Two-Sample Rank Tests for Skewness and Kurtosis

Yadolah Dodge
University of Neuchâtel, Statistics Group, Switzerland.

Valentin Rousson
University of Zurich, Department of Biostatistics, Sumatrastrasse 30, 8006 Zurich, Switzerland,rousson@ifspm.unizh.ch

1. Introduction

Consider two independent random samples $X_1, ..., X_m$ and $Y_1, ..., Y_n$ drawn from unknown continuous distributions F and G. To test the null hypothesis of the equality of the two distributions $H_0 : F(x) = G(x)\ \forall x$ against the alternative $H_1 : \exists\theta,\ F(x) = G((x-\theta))\ \forall x$, where the location shift $\theta \neq 0$, one may use the Mann and Whitney's (1947) procedure, whose test statistic is the sum of the the ranks of the Y_i in the combined sample. This statistic is distribution-free under H_0.

If we are interested to test a change of scale between F and G when the two distributions have possibly a different location, that is if we wish to test the null hypothesis H_1 against the alternative

$$H_2 : \exists\theta, \sigma > 0,\ F(x) = G((x-\theta)/\sigma)\ \forall x$$

where the scale ratio $\sigma \neq 1$, we may use the procedure proposed by Moses (1963). It consists to rearrange the X_i and Y_i in random order, to define for $i = 1, 3, 5...$

$$\begin{aligned} U_i &= |X_{i+1} - X_i|, \\ V_i &= |Y_{i+1} - Y_i|, \end{aligned}$$

and to apply the Mann and Whitney's test on the U_i and V_i. If a sample size is odd, the last observation is discarded. Again, we have here an exact test.

In this note, we are interested to define a strictly nonparametric procedure for testing the null hypothesis H_2 that F and G have the same shape, against the alternative that the two distributions differ in skewness or in kurtosis.

2. Skewness and Kurtosis Orderings

Let X and Y be two continuous random variables with distribution F and G respectively. The one-sided version of Mann and Whitney's (1947) test is known to be powerful against the alternative hypothesis that X is stochastically larger than Y, noted here $Y \prec X$. On the other hand, Oja (1981) defines X to be stochastically more skew to the right than Y if

$$\frac{Y_{(3)} - Y_{(2)}}{Y_{(3)} - Y_{(1)}} \prec \frac{X_{(3)} - X_{(2)}}{X_{(3)} - X_{(1)}}$$

where $X_{(1)}, X_{(2)}, X_{(3)}$ and $Y_{(1)}, Y_{(2)}, Y_{(3)}$ are ordered samples of size three from F and G. Similarly, X is stochastically more flattened (or has stochastically less kurtosis) than Y if

$$\frac{Y_{(3)} - Y_{(2)}}{Y_{(4)} - Y_{(1)}} \prec \frac{X_{(3)} - X_{(2)}}{X_{(4)} - X_{(1)}}$$

where $X_{(1)},...,X_{(4)}$ and $Y_{(1)},...,Y_{(4)}$ are ordered samples of size four from F and G. This suggests the following nonparametric procedure for testing whether two distributions differ in skewness or in kurtosis, similar to the rank test for dispersion proposed by Moses (1963). Note that Behboodian (1989) proposed a procedure to test the symmetry of a distribution which also considers independent triples of observations, and which requires to divide by three the sample size.

Let rearrange the X_i and Y_i in random order, and define for $i = 1, 4, 7, ...$

$$U_i = \frac{X_{(i+2)} - X_{(i+1)}}{X_{(i+2)} - X_{(i)}},$$
$$V_i = \frac{Y_{(i+2)} - Y_{(i+1)}}{Y_{(i+2)} - Y_{(i)}}.$$

The proposed procedure consists to apply the Mann and Whitney's test on the U_i and V_i. If m or n are not multiples of three, some observations are discarded. The test statistic obtained is distribution-free under the null hypothesis H_2 defined in Section 1. A rejection of H_2 using this procedure means that the two distributions differ in skewness.

Similarly one could apply the Mann and Whitney's test on the

$$U_i = \frac{X_{(i+2)} - X_{(i+1)}}{X_{(i+3)} - X_{(i)}},$$
$$V_i = \frac{Y_{(i+2)} - Y_{(i+1)}}{Y_{(i+3)} - Y_{(i)}},$$

where $i = 1, 5, 9, ...$ Again, some observations are discarded if m or n are not multiples of four. Rejecting the null hypothesis H_2 by applying this procedure means that the two distributions differ in kurtosis.

We will apply this methodolgy to find the direction of relationship between two variables in a simple regression problem as in Dodge and Rousson (2001).

REFERENCES

Behboodian, J. (1989). A Note On Skewness and Symmetry. *The Statistician*, 38, 21–23.

Dodge, Y. and Rousson, V. (2001). On Asymmetric Properties of the Correlation Coefficient in the Regression Setting. *The American Statistician*, 55, 51–54.

Mann, H.B. and Whitney, D.R. (1947). On a Test of Whether One of Two Random Variables is Larger Than the Other. *Ann. Math. Stat.*, 18, 50–60.

Moses, L. (1963). Rank Tests of Dispersion. *Ann. Math. Stat.*, 34, 973–983.

Oja, H. (1981). Two Location and Scale-Free Goodness-of-Fit Tests. *Biometrika*, 68, 637–640.

RESUME

Dans cet article nous construisons les tests de rang pour comparer le skewness et le kurtosis de deux distributions. Nos procédures sont strictement non-paramétriques.

Mots clés: Kurtosis, Nonparametric tests, Rank tests, Skewness.