Academic Journal of Agriculture 2(1):55-61 (2013)
ISSN: 2147-6403 http://azd.odu.edu.tr

Review
(Derleme)

A fast approach to select the appropriate test statistics

Fatih ÜÇKARDEŞ[1], Elif ASLAN[1], Hande KÜÇÜKÖNDER[2]

[1] Adiyaman University, Faculty of Medicine, Biostatistics and Medical Informatics, ADIYAMAN, Turkey
[2] Bartin University, Faculty of Economics and Administrative Sciences, Department of Business Administration, BARTIN, Turkey

Received: January 9, 2013, Accepted: December 17, 2013
Corresponding author: Fatih UCKARDES, e-mail: fatihuckardes@gmail.com

Abstract

The aim of this study was to develop a flow diagram showing which test statistic should be applied to the data obtained before starting the study or after the study without dealing with complex test statistics such as multivariate and experimental designs. Flow diagrams were prepared for commonly used parametric tests while for non-parametric tests a table was prepared to show differences in their own. It is aimed for researchers to be able to make a decision quickly and accurately while deciding which test statistic to select and what matters to be considered by following these flow diagrams and the table. As a result of this study, 25 non-parametric, 5 parametric test statistics and pairwisecomparison, multiplecomparison, normality and the homogeneity of variance test statistics were included as well.

Keywords: Test statistic, parametric test, non-parametric test

İstatistik test seçimine hızlı bir yaklaşım

Özet

Bu çalışmanın amacı, çok değişkenli ve deneme desenleri gibi kompleks test istatistiklerine girmeden bir araştırıcının çalışmaya başlamadan önce ya da çalışmayı bitirdikten sonra elde ettiği verilere hangi test istatistiğini uygulaması gerektiğini gösterecek akış diyagramları geliştirmekti. Akış diyagramı parametrik parametrik olmayan testlere göre hazırlanmış ancak parametrik olmayan testler kendi içindeki farklılıkları göstermek için çizelge şeklinde hazırlanmıştır. Araştırıcı bu çizelge ve akış diyagramlarını takip ederek hangi hususlara dikkat etmesi ve hangi test istatistiğini seçmesi gerektiğine hızlı ve doğru bir biçimde ulaşması hedeflenmiştir. Bu çalışma sonucunda 25 parametrik olmayan, 5 parametrik test istatistiği ve ayrıca çoklu karşılaştırma, ikili karşılaştırma, normallik ve varyans homojenlik test istatistiklerine de yer verilmiştir.

Anahtar Kelimeler: İstatistik test, parametrik test, parametrik olmayan test

Introduction

Researchers prepare tables and graphs to make some particular data obtained from the result of their experiments which are suitable for the purpose more understandable. Besides that they require generalizing the data to populations and interpreting them. For this purpose, they make use of the test statistic in order to interpret the data correctly.

Researchers can also compare their own study results with previously done studies or can defend their studies reliability (testability) owing to these tests. Therefore, the test statistics are used extensively in almost every field of science.

Researchers generally benefit from similar studies previously conducted while designing their study. The fulfillment of certain basic rules is important in the selection of a test statistic. Selection of the test statistic without knowing these basic rules can direct the researcher to erroneous conclusions.

The aim of this study was to develop a flow diagram showing which test statistic should be applied to the data obtained before starting the study or after the study without dealing with complex test statistics such as multivariate and experimental design. Flow diagrams were prepared for commonly used parametric tests while for non-parametric tests a table was prepared to show differences in their own. It is aimedforthe researcher to be able to make a decision quickly and accurately while deciding which test statistic to selectand what matters to be considered by following these flow diagrams and the table.

Method

Data is the information obtained from the researchers' study results and it is divided into two as discrete and continuous data as shown in Figure 1.

