inkMaths

ANSWER BOOK

# Confidence intervals and tests with the t-distribution

Further Statistics 2 · Year FM

6 questions · 26 marks

NAVY EDITION


inkphysics.co.uk/maths

## SECTION A · Warm-up

**A1** It is used when the population is normal but its variance is unknown and estimated by $s^2$ from the same sample, particularly when that sample is small. The t-distribution is symmetric about zero like the normal but has heavier tails, so its critical values are further out, and it approaches the normal as the degrees of freedom grow. B1 for the condition of use, B1 for the heavier tails. [2]

**A2** The sample mean is $216/9 = 24$, and the degrees of freedom are $n - 1 =$ **8**. One is lost because s is estimated from the same sample, and quoting 9 here is the commonest slip in the topic. [3]
Standard error $= s/\sqrt{n} = 4.5/3 =$ **1.5**.
$t = (24 - 22)/1.5 =$ **1.33**.
B1 for the degrees of freedom, M1 for the standard error, A1 for the statistic.

## SECTION B · Standard

**B1** $H_0$: $\mu = 50$; $H_1$: $\mu > 50$, on $16 - 1 = 15$ degrees of freedom. B1 for both hypotheses and the degrees of freedom, which is a mark before any arithmetic. [5]
Standard error $= 6/\sqrt{16} = 1.5$, so $t = (52.5 - 50)/1.5 =$ **1.667**.
Since $1.667 < 1.753$, **do** not reject $H_0$. There is insufficient evidence at the 5% level that the mean exceeds 50.
M1 for the standard error, A1 for the statistic, M1 for the comparison with the critical value, A1 for the conclusion in context. 'Accept $H_0$' does not earn that last mark, because the test never establishes that $\mu$ is 50.

**B2** Standard error $= 6/\sqrt{16} = 1.5$, so the interval is $52.5 \pm 2.131 \times 1.5 = 52.5 \pm 3.20$, that is (**49.30,** 55.70). [4]
The value 50 lies inside the interval, so the data are **consistent** with a mean of 50.
M1 for the form of the interval, A1 for the half-width, A1 for the interval, B1 for the verdict on 50.
Using 1.96 in place of 2.131 gives (49.56, 55.44), which is too narrow. The extra width is the price of estimating the variance from the sample, and it is why the t value is always the larger of the two.

**B3** Work with the differences alone, never with the raw before and after readings. [6]
Subtracting within pairs removes the variation between subjects, which is usually far larger than the effect being measured, and that is the whole reason for pairing.
$H_0$: $\mu_d = 0$; $H_1$: $\mu_d > 0$, on $8 - 1 = 7$ degrees of freedom.
The differences total 16, so the mean difference is $16/8 =$ **2**. Their squared deviations from 2 total 28, so $s^2 = 28/7 = 4$ and $s = 2$. Divide by 7, not by 8; the divisor $n - 1$ is what makes $s^2$ unbiased.
Standard error $= 2/\sqrt{8} = 0.707$, so $t = 2/0.707 =$ **2.83**.
Since $2.83 > 1.895$, **reject** $H_0$. There is evidence at the 5% level that the treatment raises the measurement.
B1 for the hypotheses and the degrees of freedom, M1 A1 for the mean difference and the sample standard deviation, M1 for the standard error, A1 for the statistic, A1 for the conclusion in context.
Treating these data as two independent samples of 8 would inflate the standard error and hide the effect.

## SECTION C · Stretch

**C1** $H_0$: $\mu_1 = \mu_2$; $H_1$: $\mu_1 \neq \mu_2$, on $n_1 + n_2 - 2 = 6 + 9 - 2 =$ **13** degrees of freedom. Two are lost, one for each mean estimated, and 14 is the usual wrong answer here. [6]
Pool the variances, weighting each by its own degrees of freedom rather than by its sample size: $s_p^2 = [5(2.4^2) + 8(3.1^2)]/13 = (28.8 + 76.88)/13 = 105.68/13 =$ **8.129**, so $s_p =$ 2.851. A plain average of 5.76 and 9.61 would give 7.685 and lose both marks for this stage.
Standard error of the difference $= s_p\sqrt{(1/n_1 + 1/n_2)} = 2.851 \times \sqrt{(1/6 + 1/9)} = 2.851 \times 0.527 =$ **1.503**.
$t = (18.2 - 15.8)/1.503 = 2.4/1.503 =$ **1.60**.
Since $1.60 < 2.160$, **do** not reject $H_0$. There is insufficient evidence at the 5% level that the population means differ.
B1 for the hypotheses and the degrees of freedom, M1 A1 for the pooled variance, M1 for the standard error, A1 for the statistic, A1 for the conclusion in context.
The equal-variance assumption is doing a lot of work, and with sample standard deviations of 2.4 and 3.1 it would be worth testing with an F test before relying on it.