

# SPSS ile Nicel Veri Analizi Rehberi

Egemen Cevahir

Kibele Yayınları No: 116  
İnceleme – Araştırma Dizisi: 23

**SPSS  
ile  
Nicel Veri Analizi Rehberi**

**Egemen Cevahir**

Genel Yayın Yönetmeni: **Çağatay Apaydın**  
Bilimsel Editör: **R. Özgür Çatar**  
Kapak Tasarımı: **Çağatay Apaydın**  
Görsel Öğeler: **Uygar Alptekin**  
Sayfa Düzeni: **Uygar Alptekin**

**Bu e-kitap meta dışıdır, satılamaz!**

Bu kitabın yayın hakları, Kibele Yayınları'na aittir. İzin alınmaksızın, paylaşılabılır.

Birinci Baskı: Temmuz 2020  
ISBN: 978-605-9467-36-0

**KIBELE YAYINLARI**  
Sertifika No: 13847

Rıhtım Caddesi, Tayyareci Sami Sokak, Demirci İş Merkezi, No: 18, Kat: 4  
Kadıköy – İstanbul  
Tel/Faks: 0216 450 65 00  
[www.kibelegrup.com](http://www.kibelegrup.com) – [kibele@kibelegrup.com.tr](mailto:kibele@kibelegrup.com.tr)

# **SPSS ile Nicel Veri Analizi Rehberi**

**Egemen Cevahir**

**Editör: R. Özgür Çatar**

*“Emek ve Demokrasi Çalışmalarına İthaf Olunur.”*



## Önsöz

Her araştırmacının, sayısal verilerin analiz edilmesi sürecinde farklı kaynaklardan kimi zaman birbiriyle çelişen, muallakta kalan ya da tam olarak açıklanamayan ifade ve yöntemlerle karşı karşıya kalması sıkça rastlanan bir durumdur. Bu tür durumlarda çoğunlukla tercih edilen yöntem ise alanda duayen kabul edilen yazar ya da hocayı rehber kabul ederek çalışmaya devam etmek olmaktadır. Uzun yıllar birlikte çalıştığım değerli meslektaşım Egemen CEVAHİR'in, istatistik ve veri analizi ile ilgili akademik anlamda çalışma yapan tüm bilim insanlarına yardımcı olmasını umduğu, bu kitabı hazırlama fikrini duyup editörlüğünü de yapmamı teklif ettiğinde büyük heyecan ve mutluluk duyduğumu açık yüreklilikle söyleyebilirim. Bu sayede araştırma görevlisi olarak görev aldığımız yıllardan beri gerek kendi araştırmalarımız gerekse de bölümümüzden istatistiksel analiz konusunda yardım isteyen araştırmacılar için yaptığımız sayısız istatistik uygulamaları sonucu edindiğimiz bilgi ve deneyimleri, bu kitaba yansıtma şansını bulmuş bulunmaktayız. Konuyla ilgili daha önce yazılmış birçok kaynak olmasına karşın bu kitabı diğerlerinden ayıran nokta, kanımca, ilgili literatürde yol gösterici olan yazar ve kaynakların anlaşılır bir derlemesi olması, gerek istatistiki analizle yeni haşır neşir olan gerekse de uzun yıllardır bu çalışmaları yapan tüm araştırmacılar için tereddütte düştükleri noktalarda önemli bir rehber olacağı düşüncesidir. Diğer yandan yazarın bu kitapta aktardığı her bir bilgi akademik bakış açısında olması gerektiği gibi literatürdeki her bir görüşü açıklayarak hangi durumda nasıl hareket edilmesi gerektiğiyle ilgili farklı görüşleri de belirterek, uygulamada kararı araştırmacıya bırakmasıdır. İstatistiksel analiz yapan tüm araştırmacılar için; Türkçe, anlaşılır, sistematik ve erişilebilir bir bilgi kaynağı olan bu kitabın, ihtiyaç duyan herkese yardımcı olmasını temenni ederim.

Sevgi ve Saygılarımla...

R. Özgür ÇATAR

Temmuz 2020

Ataşehir

ozgurcatar@gmail.com

## Teşekkür

Bu kitabın ortaya çıkması bir sürecin sonucudur. Dolayısı ile bu süreçte yer alan sayısız insanın; yanlışlarından öğrendiğimiz “hocalarımızın”, doğrular konusunda önümüzü açan hocalarımızın, akademisyen dostlarımızın, öğrencilerimizin ve öğrencilerimizin, analiz deneyimlerimizin yani sayısız özne ve nesnelerin bu kitabın ortaya çıkmasında katkısı olduğunu öncelikle belirtmek isterim.

Kitap özelinde: eleştirel okuması, uyarı ve düzeltmeleri nedeni ile değerli hocamız Fikri ÖZTÜRK’e; dil yönünden okumaları ve düzeltmeleri nedeni ile değerli akademisyen dostlarım Miriş Meryem KURTULMUŞ ve İclal Ayşe KÜÇÜKKIRCA’ya; kitap kapak tasarımı ve dizaynı nedeni ile Çağatay APAYDIN’a; görsel öğeler ve düzenleme yardımları nedeni ile Uygur ALPTEKİN’e; yayınlanması hususunda ise Kibele Yayınları’na ve Eşber YAĞMURDERELİ’ye teşekkürlerimi sunarım...

Egemen CEVAHİR

Temmuz 2020

egemencevahir@gmail.com

## İçindekiler

<b>İSTATİSTİĞE GİRİŞ VE BAZI TEMEL KAVRAMLAR .....</b>	<b>1</b>
1.1. İstatistik .....	1
1.2. Değişken, Değişken Türleri ve Veri .....	2
1.2.1. Geleneksel Sınıflandırma: Nicel (Ölçümsel), Nitel (Kategorik) .....	2
1.2.2. Ölçüm Düzeylerine Göre Değişkenlerin Sınıflandırılması: .....	2
1.2.3. SPSS'in Değişken Sınıflandırması .....	5
1.2.4. Analiz Sürecinde Bağımlı-Bağımsız Değişken .....	5
1.3. Verilerin Özetlenmesi .....	6
1.3.1. Nitel Verilerin Özetlenmesi .....	6
1.3.2. Nicel Verilerin Özetlenmesi .....	6
1.3.2.1. Merkezi Eğilim Ölçüleri .....	6
1.3.2.2. Yayılım/Dağılım Ölçüleri .....	7
1.4. Normal Dağılım .....	12
1.4.1. Normal Dağılım Özellikleri .....	13
1.4.2. Normal Dağılım Testleri .....	14
1.5. Ölçeklerin Güvenilirliği .....	22
1.6. Hipotez Testleri: Parametrik ve Non-Parametrik Testler .....	25
1.6.1. Hipotez .....	25
1.6.2. Hipotez Testleri: Parametrik ve Non-Parametrik Testler .....	25
1.6.3. Hipotez Testleri ile Hipotezin Sınanması .....	26
1.7. Hata Tipleri ve Güç .....	27
<b>ORTALAMA KARŞILAŞTIRMALARI VE NON-PARAMETRİK TESTLERİ .....</b>	<b>29</b>
2.1. Tek Örneklem t Testi .....	29
2.2. Bağımsız Gruplarda t Testi .....	34
2.3. Mann Whitney U Testi .....	39
2.4. Bağımlı Gruplarda t Testi .....	43
2.5. Wilcoxon İşaretli Sıralar Testi .....	48
2.6. Bağımsız Gruplar için Tek Yönlü Varyans Analizi .....	53
2.7. Kruskal Wallis Testi .....	61
2.8. Tekrarlayan Ölçümler için Tek Yönlü Varyans Analizi .....	68
2.9. Friedman Testi .....	77
<b>ORAN KARŞILAŞTIRMALARI .....</b>	<b>83</b>
3.1. Ki Kare Testleri .....	83
3.1.1. Tek Örneklem Ki Kare Testi .....	83
3.1.2. Ki Kare Bağımsızlık Testi .....	90
3.2. McNemar Testi .....	99
3.3. Cochran Q Testi .....	104
<b>NİCEL DEĞİŞKENLER ARASI İLİŞKİ .....</b>	<b>111</b>
4.1. Basit Korelasyon .....	111
4.1.1. Pearson Korelasyon Testi .....	111
4.1.2. Spearman Korelasyon Testi .....	116
4.2. Kısmi Korelasyon .....	120
<b>DOĞRUSAL REGRESYON ANALİZİ .....</b>	<b>125</b>
5.1. Basit Doğrusal Regresyon Analizi .....	125
5.2. Çoklu Doğrusal Regresyon Analizi .....	132
<b>Kaynakça .....</b>	<b>153</b>
<b>Ekler .....</b>	<b>157</b>





# 1.BÖLÜM

## İSTATİSTİĞE GİRİŞ VE BAZI TEMEL KAVRAMLAR

### 1.1. İstatistik

Verilerin toplanması, sınıflandırılması, özetlenmesi, analiz edilmesi, yorumlanması ile ilgili teknik ve yöntemleri içeren bir disiplindir. İstatistik, sayısal olarak ifade edilen ya da sayıya dönüştürülebilen verilerin, nicel analiz süreçlerinin bütünüdür.

Verilerin derlenip toplanması, işlenip özetlenmesi, düzenlenip sunulması, çözümlenip sonuçlara varılması aşamalarında kullanılan kuramların, tekniklerin ve yöntemlerin hepsine birden istatistik denir. İstatistiğin gerçekle ilişkisi olmadığı, zaman zaman da bir tür yalan söyleme sanatı olduğu konusundaki kanılar kesinlikle haksızdır. İstatistiği, sayıların arkasında yatan gerçekleri anlamak, bunlardan işe yarar sonuçlar çıkarmakta kullanılan bir araç olarak düşünmemiz ve bu yönde kullanmamız gerekir. Her araç iyiye de kullanılabilir, kötüye de. Elektrikle ister aydınlatma yaparsınız, ister işkence! İstatistikle doğruları gösterebileceğiniz gibi insanları yanıltabilirsiniz de ve biraz istatistik öğrenirseniz başkalarının sizi yanıltması da pek kolay olmaz (Şenesen, 2013). “İşkence” metaforu istatistiğin kötü niyetli araçsallaştırılmasının ifadesi açısından gerçekçi ve çarpıcı bir benzetmedir. Bu benzetmenin üç temel ayağı vardır; birincisi örneklemin evreni temsiliyeti, ikincisi örneklemden elde edilen veri setine etik dışı müdahaleler, üçüncüsü ise verilerin doğru yöntemlerle analiz edilmemesidir. İstatistik süreci verinin elde edildiği örneklem aşamasından başlayıp, bu örneklemden elde edilen verilerle yapılan analizlerle devam eder. Veri setinin kaynağını oluşturan örneklemin temsil ettiği evrenden doğru bir yöntemle ve yeterli büyüklükte çekilmesi önemlidir. Örneklemin evreni temsiliyeti sağlanmadığında bu örneklemin verileri üzerinden yapılacak analiz sonuçlarının doğruluğu/gerçeği işaret etmesi ve genelleştirilebilmesi imkânı da olmayacaktır. Doğru bir yöntem ve büyüklükte seçilen örneklem verilerinin ise doğru yöntemlerle analiz edilmesi gerekmektedir. Bu kitap bu yanlışın birinci ve ikinci ayağı konusunda uyarıda bulunup, üçüncü ayağı konusunda rehberlik etme kapsamında yazılmıştır.

İstatistik disiplininin ürettiği bilgiler, nicel araştırma süreçlerinde kullanılarak, bir disiplin aynı zamanda diğer disiplinler için yönetime dönüşmüştür. Bu kitabın kapsamı da bu anlamda yöntem olarak istatistiktir. Bu kitap, nicel araştırma süreçlerinde veriden başlayıp, analizle sonuçlanan bu yolculukta, istatistiğin ürettiği doğru bilgileri ulaştırma, hangi durum ve koşullarda hangi testin yapılması gerektiği ve sonuçların yorumlanması hususlarında “açık” ve “anlaşılır” bir dille yöntem olarak istatistiğe odaklanmıştır.

## 1.2. Değişken, Değişken Türleri ve Veri

Bir durumdan diğerine, katılımcıdan katılımcıya, denekten deneğe farklılık gösteren özelliklere “değişken” adı verilir. Araştırma birimlerinin sabit olmayan her bir özelliği değişkendir. Yaş, boy, kilo, cinsiyet, eğitim durumu, aylık gelir, depresyon skoru, hava sıcaklığı, vb. gibi şeyler değişken örnekleridir. Değişkenlerin aldığı değerlere ise veri denir.

Değişkenler farklı şekillerde sınıflandırılabilir.

### 1.2.1. Geleneksel Sınıflandırma: Nicel (Ölçümsel), Nitel (Kategorik)

Eğer bir değişkenin özelliği sayı, miktar olarak ölçülebiliyorsa buna nicel (ölçümsel) değişken denir. Yaş, boy, kilo, IQ puanı, aylık gelir vb. gibi değişkenler nicel değişkenlerdir.

Eğer bir değişkenin özelliği sayı, miktarla ölçülemiyor ve bir duruma göre sınıflandırılıyorsa ya da kategorize ediliyorsa buna da nitel (kategorik) değişken denir. Cinsiyet, medeni durum, doğum yeri, eğitim durumu, ölçek ifadelerinin likert tipi cevap seçeneklerini (1 Tamamen katılmıyorum, 2 Katılmıyorum, 3 Kararsızım, 4 Katılıyorum, 5 Tamamen katılıyorum) içeren değişkenler nitel değişkenlerdir.

### 1.2.2. Ölçüm Düzeylerine Göre Değişkenlerin Sınıflandırılması:

Nitel ve nicel değişken şeklinde geleneksel olarak sınıflandırılan değişkenler aynı zamanda ölçme düzeylerine göre de -yaygın olarak- sınıflandırılır: Sınıflama Ölçeği, Sıralama Ölçeği, Aralık Ölçeği, Oran Ölçeği.

#### Sınıflama Ölçeği (Nominal Scale)

Üzerinde ölçüm yapılan varlıkları, belli özelliklere göre kategorilere ayıran, sınıflayan; tanımlama ve ayırt etme amacına yönelik bir ölçektir. Cinsiyet, medeni durum, doğum yeri gibi nitel (kategorik) değişkenleri ifade eder. Bu ölçekte, ölçülen özellikler miktarı ifade etmez ve bu nedenle de veriler arasında, sıralama ve derecelendirme işlemleri yapılamaz (Can, 2019). Sayısal olmayan nitel değişkenlerin kümelenmesinde kullanılan bu ölçek, Ad ölçeği, Gruplama ölçeği olarak da adlandırılmaktadır (Şenesen, 2013). Bu ölçeğin kategorileri birbirlerine bir üstünlük göstermez, sadece belli bir özelliği gösterdiğini ifade eder. Cinsiyet değişkeni kategorilerinin 1=Kadın, 2=Erkek olması, sadece bu kategorilerin –sayısal kodları- belli bir cinsiyeti ifade ettiği anlamını taşır. Sınıflama ölçeği Nitel (Kategorik) değişkenin bir alt kategorisidir.

Sınıflama ölçeği verileri sıklık (n) ve yüzde (%) cinsinden özetlenir. Sınıflama ölçeğinin sayısal değerleri (kodları) dört işleme tabi tutulamaz. Her kümedeki sıklık sayısı, yüzdesi başka kümelerle karşılaştırılabilir ya da onlarla oranlanabilir (Şenesen, 2013). Bu ölçekle elde edilmiş veriler non-parametrik testler ile sınanır.

### Sıralama Ölçeği (Ordinal Scale)

Nitel bir değişkenin gözlemlerinin dağıtıldığı kategoriler arasında genel kabul gören bir sıra varsa, kullanılan ölçeğe Sıralama ölçeği denilir (Şenesen, 2013). Sıralama ölçeği, değişkenlerin kategorilerini küçükten büyüğe doğru sıralayabilen ölçektir. Bu ölçeğin en klasik örneği eğitim durumu değişkenidir. Eğitim durumunun 0=okuryazar değil, 1=okuryazar, 2=ilkokul, 3=ortaokul, 4=lise, 5= üniversite olarak kodlandığı/kategorize edildiği bu değişken Sıralama ölçeğindedir. Genel kabul görmüş sıralama durumu ve büyük bir sayı ile kodlanmış kategorinin, küçük sayı olarak kodlanmış kategoriye göre bir büyüklüğü/üstünlüğü vardır. Bununla birlikte sayısal olarak ifade edilen kategorilerin arasındaki farklar nesnel olarak aynı büyüklüğü ifade etmez, yani aralıkları eşit değildir. Okuryazar (1) ile ilkokul (2) mezunu arasındaki fark ( $2-1=1$ ) ile lise mezunu (4) ile üniversite mezunu (5) arasında ki fark ( $5-4=1$ ) nesnel olarak aynı büyüklüğü ifade etmez. Bu ölçeğin diğer bir klasik örneği ise likert tipi ölçeklerin ifadelerine katılma durumlarını gösteren cevapları içeren değişkenlerdir (1=tamamen katılmıyorum, 2=katılmıyorum, 3=kararsızım, 4=katılıyorum, 5=tamamen katılıyorum). Sıralama ölçeğinin sıra sayıları (kodları) bir büyüklüğü ifade eder, ancak sıra sayıları arasındaki farklar nesnel bir büyüklüğü ifade etmezler. Sıralama ölçeği verileri sıklık (n) ve yüzde cinsinden özetlenir. Sıralama ölçeği Nitel (Kategorik) değişkenin bir alt kategorisidir.

Sıralama ölçeği verileri sıklık (n) ve yüzde (%) cinsinden özetlenir. Ayrıca büyüklük sırasına dizilmiş verilerin ortancası (medyan) verilerek de özetlenebilir (bu durumda yayılım ölçüm değerleri olarak 1. ve 3. çeyrek değerleri de verilebilir). Sıralama ölçeğinin sayısal değerleri (kodları) dört işleme tabi tutulamaz, ortalaması hesaplanamaz (Şenesen, 2013). Bu ölçekle elde edilmiş veriler, n, yüzde ya da sıra sayısı değerlerini temel alan; sıra ortalamaları, sıra toplamaları, sıra fark değerleri üzerinden non-parametrik testler ile sınanır.

### Aralık Ölçeği (Interval Scale)

Aralık Ölçeği (Eşit Aralık Ölçeği olarak da adlandırılır), verilerin ölçmeyi yapana ya da kullanılan ölçme aracına göre göreceli bir başlangıç noktası saptanarak ve eşit aralıklara bölünerek sayısallaştırıldığı ölçek tipidir. Bu ölçekteki değerler, varlığın miktarını gösteren gerçek değerlerdir ve birbirini takip eden ölçek değerleri arasındaki farklar eşittir. Ancak, aralık ölçeğinde göreceli olarak alınan ve sıfır olarak tanımlanan başlangıç noktası, ölçülen nesnenin hiç olmadığını göstermez. Deniz yüzeyinin sıfır olarak alındığı yükseklik ölçümünde, deniz yüzeyinin bir yüksekliği olmadığı anlamına gelmez. Gece yarısının saat olarak 00.00 (sıfır) alındığı bir zaman ölçümünde o anda zamanın olmadığı anlamına gelmez. Suyun donmasının başlangıç (sıfır) alındığı sıcaklık ölçümlerinde, 0 °C derece sıcaklığın olmadığı anlamına gelmez. Sıcaklık, IQ puanı, Sınav başarı puanı, Depresyon puanı (ölçeklerle ulaşılan puanlar) vb. ölçekler Aralık ölçeğidir (Can, 2019). Aralık ölçeği Nicel (Ölçümsel) değişkenin bir alt kategorisidir. Aralık ölçeği verilerinin aritmetik ortalaması ve standart sapması hesaplanabilir ve normal dağılım özelliği göstermesi durumunda parametrik testler ile sınanır.

### Oran Ölçeği (Ratio Scale)

Aralık ölçeğinin bütün özelliklerini taşıyan ancak; başlangıç noktası olarak tanımlanan 0 (sıfır) noktasının ölçülen özelliğin gerçekten hiç olmadığını gösteren ölçektir (Can, 2019). Çok yaygın olarak kullanılan Kg (Kilogram), m (metre), TL (Türk Lirası), kardeş sayısı gibi ölçümler bu ölçeğin örnekleridir (Şenesen, 2013). Oran ölçeği Nicel (Ölçümsel) değişkenin bir alt

kategorisidir. Oran ölçeği verilerinin aritmetik ortalaması ve standart sapması hesaplanabilir ve normal dağılım özelliği göstermesi durumunda parametrik testler ile sınanır.

#### Şapkadan tavşan çıkarmak!

Sıralama ölçeği değişkeninin değerleri (kategori kodları) bir büyüklüğü ifade etmekle birlikte, bu büyüklük başka bir kategoriye göre genel kabul görmüş bir durumun ifadesi olan bir büyüklüktür, “nesnel” bir büyüklüğün ifadesi değildir. Eğitim değişkeninin kategori kodları olan 1 okuryazar değil, 2 okuryazar, 3 ilkokul, 4 ortaokul, 5 lise, 6 üniversite de olduğu gibi kategori kodları bir büyüklüğü ifade eder. Ancak kategorilerin büyüklük değerleri arasındaki farklar “nesnel” olarak eşit bir büyüklüğün ölçümü değildir:

2 (okuryazar) -1 (okuryazar değil)=1 ile 6 (üniversite) – 5 (lise)=1 eşit değildir.

Kategori kodların büyüklüğü aynı birimin “nesnel” ölçüsü değildir:

5 (lise)  $\neq$  1 (okuryazar değil) x 5

2 (okuryazar)  $\neq$  1 (okuryazar değil) + 1 (okuryazar değil)

Sıralama ölçeğinin bu özellikleri nedeni ile sıralama ölçeğinin sayısal değerleri (kodları) dört işleme tabi tutulamaz, ortalaması hesaplanamaz (Şenesen, 2013).

Sıralama ölçeğindeki tüm değişkenler için bu durumlar geçerlidir. Sosyal bilimlerde yaygınlıkla kullanılan, objektif bir ölçüm aracı ile ölçülemeyen durumlara yönelik -geçerlilik ve güvenilirlik analizleri ile- geliştirilen ölçekler (Depresyon Ölçeği, Kaygı Ölçeği, Özgüven Ölçeği, vb.) ile yapay/suni olarak bazı özellikler ölçülür. Bu ölçekler ölçmek istediği özelliğe yönelik ifadeleri ve bu ifadelerle katılım durumlarını gösteren likert tipi (1=tamamen katılmıyorum, 2=katılmıyorum, 3=kararsızım, 4=katılıyorum, 5=tamamen katılıyorum) cevapları içerir. Bu ölçeklerin her bir ifadesi sıralama ölçeğindedir. Dolayısı ile **ölçeğin her bir ifadesi dört işleme ilkesel olarak tabi tutulamaz**. Her bir ifadesi sıralama ölçeğinde ölçülmüş değişkenler iken **ölçeğin ifadeleri dört işleme tabi tutularak**, ölçeğe ilişkin toplam bir puan oluşturulur, elde edilen ölçeğin toplam puanı ise aralık ölçeğinde bir değişken olarak kabul edilir! Ölçeklerin geçerlilik güvenilirlik analizlerinin yapılması sıralama ölçeğindeki bir değişkeni aralık ölçeğindeki bir değişkene dönüştürülmesini sağlamaz.

Bu durum tam olarak şuna benzetilebilir, sihirbazlığa! Şapkanın içine kurdeleler soktukten sonra hokus pokus işlemi ile sihirbazın şapkadan tavşan çıkarması gibi bir durumdur. Ölçeğin her bir ifadesi sıralama ölçeğindeki bir değişken iken –dört işleme ilkesel olarak tabi tutulamazken- ölçeğin ifadelerini şapkaya koyup, ölçeğin ifadeleri sıralama ölçeğinde iken hokus pokus işlemi ile şapkadan aralık ölçeğindeki ölçeğin toplam puanının çıkarılması işlemidir. Ölçeklerin ifadeleri sıralama ölçeğinde iken toplam puanının aralık ölçeğine dönüştürülmesi ve ardından normal dağılım özelliği göstermesi durumunda, parametrik testlerin nesnesi olması, literatürde eleştirisi olmasına rağmen genel kabul görmüş **de facto** bir durumdur.

**Önemli not:** Testlerin varsayımlarında ölçüm düzeylerine göre değişkenlerin sınıflandırılması esas alınmıştır. Buna göre özellikle şu ifadelerle dikkat edilmeli:

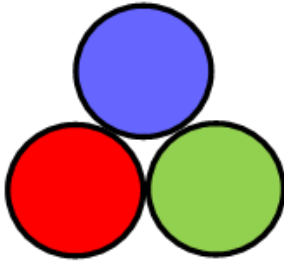
“En az sıralama ölçeğinde” ifadesi: Sırama Ölçeği, Aralık Ölçeği ve Oran Ölçeğini kapsar.

“En az aralık ölçeğinde” ifadesi: Aralık Ölçeği ve Oran Ölçeğini kapsar.

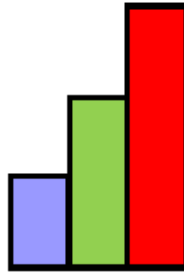
### 1.2.3. SPSS'in Değişken Sınıflandırması

SPSS değişkenleri 3 kategoride sınıflandırır. Nominal (Sınıflama), Ordinal (Sıralama) ve Scale (Ölçümsel). Nominal (Sınıflama) ve Ordinal (Sıralama) değişkenler ölçüm düzeylerine göre değişkenlerin sınıflandırılmasında aynı isimle tanımlanan değişkenlerdir. Scale (Ölçümsel) değişkeni ise ölçüm düzeyleri itibari ile Aralık (Interval) ve Oran (Ratio) ölçeğindeki değişkenleri birlikte kapsayan bir kategoridir.

**Nominal**



**Ordinal**



**Scale**



### 1.2.4. Analiz Sürecinde Bağımlı-Bağımsız Değişken

Değişkenlerin birbiri ile ilişkisinin ya da farklılığının sınındığı analiz aşamasında değişkenler; bağımlı ve bağımsız şeklinde sınıflandırılabilir. Analiz aşamasında değişkenler neden-sonuç (etkileyen-etkilenen) ilişkisinde kurgulanarak, bağımsız ve bağımlı değişken olarak sınıflandırılır ve bu kurgu içinde analiz yapılır. Bağımsız değişken, bağımlı değişken üzerindeki etkisi sınanmak istenen değişkendir.

**BAĞIMSIZ**



**BAĞIMLI**

### 1.3. Verilerin Özetlenmesi

İstatistiğin analiz süreci, verilerin özetlenmesiyle başlar. Belli bir örneklemden, anket ya da başka bir veri toplama aracı ile ilgilenilen değişkenlere ait katılımcıların değerleri toplanarak veriler oluşturulur. Ardından topladığımız veriler özetlenir. Verilerin özetlenmesi bir zorunluluk olduğu kadar, örnekleme değişkenlerinin aldığı değerlerin tanımlanması sürecidir. Verileri özetlerken değişkenin nitel ya da nicel olması (nicel değişkenlerin de normal dağılıma uyup uymadığına göre) durumuna göre farklı kriterlerle özetlenir.

#### 1.3.1. Nitel Verilerin Özetlenmesi

Nitel veriler, örneklem içindeki katılımcıların bu değişkene ilişkin gözlenen değerlerinin sayısı (**n, görölme sıklığı**) ve bu görölme sayısının örneklem büyüklüğüne (sayısına) oranı ile ulaşılan yüzde (%) değerleri ile özetlenir. Sınıflama ve sıralama ölçeği verileri, sıklık (n) ve (yüzde) cinsinden özetlenir. Ayrıca Sıralama ölçeği verileri, ortancası (%50. Değer) ve 1. Çeyrek (%25. Değer) ve 3. Çeyrek (%75. Değer) ile de özetlenebilir.

#### 1.3.2. Nicel Verilerin Özetlenmesi

Nicel veriler merkezi eğilim ve yayılım ölçüm değerleri ile birlikte özetlenir. Nicel veriler normal dağılım özelliği gösterip göstermemesi durumuna göre özetleneceği merkezi eğilim ve yayılım değerleri değişir. Nicel veriler özetlenirken merkezi eğilim ölçüleri ile birlikte yayılım/dağılım ölçülerini birlikte vermek bir zorunluluktur.

Normal dağılım özelliği gösteriyorsa: Ortalama, Standart sapma, Minimum ve Maximum değerleri ile özetlenir.

Normal dağılım özelliği göstermiyorsa: Ortanca (%50.değer), 1. Çeyrek (%25. Değer), 3. Çeyrek (%75. Değer), Minimum ve Maximum değerleri ile özetlenir.

##### 1.3.2.1.Merkezi Eğilim Ölçüleri

**Aritmetik Ortalama:** Gözlenen değerlerin tümü toplanarak gözlem sayısına (n) bölüldüğünde elde edilen değere aritmetik ortalama denir. Örneklemden her bir katılımcının söz konusu nicel değişkeninin değerleri toplanarak, örneklem büyüklüğü sayısına (n) bölünmesi ile aritmetik ortalama değeri hesaplanır. Nicel bir değişken normal dağılım özelliği gösteriyor ise merkezi eğilim ölçümü olarak aritmetik ortalaması verilir.

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \dots + X_n}{n}$$

**Ortanca (Medyan):** Gözlenen ölçüm değerleri küçükten büyüğe doğru sıralandığında ortada kalan değerdir (% 50. Değer). Örneklemden her bir katılımcının söz konusu nicel değişkeninin değerleri, küçükten büyüğe doğru sıralanır, bu sıralama sonrasında ortada kalan değer, ortancadır. Ortada tek bir değer değil de 2 değer kalırsa, bu iki değer toplanıp ikiye bölünmesi

sonucunda ulaşılan değer, ortancadır. Nicel bir değişken normal dağılım özelliği göstermiyor ise merkezi eğilim ölçümü olarak ortancası (medyanı) verilir.

Örnek: 9 kişilik bir grubun matematik ders notları 100 üzerinden şöyle olsun:

Matematik notu:	50	90	55	62	45	83	71	95	75
Matematik notlarını küçükten büyüğe doğru sıralıyoruz									
	45	50	55	62	71	75	83	90	95

Bu sıralama sonucunda ortada kalan, ortanca (%50.) değeridir. Ortanca=71

**Mod (Tepe Değer):** Gözlenen ölçüm değerleri arasında en çok tekrar eden değerdir. Eşit sayıda en çok tekrar eden değerlerin olma ihtimali de vardır, bu durumda birden çok tepe değer gerçekleşmiş olur.

Örnek: 10 kişilik bir grubun matematik ders notları 5 üzerinden şöyle olsun:

Matematik notu:	1	2	2	3	3	3	3	4	5	5
-----------------	---	---	---	---	---	---	---	---	---	---

Matematik notu	Tekrarlama sayısı
1	1
2	2
3	4
4	1
5	2

Bu grubun matematik notlarının tepe değeri: 3 notudur.

### 1.3.2.2. Yayılım/Dağılım Ölçüleri

Yayılım ölçüleri, merkezi eğilim ölçülerinin grup içinde nasıl dağıldığına ilişkin bilgi veren değerlerdir. Nicel bir değişkenin sadece merkezi eğilim ölçüsü ile özetlenmesi eksik ve yetersiz bir bilgi sunar, söz konusu değişkenin merkezi ölçümü ile yayılım ölçümlerinin birlikte sunulması teknik açıdan da bir zorunluluktur. Normal dağılıma uyan bir nicel değişkenin ortalama ve standart sapma değerleri ile birlikte verilmesi; normal dağılıma uymayan bir nicel değişkenin ise ortanca (medyan), 1. Çeyrek (%25. Değer) ve 3. Çeyrek (%75. Değer) değerleri ile birlikte verilmesi gerekir.

**Standart sapma:** Her bir gözlenen değer aritmetik ortalamaya olan farkının karelerinin toplamının, n-1'e bölümünün, kareköküdür. Standart sapma gözlenen değerlerin ortalamadan sapmasını gösteren bir yayılım ölçüsü değeridir.

$$s = \sqrt{\frac{(\bar{X} - X_1)^2 + (\bar{X} - X_2)^2 + (\bar{X} - X_3)^2 + \dots + (\bar{X} - X_n)^2}{n-1}}$$

**Varyans:** Varyans standart sapmanın karesidir. Her bir gözlenen değer aritmetik ortalamaya olan farkının karelerinin toplamının, n-1'e bölümüdür.

$$\text{Varyans} = s^2$$

**Standart hata:** Standart sapmanın, örneklem büyüklüğünün (n) kareköküne bölümü ile elde edilen bir değerdir.

$$\text{Standart hata} = \frac{s}{\sqrt{n}}$$

Standart sapma, Varyans, Standart hata, ortalamanın grup içinde nasıl yayıldığına ilişkin yayılım/dağılım ölçüm değerleridir. Nicel bir değişken normal dağılım özelliği gösteriyor ise ortalama ile birlikte çoğunlukla verilen yayılım değeri standart sapmadır.

### Çeyrek Değerler:

**1. Çeyrek (%25. Değer):** Gözlenen ölçüm değerleri küçükten büyüğe doğru sıralandığında %25. değerdir. Örneklemdeki her bir katılımcının söz konusu nicel değişkeninin değerleri, küçükten büyüğe doğru sıralanır, bu sıralama sonrasında, sıralamadaki %25. değer, 1. Çeyrek değerdir. Bu sıralamanın sonucunda %25. değer tek bir değer değil de 2 değer arasında paylaşılıyor ise, bu iki değerın toplanıp ikiye bölünmesi sonucunda ulaşılan değer, 1. Çeyrek değerdir.

**3. Çeyrek (%75. Değer):** Gözlenen ölçüm değerleri küçükten büyüğe doğru sıralandığında %75. değerdir. Örneklemdeki her bir katılımcının söz konusu nicel değişkeninin değerleri, küçükten büyüğe doğru sıralanır, bu sıralama sonrasında, sıralamadaki %75. değer, 3. Çeyrek değerdir. Bu sıralamanın sonucunda %75. değer tek bir değer değil de 2 değer arasında paylaşılıyor ise, bu iki değerın toplanıp ikiye bölünmesi sonucunda ulaşılan değer, 3. Çeyrek değerdir.

Örnek: 9 kişilik bir grubun depresyon puanları 100 üzerinden şöyle olsun:

Depresyon puanı:	50	45	42	55	62	75	83	71	75
Depresyon puanlarını küçükten büyüğe doğru sıralıyoruz									
Depresyon puanı sıralı:	42	45	50	55	62	71	75	75	83
		45	50				75	75	
		%25. Değer =(45+50)/2					%75. Değer =(75+75)/2		
		47,5					75		

Bu grubun depresyon 1. Çeyrek (%25. değeri) puanı= 47.5, 3. Çeyrek (%75. değeri) puanı= 75  
Bu grubun depresyon ortanca (%50. Değeri) puanı= 62

1.Çeyrek (%25. değer) ve 3. Çeyrek (%75. değer) değerleri, ortancanın (medyanın) grup içinde nasıl yayıldığına ilişkin yayılım/dağılım ölçüm değerleridir. Nicel bir değişken normal dağılım özelliği göstermiyor ise ortanca ile birlikte yayılım ölçümleri 1.Çeyrek (%25. değer) ve 3. Çeyrek (%75. değer) değerleri verilir.

**Minimum değer:** Gözlenen ölçüm değerleri arasındaki en küçük değerdir.

**Maximum değer:** Gözlenen ölçüm değerleri arasındaki en büyük değerdir.

**Ranj (Değişim Aralığı/Genişliği):** Gözlenen ölçüm değerleri arasındaki en büyük ve en küçük değer arasındaki farkı ifade eder.



## Değişkenlerin özetlenmesi SPSS Uygulamaları

### Nitel değişkenler için SPSS Uygulaması:

- 1- **Analyze, Descriptive Statistic, Frequencies** seçilir.
- 2- **Frequencies** penceresinde sol taraftaki değişken listesinden **özetlenecek nitel değişkenler** seçilip sağ taraftaki **Variables** penceresine aktarılır. **Frequencies** penceresinde **OK** Tıkılır.

### SPSS çıktısı:

#### Frequencies

		Statistics	
		cinsiyet	medeni durum
N	Valid	135	135
	Missing	0	0

**N Valid:** Geçerli n sayısı (geçerli cevap sayısı)

**N Missing:** Kayıp n sayısı (kayıp veri sayısı)

		Cinsiyet			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	kadın	72	53,3	53,3	53,3
	erkek	63	46,7	46,7	100,0
	Total	135	100,0	100,0	

		gelir seviyeleri			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	düşük	32	23,7	23,7	23,7
	orta	78	57,8	57,8	81,5
	yüksek	25	18,5	18,5	100,0
	Total	135	100,0	100,0	

**Frequency:** n, sıklık (söz konusu değer görülen sayısı)

**Percent:** %, yüzde (söz konusu değer görülen sayısının toplam örneklem içindeki oranı)

**Valid Percent:** Geçerli yüzde (kayıp veriler hariç yüzde)

**Cumulative Percent:** Toplam yüzde

**Tablo: Araştırma Grubunda Cinsiyet ve Gelir Seviyelerinin Dağılımı**

Değişken		n	%
Cinsiyet	Kadın	72	53,3
	Erkek	63	46,7
	Toplam	135	100,0
Gelir seviyesi	Düşük	32	23,7
	Orta	78	57,8
	Yüksek	25	18,5
	Toplam	135	100,0

Araştırma grubundaki katılımcıların %53,3'ü kadın ve %46,7'si erkektir. Katılımcıların %23,7'sinin gelir seviyesi düşük, %57,8'inin orta ve %18,5'inin ise gelir seviyesinin yüksek olduğu saptanmıştır.

#### Nicel değişkenler için SPSS Uygulaması:

- 1- **Analyze, Descriptive Statistic, Frequencies** seçilir.
- 2- **Frequencies** penceresinde sol taraftaki değişken listesinden özetlenecek **nicel değişkenler** seçilip sağ taraftaki **variables** penceresine **aktarılır**.
- 3- **Frequencies** penceresinde **Statistics** butonu **tıklanır** ve **Mean, Std. Deviation, (Normal dağılıma uyuyor ise). Median, Quartiles, (Normal dağılıma uymuyor ise). Minimum ve Maximum** seçilir. **Continue** tıklanır.
- 4- **Frequencies** penceresinde **OK** Tıklanır.

#### SPSS çıktısı:

#### Frequencies

		Statistics	
		aylık gelir	Yaş
N	Valid	135	135
	Missing	0	0
Mean		2982,9630	30,5556
Median		2500,0000	28,0000
Std. Deviation		1863,59630	9,28059
Minimum		700,00	17,00
Maximum		8000,00	50,00
Percentiles	25	1600,0000	23,0000
	50	2500,0000	28,0000
	75	4000,0000	37,0000

**Mean:** Ortalama. **Median:** Ortanca. **Std. Deviation:** Standart Sapma.

**Minimum:** En küçük değer. **Maximum:** En büyük değer.

**Percentiles 25:** %25. değer (1. Çeyrek). **Percentiles 50:** %50. değer yani Ortanca.

**Percentiles 75:** %75. değer (3. Çeyrek)

**Tablo: Araştırma Grubunda Aylık Gelir ve Yaş Dağılımı**

Değişken	n	Ortalama	Ortanca	Standart sapma	1.Çeyrek	3.Çeyrek	En küçük değer	En büyük değer
Aylık gelir (TL)	135	2.982,96	2.500,00	1.863,60	1.600,00	4.000,00	700,00	8.000,00
Yaş (Yıl)	135	30,56	28,00	9,28	23,00	37,00	17,00	50,00

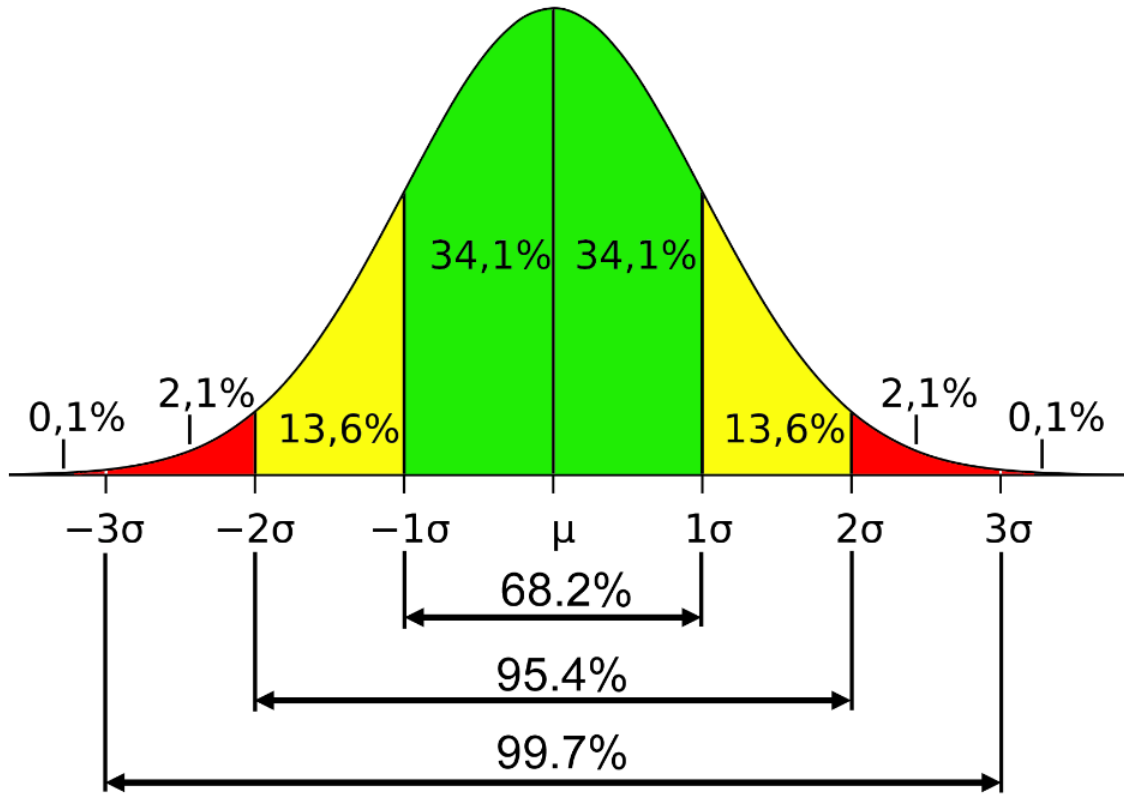
**Önemli not:** Nicel bir değişken bu şekilde birden çok merkezi eğilim ve yayılım ölçüm değerleri ile özetlenmez, buna gerek yoktur. Bu tablo SPSS çıktısının nasıl tablolandırıldığına dair bir örnektir. Nicel bir değişken normal dağılım özelliği gösteriyor ise ortalama ile birlikte standart sapma, minimum ve maximum değerleri ile özetlenmesi yeterlidir. Nicel bir değişken normal dağılım özelliği göstermiyor ise ortanca ile birlikte 1.Çeyrek (%25. değer), 3. Çeyrek (%75. değer), minimum ve maximum değerleri ile özetlenmesi yeterlidir.

Araştırma grubundaki katılımcıların aylık gelir ortalaması 2.982,96 TL±1.863,60 TL, ortancası 2.500 TL (1. Çeyrek: 1.600 TL; 3. Çeyrek: 4.000 TL) olduğu, katılımcılar arasında en az aylık gelirin 700 TL ve en çok aylık gelirin ise 8.000 TL olduğu görülmektedir.

Araştırma grubundaki katılımcıların yaş ortalamasının 30,56±9,28 yıl olduğu, yaş ortancasının 28 yıl (1. Çeyrek: 23; 3. Çeyrek: 37) olduğu, katılımcılar arasında en küçük yaştan 17 ve en büyük yaştan ise 50 olduğu görülmektedir.

#### 1.4. Normal Dağılım

Normal dağılım doğada gerçekleşen pek çok olayla uyum içindedir ve pek çok rastlantısal sürecin dağılımı olarak karşımıza çıkar (Can, 2019). Normal Dağılım, gerçek dünyadaki rastgelelik olgusunun modellenmesi sırasında en çok kullanılan dağılımdır. Ayrıca istatistik teorisinde de teorik açıdan çok önemli ve ilginç özelliklere sahiptir. Bunlardan en önemlisi Merkezi Limit Teoremidir (Öztürk ve Özbek, 2004). Normal dağılımı ilk tanımlayan Gauss'tur. Doğa ya da toplumdan rastgele seçilen belli bir sayıdaki örnekleme ait nicel bir değişkenin aldığı değerlerinin dağılımının çan eğrisi şeklinde olduğunu ve belli özellikleri taşıdığını belirtmiştir.



Kaynak: Yalçın, 2016

Evrendeki herhangi bir nicel değişkenin değerlerinin sıklık dağılımı yukarıdaki grafikte ki gibi olduğu kabul edilir ve bu dağılım Gauss tarafından “normal dağılım” olarak adlandırılmıştır. Yatay ekseninde nicel değişkenin aldığı değerler yer almaktadır (bu grafikte bu değerler yoktur, bunun yerine  $\mu$ : evrenin ortalaması ve  $\sigma$ : evrenin standart sapmasına yer verilmiştir), çan eğrisi şeklindeki çizgi ise her bir değer yoğunluğunu (olasılık yoğunluk fonksiyonu değerini) gösteren noktaların birleştirilmesi ile oluşan çizgidir. Örneklemin nicel değişken verilerinin normal dağılım göstermesi beklenir.

### 1.4.1. Normal Dağılım Özellikleri

Teorik olarak normal dağılımın ortalaması, ortancası (medyan) ve tepe değeri (mod) aynıdır. Değerlerin histogram dağılımı çan eğrisi şeklindedir ve dağılım ortalamaya göre simetrik.

- ❖ Değerlerin %68,26'sı  $\pm 1$  standart sapma içinde yer alır.
- ❖ Değerlerin %95,44'ü  $\pm 2$  standart sapma içinde yer alır.
- ❖ Değerlerin %99,74'si  $\pm 3$  standart sapma içinde yer alır.

Normal dağılım özelliği gösteren nicel değişken verilerin dağılımı çan eğrisi şeklinde olur ve çan eğrisinin tepe noktası değeri (mod) aynı zamanda, ortalamanın ve ortancanın değerini ifade eder. Verilerin dağılımı çan eğrisinin tepe değeri çevresinde yoğunlaşır ve simetrik bir şekilde yayılım gösterir. Örneklem grubunda değerleri ortalama ile ortalamadan bir standart sapma küçük değerleri olan katılımcılar toplam örneklem %34'ünü, değerleri ortalama ile ortalamadan bir standart sapma büyük değerleri olan katılımcılar toplam örneklem %34'ünü oluşturur. Ortalamanın bir standart sapma altında ve üstünde değerleri olan katılımcıların örneklem içindeki payı %68'dir. Ortalamanın iki standart sapma altında ve üstünde değerleri olan katılımcıların örneklem içindeki payı %95'tir. Ortalamanın üç standart sapma altında ve üstünde değerleri olan katılımcıların örneklem içindeki payı ise %99'dur.

Normal dağılıma ilişkin genel kabullerden biri de normal dağılımın sağlanabilmesi için en az 30 örneklem büyüklüğüne ( $n \geq 30$ ) ihtiyaç duyulduğu hususundadır. Bu kabul Merkezi Limit Teoremine dayanır. Bu teoremin iddiası evrenden rastgele seçilme şartı ile normal dağılım için örneklem büyüklüğü  $n \geq 30$  koşulunda sağlanmaya başlandığı yönündedir (Jolliffe, 1995). Bu aynı zamanda örneklem büyüklüğünün 30'dan küçük ( $n < 30$ ) olduğu durumlarda normal dağılımın sağlanamayacağı iddiasını da taşımaktadır. Bu görüş 20 yıl öncesine dek genel kabul görmüş ise de son çalışmaların kanıtları göstermiştir ki örneklem büyüklüğünün 30'dan küçük ( $n < 30$ ) olduğu durumlarda normal dağılım sağlanabildiği gibi örneklem büyüklüğünün 30'dan büyük olduğu durumlarda normal dağılımın sağlanamadığı örnekler de mevcuttur (Chang ve Diğr., 2008; Warner, 2008).

### Nicel değişkenlerin normal dağılıma uygunluğu nasıl değerlendirilir?

Herhangi bir nicel değişkenin normal dağılım sergileyip sergilemediği, normal dağılım özelliklerini temel alan kriterlere göre değerlendirilir. Bu kriterler:

- ✓ Ortalama, ortanca, mod değerlerinin birbirine yakın olması,
- ✓ Değerlerin histogram dağılımı çan eğrisi şeklinde olması,
- ✓ Çarpıklık (Skewness) ve Basıklık (Kurtosis) katsayılarının -1+1 aralığında olması,
- ✓ Normal Q-Q Plot grafiğinde noktaların 45 derecelik doğru üstünde (ya da çok yakınında) olması gerekir.

Çarpıklık (Skewness) katsayısı, histogram dağılımındaki çan eğrisinin sola ya da sağa çarpıklığını ifade eden bir katsayıdır. Çarpıklık katsayısının işareti negatif (-) ise dağılım sağa; pozitif (+) ise dağılım sola çarpıktır. Çarpıklık katsayısının negatif (-) olması katılımcı verilerinin ortalamasının sağında (büyük değerlerde) yoğunlaştığını; pozitif (+) olması ise ortalamasının solunda (küçük değerlerde) yoğunlaştığını gösterir. Basıklık (Kurtosis) katsayısı ise historgam dağılımındaki çan eğrisinin sivrililiğini ya da basıklığını ifade eden bir katsayıdır. Basıklık katsayısının işareti negatif (-) ise dağılım basık; pozitif (+) ise dağılım sivridir. Basıklık katsayısının negatif (-) olması verilerin ortalamasının etrafında yayılımına işaret eder (standart

sapmanın büyüklüğünün bir işaretidir), pozitif (+) olması verilerin ortalamasının yakınında yoğunlaştığını (standart sapmanın küçüklüğünü işaret eder), verilerin merkezde toplandığını gösterir. Çarpıklık (Skewness) ve Basıklık (Kurtosis) katsayılarının sıfır (0) olması ideal bir normal dağılım olduğunu gösterir. Bu katsayıların büyümesi normallikten uzaklaşmaya işaret eder (Pallant, 2017; Field, 2009). Normal dağılım için Çarpıklık (Skewness) ve Basıklık (Kurtosis) katsayılarının hangi aralıklarda olması konusunda literatürde üzerinde anlaşılmış bir referans aralığı yoktur (Mangiafico, 2016). Bununla birlikte Çarpıklık (Skewness) ve Basıklık (Kurtosis) katsayılarının -1+1 aralığında olması, normal dağılımla ilgili kabul edilebilir sınır değerler olarak kullanılabilir (Mangiafico, 2016; Büyüköztürk, 2010).

Normal Q-Q Plot grafiği beklenen/kuramsal normal dağılım değerleri ile gözlenen/gerçekleşen değerler arasındaki ilişkiyi gösterir. Beklenen ve gözlenen değerler arasındaki ilişki örtüştüğünde, ortaya 45 derecelik açı yapan bir doğru çıkar. Ayrıca gözlenen değerler bu grafikte noktalarla gösterilir ve noktaların 45 derecelik doğru üstünde ya da doğrunun çok yakınında olması beklenir (Can, 2019).

#### 1.4.2. Normal Dağılım Testleri

Nicel bir değişkenin normal dağılıma uygunluğu testler aracılığı ile de sınanabilir. Normal dağılıma uygunluğun sınanmasında en yaygın kullanılan iki test Kolmogorov Smirnov ve Shapiro Wilk testleridir. Örneklem büyüklüğü 30'dan küçük ( $n < 30$ ) olan durumlarda Shapiro Wilk testi ile örneklem büyüklüğü 30 ve üstü ( $n \geq 30$ ) olan durumlarda ise Kolmogorov Smirnov testi ile nicel değişkenin normal dağılıma uygunluğu sınanabilir.

- $n \geq 30$  ise Kolmogorov Smirnov testi ile,
- $n < 30$  ise Shapiro Wilk testi ile sınanabilir.

Test sonucu  $p > 0,05$  ise değişken normal dağılım özelliği gösteriyor,

$p < 0,05$  ise değişken normal dağılım özelliği göstermiyor anlamına gelir.

Nicel değişkenin normal dağılımı istatistiksel olarak test edildiğinde testin p değerinin ötesinde, normal dağılım kriterlerine göre nicel değişkenin normal dağılıma uyup uymadığına karar verebiliriz. Testin p değeri, nicel değişkenin normal dağılıma uyup uymadığına karar vermekte tek başına bizim için bağlayıcı değildir, normal dağılıma uygunluk sadece teste göre değerlendirilmez, öncelikli ya da önemli olan normal dağılım kriterleridir ve bu kriterlere göre karar vermek daha sağlıklıdır.

Literatürde Shapiro Wilk testinin normallliği sınamada daha güçlü bir test olduğu yönünde görüşler de mevcuttur. Bu görüşte olan yazarlar Shapiro Wilk testinin sadece küçük örneklemelerde değil ( $n < 30$ ) büyük örneklemelerde ( $n \geq 30$ ) de normallüğün sınanmasında kullanılmasını önermektedirler (Razali ve Wah 2011). Bunun aksi yönünde olan görüş ise küçük örneklemelerde ( $n < 30$ ) Shapiro Wilk testinin normal dağılmayan durumları tespit etmesi yönünde daha güçlü sonuçlar ürettiği ve büyük örneklemelerde ( $n \geq 30$ ) ise Tip I hatadan kaçınarak karar vermesi yönünde en iyi sonuçlar üretmesi nedeni ile Kolmogorov Smirnov testinin yapılmasını önermektedir (Öztuna ve Diğr., 2006). Field (2009) da Shapiro Wilk testinin normallikten farklılıkları tespit etmede Kolmogorov Smirnov testinden daha güçlü bir test olduğunu, bununla birlikte büyük örneklemelerde güçlü sonuçlar üretmede hataya

düşebileceğini belirtmiştir. Corder ve Foreman (2009) da normal dağılımın sınanmasında Kolmogorov Smirnov testini önermiştir. Normal dağılım testlerinin yazar tarafından önerilen tarzı, literatürde hâkim olarak kabul gören ve yaygın olarak kullanılan biçimdir.

### Normal dağılım neden önemlidir?

Nicel veriler özetlenme ve analiz süreçlerinden önce normal dağılıma uygunluğu öncelikli olarak sınanmalıdır. Nicel veriler normal dağılım özelliği gösterip göstermemesi durumuna göre özetleneceği merkezi eğilim ve yayılım ölçümleri değişir. Nicel veriler özetlenirken merkezi eğilim ölçümleri ile birlikte yayılım/dağılım ölçülerini birlikte vermek bir zorunluluktur. Normal dağılım özelliği gösteriyorsa: Ortalama, Standart sapma, Minimum ve Maximum değerleri ile özetlenir. Normal dağılım özelliği göstermiyorsa: Ortanca (%50.değer), 1. Çeyrek (%25. değer), 3. Çeyrek (%75. değer), Minimum ve Maximum değerleri ile özetlenir.

Analiz süreçlerinde ise normal dağılım özelliği gösteren nicel veriler parametrik testler ile sınanırken, normal dağılım özelliği göstermeyen nicel veriler non-parametrik testler aracılığı ile sınanır. Parametrik testlerin (testten teste değişen) birçok varsayımı vardır, temel varsayımlarından biri ise normal dağılımdır. Bu nedenle nicel değişken verilerinin normal dağılıma uygunluğuna göre analiz sürecinde kullanacağımız test belirleneceğinden, önemsenmesi ve dikkat edilmesi gereken bir husustur.

### SPSS Uygulaması:

- 1- **Analyze, Descriptive Statistic, Explore seçilir.**
- 2- **Explore** penceresinde sol taraftaki değişken listesinden **nicel değişkenler seçilip** sağ taraftaki **dependent list** penceresine **aktarılır.** (Alt kategorilere göre normal dağılım test edilecekse sol taraftaki değişken listesinden kategorik değişken **Factor list** penceresine **aktarılır.**)
- 3- **Explore** penceresinde **Plots** tıklanır ve **Histogram** ve **Normality plots with tests** seçilir. **Continue** tıklanır.
- 4- **Explore** penceresinde **OK** tıklanır.

**Not:** Explore'dan bu işlem yapıldığında mod değerini ve histogramın normal dağılım eğrisini göremeyiz. **Mod ve histogramın normal dağılım eğrisini için:**

- 1- **Analyze, Descriptive Statistic, Frequencies seçilir.**
- 2- **Frequencies** penceresinde sol taraftaki değişken listesinden özetlenecek **nicel değişkenler seçilip** sağ taraftaki **variables** penceresine **aktarılır.**
- 3- **Frequencies** penceresinde **Statistics** butonu **tıklanır** ve **Mod** seçilir. **Continue** tıklanır.
- 4- **Frequencies** penceresinde **Charts** butonu **tıklanır** ve **Histogram** ve **Show normal curve on histogram** seçilir. **Continue** tıklanır.
- 5- **Frequencies** penceresinde **OK** tıklanır.

**SPSS çıktısı:****X değişkeninin normal dağılıma uygunluğunun sınanması****Explore  
Case Processing Summary**

Cases		Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent	
X	135	100,0%	0	0,0%	135	100,0%	

X değişkeninin geçerli (valid) 135 cevabı olduğu, kayıp (missing) verisi olmadığını görüyoruz.

**Descriptives**

		Statistic	Std. Error
X	Mean	12,0593	,27933
	95% Confidence Interval for Lower Bound	11,5068	
	Mean Upper Bound	12,6117	
	5% Trimmed Mean	12,0432	
	Median	12,0000	
	Variance	10,534	
	Std. Deviation	3,24558	
	Minimum	5,00	
	Maximum	20,00	
	Range	15,00	
	Interquartile Range	4,00	
	Skewness	,058	,209
	Kurtosis	-,423	,414

**Mean:** Ortalama. **Std. Error:** Standart hata. **95% Confidence Interval for Mean Lower Bound:** %95 Güven Aralığında ortalamanın alt sınırı (mean- 2 x Std. Error). **95% Confidence Interval for Mean Upper Bound:** %95 Güven Aralığında ortalamanın üst sınırı (mean + 2 x Std. Error). **5% Trimmed Mean:** En küçük %2,5 ve En büyük %2,5'luk veriler çıkarıldıktan sonraki ortalama (yani toplamda %5'lik uç değerler çıkarıldıktan sonra geriye kalan %95 verinin ortalaması). **Median:** Ortanca. **Skewness:** Çarpıklık katsayısı (-1 +1 aralığında olmalı). **Kurtosis:** Basıklık katsayısı (-1 +1 aralığında olmalı). **Mode:** Tepe değer (en çok tekrar eden değer).

