

Inference using normal and t-distributions

A2

Read online at pastopic.com/cambridge/further-mathematics-9231/notes/inference-using-normal-and-t-distributions

Last checked 1 October 2026

© 2026 Pastopic. Free for personal study. Please share the link, not the file.

What you need to know

By the end of this topic you should be able to:

1. **Test a population mean with a t -test.** The sample is small, the population is normal, and its variance is unknown. Write the hypotheses, work out t and compare it with the t table.
2. **Pool two samples to estimate a common variance.** The data may be raw values or summaries.
3. **Test the difference between two population means**, choosing the right test for the situation:
 - a **two-sample t -test** (small samples, normal populations, equal variances, pooled estimate);
 - a **paired-sample t -test** (two measurements on the same subjects, or matched pairs);
 - a **test using the normal distribution** (large samples, or known population variances).
4. **Find a confidence interval for a population mean** from a small sample of a normal population with unknown variance, using the t -distribution.
5. **Find a confidence interval for a difference of population means**, using the t -distribution or the normal distribution, whichever fits.

You are expected to know from AS & A Level Mathematics (9709, Probability & Statistics 2): unbiased estimates, the Central Limit Theorem, z confidence intervals and the language of hypothesis tests.

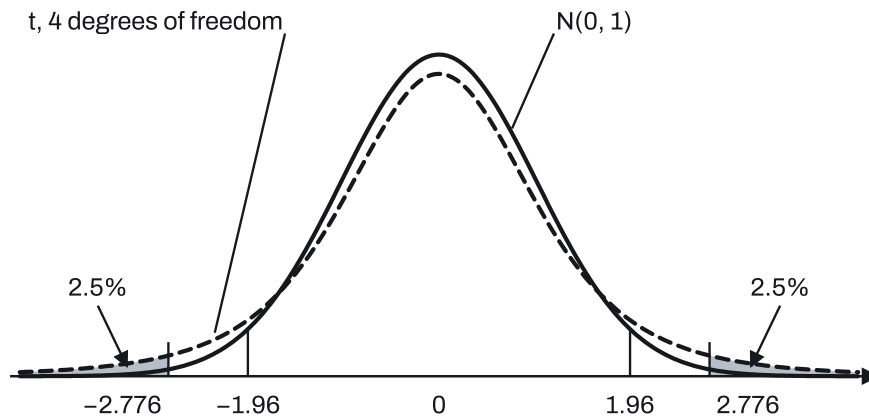
Understand it

Why a new distribution?

For a normal population, $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$. But you rarely know σ . Replace it with the sample estimate s and the ratio becomes

$$T = \frac{\bar{X} - \mu}{s/\sqrt{n}},$$

which is **not** normal for a small sample, because s itself varies from sample to sample. T has a **t -distribution with $\nu = n - 1$ degrees of freedom**. It is symmetrical about 0 and bell-shaped like $N(0, 1)$, but lower in the middle with **heavier tails**. As n grows, s becomes a reliable estimate of σ and the t -distribution gets closer and closer to $N(0, 1)$.



Heavier tails mean the cut-off values are further out: the middle 95% of $N(0, 1)$ lies within ± 1.96 , but for t with 4 degrees of freedom it lies within ± 2.776 . So a t interval is wider, which is the honest price of not knowing σ .

Reading the t table (MF19). It gives t with $P(T \leq t) = p$ for each ν . For a 95% two-sided interval or a 5% two-tailed test use $p = 0.975$; for a 5% one-tailed test use $p = 0.95$; for 90% use $p = 0.95$. For $\nu = 8$: $p = 0.95$ gives 1.860, $p = 0.975$ gives 2.306.

Unbiased estimates

Everything starts from the sample mean and the **unbiased** estimate of the population variance:

$$\bar{x} = \frac{\Sigma x}{n}, \quad s^2 = \frac{\Sigma(x - \bar{x})^2}{n - 1} = \frac{1}{n - 1} \left(\Sigma x^2 - \frac{(\Sigma x)^2}{n} \right).$$

