KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

İSTATİSTİK VE BİLGİSAYAR BİLİMLERİ ANABİLİM DALI

VERİ ANALİZİNDE DOĞRU İSTATİSTİKSEL TESTİN BİLGİSAYAR YARDIMIYLA SEÇİLMESİ

YÜKSEK LİSANS TEZİ

Musa SARİ

HAZİRAN 2017
TRABZON

KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

İSTATİSTİK VE BİLGİSAYAR BİLİMLERİ ANABİLİM DALI

VERİ ANALİZİNDE DOĞRU İSTATİSTİKSEL TESTİN BİLGİSAYAR YARDIMIYLA SEÇİLMESİ

Musa SARİ

Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsünce
"YÜKSEK LİSANS (İSTATİSTİK)"
Unvanı Verilmesi İçin Kabul Edilen Tezdir.

Tezin Enstitüye Verildiği Tarih : 29.05.2017
Tezin Savunma Tarihi : 16.06.2017

Tez Danışmanı: Yrd. Doç. Dr. Tolga BERBER

Trabzon 2017

Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsü

İstatistik ve Bilgisayar Bilimleri Anabilim Dalında

Musa SARİ Tarafından Hazırlanan

VERİ ANALİZİNDE DOĞRU İSTATİSTİKSEL TESTİN BİLGİSAYAR YARDIMIYLA SEÇİLMESİ

başlıklı bu çalışma, Enstitü Yönetim Kurulunun 16 / 05 / 2017 gün ve 1702 sayılı kararıyla oluşturulan jüri tarafından yapılan sınavda
YÜKSEK LİSANS TEZİ
olarak kabul edilmiştir.

Jüri Üyeleri

Başkan : Prof. Dr. Türkan ERBAY DALKILIÇ ………………………………

Üye : Yrd. Doç. Dr. Burcu KARTAL ………………………………

Üye : Yrd. Doç. Dr. Tolga BERBER ………………………………

Prof. Dr. Sadettin KORKMAZ

Enstitü Müdürü

ÖN SÖZ

“Veri Analizinde Doğru İstatistiksel Testin Bilgisayar Yardımıyla Seçilmesi” isimli bu tez Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstatistik ve Bilgisayar Bilimleri Anabilim Dalı, Yüksek Lisans Programı’nda hazırlanmıştır.

Başta tez çalışma süresince değerli yardım ve katkılarıyla beni yönlendiren danışman hocam sayın Yrd. Doç. Dr. Tolga BERBER ‘e, yüksek lisans eğitimi yapma fırsatı bulduğum İstatistik ve Bilgisayar Bilimleri Bölümündeki tüm hocalarıma teşekkürü borç bilirim.

Son olarak tüm hayatım boyunca maddi manevi her zaman beni destekleyen, her adımımda arkamda duran ve bugünlerimin mimarı olan aileme sonsuz teşekkürlerimi sunarım.

Bu tez çalışmasının, bundan sonraki çalışmalara katkı sağlamasını temenni ederim.

Musa SARİ

Trabzon 2017

TEZ ETİK BEYANNAMESİ

Yüksek Lisans Tezi olarak sunduğum "Veri Analizinde Doğru İstatistiksel Testin Bilgisayar Yardımıyla Seçilmesi" başlıklı bu çalışmayı baştan sona kadar danışmanım Yrd. Doç. Dr. Tolga BERBER 'in sorumluluğunda tamamladığımı, verileri kendim topladığımı, analizleri ilgili laboratuvarlarda yaptığımı, başka kaynaklardan aldığım bilgileri metinde ve kaynakçada eksiksiz olarak gösterdiğimi, çalışma sürecinde bilimsel araştırma ve etik kurallara uygun olarak davrandığımı ve aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul ettiğimi beyan ederim. 16/06/2017

Musa SARİ

İÇİNDEKİLER

Sayfa No

Yüksek Lisans Tezi

ÖZET

VERİ ANALİZİNDE DOĞRU İSTATİSTİKSEL TESTİN BİLGİSAYAR YARDIMIYLA SEÇİLMESİ

Musa SARİ

Karadeniz Teknik Üniversitesi
Fen Bilimleri Enstitüsü
İstatistik ve Bilgisayar Bilimleri Anabilim Dalı
Danışman: Yrd. Doç. Dr. Tolga BERBER
2017, 68 Sayfa

Araştırma, bir gerçeği ortaya çıkarmak, bir sorunu çözümlemek ve eldeki verileri arttırmak için bilimsel yöntem ve tekniklerden yararlanılarak yapılan düzenli çalışmadır. Araştırma sonucunda elde edilen verilerin doğru yorumlanabilmesi, araştırmanın amaca uygun şekilde yapılması ve analiz sırasında kullanılan istatistiksel testlerin doğru seçilmesine bağlıdır. Test seçimi, yapılan araştırmanın amacı ve elde edilen verinin; ölçme düzeyi, değişken türü, dağılım biçimi ile varyans homojenliği gibi özellikleri göz önünde bulundurularak yapılır. Araştırmalarda birbirinden farklı ölçme düzeyine sahip birden çok değişken olduğu durumlarda, araştırmacılar için uygun test seçimi karmaşık bir problem haline gelmektedir. Özellikle tıp alanında yapılan bazı çalışmalarda doğru istatistiksel yöntemlerin kullanılmadığını gösteren birçok araştırma mevcuttur. Literatürde test seçimi konusunda araştırmacılara yardımcı olması amacıyla yapılmış çeşitli çalışmalar mevcuttur. Bu çalışmalarda test seçimi için cevaplanması gereken bazı sorular ile karar ağaçları ve tabloların kullanılması önerilmiştir. Bunun yanında bazı istatistiksel paket programlarının veri girişi sonrasında yapılacak testi önerme özelliği bulunmaktadır. Bu çalışmada, istatistiksel test seçimi için bir yöntem önerilmiştir. Geliştirilen test seçim algoritması ile değişkenlerin ölçme düzeyi bilgileri ve örnek verileri kullanılarak araştırma amacına en uygun istatistiksel testin bilgisayar yardımıyla önerilmesi amaçlamıştır. Elde edilen çıktılar, literatürdeki çalışmalar ve test önerisi yapan paket programlar ile karşılaştırılarak sonuçlar verilmiştir.

Anahtar Kelimeler: İstatistiksel Test Seçimi, İstatistiksel Veri Analizi

Master Thesis

SUMMARY

COMPUTER-AIDED SELECTION OF CORRECT STATISTICAL TESTS IN DATA ANALYSIS

Musa SARİ

Karadeniz Technical University
The Graduate School of Natural and Applied Sciences
Statistical and Computer Science Graduate Program
Supervisor: Asst. Prof. Dr. Tolga BERBER
2017, 68 Pages

Research is the systematic work of using scientific methods and techniques to reveal a truth, to solve a problem and to increase the amount of data available. The exact clarification of the data obtained from the results of research depends on the suitability of the research methods and the correct selection of the statistical tests used during the analysis. The choice of statistical tests should be based on the purpose of the study conducted and the properties of acquired data such as measurement level, variable type, distribution form and homogeneity of variance. The choice of the proper test becomes a complex problem for researchers, when there are multiple variables with different levels of measurement in study. Especially, there are some studies in literature proving that some statistical methods are used inappropriately in medical studies. There are various studies in the literature to help researchers in selecting the proper statistical test. In these studies, it has been suggested to use some questions with decision trees and tables for statistical test selection process. In addition, some statistical package programs can suggest tests to be performed after data entry. In this study, a method for statistical test selection was proposed. Using the proposed test selection algorithm, we aimed to suggest the most suitable statistical test for the research by using measurement level and sample data with the help of computer. The obtained outputs are compared with the literature and the package programs capable to suggest statistical test, and the results are given.

Key Words: Statistical Test Selection, Statistical Data Analysis

ŞEKİLLER DİZİNİ

<u>Sayfa No</u>

TABLOLAR DİZİNİ

Sayfa No

SEMBOLLER DİZİNİ

H_0	: Yokluk hipotezi
H_1	: Karşıt hipotez
G_j	: j. sınıftaki gözlenen frekans
B_j	: j. sınıftaki beklenen frekans
χ^2	: Ki-kare test istatistiği
$F(x)$	: Birikimli dağılım fonksiyonu
X	: Rastgele değişken
$S(x)$	: Gözlenen birikimli dağılım fonksiyonu
$\bar{X}$	: Örneklem ortalaması
S^2	: Örneklem varyansının tahmin edicisi
N	: Örneklem hacmi
σ^2	: Örneklem varyansı
VAP	: Veri Analiz Platformu
$GAKT$	: Gruplar arası kareler toplamı
$GIKT$	: Grup içi kareler toplamı
$GAOK$	: Gruplar arası ortalama kare
$GIOK$	: Gruplar içi ortalama kare
$F_{\alpha,\beta}$	: α, β serbestlik dereceli F dağılımı test istatistiği
z	: Normal dağılım test istatistiği
t	: Student t dağılımı test istatistiği
$KÇÇT$	: Kareler ve çapraz çarpımlar toplamı matrisi
Ö1	: Sınıflama ölçme düzeyi
Ö2	: Sıralama ölçme düzeyi
Ö3	: Eşit aralıklı ölçme düzeyi
Ö4	: Oranlama ölçme düzeyi
$R(X_i)$	: X örneğinin i. sıra sayısı
GÜ	: Giresun Üniversitesi
KTÜ	: Karadeniz Teknik Üniversitesi
ODTÜ	: Orta Doğu Teknik Üniversitesi
$\bar{R}_{.j}$	: j. örnek için sıra sayıları ortalaması

1. GENEL BİLGİLER

1.1. Giriş

Araştırma, bir gerçeği ortaya çıkarmak, bir sorunu çözümlemek ve eldeki verileri arttırmak için bilimsel yöntem ve tekniklerden yararlanılarak yapılan düzenli çalışmalardır [1]. Araştırma sonucunda elde edilen veriler istatistiksel anlamda yorumlanırken; değişkenlerin ölçme düzeyi, değişken türü (bağımlı veya bağımsız değişkenler), dağılım biçimi ve varyans homojenliği gibi özellikleri göz önünde bulundurulur. Veriye ait bu özellikler istatistiksel test seçimini etkileyen en önemli faktörlerdir.

Birimlerin farklı değerler alabildikleri nitelik ve niceliklerine değişken denir. Belirli bir yığın için çok sayıda değişken tanımlamak mümkündür. Örneğin Trabzon ilinde kamu ve özel hastanelere başvuranlar yığınını düşünelim. Bu yığın tedavi amacıyla başvuran hastalardan oluşur. Bu hastalar için yaş, cinsiyet, hastaneye geliş sayısı, hastane hizmetlerinden beklenti, hemşirelerin tutumlarını beğenip beğenmeme durumu, meslek, eğitim, ağırlık, medeni durum ve doktor tercihi birer değişken olarak düşünülebilir. Aynı yığından değişkenlerin sayısını arttırmak da mümkündür. Araştırmalarda en önemli aşamalardan biri ilgilenilen değişkenin nasıl ölçüleceğidir. Duyarlılıklarına göre değişkenlerin ölçülmeleri dört grupta toplanabilir [2].

1.1.1. Sınıflama (Nominal) Ölçme Düzeyi

Sınıflama ölçme düzeyinde birimlere verilen sayı sadece bir ad olarak düşünülür. Bu nedenle sınıflama ölçme düzeyi çok kaba bir ölçme düzeyidir. Birimlere sayı yerine büyüklük sırasını ifade etmeyen Romen rakamları, Latin harfleri ya da isimler vermenin de hiçbir sakıncası yoktur. Öğrencilerin baba mesleklerine, bitirdikleri okullara, taraftarı oldukları spor kulüplerine ve okula gelirken kullandıkları ulaşım araçlarının türlerine göre gruplandırılması sınıflama ölçme düzeyine birer örnek olarak verilebilir.

1.1.2. Sıralama (Ordinal) Ölçme Düzeyi

Sıralama ölçme düzeyi sınıflama ölçme düzeyine göre daha hassas bir ölçme düzeyidir. Sıralama ölçme düzeyinde birimlere verilen sayılar sadece birimlerin farklı olduklarını göstermeyip, ilgili değişken bakımından birimlerin büyüklük sırasını belirlemek amacıyla da kullanılmışlardır. Örneğin Can, Cem, Ali ve Akın isimli öğrencilere bir değişken bakımından sırasıyla 4, 7, 7 ve 12 sayıları verilmiş olsun. Bu sayılar ölçülen değişken bakımından Cem ve Ali'nin aynı değerli olduğunu, Can'ın en düşük değerli olduğunu ve Akın'ın da en büyük değerli olduğunu ifade eder. Bu tip ölçmeye sıralama ölçme düzeyi denir.

1.1.3. Eşit Aralıklı (Interval) Ölçme Düzeyi

Eşit aralıklı ölçme düzeyi sınıflama ve sıralama ölçme düzeylerine göre daha hassas bir ölçmeyi sağlar. Değişkenlerin eşit aralıklı düzeyde ölçülebilmeleri için birimlere verilen sayılar arasındaki farklara matematiksel anlam kazandırmak gerekir. Eşit aralıklı ölçme düzeyinde sayılar bir birimle ifade edilir. Örneğin İstatistik Bölümü öğrencilerinden Cem'in araştırma projesi dersinden 60 puan alması gibi. Bu ölçme düzeyinde birimlere verilen sayılar matematiksel olarak önceki iki ölçme düzeyine göre daha uygundur. Eşit aralıklı ölçme düzeyindeki verilere veya değişkenlere toplama, çıkarma, çarpma ve bölme yapılabilir. Fakat önceki iki ölçme düzeyinde bu işlemler yapılamaz.

1.1.4. Oranlama (Ratio) Ölçme Düzeyi

Birimler eşit aralıklı düzeyde ölçüldükleri zaman birimlerin aldıkları değerler arasındaki farkların anlamlı oldukları belirtilmişti. Ancak birimlerin aldıkları bu değerlerin oranlanması eşit aralık ölçme düzeyinde doğru olmayabilir. Eşit aralıklı düzeyde yapılan bir ölçme mutlak bir sıfır noktasına göre yapılabiliyorsa bu ölçme düzeyine oranlama ölçme düzeyi denir. Örneğin ağırlık, uzunluk, zaman, hız gibi değişkenler için mutlak sıfır noktasının tanımı yapılmıştır. Oranlama düzeyinde ölçülebilen değişkenler için ailedeki çocuk sayısı, gelir, aylık harcama, konut sayısı, televizyon sayısı, tüketilen su miktarı örnek olarak verilebilir.

Araştırmada birden fazla değişken olduğu durumlarda araştırmacı, veri analizine başlamadan önce bağımlı ve bağımsız değişkenleri belirlemelidir. İlişkili iki değişkenden biri diğeri aracılığı ile (diğerine bağlı olarak) betimlenecek ya da açıklanacak ise açıklanan temel değişken bağımlı değişken bu değişkenle ilişkilendirilen diğer değişken ise bağımsız değişken olarak adlandırılır.

Test seçiminde diğer bir önemli faktör veri kümesinin dağılım biçimidir. Veri kümesinin normal dağılıma uygun olduğu durumlarda parametrik istatistiksel yöntemler aksi durumlarda ise parametrik olmayan istatistiksel yöntemlerin kullanılması önerilir. Ayrıca parametrik testlerin kullanılması için varyans homojenliği kontrolünün de yapılması gerekir. Parametrik olmayan testler normallik ve varyans homojenliği kontrolü gerektirmez.

Özet olarak istatistiksel test seçimini etkileyen en önemli [3] faktörler:

1) Araştırmanın amacı
2) Değişkenlerin ölçme düzeyi
3) Değişkenlerin türü (bağımlı, bağımsız değişkenler)
4) Sayısal değişkenlerin normal dağılıma uygunluğu ve varyans homojenliği

şeklinde sıralanabilir. Bu faktörlerin kontrolleri yapıldıktan sonra araştırmanın amacına uygun olan istatistiksel test seçilerek çalışma tamamlanır.

İstatistiksel bir araştırmada amaç:

1) Verinin istatistiksel olarak tanımlanması,
2) Hipotetik bir değerin veri grubuyla karşılaştırılması,
3) İki veya daha fazla grubun karşılaştırılması,
4) İki değişken arasındaki ilişkiyi ölçmek,
5) Tahminde bulunmak.

şeklinde tanımlanabilir.

2. BAZI İSTATİSTİKSEL YÖNTEMLER

Bu çalışmaya dahil edilen istatistiksel testler seçilirken, Türkiye'de istatistik bölümü olan üniversitelerde görev yapan akademisyenler ile bazı kamu kurumlarında görev yapan istatistik uzmanlarınca gerçekleştirilmesi için bir anket gerçekleştirilmiştir. Ankete davet edilen 450 kişiden, 75 kişi anketi cevaplandırmıştır. Katılımcılar yaptıkları çalışmalarda sıklıkla kullanmak durumunda kaldıkları istatistiksel testler sorulmuş ve elde edilen sonuçlar yorumlanarak bu araştırmaya dahil edilen testler belirlenmiştir. İlgili testler gerekli varsayımlar ve kullanım amaçları doğrultusunda incelenerek test seçim algoritması geliştirilmiştir. Bu testler ve test istatistikleri aşağıda açıklanmıştır.

2.1. Normallik Testleri

Parametrik testlerin varsayımlarından biri de örneklerin seçildiği yığınların dağılım biçiminin normal olduğudur. Bu nedenle parametrik testlerin kullanılması için öncelikle örnek verisine, verinin yapısına uygun bir normallik testinin uygulanması gerekmektedir [2].

2.1.1. Ki-Kare Uyum İyiliği Testi

Bazı araştırmalarda örneğe çekilen birimlerden bir değişken bakımından sınıflama düzeyinde bilgi toplanabilir. Bu tür durumlarda örneğin geldiği yığının dağılım biçiminin belirlenmesi amacıyla ki-kare uyum iyiliği testi kullanılır. Bir iş sınavına başvuranların bitirdikleri okullara göre dağılımı bu duruma örnek olarak verilebilir. Temel özellik değişkenin sınıflama ölçme düzeyinde ölçülmüş olmasıdır. Yukarıdaki örnekte değişken bitirilen okuldur. Sınıflama düzeyinde ölçülen değişkenin alabileceği değerlere "sınıflar" ya da "kategoriler" adı verilir. Sınıflar sözel ya da sayısal olabilir. Yukarıdaki örnekte iş sınavına başvuranların bitirdikleri okullar GÜ, KTÜ, ODTÜ, … vb. şeklinde olacağından değişken sözel değerler alacaktır.

Rastgele seçilen *n* hacimli örnekten sınıflama düzeyinde ölçülen ilgili değişken bakımından bilgi toplandığında veri tablosu Tablo 1'deki gibi oluştur [2].

Tablo 1. Ki-Kare Gözlenen Frekans Tablosu

Sınıflar	**1**	**2**	**3**	**…**	**j**	**…**	**c**	**Toplam**
Gözlenen frekanslar	G_1	G_2	G_3	$\dots$	G_j	$\dots$	G_c	n

c, sınıf sayısı olmak üzere, ki-kare uyum iyiliği testinde yokluk ve karşıt hipotezler genel olarak aşağıdaki gibidir.

H_0 : Örnek belirli bir dağılıma sahip yığından seçilmiştir.
H_1 : Örnek yokluk hipotezinde belirtilen dağılımdan seçilmemiştir.

Yokluk hipotezi doğru ise yığından seçilen örneğin yığının karakteristiklerini taşıması beklenir. Örnek söz edilen dağılımdan seçilmiş ise sınıfların her birinde beklenen frekanslar ile gözlenen frekansların birbirine eşit ya da yakın olması beklenir. Test istatistiği,

$$\sum_{j=1}^{c} \frac{\left(G_j - B_j\right)^2}{B_j} \sim \chi^2 \tag{1}$$

olarak tanımlanır.

