^* = j+1 | T_i > j)) \right) \\
 &= (1 - e^{\lambda(F(\xi_i)-1)}) \cdot \left( \prod_{j=0}^{t-1} (e^{\lambda(F(\xi_i)-1)}) \right) \\
 &= (1 - e^{\lambda(F(\xi_i)-1)}) e^{\lambda t(F(\xi_i)-1)}
 \end{aligned}$$

Les comportements des individus étant censément indépendants, les variables  $(T_i^*)_{i \in \llbracket 1, n \rrbracket}$  sont statistiquement indépendantes et par conséquent

$$L_{T_1^*, \dots, T_n^*}(t_1, \dots, t_n; F, \lambda) = \prod_{i=1}^n (1 - e^{\lambda(F(\xi_i)-1)}) e^{\lambda(F(\xi_i)-1) \cdot (t_i-1)} \mathbf{1}_{\min_i t_i \geq 1}$$

Q3 Soit  $(t_1, \dots, t_n) \in \llbracket 1, T \rrbracket^n$ , et soit  $\phi : \begin{pmatrix} [0, 1] & \rightarrow & \mathbb{R}^+ \\ \eta & \mapsto & (1 - \eta)^n \eta^{\sum_{i=1}^n (t_i-1)} \end{pmatrix}$ .

Alors  $\ln \phi$  est  $C^\infty$ , et de dérivée  $\frac{\partial \ln \phi}{\partial \eta}(\eta) = -\frac{n}{1-\eta} + \frac{1}{\eta} \sum_{i=1}^n (t_i - 1)$ ; donc elle est maximale  $\hat{\eta} = 1 - \frac{1}{\bar{t}}$  où  $\bar{t} = \frac{1}{n} \sum_{i=1}^n t_i$ .

Or  $L_{T_1^*, \dots, T_n^*}(t_1, \dots, t_n; F, \lambda) = \phi(e^{\lambda(F(\xi)-1)})$ , donc par conséquent  $\hat{\lambda}$  vérifie  $e^{\hat{\lambda}(F(\xi)-1)} = \hat{\eta} = 1 - \frac{1}{t}$ , c'est-à-dire

$$\hat{\lambda} = \frac{1}{1-F(\xi)} \ln\left(1 + \frac{1}{t-1}\right)$$

On a d'après la loi de grands nombres

$$\begin{aligned} \lim_{p.s., n \rightarrow \infty} \bar{T} &= \mathbb{E}(T_i^*) \\ &= \sum_{t=1}^{+\infty} \mathbb{P}(T_i^* = t) \cdot t \\ &= \sum_{t=1}^{+\infty} (1-a)ta^{t-1}, \text{ en notant } a = e^{\lambda(F(\xi)-1)} \\ &= (1-a) \cdot \left( \begin{array}{l} ]-1, 1[ \rightarrow \mathbb{R} \\ x \mapsto \sum_{t=1}^{+\infty} tx^{t-1} \end{array} \right) (a) \\ &= (1-a) \cdot \frac{\partial (x \mapsto \sum_{t=0}^{+\infty} x^t)}{\partial x} (a), \text{ car la série converge absolument sur le disque ouvert} \\ &= (1-a) \cdot \frac{\partial (x \mapsto \frac{1}{1-x})}{\partial x} (a) \\ &= (1-a) \cdot \frac{1}{(1-a)^2} \\ &= \frac{1}{1 - e^{\lambda(F(\xi)-1)}} \end{aligned}$$

*Remarque :* Pour calculer  $\sum_{k=0}^{+\infty} P(k)t^k$ , il suffit de décomposer le polynôme  $P(\mathbb{X})$  dans la base  $1, \mathbb{X} - 1, (\mathbb{X} - 1)(\mathbb{X} - 2), \dots$  ce qui conduit à

$$\begin{aligned} \sum_{k=0}^{+\infty} P(k)t^k &= \sum_{k=0}^{+\infty} \left( \sum_{i=0}^d \alpha_i \underbrace{(k-1) \cdots (k-i)}_{i \text{ termes}} \right) t^k \\ &= \sum_{i=0}^d \sum_{k=i}^{+\infty} \alpha_i \frac{\partial^i (x \mapsto x^k)}{\partial x^i} (t) \\ &= \sum_{i=0}^d \alpha_i \frac{\partial^i (x \mapsto \frac{1}{1-x})}{\partial x^i} (t) \\ &= \sum_{i=0}^d \alpha_i \frac{(-1)^i i!}{(1-t)^{i+1}} \end{aligned}$$

Par suite

$$\begin{aligned} \hat{\lambda} &\xrightarrow[p.s.]{n \rightarrow +\infty} \frac{1}{1-F(\xi)} \ln \left( 1 + \frac{1}{\frac{1}{1-e^{\lambda(F(\xi)-1)}} - 1} \right) \\ &= \frac{1}{1-F(\xi)} \ln \left( 1 + \frac{1 - e^{\lambda(F(\xi)-1)}}{e^{\lambda(F(\xi)-1)}} \right) \\ &= \frac{1}{1-F(\xi)} \ln \left( \frac{1}{e^{\lambda(F(\xi)-1)}} \right) \\ &= \lambda \end{aligned}$$

ce qui prouve que  $\hat{\lambda}$  est un estimateur convergent de  $\lambda$ .

