# STA 131A Introduction to Probability Theory

## Practice Final Exam – Version B Solution

Instructor: DS

### Problem 1. Warm-up

**(a)** Using

$$P(A \cup B) = P(A) + P(B) - P(A \cap B),$$

we have

$$P(A \cap B) = P(A) + P(B) - P(A \cup B) = \frac{3}{5} + \frac{1}{2} - \frac{4}{5} = \frac{3}{10}.$$

Then

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{3/10}{1/2} = \frac{3}{5}.$$

Moreover,

$$P(A)P(B) = \frac{3}{5} \cdot \frac{1}{2} = \frac{3}{10} = P(A \cap B).$$

Thus $A$ and $B$ are independent.

**(b)**

$$\mathbb{E}[X] = (-2) \cdot \frac{1}{6} + 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{2} = -\frac{1}{3} + \frac{1}{2} = \frac{1}{6}.$$

Also,

$$\mathbb{E}[X^2] = 4 \cdot \frac{1}{6} + 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{2} = \frac{2}{3} + \frac{1}{2} = \frac{7}{6}.$$

Therefore,

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{7}{6} - \left(\frac{1}{6}\right)^2 = \frac{42}{36} - \frac{1}{36} = \frac{41}{36}.$$

**(c)** The first digit has 9 choices, since it cannot be 0. After that, the remaining three digits must be distinct and different from the first digit, so there are

$$9, \quad 8, \quad 7$$

choices for the remaining positions. Therefore, the number of passwords is

$$9 \cdot 9 \cdot 8 \cdot 7 = 4536.$$

**(d)**

$$P(X \le 1/2) = \int_0^{1/2} 2(1-x)\,dx = \left[2x - x^2\right]_0^{1/2} = 1 - \frac{1}{4} = \frac{3}{4}.$$

Also,

$$\mathbb{E}[X] = \int_0^1 x \cdot 2(1-x)\,dx = 2\int_0^1 (x - x^2)\,dx = 2\left(\frac{1}{2} - \frac{1}{3}\right) = \frac{1}{3}.$$

**(e)** The correlation coefficient is

$$\rho(X,Y) = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \frac{3}{\sqrt{4 \cdot 9}} = \frac{1}{2}.$$

Also,

$$\text{Var}(2X - Y) = 4\text{Var}(X) + \text{Var}(Y) + 2(2)(-1)\text{Cov}(X,Y).$$

Thus,

$$\text{Var}(2X - Y) = 4(4) + 9 - 4(3) = 16 + 9 - 12 = 13.$$

**(f)** Standardize:

$$P(85 \le X \le 130) = P\left(\frac{85-100}{15} \le Z \le \frac{130-100}{15}\right) = P(-1 \le Z \le 2).$$

Thus,

$$P(85 \le X \le 130) = \Phi(2) - \Phi(-1).$$

Using symmetry,

$$\Phi(-1) = 1 - \Phi(1),$$

so

$$P(85 \le X \le 130) = \Phi(2) + \Phi(1) - 1.$$

From the table, $\Phi(2) \approx 0.9772$ and $\Phi(1) \approx 0.8413$. Therefore,

$$P(85 \le X \le 130) \approx 0.9772 + 0.8413 - 1 = 0.8185.$$

## Problem 2. Hidden state, posterior weights, and prediction

**(a)** Given success probability $p$, the probability of exactly two successes in three independent alerts is

$$\binom{3}{2} p^2(1-p).$$

Thus,

$$\begin{aligned}
P(E \mid H = A) &= 3(0.8)^2(0.2) = 0.384 = \frac{48}{125}, \\
P(E \mid H = B) &= 3(0.5)^2(0.5) = 0.375 = \frac{3}{8}, \\
P(E \mid H = C) &= 3(0.2)^2(0.8) = 0.096 = \frac{12}{125}.
\end{aligned}$$

