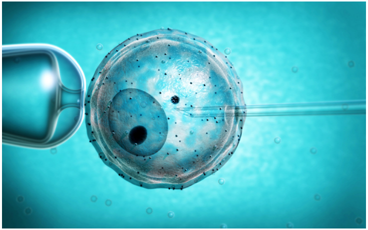
Topic9: Test for Proportion (Proportion and Sign Tests)

Outline

Example1: IVF

IVF (In Vitro Fertilisation) is a technique of assisted reproductive technology. The first IVF baby was born in Australia in 1980 under the supervision of a team of doctors at Monash University, the 3rd IVF baby in the world. The first IVF baby was born in the UK (Louise Brown, 25 July 1978, physiologist and Nobel Prize Winner Sir Robert Edwards, surgeon Dr Patrick Steptoe).

Interestingly, of the first 7 IVF births in Australia, 6 were girls. **Is there evidence of a gender bias in IVF?**



Example2: Vegetarianism in Australia

According to data from 2013, 10% of the Australian population choose a vegetarian diet. In 2015, a survey of 10,000 people resulted in 1,190 vegetarians. **Has vegetarianism increased in Australia?**

▸ Roy Morgan Stats

An Intuitive approach to the Proportion Test

Select a coin from your wallet. **Is the coin biased?**

$\boxed{\text{Hypothesis}}$ H_0: Coin is fair, $p = P(Head) = 0.5$.
$\boxed{\text{Experiment}}$ Toss the coin 20 times and see how many heads come up.
$\boxed{\text{Observed Sample}}$ This results in x heads.
$\boxed{\text{Conclusion}}$ If the number of heads x is very small or very big, then there is evidence against H_0.

Assuming H_0 is true, then
X = Number of heads $\sim Bin(n = 20, p = 0.5)$.

x	$P(X$ is x or more extreme$)$	Conclusion
0	$P(X = 0) = 0.00000095$	H_0 is highly unlikely.
1	$P(X \leq 1) = 0.00002$	H_0 is highly unlikely.
2	$P(X \leq 2) = 0.0002$	H_0 is highly unlikely.
5	$P(X \leq 5) = 0.02$	H_0 is fairly unlikely.
8	$P(X \leq 8) = 0.25$	Data consistent with H_0.

```r
x=c(0,1,2,5,8)
pbinom(x,20,0.5)
```

```
## [1] 9.536743e-07 2.002716e-05 2.012253e-04 2.069473e-02 2.517223e-01
```

Steps for Proportion Test

For a hypothesis concerning a unknown population proportion p, we follow the following steps.

$\boxed{\text{H}}$ $H_0 : p = p_0$ vs $H_1 : p < p_0$.

$\boxed{\text{A}}$ The n trials are independent with constant probability p.

$\boxed{\text{T}}$

- $\tau = X =$ number of successes $\sim Bin(n, p_0)$ under H_0.
- Small values of x will argue against H_0 for H_1.
- The observed value is x.

$\boxed{\text{P}}$ P-value $= P(X \leq x)$.

$\boxed{\text{C}}$ Weigh up the P-value. A rule of thumb is to reject H_0 for P-value $< 0.05 = \alpha$.

Notes on P-value:

- If the alternate hypotheses is $H_1 : p > p_0$, then large values of x will argue against H_0 for H_1 and the associated P-value is $P(X \geq x)$.
- If the alternate hypotheses is 2 sided $H_1 : p \neq p_0$, then both small and large values of x will argue against H_0 for H_1 and the associated P-value is $P(|X - np_0| \geq |x - np_0|)$.
- In the special case when $p_0 = \frac{1}{2}$, the 2 sided P-value reduces to
  - P-value $= 2P(X \geq x)$, for $x > \frac{n}{2}$.
  - P-value $= 2P(X \leq x)$, for $x < \frac{n}{2}$.
- If n is large, then we can use the CLT (Topic 7) to find an approximation to the P-value. We find the approximating Normal $Y \sim N(np_0, np_0(1 - p_0))$, and then either use R or standardise the Normal and look up the Normal tables.

Examples

IVF

Of the first 7 IVF births in Australia, 6 were girls. Is there evidence of a gender bias in IVF?

Let $p = P(\text{girl baby from IVF})$. Note this is a 2 sided test as the possible direction of bias is unspecified, given this was the first IVF trials.

