

Klinik araştırmalarda veri analizi nasıl yapılır?

Dr. Oktay ÖZDEMİR¹, Dr. Mutlu HAYRAN²

Hacettepe Üniversitesi Tıp Fakültesi İç Hastalıkları¹ Anabilim Dalı Hematoloji Ünitesi ve Onkoloji Enstitüsü Kanser Epidemiyolojisi² Bilim Dalı, Ankara

ARAŞTIRMACILARIN çok sık yaptığı hatalardan biri araştırmanın veri toplama aşaması biter bitmez, hemen "akla en yakın gelen" istatistik analiz yöntemiyle "meşhur" p değerlerinin hesaplanmasıdır. Oysa veri analizi, yalnızca istatistik analiz değildir. Veri analizi denince şu aşamaları düşünmek gerekir: (1) Verilerin istatistik analize hazırlanması, (2) Verilerin gözden geçirilmesi, (3) Uygun istatistik analiz yöntemlerinin seçimi, (4) İstatistik analiz, (5) Analiz sonuçlarının yorumu ve (6) İstatistik sonuçların sunumu. İstatistik analiz, bu aşamalardan yalnızca biri ve hızlı ve güçlü bilgisayarlar sayesinde belki de en kolaydır.

Veri analizi sırasında dikkat edilecek noktalar Tablo 1'de gösterilmiştir. Bu yazıda istatistik analize hazırlık ve istatistik analiz tekniklerinin seçiminde dikkat edilecek noktalar özetlenecektir. Verilerin istatistik analize hazırlanması, eksik verilere, aşırı değerlere, küçük kategorilere ne yapılacağına ilişkin kararlar ve istatistik analiz sonuçlarının yorumu ve sunumu ile ilgili ilkeler bu yazının kapsamı dışında tutulmuştur.

VERİLERE KUŞBAKIŞI

İstatistik analize geçmeden önce yapılması gereken ilk iş verilerin tanımlayıcı istatistiklerinin gözden geçirilmesidir.

Tanımlayıcı istatistik

Bir çalışma grubundaki verilerin çeşitli özelliklerinin özetlendiği ortalama, oran, standart sapma vb. rakamlar topluluğuna tanımlayıcı istatistik adı verilir. Tanımlayıcı istatistikler, (1) Merkezi eğilim ölçütleri ve (2) Yayılma ölçütleri olarak iki başlık altında ele alınabilir.

Merkezi eğilim ölçütleri

Değişken değerleri hakkında bilgi veren en önemli ölçütlerin başında merkezi eğilim ölçütleri (*measures of central tendency*) gelir. Ortalama ve ortanca en sık kullanılan merkezi eğilim ölçütleridir.

Değerlerin toplamının denek sayısına bölünmesiyle elde edilen aritmetik ortalama en sık kullanılan merkezi eğilim ölçütüdür. X üzerinde çizgiyle (X) gösterilir. Sayısal değişkenler için merkezi eğilim ölçütü olarak ortalama kullanılır. Örneğin bilirubin düzeyinin merkezi eğilim ölçütü olarak ortalama kullanılabilir. Ortalama, aşırı değerlerden etkilenir. Örneğin ortalama bilirubin değeri hesaplanacak olan 35 denneğin 34'ünün bilirubin değeri 1.0 mg/dL civarında, birinin bilirubin değeri 16.0 mg/dL ise, bu tek denneğin değeri nedeniyle ortalama yükselir. Ordinal değişkenler için ortalama genellikle kullanılamaz. Örneğin 25 hastanın CHILD sınıflamasına göre grubunun merkezi eğilim ölçütü olarak ortalama kullanılamaz.

Ortanca, küçükten büyüğe (ya da büyükten küçüğe) doğru sıralandığında tam ortadaki denneğin değeridir. Denek sayısı çiftse, ortanca, ortada yer alan iki denneğin değerlerinin ortalaması olarak hesaplanır. Ortanca, ordinal veriler için en iyi merkezi dağılım ölçütüdür. Çünkü değerlerin sıralaması dikkate alınarak elde edilir. Sayısal veriler için tercih edilen merkezi eğilim ölçütü ise ortalamadır. Simetrik dağılımlar için ortalama ve ortanca birbirine yakındır. Ortanca aşırı değerlerden etkilenmez. Bu nedenle aşırı uç değerler varsa, sayısal veriler için de ortanca tercih edilmelidir.

Ne zaman hangi merkezi eğilim ölçütünü kullanalım?

Bir değişkenin merkezi eğilimini en iyi temsil

Tablo 1. Veri analizi sırasında dikkat edilmesi gereken noktalar.**1. HAZIRLIK**

Verileri kodlayın ve bilgisayara ortamına aktarın.
Verilerde hata kontrolünü yapın.
Eksik verilere ne yapılacağına karar verin.

2. VERİLERİN GÖZDEN GEÇİRİLMESİ

Tanımlayıcı istatistikleri yapın.
Aşırı değerlere ne yapılacağına karar verin.
Küçük kategorilere ne yapılacağına karar verin.
Yeni kategoriler türetecek misiniz?
Yeni kategorileri kategorik verilerden mi türeteceksiniz?
Yeni kategorileri sayısal verilerden mi türeteceksiniz?

3. İSTATİSTİK ANALİZ VARSAYIMLARININ GÖZDEN GEÇİRİLMESİ VE UYGUN ANALİZ YÖNTEMLERİNİN SEÇİLMESİ

Çalışmanın asıl amacını tekrar hatırlayın.
Çalışmanın başlangıcında kurduğunuz hipotezlerinizi tekrar hatırlayın: Gruplar arasındaki farklarla mı ilgileniyorsunuz, yoksa kestirim mi yapmak istiyorsunuz?
Değişkenlerin ölçüm türlerini belirleyin.
Sayısal verilerin dağılımını inceleyin.
Parametrik test mi, nonparametrik test mi gerektiğine karar verin.
Gerekirse veri dönüşümü yapın.
Doğru analizi uygulayın.
Alternatif istatistik analiz yöntemleri olup olmadığını araştırın.

