

HİPOTEZ TESTLERİ ve ANLAM ÇIKARTICI İSTATİSTİK

4. SUNUM

- Gözlem ya da deneme sonucu elde edilmiş sonuçların, rastlantıya bağlı olup olmadığının incelenmesinde kullanılan istatistiksel yöntemlere **HİPOTEZ** **TESTLERİ** denir.
- Sonuçların rastlantıya bağlı olup olmadığı, evren parametreleri (ortalama, ortanca, varyans, vb.) üzerine kurulmuş hipotezlerin test edilmesi ile yapılır.

Örnek 1: A ve B diyeti arasında fark olup olmadığını araştırmak isteyen bir araştırmacı rasgele 50 kişi seçiyor ve seçtiği 50 kişiyi yine rasgele 2 diyet grubuna atıyor. Diyetisyen, her iki gruptaki kişilerin diyet uygulamadan önce ve sonraki BKİ'leri arasındaki farkları ölçüyor ve aşağıdaki gibi bir tablo elde ediyor.

Diyet	Denek Sayısı	BKİ farkı Ortalaması (kg/m ²)	BKİ farkı Standart Sapması (kg/m ²)
A	25	1.2	0.1
B	25	1.5	0.2

Acaba A ve B diyeti arasında kilo düşürme açısından fark var mıdır?

- **Örnek 2:** Kan ve kan ürünleri ile çalışan 100 hastane personelinin yapılan test sonucu 23'ünde hepatit B pozitif bulunmuştur. Bu bilgilerle kan ve kan ürünleri ile çalışan hastane personelinde hepatit B pozitif olanların oranının %15' den büyük olduğu söylenebilir mi?

- **Örnek 3:** Çalışma pozisyonunun varis oluşumu ile ilişkisini incelemek üzere yapılan bir çalışma sonucu aşağıdaki gibidir.

Çalışma Pozisyonu	Varis Oluşumu				Toplam
	VAR		YOK		
	Sayı	Yüzde	Sayı	Yüzde	
Ayakta	25	0.25	75	0.75	100
Oturarak	10	0.13	70	0.88	80
Toplam	35	0.19	145	0.81	180

Bu bilgilerle ayakta çalışanlarda varis oluşumu daha fazladır denebilir mi?

- **Örnek 4:** Farklı üç ilaç (A,B,C) kullanan üç grupta kan pıhtılaşma zamanları farklı mıdır?

İlaç	Denek Sayısı	Ortalama (sn)	Standart Sapma
A ilacı	20	40	12
B ilacı	30	56	20
C ilacı	25	50	15

- Verilen örneklerin tümünde incelenmek istenen, evren ortalaması(ları) ya da evren oran(ları) üzerine kurulmuş hipotezlerdir.
- Hipotez testlerinde iki hipotez vardır. **Birincisi, H_0 ile gösterilen (NULL) yokluk hipotezi, ikincisi H_1 ile gösterilen (ALTERNATİF) seçenek hipotezdir.**
- İstatistiksel hipotez testlerinin tümü H_0 hipotezinin doğru olduğu varsayımı altında gerçekleştirilir.

Örnek 1 (devam): A ve B diyeti arasında fark olup olmadığını araştırmak isteyen bir araştırmacı rasgele 50 kişi seçiyor ve seçtiği 50 kişiyi yine rasgele 2 diyet grubuna atıyor. Diyetisyen, her iki gruptaki kişilerin diyet uygulamadan önce ve sonraki BKİ'leri arasındaki farkları ölçüyor ve aşağıdaki gibi bir tablo elde ediyor.

Diyet	Denek Sayısı	BKİ farkı Ortalaması (kg/m ²)	BKİ farkı Standart Sapması (kg/m ²)
A	25	1.2	0.1
B	25	1.5	0.2

Araştırmanın Hipotezi:

H_0 : A ve B diyetleri arasında fark yoktur.

H_1 : B diyeti A'ya göre daha başarılıdır.

Veya;

H_1 : A ve B diyetleri arasında fark vardır.

- İstatistiksel hipotez testlerinde iki tür yanlış vardır.

Test Sonucu	Gerçek Durum	
	H_0 Doğru	H_0 Yanlış
H_0 Kabul	Doğru Karar	II. Tip Hata (β)
H_0 Red	I. Tip Hata (α)	Doğru Karar

α : Anlamlılık (manidarlık) Düzeyi $1-\alpha$ = Güven Düzeyi $1- \beta$: Testin Gücü

Örnek 1 için;

Test Sonucu	Gerçek Durum	
	A ve B diyetleri arasında fark yok	A ve B diyetleri arasında fark var
A ve B diyetleri arasında fark yok (H_0 Kabul)	Doğru Karar	II. Tip Hata (β)
A ve B diyetleri arasında fark var (H_0 Red)	I. Tip Hata (α)	Doğru Karar

- İstatistiksel hipotez testlerinin tümü H_0 hipotezinin doğru olduğu varsayımı altında gerçekleştirilir.
- Araştırmacı, çalışmasına başlamadan önce tip I hata olasılığı için belirli bir değer öngörür. Bu değer alfa (α) değeri ile gösterilir ve genellikle 0.05 veya 0.01 gibi küçük değerler olarak alınır.

Örnek 1 için;

Test Sonucu	Gerçek Durum	
	A ve B diyetleri arasında fark yok	A ve B diyetleri arasında fark var
A ve B diyetleri arasında fark yok (H_0 Kabul)	Doğru Karar	II. Tip Hata (β)
A ve B diyetleri arasında fark var (H_0 Red)	I. Tip Hata (α)	Doğru Karar

- Diyelim ki, çalışmamızın başında tip I hata olasılığını $\alpha=0.05$ olarak öngördük. Bunun anlamı H_0 gerçekte doğru iken onu yanlışlıkla red etme olasılığımız maksimum %5 olmalı.
- İstatistiksel paket programları, bir hipotez testi sonucunda gerçekleşen I. tip hata miktarını hesaplar ve bu değere **p değeri** denir. p değeri önceden belirlenmiş α değeri ile karşılaştırılarak karar verilir.

Eğer:

- $P \leq \alpha$ ise H_0 red edilir. Bunun anlamı, H_0 'ı red etmekle gerçekleşen yanılğı **öngörülenden küçüktür**. Dolayısıyla rahatlıkla H_0 red edilebilir.
- $P > \alpha$ ise H_0 kabul edilir. Bunun anlamı gerçekleşen yanılğı öngörülenden küçük olmadığı için H_0 red edilemez.

- Varsayalım ki, Örnek 1 için uygun hipotez testini kullandık ve p değerini 0.26 olarak elde ettik. Bu durumda aşağıdaki şekilde kurulan

H_0 : A ve B diyetleri arasında fark yoktur.

H_1 : B diyeti A'ya göre daha başarılıdır.

$P > \alpha$ için H_0 kabul edilir. Bunun anlamı A ve B diyeti arasında fark yoktur.

