STA 130A: Mathematical Statistics: Brief Course

Practice final A solutions

Instructor: Dogyoon Song

Problem 1 (60 points). Fundamentals

(a) Identifying the jar after a red draw. Let R denote a red draw. The two ways to draw red have probabilities $P(A \cap R) = \frac{2}{3} \times \frac{1}{4} = 1/6$ and $P(B \cap R) = \frac{1}{3} \times \frac{3}{4} = 1/4$. Thus $P(B \mid R) = (1/4)/(1/6+1/4) = 3/5$.

(b) Remaining time to the first success. Given $N > 4$, the first four trials failed. Waiting exactly k more trials means $k-1$ further failures followed by a success. By independence, $P(N-4 = k \mid N > 4) = 0.6^{k-1} \times 0.4$ for $k = 1, 2, \ldots$, and the conditional PMF is zero otherwise.

(c) A normal percentile. For the 90th percentile x, standardization gives $\Phi((x-50)/5) = 0.90$. Hence $x = 50 + 5 \times \Phi^{-1}(0.90) \approx 50 + 5 \times 1.282 = 56.41$.

(d) Variance of a difference. The covariance is $\mathrm{Cov}(X, Y) = (-1/2) \times 2 \times 3 = -3$. Thus $\mathrm{Var}(X - Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) - 2\mathrm{Cov}(X, Y) = 4 + 9 - 2 \times (-3) = 19$.

(e) Variance from an MGF. Since $M_X'(t) = e^{2t}/2$ and $M_X''(t) = e^{2t}$, evaluating at $t = 0$ gives $\mathbb{E}[X] = 1/2$ and $\mathbb{E}[X^2] = 1$. Thus $\mathrm{Var}(X) = 1 - (1/2)^2 = 3/4$.

(f) A sharp upper bound on the delay probability. By Markov's inequality, $P(D \geq 24) \leq \mathbb{E}[D]/24 = 1/8$. This bound is attained when $P(D = 24) = 1/8$ and $P(D = 0) = 7/8$, since then $\mathbb{E}[D] = 3$ and $P(D \geq 24) = 1/8$.

(g) Sample size for a target standard error. Independence gives $\mathrm{SE}(\bar{X}_n) = 4/\sqrt{n}$. The condition $4/\sqrt{n} \leq 1.25$ gives $n \geq (4/1.25)^2 = 10.24$, so the smallest sample size is 11.

(h) An interval probability for a sample mean. Independence and normality give $\bar{X} \sim N(12, 3^2/9) = N(12, 1)$. Thus

$$P(11.25 < \bar{X} < 13.40) = \Phi(1.40) - \Phi(-0.75) \approx 0.9192 - (1 - 0.7734) = 0.6926 \approx 0.693.$$

(i) The sample mean and sample variance. Since $\bar{X} \sim N(\mu, \sigma^2/13)$, the cutoff is one standard error above μ. Thus the probability is $1 - \Phi(1) \approx 0.159$. Knowing the sample variance does not change this probability because $\bar{X}$ and S^2 are independent in a normal sample.

(j) Actual coverage for a future reading. The future reading is independent of $\bar{X}_n$, so $X_{n+1} - \bar{X}_n \sim N(0, \sigma^2(1 + 1/n))$. Standardizing the event $|X_{n+1} - \bar{X}_n| \leq 1.96\sigma/\sqrt{n}$ gives

$$P(X_{n+1} \in I_n) = 2\Phi\left(\frac{1.96}{\sqrt{n+1}}\right) - 1.$$

As n grows, coverage tends to zero because the interval shrinks while future readings retain variance σ^2.

(k) Estimating the population variance. Substitute the sample averages for the corresponding expectations:

$$\widehat{\sigma}^2_{\mathrm{MM}} = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2 = 11 - 3^2 = 2.$$

(l) Likelihood ratio for an observed sequence. The observed sequence has three successes and one failure, so independence gives $L(p) = p^3(1-p)$. Thus

$$\frac{L(3/4)}{L(1/2)} = \frac{(3/4)^3 \times (1/4)}{(1/2)^4} = \frac{27}{16}.$$

Problem 2 (20 points). An error count and an audit flag

(a) Marginal distribution of the error count. Sum each column over the two possible flag values. The three column sums are $(3+1)/12$, $(2+2)/12$ and $(1+3)/12$, respectively. Thus $P(X = x) = 1/3$ for $x = 0, 1, 2$ and zero otherwise.