Cherchons enfin la loi limite de  $(\hat{\lambda} - \lambda)$ ; pour ce faire cherchons tout d'abord celle de  $(\bar{T} - \mathbb{E}(\bar{T}))$ . On a, en notant  $a = e^{\lambda(F(\xi)-1)}$

$$\begin{aligned} \mathbb{V}(T_i^*) &= \mathbb{E}(T_i^{*2}) - \mathbb{E}(T_i^*)^2 \\ &= (1-a) \sum_{t=1}^{+\infty} t^2 a^{t-1} - \frac{1}{(1-a)^2} \\ &= (1-a) \left( a \sum_{t=2}^{+\infty} t(t-1)a^{t-2} + \sum_{t=1}^{+\infty} ta^{t-1} \right) - \frac{1}{(1-a)^2} \\ &= (1-a) \left( a \frac{2}{(1-a)^3} + \frac{1}{(1-a)^2} \right) - \frac{1}{(1-a)^2}, \text{ car les séries convergent absolument} \\ &= \frac{1+a}{(1-a)^2} - \frac{1}{(1-a)^2} \\ &= \frac{a}{(1-a)^2} \end{aligned}$$

D'après le Théorème Central Limite on a donc

$$\sqrt{n} \left( \bar{T} - \frac{1}{1-a} \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left( 0, \frac{a}{(1-a)^2} \right)$$

donc d'après le théorème de Slutsky

$$\sqrt{n} (\hat{\lambda} - \lambda) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left( 0, \frac{a}{(1-a)^2} \cdot g' \left( \frac{1}{1-a} \right)^2 \right)$$

où  $g : \left( \begin{array}{l} ]0, 1[ \rightarrow \mathbb{R} \\ x \mapsto \frac{1}{1-F(\xi)} \ln\left(1 + \frac{1}{x-1}\right) \end{array} \right)$ .

On a  $\forall x, g'(x) = \frac{1}{1-F(\xi)} \frac{1}{x(x-1)}$  et donc  $g' \left( \frac{1}{1-a} \right) = \frac{1}{1-F(\xi)} \frac{(1-a)^2}{a}$  de sorte que finalement

$$\sqrt{n} (\hat{\lambda} - \lambda) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left( 0, \frac{1}{(1-F(\xi))^2} \frac{(1 - e^{\lambda(F(\xi)-1)})^2}{e^{\lambda(F(\xi)-1)}} \right)$$

☞ Q4 On a pour  $t \in \llbracket 1, T-1 \rrbracket$

$$\begin{aligned} "T_i^{**} = t + 1" &= \begin{cases} "T_i^* = t + 1" & \text{si } "T_i^* \leq T" \\ "T + 1 = t + 1" & \text{si } "T_i^* > T" \end{cases} \\ &= \begin{cases} \text{soit } "T_i^* \leq T \text{ et } T_i^* = t + 1" \\ \text{soit } "T_i^* > T \text{ et } T + 1 = t + 1" \end{cases} \\ &= \begin{cases} \text{soit } "t + 1 \leq T \text{ et } T_i^* = t + 1" \\ \text{soit } "T_i^* = T + 1 \text{ et } T + 1 = t + 1" \\ \text{soit } "T_i^* = T + 1 \text{ et } T + 1 > t + 1" \end{cases} \\ &= \begin{cases} \text{soit } "t + 1 \leq T \text{ et } T_i^* = t + 1" \\ \text{soit } "T_i^* = t + 1 \text{ et } T + 1 = t + 1" \\ \text{soit } "T_i^* = T + 1 \text{ et } T + 1 > t + 1" \end{cases} \end{aligned}$$

Donc

$$\mathbb{P}(T_i^{**} = t + 1) = \begin{cases} \mathbb{P}(T_i^* = t + 1) & \text{si } t + 1 \leq T \\ \mathbb{P}(T_i^* > T) & \text{si } t + 1 = T + 1 \\ 0 & \text{si } t + 1 > T + 1 \end{cases}$$

Or  $\mathbb{P}(T_i^* = t + 1) = (1 - e^{\lambda(F(\xi)-1)}) e^{\lambda t(F(\xi)-1)}$ ; et par ailleurs

$$\begin{aligned} \mathbb{P}(T_i^* > T) &= \sum_{\tau=T+1}^{+\infty} \mathbb{P}(T_i^* = \tau) \\ &= \sum_{\tau=T}^{+\infty} (1 - e^{\lambda(F(\xi)-1)}) (e^{\lambda(F(\xi)-1)})^\tau \\ &= e^{\lambda(F(\xi)-1) \cdot T} \end{aligned}$$