4. İSTATİSTİK ANALİZ

"Meşhur" p değeri ile yetinmeyin. Güven sınırlarını da hesaplayın.
Çoklu karşılaştırmalardan mümkün olduğunca kaçının. Mutlaka gerekliyse, nasıl yapacağınızı bir bilene danışın.

5. YORUM

$p < 0.05$ ise
Bulgularınız istatistiksel olarak önemli. Peki klinik olarak da önemli mi?
 $p < 0.05$ ise
Bulgularınız istatistiksel olarak önemsiz. Acaba çalışmanızın gücü (power) yeterli mi?

6. İSTATİSTİK VERİLERİN SUNULMASI

Hangi verileri sunmalısınız?
p değerini nasıl belirtmelisiniz? $p > 0.05$ mi, $p = 0.68$ mi?
p değerini mi, güven sınırlarını mı, ikisini de mi belirtmelisiniz?
Sayıları nasıl belirtmelisiniz?
Tablo mu, grafik mi, hiçbirini mi?

eden ölçütü seçerken, gözönüne alınması gereken iki önemli etken vardır. Bunlar değişkenin ölçüm skalası (yani değişkenin ordinal mi, sayısal mı olduğu) ve değişken değerlerinin dağılımıdır. Değişken değerlerinin dağılımı simetrik ya da basık (*skewed*) olabilir. Eğer aşırı düşük ya da yüksek değerlerden biri diğerinden daha fazla görülüyorsa, dağılım simetrik değildir. Simetrik dağılan sayısal veriler için ortalama, ordinal veriler ya da simetrik olmayan sayısal veriler için ortanca seçilmelidir.

Yayılım ölçütleri

Farklı grupların ortalama ve ortanca gibi merkezi

dağılım ölçütleri aynı olduğu halde, gruplar birbirlerinden çok farklı olabilirler. Bu nedenle merkezi eğilim ölçütleri yanında, yayılım ölçütleri (*measures of dispersion*) de çok önemlidir.

Her birinde yedi hasta bulunan iki gruptaki serum bilirubin değerleri birinci grupta 1.5, 1.6, 1.6, 1.6, 1.6, 1.7 ve 1.7 mg/dL, ikinci grupta ise 0.8, 1.0, 1.4, 1.6, 1.8, 2.3 ve 2.4 mg/dL'dir. Her iki grubun da ortalama ve ortanca bilirubin değeri 1.6 mg/dL'dir. Oysa iki gruptaki bilirubin değerlerinin dağılımının hiç de benzer olmadığı görülmektedir.

Değişken değerlerinin yayılımını belirten yararlı

bir ölçüt, en büyük ve en küçük değer arasındaki fark olan değer aralığıdır (*range*). Yukarıdaki örnekte değer aralığı, birinci grup için 0.2, ikinci grup için 1.6'dır. Değer aralığı da, ortalama gibi uç değerlerden çok etkilenir. Üstelik en uçtaki iki değer arasında kalan değerler hakkında bilgi vermez.

Tüm değerlerin yayılımı ile ilgili bilgi veren, sık kullanılan yayılım ölçütleri standart sapma (*standard deviation*) ve varyanstır (*variance*). Tüm değerler eşitse, her ikisi de sifıra eşittir. Değerler arasında farklar arttıkça standart sapma ve varyans büyür. SD, sd, s ya da σ ile gösterilen standart sapma, değişken değerlerinin ortalamasının etrafındaki yayılmasını temsil eden bir yayılım ölçütüdür. Formülü şöyledir:

$$SD = \sqrt{\frac{\sum (X - \bar{X})^2}{n - 1}}$$

$\bar{X}$ = Ortalama,
n = Denek sayısı

Standart sapmanın karesine varyans adı verilir. Standart sapma ve varyans, birçok istatistik analiz yönteminin temelini oluşturur.

Dağılım özelliği ne olursa olsun, değerlerin en az %75'i ortalama ± 2 SD içinde yer alır.

Normal dağılım gösteren değişken değerleri için aşağıdaki kurallar geçerlidir:

1. Değerlerin %67'si ortalama ± 1 SD içindedir.
2. Değerlerin %95'i ortalama ± 2 SD içindedir.
3. Değerlerin %99.7'si ortalama ± 3 SD içindedir.

Birimleri farklı olan değişkenlerin yayılmalarını karşılaştırmak için değişkenlik katsayıları kullanılabilir. Değişkenlik katsayısı (*coefficient of variation*), standart sapmanın ortalamaya oranının yüzde olarak ifadesidir. Kısaca cv ile gösterilir. Bir grup hastanın ALT düzeyi ortalaması 20 IU/L, standart sapması 5 IU/L ise, değişkenlik katsayısı $(5/20) \times 100 = \%25$ 'tir.

Değişkenlik katsayısı, biyolojik bilimlerde, genellikle bir laboratuvar işleminin ölçüm doğruluğunu test etmek için kullanılır.

Serumda albümin düzeyi tayini için geliştirilen yeni bir yöntemle hazırlanmış bir kitin güvenilirliği belirlenmek isteniyor. Aynı kişiden alınan serum örneğinde bu kit kullanılarak 20 kez albümin ölçümü yapılıyor ve 20 ölçüm sonucundaki değerlerin ortalaması 3.5 mg/dL ve SD'si 0.2 mg/dL olarak bulunuyor.

