MA206X: Probability & Statistics

Comprehensive TEE Review — Answer Key
AY26-2

How to use this review. This review covers everything on the TEE: probability rules, named distributions, one- and two-sample inference, regression, and ANOVA. Work each problem using only your reference card, calculator, and R-lite (`pnorm`/`qnorm`, `pt`/`qt`, `pbinom`, `ppois`, `pexp`, `pf`/`qf`). Solutions are shown in blue.

1 Multiple Choice

1. In hypothesis testing, the p-value is the probability of:
   (a) The null hypothesis being true given the data
   (b) Observing data at least as extreme as what was observed, assuming H_0 is true
   (c) Making a Type I error
   (d) The alternative hypothesis being true

   B. The p-value is $P(\text{test stat as or more extreme} \mid H_0)$.

2. Two events A and B are mutually exclusive. Which statement must be true?
   (a) A and B are independent
   (b) $P(A \cap B) = P(A) \cdot P(B)$
   (c) $P(A \cup B) = P(A) + P(B)$
   (d) $P(A|B) = P(A)$

   C. Mutually exclusive means $P(A \cap B) = 0$, so $P(A \cup B) = P(A) + P(B)$. Mutually exclusive events with positive probability are NOT independent.

3. For a random variable X that follows a Poisson distribution with mean μ, the variance equals:
   (a) μ^2
   (b) $\sqrt{\mu}$
   (c) μ
   (d) $1/\mu$

   C. For Poisson, $E(X) = V(X) = \mu$.

4. The Central Limit Theorem says that the sampling distribution of the sample mean $\bar{X}$ is approximately normal when:
   (a) The population is normal regardless of n
   (b) The sample size n is sufficiently large
   (c) Either (a) or (b)
   (d) The population variance is known

C. Either condition guarantees approximate normality of $\bar{X}$.

5. A 95% confidence interval for μ is reported as $(42, 58)$. Which interpretation is correct?
   (a) There is a 95% probability that the true mean lies in $(42, 58)$.
   (b) 95% of all sample means fall in $(42, 58)$.
   (c) If we repeated the sampling procedure many times, about 95% of the resulting intervals would contain the true mean.
   (d) 95% of the population values fall in $(42, 58)$.

   C. The frequentist interpretation: it's the procedure that has a 95% success rate, not the specific interval.

6. In a multiple linear regression, the coefficient on x_1 is interpreted as:
   (a) The total change in y for a one-unit increase in x_1
   (b) The average change in y per one-unit increase in x_1, holding all other predictors constant
   (c) The correlation between x_1 and y
   (d) The proportion of variability in y explained by x_1

   B. The phrase "holding other predictors constant" is essential to the partial-effect interpretation.

7. Failing to reject H_0 at significance level $\alpha = 0.05$ means:
   (a) H_0 is true.
   (b) There is insufficient evidence to reject H_0 at the 0.05 level.
   (c) The probability that H_0 is true exceeds 0.95.
   (d) The Type I error rate is below 5%.

   B. "Absence of evidence is not evidence of absence." We never accept or prove H_0; we only fail to reject.

8. Which distribution would be most appropriate for modeling the time until the next email arrives in your inbox, assuming emails arrive at a constant average rate?
   (a) Binomial
   (b) Poisson
   (c) Exponential
   (d) Normal

   C. Exponential models time *between* events in a Poisson process.

9. The standard error of the sample mean is:
   (a) σ
   (b) σ^2/n

(c) $\sigma/\sqrt{n}$
(d) s

C. $SE(\bar{X}) = \sigma/\sqrt{n}$ (estimated as $s/\sqrt{n}$).

10. In a one-way ANOVA with k groups and N total observations, the test statistic F follows an F distribution with degrees of freedom:

(a) $(N-1, k-1)$
(b) $(k-1, N-k)$
(c) (k, N)
(d) $(N-k, k-1)$

B. Numerator df $= k-1$ (treatment), denominator df $= N-k$ (error).

11. Which of the following will increase the power of a hypothesis test? *Select all that apply*, then choose the best answer.

(a) Increasing the sample size n
(b) Increasing the significance level α
(c) Increasing the effect size
(d) All of the above

D. Power $= 1-\beta$ increases with n, α, and effect size.

12. For a continuous random variable X, $P(X = c)$ for any specific value c equals:

(a) Approximately 0 for large n
(b) $f(c)$, the density at c
(c) Exactly 0
(d) It depends on the distribution

C. For continuous RVs, the probability of any single point is exactly zero. The PDF $f(c)$ is a density, not a probability.

2 Event Probability (Intersections, Unions, Complements)

Problem 2.1. A standard urn contains 8 red, 5 blue, and 7 green marbles. One marble is drawn at random.

(a) What is the probability of drawing a red OR blue marble?

(b) What is the probability the marble is NOT green?

(c) What is the probability of drawing a marble that is both red AND blue?

Setup: Total $= 8 + 5 + 7 = 20$. Each color is mutually exclusive.
(a) $P(R \cup B) = P(R) + P(B) = \frac{8}{20} + \frac{5}{20} = \frac{13}{20} = 0.650$. (Mutually exclusive, so no overlap subtracted.)
(b) By complement: $P(\text{not green}) = 1 - P(G) = 1 - \frac{7}{20} = \frac{13}{20} = 0.650$.
(c) $P(R \cap B) = 0$ since one marble cannot simultaneously be two colors — the events are mutually exclusive.

Problem 2.2. A retailer surveys 500 customers about their shopping channels. The survey finds:

- 320 customers use the mobile app
- 280 customers use the website
- 180 customers use BOTH the mobile app AND the website

(a) What is the probability a randomly selected customer uses at least one of the two channels?

