Week 5: classical laws and independence

Submission of solutions. Feedback can be given on Exercise 1 and any other exercise from the Training exercises. If you want to hand in, do it so by Monday 23/10/2023 17:00 (online) following the instructions on the course website

https://metaphor.ethz.ch/x/2023/hs/401-3601-00L/

Please pay attention to the quality, the precision and the presentation of your mathematical writing.

***

1 Exercise covered during the exercise class

The following exercise will be covered during the exercise class.

Exercise 1. Let (X, Y) be a random variable with values in $\mathbb{R}^2$ whose joint distribution has the density $f_{(X,Y)}(x, y) = \frac{1}{4}(1 + xy)\mathbb{1}_{-1 \leq x, y \leq 1}$.

(1) Find the law of X.

(2) Compute $\mathbb{E}[X\mathbb{1}_{X<1/2}]$.

(3) Compute $\mathbb{E}\left[\frac{1}{X}\right]$.

(4) Compute $\mathbb{E}[XY]$.

(5) Compute $\mathbb{P}(X \leq Y)$. Is the result surprising?

(6) Are the random variables X and Y independent? Justify your answer.

Solution:

(1) Since (X, Y) has a density, we know that X has a density, and its density is obtained by integrating $f_{(X,Y)}$ with respect to the second variable. Thus, for $-1 \leq x \leq 1$:

$$f_X(x) = \int_{-\infty}^{\infty} \mathrm{d}y\, f_{(X,Y)}(x, y) = \int_{-1}^{1} \mathrm{d}y \frac{1}{4}(1 + xy)\mathbb{1}_{-1 \leq x \leq 1} = \frac{1}{2}.$$

Therefore, X follows the uniform distribution on $[-1, 1]$.

(2) According to the transfer theorem,

$$\mathbb{E}[X\mathbb{1}_{X<1/2}] = \int_{\mathbb{R}} x\mathbb{1}_{x<1/2} f_X(x)\mathrm{d}x = \int_{-1}^{1/2} \frac{x}{2}\mathrm{d}x = -\frac{3}{16}.$$

(3) It's a trap: $\frac{1}{X}$ is not integrable, because

$$\mathbb{E}\left[\frac{1}{|X|}\right] = \int_{-1}^{1} \frac{1}{|x|}\mathrm{d}x = \infty,$$

so the expression $\mathbb{E}\left[\frac{1}{X}\right]$ doesn't make sense.

(4) According to the transfer theorem,

$$\mathbb{E}[XY] = \int_{[-1,1]^2} \mathrm{d}x\mathrm{d}y \frac{1}{4}(xy + x^2y^2) = \frac{1}{4} \cdot \left(\int_{-1}^{1} x^2 \mathrm{d}x\right)^2 = \frac{1}{9}.$$

(5) According to the transfer theorem,

$$\mathbb{P}(X \leq Y) = \mathbb{E}[\mathbb{1}_{X \leq Y}] = \int_{[-1,1]^2} \mathrm{d}x\mathrm{d}y \mathbb{1}_{x \leq y} \frac{1}{4}(1 + xy) = \int_{-1}^{1} \mathrm{d}x \int_{x}^{1} \mathrm{d}y \frac{1}{4}(1 + xy).$$

so

$$\mathbb{P}(X \leq Y) = \int_{-1}^{1} dx \left(\frac{1}{4} - \frac{x}{8} - \frac{x^3}{8}\right) = \frac{1}{2}.$$

This is not surprising because

$$1 = \mathbb{P}(X < Y) + \mathbb{P}(X = Y) + \mathbb{P}(X > Y),$$

and since (X, Y) and (Y, X) have the same density due to symmetry, we have $\mathbb{P}(X < Y) = \mathbb{P}(X > Y)$, and because (X, Y) has a density, $\mathbb{P}(X = Y) = 0$.

(6) Intuitively, X and Y are not independent because $f_{(X,Y)}$ cannot be expressed as a function of x times a function of y. Formally, if X and Y were independent, we would have $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y] = 0$, which is not the case according to question 4.

□

2 Training exercises

Exercise 2. Let U be a uniform random variable on $[-1, 1]$. Compute $\mathbb{E}[e^U]$.

Solution:

The function e^x is positive, so we can apply the transfer theorem:

$$\mathbb{E}\left[e^U\right] = \int_{-1}^{1} e^x \frac{\mathrm{d}x}{2} = \frac{e - e^{-1}}{2}.$$

□

Exercise 3. Let X be a real random variable that follows an exponential distribution with parameter 1. Let $\lambda > 0$. Show that λX follows an exponential distribution with parameter $1/\lambda$.

