15.2 The binomial distribution

The essential elements of a binomial distribution are:

- There is a fixed number of (n) trials
- Each trial has only two possible outcomes – a "success" or a "failure"
- The probability of a success (p) is constant from trail to trial.
- Trials are independent of each other.

The outcomes of a **binomial experiment** and the corresponding probabilities of these outcomes are called a binomial distribution

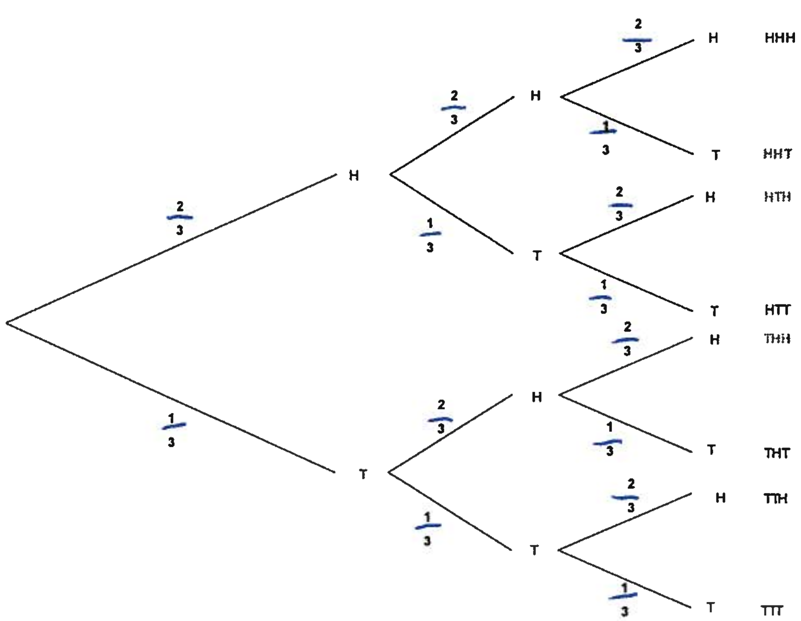
The **binomial distribution** describes the behavior of a discrete variable X if the conditions above apply.

The parameters that define a unique binomial distribution are the values of n (the number of trials) and p (the proability of a success).

Any binomial distribution is represented as $X \sim \mathrm{B}(n, p)$

Consider this problem, which you first met in Chapter 3: determine the probability of getting exactly two heads in three tosses of a biased coin for which $\mathrm{P}(\text{head}) = \frac{2}{3}$.

Let's first look at a tree diagram to help answer this question.



HHT $2/3 \cdot 2/3 \cdot 1/3 = 4/27$

HTH $2/3 \cdot 1/3 \cdot 2/3 = 4/27$

THH $1/3 \cdot 2/3 \cdot 2/3 = 4/27$

So $P(\text{2 heads in three tosses}) = P(HHT) + P(HTH) + P(THH)$

Each of the probabilities are the same.

$$P(HHT) = P(HTH) = P(THH) = \left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right) = \frac{4}{27}$$

Thus, $P(\text{2 heads in three tosses}) = 3\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right) = \frac{12}{27} = \frac{4}{9}$

<u>However, you should only use a tree diagram if the number of trials, *n*, is small.</u>

What if you were asked to find the probability of obtaining exactly two heads in six tosses of this coin? The tree diagram for this question would be too large, so we look for a formula.

> We often use a theoretical distribution, such as the binomial distribution, to describe a random variable that occurs in real life. This process is called *modeling* and enables us to carry out calculations. If the theoretical distribution matches the real-life variable perfectly, then the model is perfect. However, this is usually not the case. Generally the results of any calculations will not necessarily give a completely accurate description of the real-life situation. Does this make them any less useful?

Have the essential conditions for a binomial distribution been met?

• There is a fixed number (n) of trials.	6 trials
• Each trial has two possible outcomes – a "success" or a "failure"	Success = heads failure = tails
• The probability of a success (p) is constant from trial to trial.	Prob of success is 2/3 each time the coin is tossed
• Trials are independent of each other.	Getting a head one time will not affect getting a head next time.

What would be one combination of Hs and Ts that will produce 2 heads and 4 tails?

$B(6, 2/3)$
Binomial
# of trials
Prob of Success

H H T T T T

And the probability of that would be....

$\frac{2}{3} \cdot \frac{2}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3}$ or $\left(\frac{2}{3}\right)^2 \cdot \left(\frac{1}{3}\right)^4 = 4/729 = .00549$

And every possible combination of 2 Hs and 4 Ts will have the Same probability

> The most usual error when calculation a binomial probability is to forget that if there are exactly r successes, there must also be $n - r$ failures.

But how many combinations are there?

$\binom{n}{r}$ represents the number of ways choosing r items out of n items.

$\binom{n}{r}$ means the same as ${}_nC_r$

The number of combinations of 6 items that have 2 Hs and 4 Ts is therefore $\binom{6}{2}=\binom{6}{4}=15$

${}_6C_2$ ${}_6C_4$

You can use your GDC, or the formula $\frac{n!}{r!(n-r)!}$, or the 3rd entry in the 6th row of Pascal's triangle.

