STAT 515 — Exam 2 Practice

Questions with Worked Solutions

Review Session Handout · Instructor: Ho, Yen-Yi

Question 1. Definitions (8 pts)

a) Define what is meant by "the significance level (α-level)" of a hypothesis test.

b) Define what is meant by "Type I error."

SOLUTION

a) The significance level α is the probability of committing a Type I error — the probability of rejecting H_0 when H_0 is actually true. It is chosen by the researcher before the data are examined, and it fixes how much evidence is required to reject H_0: it determines the rejection region, and equivalently we reject H_0 whenever the p-value $\leq \alpha$.

b) A Type I error occurs when we reject H_0 even though H_0 is in fact true (a "false alarm"). Its probability equals α.

Question 2. Quality control (10 pts)

A manufacturer knows that 8% of its light bulbs are defective. Use the normal approximation to the binomial to approximate the probability that more than 25 bulbs are defective in a shipment of 250. You may leave your answer in the form of a probability statement about a standard normal (for example P(Z > 1.5)).

SOLUTION

Let X = number of defective bulbs, X ~ Binomial(n = 250, p = 0.08).

Step 1 — check the condition for the normal approximation: $np = 250(0.08) = 20 \geq 5$ and $np(1-p) = 250(0.08)(0.92) = 18.4 \geq 5$. ✓

Step 2 — find the mean and standard deviation:

$$\mu = np = 250(0.08) = 20, \quad \sigma = \sqrt{(np(1-p))} = \sqrt{18.4} \approx 4.29$$

Step 3 — standardize

$$\mathbf{P(X > 25) \approx P(Z > (25 - 20)/4.29) = P(Z > 1.17)}$$

This is an acceptable final answer. (From the table it equals $1 - 0.8790 = 0.1210$; the exact binomial value is 0.1029.)

Question 3. Normal probability (Q–Q) plot (10 pts)

A quality engineer records the fill volume of 30 randomly selected bottles and constructs a normal quantile-quantile (Q–Q) plot of the data before carrying out any inference.

a) (3 pts) What is plotted on each axis of a Q-Q plot, and what pattern should the points follow if the sample is consistent with a normally distributed population?

b) (4 pts) Instead, the plot bends away from a straight line: the smallest observations fall slightly above the line and the largest observations fall far above it, producing an upward curve. What does this pattern tell you about the shape of the underlying distribution?

SOLUTION

a) One axis holds the ordered sample values (the sample quantiles); the other holds the corresponding theoretical quantiles of the standard normal distribution. If the population is normal, the points should fall close to a straight line.

b) An upward (convex) curve indicates a distribution that is skewed to the right. The upper tail is stretched relative to a normal distribution, so the largest observations are much larger than normality would predict.

Question 4. Sampling concepts (8 pts)

State the Central Limit Theorem. Under what conditions is n generally large enough for the CLT to apply?

SOLUTION

Central Limit Theorem. For a random sample of size n drawn from any population with mean μ and standard deviation σ, the sampling distribution of the sample mean $\bar{x}$ has

$$\mu_{\bar{x}} = \mu, \quad \sigma_{\bar{x}} = \sigma/\sqrt{n},$$

and is approximately normal when n is sufficiently large — regardless of the shape of the population distribution.

When is n large enough?

- If the population is itself normal, $\bar{x}$ is exactly normal for every n (no condition needed).
- Otherwise, $n \geq 30$ is the working rule used in this course. The more skewed the population, the larger n must be.
- For a sample proportion (the binomial case), the parallel conditions are $np \geq 5$ and $n(1-p) \geq 5$.

Question 5. Interpreting a confidence interval (10 pts)

A nutritionist studied the average daily sodium intake (in mg) of adults in a city. A sample of 40 adults yielded a 95% confidence interval of (2,104 mg, 2,896 mg). Identify the error in each of the following interpretations, and briefly explain why it is an error.

a) "We are 95% confident that the average daily sodium intake of the 40 adults sampled is between 2,104 mg and 2,896 mg."

b) "There is a 95% probability that the true population mean daily sodium intake is between 2,104 mg and 2,896 mg."

SOLUTION

a) Error: the interval estimates the population mean, not the sample mean. The sample mean is already known exactly — it is the center of the interval, (2104 +

2896)/2 = 2500 mg — so there is nothing to be "confident" about. Corrected: "We are 95% confident that the mean daily sodium intake of all adults in this city is between 2,104 and 2,896 mg."

b) Error: once the data are collected and the interval is computed, μ is a fixed (unknown) constant and the two endpoints are fixed numbers. This particular interval either contains μ or it does not — the probability is 0 or 1, not 0.95. The 95% describes the procedure: in repeated sampling, 95% of intervals constructed this way would contain μ. Corrected: "We are 95% confident that …"

Question 6. Sleep habits survey (10 pts)

A university health center surveyed 980 undergraduate students and found that 412 reported getting fewer than 7 hours of sleep on a typical night. Construct a 95% confidence interval for the true proportion of all undergraduates at this university who get fewer than 7 hours of sleep per night.

