STA 130A: Mathematical Statistics: Brief Course

Midterm exam 2 solution

Instructor: Dogyoon Song

Problem 1 (40 points). Conceptual checks

(a) From correlation to the variance of a sum. First,

$$\mathrm{Cov}(X,Y) = \mathrm{Corr}(X,Y)\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)} = -\frac{1}{2}\sqrt{1\cdot 4} = -1.$$

Therefore

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y) = 1 + 4 - 2 = 3.$$

(b) A probability from a joint PDF. Inside the unit square, the event $U + V < 1/2$ is the triangle $0 < u < 1/2$, $0 < v < 1/2 - u$. Therefore,

$$P(U+V<1/2) = \int_0^{1/2}\int_0^{1/2-u} 1\, dv\, du = \int_0^{1/2}(1/2-u)\, du = \frac{1}{8}.$$

Equivalently, its probability is its area, $\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2} = \frac{1}{8}$, since the joint density is 1.

(c) Conditional variance. Given X, the value $Y = 3X - 1$ is fixed. Hence

$$\mathrm{Var}(Y \mid X) = 0 \qquad \text{with probability 1.}$$

(d) Variance from a differentiable MGF. Differentiating the given function,

$$M_X'(t) = 2(1-t)^{-3}, \qquad M_X''(t) = 6(1-t)^{-4}.$$

Thus $\mathbb{E}[X] = M_X'(0) = 2$ and $\mathbb{E}[X^2] = M_X''(0) = 6$, so

$$\mathrm{Var}(X) = M_X''(0) - \big(M_X'(0)\big)^2 = 6 - 2^2 = 2.$$

(e) A bound from limited information. Nonnegativity permits Markov's inequality:

$$P(W \geq 12) \leq \frac{\mathbb{E}[W]}{12} = \frac{3}{12} = \frac{1}{4}.$$

Optional sharpness check, not to be graded. This bound is optimal: a variable with $P(W = 12) = 1/4$ and $P(W = 0) = 3/4$ satisfies the assumptions and attains equality.

Problem 2 (30 points). Continuous distributions

The three parts use separate models; none requires an answer from another part.

(a) A marginal PDF. For fixed $0 < y < 2$, integrate over $0 < x < y$:

$$f_Y(y) = \begin{cases} \displaystyle\int_0^y \frac{1}{2}\,dx = \frac{y}{2}, & 0 < y < 2, \\ 0, & \text{otherwise.} \end{cases}$$

(b) A conditional PDF. For $0 < v < 1$, the normalizing marginal is $f_V(v) = \int_0^v 3v\,du = 3v^2$. Dividing the joint density by this marginal gives

$$f_{U|V}(u \mid v) = \begin{cases} \dfrac{1}{v}, & 0 < u < v, \\ 0, & \text{otherwise,} \end{cases} \qquad 0 < v < 1.$$

(c) A decreasing nonlinear transformation. The map $z = 1/(1+w)$ sends $0 < w < 1$ to $1/2 < z < 1$. Its inverse is $w = 1/z - 1$, with $|dw/dz| = 1/z^2$. Since $f_W(w) = 1$ on $(0,1)$,

$$f_Z(z) = \begin{cases} z^{-2}, & 1/2 < z < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Alternative: CDF-first approach. Since $0 < W < 1$, we have $1/2 < Z < 1$. For $1/2 < z < 1$, the transformation is decreasing, so

$$\frac{1}{1+W} \le z \quad \Longleftrightarrow \quad W \ge \frac{1}{z} - 1.$$

The threshold $1/z - 1$ lies in $(0,1)$. Since W is uniform on $(0,1)$,

$$F_Z(z) = P\left(W \ge \frac{1}{z} - 1\right) = 1 - \left(\frac{1}{z} - 1\right) = 2 - \frac{1}{z}.$$

Thus the complete CDF is

$$F_Z(z) = \begin{cases} 0, & z \le 1/2, \\ 2 - 1/z, & 1/2 < z < 1, \\ 1, & z \ge 1. \end{cases}$$

Differentiating on the interior of each interval gives a PDF:

$$f_Z(z) = \begin{cases} z^{-2}, & 1/2 < z < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Problem 3 (20 points). Independent sums and correlation

(a) MGF of the independent sum. The given PMF yields

$$M_R(t) = \mathbb{E}[e^{tR}] = \frac{3}{4} + \frac{1}{4}e^{2t}.$$

Using independence in the factorization,

$$\begin{aligned} M_S(t) &= e^t \mathbb{E}\big[e^{tR_1} e^{tR_2} e^{tR_3}\big] \\ &= e^t \prod_{i=1}^{3} \mathbb{E}[e^{tR_i}] \\ &= e^t \left(\frac{3}{4} + \frac{1}{4}e^{2t}\right)^3, \qquad t \in \mathbb{R}. \end{aligned}$$

The finite support makes the MGF finite for every real t.