(b) What is the probability a customer uses neither channel?

(c) What is the probability a customer uses ONLY the mobile app (not the website)?

Setup: Let M = uses mobile, W = uses website. Then $P(M) = 320/500 = 0.64$, $P(W) = 280/500 = 0.56$, $P(M \cap W) = 180/500 = 0.36$.
(a) By the addition rule:

$$P(M \cup W) = P(M) + P(W) - P(M \cap W) = 0.64 + 0.56 - 0.36 = \boxed{0.84}$$

(b) $P(\text{neither}) = 1 - P(M \cup W) = 1 - 0.84 = \boxed{0.16}$.
(c) "Mobile only" means M but not W:

$$P(M \cap W^c) = P(M) - P(M \cap W) = 0.64 - 0.36 = \boxed{0.28}$$

3 Conditional Probability (Bayes and Non-Bayes)

Problem 3.1. A medical lab has developed a new screening test for a relatively rare disease. Background information:

- Approximately 2% of the general population has the disease.
- Among those who have the disease, the test correctly returns a positive result 96% of the time (sensitivity).
- Among those who do NOT have the disease, the test correctly returns a negative result 95% of the time (specificity).

(a) What is the probability that a randomly selected person tests positive?

(b) Given that a person has tested positive, what is the probability they actually have the disease?

(c) Briefly explain why your answer to (b) is so much smaller than the 96% sensitivity figure might suggest.

Setup: Let D = has disease, D^c = no disease, T^+ = positive test. $P(D) = 0.02$, $P(D^c) = 0.98$, $P(T^+|D) = 0.96$, $P(T^+|D^c) = 1 - 0.95 = 0.05$.

(a) Law of Total Probability:

$$P(T^+) = P(T^+|D)P(D) + P(T^+|D^c)P(D^c) = (0.96)(0.02) + (0.05)(0.98)$$
$$= 0.0192 + 0.049 = \boxed{0.0682}$$

(b) Bayes' Theorem:

$$P(D|T^+) = \frac{P(T^+|D)P(D)}{P(T^+)} = \frac{(0.96)(0.02)}{0.0682} = \frac{0.0192}{0.0682} = \boxed{0.282}$$

(c) The base rate of the disease is very low (2%). Even though the test rarely flags healthy people (5% false positive rate), there are SO many more healthy people than sick ones that the false positives outnumber the true positives. This is the classic **base-rate fallacy**: high sensitivity does not guarantee that a positive result is reliable when the prior probability is small.

Problem 3.2. Two cards are drawn at random from a standard 52-card deck *without replacement*.

(a) Given that the first card drawn is a King, what is the probability the second card is also a King?

(b) What is the probability that BOTH cards are Kings?

(c) What is the probability that BOTH cards are red (hearts or diamonds)?

Setup: A standard deck has 52 cards: 4 Kings (one in each suit), 26 red cards (hearts + diamonds), 26 black.

(a) After removing one King, 3 Kings remain in 51 cards:

$$P(\text{2nd K} \mid \text{1st K}) = \frac{3}{51} = \boxed{0.0588}$$

(b) Use the multiplication rule for conditional probability:

$$P(\text{both K}) = P(\text{1st K}) \cdot P(\text{2nd K}|\text{1st K}) = \frac{4}{52} \cdot \frac{3}{51} = \frac{12}{2652} = \boxed{0.00452}$$

(c) Same logic with 26 red cards:

$$P(\text{both red}) = \frac{26}{52} \cdot \frac{25}{51} = \frac{650}{2652} = \boxed{0.245}$$

Note: drawing without replacement makes the events *dependent* — this is why we used conditional probability rather than simply $(26/52)^2$.

4 Binomial Distribution

Problem 4.1. A factory produces electronic widgets, and historical data show that 5% of widgets are defective. A quality-control inspector randomly selects $n = 20$ widgets from the production line. Let X be the number of defective widgets in the sample.

(a) Identify the distribution of X and explain why the binomial assumptions are reasonable here.

(b) What is the probability that NONE of the 20 widgets are defective?

(c) What is the probability that at most 2 widgets are defective?

(d) What is the probability that 3 or more widgets are defective?

(a) $X \sim \text{Bin}(n = 20, p = 0.05)$.
Conditions: (1) Fixed number of trials ($n = 20$); (2) Two outcomes per trial (defective / not); (3) Constant probability ($p = 0.05$); (4) Approximately independent trials — reasonable when sampling from a large production batch.
(b) $P(X = 0) = \binom{20}{0}(0.05)^0(0.95)^{20} = (0.95)^{20} = \boxed{0.358}$.
R-lite: `dbinom(0, 20, 0.05)` = 0.358.
(c) $P(X \leq 2) = \texttt{pbinom(2, 20, 0.05)} = \boxed{0.925}$.
(d) $P(X \geq 3) = 1 - P(X \leq 2) = 1 - 0.925 = \boxed{0.0755}$.
R-lite: `1 - pbinom(2, 20, 0.05)` = 0.0755.

Problem 4.2. A college basketball player is an 80% free-throw shooter. In an upcoming game she is expected to attempt 10 free throws. Let X be the number of free throws she makes.

(a) Identify the distribution of X, including parameter values.

(b) Compute $E[X]$ and $V[X]$.

(c) What is the probability she makes exactly 8 free throws?

(d) What is the probability she makes at least 9?

