



MAS1403

Quantitative Methods for Business Management

Semester 1, 2017–2018

Module leader: Dr. Lee Fawcett

Additional lecturers: Dr. Dave Walshaw and Dr. Ged Cowburn

Announcements

- This week is a **computer practical week**
- That means the standard Thursday tutorials are replaced with computer sessions:

Group	Day	Time
A	Tues	10am
B	Wed	10am
C	Thurs	11am

- All sessions take place in the **Herschel PC cluster**
- You will be introduced to some software for statistical analysis – important for the first written assignment

Announcements

- The **semester 1 written assignment** will be available to view from the course webpage later this week
 - It is worth 10% of this module, so you should treat it like a mini project
 - You will have four full weeks to complete the assignment – the deadline for submission is **4pm, Thursday 14th December 2017**
 - Some questions will require you to use the software you will be introduced to in this week's computer sessions
 - For some questions you will be allocated your own personal dataset
 - Some questions will be open-ended, so you will have to think carefully about how to tackle the problems
 - Time in tutorials will be given for support
- **CBA2** is now live in assessed mode – deadline: **23:59 this coming Friday, 17th November**

Lecture 7

DISCRETE PROBABILITY MODELS

7.1 Probability distributions

The **probability distribution** of a discrete random variable X is the list of all possible values X can take and the probabilities associated with them.

For example, if the random variable X is the outcome of a roll of a die then the probability distribution for X is:

r	1	2	3	4	5	6	Sum
$P(X = r)$	1/6	1/6	1/6	1/6	1/6	1/6	1

7.1 Probability models

In the die-rolling example, we used the interpretation of probability to obtain the probability distribution for X , the outcome of a roll on the die.

7.1 Probability models

In the die-rolling example, we used the **classical** interpretation of probability to obtain the probability distribution for X , the outcome of a roll on the die.

Consider the following **frequentist** example.

Let X be the number of cars observed in half-hour periods passing the junction of two roads. In a five hour period, the following observations on X were made:

2 3 2 5 5 3 4 5 6 7

Obtain the probability distribution of X .

7.1 Probability models

2 3 2 5 5 3 4 5 6 7

We can calculate the following probabilities:

$$P(X = 0) = \frac{0}{10} = 0$$

$$P(X = 1) = \frac{0}{10} = 0$$

$$P(X = 2) = \frac{2}{10} = 0.2$$

$$P(X = 3) = \frac{2}{10} = 0.2$$

7.1 Probability models

2 3 2 5 5 3 4 5 6 7

$$P(X = 4) = \frac{1}{10} = 0.1$$

$$P(X = 5) = \frac{3}{10} = 0.3$$

$$P(X = 6) = \frac{1}{10} = 0.1$$

$$P(X = 7) = \frac{1}{10} = 0.1$$

7.1 Probability models

Thus would give:

x	$P(X = x)$
< 2	0
2	0.2
3	0.2
4	0.1
5	0.3
6	0.1
7	0.1
> 7	0
sum	1

Does this make sense?

7.2 The binomial distribution

In many surveys and experiments data is collected in the form of **counts**. For example,

- the number of people in a survey who bought a CD
- the number of people who said they would vote Labour
- the number of defective items in a sample

All these variables have common features:

- 1 Each person/item has only **two** possible (exclusive) responses (Yes/No, Defective/Not defective etc)
 - this is referred to as a **trial** which results in a **success** or **failure**
- 2 The survey/experiment takes the form of a **random sample**
 - the responses are independent

7.2 The binomial distribution

If:

- There are a fixed number of trials or experiments (n)
- There are only two possible outcomes for each trial ('success' or 'failure')
- There is a constant probability of 'success', p
- The outcome of each trial is independent of any other trial

Then we say that the number of successes, X , follows a **binomial distribution**.

Example 2

Which of the following scenarios could be adequately modelled by a binomial distribution?

- The number of sixes on 3 rolls of a fair six-sided die.
- The number of students who pass MAS1403 this year.

Example 2

Which of the following scenarios could be adequately modelled by a binomial distribution?

- **The number of sixes on 3 rolls of a fair six-sided die.**
- ~~The number of students who pass MAS1403 this year.~~

7.2 The binomial distribution

Suppose we are interested in the **number of sixes** we get from 3 rolls of a die.

Each roll of the die is an experiment or trial which gives a “**six**” (success, or **s**) or “**not a six**” (failure, or **f**).

The probability of a success is $p = P(\text{six}) = 1/6$.

We have $n = 3$ **independent** experiments or trials (rolls of the die).

7.2 The binomial distribution

Let X be the number of sixes obtained.

