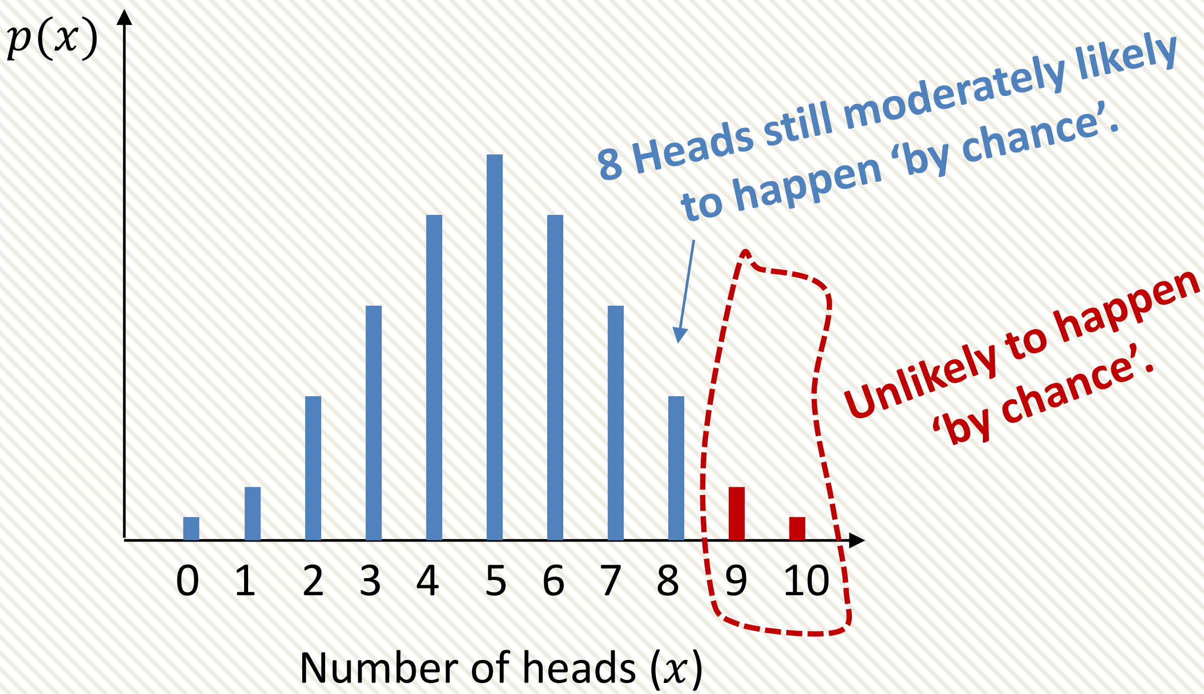
Stats1 Chapter 7: Hypothesis Tests

Null and Alternative Hypothesis

What is Hypothesis Testing?

I throw a coin 10 times. For what numbers of heads might you conclude that the coin is biased towards heads? Why?

Our intuition is that the further away we are from the 'expected' number of heads (i.e. 5 heads out of 10), the more unlikely it is.



What is Hypothesis Testing?

I throw a coin 10 times. For what numbers of heads might you conclude that the coin is biased towards heads? Why?

✎ A hypothesis is a statement made about the value of a **population parameter** that we wish to test by collecting evidence in the form of a sample.

✎ The **null hypothesis**, H_0 is the default position, i.e. that nothing has changed, unless proven otherwise.

