AS Statistics – Chapter 7 – Hypothesis Testing using a Binomial Distribution

Hypothesis tests are a way of decision whether an unusual result is **statistically significant** or likely to be the result of random chance. When testing a result from a binomial distribution:

The **null hypothesis** $\boldsymbol{H_0}$ is the assumption that the previous value of p remains true.

The **alternative hypothesis** $\boldsymbol{H_1}$ could be that the previous value of value of p has increased, decreased or changed.

The **significance level** is the probability of incorrectly rejecting the null hypothesis when it is in fact true.

We use this significance level as a cut-off point for how unlikely a result has to be before we reject H_0

Hypothesis tests can be one-tailed (has p increased/decreased?) or two-tailed (has p changed?)

Hypothesis tests need to fit a very precise template, which is set out below as a model solution:

Question

A coin is thought to be biased in favour of heads, with a 0.6 chance of getting a head on any single throw. In a sample of 20 throws, a total of 15 heads are recorded. Test at the 5% significance level whether this is evidence that the probability of getting a head is actually higher than previously thought.

Solution

- **Define your variable and its distribution**

  Let X be the number of heads recorded in 20 throws $\rightarrow$ $Assume\ X \sim B(20\,, 0.6)$

- **State your hypotheses and significance level**

  $H_0:\ p = 0.6$

  $H_1:\ p > 0.6$ (one-tailed, as we are testing whether the rate has **increased**)

  Significance level: $5\% \ = \ 0.05$

- **Write down the relevant probability**

  $P(X \geq 15) = 1 - P(X \leq 14) = 0.1256$ (15 heads, and any more extreme results in the same tail)

- **Compare this to the significance level**

  $0.1256 > 0.05$

  "Result is not significant, do not reject H_0" (15 heads or more isn't actually that unlikely to occur by chance)

- **Make a concluding statement in context**

  "The evidence doesn't suggest the probability of getting a head is higher than previously thought"

Two-Tailed Tests

The above example would be a two-tailed test if we were checking whether the probability of getting a head has "changed" rather than increased. The alternative hypothesis would be $H_1:\ p \neq 0.6$

The test follows the same steps, but we use **half of the significance level** as the 5% probability of rejecting the null hypothesis is shared between the upper and lower tails, so we would compare our probability to 0.025, not 0.05.

Critical Regions

The **critical region** is the range of possible values which would lead to you rejecting the null hypothesis – essentially, a list of which values would be considered **significant**. The value on the boundary of the critical region (the least extreme value for which H_0 is rejected) is called the **critical value**.

A **one-tailed** test has a single critical region, with one critical value.

A **two-tailed** test has two critical regions, and so has two critical values. Remember to halve the significance level!

You need to be able to find critical region for one- and two-tailed tests by showing where the cumulative probabilities cross the significance level, which can require some trial and error.

Using the previous example, with $X \sim B(20\,, 0.6)$ and a one-tailed test of the upper tail at the 5% level, we find that

$$P(X \leq 15) = 0.9491 \;\rightarrow\; P(X \geq 16) = 0.0509 \;>\; 0.05 \quad \leftarrow \quad this\ is\ the\ 5\%\ significance\ level$$

$$P(X \leq 16) = 0.9840 \;\rightarrow\; P(X \geq 17) = 0.0160 \;<\; 0.05$$

Make sure you show the values either side of the significance level! The exam boards expect to see this to confirm that you have checked that your critical value is correct.

So the critical value is 17 and the critical region is $\boldsymbol{X \geq 17}$ or $X = \{17\,, 20\}$

This means we will reject H_0 if there are 17 or more heads, and accept H_0 otherwise.

You may see reference to the **acceptance region** as well, which in this case is $\{0\,, 16\,\}$

If the previous example was two-tailed, with H_1: $p \neq 0.6$ you would show the following:

Lower tail:	Upper tail:
$P(X \leq 7) = 0.0210 \;<\; 0.025$	$P(X \geq 16) = 0.0509 \;>\; 0.025$
$P(X \leq 8) = 0.0565 \;>\; 0.025$	$P(X \geq 17) = 0.0201 \;<\; 0.025$

Note that we are comparing to 2.5%, not 5%, as the test is two-tailed.

The critical values are 7 for the lower tail and 17 for the upper tail.

The critical region is $\{X \leq 7 \;\cup\; X \geq 17\,\}$, which can just be written as " $X \leq 7$ or $X \geq 17$ "

We reject H_0 if there are either 7 heads or fewer, or 17 heads or more, and we accept H_0 otherwise.