

ÜCD Güncelleme Serileri

Temmuz 2019 • Cilt: 8 • Sayı: 3

Ürologlar için İstatistik ve Araştırma Metodolojisi

Sayı Editörü:

Dr. Fehmi NARTER

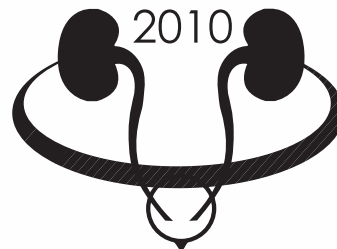
Yazarlar:

Dr. Pınar AY

Dr. Fatih TARHAN

Ertan KOÇ

Dr. Fehmi NARTER



Ürolojik
Cerrahi
Derneği



Editör:

Dr. Serdar TEKGÜL

Editör Yardımcısı ve Genel Koordinatör:

Dr. Rasin ÖZYAVUZ

Sayı Editörü :

Dr. Fehmi NARTER

Yayımlayan:

Ürolojik Cerrahi Derneği

Sorumluluk:

Bilimsel içeriğin sorumluluğu yazarlara aittir.

Bilimsel Danışma Kurulu

*Dr. Abdullah GEDİK
Dr. Ahmet Adil ESEN
Dr. Ahmet ERÖZENCİ
Dr. Ahmet METİN
Dr. Ahmet ŞAHİN
Dr. Ali ERGEN
Dr. Ali GÖKALP
Dr. Ali GÜNEŞ
Dr. Ali TEKİN
Dr. Aydın MUNGAN
Dr. Ayhan KARABULUT
Dr. Bedrettin SEÇKİN
Dr. Cavit CAN
Dr. Ceyhun ÖZYURT
Dr. Cenk Yücel BİLEN
Dr. Cüneyt ÖZKÜRKÇÜGİL
Dr. Çağ ÇAL
Dr. Erim ERDEM
Dr. Feridun ŞENGÖR
Dr. Ferruh ZORLU
Dr. Hakan GEMALMAZ
Dr. Hakan ÖZKARDEŞ
Dr. Haluk ÖZEN
Dr. Hamit ERSOY
Dr. Hayrettin ŞAHİN
Dr. İbrahim CÜREKLİBATUR
Dr. Kaan AYDOS
Dr. Kadir Emre AKKUŞ*

*Dr. Kamil ÇAM
Dr. Levent EMİR
Dr. Levent TÜRKERİ
Dr. M. Bülent ALICI
Dr. M. Zafer SINIK
Dr. Mehmet Bülent ÇETİNEL
Dr. Mesut GÜRDAL
Dr. Nihat SATAR
Dr. Oktay DEMİRKESEN
Dr. Önder KAYIGİL
Dr. Reşit TOKUÇ
Dr. Rüknettin ASLAN
Dr. Şaban SARIKAYA
Dr. Serdar TEKGÜL
Dr. Sinan Sözen
Dr. Sümer BALTACI
Dr. Tahir TURAN
Dr. Tarık ESEN
Dr. Tufan TARCAN
Dr. Turgut ALKIBAY
Dr. Uğur ALTUĞ
Dr. Uğur KUYUMCUOĞLU
Dr. Üstünoğlu KARAOĞLAN
Dr. Zühtü TANSUĞ
Dr. Veli YALÇIN
Dr. Yaşar BEDÜK
Dr. Zafer AYBEK*

İçindekiler

Hangi Epidemiyolojik Araştırma Tipini Seçmeliyim ?	5
Tıbbi İstatistik.....	11
Klinik Çalışmalarda İleri İstatistiğin Yeri	15
Sağkalım Analizleri	19
Sorular	24



Hangi Epidemiyolojik Araştırma Tipini Seçmeliyim?

Dr. Pınar AY

Epidemiyoloji; hastalık veya sağlıkla ilgili durumların toplumda dağılımını ve bu dağılımı etkileyen faktörleri inceleyerek sağlıkla ilgili sorunları kontrol etmeyi hedefleyen bir çalışma alanıdır ⁽¹⁾. Sağlık sorunlarının sıklığı, bu sorunların gelişme riskini artıran ya da azaltan faktörler ve sonuçları epidemiyolojik araştırmalar yoluyla incelenmektedir. Sağlık sorunlarının önlenmesi ve kontrolüne yönelik yapılacak müdahalelerin etkinliği ve istenmeyen etkileri de yine epidemiyolojik araştırmaların konusudur ⁽²⁻⁵⁾.

Epidemiyolojik araştırmalar; deneysel, gözlemsel ve geçerlik/güvenirlik araştırmaları olmak üzere başlıca üç ana türde incelenmektedir (Şekil 1). Bu yazıda; araştırma sorusu ve hipotezine uygun epidemiyolojik araştırma tipinin seçimi ve araştırma tasarımı, klinik örneklerle açıklanmaya çalışılacaktır. Yazıda sadece gözlemsel ve deneysel araştırmalara değinilecek olup, geçerlik ve güvenilirlik çalışmaları çok özel bir alan olduğu için kapsam dışı bırakılmıştır.

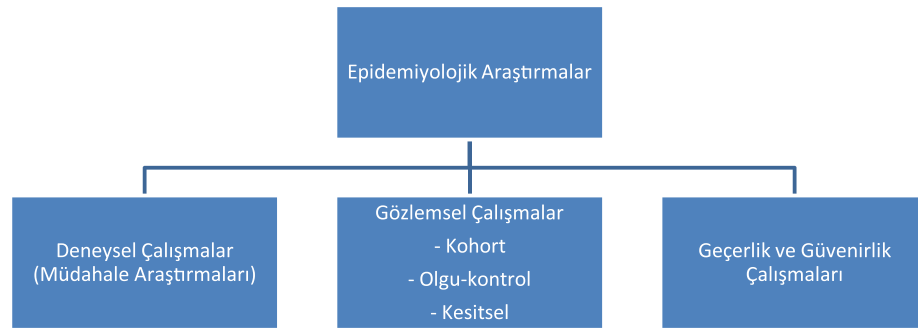
etkinliği incelenen girişim bir ilaç, ameliyat tekniği gibi tedaviye yönelik bir yöntem olabileceği gibi aşı ya da yaşam tarzı değişikliği gibi bir hastalık önleme stratejisi de olabilmektedir. Bu tür araştırmalar, sıklıkla randomizasyon yönteminin kullanımını da içerdikleri için Randomize Kontrollü Deneyler olarak sınıflandırılmaktadırlar.

Randomize Kontrollü Deneylere, homojen bir hasta grubu ile başlanması esastır. Bu amaçla, ilk aşamada çalışmaya alınma ve dışlanma kriterleri kullanılarak, benzer özelliklere sahip bir örneklem grubu oluşturulur. Daha sonraki aşamada ise örneklem randomizasyon yoluyla biri deney ve diğeri de kontrol olmak üzere en az iki gruba ayrılır. Randomizasyon, tamamen rastgelelik seçim esasları uygulanarak deney ve kontrol gruplarının seçimi anlamına gelmektedir. Bu işlem ile hedeflenen, deney ve kontrol gruplarının sonucu etkileyebilecek ve karıştırıcı olarak adlandırılan değişkenler açısından olabildiğince benzer olmasının sağlanmasıdır. Randomizasyon; örnekleme alınan bireylerin yazı tura, zar atılarak veya kura çekerek deney ve kontrol gruplarına atanması yöntemiyle yapılabilir. Ancak bugün için klinik çalışmalarda bu yöntemler yerine daha pratik olması sebebiyle sıklıkla bilgisayar programları kullanılarak randomizasyon işlemi yapılmaktadır. Randomizasyonun başarısı ise grupların çeşitli özellikleri açısından benzer dağılıdığının or-

lerde, sadece rastlantısallığa bağlı olarak karıştırıcı değişkenlerin gruplarda dağılımı farklılık gösterebilir. Örneğin bir grupta daha genç, diğer grupta ise daha yaşlı katılımcılar daha fazla oranda yer alabilirler. Veya bir gruba daha hafif diğer gruba ise daha ağır hastalar atanmış olabilir. Bu nedenle özellikle küçük örneklem grubu ile çalışırken, grupların karıştırıcı değişkenler açısından farklı dağılımını önlemek amacıyla tabakalı randomizasyon yöntemine de başvurulabilir.

Randomizasyonun ardından, deney grubuna etkinliği saptanmak istenen müdahale uygulanırken kontrol grubuna geçeri tedavi verilir. Geçeri tedavinin olmadığı durumlarda plasebo vermek, plasebo da verilemiyorsa hiçbir girişimde bulunmamak uygun olacaktır. Ancak araştırmacı ve katılımcıların hangi grubun deney hangisinin kontrol olduğunu bilmesi, incelenen sonuçların değerlendirilmesinde biasların oluşumuna yol açabilmektedir. Bu nedenle, deneysel araştırmalarda tek ya da çift kör yönteminin uygulanması önerilmektedir. Katılımcı ya da sonucu değerlendiren araştırmacı, etkinliği saptanmaya çalışılan müdahalenin hangi gruba uygulandığını bilmiyorsa buna tek kör adı verilmektedir. Hem katılımcı hem de araştırmacı, hangi grubun deney hangisinin kontrol grubu olduğunu bilmiyorsa, bu işlem çift kör olarak adlandırılır.

Bundan sonraki aşamada hem deney hem de kontrol grubu ileriye dönük olarak planlanan süre boyunca izlenir ve belli aralıklarla sonucun gelişimi açısından değerlendirilir. Çalışmada, deney ve kontrol grubu oluşan çıktılar açısından karşılaştırılarak müdahalenin etkinliği bazı epidemiyolojik ölçütler kullanılarak hesaplanır. Bu ölçütlerden en sık kullanılanları; rölatif etkinlik, mutlak risk farkı ve tedavi için gereken sayıdır. Tablo 1'de sözü edilen ölçütlerin formülleri okuyucuya sunulmuştur. Rölatif etkinlik 1'in üzerinde ise sonucun gelişme oranı deney grubunda, kontrol grubuna orana daha fazladır. Örneğin bir tedavi protokolünün etkinliğinin plasebo ile karşılaştırıldığını düşünelim ve araştırmacının temel çıktısının iyileşme olduğunu varsayalım (Tablo 2).



Şekil 1: Epidemiyolojik Araştırma Türleri

Deneysel Araştırmalar

Müdahale araştırmaları adıyla da anılan deneysel araştırmalar; bir girişimin etkinliği ile istenmeyen etkilerini incelemek amacıyla kullanılmaktadır. Bu tür araştırmalarda;

taya konması yoluyla kanıtlanır. Randomize kontrollü deneylerde genellikle bu durum, bir tablo ile okuyucuya sunulur. Ancak randomizasyonun başarılı olabilmesi için yeterli büyüklükte bir örnekte yapılması gerektiği unutulmamalıdır. Küçük örnek-

Deney grubunda iyileşme %40 iken kontrol grubunda bu oran %20 olarak bulunmuştur. Bu durumda, rölatif etkinlik 2 olarak hesaplanır. Diğer bir deyişle tedavinin uygulanması, uygulanmamasına oranla iyileşme olasılığını iki kat artırmaktadır. Rölatif etkinliğin 1 olması deney ve kontrol grupları arasında sonucun gelişme olasılığı açısından fark olmadığını gösterirken, 1'den düşük rölatif etkinlik söz konusu çıktının deney grubunda kontrol grubuna oranla daha az sıklıkta meydana geldiğini ifade eder. Örneğin incelediğimiz olayın mortalite olduğunu varsayalım. Deney grubunda mortalite gelişme oranı %10,

kontrol grubunda %20 ise rölatif etkinlik 0.5 olarak saptanır. Bu durumda yapılan girişimin, ölüm olasılığını yarıya indirdiği sonucuna varılacaktır. Mutlak risk farkı ise girişim sonucunda oluşması beklenen etki büyüklüğünü ortaya koyar. Mutlak risk farkı %20 ise girişim etkisiyle kontrol grubuna göre %20 daha fazla iyileşme elde edildiği sonucuna varılır. Bu ölçütü tedavi için gereken sayı şeklinde de ifade etmek mümkündür. Tedavi için gereken sayının 5 olması, girişimin uygulandığı her beş kişiden birinde iyileşme olacağı anlamına gelmektedir.

Marschall-Kehrel ve arkadaşları tarafından, nörojenik olmayan aşırı aktif mesane ve üriner inkontinanslı çocuklarda antimuskarinik tedavinin plaseboya karşı etkinliğini ölçmek için deneysel türde bir araştırma yürütülmüştür⁽⁶⁾. Bu hastalığa sahip çocuklarda; propiverine gibi antimuskarinik ajanlar kullanılmakla birlikte, yazarlar bu ilaçların etkinliğine ait kanıtların düşük düzeyde olduğunu ifade etmektedirler. Araştırmacılar, çalışmaya homojen bir hasta grubu ile başlamışlar ve deneye alınma/dışlanma kriterlerini net, açık ve ölçülebilir bir biçimde ortaya koymuşlardır. Yaşı 5-10 yıl arasında değişen, 17-45 kg ağırlığa sahip, miktürasyon sıklığı ≥ 8 /gün ve inkontinans epizodu ≥ 1 /hafta olan kız ve erkek çocuklar çalışmaya dahil edilmiştir. Öte yandan çalışma sırasında idrar yolu infeksiyonunun olması, böbrek, karaciğer ve kardiyovasküler hastalık varlığı da dışlama kriteri olarak belirlenmiştir. Sözü edilen çalışmada katılımcılar, iki gruba randomize edilmişlerdir. Deney grubu antimuskarinik ilaç olan propiverine, kontrol grubu ise plasebo almıştır. Çalışmada randomizasyonun başarılı olduğu; propiverine ve plasebo alan çocukların cinsiyet, yaş, vücut kütle indeksi, günlük idrar yapma sıklığı, idrar miktarı, günlük ve haftalık inkontinans epizod sayısı gibi karıştırıcı olabilecek değişkenler açısından benzer dağılıdığı okuyucuya gösterilmektedir. Çift kör yöntemin kullanıldığı çalışmada, temel çıktı inkontinans epizod sayısı, ikincil çıktı ise ilacın istenmeyen etkileridir. Deney ve kontrol grubu bu çıktılar açısından karşılaştırılmıştır ve ilacın etkinliği gösterilmiştir⁽⁶⁾. Deneysel çalışmalarda, eğer olumsuz bir durum ya da istenmeyen etkiden söz edilecekse rölatif etkinlik yerine rölatif risk kavramını kullanmak daha doğru olacaktır. Propiverine ve plasebonun karşılaştırıldığı çalışmada gastrointestinal istenmeyen etkiler; deney grubunda %9.2, kontrol grubunda ise %3.6'dır. Propiverine kullanımı, plaseboya göre gastrointestinal yan etki olasılığını yaklaşık olarak 2.5 kat artırmaktadır⁽⁶⁾.

