

BİYOİSTATİSTİK

Status: Devlet,durum

İstatistik: Herhangi bir konuyu incelemek için gerekli verilerin toplanmasını, toplanan verilerin değerlendirilmesini ve değerlendirme sonucu karara varılmasını sağlayan bilimdir.

İstatistik herhangi bir olgu ya da olayın nicesel yönünün sayılarla anlatımıdır.

Bradford-Hill ise istatistiği şöyle tanımlar:

İSTATİSTİK BİLİMİ, çeşitli etkenler altında bulunan bir olay üzerinde belirli bir etkenin etkisinin değerlendirilmesini sağlayan ve evrenden alınan bir örnek üzerinde yapılan incelemeler sonunda evren hakkında doğru bir fikir elde edilmesini sağlayan bilimdir.

İstatistiksel yöntemler sağlık ve biyoloji alanında kullanıldığında Biyoistatistik adı verilmektedir.

İstatistik konu olarak iki ana gruba ayrılır:

1. TANIMLAYICI İSTATİSTİK

Elde edilen verilerin sınıflandırılması, frekans dağılımlarının yapılması, bu dağılımların ortalamalar, çeyrek ve yüzdelikler, standart sapma vb. ölçülerle tanımlanması ve bulguların tablo ve grafiklerle okuyuculara sunulması tanımlayıcı istatistiğin konularıdır.

2. ÇIKARIMSAL İSTATİSTİK

Örneklemden elde edilen bulgularla Örneklemin çekildiği evren hakkında tahminlerde bulunma, karşılaştırmalar yapma ve kararlara varma işlemleri çıkarımsal istatistiğin konularıdır.

Evren: Belirli bir özelliğe sahip bireylerin tümünün oluşturduğu topluluk olarak tanımlanabilir. Evren büyük ya da küçük, sonsuz ya da sonlu olabilir.

Örneklem: Çekildiği evreni temsil ettiği düşünülen ve evrenden çekilen küçük bir grubun oluşturduğu topluluktur.

Örnekleme: Örnekleme seçmek için yapılan işlemlerin tümü.

Parametre: Evreni tanımlamak için kullanılan ölçülere parametre denir. Ör: Evren ortalaması (μ).

Veri: Bir olayı aydınlatmak ya da bir gerçeği ortaya çıkarmak için toplanan materyal (ölçüm, bilgi, belge, madde vb.)

Karakter: Karakter canlılar için kullanılan bir terimdir. Genel anlamda canlının herhangi bir özelliği olarak tanımlanır.

Faktör: Faktör, çevre özelliği için kullanılan bir terimdir. Ör: hava sıcaklığı vs.

Değişken: İncelenen karakter ya da faktör değişik kişilerde, yerlerde ya da durumlarda değişik değer alabilir. Bu nedenle, değişken terimi karakter ya da faktör yerine genel anlamda kullanılan bir terimdir. Başka bir deyişle incelenen özelliktir.

VERİNİN ÖLÇÜM BİÇİMİ

Verileri ölçüm biçimine göre 3 grupta toplayabiliriz:

1. Ölçümle belirtilen sürekli veriler. İki aralıkta noktalı değer alabilen değişkenlerdir. Genellikle normal dağılıma uygunluk gösterirler. Sınıflar birbirine geçişlidir. “Nicel” değişkenler de denir.
2. Sayısal olarak belirtilen kesikli veriler. İki aralıkta noktalı değerler alamayan verilerdir.
3. Nitelik (isimsel) olarak belirtilen veriler.

BİYOİSTATİSTİĞİN SAĞLIK BİLİMLERİNDE KULLANIMI

1. Hizmet planlamasında kullanım.
2. Toplumsal değişimlerin incelenmesinde kullanım.
3. Tanı ve tedavi işlemlerinde kullanım.
4. Koruyucu hizmetlerde kullanım.
5. Biyolojik, morfolojik ve fizyolojik özelliklerin tanımlanmasında kullanım.
6. Bilimsel çalışmalarda kullanım.
7. Hizmet göstergesi olarak kullanım.

FREKANS (SIKLIK) DAĞILIMLARI VE TANIMLAYICI ÖLÇÜLER

Verilerin Sınıflandırılması: Verilerin sınıflandırılması, özellikle denek sayısı fazla olduğunda, hem veriler üzerinde yapılacak harcamaları kolaylaştırır, hem de verilerin kolay anlaşılır biçimde okuyucuya sunulmasını sağlar. Niceliksel verileri sınıflandırırız.

Dağılım Aralığı (Range) (R) : En büyük değer ile en küçük değer arasındaki fark (aralık).

Sınıf Sınırı: Her sınıfın bir alt ve bir üst değeri vardır. Bunlara o sınıfın alt ve üst sınırı denir.

Sınıf Aralığı (c): Her sınıfın alt sınırı ile üst sınırı arasındaki aralıktır. Yani, ard arda gelen iki sınıfın alt ve üst sınırları arasındaki farktır (Aynı sınıfın alt sınırı – üst sınır + 1 birim). $c = \frac{R}{k}$

Sınıf Sayısı (k): Ard arda gelen bütün sınıfların kaç tane olduğunu gösteren sayıdır.

Sınıf Değeri (SD): Sınıfın ortalamasıdır. Her sınıfın alt sınırı ile üst sınırı toplanıp 2'ye bölünerek bulunur. Sınıfı yaklaşık olarak temsil eden değerdir.

Sınıf Ara Değeri (SAD): Bir üstteki sınıfın üst sınırı ile bir alttaki sınıfın alt sınırının ortalamasıdır. İkinci sınıfın SAD'den sınıf aralığı çıkarılarak ilk sınıfın SAD bulunur. Sınıfları kendinden öncekinden ayıran gerçek sınırdır.

Sınıf Frekansı (sıklık) (f): Her sınıfa giren veri sayısını ifade eder. Çetelenerek bulunur.

Yığılımlı (kümülatif) Frekans (Y_f): Her sınıfın frekansının önceki frekanslarla toplamıdır.

SINIFLANDIRMANIN KURALLARI

1. Sınıf sınırları kesin olmalıdır. Sınıflar birbirine karışmamalıdır.
2. Sınıflama bütün değerleri içine almalıdır.
3. Bütün sınıflarda sınıf aralıklarının eşit olması tercih edilir. Şart değildir. İstatistiksel amaçlı değerlendirmelerde önemlidir.
4. İncelemeyi kolaylaştırmak ve dağılım hakkında yeterli bilgiye sahip olabilmek için sınıf sayısının 8-15 arasında olması önerilir.

I- SINIFLANDIRMANIN YAPILIŞI

1. Dağılımdaki en küçük ve en büyük değer bulunur.
2. En büyük değerden en küçük değer çıkartılarak dağılım aralığı (Range)(R) bulunur.
3. Dağılım aralığı bir kez 8'e, bir kez de 15'e bölünerek sınıf aralığı saptanmaya çalışılır.

$$k_1 = 8 \quad c_1 = \frac{R}{k_1} \quad c_2 = \frac{R}{k_2}$$
$$k_2 = 15$$

c_1 ve c_2 arasındaki herhangi bir uygun değer sınıf aralığı olarak alınabilir. (Verinin birimine göre uygun olmalı.)

4. Dağılımdaki en küçük değer (min) ilk sınıfın alt sınırı olarak alınır. Bu değere her seferinde sınıf aralığı kadar ilave edilerek bundan sonraki sınıfların alt sınırları bulunur.
5. İkinci sınıfın alt sınırından 1 birim çıkarılarak ilk sınıfın üst sınırı bulunur. Bu üst sınıra her seferinde sınıf aralığı kadar ilave edilerek bundan sonraki sınıfların üst sınırları bulunur.
6. Sınıf oluşturma işlemine dağılımdaki en üst değer (veri) kapsanıncaya kadar devam edilir.

II- Sınıf Frekanslarının Belirlenmesi

Dağılımdaki herhangi bir değer tek tek gözden geçirilir ve ait olduğu sınıfın karşısına bir işaretle gösterilir (çeteleme). Bu işaretler sayılarak frekanslar belirlenir.

FREKANS DAĞILIMLARINI TANIMLAYICI ÖLÇÜLER

1. Yer gösteren ölçüler
 - a) Merkez ölçüleri: Ortalamalar
 - b) Çeyrek ve yüzdeler
2. Yaygınlık ölçüleri
 - a) Standart sapma
 - b) Varyans
 - c) Varyasyon katsayısı
 - d) Standart hata
 - e) Range

Aritmetik Ortalama

Aritmetik Ortalama deneklerin aldıkları değerlerin toplanıp denek sayısına bölünmesiyle elde edilen matematiksel gerçek bir değerdir.

Bu nedenle deneklerin aldıkları değerlerden, özellikle "aşırı" değerlerden etkilenir. Aşırı değerler dağılımdaki diğer değerlerden çok farklı olan az sayıdaki değerler olarak tanımlanabilir. Bu aşırı değerler; ölçüm hatalarına, materyalin ya da ölçüm aracının bozulmasına, yanlış kaydetmeye bağlı olarak ortaya çıkabileceği gibi, kişilik farklılıklarından dolayı ortaya çıkan gerçek değerler de olabilir. Bu değerler istatistiksel olarak fazla bir anlam taşımaz.

Bir dağılımda aşırı değerler varsa yapılacak işlemler şunlardır:

1. Kişilik farkına bağlı gerçek bir değer değilse ve olanak varsa ölçüm tekrarlanmalıdır.
2. Aşırı değerler değerlendirme dışı bırakılabilir.
3. Aşırı değerler alan bireylere diğer değerlere yakın bir değer atanabilir.
4. Bunlar yapılamıyorsa aritmetik ortalama yerine başka bir ortalama ölçüsü, örneğin ortanca kullanılmalıdır.

Aritmetik Ortalamanın Hesaplanması

Aritmetik ortalama sembolü "X" dir. Sınıflanmamış ve sınıflanmış verilerde ayrı yollarla hesaplanır.

Sınıflanmamış Verilerde Aritmetik Ortalamanın Hesaplanması

Bütün deneklerin değerlerinin toplanıp denek sayısına bölünmesiyle bulunur.

örnek : 30 kişinin hemoglobin değerleri aşağıda verilmiştir.

13.0	13.6	14.0	12.8	11.4	12.4	13.4	11.7	14.2	12.9
13.5	12.6	12.3	12.1	11.8	10.8	10.5	11.6	13.4	14.6
10.1	12.9	11.0	15.0	9.9	10.0	11.4	10.8	12.0	10.3

Bu dağılımın aritmetik ortalaması nedir?

$$\bar{X} = \frac{13.0 + 13.6 + \dots + 12.0 + 103}{30} = \frac{366.0}{30} = 12.2$$

Sınıflanmış Verilerde Aritmetik Ortalamanın Hesaplanması

Sırasıyla şu işlemler yapılır:

1. Sınıflar yazılır.
2. Sınıf değeri (SD) bulunur ve sınıfın karşısına yazılır. Sınıf değeri sınıfın ortalamasıdır. Her sınıfın alt sınıfın ile üst sınıfın toplanıp 2'ye bölünerek bulunur. Aşağıdaki örnek problemde ilk sınıfın sınıf değeri $(15 + 19) / 2 = 17$ 'dir.
3. Her sınıfın frekansı karşısına yazılır.
4. Çalışma birimi olarak adlandıracağımız b kolonu geliştirilir. Bu kolonda herhangi bir sınıfın karşısına sıfır yazıldıktan sonra üste doğru eksi olarak l'den başlanarak ve birer artırılarak, alta doğru artı olarak l'den başlanarak ve birer artırılarak çalışma birimleri yazılır, istenilen herhangi bir sınıfın karşısına sıfır konabilir. Genellikle hesaplama işlemlerini kolaylaştırmak için frekansı en büyük sınıfın karşısına konur.
5. Frekansla çalışma biriminin çarpımları (fb) alınarak her sınıfın karşısına yazılır. Sonra işaretleri dikkate alınarak toplanır. Toplam fb (Σfb) bulunur. Σfb eksi ya da artı olabilir.
6. Değerler formüle yerleştirilerek aritmetik ortalama bulunur. Sınıflanmış verilerde aritmetik ortalama formülü:

$$\bar{X} = A + \frac{\Sigma fb \cdot xc}{n}$$

A = b kolonunda karşısına sıfır konulan sınıfın sınıf değeri

C : Sınıf aralığı

n : Denek sayısı

Aritmetik Ortalamanın Özellikleri

Avantajları:

1. Hesaplanması ve anlaşılması kolaydır.
2. Her dağılımda tektir.
3. Aritmetik işlemler için elverişlidir.

Dezavantajları:

1. Dağılımdaki aşırı (uç) değerlerden ileri derecede etkilenir.

Ortanca (Medyan)

Ortanca, dağılımın orta noktasındaki değer olarak tanımlanabilir. Başka bir deyişle, ortanca öyle bir değerdir ki, dağılımdaki değerlerin % 50'si ortancaya eşit ve/veya daha küçük, % 50'si ortancaya eşit ve/veya daha büyüktür. Bu nedenle ortanca dağılımdaki aşırı değerlerden etkilenmez.

Ortanca sınıflanmamış ve sınıflanmış verilerde ayrı yollarla hesaplanır.

Sınıflanmamış Verilerde Ortancanın Hesaplanması

3'âilımdaki değerler ya küçükten büyüğe ya da büyükten küçüğe doğru sıralanarak tam ortadaki değer bulunur. Denek sayısı tek ise $(n+1) / 2$ 'nci değer tam ortadaki değerdir. Denek sayısı çift ise tam orta noktada bir değer yoktur. Bu durumda $n/2$ 'nci değer ile $(n+2)/2$ 'nci değer toplanıp 2'ye bölünerek ortanca bulunur.

Sınıflanmış Verilerde Ortancanın Hesaplanması

Sırası ile şu işlemler yapılır:

1. Sınıflar yazılır.
2. Her sınıfın frekansı yazılır.
3. Yığılımlı frekans (Yf) bulunur. Yığılımlı frekans her sınıfın frekansının önceki frekanslarla toplamıdır. Bu toplam her sınıfın karşısına yazılır.
4. Sınıflanmış verilerde ortanca formülü:

$$\text{Ortanca} = L + \left(\frac{n/2 - Yf_1}{f} \right) \times C$$

L:Ortancanın değeridir.	Ortancanın içinde bulunduğu sınıfın ara
Yfi	: Ortancanın içinde bulunduğu sınıfın bir üstündeki sınıfın yığılımlı frekansı
f	: Ortancanın içinde bulunduğu sınıfın frekansı
C	: Sınıf aralığı
n	: denek sayısı

Ortancanın Özellikleri

Avantajları:

1. Aritmetik ortalama gibi hesaplanması ve anlaşılması kolaydır.
2. Her dağılımda tektir.
3. Dağılımdaki aşırı (uç) değerlerden etkilenmez.

Dezavantajları:

1. Aritmetik işlemler için elverişli değildir.

Tepe Değeri

Sınıflanmamış verilerde tepe değeri darıımda en çok görülen, diğer bir deyişle en çok tekrarlayan değerdir. Aşağıdaki dağılımın tepe değeri en çok görülen değer olan 11.0'dır.

10.5 10.0 10.4 11.0 11.0 11.6 12.0 11.8 11.0 11.0

Bir dağılımda aynı sayıda görülen değişik değerler varsa, diğer bir deyişle tepe değeri olabilecek birden çok değer varsa tepe değeri kullanılmamalıdır.

Sınıflanmış verilerde tepe değeri en fazla frekansa sahip olan sınıfın sınıf değeridir.

Tepe Değerinin Özellikleri

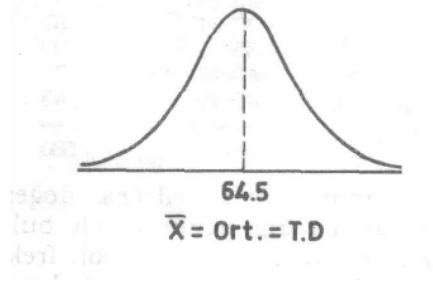
Avantajları:

1. Bulunması ve anlaşılması kolaydır.
2. Dağılımdaki aşırı (uç) değerlerden etkilenmez veya kendisi dışındaki hiçbir değerden etkilenmez.

Dezavantajları:

1. Bazı dağılımlarda tepe değeri birden fazla olabilir. O zaman tepe değeri kullanılmamalıdır.
2. Aritmetik işlemler için elverişli değildir.

Aritmetik Ortalama, Ortanca ve Tepe Değeri İlişkileri



Simetrik dağılımlarda aritmetik ortalama, ortanca ve tepe değeri birbirine eşittir.

Sağa
olduğundan
ortalamadan
aşağıdaki gibi olur:

çarpık
tepe
daha

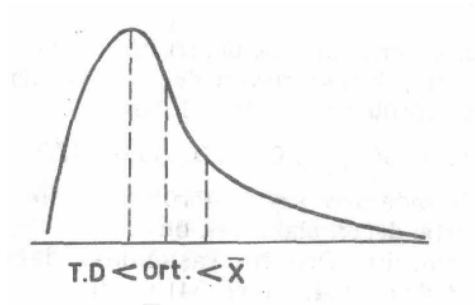
dağılımlarda
değeri
küçüktür.

küçük
ortancadan,
Dağılımın

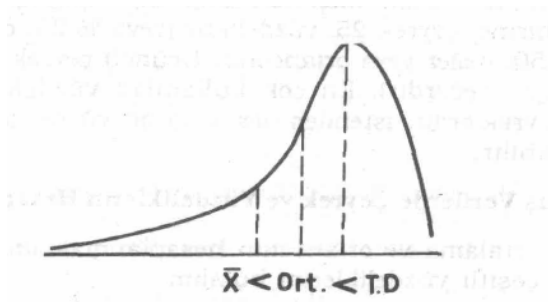
değerlerde
ortanca

bir
ise
grafikte

yığılma
aritmetik
görünümü



Sola çarpık dağılımlarda büyük değerlerde yığılma olduğundan tepe değeri ortancadan, ortanca ise aritmetik ortalamadan daha büyüktür. Dağılımın grafikte görünümü aşağıdaki gibi olur.



