



Journal of Applied Sciences

ISSN 1812-5654

science
alert

ANSI*net*
an open access publisher
<http://ansinet.com>

Generalization of Klotz's Test

F. Karaman

Department of Statistics, Faculty of Art and Science,
Yildiz Technical University, Esenler, İstanbul, Turkey

Abstract: In this study, Klotz's test for heterogeneity of variance is generalized to factorial designs. Although, Levene's test, the jackknife and Fligner and Killeen's test for heterogeneity of variance were generalized to factorial designs, Klotz's test was not studied. The performance of Klotz's test is compared with other previously studied tests. An application and comparison of these analyses to the 2 by 2 factorial design are examined in detail. A simulation study was performed for 10000 data sets to compare power and robustness properties of the tests. It is observed that as sample size increased and as the difference in variances increased all tests have higher power. According to these simulations, Klotz's test can particularly be recommended for symmetric parent distributions. For skewed parent distribution like chi-square, Klotz's test should not be used since it is not robust. Data from a multifactor agricultural system were used as an example to illustrate the usefulness of these tests.

Key words: Klotz's test, factorial design, heterogeneity of variance, test for variance

INTRODUCTION

Most frequently, statistical analysis were performed to detect the differences in the location or means of the treatment groups, but in some cases, differences in dispersion may be of as much interest. Distance-based tests were proposed by Anderson (2006) to test homogeneity of multivariate dispersions. For unreplicated two-level fractional factorial designs McGrath and Lin (2002) developed a nonparametric dispersion tests. Ozaydin *et al.* (1999) studied factor effects on the variance in studies of an agricultural system involving chile peppers and two chile-pepper pests-root-knot nematodes and yellow nutsedge. Since there exist factor effects on the variance, analysis to detect factor effects on differences in means failed. Hence, this system required methods for examining the structure of effects on the variance. Although, there are many other situations to understand factor effects to the variance, pesticide application is one example to understand factor effects on the variance. Higher overall amounts to be applied for the methods with greater variability, so that; each area will receive the minimum amount of pesticide to be effective.

Some researchers performed testing heterogeneity of variance for one-way design. Some of these researchers are: Milliken and Johnson (1984) and Miller (1968) and so on. Among them Conover *et al.* (1981) performed 56 tests in a comprehensive simulation study of tests for heterogeneity of variance in the one-way model.

The Levene and the jackknife tests to detect heterogeneity of variance as a factor effect in a replicated

two-way treatment structure were studied and then generalization of Fligner and Killeen (1976)'s test was considered by Ozaydin *et al.* (1999).

While, many tests for scale exist for the completely randomized design, fewer tests have been considered for more complex designs. In particular nonparametric rank based tests for scale have not been applied to more complex designs. In addition to Fligner and Killeen's nonparametric rank-based test, which was studied by Ozaydin *et al.* (1999), it is of interest to consider the generalization of one such test, Klotz (1962)'s nonparametric rank-based test, to a two-way treatment structure.

For one-way design according to power and robustness, performance of Klotz's test for symmetric parent distributions was quite well. It was also recommended for symmetric parent distributions by Conover *et al.* (1981).

As it mentioned on Ozaydin *et al.* (1999) study; computation of pseudo-observations reflects changes in the magnitude of the variance. Then either the usual ANOVA or an analogous chi-square analysis was used to analyze these pseudo-observations. A review of the tests was studied by Ozaydin *et al.* (1999).

Simulation studies were used to examine the power and the robustness of the generalized tests for the special case of a 2 by 2 treatment structure with balanced data. Performance of Klotz's test for the balanced two-way factorial treatment structure is compared by other previously studied tests. These simulation studies were performed for 10000 data sets.

The simulation using SAS/STAT software was performed. Uniform (short-tailed), normal, double exponential (long-tailed) and chi-square (skewed) parent distributions were used for the simulation study. Samples of size 5, 10 and 20 observations were used for the different variance combinations.