En conséquence on a

$$\begin{aligned} &L_{T_1^{**}, \dots, T_n^{**}}(t_1 + 1, \dots, t_n + 1; F, \lambda) \\ &= \prod_{i \in \llbracket 1, n \rrbracket} \mathbb{P}(T_i^{**} = t_i + 1) \\ &= (\prod_{t_i+1 \leq T} \mathbb{P}(T_i^* = t_i + 1)) \cdot (\prod_{t_i+1 > T} \mathbb{P}(T_i^* > T) \mathbf{1}_{t_i+1=T+1}) \\ &= (\prod_{t_i+1 \leq T} (1 - e^{\lambda(F(\xi)-1)}) e^{\lambda t_i(F(\xi)-1)}) (\prod_{t_i \geq T} (e^{\lambda(F(\xi)-1)})^T \cdot \mathbf{1}_{t_i+1=T+1}) \\ &= \left( (1 - e^{\lambda(F(\xi)-1)})^{|\{i/t_i \leq T\}|} \cdot (e^{\lambda(F(\xi)-1)})^{\sum_{t_i+1 \leq T} t_i} \right) \left( (e^{\lambda(F(\xi)-1)})^{\sum_{t_i \geq T} t_i} \right) \\ &= \left( (1 - e^{\lambda(F(\xi)-1)})^{|\{i/t_i < T\}|} \right) \left( (e^{\lambda(F(\xi)-1)})^{\sum_{i=1}^n t_i} \right) \\ &= \left( (1 - e^{\lambda(F(\xi)-1)})^{|\{i/t_i < T\}|} \right) \left( (e^{\lambda(F(\xi)-1)})^{n\bar{t}} \right) \end{aligned}$$

On en tire que

$$\frac{\partial \ln L_{T_1^{**}, \dots, T_n^{**}}(t_1, \dots, t_n; F, \lambda)}{\partial \lambda} = -\frac{|\{i/t_i < T\}|}{1 - e^{\lambda(F(\xi)-1)}} + \frac{n(\bar{t} - 1)}{e^{\lambda(F(\xi)-1)}}$$

de sorte que

$$e^{\widehat{\lambda(F(\xi)-1)}} = \frac{n(\bar{t} - 1)}{|\{i/t_i < T\}| + n(\bar{t} - 1)}$$

et donc

$$\widehat{\lambda^{censure}} = \frac{1}{F(\xi)-1} \ln \left( \frac{n(\bar{t}-1)}{|\{i/t_i < T\}| + n(\bar{t}-1)} \right)$$

Comme en l'absence de censure ( $T = +\infty$ ) on a  $\hat{\lambda} = \frac{1}{F(\xi)-1} \ln(1 - \frac{1}{\bar{t}})$  on a

$$\begin{aligned} \widehat{\lambda^{censure}} - \hat{\lambda} &= \frac{1}{F(\xi) - 1} \left( \ln \left( \frac{\bar{t} - 1}{\frac{|\{i/t_i < T\}|}{n} + \bar{t} - 1} \right) - \ln \left( \frac{\bar{t} - 1}{\bar{t}} \right) \right) \\ &= -\frac{1}{F(\xi) - 1} \ln \left( 1 + \frac{|\{i/t_i < T\}|}{n} \right) \end{aligned}$$

☞ Q5 La vraisemblance s'écrit cette fois

$$L_{T_1^{**}, \dots, T_n^{**}}(t_1, \dots, t_n; \gamma, \lambda) = \left(1 - e^{-\lambda e^{\gamma(\xi - \xi_0)}}\right)^n \left(e^{-\lambda e^{\gamma(\xi - \xi_0)}}\right)^{n\bar{t}} \mathbf{1}_{\xi \geq \xi_0}$$

On constate alors que  $\phi : ((\lambda, \gamma) \mapsto \lambda r^\gamma)$  n'est jamais injective (car  $\forall r \in \mathbb{R}^{+*}, \phi(1, 1) = r = \phi(\frac{1}{2}, \sqrt{r})$ ), de sorte que la vraisemblance  $((\lambda, \gamma) \mapsto L_{T_1^{**}, \dots, T_n^{**}}(t_1, \dots, t_n; \gamma, \lambda))$  ne l'est pas non plus, de sorte que le paramètre  $(\lambda, \gamma)$  n'est pas identifiable. D'où l'intérêt des calculs précédents...

\*  
\* \*

### Corrigé de l'exercice 3

☞ Q1 L'assertion est triviale lorsque  $n = 1$ .

Soit donc  $n \geq 1$  tel que l'assertion soit vraie.

Soient  $Y_1, \dots, Y_{n+1}$  i.i.d. de densité  $f_{\mathcal{E}}(\cdot, \lambda)$ .

