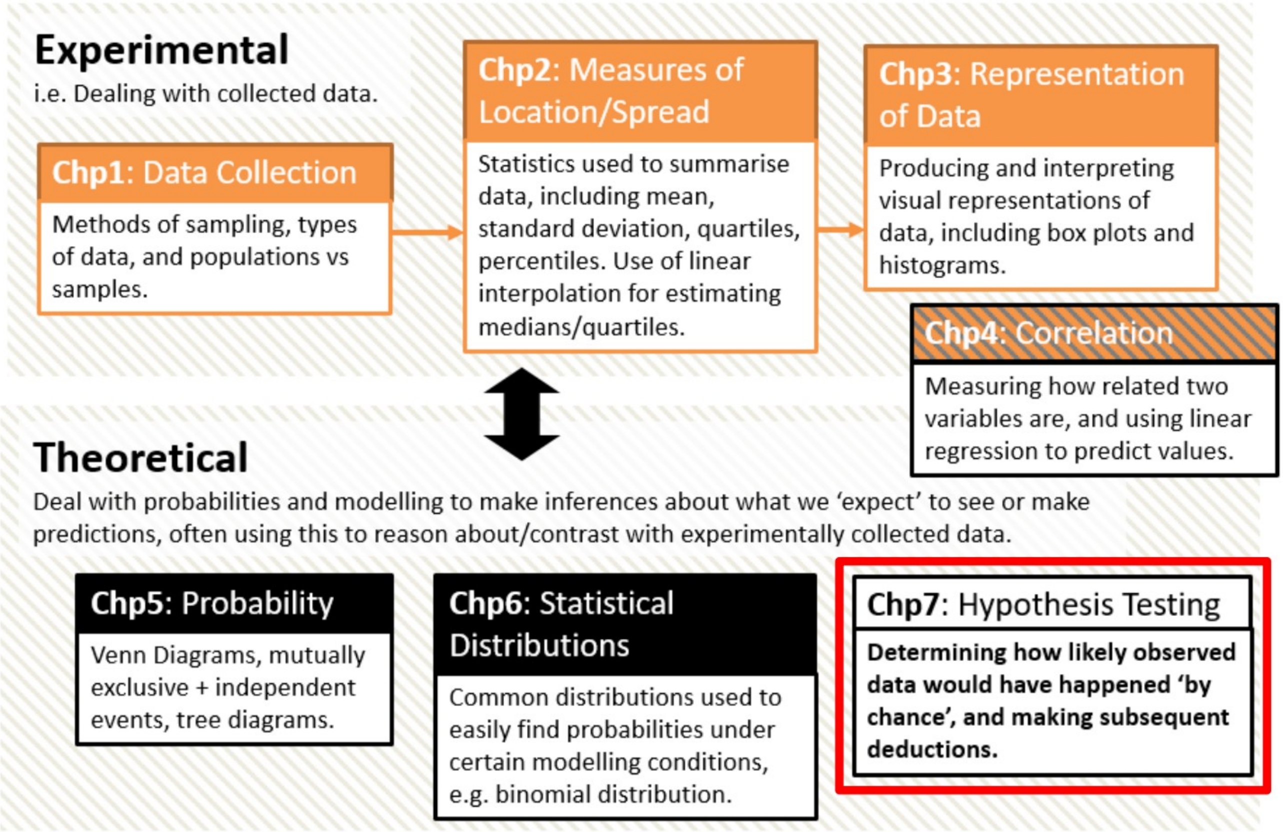
Chapter 7: Hypothesis Testing



I am playing a game that *apparently* gives out a prize 20% of the time.

I play the game 50 times.

What do we expect to happen?
Does this happen?

