Mathematical Researches
Print ISSN: 2588-2546
Online ISSN: 2588-2554
Homepage: https://jemr.khu.ac.ir/
Kharazmi University

A Test for Homogeneity of Variances Using the Jackknife Empirical Likelihood Approach and Its Comparison with Several Common Tests

Nabaz Esmailzadeh [1,2], **Parisa Hoseini** [2]

1. Corresponding Author, Department of Statistics, Faculty of Sciences, Razi University, Kermanshah, Iran. E-mail: n.esmailzadeh@razi.ac.ir
2. Department of Statistics, Faculty of Sciences, University of Kurdistan, Sanandaj, Iran. E-mail: hosiniparisa7@gmail.com

Article Info

Article type:
Research Article

Article history:
Received: 18 September 2024
Received in revised form: 11 November 2024
Accepted: 20 November 2024
Published online: 28 February 2025

Keywords:
Homogeneity of variances,
Jackknife,
Empirical likelihood,
Permutation test.

ABSTRACT

Introduction

The homogeneity of variances test is a prerequisite for many statistical methods such as analysis of variance and regression methods. This test is often used to identify potential differences and variations between traits in different populations. Given the interest and importance of researchers in increasing uniformity in quality control of production processes, tests for equality of variances have gained attention. Additionally, in many research areas, examining the equality of variances among multiple groups is significant. For instance, in various scientific fields such as biology, sociology, agriculture, and economics, studying the dispersion of traits in different communities is noteworthy. Therefore, many tests in this area have been studied and incorporated into numerous statistical software packages.

In this article, a recently introduced likelihood-based jackknife test is examined alongside common tests such as Levene's and Bartlett's tests, as well as two tests by James and Alexander-Govern, in terms of their ability to maintain the first type error rate and test power for several symmetric and asymmetric distributions. The permutation versions of these tests were also investigated.

We examine the Jackknife Empirical Likelihood ratio (JEL) test for cases where sample sizes are unequal, and we study the power and Type I error rate of the test through simulation studies. Additionally, for small sample sizes, the proposed JEL test by Sang [24] does not have high power. In this study, we investigate the permutation version of the test to assess the hypothesis of improved test performance.

Material and Methods

To examine the Type I error rate and the power of the jackknife empirical likelihood ratio test to compare these quantities with the four tests of Levene, Bartlett, James, and Alexander and Govan discussed in Section 2, we will analyze samples from three distributions with different characteristics. The simulations are based on 1000 samples from each distribution and 1200 permutations for the permutation test. The simulations were conducted using coding in R software. We consider the significance level of the test to be 0.05.

We have considered different sample sizes from the standard normal distribution (symmetric), the student’s t-distribution with four degrees of freedom (symmetric, long tail, and short kurtosis), and the chi-squared distribution with four degrees of freedom (skewed, long tail, and high kurtosis). The sample sizes for the two-group population are (3,3), (4,4), (8,8), (10,10), (11,11), (3,4), (4,5), (5,6), (7,8), (11,12); for the three-group population, they are (3,3,3), (4,4,4), (8,8,8), (10,10,10), (11,11,11), (3,4,5), (5,6,7), (7,8,9), (9,10,11), (11,12,13); and for the four-group population, they are (3,3,3,3), (4,4,4,4), (8,8,8,8), (10,10,10,10), (11,11,11,11), (3,4,5,9), (5,6,7,10), (7,8,9,12), (9,10,11,14).

Results and discussion

The simulation results related to the Type I error rate for k=2, 3, and 4 are presented in box plots in some figures. According to the charts, the JEL test exhibits a high error rate across all distributions; however, its permutation version maintains the error rate effectively. This behavior is also observed for the Bartlett test, which preserves the error rate only for the normal distribution. Other tests, when using the critical region of the approximate distribution, are conservative and estimate the error rate to be less than the nominal value of 0.05. The performance of these tests improves with the application of the permutation method, bringing the estimated error rate closer to the nominal value.

The simulation results related to the power of the test for the 2, 3, and 4 group scenarios are presented in some box plots. According to the charts, the behavior of the JEL and Bartlett tests differs from that of other tests. The JEL test has a high average power across all distributions, although its variability is quite significant. As the number of groups increases, the power of the JEL test also increases, and its variability decreases. The power of the Bartlett test behaves uniformly across all distributions and group numbers. Although it has an acceptable average power, its variability is high. Other tests exhibit nearly similar performance in both normal and permutation scenarios. In the permutation test, the variability of the JEL and Bartlett tests decreases, and their average power also diminishes.

It should be noted that simulations for large sample sizes for both equal and unequal populations were also conducted. In this case, the tests generally perform well, yielding a power close to one.

Conclusion

The study was based on simulations in various scenarios with equal and unequal sample sizes for cases where the number of groups is 2, 3, and 4. The results indicate that although there is no single test that performs best in all scenarios, the performance of the tests significantly improves in their permutation versions. To evaluate the performance of the tests in real-world situations, the tests were applied to two real datasets, and the results are presented.

How to cite: Esmailzadeh, N., Hoseini, P. (2024). A Test for Homogeneity of Variances Using the Jackknife Empirical Likelihood Approach and its Comparison with Several Common Tests, *Mathematical Researches*, **10** (4), 13 – 32.

 Publisher: Kharazmi University

شاپا چاپی: ۲۰۴۶-۲۵۸۸
شاپا الکترونیکی: ۲۵۵۴-۲۵۸۸

پژوهش‌های ریاضی

دانشگاه خوارزمی

Homepage: https://mmr.khu.ac.ir/

آزمونی برای همگنی واریانس‌ها با رویکرد درستنمایی تجربی جک نایف و مقایسه آن با چند آزمون رایج

نبز اسمعیل زاده [۱و۲] ✉، پریسا حسینی [۳]

۱. نویسنده مسئول، گروه آمار، دانشکده علوم پایه، دانشگاه رازی، کرمانشاه، ایران. رایانامه: n.esmailzadeh@razi.ac.ir
۲. گروه آمار، دانشکده علوم پایه، دانشگاه کردستان، سنندج، ایران. رایانامه: hosiniparisa7@gmail.com

اطلاعات مقاله

نوع مقاله: مقاله پژوهشی

تاریخ دریافت: ۱۴۰۳/۶/۲۸
تاریخ بازنگری: ۱۴۰۳/۸/۲۱
تاریخ پذیرش: ۱۴۰۳/۸/۳۰
تاریخ انتشار: ۱۴۰۳/۱۲/۱۰

واژه‌های کلیدی:
همگنی واریانس‌ها،
جک‌نایف،
درست نمایی تجربی،
آزمون جایگشتی.

چکیده

آزمون همگنی واریانس‌ها پیش‌فرض بسیاری از روش‌های آماری است. همچنین خیلی اوقات این آزمون به منظور شناسایی اختلافات و تفاوت های ممکن بین صفات در جوامع مختلف مورد توجه است. در این مقاله، آزمونی با رویکرد درستنمایی جک‌نایف که اخیرا معرفی شده است، همراه با آزمون‌های رایج لون و بارتلت و همچنین دو آزمون جیمز و الکساندر- گاورن از نظر توانایی حفظ نرخ خطای نوع اول و توان آزمون برای چند توزیع متقارن و نامتقارن مورد بررسی قرار گرفتند. همچنین نسخه جایگشتی این آزمون‌ها نیز بررسی شد. مطالعه براساس شبیه سازی در حالات مختلف حجم نمونه های برابر و نابرابر برای وقتی که تعداد گروه ها ۲، ۳ و ۴ است، انجام شد. نتایج نشان می‌دهد اگرچه آزمونی که در همه حالات بهترین باشد وجود ندارد، با وجود این، عملکرد آزمون‌ها در نسخه جایگشتی بطور چشمگیری بهبود پیدا می‌کند. به منظور بررسی عملکرد آزمون‌ها در دنیا واقعی، آزمون‌ها روی دو مجموعه داده واقعی اعمال شده و نتایج ارایه شده است.

استناد: اسمعیل‌زاده، نبز؛ حسینی، پریسا (۱۴۰۳). آزمونی برای همگنی واریانس‌ها با رویکرد درستنمایی تجربی جک نایف و مقایسه آن با چند آزمون رایج. *پژوهش‌های ریاضی*، **۱۰** (۴)، ۱۳ – ۳۲.

ناشر: دانشگاه خوارزمی

مقدمه

تحلیل واریانس یکی از فنون پرکاربرد در تحلیل آماری است. یکی از مفروضات زیربنایی در این تحلیل، آزمون همگنی واریانس‌ها است. یعنی فرض برابری واریانس‌های k جامعه $(\mathrm{k} \geq ۲)$ در مقابل اینکه حداقل واریانس دو جامعه از آنها برابر نباشد.

با توجه به علاقه و اهمیت محققان به افزایش یکنواختی در کنترل کیفیت فرایندهای تولید و رسیدن به این مهم آزمون‌های برابری واریانس‌ها مورد علاقه قرار گرفته‌اند. همچنین در بسیاری از حوزه‌های تحقیقاتی بررسی برابری واریانس‌های چند گروه مورد مطالعه حائز اهمیت است. به طور مثال در زمینه های علمی مختلف از جمله: زیست شناسی، جامعه شناسی، کشاورزی و اقتصاد بررسی پراکنش صفات جوامع مختلف مورد مطالعه، قابل توجه است. لذا آزمون‌های بسیاری در این زمینه مورد مطالعه قرار گرفته‌اند و در بسیاری از نرم افزارهای آماری گنجانده شده‌اند (گستویرت و همکاران [۱۱]).