Geometrik Ortalama

Mikroorganizmaların çoğalması, nüfus artışı, fiyat artışı gibi birbirinin katlan olarak çoğalan yani geometrik artış gösteren verilerde ortalama hesaplamak için kullanılan bir ortalama ölçüsüdür, n tane değerin birbiriyle çarpımlarının n'inci kökü alınarak hesaplanır. Bu nedenle dağılımda negatif ya da sıfır değerleri varsa hesaplanamaz.

ÇEYREK ve YÜZDELİKLER

Ortalamalar dağılımın orta noktasını gösteren ölçülerdir. Çeyrek ve yüzdelikler ise dağılımın herhangi bir noktasını gösterirler. Örneğin, birinci çeyrek 25. yüzdeliktir (veya % 25. değerdir), ikinci çeyrek % 50. değer veya ortancadır, üçüncü çeyrek 75. yüzdeliktir (veya % 75. değerdir). En çok kullanılan yüzdelikler birinci ve üçüncü çeyreklerdir, istenilen herhangi bir yüzde değer de kolayca hesaplanabilir.

DAĞILIMIN YAYGINLIK ÖLÇÜLERİ

Standart Sapma ve Varyans

Aritmetik ortalama dağılımın orta noktasını gösteren ve dağılımı temsil eden bir ölçüdür. Ancak dağılımın yaygınlığı hakkında bilgi vermez. Aritmetik ortalamaları aynı olan iki dağılım aynı yaygınlıkta olmayabilir. Örneğin; 10, 22, 34 değerlerini alan 3 kişilik bir dağılımda aritmetik ortalama $66/3 = 22$ 'dir. 21, 23, 22 değerlerini alan başka bir 3 kişilik dağılımda aritmetik ortalama $66/3 = 22$ 'dir. İki dağılımın aritmetik ortalaması 22 olduğu halde birinci dağılımda değerler (1 ve 3'üncü değerler) aritmetik ortalamadan çok uzakta iken ikinci dağılımdaki değerler ortalamaya çok yakındır. Bir dağılımda değerler aritmetik ortalamadan uzaklaştıkça dağılımın yaygınlığı artar. Dağılımın yaygınlığını gösteren ölçülerin en önemlisi standart sapmadır.

Standart sapmanın karesi (S^2) varyans olarak adlandırılır. Standart sapma dağılımdaki her bir değer ortalama göre ne uzaklıkta olduğunu, diğer bir deyişle dağılımın ne yaygınlıkta olduğunu gösteren bir ölçüdür. Standart sapma büyüdükçe dağılım yaygınlaşır. Standart sapma sınıflanmış ve sınıflanmamış verilerde değişik şekilde hesaplanır.

Sınıflanmamış Verilerde Standart Sapmanın Hesaplanması

Sırasıyla şu işlemler yapılır:

1. Deneklerin aldıkları değerler toplanır.
2. Deneklerin aldıkları değerlerin teker teker kareleri alınır ve toplanır.
3. Bir ve ikinci işlemde bulunan değerler formüle yerine konarak standart sapma hesaplanır. Sınıflanmamış verilerde standart sapmanın formülü:

$$S = \sqrt{\frac{\sum_{i=1}^n X_i^2 - \frac{\left(\sum_{i=1}^n X_i\right)^2}{n}}{n - 1}}$$

S : Standart sapma

$\sum_{i=1}^n X_i^2$: Deneklerin aldıkları değerlerin karelerinin toplamı (İkinci işlemde bulunan değer)

$\sum_{i=1}^n X_i$: Deneklerin aldıkları değerlerin toplamı (Birinci işlemde bulunan değer)

n : Denek sayısı

Sınıflanmış Verilerde Standart Sapmanın Hesaplanması

Sınıflanmış verilerde aritmetik ortalamayı hesaplariken yapılan işlemlerin ilk 5 maddesi burada da aynen yapılır. Buna ek olarak yapılacak işlemler şunlardır:

6. Çalışma birimi b'nin karesi alınır ve her sınıfın karşısına yazılır.
7. Her sınıfın frekansı kendi b² değeri ile çarpılır ve fb² kolonuna yazılır.
5. Bulunan değerler formüle konur. Sınıflanmış verilerde standart sapmanın formülü:

örnek;

S	: Standart sapma
C	: Sınıf aralığı
$\sum fb^2$	: Yedinci işlemde bulunan değerlerin toplamı
$(\sum fb)^2$	: Beşinci işlemde bulunan değer
n	: Denek sayısı

$$S = C \sqrt{\frac{\sum fb^2 - \frac{(\sum fb)^2}{n}}{n - 1}}$$

Standart Sapmanın Özellikleri

- Dağılımın yaygınlığını tek başına anlatmaya yetmez, aritmetik ortalama ile beraber söylenirse anlamlıdır.
- Aritmetik ortalamaya göre hem (+) hem (-) olacaktır.
- Birimi incelenen değişkenin birimidir.

Varyans

Varyans standart sapmanın karesidir.

$$v = S_D^2$$

Varyansın Standart Sapmadan Farkı:

1. Kare olduğu için değeri daima (+) dir.
2. Ölçü birimi yoktur.
3. Standart Sapma kadar anlamlı değildir.

Varyasyon Katsayısı (Değişim Katsayısı)

Standart sapma dağılımının yaygınlığını gösteren bir ölçüdür. Ancak standart sapma ile dağılım hakkında çok fazla bir şey söylemek olanaksızdır. Çünkü bulduğumuz standart sapma değeri mutlak bir değer olduğundan büyük müdür, yoksa küçük müdür karar vermemiz mümkün değildir. Örneğin, bir dağılımın standart sapması 6 olsa bu değer büyük müdür yoksa normal midir bir fikir yürütemeyiz. Buna karar verebilmek için varyasyon katsayısını hesaplamamız gerekir. Varyasyon katsayısı standart sapmanın ortalamaya göre yüzde kaçlık bir değişim gösterdiğini belirtir. Aşağıdaki formülle hesaplanır :

$$V = \frac{S}{\bar{X}} \times 100$$

Standart Hata

Aritmetik ortalama standart hata ile birlikte gösterilmelidir. Standart hata şu formülle hesaplanır:

Sınıflanmış ve sınıflanmamış verilerde aynı formül kullanılır.

$$S_{\bar{x}} = \frac{S}{\sqrt{n}}$$

Range (Dağılım Aralığı)

- Merkez ölçüsü olarak aritmetik ortalama kullanılmıyorsa standart sapma ve ona dayalı yaygınlık ölçüleri kullanılamaz.
- Dağılımın merkezinin gösterilmesi için ortancanın kullanılması halinde yaygınlık ölçüsü olarak range kullanılır.
- Range genellikle minimum ve maksimum değerlerle ifade olunur.

Niteliksel Verilerin Tanımlanması

Niteliksel veriler genellikle incelenen niteliğin görülme sıklığı ile ifade edilirler.

p = görülme sıklığı

q = görülmeme sıklığı = $1-p$

$S_D = \frac{pxq}{n}$

$S_x = \frac{pxq}{n}$

Niteliksel veriler genellikle x , S_D ve n ile ifade edilirler. x , S_D ve n biliniyorsa dağılım biliniyor demektir.

Niteliksel veriler ise sadece p ve n ile ifade edilirler. p bilinince diğerleri türetilebilir.

Evren Ortalamasının Güven Sınırları

Evreni incelemek çoğu kez imkansız olduğundan evren ortalamasını hesaplayamayız. Ancak, örnekten elde edilen x ve S_x yardımı ile belirli bir olasılık düzeyinde evren ortalamasının güven sınırlarını saptayabiliriz.

$\mu = X \pm S_x.t$ veya $X - S_x.t < \mu < X + S_x.t$

μ = Evrenin aritmetik ortalaması

t = Belirlenen yanılma düzeyi (α) (α en fazla 0.05 olmalıdır) ve $n-1$ serbestlik derecesinde (SD) t tablosundan bulunan değer.

Evren Oranının Güven Sınırları

Niteliksel verilerde hesaplanır.

p = Örnekteki görülme sıklığı (oranı)

P = Evrendeki görülme sıklığı (oranı)

$P = p \pm \frac{pxq}{n}.t$

Teorik Dağılımlar

İncelenen olaylar çeşitli teorik dağılımlara uygunluk gösterirler. Bir olayın beklenen olası sonuçlarını teorik dağılımlardan yararlanarak tahmin edebiliriz.

Binomiyal Dağılım

Sayımla belirtilen kesikli değişkenlerin dağılımıdır. Bir olayın oluş olasılığı (p) büyük, n küçük olduğunda olasılık hesaplamak için kullanılır. (n) büyüdükçe hesaplamak zorlaşır. “Yığılımlı binomiyal olasılık dağılımı” tabloları kullanılır. (p=q) olduğunda dağılım simetrik olur. P sifıra yaklaştıkça yani oluş olasılığı küçüldükçe binomiyal dağılım Poisson dağılımına yaklaşıp.

$$P(r) = \frac{n!}{(n-r)!r!} \cdot p^r \cdot q^{n-r}$$

n= Toplam olay sayısı

r=İstenen olayın oluş sayısı

n-r = İstenmeyen olayın oluş sayısı

p= İstenen olayın oluş olasılığı

q= İstenmeyen olayın oluş olasılığı

!=Faktöryel

Poisson Dağılımı

Poisson dağılımı sayımla belirtilen kesikli değişkenlerin dağılımıdır. İncelenen olayın görülüş olasılığı (p) küçük, n küçük olduğunda olasılık hesaplamak için kullanılır.

$$P(r) = \frac{X^r \cdot e^{-X}}{r!}$$

r= İstenen olayın oluş sayısı

e^x=Üstel fonksiyonlar tablosundan bulunan değer

Normal Dağılım

Ölçümle belirtilen sürekli değişkenlerin dağılımıdır. µ'nün ve σ'nın her farklı değeri farklı bir dağılım oluşturur.µ'nün farklı değeri dağılım grafiğinin yerini x eksenini boyunca σ'nın farklı değeri ise y eksenini boyunca değiştirmektedir.

Normal Dağılımın Özellikleri

1. Normal dağılım-sonsuzdan +sonsuz kadar devam eden sürekli bir dağılımdır.
2. Dağılım çan eğrisi biçiminde simetrik bir dağılımdır. Ancak her çan eğrisi biçimindeki dağılım normal dağılım değildir.
3. Normal dağılımda X=Ortanca=Tepe değeri
4. Eğriyle x eksenini arasındaki toplam alan bir birim karedir.
5. Verilerin (tüm dağılımın) %68.26'sı X±1 S_D sınırları arasında bulunur.
6. Tüm verilerin %95.44'ü X±2 S_D sınırları arasında bulunur.
7. Tüm verilerin %99.74'ü X±3 S_D sınırları arasında bulunur.
8. Dağılımın tamamı pratik olarak X±4 S_D sınırları arasında kabul edilir.

Standart Normal Dağılım

X ve S_D değiştikçe sonsuz sayıda normal dağılım ortaya çıkar. İncelemede kolaylık sağlamak için X=0 S_D=1 olan bir normal dağılım kullanılır. Buna standart normal dağılım adı verilir. Diğer dağılımlar istendiğinde

$Z = \frac{X_i - X}{S_D}$ formülü ile standart normal dağılıma dönüştürülebilir.

Normal dağılım Kullanılarak Olasılık Hesaplanması

-Z tablosu kullanılır.

-Standart normal dağılım eğrisinden yararlanılarak herhangi bir ampirik dağılımda verilerin belli değerler arasına düşme ihtimalleri hesaplanabilir.

İncelenen örnek dağılışı herhangi bir teorik dağılışa uygun olup olmadığı test edilebilir. Bu testlere genel olarak uyum iyiliği testi adı verilir.

Önemlilik Testleri Hakkında Genel Bilgiler

Önemlilik testleri elde edilen değerlerin ya da varılan sonuçların istatistiksel olarak önem taşıyıp taşımadığını ya da anlamlı olup olmadığını test etmek için başvurulan yöntemlerdir.

Önemlilik Testlerinin Kullanılış Yerleri

1. İki ya da daha fazla grup arasındaki farkın tesadüfi olup olmadığının test edilmesinde kullanılır.
2. İki değişken arasındaki ilişkinin tesadüfi olup olmadığının belirlenmesinde kullanılır.
3. Aynı grupta farklı koşullar altında elde edilen değerler arasındaki farkın tesadüfi olup olmadığının belirlenmesinde kullanılır.
4. Örnekten elde edilen veriler yardımı ile evren parametresinin belli bir değere eşit olup olmadığının test edilmesinde kullanılır.

Önemlilik Testlerinin Seçimi

I-Verinin Ölçüm Biçimi

1. Ölçümle belirtilen veriler: Niceliksel (kantitatif) verilerdir. Sürekli dağılım gösterirler. Sınıflar birbirine geçişlidir ve çoğunlukla normal dağılıma uyarlar.
2. Sayımla belirtilen veriler: Niteliksel (kalitatif) verilerdir. Kesikli dağılım gösterirler. Sınıflar birbirine geçişli değildir. Genellikle Binomiyal ya da Poisson dağılımına uyarlar.

Ölçümle belirtilen verilere parametrik test varsayımları yerine geliyorsa parametrik testler, gelmiyorsa nonparametrik testler uygulanır.

Sayımla belirtilen verilere non-parametrik testler uygulanır.

II-Örneklem Büyüklüğü (Veri Sayısı)

- Gruplardaki denek sayısı arttıkça kullanılan testin gücü ve güvenilirliği artar.
- Gruplardaki denek sayısı az olduğunda (≈ 30 dan az olduğunda) non-parametrik testler kullanılmalıdır.
- İki ya da daha çok grup karşılaştırılıyorsa gruplardaki denek sayılarının eşit olması için gerekli önlem alınmalıdır.

III- İncelenen Grupların Bağımlı ya da Bağımsız Olması

İki veya daha fazla grup karşılaştırılıyorsa gruplardaki veriler birbirinden tamamen bağımsız olmalıdır yani gruplar ayrı bireylerden oluşmalıdır ve bir kişiyi örneğe almamız başkasına bağımlı olmamalıdır.

Bir denek üzerinde birden çok gözlem yapıldığında ya da gözlem sayısı denek sayısını aştığında gruplar bağımlı olur.

İncelenen grupların bağımlı ya da bağımsız olması durumunda uygulanacak önemlilik testi farklı olacaktır.

Test Sonuçlarının Yorumlanması (Değerlendirme)

1.Hipotezler:

Önemlilik testleri bir hipotezi test etmek için yapılır. Bu yüzden bazen önemlilik testi yerine hipotez testi de denmektedir. Hipotez bir ön yargıdır.

- a) H_0 hipotezi: Farksızlık hipotezi, geçersizlik hipotezi, sıfır hipotezi, null hipotez.

Bir testte öne sürülen ve asıl test edilmek istenen hipotezdir. Örnekten elde edilen sonuçların tesadüfi yani geçersiz olduğu temeline dayanır. Gruplar arasında fark arandığında H_0 hipotezi olumsuz olarak belirlenir. Evren parametresinin belli bir değere eşit olup olmadığı test ediliyorsa H_0 hipotezi "Evren parametresi belli bir değere eşittir." diye kurulur.

b) H_1 hipotezi: Alternatif hipotez, H_A hipotezi

H_1 hipotezi H_0 hipotezine ters yönde kurulur. Test sonucunda H_0 kabul edilirse H_1 reddedilir. H_0 reddedilirse H_1 kabul edilir.

Bir hipotez tek yönlü ya da çift yönlü olarak belirtilebilir. Hipotezin tek yönlü ya da çift yönlü olduğunu H_1 hipotezi belirler. Evren parametresi belli bir değere eşit olup olmadığı test ediliyorsa H_1 hipotezi evren parametresi H_0 hipotezinde belirtilen değere eşit değildir (çift yönlü), o değerden küçüktür ya da büyüktür şeklinde (tek yönlü) kurulabilir.

Test sonucu hesapla bulunan değeri karşılaştırmak için kullanılan teorik tablolar tek yönlü ya da çift yönlü hazırlanmış olabilir. Hipotezin tek ya da çift yönlü olması testin sonucunu anlamlı ölçüde etkiler.

3. Yanılma Düzeyi:

Bir hipotez kabul ya da reddedildiğinde iki tip hata ortaya çıkabilir.

Hipotez (H_0)	Karar	
	Kabul Etme	Reddetme
Doğru	Doğru karar	Tip I hata (α)
Yanlış	Tip II hata (β)	Doğru karar

Bu iki hata hiçbir zaman birlikte ortaya çıkmaz.

Tip I hata: Yanılma olasılığı, α denir. Doğru bir hipotezin yanlışlıkla reddedilme olasılığıdır. Tip I hatayı azaltmak için α 'yı küçük seçeriz. α 0,05'den büyük olmamalıdır.

Tip II hata: β denir. β üzerinde bir denetime sahip değiliz. Ancak β 'nın α 'dan büyük olduğunu biliriz.

Test Çeşitleri ve Özellikleri

Önemlilik testleri iki ana gruba ayrılır.

1. Parametrik önemlilik testleri

2. Parametrik olmayan (nonparametrik) önemlilik testleri

-Bir testte; ortalama, varyans, oran gibi ölçüler kullanılıyorsa bu test parametrik bir testtir. Ölçü yerine sıralama, sayma, işaretlem gibi işlemler yapılıyorsa bu test parametrik olmayan bir testtir.

