

Tutorial n°1 – Probability Basics

stoehr@ceremade.dauphine.fr

Exercise 1 (*Probability or Statistics?*). Let $X_1, \dots, X_n$ be *i.i.d.* random variables and $x_1, \dots, x_n$ be realisations or observations of the later random variables. Which of the following quantities are random?

1. $\max\{x_1, \dots, x_n\}$,
2. the sample size n ,
3. $\frac{1}{n} \sum_{i=1}^n X_i$,
4. $\min\{X_1, \dots, X_n\}$,
5. $\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}_n)^2$.

Exercise 2 (*True or False?*). Let X and Y be integrable random variables. Which of the following statements are correct? Justify your answer with a brief proof or a counterexample.

1. If X is symmetric with respect to 0, then $\mathbb{E}[X] = 0$.
2. $\mathbb{E}\left[\frac{1}{X}\right] = \frac{1}{\mathbb{E}[X]}$.
3. $\mathbb{E}[X]^2 \leq \mathbb{E}[X^2]$.
4. $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.
5. $\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y]$.

Exercise 3. Let X be a real random variable with density

$$f_X(x) = \frac{5}{x^2} \mathbb{1}_{\{x > 5\}}$$

Compute the following quantities:

1. $\mathbb{P}[X > 20]$,
2. $F_X(t)$, for all $t \in \mathbb{R}$,
3. $\mathbb{E}[X]$.

Exercise 4. Let X be a random variable following the uniform distribution on $[-\pi/2, \pi/2]$. Determine the distribution of $\tan(X)$.

Exercise 5 (*Multinomial distribution*). A population is divided into K groups. We denote $p_1, \dots, p_K$ the proportions of individuals in each group, with $p_1, \dots, p_K \in [0, 1]$ and $\sum_{i=1}^K p_i = 1$. We draw in this population n individuals with replacement. Let denote N_i the number of individuals belonging to group i , $i = 1, \dots, K$, among the n individuals drawn.

1. Give, with a justification, the distribution of $(N_1, \dots, N_K)$.
2. Give the marginal distribution of N_i , $i = 1, \dots, K$.
3. Give the R command to use to run this experiment.

Exercise 6. We consider a system made of two different machines working in series, that is the system works as long as both machines work. Let denote X_1 and X_2 the lifetime of the two machines and Z the lifetime of the system. We assume that the random variables X_1 and X_2 are independent

and follow an exponential distribution with respective parameters λ_1 and λ_2 .

1. Compute the probability that the system breaks down after time $t \geq 0$ and deduce the distribution of Z .
2. Compute the probability that the break down is due to a failure of machine 1.
3. Let Y be a random variable such that $Y = 1$ if the failure is due to machine 1 and $Y = 0$ otherwise.
 - (a) Compute $\mathbb{P}[Z > t, Y = 1]$ for all $t \geq 0$.
 - (b) Deduce that Z and Y are independent.

◇ To do at Home ◇

Exercise 7 (*Characteristic function*).

1. Let Y be a real random variable and Z a random variable, independent of Y , such that

$$\mathbb{P}[Z = 1] = \mathbb{P}[Z = -1] = \frac{1}{2}.$$

- (a) Show that the law of $X = ZY$ is symmetric.
 - (b) Compute the characteristic function of X according to the characteristic function of Y .
2. Let X be a random variable following the standard Laplace law:

$$f_X(x) = 0.5 \exp(-|x|).$$

Show that, for every real t , $\Phi_X(t) = (1 + t^2)^{-1}$.

Exercise 8 (*Gamma distribution*). Given $\alpha > 0$ and $\lambda > 0$, the gamma distribution $\gamma(\alpha, \lambda)$ is defined by the density with respect to the Lebesgue measure

$$f(x; \alpha, \lambda) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \exp(-\lambda x) x^{\alpha-1} \mathbf{1}_{\{x \geq 0\}}, \quad \text{where} \quad \Gamma(\alpha) = \int_0^{+\infty} e^{-t} t^{\alpha-1} dt.$$

1. Check that $f(\cdot; \alpha, \lambda)$ is a probability density function.
2. Compute the expectation of the gamma distribution $\gamma(\alpha, \lambda)$.
3. Let $X_1, \dots, X_n$ be *i.i.d.* random variables with distribution $\mathcal{E}(\lambda)$. Show that $X_1 + \dots + X_n$ is distributed according to the gamma distribution $\gamma(n, \lambda)$.
4. Let X and Y be independent random variables with distribution $\gamma(\alpha, \lambda)$.
 - (a) Show that $X + Y$ and $X/(X+Y)$ are independent random variables.
 - (b) Give the distributions of $X + Y$ and $X/(X+Y)$.