Définissons pour  $k \in \llbracket 1, n+1 \rrbracket$   $S_k = \sum_{i=1}^k Y_i$ .

Constatant que  $S_{n+1} = S_n + Y_{n+1}$ , définissons  $\phi : \left( \begin{array}{ccc} \mathbb{R}^{+*2} & \rightarrow & \mathbb{R}^{+*} \times ]0, 1[ \\ (x, y) & \mapsto & (x + y, \frac{x}{x+y}) \end{array} \right)$ . Alors  $\phi$

un  $C^\infty$ -difféomorphisme, de jacobien  $\left| \left( \begin{array}{cc} y & x \\ 1-y & x \end{array} \right)^{-1} \right| = \frac{1}{x}$ .

Or  $L_{S_n, Y_{n+1}}(s, y) = \left( \lambda^n \frac{s^{n-1}}{(n-1)!} e^{-\lambda s} \mathbf{1}_{s>0} \right) \cdot \left( \lambda e^{-\lambda y} \mathbf{1}_{y>0} \right)$ , par hypothèse de récurrence et puisque  $S_n$  et  $Y_{n+1}$  sont indépendantes (car les  $(Y_i)_i$  sont toutes indépendantes).

En conséquence, puisque  $\phi^{-1}(u, v) = (uv, u(1-v))$ ,  $L_{\phi(S_n, Y_{n+1})}(u, v) = \lambda^{n+1} \frac{(uv)^{n-1}}{(n-1)!} e^{-\lambda uv} \mathbf{1}_{u>0 \wedge v \in ]0,1[}$ .

En particulier d'après le théorème de projection  $L_{(\phi(S_n, Y_{n+1}))_1}(u) = \lambda^{n+1} \frac{u^n}{n!} e^{-\lambda u} \mathbf{1}_{u>0}$ , c'est-à-dire

$$L_{Y_1+\dots+Y_{n+1}}(u) = \lambda^{n+1} \frac{u^n}{n!} e^{-\lambda u} \mathbf{1}_{u>0}$$

de sorte que l'assertion est vraie pour tout  $n \in \mathbb{N}^*$ .

☞ Q1 Le modèle statistique s'écrit  $\left\{ \left( \mathbb{R}^{+2} \right)^n, (\mathcal{E} \cdot \mathcal{E})^{\otimes n} \right\}$ .

On a de plus pour  $i \in \llbracket 1, n \rrbracket$   $L_{Z_i, C_i}(z_i, c_i; \lambda, \mu) = \lambda \mu e^{-\lambda z_i - \mu c_i}$ , et donc

$$L_{Z_1, \dots, Z_n, C_1, \dots, C_n}(z_1, \dots, z_n, c_1, \dots, c_n; \lambda, \mu) = (\lambda \mu)^n e^{-\lambda \sum_{i=1}^n z_i - \mu \sum_{i=1}^n c_i} \mathbf{1}_{\min_i z_i \geq 0} \mathbf{1}_{\min_i c_i \geq 0}$$

En particulier le modèle est exponentiel et une statistique exhaustive est

$$(z_1, \dots, z_n, c_1, \dots, c_n) \mapsto \left( \sum_{i=1}^n z_i, \sum_{i=1}^n c_i \right)$$

☞ Q2 On a immédiatement  $\frac{\partial \ln L_{Z_1, \dots, Z_n, C_1, \dots, C_n}}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^n z_i$ , et par suite un estimateur de  $\lambda$  et  $\hat{\lambda} = \frac{1}{\bar{z}}$  (la log-vraisemblance étant concave, elle est bien maximale en  $\bar{z}$ ). De même un estimateur de  $\mu$  et  $\hat{\mu} = \frac{1}{\bar{c}}$ .

☞ Q3 On calcule successivement

$$\frac{\partial^2 \ln L_{Z_1, \dots, Z_n, C_1, \dots, C_n}}{\partial \lambda^2} = \frac{n}{\lambda^2}$$

$$\frac{\partial^2 \ln L_{Z_1, \dots, Z_n, C_1, \dots, C_n}}{\partial \lambda \partial \mu} = 0$$

$$\frac{\partial^2 \ln L_{Z_1, \dots, Z_n, C_1, \dots, C_n}}{\partial \mu^2} = \frac{n}{\mu^2}$$

de sorte que  $I_1(\lambda, \mu) = \begin{pmatrix} \frac{1}{\lambda^2} & 0 \\ 0 & \frac{1}{\mu^2} \end{pmatrix}$  d'inverse  $I_1(\lambda, \mu)^{-1} = \begin{pmatrix} \lambda^2 & 0 \\ 0 & \mu^2 \end{pmatrix}$  et donc d'après le

Théorème Central Limite

$$\sqrt{n} \begin{pmatrix} \hat{\lambda} - \lambda \\ \hat{\mu} - \mu \end{pmatrix} \xrightarrow{n \rightarrow +\infty} \mathcal{N} \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \lambda^2 & 0 \\ 0 & \mu^2 \end{pmatrix} \right)$$

