AP STATISTICS · MIDTERM REVISION 2026

Midterm Revision C (Variant B)

Probability, Sampling Distributions and Inference

Normal calculations, sampling distributions and proportions — varied questions with condensed solutions

20	3	72
questions	themes: normal, sampling, proportion	slides in total

A varied variant of the Midterm Revision C lesson — new values, new order, condensed wording

How to use this lesson

Key points first, then question, then solution

Key points — Every section opens with a summary slide of the ideas you need, main point first.

Question — Each question sits on its own slide with the mark tariff in the corner — attempt it before turning the page.

Solution — The next slide holds the condensed solution: the reasoning that earns the mark, not just the answer.

Trap box — Each solution ends with the mistake that most often costs marks on that question.

Plenary — Eight True/False statements close the lesson — one per slide, each answer on the slide that follows.

Today's roadmap

Three pillars carry every inference question

1

Normal calculations

z-scores, percentiles and probabilities for individual values

2

Sampling distributions

Centre, spread and shape for means and proportions — plus large counts

3

Proportion inference

Intervals, margin of error and minimum sample sizes

4

Reading the curves

Interpreting sampling distribution pictures — centre, spread, tails

01

Normal calculations

z-scores · percentiles · probabilities for individuals

- Questions 1 to 4
- Standardise, then read the probability

Key points · normal calculations

Standardise first: convert to a z-score, then find the probability

1	**The z-score**	Standardise with $z = (x - \mu)/\sigma$: the number of SDs above (+) or below (−) the mean. Start every normal problem here.
2	**Percentiles**	For the 90th percentile take $z \approx 1.28$ (90% of area to the left), then back-transform $x = \mu + z\sigma$.
3	**Probability**	Area under the normal curve = probability. Find it with calculator normalcdf or a z-table.
4	**Unusual observations**	Call a value unusual when $\lvert z \rvert \geq 2$ — roughly 95% of data sit within 2 SDs of the mean.
5	**SD of the sample mean**	Averages of n values vary less than individuals: the spread is $\sigma/\sqrt{n}$.

Formulae · normal distributions

formulae

Everything you need for Questions 1 to 4

z-score	$z = (x - \mu) / \sigma$ (standard deviations from the mean)
Back-transform	$x = \mu + z\sigma$ (from percentile to value)
Percentile z-values	90th: 1.28 · 95th: 1.645 · 97.5th: 1.96 · 99th: 2.33
SD of sample mean	$\sigma(\bar{x}) = \sigma / \sqrt{n}$ (spread shrinks as n grows)
Empirical rule	68% within 1 SD · 95% within 2 SD · 99.7% within 3 SD

Sketch the curve and shade the area first: the z-score standardises, the shaded area answers the question.

Question 1

1 mark

Normal calculations

THE QUESTION

Scores on a national maths challenge are approximately normal with mean 72 points and standard deviation 9 points. Which score is closest to the 90th percentile of the distribution?

(A) 81 points

(B) 84 points

(C) 88 points

(D) 90 points

Question 1 — solution

1 mark

Normal calculations

CONDENSED SOLUTION

- Answer: (B).
- 90th percentile: $z \approx 1.28$.
- $x = \mu + z\sigma = 72 + 1.28 \times 9 = 83.5 \approx 84$ points.

TRAP Know the anchor z-values. $1.28 \times 9 = 11.5$ sits clearly nearer 84 than 81 — compute, never guess between options.

Question 2

1 mark

Normal calculations

THE QUESTION

Daily commute times in a town have mean 27.5 minutes and standard deviation 5.2 minutes. A planner will take a random sample of 60 commuters and test whether the mean commute exceeds 28.5 minutes. Assuming independence, which of the following describes the sampling distribution of the sample mean commute time for samples of size 60?

(A) Approximately normal with mean 27.5 minutes and standard deviation 5.2 minutes

(B) Approximately normal with mean 27.5 minutes and standard deviation $5.2/\sqrt{60}$ minutes

(C) Approximately normal with mean 28.5 minutes and standard deviation 5.2 minutes

(D) Binomial with mean $60(0.47)$ minutes and standard deviation $\sqrt{(60(0.47)(0.53))}$ minutes

Question 2 — solution

1 mark

Normal calculations

CONDENSED SOLUTION

- Answer: (B).
- The sampling distribution of $\bar{x}$ is centred at $\mu = 27.5$ with SD $\sigma/\sqrt{n} = 5.2/\sqrt{60} \approx 0.67$ minutes.
- Normal shape is fine: $n = 60$ is large, so the CLT applies even if individual commutes are skewed.
- (C) shifts the centre to the alternative value; (D) is binomial — commute times are not success counts.