(b) A change in the proportion of flagged files. Each original row has mass $6/12$. Normalizing within rows gives

$$\mathbb{E}[X \mid Y = 0] = \frac{0\times 3 + 1\times 2 + 2\times 1}{6} = \frac{2}{3}, \qquad \mathbb{E}[X \mid Y = 1] = \frac{0\times 1 + 1\times 2 + 2\times 3}{6} = \frac{4}{3}.$$

The follow-up audit keeps these conditional means but uses weights $1/4$ and $3/4$. Its expected count is therefore $\frac{1}{4}\times\frac{2}{3} + \frac{3}{4}\times\frac{4}{3} = 7/6$. This exceeds the original mean of 1 because the follow-up audit samples a larger fraction of files from the higher-mean group.

Problem 3 (20 points). Continuous distributions

(a) A conditional slice through the joint density. Fix $0 < x < 1$. The possible values of Y satisfy $x < y < 2$. Dividing the joint density on this slice by the given marginal gives

$$f_{Y|X}(y \mid x) = \begin{cases} \dfrac{2/3}{(2/3)\times(2-x)} = \dfrac{1}{2-x}, & x < y < 2, \\ 0, & \text{otherwise.} \end{cases}$$

The conditional density integrates to one because its interval has length $2 - x$.

(b) The absolute value of an asymmetric uniform variable. For $z \geq 0$, $F_Z(z) = P(-z \leq W \leq z)$ is the length of $[-z, z] \cap (-1, 2)$ divided by 3. That length is $2z$ for $0 \leq z \leq 1$ and $1 + z$ for $1 < z < 2$. The CDF is zero for $z \leq 0$ and one for $z \geq 2$. Differentiation gives a density

$$f_Z(z) = \begin{cases} 2/3, & 0 < z < 1, \\ 1/3, & 1 < z < 2, \\ 0, & \text{otherwise.} \end{cases}$$

For $0 < z < 1$, both z and $-z$ lie in the support of W. For $1 < z < 2$, only z does. The density values at 0, 1, and 2 do not affect probabilities.

Problem 4 (20 points). Sampling distributions and confidence intervals

(a) Normal approximation for a sum. In both studies, $\mathbb{E}[S] = 64 \times 3 = 192$ and $\mathrm{SD}(S) = \sqrt{64} \times 2 = 16$.
Study A. For continuous measurements, use the boundary 208. The standardized cutoff is $z_A = (208 - 192)/16 = 1$.
Study B. The event $S \geq 208$ includes the count 208 and excludes 207. The continuity-corrected boundary is halfway between them, at 207.5. Thus $z_B = (207.5 - 192)/16 = 31/32 \approx 0.969$.

(b) A mean interval and a feasible prediction width. (i) With σ unknown, $(\bar{X} - \mu)/(S/\sqrt{9}) \sim t_8$. The exact equal-tailed 95% interval for μ is $\bar{x} \pm t_8^* s/\sqrt{9}$. Using the supplied critical value gives $11.4 \pm 2.306 \times (2/3) \approx (9.863, 12.937)$.
(ii) By independence, $\bar{Y}_m - \bar{X}_n \sim N(0, 4/m + 4/n)$. For $w \geq 0$, the coverage $2\Phi\left(w/\sqrt{4/m + 4/n}\right) - 1$ increases strictly with w. Setting it equal to 0.95 gives

$$w_{n,m} = 2z^* \sqrt{\frac{1}{n} + \frac{1}{m}}, \qquad z^* = \Phi^{-1}(0.975) \approx 1.960.$$

For fixed n, the half-width decreases to $2z^*/\sqrt{n}$ from above as m increases. Thus some finite m gives $w_{n,m} \leq 1$ exactly when $2z^*/\sqrt{n} < 1$, or $n > (2z^*)^2 \approx 15.366$. The smallest such n is 16.

Problem 5 (20 points). Combining two instruments

(a) Maximum likelihood for a common mean. Differentiating the given log likelihood yields $\ell'(\mu) = (x - \mu) + (y - \mu)/4$. Setting this to zero gives $\hat{\mu} = (4x + y)/5$. Since $\ell''(\mu) = -5/4 < 0$ for all $\mu \in \mathbb{R}$, the stationary point is the unique global maximum. The estimator is $(4X + Y)/5$, with more weight on the more precise instrument.

(b) Choosing an estimator under bounded calibration error. (i) Since $A_m = (4\bar{X} + \bar{Y})/5$, its mean is $\mu + \delta/5$. Independence within and across pairs gives $\mathrm{Var}(A_m) = [16/m + 4/m]/25 = 4/(5m)$. The first-instrument average is unbiased. Therefore

$$\mathrm{MSE}(A_m) = \frac{4}{5m} + \frac{\delta^2}{25}, \qquad \mathrm{MSE}(\bar{X}) = \frac{1}{m}.$$

(ii) The worst MSE of A_m occurs at $\delta = \pm d$, whereas the MSE of $\bar{X}$ is constant in δ. The first worst-case risk is smaller precisely when

$$\frac{4}{5m} + \frac{d^2}{25} < \frac{1}{m} \quad \Longleftrightarrow \quad md^2 < 5.$$

Choose A_m for $md^2 < 5$, choose $\bar{X}$ for $md^2 > 5$, and either at equality. As m grows the variance saving $1/(5m)$ shrinks, while the largest squared bias stays at $d^2/25$.

