STA 131A Introduction to Probability Theory

Final Exam Solution

Instructor: Dogyoon Song

Problem 1.

(a)

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{2} + \frac{2}{5} - \frac{1}{10} = \frac{4}{5},$$
$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{1/10}{2/5} = \frac{1}{4}.$$

A and B are not independent because

$$P(A)P(B) = \frac{1}{2} \cdot \frac{2}{5} = \frac{1}{5} \neq \frac{1}{10} = P(A \cap B).$$

(b) Since $Y = X^2$,

$$P(Y = 0) = P(X = 0) = \frac{1}{2},$$
$$P(Y = 1) = P(X = -1) = \frac{1}{4},$$
$$P(Y = 4) = P(X = 2) = \frac{1}{4}.$$

Thus,

$$p_Y(y) = \begin{cases} 1/2, & y = 0, \\ 1/4, & y = 1, \\ 1/4, & y = 4, \\ 0, & \text{otherwise.} \end{cases}$$

Therefore,

$$\mathbb{E}[Y] = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} + 4 \cdot \frac{1}{4} = \frac{5}{4}.$$

(c)

$$P(X \geq 1/2) = \int_{1/2}^{1} 3x^2 \, dx = \left[x^3\right]_{1/2}^{1} = 1 - \frac{1}{8} = \frac{7}{8},$$
$$\mathbb{E}[X] = \int_0^1 x \cdot 3x^2 \, dx = 3 \int_0^1 x^3 \, dx = \frac{3}{4}.$$

(d) Choose the position of the single 6 in $\binom{4}{1} = 4$ ways. Each of the remaining three rolls can be any of the five non-6 outcomes. Thus,

$$\binom{4}{1} 5^3 = 4 \cdot 125 = 500.$$

(e)

$$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}} = \frac{-6}{\sqrt{9\cdot 16}} = -\frac{1}{2}.$$

Also,

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y) = 9 + 16 + 2(-6) = 13.$$

Problem 2.

(a) For an urn with r red balls and b blue balls, the probability of drawing exactly one red in two draws without replacement is

$$\frac{\binom{r}{1}\binom{b}{1}}{\binom{6}{2}}.$$

Thus,

$$\begin{aligned} P(E \mid H = A) &= \frac{\binom{4}{1}\binom{2}{1}}{\binom{6}{2}} = \frac{8}{15}, \\ P(E \mid H = B) &= \frac{\binom{3}{1}\binom{3}{1}}{\binom{6}{2}} = \frac{9}{15} = \frac{3}{5}, \\ P(E \mid H = C) &= \frac{\binom{1}{1}\binom{5}{1}}{\binom{6}{2}} = \frac{5}{15} = \frac{1}{3}. \end{aligned}$$

Since the urns are equally likely,

$$P(E) = \frac{1}{3}\left(\frac{8}{15} + \frac{3}{5} + \frac{1}{3}\right) = \frac{1}{3}\left(\frac{8}{15} + \frac{9}{15} + \frac{5}{15}\right) = \frac{22}{45}.$$

(b) The posterior odds satisfy

$$\frac{P(H = B \mid E)}{P(H = C \mid E)} = \frac{P(E \mid H = B)P(H = B)}{P(E \mid H = C)P(H = C)}.$$

Since the priors are equal,

$$\frac{P(H = B \mid E)}{P(H = C \mid E)} = \frac{P(E \mid H = B)}{P(E \mid H = C)} = \frac{3/5}{1/3} = \frac{9}{5}.$$

Similarly,

$$\frac{P(H = A \mid E)}{P(H = C \mid E)} = \frac{8/15}{1/3} = \frac{8}{5}.$$

Since

$$\frac{9}{5} > \frac{8}{5} > 1,$$

the most likely state after observing E is $H = B$.

(c) Use unnormalized posterior weights

$$\begin{aligned} w_A &= P(E \mid H = A)P(H = A) = \frac{8}{45}, \\ w_B &= P(E \mid H = B)P(H = B) = \frac{1}{5} = \frac{9}{45}, \\ w_C &= P(E \mid H = C)P(H = C) = \frac{1}{9} = \frac{5}{45}. \end{aligned}$$

For Action 1, the unnormalized posterior expected payoff is

$$10w_A + 8w_B + 4w_C = \frac{80 + 72 + 20}{45} = \frac{172}{45}.$$

For Action 2, it is

$$5w_A + 6w_B + 8w_C = \frac{40 + 54 + 40}{45} = \frac{134}{45}.$$

Since

$$\frac{172}{45} > \frac{134}{45},$$

Action 1 has the larger posterior expected payoff.

