**Araştırma** ; belli bir amaca yönelik sistemli bir şekilde veri toplama, verileri analiz etme ve yeni bilgiler ortaya çıkarma sürecidir. Araştırmada izlenecek yol, en başından son sonuna kadar sistemli ve planlı bir şekilde tasarlanmalıdır.

**Araştırma Türleri :**

. Nitel ve Nicel Araştırma

. Deneysel Araştırmalar

. Tanımlayıcı Araştırmalar

. Kesitsel Araştırmalar

. Kohort Araştırmaları

**İstatistik ;** olguların ve olayların sayısal verilere dönüştürülerek ifade edilmesi olarak tanımlanır.

**Evren (Ana Kitle);** araştırmanın konusuna göre değişiklik gösteren evrensel kümedir. Örneğin Türkiye’deki lise öğrencileri “X” araştırmasının evreni, Ankara’daki lise öğrencileri de “Y” araştırmasının evreni olabilir.

**Örneklem;** evreni temsil edebilecek olan daha küçük bir kümedir. Evreni oluşturan tüm elemanlara ulaşmanın mümkün olmadığı durumlarda örneklem almaya ihtiyaç duyulur.

**Anket;** belli bir amaç için sistemli bir şekilde veri toplamak üzere araştırmacı tarafından geliştirilen araçtır.

. Anketin başına araştırmanın amacını belirten kısa bir açıklama konulmalıdır.

. Anket soruları araştırmanın konusuna yönelik kısa, net ve anlaşılır olmalıdır.

. Katılımcıların anketi bırakmaması için anket çok uzun tutulmalıdır.

. Anket ön uygulama yapılıp eksiklikleri giderildikten sonra araştırma grubuna uygulanmalıdır.

. Anket sosyo-demografik bölüm ve araştırma konusu ile ilgili bölüm olmak üzere iki bölümden oluşmalıdır.

**Hipotez;** deneylerle henüz yeteri derecede doğrulanmamış ancak doğrulanması umulan teorik düşünce, varsayım olarak tanımlanır.

. Bir araştırma probleminde **bağımlı değişken** ile **bağımsız değişken** arasındaki ilişkiyi ifade eder.

. Erkekler evlenince daha tutumlu olur. (Bağımlı değişken tutumluluk, bağımsız değişken evlilik)

. **H0 (Yokluk Hipotezi);** iki grubun ortalamaları arasında fark yoktur.

. **H1 (Alternatif Hipotez);** iki grubun ortalamaları arasında fark vardır. (Çift yönlü hipotez) 🡪 Hangi grup olduğu belirtilmiyorsa.

. **H1 (Alternatif Hipotez);** birinci grubun ortalaması ikinci grubun ortalamasından daha büyüktür. (Tek yönlü hipotez)

**Değişken ;** nesnelerin ya da kişilerin birden çok değer alabilen özelliklerine denir.

. **Bağımsız Değişken;** olayları ya da durumları etkileyen değişkendir.

. **Bağımlı Değişken;** olaylardan veya durumlardan etkilenen değişkendir.

. Bağımlı Değişken (Neden Değişkeni) 🡪 Bağımlı Değişken (Sonuç Değişkeni)

Örneğin; obezite(Bağımsız Değişken) sonucu kalp ve damar rahatsızlığı(Bağımlı Değişken) meydana gelir.

**Kategorik Değişkenler**

. **Nominal Veri;** derecelendirme yoktur. (Cinsiyet, renk, meslek, kan grubu)

. **Ordinal Veri;** derecelendirme vardır. (Hastalık evresi, ağrı derecesi)

**Sayısal Değişkenler**

. **Kesikli Veri;** sayımla elde edilir. (Çocuk sayısı, kardeş sayısı)

. **Sürekli Veri;** ölçümle elde edilir. (Boy, kilo, aylık gelir)

**Bağımlı ve Bağımsız Grup**

. **Bağımlı Grup;** tek bir grup vardır.

Erkeler 🡪1.Ölçüm 🡪Eğitim 🡪2.Ölçüm 🡪 Erkeklerin 1.ve 2.ölçüm puanlarının karşılaştırma

. **Bağımsız Grup;** birden fazla grup vardır.

Erkekler

1.Ölçüm 🡪Eğitim 🡪2.Ölçüm 🡪 Kadınlar ve erkeklerin son ölçüm puanlarını karşılaştırma

Kadınlar

**İstatistiksel Anlamlılık;** gruplar arası bulunan ilişki ya da fark şans eseri mi yoksa gerçek olarak mı bulunmuştur sorusuna açıklık getirmektedir.

. **”p” değeri;** iki grubu karşılaştırmak için yapılan bir istatistiksel değerlendirmede iki grup arasında çıkan farkın şans eseri ortaya çıkma olasılığı göstermektedir. **p=0.03** ise çıkan fark %3 şans %97 araştırılan etkene bağlıdır.

**Veritabanı Oluşturma**

**Variable View:** Elimizde bulunan anketlerdeki değişkenleri temel olarak tanımladığımız yer.

****

**Name:** Kısaltılmış olarak anket içinde değişkenlerin isimlerinin girilmesi (İngilizce karakter ve boşluksuz.)

**Type :** Girilecek olan değişkenin tipi ( Sayısal, virgüllü, tarih, para birimi vs.)

**Width :** Kutucuğun genişliği

**Decimals :** Virgülden sonra gösterilecek olan basamak sayısı

**Label :** Name kısmında girilmiş olan değişken isminin detaylı olarak bu kısımda açıklamalı şekilde yazılması.

**Values :** Anket içindeki ifadelerin değer olarak tanımlanması ( Örneğin ; Erkek=1 Kadın=2 )

**Missing :** Kayıp veriler olduğu zaman o verilerin yerine gelmesi gerekenler

**Columns :** Sütun genişliği

**Align :** Yazının hangi tarafta olacağı ( Sağ ya da Sol )

**Measure :** Verinin sayısal , ordinal veya nominal oluşuna göre ifadelendirilmesi.

**Verileri Kontrol Etme ve Düzenleme**

Veritabanını Özet Olarak Görmek

**File 🡪 Display Data File İnformation 🡪 Working File**

**Utilities 🡪 Variables**

String olarak girilen ifadelerin analizi yapılamıyor.

**Hatalı Verilerin Bulunması**

Küçük olan veritabanlarında hatalı verilerin bulunması işlemi kolaydır. Sort Ascending (sütün üstünde sağ tık ) yaparak girilen verilerin kontrolü sağlanabilir. Fakat büyük verilerin olduğu datasetlerde minimum ve maksimum değer kontrolü yapılır.

**Analyze 🡪 Descriptive Statistics 🡪 Descriptives**

Çıkan kısımda anket içinde bulunan değişkenlerden kontrolünü yapılmasını istediğimiz kısımları seçiyoruz. Daha sonraki kısımda minimum ve maksimum değer kısımları seçilip işleme devam ediyoruz. Sonra karşımıza çıkan output kısmında Descriptives içinde seçmiş olduğumuz değişkenlerin maksimum ve minimum değerleri görüyoruz. Burda eğer yanlışlık içeren değerler var ise dataset içinde o kısmın kolonuna gelerek Ctrl + F ya da Edit kısmından Find diyerek gelen ekrandan yanlışlık olan değere giderek onu uygun şekilde düzenliyoruz.