Yukarıdaki verilerle hesaplanan değişkenlik katsayısı %5.7'dir. Genellikle laboratuvar yöntemlerin kabul edilebilir maksimum değişkenlik katsayısı %10'dur. Buna dayanarak bu kalsiyum kiti ile güvenilir ve tekrarlanabilir ölçümler yapılabildiği söylenebilir. Aynı serum örneğinden yapılan ölçümlerin %95'inin $3.5 \pm 0.4 = 3.1$ ila 3.9 arasında olması beklenir. Eğer SD, 0.2 değil de, örneğin 0.35 olsaydı (yani değişkenlik katsayısı %10 olsaydı), ölçümlerin %95'i $3.5 \pm 0.7 = 2.8$ ila 4.2 arasında olabilirdi ki, bu durumda sonuçlar yeterince güvenilir olarak kabul edilemezdi.

Boy, ağırlık vb. standart normların dağılımları için genellikle persentil kullanılır.

Yüz kişinin boy ortalamasının 168 cm olduğunu ve bunların içinde en kısa olanının 152 cm, en uzun olanının 191 cm olduğunu varsayalım. Bu kişileri kıstadan uzuna doğru boy sırasına dizdiğimizde, 95. kişinin boyu, bu grubun boy değerlerinin 95. persentiline (percentile) eşittir. Aynı şekilde 5. kişinin boyu da 5. persentile eşittir. Ellinci persentil tam ortadaki kişinin değeridir, yani ortancaya eşittir.

Persentilin kullanıldığı diğer bir alan ise, laboratuvar değerlerinin alt ve üst normal sınırlarının belirlenmesidir. Genellikle kabul edilen alt normal sınır 2.5 persentil ve üst normal sınır 97.5 persentildir. Değerlerin dağılımı normalse, ortalama -2SD ve ortalama +2SD, sırasıyla 2.5 ve 97.5 persentil değerlerine karşılık gelir.

25. ve 75. persentil değerleri arasındaki farka çeyreklerarası aralık (*interquartile range*) adı verilir. Yani değerlerin ortada yer alan %50'si, çeyrekler arası aralıktır.

Ne zaman hangi yayılım ölçütünü kullanalım?

Merkezi dağılım ölçütlerinin seçiminde olduğu gibi yayılım ölçütlerinin seçiminde de gözönüne alınması gereken bazı noktalar vardır. Aşağıdaki

kurallar genellikle geçerlidir.

1. Merkezi eğilim ölçütü olarak ortalama kullanıldığı zaman, yayılım ölçütü olarak da standart sapma kullanılır.

2. Persentil ve çeyrekler arası aralık iki durumda kullanılır:

(a) Merkezi eğilim ölçütü olarak ortanca kullanılmışsa;

(b) Merkezi eğilim ölçütü olarak ortalama kullanılmış, ama gözlemlerin standart normlara uygunluğu karşılaştırılmak isteniyorsa.

3. Dağılımın şekli ne olursa olsun değerlerin ortada kalan %50'si tanımlanmak isteniyorsa çeyrekler arası aralık kullanılır.

4. Sayısal verilerde aşırı uç değerler belirtilmek isteniyorsa değer aralığı kullanılır.

5. Sayısal verilerde farklı birimlerle ölçülen değerlerin dağılımları karşılaştırılmak isteniyorsa, değişim katsayısı kullanılır.

İSTATİSTİK ANALİZE HAZIRLIK

Çalışmanın amacı nedir?

"Ask a small and modest question, focus on it in laser beam fashion, and then maintain the focus until the beam burns through." Arthur Kornberg, Nobel ödüllü bilimadamı.

Sık yapılan hatalardan biri, çalışmanın amacının, çalışmanın başlangıcında kesin ve net olarak belirlenmemesidir. Örneğin peptik ülser tedavisinde H2 reseptör antagonisti ve antibiyotik kombinasyonunun plaseboyla karşılaştırıldığı bir çalışmanın amacı şunlardan herhangi biri olabilir:

1. Tedavi ile endoskopik iyileşme ortalama kaç günde sağlanmaktadır?

2. Yaş, cinsiyet, ülserin büyüklüğü vb. gibi ilaca yanıt etkileyen etkenler var mıdır?

3. Tedavi ile H. pylori eradikasyonu arasında ilişki var mıdır?

4. Tedavi ile hastaların yaşam niteliklerinde ne yönde değişim olmaktadır?

Bir çalışmada, yalnızca bir tek amaç olması şart değildir. Ancak çalışmanın amaçlarından birinin birincil amaç olarak seçilmesi ve çalışmanın özel-

likle planlama ve analiz aşamalarında buna dikkat edilmesi gereklidir. Uygun izlem ölçütlerinin belirlenmesi, ancak birincil amacın belirlenmesine bağlıdır. Örneğin yukarıdaki dört amaçtan ilkinde endoskopik izlem yeterli olacakken, üçüncüsünde H. pylori için bir tanı yönteminin uygulanması, sonuncusunda ise yaşam niteliği ölçüğü gibi farklı uygulamalar gerekecektir.

Çalışmanın birincil amacının ve buna uygun izlem ölçütlerinin geliştirilmesi, bu yazının asıl konusu olan veri analizi ile yakından ilişkilidir.

Hipotezlerin gözden geçirilmesi: Gruplar arasındaki farklarla mı ilgileniyorsunuz yoksa kestirim mi yapmak istiyorsunuz?

Çalışmanın amacı belirlendikten sonra bu amaca uygun bir hipotez cümlesi kurulmalıdır. Genel olarak ya "gruplar arasında fark bulmak" ya da "bazı verilere dayanarak diğer bazı verileri kestirmek" olarak tanımlanabilecek iki türde hipotez kurulabilir. Yukarıdaki çalışmanın amacı "Yaş, cinsiyet, ülserin büyüklüğü vb. gibi ilaca yanıt etkileyen etkenlerin olup olmadığının araştırılması" ise hipoteziniz şu cümlelerden biriyle belirtilebilir:

1. Gruplar arasında fark:

"Erkekler ve kadınlarda ilaca yanıt oranları farklıdır."

"Büyük ve küçük ülserlerde ilaca yanıt oranları farklıdır."