Solution:

We use the dummy function method. Let $f : \mathbb{R} \to \mathbb{R}_+$ be measurable. Then by the transfer theorem

$$\mathbb{E}[f(\lambda X)] = \int_0^\infty f(\lambda x) e^{-x} \mathrm{d}x.$$

By the change of variables $\lambda x = u$ we get

$$\mathbb{E}[f(\lambda X)] = \int_0^\infty f(u) e^{-u/\lambda}/\lambda \mathrm{d}u.$$

This shows that λX has density $e^{-x/\lambda}/\lambda \mathbb{1}_{x \geq 0}$ on $\mathbb{R}_+$, which is the density of an exponential random variable of parameter $1/\lambda$. □

Exercise 4. Let Z be a real random variable with density $\frac{1}{\pi} \cdot \frac{1}{1+x^2}$ (it is a so-called Cauchy random variable). For which values of $\alpha \in \mathbb{Z}$ is the random variable Z^α integrable?

Solution:

By definition, Z^α is integrable if and only if $|Z|^\alpha$ is integrable, which by the transfer theorem is equivalent to:

$$\int_{-\infty}^\infty \frac{|x|^\alpha}{1+x^2} dx = 2 \int_0^\infty \frac{x^\alpha}{1+x^2} dx < \infty.$$

Set $f(x) = x^\alpha/(1+x^2)$, which is continuous on $\mathbb{R}_+^*$.

Let's analyze this for different values of α:

Behavior at $+\infty$. As $f(x) \sim \frac{1}{x^{2-\alpha}}$ when $x \to \infty$, f is integrable at $+\infty$ if and only if $\alpha < 1$.

Behavior at 0. Since $f(x) \sim x^\alpha$ as $x \to 0$, f is integrable at 0 if and only if $\alpha > -1$.

In conclusion, $|Z|^\alpha$ is integrable if and only if $\alpha = 0$.

□

Exercise 5. Let X and Y be two independent random variables, where X follows an exponential distribution with parameter $\lambda > 0$, and Y follows a geometric distribution with parameter $p \in (0,1)$. Compute $\mathbb{P}(X > Y)$.

Solution:

We can apply the law of total probability using the complete system of events $Y = k : k \geq 1$:

$$\mathbb{P}(X > Y) = \sum_{k \geq 1} \mathbb{P}(X > Y, Y = k).$$

Now, we have the equality of events $\{X > Y, Y = k\} = \{X > k, Y = k\}$. So, using the independence of X and Y:

$$\mathbb{P}(X > Y) = \sum_{k \geq 1} \mathbb{P}(X > k, Y = k) = \sum_{k \geq 1} \mathbb{P}(X > k) \mathbb{P}(Y = k) = \sum_{k \geq 1} e^{-\lambda k} p(1-p)^{k-1}$$

Therefore

$$\mathbb{P}(X > Y) = pe^{-\lambda}\sum_{k\geq 1}\left(e^{-\lambda}(1-p)\right)^{k-1} = \frac{pe^{-\lambda}}{1-e^{-\lambda}(1-p)} = \frac{p}{e^{\lambda}+p-1}$$

□

Exercise 6. Let X and Y be two independent real variables. Let $F : \mathbb{R}\times\mathbb{R} \to \mathbb{R}_+$ be a measurable function. Show that $\mathbb{E}[F(X,Y)] = \mathbb{E}[g(Y)]$, where $g : \mathbb{R} \to \mathbb{R}$ is the function defined by $g(y) = \mathbb{E}[F(X,y)]$ for $y \in \mathbb{R}$.

Solution:

By independence of X and Y we have $\mathbb{P}_{(X,Y)}(dxdy) = \mathbb{P}_X(dx)\otimes\mathbb{P}_Y(dy)$, so by Fubini-Tonnelli's theorem:

$$\begin{aligned}\mathbb{E}[F(X,Y)] &= \int_{\mathbb{R}\times R} F(x,y)\mathbb{P}_{(X,Y)}(dxdy) = \int_{\mathbb{R}\times R} F(x,y)\mathbb{P}_X(dx)\otimes\mathbb{P}_Y(dy) \\ &= \int_{\mathbb{R}} \mathbb{P}_Y(dy)\left(\int_{\mathbb{R}} F(x,y)\mathbb{P}_X(dx)\right) = \int_{\mathbb{R}} \mathbb{P}_Y(dy)g(y) = \mathbb{E}[g(Y)].\end{aligned}$$

□

3 More involved exercises (optional, will not be covered in the exercise class)

Exercise 7. Let X be an exponential random variable with a parameter of 1, and $a > 0$. Does the random variable $\min(X,a)$ have a density?

