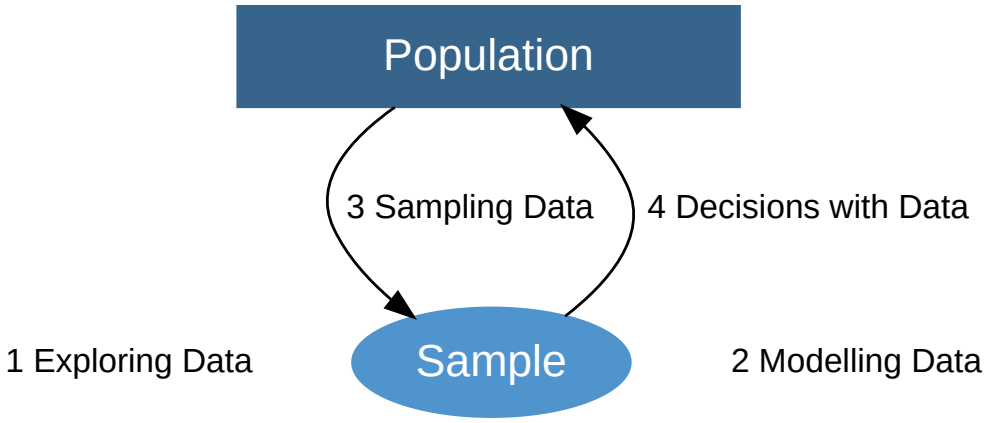
Proportion Test

Decisions with Data | Test for a Proportion

Unit Overview



Module4 Decisions with Data

Test for a Proportion

How can we make evidence based decisions? Is an observed result due to chance or something else? How can we test whether a population has a certain proportion?

Tests for a Mean

How can we test whether a population has a certain mean? Or whether 2 populations have the same mean?

Tests for a Relationship [DATA1001/MATH1115]

How can we test whether 2 variables are linearly related? How can we test whether a categorical variable is in certain proportions?

Proportion Test

Data Story

Does a probiotic treatment cure peanut allergy?

Past research

- Peanut allergy is a life threatening condition.
- Although oral immunotherapy had been shown to safely desensitise most children, only a very small proportion ever develops sustained unresponsiveness.
- Caregivers prefer lower target doses of oral immunotherapy (with the aim of achieving protection against small accidental exposures), rather than higher doses targeting sustained unresponsiveness.

Melbourne data

- The breakthough from the team in Melbourne was to show that an innovative coadministration of the probiotic Lactobacillus rhamnossus CGMCC with peanut oral immunotherapy (2g peanut protein) seemed to cause sustained unresponsiveness to peanut exposure (**desensitisation**).

Initial trial results

Group	Participants	Numbers showing desensitisation
Treatment	29	26
Placebo	28	2

Statistical Thinking

- What is your reaction to this data?
- What is the design of the study?
- Does it prove that the new treatment works?

Proportion Test

H: Hypotheses

Suppose the research team wants to claim that the new oral immunotherapy has a desensitisation rate of higher than 80%.

H_0: 80% of people respond to the treatment. [Or $H_0 : p = 0.8$, where p = Proportion of patient who respond to the treatment (desensitise to peanut allergy).]

H_1: More than 80% respond to the treatment. [Or $H_1 : p > 0.8$]

 What are other reasons that a person could show desensitisation?

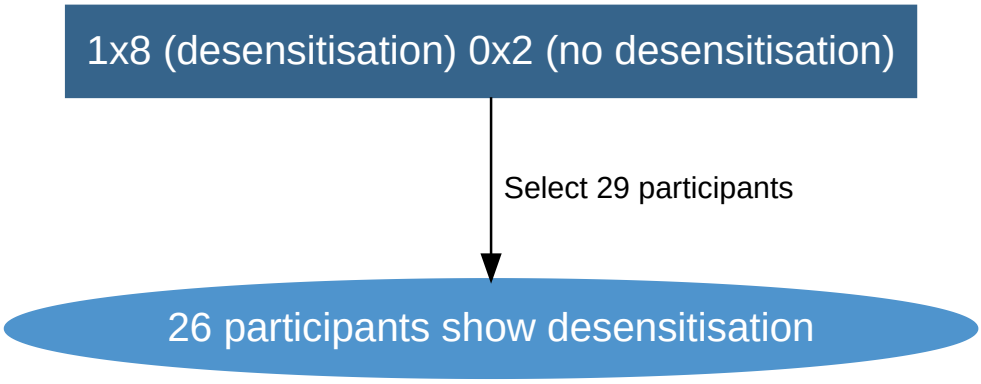
Note: Here we just consider the treatment group and illustrate with a simple null hypothesis ($H_0 = 0.8$). For better analysis, we should use the two sample proportion test which is introduced in later courses.

