BİYOİSTATİSTİK DÖNEM I

Prof.Dr.Refik Burgut

Biyoistatistik ABD
E-mail:refik@cu.edu.tr
Tel:3094-95

1

İÇERIK

- Biyoistatistik Bilimi Ne Yapar?
- Temel tanımlar
  - Veri Tipleri İle İlgili Tanımlamalar
- Tıbbi araştırmalarda veri kaynakları
  - Rutin Veri Kaynakları
  - Özel Araştırmalarla sağlanan veriler

2

- Klinik araştırmalarda Biyoistatistik
  - Tanıda
  - Prognozda
  - tedavinin etkinliklerini saptamada
- Koruyucu Hekimlikte Biyoistatistik
  - Risk Faktörlerinin saptanması
  - İlişkinin saptanması
  - Uyumluluğun saptanması

3

- Gözlemsel Araştırmalar
  - Kesitsel araştırmalar
  - Ekolojik Araştırmalar
  - Vaka-Kontrol Araştırmaları
  - Kohort Araştırmaları
- Deneysel Araştırmalar
  - Terapatik deneysel araştırmalar
    - Kontrollu randomize Kliniksel araştırmalar
  - Preventif deneysel araştırmalar

4

- Verilerin Özetlenmesi
  - Şekillerle
  - -Ölçütlerle
  - ve Olasılık Formülleri ile
- Dağılımlar
  - Normal Dağılım ve tıpta kullanımı
  - Binom ve Poisson Dağılımları ve tıpta kullanımları
  - Ve diğer dağılımlar

5

- Örnekleme ve Örneklem Dağılımı
- İstatistiksel Karar
  - Tahmin Yöntemleri (Nokta Tahmini ve Güven Aralıkları)
  - Önemlilik Testleri
- Önemlilik testlerinde kullanılan test istatistikleri

6

- Univariate Analizler
- Multivariate Analizler
- META ANALİZİ

7

BAŞVURULABILECEK KAYNAKLAR

- Jekel et al. Epidemiology, Biostatistics, and preventive Medicine, 3rd EditionSounders and Company
- Dawsan B,Trapp, RG.Basic and Clinical Biostatistics. Fourt Edition.Lange medical Books/Mc Graw Hill, London 2004.
- Piantadosi S. Clinical Trials A methodological Perspective.2nd Edition. Wiley, 2005.
- Rosner, B. Fundamentals of Biostatistics, 5th Edition,International Thompson Publishing, 2003
- Zhou X.H., Obuchowski N.A., McClish D.K.,Statistical Methods in Diagnostic Medicine, Wiley, 2002
- Woodward, M. Epidemiology: Study Design and Data Analysis. Second edition. Chapman and Hall.London, 2005
- Chow SC, Shao JC, and Wang H. Sample Size calculations in Clinical Trials.Marcel Dekker Inc. NewYork, 2003
- Hayran O.(Çeviren), Lwanga SK, Lemeshow S., Sağlık Araştırmalarında Ornek Büyüklüğünün Belirlenmesi, WHO, 1992.

8

ARAŞTIRMA

- Bilinmeyen olay ve etmenleri ortaya çıkarabilmek, bilinenleri geliştirmek,
- herhangi bir konuyu aydınlatmak,
- bir soruna çözüm getirmek,
- belirli kuram ve kavramlar yada yasalara ulaşabilmek amacı ile yapılan bir çalışma olarak tanımlanabilir

9

ARAŞTIRMA

- Bir araştırma da en önemli husus araştırmanın bilimsel yöntemlerle planlanması, yürütülmesi ve sonuçlandırılmalıdır.
- Bir araştırmanın bilimsel olmasını sağlayan en önemli unsur İstatistik yöntemlerin kullanılmasıdır.
- Bu nedenle İstatistik bilimi araştırmanın ayrılmaz bir parçası olup her araştırıcının Biyoistatistikçi ile çalışması ve kendisinin bu konularda yetiştirmesi gereklidir

10

BİYOİSTATİSTİK NİÇİN GEREKLİDİR

1. **Biyoistatistiğin prensip ve yöntemlerinin uygulamasının bilinmesi Sadece Biyolojik ve tıbbı bilimlerin anlaşılması için değil aynı zamanda sağlık personelinin etkili bir pratiğinin yapılması içinde gereklidir.**

2. **Biyolojik, laboratuvar ve klinik verilerdeki yaygınlık nedeni ile bunların anlaşılmasında ve yorumlanmasında Biyoistatistik bilimi gereklidir**

Doktor;

- Laboratuvar test sonuçlarını ,
- Klinik vizitlerde hastanın durumunu;
- Fizyolojik, gözlem ve ölçümlerdeki varyasyonları göz önüne alarak bir yorumlama getirmek zorundadırlar

12

BİYOİSTATİSTİK NİÇİN GEREKLİDİR

- Doktor;
  - laboratuvar test sonuçlarını ,
  - klinik vizitlerde hastanın durumunu;
  - fizyolojik, gözlem ve ölçümlerdeki varyasyonları gözönüne alarak bir yorumlama getirmek zorundadırlar

13

BİYOİSTATİSTİK NİÇİN GEREKLİDİR

- Aşağıdaki aşamalarda karar için mutlaka,
- biyoistatistiksel bilgi gerekmektedir
  - Tanı
  - Prognoz ve
  - Tedavilerin etkinliğinin belirlenmesinde
  - Hastalık etiyolojilerinin (Risk faktörlerinin) belirlenmesinde

14

BİYOİSTATİSTİK NİÇİN GEREKLİDİR

Doktor;

- ✓ tedavi ettiği hastalıkla ilgili istatistik ve epidemiyolojik bulguları;
  - ✓ bu hastalıktan korunmak veya kaçınabilmesi hakkında
- hastasına en iyi öneriyi verebilmesi için bilmek zorundadır

15

Sağlık istatistiklerinin elde edildiği verilerin sağlanmasında;

- Hekim temel çalışanlar olup bu veriler bir ülkede sağlık hizmetlerinin planlanması ve değerlendirmesinde kullanılırlar
- hekim bu sağlık göstergelerini nasıl değerlendirileceğini ve bunların sağlık seviyelerini göstermede nasıl kullanılacağını bilmek zorundadır.
- İyi bir biyoistatistik eğitim alan hekim, o ülkenin sağlık informasyon sistemlerine güvenilir katkılarda bulunması beklenir.

16

Bu derste;

istatistik terminolojisinin değişik açıklamalarını,

istatistiksel yöntemlerle ilişkilendirerek;

- Hastalık tanısının konulması ve toplumda sağlık problemlerinin saptanması
- Toplumlarda veya kişilerde hastalıktan korunma programlarının sonucunun predikte edilmesinde
- Bireysel hastalarda uygun tedavi formlarının belirlenmesinde
- Sağlık yönetimi ve planlamasında
- Tıbbı araştırmaların planlanması, uygulanması, analizi, yorumlanması ve rapor edilmesinde

17

BİYOİSTATİSTİK;

- Tıp ve sağlık bilimleri alanlarında veri toplanması, özetleme, analiz ve değerlendirmede istatistiksel yöntemleri kullanan bir bilim dalıdır.
- Eşdeğer veya benzer tanımlamalar Biyometri içinde yapılabilir. Yine istatistiği aynı amaçlarla kullanan bilim dalları içinde Ekonometri, psikometri vb gibi bilim dalları da mevcuttur.