Tablo 1: Deneysel çalışmalarda bazı ölçütlerin hesaplanması

	Sonuç		
	Var	Yok	Toplam
Deney grubu	a	b	a + b
Kontrol grubu	c	d	c + d
Toplam	a + c	b + d	a + b + c + d

Rölatif etkinlik

deney grubundaki sonuç gelişim oranı / kontrol grubundaki sonuç gelişim oranı

$$\text{Rölatif etkinlik} = (a/a+b) / (c/c+d)$$

Mutlak risk farkı

deney grubundaki sonuç gelişim oranı - kontrol grubundaki sonuç gelişim oranı

$$\text{Mutlak risk farkı} = (a/a+b) - (c/c+d)$$

Tedavi için gereken sayı

Tedavi için gereken sayı= 1 / Mutlak risk farkı

Tablo 2: Deneysel çalışmalarda bazı ölçütlerin hesaplanması ile ilgili bir örnek

	Sonuç		
	Var	Yok	Toplam
Deney grubu	40	60	100
Kontrol grubu	20	80	100
Toplam	60	140	200

$$\text{Rölatif etkinlik} = (40/100) / (20/100) = 2$$

$$\text{Mutlak risk farkı} = (40/100) - (20/100) = \%20$$

$$\text{Tedavi için gereken sayı} = 1 / 0.20 = 5$$

Gözlemsel Araştırmalar

Gözlemsel çalışmalarda neden ile sonuç arasındaki ilişki, araştırmacı tarafından yapılan bir müdahale olmadan incelenmektedir. Bu çalışmalar; temel olarak kohort, olgu-kontrol ve kesitsel olmak üzere üç ana türde yürütülmektedir (Şekil 1). Çalışma nedenden yola çıkarak sonuca doğru ilerlediğinde kohort, sonuçtan yola çıkarak nedene doğru bir bakış oluşturduğunda ise olgu-kontrol olarak tanımlanmaktadır. Neden ile sonuç ilişkisini aynı anda değerlendiren çalışmalar ise kesitsel olarak adlandırılmaktadır (Şekil 2).

Kohort türü çalışmalara, henüz sonucun gelişmediği bir örneklem ile başlanır. Bu tür çalışmalarda temel amaç, belirli bir izlem dönemi sonunda oluşan yeni olguları belirleyip araştırılan çıktının insidansını ortaya koymaktır. Örneğin mesane kanserinde metastaz insidansının değerlendirildiği bir çalışma, henüz metastaz gelişmemiş homojen özellikteki bir çalışma grubuyla başlatılır. Katılımcılar, ileriye dönük olarak altışar aylık dönemlerde metastaz gelişimi açısından değerlendirilir. İzlem süresinin iki yıl olduğu bu çalışmanın sonunda, iki yıllık metastaz insidansı da elde edilmiştir olacaktır.

Bu tür çalışmaların bir başka amacı da neden sonuç ilişkisini ortaya koymaktır. Örneğin sözü edilen çalışmada metastaz

gelişimini etkileyen faktörler de ortaya konmak istenmiştir.

Böyle bir çalışmada, sigaranın metastaz gelişimini etkileyebilecek bir faktör olup olmadığı değerlendirilmek istenebilir. Bu durumda katılımcılar, sigara içen ve içmeyen diğer bir deyişle maruziyeti olan ve olmayan üzere en az iki gruba ayrılarak iki yıl boyunca izlenir. Çalışmanın sonunda; sigara içen gruptaki metastaz insidansı, sigara içmeyen grup ile karşılaştırılarak rölatif risk elde edilir. Maruziyet ile sonuç arasında bir ilişki yoksa her iki gruptaki insidans benzer olacak ve rölatif risk 1 ya

da 1'e yakın olarak hesaplanacaktır. Rölatif riskin 1'den uzaklaşması risk artışına, 1'in altına düşmesi ise risk azalışına, diğer bir deyişle incelenen değişkenin koruyucu bir faktör olduğuna işaret eder. Mutlak risk farkı ise maruz kalan ve kalmayan gruplardaki insidansların farkı alınarak hesaplanır. Bu ölçüt, sonuç gelişenlerin ne kadarında maruziyetin doğrudan sorumlu olduğunu gösterir. Zarar için gereken sayı ise mutlak risk farkının 1'e bölünmesi ile edilir. Bu ölçüt, maruz kalanların kaç tanesinde sonucun gelişeceğine dair bir bilgi vermektedir (Tablo 3).

Tablo 3: Kohort türü çalışmalarda bazı ölçütlerin hesaplanması

	Sonuç		
	Var	Yok	Toplam
Maruz kalan grup	a	b	a + b
Maruz kalmayan grup	c	d	c + d
Toplam	a + c	b + d	a + b + c + d

Rölatif risk

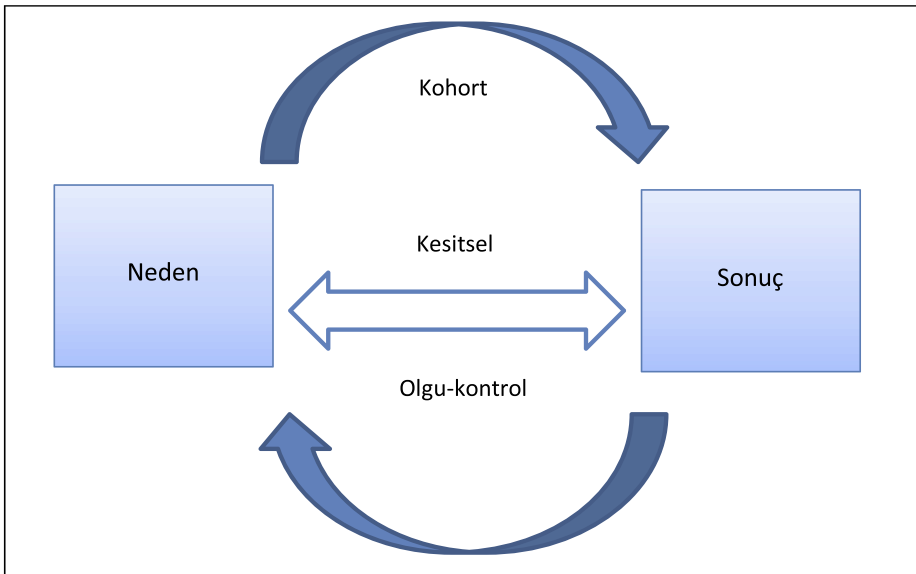
Maruz kalan gruptaki insidans / Maruz kalmayan gruptaki insidans
Rölatif etkinlik = $(a/a+b) / (c/c+d)$

Mutlak risk farkı

Maruz kalan gruptaki insidans - Maruz kalmayan gruptaki insidans
Mutlak risk farkı = $(a/a+b) - (c/c+d)$

Zarar için gereken sayı

Zarar için gereken sayı = $1 / \text{Mutlak risk farkı}$



Şekil 2: Gözlemsel araştırmaların yönü

Yukarıdaki örnekte hareketle sigara içen gruptaki metastaz insidansının %20, sigara içmeyen gruptaki insidansın ise %5 olduğunu varsayalım (Tablo 4). Bu durumda rölatif risk 4 olarak hesaplanacaktır. Diğer bir deyişle sigara içenlerde metastaz riski, sigara içmeyenlere göre iki senenin sonunda dört kat daha fazladır. Aynı çalışmada, mutlak risk farkı ise %15 olarak hesaplanmıştır. Bunun anlamı, metastaz gelişenlerin %15'inde nedenin sigara olduğudur. Bu sonuç metastaz gelişenlerde sigara içme engellenebilseydi %15 daha az metastaz gelişeceğini göstermektedir. Son olarak zarar için gerekli sayı yaklaşık olarak 7 olarak bulunur; mesane kanserine sahip her yedi sigara içen kişiden birinde metastaz gelişecektir (Tablo 4).

Tablo 4: Kohort türü çalışmalarda bazı ölçütlerin hesaplanması ile ilgili bir örnek

	Sonuç		
	Var	Yok	Toplam
Maruz kalan grup	20	80	100
Maruz kalmayan grup	5	95	100
Toplam	25	175	200

$$\begin{aligned} \text{Rölatif risk} &= (20/100) / (5/100) = 4 \\ \text{Mutlak risk farkı} &= (20/100) - (5/100) = \%15 \\ \text{Zarar için gereken sayı} &= 1 / 0.15 = 6.7 \approx 7 \end{aligned}$$

Kohort türü arařtırmalar, prospektif olduđu gibi retrospektif olarak da yürütülebilir. Retrospektif yürütülen kohorta tarihi kohort adı da verilmektedir. Tarihi kohortta tıpkı prospektif kohortta olduđu gibi nedenden sonuca doğru bir gidiş değerlendirilir. Ancak bu sefer, veriler kayıtlarda mevcuttur ve izlenen grupta sonuç hali hazırda gelişmiştir. Arařtırmacı, kayıtlardan yola çıkarak maruz kalan ve kalmayan grupların sonuç gelişimi açısından insidanslarını tespit ederek kıyaslar. Böylece tarihi kohort, prospektife göre zaman ve maliyet açısından çok daha büyük bir avantaj sunar. Yang ve arkadaşları tarafından yapılan bir çalışmada, prostat kanserli hastalardan androjen yoksunluk tedavisi alan ve almayanlar akut üriner retansiyon gelişimi açısından değerlendirilmiştir⁽⁷⁾. Bu amaç doğrultusunda, yeni prostat kanseri tanısı almış olan toplam 24 464 erkek hasta 2000 yılından 2008 yılına dek izlenmiştir. On iki yıllık bir izlemin ardından androjen yoksunluk tedavisi alanlarda, almayanlara göre akut üriner retansiyon gelişim riskinin 1.21 kat fazla olduđu saptanmıştır. Bu çalışmada rölatif riskin, zamanı da hesaba katan bir biçimi olan hazard ratio kullanılmıştır⁽⁷⁾. Kohort çalışmaları, uzun süre izlem gerektirdikleri için oldukça maliyetlidirler. Ayrıca genellikle çalışma örnekleme alınanların tamamı izlem süresi boyunca arařtırmada kalmamaktadırlar. Katılımcıların bir kısmı arařtırmadan ayrılmakta, ayrılanların oranı ise çalışma süresi uzadıkça daha da artmaktadır. Çalışmadan ayrılanlarda gelişen sonuçlar, bilinmediđi ve değerlendirilemediđi için bu durum çalışma

sonuçlarını geçerliliđine de tehlikeye düşürmektedir. İzlem kaybına bađlı yanılıđı olarak tanımlanan bu durum, kohort çalışmalarda en sık karşımıza çıkan bias türlerinden bir tanesidir. Kohort çalışmaları ile ilgili bir başka zorluk nadir ortaya çıkan hastalıklarda kullanılmaması ile ilgilidir. Gelişim sıklığı çok düşük olan hastalıklarda, nedenleri ya da maruziyetleri ortaya çıkarmak için kohort türü çalışmaların yapılması çok geniş bir örneklem gerektirmesi nedeniyle pratik olmamaktadır. Bu tür zorluklardan yola çıkarak kohort yerine olgu-kontrol adı verilen bir başka çalışma tasarımı tercih edilebilir. Olgu-kontrol türü çalışmalar, kohortun aksine sonuçtan yola çıkarak nedene doğru ilerlemektedir (Şekil 2). Yukarıda verilen örnek üzerinden gidecek olursak, mesane kanserli hastalarda sigara kullanımı ile metastaz gelişimi arasındaki ilişkinin arařtırıldıđını varsayalım. Böyle bir durumda hali hazırda metastaz gelişmiş olan katılımcılar olgu, olmayanlar ise kontrol grubunu oluşturacaktır. Arařtırmacılar, her iki grubu da geçmişe yönelik olarak sorgulayarak sigara içme durumlarını ortaya çıkarmaya çalışacaktır. Çalışmada izlem söz konusu olmadığı için kohorta göre daha kolay ve daha az maliyetle yapılabileceđi aşikardır. Öte yandan, geçmişe dayalı olarak maruziyetlerin sorgulanması, bu arařtırma tasarımının en zayıf noktasını oluşturmaktadır. Katılımcıların sadece bir kısmı maruziyetlerini hatırlayıp, bu konuda doğru bilgi verebileceklerdir. Ayrıca olgu ve kontrol grubunda doğru hatırlama oranları birbirinden farklı olacaktır. Me-

tastaz gelişen hastalarda hatırlama daha gerçeđe yakinken, bu oran metastaz gelişmeyenlerde belki daha düşük olacaktır. Bu durumda, sigara kullanımı ile metastaz gelişimi arasında beklenen ilişki, gerçekte olduđundan daha yüksek olarak tespit edilecektir. Hatırlamaya bađlı taraf tutma olarak adlandırılan bu durum, olgu-kontrol arařtırmalarda olduđu gibi tüm geçmişe yönelik veri toplanan tasarımlarda karşımıza çıkmaktadır.

