Correction de la Feuille N° 6

Exercice 1. *Moyenne et variance empiriques*

1. En utilisant que les X_i sont i.i.d., on calcule

$$\begin{aligned}
\mathbb{E}[\bar{X}_n] &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[X_i] = m,\\
\mathbf{Var}(\bar{X}_n) &= \frac{1}{n^2}\sum_{i=1}^{n}\mathbf{Var}(X_i) = \frac{\sigma^2}{n},\\
\mathbb{E}[(\bar{X}_n - m)^2] &= \mathbb{E}[(\bar{X}_n - \mathbb{E}[\bar{X}_n])^2] = \mathbf{Var}(\bar{X}_n) = \frac{\sigma^2}{n}.
\end{aligned}$$

2. Par la loi des grands nombres (puisque $\mathbb{E}[|X_1|] < \infty$), on a $\bar{X}_n \xrightarrow{P} \mathbb{E}[X_1] = m$ lorsque $n \to \infty$. D'après le théorème central limite (car $\mathbb{E}[X_1^2] < \infty$), on a

$$\sqrt{n}(\bar{X}_n - m) = \sqrt{n}(\bar{X}_n - \mathbb{E}[X_1]) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \mathbf{Var}(X_1)) = \mathcal{N}(0, \sigma^2).$$

3. On calcule

$$\begin{aligned}
\mathbb{E}[s_X^2] &= \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2\right] = \frac{1}{n}n\mathbb{E}\left[(X_1 - \bar{X}_n)^2\right]\\
&= \mathbb{E}\left[X_1^2 - 2X_1\bar{X}_n + (\bar{X}_n)^2\right]\\
&= \mathbb{E}[X_1^2] - \frac{2}{n}(\mathbb{E}[X_1^2] + (n-1)(\mathbb{E}[X_1])^2) + \frac{1}{n^2}(n\mathbb{E}[X_1^2] + n(n-1)(\mathbb{E}[X_1])^2)\\
&= \frac{n-1}{n}\left(\mathbb{E}[X_1^2] - (\mathbb{E}[X_1])^2\right)\\
&= \frac{n-1}{n}\sigma^2.
\end{aligned}$$

On en déduit que $\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$ est un estimateur sans biais de σ^2, i.e. $\mathbb{E}[\hat{\sigma}^2] = \sigma^2$.

4. On constate d'abord que

$$\begin{aligned}
s_X^2 &= \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 = \frac{1}{n}\sum_{i=1}^{n}X_i^2 - \frac{2}{n}\sum_{i=1}^{n}X_i\bar{X}_n + \frac{1}{n}\sum_{i=1}^{n}\bar{X}_n^2\\
&= \frac{1}{n}\sum_{i=1}^{n}X_i^2 - \bar{X}_n^2 + \bar{X}_n^2 = \frac{1}{n}\sum_{i=1}^{n}X_i^2 - \bar{X}_n^2.
\end{aligned}$$

D'après la LGN, $\frac{1}{n}\sum_{i=1}^{n}X_i^2 \xrightarrow{P} \mathbb{E}[X^2] = \sigma^2 + m^2$ et $\bar{X}_n \xrightarrow{P} m$. Par le théorème de continuité, $(\bar{X}_n)^2 \xrightarrow{P} m^2$. Par conséquent, $s_X^2 \xrightarrow{P} \sigma^2$.

5. Par le théorème de Slutsky,

$$T_n = \sqrt{n-1}\frac{\bar{X}_n - m}{s_X} = \underbrace{\frac{\sqrt{n-1}\sigma}{\sqrt{n}s_X}}_{\xrightarrow{P} 1} \underbrace{\sqrt{n}\frac{\bar{X}_n - m}{\sigma}}_{\xrightarrow{\mathcal{L}} \mathcal{N}(0,1)} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1).$$

