CS 6660: Mathematical Foundations of Data Science (Probability)

Instructor: Dr. Karthik P. N.

Practice Problems 02

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. All random variables appearing below are defined with respect to $\mathscr{F}$.

1. Virat and Anushka have a date at 7 pm. Each will arrive at the meeting place with a delay that is distributed uniformly randomly between 0 minutes and 60 minutes, independent of the delay of the other. The first to arrive will wait for 15 minutes and leave if the other does not arrive within 15 minutes. Find the probability that both meet.

   Solution: Let X denote the delay in Virat's arrival, and let Y denote the delay in Anushka's arrival. If both are to meet, then the event of interest is $\{|X-Y| \leq 15\}$. To compute this probability, we simply note that

   $$\mathbb{P}(\{|X-Y| \leq 15\}) = \int_{(x,y):|x-y|\leq 15} f_{X,Y}(x,y)\, \mathrm{d}x\, \mathrm{d}y = \int_{(x,y):|x-y|\leq 15} \frac{1}{60^2}\, \mathrm{d}x\, \mathrm{d}y.$$

   We now note that

   $$\begin{aligned}&\mathbb{P}(\{|X-Y| \leq 15\})\\ &= \mathbb{P}(\{|X-Y| \leq 15\} \cap \{X \leq 15\}) + \mathbb{P}(\{|X-Y| \leq 15\} \cap \{15 < X < 45\}) + \mathbb{P}(\{|X-Y| \leq 15\} \cap \{X \geq 45\}).\end{aligned} \tag{1}$$

   The first term on the right-hand side of (1) is given by

   $$\begin{aligned}\mathbb{P}(\{|X-Y| \leq 15\} \cap \{X \leq 15\}) &= \mathbb{P}(\{X \leq 15\} \cap \{0 \leq Y \leq X+15\})\\ &= \int_0^{15} \int_0^{x+15} \frac{1}{60^2}\, \mathrm{d}x\, \mathrm{d}y = \frac{3}{32}.\end{aligned}$$

   By symmetry, it follows that the last term on the right-hand side of (1) is also equal to $3/32$. The second term on the right-hand side of (1) is given by

   $$\mathbb{P}(\{|X-Y| \leq 15\} \cap \{15 < X < 45\}) = \int_{15}^{45} \int_{x-15}^{x+15} \frac{1}{60^2}\, \mathrm{d}x\, \mathrm{d}y = \frac{1}{4}.$$

   Combining the above results, we get

   $$\mathbb{P}(\{|X-Y| \leq 15\}) = \frac{3}{32} + \frac{1}{4} + \frac{3}{32} = \frac{7}{16}.$$

2. Let X and Y be continuous random variables with PDFs f_X and f_Y respectively. For any $\alpha \in [0,1]$, argue that $\alpha\, f_X + (1-\alpha)\, f_Y$ is a valid PDF. Can you think of a random variable Z whose PDF is $f_Z = \alpha\, f_X + (1-\alpha)\, f_Y$?

   Solution: Let F_X and F_Y denote the CDFs of the random variables X and Y respectively. Thus,

   $$F_X(z) = \int_{-\infty}^{z} f_X(t)\, \mathrm{d}t, \qquad F_Y(z) = \int_{-\infty}^{z} f_Y(t)\, \mathrm{d}t \quad \forall z \in \mathbb{R}.$$

   For any $\alpha \in [0,1]$, consider the function F defined as

   $$F(z) = \alpha F_X(z) + (1-\alpha) F_Y(z), \qquad z \in \mathbb{R}.$$

   Then, we observe that

   - $F(z) \in [0,1]$ for every $z \in \mathbb{R}$.

- $\lim_{z\to-\infty} F(z) = 0, \quad \lim_{z\to+\infty} F(z) = 1.$
- F is non-decreasing, i.e., $F(z) \le F(z')$ for $z \le z'$.
- F is right-continuous, i.e., $\lim_{\varepsilon\downarrow 0} F(z+\varepsilon) = F(z)$ for every $z \in \mathbb{R}$.