THE MODEL AND HYPOTHESES

For the fixed two-way design with equal replication the cell mean model is:

$$y_{ijk} = \mu_{ij} + \epsilon_{ijk} \tag{1}$$

where, $i = \dots, t_A > 1, j = 1 \dots, t_B > 1, k = 1 \dots, n > 1$. In the model, y_{ijk} is the k 'th response from the i 'th level of factor A and the j 'th level of factor B. The μ_{ij} are unknown constants representing the average response due to the i 'th level of factor A and the j 'th level of factor B. Then, the ϵ_{ijk} are the random error associated with the observation y_{ijk} and are assumed to be independently distributed with zero mean and variance σ_{ij}^2 .

To investigate the behavior of the following test statistics under less restrictive distributional assumptions, the assumption that the errors are assumed to be normal random variables is suspended.

For my studies, I use the special case $t_A = 2, t_B = 2$, the null hypothesis is $H_0: \sigma_{11}^2 = \sigma_{12}^2 = \sigma_{21}^2 = \sigma_{22}^2$, (all variances are equal).

At least one variance does not equal to others' is the alternative hypothesis. Instead of using usual one-way ANOVA analysis to test equality of variance, following the Ozaydin *et al.* (1999), I perform an analysis that is analogous to the usual ANOVA for a two-way replicated treatment structure. The following null hypothesis are used to find the differences in variances and, if they differ, the form of the difference:

- $H_0: \sigma_{11}^2 - \sigma_{12}^2 = \sigma_{21}^2 - \sigma_{22}^2$ (interaction between factor A and factor B does not exist)
- $H_0: \sigma_{11}^2 + \sigma_{12}^2 = \sigma_{21}^2 + \sigma_{22}^2$ (there is not a factor A main effect on the variance)
- $H_0: \sigma_{11}^2 + \sigma_{21}^2 = \sigma_{12}^2 + \sigma_{22}^2$ (there is not a factor B main effect on the variance)

If there exists an interaction, main effects are not considered further. The presence of interaction suggests that a simple additive main effects structure does not exist among the factor effects on the variance so that a reduction in summarizing the variances is not possible.

THE TEST PROCEDURES

Levene's test and the jackknife test are classified as tests based on modifications of the F-test for means and Klotz's test and Fligner and Killeen's test are classified as a linear rank tests (or a rank-like tests) (Conover *et al.*, 1981). For these tests, pseudo-observations that in some way reflect changes in the magnitude of the variance are computed. Then, the pseudo-observations are analyzed.

In general, the pseudo-observations are some function of the absolute deviations from either the mean or the median. For the (1) continuous pseudo-observations, the analysis will be the two-way ANOVA with F-tests. For the (2) rank-based pseudo-observations, an analogous analysis using chi-square test statistics is more appropriate. The tests considered in this study, such as Levene's test, the jackknife test, Klotz's test and Fligner and Killeen's test for the one-way model are among the tests that follow this basic test rationale.

The following seven tests are considered and compared for balanced data using the pseudo-observations in this study (Ozaydin *et al.*, 1999) considered the first five of these tests and then compared all these five tests as having the best performance on the basis of robustness and power. The last two of the seven tests are generalizations of Klotz's test.

- (1) For (Lev 1) the square of the residuals was computed, then ANOVA was performed on these pseudo-observations:

$$Z_{ijk} = (y_{ijk} - \bar{y}_{ij})^2, i = 1, \dots, t_A; j = 1, \dots, t_B; k = 1, \dots, n \tag{2}$$

where, \bar{y}_{ij} is the sample mean for the ij 'th group (Ozaydin *et al.*, 1999).

- (2) For a modified Levene's test (Lev 2), the absolute deviations from the sample medians \tilde{y}_{ij} was computed as follows, then ANOVA was performed on these pseudo-observations (Ozaydin *et al.*, 1999).

$$Z_{ijk} = |y_{ijk} - \tilde{y}_{ij}| \tag{3}$$

- (3) For the jackknife test, (Ozaydin *et al.*, 1999) performed ANOVA F-tests on the following pseudo-observations (jack):

$$U_{ijk} = n \ln(S_{ij}^2) - (n-1) \ln(S_{ij(k)}^2) \tag{4}$$

where, S_{ij}^2 is the usual sample variance for the ij th group.