Yokluk hipotezi doğru ve öngörülen dağılım için parametre tahmini yapılmadıysa yukarıda tanımlanan istatistik, yaklaşık olarak $c-1$ serbestlik dereceli ki-kare dağılımına sahiptir. Test işleminde karar şu şekilde verilir. χ^2_{c-1} istatistiği için hesaplanan değer χ^2_h ve α anlamlılık düzeyinde tablo değeri $\chi^2_{c-1,\alpha}$ olmak üzere,

$\chi^2_h > \chi^2_{c-1,\alpha}\ ise\ H_0\ reddedilir.$
$\chi^2_h \leq \chi^2_{c-1,\alpha}\ ise\ H_0\ reddedilemez.$

Test istatistiğinin yaklaşık olarak $c-1$ serbestlik dereceli ki-kare dağılımına sahip olması için örnek hacminin büyük olması gerekir. Çoğu durumda örnek hacminin en az 30 olması yeterli kabul edilmektedir. Ancak hiçbir sınıfta beklenen frekansın değeri 1'den küçük olmamalı ve değeri 1 ile 5 arasında olan beklenen frekansların oranı %20'yi geçmemelidir. Bu durum ortaya çıkarsa bazı sınıflar birleştirilebilir.

Ki-kare uyum iyiliği testi normal dağılıma uygunluk için kullanıldığında *n* hacimli örneğin parametreleri bilinmeyen bir normal dağılımdan geldiği şeklindeki bir yokluk hipotezini test edebilmek için öncelikle normal dağılımın parametrelerini tahmin etmek gerekir. Bu tahmin işlemi için örnekteki bilgi kullanılır. Normal dağılımın ortalamasının ve varyansının örnek verisiyle tahmin edilmesi bu durumda ilgilidir. Bu işlemden sonra sınıflara ilişkin olasılıkların, beklenen frekanslarını ve test istatistiğinin değerinin hesaplanmasında bir farklılık yoktur. Farklılık sadece serbestlik derecesinin değişmesindendir. Bu durumda tahmin edilen parametre sayısı k ise test istatistiği olan ki-karenin serbestlik derecesi $c - k - 1$ olacaktır.

Ki-Kare uyum iyiliği testi için varsayımlar;

1) Değişken sınıflama ya da sıralama ölçme düzeyinde ölçülmüş kategorik bir veri olmalı,
2) Her sınıfta beklenen frekans 1'den büyük olmalı ve değeri 1 ile 5 arasında olan beklenen frekansların oranı %20'yi geçmemeli,
3) Örnek hacmi en az 30 olmalı.

2.1.2. Kolmogorov-Smirnov Uyum İyiliği Testi

Kolmogorov-Smirnov uyum iyiliği testi oranlama ya da eşit aralıklı düzeyde ölçülen değişkenler için kullanılır. Bir örnek Kolmogorov-Smirnov uyum iyiliği iki birikimli dağılım fonksiyonuna dayanır [4]. Birincisi yokluk hipotezinde verilen dağılımın birikimli dağılım fonksiyonudur. İkincisi ise örnekten elde edilen gözlenen birikimli dağılım fonksiyonudur. Birikimli dağılım fonksiyonu $F(x), X$ rastgele değişkeninin değerinin *x* değerine eşit ya da daha küçük olması olasılığıdır. Bu matematiksel olarak,

$$F(x) = P(X \leq x) \tag{2}$$

biçiminde ifade edilebilir.

Birikimli dağılım fonksiyonu bilinmeyen bir yığından n hacimli rastgele bir örnek seçilmiş olsun. Bu durumda veri $x_1, x_2, \ldots, x_n$ gibi bağımsız gözlem sonuçlarından oluşacaktır. $F_0(x)$ hipotezde belirtilen dağılımın birikimli dağılım fonksiyonu iken, yokluk hipotezi ve karşıt hipotezler aşağıdaki gibidir.

$$H_0 : F(x) = F_0(x), \text{tüm } x \text{ değerleri için}$$
$$H_1 : F(x) \neq F_0(x), \text{en az bir } x \text{ değerleri için}$$

Gözlenen birikimli dağılım fonksiyonu $S(x)$ olmak üzere: $S(x)$, x değerine eşit ya da daha küçük değerli örnek birimlerinin sayısının örnek hacmine oranıdır. İki yanlı bir test için bir örnek Kolmogorov-Smirnov test istatistiği eşitlik (3) ile ifade edilmiştir.

$$D = \sup_x |S(x) - F_0(x)| \tag{3}$$

Supremum (sup) sayılabilir kümeler içinde en büyük olarak düşünülebilir. D istatistiğinin örnekleme dağılımından elde edilen kritik değerler tablosu Kolmogorov (1933) tarafından yayınlanmıştır [10]. Bu tablodan bulunacak olan kritik değer D_k olsun. D_k değeri iki-yanlı test, n ve $1-\alpha$ değerlerine göre ilgili tablodan bulunabilir. D istatistiği için örnekten hesaplanan değer D_h olmak üzere, karar kuralı aşağıdaki gibi yazılabilir.

$$D_h \geq D_k \; ise \; H_0 \; reddedilir.$$
$$D_h < D_k \; ise \; H_0 \; reddedilemez.$$

$F_0(x)$ sürekli iken $S(x)$ kesiklidir. Bu nedenle seçilen bir örnek için X'in değer almadığı herhangi bir aralığın iki ucunda $|S(x) - F_0(x)|$ için iki farklı sonuç vardır. Bu özellik nedeniyle $F_0(x)$ sürekli ve $S(x)$ kesikli iken bir örnek Kolmogorov-Smirnov istatistiğinin değeri eşitlik (4) ile ifade edilmiştir.

$$D = \sup\{|S(x_j) - F_0(x_j)|, |S(x_{j-1}) - F_0(x_j)|\} \tag{4}$$

Kolmogorov-Smirnov uyum iyiliği testinin normal dağılıma uyum için kullanılması için örneklem verileri standardize edilerek elde edilen normalleştirilmiş veriler standart normal dağılım ile karşılaştırılır. Fakat bu dönüşümden kaynaklı olarak testin gücünün azaldığı ispat edilmiştir [5]. Kolmogorov-Smirnov istatistiğine yapılan Lilliefors düzeltmesi ile bu durum iyileştirilmiştir.

Kolmogorov-Smirnov uyum iyiliği testi için varsayımlar;

1) Veri, oranlama ya da eşit aralıklı ölçme düzeyinde ölçülmüş olmalıdır,
2) Örnek hacmi N için, $N > 2000$ olduğu durumlarda önerilir.

2.1.3. Lilliefors Normallik Testi

Lilliefors testi temelde Kolmogorov-Smirnov testine dayanmaktadır. Normal dağılıma uygunluk testinde öngörülen normal dağılımın parametreleri bilinmiyorsa ve bunlar örnek verisinden tahmin ediliyorsa normal dağılıma uygunluk için Lilliefors testinin kullanımı önerilmektedir [6]. Bu test için hipotezler aşağıdaki gibi ifade edilir.

H_0 : n hacimli örnek ortalaması ve varyansı bilinmeyen bir normal dağılımdan gelmiştir.

H_1 : n hacimli örnek bir normal dağılımdan gelmemiştir.

n hacimli örnek $X_1, X_2, \dots, X_n$ olsun. H_0 hipotezinde öngörülen normal dağılımın bilinmeyen parametreleri μ ve σ^2 parametrelerinin tahmin edicileri, sırasıyla

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}, S^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1} \tag{5}$$

ile gösterilsin. H_0 hipotezinde öngörülen normal dağılımın dağılım fonksiyonu $F_0(x)$ ile ifade edilsin. Bu $F_0(x)$ fonksiyonu oluşturulurken μ ve σ^2 yerlerine bunların tahmin değerleri, yani $\bar{X}$ ve S^2 kullanılır. Buna göre dağılımın normalliği bilgisi ile,

$$F_0(x) = P(X < x) = P\left(Z < \frac{X - \bar{X}}{S}\right) \tag{6}$$

fonksiyonu elde edilir. Lilliefors testinde $S(x)$ fonksiyonu Z değerlerine dayanır. Bu nedenle,

$$S(x) = \frac{\begin{array}{c} z \text{ değerine eşit ya da daha} \\ \text{küçük değerli örnek birimlerin sayısı} \end{array}}{n} \tag{7}$$

$F_0(x)$ fonksiyonu sürekli ve $S(x)$ fonksiyonu da kesikli olduğundan normallik için Lilliefors istatistiği,

$$D_h = Sup\{|S(x_j) - F_0(x_j)|, |S(x_{j-1}) - F_0(x_j)|\} \tag{8}$$

biçiminde elde edilmiştir.

Lilliefors testi için varsayımlar;

1) Kitleye ilişkin parametreler bilinmiyorsa ve örnekten tahmin edilecekse kullanılır,
2) Örnek hacmi n için, $n < 30$ olduğu durumlarda kullanılması önerilir.

2.1.4. Shapiro-Wilk Normallik Testi

Shapiro-Wilk normallik testi örnek verisinin, parametreleri bilinmeyen bir normal dağılımdan geldiği durumlarda kullanılır [7]. Bu test için hipotezler,

H_0: $F(x)$, ortalaması ve varyansı bilinmeyen bir normal dağılım fonksiyonudur.
H_1: $F(x)$, ortalaması ve varyansı bilinmeyen bir normal dağılım fonksiyonu değildir.

n hacimli rastgele bir örnek $X_1, X_2, \dots, X_n$ olsun ve bu rastgele örnek için sıralı istatistikler de $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ olsun. Ölçüm değerleri ile aritmetik ortalamanın farklarının kareleri toplamı,

$$D = \sum_{i=1}^{n} (X_i - \bar{X})^2 \tag{9}$$

ile ifade edilsin.

$k = \frac{n}{2}$ olmak üzere $a_1, a_2, \dots, a_n$ sabitleri S. S. Shapiro ve M. Wilk (1965) tarafından yayınlanmış olan tablodan [7] bulunur ve W ile gösterilen Shapiro-Wilk istatistiği eşitlik (9)'daki gibi tanımlanır.

$$W = \frac{1}{D}\left[\sum_{i=1}^{k} a_i (X_{(n-i+1)} - X_{(i)})\right]^2 \tag{10}$$

Normallik için Shapiro-Wilk testinde, W istatistiği için hesaplanan değer W_h olmak üzere, $P(W < W_h)$ değerine göre karar verilir.

$$P(W < W_h) < \alpha \tag{11}$$

ise H_0 hipotezi reddedilir.
Shapiro-Wilk normallik testi için varsayımlar;

1) Örnek verisinin parametreleri bilinmiyorsa ve örnekten tahmin edilecekse kullanılır,
2) Örnek hacmi n için, $n > 30$ olduğu durumlarda kullanılması önerilir.

Yukarıda verilen bilgiler doğrultusunda bu çalışmada normallik testi seçimi Tablo 2'ye göre gerçekleştirilmiştir.

Tablo 2. Normallik Kontrolü için Test Seçimi

Değişkenlerin Ölçme Düzeyi	Örnek Hacmi (n)	Test
Eşit Aralıklı veya Oranlama	$n < 30$	Lilliefors Testi
	$n > 30$	Shapiro-Wilk Testi

Sınıflama veya Eşit Aralıklı	-	Ki-Kare Uyum İyiliği

2.2. Varyans Homojenliği Testleri

Varyansların homojenliği varsayımı veri analizinde uygun test istatistiğinin belirlenmesi için önemlidir. Örneğin varyans analizi problemlerinde F testinin (Anova) kullanımı varyans homojenliği varsayımına dayanır [2]. Varyans homojenliği için yapılan testlerde hipotezler,

$$H_0 : \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2 = \sigma^2$$
$$H_1 : \sigma_j^{2\prime}\text{lerin en az biri farklıdır}$$

biçiminde ifade edilir.

H_0 hipotezinin testi için k sayıda yığından birer rastgele örnek seçilsin, j. yığından seçilen örneğin hacmi $n_j, j = 1,2, \dots, k$ ve $i = 1,2, \dots, n_j$ olmak üzere bu örnekler X_{ij} ile gösterilsin. k sayıda örnek için gösterim Tablo 3'te verilmiştir.

Tablo 3. Verinin Örneklere Göre Dağılımı

Örnek No			
1	2	…	k
X_{11}	X_{12}	…	X_{1k}
X_{21}	X_{22}	…	X_{2k}
⋮	⋮	⋮	⋮
$X_{n_1 1}$	$X_{n_2 2}$	…	$X_{n_k k}$

2.2.1. Levene Testi

Bu test, j. örnekteki i. birimin değeri X_{ij} ve j. örneğin ortalaması da $\bar{X}_j$ olmak üzere,

$$Z_{ij} = \left|X_{ij} - \bar{X}_j\right|, \quad \bar{X}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} X_{ij} \tag{12}$$

olarak tanımlanan gözlem değerlerinden ortalamanın sapmalarının mutlak değerleri ile tek faktörlü varyans analizi yönteminin kullanılmasına dayanır [8].

Buna göre, j. örnekteki mutlak sapmaların ortalaması

$$\bar{Z}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} Z_{ij} \tag{13}$$

ve mutlak sapmaların genel ortalaması da,

$$\bar{Z} = \frac{1}{n_j}\sum_{j=1}^{k}\sum_{i=1}^{n_j} Z_{ij} \tag{14}$$

olmak üzere, gruplar arası kareler toplamı (GAKT) ve gruplar içi kareler toplamı (GIKT) sırasıyla eşitlik (14) ve eşitlik (15) ile hesaplanır;

$$GAKT = \sum_{j=1}^{k} n_j\left(\bar{Z}_j - \bar{Z}\right)^2 \tag{15}$$

ve

$$GIKT = \sum_{j=1}^{k}\sum_{i=1}^{n_j}\left(Z_{ij} - \bar{Z}_j\right)^2 \tag{16}$$

Gruplar arası ortalama kare (GAOK) ve gruplar içi ortalama kare (GIOK),

$$GAOK = \frac{GAKT}{k-1}\,, \qquad GIOK = \frac{GIKT}{n-k} \tag{17}$$

olmak üzere, varyansların homojenliği için Levene testi istatistiği aşağıdaki gibidir.

$$F_{k-1,n-k} = \frac{GAOK}{GIOK} \tag{18}$$

2.2.2. Brown-Forsythe Testi

Varyansların homojenliği için Brown-Forsythe testi Levene testine benzer. Levene testi ölçüm değerleri ile aritmetik ortalama arasındaki mutlak farkları kullanırken Brown-Forsythe testi ölçüm değerleri ile örnek medyanı (ortancası) arasındaki mutlak farklara dayanır [9].

$j.$ örnekteki $i.$ birimin ölçüm değeri X_{ij} ve $j.$ örneğin medyanı da m_j olsun.

$$m_j = Medyan\left(X_{1j}, X_{2j}, \dots, X_{n_j j}\right) \quad j = 1,2, \dots, k \tag{19}$$

Bilindiği gibi medyan için aşağıdaki tanımlar da yapılabilir.

1) n_j tek ise $X_{1j}, X_{2j}, \dots, X_{n_j j}$ değerleri küçükten büyüğe doğru sıralandığında ortadaki değer medyandır.
2) n_j çift ise $X_{1j}, X_{2j}, \dots, X_{n_j j}$ değerleri küçükten büyüğe doğru sıralandığında ortadaki iki değerin ortalaması medyandır.

Buna göre,

$$Z_{ij} = \left|X_{ij} - m_j\right| \tag{20}$$

olarak tanımlansın. Brown-Forsythe testi eşitlik (19) ‘daki gibi tanımlanan Z_{ij} değerleri tek faktörlü varyans analizi yönteminin kullanımına dayanır. Levene testinde verilen GAKT, GIKT, GAOK VE GIOK formülleri bu test için de geçerlidir. Brown-Forsythe test istatistiği aşağıdaki gibi tanımlanır.

$$BF = \frac{GAOK}{GIOK} \tag{21}$$

BF test istatistiği $k-1$ ve $n-k$ serbestlik dereceli F dağılımına sahiptir ve $BF \sim F_{k-1,n-k}$ biçiminde ifade edilir.

Bu sonuca göre,

$$F_h > F_{k-1,n-k,\alpha} \; ise \; H_0 \; reddedilir.$$
$$F_h \leq F_{k-1,n-k,\alpha} \; ise \; H_0 \; reddedilemez.$$

2.3. Bir Örnek İçin Testler

Bir örnek için testler, örneklemin belirli bir yığından çekilip çekilmediğinin ya da başka bir deyimle belirli bir yığına ait olup olmadığının incelenmesi amacı ile kullanılır. Yığına ilişkin aritmetik ortalama ile ilgili hipotezler için yığın varyansının bilinip bilinmemesine göre parametrik testlerden z testi ya da t testi kullanılabilir. Ancak bu iki testin kullanılabilmesi için örnek birimlerinin seçildiği yığının ilgili değişken bakımından dağılımının normal olması gerekir. Bu varsayım sağlanmadığında, özellikle örnek hacmi küçükse, parametrik testleri kullanmak doğru olmaz. Böyle durumlar için uygun analiz yöntemi parametrik olmayan testlerdir [2].

Bu çalışmaya dahil edilen bir örnek için parametrik testler z testi ve t testidir. Parametrik olmayan testler ise; İşaret testi, Binom testi, Run testi, Wilcoxon İşaretli Sıra Sayıları testi biçimindedir. Geliştirilen algoritma, ilk olarak örnek birimlerinin normal dağılıma uygunluğunu kontrol etmekte ve sonrasında ilgili testlerin özel varsayımlarını dikkate alarak öneri (ler)de bulunmaktadır.

2.3.1. *z* Testi

Örneklem ortalamasına dayalı bir testtir. Bilinen σ^2 varyanslı ve bilinmeyen μ ortalamalı normal dağılıma sahip bir kitleden alınan n rastgele değişken $X_1, X_2, \dots, X_n$ olsun. İddia edilen örneklem ortalaması, μ_0 değeri için hipotezler aşağıdaki gibidir [10].

$$H_0: \mu = \mu_0$$
$$H_1: \mu \neq \mu_0$$

Test istatistiği z ;

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \quad (22)$$

biçimindedir.

z testi için varsayımlar;

1) Örneklemin varyansı (σ^2) bilinmelidir,
2) Örneklem normal dağılıma uygunluk göstermelidir,
3) Ölçme düzeyi en az eşit aralıklıdır,
4) Örnek hacmi $n > 30$ olduğu durumlarda tercih edilir.

2.3.2. t Testi

Örneklem ortalamasına dayalı bir testtir. Bilinmeyen σ^2 varyanslı ve bilinmeyen μ ortalamalı normal dağılıma sahip bir kitleden alınan örneklem ile kullanılır [10].

İddia edilen örneklem ortalaması μ_0 olmak üzere hipotez testleri;

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \neq \mu_0$$

biçiminde tanımlanır. Test istatistiği t olmak üzere,

$$t = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \quad (23)$$

biçiminde elde edilir. Burada σ bilinmediğinden örneklemden tahmin edilir. σ ‘nın tahmin edicisi S ile ifade edilir ve,

$$S^2 = \frac{\sum_{i=0}^{n}(x-\bar{x})^2}{n-1} \tag{24}$$

biçimindedir.

t testi için varsayımlar;

1) Örneklem normal dağılıma uygunluk göstermelidir,
2) Ölçme düzeyi en az eşit aralıklıdır,
3) Örnek hacmi $n < 30$ olduğu durumlarda tercih edilmelidir.

2.3.3. Binom Testi

Bazı araştırmalarda bir karakteristik bakımından yığına ilişkin oran ilgi konusu olabilir. Örneğin bir hekim uyguladığı tedaviye olumlu tepki veren hastaların oranının 0.80 ‘den büyük olup olmadığını, bir işyeri müşterilerinde bayanların oranının 0.50 ’den farklı olup olmadığını bilmek isteyebilir. Bu tür problemlerde yığın ya soyut özellikte ya da çok sayıda birim içerdiğinden ilgi duyulan bu parametreyi (oran) hesaplamak yerine yığından seçilen n hacimli bir örnekteki bilgi kullanılarak bu parametreye ilişkin bir hipotez testi ile karar verilmeye çalışılır. Bu amaçla kullanılan parametrik olmayan test binom testi olarak bilinir [11].