Thus, F is a valid CDF. Furthermore,

$$F(z) = \int_{-\infty}^{z} (\alpha f_X(t) + (1-\alpha) f_Y(t))\, \mathrm{d}t,$$

thus proving that $\alpha f_X + (1-\alpha) f_Y$ a valid PDF.

Let $W \sim \text{Bernoulli}(\alpha)$ be independent of both X and Y. For each $\omega \in \Omega$, let

$$Z(\omega) := \begin{cases} X(\omega), & \text{if } W(\omega) = 1, \\ Y(\omega), & \text{if } W(\omega) = 0. \end{cases}$$

Then, we observe that

$$\begin{aligned} F_Z(z) &= \mathbb{P}(\{Z \le z\}) \\ &= \mathbb{P}(\{Z \le z\} \cap \{W = 1\}) + \mathbb{P}(\{Z \le z\} \cap \{W = 0\}) \\ &= \mathbb{P}(\{X \le z\} \cap \{W = 1\}) + \mathbb{P}(\{Y \le z\} \cap \{W = 0\}) \\ &\overset{(a)}{=} \mathbb{P}(\{X \le z\}) \cdot \mathbb{P}(\{W = 1\}) + \mathbb{P}(\{Y \le z\}) \cdot \mathbb{P}(\{W = 0\}) \\ &= \alpha F_X(z) + (1-\alpha) F_Y(z), \end{aligned}$$

where (a) above follows from the fact that $W \perp\!\!\!\perp X$ and $W \perp\!\!\!\perp Y$. From the above set of equalities, it follows that

$$f_Z = \alpha f_X + (1-\alpha) f_Y.$$

3. Let X and Y be jointly continuous random variables with the joint PDF

$$f_{X,Y}(x,y) = \begin{cases} cx(y-x)e^{-y}, & 0 \le x \le y < +\infty, \\ 0, & \text{otherwise.} \end{cases}$$

   (a) Determine the constant c.
   (b) Show that

$$f_{X|Y=y}(x) = \begin{cases} 6x(y-x)y^{-3}, & 0 \le x \le y, \\ 0, & \text{otherwise,} \end{cases} \qquad f_{Y|X=x}(y) = \begin{cases} (y-x)e^{x-y}, & x \le y < +\infty, \\ 0, & \text{otherwise,} \end{cases}$$

Solution: We present the solution to each of the parts below.

(a) To determine the constant c, we set the integral of the joint PDF to 1. Doing so, we obtain $c = 1$.

(b) From the joint PDF expression, we first obtain the marginal PDFs of X and Y. For any $0 \le y < +\infty$, we note that

$$f_Y(y) = \int_0^y x(y-x)\, e^{-y}\, \mathrm{d}x = \frac{y^3\, e^{-y}}{6},$$

from which it follows that

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \begin{cases} 6x(y-x)y^{-3}, & 0 \le x \le y, \\ 0, & \text{otherwise.} \end{cases}$$

Along similar lines, for any $0 \le x < +\infty$, we have

$$f_X(x) = \int_x^{\infty} x(y-x)e^{-y}\, \mathrm{d}y = xe^{-x},$$

from which it follows that for all $x > 0$,

$$f_{Y|X=x}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \begin{cases} (y-x)\, e^{-(y-x)}, & x \le y < +\infty, \\ 0, & \text{otherwise.} \end{cases}$$

4. Let X and Y be independent Poisson variables with parameters λ and μ respectively. Fix $n \in \mathbb{N}$. Determine the conditional PMF of X, conditioned on the event $\{X+Y=n\}$.

