Statistics is scored on communication: name the procedure, check the conditions against THIS study's numbers, compute, and conclude in context. These solutions show all four, because a correct number with any one missing loses points. 统计按四步给分: 方法、条件、计算、结合情境的结论。

A confidence interval

In a random sample of 400 voters, 232 support a proposal. Construct and interpret a 95% confidence interval for the population proportion. [4]

Procedure: one-sample z-interval for a proportion. Conditions: the sample was random; $n\hat{p} = 232$ and $n(1-\hat{p}) = 168$ are both at least 10; and 400 is less than 10% of all voters, so observations are independent. Calculation: $\hat{p} = 232/400 = 0.58$, and the interval is $0.58 \pm 1.96\sqrt{\frac{0.58(0.42)}{400}} = 0.58 \pm 0.0484 = (0.532, 0.628)$. Conclusion: we are 95% confident that the interval from 0.532 to 0.628 captures the true proportion of all voters who support the proposal.

Conditions must be CHECKED with this study's numbers, not merely listed. '$n\hat{p} = 232 \geq 10$' scores; 'large enough' does not.

The interpretation is about the INTERVAL capturing the parameter, not about a probability that the parameter lies in it. This is the most-lost point in the course.

Name the population and the variable. 'We are 95% confident the proportion is between...' with no context is incomplete.

A significance test

A company claims its bags contain 500 g. A random sample of 36 bags has mean 494 g and standard deviation 15 g. Test at the 5% level whether the mean is less than claimed. [5]

Hypotheses: $H_0 : \mu = 500$ against $H_a : \mu < 500$, where μ is the true mean contents of all bags. Procedure: one-sample t-test, since σ is unknown. Conditions: random sample; 36 is under 10% of production; $n = 36 \geq 30$ so the sampling distribution of the mean is approximately normal by the central limit theorem. Test statistic: $t = \frac{494 - 500}{15/\sqrt{36}} = \frac{-6}{2.5} = -2.40$ with 35 degrees of freedom, giving $p = 0.011$. Conclusion: since $0.011 < 0.05$, we reject H_0. There is convincing evidence that the true mean contents is less than 500 g.

Hypotheses are about the PARAMETER μ, not about $\bar{x}$. Writing $H_0 : \bar{x} = 500$ loses the point outright.

It is a t-test because σ is estimated from the sample. Using z here is a procedure error.

The conclusion links the p-value to α AND states what it means in context. 'Reject H_0' alone is half the point.

Interpret the p-value of 0.011 in this context. [2]

If the true mean contents were 500 g, there would be a 1.1% chance of obtaining a sample mean of 494 g or less in a random sample of 36 bags.

The conditional 'if the true mean were 500 g' is the mark. A p-value is never the probability that the hypothesis is true.

'Or less' matters – the direction must match the alternative hypothesis.

Design and inference

A study finds students who eat breakfast score higher on tests. Explain why this does not show that breakfast causes higher scores, and describe a study design that could. [4]

This is observational, so students chose whether to eat breakfast. A confounding variable such as household income could affect both breakfast frequency and test performance, and its effect cannot be separated from breakfast's. To show causation, use a randomised experiment: recruit volunteers, randomly ASSIGN half to receive breakfast and half not, hold other conditions constant, and compare mean test scores. Random assignment distributes confounding variables evenly across the groups, so a difference in outcome can be attributed to the treatment.

Name a plausible confounder specific to this context. 'There could be other factors' is not creditable.

Random ASSIGNMENT is what permits causation; random sampling permits generalisation. Confusing the two is the course's most examined error.