Non-parametric tests are more suitable for discrete data since the data is obtained by counting. Nominal scale which is a discrete data is the simplest one in scaling and consists of the binary number system as “0-1”or as “yes-no”, and non-parametric tests are suitable for data. As for ordinal scale, there is a particular ranking, andnon-parametric tests are also more suitable for such data. However, the data is turned into a continuous form while asking some questions to get an idea about a specific topic and combine them to obtain an average score by using an ordinal scale in survey studies. In this case, parametric tests can be used for this kind of data.

Continuous data which are obtained by measuring and weighing are more suitable for parametric tests. Ratio scale which is a continuous data is a scale measurement used for the measurements when the absolute initial point value is other than zero. As for interval scale, itis akind of scale measurement which used when the absolute initial point is zero.

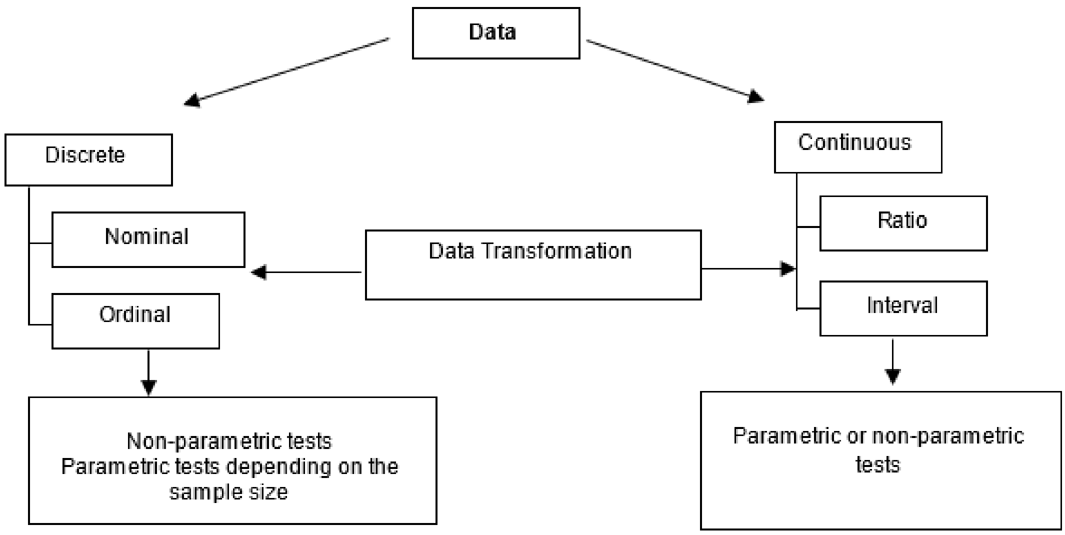

Figure 1. Classification of data

In general, the interval scale is used more widely than the ratio scale. However it is possible to evaluate the continuous measured data by dividing into specific categories. The discrete data group transforms into nominal or ordinal kind of data as a result of these categories. As it is seen a number of transformations can be done on the data. Similarly, non-parametric tests can be used depending on the sample size for such data. In this case, the continuous data is transformed into discrete data. However, instead of distribution-free non-parametric tests which are alternative to parametric tests should be used when the data is not normally distributed and the variances are not homogeneous (Sheskin, 2004; Üçkardeş, 2006).

The non-parametric tests which use discrete data are not included in the following flow diagrams. The flow diagrams created by taking continuous data into account. The way to be followed for non-parametric tests are presented in tabular form. In this study, the test designs for one-sample, two-sample and more than two samples were prepared.

Conclusion

Flow diagrams were prepared for continuous data, for one-sample, two-sample and k-sample, respectively. In selecting test, data type, sample size, whether the data is normally distributed, knowing population parameters and group variances to be homogeneous is important. The flow diagram for one-sample tests are shown in Figure 2. Normality, sample size and population variances are important criteria in test selection. These terms will be briefly described. There are various tests which vary depending on the sample to determine whether the distribution of the data is normally distributed. The Kolmogorov Smirnov test is suitable for the sample size n>50, while Shapiro-Wilks W normality test is suitable for the sample size n<50 (Uckardes, 2006). As a result of normality tests, if the data set is not distributed normally, an appropriate transformation technique is chosen depending on whether the data type is discrete or continuous. Logarithmic transformation is applied for continuous and right-skewed data, square root transformation are applied to discrete data, and arc sin transformations are applied for discrete and ratio data (Rohlf and Sokal, 1969; Duzgunes et al., 1987). Nevertheless, non-parametric tests should be preferred instead of parametric tests if the normal distribution is still not achieved as a result of these transformations (Table 1).

