Cursus Intégré
2004-2005

Rappels de statistique mathématique
Corrigé des travaux dirigés n°5

Guillaume Lacôte
Bureau **E03**
✉ Guillaume.Lacote@ensae.fr
☞ http://ensae.no-ip.com/SE222/

Version du 20050421-15h38, révisée le 21 avril 2005

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

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

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

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

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

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

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 de 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

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

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 fonctionde 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 :

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

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 $(y_1,\ldots,y_n)\in\mathbb{R}^n$ par

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

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 $(y_1,\ldots,y_n)\in\mathbb{R}^{+n}$

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

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

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

Il s'ensuit que

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

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

☞ Q3 De la condition nécessaire $\frac{\partial \ln L}{\partial a}(\widehat{a_0}) = 0$ on tire $\widehat{a_0} = \frac{1}{\bar{y}}$; réciproquement $\widehat{a_0}$ maximise bien la log-vraisemblance car celle-ci est concave.

Par ailleurs, on a

$$\begin{aligned}
\mathbb{E}\left(\widehat{a_0}\right) &= \mathbb{E}\left(\frac{1}{\bar{y}}\right) \\
&> \frac{1}{\mathbb{E}\left(\bar{y}\right)} \quad \text{(d'après l'inégalité de Jensen)} \\
&= a
\end{aligned}$$

Autrement dit, $\widehat{a_0}$ est biaisé et surestime systématiquement la vraie valeur a.

On pourrait même calculer le biais exact, en s'appuyant sur le fait que $(Y_1 + \ldots + Y_n) \rightsquigarrow \Gamma(n,a)$ et en calculant explicitement $\int_{\mathbb{R}} \frac{n}{s} f_{\Gamma(a)}(s)ds = \frac{n}{n-1}a$; on en déduirait que $\frac{n-1}{n}\widehat{a_0}$ 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

$$\boxed{\mathbb{V}_{as}\left(\widehat{a_0}\right) = \frac{a^2}{n}}$$

(Attention, c'est bien $I_n(a)^{-1}$ et non pas $I_1(a)^{-1}$ car on étudie la loi asymptotique de $(\widehat{a_0} - \mathbb{E}\left(\widehat{a_0}\right))$ et non pas celle de $\sqrt{n}(\widehat{a_0} - \mathbb{E}\left(\widehat{a_0}\right)))$

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

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

On constate donc que le modèle n'est **pas** exponentiel, car on ne peut dissocier les termes en 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 $(Y_1, \ldots, Y_n)$ 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 $(y_1, \ldots, y_n) \in {\mathbb{R}^+}^n$

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

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 $(\widehat{a},\widehat{b})$ de (a,b) vérifie

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

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

☞ Q3 On a successivement

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

En conséquence

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

Il reste alors à exprimer $\mathbb{E}\left(y_1^b \ln(y_1)^i\right)$ 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 (α,β), de densité $f_{\alpha,\beta}(x) = \alpha\,\theta\,x^{\alpha-1}\,\mathrm{e}^{-\theta x^\alpha}\mathbb{1}_{x>0}$, alors pour $p \in \mathbb{N}^*$ on a

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

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

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

Or (TD 2, exercice 2)

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

Ceci permet alors de conclure en posant $\theta = a$ et $\alpha = b$ que[1]

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

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

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

Par ailleurs,

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

En conséquence,

$$\begin{aligned}
\mathbb{V}_{as}\left(\widehat{a}\right) &= \frac{n + na\mathbb{E}\left((\ln y_1)^2 y_1\right)}{|I_n(a,1)|} \\
&= \frac{n + na\mathbb{E}\left((\ln y_1)^2 y_1\right)}{\frac{n}{a^2}\left(n + na\mathbb{E}\left((\ln y_1)^2 y_1\right)\right) - \left(n\mathbb{E}\left(y_1\ln y_1\right)\right)^2} \\
&= \frac{a^2}{n}\left(1 + \frac{\frac{a^2}{n}\left(n\mathbb{E}\left(y_1\ln y_1\right)\right)^2}{\left(n + na\mathbb{E}\left((\ln y_1)^2 y_1\right)\right) - \frac{a^2}{n}\left(n\mathbb{E}\left(y_1\ln y_1\right)\right)^2}\right) \\
&\geq \frac{a^2}{n} \\
&= \mathbb{V}_{as}\left(\widehat{a_0}\right)
\end{aligned}$$

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 de $\begin{pmatrix} \widehat{a} \\ \widehat{b} \end{pmatrix}$, de la façon suivante :

– Observant $(y_1,\ldots,y_n)$, on en tire $\begin{pmatrix} \widehat{a^0} \\ \widehat{b^0} \end{pmatrix}$ estimateur du maximum de vraisemblance de $\begin{pmatrix} a^0 \\ b^0 \end{pmatrix}$.

