Statistics for Computing MA4413

Lecture 7

Bernoulli Trials and the Binomial Distribution

Kevin Burke

kevin.burke@ul.ie

Bernoulli Trials

Many experiments only have *two* possible outcomes, i.e., where some event either happens or it does not happen:

- $\{$Windows user, Non-Windows user$\}$.
- $\{$In favour, Not in favour$\}$ (of a government policy for example).
- $\{$Head, Tail$\}$ - flipping a coin.
- $\{$Die shows a six, Die does not show a six$\}$.
- $\{$Component is defective, Component is non-defective$\}$.
- $\{$Individual on time for work, Individual not on time$\}$.

Such an experiment is called a **Bernoulli trial**.

Bernoulli Trials

Clearly these variables are *categorical* but we can code them using a *binary random variable* X where

- $X = 1$ means the event has occurred.
- $X = 0$ means the event has *not* occurred.

Thus,

- $\Pr(X = 1)$: probability that event occurs.
- $\Pr(X = 0) = 1 - \Pr(X = 1)$: probability that event does *not* occur.

Bernoulli Distribution

For simplicity we let p represent the probability that the event occurs and, hence, $1 - p$ is the probability that it does not.

The *probability distribution* is:

x	1	0
$\Pr(X = x)$	p	$1 - p$

This is known as the **Bernoulli distribution**.

Bernoulli Distribution

Note that the *probability function* can be written as

$$\Pr(X = x) = p^x\,(1-p)^{1-x}.$$

We can check that this works:

$$\Pr(X=1) \;=\; p^1\,(1-p)^{1-1} \;=\; p^1\,(1-p)^0 \;=\; p. \qquad \checkmark$$

$$\Pr(X=0) \;=\; p^0\,(1-p)^{1-0} \;=\; p^0\,(1-p)^1 \;=\; 1-p. \qquad \checkmark$$

Bernoulli Distribution

We can calculate:

$$E(X) = 1 \times p + 0 \times (1-p) = p.$$

$$E(X^2) = 1^2 \times p + 0^2 \times (1-p) = p.$$

$$Var(X) = E(X^2) - [E(X)]^2 = p - p^2 = p(1-p).$$

$$Sd(X) = \sqrt{Var(X)} = \sqrt{p(1-p)}.$$

Bernoulli Distribution

The Bernoulli distribution is summarised via the following key formulae:

$$\boxed{X \sim \text{Bernoulli}(p)}$$

(this means "the random variable X has a Bernoulli distribution with parameter p")

$$\boxed{\Pr(X = x) = p^x\,(1-p)^{1-x}}$$

where $\boxed{x \in \{0,1\}}$

$$\boxed{E(X) = p}$$

$$\boxed{Var(X) = p\,(1-p)}$$

Example: Defective Components

Consider the experiment of inspecting resistors produced in a factory.

Let's assume that the *true* proportion of defective units is 1%, i.e., the probability of defect is $p = 0.01$.

Selecting a resistor randomly from the line for inspection leads to a *Bernoulli trial* with $\Pr(X = 1) = 0.01$ and $\Pr(X = 0) = 0.99$.

$\Rightarrow \Pr(X = x) = 0.01^x \, 0.99^{1-x}$.

The *expected value* is $E(X) = p = 0.01 \Rightarrow$ if this experiment is repeated a large number of times, we expect that 1% of units tested would be defective.

Proportions: Hypothesis Testing

Rarely do we know the *true* proportion p.

However, we may *hypothesise* something about its value, e.g., we might assume that $p = 0.4$.

We can *estimate* p using a sample (as discussed in Lecture1) which gives us $\hat{p}$.

If $\hat{p}$ is close to 0.4, then we could conclude that the true value of p is as we *hypothesised*.

Using Bernoulli distribution theory (just covered) and the *central limit theorem* (still to come) we can test this formally.

Independent Bernoulli Trials

Consider the experiment of

- carrying out n Bernoulli trials

  where

- the probability of the event occurring, p, is the same in each trial

  and

- the result of each trial is independent of the other trials.

We can calculate the probability of a particular number of events occurring using the **Binomial distribution**, e.g., 5 events in 20 trials, more than 3 events in 7 trials, no events in 100 trials etc.

Example: Biased Coin

Consider flipping a *biased* coin where the Pr("coin shows head") $= 0.1$ and let $X = 1$ represent a head showing.

