Map 311 - Introduction aux probabilités et à la simulation aléatoire.
PC 7 - Année 2016

Vincent Bansaye - Lucas Gerin

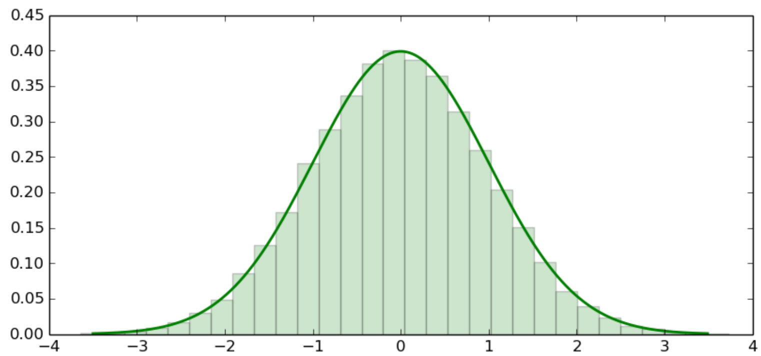
Convergence en loi, Théorème Central Limite

EXERCICE 1 -(Une somme aléatoire de variables aléatoires)
Soit $X_1, X_2, \ldots$ une suite de v.a. i.i.d. de fonction caractéristique $\Phi_X(t) = \mathbb{E}[e^{itX}]$, et $N \geq 1$ une variable aléatoire entière de fonction génératrice $g_N(z) = \mathbb{E}[z^N]$. On suppose que N est indépendante des X_k et que $\mathbb{E}[X_1] < \infty$, $\mathbb{E}[N] < \infty$. On pose

$$S = X_1 + \cdots + X_N,$$

calculer la fonction caractéristique de S et en déduire $\mathbb{E}[S]$ en fonction de $\mathbb{E}[X_1], \mathbb{E}[N]$.

EXERCICE 2 -(La "loi des douze uniformes")
Une méthode approchée (et très obsolète...) pour simuler une $\mathcal{N}(0,1)$ consiste à faire la somme de 12 uniformes indépendantes sur $[0,1]$, et soustraire 6. Pourquoi ça marche?



L'histogramme de 20000 *simulations de* $U_1 + \cdots + U_{12} - 6$*, et la densité de la* $\mathcal{N}(0,1)$.

EXERCICE 3 -(Intervalles de confiance exacts/asymptotiques)
Avant un référendum, on effectue une enquête pour estimer la proportion p de personnes votant *oui*. On interroge un échantillon représentatif de n personnes et on note F_n le nombre de réponses *oui* dans l'échantillon.

1. Quelle est la loi de F_n (en faisant des hypothèses raisonnables)? Rappeler l'énoncé du TCL dans ce cas et en déduire un intervalle de confiance I_n *asymptotique* à 95% pour p : $\lim_n \mathbb{P}(I_n \ni p) \geq 0.95$.
   (Si $Z \sim \mathcal{N}(0,1)$, $\mathbb{P}(-1.96 \leq Z \leq 1.96) = 0.95$.*)*

2. On rappelle l'inégalité de concentration (*vue en Amphi 4*) pour la loi binomiale :

$$\mathbb{P}\left(\quad |\mathrm{Binom}(n,p) - np| > r \right) \leq 2\exp\left(-2\frac{r^2}{n}\right).$$

   Utiliser cette inégalité pour obtenir un intervalle de confiance J_n *exact* à 95% pour p : pour tout n, $\mathbb{P}(J_n \ni p) \geq 0.95$.

3. Quel intervalle choisiriez-vous?

EXERCICE 4 -(Une somme positive)
Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires i.i.d., on suppose que $\mathbb{E}[X_1] = 0$ et $\mathbb{E}[X_1^2] < +\infty$. Démontrer que

$$\lim_{n\to+\infty} \mathbb{P}(X_1 + \cdots + X_n \geq 0) = 1/2.$$

EXERCICE 5 -(Réduction de variance dans une méthode de Monte Carlo). Soit g une fonction mesurable et bornée ($0 \leq g \leq 1$). On souhaite calculer de façon approchée $m = \int_0^1 g(x)dx$.
Soient X et Y des variables i.i.d. de loi uniforme sur $[0, 1]$. On pose

$$U = \mathbf{1}_{Y\leq g(X)}, \qquad V = g(X) \quad \text{et} \quad W = \frac{g(X) + g(1 - X)}{2}.$$

1. Calculer l'espérance et la variance de U, V et W. Comparer les variances.
2. Proposer trois méthodes de type Monte-Carlo pour estimer m.
3. On suppose maintenant que g est monotone. Démontrer que $\mathbb{E}(g(X)g(1 - X)) \leq m^2$. *(Indication : on pourra d'abord montrer que $(g(x) - g(y))(g(1 - x) - g(1 - y)) \leq 0$ pour tout x, y.)*
4. On considère une suite $(X_i)_{i\geq 1}$ de variables i.i.d. de loi uniforme sur $[0, 1]$. On considère les estimateurs
$$A_n = \frac{1}{2n}\sum_{i=1}^{2n} g(X_i), \qquad B_n = \frac{1}{2n}\sum_{i=1}^{n}(g(X_i) + g(1 - X_i)),$$
lequel choisiriez-vous pour approcher m?

EXERCICE 6 -(Somme normalisée)
Soit $(X_i)_{i\geq 1}$ une suite de v.a. i.i.d. d'espérance 0 et de carré intégrable. Quelle est la limite en loi de

$$\frac{\sum_{i=1}^n X_i}{\sqrt{\sum_{i=1}^n X_i^2}} \quad ?$$

EXERCICE 7 -(Pas de convergence en proba dans le TCL $\star$)
Soit $(X_n)_n$ une suite de variables aléatoires i.i.d. centrées de variance $\sigma^2 > 0$. Soit

$$Z_n = \frac{1}{\sigma\sqrt{n}}\sum_{j=1}^{n} X_j\,.$$

Par le théorème central limite, cette variable converge en loi vers la loi normale centrée réduite. L'objet de cet exercice est de montrer que la suite (Z_n) ne peut pas converger en probabilité.

1. En écrivant $Z_{2n} - Z_n$ comme une somme de variables indépendantes, démontrer que la suite $(Z_{2n} - Z_n)_{n\geq 1}$ converge en loi vers une loi normale dont on déterminera les paramètres.
2. On suppose maintenant que (Z_n) converge en probabilité vers une variables X. Déduire de la question précédente une contradiction.