2. Kestirim:

"Yaş, cinsiyet, ülserin büyüklüğü vb. veriler bilindiğinde hastanın ilaca yanıt verip vermeyeceği kestirilebilir."

Hipotezinizin önceden belirlenmesi, uygun istatistik analiz yönteminin seçilmesi için gereklidir. Yukarıdaki örneklerde hipotez "gruplar arasında fark olduğu" şeklindeyse, oranlar arası farkların test edilmesi gerekirken, "kestirim" yapmak yönünde hipotez kurulmuşsa, regresyon modellerinden birinin seçilmesi uygun olacaktır.

Değişkenlerin ölçüm türlerini belirleyin

İstatistikte değişken (*variable*) terimi, deneklere ait özellikler anlamında kullanılır. Örneğin iki gruptan birine A ilacı, diğerine B ilacının verildiği ve deneklerin ilaca verdikleri yanıtın karşılaştırıl-

dięi bir klinik alıřma düşünelim. Bu alıřmada hastaların yaşı, bir deęiřkendir. Aynı řekilde cinsiyet, boy, aęırlık, kan basıncı, hangi ila grubuna girdikleri, tedavi öncesi kan basıncı vb. hepsi birer deęiřkendir. Pratik anlamda düşünürsek, veri giriř formundaki hastalara ait her bilgi bir deęiřkendir.

Deęiřkenlerin nasıl ölçüldüğünün belirlenmesinden anlařılan "sfigmomanometre, Sahli yöntemi, termometre, otoanalizer,..." gibi cihaz ya da yöntemler ya da "santimetre, yıl, IU/mL" gibi birimler deęildir. İstatistikte ölçüm denildięi zaman anlařılan, deęiřkenin alabileceęi deęerlerle ilgili kısıtlamalardır. Örneęin bir kadının gebelik sayısı 5.5 olamaz, ama yaşı 25.64567 yıl olabilir. Ya da yeni doğan bir bebeęin APGAR skoru 0 ile 10 arasındaki tüm tam sayılar olabilir, ama 8.4 ya da 15 olamaz. Deęiřkenlerin alabileceęi deęerlerin neler olabileceęi, neler olamayacaęı, yani nasıl ölçüldüğünü belirlemek, yapılacak istatistik analizin seimi için ok önemlidir. Ölçüm özelliklerine göre deęiřkenleri başlıa üç gruba ayırabiliriz: nominal, ordinal ve sayısal deęiřkenler.

Nominal ölçüm

Nominal bir deęiřkende ölçüm düzeyleri arasında bir sıralama ya da uzaklık-yakınlık gibi belirli bir mesafe yoktur. Ölçüm düzeyleri (*levels of measurement*), deęiřkenin alabileceęi deęerler anlamındadır. Örneęin kronik aktif hepatitli, sirozlu ve normal saęlıklı deneklerin alındıęı bir alıřmayı ele alalım. Her denek için bir veri giriř formu hazırlanmıřsa, formda "HASTANIN AIT OLDUĞU GRUP" yerine üç farklı deęer girilebilir. O halde bu deęiřkenin 3 ölçüm düzeyi vardır. Normaller "0", kronik aktif hepatit "1" ve siroz "2" olarak kodlanmış olsun. Burada 0, 1 ve 2 aslında rakam olarak bir anlam ifade etmemektedir. Deneklerin ait oldukları grupları "0, 1, 2" ile deęil de "9, 1, 2" ya da "NOR, KAH, SİR" diye de kodlayabiliriz. Nominal deęiřken deęerleri, rakam olarak deęil, ad olarak anlam ifade ederler.

İki düzeyli, yani yalnız iki deęer alabilen nominal deęiřkenlere ikili deęiřkenler (*dichotomous variable*) adı verilir. Örneęin "CİNSİYET", erkek ve kadın olarak iki düzeyi olan bir ikili deęiřkendir. Aynı řekilde "evet, hayır" olarak iki deęer alabilen bir deęiřken de ikili deęiřkendir.

Ordinal ölçüm

Ordinal bir deęiřkende ölçüm düzeyleri arasında

bir sıralama vardır, ama düzeyler arasındaki mesafeler belirli deęildir. Örneęin tümörlü hastaların evresinin girildięi "EVRE" adını verdięimiz bir deęiřken olsun. Bu deęiřkene evre 1'den 4'e karřılık gelmek üzere, 1, 2, 3 ve 4 deęerleri girilebilir. Bu deęerler belirli bir sıra ifade etmektedir. Örneęin "evre 3, evre 2'den daha ileri evredir", "evre 1 en iyi, evre 4 en kötü evredir" vb. Ama deęiřkenin düzeyleri arasındaki mesafeler belirli deęildir. Örneęin matematik işlem yapıldıęında $2-1=4-3=1$ doğrudur, ama "evre 2, evre 1'den ne kadar ileriye, evre 4 de evre 3'den o kadar ileridir" denemez.

Ordinal deęiřken deęerleri yalnızca ">" ve "<" işlemleri için sayı gibi deęerlendirilir; bunlar dışındaki matematik işlemler uygulanamaz.