18

BİYOİSTATİSTİK

- Biyoistatistiksel yöntemler; hem toplum bazında hem de kişisel seviyede kullanılmaktadır.
- Toplumsal seviyede, sağlık durumlarının takibi ve değerlendirilmesi, herhangi bir korunma amaçlı uygulanmış olan programların etkinliklerini belirleme de kullanıldığı gibi kişisel seviyede en olabilir tanının konulmasında, prognostik gidişin veya değişik tedavi çeşitlerinin etkisinin belirlenmesinde kullanılır.

19

BİYOİSTATİSTİK

- Tıp literatürünün kritiğinin yapılabilmesi ve anlaşılabilmesi için biyoistatistik bilgisi gereklidir.
- Tıp ve sağlık alanlarındaki klinik, saha araştırmaları gibi her türlü araştırmaların planlanması, gerçekleştirilmesi, analiz ve yorumlanması için gereklidir.
- Biyoistatistiksel tekniklerin kullanılmadan gerçekleştirilen araştırmalar bilimsel araştırmalar değildir.

20

BİYOİSTATİSTİK

- **İstatistik, günlük yaşamda kullanılışı**
- Veri, nümeriksel sayılarla ifade edilen gözlemler veya sayısal olarak ifade edilen bilgiler olarak tanımlanır.

- Bir şehrin değişik sağlık ocağı bölgelerinde çalışan hekimler
- Belirli bir zaman dilimi içerinde bir hastanede doğan bebeklerin doğum ağırlıkları,
- Belirli bir diş hekimi kliniğinde görülen hastalarda protez gereksinimi veya dolgu gereksinimi

21

BİYOİSTATİSTİK

- Belirli bir şehirde HIV pozitif kadınla cinsel ilişkiye giren erkeklerde HIV pozitif olanların sayısı
- 24 saatlik idrarda kreatinin mg/lt miktarı
- 24 saat içinde meydana gelen trafik kazalarında ölümlü vaka sayısı gibi

22

BİYOİSTATİSTİK

- Biyoistatistiksel yöntemler genel olarak üç ana aktivite üzerine yoğunlaşır.

23

BİYOİSTATİSTİK

1. **En uygun olabilir yolla verilerin toplanmasında**

- Araştırmaya alınacak objelerin en uygun ve yapılabilir bir yöntemle seçiminin sağlanması. Örnek olarak, sağlıkla ilgili bir anket çalışmasında kişilerin uygun ve bir örnekleme yöntemi kullanılarak seçilmeleri
- Uygun veri toplama yöntemlerinin seçimi, örneğin verilerin sorgulama yöntemi ile belirlenmesi. Gebe kadınlarda anemi hakkında bilgi toplayabilmek için sorgu yöntemi geliştirme

24

BİYOİSTATİSTİK

- Toplumda hastalık taramalarında , epidemiyolojik araştırmalarda, klinik ve laboratuvar araştırmalarında veri toplama yöntemlerini organize etmek.
- Terminoloji standardizasayonu, doğum ve ölümlerle ilgili veri toplamada çalışacak elemanların eğitimi gibi

25

BİYOİSTATİSTİK

2. **Verilerin tanımlanması ve özetlenmesi**

- Tablo ve grafiklerle verilerin özetlenmesi ve sunumunun yapılması
- Verilerin ölçütlerle özetlenmesi

26

BİYOİSTATİSTİK

3. Verilerin analizi ve sonuçların çıkarılması

- Belirli bir karakterde veya faktörde; kişiden kişiye, zamandan zamana, aynı kişide ölçümde kullanılan enstrümandan enstrümana veya gözlemciden gözlemciye göre değeri değiştiğinde varyasyonlar meydana gelir.

27

BİYOİSTATİSTİK

3. Verilerin analizi ve sonuçların çıkarılması

- Sağlıkla ilgili tüm faktörler bunun gibi değerlerinde yaygınlık veya varyasyon gösterebilir. Örneğin kan basıncında kişiden kişiye, sabahtan öğleden sonraya, heyecan ve sonrasında, oturur pozisyondaki ayaktakinden farklı değerler alır.
- Bu varyasyonlar verilerin özetlenmesinde uygun yöntemlerin kullanılmasını gerektirir.

28

TEMEL TANIMLAR

Popülasyon (evren)

- Araştırıcının ilgilendiği karakterlere sahip en geniş topluluk olarak tanımlanır.
- 2003-2004 öğretim yılında Tıp fakültesi 1.sınıf öğrencileri gibi.
- Türkiyede çalışan hekimler,
- Sigara içen hekimler gibi
- Her birisi ayrı populasyonları tanımlar.

29

TEMEL TANIMLAR

Örnek

Popülasyonu temsil eden ve popülasyonun bir alt grubunu içerir.

- Tıp fakültesi 1. sınıf öğrencileri örneği,
- Bayan tıp fakültesi öğrencileri gibi.

30

TEMEL TANIMLAR

- **İnsidans:**Takip araştırmalarından elde edilen riskin ölçütü olan bir hızdır. Payda da sadece belirli zaman aralığında ortaya çıkan yeni hastaların sayısını içerir.
- **Prevelans:**Kesitsel araştırmalardan elde edilen durumun büyüklüğünü ölçmede kullanılan bir hızdır. Hastalık riskinin ölçütü değildir. İnsidansa eşdeğer değildir.
- Prevalans ≈ İnsidans x hastalığın süresi
- **Hız(Rate):** Bir olayın oluş olasılığı olarak tanımlanır

31

TEMEL TANIMLAR

- **Oran(Ratio):** İki farklı karakterin birbirilerine göre relatif sıklıklarıdır. Ör: Hekim başına düşen hasta sayısı bir orandır.
- **Oran(Proportion):**Bir karakteristiğin yüzde olarak sıklığıdır. Ör: Sigaraya bağlı ortaya çıkan kanserlerin toplam kanserler içindeki oranı.

32

TEMEL TANIMLAR

Risk faktörleri:

- Hastalığın ortaya çıkışını artan bir olasılıkla sağlayan faktör veya faktörlere risk faktörleri adı verilir

33

TEMEL TANIMLAR

Neden-sonuç ilişkisi

Hastalığın ortaya çıkışı doğrudan bu faktörün varlığına bağlı olması durumudur.

- Doğrudan ilişki
- Doğrudan olmayan ilişki
- Yanlış ilişki

34

TEMEL TANIMLAR

Epidemiyoloji

Toplumda görülen hastalıkların kişi, yer ve zaman özelliklerine göre dağılımlarını ve nedenlerini inceleyen bir bilim dalıdır. Yöntem olarak Biyoistatistik kullanır.

35

TEMEL TANIMLAR

Epidemiyoloji

- Kronik hastalıklar epidemiyolojisi
- Enfeksiyon hastalıkları epidemiyolojisi
- Kanser epidemiyolojisi
- Meslek Hastalıkları epidemiyolojisi

36

TEMEL TANIMLAR

Sağlık istatistikleri

- Toplumun sağlık düzeyini göstermede kullanılan özet ölçütlerdir.

İstatistik

- Günlük yaşamımızda kullanıldığı gibi, araştırmalarda örnek verilerinin özetlenmesinde kullanılan özet ölçütler olup hepsi birer değişkendirler.

Parametre

- Bilinen veya bilinmeyen sabitlerdir. Popülasyonun tüm bireylerinde ölçülen verileri özetleyen ölçütlerdir. Nufüs sayımları sonuçlarını özetleyen ölçütler gibi.

