Less Stress More Success Maths Leaving Cert Higher Level Paper 2

Revised pages for Chapter 13 Statistics IV: The Normal Curve, z-Scores, Hypothesis Testing and Simulation

(b) The sample proportion $= \hat{p} = \frac{960}{1,600} = 0{\cdot}6 = 60\%$

The 95% confidence interval for the proportion of people, p, who want an election now is:

$$\hat{p} - \frac{1}{\sqrt{n}} \leq p \leq \hat{p} + \frac{1}{\sqrt{n}}$$

$$0{\cdot}6 - 0{\cdot}025 \leq p \leq 0{\cdot}6 + 0{\cdot}025$$

$$0{\cdot}575 \leq p \leq 0{\cdot}625$$

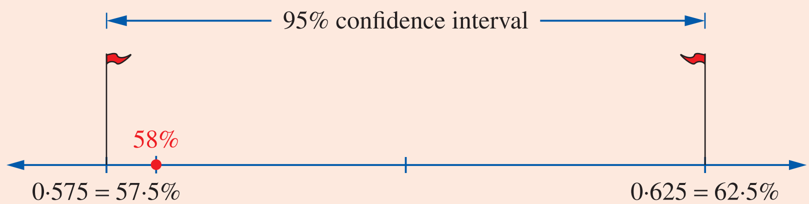
$$57{\cdot}5\% \leq p \leq 62{\cdot}5\%$$

Thus, the 95% confidence level for the proportion of people who want an election now is between 57·5% and 62·5%, inclusive.

Answers can also be given in brackets: [0·575, 0·625] or [57·5%, 62·5%].

(ii) Null hypothesis $= H_0 =$ there is no change in support for an election now.

In the second survey, $\hat{p} = \frac{696}{1,200} = 0{\cdot}58$ (58%).



58% is inside the 95% confidence interval in the first survey. Thus, we do not reject the null hypothesis.

key point

If the second sample proportion was outside the 95% confidence interval, 57·5 to 62·5, we would reject the null hypothesis.

exam focus

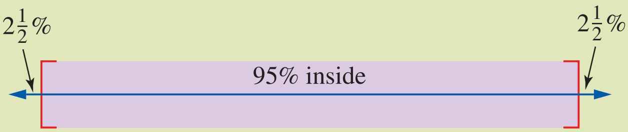
What does '95% confidence' really mean?

What do we mean when we say we have 95% confidence that our interval contains the randomly selected value? Formally, what we mean is that '95% of randomly selected values will be captured/fall into the confidence interval'. This is correct, but somewhat long winded, so we usually say 'we are 95% confident that the (randomly) selected value lies in our interval.'

Our uncertainty is about whether the particular (randomly) selected value is one of the successful ones or one of the 5% that falls outside the interval.

key point

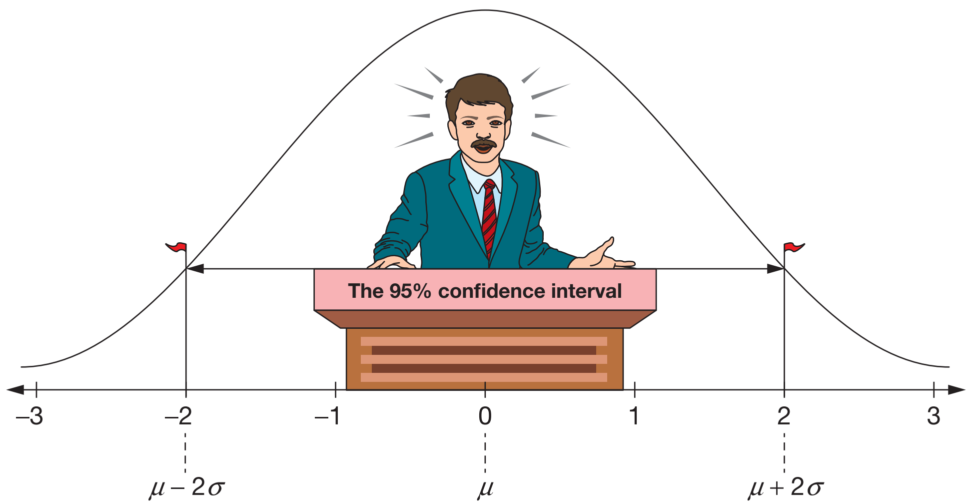
When working with levels of confidence (or levels of significance), statisticians can use percentages ambiguously. In particular, the 5% level of significance and the 95% level of confidence mean the same thing. That is to say, 5% of the time outside the confidence interval or 95% of the time inside the confidence interval.



The normal curve and confidence intervals

Given data that is approximately normal and using the mean (μ) and standard deviation (σ), we can calculate the z-score for each and every data point x, using $z = \dfrac{x - \mu}{\sigma}$.