Hipotez Testi Aşamaları:

I. Aşama: H_0 ve H_1 Hipotezlerinin Belirlenmesi ve Formüle edilmesi

II. Aşama: Test Ölçütlerinin Belirlenmesi

III. Aşama: Test İstatistiğinin Değerinin hesaplanması

IV. Aşama: İstatistiksel Kararın Verilmesi ve Yorumlanması:

Hipotez Testi Aşamaları:

I. Aşama: H_0 ve H_1 Hipotezlerinin Belirlenmesi ve Formüle edilmesi:

- **Örnek 5:** Kolesterol ortalaması 190, standart sapması 45 olan 100 kişilik bir örneklem, kolesterol yönünden normal kabul edilebilir mi?
- H_0 hipotezi, evren parametreleri cinsinden ifade edilir.

- Bu örnekte öncelikle kolesterolü normal evrenin parametrelerinin bilinmesi ya da belirlenmesi gerekir.
- Kolesterolü normal evrenin ortalaması 180 standart sapması 58 ise Örneklemin çekildiği evrenin ortalamasının 180 olup olmadığını incelemek gerekir. Bu durumda yokluk hipotezimiz;

$$H_0 : \mu = 180$$

biçiminde formüle edilir.

H_1 Seçenek (ALTERNATİF) Hipotezinin Belirlenmesi ve Formüle edilmesi:

- H_0 hipotezi, örneklemin kolesterolü normal bir evrenden çekildiği olduğuna göre H_1 seçenek hipotezi H_0 'a karşıt olarak örneklemin kolesterolü normal olmayan bir evrenden çekildiği biçiminde olacaktır.
- Bu durumda kolesterolü normal olmayan evrenin tanımlanmasına gerek vardır.

Örneklemin çekildiği evrenin ortalamasının 180'den farklıdır:	Örneklemin çekildiği evren ortalaması 180'den büyüktür:	Örneklemin çekildiği evren ortalaması 180'den küçüktür:
$H_0 : \mu = 180$ $H_0 : \mu \neq 180$	$H_0 : \mu = 180$ $H_0 : \mu > 180$	$H_0 : \mu = 180$ $H_0 : \mu < 180$

- Araştırmacı amacına ya da tanımlamalarına uygun olarak yokluk hipotezine karşıt olarak üç farklı seçenek hipotez kullanabilir.

Çift Yönlü

$$H_0: \mu = 180$$

$$H_1: \mu \neq 180$$

Tek Yönlü

$$H_0: \mu = 180$$

$$H_1: \mu > 180$$

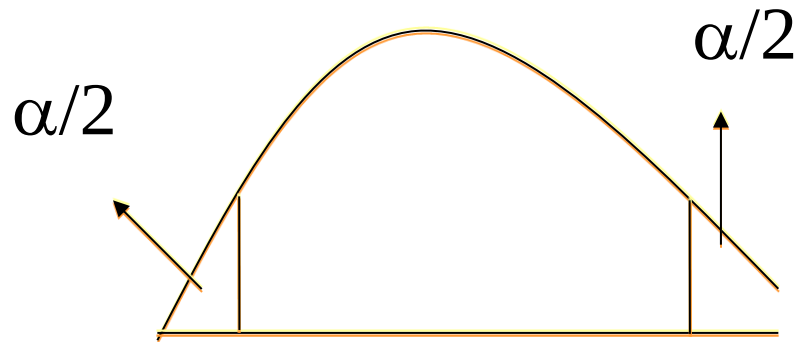
$$H_0: \mu = 180$$

$$H_1: \mu < 180$$

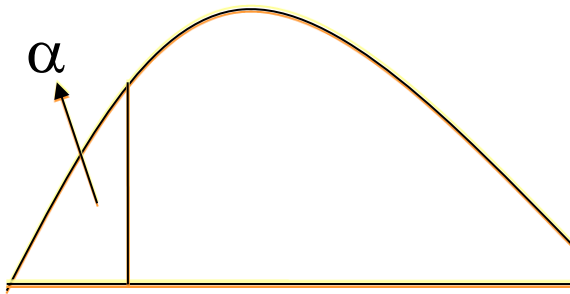
- H_1 seçenек hipotezinin iki ya da çok yönlü olması test sonucu karar verilme koşullarında farklılık yaratır öyle ki; H_1 seçenек hipotezinin iki yönlü olması 1. Tip hata α 'nın ikiye bölünmesini gerektirir. Bunun nedeni H_1 hipotezinin iki yönlü seçilmesi yanılığının her iki yönde öngörülmesi demek olacağından toplam 1. Tip hata olasılığı olarak tanımlanan α 'nın her iki yönde $\alpha/2$ olarak tanımlanmasını gerektirir.

$$H_0: \mu = 180$$

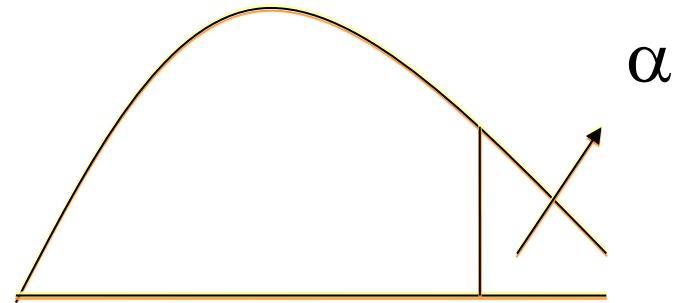
$$H_1: \mu \neq 180$$



- H_1 hipotezi tek yönlü iken gerçekleşen I. Tip hata p , α ile karşılaştırılırken H_1 hipotezi iki yönlü iken gerçekleşen I. Tip hata **p** ; **$\alpha/2$** ile karşılaştırılır.



