4 A bag contains 6 red balls and 9 green balls. A ball is chosen at random from that bag, its colour noted and the ball placed to one side. A second ball is chosen at random and its colour noted.

a Draw a tree diagram to illustrate this situation. **(2 marks)**

b Find the probability that:

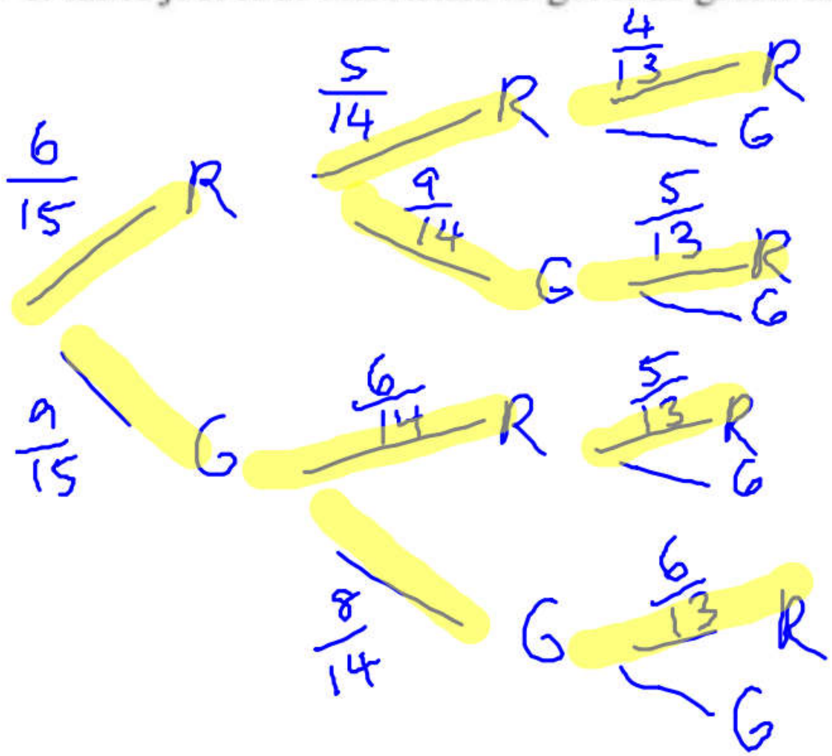
i both balls are green **(1 mark)**

ii the balls are different colours. **(2 marks)**

Further balls are drawn from the bag and not replaced. Find the probability that:

c the third ball is red **(2 marks)**

d it takes just four selections to get four green balls. **(2 marks)**



6 The colours of the paws of 75 kittens are recorded. 26 kittens have all black paws and 14 kittens have all white paws. 15 have a combination of black and white paws. One kitten is chosen at random. Find the probability that the kitten has:

a neither white nor black paws **(1 mark)**

b a combination of black and white paws given that they have some black paws. **(2 marks)**

Two kittens are now chosen. Find the probability that:

c both kittens have all black paws **(2 marks)**

d both kittens have some white paws. **(2 marks)**

a) $\frac{20}{75}$ B W B + W N

Cumulative Probabilities

Often we wish to find the probability of a range of values.
For a Binomial distribution, this was relatively easy if the range was narrow, e.g. $P(X \leq 1) = P(X = 0) + P(X = 1)$, but would be much more computationally expensive if we wanted say $P(X \leq 6)$.

If $X \sim B(10, 0.3)$, find $P(X \leq 6)$. $= 0.9894$

How to calculate on your ClassWiz:

Press Menu then 'Distributions'.
Choose "Binomial CD" (the C stands for 'Cumulative').
Choose 'Variable'.

$$x = 6$$
$$N = 10$$
$$p = 0.3$$

Pressing = gives the desired value.

The random variable $X \sim B(20, 0.4)$. Find:

$P(X \leq 7) =$ 0.4159

Look up $n = 20, p = 0.4, x = 7$

$P(X < 6) =$ $P(X \leq 5) = 0.1256$

Note that the table requires $\leq$

$P(X \geq 15) =$ $1 - P(X \leq 14) = 0.0016$

13 14 | 15 16 17 ...

To get this right, just say in your head "*What's the opposite of 'at least 15'?*". Hopefully you can see it's '*at most 14*'.