SOLUTION

Let p = the true proportion of all undergraduates at this university who get fewer than 7 hours of sleep per night.

Step 1 — point estimate:

$$\hat{p} = 412/980 \approx 0.4204$$

Step 2 — check conditions: $n\hat{p} = 412 \geq 5$ and $n(1-\hat{p}) = 568 \geq 5$. ✓

Step 3 — standard error and margin of error ($z_{0.025} = 1.96$):

$$SE = \sqrt{\hat{p}(1-\hat{p})/n} = \sqrt{0.4204 \times 0.5796 / 980} = \sqrt{0.0002487} \approx 0.01577$$

$$ME = 1.96(0.01577) \approx 0.0309$$

Step 4 — interval:

$$0.4204 \pm 0.0309 \quad \longrightarrow \quad (0.3895, 0.4513)$$

Interpretation: we are 95% confident that between 38.95% and 45.13% of all undergraduates at this university get fewer than 7 hours of sleep per night.

Question 7. Sample size determination (10 pts)

a) A researcher wants to estimate the mean commute time (in minutes) for employees at a large company to within ±2 minutes with 95% confidence. A pilot study suggests the standard deviation is approximately 10 minutes. What sample size is needed?

b) If instead the researcher had no prior estimate of σ, and wanted to estimate the proportion of employees who commute more than 30 minutes to within ±0.04 with 90% confidence, what sample size would be needed?

SOLUTION

a) Estimating a mean. Here E = 2, σ = 10, and 95% confidence gives z = 1.96:

$$\mathbf{n = (z\sigma/E)^2 = (1.96 \times 10 / 2)^2 = (9.8)^2 = 96.04 \quad \longrightarrow \quad n = 97}$$

b) Estimating a proportion. With no prior estimate, use p = 0.5, which makes p(1−p) = 0.25 as large as possible and therefore gives the most conservative (largest) n. For 90% confidence, z = 1.645 and E = 0.04:

$$\mathbf{n = (z/E)^2\, p(1-p) = (1.645/0.04)^2\, (0.25) = (41.125)^2\, (0.25) \approx 1691.3 \times 0.25 \approx 422.8 \quad \longrightarrow \quad n = 423}$$

Question 8. Setting up hypotheses (12 pts)

For each situation, state the appropriate null and alternate hypotheses. Be sure to define any parameter(s) you use in terms of the problem.

a) A coffee company claims its espresso pods contain an average of at least 75 mg of caffeine per pod. A consumer watchdog group suspects the true mean is less than 75 mg and plans to test the claim.

b) A hospital administrator wants to determine whether the average length of stay (in days) differs between two units — a standard care unit and an enhanced recovery unit. Patients are randomly assigned to one of the two units.

c) A packaging machine is set to fill bottles with exactly 500 mL. A quality inspector suspects the machine is not filling correctly (either over- or under-filling) and takes a random sample of bottles to test.

SOLUTION

a) Let μ = the true mean caffeine content (mg) per espresso pod.

$$\mathbf{H_0: \mu = 75 \quad vs. \quad H_a: \mu < 75}$$

The company's claim ("at least 75") is the status quo and goes in H_0; what the watchdog group wants to demonstrate goes in H_a.

b) Let μ_1 = the mean length of stay (days) in the standard care unit and μ_2 = the mean length of stay (days) in the enhanced recovery unit.

$$\mathbf{H_0: \mu_1 - \mu_2 = 0 \quad vs. \quad H_a: \mu_1 - \mu_2 \neq 0}$$

"Differs" names no direction, so the test is two-sided.

c) Let μ = the true mean fill volume (mL) per bottle.

$$\mathbf{H_0: \mu = 500 \quad vs. \quad H_a: \mu \neq 500}$$

"Either over- or under-filling" is explicitly two-sided.

Question 9. Hypothesis test (14 pts)

A physical therapist believes that the mean resting heart rate of patients enrolled in a cardiac rehabilitation program is different from the national average of 72 bpm. A random sample of 18 patients gave a mean of 76.4 bpm and a standard deviation of 9.2 bpm. Assume the population is approximately normal.

a) State the null and alternate hypotheses.

b) Compute the test statistic.

c) Find the rejection region using $\alpha = 0.05$.

d) State your conclusion in the context of the problem.

SOLUTION

a) Let μ = the true mean resting heart rate (bpm) of patients in the cardiac rehabilitation program.

$$\mathbf{H_0: \mu = 72 \quad vs. \quad H_a: \mu \neq 72}$$

b) σ is unknown, $n = 18$ is small, and the population is approximately normal, so use a one-sample t test with $df = n - 1 = 17$.

$$\mathbf{SE = s/\sqrt{n} = 9.2/\sqrt{18} = 9.2/4.243 \approx 2.169}$$

$$\mathbf{t = (\bar{x} - \mu_0)/SE = (76.4 - 72)/2.169 = 4.4/2.169 \approx 2.03}$$

c) Two-tailed test, $\alpha = 0.05$, $df = 17 \Rightarrow t_{0.025,17} \approx 2.110$.