Clearly this is a Bernoulli trial with $p = 0.1$.

If we flipped the coin 4 times, we might enquire about the probability of getting the sequence $HTTT = 1000$.

By independence of the trials, we can multiply probabilities:

$$\Pr(1000) = p(1)\,p(0)\,p(0)\,p(0) = 0.1 \times 0.9 \times 0.9 \times 0.9 = (0.1^1)\,(0.9^3).$$

Example: Biased Coin

What if we didn't specify the order? We wish to know the probability of obtaining one head.

In this case there are *four* possibilities $\{1000, 0100, 0010, 0001\}$.

$$
\begin{aligned}
\Rightarrow \text{Pr(“one head”)} & \\
&= \text{Pr}(1000) + \text{Pr}(0100) + \text{Pr}(0010) + \text{Pr}(0001) \\
&= 0.1(0.9)(0.9)(0.9) + 0.9(0.1)(0.9)(0.9) + \\
&\qquad 0.9(0.9)(0.1)(0.9) + 0.9(0.9)(0.9)(0.1) \\
&= (0.1^1)(0.9^3) + (0.1^1)(0.9^3) + (0.1^1)(0.9^3) + (0.1^1)(0.9^3) \\
&= 4 \times (0.1^1)(0.9^3) = 0.2916.
\end{aligned}
$$

Example: Biased Coin

Similarly, if we wish to work out the probability of two heads, there are *six* possibilities $\{1100, 1010, 1001, 0110, 0101, 0011\}$.

$$\Rightarrow \text{Pr(“two heads”)} = 6 \times (0.1^2)\,(0.9^2) = 0.0486.$$

Clearly it can be quite tedious to list various outcomes. Recall that using the *choose operator* makes things easier (Lecture5).

In the case of 2 heads above, we have 4 available positions and wish to place a “1” in 2 of these positions, i.e., we must *choose* 2 positions from 4 $\Rightarrow \binom{4}{2} = 6$ possibilities (the 0s go in the remaining positions).

Example: Biased Coin

Letting X = "the number of heads", we have

$$\Pr(X = 0) = \binom{4}{0} \times (0.1^0)\,(0.9^4) = 1\,(1)(0.9^4) = 0.6561.$$

$$\Pr(X = 1) = \binom{4}{1} \times (0.1^1)\,(0.9^3) = 4\,(0.1^1)(0.9^3) = 0.2916.$$

$$\Pr(X = 2) = \binom{4}{2} \times (0.1^2)\,(0.9^2) = 6\,(0.1^2)(0.9^2) = 0.0486.$$

$$\Pr(X = 3) = \binom{4}{3} \times (0.1^3)\,(0.9^1) = 4\,(0.1^3)(0.9^1) = 0.0036.$$

$$\Pr(X = 4) = \binom{4}{4} \times (0.1^4)\,(0.9^0) = 1\,(0.1^4)(1) = 0.0001.$$

This is the probability distribution for X. Note that $\sum p(x) = 1$.

Example: Biased Coin

The information on the previous slide can be summarised via the *probability function*:

$$p(x) = \Pr(X = x) = \binom{4}{x} 0.1^x \; 0.9^{4-x}.$$

(check: substitute different values of x into the above formula)

More generally, for n trials:

$$p(x) = \Pr(X = x) = \binom{n}{x} 0.1^x \; 0.9^{n-x}.$$

More generally still, for any value of p:

$$p(x) = \Pr(X = x) = \binom{n}{x} p^x \; (1-p)^{n-x}.$$

Binomial Distribution

The **Binomial distribution** is used for calculating the probability of x events in n independent Bernoulli trials:

$$\boxed{X \sim \text{Binomial}(n, p)}$$

$$\boxed{\Pr(X = x) = \binom{n}{x} p^x \, (1-p)^{n-x}}$$

where $\boxed{x \in \{0, 1, 2, \ldots, n\}}$

$$\boxed{E(X) = n\,p}$$

$$\boxed{Var(X) = n\,p\,(1-p)}$$

(the derivation of the $E(X)$ and $Var(X)$ formulae is beyond the scope of this course)

Example: Defective Resistors

Let's assume that 5% of all resistors manufactured by a particular company are defective. We purchase 20 resistors from this manufacturer.

Let $X =$ the number of faulty resistors received.