**SPSS’de Tanımlayıcı İstatistikleri Hesaplama**

**Tanımlayıcı İstatistikler**

**Merkezi Eğilim Ölçüleri:**

**Aritmetik Ortalama ( Mean ) :** Bir değişkenin değerlerinin toplamının gözlem sayısına bölünmesiyle elde edilir.

**Ortanca ( Median ) :** Sıralanmış veri setinde, veri setini iki eşit parçaya bölen tam ortadaki değerdir.

**Tepe Değeri ( Mod ) :** Bir veri kümesinde en sık görülen değerdir.

**Merkezi Dağılım Ölçümleri :**

**Değer Aralığı ( Range ):** En büyük ve en küçük değerler arasındaki farktır.

**Standart Sapma:** Her bir değerin ortalamaya olan uzaklığıdır.

**Standart Hata:** Aynı büyüklükte aynı popülasyondan seçilen örneklem dağılımıdır. Standart sapma değerinin denek sayısının kareköküne bölünmesi ile elde edilir.

**Varyans:** Bir veri grubundaki verilerin ortalamadan ne kadar saptığını gösterir. Standart sapmanın karesi alınarak hesaplanır.

**Yüzdelikler ve Çeyreklikler:** Bir veri kümesi değerinin göreli konumunu bulmak için kullanılır. 1.çeyrek(Q1) %25’lik dilimi, 2.çeyrek(Q2) %50’lik dilimi ve 3.çeyrek(Q3) %75’lik dilimi ifade eder.

**Çarpıklık ( Skewness ) :** Bir dağılımın simetrik olup olmadığının bir ölçüsüdür. Çarpıklık değerinin sıfır (0) olması simetrik bir dağılım olduğunu göstermektedir. Kabul edilebilir çarpıklık değeri -2 ile +2 arasındadır.

**Basıklık/Diklik ( Kurtosis ) :** Dağılımın ne kadar sivri ya da düz olduğunun bir ölçüsüdür. Tam bir çan eğrisinde diklik katsayısı sıfır (0)’dır. Kabul edilebilir çarpıklık değeri -2 ile +2 arasındadır.

**Kategorik: Analyze 🡪Descriptives Statistic 🡪Frequencies** dedikten sonra ; Reset diyerek kategorik değişkenleri seçip Statistik diyerek Mode, Skewness, Kurtosis ve Quartiles seçeneklerini işaretleyerek bir output alabiliriz.

**Sayısal: Analyze 🡪Descriptives Statistic 🡪Frequencies** dedikten sonra ; Reset diyerek sayısal değişkenleri seçip Options diyerek Mean, Std, Variance, Mode, Skewness, Kurtosis ve Min-Max seçeneklerini işaretleyerek bir output alabiliriz.

Kategorik değişkenleri seçerek görsel chartları oluşturmak için 🡪 **Analyze 🡪Descriptives Statistic 🡪Frequencies** dedikten sonra istenilen kategorik değişkenler seçildikten sonra **Charts** tıklanarak oluşturulmak istenen tür (Bar, Pie, Histograms) seçilerek output alınır.

**Yeni Veri Oluşturma**

**1.Select Cases:** Belli bir durumu seçme işlemidir. Örneğin; bir dataset içinde bulunan değişkenler içinden sadece bir durumu seçip onun analizlerini yapmaktır.

**Data 🡪 Select Cases**

Sol kısımda istenilen değişkeni seç 🡪 **if condition** 🡪 İstediğimiz durumun yazılımı (Örnek: cinsiyet =1)

Seçmiş olduğumuz değişken verilerini yeni bir dataset’e aktarmak için;

**Data 🡪 Select Cases 🡪 Copy selected cases to a new dataset + Dataset Name**

**2.Split File:** Kategorik verileri gruplara ayırır ve her analizi her grup için ayrı şekilde yapar. Örneğin; bir veri seti içinde bulunan değişkenlerden cinsiyet içindeki kadın ve erkeği ayırıp ayrı ayrı analiz değerlerini görmek.

**Data 🡪 Split File 🡪 Compare Groups** dedikten sonra sol kısımdaki değişkenlerden istenilen kategorik değişkeni seçiyoruz.

Daha sonra **Analyze 🡪 Descriptive Statistic** kısmı içinde seçilen kategorik değişkene göre sayısal değişken seçip çıktıdan sonuçları görebiliriz.

**3.Compute:** Yeni bir değişken elde etmek için kullanılan yöntemdir. Veri seti içinde bulunan birden çok değişkeni kullanarak yeni bir hesaplama yapılmak istenildiğinde kullanılır. Örneğin; boy ve kilo değişkenlerini kullanarak beden kitle endeksi hesaplamak.

**Transform 🡪 Compute** dedikten sonra yeni değişkenin ismini **Target Variable** içine yazıyoruz ve değişkenleri seçerek **Numeric Expression** içinde işlemi gerçekleştiriyoruz.

**4.Recode:** Elimizdeki bir veriyi yeniden isimlendirmek veya kodlamak için kullanılır. Örneğin; bir sayısal veri içindeki değerleri belli aralıklarla atama yaparken ( kilo < 50 =1 , kilo 50-75 =2, kilo >75 =3 ). Ya da string olarak girilmiş bir verinin analizi yapılmayacağından dolayı onu numeric olarak çevirmek istediğimiz zaman bu komutu kullanılırız.

**Transform 🡪 Recode into Same Variables** – Seçilen kolon üstünde işlem yapar.

**Recode into Different Variables** – Yeni bir kolon oluşturur.

Daha sonra istenilen değişkeni seçiyoruz ve yeni sütun adıyla birlikte açıklamasını yazıyoruz**. Old and New Values** içinden eski değişken ve oluşturmak istediğimiz yeni değişkeni ekliyoruz. Sonra **Continue** diyoruz ve **Change** diyerek dönüşüm işlemini gerçekleştiriyoruz.

Sayısal bir veriyi aralıklarla kategorik şekle dönüştürmek istersek **Range** kısmından değer girerek New Value kısmından değer atamasını yapıyoruz.

**5.Count Value:** Belli rakamsal verileri olan bir veri setinde belirlenmiş olan ölçüm ya da değerin kaç defa gerçekleştiğini saymak istediğimiz zaman kullanırız.

**Transform 🡪 Count Values** dedikten sonra oluşacak yeni sütundaki değişkene isim verip, açıklamasını giriyoruz.

Sonra **Define Values** kısmından değer aralığının tanımlamasını yapıp **Add** diyoruz ve **Continue** seçip OK ile bitiriyoruz.