همچنین آزمون برابری واریانس‌ها یک آزمون مقدماتی برای بسیاری از روش‌های آمارکلاسیک مانند آزمون تی- استیودنت و تحلیل رگرسیون می‌باشد. بنابراین بررسی برابری واریانس‌ها در جوامع مختلف از اهمیت زیادی برخوردار است و از قدیمی‌ترین مسائل مورد بررسی در استنباط آماری است.

آزمون برابری واریانس‌ها در زمینه‌های کاربردی مختلفی مورد توجه می‌باشد. به همین دلیل تعداد قابل توجهی مقاله در مورد آزمون‌های همگنی واریانس‌های چند جامعه وجود دارند، برخی از این آزمون‌ها تحت فرض نرمال بودن بررسی می‌شود که از حساسیت شدیدی نسبت به نقض این فرض رنج می‌برند مانند آزمون‌های کلاسیک F فیشر و آزمون بارتلت [۲] بارتلت و کندال [۳] با ارائه آزمونی با تکیه به این حقیقت که $\ln \mathrm{s}^2$ تقریباً نرمال است (s^2 واریانس نمونه‌ای می‌باشد) با استفاده از جدول‌هایی سعی کردند به سادگی محاسبات کمک کنند. هارتلی [۱۲] آزمونی معروف به $F - max$ را ارائه کرد که تنها یک تبدیل نمایی از آزمون کندال و بارتلت بود. مزیت این آزمون داشتن جدول‌های دقیق و قابل دسترس برای اندازه نمونه‌های برابر است.

آزمونی دیگر برای مقایسه واریانس‌ها آزمون لون [۱۸] با استفاده از تحلیل واریانس آزمون F بر روی انحراف مطلق مشاهدات از میانگین گروه خود، $Z_{ij} = |Y_{ij} - \overline{Y_i}|$ است که Y_{ij} نمونه‌های تصادفی با اندازه n_i و میانگین $\overline{Y_i}$ هستند. اصلاحات مختلفی از آزمون لون صورت گرفته است. لوه [۲۰] استواری آزمون لون را برای نمونه‌های با حجم کم بهبود بخشید. یتنوسومارتو و اونیل [۲۷] آزمون لون را با تغییر درجات آزادی آزمون F اصلاح کردند. شارما و کیبریا [۲۵] نیز با هدف مقایسه‌ی آزمون‌های مختلف از لحاظ استواری نرخ خطا و توان، متغیرهای انحراف مطلق مشاهدات از میانگین گروه خود Z_{ij}، ${Z_{ij}}^2$ و ${Z_{ij}}^{1/2}$ را با استفاده از میانگین، میانه و میانگین پیراسته‌ی ده درصد مورد بررسی قرار دادند و به این نتیجه رسیدند که در آزمون لون استفاده از میانه نتیجه بهتری می‌دهد. آزمونی که معمولاً در حالت غیرنرمال از نظر توان و استواری مناسب است، آزمون لون با استفاده از میانه‌ی نمونه‌ایی به عنوان یک براورد پارامتر مکان است [۲۳].

باکس [۵] با توسعه مفهومی که توسط کندال و بارتلت پیشنهاد شده بود، روشی بسیار جالب برای بدست آوردن آزمونی استوارتر برای واریانس که شامل استفاده از آماره‌ی F تحلیل واریانس یک‌طرفه است، ارائه داد که به وضوح استوار است. اگر چه در مطالعات بعدی استواری این روش تایید شده است اما نشان داده شده است که توان این آزمون در مقایسه با آزمون‌های

دیگری که به همین اندازه استوار هستند، کمتر است(لوی [۱۹] ،لایارد [۱۶] و گرتساید [۱۰]). براون و فورسایت [۶] با استفاده از میانه و میانگین پیراسته‌ی ده درصد به جای استفاده از میانگین، آزمون قوی‌تری برای مقایسه دو واریانس پیشنهاد کردند. چندین آزمون مهم دیگر نیز برای بررسی فرض همگنی واریانس‌های چند جامعه مانند آزمون‌های ویلچ، شومیکر، فلینگر- کیلین، جیمز و آزمون الکساندر و گاورن بررسی شده‌اند [۲۶، ۲۳].

سانگ [۲۴] براساس رویکرد درستنمایی تجربی جک‌نایف (JEL) آزمونی ناپارامتری برای بررسی همگنی واریانس‌های چند جامعه مطالعه کرد. این آزمون قابلیت اطمینان رویکرد درستنمایی تجربی و کارایی محاسباتی روش JEL را به ارث می‌برد. سانگ حالتی که حجم نمونه‌ها برابر باشند را مورد مطالعه قرار داده و توان و میزان خطای نوع اول آزمون مورد نظر را در این حالت با انجام مطالعات شبیه‌سازی به دست آورده و با چند آزمون قبلی مقایسه کرده است. با تلاش‌های صورت گرفته برای توسعه‌ی آزمون‌های برابری واریانس‌ها درمی‌یابیم که ویژگی‌های کلیدی یک آزمون مانند ویژگی‌های مجانبی جذاب، توان بالا و خطای نوع اول کنترل شده حایز اهمیت هستند.

در این مقاله رویکرد JEL را برای حالتی که حجم نمونه‌ها برابر نیستند، مورد بررسی و از طریق مطالعات شبیه‌سازی میزان توان و خطای نوع اول آزمون را مورد مطالعه قرار دهیم. همچنین برای حجم نمونه‌های کوچک آزمون پیشنهادی سانگ [۲۴] دارای توان بالایی نیست. در این مطالعه نسخه جایگشتی آزمون را به منظور بررسی فرض بهبود عملکرد آزمون بررسی می‌کنیم.

همچنین، استواری (توانایی در حفظ نرخ خطای نوع اول) آزمون JEL با آزمون‌های رایج لون و بارتلت برای بررسی همگنی واریانس‌های چند جامعه و همچنین آزمون‌های جیمز و الکساندر و گاورن که به نظر می‌رسد برای حجم نمونه‌های نابرابر کارایی و استواری خوبی دارند [۹]، با استفاده از مطالعات شبیه‌سازی بررسی و مقایسه می‌کنیم.

مقاله به این صورت سازماندهی شده است که در بخش ۲ آزمون‌های مورد مطالعه مرور می‌شوند. در بخش ۳ طراحی شبیه‌سازی و نتایج آن بیان شده است. اعمال آزمون‌ها روی دو مجموعه داده واقعی در بخش ۴ انجام می‌شود.

۱. آزمون‌های مورد مطالعه

تا به حال مطالعات زیاد و جامعی در زمینه‌ی بررسی همگنی واریانس‌ها در جوامع مختلف صورت گرفته است. برای دستیابی به آزمون‌های استوار، آزمون‌های مبتنی بر رهیافت‌های تعدیل کشیدگی، آزمون‌های مبتنی بر متغیرهای تبدیل یافته و آزمون‌های مبتنی بر روش‌های باز نمونه‌گیری ارائه شده است(بوس و براونی [۴]). در این بخش آزمون رویکرد درستنمایی تجربی جک‌نایف و آزمون‌های رایج لون و بارتلت و همچنین آزمون‌های جمیز و الکساندر و گاورن را مررو می‌کنیم.

فرض کنید k نمونه‌ی مستقل $Y_{i1},\cdots,Y_{in_i}$ $\quad i=1,\cdots,k;$ داریم که $E(Y_{ij})=\mu_i$ و $Var(Y_{ij})=\sigma_i^2$ برای $j=1,\cdots,n_i$. می‌خواهیم آزمون فرض

$$H_0\colon \sigma_1^2=\cdots=\sigma_k^2 \quad \text{vs.} \quad H_a\colon \sigma_i^2\neq\sigma_j^2$$

را برای حداقل یک ز وج (i,j) که $1\leq i<j\leq k$ انجام دهیم.

۱.۱ آزمون لون: آزمون لون یک تحلیل واریانس یک طرفه روی قدرمطلق انحراف از میانگین Y_{ij}هاست. میلر [۲۱] نشان داد اگر به جای میانگین نمونه‌ی گروه $i-$ام میانه را استفاده کنیم، می‌توان به طور مجانبی آن را برای جوامع نامتقارن نیز به کار برد. لذا، از آزمون لون که یک تحلیل واریانس یک‌طرفه روی $Z_{ij} = \left|Y_{ij} - m_i\right|$ است، استفاده می‌کنیم. در اینجا m_i میانه گروه $i-$ام است. آماره آزمون به صورت زیر است:

$$L = \frac{(N-k)\sum_{i=1}^{k}(\overline{Z_i} - \overline{Z})^2 n_i}{(k-1)\sum_{i=1}^{k}\sum_{j=1}^{n_i}(Z_{ij} - \overline{Z_i})^2},$$

که در آن $N = \sum_{i=1}^{k} n_i$، $\overline{Z_i} = \frac{\sum_{j=1}^{n_i} Z_{ij}}{n_i}$ و $\overline{Z} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i} Z_{ij}}{N}$.