Olgu-kontrol türü çalışmalarda, kontrollerin nereden ve nasıl seçildiđi çok önemlidir. İdeal olan kontrollerin hastaların geldiđi toplumdan seçilmesi, arařtırılan maruziyetlerin toplum düzeyini yansıtmasıdır. Ancak toplum tabanlı seçim, pratik açıdan maliyetli ve zaman alıcıdır. Bu nedenle, arařtırmacılar genellikle kontrolü sađlık kurumlarından arařtırılan hastalığı olmayan kişiler arasından seçmektedirler. Bu seçimde de bazı tasarım hataları gündeme gelebilmektedir. Örneđin sigara kullanımı ile mesane kanseri arasındaki ilişkinin arařtırıldıđı olgu-kontrol türü bir çalışma düşünelim. Bu çalışmada bir ya birkaç sađlık kurumundan ya da kanser kayıt sisteminden olgular, diđer bir deyişle mesane kanserli hastalar seçilebilir. Arařtırmacılar aynı kaynaklardan kontrolleri de seçmek isteyebilirler. Kontrol grubu olarak da akciđer kanserli hastaların seçildiđini varsayalım. Ancak sigara kullanımı ve akciđer kanseri arasında da bir ilişki vardır, dolayısıyla bu şekilde seçilen kontrollerin maruziyet düzeyi toplumu yansıtmamaktadır. Bu durumda, sigara kullanımı ile mesane kanseri arasında gerçekte var olan ilişki, karşılaştırma grubu olarak akciđer kanserli hastalar kullanıldıđı için yokmuş gibi görülebilecektir.

Olgu-kontrol türü çalışmalar, karıştırcılıđın kontrolü açısından bize bazı tasarım avantajları sunmaktadırlar. Eđer cinsiyet, yař gibi karıştırcı faktörlerin bir bölümünü çalışmanın başında kontrol etmek istersek, bu deđişkenleri olgu ve kontrol grupları arasında eşleřtirmek mümkün olabilmektedir. Örneđin erkeklerde mesane kanseri gelişim riski, kadınlara göre daha yüksek görülmüyor olabilir. Arařtırmacılar bu ilişkinin nedensel

olduğu çıkarımını yapmadan önce olası karıştırıcıları kontrol etmek zorundadırlar. Böylesi bir ilişkide sigara kullanımı karıştırıcı olabilir mi? Sigara kullanımı mesane kanseri riskini artırmaktadır, öte yandan erkekler kadınlara oranla daha fazla sıklıkla sigara kullanılmaktadır. Dolayısıyla ilk bakışta cinsiyet ve mesane kanseri arasında varmış gibi gözükken bir ilişki, sigaranın oluşturabileceği karıştırıcı etki nedeniyle ortaya çıkabilir. Bu durumda, olgu ve kontrol grubu karıştırıcı olduğu düşünülen değişkenler açısından eşleştirilebilir. Böylelikle karıştırıcı değişkenin etkisi daha tasarım aşamasında iken kontrol edilmiş olacaktır.

Olgu-kontrol araştırmalarda da tıpkı diğer tasarımlarda olduğu için dört göznlü tablodan hareketle risk hesabı yapılmaktadır. Ancak bu kez kohortta olduğu gibi insidans elde edilemeyeceği için rölatif risk kullanılamaz. Olgu-kontrol araştırmalarda temel risk ölçütü Odds ratio'dur. Odds ratio, bazı Türkçe kaynaklarda tahmini rölatif risk olarak da adlandırılır. Odds ratio, sonucun geliştiği grupta maruz kalanlar ile kalmayanların oranının, sonuç gelişmeyen grupta maruz kalanlar ile kalmayanların oranına bölünmesi ile elde edilir. Yorumu rölatif risk gibidir, Odds ratio 1'den uzaklaştıkça risk artışı ya da azalışı söz konusu olmaktadır (Tablo 5 ve 6).

İYİ'nuna spesifik olarak kullanılan antibiyotikler, infeksiyonun belirteci olarak kullanılmışlardır. Antibiyotik kullanımı 0-1 olanlarda, kullanımın ≥ 10 olanlara kıyasla skuamöz hücreli mesane kanseri riski odds ratio cinsinden 11.4 kat atmış olduğu tespit edilmiştir⁽⁸⁾.

Gözlemsel çalışmaların bir başka türü de kesitsel çalışmalardır. Kesitsel çalışmalarda ana amaç, toplumu temsil eden bir örnekte sağlıkla ilgili bir durumun ya da hastalığın prevalansını saptamaktır. Prevalans; incelenen toplumda var olan tüm olgular anlamına gelir, olguların eski veya yeni olmasının bir önemi yoktur. Bu ölçüt, taramanın yapıldığı toplumda saptanan olguların, tüm risk altındaki popülasyona bölünmesi ile elde edilir. Prevalans, incelenen toplumdaki hastalık yükünü göstermesini açısından da önemli bir ölçüttür.

Kesitsel çalışmaların adından da anlaşılacağı üzere evreni temsil eden bir örnekte yürütülmesi esastır. Buradaki temel amaç çalışmadan elde edilen verilerin, evrene genellenmesidir. Bu nedenle, kesitsel çalışmada örnek tasarımı büyük önem taşır ve olasılıklı örnekleme yöntemlerinin kullanımını gerektirir. Kesitsel çalışmalarda prevalansın yanı sıra maruziyet ile sonuç ilişkisi de incelenmektedir. Ancak kohort ve olgu-kontrol çalışmaların aksine kesitsel çalışmaların yönü yoktur, neden ile sonuç ilişkisine aynı zaman dilimi içinde bakılır (Şekil 2). Bu nedenle kesitsel çalışmalarda, maruziyet ve sonuç arasındaki ilişki net olarak belirlenemez; hangi etkinin maruziyet, hangisinin ise sonuç olduğu tam olarak kestirilemez.

Çin'de yapılan bir çalışmada 5-14 yaş arası çocuklarda, aşırı aktif mesane prevalansı değerlendirilmiştir⁽⁹⁾. Bu amaç doğrultusunda, beş ilde 11 anaokulu ve 25 ilkokul çalışma kapsamına alınmış, öğrenciler bu okullardan rastgele örnekleme yöntemi kullanılarak diğer bir deyişle olasılıklı örnekleme seçilmiştir. Çalışmada aşırı aktif mesane prevalansı %9 olarak saptanmıştır. Araştırmada; obezite, İYİ öyküsü, nokturnal enürezis, kabızlık, fekal inkontinans gibi bazı faktörler ile aşırı aktif mesane arasında

Tablo 5: Olgu-kontrol türü çalışmalarda bazı ölçütlerin hesaplanması

	Sonuç		
	Var	Yok	Toplam
Maruz kalan grup	a	b	a + b
Maruz kalmayan grup	c	d	c + d
Toplam	a + c	b + d	a + b + c + d

$$\text{Odds ratio} = (a/c) / (b/d) = axd / bxc$$

Tablo 6: Olgu-kontrol türü çalışmalarda bazı ölçütlerin hesaplanması ile ilgili bir örnek

	Sonuç		
	Var	Yok	Toplam
Maruz kalan grup	20	10	30
Maruz kalmayan grup	80	90	170
Toplam	100	100	200

$$\begin{aligned} \text{Odds ratio} &= (20/80) / (10/90) \\ &= 20 \times 90 / 10 \times 80 \\ &= 2.25 \end{aligned}$$

Pottegård ve arkadaşları Danimarka'da skuamöz hücreli mesane kanseri gelişiminde idrar yolu infeksiyonların (İYİ) riskini incelemek amacıyla olgu-kontrol türü bir çalışma yürütmüşlerdir⁽⁸⁾. Yazarlar ülkede 2000-2015 yılları arasında histolojik olarak tanı almış skuamöz hücreli mesane kanser olgularını, yaşa ve cinsiyete göre kanser olmayan kontroller ile eşleştirmişlerdir.

Araştırmacılar, görece nadir bir hastalıkla ilgili çalıştıkları için sadece 333 olgu bulmuşlardır. Bu durumda araştırma gücünü düşürmemek amacıyla, kontrol grubu geniş tutulmuş ve her olgu için 50 ayrı kontrol seçilmiştir. Kontroller de toplum maruziyet düzeyini yansıtabilmeleri için hastane tabanlı değil, merkezi nüfus idaresi sistemi kullanılarak seçilmiştir. Çalışmada

da ilişki gösterilmiştir⁽⁹⁾. Ancak çalışmanın zamansal yönü olmadığı için İYİ varlığının maruziyet, aşırı aktif mesanenin ise sonuç olduğu net olarak söylenememektedir. İlişki ters yönde, diğer bir deyişle aşırı aktif mesane varlığı neden, İYİ gelişimi sonuç da olabilir. Dolayısıyla nedensel ilişkilerin aydınlatılabilmesi açısından kesitsel çalışmaların kanıt düzeyi, diğer gözlemsel araştırmalara kıyasla çok daha düşüktür^(5,10).

Kesitsel çalışmalarda risk ölçütü olarak olgu-kontrol araştırmalarda olduğu gibi odds ratio kullanılır, bu ölçütünün hesabı olgu-kontrol türü araştırmalardan farklılık göstermez.

Klinikte hangi epidemiyolojik araştırma türünün seçileceğine; araştırma amacı ya da hipotezi, incelenecek hastalığın sıklığı ile doğal seyri ve eldeki kaynaklar dikkate alınarak karar verilir^(5,10). Eğer eldeki mali kaynaklar, insan gücü ve zaman imkan veriyorsa kanıt düzeyi en yüksek olan araştırma tipini seçmek, çalışmanın gücünü artırması açısından önemlidir.

Kaynaklar

1. Last JM, International Epidemiological Association. A dictionary of epidemiology. New York: Oxford University Press; 2001. 2001
2. Gordis L. Epidemiology. Philadelphia: W.B. Saunders Company, 1996.
3. Tezcan S. Epidemiyoloji: Tıbbi araştırmalarda yöntem bilim. Ankara: Hacettepe Halk Sağlığı Vakfı, 1992.
4. Aksakoğlu G. Sağlıkta araştırma teknikleri ve analiz yöntemleri. İzmir: Dokuz Eylül Üniversitesi Rektörlük Matbaası, 2001.
5. Topuzoğlu, A, Ay P. Kanıt Dayalı Tıp, Klinik Epidemiyolojik Çalışmaların Eleştirel Değerlendirilmesi, Ege Yayınları, İstanbul, 2007.
6. Marschall-Kehrel D, Feustel C, Persson de Geeter C, Stehr M, Radmayr C, Sillén U, Strugala G. Treatment with propiverine in children suffering from nonneurogenic overactive bladder and urinary incontinence: results of a randomized placebo-controlled phase 3 clinical trial. Eur Urol. 2009 Mar;55(3):729-36. doi: 10.1016/j.eururo.2008.04.062.
7. Yang TK, Wu CC, Chang CH, Muo CH, Huang CY, Chung CJ. Subsequent risk of acute urinary retention and androgen deprivation

therapy in patients with prostate cancer: A population-based retrospective cohort study. Medicine (Baltimore). 2020 Feb;99(7):e18842. doi: 10.1097/MD.00000000000018842.

8. Pottegård A, Kristensen KB, Friis S, Hallas J, Jensen JB, Nørgaard M. Urinary tract infections and risk of squamous cell carcinoma bladder cancer: A Danish nationwide case-control study. Int J Cancer. 2020 Apr 1;146(7):1930-1936. doi: 10.1002/ijc.32842.
9. Xing D, Wang YH, Wen YB, Li Q, Feng JJ, Wu JW, Jia ZM, Yang J, Sihoe JD, Song CP, Hu HJ, Franco I, Wen JG. Prevalence and risk factors of overactive bladder in Chinese children: A population-based study. Neurourol Urodyn. 2020 Feb;39(2):688-694. doi: 10.1002/nau.24251.
10. Guyatt G. Evidence based medicine. Chicago, USA, AMA Press, 2002.

Tıbbi İstatistik

Dr. Fatih TARHAN

"Araştırma yapıldıktan sonra biyoistatistikçiye başvurmak; ölüye otopsi yapmasını istemekten başka bir şey değildir. Çünkü bu aşamada biyoistatistikçi sadece, araştırmanın neden öldüğünü söyleyebilir."

R.A. Fisher

TEMEL KAVRAMLAR (1, 2, 3)

Deneklerin sahip olduğu ve farklı değerler alabilen niteliklerine "değişken" adı verilir. Deneğin ait olduğu grup, yaşı, cinsiyeti, boyu, ağırlığı, kan basıncı, serum PSA düzeyi, idrar akım hızı vb. birer değişkendir. "Bağımlı (açıklanan) değişken", sonuç değişkeni olarak da bilinir. Çalışmada incelenen ve başka değişkenler tarafından açıklanan değişkenlerdir, diğer değişkenlerden etkilenir. "Bağımsız (açıklayıcı) değişken" bağımlı değişken üzerinde etkisi araştırılan değişkenlerdir (örn. deneğin grubu). Bağımlı ve bağımsız değişkenler üzerinde etkili olabilecek değişkenler ise kontrol edilecek (etkisi arındırılacak) değişken olarak bilinir. Deneklere ait veriler, niteliksel yani sayılabilir (nominal, ordinal) veya niceliksel yani ölçülebilir (sürekli, kesikli) olabilir. Veriler ölçeklerine göre aralık, nominal, ordinal, oransal olarak da sınıflanabilirler. Nitel verilerde oranlar, nicel verilerde ise ortalamalar analiz edilir. Tıp alanında yapılan çalışmalarda bir gruba ait belirli değişkenlerin değerleri hakkında bilgi, tanımlayıcı istatistik diye adlandırılan ölçütlerle özetlenir. Merkezi eğilim ölçütleri (veri dağılımının nerede toplandığını gösterir: Tepe değeri, Ortanca, Ortalama) ve yaygınlık ölçütleri (veri dağılımının yaygınlığını gösterir: Aralık, Standart sapma, Varyans) tanımlayıcı istatistiğin iki grubudur.