**Statistics**

X		
N	Valid	135
	Missing	0
Mode		12,00

X verisinin ortalama (12,05), mod (12,00) ve ortanca (12,00) değerlerinin birbirlerine çok yakın değerler olduğunu, Çarpıklık (Skewness) ve Basıklık (Kurtosis) katsayılarının -1 +1 aralığında olduğunu görüyoruz.



## Tests of Normality

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
X	,070	135	,099	,987	135	,211

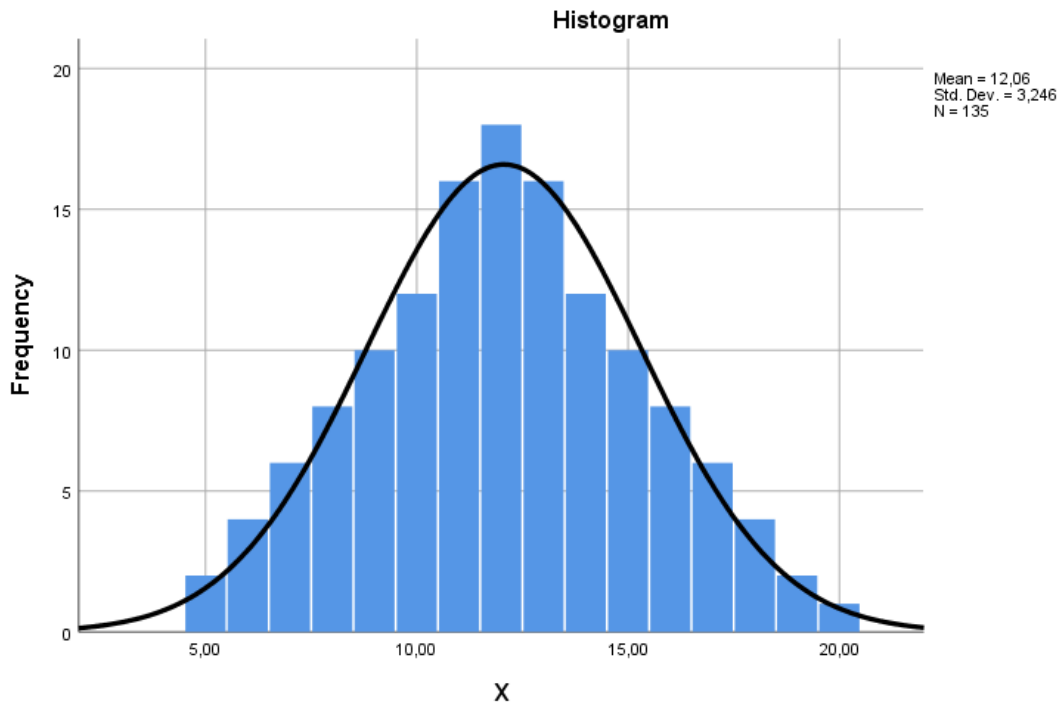
a. Lilliefors Significance Correction

**Sig.:** p değeridir.

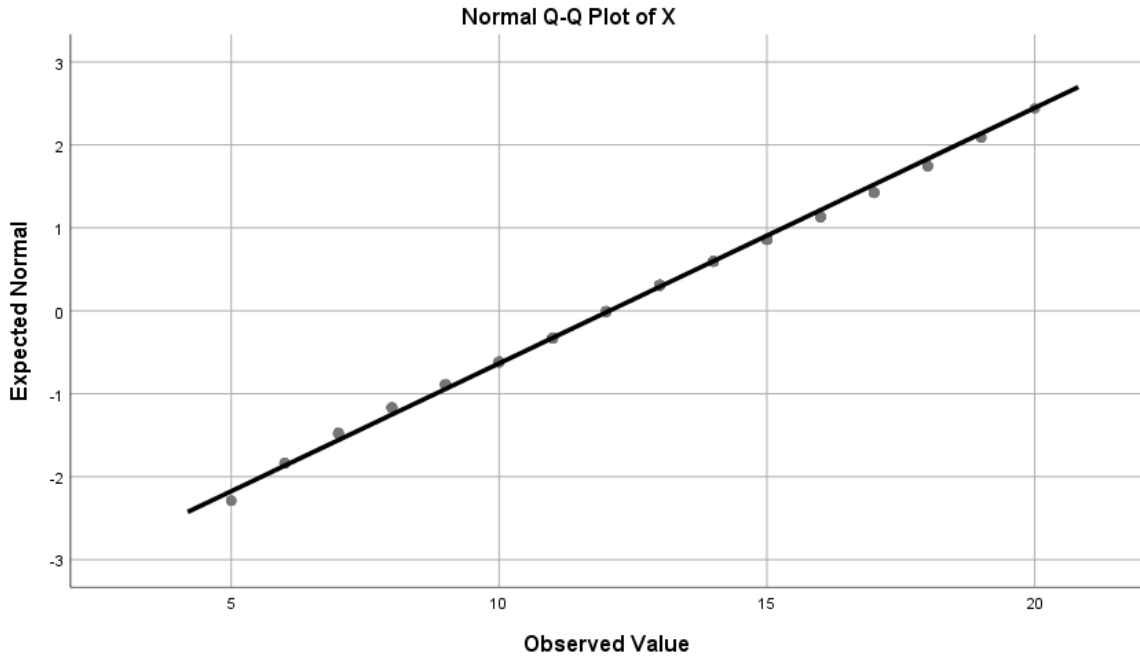
$p > 0,05$  ise değişken normal dağılım özelliği gösteriyor anlamına gelir.

$p < 0,05$  ise değişken normal dağılım özelliği göstermiyor anlamına gelir.

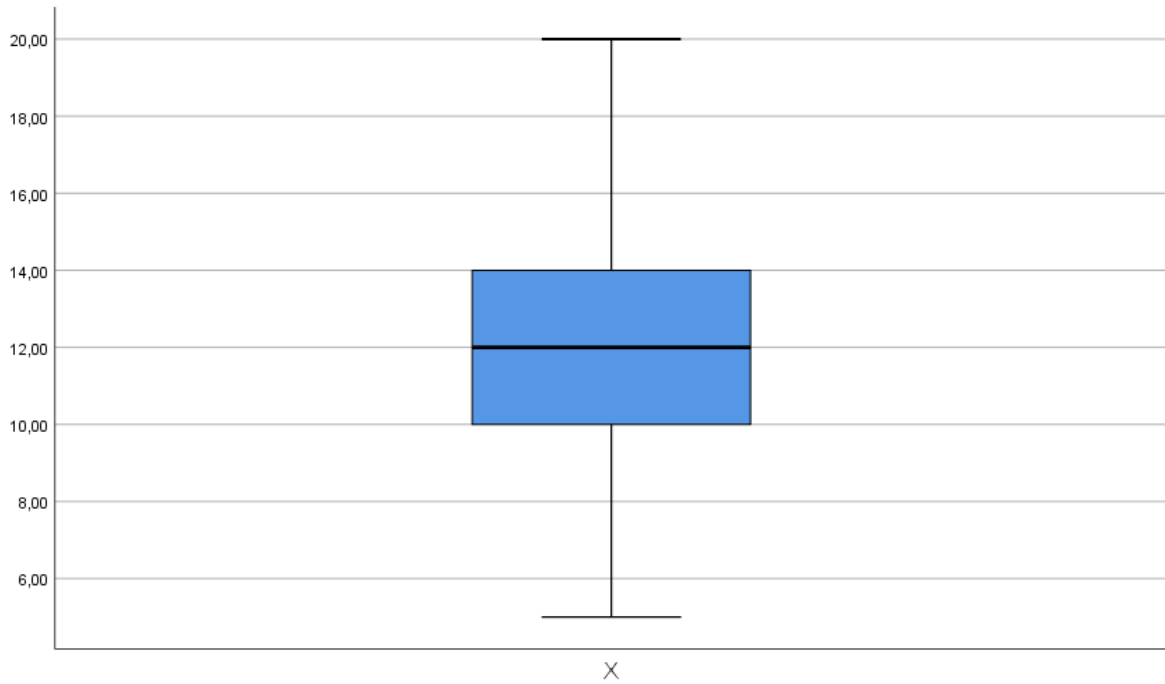
Örneklem sayısı 135 olduğu için ( $n \geq 30$ ) Kolmogorov Smirnov testinin p (Sig.) değerine bakıyoruz.  $p = 0,099$  yani  $p > 0,05$  olduğu için X değişkeni test sonucu itibarı ile normal dağılım özelliği gösteriyor.



Histogram dağılımının görüntüsü ve eğrisi çan eğrisi şeklinde, ortalamanın etrafında simetrik bir dağılım gösteriyor.



Normal Q-Q Plot grafiğinde noktaların 45 derecelik doğru üstünde ya da çok yakınında olduğunu görüyoruz.



Bu grafikteki dikdörtgenin ortasındaki koyu çizgi medyayı gösterir (12), dikdörtgenin alt çizgisi %25. değeri (10) ve dikdörtgenin üst çizgisi %75. değeri (14) gösterir. Grafiğin alt çizgisi en küçük değeri (5) ve üst çizgisi ise en büyük değeri (20) gösterir.

**Sonuç:** X değişkeni normal dağılımın bütün kriterlerine uymakta ve istatistiksel test sonucu da aynı sonuca işaret etmektedir. X değişkeni normal dağılım özelliği gösteriyor.

**Y değişkeninin normal dağılıma uygunluğunun sınanması****Case Processing Summary**

	Valid		Cases Missing		Total	
	N	Percent	N	Percent	N	Percent
Y	135	100,0%	0	0,0%	135	100,0%

Y değişkeninin geçerli (valid) 135 cevabı olduğu, kayıp (missing) verisi olmadığını görüyoruz.

**Descriptives**

		Statistic	Std. Error
Y	Mean	31,0296	,85201
	95% Confidence Interval for Mean	Lower Bound	29,3445
		Upper Bound	32,7148
	5% Trimmed Mean	30,4568	
	Median	28,0000	
	Variance	97,999	
	Std. Deviation	9,89945	
	Minimum	18,00	
	Maximum	56,00	
	Range	38,00	
	Interquartile Range	15,00	
	Skewness	,781	,209
	Kurtosis	-,331	,414

**Statistics**

Y	N	Valid	135
		Missing	0
	Mode		23,00

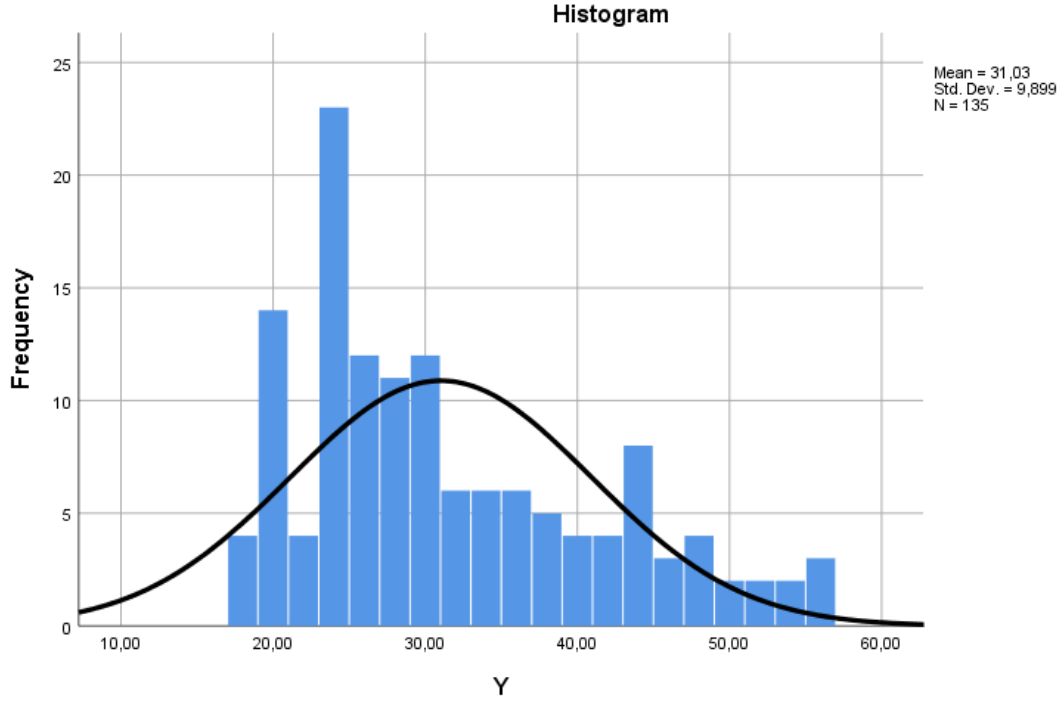
Y verisinin ortalama (31,02), mod (23) ve ortanca (28,00) değerlerinin birbirlerinden farklı olduğunu, Çarpıklık (Skewness) ve Basıklık (Kurtosis) katsayılarının -1 +1 aralığında olduğunu görüyoruz.

**Tests of Normality**

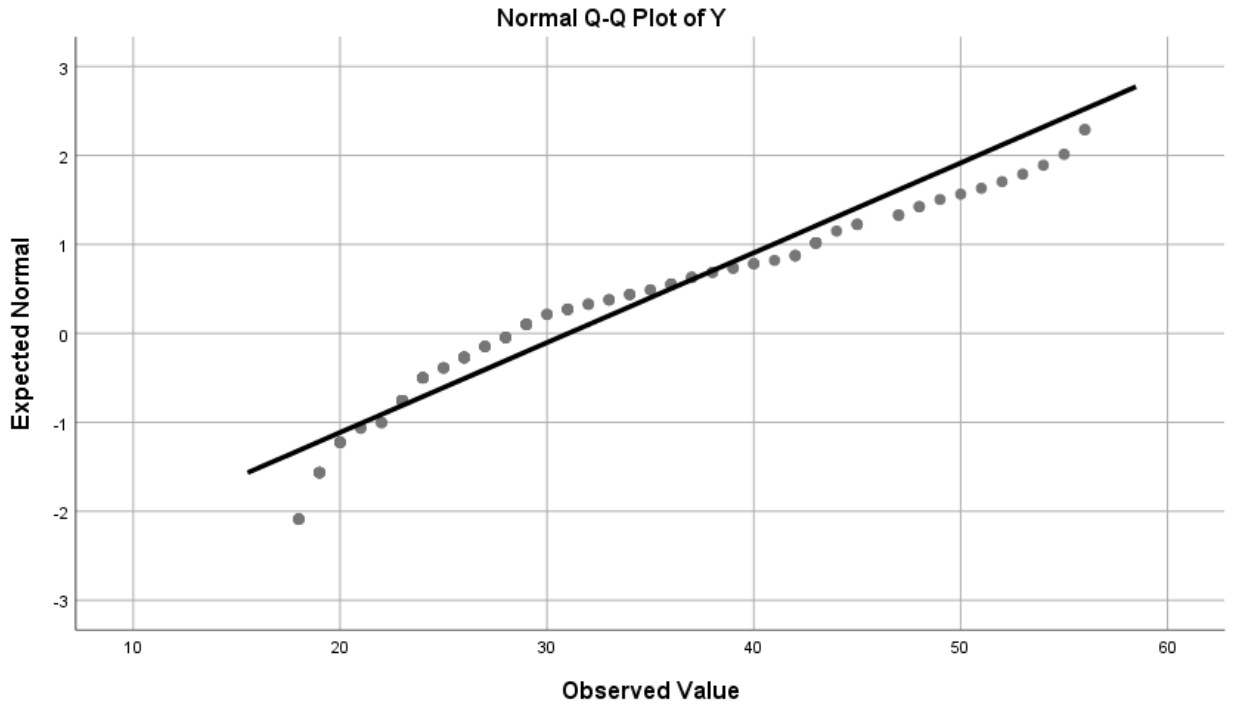
	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Y	,159	135	,000	,918	135	,000

a. Lilliefors Significance Correction

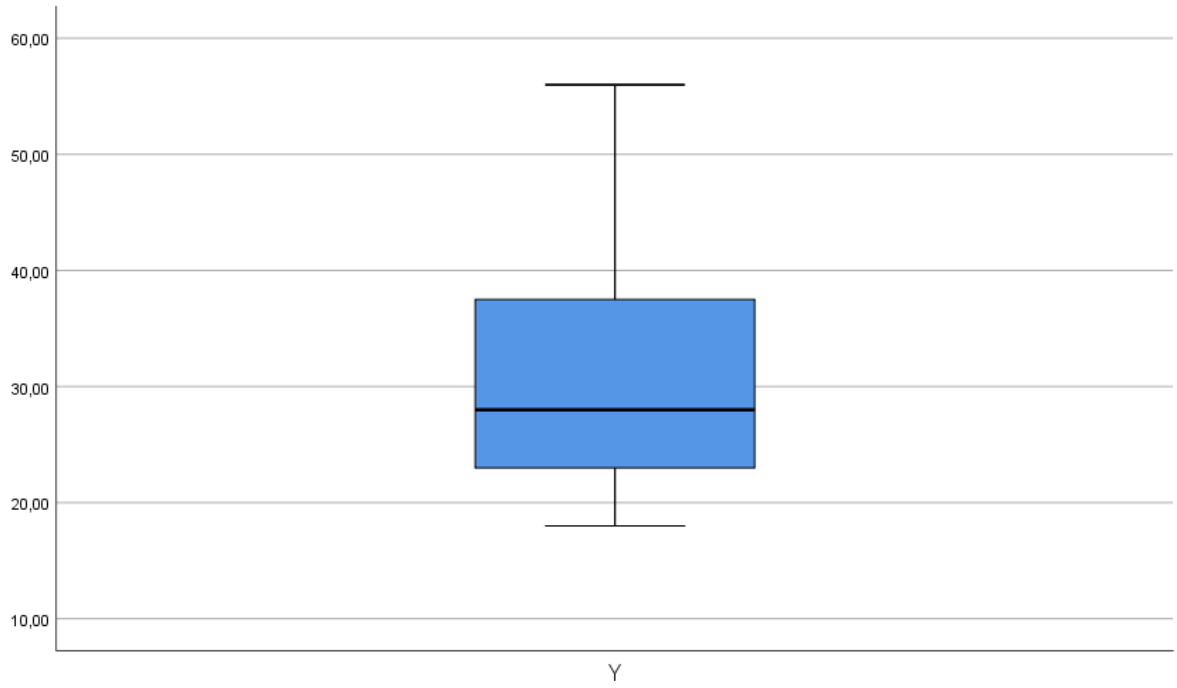
Örneklem sayısı 135 olduğu için ( $n \geq 30$ ) Kolmogorov Smirnov testinin p (Sig.) değerine bakıyoruz.  $p=0,000$  yani  $p < 0,05$  olduğu için Y değişkeni test sonucu itibari ile normal dağılım özelliği göstermiyor.



Histogram dağılımının görüntüsü ve eğrisi: çan eğrisinde basıklık ve çarpıklık (solda bir yığılma) var. Dağılım ortalamanın etrafında simetrik değil.



Normal Q-Q Plot grafiğinde noktaların büyük çoğunluğunun 45 derecelik doğrudan uzak olduğunu görüyoruz.



Bu grafikteki dikdörtgenin ortasındaki koyu çizgi medyayı gösterir (28), dikdörtgenin alt çizgisi %25. değeri (23) ve dikdörtgenin üst çizgisi %75. değeri (38) gösterir. Grafiğin alt çizgisi en küçük değeri (18) ve üst çizgisi ise en büyük değeri (56) gösterir.

**Sonuç:** Y değişkeni normal dağılımın kriterlerinin çoğunluğuna uymamakta ve istatistiksel test sonucu da aynı sonuca işaret etmektedir. Y değişkeni normal dağılım özelliği göstermiyor.

### 1.5. Ölçeklerin Güvenilirliği

Güvenilirlik, katılımcıların ölçek maddelerine verdikleri cevaplar arasındaki tutarlılık olarak tanımlanabilir. Güvenilirlik, ölçeğin ölçmek istediği özelliği ne derece doğru ölçtüğünü ifade eder (Büyüköztürk, 2010). Güvenilirlik “ölçümdeki hatalarla” ilgili bir kavramdır. Eğer ölçmede hata yoksa ya da hata miktarı az ise güvenilirlik yüksek olacaktır. Güvenilirliğin en genel tanımı “ölçmenin hatalardan arınmış olmasıdır” şeklinde yapılabilir (Can, 2019). Ölçek geliştirmenin en önemli kısmı güvenilirlik ve geçerlilik analizleridir ve olmazsa olmazdır. Analizlerin veri kaynağını oluşturan ölçeklerin güvenilirlik ve geçerlilik çalışmasının doğru yapılması çok önemlidir. Ölçek geliştirmenin kendisi başlı başına büyük ve önemli bir katkıdır! Bu kitap kapsamında ölçek geliştirmeye değinilmeyeceğinden, geçerlilik analizi kapsam dışında tutulmuştur.

Araştırmanızda kullanacağınız ölçeğin doğru bir şekilde geçerlilik ve güvenilirlik analizinin yapılmış olduğundan emin olduktan sonra bu ölçeği kendi araştırmanızda kullanabilirsiniz. Daha önce geçerlilik ve güvenilirlik analizi yapılmış bir ölçeğin; yeni bir çalışmada kullanılırken geçerlilik analizinin yapılması zorunluluk değildir, ancak güvenilirlik analizinin yapılması zorunluluktur ve yöntem kısmında belirtilmesi gerekir.

Ölçeklerin güvenilirliği farklı yöntemlerle ölçülebilir.

**Test tekrarı ile ölçülebilir:** Farklı iki zamanda aynı katılımcıların aynı ölçeğe verdiği cevaplar üzerinden yapılır. Test tekrarı arasında geçen zaman da kontrol edemeyeceğimiz koşulların, katılımcılara dolayısı ile test sonucuna etkisi söz konusu olabilir!

**Cronbach Alfa Güvenilirliği:** Tek uygulamada katılımcıların verdiği cevaplar üzerinden bir güvenilirlik katsayısı hesaplar. Cronbach Alfa katsayısı 0 ile 1 arasında bir değer alır. Cronbach Alfa katsayısının 0’a (sıfıra) yakın olması ölçeğin güvenilmez olduğunu, bu katsayının 1’e (bire) yaklaşması ise ne kadar güvenilir olduğunu ifade eder. Bu katsayının 0,7’nin üzerinde olması beklenir (Pallant, 2017). Tatmin edici Cronbach Alfa katsayısı ise 0,8 ve üzeri değerlerdir.

#### SPSS Uygulaması:

- 1- **Analyze, Scale, Reliability Analysis** seçilir.
- 2- **Reliability Analysis** penceresinde sol taraftaki değişkenler listesinden **ölçeği oluşturan ifadeler** seçilip sağ taraftaki **Items** penceresine aktarılır.  
**Model** kısmında **Alpha** seçili olmalıdır.
- 3- **Statistics** butonu tıklanır ve **Reliability Analysis: Statistics** penceresinde **Descriptives for** kısmından **Item, Scale** ve **Scale if Item Deleted** seçilir.  
**Inter-Item** kısmından **Correlations** seçilir.  
**Summaries** kısmından **Correlations** seçilir.  
**Continue** tıklanır.
- 4- **Reliability Analysis** penceresinde **OK** tıklanır.

**SPSS çıktısı:**

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,985	,985	10

**Reliability Statistics:** Güvenilirlik istatistiği

**Cronbach's Alpha:** Cronbach Alfa katsayısı (Ölçeğin toplamının Cronbach Alfa katsayısı)

**N of Items:** Ölçeğin madde/ifade sayısı

Item Statistics			
	Mean	Std. Deviation	N
depresyon ölçeği soru 1	4,1259	,94181	135
depresyon ölçeği soru 2	4,0667	1,09408	135
depresyon ölçeği soru 3	3,9630	,89282	135
depresyon ölçeği soru 4	3,9630	1,07486	135
depresyon ölçeği soru 5	4,0000	1,03664	135
depresyon ölçeği soru 6	4,1111	1,02700	135
depresyon ölçeği soru 7	4,0667	1,00892	135
depresyon ölçeği soru 8	4,0815	1,01520	135
depresyon ölçeği soru 9	4,0074	1,09610	135
depresyon ölçeği soru 10	4,0519	1,00237	135

**Item Statistics:** Ölçek ifadelerinin özet değerlerini içeren tablo: her bir ölçek ifadesine ait ortalama (**Mean**), standart sapma (**Std. Deviation**) ve cevap sayısı (**N**).

Inter-Item Correlation Matrix										
	D s1	D s2	D s3	D s4	D s5	D s6	D s7	D s8	D s9	D s10
D s1	1,000	,955	,849	,882	,856	,873	,973	,949	,903	,934
D s2	,955	1,000	,820	,884	,842	,917	,949	,983	,877	,943
D s3	,849	,820	1,000	,815	,726	,729	,856	,802	,770	,778
D s4	,882	,884	,815	1,000	,830	,849	,876	,871	,919	,937
D s5	,856	,842	,726	,830	1,000	,743	,828	,851	,854	,804
D s6	,873	,917	,729	,849	,743	1,000	,879	,886	,815	,930
D s7	,973	,949	,856	,876	,828	,879	1,000	,920	,897	,934
D s8	,949	,983	,802	,871	,851	,886	,920	1,000	,865	,927
D s9	,903	,877	,770	,919	,854	,815	,897	,865	1,000	,903
D s10	,934	,943	,778	,937	,804	,930	,934	,927	,903	1,000

**Inter-Item Correlation Matrix:** Ölçek ifadelerinin birbirleri ile ilişki düzeylerini korelasyon katsayısı (r) ile gösteren matrix tablo.

**Summary Item Statistics**

	Mean	Minimum	Maximum	Range	Maximum / Minimum	Variance	N of Items
Inter-Item Correlations	,871	,726	,983	,257	1,354	,004	10

**Summary Item Statistics:** Ölçek ifadelerinin birbirleri ile toplam ilişki düzeylerinin özetlendiği tablo.

**Inter-Item Correlations Mean:** Ölçek ifadelerinin birbirleri ile olan ilişki katsayılarının ortalaması (ortalama r).

**Item-Total Statistics**

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
depresyon ölçeği soru 1	36,3111	75,171	,969	,971	,982
depresyon ölçeği soru 2	36,3704	72,623	,969	,984	,982
depresyon ölçeği soru 3	36,4741	77,923	,836	,823	,986
depresyon ölçeği soru 4	36,4741	73,594	,929	,942	,983
depresyon ölçeği soru 5	36,4370	75,322	,861	,813	,985
depresyon ölçeği soru 6	36,3259	74,878	,898	,895	,984
depresyon ölçeği soru 7	36,3704	74,175	,961	,970	,982
depresyon ölçeği soru 8	36,3556	74,171	,954	,978	,983
depresyon ölçeği soru 9	36,4296	73,381	,922	,900	,984
depresyon ölçeği soru 10	36,3852	74,298	,960	,969	,982

**Item-Total Statistics:** Ölçek ifadelerinin istatistiksel değerleri

**Scale Mean if Item Deleted:** İlgili Ölçek ifadesi silindiğinde (çıkarıldığında) ölçek toplamının ortalama değeri.

**Cronbach's Alpha if Item Deleted:** İlgili Ölçek ifadesi silindiğinde (çıkarıldığında) ölçeğin –geriye kalan ifadelerin toplamının- Cronbach Alfa katsayısı.

**Scale Statistics**

Mean	Variance	Std. Deviation	N of Items
40,4370	91,890	9,58591	10

**Scale Statistics:** Ölçek istatistiği

**Mean:** Ölçeğin (bütün ifadelerin toplamının) puan ortalaması



## 1.6. Hipotez Testleri: Parametrik ve Non-Parametrik Testler

### 1.6.1. Hipotez

İstatistiğin analiz sürecinin başlangıç noktası hipotezdir. Araştırmacı sorunsallaştırdığı bir olgu üzerine yola çıkar ve sahadan buna yönelik verileri toplar. Toplanan veriler durumu tanımlamaya ve değişkenler arasındaki farklılığı ya da ilişkiyi açıklamaya, keşfetmeye yöneliktir. Hipotez araştırma sorununa yönelik sınanan iddiadır. Değişkenler arasındaki ilişki, farklılık ve açıklayıcılığa yönelik analizlerde hipotez söz konusudur.

İstatistiksel hipotez iddia edilen bir önermedir (Öztürk ve Özbek, 2004). Gözlenen olaylar ya da olaylar bütünü içinde olası görülen, henüz kanıtlanmamış ilişkilere ait önermelere hipotez (denence) denir. Hipotez, bilimsel araştırmanın gerekçesidir. Olaylar arasında ilişki kurmak ya da olayları bir nedene bağlamak üzere oluşturulur. Bilimsel gerçeklere, genellemelere ancak hipotezlerle ulaşmak olasıdır (Gökçe: 2018).

Araştırmacının bir kuramdan, bir olay ya da olgudan yola çıkarak, doğru olduğunu düşündüğü hipoteze araştırma hipotezi denir (Can, 2019). Araştırma hipotezine, karşıt (alternatif) hipotez de denir. Genel olarak hipotezler Yokluk (Sıfır) Hipotezi ve Araştırma (Alternatif) Hipotezi şeklinde ikiye ayrılır.

$H_0$ : Yokluk Hipotezi (Null Hypothesis)

$H_1$ : Araştırma Hipotezi

Yokluk hipotezinde ( $H_0$ ), ilgilendiğimiz değişkenler arasında anlamlı bir ilişki olmadığını ya da bağımsız değişkenin oluşturduğu alt gruplarda bağımlı değişken açısından farklılık olmadığını iddia etmekteyiz. Yokluk hipotezine alternatif olarak ortaya attığımız hipotezde ( $H_1$ ) ise anlamlı ilişki ya da farklılık olduğunu önermekteyiz (Işık, 2014).

### 1.6.2. Hipotez Testleri: Parametrik ve Non-Parametrik Testler

Hipotezler, istatistiksel teknikler (testlerle) ile sınanır. İstatistiksel testler Hipotez testleri olarak da adlandırılır. Hipotez testleri Parametrik ve Non-Parametrik testler olarak ikiye ayrılır.

“Parametrik” kelimesi “parametre” ya da “evrenin özelliğinden” gelmektedir. Parametrik testler örneklemin alındığı evren hakkında varsayımlar yapar. Örneklemin de bu varsayımları göstermesini bekler. Bu varsayımların sağlanması durumunda parametrik testler uygulanabilir. Parametrik testler, değişkenin aralık ya da oran düzeyinde ölçüldüğü ve bu ölçüm verilerinin normal dağılım sergilediği durumlarda kullanılır. Parametrik testlerin diğer bir temel varsayımı ise Örneklemin Evrenden rastgele/tesadüfi bir modelle seçilmesidir. Örneklemin rastgele bir modelle oluşturulması gerekli ise de bu durum gerçek yaşam araştırmalarında genellikle böyle değildir (Pallant, 2017). Parametrik testlerin iki temel varsayımı nicel veri (aralık ya da oran ölçeğinde) ve normal dağılımdır (Field, 2009). Bunun dışında da Parametrik testlerin, testten teste değişen ek varsayımları vardır (Bu varsayımlar testlerde verilmiştir). Parametrik testlerin uygulanabilmesi için varsayımlarının sağlanması gerekir, varsayımların sağlanamadığı durumlarda hipotezin sınanması için Non-Parametrik testler kullanılır.

Non-Parametrik testler, Parametrik testler gibi katı gerekliliklere sahip değildir, evrenle ilgili varsayımlarda bulunmaz. Non-Parametrik testler sınıflama ve sıralama ölçeklerinden elde ettiğimiz veriler için idealdir. Örneklem sayısının az olduğu ya da verilerinizin parametrik testlerin varsayımlarını yerine getirmediği durumlarda parametrik olmayan testler kullanılır

(Pallant, 2017). Non-Parametrik testler, sınıflama, sıralama ölçek verileri ve normal dağılım özelliği göstermeyen aralık ve oran ölçeği verileri için kullanılır.

Non-Parametrik testlerin bazı dezavantajları vardır. Parametrik testlere göre daha az duyarlı olma eğilimindedirler ve bu yüzden gruplar arası gerçekte var olan farkları tespit etmede başarısız olabilirler. Bu yüzden “doğru” veri türüne sahipseniz ve ilgili testin varsayımları yerine getirebildiniz takdirde her zaman parametrik testleri kullanmak daha iyi bir seçenektir (Pallant, 2017; Can, 2019).

### 1.6.3. Hipotez Testleri ile Hipotezin Sınanması

İstatistiksel testler belli bir Alfa ( $\alpha$ ) düzeyinde Yokluk Hipotezini ( $H_0$ ) sınar. Alfa ( $\alpha$ ) kabul edilen hata payını –anlamlılık düzeyini- ifade eder ve sınır değerdir. Testlerde genellikle Alfa ( $\alpha$ ) 0,05 (%5) sınır değeri üzerinden hipotezler sınanır.<sup>1</sup> İstatistiksel test sonucunda ulaştığımız p değeri 0,05’ten küçük ise Yokluk Hipotezini ( $H_0$ ) ret edip, Araştırma Hipotezini ( $H_1$ ) kabul ederiz. Test sonucunda Araştırma Hipotezinin kabul edilmesi için test sonucu ulaşılan **p** değerinin **alfa** değerinden küçük olması gerekir.

#### İstatistiksel anlamlılık

İstatistiksel anlamlılık, Alfa<sup>2</sup> ( $\alpha$ ) ve p değerleri sonucunda karar verilen bir durumdur. Alfa ( $\alpha$ ) düzeyi, anlamlılık düzeyi (level of significance) olarak da adlandırılır (Field, 2009).  $H_0$  gerçekten doğru ise araştırmacı bu doğru iddiayı testin sonucundaki hesapladığı değere (p) ve bu değerın sınıandığı Alfa ( $\alpha$ ) değerine göre ret ederse, hata yapmış olacaktır, istatistikte buna “Tip I Hata” denir. Bu hatanın yapılması olasılığı da Alfa ( $\alpha$ ) ile gösterilir. Bu olasılık “testin önem düzeyi” veya “anlamlılık düzeyi” olarak da adlandırılır. Yani  $\alpha=0,05$  önem düzeyinde test yapıldı denildiğinde, bunun anlamı; araştırmacı doğru bir  $H_0$  hipotezini reddetmek için 0,05’lik (%5’lik) bir hata yapma riskini kabulleniyor demektir. Bu hipotezin sınanmasında kabul edilen risktir (Terzi, 2019). Alfa ( $\alpha$ ) değişkenler arasında fark/ilişki sınanırken kabul edilen hata olasılığıdır. Bu olasılık araştırmanın aynı evrenden örneklerle tekrarlanması halinde ortaya çıkabilecek hata düzeyidir. Standart bir değer olarak genelde 0,01 ya da 0,05 olarak seçilir. Eğer farklı bir değer seçilecek olursa genellikle 0,05’ten daha düşük olan değerler seçilir. Alfa değeri, analizler yapılmadan seçilen kıstastır (Işık, 2014). Örneklemdeki analiz sonucunda ulaşılan **p** değeri, gözlenen bir etkinin sadece şanstın kaynaklanma olasılığını gösterir. “P” olasılık (probability) anlamına gelir ve gruplar arasında gözlemlenen herhangi bir farkın/ilişkinin şansa bağlı olma olasılığını ölçer. Bir olasılık olan **p**, 0 ile 1 arasında herhangi bir değer alabilir. 0’a yakın değerler, gözlemlenen farkın/ilişkinin şansa bağlı olma olasılığının düşük olduğunu gösterirken, 1’e yakın p değerleri ise farkın/ilişkinin olmadığını gösterir. P değeri Yokluk Hipotezine ( $H_0$ ) karşı kanıt gücünü ölçer (Whitley ve Ball, 2002). P değeri “anlamlılığın tam seviyesi”dir. Yokluk hipotezinin sınanmasında, p değerinin alfa ( $\alpha$ ) değerinden düşük olması

<sup>1</sup> İstatistiksel anlamlılığın (statistical significance) belirlenmesi için %5 düzeyinin en çok kabul edilebilir olasılık düzeyi olarak ilk kullanımı Ronald Fisher (1925) tarafından yapılmıştır (Cowles ve Davis, 1982). Fisher (1925) istatistiksel anlamlılık için kriter değerler -anlamlılık sınırları- olarak 0,01 (%1), 0,02 (%2) ve 0,05 (%5) düzeylerinin kullanılmasını ve en yüksek kabul edilebilir seviyenin ise 0,05 (%5) olduğunu ilk defa literatürde ifade etmiştir (Field, 2009).

<sup>2</sup> Alfa ( $\alpha$ ) düzeyi, hata olasılığını ifade eder. Alfa ( $\alpha$ ) düzeyinin 0,05 (%5) olarak kabul edilmesi, test sonucu kararların hatalı olma olasılığının %5 olmasını belirtir. Alfa ( $\alpha$ ) düzeyinin 0,05 olması sonuçların %95 güvenilirlikte ya da %5 şans eseri olma olasılığını gösteren bir değerdir. Test sonucu ulaşılan kararı, aynı evrenden örneklem çekerek 100 kez tekrar ettiğimizde 95’inde aynı sonuçlara ulaşacağımız varsayımına dayanır (Field, 2009).

beklenir. Bu sonuç elde edildiğinde Yokluk Hipotezi ( $H_0$ ) ret edilir ve Araştırma Hipotezi ( $H_1$ ) için destek elde edilmiş olur (Işık, 2014).

İstatistiksel analizler sadece Hipotez testleri ve bunun sonucunda ulaşılan p değeri itibari ile  $p > 0,05$  ise  $H_0$  hipotezinin kabul edilmesi ya da  $p < 0,05$  ise  $H_0$  hipotezinin ret edilmesi şeklinde raporlaştırılması yetersizdir. Hipotez testleri sonucunda ulaşılan **test değerleri** ve istatistiksel anlamlılık düzeyleri **p** ile verilirken, **testin gücü (power)**, anlamlı farklılık/ilişkinin **etki büyüklüğü (effect size)** ve **%95 güven aralığı değerleri** ile birlikte raporlaştırılması, istatistiksel açıklayıcılık için gereklidir.

### 1.7. Hata Tipleri ve Güç

Hipotez testleri sonucunda verilen kararların yanlışlığı iki şekilde sınıflandırılır. İki tip hata söz konusudur. Birincisi, Yokluk Hipotezi evrende gerçekte doğru olmasına rağmen Yokluk Hipotezinin reddedilmesi (Tip I Hata); ikincisi ise Yokluk Hipotezi evrende gerçekte yanlış olmasına rağmen Yokluk Hipotezinin kabul edilmesi (Tip II Hata) durumudur.

#### Hata Tipleri:

		Test sonucu verilmiş karar	
		Fark/ilişki yok	Fark/ilişki var
Evrende (gerçekte) var olan durum	Fark/İlişki yok	Doğru karar	Tip I hata ( $\alpha$ )
	Fark/ilişki var	Tip II hata ( $\beta$ )	Doğru karar İstatistiksel Güç (Power)

Kaynak: Can, 2019; Işık, 2014; Kul, 2014'ten uyarlanmıştır.

**Tip I Hata:** “Evrende (gerçekte) fark/ilişki yok iken test sonucunda fark/ilişki var” hatası

**Tip II Hata:** “Evrende (gerçekte) fark/ilişki var iken test sonucunda fark/ilişki yok” hatası

**Tip I hata Alfa ( $\alpha$ ) düzeyine bağlıdır.** Alfa ( $\alpha$ ) düzeyi Tip I hata yapma olasılığı için bir üst sınırdır (Field, 2009).  $\alpha = 0,05$  düzeyinde test sonuçlarını sınarken %5’lik hata payı ile çalışmayı kabul ediyoruz demektir.

**Tip II hata Beta ( $\beta$ ) düzeyine bağlıdır.**  $\beta$  düzeyi de  $\alpha$  düzeyine bağlıdır.  $\beta$  güven aralığı olarak ifade edilir ve  $\beta = 1 - \alpha$  dır. Yani testte  $\alpha = 0,05$  düzeyinde çalışıyor iken 0,95 (%95) güven aralığında karar vermiş oluyoruz (Kul, 2014).

Alfa ( $\alpha$ ) düzeyini düşürerek Tip I hatadan kaçınırken de Tip II hataya düşme ihtimalini arttırmış oluruz!

**Tip I ve Tip II hatadan kaçınmanın yolu= Gücü Arttırmak**

Araştırmacı, kullandığı testlerin (Parametrik ve Non-Parametrik testler) sınıadığı durumla ilgili fark/ilişki olup olmadığını **doğru bir biçimde tespit etmesini** bekler. Buna **testin gücü (power)** adı verilir (Pallant, 2017). Evrende (gerçekte) fark/ilişki olduğu durumda testin farklılığı/ilişkiyi tespit edip doğru karar vermesi, testin istatistiksel gücü olarak (power) olarak isimlendirilir (Işık, 2014). Genel olarak, Parametrik testler Non-Parametrik testlere göre daha güçlü testlerdir. Bu nedenle varsayımlarının sağlanması durumunda Parametrik testlerin tercih edilmesi gerekir. **Testin gücü (Power)**, örneklem büyüklüğü (n), test ettiğimiz değişkenin değerleri ve test değerleri ile hesaplanır. Testin gücünün en az %80 olması beklenir. Testin gücünü etkileyen başlıca faktörler:

**Alfa ( $\alpha$ ) düzeyi:** 0,05 genel kabul görmüş (sosyal bilimlerde) düzeydir. Bu düzeyin yukarı yönlü değiştirilmesi kabul edilebilir değildir, aşağı yönlü değiştirilmesi ise tercih edilebilir olmakla birlikte Tip II hataya düşme ihtimalini artıracaktır.

**Örneklem büyüklüğü:** Evreni temsil edecek büyüklükte bir örneklemle çalışmalı ve Evrenin temel karakterleri örnekleme eşit büyüklükte temsil edilmedi. Örneklem seçimi, Tesadüfi bir modelle gerçekleştirilmedi. Yani evrendeki her bir birimin örnekleme girme ihtimali eşit olmalıdır.

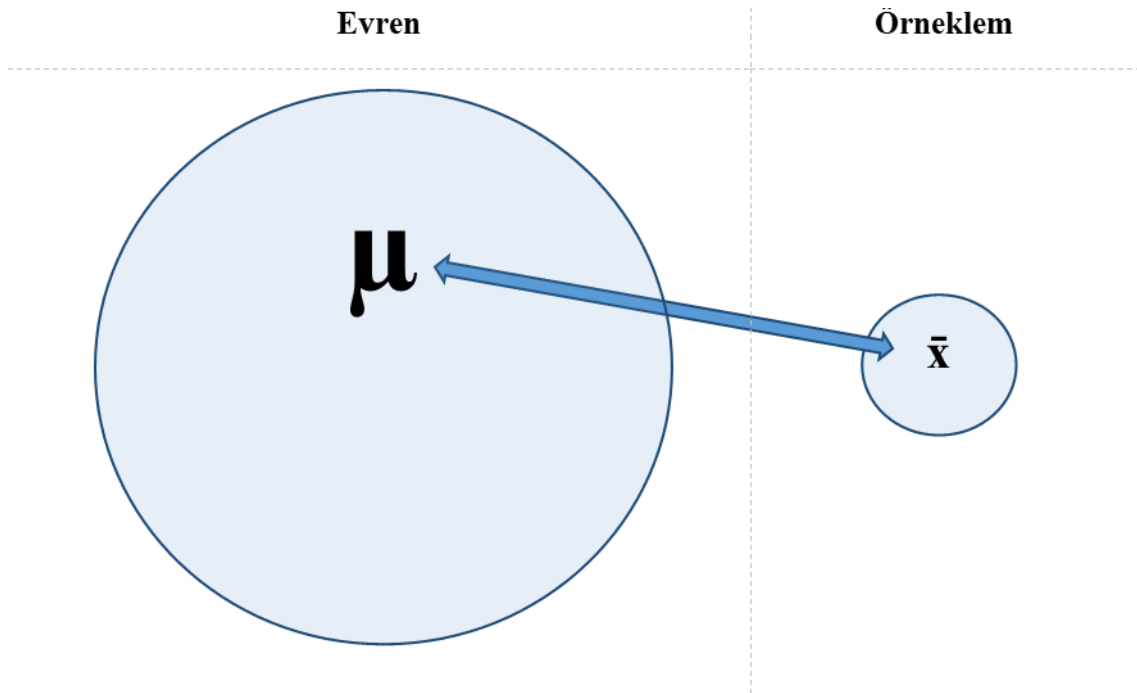
**Ölçeklerin Güvenilirliği ve Geçerliliği:** Ölçeklerin güvenilirliğinin yüksek olması testin gücünü arttıracak bir faktördür (Işık, 2014). Kullanılan ölçeklerin, geliştirilmesi ve uyarlanması aşamasında sağlıklı bir şekilde geçerlilik ve güvenilirlik çalışmalarının yapılması gerekir. Geçerliliği ve güvenilirlik analizleri yapılmış ölçekler araştırmalarda kullanılmalı ve her yeni kullanımda güvenilirlik analizi mutlaka yapılmalıdır.

Örneklem büyüklüğü ve testin gücü SPSS üzerinden maalesef hesaplanamaz. Örneklem büyüklüğünü ve ulaşılan test değerleri üzerinden testin gücünü (power) hesaplayan internet siteleri ve bilgisayar programları mevcuttur.

## 2.BÖLÜM

# ORTALAMA KARŞILAŞTIRMALARI VE NON-PARAMETRİK TESTLERİ

### 2.1. Tek Örneklem t Testi <sup>3</sup>



Nicel bir değişkenin Örneklem ortalaması ile Evren ortalaması arasında anlamlı bir fark olup olmadığını sınamak için parametrik bir test olan Tek Örneklem t Testi (One-Sample T Test) yapılır. Bu testin yapılabilmesinin varsayımları; Örneklemen Evrenden tesadüfi/ rastgele bir şekilde seçilmesine, test edilecek değişkenin en az aralık ölçeğinde bir değişken olmasına ve bu değişkenin verilerin de normal dağılım özelliği göstermesine dayanır (Can, 2019).

Bu varsayımların yerine getirilmesi testin sağlıklı sonuçlar üretmesi için önemli ve esnetilemeyecek koşullardır.

Örneklem ve Evren arasında söz konusu nicel değişken için bir uyum/farksızlık araştırmacı için önemli/anlamlı olabileceği gibi (söz konusu nicel değişken için örneklem ve evren arasında homojenlik nedeni ile), müdahaleli bir çalışma için söz konusu değişkenin Örneklem ortalamasının Evren ortalamasından farklılaşması da önemli/anlamlı olabilir. Sonuç olarak Tek Örneklem t Testi bize madalyonun her iki yüzü (fark var ya da yok) hakkında objektif bir sonuç üretmiş olur.

Müdahalesiz bir çalışmada Örnek Evrenden tesadüfi bir şekilde seçilmiş ise Evrenin temel karakterlerine ait değişkenlerin Örneklemde de aynı şekilde temsil edilmesini bekleriz. Bu

<sup>3</sup> One-Sample t Test

kurgu içinde evrenin temel karakterlerinin (örneğin; yaş, eğitim yılı, aylık gelir vs.) örneklemdaki değerleri ile farklılaşmamasını bekleriz.

Müdahaleli bir çalışmada yapılan müdahalenin değişken üzerinde bir farklılaşma yaratıp yaratmadığını da yine bu test aracılığı ile sınayabiliriz. Örneğin Kanser hastalarının X kan değeri ortalaması 100 birim iken ve X kan değerinin düşmesinin kanser hastalığının tedavisi için olumlu bir etkisinin bulunduğu durumda, X kan değerinin düşürülmesi için geliştirilen bir A ilacının kanser hastaları üzerinde uygulandığını düşünelim. Kanser hatalarında A ilacını kullanan örnek grubunda X kan değeri ortalamasının 90 birim olduğu durumda, Kanser hastalarının X kan değeri ortalaması olan 100 birim ile A ilacı ile müdahale yapılan grubun X kan değeri ortalaması olan 90 birim arasında gözlenen bu farkın anlamlı bir farklılık yaratıp yaratmadığı da yine bu testle sınıanır.

### **Test neyi yapar:**

Tek Örneklem t Testi, tek bir değişkenin örneklem ortalamasının belirtilen –evren- değerden farklı olup olmadığını test eder (IBM SPSS, 2017). Nicel (en az aralık ölçeğinde) bir değişkenin Evren ortalaması ile Örneklem ortalamasını karşılaştırılır.

### **Varsayımları:**

Örneklem Evrenden rastgele seçilmiştir.

Test ettiğimiz değişken en az aralık ölçeğindedir ve normal dağılım özelliği gösterir.

**Non-Parametrik testi:** Yok

### **Hipotezleri:**

$H_0$ : Evren ortalaması ile örneklem ortalaması arasında fark yoktur. (sınanan hipotez)

$H_1$ : Evren ortalaması ile örneklem ortalaması arasında fark vardır.

### **Test sonucu:**

$p > 0,05$  ise  $H_0$  hipotezi kabul edilir: Evren ortalaması ile örneklem ortalaması arasında fark yoktur.

$p < 0,05$  ise  $H_0$  hipotezi ret edilir: Evren ortalaması ile örneklem ortalaması arasında fark vardır.

### **Etki Büyüklüğü:**

Yapılan hipotez testi bize Evren ortalaması ile Örnek ortalaması arasında istatistiksel olarak anlamlı bir fark olup olmadığını söyler, ancak fark var ise bu farkın etkisi ya da büyüklüğü konusunda bilgi vermez. Etki büyüklüğü bize etkinin önemi konusunda objektif bir ölçüm sağladığı için faydalıdır (Field, 2009). Etki büyüklüğü sonuçların yorumlanmasında dikkate alınması gereken istatistiksel bir değerdir. Hipotez testlerinde etki büyüklüğü değerlerinin de hesaplanarak yorumlanması sonuçların anlaşılabilirliğini arttıracaktır (Büyüköztürk, 2010). Hipotez testinde fark bulunduğu durumda etki büyüklüğüne bakılır. Hipotez testleri için

hesaplanan farklı etki büyüklüğü değerleri söz konusudur.<sup>4</sup> Tek örneklem t testi için kullanılan etki büyüklüğü değerleri:

**Cohen d istatistiği (d):** Cohen'in d değeri karşılaştırılan ortalamaların birbirlerinden kaç standart sapma uzaklaştığını yorumlama imkânı verir (Büyüköztürk, 2010). Cohen'in etki büyüklüğü (d istatistiği) işaretinden bağımsız olarak değerlendirilir ve her değeri alabilir. d değerinin 0,2 olması küçük etki, 0,5 orta ve 0,8 olması ise büyük (large) etki olarak değerlendirilir (Cohen, 1992; Green ve Salkind, 2014). Cohen d istatistiği iki farklı formülle hesaplanabilir:

$$d = \frac{\text{Örneklemin ortalaması} - \text{Evrenin ortalaması}}{\text{Örneklemin standart sapması}}$$

$$d = \frac{t}{\sqrt{n}}$$

**r etki değeri:** Pearson'un korelasyon katsayısı olan **r** bir etki değeri katsayısıdır. r değeri 0 (etki yok) ile 1 (mükemmel etki/perfect effect) arasında bir değer alır. r değeri işaretinden bağımsız olarak değerlendirilir. r değerinin 0,1 olması küçük etki, 0,3 olması orta ve 0,5 olması büyük (large) etki olarak değerlendirilir (Field, 2009). r katsayısının karesi ( $r^2$ ) toplam varyansın ne kadarını açıkladığını ifade eder.  $r^2$  değeri üzerinden bağımsız değişkenin bağımlı değişken üzerindeki değişimin ne kadarını açıkladığı bilgisine de ulaşırız.

Cohen (1992) r katsayısının etki düzeylerinin yorumlanmasına ilişkin -yaygın olarak kullanılan- şu sınıflandırmayı yapmıştır:

**r = 0.10 küçük etki (small effect).** Bu durumda etki toplam varyansın %1'ini açıklar.

**r = 0.30 orta etki (medium effect).** Bu durumda etki toplam varyansın %9'unu açıklar.

**r = 0.50 büyük etki (large effect).** Bu durumda etki toplam varyansın %25'ini açıklar.

r etki değeri aşağıdaki formülle hesaplanır (Rosenthal ve Rosnow, 2008):

$$r = \sqrt{\frac{t^2}{t^2 + \text{Serbestlik Derecesi}}}$$

<sup>4</sup> **r** etki büyüklüğü doğrusal bir ilişkinin değeri/katsayısı olarak işlev görürken, Eta kare ( $\eta^2$ ) etki büyüklüğü ise her türlü ilişkinin bir değeri/katsayısı olarak işlev görmektedir (Rosenthal ve Rosnow, 2008). Rosenthal, Rosnow ve Field **r** etki değerini etki büyüklüğü katsayısı olarak kullanıp önerirken, Pallant Eta kare ( $\eta^2$ ) etki büyüklüğü katsayısını kullanıp, önermektedir. Eta kare ( $\eta^2$ ) ve **r** değerleri en yaygın kullanılan iki etki büyüklüğü değerleridir.

**Önemli not:** Bu uygulamada değişken verilerinin normal dağılım özelliği gösterdiği varsayılmıştır. Siz kendi analizlerinizde öncelikli olarak normal dağılıma uygunluğu sınamalısınız.

### SPSS Uygulaması:

- 1- **Analyze, Compare Means, One-Sample T Test** seçilir.
- 2- **One-Sample T Test** penceresinde sol taraftaki değişkenler listesinden **değişken** seçilerek sağ taraftaki **Test Variable(s)** penceresine aktarılır.  
**Test Value** kısmına -bu değişken için- **evrenin ortalama değeri** yazılır.  
**OK** tıklanır.

### SPSS çıktısı:

#### T-Test

##### One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
aylık gelir	135	2982,9630	1863,59630	160,39283

**N:** Örneklem sayısı

**Mean:** Ortalama

**Std. Deviation:** Standart Sapma

##### One-Sample Test

Test Value = 2500

				95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	
				Lower	Upper
aylık gelir	3,011	134	,003	482,96296	165,7339 800,1920

**Test Value:** Evrenin ortalaması (ilgili değişkenin evren ortalaması değeri)

**t:** t testinin değeri

**df:** Serbestlik derecesi

**Sig. (2-tailed):** t testinin p değeri

**Mean Difference:** Ortalamalar arası fark

**95% Confidence Interval of the Difference Lower-Upper:** %95 Güven Aralığında ortalamalar arası farkın alt ve üst sınırı değerleri.



**Tablo: Grubun ve Evrenin aylık gelir ortalamalarının karşılaştırılması**

Aylık gelir	n	Ortalama	Standart sapma	t	p
Test değeri=2.500	135	2.982,96	1.863,59	3,011	,003

Araştırma grubundaki katılımcıların aylık gelir ortalaması 2.983 TL±1.864 TL'dir ve içinden çekilen Evrenin aylık gelir ortalaması olan 2.500 TL ile Örneklemin aylık gelir ortalaması arasında gözlenen bu fark istatistiksel olarak anlamlı bir farklılık olarak tespit edilmiştir ( $p<0,01$ ). Tespit edilen bu farkın etki büyüklüğü  $d=0,26$  olduğu, farkın küçük bir etkiye sahip olduğunu göstermektedir. (r etki değeri üzerinden yazmak istersek: Tespit edilen bu farkın etki büyüklüğü  $r=0,25$  olduğu, farkın küçük bir etkiye sahip olduğu ve toplam varyansın %6'sını açıkladığını göstermektedir.)

**Not:** Gelir ortalaması gibi bir değişkenin Örneklem ortalamasının Evrenden farklılaşmaması beklenir. Bu veriler kurgu veriler olduğu ve uygulamaları somutlaştırmak için verildiğinden böyle bir sonuç ortaya çıkmıştır. Aylık gelir ortalaması gibi bir değişkenin Örneklem ortalaması ile Evren ortalaması arasında farklılıklar şu koşullarda beklenebilir: Evrenin (Kadın+Erkekler) aylık gelir ortalaması ile Evrenin sadece bir alt kategorisinden oluşan Örneklemin (Kadınlar) aylık gelir ortalaması ile karşılaştırıldığı vb. durumlarda olası sonuçlardır.

#### Etki büyüklüğü:

##### Cohen d istatistiği (d):

$$d = \frac{\text{Örneklemin Ortalaması} - \text{Evrenin ortalaması}}{\text{Örneklemin standart sapması}} = \frac{2.982,96 - 2.500}{1.863,59} = 0,26$$

$$d = \frac{t}{\sqrt{n}} = \frac{3,011}{\sqrt{135}} = 0,26$$

**d= 0,26** (küçük etki)

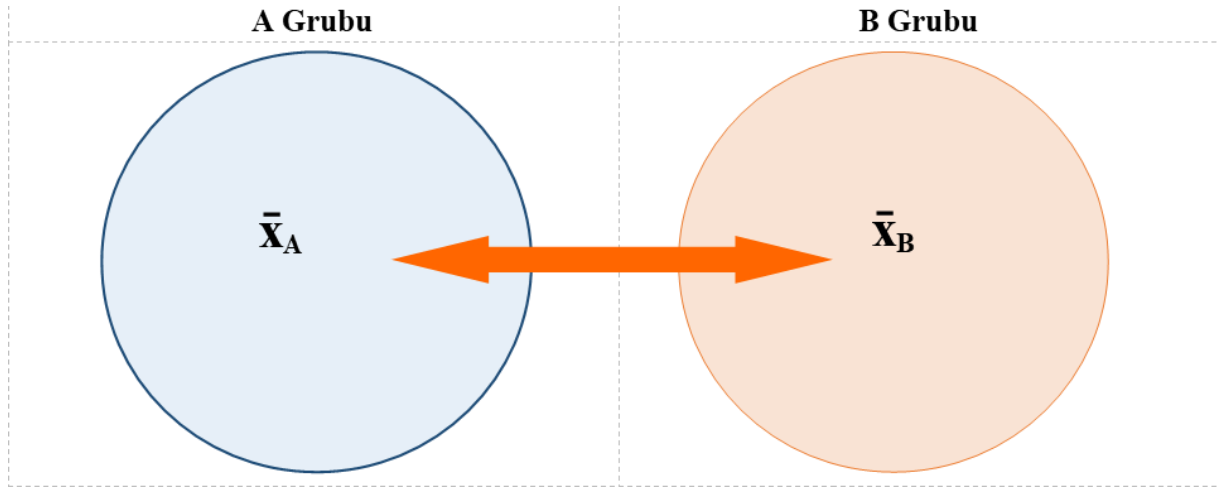
##### r etki değeri:

$$r = \sqrt{\frac{t^2}{t^2 + \text{Serbestlik Derecesi}}} = \sqrt{\frac{3,011^2}{3,011^2 + 134}} = 0,25$$

**r = 0,25** (küçük etki)

Bu durumda etki toplam varyansın %6'sını açıklar ( $r^2 = 0,25^2 = 0,0625 = \%6$ ).