37

TEMEL TANIMLAR

Sağlık istatistikleri

- ✓ Beklenen yaşam uzunluğu,
- ✓ Belirli hastalıklar nedeni ile ölüm hızları,
- ✓ Hastalanma hızları,
- ✓ Anne ölüm hızları,
- ✓ Bebek ölüm hızları,
- ✓ Doğurganlık hızları,
- ✓ Kanser insidansları,
- ✓ Okul çağı çocukları diş çürüklüğü prevalansı,
- ✓ Ağız kanserleri insidansı gibi

38

TEMEL TANIMLAR

Hastalık toplumda;

—

- Endemik: Toplumda hastalığın düzenli ve düşük seviyelerde bulunması
- Epidemik: Toplumda hastalığın beklenmiyen yüksek seviyelerde ve geniş alanlarda görülme durumu
- Pandemik: Çok daha geniş alanlarda toplumu etkileyen epidemilerdir.

39

TEMEL TANIMLAR

Değişken;

- Yapılan ölçümlerde kişilerde değişik değerler gözlenmesi durumudur. Yaş, cinsiyet, doğum sırası, vücut sıcaklığı, yaşam uzunluğu, hastalık şiddeti gibi biyolojik karakteristiklerin hepsi birer değişkendirler.

40

TEMEL TANIMLAR

Rastgele değişken;

Yapılan ölçümlerde kişilerin belirli bir olasılıkla değişik değerler alması durumunda değişken rastgele değişken adını alır

41

TEMEL TANIMLAR

Rastgele değişken

- Rastgele değişken olabilmesi için ölçümlerin yapıldığı deneklerin popülasyonu temsil edecek şekilde rastgele alınması gerekir. Rastgele alınmayan örneklerden elde edilen değişkenler rastgele değişken yerine değişken olarak alınabilirler.

42

TEMEL TANIMLAR

Rastgele değişken

- Araştırmalarda ölçülen değişkenlerin rastgele değişken olması durumunda bu verilerden elde edilen sonuçlar genelleştirilebilir.
- Yani popülasyona atfedilen kararlar ancak rastgele değişkene ait veriler kullanılması durumunda uygun olur.

43

TEMEL TANIMLAR

Değişkenleri:

- Kantitatif, ölçülebilir
- Kalitatif, ölçülemeyen

44

TEMEL TANIMLAR

Değişkenleri:

- Sürekli

  (Aldığı değerler arasında boşluk bulunmayan) cm olarak boy uzunluğu, kg olarak ağırlık gibi

- Kesikli değişkenler

  (Aldığı değerler arasında boşluk bulunan) olmak üzere sınıflandırmak olasıdır. Bir mahalledeki hane halkı sayısı, ailedeki çocuk sayısı, belirli zaman diliminde acil servise başvuran hasta sayısı gibi.

45

TEMEL TANIMLAR

Değişkenleri:

ölçüm skalalarına göre de dört alt gruba ayırabiliriz

46

TEMEL TANIMLAR

Nominal (İsimsel) skala

- Değişkenlerin aldığı değerlerin sadece isim vererek ölçülebilmesi.
  - ✓ Cinsiyetin erkek veya bayan olarak,
  - ✓ ameliyat sonucunun başarılı veya başarısız olduğu gibi.
- Burada bu değişkenin aldığı değerlerle bir aritmetiksel işlem yapma olası değildir. Farklılığın dereceside bilinmemektedir.

47

TEMEL TANIMLAR

Dizinsel (Ordinal)skala

- Değişkenlerin aldığı değerlerin dizinlenmiş değerleri içermesidir:
  - ✓ Sosyoekonomik durumun fakir, orta halli, zengin ve çok zengin olarak ölçülmesi gibi.
  - ✓ Sınıf geçme notunun zayıf, orta, iyi, pekiyi olması gibi.
- Alınan değerler arasında farklılığın olduğu bilinmesine karşın farklılığın derecesi bilinmemektedir.

48

TEMEL TANIMLAR

Aralık (interval skala)

- Mutlak bir başlangıç noktası olmamasına rağmen aldığı değerler arasındaki farklılık ve farklılığın derecesi saptanabilmekte ve her türlü aritmetiksel işlem gerçekleştirilebilmektedir.
- Sıcaklık ölçümlerinde olduğu gibi. Sıcaklık bir çok değişik birimde Fahrenheight, santigrat derece veya Celsius derecede ölçüldüğü gibi. Her birisinde 0 derece farklı sıcaklıkları göstermektedir.

49

TEMEL TANIMLAR

Oransal skala:

- Aldığı değerler arasındaki farklılık ve farklılığın derecesi belirlenebilmekte ve her türlü işlem yapma olası ve
- mutlak bir başlangıc (sıfır) değerine sahiptir.
- O nedenle en gelişmiş ölçüm çeşididir.

50

TEMEL TANIMLAR

Değişkenlerin elde edilişinde kullanılan ölçüm skalaları;

- İçerdikleri bilgi seviyeleri yönünden karşılaştırıldıklarında metinde verilen sırayı takip eder.
- sonuçta bu verilerin özetlenmesi ve analizlerinde ne tür istatistiksel değerlendirme yöntemlerinin kullanılacağını belirler.
- Bu nedenle her araştırıcı gerekli olan verilerin özelliklerini ve hangi analiz yöntemleri kullanacağını araştırmaya başlamadan saptamalıdır.
-

51

TEMEL TANIMLAR

- Ölçüm skalaları;
- Verilerin özellikleri araştırma sonuçlarının güvenilirliğini ve doğruluğunu doğrudan etkilemektedir.
- Sonuçta araştırma maliyetlerini de etkileyen en önemli faktörlerden birisi olmaktadır.

52

Veriler için ölçüm araçlarını;

A- **Mekanik araçlarla ölçümler**; Uzunluk, yükseklik, ve derinlik ölçen araçlar,Isı nem ve basınç ölçen aletler,Yoğunluk ölçen aletler,Hız ve zaman ölçen aletler,Miktar ölçen aletler örneğinde olduğu gibi

B-Sadece gözlemlenerek yapılan ölçümler:

- ✓ Bilgi, tutum ve davranışın ölçülmesi,
- ✓ Fizik muayene sonuçları gibi ölçümler,
- ✓ gözlem, hasta öyküsünün alınması
- ✓dalak veya karaciğer büyümesinin ölçülmesi gibi

C-Hem insan hem de mekanik araçlarla ölçüm yapılması

- ✓ Röntgen filmlerinin okunmasında olduğu gibi

53

Ölçüm araçlarının sahip olması arzu edilen iki özelliği;

1. Geçerlilik ve
2. güvenilirliğidir.

54

Güvenilir araç;

Aynı şartlar altında birden fazla uygulandığında aynı sonuçları vermelidir.