Binom testinde yığına ilişkin bilinmeyen oran π ve 0 ile 1 arasındaki herhangi bir sayı da π_0 olmak üzere, hipotezler;

$$H_0: \pi = \pi_0$$
$$H_1: \pi \neq \pi_0$$

biçiminde oluşturulur. Test istatistiği b olmak üzere,

$$b = \sum_{i=1}^{n} \delta_i \sim Binom \tag{25}$$

biçiminde elde edilir. Burada δ_i değişkeni eşitlik (25) ‘deki gibi tanımlanır.

$$\delta_i = \begin{cases} 1, & i.\text{ölçüm değeri ilgilenilen özellikte ise} \\ 0, & i.\text{ölçüm değeri ilgilenilen özellikte değilse} \end{cases} \tag{26}$$

Binom testi için varsayımlar;

1) Veri n tekrarlı Bernoulli deneyinin sonuçlarından oluşur. Yani her gözlem (ölçüm) sonucu "başarılı" ya da "başarısız", "sağlam" ya da "bozuk" gibi birbirinden ayrık iki sonuçtan biri ile veri setinde yer alır,
2) Tekrarlar (deneyler) birbirinden bağımsızdır,
3) İlgilenilen sonucun olasılığı p olmak üzere, bu olasılık tekrardan tekrara (deneyden deneye) bağımsızdır.

2.3.4. İşaret Testi

1710 yılında John Arbuthnot tarafından önerilen işaret testi parametrik olmayan testler içerisinde en eskisidir. Ayrıca bu test binom testinin özel bir hali olarak da düşünülebilir [12].

İşaret testinde, M yığına ilişkin bilinmeyen medyan ve M_0 herhangi bir reel sayı olmak üzere, hipotezler;

$$H_0 : M = M_0$$
$$H_1 : M \neq M_0$$

biçiminde oluşturulur. Test istatistiği k olmak üzere,

$$k = \sum_{i=1}^{n} \delta_i \tag{27}$$

biçiminde elde edilir. Burada δ_i değişkeni aşağıdaki gibi tanımlanır. İlgili yığından rastgele seçilen n hacimli örnek $X_i : X_1, X_2, \dots, X_n$ olsun.

$$D_i = X_i - M_0, \qquad i = 1,2,\dots,n$$
$$\delta_i = \begin{cases} 1, & D_i > 0 \\ 0, & D_i < 0 \end{cases} \tag{28}$$

İşaret testi için varsayımlar;

1) n hacimli örnek medyanı bilinmeyen bir yığından rastgele seçilmiştir,
2) İlgilenilen değişken en az sıralama ölçme düzeyinde ölçülmüştür,
3) İlgilenilen değişken süreklidir.

2.3.5. Wilcoxon İşaretli Sıra Sayıları Testi

İşaret testi gözlem değerleri ile hipotezde belirtilen medyan değeri arasındaki farkların sadece işaretlerini kullanır. Wilcoxon işaretli sıra sayıları testi ise yoluk hipotezinin test edilmesinde farkların büyüklüklerini de hesaba katan parametrik olmayan bir testtir [13]. Wilcoxon işaretli sıra sayıları testi D_i farklarının büyüklüklerini de hesaba kattığı için işaret testine göre daha çok bilgi kullanır ve bu nedenle işaret testinden daha güçlüdür.

Wilcoxon işaretli sıra sayıları testinde, M yığına ilişkin bilinmeyen medyan ve M_0 herhangi bir reel sayı olmak üzere, hipotezler;

$$H_0 : M = M_0$$
$$H_1 : M \neq M_0$$

biçiminde oluşturulur. Test istatistiği T^+ olmak üzere,

$$T^+ = \sum_{i=1}^{n} r(|D_i|)\delta_i \tag{29}$$

biçiminde elde edilir. Ve D_i ile δ_i istatistikleri,

$$D_i = X_i - M_0, \qquad i = 1,2,\dots,n \tag{30}$$

$$\delta_i = \begin{cases} 1, & D_i > 0 \\ 0, & D_i < 0 \end{cases}$$

biçiminde elde edilir.

$|D_i|$ için sayı-sıralı istatistik de $r(|D_i|)$ olsun.

$$r(|D_i|) = \sum_{j=1}^{n} S\left[|D_i| - |D_j|\right] \tag{31}$$

$$S(u) = \begin{cases} 1, & u \geq 0 \\ 0, & u < 0 \end{cases} \tag{32}$$

biçiminde tanımlanır.

Wilcoxon işaretli sıra sayıları testi için varsayımlar;

1) n hacimli örnek, medyanı bilinmeyen yığından rastgele seçilmiştir,
2) İlgilenilen değişken süreklidir,
3) Yığının dağılımı simetriktir,
4) Ölçme düzeyi en az eşit aralıklıdır,
5) Gözlemler birbirinden bağımsızdır.

2.3.6. Rastgelelik İçin Dizi Parçaları (Run) Testi

Rastgelelik istatistikte önemli bir kavramdır ve veri analizinde sıkça karşılaşılır. Örneğin istatistiksel testlerin çoğu ilgili yığından n hacimli rastgele örnek seçilmesine dayanır. Benzer biçimde istatistiksel kalite kontrolü süreci üretim sürecinden n hacimli ürünün rastgele seçildiği varsayımına dayanır [10]. Regresyonda da hata terimlerinin 0 ortalamalı ve σ^2 varyanslı bir rastgele değişken olduğu varsayılır.

Bu çalışmada dizi parçaları testi n hacimli örnekten derlenen ölçüm (gözlem) değerlerinin iki tür simge ile gösterildiği durum için açıklanacaktır. Bu test için ölçüm değerlerinin önce iki tür simgeden oluşan bir dizi olarak gösterilmesi gerekir. Bir kuyrukta yer alan 8 bireyin cinsiyetlerine göre sıralanışı bu duruma uygun bir örnektir ve Tablo 4‘te yer almaktadır.

Tablo 4. Bir Kuyrukta Yer Alan Bireylerin Cinsiyetlerine Göre Sıralanışı

Birey no	1	2	3	4	5	6	7	8
Cinsiyet	K	K	E	E	K	E	E	K

Bu bireylerin boy uzunlukları Tablo 5‘te verilmiştir.

Tablo 5. Bir Kuyrukta Yer Alan Bireylerin Boy Uzunlukları

Birey no	1	2	3	4	5	6	7	8
X_i	1.70	1.75	1.80	1.82	1.69	1.90	1.74	1.65

Aritmetik ortalamadan sapmaların rastgelelik gösterip göstermediği araştırılmak istenirse önce bu verideki ölçüm değerlerinin aritmetik ortalamadan sapmalarına dayalı simgeler oluşturulur. Bu veri için $\bar{X} = 1.756$ ‘dır.

Aritmetik ortalamanın altında olanlar A,
Aritmetik ortalamanın üstünde olanlar B ile ifade edildiğinde bireylere ilişkin simgeler Tablo 6‘da verildiği gibidir.

Tablo 6. Run Testi için Sembolize Edilmiş Değerler

Birey no	1	2	3	4	5	6	7	8
Simge	A	A	Ü	Ü	A	Ü	A	A

olarak ifade edilebilir.

Rastgelelik için Run testi n birimlik örnekten derlenen verinin ortalamadan, medyandan ya da herhangi bir değerden sapmalarının rastgele olup olmadığının kontrol edilmesinde kullanılır.

2.4. İki Bağımsız Örnek İçin Testler

İki bağımsız örnek için testler, iki örnek ortalamasının eşitliği hipotezini test etmek amacıyla kullanılır. Bu amaç doğrultusunda parametrik testlerden t testi kullanılabilir. Ancak bu testin kullanımı örneklerin seçildikleri yığınların dağılım biçimlerinin normal olduğu varsayımına dayanır. Eğer iki bağımsız örnekten biri bile normal dağılımdan gelmediyse ve küçük hacimli örnekler kullanılıyorsa, bu parametrik testin kullanımı doğru olmaz. Böyle durumlarda parametrik olmayan testler tercih edilmelidir [2]. Bu çalışmada iki bağımsız örnek için konum parametrelerinin eşitliği hipotezinde kullanılan parametrik testlerden t testi ve parametrik olmayan testlerden Fisher ‘ın Tam Olasılık testi, Mann-Whitney testi, Wald-Wolfowitz Dizi Parçaları testi, Kolmogorov – Smirnov testi ve dağılış parametrelerinin eşitliği hipotezinin testinde kullanılan Mood testi ile Siegel-Tukey testine yer verilmiştir.

2.4.1. t Testi

İki bağımsız örnekten derlenen n_1 ve n_2 hacimli örnek için kitle ortalaması μ_1 ve μ_2 arasındaki farkın anlamlılığını test etmek için kullanılır [10]. İki bağımsız örnek t testi parametrik bir testtir ve bu test için hipotezler;

$$H_0 : \mu_1 - \mu_2 = 0$$
$$H_1 : \mu_1 - \mu_2 \neq 0$$

biçiminde oluşturulur. Test istatistiği t olmak üzere,

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\bar{\mu}_1 - \bar{\mu}_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}} \tag{33}$$

biçimindedir. Burada $\sigma_{\bar{x}_1 - \bar{x}_2}$,

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \tag{34}$$

ile elde edilir.

Ancak kitle varyansları bilinmiyorsa $\sigma_{\bar{x}_1 - \bar{x}_2}$ ‘nın tahmin edicisi olan,

$$S_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} \tag{35}$$

kullanılır.

t testi için varsayımlar;

1) Ölçme düzeyi eşit aralıklı veya oranlama olmalıdır,
2) Örnekler normal dağılıma uygun olmalıdır.

2.4.2. Fisher ’ın Tam Olasılık Testi

İki bağımsız örnekten derlenen verideki her bir gözlem değerinin iki (ayrık) sınıftan birine sınıflandırılması problemi ile uygulamada sıkça karşılaşılır. Örneğin, A ve B ilaçlarının uygulandığı hastaların tedaviye olumlu tepki verenler ve vermeyenler olarak sınıflandırılmaları gibi [14]. Bu tür sınıflandırma ile 2×2 boyutlu bir frekans tablosu elde edilir. Fisher tam olasılık testinde hipotezler, yığınlar için ilgilenilen özellik bakımından oranlar sırasıyla π_1 ve π_2 olmak üzere Tablo 7’deki gibi ifade edilir.

$H_0: \pi_1 = \pi_2$

$H_1: \pi_1 \neq \pi_2$

Tablo 7. Fisher Tam Olasılık Testi Frekans Tablosu

Örnek No	**İlgilenilen özellikte olanlar**	**İlgilenilen özellikte olmayanlar**	**Toplam**
1	a	$n_1 - a$	n_1
2	b	$n_2 - b$	n_2
Toplam	$a + b$	$n - (a + b)$	n

Bu testte “ikinci örnekte ilgilenilen özellikte olanlar sınıfına düşen birim sayısı” test istatistiği olarak kullanılır. Yukarıda verilen tabloya göre test istatistiği b ile gösterilir ve bu test istatistiği hipergeometrik dağılıma sahiptir.

$$f(b,a) = \frac{\binom{n_2}{b}\binom{n_1}{a}}{\binom{n_1+n_2}{a+b}} \tag{36}$$

Bu testteki temel amaç iki ayrık sınıfa düşen birimlerin oranı bakımından iki grubun farklı olup olmadığının belirlenmesidir. Fisher testi bu amaçla kullanılan testlerden biridir.

Fisher ‘ın tam olasılık testi için varsayımlar;

1) Veri $X_1, X_2, \dots, X_{n_1}$ ve $Y_1, Y_2, \dots, Y_{n_2}$ gözlemlerinden oluşur,
2) Her bir gözlem iki ayrık sonuçtan birine sınıflandırılabilir,
3) İlgili değişken sınıflama ölçme düzeyinde ölçülmüş olmalıdır.

2.4.3. Mann-Whitney Testi

İki bağımsız örnek için örneklerin geldikleri yığınların konum parametrelerinin (medyan) eşitliği hipotezini test etmek amacıyla kullanılır [15]. Bu amaçla kullanılan Mann-Whitney testinin temel özelliği gözlem değerlerine atanan büyüklük sıra sayılarına dayanmasıdır.

Örneklerin seçildikleri yığınların bilinmeyen medyanları sırasıyla M_1 ve M_2 olmak üzere, Mann-Whitney testinde hipotezler;

$$H_0: M_1 = M_2$$
$$H_1: M_1 \neq M_2$$

biçiminde tanımlanır.

Mann-Whitney test istatistiği elde edilirken önce n_1 ve n_2 hacimli iki bağımsız örnek birleştirilir ve $n_1 + n_2 = n$ hacimli birleştirilmiş örnek oluşturulur. Sonra bu birleştirilmiş örnekteki birimlere küçükten büyüğe doğru sıra sayıları atanır. Bu sıra sayıları atama işleminde aynı değerli olanlar varsa bunlara ortalama sıra sayısı atanır. X gözlemlerine atanan sıra sayılarının toplamı S olmak üzere, Mann-Whitney test istatistiği,

$$T = S - \frac{n_1(n_1+1)}{1} \tag{37}$$

biçiminde tanımlanır. Test istatistiğinin alabileceği değerler $T = 0,1, \dots, n_1 n_2$ aralığındadır.

Mann Whitney testi için varsayımlar;

1) Veri $X_1, X_2, \dots, X_{n_1}$ ve $Y_1, Y_2, \dots, Y_{n_2}$ bağımsız gözlemlerinden oluşur,
2) Örnekler bağımsızdır,
3) Gözlenen değişken sürekli rasgele değişkendir,
4) Ölçme düzeyi en az sıralamadır.

2.4.4. Wald-Wolfowitz Dizi Parçaları Testi

Bu test iki bağımsız örnekten derlenen verinin Rastgelelik için Dizi Parçaları (Run) testinde kullanılabilecek biçimde düzenlenmesine dayanır [2]. Run testinde olduğu gibi Wald-Wolfowitz testinde de ölçüm değerleri simgelerle temsil edilir.
Wald-Wolfowitz dizi parçaları testinde hipotezler;

H_0 : X ve Y gözlemleri benzer dağılımlı yığınlardan gelmişlerdir.
H_1 : X ve Y gözlemleri benzer dağılımlı yığınlardan gelmemişlerdir.

biçiminde tanımlanır.

Wald-Wolfowitz dizi parçaları testi için varsayımlar;

1) Veri n_1 ve n_2 hacimli $X_1, X_2, \dots, X_{n_1}$ ve $Y_1, Y_2, \dots, Y_{n_2}$ bağımsız gözlemlerinden oluşur,
2) İlgilenilen değişken süreklidir.

2.4.5. Kolmogorov-Smirnov Testi

Kolmogorov-Smirnov testinde amaç n_1 ve n_2 hacimli iki örneğin aynı yığından mı yoksa farklı yığınlardan mı geldiğine karar vermektir [4]. Test için hipotezler,

$$H_0 : F_1(x) = F_2(X), \qquad x \text{ değerlerinin hepsi için}$$
$$H_1 : F_1(x) \neq F_2(X), \qquad x \text{ değerlerinin en az biri için}$$

biçiminde tanımlanır. Test istatistiği D ‘yi elde etmek için,

$$S_1(x) : 1. \text{ örneğe ilişkin gözlenen dağılım fonksiyonu}$$
$$S_2(x) : 2. \text{ örneğe ilişkin gözlenen dağılım fonksiyonu}$$

olmak üzere, iki yanlı test için test istatistiği Eşitlik (36) ‘daki gibi tanımlanır.

$$D = En\ büyük|S_1(x) - S_2(x)| \tag{38}$$

Kolmogorov-Smirnov testi için varsayımlar;

1) Veri, oranlama ya da eşit aralıklı ölçme düzeyinde ölçülmüş olmalıdır.

2.4.6. Mood Testi

Bazen dağılış (saçılım, yayılım) ölçüsü olan iki parametrenin eşit olup olmadığının bilinmesine ihtiyaç duyulur. İki ortalama farkına ilişkin test yaparken yığın varyansları bilinmiyorsa, bu bilgiye özellikle küçük örnek hacimlerinde ihtiyaç vardır. Çünkü t dağılımının serbestlik derecesinin belirlenebilmesi bu bilgiyi gerektirir. İki yığının dağılış parametrelerinin eşitliğinin testi için kullanılan parametrik olmayan testlerden biri Mood testidir [2]. Mood testinin kullanılabilmesi için yığınların medyanlarının aynı olması gerekir. Mood testi için hipotezler;

$$H_0 : \sigma_1 = \sigma_2$$
$$H_1 : \sigma_1 \neq \sigma_2$$

biçiminde kurulur. Test istatistiği M olmak üzere,

$$M = \sum_{i=1}^{n_1} \left(r_i - \frac{n+1}{2} \right)^2 \tag{39}$$

Burada n, birleştirilmiş örnek hacmi $n = (n_1 + n_2)$ ve r_i, birleştirilmiş örnekteki ölçüm değerlerine atanan sıra sayılarındaki X gözlemlerine karşı gelenler şeklinde tanımlanır.

Mood testi için varsayımlar;

1) $n_1 \leq n_2$ olmak üzere, veri, $X_1, X_2, \dots, X_{n_1}$ ve $Y_1, Y_2, \dots, Y_{n_2}$ gözlemlerinden oluşur,
2) İlgili değişken süreklidir,
3) Ölçme düzeyi en az sıralamadır,
4) Yığınların medyanları aynıdır.

2.4.7. Siegel-Tukey Testi

İki dağılış parametresinin eşitliği hipotezi için kullanılabilecek testlerden bir diğeri de Siegel-Tukey testidir. Varsayımları Mood testi ile aynıdır [2]. Siegel-Tukey testi için hipotezler;

$$H_0 : \sigma_1 = \sigma_2$$
$$H_1 : \sigma_1 \neq \sigma_2$$

biçiminde tanımlanır. Test istatistiği ST olmak üzere,

$$ST = \sum_{i=1}^{n} a_i \delta_i \tag{40}$$

olarak tanımlanır. Burada $n = (n_1 + n_2)$ olmak üzere,

$$\delta_i = \begin{cases} 1, & i.\text{sıradaki değer } X \text{ gözlemi ise} \\ 0, & i.\text{sıradaki değer } Y \text{ gözlemi ise} \end{cases} \tag{41}$$

$$a_i = \begin{cases} 2i, & i.\,çift\ ve\ 1 < i \leq \left(\frac{n}{2}\right)\ ise \\ 2i-1, & i.\,tek\ ve\ 1 \leq i \leq \left(\frac{n}{2}\right) ise \\ 2(n-i)+2, & i.\,çift\ ve\ \left(\frac{n}{2}\right) < i \leq n\ ise \\ 2(n-i)+1, & i.\,tek\ ve\ \left(\frac{n}{2}\right) < i < n\ ise \end{cases} \tag{42}$$

şeklinde elde edilir.

Siegel-Tukey testi için varsayımlar;

1) $n_1 \leq n_2$ olmak üzere, veri, $X_1, X_2, \dots, X_{n_1}$ ve $Y_1, Y_2, \dots, Y_{n_2}$ gözlemlerinden oluşur,
2) İlgili değişken süreklidir,
3) Ölçme düzeyi en az sıralamadır,
4) Yığınların medyanları aynıdır.