Standart sapma ("Standard deviation"-SD) (sigma- σ): Verilerin ortalamadan ne kadar uzaklıkta olduğunu gösteren bir ölçüdür. Normal dağılımlı bir sürekli değişken verisinde, SD ne kadar küçükse veriler

ortalamaya o kadar yakındır. Aksi durumda ise ortalamadan o kadar uzakta kümelendiği. Çalışma gruplarına ait veriler, sadece ilgili olduğu grubun özelliklerini gösteriyorsa "ortalama \pm standart sapma" tercih edilmelidir.

Standart hata ("Standart error of mean"-SEM): Örneklem, kendi evrenini bir yanlıgı payı ile temsil eder. Bu yanlıgı payına "standart hata" denir. Yani standart hata örnekleme hatasının ölçümüdür. Standart hata, ortalamanın dağılımındaki değişimi gösterir, örneklem sayısının artması ile küçülür. Standart hata arttıkça, örneklemin evreni temsil gücünün azaldığı, küçük olması durumunda örneklemin ana kütleliyi temsil etme yeteneği o kadar yüksek olduğu düşünülür. Gruplar arasında fark araştırıldığında "ortalama \pm standart hata" kullanılması uygundur.

İnsanlar için değişkenlerin çoğunun normal dağılım özelliği gösterdiği varsayılır. Dağılımın nasıl olduğu kullanılacak istatistik yöntemi seçmek için önemlidir. Eğer veriler normal dağılmışsa "ortalama \pm standart sapma" şeklinde, aksi durumda ise merkezi eğilim ölçütleri verilir.

Güven aralığı (confidence interval-CI): Bir parametrenin standart sapma ve ölçme sayısına bağlı gerçek ortalama değerini kapsayan **güven sınırlarının** belirlediği aralıktır. Bir popülasyon parametresinin belirli bir aralık tahminidir. Parametreyi tek bir değerle tahmin etmek yerine, muhtemel bir aralık verilir. Genellikle yüzde olarak ifade edilen belirli bir güven düzeyi ile nitelendirilir. Güven aralığının bitiş noktaları **güven sınırları** olarak adlandırılır. Güven aralığı, başlıca örneklemin büyüklüğü ve homojenliği tarafından etkilenir. Küçük örneklemelerde daha fazla örnekleme hatası olacağından daha geniş güven aralıkları olacaktır. Örneklem boyutu arttıkça, güven aralıkları genişliği azalacaktır.

Güven seviyesi: Gerçek ortalama değer, belirlenen sınırlar içerisinde olma olasılığını belirleyen büyüklüktür. Tıbbi araştırma-

larında çoğunlukla %95 güven seviyesi kullanılmaktadır.

p (probability; olasılık) değeri: Gruplar arasında değişkenler arasında saptanan farkın şans eseri ortaya çıkmış olma olasılığıdır.

Örneklem büyüklüğü: Çalışmaya alınan örneklem kümesinin ait olduğu evreni temsil etmesi gereklidir. Tıbbi araştırmalarda uygun örneklem büyüklüğünün belirlenmesi çalışma maliyeti ve etkin bir çalışma gücünün saptanması bakımından önemlidir. Çalışmanın tasarımı ve kullanılan istatistiksel yöntemlere göre örneklem büyüklüğünün hesaplanması gerekmektedir. Örneklem büyüklüğünü hesaplamak için etki büyüklüğünün, gücün, önemlilik seviyesinin, standart sapmanın ve kullanılacak istatistiksel testin bilinmesine ihtiyaç vardır. Hesaplanan örnek büyüklüğüne olabilecek kayıplar düşünülerek %10-15 oranında eklemeye yapılır.

Testin gücü: Bir testin gerçekte var olan farkı bulabilme yeteneğidir. Güç, örneklem büyüklüğü, etki büyüklüğü ve anlamlılık düzeyi ile ilişkilidir. Sağlık araştırmalarında gücün \geq %80 olması beklenir. Çalışmanın gücünün yeterli olmaması çalışmaların sonuçlarının yanlış yorumlanmasına yol açmaktadır.

Güç analizi: Genellikle çalışma öncesinde yapılır. Güç analizinin çalışma sonrasında geriye dönük yapılması tartışmalıdır. Örneklem büyüklüğü ve güç analizinin doğru olarak saptanması için biyoistatistik uzmanı ile iş birliği gereklidir.

İSTATİSTİKSEL ANALİZ (1, 2)

Bağımlı değişkenin ölçülme şekli ve çalışmanın tasarımı verileri analiz etmek için kullanılan istatistik testine karar vermede yardımcı olur. Nicel değişkenler için verilerin normal dağılıma uygunluğu ve varyansların homojenliği test edilir. Dağılımın normal olup olmadığı hipotez testleri, (Shapiro-Wilks, D'Agostino-Pearson, Kolmogorov-Smirnov testleri

vb.) varyansların homojenliği ise F-Testi (Levene Testi) ile kontrol edilir. Bağımsız iki grupta varyansların birbirine oranı 0,5 ile 2 arasında ise kabaca varyanslar eşittir denebilir. Bazen, çarpık verilerin “normalleştirilmesi” gerekli olabilir. Nicel değişkenlerde, ortalama ve standart sapma ve varyans hesaplanır ve eğer temel varsayımlar sağlanırsa parametrik testler kullanılır. Nitel değişkenlerde ise sıklık ya da her kategorideki denek sayısı gibi tanımlayıcı istatistikler verilir ve non-parametrik testler kullanılır.

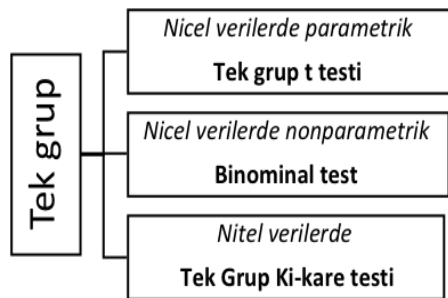
Gruplarda, deneklerin veya onlara ait verilerin sadece bir grupta yer alması, diğer bir grupta yer almaması durumunda **bağımsız veri grubundan** söz edilebilir. Aksi takdirde **bağımlı veri grubu** olur.

TEK DEĞİŞKENLİ ANALİZLER (1, 2)

Parametrik testler: Nicel verilerde, dağılım normal ve varyanslar eşit ise uygulanır. Bağımlı t-testi, Bağımsız t-testi, Varyans analizi, tekrarlayan ANOVA, korelasyon ve lineer regresyon gibi.

Nonparametrik testler: Nitel verilerle dağılımın normal ve varyansların eşit olmadığı ve denek sayısının 30'dan küçük olduğu nicel veriler için uygulanır. Wilcoxon testi, Mann-Whitney U testi, Kruskal-Wallis Varyans Analizi, ki-kare testi, nokta-çift serili korelasyon ve lojistik regresyon gibi.

Bir grupta bir kez ölçüm yapılmışsa tanımlayıcı istatistikler (ortalama, ortanca, oran, standart sapma gibi) verebilir. Tek grup ortalaması t testi, binomial test ve tek grup ki-kare testleri ile değişkenler arasında ilişki olup olmadığı incelenebilir (Şekil 1).

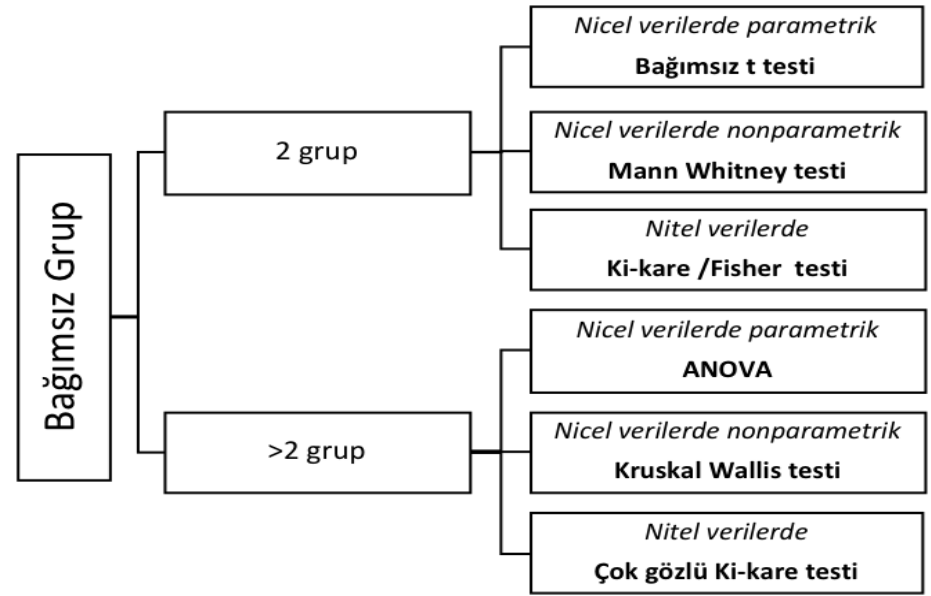


Şekil 1: Bir grupta bir kez ölçüm yapılmışsa kullanılan önemlilik testleri

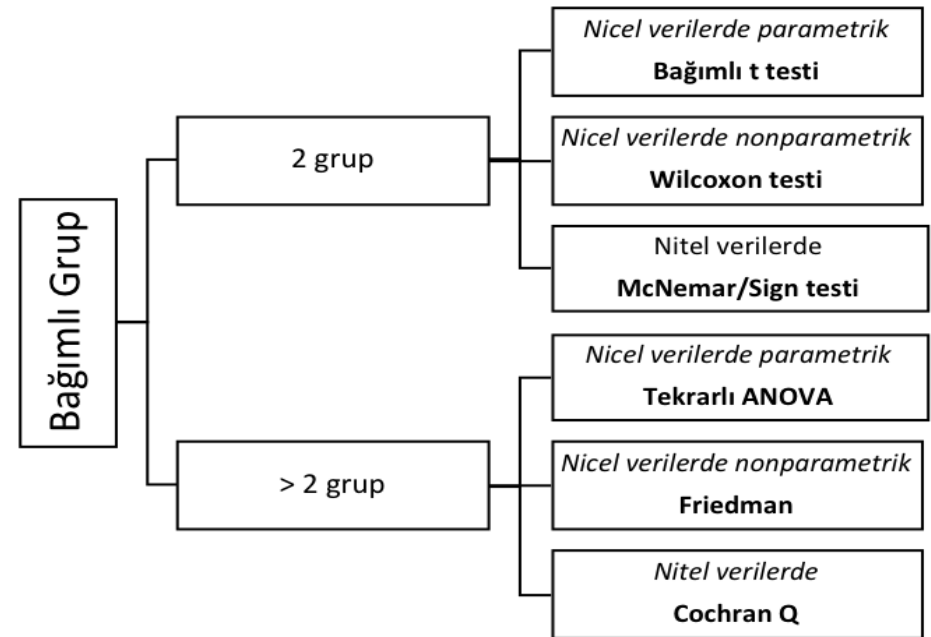
Bağımsız Gruplarda Önemlilik Testleri

Bağımsız gruplarda uygulanabilecek olan testler Şekil 2'de özetlenmiştir. Nicel verilerde temel varsayımlar geçerliyse parametrik eğer geçerli değilse nonparametrik testler uygulanır. Grup sayısı ikiden fazla olduğunda tek değişken (faktör)

analiz edilecekse tek yönlü varyans analizi (ANOVA) kullanılır. ANOVA sonucunda fark bulunursa çoklu karşılaştırma testleri (Tukey HSD, Student-Newman-Keuls, Scheffe, Dunnett, Bonferroni gibi post-hoc testler) ile hangi grupların birbirinden farklı olduğu araştırılır. Eğer iki veya daha çok değişken (faktör) analiz edilecekse iki



Şekil 2: Bağımsız gruplarda kullanılan önemlilik testleri



Şekil 3: Bağımlı gruplarda kullanılan önemlilik testleri

yönlü ANOVA testi yapılır. Pratikte ANOVA testinden kastedilen tek yönlü ANOVA testidir. Eğer dağılım normal değilse, uygun parametrik olmayan test, Kruskal-Wallis tek yönlü varyans analizidir. Kruskal-Wallis testi sonrasında ise ikili karşılaştırmalar Mann-Whitney U testi ile yapılır.

Nitel verilerde ki-kare ve Fischer Testi kullanılır. Ki-kare testinde, frekanslar azaldıkça dağılım bozulur. Gözlerdeki beklenen frekanslar 25'in üzerinde ise Pearson ki-kare testi uygulanır. Aksi durumda Yates düzeltmesi uygulanır. Herhangi bir gözdeki beklenen frekans değeri 5'in altında ise Fisher'in kesin ki-kare testi kullanılır (Şekil 2).