مقدار مشاهده شده این آماره را با مقدار چندک بالایی α ام توزیع F با درجات آزادی $k-1$ و $N-k$ مقایسه می‌کنیم. اگر مقدار L بزرگتر از چندک توزیع F باشد، فرض صفر رد می‌شود.

مطالعات اخیر رویکردهای متفاوتی برای بهبود آزمون لون با به کار بردن مقیاس‌های استوار مکان به جای میانگین را نشان می‌دهند (پارا فروتوس [۲۳]). برون و فورسیت [۶] با انجام شبیه سازی نشان داده‌اند که سطح آزمون میانگین لون برای داده‌های دارای توزیع متقارن نزدیک سطح معنی‌داری اسمی است. اما برای داده‌ها با توزیع‌های نامتقارن اینگونه نیست. آزمون‌های میانگین و میانه‌ی لون تقریبا سطح معنی‌داری اسمی را برای توزیع‌های متقارن حفظ می‌کنند. مطالعات شبیه سازی کانوور و همکارانش [۷، ۸] نشان داده است که آزمون لون از لحاظ توان و استواری سطح، یکی از بهترین آزمون‌ها در شرایط غیرنرمال است.

۲.۱ آزمون بارتلت : آزمون بارتلت آزمونی مبتنی بر رهیافت تعدیل کشیدگی است. آزمون بارتلت اصلاحی از آزمون نسبت درستنمایی مطرح شده توسط نیمن و پیرسون [۲۲] است. مهمترین ضعف این آزمون غیراستواری در برابر غیرنرمال بودن است.

آماره‌ی آزمون بارتلت به صورت زیر است:

$$B = \frac{(N-k)\ln S_a^2 - \sum_{i=1}^{k}(n_i - 1)\ln S_i^2}{1 + \frac{1}{3(k-1)}\left(\sum_{i=1}^{k}\frac{1}{n_i - 1} - \frac{1}{N-k}\right)}$$

که در آن $S_i^2 = \frac{\sum_{j=1}^{n_i}(Y_{ij} - \overline{Y_i})^2}{n_i - 1}$ و $S_a^2 = \frac{\sum_{i=1}^{k}(n_i - 1)S_i^2}{N-k}$

تحت فرض صفر، B به طور تقریبی یک متغیر کای‌دو با $k-1$ درجه آزادی است. وقتی B بزرگتر از چندک $1-\alpha$ از توزیع کای‌دو با $k-1$ درجه آزادی باشد، آزمون بارتلت فرض H_0 را رد می‌کند.

۳.۱. آزمون جیمز: آزمون جیمز آزمونی مبتنی بر رهیافت متغیرهای تبدیل یافته می‌باشد که توسط جیمز [۱۳] پیشنهاد شده است.

فرض کنید برای $Z_{ij}=|Y_{ij}-\overline{Y_i}|$ ،$j=1,\cdots,n_i;\quad i=1,\cdots,k;$ آماره‌ی آزمون جیمز به صورت زیر است:

$$J=\sum_{i=1}^{k}w_i(\overline{Z_i}-\overline{Z}^*)^2,$$

که در آن $w_i=\frac{n_i}{S_{z,i}^2}$، $\overline{Z_i}=\frac{\sum_{j=1}^{n_i}Z_{ij}}{n_i}$، $\overline{Z}^*=\frac{\sum_{i=1}^{k}w_i\overline{Z_i}}{W}$، $S_{z,i}^2=\frac{\sum_{j=1}^{n_i}(Z_{ij}-\overline{Z_i})^2}{n_i-1}$ و $W=\sum_{i=1}^{k}w_i$.

زمانی که اندازه نمونه‌ها کوچک و یا حتی نسبتا بزرگ است تقریب کای‌دو با $k-1$ درجه آزادی برای آماره J رضایت بخش نیست. بنابراین جیمز دو روش را برای اصلاح مقدار بحرانی پیشنهاد کرده است. آزمون مرتبه‌ی اول جیمز نرخ خطای نوع اول را در ناهمگنی واریانس‌ها برای اندازه نمونه‌های کوچک کنترل نمی‌کند و اغلب اوقات فرض صفر را رد می‌کند. یکی از نقص‌های عمده آزمون مرتبه‌ی دوم جیمز پیچیدگی محاسباتی آن است اما مطالعات شبیه‌سازی مختلف نشان می‌دهد که این آزمون تحت شرایط واقعی صحیح‌ترین روش به نظر می‌رسد. بنابراین فرایند مرتبه‌ی دوم جیمز به صورت گسترده‌ای توصیه می‌شود و به صورت زیر است.

$$\begin{aligned}
h_2(\alpha)=\;&c(\alpha)+\frac{1}{2}(3X_4+X_2)\sum_{i=1}^{k}\frac{(1-\frac{w_i}{W})^2}{n_i-1}+\frac{1}{16}(3X_4+X_2)^2(1-\frac{k-3}{c(\alpha)})(\sum_{i=1}^{k}\frac{(1-\frac{w_i}{W})^2}{n_i-1})^2\\
&+\frac{1}{2}(3X_4+X_2)[(8R_{23}-10R_{22}+4R_{21}-6R_{12}^2+8R_{12}R_{11}-4R_{11}^2\\
&+(2R_{23}-4R_{22}+2R_{21}-2R_{12}^2+4R_{12}R_{11}-2R_{11}^2)(X_2-1)+\frac{1}{4}(-R_{12}^2\\
&+4R_{12}R_{11}-2R_{12}R_{10}-4R_{11}^2+4R_{11}R_{10}-R_{10}^2)(3X_4-2X_2-1)\\
&+(R_{23}-3R_{22}+3R_{21}-R_{20})(5X_6+2X_4+X_2)\\
&+\frac{3}{16}(R_{12}^2-4R_{23}+6R_{22}-4R_{21}+R_{20})(35X_8-15X_6+9X_4+5X_2)\\
&+\frac{1}{16}(-2R_{22}+4R_{21}-R_{20}+2R_{12}R_{10}-4R_{11}R_{10}+R_{10}^2)(9X_8-3X_6-5X_4\\
&-X_2)+\frac{1}{4}(-R_{22}+R_{11}^2)(27X_8+3X_6+X_4+X_2)\\
&+\frac{1}{4}(R_{23}-R_{12}R_{11})(45X_8+9X_6+7X_4+3X_2)
\end{aligned}$$

$c(\alpha)$ نشان دهنده‌ی چندک $\alpha-$ ۱ توزیع χ_{k-1}^2 است و $X_{2s}=\frac{c(\alpha)^s}{(k-1)(k+1)(k+3)\cdots(k+2s-3)}$ و $R_{st}=\sum_{i=1}^{k}\frac{(\frac{w_i}{W})^t}{(n_i-1)^s}$.

فرض صفر رد می‌شود اگر $J>h_2(\alpha)$

۴.۱. آزمون الکساندر و گاورن: آزمون دیگری که در اینجا مطرح می‌کنیم آزمون الکساندر و گاورن [۱] است که مبتنی بر رهیافت متغیرهای تبدیل یافته می‌باشد. آماره این آزمون به صورت زیر است:

$$A = \sum_{i=1}^{k} g_i^2$$

که در آن $g_i = c_i + \frac{c_i^3+3c_i}{b_i} - \frac{4c_i^7+33c_i^5+240c_i^3+855c_i}{10b_i^2+8b_ic_i^4+1000b_i}$، $a_i = n_i - 1.5$، $b_i = 48a_i^2$، $t_i = \frac{(\overline{Z}_i-\overline{Z}^*)\sqrt{n_i}}{s_{z,i}}$ و $c_i = [a_i \ln(1+\frac{t_i^2}{n_i-1})]^{\frac{1}{2}}$. A تقریباً دارای توزیع χ^2_{k-1} است. هر گاه A بزرگتر از $\chi^2_{\alpha,k-1}$ باشد فرض صفر رد می‌شود. از آنجای که آزمون الکساندر و گاورن از لحاظ محاسباتی ساده‌تر و در شرایط تجربی روی هم‌رفته بهتر است، این آزمون را به عنوان بهترین جایگزین تحلیل واریانس با استفاده از آماره‌ی F تحت ناهمگنی واریانس‌ها توصیه می‌کنند [۲۳].

۵.۱. درستنمایی تجربی جک‌نایف برای همگنی واریانس‌ها : آزمون درستنمایی تجربی جک‌نایف (*JEL*) یک روش ناپارامتری مبتنی بر رهیافت باز نمونه‌گیری می‌باشد [۱۹]. این آزمون برای مقایسه همگنی واریانس با رویکرد درستنمایی تجربی اصلاح شده، توسط سانگ [۲۴]، مورد استفاده قرار گرفت. جک‌نایف یک روش بازنمونه‌گیری بدون جایگذاری است که برای برآورد مقدار اریبی و برآورد واریانس روش مناسب و مفیدی است. روش جک‌نایف توسط موریس کوئینل آمارشناس انگلیسی در (۱۹۴۹) معرفی شد ولی بعدها بردلی افرون دانشمند آمریکایی در سال (۱۹۷۹) آن را توسعه داد و روش بوت استرپ را معرفی کرد.