2.5. İki Bağımlı Örnek için Testler

İki bağımlı örneğe ilişkin veri genel olarak iki şekilde elde edilir. Bunlardan birincisi "işlem öncesi-işlem sonrası" tasarımıdır. Bu tasarımda önce örneğe seçilen birimlerin ilgili bağımlı değişken bakımından aldıkları değerler ölçülür. Daha sonra bu birimlere ilgili işlem uygulanır ve işlem uygulaması bittikten sonra aynı bağımlı değişken bakımından bu birimlerin işlemden sonraki aldıkları değerler tekrar ölçülür. Örneğin rastgele seçilen 10 kadının ağırlıklarının ölçülmüş olduğu varsayılsın. Bu kadınlara bir zayıflama rejimi uygulanmış olsun. Aynı kadınların zayıflama rejiminden 20 gün sonraki ağırlıkları da saptanabilir. Bu 10 kadının zayıflama rejiminden önceki ve sonraki ağırlıkları işlem öncesi-işlem sonrası tasarımı için bir örnektir.

İki bağımlı örnek için ikinci tasarım "eşleştirme" dir. Bu tasarımda örneğe çekilen birimler önce belirli bir dışsal değişken bakımından eşlere ayrılır. Bu değişken bağımlı değişken üzerinden etkisi arındırılmak istenen değişkendir. Eşler, ya da çiftler, oluşturulurken etkisi arındırılmak istenen bu dışsal değişken bakımından aynı değerli ya da birbirine en yakın değerli olan iki birim bir eş olarak alınır. Sonra eşlerdeki birimler rastgele, birer tane olmak üzere, iki gruptan birine atanır. Böylece iki grup oluşturulur. Gruplardan biri işlem grubu diğeri kontrol grubudur. Ya da gruplardan birine belirli bir konuda bilinen bir yöntem diğerine de geliştirilen yöntem uygulanabilir. Örneğin rastgele seçilen 10 hasta "hastalık derecesi" ya da "yaşlarına" göre eşlere ayrılabilir. Hastalar yaşlarına göre eşlere ayrılırken aynı yaşta olanlar ya da yaşları en yakın olanlar birer eş olacaktır. Bu ölçüte göre oluşturulan 5 eş ortaya çıkar. Her eşteki iki hasta rastgele, birer tane olmak üzere, iki gruptan birine atanır ve sonuçta her birinde 5'er hasta bulunan iki grup oluşturulur. Yığınların dağılımlarının normal olması durumunda bu tip veriler için bilinen parametrik yöntem Eşleştirilmiş t testidir [2].

Bu çalışmada iki bağımlı örnek için parametrik testlerden Eşleştirilmiş t testine ve parametrik olmayan testlerden, McNemar testi, İşaret testi, Wilcoxon testine yer verilmiştir.

2.5.1. Eşleştirilmiş t Testi

Eşleştirilmiş t testi aynı örnek birimlerinde yapılan iki ölçüm arasında fark olup olmadığı, diğer bir ifade ile iki ölçüm arasında fark varsa bu farkın istatistiksel olarak anlamlı olup olmadığını belirlemek için uygulanır [10]. Test hipotezleri;

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2$$

biçiminde tanımlanır. Test istatistiği t ve ölçümler arasındaki fark D_i olmak üzere,

$$D_i = X_i - Y_i$$

$$t = \frac{\sum D_i}{\sqrt{\frac{n\sum D_i^2 - (\sum D_i)^2}{n-1}}} \tag{43}$$

biçiminde elde edilir.

Eşleştirilmiş t testi için varsayımlar;

1) Ölçme düzeyi eşit aralıklı veya oranlama olmalıdır,

2) Örnekler normal dağılıma uygun olmalıdır.

2.5.2. McNemar Testi

Bu test iki bağımlı örnek problemlerinde ölçüm (işlem) sonucuna göre eşlerin kategorilere (sınıflara) ayrıldığı durumlarda kullanılır [16]. Eşlerin,

a) Tahlil sonucunun pozitif ya da negatif olmasına göre sınıflandırılması,
b) Tedavi sonucunda hastaların yaşamlarını sürdürüp sürdürememe durumlarına göre sınıflandırılması
c) Üretimde hedeflenen kaliteyi sağlayıp sağlayamama durumuna göre sınıflandırılması

durumları örnek olarak verilebilir. Benzer biçimde $i.$ birimin işlem öncesi ve işlem sonrası ölçüm sonucuna göre kategorilere ayrıldığı durumlarda da McNemar testi kullanılır. $i.$ hastanın tedavi öncesi ve tedavi sonrası tahlil sonuçlarına göre normal değerleri içinde olup olmama durumuna göre sınıflandırılması, $i.$ sporcunun klasik ve yeni antrenman programlarında gösterdikleri performansa göre olimpiyatlara katılma barajını geçip geçemediğine göre sınıflandırılması, vb. bu duruma örnek olarak verilebilir.

Burada amaç uygulanan işlemin iyileşme sağlayıp sağlamadığının belirlenmesidir. İki bağımlı örnek problemlerinde eşler ve birimler problemin yapısına uygun olarak sınıflandırıldığında $2x2$ boyutlu bir frekans tablosu elde edilir. Eşleştirme tasarımı için genel biçimi Tablo 8'de verilmiştir.

McNemar testi, iki bağımlı gruba ait farklı uygulama söz konusu olduğunda, iki ölçüm arasındaki farklılığı incelemek için kullanılmaktadır. Bu nedenle, iki bağımlı gruba ait iki farklı uygulama söz konusu olduğundan sadece $2x2$ boyutlu tablolarda kullanılır.

Tablo 8. McNemar Testi Frekans Tablosu

		İşlem Grubu		**Toplam**
		Negatif	**Pozitif**	
Kontrol Grubu	**Negatif**	A	C	$A+C$
	Pozitif	B	D	$B+D$
	Toplam	$A+B$	$C+D$	n

Bir işlem (koşul) altında ilgilenilen özellikte olanların oranı π_1 ve diğer işlem altında ilgilenilen özellikte olanların oranı da π_2 olmak üzere, McNemar testinde hipotezler,

$$H_0 : \pi_1 = \pi_2$$
$$H_1 : \pi_1 \neq \pi_2$$

biçiminde ifade edilir. İlgilenilen özellik bakımından örnek oranlarını P_1 ve P_2 ile gösterelim. Buradan,

$$P_1 = \frac{A + C}{n} \text{ ve } P_2 = \frac{A + B}{n} \tag{44}$$

olarak belirlenir. Bu örnek oranları arasındaki fark,

$$P_1 - P_2 = \frac{C - B}{n} \tag{45}$$

olur. Yokluk hipotezi doğru iken

$$E = \left(\frac{C - B}{n}\right) = 0 \tag{46}$$

dır. H_0 hipotezinin testi için McNemar test istatistiği

$$Z = \frac{C - B}{\sqrt{C + B}} \tag{47}$$

biçimindedir.

2.5.3. İşaret Testi

İki bağımlı örnek için işaret testi, bir örnek işaret testine dayanır [16]. Test hipotezleri;

H_0 : D_i farklı yığınının medyanı sıfırdır

H_1 : D_i farklı yığınının medyanı sıfırdan farklıdır

biçiminde tanımlanır.

İki bağımlı örnek için işaret testi X ve Y değişkenlerinden yeni bir değişken tanımlamayı gerektirir. Bu yeni değişken D olsun. D değişkeni Eşitlik (49) 'daki gibidir.

$$D_i = X_i - Y_i, \qquad i = 1,2,\dots,n \tag{48}$$

İşaret testinde olduğu gibi işaret testi istatistiği

$$\delta_i = \begin{cases} 1, & D_i > 0 \\ 0, & D_i < 0 \end{cases}$$
$$k = \sum_{i=1}^{n} \delta_i \tag{49}$$

olarak tanımlanır.

İşaret testi için varsayımlar;

1) Veri $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ gibi n sayıda gözlem çiftinden oluşur ve bu gözlemler ya aynı birimden alınmıştır ya da etkisi arındırılan dışsal değişken bakımından aynı değerli olan iki birimden alınmıştır,
2) İlgilenilen değişkenin ölçme düzeyi en az sıralamadır,
3) İlgilenilen değişken süreklidir.

2.5.4. Wilcoxon İşaretli Sıra Sayıları Testi

İki bağımlı örnek problemlerinde bağımlı değişken için ölçme düzeyi eşit aralıklı ya da oranlama ise işaret testinin kullanımı bilgi kaybına neden olur. Böyle durumlarda daha güçlü olan Wilcoxon işaretli sıra sayıları testi seçilmelidir [17].

Test hipotezleri aşağıdaki gibi tanımlanır.

H_0 : D_i farklı yığınının medyanı sıfırdır

$H_1 : D_i$ farklı yığınının medyanı sıfırdan farklıdır

Test istatistiği T^+ olmak üzere,

$$T^+ = \sum_{i=1}^{n} \delta_i r(|D_i|) \tag{50}$$

biçiminde elde edilir. Burada,

$$D_i = X_i - Y_i, \qquad i = 1,2,\dots,n$$
$$\delta_i = \begin{cases} 1, & D_i > 0 \\ 0, & D_i < 0 \end{cases} \tag{51}$$

biçiminde elde edilir.

Wilcoxon işaretli sıra sayıları testi için varsayımlar;

1) Veri $D_i = X_i - Y_i,\ i = 1,2,\dots,n$ olmak üzere n sayıda farktan oluşur,
2) Farklar sürekli bir rastgele değişkendir,
3) Farklar yığının dağılımı simetriktir,
4) Ölçme düzeyi en az eşit aralıklıdır.

2.6. İkiden Fazla Bağımsız Örnek İçin Testler

İkiden fazla bağımsız örneğin seçildiği yığınların konum parametrelerinin eşitliği hipotezinin testi için bilinen parametrik yöntem “F Testi (Varyans Analizi)” dir. Tek faktörlü (yönlü) varyans analizi olarak da bilinen bu yöntemde yokluk ve karşıt hipotezi, yığın sayısı k, j. yığının ortalaması μ_j ve genel ortalama da μ olmak üzere,

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k = \mu$$
$$H_1 : \mu_j \, '\mathit{lerin\ en\ az\ biri\ farklıdır}.$$

biçiminde ifade edilir.

Bu H_0 hipotezinin testi için, j. örnek hacmi n_j ve $\sum_{j=1}^{k} n_j = n$ olmak üzere, $k-1$ ve $n-k$ serbestlik dereceli F istatistiği kullanılabilir. Ancak bu test istatistiğinin kullanılabilmesi için,

1) "n_j hacimli örneklerin her biri normal dağılımdan gelmiştir"
2) "n_j hacimli örneklerden her birinin geldikleri yığınların varyansları aynıdır"

varsayımlarının sağlanması gerekir. Aynı zamanda varyans homojenliğinin de sağlanması gerekmektedir.

Bu çalışmada ikiden fazla bağımsız örnek için parametrik testlerden Tek Yönlü Varyans Analizi ve parametrik olmayan testlerden Medyan testi ile Kruskal-Wallis-H testi ve Bağımsızlık için Ki Kare testine yer verilmiştir.

2.6.1. Tek Yönlü Varyans Analizi

Varyans analizi, normal dağılım gösteren bağımlı ya da bağımsız yığınların ortalamalarına ilişkin hipotezlerin test edilmesinde yararlanılan bir yöntemdir [18]. Varyans analizi $k > 2$ olmak üzere k örnekten elde edilen veri setinde incelenen değişkene ait olan genel varyansı (genel değişimi), bu değişime katkıda bulunan öğelerine ayırarak analiz etmeyi sağlayan bir yöntemdir. Varyans analizine bağımsız örneklerde t testinin ikiden fazla grup için genellenmiş bir şekli denilebilir. Tek yönlü varyans analizi için özetle, bir bağımsız değişkenin bir bağımlı değişken üzerindeki etkisini inceler.
Tek yönlü varyans analizinde hipotezler aşağıdaki gibidir.

$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k = \mu$
$H_1 : \mu_j$ 'lerin en az biri farklıdır.

Test istatistiği F olmak üzere,

$$F = \frac{GAKT}{GİKT} \tag{52}$$

biçiminde verilir. Tek yönlü varyans analizinde test istatistiğinin hesaplanmasına yönelik varyans analizi Tablo 9‘da verildiği gibidir.

Tablo 9. Tek Yönlü Varyans Analiz Tablosu

Değişimin Kaynağı	**Kareler Toplamı**	**Serbestlik Derecesi**	**Kareler Ortalaması**	**Oran**
Gruplar Arası	$GAKT = \sum_{i=1}^{k} \left(\frac{T_{i.}^2}{n_i}\right) - \frac{T_{..}^2}{n}$	$k-1$	$GAOK = \frac{GAKT}{k-1}$	$F = \frac{GAOK}{GIOK}$
Gruplar İçi	$GIKT = \sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij}^2 - \sum_{i=1}^{k}\frac{T_{i.}^2}{n_i}$	$n-k$	$GIOK = \frac{GIKT}{n-k}$	
Toplam	$\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(x_{ij} - \bar{x}_{..}\right)^2$	$n-1$		

Burada, k, örneklem sayısı ve n, toplam örnek sayısını ifade etmektedir.

Tek yönlü varyans analizi testi için varsayımlar;

1) İncelenen değişken her bir örneklemde normal dağılım göstermelidir,
2) Grup ortalamaları ve standart sapmaları arasında bir doğrusallık olmalıdır,
3) Bağımlı değişkenin ölçüm düzeyi k sayıdaki her bir grup için oransal veya en az eşit aralıklı olmalıdır,
4) Bağımlı değişken bütün gruplarda normal dağılıma ve homojen (yaklaşık olarak eşit) varyanslara sahip olmalıdır.

2.6.2. Tek Yönlü Çok Değişkenli Varyans Analizi

Çok değişkenli varyans analizi iki ve daha fazla bağımsız ve bağımlı gruplarda çok değişkenli normal dağılıma dayalı hipotezleri test etmek için kullanılır. İki ve daha fazla ($k \geq 2$) bağımsız grupta çok değişkenli normal dağılım gösteren örneklere ilişkin kurulan hipotezlerin test edilmesinde Tek Yönlü Çok Değişkenli Varyans Analizi (Tek Yönlü Manova) kullanılır [19].

Tek yönlü Manova 'da H_0 hipotezi bağımsız k örneklem ortalamalarının birbirlerine eşit olduğunu varsayarken, H_1 hipotezi ise k örneklem ortalamalarından en az birinin diğerinden farklı olduğunu varsayar.

$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$

$H_1 :$ En az biri farklı

Tek yönlü çok değişkenli varyans analizinde test istatistiğinin hesaplanmasına yönelik varyans analizi tablosu Tablo 10'da verilmiştir.

Tablo 10. Tek Yönlü Çok Değişkenli Varyans Analizi

Değişim Kaynağı	**Kareler ve Çapraz Çarpımlar Matrisi**	**Serbestlik Derecesi**
Gruplar arası	$B = \sum_{i=1}^{k} n_i(\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})'$	$sd_1 = k - 1$
Gruplar İçi	$W = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)(x_{ij} - \bar{x}_i)'$	$sd_2 = \sum_{i=1}^{k} n_i - k$
Genel	$T = B + W$	$gsd = \sum_{i=1}^{k} n_i - 1$

Gruplar arası değişimin hataya göre önemliliğini test etmek için Wilk's Lambda istatistiğinden yararlanılır. Wilk's Lambda istatistiği $\Lambda(= L)$ aşağıdaki gibi hesaplanır.

$$\Lambda = L = \frac{|W|}{|B + W|} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)(x_{ij} - \bar{x}_i)'}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - \bar{x})(x_{ij} - \bar{x})'} \tag{53}$$

Tek yönlü çok değişkenli varyans analizi testi için varsayımlar;

1) İncelenen örneklem çok değişkenli normal dağılım göstermelidir,
2) Grup ortalamaları ve standart sapmaları arasında bir doğrusallık olmalıdır,
3) Bağımlı değişkenin ölçüm düzeyi k sayıdaki her bir grup için oransal veya en az eşit aralıklı olmalıdır,

4) Bağımlı değişken bütün gruplarda normal dağılıma ve homojen (yaklaşık olarak eşit) varyanslara sahip olmalıdır.

2.6.3. Medyan Testi

İkiden fazla bağımsız örneğin medyanları aynı olan yığınlardan gelip gelmediğinin araştırılmasında Medyan testi kullanılır. Parametrik olmayan bir testtir ve bu nedenle normallik varsayımları gerektirmez [20].

Medyan testi için hipotezler;

H_0 : Örneklerin seçildikleri yığınlar aynı medyana sahiptir.
H_1 : Örneklerin seçildikleri yığınların en az birinin medyanı farklıdır.

biçiminde tanımlanır. m, birleştirilmiş örnek medyanı ve $B_{ij} = n_j/2$ olmak üzere ve,

G_{1i} : j. örnekte m değerinden büyük değerli olanların sayısı
G_{2j} : j. örnekte m değerine eşit veya küçük değerli olanların sayısı

olmak üzere, test istatistiği,

$$\sum_{j=1}^{k}\sum_{i=1}^{2}\frac{\left(G_{ij}-B_{ij}\right)^2}{B_{ij}} \sim \chi^2_{(k-1)(2-1)} \tag{54}$$

şeklinde elde edilir.

Medyan testi için varsayımlar;

1) Veri, k sayıda yığının her birinden rastgele seçilen n_j hacimli örneklerden oluşturulur,
2) Gözlemler hem örnek içinde hem de örnekler arasında bağımsızdır,
3) Ölçme düzeyi en az sıralamadır.

2.6.4. Kruskal-Wallis H Testi

Varyans çözümlemesinde amaç, kısaca, her biri n_j hacimli k sayıda bağımsız örneğin aynı yığından gelip gelmediğine karar vermektir [21]. Kruskal-Wallis H testi tek yönlü varyans analizinin parametrik olmayan test karşılığıdır. Kruskal-Wallis H testinde hipotezler aşağıdaki gibidir.

H_0 : Örneklerin seçildiği k sayıda yığının dağılım fonksiyonları aynıdır.
H_1 : Örneklerin seçildiği k sayıda yığının hepsi aynı medyana sahip değildir.

Test istatistiği H olmak üzere,

$$H = \frac{12}{n(n+1)} \sum_{j=1}^{k} n_j \left(\bar{R}_{.j} - \frac{n+1}{2} \right)^2 \tag{55}$$

olarak tanımlanmıştır.

Kruskal Wallis H Testi için varsayımlar;

a) Veri $n_1, n_2, \ldots, n_k$ hacimli k sayıda rastgele seçilen örnekten derlenir,
b) Gözlemler hem örnek içinde hem de örnekler arasında bağımsızdır,
c) Ölçme düzeyi en az sıralamadır,
d) İlgilenilen değişken süreklidir.

2.6.5. Bağımsızlık İçin Ki Kare Testi

Bağımsızlık için ki-kare testinde biri sınıflama diğeri sıralama veya ikisi de sınıflama düzeyinde ölçülen iki değişken arasında ilişki olup olmadığı veya bu değişkenlerin bağımsız olup olmadığını test etmek için kullanılır [2]. Örneğin bir araştırmacı cinsiyet ile eğitim düzeyi arasında ilişki olup olmadığını bilmek isteyebilir. Bir eğitimci sınava hazırlık biçimi ile başarı düzeyi arasında ilişki olup olmadığını bilmek isteyebilir. Eğer bu tür iki değişken arasında ilişki yoksa, bu iki değişkenin bağımsız olduğu söylenir. Eğer iki değişken arasında ilişki yoksa yani değişkenler bağımsız ise, yığındaki birimlerden herhangi birinde değişkenlerden birinin değerini bilmek değişkenlerden diğerinin değerini tahmin etmede yardımcı olmaz.

Diğer yandan iki değişken arasında ilişki varsa, yığındaki birimlerden herhangi birinde değişkenlerden birinin değerini bilmek değişkenlerden diğerinin değerini tahmin etmede yardımcı olur. Bağımsızlık için ki-kare testinde hipotezler aşağıdaki gibidir [22].

H_0 : Değişkenler bağımsızdır.
H_1 : Değişkenler bağımsız değildir.

veya

H_0 : Değişkenler arasında ilişki yoktur.
H_1 : Değişkenler arasında ilişki vardır.