Güvenilirlik aşağıdaki durumlarda doğrudan etkilenebilir:

- Ölçüm aracının kendisinden kaynaklanan varyasyon sonucu aynı sonuçlar elde edilemeyebilir. Bir tartı cihazında değişen sıfır başlangıç noktası gibi.
- Ölçümü yapılan maddeden kaynaklanan değişimler sonucu yine aynı sonuçlar elde edilemeyebilir

55

GÜVENİLİRLİK AŞAĞIDAKİ DURUMLARDA DOĞRUDAN ETKİLENEBİLİR

- Öykü almada hastanın sorulara değişik yanıtlar vermesi gibi, hastanın soruyu nasıl anladığına göre yanıtlaması.
- Ölçümü yapan ölçümü tekrarlamasından kaynaklanan hata sonucu (gözlemciler içi)
- Aynı cihaz ile birden fazla ölçüm yapılması sonucu değişik sonuçların elde edilmesi, Örneğin tekrarlanan kan basıncı ölçümlerinde farklı sonuçların elde edilmesi
- • Farklı ölçüm yapanlardan kaynaklanan (Gözlemciler arası= interobserver) Farklı kişilerin kan basıncı ölçümler veya film okuma sonuçları gibi

56

GEÇERLİLİK

- Bir ölçümün geçerli olabilmesi için ölçmesi farz edilen durumun göstergeci olmalıdır.
- Vücut ateşi ölçümü, endemik seviyelerde görülen sıtma için geçerli bir göstergeç olmayabilir.
- Doğrudan görüşme ile elde edilen cevaplar kendiliğinden yapılan düşükler için tam bir göstergeç olmayabilir
- Çocuk sahibi olmama durumu infertiletinin göstergeci olmayabilir

57

GEÇERLİLİK

- iki ayrı elementi yani **duyarlılık** ve **seçicilik** özelliği ile değerlendirilir.
- Kanser konusunda en duyarlı aracın patolojik muayene olduğunu biliyoruz.
- Saha araştırmalarında, taramalarda kolay ve çabuk değerler veren yöntemlerin duyarlılık ve seçicilik yetenekleri patolojik sonuçlarla karşılaştırılmalıdır.
- **Duyarlılık**; konsantrasyondaki küçük değişiklikler test sonucu elde edilen ölçümlerde geniş farklılıklara neden oluyorsa bu test duyarlıdır denilir.
- **Seçicilik ise;** bilinen bir karakterin varlığında cevap veren ve diğer karakterlerin varlığında yanıt vermeme durumu o testin, yöntemin, ölçüm aracının seçiciliği olarak tanımlanır

58

•Testler ve ölçüm araçlarının geçerliliğini ölçmede değerlendirmede kullanılan **duyarlılık** ve **seçicilik** (özgünlük) ölçütleri aynı zamanda tanı ve tarama testlerinin güvenilirliklerini belirlemede kullanılırlar.

•Epidemiyolojide;

•**Duyarlılık**, hastalığın varlığında test'te pozitif değer verenlerin oranı ve

•**seçiciliği** de hasta olmayanlarda teste negatif değer verenlerin oranı olarak tanımlarız.

• Bu amaçla kullanılan diğer ölçütler yanlış negatif veya yanlış pozitif oranları gibi

59

•Bunu tablo ile özetlemek gerekirse

Test sonuçları	Gerçek durum +	Gerçek durum -	Toplam
+	a	b	a+b
-	c	d	c+d
Toplam	a+c	b+d	N

Duyarlılık=a/(a+c), seçicilik=d/(b+d)

Pozitif prediktif değeri= a/ (a+b)

Negatif prediktif değeri= d/ (c+d)

60

Tıp ve sağlık bilimlerinde veri kaynağı olarak iki kaynağa başvurulmaktadır

1. Regular veya rutin sistemle toplanan veri kaynağı

Doğumların, ölümlerin evlenme ve boşanma ile iligili ulusal hayati istatistiklerin kayıt sistemi

- Bildirimi zorunlu hastalıkların kayıtları.

Kanser ve son yıllarda SARS bildirimi zorunlu hastalılar listesine ilave edilmişlerdir.

Her ülkede WHO önerisi ile kabul edilen ve her yıl değerlendirilen bir hastalıklar listesi bulunmaktadır. Böylece ülkeleri bu hastalıklar yönünden karşılaştırmak mümkün olabilmektedir

61

.

Kanser kayıt sistemlerinde olduğu gibi bazı özel hastalık kayıt sistemleri de geliştirilmiştir (Juvenil diyabet, özel genetik hastalıklar kayıt sistemleri vb gibi).

- 1., 2. ve 3. Basamak sağlık hizmeti veren kurumlar tarafından (Hastaneler, Sağlık merkezleri, Sağlık ocakları, Dispanserler, Ana çocuk Sağlığı merkezleri gibi) hasta ile ilgili tutulan kayıtlar.

- Bu kayıtların varlığı bir avantaj olmasına rağmen aynı zamanda doğruluğu ve güvenilirliğinde eksiklikler ve farklılıklar yani heterojen olmasının yarattığı sorunlarda bulunmaktadır

62

2. Özel Araştırmalardan elde edilen Veri kaynakları

- Verilerin hangi popülasyonlardan toplanmasının belirlenmesi
- Verilerin popülasyonun tamamındanmı yoksa bir alt grubundanmı toplanacağının belirlenmesi
- Bu popülasyondan alınan kaç kişide bu verilerin toplanacağının belirlenmesi (Örnek büyüklüğünün belirlenmesi)
- Bu çalışmaya alınacak kişilerin nasıl seçileceğinin belirlenmesi
- Veri kaydı için gerekli enstrümanların tasarımının yapılması (Anket formları)
- Veri toplamada görev alacak personelin eğitimi
- Veri toplama şeklinin (Yüz yüze görüşme, telefon veya kişilerin formları kendilerinin doldurmaları)
- Seçilmiş bireylerin tespiti ve veri toplama işlemi

63

VERİLERİN ÖZETLENMESİ

- Grafik ve Tablolarla (Şekillerle)
- Ölçütlerle
- Sıklık dağılımları-Formüllerle

64

VERİ ORGANİZASYONU VE ÖZETLENMESİ

- Geniş veri kitlelerinden gerekli veya faydalı bilgilerin görülmesi açık olmayabilir.
- Toplanmış olan bu veriler öyle bir şekilde organize edilmelilerki içerdikleri bilgiler ve bu bilgilerdeki varyasyon kalıpları ortaya çıkarılabilsin.
- Bunlar için önce verilerin özelliklerinin bilinmesi gerekmektedir

Verilerin özetlenmesinde

A- **Tablo ve grafikler**

Sürekli değişkenler için

•Dizinlenmiş sıralar • Histogram • Poligon • Frekans Tabloları

kesikli değişkenler için

• Kesişim tabloları • Bar grafiği (çizgi grafiği) • Pie Grafiği (çember grafiği)

65

VERİLERİN ŞEKİLLERLE ÖZETLENMESİ

- Sadece kesikli değişkenler için; Bar ve daire grafiği
- Hem kesikli hem de sürekli değişkenler içinFrekans tabloları
- Sürekli değişkenler için Histogram ve poligon
- Sürekli değişkenler için Box Plot
- Sürekli değişkenler için Stem and leaf (sap ve yaprak plot' ı
- İki sürekli değişken arasındaki ilişkiyi göstermede Scattergram

66

VERİLERİN ŞEKİLLERLE ÖZETLENMESİ

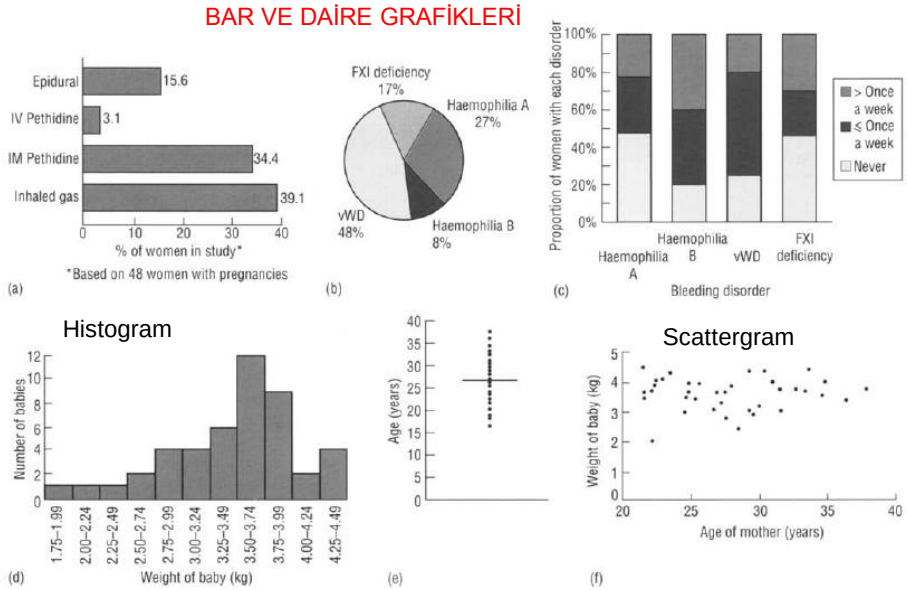
- Hem kesikli hem de sürekli değişkenler için