Therefore: $\binom{6}{2} = \frac{6!}{2!(4!)} = \frac{6\cdot5\cdot4!}{2\cdot1\cdot4!} = \frac{6\cdot5}{2\cdot1} = \frac{30}{2} = 15$

P(2 heads in six tosses) =

$\binom{6}{2}=\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)^4 = 15\left(\frac{4}{729}\right) = \frac{20}{243} = .0823$

Generalizing this method gives the binomial distribution function:

> If X is binomially distributed, $X \sim B(n, p)$, then the probability of obtaining r successes out of n independent trials, when p is the probability of success for each trial, is
>
> $$P(X=r)=\binom{n}{r}p^r(1-p)^{n-r}$$
>
> This is often shortend to
>
> $$P(X=r)=\binom{n}{r}p^r q^{n-r} \text{ where } q=1-p$$

Example: X is binomially distributed with 7 trials and a probability of success equal to $\frac{3}{8}$ each attempt. What is the probability of

$B(7, 3/8)$

$q = 5/8$

a. Exactly three successes

b. At least one success?

c. Four or fewer success?

(a) $P(x=3)=\binom{7}{3}\left(\frac{3}{8}\right)^3\left(\frac{5}{8}\right)^4 = 35\left(\frac{27}{512}\right)\left(\frac{625}{4096}\right) = .282$

(b) $P(x\geq1) = 1-P(x=0) = \binom{7}{0}\left(\frac{3}{8}\right)^0\left(\frac{5}{8}\right)^7 = .963$

(c) $P(x\leq4) = P(x=0)+P(x=1)+P(x=2)+P(x=3)+P(x=4) = .926$

on calc...

Go to DISTR

(a) binompdf (7, 3/8, 3) ~ exactly 3
or 7 trials, p=3/8, x=3

(c) binomcdf (7, 3/8, 4) ~ cumulative ≤ 4
or 7 trials, p=3/8, x=4

Exercise 15C

Example: The probability that Allison gets to school on time on any day is 0.7. What is the probability that in a regular school week of five days Allsion will make it to school on time exactly four times? Is this a binomial distribution? Why or why not? If the situation is a binomial distribution does the problem meet the conditions?

Binomial? – yes, only 2 outcomes

Fixed # of days (5)

Success = on time

failure = late

Trials are independent

$B(5, 0.7)$

$P(x=4) = \binom{5}{4}(.7)^4(.3)^1$

or

binompdf (5, .7, 4)

0.360

(3 s.f.)

Example: Jim the Amazing Hypnotist is known to be able to hypnotize 75% of his subjects. Jim meets with two groups of 12 subjects.
What is the probability that all 12 patients were hypnotized in both groups?

$P(x=12) = \binom{12}{12}(0.75)^{12}(0.25)^0 = (.75)^{12} = 0.0316$

or binompdf (12, 0.75, 12)

The two groups are independent so to find the probability of both having all 12 hypnotized would be $P(x=12) \cdot P(x=12)$

or

$(P(x=12))^2 = 0.0010033$

0.00100

(3 s.f.)

Exercise 15D

Example: A college lecture hall contains a large number of students where one quarter of the students are male. The rest are females, duh. The professor picks a random number of students from the hall to participate in a study. How many students must be picked so that the probability that there is at least one female is among them is greater than 0.90?

Let F be the random variable "The # of females" $F \sim B(n, 0.75)$

$P(F \geq 1) = 1 - P(F = 0)$

$1 - \binom{n}{0}(0.75)^0(0.25)^n$

$P(F \geq 1) = 1 - (0.25)^n$

we want an n so that $p > 0.9$

$1 - (0.25)^n > 0.9$

$-(0.25)^n > -0.1$

$(0.25)^n < 0.1$

$\log 0.25^n < \log 0.1$

$n \log 0.25 < \log 0.1$

$n > \frac{\log 0.1}{\log 0.25}$ ← negative #

$n > 1.66$ ~ At least 2 students

Exercise 15E

Expectation of a binomial distribution

Think of the example of the biased coin where $P(H) = \frac{2}{3}$.

If you toss the coin 3 times how many times would you expect to get a head?

2

Intuituley the answer is 2.

This is the same as calculationing $3 \times \frac{2}{3} = 2$

For the binomial distribution where $X \sim B(n, p)$, the expection of X, $E(X) = np$

Example: A biased dice is thrown 40 times and the number of fours is 12. The dice is thrown 15 more times. Find the expected number of fours for these 15 throws.

The expected value of X, $E(X)$, is also called the mean, μ, $E(X) = \mu$.

$P(x=4) = \frac{12}{40} = \frac{3}{10}$

$E(x) = 15(3/10) = 45/10 = 4.5$ fours

Exercise 15F

Possible Exploration pg. 535

Variance of a binomial distribution

Chapter 8 introduced the concept of the variance of a set of data, as a measure of dispersion.

The formula for the variance of the binomial distribution is given in the formula booklet.

If $X \sim \mathrm{B}(n, p)$ then $\mathrm{Var}(X) = npq$, where $q = 1 - p$

Thinking about the original example of the biased coin where, if you toss the coin 3 times you expect to get a head 2 times. However, obviously this will not happen every time. If you repeat the experiment many times you will sometimes get 0, 1, and 3 heads.

You can find the standard deviation, σ, by taking the square root of the variance.

Using the formula for variance,

$$\mathrm{Var}(X) = npq = 3 \times \frac{2}{3} \times \frac{1}{3} = \frac{2}{3}$$

Example: In a large company, 30% of the ~~works~~ workers travel to work on public transport. A random sample of 18 workers is selected. Find the expected number of workers in this sample that travel to work on public transport and the standard deviation.

Let X be the # of workers

$X \sim B(18, 0.3)$

$E(x) = 18(0.3) = 5.4$

$VAR(x) = 18(0.3)(0.7) = 3.78$

$SD = \sqrt{3.78} = 1.944$

Exercise 15G