Licence de Physique, FIP 2018-19

Mathématiques pour physiciens : TD n°7

Probabilités (distributions, et transformée de Fourier)

Guilhem Semerjian & Francesco Zamponi

Définitions

Dans ce TD, on notera $p_X(x)$ la densité de probabilité d'une variable aléatoire X, définie par la relation

$$\mathbb{P}(X \in [a,b]) = \int_a^b \mathrm{d}x \, p_X(x) \ ,$$

pour tout intervalle $[a,b]$.

Les moyennes sur les variables aléatoires seront écrites $\langle \bullet \rangle$.

La densité gaussienne centrée d'écart-type σ sera notée

$$\mathcal{G}_\sigma(x) = \frac{e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} \ .$$

La fonction caractéristique d'une variable aléatoire X est définie comme la transformée de Fourier de sa densité de probabilité :

$$G_X(t) = \langle e^{itX} \rangle = \int_{-\infty}^{\infty} \mathrm{d}x \, p_X(x) e^{itx} \ .$$

On a donc par transformée de Fourier inverse

$$p_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}t \, G_X(t) e^{-itx} \ .$$

On définit le $n^{\text{ème}}$ moment de X comme

$$M_n(X) = \langle X^n \rangle \ , \qquad G_X(t) = \sum_{n=0}^{\infty} M_n(X) \frac{(it)^n}{n!} \ ,$$

et le $n^{\text{ème}}$ cumulant de X, $M_n^c(X)$, par

$$\log G_X(t) = \sum_{n=1}^{\infty} M_n^c(X) \frac{(it)^n}{n!} \ .$$

1 Probabilité continue

1. Montrer que
$$p_X(x) = \langle \delta(x - X) \rangle \ . \tag{1}$$
2. Considérons deux variables aléatoires indépendantes X et Y, et définissons une nouvelle variable aléatoire par $Z = X + Y$. Montrer, en utilisant la fonction caractéristique, que
$$p_Z(z) = \int \mathrm{d}x \, p_X(x) p_Y(z - x) \ .$$
Montrer la même relation en utilisant l'Eq. (1).
3. Calculer explicitement $p_Z(z)$ dans les deux cas suivants :
$$\begin{array}{ll} \text{a)} & p_X(x) = \mathcal{G}_{\sigma_X}(x) \ , \qquad p_Y(y) = \mathcal{G}_{\sigma_Y}(y) \ . \\ \text{b)} & p_X(x) = \dfrac{1}{\pi} \dfrac{a}{a^2 + x^2} \ , \qquad p_Y(y) = \dfrac{1}{\pi} \dfrac{b}{b^2 + x^2} \ , \qquad a, b > 0 \ . \end{array}$$
4. Montrer, à partir de l'Eq. (1), que pour une variable $Y = f(X)$, avec $f'(x) > 0$, on a
$$p_Y(y) = p_X(f^{-1}(y)) \frac{1}{f'(f^{-1}(y))} \ .$$
5. Dans le cas où X et Y sont deux variables aléatoires indépendantes positives de densité $p_X(x) = \alpha e^{-\alpha x}$ et $p_Y(y) = \beta e^{-\beta y}$ respectivement, calculer les densités de probabilité de X^2, Y^3, $X + Y$ et Y/X.
6. On choisit au hasard X dans $[0, 1]$, puis Y au hasard dans $[0, X]$.
   (a) Déterminer la distribution jointe $p_{X,Y}(x, y)$.
   (b) Déterminer les distributions marginales $p_X(x)$ et $p_Y(y)$.
   (c) Calculer l'espérance de Y et le coefficient de corrélation ρ de (X, Y) qui est défini comme
   $$\rho(X, Y) = \frac{\langle (X - \langle X \rangle)(Y - \langle Y \rangle) \rangle}{\sigma(X)\sigma(Y)} \ . \tag{2}$$

2 Cumulants

1. Les moments d'une variable aléatoire ne sont pas toujours bien définis. On dit que X admet un moment d'ordre n si $\langle |X|^n \rangle < \infty$; donner une condition sur $p_X(x)$ pour que cela soit le cas. Que peut-on en conclure sur l'existence des autres moments de X ? et de ses cumulants ?
2. Comment s'expriment les quatre premiers cumulants en fonction des moments de X ?
3. Montrer que les cumulants d'une variable gaussienne $p_X(x) = \mathcal{G}_\sigma(x)$ sont nuls pour $n > 2$. Les cumulants pour $n > 2$ mesurent donc le degré de "non-gaussianité" d'une densité de probabilité.
4. Considérer une densité paire, $p_X(x) = p_X(-x)$. Montrer que tous les cumulants impairs sont nuls.