It is clear that $X \sim$ Binomial$(n = 20, p = 0.05)$.

$$\Rightarrow \Pr(X = x) = \binom{n}{x} p^x (1-p)^{n-x} = \binom{20}{x} (0.05^x)(0.95^{20-x}).$$

We simply plug in values for x into this formula to work out the probability of receiving that many defective resistors.

Example: Defective Resistors

What is the probability that we receive:

...No defective resistors?

$$\Pr(X = 0) = \binom{20}{0}(0.05^0)(0.95^{20-0}) = 1\,(1)\,(0.95^{20}) = 0.3585.$$

...At least one defective resistor?

$$\Pr(X \geq 1) = 1 - \Pr(X = 0) = 1 - 0.358 = 0.6415.$$

...Three defective resistors?

$$\Pr(X = 3) = \binom{20}{3}(0.05^3)(0.95^{20-3}) = 1140\,(0.05^3)\,(0.95^{17}) = 0.0596.$$

Example: Defective Resistors

What is the probability of receiving *more than one* defective resistor?

Since these are *discrete* values "more than one" means "two or more":

$$\Pr(X > 1) = \Pr(X \geq 2) = p(2) + p(3) + \ldots + p(19) + p(20).$$

We could sum all of the above probabilities but this is quite tedious. It is easier if we use the *complement rule*.

$$\begin{aligned}
\Rightarrow \Pr(X > 1) &= 1 - \Pr(X \leq 1) \\
&= 1 - [\, p(0) + p(1) \,] \\
&= 1 - \left[\binom{20}{0} (0.05^0)\,(0.95^{20}) + \binom{20}{1} (0.05^1)\,(0.95^{19}) \right] \\
&= 1 - (0.3585 + 0.3774) \;=\; 1 - 0.7359 \;=\; 0.2641.
\end{aligned}$$

Example: Defective Resistors

What is the probability of receiving *between two and four* defective resistors?

$$
\begin{aligned}
\Pr(2 \le X \le 4) &= p(2) + p(3) + p(4). \\
&= \binom{20}{2}(0.05^2)(0.95^{18}) + \binom{20}{3}(0.05^3)(0.95^{17}) + \\
&\qquad \binom{20}{4}(0.05^4)(0.95^{16}) \\
&= 0.1887 + 0.0596 + 0.0133 \\
&= 0.2616.
\end{aligned}
$$

Example: Defective Resistors

How many defective resistors will we receive on average?
(per shipment of 20 resistors)

$$E(X) = np = 20(0.05) = 1 \text{ resistor.}$$

What is the standard deviation?

$$Var(X) = np(1-p) = 20(0.05)(0.95) = 0.95 \text{ resistors}^2.$$

$$\Rightarrow Sd(X) = \sqrt{Var(X)} = \sqrt{0.95} = 0.97 \text{ resistors.}$$

Question 1

Let's assume that 10% of resistors produced by another company are defective. Assume again that we purchase 20 resistors. Let X represent the number of defective resistors received. What is the probability of receiving:

a) Two defective resistors.

b) No defective resistors.

c) Less than four defective resistors.

d) Two or more defective resistors.

e) How many defective resistors will we receive on average?

f) Calculate $Sd(X)$?

Binomial Tables

The **binomial tables** are very useful for calculating binomial probabilities quickly.

In particular, **"greater than or equal to"** probabilities are tabulated:

$$\boxed{\Pr(X \geq r)}$$

where r is the value in question.

We select the appropriate binomial distribution by finding p in the column headings and n in the row headings.

Limitation of Binomial Tables

The tables do not show *all* possible binomial distributions (obviously - since there are an infinite number of *n*-*p* combinations).

The tables can be used for binomial distributions with:

$n = \{2, 5, 10, 20, 50, 100\}$ *(see rows)*

and

$$p = \{0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, \\ 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50\} \quad \textit{(see columns)}$$

Example: Defective Resistors

Recall that in the defective resistors example
$X \sim \text{Binomial}(n = 20, p = 0.05)$.

This binomial distribution *does* appear in the tables.

$\Rightarrow$ We can find all of the probabilities again but now using the tables.

The key thing when using the tables is that we must rework the question in terms of **greater than or equal to** probabilities.