TRAP Means keep the population MEAN and shrink the SPREAD by $\sqrt{n}$. The centre never moves to 28.5.

Question 3

1 mark

Normal calculations

THE QUESTION

A ferry pontoon has a load limit of 3,600 kg and takes up to 18 passengers at a time. Passenger weights are approximately normal with $\mu = 185$ kg and $\sigma = 30$ kg. Which value best approximates the probability that a random group of 18 passengers exceeds the 3,600 kg limit?

(A) 0.004

(B) 0.017

(C) 0.098

(D) 0.221

Question 3 — solution

1 mark

Normal calculations

CONDENSED SOLUTION

- Answer: (B).
- Total weight of 18 passengers: mean 18 × 185 = 3,330 kg, SD √18 × 30 ≈ 127.3 kg.
- z = (3600 − 3330)/127.3 ≈ 2.12. P(Z > 2.12) ≈ 0.017.
- Equivalently: limit per person 3600/18 = 200 kg, z = (200 − 185)/(30/√18) ≈ 2.12 — the same answer.

TRAP Use the SUM (mean nμ, SD √nσ) or the MEAN (μ, σ/√n) — never the individual distribution alone.

Question 4

1 mark

Normal calculations

THE QUESTION

The normal curve shown represents the sampling distribution of a sample mean for samples of size n = 36, drawn at random from a population with standard deviation σx. The inflection points sit at 22 and 46 on the axis, so the centre is 34 and the standard error is about 12.
Which of the following is the best estimate of σx?
(A) 12
(B) 34
(C) 72
(D) 216

Question 4 — solution

1 mark

Normal calculations

CONDENSED SOLUTION

- Answer: (C).
- The standard error (SD of the sampling distribution) is about 12: $\sigma x/\sqrt{n} = \sigma x/6$.
- Solve: $\sigma x = 12 \times \sqrt{36} = 72$.

TRAP Centre-to-inflection distance = standard error. Multiply by √n to recover σ.

02

Sampling distributions

Means and proportions · centre · spread · large counts

- Questions 5 to 12
- Centre never moves; spread shrinks by √n; shape needs conditions

Key points · sampling distributions

Centre never changes — only spread and shape depend on the sample

1	**Centre**	Mean of $\bar{x}$ = μ and mean of $\hat{p}$ = p — sample size never shifts the centre of a sampling distribution.
2	**Spread**	SD of $\bar{x}$ = σ/√n and SD of $\hat{p}$ = √(p(1−p)/n): bigger samples, tighter distributions.
3	**Shape for means**	$\bar{x}$ is approximately normal when the population is normal or n ≥ 30 (CLT); skewed populations need larger n.
4	**Shape for proportions**	$\hat{p}$ is approximately normal when BOTH np ≥ 10 and n(1−p) ≥ 10 — the large counts condition.
5	**Simulations**	Simulated distributions centre on the true value: smaller samples add spread, never a new centre.

Formulae · sampling distributions

formulae

Everything you need for Questions 5 to 12

Mean of $\hat{p}$	mean of the sampling distribution of $\hat{p} = p$ (centre never moves)
SD of $\hat{p}$	$\sigma(\hat{p}) = \sqrt{\frac{p(1-p)}{n}}$
Large counts	$np \geq 10$ AND $n(1-p) \geq 10 \rightarrow \hat{p}$ approximately normal
Mean and SD of $\bar{x}$	mean $= \mu$; $\sigma(\bar{x}) = \sigma / \sqrt{n}$
Difference of means	mean of $\bar{x}1 - \bar{x}2 = \mu1 - \mu2$; SD $= \sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}$

Run the large counts check with the CLAIMED population p before any data exist — never the sample proportion.

Question 5

1 mark

Sampling distributions

THE QUESTION

For a sample of 55 coffee shops, the mean queue time is 2.1 minutes and the standard deviation of queue times is 1.2 minutes. Which of the following is most likely true about queue times in this sample?

(A) The distribution is approximately normal because the sample size is 55, so the central limit theorem applies.

(B) The distribution is approximately normal because the standard deviation is less than the mean.

(C) The distribution is skewed to the right because the least possible value is within 2 standard deviations of the mean.

(D) The distribution is skewed to the left because the least possible value is within 2 standard deviations of the mean.

Question 5 — solution

1 mark

Sampling distributions

CONDENSED SOLUTION

- Answer: (C).
- Queue times cannot be negative. Zero sits (2.1 − 0)/1.2 ≈ 1.75 SDs below the mean — within 2 SDs.
- Data pushed against a hard floor at zero while the mean sits at 2.1 are typically right-skewed.
- (A) misapplies the CLT to raw data; (B) is a non-sequitur.