تابع درست‌نمایی تجربی از توزیع تجربی داده‌ها ساخته می‌شود. فرض کنید $x_1, \cdots, x_n$ مونه‌ی تصادفی مستقل از توزیعی نامعلوم باشد. تابع درستنمایی تجربی به صورت زیر تعریف می‌شود:

$$L(p_1, p_2, \dots, p_n) = \prod_{i=1}^{n} p_i$$

که در آن p_i احتمال تخصیص داده شده به مشاهده‌ی x_i است و در شرایط $p_i \geq 0, \quad \sum_{i=1}^{n} p_i = 1$ صدق می‌کنند. تحت قیود $\sum_{i=1}^{n} p_i = 1$ و $p_i \geq 0$ ، بیشترین مقدار تابع $L(p_1, p_2, \dots, p_n)$ به ازای $p_i = \frac{1}{n}$ رخ می دهد و تابع توزیع تجربی $F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I(X \leq x_i)$ ست. تابع نسبت درستنمایی تجربی به صورت زیر تعریف می‌شود:

$$R(p_1, p_2, \dots, p_n) = \frac{L(p_1, p_2, \dots, p_n)}{L(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n})} = \frac{\prod_{i=1}^{n} p_i}{\prod_{i=1}^{n} \frac{1}{n}} = \prod_{i=1}^{n} np_i.$$

اینک آماره‌ی U_{n_i} را به صورت زیر تعریف می‌کنیم:

$$U_{n_i} = \binom{n_i}{2}^{-1} \sum_{1 \leq j < j' \leq n_i} \frac{(Y_{ij} - Y_{ij'})^2}{2}$$

که یک U -آماره تک نمونه‌ای درجه ۲ با تابع هسته زیر می‌باشد:

$$h_i(y_{ij}, y_{ij'}) = \frac{(y_{ij} - y_{ij'})^2}{2}, i = 1, \cdots, k.$$

۶.۱. آزمون جایگشتی: آزمون‌های جایگشتی رویکردی ناپارامتری قدرتمندی برای آزمون فرضیه‌ها ارائه می‌دهند که مانند روش‌های پارامتری به فرضیات سخت‌گیرانه درباره توزیع داده‌ها وابسته نیستند. در این مقاله برای پنج آزمون ذکر شده فوق نسخه آزمون جایگشتی را براساس مراحل زیر انجام می‌دهیم.

۱- تحت این فرض که داده‌های گروه‌های مختلف از توزیع یکسانی هستند، داده‌ها را ادغام کرده و بصورت تصادفی به گروه‌های مختلف با حجم نمونه‌های مشخص گروه‌بندی می‌کنیم.

۲- آماره آزمون مورد نظر را براساس گروه‌بندی جدید محاسبه و آنرا را با مقدار آماره آزمون براساس داده‌های گروه‌های واقعی مقایسه می‌کنیم.

۳- مراحل ۱ و ۲ را به دفعات زیاد (در این مقاله ۱۲۰۰ بار) تکرار می‌کنیم. اینک P- مقدار را محاسبه و آن را با سطح معنی‌داری اسمی $(\alpha = 0.05)$ مقایسه می‌کنیم و براساس آن در مورد رد یا پذیرش فرض صفر تصمیم می‌گیریم.

۴- مراحل ۱ تا ۳ را برای هر نمونه شبیه‌سازی شده تکرار می‌کنیم (در این مقاله ۱۰۰۰ بار).

۵- نرخ خطای نوع اول را محاسبه می‌کنیم.

۲. مطالعات شبیه سازی برای مقایسه آزمون‌ها

برای بررسی نرخ خطای نوع اول و توان آزمون درستنمایی تجربی جک‌نایف و مقایسه این کمیت‌ها با چهار آزمون لون، بارتلت، جیمز و الگساندر و گاورن مطرح شده در بخش ۲، نمونه‌هایی را از سه توزیع با ویژگی‌های مختلف مورد بررسی قرار می‌دهیم. شبیه سازی ها را بر پایه‌ی هزار نمونه از هر توزیع و هزار و دویست جایگشت برای آزمون جایگشتی انجام داده‌ایم. شبیه‌سازی براساس کدنویسی در نرم‌افزار R انجام شده است و برحسب درخواست از نویسنده مسئول قابل ارسال است. سطح معنی‌داری آزمون را برابر ۰/۰۵ در نظر می‌گیریم.

اندازه نمونه‌های مختلف را از توزیع نرمال استاندارد (متقارن)، توزیع t استودنت با چهار درجه آزادی (متقارن، دنباله‌ی بلند و کشیدگی کوتاه) و توزیع کای دو با چهار درجه آزادی (چوله، دنباله‌ی بلند و کشیدگی بالا) در نظر گرفته‌ایم. حجم نمونه‌ها برای جامعه دو گروهی برابر (۳،۳)، (۴،۴)، (۸،۸)، (۱۰،۱۰)، (۱۱،۱۱)، (۳،۴)، (۴،۵)، (۵،۶)، (۷،۸)، (۱۱،۱۲)، برای جامعه سه گروهی برابر (۳،۳،۳)، (۴،۴،۴)، (۸،۸،۸)، (۱۰،۱۰،۱۰)، (۱۱،۱۱،۱۱)، (۳،۴،۵)، (۵،۶،۷)، (۷،۸،۹)، (۹،۱۰،۱۱)، (۱۱،۱۲،۱۳) و برای جامعه چهار گروهی برابر (۳،۳،۳،۳)، (۴،۴،۴،۴)، (۸،۸،۸،۸)، (۱۰،۱۰،۱۰،۱۰)، (۱۱،۱۱،۱۱،۱۱)، (۳،۴،۵،۹)، (۵،۶،۷،۱۰)، (۷،۸،۹،۱۲)، (۹،۱۰،۱۱،۱۴) انتخاب کردیم. به منظور بررسی توان آزمون‌ها، فرض برابری واریانس‌ها را در مقابل فرض‌های

$$(\sigma_1^2, \sigma_2^2) = (۱,۱/۵)$$
$$(\sigma_1^2, \sigma_2^2, \sigma_3^2) = (۱,۱/۵,۲)$$
$$(\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2) = (۱,۱/۵,۲,۳)$$

به ترتیب برای k=2، k=3 و k=4 انجام می‌دهیم.

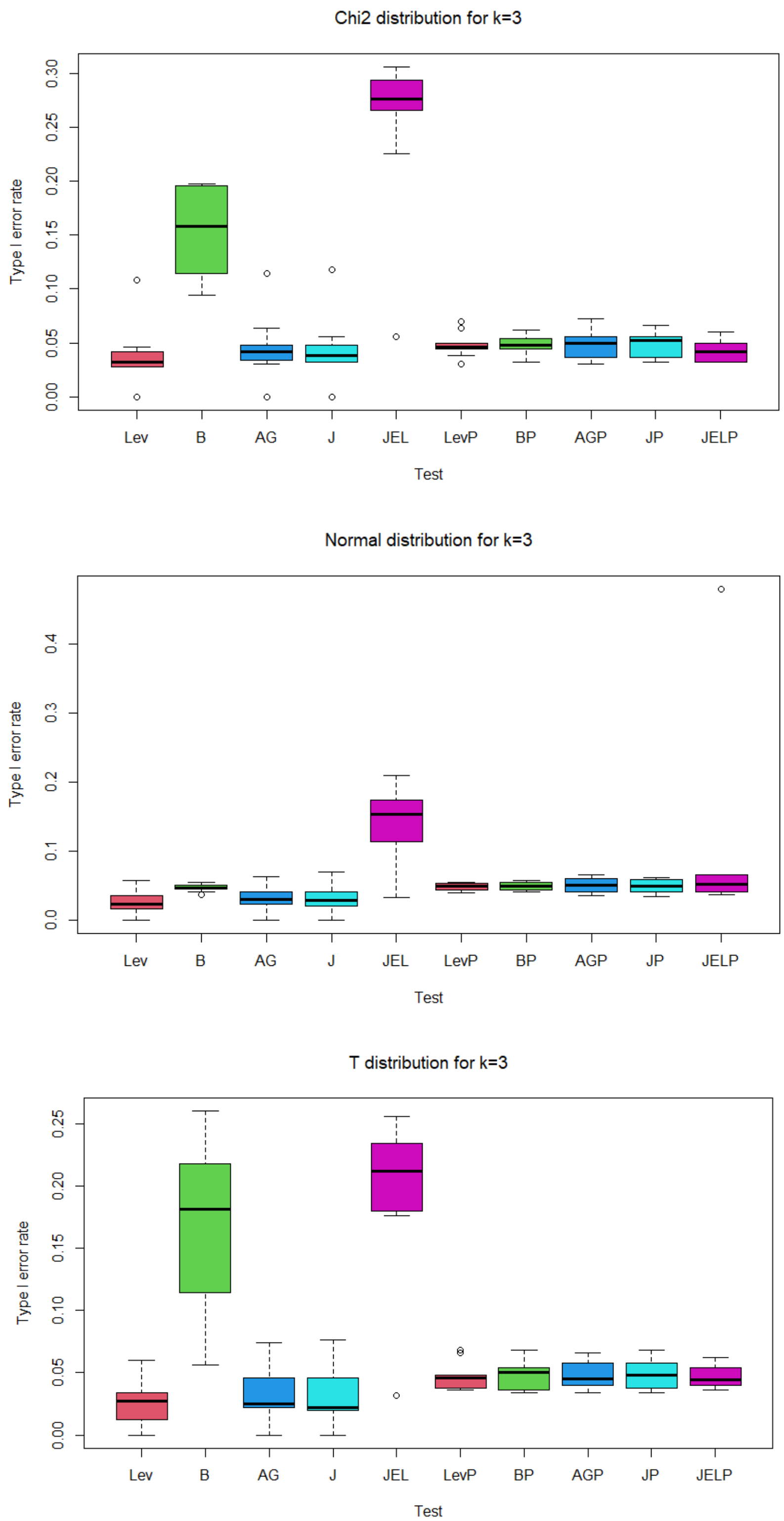
۱,۲. نتایج شبیه سازی: در این قسمت نتایج شبیه‌سازی را تحلیل می‌کنیم. در نمودارها آزمون‌های لون، بارتلت، الکساندر و گاورن، جیمز و درستمایی تجربی جک‌نایف را به ترتیب با نمادهای Lev، Bar ، AG ، J و JEL و آزمون‌های جایگشتی آنها را به ترتیب با LevP ، BarP ، AGP ، JP و JELP نشان می‌دهیم. نتایج شبیه‌سازی مربوط به نرخ خطای نوع اول برای k=2,3,4 به ترتیب در شکل‌های ۱، ۲ و ۳ بصورت نمودار جعبه‌ای برای حجم نمونه‌های شبیه‌سازی شده در هر گروه، داده شده‌اند.