A: Assumptions

- The participants in the treatment group are independent of each other.
- The chance of becoming desensitised is the same for all participants.

T: Test statistic

If H_0 is true, then we can model the Treatment participants by a simple box model where a 1 (“shows desensitisation”) and 0 (“doesn’t show desensitisation”).



For the box

- The mean is $\frac{1\times 8+0\times 2}{10}=0.8$
- The SD is $(1-0)\sqrt{0.8\times 0.2}=0.4$

For the box model (modelling the Sum of the Sample)

- EV = $29\times 0.8=23.2$
- SE = $\sqrt{29}\times 0.4\approx 2.2$

Test statistic

$$\text{test statistic} = \frac{\text{OV - EV}}{\text{SE}}$$

The observed value of the test statistic is:

$$\text{t}_{obs} = \frac{26 - 23.2}{2.2} \approx 1.3$$

What does the test statistic represent? What does its size tell you?

P: P-value

Calculating P-value

The **p-value** is the chance of observing the test statistic (or something more extreme) if H_0 really is true.

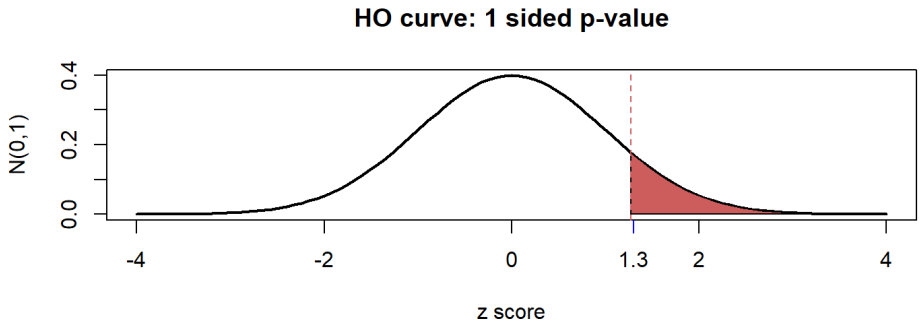
$$P(\text{test statistic} \geq 1.3)$$

Modelling by a Normal (assuming CLT), we find that

$$P(\text{test statistic} \geq 1.3) \approx 0.097$$

```
1 - pnorm(1.3)
```

```
## [1] 0.09680048
```



Note:

Due to the small sample size, the Normality assumption for the Sample Sum (total count) may not be suitable. So next, we use simulation to approximate the distribution for the Sample Sum.

Size of P-value

The P-value (0.097) is just larger than 0.05.

What does this P-value suggest?

C: Conclusion

- We conclude that the data is consistent with the null hypothesis.
- Therefore, the new treatment does not have an effectivness rate higher than 0.8.
- Note we have not proved that the effectiveness rate is 0.8. Rather, we have just failed to find sufficient evidence to claim an effectiveness rate of higher than 0.8.

 Have a go at home …

Try testing $H_0 : p = 0.7$. You will find a test statistic of 2.3, which results in a significant result.