-Niteliksel veriler için parametrik olmayan testler kullanılır. Yine, ölçümle belirtildiği halde veri parametrik test varsayımlarını yerine getiremiyorsa, denek sayısı az ise ya da değerler yerine sıraları verilmişse yine parametrik olmayan testler kullanılır.

-Veri ölçümle belirtilmişse ve parametrik test varsayımlarını yerine getirebiliyorsa parametrik testleri uygulamak daha doğru olur. Çünkü, parametrik testler parametrik olmayan testlerden daha güçlüdür.

Parametrik Test Varsayımları:

1. Örneklemin çekildiği evrenle ilgili:

a) Normal dağılıma sahip olacak

b) Varyanslar homojen olacak

2. Örneklemele ilgili

a) Denekler evrenden rastgele seçilecek

b) Denekler birbirinden bağımsız olarak seçilecek (Bir deneğin seçimi diğer deneklerin seçimini etkilemeyecek)

Evren ile ilgili varsayımların yerine getirilmesi araştırmacının elinde değildir. Örneklemele ilgili varsayımların yerine getirilmesi ise araştırmacının elindedir.

Parametrik olmayan testlerde evrenle ilgili varsayımlar aranmaz. Örneklemele ilgili varsayımlar hem parametrik hem de nonparametrik testler için geçerlidir.

Her bir parametrik teste karşı birden çok parametrik olmayan test vardır.

Parametrik ve Parametrik Olmayan Testlerin Bazı Özelliklere Göre Karşılaştırılması

Özellik		Parametrik	Parametrik Olmayan
Varsayımlar	-Örneklem çekilen evrenle ilgili	1. Normal dağılıma sahip olacak. 2. Varyanslar homojen olacak.	Çoğu kez hiçbir varsayım aranmaz
	-Örneklemeyle ilgili	1. Denekler evrenden rastgele seçilecek. 2. Denekler birbirinden bağımsız olarak seçilecek (Bir denneğin seçimi diğer denneğin seçimini etkilemeyecek).	1. Denekler evrenden rastgele seçilecek. 2. Denekler birbirinden bağımsız olarak seçilecek (Bir denneğin seçimi diğer denneğin seçimini etkilemeyecek).
Varsayımlar Bozulursa	-Örneklem çekilen evrenle ilgili	Test sonucu hatalı olur.	Varsayım aranmadığı için etkilenmez.
Teste Etkisi Ne Olur?	-Örneklemeyle ilgili	Test sonucu hatalı olur.	Test sonucu hatalı olur.
Varsayımların Bozulduğu Anlaşılabilir Mi?		1. Normal dağılıma uygunluk testi ile 2. Varyansların homojenlik testi ile	Verilerde eşit skorların ortaya çıkması ile
Verinin Ölçüm Biçimi		Ölçümle belirtilen karakterler (uzunluk, ağırlık, Hb miktarı, yaş, kolesterol miktarı vb.)	Sayımla belirtilen karakterler (saç rengi, cinsiyet, meslek, bölge vb.) Ayrıca varsayımlar yerine getirilemediğinde ölçümle belirtilen karakterler.
Örneklem Büyüklüğü	$n \leq 10$ ise	Varsayımlar büyük olasılıkla bozulur.	Çabuk ve kolay yapılır.
	$n \geq 30$ ise	Parametrik olmayan testlere üstündür.	Uygulanması güç işlemleri gerektirir.

Parametrik Test

Evren ortalaması önemlilik testi
İki ortalama arasındaki farkın önemlilik testi
İki eş arasındaki farkın önemlilik testi
İki yüzde arasındaki farkın önemlilik testi
Varyans analizi (tek yönlü)

Parametrik Olmayan Test

İşaret testi
Mann-Whitney U testi
Wilcoxon eşleştirilmiş iki örnek testi
4 gözlü ki-kare testi
Kruskal-Wallis varyans analizi

İstatistiksel Karar

Test sonucu H_0 hipotezi ya reddedilir ya kabul edilir. Bu karara hesaplanan değer ile seçilen yanılma düzeyindeki tablo değeri karşılaştırılarak varılır.

(Birkaç test dışında) test sonucunda H_0 hipotezinin kabul ya da reddedilmesinde kriter şudur:

-Test değeri tablo değerinden büyük ise H_0 reddedilir H_1 kabul edilir.

-Test değeri tablo değerinden küçük ise H_0 kabul edilir.

H_0 hipotezi kabul edilirken daima $p > 0.05$ ifadesi kullanılır.

H_0 'ın reddedilmesi ise H_0 'ın reddedildiği en düşük α sınırı ile ifade edilir.

Uygun Test Seçimi İçin Anahtar

Verinin Ölçüm Biçimi

Ölçüm

Karşılaştırılacak Gruplar

Bağımsız

Karşılaştırılacak Grup Sayısı

İki

Üç+

n=?

n=?

<30

30+

<30

30+

Mann-Whitney U testi

İki ortalama arasındaki farkın önemlilik testi

Kruskal-Wallis varyans analizi

Varyans analizi

Bağımlı

Karşılaştırılacak Grup Sayısı

İki

Üç+

n=?

<30

30+

İki ölçümde de tüm denekler incelendi mi?

Evet

Hayır

Wilcoxon eşleştirilmiş iki örnek testi

İşaret testi

İki eş arasındaki farkın önemlilik testi

Evren ortalaması önemlilik testi

Nitelik (Sayım)

Karşılaştırılacak Gruplar

Bağımsız

Karşılaştırılacak Grup Sayısı

İki

Üç+

Beklenen değerler 5'den küçük değilse

2x2 ki-kare ya da iki yüzde arasındaki farkın önemlilik testi

5'den küçükse Fisher kesin ki-kare testi

Ki-kare nx2 2xm nxm

ya da Kolmogorov Smirnov nx2 2xm

Bağımlı

Karşılaştırılacak Grup Sayısı

İki

İki gözlemde de tüm denekler incelendi mi?

Evet

Hayır

Bağımlı gruplarda ki-kare testi

ya da

Bağımlı gruplarda iki yüzde arasındaki farkın önemlilik testi

Normal Dağılıma Uygunluk Testi

Hazırlık İşlemleri:

a) Dağılımın X_{S1} ve S_D si bulunur.

b) Z değerleri bulunur.

$$Z_i = \frac{X_i - X}{S_D}$$

Z_i = i sınıfının Z değeri

X_i = i sınıfının sınıf ara değeri

c) Z değerlerine karşı gelen olasılıklar Z tablosundan bulunur.

Z değeri eksi ise tablodan bulunan değer 0,5 den çıkarılarak olasılık bulunur.

Z değeri artı ise tablodan bulunan değer 0,5 ile toplanarak olasılık bulunur.

d) Beklenen yüzde frekanslar bulunur.

Bir sınıfın beklenen yüzde frekansı, o sınıfın olasılığının bir altındaki sınıfın olasılığından çıkarılmasıyla bulunur.

Son sınıfın beklenen yüzde frekansı ise kendi olasılığının 1 den çıkarılmasıyla bulunur.

Beklenen yüzde frekanslarla toplam denek sayısı çarpılarak beklenen frekanslar bulunur.

Test İşlemleri:

a) H_0 = Normal dağılımdan ayrışık önemsizdir.

b) Test istatistiğinin hesaplanması:

$$\chi^2 = \sum \frac{(G-B)^2}{B}$$

G= Gözlenen frekans

B= Beklenen frekans

Birinci ve sonuncu sınıfların beklenen frekansları 5 den küçük olduğu için bu frekanslar bir alttaki ve bir üstteki sınıfların frekanslarıyla birleştirilerek hesaplama yapılacaktır.

c) Yanılma olasılığı $\alpha = 0,05$ seçilir.

d) Serbestlik derecesi = sınıf sayısı – 3

e) $\alpha = 0,05$ yanılma düzeyinde ve (sınıf sayısı – 3) SD de tablo χ^2 değeri bulunur.

f) Karşılaştırma= Hesapla bulunan χ^2 değeri tablo χ^2 değerinden büyükse H_0 hipotezi reddedilir, küçükse kabul edilir.

g) Karar verilir.

Elde Edilen Dağılımın Normal Dağılıma Uyup Uymadığının Pratik Olarak İncelenmesi

Normal dağılımda deneklerin % 68,26 sı $X \pm 1 S_D$ sınırları içine

% 95,44 ü $X \pm 2 S_D$ sınırları içine

% 99,74 ü $X \pm 3 S_D$ sınırları içine girmesi gerekir.

Bu nedenle dağılımın ortalama ve standart sapması hesaplanarak deneklerin yaklaşık olarak bu sınırlar içinde kalıp kalmadığı kontrol edilebilir.

Örnek: Bir dağılımın $X=103,72$ $S_D= 6,35$ dir.

Buna göre bu sınırlar içine giren denek sayısı ve yüzdeleri:

	Sınırlar	Denek Sayısı	Yüzde
$X \pm 1 S_D$	97,34 – 110,07	49	65,33
$X \pm 2 S_D$	91,02 – 116,42	73	97,33
$X \pm 3 S_D$	84,67 – 122,77	75	100,0

İncelenen dağılımdan elde edilen yüzdeler ile normal dağılımda olması gereken yüzdeler birbirine çok yakındır. Bu nedenle dağılımın normal dağılıma uyduğu kabul edilebilir.

VARYANSLARIN HOMOJENLİK TESTİ

İki Grup Olduğunda Varyansların Homojenlik Testi

1. H_0 : Varyanslar homojendir.
2. Test istatistiğinin hesaplanması:

$$F = \frac{\text{Büyük Varyans}}{\text{Küçük Varyans}}$$

3. Yanılma olasılığı seçilir ($\alpha = 0.05$).
4. n_1-1 ve n_2-1 olmak üzere 2 serbestlik derecesi vardır.
5. $\alpha = 0.05$ yanılma düzeyinde n_1-1 ve n_2-1 serbestlik derecelerinde F tablosundan tablo F değerleri bulunur. F değeri bulunurken büyük varyansın serbestlik derecesine soldan sağa doğru, küçük varyansın serbestlik derecesine yukardan aşağı doğru bakılır.
6. Karşılaştırma: Hesapla bulunan F değeri tablo F değerinden büyükse H_0 hipotezi reddedilir, küçükse kabul edilir.
7. Karar: Bu iki dağılımın varyansları homojendir (homojen değildir).

İkiden Çok Grup Olduğunda Varyansların Homojenlik Testi

Gruplardaki Denek Sayıları Eşit Olduğunda Test İşlemleri

1. Hazırlık İşlemleri:

- a. Her grubun varyansı hesaplanır.
- b. Varyansların tabii logaritması (e tabanına göre logaritması) alınır.
- c. Varyanslar toplanır.
- d. Varyansların tabii logaritmaları toplanır.
- e. Ortalama varyans ($\bar{S}^2$) hesaplanır.

$$\bar{S}^2 = \frac{\sum S_i^2}{k} \quad k: \text{grup sayısı}$$

- f. Ortalama varyansın tabii logaritması alınır. $\ln \bar{S}^2$
- g. M değeri hesaplanır.
 $M = (n-1)(k \times \ln \bar{S}^2 - \sum \ln S_i^2)$
- h. C değeri hesaplanır.

$$C = 1 + \frac{k+1}{3 \times k \times (n-1)}$$

2. Test İşlemleri:

- a. H_0 : Varyanslar homojendir.
- b. Test istatistiğinin hesaplanması:

$$\chi^2 = \frac{M}{C}$$

- c. Serbestlik derecesi = Grup sayısı – 1
- d. Yanılma olasılığı seçilir ($\alpha = 0.05$).
- e. $\alpha = 0.05$ yanılma düzeyinde (Grup sayısı – 1) serbestlik derecesinde χ^2 tablosundan tablo χ^2 değeri bulunur.
- f. Karşılaştırma: Hesapla bulunan χ^2 değeri tablo χ^2 değerinden büyükse H_0 hipotezi reddedilir, küçükse kabul edilir.
- g. Karar: Bu üç grubun varyansları homojendir (homojen değildir).

Gruplardaki Denek Sayıları Eşit Olmadığında Test İşlemleri

1. Hazırlık İşlemleri:

- Her grubun varyansı hesaplanır.
- Varyansların tabii logaritması (e tabanına göre logaritması) alınır.
- Her grubun serbestlik derecesi (m_j) bulunur.
$$m_j = n_j - 1$$
Serbestlik dereceleri toplanarak $\sum m_j$ bulunur.
- Her grubun varyansı ile serbestlik derecesi çarpılır ve toplanır
 $(\sum m_j S_j^2)$.
- Her grubun varyansının tabii logaritması ile serbestlik derecesi çarpılır ve toplanır $(\sum m_j \ln S_j^2)$.
- Her grubun serbestlik derecesinin tersi (resiprokal) alınır ve toplanır.
$$(\sum \frac{1}{m_j})$$
- Ortalama varyans ($\bar{S}^2$) bulunur:
$$\bar{S}^2 = \sum m_j S_j^2 / \sum m_j$$
- Ortalama varyansın tabii logaritması alınır:
$$\ln \bar{S}^2$$
- M değeri hesaplanır:
$$M = (\sum m_j) \ln \bar{S}^2 - \sum m_j \ln S_j^2$$
- C değeri hesaplanır.
$$C = 1 + [1/(3(k-1))] (\sum 1/m_j - 1/\sum m_j)$$

2. Test İşlemleri:

- H_0 : 4 grubun varyansları homojendir.
- Test istatistiğinin hesaplanması:
$$x^2 = \frac{M}{C}$$
- Serbestlik derecesi = Grup sayısı – 1
- Yanılma olasılığı seçilir ($\alpha = 0.05$).
- $\alpha = 0.05$ yanılma düzeyinde (Grup sayısı – 1) serbestlik derecesinde x^2 tablosundan tablo x^2 değeri bulunur.
- Karşılaştırma: Hesapla bulunan x^2 değeri tablo x^2 değerinden büyükse H_0 hipotezi reddedilir, küçükse kabul edilir.
- Karar: Bu dört grubun varyansları homojendir (homojen değildir).

PARAMETRİK ÖNEMLİLİK TESTLERİ

İki Ortalama Arasındaki Farkın önemlilik Testi (Student's t Testi)

- Parametrik test varsayımları yerine getirildiğinde, ölçümle belirtilen sürekli bir değişken yönünden bağımsız iki grup arasında fark olup olmadığını test etmek için kullanılır.

Student's t Testinin Doğru Kullanılması İçin Gerekli Koşullar

- Önemlilik testleri ile ilgili genel noktalara uyulmalıdır.
- Bu testle ilgili özel noktalara uyulmalıdır:
 - Bu testte iki grubun aritmetik ortalamaları karşılaştırılmaktadır. Bu nedenle aşırı değerlerin Xya yapacağı olumsuz etkiler göz önünde bulundurulmalıdır.
 - Parametrik test varsayımları yerine getirilmelidir.
 - Gruplar birbirinden bağımsız olmalıdır.
 - Veri ölçümle belirtilen sürekli bir karakter olmalıdır. Ayrıca, örneklem büyüklüğü (n) yeterli olduğunda sayısal olarak belirtilen (ölen, doğan, hastalanan, yaşayan sayısı gibi) sürekli olmayan değişkenlere de uygulanabilir. Niteliksel verilere uygulanamaz.
 - Her gruptaki denek sayısı 30'dan az olmamalıdır.
 - Her iki gruptaki denek sayısı birbirine eşit ya da çok yakın olmalıdır.
- Parametrik test varsayımları yerine getirilemiyorsa bu test yerine Mann-Whitney U testi kullanılmalıdır.

Test işlemleri: Önce her iki dağılımın normal dağılıma uyup uymadığı test edilir. Her ikisi de normal dağılıma uyuyorsa varyanslarının homojen olup olmadığı test edilir.

Varyanslar Homojen Olduğunda Test İşlemleri

$$H_0 = X_1 - X_2 = 0$$

$$H_1 = X_1 - X_2 \neq 0$$

$$t = \frac{X_1 - X_2}{\frac{\sqrt{V_0 + V_0}}{n_1 + n_2}}$$

$$V_0 = \text{Ortak varyans}$$

$$V_0 = \frac{(n_1 - 1)V_1 + (n_2 - 1)V_2}{n_1 + n_2 - 2}$$

Yanılma olasılığı seçilir ($\alpha = 0.05$).

Serbestlik derecesi = $n_1 + n_2 - 2$

$\alpha = 0.05$ yanılma düzeyinde $n_1 + n_2 - 2$ serbestlik derecesinde t tablosundan tablo t değeri bulunur.

Karşılaştırma: Hesapla bulunan t değeri tablo t değerinden büyükse H_0 hipotezi reddedilir, küçükse kabul edilir.

Varyanslar Homojen Olmadığında Test İşlemleri

Gruplardaki Denek Sayısı Eşit İse:

$$t = \frac{X_1 - X_2}{\frac{\sqrt{V_1 + V_2}}{n_1 + n_2}}$$

$$SD = n - 1$$

$$\frac{V_1 + V_2}{n_1 + n_2}$$

Gruplardaki Denek Sayısı Eşit Değil İse:

$$t = \frac{X_1 - X_2}{\frac{\sqrt{V_1 + V_2}}{n_1 + n_2}}$$

$$\frac{V_1 + V_2}{n_1 + n_2}$$

Gruplardaki denek sayıları eşit olmadığı için $n_1 - 1$ ve $n_2 - 1$ SD'lerindeki tablo t değerleri kullanılarak teorik t değerleri hesaplanır.