We can now obtain the full **probability distribution** of X ; a probability distribution is a list of all the possible outcomes for X with along with their associated probabilities.

7.2 The binomial distribution

For example, suppose we want to work out the probability of obtaining **three** sixes: (three “**successes**” — i.e. **sss** — or $P(X = 3)$).

Since the rolls of the die can be considered independent, we get (using the **multiplication law**):

$$P(\text{sss}) = P(s) \times P(s) \times P(s) = \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} = \left(\frac{1}{6}\right)^3$$

7.2 The binomial distribution

That one's easy!

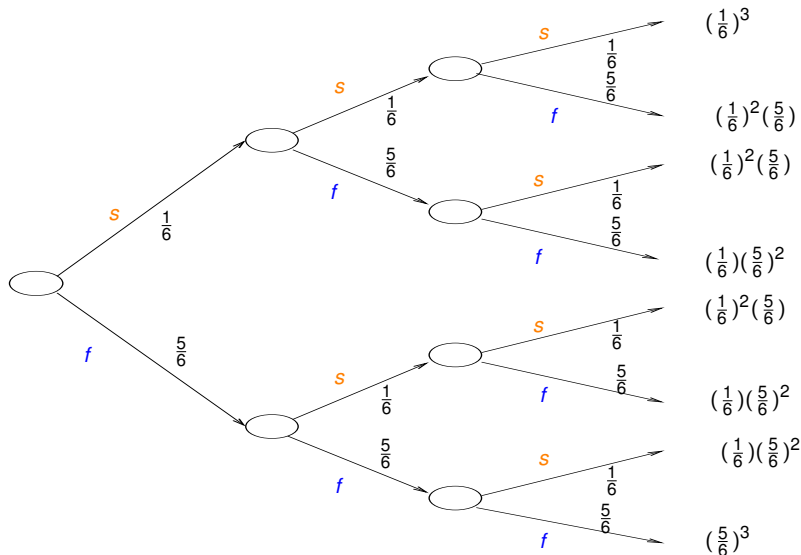
What about the probability that we get **two** sixes — i.e.
 $P(X = 2)$?

This one's a bit more tricky, because that means we need two
s's and one **f**...

...but the **f** (“not six”) could appear on the first roll, or the
second roll, or the third!

Thinking about it, there are actually **eight** possible outcomes
for the three rolls of the die:

7.2 The binomial distribution



7.2 The binomial distribution

So, for $P(X = 2)$, we *could* have:

$$P(\textcolor{blue}{f}\textcolor{brown}{ss}) = \frac{5}{6} \times \frac{1}{6} \times \frac{1}{6} = \left(\frac{1}{6}\right)^2 \times \frac{5}{6},$$

or we *could* have:

$$P(\textcolor{brown}{s}\textcolor{blue}{fs}) = \frac{1}{6} \times \frac{5}{6} \times \frac{1}{6} = \left(\frac{1}{6}\right)^2 \times \frac{5}{6},$$

or even:

$$P(\textcolor{brown}{ss}\textcolor{blue}{f}) = \frac{1}{6} \times \frac{1}{6} \times \frac{5}{6} = \left(\frac{1}{6}\right)^2 \times \frac{5}{6},$$

7.2 The binomial distribution

Can you see that we therefore get:

$$P(X = 2) = 3 \times \left(\frac{1}{6}\right)^2 \times \frac{5}{6}.$$

Which takes the form:

$$\begin{aligned} P(X = 2) &= \text{Number of ways to get two sixes} \\ &\quad \times P(2 \text{ sixes}) \times P(1 \text{ "not six"}). \end{aligned}$$

7.2 The binomial distribution

Using the same argument as above we can calculate the other probabilities:

$$P(X = 0) = \left(\frac{5}{6}\right)^3 = 0.579$$

$$P(X = 1) = 3 \times \left(\frac{1}{6}\right) \times \left(\frac{5}{6}\right)^2 = 0.347$$

$$P(X = 2) = 3 \times \left(\frac{1}{6}\right)^2 \times \frac{5}{6} = 0.069$$

$$P(X = 3) = \left(\frac{1}{6}\right)^3 = 0.005...$$

7.2 The binomial distribution

... and so the full **probability distribution** for X is:

x	0	1	2	3
$P(X = x)$	0.579	0.347	0.069	0.005

This probability distribution shows that most of the time we would get either 0 or 1 sixes and, for example, 3 sixes would be quite rare.

Try your own experiment!

Minitab

7.2 The binomial distribution

Now this is a bit long-winded . . . and that was just for **three** rolls of the die!