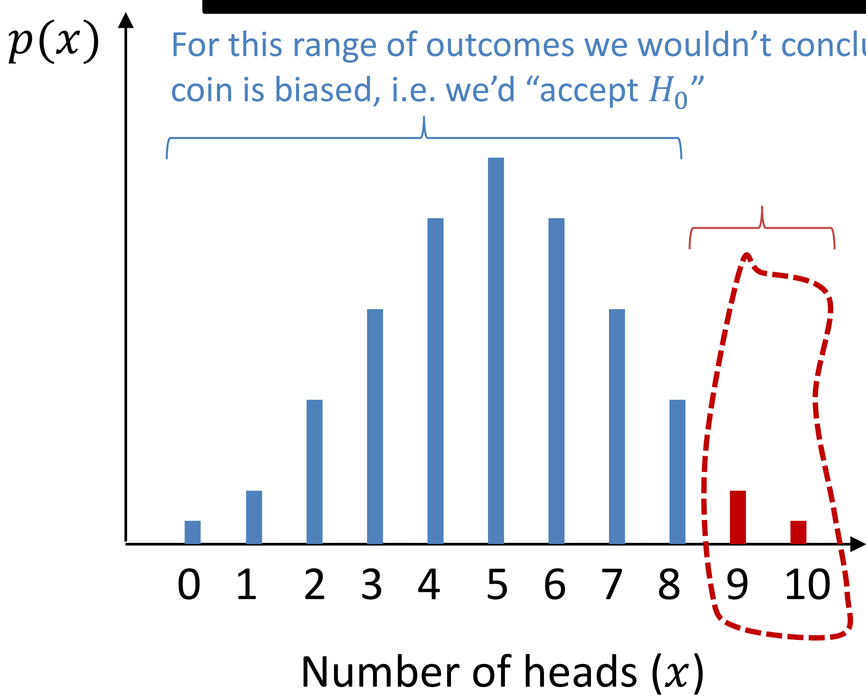
✎ The **alternative hypothesis**, H_1, is that there has been some change in the population parameter.

In this context…

We're asking "is the coin biased", so the ***population parameter*** is the probability p of getting heads (i.e. the p in $B(n,p)$)

The 'default position' is that the coin is fair, i.e. $p = 0.5$.

The 'alternative' position is that the coin is biased towards heads, i.e. p is more than 0.5.



For this range of outcomes we'd conclude that this number of heads was too unlikely to happen by chance, and hence reject H_0 (i.e. that coin was fair) and accept H_1 (i.e. that coin was biased).

What is Hypothesis Testing?

I throw a coin 10 times. For what numbers of heads might you conclude that the coin is biased towards heads? Why?

> ✎ In a hypothesis test, the evidence from the sample is a **test statistic**.

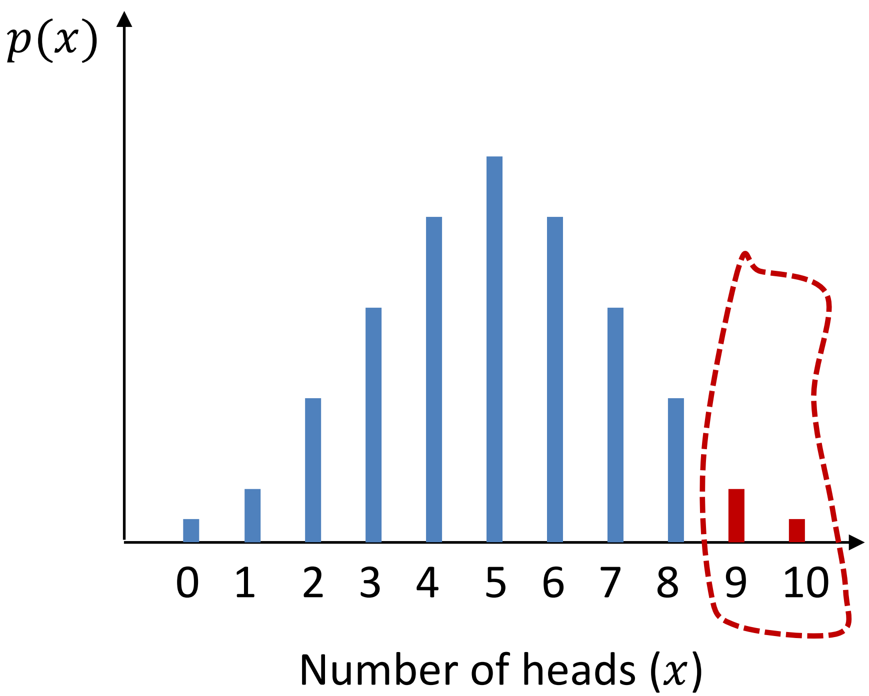
In this context…

The test statistic X is **what we observed**, in this case, X is the number of heads seen in 10 throws.