Problem 3. Conditional discrete model

(a) Since $P(G = 0) = P(G = 1) = 1/2$,

$$P(X = -1) = P(G = 0)P(X = -1 \mid G = 0) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}.$$

Similarly,

$$P(X = 1) = \frac{1}{4}, \qquad P(X = -2) = \frac{1}{4}, \qquad P(X = 2) = \frac{1}{4}.$$

Thus X is uniform on $\{-2, -1, 1, 2\}$, and the marginal PMF of X is

$$p_X(x) = \begin{cases} \frac{1}{4} & \text{if } x \in \{-2, -1, 1, 2\}, \\ 0 & \text{otherwise.} \end{cases}$$

$$\therefore \mathbb{E}[X] = \frac{-2 - 1 + 1 + 2}{4} = 0.$$

(b) Use the law of total variance:

$$\mathrm{Var}(X) = \mathbb{E}[\mathrm{Var}(X \mid G)] + \mathrm{Var}(\mathbb{E}[X \mid G]).$$

For $G = 0$ and $G = 1$,

$$\begin{aligned} \mathbb{E}[X \mid G = 0] &= 0, & \mathrm{Var}(X \mid G = 0) &= 1, \\ \mathbb{E}[X \mid G = 1] &= 0, & \mathrm{Var}(X \mid G = 1) &= 4. \end{aligned}$$

Thus,

$$\mathbb{E}[\mathrm{Var}(X \mid G)] = \frac{1}{2}(1) + \frac{1}{2}(4) = \frac{5}{2}, \qquad \text{and} \qquad \mathrm{Var}(\mathbb{E}[X \mid G]) = \mathrm{Var}(0) = 0.$$

Therefore,

$$\mathrm{Var}(X) = \frac{5}{2}.$$

(c) Since $\mathbb{E}[X] = 0$,

$$\mathrm{Cov}(X, G) = \mathbb{E}[XG] - \mathbb{E}[X]\mathbb{E}[G] = \mathbb{E}[XG].$$

Using conditioning,

$$\mathbb{E}[XG] = \mathbb{E}[\mathbb{E}[XG \mid G]] = \mathbb{E}[G\mathbb{E}[X \mid G]].$$

But $\mathbb{E}[X \mid G] = 0$ for both values of G. Therefore,

$$\mathrm{Cov}(X, G) = 0.$$

However, X and G are not independent. For example,

$$P(X = 2 \mid G = 1) = \frac{1}{2}, \qquad \text{but} \qquad P(X = 2) = \frac{1}{4}.$$

Thus conditioning on G changes the distribution of X, so X and G are not independent.

Problem 4. Task timing

(a) The support is the parallelogram with vertices

$$(0,0), \quad (1,1), \quad (1,2), \quad (0,1).$$

Its area is 1, so normalization gives

$$c = 1.$$

Thus

$$f_{X,Y}(x,y) = 1 \qquad \text{on } R.$$

The random variables X and Y are not independent because the joint support is not rectangular. For example, if $X = x$, then Y must lie in $[x, x+1]$, so the possible values of Y depend on x.

(b) For fixed y, the allowable values of x satisfy

$$0 \le x \le 1, \qquad x \le y \le x+1.$$

Equivalently,

$$\max(0, y-1) \le x \le \min(1, y).$$

Thus,

$$f_Y(y) = \begin{cases} y, & 0 \le y \le 1, \\ 2-y, & 1 < y \le 2, \\ 0, & \text{otherwise.} \end{cases}$$

Since the joint density is constant, $X \mid Y = y$ is uniform over the corresponding interval. Hence,

$$\mathbb{E}[X \mid Y = y] = \begin{cases} \dfrac{y}{2}, & 0 < y \le 1, \\ \dfrac{(y-1)+1}{2} = \dfrac{y}{2}, & 1 < y < 2. \end{cases}$$

Therefore,

$$\mathbb{E}[X \mid Y = y] = \frac{y}{2}, \qquad 0 < y < 2.$$

(c) Let $D = Y - X$. Since $R = \{(x,y) : 0 \le x \le 1,\ x \le y \le x+1\}$, for each $x \in [0,1]$, the duration $D = Y - X$ ranges from 0 to 1.