$t_1 = n_1 - 1$ SD'deki tablo t değeri

$t_2 = n_2 - 1$ SD'deki tablo t değeri

$$t_1 = \frac{(V_1/n_1)t_1 + (V_2/n_2)t_2}{\frac{V_1 + V_2}{n_1 + n_2}}$$

$$\frac{V_1 + V_2}{n_1 + n_2}$$

İki Eş Arasındaki Farkın Önemlilik Testi (Paired T Testi)

Parametrik test varsayımları yerine getirildiğinde, ölçümle belirtilen sürekli bir değişken yönünden iki eş grubu karşılaştırmak için bu test kullanılır (bağımlı verilerde).

- Veri ölçümle belirtilmiş olacak.
- Gruplar bağımlı olacak:
 - Aynı birey üzerinde aynı konuda iki kez ölçüm yapılacak
 - Tek yumurta ikizleri
 - Suni eşleme (birebir eşleme)

Doğru Kullanım için Gerekli Koşullar:

- Önemlilik testleri ile ilgili genel noktalara dikkat edilmelidir.
- Bu test ile ilgili özel noktalara dikkat edilmelidir.
 - İki eş arasındaki farkın önemlilik testi iki ortalama arasındaki farkın önemlilik testi ile karıştırılmamalıdır. İki eş arasındaki farkın önemlilik testinde farkların ortalamasının, iki ortalama arasındaki farkın önemlilik testinde ise iki ortalamanın farkının istatistiksel olarak anlamlı olup olmadığı test edilir.
 - Parametrik testlerle ilgili varsayımlar yerine getirilmelidir. Varyansların homojenlik koşulu bu test için geçerli değildir.
 - Gruplar bağımlıdır.
 - Veri ölçümle belirtilen sürekli bir karakter olmalıdır. Ayrıca, örneklem büyüklüğü (n) yeterli olduğunda sayısal olarak belirtilen sürekli olmayan karakterler de kullanılabilir. Niteliksel verilere uygulanamaz.
 - Denek sayısı 30'dan az olmamalıdır.
- Parametrik test varsayımları yerine getirilemiyorsa bu test yerine 'Wilcoxon Eşleştirilmiş İki Örnek Testi' kullanılmalıdır.
- Aynı birey üzerinde aynı konuda ikiden çok ölçüm yapıldığında, ölçümler arası farklılık 'Tekrarlı Ölçümlerde Varyans Analizi' yöntemiyle aranmalıdır.

Test İşlemleri:

1. Hazırlık İşlemleri:

- Bireylerin önceki ve sonraki değerleri arasındaki farklar ($D = \text{önce} - \text{sonra}$) bulunur ve işaretleri dikkate alınarak toplanır ($\sum D$).
- Farkların kareleri (D^2) alınır ve toplanır ($\sum D^2$).
- Farkların ortalaması bulunur:

$$\bar{D} = \frac{\sum D}{n}$$

- Farkların ortalamasının standart sapması bulunur.

$$S = \sqrt{\frac{\sum D^2 - \frac{(\sum D)^2}{n}}{n - 1}}$$

- Standart hata bulunur:

$$S_{\bar{D}} = \frac{S}{\sqrt{n}}$$

2. Test İşlemleri:

- Veri ölçümle belirtilmiştir, parametrik test varsayımları yerine gelmektedir.
- H_0 : Önceki ve sonraki değerler farksızdır.
- H_1 : Sonraki değerler daha yüksektir (tek yönlü hipotez).
- Test istatistiğinin hesaplanması:

$$t = \frac{\bar{D}}{S_D}$$

- e. Yanılma olasılığı seçilir ($\alpha = 0.05$).
- f. Serbestlik derecesi = $n - 1$
- g. $\alpha = 0.05$ yanılma düzeyinde ($n - 1$) serbestlik derecesinde tek yönlü t tablosundan tablo t değeri bulunur.
- h. Karşılaştırma: Hesapla bulunan t değeri tablo t değerinden büyükse H_0 hipotezi reddedilir, küçükse kabul edilir.
- i. Karar: ... verilmesi ... değerlerini yükseltmiştir ($t = \dots p < 0.05$) veya tam tersi.

İKİ YÜZDE ARASINDAKİ FARKIN ÖNEMLİLİK TESTİ

Niteliksel bir değişken yönünden iki gruptan elde edilen yüzdelerin farklı olup olmadığını test etmek için kullanılan bir önemlilik testidir. Bu test bağımlı ve bağımsız gruplarda farklı şekilde uygulanır.

Bağımsız Gruplarda İki Yüzde Arasındaki Farkın Önemlilik Testi

Şu noktalara dikkat etmek gerekir:

- a. İki grup karşılaştırılmaktadır.
- b. Karşılaştırılan bu iki grup birbirinden bağımsızdır.
- c. Bu iki grup arasında farklı olup olmadığı test edilen değişken aslında sayımla belirtilen niteliksel bir karakterdir. Ör: var-yok, iyileşti-iyileşmedi gibi. Niteliksel bu veri sonradan yüzdeye dönüştürülerek işlem yapılmaktadır.

Test İşlemleri:

- 1. Veri sayımla belirtilen niteliksel bir değişkendir. Sonradan yüzde olarak ölçüme dönüştürülmüştür.
- 2. H_0 :yüzdeleri farksızdır ($p_1 = p_2 = 0$)
 H_1 : $p_1 = p_2 \neq 0$
- 3. Test istatistiğinin hesaplanması:

$$t = \frac{p_1 - p_2}{S_d} \quad S_d = \sqrt{\frac{pq}{n_1} + \frac{pq}{n_2}}$$

- 4. Yanılma olasılığı seçilir ($\alpha = 0.05$).
- 5. Serbestlik derecesi = $n_1 + n_2 - 2$
- 6. $\alpha = 0.05$ yanılma düzeyinde ($n_1 + n_2 - 2$) serbestlik derecesinde t tablosundan tablo t değeri bulunur.
- 7. Karşılaştırma: Hesapla bulunan t değeri tablo t değerinden büyükse H_0 hipotezi reddedilir, küçükse kabul edilir.
- 8.yüzdeleri arasında istatistiksel olarak anlamlı bir fark bulunmuştur.
($t = \dots p < 0.05$) veya tam tersi.

Bağımlı Gruplarda İki Yüzde Arasındaki Farkın Önemlilik Testi

Şu noktalara dikkat etmek gerekir:

- Karşılaştırılan bu iki grup birbirine bağımlıdır.
- Bu iki grup arasında farklı olup olmadığı test edilen değişken aslında sayımla belirtilen niteliksel bir karakterdir. Ör: var-yok, iyileşti-iyileşmedi gibi. Niteliksel bu veri sonradan yüzdeye dönüştürülerek işlem yapılmaktadır.

1. Hazırlık İşlemleri:

Genel toplama göre her göze düşen yüzdeler bulunur.

2. Test İşlemleri:

- Veri sayımla belirtilen niteliksel bir değişkendir. Sonradan yüzde olarak ölçüme dönüştürülmüştür.
- H_0 :yüzdeleri farksızdır ($p_1 = p_2 = 0$)
 H_1 : $p_1 = p_2 \neq 0$
- Test istatistiğinin hesaplanması:

$$t = \frac{p_1 - p_2}{S_d} \quad S_d = \sqrt{\frac{b+c}{n}}$$

Bağımlı örnekler tablosu:

TeknikII			
Teknik I	+	-	Toplam
+	a	b	P_1
-	c	d	
Toplam	P_2		n

- Yanılma olasılığı seçilir ($\alpha = 0.05$).
- Serbestlik derecesi = $n - 2$
- $\alpha = 0.05$ yanılma düzeyinde ($n - 2$) serbestlik derecesinde t tablosundan tablo t değeri bulunur.
- Karşılaştırma: Hesapla bulunan t değeri tablo t değerinden büyükse H_0 hipotezi reddedilir, küçükse kabul edilir.
-yüzdeleri arasında istatistiksel olarak anlamlı bir fark bulunmuştur.
($t = p < 0.05$) veya tam tersi.

EVREN ORTALAMASI ÖNEMLİLİK TESTİ

1. Bir evrende herhangi bir olayın ortalamasının hipotezsel olarak düşünülen bir değere eşit olup olmadığının test edilmesinde kullanılır (Evrenden çekilen bir örneklemin çekildiği evreni temsil edip etmediğini test etmede kullanılır).

2. Bir evrenden ya da bir örneklemden daha önce elde edilen ortalama ile bu evrenin ya da örneklemin bir kısmı üzerinde aynı konuda ikinci kez elde edilen ortalamanın farklı olup olmadığının test edilmesinde kullanılır.

Doğru Kullanım için Gerekli Koşullar:

- Önemlilik testleri ile ilgili genel noktalara dikkat edilmelidir.
- Bu test ile ilgili özel noktalara dikkat edilmelidir.
 - Evren ortalaması önemlilik testi iki ortalama arasındaki farkın önemlilik testi ile karıştırılmamalıdır. İki ortalama arasındaki farkın önemlilik testi iki bağımsız gruptan elde edilen iki ortalamanın farklı olup olmadığının test edilmesinde kullanılır. Oysa özellikle ikinci durumda gruplar bağımsız değildir. Aynı bireylerin bir kısmı yeniden incelenmektedir.
 - Örnekleme ortalamasına aşırı değerlerin yapacağı olumsuz etkiler gözönünde bulundurulmalıdır.
 - Parametrik testlerle ilgili varsayımlar yerine getirilmelidir. Tek grup olduğu için varyansların homojenlik koşulu bu test için geçerli değildir.
 - Gruplar bağımlıdır.
 - Veri ölçümle belirtilen sürekli bir karakter olmalıdır. Ayrıca, örneklem büyüklüğü (n) yeterli olduğunda sayısal olarak belirtilen sürekli olmayan karakterler de kullanılabilir. Niteliksel verilere uygulanamaz.
 - Denek sayısı 30'dan az olmamalıdır.

3. Parametrik test varsayımları yerine getirilemiyorsa bu test yerine ‘İşaret Testi’ kullanılmalıdır.

Test İşlemleri:

1. Veri ölçümle belirtilen bir karakter mi, örneklem büyüklüğü yeterli mi, örneklemdeki denekler rastgele ve birbirinden bağımsız olarak seçilmiş mi, dağılım normal dağılıma uymakta mı test edilmelidir.

2. $H_0: \mu =$

$H_1: \mu \neq$

3. Test istatistiğinin hesaplanması:

$$t = \frac{\bar{X} - \mu}{S_x}$$

4. Yanılma olasılığı seçilir ($\alpha = 0.05$).

5. Serbestlik derecesi = $n - 1$

6. $\alpha = 0.05$ yanılma düzeyinde ($n - 1$) serbestlik derecesinde t tablosundan tablo t değeri bulunur.

7. Karşılaştırma: Hesapla bulunan t değeri tablo t değerinden büyükse H_0 hipotezi reddedilir, küçükse kabul edilir.

8. Karar:ortalaması.....dan farklıdır. ($t =$ p<0.05) veya tam tersi.

EVREN ORANI ÖNEMLİLİK TESTİ

1. Bir evrende herhangi bir olayın oranının, hızının ya da görülüş sıklığının hipotezsel olarak düşünülen bir değere eşit olup olmadığının test edilmesinde kullanılır (Evrenden çekilen bir örneklemin çekildiği evreni temsil edip etmediğini test etmede kullanılır).

2. Bir evrenden ya da bir örneklemden daha önce elde edilen oran ile bu evrenin ya da örneklemin bir kısmı üzerinde aynı konuda ikinci kez elde edilen oranın farklı olup olmadığının test edilmesinde kullanılır.

- Evren oranı önemlilik testi bağımsız gruplarda iki yüzde arasındaki farkın önemlilik testi ile karıştırılmamalıdır. Bağımsız gruplarda iki yüzde arasındaki farkın önemlilik testi iki bağımsız gruptan elde edilen iki yüzdenin farklı olup olmadığının test edilmesinde kullanılır. Oysa özellikle ikinci durumda gruplar bağımsız değildir. Aynı bireylerin bir kısmı yeniden incelenmektedir.

Test İşlemleri:

1. Veri niteliksel olarak belirtilmiştir. Ancak yüzdeye dönüştürüldüğü için parametrik bir test uygulanacaktır.

2. $H_0: P =$

$H_1: P >$

3. Test istatistiğinin hesaplanması:

$$t = \frac{p - P}{\sqrt{\frac{pq}{n}}}$$

4. Yanılma olasılığı seçilir ($\alpha = 0.05$).

5. Serbestlik derecesi = $n - 1$

6. $\alpha = 0.05$ yanılma düzeyinde ($n - 1$) serbestlik derecesinde t tablosundan tablo t değeri bulunur.

7. Karşılaştırma: Hesapla bulunan t değeri tablo t değerinden büyükse H_0 hipotezi reddedilir, küçükse kabul edilir.

8. Karar: Evren oranıdan farklı değildir ($t =$ p>0.05) veya tam tersi.

VARYANS ANALİZİ

Parametrik test varsayımları yerine getirildiğinde ölçümle belirtilen bir değişken yönünden ikiden çok bağımsız grup arasında fark olup olmadığını test etmek için kullanılan bir yöntemdir.

1. Varyans analizi tek yönlü, iki yönlü ya da daha çok yönlü olarak uygulanabilir.
2. İki ve daha çok yönlü varyans analizinde değişkenler arası etkileşim (interaksiyon) incelenebilir.
3. Grupların farklı ya da farksız bulunması durumunda verinin böyle bir karara varılmada yeterli olup olmadığı, başka bir ifade ile ‘deneysel hata’ araştırılabilir.
4. Çeşitli varyasyon kaynaklarının ‘toplam varyasyona’ yaptıkları katkı gösterilebilir.
5. Varyans analizi ‘bağımsız gruplarda’ uygulanır. Ancak, bağımlı gruplarda yani, aynı bireylerde aynı konuda değişik zaman ya da durumlarda yapılan ikiden çok sayıda ölçümler için de ‘Tekrarlı Ölçümlerde Varyans Analizi’ uygulanır.

Doğru Kullanım için Gerekli Koşullar:

1. Önemlilik testleri ile ilgili genel noktalara dikkat edilmelidir.
2. Bu yöntem ile ilgili özel noktalara dikkat edilmelidir.
 - a. Deney düzenlenirken karşılaştırılacak gruplardaki bireylerin birbirine benzer olmasına dikkat edilmelidir. Yani, karşılaştırılacak gruplar homojen olmalıdır ki, gruplar arasında bir farklılık ortaya çıkarsa bu fark uygulanan etkene bağlanabilsin.
 - b. Parametrik testlerle ilgili varsayımlar yerine getirilmelidir. Örneklem çekilen evrenle ilgili varsayımlar yerine getirilemediğinde elde edilen verilere dönüşüm (transformasyon) uygulanarak varsayımlar yerine getirilmeye çalışılır.
 - c. Gruplar birbirinden bağımsız olmalıdır..
 - d. Veri ölçümle belirtilen sürekli bir karakter olmalıdır. Ayrıca, örneklem büyüklüğü (n) yeterli olduğunda sayısal olarak belirtilen sürekli olmayan karakterler de kullanılabilir. Niteliksel verilere uygulanamaz.
 - e. Gruplardaki denek sayısı 30’dan az olmamalıdır
 - f. Gruplardaki denek sayıları birbirine eşit ya da çok yakın olmalıdır.
 - g. Analizin tek yönlü mü, çift yönlü mü olduğu dikkate alınarak uygun yöntem seçilmelidir.

Varyans Analizi İle İlgili Bazı Terimlerin Tanımlanması

Genel Ortalama: İncelene bütün bireylerin aldıkları değerlerin toplanıp incelenen birey sayısına bölünmesiyle elde edilir.

Grup Ortalaması: İncelenen grupların ayrı ayrı ortalamalarıdır. Her gruptaki bireylerin aldıkları değerler toplanıp o gruptaki birey sayısına bölünerek elde edilir.

GnKT: Genel kareler toplamıdır. İncelenen bütün bireylerin aldıkları değerlerin genel ortalamadan farklarının kareleri toplanarak elde edilir.

GAKT: Gruplararası kareler toplamıdır. Her grubun ortalamasının genel ortalamadan farklarının kareleri toplanarak elde edilir.

GİKT: Grup içi kareler toplamıdır. Her bireyin değerinin içinde bulunduğu grubun ortalamasından farklarının kareleri toplanarak elde edilir.

Tek Yönlü Varyans Analizi

Parametrik test varsayımları yerine getirildiğinde, ölçümle belirtilen bir değişken yönünden ikiden çok bağımsız grup arasında fark olup olmadığını test etmek için kullanılan bir yöntemdir.

1. İki grup karşılaştırılmaktadır.
2. Karşılaştırılacak gruplar birbirinden bağımsızdır.
4. Gruplar arasında fark olup olmadığı test edilen değişken ölçümle belirtilen sürekli bir değişkendir.

Test İşlemleri

- a) Veri ölçümle belirtilmiştir. Parametrik test varsayımlarının yerine getirildiği kabul edilmiştir.
- b) H_0 : Gruplar arasında fark yoktur.
 H_1 : En az bir grup diğerlerinden farklıdır.
- c) Test istatistiğinin hesaplanması:
Varyans analizinde F istatistiği kullanılır.
$$F = \frac{GAKO}{GİKO}$$
- d) Yanılma olasılığı olarak $\alpha=0.05$ seçilmiştir.
- e) Varyans analizinde iki serbestlik derecesi vardır. GASD ve GİSD
- f) Tablo değerinin bulunması: F formülünde iki değer vardır (GAKO ve GİKO). Bu değerlerden büyük olanın serbestlik derecesine yukardan aşağı doğru bakılır ve iki serbestlik derecesinin kesiştiği yerdeki değer tablo F değeridir.
- g) Karşılaştırma: Hesapla bulunan F değeri tablo F değerinden büyükse H_0 hipotezi reddedilir, küçükse kabul edilir.
- h) Karar: açısından gruplar farklı bulunmuştur (veya tersi).