Note that the test statistic **is a distribution** (i.e. across the possible things we might observe).

$$X \sim B(10, p)$$

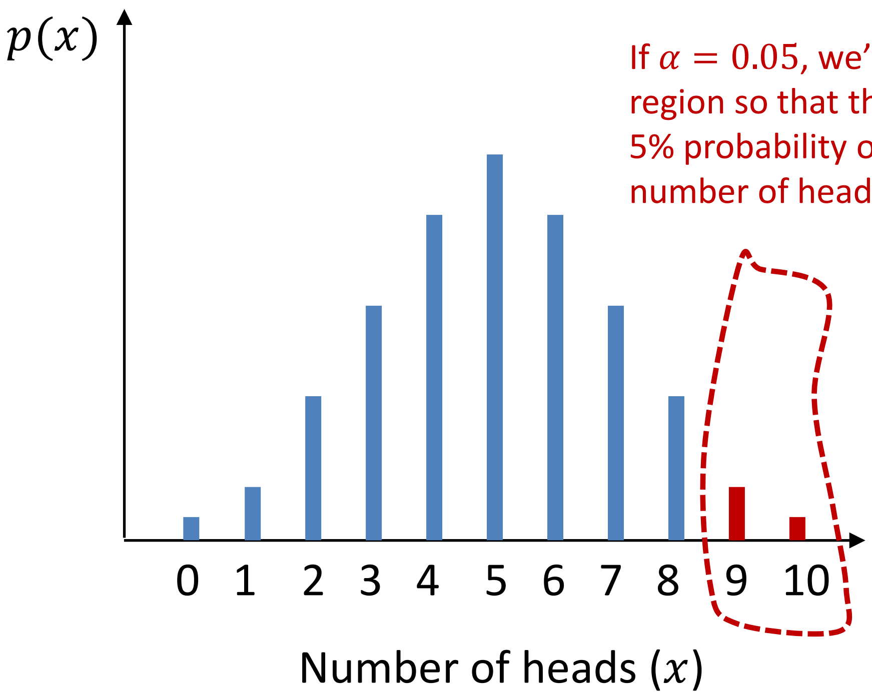
noting that p is not known until we start making assumptions.



What is Hypothesis Testing?

I throw a coin 10 times. For what numbers of heads might you conclude that the coin is biased towards heads? Why?

> ✎ The **level of significance α** is the maximum probability where we would reject the null hypothesis.
> This is usually 5% or 1%.



If $\alpha = 0.05$, we've set this region so that there's at most a 5% probability of seeing these number of heads 'by chance'.

In this context...

We said that if we saw a number of heads within ranges of outcomes that were sufficiently unlikely, then we'd rule out that the coin is fair and conclude it was in fact biased.

But how unlikely is 'sufficiently unlikely'? If $\alpha = 5\%$, then we'd find a region of outcomes where there's (at most) a 5% chance of one of these extreme values happening 'by chance' (i.e. if the coin was fair).

What is Hypothesis Testing?

Hypothesis testing in a nutshell then is:

1. We have some hypothesis we wish to see if true (e.g. coin is biased towards heads), so…
2. We collect some sample data by throwing the coin (giving us our 'test statistic') and…
3. If that number of heads (or more) is sufficiently unlikely to have emerged 'just by chance', then we conclude that our (alternative) hypothesis is correct, i.e. the coin is biased.

Null Hypothesis and Alternative Hypothesis

[Textbook] John wants to see whether a coin is unbiased or whether **it is biased towards coming down heads**. He tosses the coin 8 times and counts the number of times X, it lands head uppermost.

We said that our two hypotheses are about the population parameter.

Suppose p is the probability of a coin landing heads.

Null hypothesis: $H_0: p = 0.5$
Alternative hypothesis: $H_1: p > 0.5$

Under the **null hypothesis** H_0, we **assume that the population parameter is correct**, in this case, that it is a normal coin and the probability of heads is 0.5

Under the **alternative hypothesis** H_1, there has been an underlying change in the population parameter, in this case that the coin is actually biased towards heads.

