Lecture 49: Runs up and down; the Wald–Wolfowitz two-sample runs 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

This lecture

Previous lecture

Lecture 48: Mean and variance of the number of runs; the one-sample runs test.

Today

- Tests for Randomness Based on Runs
  - Runs Up and Down
  - Two-Sample Wald–Wolfowitz Runs Test

Runs up and down

For quantitative observations in their recorded order, form the signs

$$S_i = \begin{cases} +, & X_{i+1} > X_i, \\ -, & X_{i+1} < X_i, \end{cases} \qquad i = 1, \ldots, n-1.$$

Repeated observations require a stated convention. Let R_{ud} be the number of runs in this sequence of directions. For independent observations from a continuous distribution,

$$\mathbb{E}(R_{ud}) = \frac{2n-1}{3}, \qquad \operatorname{Var}(R_{ud}) = \frac{16n-29}{90} \quad (n \geq 4).$$

The standardised statistic is approximately normal for sufficiently large n. This version is particularly sensitive to monotone trend.

Example: runs up and down

Example 49.1 (Runs up and down)

The output (in hundreds of units) of a factory in 14 consecutive months was

$$50,\ 53,\ 57,\ 58,\ 62,\ 60,\ 56,\ 61,\ 64,\ 66,\ 70,\ 67,\ 71,\ 74.$$

Test the hypothesis of randomness against a trend.

Solution

There are no repeated adjacent values. The 13 directions of change are

$$+\ +\ +\ +\ |\ -\ -\ |\ +\ +\ +\ +\ |\ -\ |\ +\ +,$$

so $R_{ud} = 5$.

Example: runs up and down (cont.)

Solution (cont.)

With $n = 14$,

$$\mathbb{E}(R_{ud}) = \frac{2(14) - 1}{3} = 9, \qquad \mathrm{Var}(R_{ud}) = \frac{16(14) - 29}{90} = \frac{195}{90} = 2.1667,$$

and $\sqrt{\mathrm{Var}(R_{ud})} = 1.4720$. A trend produces long monotone stretches and hence few runs, so the lower tail is relevant. With the continuity correction,

$$z = \frac{5 + 0.5 - 9}{1.4720} = -2.378, \qquad \mathbb{P}(R_{ud} \le 5) \approx \Phi(-2.378) = 0.0087.$$

Randomness is rejected at the 5% (indeed the 1%) level: the long rising stretches indicate an upward trend.

Example: runs up and down (cont.)

Solution (cont.): comment

Here the normal approximation is adequate: the exact probability is

$$\mathbb{P}(R_{ud} \leq 5) = 0.0079$$

(Lecture Notes, Appendix A). The test uses only the directions of successive changes, not their sizes. A series that rises overall but with frequent small reversals, such as the series of Exercise 3.11, can have R_{ud} close to its mean; such gradual trends are better detected by a rank correlation with time.

Two-sample Wald–Wolfowitz runs test

Let $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ be independent samples from continuous distributions. Pool and order all observations, then replace each ordered value by its sample label X or Y. Under

$$H_0 : F_X = F_Y,$$

all allocations of the m labels X among the $m+n$ ordered positions are equally likely. Therefore the theorems on the null distribution and on the mean and variance of the number of runs apply with $n_1 = m$, $n_2 = n$ and $N = m + n$ in place of n.

Interpreting the two-sample runs test

Very few runs indicate that one sample tends to occupy one part of the ordered scale and the other sample another part; an unusually large number indicates excessive interlacing.

The test is sensitive to broad distributional differences, not only to a location shift. The usual test rejects H_0 for small values of R.

Remark 49.2

The two-sample runs test requires the observations to be pooled in increasing order. A one-sample time-order runs test instead preserves the original order. Confusing these two orderings changes the problem completely.

Example: Wald–Wolfowitz runs test

Example 49.3 (Wald–Wolfowitz runs test)

Diameters (mm) of bolts produced by two machines were

$$X \text{ (machine 1)}: \ 9.62,\ 10.24,\ 9.75,\ 10.36,\ 9.71,\ 10.29;$$
$$Y \text{ (machine 2)}: \ 10.02,\ 9.93,\ 10.41,\ 9.88,\ 10.11,\ 9.96,\ 10.05.$$

Test $H_0 : F_X = F_Y$ by the two-sample runs test at the 5% level.

Solution

Here $m = 6$, $n = 7$, $N = 13$. The pooled ordered sample and its labels are

9.62	9.71	9.75	9.88	9.93	9.96	10.02	10.05	10.11	10.24	10.29	10.36	10.41
X	X	X	Y	Y	Y	Y	Y	Y	X	X	X	Y

so $XXX\,|\,YYYYYYY\,|\,XXX\,|\,Y$ has $R = 4$ runs.

Example: Wald–Wolfowitz runs test (cont.)

Solution (cont.)

By the Theorem on the null distribution of the number of runs, with $n_1 = 6$, $n_2 = 7$ and $\binom{13}{6} = 1716$,

$$\mathbb{P}(R = 2) = \frac{2\binom{5}{0}\binom{6}{0}}{1716} = \frac{2}{1716}, \qquad \mathbb{P}(R = 3) = \frac{\binom{5}{1}\binom{6}{0} + \binom{5}{0}\binom{6}{1}}{1716} = \frac{11}{1716},$$

$$\mathbb{P}(R = 4) = \frac{2\binom{5}{1}\binom{6}{1}}{1716} = \frac{60}{1716},$$

so the exact p-value is

$$\mathbb{P}(R \leq 4) = \frac{2 + 11 + 60}{1716} = \frac{73}{1716} = 0.0425 < 0.05.$$

H_0 is rejected at the 5% level.

Example: Wald–Wolfowitz runs test (cont.)

Solution (cont.)

For comparison, by the Theorem on the mean and variance of the number of runs,

$$\mathbb{E}(R) = 1 + \frac{2(42)}{13} = 7.4615, \qquad \mathrm{Var}(R) = \frac{2(42)(84-13)}{13^2(12)} = 2.9408,$$

so $z = (4 + 0.5 - 7.4615)/1.7149 = -1.727$ and $\Phi(-1.727) = 0.0421$.

Comment. The X observations occupy both ends of the pooled ordering and the Y observations the middle: machine 1 is more variable. The runs test detects such differences of spread or shape; a test aimed only at location can miss them (see the remark "Location versus spread" in the Mann–Whitney section).

After this lecture

Practice (Lecture Notes, Chapter 3)

- Exercises (Section 3.15): 3.11.

Next lecture

Lecture 50: One-sample sign test.