(a) $X \sim \text{Bin}(n = 10, p = 0.80)$, assuming independent shots and constant success probability.
(b) $E[X] = np = 10(0.80) = 8.0$ and $V[X] = np(1 - p) = 10(0.80)(0.20) = 1.6$. Std. dev. $\sigma = \sqrt{1.6} \approx 1.265$.
(c) $P(X = 8) = \binom{10}{8}(0.80)^8(0.20)^2 = 45 \cdot 0.1678 \cdot 0.04 = \boxed{0.302}$.
R-lite: `dbinom(8, 10, 0.80)` = 0.302.
(d) $P(X \geq 9) = P(X = 9) + P(X = 10) = 1 - P(X \leq 8) = 1 - \texttt{pbinom(8, 10, 0.80)} = 1 - 0.6242 = \boxed{0.376}$.

5 Poisson Distribution

Problem 5.1. Customers arrive at a coffee shop during the morning rush at an average rate of 12 per hour, and arrivals can be modeled as a Poisson process.

(a) What is the probability that exactly 10 customers arrive during a randomly chosen 1-hour interval?

(b) What is the probability that exactly 20 customers arrive during a randomly chosen 2-hour interval?

(c) What is the probability that 15 or more customers arrive in a 1-hour interval?

Setup: Let X = number of arrivals in the relevant interval. Poisson rate scales with the length of the interval.

(a) For 1 hour, $X \sim \text{Pois}(\mu = 12)$:

$$P(X = 10) = \frac{e^{-12} \cdot 12^{10}}{10!} = \boxed{0.105}$$

R-lite: `dpois(10, 12)` = 0.105.

(b) For 2 hours, scale the rate: $X \sim \text{Pois}(\mu = 24)$:

$$P(X = 20) = \frac{e^{-24} \cdot 24^{20}}{20!} = \boxed{0.0624}$$

R-lite: `dpois(20, 24)` = 0.0624.

(c) For 1 hour, $X \sim \text{Pois}(12)$:

$$P(X \geq 15) = 1 - P(X \leq 14) = 1 - \texttt{ppois(14, 12)} = 1 - 0.7720 = \boxed{0.228}$$

Problem 5.2. A copy editor finds that her typewritten manuscripts contain typos at an average rate of 1.5 typos per page. The number of typos on any given page can be modeled as Poisson.

(a) What is the probability that a single randomly chosen page contains AT LEAST one typo?

(b) For a 5-page chapter, what is the probability the chapter contains exactly 8 typos?

(c) For a 5-page chapter, what is the probability the chapter contains 5 or fewer typos?

(a) For one page, $X \sim \text{Pois}(\mu = 1.5)$:

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{e^{-1.5} \cdot 1.5^{0}}{0!} = 1 - e^{-1.5} = 1 - 0.2231 = \boxed{0.777}$$

R-lite: `1 - dpois(0, 1.5)` = 0.777.

(b) For 5 pages, scale the rate: $\mu = 5 \cdot 1.5 = 7.5$, so $X \sim \text{Pois}(7.5)$:

$$P(X = 8) = \frac{e^{-7.5} \cdot 7.5^{8}}{8!} = \boxed{0.137}$$

R-lite: `dpois(8, 7.5)` = 0.137.

(c) $P(X \leq 5) = \texttt{ppois(5, 7.5)} = \boxed{0.241}$.

6 Exponential Distribution

Problem 6.1. At a customer-service call center, the time between consecutive incoming calls is exponentially distributed with mean 4 minutes. Let T be the time (in minutes) until the next call.

(a) Identify the distribution of T and state λ.

(b) What is the probability the next call arrives within the next 3 minutes?

(c) What is the probability that more than 5 minutes pass before the next call?

(d) What is the probability the next call arrives between 2 and 6 minutes from now?

(a) $T \sim \text{Exp}(\lambda)$ with mean $1/\lambda = 4$, so $\lambda = 1/4 = 0.25$ per minute. CDF: $F(t) = 1 - e^{-t/4}$.
(b) $P(T < 3) = 1 - e^{-3/4} = 1 - 0.4724 = \boxed{0.528}$.
R-lite: `pexp(3, rate = 1/4)` = 0.528.
(c) $P(T > 5) = e^{-5/4} = \boxed{0.287}$.
R-lite: `1 - pexp(5, rate = 1/4)` = 0.287.
(d) $P(2 < T < 6) = F(6) - F(2) = (1 - e^{-6/4}) - (1 - e^{-2/4}) = e^{-0.5} - e^{-1.5} = 0.6065 - 0.2231 = \boxed{0.383}$.

Problem 6.2. A manufacturer claims that the lifetime of a particular model of LED bulb is exponentially distributed with mean 8,000 hours. Let T denote the lifetime of a randomly chosen bulb.

(a) What is the probability that a randomly chosen bulb lasts more than 10,000 hours?

(b) Find the median lifetime of these bulbs.

(c) For warranty purposes, the manufacturer wants to identify the time t^* such that only 10% of bulbs fail before t^*. Find t^*.

Setup: $T \sim \text{Exp}(\lambda)$ with $\lambda = 1/8000$.
(a) $P(T > 10{,}000) = e^{-10000/8000} = e^{-1.25} = \boxed{0.287}$ hours.
(b) The median m satisfies $F(m) = 0.5$:

$$1 - e^{-m/8000} = 0.5 \;\Rightarrow\; e^{-m/8000} = 0.5 \;\Rightarrow\; m = 8000 \ln 2 \approx \boxed{5545 \text{ hours}}$$

Note that the median is less than the mean (8,000) — the exponential is right-skewed.
(c) Find t^* such that $P(T < t^*) = 0.10$:

$$1 - e^{-t^*/8000} = 0.10 \;\Rightarrow\; e^{-t^*/8000} = 0.90 \;\Rightarrow\; t^* = -8000 \ln(0.90) \approx \boxed{843 \text{ hours}}$$

R-lite: `qexp(0.10, rate = 1/8000)` = 843.