## 2.2. Bağımsız Gruplarda t Testi <sup>5</sup>



İki bağımsız grubun bağımlı bir nicel değişkeninin ortalamaları arasında anlamlı bir fark olup olmadığını sınamak için parametrik bir test olan Bağımsız Gruplarda t Testi (Independent Samples T-Test) yapılır. Bu testin yapılabilmesinin varsayımları; Örneklemin Evrenden tesadüfi/rastgele bir şekilde seçilmesine, ortalamaları kıyaslanacak en az aralık ölçeğinde olan değişken verilerinin her iki grupta da normal dağılım özelliği göstermesi, iki grubun birbirinden bağımsız olması ve grupların varyanslarının eşit olması beklenir (Can, 2019; Büyüköztürk, 2010; Field, 2009).

Bu varsayımlardan; 1. Söz konusu değişken verilerinin her iki gruptaki dağılımının normal dağılım özelliği göstermesi varsayımı, örneklemin evrenden tesadüfi bir şekilde seçilmesi ve her bir grupta en az 10 örneklem sayısı ( $n_1 \geq 10$  ve  $n_2 \geq 10$ ) olması koşulu ile esnetilebilecek bir varsayımdır. Bu koşulla bağımlı nicel değişken verilerinin her bir grupta değil de tüm örnekleme normal dağılımı sağlaması yeterlidir. 2. Nicel değişken verilerinin iki gruptaki varyanslarının eşit olması zorunluluğu da yoktur. Bu test varyansların eşitliğini “Levene's Test for Equality of Variances” ile sınar ve varyansların eşit olması ve olmaması durumuna göre iki adet test değeri (t ve p) üretir.

### Test neyi yapar:

İki grubun bir değişkeninin ortalamalarını karşılaştırır (IBM SPSS, 2017). İki bağımsız grubun bir nicel (en az aralık ölçeğinde) değişkeninin ortalamalarını karşılaştır.

### Varsayımları:

Örnekleme Evrenden rastgele seçilmiştir.

Gruplar birbirinden bağımsızdır (iki bağımsız grup).

Bağımlı değişken en az aralık ölçeğinde olan nicel bir değişkendir.

Bağımlı değişkene ilişkin ölçümlerin dağılımı her iki grupta da normal dağılım özelliği gösterir.

Grupların varyansları eşittir.

<sup>5</sup> Independent Samples T-Test

**Non-Parametrik testi: Mann Whitney U Testi****Hipotezleri:**

$H_0$  : İki grubun ortalaması arasında fark yoktur. (sınanan hipotez)

$H_1$ : İki grubun ortalaması arasında fark vardır.

**Test sonucu:**

$p > 0,05$  ise  $H_0$  hipotezi kabul edilir: İki grubun ortalaması arasında fark yoktur.

$p < 0,05$  ise  $H_0$  hipotezi ret edilir: İki grubun ortalaması arasında fark vardır.

**Etki Büyüklüğü:**

Yapılan hipotez testi bize bağımsız iki grubun ortalamaları arasında istatistiksel olarak anlamlı bir fark olup olmadığını söyler, ancak fark var ise bu farkın etkisi ya da büyüklüğü konusunda bilgi vermez. Etki büyüklüğü bize etkinin önemi konusunda objektif bir ölçüm sağladığı için faydalıdır (Field, 2009). Etki büyüklüğü sonuçların yorumlanmasında dikkate alınması gereken istatistiksel bir değerdir. Hipotez testlerinde etki büyüklüğü değerlerinin de hesaplanarak yorumlanması sonuçların anlaşılabilirliğini arttıracaktır (Büyüköztürk, 2010). Hipotez testinde fark bulunduğu durumda etki büyüklüğüne bakılır. Hipotez testleri için hesaplanan farklı etki büyüklüğü değerleri söz konusudur.<sup>6</sup> Bağımsız Gruplarda t Testi için kullanılan etki büyüklüğü değerleri:

**Eta kare ( $\eta^2$ ):** Eta kare korelasyon katsayısı. Eta kare bağımsız değişkenin bağımlı değişken üzerinde ne derecede etkili olduğunu gösterir. Eta kare bağımsız değişkenin (ya da faktörün) bağımlı değişkendeki toplam varyansın (değişimin) ne kadarını açıkladığını gösterir ve 0 ile 1 arasında bir değer alır (Pallant, 2017; Büyüköztürk, 2010). 0,01 küçük, 0,06 orta ve 0,14 düzeyindeki eta kare ise büyük (large) etki olarak yorumlanır (Cohen, 1992). Eta kare değerinin kendisi bağımsız değişkenin bağımlı değişkendeki varyansın ne kadarını (%) açıkladığını ifade eder (Rosenthal ve Rosnow, 2008). Eta kare 0,01 ise varyansın %1'inin bağımsız değişken tarafından açıklandığını, Eta kare 0,06 ise varyansın %6'sının açıklandığını ve Eta kare 0,14 ise varyansın %14'ünün bağımsız değişken tarafından açıklandığı anlamına gelir.

$$\text{Eta kare } (\eta^2) = \frac{t^2}{t^2 + (n_1 + n_2 - 2)}$$

**r etki değeri:** Pearson'un korelasyon katsayısı olan **r** bir etki değeri katsayısıdır. r değeri 0 (etki yok) ile 1 (mükemmel etki/perfect effect) arasında bir değer alır. r değeri işaretinden bağımsız olarak değerlendirilir. r değerinin 0,1 olması küçük etki, 0,3 olması orta ve 0,5 olması büyük (large) etki olarak değerlendirilir (Field, 2009). r katsayısının karesi ( $r^2$ ) toplam varyansın ne

<sup>6</sup> **r** etki büyüklüğü doğrusal bir ilişkinin değeri/katsayısı olarak işlev görürken, Eta kare ( $\eta^2$ ) etki büyüklüğü ise her türlü ilişkinin bir değeri/katsayısı olarak işlev görmektedir (Rosenthal ve Rosnow, 2008). Rosenthal, Rosnow ve Field **r** etki değerini etki büyüklüğü katsayısı olarak kullanıp önerirken, Pallant Eta kare ( $\eta^2$ ) etki büyüklüğü katsayısını kullanıp, önermektedir. Eta kare ( $\eta^2$ ) ve **r** değerleri en yaygın kullanılan iki etki büyüklüğü değeridir.

kadarını açıkladığını ifade eder.  $r^2$  değeri üzerinden bağımsız değişkenin bağımlı değişken üzerindeki değişimin ne kadarını açıkladığı bilgisine de ulaşırız.

Cohen (1992) r katsayısının etki düzeylerinin yorumlanmasına ilişkin -yaygın olarak kullanılan- şu sınıflandırmayı yapmıştır:

**r = 0.10 küçük etki (small effect).** Bu durumda etki toplam varyansın %1'ini açıklar.

**r = 0.30 orta etki (medium effect).** Bu durumda etki toplam varyansın %9'unu açıklar.

**r = 0.50 büyük etki (large effect).** Bu durumda etki toplam varyansın %25'ini açıklar.

r etki değeri aşağıdaki formülle hesaplanır (Rosenthal ve Rosnow, 2008):

$$r = \sqrt{\frac{t^2}{t^2 + \text{Serbestlik Derecesi}}}$$

**Cohen d istatistiği (d):** Cohen'in d değeri karşılaştırılan ortalamaların birbirlerinden kaç standart sapma uzaklaştığını yorumlama imkanı verir (Büyüköztürk, 2010). Cohen'in etki büyüklüğü (d istatistiği) işaretinden bağımsız olarak değerlendirilir ve her değeri alabilir. d değerinin 0,2 olması küçük etki, 0,5 orta, 0,8 olması büyük (large) etki olarak değerlendirilir (Cohen, 1992; Green ve Salkind, 2014). Cohen d istatistiği aşağıdaki formülle hesaplanabilir:

$$d = t \times \sqrt{\frac{n_1 + n_2}{n_1 \times n_2}}$$

**Bu testin uygulamasına örnek 1:** Kadınların ve erkeklerin depresyon puan ortalamalarını karşılaştırmak istediğimizde depresyon verileri normal dağılım özelliği gösteriyor ise (ve kadınların örneklem sayısı  $\geq 10$  ve erkeklerin örneklem sayısı  $\geq 10$  ise) Bağımsız Gruplarda t Testi yapılır.

**Bu testin uygulamasına örnek 2:** Kadınların ve erkeklerin aylık gelir ortalamalarını karşılaştırmak istediğimizde aylık gelir verileri normal dağılım özelliği gösteriyor ise (ve kadınların örneklem sayısı  $\geq 10$  ve erkeklerin örneklem sayısı  $\geq 10$  ise) Bağımsız Gruplarda t Testi yapılır.

**Önemli not:** Bu uygulamada karşılaştırılan bağımlı değişken verilerinin normal dağılım özelliği gösterdiği varsayılmıştır. Siz kendi analizlerinizde öncelikli olarak normal dağılıma uygunluğu sınamalısınız.

**SPSS Uygulaması:**

- 1- **Analyze, Compare Means, Independent Samples T Test** seçilir.
- 2- **Independent Samples T Test** penceresinde sol taraftaki değişkenler listesinden **bağımlı nicel değişken** seçilerek sağ taraftaki **Test Variable(s)** penceresine aktarılır. Sol taraftaki değişkenler listesinden **bağımsız kategorik değişken** seçilerek sağ taraftaki **Grouping Variable** kısmına aktarılır. **Define Groups** tıklanır ve **bağımsız değişken kategorilerinin ne olarak kodlandığı** yazılır ve **Continue** tıklanır. **Independent Samples T Test** penceresinde **OK** tıklanır.

**SPSS çıktısı:****T-Test**

Group Statistics					
	cinsiyet	N	Mean	Std. Deviation	Std. Error Mean
aylık gelir	kadın	72	2211,1111	1674,83912	197,38168
	erkek	63	3865,0794	1676,70556	211,24504

**N:** Örneklem sayısı. **Mean:** Ortalama. **Std. Deviation:** Standart Sapma

Independent Samples Test										
			Levene's Test for Equality of Variances		t-test for Equality of Means					95% Confidence Interval of the Difference
			F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	
aylık gelir	Equal variances assumed		,909	,342	-5,721	133	,000	1653,96	289,08	2225,77 1082,16
	Equal variances not assumed				-5,721	130,59	,000	1653,96	289,10	2225,91 1082,02

**Levene's Test for Equality of Variances:** Grupların varyanslarının eşitliğinin sınandığı testtir. Bu testin sonucuna göre t testinin ilgili sonuçları alınır (ilgili satırdan).

**F:** Levene testinin değeri

Levene's Test for Equality of Variances Sig. değeri  $p > 0,05$  iki grubun varyansları eşittir ve üst satırdaki test değerleri (t ve Sig. 2-tailed) kullanılır.

Levene's Test for Equality of Variances Sig. değeri  $p < 0,05$  iki grubun varyansları eşit değildir ve alt satırdaki test değerleri (t ve Sig. 2-tailed) kullanılır.

**t:** t testinin değeri

**df:** Serbestlik derecesi

**Sig. (2-tailed):** t testine ait p değeri (gruplar arasında fark bu değere göre değerlendirilir).

**95% Confidence Interval of the Difference Lower-Upper:** %95 Güven Aralığında ortalamalar arası farkın alt ve üst sınırı değerleri.

**Tablo: Araştırma grubunda kadın ve erkeklerin aylık gelir ortalamalarının karşılaştırılması**

Değişken	n	Ortalama	Standart sapma	t	p
Kadın	72	2.211,11	1.674,84	-5,721	,000
Erkek	63	3.865,08	1.676,71		

Araştırma gurubundaki kadınların aylık gelir ortalaması 2.211 TL±1.675 TL iken erkeklerin 3.865 TL±1.677 TL'dir. Kadın ve erkeklerin aylık gelir ortalamaları arasında istatistiksel olarak anlamlı bir farklılık tespit edilmiştir ( $p<0,01$ ). Tespit edilen bu farkın etki büyüklüğü Eta kare ( $\eta^2$ )= 0,19 olduğu, farkın büyük bir etkiye sahip olduğu ve toplam varyansın %19'unun bağımsız değişken (cinsiyet) tarafından açıklandığını göstermektedir. (Diğer etki değerleri ile açıklamaları: Tespit edilen bu farkın etki büyüklüğü  $r= 0,44$  olduğu, farkın büyük bir etkiye sahip olduğu ve toplam varyansın %19'unun bağımsız değişken (cinsiyet) tarafından açıklandığını göstermektedir. Tespit edilen bu farkın etki büyüklüğü  $d=0,98$  olduğu, farkın büyük bir etkiye sahip olduğunu göstermektedir.)

**Etki büyüklüğü:****Eta kare ( $\eta^2$ ):**

$$\text{Eta kare} = \frac{t^2}{t^2 + (n_1 + n_2 - 2)} = \frac{-5,721^2}{-5,721^2 + (72 + 63 - 2)} = 0,19$$

**Eta kare ( $\eta^2$ )=0,19** (büyük etki). Bu durumda etki toplam varyansın %19'unu açıklar.

**r etki değeri:**

$$r = \sqrt{\frac{t^2}{t^2 + \text{Serbestlik Derecesi}}} = \sqrt{\frac{-5,721^2}{-5,721^2 + 133}} = 0,44$$

**r=0,44** (büyük etki)

Bu durumda etki toplam varyansın %19'unu açıklar ( $r^2=0,44^2=0,19$ ).

**Cohen d istatistiği (d):**

$$d = t \times \sqrt{\frac{n_1 + n_2}{n_1 \times n_2}} = -5,721 \times \sqrt{\frac{72 + 63}{72 \times 63}} = -0,98$$

**d= 0,98** (büyük etki)

### 2.3. Mann Whitney U Testi

Bağımsız Gruplarda t Testinin Non-Parametrik testidir. Bağımsız Gruplarda t Testinin varsayımlarının sağlanamaması durumunda, iki bağımsız grubun puanları arasında istatistiksel olarak anlamlı bir farklılığın olup olmadığı bu test ile sınanabilir.

Mann Whitney U Testi, en az sıralama ölçeğindeki verileri grubuna bakılmaksızın -bütün verileri- küçükten büyüğe doğru sıralar ve 1'den başlayarak sıra sayısı verilerek sıra sayısına dönüştürür. Bu en az aralık ölçeğindeki verilerin basitçe sıralama ölçeğine dönüştürülmesi işlemidir. Ardından veriler karşılaştırılacak iki gruba ayrılır ve grupların sıra sayılarının ortalaması (sıra ortalaması/mean rank) hesaplanır (Can,2019; Büyüköztürk, 2010). Mann Whitney U Testi, bağımsız iki grubun sıra ortalamasını (mean rank) karşılaştırır ve bunun sonucunda iki grubun puanları arasında istatistiksel olarak anlamlı bir fark olup olmadığı sonucuna ulaşır.

**Mann Whitney U Testinin sıra değerlerini üretme ve işlem süreci örneği:**

Veri seti			Mann Whitney U testinin işlem süreci			
Anket No	Grup	Puan	Anket No	Grup	Puan	Sıra değeri
1	A	63	6	A	62	1
2	A	65	1	A	63	2
3	B	87	2	A	65	3
4	B	89	12	B	67	4
5	A	95	8	B	73	5
6	A	62	7	B	75	6
7	B	75	13	A	76	7
8	B	73	9	B	77	8
9	B	77	11	A	81	9
10	B	83	10	B	83	10
11	A	81	3	B	87	11
12	B	67	4	B	89	12
13	A	75	5	A	95	13

	n	Sıra Ortalaması (Mean Rank)	Sıralar Toplamı (Sum of Ranks)
B Grubu	7	8,00	56
A Grubu	6	5,83	35

Mann Whitney U Testi iki grubun medyanlarını karşılaştırmaz.<sup>7</sup> Farklılığın hesaplanmasında medyanların bir işlevi yoktur, dolayısı ile bu testin sonucunda ulaşılan grupların birbirinden farklılaşmasını ya da farklılaşmamasını, medyanlar üzerinden açıklanması ve yorumlanması yanlıştır! Özet değeri olarak medyan verilebilir ancak farklılaşmanın medyan değerleri üzerinden açıklanması ve yorumlanması yanlıştır.

<sup>7</sup> Mann Whitney U testi iki grubun medyanlarını değil sıra ortalamalarını karşılaştırır (Bryman ve Cramer, 1999:136; Carver ve Nash, 2012:250; Corder ve Foreman, 2009:58; Landau ve Everitt, 2004:41-53; Connolly, 2007:195).

**Test neyi yapar:**

İki bağımsız grubun sıra ortalamasını (mean rank) karşılaştırır ve bunun sonucunda iki grubun puanları arasında istatistiksel olarak anlamlı bir farklılık olup olmadığı sınırlar.

**Varsayımları:**

Gruplar birbirinden bağımsızdır (iki bağımsız grup).

Bağımlı değişken en az sıralama ölçeğinde bir değişkendir.

**Parametrik testi:** Bağımsız Gruplarda t Testi

**Hipotezleri:**

$H_0$ : İki grubun puanları arasında fark yoktur. (sınanan hipotez)

$H_1$ : İki grubun puanları arasında fark vardır.

**Test sonucu:**

$p > 0,05$  ise  $H_0$  hipotezi kabul edilir: İki grubun puanları arasında fark yoktur.

$p < 0,05$  ise  $H_0$  hipotezi ret edilir: İki grubun puanları arasında fark vardır.

**Etki Büyüklüğü:**

Yapılan hipotez testi bize iki grubun sıra ortalaması arasında istatistiksel olarak anlamlı bir fark olup olmadığını söyler, ancak fark var ise bu farkın etkisi ya da büyüklüğü konusunda bilgi vermez. Etki büyüklüğü bize etkinin önemi konusunda objektif bir ölçüm sağladığı için faydalıdır (Field, 2009). Etki büyüklüğü sonuçların yorumlanmasında dikkate alınması gereken istatistiksel bir değerdir. Hipotez testlerinde etki büyüklüğü değerlerinin de hesaplanarak yorumlanması sonuçların anlaşılabilirliğini arttıracaktır (Büyüköztürk, 2010). Hipotez testinde fark bulunduğu durumda etki büyüklüğüne bakılır. Hipotez testleri için hesaplanan farklı etki büyüklüğü değerleri söz konusudur. Mann Whitney U Testi için kullanılan etki büyüklüğü değeri:

**r etki değeri:** Pearson'un korelasyon katsayısı olan **r** bir etki değeri katsayısıdır. r değeri 0 (etki yok) ile 1 (mükemmel etki/perfect effect) arasında bir değer alır. r değeri işaretinden bağımsız olarak değerlendirilir. r değerinin 0,1 olması küçük etki, 0,3 olması orta ve 0,5 olması büyük (large) etki olarak değerlendirilir (Field, 2009). r katsayısının karesi ( $r^2$ ) toplam varyansın ne kadarını açıkladığını ifade eder.  $r^2$  değeri üzerinden bağımsız değişkenin bağımlı değişken üzerindeki değişimin ne kadarını açıkladığı bilgisine de ulaşırız.

Cohen (1992) r katsayısının etki düzeylerinin yorumlanmasına ilişkin -yaygın olarak kullanılan- şu sınıflandırmayı yapmıştır:

**r = 0.10 küçük etki (small effect).** Bu durumda etki toplam varyansın %1'ini açıklar.

**r = 0.30 orta etki (medium effect).** Bu durumda etki toplam varyansın %9'unu açıklar.

**r = 0.50 büyük etki (large effect).** Bu durumda etki toplam varyansın %25'ini açıklar.



r etki değeri aşağıdaki formülle hesaplanır (Corder ve Foreman, 2009):

$$r = \frac{Z}{\sqrt{n}}$$

Z değeri SPSS analiz çıktısındaki Z (eski versiyonlarda Standardized Test Statistic) değeridir.

**Bu testin uygulamasına örnek:** Kadınların ve erkeklerin aylık gelir ortalamalarını karşılaştırmak istediğimiz de aylık gelir verileri normal dağılım özelliği göstermiyor ise Mann Whitney U Testi ile farklılık sınanabilir.

#### SPSS Uygulaması:

- 1- **Analyze, Nonparametric Tests, Legacy Dialogs, 2 Independent Samples** seçilir.
- 2- **Two Independent Samples Test** penceresinde sol taraftaki değişken listesinden **bağımlı nicel değişken Test Variable List'e** taşınır.  
Sol taraftaki değişkenler listesinden **bağımsız kategorik değişken** seçilerek sağ taraftaki **Grouping Variable** kısmına aktarılır. **Define Groups** tıklanır ve **bağımsız değişken kategorilerinin ne olarak kodlandığı** yazılır ve **Continue** tıklanır.  
**Two Independent Samples Test** penceresinde Test Type **Mann-Whitney U** testi otomatik seçilidir.  
**OK** Tıklanır.

#### SPSS çıktısı:

#### NPar Tests

#### Mann-Whitney Test

		Ranks		
	cinsiyet	N	Mean Rank	Sum of Ranks
aylık gelir	kadın	72	46,42	3342,00
	erkek	63	92,67	5838,00
	Total	135		

**N:** Örneklem sayısı. **Mean rank:** Sıra ortalaması. **Sum of Ranks:** Sıra değerlerinin toplamı.

#### Test Statistics<sup>a</sup>

	aylık gelir
Mann-Whitney U	714,000
Wilcoxon W	3342,000
Z	-6,872
Asymp. Sig. (2-tailed)	,000

a. Grouping Variable: cinsiyet

**Mann-Whitney U:** Mann Whitney U testinin değeri. **Asymp. Sig. (2-tailed):** testin p değeri.

**Tablo: Araştırma grubunda kadın ve erkeklerin aylık gelirlerinin karşılaştırılması**

Değişken	n	Sıra Ortalaması	U	p
Kadın	72	46,42	714,000	,000
Erkek	63	92,67		

Araştırma grubunda kadınların aylık gelir sıra ortalaması 46 iken erkeklerin 93'tür. Kadın ve erkeklerin aylık gelir sıra ortalamaları arasında istatistiksel olarak anlamlı bir farklılık tespit edilmiştir ( $p < 0,01$ ). Tespit edilen bu farkın etki büyüklüğünün  $r = 0,59$  olduğu, farkın büyük bir etkiye sahip olduğu ve toplam varyansın %35'inin bağımsız değişken (cinsiyet) tarafından açıklandığını göstermektedir.

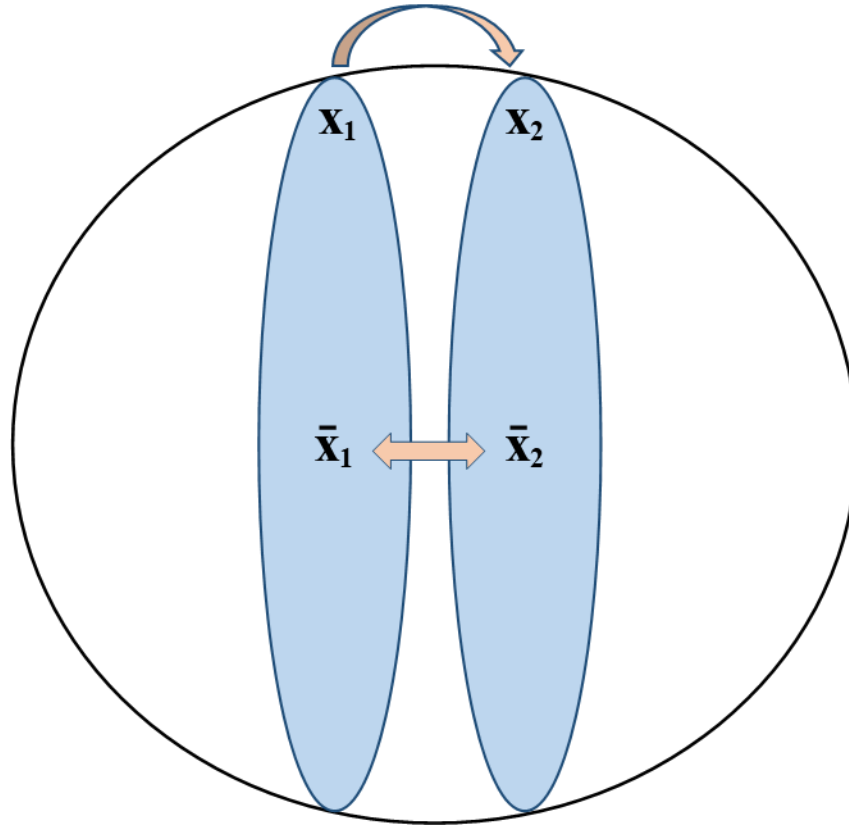
**Etki büyüklüğü:**

$$r = \frac{Z}{\sqrt{n}} = \frac{-6,872}{\sqrt{135}} = -0,59$$

**r = 0,59** (büyük etki)

Bu durumda etki toplam varyansın %35'ini açıklar ( $r^2 = 0,59^2 = 0,35$ ).

## 2.4. Bağımlı Gruplarda t Testi <sup>8</sup>



Bir grubun tekrarlayan iki ölçüm ortalaması arasında istatistiksel olarak anlamlı bir fark olup olmadığını sınamak için parametrik bir test olan Bağımlı Gruplarda t Testi (Paired Samples T-Test) yapılır. Bu testin yapılabilmesinin varsayımları; Örneklemin Evrenden tesadüfî bir şekilde seçilmesi, ortalamaları kıyaslanacak en az aralık ölçeğinde olan tekrarlayan ölçüm değerlerinin fark puanlarının normal dağılım özelliği göstermesi beklenir (Can, 2019; Field, 2009; Pallant, 2017).

Bu testin güçlü bir test olması nedeni ile Örneklem büyüklüğünün  $n \geq 30$  olması koşulu ile fark puanlarının normal dağılmamasının test için ciddi sorun yaratmayacağı dair liberal iddialar mevcuttur (Pallant, 2017; Green ve Salkind, 2014).

### Test neyi yapar:

Tek bir grubun iki değişkenin ortalamalarını karşılaştırır. Bu test aynı zamanda eşleştirilmiş çiftler veya vaka kontrol çalışması tasarımları için de geçerlidir (IBM SPSS, 2017). Bir grubun en az aralık ölçeğindeki aynı değişkenine ait tekrarlayan iki ölçüm ortalamasını karşılaştırır. Örnek: ön test, son test puanlarının karşılaştırmalarında bu test kullanılır.

<sup>8</sup> Paired Samples T-Test

**Varsayımları:**

Örneklem Evrenden rastgele seçilmiştir.

Bağımlı değişken en az aralık ölçeğinde bir değişkendir.

Tekrarlayan iki ölçüm değeri aynı gruba aittir.

Tekrarlayan iki ölçüm değerinin fark puanlarının normal dağılım özelliği göstermesi gerekir.

**Non-Parametrik testi:** Wilcoxon İşaretli Sıralar Testi

**Hipotezleri:**

$H_0$ : İki ölçüm ortalaması arasında fark yoktur. (sınanan hipotez)

$H_1$ : İki ölçüm ortalaması arasında fark vardır.

**Test sonucu:**

$p > 0,05$  ise  $H_0$  hipotezi kabul edilir: İki ölçüm ortalaması arasında fark yoktur.

$p < 0,05$  ise  $H_0$  hipotezi ret edilir: İki ölçüm ortalaması arasında fark vardır.

**Etki Büyüklüğü:**

Yapılan hipotez testi bize tekrarlayan iki ölçümün ortalamaları arasında istatistiksel olarak anlamlı bir fark olup olmadığını söyler, ancak fark var ise bu farkın etkisi ya da büyüklüğü konusunda bilgi vermez. Etki büyüklüğü bize etkinin önemi konusunda objektif bir ölçüm sağladığı için faydalıdır (Field, 2009). Etki büyüklüğü sonuçların yorumlanmasında dikkate alınması gereken istatistiksel bir değerdir. Hipotez testlerinde etki büyüklüğü değerlerinin de hesaplanarak yorumlanması sonuçların anlaşılabilirliğini arttıracaktır (Büyüköztürk, 2010). Hipotez testinde fark bulunduğu durumda etki büyüklüğüne bakılır. Hipotez testleri için hesaplanan farklı etki büyüklüğü değerleri söz konusudur.<sup>9</sup> Bağımlı Gruplarda t Testi için kullanılan etki büyüklüğü değerleri:

**Eta kare ( $\eta^2$ ):** Eta kare korelasyon katsayısı. Eta kare bağımsız değişkenin ya da faktörün bağımlı değişken üzerinde ne derecede etkili olduğunu gösterir. Eta kare bağımsız değişkenin (ya da faktörün) bağımlı değişkendeki toplam varyansın (değişimin) ne kadarını açıkladığını gösterir ve 0 ile 1 arasında bir değer alır (Pallant, 2017; Büyüköztürk, 2010). 0,01 küçük, 0,06 orta ve 0,14 düzeyindeki eta kare ise büyük (large) etki olarak yorumlanır (Cohen, 1992). Eta kare değerinin kendisi bağımsız değişkenin bağımlı değişkendeki varyansın ne kadarını (%) açıkladığını ifade eder (Rosenthal ve Rosnow, 2008). Eta kare 0,01 ise varyansın %1'inin bağımsız değişken tarafından açıklandığını, Eta kare 0,06 ise varyansın %6'sının açıklandığını ve Eta kare 0,14 ise varyansın %14'ünün bağımsız değişken tarafından açıklandığı anlamına gelir.

<sup>9</sup>  $r$  etki büyüklüğü doğrusal bir ilişkinin değeri/katsayısı olarak işlev görürken, Eta kare ( $\eta^2$ ) etki büyüklüğü ise her türlü ilişkinin bir değeri/katsayısı olarak işlev görmektedir (Rosenthal ve Rosnow, 2008). Rosenthal, Rosnow ve Field  $r$  etki değerini etki büyüklüğü katsayısı olarak kullanıp önerirken, Pallant Eta kare ( $\eta^2$ ) etki büyüklüğü katsayısını kullanıp, önermektedir. Eta kare ( $\eta^2$ ) ve  $r$  değerleri en yaygın kullanılan iki etki büyüklüğü değeridir.

$$\text{Eta kare } (\eta^2) = \frac{t^2}{t^2 + (n-1)}$$

**r etki değeri:** Pearson'un korelasyon katsayısı olan **r** bir etki değeri katsayısıdır. **r** değeri 0 (etki yok) ile 1 (mükemmel etki/perfect effect) arasında bir değer alır. **r** değeri işaretinden bağımsız olarak değerlendirilir. **r** değerinin 0,1 olması küçük etki, 0,3 olması orta ve 0,5 olması büyük (large) etki olarak değerlendirilir (Field, 2009). **r** katsayısının karesi ( $r^2$ ) toplam varyansın ne kadarını açıkladığını ifade eder.  $r^2$  değeri üzerinden bağımsız değişkenin bağımlı değişken üzerindeki değişimin ne kadarını açıkladığı bilgisine de ulaşırız.

Cohen (1992) **r** katsayısının etki düzeylerinin yorumlanmasına ilişkin -yaygın olarak kullanılan- şu sınıflandırmayı yapmıştır:

**r = 0.10 küçük etki (small effect).** Bu durumda etki toplam varyansın %1'ini açıklar.

**r = 0.30 orta etki (medium effect).** Bu durumda etki toplam varyansın %9'unu açıklar.

**r = 0.50 büyük etki (large effect).** Bu durumda etki toplam varyansın %25'ini açıklar.

**r** etki değeri aşağıdaki formülle hesaplanır (Rosenthal ve Rosnow, 2008):

$$r = \sqrt{\frac{t^2}{t^2 + \text{Serbestlik Derecesi}}}$$

**Cohen d istatistiği (d):** Cohen'in **d** değeri karşılaştırılan ortalamaların birbirlerinden kaç standart sapma uzaklaştığını yorumlama imkanı verir (Büyüköztürk, 2010). Cohen'in etki büyüklüğü (**d** istatistiği) işaretinden bağımsız olarak değerlendirilir ve her değeri alabilir. **d** değerinin 0,2 olması küçük etki, 0,5 orta ve 0,8 olması ise büyük (large) etki olarak değerlendirilir (Cohen, 1992; Green ve Salkind, 2014). Cohen **d** istatistiği aşağıdaki formülle hesaplanabilir:

$$d = \frac{t}{\sqrt{n}}$$

**Bu testin uygulamasına örnek 1:** Bir grubun depresyona yönelik psikoterapi seansına başlamadan önceki ve sonraki depresyon puan ortalamalarını karşılaştırmak istediğimizde depresyon puan ölçümlerinin fark puanları normal dağılım özelliği gösteriyor ise Bağımlı Gruplarda **t** Testi (Paired Samples T-Test) yapılır.

**Bu testin uygulamasına örnek 2:** Bir grubun eğitimden öncesi ve sonrası **X** Bilgi puan ortalamalarını karşılaştırmak istediğimizde **X** Bilgi puan ölçümlerinin fark puanları normal dağılım özelliği gösteriyor ise Bağımlı Gruplarda **t** Testi (Paired Samples T-Test) yapılır.

**Önemli not:** Bu uygulamada karşılaştırılan tekrarlayan ölçümlerin fark puanlarının normal dağılım özelliği gösterdiği varsayılmıştır. Siz kendi analizlerinizde öncelikli olarak fark puanlarını hesaplayıp, ardından fark puanlarının normal dağılıma uygunluğunu sınamalısınız.

**SPSS Uygulaması:**

- 1- **Analyze, Compare Means, Paired Samples T Test** seçilir.
- 2- **Paired Samples T Test** penceresinde sol taraftaki değişkenler listesinden **nicel** değişkenler seçilerek sağ taraftaki **Paired Variable(s)** kutusuna –eşleştirilerek- aktarılır.  
**OK** tıklanır.

**SPSS çıktısı:****T-Test****Paired Samples Statistics**

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	X bilgi puanı eğitimden önce	56,5037	135	15,67243	1,34887
	X bilgi puanı eğitimden sonra	80,0815	135	10,18123	,87626

**Mean:** Ortalama. **N:** örneklem sayısı. **Std. Deviation:** Standart Sapma

**Paired Samples Correlations**

		N	Correlation	Sig.
Pair 1	X bilgi puanı eğitimden önce & X bilgi puanı eğitimden sonra	135	,831	,000

**Paired Samples Correlations:** Eşleştirme dizaynındaki iki ölçüm arasındaki ilişkiyi gösteren korelasyon analizi sonuçları. SPSS bu analizi otomatik olarak yapar. Biz bu tablo ile ilgilenmiyoruz.

**Paired Samples Test**

		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
X bilgi eğitimden önce -	X bilgi eğitimden sonra	-23,57778	9,16472	,78877	-25,13783	-22,01772	-29,892	134	,000

**Tablo: Araştırma grubunda Eğitimden Önce ve Sonra X bilgi puan ortalamalarının karşılaştırılması**

Ölçüm	n	Ortalama	Standart sapma	t	p
Eğitimden önce	135	56,50	15,67	-29,892	,000
Eğitimden sonra	135	80,08	10,18		

Araştırma grubunda eğitim öncesi X bilgi puanı ortalaması  $57 \pm 16$  iken eğitim sonrası  $80 \pm 10$ 'dur. Eğitim öncesi ve sonrası X bilgi puanı ortalamaları arasında istatistiksel olarak anlamlı bir farklılık tespit edilmiştir ( $p < 0,01$ ). Tespit edilen bu farkın etki büyüklüğü Eta kare ( $\eta^2$ )= 0,87 olduğu, farkın büyük bir etkiye sahip olduğu ve toplam varyansın %87'sinin faktör (eğitim) tarafından açıklandığını göstermektedir. (Diğer etki değerleri ile açıklamaları: Tespit edilen bu farkın etki büyüklüğü  $r = 0,93$  olduğu, farkın büyük bir etkiye sahip olduğu ve toplam varyansın %87'sinin faktör (eğitim) tarafından açıklandığını göstermektedir. Tespit edilen bu farkın etki büyüklüğü  $d = 2,57$  olduğu, farkın büyük bir etkiye sahip olduğunu göstermektedir.)

**Etki büyüklüğü:****Eta kare ( $\eta^2$ ):**

$$\text{Eta kare} = \frac{t^2}{t^2 + (n - 1)} = \frac{-29,892^2}{-29,892^2 + (135 - 1)} = 0,87$$

**Eta kare ( $\eta^2$ )=0,87** (büyük etki). Bu durumda etki toplam varyansın %87'sini açıklar.

**r etki değeri:**

$$r = \sqrt{\frac{t^2}{t^2 + \text{Serbestlik Derecesi}}} = \sqrt{\frac{-29,892^2}{-29,892^2 + 134}} = 0,93$$

**r=0,93** (büyük etki)

Bu durumda etki toplam varyansın %87'sini açıklar ( $r^2 = 0,93^2 = 0,87$ ).

**Cohen d istatistiği (d):**

$$d = \frac{t}{\sqrt{n}} = \frac{-29,892}{\sqrt{135}} = -2,57$$

**d= 2,57** (büyük etki)

## 2.5. Wilcoxon İşaretli Sıralar Testi <sup>10</sup>

Bağımlı Gruplarda t Testinin Non-Parametrik testidir. Bağımlı Gruplarda t Testinin varsayımlarının sağlanamaması durumunda bir grubun tekrarlayan iki ölçüm puanları arasında istatistiksel olarak anlamlı bir farklılığın olup olmadığı bu test ile sınanabilir.

Wilcoxon İşaretli Sıralar Testi, tekrarlayan ikinci ölçümünden birinci ölçümü çıkararak aradaki fark puanlarını hesaplar, fark olmayanlar (2. Ölçüm-1 ölçüm=0 olanlar) hariç tutularak, fark puanları mutlak değerine göre küçükten büyüğe doğru sıralanır ve 1’den başlayarak sıra sayısı verilerek, fark puanları sıralama ölçeğine dönüştürülür. Ardından, sıralama ölçeği puanlarını, bu puanın karşılığı olan fark puanlarının işaretine göre gruplar ve negatif işaretli sıralar ile pozitif işaretli sıraların (sıralama ölçeği puanlarının) ortalamalarını alır. Test edilen durum negatif sıra sayıları ile pozitif sıra sayılarının toplamı arasındaki farkın istatistiksel analizidir. Bu testin sonuç tablosundaki pozitif sıralar ikinci ölçümün büyük olduğu vakaları, negatif sıralar ise birinci ölçümün büyük olduğu vakaları gösterir (Can,2019; Büyüköztürk, 2010).

### Wilcoxon testinin sıra değerlerini üretme ve işlem süreci örneği:

Veri setimiz			Wilcoxon testinin işlem süreci		
Anket no	Ön test	Son test	Fark Puanı (Son test-İlk Test)	Sıra değeri (Fark puanı sıfır hariç)	İşareti
1	62	62	0	Hariç	
2	53	65	+12	5	+
3	48	47	-1	1	-
4	63	70	+7	3	+
5	65	76	+11	4	+
6	56	51	-5	2	-
7	73	88	+15	8	+
8	74	90	+16	9	+
9	58	72	+14	7	+
10	62	75	+13	6	+

	n	Sıra Ortalaması (Mean Rank)	Sıralar Toplamı (Sum of Ranks)
Negatif sıralar	2	1,50	3
Pozitif sıralar	7	6,00	42
Eşit	1		

Wilcoxon İşaretli Sıralar Testi, iki ölçümün medyanlarını karşılaştırmaz. Farklılığın hesaplanmasında medyanların bir işlevi yoktur, dolayısı ile bu testin sonucunda ulaşılan tekrarlayan ölçümlerin birbirinden farklılaşmasını ya da farklılaşmamasını, medyanlar üzerinden açıklanması ve yorumlanması yanlıştır!

<sup>10</sup> Wilcoxon Singed Ranks Test



**Test neyi yapar:**

Bir grubun aynı değişkenine ait tekrarlayan iki ölçüm puanının farklılığını, fark puanlarının sıra değerleri üzerinden karşılaştırır. Örnek: ön test, son test puanlarının karşılaştırmalarında bu test kullanılır.

**Varsayımları:**

Bağımlı değişken en az sıralama ölçeğinde bir değişkendir.

Tekrarlayan iki ölçüm değeri aynı gruba aittir.

**Parametrik testi:** Bağımlı Gruplarda t Testi**Hipotezleri:**

$H_0$ : İki ölçüm puanları arasında fark yoktur. (sınanan hipotez)

$H_1$ : İki ölçüm puanları arasında fark vardır.

**Test sonucu:**

$p > 0,05$  ise  $H_0$  hipotezi kabul edilir: İki ölçüm puanları arasında fark yoktur.

$p < 0,05$  ise  $H_0$  hipotezi ret edilir: İki ölçüm puanları arasında fark vardır.

**Etki Büyüklüğü:**

Yapılan hipotez testi bize tekrarlayan iki ölçüm puanları arasında istatistiksel olarak anlamlı bir fark olup olmadığını söyler, ancak fark var ise bu farkın etkisi ya da büyüklüğü konusunda bilgi vermez. Etki büyüklüğü bize etkinin önemi konusunda objektif bir ölçüm sağladığı için faydalıdır (Field, 2009). Etki büyüklüğü sonuçların yorumlanmasında dikkate alınması gereken istatistiksel bir değerdir. Hipotez testlerinde etki büyüklüğü değerlerinin de hesaplanarak yorumlanması sonuçların anlaşılabilirliğini arttıracaktır (Büyüköztürk, 2010). Hipotez testinde fark bulunduğu durumda etki büyüklüğüne bakılır. Hipotez testleri için hesaplanan farklı etki büyüklüğü değerleri söz konusudur. Wilcoxon İşaretli Sıralar Testi için kullanılan etki büyüklüğü değeri:

**r etki değeri:** Pearson'un korelasyon katsayısı olan **r** bir etki değeri katsayısıdır. r değeri 0 (etki yok) ile 1 (mükemmel etki/perfect effect) arasında bir değer alır. r değeri işaretinden bağımsız olarak değerlendirilir. r değerinin 0,1 olması küçük etki, 0,3 olması orta ve 0,5 olması büyük (large) etki olarak değerlendirilir (Field, 2009). r katsayısının karesi ( $r^2$ ) toplam varyansın ne kadarını açıkladığını ifade eder.  $r^2$  değeri üzerinden bağımsız değişkenin bağımlı değişken üzerindeki değişimin ne kadarını açıkladığı bilgisine de ulaşırız. Cohen (1992) r katsayısının etki düzeylerinin yorumlanmasına ilişkin -yaygın olarak kullanılan- şu sınıflandırmayı yapmıştır:

**r = 0.10 küçük etki (small effect).** Bu durumda etki toplam varyansın %1'ini açıklar.

**r = 0.30 orta etki (medium effect).** Bu durumda etki toplam varyansın %9'unu açıklar.

**r = 0.50 büyük etki (large effect).** Bu durumda etki toplam varyansın %25'ini açıklar.

r etki değeri aşağıdaki formülle hesaplanır (Connolly, 2007):

$$r = \frac{Z}{\sqrt{n}}$$

Z değeri SPSS analiz çıktısındaki Z değeridir.

**Bu testin uygulamasına örnek:** Bir grubun eğitimden öncesi ve sonrası X Bilgi puan ortalamalarını karşılaştırmak istediğimizde X Bilgi puan ölçümlerinin fark puanları normal dağılım özelliği göstermiyor ise Wilcoxon İşaretli Sıralar Testi (Paired Samples T-Test) ile farklılık sınanabilir.

#### **SPSS Uygulaması:**

- 1- Analyze, Nonparametric Tests, Legacy Dialogs, 2 Related Samples –seçilir-**
- 2- Two Related Samples Test** penceresinde sol taraftaki değişken listesinden tekrarlayan ölçüm değişkenleri **Test Pairs**'e eşleştirilerek **taşınır**.
- 3- (İstenirse Options** butonu tıklanıp açılan pencereden **Quartiles** seçilir)
- 4- Two Related Samples Test** penceresinde **Test Type Wilcoxon** testi otomatik seçilidir. **OK** Tıklanır.

## SPSS çıktısı:

## NPar Tests

## Descriptive Statistics

	N	Percentiles		
		25th	50th (Median)	75th
X bilgi puanı eğitimden önce	135	45,0000	55,0000	65,0000
X bilgi puanı eğitimden sonra	135	70,0000	80,0000	85,0000

**Descriptive Statistics:** Tanımlayıcı İstatistikler

**N:** örneklem sayısı. **25th:** %25. Değer. **50th (Median):** %50. Değer (Ortanca). **75th:** %75. Değer

## Wilcoxon Signed Ranks Test

## Ranks

		N	Mean Rank	Sum of Ranks
X bilgi puanı eğitimden sonra - X bilgi puanı eğitimden önce	Negative Ranks	0 <sup>a</sup>	,00	,00
	Positive Ranks	131 <sup>b</sup>	66,00	8646,00
	Ties	4 <sup>c</sup>		
	Total	135		

a. X bilgi puanı eğitimden sonra &lt; X bilgi puanı eğitimden önce

b. X bilgi puanı eğitimden sonra &gt; X bilgi puanı eğitimden önce

c. X bilgi puanı eğitimden sonra = X bilgi puanı eğitimden önce

**Negative Ranks:** Negatif sıralar (a eğitimden sonra < eğitimden önce)**Positive Ranks:** Pozitif sıralar (b eğitimden sonra > eğitimden önce)**Ties:** Eşit (c X bilgi puanı eğitimden sonra = X bilgi puanı eğitimden önce)**Mean Rank:** Sıra ortalaması**Sum of Ranks:** Sıra değerlerinin toplamıTest Statistics<sup>a</sup>

X bilgi puanı eğitimden sonra - X bilgi puanı eğitimden önce	
Z	-9,973 <sup>b</sup>
Asymp. Sig. (2-tailed)	,000

a. Wilcoxon Signed Ranks Test

b. Based on negative ranks.

**Z:** test değeri**Asymp. Sig. (2-tailed):** testin p değeri

**Tablo: Araştırma grubunda Eğitimden Önce ve Sonra X bilgi puanlarının karşılaştırılması**

X bilgi puanı	n	Sıra Ortalaması	Sıralar toplamı	Z	p
Negatif sıralar	0	,00	0,00	-9,973	0,000
Pozitif sıralar	131	66,00	8.646,00		
Eşit	4				

Araştırma grubunda eğitim öncesine göre eğitim sonrasında bilgi puanı düşen herhangi bir katılımcı yoktur, katılımcılardan 4'ünün eğitim öncesi ve sonrası bilgi puanlarının değişmediği ve eğitim sonrasında öncesine göre bilgi puanları artan 131 katılımcı vardır. Eğitim sonrası bilgi puanı artan 131 katılımcının artan (fark) bilgi puanlarının sıra ortalaması değeri 66 olarak tespit edilmiştir. Araştırma grubundaki katılımcıların eğitim öncesi ve sonrası X bilgi puanı sıra ortalamaları arasında istatistiksel olarak anlamlı bir farklılık tespit edilmiştir ( $p < 0,01$ ). Tespit edilen bu farkın etki büyüklüğünün  $r = 0,85$  olduğu, farkın büyük bir etkiye sahip olduğu ve toplam varyansın %72'sinin bağımsız faktör (eğitim) tarafından açıklandığını göstermektedir.

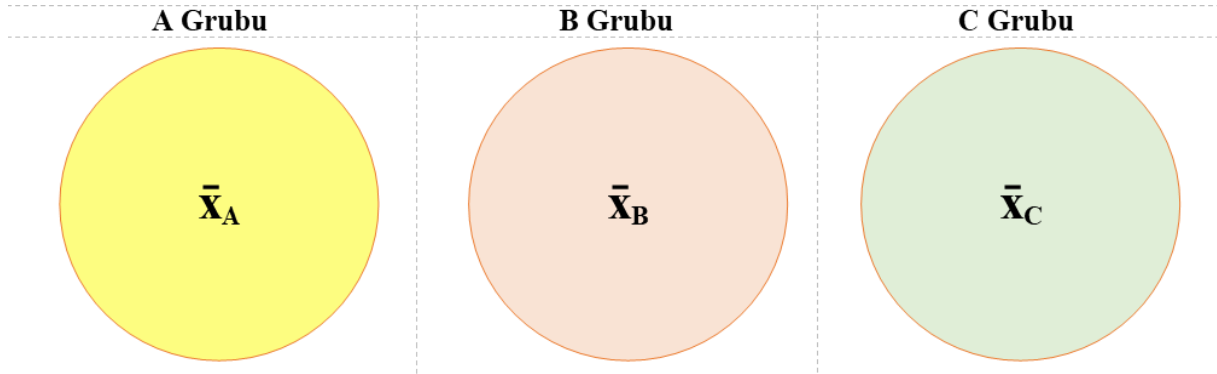
**Etki büyüklüğü:**

$$r = \frac{Z}{\sqrt{n}} = \frac{-9,973}{\sqrt{135}} = -0,85$$

**$r = 0,85$**  (büyük etki)

Bu durumda etki toplam varyansın %72'sini açıklar ( $r^2 = 0,85^2 = 0,72$ ).

## 2.6. Bağımsız Gruplar için Tek Yönlü Varyans Analizi <sup>11</sup>



Bağımsız Gruplar için Tek Yönlü Varyans Analizi, ikiden fazla bağımsız grubun ortalamaları arasında istatistiksel olarak anlamlı bir fark olup olmadığını sınavan parametrik bir testtir. Bu testin yapılabilmesinin varsayımları; Örneklemin Evrenden tesadüfî bir şekilde seçilmesi, ortalamaları kıyaslanacak en az aralık ölçeğinde olan değişken verilerinin dağılımının her grupta normal dağılım özelliği göstermesi, grupların birbirinden bağımsız olması ve grupların varyanslarının eşit olması beklenir (Can, 2019; Field 2009).

ANOVA testi büyük örneklerle çalışıldığında normalliğin ihlalden etkilenmez. ANOVA testi normalliğin ihlali durumunda her bir grubun örneklem sayısı 15'e düştüğü durumlarda da oldukça doğru p değerleri üretir (Green ve Salkind, 2014). Bu ANOVA'nın güçlü (robust) bir test olması nedeni ile testin varsayımlarını ihlal etmenin önemi olmadığı gibi yorumlansa da bu aynı zamanda durumun aşırı derecede basitleştirilmesidir (Field, 2009). Kanıtlar grup boyutlarının (grupların n sayıları) eşit olduğu durumlarda ANOVA'nın normallik ihlallerine karşı oldukça güçlü (robust) olduğunu; bununla birlikte, grup boyutları eşit olmadığında, ANOVA testinin sonuçlarının normallik ihlalden ciddi şekilde etkilendiği yönündedir (Wilcox, 2005).

ANOVA varyans homojenliği varsayımı ihlalleri açısından, örneklem boyutları (grupların n sayıları) eşit olduğunda oldukça güçlüdür (robust). Bununla birlikte, örneklem boyutları eşit olmadığında, ANOVA varyans homojenliği ihlallerine karşı güçlü (robust) değildir ya da bu durumda test sonuçları güvenilmezdir (untrustworthy) (Field, 2009; Green ve Salkind, 2014).

Testin uygulanmasına yönelik bu varsayımları kuramsal çerçevede değerlendirdiğimizde: Her ne kadar karşılaştırılacak en az aralık ölçeğinde olan bağımlı değişken verilerinin, her grupta normal dağılımı beklensede bağımlı değişken verilerinin tüm örneklem için normal dağılımı sağlaması ve her bir grupta en az 10 katılımcının olması koşulu ile bu test yapılabilir. Ayrıca grupların varyansların eşitliği varsayımı testin uygulanması ile bağlayıcı değildir (örneklem büyüklüklerinin eşit olmadığı durumlar hariç). Grupların varyansları eşit olmasa da test yapılabilir, bu varsayım Tek Yönlü ANOVA'da farklılık tespit edilmesi durumunda yapılacak post-hoc testlerin seçiminde bağlayıcı bir yönlendiricilik yapar.

<sup>11</sup> One-Way ANOVA

Bağımsız Gruplar için Tek Yönlü Varyans Analizi, grupların ortalamaları arasında istatistiksel olarak anlamlı bir fark olup olmadığını sınar. Bu testin sonucunda farklılık olduğu sonucuna ulaştığımızda, grupların ortalamaları arasında bir fark olduğu sonucuna ulaşırız. Ancak Tek Yönlü Varyans Analizi bize bu farkın hangi grupların birbirinden farklılaştığından kaynaklandığını vermez. Tek Yönlü Varyans Analizi sonucunda farklılık olduğu sonucuna ulaştığımızda, farkın kaynağını tespit etmek için post-hoc test yapmamız gerekir. Yapılan post-hoc test bize hangi grupların birbirinden farklılaştığını gösterecektir.

### Test neyi yapar:

Tek Yönlü ANOVA, tek faktörlü (bağımsız) değişken aracılığıyla nicel bağımlı bir değişken için tek yönlü varyans analizi yapar. Varyans analizi ikiden çok ortalamanın eşit olduğu hipotezini test eder. Bu test Bağımsız Gruplarda t Testinin bir uzantısıdır (extension) (IBM SPSS, 2017). Tek faktörlü ANOVA, üç ya da daha fazla bağımsız grup için en az aralık ölçeğinde olan bağımlı değişkenin puan ortalamaları arasında anlamlı bir farklılık olup olmadığını sınar. Farklılık tespit edilmesi durumunda yapılacak post-hoc testler ile farklılığın hangi grupların birbirinden farklılaşmasından kaynaklandığı tespit edilir (Pallant, 2017).

### Varsayımları:

Örnekleme Evrenden rastgele seçilmiştir.

Gruplar birbirinden bağımsızdır (3 ya da daha fazla bağımsız grup).

Bağımlı değişken en az aralık ölçeğinde olan bir değişkendir.

Bağımlı değişkene ilişkin ölçümlerin dağılımı her bir grupta normal dağılım özelliği gösterir.

Bağımlı değişkene ilişkin varyanslar gruplar arasında eşittir.<sup>12</sup>

### Non-Parametrik testi: Kruskal Wallis Testi

#### Hipotezleri:

$H_0$ : Grupların puan ortalamaları arasında fark yoktur. (sınanan hipotez)

$H_1$ : Grupların puan ortalamaları arasında fark vardır.

#### Test sonucu:

$p > 0,05$  ise  $H_0$  hipotezi kabul edilir: Grupların puan ortalamaları arasında fark yoktur.

$p < 0,05$  ise  $H_0$  hipotezi ret edilir: Grupların puan ortalamaları arasında fark vardır.

<sup>12</sup> SPSS'te ANOVA testi yapılırken, "Homogeneity of variance test" seçildiğinde, grupların varyanslarının eşitliği Levene testi ile sınanır. SPSS test çıktısında "Test of Homogeneity of Variances" tablosu bize grupların varyanslarının eşit olup olmadığını belirtir. Bu tablodaki "Based on Mean" satırındaki "Sig." yani p değerine bakarız. p değerinin 0,05'den büyük olması varyansların eşit olduğunu, p değerinin 0,05'den küçük olması ise varyansların eşit olmadığı anlamına gelir. Varyansların eşit olup olmadığına göre post-hoc test seçenekleri değişecektir.

### Etki Büyüklüğü:

Yapılan hipotez testi bize grupların ortalamaları arasında istatistiksel olarak anlamlı bir fark olup olmadığını söyler, ancak fark var ise bu farkın etkisi ya da büyüklüğü konusunda bilgi vermez. Etki büyüklüğü bize etkinin önemi konusunda objektif bir ölçüm sağladığı için faydalıdır (Field, 2009). Etki büyüklüğü sonuçların yorumlanmasında dikkate alınması gereken istatistiksel bir değerdir. Hipotez testlerinde etki büyüklüğü değerlerinin de hesaplanarak yorumlanması sonuçların anlaşılabilirliğini arttıracaktır (Büyüköztürk, 2010). Hipotez testinde fark bulunduğu durumda etki büyüklüğüne bakılır. Hipotez testleri için hesaplanan farklı etki büyüklüğü değerleri söz konusudur.<sup>13</sup> Bağımsız Gruplar için Tek Yönlü Varyans Analizi için kullanılan etki büyüklüğü değerleri:

**Eta kare ( $\eta^2$ ):** Eta kare korelasyon katsayısı. Eta kare bağımsız değişkenin (faktör) bağımlı değişken üzerinde ne derecede etkili olduğunu gösterir. Eta kare bağımsız değişkenin (ya da faktörün) bağımlı değişkendeki toplam varyansın (değişimin) ne kadarını açıkladığını gösterir ve 0 ile 1 arasında bir değer alır (Pallant, 2017; Büyüköztürk, 2010). 0,01 küçük, 0,06 orta ve 0,14 düzeyindeki eta kare ise büyük (large) etki olarak yorumlanır (Cohen, 1992). Eta kare değerinin kendisi bağımsız değişkenin bağımlı değişkendeki varyansın ne kadarını (%) açıkladığını ifade eder (Rosenthal ve Rosnow, 2008). Eta kare 0,01 ise varyansın %1'inin bağımsız değişken (faktör) tarafından açıklandığını, Eta kare 0,06 ise varyansın %6'sının açıklandığını ve Eta kare 0,14 ise varyansın %14'ünün bağımsız değişken tarafından açıklandığı anlamına gelir.

