

Exercise walk-through

4.4

Non-parametric tests

A-Level Further Mathematics

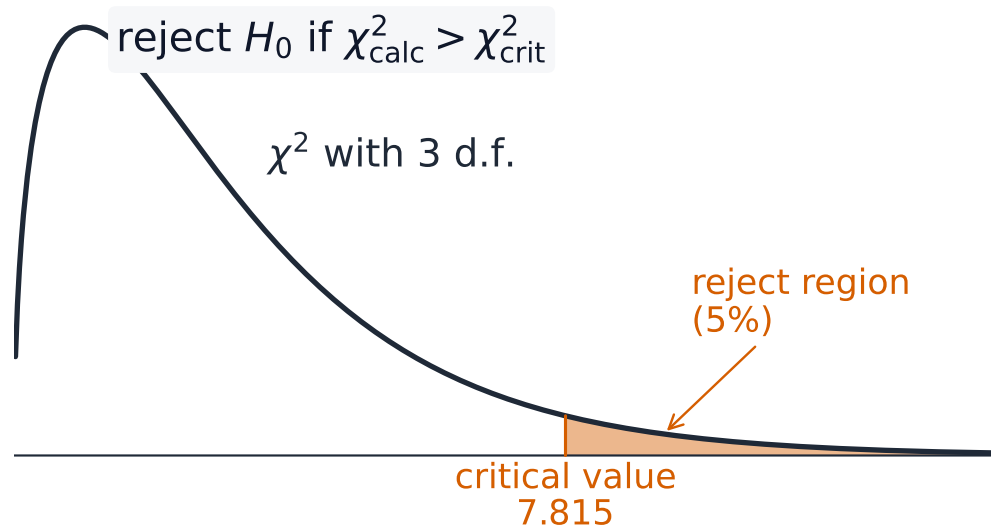
Chi-squared tests

Compare observed counts O with expected E :

$$\chi^2 = \sum \frac{(O - E)^2}{E}.$$

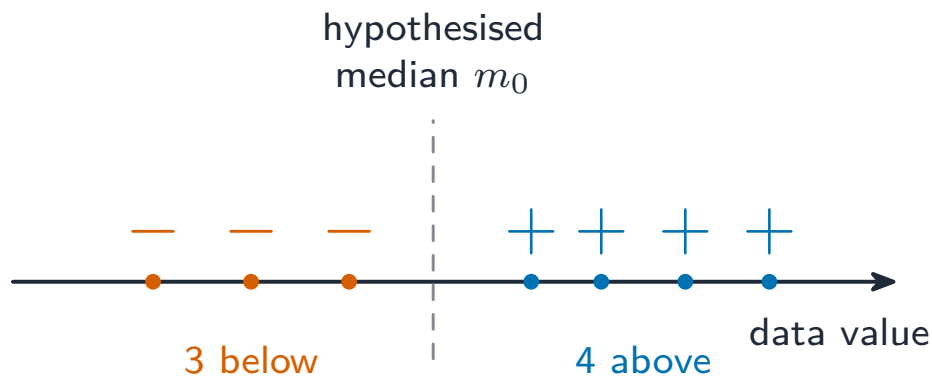
Large $\chi^2 \Rightarrow$ reject H_0 . Degrees of freedom = categories $- 1$, minus any estimated parameters.

Chi-squared tests, in detail



The test rejects the model when χ^2 exceeds the critical value, landing in the shaded 5%

Non-parametric tests



No assumption that the data is normal:

- **sign test** – counts above/below a median
- **Wilcoxon** signed-rank & rank-sum – use the sizes too

Worked example – a sign test

Ten values are tested against H_0 : median = 5. Nine fall above it. Is that evidence against H_0 ?

- under H_0 , each value is above with probability 0.5, so the count $X \sim B(10, 0.5)$
- $P(X \geq 9) = \binom{10}{9}(0.5)^{10} + \binom{10}{10}(0.5)^{10}$
- $= \frac{10 + 1}{1024} = 0.0107$
- $0.0107 < 0.05$, so **reject** H_0 at the 5% level

The whole test is a binomial tail. Note it uses *only* the signs – discard how far above the median each value was, and a great deal of information with it.

The three non-parametric tests, and what each uses

sign test

uses only **which side** of the hypothesised median each value falls – the weakest, and the most assumption-free

Wilcoxon signed-rank

the **matched-pairs** test: uses the **sizes** of the differences as well as their signs, so it is more powerful

Wilcoxon rank-sum

for **two independent** samples – ranks all the data together and compares the rank totals

when to use them

when normality cannot be assumed, or the data are ordinal

goodness of fit

χ^2 compares observed with expected frequencies; a **contingency table** tests independence

Matched pairs → signed-rank. Two independent samples → rank-sum. Choosing the wrong one loses the whole question, so read for **paired or not** first.

Exercises 2.1·3.1

2.1 State when a non-parametric test is preferred over a t-test. [1]

2.2 In a sign test, state the distribution of the number of positive signs under H_0 . [1]

3.1 The Wilcoxon signed-rank test uses, in addition to the signs of the differences: [1]

☐

A nothing else

B the sizes (ranks) of the differences

C the sample mean

D the population variance

Exercises 2.1-3.1

2.1 State when a non-parametric test is preferred over a t-test. [1]

2.1

- when the data cannot be assumed to be normally distributed

2.2 In a sign test, state the distribution of the number of positive signs under H_0 . [1]

2.2

- $\sim B(n, 0.5)$

3.1 The Wilcoxon signed-rank test uses, in addition to the signs of the differences: [1]

☐

A nothing else

B the sizes (ranks) of the differences

C the sample mean

D the population variance

3.1 B

- The signed-rank test also uses the ranks (sizes) of the differences

Exercise 3.2 (a)

3.2 Ten people rate a product before and after a change. Eight rate it higher (+) and two lower (−).

(a) State H_0 and H_1 for a sign test of "no change" vs "an improvement". [2]

Exercise 3.2 (a)

3.2 Ten people rate a product before and after a change. Eight rate it higher (+) and two lower (−).

(a) State H_0 and H_1 for a sign test of "no change" vs "an improvement". [2]

(a)

- H_0 : median difference = 0 (no change); H_1 : median difference > 0 (improvement)

Exercise 3.2 (b)

3.2 Ten people rate a product before and after a change. Eight rate it higher (+) and two lower (−).

(b) Using $X \sim B(10, 0.5)$, find $P(X \geq 8)$ and state the conclusion at the 5% level. [4]

Exercise 3.2 (b)

3.2 Ten people rate a product before and after a change. Eight rate it higher (+) and two lower (−).

(b) Using $X \sim B(10, 0.5)$, find $P(X \geq 8)$ and state the conclusion at the 5% level. [4]

(b)

- $P(X \geq 8) = P(8) + P(9) + P(10) = \frac{45 + 10 + 1}{1024} = \frac{56}{1024} = 0.0547$; $0.0547 > 0.05$, so do not reject H_0 – insufficient evidence of improvement

Exercise 3.3

3.3 Explain why a non-parametric test might be chosen instead of a t-test for a small sample. [2]

Exercise 3.3

3.3 Explain why a non-parametric test might be chosen instead of a t-test for a small sample. [2]

3.3

- a t-test assumes the population is normal; for a small sample that assumption cannot be checked reliably; a non-parametric test makes no such assumption, so it is safer when normality is doubtful