Hypothesis testing

O1	Extend *AS content* to correlation coefficients as measures of how close data points lie to a straight line and be able to interpret a given correlation coefficient using a given p-value or critical value (calculation of correlation coefficients is excluded)
O3	Conduct a statistical hypothesis test for the mean of a Normal distribution with known, given or assumed variance and interpret the results in context

Commentary

The terminology and structure of hypothesis tests is covered at AS level, here the method is adapted to test the strength of a linear correlation and to determine if the evidence supports a suggested population mean for a normal distribution.

The correlation coefficient r summarizes the strength of a linear relationship in samples only. If we obtained a different sample, we would obtain different correlations, and therefore potentially different conclusions. As always, we want to draw conclusions about populations, not just samples. The true correlation coefficient of the population is ρ. We test the existence of correlation in the population by setting $H_0 : \rho = 0$ and H_1 as $\rho > 0, \rho < 0$ or $\rho \neq 0$ accordingly.

As the size of the sample increases it is less likely that the correlation coefficient will be close to 1 by chance (think of $n = 2$, it is certain that $r = 1$) so the minimum correlation coefficient needed (a critical value) to be convinced of a correlation in the population will be lower for larger sample sizes. Equivalently, the probability that a randomly generated sample (from a population with zero correlation) gives a correlation at least as extreme as any observed correlation coefficient (the p-value) will decrease as n increases.

In order to make any sense of a hypothesis test for the mean of a normal distribution, X, it is necessary to know that any calculated sample mean, $\bar{x}$, (from a random sample drawn from a normal distribution) is itself drawn from a population, $\bar{X}$. This population is known as the sampling distribution of the mean. For samples of size n drawn from $X \sim N(\mu, \sigma^2)$ this sampling distribution of the mean is $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$. Notice how larger sample sizes create sampling distributions with smaller variances – this is justifies our intuition that larger sample sizes tend to give better estimates of population parameters. An important assumption is that, whilst we are testing the mean we are assuming the variance of the parent population is unchanged.

If an observed sample mean m falls sufficiently far from the population mean μ suggested by the null hypothesis, then we have statistical evidence that the null hypothesis is unlikely to be true. This is done in a very similar way to how we test a proportion given by the Binomial distribution – assuming the null hypothesis is true we calculate the probability of a result at least as extreme as m (the p-value) or create a critical region of extreme values based on the significance level. Working with continuous data is actually an advantage as the critical region can be made to fit the significance level exactly (subject to the rounding of a z value).

Sample resource

'Correlation game' (which can be found at https://my.integralmaths.org/integral/sow-resources.php) is a simple game where players are invited to guess the correlation in a given scatter diagram. Good for building an understanding of the visual representation of different values of correlation.
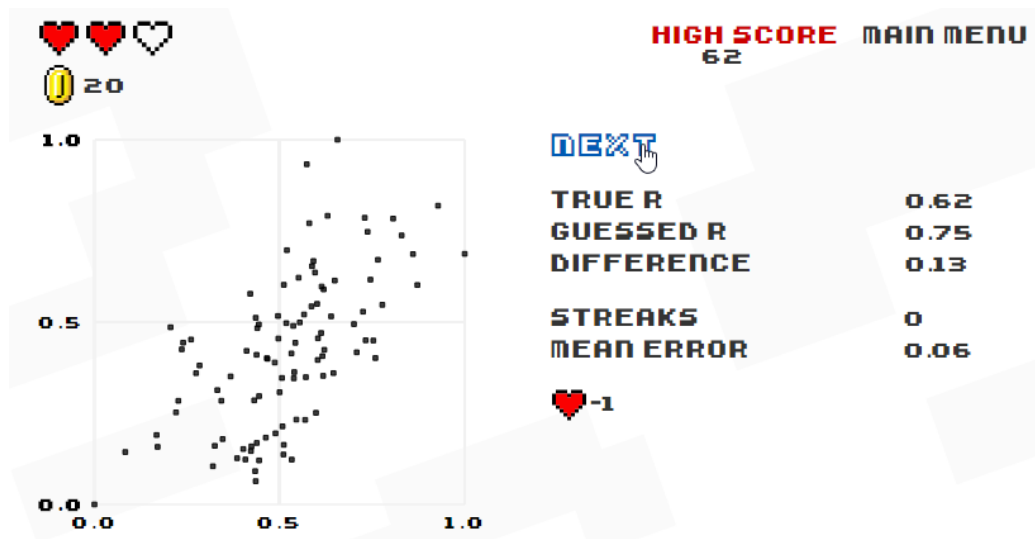
Effective use of technology

