Name	Date started	Target end date

GCE A LEVEL — FURTHER STATISTICS B QUESTION PACK

0984-01 (Legacy S2) & 0985-01 (Legacy S3) · New spec A2 Unit 5 Topic 4

REVISE
.wales

FURTHER MATHS — FS B · CONFIDENCE INTERVALS

Confidence Intervals & Estimation — Mean, Difference of Means & Proportion

Every confidence-interval question from the legacy WJEC S2 and S3 papers (June 2005 – June 2017 + Specimen) that maps onto the new-spec A2 Unit 5 Topic 4.

LEGACY 2008 SPECIFICATION

Estimated time for entire question pack: ~2 hours 54 minutes

*Derived from the legacy S2/S3 paper's pace of **~1.3 min/mark** (134 marks over 17 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 question from the legacy WJEC S2/S3 papers (2008 modular spec) that maps onto new-spec A2 Unit 5 Topic 4 (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.

For Examiner's use only

Q	Source	Max	Mark
1	Jun 05 Q1	6	
2	Jun 06 Q1	6	
3	Jun 08 Q4	8	
4	Jun 09 Q3	9	
5	Jun 11 Q2	9	
6	Jun 12 Q3	6	
7	Jun 14 Q1	6	
8	Jun 16 Q2	10	
9	Spec. Q3	8	

Q	Source	Max	Mark
10	Jun 07 Q2	9	
11	Jun 15 Q2	10	
12	Jun 17 Q5	9	
13	Jun 06 Q2	8	
14	Jun 10 Q1	7	
15	Jun 13 Q1	6	
16	Jun 17 Q4	8	
17	Jun 06 Q3	9	
Total		134	

Confidence Intervals & Estimation — Mean, Difference of Means & Proportion — what the new spec asks

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

CI for μ — known σ 2.5.4

- Sample $X_1, \ldots, X_n$ from $N(\mu, \sigma^2)$ with σ known.
- $(1-\alpha)\cdot 100\%$ CI: $\bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$.
- Standard z-values: $z_{0.05} \approx 1.645$, $z_{0.025} \approx 1.96$, $z_{0.005} \approx 2.576$.
- Interpretation: *If we repeated the sampling many times,* $(1-\alpha)\cdot 100\%$ *of intervals would contain* μ.

CI for μ — unknown σ 2.5.4

- Sample from $N(\mu, \sigma^2)$, σ unknown.
- $(1-\alpha)\cdot 100\%$ CI uses Student's t-distribution: $\bar{x} \pm t_{n-1,\alpha/2} \cdot \frac{s}{\sqrt{n}}$.
- Compute $s^2 = \frac{1}{n-1}\sum(x_i - \bar{x})^2$ from the data.
- For large n (≥ 30), the t-distribution approaches $N(0,1)$.

CI for difference of two means 2.5.4

- Independent samples from $N(\mu_X, \sigma_X^2)$ and $N(\mu_Y, \sigma_Y^2)$ with known variances.
- CI for $\mu_X - \mu_Y$: $(\bar{x} - \bar{y}) \pm z_{\alpha/2} \cdot \sqrt{\frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}}$.
- Large-sample version: replace σ^2 by s^2 when populations are not normal and samples are large.
- If the CI includes 0, the data are consistent with $\mu_X = \mu_Y$.

CI for proportion 2.5.4

- Large-sample CI for a proportion p based on $X \sim B(n,p)$ with $\hat{p} = X/n$:
- $\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ — uses normal approximation.
- Approximate because (i) we use a normal approximation to the binomial, and (ii) we estimate the SE.
- Valid when n is large enough that $n\hat{p} \geq 5$ and $n(1-\hat{p}) \geq 5$.

Confidence Intervals in one page

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

95% CI for μ, known σ

For sample $X_1, \dots, X_n$ from $N(\mu, \sigma^2)$ with σ known:

$$\bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

For 95%: $z_{0.025} = 1.96$.

z-values for common levels

90%: $z = 1.645$

95%: $z = 1.96$

98%: $z = 2.326$

99%: $z = 2.576$

These are *two-sided* critical values (i.e. $z_{\alpha/2}$).

CI with t-distribution

For sample from $N(\mu, \sigma^2)$ with σ **unknown**:

$$\bar{x} \pm t_{n-1,\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

Use t-tables with $n-1$ degrees of freedom.

s is the unbiased sample standard deviation.

CI for two-mean difference

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

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

Large-sample version: replace σ^2 with s^2.

CI for 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 — valid when $n\hat{p} \geq 5$ and $n(1-\hat{p}) \geq 5$.

Margin of error

The half-width of a CI is the **margin of error**:

$$\mathrm{ME} = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

To halve the width: quadruple n (because $\sqrt{n}$ goes into the denominator).

To find n for a target ME: solve $z\sigma/\sqrt{n} \leq \mathrm{ME}$.

Interpretation

A 95% CI does **not** mean "there is a 95% probability that μ lies in the interval".

It means: *if we repeated the sampling many times, about 95% of resulting CIs would contain the true μ.*

Once an interval is computed, the parameter either is or isn't in it.

Common pitfalls

- Using σ when only s is known — switch to t-distribution.
- Forgetting why proportion-CIs are *approximate* (binomial approximated by normal + estimated SE).
- Wrong z-value: 1.96 is for 95%, not 99%.
- Misinterpreting the meaning of a CI (probability vs frequentist coverage).

Strategy

1. Decide: is σ known or estimated from data?
2. Pick z (known σ / large n) or t-value (unknown σ, small n).
3. Compute $\bar{x} \pm \text{critical value} \cdot \mathrm{SE}$.
4. For proportion or 2-sample: use the corresponding SE formula.