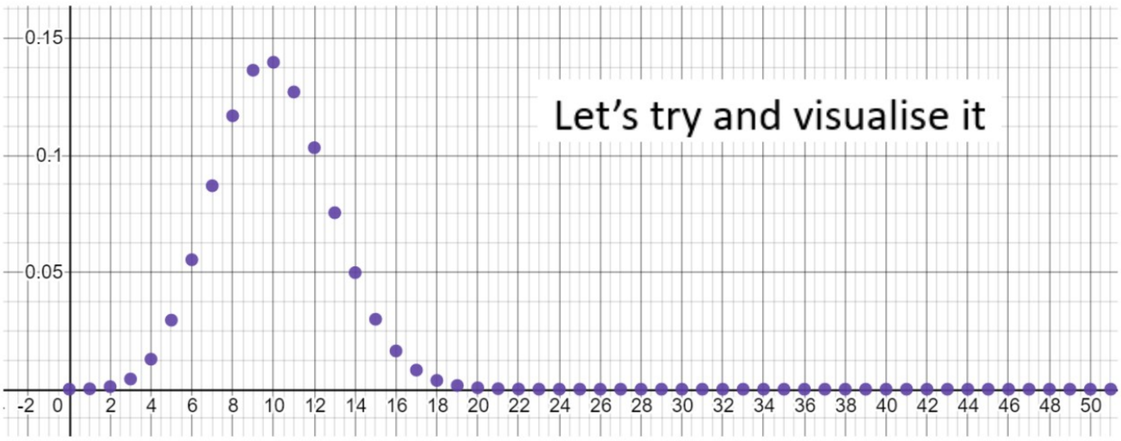
As we know, we won't always win 20% of the games. In fact, the probability of winning 20% of the games is ...

$X \sim B(50, 0.2)$

$P(X = 10) =$

We'd expect most of the time that our results were around 10 games – what's the probability of say...

$P(7 \leq X \leq 13) =$



https://www.desmos.com/calculator/0fnthepm9x

Here's actually what happened from playing the game 50 times (the observed data)

Game number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Result	x	x	x	x	x	x	x	P	x	x	x	x	x	x	P	x	P	x	x	x	x	x	x	x	x
Game number	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
Result	x	x	x	P	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	P	x	x	x

P	means wins a prize
x	means no prize

I ***only*** won 5 TIMES?! The game must be rigged, there's no way it gives a prize 20% of the time!

There are 2 things that getting a 5 or even lower could mean...

So how do we decide?

We set a probability at a particular level (usually 5% or 1%) called a *significance level.* If our observed data is *very unlikely* to have happened 'by chance' then we are suspicious and doubt the information we have been given. If there's a good chance our observed data could have occurred anyway, we're not suspicious.

This process is called hypothesis testing.

<u>**Keywords**</u>:

H_0 - the 'null hypothesis' – the thing that we assume to be true

H_1 - the 'alternative hypothesis' – what we think *could* be true given our observed data

X - the 'test statistic' – the thing that we are observing

Hypothesis test for what we just discussed:

FORMAL LANGUAGE:

WHAT THIS MEANS IN CONTEXT:

(Significance level is 5%)

H_0: $p = 0.2$

H_1: $p < 0.2$

Let X be the number of prizes won.

Assume that $\mathrm{H}_0$is true $-$ i.e. $\mathrm{X} \sim B(50, 0.2)$

Then $P(X \leq 5) = 0.0480 < 0.05$

This suggests that there is evidence to reject H_0.

The probability of winning a prize is less than 20%

Null Hypothesis and Alternative Hypothesis

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.

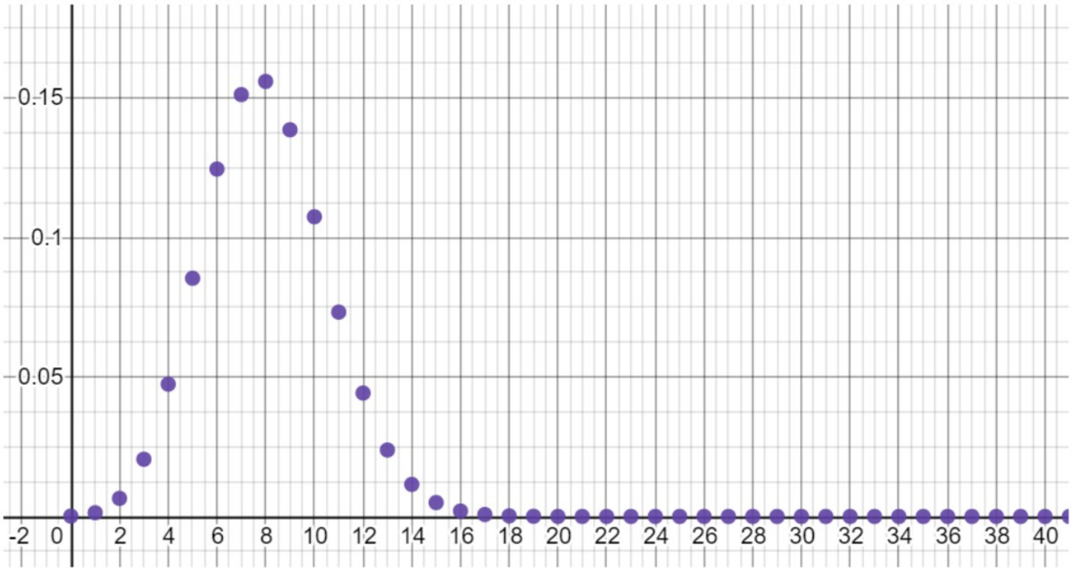
In a hypothesis test, the evidence from the sample is a **test statistic**. For binomial, the test statistic is always the **count of the successes.**

In the UK, 5% of students turn up late to school each day. Mr Bicen wishes to determine, to a 10% significance level if his school, Morpeth 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.