TRAP The CLT describes sample MEANS, never raw data. Floor-at-zero plus modest SD is the classic right-skew setup.

Question 6

1 mark

Sampling distributions

THE QUESTION

Country P has 9 million commuters; Country Q has 0.4 million. Researchers believe the proportion of Country Q commuters who work from home is between 10% and 30%, and the proportion in Country P is between 15% and 40%. Independent random samples will be taken, one from each country. Based on these population proportions, which sample sizes guarantee that the sampling distribution of the difference in sample proportions is approximately normal?

(A) 40 in Country P and 70 in Country Q

(B) 70 in Country P and 40 in Country Q

(C) 70 in Country P and 100 in Country Q

(D) 100 in Country P and 70 in Country Q

Question 6 — solution

1 mark

Sampling distributions

CONDENSED SOLUTION

- Answer: (C).
- Worst cases: Country P needs $n \times 0.15 \geq 10 \rightarrow n \geq 66.7$, so 70. Country Q needs $n \times 0.10 \geq 10 \rightarrow n \geq 100$.
- Only option (C), 70 and 100, meets both (and the failure counts $n(1-p)$ are far larger).

TRAP Check the SMALLEST plausible proportion for each sample — the worst case binds, not the middle of the range.

Question 7

1 mark

Sampling distributions

THE QUESTION

In a population of adults, 15 percent cycle to work. A random sample of 60 adults will be selected. What is the standard deviation of the sampling distribution of the sample proportion of adults who cycle to work for samples of size 60?
(A) $\sqrt{(0.15(1 - 0.15)/60)}$
(B) $\sqrt{(15(60 - 15)/100)}$
(C) $\sqrt{(0.15(1 - 0.15))}$
(D) $0.15(1 - 0.15)/60$

Question 7 — solution

1 mark

Sampling distributions

CONDENSED SOLUTION

- Answer: (A).
- SD of $\hat{p}$ = $\sqrt{p(1-p)/n}$ with p = 0.15 and n = 60: $\sqrt{0.15 \times 0.85/60} \approx 0.046$.
- Convert 15% to 0.15 and divide by n — the square root covers the whole fraction.

TRAP Convert 15% to 0.15 first and keep the denominator as n = 60. Percents have no place in this formula.

Question 8

1 mark

Sampling distributions

THE QUESTION

A school canteen recycles 85 percent of its food waste. A random sample of 120 days will be audited. Which of the following best describes the shape of the sampling distribution of the sample proportion of days on which the target was met?

(A) Bimodal

(B) Approximately normal

(C) Strongly skewed to the left

(D) Strongly skewed to the right

Question 8 — solution

1 mark

Sampling distributions

CONDENSED SOLUTION

- Answer: (B).
- Large counts with $p = 0.85$, $n = 120$: $np = 102 \geq 10$ and $n(1-p) = 18 \geq 10$.
- Both pass, so the sampling distribution of $\hat{p}$ is approximately normal.

TRAP Run BOTH large counts. A high success rate alone does not create skew — n = 120 keeps the failure count at 18.

Question 9

1 mark

Sampling distributions

THE QUESTION

Two phone manufacturers, A and B, sell flagship models. About 6 percent of buyers of either brand return the phone within a year. A random sample of 120 buyers of brand A and a random sample of 250 buyers of brand B are selected. Are the sample sizes large enough to treat the sampling distribution of the difference in sample proportions as approximately normal?

(A) Yes. Both sample sizes are large enough.

(B) No. The sample size for brand A is large enough, but the sample size for brand B is not.

(C) No. The sample size for brand B is large enough, but the sample size for brand A is not.

(D) There is not enough information to decide whether the sampling distribution is approximately normal.

Question 9 — solution

1 mark

Sampling distributions

CONDENSED SOLUTION

- Answer: (C).
- Brand A: np = 120 × 0.06 = 7.2 < 10 — fails. Brand B: np = 250 × 0.06 = 15 ≥ 10 — passes (failure counts are far above 10 for both).
- A difference of proportions needs BOTH samples to pass, so overall the answer is no.

TRAP Small success probabilities make the success count binding. One failing sample sinks the whole difference.

Question 10

1 mark

Sampling distributions

THE QUESTION

The distribution of the number of years residents have lived in a village is moderately skewed to the right, with mean 9.1 years and standard deviation 4.7 years. A researcher will survey a random sample of 90 residents. Which of the following best describes the sampling distribution of the sample means of years lived for samples of size 90?