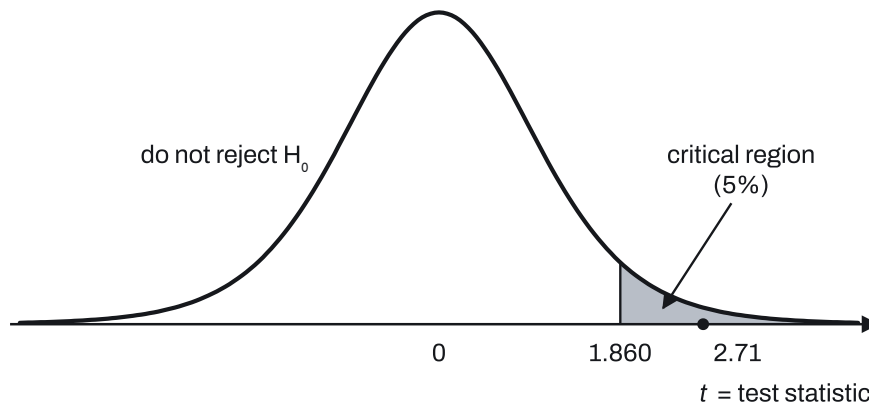
For 4, 7, 9, 12: $\bar{x} = 8$, $\Sigma(x - \bar{x})^2 = 34$, so $s^2 = \frac{34}{3} = 11.33$. Dividing by $n = 4$ instead gives 8.5, the biased value, which loses the method mark. Calculator " σ_n " is biased; " s " or " σ_{n-1} " is unbiased.

One sample: test and interval

To test $H_0 : \mu = \mu_0$:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}, \quad \nu = n - 1.$$

If t falls in the **critical region** (beyond the table value, on the side H_1 points to), reject H_0 .



A **confidence interval** for μ is every value that is not "too far" from $\bar{x}$:

$$\bar{x} \pm t \times \frac{s}{\sqrt{n}}.$$

Both need the population to be **normal**; the sample being small is why you need t rather than z .

Two samples: which method?

Ask three questions.

1. **Are the data paired?** The same people measured twice (before and after), or items matched up one-to-one (pair 1, pair 2, ...). Then work with the **differences** d and do a one-sample t on them: a **paired t -test**. The two columns are not independent samples, so a two-sample test would be wrong.
2. **Are the samples large** (roughly 30 or more each)? Then by the Central Limit Theorem $\bar{X}$ and $\bar{Y}$ are approximately normal whatever the populations look like, and s^2 is a good estimate of σ^2 . Use z with **separate** variances:

$$z = \frac{(\bar{x} - \bar{y}) - c}{\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}}.$$

If the population variances are **known**, use them in place of s_x^2 , s_y^2 , and then small samples from normal populations are fine too.

3. **Small independent samples?** Then you need both populations **normal with equal variances**. Pool the two samples to estimate the common variance:

$$s_p^2 = \frac{\Sigma(x - \bar{x})^2 + \Sigma(y - \bar{y})^2}{n_x + n_y - 2}, \quad t = \frac{(\bar{x} - \bar{y}) - c}{\sqrt{s_p^2 \left(\frac{1}{n_x} + \frac{1}{n_y} \right)}}, \quad \nu = n_x + n_y - 2.$$

Pooling is only allowed when the question says (or tells you to assume) the variances are equal. With large samples and no such statement, **don't pool**.

Useful link: $\Sigma(x - \bar{x})^2 = (n - 1)s^2$. So a sample given as " $n = 7$, unbiased variance 3.2" contributes $6 \times 3.2 = 19.2$ to the numerator of s_p^2 .

Tests and intervals are two views of one idea

A 95% interval for $\mu_x - \mu_y$ holds exactly the values c that a **two-tailed** 5% test of $H_0 : \mu_x - \mu_y = c$ would **not** reject. So if 0 is inside the interval, a two-tailed test at the matching level finds no evidence of a difference; if 0 is outside, it does. A one-tailed 5% test matches a **90%** interval, not a 95% one (Example 3 shows this).

What a conclusion can say

A test never proves anything. It either finds **sufficient evidence** to reject H_0 or **insufficient evidence**. The conclusion must be in the words of the question: "There is sufficient evidence to suggest that the mean mass of honey in a jar is more than 500 g."

Key facts and formulas

"Given" means it is in the MF19 list of formulae. "Learn it" means you must remember it. The z and t tables are in MF19.