The empirical rule tells us that about 95% of all randomly selected data points will be within $\pm 2\sigma$ from μ or within ± 2 from 0. That is to say, if we reach out 2 standard deviations on both sides from the mean, we are sure to 'trap' 95% of the data.



exam Q

A scientific expedition discovers a large colony of birds. The weights x kg of a random sample of 200 of these birds are measured and the following results obtained:

$$\sum x = 224{\cdot}4, \quad \sum (x - \mu)^2 = 5{\cdot}823$$

(i) Calculate unbiased estimates of the mean, μ, and the standard deviation, σ, of the weights of these birds.

(ii) Based on previous experience, a teacher has claimed that in these circumstances, half of all students will measure the angle correctly to within two degrees. Taking these students to be a simple random sample, and assuming the true value of the angle is the one you calculated in part **(i)**, is there sufficient evidence to reject the teacher's claim at the 5% level of significance?

Solution

(i) There are several methods that will lead to success in this question.

One method is to draw an ordered stem and leaf plot and use the median as the best estimate.

Median value = 18
Good idea because median
is not influenced by outliers.

1	2 4 5 5 5 5 6 6 6 6 8 8 8 8 8
2	0 0 0 1 1 2 2 4
3	
4	
5	
6	
7	0 3

Key $2|1 = 21°$

Note: Two students seem to have very atypical answers. However, on reflection they may have simply misread the clinometer, reading the complementary angle:

i.e. $90 - 70 = 20°$ and $90 - 73 = 17°$

(ii) Taking 18° from **(i)**, then 12 students measured correctly to within 2°. You may check this yourself by counting from the stem and leaf plot.

How many measured between 18 ± 2?

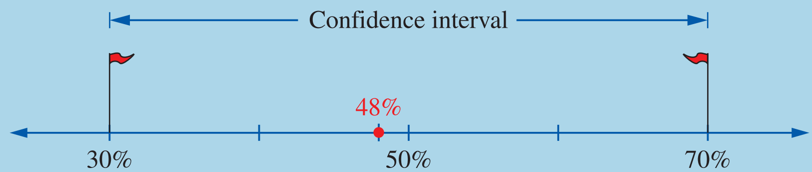
How many measured between 16 and 20 inclusive?

Answer $= \frac{12}{25} = 48\%$ of students.

We must know the 95% margin of error for a sample of size $n = 25$ is given by
$\dfrac{1}{\sqrt{n}} = \dfrac{1}{\sqrt{25}} = \dfrac{1}{5} = 0{\cdot}2 = 20\%$.

In the question, the teacher claims that half (= 50%) of the students will measure the angle correctly to within 2°.

Since the margin of error is $\pm 20\%$, we expect $50\% \pm 20\%$ of students to measure correctly. That is, from 30% to 70% of students to measure correctly. Since 48% 'fits' between 30% to 70% $\Rightarrow$ there is not sufficient evidence to reject the teacher's claim.



Example

(i) The breaking strengths of cables produced by a manufacturer have a mean of 1,800 N (Newtons) and a standard deviation of 100 N. By a new technique in the manufacturing process, it is claimed that the breaking strength can be increased. To test this claim, a sample of 50 new cables is tested and it is found that their mean breaking strength is 1,825 N.

(a) State the null hypothesis.

(b) Is there evidence to reject or accept the null hypothesis, to a 5% level of significance?

(ii) A doctor claims that 17-year-olds have an average body temperature that is higher than the commonly accepted average adult human temperature of 37 degrees Celsius, with a standard deviation of 0·4 degrees. A simple random statistical sample of 25 people, each of age 17, is selected. The average temperature of the 17-year-olds is found to be 37·167 degrees.

(a) State the null hypothesis.

(b) Is there evidence to reject or accept the null hypothesis, to a 5% level of significance?

Solution

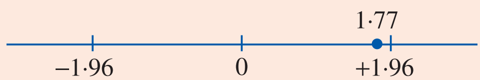
(i) (a) The null hypothesis, H_0, states the breaking strength of the cables is not increased.

(b) The 5% level of significance $\Rightarrow z = \pm 1{\cdot}96$

The standard error for this sample is given by

$$\frac{\sigma}{\sqrt{n}} = \frac{100}{\sqrt{50}} = 14{\cdot}14$$

$$z = \frac{x - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{1{,}825 - 1{,}800}{14{\cdot}14} = 1{\cdot}77$$



The z-score 1·77 falls inside the required confidence interval.

$\therefore$ We do not reject the null hypothesis, H_0, at the 5% level of significance.

(ii) (a) The null hypothesis, H_0, states the average temperature of 17-year-olds is the same as the average temperature of adult humans.