İKİ YÖNLÜ VARYANS ANALİZİ

Tek yönlü varyans analizinde ölçümle belirtilen bir değişken yönünden ikiden çok bağımsız grup arasında fark olup olmadığı araştırılır. İki yönlü varyans analizinde ise, ölçümle belirtilen bir değişkene (kolay anlaşılması için bu değişkene bağımlı değişken diyelim) bağımsız iki değişkenin birlikte yaptıkları etki incelenir.

Ör: Gaziantep Ün. Tıp Fak. 1. sınıfa kaydolan öğrencilerin ÖSS puanlarını etkileyen bir faktör mezun oldukları lisenin çeşidi, diğer bir faktör sınavdan önce dershaneye devam edip etmemesi ise uygulanır.

Varyans analizinde H_0 hipotezi reddedildiğinde yani gruplar farklı bulunduğunda farklılığı hangi grubun ortaya çıkardığı incelenmelidir. Bu amaç için geliştirilen yöntemlerden en çok kullanılanlar şunlardır:

- 1) En küçük önemli fark yöntemi
- 2) Tukey yöntemi
- 3) Dunnett yöntemi

PARAMETRİK OLMAYAN ÖNEMLİLİK TESTLERİ

Kİ-KARE TESTİ

Kullanıldığı Yerler

- İki ya da daha çok grup arasında fark olup olmadığının testinde
- İki değişken arasında bağ olup olmadığının testinde
- Gruplar arası homojenlik testinde
- Örneklemden elde edilen dağılımın istenen herhangi bir teorik dağılıma uyup uymadığının testinde (uyum iyiliği testinde) kullanılır.

Uygulandığı Düzenler

- Dört gözlü düzen (2 x 2 düzeni): İncelenen değişkenlerin ikisi de iki nitelik halinde sınıflandırılır ve ikili çapraz tablo haline getirilirse bu isim verilir.
- Çok gözlü düzenler (2 x m, n x 2, n x m düzeni)
- Tek değişkenli düzen: İncelenen bir değişken belirli bir kritere göre sınıflandırılır ve marjinal tablo haline getirilirse bu isim verilir.

Varsayımlar

- Gruplar birbirinden bağımsız olmalıdır.
Beklenen frekanslardan herhangi biri 5’den küçük ise:
a) 2 x 2 düzeninde: Fisher kesin Ki-kare testi uygulanır.
b) 2 x m ya da n x 2 düzeninde: Satır ve kolonlar birleştirilerek 5’den küçük değerlerin ortadan kaldırılmasına çalışılır. Bu mümkün olmazsa Kolmogorov-Smirnov testi uygulanır.
c) n x m düzeninde: Satır ve kolonlar birleştirilerek 5’den küçük değerlerin ortadan kaldırılması daha uygun olur.

4 Gözlü (2 x 2) Düzende Ki-Kare Testi

- Ör: Bir ilaç firması A hastalığına karşı yeni bir ilaç bulmuştur. Bir kısım hastayı bu yeni ilaçla, bir kısım hastayı da eski ilaçla tedavi altına alarak kendi ilacının etkinliğini araştırmıştır.
- H_0 : İyileştirme yönünden eski ilaçla yeni ilaç arasında fark yoktur.
- Beklenen frekansların bulunması: Beklenen frekanslar orantı yolu ile bulunur. Orantı ‘Tedavi edilen toplam 88 hastadan 32 hasta iyileşirse, yeni ilaçlarla tedavi edilen 48 hastadan kaç hastanın iyileşmesi beklenir?’ diye kurulur. Diğer gözlerdeki beklenen frekanslar da benzer orantılarla bulunabilir. Ancak, ki-kare düzenlerinde beklenen frekansların satır ve kolon toplamaları, gözlenen frekansların satır ve kolon toplamalarına eşit olur. Bu yüzden 4 gözlü düzende bir gözün beklenen frekansının bulunması diğer gözlerin beklenen frekanslarının bulunmasını sağlar.

Gözlenen ve Beklenen Frekansların Tabloda Gösterilmesi

İlaç	İyileşen	İyileşmeyen	Toplam
Yeni	21 (17.5)	27 (30.5)	48
Eski	11 (14.5)	29 (25.5)	40
Toplam	32	56	88

$$\chi^2 = \sum \frac{(G - B)^2}{B} = 2.427$$

- Yanılma olasılığı olarak $\alpha = 0.05$ seçilmiştir.
- Serbestlik derecesi = (satır sayısı - 1)(kolon sayısı - 1)
- $\alpha = 0.05$ yanılma düzeyinde (satır sayısı - 1)(kolon sayısı - 1) SD’de tablo χ^2 değeri bulunur (3.841).
- Karşılaştırma: Hesapla bulunan χ^2 değeri tablo χ^2 değerinden büyükse H_0 hipotezi reddedilir, küçükse kabul edilir. H_0 hipotezi kabul edilir.
- Karar: İyileştirme yönünden eski ilaçla yeni ilaç arasında fark bulunamamıştır ($\chi^2 = 2.427$, $p > 0.05$)

YATES DÜZELTMESİ

- Dört gözlü düzenlerde genellikle Yates düzeltmesi yapılır. Çok gözlü düzenlerde Yates düzeltmesi yapılmaz.
- Gözlerde 25’den küçük gözlenen frekans varsa Yates düzeltmesi mutlaka yapılmalıdır.

$$\chi^2 = \sum \frac{[|G - B| - 0.5]^2}{B}$$

FİŞHER KESİN Kİ-KARE TESTİ

- ω 4 gözlü düzende gözlerden herhangi birisinde beklenen frekans 5'den küçükse Fisher kesin ki-kare testi kullanılır.
- ω Ör: Aynı hastalığa karşı değişik iki yöntem uygulayan bir hekim aşağıdaki bulguları elde etmiştir. Tedavi yönünden yöntemler arasında fark var mıdır?

<u>Yöntem</u>	<u>İyileşen</u>	<u>İyileşmeyen</u>	<u>Toplam</u>
I	8 (a)	2 (b)	10(A)
II	4 (c)	14 (d)	18(B)
Toplam	12(C)	16 (D)	28(n)

* Tablo incelenerek gözlerdeki en küçük frekans seçilir ve bu frekans sıfır oluncaya kadar her seferinde 1 eksilterek yeni tablolar yapılır.

* Bu işlem yapılırken satır ve kolon toplamaları değiştirilmez. En küçük frekans her seferinde 1 eksiltirirken diğer gözlerdeki frekansların bazıları 1 eksiltir, bazıları 1 arttırılır. Böylece biri ana tablo olmak üzere 'en küçük frekans+1' tane tablo elde edilir.

* Örneğimizde en küçük frekans, b gözündeki 2'dir. Bu nedenle ana tablo dahil 3 tablo yapmamız gerekmektedir.

Birinci tablo: Ana tablo aynen alınır. İkinci tablo:

<u>Yöntem</u>	<u>İyileşen</u>	<u>İyileşmeyen</u>	<u>Toplam</u>
I	9 (a)	1 (b)	10 (A)
II	3 (c)	15 (d)	18 (B)
Toplam	12(C)	16 (D)	28 (n)

Üçüncü tablo:

<u>Yöntem</u>	<u>İyileşen</u>	<u>İyileşmeyen</u>	<u>Toplam</u>
I	10 (a)	0 (b)	10 (A)
II	2 (c)	16 (d)	18 (B)
Toplam	12(C)	16 (D)	28 (n)

Test İşlemleri

- ω Ho: Tedavi etme yönünden iki yöntem arasında fark yoktur.
- ω Yanılma olasılığı olarak $\alpha=0.05$ seçilmiştir.
- ω Kesin ki-kare testinde hesapla bulunan değer ki-kare değeri değil, direkt olarak α ile karşılaştırılacak olasılık değeri p'dir.
- ω $p = \frac{\sum A! B! C! D!}{a! b! c! d!}$
- ω Karşılaştırma: Hesapla bulunan olasılık değeri saptanan yanılma olasılığından küçükse H_0 hipotezi reddedilir, büyükse kabul edilir. H_0 hipotezi reddedilecektir.
- ω Karar: Tedavi etme yönünden iki yöntem farklıdır. Birinci yöntemde iyileşen yüzdesi (% 80), ikinci yöntemdekinden (% 22.2) daha büyük olduğu için birinci yöntem daha etkindir ($p<0.05$).

ÇOK GÖZLÜ DÜZENLERDE Kİ-KARE TESTİ

- ω 4 Gözlü düzenlerde ki-kare testi için yapılan işlemler çok gözlü düzenlere de aynen uygulanır. Ki-kare değerleri her satır için ayrı ayrı hesaplanır ve sonra bunlar toplanarak toplam ki-kare değeri elde edilir.
- ω Çok gözlü düzenlerde test sonucu H_0 hipotezi kabul edildiğinde 'Gruplar arasında fark bulunamamıştır' kararına varılarak analiz bitirilir.
- ω Ancak, test sonucu H_0 hipotezi reddedildiğinde farklılığın nereden ileri geldiği araştırılmalıdır. Bu amaçla ileri analiz yapılır. Ki-kare değeri en büyük olan satır analiz dışı bırakılarak işlem tekrar yapılır. Testi farklı çıkaran faktör bulununca da yüzdelere bakılarak yorum yapılır.

TEK DEĞİŞKENLİ DÜZENLERDE Kİ-KARE TESTİ

⊗ Bazı durumlarda incelenen değişkenin sadece bir özelliği bilinir. Ör: bir bölgede bir hastalığa yakalananlar bilinir, fakat yakalanmayanlar bilinmeyebilir. Böyle durumlarda bilinen özelliğin değişik durum ya da zamanlarda farklılık gösterip göstermediği tek değişkenli düzenlerde ki-kare testi ile araştırılabilir.

⊗ Ör: Bir bölgede görülen kızamık vakalarının mevsimlere göre dağılımı aşağıda verilmiştir. Kızamık görülme bakımından mevsimler arasında fark var mıdır?

Mevsimler	Görülen Vaka Sayısı
İlkbahar	210 (225)
Yaz	165 (225)
Sonbahar	225 (225)
Kış	300 (225)
Toplam	900

Test İşlemleri

⊗ H_0 : Kızamık görülme sayısı bakımından mevsimler arasında fark yoktur.

⊗ Beklenen frekansların bulunması: H_0 hipotezi 'Kızamık görülme sayısı bakımından mevsimler arasında fark yoktur.' olduğuna göre kızamık vakalarının her mevsimde aynı sayıda görülmesi beklenir. Bu nedenle toplam 900 vaka 4'e (grup sayısına) bölünerek beklenen frekanslar bulunur.

$$\chi^2 = \sum \frac{(G - B)^2}{B} = 42$$

⊗ Yanılma olasılığı olarak $\alpha = 0.05$ seçilmiştir.

⊗ Serbestlik derecesi = Grup sayısı - 1 = 3

⊗ $\alpha = 0.05$ yanılma düzeyinde 3 SD'de tablo χ^2 değeri bulunur (7.82).

⊗ Karşılaştırma: Hesapla bulunan χ^2 değeri tablo χ^2 değerinden büyükse H_0 hipotezi reddedilir, küçükse kabul edilir. H_0 hipotezi reddedilecektir

⊗ Karar: Kızamık görülme sayısı bakımından mevsimler arasında fark vardır ($\chi^2 = 42$, $p < 0.05$).

⊗ İstenirse farklılığın hangi mevsimden kaynaklandığı ileri analizle test edilebilir.

BAĞIMLI GRUPLARDA Kİ-KARE TESTİ

⊗ Bu test, nitelik olarak belirtilen bir değişken yönünden aynı bireylerden değişik zaman ya da durumda elde edilen iki gözlemin farklı olup olmadığını test etmek için kullanılır.

⊗ Ör: 1984 yılında 315 kadın incelenmiş ve 200 kadının doğum kontrol yöntemi kullandığı görülmüş. Aynı kadınlar 1985 yılında tekrar incelendiğinde bu kez 225 kadının yöntem kullandığı saptanmış. Bu durumda aradaki farkı test etmek için aşağıdaki gibi bir tablo geliştirilir.

1985	1984		
	Kullanan	Kullanmayan	Toplam
Kullanan	150 (a)	50 (b)	200
Kullanmayan	75 (c)	40 (d)	115
Toplam	225	90	315

⊗ H_0 : Yöntem kullanma bakımından iki yıl arasında fark yoktur.

$$\chi^2 = \frac{(c - b)^2}{c + b} = \frac{(75 - 50)^2}{75 + 50} = 5$$

⊗ Yanılma olasılığı seçilir ($\alpha = 0.05$).

⊗ Serbestlik derecesi = 1

⊗ $\alpha = 0.05$ yanılma düzeyinde 1 SD'de tablo χ^2 değeri = 3.841

⊗ Karşılaştırma: Hesapla bulunan χ^2 değeri tablo χ^2 değerinden büyükse H_0 hipotezi reddedilir, küçükse kabul edilir. H_0 hipotezi reddedilecektir.

⊗ Karar: Yöntem kullanma yönünden iki yıl farklıdır ($\chi^2 = 5$, $p < 0.05$)

İŞARET TESTİ

- Evren ortancasının belirli bir değere eşit olup olmadığının test edilmesinde kullanılır.
- 'Evren ortalaması önemlilik testi'nin nonparametrik karşılığıdır.
- Bu teste işaret testi denmesinin nedeni değerlerin + ve - işaretlerine dönüştürülmesidir.
- İncelenen denek sayısının 25'den az olup olmama durumuna göre test ayrı işlemlerle yapılır.

Denek Sayısı 25'den Az Olduğunda Test İşlemleri

- Her veri (Xi) ortancadan çıkarılır ve sonuç pozitif ise '+', negatif ise '-' işareti ile gösterilir. Eğer $X_i - \text{Ort.} = 0$ ise bu değer analizden çıkarılır ve denek sayısı analizden çıkarılan denek sayısı kadar azaltılır.
- $H_0 = \text{Ortanca} =$
 $H_1 = \text{Ortanca} <$
- Yanılma olasılığı seçilir ($\alpha = 0.05$).
- Tablo olasılık değerinin bulunması: n' ye ve işaret dizisinde az sayıda olan işaret sayısına (k) göre İşaret testi tablosundan olasılık değeri bulunur.
- Karşılaştırma: Tablodan bulunan olasılık değeri saptanan yanılma olasılığından küçükse H_0 hipotezi reddedilir, büyükse kabul edilir.
- Karar: Ortancadan farklı bulunamamıştır ($p > 0.05$).

Denek Sayısı 25'den Daha Fazla Olduğunda Test İşlemleri

- Her veri (Xi) ortancadan çıkarılır ve sonuç pozitif ise '+', negatif ise '-' işareti ile gösterilir. Eğer $X_i - \text{Ort.} = 0$ ise bu değer analizden çıkarılır ve denek sayısı analizden çıkarılan denek sayısı kadar azaltılır.
- $H_0 = \text{Ortanca} =$
 $H_1 = \text{Ortanca} <$
- Yanılma olasılığı seçilir ($\alpha = 0.05$).
- $$z = \frac{k - n/2}{n/2}$$
- Karşılaştırma: Z tablosundan bulunan olasılık değeri saptanan yanılma olasılığından küçükse H_0 hipotezi reddedilir, büyükse kabul edilir.
- Karar: Ortancadan farklı bulunamamıştır ($p > 0.05$).

MANN - WHITNEY U TESTİ

- 'İki ortalama arasındaki farkın önemlilik testi' parametrik test varsayımları yerine getirildiğinde ölçümle belirtilen sürekli bir değişken yönünden bağımsız iki grup arasında fark olup olmadığını test etmek için kullanılır. Mann-Whitney U testi 'İki ortalama arasındaki farkın önemlilik testi' nin non-parametrik karşılığıdır.
- Her iki gruptaki denek sayısı 20 ya da daha fazla ise, grupların birindeki ya da ikisindeki denek sayısı 20 'den fazla ise test ayrı işlemlerle yapılır.

Her İki Gruptaki Denek Sayısı 20 ya da Daha Az Olduğunda Test İşlemleri

Ör: A sınıfından 7, B sınıfından 8 öğrencinin vücut ağırlıkları ölçülmüş ve bulgular aşağıda verilmiştir. İki grup öğrencinin vücut ağırlıkları farklı mıdır?

A sınıfı: 30, 38, 32, 34, 40, 37, 39

B sınıfı: 29, 35, 36, 42, 41, 31, 33, 43

* İki sınıftaki öğrencilerin ağırlık değerleri küçükten büyüğe doğru tek bir dizi halinde sıralanır. Böylece 2 grup tek grup haline dönüşür.

* Bu dizideki değerlere küçükten büyüğe doğru, 1 'den başlanarak ve birer arttırılarak sıra numarası verilir. Dizide aynı değeri alan birden fazla denek varsa bunlara denk gelen sıra numaralarının ortalaması sıra numarası olarak verilir.

Sınıf	B	A	B	A	B	A	B	B	A	A	A	A	B	B	B
Ağırlık	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43
Sıra No	1	(2)	3	(4)	5	(6)	7	8	(9)	(10)	(11)	(12)	13	14	15

ω Ho: İki dağılım arasında fark yoktur.

$$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

R_1 = Birinci gruptaki deneklerin sıra numaraları toplamı

$$U_2 = n_1 n_2 - U_1$$

$$U_1 = 30 \quad U_2 = 26$$

ω Yanılma olasılığı seçilir ($\alpha = 0.05$).

ω Mann-Whitney kritik U değeri tablosunda n_1 ve n_2 değerlerinin kesiştiği yerdeki değer tablo U değeridir (43).

ω Karşılaştırma: Hesapla bulunan U değerlerinden büyük olanı tablo U değeri ile karşılaştırılır. Bu değer tablo U değerinden büyükse H_0 hipotezi reddedilir, küçükse kabul edilir. H_0 hipotezi kabul edilecektir.