$$\bar{y}_{ij(k)} = \sum_{i \neq k} \frac{y_{ij}}{n-1}$$

was calculated and then:

$$S_{y(k)}^2 = \sum_{i \neq k} \frac{(y_{ij} - \bar{y}_{ij(k)})^2}{n-2}$$

- (4) The first generalization of Fligner and Killeen's test applies an F-test analysis to rank-based pseudo-observations. Following (Conover *et al.*, 1981; Ozaydin *et al.*, 1999) used absolute deviations from the median:

$$Z_{ijk} = |y_{ijk} - \tilde{y}_{ij}|$$

The Z_{ijk} 's were then ranked from smallest to largest (i.e., from 1 to $N = nt_A t_B$). The ranks are denoted by R_{ijk} . Following Conover (1980), if there were ties in my sample, I calculated the average of the ranks of the tied values and used the average as the R_{ijk} value for the tied Z_{ijk} 's.

Then to modify the ranks, a score function was used. The score function (denoted FK) was used by Fligner and Killeen (1976):

$$a(R_{ijk}) = \left[\Phi^{-1} \left(\frac{N+1+R_{ijk}}{2(N+1)} \right) \right]^2 \quad (5)$$

where, Φ is the standard normal cumulative distribution function. Following (Ozaydin *et al.*, 1999), I used Fligner and Killeen's original score function and denote the values $a(R_{ijk})$ by a_{ijk} .

Conover *et al.* (1981) used the positive square root of this score function. Ozaydin *et al.* (1999) used Fligner and Killeen's original score function since it has better power and is more robust than using the positive square root of this score function. The ANOVA F-test analysis was performed to the scores a_{ijk} . The procedure obtained by applying F-tests to the pseudo-observations FK is denoted as FK-F, which is same as by Ozaydin *et al.* (1999).

- (5) The second generalization of Fligner and Killeen's test performs a chi-square analysis on the pseudo-observations FK. For the overall test of equality of variances, the one-way test statistic presented by Conover *et al.* (1981) was used:

$$X = n \sum_{i=1}^{t_A} \sum_{j=1}^{t_B} (\bar{a}_{ij} - \bar{a}_{...})^2 / D^2 \quad (6)$$

where,

$$D^2 = \frac{1}{N-1} \sum_{i=1}^{t_A} \sum_{j=1}^{t_B} \sum_{k=1}^n (a_{ijk} - \bar{a}_{...})^2$$

where, \bar{a}_{ij} is the mean score in the ij 'th sample and $\bar{a}_{...}$ is the overall mean score. Under the null hypothesis of equal variances, D^2 is the known variance of the scores. Ozaydin *et al.* (1999) partitioned X into three pieces, corresponding to each of the three hypotheses mentioned earlier as:

$$X_1 = n \sum_{i=1}^{t_A} \sum_{j=1}^{t_B} (\bar{a}_{ij} - \bar{a}_{i.} - \bar{a}_{.j} + \bar{a}_{...})^2 / D^2 \quad (7)$$

$$X_2 = nt_B \sum_{i=1}^{t_A} (\bar{a}_{i.} - \bar{a}_{...})^2 / D^2 \quad (8)$$

where, $\bar{a}_{i.}$ is the mean score in the i th sample

$$X_3 = nt_A \sum_{i=1}^{t_B} (\bar{a}_{.j} - \bar{a}_{...})^2 / D^2 \quad (9)$$

where, $\bar{a}_{.j}$ is the mean score in the j th sample

Under the null of no differences, each of the above test statistics is approximately distributed as a chi-square random variable with degrees of freedom corresponding to the numerator sum of squares:

$$X - \chi_{t_A t_B - 1}^2, X_1 - \chi_{(t_A - 1)(t_B - 1)}^2, X_2 - \chi_{(t_A - 1)}^2 \text{ and } X_3 - \chi_{(t_B - 1)}^2$$

where, I denote the procedure obtained by applying the chi-square tests to the pseudo-observations FK by FK-Chi.

- (6) The first generalization of Klotz's test applies an F-test analysis to rank-based pseudo-observations. Following the same procedures as in (4), values of R_{ijk} were calculated

A score function was then used to modify the ranks. Klotz (1962) used the score function (denoted Klotz's)

$$a(R_{ijk}) = \left[\Phi^{-1} \left(\frac{R_{ijk}}{N+1} \right) \right]^2 \quad (10)$$

where, Φ is the standard normal cumulative distribution function. The values $a(R_{ijk})$ are denoted by a_{ijk} . The ANOVA F-test analysis was applied to the scores a_{ijk} . The

procedure obtained by applying F-tests to the pseudo-observations Klotz is denoted as Klotz-F.