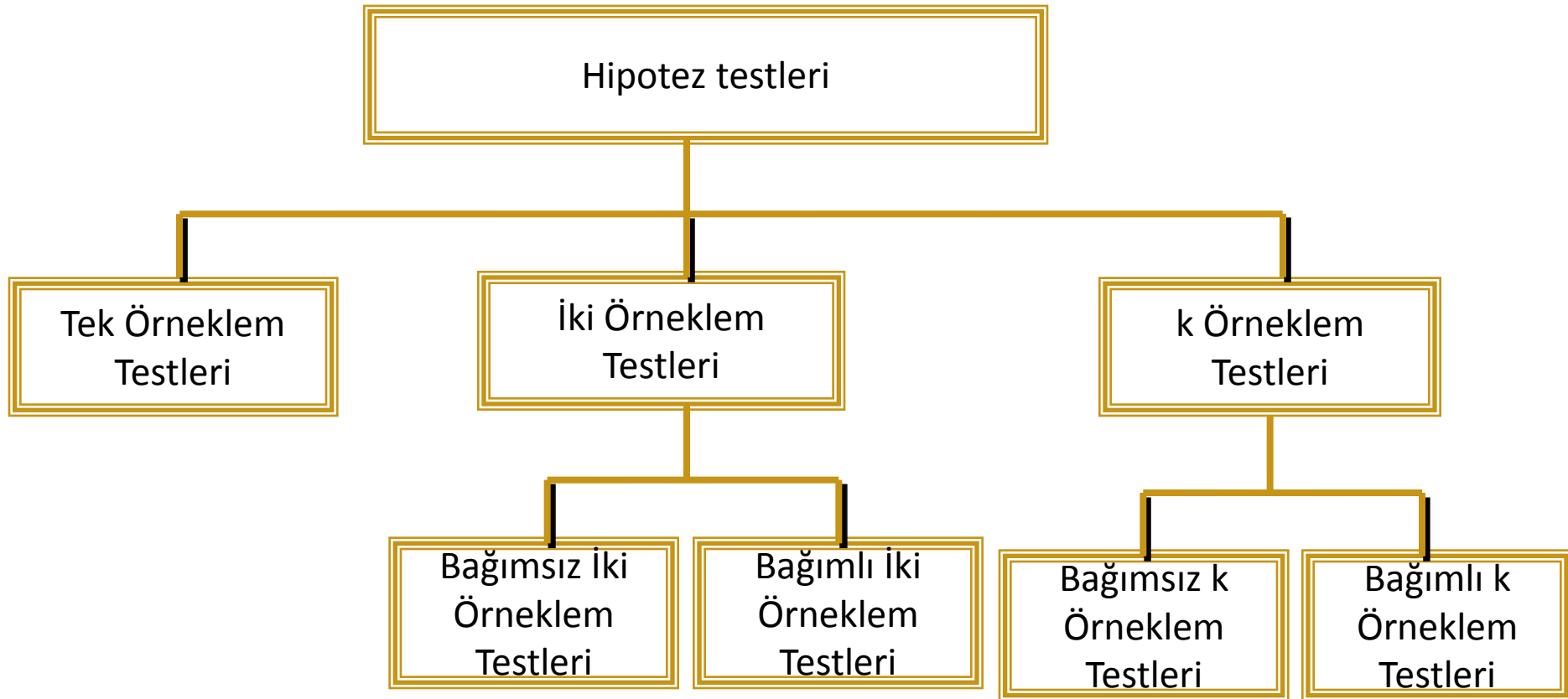
$$H_0: \mu = 180$$
$$H_1: \mu < 180$$



$$H_0: \mu = 180$$
$$H_1: \mu > 180$$

II. Aşama (a): Hipotezler için uygun test veya test istatistiğinin belirlenmesi:

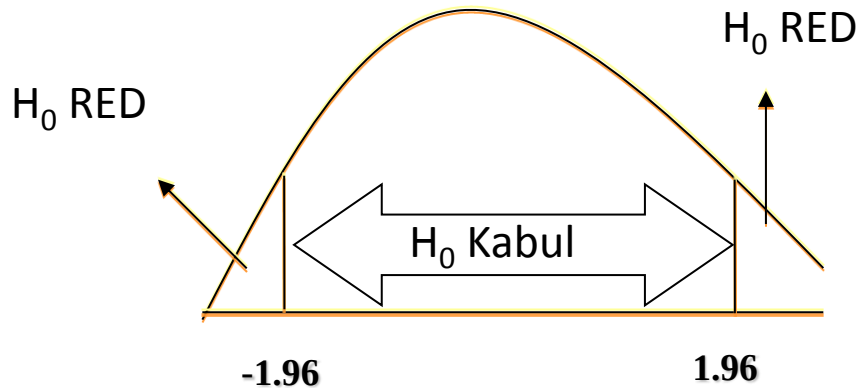
- Farklı hipotez testleri için değişik test istatistiklerinden yararlanılır.
- Örneğin iki örneklem ortalamasını karşılaştırmak için t test istatistiğinden yararlanırken, ikiden fazla örneklem ortalamasının birbirinden farklı olup olmadığını karşılaştırmada F test istatistiği kullanılmaktadır. Uygun testi dolayısıyla test istatistiğini seçmek hipotez testlerinin en önemli adımıdır. **Bu ders kapsamında test istatistiklerinin nasıl hesaplandığı hakkında bilgi ileride verilecektir.**



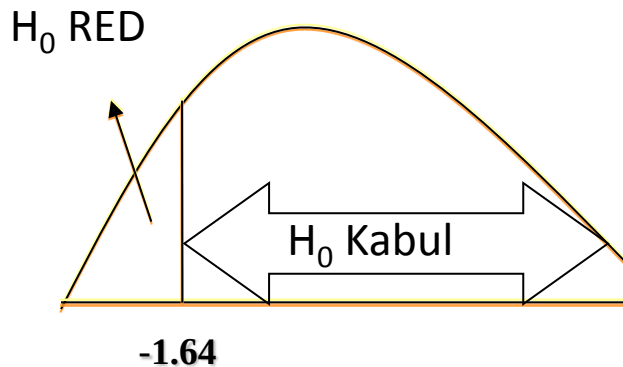
II. Aşama (b): İstatistiksel test için I. Tip hatanın olasılığı olan α 'nın belirlenmesi:

- Çalışmalarda genellikle $\alpha=0.05$, 0.01 gibi küçük değerler alınır.

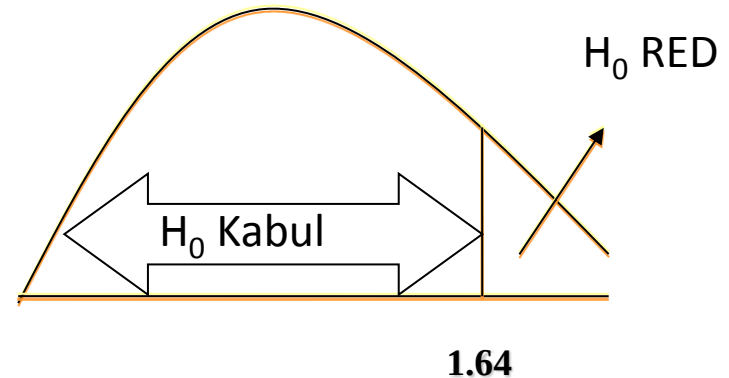
II. Aşama (c): Belirlenen I. Tip hataya Bağlı Olarak H_0 Hipotezi için Kabul ve Red Bölgelerinin Saptanması:



$$H_0: \mu = 180$$
$$H_1: \mu \neq 180$$



$$H_0: \mu = 180$$
$$H_1: \mu < 180$$



$$H_0: \mu = 180$$
$$H_1: \mu > 180$$

III. Aşama: İstatistiksel test değerinin hesaplanması:

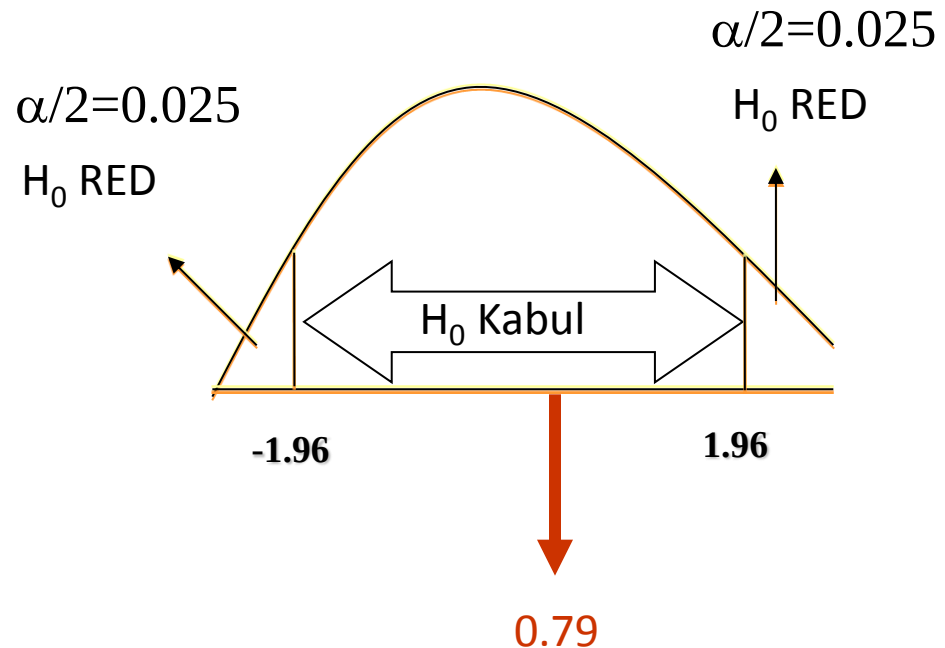
- Bu adımda daha önce kullanılmasına karar verilen test istatistiği (t, F, ki-kare vb.) değeri gözlemlerden elde edilen verilere dayalı olarak hesaplanır.

IV. Aşama: İstatistiksel Kararın verilmesi ve Yorumlama:

- Yapılacak test sonucunda hesapla bulunan test istatistiği değeri belirli bir teorik dağılıma uyar (örneğin standart normal dağılım veya t dağılımı gibi). Eğer hesapla bulunan test istatistiği değeri teorik tablo değerine eşit ya da büyük ise H_0 **RED** edilir.
- Hesapla bulunan test istatistiği teorik tablo değerinden küçük ise H_0 **KABUL** edilir.
- Diğer bir yol ise daha önce bahsedildiği gibi test sonucunda elde edilen p değeri ile karar vermektir.
- P değeri, daha önce belirlediğimiz yanılma düzeyinden küçük ise H_0 **RED edilir**, eğer p değeri belirlenen hata düzeyinden büyük ise H_0 **KABUL** edilir.

- Örnek 5 için $\alpha=0.05$ olarak alalım ve çift yönlü hipotez kurmuş olalım.
- Yapılan hipotez testi sonucunda hesaplanan z test istatistiği 0.79 olsun.

$$H_0: \mu = 180$$
$$H_1: \mu \neq 180$$



Dolayısıyla H_0 KABUL edilir.

- Yorum: Örneklemin çekildiği evrenin kolesterol ortalaması 180'e eşittir. Dolayısıyla normal olarak kabul edilebilir.

SERBESTLİK DERECESESİ

- Örneklemden hesaplanan bir istatistiğin evren değerini kestirmek amacıyla yapılan hesaplamalarda ya da test istatistiğinin tablo değerlerini belirlemede serbestlik derecesinin (degrees of freedom) ihtiyaç vardır.
- Serbestlik derecesinin sayısı, bir istatistiğin nihai hesaplamasında, değişiklik göstermekte serbest olan sayıdır. Bir değişkene ilişkin elde edilen puanların değişiklik gösterebilme serbestliği olarak ifade edilebilir.

SERBESTLİK DERECESESİ

- 8 kişilik bir öğrenci grubundan eldeki 8 balonu seçmeleri istendiğini varsayalım. Burada ilk 7 öğrenci seçme hakkına sahip iken sona kalan öğrencinin seçme serbestliği yani seçme hakkı kalmamış olur. Bu durumda serbestlik derecesi $8-1=7$ olarak alınabilir.



SERBESTLİK DERECESESİ

- Toplamlarının 200 olduğunu bildiğimiz 3 sayı olduğunu düşünün. İlk sayıyı seçmekte özgürüz. İkinci sayıyı da seçmede özgürüz. Fakat üçüncü sayıyı seçme gibi bir özgürlüğümüz yoktur. Seçilen ilk iki sayıya göre üçüncü sayı kendiliğinden belirlenecektir. Bu durumda da serbestlik derecemiz $3-1=2$ 'dir.

ÖRNEKLEME DAĞILIMI ve ORTALAMANIN GÜVEN ARALIĞI TAHMİNİ

- Örneklemenin dağılımı, örneklemelerden hesaplanan istatistiklerin dağılımını tanımlar. Araştırmacı, örneklemde hesaplanan bir istatistiğin evren değerini tahmin etmek için örneklem dağılımının özelliklerini kullanır.
- Örneklemelerin ortalaması, evren ortalamasının yansız bir tahmini olarak düşünülür.

ÖRNEKLEME DAĞILIMI

- Örneklemelerden hesaplanan istatistiklerin dağılımına örneklem dağılımı denir.
- N genişliğindeki evrenden n genişliğinde elde edilen tüm örneklemelerden birer ortalama hesaplanabilir ve bu örneklem ortalamalarının bir dağılımı elde edilebilir.
- Buna ortalamanın örneklem dağılımı adı verilir.

ÖRNEKLEME DAĞILIMI

- Aynı büyüklükteki örneklemlerden elde edilen ortalamalar ne kadar birbirine yakınsa (örneklemden örneklem değişim ne kadar azsa) herhangi bir örneklem sonucu o kadar güvenilirdir ya da kesindir.
- Eğer hesaplanan ortalamalar, bir örneklemden diğerine çok farklılık gösteriyorsa, çekilen herhangi bir örneklemden elde edilen ortalama (kestirim) o derece az güvenilir ya da kesindir.
- Bu nedenle örneklem dağılışının standart sapması kesinliğin ya da hatanın bir ölçüsü olarak kullanılır.

ÖRNEKLEME DAĞILIMI

- Pratikte hiçbir zaman olası tüm örneklemeleri ya da bir evrenden bir çok örnekleme çekmeyiz. İstatistik kuramı elimizdeki bir örneklemden yararlanarak örnekleme dağılımının özelliklerini bulmamıza yardımcı olur.
- Normal dağılımı tanımlayan parametreler, dağılımın ortalaması ve standart sapması olduğundan bu parametrelerin özelliklerinin bilinmesi gerekir.

Evren Ortalamasının Tahmini

- Evren ortalaması iki farklı yaklaşımla tahmin edilebilir: **Nokta tahmini** ve **aralık tahmini**.
- Evren ortalamasının nokta tahmini, örneklemden hesaplanan ortalamadır. Bu ortalamanın verilmesi standart hatanın verilmesiyle anlam kazanır.
- Aralık tahmini, evren değeri belli olasılıklarla (ör. %95) kapsayan bir güven aralığının oluşturulması işlemidir.