Notons q_γ^N le quantile d'ordre γ de la loi normale standard $\mathcal{N}(0,1)$. Soit Z une v.a. de loi $\mathcal{N}(0,1)$. On a alors,

$$\mathbb{P}(Z \in [q_{\alpha/2}^N, q_{1-\alpha/2}^N]) = 1 - \alpha.$$

Par la convergence en loi de T_n vers la loi normale standard (i.e. $T_n \xrightarrow{\mathcal{L}} Z$) on a

$$\begin{aligned} 1 - \alpha &= \lim_{n\to\infty} \mathbb{P}(T_n \in [q_{\alpha/2}^N, q_{1-\alpha/2}^N]) \\ &= \lim_{n\to\infty} \mathbb{P}\left(\sqrt{n-1}\frac{\bar{X}_n - m}{s_X} \in [q_{\alpha/2}^N, q_{1-\alpha/2}^N]\right) \\ &= \lim_{n\to\infty} \mathbb{P}\left(m \in \left[\bar{X}_n - \frac{s_X}{\sqrt{n-1}} q_{1-\alpha/2}^N, \bar{X}_n - \frac{s_X}{\sqrt{n-1}} q_{\alpha/2}^N\right]\right). \end{aligned}$$

On a montré que $\left[\bar{X}_n - \frac{s_X}{\sqrt{n-1}} q_{1-\alpha/2}^N, \bar{X}_n - \frac{s_X}{\sqrt{n-1}} q_{\alpha/2}^N\right]$ est un intervalle de confiance asymptotique de niveau $1-\alpha$ pour m.

6. Si $X_i \sim \mathcal{N}(m, \sigma^2)$ i.i.d., alors $\bar{X}_n$ suit une loi normale en tant que combinaison linéaire de variables aléatoires gaussiennes indépendantes. Plus précisément, $\bar{X}_n \sim \mathcal{N}(\mathbb{E}[\bar{X}_n], \mathbf{Var}(\bar{X}_n)) = \mathcal{N}(m, \sigma^2/n)$.

7. Rappel : Soit $(U_1, \ldots, U_n)^T$ un vecteur gaussien. On a $\mathbf{Cov}(U_i, U_j) = 0$ si et seulement si U_i et U_j sont indépendants. On a le même résultat pour des sous-vecteurs d'un vecteur gaussien. Comme $X_i \sim \mathcal{N}(m, \sigma^2)$ i.i.d., le vecteur $\mathbf{X} = (X_1, \ldots, X_n)^T$ est un vecteur gaussien de loi

$$\mathbf{X} \sim \mathcal{N}_n\left(\begin{pmatrix} \mu \\ \vdots \\ \mu \end{pmatrix}, \begin{pmatrix} \sigma^2 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma^2 \end{pmatrix}\right).$$

Le vecteur

$$\mathbf{Y} = \begin{pmatrix} \bar{X}_n \\ X_1 - \bar{X}_n \\ \vdots \\ X_n - \bar{X}_n \end{pmatrix}$$

est aussi vecteur gaussien, car il peut s'écrire comme un transformation linéaire du vecteur gaussien $\mathbf{X}$. Plus précisément, on a

$$\mathbf{Y} = A\mathbf{X} \quad \text{avec} \quad A = \begin{pmatrix} \frac{1}{n} & \cdots & \cdots & \frac{1}{n} \\ 1 - \frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ -\frac{1}{n} & 1 - \frac{1}{n} & \ddots & \vdots \\ \vdots & \ddots & \ddots & -\frac{1}{n} \\ -\frac{1}{n} & \cdots & -\frac{1}{n} & 1 - \frac{1}{n} \end{pmatrix}.$$

On calcule les covariances

$$\begin{aligned} \mathbf{Cov}(\bar{X}_n, X_i - \bar{X}_n) &= \mathbf{Cov}(\bar{X}_n, X_i) - \mathbf{Cov}(\bar{X}_n \bar{X}_n) \\ &= \frac{1}{n}\sum_{j=1}^n \mathbf{Cov}(X_j, X_i) - \mathbf{Var}(\bar{X}_n) \\ &= \frac{1}{n}\mathbf{Var}(X_i) - \frac{1}{n}\mathbf{Var}(X_1) \\ &= 0, \end{aligned}$$

en utilisant l'indépendance des X_i. Donc, $\bar{X}_n$ et $X_i-\bar{X}_n$ sont des v.a. mutuellement indépendantes pour $i=1,\dots,n$. Il s'ensuit que $\bar{X}_n$ et $(X_1-\bar{X}_n,\dots,X_n-\bar{X}_n)$ sont indépendants. En remarquant que s_X^2 est une fonction de $(X_1-\bar{X}_n,\dots,X_n-\bar{X}_n)$, on obtient l'indépendance de $\bar{X}_n$ et s_X^2.