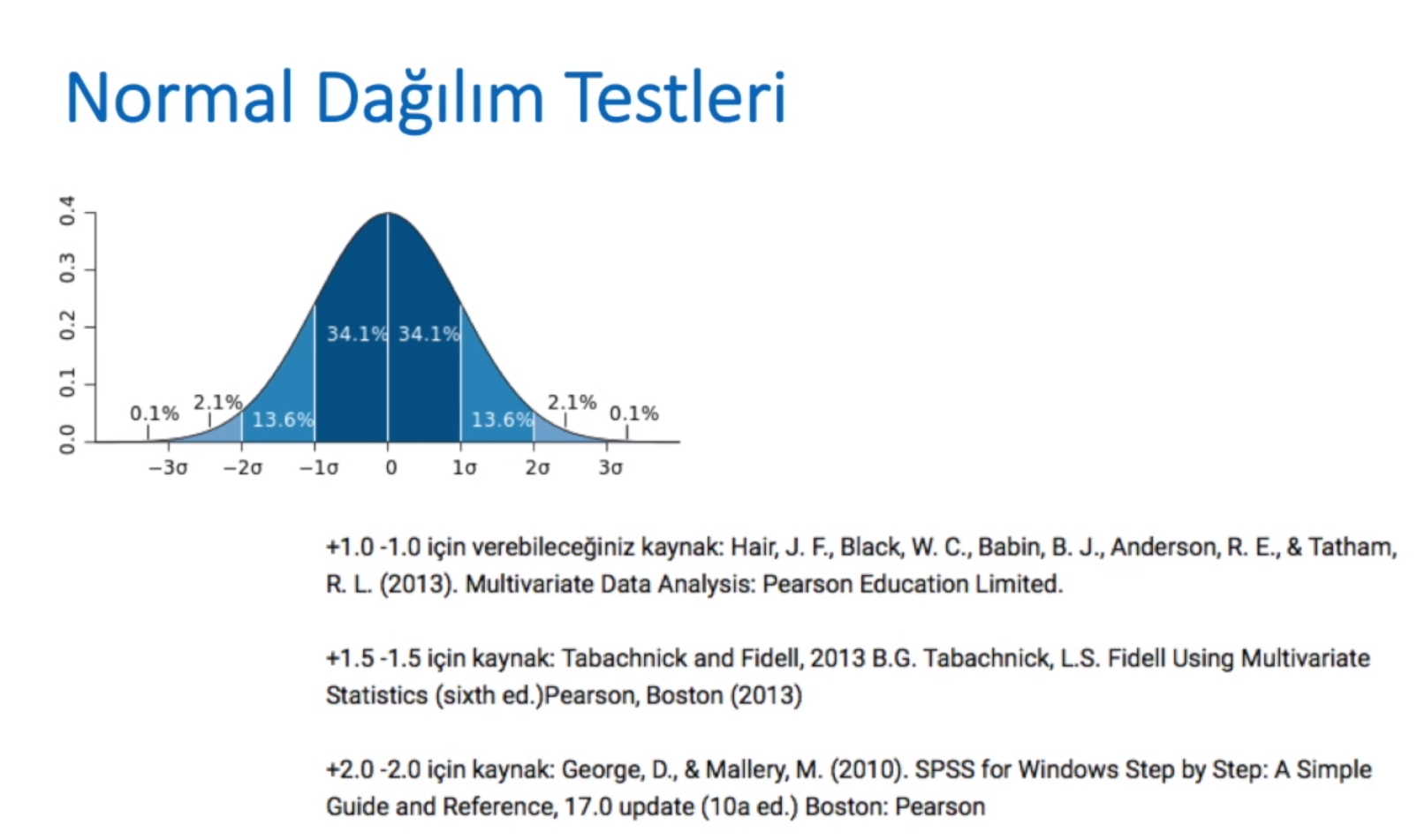
**6.Transpose:** Satırları sütuna, sütunları satıra çevirmek için kullanılır.

**Data 🡪 Transpose**

Değiştirilecek satır ve sütunları **Variable** ve **Name Variable** içine giriyoruz.

**Normal Dağılım Testleri**

Günlük hayatımızdaki pek çok özelliğin ortalama etrafında dağılım göstermesi **Normal Dağılım** olarak adlandırılmaktadır. Örneğin; ülkemizdeki 18 yaş üstü boy dağılımı, 1.60 ile 1.90 arasında değişmektedir. Bu belirtilen sınırlar dışından değerler olsa dahi çoğunluk kümelenmesi belirlenmiş değer arasındadır. Örneklem sayısının **30’**un üzerinde olması normal dağılımı ifade etmektedir.



**Analyze 🡪 Descriptive Statistic 🡪 Explore** dedikten sonra

**Plots 🡪 Normality plots with test** ve **Histogram** seçenekleri tıklayarak istenilen değişkenin normal dağılım olup olmadığına bakabiliriz.

Bir testin normal dağılım olup olmadığını kontrol ettiğimiz durumları vardır. Yapılan analizler sonucu ve histogram grafiği dağılımı ile bunları görebiliriz. Bakılacak değerler;

**. Çarpıklık ( Skewness )**

**. Basıklık/Diklik ( Kurtosis )**

**. p value değeri > 0.05**

**Temel SPSS Testleri**

**. One Simple T Test ( Evren Ortalaması Önemlilik Testi ) 🡪** Örneğin ; bir ekonomiste göre, bir bölgede kişi başına düşen gelir ülke ortalamasından azdır.

**. Independent Samples T Test ( Bağımsız Gruplarda T Testi )** 🡪 Örneğin ; bir ekonomiste göre, farklı iki bölgenin kişi başına düşen geliri arasında fark vardır. Nonparametrik Testi 🡪 **Mann-Whitney U Test**

**. Paired Samples T Test ( Bağımlı Gruplarda T Testi )** 🡪 Örneğin ; bir insan kaynakları yöneticisine göre, belirli bir eğitim programı çalışanların motivasyonunu artırıcı bir etkiye sahiptir. Nonparametrik Testi 🡪 **Wilcoxon Test**

**. Chi Square ( Ki Kare Test )**

**. One – Way Anova Test**

**. Kruskal-Wallis Test**

**. Korelasyon 🡪** Örneğin; Pazarlama Müdürü, fiyat indiriminin satış artışı ile ilgisi olduğunu iddia etmektedir.

**Uygun Teste Karar Verme**

**.** Değişkenin tipi sürekli mi (ölçümle elde edilmiş örneğin; boy) yoksa kategorik mi (örneğin; kadın-erkek)

**.** Gruplar bağımlı mı bağımsız mı?

**.** Değişkenin dağılımı normal mi değil mi?

**.** Ölçülmek istenen ne? (Gruplar arası fark mı, gruplar arasındaki ilişki mi?)

**.** Gruplardaki kişi sayısı 30’un üzerinde mi altında mı? (30’un üzerinde yapılan ölçümlerin normal dağıldığı varsayılır.)