☞ Q4 Déterminons la loi de  $\hat{\lambda}$  : soit donc  $h : (\mathbb{R} \rightarrow \mathbb{R})$  mesurable bornée, et calculons  $\mathbb{E} \left( h(\hat{\lambda}) \right)$ . On a

$$\begin{aligned} \mathbb{E} \left( h(\hat{\lambda}) \right) &= \mathbb{E} \left( h \left( \frac{n}{\sum_{i=1}^n z_i} \right) \right) \\ &= \int_{\mathbb{R}} h \left( \frac{n}{s} \right) \lambda^n \frac{s^{n-1}}{(n-1)!} e^{-\lambda s} \mathbf{1}_{s>0} ds \\ &= - \int_{0+}^{+\infty} h(u) \lambda^n \frac{\left( \frac{n}{u} \right)^{n-1}}{(n-1)!} e^{-\lambda \frac{n}{u}} \mathbf{1}_{u>0} \left( -\frac{n}{u^2} \right) du \end{aligned}$$

En conséquence,  $L_{\hat{\lambda}}(u) = \frac{\lambda^n}{n!} \left( \frac{n}{u} \right)^{n+1} e^{-\lambda \frac{n}{u}} \mathbf{1}_{u>0}$ .

De façon similaire,  $L_{\hat{\mu}}(v) = \frac{\mu^n}{n!} \left( \frac{n}{v} \right)^{n+1} e^{-\mu \frac{n}{v}} \mathbf{1}_{v>0}$ .

Ainsi,

$$L_{\hat{\lambda}, \hat{\mu}}(l, m) = \frac{(\lambda \mu)^n}{(n!)^2} \left( \frac{n^2}{lm} \right)^{n+1} e^{-n \frac{\lambda}{l} - n \frac{\mu}{m}}$$

Par ailleurs on a pour  $n \geq 2$

$$\begin{aligned} \mathbb{E} \left( \hat{\lambda} \right) &= \int_{\mathbb{R}} l \frac{\lambda^n}{n!} \left( \frac{n}{l} \right)^{n+1} e^{-\lambda \frac{n}{l}} \mathbf{1}_{l>0} dl \\ &= \int_{0+}^{+\infty} \frac{\lambda^n}{(n-1)!} \left( \frac{n}{l} \right)^n e^{-\lambda \frac{n}{l}} dl \\ &= \int_{0+}^{+\infty} \frac{\lambda^n}{(n-1)!} u^{n-2} e^{-\lambda u} n du \\ &= \frac{n \lambda}{(n-1)} \int_{0+}^{+\infty} \frac{1}{(n-2)!} \lambda^{n-2} u^{n-2} e^{-\lambda u} \lambda du \\ &= \frac{n \lambda}{(n-1)} \underbrace{\int_{0+}^{+\infty} \frac{1}{(n-2)!} v^{n-2} e^{-v} dv}_{\int_{0+}^{+\infty} f_{\Gamma(n-2, \lambda)}(v) dv = 1} \\ &= \frac{n}{n-1} \lambda \end{aligned}$$

de sorte que

$$\mathbb{E} \left( \begin{pmatrix} \hat{\lambda} \\ \hat{\mu} \end{pmatrix} \right) = \frac{n}{n-1} \begin{pmatrix} \lambda \\ \mu \end{pmatrix}$$

Autrement dit,  $\begin{pmatrix} \hat{\lambda} \\ \hat{\mu} \end{pmatrix}$  est biaisé à distance finie.

Enfin pour  $n \geq 3$

$$\begin{aligned} \mathbb{E} \left( \hat{\lambda}^2 \right) &= \int_{\mathbb{R}} l^2 \frac{\lambda^n}{n!} \left( \frac{n}{l} \right)^{n+1} e^{-\lambda \frac{n}{l}} \mathbf{1}_{l>0} dl \\ &= - \int_{0+}^{+\infty} \left( \frac{n}{u} \right)^2 \frac{\lambda^n}{n!} u^{n+1} e^{-\lambda u} \left( -\frac{n}{u^2} \right) du \\ &= \frac{n^2 \lambda^2}{(n-1)(n-2)} \int_{0+}^{+\infty} \frac{1}{(n-3)!} \lambda^{n-3} u^{n-3} e^{-\lambda u} \lambda du \\ &= \frac{n^2 \lambda^2}{(n-1)(n-2)} \underbrace{\int_{0+}^{+\infty} \frac{1}{(n-3)!} v^{n-3} e^{-v} dv}_{\int_{0+}^{+\infty} f_{\Gamma(n-3, \lambda)}(v) dv = 1} \\ &= \frac{n^2}{(n-1)(n-2)} \lambda^2 \end{aligned}$$



et donc

$$\begin{aligned} \mathbb{V}(\widehat{\lambda}) &= \mathbb{E}(\widehat{\lambda}^2) - \mathbb{E}(\widehat{\lambda})^2 \\ &= \frac{n^2}{(n-1)(n-2)}\lambda^2 - \left(\frac{n}{n-1}\lambda\right)^2 \\ &= \frac{n^2(n-1) - (n-2)n^2}{(n-1)^2(n-2)}\lambda^2 \\ &= \frac{n^2}{(n-1)^2(n-2)}\lambda^2 \end{aligned}$$

(on vérifie que l'e.m.v. est asymptotiquement efficace :  $\mathbb{V}(\widehat{\lambda}) \underset{n \rightarrow \infty}{\sim} \frac{1}{n}\lambda^2$ ).