Test istatistiği aşağıdaki gibi ifade edilir.

$$\sum_{j=1}^{c}\sum_{i=1}^{r}\frac{\left(G_{ij}-B_{ij}\right)^2}{B_{ij}} \sim \chi^2_{(r-1)(c-1)} \tag{56}$$

Buradaki B_{ij} eşitlik (57) ‘de verilmiştir.

$$B_{ij}=\frac{\left(n_{.j}\right)\left(n_{i.}\right)}{n} \tag{57}$$

c, örneklem sayısı ve r, örneklem hacmini ifade eder.

Bağımsızlık için Ki-Kare testi varsayımları;

1) İlgilenilen yığından n hacimli örnek rastgele seçilmeli ve bu n sayıda örnek biriminde değişkenlerin değerleri saptanmalıdır,
2) Değişkenler sınıflama veya sıralama ölçme düzeyinde ölçülmüş olmalıdır.

2.7. İkiden Fazla Bağımlı Örnek İçin Testler

Aynı örnek birimlerinde bağımlı değişken bakımından ölçüm (gözlem) değerleri çok fazla değişkenlik gösterirse Kruskal-Wallis H testi ile gruplar arasındaki farkı ortaya çıkarmak kolay olmayabilir. Bazen ilgili bağımlı değişken bakımından gruplar arasındaki farklılık, aynı grup içindeki birimler arasındaki değişkenlik biçiminde gizlenebilir. Aynı grup içindeki birimler arasındaki değişkenlik etkisi "blok etkisi" olarak adlandırılır.

Örnek birimlerini "bloklar" olarak adlandırılan homojen alt gruplara ayırdıktan sonra ilgili bağımlı değişken bakımından gruplar (faktör düzeyleri) arasında farklılık olup olmadığının araştırılması problemi ile birçok alanda karşılaşılabilinir. Örneğin hastaları yaş gruplarına ayırdıktan sonra ilaçların etkilerinin farklı olup olmadığının araştırılması gibi. Burada yaş grupları bloklar olarak adlandırılır ve her bloktaki hastaların yaşları aynıdır. Bu yöntem istatistiksel deney tasarımında "Rastgele Tamamlanmış Blok Tasarımı" olarak da bilinir. Bu tür tasarımlarda bloklar; aynı yaş grubundan seçilen hastalar, aynı toprak türünden seçilen tarlalar, ağırlıkları aynı olan hastalar vb. olabilir. Rastgele tamamlanmış blok tasarımından sağlanan veriyi analiz etmek için kullanılan parametrik yöntem iki yönlü (faktörlü) varyans analizidir [10].

Bu çalışmada ikiden fazla bağımlı örnek için parametrik testlerden iki yönlü varyans analizi, iki yönlü çok değişkenli varyans analizi ve parametrik olmayan testlerden Friedman 'ın S testi ile Cochran 'ın Q testine yer verilmiştir.

2.7.1. İki Yönlü Varyans Analizi

İki bağımsız değişkenin bir bağımlı değişken üzerine etkisini araştırırken bağımsız değişkenlerin bağımlı değişken üzerine etkilerini ayrı ayrı araştırmak yerine, ikisinin etkileşiminin ortak etkisini araştırmak için iki yönlü varyans analizi kullanılır [10].

İki yönlü varyans analizinde n birimlik bir grupta; b blok, t işlem uygulaması sonuçlarının önemliliği test edilmektedir. t işlem sadece n birimlik grup (blok) üzerinde uygulanarak t kez deneme tekrarlanarak t bağımlı grupta elde edilen sonuçlar analiz edilmektedir. İki yönlü varyans analizi iki bağımsız değişkenin bir bağımlı değişken üzerindeki etkisini inceler [23]. İki yönlü varyans analizinde test istatistiğinin hesaplanmasına yönelik varyans analizi Tablo 11'de verilmiştir [19].

Tablo 11. İki Yönlü Varyans Analiz Tablosu

Değişimin Kaynağı	**Kareler Toplamı**	**Serbestlik Derecesi**	**Kareler Ortalaması**	**Oran**
Satırlar	$KT_A = \frac{1}{b}\sum_{i=1}^{a} T_{i.}^2 - \frac{1}{ab}T_{..}^2$	$a-1$	$KO_A = \frac{KT_A}{a-1}$	$F_A = \frac{KO_A}{KO_E}$
Sütunlar	$KT_B = \frac{1}{a}\sum_{i=1}^{b} T_{.j}^2 - \frac{1}{ab}T_{..}^2$	$b-1$	$KO_B = \frac{KT_B}{b-1}$	$F_B = \frac{KO_B}{KO_E}$
Hata	$KT_E = \sum_{i=1}^{a}\sum_{j=1}^{b} x_{ij}^2 - \frac{1}{ab}T_{..}^2$	$(a-1)(b-1)$	$KO_E = \frac{KT_E}{(a-1)(b-1)}$	
Toplam	$ab-1$			

İki yönlü varyans analizi testi için varsayımlar;

1) İncelenen değişken her bir örneklemde normal dağılım göstermelidir,
2) Grup ortalamaları ve standart sapmaları arasında bir doğrusallık olmalıdır,
3) Bağımlı değişkenin ölçüm düzeyi k sayıdaki her bir grup için oransal veya en az eşit aralıklı olmalıdır,
4) Bağımlı değişken bütün gruplarda normal dağılıma ve homojen (yaklaşık olarak eşit) varyanslara sahip olmalıdır.

2.7.2. İki Yönlü Çok Değişkenli Varyans Analizi

İki yönlü çok değişkenli varyans analizi, iki yönlü varyans analizinin bağımlı değişken sayısı p olmak üzere, $p > 2$ olduğu durumlar için genelleştirilmiş halidir.

$p \geq 2$ değişken sayısı, i işlem sayısı, j birim sayısı olmak üzere bağımlı gözlem;

$$X_{ijk} = \mu + \sigma_i + \beta_j + \alpha\beta_{ij} + e_{ijk} \tag{58}$$

biçiminde ifade edilir.

Burada μ, genel ortalama vektörünü; σ_i, işlem etkisini; β_j, birim etkisini; $\alpha\beta_{ij}$, birim ile işlem etkileşimini ve e_{ijk} ise rastgele hatayı belirtmektedir. Burada,

$i = 1,2, \dots, t; j = 1,2, \dots, b; k = 1,2, \dots, n$

olarak değer almaktadır. İki yönlü çok değişkenli varyans analizinde test istatistiğinin elde edilmesi Tablo 12'de verilmiştir [19].

Tablo 12. İki Yönlü Çok Değişkenli Varyans Analizi

Değişim Kaynağı	**KÇÇT Matrisi**	**Serbestlik Derecesi**
Birim (Blok)	$KÇÇT_{BR} = \sum_{j=1}^{b} tn(\bar{X}_{.j} - \bar{X})(\bar{X}_{.j} - \bar{X})'$	$t-1$
İşlem	$KÇÇT_{TR} = \sum_{i=1}^{t} bn(\bar{X}_{i.} - \bar{X})(\bar{X}_{i.} - \bar{X})'$	$b-1$
Etkileşim	$KÇÇT_{ETK} = \sum_{i=1}^{t}\sum_{j=1}^{b} n(\bar{X}_{ij} - \bar{X}_{i.} - \bar{X}_{.j} + \bar{X})(\bar{X}_{ij} - \bar{X}_{i.} - \bar{X}_{.j} + \bar{X})'$	$(t-1)$ $(b-1)$
Hata	$KÇÇT_{HATA} = \sum_{i=1}^{t}\sum_{j=1}^{b}\sum_{k=1}^{n} n(\bar{X}_{ijk} - \bar{X}_{.j})(\bar{X}_{ijk} - \bar{X}_{ik})'$	$tb(n-1)$
Genel	$KÇÇT_{GENEL} = \sum_{i=1}^{t}\sum_{j=1}^{b}\sum_{k=1}^{n} n(\bar{X}_{ijk} - \bar{X}_{.j})(\bar{X}_{ijk} - \bar{X}_{ik})'$	$tbn-1$

İki yönlü çok değişkenli varyans analiz testi için varsayımlar;

1) İncelenen örneklem çok değişkenli normal dağılım göstermelidir,
2) Grup ortalamaları ve standart sapmaları arasında bir doğrusallık olmalıdır,
3) Bağımlı değişkenin ölçüm düzeyi k sayıdaki her bir grup için oransal veya en az eşit aralıklı olmalıdır,
4) Bağımlı değişken bütün gruplarda normal dağılıma ve homojen (yaklaşık olarak eşit) varyanslara sahip olmalıdır.

2.7.3. Friedman 'ın S Testi

Friedman' ın S Testi parametrik iki-yönlü (faktörlü) varyans analizi yönteminin parametrik olmayan alternatifidir [16]. Bu testin temel özelliği normallik ve homojenlik varsayımı gerektirmemesi ve ölçüm değerlerine atanan büyüklük sıra sayılarına dayanmasıdır.

Friedman ’ın S testi için varsayımlar;

1) Veri, her biri c hacimli n sayıda bağımsız bloklardan derlenir,
2) İlgilenilen bağımlı değişken süreklidir,
3) Bloklar ve işlemler arasında etkileşim yoktur,
4) Ölçme düzeyi en az sıralamadır.

2.7.4. Cochran ‘ın Q Testi

Bağımlı değişkenin ölçüm değeri, iki değer alabilen bir değişken özelliği taşıyorsa işlemlerin etki bakımından aynı olup olmadıklarını belirlemek için kullanılabilecek testlerden biri de Cochran ‘ın Q testidir [2]. Örneğin A, B, C, D ve E ilaçları farklı yaş gruplarında birer hastaya uygulandıktan sonra, hastaların “tedaviye olumlu tepki vermedi” veya “tedaviye olumlu tepki verdi” biçiminde değerlendirilmesi gibi. Burada “tedaviye olumlu tepki vermedi” için “0” ve “tedaviye olumlu tepki verdi” için de “1” verilebilir.

Cochran ‘ın Q testi için varsayımlar;

a) Veri, her biri c hacimli n sayıda bağımsız bloklardan derlenir,
b) Bloklar ve işlemler arasında etkileşim yoktur,
c) Ölçme düzeyi en az sınıflamadır.

2.8. İlişki Analizi

Biri sınıflama diğeri sıralama ölçme düzeyinde ya da ikisi de sınıflama ölçme düzeyinde ölçülen iki değişken arasında ilişki olup olmadığının belirlenmesi için ki kare testleri kullanıldığına daha önce değinilmişti. Eğer aralarında ilişki olup olmadığı araştırılan iki değişken de oranlama ve/veya eşit aralıklı ölçme düzeyinde ölçülmüşse, yani değişkenlerin ikisi de sayısal değişkenler ise, bu amaçla kullanılabilen istatistiklerden biri örnek korelasyon katsayısıdır [2, 24–26]. Hesaplama sonucu elde edilen katsayı $[-1,1]$

aralığında değer alır. Hesaplanan değer “-1” ise ters yönlü ilişki, “0” ise ilişki yok, “1” ise aynı yönlü ilişki vardır denir.

Bu çalışmada normallik varsayımı gerektiren Pearson Korelasyon ile normallik varsayımı gerektirmeyen Spearman Korelasyonu ve Kendall ‘ın τ katsayısına yer verilmiştir.

2.8.1. Pearson Korelasyon Katsayısı

Pearson korelasyon katsayısı iki değişken arasında ilişki olup olmadığının kontrolü için kullanılır [10, 27].

$$r_{xy} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \tag{59}$$

Pearson Korelasyon katsayısı için varsayımlar;

1) Veri eşit aralıklı veya oranlama ölçme düzeyinde ölçülmüş olmalıdır,
2) Her iki örneğin de normallik varsayımını sağlaması gerekmektedir.

2.8.2. Spearman Sıra Korelasyon Katsayısı

Spearman sıra korelasyon katsayısı iki değişken arasında ilişki olup olmadığının kontrolü için kullanılır [2, 28]. Veriye atanan büyüklük sıra sayılarına dayanır ve

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2 - 1)} \tag{60}$$

ile elde edilir. Burada;

$$d_i = R(X_i) - R(Y_i) \tag{61}$$

dir.

Spearman sıra korelasyon katsayısı için varsayımlar;

1) Veri en az sıralama ölçme düzeyinde ölçülmüş olmalıdır,
2) Her iki örnek için de normallik varsayımı gerektirmez,
3) Doğrusallık ve eş varyanslılık gerektirir.

2.8.3. Kendall 'ın τ İlişki Katsayısı

Kendall 'ın τ ilişki katsayısı iki değişken arasında ilişki olup olmadığının kontrolü için kullanılır [2, 29]. X_i ve Y_i aralarında ilişki olup olmadığı araştırılan iki örnek ve,

p_i : Her Y_i değeri için doğal sıranın sağlandığı durum sayısı

q_i : Her Y_i değeri için doğal sıranın sağlanmadığı durum sayısı

olmak üzere, τ ilişki katsayısı;

$$\hat{\tau} = \frac{P - Q}{n(n-1)/2} \tag{62}$$

ile elde edilir. Burada P ve Q;

$$P = \sum p_i \text{ ve } Q = \sum q_i \tag{63}$$

dir.

Kendall 'ın τ ilişki katsayısı için varsayımlar;

1) Veri en az eşit aralıklı ölçme düzeyinde ölçülmüş olmalıdır,
2) Her iki örnek için de normallik varsayımı gerektirmez.

3. ÖNCEKİ ÇALIŞMALAR

Andy Field & Graham Hole [30] istatistiksel test seçimi için beş sorunun cevaplanması gerektiğini ileri sürmüş ve karar ağaçları yayınlamışlardır. Bu sorular aşağıda verilmiştir.

1) Verilerin türü nedir?
2) Kaç tane bağımsız değişken kullanılacak?
3) Ne tür bir tasarım (deneysel ya da ilişkisel) kullanılacak?
4) Elde edilen ölçümler tekrarlı ölçümler mi, bağımsız ölçümler mi?
5) Veri parametrik mi yoksa parametrik olmayan yapıda mı?

Nancy L. Leech, Karen C. Barrett, George A. Morgan [31] istatistiksel test seçimi için üç temel adım belirlemiş ve buna uygun karar ağaçları ile tablolar yayınlamışlardır.

1) Araştırma sorusunda ya da hipotezinde kaç değişken (bağımlı, bağımsız ayrımı yapılmadan) olduğu belirlenir.
2) İki değişken söz konusu ise bağımsız değişkenlerin türü belirlenir.
3) Üç veya daha fazla değişken söz konusu ise bağımlı değişken sayısı bulunur ve türleri belirlenir.

Mertler, C.A., Vannatta, R.A. [32] istatistiksel test seçimi için dört temel adım belirlemişlerdir.

1) Araştırma sorusundaki değişkenler tanımlanır.
2) Hangi değişken ya da değişkenlerin bağımsız, hangilerinin bağımlı değişken olduğunu ve kontrol (kovaryant) değişken olup olmadığı belirlenir.
3) Tüm değişkenlerin (nitel ve nicel) türü belirlenir.
4) Araştırma sorusunun amacı belirlenir (ilişki derecesi, grup farklılıkları, grup üyelik tahmini, yapı).

Dr. Jaykaran Charan [3] uygun istatistiksel testi belirlemenin üç soruya bağlı olduğunu ileri sürmüş ve karar ağacı yayınlamıştır.

1) Çalışmada kullanılan verinin türü nedir?
2) Veri normal dağılıma uygun mudur, değil midir?
3) Yapılan çalışmanın amacı nedir?

Doç. Dr. Seval Kul [24] istatistiksel test seçimini etkileyen en önemli faktörleri aşağıdaki şekilde sıralamıştır.

1) Hipotezin türü: ilişki mi, fark mı araştırılıyor?
2) Bağımlı değişkenin ölçme düzeyi: sayısal ya da sözel ifade edilen değişken.
3) Bağımsız değişkenin ölçme düzeyi: sayısal ya da sözel ifade edilen değişken.
4) Sayısal değişkenin normal dağılıma uygunluğu.

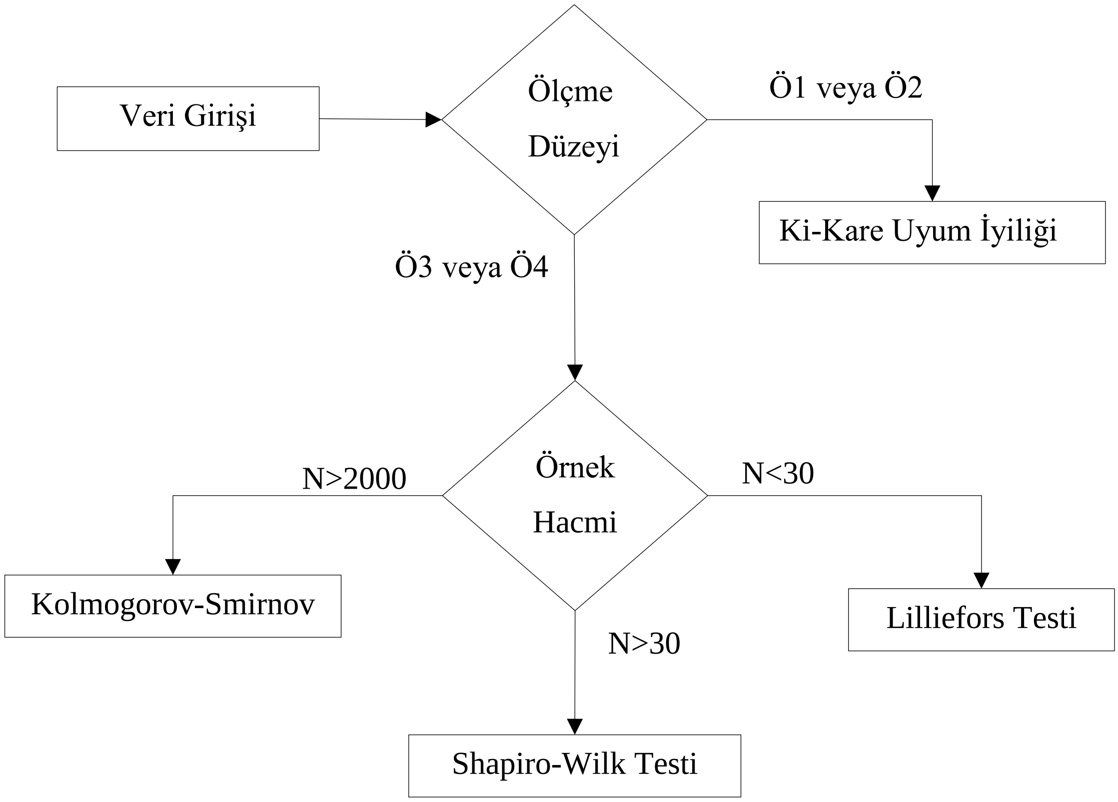
4. YAPILAN ÇALIŞMA

Yapılan çalışmada yaygın olarak kullanılan bazı istatistiksel analiz yöntemleri incelenerek her bir testin gerektirdiği özel varsayımları dikkate alan bir istatistiksel test seçim sihirbazı geliştirilmiştir. Bu sihirbaz, Veri Analiz Platformu (VAP) [33] isimli yazılım programına dahil edilerek geliştirilmiştir. Seçim sihirbazı, kullanıcı tarafından programa girilen veri setlerine gerekli varsayım kontrollerini yaparak verinin yapısına ve araştırmanın amacına en uygun testleri önerecek şekilde çalışmaktadır.

Sihirbaz aşağıdaki dört temel adımda sonuca ulaşmaktadır.

1) Değişkenlerin ölçme düzeyi kontrolü.
2) Normallik ve homojenlik kontrolü.
3) Bağımlı ve bağımsız değişkenlerin belirlenmesi.
4) Araştırmanın amacına (ilişki derecesi, gruplar arası farklılıklar, grup tahmini) yönelik test önerisi.

