Warm-up

Soit $(X_1, \dots, X_n)$ des variables alàtoires i.i.d. réelles d'espérance μ et de variance σ^2, on suppose que X_1 a un moment d'ordre 4 fini. On note

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i\,, \quad s_n^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X}_n)^2\,, \quad T_n = \sqrt{n-1}\frac{\bar{X}_n - \mu}{s_n}\,.$$

1. Calculer $\mathbb{E}[\bar{X}_n]$, $\mathbb{V}[\bar{X}_n]$ et $\mathbb{E}[(\bar{X}_n - \mu)^2]$.

   En utilisant le fait que les variables $(X_1, \dots, X_n)$ sont i.i.d. on obtient

$$\begin{aligned}
\mathbb{E}[\bar{X}_n] &= \mathbb{E}[X_1] = \mu\,. \\
\mathbb{V}[\bar{X}_n] &= \frac{1}{n^2} n \mathbb{V}[X_1] = \frac{\sigma^2}{n}\,. \\
\mathbb{E}[(\bar{X}_n - \mu)^2] &= \mathbb{E}[(\bar{X}_n - \mathbb{E}[\bar{X}_n])^2] = \mathbb{V}[\bar{X}_n] = \frac{\sigma^2}{n}\,.
\end{aligned}$$

2. Montrer que $(\bar{X}_n)_{n\geq 0}$ converge en probabilité vers μ et que $(\sqrt{n}(\bar{X}_n - \mu))_{n\geq 0}$ converge en loi vers Z où $Z \sim \mathcal{N}(0, \sigma^2)$.

   Il suffit d'appliquer la loi des grands nombres et le théorème central limite.

3. Calculer $\mathbb{E}[s_n^2]$ et montrer que $(s_n^2)_{n\geq 0}$ converge en probabilité vers σ^2.

   Par définition,

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

   On en déduit que $ns_n^2/(n-1)$ est un estimateur sans biais de σ^2. On peut également montrer aisément que s_n^2 converge en probabilité vers σ^2.

4. Déterminer la limite en loi de $(T_n)_{n\geq 0}$.

   Par définition,

$$T_n = \sqrt{n-1}\frac{\bar{X}_n - \mu}{s_n} = \frac{\sqrt{n-1}\sigma}{\sqrt{n}s_n}\sqrt{n}\frac{\bar{X}_n - \mu}{\sigma}\,.$$

   En utilisant les questions précédentes et en appliquant le lemme de Slutsky, on en déduit que $(T_n)_{n\geq 0}$ converge en loi vers Z où $Z \sim \mathcal{N}(0, 1)$.

5. Nous supposons dans la suite que les $(X_i)_{1\le i\le n}$ sont gaussiennes. Quelle est la loi de $\bar{X}_n$? Montrer que $\bar{X}_n$ et s_n^2 sont indépendantes.

   Dans ce cas, $\bar{X}_n$ est une combinaison linéaire de variables aléatoires gaussiennes indépendantes. On en déduit que $\bar{X}_n \sim \mathcal{N}(\mu, \sigma^2/n)$. Pour la seconde question, notons que $s_n^2 = \|Y\|_2^2/n$ où Y est le vecteur de $\mathbb{R}^n$ dont la i-ème composante est $X_i - \bar{X}_n$, pour $1 \le i \le n$. Pour obtenir le résultat, il suffit donc de montrer que Y et $\bar{X}_n$ sont indépendantes. Pour cela il suffit de remarquer que le vecteur $(\bar{X}_n, Y)^\top$ est un vecteur gaussien (comme transformation linéaire d'un vecteur gaussien). Ainsi, $\bar{X}_n$ et Y sont indépendantes si et seulement si pour tout $1 \le i \le n$, $\mathrm{Cov}(\bar{X}_n, Y_i) = 0$. Or, pour tout $1 \le i \le n$,

   $$\mathrm{Cov}(\bar{X}_n, Y_i) = \mathrm{Cov}(\bar{X}_n, X_i - \bar{X}_n) = \mathrm{Cov}(\bar{X}_n, X_i) - \mathbb{V}[\bar{X}_n] = \frac{1}{n}\sum_{j=1}^n \mathrm{Cov}(X_j, X_i) - \frac{\sigma^2}{n} = 0\,,$$

   ce qui permet de conclure.

   Ce résultat peut également être obtenu directement par application du théorème de Cochran.

Loi Gamma

Soit $a > 0$ et $b > 0$. La densité de la loi Gamma de paramètres a et b est définie sur $\mathbb{R}_+^*$ par :

$$f_{a,b} : x \mapsto \frac{b^a}{\Gamma(a)} x^{a-1} \mathrm{e}^{-bx}\,,$$

où Γ est la fonction gamma. Soit $(X_i)_{1\le i\le n}$ des variables i.i.d. de loi Gamma de paramètres a et b.

