ABOUT ROBUSTNESS AND POWER OF VARIANCE HOMOGENEITY TESTS

A.A. Gorbunova, B.Yu. Lemeshko, S.B. Lemeshko
Novosibirsk State Technical University
Novosibirsk, Russia
e-mail: gorbunova.alisa@gmail.com

Abstract

Classical tests for homogeneity of variances (Fisher's, Bartlett's, Cochran's, Hartley's and Levene's tests) and nonparametric tests (Ansari-Bradley's, Mood's, Siegel-Tukey's, Capon's and Klotz's tests) have been considered. Distributions of classical tests statistics have been investigated under violation of assumption that samples are normally distributed. The comparative analysis of power of classical tests with power of nonparametric tests has been carried out. Tables of percentage points for Cochran's test have been made for distributions which are different from normal. Software, that allows us to correctly apply tests, has been developed.

1 Introduction

Testing for samples homogeneity is frequently of interest in a number of research areas. The question can be about homogeneity of samples distributions, population means or variances. Of course, conclusions in full measure can be made in the first case. However, researcher can be interested in possible deviations in the sample mean values or differences in variances of measurements.

One of the basic assumptions to formulate classical tests for comparing variances is normal distribution of samples. It is well known, that classical tests are very sensitive to departures from normality. Therefore, the application of classical criteria always involves the question of how valid the obtained results are in this particular situation.

This work continues the research of criteria stability for testing hypotheses about the equality of variances [1, 2]. Classical Bartlett's, Cochran's, Fisher's, Hartley's, Levene's tests are compared, nonparametric (rank) Ansari-Bradley's, Mood's, Siegel-Tukey's, Capon's and Klotz's tests are considered [3].

The purpose of our study was to:

- investigate distributions of the statistics for several tests when samples are not normally distributed;
- make a comparative analysis of the criteria power for concrete competing hypotheses;
- give a possibility to apply classical tests when the normality assumption may not be true.

A null hypothesis of equal variances for m samples is given by

$$H_0 : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_m^2 \quad (1)$$

and the alternative hypothesis is

$$H_1 : \sigma_{i_1}^2 \neq \sigma_{i_2}^2 \quad (2)$$

where the inequality holds at least for one pair of subscripts i_1, i_2.

Statistical simulation methods and developed software were used to investigate statistics distributions, to calculate percentage points and to estimate tests power for different competing hypotheses. Each test statistic was computed $N = 10^6$ times. In this case an absolute value of the difference between the true law of statistics distribution and a simulated empirical distribution does not exceed 10^{-3}.

Distributions of the statistics were investigated using various distributions, in particular, in the case when simulated samples are in the family of distributions with the density

$$De(\theta_0) = f(x; \theta_0, \theta_1, \theta_2) = \frac{\theta_0}{2\theta_1 \Gamma(1/\theta_0)} exp\left(-\left(\frac{|x - \theta_2|}{\theta_1} \right)^{\theta_0} \right) \quad (3)$$

using different values of the shape parameter θ_0. This family can be a good model for error distributions of many measuring systems. Special cases of the family $De(\theta_0)$ are the Laplace ($\theta_0 = 1$) and the normal ($\theta_0 = 2$) distributions. This family makes it possible to set various symmetric distributions that differ from the normal distribution. That is a smaller value of the shape parameter θ_0 leads to a "heavier" tails of the distribution.

In the comparative analysis of the tests power we consider the competing hypotheses of the form $H_1 : \sigma_m = d\sigma_0$ ($d \neq 1$). That is, an alternative hypothesis presents the situation when $m-1$ samples are from the population with the variance $\sigma = \sigma_0$, while one of the samples, for example, with the number m has some different variance. A null hypothesis is $H_0 : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_m^2 = \sigma_0^2$.

2 Comparative analysis of power

At the given probability of a type I error α (to reject the null hypothesis when it is true) it is possible to judge about the advantages of the test by the value of power $1-\beta$, where β - probability of type II error (not to reject the null hypothesis when alternative is true). In [4] it is definitely said that Cochran's test power is lower in comparison with Bartlett's test. In [1] it was shown that Cochran's test has a greater power by the example of testing hypothesis about variances homogeneity for *five* samples.

The study of power of Bartlett's, Cochran's, Hartley's, Fisher's and Levene's tests for several competing hypotheses $H_1 : \sigma_2 = d\sigma_1$ ($d \neq 1$) has shown that Bartlett's, Cochran's, Hartley's and Fisher's tests have equal power for two normal samples and Levene's test power is much less in this case.

As for non-normal distributions, for example, family of distributions with density (3), Bartlett's, Cochran's, Hartley's and Fisher's tests remain equal in power, and

Levene's test power is also much less. However, for heavy-tailed distributions (for example, the Laplace distribution) Levene's test is more powerful than the others.

Bartlett's, Cochran's, Hartley's and Levene's tests can be applied when number of samples $m > 2$. In such situations the power of these tests is different. If $m > 2$ and normality assumption is true, these tests can be ordered according to the decrease of power in the following way:

$$Cochran's \succ Bartlett's \succ Hartley's \succ Levene's.$$

The preference order also remains in case of violation of a normality assumption. When samples are from heavy-tailed distributions, this preference order changes. For example, in the case of the Laplace distribution Levene's test has a greater power.