7 Normal Distribution

Problem 7.1. The heights of adult women in the United States are approximately normally distributed with a mean of 64 inches and a standard deviation of 2.5 inches. Let X denote the height of a randomly chosen adult woman.

(a) What is the probability that a randomly chosen woman is less than 60 inches tall?

(b) What is the probability that her height is between 60 and 68 inches?

(c) What is the probability that she is taller than 70 inches?

Setup: $X \sim N(\mu = 64, \sigma = 2.5)$. Standardize using $Z = (X - 64)/2.5$.
(a) $P(X < 60) = P\left(Z < \frac{60 - 64}{2.5}\right) = P(Z < -1.6) = \boxed{0.0548}$.
R-lite: `pnorm(60, 64, 2.5)` = 0.0548.
(b) $P(60 < X < 68) = P(-1.6 < Z < 1.6) = \Phi(1.6) - \Phi(-1.6) = 0.9452 - 0.0548 = \boxed{0.890}$.
(c) $P(X > 70) = P(Z > 2.4) = 1 - \Phi(2.4) = 1 - 0.9918 = \boxed{0.00820}$.
R-lite: `1 - pnorm(70, 64, 2.5)` = 0.0082.

Problem 7.2. SAT total scores are approximately normally distributed with mean 1050 and standard deviation 200.

(a) What proportion of test takers score above 1300?

(b) What is the 90th percentile of SAT scores — i.e., the score below which 90% of students fall?

(c) What range of scores corresponds to the middle 50% of all test takers?

Setup: $X \sim N(1050, 200^2)$.
(a) $P(X > 1300) = P(Z > (1300 - 1050)/200) = P(Z > 1.25) = 1 - \Phi(1.25) = 1 - 0.8944 = \boxed{0.106}$.
(b) 90th percentile: $z_{0.90} = 1.282$ (or use `qnorm(0.90)`).

$$x = \mu + z\sigma = 1050 + (1.282)(200) = 1050 + 256.4 = \boxed{1306}$$

R-lite: `qnorm(0.90, 1050, 200)` = 1306.
(c) Middle 50% lies between 25th and 75th percentiles. $z_{0.25} = -0.674$, $z_{0.75} = 0.674$.

$$x_{25} = 1050 + (-0.674)(200) = 915 \qquad x_{75} = 1050 + (0.674)(200) = 1185$$

The middle 50% spans $\boxed{(915, 1185)}$. **R-lite:** `qnorm(0.25, 1050, 200)` and `qnorm(0.75, 1050, 200)`.

8 One-Sample t-Test

Problem 8.1. A cereal manufacturer claims that the mean weight of cereal in its 16-ounce boxes is 16 oz. A consumer-protection agency suspects boxes are systematically underfilled or overfilled. The agency randomly samples $n = 25$ boxes and measures their contents. The sample mean weight is $\bar{x} = 15.7$ oz with standard deviation $s = 0.5$ oz.

(a) State the hypotheses for testing the agency's suspicion.

(b) Conduct the test at $\alpha = 0.05$. Compute the test statistic, p-value, and state your conclusion in context.

(c) Construct and interpret a 95% confidence interval for the true mean weight.

(a) Let μ = true mean weight (oz). The agency suspects either over- or under-filling, so test two-sided:

$$H_0 : \mu = 16 \quad \text{vs.} \quad H_a : \mu \neq 16$$

(b) Conditions: random sample, $n = 25$ borderline for CLT — assume box weights are approximately normal.

Test statistic:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{15.7 - 16.0}{0.5/\sqrt{25}} = \frac{-0.3}{0.10} = \boxed{-3.00} \quad (df = 24)$$

p-value (two-sided): $p = 2 \cdot P(T_{24} > 3.00) =$ `2*(1 - pt(3, 24))` $\approx \boxed{0.00619}$.

Conclusion: With a p-value of 0.0062 which is less than $\alpha = 0.05$, we reject H_0. There is strong evidence that the true mean box weight differs from 16 oz; in fact the sample suggests boxes are underfilled.

(c) 95% CI: $t_{0.025,24} = 2.064$.

$$15.7 \pm 2.064 \cdot \frac{0.5}{\sqrt{25}} = 15.7 \pm 2.064(0.10) = 15.7 \pm 0.206$$

$$\boxed{95\% \text{ CI} : (15.49, 15.91)}$$

We are 95% confident that the true mean box weight lies between 15.49 and 15.91 oz. Since 16 is NOT in the interval, this is consistent with rejecting H_0.

Problem 8.2. A laptop manufacturer advertises that its new battery lasts at least 50 hours under normal usage. A reviewer samples $n = 15$ laptops and finds a sample mean lifetime of $\bar{x} = 47.5$ hours with $s = 4.2$ hours. Assume battery lifetimes are approximately normal.

(a) State the appropriate hypotheses for testing whether the manufacturer's claim is overstated.

(b) Conduct the test at $\alpha = 0.05$. State your conclusion in context.

(a) Let μ = true mean battery life (hours). The reviewer suspects the actual life is *less than* the claim:

$$H_0 : \mu = 50 \quad \text{vs.} \quad H_a : \mu < 50$$

(b) Conditions: random sample, normality assumed (small n). $df = 14$.

$$t = \frac{47.5 - 50}{4.2/\sqrt{15}} = \frac{-2.5}{1.0844} = \boxed{-2.305}$$

p-value (lower tail): $p = P(T_{14} < -2.305) =$ `pt(-2.305, 14)` $\approx \boxed{0.0185}$.