Eta kare ( $\eta^2$ ) değeri aşağıdaki formülle hesaplanır (Yockey, 2016):

$$\text{Eta kare } (\eta^2) = \frac{\text{Gruplar arası kareler toplamı}}{\text{Toplam kareler toplamı}}$$

SPSS çıktısında

$$\text{Eta kare } (\eta^2) = \frac{\text{Sum of Squares Between Groups}}{\text{Sum of Squares Total}}$$

**r etki değeri:** Pearson'un korelasyon katsayısı olan **r** bir etki değeri katsayısıdır. **r** değeri 0 (etki yok) ile 1 (mükemmel etki/perfect effect) arasında bir değer alır. **r** değeri işaretinden bağımsız olarak değerlendirilir. **r** değerinin 0,1 olması küçük etki, 0,3 olması orta ve 0,5 olması büyük (large) etki olarak değerlendirilir (Field, 2009). **r** katsayısının karesi ( $r^2$ ) toplam varyansın ne kadarını açıkladığını ifade eder.  $r^2$  değeri üzerinden bağımsız değişkenin bağımlı değişken üzerindeki değişimin ne kadarını açıkladığı bilgisine de ulaşırız. Cohen (1992) **r** katsayısının etki düzeylerinin yorumlanmasına ilişkin -yaygın olarak kullanılan- şu sınıflandırmayı yapmıştır:

**r = 0.10 küçük etki (small effect).** Bu durumda etki toplam varyansın %1'ini açıklar.

**r = 0.30 orta etki (medium effect).** Bu durumda etki toplam varyansın %9'unu açıklar.

**r = 0.50 büyük etki (large effect).** Bu durumda etki toplam varyansın %25'ini açıklar.

<sup>13</sup> **r** etki büyüklüğü doğrusal bir ilişkinin değeri/katsayısı olarak işlev görürken, Eta kare ( $\eta^2$ ) etki büyüklüğü ise her türlü ilişkinin bir değeri/katsayısı olarak işlev görmektedir (Rosenthal ve Rosnow, 2008). Rosenthal, Rosnow ve Field **r** etki değerini etki büyüklüğü katsayısı olarak kullanıp önerirken, Pallant Eta kare ( $\eta^2$ ) etki büyüklüğü katsayısını kullanıp, önermektedir. Eta kare ( $\eta^2$ ) ve **r** değerleri en yaygın kullanılan iki etki büyüklüğü değeridir.

r etki değeri aşağıdaki formülle hesaplanır (Field, 2009):

$$r = \sqrt{\frac{\text{Gruplar arası kareler toplamı}}{\text{Toplam kareler toplamı}}}$$

SPSS çıktısında

$$r = \sqrt{\frac{\text{Sum of Squares Between Groups}}{\text{Sum of Squares Total}}}$$

### Bağımsız Gruplar için Tek Yönlü Varyans Analizi için Post-Hoc Test:

#### Post-Hoc testler hangi anlamlılık düzeyi üzerinden değerlendirilir?

ANOVA testinin sonucu bize gruplar arasında farkın olup olmadığını söyler. Farkın kaynağını yani hangi grupların birbirinden farklılaştığını söylemez! Farkın kaynağını tespit etmek için post-hoc test yapılması gerekir. Post-Hoc test sonucunda hangi grupların birbirinden farklılaştığını ya da farklılaşmadığını tespit ederiz. Post-Hoc testte Tip I hatadan kaçınmak için Alfa ( $\alpha$ ) anlamlılık düzeyi eşleştirilen grup sayısına (test sayısına) bölünür ve bulunan yeni Alfa ( $\alpha$ ) değeri post-hoc testlerdeki yeni anlamlılık düzeyini ifade eder. Post-Hoc testlerde Tip I hatadan kaçınmak için yeni bir anlamlılık düzeyinin/sınırının hesaplanması ve kullanılmasına Bonferroni düzeltmesi denir (Connolly, 2007; Field, 2009; Pallant, 2017). Bonferroni düzeltmesi post-hoc testlerde Tip I hatadan kaçınmak için anlamlılık sınırında yapılan bir düzeltmedir. Post-Hoc testlerde anlamlılık düzeyinde yapılan bu düzeltme Tip I hatadan kaçınmayı sağlarken, Tip II hataya düşme ihtimalini de arttıracaktır<sup>14</sup> Pos-Hoc testlerde Bonferroni düzeltmesi yaygın olarak kullanılan bir yöntem olsa da bir zorunluluk değildir.

#### Post-Hoc testler için Bonferroni düzeltmesi Örnekleri:

##### 3 grup için post-hoc test yapıldığında

Eşleşen grup sayısı (test sayısı):

1-2, 1-3, 2-3; 3 eşleşen/karşılaştırılan grup olacaktır. Yani post-hoc 3 test yapılacaktır.

$\alpha/3 = 0,05/3 = 0,017$  olacaktır. **Post-Hoc testlerde farkın kaynağı  $\alpha$ : 0,017 düzeyinde sınanacaktır.**  $p > 0,017$  ise  $H_0$  hipotezi kabul edilir.  $p < 0,017$  ise  $H_0$  hipotezi reddedilir.

##### 4 grup için post-hoc test yapıldığında

Eşleşen grup sayısı (test sayısı):

1-2, 1-3, 1-4, 2-3, 2-4, 3-4; 6 eşleşen/karşılaştırılan grup olacaktır. Yani post-hoc 6 test yapılacaktır.

$\alpha/6 = 0,05/6 = 0,008$  olacaktır. **Post-Hoc testlerde farkın kaynağı  $\alpha$ : 0,008 düzeyinde sınanacaktır.**  $p > 0,008$  ise  $H_0$  hipotezi kabul edilir.  $p < 0,008$  ise  $H_0$  hipotezi reddedilir.

##### 5 grup için post-hoc test yapıldığında

Eşleşen grup sayısı (test sayısı):

1-2, 1-3, 1-4, 1-5, 2-3, 2-4, 2-5, 3-4, 3-5, 4-5; 10 eşleşen/karşılaştırılan grup olacaktır. Yani post-hoc 10 test yapılacaktır.

<sup>14</sup> Tip I ve Tip II hata için bakınız sayfa:27



$\alpha / 10 = 0,05/10 = 0,005$  olacaktır. **Post-Hoc testlerde farkın kaynağı  $\alpha$ : 0,005 düzeyinde sınanacaktır.**  $p > 0,005$  ise  $H_0$  hipotezi kabul edilir.  $p < 0,005$  ise  $H_0$  hipotezi ret edilir.

### Post-Hoc test seçenekleri:

Tek Yönlü Varyans Analizinde grupların ortalamaları arasında istatistiksel olarak anlamlı bir farklılık tespit ettiğimizde, bu farkın kaynağını (hangi grupların birbirinden farklılaştığını) tespit etmek için post-hoc test yaparız. Post-Hoc test bize hangi grupların birbirinden farklılaştığını ya da farklılaşmadığını gösterecektir.

Yapılacak olan post-hoc test seçimi grupların varyanslarının eşit olup olmadığına bağlıdır. SPSS'te ANOVA testi yapılırken, "Homogeneity of variance test" seçildiğinde, grupların varyanslarının eşitliği Levene testi ile sınanır. SPSS test çıktısında "Test of Homogeneity of Variances" tablosu bize grupların varyanslarının eşit olup olmadığını belirtir. Bu tablodaki "Based on Mean" satırındaki "Sig." yani p değerine bakarız. p değerinin 0,05'den büyük olması varyansların eşit olduğunu, p değerinin 0,05'den küçük olması ise varyansların eşit olmadığı anlamına gelir. Post-Hoc testlerin seçimi varyansların eşit olup olmadığını sınavan bu teste göre yapılmalıdır.

### Varyanslar eşitse:

Grupların varyanslarının eşit olduğu durumda Field (2009) tarafından koşullara göre önerilen post-hoc testler:

**Tukey:** Örneklem sayıları eşitse/yakınsa, yaygın olarak kullanılan, Tip I hatadan uzak tutucu (conservative) bir testtir.

**Scheffe:** Örneklem sayıları birbirinden farklı ise yaygın olarak kullanılan, Tip I hatadan uzak tutucu (conservative) bir testtir.

**Bonferroni:** En güvenilir, yaygın olarak kullanılan, Tip I hatadan en uzak tutucu (conservative) bir testtir.

**Gabriel:** Örneklem sayıları eşitse/yakınsa güvenilir alternatif bir post-hoc testtir.

**Hochberg's GT2:** Örneklem sayıları birbirinden farklı ise güvenilir alternatif bir post-hoc testtir.

**LSD:** Fark çıkarma eğiliminde liberal bir testtir. Tip I hataya düşme ihtimali yüksek bir testtir.<sup>15</sup>

**S-N-K:** Çok liberal bir testtir ve Tip I hatanın kontrolünden yoksundur (lacks).

### Varyanslar eşit değilse:

Grupların varyanslarının eşit olmadığı durumda Field (2009) tarafından koşullara göre önerilen post-hoc testler:

**Tamhane'in T2:** Tutucu (conservative) bir testtir.

**Dunnett's T3:** Tutucu (conservative) ve çok sıkı (tight) Tip I hata kontrolü yapan bir testtir.

**Games-Howell:** Games-Howell prosedürü en güçlü olan testtir fakat örneklem boyutları küçük olduğunda liberal olabilir. Bununla birlikte Games-Howell grupların örneklem sayıları eşit

<sup>15</sup> Karşılaştırılan grup sayısı 3'ün üzerinde ise hiç önerilmez!

olmadığında da doğrudur (accurate). Varyansların eşit olmadığı durumda yaygın olarak kullanılan, prosedürü en güçlü testtir.

**Dunnet's C:** Tutucu (conservative) ve çok sıkı (tight) Tip I hata kontrolü yapan bir testtir.

**Bu testin uygulamasına örnek:** Gelir seviyesi düşük, orta ve yüksek olan grupların depresyon puan ortalamalarını karşılaştırmak istediğimizde, depresyon puanı tüm örneklem için normal dağılım özelliği gösteriyor ve her bir grupta en az 10 katılımcının olması durumunda Bağımsız Gruplar için Tek Yönlü Varyans Analizi yapılır.

**Önemli not:** Bu uygulamada karşılaştırılan nicel verilerin normal dağılım özelliği gösterdiği varsayılmıştır. Siz kendi analizlerinizde öncelikli olarak normal dağılıma uygunluğu sınamalısınız.

#### SPSS Uygulaması:

- 1- **Analyze, Compare Means, One-way ANOVA** seçilir.
- 2- **One-way ANOVA** penceresinde sol taraftaki değişkenler listesinden **nicel değişkenler** seçilerek sağ taraftaki **Dependent List** kutusuna **aktarılır**.  
Sol taraftaki değişkenler listesinden **kategorik (bağımsız) değişken** seçilerek sağ taraftaki **Factor** kutusuna **aktarılır**.
- 3- **One-way ANOVA** penceresinde **Options** tıklanır ve **Descriptive Homogeneity of variance test** seçilir. **Continue** tıklanır.
- 4- **One-way ANOVA** penceresinde (ANOVA testinde fark çıkar ise) **Post-Hoc** tıklanır ve **Varyanslar eşit ise Equal Variances Assummed** kısmından bir test seçilir. **Varyanslar eşit değil ise Equal Variances Not Assummed** kısmından bir test seçilir. **Continue** tıklanır.
- 5- **One-way ANOVA** penceresinde **OK** tıklanır.

#### Bağımsız Gruplar için Tek Yönlü Varyans Analizi (One-Way ANOVA) SPSS çıktısı:

##### Oneway

##### Descriptives

depresyon 1. ölçüm puanı

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
düşük	32	74,2500	14,26252	2,52128	69,1078	79,3922	48,00	90,00
orta	78	54,4615	14,44892	1,63602	51,2038	57,7193	30,00	75,00
yüksek	25	36,8400	4,81041	,96208	34,8544	38,8256	30,00	45,00
Total	135	55,8889	17,90425	1,54095	52,8412	58,9366	30,00	90,00

**Descriptives:** Tanımlayıcı özet değerler.

**N:** örneklem sayısı. **Mean:** Ortalama. **Std. Deviation:** Standart Sapma.

**95% Confidence Interval for Mean Lower-Upper Bound:** %95 Güven Aralığında ortalamanın Alt-Üst sınırı değerleri.

### Test of Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.
depresyon 1. ölçüm puanı	Based on Mean	14,716	2	132	,000
	Based on Median	11,033	2	132	,000
	Based on Median and with adjusted df	11,033	2	112,247	,000
	Based on trimmed mean	14,148	2	132	,000

**Test of Homogeneity of Variances:** Grupların varyanslarının eşitliğinin sınıandığı test. Based on Mean satırındaki Sig. yani p değerine bakarız. Bu tabloda  $p:0,000$  yani  $p<0,05$  olduğu için gruplar arasındaki varyanslar eşit değildir. Post-Hoc test seçimi Varyansların eşit olmama durumunda önerilen testler arasından seçilecektir.

### ANOVA

depresyon 1. ölçüm puanı

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	20018,589	2	10009,294	57,603	,000
Within Groups	22936,745	132	173,763		
Total	42955,333	134			

**ANOVA:** Tek Yönlü Varyans Analizi Sonuçları

**Between Groups:** Gruplar arasında. **Within Groups:** Gruplar içi. **Total:** Toplam.

**Sum of Squares:** Kareler toplamı. **df:** Serbestlik derecesi. **Mean Square:** Kareler ortalaması. **F:** test değeri. **Sig. :** testin p değeri.

### Post Hoc Tests

Varyanslar eşit olmadığı için post-hoc test olarak Games-Howell tercih edildi. Post-Hoc testteki anlamlılık düzeyi  $\alpha: 0,017$  düzeyinde sınanacaktır (aşağıdaki SPSS tablosu).

### Multiple Comparisons

Dependent Variable: depresyon 1. ölçüm puanı

Games-Howell

		Mean		Sig.	95% Confidence Interval	
(I) gelir seviyeleri	(J) gelir seviyeleri	Difference (I-J)	Std. Error		Lower Bound	Upper Bound
düşük	orta	19,78846*	3,00556	,000	12,5605	27,0164
	yüksek	37,41000*	2,69860	,000	30,8393	43,9807
orta	düşük	-19,78846*	3,00556	,000	-27,0164	-12,5605
	yüksek	17,62154*	1,89793	,000	13,1067	22,1364
yüksek	düşük	-37,41000*	2,69860	,000	-43,9807	-30,8393
	orta	-17,62154*	1,89793	,000	-22,1364	-13,1067

\*. The mean difference is significant at the 0.05 level.

**Mean Difference:** Ortalamalar arası fark. **Sig. :** testin p değeri.

**Tablo: Gelir seviyesi gruplarının depresyon puan ortalamalarının karşılaştırılması**

	n	Ortalama	Standart sapma	F	p	Anlamlı fark*
Düşük	32	74,25	14,26	57,603	,000	Düşük-Orta
Orta	78	54,46	14,45			Düşük-Yüksek
Yüksek	25	36,84	4,81			Orta-Yüksek
Toplam	135	55,89	17,91			

\* Games-Howell testi yapılmıştır. Bonferroni düzeltmesi sonrasında  $p < 0,017$  düzeyinde fark çıkan sonuçlar.

Araştırma grubundaki düşük gelir seviyesi kategorisindeki katılımcıların depresyon puan ortalaması  $72 \pm 14$ , orta gelir seviyesi kategorisinin ortalaması  $55 \pm 15$  iken yüksek gelir seviyesi kategorisindeki katılımcıların depresyon puan ortalaması  $37 \pm 5$ 'dir. Gelir seviyesi gruplarının depresyon puan grupları arasında istatistiksel olarak anlamlı bir farklılık tespit edilmiştir ( $p < 0,01$ ). Tespit edilen bu farkın etki büyüklüğü Eta kare ( $\eta^2$ )= 0,47 olduğu, farkın büyük bir etkiye sahip olduğu ve toplam varyansın %47'sinin bağımsız değişken (gelir) tarafından açıklandığını göstermektedir. (Diğer etki değerleri ile açıklama: Tespit edilen bu farkın etki büyüklüğü  $r = 0,68$  olduğu, farkın büyük bir etkiye sahip olduğu ve toplam varyansın %47'sinin bağımsız değişken (gelir) tarafından açıklandığını göstermektedir.) Farklılığın düşük ile orta ve yüksek gelir gruplarının; orta ile yüksek gelir gruplarının puan ortalamalarının birbirinden farklılaşmasından kaynaklandığı tespit edilmiştir ( $p < 0,01$ ).

#### Etki büyüklüğü:

##### Eta kare ( $\eta^2$ ):

$$\text{Eta kare} = \frac{\text{Gruplar arası kareler toplamı}}{\text{Toplam kareler toplamı}} = \frac{20.018,589}{42.955,333} = 0,47$$

**Eta kare ( $\eta^2$ )= 0,47** (büyük etki) Bu durumda etki toplam varyansın %47'sini açıklar.

##### r etki değeri:

$$r = \sqrt{\frac{\text{Gruplar arası kareler toplamı}}{\text{Toplam kareler toplamı}}} = \sqrt{\frac{20.018,589}{42.955,333}} = 0,685$$

**r= 0,68** (büyük etki)

Bu durumda etki toplam varyansın %47'sini açıklar ( $r^2 = 0,685^2 = 0,47$ ).

## 2.7. Kruskal Wallis Testi

Bağımsız Gruplar için Tek Yönlü Varyans Analizi'nin (One-Way ANOVA) Non-Parametrik testidir. Bağımsız Gruplar için Tek Yönlü Varyans Analizi'nin varsayımlarının sağlanamaması durumunda üç ya da daha fazla bağımsız grubun puanları arasında istatistiksel olarak anlamlı bir farklılığın olup olmadığı bu test ile sınanabilir.

Kruskal Wallis Testi, en az sıralama ölçeğindeki verileri grubuna bakılmaksızın -bütün verileri- küçükten büyüğe doğru sıralar ve 1'den başlayarak sıra sayısı verilerek, sıra sayısına dönüştürür. Bu en az aralık ölçeğindeki verilerin basitçe sıralama ölçeğine dönüştürülmesi işlemidir. Ardından veriler karşılaştırılacak gruplara ayrılır ve grupların sıra sayılarının ortalaması (sıra ortalaması/mean rank) hesaplanır (Can,2019; Büyüköztürk, 2010). Kruskal Wallis Testi, bağımsız üç ya da daha fazla grubun sıra ortalamasını (mean rank) karşılaştırır ve bunun sonucunda grupların puanları arasında istatistiksel olarak anlamlı bir fark olup olmadığı sonucuna ulaşır.

**Kruskal Wallis Testinin sıra değerlerini üretme ve işlem süreci örneği:**

Veri seti			Kruskal Wallis testinin işlem süreci			
Anket no	Grup	Puan	Anket no	Grup	Puan	Sıra değeri
1	A	63	17	C	61	1
2	A	65	6	A	62	2
3	B	87	1	A	63	3
4	B	89	14	C	64	4
5	A	95	2	A	65	5
6	A	62	12	B	67	6
7	B	75	8	B	73	7
8	B	73	7	B	75	8
9	B	77	13	A	76	9
10	B	83	9	B	77	10
11	A	81	15	C	78	11
12	B	67	16	C	79	12
13	A	76	11	A	81	13
14	C	64	10	B	83	14
15	C	78	19	C	85	15
16	C	79	3	B	87	16
17	C	61	4	B	89	17
18	C	93	18	C	93	18
19	C	85	5	A	95	19

	n	Sıra Ortalaması (Mean Rank)	Sıralar Toplamı (Sum of Ranks)
B Grubu	7	11,14	78
A Grubu	6	8,50	51
C Grubu	6	10,17	61

Kruskal Wallis Testi grupların medyanlarını karşılaştırmaz. Farklılığın hesaplanmasında medyanların bir işlevi yoktur, dolayısı ile bu testin sonucunda ulaşılan grupların birbirinden farklılaşmasını ya da farklılaşmamasını, medyanlar üzerinden açıklanması ve yorumlanması yanlıştır!

Kruskal Wallis Testi grupların puanları arasında istatistiksel olarak anlamlı bir fark olup olmadığını sınar. Bu testin sonucunda farklılık olduğu sonucuna ulaştığımızda, grupların puanları arasında bir fark olduğu sonucuna ulaşırız. Ancak Kruskal Wallis Testi bize bu farkın hangi grupların birbirinden farklılaştığından kaynaklandığını vermez. Kruskal Wallis Testi sonucunda farklılık olduğu sonucuna ulaştığımızda farkın kaynağını tespit etmek için post-hoc test yapmamız gerekir. Yapılan post-hoc test bize hangi grupların birbirinden farklılaştığını gösterecektir.

**Test neyi yapar:**

Üç ya da daha fazla bağımsız grubun sıra ortalamasını (mean rank) karşılaştırır ve bunun sonucunda grupların puanları arasında istatistiksel olarak anlamlı bir farklılık olup olmadığını sınar. Farklılık tespit edilmesi durumunda yapılacak post-hoc testler ile farklılığın hangi grupların birbirinden farklılaşmasından kaynaklandığı tespit edilir

**Varsayımları:**

Gruplar birbirinden bağımsızdır (3 ya da daha fazla bağımsız grup).

Bağımlı değişken en az sıralama ölçeğinde bir değişkendir.

**Parametrik testi:** Bağımsız Gruplar için Tek Yönlü Varyans Analizi

**Hipotezleri:**

$H_0$ : Grupların puanları arasında fark yoktur. (sınanan hipotez)

$H_1$ : Grupların puanları arasında fark vardır.

**Test sonucu:**

$p > 0,05$  ise  $H_0$  hipotezi kabul edilir: Grupların puanları arasında fark yoktur.

$p < 0,05$  ise  $H_0$  hipotezi ret edilir: Grupların puanları arasında fark vardır.

### Etki Büyüklüğü:

Yapılan hipotez testi bize grupların puanları arasında istatistiksel olarak anlamlı bir fark olup olmadığını söyler, ancak fark var ise bu farkın etkisi ya da büyüklüğü konusunda bilgi vermez. Etki büyüklüğü bize etkinin önemi konusunda objektif bir ölçüm sağladığı için faydalıdır (Field, 2009). Etki büyüklüğü sonuçların yorumlanmasında dikkate alınması gereken istatistiksel bir değerdir. Hipotez testlerinde etki büyüklüğü değerlerinin de hesaplanarak yorumlanması sonuçların anlaşılabilirliğini arttıracaktır (Büyüköztürk, 2010). Hipotez testinde fark bulunduğu durumda etki büyüklüğüne bakılır. Hipotez testleri için hesaplanan farklı etki büyüklüğü değerleri söz konusudur. Kruskal Wallis testi için kullanılan etki büyüklüğü değeri:

**Eta kare ( $\eta^2$ ):** Eta kare korelasyon katsayısı. Eta kare bağımsız değişkenin bağımlı değişken üzerinde ne derecede etkili olduğunu gösterir. Eta kare bağımsız değişkenin (ya da faktörün) bağımlı değişkendeki toplam varyansın (değişimin) ne kadarını açıkladığını gösterir ve 0 ile 1 arasında bir değer alır (Pallant, 2017; Büyüköztürk, 2010). 0,01 küçük, 0,06 orta ve 0,14 düzeyindeki eta kare ise büyük (large) etki olarak yorumlanır (Cohen, 1992). Eta kare değerinin kendisi bağımsız değişkenin bağımlı değişkendeki varyansın ne kadarını (%) açıkladığını ifade eder (Rosenthal ve Rosnow, 2008). Eta kare 0,01 ise varyansın %1'inin bağımsız değişken tarafından açıklandığını, Eta kare 0,06 ise varyansın %6'sının açıklandığını ve Eta kare 0,14 ise varyansın %14'ünün bağımsız değişken tarafından açıklandığı anlamına gelir.

Eta kare ( $\eta^2$ ) değeri aşağıdaki formülle hesaplanır (Green ve Salkind, 2014):

$$\text{Eta kare } (\eta^2) = \frac{\chi^2}{n - 1}$$

$\chi^2$  : Ki-kare değeri, SPSS analiz çıktısındaki Kruskal-Wallis H değeridir (eski versiyonlarda Chi-Square değeri).

### Kruskal Wallis Testi için Post-Hoc Test:

#### Post-Hoc testler hangi anlamlılık düzeyi üzerinden değerlendirilir?

Kruskal Wallis testinin sonucu bize gruplar arasında farkın olup olmadığını söyler. Farkın kaynağını yani hangi grupların birbirinden farklılaştığını söylemez! Farkın kaynağını tespit etmek için post-hoc test yapılması gerekir. Post-Hoc test sonucunda hangi grupların birbirinden farklılaştığını ya da farklılaşmadığını tespit ederiz. Post-Hoc testte Tip I hatadan kaçınmak için Alfa ( $\alpha$ ) anlamlılık düzeyi eşleştirilen grup sayısına (test sayısına) bölünür ve bulunan yeni Alfa ( $\alpha$ ) değeri post-hoc testlerdeki yeni anlamlılık düzeyini ifade eder. Post-Hoc testlerde Tip I hatadan kaçınmak için yeni bir anlamlılık düzeyinin/sınırının hesaplanması ve kullanılmasına Bonferroni düzeltmesi denir (Connolly, 2007; Field, 2009; Pallant, 2017). Bonferroni düzeltmesi post-hoc testlerde Tip I hatadan kaçınmak için anlamlılık sınırında yapılan bir düzeltmedir. Post-Hoc testlerde anlamlılık düzeyinde yapılan bu düzeltme Tip I hatadan kaçınmayı sağlarken, Tip II hataya düşme ihtimalini de arttıracaktır.<sup>16</sup> Pos-Hoc testlerde Bonferroni düzeltmesi yaygın olarak kullanılan bir yöntem olsa da bir zorunluluk değildir.

### Kruskal Wallis'in Post-Hoc testi: Mann Whitney U testidir

<sup>16</sup>Post-Hoc testlerde Alfa ( $\alpha$ ) anlamlılık düzeyi Bonferroni düzeltmesi hesaplamaları için sayfa:56 ve Tip I ve Tip II hata için bakınız sayfa:27

**Bu testin uygulamasına örnek:** Gelir seviyesi düşük, orta ve yüksek olan grupların depresyon puan ortalamalarını karşılaştırmak istediğimizde, depresyon puanı tüm örneklem için normal dağılım özelliği göstermiyor ise bu durumda Kruskal Wallis Testi ile farklılık sınanabilir.

### SPSS Uygulaması:

- 1- **Analyze, Nonparametric Tests, Legacy Dialogs, K Independent Samples** seçilir.
- 2- **Test for Several Independent Samples** penceresinde sol taraftaki değişken listesinden **bağımlı nicel değişken Test Variable List'e** taşınır.  
Sol taraftaki değişkenler listesinden **bağımsız kategorik değişken** seçilerek sağ taraftaki **Grouping Variable** kısmına **aktarılır. Define Range** tıklanır ve bağımsız değişken kategorilerinin aralığı (hangi kod aralığında olduğu) yazılır ve **Continue** tıklanır. **Two Independent Samples Test** penceresinde **Kruskal-Wallis Test H** testi otomatik seçilidir.  
**OK** Tıklanır.

### SPSS çıktısı:

#### NPar Tests

#### Kruskal-Wallis Test

Ranks			
	gelir seviyeleri	N	Mean Rank
depresyon 1. ölçüm puanı	düşük	32	105,69
	orta	78	65,59
	yüksek	25	27,28
	Total	135	

**N:** örneklem sayısı

**Mean Rank:** Sıra ortalaması

#### Test Statistics<sup>a,b</sup>

depresyon 1. ölçüm puanı	
Kruskal-Wallis H	57,362
df	2
Asymp. Sig.	,000

a. Kruskal Wallis Test

b. Grouping Variable: gelir seviyeleri

**Kruskal-Wallis H:** test değeri

**Asymp. Sig. :** testin p değeri

**df:** serbestlik derecesi



Fark çıktığı için post-hoc test olarak Mann Whitney U testi ile ikili grup karşılaştırmaları yapıyoruz.

### SPSS Uygulaması:

**1- Analyze, Nonparametric Tests, Legacy Dialogs, 2 Independent Samples seçilir.**

**2- Two Independent Samples Test** penceresinde

Sol taraftaki değişken listesinden **bağımlı nicel değişken Test Variable List'e** taşınır.

Sol taraftaki değişkenler listesinden **bağımsız kategorik değişken** seçilerek sağ taraftaki **Grouping Variable** kısmına aktarılır. **Define Groups** tıklanır ve **bağımsız değişken kategorilerinin ne olarak kodlandığı** yazılır ve **Continue** tıklanır.

**Two Independent Samples Test** penceresinde Test Type **Mann-Whitney U** testi otomatik seçilidir.

**OK** Tıklanır.

Bu uygulamayı/karşılaştırmayı bu örnek için 3 kez yapıyoruz: 1. Grup-2. Grup karşılaştırması, 1. Grup-3. Grup karşılaştırması ve 2. Grup-3. Grup karşılaştırması için.

**Post-Hoc test Mann Whitney U testi SPSS çıktısı:**

### NPar Tests

#### Mann-Whitney Test

Ranks				
	gelir seviyeleri	N	Mean Rank	Sum of Ranks
depresyon 1. ölçüm puanı	düşük	32	80,69	2582,00
	orta	78	45,17	3523,00
	Total	110		

#### Test Statistics<sup>a</sup>

depresyon 1. ölçüm puanı	
Mann-Whitney U	442,000
Wilcoxon W	3523,000
Z	-5,316
Asymp. Sig. (2-tailed)	,000

## Mann-Whitney Test

Ranks				
	gelir seviyeleri	N	Mean Rank	Sum of Ranks
depresyon 1. ölçüm puanı	düşük	32	41,50	1328,00
	yüksek	25	13,00	325,00
	Total	57		

### Test Statistics<sup>a</sup>

depresyon 1.

ölçüm puanı

Mann-Whitney U	,000
Wilcoxon W	325,000
Z	-6,457
Asymp. Sig. (2-tailed)	,000

## Mann-Whitney Test

Ranks				
	gelir seviyeleri	N	Mean Rank	Sum of Ranks
depresyon 1. ölçüm puanı	orta	78	59,92	4674,00
	yüksek	25	27,28	682,00
	Total	103		

### Test Statistics<sup>a</sup>

depresyon 1.

ölçüm puanı

Mann-Whitney U	357,000
Wilcoxon W	682,000
Z	-4,777
Asymp. Sig. (2-tailed)	,000

**Tablo: Gelir seviyesi gruplarının depresyon puanlarının karşılaştırılması**

	n	Sıra Ortalama	$\chi^2$	p	Anlamlı fark*
Düşük	32	105,69	57,362	,000	Düşük-Orta
Orta	78	65,59			Düşük-Yüksek
Yüksek	25	27,28			Orta-Yüksek
Toplam	135				

\* Mann Whitney U testi yapılmıştır. Bonferroni düzeltmesi sonrasında  $p < 0,017$  düzeyinde fark çıkan sonuçlar.

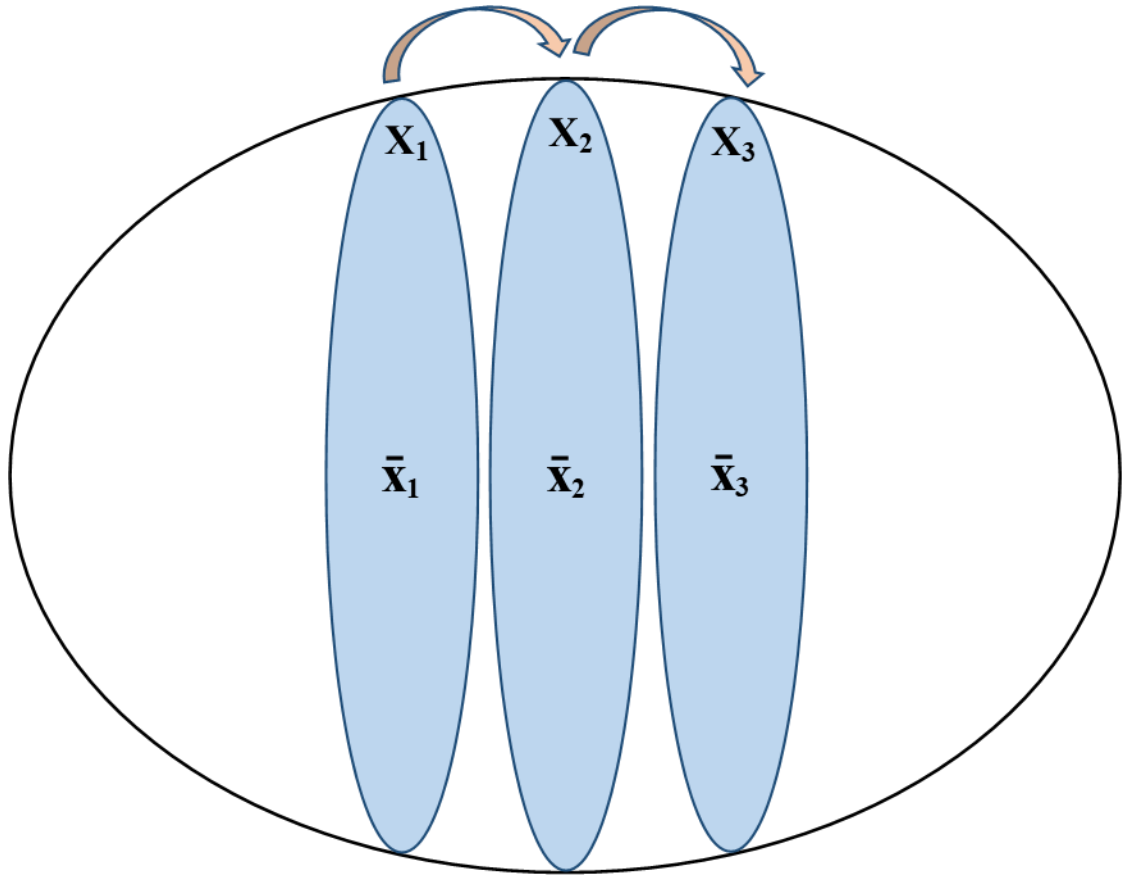
Araştırma grubundaki düşük gelir seviyesi kategorisindeki katılımcıların depresyon puanı sıra ortalaması 106, orta gelir seviyesi grubunun 66 iken yüksek gelir kategorisindeki katılımcıların depresyon puanı sıra ortalaması 27'dir. Gelir seviyesi gruplarının depresyon puanı sıra ortalamaları arasında istatistiksel olarak anlamlı bir farklılık tespit edilmiştir ( $p < 0,01$ ). Tespit edilen bu farkın etki büyüklüğü Eta kare ( $\eta^2$ )= 0,43 olduğu, farkın büyük bir etkiye sahip olduğu ve toplam varyansın %43'ünün bağımsız değişken (gelir) tarafından açıklandığını göstermektedir. Farklılığın düşük ile orta ve yüksek gelir gruplarının; orta ile yüksek gelir gruplarının sıra ortalamalarının birbirinden farklılaşmasından kaynaklandığı tespit edilmiştir ( $p < 0,01$ ).

#### Etki Büyüklüğü:

$$\text{Eta kare} = \frac{\chi^2}{n - 1} = \frac{57,362}{135 - 1} = 0,43$$

**Eta kare ( $\eta^2$ )= 0,43 (büyük etki).** Bu durumda etki toplam varyansın %43'ünü açıklar.

## 2.8. Tekrarlayan Ölçümler için Tek Yönlü Varyans Analizi <sup>17</sup>



Bir grubun tekrarlayan ikiden fazla ölçüm ortalamaları arasında istatistiksel olarak anlamlı bir fark olup olmadığını sınamak için parametrik bir test olan Tekrarlayan Ölçümler için Tek Yönlü Varyans Analizi (One Way ANOVA for Repeated Measures) yapılır. Bu testin yapılabilmesinin varsayımları (Can, 2019); Örneklemin Evrenden tesadüfî bir şekilde seçilmesi, ortalamaları kıyaslanacak en az aralık ölçeğinde olan tekrarlayan ölçüm değerlerinin normal dağılım özelliği göstermesi, tekrarlayan ölçüm farklarının varyanslarının eşit olması (Sphercity/Küresellik) beklenir.

ANOVA testi büyük örneklemlemlerle çalışıldığında normalliğin ihlalden etkilenmez. ANOVA testinin örneklem sayısı 30'un üzerine çıktığı durumlarda test sonucunun normalliğin ihlalden ciddi bir şekilde etkilenmeyeceği ve oldukça doğru p değerleri üreteceğine dair liberal görüşler mevcuttur (Green ve Salkind, 2014, Pallant, 2017). Bu ANOVA'nın güçlü (robust) bir test olması nedeni ile testin varsayımlarını ihlal etmenin önemli olmadığı gibi yorumlansa da bu aynı zamanda durumun aşırı derecede basitleştirilmesidir (Field, 2009). Nicel verilerin normal dağılımı parametrik testlerin sağlıklı, güvenilir sonuçlar üretmesi için önemli bir koşul olduğu göz önüne alındığında, normalliğin ihlali durumunda bu testin yapılmaması önerilir.

<sup>17</sup> One Way ANOVA for Repeated Measures

Tekrarlayan ölçüm farklarının varyanslarının eşit olması Küresellik (Sphercity) olarak adlandırılır ve Küresellik varsayımı kabaca şu formülle gösterilebilir (Field, 2009):

$$\text{Varyans } X_1-X_2 \approx \text{Varyans } X_1-X_3 \approx \text{Varyans } X_2-X_3$$

Tekrarlayan ölçüm farklarının varyanslarının eşit olması (Sphercity/Küresellik) Tekrarlayan Ölçümler için ANOVA'nın bir varsayımı olsa da bir gereklilik değildir (Field, 2009). Küresellik varsayımı sağlanmadığında da analiz yapılır ve güvenilir, güçlü sonuçlar verir. ANOVA testi içinde Küreselliğin sağlanıp sağlanmadığı “Mauchly's Sphericity Test” ile sınıanır. “Mauchly's Sphericity Test”inin p değeri  $p > 0,05$  ise Küresellik varsayımı sağlanmıştır. Bu durumda ANOVA test değerleri olarak “Tests of Within-Subjects Effects” tablosundan ‘Sphericity Assumed’in –sətırında ki- F ve p değeri alınır. “Mauchly's Sphericity Test”inin p değeri  $p < 0,05$  ise Küresellik varsayımı sağlanmamıştır. Bu durumda ANOVA test değeri olarak “Tests of Within-Subjects Effects” tablosundan Greenhouse-Geisser ya da Huynh-Feldt –sətırında ki- F ve p değeri alınır.<sup>18</sup>

“Mauchly's Sphericity Test”inin p değeri  $p < 0,05$  yani Küresellik varsayımı sağlanmadığında “Mauchly's Sphericity Test” tablosunda ki Epsilon değeri bakılır. “Mauchly's Sphericity Test” tablosunda ki Epsilon'un altındaki Greenhouse-Geisser ve Huynh-Feldt değeri (epsilon değeri) 0,75'ten küçük ise “Tests of Within-Subjects Effects” tablosundan Greenhouse-Geisser'in F ve p değeri ANOVA test değeri olarak kullanılır. “Mauchly's Sphericity Test” tablosunda ki Epsilon'un altındaki Greenhouse-Geisser ve Huynh-Feldt değeri (epsilon değeri) 0,75'ten büyük ise “Tests of Within-Subjects Effects” tablosundan Huynh-Feldt'in F ve p değeri ANOVA test değeri olarak kullanılır (Field, 2009).

ANOVA testinin sonucu bize tekrarlayan ölçümler arasında farkın olup olmadığını söyler. Farkın kaynağını yani hangi ölçümlerin birbirinden farklılaştığını söylemez! Farkın kaynağını tespit etmek için post-hoc test yapmak gerekir. Post-Hoc test sonucunda hangi ölçümlerin birbirinden farklılaştığını ya da farklılaşmadığını tespit ederiz.

### Test neyi yapar:

Bir grubun üç ya da daha fazla tekrarlayan ölçüm ortalamalarını arasında anlamlı bir farklılık olup olmadığını sınıar. Farklılık tespit edilmesi durumunda yapılacak post-hoc testler ile farklılığın hangi ölçüm düzeylerinin birbirinden farklılaşmasından kaynaklandığı tespit edilir.

### Varsayımları:

Örnekleme Evrenden rastgele seçilmiştir.

Bağımlı değişken en az aralık ölçeğinde olan bir değişkendir.

Tekrarlayan (üç ya da daha fazla) ölçümler bir (aynı) gruba aittir.

Tekrarlayan ölçüm düzeylerinin hepsinde veriler normal dağılım özelliği gösterir.

Tekrarlayan ölçüm düzeyleri farklarının varyansları eşittir. (Küresellik/Sphercity Varsayımı).

<sup>18</sup> Küresellik varsayımı sağlanmadığında alternatif olarak “Multivariate Tests” tablosundan Wilks'Lambda'nın F ve p değeri de kullanılabilir (Pallant, 2017). Ancak bu kullanım yaygın değildir.

**Non-Parametrik testi: Friedman Testi****Hipotezleri:**

$H_0$ : Tekrarlayan ölçüm ortalamaları arasında fark yoktur. (sınanan hipotez)

$H_1$ : Tekrarlayan ölçüm ortalamaları arasında fark vardır.

**Test sonucu:**

$p > 0,05$  ise  $H_0$  hipotezi kabul edilir: Tekrarlayan ölçüm ortalamaları arasında fark yoktur.

$p < 0,05$  ise  $H_0$  hipotezi ret edilir: Tekrarlayan ölçüm ortalamaları arasında fark vardır.

**Etki Büyüklüğü:**

Yapılan hipotez testi bize Tekrarlayan ölçüm ortalamaları arasında istatistiksel olarak anlamlı bir fark olup olmadığını söyler, ancak fark var ise bu farkın etkisi ya da büyüklüğü konusunda bilgi vermez. Etki büyüklüğü bize etkinin önemi konusunda objektif bir ölçüm sağladığı için faydalıdır (Field, 2009). Etki büyüklüğü sonuçların yorumlanmasında dikkate alınması gereken istatistiksel bir değerdir. Hipotez testlerinde etki büyüklüğü değerlerinin de hesaplanarak yorumlanması sonuçların anlaşılabilirliğini arttıracaktır (Büyüköztürk, 2010). Hipotez testinde fark bulunduğu durumda etki büyüklüğüne bakılır. Hipotez testleri için hesaplanan farklı etki büyüklüğü değerleri söz konusudur.<sup>19</sup> Tekrarlayan Ölçümler için Tek Yönlü Varyans Analizi testi için kullanılan etki büyüklüğü değeri:

**Eta kare ( $\eta^2$ ):** Eta kare korelasyon katsayısı. Eta kare bağımsız değişkenin bağımlı değişken üzerinde ne derecede etkili olduğunu gösterir. Eta kare bağımsız değişkenin (ya da faktörün) bağımlı değişkendeki toplam varyansın (değişimin) ne kadarını açıkladığını gösterir ve 0 ile 1 arasında bir değer alır (Pallant, 2017; Büyüköztürk, 2010). 0,01 küçük, 0,06 orta ve 0,14 düzeyindeki eta kare ise büyük (large) etki olarak yorumlanır (Cohen, 1992). Eta kare değerinin kendisi bağımsız değişkenin bağımlı değişkendeki varyansın ne kadarını (%) açıkladığını ifade eder (Rosenthal ve Rosnow, 2008). Eta kare 0,01 ise varyansın %1'inin bağımsız değişken tarafından açıklandığını, Eta kare 0,06 ise varyansın %6'sının açıklandığını ve Eta kare 0,14 ise varyansın %14'ünün bağımsız değişken tarafından açıklandığı anlamına gelir. Kısmi Eta karenin yorumlanması da Eta kare değerleri ile aynıdır.

1- Formül aracılığı ile manuel Kısmi Eta Kare'nin hesaplanması:

$$\text{Eta kare (kısmi)} = \frac{\text{Kareler toplamı (ölçüm)}}{\text{Kareler toplamı (ölçüm)} + \text{Kareler toplamı (ölçüm hatası)}}$$

SPSS çıktısında Test of Within Subject Effects kısmında:

$$\text{Eta kare}_{(\text{kısmi})} = \frac{\text{Sum of Squares}}{\text{Sum of Squares} + \text{Sum of Squares (Error)}}$$

<sup>19</sup> Eta karenin dışında bu analiz için kullanılan diğer etki büyüklüğü değerleri omega kare ve r etki katsayısıdır. Omega kare ve r değerinin hesaplanması için Field (2009:480) bakınız. Eta kare ( $\eta^2$ ) etki büyüklüğü her türlü ilişkinin bir değeri/katsayısı olarak işlev görmektedir (Rosenthal ve Rosnow, 2008). Eta kare ( $\eta^2$ ) yaygın kullanılan etki büyüklüğü değerlerinden biridir.

- 2- Ya da SPSS aracılığı ile Kısmi Eta Kare hesaplanabilir ☺  
Tekrarlayan Ölçümler için Tek Yönlü Varyans Analizi yapılırken  
Repeated Measures (penceresinde)  
**Options'a tıklanıp, Estimates of effect size seçilir.**

### **Tekrarlayan Ölçümler için Tek Yönlü Varyans Analizi için Post-Hoc Test:**

#### **Post-Hoc testler hangi anlamlılık düzeyi üzerinden değerlendirilir?**

Tekrarlayan Ölçümler için Tek Yönlü Varyans Analizinin sonucu bize tekrarlayan ölçümler arasında farkın olup olmadığını söyler. Farkın kaynağını yani hangi ölçümlerin birbirinden farklılaştığını söylemez! Farkın kaynağını tespit etmek için post-hoc test yapmak gerekir. Post-Hoc test sonucunda hangi ölçümlerin birbirinden farklılaştığını ya da farklılaşmadığını tespit ederiz. Post-Hoc testte Tip I hatadan kaçınmak için Alfa ( $\alpha$ ) anlamlılık düzeyi eşleştirilen grup sayısına (test sayısına) bölünür ve bulunan yeni Alfa ( $\alpha$ ) değeri post-hoc testlerdeki yeni anlamlılık düzeyini ifade eder. Post-Hoc testlerde Tip I hatadan kaçınmak için yeni bir anlamlılık düzeyinin/sınırının hesaplanması ve kullanılmasına Bonferroni düzeltmesi denir (Connolly, 2007; Field, 2009; Pallant, 2017). Bonferroni düzeltmesi post-hoc testlerde Tip I hatadan kaçınmak için anlamlılık sınırında yapılan bir düzeltmedir. Post-Hoc testlerde anlamlılık düzeyinde yapılan bu düzeltme Tip I hatadan kaçınmayı sağlarken, Tip II hataya düşme ihtimalini de arttıracaktır.<sup>20</sup> Pos-Hoc testlerde Bonferroni düzeltmesi yaygın olarak kullanılan bir yöntem olsa da bir zorunluluk değildir.

#### **Tekrarlayan Ölçümler için Tek Yönlü Varyans Analizinin Post-Hoc testleri (Field, 2009):**

**Bonferroni:** Küresellik varsayımı ihlal edildiğinde, özellikle testin gücü ve Tip I hatanın kontrolü açısından tercih edilmelidir. Bonferroni testi oldukça güçlü (robust) bir yöntemdir.

**Sidak:** Bonferroni testine göre daha az tutucu (less conservative) bir testtir.

**LSD:** Fark çıkarma eğiliminde liberal bir testtir. Tip I hataya düşme ihtimali yüksek bir testtir. Field tarafından önerilmeyen bir testtir.

---

<sup>20</sup> Post-Hoc testlerde Alfa ( $\alpha$ ) anlamlılık düzeyi Bonferroni düzeltmesi hesaplamaları için sayfa:56 ve Tip I ve Tip II hata için bakınız sayfa:27

**Bu testin uygulamasına örnek:** Bir grubun depresyona yönelik psikoterapi seansına başladıktan sonra ki 1. ay, 3. ay ve 6. ay depresyon puan ortalamalarını karşılaştırmak istediğimizde, depresyon puan ölçümlerinin hepsi de normal dağılım özelliği gösteriyor ise Tekrarlayan Ölçümler için Tek Yönlü Varyans Analizi yapılır.

**Önemli not:** Bu uygulamada tekrarlayan ölçüm verilerinin normal dağılım özelliği gösterdiği varsayılmıştır. Siz kendi analizlerinizde öncelikli olarak normal dağılıma uygunluğu sınamalısınız.

### SPSS Uygulaması:

- 1- **Analyze, General Linear Model, Repeated Measures –seçilir-**
- 2- **Repeated Measures Define Factor** penceresinde  
**Within\_Subjects Factor Name**'e tekrarlayan ölçüm değişkeninin adı yazılır  
**Number of Levels**'a ölçüm tekrar sayısı yazılır ve **Add** tıklanır.  
Pencerenin sol alt köşesindeki **Define** tıklanır
- 3- **Repeated Measures** penceresinde  
**Within\_Subjects Variables** kutusuna sol taraftaki değişken listesinden tekrarlayan ölçüm değişkenleri taşınır.  
**Options**'a tıklanıp, **Descriptive Statistics** ve **Estimates of effect size** seçilir, **Continue** tıklanır
- 4- **EM Means** tıklanıp, sol taraftan **faktör sağ pencereye (Display means for)** taşınır, **Compare main effects** seçilir, **Confidence interval adjustment**'ten **Bonferroni** ya da **Sidak** seçilir.  
**Continue** tıklanır
- 5- **Repeated Measures** penceresinde  
**OK** Tıklanır.

### SPSS çıktısı:

#### General Linear Model

##### Within-Subjects Factors

Measure: MEASURE\_1

factor1	Variable
1	depresyon1
2	depresyon2
3	depresyon3

##### Descriptive Statistics

	Mean	Std. Deviation	N
depresyon 1. ölçüm puanı	55,8889	17,90425	135
depresyon 2. ölçüm	55,0370	17,56540	135
depresyon 3.ölçüm	51,5704	15,91045	135

##### Descriptive Statistics: Tanımlayıcı İstatistikler

**Mean:** Ortalama. **Std. Deviation:** Standart Sapma. **N:** örneklem sayısı.



Multivariate Tests <sup>a</sup>							
Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
factor1	Pillai's Trace	,670	135,256 <sup>b</sup>	2,000	133,000	,000	,670
	Wilks' Lambda	,330	135,256 <sup>b</sup>	2,000	133,000	,000	,670
	Hotelling's Trace	2,034	135,256 <sup>b</sup>	2,000	133,000	,000	,670
	Roy's Largest Root	2,034	135,256 <sup>b</sup>	2,000	133,000	,000	,670

a. Design: Intercept

Within Subjects Design: factor1

b. Exact statistic

**Multivariate Tests<sup>a</sup>:** Çok değişkenli karşılaştırma. Tekrarlayan ölçüm değerlerinin karşılaştırılması. Küresellik testi sonucunda (aşağıdaki tabloda) p (Sig.) değeri  $p < 0,05$  ise küresellik varsayımı yerine getirilmemiştir ve ANOVA test değerleri olarak Multivariate Tests tablosundan Wilks' Lambda'nın değerleri alınabilir (F değeri ve p değeri). Ancak yaygın olarak test değerleri (F ve p) Tests of Within-Subjects Effects tablosundan alınır.

Mauchly's Test of Sphericity <sup>a</sup>								
Measure: MEASURE_1								
Within Effect	Subjects	Mauchly's W	Approx. Chi-Square	df	Sig.	Greenhouse-Geisser	Epsilon <sup>b</sup> Huynh-Feldt	Lower-bound
factor1		,235	192,623	2	,000	,567	,568	,500

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept

Within Subjects Design: factor1

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

**Mauchly's Test of Sphericity<sup>a</sup>:** Küresellik testi sonuçları.

**Küresellik Testi (Mauchly's Sphericity Test):** Küresellik varsayımının sınandığı test. Ölçümler arası fark varyansların eşit olup olmamasına Mauchly testi ile sınırlanır. Eğer test sonucunda  $p > 0,05$  ise ölçümler arası fark varyansları eşit demektir.

p (Sig.) değeri  $p > 0,05$  ise küresellik varsayımı yerine getirilmiştir ve ANOVA test değerleri olarak Tests of Within-Subjects Effects tablosundan Sphericity Assumed'in değerleri alınır (F değeri ve p değeri).

p (Sig.) değeri  $p < 0,05$  ise küresellik varsayımı yerine getirilmemiştir ve ANOVA test değerleri olarak Tests of Within-Subjects Effects tablosundan Greenhouse-Geisser ya da Huynh-Feldt'in değerleri alınır (F değeri ve p değeri).

Within-Subjects Effects tablosundan Greenhouse-Geisser'in mı yoksa Huynh-Feldt'in değerlerini mi (F değeri ve p değeri) kullanıp kullanmayacağımıza da yukarıda ki tablodaki (Mauchly's Sphericity Test) Epsilon'un Greenhouse-Geisser değerine göre karar veririz. Epsilon'un Greenhouse-Geisser değeri 0,75'ten küçük ise "Tests of Within-Subjects Effects" tablosundan Greenhouse-Geisser'in F ve p değerleri ANOVA test değerleri olarak kullanılır. "Mauchly's Sphericity Test" tablosunda ki Epsilon'un Greenhouse-Geisser değeri 0,75'ten büyük ise "Tests of Within-Subjects Effects" tablosundan Huynh-Feldt'in F ve p değerleri ANOVA test değerleri olarak kullanılır.

Bu analiz için küresellik testi sonucu p (Sig.) değeri  $p < 0,05$  olduğu için küresellik sağlanmamıştır. Elipson'un Greenhouse-Geisser değeri 0,75'ten küçük olduğu için "Tests of Within-Subjects Effects" tablosundan Greenhouse-Geisser'in F ve p değerleri ANOVA test değerleri olarak kullanılır.

#### Tests of Within-Subjects Effects

Measure: MEASURE\_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
factor1	Sphericity Assumed	1412,686	2	706,343	231,235	,000	,633
	Greenhouse-Geisser	1412,686	1,133	1246,716	231,235	,000	,633
	Huynh-Feldt	1412,686	1,136	1243,271	231,235	,000	,633
	Lower-bound	1412,686	1,000	1412,686	231,235	,000	,633
Error(factor1)	Sphericity Assumed	818,647	268	3,055			
	Greenhouse-Geisser	818,647	151,839	5,392			
	Huynh-Feldt	818,647	152,260	5,377			
	Lower-bound	818,647	134,000	6,109			

**Tests of Within-Subjects Effects:** Tekrarlayan ölçümlerin grup içi karşılaştırması. Küresellik testi sonuçlarına göre ANOVA test değerleri buradan alınır.

**factor1:** tekrarlayan ölçümler (faktör)

**Error(factor1):** tekrarlayan ölçümlerin hatası

**Type III Sum of Squares:** 3 ölçümün kareler toplamı

**df:** Serbestlik derecesi

**Mean Square:** Kareler ortalaması

**F:** test değeri

**Sig. :** testin p değeri

**Partial Eta Squared:** Eta kare (kısmi)

### Tests of Within-Subjects Contrasts

Measure: MEASURE\_1

Source	factor1	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
factor1	Linear	1258,848	1	1258,848	260,860	,000	,661
	Quadratic	153,838	1	153,838	119,854	,000	,472
Error(factor1)	Linear	646,652	134	4,826			
	Quadratic	171,995	134	1,284			

**Tests of Within-Subjects Contrasts:** Tekrarlayan ölçümlerin zıtlığının karşılaştırması.

### Estimates

Measure: MEASURE\_1

factor1	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	55,889	1,541	52,841	58,937
2	55,037	1,512	52,047	58,027
3	51,570	1,369	48,862	54,279

**Estimates:** Tahminler

**Mean:** Ortalama. **Std. Error:** Standart hata. **95% Confidence Interval for Mean Lower-**

**Upper Bound:** %95 Güven Aralığında ortalamanın Alt-Üst sınır değeri

### Pairwise Comparisons

Measure: MEASURE\_1

(I) factor1	(J) factor1	Mean Difference (I-J)	Std. Error	Sig. <sup>b</sup>	95% Confidence Interval for Difference <sup>b</sup>	
					Lower Bound	Upper Bound
1	2	,852*	,080	,000	,657	1,047
	3	4,319*	,267	,000	3,670	4,967
2	1	-,852*	,080	,000	-1,047	-,657
	3	3,467*	,240	,000	2,884	4,050
3	1	-4,319*	,267	,000	-4,967	-3,670
	2	-3,467*	,240	,000	-4,050	-2,884

Based on estimated marginal means

\*. The mean difference is significant at the ,05 level.

b. Adjustment for multiple comparisons: Bonferroni.

**Pairwise Comparisons:** Tekrarlayan Ölçümlerde Tek Yönlü Varyans Analizinin post-hoc test sonuçlarını gösterir. ANOVA’da fark çıkmış ise hangi ölçüm düzeyleri arasında fark olup olmadığını bu tablodaki p (Sig.) değerinden anlarız.

**Tablo: Araştırma grubunda tekrarlayan depresyon puanlarının karşılaştırılması**

	n	Ortalama	Standart sapma	F	p	Anlamlı fark*
Depresyon 1. Ölçüm	135	55,8889	17,90425	231,235	0,000	1-2
Depresyon 2. Ölçüm	135	55,0370	17,56540			1-3
Depresyon 3.ölçüm	135	51,5704	15,91045			2-3

\* Bonferroni testi yapılmıştır. Bonferroni düzeltmesi sonrasında  $p < 0,017$  düzeyinde fark çıkan sonuçlar.

Araştırma grubunda katılımcıların Birinci depresyon ölçüm puan ortalaması  $56 \pm 18$ , İkinci depresyon ölçüm puan ortalaması  $55 \pm 18$  iken üçüncü depresyon ölçüm puan ortalaması ise  $52 \pm 16$ 'dır. Araştırma grubunda katılımcıların tekrarlayan depresyon puan ortalamaları arasında istatistiksel olarak anlamlı bir farklılık tespit edilmiştir ( $p < 0,01$ ). Tespit edilen bu farkın etki büyüklüğü Eta kare<sub>kısmi</sub> ( $\eta^2_{kısmi}$ ) = 0,63 olduğu, farkın büyük bir etkiye sahip olduğu ve toplam varyansın %63'ünün faktör (psikoterapi) tarafından açıklandığını göstermektedir. Tekrarlayan ölçüm düzeyleri puan ortalamalarının hepsinin birbirinden istatistiksel olarak anlamlı bir şekilde farklılaştığı tespit edilmiştir ( $p < 0,01$ ).

#### Etki büyüklüğü:

##### Manuel hesaplanması:

$$\text{Eta kare}_{(kısmi)} = \frac{\text{Kareler toplamı (ölçüm)}}{\text{Kareler toplamı (ölçüm)} + \text{Kareler toplamı (ölçüm hatası)}}$$

$$\text{Eta kare}_{(kısmi)} = \frac{1.412,686}{1.412,686 + 818,647} = 0,63$$

**Eta kare**<sub>(kısmi)</sub> ( $\eta^2_{kısmi}$ ) = 0,63 (büyük etki) Bu durumda etki toplam varyansın %43'ünü açıklar.

##### SPSS hesaplaması:

Tests of Within-Subjects Effects tablosundaki **Partial Eta Squared** = 0,633

**Eta kare**<sub>(kısmi)</sub> = 0,63 (büyük etki)

$\eta^2_{kısmi}$  = 0,63 (büyük etki) Bu durumda etki toplam varyansın %63'ünü açıklar.

## 2.9. Friedman Testi

Tekrarlayan Ölçümler için Tek Yönlü Varyans Analizi'nin Non-Parametrik testidir. Tekrarlayan Ölçümler için Tek Yönlü Varyans Analizi'nin varsayımlarının sağlanamaması durumunda bir grubun tekrarlayan ikiden fazla ölçüm puanları arasında istatistiksel olarak anlamlı bir farklılığın olup olmadığı bu test ile sıranabilir.

Friedman Testi, her bir vaka için tekrarlayan üç ya da daha fazla ölçüm düzeyleri puanlarını en küçüğünden (1'den başlayarak) en büyüğüne doğru sıra sayısına dönüştürür. Her bir vaka için ölçümler sıra sayısına dönüştürüldükten sonra her bir ölçüm düzeyinin sıra ortalaması değerleri hesaplanır. Test edilen durum ölçüm düzeyleri sıra ortalaması değerleri arasındaki farkın istatistiksel analizidir (Corder ve Foreman, 2009).

