ENSTA ParisTech 1ère année
Statistiques (MA101), 2017-2018
christine.keribin@math.u-psud.fr, Ennio Fedrizzi, Jade Giguelay, Cyril Marzouk, Loïc Richier, Morgan Schmitz

TD2: Convergence d'estimateurs

Exercice 1.
Soit $(X_1, Y_1), \dots, (X_n, Y_n)$ un n-échantillon iid de loi mère telle que $\mathbb{E}(X_1^2)$ et $\mathbb{E}(Y_1^2)$ soient finis. Soit

$$C_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})$$

1. Montrer que $C_n = \frac{1}{n}\sum_{i=1}^{n} X_i Y_i - \bar{X}\bar{Y}$.
2. Montrer que C_n est biaisé pour estimer $\mathrm{cov}(X_1, Y_1)$.
3. Montrer que C_n est un estimateur consistant de $\mathrm{cov}(X_1, Y_1)$.
4. Proposer un estimateur consistant et sans biais de $\mathrm{cov}(X_1, Y_1)$.

Correction. *1.*

$$\begin{aligned}
C_n &= \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y}) \\
&= \frac{1}{n}\left(\sum_{i=1}^{n} X_i Y_i - \sum_i X_i \bar{Y} - \sum_i Y_i \bar{X} + n\bar{X}\bar{Y}\right) \\
&= \frac{1}{n}\sum_{i=1}^{n} X_i Y_i - 2\bar{X}\bar{Y} + \bar{X}\bar{Y} \\
&= \frac{1}{n}\sum_{i=1}^{n} X_i Y_i - \bar{X}\bar{Y}
\end{aligned}$$

2. Par linéarité de l'espérance et indépendance des $(X_1Y_1, \dots, X_nY_n)$:

$$\begin{aligned}
\mathbb{E}(C_n) &= \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(X_i Y_i) - \mathbb{E}(\bar{X}\bar{Y}) \\
&= \frac{n}{n}\mathbb{E}(X_1 Y_1) - \frac{1}{n^2}\mathbb{E}\left(\sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_j\right) \\
&= \mathbb{E}(X_1 Y_1) - \frac{1}{n^2}\left(\sum_{i=1}^{n} \mathbb{E}(X_i Y_i) + \sum_{i \neq j} \mathbb{E}(X_i Y_j)\right) \\
&= \mathbb{E}(X_1 Y_1) - \frac{n}{n^2}\mathbb{E}(X_1 Y_1) - \frac{n(n-1)}{n^2}\mathbb{E}(X_i)\mathbb{E}(Y_j) \\
&= \frac{n-1}{n}\mathrm{cov}(X_1, Y_1)
\end{aligned}$$

L'estimateur est donc biaisé, mais asymptotiquement sans biais.

3. *Les v.a. X_i et Y_i admettent une espérance donc* $\mathbb{E}[|X_iY_i|] \leqslant \frac{1}{2}\left(\mathbb{E}[X_i^2] + \mathbb{E}[Y_i^2]\right) < +\infty$. *Par indépendance des X_i entre eux, des Y_i entre eux et des X_iY_i entre eux, $1 \leqslant i \leqslant n$, la loi des grands nombres entraîne que*

$$\bar{X}_n \xrightarrow[n\to+\infty]{\mathbb{P}} \mathbb{E}[X], \quad \bar{Y}_n \xrightarrow[n\to+\infty]{\mathbb{P}} \mathbb{E}[Y], \quad et \quad \frac{1}{n}\sum_{i=1}^{n} X_iY_i \xrightarrow[n\to+\infty]{\mathbb{P}} \mathbb{E}[XY].$$

Nous en déduisons que $(\overline{C}_n)_{n\geqslant 1}$ converge en probabilité vers $\mathrm{cov}(X,Y)$, *d'où la consistance.*

4. *$\tilde{C}_n = (n-1)/nC_n$ est sans biais, et consistant car si T_n est une suite d'estimateur consistant de θ et a_n une suite de réels convergeant vers a, alors a_nT_n est un estimateur consistant de $a\theta$. Ce résultat se déduit du théorème de Slutsky qui est un peu plus général ou se démontre directement. Il sera considéré comme acquis par la suite.*