Solution:

Let's compute the cumulative distribution function of $Z = \min(X,a)$. For $u \geq a$, we have $\mathbb{P}(Z \leq u) = 1$. For $u < 0$, we have $\mathbb{P}(Z < 0) = 0$. For $0 \leq u < a$, we have $\mathbb{P}(Z \leq u) = \mathbb{P}(X \leq u \text{ et } u < a) = \mathbb{P}(X \leq u) = 1 - e^{-\lambda u}$. The cumulative distribution function of Z is not continuous at a, so $\min(X,a)$ is not a random variable with density. □

Exercise 8. Let T be an exponential random variable and U an independent uniform random variable on $[0,1]$. Set $X = \sqrt{T}\cos(2\pi U)$ and $Y = \sqrt{T}\sin(2\pi U)$. Find the law of (X,Y).

Solution:

We apply the "dummy function method": let $f,g : \mathbb{R} \to \mathbb{R}_+$ be measurable functions. We compute $\mathbb{E}[f(X)g(Y)]$:

$$\begin{aligned}\mathbb{E}[f(X)g(Y)] &= \mathbb{E}\left[f(\sqrt{T}\cos(2\pi U))g(\sqrt{T}\sin(2\pi U))\right] \\ &= \int_0^{\infty} dte^{-t}\int_0^1 du f(\sqrt{t}\cos(2\pi u))g(\sqrt{t}\sin(2\pi u)) \\ &= \int_0^{\infty} dr2re^{-r^2}\int_0^1 du f(r\cos(2\pi u))g(r\sin(2\pi u))\end{aligned}$$

by using the change of variables $t = r^2$. We then use the change of variables $r\cos(2\pi u) = x$ and

$r\sin(2\pi u) = y$. We have

$$\begin{vmatrix} \cos(2\pi u) & -2\pi r\sin(2\pi u) \\ \sin(2\pi u) & 2\pi r\cos(2\pi u) \end{vmatrix} = 2\pi r$$

so that $2\pi r dr du = dxxy$. Hence

$$\mathbb{E}[f(X)g(Y)] = \frac{1}{\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}x\mathrm{d}y e^{-x^2-y^2} f(x)g(y).$$

By taking $f \equiv g \equiv 1$, we get $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}x\mathrm{d}y e^{-x^2-y^2} = \pi$, so that

$$\mathbb{E}[f(X)g(Y)] = \left(\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} f(x)e^{-x^2}\mathrm{d}x\right)\left(\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} g(y)e^{-y^2}\mathrm{d}y\right).$$

We conclude that X and Y are independent standard Gaussian random variables. □

Exercise 9. Let (X_n) be a sequence of independent real random variables.

(1) Show that the radius of convergence R of the power series $\sum_{n\geq 0} X_n z^n$ is almost surely constant.

(2) Now assume that the random variables $(X_n)_{n\geq 0}$ have the same law. Show that if $\mathbb{E}[\ln(|X_1|)^+] = \infty$, then $R = 0$ almost surely, and if $\mathbb{E}[\ln(|X_1|)^+] < \infty$, then $R \geq 1$ almost surely (here $x^+ = \max(x, 0)$ represents the positive part of a real number x).

Solution:

a) The radious of convergence R is given by the formula

$$R = \frac{1}{\limsup_{n\to\infty} |X_n|^{1/n}}.$$

But the random variable $\limsup_{n\to\infty} |X_n|^{1/n}$ is measurable with respect to the tail σ-algebra of $(X_n)_{n\geq 1}$, it is therefore almost surely constant by an Exercise 4 (1) of Exercise Sheet 4.

b) Write

$$|X_n|^{1/n} = \exp\left(\frac{\ln(|X_n|)^+}{n}\right)\exp\left(-\frac{\ln(|X_n|)^-}{n}\right).$$

If $\mathbb{E}[\ln(|X_1|)^+] < \infty$, then by the previous exercise $\limsup_{n\to\infty} \ln(|X_n|)^+/n = 0$ so that we have $\ln(|X_n|)^+/n \to 0$. Hence $R \geq 1$ since $\exp\left(-\frac{\ln(|X_n|)^-}{n}\right) \leq 1$.

If $\mathbb{E}[\ln(|X_1|)^+] = \infty$, then by a result seen in the lecture, $\limsup_{n\to\infty} \ln(|X_n|)^+/n = \infty$. This implies that almost surely $\limsup_{n\to\infty} \ln(|X_n|)/n = \infty$ so that $R = 0$ almost surely.