Since the three hidden states are equally likely,

$$\begin{aligned}
P(E) &= \frac{1}{3}\left(\frac{48}{125} + \frac{3}{8} + \frac{12}{125}\right) \\
&= \frac{1}{3}\left(\frac{60}{125} + \frac{3}{8}\right) = \frac{1}{3}\left(\frac{12}{25} + \frac{3}{8}\right) \\
&= \frac{1}{3} \cdot \frac{171}{200} = \frac{57}{200}.
\end{aligned}$$

**(b)** The unnormalized posterior weights are

$$w_h = P(E \mid H = h)P(H = h).$$

Thus,

$$\begin{aligned} w_A &= \frac{48}{125} \cdot \frac{1}{3} = \frac{16}{125}, \\ w_B &= \frac{3}{8} \cdot \frac{1}{3} = \frac{1}{8}, \\ w_C &= \frac{12}{125} \cdot \frac{1}{3} = \frac{4}{125}. \end{aligned}$$

Since

$$\frac{16}{125} = 0.128, \qquad \frac{1}{8} = 0.125, \qquad \frac{4}{125} = 0.032,$$

the posterior ranking is

$$A, \quad B, \quad C.$$

Thus $H = A$ is most likely after observing $E$, though $A$ and $B$ are very close.

**(c)** Given state $H = h$, the probability of exactly one success among the next two alerts is

$$2p_h(1 - p_h),$$

where $p_h$ is the success probability under state $h$. Therefore,

$$\begin{aligned} q_A &= 2(0.8)(0.2) = 0.32 = \frac{8}{25}, \\ q_B &= 2(0.5)(0.5) = 0.5 = \frac{1}{2}, \\ q_C &= 2(0.2)(0.8) = 0.32 = \frac{8}{25}. \end{aligned}$$

Using the posterior weights,

$$P(\text{exactly one success among next two} \mid E) = \frac{w_A q_A + w_B q_B + w_C q_C}{w_A + w_B + w_C}.$$

The denominator is

$$w_A + w_B + w_C = P(E) = \frac{57}{200}.$$

The numerator is

$$\frac{16}{125} \cdot \frac{8}{25} + \frac{1}{8} \cdot \frac{1}{2} + \frac{4}{125} \cdot \frac{8}{25} = \frac{128}{3125} + \frac{1}{16} + \frac{32}{3125}.$$

This equals

$$0.04096 + 0.0625 + 0.01024 = 0.1137.$$

Therefore,

$$P(\text{exactly one success among next two} \mid E) = \frac{0.1137}{0.285} \approx 0.399.$$

Equivalently, the exact value is

$$\frac{379}{950}.$$

## Problem 3. Conditional discrete model

**(a)** Using total probability,

$$P(X=0)=P(G=0)P(X=0\mid G=0)+P(G=2)P(X=0\mid G=2)=\frac{1}{2}\cdot\frac{1}{2}+\frac{1}{6}\cdot\frac{1}{2}=\frac{1}{3}.$$

Also,

$$\begin{aligned}
P(X=1)&=P(G=1)P(X=1\mid G=1)=\frac{1}{3},\\
P(X=2)&=P(G=0)P(X=2\mid G=0)=\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4},\\
P(X=4)&=P(G=2)P(X=4\mid G=2)=\frac{1}{6}\cdot\frac{1}{2}=\frac{1}{12}.
\end{aligned}$$

Thus,

$$p_X(x)=\begin{cases}1/3, & x=0,\\ 1/3, & x=1,\\ 1/4, & x=2,\\ 1/12, & x=4,\\ 0, & \text{otherwise.}\end{cases}$$

**(b)** Using the PMF from part (a),

$$\mathbb{E}[X]=0\cdot\frac{1}{3}+1\cdot\frac{1}{3}+2\cdot\frac{1}{4}+4\cdot\frac{1}{12}=\frac{1}{3}+\frac{1}{2}+\frac{1}{3}=\frac{7}{6}.$$

