Probability and Statistics

Exercise sheet 7 - Solutions

MC 7.1. Let Z be a random variable with distribution function:

$$F_Z(z) = \begin{cases} 0, & \text{if } z < 0, \\ 0.1, & \text{if } 0 \le z < 1, \\ 0.5, & \text{if } 1 \le z < 3, \\ 0.8, & \text{if } 3 \le z < 5, \\ 1, & \text{if } z \ge 5. \end{cases}$$

(Exactly one answer is correct in each question.)

1. Is $\mathbb{E}[Z] \ge 3$?
   (a) Yes.
   (b) No.
2. Is $\mathbb{P}[Z \le 3] = \mathbb{P}[Z \ge 3]$?
   (a) Yes.
   (b) No.
3. Is $\mathbb{P}[3.5 \le Z \le 5.5] = 0.2$?
   (a) Yes.
   (b) No.
4. What is $\mathbb{E}[Z^2]$?
   (a) $\mathbb{E}[Z^2] = 8.1$.
   (b) $\mathbb{E}[Z^2] = 3.5$.
   (c) $\mathbb{E}[Z^2] = 21$.
   (d) Doesn't exist.
   (e) $\mathbb{E}[Z^2] = \infty$.
5. Is $\mathbb{P}[Z = 0] = 0$?
   (a) Yes.
   (b) No.

Solution:

1. (b). Z is a discrete random variable and the probabilities are: $\mathbb{P}[Z = 0] = 0.1$, $\mathbb{P}[Z = 1] = 0.4$, $\mathbb{P}[Z = 3] = 0.3$, $\mathbb{P}[Z = 5] = 0.2$. Thus,

$$\mathbb{E}[Z] = 0 \times 0.1 + 1 \times 0.4 + 3 \times 0.3 + 5 \times 0.2 = 2.3 < 3.$$

2. (b). We have $\mathbb{P}[Z \leq 3] = 0.8$, but $\mathbb{P}[Z \geq 3] = 1 - \mathbb{P}[Z < 3] = 1 - 0.5 = 0.5$.
3. (a). We have $\mathbb{P}[3.5 \leq Z \leq 5.5] = \mathbb{P}[Z = 5] = 0.2$.
4. (a). We compute:
$$\mathbb{E}[Z^2] = 0^2 \times 0.1 + 1^2 \times 0.4 + 3^2 \times 0.3 + 5^2 \times 0.2 = 0 + 0.4 + 2.7 + 5 = 8.1.$$
5. (b). $\mathbb{P}[Z = 0] = 0.1 \neq 0$.

Exercise 7.2. Let $r > 1$ and define $f : \mathbb{R} \to \mathbb{R}$ by
$$f(x) = \begin{cases} 0 & \text{for } x \leq 1, \\ cx^{-r} & \text{for } x > 1, \end{cases}$$
for some constant $c \in \mathbb{R}$.

(a) Determine the constant c such that f is a density.

(b) Let X be a random variable with density $f_X = f$. Compute the cumulative distribution function of X.

(c) Compute the expected value of X. For which values of r is the expected value finite?

Solution:

(a) To make f a probability density function, we require $\int f(x)\mathrm{d}x = 1$ and $f \geq 0$. We have
$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = \int_1^{\infty} cx^{-r}\mathrm{d}x = c\frac{x^{-r+1}}{-r+1}\Big|_{x=1}^{\infty} = \frac{c}{r-1}.$$
Thus, $c = r - 1$. Checking that $f(x) \geq 0$, $x \in \mathbb{R}$, is satisfied for $c = r - 1$ is straightforward.

