

## P6-4 Binomial Distributions

- binomial probabilities

[Summary](#) [Learn](#) [Solve](#) [Revise](#) [Answers](#)

---

---

### Summary

---

---

A binomial distribution arises:

- where we do a number of trials,  $n$ ,
- where the outcomes can be classified as either a success or a failure,
- where the random variable is the number of successes from the  $n$  trials, and
- where the probability of success,  $p$ , on every trial is the same, irrespective of the outcomes of the other trials.

$$P(X = x) = {}^n C_x \times p^x \times q^{n-x}, \text{ where } q = 1 - p.$$

The probabilities can also be found more directly using a graphics calculator and the binomial probability distribution option.

The probability of  $x$  or fewer successes can be found using the binomial cumulative distribution option on the calculator. From this the probabilities of more than  $x$  etc. can be calculated.

The number of trials needed to have a certain probability of at least one success can be found by using the formula for  $P(X = x)$  above, but assuming that the number of successes is 0 and taking the complement.

The mean number of successes in a binomial experiment is  $np$ ; the variance is  $npq$  and the standard deviation is  $\sqrt{npq}$ .

The shape of a binomial distribution is a bell shape if  $n$  is large and  $p$  is close to  $\frac{1}{2}$ . It tends to be skewed if  $n$  is smaller and  $p$  is closer to 0 or 1.

---

---

### Learn

---

---

#### Binomial Distributions

In Module P6-3 (Discrete Random Variables) we looked at the probability distribution for the discrete random variable '*the number of heads obtained when tossing 8 coins*'.

The distribution looked like this:

Number of Hs	0	1	2	3	4	5	6	7	8
Probability	0.004	0.031	0.109	0.219	0.273	0.219	0.109	0.031	0.004

and, as a graph, like this:



We noted that this was an example of a binomial probability distribution.

A binomial distribution arises:

- where we do a number of trials,  $n$  (in the coin example  $n = 8$ ),
- where the outcomes can be classified as either a success or a failure (in this case getting a head is a success, getting a tail is a failure),
- where the random variable is the number of successes from the  $n$  trials, and
- where the probability of success,  $p$ , on every trial is the same, irrespective of the outcomes of the other trials (in the coin example  $p = \frac{1}{2}$ ).

We can say that this distribution is binomial in a number of ways:

- the probability distribution for *the number of heads from tossing 8 coins* is binomial;
- the number of heads from tossing 8 coins is a binomial probability distribution;
- if  $X$  is the random variable '*the number of heads from tossing 8 coins*', then  $X$  has a binomial distribution;

We can also say that the random variable is binomial

- the random variable '*the number of heads from tossing 8 coin*' is binomial;
- if  $X$  is the random variable '*the number of heads from tossing 8 coins*', then  $X$  is binomial;
- if  $X$  is the random variable '*the number of heads from tossing 8 coins*', then  $X \sim B$ .

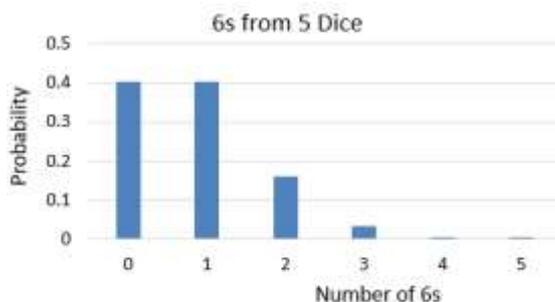
You don't have to memorise all these: they are all pretty obvious when heard in context.

Another example of a binomial random variable is the number of sixes obtained when 5 dice are rolled. Here  $n = 5$ ; a 6 is a success, any other number is a failure, and  $p = 1/6$ .

The probability distribution looks like this:

Number of 6s	0	1	2	3	4	5
Probability	0.40188	0.40188	0.16075	0.03215	0.00322	0.00013

and, as a graph, like this:



## Calculating Binomial Probabilities

How do we calculate the probabilities for a binomial random variable? Let's look at the dice situation as an example.

We roll 5 dice. Let's work out the probability of getting say 3 sixes?



We can work this out using a combination of the addition and multiplication rules.

Let 6 be the event 'getting a 6'; let N be the event 'not getting a 6'.

We need the probability of getting 666NN, 66N6N, 66NN6, 6N66N, 6N6N6, 6NN66, N666N, N66N6, N6N66 or NN666

Using the multiplication principle, each of these outcomes has a probability of  $(1/6)^3 \times (5/6)^2$ .

