Sampling and confidence intervals - Paper 6

10 questions | 60 marks | 75 minutes

Combine model checks, inverse calculations and applied evaluation in a demanding review. Use supplied two-sided critical values: 90% 1.645, 95% 1.96, 98% 2.326 and 99% 2.576. Retain intermediate precision; give final probabilities to four decimal places and state assumptions.

Equivalent correct methods earn the corresponding marks.

1. A researcher must sample 80 workers from shifts with populations 250, 150 and 100. [6 marks]

(a) Give proportional allocations and describe the random step. [3 marks]

Total = 500; multiply each population fraction by 80.

Answer: 40, 24 and 16; randomly select distinct labels within each complete shift list.

(b) A manager supplies only volunteers from each shift. Does retaining those same allocations remove self-selection bias? [3 marks]

Represented group sizes and the within-group selection mechanism are different issues.

Answer: No. Correct quotas do not make volunteer selection random within strata.

2. A right-skewed finite-variance population has mean 60 and SD 15. An independent sample has n = 225. [6 marks]

(a) Find the centre and standard error of the sample mean and state the normality status. [3 marks]

The individual population is not normal, so sample-mean normality is not exact.

Answer: Centre 60; SE 1; approximately normal if the CLT approximation is adequate.

(b) Use that model to approximate the probability that the mean is between 58.5 and 61.5. [3 marks]

Standardised bounds are -1.5 and 1.5.

Answer: 0.8664

3. A mean estimate is 75.6 and the known-spread standard error is 0.8. [6 marks]

(a) Find the 95% and 99% intervals. [3 marks]

Margins are 1.96 × 0.8 and 2.576 × 0.8.

Answer: 95% (74.032, 77.168); 99% (73.5392, 77.6608).

(b) A target of 77.3 lies outside one interval but inside the other. Explain without contradiction. [3 marks]

Compatibility depends on the chosen confidence level; neither interval proves a parameter exactly.

Answer: The wider 99% procedure allows more values to gain greater coverage.

4. A random representative sample of 800 people has 432 successes. [6 marks]

(a) Find the estimate and 95% proportion interval. [3 marks]

Estimated SE = $\sqrt{(0.54 \times 0.46 / 800)}$.

Answer: Estimate 0.54; interval (0.5055, 0.5745).

(b) Assess a claim of more than half the population under the interval model. [3 marks]

State approximate coverage and design assumptions rather than certainty.

Answer: The 95% interval lies above 0.5, supporting a majority under the assumptions.

5. A data collector expects 60% of contacted residents to give a valid response. It needs a conservative 95% proportion margin at most 0.035. [6 marks]

(a) Find the required number of valid responses. [3 marks]

Use $p_0 = 0.5$: bound = 784; hence 784 valid responses.

Answer: 784

(b) Find a contact-budget estimate and explain one reason it is not a guarantee. [3 marks]

Divide required valid responses by 0.6, then round up.

Answer: 1307 contacts. Response rate and composition may differ from expectations.

6. A mean study has known population SD 10 and n = 100. A new population has SD 15. [6 marks]

(a) Find the sample size needed in the new population to keep standard error unchanged. [3 marks]

Original SE = 1. Require $15/\sqrt{n} = 1$.

Answer: 225.

(b) If confidence also rises from 95% to 99%, will the same n keep interval width unchanged? Explain. [3 marks]

A further sample-size increase would be needed to offset the critical-value ratio.

Answer: No; the larger critical value widens the interval even though SE is unchanged.

7. A trial records one success in 40 independent observations. [6 marks]

(a) Calculate the estimate and explain why the usual symmetric interval is questionable. [3 marks]

The count condition fails, and near-boundary intervals can have bad coverage.

Answer: Estimate 0.025; only one observed success, so normal/Wald approximation is poor.

(b) Would a zero-success follow-up prove the population probability is zero? [3 marks]

A suitable binomial inference method is needed; a zero estimated SE is not proof.

Answer: No. Finite samples can miss a rare success.

8. A claimed 95% procedure is applied repeatedly under its own valid assumptions. [6 marks]

(a) In 2,000 independent repetitions, give expected misses and explain the role of random variation. [3 marks]

$2{,}000 \times 0.05 = 100$.

Answer: 100 expected misses; realised misses vary.

(b) A reader says each realised interval has a fixed parameter that is “95% random”. Correct the language. [3 marks]

After observing an interval, containment is true or false, although unknown to the analyst.

Answer: The parameter is fixed; the procedure’s random intervals capture it about 95% of the time.

9. A report compares two treatments using separate intervals: A (20.1,21.3), B (20.7,21.9). [6 marks]

(a) Find the estimates and margins. [3 marks]

Use midpoints and half-widths.

Answer: A estimate 20.7; B estimate 21.3; margins both 0.6.

(b) Explain why overlap alone establishes neither equality nor a formal conclusion about the difference. [3 marks]

Separate coverage intervals do not directly calculate uncertainty in the difference.

Answer: The difference is another estimator requiring its own sampling model.

10. A large, carefully randomised 2026 sample estimates average screen time with a narrow interval. Measurements come from self-reported recollections. [6 marks]

(a) Name a remaining uncertainty not captured by the sampling margin. [3 marks]

Random sampling does not guarantee accurate reported measurements.

Answer: Measurement/recall bias.

(b) Give a responsible reporting sentence and one possible validation improvement. [3 marks]

Keep sampling precision separate from measurement accuracy and avoid implying all error is in the margin.

Answer: Report the estimate and interval with the population/date and self-report limitation; validate a consented subsample against a consistent recorded measure.