Next,

$$\mathbb{E}[X^2]=0^2\cdot\frac{1}{3}+1^2\cdot\frac{1}{3}+2^2\cdot\frac{1}{4}+4^2\cdot\frac{1}{12}=\frac{1}{3}+1+\frac{4}{3}=\frac{8}{3}.$$

Therefore,

$$\operatorname{Var}(X)=\mathbb{E}[X^2]-(\mathbb{E}[X])^2=\frac{8}{3}-\left(\frac{7}{6}\right)^2=\frac{96}{36}-\frac{49}{36}=\frac{47}{36}.$$

Alternatively, using total variance:

$$\begin{aligned}
\mathbb{E}[X\mid G=0]&=1, & \operatorname{Var}(X\mid G=0)&=1,\\
\mathbb{E}[X\mid G=1]&=1, & \operatorname{Var}(X\mid G=1)&=0,\\
\mathbb{E}[X\mid G=2]&=2, & \operatorname{Var}(X\mid G=2)&=4.
\end{aligned}$$

Then

$$\begin{aligned}
\mathbb{E}[\operatorname{Var}(X\mid G)]&=\frac{1}{2}(1)+\frac{1}{3}(0)+\frac{1}{6}(4)=\frac{7}{6},\\
\operatorname{Var}(\mathbb{E}[X\mid G])&=\operatorname{Var}(m(G)),
\end{aligned}$$

where

$$m(0)=1,\qquad m(1)=1,\qquad m(2)=2.$$

Thus,

$$\mathbb{E}[m(G)]=\frac{7}{6},\qquad\text{and}\qquad \mathbb{E}[m(G)^2]=1\cdot\frac{1}{2}+1\cdot\frac{1}{3}+4\cdot\frac{1}{6}=\frac{3}{2}.$$

Hence

$$\operatorname{Var}(m(G))=\frac{3}{2}-\left(\frac{7}{6}\right)^2=\frac{5}{36}.$$

Therefore,

$$\operatorname{Var}(X)=\frac{7}{6}+\frac{5}{36}=\frac{47}{36}.$$

**(c)** First,

$$\mathbb{E}[G] = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{3} + 2 \cdot \frac{1}{6} = \frac{2}{3}.$$

Since $\mathbb{E}[X \mid G = 0] = 1$, $\mathbb{E}[X \mid G = 1] = 1$, $\mathbb{E}[X \mid G = 2] = 2$, we get

$$\begin{aligned}\mathbb{E}[XG] &= \mathbb{E}[G\mathbb{E}[X \mid G]] \\ &= 0 \cdot 1 \cdot \frac{1}{2} + 1 \cdot 1 \cdot \frac{1}{3} + 2 \cdot 2 \cdot \frac{1}{6} = \frac{1}{3} + \frac{2}{3} = 1.\end{aligned}$$

Thus,

$$\mathrm{Cov}(X, G) = \mathbb{E}[XG] - \mathbb{E}[X]\mathbb{E}[G] = 1 - \frac{7}{6} \cdot \frac{2}{3} = 1 - \frac{7}{9} = \frac{2}{9}.$$

Now

$$\mathbb{E}[G^2] = 0^2 \cdot \frac{1}{2} + 1^2 \cdot \frac{1}{3} + 2^2 \cdot \frac{1}{6} = 1, \qquad \text{and hence,} \qquad \mathrm{Var}(G) = 1 - \left(\frac{2}{3}\right)^2 = \frac{5}{9}.$$

Therefore,

$$\rho(X, G) = \frac{\mathrm{Cov}(X, G)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(G)}} = \frac{2/9}{\sqrt{(47/36)(5/9)}} = \frac{4}{\sqrt{235}} \approx 0.261.$$

The random variables $X$ and $G$ are not independent. For example,

$$P(X = 4 \mid G = 2) = \frac{1}{2}, \qquad P(X = 4) = \frac{1}{12}.$$

Since conditioning on $G$ changes the distribution of $X$, $X$ and $G$ are not independent.