Conclusion: With a p-value of 0.0185 which is less than $\alpha = 0.05$, we reject H_0. There is sufficient evidence that the true mean battery lifetime is below the manufacturer's claim of 50 hours.

9 One-Proportion z-Test

Problem 9.1. A company is considering replacing its current logo. Before committing, it runs a survey of 400 randomly chosen customers, asking whether they prefer the new logo over the current one. Of the 400 surveyed, 240 prefer the new logo. The marketing team will adopt the new logo only if more than half of customers prefer it.

(a) State the appropriate hypotheses.

(b) Conduct the test at $\alpha = 0.05$. Compute the test statistic, p-value, and state your conclusion in context.

(c) Construct a 95% confidence interval for the true proportion who prefer the new logo.

(a) Let p = true proportion of customers who prefer the new logo:

$$H_0 : p = 0.50 \quad \text{vs.} \quad H_a : p > 0.50$$

(b) Conditions: $np_0 = 400(0.50) = 200 \geq 10$ and $n(1 - p_0) = 200 \geq 10$ ✓.
$\hat{p} = 240/400 = 0.60$.

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{0.60 - 0.50}{\sqrt{(0.50)(0.50)/400}} = \frac{0.10}{0.025} = \boxed{4.00}$$

p-value (upper tail): $p = 1 - \Phi(4.00) =$ `1 - pnorm(4)` $\approx \boxed{0.0000317}$.

Conclusion: With a p-value of ≈ 0.00003 which is less than $\alpha = 0.05$, we reject H_0. There is overwhelming evidence that more than half of customers prefer the new logo, supporting adoption.

(c) 95% CI uses $\hat{p}$ in the SE:

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.60 \pm 1.96\sqrt{\frac{(0.60)(0.40)}{400}} = 0.60 \pm 1.96(0.0245)$$
$$= 0.60 \pm 0.048 \Rightarrow \boxed{95\% \text{ CI} : (0.552, 0.648)}$$

We are 95% confident the true proportion of customers preferring the new logo lies between 55.2% and 64.8%. Since 0.50 falls below this interval, this agrees with rejecting H_0.

Problem 9.2. A factory's stated defect rate for one of its products is 5%. A QA inspector wants to know whether the actual rate is higher than the stated rate. She randomly samples $n = 500$ units and finds 32 defective.

(a) State the hypotheses.

(b) Conduct the test at $\alpha = 0.05$. Compute the test statistic, p-value, and state your conclusion in context.

(a) Let p = true defect rate:

$$H_0 : p = 0.05 \quad \text{vs.} \quad H_a : p > 0.05$$

(b) Conditions: $np_0 = 500(0.05) = 25 \geq 10$, $n(1 - p_0) = 475 \geq 10$ ✓.
$\hat{p} = 32/500 = 0.064$.

$$z = \frac{0.064 - 0.05}{\sqrt{(0.05)(0.95)/500}} = \frac{0.014}{0.00975} = \boxed{1.44}$$

p-value (upper tail): $p = 1 - \Phi(1.44) =$ `1 - pnorm(1.4364)` $\approx \boxed{0.0754}$.

Conclusion: With a p-value of 0.0754 which is greater than $\alpha = 0.05$, we fail to reject H_0. There is insufficient evidence to conclude the true defect rate exceeds the stated 5%. (Note: it would be significant at $\alpha = 0.10$.)

10 Two-Sample t-Test (Independent)

Problem 10.1. A horticulturist wants to compare the effect of two fertilizers on tomato yield (kg per plot). She randomly assigns plots to receive either Fertilizer A or Fertilizer B. Results:

Fertilizer	n	$\bar{x}$ (kg)	s (kg)
A	20	45	4
B	22	42	5

(a) Why are the two samples treated as independent (not paired)?

(b) Test at $\alpha = 0.05$ whether the two fertilizers produce different mean yields. State the hypotheses, test statistic (using conservative df), p-value, and conclusion.

(c) Construct a 95% confidence interval for $\mu_A - \mu_B$ and interpret.

(a) The plots assigned to A and the plots assigned to B are different physical plots; there is no natural one-to-one matching. Treating them as independent samples is appropriate.

(b) Hypotheses:

$$H_0 : \mu_A = \mu_B \quad \text{vs.} \quad H_a : \mu_A \neq \mu_B \quad (\text{or } H_0 : \mu_A - \mu_B = 0)$$

Conditions: independent random samples, both n's near or above 20 (CLT plausible).

Standard error:

$$SE = \sqrt{\frac{4^2}{20} + \frac{5^2}{22}} = \sqrt{0.80 + 1.1364} = \sqrt{1.9364} = 1.392$$

Test statistic:

$$t = \frac{\bar{x}_A - \bar{x}_B}{SE} = \frac{45 - 42}{1.392} = \boxed{2.156}$$

Conservative df: $\min(20 - 1, 22 - 1) = 19$.

p-value (two-sided): $p = 2(1 - P(T_{19} \leq 2.156)) =$ `2*(1 - pt(2.156, 19))` $\approx \boxed{0.0441}$.

Conclusion: With a p-value of 0.0441 which is less than $\alpha = 0.05$, we reject H_0. There is evidence that the two fertilizers produce different mean yields; Fertilizer A appears to outperform B.

