STA 131A Introduction to Probability Theory

Midterm exam 2 solution

Instructor: Dogyoon Song

Problem 1.

(a) We need

$$\int_{-\infty}^{\infty} f_X(x)\,dx = 1.‘$$

Thus,

$$\begin{aligned}
1 &= \int_0^1 c(x+1)\,dx + \int_1^2 c(3-x)\,dx \\
&= c\left[\frac{x^2}{2}+x\right]_0^1 + c\left[3x-\frac{x^2}{2}\right]_1^2 \\
&= c\cdot\frac{3}{2} + c\cdot\frac{3}{2} = 3c.
\end{aligned}$$

Therefore,

$$c = \frac{1}{3}.$$

Since $x+1 \geq 0$ on $[0,1]$ and $3-x \geq 0$ on $(1,2]$, we have $f_X(x) \geq 0$ everywhere. Also, with $c = 1/3$, the total integral is 1. Hence f_X is a valid PDF.

(b) Using $c = 1/3$,

$$\begin{aligned}
P\left(\frac{1}{2} < X \leq \frac{3}{2}\right) &= \int_{1/2}^1 \frac{1}{3}(x+1)\,dx + \int_1^{3/2} \frac{1}{3}(3-x)\,dx \\
&= \frac{1}{3}\left[\frac{x^2}{2}+x\right]_{1/2}^1 + \frac{1}{3}\left[3x-\frac{x^2}{2}\right]_1^{3/2} \\
&= \frac{7}{24} + \frac{7}{24} = \frac{7}{12}.
\end{aligned}$$

(c) Using $c = 1/3$,

$$\begin{aligned}
\mathbb{E}[X] &= \int_0^1 x\cdot\frac{1}{3}(x+1)\,dx + \int_1^2 x\cdot\frac{1}{3}(3-x)\,dx \\
&= \frac{1}{3}\int_0^1 (x^2+x)\,dx + \frac{1}{3}\int_1^2 (3x-x^2)\,dx \\
&= \frac{1}{3}\left(\frac{1}{3}+\frac{1}{2}\right) + \frac{1}{3}\left[\frac{3x^2}{2}-\frac{x^3}{3}\right]_1^2 \\
&= \frac{5}{18} + \frac{13}{18} = 1.
\end{aligned}$$

Alternatively, we can simply observe the symmetry of $f_X(x)$ with respect to $x = 1$ and conclude $\mathbb{E}[X] = 1$.

Next,

$$\begin{aligned}
\mathbb{E}[X^2] &= \int_0^1 x^2 \cdot \frac{1}{3}(x+1)\,dx + \int_1^2 x^2 \cdot \frac{1}{3}(3-x)\,dx \\
&= \frac{1}{3}\int_0^1 (x^3+x^2)\,dx + \frac{1}{3}\int_1^2 (3x^2-x^3)\,dx \\
&= \frac{1}{3}\left(\frac{1}{4}+\frac{1}{3}\right) + \frac{1}{3}\left[x^3 - \frac{x^4}{4}\right]_1^2 \\
&= \frac{7}{36} + \frac{13}{12} = \frac{23}{18}.
\end{aligned}$$

Thus,

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{23}{18} - 1 = \frac{5}{18}.$$

Problem 2. Normal random variables

(a) Since $X \sim N(50, 8^2)$, we standardize using

$$Z = \frac{X-50}{8} \sim N(0,1).$$

Then

$$\begin{aligned}
P(44 < X \le 62) &= P\left(\frac{44-50}{8} < Z \le \frac{62-50}{8}\right) \\
&= P(-0.75 < Z \le 1.50) \\
&= \Phi(1.50) - \Phi(-0.75).
\end{aligned}$$

Using symmetry, $\Phi(-0.75) = 1 - \Phi(0.75)$. Thus,

$$P(44 < X \le 62) = \Phi(1.50) + \Phi(0.75) - 1.$$

From the table, $\Phi(1.50) \approx 0.9332$ and $\Phi(0.75) \approx 0.7734$. Therefore,

$$P(44 < X \le 62) \approx 0.9332 + 0.7734 - 1 = 0.7066 \approx 0.707.$$

(b) We need $P(X > q) = 0.10$, or equivalently, $P(X \le q) = 0.90$. Thus,

$$P\left(Z \le \frac{q-50}{8}\right) = 0.90.$$

From the table, $\Phi(1.28) \approx 0.8997$, so we may take

$$\frac{q-50}{8} \approx 1.28.$$

Therefore, $q \approx 50 + 8(1.28) = 60.24$.