$P(5 < X \leq 8) =$ $P(X \leq 8) - P(X \leq 5)$

6 7 8 $= 0.5956 - 0.1256 = 0.4700$

Quickfire Questions

Write the following in terms of cumulative probabilities, e.g. $P(X < 7) = P(X \leq 6)$

$P(X < 5) =$ $P(X \leq 4)$

$P(X \geq 7) =$ $1 - P(X \leq 6)$

7 8 9 10

$P(X > 7) =$ $1 - P(X \leq 7)$

8 9 10

$P(10 \leq X < 20) =$ $P(X \leq 19) - P(X \leq 9)$

10 11 12 ... 19

$P(10 \leq X \leq 20) =$ $P(X \leq 20) - P(X \leq 9)$

$P(X = 100) =$ $P(X \leq 100) - P(X \leq 99)$

100

$P(20 < X < 30) =$ $P(X \leq 29) - P(X \leq 20)$

21 22 ... 29

"at least 30" = $P(X \geq 30) = 1 - P(X \leq 29)$

30 31 32

"greater than 30" = $P(X > 30) = 1 - P(X \leq 30)$

31 32

In a computer game, you have 20 attempts to try and knock some monkeys off a tree branch.
The probability of knocking a monkey off a tree branch is 0.3.
Determine the probability that someone:

a) Knocks less than 6 monkeys off the branch.
b) Knocks at least 9 monkeys off the branch.

The game gives you a prize (one banana) if the you knock at least 9 monkeys off the branch.
A student plays the game 5 times.
c) Calculate the probability that they win at least 4 bananas.

Let X be the number of monkeys knocked off the tree

$X \sim B(20, 0.3)$

a) $P(X<6) = P(X \leq 5) = 0.4164$

b) $P(X \geq 9) = 1 - P(X \leq 8) = 1 - 0.8867$

$= 0.1133$

c) Let Y be the number of games I win a banana in

$Y \sim B(5, 0.1133)$

$P(Y \geq 4) = 1 - P(Y \leq 3)$

$= 1 - 0.99925075$

$= \underline{\underline{0.00075}} = \underline{0.075\%}$

Ex 6C
Q1-6
Mixed Ex 6 Q15

Dealing with Probability Ranges

A spinner is designed so that probability it lands on red is 0.3. Jane has 12 spins.

a) Find the probability that Jane obtains at least 5 reds.

Jane decides to use this spinner for a class competition. She wants the probability of winning a prize to be < 0.05. Each member of the class will have 12 spins and the number of reds will be recorded.

b) Find how many reds are needed to win the prize.

a)

Let X be the number of times it lands on red.

$X \sim B(12, 0.3)$

$P(X \geq 5) = 1 - P(X \leq 4)$

$= 1 - 0.7237$

$= \underline{0.2763}$.

b) The number of reds needed to win is r

$P(X \geq r) < 0.05$

$r \quad r+1 \quad r+2$

$1 - P(X \leq r-1) < 0.05$

$0.95 < P(X \leq r-1)$

$0.95 < P(X \leq 6)$

$r - 1 = 6$

$\underline{r = 7}$

At Morpeth University, students have 20 exams at the end of the year. All students pass each individual exam with probability 0.45. Students are only allowed to continue into the next year if they pass some minimum of exams out of the 20.
What do the university administrators set this minimum number such that the probability of continuing to next year is at least 90%?

Let X be the number of exams passed by a student

$X \sim B(20, 0.45)$ k is the number of passes needed to continue.

$P(X \geq k) \geqslant 0.9$

$1 - P(X \leq k-1) \geqslant 0.9$

$0.1 \geqslant P(X \leq k-1)$

$0.1 \geqslant P(X \leq 5)$

$k - 1 = 5$

$k = 6$

The minimum number of exams to pass is 6.

Ex 6C
Q7-10

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.