با توجه به نمودارها آزمون JEL در همه توزیع‌ها نرخ خطای بالایی دارد، اما در نسخه جایگشتی آن نرخ خطا را به خوبی حفظ می‌کند. این رفتار برای آزمون بارتلت بجز برای توزیع نرمال که در آن نرخ خطا را حفظ می‌کند، نیز تکرار می‌شود. سایر آزمونها در حالت استفاده از ناحیه بحرانی توزیع تقریبی محافظه‌کار هستند و نرخ خطا را کمتر از مقدار اسمی ۰/۰۵ برآورد کردند. عملکرد این آزمون‌ها با اعمال روش جایگشت بهبود می‌یابد و برآورد نرخ خطا به مقدار اسمی نزدیکتر است.

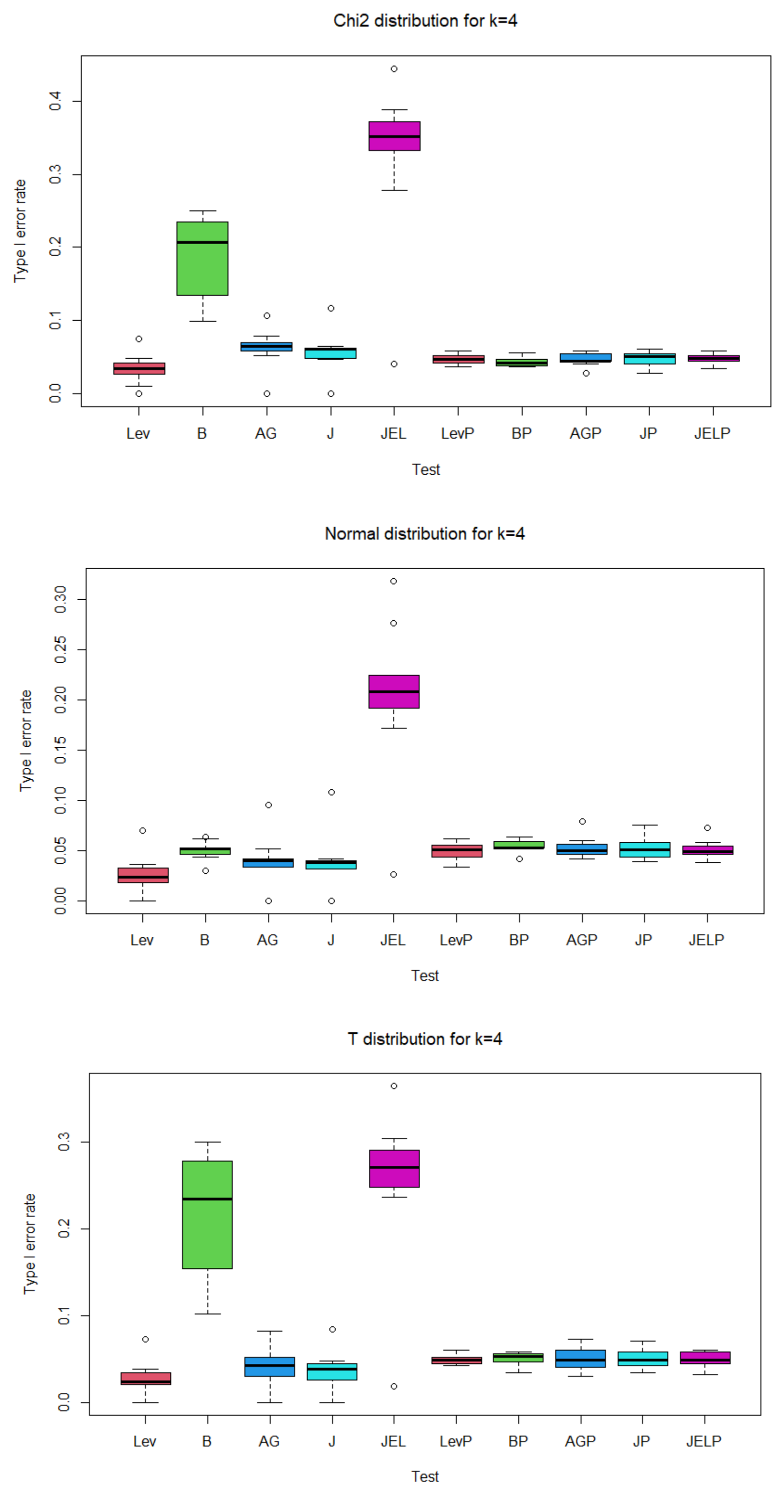
نتایج شبیه سازی مربوط به توان آزمون برای حالت ۲، ۳ و ۴ گروهی به ترتیب در شکل‌های ۴، ۵ و ۶ بصورت نمودار جعبه‌ای داده شده‌اند. با توجه به نمودارها رفتار آزمون JEL و بارتلت از سایر آزمون‌ها متفاوت است. در همه توزیع‌ها آزمون JEL متوسط توان بالایی دارد اگرچه پراکندگی آن بسیار چشمگیر است. هر چه تعداد گروه‌ها افزایش می‌یابد توان آزمون JEL نیز افزایش می‌یابد و از پراکندگی آن نیز کاسته می‌شود.

توان آزمون بارتلت برای همه توزیع‌ها و تعداد گروها رفتار یکسانی دارد. اگرچه بطور متوسط دارای توان قابل قبولی است ولی پراکندگی آن زیاد است. سایر آزمون‌ها در هر دو حالت معمولی و جایگشتی عملکرد تقریبا مشابهی دارند. در آزمون جایگشتی از میزان پراکندگی آزمون‌های JEL و بارتلت کاسته می‌شود و همچنین متوسط توان آنها نیز کاهش می‌یابد.