(A) Moderately skewed to the right, with mean 9.1 years and standard deviation 4.7 years

(B) Moderately skewed to the right, with mean 9.1 years and standard deviation less than 4.7 years

(C) Approximately normal, with mean less than 9.1 years and standard deviation 4.7 years

(D) Approximately normal, with mean 9.1 years and standard deviation less than 4.7 years

Question 10 — solution

1 mark

Sampling distributions

CONDENSED SOLUTION

- Answer: (D).
- n = 90 is large, so the CLT washes out the moderate skew: $\bar{x}$ is approximately normal.
- Centre: mean 9.1 (unchanged). Spread: $4.7/\sqrt{90} \approx 0.50$, which is less than 4.7.

TRAP The CLT trade: large n buys normal SHAPE and shrinks SPREAD, but the centre stays at μ. Options keeping the skew ignore the CLT.

Question 11

1 mark

Sampling distributions

THE QUESTION

In a study, the proportion of adults who keep a daily habit tracker was 0.42. Two simulations will model the sampling distribution of a sample proportion from a population with true proportion 0.42. Simulation A will run 2,500 trials with a sample size of 80; simulation B will run 1,800 trials with a sample size of 40.

Which of the following describes the centre and variability of simulation A and simulation B?

(A) The centres will roughly be equal, and the variabilities will roughly be equal.

(B) The centres will roughly be equal, and the variability of simulation A will be greater than the variability of simulation B.

(C) The centres will roughly be equal, and the variability of simulation A will be less than the variability of simulation B.

(D) The centre of simulation A will be greater than the centre of simulation B, and the variabilities will roughly be equal.

Question 11 — solution

1 mark

Sampling distributions

CONDENSED SOLUTION

- Answer: (C).
- Centre: both simulations draw from p = 0.42, so both centre near 0.42 — the trial counts (2,500 vs 1,800) do not move it.
- Variability: SD = √(p(1−p)/n). Simulation A has n = 80, B has n = 40, so A's SD (≈ 0.055 vs ≈ 0.078) is SMALLER — less variability.

TRAP Two different roles: the number of TRIALS sets how smooth the simulation looks; the sample SIZE sets the spread.

Question 12

1 mark

Sampling distributions

THE QUESTION

A leisure centre compares gym habits of members and non-members. Let pM be the population proportion of members who exercise at least four times a week and pN the population proportion of non-members who do. Random samples of 150 members and 150 non-members will be selected. Which of the following is the best interpretation of $P(\hat{p}M - \hat{p}N > 0) \approx 0.041$?

(A) The difference between the population proportions is approximately 0.041.

(B) For all random samples of 150 members and 150 non-members, the difference between the sample proportions will be 0.041.

(C) For all random samples of 150 members and 150 non-members, the difference between the sample proportions will be 0.

(D) For all random samples of 150 members and 150 non-members, the sample proportion of members exercising four or more times a week will exceed the non-members' sample proportion in about 4.1% of samples.

Question 12 — solution

1 mark

Sampling distributions

CONDENSED SOLUTION

- Answer: (D).
- P($\hat{p}_M - \hat{p}_N > 0$) is a probability about SAMPLES: in roughly 4.1% of all random samples of this size, the members' sample proportion would exceed the non-members'.
- (A) confuses a probability with a parameter difference; (B) and (C) misread a probability as a fixed value.

TRAP Probability statements describe what happens across random samples — “in about X% of samples”. Never turn them into parameter claims.

03

Proportion inference

Intervals · margin of error · minimum sample size

- Questions 13 to 18
- Margin of error shrinks with √n — plan samples backwards from it

Key points · proportion inference

Every inference answer lives in context: parameter, method, conclusion

1	**Margin of error**	ME = $z^* \cdot \sigma/\sqrt{n}$ for means, $z^* \cdot \sqrt{p(1-p)/n}$ for proportions; confidence is the method's capture rate.
2	**Minimum sample size**	Set $n \geq (z^*\sigma/ME)^2$ and ROUND UP — rounding down lets the margin drift past the target.
3	**Effect of n on variance**	Sampling variance scales as 1/n: to quarter the variance, collect four times the data.
4	**Interpreting the mean of $\hat{p}$**	The centre statement: the mean of all sample proportions, over all random samples of size n, equals p.
5	**Hypothesis-test curves**	The test curve is centred at the CLAIMED p with SD $\sqrt{p(1-p)/n}$; the observed $\hat{p}$ is placed on it.

Formulae · inference

formulae

Everything you need for Questions 13 to 18

Confidence interval	estimate ± margin of error — e.g. $\hat{p} \pm z^*\sqrt{\hat{p}(1-\hat{p})/n}$
Margin of error (known σ)	$ME = z^* \cdot \sigma / \sqrt{n}$ 90%: $z^* = 1.645$ · 95%: 1.96
Minimum sample size	$n = (z^* \sigma / ME)^2$ → always ROUND UP
z for one value	$z = (x - \mu) / \sigma$ — then normalcdf for the probability
Normality check (p̂)	$np \geq 10$ and $n(1-p) \geq 10$ with the claimed p

Sample-size questions run the algebra backwards: start from the tolerable margin of error and solve for n.