1. Calculer $\mathbb{E}[X_1^k]$ pour $k \ge 1$. Notons $m_1(a,b) = \mathbb{E}[X_1]$ et $m_1(a,b) = \mathbb{E}[X_1^2]$.

   Pour tout $k \ge 1$,

   $$\begin{aligned}
   \mathbb{E}[X_1^k] &= \int_0^\infty x^k \frac{b^a}{\Gamma(a)} x^{a-1}\mathrm{e}^{-bx}\mathrm{d}x\,,\\
   &= \frac{b^a}{\Gamma(a)}\int_0^\infty x^{k+a-1}\mathrm{e}^{-bx}\mathrm{d}x\,,\\
   &= \frac{b^{a-1}}{\Gamma(a)}\int_0^\infty \left(\frac{y}{b}\right)^{k+a-1}\mathrm{e}^{-y}\mathrm{d}y\,,\\
   &= \frac{1}{b^k\Gamma(a)}\int_0^\infty y^{k+a-1}\mathrm{e}^{-y}\mathrm{d}y\,,\\
   &= \frac{\Gamma(a+k)}{b^k\Gamma(a)}\,,\\
   &= \frac{\prod_{j=0}^{k-1}(a+j)}{b^k}\,.
   \end{aligned}$$

2. Calculer un estimateur de a et b par la méthode des moments en résolvant le système

   $$\begin{aligned}
   &\frac{1}{n}\sum_{i=1}^n X_i = m_1(a,b)\,,\\
   &\frac{1}{n}\sum_{i=1}^n X_i^2 = m_2(a,b)\,.
   \end{aligned}$$

D'après la question précédente, le système devient :

$$\frac{1}{n}\sum_{i=1}^{n} X_i = \frac{a}{b}\,,$$
$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 = \frac{a(a+1)}{b^2}\,.$$

En utilisant les notations de l'exercice précédent, nous obtenons les estimateurs

$$\widehat{a}_n = \frac{\bar{X}_n^2}{s_n^2} \quad \text{et} \quad \widehat{b}_n = \frac{\bar{X}_n}{s_n^2}\,.$$

3. Calculer la logvraisemblance $\ell : (a,b) \mapsto \log p_\theta(X_1,\ldots,X_n)$ où $\theta = (a,b)$ et où p_θ est la densité jointe des variables $(X_i)_{1\le i\le n}$.

   Pour tout $a > 0$ et $b > 0$, puisque les $(X_i)_{1\le i\le n}$ sont i.i.d. de loi gamma de paramètres a et b,

$$\ell(a,b) = an\log b - n\log\Gamma(a) + (a-1)\sum_{i=1}^{n}\log(X_i) - b\sum_{i=1}^{n} X_i\,.$$

4. Calculer le gradient et la matrice hessienne de ℓ et en déduire un algorithme itératif pour estimer θ par la méthode de Newton-Raphson.

   D'après la question précédente, pour tout $a > 0$ et $b > 0$,

$$\frac{\partial}{\partial a}\ell(a,b) = n\log b - n\frac{\Gamma'(a)}{\Gamma(a)} + \sum_{i=1}^{n}\log(X_i)\,,$$
$$\frac{\partial}{\partial b}\ell(a,b) = \frac{an}{b} - \sum_{i=1}^{n} X_i\,.$$

   Le gradient de ℓ est donc

$$\nabla\ell(a,b) = \begin{pmatrix} n\log b - n\frac{\Gamma'(a)}{\Gamma(a)} + \sum_{i=1}^{n}\log(X_i) \\ \frac{an}{b} - \sum_{i=1}^{n} X_i \end{pmatrix}.$$

   Et la matrice hessienne de ℓ est

$$H(a,b) = \begin{pmatrix} -n\frac{\Gamma''(a)\Gamma(a)-(\Gamma'(a))^2}{\Gamma(a)^2} & \frac{n}{b} \\ \frac{n}{b} & -\frac{an}{b^2} \end{pmatrix}.$$

Nous pouvons approcher l'estimateur du maximum de vraisemblance en proposant une méthode itérative en initialisant l'estimateur $\theta_0 = (a_0,b_0)^\top$ aléatoirement puis en définissant pour tout $k \ge 0$,

$$\theta_{k+1} = \theta_k - H(\theta_k)^{-1}\nabla\ell(\theta_k)\,.$$

L'algorithme est en général arrêté après un nombre maximum d'itérations ou si $\|\theta_{k+1}-\theta_k\|_2 < \varepsilon$ pour un seuil ε fixé.