- (7) A second variant of Klotz's test is a generalization of the usual rank-based nonparametric approach; i.e., a chi-square analysis was performed on the pseudo-observations Klotz. Following the same steps explained as in (5), the chi-square tests were performed. The procedure obtained by applying the chi-square tests to the pseudo-observations Klotz is denoted by Klotz-Chi

The last four tests are variations of Fligner and Killeen's and Klotz's nonparametric rank-based approach. Chi-square tests are more appropriate and have traditionally been used for rank-based tests. Following Conover and Iman (1981), rank transformation procedures applying the usual ANOVA F-tests to the ranks are easier to obtain. The fourth and sixth tests use this notion but apply the usual ANOVA F-tests to scores based on ranks. The fifth and seventh tests use a traditional chi-square analysis of the scores.

All of the above tests use pseudo-observations in different way. For example, pseudo-observations which are used in Lev1 are on the scale of variance, pseudo-observations which are used in Lev2 are on the scale of standard deviation, in Jack are on a log transformed scale. The FK and Klotz pseudo-observations are based on ranks.

RESULTS

It was shown by Conover *et al.* (1981) study that performance of Klotz's test for symmetric parent distributions was quite well according to power and robustness for one-way design. Therefore, in this study performance of Klotz's test statistics was studied for two-way design. Lev1 and Lev2 are modifications of the Levene's test. ANOVA performed on any monotonically increasing function of the absolute value of deviations from the sample means. To obtain more robust tests Levene's idea can be applied to deviations from the median (Miller, 1968). The third test is the jackknife test. The fourth and fifth tests are variations of Fligner and Killeen's and the sixth and seventh tests are variations of Klotz's nonparametric rank-based approach. For rank-based tests, chi-square tests are more appropriate than ANOVA. In this study, usual ANOVA F-tests applied to the ranks were also studied since it is easier to obtain (Conover and Iman, 1981). The fifth (FK-chi) and seventh

Klotz-chi tests use traditional chi-square analysis of the scores. The fourth (FK) and sixth (Klotz) tests use ANOVA F-tests to scores based on the ranks.

For the special case of a 2 by 2 treatment structure, simulations using SAS/STAT software and procedures compared the power and robustness of the seven proposed test procedures. Variance configurations, a range of parent distributions and sample sizes were indicated.

Uniform (short-tailed), normal, double exponential (long-tailed) and Chi-square (skewed) parent distributions were used for the simulation study. When sampling from a chi-square distribution to maintain the level of skewness, 1 d.f. chi-square random variables were generated and multiplied by the appropriate constant to attain the indicated variability. Samples of size 5, 10 and 20 observations were used for each of the following variance combinations indicated by ($\sigma_{11}^2, \sigma_{12}^2, \sigma_{21}^2, \sigma_{22}^2$):

- (1) (1,1,1,1)
- (2) (1,1,2,2), (1,1,4,4), (1,1,8,8)
- (3a) (1,2,2,1), (1,4,4,1), (1,8,8,1)
- (3b) (1,1,1,2), (1,1,1,4), (1,1,1,8)
- (4) (1,2,2,3), (1,4,4,7), (1,8,8,15)

Variance combination (1) was used to approximate the true type I error rate. (i.e., the observed significance level). The set of variance combinations (2) was used to investigate the power of the tests when only a factor main effect is present. The set of variance combinations (3a) was used to investigate the power of the tests when an interaction effect is present. The set of variance combinations (3b) was used to investigate the power of the tests when an interaction effect and both main effects are present. The set of variance combinations (4) was used to investigate the power of the tests when both A and B main effects exist.

The simulation results are shown in the Table 1-3. The simulation results of first five tests are similar to (Ozaydin *et al.*, 1999) results. It is observed that as sample size increased and as the difference in variances increased all tests have higher power.

All of the tests are robust with the exception of Jack, Lev 1, Klotz-F and Klotz-Chi. Jack, Klotz-F and Klotz-Chi are never robust when applied to samples from a chi-square parent distribution. Lev 1 is not robust when used with small samples (n = 5) from a chi-square parent distribution. Lev 2, FK-F and FK-Chi suffer

Table 2: Results of moderate samples (n = 10).