Adım 1: Normallik kontrolü, veri girişi sonrasında değişkenlerin ölçme düzeyleri, örnek sayısı ve örneklem hacmi (N) dikkate alınarak verinin yapısına uygun olan normallik testleri ile sağlanır. Uygulanan yöntemin algoritması aşağıda verilmiştir.



Şekil 1. Normallik Kontrolü

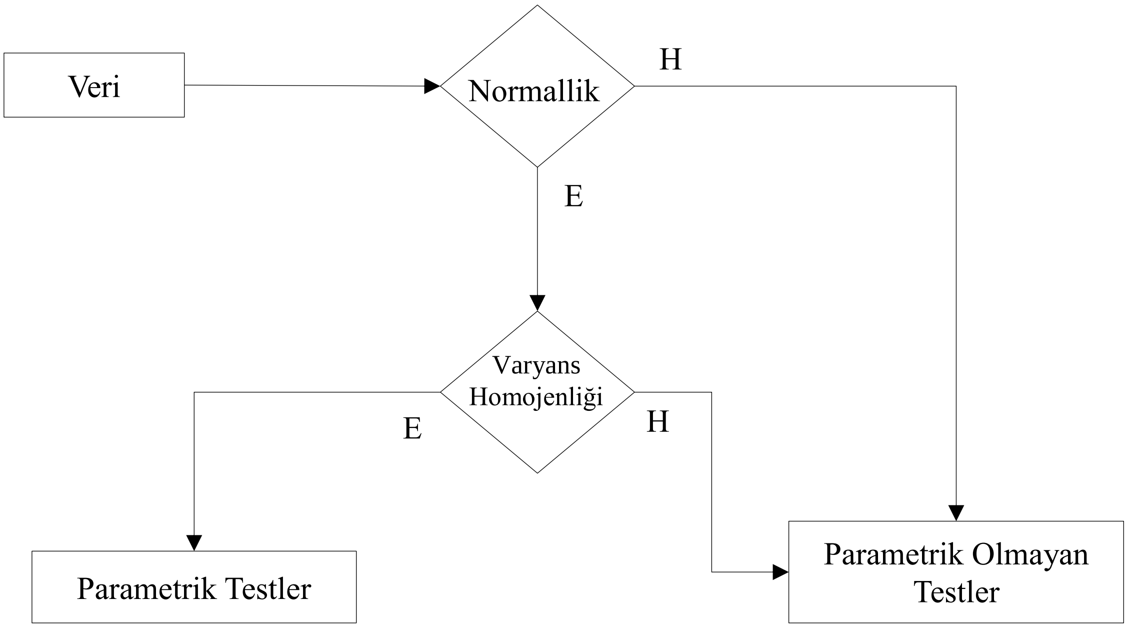
Adım 2: Homojenlik kontrolü, gerekli olduğu durumlarda, Levene testi veya çok değişkenli analizler için Box M testi ile sağlanır.

Adım 3: Değişkenlerin ölçme düzeyleri iki şekilde belirlenebilir. Bunlardan birincisi ölçme düzeylerinin araştırmacı tarafından sihirbaza girişinin yapılmasıdır. İkincisi ise değişkenlerin ölçme düzeylerinin test seçim sihirbazı tarafından tahmin edilmesidir. Bu yöntemle ikili (dikotom) veriler, sınıflama ölçme düzeyi ile ölçülmüş sembollerden oluşan veriler ve oranlama ölçme düzeyi ile ölçülmüş sayısal veriler tahmin edilebilmektedir. Ancak değişkenlerin ölçme düzeyi bilgisi, araştırmanın konusuna ve kullanılan ölçüm yöntemine bağlı olduğu için araştırmacı tarafından sihirbaza giriş yapılması daha doğru sonuçlar vermektedir.

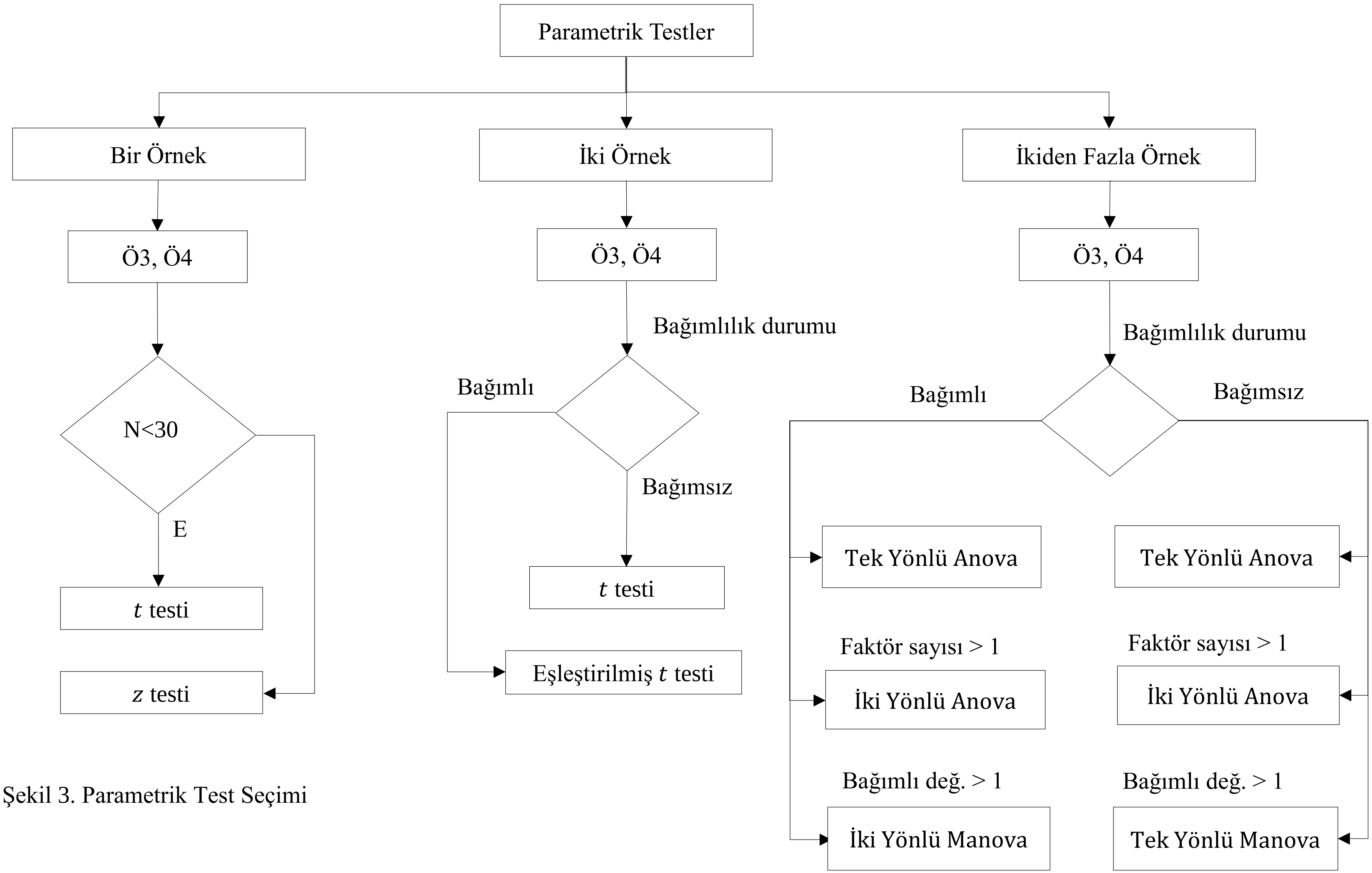
Adım 4: Bağımlı ve bağımsız değişkenler ilgilenilen problemin yapısına göre değişiklik gösterdiği için araştırmacı tarafından test sihirbazına girişi yapılmalıdır.

Adım 5: Önceki adımlarda elde edilen bilgiler doğrultusunda örnek sayısına göre en uygun testler araştırmacıya önerilir. Bu aşamada; ilişki derecesi, gruplar arası farklılıklar, grup üyelik tahmini vb. amaçlar için kullanılabilecek olan testler araştırmacıya önerilerek seçim sihirbazı tamamlanır.

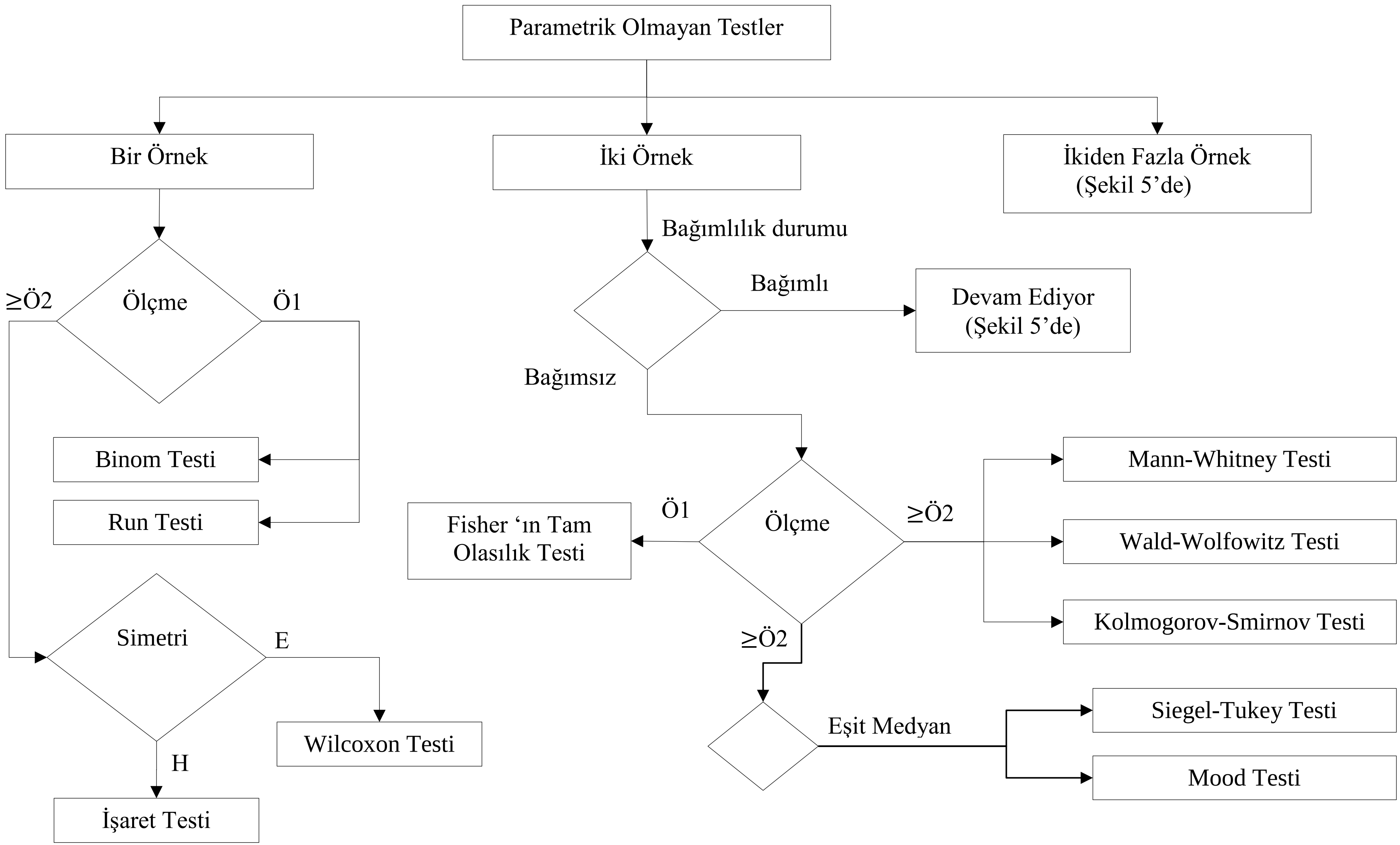
Bu çalışmada hipotez testleri ve ilişki analizine yönelik geliştirilen test seçim sihirbazının algoritması aşağıda verilmiştir.



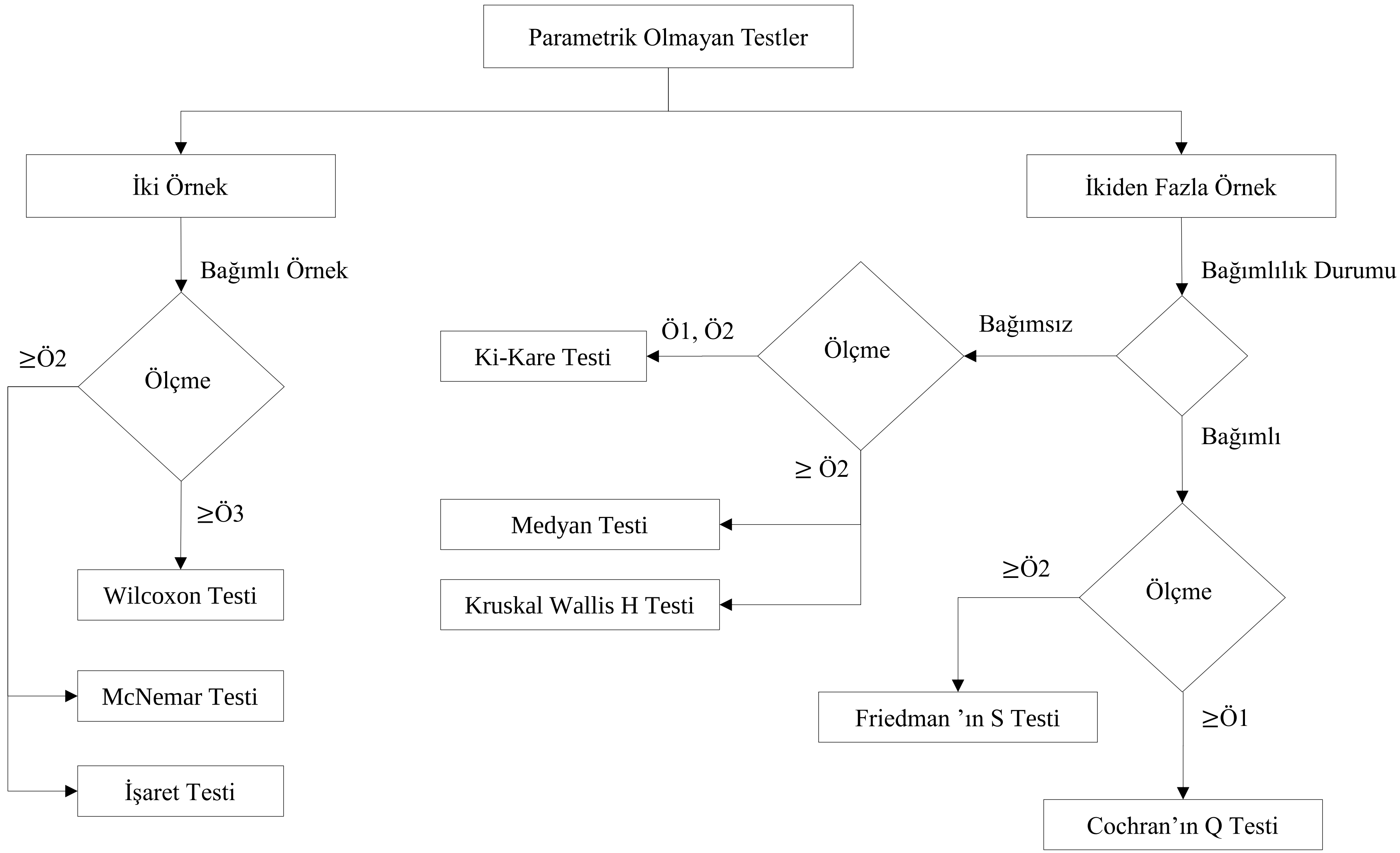
Şekil 2. Parametrik ve Parametrik Olmayan Test Seçimi



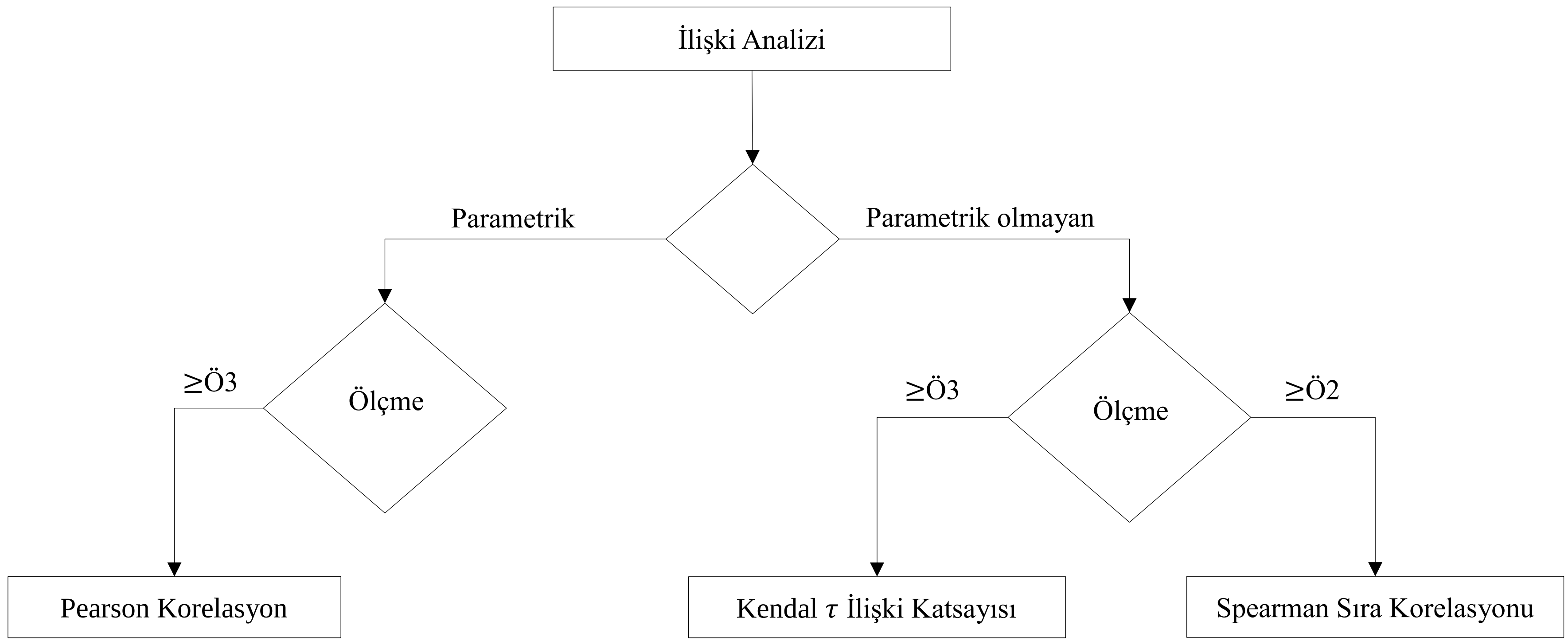
Şekil 3. Parametrik Test Seçimi



Şekil 4. Parametrik Olmayan Test Seçimi 1



Şekil 5. Parametrik Olmayan Test Seçimi 2



Şekil 6. İlişki Analizi

4.1. Örnek Uygulamalar

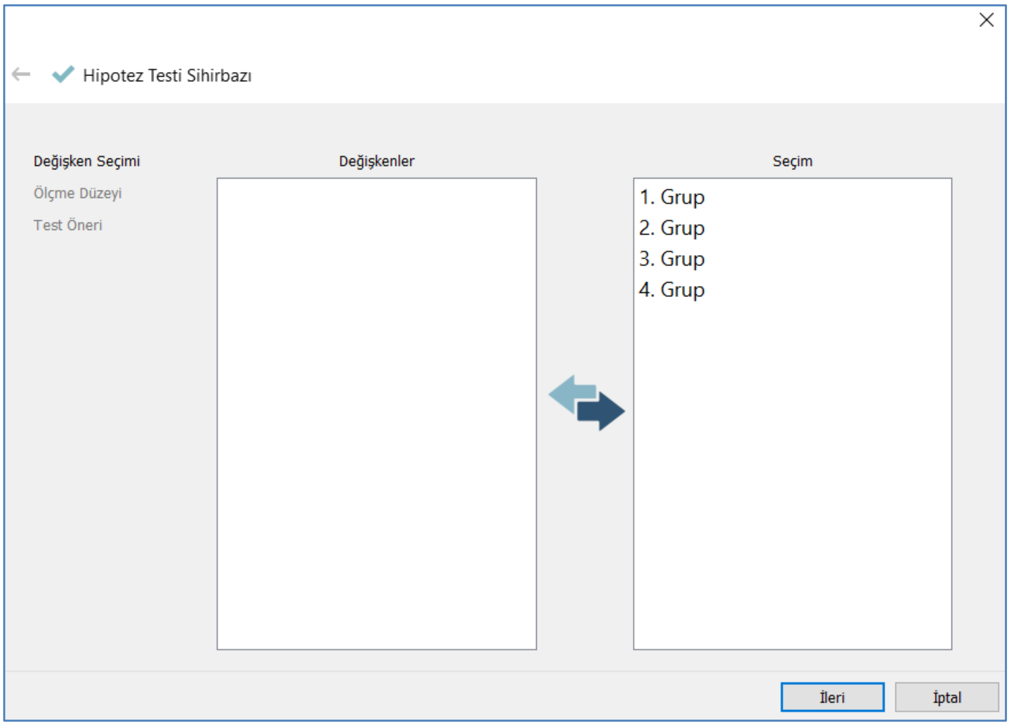
Geliştirilen yazılım birçok farklı veri kullanılarak test edilmiştir. Bunlardan biriyle yapılan seçim aşamaları aşağıdaki gibi gerçekleştirilmiştir.

