Name	Date started	Target end date

GCE A LEVEL — FURTHER STATISTICS B QUESTION PACK

0985-01 (Legacy S3) · New spec A2 Unit 5 Topic 6

REVISE
.wales

FURTHER MATHS — FS B · CONFIDENCE INTERVALS — PROPORTIONS & DIFFERENCE OF MEANS

Large-Sample Confidence Intervals for a Proportion & for the Difference of Two Means

Every large-sample CI question on a binomial proportion or a difference of two means from the legacy WJEC S3 paper (2006 – 2017) that maps onto the new-spec A2 Unit 5 Topic 4. Four proportion CIs + one wine-A-vs-wine-B difference of means CI.

LEGACY 2008 SPECIFICATION

Estimated time for entire question pack: ~0 hours 49 minutes

*Derived from the legacy S3 paper's pace of **~1.3 min/mark** (38 marks over 5 questions). The full Unit 5 exam is **1 hour 45 minutes for 80 marks** (25% of the A-level qualification, A2 optional paper alongside Unit 6 Further Mechanics B).*

*You are advised to **not** attempt to complete all of this in one sitting.*

ABOUT THIS QUESTION PACK

This is a **comprehensive practice question pack**, not a single mock paper. It contains every confidence intervals — proportions & difference of means question from the legacy WJEC S3 papers (2008 modular spec) that maps onto new-spec A2 Unit 5 Topic 6 (2.5.4). Unit 5 (Further Statistics B) is one of two **80-mark A2 optional papers** (the other being Unit 6 Further Mechanics B), each worth 25% of the A-level qualification.

Questions are ordered roughly by topic / difficulty.

INSTRUCTIONS

Use black ink or black ball-point pen. Show all working — method marks are awarded for clear setup.

A calculator is allowed. The WJEC Statistical Tables and Formula Booklet may be referred to.

All question content is © WJEC CBAC Ltd. and reproduced for revision purposes.

For Examiner's use only

Q	Source	Max	Mark
1	Jun 06 Q2	8	
2	Jun 10 Q1	7	
3	Jun 13 Q1	6	
4	Jun 17 Q4	8	
5	Jun 06 Q3	9	
	Total	38	

Large-Sample Confidence Intervals for a Proportion & for the Difference of Two Means — what the new spec asks

WJEC GCE A Level Further Mathematics (from 2017) · Unit 5: Further Statistics B · Topic 2.5.4.

CI for a proportion 2.5.4

- For $X \sim B(n,p)$ with $\hat{p} = X/n$, large-sample $(1-\alpha)\cdot 100\%$ CI:
- $\hat{p} \pm z_{\alpha/2}\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$.
- Uses the normal approximation to the binomial $X \approx N(np, np(1-p))$.
- Valid when $n\hat{p} \geq 5$ and $n(1-\hat{p}) \geq 5$.

Why is the proportion CI approximate? 2.5.4

- Two distinct approximations are made:
- (i) We approximate the discrete binomial distribution by a continuous normal distribution.
- (ii) We estimate the standard error from the data (using $\hat{p}$ in place of p in $\sqrt{p(1-p)/n}$).
- S3 June 2006 Q2(d) and S3 June 2017 Q4(b) explicitly ask candidates to justify these two approximations.

CI for difference of two means (large samples) 2.5.4

- Independent samples from two populations; either both normal with known variances, or both large.
- CI for $\mu_X - \mu_Y$: $(\bar{x}-\bar{y}) \pm z_{\alpha/2}\sqrt{\dfrac{\sigma_X^2}{m} + \dfrac{\sigma_Y^2}{n}}$.
- Large-sample version: replace σ^2 with s^2.
- If the CI includes 0, the data are consistent with $\mu_X = \mu_Y$ (no evidence of a real difference).
- WJEC S3 June 2006 Q3: 100 bottles of wine A + 150 bottles of wine B $\rightarrow$ approximate 95% CI for $\mu_A - \mu_B$.

Misinterpretation pitfalls 2.5.4

- A 95% CI is *not* a 95% probability statement about a fixed parameter.
- For proportion CIs: distinguish between the SE of the *sampling distribution* of $\hat{p}$ and the estimated SE used in the interval.
- For difference-of-means: the CI for $\mu_X - \mu_Y$ is symmetric about $(\bar{x}-\bar{y})$, not about 0.
- The data being “consistent with $\mu_X = \mu_Y$” (CI contains 0) is not the same as “evidence that $\mu_X = \mu_Y$”.

© WJEC CBAC Ltd. (spec text). Compilation for revision use only.

Confidence Intervals — Proportions & Difference of Means in one page

Quick-reference notes — revisit before each section. Don't use during questions.

CI for a proportion

For $X \sim B(n,p)$ with $\hat{p} = X/n$:

$$\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Uses normal approximation to binomial.

SE of $\hat{p}$

Exact: $\mathrm{SE}(\hat{p}) = \sqrt{\dfrac{p(1-p)}{n}}$.