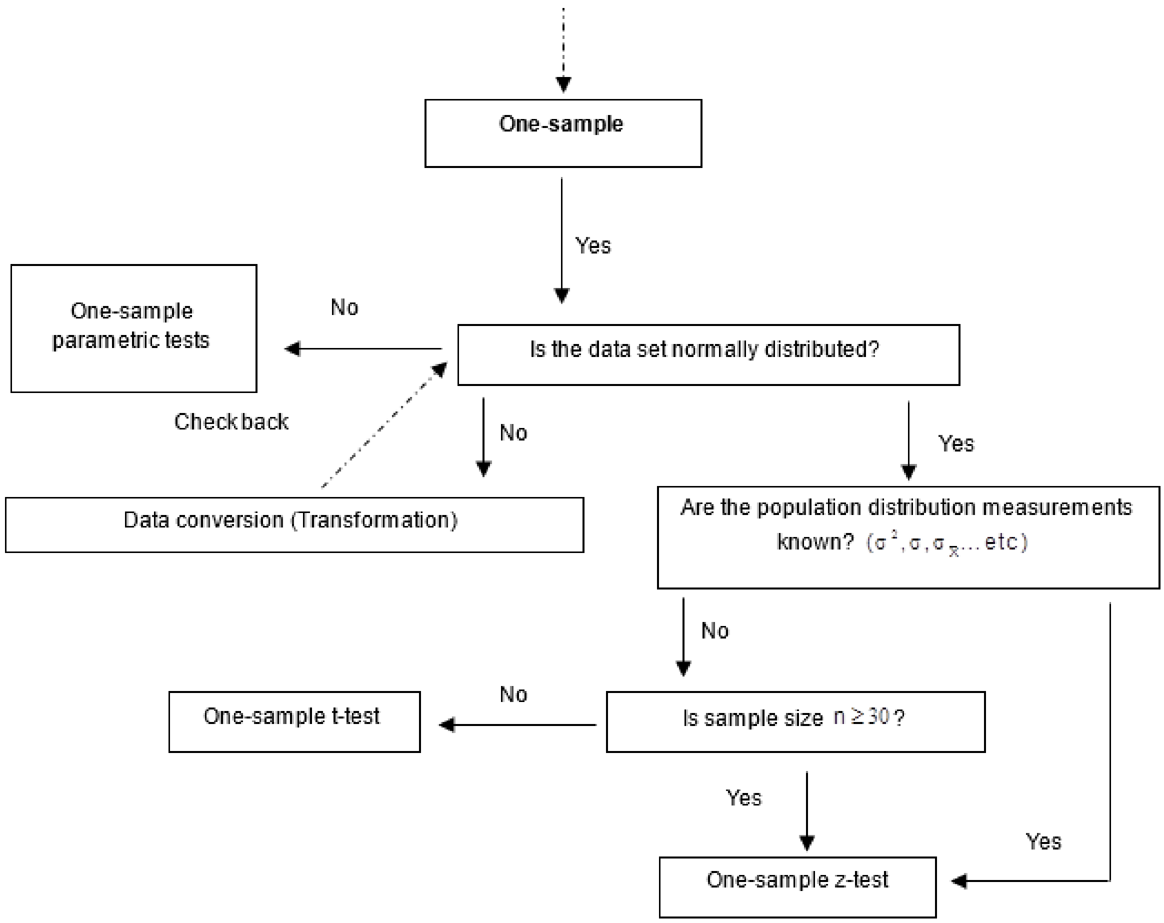

Figure 2. Flow diagram for one-sample tests

The flow diagram for two-sample tests are shown in Figure 3. Different from Figure 2, group variances and homogeneity are shown in Figure 3. Homogeneity of variance tests are used to determine whether the distributions of these groups are similar. As an example to these tests; Hartley F max, Cochran, Bartlett and Levene tests can be given, respectively (Snedecor and Cochran, 1967; Rohlf and Sokal, 1969; Uckardes, 2006). As a result of these tests, suitable transformations are applied if the homogeneity of the data is not achieved. Non-parametric tests are used if the homogeneity is still not achieved even after this transformation. It is found that the group with bigger mean is more effective when there is a significant difference between the groups as a result of the test statistic. However, the values of unwanted results may have bigger mean in some studies. In this case, the researchers need to be careful when interpreting the results. In addition to these, in the case of taking multiple measurements of the same experimental subject or multiple different applications on the same subject, a dependency occurs in obtained data. Therefore, the tests required for dependent samples should be selected. In independent two-sample tests, there is a comparison of two different groups with each other and there is independence between the groups (Figure 3).