Effect of H1

Effect of H_1 on P-value

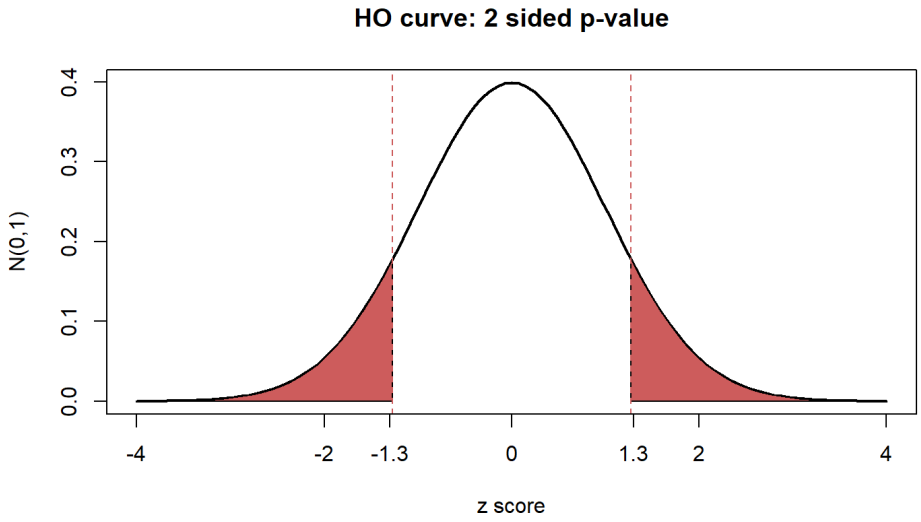
We can have 2 types of alternative hypotheses:

1 sided:

Specifies the direction of the altenative hypothesis. Eg $H_1 : p > 0.8$.

2 sided:

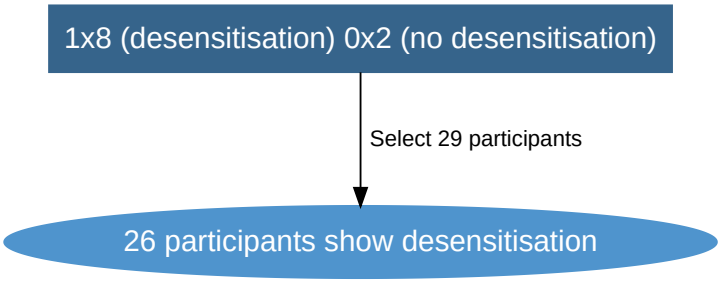
- Does not specify the direction of the alternative hypothesis. Eg $H_1 : p \neq 0.8$.
- In this case we (usually) double the p-value.


HO curve: 2 sided p-value
N(0,1)
0.0
0.1
0.2
0.3
0.4
-4
-2
-1.3
0
1.3
2
4
z score

Simulating P-values

Simulating P-values

- It is always possible that the 26 participants who show desensitisation was just a rare result.



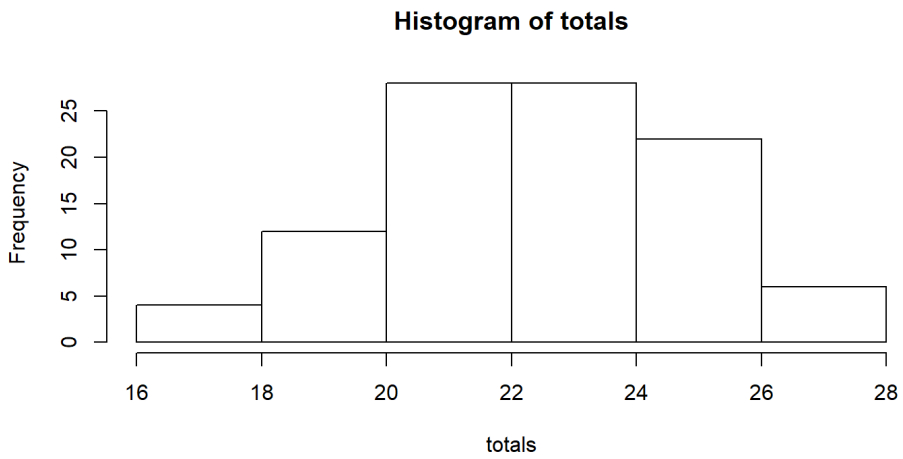
Simulate 100 times

```
set.seed(1)
box = c(1, 1, 1, 1, 0)  # using proportions of 80% and 20% in box
totals = replicate(100, sum(sample(box, 29, rep = T)))
table(totals)
```

```
## totals
## 16 17 18 19 20 21 22 23 24 25 26 27 28
##  1  2  1  7  5 12 16 15 13 16  6  4  2
```

```
hist(totals)
abline(v = 28, col = "green")
```

```
## totals
## 16 17 18 19 20 21 22 23 24 25 26 27 28
##  1  2  1  7  5 12 16 15 13 16  6  4  2
```



The estimated p-value here would be (6+4+2)/100.