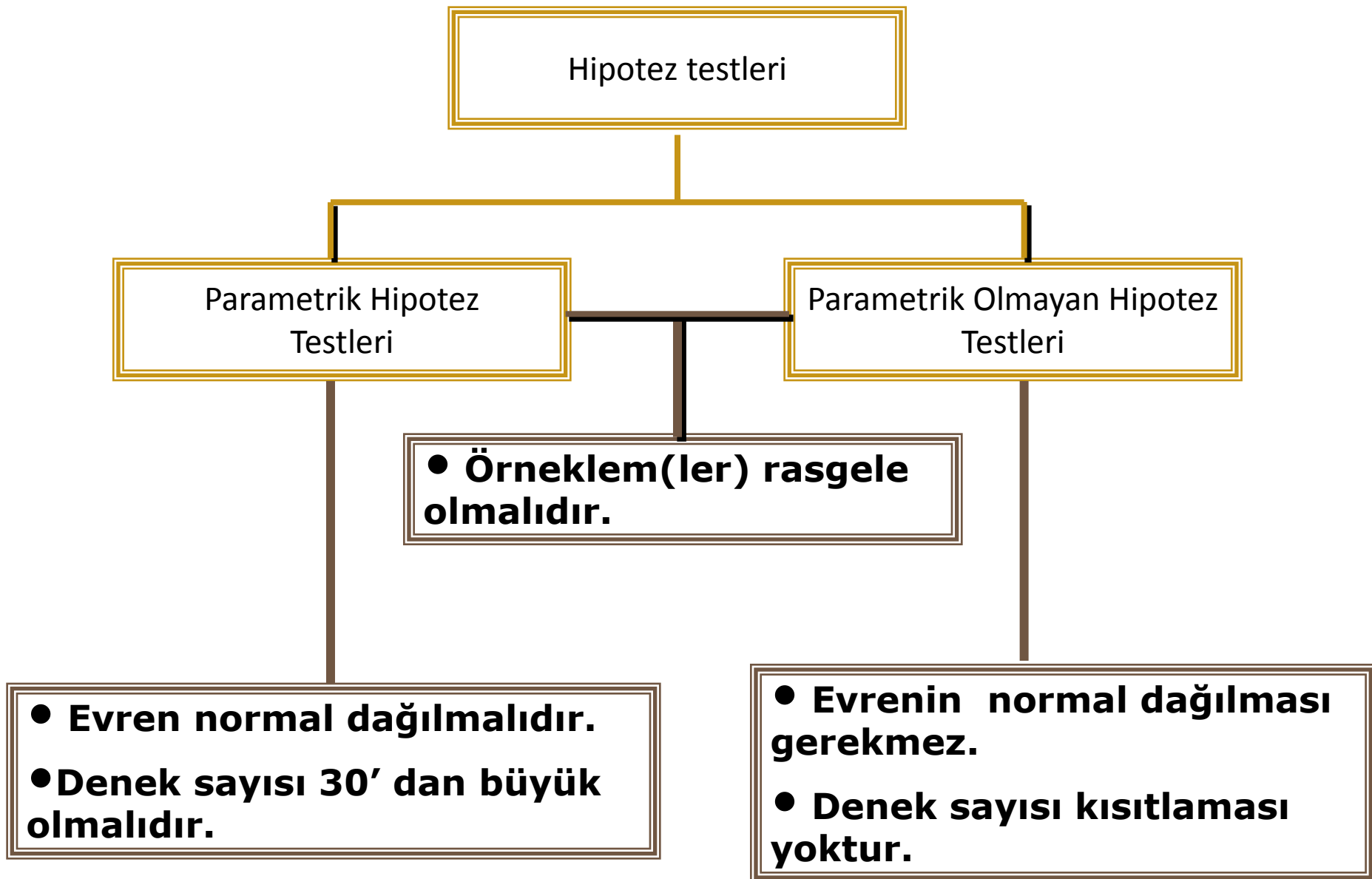
Aralık Tahmini

- Aralık tahmini, bir evren değerinin tahmininde kullanılan, örneklemden hesaplanan istatistiğe dayalı olarak bulunan ve iki değerle sınırlandırılan bir aralıktır. Örneğin aşağıda ortalama için %95 güven aralığı verilmiştir.

$$\bar{x} \pm 1,96 \times \sigma_{\bar{x}}$$

UYGUN İSTATİSTİĞİN SEÇİMİ

- Bir istatistiksel tekniğin seçiminde dikkate alınması gereken noktalar şöyle özetlenebilir:
 - Bağımlı değişkenin ölçme düzeyi
 - Alt örneklemelerin sayısı ve büyüklükleri
 - Deneysel ya da istatistiksel kontrol



ANALİZ TÜRLERİ

Bağımlı Değ.	Bağımsız Değ.	Analiz
Sürekli	İki kategorili	t-testi, Wilcoxon testi
Sürekli	Kategorik	ANOVA, doğrusal regresyon
Sürekli	Sürekli	Korelasyon, doğrusal regresyon
İki kategorili	Sürekli	Lojistik regresyon
İki kategorili	İki kategorili	Ki-Kare testi, lojistik regresyon

Uygun Analiz Yönteminin Seçilmesi

Uygun Analiz Yönteminin Seçilmesi

- Uygun analiz türünün belirlenmesinde ilk kriter verilerin türüdür. Analiz yöntemleri verilerin özelliklerine göre iki temel gruba ayrılır. Bu gruplarda yer alan temel analiz yöntemleri aşağıdaki gibidir.
- (1) Parametrik veriler için kullanılan analiz yöntemleri; Varyans Analizi, t-testi, Pearson Korelasyonu.
- (2) Parametrik olmayan veriler için kullanılan analiz yöntemleri; Mann-Whitney, Ki-Kare Testleri, Spearman Korelasyonu.



Parametrik Testlerin Varsayımları

	Parametrik	Parametrik olmayan
Varsayılan dağılım	Normal	Yok
Varsayılan varyans	Homojen	Yok
Veri tipi	Eşit oranlı ve eşit aralıklı ölçekleri	Sınıflama ve sıralama ölçekleri
Data set relationships	Independent	Any
Genel merkezi eğilim ölçüsü	Aritmetik ortalama	Medyan
Artıları	Daha fazla sonuç çıkarılabilir	Basitlik ve uç değerlerden az etkilenme
Seçim yaparken kullanılan test türü	Parametrik test	Parametrik olmayan test
Korelasyon türü	Pearson	Spearman
2 gruplu, bağımsız örneklerde	Bağımsız örnekler t-testi	Mann-Whitney testi
Bağımsız örneklerde, 2'den fazla grupta	Tek yönlü ANOVA	Kruskal-Wallis testi

Varsayımlar

- Parametrik vs. Parametrik olmayan testler
- varsayımlar:
- Veri dağılımının normalliği: bazen bağımlı değişken bazen de hataların dağılımını kastetmektedir. t-testi, ANOVA, regresyonda kontrol edilir.
- Varyans homojenliği: Verideki değişkenler boyunca varyansların aynı olması manasına gelir. t-testi ve ANOVA
- Eşit aralıklı veri: Değişkenler en az eşit aralık düzeyinde olmalıdır. Test etme yolu yoktur sadece değişkenin puanlamasını inceleyerek anlaşılır.
- Bağımsızlık: Yapılan analize göre değişmekle beraber varsayımda belirtilen değişkenler arasında bağımlılık olmaması durumunu kastetmektedir.