Formula	Status
$\bar{x} = \frac{\Sigma x}{n}, s^2 = \frac{\Sigma(x - \bar{x})^2}{n - 1} = \frac{1}{n - 1} \left(\Sigma x^2 - \frac{(\Sigma x)^2}{n} \right)$	Given
Central Limit Theorem: $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ approximately, for large n	Given
Pooled estimate: $s_p^2 = \frac{\Sigma(x_1 - \bar{x}_1)^2 + \Sigma(x_2 - \bar{x}_2)^2}{n_1 + n_2 - 2}$	Given
$\Sigma(x - \bar{x})^2 = (n - 1)s^2$	Learn it
One-sample t : $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}, \nu = n - 1$	Learn it
t interval for μ : $\bar{x} \pm t \frac{s}{\sqrt{n}}, \nu = n - 1$	Learn it
Paired t : $t = \frac{\bar{d} - k}{s_d/\sqrt{n}}, \nu = n - 1$ (n pairs)	Learn it
Two-sample t : $t = \frac{(\bar{x} - \bar{y}) - c}{\sqrt{s_p^2 \left(\frac{1}{n_x} + \frac{1}{n_y} \right)}}, \nu = n_x + n_y - 2$	Learn it
Pooled t interval: $(\bar{x} - \bar{y}) \pm t \sqrt{s_p^2 \left(\frac{1}{n_x} + \frac{1}{n_y} \right)}$	Learn it
Large samples: $z = \frac{(\bar{x} - \bar{y}) - c}{\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}}$; interval $(\bar{x} - \bar{y}) \pm z \sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}$	Learn it
Common z values: 90% or one-tailed 5%: 1.645; 95% or two-tailed 5%: 1.960; one-tailed 1%: 2.326; two-tailed 1%: 2.576	Learn it (or read from the table)

How to do it

In the Paper 4 questions we have tagged for this topic (66 questions, 45 different, 2021–2025), the types came up this often. 32 of the 66 questions mix types, so the counts add up to more than 66.

Question type	Questions
State or criticise the assumptions a test or interval needs	20
Confidence interval for a population mean using t (one sample)	14
Paired-sample t -test (difference 0 or "more than k ")	12
Two-sample t -test using a pooled variance	10
Large-sample (z) confidence interval for a difference of means	9
Large-sample (z) test for a difference of means	9
One-sample t -test for a population mean	8
Boundary value (k or significance level $\alpha\%$) where a conclusion changes	6
Interpret or evaluate a result	6
Work back from a t confidence interval to Σx and Σx^2	4
Confidence interval for a difference of means using a pooled t	2
Pooled estimate of a common variance	2

Every test follows the same five steps: **hypotheses** → **estimates** → **test statistic** → **compare with the critical value** → **conclusion in context**. The types below say what changes.

1. State or criticise the assumptions (20 questions)

One mark each, and the wording is marked strictly. Match the assumption to the method:

Method	Assumption to state
One-sample t test or interval	The population (of whatever is measured, e.g. masses) is normally distributed.
Paired t -test	The population of differences is normally distributed.
Two-sample pooled t	Both populations are normal and have equal variances .
Large-sample z	None about the populations: the samples are large, so by the Central Limit Theorem the sample means are approximately normal.

- "The data are normal", "the sample is normal" or "the population **mean** is normal" all score nothing. Say *population* (or *underlying distribution*) and name the variable.

- **Why an assumption may fail** (give a reason in context): the values are bounded or bunched against a limit so the distribution is skewed (waiting times at a clinic, where most are short but a few are very long, are skewed to the right); observations may not be independent (readings taken by one gauge that drifts during the day); the sample may not be random. For a paired test: the differences may not be symmetrical, for example if the effect is bigger for some subjects than others.
- **"Why is no assumption about the distributions needed?"** Because both samples are large, so the Central Limit Theorem makes $\bar{X} - \bar{Y}$ approximately normal.

2. Confidence interval for a population mean using t (14 questions)

1. Find $\bar{x}$ and the **unbiased** s^2 . The data may be a raw list (find Σx , Σx^2 first), Σx and Σx^2 , or $\bar{x}$ with $\Sigma(x - \bar{x})^2$ (then $s^2 = \frac{\Sigma(x - \bar{x})^2}{n-1}$).
2. Read t with $\nu = n - 1$: $p = 0.95$ for 90%, 0.975 for 95%, 0.99 for 98%, 0.995 for 99%.
3. Interval $\bar{x} \pm t \frac{s}{\sqrt{n}}$. Give it as $[a, b]$ to 3 significant figures.

A z value here scores no method mark, and " $\bar{x} \pm$ something" is not accepted as the final answer: work out both ends.