(b) Correlation between a component and the total. Let $v = \mathrm{Var}(R) > 0$. Independence and the common distribution give

$$\mathrm{Cov}(R_1, S) = v, \qquad \mathrm{Var}(R_1) = v, \qquad \mathrm{Var}(S) = 3v.$$

The constant 1 in S affects neither covariance nor variance. Hence

$$\mathrm{Corr}(R_1, S) = \frac{v}{\sqrt{v \cdot 3v}} = \frac{1}{\sqrt{3}}.$$

It is equally valid to use the numerical moments $\mathbb{E}[R] = 1/2$, $\mathbb{E}[R^2] = 1$, and $v = 3/4$. This part does not require the MGF from part (a).

Problem 4 (30 points). Planning deliveries

(a) Overall delivery-time variance. The conditional means are 10 and 14 with equal probabilities, so

$$\mathbb{E}[T_1] = 12, \qquad \mathrm{Var}(\mathbb{E}[T_1 \mid G]) = \frac{1}{2}(10-12)^2 + \frac{1}{2}(14-12)^2 = 4.$$

The average within-route variance is

$$\mathbb{E}[\mathrm{Var}(T_1 \mid G)] = \frac{1}{2}(9) + \frac{1}{2}(4) = \frac{13}{2}.$$

The law of total variance therefore gives

$$\mathrm{Var}(T_1) = \mathbb{E}[\mathrm{Var}(T_1 \mid G)] + \mathrm{Var}(\mathbb{E}[T_1 \mid G]) = \frac{13}{2} + 4 = \frac{21}{2} = 10.5 \quad \text{(squared minutes)}.$$

Alternatively, the between-route contribution is $(14-10)^2/4 = 4$.

(b) Covariance of the on-time indicators. The marginal on-time probability is

$$\mathbb{E}[I_j] = P(T_j \le 12) = \frac{1}{2} \times \frac{3}{4} + \frac{1}{2} \times \frac{1}{4} = \frac{1}{2}, \qquad j = 1, 2.$$

Conditional independence applies within each route, so

$$\mathbb{E}[I_1 I_2] = P(I_1 = 1, I_2 = 1) = \frac{1}{2}\left(\frac{3}{4}\right)^2 + \frac{1}{2}\left(\frac{1}{4}\right)^2 = \frac{5}{16}.$$

Consequently,

$$\mathrm{Cov}(I_1, I_2) = \mathbb{E}[I_1 I_2] - \mathbb{E}[I_1]\mathbb{E}[I_2] = \frac{5}{16} - \frac{1}{4} = \frac{1}{16}.$$

The positive covariance comes from the shared route; conditional independence does not imply unconditional independence.

(c) Planning an independent route-B study. Each U_i has the route-B mean 14 and variance 4. Independence gives

$$\mathbb{E}[\overline{U}_n] = 14, \qquad \mathrm{Var}(\overline{U}_n) = \frac{4}{n}.$$

Applying Chebyshev at deviation threshold 2,

$$P\big(|\overline{U}_n - 14| \ge 2\big) \le \frac{4/n}{2^2} = \frac{1}{n}, \qquad P\big(|\overline{U}_n - 14| < 2\big) \ge 1 - \frac{1}{n}.$$

The requested certificate holds when $1 - 1/n \ge 0.95$, or $n \ge 20$. Thus the smallest positive integer certified by this direct bound is

$$n = 20.$$

The same bound follows by applying Markov to $(\overline{U}_n - 14)^2$. The value 20 is the smallest integer certified by the displayed bound, not a claim that every compatible distribution needs at least 20 observations.

Problem 5 (10 bonus points). Optional bonuses

(a) The one-sided tail conditions are impossible. No. The first two conditions imply $\mathbb{E}[X^2] = 1$. Let p_X be a proposed PMF. Then

$$1 = \sum_x x^2 p_X(x) \geq \sum_{x \geq 2} x^2 p_X(x) \geq 4 \sum_{x \geq 2} p_X(x) = 1.$$

Both inequalities must be equalities. The first forces every value below 2 with positive probability to be 0; the second forces every value at least 2 with positive probability to be 2. Thus the only possible PMF is

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

But this gives $\mathbb{E}[X] = 2(1/4) = 1/2$, contradicting $\mathbb{E}[X] = 0$. Therefore no such discrete random variable exists.

(b) Zero covariance with binary marginals. No. Because $X, Y \in \{0, 1\}$,

$$\mathbb{E}[XY] = P(X = 1, Y = 1).$$

Zero covariance forces

$$P(X = 1, Y = 1) = \mathbb{E}[X]\mathbb{E}[Y] = \frac{1}{4}.$$

The Bernoulli(1/2) marginals then force all other joint probabilities to equal 1/4:

$p_{X,Y}(x, y)$	$y = 0$	$y = 1$
$x = 0$	1/4	1/4
$x = 1$	1/4	1/4

The joint PMF factors into its marginals, so the variables must be independent.