$\boxed{\text{H}}$ Without prior evidence, we would assume no gender bias, hence $p_0 = 0.5$. If there is a gender bias, then $p_0 \neq 0.5$. Hence, $H_0 : p = 0.5$ vs $H_1 : p \neq 0.5$.

$\boxed{\text{A}}$ The $n = 7$ births are independent with constant probability p.

$\boxed{\text{T}}$

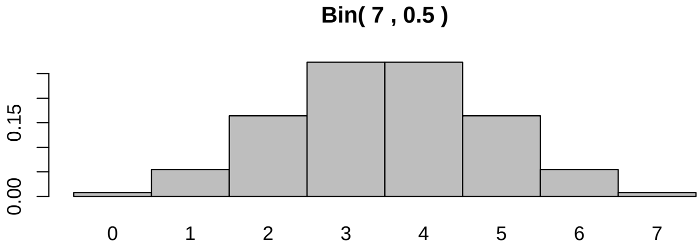
- $\tau = X =$ number of girl births from 1st 7 IVF births $\sim Bin(7, 0.5)$ under H_0.
- Small and large values of x will argue against H_0 for H_1. (ie if H_0 is not true, then we would expect x to indicate either a bias towards girls (large) or bias towards boys (small)).
- The observed value is $x = 6$ (large).

$\boxed{\text{P}}$ P-value = 2 $P(X \geq 6)$ = 0.125.

```
dbinom(6,7,0.5) + dbinom(7,7,0.5)

## [1] 0.0625

1-pbinom(5,7,0.5)

## [1] 0.0625
```

$\boxed{\text{C}}$ As the P-value is 12.5%, we would say that the data are consistent with H_0. Hence, it appears that there is not a gender bias in IVF.

Note: What the P-value indicates, is that getting $x = 6$ or greater ($x = 7$) on a Binomial(7,0.5) distribution is not that uncommon, which is clear from the Binomial probability distribution function.



Vegetarianism

Has vegetarianism increased in Australia?

Let $p = P(\text{Vegetarian preference in Australia})$. Note this is a 1 sided test as we are testing whether vegetarianism has 'increased'.

$\boxed{\text{H}}$ From the 2013 data, 10% of the Australian population choose a vegetarian diet, hence $p_0 = 0.1$.
$H_0 : p = 0.1$ vs $H_1 : p > 0.1$.

$\boxed{\text{A}}$ The $n = 10000$ people in the 2015 survey are independent with constant probability p.

$\boxed{\text{T}}$

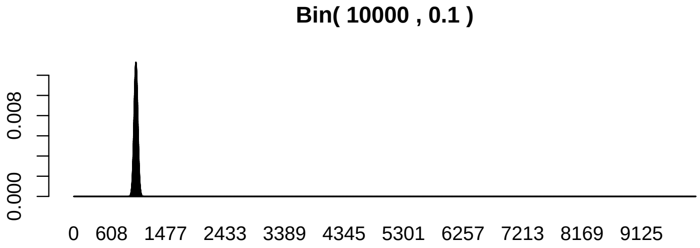
- $\tau = X =$ Number of vegetarians in the survey $\sim Bin(10000, 0.1)$ under H_0.
- Large values of x will argue against H_0 for H_1.
  (ie if H_0 is not true, and H_1 is, then we would expect x to be larger, as $p > 0.1$.).
- The observed value is $x = 1190$.

$\boxed{\text{P}}$ P-value $= P(X \geq 1190) \approx 0$.

```
1-pbinom(1189,10000,0.1)

## [1] 3.711557e-10
```

$\boxed{\text{C}}$ As the P-value is so small, we would say that the data are not consistent with H_0. Hence, it appears that there has been an increase in vegetarianism in Australia.

Note: What the P-value indicates, is that getting $x = 1190$ or greater on a Binomial(10000,0.1) distribution is rare, which is clear from the Binomial probability distribution function.



Sign Test

The Sign Test is a clever trick which effectively widens the applicability of the Proportion Test. It allows the 'Proportion' Test to be used for testing for a hypothesis about a mean or median.

We change from $H_0 : \mu = \mu_0$ (or $H_0 : \tilde{\mu} = \tilde{\mu}_0$) to $H_0 : p_+ = 0.5$, by considering the proportion of the signs of differences $\{sign(x_i - \mu_0)\}$ which are positive.