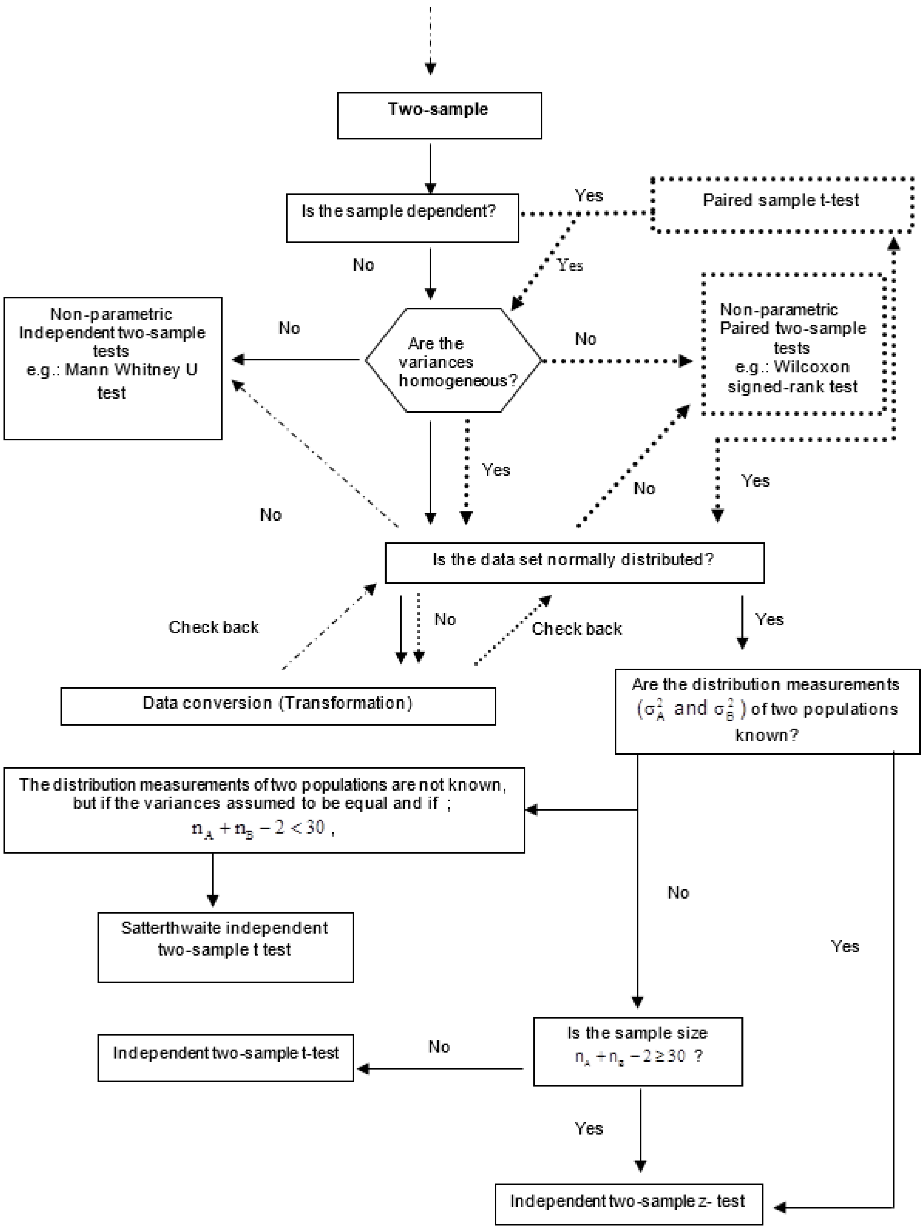

Figure 3. Flow diagram for two-sample tests

The flow diagram for k-sample tests are shown in Figure 4. When there are more than two groups to compare, the analysis of variance test (ANOVA) is used to determine significant differences among the means. To perform this analysis, homogeneity, normality and additivity assumptions should be provided (Winer, 1971; Ferguson, 1981; Bek and Efe, 1988). In the case of finding differences between groups as a result of analysis of variance, the multiple comparison tests (Duncan and SNK), or a pair wise comparison tests (Tukey and LSD) are used to determinate which group is different (Roscoe, 1975; Bek and Efe, 1988; Efe et al., 2000; Orhan et al., 2004; Köklü et al., 2006; Uckardes, 2006). Different group or groups can be determined as a result of these tests.