Critical Regions and Values

Back to our game which says a prize is given out 1 in 5 plays.
This time I play the game 40 times.
What observed data would make me suspicious that the game was giving out a prize
a) less than 20% of the time?
b) more than 20% of the time?
(Use a 5% significance level)



Less than 20% of the time? (*Use a 5% significance level)*

$p =$	0.20
$n = 40,\ x = 0$	0.0001
1	0.0015
2	0.0079
3	0.0285
4	0.0759
5	0.1613
6	0.2859
7	0.4371
8	0.5931
9	0.7318
10	0.8392
11	0.9125
12	0.9568
13	0.9806
14	0.9921

✏ The value(s) on the boundary of the critical region are called **critical value(s)**.

✏ The **critical region** is the range of values of the test statistic that would lead to you rejecting H_0

Tip: The low critical value is always the first value in the table that falls below 5% (or whatever the significance level)

✏ The **actual significance level** is the actual probability of being in the critical region.

More than 20% of the time? i.e. the game is giving out too many prizes!
(Use a 5% significance level)

$p =$	0.20
$n = 40, x = 0$	0.0001
1	0.0015
2	0.0079
3	0.0285
4	0.0759
5	0.1613
6	0.2859
7	0.4371
8	0.5931
9	0.7318
10	0.8392
11	0.9125
12	0.9568
13	0.9806
14	0.9921

Tip: The high critical value is always the first value AFTER the one in the table that exceeds 95% (or 100% - significance level)

Determine the critical region when we throw a coin where we're trying to establish if there's the specified bias, given the specified number of throws, when the level of significance is 5%.

$H_0: p = 0.5$
$H_1: p > 0.5$

$p = 0.5, n = 5$

x	$P(X \leq x)$
0	0.0312
1	0.1875
2	0.5000
3	0.8125
4	0.9688

$H_0: p = 0.5$
$H_1: p > 0.5$

$p = 0.5, n = 10$

x	$P(X \leq x)$
0	0.0010
1	0.0107
2	0.0547
...	...
7	0.9453
8	0.9893
9	0.9990

$H_0: p = 0.5$
$H_1: p < 0.5$

$p = 0.5, n = 10$

x	$P(X \leq x)$
0	0.0010
1	0.0107
2	0.0547
...	...
7	0.9453
8	0.9893
9	0.9990

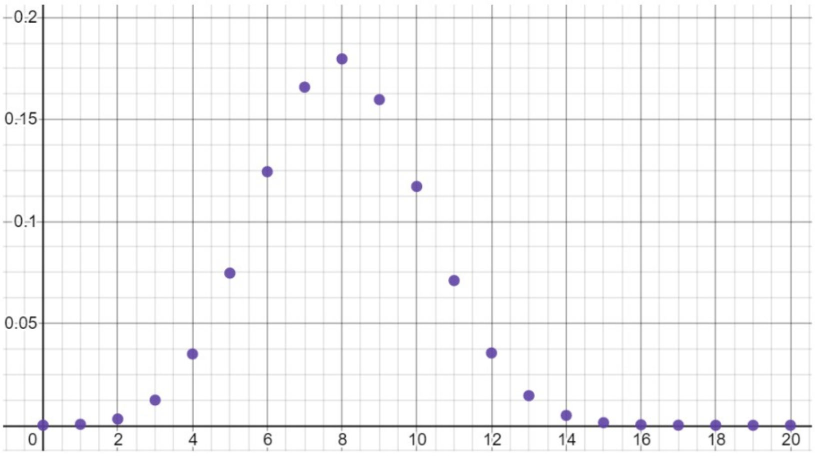
Tip: The low critical value is always the first value in the table that falls below 5% (or whatever the significance level)

Tip: The high critical value is always the first value AFTER the one in the table that exceeds 95% (or 100% - significance level)

