STA 130A: Mathematical Statistics: Brief Course

Practice midterm exam 2 — Set B solutions

Instructor: Dogyoon Song

Problem 1. Conceptual checks

(a) Support and independence. No: both $P(X > 1/2)$ and $P(Y > 1/2)$ are positive, whereas $P(X > 1/2, Y > 1/2) = 0$ because the joint support requires $X + Y < 1$; this contradicts independence.

(b) Comparing Markov and Chebyshev. Markov gives $P(W \geq 6) \leq 2/6 = 1/3$. Since $W \geq 6$ implies $|W - 2| \geq 4$, Chebyshev gives

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

Answer: Chebyshev gives the smaller upper bound, $1/16$.

(c) Recovering correlation from a difference. Since both variances equal 1,

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

Thus $\mathrm{Cov}(X, Y) = 3/4$, and $\mathrm{Corr}(X, Y) = \mathrm{Cov}(X, Y)/\sqrt{1 \cdot 1} = 3/4$.

(d) A variance constraint from conditional means. No. By total variance and nonnegativity of conditional variance,

$$\mathrm{Var}(Y) = \mathbb{E}[\mathrm{Var}(Y \mid X)] + \mathrm{Var}(2X) \geq 4\mathrm{Var}(X) = 4 > 3.$$

Variation in the conditional mean alone already contributes variance 4.

(e) An MGF obtained by conditioning. Conditioning on the independent fair-coin outcome gives

$$M_Z(t) = \frac{1}{2}\mathbb{E}[e^{tX}] + \frac{1}{2}\mathbb{E}[e^{-tX}] = \frac{M_X(t) + M_X(-t)}{2},$$

for t sufficiently close to 0, so both terms are finite. The coin selects one of two outcomes; the conditional MGFs are averaged, not multiplied.

(f) What MGF uniqueness does and does not imply. No. Equal MGFs near zero imply that X and Y have the same distribution, not that they are equal with probability one. For example, take $X \sim \mathrm{Bernoulli}(1/2)$ and $Y = 1 - X$. Then

$$M_X(t) = M_Y(t) = \frac{1 + e^t}{2}, \qquad P(X = Y) = 0.$$

Problem 2. Continuous distributions

Each part uses its own probability model.

(a) Marginal density of X. For $0 < x < 1$,

$$f_X(x) = \int_0^1 (x + y)\, dy = x + \frac{1}{2}.$$

Therefore

$$f_X(x) = \begin{cases} x + \dfrac{1}{2}, & 0 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

(b) Conditional density from a separate joint model. For $0 < v < 1$, the allowable values of u are $0 < u < v$. The denominator is

$$f_V(v) = \int_0^v 6u\, du = 3v^2.$$

Consequently, for each $0 < v < 1$,

$$f_{U|V}(u \mid v) = \frac{f_{U,V}(u, v)}{f_V(v)} = \begin{cases} \dfrac{2u}{v^2}, & 0 < u < v, \\ 0, & \text{otherwise.} \end{cases}$$

This density integrates to 1 over $(0, v)$; both its formula and its support depend on v.

(c) A nonmonotone transformation of a separate variable. For $W \sim \text{Uniform}(-2, 2)$ and $0 < z < 4$,

$$F_Z(z) = P(W^2 \le z) = P(-\sqrt{z} \le W \le \sqrt{z}) = \frac{2\sqrt{z}}{4} = \frac{\sqrt{z}}{2}.$$

Also $F_Z(z) = 0$ for $z \le 0$ and $F_Z(z) = 1$ for $z \ge 4$. Differentiating gives

$$f_Z(z) = \begin{cases} \dfrac{1}{4\sqrt{z}}, & 0 < z < 4, \\ 0, & \text{otherwise.} \end{cases}$$

Both the negative and positive values of W contribute; retaining only one branch would give half the correct density.

Problem 3. MGFs and covariance of linear combinations

(a) Distribution of $H = 2N_1 - N_2$ using MGFs. Using independence and the supplied normal MGF,

$$\begin{aligned} M_H(t) &= M_{N_1}(2t) M_{N_2}(-t) \\ &= \exp\left(2t + \frac{4(2t)^2}{2}\right) \exp\left(t + \frac{(-t)^2}{2}\right) \\ &= \exp\left(3t + \frac{17t^2}{2}\right), \qquad t \in \mathbb{R}. \end{aligned}$$

The product in the first line uses independence. This is the MGF of a normal variable with mean 3 and variance 17. By uniqueness of MGFs finite near zero,

$$H \sim N(3, 17).$$

(b) Covariance of two overlapping combinations. By bilinearity and $\mathrm{Cov}(N_1, N_2) = 0$,

$$\begin{aligned}\mathrm{Cov}(N_1 + N_2, H) &= \mathrm{Cov}(N_1 + N_2, 2N_1 - N_2) \\ &= 2\mathrm{Var}(N_1) - \mathrm{Cov}(N_1, N_2) + 2\mathrm{Cov}(N_2, N_1) - \mathrm{Var}(N_2) \\ &= 2(4) - 1 = 7.\end{aligned}$$

This calculation uses the given component variances, not the answer to part (a).

