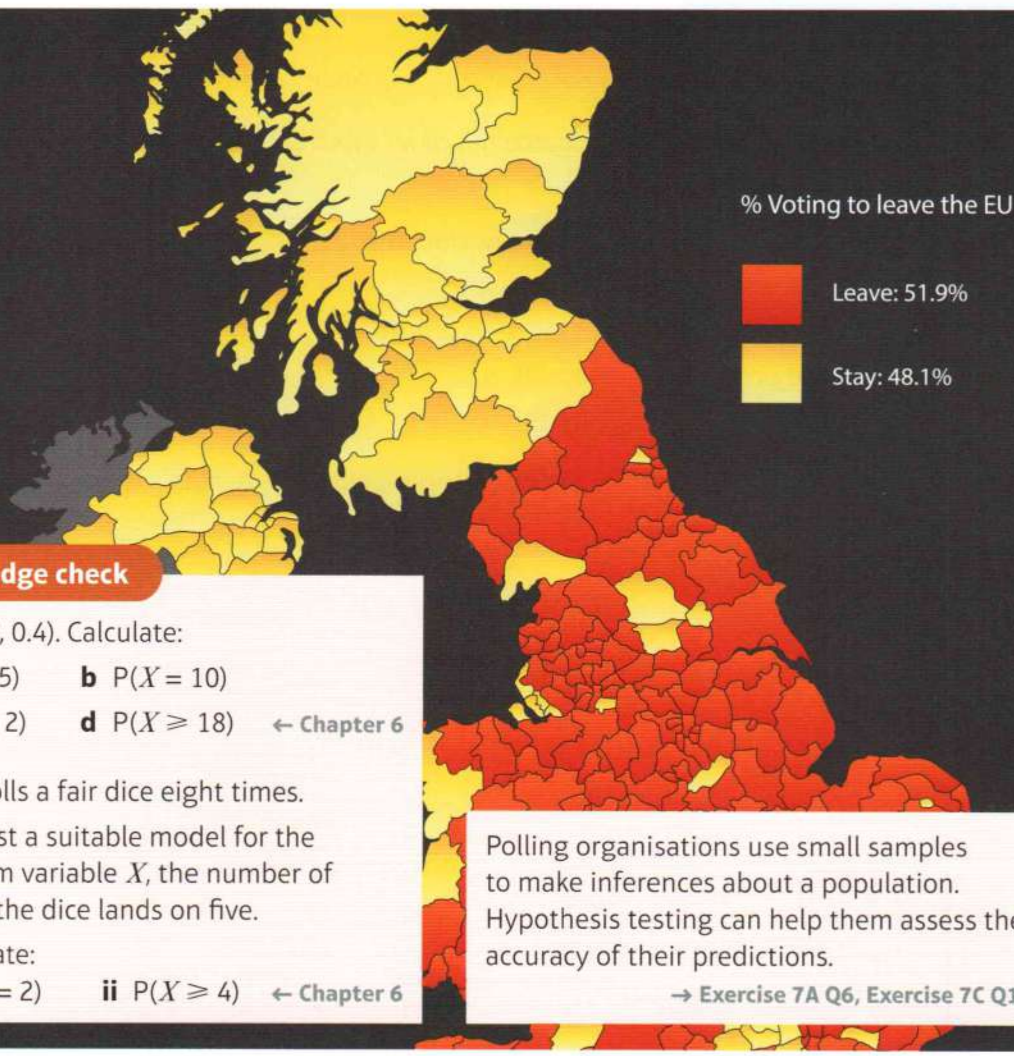
7 Hypothesis testing

Objectives

After completing this chapter you should be able to:

- Understand the language and concept of hypothesis testing → pages 99–101
- Understand that a sample is used to make an inference about a population → pages 99–101
- Find critical values of a binomial distribution using tables → pages 101–105
- Carry out a one-tailed test for the proportion of the binomial distribution and interpret the results → pages 105–107
- Carry out a two-tailed test for the proportion of the binomial distribution and interpret the results → pages 107–109



Prior knowledge check

1 $X \sim \mathrm{B}(20, 0.4)$. Calculate:

a $\mathrm{P}(X = 5)$ **b** $\mathrm{P}(X = 10)$

c $\mathrm{P}(X \leqslant 2)$ **d** $\mathrm{P}(X \geqslant 18)$ ← Chapter 6

2 Wanda rolls a fair dice eight times.

a Suggest a suitable model for the random variable X, the number of times the dice lands on five.

b Calculate:

i $\mathrm{P}(X = 2)$ **ii** $\mathrm{P}(X \geqslant 4)$ ← Chapter 6

Polling organisations use small samples to make inferences about a population. Hypothesis testing can help them assess the accuracy of their predictions.