3. Paired-sample t -test (12 questions)

1. **Choose the order** of subtraction to suit the claim, and say it: " d = after – before". Work out every **signed** difference (keep the minus signs).
2. **Hypotheses** in terms of population means: $H_0 : \mu_d = k$, $H_1 : \mu_d > k$ (define μ_d as the population mean difference). Most claims are "reduced (or increased) by **more than** k ", which puts k in H_0 ; with no k , use 0.
3. $\bar{d}$ and $s_d^2 = \frac{1}{n-1} \left(\Sigma d^2 - \frac{(\Sigma d)^2}{n} \right)$, with n the number of **pairs**.
4. $t = \frac{\bar{d} - k}{s_d / \sqrt{n}}$ and compare with t on $n - 1$ degrees of freedom.
5. Conclude in context, with doubt in the wording.

Tip: subtract so that the differences are positive when the claim is true. For a "reduces" claim use before – after; for an "increases" claim use after – before (as in Example 2). Then H_1 is always $\mu_d > k$.

4. Two-sample t -test using a pooled variance (10 questions)

Used when the question says (or tells you to assume) both populations are normal with the **same variance**, and the samples are small and independent.

1. $H_0 : \mu_x = \mu_y$, $H_1 : \mu_x > \mu_y$ (or $<$, or $\neq$, as the claim says).
2. For each sample find $\Sigma(x - \bar{x})^2 = \Sigma x^2 - \frac{(\Sigma x)^2}{n}$. One sample is often a raw list and the other a summary; a sample given by its unbiased variance s^2 contributes $(n - 1)s^2$.
3. $s_p^2 = \frac{\Sigma(x - \bar{x})^2 + \Sigma(y - \bar{y})^2}{n_x + n_y - 2}$.

$$4. t = \frac{\bar{x} - \bar{y}}{\sqrt{s_p^2 \left(\frac{1}{n_x} + \frac{1}{n_y} \right)}}, \text{ compared with } t \text{ on } n_x + n_y - 2 \text{ degrees of freedom.}$$

5. Conclude in context.

Variants: one sample's statistics come from an earlier part, for example worked back from a confidence interval (type 10).

5. Large-sample (z) confidence interval for a difference of means (9 questions)

1. Find $\bar{x}$, $\bar{y}$ and **both** unbiased variances s_x^2 , s_y^2 (not pooled).

$$2. \text{variance of } \bar{X} - \bar{Y} \approx \frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}.$$

$$3. \text{Interval: } (\bar{x} - \bar{y}) \pm z \sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}, \text{ with } z = 1.645 (90\%), 1.960 (95\%), 2.326 (98\%) \text{ or } 2.576 (99\%).$$

Keep the order of subtraction the question uses. " $\mu_A - \mu_B$ " gives $[a, b]$; " $\mu_B - \mu_A$ " gives $[-b, -a]$; both are fine if you say which.

6. Large-sample (z) test for a difference of means (9 questions)

1. $H_0 : \mu_x - \mu_y = c$ (often $c = 0$), H_1 one- or two-tailed from the wording.

2. Separate unbiased variances, as in type 5. If you are given $\Sigma(x - \bar{x})^2$, divide by $n - 1$.

$$3. z = \frac{(\bar{x} - \bar{y}) - c}{\sqrt{s_x^2/n_x + s_y^2/n_y}}.$$

4. Compare with the z critical value (for example 1.645 one-tailed 5%, 2.576 two-tailed 1%), or compare $P(Z > |z|)$ with the tail area.

5. Conclude in context.

A **nonzero** difference c ("the mean for X is more than c greater than the mean for Y ") goes into the numerator, exactly like k in a paired test. A pooled variance here scores nothing: use it only when the question gives equal variances.

7. One-sample t -test for a population mean (8 questions)

1. $H_0 : \mu = \mu_0$, $H_1 : \mu < \mu_0$, $> \mu_0$ or $\neq \mu_0$. "Improved" times means **smaller**; read the context.

2. $\bar{x}$ and unbiased s^2 (sometimes the question hands you both).

$$3. t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}, \nu = n - 1.$$

4. Compare with the critical value **on the same side**: for $H_1 : \mu < \mu_0$, reject if $t < -t_{\text{crit}}$. Keep signs consistent.

5. Conclude in context.

It often follows a t interval on the same data (type 2), so reuse $\bar{x}$ and s .

8. Boundary value where a conclusion changes (6 questions)

Write the rejection condition as an inequality and solve it for the unknown.