The number of outcomes giving 3 sixes is the number of ways of choosing 3 things from 5 (choosing 3 of the 5 rolls to be sixes). This is  ${}^5C_3$ .

So the probability we want is  ${}^5C_3 \times (1/6)^3 \times (5/6)^2$ , which we can enter into the calculator

to get approximately 0.03215.

The general formula for a binomial probability is

$$P(X = x) = {}^n C_x \times p^x \times q^{n-x},$$

where  $X$  is the binomial random variable,  $n$  is the number of trials,  $x$  is the number of successes,  $p$  is the probability of success and  $q$  is the probability of failure ( $= 1 - p$ ).

## Practice

- Q1 For each of the following questions, say whether or not it is a binomial situation. If not, explain why not. If it is,
- state what constitutes a success and what constitutes a failure
  - find the number of trials,  $n$ , and the number of successes required,  $x$
  - find  $p$  and  $q$
  - calculate the probability.
- The probability that, if we roll 6 dice, we will get exactly 2 sixes
  - The probability that, if we roll 10 dice, we will get exactly 3 ones
  - The probability that, if we toss a coin 20 times, we will get exactly 7 heads
  - The probability that, if we deal a hand of 7 cards from a standard pack of 52, the hand will include 5 spades
  - The probability that you would get 2 heads and a tail if you tossed three coins
  - The probability that, if you tossed a 50c coin a 20c coin, a 10c coin and a 5c coin, you would get a head on the 20c coin and tails on the others
  - The probability that you would get a total of 21 if you rolled 8 dice
  - Assuming that workers at the Happy Bee Plantation are on strike 5% of the time, the probability that, of the 30 workers, 5 will be on strike next Tuesday
  - The probability that a couple who have 8 children get 6 boys and 2 girls
  - The probability that, if 7 packs of cards are shuffled, 3 of them will have an ace as the top card
  - The probability that, if one pack of cards is shuffled, the top 7 cards will contain 1 ace
  - The probability that Grandma will eat 7 kg of beans before she dies
  - The probability that, if 6 dice are rolled, all of them will come up 5 or 6
  - Assuming that 18% of the pens coming off a conveyor belt are faulty, the

probability that, of the next 4, 3 are faulty

- (o) Assuming that the first frost always occurs in either May, June or July and that it is equally likely to occur in each of those three months, the probability that it will occur in May in 2 of the next 6 years.

Your calculator allows you to perform binomial probability calculations more directly. With a Casio, go into STAT mode (MENU, 2), then choose DIST (distribution), BINM (binomial), Bpd (binomial probability distribution).

To find the probability of getting 3 sixes from 5 dice rolls, set the options as follows:

Data	:Variable
x	:3
Numtrial	:5
p	:1÷6

Then press EXE to perform the calculation. This should give you  $p = 0.0321502$ .

## Practice

- Q2 For the questions in Q1 that were binomial, repeat the calculation using this calculator method. Check that you get the same answers.

### ***Bernoulli Random Variables***

You may need to know what a Bernoulli random variable is. If you don't, then you don't need to read this box.

A Bernoulli random variable is just a binomial random variable with  $n = 1$ , i.e. it is the number of successes from a single trial. For tossing a bent coin with  $P(H) = 0.6$ , the Bernoulli probability distribution would look like this:

Number of Hs	0	1
Probability	0.4	0.6

Probability calculations in Bernoulli experiments are trivial.

The single trial is called a Bernoulli trial. A binomial experiment, like tossing 8 bent coins can be considered as a series of Bernoulli trials, which can be called a Bernoulli sequence. But only do this if you want to confuse yourself.

## Binomial Probabilities - Cumulative

Suppose we tossed 20 coins and wanted to know the probability of getting 12 or more heads. We could do the probability calculations for 12 heads, 13 heads, 14 heads etc. all the way up to 20, then add them. But the calculator provides a quicker way.

Instead of choosing Bpd, choose Bcd (binomial cumulative distribution). This works the same as Bpd except that, instead of calculating the probability of  $x$  successes, it calculates the probability of  $x$  or fewer successes. In other words, the probability of anything up to and including  $x$  successes.

If we set  $x$  to 11, it will calculate the probability of 0 to 11 successes. This comes to 0.748. The probability of 12 or more is then the complement of this,  $1 - 0.748$ , i.e. 0.252.