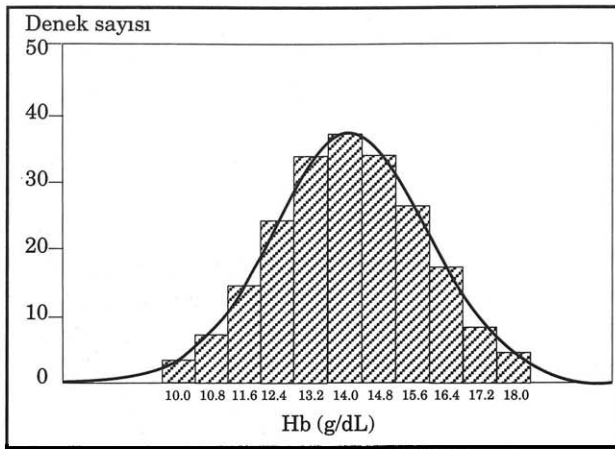
Yukarıdaki örnekte denekleri "1, 2, 3, 4" yerine "1, 4, 5, 9" ya da "8, 6, 4, 2" řeklinde de kodlasak, sıralama deęiřmeyeceęi için farketmez. Ama "1, 0, 2, 3" řeklinde kodlamak hatalı olur, ünkü bu durumda ordinal deęiřkenin ölçüm düzeyleri arasında sıralama olması koşuluna uyulmamıř olur. Ordinal deęiřken deęerleri arasında ıkartma işlemi yapılamayacaęı için, sıralamayı korumak koşuluyla istenilen rakamlarla kodlanabilir.

Nominal ve ordinal ölçümle belirtilen deęiřkenlere kategorik deęiřkenler adı da verilmektedir.

Sayısal ölçüm

Bir deęiřkenin aldıęı deęerler, nominal ya da ordinal deęiřkenlerde olduęu gibi arařtırıcı tarafından belirlenmiř kodlar deęil de gerek rakamlarsa, o deęiřkenin sayısal ölçüm skalasında ölçüldüğü söylenebilir. Sayısal ölçümle belirlenen deęiřkende, deęiřken düzeyleri arasında hem sıralama, hem de belirli bir mesafe vardır. Sayısal deęiřken deęerlerine realsayılara uygulanan her türlü matematik işlem uygulanabilir.

Sayısal deęiřkenler iki řekilde olabilir: kesikli ya da sayımla belirtilen sayısal deęiřken (*interval variable*) ve sürekli ya da ölçümle belirtilen sayısal deęiřken (*continuous variable*). Eęer bir sayısal deęiřken her deęeri alamıyor, ancak belirli deęerleri alabiliyorsa, o deęiřkene "kesikli" deęiřken adı verilir. Sayımla belirtilen deęerleri alabilen deęiřkenler kesiklidir. Her türlü deęeri alabilen deęiřkenler ise, "sürekli" ya da ölçümle belirtilen deęiřkenlerdir. Örneęin, bir ailedeki ocuk sayısı 1, 2, 3, ... gibi yalnızca tam sayı řeklindeki deęerleri alabilir. Bu nedenle "OCUK SAYISI" kesikli



Grafik 1. Normal dağılım eğrisi

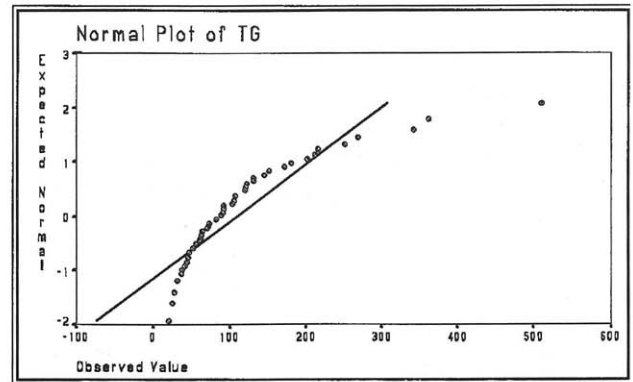
ya da sayımla belirtilen bir değişkendir. Bir insanın yaşı ise genellikle tam sayılarla belirtilse bile, teorik olarak 24.5 yıl, 34.56 gün ondalıklı değerleri alabilir, bu nedenle "YAŞ" sürekli bir değişkendir.

Sayısal değişkenlerin dağılımı

500 kişinin hemoglobinin düzeyleri ölçülüp, düşüktükten yükseğe doğru sıralanıyor. En düşük değerin 9.1 g/dL, en yüksek değerin 18.2 g/dL ve ortalama değerin 14.0 g/dL olduğu saptanıyor.

Bu örnekteki 500 kişinin hemoglobinin (Hb) değerlerinin çoğunun ortalamaya yakın olduğunu, düşük ve yüksek değerlere gidildikçe, sayının azaldığını görürüz. Örneğin Hb'i 13.6-14.4 g/dL arasında olan 36 kişi varken, 12.0-12.8 g/dL arasında olan 26, 10.4-11.2 g/dL arasında olan 7 ve 10.4 g/dL'den düşük olan 3 kişi olduğunu görebiliriz. Yatay ekseninde Hb değeri, dikey ekseninde de denek sayısı olmak koşuluyla çubuk grafik çizilirse (örneğin Hb değeri 13.6-14.4 olan 36 denek olduğunu varsayarsak, yatay eksenindeki 14.0'dan yukarı dikey eksenin 36 birimi uzunluğunda çubuk çizilip, tüm Hb değerleri için bu şekilde grafik tamamlandığında), çubukların tepe noktalarını birleştiren eğri, Hb değerlerinin dağılım eğrisidir (Grafik 1).

Doğadaki ve toplumdaki çok şeyin dağılımı buna benzer. Normal dağılım adı verilen bu dağılım şekli simetriktr. Normal dağılım eğrisi, çan şeklinde olduğu için çan eğrisi olarak da adlandırılmıştır. Normal dağılım simetrik olduğu için normal dağılım gösteren değişkenlerin ortalama ve ortancaları eşittir.

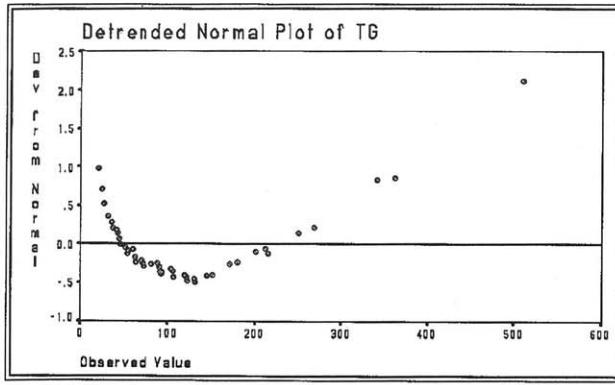


Grafik 2. Bir çalışma grubundaki trigliserid değerlerinin normal olasılık grafiği.

