Semaine 6

Exercice 42.

(i) L'estimateur Y/n est non-biaisé car

$$\mathbb{E}\left(\frac{Y}{n}\right) = \frac{np}{n} = p.$$

(ii) On cherche une fonction U telle que

$$\frac{1}{p} = \mathbb{E}_p[U(Y)] = \sum_{k=0}^{n} \binom{n}{k} U(k) p^k (1-p)^{n-k}, \qquad \forall p \in]0,1[.$$

Or, le membre droite de l'équation est un polynôme alors que le membre gauche ne l'est pas. Ainsi, une telle fonction U ne peut pas exister. (Un autre raisonnement serait de dire que la limite du membre gauche de l'équation lorsque $p \searrow 0$ est ∞.)

(iii) Pareil qu'en (ii) : supposons que $V(Y)$ soit un estimateur non biaisé de ϕ, c'est-à-dire que $\mathbb{E}_p(V(Y)) = \phi$. Nous avons alors

$$\sum_{k=0}^{n} \binom{n}{k} V(k) p^k (1-p)^{n-k} = E_p[V(Y)] = \phi = \log\left(\frac{p}{1-p}\right).$$

Le polynôme ci-dessus est de degré inférieur ou égale à n, tandis que ϕ n'est pas un polynôme de degré fini, nous obtenons donc une contradiction.

Exercice 43. Remarquons que $\overline{X}_n$ est un estimateur non biaisé pour λ, puisque

$$\mathbb{E}_\lambda(\overline{X}_n) = \frac{n\lambda}{n} = \lambda.$$

On va montrer que $\overline{X}_n$ atteint la borne de Cramér-Rao. Il suffit de calculer le logarithme de la loi de probabilité de Poisson, et de dériver :

$$I(\lambda) = \mathbb{E}\left(\frac{\partial \log f_\lambda(X)}{\partial \lambda}\right)^2 = \mathbb{E}\left(\frac{X}{\lambda} - 1\right)^2 = \frac{\mathbb{E}\, X^2}{\lambda^2} - 2\frac{\mathbb{E}\, X}{\lambda} + 1 = \frac{\lambda^2 + \lambda - 2\lambda^2 + \lambda^2}{\lambda^2} = \frac{1}{\lambda}.$$

Ainsi $I(\lambda) = 1/\lambda$. Comme $\overline{X}_n$ est un estimateur non biaisé de λ, la borne de Cramér–Rao est

$$\mathrm{Var}_\lambda(\overline{X}_n) \geq \frac{1}{nI(\lambda)} = \frac{\lambda}{n}.$$

Or $\mathrm{Var}_\lambda(\overline{X}_n) = \mathrm{Var}(X)/n = \lambda/n$, donc $\overline{X}_n$ atteint cette borne.
Pour S_n^2, on effectue la manipulation suivante (voir l'exercice 4, série 3) :

$$\begin{aligned} S_n^2 &= \frac{1}{n-1} \sum_{i=1}^{n} X_i^2 - 2X_i\overline{X} + (\overline{X})^2 \\ &= \frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n(\overline{X})^2\right) \\ &= \frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - \frac{(\sum_{i=1}^n X_i)^2}{n}\right). \end{aligned}$$

Puisque $X_i \sim Poi(\lambda)$,

$$\mathbb{E}_\lambda(X_i^2) = \mathrm{Var}_\lambda(X_i) + (\mathbb{E}_\lambda(X_i))^2 = \lambda + \lambda^2.$$

D'après exercice 2, série 4 (exercice 22 du livre), on sait que $Z = \sum_{i=1}^n X_i \sim Poi(n\lambda)$, et donc $\mathbb{E}_\lambda(Z^2) = \mathrm{Var}_\lambda(Z) + (\mathbb{E}_\lambda Z)^2 = n\lambda + (n\lambda)^2$.
Par la linéarité de l'espérance, on écrit

$$\begin{aligned}
\mathbb{E}_\lambda(S_n^2) &= \frac{1}{n-1}\left(\sum_{i=1}^n \mathbb{E}_\lambda(X_i^2) - \frac{1}{n}\,\mathbb{E}_\lambda\left(\sum_{i=1}^n X_i\right)^2\right) \\
&= \frac{1}{n-1}\left(n(\lambda+\lambda^2) - \frac{n\lambda + n^2\lambda^2}{n}\right) \\
&= \frac{1}{n-1}\left((n-1)\lambda + n\lambda^2 - n\lambda^2\right) \\
&= \lambda.
\end{aligned}$$

Autrement dit, S_n^2 est lui aussi un estimateur non biaisé pour λ. Puisque $\overline{X}_n$ atteint la borne de Cramér-Rao, on sait que

$$\mathrm{Var}\,\widehat{\lambda}_n \leq \mathrm{Var}\,S_n^2.$$

Un calcul exacte de la variance de S_n^2 est possible, en utilisant $\mathrm{Var}\,S_n^2 = \mathbb{E}[S_n^2]^2 - [\mathbb{E}S_n^2]^2$, mais fastidieux.