Imagine what it would be like to calculate for **100** rolls of the die!

We would like a more concise way of working these probabilities out without having to list all the possible outcomes as we did above.

7.2.1 Calculating probabilities

You should see from the tree diagram that we can construct a **general formula**, taking the form:

$$P(X = r) = \# \text{ ways to get } r \text{ successes out of } n \text{ trials} \\ \times P(r \text{ successes}) \times P(n - r \text{ failures})$$

We can write this more succinctly as

$$P(X = r) = {}^nC_r \times p^r \times (1 - p)^{n-r}, \quad r = 0, 1, \dots, n.$$

The **binomial coefficient** ${}^nC_r$ works out how many ways we can choose r objects out of n , and so is commonly read as “ n choose r ”: **button on the calculator!**

Example 3

What is the probability of getting 2 sixes from three rolls of a fair six-sided die?

We can just use our table of derived results from earlier... but let's use the binomial formula directly!

We have X : Number of sixes on three rolls of the die, and $X \sim \text{Bin}(3, 1/6)$. Thus

$$P(X = r) = {}^nC_r \times p^r \times (1 - p)^{n-r}$$

$$\begin{aligned} P(X = 2) &= {}^3C_2 \times 1/6^2 \times (1 - 1/6)^{3-2} \\ &= 3 \times \frac{1}{36} \times \frac{5}{6} \\ &= \frac{5}{72} = 0.069. \end{aligned}$$

Example 4

If $X \sim \text{Bin}(10, 0.2)$ calculate:

(a) $P(X = 2)$

(b) $P(X \leq 2)$

(c) $P(X < 3)$

(d) $P(X > 1)$

Example 4(a): $P(X = 2)$

$$P(X = 2) = {}^{10}C_2 \times 0.2^2 \times 0.8^8 = 0.302.$$

Example 4(b): $P(X \leq 2)$

For $P(X \leq 2)$, we need to add the answers to $P(X = 0)$, $P(X = 1)$ and $P(X = 2)$.

$$P(X = 0) = {}^{10}C_0 \times 0.2^0 \times 0.8^{10} = 0.107$$

$$P(X = 1) = {}^{10}C_1 \times 0.2^1 \times 0.8^9 = 0.268$$

$$P(X = 2) = 0.302 \quad \text{from part (a)}$$

So

$$P(X \leq 2) = 0.107 + 0.268 + 0.302 = 0.677.$$

Example 4(c): $P(X < 3)$

The possible outcomes are:

0 1 2 3 4 5 6 7 8 9 10

Example 4(c): $P(X < 3)$

The possible outcomes are:

0 1 2 3 4 5 6 7 8 9 10

Therefore

$$\begin{aligned}P(X < 3) &= P(X = 0) + P(X = 1) + P(X = 2) \\&= P(X \leq 2) \\&= 0.677.\end{aligned}$$

Example 4(d): $P(X > 1)$

The possible outcomes are:

0 1 2 3 4 5 6 7 8 9 10

Example 4(d): $P(X > 1)$

The possible outcomes are:

0 1 2 3 4 5 6 7 8 9 10

Therefore

$$P(X = 0) = {}^{10}C_0 \times 0.2^0 \times 0.8^{10} = 0.107$$

$$P(X = 1) = {}^{10}C_1 \times 0.2^1 \times 0.8^9 = 0.268$$

Therefore

$$P(X > 1) = 1 - (0.107 + 0.268) = 0.625.$$

7.2.2 Mean and variance

If $X \sim \text{Bin}(n, p)$, then its mean (or “expected value”) and variance are

$$\begin{aligned} E[X] &= n \times p & \text{and} \\ \text{Var}(X) &= n \times p \times (1 - p). \end{aligned}$$

Example 5

If $X \sim \text{Bin}(10, 0.2)$ calculate:

- (a) $E[X]$
- (b) $\text{Var}(X)$
- (c) $\text{SD}(X)$

$$E[X] = 10 \times 0.2 = 2$$

$$\text{Var}(X) = 10 \times 0.2 \times 0.8 = 1.6$$

$$\text{SD}(X) = \sqrt{1.6} = 1.265.$$

Example 6

A salesperson has a 50% chance of making a sale on a customer visit and she arranges 6 visits in a day.

- (a) Assuming sales at each visit are independent, suggest an appropriate distribution for the number of sales she makes in a day.
- (b) Calculate her expected number of sales.

Example 6

(a) X : Number of sales per day; $X \sim \text{Bin}(6, 0.5)$.

(b)

$$E[X] = 6 \times 0.5 = 3 \text{ sales.}$$