(c) 95% CI: $t_{0.025,19} = 2.093$.

$$(45 - 42) \pm 2.093(1.392) = 3 \pm 2.914 \Rightarrow \boxed{95\% \text{ CI for } \mu_A - \mu_B : (0.09, 5.91) \text{ kg}}$$

We are 95% confident that the true difference in mean yields ($\mu_A - \mu_B$) is between 0.09 and 5.91 kg per plot. Since the interval excludes 0 (just barely), this agrees with rejecting H_0.

Problem 10.2. An education researcher compares two teaching methods for an introductory statistics course. Random samples of students are assigned to each method, and final exam scores are recorded:

Method	n	$\bar{x}$	s
1 (lecture)	15	78	8
2 (flipped)	18	72	10

The researcher hypothesizes that students taught with lecture (Method 1) score higher on average than those taught with the flipped classroom (Method 2).

Test at $\alpha = 0.05$. State the hypotheses, test statistic (with conservative df), p-value, and conclusion.

Hypotheses:

$$H_0 : \mu_1 = \mu_2 \quad \text{vs.} \quad H_a : \mu_1 > \mu_2$$

Conditions: independent random samples; assume approximate normality given small n's.

Standard error:

$$SE = \sqrt{\frac{8^2}{15} + \frac{10^2}{18}} = \sqrt{4.267 + 5.556} = \sqrt{9.822} = 3.134$$

Test statistic:

$$t = \frac{78 - 72}{3.134} = \boxed{1.915}$$

Conservative df: $\min(15 - 1, 18 - 1) = 14$.

p**-value** (upper tail): $p = 1 - P(T_{14} \leq 1.915) =$ `1 - pt(1.915, 14)` $\approx \boxed{0.0381}$.

Conclusion: With a p-value of 0.0381 which is less than $\alpha = 0.05$, we reject H_0. There is sufficient evidence that the lecture method produces higher mean exam scores than the flipped classroom method in this study.

11 Two-Proportion z-Test

Problem 11.1. A national retailer surveys customers in two regions about which of two product designs they prefer. Of 200 randomly sampled customers in Region 1, 140 preferred Design A. Of 250 customers in Region 2, 145 preferred Design A. Test at $\alpha = 0.05$ whether the proportion preferring Design A differs between regions.

(a) State the hypotheses.

(b) Conduct the test. Report the test statistic, p-value, and conclusion in context.

(c) Construct a 95% CI for $p_1 - p_2$.

(a) Let p_1, p_2 = true proportions preferring Design A in Regions 1 and 2.

$$H_0 : p_1 = p_2 \quad \text{vs.} \quad H_a : p_1 \neq p_2$$

(b) Conditions: $n_1\hat{p}_1 = 140, n_1(1-\hat{p}_1) = 60, n_2\hat{p}_2 = 145, n_2(1-\hat{p}_2) = 105$ — all ≥ 10 ✓.
$\hat{p}_1 = 140/200 = 0.70, \quad \hat{p}_2 = 145/250 = 0.58.$
Pooled (because we are testing H_0):

$$\hat{p} = \frac{140+145}{200+250} = \frac{285}{450} = 0.6333$$

$$SE_{HT} = \sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} = \sqrt{(0.6333)(0.3667)\left(\frac{1}{200}+\frac{1}{250}\right)} = \sqrt{0.002089} = 0.04571$$

$$z = \frac{0.70-0.58}{0.04571} = \boxed{2.625}$$

p-value (two-sided): $p = 2(1-\Phi(2.625)) = \texttt{2*(1 - pnorm(2.625))} \approx \boxed{0.00866}$.

Conclusion: With a p-value of 0.00866 which is less than $\alpha = 0.05$, we reject H_0. There is strong evidence that the proportion of customers preferring Design A differs between the two regions, with Region 1 having the higher rate.

(c) 95% CI uses *unpooled* SE:

$$SE_{CI} = \sqrt{\frac{(0.70)(0.30)}{200} + \frac{(0.58)(0.42)}{250}} = \sqrt{0.001050 + 0.000974} = \sqrt{0.002024} = 0.04499$$

$$(0.70 - 0.58) \pm 1.96(0.04499) = 0.12 \pm 0.0882 \;\Rightarrow\; \boxed{(0.032, 0.208)}$$

We are 95% confident the true difference $p_1 - p_2$ lies between 0.032 and 0.208. Since 0 is not in the interval, this is consistent with rejecting H_0.

Problem 11.2. An online advertising team is comparing two ad campaigns. Campaign 1 was shown to 1,000 users, of whom 35 clicked. Campaign 2 was shown to 1,200 users, of whom 60 clicked. The team wants to test at $\alpha = 0.05$ whether Campaign 2 has a higher click-through rate than Campaign 1.

State the hypotheses, conduct the test, and state your conclusion in context.

Hypotheses:

$$H_0 : p_2 = p_1 \quad \text{vs.} \quad H_a : p_2 > p_1$$

Conditions: all four counts $(35, 965, 60, 1140)$ are ≥ 10 ✓.
$\hat{p}_1 = 35/1000 = 0.035, \quad \hat{p}_2 = 60/1200 = 0.050.$

Pooled:

$$\hat{p} = \frac{35 + 60}{1000 + 1200} = \frac{95}{2200} = 0.04318$$

$$SE = \sqrt{(0.04318)(0.9568)\left(\frac{1}{1000} + \frac{1}{1200}\right)} = \sqrt{0.0000757} = 0.008702$$

$$z = \frac{0.050 - 0.035}{0.008702} = \boxed{1.724}$$

p**-value** (upper tail): $p = 1 - \Phi(1.724) =$ `1 - pnorm(1.724)` $\approx \boxed{0.0424}$.

