Licence de Physique, FIP

Mathématiques pour physiciens : TD n°6

Probabilités (distributions, et transformée de Fourier)

Emmanuel Baudin & Francesco Zamponi

Definitions

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

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

pour tout intervalle $[a,b]$.

Aussi, on denotera

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

une densité gaussienne centrée d'écart type σ.

La fonction génératrice est définie comme étant la transformée de Fourier de la densité de probabilité :

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

La densité de probabilité est la transformée de Fourier inverse de la fonction génératrice,

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

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. Considerer deux variables aléatoires indépendantes X et Y, et $Z = X + Y$. Montrer, en utilisant la fonction génératrice, 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 cas suivants :
$$p_X(x) = \mathcal{G}_{\sigma_X}(x) \ , \qquad p_Y(y) = \mathcal{G}_{\sigma_Y}(y) \ .$$
$$p_X(x) = p_Y(x) = \frac{1}{\pi} \frac{1}{1 + x^2} \ .$$
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 des variables aléatoires positives de densité $\alpha e^{-\alpha x}$ et $\beta e^{-\beta x}$ 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 cumulants ne sont pas toujours bien définis. Donner la condition sur $P_X(x)$ pour que les premiers n cumulants soient bien définis.
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. Généraliser le calcul pour obtenir (on pourra utiliser une intégration par parties)

$$\langle X^{2n} \rangle = \frac{(2n)!}{2^n n!} \langle X^2 \rangle^n \ .$$

Ceci est un exemple du "théoreme de Wick-Isserlis" qui a des applications très importantes en physique.

3 Somme de variables aléatoires et théorème de la limite centrale

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 génératrice de S_N en fonction de celle de X ? Que valent les cumulants de S_N en fonction de ceux de X ?
2. *Théorème de la limite centrale (TLC)* - On définit $s_N = S_N/\sqrt{N}$. Calculer la fonction génératrice de s_N quand N est très grand. En déduire que la distribution de s_N converge vers une distribution gaussienne $\mathcal{G}_\sigma(x - x_0)$ de variance $\sigma^2 = M_2^c(X) = \langle X^2 \rangle$ et de moyenne $x_0 = \sqrt{N} M_1^c(X) = \sqrt{N}\langle X \rangle$.
3. Par simplicité, nous supposerons à présent la loi de X symétrique, $p_X(x) = p_X(-x)$, de sorte que le $n^{\text{ème}}$ moment $M_n(X) = \langle X^n \rangle$ soit nul pour n impair. À l'aide d'une expansion de la fonction génératrice de s_N à l'ordre 4 en t, monter que

$$p_{s_N}(x) = \mathcal{G}_\sigma(x) \left[1 + \frac{M_4^c(X)}{24 N M_2^c(X)^2} \left(3 - \frac{6x^2}{M_2^c(X)} + \frac{x^4}{M_2^c(X)^2} \right) + o(1/N) \right]$$

   En déduire que $p_{s_N}(x)$ peut être approximée par une loi gaussienne de largeur σ dans un intervalle de largeur d'ordre N^α et donner la valeur de α.
4. Est-il nécessaire que tous les X_i aient la même distribution ? Si possible, généraliser le TLC au cas où chaque X_i a une distribution différente.

4 Fonction de repartition

On definit la fonction de repartition

$$F_X(x) = \mathbb{P}(X < x) = \langle \theta(X - x) \rangle = \int_{-\infty}^{x} \mathrm{d}x \, p_X(x) \ .$$

On a donc $p_X(x) = F'_X(x)$.

1. On considère une variable aléatoire X qui représente le résultat du tir 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. Soit deux variables aléatoires indépendantes X et Y. Calculer la fonction de repartition 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 $\alpha e^{-\alpha x}$ et $\beta e^{-\beta x}$ respectivement.
4. On considère maintenant N variables aléatoires indépendantes et positives de densité $\alpha e^{-\alpha x}$ qu'on denote $X_1, \cdots, X_N$ (vous pouvez penser par exemple qu'on a choisi N personnes au hasard et que X_i est la taille de la personne i). Calculer la fonction de répartition de $Z = \max_{i=1,\cdots,N}\{X_i\}$. Montrer que si z reste fini quand $N \to \infty$, on a $F_Z(z) \to 0$. Montrer ensuite que dans la limite $N \to \infty$, si on choisit $z = (\ln N - \ln w)/\alpha$, avec w fini, 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 d'ordre $\ln N$ quand $N \to \infty$.