If we want the probability of between 8 and 13 heads (inclusive), we would find the probability of 13 or fewer, then subtract the probability of 7 or fewer.

Be careful to distinguish between 'up to', 'fewer than', 'less than or equal to', 'at least', 'no less than', 'more than', 'no more than' and similar wording in the question. Think carefully about what set of numbers you need for  $x$ .

### Practice

- Q3 If we toss a coin 30 times, find the probability that we will get:
- (a) 18 or fewer heads
  - (b) more than 18 heads
  - (c) fewer than 18 heads
  - (d) at least 18 heads
  - (e) 18 or more heads
  - (f) not less than 18 heads
  - (g) no more than 18 heads
  - (h) no fewer than 14 heads
- Q4 If a couple have 6 children, find the probability that:
- (a) no more than 4 will be girls
  - (b) more than 4 will be girls
  - (c) at least 5 will be girls
  - (d) less than 2 will be boys
  - (e) no more than 1 will be a boy
  - (f) all will be boys
- Q5 If you roll a die 10 times, find the probability that you will get
- (a) 4 or fewer sixes
  - (b) 4 sixes
  - (c) fewer than 4 sixes
  - (d) 4 or more sixes

- (e) more than 4 sixes
- (f) no more than 4 sixes
- (g) only one 6

Q6 If you toss a coin 25 times, find the probability of getting

- (a) between 10 and 15 (inclusive) heads
- (b) between 12 and 20 (inclusive) heads
- (c) more than 12 heads, but fewer than 18
- (d) at least 8 heads, but no more than 13
- (e) more heads than tails
- (f) between 13 and 17 (exclusive) heads
- (g) fewer than 10 tails, but more than 5

## Calculating the Number of Trials Necessary

Clare has a 25% chance of getting the ball in the basket from 10 metres. How many shots would she need to take in order to be 90% certain of getting at least one in?

We could solve this by a long process of guess and check, calculating for different numbers of shots. But there is a shorter way. The complement of 'getting at least one in' is getting none in. So we can say that the probability of getting zero in is less than or equal to 10%.

The probability of 0 successes from  $n$  trials is  ${}^n C_0 \times 0.25^0 \times 0.75^n$ .  ${}^n C_0$  is the number of ways of choosing 0 successes from  $n$  trials and this will always be 1.  $0.25^0$  is of course 1, so we get

$$1 \times 1 \times 0.75^n = 0.1 \quad \text{Solving this using logs, we get } n = 8.004$$

Now, theoretically 8.004 shots will give her a 90% chance of getting at least one in. But she can't have 8.004 shots. 8 won't be quite enough, so she has to take 9 shots.

## Practice

Q7 Solve these using the complement method above:

- (a) Joe has a 15% chance of getting a basket from 12 m. How many shots would he need to take to give himself a 75% chance of getting at least one in?
- (b) How many coins would I need to toss to have a 99% probability of getting at least one tail?
- (c) How many dice would I need to roll to have a 75% probability of getting at least one six?
- (d) How many times would I have to roll 6 dice to have an 80% probability

of getting 3 or more sixes at least once?

- (e) I need to roll 11 on two dice to get out of jail. How many times would I need to roll to have a 95% chance of getting out?
- (f) A toy car has 16 parts. Each part has a 2% chance of being defective. If any part is defective, the car won't work. How many cars would I need to buy to have a 99% chance of getting one that works?

## Mean and Standard Deviation

Suppose we toss a coin 100 times. The expected number of heads would be 50. In other words, the mean number of heads if we repeated the experiment lots of times would be 50.

Fairly intuitively, we get this by multiplying the number of trials by the probability of success.

For a binomial experiment, the mean number of successes =  $np$ . In other words the mean value of a binomial random variable is  $np$ .

But we wouldn't always get 50 heads: there would be some variation either side of that expected or mean value. This variation can be represented by the variance or standard deviation. If we repeated the experiment lots of times and recorded the number of heads, the variance of the results would be given by the expression  $npq$ .

So for a binomial experiment, the variance of the number of successes =  $npq$ .

The standard deviation is just the square root of the variance, so is given by  $\sqrt{npq}$ .