□

Exercise 10. Let $(X_n)_{n\geq 1}$ be a sequence of i.i.d. random variables with law given by

$$\mathbb{P}(X_n = 1) = \mathbb{P}(X_n = -1) = 1/2, \qquad n = 0, 1, \ldots.$$

Show that with probability 1, there is no point z_0 on the unit cercle such that the power series $F(z) = \sum_{n\geq 0} X_n z^n$ can be extended in an open ball around z_0 into a function which can be expanded in a power series around z_0.

Solution:

We say that a complex-valued function defined on an open set $U \subset \mathcal{C}$ is analytic on U if it can be expanded in a power series at each point of U. Recall that a power series is analytic in every point of its open disk of convergence.

Set $\mathbb{S} = \{z \in \mathbb{C}; |z| = 1\}$, $D = \{z \in \mathbb{C}; |z| < 1\}$ and for $\zeta \in D, r > 0$ write $D_\zeta(r) = \{z \in \mathbb{C}; |z - \zeta| < r\}$. Finally set

$$\mathcal{A}_F = \{z \in \mathbb{S}; F \text{ can be extended around } z \text{ into a function expandable in a power series around } z\}.$$

We reason by contradiction and assume that $\mathbb{P}(\mathcal{A}_F \neq \emptyset) > 0$. By a density argument, we start by showing that it is enough to show an almost sure property for **one** point and not all points of $\mathbb{S}$. To this end, let $(q_n)_{n\geq 1}$ be a dense sequence in $\mathbb{S}$. By the first paragraphe, almost surely $\mathcal{A}_F$ is open, so that

$$\{\omega; \mathcal{A}_F \neq \emptyset\} \subset \{\omega; \exists q_n \text{ tq } q_n \in \mathcal{A}_F\}.$$

Therefore

$$\mathbb{P}(\exists q_n \text{ tq } q_n \in \mathcal{A}_F) > 0.$$

But

$$\mathbb{P}(\exists q_n \text{ tq } q_n \in \mathcal{A}_F\}) \leq \sum_{n\geq 1} \mathbb{P}(q_n \in \mathcal{A}_F)$$

It follows that there exists $n \geq 1$ such that$\mathbb{P}(q_n \in \mathcal{A}_F) > 0$. To simplify, set $q_n = q$.

But if F can be extended in a function expandable in a power series around q, one can find a sequence of points $r_n \in \mathcal{C}$ such that $|r_n| < 1$ and q belongs to open disk of convergence of the expansion around r_n. By the same reasoning as in the previous paragraph, we get the existence of $\zeta \in D$ and $r > 0$ such that $D_\zeta(r) \not\subset D$ and:

$$\mathbb{P}(F \text{ extends to an analytic function } D \cup D_\zeta(r)) > 0.$$

To simplify, set $\mathcal{A} = \{\omega; F \text{ extends to an analytic function on } D \cup D_\zeta(r)\}$.

Let us first show that $\mathbb{P}(\mathcal{A}) = 1$.To this end, for $|u| < 1 - |\zeta|$ write:

$$F(\zeta + u) = \sum_{n=0}^{\infty} X_n (\zeta + u)^n = \sum_{m=0}^{\infty} u^m \left(\sum_{n=m}^{\infty} \binom{n}{m} X_n \zeta^{n-m} \right).$$

To simplify, let $a_m = \sum_{n=m}^{\infty} \binom{n}{m} X_n \zeta^{n-m}$ be the coefficients of the expansion of F around ζ. The function F is analytic on $D_\zeta(r)$ if the radius of convergence of this power series is at least r. It follows that

$\{F$ is analytic on $D_\zeta(r)\}$ belongs to the tail σ-algebra of $(X_n)_{n\geq 1}$, implying that $\mathbb{P}(\mathcal{A}) = 0$ or 1 be the Kolmogorov -1 law. Since$\mathbb{P}(\mathcal{A}) > 0$, we must have $\mathbb{P}(\mathcal{A}) = 1$.