Problem 6 (40 points). An anonymous survey

(a) The probability of a yes answer. A truthful respondent answers yes with probability p, and a respondent assigned to the coin answers yes with probability 1/2. Weighting the two branches gives $q = P(Y = 1) = \frac{2}{3} \times p + \frac{1}{3} \times \frac{1}{2} = 2p/3 + 1/6$.

(b) Conditional response means and variances. When $T = 0$, a yes requires the coin rule and heads, so $\mathbb{E}[Y \mid T = 0] = \frac{1}{3} \times \frac{1}{2} = 1/6$. When $T = 1$, a truthful response also gives a yes, so $\mathbb{E}[Y \mid T = 1] = 2/3 + 1/6 = 5/6$. Since $Y^2 = Y$, a conditional mean r gives conditional variance $r - r^2$. In either group the variance is $\frac{1}{6} \times \frac{5}{6} = 5/36$.

(c) Moment estimation and squared error. (i) Setting $\bar{Y} = 2p/3 + 1/6$ gives $\hat{p} = (3/2)\bar{Y} - 1/4$. Its bias is zero because $\mathbb{E}[\bar{Y}] = q = 2p/3 + 1/6$. Each Y is Bernoulli(q), so independence gives

$$\mathrm{Var}(\hat{p}) = \frac{9q(1-q)}{4n} = \frac{p(1-p) + 5/16}{n}.$$

Equivalently, part (b) gives $\mathrm{Var}(Y) = 5/36 + (4/9)p(1-p)$ by the law of total variance.
(ii) Yes. When $\hat{p}$ is outside $[0,1]$, its nearest endpoint is closer to the true p. For estimates in the interval, $\tilde{p} = \hat{p}$. Hence $(\tilde{p} - p)^2 \le (\hat{p} - p)^2$ on every sample. With probability $(1-q)^n > 0$, all responses are no, so $\hat{p} = -1/4$ and $\tilde{p} = 0$. The inequality is strict on this event. Taking expectations gives a strictly smaller MSE.

(d) A sample size valid across prevalences. Write $v = p(1-p)$ and $b = 5/16$ from (c)(i). The direct proportion has variance $v/64$, and the anonymous estimator has variance $(v+b)/n$. Comparing standard errors is equivalent to requiring

$$\frac{v+b}{n} \le \frac{v}{64} \quad \Longleftrightarrow \quad n \ge 64 \times \left(1 + \frac{b}{v}\right).$$

The right side is largest when $v = 1/4 - (p - 1/2)^2$ is smallest, at either endpoint $p = 1/4, 3/4$. There $v = 3/16$, so $n_{\min} = \lceil 64 \times (1 + 5/3) \rceil = \lceil 512/3 \rceil = 171$.

Problem 7 (bonus, 10 points)

(a) Centering a square to cancel covariance. By symmetry, $\mathbb{E}[X] = \mathbb{E}[X^3] = 0$, and $\mathbb{E}[X^2] = \int_{-1}^{1} x^2/2 \, dx = 1/3$. Hence

$$\mathrm{Cov}\big(X, (X-a)^2\big) = \mathbb{E}[X^3 - 2aX^2 + a^2 X] = -2a/3.$$

Only $a = 0$ works. Then $\{|X| < 1/2\}$ and $\{Y < 1/4\}$ are the same event, of probability $1/2$. Their intersection has probability $1/2$ while the product of their probabilities is $1/4$, proving dependence.

(b) Estimating the probability of two successes. Average the pair product over all ordered pairs of distinct observations. Exactly $K(K-1)$ of these pairs contain two successes, so the estimator is

$$V_n = \frac{1}{n(n-1)} \sum_{i \ne j} X_i X_j = \frac{K(K-1)}{n(n-1)}.$$

Unbiasedness. Distinct observations are independent, so each product has mean p^2. Linearity of expectation applies even to products that share an observation, giving $\mathbb{E}[V_n] = p^2$.
Consistency. Put $\hat{p} = K/n$. Removing the diagonal pairs gives the correction

$$V_n = \hat{p}^{\,2} - \frac{\hat{p}(1-\hat{p})}{n-1}.$$

Thus $0 \le \hat{p}^{\,2} - V_n \le 1/[4 \times (n-1)]$. The weak law gives $\hat{p} \to p$ in probability, so $V_n \to p^2$ in probability.