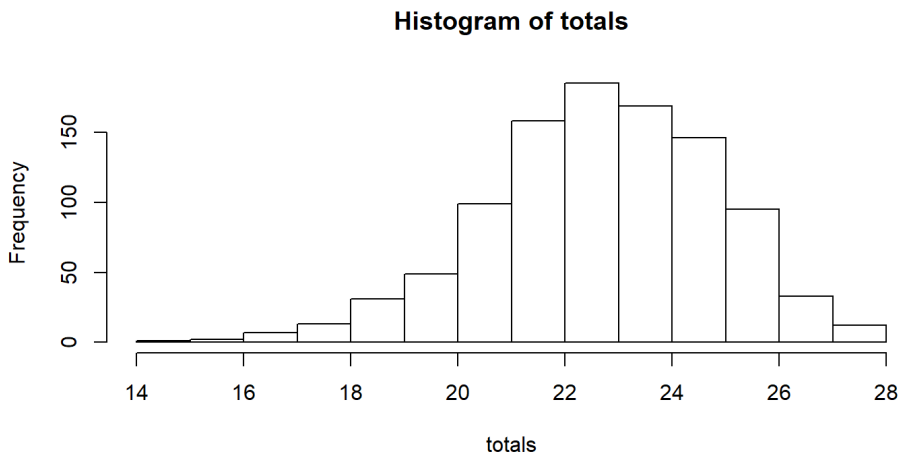
Simulate 1000 times

```
set.seed(1)
box = c(1, 1, 1, 1, 0)
totals = replicate(1000, sum(sample(box, 29, rep = T)))
table(totals)
```

```
## totals
##  14  16  17  18  19  20  21  22  23  24  25  26  27  28
##   1   2   7  13  31  49  99 158 185 169 146  95  33  12
```

```
hist(totals)
```

```
## totals
##  14  16  17  18  19  20  21  22  23  24  25  26  27  28
##   1   2   7  13  31  49  99 158 185 169 146  95  33  12
```



What would be the approximate p-value here?

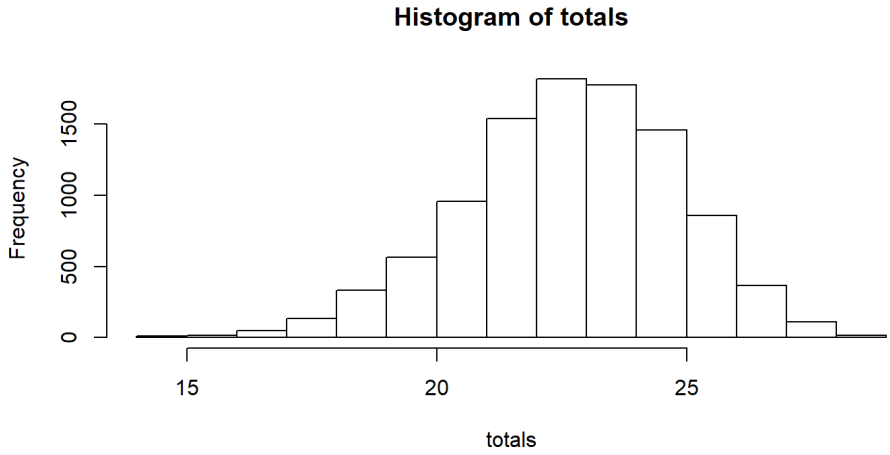
Simulate 10,000 times

```
set.seed(1)
box = c(1, 1, 1, 1, 0)
totals = replicate(10000, sum(sample(box, 29, rep = T)))
table(totals)
```

```
## totals
##   14   15   16   17   18   19   20   21   22   23   24   25   26   27   28
##    2    7   14   51  134  334  565  958 1538 1816 1777 1457  858  365  110
##   29
##   14
```

```
hist(totals)
```

```
## totals
##   14   15   16   17   18   19   20   21   22   23   24   25   26   27   28
##    2    7   14   51  134  334  565  958 1538 1816 1777 1457  858  365  110
##   29
##   14
```



What would be the approximate p-value here?