**Friedman testinin sıra değerlerini üretme ve işlem süreci örneği:**

Veri setimiz				Friedman testinin işlem süreci		
Anket no	I.Ölçüm	II. Ölçüm	III. Ölçüm	I.Ölçüm Sıra Değeri	II. Ölçüm Sıra Değeri	III. Ölçüm Sıra Değeri
1	62	60	58	3	2	1
2	65	67	70	1	2	3
3	48	47	45	3	2	1
4	63	70	75	1	2	3
5	65	76	86	1	2	3
6	56	51	57	2	1	3
7	73	88	95	1	2	3
8	74	90	91	1	2	3
9	58	72	75	1	2	3
10	62	75	74	1	3	2

	n	Sıra Ortalaması (Mean Rank)	Sıralar Toplamı (Sum of Ranks)
I Ölçüm	10	1,5	15
II. Ölçüm	10	2,0	20
III. Ölçüm	10	2,5	25

Friedman Testi tekrarlayan ölçümlerin medyanlarını karşılaştırmaz. Farklılığın hesaplanmasında medyanların bir işlevi yoktur, dolayısı ile bu testin sonucunda ulaşılan tekrarlayan ölçümlerin birbirinden farklılaşmasını ya da farklılaşmamasını, medyanlar üzerinden açıklanması ve yorumlanması yanlıştır!

Friedman testinin sonucu bize tekrarlayan ölçümler arasında fark olup olmadığını söyler. Farkın kaynağını yani hangi ölçümlerin birbirinden farklılaştığını söylemez! Farkın kaynağını tespit etmek için post-hoc test yapmak gerekir. Post-Hoc test sonucunda hangi ölçümlerin birbirinden farklılaştığını ya da farklılaşmadığını tespit ederiz.

**Test neyi yapar:**

Bir grubun tekrarlayan üç ya da daha fazla ölçüm puanlarını karşılaştırır.

**Varsayımları:**

Bağımlı değişken en az sıralama ölçeğinde bir değişkendir.

Tekrarlayan (üç ya da daha fazla) ölçümler bir (aynı) gruba aittir.

**Parametrik Testi:** Tekrarlayan Ölçümler için Tek Yönlü Varyans Analizi

**Hipotezleri:**

$H_0$ : Tekrarlayan ölçüm puanları arasında fark yoktur. (sınanan hipotez)

$H_1$ : Tekrarlayan ölçüm puanları arasında fark vardır.

**Test sonucu:**

$p > 0,05$  ise  $H_0$  hipotezi kabul edilir: Tekrarlayan ölçüm puanları arasında fark yoktur.

$p < 0,05$  ise  $H_0$  hipotezi ret edilir: Tekrarlayan ölçüm puanları arasında fark vardır.

**Etki Büyüklüğü:**

Yapılan hipotez testi bize tekrarlayan ölçümlerin sıra ortalaması arasında istatistiksel olarak anlamlı bir fark olup olmadığını söyler, ancak fark var ise bu farkın etkisi ya da büyüklüğü konusunda bilgi vermez. Etki büyüklüğü bize etkinin önemi konusunda objektif bir ölçüm sağladığı için faydalıdır (Field, 2009). Etki büyüklüğü sonuçların yorumlanmasında dikkate alınması gereken istatistiksel bir değerdir. Hipotez testlerinde etki büyüklüğü değerlerinin de hesaplanarak yorumlanması sonuçların anlaşılabilirliğini arttıracaktır (Büyüköztürk, 2010). Hipotez testinde fark bulunduğu durumda etki büyüklüğüne bakılır. Hipotez testleri için hesaplanan farklı etki büyüklüğü değerleri söz konusudur. Friedman Testi için kullanılan etki büyüklüğü değeri:

**Kendall'in W Katsayısı:** Friedman testinin etki büyüklüğü istatistiği olarak Kendall'in W (Kendal'in uyum katsayısı/ Kendall's coefficient of concordance) değeri kullanılabilir (Green ve Salkind, 2014; Rovai ve Diğr., 2014; Tomczak ve Tomczak, 2014; Mangiafico, 2016; Malhotra, 2018). Kendal'in W değeri 0 ile 1 arasında değer almaktadır, yüksek değerler daha güçlü ilişkiyi göstermektedir (Green ve Salkind, 2014). Kendal'in W katsayısı şu şekilde yorumlanabilir: 0,00 – 0,20 = çok zayıf etki, 0,21 – 0,40 = zayıf etki, 0,41 – 0,60 = orta (moderate) etki, 0,61 – 0,80 = güçlü etki, 0,81 – 1,00 = çok güçlü etki (Rovai ve Diğr., 2014).

SPSS aracılığı ile Kendall'in W Katsayısının hesaplanması:

Friedman testini yaparken **Test for Several Related Samples** penceresinde **Kendall's W** seçilir.

## Friedman Testinin Post-Hoc Testi:

### Post-Hoc testler hangi anlamlılık düzeyi üzerinden değerlendirilir?

Friedman Testinin sonucu bize tekrarlayan ölçümler arasında farkın olup olmadığını söyler. Farkın kaynağını yani hangi ölçümlerin birbirinden farklılaştığını söylemez! Farkın kaynağını tespit etmek için post-hoc test yapmak gerekir. Post-Hoc test sonucunda hangi ölçümlerin birbirinden farklılaştığını ya da farklılaşmadığını tespit ederiz. Post-Hoc testte Tip I hatadan kaçınmak için Alfa ( $\alpha$ ) anlamlılık düzeyi eşleştirilen grup sayısına (test sayısına) bölünür ve bulunan yeni Alfa ( $\alpha$ ) değeri post-hoc testlerdeki yeni anlamlılık düzeyini ifade eder. Post-Hoc testlerde Tip I hatadan kaçınmak için yeni bir anlamlılık düzeyinin/sınırının hesaplanması ve kullanılmasına Bonferroni düzeltmesi denir (Connolly, 2007; Field, 2009; Pallant, 2017). Bonferroni düzeltmesi post-hoc testlerde Tip I hatadan kaçınmak için anlamlılık sınırında yapılan bir düzeltmedir. Post-Hoc testlerde anlamlılık düzeyinde yapılan bu düzeltme Tip I hatadan kaçınmayı sağlarken, Tip II hataya düşme ihtimalini de arttıracaktır.<sup>21</sup> Post-Hoc testlerde Bonferroni düzeltmesi yaygın olarak kullanılan bir yöntem olsa da bir zorunluluk değildir.

### Friedman Testinin Post-Hoc testi: Wilcoxon İşaretli Sıralar Testi

**Bu testin uygulamasına örnek:** Bir grubun depresyona yönelik psikoterapi seansına başladıktan sonra ki 1. ay, 3. ay ve 6. ay depresyon puan ortalamalarını karşılaştırmak istediğimizde, depresyon puan ölçümleri normal dağılım özelliği göstermiyor ise Friedman Testi yapılır.

### SPSS Uygulaması:

- 1- **Analyze, Nonparametric Tests, Legacy Dialogs, K Related Samples** seçilir.
- 2- **Test for Several Related Samples** penceresinde  
**Test Variables**'e sol taraftaki değişken listesinden tekrarlayan ölçüm değişkenleri taşınır.
- 3- (İstenirse **Statistisc** butonu tıklanıp açılan pencereden **Quartiles** seçilir)
- 4- **Test for Several Related Samples** penceresinde  
–**Friedman** testi otomatik seçilidir-  
**Kendall's W** seçilir (fark çıkması durumunda etki büyüklüğünü görmek için).  
**OK** Tıklanır.

<sup>21</sup> Post-Hoc testlerde Alfa ( $\alpha$ ) anlamlılık düzeyi Bonferroni düzeltmesi hesaplamaları için sayfa:56 ve Tip I ve Tip II hata için bakınız sayfa:27

## SPSS çıktısı:

## NPar Tests

## Descriptive Statistics

	N	Percentiles		
		25th	50th (Median)	75th
depresyon 1. ölçüm puanı	135	38,0000	55,0000	70,0000
depresyon 2. Ölçüm	135	36,0000	55,0000	67,0000
depresyon 3.ölçüm	135	35,0000	52,0000	62,0000

## Descriptive Statistics: Tanımlayıcı İstatistikler

**N:** örneklem sayısı. **25th:** %25. Değer. **50th (Median):** %50. Değer (Ortanca). **75th:** %75. Değer.

## Friedman Test

## Ranks

	Mean Rank
depresyon 1. ölçüm puanı	2,77
depresyon 2. Ölçüm	2,12
depresyon 3.ölçüm	1,11

**Mean Rank:** Sıra Ortalaması

Test Statistics<sup>a</sup>

N	135
Chi-Square	222,463
df	2
Asymp. Sig.	,000

a. Friedman Test

**Chi-Square:** Ki kare değeri (Friedman testinin değeri)

**df:** serbestlik derecesi. **Asymp. Sig. :** testin p değeri.

## Kendall's W Test

## Ranks

	Mean Rank
depresyon 1. ölçüm puanı	2,77
depresyon 2. Ölçüm	2,12
depresyon 3.ölçüm	1,11



**Test Statistics**

N	135
Kendall's W <sup>a</sup>	,824
Chi-Square	222,463
df	2
Asymp. Sig.	,000

a. Kendall's Coefficient of Concordance

**Kendall's W:** Kendall'in W Katsayısı (Friedman testinin etki büyüklüğü değeri)

Fark çıktığı için post-hoc test olarak Wilcoxon İşaretli Sıralar Testi (Wilcoxon Signed Ranks Test) yapıyoruz:

**SPSS Uygulaması:**

- 1- **Analyze, Nonparametric Tests, Legacy Dialogs, 2 Related Samples** seçilir.
- 2- **2 Related Samples Test** penceresinde sol taraftaki değişken listesinden tekrarlayan ölçüm değişkenleri **Test Pairs'e** –eşleştirilerek- **taşınır**.
- 3- (İstenirse **Options** butonu tıklanıp açılan pencereden **Quartiles** seçilir)
- 4- **2 Related Samples Test** penceresinde – Wilcoxon testi otomatik seçilidir- **OK** Tıklanır.

**SPSS çıktısı:****NPar Tests****Wilcoxon Signed Ranks Test**

		Ranks		
		N	Mean Rank	Sum of Ranks
depresyon 2. ölçüm - depresyon 1. ölçüm puanı	Negative Ranks	85 <sup>a</sup>	44,26	3762,00
	Positive Ranks	2 <sup>b</sup>	33,00	66,00
	Ties	48 <sup>c</sup>		
	Total	135		
depresyon 3.ölçüm - depresyon 1. ölçüm puanı	Negative Ranks	125 <sup>d</sup>	63,00	7875,00
	Positive Ranks	0 <sup>e</sup>	,00	,00
	Ties	10 <sup>f</sup>		
	Total	135		
depresyon 3.ölçüm - depresyon 2. ölçüm	Negative Ranks	117 <sup>g</sup>	60,86	7121,00
	Positive Ranks	2 <sup>h</sup>	9,50	19,00
	Ties	16 <sup>i</sup>		
	Total	135		

a. depresyon 2. ölçüm < depresyon 1. ölçüm puanı

Test Statistics <sup>a</sup>			
	depresyon 2. ölçüm - depresyon 1. ölçüm puanı	depresyon 3.ölçüm - depresyon 1. ölçüm puanı	depresyon 3.ölçüm - depresyon 2. ölçüm
Z	-8,263 <sup>b</sup>	-9,761 <sup>b</sup>	-9,453 <sup>b</sup>
Asymp. Sig. (2-tailed)	,000	,000	,000

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

**Tablo: Araştırma grubunda tekrarlayan depresyon puanlarının karşılaştırılması**

Depresyon puanları	n	Sıra ortalaması	p	$\chi^2$	Anlamlı fark*
1. Ölçüm	135	2,77	0,000	222,463	1-2
2. Ölçüm	135	2,12			1-3
3. Ölçüm	135	1,11			2-3

\* Wilcoxon İşaretli Sıralar Testi yapılmıştır. Bonferroni düzeltmesi sonrasında  $p < 0,017$  düzeyinde fark çıkan sonuçlar.

Araştırma grubunda katılımcıların Birinci depresyon ölçüm puanı sıra ortalaması 2,77, İkinci depresyon ölçüm puanı sıra ortalaması 2,12 iken Üçüncü depresyon ölçüm puanı sıra ortalaması ise 1,11'dir. Araştırma grubunda katılımcıların tekrarlayan depresyon puanı sıra ortalamaları arasında istatistiksel olarak anlamlı bir farklılık tespit edilmiştir ( $p < 0,01$ ). Tespit edilen bu farkın etki büyüklüğü  $W = 0,82$  olduğu, farkın çok güçlü bir etkiye sahip olduğunu göstermektedir. Tekrarlayan ölçüm düzeyleri puanları sıra ortalamalarının hepsinin birbirinden istatistiksel olarak anlamlı bir şekilde farklılaştığı tespit edilmiştir ( $p < 0,01$ ).

### Etki büyüklüğü:

Friedman testi için Kendall'in W Katsayısı

Kendall's W: ,824 (Çok güçlü etki)

## 3.BÖLÜM

# ORAN KARŞILAŞTIRMALARI

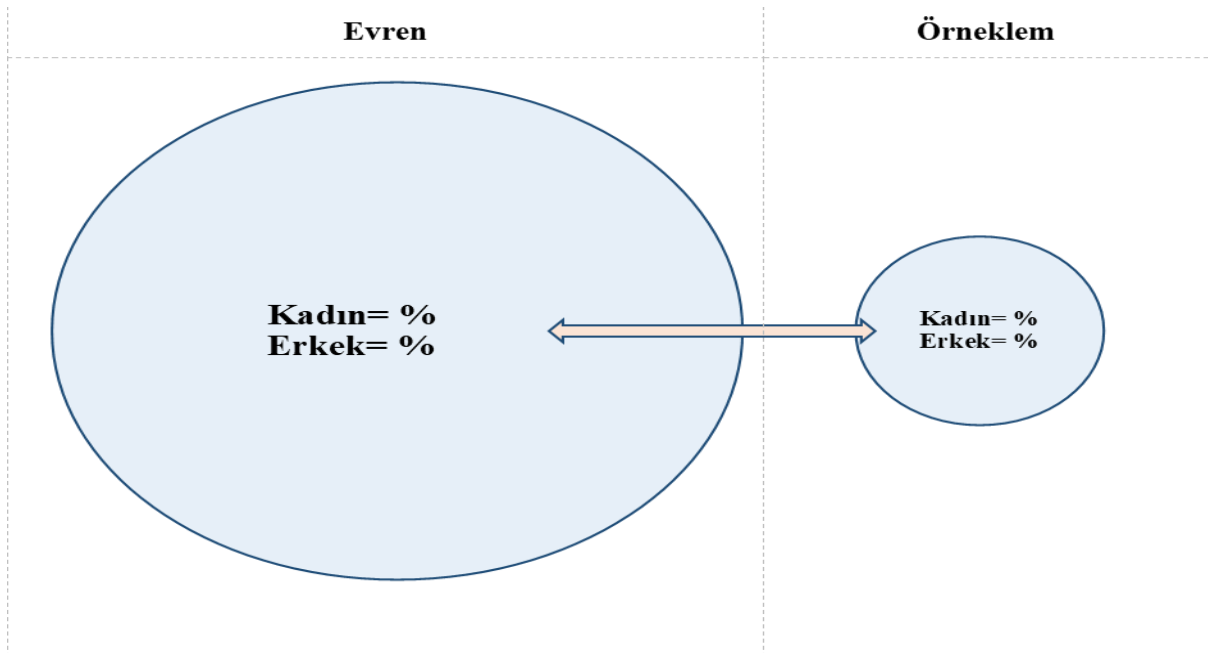
### 3.1. Ki Kare Testleri

#### 3.1.1. Tek Örneklem Ki Kare Testi <sup>22</sup>

Tek Örneklem Ki Kare Testi, bir kategorik değişkenin alt kategorilerinin örneklemde görülme oranları ile beklenen oranlarını ya da Evrende görülme oranlarını karşılaştırmak için kullanılan bir iyi uyum testidir (Pallant, 2005). Test kategorik değişkenin her bir alt kategorisinde gözlenen sayıların, alt kategoriler için beklenen sayılardan<sup>23</sup> farkın anlamlılığını sınar (Büyüköztürk, 2010).

#### Test neyi yapar:

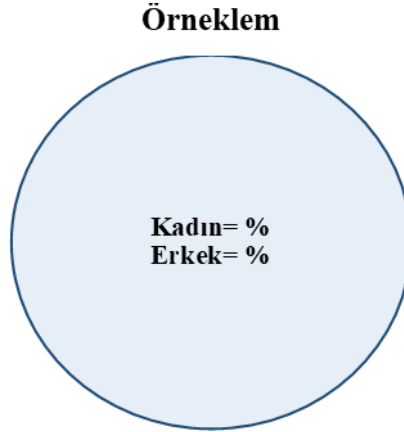
İki amaçla kullanılır: I- Herhangi bir kategorik değişkenin Evrende (ya da bir başka çalışmada) görülme oranları ile Örneklemde görülme oranları arasında uyum sınanmak istendiğinde:



<sup>22</sup> Ki Kare Uyum Testi/ Chi-Square test for goodness of fit

<sup>23</sup> Tek örneklem kategorik değişkeninin alt kategorilerinin eşit dağılımı sınıandığında beklenen değer, örneklem sayısının alt kategori sayısına bölünmesi ile bulunur. Kategorik değişkenin örneklemde görülme sayıları ile evrende görülme oranları karşılaştırıldığında, beklenen değer kategorik değişkenin evrende görülme oranları ile örneklem sayısı üzerinden hesaplanır.

- II- Herhangi bir kategorik değişkenin alt kategorilerinin örneklem içindeki oranlarının eşit dağılıp dağılmadığı sınanmak istendiğinde:



### Varsayımları:

Test edilecek değişken kategorik bir değişkendir.

Beklenen değeri 5'ten küçük olan kategori sayısının toplam kategori sayısının %20'sini aşmaması ve tüm kategorilerde bu sayının 1'den büyük olması gerekir. Bu koşulun sağlanmadığı durumlarda, mantıklı ise kategori birleştirme yoluna gidilebilir (Büyüköztürk, 2010; Mangiafico, 2016).

**Parametrik testi:** Yok

- I- Herhangi bir kategorik değişkenin Evrende (ya da bir başka çalışmada) görülme oranları ile Örneklemde görülme oranları arasında uyum sınanmak istendiğinde:

### Hipotezleri:

$H_0$ : Kategorik değişkenin Evrende ve Örneklemde ki görülme oranları arasında fark yoktur.

$H_1$ : Kategorik değişkenin Evrende ve Örneklemde ki görülme oranları arasında fark vardır.

### Test sonucu:

$p > 0,05$  ise  $H_0$  hipotezi kabul edilir: Kategorik değişkenin Evrende ve Örneklemde ki görülme oranları arasında fark yoktur.

$p < 0,05$  ise  $H_0$  hipotezi ret edilir: Kategorik değişkenin Evrende ve Örneklemde ki görülme oranları arasında fark vardır.

### Etki Büyüklüğü:

Yapılan hipotez testi bize kategorik değişken için Örneklemde gözlenen değerler (oranlar) ile Evrendeki değerler (oranlar) arasında istatistiksel olarak anlamlı bir fark olup olmadığını söyler, ancak fark var ise bu farkın etkisi ya da büyüklüğü konusunda bilgi vermez. Etki büyüklüğü bize etkinin önemi konusunda objektif bir ölçüm sağladığı için faydalıdır (Field,

2009). Etki büyüklüğü sonuçların yorumlanmasında dikkate alınması gereken istatistiksel bir değerdir. Hipotez testlerinde etki büyüklüğü değerlerinin de hesaplanarak yorumlanması sonuçların anlaşılabilirliğini arttıracaktır (Büyüköztürk, 2010). Hipotez testinde fark bulunduğu durumda etki büyüklüğüne bakılır. Hipotez testleri için hesaplanan farklı etki büyüklüğü değerleri söz konusudur. Tek Örneklem Ki Kare Testi için kullanılan etki büyüklüğü değeri:

Tek Örneklem Ki Kare Testinin etki büyüklüğü 0 ile 1 arasında bir değer alır. Etki büyüklüğünün 0 olması gözlenen dağılımın beklenen dağılım ile tam olarak aynı olduğunu, 1 olması ise iki dağılımın birbirinden mümkün olduğunca farklı olduğunu gösterir. Etki Büyüklüğü aşağıdaki formülle hesaplanır (Green ve Salkind, 2014):

$$\text{Etki Büyüklüğü} = \frac{\chi^2}{n \times (\text{kategori sayısı}-1)}$$

**Bu testin uygulamasına örnek:** Örneklemde görülen sigara içme oranları ile Evrende ki sigara içme oranlarını (Evrende sigara içmeyenlerin %60, içenlerin %40 olduğunu farz edelim) karşılaştırmak istediğimizde Tek Örneklem Ki Kare Testi yaparız.

#### SPSS Uygulaması:

- 1- **Analyze, Nonparametric Tests, Legacy Dialogs, Chi-Square** seçilir.
- 2- **Chi-Square Test** penceresinde  
Sol taraftaki değişken listesinden **kategorik değişken Test Varibale List'e taşınır.**
- 3- **Expected Values** kısmında **Values seçeneği tıklanır.** Values kutusuna sırasıyla kategorik değişkenin 1. alt kategorisinin (1 kodu olarak girdiğimiz kategorinin) Evrende görülme oranı yazılır (bu örnekte: 0,60 yani evrende görülme oranı %60). **Add tıklanır.**  
**Ardından**  
kategorik değişkenin 2. alt kategorisinin (2 kodu olarak girdiğimiz kategorinin) Evrende görülme oranı yazılır (bu örnekte 0,40 yani evrende görülme oranı %40). **Add tıklanır.** (Eğer 2'den fazla alt kategori var ise bu şekilde evrende görülme oranları sırası ile girilmeye devam edilir.)
- 4- **OK Tıklanır.**

**SPSS çıktısı:**  
**NPar Tests**

sigara oyküsü			
	Observed N	Expected N	Residual
yok	73	81,0	-8,0
var	62	54,0	8,0
Total	135		

**Observed N:** Gözlenen sayı

**Expected N:** Beklenen sayı

**Residual:** Fark

**Test Statistics**

sigara oyküsü	
Chi-Square	1,975 <sup>a</sup>
df	1
Asymp. Sig.	,160

a. 0 cells (0,0%) have expected frequencies less than 5. The minimum expected cell frequency is 54,0.

**Chi-Square:** test değeri (Ki kare değeri)

**df:** serbestlik derecesi

**Asymp. Sig. :** testin p değeri

**Tablo: Örneklemde ki sigara içme oranların Evren oranları ile karşılaştırılması**

Sigara	Gözlenen		Beklenen	
	n	% <sup>24</sup>	n	%
İçiyor	73	54,07	81	60,00
İçmiyor	62	45,93	54	40,00
<b>Toplam</b>	135	100,00		100,00
$\chi^2=1,975$ $p=0,160$				

Araştırma grubundaki katılımcıların %54,07'si sigara içiyor, %45,93'ü sigara içmiyor iken Evrendeki sigara içme oranı %60 ve içmeme oranı ise %40'tır. Araştırma grubunun gözlenen sigara içme oranı ile Evrenin sigara içme oranı arasında istatistiksel olarak anlamlı bir farklılık tespit edilmemiştir ( $p>0,05$ ).

**Etki Büyüklüğü:**

Etki büyüklüğü fark çıkan durumlarda hesaplanır! Burada yapılan sadece formülün işleyişine (hangi değerin nereden alındığına) yönelik bir örnek hesaplama yapılmıştır.

$$\text{Etki Büyüklüğü} = \frac{\chi^2}{n \times (\text{kategori sayısı} - 1)} = \frac{1,975}{135 \times (2-1)} = 0,015$$

<sup>24</sup> Bu tablodaki yüzdeler sonradan hesaplanmıştır. SPSS bu yüzdeleri üretmez.

## II- Herhangi bir kategorik değişkenin alt kategorilerinin örneklem içindeki oranlarının eşit dağılıp dağılmadığı test edilmek istendiğinde:

### Hipotezleri:

$H_0$ : Kategorik değişkenin Örneklemdeki görülme oranları arasında fark yoktur.

$H_1$ : Kategorik değişkenin Örneklemdeki görülme oranları arasında fark vardır.

### Test sonucu:

$p > 0,05$  ise  $H_0$  hipotezi kabul edilir: Kategorik değişkenin Örneklemdeki görülme oranları arasında fark yoktur.

$p < 0,05$  ise  $H_0$  hipotezi ret edilir: Kategorik değişkenin Örneklemdeki görülme oranları arasında fark vardır.

### Etki Büyüklüğü:

Yapılan hipotez testi bize gözlenen ve beklenen değerler arasında istatistiksel olarak anlamlı bir fark olup olmadığını söyler, ancak fark var ise bu farkın etkisi ya da büyüklüğü konusunda bilgi vermez. Etki büyüklüğü bize etkinin önemi konusunda objektif bir ölçüm sağladığı için faydalıdır (Field, 2009). Etki büyüklüğü sonuçların yorumlanmasında dikkate alınması gereken istatistiksel bir değerdir. Hipotez testlerinde etki büyüklüğü değerlerinin de hesaplanarak yorumlanması sonuçların anlaşılabilirliğini arttıracaktır (Büyüköztürk, 2010). Hipotez testinde fark bulunduğu durumda etki büyüklüğüne bakılır. Hipotez testleri için hesaplanan farklı etki büyüklüğü değerleri söz konusudur. Tek Örneklem Ki Kare Testi için kullanılan etki büyüklüğü değeri:

Tek Örneklem Ki Kare Testinin etki büyüklüğü 0 ile 1 arasında bir değer alır. Etki büyüklüğünün 0 olması gözlenen dağılımın beklenen dağılım ile tam olarak aynı olduğunu, 1 olması ise iki dağılımın birbirinden mümkün olduğunca farklı olduğunu gösterir. Etki Büyüklüğü aşağıdaki formülle hesaplanır (Green ve Salkind, 2014):

$$\text{Etki Büyüklüğü} = \frac{\chi^2}{n \times (\text{kategori sayısı} - 1)}$$

**Bu testin uygulamasına örnek:** Örneklemde görülen cinsiyet dağılımının (kadın ve erkek oranının) eşit dağılıp dağılmadığını sınamak istediğimizde Tek Örneklem Ki Kare Testi yaparız.

#### SPSS Uygulaması:

- 1- **Analyze, Nonparametric Tests, Legacy Dialogs, Chi-Square** seçilir.
- 2- **Chi-square Test** penceresinde  
Sol taraftaki değişken listesinden **kategorik değişken Test Varibale List'e** taşınır.
- 3- **Expected Values** kısmında **All categories equal** seçilidir.
- 4- **OK** Tıklanır.

#### SPSS çıktısı:

#### NPar Tests

#### Chi-Square Test

#### Frequencies

cinsiyet			
	Observed N	Expected N	Residual
kadın	72	67,5	4,5
erkek	63	67,5	-4,5
Total	135		

**Observed N:** Gözlenen sayı

**Expected N:** Beklenen sayı

**Residual:** Fark

#### Test Statistics

cinsiyet	
Chi-Square	,600 <sup>a</sup>
df	1
Asymp. Sig.	,439

a. 0 cells (0,0%) have expected frequencies less than 5. The minimum expected cell frequency is 67,5.

**Chi-Square:** test değeri (Ki kare değeri)

**df:** serbestlik derecesi

**Asymp. Sig. :** testin p değeri



**Tablo: Örneklem grubunda cinsiyet dağılımının karşılaştırılması**

Cinsiyet	Gözlenen		Beklenen	
	n	% <sup>25</sup>	n	%
<b>Kadın</b>	72	53,33	67,5	50,00
<b>Erkek</b>	63	46,67	67,5	50,00
<b>Toplam</b>	135	100,00	135	100,00
$\chi^2=0,600$ $p=0,439$				

Araştırma grubundaki katılımcıların %53,33'ü kadın ve %46,67'si erkektir, bu örneklemde ki kadın ve erkeklerin eşit dağılımının %50 %50 olması beklenirdi. Araştırma grubunun gözlenen ve beklenen cinsiyet dağılımı oranları arasında istatistiksel olarak anlamlı bir farklılık tespit edilmemiştir ( $p>0,05$ ).

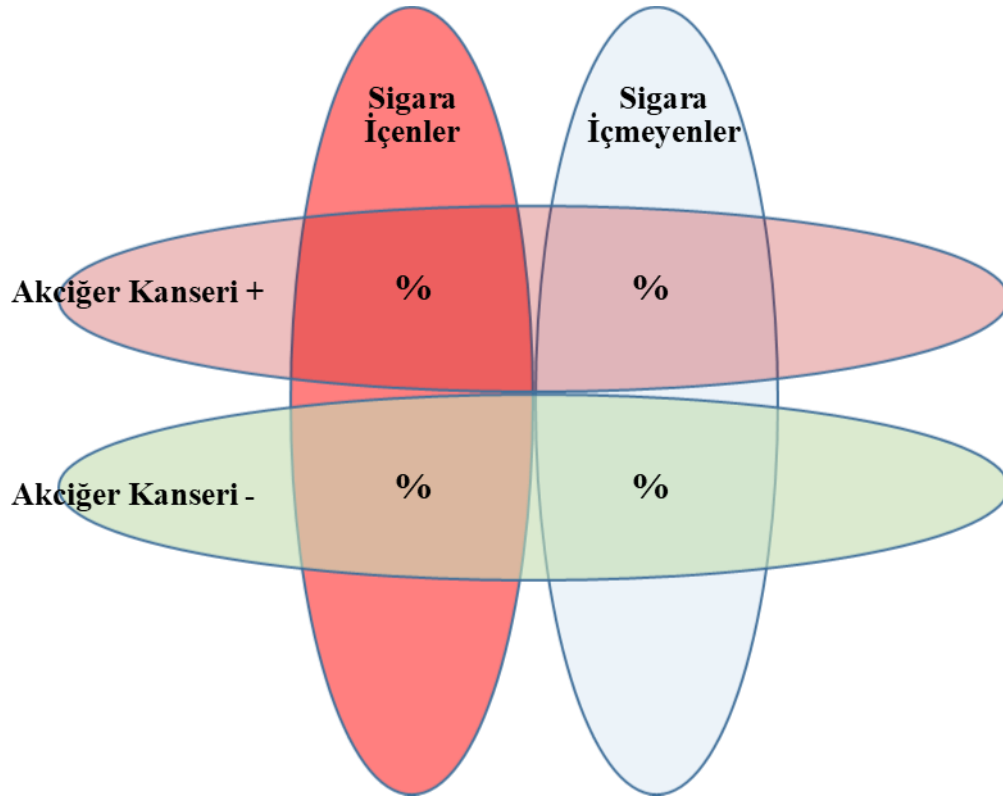
### Etki Büyüklüğü:

Etki büyüklüğü fark çıkan durumlarda hesaplanır! Burada yapılan sadece formülün işleyişine (hangi değerin nereden alındığına) yönelik bir örnek hesaplama yapılmıştır.

$$\text{Etki Büyüklüğü} = \frac{\chi^2}{n \times (\text{kategori sayısı} - 1)} = \frac{0,6}{135 \times (2-1)} = 0,004$$

<sup>25</sup> Bu tablodaki yüzdeler sonradan hesaplanmıştır. SPSS bu yüzdeleri üretmez.

### 3.1.2. Ki Kare Bağımsızlık Testi



Ki Kare Bağımsızlık Testi, iki kategorik<sup>26</sup> değişken arasında ki ilişkiyi sınavan bir testtir. Bu değişkenlerin her biri iki ya da daha fazla kategoriye sahip olabilir. Bu test kategorilerin her birinde yer alan gözlenen değerleri (frekansları), beklenen değerlerle (frekanslarla) kıyaslar (Pallant, 2017). Test iki kategorik değişkenin kesişim kümelerini ifade eden çapraz tablolara dayanır. Bu çapraz tablonun her bir kesişen kategorisi bu tablonun hücrelerini oluşturur (Can, 2019). Test edilen durum ise her bir hücrede gözlenen değerler (frekanslar) ile beklenen değerlerin (frekansların) karşılaştırılmasıdır. Oluşturulan çapraz tablonun her bir hücresindeki gözlenen değerlerle, beklenen değerlerin birbirlerinden anlamlı bir şekilde farklılık gösterip göstermediği test edilir. Buna göre iki değer (gözlenen-beklenen) arasındaki fark arttıkça değişkenler arasında ki ilişkinin anlamlı çıkma olasılığı artar (Büyüköztürk, 2010).

Gözlenen değerler örneklemdaki gerçek durumu ifade eden katılımcı sayılarını ifade eder. Beklenen değerler ise  $H_0$  hipotezinin iddiası doğrultusunda çapraz tablonun her bir hücresine “bir nevi” eşit dağılım değerlerinin atanmasıdır. Beklenen değerler tam olarak eşit dağılım değerleri olmamakla birlikte değişkenler arasında bağımsızlığı/ilişkisizliği sağlayacak teorik değerlerdir. Gözlenen değerlerin beklenen değerlerden farkını sınavan bu test, farklılığı anlamlı bulduğunda iki kategorik değişken arasında anlamlı bir ilişki olduğunu gösterir. Bu durum aynı zamanda bir kategorik değişkenin (bağımsız) diğer kategorik değişken (bağımlı) üzerinde etkisi olduğunu da göstermektedir.

<sup>26</sup> Sınıflama ve sıralama ölçeğindeki kategorik değişkenlerin her ikisi de bu testin değişkeni olabilir (Büyüköztürk, 2010).

**Beklenen değerlerin hesaplanması:**

Örneklemdaki gözlenen (tespit edilen) değerlerimiz şöyle olsun:

		Sigara öyküsü		Toplam
		Var	Yok	
Akciğer Kanseri	Pozitif	8	4	12
	Negatif	3	10	13
Toplam		11	14	25

Her bir hücrenin beklenen değeri, o hücrenin satır ve sütun toplamalarının çarpımının, örneklem sayısına bölümüdür. Beklenen değerlerin hesaplanması:

Akciğer kanseri Pozitif ve Sigara öyküsü Var hücresinin beklenen değeri= (Akciğer Kanseri Pozitif Toplam X Sigara Öyküsü Var Toplam)/ Örneklem sayısı= (12 X 11) / 25= 5,28

Akciğer kanseri Pozitif ve Sigara öyküsü Yok hücresinin beklenen değeri= (Akciğer Kanseri Pozitif Toplam X Sigara Öyküsü Yok Toplam)/ Örneklem sayısı= (12 X 14) / 25= 6,72

Akciğer kanseri Negatif ve Sigara öyküsü Var hücresinin beklenen değeri= (Akciğer Kanseri Negatif Toplam X Sigara Öyküsü Var Toplam)/ Örneklem sayısı=(13 X 11) / 25= 5,72

Akciğer kanseri Negatif ve Sigara öyküsü Yok hücresinin beklenen değeri= (Akciğer Kanseri Negatif Toplam X Sigara Öyküsü Yok Toplam)/ Örneklem sayısı= (13 X 14) / 25= 7,28

Bu hesaplamalar sonucunda

Örneklemdaki beklenen değerlerimiz şöyle olacaktır:

		Sigara öyküsü		Toplam
		Var	Yok	
Akciğer Kanseri	Pozitif	5,28	6,72	12
	Negatif	5,72	7,28	13
Toplam		11	14	25

Beklenen değerler; kategorilerin toplam vaka sayıları ve örneklemin toplam sayısı dikkate alınarak her bir hücreye  $H_0$  hipotezinin ilişkisizliği iddiası doğrultusunda “bir nevi” eşit dağılımı sağlayan değerlerin atanması işlemidir. Ki Kare testi gözlenen ve beklenen değerler arasındaki farkın anlamlılığını istatistiksel olarak sınar. Gözlenen değerler ile beklenen değerler arasında anlamlı bir farkın olmaması iki kategorik değişken arasında anlamlı bir ilişkinin olmadığını; gözlenen değerler ile beklenen değerler arasında anlamlı bir farkın olması ise iki kategorik değişken arasında anlamlı bir ilişkinin olduğunu gösterir.

**Test neyi yapar:**

İki kategorik değişken arasında anlamlı bir ilişkinin olup olmadığını sınar. Bu aynı zamanda kategorik bir değişkenin(bağımsız) diğer bir kategorik değişken(bağımlı) üzerinde etkisinin sınanmasını da ifade eder.

**Varsayımları:**

Karşılaştırılan her iki değişken de kategoriktir (Sınıflama ya da sıralama ölçeğinde).

Karşılaştırılan gruplar bağımsızdır.

Beklenen değerleri 5'ten küçük olan kategori sayısının toplam kategori sayısının %20'sini aşmaması ve tüm kategorilerde bu sayının 1'den büyük olması gerekir (Büyüköztürk, 2010; Mangiafico, 2016). Bu koşulun sağlanmadığı durumlarda, mantıklı ise kategori birleştirme yoluna gidilebilir.

**Parametrik Testi: Yok****Hangi durumlarda hangi Ki Kare değeri kullanılır:**

2x2 tablolarda (her bir değişkenin iki kategorisi var ise) beklenen değerleri 5'ten az olan kategori sayısının toplam kategori sayısının %20'sini aşmıyorsa Ki Kare Analizinde Süreklilik için Yates Düzeltmesi (Continuity Correction) değerleri kullanılır. 2x2 bir tabloda beklenen değerleri 5'ten az olan kategori sayısının toplam kategori sayısının %20'sini aşması durumunda Fisher'in Tam Olasılık Testi (Fisher's Exact test) değeri kullanılır (Pallant, 2017; Mangiafico, 2016).

2x2'den büyük tablolarda (değişkenlerden birinin ya da her ikisinin ikiden fazla kategorisi var ise) Pearson Chi-Square değerleri kullanılır (Pallant, 2017). Beklenen değerleri 5'ten küçük olan alt kategorilerin oranının %20'yi geçmemesi gerekir. Bu koşul ihlal edildiği durumda bu dağılım Ki Kare analizine uygun değildir. Bu durumda, mantıklı ise kategori birleştirme yoluna gidilerek yeniden analiz yapılır (Büyüköztürk, 2010). 2x2'den büyük tablolarda fark çıkması durumunda farkın kaynağını tespit etmek için post-hoc test yapılması gerekir.

**Hipotezleri:**

$H_0$ : Değişkenler birbirinden bağımsızdır. (Değişkenler arasında ilişki yoktur.)

$H_1$ : Değişkenler birbirinden bağımsız değildir. (Değişkenler arasında ilişki vardır.)

**Örnek Hipotez:**

$H_0$ : Sigara içme ile Akciğer kanseri olma durumu arasında ilişki yoktur.

$H_1$ : Sigara içme ile Akciğer kanseri olma durumu arasında ilişki vardır.

**Test sonucu:**

$p > 0,05$  ise  $H_0$  hipotezi kabul edilir: Sigara içme ile Akciğer kanseri olma durumu arasında ilişki yoktur

$p < 0,05$  ise  $H_0$  hipotezi ret edilir: Sigara içme ile Akciğer kanseri olma durumu arasında ilişki vardır.

### Etki Büyüklüğü:

Yapılan hipotez testi bize kategorik iki değişken arasında istatistiksel olarak anlamlı bir ilişkinin olup olmadığını söyler, ancak ilişki var ise bu ilişkinin gücü ya da büyüklüğü konusunda bilgi vermez. Etki büyüklüğü bize etkinin önemi konusunda objektif bir ölçüm sağladığı için faydalıdır (Field, 2009). Etki büyüklüğü sonuçların yorumlanmasında dikkate alınması gereken istatistiksel bir değerdir. Hipotez testlerinde etki büyüklüğü değerlerinin de hesaplanarak yorumlanması sonuçların anlaşılabilirliğini arttıracaktır (Büyüköztürk, 2010). Hipotez testinde fark bulunduğu durumda etki büyüklüğüne bakılır. Hipotez testleri için hesaplanan farklı etki büyüklüğü değerleri söz konusudur. Ki Kare Bağımsızlık Testi için kullanılan etki büyüklüğü değerleri:

### Phi ve Cramer'in V Etki Değerleri:

2x2 tablolar için yaygın olarak kullanılan etki büyüklüğü katsayısı Phi değeri<sup>27</sup> ve 2x2'den büyük tablolarda ise yaygın olarak kullanılan etki büyüklüğü değeri ise Cramer's V değeridir (Pallant, 2017; Green ve Salkind 2014; Mangiafico, 2016).

Phi değeri -1 ile +1 arasında değer alır ve Cramer'in V değeri ise 0 ile 1 arasında değer alır. Phi ve Cramer'in V değerlerinin 0'a yakın olması çok zayıf bir ilişkiyi, 1'e yakın olması ise çok güçlü bir ilişkiyi/etkiyi gösterir (Green ve Salkind 2014; Mangiafico, 2016).

Etki değerleri aşağıdaki şekilde yorumlanabilir (Cohen, 1988):

Etki katsayısı	Etki değeri ve yorumu		
	Küçük	Orta	Büyük (Large)
<b>Phi</b>	0.10 – < 0.30	0.30 – < 0.50	≥ 0.50
<b>Cramer's V, k = 2*</b>	0.10 – < 0.30	0.30 – < 0.50	≥ 0.50
<b>Cramer's V, k = 3*</b>	0.07 – < 0.20	0.20 – < 0.35	≥ 0.35
<b>Cramer's V, k = 4*</b>	0.06 – < 0.17	0.17 – < 0.29	≥ 0.29

\* k: satır veya sütunlardaki minimum kategori sayısıdır.<sup>28</sup>

SPSS aracılığı ile Phi ve Cramer's V değerlerinin hesaplanması:

Ki Kare testi yapılırken **Statistics** butonu tıklanır ve **Phi and Cramer's V** seçilir.

Bu seçim bize analiz sonuçlarında Symmetric Measures tablosu içinde Phi ve Cramer's V değerlerini (etki büyüklüğü katsayılarını) verir.

<sup>27</sup> Phi etki değeri yalnızca 2x2 tablolar için kullanılır.

<sup>28</sup> Cramer'in V değerini yorumlarken, karşılaştırılan iki kategorik değişkenden kategori sayısı en küçük olanının kategori sayısını k değeri olarak alıp yorumlayınız.

## 2X2'den büyük tablolar da Ki Kare Bağımsızlık Testinin Post-Hoc Testi:

### Post-Hoc testler hangi anlamlılık düzeyi üzerinden değerlendirilir?

2X2'den büyük tablolarda Ki Kare Bağımsızlık Testinin sonucu bize iki kategorik değişken arasında ilişki olup olmadığını söyler. Farkın kaynağını yani hangi kategorilerin (bağımsız değişken kategorilerinin) birbirinden farklılaştığını söylemez! Farkın kaynağını tespit etmek için post-hoc test yapmak gerekir. Post-Hoc test sonucunda hangi kategorilerin (bağımsız değişken kategorilerinin) birbirinden farklılaştığını ya da farklılaşmadığını tespit ederiz. Post-Hoc testte Tip I hatadan kaçınmak için Alfa ( $\alpha$ ) anlamlılık düzeyi eşleştirilen grup sayısına (test sayısına) bölünür ve bulunan yeni Alfa ( $\alpha$ ) değeri post-hoc testlerdeki yeni anlamlılık düzeyini ifade eder. Post-Hoc testlerde Tip I hatadan kaçınmak için yeni bir anlamlılık düzeyinin/sınırının hesaplanması ve kullanılmasına Bonferroni düzeltmesi denir (Connolly, 2007; Field, 2009; Pallant, 2017). Bonferroni düzeltmesi post-hoc testlerde Tip I hatadan kaçınmak için anlamlılık sınırında yapılan bir düzeltmedir. Post-Hoc testlerde anlamlılık düzeyinde yapılan bu düzeltme Tip I hatadan kaçınmayı sağlarken, Tip II hataya düşme ihtimalini de arttıracaktır.<sup>29</sup> Pos-Hoc testlerde Bonferroni düzeltmesi yaygın olarak kullanılan bir yöntem olsa da bir zorunluluk değildir.

### Ki Kare Bağımsızlık Testinin Post-Hoc Testi:

Ki Kare Bağımsızlık Testinin post-hoc testi Z testidir. Z testi, iki grubun oranlarının bir özelliğe göre önemli ölçüde farklılık gösterip göstermediğini test etmek için kullanılır (Marshall ve Boggis, 2016). Z testi iki bağımsız grubun bağımlı kategorik değişkeninin oranları arasındaki önemli bir farklılık olup olmadığını sınar. Z testi bağımsız değişkenin alt kategorilerini ikili, ikili karşılaştırarak hangi bağımsız grupların birbirinden farklılaştığını tespit eder.

SPSS aracılığı ile Ki Kare testi yapılırken, bağımsız değişkenin sütunda (Column) olmasına dikkat ederek,

**Cells butonu tıklanır**

**Z-test'in altındaki Compare column proportions ve Adjust p-values seçilir.**

Adjust p-values'un seçilmesi Alfa ( $\alpha$ ) anlamlılık düzeyinde Bonferroni düzeltmesi yapılarak analiz yapılmasıdır.

---

<sup>29</sup> Post-Hoc testlerde Alfa ( $\alpha$ ) anlamlılık düzeyi Bonferroni düzeltmesi hesaplamaları için sayfa:56 ve Tip I ve Tip II hata için bakınız sayfa:27

**SPSS Uygulaması:**

- 1- **Analyze, Descriptive Statistic, Crosstabs** seçilir.
- 2- **Crosstabs** penceresinde sol taraftaki değişkenler listesinden **karşılaştırılacak kategorik değişkenin biri Row(s)'a diğeri ise Column(s)'a taşınır.** (Tercihen etki edeceği düşünülen değişken Column'a taşınır)
- 3- **Statistics** butonu **tıklanır** ve **Chi-square** ve **Phi and Cramer's V** seçilir.  
**Continue** tıklanır.
- 4- **Cells** butonu **tıklanır** ve **Counts** kısmında **Observed** seçilidir.  
(2x2'den büyük bir tablo ise fark çıkma durumunda **farkın kaynağını tespit etmek için Cells** kısmında **z-test'in** altındaki **Compare column proportions** ve **Adjust p-values** seçilir)  
**Percentages** kısmında etki edeceği düşünülen değişkenin konumuna göre **Row** ya da **Column** seçilir. (Etki edeceği düşünülen değişken Column'a taşınmışsa burada Column seçilir)  
**Continue** tıklanır.
- 5- **Crosstabs** penceresinde **OK** tıklanır.

**2x2 tablo örneği SPSS çıktısı:****akciger CA yakalanma durumu \* sigara oyküsü Crosstabulation**

			sigara oyküsü		Total
			yok	var	
akciger CA yakalanma durumu	hayır	Count	71	42	113
		% within sigara oyküsü	97,3%	67,7%	83,7%
	evet	Count	2	20	22
		% within sigara oyküsü	2,7%	32,3%	16,3%
Total		Count	73	62	135
		% within sigara oyküsü	100,0%	100,0%	100,0%

**Count:** n % : yüzde

## Chi-Square Tests

	Value	df	Asymptotic Significance (2- sided)	Exact Sig. (2- sided)	Exact Sig. (1- sided)
Pearson Chi-Square	21,416 <sup>a</sup>	1	,000		
Continuity Correction <sup>b</sup>	19,306	1	,000		
Likelihood Ratio	23,724	1	,000		
Fisher's Exact Test				,000	,000
Linear-by-Linear Association	21,257	1	,000		
N of Valid Cases	135				

a. 0 cells (**0,0%**) have expected count less than 5. The minimum expected count is 10,10.

b. Computed only for a 2x2 table

Beklenen değerleri 5'den az olan hücrelerin oranı **%0,0** olduğu için dağılım analize uygun. Bu durumda **Continuity Correction değerleri kullanılır**. Eğer beklenen değerleri 5'ten az olan hücre sayısının oranı %20'yi geçerse 2X2 tablolarda Fisher's Exact Test değerleri kullanılır.

**Continuity Correction:** 2x2 tablolarda ki kare değeri olarak (bu örnekte  $\chi^2 = 19,306$ ) Continuity Correction'ın Value değeri alınır.

**Asymptotic Significance (2-sided) :** 2x2 tablolarda tablolarda Continuity Correction satırının değeri testin p değeridir.

## Symmetric Measures

	Value	Approximate Significance
Nominal by Nominal		
Phi	,398	,000
Cramer's V	,398	,000
N of Valid Cases	135	

2x2 tablolarda etki katsayısı olarak **Phi Value** kullanılır. Phi= ,398 (orta etki)

**Tablo: Sigara öyküsüne göre Akciğer kanserine yakalanma oranlarının karşılaştırılması**

Akciğer Kanseri	Sigara Öyküsü				Toplam	
	Yok		Var			
	n	%	n	%	n	%
<b>Negatif</b>	71	97,3	42	67,7	113	83,7
<b>Pozitif</b>	2	2,7	20	32,3	22	16,3
<b>Toplam</b>	73	100,00	62	100,00	135	100,0
$\chi^2=19,306$ $p=0,000$						

Araştırma grubundaki katılımcılarda sigara öyküsü olmayanların %2,7'sinde akciğer kanseri görülürken bu oran sigara öyküsü olanlarda %32,3'tür. Sigara öyküsü olan ve olmayanlarda akciğer kanseri görülme oranları arasında istatistiksel olarak anlamlı bir farklılık tespit edilmiştir ( $p<0,01$ ). Tespit edilen bu farkın etki büyüklüğü Phi= 0,39 olduğu, ilişkinin orta güçte bir etkiye işaret ettiğini göstermektedir. Sigara öyküsünün akciğer kanserine yakalanma üzerinde orta güçte bir etkiye sahip olduğu (Phi=0,39) tespit edilmiştir.



**3x2 tablo örneği I SPSS çıktısı:****sigara oyküsü \* gelir seviyeleri Crosstabulation**

			gelir seviyeleri			
			düşük	orta	yüksek	Total
sigara oyküsü	yok	Count	8 <sub>a</sub>	40 <sub>b</sub>	25 <sub>c</sub>	73
		% within gelir seviyeleri	25,0%	51,3%	100,0%	54,1%
	var	Count	24 <sub>a</sub>	38 <sub>b</sub>	0 <sub>c</sub>	62
		% within gelir seviyeleri	75,0%	48,7%	0,0%	45,9%
Total		Count	32	78	25	135
		% within gelir seviyeleri	100,0%	100,0%	100,0%	100,0%

Each subscript letter denotes a subset of gelir seviyeleri categories whose column proportions do not differ significantly from each other at the ,05 level.

**Count:** n, **% :** yüzde. Ki kare de fark çıktığında Z testi sonuçlarına göre: Aynı harflerle kodlanmış sütunlar (gruplar) birbirinden farklılaşmamaktadır.

**Chi-Square Tests**

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	32,370 <sup>a</sup>	2	,000
Likelihood Ratio	42,183	2	,000
Linear-by-Linear Association	30,488	1	,000
N of Valid Cases	135		

a. 0 cells (**0,0%**) have expected count less than 5. The minimum expected count is 11,48.

Beklenen değerleri 5'den az olan hücrelerin oranı **%0,0** olduğu için dağılım analize uygun. Ki kare analizi sonuçları geçerlidir.

**Pearson Chi-Square:** 2x2'den büyük tablolarda ki kare değeri olarak (bu örnekte  $\chi^2 = 32,370$ ) Pearson Chi-Square'in Value değeri alınır.

**Asymptotic Significance (2-sided) :** 2x2'den büyük tablolarda tablolarda Pearson Chi-Square satırının değeri testin p değeri.

**Symmetric Measures**

		Value	Approximate Significance
Nominal by Nominal	Phi	,490	,000
	Cramer's V	,490	,000
N of Valid Cases		135	

2x2'den büyük tablolarda etki katsayısı olarak Cramer's V Value kullanılır. Cramer's V = 0,490 (orta etki).

**Tablo: Gelir seviyesine göre Sigara içme oranlarının karşılaştırılması**

Sigara öyküsü	Gelir Seviyesi						Toplam	
	Düşük		Orta		Yüksek			
	n	%	n	%	n	%	n	%
Yok	8	25,0	40	51,3	25	100,0	73	54,1
Var	24	75,0	38	48,7	0	0,0	62	45,9
Toplam	32	100,00	78	100,00	25	100,0	135	100,0
$\chi^2=32,370$ $p=0,000$								

Araştırma grubundaki katılımcılarda gelir seviyesi düşük olanların %25'inin sigara içmediği, gelir seviyesi orta olanların %51,3'ünün sigara içmediği ve yüksek gelir seviyesinde olanların ise %100'ünün sigara içmediği görülmüştür. Gelir seviyesi gruplarının sigara içme oranları arasında istatistiksel olarak anlamlı bir farklılık tespit edilmiştir ( $p<0,01$ ). Tespit edilen bu farkın etki büyüklüğü  $V=0,49$  olduğu, ilişkinin orta güçte bir etkiye işaret ettiğini göstermektedir. Gelir seviyesinin sigara içme durumu üzerinde orta güçte bir etkiye sahip olduğu ( $V=0,49$ ) tespit edilmiştir. Yapılan post-hoc Z testi sonucunda ise tüm grupların birinden farklılaştığı tespit edilmiştir ( $p<0,01$ ).

### 3x2 tablo örneği II SPSS çıktısı:

#### uyuşturucu öyküsü \* gelir seviyeleri Crosstabulation

			gelir seviyeleri			
			düşük	orta	yüksek	Total
uyuşturucu öyküsü	yok	Count	24	76	25	125
		% within gelir seviyeleri	75,0%	97,4%	100,0%	92,6%
	var	Count	8	2	0	10
		% within gelir seviyeleri	25,0%	2,6%	0,0%	7,4%
Total		Count	32	78	25	135
		% within gelir seviyeleri	100,0%	100,0%	100,0%	100,0%

Count: n, % : yüzde

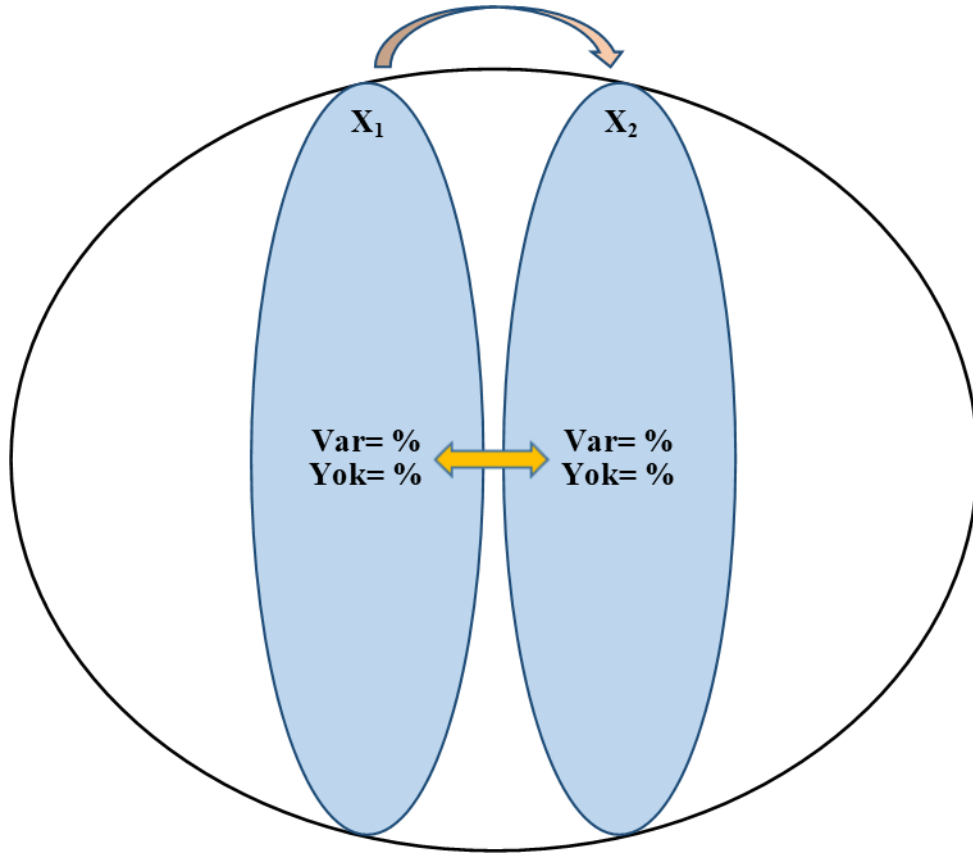
#### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	19,108 <sup>a</sup>	2	,000
Likelihood Ratio	16,702	2	,000
Linear-by-Linear Association	14,302	1	,000
N of Valid Cases	135		

a. 2 cells (33,3%) have expected count less than 5. The minimum expected count is 1,85.

Beklenen değerleri 5'den az olan hücrelerin oranı **%33** olduğu için **dağılım analize uygun değil**. Ki kare analizi sonuçları geçerli değildir. Kategori birleştirme seçeneği ile analiz yeniden yapılabilir.

### 3.2. McNemar Testi <sup>30</sup>



McNemar testi, bağımsız (ilişkisiz) gruplar yerine bağımlı (yani ilişkili veya eşleştirilmiş) veriler (iki kategorili/dichotomous variable) için kullanan bir tür ki-kare testi olarak görülebilir (Adedokun ve Burgess, 2012). McNemar testi, tekrarlayan veya eşleştirilmiş ölçümler arasındaki farkları değerlendirir. Tekrarlayan veya eşleştirilen değişkenler kategorik bir değişkendir ve yalnızca iki kategorisi vardır (Evet/Hayır; Var/Yok vb). McNemar testi kategorik verilerin tekrarlayan ya da eşleştirilmiş I. ölçüm oranı ile II. ölçüm oranlarını karşılaştır (Green ve Salkind, 2014).  $H_0$  (yokluk hipotezi/ null hypothesis) hipotezi I. ölçüm oranı ile II. ölçüm oranlarının aynı olduğunu iddia eder (Adedokun ve Burgess, 2012). McNemar Ki Kare testi iki dikotom (iki kategorili) cevabın ölçümlerinin uyumsuzluğunu karşılaştırır (Trajman ve Luis, 2008).

<sup>30</sup> McNemar testi literatürde McNemar Ki Kare testi (Rufibach, 2011; Trajman ve Luis, 2008; Sun ve Yang, 2008) ya da Bağımlı Gruplarda Ki Kare testi (Adedokun ve Burgess, 2012; Terzi, 2019) olarak da adlandırılmaktadır.