8. Rappel : Définition de la loi de khi-deux : Soient $U_1,\dots,U_d$ des v.a. indépendantes de loi normale standard $\mathcal{N}(0,1)$. On dit que $S=\sum_{k=1}^d U_k^2$ suit la loi de khi-deux à d degrés de liberté et on note $S\sim\chi_d^2$.

   Remarquons que ns_X^2/σ^2 est le carré de la norme euclidienne du vecteur

$$\mathbf{Z} \;=\; \frac{1}{\sigma}\begin{pmatrix} X_1-\bar{X}_n \\ \vdots \\ X_n-\bar{X}_n \end{pmatrix} \;=\; P_{e^\perp}\begin{pmatrix} \frac{1}{\sigma}(X_1-m) \\ \vdots \\ \frac{1}{\sigma}(X_n-m) \end{pmatrix}$$

   où $P_{e^\perp}$ est la projection orthogonale sur l'orthogonal du vecteur $e=(1,1,\dots,1)^T/\sqrt{n}$ (de norme 1), c'est-à-dire que $P_{e^\perp}x=x-(x\cdot e)e$. Notons $G_i=(X_i-m)/\sigma$, de sorte que $\mathbf{G}=(G_1,\dots,G_n)^T$ est un vecteur gaussien standard. Pour toute matrice orthogonale Q (telle que $Q^TQ=I_n$), la loi de $Q\mathbf{G}$ est $\mathcal{N}(Q0,Q^TI_nQ)=\mathcal{N}(0,I_n)$. En effet, l'image par une application linéaire d'un vecteur gaussien étant gaussien, pour connaître sa loi il suffit de calculer son espérance et sa variance, or $\mathbb{E}(Q\mathbf{G})=Q\mathbb{E}(\mathbf{G})$ et

$$\mathbb{E}(Q\mathbf{G}(Q\mathbf{G})^T)=Q^T\mathbb{E}(\mathbf{G}^T\mathbf{G})Q=Q^TI_nQ=I_n\,.$$

   Autrement dit $Q\mathbf{G}$ a la même loi que $\mathbf{G}$ (c'est-à-dire que la loi normale standard est invariante par rotations). Considérons Q une rotation telle que $e=Qe_1$ avec $e_1=(1,0,\dots,0)^T$ premier vecteur de la base canonique (l'action des matrices orthogonales sur la sphère unité est transitive, autrement dit pour tout couple de vecteurs de norme 1 il existe une rotation qui envoit le premier sur le second). Alors

$$\mathbf{Z} \;\stackrel{loi}{=}\; P_{e^\perp}Q\mathbf{G} \qquad \text{et donc} \qquad |\mathbf{Z}|^2 \;\stackrel{loi}{=}\; |Q^TP_{e^\perp}Q\mathbf{G}|^2\,,$$

   puisque les transformations orthogonales conservent la norme euclidienne. Or, en utilisant que $Q^Te=e_1$,

$$Q^TP_{e^\perp}Qx=x-(Qx\cdot e)Q^Te=x-(x\cdot e_1)e_1=P_{e_1^\perp}x\,.$$

   On a donc obtenu

$$\frac{ns_X^2}{\sigma^2} \;\stackrel{loi}{=}\; |P_{e_1^\perp}\mathbf{G}|^2 \;=\; \sum_{k=2}^n G_i^2\,,$$

   qui suit bien la loi χ_{n-1}^2.

9. Rappel : Définition de la loi de Student : Soient $U\sim\chi_d^2$ et $V\sim\mathcal{N}(0,1)$ deux v.a. indépendantes. On dit que $T=V/\sqrt{U/d}$ suit la loi de Student à d degrés de liberté et on note $T\sim t_d$.

   On a

$$T_n=\sqrt{n-1}\frac{\bar{X}_n-m}{s_X}=\frac{\frac{\sqrt{n}(\bar{X}_n-m)}{\sigma}}{\sqrt{\frac{\frac{ns_X^2}{\sigma^2}}{n-1}}}.$$

   On a bien l'indépendance entre le numérateur et le dénominateur, et leurs distributions sont en sorte que T_n suit la loi de Student à $n-1$ degrés de liberté : $T_n\sim t_{n-1}$.