Örnek 1: Aynı hastalığa yakalanmış hastalar içinden 23 hasta rastgele seçilmiş ve yine rastgele 4 gruba ayrılmıştır. Gruplardaki hasta sayıları, sırasıyla 5,6,5 ve 7 'dir. 1. gruptaki hastalara A, 2. gruptaki hastalara B, 3. gruptaki hastalara C ve 4. gruptaki hastalara D tedavi yöntemi uygulanmıştır [2]. Hastaların kaç günde iyileştikleri Tablo 13'de verilmiştir.

Tablo 13. Hastaların İyileşleşme Süreleri

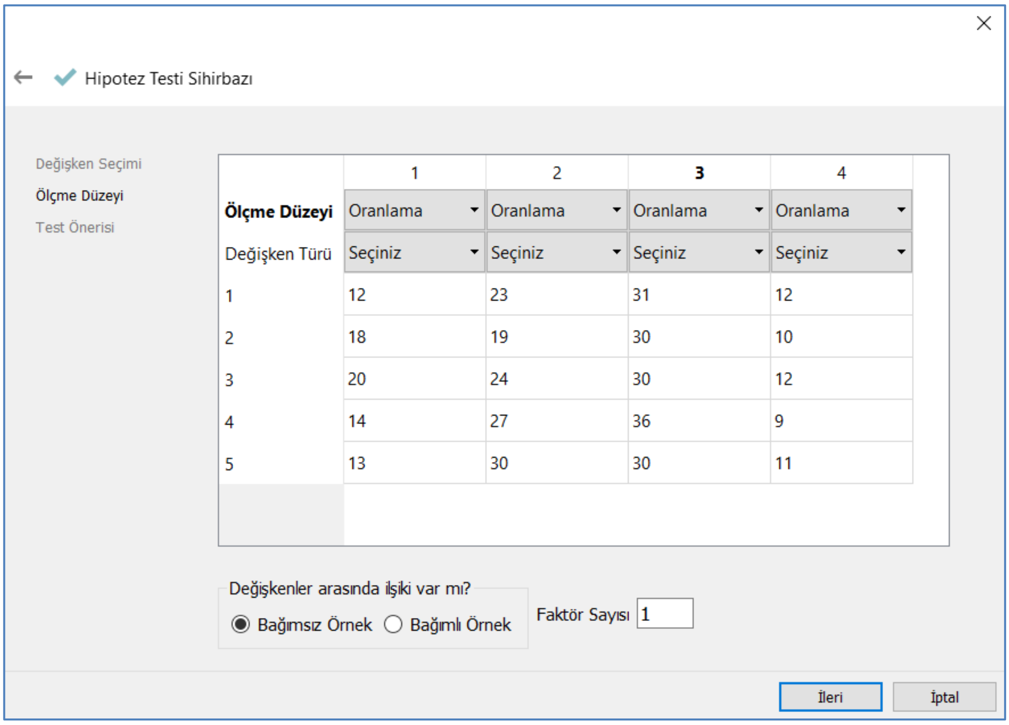
Tedavi yöntemi			
1.grup (A)	2.grup (B)	3.grup (C)	4.grup (D)
12	23	31	12
18	19	30	10
20	24	30	12
14	27	36	9
13	30	30	11
	30		7
			6

Örnek verileri VAP programına girilerek hipotez seçim sihirbazı çalıştırıldığında ilk olarak Şekil 7'deki ekran açılır. Burada kullanılacak örneklerin seçilmesi istenir.

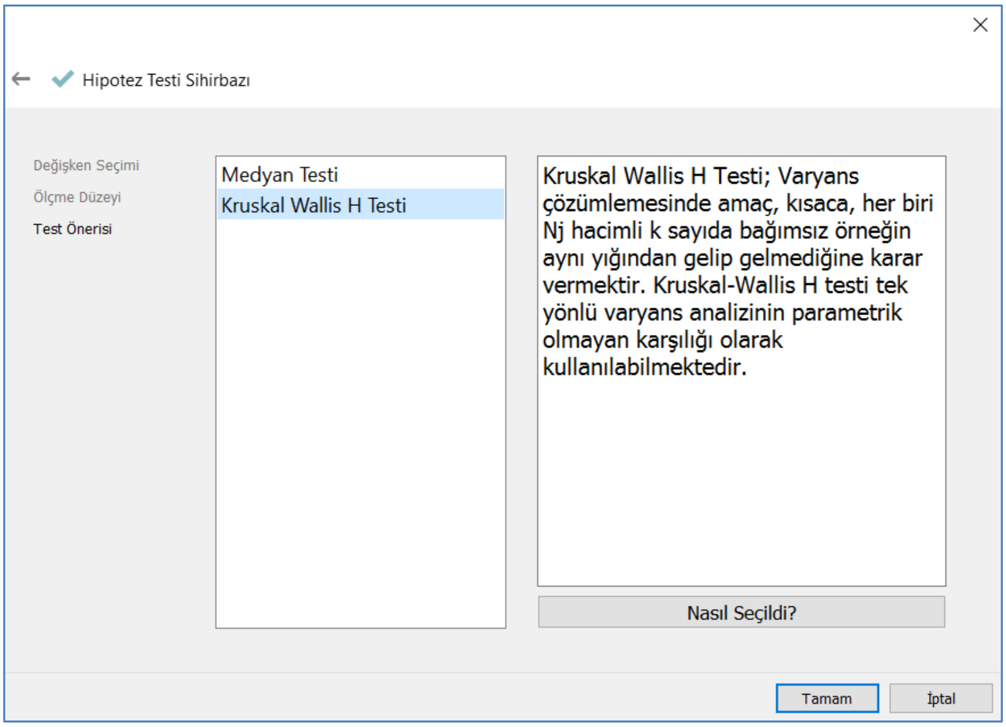


Şekil 7. Test Seçim Sihirbazı Değişken Seçim Ekranı (Örnek 1)

Değişkenler bölümünden sihirbazda kullanılacak örnekler seçildikten sonra "İleri" tuşu tıklanır ve sihirbaz tarafından tahmin edilen ölçme düzeyi bilgileri kullanıcı tarafından onaylanır. Bunun yanında problemin yapısına bağlı olarak varsa "değişken türü" seçeneğinde "bağımlı değişkenler, bağımsız değişkenler ve etiket verileri" belirlenir. Son olarak Şekil 8'de görüldüğü gibi örneklerin bağımlı ya da bağımsız olma durumları işaretlenerek "ileri" tuşu tıklanır ve sihirbaz tamamlanır.

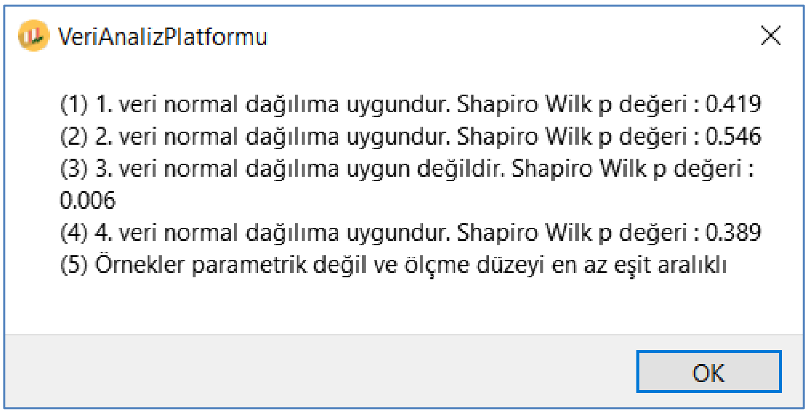


Şekil 8. Test Seçim Sihirbazı Ölçme Düzeyi Belirleme Ekranı (Örnek 1)



Şekil 9. Test Seçim Sihirbazı Sonuç Ekranı (Örnek 1)

Şekil 9‘da görülen sonuç ekranında kullanıcıya verinin yapısına bağlı olarak kullanabileceği testlerin listesi ve bu listedeki herhangi bir teste tıklandığında, testle ilgili kısa bilgilerin verildiği açıklama bölmesi gösterilmektedir. “Nasıl Seçildi?” tuşuna tıklandığında sihirbazın yaptığı seçimi neye göre gerçekleştiğinin açıklaması gösterilmektedir. Bu açıklama Şekil 10’da gösterilmiştir.



Şekil 10. Örnek 1'e Ait Test Seçim Açıklaması

Bu aşamadan sonra kullanıcı “Tamam” tuşuna basarak sihirbazı tamamladığında, seçilen test VAP üzerinde gerçekleştirilerek kullanıcıya sunulmaktadır. Elde edilen sonuç tablosu Şekil 11’da verilmiştir.

Kruskal Wallis Testi

İstatistikler	Sonuçlar
Toplam Veri Sayısı	23
Test İstatistiği	19.648
Serbeslik Derecesi	3
p değeri	0.000

Şekil 11. Kruskal Wallis Testi Sonuç Tablosu

Kruskal Wallis testi sonucunda $p < 0.05$ olarak elde edilmiştir. Bu sonuca göre gruplar arasında anlamlı bir farklılık vardır denir ve çoklu karşılaştırma yapmak gerekmektedir. VAP üzerinde yapılan çoklu karşılaştırma tablosu Şekil 12’de verilmiştir.

Çoklu Karşılaştırma

Gruplar	Test İstatistiği	Standart Hata	Standart Test ist.	p değeri	Düzeltilmiş p değeri	Etki Boyutu
1. Grup - 2. Grup	-5.833	4.107	-1.420	0.155	0.933	-0.731
1. Grup - 3. Grup	-10.400	4.290	-2.425	0.015	0.092	0.466
1. Grup - 4. Grup	5.857	3.971	1.475	0.140	0.842	-0.321
2. Grup - 3. Grup	-4.567	4.107	-1.112	0.266	1.000	0.934
2. Grup - 4. Grup	11.690	3.773	3.098	0.002	0.012	1.135
3. Grup - 4. Grup	16.257	3.971	4.094	0.000	0.000	0.000

Şekil 12. Kruskal Wallis Testi Çoklu Karşılaştırma Tablosu

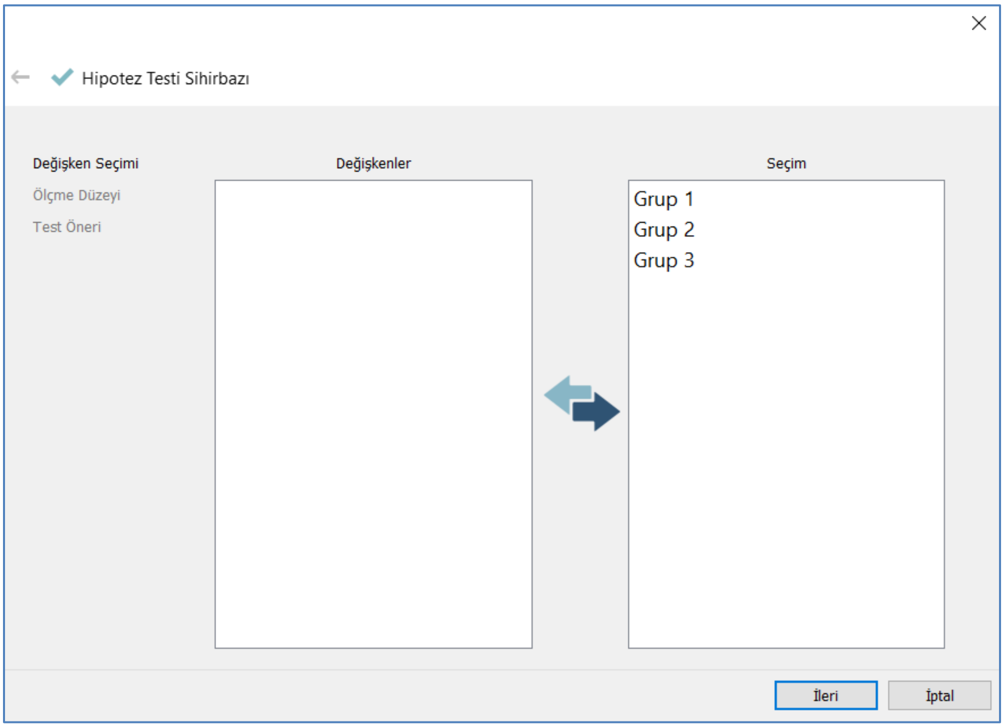
Çoklu karşılaştırma tablosunda $p < 0.05$ olan “2. Grup ile 4. Grup” ve “3. Grup ile 4. Grup” arasında anlamlı bir farklılık olduğu %5 güven düzeyi ile söylenebilir.

Örnek 2: Üç öğrenci grubuna üç farklı matematik öğretim tekniği uygulanmakta ve bu farklı tekniklerin elde edilen notlar üzerinde etkili olup olmadıkları %5 anlamlılık düzeyinde belirlenmek istenmektedir [34]. Üç öğrenci grubuna ait notlar Tablo 14’te verilmiştir.

Tablo 14. Öğrencilerin Elde Ettikleri Notlar

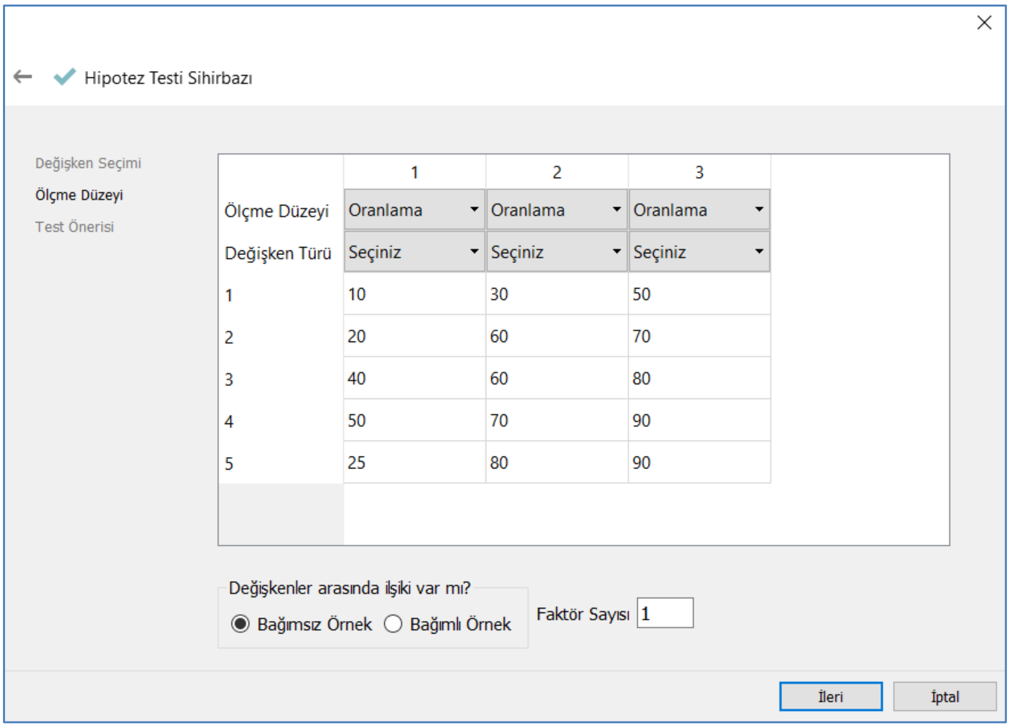
	Öğretim Tekniği	
1.grup	2.grup	3.grup
10	30	50
20	60	70
40	60	80
50	70	90
25	80	90
30		100

Örnek verileri VAP programına girilerek hipotez seçim sihirbazı çalıştırıldığında ilk olarak Şekil 13’deki ekran açılır ve burada kullanılacak örneklerin seçilmesi istenir.

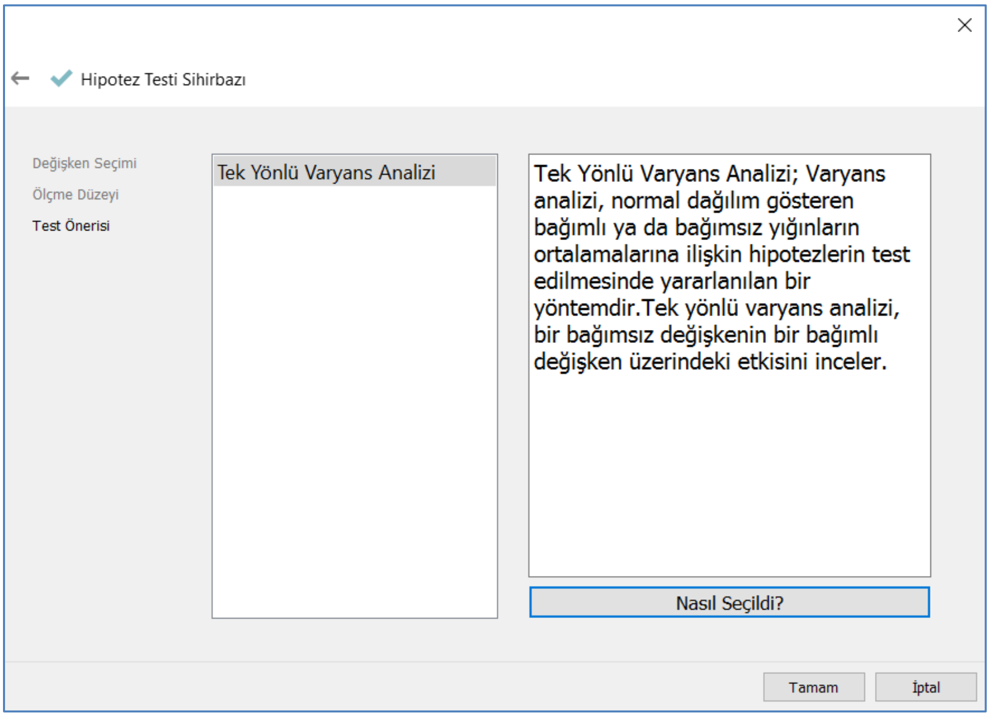


Şekil 13. Test Seçim Sihirbazı Değişken Seçim Ekranı (Örnek 2)

Sihirbazda kullanılacak örnekler seçildikten sonra değişkenlerin ölçme düzeyi tahmin edilerek kullanıcıdan bu tahmini onaylaması ve değişkenlerin türünü belirlemesi istenir. Bu işlem Şekil 14’te gösterilmiştir.