## Problem 4. Joint PDF and derived distribution

**(a)** The region is

$$R = \{(x, y) : 0 \le x \le 1,\ 0 \le y \le 1 + x\}.$$

For each $x \in [0, 1]$, the vertical length is $1 + x$, so the area is

$$\int_0^1 (1 + x)\,dx = \left[x + \frac{x^2}{2}\right]_0^1 = \frac{3}{2}.$$

Since the density is constant on $R$, we need

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

**(b)** For fixed $y$, the constraints are

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

Thus $x \ge y - 1$. Therefore,

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

The possible values of $y$ range from 0 to 2.

For $0 \le y \le 1$, the allowed $x$-interval is $[0, 1]$, so

$$f_Y(y) = \int_0^1 \frac{2}{3}\,dx = \frac{2}{3}.$$

For $1 < y \le 2$, the allowed $x$-interval is $[y - 1, 1]$, so

$$f_Y(y) = \int_{y-1}^1 \frac{2}{3}\,dx = \frac{2}{3}(2 - y).$$

Thus,

$$f_Y(y) = \begin{cases} \frac{2}{3}, & 0 \le y \le 1, \\ \frac{2}{3}(2-y), & 1 < y \le 2, \\ 0, & \text{otherwise.} \end{cases}$$

Since the joint density is constant, $X \mid Y = y$ is uniform on the allowable $x$-interval. Hence,

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

The corresponding conditional densities are

$$f_{X|Y}(x \mid y) = \begin{cases} 1, & 0 \le x \le 1, \quad 0 < y \le 1, \\ \dfrac{1}{2-y}, & y-1 \le x \le 1, \quad 1 < y < 2, \\ 0, & \text{otherwise.} \end{cases}$$

**(c)** Let

$$T = X + Y.$$

For a given $t$, set $y = t - x$. The constraints are

$$0 \le x \le 1, \qquad 0 \le t - x, \qquad t - x \le 1 + x.$$

Equivalently,

$$0 \le x \le 1, \qquad x \le t, \qquad x \ge \frac{t-1}{2}.$$

Thus the allowable $x$-interval has endpoints

$$\max\left(0, \frac{t-1}{2}\right) \quad \text{and} \quad \min(1, t).$$

Since the joint density is $2/3$, the PDF of $T$ is

$$f_T(t) = \frac{2}{3}\left[\min(1,t) - \max\left(0, \frac{t-1}{2}\right)\right],$$

when the bracket is positive.

Therefore,

$$f_T(t) = \begin{cases} \dfrac{2t}{3}, & 0 < t < 1, \\ \dfrac{3-t}{3}, & 1 \le t < 3, \\ 0, & \text{otherwise.} \end{cases}$$

Integrating this PDF gives the CDF:

$$F_T(t) = \begin{cases} 0, & t < 0, \\ \dfrac{t^2}{3}, & 0 \le t \le 1, \\ 1 - \dfrac{(3-t)^2}{6}, & 1 < t \le 3, \\ 1, & t > 3. \end{cases}$$

## Problem 5. Tail bounds, convergence, and CLT

**(a)** Using the given tail integral,

$$P(X \geq 5) = \int_5^\infty xe^{-x}\,dx = (5+1)e^{-5} = 6e^{-5} \approx 0.040.$$

Markov's inequality gives

$$P(X \geq 5) \leq \frac{\mathbb{E}[X]}{5} = \frac{2}{5} = 0.4.$$

For Chebyshev, since $\mathbb{E}[X] = 2$ and $X \geq 5$ implies $|X-2| \geq 3$,

$$P(X \geq 5) \leq P(|X-2| \geq 3) \leq \frac{\mathrm{Var}(X)}{3^2} = \frac{2}{9} \approx 0.222.$$

**(b)** Fix $\epsilon > 0$. Then

$$P(|Y_n| \geq \epsilon) = P(nU^n \geq \epsilon).$$

For $n > \epsilon$,

$$P(nU^n \geq \epsilon) = P\left(U \geq \left(\frac{\epsilon}{n}\right)^{1/n}\right) = 1 - \left(\frac{\epsilon}{n}\right)^{1/n}.$$