(b) The cumulative distribution function F_X is:
$$F_X(t) = \int_{-\infty}^{t} f_X(x)\mathrm{d}x, \quad t \in \mathbb{R}.$$
For $t \leq 1$, we have $F_X(t) = 0$ (since $f_X(x) = 0$ on $(-\infty, 1]$). For $t > 1$:
$$F_X(t) = \int_{-\infty}^{t} f_X(x)\mathrm{d}x = \int_1^{t} (r-1)x^{-r}\mathrm{d}x = (r-1)\frac{x^{-r+1}}{-r+1}\Big|_{x=1}^{t} = 1 - t^{-r+1}.$$
So the cumulative distribution function is:
$$F_X(t) = \begin{cases} 0 & \text{for } t \leq 1, \\ 1 - t^{-r+1} & \text{for } t > 1. \end{cases}$$

(c) Since $\mathbb{P}[X \geq 1] = 1$, the expectation is defined. For $r \neq 2$:

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x \cdot f_X(x)\mathrm{d}x = \int_1^{\infty} x \cdot (r-1)x^{-r}\mathrm{d}x = \int_1^{\infty} (r-1)x^{-r+1}\mathrm{d}x$$
$$= (r-1)\frac{x^{-r+2}}{-r+2}\bigg|_{x=1}^{\infty} = \begin{cases} \frac{r-1}{r-2}, & \text{if } r > 2, \\ \infty, & \text{if } 1 < r < 2. \end{cases}$$

For $r = 2$, we compute:

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x \cdot f_X(x)\mathrm{d}x = \int_1^{\infty} x \cdot x^{-2}\mathrm{d}x = \int_1^{\infty} x^{-1}\mathrm{d}x = \infty.$$

Therefore, the expected value is finite if and only if $r > 2$.

Exercise 7.3. $k \in \mathbb{N}$ hunters each shoot once simultaneously at a flock of $m \in \mathbb{N}$ ducks. They independently choose which duck to aim at, and they hit their chosen duck independently of each other and independently of the duck selected, with probability $p \in (0, 1)$.

Introduce for each duck $n \in \{1, \ldots, m\}$ a random variable X_n indicating whether the duck was hit (by at least one hunter) or not. We define $\{X_n = 1\}$ = "n-th duck not hit" and $\{X_n = 0\}$ = "n-th duck hit".

(a) What is the distribution of X_n for $n = 1, \ldots, m$?

(b) What is the expected number of unharmed ducks?

(c) Are the events $\{X_n = 0\}$, $n \in \{1, \ldots, m\}$ independent? Consider only the case $k < m$ for simplicity.

Solution:

(a) X_n can take only values in $\{0, 1\}$. The probability that the n-th duck is not hit by the ℓ-th hunter is $1 - p/m$. Since the hunters shoot independently, the probability that the n-th duck, and thus any duck, is unharmed is

$$\mathbb{P}[X_n = 1] = \left(1 - \frac{p}{m}\right)^k.$$

Hence, all X_n follow a binomial distribution with parameters $\tilde{n} = 1$ and $\tilde{p} = (1 - p/m)^k$, i.e., a Bernoulli distribution with parameter $\tilde{p}$.

(b) The total number X of unharmed ducks is $X = X_1 + X_2 + \cdots + X_m$. Due to the linearity of expectation,

$$\mathbb{E}[X] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \cdots + \mathbb{E}[X_m].$$

Since the X_n are Bernoulli variables, we have $\mathbb{E}[X_n] = \mathbb{P}[X_n = 1] = (1 - p/m)^k$ for all $n \in \{1, \ldots, m\}$, and thus

$$\mathbb{E}[X] = m\left(1 - \frac{p}{m}\right)^k.$$

(c) We consider only the case $k < m$, i.e., fewer hunters than ducks. Then $X_1, \ldots, X_m$ are not independent because

$$\mathbb{P}[X_1 = \cdots = X_m = 0] = 0 < \left(1 - \left(1 - \frac{p}{m}\right)^k\right)^m = \prod_{n=1}^{m} \mathbb{P}[X_n = 0].$$

Remark: The random variables $X_1, \ldots, X_m$ are also not independent if $k \geq m$ and $m > 1$. In particular, X is not binomially distributed.