→ Exercise 7A Q6, Exercise 7C Q10

7.1 Hypothesis testing

A hypothesis is a statement made about the value of a **population parameter**. You can test a hypothesis about a population by carrying out an experiment or taking a sample from the population.

Links In this chapter the population parameter you will be testing will be the probability, p, in a binomial distribution $B(n, p)$. ← Chapter 6

The result of the experiment or the statistic that is calculated from the sample is called the **test statistic**.

In order to carry out the test, you need to form two hypotheses:

- **The null hypothesis, H_0, is the hypothesis that you assume to be correct.**
- **The alternative hypothesis, H_1, tells you about the parameter if your assumption is shown to be wrong.**

Notation In this chapter you should always give H_0 and H_1 in terms of the population parameter p.

Example 1

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, that it lands head uppermost.

a Describe the test statistic.

b Write down a suitable null hypothesis.

c Write down a suitable alternative hypothesis.

a The test statistic is X (the number of heads in 8 tosses).

The test statistic is calculated from the sample or experiment.

b If the coin is unbiased the probability of a coin landing heads is 0.5 so H_0: $p = 0.5$ is the null hypothesis.

You always write the null hypothesis in the form H_0: $p = \dots$

c If the coin is biased towards coming down heads then the probability of landing heads will be greater than 0.5. H_1: $p > 0.5$ is the alternative hypothesis.

If you were testing the coin for bias towards tails your alternative hypothesis would be H_1: $p < 0.5$. If you were testing the coin for bias in either direction your alternative hypothesis would be H_1: $p \neq 0.5$.

- **Hypothesis tests with alternative hypotheses in the form H_1: $p < \dots$ and H_1: $p > \dots$ are called one-tailed tests.**
- **Hypothesis tests with an alternative hypothesis in the form H_1: $p \neq \dots$ are called two-tailed tests.**

Hint You can think of a two-tailed test such as H_1: $p \neq 0.5$ as **two tests**, H_1: $p > 0.5$ or $p < 0.5$.

To carry out a hypothesis test you **assume the null hypothesis is true**, then consider how likely the observed value of the test statistic was to occur. If this likelihood is less than a given threshold, called the **significance level** of the test, then you reject the null hypothesis.

Typically the significance level for a hypothesis test will be 10%, 5% or 1% but you will be told which level to use in the question.

Example 2

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 that 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 test statistic is the number of people who say they support the candidate.

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

c The null hypothesis will be rejected if the probability of 3 or fewer people saying they support the candidate is less than 5%, given that $p = 0.4$.

This is a one-tailed test – the researcher wants to see if the candidate is **over-estimating** her support – if she is then the actual proportion of residents who support her will be **less than** 40%.

Watch out You are testing to see whether the actual probability is **less** than 0.4, so you would need to calculate the probability that the observed value of the test statistic is 3 or **fewer**.

Exercise 7A

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 $\frac{1}{6}$.

(P) 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.

(P) **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.

(P) **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.

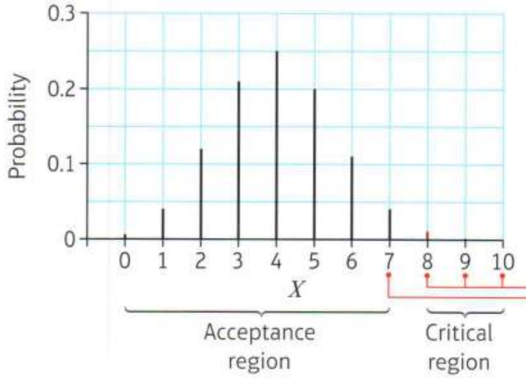
7.2 Finding critical values

When you carry out a hypothesis test, you need to be able to calculate the probability of the test statistic taking particular values given that the null hypothesis is true.

In this chapter you will assume that the test statistic can be modelled by a binomial distribution. You will use this to calculate probabilities and find critical regions.

- **A critical region is a region of the probability distribution which, if the test statistic falls within it, would cause you to reject the null hypothesis.**

A test statistic is modelled as B(10, p) and a hypothesis test at the 5% significance level uses H_0: $p = 0.4$, H_1: $p > 0.4$. Assuming H_0 to be true, X has the following distribution: $X \sim B(10, 0.4)$:



$P(X = 10) = 0.0001$, $P(X = 9) = 0.0016$ and $P(X = 8) = 0.0106$. Hence $P(X \geqslant 8) = 0.0123$. This is less than the significance level of 5%.
A test statistic of 10, 9 or 8 would lead to the null hypothesis being rejected.