Question 13

1 mark

Proportion inference

THE QUESTION

A council report states that 27 percent of households in a district own a heat pump. For random samples of 75 households, which of the following best describes the sampling distribution of the sample proportion of households that own a heat pump?

(A) The sampling distribution is approximately normal, with mean 0.27 and standard deviation 0.051.

(B) The sampling distribution is not necessarily approximately normal, with mean 0.27 and standard deviation 0.051.

(C) The sampling distribution is approximately normal, with mean 20.25 and standard deviation 0.0026.

(D) The sampling distribution is not necessarily approximately normal, with mean 20.25 and standard deviation 0.0026.

Question 13 — solution

1 mark

Proportion inference

CONDENSED SOLUTION

- Answer: (A).
- Centre and spread: mean p = 0.27 and SD = √(0.27 × 0.73/75) ≈ 0.051 — options (A) and (B) share them, so the decision is normality.
- Large counts: np = 75 × 0.27 = 20.25 ≥ 10 and n(1−p) = 75 × 0.73 ≈ 54.8 ≥ 10. Both pass, so the distribution IS approximately normal.
- (C) and (D) use the count 20.25 as a mean and the variance 0.0026 as an SD — wrong units.

TRAP Run BOTH counts: 75×0.27 = 20.25 and 75×0.73 ≈ 54.8. Both ≥ 10, so the distribution IS approximately normal — answer (A).

Question 14

1 mark

Proportion inference

THE QUESTION

In a region, about 55 percent of adults hold a season ticket for the local football club. A random sample of 250 adults and a random sample of 100 adults will be selected. For which sample size is it more likely that more than 65 percent of the sample hold a season ticket?

(A) It is more likely for a sample size of 250, because there is more variability in the sampling distribution for larger sample sizes.

(B) It is more likely for a sample size of 250, because there is less variability in the sampling distribution for larger sample sizes.

(C) It is more likely for a sample of size 100, because there is more variability in the sampling distribution for smaller sample sizes.

(D) It is more likely for a sample size of 100, because there is less variability in the sampling distribution for smaller sample sizes.

Question 14 — solution

1 mark

Proportion inference

CONDENSED SOLUTION

- Answer: (C).
- 65% is ABOVE the population value of 55%. Reaching it needs a big swing from p — more likely when the sampling distribution has MORE spread.
- Smaller samples have larger SD: $\sqrt{(0.55\times0.45/100)} \approx 0.050$ vs $\sqrt{(0.55\times0.45/250)} \approx 0.031$. The n = 100 distribution is wider, so extreme values are more likely.

TRAP Cut-offs beyond the centre are reached more often by wider distributions. Smaller n = wider = more likely to land far from p.

Question 15

1 mark

Proportion inference

THE QUESTION

A transport study reports that 22% of commuters in a city take a bicycle to work. Random samples of 150 commuters and 400 commuters are selected from this population. For which sample size — 150 or 400 — is the probability lower that fewer than 18% of the sampled commuters cycle to work, and why?

(A) The probability is the same for both sample sizes, because sample size does not affect that outcome.

(B) The probability is lower for the sample of 150, because 150 is smaller than 400.

(C) The probability is lower for the sample of 150. A smaller sample gives a larger standard deviation, a z-score closer to zero, and so a smaller tail probability.

(D) The probability is lower for the sample of 400. A larger sample gives a smaller standard deviation, a larger $|z|$, and so a smaller tail probability.

Question 15 — solution

1 mark

Proportion inference

CONDENSED SOLUTION

- Answer: (D).
- 18% sits BELOW p = 0.22. With n = 400, SD = √(0.22×0.78/400) ≈ 0.021 vs ≈ 0.034 for n = 150.
- Smaller SD → z = (0.18 − 0.22)/SD is more extreme (−1.93 vs −1.18) → less area below 0.18. Lower probability for n = 400.
- (C) has the mechanism backwards: a larger SD gives a z closer to zero, which INCREASES the tail probability.

TRAP Tail probabilities below a fixed cut-off shrink as the distribution tightens. Trace z = (cut-off − p)/SD for both n values.

Question 16

1 mark

Proportion inference

THE QUESTION

A student union publishes the average term-time rent in the city each year. It will collect a random sample of listings and estimate the mean rent with a 95 percent confidence interval. Over recent years the standard deviation has stayed close to £210. Assuming the population standard deviation is £210, what is the minimum sample size needed for a margin of error of no more than £25?