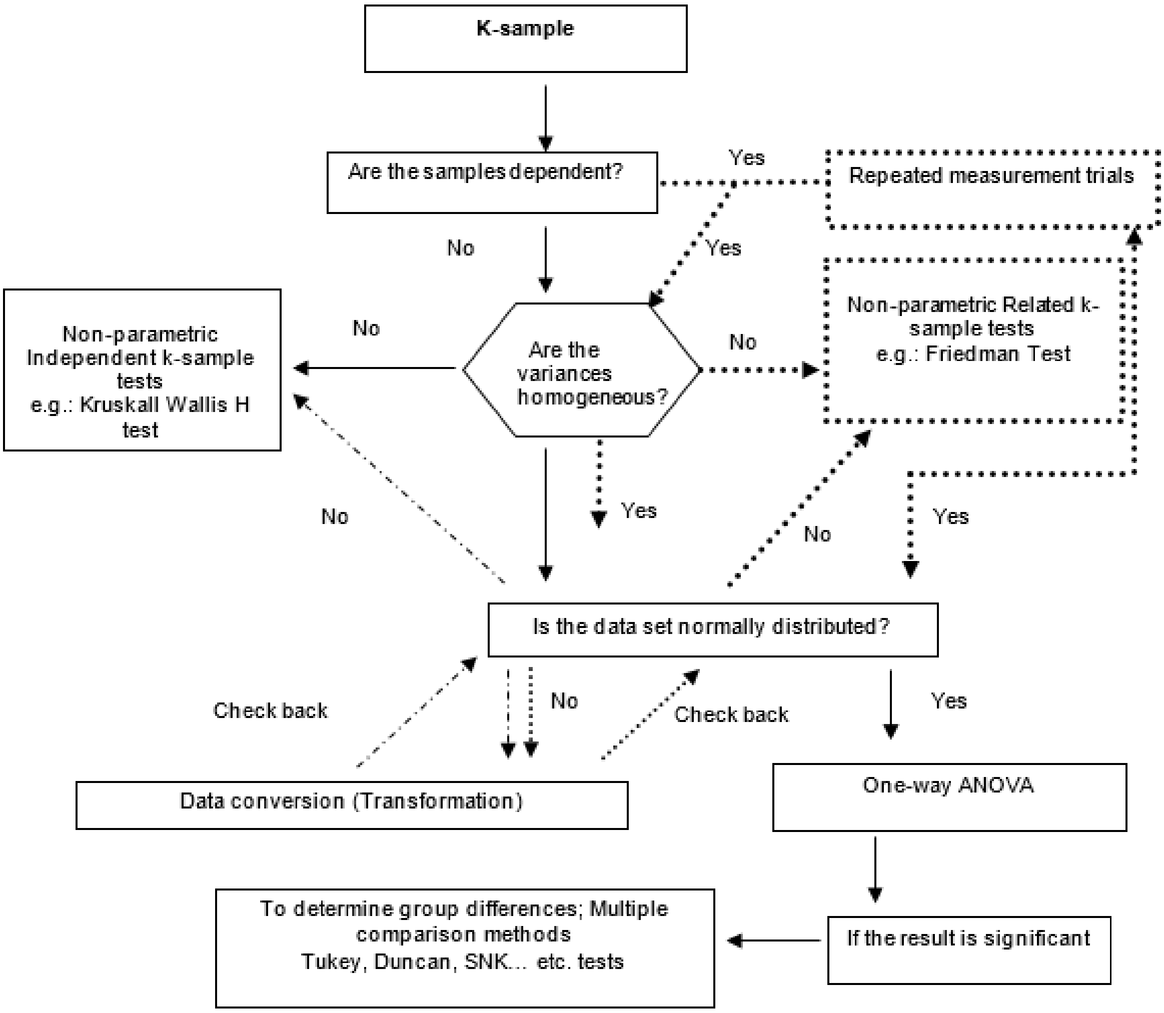

Figure 4. Flow diagram for k-sample tests

Non-parametric tests

To use non-parametric tests, the assumption of normality and homogeneity of variance requirement is not needed. However, there are given below some important explanation in selecting non-parametric tests, respectively.

1. The test statistic varies depending on the data scale to beused.

2. The test statistic varies depending on whether the samples are dependent and the number of the groups.

The test statistic to be selected for the non-parametric tests depending on the data scale and the number of the samples is given in Table 1. Tests specified in bold indicates widely preferred tests which are used when parametric tests can not be

3. The test statistic also depends on the non-parametric hypothesis tests to be established.

To explain this with an example; whena researcher investigating the effects of four different doses of a drug establishes his hypothesis on the assumption of learning will increase; to test this hypothesis, using Jonckheere Terpstra rank test as an alternative test rather than using Kruskal-Wallis H test is going to be more accurate since Kruskal-Wallis H test remains very general for this hypothesis (Siegel and Castellan, 1988). For more detailed information, differences are tried to be summarized in Table 1. applied. In Table 1, the test criteria in accordance with the established hypothesis is not included. More detailed informationcan be obtained from the published data.