$P(X = 7) = 0.0425$. Adding this probability to $P(X \geqslant 8)$ takes the probability over 0.05 so a test statistic of 7 or less would lead to the null hypothesis being accepted.

- **The critical value is the first value to fall inside of the critical region.**

In this example, the critical value is 8 and the critical region is 8, 9 or 10. 7 falls in the **acceptance region**, the region where we accept the null hypothesis.

The critical value and hence the critical region can be determined from binomial distribution tables, or by finding cumulative binomial probabilities using your calculator.

In the $n = 10$ table, the critical region is found by looking for a probability such that $P(X \geqslant x) < 0.05$.

$p =$	0.20	0.25	0.30	0.35	0.40	0.45	0.50
$n = 10, x = 0$	0.1074	0.0563	0.0282	0.0135	0.0060	0.0025	0.0010
1	0.3758	0.2440	0.1493	0.0860	0.0464	0.0233	0.0107
2	0.6778	0.5256	0.3828	0.2616	0.1673	0.0996	0.0547
3	0.8791	0.7759	0.6496	0.5138	0.3823	0.2660	0.1719
4	0.9672	0.9219	0.8497	0.7515	0.6331	0.5044	0.3770
5	0.9936	0.9803	0.9527	0.9051	0.8338	0.7384	0.6230
6	0.9991	0.9965	0.9894	0.9740	0.9452	0.8980	0.8281
7	0.9999	0.9996	0.9984	0.9952	0.9877	0.9726	0.9453
8	1.0000	1.0000	0.9999	0.9995	0.9983	0.9955	0.9893
9	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9990

$P(X \leqslant 6) = 0.9452$ so $P(X \geqslant 7) = 0.0548$ (>0.05)

$P(X \leqslant 7) = 0.9877$ so $P(X \geqslant 8) = 0.0123$ (<0.05)

This shows that $x = 7$ is not extreme enough to lead to the rejection of the null hypothesis but that $x = 8$ is. Hence $x = 8$ is the critical value and $x \geqslant 8$ is the critical region.

The probability of the test statistic falling within the critical region, given that H_0 is true is 0.0123 or 1.23%. This is sometimes called the **actual significance level** of the test.

- **The actual significance level of a hypothesis test is the probability of incorrectly rejecting the null hypothesis.**

Watch out The threshold probability for your test (1%, 5%, 10%) is often referred to as the level of significance for your test. This might be different from the actual significance level, which is the probability that your test statistic would fall within the critical region even if H_0 is true.

Example 3

A single observation is taken from a binomial distribution $B(6, p)$. The observation is used to test $H_0: p = 0.35$ against $H_1: p > 0.35$.

a Using a 5% level of significance, find the critical region for this test.

b State the actual significance level of this test.

a Assume H_0 is true then $X \sim B(6, 0.35)$

$P(X \geqslant 4) = 1 - P(X \leqslant 3) = 1 - 0.8826$
$= 0.1174$

Use tables or your calculator to find the first value of x for which $P(X \geqslant x) < 0.05$.

$P(X \geqslant 5) = 1 - P(X \leqslant 4) = 1 - 0.9777$
$= 0.0223$

$P(X \geqslant 4) > 0.05$ but $P(X \geqslant 5) < 0.05$ so 5 is the critical value.

The critical region is 5 or 6.

Online Find the critical region using your calculator.

b The actual significance level is the probability of incorrectly rejecting the null hypothesis:
P(reject null hypothesis) $= P(X \geqslant 5)$
$= 0.0223$
$= 2.23\%$

This is the same as the probability that X falls within the critical region.

- **For a two-tailed test the critical region is made up of two parts, one at each end of the distribution.**

Example 4

A random variable X has binomial distribution $B(40, p)$. A single observation is used to test $H_0: p = 0.25$ against $H_1: p \neq 0.25$.

a Using a 2% level of significance, find the critical region of this test. The probability in each tail should be as close as possible to 0.01.

b Write down the actual significance level of the test.

a Assume H_0 is true then $X \sim B(40, 0.25)$
Consider the lower tail:
$P(X \leqslant 4) = 0.0160$
$P(X \leqslant 3) = 0.0047$
Consider the upper tail:
$P(X \geqslant 18) = 1 - P(X \leqslant 17) = 1 - 0.9953$
$= 0.0047$
$P(X \geqslant 17) = 1 - P(X \leqslant 16) = 1 - 0.9884$
$= 0.0116$
The critical regions is $0 \leqslant X \leqslant 3$ and $17 \leqslant X \leqslant 40$.

b The actual significance level is $0.0047 + 0.0116 = 0.0163 = 1.63\%$

$P(X \leqslant 3)$ is closest to 0.01 so 3 is the critical value for this tail.