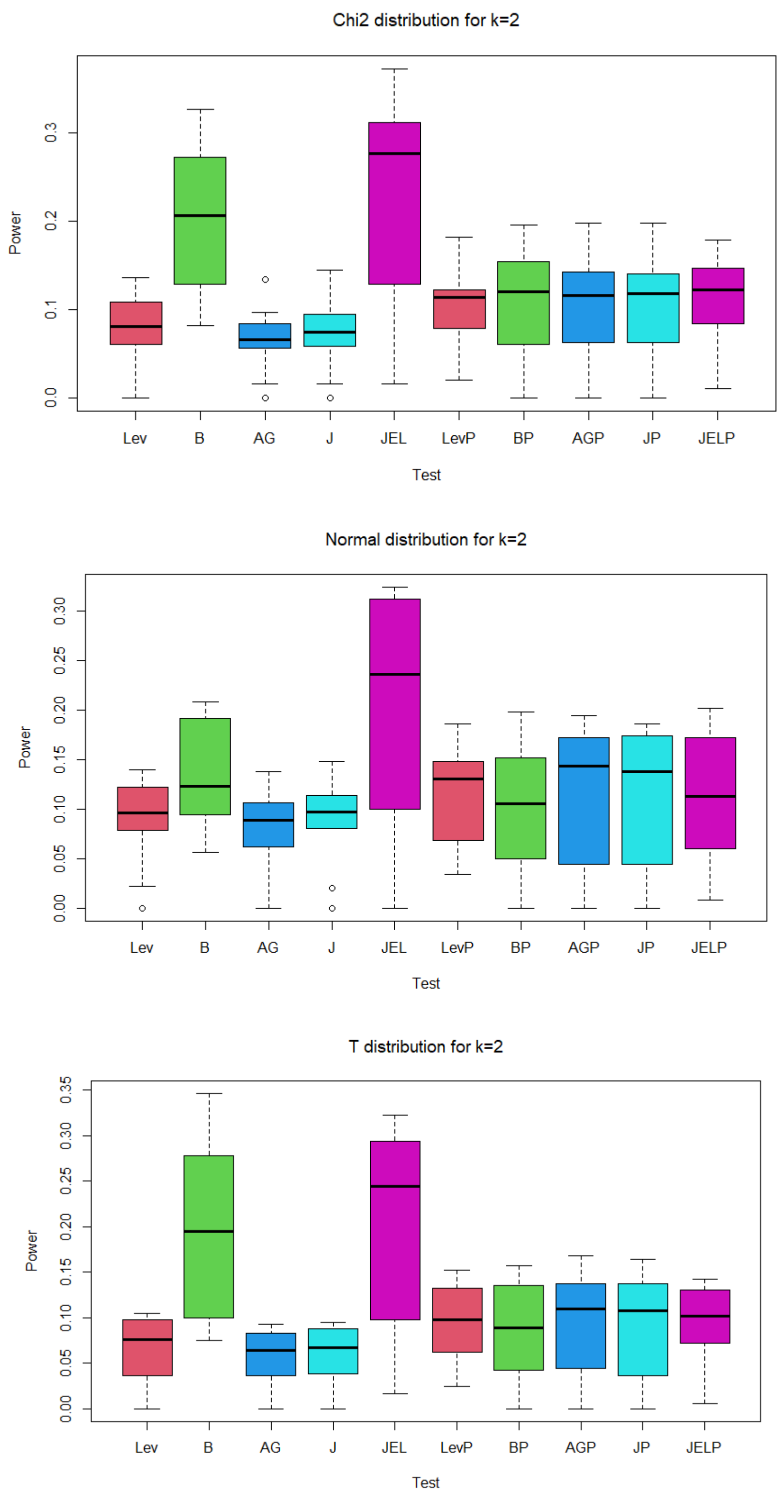
لازم است ذکر شود شبیه‌سازی برای حجم نمونه‌های بزرگ برای جوامع برابر و نابرابر نیز انجام شد. در این حالت عموما آزمون‌ها عملکرد خوبی دارند و توان نزدیک به یک حاصل می‌شود.



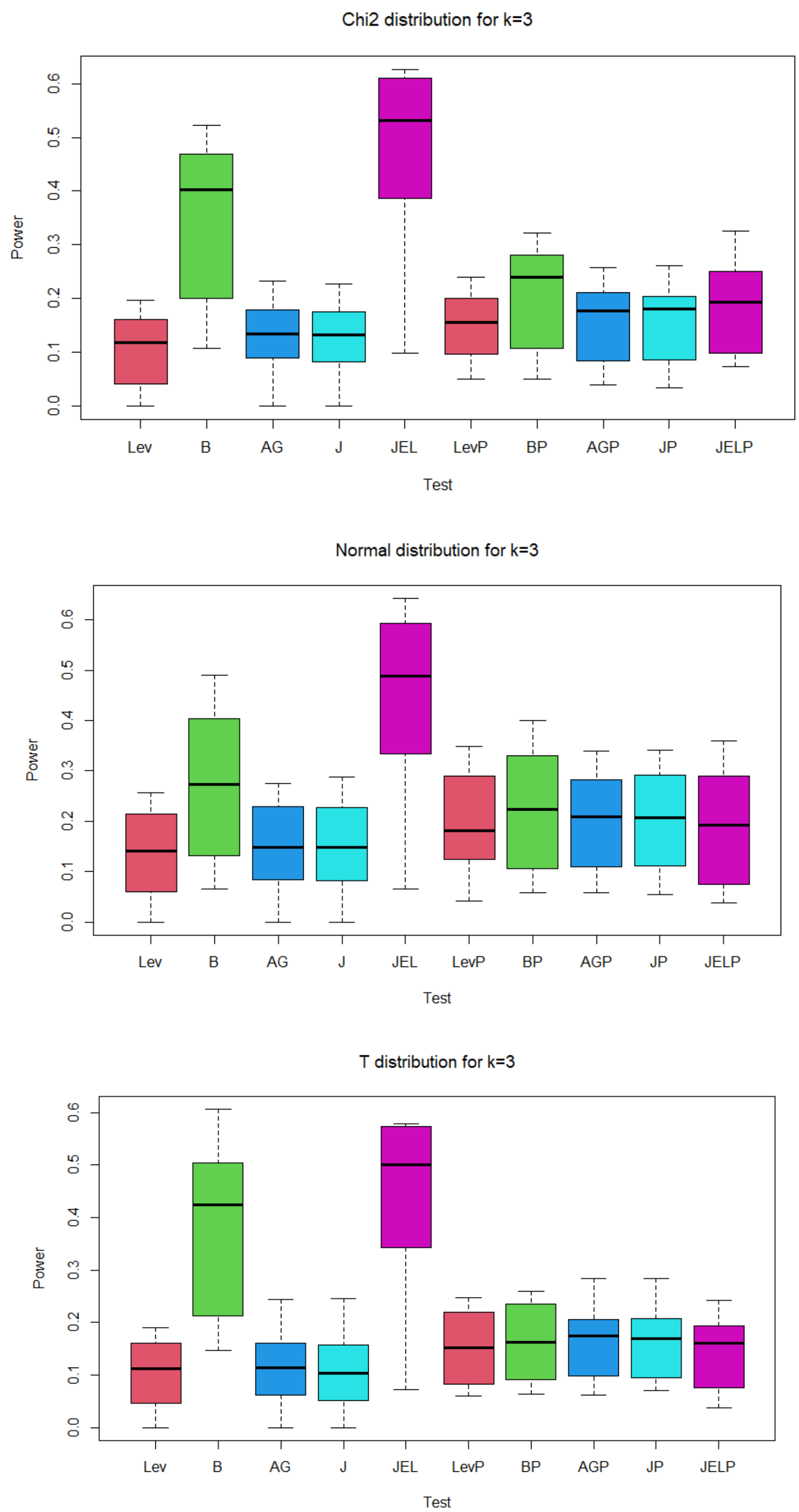
شکل ۲. نمودار جعبه‌ی نرخ خطای نوع اول در حالت سه گروهی



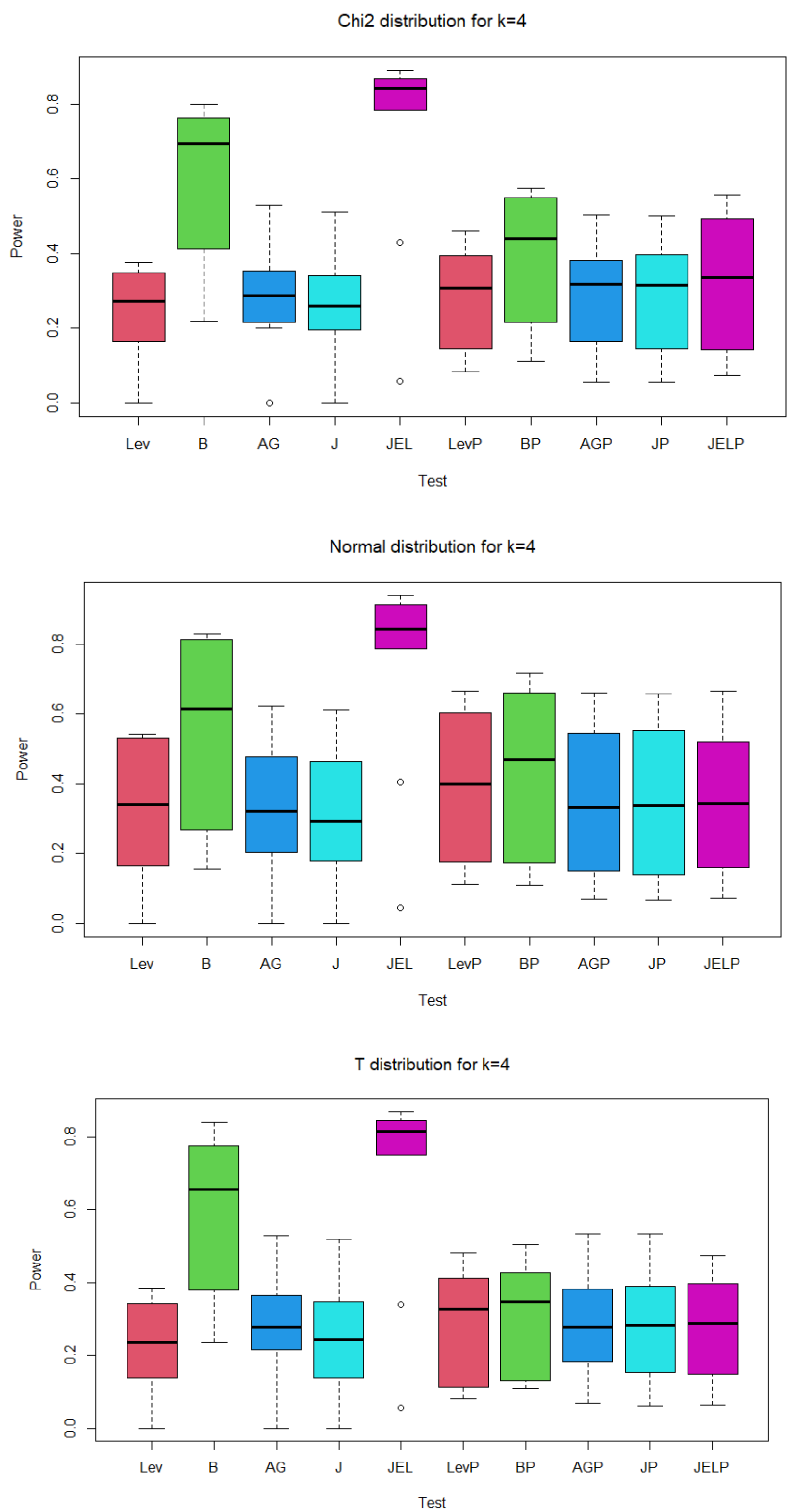
شکل ۳. نمودار جعبه‌ی نرخ خطای نوع اول در حالت چهار گروهی