Si W_n est un estimateur consistant de θ et si $(a_n)_{n\geqslant 1}$ est une suite de réels convergeant vers a alors a_nW_n est un estimateur consistant de $a\theta$. En effet

$$\forall n \geqslant 1 \qquad |a_nW_n - a\theta| = |a_n(W_n-\theta) + (a_n - a)\theta)| \leqslant |a_n||W_n - \theta| + |a_n - a||\theta|.$$

Soit $\varepsilon > 0$. Il existe $n_\varepsilon \geqslant 1$ tel que, pour tout $n \geqslant n_\varepsilon$, $|a_n - a||\theta| \leqslant \varepsilon/2$ et $|a_n| \leqslant |a| + 1$. Donc

$$\forall n \geqslant n_\varepsilon \qquad |a_nW_n - a\theta| \leqslant (|a|+1)|W_n - \theta| + \frac{\varepsilon}{2}.$$

Par conséquent

$$\forall n \geqslant n_\varepsilon \qquad \mathbb{P}\big(|a_nW_n - a\theta| \geqslant \varepsilon\big) \leqslant \mathbb{P}\left((|a|+1)|W_n - \theta| \geqslant \frac{\varepsilon}{2}\right) \xrightarrow[n\to+\infty]{} 0.$$

Variante avec une hypothèse supplémentaire: L'estimateur $\tilde{C} = \frac{n}{n-1}C_n$ est sans biais. Si $E(X_i^4) < \infty$ et $\mathbb{E}(Y_i^4) < \infty$, $\mathrm{Var}(C_n) < \infty$. *On a:*

$$\mathrm{Var}(\tilde{C}) = \mathrm{Var}\left(\frac{n}{n-1}C_n\right) = \left(\frac{n}{n-1}\right)^2 \mathrm{Var}(C_n)$$

a le même comportement que celle de C_n et tend donc vers 0. *L'estimateur converge en moyenne quadratique, il est donc consistant.*

Exercice 2.
Soit un n-échantillon de loi exponentielle $\mathcal{E}(\lambda)$ de paramètre $\lambda > 0$ et de densité définie pour tout $x \geq 0$ par $f(x) = \lambda\exp(-\lambda x)$. On souhaite estimer λ.

1. Construire un estimateur $\widehat{\lambda}$ de λ par la méthode des moments. Est-il consistant?
2. Calculer l'estimateur du maximum de vraisemblance.
3. On rappelle que la somme $S = \sum_{i=1}^n X_i$ de n variables indépendantes de loi exponentielle $\mathcal{E}(\lambda)$ suit une loi Gamma $\mathcal{G}(n, \lambda)$ de densité définie pour $s \geq 0$ par
$$g(s) = \frac{\lambda^n}{\Gamma(n)} s^{n-1} \exp(-\lambda x)$$
Calculer le biais de l'estimateur $\widehat{\lambda}$.
4. Déterminer un estimateur $\widehat{\lambda}^*$ sans biais et un estimateur $\widehat{\lambda}_m$ qui minimise le risque quadratique parmi les estimateurs de la forme
$$\widehat{\lambda}_c = \frac{c}{\sum_{i=1}^n X_i}, \text{ où } c > 0.$$
$\widehat{\lambda}^*$ est-il préférable à $\widehat{\lambda}$? Est-il admissible?
5. $\widehat{\lambda}^*$ est-il consistant?
6. Montrer que $\sqrt{n}(\lambda/\widehat{\lambda} - 1) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$

Correction. *1. LGN ($\mathrm{Var}(X) = 1/\lambda^2 < \infty$) donc $\bar{X}$ l'estimateur empirique de l'espérance converge ps vers $\mathbb{E}(X_1) = 1/\lambda$. Par continuité de la fonction inverse $x \to 1/x$ sur $]0, +\infty[$, l'estimateur $\widehat{\lambda} = 1/\bar{X}$ converge vers λ.*

2. La log-vraisemblance des observations vaut $L(\lambda; X) = n \log \lambda - \lambda \sum_{i=1}^n X_i$. En la dérivant par rapport à λ, on trouve $\widehat{\lambda} = 1/\bar{X}$. La log-vraisemblance étant strictement concave, $dL(\lambda; X)/d\lambda^2 = -n/\lambda^2 < 0$, c'est bien un maximum. L'estimateur des moments et l'estimateur du maximum de vraisemblance sont identiques.