Chi-square distribution							Double exponential distribution						
Tests	VC	Model	Only A	Only B	Both A and B	INT	Tests	VC	Model	Only A	Only B	Both A and B	INT
Lev 1	1	0.0907	0.0500	0.0539	0.0040	0.0690	Lev 1	1	0.0484	0.0461	0.0450	0.0027	0.0514
	2	0.1875	0.1817	0.0321	0.0152	0.0752		2	0.2232	0.3694	0.0190	0.0132	0.0497
	3a	0.1856	0.0311	0.0357	0.0037	0.2329		3a	0.2241	0.0201	0.0179	0.0017	0.4113
	3b	0.1894	0.0696	0.0643	0.0117	0.1242		3b	0.2703	0.0796	0.0858	0.0188	0.1709
Lev 2	1	0.1628	0.0964	0.0945	0.0147	0.0761	Lev 2	1	0.1759	0.1409	0.152	0.0249	0.0551
	2	0.0519	0.0376	0.0432	0.0033	0.0496		2	0.0358	0.0376	0.0366	0.0010	0.0388
	3a	0.1468	0.2149	0.0282	0.0131	0.0539		3a	0.2452	0.4175	0.0156	0.0158	0.0385
	3b	0.1485	0.0262	0.0285	0.0027	0.2478		3b	0.2453	0.0172	0.0157	0.0013	0.4560
Jack	1	0.1518	0.0711	0.0678	0.0116	0.1027	Jack	1	0.2627	0.0904	0.0943	0.0222	0.1676
	2	0.1269	0.1037	0.0999	0.0143	0.0592		2	0.2015	0.1608	0.1686	0.0352	0.0498
	3a	0.1451	0.0709	0.0738	0.0123	0.1023		3a	0.0852	0.0572	0.0561	0.0057	0.0685
	3b	0.2650	0.2222	0.0504	0.0357	0.1023		3b	0.3403	0.3937	0.0272	0.0346	0.0685
FK-F	1	0.2723	0.0489	0.0543	0.0077	0.3030	FK-F	1	0.3407	0.0272	0.0258	0.0027	0.4654
	2	0.2376	0.0897	0.0937	0.027	0.1571		2	0.2823	0.1028	0.1037	0.0312	0.1822
	3a	0.2774	0.1258	0.1225	0.0423	0.1181		3a	0.3637	0.1545	0.1732	0.0814	0.1044
	3b	0.0718	0.0504	0.0538	0.0041	0.0671		3b	0.0348	0.0396	0.0385	0.0016	0.0414
FK-Chi	1	0.2125	0.2605	0.0359	0.0194	0.0723	FK-Chi	1	0.2490	0.4191	0.0175	0.0159	0.0442
	2	0.2077	0.0345	0.0352	0.0041	0.3043		2	0.2530	0.0194	0.0183	0.0013	0.4616
	3a	0.2004	0.0847	0.0871	0.0162	0.1385		3a	0.2574	0.0947	0.094	0.0229	0.1708
	3b	0.1843	0.1312	0.1304	0.0216	0.0786		3b	0.1905	0.1621	0.1673	0.0338	0.0501
Klotz-F	1	0.0628	0.0523	0.0562	0.0029	0.0628	Klotz-F	1	0.0280	0.0418	0.0389	0.0009	0.0403
	2	0.1927	0.2699	0.0397	0.0133	0.0617		2	0.2244	0.4326	0.0186	0.0096	0.0327
	3a	0.1913	0.0366	0.0388	0.0038	0.2968		3a	0.2251	0.0204	0.0201	0.0007	0.4590
	3b	0.1838	0.0905	0.0937	0.0124	0.1252		3b	0.2364	0.1006	0.1023	0.0174	0.1454
Klotz-Chi	1	0.1673	0.1348	0.1365	0.0149	0.0708	Klotz-Chi	1	0.1675	0.1690	0.1733	0.0207	0.0440
	2	0.7296	0.1749	0.1778	0.0729	0.4443		2	0.0835	0.0695	0.0643	0.0048	0.0774
	3a	0.8593	0.3398	0.0765	0.1008	0.4204		3a	0.3989	0.4865	0.0275	0.0319	0.0776
	3b	0.8593	0.0788	0.0797	0.0301	0.7514		3b	0.4122	0.0282	0.0271	0.0027	0.5718
Normal distribution	1	0.8371	0.1474	0.1423	0.0801	0.5553	Uniform distribution	1	0.3763	0.1156	0.1181	0.0334	0.2547
	2	0.8477	0.2025	0.1893	0.1238	0.4199		2	0.3312	0.2000	0.2086	0.0612	0.0847
	3a	0.7076	0.2063	0.2117	0.0636	0.3801		3a	0.0736	0.0718	0.067	0.0030	0.0727
	3b	0.8440	0.4223	0.0975	0.1035	0.3107		3b	0.3745	0.5086	0.0283	0.0225	0.0587
Uniform distribution	1	0.8412	0.0984	0.1015	0.0259	0.7101	Normal distribution	1	0.3846	0.0307	0.0289	0.0020	0.5632
	2	0.8188	0.1867	0.1815	0.0851	0.4663		2	0.3530	0.1294	0.1315	0.0276	0.2209
	3a	0.8292	0.2504	0.2382	0.1213	0.3201		3a	0.3028	0.2105	0.2198	0.0431	0.0717
	3b							3b					