(c) A sum of independent Poisson variables. Independence gives, for every real t,

$$M_{A+B}(t) = M_A(t)M_B(t) = e^{2(e^t-1)}e^{3(e^t-1)} = e^{5(e^t-1)}.$$

This is the MGF of a Poisson variable with parameter 5. By MGF uniqueness,

$$A + B \sim \text{Poisson}(5).$$

Problem 4. A shared calibration error

Here $W_i = \theta + B + \varepsilon_i$, with a fixed θ, a common offset B, and mutually independent $B, \varepsilon_1, \varepsilon_2, \ldots$.

(a) Correlation of two readings. Constants do not affect covariance, and mutual independence removes the cross terms:

$$\mathrm{Cov}(W_1, W_2) = \mathrm{Cov}(B + \varepsilon_1, B + \varepsilon_2) = \mathrm{Var}(B) = 4.$$

For each i, $\mathrm{Var}(W_i) = \mathrm{Var}(B) + \mathrm{Var}(\varepsilon_i) = 4 + 1 = 5$. Therefore

$$\mathrm{Corr}(W_1, W_2) = \frac{4}{\sqrt{5 \cdot 5}} = \frac{4}{5}.$$

The shared offset creates positive correlation despite independence of the individual errors.

(b) Variance of the average by conditioning. Let $\bar{\varepsilon}_n = n^{-1}\sum_{i=1}^n \varepsilon_i$, so that $\overline{W}_n = \theta + B + \bar{\varepsilon}_n$. Conditional on B,

$$\mathbb{E}[\overline{W}_n \mid B] = \theta + B, \qquad \mathrm{Var}(\overline{W}_n \mid B) = \frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}(\varepsilon_i) = \frac{1}{n}.$$

Total variance therefore gives

$$\mathrm{Var}(\overline{W}_n) = \mathbb{E}[\mathrm{Var}(\overline{W}_n \mid B)] + \mathrm{Var}(\mathbb{E}[\overline{W}_n \mid B]) = \frac{1}{n} + 4.$$

Averaging reduces the individual noise but not the shared random offset.

(c) What Chebyshev can certify. First, $\mathbb{E}[\overline{W}_n] = \theta$. Thus Chebyshev gives

$$P\big(|\overline{W}_n - \theta| \geq 4\big) \leq \frac{4 + 1/n}{16}, \qquad P\big(|\overline{W}_n - \theta| < 4\big) \geq \frac{3}{4} - \frac{1}{16n}.$$

Interpretation: This lower bound is below 3/4 for every positive integer n, so it never certifies probability 0.95; it does *not* imply that the actual probability is below 0.95.

Problem 5. Optional bonuses

(a) The smallest common pairwise correlation. Each X_i has mean 0 and variance 1, so $\mathbb{E}[X_iX_j] = \rho$ for $i \neq j$. Let $S = X_1 + X_2 + X_3$. It is a sum of three signs, hence an odd integer; therefore $S^2 \geq 1$. Consequently,

$$1 \leq \mathbb{E}[S^2] = \sum_{i=1}^{3} \mathbb{E}[X_i^2] + 2\sum_{i<j} \mathbb{E}[X_iX_j] = 3 + 6\rho, \qquad \text{so} \qquad \rho \geq -\frac{1}{3}.$$

Define the joint PMF by

$$p(x_1, x_2, x_3) = \begin{cases} 1/6, & x_i \in \{-1, 1\} \text{ for all } i, \quad |x_1 + x_2 + x_3| = 1, \\ 0, & \text{otherwise.} \end{cases}$$

There are six such triples. In each coordinate, three have sign 1 and three have sign -1, so the marginals are correct. For every pair of coordinates, two triples have equal signs and four have opposite signs. Thus

$$\mathrm{Corr}(X_i, X_j) = \mathbb{E}[X_iX_j] = \frac{2}{6} - \frac{4}{6} = -\frac{1}{3}.$$

The bound is attained, and the smallest possible common correlation is $-1/3$.

(b) Two conditional-mean identities force equality. Yes. Finite second moments ensure that all products below are integrable. Conditioning in the two directions gives

$$\mathbb{E}[XY] = \mathbb{E}[X\mathbb{E}[Y \mid X]] = \mathbb{E}[X^2], \qquad \mathbb{E}[XY] = \mathbb{E}[Y\mathbb{E}[X \mid Y]] = \mathbb{E}[Y^2].$$

Therefore

$$\mathbb{E}[(X - Y)^2] = \mathbb{E}[X^2] + \mathbb{E}[Y^2] - 2\mathbb{E}[XY] = 0.$$

The nonnegative variable $(X - Y)^2$ must be zero with probability 1, so $P(X = Y) = 1$.