شکل ۴. نمودار جعبه‌ی توان در حالت دو گروهی



شکل ۵. نمودار جعبه‌ی توان در حالت سه گروهی



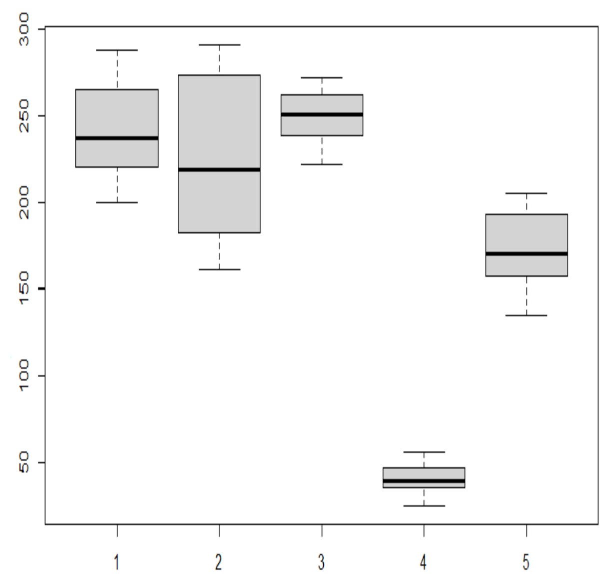
شکل ۶. نمودار جعبه‌ی توان در حالت چهار گروهی

۳. مطالعه‌ی موردی

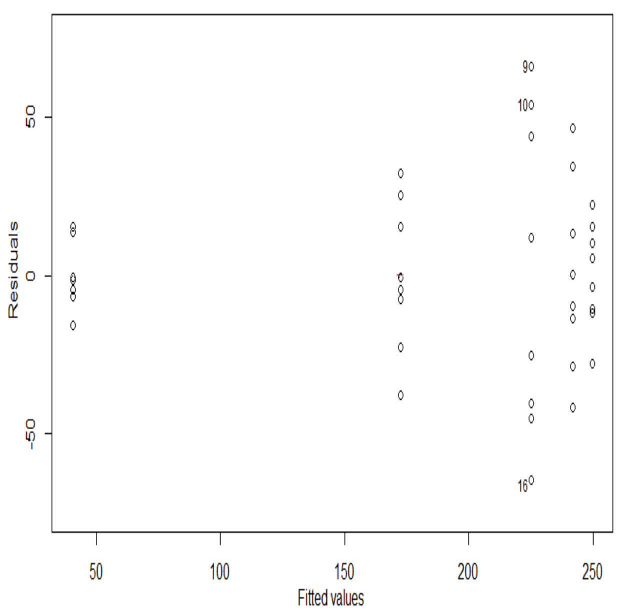
در این بخش با استفاده از داده‌های دو مثال واقعی، آزمون‌های مطرح شده را برای بررسی فرض همگنی واریانسها اعمال می‌کنیم.

مثال ۱: این مثال آزمون حافظه است که بر روی موش‌ها انجام گرفته است. در این آزمایش از دو جعبه‌ی سیاه و سفید که درب کوچکی بین آنها وجود دارد و باز شدن در به وسیله‌ی محقق کنترل می‌شود استفاده شده است. این آزمایش حافظه در دو روز انجام می‌شود. در روز اول حیوان در خانه‌ی سفید قرار دارد و یاد می‌گیرد که اگر درب باز شود و از خانه‌ی سفید به خانه‌ی سیاه برود شوک الکتریکی دردناکی به کف پایش داده می‌شود. بعد از ۲۴ ساعت آزمایش دوباره تکرار می‌شود؛ اگر حافظه‌ی موش سالم باشد در روز دوم با وجود باز بودن درب بین دو خانه وارد خانه سیاه نمی‌شود. لازم به ذکر است که موش‌ها به طور طبیعی خانه‌ی سیاه و تاریک را بیشتر دوست دارند.

در این بررسی متغیر مستقل پنج گروه آزمایشی و متغیر وابسته میزان حافظه است که با میزان به یاد آوردن شوک یا صبر کردن و نرفتن از خانه‌ی سفید به خانه‌ی سیاه اندازه‌گیری می‌شود. سقف زمان را ۳۰۰ ثانیه در نظر گرفته‌اند. گروه‌های آزمایشی به صورت کنترل عادی، کنترل دریافت کننده حلال داروی اول، کنترل دریافت کننده حلال داروی دوم، تیمار شده با داروی اول و حلال داروی دوم و تیمار شده با داروی اول و دوم در نظر گرفته شده‌اند که در اینجا آنها را به ترتیب با ۱ ، ۲ ، ۳ ، ۴ و ۵ نشان می‌دهیم. برای این آزمایش ۸ موش در هر گروه آزمایشی مورد بررسی قرار گرفته‌اند.



شکل ۷. نمودار جعبه‌ای آزمایش حافظه‌ی موش‌ها



شکل ۸. مانده‌ها در مقابل مقادیر برازش شده آزمون حافظه‌ی موش‌ها

این داده‌ها در آزمایشگاه گروه زیست‌شناسی دانشگاه کردستان تهیه شده است و توسط آقای دکتر شمس الدین احمدی عضو هیئت علمی گروه زیست‌شناسی در اختیار قرار گرفته است. از همکاری و حسن نظر ایشان کمال تشکر و قدردانی را می‌نماییم.

برای داده‌های آزمون حافظه‌ی موش‌ها، نمودار جعبه‌ی در شکل ۷ نشان داده شده است. شکل ۸ نمودار مانده‌ها در مقابل مقادیر برازش شده در مدل تحلیل واریانس یک‌طرفه مربوط به داده‌های آزمون حافظه موش‌ها می‌باشد که ناهمگنی واریانس‌ها را نشان می‌دهد.

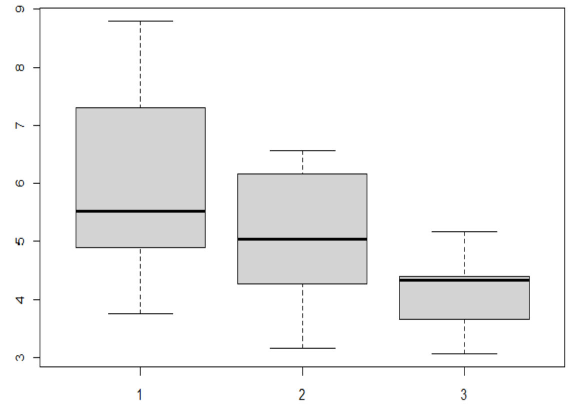
در جدول ۱ نتایج تحلیل داده‌های مورد نظر آورده شده است که مشاهده می‌کنیم آزمون‌های لون، بارتلت، الکساندر و گاورن، جیمز و درستنمایی تجربی جک‌نایف هم بر اساس توزیع تقریبی آماره و هم در حالت روش جایگشت فرض همگنی واریانس‌ها را رد می‌کنند.

جدول ۱. نتایج داده‌های مربوط به آزمون حافظه‌ی موش‌ها

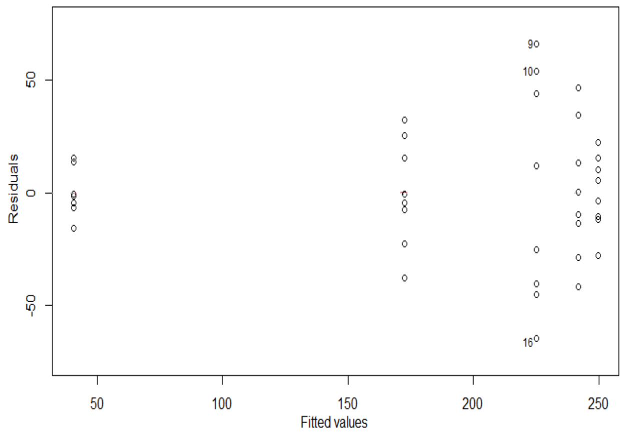
K=5	Lev	Bar	AG	J	JEL	LevP	BarP	AGP	JP	JELP
p- مقدار	۰/۰۰۰۱	۰/۰۰۱۵	۰/۰۰۰۳	-	۰/۰۱۱	۰/۰۰۰۸	۰/۰۰۱۴	۰/۰۳۵	۰/۰۳۰۲	۰
آماره مشاهده شده	۷/۸۲۴	۱۷/۵۰۷	۱۵/۸۸۹	۲۹/۰۵۷	۱۳/۰۶۳۵	-	-	-	-	-
مقدار بحرانی	۲/۶۴۲	۹/۴۸۸	۹/۴۸۷	۱۳/۲۹۱	۹/۴۸۷	-	-	-	-	-

مثال ۲: داده‌های این مثال از مقاله‌ی جاف و همکاران [۱۴] گرفته شده است. اندازه‌گیری‌ها در پایین دست یک زباله‌دان متروکه در رودخانه‌ی ولف در تسنی انجام شده‌اند که قبلاً برای دفع ضایعات آفت‌کش‌ها استفاده شده است. در اینجا متغیر مستقل سه عمق (سطح، میانی و پایین) و متغیر وابسته غلظت آلدرین بر حسب نانوگرم در لیتر است که به صورت ۱۰ مشاهده در هر عمق مورد بررسی قرار گرفته‌اند.