The study of the nonparametric criteria power has shown that Mood's test power is the highest. And other nonparametric tests, as Siegel-Tukey's, Ansari-Bradley's, Capon's and Klotz's have practically equal power.

Figure 1 shows graphs of criteria power for competing hypotheses $H_1^1 : \sigma_2 = 1.1\sigma_1$ and $H_1^2 : \sigma_2 = 1.5\sigma_1$ depending on sample size n_i when $\alpha = 0.1$ and samples are normally distributed. Advantage in power of Cochran's test is rather significant in comparison with Mood's test - most powerful among nonparametric criteria. Let's remind that Bartlett's, Cochran's, Hartley's and Fisher's tests have equal power for two samples.

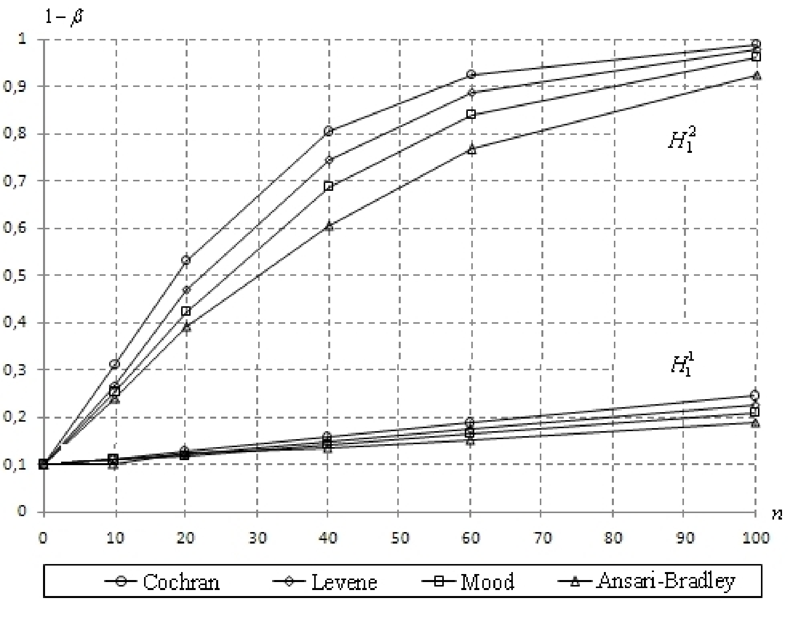

Figure 1: Power of tests for competing hypotheses H_1^1 and H_1^2 depending on sample size n when $\alpha = 0.1$ and samples are normally distributed

The main and valid reason for using nonparametric tests is based on the fact that these test statistics are distribution-free. But this is true if both samples are from the same population. If samples are not identically distributed, *nonparametric tests depend on both sample laws and even the order in which these laws are used.*

3 Cochran's test for non-normal distributions

Classical tests have a great advantage in power over nonparametric tests. This advantage remains when samples are not normally distributed. Therefore, there is every reason to study distributions of *classical* criteria for testing variances homogeneity. To study distributions means to develop distribution models or tables of percentage points. It should be done for non-normal distributions mostly used in practice. Among the tests studied Cochran's test seems to be the most suitable for this purpose.

Tables of upper percentage points (1%, 5%, 10%) for Cochran's test were made using statistical simulation for the number of samples $m = 2 \div 5$ when simulated samples were taken from an exponential family of distributions (3) with shape parameter $\theta_0 = 1, 2, 3, 4, 5$. The results obtained can be used in situations when distribution from an exponential family (3) with an appropriate parameter θ_0 is a good model for the observed variables. Computed percentage points improve results presented in [1] and expand possibilities to apply Cochran's test.

We have developed software that allows us to correctly apply tests for comparing variances when samples are from any distributions. We can choose any distribution from the list and simulate a distribution of the statistic to obtain percentage points. And then use these points to test the hypothesis of equal variances.

4 Acknowledgments

This research was supported by the Russian Foundation for Basic Research (project no. 09-01-00056a), by the Federal Agency for Education within the framework of the analytical domestic target program "Development of the scientific potential of higher schools" and federal target program of the Ministry of Education and Science of the Russian Federation "Scientific and scientific-pedagogical personnel of innovative Russia".

References

[1] Lemeshko B.Yu., Mirkin E.P. (2004). Bartlett and Cochran tests in measurements with probability laws different from normal. *Measurement Techniques*. Vol. **47**, pp. 960-968.

[2] Lemeshko B.Yu., Ponomarenko V.M. (2006). Research of statistics distributions used to test hypotheses about the equality of variances in case of laws of errors different from normal (in Russian). *Scientific Journal of NSTU*. Vol. **2(23)**, pp. 21-33.

[3] Kobzar A.I. (2006). *Applied mathematical statistics. For engineers and scientists* (in Russian). FIZMATLIT Publishing House, Moscow.

[4] Bolshev L.N., Smirnov N.V. (1983). *Tables of Mathematical Statistics* (in Russian). Nauka, Moscow.