Exercice 44.

(i) On a que $X_i,\ i = 1, \ldots, n$ sont i.i.d, et $\mathbb{E}_\lambda(X_i) = \frac{1}{\lambda}$. Donc, par la loi des grands nombres,

$$\overline{X} = \frac{1}{n}X_i \xrightarrow{p} \frac{1}{\lambda}.$$

En utilisant la theoréme de l'application continue avec la fonction $g(x) = \frac{1}{x}$, on a

$$\overline{X}^{-1} \xrightarrow{p} \lambda,$$

ce qui dit justement que $\widehat{\lambda}_n = \overline{X}^{-1}$ est consistent pour λ.

(ii) On utilise le fait que $Z = \sum_{i=1}^n X_i \sim Gamma(n, \lambda)$, et on écrit $\widehat{\lambda}_n = n/Z$. Ainsi

$$\begin{aligned}
\mathbb{E}_\lambda(\hat{\lambda}_n) &= n\int_0^\infty \frac{1}{z} f_{\lambda,n}(z)dz = n\int_0^\infty \frac{1}{z}\cdot\frac{1}{\Gamma(n)}\lambda \mathrm{e}^{-\lambda z}(\lambda z)^{n-1}dz \\
&= \frac{n\Gamma(n-1)}{\Gamma(n)}\lambda\int_0^\infty \frac{1}{\Gamma(n-1)}\lambda\mathrm{e}^{-\lambda z}(\lambda z)^{n-2}dz = \frac{n}{n-1}\lambda.
\end{aligned}$$

L'estimateur $\hat{\lambda}_n^{NB} = \frac{n-1}{n}\hat{\lambda}_n$ est donc non biaisé. La dernière égalité vient du fait que l'intégrale vaut 1, et que $\Gamma(x) = (x-1)\Gamma(x-1)$ pour tout $x > 1$.

(iii) Calculons

$$\begin{aligned}
\mathbb{E}_\lambda(\hat{\lambda}_n^2) &= n^2\int_0^\infty \frac{1}{z^2} f_{\lambda,n}(z)dz = n^2\int_0^\infty \frac{1}{z^2}\cdot\frac{1}{\Gamma(n)}\lambda\mathrm{e}^{-\lambda z}(\lambda z)^{n-1}dz \\
&= \frac{n^2\Gamma(n-2)}{\Gamma(n)}\lambda^2\int_0^\infty \frac{1}{\Gamma(n-2)}\lambda\mathrm{e}^{-\lambda z}(\lambda z)^{n-3}dz = \frac{n^2}{(n-1)(n-2)}\lambda^2.
\end{aligned}$$

Ainsi

$$\begin{aligned}\mathrm{Var}_\lambda(\hat{\lambda}_n) &= \mathbb{E}_\lambda(\hat{\lambda}_n^2) - [\mathbb{E}_\lambda(\hat{\lambda}_n)]^2 \\ &= \frac{n^2}{(n-1)(n-2)}\lambda^2 - \frac{n^2}{(n-1)^2}\lambda^2 \\ &= \frac{n^2}{(n-1)^2(n-2)}\lambda^2.\end{aligned}$$

(iv) L'information de Fisher $I(\lambda)$ est

$$\begin{aligned}I(\lambda) &= \mathbb{E}\left[\left\{\frac{\partial}{\partial\lambda}\log\left(\lambda\exp\left(-\lambda X_1\right)\right)\right\}^2\right] \\ &= \mathbb{E}\left[\left\{\frac{1}{\lambda} - X_1\right\}^2\right] \\ &= \mathbb{E}\left[\frac{1}{\lambda^2} - \frac{2}{\lambda}X_1 + X_1^2\right] = \frac{1}{\lambda^2},\end{aligned}$$

car $X_1 \sim Exp(\lambda)$ dont l'éspérance est $1/\lambda$ et la variance $1/\lambda^2$. La borne de Cramér–Rao est donc $(nI(\lambda))^{-1} = \lambda^2/n$.
Comme $\mathrm{Var}_\lambda(\hat{\lambda}_n^{\mathrm{NB}}) = \lambda^2/(n-2) > \lambda^2/n$, l'estimateur $\hat{\lambda}_n^{\mathrm{NB}}$ n'atteint (tout juste) pas la borne de Cramér-Rao.