Journal of the Korean Statistical Society 44 (2015) 592–605

Contents lists available at ScienceDirect

Journal of the Korean Statistical Society

journal homepage: www.elsevier.com/locate/jkss

The modified Mood test for the scale alternative and its numerical comparisons

Hidetoshi Murakami[a], Hyung-Tae Ha[b,*]

[a] Department of Mathematical Information Science, Tokyo University of Science, 1-3 Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan

[b] Department of Applied Statistics, Gachon University, Sungnam-ci, Kyunggi-do 461-701, Republic of Korea

ARTICLE INFO

Article history:
Received 13 May 2014
Accepted 24 March 2015
Available online 22 April 2015

AMS 2000 subject classifications:
primary 62G10
secondary 62E17

Keywords:
Approximation
Critical values
Asymptotic efficiency
Modified Mood test
Power comparison

ABSTRACT

On statistical hypotheses testing in two-sample problems, the Mood test is popular as one of the most efficient nonparametric tests for dispersion differences. In this paper, we show that a modification of the Mood test proposed by Tamura (1963) can gain even more efficiency and power under various distributional assumptions. The accuracy of the proposed approximations to the tail probabilities and critical values of the modified Mood test, namely M_p, were investigated. Our results showed that the Edgeworth expansion was more accurate than the other approximations. Asymptotic efficiencies and the optimal value p of the modified Mood test under various distributional assumptions were examined, with the results revealing that the large (small) value of p was useful for light (heavy) tail distributions. Additionally, the power of the modified Mood test for the one-sided alternative with various population distributions for small sample sizes was investigated via Monte Carlo simulations. Finally, the proposed method was demonstrated using real data.

1. Introduction

Testing the differences in dispersion as a measure of spread in two-sample problems has been extensively discussed in statistical literature. In parametric hypothesis testing, the F-test is widely utilized for testing the equality of variances when it is reasonable to assume that populations are normally distributed. However, it is well known that the F-test is not robust with respect to the normality assumption, but is very sensitive to kurtosis of the sampling distributions of interest. Given the circumstances that the underlying distribution is not adequately understood to assume normality or some other specific distribution, the nonparametric approach has been considered as a viable alternative to overcome the sensitivity of the normality assumption.

Many nonparametric tests of dispersion have been proposed over the last half century, based on the weighted summation of ranks of the combined observations of the two samples. Popular nonparametric tests of dispersion include the Mood (Mood, 1954), Freund–Ansary–Bradley (Ansari & Bradley, 1960; Freund & Ansari, 1957), and Siegel–Tukey (Siegel & Tukey, 1960) tests. Among the many nonparametric tests of dispersions, the Mood test is one of the most powerful, such that it is commonly utilized for nonparametric scale tests. In fact, the Mood test provides better performance than the others. For instance, with respect to the performance measure of the asymptotic relative efficiency (ARE) of the F test, under the

* Corresponding author.
E-mail addresses: murakami@gug.math.chuo-u.ac.jp (H. Murakami), htha@gachon.ac.kr (H.-T. Ha).

http://dx.doi.org/10.1016/j.jkss.2015.03.005

assumption of normal populations differing only in variance, the Mood test has a higher value than the others, meaning that the Mood test requires a smaller number of sample observations to provide the same power. The AREs of the F test, under the assumption of normal populations differing only in variance, are 0.76 for the Mood test and 0.608 for both the Freund–Ansary–Bradley and Siegel–Tukey tests. Recent reports have outlined the development of a substantial number of nonparametric control charts, in which no underlying distribution is assumed for the process output. Additionally, Chakraborti and van de Wiel (2008) indicated the importance of considering the charting statistics for dispersion based on a nonparametric test, because the control chart based on Wilcoxon–Mann–Whitney statistics had difficulty resolving the scale shift. Murakami and Matsuki (2010) presented the importance of univariate nonparametric control charts, with respect to the Mood test; the control limits of the Mood test were computed using the saddlepoint approximation. Abd-Elfattah (2013) also considered the double saddlepoint approximation to calculate the mid p-value of the Mood test. Recent progress in computerized measurement technology has permitted the accumulation of multivariate high-dimensional data. The analysis of high-dimensional data is particularly important in studies of finance data, functional data, image data, and other data of this type. Principal component analysis is an extremely useful and important technique in multivariate analysis from both practical and theoretical viewpoints. However, it is very difficult to obtain the precise distribution of eigenvalues of a covariance matrix for a non-normal population. Murakami, Hino, and Tsukada (2007) proposed a procedure to test the equality of eigenvalues that used principal component scores by applying the Mood test. Additionally, unreplicated two-level fractional factorial designs are often used in industry to study the effects of factors on the mean of a response. For example, McGrath and Lin (2002) investigated the power of the Mood test for unreplicated two-level fractional factorial designs.