ω Karar: İki grubun dağılımları arasında fark yoktur. ($U = 30, p > 0.05$)

Grupların Birinde ya da İkisindeki Denek Sayısı 20'den Fazla Olduğunda Test İşlemleri

ω U değerleri aynı şekilde hesaplanır.

$$z = \frac{U - \frac{(n_1 n_2)}{2}}{\sqrt{\frac{(n_1 + n_2 + 1) n_1 n_2}{12}}}$$

U yerine U_1 veya U_2 'den herhangi birisi kullanılabilir. Sadece bulunacak z değerinin işareti farklı olur.

ω Karşılaştırma: Z tablosundan bulunan olasılık değeri saptanan yanılma olasılığından küçükse H_0 hipotezi reddedilir, büyükse kabul edilir.

WILCOXON EŞLEŞTİRİLMİŞ İKİ ÖRNEK TESTİ

ω 'İki eş arasındaki farkın önemlilik testinin non-parametrik karşılığıdır.

ω İncelenen denek sayısının 25'den az olup olmama durumuna göre test ayrı işlemlerle yapılır.

Denek Sayısı 25'den Az Olduğunda Test İşlemleri

ω Ör: Kandaki kolesterol düzeyinin düşürülmesinde diyetin etkinliğini incelemek için 12 kişinin diyetten önceki ve diyetten sonraki kolesterol düzeylerinin ölçüldüğünü ve bulguların aşağıdaki gibi olduğunu kabul edelim. Bu verilere göre diyetin kolesterol düzeyini düşürdüğünü söyleyebilir miyiz?

ω Her kişinin değerleri önce ve sonra kolonlarına yazılır.

ω İki ölçüm arasındaki farklar (önce-sonra) alınır ve fark kolonuna yazılır.

ω Farklara küçükten büyüğe doğru sıra numarası verilir. Aynı değeri alan birden çok fark varsa onlara denk gelecek sıra numaralarının ortalaması bu değerlerin sıra numarası olur.

ω Sıfır değeri alan fark ya da farklar varsa:

1 tane sıfır varsa

değerlendirmeden çıkarılır ve denek sayısı 1 azaltılır. Sıfır sayısı çift ise önce sıfırlar sıralanır. Sıfırlara denk gelen sıra numaraları ortalaması sıfırların sıra numarası olur. Sıfırların sıra numaralarının yarısına +, yarısına - işareti konur.

Sıfır sayısı tek ise sıfırların herhangi birisi

değerlendirmeden çıkarılır, denek sayısı 1 azaltılır ve işlem sıfır sayısı çift olduğu zamanki gibi yapılır.

ω Farkların işareti sıra numarasının önüne konur.

Diyetten Önce	Diyetten Sonra	Fark	Sıra No	İşaretili Sıra No
201	200	1	1	1
231	236	-5	3.5	- 3.5
221	216	5	3.5	3.5

..

* Artı ve eksi işaretlerinin hangisi daha az ise onun sıra numaraları toplanır.

$$T = 5 + 3.5 = 8.5$$

* H_0 = İki eş arasında fark yoktur.

* Yanılma olasılığı olarak $\alpha = 0.05$ seçilmiştir.

* $\alpha = 0.05$ yanılma düzeyinde $n = 12$ 'deki tablo T değeri Wilcoxon eşleştirilmiş iki örnek testi tablosundan bulunur (14).

• Karşılaştırma: Hesapla bulunan T değeri tablo T değerinden küçükse H_0 hipotezi reddedilir, büyükse kabul edilir. H_0 hipotezi reddedilecektir.

• Karar: Diyet kolesterol düzeyini düşürmüştür ($T = 14, p < 0.05$)

Denek Sayısı 25 ya da Daha Fazla Olduğunda Test İşlemleri

ω Hazırlık işlemleri aynı şekilde yapılır.

$$\omega \quad z = \frac{T - \frac{n(n+1)}{4}}{\frac{n(n+1)(2n+1)}{24}}$$

ω Hesaplanan z değerine göre z tablosundan olasılık değeri bulunur. Tablo olasılık değeri saptanan yanılma olasılığından küçükse Ho hipotezi reddedilir, büyükse kabul edilir.

KRUSKAL WALLİS VARYANS ANALİZİ

ω ‘Varyans analizi’ parametrik test varsayımları yerine getirildiğinde ölçümle belirtilen sürekli bir değişken yönünden ikiden çok bağımsız grup arasında farklılık olup olmadığını incelemek için kullanılır. Kruskal Wallis Varyans Analizi Varyans Analizinin non-parametrik karşılığıdır.

ω Ör: 18 deney hayvanı rastgele 3 gruba ayrılmış, her gruptaki hayvanlara değişik dozda toksik madde verilmiş ve ölüm süreleri dakika cinsinden ölçülmüştür. Bulgular aşağıda verilmiştir. Dozlar arasında fark var mıdır?

ω Bütün gruptaki değerler küçükten büyüğe doğru tek dizi halinde sıralanır ve her bir değere sıra numarası verilir. Dizide aynı değeri alan birden çok denek varsa bunlara denk gelen sıra numaralarının ortalaması bu değerlerin sıra numarası olur.

ω Ölüm Süresi: 3 3 4 5 5

ω Sıra No : 1.5 1.5 3 4.5 4.5

* Her gruptaki değerlerin sıra numaraları yanlarına yazılır. Önce her gruptaki sıra numaraları toplanır (Tj). Sonra bütün gruptaki Tj’ler toplanarak T değeri bulunur.

DOZLAR

0.50 mgr		0.75 mgr		1.00 mgr	
Ölüm Süresi(dk)	Sıra No	Ölüm Süresi(dk)	Sıra No	Ölüm Süresi(dk)	Sıra No
10	13.5	7	7.5	4	3
11	15	8	9.5	5	4.5
10	13.5	9	11.5	6	6
12	16	8	9.5	3	1.5
13	17	7	7.5	5	4.5
14	18	9	11.5	3	1.5
Tj	93		57		21
nj	6		6		6

Toplam (T): 171

(n): 18

* Eğer her bir değerlerin sıra numarası doğru olarak verilmişse bulunan T değeri aşağıdaki formülde elde edilen değere eşit olur: $T = \frac{n(n+1)}{2}$

* Ho: Gruplar arasında fark yoktur.

H1: En az bir grup diğerlerinden farklıdır.

$$* KW = \frac{12}{n(n+1)} \left(\sum \frac{T_j^2}{n_j} \right) - 3(n+1) = 15.16$$

* Yanılma olasılığı olarak $\alpha = 0.05$ seçilmiştir.

* İncelenen gruptaki denek sayısı 5’ ten fazla ise ya da grup sayısı 3’den fazla ise tablo değeri $\alpha = 0.05$ yanılma düzeyinde Grup sayısı-1 SD’de ki-kare tablosundan bulunur. Hesapla bulunan KW değeri tablo ki-kare değerinden büyükse H0 hipotezi reddedilir, küçükse kabul edilir. Örneğimizde ki-kare değeri 5.99’ dur. Ho hipotezi reddedilecektir.

Karar: Ölüm süresi bakımından dozlar arasında fark vardır. (KW=15.16, p<0.05)

* Grup sayısı 3 ise ve her gruptaki denek sayısı 5 ya da daha küçükse tablo olasılık değeri Kruskal Wallis tablosundan bulunur. Bu tabloda her 3 gruptaki denek sayıları ve KW değerine göre olasılık değeri verilmiştir. Tablodaki ilk kolondan n1 n2 n3 sırası bulunur. Bu bloktaki KW değerleri incelenir, hesaplanan KW değerine en yakın küçük değer karşısındaki olasılık değeri (p) bulunur. Tablo olasılık değeri saptanan yanılma düzeyinden küçükse Ho hipotezi reddedilir, büyükse kabul edilir.

KOLMOGOROV SMİRNOV İKİ ÖRNEK TESTİ

- Herhangi bir değişken yönünden birbirinden bağımsız iki grubun dağılımlarının benzer olup olmadığını, başka bir deyişle bu iki grubun aynı evrenden ya da aynı dağılıma sahip değişik evrenlerden çekilip çekilmediğini test etmek için kullanılan bir yöntemdir.
- Grup sayısı ikiden fazla olursa kullanılamaz.
- Gözlerdeki küçük frekanslardan etkilenmez. Bu nedenle nx2 ve mx2 düzeninde gözlerden birisinde 5’den küçük beklenen frekans olduğunda ki-kare testi yerine Kolmogorov-Smirnov testi uygulanabilir.
- İncelenen değişkenin ayrıntılı olarak sınıflandırılması testin gücünü artırır.
- Bu test denek sayısı 40’dan büyük ya da küçükse ve hipotez çift yönlü ya da tek yönlü ise ayrı işlemlerle yapılır.
- İşlemler her iki grubun yığılımlı dağılımları üzerinden yapılır. Eğer iki grup aynı evrenden ya da aynı dağılıma sahip ayrı evrenlerden çekilmiş ise bu iki grubun yığılımlı dağılımlarının da benzer olması gerekir.

Her İki Gruptaki Denek Sayısı 40’dan Büyük Olduğunda

- Örnek:** 113 kişinin eğitim düzeylerine göre bulaşıcı hastalıklar bilgisi aşağıda belirtilmiştir. Bulaşıcı hastalıklar bilgisi eğitim düzeyine göre farklılık göstermekte midir?

Eğitim Düzeyi	Bulaşıcı Hastalıklar Bilgisi	
	Yeterli	Yetersiz
Okur-yazar değil	3	13
Okur-yazar	4	12
İlkokul	6	8
Ortaokul	14	6
Lise	20	3
Yüksek okul	22	2

1. Hazırlık İşlemleri:

Eğitim Düzeyi	Bulaşıcı Hastalıklar Bilgisi						
	Yeterli			Yetersiz			
	f	%	Yığılımlı %	f	%	Yığılımlı %	Fark
	(1)	(2)	(3)	(4)	(5)	(6)	(6)-(3)=(7)
Okur-yazar değil	3	0.043	0.043	13	0.295	0.295	0.252
Okur-yazar	4	0.058	0.101	12	0.273	0.568	0.467
İlkokul	6	0.087	0.188	8	0.182	0.750	(0.562)
Ortaokul	14	0.203	0.391	6	0.136	0.886	0.495
Lise	20	0.290	0.681	3	0.068	0.955	0.274
Yüksek okul	22	0.319	1.000	2	0.045	1.000	0.000
Toplam	69	1.000		44	1.000		

Her grup için ayrı ayrı:

- Frekanslar yazılır (kolon 1 ve 4).
- Frekanslar gruptaki toplam denek sayısına bölünerek yüzdelere hesaplanır (kolon 2 ve 5).
- Yüzdelerin yığılımlı toplamaları alınır (kolon 3 ve 6).

Her eğitim düzeyinde birinci grubun yığılımlı yüzdesi ikinci grubun yığılımlı yüzdesinden çıkarılarak 7’inci kolona yazılır (Farkların mutlak değerleri önemlidir, işareti dikkate alınmaz).

2. Test İşlemleri:

Hipotezin çift yönlü ya da tek yönlü kurulmasına göre test işlemleri ayrı yolla yapılır.

Hipotez Çift Yönlü İse Test İşlemleri:

- H_0 : Bulaşıcı hastalıklar bilgisi bakımından eğitim grupları arasında fark yoktur.
 H_1 : Eğitim grupları farklıdır.
- Yanılma olasılığı olarak $\alpha = 0.05$ seçilmiştir.
- Gözlenen D değerinin bulunması: Fark kolonundaki mutlak değeri en büyük olan değer gözlenen D değeridir. Örneğimizde bu değer 0.562'dir.
- Beklenen D değerinin bulunması:

$$\text{Beklenen D} = K \sqrt{\frac{n_1 + n_2}{n_1 n_2}}$$

K= Saptanan yanılma olasılığında sabit değer. Çeşitli yanılma düzeylerine göre K Değerleri aşağıda verilmiştir.

Yanılma Olasılığı	K değeri
0.10	1.22
0.05	1.36
0.025	1.48
0.01	1.63
0.001	1.95

Örneğimizde $n_1 = 69$, $n_2 = 44$ ve $\alpha = 0.05$ düzeyinde K:1.36'dır. Buna göre;

$$\text{Beklenen D} = 1.36 \sqrt{\frac{69+44}{69 \times 44}} = 0.262 \text{ dir.}$$

Karşılaştırma: Gözlenen D değeri, beklenen D değerinden büyükse H_0 hipotezi reddedilir, küçükse ya da eşitse kabul edilir. Gözlenen D (0.562) beklenen D değerinden (0.262) büyük olduğu için H_0 hipotezi reddedilecektir.

Karar: Bulaşıcı hastalıklar bilgisi bakımından eğitim grupları arasında fark bulunmuştur ($p < 0.05$).

Hipotez Tek Yönlü İse Test İşlemleri:

- H_0 : Bulaşıcı hastalıklar bilgisi bakımından eğitim grupları arasında fark yoktur.
 H_1 : Eğitim düzeyi yüksek olanların bulaşıcı hastalıklar bilgisi yeterlidir.
- Yanılma olasılığı olarak $\alpha = 0.05$ seçilmiştir.
- Gözlenen D değeri aşağıdaki formülle ki-kareye dönüştürülür.

$$x^2 = 4D^2 \frac{n_1 n_2}{n_1 + n_2}$$

D= Tabloda fark kolonundaki mutlak değeri en büyük olan değer.

Örneğimizde $n_1 = 69$, $n_2 = 44$ ve $D = 0.562$ 'dir. Buna göre:

$$x^2 = 4 \times (0.562)^2 \times \frac{69 \times 44}{69 + 44} = 33.94$$

- Serbestlik derecesi: Kolmogorov-Smirnov testinde serbestlik derecesi her zaman 2'dir.
- $\alpha = 0.05$ düzeyinde ve 2 serbestlik derecesindeki tablo x^2 değeri 5.99'dur.
- Karşılaştırma: Hesapla bulunan ki-kare değeri tablo ki-kare değerinden büyükse H_0 hipotezi reddedilir, küçükse kabul edilir. Hesapla bulunan ki-kare değeri (33.94) tablo ki-kare değerinden (5.99) büyük olduğu için H_0 hipotezi reddedilir.
- Karar: Eğitim düzeyi yüksek olanların bulaşıcı hastalıklar bilgisi yeterlidir ($x^2 = 33.94$, $p < 0.05$).

Her İki Gruptaki Denek Sayısı 40 ve Daha Az ve Birbirine Eşitse

Denek sayıları 40 ve daha az olduğunda her iki grupta aynı sayıda denek olmalıdır. Aksi halde bu test uygulanamaz.

Örnek: A ve B bölgelerindeki sağlık ocaklarından elde edilen bebek ölüm hızları aşağıda gösterilmiştir. Bölgeler farklı mıdır?

BÖH (Binde)	A Bölgesi	B Bölgesi
91-100	1	0
101-110	3	1
111-120	5	6
121-130	3	4
131-140	0	2
141-150	1	0
Toplam	13	13

1. 1. Hazırlık İşlemleri:

- Frekanslar yazılır.
 - Frekansların yığılımlı toplamı alınır.
- Bu iki işlem her iki grup için ayrı ayrı yapılır.
- Yığılımlı frekanslar arasındaki farklar alınır ve fark kolonuna yazılır.

Bebek Ölüm Hızı (Binde)	A Bölgesi		B Bölgesi		Fark (2)-(4)=(5)
	f (1)	Yığılımlı f (2)	f (3)	Yığılımlı f (4)	
91-100	1	1	0	0	1
101-110	3	4	1	1	(3)
111-120	5	9	6	7	2
121-130	3	12	4	11	1
131-140	0	12	2	13	-1
141-150	1	13	0	13	0

2. Test İşlemleri:

- H_0 : Bebek ölüm hızı bakımından bölgeler arasında fark yoktur.
 H_1 : Bebek ölüm hızı bakımından bölgeler arasında fark vardır.
- Yanılma olasılığı olarak $\alpha = 0.05$ seçilmiştir.
- Gözlenen D değerinin bulunması: Fark kolonundaki mutlak değeri en büyük olan değer gözlenen D değeri idi. Örneğimizde bu değer 3'tür.
- $\alpha = 0.05$ yanılma düzeyinde $n_1 = n_2 = n = 13$ 'deki tablo D değeri 7'dir (iki yönlü hipotezde).
- Karşılaştırma: Gözlenen D değeri tablo D değerinden büyükse H_0 hipotezi reddedilir, küçükse kabul edilir. Gözlenen D değeri (3) tablo D değerinden (7) küçük olduğu için H_0 hipotezi kabul edilecektir.
- Karar: Bebek ölüm hızı bakımından bölgeler arasında fark bulunamamıştır ($p > 0.05$).