Estimated: $\widehat{\mathrm{SE}}(\hat{p}) = \sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$.

Substitute $\hat{p}$ for the unknown p.

Why is the proportion CI approximate?

Two reasons:

(i) We use a normal approximation to the binomial.

(ii) We estimate the SE using $\hat{p}$ in place of p.

S3 June 2006 Q2(d) and S3 June 2017 Q4(b) both ask for this.

CI for $\mu_X - \mu_Y$

Independent samples from $N(\mu_X, \sigma_X^2)$, $N(\mu_Y, \sigma_Y^2)$:

$$(\bar{x} - \bar{y}) \pm z_{\alpha/2}\sqrt{\frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}}$$

Large-sample version: replace σ^2 with s^2 (S3 June 2006 Q3 with $m = 100, n = 150$).

Validity of normal approximation

For binomial $\to$ normal: need $n\hat{p} \geq 5$ *and* $n(1-\hat{p}) \geq 5$.

For $\bar{X} - \bar{Y}$ CI to be valid with s^2 substitution: need m, n large (rule of thumb ≥ 30 each).

Otherwise the CI may not have the nominal coverage.

Interval interpretation

If 95% proportion CI is $[0.32, 0.44]$, do *not* say:

"There is a 95% probability that p is in [0.32, 0.44]."

Correct: *95% of intervals constructed this way would contain the true p on repeated sampling.*

CI contains d_0 test

If the CI for $\mu_X - \mu_Y$ contains 0, the data are consistent with $\mu_X = \mu_Y$.

If the CI contains a specified d_0, the data are consistent with $\mu_X - \mu_Y = d_0$.

This is the dual of a two-sided hypothesis test at level α.

Common pitfalls

- Forgetting to estimate SE using $\hat{p}$ — using exact p when it's unknown.
- Misinterpreting CI as a probability statement about a fixed parameter.
- Subtracting variances when constructing the SE for $\bar{X} - \bar{Y}$.
- Using z-table for 95% as 1.645 instead of 1.96.

Strategy

1. Identify proportion CI vs difference-of-means CI.
2. Compute $\hat{p}$ (or $\bar{x}, \bar{y}, s_X^2, s_Y^2$).
3. Plug into the appropriate formula with $z_{\alpha/2}$.
4. Interpret in context; explain "approximate" if asked.

SECTION T6

Large-Sample Confidence Intervals for a Proportion & for the Difference of Two Means

Questions 1–5 · 38 marks

2. In order to estimate the proportion p of a certain population who are bilingual, a random sample of 1200 members of the population is questioned. It is found that 498 of them are bilingual.

(a) Calculate an unbiased estimate of p. [1]

(b) Estimate the standard error of your estimate. [2]

(c) Calculate an approximate 90% confidence interval for p. [3]

(d) Give **two** reasons why your interval is approximate. [2]

1. Jamie is given a coin and he wishes to estimate p, the probability of its landing 'heads' when tossed. He therefore tosses the coin 250 times and obtains 140 'heads'.

(a) Calculate an unbiased estimate of p. [1]

(b) Calculate an approximate 99% confidence interval for p. [5]

(c) State, with a reason, whether or not your results suggest that the coin is biased. [1]

1. A university Vice Chancellor wishes to estimate the proportion of students in the university who are fluent in Welsh. She therefore contacts a random sample of 300 students and finds that 87 of them are fluent in Welsh.
Determine an approximate 95% confidence interval for the proportion of the students in the university who are fluent in Welsh. [6]

4. A mathematics teacher takes a biased dice to his class, wishing to estimate p, the probability of throwing a 'six'. He throws it 75 times and obtains 24 'sixes'.

(a) Calculate an approximate 95% confidence interval for p. [6]

(b) The teacher calculates this interval and he asks Tom to interpret it. Tom states that 'There is, approximately, a 0·95 probability that the interval that the teacher has calculated contains the unknown value of p'. Explain why this statement is incorrect and give a correct interpretation. [2]

© WJEC CBAC Ltd.

3. Two different varieties of wine, A and B, are sold in 1 litre bottles. A consumer organisation measured the amount of wine, x litres, in each of 100 randomly selected bottles of variety A. The results are summarised below.

$$\sum x = 103{\cdot}4; \sum x^2 = 106{\cdot}95$$

The organisation also measured the amount of wine, y litres, in each of 150 randomly selected bottles of variety B. The results are summarised below.

$$\sum y = 152{\cdot}4; \sum y^2 = 154{\cdot}86$$

Determine an approximate 95% confidence interval for the difference in the population means of the amounts of wine in 1 litre bottles of the two varieties. [9]

END OF CONFIDENCE INTERVALS — PROPORTIONS & DIFFERENCE OF MEANS PACK

Source: WJEC S3 (2008 modular spec) · 2005–2017
Curated for WJEC FM 2017 spec A2 Unit 5 — Topic 6 (2.5.4)

© WJEC CBAC Ltd. Pack layout © revise.wales for revision purposes only.