Using the hint,

$$\left(\frac{\epsilon}{n}\right)^{1/n} \to 1, \qquad \text{so} \qquad P(|Y_n| \geq \epsilon) \to 0.$$

Therefore,

$$Y_n \xrightarrow{P} 0.$$

Next,

$$\mathbb{E}[Y_n] = n\mathbb{E}[U^n] = n\int_0^1 u^n\,du = n \cdot \frac{1}{n+1} = \frac{n}{n+1}, \qquad \text{and thus,} \qquad \mathbb{E}[Y_n] \to 1.$$

So $\mathbb{E}[Y_n]$ does not converge to 0, even though $Y_n \to 0$ in probability.

**(c)** Let $W_i = U_i^2$. Then

$$\mathbb{E}[W_i] = \frac{1}{3}, \qquad \mathrm{Var}(W_i) = \frac{4}{45}.$$

Also,

$$A_{500} = \frac{1}{500}\sum_{i=1}^{500} W_i.$$

Therefore,

$$\mathbb{E}[A_{500}] = \frac{1}{3}, \qquad \text{and} \qquad \mathrm{Var}(A_{500}) = \frac{1}{500}\cdot\frac{4}{45} = \frac{4}{22500}.$$

Hence

$$\mathrm{SD}(A_{500}) = \sqrt{\frac{4}{22500}} = \frac{2}{150} = \frac{1}{75}.$$

By the CLT,

$$P\left(\left|A_{500} - \frac{1}{3}\right| \leq 0.02\right) \approx P\left(|Z| \leq \frac{0.02}{1/75}\right) = P(|Z| \leq 1.5).$$

Thus,

$$P\left(\left|A_{500} - \frac{1}{3}\right| \leq 0.02\right) \approx 2\Phi(1.5) - 1.$$

Using $\Phi(1.5) \approx 0.9332$,

$$2\Phi(1.5) - 1 \approx 2(0.9332) - 1 = 0.8664.$$

## Problem 6. Inverse problem and statistical estimation

**(a)** Since $Y = \Theta + \varepsilon$, and $\Theta$ and $\varepsilon$ are independent,

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

Here, $\mathbb{E}[\Theta] = 1$ and $\mathbb{E}[\varepsilon] = 0$. Therefore,

$$\mathbb{E}[Y] = 1.$$

Also,

$$\text{Var}(Y) = \text{Var}(\Theta) + \text{Var}(\varepsilon).$$

Since $\Theta \sim \text{Uniform}(0, 2)$ and $\varepsilon \sim \text{Uniform}(-1, 1)$, we have

$$\text{Var}(\Theta) = \frac{(2-0)^2}{12} = \frac{1}{3}, \qquad \text{and} \qquad \text{Var}(\varepsilon) = \frac{(1-(-1))^2}{12} = \frac{1}{3}.$$

Thus,

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

**(b)** We first compute

$$P(B \mid \Theta = \theta) = P(0 \le \theta + \varepsilon \le 1/2) = P(-\theta \le \varepsilon \le 1/2 - \theta).$$

Since $\varepsilon \sim \text{Uniform}(-1, 1)$, this probability equals half the length of the intersection

$$[-1, 1] \cap [-\theta, 1/2 - \theta].$$

- For $0 \le \theta \le 1$, the interval $[-\theta, 1/2 - \theta]$ lies inside $[-1, 1]$, so its length is $1/2$. Hence

$$P(B \mid \Theta = \theta) = \frac{1}{4}.$$

- For $1 < \theta \le 3/2$, the intersection is $[-1, 1/2 - \theta]$, whose length is $3/2 - \theta$. Hence

$$P(B \mid \Theta = \theta) = \frac{3/2 - \theta}{2}.$$

- For $3/2 < \theta \le 2$, the intersection is empty, so

$$P(B \mid \Theta = \theta) = 0.$$

Thus,

$$P(B \mid \Theta = \theta) = \begin{cases} \frac{1}{4}, & 0 \le \theta \le 1, \\ \frac{3/2-\theta}{2}, & 1 < \theta \le 3/2, \\ 0, & 3/2 < \theta \le 2. \end{cases}$$