Bağımlı Gruplarda Önemlilik Testleri

Bağımlı değişken nicel veri ise ve temel varsayımlar sağlanıyorsa parametrik testler kullanılır. Aksi durumda parametrik olmayan testler uygulanır. Bağımlı gruplarda uygulanabilecek olan testler Şekil 3'te özetlenmiştir (Şekil 3).

BASİT İLİŞKİ ANALİZİ (1, 2)

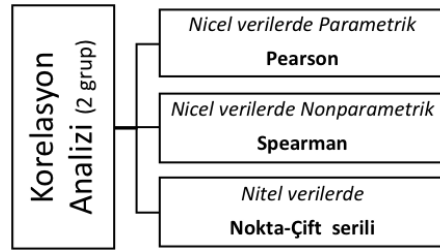
İki ya da daha çok değişken arasında ilişki olup olmadığını, ilişki varsa yönünü ve gücünü korelasyon analizi ile değişkenlerden birisi belirli bir birim değiştiğinde diğerinin nasıl bir değişim gösterdiğini ise regresyon analizi ile inceleyebiliriz.

Korelasyon analizinde, değişkenlerin bağımlı ve bağımsız değişken olarak belirlenmesi hesaplamaların sonucu açısından önemli değilken, regresyon analizinde değişkenlerin hangisinin bağımlı hangisinin bağımsız değişken olduğunu tespit etmek çok önemlidir.

Korelasyon analizi

Korelasyon analizi, bir gruptaki iki değişken arasında ilişkinin belirlenmesi için yapılır. Korelasyon analizi ile neden sonuç ilişkisi gösterilemez, iki değişken arasındaki ilişkinin yönü ve derecesi gösterilir. Değişkenlerin her ikisi de normal dağılım gösteriyorsa yapılan Pearson korelasyon analizinde pozitif tam ilişki: $r=+1$; negatif tam ilişki: $r=-1$ dir. Korelasyonun gücü ise $r=0.0-0.24$ zayıf; $r=0.25-0.49$ orta; $r=0.50-0.74$ güçlü; $r=0.75-1.00$ çok güçlü olarak

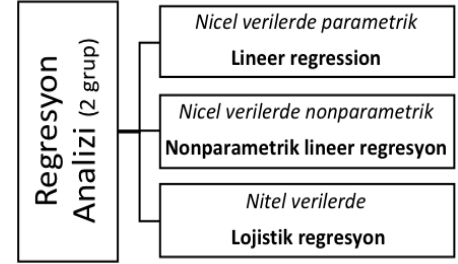
değerlendirilir. Değişkenler arasındaki ilişkinin önemini yorumlamanın bir başka yolu da R^2 (korelasyon katsayısının karesi-tamamlayıcılık katsayısı) hesaplamaktır. Değişkenler, gerekli şartları sağlamıyorsa parametrik olmayan Spearman korelasyon kullanılır (r_s). Nitel verilerde ise nokta-çift serili korelasyon uygulanır. Bu Pearson korelasyonuna benzer ve aynı şekilde yorumlanır (Şekil 4).



Şekil 4: Korelasyon analizi

Regresyon analizi

Regresyon analizi bir değişkenden yararlanılarak bir diğer değişkenin değerlerinin tahmin edilmesi istendiğinde kullanılır. Yani, bağımlı ve bağımsız değişken arasındaki ilişkinin matematiksel bir fonksiyonla da ifade edilmesini sağlar. Nedenellik sorgulanmaz. Basit doğrusal regresyon analizinde bir bağımlı ve bir bağımsız değişken söz konusudur. Değişkenler arasında doğrusal ilişki olabileceği gibi, doğrusal olmayan bir ilişki de olabilir. Kestirim yapabilmemizi sağlayan formüle veya eşitliğe regresyon eşitliği denir. Regresyon analizinde ikinci değişkenin sıfırdan farklı olup olmadığı önemlidir. Eğer sıfıra yakınsa $p>0.005$ dir ve iki değişken arasında bağıntı önemsizdir. Değişkenlerde normal dağılım ve homojenlik yoksa nonparametrik test ile analiz yapılır. Eğer ilişki doğrusal değilse ilişki doğrusallaştırılmalı veya lineer olmayan regresyon analizi yapılmalıdır. Tek değişkenli analizde bağımlı değişken nitel olduğu durumlarda ise lojistik regresyon (binominal) kullanılır (Şekil 5).



Şekil 5: Regresyon analizi

ÇOK DEĞİŞKENLİ ANALİZLER (1, 4)

Çoklu korelasyon ve regresyon analizi: Bağımlı değişkeni birden fazla değişken etkilediğinde çoklu korelasyon ve regresyon analizi yapılmalıdır. Hesaplama ve yorum bakımından tek değişkenli regresyon analizine benzemektedir. Çoklu regresyon analizinde denek sayısının bağımsız değişken sayısının 10 katı kadar olması kabul edilir. Parametrik test için basit analizdeki varsayımlar sağlanmış olmalıdır.

Lojistik regresyon: Çoklu lineer regresyondan farkı bağımlı değişkenin nitel olmasıdır. Bağımsız değişkenler nitel ve/veya nicel olabilir. Normal dağılım şartı gerekli değildir. Bağımlı değişken ikiden çok sayıda kategorili ise multinominal lojistik regresyon yöntemi, bağımlı değişken ikiden çok sayıda kategorili ve ve sıralı olarak belirtiliyorsa ordinal lojistik regresyon kullanılır. Bağımlı değişken kategori sayısı iki olduğunda, bağımsız değişken sayısı tek ise binominal lojistik regresyon, bağımsız değişken iki veya daha fazla sayıda olduğunda ise çok değişkenli lojistik regresyon yöntemi kullanılır. Lojistik regresyon analizi sonucunda elde edilen modelin uygun olup olmadığı "model ki-kare" testi ile her bir bağımsız değişkenin modelde varlığının anlamlı olup olmadığı ise Wald istatistiği ile test edilir. Kural olarak, modeldeki her bağımsız değişken için en az 10 denek olması önerilmektedir.

Çok değişkenli varyans analizi (MANOVA): Bağımlı değişken sayısının birden fazla olduğu bir varyans analizidir. MANOVA'nın ANOVA'dan farkı, iki veya daha fazla bağımlı değişkene sahip olmasıdır. ANOVA gibi, MANOVA'nın hem tek yönlü hem de iki yönlü analizi vardır. Tek yönlü MANOVA,



iki veya daha fazla bağımlı değişkeni tek faktörlü bir değişkenle karşılaştırır. İki yönlü MANOVA ise, iki veya daha fazla bağımlı değişkeni iki veya daha fazla faktör değişkeni ile karşılaştırır. MANOVA'nın parametrik olmayan karşılığı PERMANOVA'dır.

Kovaryans analizi (ANCOVA): İki veya daha çok grupta bir sayısal değişkenin ortalamalarının karşılaştırılması sırasında, bu değişkene etki eden başka bir sayısal değişkenin etkisinin ortadan kaldırılması amaçlanır. ANCOVA'da kullanılan sürekli bağımsız değişken için terim "kovaryat"tır (ortak değişken). Regresyon analizi ile varyans analizinin bir karışımıdır. Parametrik test varsayımları olmalıdır. ANCOVA ile ANOVA arasındaki fark kovaryansdır. ANOVA'nın aksine, ANCOVA bir bağımlı değişkeni hem bir faktör hem de sürekli bir bağımsız değişkenle karşılaştırır.

Çok değişkenli kovaryans analizi (MANCOVA): MANCOVA, ANCOVA'nın çok değişkenli karşılığıdır. ANCOVA'dan farkı kovaryatın çoklu olmasıdır. MANCOVA, bir veya daha fazla ortak değişkenin etkilerini modelden kaldırır. Bu, bağımsız değişkenlerin bağımlı değişkenleriniz üzerindeki gerçek etkisini görmenizi sağlar. Ancak çok daha büyük bir örneklem büyüklüğüne ihtiyaç vardır. MANCOVA'nın da hem tek yönlü hem de iki yönlü analizi vardır.

Çok değişkenli kovaryans analizinde (MANCOVA), tüm varsayımlar MANOVA ile aynıdır. MANOVA ve MANCOVA arasındaki fark ise, "kovaryans" dır. MANOVA sadece faktörleri içerebilirken, bir veya daha fazla kovaryat eklendiğinde analiz MANOVA'dan MANCOVA'ya dönüşür.

Log-lineer analiz: İki den çok sırasız nitel değişken arasındaki ilişkiyi incelemenin yolu değişkenler arasındaki ilişkiyi ikiye bölerek ki-kare analizi ile incelemek değil, log-lineer analizi yapmaktır. Bu analizde değişkenler bağımlı veya bağımsız olarak sınıflandırılmaz. Ki-karede olduğu gibi frekans sayıları üzerinde bazı şartları gerektirir.

Çaprazlık tablosunun her hücresi 5'ten büyük frekansa sahip olmalıdır. Tablodaki her hücrede 1'den küçük hiç değer olmaması ve 5'ten küçük frekansa sahip hücrelerin verinin %20'sini geçmemesi gerekir.

Diskriminant analiz: Temel amacı, grup farklılıklarını açıklamak ve birçok metrik bağımsız değişkene dayalı olarak bir varlığın belli bir sınıfa ya da gruba ait olup olmadığını tahmin etmektir. Hastaların hangi gruba ait olduklarını kestirmek için hangi değişkenlerin önemli kestirici olduğu araştırılır yani birden çok değişkenin değeri bilindiğinde hastaların bir gruba ait olma olasılıkları hesaplanır. Normal dağılım şartı gereklidir.

Faktör analizi: Değişkenlerin hepsi bağımsız değişkendir. Bir grup denekte ölçülen sayısal değişkenler arasındaki ilişkiyi inceleyerek, deneklerdeki farklılıkları en çok belirleyen değişkenlerin ortaya konulmasına yarar. Yani çok sayıda değişkenden deneklerdeki farklılığı en iyi özetleyen değişkenlerin belirlenmesini sağlar. Değişkenler arasındaki ilişkinin derecesi veya birbirine yakınlığı incelenir.

Temel bileşenler analizi: Çok sayıda değişkeni daha az sayıda değişkene indirgeyerek, indirgenen değişkenlerle ileri analiz yapma olanağını sağlayan bir yöntemdir. Özgün verileri bir dizi doğrusal değişken olarak kabul edip her değişkenin temel bileşene katkısını ortaya çıkarmaya çalışır.

Kümeleme analizi: Değişkenlerin hepsi bağımsız değişkendir. Küme analizi ve diskriminant analizine benzer. Deneklerin birbirine yakınlığı incelenir. Hastalar özelliklerine göre gruplara veya kümelere ayrılır. Deneklerin ait olduğu gruplar baştan belirli olmadığı bir diskriminant analizi gibidir. Diskriminant analizinden farklı olarak gruplar önceden tanımlanmaz.

Poisson regresyon analizi: Kesikli sayısal bir bağımlı değişkenle "y", bir veya daha fazla bağımsız değişken arasında ilişki arandığında kullanılacak bir yöntemdir.

Kanonik korelasyon analizi: İki değişken seti arasındaki korelasyonu inceleyen bir yöntemdir. Değişken setinin birisinde bağımsız diğerinde bağımlı değişkenler vardır. Çok sayıda bağımsız değişkene karşı çok sayıda bağımlı değişken incelenebilir. Çoklu regresyon analizinin uzantısıdır.

Çok değişkenli analizlerin doğru uygulanması için ilgili uzmandan destek alınmalıdır.

Kaynaklar

1. Özdemir O. Medikal İstatistik.1.baskı. İstanbul: İstanbul Medikal Yayıncılık; 2006
2. Sümbüloğlu K, Sümbüloğlu V. Biyoistatistik.7.baskı. Ankara: Hatipoğlu Yayınları; 1997.
3. Sümbüloğlu V, Sümbüloğlu K. Klinik ve Saha Araştırmalarında Örneklem Yöntemleri ve Örneklem Büyüklüğü. 1.baskı. Ankara: Hatipoğlu Yayınları; 2005.
4. Sümbüloğlu K, Akdağ B. İleri Biyoistatistiksel Yöntemler. 1.baskı. Ankara: Hatipoğlu Yayınları; 2009.

Klinik Çalışmalarda İleri İstatistiğin Yeri

Ertan KOÇ

Literatürde veya birçok istatistik kitabında istatistik Temel ve İleri olarak ikiye ayrılarak incelenmiştir. Temel İstatistikte hipotez testleri bulunur. İleri İstatistikte ise Regresyon, ROC Eğrisi, Sağkalım analizi gibi testler bulunur.

İleri İstatistik testleri temel İstatistiğin tamamlayıcısıdır. Sağlaması yapılabilir, net yargılarda bulunmamıza fırsat verir. Klinikte kullanılan çoğu tanı algoritmaları, tedavi algoritmaları, kesme (cut off) değerleri İleri İstatistik ile belirlenir.

İleri İstatistiğin avantajlarını vurgulamadan önce temel İstatistiğin eksiklerini vurgulamak gerekir. Klinik çalışmalarda farkın anlamlı olmasını vurgularız. Bunun içinde p değeri kullanılır. P değeri yaptığımız kıyaslamada farkın tesadüf olup olmadığını belirtir. Tıp çalışmalarında genellikle p değerinin 0.05 in altında ($p < 0.05$) olmasını anlamlı sayarız. P değeri çok önemlidir ama aynı zamanda bizi yanıltabilir. P değeri vaka sayısından etkilenir. A hastalığında X semptomunun fazla olduğu ispatlamaya çalışalım. A hastalığı olanlarda X semptomu %20 görülsün, hastalık olmayan normal popülasyonda X semptomu %10 görülsün. Eğer her bir gruba 100 er vaka alırsak $p > 0.05$, her bir gruba 200 er vaka alırsak $p < 0.05$, her gruba 250 şer vaka alırsak $p < 0.01$, her gruba 300 er vaka alırsak $p < 0.001$ olur. Oran hep aynı ama vaka sayısı arttıkça %10'luk fark anlamlı olmaya başlıyor. Temel İstatistik sadece p değerine odaklandığı için yanıltıcı olabilir. İleri İstatistik sadece p değerine odaklanmaz. Bulduğumuz sonucun ayırıcı tanı olup olmayacağını sağlamasını yaparak inceler.