Interestingly, even with the high level of performance of the Mood test and its popularity over the course of many years, new versions of the modified Mood test after Tamura (1963) are very rarely discussed in the statistical literature; whereas several types of nonparametric test modifications, such as the Wilcoxon rank sum test and Mann–Whitney test have been proposed. Tamura's modification of the Mood test, namely M_p, raises the ARE for the scale parameter by providing flexibility with respect to the power of the linear rank sum weight. Tamura (1963) focused on the ARE only, due to the difficulty associated with simulation of specific topics at that time. Note that the original Mood test is a special case of the modified Mood test with $p = 2$, that is, M_2. It should be mentioned that utilization of the modified Mood test can be very limited, depending on the details available with respect to its relevant characteristics and performance improvements.

Herein, we have attempted to overcome the limitations associated with the Mood test. First, we propose several approximation methods to obtain its distribution and critical values, and then show numerical examples for the optimal degree p of the weights of the linear rank sum and statistical power in various cases of the underlying sampling distributions. Calculation of the exact critical value of the M_p test is an important task in testing hypotheses. However, it is challenging to evaluate the exact critical value of the M_p test when the sample size increases. To illustrate this problem, consider the case of $n_1 = n_2 = 20$. The number of all possible combinations to calculate the exact critical value of the test statistic is ${}_{40}C_{20} = 137,846,528,820$. For most modern computing environments, this computation can be expensive. Many linear rank tests satisfy the asymptotic normality theorem only when $\min(n_1, n_2) \to \infty$; the normal approximation does not provide a good approximation for small or moderate finite sample sizes. Under these circumstances, the approximation method must be modified to estimate the exact critical values and the density or distribution function under finite sample sizes. Second, we are interested in determining the optimal value p of the M_p test via maximizing the asymptotic efficiency in various cases to present typical distributional characteristics, such as tail behavior, support, and modality. An analysis of the optimal value of p is particularly important in a wide array of applications, such as finance, functional data, and image data, because the modified Mood test is well designed to adapt the distributional characteristics arising in certain circumstances. For example, many financial data are likely to exhibit behavior similar to that of heavy-tailed distributions, and, in such cases, the smaller value of p for the M_p test is more desirable than the original Mood test. Finally, the power of the nonparametric test is discussed as the selection criterion of the most appropriate test. A numerical comparison of the statistical power of the M_p test under the assumption of various distributions was performed.

The organization of this paper is as follows. We first introduce Tamura's modification of the Mood test in Section 2. The results, including the approximation methods, optimal p, and asymptotic efficiency, are shown in Section 3; the liaison between Bayesian methodology and the asymptotic efficiency of Tamura's modification of the Mood test is also described in Section 3. In Section 4, statistical power comparisons were examined via simulation; additionally, the M_p test was used to analyze data from the study by Davis and Lawrance (1989). Section 5 provides a discussion and conclusions.

2. The modified Mood test

In this section, we introduce the modified Mood test, namely M_p, suggested by Tamura (1963). Let $X = (X_1, \dots, X_{n_1})$ and $Y = (Y_1, \dots, Y_{n_2})$ be two random samples of size n_1 and n_2 independent observations, each of which has a distribution described as F_1 and F_2, respectively. We are interested in the following hypotheses:

$$H_0 : F_1(x) = F_2(x) \quad \text{against } H_1 : F_1(x) = F_2(x/\theta), \quad \theta \neq 1,\ \theta > 0.$$

If the population medians are known, then the sample observations should be adjusted by the median of X and Y, namely Med_X and Med_Y, respectively, as follows:

$$X_i' = X_i - \text{Med}_X, \quad i = 1, \dots, n_1 \quad \text{and} \quad Y_j' = Y_j - \text{Med}_Y, \quad j = 1, \dots, n_2.$$