Frekans tabloları

	Sayı	%	Geçerli %	Kumulative %
Komplikasyon				
Yok	82	92,1	92,1	92,1
DVT	3	3,4	3,4	95,5
Pl.hem	4	4,5	4,5	100,0
Toplam	89	100,0	100,0	

VERİLERİN ŞEKİLLERLE ÖZETLENMESI

BAR VE DAİRE GRAFİKLERİ



VERİLERİN ŞEKİLLERLE ÖZETLENMESİ

Graph and Table

Sürekli veriler için Sap ve Yaprak(Stem and leaf plot)

Stem Leaf	#	Boxplot
9 0	1	0
8 1	1	|
7 07	2	|
6 0023479	7	+-----+
5 234788	6	*-----*
4 14677889	8	+-----+
3 1	1	|
----+----+----+----+		

Multiply Stem.Leaf by 10**+ 1

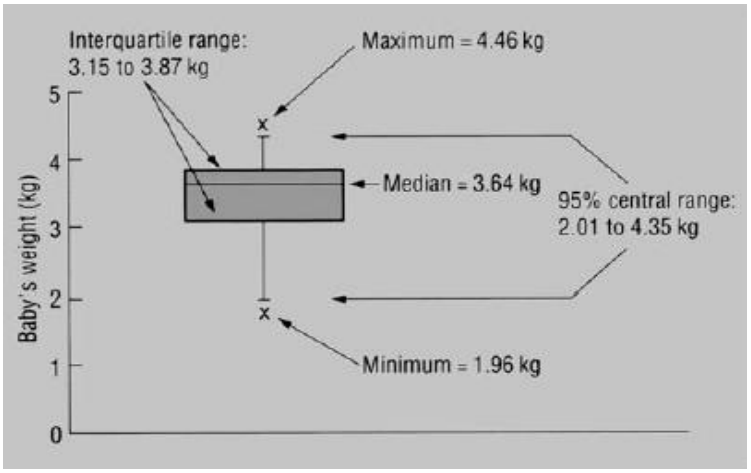
Quantiles (Percentiles)

100% Max	90	99%	90
75% Q3	64	95%	81
50% Med	57.5	90%	77
25% Q1	48	10%	44
0% Min	31	5%	41
		1%	31
Range	59		
Q3 - Q1	16		
Mode	47		

© Elsevier. Jekel et al: Epidemiology, Biostatistics and Preventive Medicine 3e - www.studentconsult.com

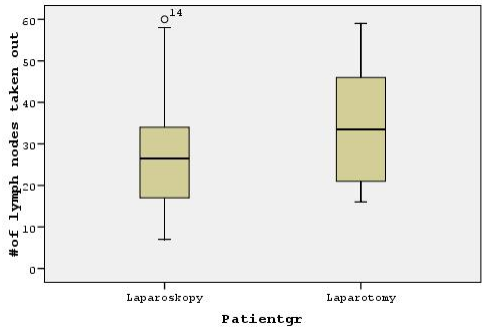
VERİLERİN ŞEKİLLERLE ÖZETLENMESI

BOX PLOT



VERİLERİN ŞEKİLLERLE ÖZETLENMESI

BOX PLOT



71

VERİLERİN ÖLÇÜTLERLE ÖZETLENMESİ

a) Eğilim Ölçütleri

Ortalamalar(μ):Aritmetik, Geometrik,Trimmed Medyan, Mod, persentil, Quartil,Oran,orantı (p), hızlar, mortalite veya morbidite hızları, Insidans, prevalans hızları

a) Yayılım Ölçütleri

Varyans, Standard Sapma, aralık, varyasyon katsayısı ve interquartil aralığı

72

B-Ölçütlerle

B1-Eğilim ölçütleri

sürekli değişkenlerde:

- Aritmetik ortalama
- geometrik ortalama
- Ağırlıklı ortalama
- Medyan (Ortanca)
- Mode (tepe değeri)
- Quartiller
- **P**ersentiller

Kesikli değişkenlerde:

- Oranlar (Yüzdeler)
- Hızlar (Doğurganlık hızı, bebek Ölüm hızı, kanser insidans hızı gibi)
- Orantılar (Bir çalışmada erkek kadın oranı1:1, toplumda arkek ve kadında akciğer kanseri orantısı 4:1, erkek ve kadında meme kanseri insidansı orantısı 1: 20)

- Mode (tepe değeri)

73

VERİLERİN ÖZETLENMESİ-EĞİLİM ÖLÇÜTLERİ

- 18 rastgele seçilmiş hastada Şok skorları aşağıdaki gibi ölçülmüştür:
- 0.61,0.56,0.52,0.33,0.45,0.74,0.73,0.92,0.42,0.63,0.55,0.50,0.75,0.82,1.30,1.29,0.85,0.44
- Ölçümlerin aritmetik ortalaması:. Rastgele değişkeni X olarak tanımlarsak,
- Örneğe ait aritmetik ortalama yı, $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ $\bar{X} = \frac{0.61+0.56+...+0.44}{18} = 0.689$
- Populasyona ait aritmetik ortalama: $\mu = 1/N \sum_{1}^{N} x_i$
- Ortalamalar uç değerlere hassas ölçütlerdir.
- Medyan değeri ise ortada bulunan hastanın değeri olup, bu değer örneğe giren kişileri eşit olarak ikiye bölen değerdir. Yani hastaların yarısı bu değerden büyük diğer yarısı da küçük değerler alır.

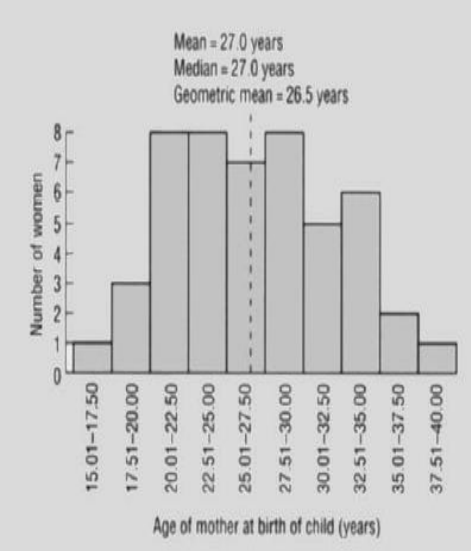
  0.33, 0.42, 0.44, 0.45 ,0.50, 0.52, 0.55, 0.56, 0.61, 0.63, 0.73, 0.74, 0.75, 0.82, 0.85, 0.92, 1.29, 1.30- medyan değeri ,9th ve 10th hastaların ortalama değerleridir. yani 0.62 dir.

