Verified exercise answer guide

Results from a fixed 1,000-student fictional population and 100-student SRS.

1. Define descriptive statistics

Answer: It summarizes and displays observed data: for example, the mean of the 100 sampled scores is 66.953. This alone is not a justified claim about all 1,000.

Common mistake: Do not treat a sample mean as a known population mean.

2. Define inferential statistics

Answer: Inference uses the probability sample and assumptions to estimate or test a population quantity. The sample mean is an estimate of the unknown population mean.

Common mistake: Do not assert certainty from a sample.

3. Identify the target population

Answer: The 1,000 generated fictional students are the finite target population for this teaching exercise.

Common mistake: A class population differs from a school population.

4. Identify the sample

Answer: There are 100 unique randomly selected StudentCode values, chosen without replacement from 1,000.

Common mistake: An arbitrary first 100 rows would not be the specified SRS.

5. Separate parameter from statistic

Answer: The known fictional population mean is 66.0142 (parameter); the 100-student sample mean is 66.953 (statistic).

Common mistake: Never use x-bar and μ as interchangeable symbols.

6. Calculate the eight-score mean

Answer: Scores sum to 600; mean = 600/8 = 75.

Common mistake: Do not divide by the number of distinct score values.

7. Find the eight-score median

Answer: Middle values are 70 and 75, so median = (70+75)/2 = 72.5.

Common mistake: For even n, there are two central values.

8. Find the eight-score mode

Answer: Score 70 occurs twice; every other score occurs once. The mode is 70.

Common mistake: The mode need not equal the mean or median.

9. Compute range

Answer: Range = 95 - 60 = 35 score points.

Common mistake: Range is sensitive to extremes.

10. Compute sample variance and SD

Answer: Sum of squared deviations is 900; sample variance = 900/7 = 128.5714; sample SD = 11.3389.

Common mistake: Population variance would be 900/8 = 112.5.

11. Explain simple random sampling

Answer: The fixed-seed generator selects 100 indices uniformly without replacement from 1,000; no index can repeat.

Common mistake: Do not call a convenience subset an SRS.

12. Explain sampling variability

Answer: The simulated mean of 2,000 sample means is 66.052; their empirical SD is 0.953 points.

Common mistake: A different seed gives a different realized simulation.

13. Calculate the finite-population standard error

Answer: With s=9.9852, n=100 and N=1000, estimated SE = 0.9478 points.

Common mistake: Do not confuse SE with SD or omit design assumptions.

14. Interpret the 95% confidence interval

Answer: Approximate 95% t interval with FPC: 65.07 to 68.83 points. Under assumptions, the procedure covers the fixed μ at approximately the nominal long-run rate.

Common mistake: Do not claim P(μ inside the observed fixed interval) = 0.95.

15. Define a two-sided test

Answer: H0: $\mu=65$; H1: $\mu\neq65$. Approximate FPC t statistic=2.061, df 99, two-sided model-based $p\approx0.0420$.

Common mistake: Not a causal claim or proof of a fixed null being true/false.

16. Explain the p-value

Answer: $p\approx0.0420$ is a null-model tail probability for the specified test statistic; it is not P(H0 is true).

Common mistake: p-value is not the probability the result is due to chance alone.

17. Significance versus practical value

Answer: A difference can produce a relatively small p-value yet be educationally trivial; report point difference, interval, data quality and study design.

Common mistake: Do not use 0.05 as an automatic policy rule.

18. Detect selection bias

Answer: Such a sample could systematically overestimate programme mean, even with a large n; increasing n alone does not fix coverage bias.

Common mistake: Random draw within a biased frame is not representative of the desired population.

19. Correlation versus causation

Answer: Yes: prior learning, resources and selection may influence both. An observed correlation does not by itself show that study time caused score differences.

Common mistake: Association is not causal identification.

20. Write a responsible conclusion

Answer: In a fictional population of 1,000, a 100-student SRS produced mean 66.95 points. Approximate design-aware 95% interval: 65.07–68.83; sampling and measurement assumptions apply.

Common mistake: Do not generalize to real learners, countries or schools.