Watch out Read the question carefully. Even though $P(X \geqslant 17)$ is greater than 0.01 it is still the closest value to 0.01. The critical value for this tail is therefore 17.

Online Use technology to explore the locations of the critical values for each tail in this example.

Exercise 7B

1 Explain what you understand by the following terms:
a critical value **b** critical region **c** acceptance region.

2 A test statistic has a distribution $B(10, p)$. Given that $H_0: p = 0.2$, $H_1: p > 0.2$, find the critical region for the test using a 5% significance level.

3 A random variable has a distribution $B(20, p)$. A single observation is used to test $H_0: p = 0.15$ against $H_1: p < 0.15$. Using a 5% level of significance, find the critical region of this test.

(E) **4** A random variable has distribution B(20, p). A single observation is used to test H_0: $p = 0.4$ against H_1: $p \neq 0.4$.

a Using the 5% level of significance, find the critical region of this test. **(3 marks)**

b Write down the actual significance level of the test. **(1 mark)**

5 A test statistic has a distribution B(20, p). Given that H_0: $p = 0.18$, H_1: $p < 0.18$, find the critical region for the test using a 5% level of significance.

Watch out These probabilities are not found in statistical tables. You can use your calculator to find cumulative probabilities for B(n, p) with any values of n and p.

(E/P) **6** A random variable has distribution B(10, p). A single observation is used to test H_0: $p = 0.22$ against H_1: $p > 0.22$.

a Using a 5% level of significance, find the critical region of this test. **(3 marks)**

b Write down the actual significance level of the test. **(2 marks)**

(E/P) **7** A mechanical component fails, on average, 3 times out of every 10. An engineer designs a new system of manufacture that he believes reduces the likelihood of failure. He tests a sample of 20 components made using his new system.

a Describe the test statistic. **(1 mark)**

b State suitable null and alternative hypotheses. **(2 marks)**

c Using a 5% level of significance, find the critical region for a test to check his belief, ensuring the probability is as close as possible to 0.05. **(3 marks)**

d Write down the actual significance level of the test. **(1 mark)**

(E/P) **8** Seedlings come in trays of 36. On average, 12 seedlings survive to be planted on. A gardener decides to use a new fertiliser on the seedlings which she believes will improve the number that survive.

a Describe the test statistic and state suitable null and alternative hypotheses. **(3 marks)**

b Using a 10% level of significance, find the critical region for a test to check her belief. **(3 marks)**

c State the probability of incorrectly rejecting H_0 using this critical region. **(1 mark)**

(E/P) **9** A restaurant owner notices that her customers typically choose lasagne one fifth of the time. She changes the recipe and believes this will change the proportion of customers choosing lasagne.

a Suggest a model and state suitable null and alternative hypotheses. **(3 marks)**

She takes a random sample of 25 customers.

b Find, at the 5% level of significance, the critical region for a test to check her belief. The probability in each tail should be as close as possible to 0.025. **(4 marks)**

c State the probability of incorrectly rejecting H_0. **(1 mark)**

Challenge

A test statistic has binomial distribution $B(50, p)$. Given that:

$H_0: p = 0.7$, $H_1: p \neq 0.7$:

a find the critical region for the test statistic such that the probability in each tail is close as possible to 5%.

Chloe takes two observations of the test statistic and finds that they both fall inside the critical region. Chloe decides to reject H_0.

b Find the probability that Chloe has incorrectly rejected H_0.

7.3 One-tailed tests

If you have to carry out a one-tailed hypothesis test you need to:

- Formulate a model for the test statistic
- Identify suitable null and alternative hypotheses
- Calculate the probability of the test statistic taking the observed value (or higher/lower), assuming the null hypothesis is true
- Compare this to the significance level
- Write a conclusion in the context of the question

Notation This probability is known as the **p-value** for that observation.

Alternatively, you can find the critical region and see whether the observed value of the test statistic lies inside it.

Example 5

The standard treatment for a particular disease has a $\frac{2}{5}$ probability of success. A certain doctor has undertaken research in this area and has produced a new drug which has been successful with 11 out of 20 patients. The doctor claims that the new drug represents an improvement on the standard treatment.

Test, at the 5% significance level, the claim made by the doctor.