Temel İstatistiğin bir diğer eksiği ise asıl farkı yaratan değişken veya değişkenleri göstermez. Tek bir sonuç vardır, gruplar arasında anlamlı fark vardır ya da yoktur. İleri İstatistik değişkenlerin etki düzeyini gösterir bu sayede değişkenleri birbiri ile de kıyaslar.

Temel İstatistiği tamamlayan ve klinik çalışmalarda en çok ihtiyaç duyulan testler Lojistik Regresyon ve ROC eğrisidir. Lojistik Regresyon ve ROC eğrisi gibi testler için iki uçlu olay (event) kavramına ihtiyacımız var. Böbrek taşı olan ve böbrek taşı olmayan iki grupta olay (event) böbrek taşı olmasıdır. Komplikasyon oluşan ve komplikasyon oluşmayan iki grupta olay (event) komplikasyondur. Tedavi başarısı olan olmayan, nüks olan olmayan, tedavi başarısızlığı olan olmayan, mortalite olan olmayan gibi birçok iki grup üzerinden İleri İstatistik analizi (regresyon-ROC eğrisi) yapabiliriz. A tedavisi ve B tedavisi bu da iki grup olur ama İleri İstatistik için uygun olmaz, çünkü olay (event) kavramı yoktur.

Tablo 1 de temel İstatistik ile yapılan örnek bir kıyaslama mevcut. A hastalığında Marker-A, Marker-B, Marker-C, Marker E değerleri kontrol grubundan anlamlı ($p < 0.05$) olarak daha yüksekti. A hastalığında Marker-D, Marker-F değerleri kontrol grubundan anlamlı ($p < 0.05$) olarak daha düşüktü. Temel İstatistik kıyaslamasında sadece hangisi farklı onu söyleyebiliyoruz. Farkın düzeyini ve asıl odaklanmamız gereken marker ı söyleyemiyoruz.

Tablo-1

	Kontrol Grubu	A Hastalığı	p
	Ort.±ss(Medyan)	Ort.±ss(Medyan)	
Marker A	3.2±1.2(3.1)	6.2±1.8(5.2)	$p < 0.001$
Marker B	6.2±2.1(4.9)	7.9±3.1(6.2)	$p < 0.001$
Marker C	6.6±2.8(5.1)	7.2±2.5(7.2)	$p < 0.001$
Marker D	9.2±3.2(8.5)	3.5±2.5(3.8)	$p < 0.001$
Marker E	5.2±2.1(4.9)	7.2±5.6(6.9)	$p < 0.001$
Marker F	7.4±2.5(7.4)	6.2±2.6(6.1)	$p < 0.001$

Tablo-2

	Tek Değişkenli Model		Çok Değişkenli Model	
	OR (%95 GA)	p	OR (%95 GA)	p
Marker A	1.52 (1.2 - 2.6)	$p < 0.001$	1.64 (1.2 - 2.8)	$p < 0.001$
Marker B	1.82 (1.4 - 3.1)	$p < 0.001$		
Marker C	1.69 (1.2 - 3.2)	$p < 0.001$		
Marker D	0.76 (0.6 - 0.9)	$p < 0.001$		
Marker E	2.74 (2.2 - 4.7)	$p < 0.001$		
Marker F	0.30 (0.2 - 0.5)	$p < 0.001$		

Tablo 2' de lojistik regresyon sonucu mevcut. Tek değişkenli modelde temel İstatistik sonucunun benzerini görüyoruz. Tablo da OR (ODSS Ratio) değeri olasılık oranı demektir. Eğer OR değeri 1 in altındaysa A hastalığında o marker düşüktür, eğer OR değeri 1 in üzerinde ise A hastalığında o marker yüksek demektir. Lojistik regresyonun asıl faydasını çok değişkenli modelde görüyoruz. Çok değişkenli model sadece iki marker kalmış. Bu sonuç Marker B ve Marker D nin A hastalığını öngörmede bağımsız etken olduğunu gösterir. Eğer bir tanı algoritması yapacak isek bütün belirteçleri sorgulamak yerine Marker B ve Marker D yi sorgulamak yeterli olacaktır. Lojistik regresyonun ilk avantajı çok değişkenli modelde odaklanmamız gereken bağımsız etkenleri ortaya çıkarmasıdır.

Lojistik regresyonun diğer bir faydası ise bir model kurmasıdır. Lojistik regresyon kurulan modelin çarpan değerlerini β (Beta) simgesi ile verir. Bu sayede B marker ı ve D marker larına bakarak A hastası olma olasılığı belirlenir. Birçok tanı algoritması bu yöntem ile ortaya çıkar. Önce bağımsız etkenler bulunur, çarpanlar belirlenir, çarpanlar klinikte kullanılabilir şekilde düzeltilir. Lojistik regresyonun en büyük artışı ise kurulan bu modellerin doğruluğunu verir. Yani B ve D Marker ı ile bir model kurulursa A hastalığını doğru tespit oranı, normal grubu doğru tespit oranı, yanlışma oranlarını verir. Sonucun sağlaması da yapılmış olur.

Bir sonucun doğru olup olmadığını sağlaması 5 temel kavram ile yapılır. Duyarlılık (Sensitivity); markerların hasta olanı doğru bulma oranıdır. Özgüllük (Specificity); markerların hasta olmayanı doğru bulma

oranıdır. Pozitif kestirim; markerların hasta diye öngördüklerinin doğru çıkma oranıdır. Negatif kestirim; markerların hasta değil diye öngördüklerinin doğru çıkma oranıdır. Doğruluk (Accuracy); markerların hasta olanı ve hasta olmayanı doğru bulma oranıdır. İleri istatistik analizlerinde bu parametreler ile kurulan modelin doğruluğu gösterilir. Lojistik regresyonda kurulan modelin kullanılabilir olduğu bazı testler (Hosmer and Lemeshow Test vs) ile de kontrol edilebilir. Modelin doğruluğunu en iyi kontrol etme yöntemi ise modelin başka kurumlarda test edilmesidir. Kendi hastanemizdeki veriler ile bir tanı test bulduk diyelim bunu başka hastanelerde doğru tanı koyabiliyor mu diye test edebiliriz. Kendi ülkemizde bulduğumuz modelin doğruluğunu başka ülke verileri ile test edebiliriz. Temel istatistik ile ileri istatistik arasındaki en büyük fark burada oluşmaktadır. Ortaya çıkardığımız bilginin doğruluğu, tekrarlanabilmesi test edilebilir. P değerine bağlı değiliz. Sadece düşük ya da yüksek gibi sınırlı yorum yerine koşulları belirterek klinikte tanı koymamızı sağlayacak net yargılarda bulunabiliriz.

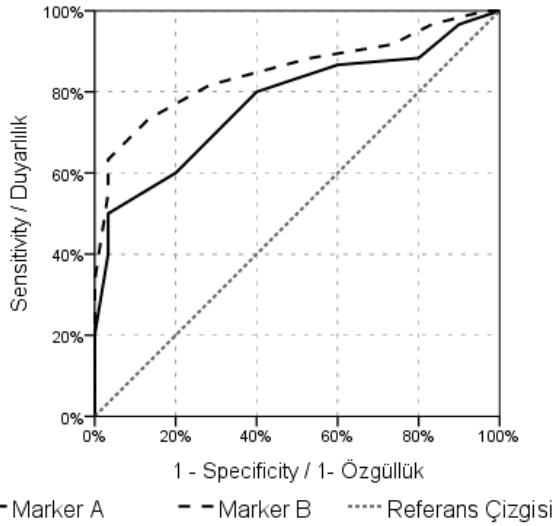
ROC eğrisi hem temel istatistiği hem de lojistik regresyonu tamamlayan analiz yöntemidir. ROC eğrisinde kutu ana kitleyi temsil eder. Hasta olanı ve olmayanı ayırma gücüne göre her bir değişken için bir çizgi çizer. Bu çizginin (eğrinin) altındaki kalan alan eğri altı alan olarak adlandırılır. Marker B de eğri altı alan 0.850 (0.779-0.922) olarak bulunmuştur. Bulduğumuz eğri altı alan Marker B nin hastalığı öngörmeye yaklaşık olarak %85 (%77.9-%92.2) lik bir doğruluğa sahip olacağına işaret eder. Marker B nin eğri altı alan büyüklüğü marker A dan daha yüksektir. Bu Marker B nin daha iyi bir öngörücü olduğuna işaret eder. ROC eğri grafiğinde referans çizgisi mevcuttur. P değerinin bazen yanıltıcı olabileceğine değinmiştik, bunu özetleyen bir noktadır. Eğer bir hasta olan ve olmayanı ayırmak istediğimizde yazı tura atarsak %50 doğruluğa sahip oluruz. Referans çizgisi de bunu belirtir. P değerinin anlamlılığında ziyade bizim öngörü düzeyimizin yazı tura atmaktan farklı olması önemlidir. ROC eğrisinde eğri altı alanın büyüklüğü çalışmaya göre değer kazanır. Daha önceden

çıkartarak tanıyı kolaylaştırmak mümkün mü diye bakılabilir. Farklı kesme (cut off) değerlerinin öngörü düzeylerini kıyaslamak için kullanılabilir. Öngörü düzeyini referans aldığımız zaman ROC eğrisini çok farklı amaçlar için kullanabiliriz.

ROC eğrisinin kesme (Cut off) değeri verdiğine dair yaygın kanı vardır. Ancak ROC eğrisi bir kesme (Cut off) değeri vermez. ROC eğrisi farklı cut off değerlerinde eğri altı alanı vererek öngörü düzeyini gösterebilir. Eğri altı alanı en yüksek kesme (Cut off) değeri en doğru kesme değeri değildir. Her çalışmanın amacına göre kesme değeri ayrı değerlendirilmelidir. Kesme (cut off) değerini hasta taramasında kullanacaksa duyarlılığı yüksek kesme değeri önemlidir. Eğer hasta yoğunluğunda hasta olmayanı ayırmak için kullanılacaksa özgüllüğü yüksek kesme değeri faydalı olabilir. Tanıda kullanılacak ise tek bir kesme değeri yerine birden fazla kesme değeri ile hafif-orta-yüksek risk gibi değerlendirmek daha doğru sonuç verecektir.

ROC eğrisi ve lojistik regresyonun birçok avantajı vardır. Çalışmaya başlamadan bazı planlamaları doğru yapmak gerekir. Her iki test içinde olaya (event) ihtiyaç vardır. Üç ve üstü grup arasında bu analizleri kullanamayız. İki grubu olması gerekir. Ancak iki grup olması da her zaman uygun değildir. Örnek olarak A tedavisi ve B tedavisini içeren iki grup olay (event) kavramına girmez. A tedavisi veya B tedavisini öngörmek diye bir amaç olamaz. Hasta olan ve olmayan grupta hastalığı öngörmeye çalışılabilir. Özellikle lojistik regresyon için sebep-sonuç kavramları önemlidir. Komplikasyona etki eden faktörleri araştırmaya çalışalım. Hedef nokta komplikasyonu öngörmek ise lojistik regresyona mortaliteyi alamayız. Çünkü mortalite komplikasyon sebebi olamaz. Mortalite komplikasyonun sonucu olabilir. Bu sebeple analize katacağımız her değişkenin sebep mi sonuç mu olduğunu değerlendirmemiz gerekir.

En sık kullandığımız paket programları bu tarz analizleri binom dağılım üzerine kurgulamıştır. Bu yüzden daha doğru yorum yapabilmek için verilerimizi binom dağılıma uygun olacak şekilde 0-1 olarak



öngörü düzeyi yüksek marker lar mevcut ise yeni bulduğumuz markerin eğri altı alanının yüksek olması gerekir ki ek fayda sağlasın. Tanı için önceden hiç belirleyici yok ise 0.600 gibi düşük bir eğri altı alan bile değerli olabilir. ROC eğrisi tek bir değişkenin etkisini göstermek için kullanılabilir. ROC eğrisi birden fazla değişkenin öngörü düzeylerini kullanmak için kullanılabilir. Birden fazla değişkenin birlikte değerlendirilmesi ile öngörü düzeyinde artış var mı, bunu değerlendirmek için kullanılabilir. Daha önceden kendini ispatlamış bir ölçüğe değişken ekleyip, değişken

Tablo-3

	Eđri AltıAlan (AUC) (%95 GA)		P
Marker A	0.774	0.689 - 0.859	0.000
Marker B	0.850	0.779 - 0.922	0.000

kodlamamız yorumlarımızı doğru yapmamıza fırsat sağlayacaktır. Komplikasyon olmayan grup 0, komplikasyon olan grup 1 olarak kodlanabilir.

Lojistik regresyonda en sık yapılan hata modeli indirgememektir. Model bağımlı etkenleri ayıklayarak, bağımsız etkenleri karşımıza çıkartmalıdır. Tablo 2 de olduğu gibi tek değişkenli modelde sonuçları ve çok değişkenli modelde bağımlı değişkenlerin elendiği bir kurgu olmalıdır. Bağımlı değişkenler elenmediği takdirde kurulan modelde gereksiz çok sayıda değişken olacaktır. Eğer indirgeme yapmazsak hastalığın tanısını koymak için birkaç soru yerine onlarca gereksiz soru sormamız gerekir. Bu ölçek yapma, algoritma yapma mantığına aykırıdır.