Example: Defective Resistors

...No defective resistors? We need $Pr(X = 0)$. Note that:

$$
\begin{array}{rl}
\Pr(X \geq 0) = & p(0)+p(1)+p(2)+\ldots+p(19)+p(20) \\
\Pr(X \geq 1) = & \quad\quad\;\; p(1)+p(2)+\ldots+p(19)+p(20) \\
\hline
\Pr(X \geq 0) - \Pr(X \geq 1) = & p(0)
\end{array}
$$

(think of this as using $\Pr(X \geq 1)$ to "chop off" all unwanted probabilities from $\Pr(X \geq 0)$ leaving $p(0)$ as required

$$\Rightarrow \Pr(X = 0) = \Pr(X \geq 0) - \Pr(X \geq 1) = 1.0000 - 0.6415 = 0.3585.$$

The probabilities $\Pr(X \geq 0) = 1.0000$ and $\Pr(X \geq 1) = 0.6415$ were found in column $p = 0.05$, row $n = 20$.

Example: Defective Resistors

...At least one defective resistor?

This is $\Pr(X \geq 1)$ which is already a greater than or equal to probability $\Rightarrow$ look it up directly:

$$\Pr(X \geq 1) = 0.6415.$$

...Three defective resistors?

From $\Pr(X \geq 3)$ we subtract $\Pr(X \geq 4)$ to chop off all but $\Pr(X = 3)$:

$$\Pr(X = 3) = \Pr(X \geq 3) - \Pr(X \geq 4) = 0.0755 - 0.0159 = 0.0596.$$

Example: Defective Resistors

What is the probability of receiving *more than one* defective resistors?

$$\Pr(X > 1) = \Pr(X \geq 2) = 0.2642.$$

What is the probability of receiving *between two and four* defective resistors?

$$\Pr(2 \leq X \leq 4) = \Pr(X \geq 2) - \Pr(X \geq 5) = 0.2642 - 0.0026 = 0.2616.$$

Check that the answers are the same as those found using the probability function.

Question 2

Assume that $X =$ the number of defective resistors where $X \sim \text{Binomial}(n = 20, p = 0.1)$.

Using the binomial tables, calculate the probability of:

a) Two defective resistors.

b) No defective resistors.

c) Less than four defective resistors.

d) Two or more defective resistors.

Note: you calculated these in Question 1 using the *formula* for the probability function.

R Code

R has various probability distributions built in. The function `dbinom(x,size,prob)` is $\Pr(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$ where `size` is n and `prob` is p.

For example:

`dbinom(0,size=20,prob=0.05)` gives `0.3584859`

and

`dbinom(3,size=20,prob=0.05)` gives `0.05958215`

Compare these with the values calculated previously on slide 18.

R Code

We evaluate a range of probabilities at once.

For example:

`dbinom(2:4,size=20,prob=0.05)`
gives `0.18867680 0.05958215 0.01332759`.

We can also `sum` these:

`sum(dbinom(2:4,size=20,prob=0.05))`
which gives `0.2615865`

Compare this with slide 20.

R Code

Greater than probabilities, i.e., $\Pr(X > x)$, can be calculated using the `pbinom` function.

It is important to note that this differs from the binomial tables which (as we saw) provide *greater than or equal to* probabilities.

For example:

`pbinom(0,size=20,prob=0.05,lower=F)`
gives `0.6415141` which is $\Pr(X > 0) = \Pr(X \geq 1)$.

`pbinom(2,size=20,prob=0.05,lower=F)`
gives `0.07548367` which is $\Pr(X > 2) = \Pr(X \geq 3)$.

`pbinom(3,size=20,prob=0.05,lower=F)`
gives `0.01590153` which is $\Pr(X > 3) = \Pr(X \geq 4)$.

Compare this with slide 27.

R Code

We can *generate* binomial random variables using `rbinom`.

For example:

```
rbinom(100,size=20,prob=0.05)
```

generates 100 binomial variables from the Binomial($n = 20, p = 0.05$) distribution.

This represents getting 100 shipments of 20 resistors and counting the number of defective resistors in the first shipment, second, third etc.

R Code

Since the Bernoulli distribution is a binomial with $n = 1$ we can generate Bernoulli variables (i.e., binary random variables) by setting `size = 1`.

For example:

```
rbinom(250,size=1,prob=0.1)
```

generates 250 binary variables where the probability of getting a 1 is 0.1.