   Notons t_{n-1}^γ le quantile d'ordre γ de la loi de Student t_{n-1}. On a alors

$$\begin{aligned}
1-\alpha &= \mathbb{P}(T_n\in[t_{n-1}^{\alpha/2},t_{n-1}^{1-\alpha/2}]) \\
&= \mathbb{P}\left(\sqrt{n-1}\frac{\bar{X}_n-m}{s_X}\in[t_{n-1}^{\alpha/2},t_{n-1}^{1-\alpha/2}]\right) \\
&= \mathbb{P}\left(m\in\left[\bar{X}_n-\frac{s_X}{\sqrt{n-1}}t_{n-1}^{1-\alpha/2},\bar{X}_n-\frac{s_X}{\sqrt{n-1}}t_{n-1}^{\alpha/2}\right]\right).
\end{aligned}$$

Un intervalle de confiance (non-asymptotique) de niveau $1-\alpha$ pour m est donné par $\left[\bar{X}_n - \frac{s_X}{\sqrt{n-1}} t_{n-1}^{1-\alpha/2}, \bar{X}_n - \frac{s_X}{\sqrt{n-1}} t_{n-1}^{\alpha/2}\right]$.

Exercice 2. *Convergence de la loi de Student*
Afin de montrer ce résultat, considérons une suite de v.a. T_n définies de façon suivante. Soient $V, U_1, U_2, \ldots,$ des v.a. i.i.d. de loi normale standard $\mathcal{N}(0,1)$. Posons

$$T_n = \frac{V}{\sqrt{\frac{1}{n}\sum_{k=1}^n U_k^2}}.$$

Il est clair que $T_n \sim t_n$. Par ailleurs, par la loi des grands nombres, on a $\frac{1}{n}\sum_{k=1}^n U_k^2 \xrightarrow{P} \mathbb{E}[U_1^2] = \mathbf{Var}(U_1) = 1$ lorsque $n \to \infty$. Donc, $1/\sqrt{\frac{1}{n}\sum_{k=1}^n U_k^2} \xrightarrow{P} 1$. Or, $V \sim \mathcal{N}(0,1)$ et donc d'après le théorème de Slutzky, $T_n \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$.

Exercice 3. *Loi Gamma*

1. En effectuant le changement de variable $y = bx$, on obtient

$$\begin{aligned}\mathbb{E}[X^k] &= \int_{\mathbb{R}} x^k f(x)\mathrm{d}x = \frac{b^a}{\Gamma(a)}\int_0^\infty x^{k+a-1}\mathrm{e}^{-bx}\mathrm{d}x = \frac{b^a}{\Gamma(a)}\int_0^\infty \left(\frac{y}{b}\right)^{k+a-1}\mathrm{e}^{-y}\frac{\mathrm{d}y}{b}\\ &= \frac{1}{b^k\Gamma(a)}\int_0^\infty y^{k+a-1}\mathrm{e}^{-y}\mathrm{d}y = \frac{\Gamma(k+a)}{b^k\Gamma(a)} = \frac{(k+a-1)\cdot(k+a-2)\cdot\cdots\cdot a}{b^k},\end{aligned}$$

pour tout $k \geq 1$.

2. Par la question précédente on a $m_1(a,b) = \mathbb{E}[X] = a/b$ et $m_2(a,b) = \mathbb{E}[X^2] = (a+1)a/b^2$. Ainsi, on pose

$$\begin{aligned}\begin{cases}\bar{X}_n = \frac{a}{b}\\ \frac{1}{n}\sum_{i=1}^n X_i^2 = \frac{a(a+1)}{b^2}\end{cases} &\iff \begin{cases}\bar{X}_n = \frac{a}{b}\\ \frac{1}{n}\sum_{i=1}^n X_i^2 = \frac{(a+1)}{a}(\bar{X}_n)^2\end{cases}\\ &\iff \begin{cases}b = \frac{a}{\bar{X}_n}\\ a = \frac{(\bar{X}_n)^2}{\frac{1}{n}\sum_{i=1}^n X_i^2 - (\bar{X}_n)^2}\end{cases}\\ &\iff \begin{cases}b = \frac{\bar{X}_n}{s_X^2}\\ a = \frac{(\bar{X}_n)^2}{s_X^2}\end{cases}\end{aligned}$$

où $s_X^2 = \frac{1}{n}\sum_{i=1}^n X_i^2 - (\bar{X}_n)^2$. On en déduit que l'estimateur par la méthode des moments de a et b est donné par

$$\hat{a} = \frac{(\bar{X}_n)^2}{s_X^2}, \qquad \hat{b} = \frac{\bar{X}_n}{s_X^2}.$$