**Parametrik Test Varsayımları**

**.** Veriler ölçüm ile elde edilmiş (aylık gelir durumu, boy uzunluğu)

**.** Örneklem sayısı > 30

**.** Verilerin normal dağılım göstermesi

**.** Varyanslar homojen

**(Tüm aşamaların sağlanması durumunda parametrik test uygulaması yapılır aksi durumda nonparametrik test uygulanır.)**

**Nonparametrik Test Varsayımları**

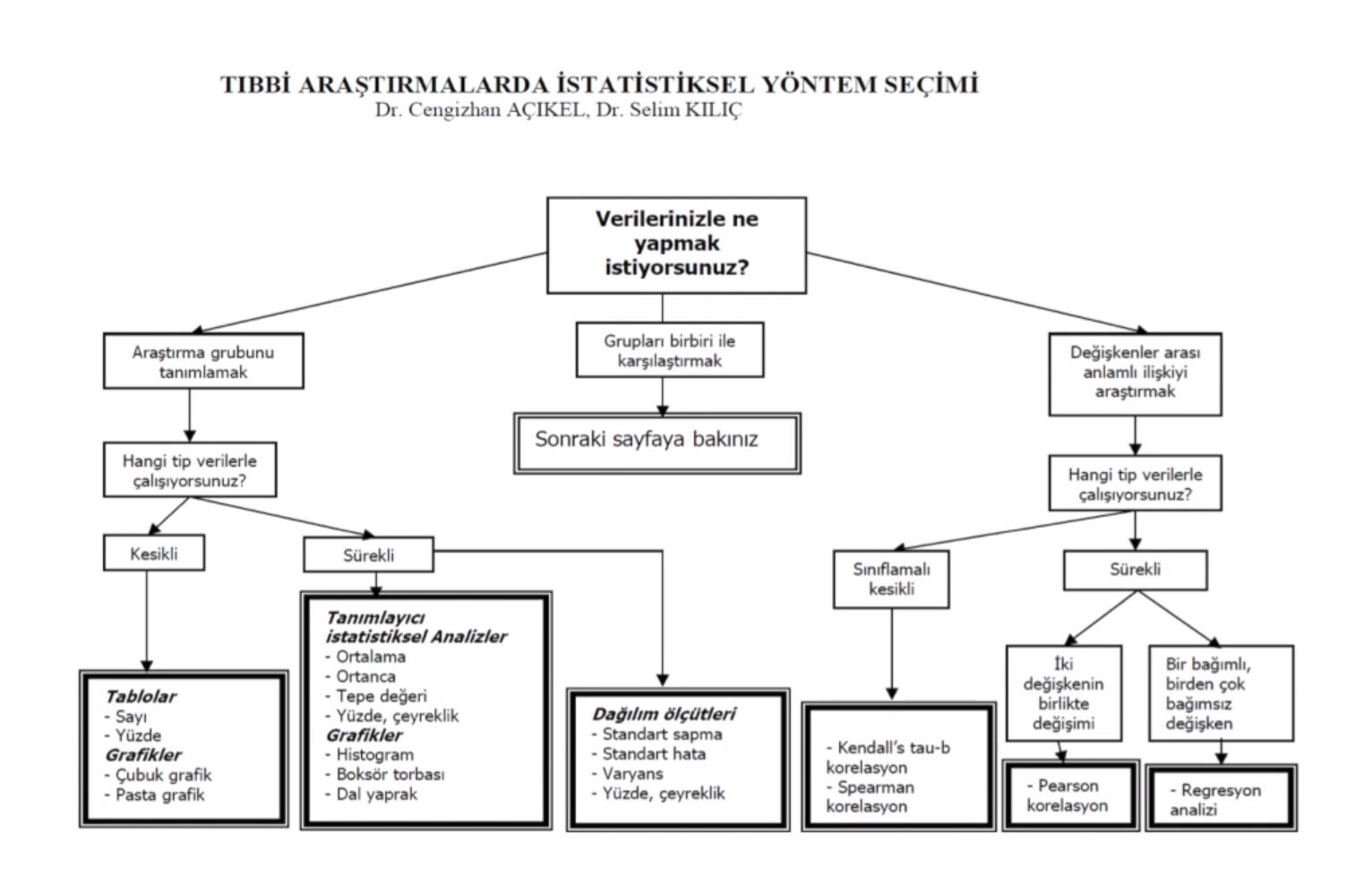
**.** Veriler sayım ile elde edilmiş ( Kategorik veri gibi Kadın-Erkek )

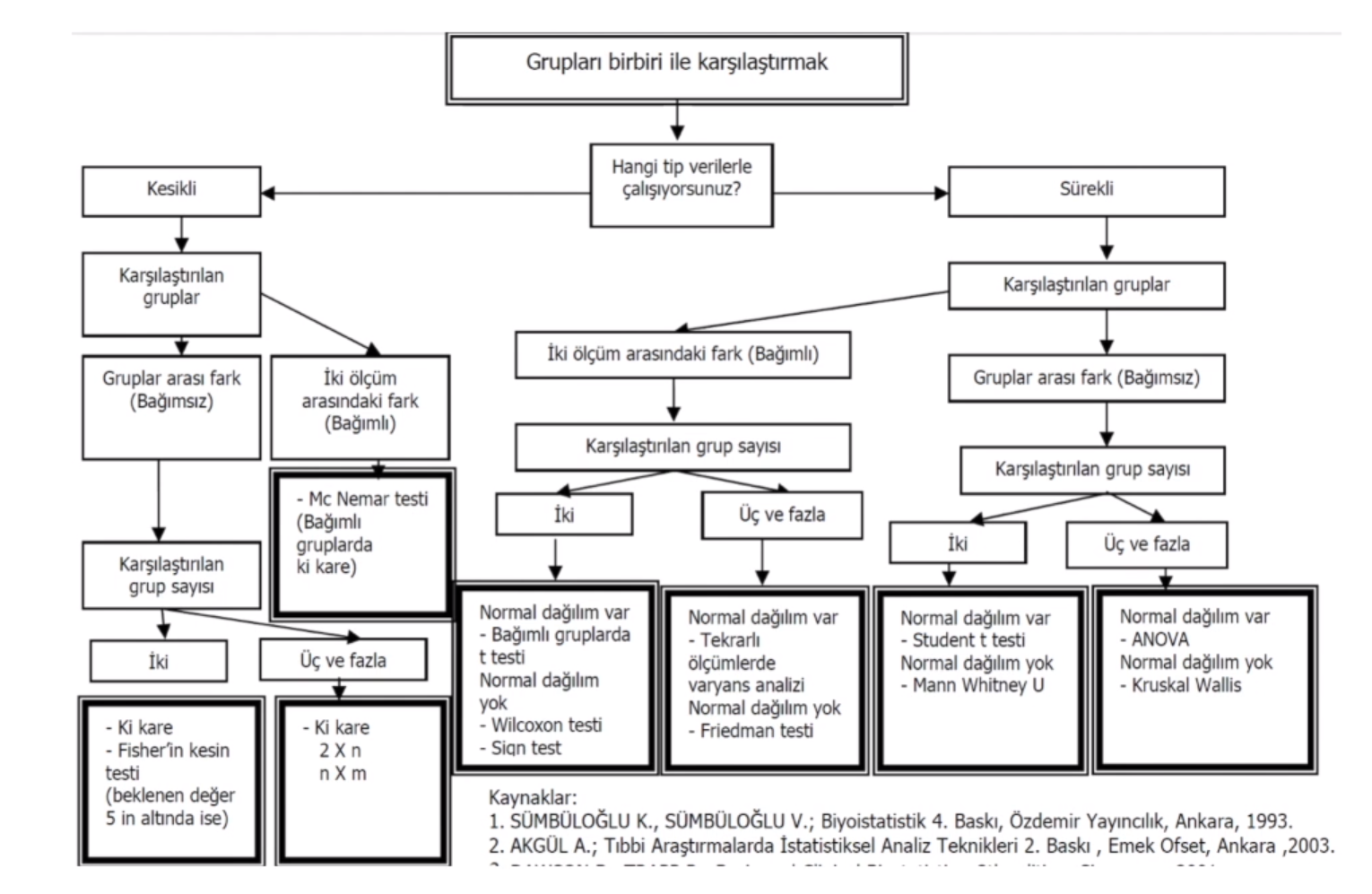
**.** Örneklem sayısı < 30

**.** Verilerin normal dağılım göstermemesi

**.** Varyanslar homojen değil







**TESTLER**

**1.One Sample T Test ( Evren Ortalaması Önemlilik Testi )**

Bu test, bir grupta ölçüm verisi ortalamasının, belli bir evren verisi ortalamasından farkının anlamlı olup olmadığını incelemek için kullanılır.

