Statistical hypothesis testing using the binomial distribution (AS)

O1	Understand and apply the language of statistical hypothesis testing, developed through a binomial model: null hypothesis, alternative hypothesis, significance level, test statistic, 1-tail test, 2-tail test, critical value, critical region, acceptance region, *p*-value
O2	Conduct a statistical hypothesis test for the proportion in the binomial distribution and interpret the results in context Understand that a sample is being used to make an inference about the population and appreciate that the significance level is the probability of incorrectly rejecting the null hypothesis

Commentary

There is a substantial amount of vocabulary and terminology which students must understand. Students can be introduced to this through a variety of experiments to investigate the truth of a statement. Consider a situation where students are asked "is the coin I have given you biased?" The first question you need to ask yourself is how convincing does the evidence need to be? This is measured using probability. If you flipped a fair coin 10 times it could show heads every time, but the probability of that happening would be less than 1 in 1000, or 0.1%. For most purposes this would be considered so unlikely that our conclusion would be that the sequence of coin flips did not belong to a fair coin. However if I only flipped the coin three times then the probability of it showing heads every time would be 12.5% - this would not usually be considered something which was sufficiently unlikely to make someone think the coin is biased. The maximum probability at which we start to decide if something is unusual is called the significance level; 5% is often used, but other values can be used depending on the situation being considered.

At AS Level all Hypothesis Tests will relate to the probability of success for a Binomial Distribution.

The null hypothesis, H_0, is the default position in our example we start by assuming that the coin is fair and the probability it shows heads, $p = \frac{1}{2}$. The alternative hypothesis, H_1, is that there has been a change from the position described by the null hypothesis. The alternative hypothesis may be $p > \frac{1}{2}$ if we were suspicious that the coin may be biased towards heads (a 1 tailed test), or $p \neq \frac{1}{2}$ if we are not sure which way the coin is biased (a 2 tailed test). Data must be collected using a random sampling procedure that ensures the outcomes are independent.

The probability of rejecting the null hypothesis when it is true, sometimes called a Type I error, is equal to the significance level by definition – there is always a chance that a fair coin might do something unusual!

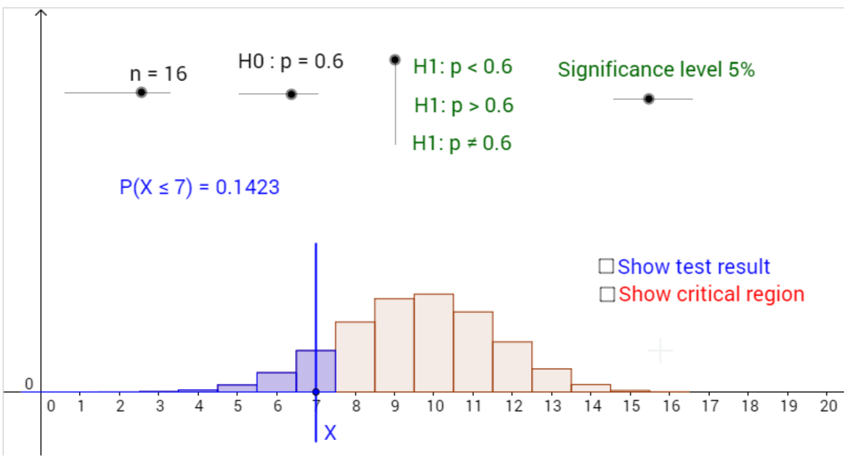
Sample MEI resource

'Matching Critical Regions' (which can be found at https://integralmaths.org/sow-resources.php) is designed to help students get a feel for one and two tailed tests and develop the idea of critical regions.
The 8 hypothesis cards and the 8 critical region cards should be cut up – they both have missing parts to them, which should be filled in for practicing

$H_0: p = 0.4$ H_1: $\boxed{}$ If H_0 is true, $X \sim B(20, 0.4)$ Test at 1 % significance level	$H_0: p = 0.4$ $H_1: p < 0.4$ If H_0 is true, $X \sim B(20, 0.4)$ Test at 5 % significance level
$H_0: p = 0.4$ $H_1: p < 0.4$ If H_0 is true, $X \sim B(20, 0.4)$ Test at 2 % significance level	$H_0: p = 0.4$ H_1: $\boxed{}$ If H_0 is true, $X \sim B(20, 0.4)$ Test at 1 % significance level

Effective use of technology

'Critical regions' (which can be found at http://mei.org.uk/integrating-technology) is designed to allow students to visualise the critical region of a Hypothesis Test. It is interesting to observe how even the most extreme number of successes is not significant for small n, like the example of the coin in the commentary.



Statistical hypothesis testing using the binomial distribution (AS)	Time allocation:

Pre-requisites

- AS: Binomial distribution
- GCSE/AS: Writing inequalities for critical regions
-

Links with other topics

- A-level: Hypothesis tests relating to the Normal Distribution
- Functions: Inequalities of the form $|x - a| > b$ for 2-tailed tests
-

Questions and prompts for mathematical thinking

- The historical probability that a student at a certain school passes a statistics module is 0.75. A new teacher is appointed and takes the next group. Out of a group of 10 students only 6 students now pass. Should the Head of Maths be concerned? Are the assumptions necessary for a Hypothesis Test fully met? Assuming they are, change one aspect so the Hypothesis test has a different result.
-

Applications and modelling

- Test of ESP. In a class of 20 students you would expect, on average, that one student could produce a result that is unusual at a 5% significance level. Simple experiments like getting students to predict the roll of the dice, or flip of a coin, eliminating students who get it wrong. Does the fact that a student correctly guessed the roll 3 times suggest heightened powers of ESP?
- A company uses a multiple choice test to see whether employees are suitable for promotion. The test has 30 questions, each with 4 answers. What kind of pass mark would ensure that there was very little chance of someone who guess them all passing? Being promoted means working longer hours. The company monitors test results to check for people deliberately getting questions wrong even though they know the answer – what kind of marks would be evidence of this?
-

Common errors

- Forgetting that a hypothesis test does not prove anything, rather it provides evidence.
- Looking at the probability of one outcome (e.g. getting exactly 8 heads from 10 coin tosses) rather than a critical region (8 or more heads from 10 coin tosses).
- Failure to define ‘p’ in the null hypothesis.
- Confusing a one-tailed test for a two-tailed test.
-