PARAMETRİK ANALİZLERE ÖRNEKLER

- t-testi
- Tek yönlü ANOVA (Varyans analizi)
- Pearson Korelasyonu
- Basit Doğrusal Regresyon
- Çoklu Doğrusal Regresyon



Parametrik Testlerin Varsayımları (sayıltıları)

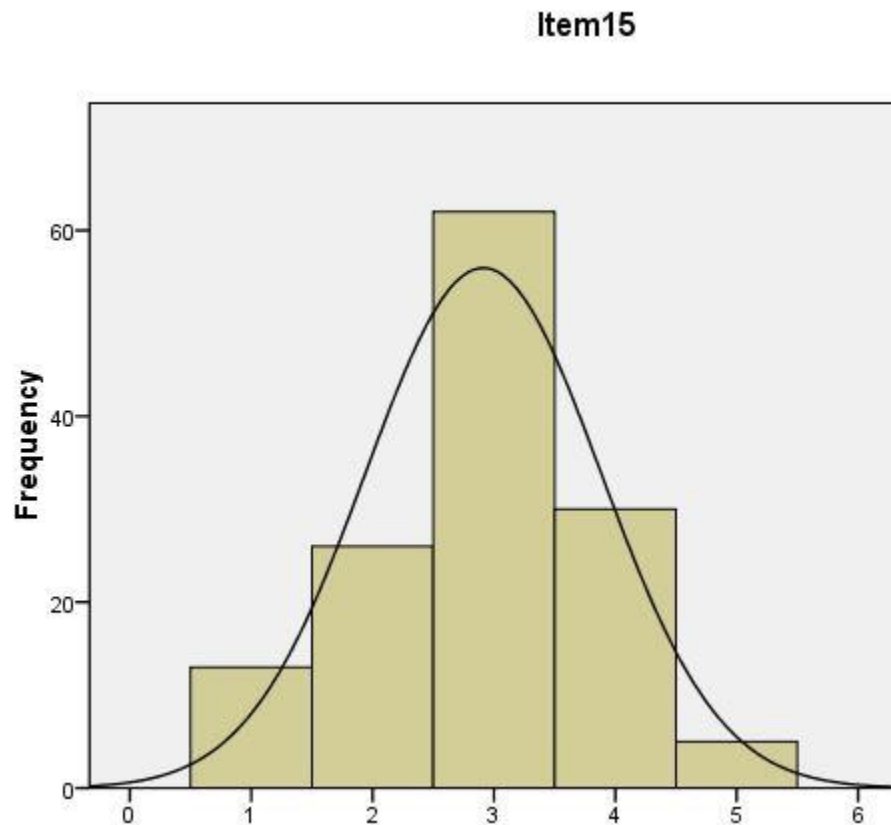
- Kullanılması için şartlar:
- Evrenle ilgili
 - Normal dağılım
 - Homojen varyanslar
- Örneklemle ilgili
 - Denekler rasgele seçilmeli
 - Denekler birbirinden bağımsız seçilmeli
 - Veriler en azından Interval (eşit aralıklı) düzeyde ölçülmeli

Normallik Varsayımı

- Normallik görsel olarak kontrol edilebildiği gibi sayısal verilerle de kontrol edilebilir. Ayrıca normallik test etmek için 2 tane de test üretilmiştir.
- Görsel olarak histogram ve P-P plot yardımıyla
- Sayısal olarak çarpıklık ve basıklık değerleri yardımıyla
- Test olarak Kolmogorov–Smirnov test ve Shapiro–Wilk testleri kullanarak test edilebilir.

Normallik Varsayımı Nasıl Test Edilir?

- Frekans dağılımına (histogram oluşturarak) bakarız



Normallik Varsayımı Nasıl Test Edilir?

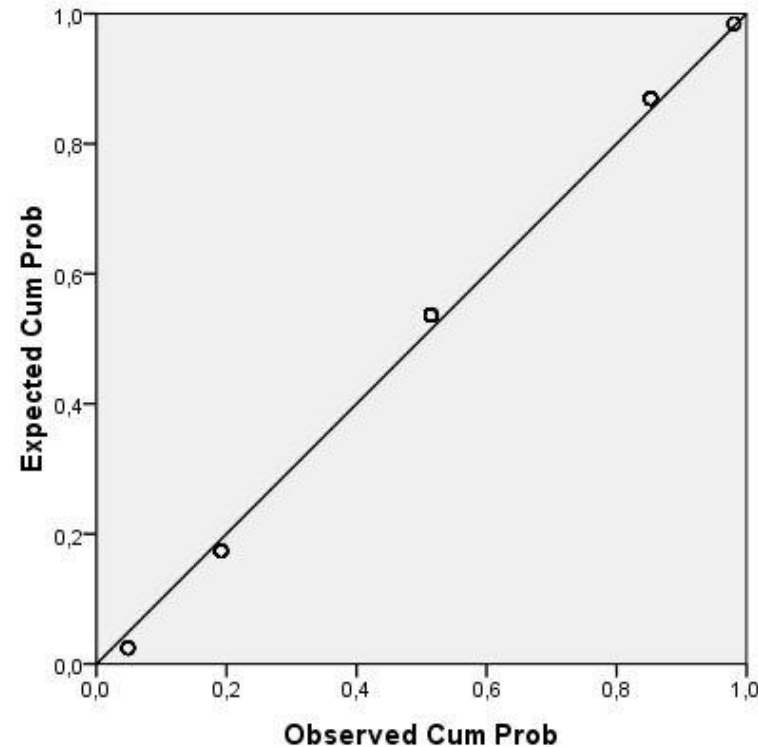
Normal P-P Plot of Item15

- **P-P plot**
- **Analyze>Descriptive>P-P plot**
- Örnek hsb2 data

- **Q-Q Plot**

- Q-Q plot ile P-P plot'un farklı grafiği oluştururken kullandıkları veri. P-P plot