KORELASYON VE REGRESYON ANALİZİ

- ◆ İki ya da daha çok değişken arasında ilişki olup olmadığını, ilişki varsa yönünü ve gücünü inceleyen korelasyon analizi ile değişkenlerden birisi belirli bir birim değiştiğinde diğerinin nasıl bir değişim gösterdiğini inceleyen regresyon analizi sağlık bilimlerinde çok kullanılan istatistiksel yöntemlerdir.
 - ◆ Değişkenler arasındaki ilişkilerden yararlanılarak geliştirilecek matematiksel modeller yardımıyla tahminler yapılabilir.
 - ◆ Korelasyon ve regresyon analiziyle yalnız bireylere ait değişkenlerin kendi aralarındaki ilişkileri değil, aynı zamanda bireylere ait değişkenlerle bazı çevresel değişkenlerin ilişkileri ya da çevresel değişkenlerin kendi aralarındaki ilişkileri incelenebilir.
 - *Korelasyon ve regresyon analizi değişkenler arasındaki ilişkinin neden-sonuç ilişkisi olup olmadığını gösteremez.*
 - Gerçekte iki değişken arasında ilişki olmamasına rağmen analiz sonucunda iki değişken arasında ilişki varmış gibi bir sonuç elde edilebilir. Bunun nedeni bu iki değişkenin her ikisinin de başka bir değişkenden etkilenmeleridir.
 - Bu nedenle korelasyon ve regresyon analizinde iki değişken arasında ilişki saptanırsa bu ilişkinin neden-sonuç ilişkisi olup olmadığı mutlaka incelenmelidir.
 - İki değişken arasındaki ilişkinin incelenmesine basit korelasyon-regresyon analizi, ikiden çok değişken arasında ilişkinin incelenmesine ise çoklu korelasyon-regresyon analizi adı verilir.
- Bir değişkeni etkileyen değişkenlere bağımsız değişkenler denir ve 'x' ile sembolize edilir. Bağımsız değişkenlerden etkilenen değişkene ise bağımlı değişken denir ve 'y' ile sembolize edilir.

BASİT KORELASYON VE REGRESYON ANALİZİ

KORELASYON ANALİZİ

Basit korelasyon analizi ile iki değişken arasındaki ilişkinin yönü ve gücü incelenir.

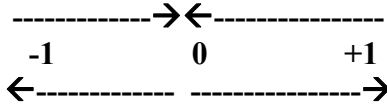
İki değişken arasındaki ilişki üç şekilde ortaya çıkabilir:

- İki değişken arasında pozitif bir ilişki vardır: Bir değişken artarken diğeri de artıyorsa ya da biri azalırken diğeri de azalıyorsa.
- İki değişken arasında negatif bir ilişki vardır: Bir değişken artarken diğeri azalıyorsa ya da biri azalırken diğeri artıyorsa.
- İki değişken arasında ilişki yoktur: İki değişken birbirinden tamamen bağımsızdır ve birbirini etkilememektedir. Değişkenlerden biri yüksek bir değer alırken diğeri yüksek ya da düşük bir değer alabilmektedir.

KORELASYON KATSAYISI

- İki değişken arasındaki ilişkinin gücünü gösteren ölçü korelasyon katsayısıdır. ‘r’ sembolü ile gösterilir.
- Değişkenler arasındaki ilişki pozitif ise işareti artı, negatif ise eksi olur.
- Korelasyon katsayısı -1 ile $+1$ arasında herhangi bir değer alabilir ($-1 \leq r \leq +1$). Hiçbir zaman -1 ’den küçük, $+1$ ’den büyük olamaz. Her iki yönde sıfırdan 1 ’e yaklaştıkça ilişkinin kuvveti azalır, sıfıra gelince kaybolur.

İlişkinin kuvveti azalır



İlişkinin kuvveti artar

* İki değişken arasında pozitif tam bir ilişki varsa korelasyon katsayısı $+1$, negatif tam bir ilişki varsa -1 , ilişki yoksa sıfır olur. Korelasyon katsayısı $0.0 - 0.50$ arasında ise ilişkinin zayıf, $0.50 - 1.0$ arasında ise ilişkinin kuvvetli olduğu kabul edilir.

Korelasyon katsayısının hesaplanması

- Ör: Gebelik sayısı ile Hb değeri arasında ilişki olup olmadığını saptamak için 20 kadın üzerinde yapılan bir araştırmada aşağıdaki sonuçlar bulunmuştur. Gebelik sayısı x, Hb değeri y ile gösterilmiştir.

x	y	x^2	y^2	xy
1	14	1	196	14
2	13	4	169	26
3	12	9	144	36
3	13	9	169	39
2	11	4	121	22
1	12	1	144	12
4	12	16	144	48
5	11	25	121	55
4	14	16	196	56
3	13	9	169	39
6	12	36	144	72
5	12	25	144	60
10	10	100	100	100
9	11	81	121	99
1	14	1	196	14
8	11	64	121	88
9	10	81	100	90
7	9	49	81	63
6	12	36	144	72
7	10	49	100	70
$\Sigma x=96$	$\Sigma y=236$	$\Sigma x^2=616$	$\Sigma y^2=2824$	$\Sigma xy=1075$

* Her deneğin x ve y değerleri toplanır. Buradan Σx ve Σy bulunur.

* Her deneğin x ve y değerlerinin ayrı ayrı kareleri alınır ve kareler toplanır. Buradan Σx^2 ve Σy^2 bulunur.

* Her deneğin x ve y değerleri çarpımları toplanır. Buradan Σxy bulunur.

$$r = \frac{\Sigma xy - \frac{(\Sigma x)(\Sigma y)}{n}}{n}$$

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{(\sum (x - \bar{x})^2)(\sum (y - \bar{y})^2)}}$$

$$r = -0.74$$

Gebelik sayısı ile Hb değeri arasında - 0.74 'lük bir ilişki vardır. Bu ilişki negatif yöndedir. Gebelik sayısı arttıkça Hb değeri azalmaktadır.

Korelasyon Katsayısı Önem Kontrolü

- ♦ Bulunan korelasyon katsayısının önemli bir katsayı mı yoksa tesadüfe mi bağlı olduğu test edilebilir.
- ♦ Ho: Korelasyon katsayısı tesadüfe bağlı bir değerdir. (r= 0)
- ♦ $t = \frac{r}{S_r} = -4.68$
- ♦ S_r korelasyon katsayısının standart hatasıdır.
- ♦ $S_r = \frac{1 - r^2}{n - 2} = 0.158$
- ♦ Yanılma olasılığı olarak $\alpha=0.05$ seçilmiştir.
- ♦ Serbestlik derecesi = n - 2
- ♦ $\alpha=0.05$ düzeyinde ve n - 2 SD'de tablo t değeri 2.10'dur.
- ♦ Karşılaştırma: Hesapla bulunan t değeri tablo t değerinden büyükse Ho hipotezi reddedilir, küçükse kabul edilir. Ho hipotezi reddedilecektir.
- ♦ Karar: Korelasyon katsayısı önemli bir değerdir. Tesadüfen bulunmuş bir değer değildir. (t= 4.68, p<0.05)

Tamamlayıcılık Katsayısı(R²)

- ♦ Tamamlayıcılık katsayısı korelasyon katsayısının karesidir. Bağımlı değişkendeki değişimin yüzde ne kadarının bağımsız değişken tarafından tanımlanabildiğini gösteren bir ölçüdür.
- ♦ Eğer iki değişken arasında tam bir ilişki varsa korelasyon katsayısı -1 veya +1, dolayısı ile karesi de 1 olacaktır. Böyle bir durumda bağımsız değişkenin bağımlı değişkeni tümüyle tanımladığı yani bağımlı değişkenin başka hiçbir değişken tarafından etkilenmediği kararına varılır.

REGRESYON ANALİZİ

- ♦ İki değişken arasında belirgin bir ilişki olduğunda, bu ilişki dağılım grafiğindeki noktalar arsından geçen uygun bir doğru ile tanımlanabilir. Bu doğruya regresyon doğrusu denir ve matematiksel olarak bir denklem ile gösterilebilir. Bu denkleme de regresyon denklemi denir.
- ♦ Regresyon denklemi yardımı ile bağımsız değişkene verilen herhangi bir değere karşı bağımlı değişkenin alacağı değer hesaplanabilir.
- ♦ İki değişken arasında tam bir ilişki varsa (r= ±1) dağılımdaki bütün noktalar regresyon doğrusu üzerine düşer. Ancak, çoğu kez regresyon doğrusu bütün noktaların üzerinden geçemez ve dağılımı yaklaşık olarak temsil edebilir.
- ♦ 'x' değerlerine karşı 'y' değerleri bir grafik üzerinde noktalandıktan sonra ilişkinin bir doğru ile tanımlanması istenirse doğrunun göz kararı ile çizilmesi isabetli olamaz. Noktaları en iyi temsil eden doğru en küçük kareler yöntemi (EKKY) ile çizilir.
- ♦ Regresyon doğrusunun denklemi: y= a + bx dir.
- ♦ Denklemde: y : Bağımlı değişken
a : Doğrunun y eksenini kestiği nokta
b : Regresyon katsayısı

x : Bağımsız değişken

Korelasyon katsayısının hesaplanmasında verilen örnek problemi kullanarak regresyon denklemini bulalım ve regresyon doğrusunu çizelim:

1. Regresyon katsayısı (b) bulunur.

$$\frac{\sum x \sum y}{n}$$

$$b = \frac{\sum xy - \frac{n}{\sum x^2 - (\sum x)^2}}{n} = -0.37$$

Regresyon katsayısı her değeri alabilir. İşareti artı ya da eksi olabilir.

2. x ve y değerlerinin ortalamaları bulunur:

$$\bar{X} = \frac{96}{20} = 4.8 \quad \bar{Y} = \frac{236}{20} = 11.8$$

3. a değeri bulunur.

$$a = \bar{Y} - b \bar{X} = 13.57$$

4. Bulgular denkleme yerleştirilir. $y = a + bx$

$$y = 13.57 - 0.37x$$

5. Regresyon Doğrusunun çizimi: Çizim için doğrunun y eksenini nokta (a) ile X ve Y'nin kesiştiği noktanın bilinmesi yeterlidir.

Bulunan denklem yardımıyla bağımsız değişken x'e değişik değerler verilerek, bağımlı değişken y'nin alacağı değerler tahmin edilebilir.

Örneğin x'in 0, 5 ve 10 değerlerine karşı y'nin alacağı değerler aşağıda gösterilmiştir.

x	0	5	10
y	13.57	11.72	9.87

Bu tahminlere göre; hiç gebe kalmayan bir kadının Hb değeri 13.57 mgr., beş kez gebe kalanınki 11.72 mgr., ve 10 kez gebe kalanınki ise 9.87 mgr.olacaktır.

DOĞRUSALLIKTAN AYRILIŞ ÖNEM KONTROLÜ

- ◆ Regresyon doğrusu çizmeden önce doğrusallıktan ayrılış önem kontrolü yapmak gerekir.
- ◆ Eğer doğrusallıktan ayrılış önemli ise iki değişken arasındaki ilişki doğrusal değildir. Başka bir deyişle ilişki $y = a + bx$ doğrusal denklemi ile ifade edilemez. Bu nedenle regresyon doğrusu çizilemez.

Çoklu Korrelasyon ve Regresyon Analizi

- ◆ İncelenen bağımlı değişkeni sadece bir değişken etkilediği zaman bu iki değişken arasındaki ilişki basit korrelasyon ve regresyon analiziyle incelenir.
- ◆ Bazı durumlarda ise incelenen bağımlı değişkeni birden çok sayıda değişken etkileyebilir. Bu değişkenler arasındaki ilişkiler ise çoklu korrelasyon ve regresyon analiziyle incelenir.
- ◆ Ör: Doğan bebeğin doğumdaki ağırlığını (bağımlı değişken), gebelik haftası, anne yaşı, annenin gebelik süresince beslenme durumu, gebelik sayısı gibi birçok değişken etkileyebilir.
- ◆ Çoklu korrelasyon ve regresyon analizinin hesaplanma işlemleri elle yapılamayacak kadar güç ve zaman alıcıdır.

* Çoklu korellasyon ve regresyon analizinden olumlu sonuç alabilmek için bağımlı ve bağımsız değişkenlerin tümünün ölçümle belirtilmiş olması, en azından sayısal olarak gösterilmesi gerekir. İyi-kötü, az-çok, zengin-fakir gibi niteliksel verilerde bu analiz iyi sonuç vermemektedir.

* Bu güçlüğü gidermek için bazen niteliksel özellikler sayıya çevrilmektedir. Ancak, bu bile çoğu kez yeterli olmamaktadır. Ör: Zengine 1, fakire 0 kodu verilirse analizde 1'le 0 arasındaki sayısal fark, zenginle fakir arasındaki farkı ne denli yansıtabilir?

* Bazı niteliksel özellikleri sayıya çevirmek bile anlamsız olabilir. Ör: Mesleği bağımsız olarak analize katmak istersek, meslekleri nasıl sayıya çevirebiliriz?

* Çoklu korellasyon ve regresyon analizinden iyi sonuç elde edilememesinin nedenlerinden birisi incelenen bağımlı değişken etkileyebilecek bağımsız değişkenleri uygun ve yeterli sayıda seçememektir. Seçim ya kişisel deneyim ve bilgilere ya da daha önce yapılmış araştırmalardan elde edilen sonuçlara göre yapılır.

* Bağımsız değişkenleri seçerken dikkat edilecek diğer bir husus, bu değişkenlerin kendi aralarında yüksek derecede korrelasyona sahip olmamalarıdır. Ör: Kadının yaşı, gebelik sayısı, doğum sayısı ayrı değişkenlermiş gibi görünürse de aslında her üçü de birbiri ile ilişkilidir. Bu bakımdan bir çalışmada üçünün birden bağımsız değişken olarak kullanılması yerine birisinin seçilmesi yeterlidir.

Hangi bağımsız değişkenin hangisi ile ilişkili olduğu bilinmiyorsa önce bütün bağımsız değişkenlerin birbiriyle olan korrelasyon katsayılarına bakılmalı ve bunlardan birbiriyle yüksek korrelasyon katsayısına sahip olanlar arasından bir tanesi seçilmelidir

ÇOKLU REGRESYON VE KORRELASYON ANALİZİ YÖNTEMLERİ

- ◆ Çoklu korrelasyon ve regresyon analizi sonucunda matematiksel bir model elde edilir. Bu modele regresyon eşitliği de denir.
- ◆ Bu model 'a' katsayısı ile bağımlı değişkeni tanımlayan bağımsız değişkenleri ve katsayılarını gösteren matematiksel bir denklemdir.
- ◆ Analiz sonucunda çoklu korelasyon katsayısı (r)ve tamamlayıcılık katsayısı da (R^2) hesaplanabilir.
- ◆ Buna ek olarak bağımlı değişkende olan değişimin ne kadarının hangi bağımsız değişken tarafından ortaya çıkarıldığı ve her aşamada modele giren bağımsız değişkenin r ve R^2 'ye ne kadar katkıda bulunduğu da incelenebilir.
- ◆ Model ne kadar iyi olursa bağımsız değişkenlerin bağımlı değişkeni tanımlama yüzdeleri (R^2) de o kadar büyük olur.

Çoklu Korrelasyon ve Regresyon Analiziyle Model Geliştirmede Kullanılabilecek Yöntemler:

- ◆ 1. Tüm olası regresyon eşitlikleri (all possible regression)
- ◆ 2. Geriye doğru eleme (backward selection)
- ◆ 3. İleriye doğru seçim (forward selection)
- ◆ 4. Adım adım regresyon (Stepwise regression)
- ◆ Bütün bu yöntemler için bilgisayar programları geliştirilmiştir. Bu sayıda çok sayıda bağımsız değişken incelenebilir.

Adım Adım Regresyon

- ◆ En gelişmiş yöntem olarak bilinir.
- ◆ 1. Bağımlı değişkenle en yüksek korrelasyon katsayısına sahip bağımsız değişken ilk olarak denkleme alınır.
- ◆ 2. Geriye kalan değişkenlerle bağımlı değişken arasında kısmi korrelasyon katsayısı hesaplanır ve katsayısı en yüksek olan değişken denkleme alınır.
- ◆ 3. Her aşamada denkleme son giren değişkenin modelde kalmasına gerek olup olmadığı ve son giren değişken modelde iken önceki değişkenlerin modelde kalmasına gerek olup olmadığı test edilir. Gereksiz olan model dışı bırakılır. Denkleme herhangi bir adımda alınan bir değişken herhangi bir adımda denklem dışı bırakılabilir.
- ◆ 4. Sonuçta denkleme kalması gereken değişkenler ve katsayıları hesaplanarak işlem bitirilir.

Kısmi Korrelasyon Analizi

- ◆ Bir değişken birden çok değişkeni etkileyebilir. Ör: Yaş değişkeni; hem çocuğun boyunu, hem ağırlığını, hem zeka düzeyini, hem de psikomotor yeteneğini etkiler.
- ◆ Böyle bir durumda bir araştırmacı yaş değişkenini düşünmeden zeka düzeyi ile psikomotor yetenek arasında bir ilişki aradığında pozitif yönde yüksek bir ilişki bulabilir.
- ◆ Bu değişkenler arasındaki ilişki neden-sonuç ilişkisi değildir, ikincil bir ilişkidir. Yani bu değişkenler yaşa bağlı olarak artmaktadır.
- ◆ Bu tür ilişkilerde etkileyici değişkenin etkisi ortadan kaldırılarak diğer değişkenler

arasındaki ilişkiler incelenebilir.

- ◆ Bu inceleme ‘Kısmi Korrelasyon Analizi’ ile yapılır.

SPERMAN SIRA (RANK) KORRELASYONU

- ◆ Parametrik olmayan bir korrelasyon yöntemidir.
- ◆ Bu yöntemde korrelasyon katsayısı x ve y değişkenlerinin gerçek değeriyle değil, bu değerlerin sıra numaralarıyla hesaplanır.
- ◆ x değişkenindeki sıralamanın y değişkenindeki sıralamaya nasıl bir uygunluk sağladığı aranır.

ZAMAN SERİLERİ ANALİZİ

- Bir olay hakkında belirli zaman aralıklarında elde edilen gözlemler ‘Zaman Serileri’, bu gözlemlerin zaman içerisinde değişimlerinin incelenmesi de ‘Zaman Serileri Analizi’ olarak tanımlanır.
- Bazı hastalıkların görülme sıklıklarının, ölüm ve doğum hızlarının, doktora ya da hastaneye başvuru sayılarının belirli zaman birimleri içinde nasıl bir değişime uğradığının incelenmesi zaman serileri analizinin sağlık bilimlerinde kullanımı için verilebilecek birkaç örnektir.
- Zaman serileri analizi ile hem bir olayın belirli zaman birimleri içinde nasıl bir değişime uğradığı incelenebilir, hem de bu değişimlerden yararlanılarak ileriye yönelik tahminler yapılabilir.
- Bu tahminler sağlık hizmetlerinin planlanmasında kullanılır.