Note: If any observations are equal to the null hypothesis value $x_i = \mu_0$ then we eliminate them from the sample. This assumes there are only a few 'zeroes', as it effectively reduces the sample size.

Context

- Suppose a single sample $x_1, x_2, \ldots, x_n$ is taken from a continuous distribution of unknown type.
- We want to test $H_0 : \mu = \mu_0$ (or $H_0 : \tilde{\mu} = \tilde{\mu}_0$).
- If we assume that the distribution is symmetric, then if H_0 holds, every observation is equally likely to be above or below μ_0.
- Consider the set of signs of differences

$$sign(x_1 - \mu_0), sign(x_2 - \mu_0), \ldots, sign(x_n - \mu_0)$$

- Define $X =$ the number of $+$ signs $\sim Bin(n, p_+)$, where $p_+ = P(+$ difference$)$.
- Then $H_0 : \mu = \mu_0$ is equivalent to $H_0 : p_+ = 0.5$.

Steps for Sign Test

$\boxed{\text{Preparation}}$ For $H_0 : \mu = \mu_0$, calculate the signs of differences $\{sign(x_i - \mu_0)\}$ and count the number of positive signs x.

$\boxed{\text{H}}$ $H_0 : p_+ = 0.5$ vs $H_1 : p_+ < 0.5$.

$\boxed{\text{A}}$ The distribution is continuous and symmetric.

Note: For $H_0 : \tilde{\mu} = \tilde{\mu}_0$, we only need to assume that the distribution is continuous.

$\boxed{\text{T}}$

- $\tau = X =$ number of positive signs $\sim Bin(n, p_+)$ (under H_0)
- Small values of x will argue against H_0 for H_1.
- The observed value is x.

$\boxed{\text{P}}$ P-value $= P(X \leq x)$.

$\boxed{\text{C}}$ Weigh up the size of P-value.

Example: Freeze Dried Coffee

Freeze drying and spray drying are 2 different methods of producing instant coffee. The consumer is interested in the amount of caffeine residue. It is known that the median residue for freeze drying is 3.55g/100g of dry matter. This means that if R is the residue (and is continuous and symmetric) that
$P(R < 3.55) = P(R > 3.55) = 0.5$.
In 8 samples of spray dried coffee, the residue was

$$4.8, 4.0, 3.8, 4.3, 3.9, 4.6, 3.1, 3.7 \text{ (g/100g dry matter)}$$

Is there any difference in the caffeine residue between the 2 methods?

$\boxed{\text{Preparation}}$ For $H_0 : \mu = 3.55$, we calculate the signs of differences $\{sign(x_i - 3.55)\}$ and count the number of positive signs $x = 7$.

```
x=c(4.8, 4.0, 3.8, 4.3, 3.9, 4.6, 3.1, 3.7)
x-3.55

## [1]  1.25  0.45  0.25  0.75  0.35  1.05 -0.45  0.15

sign(x-3.55)

## [1]  1  1  1  1  1  1 -1  1
```

$\boxed{\text{H}}$ $H_0 : p_+ = 0.5$ vs $H_1 : p_+ \neq 0.5$.
This means that if the median caffeine residue is 3.55, then we should expect half the signs of differences to be positive.
We choose a 2 sided test, as there is no prior evidence about which method has higher caffeine residue.

A The distribution of caffeine residue is continuous and symmetric.

T

- $\tau = X =$ number of positive signs $\sim Bin(8, 0.5)$ (under H_0)
- Large and small values of x will argue against H_0 for H_1.
- The observed value is $x = 7$.

P P-value $= 2\ P(X \geq 7) \approx 0.07$. (As our observed value $x = 7$ is in the upper tail of the distribution, we consider $P(X \geq 7)$.

```
dbinom(7,8,0.5) + dbinom(8,8,0.5)

## [1] 0.03515625

1-pbinom(6,8,0.5)

## [1] 0.03515625
```

$\boxed{\text{C}}$ As the P-value is 7%, the data are consistent with H_0.
ie If H_0 is true, we would expect to see these data (or more extreme) 7% of the time. It appears there is not a difference in the caffeine residue between the 2 methods of coffee production.

