

Exercices de statistique avec le logiciel R (S4)

Exercice 5. Loi des grands nombres

Il s'agit d'illustrer la loi forte des grands nombres : si X_1, X_2, \dots, X_n est une suite de variables aléatoires i.i.d. avec $\mathbb{E}[X_1] < \infty$, alors

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{p.s.} \mathbb{E}[X_1].$$

- Générer une réalisation x_1, \dots, x_n de la suite X_1, \dots, X_n avec $n = 50000$ dans le cas d'une loi normale $X_1 \sim \mathcal{N}(0, 1)$.
- Calculer $\bar{x}_m = \frac{1}{m} \sum_{i=1}^m x_i$ pour $m = 1, \dots, n$.
- Tracer \bar{x}_m en fonction de m . Conclusion ?
- Procéder de même avec la loi de Cauchy. Conclusion ?

Exercice 6. Théorème central limite (TCL)

Il s'agit d'illustrer maintenant le théorème central limite : si X_1, X_2, \dots, X_n est une suite de variables aléatoires i.i.d. avec $\mathbb{E}[X_1] < \infty$ et $\mathbb{E}[X_1^2] < \infty$, alors

$$\sqrt{n} \frac{\frac{1}{n} \sum_{i=1}^n X_i - \mathbb{E}[X_1]}{\sqrt{\text{Var}[X_1]}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1)$$

- Générer R réalisations $x_{1,r}, \dots, x_{n,r}$ de la suite X_1, \dots, X_n dans le cas d'une loi gamma $X_1 \sim \mathcal{G}(1, 1)$, ($r = 1, \dots, R$). On a donc $\mathbb{E}[X_1] = 1$ et $\text{Var}[X_1] = 1$. On prendra $n = 1000$ et $R = 500$.
- Calculer $\bar{x}_{n,r} = \frac{1}{n} \sum_{i=1}^n x_{i,r}$ pour chaque r .
- Tracer (1) un histogramme de $\sqrt{(n)}(\bar{x}_{n,r} - 1)$ sur le même graphique que la fonction de densité de la loi $\mathcal{N}(0, 1)$ et (2) la fonction de répartition empirique de cette suite sur le même graphique que la fonction de répartition de la loi $\mathcal{N}(0, 1)$. Essayer avec d'autres valeurs de n . Conclusion ?
- Procéder de même avec la loi binomiale $\mathcal{B}(10, 0.8)$.

Exercice 7. TCL généralisé

Si X_1, X_2, \dots, X_n est une suite de variables aléatoires i.i.d. vérifiant la condition de régularité $\mathbb{P}(|X_i| > x) \sim x^{-\alpha}$ avec $\alpha \in (0, 2)$ pour x suffisamment grand, alors $\exists C > 0$ et une suite réelle D_n t.q.

$$\frac{\sum_{i=1}^n X_i + D_n}{n^{1/\alpha} C} \xrightarrow{\mathcal{D}} \alpha\text{-stable}$$

- a) Générer R réalisations $x_{1,r}, \dots, x_{n,r}$ de la suite X_1, \dots, X_n dans le cas d'une loi Cauchy $X_1 \sim \text{Cauchy}(0, 1)$, ($r = 1, \dots, R$). Dans ce cas, on a la convergence du TCL généralisé avec $\alpha = 1$, $D_n = 0$ et $C = 1$. On prendra $n = 1000$ et $R = 5000$.
- b) Calculer $\bar{x}_{n,r} = \frac{1}{n} \sum_{i=1}^n x_{i,r}$ pour chaque r .
- c) Tracer (1) un histogramme de $\bar{x}_{n,r}$ sur le même graphique que la fonction de densité de la loi $\text{Cauchy}(0, 1)$ et (2) la fonction de répartition empirique de cette suite sur le même graphique que la fonction de répartition de la loi $\text{Cauchy}(0, 1)$.

Exercice 8. Lois des valeurs extremes

voir dans "notes-de-cours-2013.pdf" page 24/27

Exemples de convergence du maximum renormalisé et Exercice