Conclusion: With a p-value of 0.0424 which is less than $\alpha = 0.05$, we reject H_0. There is sufficient evidence that Campaign 2 has a higher true click-through rate than Campaign 1.

12 Simple Linear Regression

Problem 12.1. A statistics professor fits a simple linear regression of final exam score (y, out of 100) on the number of hours per week the student studied (x). Using data from $n = 30$ students, the fitted regression equation is:

$$\hat{y} = 50 + 4.2\,x$$

The model has $R^2 = 0.65$, and the slope is statistically significant at $\alpha = 0.05$.

(a) Interpret the slope coefficient in context.

(b) Interpret the intercept. Is its interpretation meaningful in this context?

(c) Interpret the value of R^2 in context.

(d) Predict the exam score of a student who studies 8 hours per week.

(a) For each additional hour of weekly study time, predicted final exam score increases by an average of **4.2 points**. Because the slope is statistically significant, the data provide evidence that there is a non-zero linear relationship.

(b) The intercept (50) is the predicted exam score for a student who studies *0 hours* per week. While this is mathematically the y-intercept, it is on the boundary or outside the range of typical study habits, so the interpretation is at best a baseline reference and at worst an extrapolation. It does not predict a real student well.

(c) $R^2 = 0.65$ means that approximately **65% of the variability** in final exam scores in this sample is explained by the linear relationship with weekly study hours. The remaining 35% is due to other factors not in the model (innate ability, sleep, prior preparation, test anxiety, etc.).

(d) For $x = 8$:

$$\hat{y} = 50 + 4.2(8) = 50 + 33.6 = \boxed{83.6 \text{ points}}$$

Problem 12.2. A real-estate analyst regresses home sale price (y, in thousands of dollars) on home size (x, in square feet) using $n = 50$ recently sold homes. The fitted equation is:

$$\hat{y} = 80 + 0.15\,x$$

with $R^2 = 0.78$. The data set contains homes ranging from 1,000 to 3,500 square feet.

(a) Interpret the slope coefficient in dollars per square foot.

(b) Predict the sale price of a 2,000-square-foot home.

(c) A developer asks you to use the model to predict the price of a 5,000-square-foot luxury home. Why should you be cautious?

(a) Each additional square foot of home size is associated with a predicted increase of \$0.15 thousand $=$ \$**150** in sale price, on average.

(b) For $x = 2000$:

$$\hat{y} = 80 + 0.15(2000) = 80 + 300 = 380 \text{ thousand dollars} = \boxed{\$380{,}000}$$

(c) Extrapolation. 5,000 sq ft falls well outside the data range (1,000–3,500 sq ft) used to fit the model. The relationship between size and price may not remain linear at much larger sizes (luxury homes often command non-linear premiums for amenities, lot, location). Predicting outside the range of observed predictor values is risky — the true relationship may bend in ways the linear model cannot detect.

13 Multiple Linear Regression

Problem 13.1. A labor economist fits an MLR model predicting annual salary (y, in \$1,000s) from years of education (x_1), years of work experience (x_2), and average weekly hours (x_3), using a random sample of $n = 100$ workers in a metropolitan area. Selected R output:

$$\hat{y} = 35 + 2.5\,x_1 + 1.8\,x_2 - 0.4\,x_3$$

Predictor	Estimate	Std. Error	t-value	p-value
(Intercept)	35.00	4.20	8.33	< 0.001
Education (x_1)	2.50	0.55	4.55	< 0.001
Experience (x_2)	1.80	0.30	6.00	< 0.001
Hours (x_3)	-0.40	0.35	-1.14	0.257

$R^2 = 0.72$, Adjusted $R^2 = 0.71$.

(a) Identify which predictors are statistically significant at $\alpha = 0.05$.

(b) Interpret the coefficient on Education in context.

(c) Interpret R^2 in context.

(d) Predict the annual salary of a worker with 16 years of education, 10 years of experience, working 45 hours per week.

(a) A predictor is significant if its p-value is less than α.

- Education: $p < 0.001 \;\Rightarrow$ **significant**
- Experience: $p < 0.001 \;\Rightarrow$ **significant**
- Hours: $p = 0.257 > 0.05 \;\Rightarrow$ **not significant**

Education and Experience are significant predictors; Hours is not.

(b) Holding Experience and Hours constant, each additional year of education is associated with an estimated increase of \$2,500 in annual salary, on average.

(c) Approximately 72% of the variability in annual salary in this sample is explained by Education, Experience, and Hours together. The remaining 28% is due to factors not captured (industry, employer size, location, individual ability, networking, etc.).

(d) For $x_1 = 16$, $x_2 = 10$, $x_3 = 45$:

$$\hat{y} = 35 + 2.5(16) + 1.8(10) - 0.4(45) = 35 + 40 + 18 - 18 = \boxed{75 \text{ (i.e., \$75,000)}}$$

Problem 13.2. A property-management firm fits an MLR model predicting monthly rent (y, in dollars) from the number of bedrooms (x_1), bathrooms (x_2), and a categorical variable for neighborhood. The neighborhood variable has three levels: Suburbs (reference), Midtown, Downtown. The fitted model is:

$$\hat{y} = 800 + 250\,x_1 + 100\,x_2 + 150\,I_{\text{Midtown}} + 200\,I_{\text{Downtown}}$$

where I_{Midtown} and I_{Downtown} are indicator variables.

(a) Interpret the coefficient on Bedrooms (x_1).

(b) Interpret the coefficient on I_{Downtown}.

(c) Predict the monthly rent of a 2-bedroom, 2-bathroom apartment in Downtown.

(d) Predict the monthly rent of a 2-bedroom, 2-bathroom apartment in the Suburbs.