(A) 196

(B) 245

(C) 272

(D) 314

Question 16 — solution

1 mark

Proportion inference

CONDENSED SOLUTION

- Answer: (C).
- $n \geq (z^*\sigma/ME)^2 = (1.96 \times 210/25)^2 = (16.464)^2 \approx 271.1$.
- Round UP: n = 272. Rounding down to 271 would let the margin of error drift just past £25.

TRAP Sample size always rounds UP. The inequality is n ≥ value, so the smallest whole number satisfying it is 272.

Question 17

1 mark

Proportion inference

THE QUESTION

A survey of 120 randomly selected veterinary clinics in Country Z found that 70% stock a certain pet food brand. The population proportion is known to be p = 0.65. For samples of size 120, which of the following best interprets the mean of the sampling distribution of the sample proportion of clinics stocking the brand?

(A) For all random samples of 120 clinics in Country Z, the sample proportion will be 0.65.

(B) For all random samples of 120 clinics in Country Z, the sample proportion will be 0.70.

(C) The mean of all sample proportions from all random samples of 120 clinics in Country Z is equal to 0.65.

(D) The mean of all sample proportions from all random samples of 120 clinics in Country Z is equal to 0.70.

Question 17 — solution

1 mark

Proportion inference

CONDENSED SOLUTION

- Answer: (C).
- The mean of the sampling distribution of $\hat{p}$ equals the POPULATION proportion p = 0.65 — not the observed sample value 0.70.
- (A) and (B) wrongly claim every sample lands on the centre. Samples vary; the AVERAGE across samples is 0.65.

TRAP Centre of a sampling distribution = parameter (0.65). The 0.70 is one realised sample — it never defines the centre.

Question 18

1 mark

Proportion inference

THE QUESTION

A rail company claims that 25 percent of its morning trains arrive late. A passengers' group doubts the figure. The group records a random sample of 80 morning trains and finds that 28 of them arrived late. To test whether the true proportion differs from 25 percent, which sampling distribution should be used?

(A) A normal curve centred at 0.25 with standard deviation $\sqrt{(0.25 \times 0.75/80)} \approx 0.048$, with 0.35 in its upper tail

(B) A normal curve centred at 0.35 with standard deviation $\sqrt{(0.35 \times 0.65/80)} \approx 0.053$

(C) A normal curve centred at 0.35 with standard deviation $\sqrt{(0.25 \times 0.75/80)} \approx 0.048$

(D) A normal curve centred at 0.25 with standard deviation $\sqrt{(0.35 \times 0.65/80)} \approx 0.053$

Question 18 — solution

1 mark

Proportion inference

CONDENSED SOLUTION

- Answer: (A).
- Under the hypothesis, the sampling distribution of $\hat{p}$ is centred at the CLAIMED p = 0.25 with SD $\sqrt{(0.25 \times 0.75/80)} \approx 0.048$.
- The observed 0.35 is then located in the tail — that tail area is the evidence against the claim.
- Mixing the observed 0.35 into the centre or the SD (B, C, D) double-counts the data.

TRAP Build the null model from the CLAIMED parameter. The data (0.35) sit ON that model, never inside it.

Question 19

5 marks

Inference FRQ · consolidated

THE QUESTION

A school claims that 55% of students eat a cooked breakfast on exam days. A researcher selects a random sample of 160 students and asks each one.

(a) Check whether the sampling distribution of $\hat{p}$ is approximately normal for samples of size 160.
(b) Calculate the mean and standard deviation of the sampling distribution of $\hat{p}$.
(c) Find the probability that in a random sample of 160 students, fewer than 50% ate a cooked breakfast.
(d) The researcher wants the standard deviation of $\hat{p}$ to be no more than 0.035. Determine the minimum sample size required.

Question 19 — solution

5 marks

Inference FRQ · consolidated

CONDENSED SOLUTION

- (a) $np = 160 \times 0.55 = 88 \geq 10$ and $n(1-p) = 160 \times 0.45 = 72 \geq 10$ (1). Both large counts pass, so $\hat{p}$ is approximately normal.
- (b) Mean = 0.55; SD = $\sqrt{(0.55 \times 0.45/160)} \approx 0.0393$ (1).
- (c) $z = (0.50 - 0.55)/0.0393 \approx -1.27$; $P(Z < -1.27) \approx 0.1018$, so about a 10.2% chance (2: 1 method, 1 answer).
- (d) $\sqrt{(0.2475/n)} \leq 0.035 \rightarrow n \geq 0.2475/0.001225 \approx 202.04 \rightarrow n = 203$ students (1).