'Sampling distributions' (which can be found at http://www.mei.org.uk/integrating-technology) is designed for demonstrating to students how a distribution of sample means can be generated from any population. Click the link, click the Begin button, then experiment with the buttons and drop down menus.
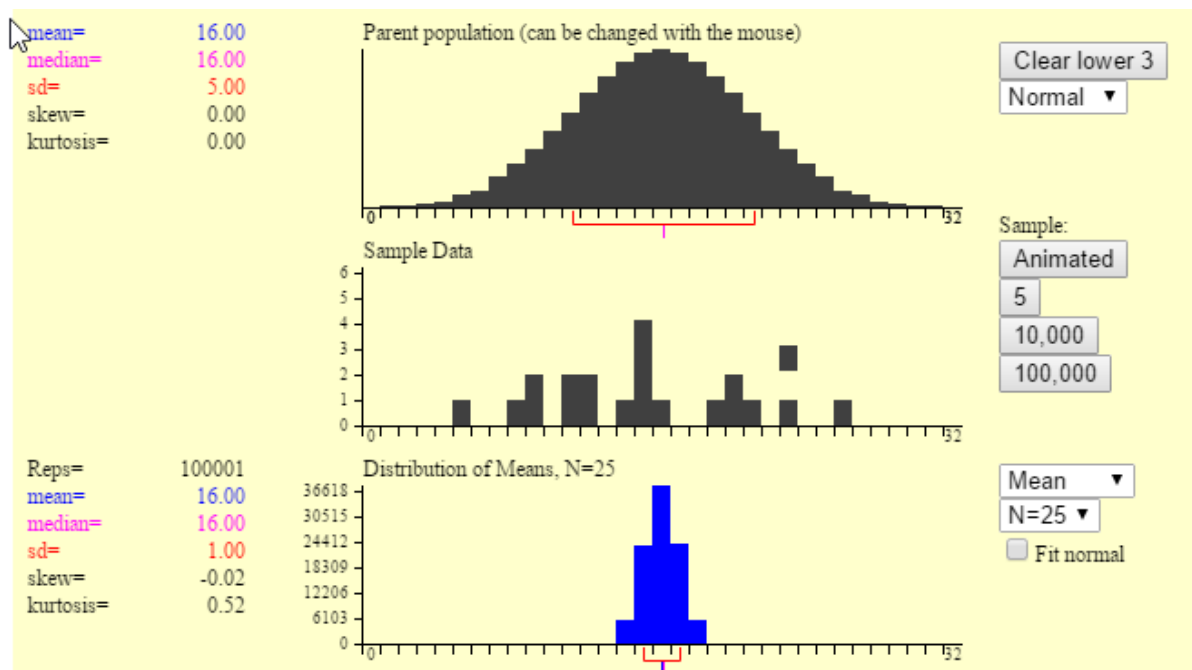
Hypothesis testing	Time allocation:

Pre-requisites

- Hypothesis testing (AS)
- Probability distributions
- Data processing, presentation and interpretation (AS)
-

Links with other topics

- Functions: Inequalities of the form $|x - a| > b$ for 2-tailed tests
-

Questions and prompts for mathematical thinking

- Present a scatter diagram with a small number of points – add a point to increase correlation (or to decrease correlation) (or remove a point)
- Two variables have positive correlation – what effect does adding an outlier have – does it always sometimes or never decrease the correlation?
- Think of two variables where you would expect there to be a correlation but there is no causal relationship
- Show two overlapping Normal curves with equal variance but different means– one is distribution under null hypothesis, other is actual distribution – what values of $\bar{x}$ cause acceptance/rejection of null hypothesis? What happens as the mean of the actual distribution gets nearer to/further from the hypothesised mean?
- In hypothesis testing for a the mean of a random sample drawn from a large population we have assumed that the population is normal and the variance is known. Often in practice we know nothing about the population from which we are sampling. Can we still use the same methods?
-

Applications and modelling

- Are you Clairvoyant? Try the test http://www.psychicscience.org/esp3.aspx which uses a Normal approximation to the Binomial to give z and p values for the number of cards predicted after different numbers of trials.
- Have you ever seen the e-mark symbol on a packet of crisps? The average quantity of product in a batch of packages shall not be less than the nominal quantity stated on the label. Find out more online – search for "e-mark".
-

Common errors

- Thinking that a p-value tells you how likely the null hypothesis is.
- Using the parent population's standard deviation rather than the standard error (in effect using $n = 1$ when the sample is larger)
- Confusing a sample statistic with a population parameter.
-

MEI Innovators in Mathematics Education