[1]Bien que la tentation soit forte, nous résisterons ici au plaisir d'inverser cette matrice.

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

Bien-entendu, à la limite lorsque $N \to +\infty$ on a $\begin{pmatrix} \widehat{a^1} \\ \widehat{b^1} \end{pmatrix} = \begin{pmatrix} \widehat{a^0} \\ \widehat{b^0} \end{pmatrix}$; 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 $\begin{pmatrix} \widehat{a^1} \\ \widehat{b^1} \end{pmatrix}$.

On considère maintenant le cas de T observations $Y_1, \ldots, Y_T$ 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 $(y_1, \ldots, y_n) \in \mathbb{R}^n$ par

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

Une condition nécéssaire sur l'estimateur $\widehat{\alpha}$ de α est donc

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

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

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

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

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

de sorte que

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

Posons donc

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

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

$$\frac{\partial \ln L}{\partial \alpha}(y,\widehat{\alpha}) = 0 \quad \Leftrightarrow \quad < (1,\ldots,T) | u) >_{e^{\widehat{\alpha}}} = 0$$

⋆

⋆ ⋆

Corrigé de l'exercice 2

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

On a pour $t \in [\![1, T-1]\!]$

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

☞ Q2 On a pour $t \in [\![1, T-1]\!]$

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

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

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

avec $u_1 = \mathbb{P}\left(T_i^* = 1\right) = \mathbb{P}\left(T_i^* = 1 \,|\, T_i > 0\right) = \rho$.
Alors, $u_2 = \rho(1-\rho)$, puis $u_3 = \rho\left(1 - \rho - \rho(1-\rho)\right) = \rho(1-\rho)^2$.

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

$$\boxed{\forall t \in [\![1, T-1]\!],\ \mathbb{P}\left(T_i^* = t+1\right) = \left(1 - e^{\lambda(F(\xi_i)-1)}\right) e^{\lambda t(F(\xi_i)-1)}}$$

<u>*Autre méthode :*</u> On a directement pour $t \in [\![1, T-1]\!]$

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

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

$$L_{T_1^*,\ldots,T_n^*}\left(t_1, \ldots, t_n \ ;\ F, \lambda\right) = \Pi_{i=1}^n \left(1 - e^{\lambda(F(\xi_i)-1)}\right) e^{\lambda(F(\xi_i)-1)\cdot(t_i-1)} \mathbb{1}_{\min_i t_i \geq 1}$$

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

Alors $\ln\phi$ est C^∞, 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 en $\widehat{\eta} = 1 - \frac{1}{\bar{t}}$ où $\bar{t} = \frac{1}{n}\sum_{i=1}^n t_i$.

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

$$\boxed{\widehat{\lambda} = \frac{1}{1-F(\xi)}\ln(1+\frac{1}{\bar{t}-1})}$$

On a d'après la loi de grands nombres

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

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

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

Par suite

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

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

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

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

D'après le Théorème Central Limite on a donc

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

donc d'après le théorème de Slutsky

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

où $g:\begin{pmatrix}]0,1[& \to & \mathbb{R}\\ x & \mapsto & \frac{1}{1-F(\xi)}\ln(1+\frac{1}{x-1})\end{pmatrix}$.

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

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

☞ Q4 On a pour $t \in [\![1, T-1]\!]$

$$\begin{aligned}
\text{“}T_i^{**} = t+1\text{”} &= \begin{cases} \text{“}T_i^* = t+1\text{”} & \text{si “}T_i^* \leq T\text{”} \\ \text{“}T+1 = t+1\text{”} & \text{si “}T_i^* > T\text{”} \end{cases} \\
&= \begin{cases} \text{soit} & \text{“}T_i^* \leq T \text{ et } T_i^* = t+1\text{”} \\ \text{soit} & \text{“}T_i^* > T \text{ et } T+1 = t+1\text{”} \end{cases} \\
&= \begin{cases} \text{soit} & \text{“}t+1 \leq T \text{ et } T_i^* = t+1\text{”} \\ \text{soit} & \text{“}T_i^* = T+1 \text{ et } T+1 = t+1\text{”} \\ \text{soit} & \text{“}T_i^* = T+1 \text{ et } T+1 > t+1\text{”} \end{cases} \\
&= \begin{cases} \text{soit} & \text{“}t+1 \leq T \text{ et } T_i^* = t+1\text{”} \\ \text{soit} & \text{“}T_i^* = t+1 \text{ et } T+1 = t+1\text{”} \\ \text{soit} & \underbrace{\text{“}T_i^* = T+1 \text{ et } T+1 > t+1\text{”}}_{\text{événement impossible}} \end{cases}
\end{aligned}$$