Exercise 7.4. We consider a circle which has a random radius R. The radius R is exponentially distributed with expectation $1/\lambda$ for some $\lambda > 0$. Let us denote by A the (random) area of this circle. Determine:

(a) The distribution function and the density function of A;

(b) The expected value of A.

Solution:

(a) Let X be exponentially distributed with parameter μ, i.e., the density of X is $f_X(x) = \mu e^{-\mu x}$ for $x \geq 0$, and 0 otherwise. By integration by parts, we have:

$$\begin{aligned}\mathbb{E}[X] &= \int_0^\infty x\mu e^{-\mu x}\mathrm{d}x = -xe^{-\mu x}\Big|_{x=0}^{\infty} - \int_0^\infty (-e^{-\mu x})\mathrm{d}x \\ &= 0 + \frac{1}{\mu}\int_0^\infty \mu e^{-\mu x}\mathrm{d}x = \frac{1}{\mu}\int_{-\infty}^\infty f_X(x)\mathrm{d}x = \frac{1}{\mu}\cdot 1 = \frac{1}{\mu}.\end{aligned}$$

Thus, since R has expectation $1/\lambda$, we conclude that R is exponentially distributed with parameter $\mu = \lambda$.

The area of the circle with radius R is given by the random variable $A = \pi R^2$. The distribution function of A is:

$$\begin{aligned}F_A(x) &= \mathbb{P}[A \leq x] = \mathbb{P}[\pi R^2 \leq x] = \mathbb{P}\Big[R \leq \sqrt{x/\pi}\Big] = F_R\Big(\sqrt{x/\pi}\Big) = \int_{-\infty}^{\sqrt{x/\pi}} f_X(t)\mathrm{d}t \\ &= \int_0^{\sqrt{x/\pi}} \lambda e^{-\lambda t}\mathrm{d}t = -e^{-\lambda t}\Big|_{t=0}^{\sqrt{x/\pi}} = 1 - e^{-\lambda\sqrt{x/\pi}}, \quad \text{for } x \geq 0,\end{aligned}$$

and 0 otherwise.

The density function is then given by:

$$f_A(x) = \frac{\mathrm{d}}{\mathrm{d}x}F_A(x) = \frac{\lambda}{2\sqrt{\pi x}}e^{-\lambda\sqrt{x/\pi}}, \quad \text{for } x \geq 0,$$

and 0 otherwise.

Alternatively, since $F_A(x) = F_R\big(\sqrt{x/\pi}\big)$, applying the chain rule yields:

$$f_A(x) = \frac{\mathrm{d}}{\mathrm{d}x}F_A(x) = \frac{\mathrm{d}}{\mathrm{d}x}F_R\big(\sqrt{x/\pi}\big) = f_R\Big(\sqrt{x/\pi}\Big)\frac{\mathrm{d}}{\mathrm{d}x}\sqrt{\frac{x}{\pi}} = \frac{1}{2\sqrt{x\pi}}\lambda e^{-\lambda\sqrt{x/\pi}} \quad \text{for } x \geq 0,$$

and 0 otherwise.

Alternatively, we may use substitution as in the sample solution to Exercise 7.5.(c) to obtain the density.

(b) Using integration by parts as in (a):

$$\begin{aligned}\mathbb{E}[A] = \mathbb{E}[\pi R^2] &= \int_0^\infty \pi t^2 f_R(t)\mathrm{d}t = \pi\lambda \int_0^\infty t^2 e^{-\lambda t}\mathrm{d}t \\ &= \pi\lambda\Big(t^2\frac{e^{-\lambda t}}{-\lambda}\Big|_{t=0}^\infty - \int_0^\infty 2t\frac{e^{-\lambda t}}{-\lambda}\mathrm{d}t\Big) = 2\pi\int_0^\infty te^{-\lambda t}\mathrm{d}t = \frac{2\pi}{\lambda^2}.\end{aligned}$$

Remark: Of course, it is also possible to determine the expected value using the density f_A from part (a).