Autre méthode (plus rapide) :

$$\begin{aligned} \mathbb{E}(\widehat{\lambda}) &= \int_{\mathbb{R}} \frac{n}{s} \lambda^n \frac{s^{n-1}}{(n-1)!} e^{-\lambda s} ds \\ &= \frac{n}{n-1} \lambda \int_{\mathbb{R}} \lambda^{n-1} \frac{s^{n-2}}{(n-2)!} e^{-\lambda s} ds \\ &= \frac{n}{n-1} \lambda \int_{\mathbb{R}} \underbrace{\lambda^{n-1} \frac{s^{n-2}}{(n-2)!} e^{-\lambda s}}_{f_{\Gamma(n-1, \lambda)}(s)} ds \\ &= \frac{n}{n-1} \lambda \end{aligned}$$

De même

$$\begin{aligned} \mathbb{E}(\widehat{\lambda}^2) &= \frac{n^2}{(n-1)(n-2)} \lambda^2 \int_{\mathbb{R}} \underbrace{\lambda^{n-2} \frac{s^{n-3}}{(n-3)!} e^{-\lambda s}}_{f_{\Gamma(n-2, \lambda)}(s)} ds \\ &= \frac{n^2}{(n-1)(n-2)} \lambda^2 \end{aligned}$$

Enfin, remarquant que toutes les  $Z_1, \dots, Z_n, C_1, \dots, C_n$  sont indépendantes, on a  $\text{Cov}(\widehat{\lambda}, \widehat{\mu}) = 0$  de sorte que

$$\mathbb{V}\left(\begin{pmatrix} \widehat{\lambda} \\ \widehat{\mu} \end{pmatrix}\right) = \frac{n^2}{n-1} \begin{pmatrix} \frac{1}{n-2} & -\frac{1}{n-1} \\ 0 & \mu^2 \end{pmatrix} \begin{pmatrix} \lambda^2 & 0 \\ 0 & \mu^2 \end{pmatrix}$$

Q5 Posons  $\widehat{\lambda}^* = \frac{n-1}{n}\widehat{\lambda}$  et  $\widehat{\mu}^* = \frac{n}{n-1}\widehat{\mu}$ .

D'après le résultat précédent,  $\begin{pmatrix} \widehat{\lambda}^* \\ \widehat{\mu}^* \end{pmatrix}$  est un estimateur sans biais de  $\begin{pmatrix} \lambda \\ \mu \end{pmatrix}$ . Or  $T(Z, C) =$

$(\frac{1}{n} \sum_{i=1}^n Z_i, \frac{1}{n} \sum_{i=1}^n C_i)$  est une statistique exhaustive complète (car canonique dans un modèle exponentielle); en conséquence, d'après le théorème de Lehman-Scheffé,  $\begin{pmatrix} \widehat{\lambda}^* \\ \widehat{\mu}^* \end{pmatrix}$  est optimal

parmi les estimateurs sans biais de  $\begin{pmatrix} \lambda \\ \mu \end{pmatrix}$ .

Par ailleurs,

$$\begin{aligned} \text{Cov}\left(\begin{pmatrix} \widehat{\lambda}^* \\ \widehat{\mu}^* \end{pmatrix}\right) &= \left(\frac{n-1}{n}\right)^2 \text{Cov}\left(\begin{pmatrix} \widehat{\lambda} \\ \widehat{\mu} \end{pmatrix}\right) \\ &= \frac{1}{n-2} \begin{pmatrix} \lambda^2 & 0 \\ 0 & \mu^2 \end{pmatrix} \end{aligned}$$

Or la borne FDCR du modèle est  $\frac{1}{n} \begin{pmatrix} \lambda^2 & 0 \\ 0 & \mu^2 \end{pmatrix}$ .

En conséquence,  $\begin{pmatrix} \widehat{\lambda}^* \\ \widehat{\mu}^* \end{pmatrix}$  n'est pas un estimateur efficace de  $\begin{pmatrix} \lambda \\ \mu \end{pmatrix}$ .

On suppose dans cette seconde partie que les seules observations disponibles portent sur  $X = \text{Min}(Z_i, C_i)$ ,  $i = 1, \dots, n$ .