For $0 \le d \le 1$,

$$F_D(d) = P(D \le d) = P(Y - X \le d).$$

Within the parallelogram, for each $x \in [0,1]$, the allowed range of y satisfying $Y - X \le d$ has length d, namely

$$x \le y \le x + d.$$

Since the joint density is 1,

$$F_D(d) = \int_0^1 \int_x^{x+d} 1 \, dy \, dx = d.$$

Thus,

$$F_D(d) = \begin{cases} 0, & d < 0, \\ d, & 0 \le d \le 1, \\ 1, & d > 1. \end{cases}$$

Differentiating,

$$f_D(d) = \begin{cases} 1, & 0 < d < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Thus $D \sim \text{Uniform}(0,1)$.

Problem 5. Tail bounds, CLT, and convergence

(a) For $X \sim \text{Exponential}(1/2)$,

$$P(X \geq 6) = e^{-(1/2)6} = e^{-3} \approx 0.050.$$

Using Markov's inequality,

$$P(X \geq 6) \leq \frac{\mathbb{E}[X]}{6} = \frac{2}{6} = \frac{1}{3}.$$

Using Chebyshev's inequality, since $X \geq 6$ implies $|X-2| \geq 4$,

$$P(X \geq 6) \leq P(|X-2| \geq 4) \leq \frac{\text{Var}(X)}{4^2} = \frac{4}{16} = \frac{1}{4}.$$

(b) For $X_i \sim \text{Exponential}(4)$,

$$\mathbb{E}[X_i] = \frac{1}{4}, \qquad \text{Var}(X_i) = \frac{1}{16}.$$

Thus,

$$\text{SD}(\bar{X}_{64}) = \sqrt{\frac{1/16}{64}} = \frac{1}{32}.$$

By the CLT,

$$P(0.20 \leq \bar{X}_{64} \leq 0.30) \approx P\left(\frac{0.20 - 0.25}{1/32} \leq Z \leq \frac{0.30 - 0.25}{1/32}\right).$$

The standardized bounds are

$$-1.6 \qquad \text{and} \qquad 1.6.$$

Therefore,

$$P(0.20 \leq \bar{X}_{64} \leq 0.30) \approx \Phi(1.6) - \Phi(-1.6) = 2\Phi(1.6) - 1.$$

Using $\Phi(1.6) \approx 0.9452$,

$$P(0.20 \leq \bar{X}_{64} \leq 0.30) \approx 2(0.9452) - 1 = 0.8904.$$

(c) For $0 \leq x \leq 1$,

$$P(M_n \leq x) = P(U_1 \leq x, \ldots, U_n \leq x) = x^n.$$

To check convergence in probability to 1, fix $\epsilon > 0$. For $0 < \epsilon < 1$,

$$P(|M_n - 1| \geq \epsilon) = P(M_n \leq 1 - \epsilon) = (1-\epsilon)^n \to 0.$$

For $\epsilon \geq 1$, the event is impossible since $M_n \in [0,1]$. Hence

$$M_n \xrightarrow{P} 1.$$

(d*) For fixed $t > 0$,

$$P(n(1 - M_n) > t) = P\left(M_n < 1 - \frac{t}{n}\right).$$

For $n > t$,

$$P\left(M_n < 1 - \frac{t}{n}\right) = \left(1 - \frac{t}{n}\right)^n.$$

Therefore,

$$\lim_{n\to\infty} P(n(1 - M_n) > t) = \lim_{n\to\infty} \left(1 - \frac{t}{n}\right)^n = e^{-t}.$$

Problem 6. Measurement model

(a) Conditional on $\Theta = \theta$,

$$Y = \theta + \varepsilon.$$

Thus,

$$\mathbb{E}[Y \mid \Theta = \theta] = \theta + \mathbb{E}[\varepsilon] = \theta, \qquad \text{and} \qquad \operatorname{Var}(Y \mid \Theta = \theta) = \operatorname{Var}(\varepsilon) = \frac{1}{6}.$$

Therefore,

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

Since $\Theta \sim \text{Uniform}(0, 2)$,

$$\mathbb{E}[\Theta] = 1, \qquad \text{and thus,} \qquad \mathbb{E}[Y] = 1.$$

By the law of total variance,

$$\operatorname{Var}(Y) = \mathbb{E}[\operatorname{Var}(Y \mid \Theta)] + \operatorname{Var}(\mathbb{E}[Y \mid \Theta]).$$

The two terms are respectively

$$\mathbb{E}[\operatorname{Var}(Y \mid \Theta)] = \frac{1}{6}, \qquad \operatorname{Var}(\mathbb{E}[Y \mid \Theta]) = \operatorname{Var}(\Theta).$$