TRAP State the large counts check with numbers before any normal probability. For sample size, rearrange the SD formula and round up.

Question 20

5 marks

Inference FRQ · consolidated

THE QUESTION

The lifetimes of rechargeable batteries from a factory are skewed to the right with mean 12.5 hours and standard deviation 1.6 hours. A quality engineer tests a random sample of 64 batteries and computes the sample mean lifetime.

(a) Describe the sampling distribution of the sample mean (shape, centre, spread), with justification.
(b) Find the probability that the sample mean exceeds 12.8 hours.
(c) Explain why the probability in (b) is much smaller than the probability that a single randomly selected battery lasts more than 12.8 hours.
(d) The factory improves its process and the standard deviation falls to 1.2 hours. Without recomputing fully, state and justify the effect on the probability that the sample mean of 64 batteries exceeds 12.8 hours.

Question 20 — solution

5 marks

Inference FRQ · consolidated

CONDENSED SOLUTION

- (a) Approximately normal (1) by the CLT: n = 64 ≥ 30 even though the population is skewed. Centre μ = 12.5 hours; SD = 1.6/√64 = 0.2 hours (1).
- (b) z = (12.8 − 12.5)/0.2 = 1.50; P(Z > 1.50) ≈ 0.0668, about 6.7% (1).
- (c) A single battery has SD 1.6, so 12.8 is only z ≈ 0.19 above the mean — common. Averages of 64 have SD 0.2, so the same cut-off is z = 1.50 — much rarer. Averaging cancels individual extremes (1).
- (d) The probability gets SMALLER (1): with σ = 1.2, the SD of $\bar{x}$ drops to 0.15 hours, so 12.8 is now z = 2.00 — an even smaller upper-tail area.

TRAP Shape justifications must name the mechanism (CLT with n ≥ 30, or normal population). Spread of the mean is always σ/√n, never σ.

Mark scheme · multiple choice

One mark each · 18 MCQ marks

Q1	**B**	72 + 1.28·9 = 83.5 ≈ 84 points
Q2	**B**	Mean 27.5, SD 5.2/√60 ≈ 0.67, normal by CLT (n = 60)
Q3	**B**	z = (3600 − 3330)/127.3 ≈ 2.12 → 0.017
Q4	**C**	σx = 12 × √36 = 72
Q5	**C**	Zero within 2 SD of mean 2.1 → right skew
Q6	**C**	P: n·0.15 ≥ 10 → 70; Q: n·0.1 ≥ 10 → 100
Q7	**A**	√(0.15×0.85/60) ≈ 0.046
Q8	**B**	np = 102, n(1−p) = 18 → approximately normal
Q9	**C**	Brand A 7.2 < 10 fails; brand B 15 ≥ 10 passes
Q10	**D**	CLT: normal, mean 9.1, SD ≈ 0.50 < 4.7
Q11	**C**	Same centre 0.42; n = 80 less variable than n = 40
Q12	**D**	In about 4.1% of samples $\hat{p}_M > \hat{p}_N$
Q13	**A**	np = 20.25 and n(1−p) ≈ 54.8 both ≥ 10 → normal
Q14	**C**	Smaller n = wider distribution = extremes more likely
Q15	**D**	n = 400: smaller SD, larger |z| (−1.93 vs −1.18), lower probability
Q16	**C**	$(1.96\times210/25)^2 = 271.1$ → round up to 272
Q17	**C**	Mean of all $\hat{p}$ = population p = 0.65
Q18	**A**	Null model: centred at claimed 0.25, SD 0.048

Multiple choice total: 18 marks. FRQ Q19 + Q20: 10 marks.

Mark scheme · free response

Q19 and Q20 — mark by part, award for reasoning

Q19a	**1**	1 for both large counts (88 and 72) ≥ 10 → approximately normal
Q19b	**1**	1 for mean 0.55 and SD ≈ 0.0393
Q19c	**2**	1 for z ≈ −1.27 · 1 for probability ≈ 0.102 (10.2%)
Q19d	**1**	1 for n = 203 with working from SD ≤ 0.035
Q20a	**2**	1 for approximately normal with CLT justification · 1 for mean 12.5, SD = 0.2
Q20b	**1**	1 for z = 1.50 and probability ≈ 0.067
Q20c	**1**	1 for comparing individual SD 1.6 (z ≈ 0.19) with SD of mean 0.2 (z = 1.50)
Q20d	**1**	1 for probability decreases: SE 1.2/√64 = 0.15, so z = 2.00 — smaller upper-tail area

Free response total: 10 marks. Lesson total: 28 marks (18 MCQ + 10 FRQ).