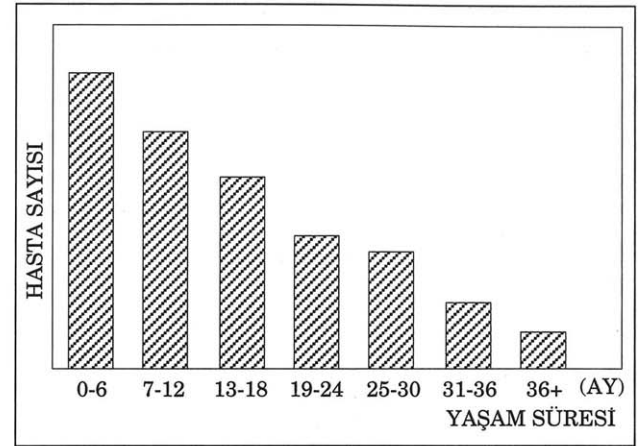
İstatistikte dağılımın normal olup olmadığının belirlenmesi çok önemlidir. Çünkü farklı dağılım gösteren verilere uygulanacak tanımlayıcı ve analitik istatistik yöntemleri de farklıdır. Dağılımın normal olup olmadığı grafik ve istatistik analiz yöntemleri ile araştırılır. Değerlerin histogramının çizilmesi ile dağılımda normalden sapma olup olmadığının kabaca anlaşılması dışında diğer bir grafik yöntem de "normal olasılık grafiği" (*normal probability plot*) adı verilen grafiğin çizilmesidir. Bu grafikte bir eksen gözlenen değerler, diğer eksen ise normal dağılım olduğu varsayımına göre beklenen değerler eksenidir. Veriler normal dağılım gösteriyorsa, noktaların bir doğru üzerinde yer alması ya da etrafında belirli bir pattern göstermeden dağılması gerekir. Grafik 2'de bir çalışmadaki hastaların trigliserid değerlerine ait normal olasılık grafiği örneği görülmektedir. Noktalar bir doğru oluşturmamaktadır ve eğrisel bir pattern göstermektedir. Bu nedenle normal dağılım göstermediği rahatlıkla söylenebilir.

Her değer gözlenen değeri ile normal dağılım olduğu takdirde beklenen değeri arasındaki farkların noktalandığı "detrended" normal olasılık grafiğinde ise beklenen, noktaların dikey eksenindeki "0" dan çizilen yatay çizgi etrafında dağınık olarak yer almalarıdır. Böyle değilse, yine normal dağılım olmadığı varsayılabilir. Grafik 3'de trigliserid değerlerine ait "detrended" normal olasılık örneği görülmektedir. Bu grafikten de dağılımın normal olmadığı söylenebilir.

Grafik yöntemlerden dağılımın normal olup olmadığı hakkında fikir edinilebilir. Ama bu izlenimin istatistik yöntemlerle de test edilmesi gere-



Grafik 3. Bir çalışma grubundaki trigliserid deęerlerinin "detrended" normal olasılık grafięi.



Grafik 4. Kanser hastalarında yaşam süresinin histogramı.

kir. Tek örneklem Kolmogorov-Smirnov testi, Shapiro-Wilks' ve Lilliefors testleri bu amaç için sıklıkla kullanılan testlerdir. Bu testlerde p değeri <0.05 ise, daęılımın normal olmadığı sonucuna varılır. Ancak tamamen normal daęılıma uyan verilerle çalışmak hemen hemen imkansızdır. Özellikle denek sayısı fazlaysa, normal daęılım testleri ile " $p<0.05$ " bulunsu bile "detrended normal plot" gibi grafik deęerlendirmelerde normalden çok ciddi sapmalar yoksa (özellikle belli bir pattern çizerek), daęılımın normal olduğu varsayılmalıdır. Daęılımın normal olup olmadığı kararını verirken grafik yöntemler, istatistik analiz yöntemlerinden daha yararlı olabilir.

Parametrik ya da nonparametrik analiz seçimi

İstatistik analiz tekniklerinin uygulanabilmesi için analize alınacak verilerin bazı koşullara uyması gerekir. Analiz teknikleri uygulanabilme koşullarına göre parametrik ve nonparametrik teknikler olarak ikiye ayrılır. Araştırma düzeni, deneklerin sayısı, verilerin daęılımı, verilerin yayılım ölçütleri (örneğin varyanslar) gibi çeşitli özelliklerle ilgili koşullar sağlanabildiği takdirde uygulanabilecek testler parametrik testlerdir. Parametrik testler için sağlanması gereken koşulların başında verilerin normal daęılıma uyması gelir. Verilerin normal daęılım göstermediği, denek sayısının az olduğu ya da diğer parametrik test uygulayabilme koşullarından herhangi birinin sağlanamadığı durumlarda, parametrik testlerin eşdeęeri ya da alternatifi olan nonparametrik testler uygulanabilir.

Veri dönüşümü

Verilerin daęılımı normal değilse uygulanacak istatistik analiz yöntemi deęişecektir (Parametrik

test yerine nonparametrik test uygulamak gerekecektir). Bu nedenle deęerlerin daęılımının normal olup olmadığı incelenmelidir, çünkü normal daęılım göstermeyen verilere, nonparametrik test yerine parametrik test uygulandığı zaman gerçekte olduğundan daha farklı (genellikle daha küçük) bir p değeri hesaplanır. Bu durumda gerçekte fark olmadığı halde, fark varmış gibi sonuç elde edilebilir.