For VC see the variance combinations. As for the proportions in the columns Model, Only A, Only B, Both A and B, INT are based on 10,000 simulations for VC (1) and 30,000 for all other VC's, since these combine all three subcombinations for combinations (2), (3a), (3b), and (4). VC (1) exploring robustness, while VC (2) explores power in the presence of an A main effect, VC (3a) and (3b) explore power in the presence of interaction, VC (4) investigating power in the presence of A and B main effects. The column Model corresponds to the overall test of homogeneity of variance. Categories Only A, Only B, Both A and B, and INT are mutually exclusive and correspond to a factorial ANOVA.

Table 3: Result of Large samples (n =20)

Table with 14 columns: Tests, VC, Model, Only A, Only B, Both A and B, INT, Tests, VC, Model, Only A, Only B, Both A and B, INT. It is divided into four sections: Chi-square distribution, Double exponential distribution, Normal distribution, and Uniform distribution.

For VC see the variance combinations. As for the proportions in the columns Model, Only A, Only B, Both A and B, INT pare based on 10,000 simulations for VC (1) and 30,000 for all other VC's, since these combine all three subcombinations for combinations (2), (3a), (3b), and (4). VC (1) exploring robustness, while VC (2) explores power in the presence of an A main effect, VC (3a) and (3b) explore power in the presence of interaction, VC (4) investigating power in the presence of A and B main effects. The column Model corresponds to the overall test of homogeneity of variance. Categories Only A, Only B, Both A and B, and INT are mutually exclusive and correspond to a factorial ANOVA.

from being too conservative and also lack power when applied to small samples (n = 5).

For skewed parent distribution like chi-square Klotz-F and Klotz-Chi tests are not robust and strongly recommended that these two tests should not be used for skewed parent distributions. For symmetric parent distributions performance of Klotz-F and Klotz-Chi is quite well. Both Klotz-F and Klotz-Chi tests show the most power. They are also robust. Klotz-F and Klotz-Chi tests can strongly be recommended when applied to samples from symmetric parent distributions.

NUMERICAL EXAMPLE

An example of a greenhouse experiment is used from Ozaydin *et al.* (1999). The treatment structure for experiment is a 2 by 2 factorial. The application of the seven test procedures to a two-way design is studied by the data. The first factor was chile (absent, present) and the other factor was source of nematodes (nematode from chile or from tomato) (Table 4). Germination of the yellow nutsedge tubers produced during the experiment was the response variable.

Regardless of nematode source, variance in yellow nutsedge tuber germination was higher when chile was

present (Table 5). To test for the presence and form of heterogeneity of variance for two-way structures all seven of the analysis for two-way structures were used (Table 6). The conclusion of all the analyses was: Absence or presence of chile affects the variance of nutsedge tuber germination (means statistically significant chile main effect).