5. Dans le même cas, montrer que
$$M_4^c(X) = \langle X^4 \rangle - 3\langle X^2 \rangle^2 \ .$$
En déduire que pour une variable gaussienne centrée, $\langle X^4 \rangle = 3\langle X^2 \rangle^2$.
6. Montrer plus généralement que les moments d'ordre pair d'une variable gaussienne centrée sont
$$\langle X^{2n} \rangle = \frac{(2n)!}{2^n n!} \langle X^2 \rangle^n \ .$$
Ceci est un exemple du "théoreme de Wick-Isserlis" qui a des applications très importantes en physique.

3 Somme de variables aléatoires et théorème limite central

Le but de cet exercice est d'étudier le comportement limite de la somme d'un grand nombre de variables aléatoires.

1. On considère $S_N = X_1 + \cdots + X_N$, où les X_i sont N variables aléatoires indépendantes distribuées avec la même loi que X. Que vaut la fonction caractéristique de S_N en fonction de celle de X ? Que valent les cumulants de S_N en fonction de ceux de X ?
2. *Théorème limite central (TLC)* - On définit $s_N = \frac{S_N - N\langle X \rangle}{\sqrt{N}}$. Calculer la limite de la fonction caractéristique de s_N quand N tend vers l'infini. En déduire que la distribution de s_N converge vers une distribution gaussienne $\mathcal{G}_\sigma(x)$ de variance $\sigma^2 = M_2^c(X) = \langle X^2 \rangle$ et de moyenne nulle.
3. Nous supposerons à présent la loi de X symétrique, $p_X(x) = p_X(-x)$, de sorte que le $n^{\text{ème}}$ moment $M_n(X) = \langle X^n \rangle$ est nul pour n impair, et nous supposerons aussi que le quatrième moment de X existe. A l'aide d'un développement de la fonction caractéristique de s_N à l'ordre $1/N$ et d'une transformée de Fourier inverse, montrer que
$$p_{s_N}(x) = \mathcal{G}_\sigma(x) \left[1 + \frac{1}{N} \frac{M_4^c(X)}{24\,\sigma^4} \left(3 - \frac{6x^2}{\sigma^2} + \frac{x^4}{\sigma^4} \right) + o(1/N) \right] \ .$$
En déduire que $p_{s_N}(x)$ peut être approximée par une loi gaussienne d'écart-type σ dans un intervalle de largeur d'ordre N^α et donner la valeur de α.
4. Généraliser vos réponses aux deux premières questions de cet exercice quand les X_i sont des variables aléatoires indépendants, mais pas forcément identiquement distribuées.

4 Fonction de répartition

On définit la fonction de répartition d'une variable aléatoire X par

$$F_X(x) = \mathbb{P}(X \leq x) = \langle H(x-X) \rangle = \int_{-\infty}^{x} \mathrm{d}y \, p_X(y) \ ,$$

où H est la fonction de Heaviside. On a donc $p_X(x) = F'_X(x)$.

1. On considère une variable aléatoire X qui représente le résultat du jet d'un dé à 6 faces. Tracer la fonction de répartition $F_X(x)$ pour $x \in \mathbb{R}$. En déduire la densité $p_X(x)$.
2. Soient deux variables aléatoires indépendantes X et Y. Calculer la fonction de répartition de $Z = \max(X, Y)$.
3. Calculer la densité de $Z = \max(X, Y)$ dans le cas où X et Y sont des variables aléatoires indépendantes et positives de densité $p_X(x) = \alpha e^{-\alpha x}$ et $p_Y(y) = \beta e^{-\beta y}$ respectivement.
4. On considère maintenant N variables aléatoires $X_1, \cdots, X_N$, indépendantes et positives de densité $\alpha e^{-\alpha x}$.

   Calculer la fonction de répartition de $Z = \max(X_1, \ldots, X_N)$.

   Montrer que si z reste fini quand $N \to \infty$, on a $F_Z(z) \to 0$.

   On définit une nouvelle variable aléatoire W selon $Z = (\ln N - \ln W)/\alpha$. Montrer que dans la limite $N \to \infty$ on obtient

   $$F_W(w) = 1 - e^{-w} \ , \qquad p_W(w) = e^{-w} \ .$$

   On en déduit que le maximum de $X_1, \cdots, X_N$ est ici d'ordre $\ln N$ quand $N \to \infty$.