Verilerin normal daęılmadığı durumlarda genel olarak iki çözüm vardır:

1. Verilere dönüşüm (transformation) uygulayarak, dönüştürülmüş verilerin normal daęılıma uymasının sağlanması ve bu verilere parametrik test uygulanması.
2. Verilerin daęılımı ne olursa olsun uygulanabilen eşdeęer nonparametrik testlerin uygulanması.

Normal daęılımda deęişken değeri ortalamaya yakın olan denek sayısının fazla olduğu ve ortalamadan daha küçük ya da büyük deęerlere doğru gidildikçe, denek sayılarının azaldığını biliyoruz. Bir gruptaki deęerlerin çoğunun ortalamadan küçük olduğu, yalnızca birkaç deęerin ortalamadan büyük olduğu durumda daęılım eğrisi çok farklı olacaktır. Örneğin 100 kanser hastasının tanıdan sonraki yaşam sürelerine bakıldığında, hastaların çoğunun yaşam süresinin kısa olduğu ve süre uzadıkça, yaşayan hasta sayısının azaldığı görülecektir (Grafik 4). Yaşam süresinin histogramı çizildiğinde soldaki (başlangıçta) çubuğun tepe noktasının en üst noktada olduğu ve sağa doğru gittikçe tepe noktalarının aşağıya doğru indiği izlenecektir. Bunun gibi normal dışı daęılım gösteren verilere ilişkin ortalamaları test

Tablo 2. İki değişken arasındaki ilişki ya da bağıntıyı incelemek için uygun analiz teknikleri.

Değişkenlerin ölçüm türleri		Uygun analiz teknikleri
Sayısal-dağılımı normal	Sayısal-dağılımı normal	Pearson basit bağıntı analizi Regresyon analizi
Sayısal-dağılımı normal	Sayısal-dağılımı normal değil ya da ordinal	Spearman nonparametrik bağıntı analizi
Sayısal-dağılımı normal değil	Sayısal-dağılımı normal değil ya da ordinal	Spearman nonparametrik bağıntı analizi
Ordinal	Ordinal	Spearman nonparametrik bağıntı analizi
Nominal ya da dikotom	Ordinal, nominal ya da dikotom	Ki-kare testi

ederken, verileri olduğu gibi kullanmak hatalı olacaktır. Bu durumlarda değerlerin logaritması, karekökü, vb. matematik bir işlem uygulandıktan sonraki değerlerin dağılımı normal dağılıma yaklaşır.

En sık kullanılan dönüşüm yöntemlerinden biri de sıra dönüşümüdür (*rank transformation*). Sıra dönüşümünde değişken değerleri küçükten büyüğe doğru dizildikten sonra, en küçük değere 1, ikinci değere 2, ... gibi sıra numaraları verilir ve testler sırasında verilerin gerçek değeri yerine bu sıra numaraları kullanılır. Özellikle bir taraftan basık dağılım gösteren değişkenler için sıra dönüşümü oldukça uygundur.

HANGİ DURUMLARDA HANGİ ANALİZ?

Uygun analiz yöntemini seçebilmek için yanıtlandırılması gereken bazı önemli sorular vardır. Bu sorular ve yanıtları şöyle özetlenebilir:

1. Analizi yapılacak çalışmada kaç grup denek var?

- Bir grup
- İki grup
- İkiden çok

2. İkiden çok grup varsa grupları belirleyen ya da birbirinden ayıran özelliğin (yani grup değişkeninin) ölçüm skalası nasıl?

- Nominal (Örnek: Antiasit grubu-Proton pompa inhibitörü grubu-Kontrol grubu)
- Ordinal (Örnek: Plasebo grubu-5 mg ilaç grubu- 10 mg ilaç grubu- 20 mg ilaç grubu)

3. Ölçtüğünüz değişkenin (bağımlı değişkenin) ölçüm skalası nasıl?

- Dikotom (Örnek: Tedaviye yanıt var- Tedaviye yanıt yok)
- Nominal (Örnek: Ölüm nedeni hepatik koma-hepatorenal sendrom-infeksiyon-diğer)
- Ordinal (Örnek: Tam iyileşme var- Kısmi iyileşme var- İyileşme yok)
- Sayısal (Örnek: Hb, bilirubin vb.)

Tablo 3. Tek grupta iki ölçüm yapılan durumlarda ölçümler arasındaki farkı ya da uyumu incelemek için uygun analiz teknikleri.

Değişkenlerin ölçüm türleri	Uygun analiz teknikleri	
	Değişkenler arasındaki fark için	Değişkenlerin ölçüm türleri
Sayısal-dağılımı normal	Bağımlı gruplar için (paired) Student t testi	ROC (receiver operator curve) analizi Duyarlılık ve özgüllük analizi
Sayısal-dağılımı normal değil	Wilcoxon testi	ROC (receiver operator curve) analizi Duyarlılık ve özgüllük analizi
Ordinal	Wilcoxon testi	Kappa analizi Duyarlılık ve özgüllük analizi
Dikotom	Mc Nemar testi	Kappa analizi Duyarlılık ve özgüllük analizi

Tablo 4. Tek grupta ikiden çok ölçüm yapılan durumlarda ölçümler arasındaki farkı incelemek için uygun analiz teknikleri.

Değişkenlerin ölçüm türleri	Uygun analiz teknikleri
Sayısal-dağılımı normal Sayısal-dağılımı normal değil ya da ordinal Dikotom	Tekrarlı ölçümler varyans analizi Friedman nonparametrik varyans analizi Cochran Q testi

4. Bağımlı değişkeni kaç kez ölçtünüz?

5. Bağımlı değişken sayıysa, dağılımı normal mi, değil mi?

6. İlgilendiğiniz konu gruplar ya da ölçümler arasındaki fark mı, yoksa birden fazla değişken arasında ilişki olup olmadığı ile mi ilgileniyorsunuz?