Table 1. A fast approach to the non-parametric tests (Siegel and Castellan, 1988; Uckardes, 2006)

<table>
<tr><th rowspan="2">Scale type</th><th rowspan="2">One-sample</th><th colspan="2">Two-sample</th><th colspan="2">K-sample</th></tr>
<tr><th>Paired</th><th>Independent</th><th>Paired</th><th>Independent</th></tr>
<tr><td rowspan="2">Nominal</td><td>Binomial Test</td><td rowspan="2">McNemar Test</td><td rowspan="2">Chi-square test for rx2 tables</td><td rowspan="2">Cochran Q Test</td><td rowspan="2">Chi-square test for rxc tables</td></tr>
<tr><td>Chi-square Goodness of Fit Test</td></tr>
<tr><td rowspan="6">Ordinal</td><td>Kolmogorov Smirnov One-sample Test</td><td rowspan="6">Sign Test</td><td>Median test</td><td>Friedman Test</td><td>Multi-Sample Median Test</td></tr>
<tr><td rowspan="5">One-sample runs test</td><td>Robust Rank Order Test</td><td rowspan="5">Page Test</td><td rowspan="5">Jonkheere-Terpstra Test</td></tr>
<tr><td>Kolmogorov-Smirnov Two-sample Test</td></tr>
<tr><td>For measurement differences Siegel-Tukey Test</td></tr>
<tr><td>Wald-Wolfowitz Runs Test</td></tr>
<tr><td>Tukey Duckworth Test</td></tr>
<tr><td rowspan="2">Nominal, Ordinal , Ratio or Interval</td><td rowspan="2">Kolmogorov Smirnov One-sample Test</td><td rowspan="2">Wilcoxon signed-rank test</td><td>Gehan Test</td><td rowspan="2">Friedman Test</td><td rowspan="2">Kruskal-Wallis H Test</td></tr>
<tr><td>Mann-Whitney U Test</td></tr>
</table>