Answer key

Fast self-check across all questions

MULTIPLE CHOICE

- Q1 B
- Q2 B
- Q3 B
- Q4 C
- Q5 C
- Q6 C
- Q7 A
- Q8 B
- Q9 C
- Q10 D
- Q11 C
- Q12 D
- Q13 A
- Q14 C
- Q15 D
- Q16 C
- Q17 C
- Q18 A

FREE RESPONSE

- Q19 large counts pass; SD ≈ 0.0393; $P(\hat{p} < 0.50) \approx 0.102$; n = 203
- Q20 normal by CLT; SD = 0.2; $P(\bar{x} > 12.8) \approx 0.067$;
- individual $z \approx 0.19$ vs mean $z = 1.50$;
- smaller σ → smaller standard error → smaller probability

True or false · 1

T / F

Decide, then justify your decision — answer on the next slide

For the sampling distribution of the sample mean, the centre equals the population mean μ, whatever the sample size.

True or false · 1 — answer

TRUE

TRUE

Centre never moves: mean of $\bar{x} = \mu$. Sample size only shrinks the spread ($\sigma/\sqrt{n}$) and improves the shape.

True or false · 2

T / F

Decide, then justify your decision — answer on the next slide

The standard deviation of the sampling distribution of $\hat{p}$ is $\sqrt{p(1-p)/n}$, so samples of 90 give three times the standard deviation of samples of 10.

True or false · 2 — answer

FALSE

FALSE

The ratio of SDs is √(90/10) ≈ 3 in the OTHER direction: samples of 10 vary √9 = 3 times as much as samples of 90. SD scales with 1/√n.

True or false · 3

T / F

Decide, then justify your decision — answer on the next slide

If np = 7.5 < 10 for one sample while the other sample passes the large counts condition, the sampling distribution of the difference in sample proportions is still approximately normal.

True or false · 3 — answer

FALSE

FALSE

Both samples must pass np ≥ 10 and n(1−p) ≥ 10. One failure — like the brand A sample with np = 7.2 — blocks the normal model.

True or false · 4

T / F

Decide, then justify your decision — answer on the next slide

The sampling distribution of $\hat{p}$ for $n = 120$ and $p = 0.65$ is approximately normal because $np = 78$ and $n(1-p) = 42$ both exceed 10.

True or false · 4 — answer

TRUE

TRUE

Both large counts pass comfortably, so normality holds — the heat-pump question is exactly this check.

True or false · 5

T / F

Decide, then justify your decision — answer on the next slide

To guarantee a margin of error of no more than £25 with σ = 210 at 95% confidence, a sample of 271 listings is sufficient.

True or false · 5 — answer

FALSE

FALSE

$n \geq (1.96 \times 210/25)^2 = 271.1$, so 271 falls just short — the margin would exceed £25. Always round UP to 272.

True or false · 6

T / F

Decide, then justify your decision — answer on the next slide

In a hypothesis test for a proportion, the sampling distribution used should be centred at the claimed value p = 0.25, not at the observed sample proportion.

True or false · 6 — answer

TRUE

TRUE

The null model is built from the claimed parameter. The observed 0.35 is then located on that curve to measure surprise.

True or false · 7

T / F

Decide, then justify your decision — answer on the next slide

A moderately right-skewed population with mean 9.1 and SD 4.7 produces sample means that are approximately normal for samples of size 90, with mean 9.1 and SD about 0.50.

True or false · 7 — answer

TRUE

TRUE

CLT with n = 90: normal shape, centre 9.1 unchanged, spread $4.7/\sqrt{90} \approx 0.50$. All three parts matter.

True or false · 8

T / F

Decide, then justify your decision — answer on the next slide

For which sample size is it more likely that the sample proportion exceeds 65% when the population proportion is 55%: 100 or 250? The answer is 250, because larger samples vary less.

True or false · 8 — answer

FALSE

FALSE

Reaching 65% requires swinging far ABOVE 55% — more likely with MORE variability. The smaller sample (100) has the wider distribution, so 100 is the answer.

FIVE THINGS TO DO IN THE EXAM

Do these and you will not lose inference marks

- Standardise first: $z = (x - \mu)/\sigma$ for individuals; $\sigma/\sqrt{n}$ for means.
- Run BOTH large counts, $np \geq 10$ and $n(1-p) \geq 10$, before calling $\hat{p}$ normal.
- Centre never moves: mean of $\bar{x} = \mu$, mean of $\hat{p} = p$. Sample size only shrinks spread.
- Sample size problems: $n \geq (z^*\sigma/ME)^2$, always round UP.
- Hypothesis-test curves are centred at the CLAIMED p — the data sit on the curve, not in it.

Varied questions in the style of Midterm 2026 versions A and B.