**McNemar Testinin işlem süreci örneği:**

		Son Test		
		Doğru	Yanlış	Toplam
Ön Test	Doğru	a= 55	b= 35	n <sub>1</sub> = 90
	Yanlış	c= 199	d= 217	416
Toplam		n <sub>2</sub> = 254	252	n= 506

Bir öğrenci grubunun ön test ve son testteki doğru ve yanlış sayıları yukarıdaki gibidir.

a = ön test ve son testte doğru cevap veren öğrencilerin sayısı

b = ön testte doğru fakat son testte yanlış veren öğrencilerin sayısı

c= ön teste yanlış cevap verip son teste doğru cevap veren öğrencilerin sayısı

d = ön test ve son testte yanlış cevap veren öğrencilerin sayısı

n = toplam örneklem sayısı

n<sub>1</sub> = ön teste doğru cevap veren öğrenci sayısı (yani a + b)

n<sub>2</sub> = son teste doğru cevap veren öğrenci sayısı (yani a + c)

p<sub>1</sub> = ön testte doğru yanıtların oranı, yani n<sub>1</sub> / n veya (a + b) / n

p<sub>2</sub> = son testteki doğru yanıtların oranı, yani n<sub>2</sub> / n veya (a +c) / n

Ön test ve son testte doğru cevap veren öğrencilerin oranlarındaki değişiklikleri sınamak istediğimizi varsayalım, bu durumda H<sub>0</sub>: p<sub>1</sub> = p<sub>2</sub> hipotezini test edeceğiz. McNemar Testi H<sub>0</sub> hipotezini sınarken, iki oranın eşitliğini (equivalence) test etmek için değişikliğin meydana geldiği iki uyumsuz (discordant) hücreden b & c'deki (yukarıdaki tablodaki) verileri kullanır (Adedokun ve Burgess, 2012).

McNemar testinin güvenilir sonuçlar üretebilmesi için örneklem büyüklüğünün yeterli sayıda olması ve ayrıca uyumsuz (discordant) hücrelerdeki vaka sayılarının küçük olmaması gerekir, yazarlar uyumsuz hücrelerin toplamında en az 5, 10 veya 25 sayısının olmasını tavsiye etmektedir (Mangiafico, 2016). McNemar testinin geçerliliği (validity) için yaygın olarak kabul edilen kural, uyumsuz (discordant) hücrelerdeki sayıların toplamının 10'dan büyük olmasıdır (Rufibach, 2011).

**McNemar testinin düzeltilmemiş değeri aşağıdaki formülle hesaplanır:**

Bu formül büyük örneklem için kullanılır.

$$\chi^2 = \frac{(b - c)^2}{(b + c)} = \frac{(35 - 199)^2}{(35 + 199)} = 114,931$$

$$\chi^2 = 114,931 \text{ Serbestlik Derecesi 1 için } p=0,000$$

**McNemar testinin düzeltilmiş değeri aşağıdaki formülle hesaplanır:**

Bu formül küçük örneklem de kullanılmalıdır. SPSS McNemar testinde bu formülü kullanarak bir p değeri üretir (Exact Sig. 2-sided değeri p değeridir).

$$\chi^2 = \frac{(|b - c| - 1)^2}{(b + c)} = \frac{(|35 - 199| - 1)^2}{(35 + 199)} = 113,543$$

$$\chi^2 = 113,543 \text{ Serbestlik Derecesi 1 için } p=0,000$$

**Test neyi yapar:**

Bir grubun tekrarlayan ya da eşleştirilen kategorik ölçümlerinde katılımcıların cevaplarını bir yönde değiştirenlerin sayısı ile karşı yönde değiştirenlerin sayısı arasında anlamlı bir fark olup olmadığını sınar. Ön test son test dizaynı içinde eşleştirme ya da tekrarlayan ölçüm söz konusudur. Söz konusu kategorik değişkenin yalnızca iki alt kategorisi olmalıdır (0=yok, 1 =var ya da 1=hayır, 2=evet vb. Kategori kodlamalarının 0-1 şeklinde yapılması zorunluluğu yoktur, 1-2 şeklinde de kategoriler tanımlanabilir).

**Varsayımları:**

Karşılaştırılan gruplar bağımlıdır (aynı grup).

Karşılaştırılan değişken kategoriktir (iki kategorisi olan).

Tekrarlayan ya da eşleştirilmiş ölçümler birbirinden bağımsızdır (iki ayrı ölçüm).

**Parametrik testi: Yok****Hipotezleri:**

H<sub>0</sub>: Tekrarlayan/eşleştirilen ölçüm oranları arasında fark yoktur.

H<sub>1</sub>: Tekrarlayan/eşleştirilen ölçüm oranları arasında fark vardır.

**Test sonucu:**

$p > 0,05$  ise H<sub>0</sub> hipotezi kabul edilir: Tekrarlayan/eşleştirilen ölçüm oranları arasında fark yoktur.

$p < 0,05$  ise H<sub>0</sub> hipotezi ret edilir: Tekrarlayan/eşleştirilen ölçüm oranları arasında fark vardır.

**Bu testin uygulamasına örnek:** Bir örneklemin tekrarlayan kategorik ölçümlerinde; referandum öncesi ve sonrası katılımcıların siyasal iktidara verdikleri desteğin (var/yok) yönündeki değişimi sınamak istediğimizde McNemar testi yaparız.

**SPSS Uygulaması:****Yöntem 1**

- 1- **Analyze, Descriptive Statistic, Crosstabs** seçilir.
- 2- **Crosstabs** penceresinde sol taraftaki değişkenler listesinden **karşılaştırılacak kategorik değişkenin biri Row(s)'a diğeri ise Column(s)'a taşınır.**
- 3- **Cells** butonu **tıklanır**  
**Percentages** kısmında **Row** seçilir  
**Continue** **tıklanır.**
- 4- **Statistic** butonu **tıklanır**  
**McNemar** seçilir  
**Continue** **tıklanır.**
- 5- **Crosstabs** penceresinde  
**OK** **tıklanır.**

**Yöntem 2**

- 1- **Analyze, Nonparametric Tests, Legacy Dialogs, 2 Related Sampes** seçilir.
- 2- **Two Related Sampes** penceresinde  
Sol taraftaki değişken listesinden **kategorik değişkenler** eşleştirilerek **Test Pairs'e** taşınır.  
**Test Type** kısmından **McNemar** seçilidir.  
**OK** Tıklanır.

SPSS çıktısı:

**Yöntem 1:****Crosstabs****referandumdan önce destek \* referandumdan sonra destek Crosstabulation**

			referandumdan sonra destek		
			yok	var	Total
referandumdan önce destek	yok	Count	60	0	60
		% within referandumdan önce destek	100,0%	0,0%	100,0%
	var	Count	11	64	75
		% within referandumdan önce destek	14,7%	85,3%	100,0%
Total		Count	71	64	135
		% within referandumdan önce destek	52,6%	47,4%	100,0%

**Chi-Square Tests**

	Value	Exact Sig. (2-sided)
McNemar Test		,001 <sup>a</sup>
N of Valid Cases	135	

a. Binomial distribution used.

**Exact Sig. (2-sided)** : testin p değeri

**Tablo: Siyasal İktidara Referandum Öncesi ve Sonrası Destek Durumundaki Değişim**

			Referandumdan Sonra		Toplam
			Yok	Var	
Referandumdan Önce	Yok	n	60	0	60
		%	100,0	0,0	100,0
	Var	n	11	64	75
		%	14,7	85,3	100,0
Toplam		n	71	64	135
		%	52,6	47,4	100,0
p= 0,001					

Araştırma grubunda referandum öncesinde siyasal iktidara destek vermeyen 60 katılımcının referandum sonrası görüşlerinde hiçbir değişiklik (%0) olmamıştır, referandum öncesi siyasal iktidara destek veren 75 katılımcıdan 11'inin (%15'inin) ise referandum sonrası görüşlerinde bir değişiklik olmuş ve siyasal iktidara verdikleri destekten vazgeçmişlerdir. Araştırma grubundaki katılımcıların, referandum öncesi ve sonrası siyasal iktidara destek verme durumlarının değişimi istatistiksel olarak anlamlı bulunmuştur ( $p<0,01$ ).

## Yöntem 2 SPSS çıktısı:

### McNemar Test

referandumdan önce destek & referandumdan sonra destek

referandumdan önce destek	referandumdan sonra destek	
	yok	var
Yok	60	0
Var	11	64

### Test Statistics<sup>a</sup>

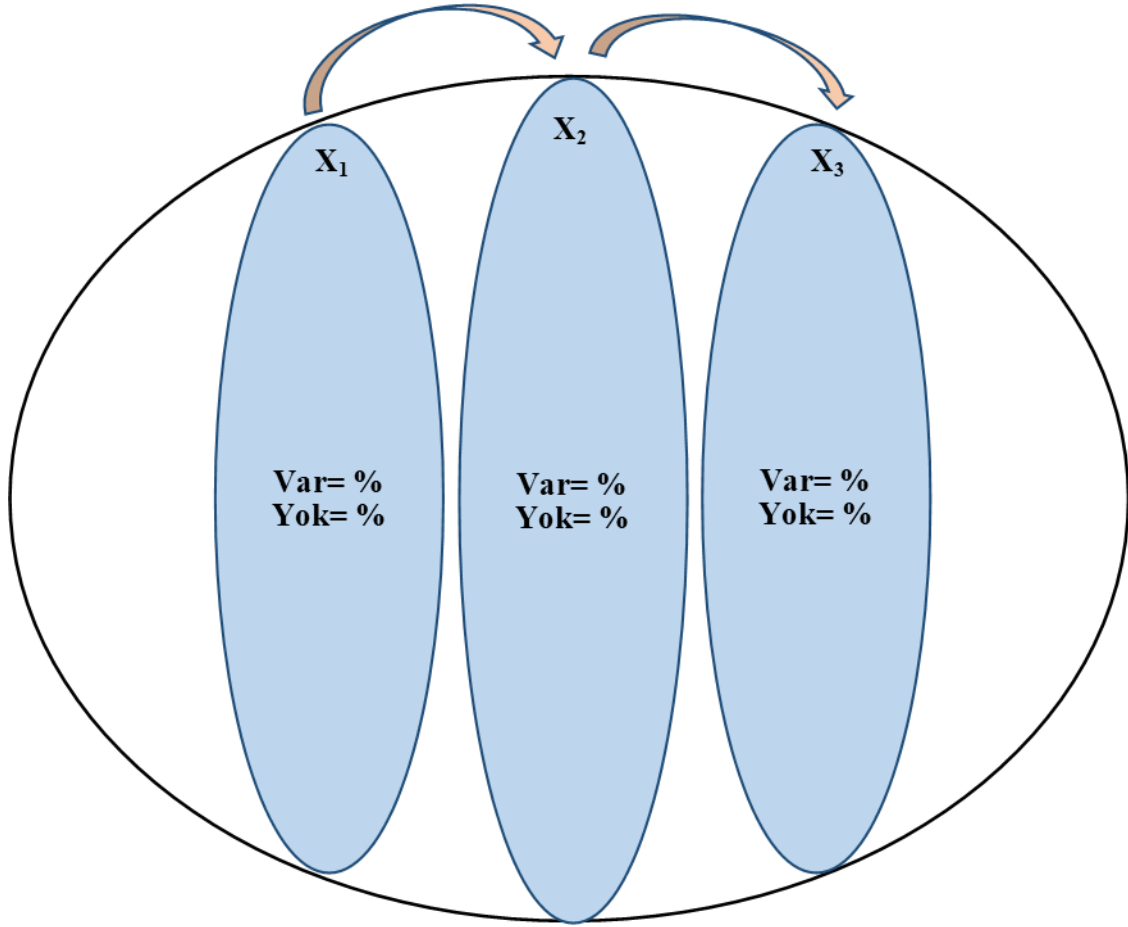
referandumdan önce destek & referandumdan sonra destek	
N	135
Exact Sig. (2-tailed)	,001 <sup>b</sup>

a. McNemar Test

b. Binomial distribution used.

**Exact Sig. (2-sided) :** testin p değeri

### 3.3. Cochran Q Testi <sup>31</sup>



Cochran testi, tekrarlayan ölçüm veya eşleştirilen konu tasarımları ile ilgili problemlerde uygulanabilir. Cochran testi bir grubun tekrarlayan ya da eşleştirilmiş üç ya da daha fazla ölçüm verisinin (dikotom/iki kategorili) karşılaştırılmasını sağlar (Green and Salkind, 2014; McCrum-Gardner, 2008). Bu test McNemar'ın testinin bir uzantısıdır (extension) ve temelde dikotom (iki kategorili) veriler için yapılan bir Friedman testidir (Field, 2009). Tekrarlayan ya da eşleştirilen gözlemler bağımlıdır (bir gruba aittir), ancak tekrarlayan/eşleştirilen gözlemler birbirinden bağımsızdır (yani birbirinden ayrı ölçümleri ifade eder). Cochran testi için bağımlı değişkenin iki kategorisi vardır ve  $H_0$  hipotezi (yokluk hipotezi) popülasyon oranların zaman ya da eşleştirilen faktör seviyeleri için eşit olduğunu belirtir (Green and Salkind, 2014). Başka bir ifade ile bu testin  $H_0$  hipotezi (yokluk hipotezi) birden çok ilişkili ölçüm oranlarının aynı olduğudur (Sing ve Diğr., 2013).

Cochran testi tekrarlayan ölçümler ya da eşleştirilen konular arasında katılımcıların verdikleri cevaplarda önemli ölçüde bir farklılık gösterip göstermediğini değerlendirir. Test, tasarımın tekrarlanan ölçümleri veya eşleşen konu özellikleri için verilen cevapların bağımlılığını dikkate alır. Örneklem büyüklüğü 30'dan büyükse test sonuçları oldukça doğru sonuçlar üretir (Green

<sup>31</sup> Bu test Cochran Ki Kare Testi olarak da adlandırılmaktadır (Field, 2009).



and Salkind, 2014).<sup>32</sup> Test sonucu tekrarlayan ya da eşleştirilen dikotom (iki kategorili) verilerin dağılımı arasında fark olup olmadığını söyler, fark çıkması durumunda farkın kaynağını tespit etmek için post-hoc test yapmak gerekir.

### **Test neyi yapar:**

Aynı grubun üç ya da daha fazla tekrarlayan/eşleştirilen dikotom (iki kategorili) değişkeninin ölçümleri arasında anlamlı bir fark olup olmadığını sınamak için uygulanır. Aynı grubun tekrarlayan 3 ya da fazla ölçümü söz konusudur. Söz konusu kategorik değişkenin yalnızca iki alt kategorisi olmalıdır (0=yok, 1 =var ya da 1=hayır, 2=evet vb. Kategori kodlamalarının 0-1 şeklinde yapılması zorunluluğu yoktur, 1-2 şeklinde de kategoriler tanımlanabilir).

### **Varsayımları:**

Karşılaştırılan gruplar bağımlıdır (aynı grup).

Karşılaştırılan değişken kategoriktir (iki kategorisi olan).

Tekrarlayan ölçüm sayısı 3 ya da daha fazladır.

Tekrarlayan ya da eşleştirilmiş ölçümler birbirinden bağımsızdır (ayrı ölçümler).

### **Parametrik testi: Yok**

### **Hipotezleri:**

$H_0$ : Tekrarlayan/eşleştirilen ölçüm oranları arasında fark yoktur.

$H_1$ : Tekrarlayan/eşleştirilen ölçüm oranları arasında fark vardır.

### **Test sonucu:**

$p > 0,05$  ise  $H_0$  hipotezi kabul edilir: Tekrarlayan/eşleştirilen ölçüm oranları arasında fark yoktur.

$p < 0,05$  ise  $H_0$  hipotezi ret edilir: Tekrarlayan/eşleştirilen ölçüm oranları arasında fark vardır.

### **Etki Büyüklüğü:**

Yapılan hipotez testi bize tekrarlayan ya da eşleştirilen ölçümlerin oranları arasında istatistiksel olarak anlamlı bir fark olup olmadığını söyler, ancak fark var ise bu farkın etkisi ya da büyüklüğü konusunda bilgi vermez. Etki büyüklüğü bize etkinin önemi konusunda objektif bir ölçüm sağladığı için faydalıdır (Field, 2009). Etki büyüklüğü sonuçların yorumlanmasında dikkate alınması gereken istatistiksel bir değerdir. Hipotez testlerinde etki büyüklüğü değerlerinin de hesaplanarak yorumlanması sonuçların anlaşılabilirliğini arttıracaktır (Büyüköztürk, 2010). Hipotez testinde fark bulunduğu durumda etki büyüklüğüne bakılır. Hipotez testleri için hesaplanan farklı etki büyüklüğü değerleri söz konusudur. Cochran Testi için kullanılan etki büyüklüğü değeri:

<sup>32</sup> Örneklem büyüklüğü ile ilgili daha liberal yaklaşımlarda mevcuttur. Kanji (2006) bu test için örneklem büyüklüğünün 10 ve üzeri olmasının yeterli olacağını iddia etmektedir.

**Kendall'in W Katsayısı:** Cochran testinin etki büyüklüğü istatistiği olarak Kendall'in W (Kendal'in uyum katsayısı/ Kendall's coefficient of concordance) değeri kullanılır (Green ve Salkind, 2014). Kendal'in W değeri 0 ile 1 arasında değer almaktadır, yüksek değerler daha güçlü ilişkiyi göstermektedir (Green ve Salkind, 2014). Kendal'in W katsayısı şu şekilde yorumlanabilir: 0,00 – 0,20 = çok zayıf etki, 0,21 – 0,40 = zayıf etki, 0,41 – 0,60 = orta (moderate) etki, 0,61 – 0,80 = güçlü etki, 0,81 – 1,00 = çok güçlü etki (Rovai ve Diğr., 2014).

SPSS aracılığı ile Kendall'in W Katsayısının hesaplanması:

Cochran testini yaparken **Test for Several Related Samples** penceresinde **Kendall's W** seçilir.

### Cochran Testinin Post-Hoc Testi: McNemar Testi

#### Post-Hoc testler hangi anlamlılık düzeyi üzerinden değerlendirilir?

Cochran Testinin sonucu bize tekrarlayan ya da eşleştirilen ölçümler arasında farkın olup olmadığını söyler. Farkın kaynağını yani hangi ölçümlerin birbirinden farklılaştığını söylemez! Farkın kaynağını tespit etmek için post-hoc test yapmak gerekir. Post-Hoc test sonucunda hangi ölçümlerin birbirinden farklılaştığını ya da farklılaşmadığını tespit ederiz. Post-Hoc testte Tip I hatadan kaçınmak için Alfa ( $\alpha$ ) anlamlılık düzeyi eşleştirilen grup sayısına (test sayısına) bölünür ve bulunan yeni Alfa ( $\alpha$ ) değeri post-hoc testlerdeki yeni anlamlılık düzeyini ifade eder. Post-Hoc testlerde Tip I hatadan kaçınmak için yeni bir anlamlılık düzeyinin/sınırının hesaplanması ve kullanılmasına Bonferroni düzeltmesi denir (Connolly, 2007; Field, 2009; Pallant, 2017). Bonferroni düzeltmesi post-hoc testlerde Tip I hatadan kaçınmak için anlamlılık sınırında yapılan bir düzeltmedir. Post-Hoc testlerde anlamlılık düzeyinde yapılan bu düzeltme Tip I hatadan kaçınmayı sağlarken, Tip II hataya düşme ihtimalini de arttıracaktır.<sup>33</sup> Pos-Hoc testlerde Bonferroni düzeltmesi yaygın olarak kullanılan bir yöntem olsa da bir zorunluluk değildir.

**Bu testin uygulamasına örnek:** Bir örneklemin tekrarlayan kategorik ölçümlerinde; Ocak, Haziran ve Aralık aylarında siyasal iktidara verdikleri desteğin (var/yok) yönündeki oranları arasındaki farkı/değişimi sınamak istediğimizde Cochran Testi testi yaparız.

#### SPSS Uygulaması:

##### Yöntem 1

- 1- **Analyze, Nonparametric Tests, Legacy Dialogs, K Related Sampes** seçilir.
- 2- **Test for Several Related Sampes** penceresinde sol taraftaki değişken listesinden **kategorik değişkenler Test Variables'e** taşınır.  
**Test Type** kısmından **Cochran's Q Testi** seçilidir.  
**Kendall's W** seçilir (Etki büyüklüğü değeri için).  
**OK** Tıklanır.

**Fark çıktığında post-hoc test olarak McNemar testi yapılır.**

<sup>33</sup> Post-Hoc testlerde Alfa ( $\alpha$ ) anlamlılık düzeyi Bonferroni düzeltmesi hesaplamaları için sayfa:56 ve Tip I ve Tip II hata için bakınız sayfa:27

## Yöntem 2

- 1- **Analyze, Nonparametric Tests, Related Samples** seçilir.
- 2- **Objective** sekmesinde **Customize Analysis** seçilir.
- 3- **Fields** sekmesi **tıklanır**.  
Sol taraftaki değişken listesinden **analiz edilecek kategorik değişkenler Test Fields'e taşınır**.
- 4- **Settings** sekmesinde **tıklanır**.  
**Customize tests** seçilir.  
**Cochran's Q (k samples)** seçilir.  
**Multiple comparisons** kısmında **All pairwise** seçilidir.
- 5- **Run** Tıklanır.

Yöntem 1 SPSS çıktısı:

### NPar Tests

### Cochran Test

#### Frequencies

	Value	
	0	1
iktidara destek ocak	60	75
iktidara destek haziran	71	64
iktidara destek aralık	75	60

**Value:** Değer, yani kategori kodları (bu örnekte 0=destek yok, 1=destek var).  
Tablonun her bir hücrelerinde yer alan sayılar ise o kesişime düşen örneklem sayısıdır.

#### Test Statistics

N	135
Cochran's Q	24,133 <sup>a</sup>
df	2
Asymp. Sig.	,000

a. 1 is treated as a success.

**N:** Örneklem sayısı

**Cochran's Q:** test değeri

**df:** serbestlik derecesi

**Asymp. Sig. :** testin p değeri

## Kendall's W Test

### Ranks

	Mean Rank
iktidara destek ocak	2,10
iktidara destek haziran	1,97
iktidara destek aralık	1,93

### Test Statistics

N	135
Kendall's W <sup>a</sup>	,089
Chi-Square	24,133
df	2
Asymp. Sig.	,000

a. Kendall's Coefficient of Concordance

**Kendall's W:** Etki büyüklüğü değeri

**Tablo: Ocak, Haziran ve Aralık Aylarında Siyasal İktidara Destek**

Ölçüm zamanı	İktidara Destek				Toplam	
	Yok		Var			
	n	%	n	%	n	%
Ocak	60	44,4	75	55,6	135	100,0
Haziran	71	52,6	64	47,4	135	100,0
Aralık	75	55,6	60	44,4	135	100,0
$\chi^2=24,133$ $p=0,000$						

Araştırma grubunda katılımcıların siyasal iktidara verdikleri destek oranlarının Ocak ayında %55,6, Haziran ayında %47,4 ve Aralık ayında ise %44,4 olduğu tespit edilmiştir. Katılımcıların Ocak, Haziran ve Aralık aylarında ki siyasal iktidarı destekleme oranları arasında istatistiksel olarak anlamlı bir farklılık tespit edilmiştir ( $p<0,01$ ). Tespit edilen bu farkın etki büyüklüğü  $W=0,09$  olduğu, çok zayıf bir etkiye işaret ettiğini göstermektedir. Zaman içinde siyasal iktidara olan destekte anlamlı bir düşüş görülse de bu düşüşün çok zayıf bir etkiye sahip olduğu ( $W=0,09$ ) tespit edilmiştir.

Yapılan post-hoc McNemar Testi sonucunda ise Ocak ve Haziran aylarındaki oranlar ile Ocak ve Aralık aylarındaki siyasal iktidara destek oranlarının birbirlerinden istatistiksel olarak anlamlı bir şekilde farklılaştıkları ( $p<0,01$ ); Haziran ve Aralık aylarındaki siyasal iktidara destek oranlarının ise istatistiksel olarak anlamlı bir şekilde farklılaşmadığı ( $p>0,017$ ) tespit edilmiştir.

## Yöntem 2 SPSS çıktısı:

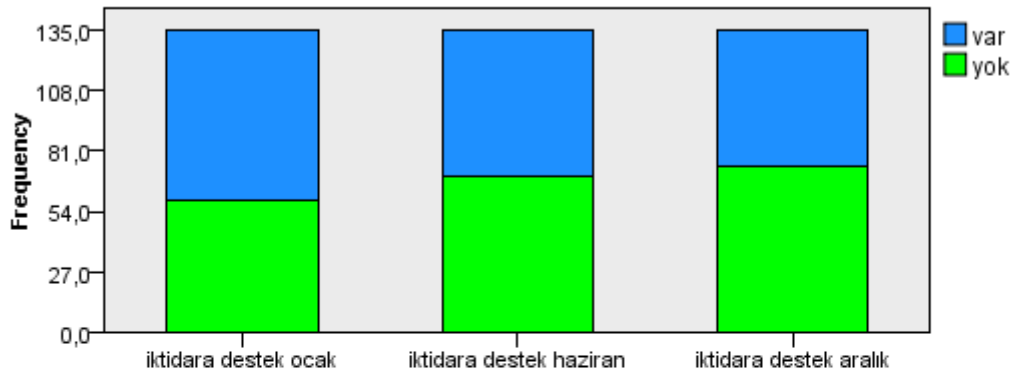
## Nonparametric Tests

## Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distributions of iktidara destek ocak, iktidara destek haziran and iktidara destek aralık are the same.	Related-Samples Cochran's Q Test	,000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

## Related-Samples Cochran's Q Test



<b>Total N</b>	135
<b>Test Statistic</b>	24,133
<b>Degrees of Freedom</b>	2
<b>Asymptotic Sig. (2-sided test)</b>	,000

**Total N:** Örneklem sayısı

**Test Statistic:** Cochran's Q test değeri

**Degrees of Freedom:** Serbestlik derecesi

**Asymp. Sig. (2 sided test):** testin p değeri

**Post-Hoc test sonuçları:****Pairwise Comparisons**

Each node shows the sample number of successes.

Sample1-Sample2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj.Sig.
iktidara destek ocak-iktidara destek haziran	-,081	,023	-3,479	,001	,002
iktidara destek ocak-iktidara destek aralık	-,111	,023	-4,743	,000	,000
iktidara destek haziran-iktidara destek aralık	-,030	,023	-1,265	,206	,618

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is ,05.

Significance values have been adjusted by the Bonferroni correction for multiple tests.

## 4.BÖLÜM

# NİCEL DEĞİŞKENLER ARASI İLİŞKİ

### 4.1. Basit Korelasyon <sup>34</sup>

#### 4.1.1. Pearson Korelasyon Testi

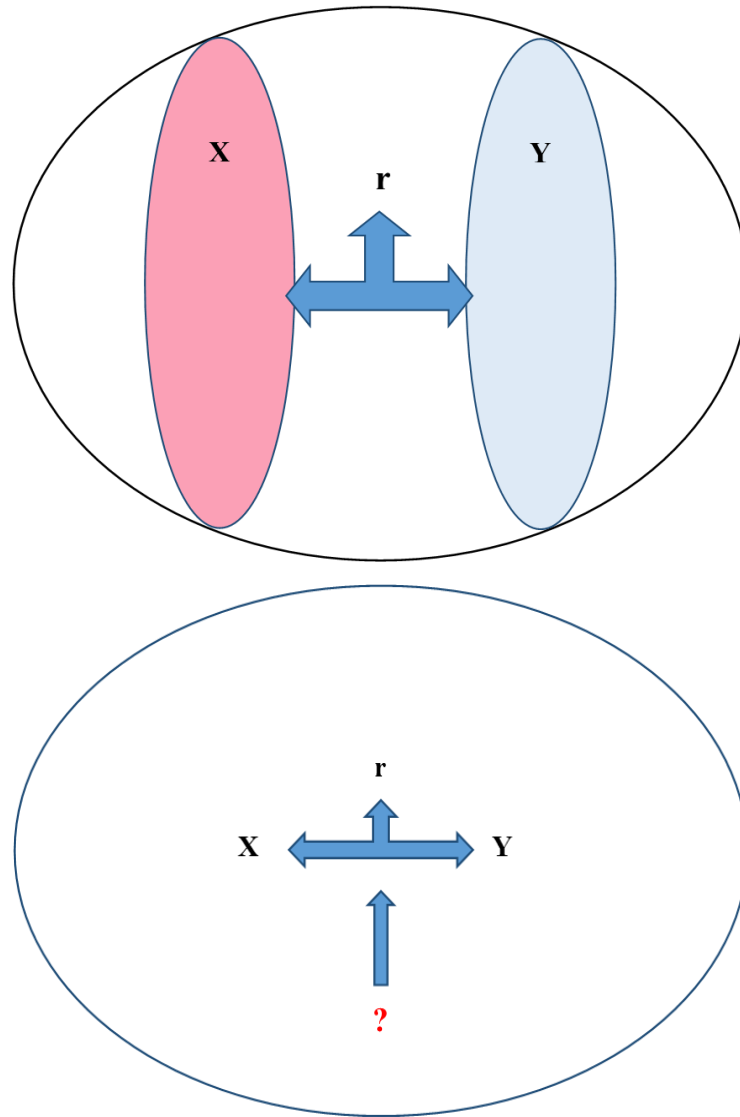
Basit Korelasyon analizi iki nicel değişken arasındaki ilişkiyi sınar ve yapılabilmemesinin ön koşulu iki nicel verinin aynı kaynaktan elde edilmiş ve eşleştirilmiş olmasıdır (Can, 2019). Korelasyon analizi, iki nicel değişken arasındaki doğrusal ilişkinin<sup>35</sup> gücü ve yönü hakkında sayısal bilgi verir. Pearson korelasyon katsayısı ( $r$ ) -1 ile +1 arasında bir değer alır. Katsayının önündeki işaret ilişkinin yönünü: (+) pozitif (bir değişken artarken diğerinin de arttığını ya da bir değişkenin değeri azalırken diğerinin de azaldığı), (-) negatif (bir değişken artarken diğerinin azaldığı ya da biri azalırken diğerinin arttığı) gösterir. Korelasyon katsayısının mutlak (işaretinden bağımsız olarak) değerinin büyüklüğü ise değişkenler arasındaki ilişkinin gücü hakkında bilgi verir. Korelasyon katsayısının değerinin 0 olması iki değişken arasında doğrusal bir ilişkinin olmadığını, +1 ya da -1 olması ise iki değişken arasında mükemmel bir doğrusal bir ilişkinin olduğunu gösterir (Pallant, 2017; Field, 2009).

Basit korelasyon iki değişken arasındaki ilişkinin bir göstergesidir ve istatistiksel olarak bir değişkenin diğerinin nedeni olduğu yönünde bir analizi içermez (Pallant, 2017). Korelasyon sonuçlarının yorumlanmasında yaygın olarak yapılan bir yanlış, ilişkinin bir neden-sonuç ilişkisi içinde yorumlanmasıdır. Korelasyon analizinde elde edilen sonuç, bir karşılıklılık ilişkisinin sonucudur (Can, 2019). Korelasyon analizinin neden-sonuç şeklinde yorumlanmasının yanlış olmasının iki nedeni vardır: 1. Korelasyon iki nicel değişken arasındaki karşılıklı bir ilişkinin ölçümüdür, nedenselliğin yönünü vermez, 2. İki nicel değişken arasında neden-sonuç ilişkisini etkileyebilecek diğer değişkenlerin etkisi (üçüncü değişken sorunu) göz ardı edilmiş olunacaktır (Pallant, 2017, Field, 2009).<sup>36</sup>

<sup>34</sup> Simple/Bivariate Correlation

<sup>35</sup> Doğrusal ilişki, bir değişkende sabit bir değişimin her zaman diğer değişkende sabit bir değişimle bağlantılı olduğu ilişki türüdür (Rosenthal ve Rosnow, 2008).

<sup>36</sup> Nedensellik yönü: Korelasyon analizi, hangi değişkenin diğerinin değişmesine neden olduğu hakkında hiç bir şey söylemez. Üçüncü değişken sorunu göz ardı edildiğinde ve iki ilişkili değişkenin tek önemli değişken olduğu varsayıldığında da korelasyon katsayısı nedenselliğin hangi yönde çalıştığını göstermez (Field, 2009).



İki nicel değişken arasındaki karşılıklı ilişkinin, bir neden-sonuç ilişkisi olarak yorumlanabilmesi; alan bilgisinde tartışılmayacak değerdeki doğrulara dayanması ve ilişkisi sınanan iki değişken üzerinde etkisi olabilecek diğer değişkenlerin bu ilişkiye olan etkilerinin kontrol altına alınması ile mümkün olabilir. Örneğin IQ puanı ile sınav başarı puanı arasında bir ilişkinin korelasyon analizinde, neden sonuç ilişkisi içinde sonuçların yorumlanmasında bir sakınca olmayacağı düşünülebilir. Yine de bu durumda neden-sonuç ilişkisi olarak yorumlanmasının ikinci sakıncası doğacaktır; çünkü iki değişken arasındaki ilişkiyi etkileyebilecek diğer değişkenler –üçüncü değişken sorunu- (çalışma süresi, özgüven, depresyon vb.) göz ardı edilmiş olunacaktır. Basit korelasyon iki nicel değişken arasındaki karşılıklı ilişkinin yönü ve gücü konusunda bir bilgi üretir, bunu yaparken bu iki değişkeni etkileyebilecek (dolayısı ile iki değişken arasındaki ilişkiyi etkileyebilecek) diğer değişkenleri önemsemez ve göz ardı eder. İki değişken arasındaki ilişkiyi sınarken bu iki değişkeni, dolayısı ile aralarındaki ilişki katsayısına etki edebilecek diğer değişken ya da değişkenlerin de araştırma sürecinde ölçülmesi önemlidir. İki değişken arasındaki ilişkiyi sınarken bunların üzerinde etkisi olabilecek diğer değişkenlerin etkisinin kontrol altına alınması ile Kısmi Korelasyon



analizi yapılır, bunun sonucunda ulaşılan korelasyon katsayısı iki değişken arasındaki ilişkinin daha objektif bir ölçümünü sunması açısından değerlidir.

**Test neyi yapar:**

Bir grubun iki nicel değişkeni arasında istatistiksel olarak anlamlı –doğrusal- bir ilişkinin olup olmadığını sınırlar, ilişkinin varlığı durumunda ilişkinin yönünü ve gücünü bize verir. Basit korelasyon X ve Y değişkenleri arasındaki doğrusal ilişkiyi analiz ederken, bu değişkenler arasındaki ilişkiye etki edebilecek diğer değişkenlerin etkisini göz ardı eder. Bir anlamda X ve Y değişkenlerini içinde bulunduğu çevresel koşullardan soyutlayarak X ve Y arasındaki doğrusal ilişkiye yönelik bir bilgi verir.

**Varsayımları:**

Örneklem Evrenden rastgele seçilmiştir.

Değişkenler en az aralık ölçeğindedir ve normal dağılım özelliği gösterir.

Karşılaştırılan gruplar bağımlıdır (ölçümler aynı gruba aittir).

Değişkenler birbirinden bağımsızdır (iki ayrı ölçüm ve ölçümler birinden etkilenmez).

**Non-Parametrik testi: Spearman Korelasyon Testi****Hipotezleri:**

$H_0$ : X ve Y nicel değişkenleri arasında ilişki yoktur.

$H_1$ : X ve Y nicel değişkenleri arasında ilişki vardır.

**Test sonucu:**

$p > 0,05$  ise  $H_0$  hipotezi kabul edilir: X ve Y nicel değişkenleri arasında ilişki yoktur.

$p < 0,05$  ise  $H_0$  hipotezi ret edilir: X ve Y nicel değişkenleri arasında ilişki vardır.

**Korelasyon testi sonuçlarının değerlendirilmesi:**

İlişkinin olup olmadığına **p değerine** göre karar verilir.  $p < 0,05$  ise iki değişken arasında ilişki vardır,  $p > 0,05$  ise iki değişken arasında ilişki yoktur.

**İlişki olması durumunda ise r değerine bakılır: r, -1 ile +1 aralığında değer alır.**

**İlişkinin yönü:**

**r işareti + ise Pozitif yönde ilişki:** değişkenlerden biri artarken diğerinin de artması ya da değişkenlerden birinin değeri azalırken diğerinin de değerinin düştüğü anlamına gelir.

**r işareti – ise Negatif yönde ilişki:** değişkenlerden biri artarken diğerinin değerinin düştüğü ya da değişkenlerden birinin değeri düşerken diğerinin değerinin arttığı anlamına gelir.

### İlişkinin gücü:

İlişkinin gücünü  $r$  katsayısının (işaretinden bağımsız olarak) sayısal değeri gösterir. Korelasyon katsayısının ( $r$ ) büyüklüğüne göre yorumlanması evrensel değildir, disiplinden disipline farklılık gösterir (Mangiafico, 2016). Korelasyon katsayıları işareten bağımsız olarak: 0,10-0,29 küçük, 0,30-0,49 orta, 0,50-1,00 ise büyük (large) etki olarak yorumlanabilir<sup>37</sup> (Cohen, 1988 ve 1992; Field, 2009; Green ve Salkind 2014). İlişkinin gücünü gösteren  $r$  katsayısı aynı zamanda etki büyüklüğü değeridir.

**Determinasyon katsayısı:** Açıklanan varyans, değişkenlerden birinde gözlenen değişkenliğinin ne kadarının diğer değişken tarafından açıklandığını ifade eder ve determinasyon katsayısı olarak isimlendirilen bu katsayı korelasyon katsayısının karesine eşittir ( $r^2$ ) (Büyüköztürk, 2010). Determinasyon katsayısı ( $r^2$ ), bir değişkenin diğeri tarafından paylaşılan değişkenlik miktarının bir ölçüsüdür (Field, 2009). Determinasyon katsayısı bir değişkendeki değişimin yüzde kaçının diğer değişken tarafından açıklanabildiğini gösterir. Korelasyon teknik olarak bir neden-sonuç ilişkisi analizi olmadığı için açıklanan varyans karşılıklı olarak her iki değişkenin birbirindeki değişimin ne kadarını açıkladığı şekilde değerlendirilir (Can, 2019). Bu katsayı iki değişkenin karşılıklı olarak birbirlerinin ne kadarını açıkladığını ifade eder. **Determinasyon katsayısı=  $r^2$**

**Bu testin uygulamasına örnek:** Bir grubun aylık gelirleri ile eğitim aldıkları yıl arasında bir ilişki olup olmadığı sınanmak istendiğinde, her iki değişkeninde verileri normal dağılım özelliği gösterdiğinde ilişki Pearson Korelasyon testi ile sınanır.

**Önemli not:** Bu uygulamada verilerin normal dağılım özelliği gösterdiği varsayılmıştır. Siz kendi analizlerinizde öncelikli olarak normal dağılıma uygunluğu sınamalısınız.

### SPSS Uygulaması:

- 1- **Analyze, Correlate, Bivariate seçilir.**
- 2- **Bivariate Correlations** penceresinde  
Sol taraftaki değişken listesinden **karşılaştırılacak nicel değişkenler Variables'e taşınır.**  
**Correlations Coefficients** kısmından  
**Pearson seçilir.**  
**OK Tıklanır.**

<sup>37</sup> Korelasyon katsayısının yorumlanmasında yaygın olarak Cohen'in (1988 ve 1992) sınıflandırılması kullanılsa da Gignac ve Szodorai (2016) korelasyon analizi yapılmış 708 çalışmanın meta analizi sonucunda,  $r$  katsayısının 0,10-0,19 küçük, 0,20-0,29 orta, 0,30 ve üstü değerlerin ise büyük etki olarak yorumlanmasını önermiştir. Cohen ve Holliday (1982) ise  $r$  katsayısının: 0,19 ve altı çok düşük (very low), 0,20-0,39 düşük (low), 0,40-0,69 orta (modest), 0,70-0,89 yüksek (high), 0,90-1 çok yüksek (very high) etki olarak yorumlanmasını önermişlerdir.

## SPSS çıktısı:

## Correlations

Correlations			
		aylık gelir	eğitim aldığı yıl
aylık gelir	Pearson Correlation	1	,857**
	Sig. (2-tailed)		,000
	N	135	135
eğitim aldığı yıl	Pearson Correlation	,857**	1
	Sig. (2-tailed)	,000	
	N	135	135

\*\* . Correlation is significant at the 0.01 level (2-tailed).

**Pearson Correlation:** r katsayısı

**Sig. (2-tailed):** testin p değeri

**N:** Örneklem sayısı

$$r^2 = 0,857 \times 0,857 = 0,73$$

$$r^2 = \%73$$

**Tablo: Aylık Gelir ve Eğitim Yılı Arasındaki İlişki**

	Eğitim Yılı	
Aylık Gelir	r	0,857
	p	0,000
	n	135

Araştırma grubundaki katılımcıların aylık gelirleri ile eğitim aldıkları yıl arasında pozitif yönde güçlü ve istatistiksel olarak anlamlı bir ilişki tespit edilmiştir ( $r=0,857$ ,  $p<0,01$ ). Eğitim alınan yılın, aylık gelir üzerinde pozitif yönde ve büyük bir etkisinin olduğu görülmüştür. Aylık gelirdeki toplam varyansın (değişimin) %73'ü eğitim yılı ile açıklanmaktadır. (Bu tablonun açıklama metninin ilk cümlesi teknik olarak doğrudur, ardından gelen cümleler neden-sonuç kurgusu içinde yazılmıştır dolayısı ile neden-sonuç ilişkisinin handikaplarını taşır!)

#### 4.1.2. Spearman Korelasyon Testi

Spearman Korelasyon Testi, Pearson Korelasyon Testinin Non-Parametrik testidir. Spearman Korelasyon Testi, Pearson Korelasyon testi gibi Basit Korelasyon analizi yapar. Basit Korelasyon Analizi için Pearson Korelasyon Testinde verilen bilgiler bu test içinde geçerlidir. Spearman Korelasyon Testi, ilişkisi sınanacak iki değişkenin sıralama ölçeğinde olması ya da nicel (en az aralık ölçeğindeki) değişken verilerinin normal dağılım özelliği göstermemesi durumunda başvurulmuş bir yöntemdir.

Spearman Korelasyon Testi en az aralık ölçeğindeki X ve Y değişkenleri verilerinin normal dağılım özelliği göstermemesi durumunda, X ve Y veri gruplarını kendi içinde değerlerini küçükten büyüğe doğru sıralar ve 1'den başlayarak sıra sayısı vererek sıra sayısına dönüştürür. Bu en az aralık ölçeğindeki verilerin basitçe sıralama ölçeğine dönüştürülmesi işlemidir. Ardından X ve Y değişkeninin sıra sayısı değerlerinin her bir katılımcı için sıra fark değerlerini hesaplar ve bu fark değerleri üzerinden X ve Y değişkenleri arasında ki doğrusal ilişkiyi sınırlar (Corder ve Foreman, 2009).

#### Spearman Korelasyon Testinin sıra değerlerini üretme ve işlem süreci örneği:

Veri seti			Spearman Korelasyon Testinin işlem süreci			
Gerçek Değerler			Sıra Değerleri		Sıra Farkları	
Anket No	X Değişkeni	Y Değişkeni	X Değişkeni	Y Değişkeni	(Y-X)	(Y-X) <sup>2</sup>
1	5	100	2	7	5	25
2	12	89	5	6	1	1
3	7	78	3	5	2	4
4	14	66	6	2	-4	16
5	2	77	1	4	3	9
6	8	105	4	9	5	25
7	15	67	7	3	-4	16
8	25	63	9	1	-8	64
9	17	115	8	10	2	4
10	30	103	10	8	-2	4
			$\sum (Y-X)^2 = 168$			

#### Test neyi yapar:

Spearman korelasyon katsayısı değişkenlerin sıra değerleri arasındaki doğrusal bir ilişkinin ölçümüdür (IBM SPSS, 2017). İki değişken arasındaki doğrusal ilişki incelenirken, değişkenlerin sıralama ölçeğinde olması ya da en az aralık ölçeğindeki nicel verilerin biri veya her ikisi de normal dağılım özelliği göstermiyorsa Spearman korelasyon testi yapılır. Aynı grubun iki nicel değişkeni arasında istatistiksel olarak anlamlı –doğrusal- bir ilişkinin olup olmadığını sınırlar, ilişkinin varlığı durumunda ilişkinin yönünü ve gücünü bize verir. Basit korelasyon X ve Y değişkenleri arasındaki doğrusal ilişkiyi bize verirken, bu değişkenler arasındaki ilişkiye etki edecek diğer değişkenlerin etkisini göz ardı eder. Bir anlamda X ve Y değişkenlerini içinde bulunduğu çevresel koşullardan soyutlayarak X ve Y arasındaki ilişkiye yönelik bir bilgi verir.

**Varsayımları:**

Değişkenler en az sıralama ölçeğindedir.

Karşılaştırılan gruplar bağımlıdır (ölçümler aynı gruba aittir).

Değişkenler birbirinden bağımsızdır (iki ayrı ölçüm ve ölçümler birinden etkilenmez).

**Parametrik testi: Pearson Korelasyon Testi****Hipotezleri:**

$H_0$ : X ve Y nicel değişkenleri arasında ilişki yoktur.

$H_1$ : X ve Y nicel değişkenleri arasında ilişki vardır.

**Test sonucu:**

$p > 0,05$  ise  $H_0$  hipotezi kabul edilir: X ve Y nicel değişkenleri arasında ilişki yoktur.

$p < 0,05$  ise  $H_0$  hipotezi ret edilir: X ve Y nicel değişkenleri arasında ilişki vardır.

**Korelasyon testi sonuçlarının değerlendirilmesi:**

İlişkinin olup olmadığına **p değerine** göre karar verilir.  $p < 0,05$  ise iki değişken arasında ilişki vardır,  $p > 0,05$  ise iki değişken arasında ilişki yoktur.

**İlişki olması durumunda ise r değerine bakılır: r, -1 ile +1 aralığında değer alır.**

**İlişkinin yönü:**

**r işareti + ise Pozitif yönde ilişki:** değişkenlerden biri artarken diğerinin de artması ya da değişkenlerden birinin değeri azalırken diğerinin de değerinin düştüğü anlamına gelir.

**r işareti – ise Negatif yönde ilişki:** değişkenlerden biri artarken diğerinin değerinin düştüğü ya da değişkenlerden birinin değeri düşerken diğerinin değerinin arttığı anlamına gelir.

**İlişkinin gücü:**

İlişkinin gücünü r katsayısının (işaretinden bağımsız olarak) sayısal değeri gösterir. Korelasyon katsayısının (r) büyüklüğüne göre yorumlanması evrensel değildir, disiplinden disipline farklılık gösterir (Mangiafico, 2016). Korelasyon katsayıları işareten bağımsız olarak: 0,10-0,29 küçük, 0,30-0,49 orta, 0,50-1,00 ise büyük (large) etki olarak yorumlanabilir<sup>38</sup> (Cohen, 1992; Field, 2009; Green ve Salkind 2014). İlişkinin gücünü gösteren r katsayısı aynı zamanda etki büyüklüğü değeridir.

<sup>38</sup> Korelasyon katsayısının yorumlanmasında yaygın olarak Cohen'in (1988 ve 1992) sınıflandırılması kullanılsa da Gignac ve Szodorai (2016) korelasyon analizi yapılmış 708 çalışmanın meta analizi sonucunda, r katsayısının 0,10-0,19 küçük, 0,20-0,29 orta, 0,30 ve üstü değerlerin ise büyük etki olarak yorumlanmasını önermiştir. Cohen ve Holliday (1982) ise r katsayısının: 0,19 ve altı çok düşük (very low), 0,20-0,39 düşük (low), 0,40-0,69 orta (modest), 0,70-0,89 yüksek (high), 0,90-1 çok yüksek (very high) etki olarak yorumlanmasını önermişlerdir.

**Determinasyon katsayısı:** Açıklanan varyans, değişkenlerden birinde gözlenen değişkenliğinin ne kadarının diğer değişken tarafından açıklandığını ifade eder ve determinasyon katsayısı olarak isimlendirilen bu katsayı korelasyon katsayısının karesine eşittir ( $r^2$ ) (Büyüköztürk, 2010). Determinasyon katsayısı ( $r^2$ ), bir değişkenin diğeri tarafından paylaşılan değişkenlik miktarının bir ölçüsüdür (Field, 2009). Determinasyon katsayısı bir değişkendeki değişimin yüzde kaçının diğer değişken tarafından açıklanabildiğini gösterir. Korelasyon teknik olarak bir neden-sonuç ilişkisi analizi olmadığı için açıklanan varyans karşılıklı olarak her iki değişkenin birbirinde ki değişimin ne kadarını açıkladığı şekilde değerlendirilir (Can, 2019). Bu katsayı iki değişkenin karşılıklı olarak birbirlerinin ne kadarını açıkladığını ifade eder. **Determinasyon katsayısı=  $r^2$**

**Bu testin uygulamasına örnek:** Bir grubun aylık gelirleri ile eğitim aldıkları yıl arasında bir ilişki olup olmadığı sınanmak istendiğinde, değişkenlerin verileri normal dağılım özelliği göstermediğinde ilişki Spearman Korelasyon testi ile sınanır.

#### SPSS Uygulaması:

- 1- **Analyze, Correlate, Bivariate** seçilir.
- 2- **Bivariate Correlations** penceresinde  
Sol taraftaki değişken listesinden **karşılaştırılacak nicel değişkenler Variables'e** taşınır.  
**Correlations Coefficients** kısmından  
**Spearman** seçilir.  
**OK** Tıklanır.

#### SPSS çıktısı:

#### Nonparametric Correlations

Correlations			aylık gelir	eğitim aldığı yıl
Spearman's rho	aylık gelir	Correlation Coefficient	1,000	,916**
		Sig. (2-tailed)	.	,000
		N	135	135
	eğitim aldığı yıl	Correlation Coefficient	,916**	1,000
		Sig. (2-tailed)	,000	.
		N	135	135

\*\*. Correlation is significant at the 0.01 level (2-tailed).

**Correlation Coefficient:** r katsayısı

**Sig. (2-tailed):** testin p değeri

**N:** Örneklem sayısı

**$r^2=0,916 \times 0,916=0,84$**

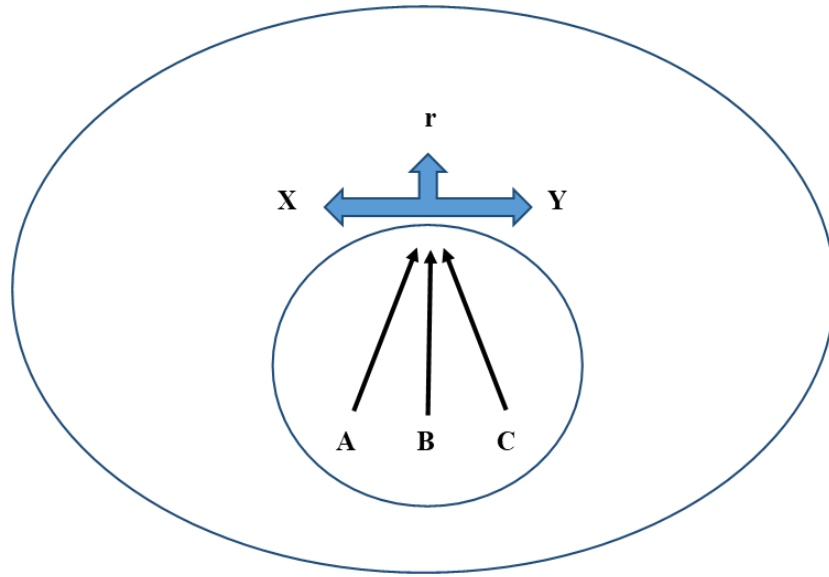
**$r^2= \%84$**

**Tablo: Aylık Gelir ve Eğitim Yılı Arasındaki İlişki**

Aylık Gelir	Eğitim Yılı	
	<b>r</b>	0,916
	<b>p</b>	0,000
	<b>n</b>	135

Araştırma grubundaki katılımcıların aylık gelirleri ile eğitim aldıkları yıl arasında pozitif yönde güçlü ve istatistiksel olarak anlamlı bir ilişki tespit edilmiştir ( $r=0,916$ ,  $p<0,01$ ). Eğitim alınan yılın, aylık gelir üzerinde pozitif yönde ve büyük bir etkisinin olduğu görülmüştür. Aylık gelirdeki toplam varyansın (değişimin) %84'ü eğitim yılı ile açıklanmaktadır. (Bu tablonun açıklama metninin ilk cümlesi teknik olarak doğrudur, ardından gelen cümleler neden-sonuç kurgusu içinde yazılmıştır dolayısı ile neden-sonuç ilişkisinin handikaplarını taşır!)

#### 4.2. Kısmi Korelasyon <sup>39</sup>



İki tür korelasyon vardır: basit ve kısmi. Basit korelasyon, iki değişken arasındaki ilişkidir, kısmi korelasyon ise iki değişken arasındaki ilişkiyi sınavken bir veya daha fazla ilave değişkenin etkisini “kontrol eder”. Kısmi korelasyonda iki nicel değişken arasındaki ilişki analiz edilirken, bu iki değişkene dolayısı ile bunların arasındaki ilişkiyi etkileyebilecek diğer değişkenlerin etkisini (üçüncü değişken sorunu) kontrol ederek (arındırarak) daha objektif bir korelasyon katsayısı üretir (Field, 2009). Basit Korelasyona göre Kısmi Korelasyon iki değişken arasındaki ilişkinin sınavmasında daha objektif bir analiz yöntemidir.

İki değişken arasındaki ilişkiyi sınavken, bu iki değişkene dolayısı ile aralarındaki ilişki katsayısına etki edebilecek diğer değişken ya da değişkenlerin de araştırma sürecinde ölçülmesi önemlidir. İki değişken arasındaki ilişkiyi sınavken bunların üzerinde etkisi olabilecek diğer değişkenlerin etkisinin kontrol altına alınması ile Kısmi Korelasyon analizi yapılır, bunun sonucunda ulaşılan korelasyon katsayısı iki değişken arasındaki ilişkinin daha objektif bir ölçümünü sunması açısından değerlidir. Kısmi Korelasyon da Basit Korelasyon gibi iki nicel değişken arasındaki karşılıklı bir ilişki analizidir, teknik olarak nedenselliği yönünde bir sonuç üretmez. Bununla birlikte Kısmi Korelasyonda bu ilişki üzerinde etkisi olabilecek üçüncü değişken sorunu çözüldüğü için alan bilgisi dahilinde X ve Y değişkenleri arasındaki ilişki bir neden-sonuç ilişkisi şeklinde yorumlanabilir.

Kısmi Korelasyon analizi, iki nicel değişken arasındaki ilişkiyi sınavken üçüncü değişkenlerin bu ilişkiye etkisini “kontrol altına” aldıktan sonra, iki değişken arasındaki doğrusal ilişkinin<sup>40</sup> gücü ve yönü hakkında sayısal bilgi verir. Korelasyon katsayısı ( $r$ ) -1 ile +1 arasında bir değer alır. Katsayının önündeki işaret ilişkinin yönünü: (+) pozitif (bir değişken artarken diğerinin de arttığını ya da bir değişkenin değeri azalırken diğerinin de azaldığı), (-) negatif (bir değişken artarken diğerinin azaldığı ya da biri azalırken diğerinin arttığı) gösterir. Korelasyon katsayısının mutlak (işaretinden bağımsız olarak) değerinin büyüklüğü ise değişkenler arasındaki ilişkinin gücü hakkında bilgi verir. Korelasyon katsayısının değerinin 0 olması iki değişken arasında

<sup>39</sup> Partial Correlation

<sup>40</sup> Doğrusal ilişki, bir değişkende sabit bir değişimin her zaman diğer değişkende sabit bir değişimle bağlantılı olduğu ilişki türüdür (Rosenthal ve Rosnow, 2008).



doğrusal bir ilişkinin olmadığını, +1 ya da -1 olması ise iki değişken arasında mükemmel bir doğrusal bir ilişkinin olduğunu gösterir (Pallant, 2017; Field, 2009).

**Test neyi yapar:**

Kısmi korelasyon iki değişken arasındaki doğrusal ilişkiyi tanımlarken, bu ilişkiye bir ya da daha fazla değişkenin etkilerini kontrol altına aldıktan sonra bir korelasyon katsayısı hesaplar (IBM SPSS, 2017). İki nicel değişken arasındaki ilişki sınanırken, bu iki değişken arasındaki ilişki üzerinde etkisi olabileceğini düşündüğümüz (şüphelendiğimiz) başka nicel değişkenin/değişkenlerin etkisini kontrol altına alarak (arındırarak) yapılan ilişki analizidir. X ve Y arasındaki ilişkiyi sınarken X ve/veya Y üzerinde etkisi olabilecek A değişkeninin etkisini kontrol altına aldıktan sonra X ve Y arasında bir ilişkinin olup olmadığını söyler. A'nın etkisi arındırıldıktan sonra X ile Y arasında bir ilişki var mı?

**Varsayımları:**

Örnekleme Evrenden rastgele seçilmiştir.

Değişkenler en az aralık ölçeğindedir ve normal dağılım özelliği gösterir.

Karşılaştırılan gruplar bağımlıdır (ölçümler aynı gruba aittir).

Değişkenler birbirinden bağımsızdır (her biri ayrı bir ölçümü ifade eder ve birbirlerinden etkilenmez).

**Non-Parametrik testi:** Yok

**Hipotezleri:**

$H_0$ : X ve Y nicel değişkenleri arasında ilişki yoktur.

$H_1$ : X ve Y nicel değişkenleri arasında ilişki vardır.

**Test sonucu:**

$p > 0,05$  ise  $H_0$  hipotezi kabul edilir: X ve Y nicel değişkenleri arasında ilişki yoktur.

$p < 0,05$  ise  $H_0$  hipotezi ret edilir: X ve Y nicel değişkenleri arasında ilişki vardır.

**Korelasyon testi sonuçlarının değerlendirilmesi:**

İlişkinin olup olmadığına **p değerine** göre karar verilir.  $p < 0,05$  ise iki değişken arasında ilişki vardır,  $p > 0,05$  ise iki değişken arasında ilişki yoktur.

**İlişki olması durumunda ise r değerine bakılır: r, -1 ile +1 aralığında değer alır.**

**İlişkinin yönü:**

**r işareti + ise Pozitif yönde ilişki:** değişkenlerden biri artarken diğerinin de artması ya da değişkenlerden birinin değeri azalırken diğerinin de değerinin düştüğü anlamına gelir.

**r işareti – ise Negatif yönde ilişki:** değişkenlerden biri artarken diğerinin değerinin düştüğü ya da değişkenlerden birinin değeri düşerken diğerinin değerinin arttığı anlamına gelir.

### İlişkinin gücü:

İlişkinin gücünü  $r$  katsayısının (işaretinden bağımsız olarak) sayısal değeri gösterir. Korelasyon katsayısının ( $r$ ) büyüklüğüne göre yorumlanması evrensel değildir, disiplinden disipline farklılık gösterir (Mangiafico, 2016). Korelasyon katsayıları işareten bağımsız olarak: 0,10-0,29 küçük, 0,30-0,49 orta, 0,50-1,00 ise büyük (large) etki olarak yorumlanabilir<sup>41</sup> (Cohen, 1992; Field, 2009; Green ve Salkind 2014). İlişkinin gücünü gösteren  $r$  katsayısı aynı zamanda etki büyüklüğü değeridir.