   Solution: Conditioned on the event $\{X+Y=n\}$, it follows that X takes values in the set $\{0,\ldots,n\}$. For any $n \in \mathbb{N}\cup\{0\}$,

$$\begin{aligned}\mathbb{P}(\{X=k\}|\{X+Y=n\}) &= \frac{\mathbb{P}(\{X=k\}\cap\{X+Y=n\})}{\mathbb{P}(\{X+Y=n\})} \\ &= \frac{\mathbb{P}(\{X=k\}\cap\{Y=n-k\})}{\mathbb{P}(\{X+Y=n\})}.\end{aligned}$$

   We first proceed to derive the denominator probability term in closed form. Observe that

$$\begin{aligned}\mathbb{P}(X+Y=n) &= \sum_{(k,l):k+l=n} p_{X,Y}(k,l) \\ &= \sum_{k=0}^{n} p_X(k)\cdot p_Y(n-k) \\ &= \sum_{k=0}^{n} e^{-\lambda}\cdot\frac{\lambda^k}{k!}\times e^{-\mu}\cdot\frac{\mu^{n-k}}{(n-k)!} \\ &= \frac{e^{-(\lambda+\mu)}}{n!}\sum_{k=0}^{n}\binom{n}{k}\lambda^k\mu^{n-k} \\ &= e^{-(\lambda+\mu)}\cdot\frac{(\lambda+\mu)^n}{n!}.\end{aligned}$$

   That is $X+Y \sim$ Poisson$(\lambda+\mu)$. We then have

$$\begin{aligned}\mathbb{P}(\{X=k\}|\{X+Y=n\}) &= \frac{e^{-\lambda}\cdot\frac{\lambda^k}{k!}\times e^{-\mu}\cdot\frac{\mu^{n-k}}{(n-k)!}}{e^{-(\lambda+\mu)}\cdot\frac{(\lambda+\mu)^n}{n!}} \\ &= \binom{n}{k}\left(\frac{\lambda}{\lambda+\mu}\right)^k\left(\frac{\mu}{\lambda+\mu}\right)^{n-k},\end{aligned}$$

   thereby demonstrating that $X|\{X+Y=n\} \sim$ Binomial $\left(n,\ \frac{\lambda}{\lambda+\mu}\right)$.

5. Suppose that two batteries are chosen simultaneously and uniformly at random from the following group of 12 batteries : 3 new, 4 used (yet working), 5 defective. You may assume that all batteries within a particular group are identical. Let X denote the number of new batteries chosen, and let Y denote the number of used batteries chosen. Determine the joint PMF of X and Y, and compute $\mathbb{P}(\{|X-Y|\leq 1\})$.

Solution: Observe that $X \in \{0,1,2\}$, $Y \in \{0,1,2\}$, and $X + Y \le 2$. Furthermore,

$$p_{X,Y}(x,y) = \begin{cases} \dfrac{\binom{5}{2}}{\binom{12}{2}}, & x=0, y=0, \\[2ex] \dfrac{\binom{4}{1}\cdot\binom{5}{1}}{\binom{12}{2}}, & x=0, y=1, \\[2ex] \dfrac{\binom{4}{2}}{\binom{12}{2}}, & x=0, y=2, \\[2ex] \dfrac{\binom{3}{1}\cdot\binom{5}{1}}{\binom{12}{2}}, & x=1, y=0, \\[2ex] \dfrac{\binom{3}{1}\cdot\binom{4}{1}}{\binom{12}{2}}, & x=1, y=1, \\[2ex] \dfrac{\binom{3}{2}}{\binom{12}{2}}, & x=2, y=0, \\[2ex] 0, & \text{otherwise.} \end{cases}$$

We then note that

$$\mathbb{P}(\{|X-Y| \le 1\}) = p_{X,Y}(0,0) + p_{X,Y}(1,0) + p_{X,Y}(1,1) + p_{X,Y}(0,1) = \frac{57}{66} = \frac{19}{22}.$$

6. Suppose that X and Y have the joint PDF

$$f_{X,Y}(x,y) = \begin{cases} cy, & -1 \le x \le 1,\ 0 \le y \le |x|, \\ 0, & \text{otherwise.} \end{cases}$$

   (a) Determine the constant c.
   (b) Are X and Y independent?
   (c) Evaluate $\mathbb{P}(\{X \ge Y + 0.5\})$.
   (d) Evaluate $\mathbb{P}(\{X > 0.75\} | \{Y > 0.5\})$.