- **An unknown $\mu_0 = k$ in a one-sample test.** "The test of $\mu = k$ does not support $\mu < k$ " means $\frac{\bar{x} - k}{s/\sqrt{n}} \geq -t_{\text{crit}}$, so $k \leq \bar{x} + t_{\text{crit}} \frac{s}{\sqrt{n}}$. The greatest k is that bound. (For "does not support $\mu > k$ ", flip the signs.)
- **An unknown k in a paired claim "more than k ".** Rejection needs $\frac{\bar{d} - k}{s_d/\sqrt{n}} \geq t_{\text{crit}}$, so $k \leq \bar{d} - t_{\text{crit}} \frac{s_d}{\sqrt{n}}$ (Example 2(b)).
- **An unknown significance level $a\%$.** Work out z (or use the z you found), then the tail probability $p = 1 - \Phi(|z|)$ (double it for a two-tailed test). The test gives evidence whenever $a\% \geq p$, so $a \geq 100p$ (Example 4(c)).

Strict or non-strict inequalities are both accepted at the boundary.

9. Interpret or evaluate a result (6 questions)

- **A confidence interval and a test.** If a 95% interval for $\mu_x - \mu_y$ contains 0, a two-tailed 5% test finds insufficient evidence of a difference; if 0 is outside, it finds evidence. Say "because 0 lies inside (outside) the interval".
- **Changed data in a paired test.** If both values of one pair change by the same amount, the difference for that pair is unchanged, so $\bar{d}$, s_d , t and the conclusion are all unchanged. Otherwise redo the calculation.
- **An alternative test** when the differences may not be normal: a **Wilcoxon matched-pairs signed-rank test** or a **paired-sample sign test** (the word "paired" or "matched" matters). For two independent samples that may not be normal: a Wilcoxon rank-sum test.
- **Would a t -test and a Wilcoxon test agree?** Usually, but not always: they use different information (the t -test uses the actual values and the normal assumption; Wilcoxon uses only ranks), so for data near the boundary they can disagree.

10. Work back from a t confidence interval to Σx and Σx^2 (4 questions)

1. **Midpoint:** $\bar{x} = \frac{1}{2}(\text{lower} + \text{upper})$, so $\Sigma x = n\bar{x}$.
2. **Half-width:** $\frac{1}{2}(\text{upper} - \text{lower}) = t \frac{s}{\sqrt{n}}$ with t on $n - 1$ degrees of freedom at the stated level. Solve for s , then s^2 .
3. **Back to Σx^2 :** $\Sigma x^2 = (n - 1)s^2 + \frac{(\Sigma x)^2}{n}$.

Work numerically, one quantity at a time; don't combine the formulas into one expression first. Using a z value, or t for the wrong p or ν , is the usual error. Quick check 2 practises this.

11. Confidence interval for a difference of means using a pooled t (2 questions)

Same conditions as type 4 (small, normal, equal variances).

1. Find both $\Sigma(x - \bar{x})^2$ and s_p^2 .
2. t on $n_x + n_y - 2$ degrees of freedom.
3. $(\bar{x} - \bar{y}) \pm t \sqrt{s_p^2 \left(\frac{1}{n_x} + \frac{1}{n_y} \right)}$, both ends worked out (Example 3(b)).

12. Pooled estimate of a common variance (2 questions)

1. For each sample find $\Sigma(x - \bar{x})^2$: from a list directly, or from a given s as $(n - 1)s^2$.
2. $s_p^2 = \frac{\text{sum of both}}{n_1 + n_2 - 2}$.
3. Read the question: if it asks for an estimate of the **standard deviation** σ , take the square root (Quick check 1).

Worked examples

The bold codes show where marks are earned, as in Cambridge mark schemes. The number is how many marks.

- **M1** method mark: for a correct method, even if the numbers slip.
 - **A1** accuracy mark: for a correct answer; it needs the M mark before it.
 - **B1** mark for a correct result on its own, with no method needed.
-

Example 1

A company sells honey in jars labelled "500 g". The masses, in grams, of the honey in a random sample of 9 jars are:

503 498 507 501 505 499 504 506 502

The masses are normally distributed.

- (a) Find a 95% confidence interval for the population mean mass. **[4]**
- (b) Test, at the 5% significance level, whether the mean mass is more than 500 g. **[4]**

(a) $\Sigma x = 4525$, $\Sigma x^2 = 2\,275\,145$, so $\bar{x} = 502.78$ (2 d.p.) and

$$s^2 = \frac{1}{8} \left(2\,275\,145 - \frac{4525^2}{9} \right) = \frac{85}{9} = 9.444.$$

M1 for the unbiased estimate.