**Örnek Hipotez :**

**.** Bir ekonomiste göre, bir bölgedeki kişi başına düşen gelir, tüm ülkedeki kişi başına düşen ortalama gelirden azdır.

**.** Kalite kontrol departmanı, bir ürünün ortalama boyutlarının orijinal ürünle aynı olduğunu söylemektedir.

**.** Araştırma grubundaki kişilerin ortalama kolesterol düzeyi ile evren ortalaması olan 160 mg/dl kolesterol düzeyi arasında fark yoktur.

**Test Gereklilikleri :**

. Tek grup olmalıdır.

. Evren ortalaması bilinmelidir.

. Örnekle sayısı 30’dan fazla olmalıdır.

. Normal Dağılım olması gerekmektedir.

**Analyze 🡪Compare Means 🡪One Sample T Test**

İstenilen değişkeni **Test Variable** içine aktarıyoruz ve bilinen evren ortalaması değerini **Test Value** kısmına giriyoruz. Testi uygulama işleminden sonraki çıktı ekranında bulunan tablodan **Sig(2-Tailed)** değerine bakarak hipotezi reddetme ya da kabul etme durumunu görebiliriz.

**p value değeri < 0.05** ise hipotezi reddediyoruz ve farklılık olduğunu söylüyoruz.

**Araştırma Sorusu ?:**

Araştırma grubundaki kişilerin kolesterol düzeyi 160 mg/dl düzeyinde farklı mıdır?

**H0 Hipotezi :** Araştırma grubundaki kişilerin ortalama kolesterol düzeyi ile evren ortalaması bilinen kolesterol düzeyin arasında fark yoktur.

**2.Independent Samples T Test**

Bu test, iki ayrı bağımsız gruptan aynı konuda elde edilen iki ortalamanın karşılaştırılması amacı ile kullanılır.

**Örnek Hipotez:**

**.** Bir ekonomiste göre farklı iki bölgenin kişi başına düşen geliri arasında fark vardır.

**.** İki farklı işletme okulundan mezun olan öğrencilerin aldıkları maaş arasında fark vardır.

**.** Kadın ve erkeklerin kolesterol düzeyleri arasında fark vardır.

**Test Gereklikleri:**

**.** İki bağımsız grup

**.** Ölçümle elde edilmiş veri

**.** Her iki grupta da 30’dan fazla örneklem

**.** Normal Dağılım

**Analyze 🡪Compare Means 🡪Independent T Test**

**Test Variable** 🡪 İstenilen değişken

**Grouping Variable** 🡪 Karşılaştırılacak grup

Daha sonra **Define Group** içine datasette tanımlı olan grup değerlerini giriyoruz.

**Levene Test’teki** Sig > 0.05 ise gruplarının varyansı eşit kabul edilir ve t-test içindeki Sig(2-tailed) değerinin üst kısmına bakıyoruz.

**3.Manny Whitney U Test**

Bu test, Independent Sample T Test’in nonparametriğidir. Parametrik test varsayımını sağlamayan iki ayrı bağımsız gruptan aynı konuda elde edilen iki ortalamanın karşılaştırılması amacı ile kullanılır.

**Test Gereklilikleri :**

**.** İki bağımsız grup

**.** Ölçümle elde edilmiş veri

**.** Herhangi bir grupta 30’dan az örneklem sayısı

**.** Normal Dağılım yok.

**Araştırma Sorusu ? :**

Araştırmada koroner kalp rahatsızlığı olan erkekler ve kadınlar arasında kolesterol düzeyi açısından fark var mıdır ?

İlk olarak araştırma sorusunda istenilen özel duruma göre Dataset üzerinde bir indirgeme durumu gerçekleştiriyoruz. **Data 🡪 Select Cases 🡪 If Condition**

İndirgeme yani istenilen özel durumu seçtikten sonra sonuçları yeni bir Dataset’e kaydediyoruz. Karşımıza çıkan yeni Dataset üzerinde test işlemini gerçekleştirmek için;

**Analyze 🡪NonParametrik Test 🡪Legacy Dialogs 🡪 2 Independent Sample T Test**

Test işlemini yaptıktan sonra çıkan çıktı ekranında Sig(2-Tailed) değeri 0.05 den büyük ise ; araştırma sorusuna göre fark yoktur.

**4.Paired Sample T Test ( Bağımlı Gruplarda T Testi )**

Bu test, tek bir gruptan elde edilen iki ölçüm arasında fark olup olmadığına bakmak için kullanılır.

**Örnek Hipotez:**

**.** İnsan Kaynakları yöneticisine göre belirli bir eğitim programı çalışanların motivasyonunu artırıcı etkiye sahiptir.

**.** Uygulanan diyet programı hastaların total kolesterol düzeyini düşürücü etkiye sahiptir.

**.** Bir firmaya göre açık hava reklamının satışlar üzerinde etkisi yoktur.

**Test Gereklilikleri :**

**.** Tek bir grup

**.** Ölçümle elde edilmiş veri

**.** 30’dan fazla örneklem sayısı

**.** Normal Dağılım

**Araştırma Sorusu ? :**

Araştırmadaki kişiler 6 aylık sıkı bir diyet programına alınmışlar ve kolesterol ölçümleri tekrarlanmıştır. Diyet öncesi ve sonrası yapılan kolesterol ölçümleri arasında istatistiksel olarak anlamlı bir fark var mıdır?

**H0 Hipotez :** Diyet öncesi ve sonrası yapılan kolesterol ölçümleri arasında istatistiksel olarak anlamlı bir fark yoktur.

**Analyze 🡪Compare Means 🡪Paired Sample T Test**

**Variable 1** ve **Variable 2** --- Karşılaştırılacak olan değişkenler

**5.Wilcoxon Signed Ranks Test**

Bu test, Paired Sample T Test’nin nonparametriğidir. Parametrik test varsayımlarını sağlamayan bağımlı (tek) gruptan aynı konuda elde edilen iki ortalamanın karşılaştırılması ile kullanılır.

**Test Gereklilikleri :**

**.** Bağımlı tek grup

**.** Ölçümle elde edilmiş veri

**.** 30’dan küçük örneklem sayısı

**.** Normal Dağılım yok

**Araştırma Sorusu ?:**

Araştırmada diyabeti olan kadınların birinci ve ikinci sistolik kan basıncı ölçümleri arasında istatistiksel olarak anlamlı bir fark var mıdır?

