

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

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 51: Paired sign test; confidence interval for a median.

Today

- The Median Test for Two or More Independent Samples

The median test

The median test compares the proportions of observations above and below a pooled median. It applies to two or more independent samples and therefore provides a simple k -sample nonparametric procedure.

Suppose there are k independent samples with sizes $n_1, \dots, n_k$. Let M be the median of all $N = \sum n_j$ observations pooled together. For each sample, count the numbers below and above M and form the $2 \times k$ table

	1	2	$\dots$	k	Total
Below M	O_{11}	O_{12}	$\dots$	O_{1k}	O_{1+}
Above M	O_{21}	O_{22}	$\dots$	O_{2k}	O_{2+}
Total	O_{+1}	O_{+2}	$\dots$	O_{+k}	N

Here $O_{+j} = n_j$ when no observation is deleted.

Observations tied at M (there is always at least one when N is odd) are commonly deleted; the resulting effective column totals O_{+j} and effective total N must then be used.

An alternative convention keeps every observation and classifies it as “at or below M ” or “above M ”, i.e. ties are counted in the first row; this convention is used in Exercise 3.15.

The rule must be fixed before the data are examined.

The median test: Pearson statistic

Under the null hypothesis that the samples have a common median and under the usual common-shape interpretation, the expected frequencies are

$$E_{ij} = \frac{O_{i+}O_{+j}}{N},$$

and the Pearson statistic is

$$X^2 = \sum_{i=1}^2 \sum_{j=1}^k \frac{(O_{ij} - E_{ij})^2}{E_{ij}}.$$

For sufficiently large expected counts, $X^2 \sim \chi_{k-1}^2$. For $k = 2$ this is the ordinary 2×2 chi-square test, and an exact hypergeometric/Fisher calculation may be used when expected counts are small.

More generally, if the k populations are identical and continuous, then, conditionally on all the margins, the first-row counts have the multivariate hypergeometric distribution

$$\mathbb{P}(O_{11} = a_1, \dots, O_{1k} = a_k) = \frac{\prod_{j=1}^k \binom{O_{+j}}{a_j}}{\binom{N}{O_{1+}}}, \quad a_1 + \dots + a_k = O_{1+},$$

and the exact p -value is the total probability of the tables whose X^2 is at least the observed value.

Median test

1. Pool the samples and find the pooled median.
2. Count observations above and below it in each sample.
3. State how ties at the pooled median are handled.
4. Compute the expected counts and Pearson's chi-square statistic.
5. Compare with χ^2_{k-1} , or use an exact method when the table is sparse.

Remark 52.1

A significant median test shows that the proportions above the pooled median are not homogeneous. Interpreting this solely as a difference in medians is most defensible when the population distributions have similar shapes and differ primarily in location.

Example: median test for three samples

Example 52.2 (Median test for three samples)

The times (minutes) taken to complete a task by workers trained by three methods were:

Method A: 18, 21, 15, 24, 33, 20

Method B: 26, 19, 35, 23, 30, 37, 22

Method C: 34, 41, 29, 38, 31

Apply the median test at the 5% level.

Solution

Pooled median. The $N = 18$ pooled observations in increasing order are

15, 18, 19, 20, 21, 22, 23, 24, 26, 29, 30, 31, 33, 34, 35, 37, 38, 41,

so $M = (26 + 29)/2 = 27.5$. No observation equals M , so nothing is deleted.

Example: median test for three samples (cont.)

Solution (cont.)

Table and expected counts. Counting in each sample,

	A	B	C	Total
Below M	5	4	0	9
Above M	1	3	5	9
Total	6	7	5	18

$$E_{1j} = E_{2j} = \frac{9 O_{+j}}{18} = 3, 3.5, 2.5.$$

Pearson statistic. The two rows contribute equally, so

$$X^2 = 2 \left\{ \frac{(5 - 3)^2}{3} + \frac{(4 - 3.5)^2}{3.5} + \frac{(0 - 2.5)^2}{2.5} \right\} = \frac{164}{21} = 7.8095.$$

With $k - 1 = 2$ degrees of freedom, $\chi^2_{2;0.05} = 5.991$ and $\mathbb{P}(\chi^2_2 \geq 7.8095) = e^{-7.8095/2} = 0.0201$.

Example: median test for three samples (cont.)

Solution (cont.)

Exact conditional test. All expected counts are below 5, so the χ^2 approximation is doubtful. The exact conditional test, which enumerates all tables with the observed margins (Lecture Notes, Appendix A), gives $p_{\text{exact}} = 0.0253$.

Conclusion. Both p -values are below 0.05: the hypothesis of a common median is rejected. Method A tends to give shorter and method C longer times; a statement about medians presumes similar shapes for the three time distributions.

Practice (Lecture Notes, Chapter 3)

- Short-answer questions (Section 3.14): Q12.
- Exercises (Section 3.15): 3.15.

Next lecture

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