(c) Since $Y = 2X + 5$ and $X \sim N(50, 8^2)$, Y is also normal. Its mean is

$$\mathbb{E}[Y] = 2\mathbb{E}[X] + 5 = 2(50) + 5 = 105.$$

Its variance is

$$\mathrm{Var}(Y) = 2^2\mathrm{Var}(X) = 4 \cdot 8^2 = 256.$$

Therefore,

$$Y \sim N(105, 256), \qquad \text{or equivalently,} \qquad Y \sim N(105, 16^2).$$

Problem 3. Two-stage completion times

(a) We first normalize:

$$1 = \int_0^1 \int_0^x c(x+y)\,dy\,dx.$$

Compute the inner integral:

$$\int_0^x (x+y)\,dy = x^2 + \frac{x^2}{2} = \frac{3}{2}x^2.$$

Thus,

$$1 = c\int_0^1 \frac{3}{2}x^2\,dx = c \cdot \frac{1}{2}, \qquad \text{and therefore} \qquad c = 2.$$

Hence

$$f_{X,Y}(x,y) = 2(x+y), \qquad 0 \le y \le x \le 1.$$

For the marginal density of X, for $0 \le x \le 1$,

$$\begin{aligned} f_X(x) &= \int_0^x 2(x+y)\,dy \\ &= 2x^2 + x^2 = 3x^2. \end{aligned}$$

Thus,

$$f_X(x) = \begin{cases} 3x^2, & 0 \le x \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

(b) For $0 < x \le 1$ and $0 \le y \le x$,

$$f_{Y|X}(y \mid x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{2(x+y)}{3x^2}.$$

Thus,

$$f_{Y|X}(y \mid x) = \begin{cases} \dfrac{2(x+y)}{3x^2}, & 0 \le y \le x, \\ 0, & \text{otherwise.} \end{cases}$$

Then

$$\begin{aligned} \mathbb{E}[Y \mid X = x] &= \int_0^x y\frac{2(x+y)}{3x^2}\,dy = \frac{2}{3x^2}\int_0^x (xy + y^2)\,dy \\ &= \frac{2}{3x^2}\left(\frac{x^3}{2} + \frac{x^3}{3}\right) \\ &= \frac{5x}{9}. \end{aligned}$$

(c) The random variables X and Y are not independent.

One way to see this is from the support. The support of the joint density is triangular:

$$0 \le y \le x \le 1.$$

However, the marginal supports of both X and Y lie in $[0,1]$. If X and Y were independent, then

$$f_{X,Y}(x,y) = f_X(x)f_Y(y), \quad \forall (x,y),$$

and thus, the joint support would necessarily be rectangular, up to sets of area zero. Since values with $0 < x < y < 1$ are impossible under the joint density, X and Y are not independent.

(d) We compute

$$P\left(Y \le \frac{X}{2}\right) = \int_0^1 \int_0^{x/2} 2(x+y)\,dy\,dx.$$

For fixed x,

$$\int_0^{x/2} 2(x+y)\,dy = \left[2xy + y^2\right]_0^{x/2} = x^2 + \frac{x^2}{4} = \frac{5}{4}x^2.$$

Thus,

$$P\left(Y \le \frac{X}{2}\right) = \int_0^1 \frac{5}{4}x^2\,dx = \frac{5}{12}.$$

(e*) Recall $U = Y$ and $V = X - Y$. Equivalently, when $U = u$, we have $Y = u$ and $V = X - u$.

The support is obtained directly from $0 \le Y \le X \le 1$:

$$u \ge 0, \qquad v \ge 0, \qquad u + v \le 1.$$

Now use the conditional-density factorization

$$f_{U,V}(u,v) = f_U(u) f_{V|U}(v \mid u).$$

Since $U = Y$, we have $f_U(u) = f_Y(u)$. Also, given $U = u$, the variable $V = X - u$ is just a shifted version of $X \mid Y = u$, so

$$f_{V|U}(v \mid u) = f_{X|Y}(u + v \mid u).$$

Therefore,

$$\begin{aligned} f_{U,V}(u,v) &= f_Y(u)\, f_{X|Y}(u+v \mid u) \\ &= f_Y(u)\, \frac{f_{X,Y}(u+v,u)}{f_Y(u)} \\ &= f_{X,Y}(u+v,u). \end{aligned}$$

Using $f_{X,Y}(x,y) = 2(x+y)$, we get

$$f_{U,V}(u,v) = 2\big((u+v)+u\big) = 4u + 2v$$

on the support $u \ge 0,\ v \ge 0,\ u + v \le 1$. Thus,

$$f_{U,V}(u,v) = \begin{cases} 4u + 2v, & u \ge 0,\ v \ge 0,\ u+v \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

The random variables U and V are not independent. Indeed, U and V can each be large individually, but they cannot be large simultaneously because

$$U + V = X \le 1.$$

For example, both events $\{U > 0.9\}$ and $\{V > 0.2\}$ have positive marginal probability, but

$$P(U > 0.9,\ V > 0.2) = 0.$$

Thus U and V are not independent.