The latter is known as a '**one-tailed test**' because we're saying the coin is biased one way or the other (i.e. $p > 0.5$ or $p < 0.5$).
But we could also have had the hypothesis 'the coin is biased (either way)', i.e. $p \neq 0.5$. This is known as a **two-tailed test**.

Further Example

[Textbook] An election candidate believes she has the support of 40% of the residents in a particular town. A researcher wants to test, at the 5% significance level, whether the candidate is over-estimating her support. The researcher asks 20 people whether they support the candidate or not. 3 people say they do.

a) Write down a suitable test statistic.
b) Write down two suitable hypotheses.
c) Explain the condition under which the null hypothesis would be rejected.

a

?

For a hypothesis test involving the binomial distribution, the test statistic is always the **count of successes**.

b

?

The alternative hypothesis is that the candidate is **overestimating** her support, so we're interested where **less than 40%** support them (more than 40% would not undermine the candidate's claim).

c

?

This is the hard bit!
We always calculate the probability of seeing this outcome **<u>or more extreme</u>** (in this case, 'more extreme' meaning even fewer the 3 people, because this takes us even further from the expected number of people out of the 20 (i.e. 8) who would support them.
The "$p = 0.4$" bit is because, as discussed before, we calculate the probability of seeing the observed outcome of 3 people (or more extreme) if it occurred **purely by chance** (the null hypothesis), i.e. if the candidate **did** have 40% support.

Further Example

[Textbook] An election candidate believes she has the support of 40% of the residents in a particular town. A researcher wants to test, at the 5% significance level, whether the candidate is over-estimating her support. The researcher asks 20 people whether they support the candidate or not. 3 people say they do.

a) Write down a suitable test statistic.
b) Write down two suitable hypotheses.
c) Explain the condition under which the null hypothesis would be rejected.

a The number of people who say they support the candidate.

For a hypothesis test involving the binomial distribution, the test statistic is always the **count of successes**.

b $H_0: p = 0.4$
$H_1: p < 0.4$

The alternative hypothesis is that the candidate is **overestimating** her support, so we're interested where **less than 40%** support them (more than 40% would not undermine the candidate's claim).

c Null hypothesis would be rejected if the probability of 3 **or fewer people** supporting the candidate is less than 5%, given that $p = 0.4$

This is the hard bit!
We always calculate the probability of seeing this outcome **or more extreme** (in this case, 'more extreme' meaning even fewer the 3 people, because this takes us even further from the expected number of people out of the 20 (i.e. 8) who would support them.
The "$p = 0.4$" bit is because, as discussed before, we calculate the probability of seeing the observed outcome of 3 people (or more extreme) if it occurred **purely by chance** (the null hypothesis), i.e. if the candidate **did** have 40% support.

Test Your Understanding

In the UK, 5% of students turn up late to school each day. Mr Hameed wishes to determine, to a 10% significance level if St Trinian's School has a problem with attendance. He stands at the front gate one day and finds that 6 of the 40 students who pass him are late.

a) Write down a suitable test statistic.
b) Write down two suitable hypotheses.
c) Explain the condition under which the null hypothesis would be rejected.

a ?

b ?

If there is indeed a <u>problem</u> with attendance (i.e. the school is not the norm) then we expect the proportion p of students late at St Trinian's to be higher than 5%.

c ?

Just to reiterate again what's going on here:

a) We assume that the school is the norm (i.e. 5% of students are late), and calculate the probability that 6 or more students would be late under this assumption.
b) If the probability of this happening by chance is sufficiently unlikely (less than the 10% significance level), we conclude that it probably isn't just by chance, and the school's lateness rate is worse than 5%.