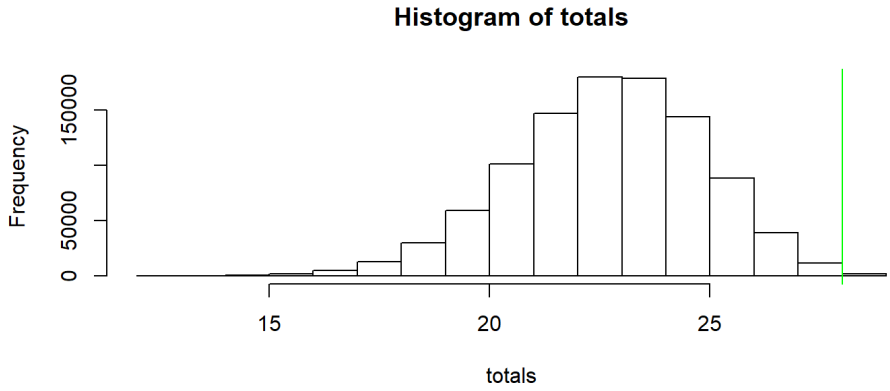
Simulate 1,000,000 times

```
set.seed(1)
box = c(1, 1, 1, 1, 0)
totals = replicate(1e+06, sum(sample(box, 29, rep = T)))
table(totals)
```

```
## totals
##     12     13     14     15     16     17     18     19     20     21
##      5     20    119    428   1576   4716  12747  29659  59001 101204
##     22     23     24     25     26     27     28     29
## 147208 180209 178786 143920  88242  39131  11517   1512
```

```
hist(totals)
```

```
## totals
##      12      13      14      15      16      17      18      19      20      21
##       5      20     119     428    1576    4716   12747   29659   59001  101204
##      22      23      24      25      26      27      28      29
##  147208  180209  178786  143920   88242   39131   11517    1512
```



What would be the approximate p-value here?

- The p-value = (88045 + 39175 + 11307 + 1494)/1000000 = 0.140021.
- This is the approximation/estimate of the chance of getting 26 or more participants who show desensitisation.
- Compare this with 0.09, using the CLT.
- Why the difference? The much higher proportion of 1s vs. 0s (4-fold ratio) combined with a relatively small sample size (29) implies that the distribution of the number of 1s in the sample is left skewed.
- Extension: The exact p-value is based on the Binomial model.

```
x = c(26, 27, 28, 29)
sum(dbinom(x, 29, 0.8))
```

```
## [1] 0.1403805
```

Further research on peanut allergies

Current breakthrough?

Statistical Thinking

A research editorial entitled *Probiotic and peanut oral immunotherapy: a breakthrough for allergy treatment* concluded: “The therapeutic effect shown is remarkable and redefines the notion of sustained unresponsiveness.”

Do you agree?

Further research

Despite achieving an important milestone, this study included small patient numbers, did not retain all participants in long-term follow-up , and did not include entry peanut challenge to show participants' baseline threshold (which was not standard practice in early oral immunotherapy studies, but is routinely done now) or comparison of oral immunotherapy with PPOIT to delineate the effect that is attributable to the probiotic. These limitations are being addressed in a multicentre trial aiming to replicate these findings. The field of allergy and immunology anxiously awaits these results. Successful replication of these data could lead to broader questions of generalisability of the probiotic plus oral immunotherapy approach to other allergens

Statistical Thinking

What were the limitations of the study? How are these limitations being addressed?

Summary

Key Words

hypotheses, assumptions, test statistic, p-value