

Lecture 53: Median test for two samples; Fisher's exact test

Statistical Inference–II ■ Units IV–V: Nonparametric Inference

Kapil Kumar

Department of Statistics
Chaudhary Charan Singh University, Meerut

M.Sc. Statistics, Semester III
Academic Session 2026–2027

Previous lecture

Lecture 52: Median test: construction; a k -sample example.

Today

- The Median Test for Two or More Independent Samples
 - Connection with Pearson Chi-Square Procedures

Example: median test for two samples

Example 53.1 (Median test for two samples)

The times (days) taken to heal after two surgical procedures were:

Procedure A: 23, 31, 18, 27, 35, 22, 29, 20, 26, 24

Procedure B: 33, 28, 41, 37, 30, 45, 34, 25, 39, 36

Test at the 5% level whether the two procedures have the same median healing time.

Solution

Pooled median. The $N = 20$ pooled observations in increasing order are

18, 20, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 33, 34, 35, 36, 37, 39, 41, 45,

so $M = (29 + 30)/2 = 29.5$. No observation equals M .

Example: median test for two samples (cont.)

Solution (cont.)

Table.

	A	B	Total
Below M	8	2	10
Above M	2	8	10
Total	10	10	20

$$E_{ij} = \frac{10 \times 10}{20} = 5 \quad \text{for every cell.}$$

Chi-square test.

$$X^2 = \sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = 4 \times \frac{3^2}{5} = 7.2,$$

$$X^2_{\text{Yates}} = \sum \frac{(|O_{ij} - E_{ij}| - \frac{1}{2})^2}{E_{ij}} = 4 \times \frac{2.5^2}{5} = 5.0.$$

Example: median test for two samples (cont.)

Solution (cont.)

With 1 degree of freedom, $\chi^2_{1;0.05} = 3.841$; the p -values are 0.0073 and 0.0253.

Exact (Fisher) test. Given the margins, the number O_{11} of A's below M is hypergeometric: $\mathbb{P}(O_{11} = k) = \binom{10}{k} \binom{10}{10-k} / \binom{20}{10}$, with $\binom{20}{10} = 184756$. Hence

$$\mathbb{P}(O_{11} \geq 8) = \frac{\binom{10}{8} \binom{10}{2} + \binom{10}{9} \binom{10}{1} + \binom{10}{10} \binom{10}{0}}{184756} = \frac{2025 + 100 + 1}{184756} = 0.0115,$$

and, by symmetry, the two-sided exact p -value is $2(0.0115) = 0.0230$.

Example: median test for two samples (cont.)

Solution (cont.)

Conclusion. H_0 is rejected at the 5% level: procedure A tends to give shorter healing times. Note that the Yates-corrected statistic is much closer to the exact test than the uncorrected one; for a 2×2 table with small counts, Fisher's exact test should be used.

The same Pearson statistic is used more generally for goodness of fit and for independence or homogeneity in contingency tables.

- For c categories with fully specified probabilities, a goodness-of-fit statistic has asymptotic χ^2_{c-1} distribution;
- if q parameters are estimated efficiently from the grouped frequencies (e.g. by multinomial maximum likelihood or minimum χ^2), the degrees of freedom are $c - 1 - q$ (a caution about estimating the parameters from the ungrouped data: Lecture Notes, Appendix A);
- for an $r \times c$ table, the asymptotic degrees of freedom are $(r - 1)(c - 1)$.

Expected frequencies should not be too small for the chi-square approximation to be reliable.

Practice (Lecture Notes, Chapter 3)

- Exercises (Section 3.15): 3.16.

Next lecture

Lecture 54: Wilcoxon signed-rank test: null distribution.