Discussion

In this study, it is tried to provide concise information on matters to be considered and about walkthroughswhile selecting the test statistic. No matter which study it is (appropriate to the content and the purpose of the study), the subject should be described well, hypotheses should be established and the sample size should be determined while selecting an appropriate test statistic. In addition to these, the experimental subjects taken into the study should be similar to each other. Moreover, in the case that homogeneity is not achieved in experiment material the data will not show normal distribution and the intragroup distributions will be bigger than the inter-group distributions. As a result, the results will be non-significant. When homogeneity is not achieved in experimental material, an appropriate experiment should be chosen from the experimental designs which are advanced statistical methods (Randomized blocks, Latin square experimental models, etc.) and then the study must be set up according to the experimental design requirements (Duzgunes et al., 1987; Bek and Efe, 1998, Efe et al., 2000). This topic is beyond the scope of this study; therefore, the necessary information be obtained from reference sources.

Kaynaklar

Bek, Y. ve Efe, E. 1988. Araştırma Deneme Metodları 1. Çukurova Üniversitesi, Ziraat Fakültesi, Ofset ve Teksir Atölyesi, Adana, 395s.

Duzgunes, O., Kesici, T., Kavuncu, O., Gurbuz, F. 1987. Araştırma ve Deneme Metodları. Ankara Üniversitesi, Ziraat Fakültesi Yayınları, Ankara, 381s.

Efe, E., Bek, Y., Şahin, M. 2000. SPSS'te Çözümleri ile İstatistik Yöntemler II. Kahramanmaraş Sütçü İmam Üniversitesi Rektörlüğü Yayınları, Kahramanmaraş, 223s.

Ferguson, G. A. 1981. Statistical analysis in psychology and education. McGraw-Hill Book Company, New York, 587p.

Orhan, H., Efe, E., Şahin, M. 2004. SAS Yazılımı ile İstatistiksel Analizler. Tuğra Ofset, Isparta, 122s.

Köklü, N., Büyüköztürk Ş., Bökeoğlu, Ç.Ö. 2006. Sosyal bilimler için istatistik. Pegema Yayıncılık, Ankara, 203 s.

Roscoe, J. T. 1975. Fundamental research statistics for the behavioral sciences. Holt, Rinehart and Winston, Inc. New York, 483 p.

Siegel, S., Castellan, N.J. 1988. Nonparametric Statistics for the Behavioral Sciences. McGraw-Hill book Company, United States of America, 399s.

Sheskin, D., J,. 2004. Parametric and Nonparametric Statistical Procedures. CRC Press New York, 1193s.

Snedecor, W.G., Cochran, G.W., 1967. Statistical Methods. The lowa State University Press, U.S.A, 593s.

Rohlf, J.F., Sokal, R.R., 1969. Biometry. W.H. Freeman and Company, San Francisco, 776s.

Üçkardeş, F. 2006. İstatistik Testler Üzerine Bir Çalışma. Fen bilimler Enstitüsü. Kahraman Maraş Sütçü İmam Üniversitesi, 249s.

Winer, B. J. 1971. Statistical principles in experimental design. McGraw-Hill Book Company, New York, 928 p.