$\nu = 8$, $t = 2.306$ ($p = 0.975$). **B1**

$$502.78 \pm 2.306 \times \sqrt{\frac{9.444}{9}} = 502.78 \pm 2.306 \times 1.0244 = 502.78 \pm 2.36.$$

M1 for the form with a t value. The interval is $[500.4, 505.1]$ g (4 s.f.). **A1**

(b) $H_0 : \mu = 500, H_1 : \mu > 500$. **B1**

$$t = \frac{502.78 - 500}{1.0244} = 2.71.$$

M1 A1. The critical value for a one-tailed 5% test with $\nu = 8$ is 1.860. $2.71 > 1.860$, so reject H_0 . There is sufficient evidence to suggest that the mean mass of honey in a jar is more than 500 g. **A1**

The second diagram in "Understand it" shows this test.

Example 2

A college runs a typing course. The typing speeds, in words per minute, of 9 randomly chosen students before and after the course are:

Student	A	B	C	D	E	F	G	H	I
Before	42	55	38	61	47	50	44	58	52
After	47	58	41	67	49	56	46	63	55

(a) The college claims that the course increases a student's typing speed by more than 3 words per minute on average. Test this claim at the 10% significance level. **[7]**

(b) Find the largest value of k for which a test at the 10% level would support the claim that the mean increase is more than k words per minute. **[2]**

(c) State an assumption needed for your test. **[1]**

(a) Let d = after – before and μ_d the population mean increase. $H_0 : \mu_d = 3, H_1 : \mu_d > 3$. **B1**

Differences: 5, 3, 3, 6, 2, 6, 2, 5, 3. **M1**

$\Sigma d = 35, \Sigma d^2 = 157$, so $\bar{d} = 3.8889$ and $s_d^2 = \frac{1}{8} \left(157 - \frac{35^2}{9} \right) = \frac{47}{18} = 2.611$. **M1**

$$t = \frac{3.8889 - 3}{\sqrt{2.611/9}} = \frac{0.8889}{0.5386} = 1.65.$$

M1 A1. With $\nu = 8$ the one-tailed 10% critical value is 1.397 ($p = 0.9$). $1.65 > 1.397$, so reject H_0 . **M1**

There is sufficient evidence to support the claim that the course increases typing speed by more than 3 words per minute on average. **A1**

(b) Need $\frac{3.8889 - k}{0.5386} \geq 1.397$, so $k \leq 3.8889 - 1.397 \times 0.5386 = 3.14$. **M1 A1**

The largest value is $k = 3.14$ (3 s.f.).

(c) The population of differences (increases in typing speed) is normally distributed. **B1**

Example 3

The lifetimes, in hours, of batteries of brand P and brand Q are normally distributed with the same variance. A random sample of 7 brand P batteries have these lifetimes:

12.4 13.1 11.8 12.9 13.6 12.2 12.7

A random sample of 9 brand Q batteries gives $\Sigma y = 108.0$ and $\Sigma y^2 = 1300.5$.

(a) Test, at the 5% significance level, whether brand P batteries last longer on average than brand Q batteries. **[8]**

(b) Find a 95% confidence interval for $\mu_P - \mu_Q$. **[3]**

(c) Explain why your answers to (a) and (b) do not contradict each other. **[1]**

(a) $H_0 : \mu_P = \mu_Q$, $H_1 : \mu_P > \mu_Q$. **B1**

Brand P : $\Sigma x = 88.7$, $\Sigma x^2 = 1126.11$, $\bar{x} = 12.6714$, $\Sigma(x - \bar{x})^2 = 1126.11 - \frac{88.7^2}{7} = 2.1543$. **M1**

Brand Q : $\bar{y} = 12.0$, $\Sigma(y - \bar{y})^2 = 1300.5 - \frac{108^2}{9} = 4.5$. **M1**

$$s_p^2 = \frac{2.1543 + 4.5}{7 + 9 - 2} = 0.4753.$$

M1 A1

$$t = \frac{12.6714 - 12.0}{\sqrt{0.4753 \left(\frac{1}{7} + \frac{1}{9}\right)}} = \frac{0.6714}{0.3474} = 1.93.$$

M1 A1. With $\nu = 14$, the one-tailed 5% critical value is 1.761. $1.93 > 1.761$, so reject H_0 : there is sufficient evidence to suggest that brand P batteries last longer on average than brand Q batteries. **A1**