نمودار جعبه‌ی این داده‌ها در شکل ۹ نشان داده شده است. در شکل ۱۰ نمودار مانده‌ها در مقابل مقادیر برازش شده مربوط به داده‌های غلظت آلدرین را رسم کرده‌ایم که ناهمگنی واریانس‌ها را نشان می‌دهد. در جدول ۲ نتایج تحلیل داده‌های بررسی میزان غلظت آلدرین، با آزمون‌های مورد بررسی را نشان داده‌ایم. مشاهده می‌شود آزمون درستنمایی تجربی جک‌نایف بر اساس توزیع تقریبی آماره با تفاوت زیاد p- مقدار از سطح معناداری ۰/۰۵ فرض همگنی واریانس‌ها را رد می‌کند. اما سایر آزمون‌های مورد بررسی فرض صفر را رد نمی‌کنند. این داده‌ها در شومیکر [۲۶] نیز مورد مطالعه قرار گرفته‌اند که در آن دو آزمون معرفی شده F_{max} و کای ۲، فرض برابری واریانس‌ها را رد کرده‌اند.



شکل ۹. نمودار جعبه‌ای برای غلظت آلدرین



شکل ۱۰. مانده‌ها در مقابل مقادیر برازش شده برای غلظت آلدرین

جدول ۲. نتایج داده‌های مربوط به آزمون غلظت آلدرین

K=3	Lev	Bar	AG	J	JEL	LevP	BarP	AGP	JP	JELP
p- مقدار	۰/۱۴۰۷	۰/۰۶۰۱	۰/۱۳۰۳	-	۰/۰۰۵۴	۰/۱۲۸۲	۰/۱۲۳	۰/۱۲۷۲	۰/۱۳۱۶	۰/۰۸۳۸
آماره مشاهده شده	۲/۱۱۰۶	۵/۷۶۴۳	۴/۰۷۵۹	۴/۷۳۵	۱۰/۴۳۸	-	-	-	-	-
مقدار بحرانی	۳/۳۱۵۳	۵/۹۹۱۴	۵/۹۹۱۴	۷/۴۹۱	۵/۹۹۱۴	-	-	-	-	-

۴. نتیجه‌گیری

در این مقاله عملکرد پنج آزمون لون، بارتلت، الکساندر و گاورن، جیمز و آزمونی براساس شبه درست‌نمایی جک نایف مورد مطالعه قرار گرفتند. نرخ خطای نوع اول و توان آزمون برای حجم نمونه‌های کوچک در اندازه گروه‌های ۲، ۳ و ۴ با استفاده از شبیه‌سازی محاسبه و با سطح معنی‌داری ۵ درصد اسمی از لحاظ استواری خطای نوع اول برای توزیع‌های نرمال استاندارد، تی و کای دو مقایسه شدند. همچنین نسخه جایگشتی پنج آزمون نیز مورد بررسی قرار گرفت. نتایج نشان داد که آزمون‌های جایگشتی تاثیر زیادی در بهبود نرخ خطا و توان آزمون دارند. این بهبود در دو آزمون بارتلت و جک نایف چشم‌گیر است.

در کل نتایج بیان می‌کنند که آزمونی که درهمه شرایط حجم و اندازه گروه‌ها و توزیع‌ها بهترین باشد وجود ندارد. با این حال، نسخه‌های جایگشتی آزمون‌ها به‌ویژه برای حجم نمونه‌های کوچک پیشنهاد می‌شود. در استفاده از آزمون بارتلت باید دقت زیادی شود، زیرا این آزمون به شدت در مقابل نرمال نبودن حساس است.

تشکر و قدردانی

از داوران محترم که با پیشنهادات ارزنده موجب بهبود در متن اولیه مقاله شدند، سپاسگزاریم. همچنین از همه مسئولین محترم مجله پژوهشهای ریاضی بخاطر زحماتشان قدردانی می نماییم.

References

1. R. A. Alexander, D. M. Govern, A new and simpler approximation for ANOVA under variance heterogeneity. J Educ Stat. **19** (1994), 91–101.

2. M. S. Bartlett, Properties of sufficiency and statistical tests. Proc R Soc Lond, **160** (1937), 268–282.

3. M. S. Bartlett, D. G. Kendall, The Statistical Analysis of Variances-Heterogeneity and the Logarithmic Transformation. J Roy Statist Soc, **8** (1946),128-138.

4. D. D. Boos, C. Brownie, Comparing variances and other measures of dispersion. Stat Sci, **19** (2004), 571–578.

5. G. E. P. Box, Non-Normality and Tests on Variances. Biometrika. **40** (1953), 318.

6. M. B. Brown, A. B. Forsythe, Robust tests for the equality of variances. J Am Stat Assoc. **69** (1974), 364–367.

7. W. J. Conover, M. E. Johnson, M. M. Johnson, A Comparative Study of Tests for Homogeneity of Variances, with Applications to the Outer Continental Shelf Bidding Data. Technometrics. **23** (1981), 351–361.

8. W. J. Conover, A.J. Guerrero-Serrano, Tercero-Gómez VG. An update on “a comparative study of tests for homogeneity of variance.” J Stat Comput Simul, **88** (2018), 1454–1469.

9. N. Esmailzadeh, A Comparison of Five Bootstrap and Non-Bootstrap Levene-Type Tests of Homogeneity of Variances. Iran. J Sci Technol Trans A: Sci, **43** (2018), 979–989.

10. P. S. Gartside, A Study of Methods for Comparing Several Variances. J Am Stat Assoc, **67** (1972), 342–346.

11. W. Miao, J. L. Gastwirth, New Statistical Tests for Detecting Disparate Impact Arising from Two-Stage Selection Processes. Am Stat, **68** (2014), 146–157.

12. H. O. Hartley, The Maximum F-Ratio as a Short-Cut Test for Heterogeneity of Variance. Biometrika, **37** (1950), 308.

13. G. S. James, The Comparison of Several Groups of Observations When the Ratios of the Population Variances are Unknown. Biometrika, **38** (1951), 324.

14. P. R. Jaffe, Parker FL, Wilson DJ. Distribution of Toxic Substances in Rivers. J Environ Eng, **108** (1982), 639–649.

15. B. Y. Jing, J. Yuan, Z. Wang, Jackknife Empirical Likelihood. J Am Stat Assoc, **104** (2009),1224–1232.

16. M. W. J. Layard, Robust Large-Sample Tests for Homogeneity of Variances, J Am Stat Assoc, **68** (1973), 195–8.

17. E. L. Lehmann, Consistency and Unbiasedness of Certain Nonparametric Tests. Ann Math Statist, **22** (1951),165–79.

18. H. Levene, Robust tests for equality of variances. In: I. Olkin, S.G. Ghurye, W. Hoeding, W.G. Madow and H.B. Mann (Eds.), Contributions to Probability and Statistic. Stanford. Univ. Press. Stanford, (1960), 278-292.

19. K. J. Levy, An Empirical Comparison of Several Multiple Range Tests for Variances. J Am Stat Assoc, **70** (1975), 180–3.

20. W. Loh, Some modifications of Levene's test of variance homogeneity. J Stat Comput Simul, **28** (1987), 213–26.

21. R. G. Miller Jackknifing Variances. Ann Math Statist. **39** (1968), 567–82.

22. J. Neyman, E. S. Pearson, On the problem of k samples. Bull. Int. Acad. Cracovie, A. (1931), 460-481.

23. I. Parra-Frutos, Testing homogeneity of variances whit unequal sample sizes. Comput. Stat, **28** (2013), 1269-1297.

24. Y. Sang, A Jackknife Empirical Likelihood Approach for Testing the Homogeneity of K Variances. Metrika, **84** (2021),1025–48.

25. D. Sharma, B. M. G. Kibria, On some test statistics for testing homogeneity of variances: a comparative study. J Stat Comput Simul, **83** (2013), 1944–63.

26. L. H. Shoemaker, Fixing the F Test for Equal Variances. Am Stat, **57** (2003), 105–14.

27. S. Yitnosumarto, O'Neil ME. On Levene's test of variance homogeneity. Austral. J. Statist, **28** (1975), 230-241.