Donc

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

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

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

En conséquence on a

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

On en tire que

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

de sorte que

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

et donc

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

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

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

☞ Q5 La vraisemblance s'écrit cette fois

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

On constate alors que $\phi : ((\lambda,\gamma)\mapsto\lambda r^\gamma)$ n'est jamais injective (car $\forall r\in\mathbb{R}^{+*}, \phi(1,1)=r=\phi\left(\frac{1}{2},\sqrt{r}\right)$), de sorte que la vraisemblance $\left((\lambda,\gamma)\mapsto L_{T_1^{**},\ldots,T_n^{**}}(t_1,\ldots,t_n\ ;\ \gamma,\lambda)\right)$ ne l'est pas non plus, de sorte que le paramètre (λ,γ) 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 $Y_1,\ldots,Y_{n+1}$ i.i.d. de densité $f_{\mathcal{E}}(\cdot,\lambda)$.
Définissons pour $k\in[\![1,n+1]\!]$ $S_k=\sum_{i=1}^k Y_i$.
Constatant que $S_{n+1}=S_n+Y_{n+1}$, définissons $\phi : \begin{pmatrix} \mathbb{R}^{+*2} & \to & \mathbb{R}^{+*}\times]0,1[\\ (x,y) & \mapsto & (x+y, \frac{x}{x+y}) \end{pmatrix}$. Alors ϕ est un C^∞-difféomorphisme, de jacobien $\left|\begin{pmatrix} \text{y} & \text{x} \\ \text{1-y} & \text{x}\end{pmatrix}^{-1}\right| = \frac{1}{x}$.

Or $L_{S_n,Y_{n+1}}(s,y) = \left(\lambda^n\frac{s^{n-1}}{(n-1)!}e^{-\lambda s}\mathbb{1}_{s>0}\right)\cdot\left(\lambda e^{-\lambda y}\mathbb{1}_{y>0}\right)$, 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 $\phi^{-1}(u,v) = (uv, u(1-v))$, $L_{\phi(S_n,Y_{n+1})}(u,v) = \lambda^{n+1}\frac{(uv)^{n-1}}{(n-1)!}e^{-\lambda u}u\mathbb{1}_{u>0\ \wedge\ v\in]0,1[}$
.
En particulier d'après le théorème de projection $L_{(\phi(S_n,Y_{n+1}))_1}(u) = \lambda^{n+1}\frac{u^n}{n!}e^{-\lambda u}\mathbb{1}_{u>0}$, c'est-à-dire

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

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 $i \in [\![1,n]\!]$ $L_{Z_i,C_i}(z_i,c_i\ ;\ \lambda,\mu) = \lambda\mu e^{-\lambda z_i-\mu c_i}$, et donc

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

En particulier le modèle est exponentiel et une statistique exhaustive est

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

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

☞ Q3 On calcule successivement

$$\begin{aligned}
\frac{\partial^2 \ln L_{Z_1,\ldots,Z_n,C_1,\ldots,C_n}}{\partial\lambda^2} &= \frac{n}{\lambda^2}\\
\frac{\partial^2 \ln L_{Z_1,\ldots,Z_n,C_1,\ldots,C_n}}{\partial\lambda\partial\mu} &= 0\\
\frac{\partial^2 \ln L_{Z_1,\ldots,Z_n,C_1,\ldots,C_n}}{\partial\mu^2} &= \frac{n}{\mu^2}
\end{aligned}$$

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

$$\boxed{\sqrt{n}\begin{pmatrix}\widehat{\lambda}-\lambda \\ \widehat{\mu}-\mu\end{pmatrix} \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}\lambda^2 & 0\\ 0 & \mu^2\end{pmatrix}\right)}$$

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

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

En conséquence, $L_{\widehat{\lambda}}(u) = \frac{\lambda^n}{n!}\left(\frac{n}{u}\right)^{n+1} e^{-\lambda\frac{n}{u}}\mathbb{1}_{u>0}$.
De façon similaire, $L_{\widehat{\mu}}(v) = \frac{\mu^n}{n!}\left(\frac{n}{v}\right)^{n+1} e^{-\mu\frac{n}{v}}\mathbb{1}_{v>0}$.
Ainsi,

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

Par ailleurs on a pour $n \geq 2$

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

de sorte que

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

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

Enfin pour $n \geq 3$

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

et donc

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

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

<u>*Autre méthode (plus rapide) :*</u>

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

De même

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

Enfin, remarquant que toutes les $Z_1, \ldots, Z_n, C_1, \ldots, C_n$ sont indépendantes, on a $\mathbb{C}ov\left(\widehat{\lambda}, \widehat{\mu}\right) = 0$ de sorte que

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

☞ 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 $T(Z, C) =$

$\left(\frac{1}{n}\sum_{i=1}^n Z_i, \frac{1}{n}\sum_{i=1}^n C_i\right)$ 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,

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

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 $X_i = \mathrm{Min}\,(Z_i, C_i)\,,\ i = 1, ..., n.$

☞ 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 $X_i \rightsquigarrow \mathcal{E}(\lambda + \mu)$.