## Practice

Q8 Solve these:

- (a) If you tossed 60 fair coins numerous times, what would be the mean and standard deviation of the number of heads obtained?
- (b) A multiple-choice quiz has 20 questions each with choices a, b, c and d. If you answered the questions with random answers, what would be your expected score? If you did this lots of times, what would be the variance of your scores?
- (c) If you roll a die 30 times, how many 4s would you expect? What would be the standard deviation?
- (d) 3% of Kanga scratch lotto tickets pay \$500 and the rest pay nothing. If you bought 12 tickets, what would be the expected winnings. What would be the standard deviation?

## Parameters and Shapes

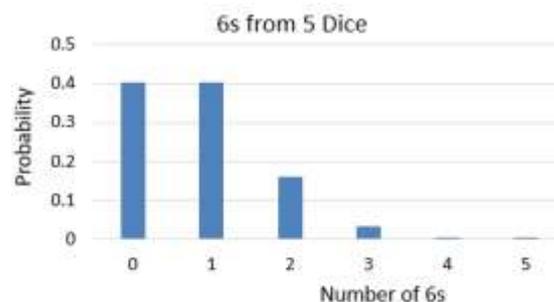
Exponential functions have a general form  $A = Pe^{rt}$ . They all have a similar shape, but the shape varies in detail depending on the values of the parameters  $P$  and  $r$ .

In the same way, binomial distributions all have a similar shape, but they vary in detail depending on the values  $n$  (the number of trials) and  $p$  (the probability of success on a given trial).

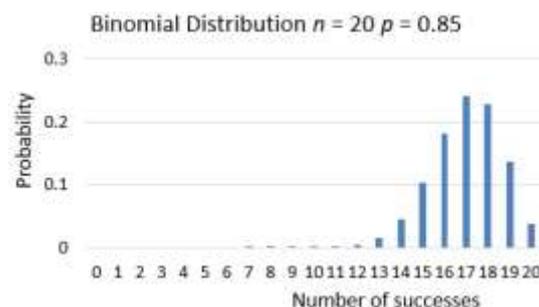
With the number of heads from tossing 8 coins,  $n = 8$ ,  $p = \frac{1}{2}$ , and the shape was a symmetrical bell-shaped curve –



With the number of sixes from rolling 5 dice,  $n = 5$ ,  $p = \frac{1}{6}$ , and the shape was centred to the left and positively skewed –



If  $p$  were close to 1, say 0.85, the distribution would be centred to the right and negatively skewed like this one with  $n = 20$  –



In general, the larger  $n$ , the closer the distribution is to a symmetrical bell curve; the further  $p$  is from  $\frac{1}{2}$ , the more skewed the distribution is.

The information in the box below is probably not needed for your studies in maths, but is an interesting application and may be of interest if you are also studying chemistry or physics.

## Entropy

The concept of entropy is met in chemistry and in physics, but is traditionally a very difficult one to get an understanding of. It is sometimes described as dispersal – of energy or particles. The second law of thermodynamics says that all natural processes lead to an increase in entropy, i.e. greater dispersal of energy or particles.

Combinations can be used to get a bit of a feel for the idea.

Consider a system consisting of two identical jars connected by a tube. 100 nitrogen molecules are put into the left jar contains, 100 oxygen molecules are put into the right jar.

Molecules can pass through the tube and so move from one jar to the other. In time, the random motion of the molecules will cause the nitrogen molecules to disperse to be more evenly distributed between the two jars; likewise with the oxygen molecules. This process is the increase in entropy of the system.

But the molecules don't know that they have to move to increase the entropy, nor do they know where they should go even if they did. It all just happens. Their movement is random.

Given enough time, the probability than any nitrogen molecule will be in the left jar will be 0.5. The probability that there will be 50 nitrogen molecules in the left jar will be  ${}^{100}C_{50} 0.5^{50} 0.5^{50}$ , which is 0.08. For other numbers, the probabilities will be as follows:

No of N <sub>2</sub> molecules in left jar	Probability	Probability
0	${}^{100}C_0 0.5^0 0.5^{100}$	$8 \times 10^{-31}$
20	${}^{100}C_{20} 0.5^{20} 0.5^{80}$	$4 \times 10^{-10}$
40	${}^{100}C_{40} 0.5^{40} 0.5^{60}$	0.01
50	${}^{100}C_{50} 0.5^{50} 0.5^{50}$	0.08
60	${}^{100}C_{60} 0.5^{60} 0.5^{40}$	0.01
80	${}^{100}C_{80} 0.5^{80} 0.5^{20}$	$4 \times 10^{-10}$
100	${}^{100}C_{100} 0.5^{100} 0.5^0$	$8 \times 10^{-31}$

As you can see, the probability that the nitrogen molecules aren't reasonably dispersed between the two jars is very small. Thus they will generally become dispersed.