Q1 On a pour  $t \in \mathbb{N}$

$$\begin{aligned} \mathbb{P}(X_i > t) &= \mathbb{P}(Z_i > t \wedge C_i > t) \\ &= \mathbb{P}(Z_i > t) \cdot \mathbb{P}(C_i > t) \quad \text{par indépendance} \\ &= \left(\int_t^{+\infty} \lambda e^{-\lambda u} du\right) \left(\int_t^{+\infty} \lambda e^{-\mu v} dv\right) \\ &= e^{-(\lambda+\mu)t} \end{aligned}$$

de sorte que  $X_i \sim \mathcal{E}(\lambda + \mu)$ .

Q2 Le modèle statistique s'écrit ici  $\{\mathbb{R}^{+n}, \mathcal{E}^{\otimes n}\}$ .

Sa vraisemblance étant  $L_{X_1, \dots, X_n}(x_1, \dots, x_n; \lambda, \mu) = (\lambda + \mu)^n e^{-(\lambda+\mu)\sum_{i=1}^n x_i}$ , la seule fonction identifiable de  $(\lambda, \mu)$  est  $\lambda + \mu$ .

Q3 Cherchons à estimer  $\gamma = \lambda + \mu$ .

(i) La maximisation de  $L_{X_1, \dots, X_n}(x_1, \dots, x_n; \gamma) = (\gamma)^n e^{-\gamma \sum_{i=1}^n x_i}$  conduit à l'estimateur  $\widehat{\gamma}_X =$

$$\frac{1}{\bar{x}} = \frac{n}{\sum_{i=1}^n \min(z_i, c_i)}$$

(ii) tandis que celle de  $L_{Z_1, \dots, Z_n, C_1, \dots, C_n}(z_1, \dots, z_n, c_1, \dots, c_n; \lambda, \mu)$  conduit à estimer  $\lambda$  et  $\mu$  individuellement, ce dont on tire une estimation convergente de  $\lambda + \mu$ , à savoir  $\widehat{\gamma}_{Z, C} = \widehat{\lambda} + \widehat{\mu}$ .

Bien entendu il n'y a aucune raison pour que ces estimateurs, fondés sur des observations différentes (c'est-à-dire, en définitive, définis dans des modèles statistiques différents), soient les mêmes. Comme il est montré ci-après ils ont même des variances asymptotiques différentes.

Q4 On a alternativement

$$\begin{aligned} \text{(i)} \quad \frac{\partial \ln L_{X_1, \dots, X_n}(x_1, \dots, x_n; \gamma)}{\partial \gamma} &= \frac{n}{\gamma} - \sum_{i=1}^n x_i, \\ \text{puis} \quad \frac{\partial^2 \ln L_{X_1, \dots, X_n}(x_1, \dots, x_n; \gamma)}{\partial \gamma^2} &= -\frac{n}{\gamma^2}, \end{aligned}$$

donc

$$\sqrt{n}(\widehat{\gamma}_X - \gamma) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(O, (\lambda + \mu)^2)$$

(ii) Soit  $g : \begin{pmatrix} \mathbb{R}^{+*2} & \rightarrow & \mathbb{R}^{+*2} \\ (x, y) & \mapsto & (x + y, y) \end{pmatrix}$  qui est de classe  $C^\infty$ , et de jacobienne  $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ . Alors d'après le théorème de Slutsky

$$\sqrt{n} \left( g \begin{pmatrix} \widehat{\lambda} \\ \widehat{\mu} \end{pmatrix} - g \begin{pmatrix} \lambda \\ \mu \end{pmatrix} \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \frac{\partial g}{\partial(x, y)}(\lambda, \mu) \begin{pmatrix} \lambda^2 & 0 \\ 0 & \mu^2 \end{pmatrix} \frac{\partial g}{\partial(x, y)}(\lambda, \mu) \right)$$

En particulier selon la première composante,<sup>2</sup>

$$\sqrt{n}(\widehat{\gamma}_{Z,C} - \gamma) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(O, \lambda^2 + \mu^2)$$

On constate que  $(\lambda + \mu)^2 \geq \lambda^2 + \mu^2$ , avec égalité ssi  $\lambda = 0$  ou  $\mu = 0$  (situations dégénérées et sans intérêt). Autrement dit, asymptotiquement  $\widehat{\gamma}_{Z,C}$  est préférable à  $\widehat{\gamma}_X$ , ce que l'on écrira sous la forme publicitaire

$$\widehat{\lambda} + \widehat{\mu} \gg \widehat{\lambda + \mu}$$

Ce résultat est bien-sûr naturel dans la mesure où le premier estimateur est fondé sur une information plus fine que le second.