ROC eğrisi ve lojistik regresyonun doğru sonuç vermesi için dikkat edilmesi gereken diğer bir konu ise eksik (missing value) veridir. Eğer eksik veriler varsa yapacağımız çok değişkenli analizler yanlış sonuç verir. A marker ı 10 hasta bakılmadıysa, B markerı 20 hastaya bakılmadıysa analiz 30 hasta eksik yapılabilir. Bu sonuç tek değişkenli analiz veya temel istatistik sonuçlarından farklılaşabilir.

Özet olarak lojistik regresyon ve ROC eğrisi çok sık kullanmamız gereken analiz yöntemleridir. Bu analiz yöntemlerini ileri istatistik olarak değil. **GEREKLİ İSTATİSTİK** olarak görmek doğru olacaktır. Çalışmaların planlanmasında bu analiz yöntemlerini nasıl kullanabileceğimizi değerlendirmek faydalı olacaktır.

Kaynaklar

1. İstatistik kitabı, Ertan Koç-İstatistik Akademisi



Sağkalım Analizleri

Dr. Fehmi Narter

Sağkalım analizleri özellikle tıp alanında istatistiksel olarak büyük önem taşırlar. Çalışmalarda bu konuda birçok tanıma rastlıyoruz ve bunları birbirleriyle karşılaştırarak tedavi metotları hakkında bir sonuca ulaşmaya çalışıyoruz. Bu konuda sık rastladığımız bazı kavramlar aşağıda listelenmiştir.

- Progresyonsuz sağkalım (PFS, progression free survival)
- Hastalığa özgü sağkalım (DSS, disease specific survival)
- Hastaliksiz sağkalım (DFS, disease free survival)
- Genel sağkalım (OS, overall survival)
- Olaysız (semptomsuz) sağkalım (EFS, event free survival)
- Yetersizlik olmadan sağkalım (FFS, failure free survival)
- Metastazsız sağkalım (MFS, metastasis free survival)
- Kansere özgü sağkalım (CSS, cancer specific survival)
- Biyokimyasal rekürrensiz sağkalım (BRFS, biochemical recurrence free survival)
- Komplikasyonsuz sağkalım (CFS, complication free survival)
- Relapsız sağkalım (RFS, relaps free survival)

Çok sayıda bu şekilde istatistiksel tanımlama yapılabilir. Ancak bunlardan ilk dört tanesi daha sık karşımıza çıkan kavramlardır. Sağkalım analizi çeşitli konularda risk etkenlerini, prognostik öğeleri, tedavi başarılarını vb ortaya koyabilmesi açısından önemlidir. Sağkalım analizi başarısızlığı tanımlayan ölüm gibi olayların ortaya çıkmasına kadar ki süreyi analiz eder (ölüm, metastaz, nüks veya tedavi ile değişen bir olay). Bu konunun kavramları arasında Hazard oranı (HR), Kaplan-Meier analizi (product limit), Cox regresyon yöntemi, yaşam tabloları (life table) yer almaktadır.

Genellikle kohort ya da klinik deneme çalışmalarında bu analizlerle karşılaşırız. Sağkalım verileri, yanıt değişkeni olan olay meydana gelene kadar geçen sürenin verisidir. Olay kavramı (event, end point) ölüm, hastalık, tekrarlama, iyileşme gibi araştırmacının ilgilendiği herhangi bir olay olabilir, çok değişkendir ve yukarıdaki tanımlarda belirtilen noktaya kadar geçen süre tanımlanmaktadır (time to event). Sağkalım eğrileri, sağkalım değerinin Y ekseninde % olarak, X ekseninde süreye göre temsil edildiği basit grafiklerdir.

Genel sağkalım için, başlangıç noktası tanı tarihi, ilk tedavi tarihi veya randomizasyon tarihidir. Anlaşılacağı gibi araştırma sonlandırıldığı zaman, izlenen tüm olguların mutlaka belli bir özel konuma (ölüm gibi) erişmiş olmaları beklenemez. Üzerinde değerlendirme yapılacak tüm olguların araştırmaya aynı anda, hepsi için geçerli bir başlangıç noktası ile katılmaları olanaksızdır. Bu dinamik süreçte araştırma için belirlenmiş belli bir süre içinde, çeşitli anlarda araştırmaya izlenmeye katılan yeni olgular olacağı gibi, belli bir özelliğe erişmesi (ölmesi) nedeni ile izlemeden çıkan örnekler de bulunur. Sağkalım çözümlemesini gerçekleştirirken kullanılacak olan yöntem hangisi olursa olsun öncelikle bu farklı zamanlarda çalışmaya giren, farklı zamanlarda farklı nedenlerle sonlanan hatta araştırma bitirildiğinde izlemesi süren tüm olgular, tek bir anda izlenmeye başlamış gibi bir düzenleme yapılır. Bazı çalışmalarda her olguyu kendi izleme süresi ile belirtmek ve değerlendirmeleri bu izleme süreleri açısından yapmak gerekir. Buna karşılık özellikle olgu sayısının çok yüksek olduğu konularda araştırmacının kendi öngöreceği süre aralıkları için değerlendirmeler yapılması söz konusudur. Bu iki ayrı değerlendirme tipi, sağkalım çözümlerinin iki ayrı ana yöntemini de belirtmektedir; ilki sınır çarpım (*product limit*) kestirimini amaçlayan "Kaplan Meier" yöntemi, diğeri ise "Yaşam Tablosu" (Cutler Ederer) yöntemidir. Bunlar dışında, süreç içinde kaybolan olguları hesaplardan çıkartmadan, araştırma başlangıç toplamındaki izleme

zamanları düzeltilerek araştırmaya katılan tüm olgular sayı üzerinden de hesaplama yapılabilir. "Araştırmaya alınan" (*intention to treat*) tüm olguların bu yaklaşımı belli ölçülerde gerçek yaşamdaki koşulları daha gerçekçi yansıtabilmektedir.

Sağkalım çalışmalarının amaçları: 1) Bir grup hasta için olay olana kadar geçen süreyi tahmin etmek (Tedavi sonrası beklenen yaşam sürelerini tahmin etmek, hastalıkların doğal seyirlerinin izlemek) 2) Gruplar arası sağkalım sürelerinin karşılaştırılması (tedavi etkinliği) 3) Risk faktörlerinin sağkalım süresine etkisinin belirlenmesidir. Sansürlü verileri yok saydığı için t-testi veya doğrusal regresyon ile gruplar arası sağkalım sürelerinde fark analiz edilemez. Yine benzer şekilde, zamanı yok saydığından RR (rölatif risk) / OR (Odds oranı) veya lojistik regresyonla grupların olay oranları araştırılmaz.

Takip süresi (follow-up) ve kullanılan istatistiksel yöntemlerin ve hedef/ hedeflerin tanımı (end-point), olayların (event) tanımı yapılmalıdır (örn: sağkalım için ölüm; hastaliksiz sağkalım için nüks veya hastaliksiz ölüm). Prospektif (ileriye dönük) izlemlerde en büyük sorunlardan biri zaman içinde takibin tam olmamasıdır.

1. Olayın gerçekleştiği olgular dışarıda bırakılır (örn: ölen olguların dışlanması)
2. Olayın gerçekleşmediği olgular kullanılarak (patients at risk) minimum, maksimum, ortalama, ortanca, vs. takip süresi değerleri tanımlanır. Sağkalım analizleri için en az iki değişkene ihtiyaç vardır;

1. Süre: (t_i = ilgilenilen olay gerçekleşene kadar geçen süre), yıl, ay, hafta, gün vb olabildiği gibi olgunun ortaya çıktığı andaki yaş da olabilir. Bu süre *sağkalım* süresi ya da *başarısızlık süresi* olarak da adlandırılır. Başlangıç noktasının her denek için aynı tarihte olması gerekmez. Farklı tarihte de olabilir. Önemli olan izlemeye başlanıldığı an olup bunun her denek için aynı olmasıdır (*Tanı tarihi, Tedaviye başlama tarihi, Cerrahi müdahale vb*).

2. Sansürlü veri: ($c_i=1$ t_i anında olay gerçekleşmişse; $c_i=0$ t_i anında olay gerçekleşmemişse), sağkalım çözümlerinde sağkalım süresi tam bilinmeyen vakalar

sansürlü (kayıbolan, tamamlanmamış, censored) olarak adlandırılır. Genelde üç farklı nedenden oluşan sansürlü gözlemler vardır. Bu kavram sağkalım süresi kesin olarak bilinmeyen gözlemlerin verileridir (çalışma süresince olayın gerçekleşmemesi, hastanın takipten çıkması, başka bir olayın yaşanması ile takipten çıkarılması).

Sağkalım analizi sonuçlarının sunulması;

- Medyan izleme süresi (min.-max.)
- Olay/denek sayısı
- Ortalama sağkalım süresi (Mean survival time) \pm standart hata (SE) (%95 CI:...) N sayılı dizide ortalama yaşam süresi tahmini (μ) ile gösterilir.
- Medyan sağkalım süresi (Median survival time) \pm standart hata (SE) (%95 CI:...)
- x yılda kümülatif sağkalım hızı $R \pm SE$ (%95 CI:...)
- Sağkalım eğrilerinin istatistiksel karşılaştırma sonucu (Log rank = ...; d.f. = ...; p=...) şeklindedir.

Sağkalım eğrileri, *Kaplan-Meier yöntemi* ile hesaplanır. Eğrilerin karşılaştırılması Tek değişkenli analizler (logrank veya Wilcoxon [Gehan] testi) 2 veya daha fazla sağkalım eğrilerinin karşılaştırılması (*hangi olay olursa olsun*) (örn: A ve B tedavileri) *logrank testi* ile gerçekleştirilir. Çok değişkenli analizler için ise *Cox modeli* kullanılır, olayı etkileyen birçok faktörün aynı anda değerlendirilerek, bunlardan anlamlı olarak bağımsız faktörlerin belirlenmesi (ör: aynı anda yaş, TNM, tedavi tipi, vs. gibi) ve bunların arasında olayı etkileyen bağımsız faktörün saptanması amaçlanır.

Sağkalım fonksiyonu:

$$S(t) = 1 - P(T \leq t)$$

En fazla t süresi kadar yaşama olasılığı

En az t süresi kadar yaşama olasılığı.

Örnek: Eğer $t=100$ yıl ise, $S(t=100)$ = en az 100 yıl yaşama olasılığı

N birimin gözlenen yaşam süreleri $t_1 < t_2 < t_3 < \dots < t_N$ biçiminde sıralanmış olsun. Formülün uygulanması ile N bireyin yaşam fonksiyonu hesaplanır.

Hazard fonksiyonu:

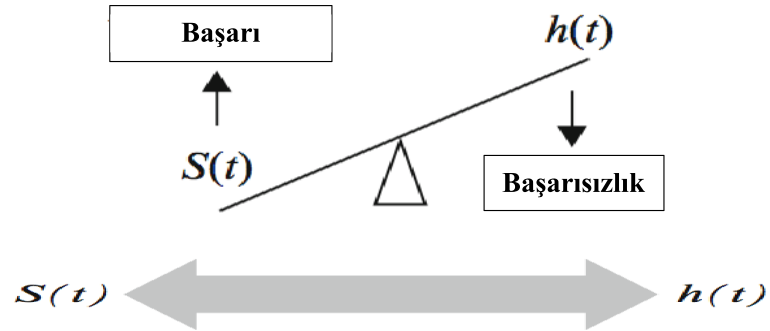
$$h(t) = \frac{f(t)}{S(t)}$$

t anında ilgili olayın (ölüm) meydana gelme olasılığı

En az t süresi kadar yaşama olasılığı

En az t süresi kadar yaşamış bir kişinin t anındaki anlık ölüm riski

Sağkalıma karşı Hazard



Yaşam tablosu (Actuarial tables) yöntemi

Yaşam tablosu yöntemi, araştırmacının olgu sonuçlarını, araştırmacı tarafından belirlenen zaman aralıkları çerçevesinde gruplayarak değerlendiren bir yöntemdir. Belli aralıklar içinde değerlendirme yapmanın getirdiği bazı sakıncalarına karşın, diğer yöntemlere (Kaplan Meier) göre de getirdiği bazı özel ölçütleri vardır. Araştırmacı sağkalım aralıklarını amacına göre düzenleyebilir.

- Her aralıkta olgu ve kayıpların tek düze dağıldığı varsayılır.
- Sağkalım olasılıklarının hesaplanmasında kayıpların yarısında aralığın ortasına kadar olgu gözlenmediği varsayımı kullanılır.

olgular çalışmaya araştırmacının sürdüğü herhangi bir zamanda katılabilirler, birim olarak düşünülen sürecin (yıl, ay, vs) sonuna kadar normal olarak izlenebilecekleri gibi belli bir anda kaybolabilir, çalışmadan çıkarılabilir veya araştırılan özelliği gösterir duruma girebilirler (ölüm, remisyon, ilaca yanıt vs.). Yöntem başlangıçta % 100 olan yaşama olasılığının belirli süreç aralıkları ile eriştiği düşüğü yeni yaşama olasılık düzeylerini belirler. Değerlendirme, temelde belli özel duruma giren olguları, bir özel birim süre içinde değerlendirmek yönünde yapılır, ancak bir döneme normal olarak başlayan tüm olguları bu dönemin bitimine değin, özel durumların dışında da, izlemek olanaklı değildir dolayısı ile aslında bir „risk grubu“ndan söz edilir. Bu risk grubu başlayan olgulardan „tamamlanamayan“ olguların çıkarılması ile elde edilen bir sayıdır. Dönem sonunun ölüm olasılığının hesabında ise ölenler / risk grubu oranı kullanılır. Yaşama olasılığı ise doğal olarak 1-ölüm olasılığıdır. Çalışmanın başından beri yaşama olasılığı ise birbirini izleyen dönemlerin yaşama olasılıkları çarpımına eşittir. Grafik gösterimde de araştırmacının seçtiği dönemlere göre,

Yaşam tablosu yönteminin toplam olgu sayısının oldukça yüksek olduğu çalışma örnekleri için (n100) kullanılması önerilmektedir. Daha önce vurgulandığı gibi

dönem sonu yaşam olasılıkları belirtilir. Başlangıç anı yaşam olasılığı "1" (% 100)'dir. Dönem sonu olasılık değer noktaları birleştirilerek çizim yapılır ve dönem içinde ölümler olması nedeni ile dönem içi yaşam olasılıklarının azalarak sürmesi gösterilmiş olur.