If we repeated the calculation with more molecules (say 1000 of each), the probabilities become very much smaller still. Actual jars of gas would contain billions of billions of molecules and the probability that more than 51% of the nitrogen molecules would be in the left jar would be so small that we could wait billions of years and never see it happen.

Even though the dispersal is a probabilistic process, the probabilities are so extreme as to constitute practical certainties. Thus we say that the molecules *will* disperse and entropy *will* increase.

Sometimes, entropy is thought of as disorder. This works as long as we consider having all the nitrogen molecules in the same jar as being in order and having them dispersed between the two jars as being disordered.

## Other Discrete Random Variables

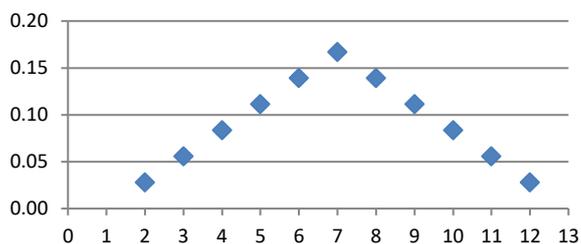
So far we have talked about just two types discrete random variable – *uniform* random variables with uniform probability distributions, for example the score obtained from rolling a die, and *binomial* random variables with binomial distributions, for example the number of sixes obtained when you roll 5 dice.

You might not need to know about any others at this stage, but being aware of them can help put what you know into perspective.

Another discrete random variable which is neither uniform nor binomial is the sum obtained from rolling two dice. This looks like this:

Score	2	3	4	5	6	7	8	9	10	11	12
Probability	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

and, as a graph, like this:



Others, more widely used in statistics include:

- Geometric random variables: the number of trials up to and including the first success in a binomial situation with probability of success  $p$ . One parameter:  $p$ .
- Negative binomial random variable: the number of trials up to and including the  $r^{\text{th}}$  success in a binomial situation with probability of success  $p$ . Two parameters:  $r$  and  $p$ .
- Poisson random variable: the number of events in time  $t$  when the events happen at random (e.g. radioactive atoms decaying or people arriving at a supermarket check-out) with an average rate,  $\lambda$ . Two parameters:  $t$  and  $\lambda$ .

---

---

## Solve

---

---

Q51 A multiple-choice test has 6 questions with 5 options and 10 questions with 4 options. If you answer at random, what is the probability that you will get just 3 correct in total?

Q52 Rita sits a multiple choice test with 20 questions, each with 5 choices. She

knows the answers to 8 of the questions and chooses random answers for the others. Sadie sits the same test and, on each question, has a 40% chance of being correct. What is the probability that both girls get more than 10 correct?

---

---

## Revise

---

---

### Revision Set 1

- Q61 What are the conditions that must be met for a binomial calculation to be used?
- Q62 You roll a die 12 times. You want to know the probability of getting a number greater than 4 exactly 5 times. Can a binomial calculation be used? If not, explain why not. If so, write an expression for the probability then evaluate the expression to find the probability.
- Q63 You are dealt 5 cards from a standard pack. You want to know the probability that you will have at least 3 aces. Can a binomial calculation be used? If not, explain why not? If so, find the probability.
- Q64 4% of people in Harlech have red hair. Find the probability that, of the 124 students at Harlech Primary School, 7 will be redheads. What is the probability that at least 10 of the students will be redheads?
- Q65 97% of Euripides beetles die before they are 5 years old. How many would I have to have to be 95% certain of having at least one alive when they turn 5?
- Q66 In a multiple choice test, there are 5 choices for each of the 20 questions. If chimpanzees sit the test and each gives random answers to all questions, find the mean and standard deviation of their scores.
- Q67 Without calculating it, describe the shape of the binomial distribution with  $n = 50$  and  $p = 0.4$ .

### Revision Set 2

- Q71 What are the conditions that must be met for a binomial calculation to be used?
- Q72 You toss a coin 100 times.
- Write an expression for the probability of getting 55 heads, then evaluate it to get the probability.
  - Use the distribution function on your calculator to check your answer to the last question.
  - Find the probability of getting more than 52 heads, but less than 60.
  - Find the mean number of heads and the variance.
  - How many times would you need to conduct the experiment to have an 80% chance of getting 60 or more heads at least once?
- Q73 Without calculating it, describe the shape of the binomial distribution with  $n = 200$  and  $p = 0.9$ .