(a) Holding bathrooms and neighborhood constant, each additional bedroom is associated with an estimated \$250 increase in monthly rent, on average.

(b) Holding bedrooms and bathrooms constant, a Downtown apartment is predicted to cost an average of \$200 more per month than an otherwise-identical apartment in the **Suburbs** (the reference category).

(c) For $x_1 = 2, x_2 = 2, I_{\text{Downtown}} = 1, I_{\text{Midtown}} = 0$:

$$\hat{y} = 800 + 250(2) + 100(2) + 150(0) + 200(1) = 800 + 500 + 200 + 0 + 200 = \boxed{\$1{,}700}$$

(d) For $x_1 = 2, x_2 = 2$, both indicators $= 0$ (Suburbs is the reference):

$$\hat{y} = 800 + 250(2) + 100(2) = 800 + 500 + 200 = \boxed{\$1{,}500}$$

The \$200 difference between (c) and (d) is exactly the Downtown-vs.-Suburbs adjustment from part (b).

14 One-Way ANOVA

Problem 14.1. An agricultural researcher is comparing the effect of $k = 4$ fertilizer formulations (A, B, C, D) on tomato yield (kg per plot). She randomly assigns 10 plots to each formulation ($N = 40$ total). After harvest, an R analysis returns:

$$\text{SSTr} = 240 \qquad \text{SSE} = 540$$

(a) State the hypotheses for the ANOVA F-test.

(b) Calculate the F test statistic.

(c) Find the p-value using R-lite, then state your conclusion at $\alpha = 0.05$ in context.

(a) Let $\mu_A, \mu_B, \mu_C, \mu_D$ be the true mean yields for each formulation.

$$H_0\colon \mu_A = \mu_B = \mu_C = \mu_D \quad \text{vs.} \quad H_a\colon \text{at least one differs}$$

(b) Degrees of freedom: $\text{df}_{\text{Tr}} = k - 1 = 3, \quad \text{df}_{\text{E}} = N - k = 36.$

$$\text{MSTr} = \frac{\text{SSTr}}{k-1} = \frac{240}{3} = 80 \qquad \text{MSE} = \frac{\text{SSE}}{N-k} = \frac{540}{36} = 15$$

$$F = \frac{\text{MSTr}}{\text{MSE}} = \frac{80}{15} = \boxed{5.333}$$

(c) p-value:

$$p = P(F_{3,36} > 5.333) = \texttt{1 - pf(5.333, 3, 36)} \approx \boxed{0.00382}$$

Conclusion: With a p-value of ≈ 0.0038 which is less than $\alpha = 0.05$, we reject H_0. There is strong evidence that at least one of the four fertilizer formulations produces a different mean yield than the others. Pairwise comparisons (e.g., Tukey HSD) would be needed to identify which specific formulations differ.

Problem 14.2. A health researcher randomly assigns 36 sedentary adults equally to one of $k = 3$ exercise programs ($n = 12$ per program). After 12 weeks, weight loss (in lb) is recorded. An R analysis returns:

$$\text{SSTr} = 180 \qquad \text{SSE} = 396$$

(a) State the hypotheses.

(b) Compute the F test statistic.

(c) Find the p-value via R-lite and state your conclusion at $\alpha = 0.05$.

(d) State the three conditions required for the ANOVA F-test to be valid, and briefly comment on whether they appear reasonable here.

(a) Let μ_1, μ_2, μ_3 be the true mean weight losses across the three programs.

$$H_0\colon \mu_1 = \mu_2 = \mu_3 \quad \text{vs.} \quad H_a\colon \text{at least one differs}$$

(b) $\text{df}_{\text{Tr}} = k - 1 = 2, \quad \text{df}_{\text{E}} = N - k = 33.$

$$\text{MSTr} = \frac{180}{2} = 90 \qquad \text{MSE} = \frac{396}{33} = 12$$

$$F = \frac{90}{12} = \boxed{7.50}$$

(c) p-value:

$$p = P(F_{2,33} > 7.50) = \texttt{1 - pf(7.50, 2, 33)} \approx \boxed{0.00207}$$

Conclusion: With a p-value of ≈ 0.0021 which is less than $\alpha = 0.05$, we reject H_0. There is strong evidence that the mean weight loss differs across at least two of the three exercise programs.

(d) ANOVA conditions:

1. *Independent random samples* within and across groups — satisfied by random assignment.
2. *Approximately normal populations* within each group — with only $n = 12$ per group, this is hard to verify rigorously, but checking residual histograms / QQ-plots is the standard approach.
3. *Equal variances* across groups — check via the rule of thumb $\max(s_i)/\min(s_i) < 2$.

The randomized design satisfies (1) by construction. Conditions (2) and (3) would need to be verified from the actual data; assuming no severe outliers and roughly comparable spread, the ANOVA result is trustworthy.

End of Review.
If you can confidently work each of these 26 free-response problems and the 12 multiple-choice items, you have demonstrated mastery of every TEE topic: probability and conditional probability, the four named distributions (Binomial, Poisson, Exponential, Normal), one- and two-sample inference for means and proportions, simple and multiple linear regression, and one-way ANOVA.

Final reminders:

- State hypotheses about *population parameters* (μ, p), not about sample statistics ($\bar{x}$, $\hat{p}$).
- Always check conditions before applying a test.
- Use the R-lite functions (`p`*`dist`*, `q`*`dist`*) for p-values and critical values — and remember the always-upper-tailed structure of ANOVA p-values.
- Conclusions need to be stated in context, not just "reject H_0."
- Confidence intervals and hypothesis tests are complementary — use both when asked.