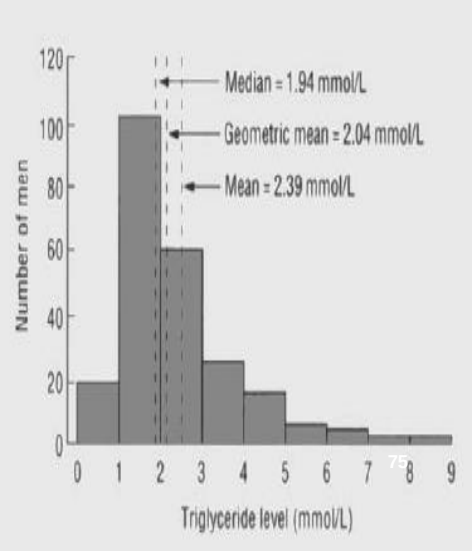
Medyan değeri uç değerlerin etkisinden ortalamalar kadar etkilenmez.

Mod değeri ise grupta en sık alınan değerdir. Bu veri setinde ölçümler hep aynı sıklıkla değer almakta olduğundan tek bir mod değerinden bahsedilemez. Eğer veri seti iki mod içeriyorsa bu veri seti bimodal dır denilir.

74

Verilerin Özetlenmesi (Merkezi Eğilim)





Merkezi Eğilim ölçütleri: Olumlu ve olumsuz yanları

Ölçüt	Olumlu	Olumsuz
Ortalama	•Bütün verileri kullanır •Cebirsel Tanımlıdır böylece matemetiksel işleme uygundur •Örneklem dağılımı belidir.	•Uç derlerden etkilenir •Dağılımın çarpık olmasından etkilenir
Medyan (Ortanca)	•Uç derlerden etkilenmez •Dağılımın çarpıklığından etkilenir	•Bilginin büyük bir kısmını göz ardı eder •Cebirsel Tanımlı değildir •Örneklem dağılımı komplikedir.
Mod	•Kategorik veride kolayca belirlenir	•Bilginin büyük bir kısmını göz ardı eder •Cebirsel Tanımlı değildir •Örneklem dağılımı komplikedir
Geometrik Ortalama	Geri transformasyonundan önce ortalama ile aynı özellikere sahiptir Sağa çarpık veri için uygundur	Sadece Log transformasyonu sonucunda simetrik dağılım veriyorsa uygundur
Ağırlıklı Ortalama	Genel olarak ortalamanın olumlu yanlarına sahip Her gözlem ağırlığı kadar temsil edilir	Ağırlığı biliniyor olmalı yada hesaplanabilmeli

VERİLERİN ÖLÇÜTLERLE ÖZETLENMESİ

YAYILIM ölçütleri

- Aralık (Range)
- İnterQuartil aralığı $(Q_3 - Q_1)/2$
- Varyans
- Standart Sapma
- Varyasyon katsayısı

77

VERİLERİN BİLGİSAYAR ORTAMINDA ANALİZ İÇİN HAZIRLANMASI

HasNo	yas	sex	nobett	episend	nobettp1	nobettp2	etiyoloji
1	66	2	3	1	1	1	2
2	6	2	5	2	2	2	1
3	39	1	5	2	2	2	1
4	108	2	5	2	1	2	2
5	102	1	2	1	1	1	2
6	24	2	2	2	1	1	2
7	66	1	5	2	1	2	2
8	72	1	1	1	1	1	2
9	84	2	2	1	1	1	2
10	60	2	2	2	2	1	1
11	10	2	5	2	1	2	2

78

VERİLERİN ÖLÇÜTLERLE ÖZETLENMESİ

B. Yayılım ölçütleri;

Bir örnekte;

- Varyans, örnek varyansı

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

=0.071

S ise örnek standard sapması olarak adlandırılır ve

S^2 = 0.071 in karekökü olaınarak bulunur (S=0.27)

Popülasyonda ise;

- Populasyon varyansı

$$\sigma^2 = 1/N\sum_{i=1}^{N}(x_i - \mu)^2$$

σ ise populasyon standard sapması olarak tanımlanır populasyon varyansı ,σ^2 nin karekökü alınarak bulunur.

79

VERİLERİN ÖLÇÜTLERLE ÖZETLENMESİ

Yayılım ölçütü

Aralık, Veri setinde alınan en yüksek değerin (maximum) en düşük değerden (Minimum) çıkarılması ile elde edilir.

- İnterquartil aralığı

3. quartil değeri ile 1. quartil değeri farkıdır. $(Q_3 - Q1)/2$

Varyasyon katsayısı (Coefficient of Variation,CV),

Veri setinde ortalama ve varyans ilişkili veya biribirinin fonksiyonu ise yayılım için uygun yöntem varyasyon katsayısıdır. Standard sapmanın ortalmaya göre yüzdesi olarak tanımlanır.

CV= (σ/μ)*100% populasyona ait CV

CV= örneğe ait $(\frac{S}{\bar{X}})*100\%$

- İki ayrı ölçüm birim ile elde edilmiş rastgele değişkenlerde yangınlık karşılaştırılmasında en uygun yöntem VC ölçütüdür.

80

YAYILIM ÖLÇÜTLERİ: OLUMLU VE OLUMSUZ YANLARI

Ölçüt	Olumlu	Olumsuz
Aralık (Range)	•Kolayca elde edilir	•Sadece iki gözlem kullanıyor •Uç derlerden etkilenir •Örneklem büyüklüğü artıkca büyüme eğilimi gösterir
Quartillere dayalı Aralık(Range)	•Uç derlerden etkilenmez •Örneklem büyüklüğüden etkilenmez •çarpık veri için uygundur	•Hesaplanması sıkıntılı •Küçük örneklemler için hesaplanamaz •Sadece iki gözlem kullanır •Algebraic? Tanımlı değildir
Varyans	•Her gözlem kullanılır •Algebraic? Tanımlıdır	•Ölçüm birimi ham veridekinin karesi şeklindedir. •Uç derlerden etkilenir •Çarpık dağılımlar için uygun değildir
Standard Sapma	Variyansınsa benzer şekilde Ölçüm birimi ham veriyle aynıdır Anlaşılması kolay	•Uç derlerden etkilenir •Çarpık dağılımlar için uygun değildir

81

VERİLERİN FORMÜLLERLE ÖZETLENMESİ

Sıklık veya olasılık dağılımları

- Bir rastgele değişkenin aldığı değerlerle bunların oluş olasılıkları arasındaki ilişki olasılık değılımı ile özetlenir.
- Kesikli rastgele değişkenler için olasılık dağılımları
- Sürekli rastgele değişkenler için olasılık dağılımları
- Rastgele bir değişkenin olasılık dağılımının bilinmesi, klinisyen ve araştırıcılara veri setini özetlemede ve tanımlamada güçlü bir olanak sunduğu gibi örnek verilerine dayalı olarak geldiği popülasyonla ilgili sonuca ulaşmasını sağlar.