Ana veri seti üstünde istenilen duruma göre **Select Cases** yaptıktan sonra oluşan yeni veri setinde;

**Analyze 🡪NonParametric Test 🡪Legacy Dialogs 🡪2 Related Sample**

Karşılaştırılması yapılacak olan değişken değerleri **Variable 1** ve **Variable 2** içine tanımlıyoruz.

**6.Chi – Square Test**

Bu test, iki nitel değişken arasında ( kadın-erkek ve sigara içiyor-sigara içmiyor) istatistiksel bir ilişki bulunup bulunmadığını incelemek için kullanılır. Bu testte kategorik veriler (nominal-ordinal) arasındaki ilişkiye bakılır.

**Örnek Hipotez:**

. Grip aşısı olan ve olmayan kişilerin grip olma ve olmama durumunun karşılaştırılması.

. Yaş gruplarına göre ( 30 yaş ve altı, 30-50 yaş, 50 yaş üstü ) depresyon sıklığı ( depresyon var-yok) farklı mıdır ?

**Pearson Ki – Kare :** Gözlerdeki gözlem sayısının 25’in üstünde olduğu zaman

**Yates Düzeltmeli Ki – Kare (Continuity Correction) :**Herhangi bir gözdeki gözlem sayısının 25’in altında olması durumunda kullanılır.(4 gözlü ki-kare için)

**Fisher Kesin Ki – Kare :** Herhangi bir gözdeki beklenen frekans değeri 5’in altında ise kullanılır.(4 gözlü ki-kare için)

**Beklenen frekans değeri** = Alt kolon toplamı x Satır toplamı / Örneklem sayısı

**Test Gereklilikleri :**

**.** Karşılaştırılan değişkenler kategorik olmalıdır.

**.** Karşılaştırılan gruplar bağımsız olmalıdır.

**.** Tablolarda herhangi bir hücrenin beklenen sıklığı 1’in altında olmaması veya beklenen sıklıkların %20’sinden fazlasının 5’in altında olmaması gerekir.

**Bağımlı Gruplarda Ki – Kare Testi ( McNemar Test )**

Bu test kategorik bir değişken yönünden aynı bireyler ya da deneklerden farklı zamanda yada durumlarda elde edilen iki gözlemin farkını analiz etmek için kullanılır. Yalnızca 4 gözlü düzenlerde uygulanabilir.

**Örnek :** Danışmanlık alan 100 diyabet hastasının hap yerine insülin kullanmayı tercih etme durumlarının karşılaştırılması ( danışmanlık öncesi insülin kullanan-hap kullanan ve danışmanlık sonrası insülin kullanan-hap kullanan)

**Araştırma Sorusu ? :**

1.Cinsiyetler arasında diyabet sıklığı açısından fark var mıdır ? ( 4 gözlü ki-kare )

2.Araştıramaya katılanların açlık kan şekerlerini 100 ve altı = 1 , 100-200 arası = 2 ve 200 üstü =3 olarak gruplayın. Daha sonra bu gruplar arasında diyabet sıklığı açısından fark olup olmadığına bakınız. ( Çok gözlü ki-kare )

**Analyze 🡪 Descriptive Statistic 🡪Crosstabs**

**Bağımsız deeğişken** – Row

**Bağımlı Değişken** – Columns

Daha sonra **Statistic** içinden **Chi-Square** işaretliyoruz ve **Cell** içinden de **expected** ile **percentages** kısmından da yüzdeliğini istediğimiz yeri seçiyoruz.

Çıkan **Chi-Square** testi sonucundaki tabloda bakmamız gereken kısım tablonun alt kısmıdır.

Hücrelerin %20’den fazlasında 5’ten küçük frekans olup olmadığı incelenecektir.

(0 cells,(0.0 %) have expected count less than 5)

**Eğer bu değer %20 den küçük ise ; Pearson Chi-Square** değeri okunabilir. Ancak bu değer %20’yi geçmiş ise bu durumda cümlenin devamı (the minimum expected count ….) okunmalıdır ve ;

**Eğer bu değer < 5 ise** 🡪Fisher’s Exact Test

**5 ≤ değer < 25 ise 🡪** Continuity Correction

**Değer ≥ 25 ise 🡪** Pearson Chi-Square

**Varyans Analizi**

Bu test ikiden fazla grubun karşılaştırılmasında kullanılır. Örneğin ; öğretmen, mühendis ve hemşirelerin aylık gelirinin karşılaştırılması. Biz iki grubu sayısal ölçüm açısından karşılaştırmak için t testlerini kullanıyorduk. Örneğin ; öğretmen ile mühendisin aylık geliri karşılaştırması.

F değerinin artması gruplar arasında farkın arttığını göstermektedir. F değeri 1’e eşik veya küçükse “gruplar arasında fark yoktur.” sonucuna ulaşılmaktadır.

Bağımlı ve bağımsız değişkenlerin sayısına göre varyans analizinin türü değişmektedir. Bağımlı ve bağımsız değişken bir tane ise; **Tek Yönlü Anova ( One-Way ) ----- Nonparametriği olan test Kruskal-Wallis**

Bağımlı değişken bir, bağımsız değişken iki tane ise; **İki Yönlü Anova** ( meslek ve çalışma alanına (kamu,özel,serbest) göre aylık gelir miktarının karşılaştırılması)

**One – Way Anova Test ( Tek Yönlü Varyans Analizi )**

Bu test ile ölçümle belirlenen bir değişken yönünden ( kilo, hemoglabin düzeyi vb ) ikiden çok bağımsız grup arasındaki fark ve bu farkın nedeni analiz edilir.

**Örnek ;**

A,B, ve C mamaları ile belirlenen çocuklar 1 yıl boyunca takip ediliyor. Yılın sonunda kilo alımları arasında fark olup olmadığına bakılmak isteniyor.

Bağımsız değişken 🡪Mama

Bağımlı değişken 🡪 Kilo Alımı

**H0 Hipotez :** Ortalama kilo alımları bütün gruplar için aynıdır.

**Test Gereklilikleri :**

**.** İkiden fazla bağımsız grup

**.** Bağımlı değişkenin ölçümle elde edilmiş olması

**.** Normal Dağılım

**Varyans Homojenliği**

Eğer varyanslar homejen ise Test of Homogenity of Variance testinin sonucu p > 0.05 olmalıdır. Homojenlik testi sonucuna göre de gruplar arasındaki farkın hangi gruptan kaynaklandığını bulmak için Post Hoc testler uygulanır.

**Araştırma Sorusu ? :**

1.Kilo grupları arasında ( 0-50 kg=1, 50-70 kg=2, 70-90 kg=3, 90 kg üstü=4 ) total kolesterol düzeyi açısından fark var mıdır ?

Bağımsız değişken 🡪kilo grupları

Bağımlı değişken 🡪 total kolesterol düzeyi

**Analyze 🡪 Compare Means 🡪 One-Way Anova**

**Factor** kısmına bağımsız değişken , **Dependent List** kısmına ise bağımlı değişkenleri koyuyoruz.

2. Kilo grupları arasında ( 0-50 kg=1, 50-70 kg=2, 70-90 kg=3, 90 kg üstü=4 ) sistolik kan basıncı açısından fark var mıdır ?