**Determinasyon katsayısı:** Açıklanan varyans, değişkenlerden birinde gözlenen değişkenliğinin ne kadarının diğer değişken tarafından açıklandığını ifade eder ve determinasyon katsayısı olarak isimlendirilen bu katsayı korelasyon katsayısının karesine eşittir ( $r^2$ ) (Büyüköztürk, 2010). Determinasyon katsayısı ( $r^2$ ), bir değişkenin diğeri tarafından paylaşılan değişkenlik miktarının bir ölçüsüdür (Field, 2009). Determinasyon katsayısı bir değişkendeki değişimin yüzde kaçının diğer değişken tarafından açıklanabildiğini gösterir. Korelasyon teknik olarak bir neden-sonuç ilişkisi analizi olmadığı için açıklanan varyans karşılıklı olarak her iki değişkenin birbirindeki değişimin ne kadarını açıkladığı şekilde değerlendirilir (Can, 2019). Bu katsayı iki değişkenin karşılıklı olarak birbirlerinin ne kadarını açıkladığını ifade eder. Basit korelasyondan farklı olarak Kısmi korelasyonda üçüncü değişkenlerin etkisi kontrol altına alındığı için, alan bilgisi dahilinde neden-sonuç ilişkisi kurulabileceğinden determinasyon katsayısı bir değişkenin diğerinin ne kadarını açıkladığı yönünde yorumlanabilir. **Determinasyon katsayısı=  $r^2$**

**Bu testin uygulamasına örnek:** Bir grubun IQ puanları ile Sınav Başarı Puanları arasındaki ilişki Sınav Kaygısı'nın etkisi kontrol altına alınarak sınanmak istendiğinde, söz konusu tüm değişkenlerin verileri normal dağılım özelliği gösterdiğinde ilişki Kısmi Korelasyon testi ile sınanır.

**Önemli not:** Bu uygulamada ilişkisi sınanan değişkenlerin verilerinin normal dağılım özelliği gösterdiği varsayılmıştır. Siz kendi analizlerinizde öncelikli olarak verilerin normal dağılıma uygunluğunu sınamalısınız.

### SPSS Uygulaması:

- 1- **Analyze, Correlate, Partial** seçilir.
- 2- **Partial Correlations** penceresinde sol taraftaki değişken listesinden **karşılaştırılacak nicel değişkenler Variables'e taşınır. Kontrol altında tutmak istediğimiz** (etkisini arındırmak istediğimiz) **nicel değişkenler ise Controlling for'a taşınır.**
- 3- **Option** butonu tıklanır. **Statistics** bölümünden **Zero order correlations** seçilir. **Continue** tıklanır. **OK** tıklanır.

<sup>41</sup> Korelasyon katsayısının yorumlanmasında yaygın olarak Cohen'in (1988 ve 1992) sınıflandırılması kullanılsa da Gignac ve Szodorai (2016) korelasyon analizi yapılmış 708 çalışmanın meta analizi sonucunda,  $r$  katsayısının 0,10-0,19 küçük, 0,20-0,29 orta, 0,30 ve üstü değerlerin ise büyük etki olarak yorumlanmasını önermiştir. Cohen ve Holliday (1982) ise  $r$  katsayısının: 0,19 ve altı çok düşük (very low), 0,20-0,39 düşük (low), 0,40-0,69 orta (modest), 0,70-0,89 yüksek (high), 0,90-1 çok yüksek (very high) etki olarak yorumlanmasını önermişlerdir.

## SPSS çıktısı:

## Partial Corr

Correlations				sınav başarı puanı	sınav kaygısı puanı
Control Variables			IQ puanı		
-none- <sup>a</sup>	IQ puanı	Correlation	1,000	,677	-,526
		Significance (2-tailed)	.	,000	,000
		df	0	133	133
	sınav başarı puanı	Correlation	,677	1,000	-,792
		Significance (2-tailed)	,000	.	,000
		df	133	0	133
	sınav kaygısı puanı	Correlation	-,526	-,792	1,000
		Significance (2-tailed)	,000	,000	.
		df	133	133	0
sınav kaygısı puanı	IQ puanı	Correlation	1,000	,502	
		Significance (2-tailed)	.	,000	
		df	0	132	
	sınav başarı puanı	Correlation	,502	1,000	
		Significance (2-tailed)	,000	.	
		df	132	0	

a. Cells contain zero-order (Pearson) correlations.

**Correlation:** r katsayısı

**Significance (2-tailed):** testin p değeri

**df:** serbestlik derecesi

$$r^2 = 0,502 \times 0,502 = 0,25$$

$$r^2 = \%25$$

**Tablo: Sınav Puanı, IQ Puanı ve Kaygı Puanı Arasındaki İlişki**

			<b>Sınav Puanı</b>	<b>Kaygı Puanı</b>
<b>Basit korelasyon</b>	<b>IQ Puanı</b>	<b>r</b>	,677	-,526
		<b>p</b>	,000	,000
		<b>sd</b>	133	133
	<b>Sınav Puanı</b>	<b>r</b>		-,792
		<b>p</b>		,000
		<b>sd</b>		133
<b>Kısmi korelasyon</b> <b>Kontrol değişkeni=Kaygı puanı</b>	<b>IQ Puanı</b>	<b>r</b>	,502	
		<b>p</b>	,000	
		<b>sd</b>	132	

Araştırma grubundaki katılımcıların Sınav Puanı ile IQ Puanı arasında pozitif yönde güçlü ( $r=0,677$ ,  $p<0,01$ ); Sınav Puanı ile Kaygı Puanı arasında negatif yönde güçlü ( $r=-0,792$ ,  $p<0,01$ ); IQ Puanı ile Kaygı Puanı arasında ise negatif yönde güçlü ( $r=-0,526$ ,  $p<0,01$ ) ve istatistiksel olarak anlamlı ilişkilerin olduğu tespit edilmiştir. IQ puanı Sınav Puanındaki toplam varyansın (değişimin) %46'sını açıklamaktadır (son cümle neden-sonuç ilişkisinin handikaplarını taşır!).

Araştırma grubunda Kaygı Puanının, IQ Puanı ve Sınav Puanı üzerinde yol açtığı değişim kontrol altına alındığında, IQ Puanı ile Sınav Puanı arasında pozitif yönde güçlü ( $r=0,502$ ,  $p<0,01$ ) ve istatistiksel olarak anlamlı bir ilişkinin olduğu tespit edilmiştir. IQ Puanı Sınav Puanındaki toplam varyansın (değişimin) %25'ini açıklamaktadır (son cümle neden-sonuç ilişkisinin handikaplarını halen taşır! Çünkü Sınav puanı üzerinde etkisi olabilecek tüm değişkenler kısmi korelasyonda kontrol değişkeni olarak dahil edilmemiştir.).

## 5.BÖLÜM

# DOĞRUSAL REGRESYON ANALİZİ <sup>42</sup>

Regresyon, sosyal bilimlerde verilerin analizinde yaygın olarak kullanılan tekniklerden biri haline gelmiştir. Birkaç noktada belirginleşeceği gibi Pearson'un Korelasyon Testi ile yakından bağlantılıdır. Değişkenler arasındaki ilişkilerin doğrusal olması ve değişkenlerin en az eşit aralık ölçeğinde olması gibi Pearson'un Korelasyon Testi ile aynı varsayımların çoğunu paylaşır. Regresyon değişkenler<sup>43</sup> arasındaki ilişkinin doğasını özetlemek ve bağımlı değişkenin olası değerlerini tahmin etmek için güçlü bir araçtır (DeCoster ve Claypool 2004; Bryman ve Cramer, 1999). Regresyon analizi, bir bağımsız (yordayıcı) değişkenden (basit regresyon) veya birkaç bağımsız değişkenden (çoklu regresyon) bir bağımlı (yordanan) değişkeni tahmin etmenin bir yoludur. Bu araç inanılmaz derecede kullanışlıdır, çünkü topladığımız verilerin ötesine geçmemize izin verir (Field, 2009). Doğrusal regresyon, aralarında doğrusal ilişki olan, en az aralık ölçeğindeki bağımlı ve bağımsız değişkenler arasındaki ilişkinin matematiksel bir eşitlik denklemi ile açıklanmasını ve bağımlı değişkenin alacağı değerlere göre bağımlı değişkenin tahminini sağlayan bir tekniktir.

### 5.1. Basit Doğrusal Regresyon Analizi <sup>44</sup>

Basit Regresyon, bir değişkenin değerini diğerinden tahmin etmenin bir yoludur (Field, 2009). Basit Doğrusal Regresyon, aralarında doğrusal ilişki olan iki nicel değişkenden, birinin bağımlı diğerinin bağımsız değişken olarak kurgulandığı durumda, aralarındaki ilişkinin bir matematiksel eşitlik ile açıklanması sürecidir. Bağımsız değişkenin bağımlı değişkenin ne kadarını açıkladığını ve aralarında nasıl bir matematiksel eşitlik olduğunu ifade eder (Büyüköztürk, 2010). Bir bağımsız değişken ile bir bağımlı değişken arasındaki doğrusal ilişkinin sınıandığı Regresyon Analizine, Basit Doğrusal Regresyon Analizi denir.

#### Basit Doğrusal Regresyon Analizi Neyi Yapar:

- 1- Bağımsız (yordayıcı) değişkenin, bağımlı (yordanan) değişkeni anlamlı bir şekilde açıklayıp açıklamadığını, ne şekilde etkilediğini (pozitif/negatif) saptar.
- 2- Bağımsız (yordayıcı) değişkenin bağımlı (yordanan) değişkenin ne kadarının açıkladığını, determinasyon katsayısı ( $R^2$ ) ile belirler.

<sup>42</sup> Linear Regression

<sup>43</sup> Field (2009) Regresyon analizinde bağımsız (yordayıcı) değişken yerine tahmin değişkeni (predictor variable) ve bağımlı değişken yerine de çıktı değişkeni (outcome variable) kavramlarını kullanmaktadır. Stevens (2009) bağımsız (yordayıcı) değişken yerine tahmin değişkeni (predictor variable) ve bağımlı değişken (aynı şekilde) kullanmaktadır. Green ve Salkind (2014), Tabachnick ve Fidell (2007), Coakes, (2005), DeCoster ve Claypool (2004), Kerr ve Diğr., (2002), Bryman ve Cramer (1999) ise bağımsız ve bağımlı değişken kavramlarını tercih etmektedir.

<sup>44</sup> Simple Linear Regression

- 3- Bağımlı (yordanan) değişken ile bağımsız (yordayıcı) değişken arasındaki ilişkiyi regresyon eşitliği ile açıklar. Böylelikle bağımsız (yordayıcı) değişkenin alacağı değere göre bağımlı (yordanan) değişkenin alacağı değerin tahmin edilmesini sağlar.

Basit Doğrusal Regresyon Analizi, bağımlı ve bağımsız değişken arasında anlamlı bir matematiksel eşitlik modeli olup olmadığını sınar. Bu modelin anlamlı bulunması durumunda, bağımlı ve bağımsız değişken arasında nasıl bir ilişki olduğunu (ilişkinin yönü ve gücü) R katsayısı ile belirler. Bağımsız (yordayıcı) değişkenin bağımlı (yordanan) değişkenin ne kadarının açıkladığını, determinasyon katsayısı ( $R^2$ ) ile belirler.  $R^2$  bağımsız değişkenin bağımlı değişkendeki varyansın ne kadarını açıkladığını gösteren bir değerdir, yüzde olarak ifade edilir. Modelin anlamlı bulunması durumunda, matematiksel eşitlik denklemindeki katsayılar ile bağımsız değişkenin alacağı değere göre bağımlı değişkenin alacağı değerin tahmin edilmesini sağlar.

$R^2$  bağımlı değişkenin ne kadarının bağımsız değişken tarafından açıklandığı şeklinde yorumlansa da aslında (teknik olarak)  $R^2$  bağımsız ve bağımlı değişken tarafından paylaşılan varyans oranıdır (Field, 2009). Basit Doğrusal Regresyon Analizi de (Pearson'un Korelasyon Testi gibi) teknik olarak bir neden-sonuç ilişkisi analizi değildir. Neden sonuç ilişkisi kurmak ya da bu şekilde yorumlamak üçüncü değişken sorunu nedeni ile sakıncalı olabilir (özellikle sosyal bilimlerde). Ayrıca iki değişken arasındaki ilişkinin, neden sonuç ilişkisi şeklinde değerlendirilebilmesi için kanıt düzeyinde alan bilgisine ihtiyaç vardır.<sup>45</sup>

#### **Basit Doğrusal Regresyon Analizi Denklemi:**

$$Y=A+ BX$$

Y: Bağımlı (yordanan) değişken

A: Sabit katsayı (X'den bağımsız olarak Y'nin var olduğu değer)

X: Bağımsız (yordayıcı) değişken

B: X değişkeninin regresyon denklemindeki katsayısı

#### **Basit Doğrusal Regresyon Analizi Varsayımları:**

- I- Değişkenler niceldir (en az aralık ölçeğinde) ve normal dağılım özelliği gösterirler.
- II- Bağımlı (yordanan) ve bağımsız (yordayıcı) değişkenler arasında doğrusal bir ilişki olmalıdır.
- III- Regresyon analizi sonucunda, tahmin edilen değerlerle gözlenen değerler arasındaki farklar/artıklar (tahmin hataları) normal dağılım sergilemelidir.

Basit Doğrusal Regresyon Analizinin varsayımları ihlal edilemeyecek kriterlerdir. Bu varsayımların yerine getirilememesi durumunda analiz yapılamaz. İki değişken arasındaki doğrusal ilişkinin varlığı Pearson Korelasyon testi ile sınanabilir (bu testin sonucunda  $p<0,05$  ise iki değişken arasında doğrusal ilişkinin var olduğu tespit edilmiş olur).

Regresyon analizi sonucunda, tahmin edilen değerlerle gözlenen değerler arasındaki farklar (tahmin hataları) normal dağılım sergilemelidir. Bu varsayıma kısaca **Farkların Normalliği** diyebiliriz. Normallik varsayımı, tahmin hatalarının, tahmin edilen her bağımlı değişken skoru etrafında normal olarak dağılması durumudur (Tabachnick ve Fidell, 2007). Bu durumun sınanması için, modele bağlı olarak bağımlı değişkenin standartlaştırılmış tahmini değerleri

<sup>45</sup> Üçüncü değişken sorunu ve neden sonuç ilişkisi sorunu için bakınız sayfa:111-112

(**ZPRED**) ile standartlaştırılmış artık değerlerin (bu değerler, gözlemlenen veriler ile modelin öngördüğü değerler arasındaki standartlaştırılmış farklardır) (**ZRESID**) saçılım grafiğine, histogram dağılımına ve Normal P-P Grafiğine bakılır. Saçılım grafikleri: Bu grafiklerde noktaların yatay sıfır çizgisinin etrafında ve rastgele dağılmasını bekleriz, bu durum farkların normal dağılımına işaret eder (Field, 2009; Stevens, 2009)<sup>46</sup>. Farklar normal dağılmış ise histogram dağılımı çan eğrisi şeklinde olmalıdır. Farklar normal dağılmış ise “Normal P-P” Grafiğinde noktalar 45° (45 derecelik) çizginin üstünde ya da çok yakınında olmalıdır (Tranmer ve Diğr., 2020). Farkların normal dağılımı; saçılım grafiği, histogram dağılımı ve Normal P-P Grafiğine göre değerlendirilir. (Saçılım grafiği, histogram dağılımı ve Normal P-P Grafiğinin oluşturulması: Linear Regression Plots penceresinde, Y eksenine ZRESID, X eksenine ZPRED aktarılır; Histogram ve Normal probability plot seçilir.)

### **Basit Doğrusal Regresyon Analizi yapılırken:**

Değişkenlerin normal dağılım özelliği göstermesi gerekmektedir (Bu durum burada sınanmamış, öyle olduğu varsayılmıştır, siz kendi analizinizi yapmadan önce mutlaka bu durumu sınamalısınız).

Bağımlı ve bağımsız değişkenler arasında doğrusal bir ilişki (r) olması gerekmektedir.

Artıkların/farkların normal dağılımı sınanmalıdır.

### **Basit Doğrusal Regresyon Analizi SPSS Uygulaması:**

- 1- **Analyze, Regression, Linear** seçilir.
- 2- **Linear Regression** penceresinde sol taraftaki değişken listesinden **bağımlı değişken Dependent** kutusuna taşınır.  
**Bağımsız değişken Independent** kutusuna taşınır.  
**Method** kısmında **Enter** seçilidir.
- 3- **Plots** butonu tıklanır.  
**Y eksenine ZRESID, X eksenine ZPRED** aktarılır  
**Histogram** ve **Normal probability plot** seçilir ve **Continue** tıklanır
- 4- **Linear Regression** penceresinde  
**OK** Tıklanır.

<sup>46</sup> Saçılım grafiklerin yorumlanması (farkların normalliği, doğrusallığı ve eş varyanslılığı /homoscedasticity) ve örnekleri için EK 2'ye bakınız sayfa: 160

## SPSS çıktısı:

## Regression

Variables Entered/Removed<sup>a</sup>

Model	Variables Entered	Variables Removed	Method
1	IQ puanı <sup>b</sup>	.	Enter

a. Dependent Variable: sınav başarı puanı

b. All requested variables entered.

## Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,677 <sup>a</sup>	,459	,455	10,05222

a. Predictors: (Constant), IQ puanı

**Model Summary:** Regresyon Model Özeti

**R:** Bağımsız ve bağımlı değişken arasındaki basit korelasyon katsayısını ifade eder. Basit Doğrusal Regresyon varsayımlarından biri bağımsız ve bağımlı değişken arasında doğrusal bir ilişkin varlığıdır, bu durum öncesinde Pearson Korelasyon testi ile sınanabileceği gibi Regresyon analizi çıktılarında ki Model Summary'de ki R değeri Pearson Korelasyon testinin r değeridir.

**R Square:**  $R^2$ , Determinasyon katsayısı (korelasyon katsayısının karesi), bağımsız değişkenin bağımlı değişkenin ne kadarını açıkladığının oranını (% olarak) ifade eder. Basit Doğrusal Regresyonda  $R^2$  değeri olarak R Square değeri verilir.

**Adjusted R Square:** Düzeltilmiş  $R^2$  anlamına gelir. Çoklu Doğrusal Regresyon Analizlerinde  $R^2$  olarak bu değer verilir.

**Std. Error of the Estimate:** Tahminin standart hatası.

ANOVA<sup>a</sup>

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	11383,283	1	11383,283	112,653	,000 <sup>b</sup>
	Residual	13439,265	133	101,047		
	Total	24822,548	134			

a. Dependent Variable: sınav başarı puanı

b. Predictors: (Constant), IQ puanı

**ANOVA<sup>a</sup>:** Modelin Anlamlılığının Sınıandığı Test

**df:** Serbestlik derecesi

**F:** Test değeri (Modelin F değeri)

**Sig. :** Modelin p değeri (Regresyon Modelin Anlamlılığının p değeri!)  $p < 0,05$  ise regresyon modeli anlamlıdır.



**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-25,587	9,600		-2,665	,009
	IQ puanı	,990	,093	,677	10,614	,000

a. Dependent Variable: sınav başarı puanı

**Coefficients:** Regresyon denklemindeki katsayılar ve önemlilik testi sonuçları

**(Constant) B:** Sabit katsayı (Denklemdaki A katsayısı, sabit katsayı)

**IQ puanı B:** Denklemdaki bağımsız (yordayıcı) değişkenin (bu örnekte IQ puanının) regresyon denklemindeki katsayısı

**Standardized Coefficients Beta:** Standardize Edilmiş Beta Katsayısı. Bağımsız değişkende 1 standart sapmalık değişimin, Bağımlı değişkeninin 1 standart sapması x Beta katsayısı kadar bağımlı değişkende bir değişime neden olacağını gösterir.

**t:** Regresyon katsayılarının önemlilik testi değeri

**Sig. :** Regresyon katsayılarının önemlilik testi p değeri

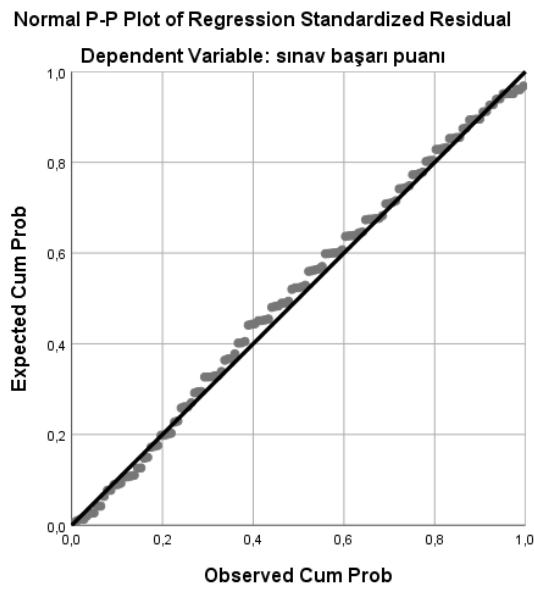
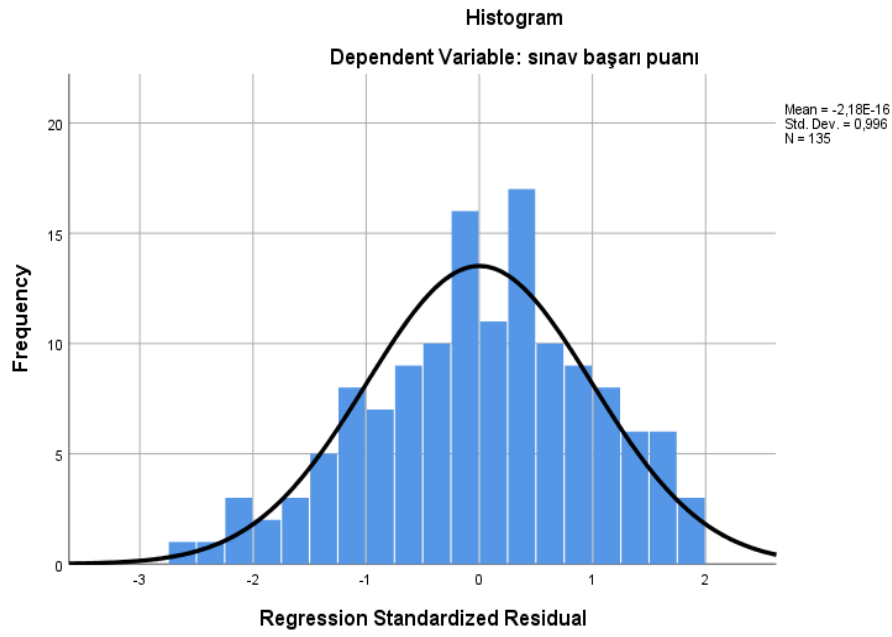
**Residuals Statistics<sup>a</sup>**

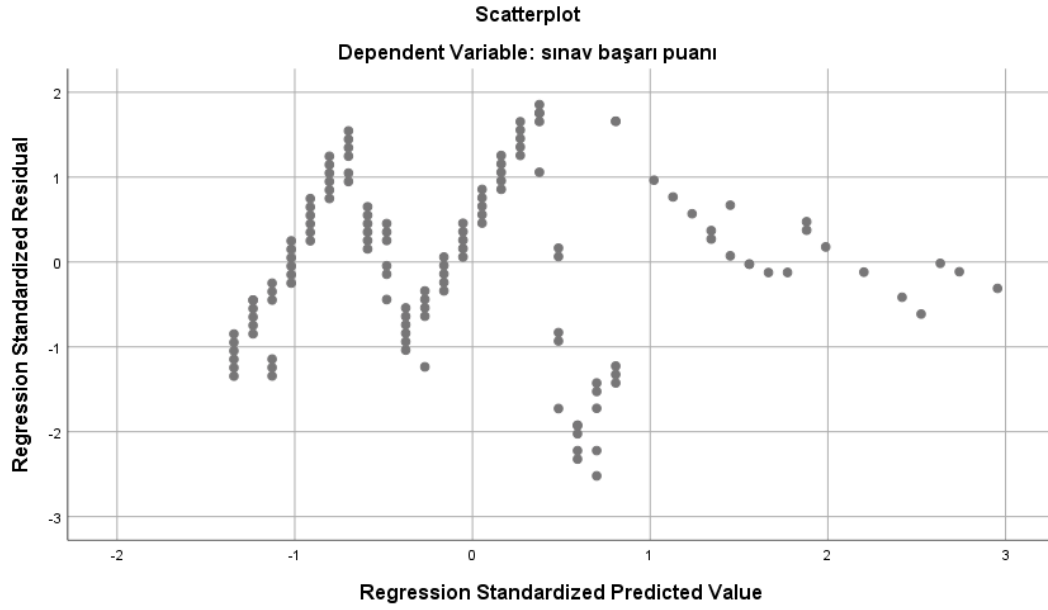
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	63,5235	103,1281	75,8963	9,21683	135
Residual	-25,33570	18,63464	,00000	10,01464	135
Std. Predicted Value	-1,342	2,955	,000	1,000	135
Std. Residual	-2,520	1,854	,000	,996	135

a. Dependent Variable: sınav başarı puanı

**Residuals Statistics:** Fark/Artık İstatistikleri

Regresyon modelinin geçerli olması artıkların/farkların normal dağılımı varsayımının yerine getirilmesine bağlıdır. Farkların/artıkların normalliği; histogram dağılımı, Normal P-P Grafiği ve saçılım grafiğine göre değerlendirilir. Bu analizin, artıkların/farkların histogram dağılımına baktığımızda normal dağılıma çok yakın bir dağılım olduğunu; “Normal P-P” Grafiğinde noktaların 45<sup>0</sup> (45 derecelik) çizginin üstünde ya da çok yakınında olduğunu (normal dağıldığını ifade eder); saçılım grafiğinde de çok belirgin bir ihlalin olmadığını görüyoruz. Üç göstergeyi birlikte değerlendirdiğimizde farkların/artıkların normal dağılım varsayımının sağlandığını görmekteyiz. Dolayısı ile regresyon modeli geçerlidir.



**Sonuç:**

Her bir değişkenin normal dağılım özelliği gösterdiği varsayıldı.

Bağımlı ve bağımsız değişkenler arasında doğrusal bir ilişki var.

Artıklar/farklar normal dağılım özelliği gösteriyor.

**Sonuç: Varsayımlar sağlandığı için Regresyon Analizi sonuçları geçerlidir.**

**Tablo: Sınav başarı puanının Zekâ ile yordanmasına ilişkin Regresyon Analizi**

Değişken	B	Standart Hata B	$\beta$	t	p
Sabit	-25,587	9,600		-2,665	,009
IQ puanı	,990	,093	,677	10,614	,000
R=,677      R <sup>2</sup> =,459					
F(1,133)= 112,653      p= ,000					

IQ puanının Sınav başarı puanının istatistiksel olarak anlamlı bir yordayıcısı (açıklayıcısı) olduğu görülmektedir (R= ,677, R<sup>2</sup>= ,459, F(1,133)= 112,653 p<0,01). Sınav başarı puanına ilişkin toplam varyansın (değişimin) %46'sı IQ puanı ile açıklanabilmektedir.

$$Y=A+ BX$$

$$\text{Sınav başarı Puanı} = -25,587 + (0,990 \times \text{IQ puanı})$$

Y: Sınav başarı Puanı

A: Sabit katsayı (yukarıdaki tabloda Sabit'in B değeri)

B: IQ puanının regresyon denklemindeki katsayısı (yukarıdaki tabloda IQ puanının B değeri)

X: IQ puanı

## 5.2. Çoklu Doğrusal Regresyon Analizi <sup>47</sup>

Çoklu Doğrusal Regresyon Analizi, Basit Doğrusal Regresyon Analizinin bir uzantısıdır (extension) (Coakes, 2005). Çoklu Doğrusal Regresyon, aralarında doğrusal ilişki olan ikiden fazla nicel değişkenden, birinin bağımlı değişken diğerlerinin bağımsız değişkenler olarak kurgulandığı durumda, aralarındaki ilişkinin matematiksel bir eşitlik ile açıklanması sürecidir. Bağımsız değişkenlerin bağımlı değişkenin ne kadarını açıkladığını ve aralarında nasıl bir matematiksel eşitlik olduğunu ifade eder (Büyüköztürk, 2010).

### Çoklu Doğrusal Regresyon Analizi Neyi Yapar:

- 1- Bağımsız (yordayıcı) değişkenler birlikte, bağımlı (yordanan) değişkeni anlamlı bir şekilde açıklayıp açıklamadığını, ne şekilde etkilediğini (pozitif/negatif) ve bağımsız değişkenlerin bağımlı değişken üzerindeki önemliliklerini sıralar.
- 2- Bağımsız (yordayıcı) değişkenlerin, bağımlı (yordanan) değişkenin ne kadarının açıkladığını, determinasyon katsayısı ( $R^2$ ) ile belirler.
- 3- Bağımlı (yordanan) değişken ile bağımsız (yordayıcı) değişkenler arasındaki ilişkiyi regresyon eşitliği ile açıklar. Böylelikle bağımsız (yordayıcı) değişkenlerin alacağı değere göre bağımlı (yordanan) değişkenin alacağı değerin tahmin edilmesini sağlar.

Çoklu Doğrusal Regresyon Analizi, bağımlı ve bağımsız değişkenler arasında anlamlı bir matematiksel eşitlik modeli olup olmadığını sınar. Bu modelin anlamlı bulunması durumunda, bağımlı ve bağımsız değişkenler arasında nasıl bir ilişki olduğunu (ilişkinin yönü ve gücü) R katsayısı ile belirler. Ayrıca her bir bağımsız değişken ile bağımlı değişken arasındaki basit (Zero-order), kısmi (Partial) ve kısım (Part) korelasyon katsayılarını ( $r$ ) da verir. Bağımsız değişkenlerin bağımlı değişken üzerindeki önemliliklerini, Standardize Edilmiş Beta Katsayılarının (Standardized Coefficients Beta) büyüklüğüne göre belirler.

Regresyon denklemindeki Standardize Edilmiş Beta Katsayıları aynı ölçüm standardını kullanır ve bu nedenle bağımlı değişkenle ilişkili olarak iki veya daha fazla bağımsız değişkenden hangisinin daha önemli olduğunu gösterir. Temel olarak bize, bağımsız değişkende bir standart sapma değişimin, bağımlı değişkenin standart sapma birimiyle ne kadar değişeceğini ifade eder (Bryman ve Cramer, 1999).

$R^2$  bağımsız değişkenlerin hepsinin bağımlı değişken üzerindeki kolektif etkisini gösteren bir değerdir (Bryman ve Cramer, 1999). Bağımsız (yordayıcı) değişkenlerin bağımlı (yordanan) değişkenin ne kadarının açıkladığını, determinasyon katsayısı (Düzeltilmiş/Adjusted  $R^2$ ) ile belirler.  $R^2$  bağımsız değişkenlerin bağımlı değişkendeki varyansın ne kadarını açıkladığını gösteren bir değerdir, yüzde olarak ifade edilir (Green ve Salkind, 2014; Kerr ve Diğr., 2002).

Modelin anlamlı bulunması durumunda matematiksel eşitlik denklemindeki sabit ve bağımsız değişkenlerin katsayılarını vererek, bağımsız değişkenlerin alacağı değerlere göre bağımlı değişkenin alacağı değerlerin tahmin edilmesini sağlar.

$R^2$  bağımlı değişkenin ne kadarının bağımsız değişkenler tarafından açıklandığı şeklinde yorumlansa da aslında (teknik olarak)  $R^2$  bağımlı ve bağımsız değişkenler tarafından paylaşılan varyans oranıdır. Çoklu Doğrusal Regresyon Analizi de teknik olarak bir neden-sonuç ilişkisi analizi değildir. Neden sonuç ilişkisi kurmak ya da bu şekilde yorumlamak üçüncü değişken

<sup>47</sup> Multiple Linear Regression

sorunu nedeni ile sakıncalı olabilir (Field, 2009; Tabachnick ve Fidell, 2007). Araştırma sürecinde bağımlı değişken üzerinde etkili olabilecek tüm değişkenlerin ölçülmesi önemlidir, bu şekilde üçüncü değişken sorunu çözülebilir. Ayrıca bağımlı ve bağımsız değişkenler arasındaki ilişkinin, neden sonuç ilişkisi şeklinde değerlendirilebilmesi için kanıt düzeyinde alan bilgisine ihtiyaç vardır.<sup>48</sup> Bağımlı değişken üzerinde etkili olabilecek olası tüm değişkenlerin modele sokulması ve aralarındaki ilişkinin neden-sonuç ilişkisi şeklinde yorumlanmasına imkân tanıyan, kanıt düzeyinde alan bilgisi olması koşulunda bağımsız ve bağımlı değişkenler arasındaki regresyon denklemi neden-sonuç ilişkisi çerçevesinde yorumlanabilir.

### Çoklu Doğrusal Regresyon Analizi Denklemi:

$$Y=A+ B_1 X_1 + B_2 X_2 + B_3 X_3 +.....$$

Y: Bağımlı (yordanan) değişken

A: Sabit katsayı (yordayıcı değişkenlerden bağımsız olarak Y'nin var olduğu değer)

X<sub>1</sub>: 1. Bağımsız (yordayıcı) değişken

B<sub>1</sub>: X<sub>1</sub> değişkeninin regresyon denklemindeki katsayısı

X<sub>2</sub>: 2. Bağımsız (yordayıcı) değişken

B<sub>2</sub>: X<sub>2</sub> değişkeninin regresyon denklemindeki katsayısı

X<sub>3</sub>: 3. Bağımsız (yordayıcı) değişken

B<sub>3</sub>: X<sub>3</sub> değişkeninin regresyon denklemindeki katsayısı

### Çoklu Doğrusal Regresyon Analizi Varsayımları:

- I- Değişkenler niceldir<sup>49</sup> (en az aralık ölçeğinde) ve normal dağılım özelliği gösterir.
- II- Bağımlı ve bağımsız değişkenler arasında doğrusal bir ilişki olmalıdır.
- III- Bağımsız değişkenler arasında **çoklu ortak doğrusallık** (multicollinearity) **olmamalıdır** (Bağımsız değişkenler arasında çok yüksek bir ilişki olmamalıdır).
- IV- Regresyon analizi sonucunda, tahmin edilen değerlerle gözlenen değerler arasındaki farklar/artıklar (tahmin hataları) normal dağılım sergilemelidir.

Çoklu Doğrusal Regresyon Analizinin varsayımları ihlal edilemeyecek kriterlerdir. Bu varsayımların yerine getirilememesi durumunda yapılan analiz geçerli olmaz.

Değişkenlerin normal dağılıma uygunluğu normal dağılım kriterlerine göre sınanmalıdır. Bağımlı ve bağımsız değişkenler arasındaki doğrusal ilişkinin varlığı Pearson Korelasyon testi ile sınanabilir, bu testin sonucunda  $p < 0,05$  ise iki değişken arasında doğrusal ilişkinin var olduğu tespit edilmiş olur. Her bir bağımsız değişken ile bağımlı değişken arasında doğrusal bir ilişkinin olması gerekir.

Çoklu ortak doğrusallık (multicollinearity) bağımsız değişkenler arasındaki yüksek bir ilişkinin olması durumudur. Çoklu ortak doğrusallık  $R^2$  sınırlar, ayrıca regresyon katsayılarının güvenilmez (unstable) olabileceği anlamına gelir (DeCoster ve Claypool 2004; Bryman ve

<sup>48</sup> Üçüncü değişken sorunu ve neden sonuç ilişkisi sorunu için bakınız sayfa:111-112

<sup>49</sup> Bu durumun bir istisnası mevcuttur, bağımsız değişken iki kategoriden; bir durumun varlığı (1) ve yokluğu (0) şeklinde kodlanarak (**dummy/kukla değişken**), oluşuyorsa doğrusal regresyona bağımsız (açıklayıcı) bir değişken olarak sokulabilir. Ancak tüm bağımsız değişkenler arasında ki dummy değişkenlerin oranı çok az olmalıdır.

Cramer, 1999). Bağımsız değişkenler arasında çoklu ortak doğrusallığın (multicollinearity) olması durumunda yapılan analiz geçerli olmaz. Çoklu ortak doğrusallık iki kriterle değerlendirilir<sup>50</sup> (Field, 2009):

1-Bağımsız değişkenler arasındaki yüksek düzeyde doğrusal ilişki çoklu ortak doğrusallık göstergesidir. Bağımsız değişkenler arasındaki  $r$  katsayısının 0,8 ve üzerinde olması çoklu ortak doğrusallık (multicollinearity) olabileceğini,  **$r$  değerinin 0,9 ve üzerinde olması ise ciddi çoklu ortak doğrusallık (multicollinearity) sorunu olduğunu gösterir.** Kıstas olarak  $r$  değerinin 0,9 ve üstünde olmaması gerekir<sup>51</sup> (Field, 2009; Tabachnick ve Fidell, 2007).

2-Doğrusal bir regresyonun Varyans Büyütme Faktörü (Variance Inflation Factor/VIF) bize, çoklu ortak doğrusallık (multicollinearity) nedeniyle regresyon tahminlerinin varyansının ne kadar arttığı hakkında bir fikir verir (Tranmer ve Diğr., 2020). Her bir bağımsız değişkenin Varyans Büyütme Faktörü (VIF) değerinin 10'dan (bazı kaynaklarda 12'den) küçük olması; Tolerans değerinin ise 0,2'den (bazı kaynaklarda 0,1'den) büyük olması gerekir. Tolerans her bir değişken için  $1/VIF$  işleminin sonucunda ulaşılan bir değerdir (Field, 2009; Stevens, 2009). VIF değeri bağımsız değişkenlerin birbirleriyle paylaştıkları varyansla ilgili çoklu ortak doğrusallık (multicollinearity) göstergesidir. VIF değerinin 10'dan büyük olması, Tolerans değerinin 0,2'den küçük olması bağımsız değişkenler arasında çoklu ortak doğrusallığa (multicollinearity) işaret eder.

Regresyon analizi sonucunda, tahmin edilen değerlerle gözlenen değerler arasındaki farklar (tahmin hataları) normal dağılım sergilemelidir. Bu varsayıma kısaca **Farkların Normalliği** diyebiliriz. Normallik varsayımı, tahmin hatalarının, tahmin edilen her bağımlı değişken skoru etrafında normal olarak dağılması durumudur (Tabachnick ve Fidell, 2007). Bu durumun sınanması için, modele bağlı olarak bağımlı değişkenin standartlaştırılmış tahmini değerleri (**ZPRED**) ile standartlaştırılmış artık değerlerin (bu değerler, gözlemlenen veriler ile modelin öngördüğü değerler arasındaki standartlaştırılmış farklardır) (**ZRESID**) saçılma grafiğine, histogram dağılımına ve Normal P-P Grafiğine bakılır. Saçılım grafiği: Bu grafikte noktaların yatay sıfır çizgisinin etrafında ve rastgele dağılmasını bekleriz, bu durum farkların normal

<sup>50</sup> Çoklu ortak doğrusallık sorununa Durum Endeksi değeri üzerinden de karar verilebilir. Her bir bağımsız değişkenin Durum Endeksi (Condition indices/CI) değerinin 30'dan küçük olması gerekir. Durum Endeksi, bir değişkenin diğerlerine karşı bağımlılığının bir ölçüsüdür. Yüksek bir durum endeksi varyans büyütmeyle (variance inflation) ile ilişkilidir. Her bir bağımsız değişkenin durum endeksi değeri bağımlı değişken için tahmininin standart hatasını ifade eder. Standart hatası çok büyük olduğunda, parametre tahmini oldukça belirsizdir. Her bir bağımsız değişken, tahmin edilen her bir bağımlı değişkenin varyansının bir kısmını oluşturur. Yüksek durum endeksine sahip bir bağımsız değişken, iki veya daha fazla değişkenin varyansına güçlü bir şekilde katkıda bulunursa (yüksek bir varyans oranına sahip olduğunda) çoklu ortak doğrusallık sorunu oluşur. Bağımsız değişkenler arasında paylaşılan varyans oranı arttığında Durum Endeksi değeri 30'un üzerine çıkar (Belsely ve Diğr., 1980). Herhangi bir bağımsız değişkenin Durum Endeksi değerinin 30'un üzerine çıkması, bu değişkenin diğer bağımsız değişkenlerle arasında çoklu ortak doğrusallığın olduğunu gösterir. Bu durum görmezden gelinebilir ya da bu değişken analiz dışında tutulur (Tabachnick ve Fidell, 2007). Günümüz literatüründe çoklu ortak doğrusallığa karar vermede Durum Endeksi (Condition indices/CI) artık kullanılmamaktadır. Bunun yerine bağımsız değişkenler arasındaki  $r$  katsayısına, her bir bağımsız değişkenin VIF ve Tolerans değerlerine göre karar verilir. Bağımsız değişkenler arasında paylaşılan varyansla ilgili Durum Endeksi yerine VIF değeri karar vermede yeterli görülmektedir.

<sup>51</sup> Literatürde çoklu ortak doğrusallık çerçevesinde  $r$  değeri ile farklı yaklaşımlar mevcuttur; bazı yazarlar bağımsız değişkenler arası  $r$  katsayısının 0,7'yi geçmemesi gerektiğini (Kerr ve Diğr., 2002), bazıları ise 0,8'i geçmemesi gerektiğini önermektedir (Bryman ve Cramer, 1999; Stevens, 2009).

dağılımına işaret eder (Field, 2009; Stevens, 2009)<sup>52</sup>. Farklar normal dağılmış ise histogram dağılımı çan eğrisi şeklinde olmalıdır. Farklar normal dağılmış ise “Normal P-P” Grafiğinde noktalar 45<sup>0</sup> (45 derecelik) çizginin üstünde ya da çok yakınında olmalıdır (Tranmer ve Diğr., 2020). Farkların normal dağılımı; saçılım grafiği, histogram dağılımı ve Normal P-P Grafiğine göre değerlendirilir. (Saçılım grafiği, histogram dağılımı ve Normal P-P Grafiğinin oluşturulması: Linear Regression Plots penceresinde, Y eksenine ZRESİD, X eksenine ZPRED aktarılır; Histogram ve Normal probability plot seçilir.)

Çoklu Doğrusal Regresyonun Temel Varsayımlarına ek olarak bazı varsayımlarda literatürde kabul görmeye başlamıştır. Yukarıdaki temel varsayımlar sağlıklı bir Çoklu Doğrusal Regresyon analizi için yeterli olsa da literatürde kabul görmeye başlayan varsayımlar<sup>53</sup> şunlardır (temel varsayımlara ek olarak):

**Farkların Doğrusallığı:** Regresyon analizi sonucunda, tahmin edilen değerlerle gözlenen değerler arasındaki farklar (tahmin hataları) arasındaki ilişki doğrusal olmalıdır (Field, 2009; Tabachnick ve Fidell, 2007; Coakes, 2005). Farkların doğrusallığı varsayımı farkların saçılım grafiği ile sınanır.

**Farkların Eş Varyanslığı (Homoscedasticity):** Regresyon analizi sonucunda, tahmin edilen değerlerle gözlenen değerler arasındaki farkların (tahmin hataları) varyansı tüm tahmin edilen değerlerde homojen/eşit olmalıdır (Mangiafico, 2016; Field, 2009; Tabachnick ve Fidell, 2007; Coakes, 2005; Landau ve Everitt 2004). Homoscedasticity farkların/artıkların dağılımını veya hata terimlerini ifade eder. Bu varsayım geçerliyse, tahmin hatalarının sabit bir varyansı vardır -başka bir deyişle, her gözlem için hata model içindeki herhangi bir değişkene bağlı değildir. Bunu söylemenin başka bir yolu, hata terimlerinin standart sapmasının sabit olması ve bağımsız değişken değerlerine bağlı olmamasıdır (Tranmer ve Diğr., 2020). Homoscedasticity varsayımı farkların saçılım grafiği ile sınanır.

Regresyon analizi sonucunda, tahmin edilen değerlerle gözlenen değerler arasındaki farkların (tahmin hataları) doğrusallığı ve eş varyanslığın sınanması için; modele bağlı olarak bağımlı değişkenin standartlaştırılmış tahmini değerleri (**ZPRED**) ile standartlaştırılmış artık değerlerin (bu değerler, gözlemlenen veriler ile modelin öngördüğü değerler arasındaki standartlaştırılmış farklardır) (**ZRESID**) saçılım grafiğine bakılır. Saçılım grafikleri: Bu grafiklerde noktaların yatay sıfır çizgisinin etrafında eşit ve rastgele dağılmasını bekleriz (bu durum farkların normal dağılımını gösterdiği gibi aynı zamanda farkların doğrusallığı ve eş varyanslığını da gösterir). Bu grafiklerde noktalar huni şeklinde belirginleşirse bu heteroscedasticity işaretidir ve homoscedasticity varsayımının ihlal edildiğini gösterir. Bu grafiklerde noktalar eğrilik

<sup>52</sup> Saçılım grafiklerin yorumlanması (farkların normallığı, doğrusallığı ve eş varyanslığı /homoscedasticity) ve örnekleri için EK 2’ye bakınız sayfa:160

<sup>53</sup> Bu kitapta ek varsayımlar olarak yer verilen bu varsayımlar, Tranmer ve Diğr. (2020), Field, (2009), Tabachnick ve Fidell (2007) tarafından temel varsayımlar olarak değerlendirilmektedir. Detaylı bilgi için ilgili kaynaklara bakılabilir.

oluşturacak şekilde belirginleşiyorsa bu da –farkların- doğrusallık varsayımının ihlal edildiğini işaret eder (Field, 2009; Stevens, 2009).<sup>54</sup>

**Regresyon analizinin varsayımları olmamakla birlikte örneklem büyüklüğü ve uç değerler literatürde tartışılan hususlardır.**

#### **Yeterli Örneklem büyüklüğü:**

Güvenilir bir regresyon modeli elde etmek ve sonuçlarını genelleştirebilmek için yeterli bir örneklem büyüklüğü ile çalışmak önemlidir. Regresyon modeli için örneklem büyüklükleri ile ilgili farklı yaklaşımlar mevcuttur. Bu kitapta sadece çapraz geçerlilik yaklaşımı (cross validation) çerçevesinde örneklem büyüklüğü hesaplanmasına yer verilmiştir. Bu yaklaşıma göre örneklem büyüğünün (n) hesaplanmasında, bağımsız değişken sayısı (m) önemlidir ve n/m oranı örneklem büyüklüğünün hesaplanmasında belirleyicidir (Tabachnick ve Fidell 2007). Stevens (2009), n/m oranının 15 olmasının, örneklem büyüklüğü için minimum  $n \geq 15m$  önermiştir. Stevens her bir bağımsız değişken için 15 katılımcının olmasının yeterli olduğunu ileri sürmüştür (örneğin, 5 bağımsız değişkenin olduğu bir regresyonda minimum 75 örneklem büyüklüğü gerekir). Green (1991) ise örneklem büyüklüğü için  $n \geq 50 + 8m$  ya da  $n \geq 104 + m$  formüllerini önermiştir (bu formüllerin sonucunda büyük olan örneklem sayısının tercih edilmesini). Green'in minimum örneklem büyüklüğü için 5 bağımsız değişkenin olduğu durumda  $n \geq 50 + 8m = 50 + (8 \times 5) = 90$  ya da  $n \geq 104 + m = 104 + 5 = 109$  örneklem büyüklüklerine ulaşp, Green'in önerisi doğrultusunda minimum 109 örneklem büyüklüğünü tercih ederiz. Stevens (2009) ve Green'in (1991) minimum örneklem büyüklüğü hesaplamaları  $\alpha:0,05$  ve güç (power) 0,80 kapsamındadır.

Çapraz geçerlilik yaklaşımı (cross validation) çerçevesinde örneklem büyüklüğü hesaplamalarına ilişkin literatürde eleştiriler de mevcuttur. Eleştiriler, çapraz geçerlilik yaklaşımı (cross validation) ile ulaşılan örneklem büyüklüklerinin determinasyon katsayısının ( $R^2$ ) düşük olduğu durumlarda bu yaklaşımın örneklem büyüklüklerinin regresyon modeli sonuçlarını değersiz ve güvenilmez kılabilir kadar küçük örneklem sayıları olduğu yönündedir -determinasyon katsayısının ( $R^2$ ) orta derece ve üstü olduğu (0,40 ve üstü) durumlarda güvenilir olduğu- (Brooks ve Barcikowski, 2012; Knofczynski ve Mundfrom, 2008).<sup>55</sup> Güvenilir bir Regresyon Analizi için beklenen  $R^2$  ihtimalleri göz önüne alınarak, bağımsız değişken sayısı ve istenen güce göre minimum örneklem büyüklüğünün hesaplanmasında fayda vardır (İnternet ortamında bu hesaplamaları yapan birçok web sitesi mevcuttur). Ayrıca herhangi bir Örneklemin veri seti/datası üzerinden yapılan test sonucunun genelleştirilebilmesi için Örneklemin Evreni temsil edebilecek bir niceliksel büyüklükte ve nitelikte olması gerekmektedir.

<sup>54</sup> Saçılım grafiklerin yorumlanması (farkların normalliği, doğrusallığı ve eş varyanslığı /homoscedasticity) ve örnekleri için EK 2'ye bakınız sayfa:160

<sup>55</sup> Eleştirilerin detayı ve diğer örneklem büyüklüğü hesaplama yaklaşımları ile ilgili bilgi için Knofczynski ve Mundfrom (2008) ve Brooks ve Barcikowski (2012)'ye bakınız.



**Uç değerlerin olmaması beklenir:**

Değişkenlerin uç değerlerinin regresyon analizi üzerinde çok fazla etkisi vardır ve ayrıca regresyon katsayılarının tahmininin kesinliğini de etkiler. Uç değerler regresyon katsayılarının standart hatalarını artırır, dolayısı ile regresyon denklemindeki standart hatayı artırır. Regresyon analizi tahmininin doğruluğunu artırmak ve hatayı minimize etmek için uç değerlerin analiz dışında tutulması, veri setinden silinmesi önerilmektedir (Tabachnick ve Fidell, 2007; Coakes, 2005). Kast edilen aşırı uç değerler ortalama 3 standart sapma büyük/küçük olan değerlerdir. Bu öneri gerekçeleri açısından doğru olmakla birlikte, örneklem veri setine müdahalenin kapısını açtığı için sakıncalı olabilir. Öncelikle, Örneklemin Evrenden rastgele bir modelle seçildiği ve ardından regresyon analizine dahil edilecek değişken verilerinin normal dağılımı sınanıp, değişkenlerin normal dağılım özelliği gösterdiği bir durumda (bu durumda aşırı uç değerler olsa dahi normalliğin ihlal edilmediğine karar verilmiştir), aşırı uç değerlerin silinerek regresyon analizi yapılması etik açıdan sakıncalı olabilir. Evrendeki her bir birimin Örnekleme dahil olma ihtimalini eşit kılan (rastgele/tesadüfi) bir modelle Örneklemin belirlenmesinden sonra Örneklemin Evreni temsiliyeti sağlanmış olur (niceliksel örneklem büyüklüğü de sağlanmış ise). Örneklemden elde edilen veriler elektronik ortama aktarıldıktan sonra verilerin kontrolü, toplanan verinin elektronik ortama doğru aktarılıp aktarılmadığı hususunda yapılması gerekir. Örnekleme ait veriler doğru bir şekilde elektronik ortama aktarılıp veri setine dönüştürüldükten sonra analiz aşamasına geçilince veri setine müdahale edilmesi, Örneklemin Evrendeki gerçekliğinden uzaklaşmasını (muhtemelen Örnekleme var olan bu uç değerler Evrende de vardır<sup>56</sup>) sağlayabileceği nedeni ile de yazar tarafından tavsiye edilmemektedir.

---

<sup>56</sup> İdealize edilen normal dağılımda değerlerin %99,74'ü ortalamanın  $\pm 3$  standart sapma aralığında yer alır. İdealize edilen normal dağılımda değerlerin %0,26'sı (binde 26'sı) ortalamanın  $\pm 3$  standart sapma aralığının dışında yer alır. Yani idealize edilen normal dağılımda değerlerin yaklaşık %0,3 (binde 3'ü) uç değerlerdir. Bu durum idealize edilen normal dağılıma ait teorik beklenen bir durumdur. Alan araştırmalarında Örnekleme aracılığı ile nicel veriler toplanırken söz konusu nicel veriler normal dağılım özelliği gösterdiğinde içinde idealize edilen normal dağılımdan daha fazla uç değerlerin olması muhtemeldir.

**Çoklu Doğrusal Regresyon Analizi Yapılırken:**

Her bir değişkenin normal dağılım özelliği göstermesi gerekmektedir (Bu durum burada sınanmamış, öyle olduğu varsayılmıştır, siz kendi analizinizi yapmadan önce mutlaka bu durumu sinamalısınız).

Bağımlı ve bağımsız değişkenler arasında doğrusal bir ilişki (r) olmalı,

Bağımsız değişkenler arasında çoklu ortak doğrusallık (Multicollinearity) olmamalı,

Artıkların/farkların normal dağılımı, sınanmalıdır.

**Çoklu Doğrusal Regresyon Analizi SPSS Uygulaması:**

**1-Analyze, Regression, Linear** seçilir.

**2-Linear Regression** penceresinde sol taraftaki değişken listesinden **bağımlı değişken** **Dependent** kutusuna taşınır.

**Bağımsız değişkenler Independent** kutusuna taşınır.

**Method** kısmında **Enter** seçilidir.

**3-Statistics** butonu tıklanır.

**Estimates**

**Model fit**

**Descriptive**

**Part and partial correlations**

**Collinearity diagnostics** seçilir

ve **Continue** tıklanır

**4-Plots** butonu tıklanır.

**Y** eksenine **ZRESID**, **X** eksenine **ZPRED** aktarılır

**Histogram** ve **Normal probability plot** seçilir ve **Continue** tıklanır

**5- Linear Regression** penceresinde

**OK** Tıklanır

**SPSS çıktısı:**

**Analiz çıktısında öncelikli olarak varsayımlar sınanmalıdır.**

**Regression****Descriptive Statistics**

	Mean	Std. Deviation	N
sınav başarı puanı	75,8963	13,61040	135
özgüven puanı	70,9852	17,80491	135
mutluluk puanı	69,7481	17,47822	135
sınav kaygısı puanı	77,4000	10,43902	135
günlük harçlık TL	32,8593	9,91393	135
haftalık çalışma saati	52,2519	12,37638	135
IQ puanı	102,4963	9,30886	135

**Correlations**

		sınav başarı puanı	özgüven puanı	mutluluk puanı	sınav kaygısı puanı	günlük harçlık TL	haftalık çalışma saati	IQ puanı
Pearson Correlation	sınav başarı puanı	1,000	,980	,879	-,792	-,604	,910	,677
	özgüven puanı	,980	1,000	,894	-,776	-,597	,910	,630
	mutluluk puanı	,879	,894	1,000	-,677	-,621	,809	,560
	sınav kaygısı puanı	-,792	-,776	-,677	1,000	,373	-,741	-,526
	günlük harçlık TL	-,604	-,597	-,621	,373	1,000	-,581	-,351
	haftalık çalışma saati	,910	,910	,809	-,741	-,581	1,000	,472
	IQ puanı	,677	,630	,560	-,526	-,351	,472	1,000
Sig. (1-tailed)	sınav başarı puanı	.	,000	,000	,000	,000	,000	,000
	özgüven puanı	,000	.	,000	,000	,000	,000	,000
	mutluluk puanı	,000	,000	.	,000	,000	,000	,000
	sınav kaygısı puanı	,000	,000	,000	.	,000	,000	,000
	günlük harçlık TL	,000	,000	,000	,000	.	,000	,000
	haftalık çalışma saati	,000	,000	,000	,000	,000	.	,000
	IQ puanı	,000	,000	,000	,000	,000	,000	.

**Correlations:** Değişkenler arası korelasyon. Bu tablo bize iki şeyi söyler:

1.si bağımlı ve bağımsız değişkenler arasındaki doğrusal bir ilişkiyi (dolgu rengi yeşil olan sütun). Bağımlı ve bağımsız değişkenler arasında p değerlerinin 0,05'ten küçük olmasına göre bağımlı ve bağımsız değişkenler arasındaki doğrusal ilişkiye karar verilir. Bu analizdeki bağımlı ve bağımsız değişkenler arasındaki tüm p değerleri 0,05'ten küçüktür. Dolayısı ile bağımlı ve bağımsız değişkenler arasında doğrusal bir ilişki vardır.

2.si bağımsız değişkenler arasındaki ilişki düzeyini vererek (dolgu rengi kahve), bağımsız değişkenler arasında çoklu ortak doğrusallığın olup olmadığını söyler. Bağımsız değişkenlerin korelasyon katsayıları itibari ile:

Özgüven ile Mutluluk arasında ( $r=0,89$ ) ve Özgüven ile Haftalık çalışma ( $r=0,91$ )

Mutluluk ile Özgüven arasında ( $r=0,89$ )

Bu durumda **Özgüven** değişkenini model dışında tutmalıyız.

**Not:** Correlations tablosunun dolgu renklendirmeleri yazar tarafından yapılmıştır.

#### Variables Entered/Removed<sup>a</sup>

Model	Variables Entered	Variables Removed	Method
1	IQ puanı, günlük harçlık TL, sınav kaygısı puanı, mutluluk puanı, haftalık çalışma saati, özgüven puanı <sup>b</sup>	.	Enter

a. Dependent Variable: sınav başarı puanı

b. All requested variables entered.

#### Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,987 <sup>a</sup>	,974	,973	2,25017

a. Predictors: (Constant), IQ puanı, günlük harçlık TL, sınav kaygısı puanı, mutluluk puanı, haftalık çalışma saati, özgüven puanı

#### ANOVA<sup>a</sup>

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	24174,453	6	4029,076	795,750	,000 <sup>b</sup>
	Residual	648,095	128	5,063		
	Total	24822,548	134			

a. Dependent Variable: sınav başarı puanı

b. Predictors: (Constant), IQ puanı, günlük harçlık TL, sınav kaygısı puanı, mutluluk puanı, haftalık çalışma saati, özgüven puanı

Coefficients <sup>a</sup>											
		Unstandardized Coefficients		Standardized Coefficients			Correlations			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	17,959	4,936		3,638	,000					
	özgüven puanı	,521	,040	,682	13,020	,000	,980	,755	,186	,074	13,443
	mutluluk puanı	,008	,026	,010	,306	,760	,879	,027	,004	,188	5,324
	sınav kaygısı puanı	-,077	,031	-,059	-2,524	,013	-,792	-,218	-,036	,369	2,707
	günlük harçlık TL	-,043	,026	-,031	-1,671	,097	-,604	-,146	-,024	,574	1,741
	haftalık çalışma saati	,177	,041	,161	4,336	,000	,910	,358	,062	,147	6,789
	IQ puanı	,181	,029	,124	6,326	,000	,677	,488	,090	,534	1,872

a. Dependent Variable: sınav başarı puanı

**Coefficients** tablosunda **Collinearity Statistics** değerleri (**Tolerance ve VIF**) çoklu ortak doğrusallığın sınıandığı değerlerdir: Varyans Büyütme Faktörü (VIF) değerlerinin 10'dan küçük olması, Tolerans değerlerinin ise 0,2'den büyük olması gerekir.