82

VERİLERİN FORMÜLLERLE ÖZETLENMESİ

- Kesikli rastgele değişkenler için olasılık dağılımları

1. Dikotom-Binary (var-yok, hasta-hasta değil, infekte-noninfekte, gibi) veriler için Binom dağılımının iki parametresi vardır. n (örnek büyüklüğü) ve p (oran)

- Formüla ile gösterilebilir
- Tablo ile gösterilebilir
- Grafik ile gösterilebilir
- X~ B(n,p) ve P(X=x | n,p)= $\binom{n}{x} p^X (1-p)^{n-x}, x = 1,2,\ldots,n$

- Binom olasılıkları yukarıdaki formülle hesaplanabildiği gibi kitapların sonunda değişen p ve n değerleri için liste halinde Tablolarda verilmiştir. (Rosner,B, 2000)

83

BİNOM DAĞILIMI

- Araştırmada, her iki ebeveynin kronik bronşit olduğu evlerde yaşayan her 20 çocuktan 3 ünün kronik bronşit olduğu saptanmmış ve ülke kronik bronşit insidansı, %5 ile karşılaştırılmıştır. Bu farklılık gerçek farklılıkmı yoksa şansa bağlanabilen bir farklılıkmı?
- Bilhassa, herhangi bir hanede yaşayan bebeklerde kronik bronşit geçirme olasılığı %5 (p=0.05) olması durumunda 20 haneden en az 3 ünde yaşayan çocukların kronik bronşet geçirmeleri ne kadar olasıdır?.

84

BİNOM DAĞILIMI

- Bebekler arasındaki bu hastalık oranının 5% olduğunu varsayalım. Bu varsayımla, bebeklerin kronik bronsit gelişeceği toplam hane sayısı dağılımı aşağıdaki parametrelerle binom dağılımı olacaktır. n=20, p=0.05
- Bu hanelerin en az 3 ünde kronik bronşitli çocuk görülme olasılığı nedir sorusunun cevaplanmasıdır
- Cevap ise $\Pr(X \geq 3 \mid 20, 0.05) = 1 - \sum_{3}^{20} \binom{20}{3} 0.05^{3} (0.95)^{17}, x = 3, \ldots, 20$

$= 0.754$

Sonuçta , x≥3 büyük olma olasılığı bir hayli yüksek olduğu görüldüğünden bu hastalıkta hanelerde hastalık kümelenmesinin varlığından söz edilebilmesi için daha yüksek sayıda hane içeren çalışmalara gereksinim vardır.

85

POİSSON DAĞILIMI

✓ Kesikli rastgele değişkenler için olasılık dağılımıdır Birim zaman ve birim alanda meydana gelen olayları tanımlayan değişkenlerin olasılık dağılımıdır. Örneğin gece 11 ile 12 arasında acil servise başvuran hasta sayısı, belirli bir bölgede görülen akciğer kanseri nedeni ile ölümler, Gölde veya havuzda veya bir kapta görülen bakteri sayısı gibi.

Olaylar için Poisson dağılımı bir parametre ile tanımlanır- μ (mean and variance value)-Meydan gelen ortalama olay sayısı (ölümler, hastalanmalar, kazalar vb)

- Formüla ile gösterilebilir
- Tablo ile gösterilebilir
- Grafik ile gösterilebilir
- X~ P(μ) ve P(X=x | μ)= $\frac{e^{-\mu}\mu^{x}}{x!}, x = 0,1,2,\ldots.$

86

POİSSON DAĞILIMI

Olay sayısı ile ilgili veriler formüla ile verilen olasılık dağılımı ile tamamı ile tanımlanmıştır.

X~ P(μ) ve P(X=x | μ)= $\frac{e^{-\mu}\mu^{x}}{x!}, x = 0,1,2,....$

87

Uzun yıllar içinde Tifo nedeni ile yıllık ölümler. Tifo nedeni ile her hangi verilen bir günde yeni bir ölüm olasılığı küçük olacağı gibi ayrık zaman dilimlerinde meydana gelen

ölümler de biribirinden bağımsız rastgele değişkenler olacaktır. Bir sene boyunca meydan gelen ölümlerin dağılımı Poisson dağılımı olacaktır.

POİSSON DAĞILIMI

Tifo nedeni ile bir yıl boyunca meydana gelen ölümlerin dağılımları ortalama, μ=4.6, olan Poisson dağılımı olsun.

6 aylık dönemde Tifo nedeni ile meydana gelen ölümlerin dağılımları ne olacaktır?

- X~ P(μ) where P(X=x | μ)= $\frac{e^{-\mu}\mu^{x}}{x!}, x = 0,1,2,....$

88

X in 6 aylık dönemdeki ölümleri temsil eden rastgele değişken olsun.

Ortalama, μ=4.6, t=1 olduğundan . λ=4.6. 6 aylık bir dönem için, λ=4.6, t=0.5 μ=λt olduğundan ,

6 aylık dönemde hiç ölüm gözlememe olasılığı :

P(X=0)= $e^{-2.3}$ =0.1

6 aylık dönemde1 ölüm gözleme olasığı, P(X=1)= $\frac{e^{-2.3}2.3^{1}}{1!}$

=0.23

SÜREKLİ DEĞİŞKENLER İÇİN DAĞILIMLAR
NORMAL DAĞILIM

Sürekli değişken için olasılık dağılımlarından normal dağılım;

- Formula ile
- Table ile
- Grafik ile gösterilebilir

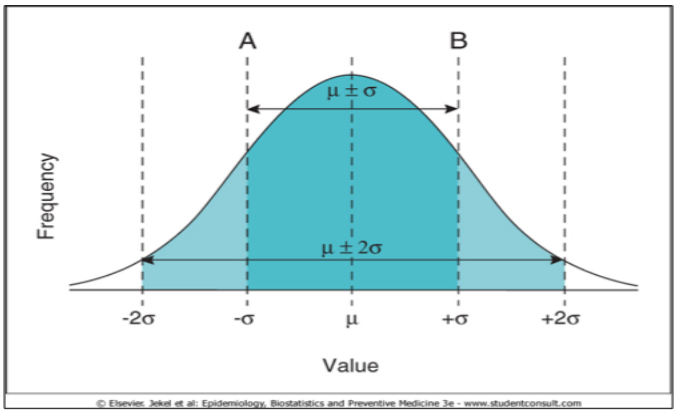
Özellikleri

- Dağılım simetriktir
- Dağılımın ortalama,medyan ve mod değerleri aynıdır.
- Şekli çan eğrisine benzer
  - İki parametre tarafından tanımlanır, μ –ortalama ve standard sapma σ.
  - μ ± σ %68
  - μ ± 2σ aralığı değişkenin değerlerinin %95.4 ünü içerir

89

NORMAL DAĞILIM

- Grafik



90

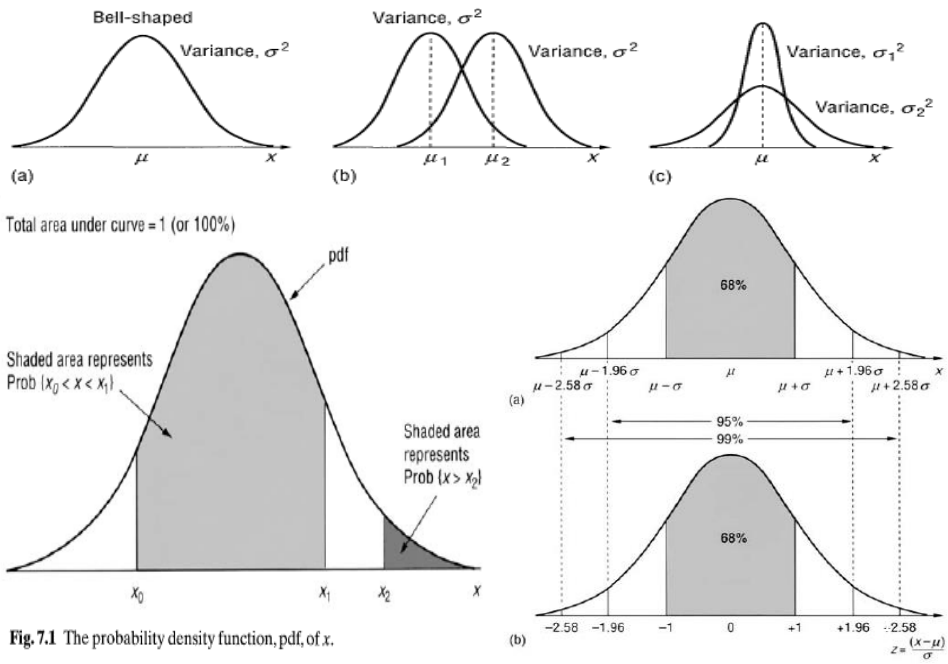


Fig. 7.1 The probability density function, pdf, of x.