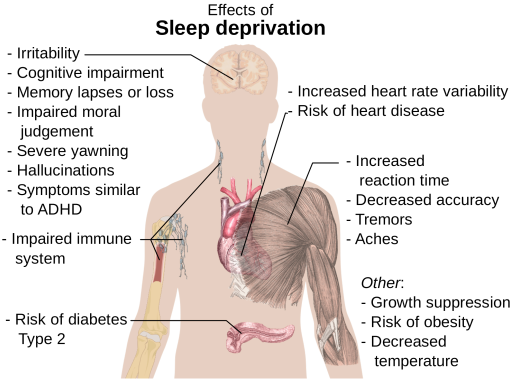
Note: To calculate the P-value, we can use the Binomial formula, the Binomial table, the Binomial histogram, or R.

Example: Student's Sleep Study

The Sign Test can also be applied to **paired data**. Paired data is very common, for example before and after trials and studies on twins. The sign test can be used for paired data, by applying the sign test to the set of differences.

A famous data set is Student (1908). For each of 10 patients, the amount of extra hours sleep was measured, after the administrating of 2 drugs A and B. (A negative result indicates that the patient's sleep reduced). ▸ Sleep Data

Patient	1	2	3	4	5	6	7	8	9	10
A	0.7	-1.6	-0.2	-1.2	-0.1	3.4	3.7	0.8	0.0	2.0
B	1.9	0.8	1.1	0.1	-0.1	4.4	5.5	1.6	4.6	3.4



Is there a difference between the affect of drugs on sleep?

Preparation

We fill out the following table to find the differences of the paired data, and then find the signs of those differences. Notice we eliminate the 5th readings, as the difference is 0.

Patient	1	2	3	4	5	6	7	8	9	10
A	0.7	-1.6	-0.2	-1.2	-0.1	3.4	3.7	0.8	0.0	2.0
B	1.9	0.8	1.1	0.1	-0.1	4.4	5.5	1.6	4.6	3.4
$B - A$										
Sign$(B - A)$					X					

```
a=c(0.7,-1.6,-0.2,-1.2,-0.1,3.4,3.7,0.8,0.0,2.0)
b=c(1.9,0.8,1.1,0.1,-0.1,4.4,5.5,1.6,4.6,3.4)
diff=b-a
diff
```

```
##  [1] 1.2 2.4 1.3 1.3 0.0 1.0 1.8 0.8 4.6 1.4
```

```
sign(diff)
```

```
##  [1] 1 1 1 1 0 1 1 1 1 1
```

For $H_0 : \mu_{diff} = 0$, we calculate the signs of differences $\{sign(diff_i - 0)\}$ and count the number of positive signs $x = 9$.

$\boxed{\text{H}}$ $H_0 : p_+ = 0.5$ vs $H_0 : p_+ \neq 0.5$.
This means that if the difference is 0, then we should expect half the signs of differences to be positive.
We choose a 2 sided test, as there is no prior evidence about which drug effects sleep more.

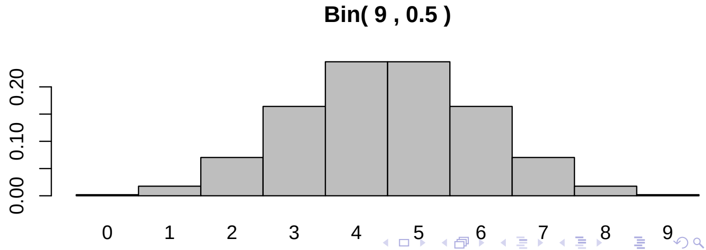
$\boxed{\text{A}}$ The set of differences is continuous and symmetric.

$\boxed{\text{T}}$

- $\tau = X =$ number of positive signs $\sim Bin(9, 0.5)$ (under H_0)
- Large and small values of x will argue against H_0 for H_1.
- The observed value is $x = 9$.

$\boxed{\text{P}}$ P-value $= 2\ P(X \geq 9) = 2P(X = 9) \approx 0.004.$
(As the observed value $x = 9$ is in the upper tail of the distribution, we consider $P(X \geq 9)$.

```
2* dbinom(9,9,0.5)

## [1] 0.00390625

2* (1-pbinom(8,9,0.5))

## [1] 0.00390625
```



$\boxed{\text{C}}$ As the P-value is so small (0.4%), there is strong evidence against H_0.
ie If H_0 is true, we would only expect to see this pattern (or more extreme) 0.4% of the time. It appears there is a difference in the effect of the 2 drugs on sleep.