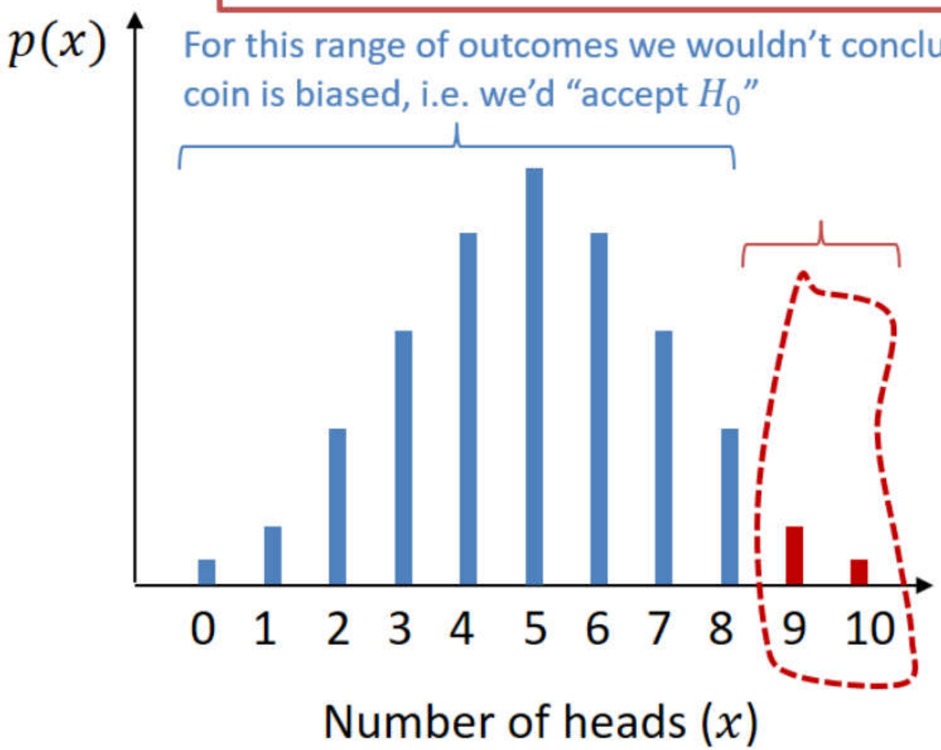
✎ The **alternative hypothesis**, H_1, is that there has been some change in the population parameter.

In this context...

We're asking "is the coin biased". This is making a statement about 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.



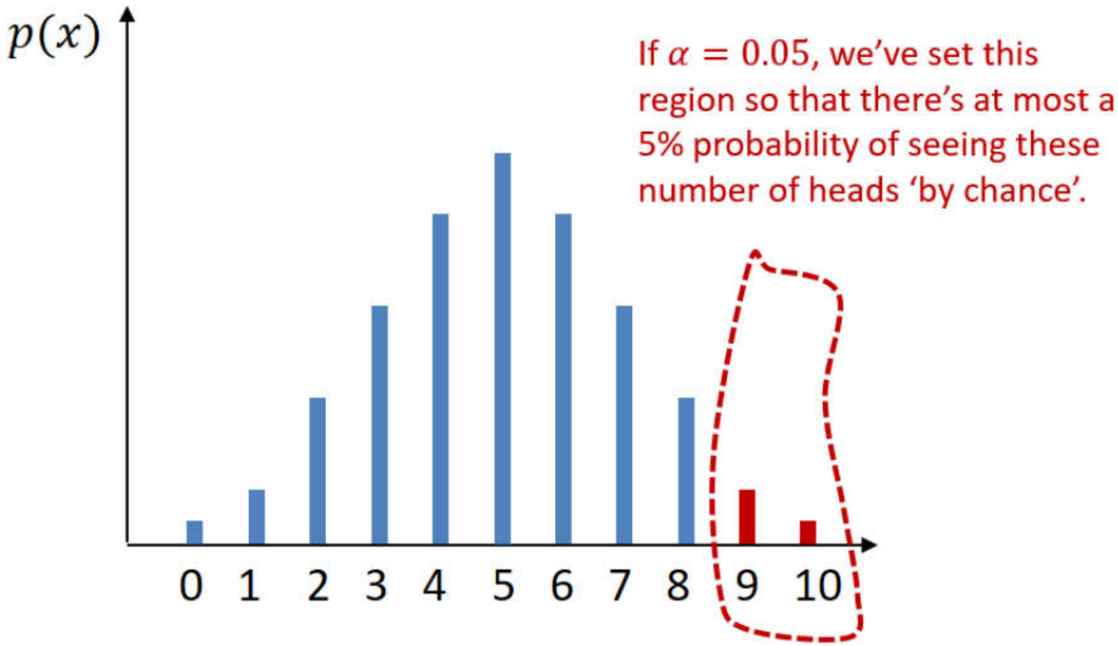
For this range of outcomes we'd conclude that this number of heads was too unlikely to happen by chance, and hence reject H_0 (i.e. that coin was fair) and accept H_1 (i.e. that coin was biased).

In this context...

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

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).



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

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.

The number of people who say they support the candidate.

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.**

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

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).

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

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.

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.

a) The number of people arriving late

$P(X \geq 4) = 1 - P(X \leq 3)$
$= 1 - 0.8619$
$= 0.1381$

b) $H_0: p = 0.05$

$H_1: p > 0.05$

$X \sim B(40, 0.05)$

c) $P(X \geq 6) < 0.1$

$0.0139 < 0.1$

So, there is enough evidence to reject the null hyp.

$P(X \geq 6) = 1 - P(X \leq 5)$
$= 1 - 0.9861$
$= 0.0139$

1.39%