3. La fonction caractéristique de la loi exponentielle est
$$\psi_X(t) = \mathbb{E}(e^{itX}) = \int_0^\infty e^{itx} \lambda e^{-\lambda x}\, dx = (1 - it/\lambda)^{-1}$$
La fonction caractéristique de la somme de n variables indépendantes est le produit des fonctions caractéristiques de chaque variable
$$\psi_S(t) = \prod_{j=1}^n \psi_{X_i}(t) = \prod_{j=1}^n \left(\frac{\lambda}{\lambda - it}\right) = \left(\frac{\lambda}{\lambda - it}\right)^n$$
On reconnait la fonction caractéristique d'une loi $\Gamma(\lambda, n)$ de densité
$$f_S(s) = \frac{\lambda^n}{\Gamma(n)} s^{n-1} e^{-\lambda s} \mathbb{1}_{(s>0)}$$

$$\begin{aligned}
\mathbb{E}_\lambda\left(\frac{1}{S}\right) &= \int_0^\infty \frac{1}{s}\frac{\lambda^n}{\Gamma(n)} s^{n-1}e^{-\lambda s}\,ds \\
&= \frac{\lambda}{n-1}\int_0^\infty \frac{\lambda^{n-1}}{(n-2)!} s^{n-2}e^{-\lambda s}\,ds \\
&= \frac{\lambda}{n-1}
\end{aligned}$$

On en déduit le biais de $\widehat{\lambda} = \mathbb{E}(\widehat{\lambda}) - \lambda = \lambda(n/(n-1)-1) = \lambda/(n-1)$ *si* $n > 1$. *Sinon,* $\mathbb{E}(\widehat{\lambda}) = \infty$.
Remarque: $\bar{X}$ *est non biaisé pour estimer* $\mathbb{E}(X) = 1/\lambda$*, mais* $1/\bar{X}$ *est biaisé pour estimer* λ *.*

4. *On propose l'estimateur sans biais* $\widehat{\lambda}^* = (n-1)/S$. *On calcule le moment d'ordre 2 de* S

$$\mathbb{E}_\lambda\left(\frac{1}{S^2}\right) = \frac{\lambda^2}{(n-1)(n-2)}\int_0^\infty \frac{\lambda^{n-2}}{(n-3)!} s^{n-3}e^{-\lambda s}\,ds = \frac{\lambda^2}{(n-1)(n-2)}.$$

soit

$$\mathrm{Var}\left(\frac{1}{S_n}\right) = \frac{\lambda^2}{(n-1)(n-2)} - \left(\frac{\lambda}{n-1}\right)^2 = \frac{\lambda^2}{(n-1)^2(n-2)}$$

Donc, pour tout $c > 0$,

$$\begin{aligned}
R(\widehat{\lambda}_c) &= \mathbb{E}_\lambda((S-\lambda)^2) = \mathrm{Var}(\widehat{\lambda}_c) + [B(\widehat{\lambda}_c)]^2 \\
&= \frac{\lambda^2 c^2}{(n-1)^2(n-2)} + \lambda^2\left(\frac{c}{n-1} - 1\right)^2 \\
&= \frac{\lambda^2}{(n-1)(n-2)}(c^2 - 2(n-2)c + (n-1)(n-2))
\end{aligned}$$

est minimum en pour $2c - 2(n-2) = 0$*, soit* $c = n-2$. *On a:*

$$\begin{aligned}
R(\widehat{\lambda}_m) &= R(\widehat{\lambda}_{n-2}) = \lambda^2/(n-1) \\
R(\widehat{\lambda}^*) &= R(\widehat{\lambda}_{n-1}) = \lambda^2/(n-2) \\
R(\widehat{\lambda}) &= R(\widehat{\lambda}_n) = \lambda^2(n+2)/[(n-2)(n-1)]
\end{aligned}$$

soit

$$R(\widehat{\lambda}) > R(\widehat{\lambda}^*) > R(\widehat{\lambda}_m)$$

Donc $\widehat{\lambda}^*$ *est préférable à* $\widehat{\lambda}$*, mais il est dominé par* $\widehat{\lambda}_m = (n-2)/S$ *qui est de risque strictement inférieur, quelque soit* λ. *Il n'est donc pas admissible.*

5. *Son risque tend vers 0, donc il est convergent en moyenne quadratique, donc consistant.*

   Rem: Etant préférable à $\widehat{\lambda}$ *et* $\widehat{\lambda}$ *étant consistant,* $\widehat{\lambda}^*$ *est aussi consistant*

6. *Application directe du TCL (somme des composantes d'un échantillon iid de loi mère de variance finie),* $\sqrt{n}(\bar{X}-1/\lambda)/\sqrt{\mathrm{Var}(X_1)} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$ *avec* $\mathrm{Var}(X_1) = 1/\lambda^2$.