Bir grupta bir kez ölçüm yapılmışsa

Bu durumda tanımlayıcı istatistikleri (ortalama, ortanca, oran, standart sapma, IQR gibi) verebilirsiniz. Bunun dışında birden çok değişken arasında ilişki olup olmadığı incelenebilir. Örneğin "Albümin ile bilirubin arasında ilişki var mıdır?" ya da "Cinsiyetle hepatitin itipi arasında ilişki var mıdır" sorularının yanıtlarını arayabilirsiniz (Tablo 2).

Bir grupta iki kez ölçüm yapılmışsa

Bu türdeki çalışmalara önce-sonra (*pre-post*) türünde ya da kendi-kendine kontrollü (*self-control*) çalışmalar adı verilir. Asıl ilgilenilen konu genellikle önceki ve sonraki ölçümler arasında fark olup olmadığıdır. Ayrıca iki farklı tanısai yöntemin performansları arasındaki uyumun test edildiği çalışmalar da bu gruba girebilir. Bu türdeki çalışmalarda ise asıl ilgilenilen konu, iki yöntem arasında uyum olup olmadığıdır. Bu grup çalışmaların analizinde de ölçülen değişkenin ölçüm skalası ve sayısal değişkenlerin dağılımı önemlidir (Tablo 3).

Bir grupta ikiden çok kez ölçüm yapılmışsa

Bu türdeki çalışmalar da önce-sonra çalışmalarıdır. Bir önceki paragraftan farklı olarak iki kez değil, ikiden çok kez ölçüm yapılmıştır. Asıl ilgilenilen konu genellikle önceki ve sonraki ölçümler arasında fark olup olmadığıdır. Tanısai yöntemlerin performansları arasındaki uyumun test edildiği çalışmalarda genellikle aynı çalışmada ikiden çok yöntem değerlendirilmediği için bu grup çalışmalarda asıl ilgilenilen konu genellikle değişik zamanlarda yapılan ölçümler arasında fark olup olmadığıdır. Bu grup çalışmalara uygun analiz teknikleri Tablo 4'te gösterilmiştir.

İki ya da daha çok grupta bir kez ölçüm yapılmışsa

Bu türdeki çalışmalarda asıl ilgilenilen konu gruplar arasında fark olup olmadığıdır. Bu grup çalışmalara uygun analiz teknikleri Tablo 5'te gösterilmiştir.

İki ya da daha çok grupta birden çok kez ölçüm yapılmışsa

Bu türdeki çalışmalarda hem birbirinden bağımsız gruplarda, hem de birden çok kez ölçüm yapılmıştır. Eğer ölçülen değişken dağılımı normal olan bir sayısal değişkense, tekrarlı ölçümler varyans analizi uygulanabilir. Aksi takdirde hem gruplar arasındaki farkı, hem de farklı zamanlardaki ölçümler arasındaki farkı incelemek çok iyi planlanmış çalışmalar dışında tek bir analizle pek

Tablo 5. İki ya da daha çok grupta bir kez ölçüm yapılan durumlarda ölçümler arasındaki farkı incelemek için uygun analiz teknikleri.

Değişkenlerin ölçüm türleri	Uygun analiz teknikleri
Grup Sayısı= 2 Sayısal- dağılımı normal Sayısal-dağılımı normal değil ya da ordinal Nominal ya da dikotom	Bağımsız gruplar için (unpaired) Student t testi Mann-Whitney U testi Ki-kare testi
Grup Sayısı> 2 Sayısal- dağılımı normal Sayısal-dağılımı normal değil ya da ordinal Nominal ya da dikotom	Tek-yönlü varyans analizi Kruskal-Wallis nonparametrik varyans analizi Ki-kare testi

mümkün değildir. Ya zamanlar arasındaki farkı her grup için ayrı ayrı uygun testlerle (Tablo 4) incelemeniz, ya da gruplar arasındaki farkı her zaman kesiti için ayrı ayrı uygun testlerle (Tablo 5) incelemeniz gerekebilir.

SON SÖZ

Anlamli bir p sayısına ulaşmak araştıracının ana hedefi olmamalıdır. Unutmamalıdır ki veriye yeteri kadar işkence edilirse (!) size anlamlı bir p değeri verecektir. Dahası, istatistik analiz, veri

analizinin çok sayıdaki aşamalarından yalnızca biridir. İstatistik analizden önce verilerin istatistik analize hazırlanması ve istatistik analizden sonra sonuçların doğru yorumu ve doğru sunumu da en az analiz kadar önemlidir. İyi planlanmış ve yürütülmüş, verilerin analizi iyi yapılmış bir araştırmanın sunumu uygun şekilde yapılmazsa, araştırma sonuçları hak ettiği ilgiyi görmez ve bir anlamda emekler boşa gitmiş olur. Öte yandan planlama, yürütme ve analiz aşamalarında hatalar olan bir araştırma ne kadar allanıp pullansa da, ancak hak ettiği (!) yeri alabilir.

KAYNAKLAR

1. Fleiss JL. The Design and Analysis of Clinical Experiments. New York: John Wiley Sons Inc., 1986.
2. Friedman LM, Furberg CD, DeMets DL. Fundamentals of Clinical Trials. 2'nci baskı, St. Louis: Mosby Year Book, 1985; s. 241-265.
3. Hayran M, Özdemir O. Bilgisayar, İstatistik ve Tıp. Ankara: Hekimler Yayın Birliği, 1995.
4. Heaney RP, Dougherty CJ. Research for Health Professionals. Iowa: Iowa State University Press, 1988; s. 188-203.
5. Kahn HA, Sempos CT. Statistical Methods in Epidemiology. New York: Oxford University Press, 1989; s. 230-244.
6. Matthews DE, Farewell VT. Using and Understanding Medical Statistics. 2'nci baskı, Basel: Karger AG, 1988; s. 167-181.
7. Spilker B. Guide to Clinical Interpretation of Data. New York: Raven Press, 1986; s. 37-43.