Test Your Understanding

In the UK, 5% of students turn up late to school each day. Mr Hameed wishes to determine, to a 10% significance level if St Trinian's School has a problem with attendance. He stands at the front gate one day and finds that 6 of the 40 students who pass him are late.

a) Write down a suitable test statistic.
b) Write down two suitable hypotheses.
c) Explain the condition under which the null hypothesis would be rejected.

a The number of students who are late to school that day.

b $H_0: p = 0.05$
$H_1: p > 0.05$

If there is indeed a <u>problem</u> with attendance (i.e. the school is not the norm) then we expect the proportion p of students late at St Trinian's to be higher than 5%.

c Null hypothesis would be rejected if the probability of 6 **or more** students being late to school is less than 10%, given that $p = 0.05$

Just to reiterate again what's going on here:

a) We assume that the school is the norm (i.e. 5% of students are late), and calculate the probability that 6 or more students would be late under this assumption.
b) If the probability of this happening by chance is sufficiently unlikely (less than the 10% significance level), we conclude that it probably isn't just by chance, and the school's lateness rate is worse than 5%.

Exercise 7.1

Pearson Applied Year 1/AS

Pages 45

Homework Exercise

1 **a** Explain what you understand by a hypothesis test.

b Define a null hypothesis and an alternative hypothesis and state the symbols used for each.

c Define a test statistic.

2 For each of these hypotheses, state whether the hypotheses given describe a one-tailed or a two-tailed test:

a $H_0: p = 0.8$, $H_1: p > 0.8$ **b** $H_0: p = 0.6$, $H_1: p \neq 0.6$ **c** $H_0: p = 0.2$, $H_1: p < 0.2$

3 Dmitri wants to see whether a dice is biased towards the value 6. He throws the dice 60 times and counts the number of sixes he gets.

a Describe the test statistic.

b Write down a suitable null hypothesis to test this dice.

c Write down a suitable alternative hypothesis to test this dice.

Hint If the dice is biased towards 6 then the probability of landing on 6 will be greater than $\frac{1}{6}$.

4 Shell wants to test to see whether a coin is biased. She tosses the coin 100 times and counts the number of times she gets a head. Shell says that her test statistic is the probability of the coin landing on heads.

a Explain the mistake that Shell has made and state the correct test statistic for her test.

b Write down a suitable null hypothesis to test this coin.

c Write down a suitable alternative hypothesis to test this coin.

Homework Exercise

Ⓟ **5** In a manufacturing process the proportion (p) of faulty articles has been found, from long experience, to be 0.1.
A sample of 100 articles from a new manufacturing process is tested, and 8 are found to be faulty.
The manufacturers wish to test at the 5% level of significance whether or not there has been a reduction in the proportion of faulty articles.

a Suggest a suitable test statistic.

b Write down two suitable hypotheses.

c Explain the condition under which the null hypothesis is rejected.

Ⓟ **6** Polls show that 55% of voters support a particular political candidate. A newspaper releases information showing that the candidate avoided paying taxes the previous year. Following the release of the information, a polling company asked 20 people whether they support the candidate. 7 people said that they did. The polling company wants to test at the 2% level of significance whether the level of support for the candidate has reduced.

a Write down a suitable test statistic.

b Write down two suitable hypotheses.

c Explain the condition under which the null hypothesis would be accepted.

Homework Answers

1 a A hypothesis is a statement made about the value of a population parameter. A hypothesis test uses a sample or an experiment to determine whether or not to reject the hypothesis.

b The null hypothesis (H_0) is what we assume to be correct and the alternative hypothesis (H_1) tells us about the parameter if our assumption is shown to be wrong.

c The test statistic is used to test the hypothesis. It could be the result of the experiment or statistics calculated from a sample.

2 a One-tailed test

b Two-tailed test

c One-tailed test

3 a The test statistic is N – the number of sixes.

b H_0: $p = \frac{1}{6}$ **c** H_1: $p > \frac{1}{6}$

4 a Shell is describing what her experiment wants to test rather than the test statistic. The test statistic is the proportion of times you get a head.

b H_0: $p = \frac{1}{2}$ **c** H_1: $p \neq \frac{1}{2}$

5 a A suitable test statistic is p – the proportion of faulty articles in a batch.

b H_0: $p = 0.1$, H_1: $p < 0.1$

c If the probability of the proportion being 0.08 or less is 5% or less the null hypothesis is rejected.

6 a A suitable test statistic is p – the proportion of people that support the candidate.

b H_0: $p = 0.55$, H_1: $p < 0.55$

c If the probability of the proportion being $\frac{7}{20}$ is 2% or more, the null hypothesis is accepted