☞ Q1 Pour calculer  $\mathbb{E} \left( \frac{1}{X_1 + \dots + X_n} \mid Z_1 + \dots + Z_n, C_1 + \dots + C_n \right)$ , il semble naturel de déterminer tout d'abord

$$\begin{aligned} & L_{X_1 + \dots + X_n \mid Z_1 + \dots + Z_n, C_1 + \dots + C_n} \\ &= \frac{L_{X_1 + \dots + X_n, Z_1 + \dots + Z_n, C_1 + \dots + C_n}}{L_{Z_1 + \dots + Z_n, C_1 + \dots + C_n}} \end{aligned}$$

La difficulté ici est que  $X_i$  n'est pas indépendante de  $(Z_i, C_i)$ , puisque  $X_i = \min(Z_i, C_i)$ , ce qui complique singulièrement le calcul.

En fait, le résultat s'exprime très simplement et peut être démontré de la façon suivante, astucieuse quoique moins naturelle.

Notons en effet  $\widehat{\theta} = \mathbb{E} \left( \frac{1}{X_1 + \dots + X_n} \mid Z_1 + \dots + Z_n = z, C_1 + \dots + C_n = c \right)$ ; alors d'après la partie II  $\widehat{\theta}$  est un estimateur sans biais de  $\frac{\lambda + \mu}{n-1}$ . La statistique en  $\widehat{\theta}(X_1, \dots, X_n)$  étant régulière (car le modèle est exponentiel), d'après le théorème de Lehman-Scheffé  $\widehat{\theta}$  est en outre optimal.

Mais d'après la partie I,  $\frac{\widehat{\lambda} + \widehat{\mu}}{n}$  est un autre estimateur sans biais de  $\frac{\lambda + \mu}{n-1}$ , lui aussi optimal.

En conséquence,  $\widehat{\theta}$  et  $\widehat{\theta} - \frac{\widehat{\lambda} + \widehat{\mu}}{n}$  sont décorrélés, ainsi que  $\frac{\widehat{\lambda}}{n}$  et  $(\widehat{\theta} - \frac{\widehat{\lambda}}{n}) - \frac{\widehat{\lambda}}{n}$ , et que  $\frac{\widehat{\mu}}{n}$  et  $(\widehat{\theta} - \frac{\widehat{\mu}}{n}) - \frac{\widehat{\mu}}{n}$ , et ce en vertu du lemme suivant :

<sup>2</sup>Il n'est pas nécessaire pour pratiquer la delta méthode que  $g$  soit inversible, et il aurait suffi ici de considérer  $g : \begin{pmatrix} \mathbb{R}^{+*2} & \rightarrow & \mathbb{R}^{+*} \\ (x, y) & \mapsto & x + y \end{pmatrix}$ .

Si  $\widehat{a}$  et  $\widehat{b}$  sont deux estimateurs sans biais optimaux du même paramètre, alors  $\text{Cov}(\widehat{a}, \widehat{a} - \widehat{b}) = 0$

En effet, définissons pour  $t \in \mathbb{R}$   $c_t = (1-t)\widehat{a} + t\widehat{b} = \widehat{a} + t(\widehat{b} - \widehat{a})$ .

Alors  $\forall t \in \mathbb{R}, \mathbb{V}(c_t) \geq \mathbb{V}(\widehat{b})$  puisque  $\widehat{b}$  est optimal;

ceci s'écrit  $\forall t \in \mathbb{R}, \mathbb{V}(\widehat{a}) + 2t\text{Cov}(\widehat{a}, \widehat{b} - \widehat{a}) + t^2\mathbb{V}(\widehat{b} - \widehat{a})$  qui est une inégalité polynôme de celui-ci est négatif ou nul, ce dont on tire le résultat.

Par suite,

$$\begin{aligned} \mathbb{V}(\widehat{\theta}) &= \mathbb{V} \left( \frac{\widehat{\lambda}}{n} \right) + \mathbb{V} \left( \frac{\widehat{\mu}}{n} \right) \\ &= \mathbb{V} \left( \frac{\widehat{\lambda} + \widehat{\mu}}{n} \right) \quad \text{car } \widehat{\lambda} \text{ et } \widehat{\mu} \text{ sont indépendants} \end{aligned}$$

Dans ces conditions, les deux estimateurs  $\widehat{\theta}$  et  $\frac{\widehat{\lambda} + \widehat{\mu}}{n}$  qui ont même espérance, même variance sont exactement corrélés, sont égaux presque sûrement :<sup>3</sup>

$$\widehat{\theta} = \frac{\widehat{\lambda} + \widehat{\mu}}{n} \text{ p.s.}$$

c'est-à-dire

$$\mathbb{E} \left( \frac{1}{\sum_{i=1}^n \min(Z_i, C_i)} \mid \sum_{i=1}^n Z_i \wedge \sum_{i=1}^n C_i \right) = \frac{1}{\sum_{i=1}^n Z_i} + \frac{1}{\sum_{i=1}^n C_i} \quad \text{p.s.}$$

\*  
\* \*

<sup>3</sup>Ce résultat est généralement connu comme l'*unicité presque sûre* d'un estimateur optimal fonction d'une statistique totale.