X is the number of patients in the trial for whom the drug was successful.

p is the probability of success for each patient.

Define your test statistic, X, and parameter, p.

$X \sim B(20, p)$

$H_0: p = 0.4$ $\quad H_1: p > 0.4$

Write down the model for your test statistic, and your hypotheses. The doctor claims the drug represents an improvement so the alternative hypothesis is $p > 0.4$.

Method 1:

Assume H_0 is true, so $X \sim B(20, 0.4)$

Assume the null hypothesis is true, and calculate the probability of 11 or more successful treatments.

$P(X \geqslant 11) = 1 - P(X \leqslant 10)$

$= 1 - 0.8725$

$= 0.1275$

$= 12.75\%$

Use the cumulative binomial tables or your calculator to find $P(X \leq 10)$. The p-value for this observation is 12.75%.

12.75% > 5% so there is not enough evidence to reject H_0.

Compare the probability to the significance level of your test.

The new drug is no better than the old one.

Make sure you write a conclusion in context.

Method 2:
$P(X \geqslant 13) = 1 - P(X \leqslant 12) = 0.021$
$P(X \geqslant 12) = 1 - P(X \leqslant 11) = 0.0565$
The critical region is 13 or more.
Since 11 does not lie in the critical region, we accept H_0.
There is no evidence that the new drug is better than the old one.

Work out the critical region and see if 11 lies within it.

Unless you are specifically instructed as to which method to use, you can use the one you prefer.

Exercise 7C

1 A single observation, x, is taken from a binomial distribution $B(10, p)$ and a value of 5 is obtained. Use this observation to test H_0: $p = 0.25$ against H_1: $p > 0.25$ using a 5% significance level.

2 A random variable has distribution $X \sim B(10, p)$. A single observation of $x = 1$ is taken from this distribution. Test, at the 5% significance level, H_0: $p = 0.4$ against H_1: $p < 0.4$.

3 A single observation, x, is taken from a binomial distribution $B(20, p)$ and a value of 10 is obtained. Use this observation to test H_0: $p = 0.3$ against H_1: $p > 0.3$ using a 5% significance level.

4 A random variable has distribution $X \sim B(20, p)$. A single observation of $x = 3$ is taken from this distribution. Test, at the 1% significance level, H_0: $p = 0.45$ against H_1: $p < 0.45$.

5 A single observation, x, is taken from a binomial distribution $B(20, p)$ and a value of 2 is obtained. Use this observation to test H_0: $p = 0.28$ against H_1: $p < 0.28$ using a 5% significance level.

6 A random variable has distribution $X \sim B(8, p)$. A single observation of $x = 7$ is taken from this distribution. Test, at the 5% significance level, H_0: $p = 0.32$ against H_1: $p > 0.32$.

(P) 7 A dice used in playing a board game is suspected of not giving the number 6 often enough. During a particular game it was rolled 12 times and only one 6 appeared. Does this represent significant evidence, at the 5% level of significance, that the probability of a 6 on this dice is less than $\frac{1}{6}$?

(P) 8 The success rate of the standard treatment for patients suffering from a particular skin disease is claimed to be 68%.

a In a sample of n patients, X is the number for which the treatment is successful. Write down a suitable distribution to model X. Give reasons for your choice of model.

A random sample of 10 patients receives the standard treatment and in only 3 cases was the treatment successful. It is thought that the standard treatment was not as effective as it is claimed.

b Test the claim at the 5% level of significance.

(E/P) 9 A plant germination method is successful on average 4 times out of every 10. A horticulturist develops a new technique which she believes will improve the number of plants that successfully germinate. She takes a random sample of 20 seeds and attempts to germinate them.

a Using a 5% level of significance, find the critical region for a test to check her belief **(4 marks)**

b Of her sample of 20 plants, the horticulturalist finds that 14 have germinated. Comment on this observation in light of the critical region. **(2 marks)**

Problem-solving

In this question you are told to find the critical region in part **a**. You will save time by using your critical region to answer part **b**.

(P) **10** A polling organisation claims that the support for a particular candidate is 35%. It is revealed that the candidate will pledge to support local charities if elected. The polling organisation think that the level of support will go up as a result. It takes a new poll of 50 voters.

a Describe the test statistic and state suitable null and alternative hypotheses. **(2 marks)**

b Using a 5% level of significance, find the critical region for a test to check the belief. **(4 marks)**

c In the new poll, 28 people are found to support the candidate. Comment on this observation in light of the critical region. **(2 marks)**

7.4 Two-tailed tests

A one-tailed test is used to test when it is claimed that the probability has either gone up, or gone down. A two-tailed test is used when it is thought that the probability has changed in either direction.