By Bayes' rule for conditioning on an event,

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

Since $f_\Theta(\theta) = 1/2$ on $[0, 2]$, the prior density is constant. The normalizing constant is

$$P(B) = \int_0^2 \frac{1}{2} P(B \mid \Theta = \theta)\, d\theta.$$

Compute

$$\int_0^1 \frac{1}{4}\,d\theta = \frac{1}{4}, \qquad \text{and} \qquad \int_1^{3/2} \frac{3/2-\theta}{2}\,d\theta = \frac{1}{16}.$$

Therefore,

$$P(B) = \frac{1}{2}\left(\frac{1}{4}+\frac{1}{16}\right) = \frac{1}{2}\cdot\frac{5}{16} = \frac{5}{32}.$$

Hence,

$$f_{\Theta|B}(\theta) = \begin{cases} \frac{4}{5}, & 0 \le \theta \le 1, \\ \frac{8}{5}\left(\frac{3}{2}-\theta\right), & 1 < \theta \le 3/2, \\ 0, & \text{otherwise.} \end{cases}$$

Now compute the posterior mean:

$$\begin{aligned} \mathbb{E}[\Theta \mid B] &= \int_0^1 \theta\cdot\frac{4}{5}\,d\theta + \int_1^{3/2} \theta\cdot\frac{8}{5}\left(\frac{3}{2}-\theta\right)\,d\theta \\ &= \frac{4}{5}\cdot\frac{1}{2} + \frac{8}{5}\int_1^{3/2}\left(\frac{3}{2}\theta-\theta^2\right)d\theta \\ &= \frac{2}{5}+\frac{7}{30} = \frac{19}{30}. \end{aligned}$$

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

$$Y_i = \theta + \varepsilon_i,$$

where $\mathbb{E}[\varepsilon_i] = 0$ and $\mathrm{Var}(\varepsilon_i) = \frac{1}{3}$. Thus, $\mathbb{E}[Y_i \mid \Theta = \theta] = \theta$ and $\mathrm{Var}(Y_i \mid \Theta = \theta) = \frac{1}{3}$. Therefore,

$$\mathbb{E}[\bar{Y}_n \mid \Theta = \theta] = \theta, \qquad \mathrm{Var}(\bar{Y}_n \mid \Theta = \theta) = \frac{1}{3n}.$$

Using Chebyshev's inequality,

$$P(|\bar{Y}_n - \theta| \ge 0.1 \mid \Theta = \theta) \le \frac{1/(3n)}{0.1^2} = \frac{100}{3n}.$$

To guarantee

$$P(|\bar{Y}_n - \theta| \le 0.1 \mid \Theta = \theta) \ge 0.95,$$

it is enough to require

$$\frac{100}{3n} \le 0.05, \qquad \text{and thus,} \qquad n \ge \frac{100}{0.15} = \frac{2000}{3} \approx 666.67.$$

So the Chebyshev sufficient integer is

$$n = 667.$$

Using the CLT,

$$\bar{Y}_n \approx N\left(\theta, \frac{1}{3n}\right).$$

Thus,

$$P(|\bar{Y}_n - \theta| \le 0.1 \mid \Theta = \theta) \approx P\left(|Z| \le \frac{0.1}{\sqrt{1/(3n)}}\right) = P(|Z| \le 0.1\sqrt{3n}).$$

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

$$0.1\sqrt{3n} \ge 1.96, \qquad \text{and hence,} \qquad n \ge \frac{1.96^2}{3(0.1)^2} = \frac{3.8416}{0.03} \approx 128.05.$$

So the CLT-based approximate integer is

$$n = 129 < 667.$$