Problem 4. Derived distributions and inverse problems

(a) Since $U \sim \text{Uniform}(0,1)$, we have $0 < U < 1$, so

$$T = -\log U \geq 0.$$

- For $t < 0$,

$$F_T(t) = 0.$$

- For $t \geq 0$,

$$\begin{aligned} F_T(t) &= P(T \leq t) = P(-\log U \leq t) \\ &= P(\log U \geq -t) = P(U \geq e^{-t}) \\ &= 1 - e^{-t}. \end{aligned}$$

Thus,

$$F_T(t) = \begin{cases} 0, & t < 0, \\ 1 - e^{-t}, & t \geq 0. \end{cases}$$

Differentiating,

$$f_T(t) = \begin{cases} e^{-t}, & t \geq 0, \\ 0, & t < 0. \end{cases}$$

Thus $T \sim \text{Exponential}(1)$.

(b) Conditional on $\Theta = \theta$, the density of Y is symmetric around θ. Therefore,

$$\mathbb{E}[Y \mid \Theta = \theta] = \theta.$$

Thus, as a random variable,

$$\mathbb{E}[Y \mid \Theta] = \Theta.$$

By the law of total expectation,

$$\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y \mid \Theta]] = \mathbb{E}[\Theta].$$

Since $\Theta \sim \text{Uniform}(-1,1)$,

$$\mathbb{E}[\Theta] = 0.$$

Therefore,

$$\mathbb{E}[Y] = 0.$$

(c) By Bayes' rule,

$$f_{\Theta|Y}(\theta \mid 1) = \frac{f_\Theta(\theta) f_{Y|\Theta}(1 \mid \theta)}{\int_{-1}^{1} f_\Theta(t) f_{Y|\Theta}(1 \mid t)\, dt}.$$

For $-1 \leq \theta \leq 1$,

$$f_\Theta(\theta) = \frac{1}{2}.$$

Also, since $1 - \theta \geq 0$ on $[-1, 1]$,

$$f_{Y|\Theta}(1 \mid \theta) = \frac{1}{2} e^{-|1-\theta|} = \frac{1}{2} e^{-(1-\theta)} = \frac{1}{2} e^{\theta - 1}.$$

Thus

$$\begin{aligned} f_\Theta(\theta) f_{Y|\Theta}(1 \mid \theta) &= \frac{1}{4} e^{\theta-1}, \\ \int_{-1}^{1} f_\Theta(t) f_{Y|\Theta}(1 \mid t)\, dt &= \int_{-1}^{1} \frac{1}{4} e^{t-1}\, dt = \frac{1}{4}\left[e^{t-1}\right]_{-1}^{1} = \frac{1}{4}(1 - e^{-2}). \end{aligned}$$

Therefore,

$$f_{\Theta|Y}(\theta \mid 1) = \frac{e^{\theta-1}}{1-e^{-2}}, \qquad -1 \le \theta \le 1.$$

Now

$$P(\Theta \ge 0 \mid Y = 1) = \int_0^1 \frac{e^{\theta-1}}{1-e^{-2}}\, d\theta = \frac{1-e^{-1}}{1-e^{-2}} = \frac{e}{e+1} \approx 0.731.$$

(d*) Let

$$D = A - B.$$

- For $d \ge 0$, write $A = d + B$. Then

$$\begin{aligned} f_D(d) &= \int_0^\infty f_A(d+b) f_B(b)\, db \\ &= \int_0^\infty e^{-(d+b)} e^{-b}\, db \\ &= e^{-d} \int_0^\infty e^{-2b}\, db = \frac{1}{2} e^{-d}. \end{aligned}$$

- For $d < 0$, write $B = A - d$. Then

$$\begin{aligned} f_D(d) &= \int_0^\infty f_A(a) f_B(a-d)\, da \\ &= \int_0^\infty e^{-a} e^{-(a-d)}\, da \\ &= e^{d} \int_0^\infty e^{-2a}\, da = \frac{1}{2} e^{d}. \end{aligned}$$

Combining the two cases,

$$f_D(d) = \frac{1}{2} e^{-|d|}, \qquad d \in \mathbb{R}.$$