- **For a two-tailed test, either double the *p*-value for your observation, or halve the significance level at the end you are testing.**

You need to know which tail of the distribution you are testing. If the test statistic is $X \sim B(n, p)$ then the **expected** outcome is np. If the observed value, x, is lower than this then consider $P(X \leqslant x)$. If the observed value is higher than the expected value, then consider $P(X \geqslant x)$. In your exam it will usually be obvious which tail you should test.

Example 6

Over a long period of time it has been found that in Enrico's restaurant the ratio of non-vegetarian to vegetarian meals is 2 to 1. In Manuel's restaurant in a random sample of 10 people ordering meals, 1 ordered a vegetarian meal. Using a 5% level of significance, test whether or not the proportion of people eating vegetarian meals in Manuel's restaurant is different to that in Enrico's restaurant.

The proportion of people eating vegetarian meals at Enrico's is $\frac{1}{3}$.

X is the number of people in the sample at Manuel's who order vegetarian meals.

p is the probability that a randomly chosen person at Manuel's orders a vegetarian meal.

$H_0 : p = \frac{1}{3}$ $\quad H_1 : p \neq \frac{1}{3}$

Significance level 5%

If H_0 is true $X \sim B(10, \frac{1}{3})$

Hypotheses. The test will be two-tailed as we are testing if they are different.

Problem-solving

You can use the same techniques to hypothesis-test for the **proportion** of a population that have a given characteristic. You could equivalently define the test parameter as the proportion of diners at Manuel's that order a vegetarian meal.

Method 1:

$P(X \leqslant 1) = P(X = 0) + P(X = 1)$

$= \left(\frac{2}{3}\right)^{10} + 10\left(\frac{2}{3}\right)^{9}\left(\frac{1}{3}\right)$

$= 0.017\,34\ldots + 0.086\,70\ldots$

$= 0.104$ (3 s.f.)

$0.104 > 0.025$

There is insufficient evidence to reject H_0.

There is no evidence that proportion of people eating vegetarian meals at Manuel's restaurant is different to Enrico's.

Method 2:

Let c_1 and c_2 be the two critical values.

$P(X \leqslant c_1) \leqslant 0.025$ and $P(x \geqslant c_2) \leqslant 0.025$

For the lower tail:

$P(X \leqslant 0) = 0.017\,341\ldots < 0.025$

$P(X \leqslant 1) = 0.104\,04\ldots > 0.025$

So $c_1 = 0$

For the upper tail:

$P(X \geqslant 6) = 1 - P(X \leqslant 5)$

$= 0.076\,56\ldots > 0.025$

$P(X \geqslant 7) = 1 - P(X \leqslant 6)$

$= 0.019\,66\ldots < 0.025$

So $c_2 = 7$

The observed value of 1 does not lie in the critical region so H_0 is not rejected. There is no evidence that the proportion of people eating vegetarian meals at Manuel's is different to Enrico's.

The expected value would be $10 \times \frac{1}{3} = 3.333\ldots$ The observed value, 1, is less than this so consider $P(X \leqslant 1)$.

You can calculate the probabilities long-hand, like this, or use your calculator with $p = 0.3333$.

We use 0.025 because the test is two-tailed. You could have also doubled your p-value to get 0.208, and then compared this with the original significance level of 0.05.

Conclusion and what it means in context.

The probability in each tail of the critical region should be less than or equal to $0.05 \div 2 = 0.025$.

Use the cumulative binomial function on your calculator, with $n = 10$ and $p = 0.3333$.

Write down a value on either side of the boundary to show that you have determined the correct critical values.

Remember to write a conclusion in the context of the question.

Exercise 7D

1 A single observation, x, is taken from a binomial distribution $B(30, p)$ and a value of 10 is obtained. Use this observation to test $H_0: p = 0.5$ against $H_1: p \neq 0.5$ using a 5% significance level.

2 A random variable has distribution $X \sim B(25, p)$. A single observation of $x = 10$ is taken from this distribution. Test, at the 10% significance level, $H_0: p = 0.3$ against $H_1: p \neq 0.3$.

3 A single observation, x, is taken from a binomial distribution $B(10, p)$ and a value of 9 is obtained. Use this observation to test $H_0: p = 0.75$ against $H_1: p \neq 0.75$ using a 5% significance level.

4 A random variable has distribution $X \sim B(20, p)$. A single observation of $x = 1$ is taken from this distribution. Test, at the 1% significance level, $H_0: p = 0.6$ against $H_1: p \neq 0.6$.