SECTION T4

Confidence Intervals & Estimation

Questions 1–17 · 134 marks

1. Ann records the times taken (in minutes) to drive to work on twelve consecutive days with the following results.

73·3 74·7 71·2 75·8 70·0 74·0 80·5 70·5 72·6 73·4 69·1 76·9

Assume that these observations form a random sample from a normal distribution with standard deviation 4·0.

(a) Calculate a 95% confidence interval for the mean time taken by Ann to drive to work. [5]

(b) Ann believed that the mean time was 75 minutes. State, with a reason, whether or not your result supports this belief. [1]

1. Bill is an athlete specialising in the long-jump. He does 10 jumps at a training session with the following results (in metres).

6·21 6·33 6·02 6·11 6·13 6·40 6·51 6·29 6·16 6·44

You may assume that the above distances form a random sample from a normal distribution with unknown mean μ metres and standard deviation 0·1 metres.

Calculate a 95% confidence interval for μ.

Bill had stated beforehand that the mean length of his jumps was 6·3 metres.

Giving a reason, state whether or not your interval supports this claim. [6]

4. When a certain instrument is used to measure the concentration of a solution, the reading obtained (in appropriate units) is a normally distributed random variable with mean equal to the actual concentration and standard deviation 0·05. The concentration of a particular solution was measured eight times and the following independent readings were obtained.

6·02 6·10 5·98 6·04 6·07 5·94 6·05 6·12

(a) Calculate a 99% confidence interval for the concentration of this solution. [5]

(b) Find the minimum number of measurements required to obtain a 99% confidence interval of width less than 0·04. [3]

3. A gardener wishes to estimate the acidity level of the soil in his garden. He therefore takes ten readings of the acidity level with the following results.

$$6{\cdot}62,\quad 6{\cdot}84,\quad 6{\cdot}77,\quad 6{\cdot}52,\quad 6{\cdot}86,\quad 6{\cdot}51,\quad 6{\cdot}82,\quad 6{\cdot}71,\quad 6{\cdot}49,\quad 6{\cdot}66$$

You may assume that this is a random sample from a normal distribution with standard deviation 0·1

(a) Calculate a 99% confidence interval for the acidity level of his soil. [5]

(b) A friend uses the same data to calculate a confidence interval and obtains the following result.

$$[6{\cdot}62, 6{\cdot}74]$$

Calculate the confidence level of this interval. [4]

2. The random variable X has a normal distribution with unknown mean μ and standard deviation 0·5.

(a) A random sample of 60 values of X was taken and it was found that $\sum x = 1290$. Calculate a 95% confidence interval for μ, giving the end-points of your interval correct to two decimal places. [5]

(b) Determine the minimum sample size required for the width of a 95% confidence interval for μ to be less than 0·1. [4]

3. The lifetime, X thousand hours, of a certain type of electric light bulb may be assumed to be normally distributed with unknown mean μ and standard deviation 0·1. The lifetimes of a random sample of 75 of these bulbs were measured and it was found that $\sum x = 69{\cdot}9$.

(a) Find a 90% confidence interval for μ. [5]

(b) Give an interpretation of this confidence interval. [1]

1. The times taken, in minutes, for trains to travel between two stations on a particular day were recorded and are given below.

48·2 49·4 56·2 44·6 47·3 55·2 50·8 53·9

It may be assumed that this is a random sample from a normal distribution with mean μ mins and standard deviation 4 mins. Determine a 90% confidence interval for μ. [6]

2. Sue keeps chickens in her garden. She selects, at random, 10 of the eggs produced and weighs them. The results, in grams, are shown below.

62·5 64·2 61·5 65·2 66·2 63·8 60·1 63·2 64·4 66·1

You may assume that this is a random sample from a normal distribution with a standard deviation of 1·9.

(a) Determine a 95% confidence interval for the mean weight of eggs produced by Sue's chickens. [6]

(b) Sue was hoping to obtain a 95% confidence interval of width 1 at most. Calculate the minimum sample size necessary to achieve this. [4]

3. Ann drives to work and she records the times taken over a 10-day period with the following results (in minutes):

72 76 69 77 81 74 71 69 72 74

You may assume that the time taken to drive to work can be modelled by a normally distributed random variable with standard deviation 4 minutes,

(a) Calculate a 95% confidence interval for the mean time. [6]

(b) Explain what is meant by a 95% confidence interval. [2]

2. When Bill throws the discus, the distance thrown (in metres) can be assumed to be normally distributed with mean μ and variance σ^2. He throws the discus 10 times with the following results.

25·3 23·8 24·7 24·9 23·7 25·6 24·6 24·0 25·3 24·1

(a) Calculate unbiased estimates of μ and σ^2. [5]

(b) Calculate a 95% confidence interval for μ. [4]

2. Emlyn solves the crossword in the Daily Bugle every day. He records the time taken to do this on 12 randomly chosen days with the following results (in minutes).

16·3 17·4 14·3 15·6 16·4 13·9 16·9 17·4 17·9 15·3 16·6 14·9

You may assume that these times are normally distributed with mean μ and variance σ^2.

(a) Calculate unbiased estimates of μ and σ^2. [5]

(b) Determine a 99% confidence interval for μ. [5]

5. When Dawn throws the javelin, the distance thrown (in metres) can be assumed to be normally distributed with mean μ and variance σ^2. She throws the javelin 9 times with the following results.

33·5 34·6 33·3 34·3 34·6 34·0 33·1 35·0 33·6

(a) Calculate unbiased estimates of μ and σ^2. [5]

(b) Calculate a 95% confidence interval for μ. [4]

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]

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 PACK

Source: WJEC S2/S3 (2008 modular spec) · 2005–2017
Curated for WJEC FM 2017 spec A2 Unit 5 — Topic 4 (2.5.4)