3. Pour des réalisations $x_1, \ldots, x_n$ i.i.d. de la loi Gamma $\Gamma(a,b)$ la fonction de vraisemblance vaut

$$\begin{aligned}\mathcal{L}(x_1,\ldots,x_n;a,b) &= \prod_{i=1}^n f_{(a,b)}(x_i) = \prod_{i=1}^n \frac{b^a}{\Gamma(a)} x_i^{a-1}\mathrm{e}^{-bx_i}\\ &= \frac{b^{na}}{(\Gamma(a))^n}\left(\prod_{i=1}^n x_i\right)^{a-1}\exp\left\{-b\sum_{i=1}^n x_i\right\}.\end{aligned}$$

La fonction de log-vraisemblance et ses dérivées partielles sont données par

$$\begin{aligned}
\ell(a,b) &= \log(\mathcal{L}(x_1,\dots,x_n;a,b)) \\
&= na\log b - n\log(\Gamma(a)) + (a-1)\sum_{i=1}^n \log x_i - b\sum_{i=1}^n x_i \\
\frac{\partial}{\partial a}\ell(a,b) &= n\log b - \frac{n\Gamma'(a)}{\Gamma(a)} + \sum_{i=1}^n \log x_i \\
\frac{\partial}{\partial b}\ell(a,b) &= \frac{na}{b} - \sum_{i=1}^n x_i,
\end{aligned}$$

où $\Gamma'(\cdot)$ désigne la dérivée de la fonction $\Gamma(\cdot)$. Il est clair que $\nabla\ell(a,b)=0$ n'a pas de solution explicite, notamment parce que les fonctions $\Gamma(\cdot)$ est donnée par un intégrale et $\Gamma'(\cdot)$ n'a pas d'expression explicite. Néanmoins, cela ne met pas en question l'existence de l'EMV ! Mais pour le calculer il faut utiliser des méthodes numériques comme par exemple la méthode de Newton-Raphson.

4. D'après la question précédente, le gradient est donné par

$$\nabla\ell(a,b) = \left(\frac{\partial}{\partial a}\ell(a,b), \frac{\partial}{\partial b}\ell(a,b)\right)^T = \left(n\log b - n(\log\Gamma(a))' + \sum_{i=1}^n \log x_i, \quad \frac{na}{b} - \sum_{i=1}^n x_i\right)^T.$$

Pour l'inverse de la Hessienne on obtient détermine d'abord les dérivées secondes :

$$\begin{aligned}
\frac{\partial^2}{\partial^2 a}\ell(a,b) &= -n\frac{\Gamma''(a)\Gamma(a)-(\Gamma'(a))^2}{(\Gamma(a))^2} = -n\left(\log\Gamma(a)\right)'' \\
\frac{\partial^2}{\partial^2 b}\ell(a,b) &= -\frac{na}{b^2} \\
\frac{\partial^2}{\partial a\partial b}\ell(a,b) &= \frac{n}{b},
\end{aligned}$$

Par la formule générale

$$\mathbf{A}^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{\det\mathbf{A}}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix},$$

on obtient

$$[H(a,b)]^{-1} = \begin{pmatrix} \frac{\partial^2}{\partial^2 a}\ell(a,b) & \frac{\partial^2}{\partial a\partial b}\ell(a,b) \\ \frac{\partial^2}{\partial a\partial b}\ell(a,b) & \frac{\partial^2}{\partial^2 b}\ell(a,b) \end{pmatrix}^{-1} = \frac{1}{n\left(1-a(\log\Gamma(a))''\right)}\begin{pmatrix} a & b \\ b & b^2(\log\Gamma(a))'' \end{pmatrix}$$

La méthode de Newton-Raphson est un algorithme itératif. Notons $(a^{(t)},b^{(t)})^T$ les valeurs actuelles de paramètres. Une itération de l'algorithme consiste à calculer les nouvelles valeurs des paramètres $(a^{(t+1)},b^{(t+1)})^T$ par la formule

$$(a^{(t+1)},b^{(t+1)})^T = (a^{(t)},b^{(t)})^T - [H(a^{(t)},b^{(t)})]^{-1}\nabla\ell(a^{(t)},b^{(t)}).$$