ZAMAN BİRİMLERİ İÇİNDE GÖRÜLEBİLECEK DEĞİŞİMLER

- Uzun Sürede Olan Değişimler
- Bir olayın uzun bir süre içinde gösterdiği değişimlerdir.
- Ör: Bir bölgedeki bebek ölüm hızının, herhangi bir bulaşıcı hastalığın görülme sıklığının 10 - 20 yıl içinde nasıl bir değişim gösterdiğinin incelenmesi gibi.

Mevsimlik Değişimler

- Bir olayın bir yıl içinde mevsimlere göre gösterdiği değişimlerdir.
- Ör: Kızamık hastalığının aylık dağılımı vs...

Dönem Dönem Görülen Değişimler

- Bir olayın birkaç yıllık bir sürede; düzenli ya da düzensiz aralıklarla artması, durgunlaşması, tekrar artması gibi değişimler göstermesidir.
- Ör: Türkiye’de sıtma hastalığı

Geçici Ya Da Rastlantıya Bağlı Değişimler

- Bu değişimler her zaman görülmez. Arada sırada ortaya çıkar ya da bir kez görülebilir.
- Deprem, sel, harpler, grevler vb. olaylara bağlı olarak görülen değişimler bu sınıfta toplanır.
- Ör: İkinci dünya savaşında Avrupa ülkelerinde ölümlerin artması, doğumların çok azalması

ANALİZ YÖNTEMLERİ

- Elde edilen veriler önce bir grafik üzerinde noktalanır ve noktalar kırık çizgi ya da eğrilerle birleştirilir.
- Grafiğin x eksenine süre (yıl, ay, mevsim, hafta vb.), y eksenine ise incelenen değişkene ait sayı, oran ya da hızlar konur.
- İncelenen süre içinde bir olayda görülen değişimler, düzensizlikler (inişler, çıkışlar, durgunluk vb.) gösterebilir. Bu nedenle ilerisi için yapılan tahminlerde bu düzensizliklerin yumuşak bir eğri ya da doğru ile ifade edilmesine gereksinim vardır. Bunu sağlamak için çeşitli yöntemler

geliştirilmiştir.

Göz Kararıyla Çizim

- Grafik üzerinde işaretlenen ve birleştirilen noktaların durumuna göre iniş ve çıkışlar dikkate alınarak göz kararı ile noktaları en iyi temsil edeceği düşünülen bir eğri ya da doğru çizilir.
- Bu yöntemde doğrunun geçeceği yerler için hiçbir hesaplama yapılmadığından aynı dağılım için değişik kişilerce değişik çizimler yapılabilir. Bu da her seferinde değişik tahminler elde edilmesine neden olur.
- Pratik olmasına rağmen bilimsel olmayan bir yöntemdir.

Hareketli Ortalama Yöntemi

- Bu yöntemde, elde edilen gözlemlerin ardışık gruplar halinde ortalamaları alınarak iniş ve çıkışlar daha düzgün bir biçime getirilmeye çalışılır.
- Ör: 1964-1978 yıllarında sağlık ve Sosyal Yardım Bakanlığına bildirilen kızamık vaka sayıları şöyledir:

Yıl	Kızamık Sayısı	Yıl	Kızamık Sayısı	Yıl	Kızamık Sayısı
1964	17312	1969	66111	1974	12836
1965	52617	1970	46761	1975	24347
1966	50614	1971	43002	1976	21740
1967	42906	1972	23601	1977	16123
1968	38266	1973	43271	1978	12517

- Yıllar kolon 1'e, kızamık sayıları kolon 2'ye yazılır.
- Kaçar yıllık gruplar halinde ortalamalar alınacağına karar verilir. Örneğimizde 3'er yıllık ortalama alınmıştır.
- İlk adımda 1964, 65 ve 66 yılları vaka sayıları toplanır ve 3'e bölünerek elde edilen ortalama bu yılların ortası olan 1965 yılının karşısına yazılır.
- İkinci adımda 1964 yılı dışarıda bırakılarak 1965, 66 ve 67 yılları vaka sayıları toplanır ve 3'e bölünür. Elde edilen ortalama 1966 yılının karşısına yazılır.
- Diğer adımlarda her seferinde bir üstteki gözlem dışarıda bırakılarak ve bir sonraki gözlem ortalamaya dahil edilerek işleme devam edilir.

- Kolon 3'e üç yıllık kızamık sayısı toplamı, kolon 4'e üç yıllık ortalamalar yazılır.

Yıl	Kızamık Sayısı	3 Yıllık Toplam	3 Yıllık Ortalama
1964	17312		
1965	52617	120543	40181
1966	50614	146137	48712

- Ortalamalar grafikte karşılarında bulundukları yılların hizasına noktalanır ve noktalar düz çizgilerle birleştirilir. 1964 ve 1978 yıllarındaki gözlemler analiz dışında kalmıştır.

Bu Yöntemin Kullanılmasında Şu Noktalar Gözönünde Bulundurulmalıdır:

- 1. Aritmetik ortalama kullanıldığı için aşırı değerlerden etkilenebilir. Bu nedenle hareketli ortalamalar da inişli çıkışlı bir durum gösterebilir.
- 2. Baştan ve sondan bir kısım bilgiler dışarıda bırakılmaktadır. Üç yıllık ortalamalar alınıyorsa baştan ve sondan 1'er yıl, 5 yıllık ortalamalarda 2'şer yıl, 7 yıllık ortalamalarda 3'er yıllık gözlem dışarıda kalmaktadır.
- Hareketli ortalama ile ham verilerin yumuşatılması amaçlanmaktadır.

Yarım Ortalama Yöntemi

- Aynı örnek üzerinde bu yöntemi inceleyelim.
- Gözlemler iki eşit gruba ayrılır. Grup sayısı çift ise iki eşit gruba ayrılır. Grup sayısı tek ise, bu

kez en ortadaki gözlem hesaba katılmaz. Diğer alt ve üst gözlemler iki gruba ayrılır.

- Örneğimizde 15 yıllık gözlem var. Bunun için en ortadaki yıl (1971 yılı) hesaba katılmayacaktır. 1964-1970 yılları bir grup, 1972-1978 yılları diğer grup olacaktır.
- Her iki gruptaki vaka sayıları toplanıp grup sayısına bölünerek her iki grubun ayrı ayrı ortalaması alınır.
- 1964-1970 yıllarında toplam vaka sayısı= 314587, yıl sayısı= 7, ortalama= 44941
- 1972-1978 yıllarında toplam vaka sayısı= 154435, yıl sayısı= 7, ortalama= 22062
- Her iki grubun ortalaması kendi grubunun ortasına gelen yılın karşısına grafikte noktalanır ve bu iki noktadan geçecek bir doğru çizilir.
- 1964-1970 grubunun orta yılı= 1967, 1972-1978 grubunun orta yılı= 1975
- Bu yöntemle hesaplama işlemleri kolaydır. Eğilim doğrusal ya da doğrusala yakın ise uygulanabilir. Aritmetik ortalamanın aşırı değerlerden etkilenebileceği unutulmamalıdır.

En Küçük Kareler Yöntemi (Regresyon Analizi)

- Bu yöntemde eğilim matematiksel olarak regresyon eşitliği ile tanımlanır. Bu nedenle tümüyle objektif ve güvenilir bir yöntemdir.
- Aynı örnek üzerinde işlemleri yapalım:
- 1. İlk kolona yıllar yazılır.
- 2. Sıfırdan başlanarak ve 1'er arttırılarak yıllara numara verilir ve x kolonuna yazılır. Bu sayılar toplanır ($\sum x$).
- Her yıldaki kızamık sayısı (y) 3'üncü kolona yazılır ve toplanır ($\sum y$).
- İkinci kolondaki x değerlerinin karesi alınarak kolon 4'e yazılır ve toplanır ($\sum x^2$)
- x ve y değerleri çarpılarak kolon 5'e yazılır ve toplanır ($\sum xy$).

En Küçük Kareler Yöntemi İle Hesaplama İşlemleri

Yıl	x	y	x^2	xy
1964	0	17312	0	0
1965	1	52617	1	52617
1966	2	50614	4	101228
1967	3	42906	9	128718
.....	...	..		...
1978	14	12517	196	175238
	105	512024	1015	2967903

- $y = a + bx$ regresyon denklemindeki a ve b katsayıları hesaplanır ve regresyon doğrusu çizilir.

• Örneğimizde $a = 49542$ ve $b = -2201$ bulunmuştur. Buna göre doğrunun denklemi $y = 49542 - 2201x$ 'dir.

BİYOMEDİKAL DENEYLERDE İSTATİSTİKSEL YÖNTEMLER

- Verilen bir uyarıcıya (ilaç, vitamin vb.) canlının verdiği cevap yardımıyla uyarıcının tabiatı, yapısı ve kuvvetinin tahmin işlemlerine biyolojik deney denir.
- Değişik dozların ya da etkenlerin etki değişikliklerini incelemek için yapılan deneyler biyolojik deney değildir. Ör: Bir ilacın dört değişik dozunun ortaya çıkaracağı değişik cevapların incelenmesi ve karşılaştırılması biyolojik deney değildir. Ancak bu ilacın herhangi bir dozunun ortaya çıkarttığı etkinin o ilacın standart dozunun ortaya çıkarttığı etki ile karşılaştırılması biyolojik deneydir.

Bir biyolojik deneyde üç unsur vardır:

- **1. Uyarıcı (etken):** Etkisi ölçülmek üzere canlıya verilen maddedir. Uyarıcı; bir ilaç, toksik bir madde, vitamin, hormon vb. olabilir. Uyarıcının dozu ve yoğunluğu araştırmacının isteğine ve incelenen konunun özelliğine göre değişik olabilir.
- **2. Denek:** Uyarıcının uygulandığı canlıdır. Denek; bir hayvan, bir bitki, bir bakteri, bir doku, bir bakteri kültürü olabilir.
- **3. Cevap:** Uyarıcının denekte ortaya çıkardığı reaksiyondur. Bu cevap ölçümle belirtilebildiği gibi nitelik olarak da (evet-hayır, canlı-ölü gibi) belirtilebilir.
- Bir biyolojik deneyin sonuçlarının geçerli olabilmesi için deneyin daha çok iyi planlanması, yürütülmesi ve uygun istatistiksel yöntemlerle değerlendirilmesi gerekir.

POTANS TAHMİN YÖNTEMLERİ

- Potans tahmini, test preparatının standart preparata göre etkinliğinin saptanmasıdır.
- Diğer bir deyişle, canlının test preparatının belirli bir dozuna verdiği cevabın, o uyarıcının bilinen standart dozuna verdiği cevapla karşılaştırılarak her iki preparatın aynı miktarda etkin madde içerip içermediğinin incelenmesidir.
- Potans tahmini için değişik deney düzenleri vardır:

EŞİK DOZ DENEYİ

- Eşik doz, bir uyarıcının uyarıyı ortaya çıkardığı noktadaki değer (doz, miktar vb.) dir.
- Test preparatı ve standart preparatın uygun görülen dilüsyonları hazırlanır. Bu dilüsyonlar hayvanlara beklenen uyarı ortaya çıkıncaya kadar enjekte edili ve uyarı ortaya çıktığında verilen miktar kaydedilir. Hayvanlar değişik ağırlıkta olabileceği için her hayvana verilen miktar kendi vücut ağırlığına bölünerek düzeltilir ve potans tahmini bu düzeltilmiş değerler üzerinden yapılır. Potansın güven sınırları bulunur.

ÜÇ NOKTA DENEYİ

- Üç nokta deneyinde 2 standart ve 1 test preparatı kullanılır.
- Standart preparatın birisi test preparatından yüksek dozda, diğeri düşük dozda seçilir. Pratik olarak test preparatı dozunu iki standart dozun ortalaması gibi düşünebiliriz.
- Her üç dozdan elde edilen cevap ortalamaları grafikte dozların logaritmalarına karşı noktalanır.
- Üç nokta deneyinde test preparatı tek dozda kullanıldığından deneyin güvenilirliği azalır. Bu nedenle yüksek düzeyde doğruluk aranmadığında ya da test solüsyonları ve materyali çok az miktarda olduğunda kullanılmalıdır. Dört nokta deneyi daha çok kullanılan ve daha güvenilir sonuç veren yöntemdir.

DÖRT NOKTA DENEYİ

- Dört nokta deneyinde 2 standart ve 2 test preparatı kullanılır.
- Her dört dozdan elde edilen cevap ortalamaları grafikte dozların logaritmalarına karşı noktalanırsa, birisi test preparatları, diğeri standart preparatlar için olmak üzere iki doğru elde edilir.
- Bu doğruların birbirine paralel olması beklenir. Ancak pratikte bu iki doğrunun eğimleri çoğu kez paralel değildir.
- Hesaplama işlemlerini kolaylaştırmak için test preparatlarının dozları arasındaki oran standart preparatların dozları arasındaki orana eşit olmalıdır.

ALTI NOKTA DENEYİ

- Bu deneyde üç değişik dozda test preparatı ve aynı dozlarda üç standart preparat kullanılır.
- Her doza alınan cevap ortalamaları grafikte log dozlara karşı noktalanırsa birisi test preparatları ve diğeri standart preparatları için iki doğru elde edilir.
- Bu doğruların paralel olması, başka bir deyişle eğimlerinin aynı olması beklenir. Ancak çoğu kez bu doğrular paralel olmayabilmektedir.
- **Bu deneyde dikkat edilmesi gereken en önemli iki nokta şudur:**
- 1. Test ve standart preparatların dozları birbirine eşit olmalıdır.
- 2. Yüksek dozun orta doza oranı ile orta dozun düşük doza oranı birbirine eşit olmalıdır.

LD₅₀ TAHMİN YÖNTEMLERİ

- LD₅₀ insektlerin % 50'sini öldüren doz (lethal dose 50) anlamına gelir. LC₅₀ olarak da adlandırılmaktadır (lethal concentration 50).
- LD₅₀'nin ölçü olarak seçilmesinin nedeni, diğer dozlardan daha doğru sonuç vermesidir. Başka bir deyişle, deneysel hata LD₅₀'de diğerlerinden daha azdır.
- Bir bölgede insektisit uygulamasından önce insektlerin insektisitlere karşı ne derece duyarlı oldukları ölçülmelidir. Bu ölçüm yapılırsa hem insektisit uygulamasından önceki ve sonraki duyarlılık, hem de değişik bölgelerdeki ve değişik türdeki insektlerin duyarlılıkları karşılaştırılabilir. Böyle bir karşılaştırma yapabilmek için; deneylerin düzenlenmesinde, yöntemlerde, verilerin kayıt edilmesinde ve ölçümlerde birlik sağlanmalıdır. Bu yüzden Dünya Sağlık Örgütü sivrisinekler için yapılan duyarlılık deneylerinde Busvine ve Nash metodunu standart metod olarak kullanılmasını önermiştir.

Busvine ve Nash metodu

- Busvine ve Nash metodunda sivrisinekler, içinde değişik dozda insektisit emdirilmiş filtre kağıdı bulunan tüplerde 27 °C'de 1 saat bekletilir, sonra tüplerden alınarak 27 °C'de tutulan kafeslere konur. 24 saat sonra ölenler sayılır. Busvine ve Nash metoduyla yapılan deneyde kontrol grubunun da kullanılması uygun olur. Bu nedenle bir grup sivrisinek de içinde insektisit bulunmayan bir madde emdirilmiş filtre kağıdıyla tüp içinde ve kafeslerde deney gruplarıyla aynı süre bekletilir. Kontrol grubunda olacak ölümler insektisite bağlı değil, doğal ölüme bağlı olacaktır. Bazı durumlarda kontrol grubundaki ölüm hızı, deney grubunun en düşük dozundaki ölüm hızından daha büyük olabilir. Böyle bir durum kontrol grubu ölüm hızı yerine en düşük dozdaki ölüm hızı düzeltme için kullanılmalıdır.

Grafik Yöntemiyle LD₅₀ Tahmini

- LD₅₀ tahmininde en kolay yol, elde edilen ölüm hızlarını bir grafik üzerinde noktalamak ve aşağıda açıklanan biçimde sivrisineklerin % 50'sini öldüren dozu tahmin etmektir.
- Bu yöntemle LD₅₀'nin tahmin edilmesi kolaydır. Ancak eğri elle çizildiği için aynı veri kullanıldığı halde değişik kişilerce değişik eğriler çizilebilir. Bu nedenle her seferinde değişik bir tahmin elde edilebilir.

Probit Log Grafik Yöntemiyle LD₅₀ Tahmini

- Kişisel farklılıkları ortadan kaldıran bu yöntemde "Probit Log" grafik kağıdı kullanılır.

- Grafikte x eksenine dozların logaritmaları konur.
- Ölüm oranları probit dönüşüm tablosundan yararlanılarak probit değere dönüştürülür.
- Probit log grafik yöntemi kişisel farklılıkları büyük oranda ortadan kaldırdığı için grafik yönteminden daha güvenilirdir.

REGRESYON ANALİZİYLE LD₅₀ TAHMİNİ

- Bu analizde LD₅₀'nin tahmini doğal ölüm olup olmamasına göre değişik işlemlerle yapılır.
- Doğal ölüm yoksa LD₅₀ tahmini sonucunda elde edilen değer grafik ve probit log grafikten elde edilen değerlere çok yakındır.
- Doğal ölümler olduğunda LD₅₀ tahmini sonucunda elde edilen değer probit log grafikten elde edilen değere çok yakındır.