(b) $\nu = 14$, $p = 0.975$: $t = 2.145$. **B1**

$$0.6714 \pm 2.145 \times 0.3474 = 0.6714 \pm 0.745.$$

M1 The interval is $[-0.074, 1.42]$ hours. **A1**

(c) The interval contains 0, which matches a **two-tailed** 5% test finding no evidence of a difference. Part (a) was a **one-tailed** 5% test, which matches a 90% interval: $0.6714 - 1.761 \times 0.3474 = 0.060 > 0$, so 0 lies outside the 90% interval, as the rejection in (a) requires. **B1**

Example 4

A researcher records the daily screen time, in hours, of random samples of students at two colleges, X and Y .

	n	Σ	Σ of squares
College X (x)	70	290.5	1274.2

	n	Σ	Σ of squares
College Y (y)	90	333	1310.4

(a) Find a 95% confidence interval for $\mu_X - \mu_Y$. **[5]**

(b) Test, at the 5% significance level, whether the mean daily screen time at college X exceeds that at college Y by more than 0.12 hours. **[4]**

(c) Find the set of values of α for which the test in (b) at the $\alpha\%$ significance level gives evidence for this claim. **[2]**

(a) $\bar{x} = 4.15, \bar{y} = 3.7$.

$$s_x^2 = \frac{1}{69} \left(1274.2 - \frac{290.5^2}{70} \right) = 0.9946, \quad s_y^2 = \frac{1}{89} \left(1310.4 - \frac{333^2}{90} \right) = 0.8798.$$

M1 A1. The samples are large, so no pooling and no normality assumption are needed.

$$\frac{0.9946}{70} + \frac{0.8798}{90} = 0.02398, \quad \sqrt{0.02398} = 0.1549.$$

M1 Interval: $0.45 \pm 1.96 \times 0.1549$, **B1** for 1.96, giving $[0.146, 0.754]$ hours. **A1**

(b) $H_0 : \mu_X - \mu_Y = 0.12, H_1 : \mu_X - \mu_Y > 0.12$.

$$z = \frac{0.45 - 0.12}{0.1549} = 2.131.$$

M1 A1. $2.131 > 1.645$, so reject H_0 . **M1** There is sufficient evidence to suggest that the mean daily screen time at college X is more than 0.12 hours greater than at college Y . **A1**

(c) $P(Z > 2.131) = 1 - 0.9835 = 0.0165$. **M1** The test gives evidence when $\alpha\% \geq 1.65\%$, so $\alpha \geq 1.65$. **A1**

Common mistakes

- **Using the biased variance.** Dividing by n instead of $n - 1$ loses the mark straight away and spoils everything after it. Examiners also see variance and standard deviation mixed up: $s/\sqrt{n}$ needs s , not s^2 .
- **Wrong table value.** The commonest error in t intervals: the wrong row ($\nu = n$ instead of $n - 1$, or forgetting $n_x + n_y - 2$ for pooled) or the wrong column (one-tailed p for a two-sided interval). Write " $\nu = \dots, p = \dots$ " before you read the table.
- **z where t is needed, and t where z is needed.** Small sample, unknown variance: t . Large samples: z . Using 1.645 or 1.96 in a small-sample interval scores no method mark, and so does a t value with two large samples.
- **Pooling when you shouldn't (and not pooling when you should).** Pool only when the question says the population variances are equal (or tells you to assume it). Every year some candidates pool in large-sample questions with no such statement, and some ignore an equal-variance statement in a small-sample question.

- **Treating paired data as two samples.** Before-and-after tables or numbered pairs need the differences. Examiners also see arithmetic slips in the subtractions: check each one, and keep negative signs.
- **Muddled direction in a paired test.** Decide "before — after" or "after — before" at the start and keep it in the hypotheses, the calculation and the conclusion.
- **Loose assumptions.** "The data are normal" or "the population mean is normal" score nothing; for a paired test the assumption must be about the **population of differences**. Questions that say "state any assumptions" are often answered without any.
- **Hypotheses in words that miss "population".** Use μ (defined) or say "population mean".
- **Conclusions that are too certain or out of context.** "This proves the mean has increased" or "reject H_0 " alone lose the last mark. Write "there is sufficient (insufficient) evidence to suggest..." in the words of the question.
- **Working back from an interval with algebra.** Combining every formula into one expression before substituting leads to pages of errors. Find $\bar{x}$, then s , then Σx^2 , one number at a time.
- **Rounding too early.** Keep at least 4 significant figures in s^2 , s_p^2 and standard errors so the final 3 s.f. answer is right.