Bu koşulları sağlamayan bağımsız değişkenleri Modelden çıkarmalıyız (çoklu ortak doğrusallık olmaması için).

Buna göre:

**Özgüven'in her iki değeri de kritere uymuyor! (Tolerance ve VIF)**

**Mutluluk puanı bir kritere uymuyor (Tolerance)**

**Haftalık çalışma saati bir kritere uymuyor (Tolerance)**

**Collinearity Diagnostics<sup>a</sup>**

Model	Dimension	Eigenvalue	Variance Proportions							
			Condition Index	(Constant)	özgüven puanı	mutluluk puanı	sınav kaygısı puanı	günlük harçlık TL	haftalık çalışma saati	IQ puanı
1	1 (Constant)	6,758	1,000	,00	,00	,00	,00	,00	,00	,00
	2 özgüven	,191	5,949	,00	,00	,01	,00	,11	,00	,00
	3 mutluluk	,030	15,095	,00	,01	,00	,10	,63	,01	,00
	4 kaygısı	,011	25,252	,00	,00	,51	,00	,03	,34	,01
	5 harçlık	,007	31,412	,00	,04	,27	,09	,06	,09	,37
	6 çalışma	,003	49,306	,02	,92	,19	,22	,01	,33	,19
	7 IQ	,001	76,357	,98	,02	,02	,58	,15	,22	,43

a. Dependent Variable: sınav başarı puanı

Not: Bu tablonun SPSS orijinal çıktısında Dimension sütunun altındaki satırlarda yalnızca 1,2,3,4,5,6,7 yazıyordu. Bu rakamların anlamı 1 Constant ve sırası ile Constant'ın yanında (sağında) yer alan bağımsız değişkenler, özgüven puanı 2, Mutluluk puanı 3 şeklinde yazar tarafından yazıldı (dimension sütunundaki değerlerin yanına).

**Durum Endeksi (Condition indices CI)** değerlerinin 30'dan küçük olması beklenir: Harçlık, çalışma saati ve IQ puanlarının CI değerleri 30'dan büyük! Bununla birlikte Durum Endeksi değerleri göz ardı edilebilecek ve günümüz literatüründe çoklu ortak doğrusallık için bir "altın standardı" özelliği taşımaz.

**Çoklu ortak doğrusallık sorunu gösteren bağımsız değişkenler:**

**Mutluluk puanı** bir kritere uymuyor (**Tolerance**)

**Haftalık çalışma saati** bir kritere uymuyor (**Tolerance**)

**Özgüven'in Tolerance, VIF ve Korelasyon ilişki düzeyi (r) uygun değil!**

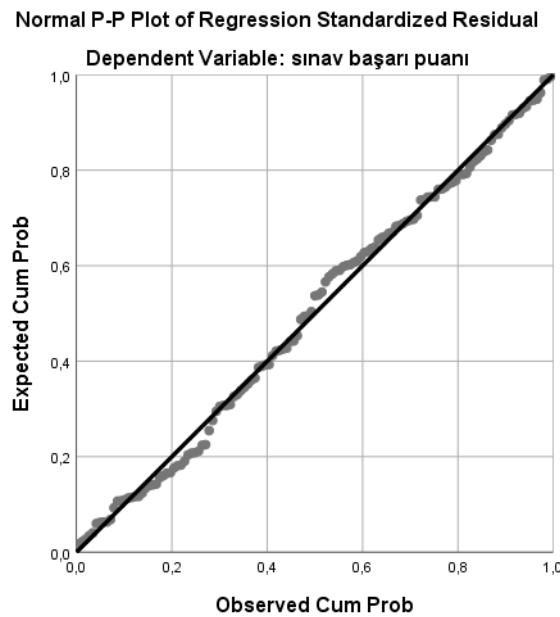
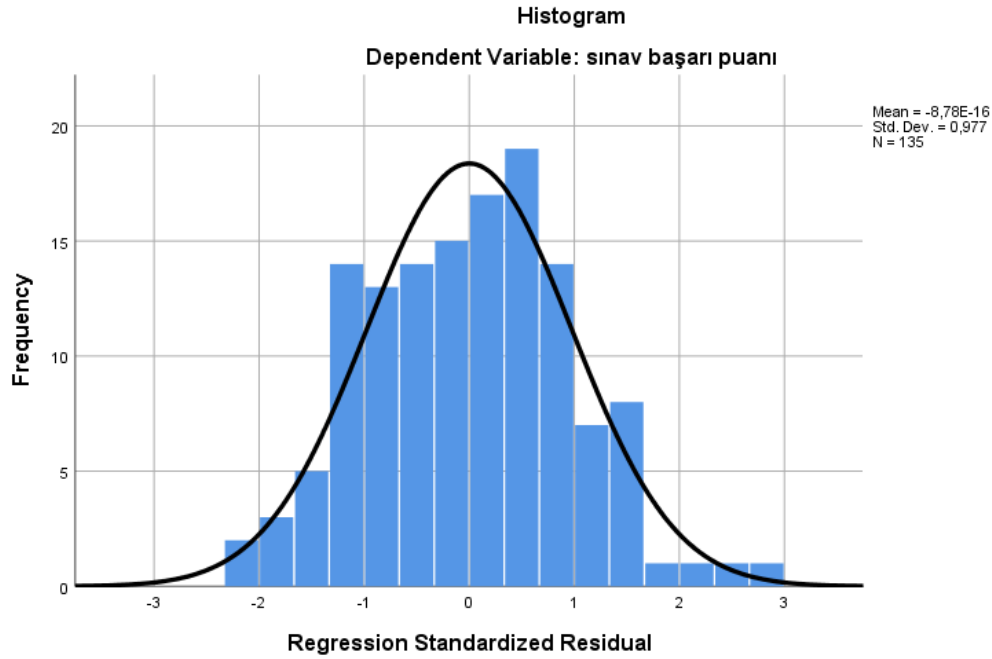
**Residuals Statistics<sup>a</sup>**

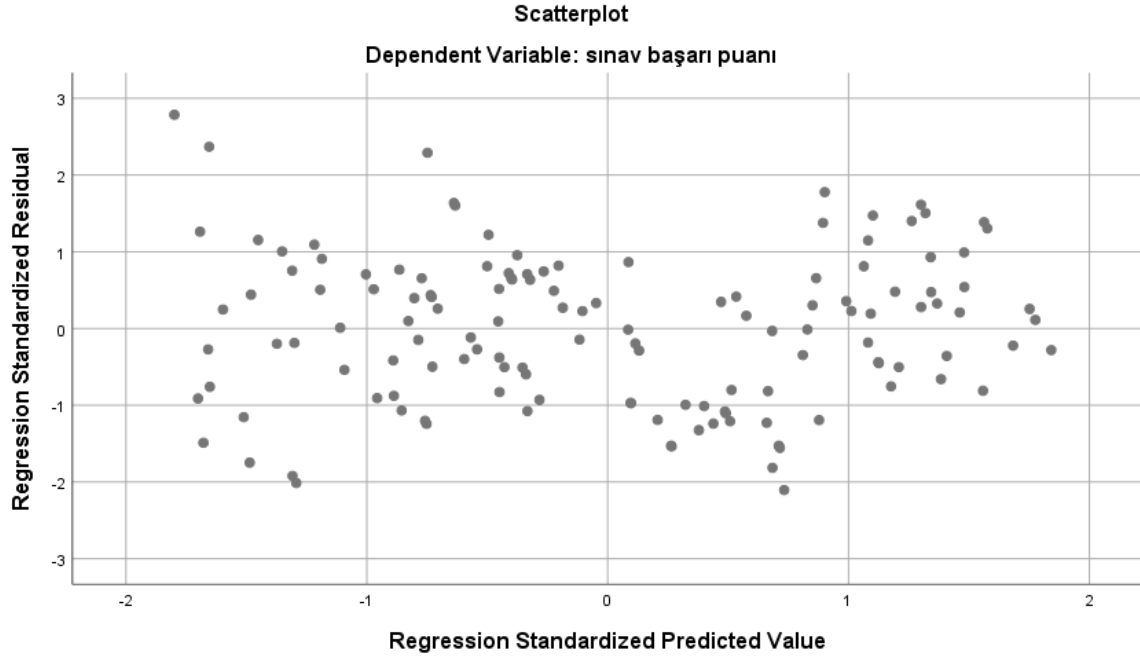
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	51,7322	100,6319	75,8963	13,43154	135
Residual	-4,73631	6,26784	,00000	2,19921	135
Std. Predicted Value	-1,799	1,842	,000	1,000	135
Std. Residual	-2,105	2,786	,000	,977	135

a. Dependent Variable: sınav başarı puanı

**Residuals Statistics:** Fark/Artık İstatistikleri

Regresyon modelinin geçerli olması artıkların/farkların normal dağılımı varsayımının yerine getirilmesine bağlıdır. Farkların/artıkların normalliği; histogram dağılımı, Normal P-P Grafiği ve saçılım grafiğine göre değerlendirilir. Bu analizde artıkların/farkların histogram dağılımına baktığımızda normal dağılıma çok yakın bir dağılım olduğunu; “Normal P-P” Grafiğinde noktaların  $45^0$  (45 derecelik) çizginin üstünde ya da çok yakınında olduğunu (normal dağıldığını ifade eder); saçılım grafiğinde de çok belirgin bir ihlalin olmadığını görüyoruz. Üç göstergeyi birlikte değerlendirdiğimizde farkların/artıkların normal dağılım varsayımının sağlandığını görmekteyiz.





### Sonuç:

Her bir değişkenin normal dağılım özelliği gösterdiği varsayıldı.

Bağımlı ve bağımsız değişkenler arasında doğrusal bir ilişki var.

Bağımsız değişkenler arasında çoklu ortak doğrusallık (Multicollinearity) sorunu var!

Artıklar/farklar normal dağılım özelliği gösteriyor.

### Sonuç: Çoklu ortak doğrusallık sorunu nedeni ile Özgüven değişkenini çıkararak analizi yeniden yapılacaktır!

Not: Mutluluk puanı ve Haftalık çalışma saati değişkenlerinin de Tolerance değerleri uygun değil farkındayız. Ayrıca Mutluluk ve Özgüven arasında  $r$ 'de yüksek (ancak özgüveni çıkaracağımız için bu sorun değil). Tek seferde çok değişken kaybetmek istemediğimizden ve ayrıca bir değişken çıkınca, çıkan değişkenin etkisi ile diğer değişkenlerin VIF ve Tolerance değerlerinin de yeni analizde değişeceğini bildiğimizden şu aşamada analizden çıkarmayı tercih etmiyoruz. Ayrıca yeni analizde her bir değişkenin VIF ve Tolerance değerleri yeniden değerlendirilecektir.



## Özgüven değişkenini çıkarıldıktan sonra

## Çoklu Doğrusal Regresyon Analizi SPSS Çıktısı

## Regression

## Descriptive Statistics

	Mean	Std. Deviation	N
sınav başarı puanı	75,8963	13,61040	135
mutluluk puanı	69,7481	17,47822	135
sınav kaygısı puanı	77,4000	10,43902	135
günlük harçlık TL	32,8593	9,91393	135
haftalık çalışma saati	52,2519	12,37638	135
IQ puanı	102,4963	9,30886	135

## Descriptive Statistics: Tanımlayıcı İstatistikler

**Mean:** Ortalama**Std. Deviation:** Standart Sapma**N:** örneklem sayısı

## Correlations

		sınav başarı puanı	mutluluk puanı	sınav kaygısı puanı	günlük harçlık TL	haftalık çalışma saati	IQ puanı
Pearson Correlation	sınav başarı puanı	1,000	,879	-,792	-,604	,910	,677
	mutluluk puanı	,879	1,000	-,677	-,621	,809	,560
	sınav kaygısı puanı	-,792	-,677	1,000	,373	-,741	-,526
	günlük harçlık TL	-,604	-,621	,373	1,000	-,581	-,351
	haftalık çalışma saati	,910	,809	-,741	-,581	1,000	,472
	IQ puanı	,677	,560	-,526	-,351	,472	1,000
Sig. (1-tailed)	sınav başarı puanı	.	,000	,000	,000	,000	,000
	mutluluk puanı	,000	.	,000	,000	,000	,000
	sınav kaygısı puanı	,000	,000	.	,000	,000	,000
	günlük harçlık TL	,000	,000	,000	.	,000	,000
	haftalık çalışma saati	,000	,000	,000	,000	.	,000
	IQ puanı	,000	,000	,000	,000	,000	.

**Correlations:** Değişkenler arası korelasyon. Bu tablo bize iki şeyi söyler:

1.si bağımlı ve bağımsız değişkenler arasındaki doğrusal bir ilişkiyi (dolgu rengi yeşil olan sütun). Bağımlı ve bağımsız değişkenler arasında p değerlerinin 0,05'ten küçük olmasına göre bağımlı ve bağımsız değişkenler arasındaki doğrusal ilişkiye karar verilir. Bu analizdeki bağımlı ve bağımsız değişkenler arasındaki tüm p değerleri 0,05'ten küçüktür. Dolayısı ile bağımlı ve bağımsız değişkenler arasında doğrusal bir ilişki vardır.

2.si bağımsız değişkenler arasındaki ilişki düzeyini vererek (dolgu rengi kahve), bağımsız değişkenler arasında çoklu ortak doğrusallığın olup olmadığını söyler. Bağımsız değişkenlerin korelasyon katsayıları itibari ile:

Bağımsız değişkenlerin birbirleri ile ilişki düzeylerinde bir sorun yok (r değerlerinin hepsi 0,9 değerinin altında).

**Not:** Correlations tablosunun dolgu renklendirmeleri yazar tarafından yapılmıştır.

#### Variables Entered/Removed<sup>a</sup>

Model	Variables Entered	Variables Removed	Method
1	IQ puanı, günlük harçlık TL, sınav kaygısı puanı, mutluluk puanı, haftalık çalışma saati <sup>b</sup>	.	Enter

a. Dependent Variable: sınav başarı puanı

b. All requested variables entered.

**Variables Entered/Removed<sup>a</sup>:** Regresyon yöntemi gösterir, Enter yöntemi ile Regresyon Analizi yapılmıştır.

#### Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,969 <sup>a</sup>	,939	,937	3,41733

a. Predictors: (Constant), IQ puanı, günlük harçlık TL, sınav kaygısı puanı, mutluluk puanı, haftalık çalışma saati

**Model Summary:** Regresyon Model Özeti

**R değeri:** Bağımsız değişkenler ile bağımlı değişken arasındaki kolektif korelasyon katsayısını ifade eder.

**R Square:** R<sup>2</sup>: Bağımsız değişkenlerin bağımlı değişkenin ne kadarını açıkladığını (% olarak) ifade eder.

**Adjusted R Square:** R<sup>2</sup>: Düzeltilmiş R<sup>2</sup> anlamına gelir. Çoklu Doğrusal Regresyon Analizlerinde R<sup>2</sup> olarak bu değer verilir.

**ANOVA<sup>a</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	23316,066	5	4663,213	399,311	,000 <sup>b</sup>
	Residual	1506,483	129	11,678		
	Total	24822,548	134			

a. Dependent Variable: sınav başarı puanı

b. Predictors: (Constant), IQ puanı, günlük harçlık TL, sınav kaygısı puanı, mutluluk puanı, haftalık çalışma saati

**ANOVA<sup>a</sup>:** Modelin Anlamlılığının Sınındığı Test

**F:** Test değeri (Modelin F değeri)

**df:** Serbestlik derecesi

**Sig. :** Testin p değeri (Modelin Anlamlılığının p değeri!)

ANOVA tablosu modelimizin istatistiksel olarak anlamlı olup olmadığını söyler.  $p < 0,05$  olduğu için modelimizdeki bağımsız değişkenlerin toplamı-ının etkisinin- bağımlı değişken için açıklayıcı bir modeldir. Bağımsız değişkenlerin birlikte bağımlı değişken üzerinde anlamlı ve açıklayıcı olduğunu söyler.

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients		Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta	t		Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	16,011	7,493		2,137	,035					
	mutluluk puanı	,193	,032	,248	5,944	,000	,879	,464	,129	,271	3,691
	sınav kaygısı puanı	-,173	,045	-,133	-3,827	,000	-,792	-,319	-,083	,392	2,551
	günlük harçlık TL	-,058	,039	-,042	-1,466	,145	-,604	-,128	-,032	,575	1,738
	haftalık çalışma saati	,528	,047	,480	11,277	,000	,910	,705	,245	,260	3,852
	IQ puanı	,333	,040	,228	8,407	,000	,677	,595	,182	,641	1,559

a. Dependent Variable: sınav başarı puanı

**Coefficients<sup>a</sup>:** Regresyon denklemindeki katsayılar ve önemlilik testi sonuçları

**Unstandardized Coefficients B değerleri:** Sabit katsayısı (Constant) ve her bir bağımsız değişkenin bağımlı değişkeni etkileme katsayılarını verir. Regresyon denkleminin değerleri (katsayıları) bu sütundaki değerlerdir.

**Standardized Coefficients Beta:** Standardize Edilmiş Beta Katsayısı. Bağımsız değişkenin bağımlı değişkeni etkileme Beta katsayılarını ifade eder (bu aynı zamanda hangi bağımsız değişkenlerin bağımlı değişken üzerinde daha çok önemli etki yaptığını da bize verir). Aynı

zamanda Beta değeri: Bağımsız değişkendeki 1 standart sapmalık değişimin, Bağımlı değişkeninin 1 standart sapması x Beta katsayısı kadar bağımlı değişkende bir değişime neden olacağını gösterir.

**t:** Her bir değişkenin regresyon katsayılarının önemlilik test değerleri

**Sig. Değerleri (p):** Her bir değişkenin model içinde istatistiksel olarak anlamlı/önemli bir etkiye sahip olup olmadığını ifade eder. Regresyon katsayılarının önemlilik testi p değeridir.

Regresyon denklemindeki sabit katsayısının (Constant) önemlilik testi sonucu burada anlamlı bulunmuş ( $p < 0,05$ ). Sabit katsayısının önemlilik testi sonucu  $p > 0,05$  olması durumunda ise model anlamlı (ANOVA tablosunun p değeri,  $p < 0,05$ ) ise analiz geçerlidir.

Her bir bağımsız değişkenin önemlilik testi sonucundaki p değerleri değişkenin regresyon denklemi için anlamlı olup olmadığını ifade eder. Günlük harçlık hariç tüm bağımsız değişkenler regresyon modeli için anlamlıdır ( $p < 0,05$ ). Günlük harçlık değişkeninin önemlilik testi sonucu anlamsız bulunmuştur ( $p > 0,05$ ). Bu durumda Günlük harçlık değişkeninin sonuçlarının çıkarılıp, geri kalan değişkenlere ait değerlerin verilmesi doğru değildir. Çünkü regresyon analizinde tek başına bu değişkenin denklem içindeki katsayısı anlamlı bulunmasa da, regresyon analizindeki R,  $R^2$  değerleri tüm bağımsız değişkenlerin ortak ilişkisi/etkisi ile ortaya çıkmıştır (günlük harçlık dahil). Ayrıca tüm değişkenler birlikte regresyon analizine sokulduğundan, bir değişkenin regresyon katsayısı önemlilik sonucu anlamlı bulunmasa da bu değişkenin, sabit katsayı ve diğer değişkenlerin katsayıları üzerinde etkisi vardır.

Bu durumda iki şey yapabilirsiniz:

- 1- Önemlilik testi sonucu anlamlı bulunmayan değişkenle birlikte regresyon sonuçlarını vermek,
- 2- Önemlilik testi sonucu anlamlı bulunmayan değişkeni çıkarmak istiyorsanız, bu değişken çıkarıldıktan sonra analiz yeniden yapılmalı ve yeni yapılan regresyon analizinin sonuçları değerlendirilmelidir.

## Correlations

**Zero-order:** Bağımsız değişken ile bağımlı değişken arasındaki basit korelasyon katsayısı.

**Partial:** Bağımsız değişken ile bağımlı değişken arasındaki kısmi korelasyon katsayısı.

**Collinearity Statistics** değerleri (**Tolerance ve VIF**) çoklu ortak doğrusallığın sınındığı değerlerdir:

**Tolerance:** Her bir bağımsız değişkenin tolerans değerlerinin 0,2'den büyük olması gerekir.

**VIF:** VIF değerlerinin 10'dan küçük olması gerekir.

Bu koşulları sağlamayan bağımsız değişkenleri Modelden çıkarmalıyız (çoklu ortak doğrusallık olmaması için).

Bu analizde bağımsız değişkenlerin hepsinin VIF ve Tolerance değerleri uygundur, yani çoklu ortak doğrusallık sorunu göstermiyor.

**Collinearity Diagnostics<sup>a</sup>**

Model	Dimension	Eigenvalue	Condition Index	(Constant)	Variance Proportions				
					mutluluk puanı	sınav kaygısı puanı	günlük harçlık TL	haftalık çalışma saati	IQ puanı
1	1 Constant	5,796	1,000	,00	,00	,00	,00	,00	,00
	2 mutluluk	,158	6,061	,00	,02	,00	,14	,02	,00
	3 kaygı	,028	14,427	,00	,01	,12	,61	,04	,00
	4 harçlık	,011	23,406	,00	,69	,00	,03	,65	,01
	5 çalışma	,006	29,961	,00	,26	,17	,06	,04	,56
	6 IQ	,001	70,189	,99	,00	,71	,16	,25	,43

a. Dependent Variable: sınav başarı puanı

**Not:** Bu tablonun SPSS orijinal çıktısında Dimension sütunun altındaki satırlarda yalnızca 1,2,3,4,5,6,7 yazıyordu. Bu rakamların anlamı 1 Constant ve sırası ile Constant'ın yanında (sağında) yer alan bağımsız değişkenler, özgüven puanı 2, Mutluluk puanı 3 şeklinde yazar tarafından yazıldı (dimension sütunundaki değerlerin yanına).

**Durum Endeksi (Condition indices CI)** değerlerinin 30'dan küçük olması beklenir: IQ puanlarının CI değerleri 30'dan büyük! Bununla birlikte Durum Endeksi değerleri göz ardı edilebilecek ve günümüz literatüründe çoklu ortak doğrusallık için bir "altın standardı" özelliği taşımaz.

IQ puanı Sorunlu bir CI değerine sahip gibi görünse de IQ puanının diğer bağımsız değişkenlerle olan ilişki katsayıları (r değerleri 0,9'dan küçük olduğu için), Tolerans ve VIF değerlerinde de bir sorun olmadığı için modelimiz içinde tutmaya devam edebiliriz. Bunun nedeni: Durum Endeksinin (Condition indices CI) çoklu ortak doğrusallık sorunu için bir "altın standart" olmaması ve bu nedenle günümüz literatüründe referans bir kıstas olmamasıdır.

**Residuals Statistics<sup>a</sup>**

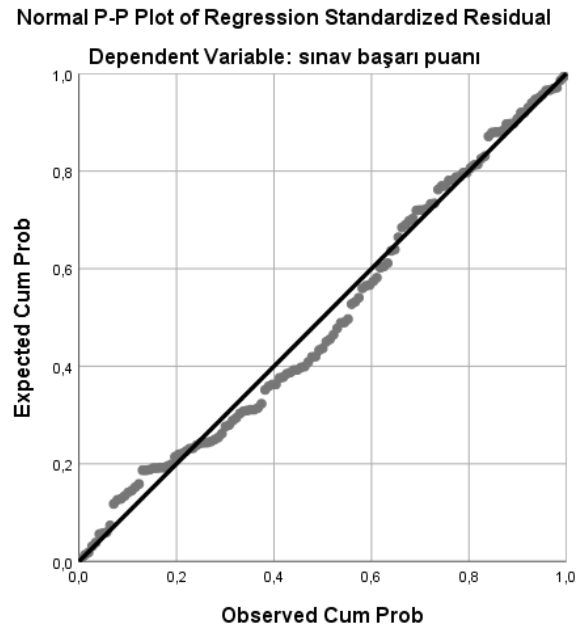
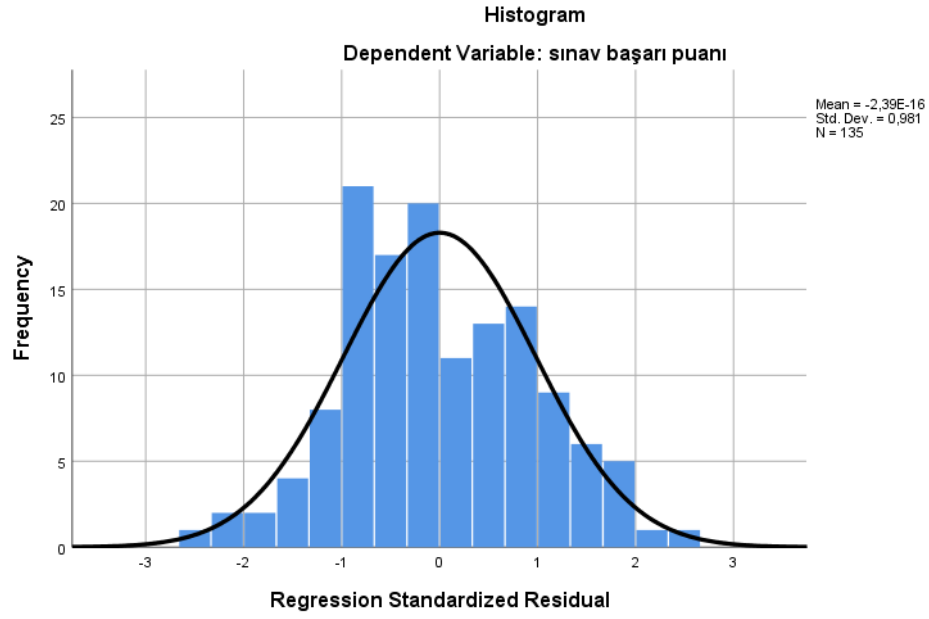
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	50,4546	103,0414	75,8963	13,19092	135
Residual	-8,84338	8,63479	,00000	3,35297	135
Std. Predicted Value	-1,929	2,058	,000	1,000	135
Std. Residual	-2,588	2,527	,000	,981	135

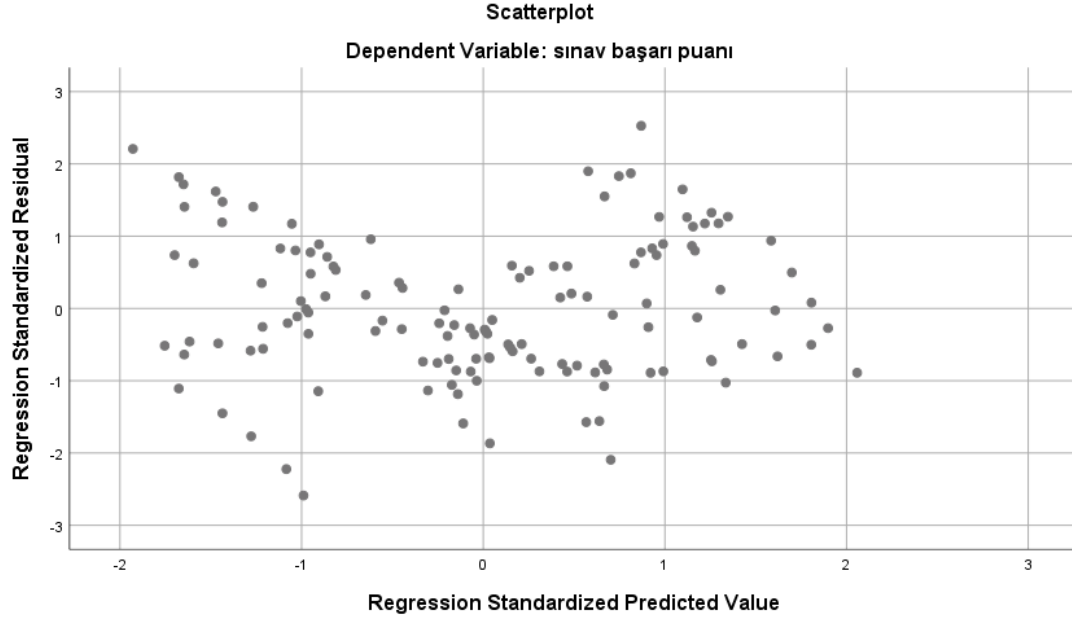
a. Dependent Variable: sınav başarı puanı

**Residuals Statistics: Fark/Artık İstatistikleri**

Regresyon modelinin geçerli olması artıkların/farkların normal dağılımı varsayımının yerine getirilmesine bağlıdır. Farkların/artıkların normalliği; histogram dağılımı, Normal P-P Grafiği ve saçılım grafiğine göre değerlendirilir. Bu analizde artıkların/farkların histogram dağılımına baktığımızda normal dağılıma yakın bir dağılım olduğunu; "Normal P-P" Grafiğinde noktaların 45<sup>0</sup> (45 derecelik) çizginin üstünde ya da yakınında olduğunu (normal dağıldığını ifade eder); saçılım grafiğinde de çok belirgin bir ihlalin olmadığını görüyoruz. Üç göstergeyi birlikte değerlendirdiğimizde farkların/artıkların normal dağılım varsayımının sağlandığını görmekteyiz.

## Charts



**Sonuç:**

Her bir değişkenin normal dağılım özelliği gösterdiği varsayıldı.

Bağımlı ve bağımsız değişkenler arasında doğrusal bir ilişki var.

Bağımsız değişkenler arasında çoklu ortak doğrusallık (Multicollinearity) sorunu yok!

Artıklar/farklar normal dağılım özelliği gösteriyor.

**Sonuç: Varsayımlar sağlandığı için Regresyon Analizi sonuçları geçerlidir.**

**Tablo: Sınav başarı puanının yordanmasına ilişkin Çoklu Regresyon Analizi**

Değişken	B	Standart Hata B	$\beta$	t	p	İkili r	Kısmi r
Sabit	16,011	7,493		2,137	,035		
Mutluluk puanı	,193	,032	,248	5,944	,000	,879	,464
Sınav kaygısı puanı	-,173	,045	-,133	-3,827	,000	-,792	-,319
Günlük harçlık TL	-,058	,039	-,042	-1,466	,145	-,604	-,128
Haftalık çalışma saati	,528	,047	,480	11,277	,000	,910	,705
IQ puanı	,333	,040	,228	8,407	,000	,677	,595
R= ,969      R <sup>2</sup> = ,937 F(5, 129)= 399,311      p= ,000							

Mutluluk puanı, Sınav kaygısı puanı, Günlük harçlık miktarı (TL), Haftalık çalışma saati, IQ puanı değişkenleri **birlikte**, Sınav başarı puanı ile anlamlı bir ilişki göstermiştir (R= ,969, R<sup>2</sup>= ,937, F(5, 129)= 399,311 p<0,01). Söz konusu beş değişken **birlikte**, Sınav başarı puanlarındaki değişimin %94'ünü açıklamaktadır. Standartlaştırılmış regresyon katsayılarına göre, yordayıcı (bağımsız) değişkenlerin Sınav Başarı Puanı üzerindeki önem sırası; Haftalık çalışma süresi ( $\beta$ =,480), Mutluluk puanı ( $\beta$ =,248), IQ puanı ( $\beta$ =,228), Sınav kaygısı puanı ( $\beta$ = -,133), Günlük harçlık TL ( $\beta$ = -,042) olduğu tespit edilmiştir.

Regresyon katsayılarının anlamlılığına ilişkin t testi sonuçları incelendiğinde haftalık çalışma süresi, Mutluluk puanı, IQ puanı, Sınav kaygısı puanı değişkenlerinin, Sınav başarı puanı üzerinde anlamlı (önemli) bir yordayıcı olduğu görülmektedir (p<0,01). Günlük harçlık TL değişkeninin ise Sınav Başarı Puanı üzerinde anlamlı (önemli) bir yordayıcı olmadığı görülmüştür (p>0,05).

Yordayıcı değişkenlerle Sınav başarı puanı arasında ilişkilere bakıldığında; mutluluk puanı ile r=,879 diğer yordayıcı değişkenlerin etkisi kontrol edildiğinde r=,464, Sınav kaygısı puanı ile r= -,792 diğer yordayıcı değişkenlerin etkisi kontrol edildiğinde r= -,319, Günlük harçlık TL ile r= -,604 diğer yordayıcı değişkenlerin etkisi kontrol edildiğinde r= -,128, Haftalık çalışma saati ile r=,910 diğer yordayıcı değişkenlerin etkisi kontrol edildiğinde r=,705, IQ puanı ile r=,677 diğer yordayıcı değişkenlerin etkisi kontrol edildiğinde r=,595 düzeylerinde korelasyon ilişkisi olduğu gözlenmiştir.

Regresyon analizi sonuçlarına göre Sınav başarı puanının yordanmasına ilişkin regresyon denklemi şu şekildedir:

$$\text{Sınav Başarı Puanı} = 16,011 + (0,193 \times \text{Mutluluk puanı}) - (0,173 \times \text{Sınav kaygısı puanı}) - (0,058 \times \text{Günlük harçlık TL}) + (0,528 \times \text{Haftalık çalışma saati}) + (0,333 \times \text{IQ puanı})$$



## Kaynakça

- Adedokun, Omolola A. ve Burgess, Wilella D. (2012) “Analysis of Paired Dichotomous Data: A Gentle Introduction to the McNemar Test in SPSS”, Journal of MultiDisciplinary Evaluation, Volume 8, Number 17, January 2012, pages:125-131
- Brooks Gordon P. ve Barcikowski Robert S. (2012) “The PEAR Method for Sample Sizes in Multiple Linear Regression”, Multiple Linear Regression Viewpoints, 2012, Vol. 38(2)
- Bryman, Alan ve Cramer, Duncan. (1999) Quantitative Data Analysis with SPSS Release 8 for Windows, (First published), Routledge, London
- Büyüköztürk, Şener. (2010) Sosyal Bilimler İçin Veri Analizi El Kitabı, 11. Baskı, Pegem Akademi, Ankara
- Can, Abdullah. (2019) SPSS ile Bilimsel Araştırma Sürecinde Nicel Veri Analizi, 7. Baskı, Pegem Akademi, Ankara
- Carver, Robert H. ve Nash Gradwohl Jane. (2012) Doing Data Analysis with SPSS Version 18, Cole Cengage Learning, Boston
- Chang, Horng-Jinh., Wu, Chao-Hsien., Ho, Jow-Fei., Chen, Po-yu. (2008) “On Sample Size in Using Central Limit Theorem for Gamma Distribution”, Information and Management Sciences, Volume 19, Number 1, pages: 153-174
- Coakes, Sheridan J. (2005) SPSS Analysis Without Anguish: Version 12.0 for Windows, (First Published), Wiley, New Jersey
- Cohen, Jacob. (1988) Statistical Power Analysis for the Behavioral Sciences, (Second Edition), LEA Publishers, New Jersey
- Cohen, Jacob. (1992) “A power primer”, Psychological Bulletin, 1992, Vol. 112. No. 1, pages: 155-159
- Cohen, Louis. ve Holliday, Michael. (1982) Statistics for Social Scientists, Harper & Row Publishers, London
- Connolly, Paul (2007) Quantitative Data Analysis in Education: A critical introduction using SPSS, Routledge, London
- Corder, Gregory W. ve Foreman, Dale I. (2009) Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach, A John Wiley & Sons, Inc., Publication, New Jersey
- Cowles, Michael. ve Davis, Caroline. (1982), “On the Origins of the .05 Level of Statistical Significance” American Psychologist, May 1982, Vol.37, No. 5, pages:553-558

- DeCoster, Jamie ve Claypool, Heather M. (2004) Data Analysis in SPSS, (e-book), <http://www.stat-help.com/spss.pdf>
- Field, Andy. (2009) Discovering Statistics Using SPSS, (Third Edition), Sage Publications Ltd., London
- Gignac, Gilles E. ve Szodorai, Eva T. (2016) “Effect size guidelines for individual differences researchers”, *Personality and Individual Differences*, 102 (2016), pages:74-78
- Gökçe, Birsen. (2018) Toplumsal Bilimlerde Araştırma, 7. Baskı, Savaş Yayınları, Ankara
- Green, Samuel B. (1991) “How Many Subjects Does It Take To Do A Regression Analysis?”, *Multivariate Behavioral Research*, 26 (3), pages: 449-510
- Green, Samuel B. ve Salkind, Neil J. (2014) Using SPSS for Windows and Macintosh: Analyzing and Understanding Data, (Seventh Edition), Pearson, New Jersey
- IBM SPSS (2017) IBM SPSS Statistics Base 25, New York
- Işık, İdil (2014) “Yokluk Hipotezi Anlamlılık Testi ve Etki Büyüklüğü Tartışmalarının Psikoloji Araştırmalarına Yansımaları”, *Eleştirel Psikoloji Bülteni*, Sayı:5, Nisan 2014, sayfa:55-80
- Jolliffe, Ian T. (1995) “Sample Sizes and the Central Limit Theorem: The Poisson Distribution as an Illustration” *The American Statistician*, Volume 49, 1995, Issues 3, page: 269
- Kanji, Gopal K. (2006) 100 Statistical Tests, (3rd Edition), SAGE Publications Ltd, London
- Kerr, Alistair W., Hall, Howard K. ve Kozub Stephen A. (2002) Doing Statistics with SPSS, (First Published), Sage, London
- Knofczynski, Gregory T. ve Mundfrom, Daniel. (2008) “Sample Sizes When Using Multiple Linear Regression for Prediction”, *Educational and Psychological Measurement*, 2008 (68), pages: 431-442
- Kul, Seval. (2014) “İstatistik Sonuçlarının Yorumu: P Değeri ve Güven Aralığı Nedir?” *Plevra Bülteni*, Ocak 2014, Cilt:8 Sayı:1, sayfa:11-13
- Landau, Sabine ve Everitt, Brian S. (2004) A Handbook of Statistical Analyses using SPSS, Chapman & Hall/CRC Press LLC, New York
- Malhotra, Ruchika. (2018) “An extensive analysis of search-based techniques for predicting defective classes” *Computers & Electrical Engineering*, Volume 71, October 2018, Pages: 611-626
- Mangiafico, Salvatore S. (2016) Summary and Analysis of Extension Program Evaluation in R, version 1.18.1., Rutgers Cooperative Extension, New Brunswick, NJ.

- Marshall, Ellen ve Boggis, Elizabeth. (2016) The Statistics Tutor's Quick Guide to Commonly Used Statistical Tests, Statstutor Community Project, Reviewer: Jean Russell University of Sheffield, Sheffield
- McCrum-Gardner, Evie. (2008) "Which is the correct statistical test to use?", British Journal of Oral and Maxillofacial Surgery, 46 (2008) pages:38–41
- Öztuna, Derya., Elhan, Atilla Halil. ve Tüccar, Ersöz. (2006)" Investigation of Four Different Normality Tests in Terms of Type 1 Error Rate and Power under Different Distributions", Turk J Med Sci, 2006; 36 (3), pages:171-176
- Öztürk, Fikri ve Özbek, Levent. (2004) Matematiksel Modelleme ve Simülasyon, Gazi Kitapevi, Ankara
- Pallant, Jullie. (2005) SPSS Survival manual: A step by step guide to data analysis using SPSS, (2nd edn), Allen & Unwin, Sydney
- Pallant, Jullie. (2017) SPSS Kullanma Kılavuzu, Çev: Sibel Balcı ve Berat Ahi, 2. Baskı, Anı Yayıncılık, Ankara
- Razali, Nornadiah Mohd. ve Wah, Yap Bee. (2011) "Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests", Journal of Statistical Modeling and Analytics, 2011, Vol.2 No.1, pages:21-33
- Rosenthal, Robert ve Rosnow, Ralph L. (2008) Essentials of Behavioral Research: Methods and Data Analysis, (Third Edition), Published by McGraw-Hill, New York
- Rovai, A. P, Baker, J. D., & Ponton, M. K. (2014) Social Sci. Research Design and Statistics: A Practitioner's Guide to Research Methods and IBM SPSS Analysis. Chesapeake, VA: Watertree Press LLC.
- Rufibach, Kaspar. (2011) "Assessment of paired binary data", Skeletal Radiol, 40:1–4
- Singh, N. Uttam., Roy, Aniruddha ve Tripathi, A. K. (2013) Non Parametric Tests: Hands on SPSS, ICAR Research Complex for NEH Region, Umiam, Meghalaya
- Stevens, James P. (2009) Applied Multivariate Statistics For The Social Sciences, (Fifth Edition), Routledge, New York
- Sun, Xuezheng ve Yang, Zhao (2008) "Generalized McNemar's Test for Homogeneity of the Marginal Distributions", SAS Global Forum 2008, Paper:382-2008
- Şenesen, Ümit. (2013) İstatistik Sayıların Arkasını Anlamak, (Dördüncü Basım), Literatür Yayıncılık, İstanbul
- Tabachnick, Barbara G. ve Fidell, Linda S. (2007) Using Multivariate Statistics, (Fifth Edition), Pearson Education. Inc., Boston
- Terzi, Yüksel. (2019) SPSS ile İstatistiksel Veri Analizi (Ders Notları), Ondokuz Mayıs Üniversitesi, Samsun

- Tomczak, Maciej ve Tomczak, Ewa. (2014) “The need to report effect size estimates revisited. An overview of some recommended measures of effect size”, *TRENDS in Sport Sciences*, 2014; 1(21), Pages: 19-25.
- Trajman, A. ve Luiz, R.R. (2008) “McNemar  $\chi^2$  test revisited: comparing sensitivity and specificity of diagnostic examinations”, *Scandinavian Journal of Clinical and Laboratory Investigation*, Volume 68, Issue 1, pages:77-80
- Tranmer, M., Murphy, J., Elliot, M., ve Pampaka, M. (2020) *Multiple Linear Regression* (2nd Edition), Cathie Marsh Institute Working Paper 2020-01.
- Warner, Rebecca M. (2008). *Applied Statistics: From Bivariate Through Multivariate Techniques*, (First Edition), Sage, London
- Whitley, Elise ve Ball, Jonathan. (2002) “Statistics review 3: Hypothesis testing and P values” *Critical Care*, June, Vol:6 No:3, pages: 222-225
- Wilcox, Rand R. (2005) *Introduction to Robust Estimation and Hypothesis Testing*, (Second Edition.), Elsevier Academic Press, Burlington
- Yalçın, Seher. (2016) *İstatistik (Ders Notları)*, Ankara Üniversitesi, Ankara <https://acikders.ankara.edu.tr/course/view.php?id=108#section-0> (Erişim:12 Mayıs 2020)
- Yockey, Ronald D. (2016) *SPSS Demystified: A Step-by-Step Guide to Successful Data Analysis For SPSS Version 18.0*, (Second Edition), Routledge, New York

## Ekler

### EK 1: TEST KILAVUZU

I- Normal dağılım Testleri ve Normal dağılım Kriterleri:		
Örneklem büyüklüğü=	n≥30	n<30
Nicel bir değişken verilerinin normal dağılıma uygunluğu	Kolmogorov Smirnov testi ile sınanabilir.	Shapiro Wilk testi ile sınanabilir.
Test sonucu:	p>0,05 ise nicel veriler normal dağılım özelliği gösteriyor. p<0,05 ise nicel veriler normal dağılım özelliği göstermiyor.	
Normal dağılım kriterleri:	Ortalama, ortanca, mod değerleri birbirine yakın olmalı. Değerlerin histogram dağılımı çan eğrisi şeklinde olmalı. Çarpıklık (Skewness) ve Basıklık (Kurtosis) katsayıları -1+1 aralığında olmalı. Normal Q-Q Plot grafiğinde noktalar 45 derecelik doğru üstünde ya da çok yakınında olmalı.	
Nicel veriler normal dağılım özelliği gösteriyor ise Parametrik testler, normal dağılım özelliği göstermiyor ise Non-Parametrik testler tercih edilir.		

<b>II- Gruplar arasında ya da Grup içi Ortalama Karşılaştırmaları:</b>		
Koşullar	Parametrik test	Non-Parametrik test
Nicel bir değişkenin Örneklem ortalaması ile Evren ortalamasının karşılaştırılması	<b>Tek Örneklem t Testi (One-Sample T Test)</b>	Yok
İki bağımsız grubun bir nicel değişkeninin–ortalamasının- karşılaştırılması	<b>Bağımsız Gruplarda t testi (Independent Samples T-Test)</b>	<b>Mann Whitney U testi</b>
Bir grubun tekrarlayan iki ölçüm –ortalamasının- karşılaştırılması. Örnek: ön test, son test.	<b>Bağımlı Gruplarda t Testi (Paired Samples T-Test)</b>	<b>Wilcoxon İşaretli Sıralar Testi</b>
Üç ya da daha fazla bağımsız grubun bir nicel değişkeninin –ortalamasının- karşılaştırılması	<b>Bağımsız Gruplar için Tek Yönlü Varyans Analizi (One-Way ANOVA)</b> Fark çıkarsa farkın kaynağını tespit etmek için post-hoc test yapılır.	<b>Kruskal Wallis Testi</b>  Fark çıkarsa farkın kaynağını tespit etmek için <b>post-hoc test</b> olarak <b>Mann Whitney U testi</b> yapılır.
Bir grubun 3 ya da daha fazla tekrarlayan ölçüm –ortalamasının- karşılaştırılması	<b>Tekrarlayan Ölçümler için Tek Yönlü Varyans Analizi (One Way ANOVA for Repeated Measures)</b> Fark çıkarsa farkın kaynağını tespit etmek için post-hoc test yapılır.	<b>Friedman Testi</b>  Fark çıkarsa farkın kaynağını tespit etmek için <b>post-hoc test</b> olarak <b>Wilcoxon İşaretli Sıralar Testi</b> yapılır.

<b>III- Gruplar arası ya da Grup içi Oran Karşılaştırmaları:</b>		
<b>Koşullar</b>	<b>Parametrik test</b>	<b>Non-Parametrik test</b>
Kategorik bir değişkenin Evrende (ya da bir başka çalışmada) görülme oranları ile Örneklemde görülme oranları arasında uyum ya da farklılığın sınanmasında	Yok	<b>Tek Örneklem Ki Kare Testi</b>
Kategorik bir değişkenin kategorilerinin Örneklem içindeki oranlarının eşit dağılıp dağılmadığı test edilmek istendiğinde	Yok	<b>Tek Örneklem Ki Kare Testi</b>
İki kategorik değişken arasında anlamlı bir ilişki olup olmadığını sınamak için uygulanır. Kategorik bir değişkenin diğer kategorik değişken üzerinde etkisinin sınıandığı durumlarda uygulanır.	Yok	<b>Ki Kare Bağımsızlık Testi</b>
Bir grubun tekrarlayan iki kategorik ölçümü arasında bir fark olup olmadığını sınamak için uygulanır. Ön test son test dizaynı içinde tekrarlayan kategorik ölçüm söz konusudur. Söz konusu kategorik değişkenin iki alt kategorisi vardır (0=yok, 1 =var ya da 0=hayır, 1=evet gibi).		<b>McNemar Testi</b>
Bir grubun tekrarlayan üç kategorik ölçümü arasında bir fark olup olmadığını sınamak için uygulanır. Aynı grubun tekrarlayan 3 ya da fazla ölçümü söz konusudur. Söz konusu kategorik değişkenin iki alt kategorisi vardır (0=yok, 1 =var ya da 0=hayır, 1=evet gibi).		<b>Cochran Q Testi</b>  Fark çıkar ise farkın kaynağını tespit etmek için <b>post-hoc test</b> olarak <b>McNemar testi</b> yapılır.

<b>IV- Nicel Değişkenler Arası İlişki:</b>		
<b>Koşullar</b>	<b>Parametrik test</b>	<b>Non-Parametrik test</b>
İki nicel değişken arasında doğrusal bir ilişkinin olup olmadığını sınıandığında	Basit Korelasyon <b>Pearson Korelasyon Testi</b>	Basit Korelasyon <b>Spearman Korelasyon testi</b>
İki nicel değişken arasında bir ilişkinin olup olmadığını sınarken, bu iki değişken üzerinde etkisi olabileceğini düşündüğümüz (şüphelendiğimiz) üçüncü bir nicel değişkenin etkisi arındırarak ilişki analizi yapılmak istendiğinde	<b>Kısmi Korelasyon</b>	Yok

<b>V- Nicel Değişkenler Arasında Açıklayıcılık Analizi:</b>		
<b>Koşullar</b>	<b>Parametrik test</b>	<b>Non-Parametrik test</b>
İki nicel değişkenden birinin bağımsız (X), diğerinin bağımlı değişken (Y) olarak kurgulandığı durumda; bağımsız değişkenin bağımlı değişkenin ne kadarını açıkladığı ve aralarında nasıl bir matematiksel eşitlik olduğu sınanmak istendiğinde <b><math>Y=A+ BX</math></b>	<b>Basit Doğrusal Regresyon Analizi</b>	Yok
İkiden çok nicel değişkenden birinin bağımlı değişken(Y), diğerlerinin bağımsız değişkenler ( $X_1, X_2, X_3, \dots$ ), olarak kurgulandığı durumda; bağımsız değişkenlerin bağımlı değişkenin ne kadarını açıkladığı ve aralarında nasıl bir matematiksel eşitlik olduğu sınanmak istendiğinde <b><math>Y=A+ B_1 X_1 + B_2 X_2 + B_3 X_3 + \dots</math></b>	<b>Çoklu Doğrusal Regresyon Analizi</b>	Yok

## EK 2: Çoklu Doğrusal Regresyon için Artık/Fark Değerlerinin Saçılım Grafikleri İle Yorumlanması

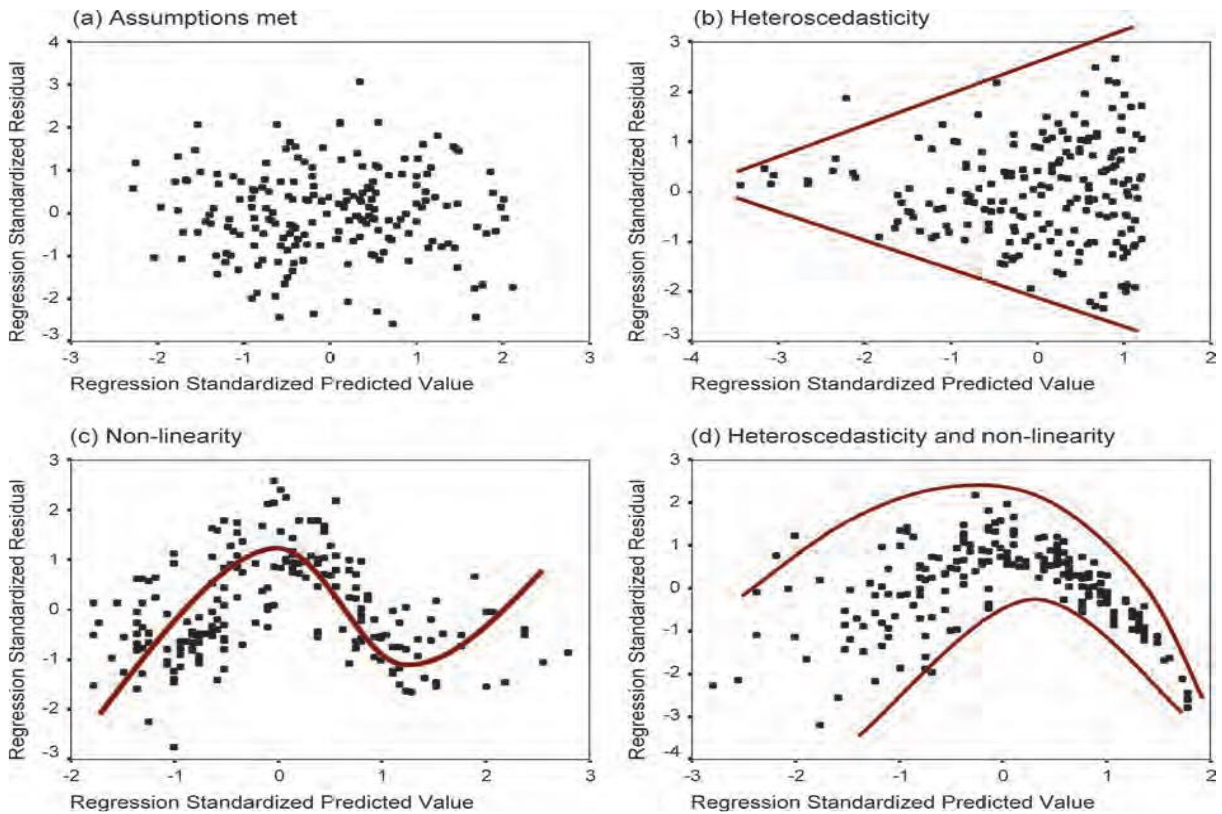
Standart tahmin edilen değerlere karşı standardize edilmiş artıkların/farkların Saçılım Grafiğinin birkaç örneği gösterilmiştir. Bu grafiklerde noktaların yatay sıfır çizgisinin etrafında ve rastgele dağılmasını bekleriz. Bu grafiklerde noktalar huni şeklinde belirginleşirse bu heteroscedasticity işaretidir ve homoscedasticity varsayımının ihlal edildiğini gösterir. Bu grafiklerde noktalar eğrilik oluşturacak şekilde belirginleşiyorsa bu da –farkların- doğrusallık varsayımının ihlaline işaret eder (Field, 2009):

Panel (a): Noktaların çizim boyunca rastgele ve eşit olarak nasıl dağıldığına dikkat edin. Bu örüntü, doğrusallık ve homoscedasticity varsayımlarının karşılandığı bir durumu gösterir.

Panel (b), homoscedasticity varsayımını ihlal eden bir veri kümesi için tipik bir grafik göstermektedir. Noktaların bir huni şeklini oluşturduğunu ve bu şekilde grafik de yayıldığına dikkat edin. Bu huni şekli, heteroscedasticity için tipiktir ve artıklar/farklar arasında artan varyansı gösterir.

Panel (c), bağımlı ile bağımsız değişkenler arasında doğrusal olmayan bir ilişkinin olduğu verilerin grafiğini göstermektedir. Bu örüntü (pattern) artıklar/farklar tarafından gösterilmiştir. Verilerdeki eğilimi göstermek için eğrisel ilişkiyi gösteren bir çizgi grafiğin üstünden çizilmiştir.

Panel (d) verilerin sadece doğrusal olmayan bir ilişkiyi temsil etmediği, aynı zamanda heteroscedasticity de gösterdiği bir durumu temsil etmektedir. Önce verilerdeki eğri eğilime dikkat edin ve daha sonra grafiğin bir ucunda noktaların birbirine çok yakın olduğunu, diğer ucunda ise yaygın olarak dağıldığını not edin.



Kaynak: Field, 2009



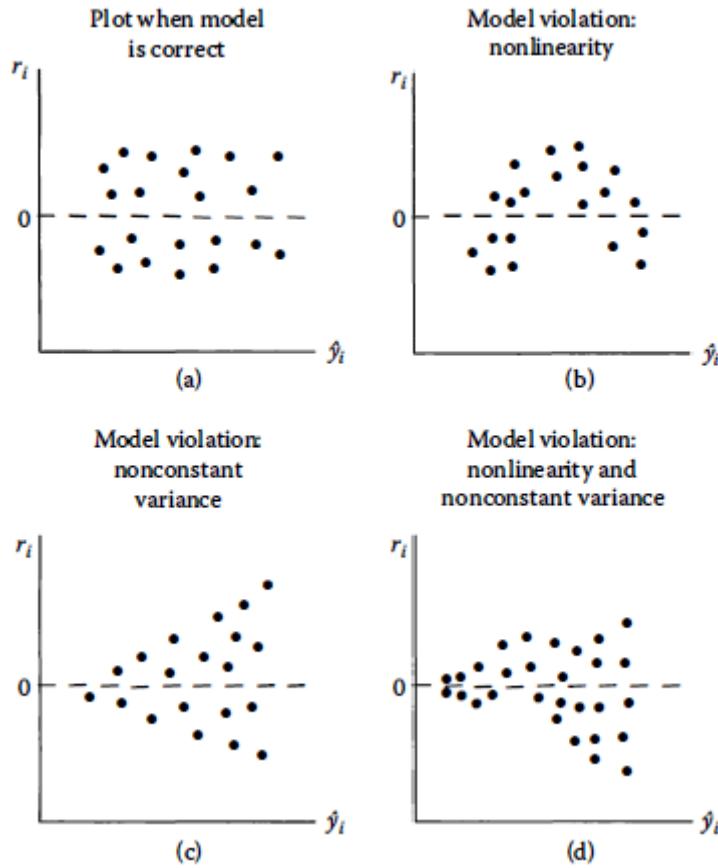
Standart farkların/artıkların ( $r_i$ ), tahmin edilen değerlere ( $y_i$ ) karşı en kullanışlı grafiklerinden biri Saçılım Grafiğidir. Doğrusal regresyon modelinin varsayımları yerine getirilmiş ise, standartlaştırılmış artıklar, panel (a)'da gösterildiği gibi  $r_i = 0$  ile tanımlanan yatay bir çizgi etrafında rastgele dağılmalıdır (Stevens, 2009):

Panel (a): Noktalar çizgi (yatay sıfır çizgisi) boyunca rastgele ve eşit olarak dağılmıştır. Bu örüntü, doğrusallık ve homoscedasticity varsayımlarının karşılandığı bir durumu gösterir.

Panel (b), bağımlı ile bağımsız değişkenler arasında doğrusal olmayan bir ilişkinin olduğu verilerin grafiğini göstermektedir. Bu örüntü (pattern) artıklar/farklar tarafından gösterilmiştir. Noktalar bir eğriselliği işaret eder.

Panel (c), homoscedasticity varsayımını ihlal eden bir grafik örneğidir. Noktalar bir huni şeklini oluşturmakta ve bu şekilde grafik de yayılmaktadır. Bu huni şekli, heteroscedasticity için tipiktir ve artıklar/farklar arasında artan varyansı gösterir.

Panel (d) verilerin sadece doğrusal olmayan bir ilişkiyi temsil etmediği, aynı zamanda heteroscedasticity de gösterdiği bir durumu temsil etmektedir. Önce verilerdeki eğri eğilime dikkat edin ve daha sonra grafiğin bir ucunda noktaların birbirine çok yakın olduğunu, diğer ucunda ise yaygın olarak dağıldığını not edin.



Kaynak: Stevens, 2009