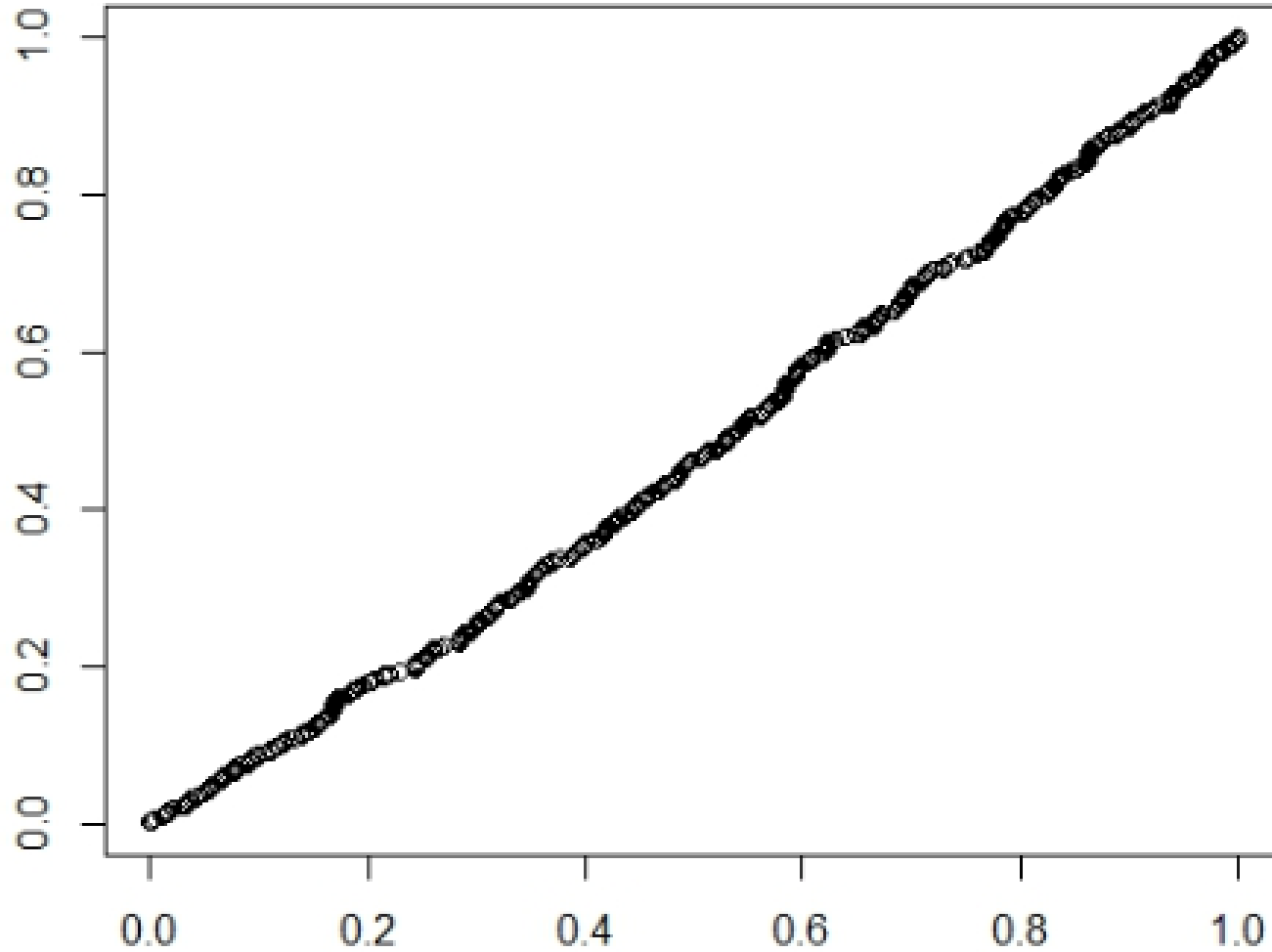
örneklemdaki her bir değerin kendisini kullanırken Q-Q plot bu değerlerden üretilen quantile çeyreklikleri kullanıyor.



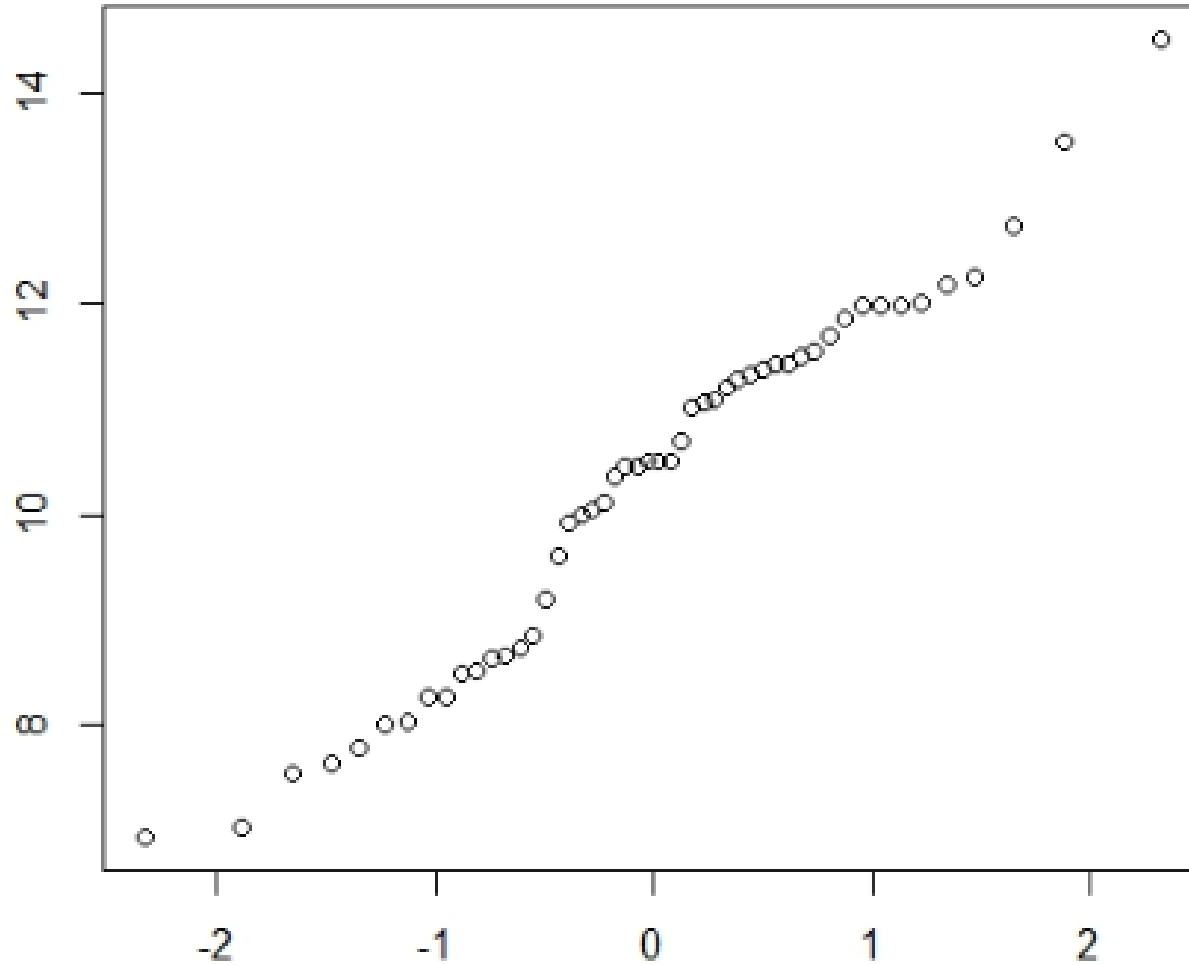
P-P Plot ve Q-Q Plot farkı

- Temel fark: Q-Q grafiği, bir veri dağılımının çeyrekliklerini (quantiles), belirli bir dağılım ailesinden standartlaştırılmış bir teorik dağılımın nicelleri ile karşılaştırır, fakat P-P, bir veri kümesinin ampirik kümülatif dağılım fonksiyonunu, belirtilen teorik kümülatif dağılım fonksiyonu $F(.)$ ile karşılaştırır.
- P-P grafiği önerilen dağılımdan orta kısımdaki sapmaları ölçerken ve Q-Q ise önerilen dağılımdan kuyruk kısımlarındaki (uç kısımlarda) sapmaları ölçer.