One-tailed vs Two-tailed tests

A **'one-tailed test'** is where the alternative hypothesis is either $p > k$ or $p < k$

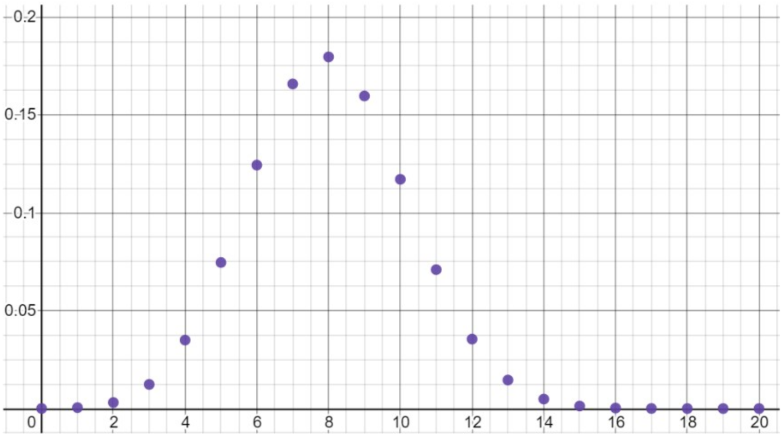
A **'two-tailed test'** is where the alternative hypothesis is $p \neq k$



Suppose I win a game with success rate 40%, and play 20 times.

I'm interested to know the critical regions if my success rate *has changed from 40%.*
(using a 5% significance level)

How do we do this?
Split the 5% significance level so that it is 2.5% for each 'tail'.
Some questions will ask for the actual significance level to be ***closest*** *to 2.5% at each end, rather than* ***below****.*



$p =$	0.40
$n = 20, x = 0$	0.0000
1	0.0005
2	0.0036
3	0.0160
4	0.0510
5	0.1256
6	0.2500
7	0.4159
8	0.5956
9	0.7553
10	0.8725
11	0.9435
12	0.9790
13	0.9935
14	0.9984
15	0.9997
16	1.0000
17	1.0000
18	1.0000

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 the 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.

The $\neq$ indicates bias either way, i.e. two-tailed.

This means you find the closest to 0.01 (even if slightly above) rather than the closest under 0.01

C.D.F. Binomial table: $p = 0.25, n = 40$

x	$P(X \leq x)$
2	0.0010
3	0.0047
4	0.0160
5	0.0433
16	0.9884
17	0.9953
18	0.9983
19	0.9994

Ex 7B

Doing a full one-tailed hypothesis test

John tosses a coin 8 times and it comes up heads 6 times. He claims the coin is **biased towards heads**. With a significance level of 5%, test his claim.

STEP 1: Define test statistic X (stating its distribution), and the parameter p.

STEP 2: Write null and alternative hypotheses.

STEP 3: Determine probability of observed test statistic (or 'more extreme'), assuming null hypothesis.
i.e. Determine probability we'd see this outcome just by chance.

STEP 4: Two-part conclusion:
1. Do we reject H_0 or not?
2. Put in context of original problem.

C.D.F. Binomial table: $p = 0.5, n = 8$

x	$P(X \leq x)$
0	0.0039
1	0.0352
2	0.1445
3	0.3633
4	0.6367
5	0.8555
6	0.9648
7	0.9961

Alternative method using critical regions

We can also find the critical region and see if the test statistic lies within it.

John tosses a coin 8 times and it comes up heads 6 times. He claims the coin is **biased towards heads**. With a significance level of 5%, test his claim.

X is number of heads.
p is probability of heads.
$X \sim B(8, p)$

$H_0: p = 0.5$
$H_1: p > 0.5$

STEP 1: Define test statistic X (stating its distribution), and the parameter p.

STEP 2: Write null and alternative hypotheses.

STEP 3 (Alternative): Determine critical region.

STEP 4: Two-part conclusion:
1. Do we reject H_0 or not?
2. Put in context of original problem.

C.D.F. Binomial table: $p = 0.5, n = 8$

x	$P(X \leq x)$
0	0.0039
1	0.0352
2	0.1445
3	0.3633
4	0.6367
5	0.8555
6	0.9648
7	0.9961

The standard treatment for a particular disease has a $\frac{2}{5}$ probability of success. A certain doctor has undertake research in this area and has produced a new drug which has been successful with 11 out of 20 patients. The doctor claims the new drug represents an improvement on the standard treatment.
Test, at the 5% significance level, the claim made by the doctor.

STEP 1: Define test statistic X (stating its distribution), and the parameter p.