How to get the marks

- **Hypotheses (B1)** are in terms of population parameters: μ , $\mu_x - \mu_y$ or μ_d , with the symbols defined if they aren't the question's. " $H_0 : \mu_d = k$ " must use the question's k .
- **Show the estimate.** Write s^2 (or s_p^2 , or the variance of $\bar{X} - \bar{Y}$) as a number. A correct value implies the method; a wrong one with no working earns nothing.
- **Show the critical value** you use and the comparison, for example " $2.71 > 1.860$, reject H_0 ". The comparison mark needs consistent signs: compare a negative t with a negative critical value.
- **The conclusion (A1)** comes from correct working, is in context, and has doubt in it ("sufficient evidence to suggest", never "proves"). For a claim, "sufficient evidence to support the claim" is fine.
- **Intervals** are given as two numbers, $[a, b]$ or $a < \mu < b$, to 3 significant figures. The reversed order $[b, a]$ is usually condoned; " $\bar{x} \pm e$ " is not.
- **"State an assumption"** means one precise sentence naming the population and the distribution. For a given reason in context, refer to the situation in the question, not a general remark.
- **"Use your confidence interval to explain"** means quote the interval and say whether 0 (or the claimed value) lies inside it.
- **"Find the set of values"** wants an inequality, such as $a \geq 1.65$.
- **Method marks follow through.** A wrong s^2 used correctly still earns the later method marks, so always finish the test.

Quick check

1. A random sample of 7 values from population A gives an unbiased variance estimate of 3.2. A random sample from population B is 14, 17, 12, 16, 11. The populations have the same variance σ^2 . Find a pooled estimate of σ .
2. A random sample of 16 observations from a normal population gives the 95% confidence interval $[23.4, 26.2]$ for the population mean, using the t -distribution. Find Σx and Σx^2 .
3. A random sample of 11 values from a normal population gives $\Sigma x = 583.0$ and $\Sigma x^2 = 30\,937.4$. Test, at the 10% significance level, whether the population mean differs from 52.
4. Random samples of 45 from population X and 64 from population Y give $\bar{x} = 18.6$, $s_x^2 = 5.2$, $\bar{y} = 17.1$, $s_y^2 = 6.9$. Find a 90% confidence interval for $\mu_X - \mu_Y$.
5. Use your interval from question 4 to say what a two-tailed test at the 10% level of $H_0 : \mu_X = \mu_Y$ would conclude, and why.

Answers

- Sample A: $\Sigma(x - \bar{x})^2 = 6 \times 3.2 = 19.2$. Sample B: $\bar{x} = 14$, $\Sigma(x - \bar{x})^2 = 0 + 9 + 4 + 4 + 9 = 26$.
 $s_p^2 = \frac{19.2 + 26}{7 + 5 - 2} = 4.52$, so the estimate of σ is $\sqrt{4.52} = 2.13$.
- $\bar{x} = 24.8$, so $\Sigma x = 16 \times 24.8 = 396.8$. $\nu = 15$, $t = 2.131$; half-width $1.4 = 2.131 \times \frac{s}{4}$, so $s = 2.6279$
 and $s^2 = 6.906$. $\Sigma x^2 = 15 \times 6.906 + \frac{396.8^2}{16} = 9944.2$.
- $H_0 : \mu = 52$, $H_1 : \mu \neq 52$. $\bar{x} = 53$, $s^2 = \frac{1}{10} \left(30\,937.4 - \frac{583^2}{11} \right) = 3.84$. $t = \frac{53 - 52}{\sqrt{3.84/11}} = 1.693$. Two-tailed 10% with $\nu = 10$: critical value 1.812. $1.693 < 1.812$, so do not reject H_0 : there is insufficient evidence that the population mean differs from 52.
- $\frac{5.2}{45} + \frac{6.9}{64} = 0.22337$, $\sqrt{0.22337} = 0.4726$. $1.5 \pm 1.645 \times 0.4726 = 1.5 \pm 0.777$, giving $[0.723, 2.28]$.
- 0 is not in the 90% interval, so the two-tailed 10% test would reject H_0 : there is sufficient evidence of a difference between the population means.

Past questions to try

- [9231/44/M/J/25 Q6](#): work back from a t interval, then a pooled two-sample test.
- [9231/43/M/J/25 Q5](#): a paired test and the range of k for which it rejects.
- [9231/43/M/J/24 Q6](#): a large-sample interval, then a test with a nonzero difference.
- [9231/43/O/N/25 Q3](#): a pooled estimate from a given s and a raw list.