Now, by construction, the arc $D_\zeta(r) \cap \mathbb{S}$ is non empty. We can therefore fix an integer $k \geq 1$ sufficiently large so that this arc has length at least $2\pi/k$. Then set

$$Y_n(\omega) = \begin{cases} X_n(\omega) & \text{if } n \not\equiv 0 \mod k \\ -X_n(\omega) & \text{if } n \equiv 0 \mod k \end{cases}$$

and introduce

$$G(z) = \sum_{n=0}^{\infty} Y_n z^n.$$

Since the two sequences $(Y_n)_{n\geq 1}$ and $(X_n)_{n\geq 1}$ have the same distribution, we have

$$\mathbb{P}(G \text{ extends to an analytic function on} D \cup D_\zeta(r)) = 1.$$

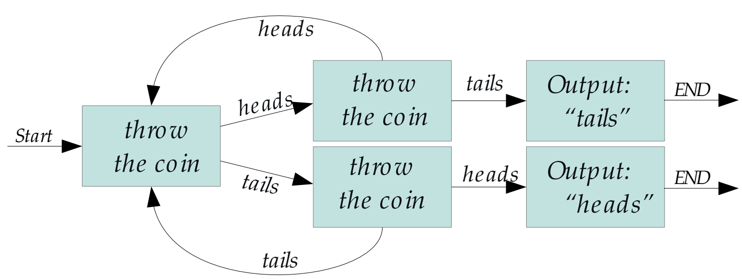
But

$$F(z) - G(z) = 2\sum_{m=0}^{\infty} X_{mk} z^{mk}.$$

By replacing z with $ze^{2\pi i/k}$, this expression does not change. Therefore, by setting $D_\zeta^{(l)}(r) = \{ze^{2\pi i l/k}; z \in D_\zeta^{(l)}\}$ for every $l \geq 1$, it follows that $F(z) - G(z)$ can almost surely be extended into an analytic function on $\{|z| < 1 + \epsilon\}$ for a certain $\epsilon > 0$ (here we use the fact that a finite union of events with probability 1 has probability 1). This is a contradiction, because the radius of convergence of $F - G$ is almost surely 1. □

4 Fun exercise (optional, will not be covered in the exercise class)

Exercise 11. We have a biased coin that comes up heads with a probability of p, and we want to use it to generate a fair coin toss. John von Neumann came up with the following algorithm:



Show that this works and compute the average number of times the coin is tossed.

Solution:

We denote by $T \in \{2,4,6,\ldots\}$ the random variable representing the number of tosses required for the algorithm to terminate, and $R \in \{\text{heads}, \text{tails}\}$ the outcome.

Observe that the algorithm comes back to the beginning when one gets the same result twice in a row.

We first show that the algorithm terminates almost surely. Let $(X_i)_{i\geq 1}$ be a sequence of i.i.d. random variables that are independent and follow a Bernoulli distribution with parameter p (representing a "heads" result on the i-th toss if $X_i = 1$). First, let's compute $A = \mathbb{P}(X_1 = X_2)$ and $B = \mathbb{P}(X_1 \neq X_2)$. We have

$$A = \mathbb{P}(X_1 = 1, X_1 = 1) + \mathbb{P}(X_1 = 0, X_2 = 0) = p^2 + (1-p)^2, \qquad B = 1 - A = 2p(1-p).$$

Then

$$\mathbb{P}(T = 2k) = \mathbb{P}\Big(X_1 = X_2, X_3 = X_4, \ldots, X_{2k-3} = X_{2k-2}, X_{2k-1} \neq X_{2k}\Big).$$

By independence it follows that

$$\mathbb{P}(T = 2k) = A^{k-1}B = (p^2 + (1-p)^2)^{k-1}\, 2p(1-p).$$

Since $\sum_{k\geq 1} \mathbb{P}(T = 2k) = 1$, we indeed have $\mathbb{P}(T < \infty) = 1$.

Now let us check that we get a fair coin toss in the end. By the formula of total probability,

$$\mathbb{P}(R = \text{ heads }) = \sum_{k\geq 1} \mathbb{P}(T = 2k, X_{2k} = 1) = \sum_{k\geq 1} A^{k-1}\mathbb{P}(X_{2k-1} = 0)\mathbb{P}(X_{2k} = 1) = \frac{p(1-p)}{1-A}.$$

Thus $\mathbb{P}(R = \text{ heads }) = 1/2$. Since $\mathbb{P}(R = \text{ heads }) + \mathbb{P}(R = \text{ tails }) = 1$, the result follows.

Let us finally compute $\mathbb{E}[T]$. Let Y be the random variable such that

$$\mathbb{P}(Y = k) = (p^2 + (1-p)^2)^{k-1}\, 2p(1-p)$$

for $k \geq 1$. We recognize a geometric random variable with parameter $2p(1-p)$, so $\mathbb{E}[Y] = \frac{1}{2p(1-p)}$. Thus,

$$\mathbb{E}[T] = \sum_{k\geq 1} 2k\mathbb{P}(T = 2k) = 2\mathbb{E}[Y] = \frac{1}{p(1-p)}.$$

□