Leçon 249 : Loi normale en probabilités et statistiques

Rappels de théorie

Une variable aléatoire réelle (v.a.r.) X suit une loi normale de moyenne m et variance σ^2 si elle admet la densité

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\,\mathrm{e}^{-(x-m)^2/(2\sigma^2)} \;.$$

On écrit alors $X \sim \mathcal{N}(m, \sigma^2)$. Attention, certains auteurs écrivent $X \sim \mathcal{N}(m, \sigma)$.

Si $X \sim \mathcal{N}(0, 1)$, on dit que X suit une loi normale centrée réduite, ou encore une loi normale standard.

Théorème 1 (Théorème central limite)**.** *Soit $(X_n)_{n\geqslant 1}$ une suite de v.a.r. indépendantes, identiquement distribuées, admettant une espérance μ et une variance σ^2. Pour $n \in \mathbb{N}^*$, soit*

$$S_n = \sum_{i=1}^{n} X_n \;, \qquad \widehat{S}_n = \frac{S_n - n\mu}{\sqrt{n\sigma^2}} \;.$$

Alors S_n converge en loi, lorsque $n \to \infty$, vers une v.a.r. de loi normale standard. En particulier, pour tout intervalle $[a, b] \subset \mathbb{R}$, on a

$$\lim_{n\to\infty} \mathbb{P}\bigl\{a \leqslant \widehat{S}_n \leqslant b\bigr\} = \int_a^b \frac{\mathrm{e}^{-x^2/2}}{\sqrt{2\pi}}\,\mathrm{d}x \;.$$

Exercices

Exercice 1 (Fonction caractéristique, voir aussi la leçon 232)**.** Soit X une v.a.r. de loi normale centrée réduite. Montrer que sa fonction caractéristique est donnée par

$$\mathbb{E}\bigl(\mathrm{e}^{\mathrm{i}\,tX}\bigr) = \mathrm{e}^{-t^2/2} \;.$$

En déduire la fonction caractéristique de $Y \sim \mathcal{N}(m, \sigma^2)$.

Exercice 2 (Somme de variables aléatoires normales)**.** Soient $X_1 \sim \mathcal{N}(m_1, \sigma_1^2)$ et $X_2 \sim \mathcal{N}(m_2, \sigma_2^2)$ des v.a.r. normales indépendantes.

1. Calculer $\mathbb{E}\bigl(\mathrm{e}^{\mathrm{i}\,tX_1}\bigr)$, $\mathbb{E}\bigl(\mathrm{e}^{\mathrm{i}\,tX_2}\bigr)$ et $\mathbb{E}\bigl(\mathrm{e}^{\mathrm{i}\,t(X_1+X_2)}\bigr)$.
2. En déduire la loi de $X_1 + X_2$.

Exercice 3 (Loi normale bivariée)**.** Soient X_1, X_2 deux v.a.r. t.q. $\mathbb{E}(\mathrm{e}^{\mathrm{i}(t_1X_1+t_2X_2)}) = \mathrm{e}^{-\langle t, Ct\rangle/2}$, où C est une matrice symétrique 2×2 et $\langle t, Ct\rangle = \sum_{i,j=1}^{2} C_{ij}t_it_j$.

1. Quelle est la loi de X_1 et de X_2 ?
2. Calculer $\mathbb{E}(X_1X_2)$. En déduire que X_1 et X_2 sont indépendantes si et seulement si elles sont non corrélées. (On admettra que si $\mathbb{E}(\mathrm{e}^{\mathrm{i}(t_1X_1+t_2X_2)})$ est un produit d'une fonction de t_1 et d'une fonction de t_2, alors X_1 et X_2 sont indépendantes).

Exercice 4 (Une preuve du théorème central limite ☠)**.** Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires indépendantes, identiquement distribuées, admettant une espérance μ et une variance σ^2.

1. Pour tout $n \in \mathbb{N}^*$, on pose
$$Y_n = \frac{X_n - \mu}{\sigma} .$$
Calculer $\mathbb{E}(Y_n)$, $\mathbb{E}(Y_n^2)$ et $\operatorname{Var}(Y_n)$.

2. Exprimer
$$\widehat{S}_n = \frac{S_n - n\mu}{\sqrt{n\sigma^2}} \qquad \text{où } S_n = \sum_{i=1}^{n} X_n$$
en fonction des Y_i. Calculer $\mathbb{E}(\widehat{S}_n)$ et $\operatorname{Var}(\widehat{S}_n)$.

3. Soit $\phi(t)$ la valeur commune des fonctions caractéristiques $\mathbb{E}(\mathrm{e}^{\mathrm{i}\,tY_i})$. Déterminer le développement limité de ϕ à l'ordre 2 en 0.

4. Exprimer $\mathbb{E}(\mathrm{e}^{\mathrm{i}\,t\widehat{S}_n})$ en fonction de $\phi(t)$.

5. Calculer le développement limité à l'ordre 2 de $\log \mathbb{E}(\mathrm{e}^{\mathrm{i}\,t\widehat{S}_n})$ en 0, et calculer sa limite lorsque $n \to \infty$.

6. Conclure en appliquant le théorème de Lévy (admis) : Soit $(\widehat{S}_n)_{n\geqslant 1}$ une suite de variables aléatoires réelles admettant les fonctions caractéristiques $(\phi_n)_{n\geqslant 1}$. On suppose que la suite des $\phi_n(t)$ converge simplement vers une fonction continue $\phi_*(t)$. Alors $\phi_*(t)$ est la fonction caractéristique d'une variable aléatoire réelle $\widehat{S}$, et $\widehat{S}_n$ converge en loi vers $\widehat{S}$.

Exercice 5. Parmi les pièces produites par une chaîne de montage, 10% sont défectueuses. Estimer la probabilité que parmi 400 pièces, plus de 50 pièces soient défectueuses.

Exercice 6 (Voir Chafaï et Zitt, Section 5.3, page 62)**.** On jette $n = 1000$ fois une pièce de monnaie équilibrée. Construire un intervalle symétrique I_α autour de 500 tel que la probabilité que le nombre de Pile obtenus n'appartienne pas à I_α soit inférieure à $\alpha = 0{,}05$.

Exercice 7 (Voir `http://exo7.emath.fr/ficpdf/fic00153.pdf`)**.** Sur un grand nombre de personnes on a constaté que la répartition du taux de cholestérol suit une loi normale avec les résultats suivants :

- 56% ont un taux inférieur à 165 cg ;
- 34% ont un taux compris entre 165 cg et 180 cg ;
- 10% ont un taux supérieur à 180 cg.

Quel est le nombre de personnes qu'il faut prévoir de soigner dans une population de 10 000 personnes, si le taux maximum toléré sans traitement est de 182 cg ?