Şekil 14. Test Seçim Sihirbazı Ölçme Düzeyi Belirleme Ekranı (Örnek 2)



Şekil 15. Test Seçim Sihirbazı Sonuç Ekranı (Örnek 3)

Şekil 15’te görülen sonuç ekranında sihirbaz öğretim tekniklerinin notlar üzerinde etkisi olup olmadığının kontrolü için tek yönlü varyans analizini önermiştir. Sihirbaz tamamlanıp varyans analizi yapıldığında Şekil 16’te görülen tablo elde edilir.

Tek Yönlü Varyans Analizi

	Kareler Toplamı	SD	Kareler Ortalaması	p değeri
Gruplar Arası	15712.809	5	3142.562	0.000
Gruplar İçi	8041.667	28	287.202	
Toplam	23754.477	33		

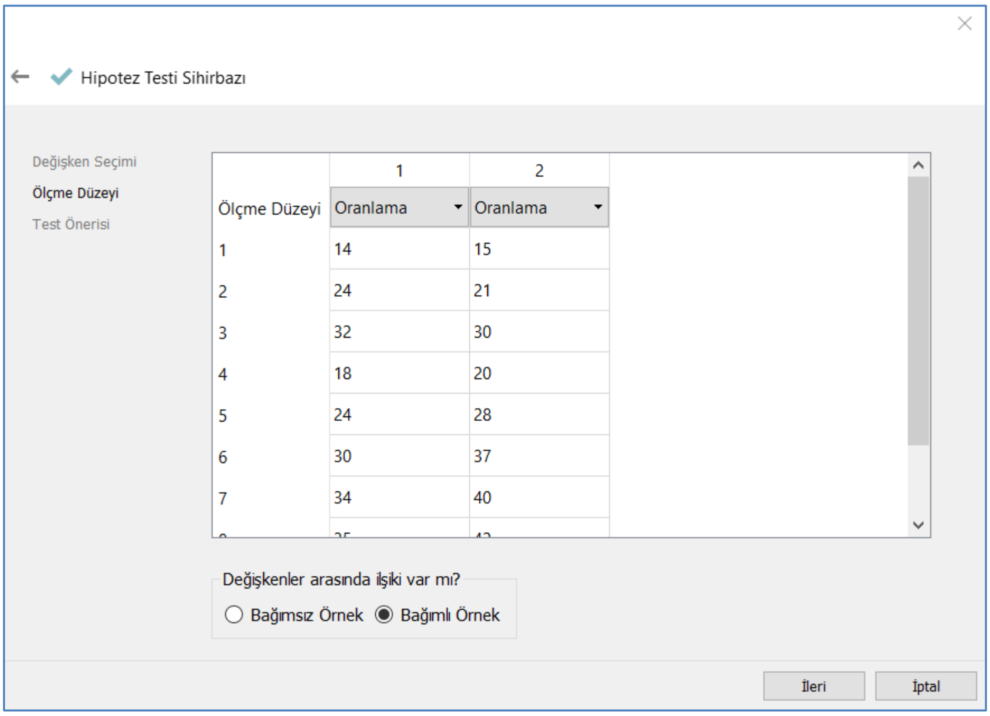
Şekil 16. Örnek 2'ye Ait Tek Yönlü Varyans Analizi Tablosu

Örnek 3: Rastgele seçilen 9 hastanın işlem öncesi ve işlem sonrası tahlil sonuçları Tablo 15’de verilmiştir. Anlamlılık düzeyi $\alpha = 0.05$ alınırsa ilgili işlemin tahlil sonucunda fark oluşturduğu söylenebilir mi?

Tablo 15. Hastaların İyileşme Süreleri

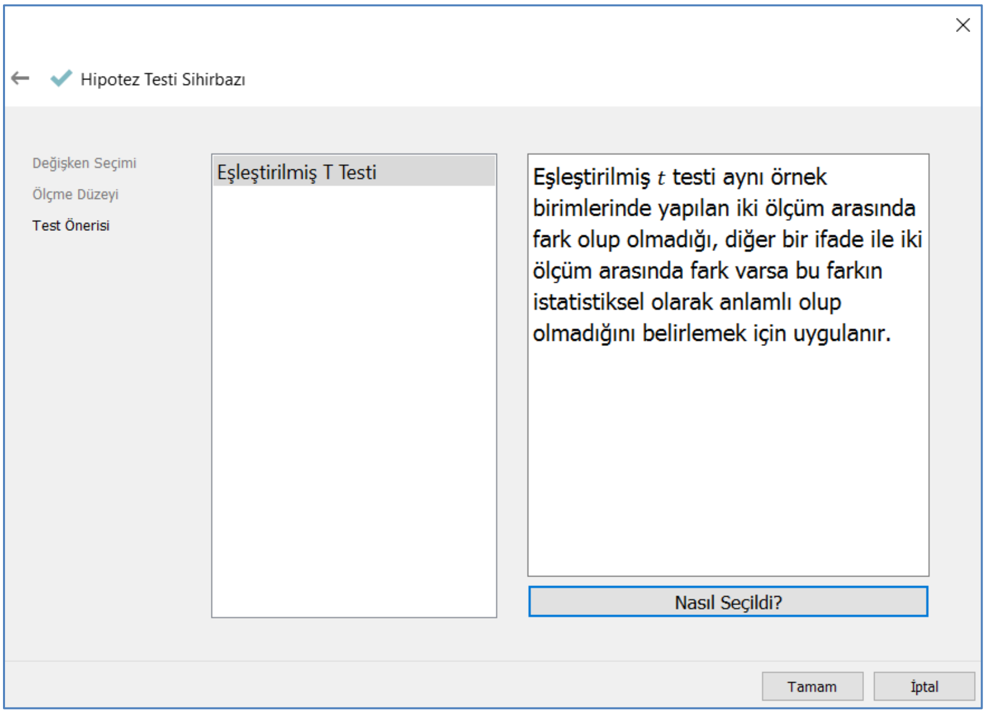
Hasta No	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
X (önce)	14	24	32	18	24	30	34	35	38
Y (sonra)	15	21	30	20	28	37	40	42	39

Örnek verileri VAP programına girilerek hipotez seçim sihirbazında ölçme düzeyi belirleme ekranında Şekil 17'daki gibi oluşmaktadır.



Şekil 17. Test Seçim Sihirbazı Ölçme Düzeyi Belirleme Ekranı (Örnek 3)

"İleri" tuşuna tıklandığında sonuç ekranı Şekil 18'deki gibi oluşmaktadır.



Şekil 18. Test seçim sihirbazı sonuç ekranı (Örnek 3)

Şekil 18'de görülen sonuç ekranında sihirbaz, tahlil sonuçlarında fark olup olmadığının kontrolü için Eşleştirilmiş t Testini önermiştir. Sihirbaz tamamlanarak önerilen test yapıldığında Şekil 19'daki çıktı oluşur.

Eşleştirilmiş T Testi

	Ortalama	Standart Sapma	Standart Hata	t hesap	SD	p değeri (iki yanlı)
X - Y	-2.556	3.712	1.237	-2.065	8.000	0.073

İstatistikler

	Ortalama	N	Standart Sapma	Standart Hata
X	27.667	9.000	8.155	2.718
Y	30.222	9.000	9.897	3.299

Şekil 19. Örnek 3'e Ait Eşleştirilmiş t Testi Tablosu

Bu sonuca göre işlem öncesi ve sonrasına ait tahlil sonuçları arasında fark olmadığı %5 anlamlılık düzeyi ile söylenebilir.

4.2. Karşılaştırma

Günümüzde istatistiksel analiz alanında birçok yazılım bulunmaktadır. Bu yazılımların çoğu araştırmacının istatistiksel analiz konusunda bilgi sahibi olduğu varsayımı ile geliştirilmişlerdir. Yakın zamanda yapılan çalışmalar ışığında bu yazılımlara otomatik test seçim mekanizmaları eklenmeye çalışılmıştır. Bu çalışma kapsamında geliştirilen VAP yazılımının otomatik test seçim sihirbazı, günümüzde en çok tercih edilen istatistiksel analiz yazılımı olan SPSS [35] paket programında yer alan otomatik test seçim yöntemi ile karşılaştırılmıştır [33].

Örnek 1 ‘e ait verilerin SPSS programına girişi yapılmış ve sonrasında “Analyze >> Nonparametric Tests >> Independent Samples” penceresi üzerinde “Groups” ve “Test Fields” alanları doldurularak “Run” düğmesine tıklanarak seçim tamamlanmıştır. Bu işlemler sonucunda SPSS veriye uygun olan Kruskal Wallis H testini kendisi seçerek gerçekleştirmiştir. SPSS üzerinde yapılan seçim işlemi bu çalışmada geliştirilen seçim sihirbazı ile karşılaştırıldığında ortaya çıkan en belirgin fark, SPSS ‘te gerçekleştirilen otomatik seçim işleminin örnek sayısına, parametrik olup olmama durumuna ilişkinin seçim sonrasında gerçekleştiriliyor olmasıdır. Daha da önemlisi, parametrik veya parametrik olmayan test kullanımının yine kullanıcıya bırakılmasıdır. Bu çalışmada önerilen yöntemde ise tamamıyla otomatik bir test seçim sistemi oluşturulmuş ve başarıyla gerçekleştirilmiştir.

5. SONUÇLAR VE ÖNERİLER

Bu çalışmada sıklıkla kullanılmak durumunda kalınan istatistiksel analiz yöntemleri incelenmiş ve her bir yöntemin özel varsayımları dikkate alınarak bir test seçim sihirbazı algoritması geliştirilmiştir. Bu algoritmanın iki temel amacı vardır. Bunlardan birincisi; test seçimi konusunda kararsız kalan araştırmacılara yardımcı olmak, ikincisi; araştırmacı tarafından seçilen testlerin varsayım kontrolünü otonom olarak gerçekleştiren bir mekanizma şeklinde çalışarak yanlış test seçimi yapıldığında kullanıcıyı uyarmaktır. Kullanıcı sihirbazdan bağımsız olarak veri setine analiz işlemleri gerçekleştirirken, kullandığı verinin yapısına uygun olmayan yöntem seçtiğinde program tarafından uyarılmaktadır. Örneğin, normallik varsayımını sağlamayan bir veri setine parametrik bir test yapılmak istendiğinde, sihirbaz gerekli normallik kontrollerini otonom olarak gerçekleştirmekte ve kullanıcıyı uyarmaktadır. Geliştirilen sihirbaz algoritması VAP programına bu amaçlar doğrultusunda dahil edilmiştir.

Bu çalışma boyunca yapılan araştırmalardan elde edilen bilgiler doğrultusunda, istatistiksel test seçimini etkileyen temel faktörler aşağıdaki gibi elde edilmiştir.

1) Değişkenlerin ölçme düzeyi
2) Örneklem hacmi
3) Örneklerin normallik varsayımını sağlayıp sağlamadığı
4) Bağımlı ve bağımsız değişkenlerin sayısı

Bu faktörlerden "bağımlı ve bağımsız değişkenlerin sayısı" ilgilenilen problemin yapısına bağlı olduğundan doğru sonuçlar alabilmek için araştırmacı tarafından belirlenmelidir. Diğer faktörler geliştirilen yazılım tarafından hesaplanmakta veya uygun yöntemlerle tahmin edilmektedir. Ancak değişkenlerin ölçme düzeyi bilgisi her ne kadar algoritma tarafından tahmin ediliyor olsa dahi, bu tahmin kesinlik göstermemektedir. Çünkü değişkenlerin ölçme düzeyi bilgisi de problemin ve gerçekleştirilen ölçmenin yapısına bağlıdır. Bu noktada araştırmacı tarafından girilmesi gereken bilgilerin de bilgisayar tarafından yeterli bir doğrulukla tahmin edilebilmesi için test seçim algoritmasına Veri Madenciliği yöntemlerin dahil edilmesinin faydalı olacağı öngörülmektedir.

Bu çalışma kapsamında geliştirilen seçim sihirbazının, hem istatistik eğitimi veren kurumların ilgili derslerine yardımcı olması hem de başlangıç seviyesinde olan araştırmacılara yol göstermesi amaçlanmıştır. Elde edilen sonuçlar göz önüne alındığında, önerilen yöntem birçok araştırmacının gerçekleştirdiği istatistiksel analizlerin doğruluğunu arttıracaktır. Özellikle, gerçekleştirilen analizlerde kullanılan testlerin varsayımlarının otomatik olarak kontrol edilmesi, gerçekleştirilen analizlerin doğruluğunu arttıracaktır. Ayrıca, istatistikçilerin kontrol etmesi gereken testlerin otomatik olarak gerçekleştirilmesi sayesinde istatistiksel analizler için gereken süre de azalacaktır.

İlerleyen çalışmalarda sihirbaza veri madenciliği yöntemleri eklenerek tahmin etme gücü arttırılabilir. Bu anlamda veri madenciliği yöntemleri özellikle değişkenlerin ölçme düzeyi bilgisini doğru tahmin etmek ve hangi veriye ne tür analizler uygulandığını VAP üzerinde takip ederek benzer veriye sahip kullanıcılara önerilerde bulunmak için kullanılabilir. Ayrıca, geliştirilen yöntem çok değişkenli analizler, regresyon analizi, yöneylem araştırması gibi alanlarda gerçekleştirilecek analizlere yardımcı olacak şekilde genişletilebilir. Özellikle farklı disiplinlerde çalışan araştırmacıların kullandıkları istatistiksel yöntemlerin doğrulanması için çalışmanın genişletilmesi gerekmektedir.

Önerilen yöntemin başarısının arttırılması için, araştırmacıların deney tasarımları ve örneklem seçimi yaklaşımlarının da sürece dahil edilmesi gerekmektedir. Bu amaçla, geliştirilen yaklaşıma deney tasarımı ve örneklem seçimi için kullanılan istatistiksel yöntemlerin eklenerek anketlerin otonom olarak analizi sağlanabilir. Özellikle veriye ait ölçme düzeyleri ve veri türleri bu sayede otomatik olarak belirlenebilecek ve araştırmanın sonuçları daha doğru bir şekilde elde edilecektir.

6. KAYNAKLAR

1. Baldick, C., The Oxford Dictionary of Literary Terms, OUP Oxford, 2015.

2. Gamgam, H. and Altunkaynak, B., Parametrik Olmayan İstatistiksel Yöntemler, 2012.

3. Not, D. O. R., How to Select Appropriate Statistical Test?, Journal of Pharmaceutical Negative Results| October, 1,2 (2010) 61.

4. Massey Jr, F.J., The Kolmogorov-Smirnov Test for Goodness of Fit, Journal of the American Statistical Association, 46,253 (1951) 68–78.

5. Stephens, M.A., EDF Statistics for Goodness of Fit and Some Comparisons, Journal of the American Statistical Association, 69,347 (1974) 730–737.

6. Lilliefors, H.W., On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown, Journal of the American Statistical Association, 62,318 (1967) 399–402.

7. Shapiro, S.S. and Wilk, M.B., An Analysis of Variance Test for Normality (Complete Samples), Biometrika, 52,3–4 (1965) 591–611.

8. Glass, G. V, Testing Homogeneity of Variances, American Educational Research Journal, 3,3 (1966) 187–190.

9. Brown, M.B. and Forsythe, A.B., The Small Sample Behavior of Some Statistics Which Test the Equality of Several Means, Technometrics, 16,1 (1974) 129–132.

10. Güriş, S. and Astar, M., Bilimsel Araştırmalarda SPSS Ile Istatistik, 2014.

11. Wadsworth, G.P. and Bryan, J.G., Introduction to Probability and Random Variables, 7, McGraw-Hill New York:, 1960.

12. Lehmann, E.L. and D’Abrera, H.J.M., Nonparametrics: Statistical Methods Based on Ranks, Springer New York, 2006.

13. Wilcoxon, F. and Wilcox, R.A., Some Rapid Approximate Statistical Procedures, Lederle Laboratories, 1964.

14. Upton, G.J.G., Fisher’s Exact Test, Journal of the Royal Statistical Society. Series A (Statistics in Society) (1992) 395–402.

15. Birnbaum, Z.W., On a Use of the Mann-Whitney Statistic, 1956, Proceedings of the third Berkeley symposium on mathematical statistics and probability,13–17.

16. Gibbons, J.D. and Chakraborti, S., Nonparametric Statistical Inference, Springer, 2011.

17. Slud, E. and Wei, L.J., Two-Sample Repeated Significance Tests Based on the Modified Wilcoxon Statistic, Journal of the American Statistical Association, 77,380 (1982) 862–868.

18. Christensen, R., Plane Answers to Complex Questions, Springer New York, New York, NY, 2011.

19. Kazım, Ö., Paket Programlar İle İstatistiksel Veri Analizi, Kaan Kitabevi, Eskişehir, 2004.

20. Siegel, S., Nonparametric Statistics for the Behavioral Sciences., 1956.

21. Kruskal, W.H., A Nonparametric Test for the Several Sample Problem, The Annals of Mathematical Statistics (1952) 525–540.

22. Toutenburg, H., Lehmann, EL, Nonparametrics: Statistical Methods Based on Ranks, San Francisco. Holden-Day, Inc., 1975. 480 S., Zamm-Journal of Applied Mathematics and Mechanics/Zeitschrift Für Angewandte Mathematik Und Mechanik, 57,9 (1977) 562.

23. Ananda, M.M.A. and Weerahandi, S., Two-Way Anova with Unequal Cell Frequencies and Unequal Variances, Statistica Sinica (1997) 631–646.

24. Kum, S., Guideline For Suitable Statistical Test Selection, Plevra Bulteni, 8,2, 2014, 26–29.

25. Lang, T., Twenty Statistical Errors Even You Can Find in Biomedical Research Articles, 2004.

26. Akdeniz, F., Olasılık ve Istatistik, Nobel Kitapevi, Adana, 1995.

27. Sedgwick, P., Pearson's Correlation Coefficient, Bmj, 345,7 (2012)

28. Myers, L. and Sirois, M.J., Spearman Correlation Coefficients, Differences between, Wiley StatsRef: Statistics Reference Online, 2006.

29. Sen, P.K., Estimates of the Regression Coefficient Based on Kendall's Tau, Journal of the American Statistical Association, 63,324 (1968) 1379–1389.

30. Field, A. and Hole, G., How to Design and Report Experiments, Sage, 2002.

31. Leech, N.L., Barrett, K.C. and Morgan, G.A., IBM SPSS for Intermediate Statistics: Use and Interpretation, Routledge, 2014.

32. Mertler, C.A. and Reinhart, R.V., Advanced and Multivariate Statistical Methods: Practical Application and Interpretation, Routledge, 2016.

33. Code Team Work, Veri Analiz Platformu, Yayın No: 1.0., Trabzon, 2016.

34. Büyüköztürk, Ş., Sosyal Bilimler Için Veri Analizi El Kitabı, 2002.

35. Ibm Corp., Ibm Spss Statistics for Windows, Yayın No: 21.0. Armonk, NY, 2012.

36. The Popularity of Data Science Software, http://r4stats.com/articles/popularity/ May 15 2017.

ÖZ GEÇMİŞ

Musa SARİ, 10 Ağustos 1988 tarihinde Trabzon'da doğdu. Orta öğrenimini İstanbul Barbaros Lisesi'nde tamamladıktan sonra 2014 yılında Karadeniz Teknik Üniversitesi, Fen Fakültesi, İstatistik ve Bilgisayar Bilimleri Bölümünden mezun oldu. 2014 yılında Karadeniz Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik ve Bilgisayar Anabilim Dalı'nda tezli yüksek lisans programına başladı.