(P) **5** A random variable has distribution $X \sim B(50, p)$. A single observation of $x = 4$ is taken from this distribution. Test, at the 2% significance level, $H_0: p = 0.02$ against $H_1: p \neq 0.02$.

Watch out Although the observed value of 4 appears to be small, the expected value of X is actually $50 \times 0.02 = 1$. You need to consider the upper tail of the distribution: $P(X \geqslant 4)$.

(P) **6** A coin is tossed 20 times, and lands on heads 6 times. Use a two-tailed test with a 5% significance level to determine whether there is sufficient evidence to conclude that the coin is biased.

(P) **7** The national proportion of people experiencing complications after having a particular operation in hospitals is 20%. A hospital decides to take a sample of size 20 from their records.

a Find the critical region, at the 5% level of significance, to test whether or not their proportion of complications differs from the national proportion. The probability in each tail should be as close to 2.5% as possible. **(5 marks)**

b State the actual significance level of the test. **(1 mark)**

The hospital finds that 8 of their 20 patients experienced complications.

c Comment on this finding in light of your critical region. **(2 marks)**

(P) **8** A machine makes glass bowls and it is observed that one in ten of the bowls have hairline cracks in them. The production process is modified and a sample of 20 bowls is taken. 1 of the bowls is cracked. Test, at the 10% level of significance, the hypothesis that the proportion of cracked bowls has changed as a result of the change in the production process. State your hypotheses clearly. **(7 marks)**

(P) **9** Over a period of time, Agnetha has discovered that the carrots that she grows have a 25% chance of being longer than 7 cm. She tries a new type of fertiliser. In a random sample of 30 carrots, 13 are longer than 7 cm. Agnetha claims that the new fertiliser has changed the probability of a carrot being longer than 7 cm. Test Agnetha's claim at the 5% significance level. State your hypotheses clearly. **(7 marks)**

(P) **10** A standard blood test is able to diagnose a particular disease with probability 0.96.
A manufacturer suggests that a cheaper test will have the same probability of success.
It conducts a clinical trial on 75 patients. The new test correctly diagnoses 63 of these patients.
Test the manufacturer's claim at the 10% level, stating your hypotheses clearly. **(7 marks)**

Mixed exercise 7

(P) **1** Mai commutes to work five days a week on a train. She does two journeys a day.
Over a long period of time she finds that the train is late 20% of the time.
A new company takes over the train service Mai uses. Mai thinks that the service will be late more often. In the first week of the new service the train is late 3 times.
You may assume that the number of times the train is late in a week has a binomial distribution.
Test, at the 5% level of significance, whether or not there is evidence that there is an increase in the number of times the train is late. State your hypothesis clearly. **(7 marks)**

E/P **2** A marketing company claims that Chestly cheddar cheese tastes better than Cumnauld cheddar cheese.
Five people chosen at random as they entered a supermarket were asked to say which they preferred. Four people preferred Chestly cheddar cheese.
Test, at the 5% level of significance, whether or not the manufacturer's claim is true.
State your hypothesis clearly. **(7 marks)**

E/P **3** Historical information finds that nationally 30% of cars fail a brake test.

a Give a reason to support the use of a binomial distribution as a suitable model for the number of cars failing a brake test. **(1 mark)**

b Find the probability that, of 5 cars taking the test, all of them pass the brake test. **(2 marks)**

A garage decides to conduct a survey of their cars. A randomly selected sample of 10 of their cars is tested. Two of them fail the test.

c Test, at the 5% level of significance, whether or not there is evidence to support the suggestion that cars in this garage fail less than the national average. **(7 marks)**

E/P **4** The proportion of defective articles in a certain manufacturing process has been found from long experience to be 0.1.
A random sample of 50 articles was taken in order to monitor the production. The number of defective articles was recorded.

a Using a 5% level of significance, find the critical region for a two-tailed test of the hypothesis that 1 in 10 articles has a defect. The probability in each tail should be as near 2.5% as possible. **(4 marks)**

b State the actual significance level of the above test. **(2 marks)**

Another sample of 20 articles was taken at a later date. Four articles were found to be defective.

c Test, at the 10% significance level, whether or not there is evidence that the proportion of defective articles has increased. State your hypothesis clearly. **(5 marks)**

E/P **5** It is claimed that 50% of women use Oriels powder. In a random survey of 20 women, 12 said they did not use Oriels powder.
Test, at the 5% significance level, whether or not there is evidence that the proportion of women using Oriels powder is 0.5. State your hypothesis carefully. **(6 marks)**