## Answers

- Q1 (a) Success – a six; failure – any other number;  $n = 6$ ;  $x = 2$ ;  $p = 1/6$ ;  $q = 5/6$ ; Probability = 0.201.  
 (b) Success – a one; failure – anything else;  $n = 10$ ;  $x = 3$ ;  $p = 1/6$ ;  $q = 5/6$ ; Probability = 0.155.  
 (c) Success – a head; failure – a tail;  $n = 20$ ;  $x = 7$ ;  $p = 0.5$ ;  $q = 0.5$ ; Probability = 0.074.  
 (d) Not binomial because the probability of a spade for the second card depends on whether the first was a spade.  
 (e) Success – a head; failure – a tail;  $n = 3$ ;  $x = 2$ ;  $p = 0.5$ ;  $q = 0.5$ ; Probability = 0.375.  
 (f) Not binomial because we are not counting successes.  
 (g) Not binomial because we are not counting successes.  
 (h) Probably not binomial because the probability of a worker being on strike is generally not independent of whether the other workers are on strike.  
 (i) Success – a boy; failure – a girl; (or the other way round);  $n = 8$ ;  $x = 6$ ;  $p = 0.5$ ;  $q = 0.5$ ; Probability = 0.109.  
 (j) Success – ace on top; failure – anything else on top;  $n = 7$ ;  $x = 3$ ;  $p = 1/13$ ;  $q = 12/13$ ; Probability = 0.0116.  
 (k) Not binomial because the probability that one of the cards is an ace depends on if the others are aces.  
 (l) Not binomial – none of the conditions met.  
 (m) Success – 5 or 6; failure – 1, 2, 3 or 4;  $n = 6$ ;  $x = 6$ ;  $p = 1/3$ ;  $q = 2/3$ ; Probability = 0.00137.  
 (n) Success – being faulty; failure – not being faulty;  $n = 4$ ;  $x = 3$ ;  $p = 0.18$ ;  $q = 0.82$ ; Probability = 0.019.  
 (o) Success – occurring in May – not occurring in May;  $n = 6$ ;  $x = 2$ ;  $p = 1/3$ ;  $q = 2/3$ ; Probability = 0.329.
- Q3 (a) 0.900 (b) 0.100 (c) 0.819 (d) 0.191  
 (e) 0.191 (f) 0.191 (g) 0.100 (h) 0.708
- Q4 (a) 0.891 (b) 0.109 (c) 0.109 (d) 0.109 (e) 0.109 (f) 0.016
- Q5 (a) 0.985 (b) 0.054 (c) 0.930 (d) 0.070  
 (e) 0.015 (f) 0.985 (g) 0.323
- Q6 (a) 0.770 (b) 0.655 (c) 0.478 (d) 0.633  
 (e) 0.5 (f) 0.291 (g) 0.113
- Q7 (a) 9 (b) 7 (c) 8 (d) 26 (e) 53 (f) 4
- Q8 (a) 30, 3.87 (b) 5, 3.75 (c) 5, 2.04 (d) \$180, \$106
- Q51 22.7%  
 Q52 5.63%
- Q61 1. the situation must consist of a number of trials  
 2. the outcome of each trial can be regarded as a success or a failure  
 3. the probability of success in each trial is the same, irrespective of what has happened in previous trials  
 4. we wish to calculate the number of successes
- Q62 Yes.  ${}^{12}C_5 \times (1/3)^5 \times (2/3)^7 = 0.191$
- Q63 No, because the cards are not replaced, the probability of success is not the same for each trial.
- Q64 10.4%, 2.76%
- Q65 99
- Q66 Mean = 4, sd = 1.79
- Q67 A bell curve centred just left of centre and slightly positively skewed
- Q71 1. the situation must consist of a number of trials  
 2. the outcome of each trial can be regarded as a success or a failure  
 3. the probability of success in each trial is the same, irrespective of what has happened in previous trials

4. we wish to calculate the number of successes

Q72 (a)  ${}^{100}C_{55} \times (1/2)^{55} \times (1/2)^{45} = 0.048$  (c) 0.280 (d) 50, 25 (e) 56

Q73 A bell curve centred near the right hand end and with a distinct negative skew