Statistical Theory MATH-442
Spring 2025 Ho Yun & Yoav Zemel

Exercise sheet 1

You may wish to refer to the probability recap document on Moodle about definitions of some distributions, if needed.

Exercise 1 If X is exponentially distributed with intensity λ, what is the distribution of $Y = \lfloor X \rfloor$, the largest integer less than or equal to X?

Exercise 2 Suppose that $S \sim \text{Exp}(\lambda)$, $C \sim \text{Exp}(\gamma)$ are independent. Define $T = \min(S, C)$ and $D = \mathbf{1}[T = S]$. Find the joint distribution of T and D and their marginal distributions. Are T and D independent?

Solution 1 Y has a discrete distribution. For $k = 0, 1, \ldots$, we compute:

$$\begin{aligned} P(Y = k) &= P(\lfloor X \rfloor = k) = P(X \in [k, k+1)) \\ &= \int_k^{k+1} \lambda e^{-\lambda x} dx = e^{-k\lambda}(1 - e^{-\lambda}) \end{aligned}$$

Thus, Y follows a geometric distribution with success probability $1 - e^{-\lambda}$.

Solution 2 T has a continuous distribution, while D is discrete (taking values 0 and 1). We compute:

$$\begin{aligned} P(T \leq t, D = 1) = P(S \leq t, S \leq C) &= \int_0^t \left(\int_s^\infty \lambda e^{-\lambda s} \gamma e^{-\gamma c} dc \right) ds \\ &= \tfrac{\lambda}{\lambda + \gamma}(1 - e^{-(\lambda+\gamma)t}). \end{aligned}$$

Similarly,

$$P(T \leq t, D = 0) = \tfrac{\gamma}{\lambda + \gamma}(1 - e^{-(\lambda+\gamma)t}).$$

Marginally, $D \sim \text{Bernoulli}(\lambda/(\lambda + \gamma))$ and $T \sim \text{Exp}(\lambda + \gamma)$. Since the product of marginals equals the joint distribution, T and D are independent.

Exercise 3 Consider a random vector $(X, Z)^T$. Let the marginal distribution of Z be exponential with parameter γ, i.e., with density $f_Z(z) = \gamma e^{-\gamma z} 1_{(0,\infty)}(z)$. Suppose that the conditional distribution of X given $Z = z$ is Poisson with parameter λz, i.e.,

$$P(X = x|Z = z) = \tfrac{(\lambda z)^x}{x!} e^{-\lambda z}, \quad x = 0, 1, \ldots, z > 0.$$

1. Find the joint density $f_{X,Z}(x, z)$.
2. Compute the marginal (unconditional) distribution of X. Which known distribution is it?
3. Find $E[X|Z]$.
4. Find $E[X]$.
5. Find $\text{var}(X|Z)$.
6. Compute $\text{var}(X)$.
7. Compute the conditional density $f_{Z|X}(z|x)$. Which known distribution is it?

Solution 3

1. $f_{X,Z}(x, z) = \gamma e^{-(\lambda+\gamma)z} \frac{(\lambda z)^x}{x!} 1_{\mathbb{N}_0 \times \mathbb{R}^+}(x, z)$.
2. X is geometric with success probability $\gamma/(\lambda + \gamma)$.
3. $E[X|Z] = \lambda Z$.
4. $E[X] = \lambda/\gamma$.
5. $\text{var}(X|Z) = \lambda Z$.
6. $\text{var}(X) = \lambda/\gamma + \lambda^2/\gamma^2$.
7. $Z|X = x \sim \Gamma(\lambda + \gamma, x + 1)$.

Exercise 4 Show that for a nonnegative random variable X,

$$\mathbb{E}X = \int_0^\infty P(X > t)dt \in [0, \infty].$$

Try to prove it without assuming the existence of a density/mass function. *Hint: Use Fubini/Tonelli theorem.*

Remark. This result is true even if $\mathbb{E}[X]$ is finite, since we are applying the theorem over σ-finite measures: Lebegue meausure on $\mathbb{R}$ and the finite probability measure corresponding to X.

Exercise 5 Note: hint on the following page! This is a nice (in the lecturer's opinion!) geometrical result we shall need in the part of the course about classification. It can be interpreted as the triangle inequality for angles. For $d \geq 2$ and nonzero vectors $y, z \in \mathbb{R}^d$ define the angle between them by

$$\theta(y, z) = \cos^{-1} \tfrac{y^\top z}{\|y\|\|z\|} \in [0, \pi].$$

Let $v \in \mathbb{R}^d \setminus \{0\}$ be an additional vector such that $\theta(y, v) \leq \pi/2$ and $\theta(z, v) \leq \pi/2$. Show that

$$\theta(y, z) \leq \theta(y, v) + \theta(v, z).$$

Hint: it might be easier to restrict (with justification) to the case where v is the first unit vector, and all vectors have norm one. You may also recall the formula $\cos(\alpha + \beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$, and use that cos is decreasing on $[0, \pi]$.

Remark. This is also true, but not really interesting, if $d = 1$.

Solution 4 Compute (using Fubini's theorem)

$$\int_0^\infty P(X > t)dt = \int_0^\infty \mathbb{E}1(X > t)dt = \mathbb{E}\int_0^\infty 1(X > t)dt = \mathbb{E}\int_0^X 1dt = \mathbb{E}X.$$

Equivalently, if you prefer such notation, write

$$\int_0^\infty P(X > t)dt = \int_0^\infty \left(\int_t^\infty dP_X(x)\right)dt = \int_0^\infty \left(\int_0^x dt\right)dP_X(x) = \int_0^\infty xdP_X(x) = \mathrm{E}X.$$

Solution 5 Dividing v, y and z by their (nonzero) norms does not change any of the angles, so we may assume that $\|v\| = \|y\| = \|z\| = 1$. Multiplying the vectors by an orthogonal transformation does not change the angles either, so we may assume that $v = (1, 0, \ldots, 0)^\top$ is the unit vector in $\mathbb{R}^d$. Then we know that

$$0 \leq \cos\theta(y, v) = y^\top v = y_1;$$
$$0 \leq \cos\theta(z, v) = z^\top v = z_1,$$

and by Cauchy(–Bunyakovsky)–Schwarz inequality

$$\cos\theta(y, z) = y^\top z = y_1 z_1 + \sum_{k=2}^d y_k z_k \geq y_1 z_1 - \sqrt{\sum_{k=2}^d y_k^2}\sqrt{\sum_{k=2}^d z_k^2}$$

$$= y_1 z_1 - \sqrt{1 - y_1^2}\sqrt{1 - z_1^2} = \cos\theta(y, v)\cos\theta(z, v) - \sin\theta(y, v)\sin\theta(z, v) = cos(\theta(y, v) + \theta(z, v)).$$

Since the cosine function is strictly decreasing on $[0, \pi]$, the result follows.