Bağımsız değişken 🡪kilo grupları

Bağımlı değişken 🡪 sistolik kan basıncı

Varyansların Homojenliği Test Sonucu -- **Analyze 🡪 Compare Means 🡪 One-Way Anova 🡪 Options 🡪 Homogenity of Variance**

Sig. değeri p > 0.05 den büyük ise grupların varyanslarının homejen dağıldığı varsayılmaktadır.

Farkın hangi gruptan kaynakladığını bulmak için ise ; Post Hoc testi seçilir.

**Analyze 🡪 Compare Means 🡪 One-Way Anova 🡪 Post Hoc**

Eşit ise ( p > 0.05 ) ; Tukey, Bonferroni

Eşit değil ise ( p < 0.05 ) ; Tamhane T2

Çıkan tablodaki **yıldız (\*)** ile işaretli olan yerlerde hangi gruplar arsında fark olduğu gösterilmektedir.( p < 0.05 )

**Kruskal – Wallis Testi**

Bu test tek yönlü varyans analizi olan One-Way Anova testinin nonparemetrik eşdeğeridir. Üç ya da daha fazla bağımsız grup verisi parametrik test varsayımlarını karşılamıyorsa bu test kullanılır. Bu test ile ortalamalar değil, ortancalar karşılaştırılır.

Eğer test sonucu p < 0.05 çıkarsa gruplar arasında fark olduğu söylenir ve farkın hangi gruptan kaynaklandığını bulabilmek için gruplar ikişerli olarak Manny-Whitney U testi ile karşılaştırılır.

**Araştırma Sorusu ? :**

Kilo grupları arasında ( 0-50 kg=1, 50-70 kg=2, 70-90 kg=3, 90 kg üstü=4 ) sistolik kan basıncı düzeyi açısından istatistiksel olarak anlamlı bir fark var mıdır ?

Bağımsız değişken 🡪 kilo grupları

Bağımlı değişken 🡪 sistolik kan basıncı düzeyi

**Anlayze 🡪 NonParametric Test 🡪 Legacy Dialogs 🡪 K Independent Samples**

**Test Variable** yazan kısıma Bağımlı değişken, **Grouping** kısmınaiseBağımsız değişken

**Benferroni Düzeltmesi**

Gruplar arasındaki farkın hangi grup ya da gruplardan kaynaklandığını saptamak için ikili grup karşılaştırmalarında p değeri için artık sınır değeri 0.05 değildir. **Yeni değer** 🡪 0.05 / Kaç farklı ikili grup karşılaştırması

**Bağımsız Gruplarda İki Yönlü Anova Testi**

Bu test, bir bağımlı değişken üzerinde ( aylık miktarı vb.) birden fazla bağımsız değişkenin ( bulunduğu bölge (a, b, c), alanı (otomotiv, ilaç, tarım vb.), deneyim yılı ( 10 yıldan az , 10-20 yıl, 20 yıldan fazla ) tek başına ve etkileşimli etkilerini değerlendirmek için kullanılır.

**Test Gereklilikleri :**

**.** Grupların birbirinden bağımsız olması

**.** Normal Dağılım

**.** Varyansların homojen olması

**Araştırma Sorusu ? :**

Cinsiyet ve diyabet varlığının sistolik kan basıncı düzeyi üzerinde bir etkisi var mıdır ?

**H0 Hipotez :** Cinsiyet ve diyabet varlığının sistolik kan basıncı düzeyi üzerinde bir etkisi yoktur.

**Analyze 🡪 General Linear Model 🡪 Univarriate**

**Fixed Group** 🡪 Bağımsız değişken ( kategorik )

**Dependent Value** 🡪 Bağımlı değişken ( ölçümle elde edilmiş )

Options 🡪 Descriptive ve Homogenity

**Korelasyon Analizi**

Bu test, değişkenler arasında ilişki olup olmadığını ve ilişki varsa bu ilişkinin yönünü ve gücünü analiz eder. Çok geniş uygulama alanı vardır. Eğer değişkenlerin ikisinin birlikte hareket ettiği yani birbiri ile ilişkisi olduğu varsayılıyorsa korelasyon analizi yapmak gerekir. Korelasyon neden-sonuç ilişkisini işaret etmez.

**Örnek Hipotez :**

**1.** Fiyat indirimin, satış artışı ile ilişkisi olduğu varsayılmaktadır.

**2.** Çalışma saatlerinin artmasıyla çalışanların verimliliğinin azaldığı varsayılmaktadır.

**3.** Hastanın kilosu ile sistolik kan basıncı arasında bir ilişki olduğu varsayılmaktadır.

**4.** Matematik notlarının artmasıyla fizik notlarının da arttığı varsayılmaktadır.

1. ve 2. 🡪 İki Yönlü

3. ve 4. 🡪 Tek Yönlü

İlişki olduğunu söylersek 🡪 İki Yönlü

İlişkini yönünü söylüyorsak 🡪 Tek Yönlü ( artıyor ve azalıyor gibi )