The variance of the tuber germination with respect to presence and absence of chile was pooled since the data suggested that there are two distinct variances, one when chile is present and one when it is absent. A two sample t-test comparing the means for the treatments (chile absent, nematode from chile) and (chile absent, nematode from tomato) was performed. The result was there is no significant difference in the mean tuber germination (p = 0.4609). Another two sample t-test comparing the means for the treatments (chile present, nematode from chile) and (chile present, nematode from tomato) was performed. The result was there is no significant difference in the mean tuber germination (p = 0.7856). Therefore the means within chile levels was also pooled (Table 7).

Although, in practice the most appropriate test should be chosen and applied, for purposes of example, all seven tests were applied to the data.

Table 4: Nutsedge tuber germination by chile and nematode source^a

Chile	Nematode source											
	Chile				Tomato							
Absent	86	100	70	80	88	70	100	60	67	71	70	90
Present	30	100	100	63	100	30	40	100	44	100	29	78

^aData from (Ozaydin *et al.*, 1999)

Table 5: Means and variances of nutsedge tuber germination by chile and nematode source^a

Chile	Nematode source	Mean	Variance
Absent	Chile	82.3333	133.4667
	Tomato	76.3333	233.8667
Present	Chile	70.5000	1,189.5000
	Tomato	65.1666	996.1667

^aData from Ozaydin *et al.* (1999)

Table 6: Summary of analyses of the effect of chile nematode (N) source on nutsedge tuber germination variability

Source	d.f.	p-values						
		Lev 1	Lev 2	Jack	FK-F	FK-Chi	Klotz-F	Klotz-Chi
Model	3	0.0016	0.0103	0.0210	0.0333	0.0465	0.0055	0.01414
Chile	1	0.0002	0.0011	0.0029	0.0049	0.0061	0.0005	0.00115
N. Source	1	0.8145	0.9732	0.7082	0.9752	0.9782	0.8648	0.89194
C×N	1	0.4611	0.7119	0.4658	0.4380	0.7749	0.9483	0.95873

Table 7: Means and variances of nutsedge tuber germination by Chile^a

Chile	Mean	Variance
Absent	79.3333	176.7879
Present	67.8333	1,001.2400

^aData from (Ozaydin *et al.* (1999)

CONCLUSION

Although performance of Klotz's tests for symmetric distributions were quite well according to power and robustness in a one-way design, Klotz's test statistics was not studied in a two-way design. This is the reason why in this study the performance of the Klotz's test statistics was studied and compared with the other five tests. Also, a simulation study was performed for 10000 times instead 1000 times. So, the performance of other previously studied tests were restudied based on 10000 simulations. For future studies, other tests examined by (Conover *et al.*, 1981) for one-way design could be tested and compared for more complex designs to evaluate their performance.

REFERENCES

- Anderson, M.J., 2006. Distance-based tests for homogeneity of multivariate dispersions. *Biometrics*, 62: 245-253.
- Conover, W.J., 1980. *Practical Nonparametric Statistics*. 2nd Edn., John Wiley and Sons, New York, ISBN-10: 0471028673.
- Conover, W.J. and R.L. Iman, 1981. Rank transformations as a bridge between parametric and nonparametric statistics. *Am. Statistician*, 35: 124-129.
- Conover, W.J., M.E. Johnson and M.M. Johnson, 1981. A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data. *Technometrics*, 23: 351-361.
- Fligner, M.A. and T.J. Killeen, 1976. Distribution-free two sample tests for scale. *J. Am. Statistical Assoc.*, 71: 210-213.
- Klotz, J., 1962. Nonparametric tests for scale. *Ann. Math. Stat.*, 33: 498-512.
- Miller, R.G., 1968. Jackknifing variances. *Ann. Math. Stat.*, 39: 567-582.
- Milliken, G.A. and D.E. Johnson, 1984. *Analysis of Messy Data*. 1st Edn., Chapman and Hall, New York, ISBN-0: 158488083X.
- McGrath, R. and D. Lin, 2002. A nonparametric dispersion test for unreplicated two-level fractional factorial designs. *J. Nonparamet. Statist.*, 14: 699-714.
- Ozaydin, F., D.M. VanLeeuwen, C.S. Miller and J. Schroeder, 1999. Factor effects on the variance in a replicated two-way treatment structure in an agricultural system. *J. Agric. Biol. Environ. Stat.*, 4: 166-184.