STEP 2: Write null and alternative hypotheses.

STEP 3: Determine probability of observed test statistic (or 'more extreme'), assuming null hypothesis.

STEP 4: Two-part conclusion:
1. Do we reject H_0 or not?
2. Put in context of original problem.

Your Turn

Edexcel S2 Jan 2011 Q2

A student takes a multiple choice test. The test is made up of 10 questions each with 5 possible answers. The student gets 4 questions correct. Her teacher claims she was guessing the answers. Using a one tailed test, at the 5% level of significance, test whether or not there is evidence to reject the teacher's claim.
State your hypotheses clearly. **(6)**

Ex 7C

Two-Tailed Tests

We have already seen that if we're interest in bias 'either way', we have two tails, and therefore have to split the critical region by **halving the significance level at each end**.

Over a long period of time it has been found that in Enrico's restaurant the ratio of non-veg to veg 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 veg meals in Manuel's restaurant is different to that in Enrico's restaurant.

Edexcel S2 Jan 2006 Q7a

A teacher thinks that 20% of the pupils in a school read the Deano comic regularly.

He chooses 20 pupils at random and finds 9 of them read the Deano.

(*a*) (i) Test, at the 5% level of significance, whether or not there is evidence that the percentage of pupils that read the Deano is different from 20%. State your hypotheses clearly.

(ii) State all the possible numbers of pupils that read the Deano from a sample of size 20 that will make the test in part (*a*)(i) significant at the 5% level. **(9)**

Ex 7D

Exam Questions

5. (a) The discrete random variable $X \sim B(40, 0.27)$

Find $P(X \geqslant 16)$

(2)

Past records suggest that 30% of customers who buy baked beans from a large supermarket buy them in single tins. A new manager suspects that there has been a change in the proportion of customers who buy baked beans in single tins. A random sample of 20 customers who had bought baked beans was taken.

(b) Write down the hypotheses that should be used to test the manager's suspicion.

(1)

(c) Using a 10% level of significance, find the critical region for a two-tailed test to answer the manager's suspicion. You should state the probability of rejection in each tail, which should be less than 0.05

(3)

(d) Find the actual significance level of a test based on your critical region from part (c).

(1)

(e) Comment on the manager's suspicion in the light of this observation.

(1)

Later it was discovered that the local scout group visited the supermarket that afternoon to buy food for their camping trip.

(f) Comment on the validity of the model used to obtain the answer to part (e), giving a reason for your answer.

(1)

2. ~~The discrete random variable $X \sim B(30, 0.28)$~~

~~(a) Find $P(5 \leqslant X < 12)$~~

(2)

Past records from a large supermarket show that 25% of people who buy eggs, buy organic eggs. On one particular day, a random sample of 40 people is taken from those that had bought eggs and 16 people are found to have bought organic eggs.

(b) Test, at the 1% significance level, whether or not the proportion, p, of people who bought organic eggs that day had increased. State your hypotheses clearly.

(5)

(c) State the conclusion you would have reached if a 5% significance level had been used for this test.

(1)

Brad planted 25 seeds in his greenhouse. He has read in a gardening book that the probability of one of these seeds germinating is 0.25. Ten of Brad's seeds germinated. He claimed that the gardening book had underestimated this probability.

Test, at the 5% level of significance, Brad's claim. State your hypotheses clearly.

(Total 6 marks)

The proportion of houses in Radville which are unable to receive digital radio is 25%. In a survey of a random sample of 30 houses taken from Radville, the number, X, of houses which are unable to receive digital radio is recorded.

(a) Find $P(5 \leqslant X < 11)$

(3)

A radio company claims that a new transmitter set up in Radville will reduce the proportion of houses which are unable to receive digital radio. After the new transmitter has been set up, a random sample of 15 houses is taken, of which 1 house is unable to receive digital radio.

(b) Test, at the 10% level of significance, the radio company's claim. State your hypotheses clearly.

(5)

3. Naasir is playing a game with two friends. The game is designed to be a game of chance so that the probability of Naasir winning each game is $\frac{1}{3}$

Naasir and his friends play the game 15 times.

(a) Find the probability that Naasir wins

(i) exactly 2 games,

(ii) more than 5 games.

(3)

Naasir claims he has a method to help him win more than $\frac{1}{3}$ of the games. To test this claim, the three of them played the game again 32 times and Naasir won 16 of these games.

(b) Stating your hypotheses clearly, test Naasir's claim at the 5% level of significance.

(4)