☞ Q2 Le modèle statistique s'écrit ici $\{\mathbb{R}^{+n}\ ,\ \mathcal{E}^{\otimes n}\}$.
Sa vraisemblance étant $L_{X_1,\ldots,X_n}(x_1, \ldots, x_n\ ;\ \lambda, \mu) = (\lambda+\mu)^n e^{-(\lambda+\mu)\sum_{i=1}^n x_i}$, la seule fonction identifiable de (λ, μ) est $\lambda + \mu$.

☞ Q3 Cherchons à estimer $\gamma = \lambda + \mu$.
(i) La maximisation de $L_{X_1,\ldots,X_n}(x_1, \ldots, x_n\ ;\ \gamma) = (\gamma)^n e^{-(\gamma)\sum_{i=1}^n x_i}$ 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 $L_{Z_1,\ldots,Z_n,C_1,\ldots,C_n}(z_1, \ldots, z_n, c_1, \ldots, c_n\ ;\ \lambda, \mu)$ conduit à estimer λ et μ individuellement, ce dont on tire une estimation convergente de $\lambda+\mu$, à savoir $\widehat{\gamma_{Z,C}} = \widehat{\lambda}+\widehat{\mu} = \frac{n}{\sum_{i=1}^n Z_i} + \frac{n}{\sum_{i=1}^n C_i}$
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
(i) $\frac{\partial \ln L_{X_1,\ldots,X_n}}{\partial \gamma}(x_1, \ldots, x_n\ ;\ \gamma) = \frac{n}{\gamma} - \sum_{i=1}^n x_i$,
puis $\frac{\partial^2 \ln L_{X_1,\ldots,X_n}}{\partial \gamma^2}(x_1, \ldots, x_n\ ;\ \gamma) = -\frac{n}{\gamma^2}$,

donc

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

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

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

En particulier selon la première composante,[2]

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

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, asympotiquement $\widehat{\gamma^{Z,C}}$ est préférable à $\widehat{\gamma^X}$, ce que l'on écrira sous la forme publicitaire

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

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

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

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 $\widehat{\theta}=\mathbb{E}\left(\frac{1}{X_1+\cdots+X_n} \mid Z_1+\cdots+Z_n=z, C_1+\cdots+C_n=c\right)$; alors d'après la partie **II** $\widehat{\theta}$ est un estimateur sans biais de $\frac{\lambda+\mu}{n-1}$. La statistique en $\widehat{\theta}(X_1,\ldots,X_n)$ 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 décorréllés, ainsi que $\frac{\widehat{\lambda}}{n}$ et $(\widehat{\theta}-\frac{\widehat{\mu}}{n})-\frac{\widehat{\lambda}}{n}$, et que $\frac{\widehat{\mu}}{n}$ et $(\widehat{\theta}-\frac{\widehat{\lambda}}{n})-\frac{\widehat{\mu}}{n}$, et ce en vertu du lemme suivant :

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

Si $\widehat{a}$ et $\widehat{b}$ sont deux estimateurs sans biais optimaux du même paramètre, alors $\mathbb{C}ov\left(\widehat{a}, \widehat{a-b}\right) = 0$

En effet, définissons pour $t \in \mathbb{R}$ $c_t = (1-t)\widehat{a} + t\widehat{b} = \widehat{a} + t\left(\widehat{a} - \widehat{b}\right)$.
Alors $\forall t \in \mathbb{R},\ \mathbb{V}(c_t) \geq \mathbb{V}\left(\widehat{b}\right)$ puisque $\widehat{b}$ est optimal ;
ceci s'écrit $\forall t \in \mathbb{R},\ \mathbb{V}(\widehat{a}) + 2t\mathbb{C}ov\left(\widehat{a}, \widehat{a} - \widehat{b}\right) + t^2\mathbb{V}\left(\widehat{a} - \widehat{b}\right)$ qui est une inégalité polynômiale en t qui ne p
de celui-ci est négatif ou nul, ce dont on tire le résultat.

Par suite,

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

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

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

c'est-à-dire

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

⋆
⋆ ⋆

[3]Ce résultat est généralement connu comme l'*unicité presque sûre* d'un estimateur optimal fonction d'une statistique totale.