Q-Q Plot Normal Dağılım



Q-Q Plot Normal Dağılım



Normalliğin Çarpıklık ve Basıklık Değerleri ile Kontrolü

- Basıklık (kurtosis) ve çarpıklık (skewness) için sıfır ve sıfıra yakın değerler normalliği sıfırdan uzak değerler normallikten uzaklaştığını gösterir. Basıklık ve çarpıklık değerleri ± 2 'den (ya da ± 3 'ten) yüksek olduğu durumlar normalliğin bozulduğuna işaret eder.
- $Z = S - 0 / (se)$ skewness ve kurtosis için kullanılabilir.
- 1.96'dan büyük değerler 0.05 derecesinde anlamlı sayılmakta ve normalliğin bozulduğuna işaret etmektedir.

Statistics

vas		
N	Valid	131
	Missing	5
Mean		26,20
Median		25,00
Mode		25
Std. Deviation		3,594
Variance		12,914
Skewness		1,124
Std. Error of Skewness		,212
Kurtosis		,891
Std. Error of Kurtosis		,420
Range		18
Minimum		20
Maximum		38

Normallik Varsayımı Nasıl Test Edilir?

- **Kolmogorov–Smirnov test ve Shapiro–Wilk test**
- Bu testler örneklemdaki veriyi normal olan başka bir veriyle karşılaştırıyor. Bu karşılaştırma sonucunda bir test istatistiği ortaya çıkıyor. Eğer bu istatistik anlamlı değil ise ($p > .05$), örneklemdaki veri normal olan veriden istatistiksel olarak farklı değildir yani normaldir ya da normale yakındır diyebiliriz.
- **Analyze>Descriptive Statistics>Explore>Plots>Normality plots with tests**

Tests of Normality						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
science	,076	200	,007	,985	200	,035

a. Lilliefors Significance Correction

Varyansların Homojenliği Testi

- SPSS'te varyansların homojen olup olmadığı **Levene's Test** kullanılarak test edilebilir.
- Levene's Testte null hipotez farklı grupların varyanslarının eşit olduğudur. Bu test anlamlı olmadığı zaman ($p > .05$) varyansların eşit olduğu ve

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$$
$$H_a: \sigma_i^2 \neq \sigma_j^2 \text{ for at least one pair } (i,j)$$



Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
TOTALDEP	Based on Mean	,187	1	110	,666
	Based on Median	,032	1	110	,859
	Based on Median and with adjusted df	,032	1	103,037	,859
	Based on trimmed mean	,136	1	110	,713

Parametrik test varsayımları

- Eğer bu varsayımların hiçbirisi yerine gelmemişse parametrik olmayan testler tercih edilebilir.
- Eğer bazıları yerine gelmişse, normallik ya da varyans homojenliğini sağlamak için transformasyon (dönüşüm) yapılabilir.
- Log-karekök- $1/n$ vb. dönüşümlerle veri normal hale getirilebilir.

Hangi İstatistik Analizini Seçeceğim?

- Parametrik ya da parametrik olmayan istatistik yöntemlerden hangisi kullanacağımıza karar vermeden önce araştırma sorumuza uygun istatistik yöntemi hangisidir buna karar vermeliyiz.
- Bilimsel araştırma sürecinde olduğu gibi her şey araştırma sorusu/sorularıyla başlar. Analiz yöntemimize de karar verirken eldeki veriden sonuca gitmek yerine araştırma sorusu ışığında veri toplayıp araştırma sorusunu cevaplayacak istatistiksel yöntemleri seçmemiz gerekiyor.
- Örneğin 2 değişken arasındaki farkı sorgulayan bir araştırma sorusunu 2 değişken arasındaki ilişkiyi gösteren bir istatistik yöntemiyle (korelasyonla) cevaplayamayız!

Hangi İstatistik Analizini Seçeceğim?

- Eğer iki grubun ortalamaları arasında anlamlı bir fark var mıdır diye merak ediyorsak seçmemiz gereken istatistik yöntemi bağımsız örneklem t-testidir.
- Eğer bir grubun iki farklı koşuldaki (eğitim öncesi/sonrası) durumları arasında anlamlı bir fark var mıdır diye merak ediyorsak kullanmamız gereken istatistik yöntemi bağımlı örneklem t-testidir.
- Bu iki parametrik testte bağımlı değişkenin normal dağılıma sahip olduğu varsayılır. Eğer bağımlı değişken çok az sayıda ise ve normal dağılım göstermiyorsa parametrik olmayan versiyonları **Mann-Whitney U ve Wilcoxon Signed Testi** seçmeliyiz.

Hangi İstatistik Analizini Seçeceğim?

- T-testini 2 grubu (durumu) karşılaştırırken kullanıyoruz. Eğer 2'den fazla grubun arasında anlamlı bir farklılık var mıdır diye merak ediyorsak ANOVA tercih edeceğimiz istatistik yöntemi olacaktır. ANOVA da diğer parametrik istatistikler gibi varsayımlara sahiptir. Varsayımlar yerine gelmemişse parametrik olamayan versiyonu **Kruskal Wallis one way ANOVA** tercih edilebilir.

Hangi İstatistik Analizini Seçeceğim?

- Eğer iki değişken arasındaki ilişki nedir sorusunu merak ediyorsak uygulamamız gereken korelasyondur. **Pearson korelasyon** katsayısı 2 sürekli değişken arasındaki ilişkiyi gösterirken kullanılır. Eğer 2 değişkene ait veriler sıralama tarzında ise ya da sıralanabiliyorsa bu 2 değişken arası ilişki **Spearman's sıra korelasyonu** adlı parametrik olmayan korelasyon yöntemi kullanarak incelenir.

Hangi İstatistik Analizini Seçeceğim?

- Eğer bir bağımsız değişkenin bağımlı değişkeni ne kadar açıkladığını ya da ne kadar yordadığını (tahmin) merak ediyorsanız yapacağınız iş regresyon analizi yapmaktır. Eğer 1 tane bağımsız değişkenden bahsediyorsak basit doğrusal regresyon kullanılır eğer 2 veya daha fazla bağımsız değişkenin 1 bağımlı değişken üzerindeki etkisini inceliyorsak çoklu doğrusal regresyon analizi tercihimiz olacaktır. Örneğin ders başarısı (sınav puanı) üzerinde haftalık ders çalışma saati ve okunan kitap sayısının etkisine bakmak için çoklu regresyon kullanılır.