**Korelasyon Katsayısı ( r )**

**.** -1 ile +1 arasında değer alır.

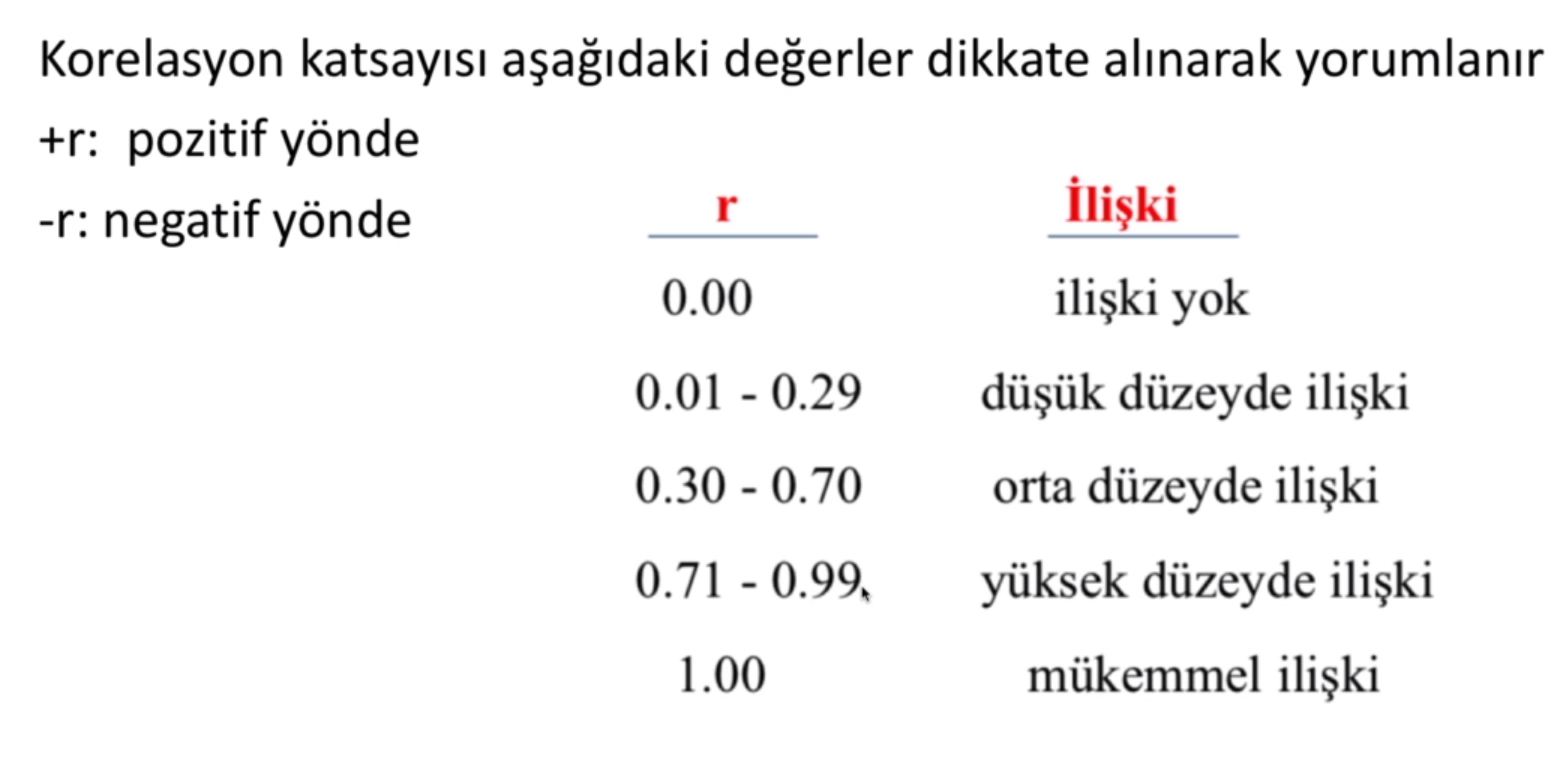
**.** +1 olması iki değişken arasında pozitif yönde tam bir doğrusal ilişki olduğunu göstermektedir. ( x değişkeni artarken, y değişkeni de artar.)

**.** -1 olması iki değişken arasında negatif yönde tam bir doğrusal ilişki olduğunu göstermektedir. ( x değişkeni artarken, y değişkeni azalır.)

**.** 0 olması demek doğrusal bir ilişkinin olmadığını göstermektedir. ( x ile y değişkeni arasında herhangi bir ilişki yoktur.)

**.** 0’dan +1 ya da -1 ile yaklaştıkça ilişkinin gücü artar.

**.** +1 ya da -1’den 0’a yaklaştıkça ilişkinin gücü azalır.



**Korelasyon Çeşitleri**

**Pearson Korelasyonu ( Parametrik )**

Parametrik bir testtir. Eğer normal dağılım yoksa nonparametiriği olan Spearman’s kullanılır.

**Test Gereklilikleri:**

**.** Normal Dağılım

**.** Bağımlı ve bağımsız değişken ölçümle elde edilmiş **sürekli veri** ( boy uzunluğu, kilo vb.)

**.** Örneklem sayısı 30’dan fazla olmalıdır.

**Spearman’s Korelasyonu ( Nonparametrik )**

**Test Gereklilikleri :**

**.** Normal olmayan dağılım

**.** Örneklem sayısı 30’dan az

**.** Sürekli veriler önce sıralanır daha sonra bu sıralanan veriler arasında korelasyona bakılır. ( **sıralı veri** )

Eğer örneklem sayısı çok küçükse Spearman’s yerine nonparametrik bir test olan Kendall’s Tau-b kullanılır. Tüm korelasyon çeşitlerinin sonuçları aynı şekilde yorumlanır.

**Serpme Grafikleri**

Korelasyon katsayısını hesaplamadan önce verinin dağılımı grafik olarak incelenebilir. Grafikte bağımlı değişken “Y”, bağımsı değişken “X” ekseni üzerinde işaretlenir.



**Pearson Korelasyonu**

**Araştırma Sorusu ? :**

.Erkeklerde başlangıçtaki tümör boyutu ile başlangıçtaki ağrı düzeyi arasında bir ilişki var mıdır ?

İlk olarak veri setinin normallik testine bakıyoruz.( **Analyze 🡪Descriptive Statistic 🡪 Explore** dedikten sonra değişkenlerimizi **Dependent List** içine aktarıyoruz**. Plots** içinden Histogram ve Normality Test kısmını seçiyoruz.)

Sonrasında istenilen kısım olan erkekleri verisetinde ayırıyoruz.( **Data 🡪Select Cases 🡪 If conditon 🡪 Cinsiyet=1** )

Daha sonra **Graph 🡪 Legacy Dialogs 🡪 Scatter / Dot 🡪 Simple Scatter** diyerek verinin grafiğini göürüyoruz. ( Bağımlı değişken Y ekseni, bağımsız değişken X ekseni )

**Analyze 🡪 Correlate 🡪 Bivariate** dedikten sonra istenilen değişkenleri **Variable** kısmına atıyoruz. Sonra alt kısımdan **Pearson** ve iki yönlü ilişki olduğunda dolayı **Two-Tailed** seçerek test işlemini gerçekleştiriyoruz.

**Spearman Korelasyonu**

**Araştırma Sorusu ? :**

Öğrencilerin fizik, kimya ve matematik notları arasında istatistiksel olarak anlamlı bir ilişki var mıdır ?

**Graphs 🡪 Legacy 🡪 Scatter 🡪Matrix**

**Analyze 🡪 Correlate 🡪 Bivariate 🡪 Spearman**

**Kendall’s Tau-b Korelasyonu**

**Araştırma Sorusu ? :**

Erkeklerde ağrı palyasyonu olarak sinir bloğu tercih edilmiş olan hastaların tedavi sonrası ölçülen tümör boyutu ve ağrı düzeyi arasında ilişki var mıdır ?

Soruda belirtilen kısımlar olan erkek ve ağrı palyasyonu sinir bloğu olması dolayısıyla ilk olarak veriseti içinde bu kısımların ayrıştırmasını Data 🡪 Select Cases kısmından yaparak yeni bir veriseti oluşturarak korelasyon testimizi yapıyoruz.

**Analyze 🡪 Correlate 🡪 Bivariate 🡪 Kendall’s**

**Kısmi ( Partial ) Korelasyon**

İki değişken arasında ilişki hesaplanırken ( bu iki değişken üzerinde etkisi olduğu düşünülen ) üçüncü bir değişkenin etkisinin sabit tutularak kontrol edildiği durumda hesaplanır.

**Araştırma Sorusu ? :**

Yaş değişkeninin hastaların başlangıçtaki tümör boyutu ve başlangıçtaki ağrı düzeyi üzerinde etkisi olduğu varsayılmaktadır. Yaş değişkenini sabit tutarak tutarak hastaların başlangıçtaki tümör boyutu ve başlangıçtaki ağrı düzeyi arasındaki ilişkiyi inceleyiniz.

**Analyze 🡪 Correlate 🡪 Partial**

**Controlling for :** Kontrol edilmek yani sabit tutulması istenilen değişken

**Variables :** İncelenmesi istenilen diğer değişkenler