Since $\Theta \sim \text{Uniform}(0, 2)$,

$$\operatorname{Var}(\Theta) = \frac{(2-0)^2}{12} = \frac{1}{3}.$$

Therefore,

$$\operatorname{Var}(Y) = \frac{1}{6} + \frac{1}{3} = \frac{1}{2}.$$

(b) By Bayes' rule,

$$f_{\Theta|Y}(\theta \mid 1.5) = \frac{f_\Theta(\theta) f_{Y|\Theta}(1.5 \mid \theta)}{\int_0^2 f_\Theta(u) f_{Y|\Theta}(1.5 \mid u)\, du}.$$

Since $\Theta \sim \text{Uniform}(0, 2)$, $f_\Theta(\theta) = 1/2$ on $[0, 2]$. The likelihood is positive when $|1.5 - \theta| \le 1$, so

$$0.5 \le \theta \le 2.$$

On this interval,

$$f_{Y|\Theta}(1.5 \mid \theta) = 1 - |1.5 - \theta| = \begin{cases} \theta - 0.5, & 0.5 \le \theta \le 1.5, \\ 2.5 - \theta, & 1.5 < \theta \le 2. \end{cases}$$

The prior density $f_\Theta(\theta) = 1/2$ is constant on $[0, 2]$, so it cancels in the normalized posterior density. Define

$$g(\theta) = \begin{cases} \theta - 0.5, & 0.5 \le \theta \le 1.5, \\ 2.5 - \theta, & 1.5 < \theta \le 2, \\ 0, & \text{otherwise.} \end{cases}$$

The normalizing constant is

$$\int_{0.5}^{1.5} (\theta - 0.5)\, d\theta + \int_{1.5}^{2} (2.5 - \theta)\, d\theta = \frac{1}{2} + \frac{3}{8} = \frac{7}{8}.$$

Therefore,

$$f_{\Theta|Y}(\theta \mid 1.5) = \begin{cases} \frac{8}{7}(\theta - 0.5), & 0.5 \le \theta \le 1.5, \\ \frac{8}{7}(2.5 - \theta), & 1.5 < \theta \le 2, \\ 0, & \text{otherwise.} \end{cases}$$

Now

$$P(\Theta \geq 1 \mid Y = 1.5) = \frac{\int_1^{1.5}(\theta - 0.5)\,d\theta + \int_{1.5}^{2}(2.5 - \theta)\,d\theta}{7/8}.$$

The numerator is

$$\frac{3}{8} + \frac{3}{8} = \frac{3}{4}.$$

Thus,

$$P(\Theta \geq 1 \mid Y = 1.5) = \frac{3/4}{7/8} = \frac{6}{7}.$$

(c) Conditional on $\Theta = \theta$,

$$\mathbb{E}[Y_i \mid \Theta = \theta] = \theta, \qquad \mathrm{Var}(Y_i \mid \Theta = \theta) = \mathrm{Var}(\varepsilon) = \frac{1}{6}.$$

Thus, by the CLT,

$$\bar{Y}_n \approx N\left(\theta, \frac{1}{6n}\right).$$

Therefore,

$$\begin{aligned} P(|\bar{Y}_n - \theta| \leq 0.05 \mid \Theta = \theta) &\approx P\left(|Z| \leq \frac{0.05}{\sqrt{1/(6n)}}\right) \\ &= P\left(|Z| \leq 0.05\sqrt{6n}\right). \end{aligned}$$

Using $P(|Z| \leq 1.96) \approx 0.95$, we require

$$0.05\sqrt{6n} \geq 1.96.$$

Equivalently,

$$n \geq \frac{1.96^2}{6(0.05)^2} \approx 256.11.$$

Hence the smallest integer satisfying the CLT-based condition is

$$n = 257.$$

(d*) The posterior density is proportional to a triangular shape with peak at 1.5, but it is truncated on the right at 2. Thus, it has a longer left side $[0.5, 1.5]$ than right side $[1.5, 2]$, and there is more posterior mass to the left of 1.5 than to the right. So

$$\mathbb{E}[\Theta \mid Y = 1.5] < 1.5.$$

Alternatively, one can directly compute

$$\begin{aligned} \mathbb{E}[\Theta \mid Y = 1.5] &= \frac{\int_{0.5}^{1.5} \theta(\theta - 0.5)\,d\theta + \int_{1.5}^{2} \theta(2.5 - \theta)\,d\theta}{7/8} \\ &= \frac{59/48}{7/8} = \frac{59}{42} \approx 1.405. \end{aligned}$$

Thus,

$$\mathbb{E}[\Theta \mid Y = 1.5] < 1.5.$$