Kaplan-Meier yöntemi (Varsayımlar)

Yaşam sürelerine ilişkin, verileri zaman aralıklarına bölmeden yaşam ve ölüm fonksiyonlarını hesaplamayı sağlayan bir yöntemdir. Kaplan-Meier yönteminde sadece ölümler üzerinden yaşam olasılığı hesaplanır. Yaşam olasılığı ise ölüm olayının gerçekleştiği ana ilişkin olarak hesaplanır. Kaplan Meier yönteminde kesin ölüm tarihi kullanıldığı için nokta yaşam olasılığı bulunur.

Bu Yöntem,

- Deneklerin çalışmaya katılma tarihlerinin kesinlikle bilinmesini
- Olgunun ortaya çıktığı zamanın kesinlikle bilinmesini
- Kayıpların (sansürün) ortaya çıktığı zamanın kesinlikle bilinmesini
- Olgunun ve kayıp(lar)ın aynı anda ortaya çıkmamasını gerektirir.

Kaplan-Meier yöntemini diğer parametrik yöntemden ayıran önemli özelliği, sağkalım olasılıklarının hesaplanmasında sansürlü denekleri işleme katmamasıdır. Yaşam tablosu yöntemine göre temel fark araştırmacının belirlediği zaman aralıkları çerçevesinde değil, gerçekleşen her ölümden beliren yeni eşit olmayan zamanlarda kestirim yapılmasıdır. Özetle her ölüm olgusu zamanı kadar aynı andaki ölümler tek zaman belirler kestirim noktası oluşturulur. Buna göre, herhangi bir ölümün olduğu andaki ölüm olasılığı; o andaki ölenlerin sayısı / o anda izlenenler (ölen+izlemi devam edenler) şeklinde hesaplanır. O ana ilişkin yaşama olasılığı yine 1-ölüm olasılığıdır ve baştan beri yaşama olasılığı da geride kalan tüm dönemlerin yaşama olasılıkları çarpımına eşittir. En uzun süreli izlemeye sahip olan olgu bir ölüm araştırılan özel durum değil de, "tamamlanmamış" bir olgu ise, ortalama

yaşam kestirimi, gerçek ortalama dan daha düşük bir değeri yansıtmaktadır. Bu tür durumlarda tüm çalışma için bir „genel“ yaşam süresi kestiriminden kaçınılmalı, sadece belli bir "süre" içinde ortalama yaşam beklentisinden söz edilmesi önerilmektedir. Yüzdeler kestiriminde birbirini izleyen ölüm anları arasında yaşam olasılığı değişmez varsayıldığından istenen yüzdeliğin (yaşam olasılığı olarak) altına düşülen ilk ölüm zamanı ilgili yüzdelik kabul edilir.

Yaşam tablosu yönteminden farklı olarak, her ölüm zamanını ayrıca değerlendirmesi nedeniyle, Kaplan Meier yönteminde sağkalım olasılıkları, adımsal (*step*) fonksiyon biçiminde belirir. Bir önceki (ölüm) kestirim anından sonra, yeni ölüm zamanına dek, başka ölüm olmaması nedeni ile olasılık değişmez kalmaktadır ve grafiksel gösterimde bir merdiven görüntüsü ortaya çıkmaktadır.

Sağkalım Eğrilerinin Karşılaştırılması

Birden fazla kümede sağkalım dağılımlarının genel olarak tüm izleme sürecini göz önünde bulundurur şekilde kıyaslanması tıpta sık kullanılır. Sağkalım dağılımlarının tipi bazen bilinmediği, bazen ise bilinen ve rahat çözülen kalıplara uymadığı için genellikle parametrik olmayan testler kullanılır. Çok özel koşullar dışında, hemen tüm araştırmalarda "tamamlanmamış" olgular bulunduğu parametrik olmayan testler bu duruma göre tasarlanmıştır. Tamamlanmamış olgusu bulunmayan iki bağımsız topluluğun yaşam süreleri standart nonparametrik testlerle kıyaslanabilir (*MannWhitney U* gibi). Aynı koşullarda bağımlı örnekler için ise Wilcoxon Signed Rank Testi (işaret testi) kullanılabilir. Genelde rastlanan; "tamamlanmamış" örneklerin bulunduğu 2 ayrı grubun kıyaslanması ise şu şekilde yapılmaktadır; *Mantel Haenszel* Ki-kare vb testi.

İkiden fazla örnek grubunun kıyaslanması için ise *Kruskal Wallis* H testinin değişimlenmiş bir uygulamasından yararlanılmaktadır. Yukarıda adı geçen tüm testlerden, araştırmacının tek başına uygulaması en kolay olanı *Mantel Haenszel* yöntemidir.

Genellikle hazır paket programlarda ise Logrank yöntemi uygulanır. Bazı özel hesaplamalarla iki kıyas kümesinin sadece belli bir dönemdeki yaşama olasılıklarını değerlendirmek de olanaklıdır.

Bazı durumlarda aynı hastalığa yakalanan bireyler iki ya da daha fazla gruba ayrılarak her gruba farklı bir tedavi yöntemi uygulanabilir. Farklı yöntemlerle tedavi edilen hastalar için birden fazla yaşam fonksiyonu hesaplanabilir. Bu yaşam fonksiyonlarının birbirleriyle olan farklılıkları test edilebilir. Farklı gruplara göre elde edilen yaşam fonksiyonlarının karşılaştırılmasını yapmak için yaygın olarak kullanılan test yöntemleri:

1. Mantel- Haenzel (log Rank), Cox *Mantel* testi
2. Gehan Breslow-Wilcoxon Testi
3. Tarone-Ware Testi
4. Peto-Prentice Testi'dir.

Cox Regresyon yöntemi

Regresyon analizinde sonuç değişken ile açıklayıcı değişkenler arasındaki neden-sonuç ilişkisi kurulurken bu değişkenlerin uyması gereken koşullar vardır. Yaşamsal verilerde birçok bağımlı değişken normal dağılmakta, bağımsız değişkenler birbirleri ile korelasyon göstermektedir. Bu nedenle yaşamsal verilerin neden-sonuç ilişkilerini incelerken çoklu regresyon analizi uygulanmamaktadır. Yaşamsal verilerin nedensellik analizlerinde Cox Regresyon Yöntemi'nden yararlanılır.

Bir izlem araştırmasında incelenen bağımlı değişken (Yaşam Süresi), bir hastalığa yakalanan bireylerin ölüm zamanlarına kadar geçen izlem süreleri (ölüm) ise; açıklayıcı değişkenler, bu değişken üzerinde etkide bulunan faktör değişkenler (Yaş, cins, tedavi türü vb) olur. Tamamlanmamış izlem verilerinde, bağımlı değişken ile bağımsız değişkenler (prognostic factors, covariates) arasındaki neden-sonuç bağıntısını ortaya koymak için yararlanılan regresyon yöntemine Cox regresyon yöntemi adı verilir. Mümkün olan tüm parametreler ile tek değişkenli analizle (logrank testi) iki tedavi grubu karşılaştırılır (örn: A ve B tedavisi, yaş < 50 ve > 50, erkek/kadın, vs. Sonra, örnek büyüklüğüne bağlı olarak sadece



belli bir anlamlılık gösteren (örn: $p < 0.10$; denek ve yöntemde hangi değer alındığı mutlaka belirtilmeli) faktörler kullanılarak çok işlemlili analiz yapılır. Genel olarak, Cox modelinde sayısal sürekli veri kullanımından kaçınılmalıdır. Bunun için, bu tür veriler Evet/Hayır verisi şekline dönüştürülmelidir. Örneğin, yaş değişkenini sürekli veri olarak kullanmaktansa ortanca (veya çeyrek) değer üstü/altı ($yaş < 40$ ve > 40) olarak kullanılmalıdır. Hangi tür çok değişkenli analiz yapılırsa yapılsın, veri tabanımızda eksik veri olmamalıdır. Bazı olgularda veri eksikse, çok değişkenli analizde sadece verileri tam olan olgular hesaba katılmalıdır. Burada unutulmaması gereken bir nokta, analize alınan ve alınmayan olguların akibetlerinin karşılaştırılması: iki grup arasında fark olmamalıdır.

Regresyon analizinde sonuç değişken ile açıklayıcı değişkenler arasındaki neden-sonuç ilişkisi kurulurken bu değişkenlerin uyması gereken koşullar vardır. Yaşamsal verilerde birçok bağımlı değişken normal dağılmakta, bağımsız değişkenler birbirleri ile korelasyon göstermektedir. Bu nedenle yaşamsal verilerin neden sonuç ilişkilerini incelerken çoklu regresyon analizi uygulanamamaktadır. Yaşamsal verilerin nedensellik analizlerinde Cox Regresyon yönteminden yararlanılır.

Cox yönteminde, bir riske maruz kalan kişinin izlendiği T zaman süresinin herhangi bir t zamanında, ölümlerle karşılaşacağı varsayılır. Bireylerin t_0 zamanındaki yaşam süreleri, logaritmik olarak azalan bir fonksiyona sahip olur. Eğer Z_1 kadar değişken, yaşam süresinin loglinear bir fonksiyonu olarak, loglinear modellerle incelemeye alınabiliyorsa, Cox'un Proportional Hazard Regression yönteminden yararlanarak her bir prognostik değişkenin yaşam süresi üzerine etkisi incelenebilir.

Z_1 bağımsız değişkenler vektörü Z ve yaşam süresi t olsun. Bir bireyin Z_1 ortak değişkenlerine göre ölüm fonksiyonu $h(t; z)$ olarak alınabilir. Buna göre orantısal ölüm modeli (Proportional hazard modeli, bir başka anlatımla Z açıklayıcı değişkenlere göre regresyon modeli), tek değişkenli ve

çoklu Cox regresyon modeli ihtiyaca göre hesaplamalarda kullanılır.

Kaynaklar

1. Sağlıkım (yaşam) çözümlemesi. Yrd. Doç. Dr. Ünal Erkorkmaz Sakarya Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı, sunumu.
2. Temel kavramlar 8/8/2014 Sağlıkım çözümlemesi ilkeleri; Şenocak
3. <http://www.ctf.edu.tr/dergi/online/1999v30/s4/994s1.htm> 2/4
4. Sağlıkım Analizleri. Çocuk Nefroloji Derneği 14-15 nisan 2018, Ankara



SORULAR

1. Aşağıdakilerden hangisi ileri istatistik konusu değildir?
 - a. Sağkalım analizleri
 - b. ROC eğrileri
 - c. Korelasyon analizleri
 - d. Regresyon analizleri
 - e. Hepsi
2. Ortalama hangi veriler için kullanılır?
 - a. Nominal
 - b. Nicel
 - c. Ordinal
 - d. Nitel
 - e. Hepsi
3. Hangisi normal dağılım testi değildir?
 - a. Shapiro-Wilks,
 - b. D'Agostino-Pearson
 - c. Levene Testi
 - d. Kolmogorov-Smirnov
 - e. Hiçbiri
4. Hangisi tanımlayıcı istatistiki değer değildir?
 - a. Olasılık değeri
 - b. Ortalama değeri
 - c. Standart sapma değeri
 - d. Oran değeri
 - e. Ortanca değeri
5. Bağımsız gruplarda kullanılmayan test hangisidir?
 - a. Kruskal Wallis testi
 - b. Fisher testi
 - c. Mann Whitney testi
 - d. Wilcoxon testi
 - e. Ki-kare testi
6. Epidemiyolojik araştırma türü olmayanı işaretleyiniz?
 - a. Kohort çalışma
 - b. Olgu kontrol çalışması
 - c. Deneysel çalışmalar
 - d. Hiçbiri
 - e. Hepsi
7. Uygun testi seçerken aşağıdakilerden hangisinin göz önünde bulundurulması gerekmez?
 - a. Verilerin normal dağılıma uygunluğu
 - b. p değeri
 - c. Grupların bağımlı mı bağımsız mı olduğunu,
 - d. Test edilen değişkenin tipi
 - e. Bazı gruplardaki örnek sayıları
8. Hangisi parametrik test değildir?
 - a. Varyans analizi
 - b. Mann-Whitney U testi
 - c. ANOVA testi
 - d. t-testi
 - e. Hiçbiri
9. Hangisi ANOVA sonrası çoklu karşılaştırma testi değildir?
 - a. Tukey HSD
 - b. Student-Newman-Keuls
 - c. Kruskal Wallis
 - d. Scheffe
 - e. Bonferroni
10. Hangisi parametrik korelasyon analiz testidir?
 - a. Spearman
 - b. ROC analizi
 - c. Pearson testi
 - d. Nokta çift serili
 - e. Hiçbiri