E/P **6** The manager of a superstore thinks that the probability of a person buying a certain make of computer is only 0.2.
To test whether this hypothesis is true the manager decides to record the make of computer bought by a random sample of 50 people who bought a computer.

a Find the critical region that would enable the manager to test whether or not there is evidence that the probability is different from 0.2. The probability of each tail should be as close to 2.5% as possible. **(4 marks)**

b Write down the significance level of this test. **(2 marks)**

15 people buy that certain make.

c Comment on this observation in light of your critical region. **(2 marks)**

/P **7** **a** Explain what is meant by:

i a hypothesis test **ii** a critical value **iii** an acceptance region. **(3 marks)**

Johan believes the probability of him being late to school is 0.2. To test this claim he counts the number of times he is late in a random sample of 20 days.

b Find the critical region for a two-tailed test, at the 10% level of significance, of whether the probability he is late for school differs from 0.2. **(5 marks)**

c State the actual significance level of the test. **(1 mark)**

Johan discovers he is late for school in 7 out of the 20 days.

d Comment on whether Johan should accept or reject his claim that the probability he is late for school is 0.2. **(2 marks)**

/P **8** From the large data set, the likelihood of a day with either zero or trace amounts of rain in Hurn in June 1987 was 0.5.

Poppy believes that the likelihood of a rain-free day in 2015 has increased.

In June 2015 in Hurn, 21 days were observed as having zero or trace amounts of rain.

Using a 5% significance level, test whether or not there is evidence to support Poppy's claim. **(6 marks)**

/P **9** A single observation x is to be taken from a binomial distribution B(30, p). This observation is used to test H_0: $p = 0.35$ against H_1: $p \neq 0.35$.

a Using a 5% level of significance, find the critical region for this test. The probability of rejecting either tail should be as close as possible to 2.5%. **(3 marks)**

b State the actual significance level of this test. **(2 marks)**

The actual value of X obtained is 4.

c State a conclusion that can be drawn based on this value giving a reason for your answer. **(2 marks)**

/P **10** A pharmaceutical company claims that 85% of patients suffering from a chronic rash recover when treated with a new ointment.

A random sample of 20 patients with this rash is taken from hospital records.

a Write down a suitable distribution to model the number of patients in this sample who recover when treated with the new ointment. **(2 marks)**

b Given that the claim is correct, find the probability that the ointment will be successful for exactly 16 patients. **(2 marks)**

The hospital believes that the claim is incorrect and the percentage who will recover is lower. From the records an administrator took a random sample of 30 patients who had been prescribed the ointment. She found that 20 had recovered.

c Stating your hypotheses clearly, test, at the 5% level of significance, the hospital's belief. **(6 marks)**

Large data set

You will need access to the large data set and spreadsheet software to answer these questions.

1 The proportion of days with a recorded daily mean temperature greater than 15 °C in Leuchars between May 1987 and October 1987 was found to be 0.163 (3 s.f.).

A meteorologist wants to use a randomly chosen sample of 10 days to determine whether the probability of observing a daily mean temperature greater than 15 °C has increased significantly between 1987 and 2015.

a Using a significance level of 5%, determine the critical region for this test.

b Select a sample of 10 days from the 2015 data for Leuchars, and count the number of days with a mean temperature of greater than 15 °C.

c Use your observation and your critical region to make a conclusion.

2 From the large data set, in Beijing in 1987, 23% of the days from May to October had a daily mean air temperature greater than 25 °C. Using a sample of size 10 from the data for daily mean air temperature in Beijing in 2015, test, at the 5% significance level, whether the proportion of days with a mean air temperature greater than 25 °C increased between 1987 and 2015.

Summary of key points

1 The null hypothesis, H_0, is the hypothesis that you assume to be correct.

2 The alternative hypothesis, H_1, tells us about the parameter if your assumption is shown to be wrong.

3 Hypothesis tests with alternative hypotheses in the form $H_1: p < \ldots$ and $H_1: p > \ldots$ are called one-tailed tests.

4 Hypothesis tests with an alternative hypothesis in the form $H_1: p \neq \ldots$ are called two-tailed tests.

5 A critical region is a region of the probability distribution which, if the test statistic falls within it, would cause you to reject the null hypothesis.

6 The critical value is the first value to fall inside of the critical region.

7 The actual significance level of a hypothesis test is the probability of incorrectly rejecting the null hypothesis.

8 For a two-tailed test the critical region is split at either end of the distribution.

9 For a two-tailed test, either double the *p*-value for your observation, or halve the significance level at the end you are testing.