Exercise 7.5. A random variable X has the density function:

$$f(x) = \begin{cases} \dfrac{c}{(1+x)^5}, & x > 0, \\ 0, & x \le 0. \end{cases}$$

(a) Find the value of c and the distribution function of X.

(b) Find $\mathbb{E}[X]$ and $\mathbb{E}[X^2]$.

Hint: It might be easier to first compute $\mathbb{E}[1+X]$ and $\mathbb{E}[(1+X)^2]$ and then use linearity.

(c) What are the distribution function and the density of $Y := e^X$?

Solution:

(a) We require that $\int_{-\infty}^\infty f(x)\mathrm{d}x = 1$ and $f \ge 0$. We compute:

$$\int_{-\infty}^\infty f(x)\mathrm{d}x = \int_0^\infty \frac{c}{(1+x)^5}\mathrm{d}x = c\Big(-\frac{1}{4}(1+x)^{-4}\Big)\Big|_{x=0}^\infty = \frac{c}{4}$$

yielding $c = 4$. Checking $f \ge 0$ for $c = 4$ is straightforward.

The distribution function is:

$$F_X(x) = \mathbb{P}[X \le x] = \int_{-\infty}^x f(y)\mathrm{d}y = \int_0^x \frac{4}{(1+y)^5}\mathrm{d}y = -(1+y)^{-4}\big|_{y=0}^x = 1 - \frac{1}{(1+x)^4}, \text{ for } x \ge 0,$$

and 0 otherwise.

(b) We first compute:

$$\begin{aligned}\mathbb{E}[1+X] &= \int_0^\infty (1+x)\cdot\frac{4}{(1+x)^5}\mathrm{d}x = \int_0^\infty \frac{4}{(1+x)^4}\mathrm{d}x = 4\Big(-\frac{1}{3}(1+x)^{-3}\Big)\Big|_{x=0}^\infty = \frac{4}{3}, \\ \mathbb{E}[(1+X)^2] &= \int_0^\infty \frac{4}{(1+x)^3}\mathrm{d}x = 4\Big(-\frac{1}{2}(1+x)^{-2}\Big)\Big|_{x=0}^\infty = 2.\end{aligned}$$

Thus,

$$\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}[1 + X - 1] = \mathbb{E}[1 + X] - 1 = \frac{1}{3}, \\
\mathbb{E}[X^2] &= \mathbb{E}[(1 + X)^2 - 2X - 1] = \mathbb{E}[(1 + X)^2] - 2\mathbb{E}[X] - 1 = 2 - \frac{2}{3} - 1 = \frac{1}{3}.
\end{aligned}$$

(c) Since $X \geq 0$, it follows that $Y = e^X \geq 1$. For $y < 1$, the distribution function F_Y of Y thus satisfies:

$$F_Y(y) = \mathbb{P}[Y \leq y] = \mathbb{P}[e^X \leq y] \leq \mathbb{P}[e^X < 1] = \mathbb{P}[X < 0] = 0.$$

For $y \geq 1$, we get:

$$F_Y(y) = \mathbb{P}[e^X \leq y] = \mathbb{P}[X \leq \log y] = F_X(\log y) = 1 - \frac{1}{(1 + \log y)^4}.$$

By differentiating the distribution function, we get the density f_Y of Y:

$$f_Y(y) = \frac{\mathrm{d}}{\mathrm{d}y} F_Y(y) = \begin{cases} 0 & \text{for } y < 1, \\ \frac{4}{y(1+\log y)^5} & \text{for } y \geq 1. \end{cases}$$

Alternatively, we have for $y > 1$ that

$$F_Y(y) = \mathbb{P}[e^X \leq y] = \mathbb{P}[X \leq \log y] = \int_0^{\log(y)} \frac{4}{(1 + x)^5} \mathrm{d}x = \int_1^y \frac{4}{t(1 + \log t)^5} \mathrm{d}t = \int_{-\infty}^y f_Y(t) \mathrm{d}t,$$

where we have used the substitution $t = e^x$ in the second-to-last equality. We thus obtain the same result as above.