Rejection region: $|t| > 2.110$ (that is, $t < -2.110$ or $t > 2.110$)

d) $|t| = 2.03 < 2.110$, so we fail to reject H_0 at $\alpha = 0.05$. There is insufficient evidence at the 5% level to conclude that the mean resting heart rate of cardiac rehabilitation patients differs from 72 bpm.

(For reference, the two-sided p-value is about 0.058 — just above 0.05, which agrees with the decision.)

Question 10. p-value and decision (8 pts)

Consider testing H_0: $\mu = 50$ vs. H_a: $\mu \neq 50$. A sample of size 22 gave $\bar{x} = 47.3$ and $s = 6.8$. The resulting two-sided p-value is 0.0761.

a) (Circle the correct response.) For an α-level of 0.05 we would fail to reject / reject the null hypothesis.

b) (Circle the correct response.) For an α-level of 0.10 we would fail to reject / reject the null hypothesis.

c) What would the p-value have been if we were instead testing H_0: $\mu = 50$ vs. H_a: $\mu < 50$? Briefly explain your reasoning.

SOLUTION

a) Fail to reject. The decision rule is "reject when p-value $\leq \alpha$," and $0.0761 > 0.05$.

b) Reject. Now $0.0761 < 0.10$.

c) $\bar{x} = 47.3 < 50$, so the observed result falls in the lower tail — the same direction the alternative H_a: $\mu < 50$ predicts. The two-sided p-value splits the same probability across both tails, so the one-sided p-value is half of it:

$p = 0.0761 / 2 = 0.03805$

Question 11. Confidence intervals for a proportion with small n (10 pts)

Answer the following questions (each counts as 5 pts).

a) A colleague constructs a confidence interval for a proportion using $\hat{p} = X/n$ directly, where $n = 12$ and $X = 1$. Give one reason why the standard formula $\hat{p} \pm z^*\sqrt{\hat{p}(1-\hat{p})/n}$ may be unreliable here, even though they calculated it correctly.

b) Explain why Wilson's adjusted interval (which adds two successes and two failures to the data before computing the interval) tends to perform better than the standard formula when $\hat{p}$ is close to 0 or 1, or when n is small.

SOLUTION

a) Here $\hat{p} = 1/12 \approx 0.083$. The large-sample condition is $n\hat{p} \geq 5$ and $n(1-\hat{p}) \geq 5$, but $n\hat{p} = 1$, far short of 5. The sampling distribution of $\hat{p}$ is nowhere near normal, so the z-based formula does not apply — the arithmetic is right but the method is not valid. The symptom is visible in the answer itself:

$$\mathbf{0.083 \pm 1.96\sqrt{(0.083 \times 0.917 / 12)} = 0.083 \pm 0.156 = (-0.073, 0.240)}$$

A negative lower endpoint is impossible for a proportion.

b) Wilson's adjustment (also called the plus-four or Agresti–Coull interval) replaces $\hat{p}$ with $\tilde{p} = (X+2)/(n+4)$ and uses $SE = \sqrt{\tilde{p}(1-\tilde{p})/(n+4)}$. Adding two successes and two failures pulls the estimate away from 0 or 1 and toward 0.5, which does two things at once: it moves the center off the boundary, and it keeps the variance estimate $\hat{p}(1-\hat{p})$ from collapsing toward 0. That collapse is the standard interval's fatal flaw at extreme $\hat{p}$ — the interval becomes far too narrow precisely when the estimate is least trustworthy. The adjusted interval's actual coverage is therefore much closer to the advertised 95% for small n or extreme p. Here:

$$\mathbf{\tilde{p} = 3/16 = 0.1875, \quad SE = \sqrt{(0.1875 \times 0.8125 / 16)} \approx 0.0976}$$

$$\mathbf{0.1875 \pm 1.96(0.0976) \approx (-0.004, 0.379) \quad \longrightarrow \quad \text{truncate to } (0, 0.379)}$$

— wider, centered more sensibly, and a far more honest summary of what one success in twelve trials tells us.

Quick reference: which procedure do I use?

Situation	Procedure	Conditions to check
Mean μ, σ known	z interval / z test	Normal population, or $n \geq 30$ (CLT)
Mean μ, σ unknown, n small	t, $df = n - 1$	Population approximately normal; t is robust to mild departures
Mean μ, σ unknown, n large	t (or z), $df = n - 1$	$n \geq 30$ (CLT); no normality needed
Proportion p	z interval / z test	$n\hat{p} \geq 5$ and $n(1-\hat{p}) \geq 5$
Proportion p, small n or extreme $\hat{p}$	Adjusted (plus-four) interval	Use $\tilde{p} = (X+2)/(n+4)$; no large-sample condition
Count $X \sim Bin(n, p)$	Normal approximation	$np \geq 5$ and $np(1-p) \geq 5$

Points total: 110.