Exercice 3.
L'angle θ selon lequel sont émis des électrons lors de désintégrations est tel que $\cos\theta$ suit une distribution de densité

$$f : x \in \mathbb{R} \longmapsto \frac{1+\alpha x}{2}\, \mathbb{1}[-1,1](x).$$

Le paramètre α est un coefficient de polarisation; pour des raisons physiques, $|\alpha| \leq 1/3$.

1. Vérifier à quelle condition f est une densité.

2. Déterminer l'équation permettant le calcul de l'estimateur du maximum de vraisemblance de α à partir de n mesures X_i. Comment la résoudre?

3. Construire un estimateur $\widehat{\alpha}_M$ de α par la méthode des moments et montrer qu'il est consistant. Déterminer la loi asymptotique de $\widehat{\alpha}_M$.

Correction. 1. *$f(x) \geq 0$ et $\int_{-1}^{1} f(x)dx = 1$*

2. *La log-vraisemblance vaut $L(\alpha; X) = \sum_{i=1}^{n} \log((1+\alpha X_i)/2)$. On résout $dL(\alpha; X)/d\alpha = 0$, soit $\sum_{i=1}^{n} X_i/(1+\alpha X_i) = 0$ qui est une équation non linéaire en α. La drivée seconde est bien négative. Elle n'a pas de solution explicite, et on utilise un schéma numérique (par ex un algorithme de Newton) pour calculer l'EMV. La log-vraisemblance étant strictement concave, le point d'initialisation de l'algorithme n'a pas d'importance.*

3. *$\mathbb{E}(X) = \alpha/3$ donc $\widehat{\alpha}_M = 3\bar{X}$. LGN pour consistance (la loi mère est de carré intégrable $\mathrm{Var}(X) = (3-\alpha^2)/9 < \infty$).*

   TCL pour obtenir la loi asymptotique. Pour n suffisamment grand:

   $$\widehat{\alpha}_M \overset{appr}{\sim} \mathcal{N}\left(\alpha, \frac{3-\alpha^2}{n}\right)$$

   C'est une loi approchée à distance finie.

Exercice 4.
On considère $X_1, \ldots, X_n$ un échantillon dont la loi a pour densité

$$f(x) = \left((1-\theta) + \frac{\theta}{2\sqrt{x}}\right) 1_{[0,1]}(x), \quad \theta \in [0,1].$$

1. Déterminer par la méthode des moments l'estimateur $\widehat{\theta}$ de θ. Quel est son risque?

2. Etudier sa consistance et sa convergence en loi.

christine.keribin@math.u-psud.fr, Ennio Fedrizzi, Jade Giguelay, Cyril Marzouk, Loïc Richier, Morgan Schmitz

3. Proposer un estimateur consistant $\widehat{V}_n$ de la variance de $\widehat{\theta}$. Quelle est la loi asymptotique de la statistique $T_n = (\widehat{\theta} - \theta)/\sqrt{\widehat{V}_n}$?

Correction. *1.* $\mathbb{E}(X) = \frac{1}{2} - \frac{1}{6}\theta$. *Si l'on dispose d'un n-échantillon de même loi que* X, $\widehat{\theta} = 6(\frac{1}{2} - \overline{X})$.

$$R(\widehat{\theta}, \theta) = \mathrm{Var}(\widehat{\theta}) = 36\,\mathrm{Var}(\overline{X}) = \frac{36}{n}\mathrm{Var}(X) = \frac{36}{n}\left(\frac{1}{12} + \frac{\theta}{30} - \frac{\theta^2}{36}\right) = \frac{1}{n}\left(3 + \frac{6}{5}\theta - \theta^2\right)$$

2. Applications de la LGN et du TLC + lemme de l'application continue (transf. linéaire).

3. On estime la variance de $\widehat{\theta}$ *par* $\widehat{V}_n = \frac{1}{n}\left(3 + \frac{6}{5}\widehat{\theta} - \widehat{\theta}^2\right)$, *on obtient un estimateur consistant par application de LGN et du lemme de l'application continue.*

La statistique normalisée $\dfrac{\widehat{\theta} - \theta}{\sqrt{\widehat{V}_n}}$ *a encore pour loi* asymptotique *une loi gaussienne centrée réduite par le lemme de Slutsky.*