Solution: We present the solution to each part below.

(a) To determine the constant c, we integrate the joint PDF and set the integral to 1. Doing so, we get

$$1 = \int_{-1}^{1} \int_{0}^{|x|} cy \, \mathsf{d}y \, \mathsf{d}x = \int_{-1}^{1} c\,\frac{x^2}{2} \, \mathsf{d}x = \frac{c}{3},$$

from which it follows that $c = 3$.

(b) To determine if X is independent of Y, or otherwise, we first compute the marginal PDFs of X and Y. For any $x \in [-1,1]$, we have

$$f_X(x) = \int_0^{|x|} 3y \, \mathsf{d}y = \frac{3x^2}{2}.$$

Similarly, for any $y \in [0,1]$, we have

$$f_Y(y) = \int_{-1}^{-y} 3y \, \mathsf{d}x + \int_y^1 3y \, \mathsf{d}x = 6y(1-y).$$

Clearly, $f_{X,Y}(1,1) = 3 \ne 0 = f_X(1) f_Y(1)$, thereby proving that $X \not\perp\!\!\!\perp Y$.

(c) The desired probability is given by

$$\mathbb{P}(\{X \geq Y + 0.5\}) = \int_{0.5}^{1} \int_{0}^{x-0.5} 3y \, \mathsf{d}y \, \mathsf{d}x = \int_{0.5}^{1} \frac{3(x-0.5)^2}{2} \, \mathsf{d}x = \frac{1}{16}.$$

(d) We first compute the conditional CDF of X, conditioned on the event $A = \{Y > 0.5\}$. First, we note that

$$\mathbb{P}(A) = \int_{0.5}^{1} 6y(1-y) \, \mathsf{d}y = \frac{1}{2}.$$

Next, we note that

$$F_{X|A}(x) = \frac{\mathbb{P}(\{X \leq x\} \cap A)}{\mathbb{P}(A)} = \begin{cases} 0, & x < -1, \\ \dfrac{\int_{-1}^{-x} \int_{0.5}^{|u|} 3v \, \mathsf{d}v \, \mathsf{d}u}{1/2}, & -1 \leq x < -\frac{1}{2}, \\ \dfrac{\int_{-1}^{-0.5} \int_{0.5}^{|u|} 3v \, \mathsf{d}v \, \mathsf{d}u}{1/2}, & -\frac{1}{2} \leq x < \frac{1}{2}, \\ \dfrac{\int_{-1}^{-0.5} \int_{0.5}^{|u|} 3v \, \mathsf{d}v \, \mathsf{d}u + \int_{0.5}^{x} \int_{0.5}^{|u|} 3v \, \mathsf{d}v \, \mathsf{d}u}{1/2}, & \frac{1}{2} \leq x < 1, \\ 1, & x \geq 1. \end{cases}$$

Simplifying the integrals in the above expression, we get

$$F_X(x) = \begin{cases} 0, & x < -1, \\ 1 + x^3 - \frac{3}{4}(1+x), & -1 \leq x < -\frac{1}{2}, \\ \frac{1}{2}, & -\frac{1}{2} \leq x < \frac{1}{2}, \\ \frac{1}{2} + x^3 - \frac{3x}{4} + \frac{1}{4}, & \frac{1}{2} \leq x < 1, \\ 1, & x \geq 1. \end{cases}$$

Differentiating the above CDF expression with respect to x, we get

$$f_{X|A}(x) = \begin{cases} 3x^2 - \frac{3}{4}, & x \in [-1, -0.5] \cup [0.5, 1], \\ 0, & \text{otherwise.} \end{cases}$$

We then have

$$\mathbb{P}(\{X > 0.75\} | \{Y > 0.5\}) = \int_{0.75}^{1} f_{X|A}(x) \, \mathsf{d}x = \int_{0.75}^{1} \left(3x^2 - \frac{3}{4}\right) \, \mathsf{d}x = \frac{25}{64}.$$