Örnekleme ve Örneklem Dağılımı

- Neden örnekleme yapıyoruz
- İyi bir örneklem nasıl olmalı
- Nokta tahmin edicileri
- Ortalamanın örneklem dağılımı
- Standart hatanın (SE) yorumlanması
  - Büyük SE değerleri nasıl yorumlanır
  - Küçük SE değerleri nasıl yorumlanır
- SD veya SE hangisi ne amaçla kullanılır

92

NORMAL DAĞILIM

- Formula ile gösterimi
  - $F(X<x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2} \quad -\infty \le x \le \infty$
  - X normal dağılım gösteren bir rastgele değişkendir.

93

NORMAL DAĞILIM

Standard normal dağılım

- Formula
  - $F(X<z \mid \mu=0, \sigma^2=1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(z)^2/2} \quad -\infty \le z \le \infty$

$$Z = \frac{X-\mu}{\sigma}$$

  - Z rastgele değişken
    - Where $X \sim N(\mu, \sigma^2)$
    - $Z \sim N(0,1)$

Değerleri ve oluş olasılıkları normal dağılımlar tablosu olarak verilmiştir. (Sekel, p-394-95)

94

SUMMARY OF DATA

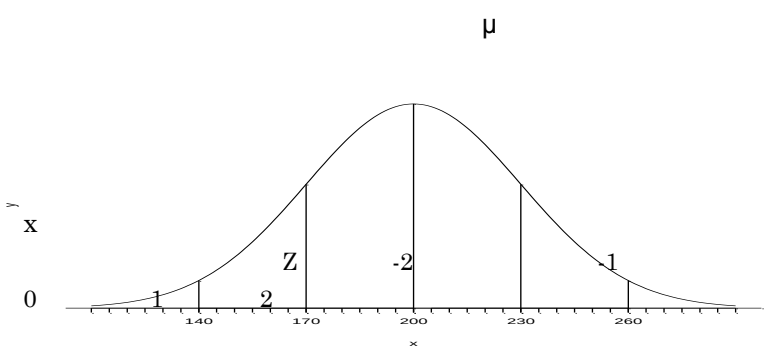
- Standard Normal Table

Table A Areas Under the Normal Curve

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4350	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

95

AĞIRLIK DEĞİŞKENİNİN DAĞILIMI

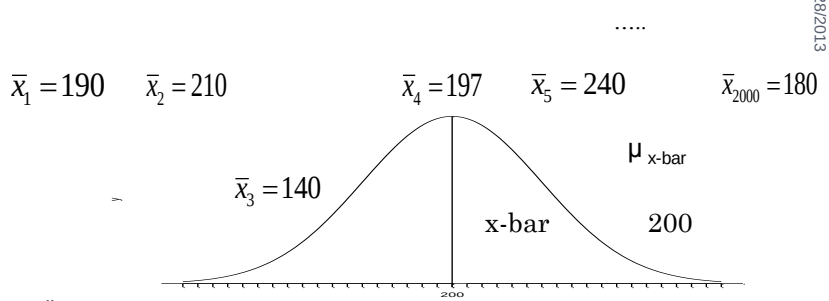


$$z = \frac{x-\mu}{\sigma} = \frac{235-200}{30} = 1.17$$

3/28/2013 96

ORTALAMANIN ÖRNEKLEME DAĞILIMI

- μ = 200 lbs.
- Örnek büyüklüğü= n = 30



Örnek büyüklüğü 30 olan her türlü olası örnek ortalamaları

MERKEZİ LİMİT TEOREMİ

1. POPULASYON DAĞILIMI NORMAL OLSUN VEYA OLMASIN ÖRNEK ORTALMALARININ DAĞILIMI YAKLAŞIK NORMAL DAĞILIMDIR.

Aşağıdaki iki ilave prensipler istatistik teorisinden gelmektedir. Bunlar:

2. Örnek ortalamalarının beklenen değeri yine popülasyon ortalmasına eşittir. $\mu_{\bar{x}} = \mu$

3. Örnek ortalmalarının standard sapması popülasyon standard sapmasının örnek büyüklüğünün kare köküne bölünmesi ile elde edilen değere eşittir. Yani;

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

= ortalamanın standard hatası.

İSTATİSTİKSEL KARAR

- İstatistiksel karar popülasyondan alınan örnek verilerinden elde edilen bilgileri kullanarak geldiği popülasyonla ilgili sonuçlara ulaşmak yöntemidir.
- Sağlık bilimlerinde tahmin etmenin arkasında yatan rasyonel bu alanda çalışanların parametrelere ilgisinin olduğu varsayımına dayanır.
- İlgilenilen parametre ortalama veya popülasyona ait birçok sağlık olabilir. Çocuklar arasında ortalama ağırlık, obez çocukların oranı, obez kadın ve erkeklerin populasyondaki oranlarının bilinmesi belirli hastalıkların önlenmesi için çok önemlidir.

3/28/2013 99

TAHMİN YÖNTEMLERİ

- Verilen örnekten belirtilen parametreler hakkında bilgi sahibi olabilmek için bir tür yöntemlere gereksinim vardır. Bu tür yöntemlere tahmin yöntemleri adı verilir.

✓ Nokta tahmini
- Populasyon parametresini tahmin etmek için kullanılan tek bir değerdir. Örneğin,
- popülasyon ortalaması için örnek ortalaması $\bar{X}$

, σ^2 (popülasyon varyansı) için örnek varyansı, S^2 ,

Populasyon korelasyon katsayısı ρ (rho) için örnek korelasyon katsayısı, r.

✓ Aralık tahmini
- Aralık tahmini, belirli bir güvenilirlikle parametrenin değerlerini içeren alt ve üst sınırları güvenilirlik derecesi ile belirlenmiş bir aralıktır ve bu aralığın bilinmeyen parametreyi içerdiği kabul edilir.

3/28/2013 100

ÖNEMLİLİK TESTLERİ

- Önemlilik testinde amaç, o popülasyondan alınan örneğe ait verileri değerlendirerek klinisyen, araştırıcı veya yöneticiye popülasyon hakkında bir sonuca ulaşabilmesine yardımcı olmaktır
- Bir hipotez, bir veya birden fazla popülasyonlar hakkında basit bir beyan olarak tanımlanır
- Bir araştırma hipotezi araştırma motivasyonu için bir A research hypothesis is the faraziye veya varsayımdır.
- Bir hekim değişik tedavi kombinasyonlarının birinin diğerlerinden yalnız başlarına etkilerinden daha etkili olduklarını hatırlayabilir.

3/28/2013 101

ÖNEMLİLİK TESTLERİ

- Araştırma projeleri araştırıcıların şüphelerinin veya teorilerinin bilimsel çalışmalardan elde edilen bulgularla desteklenmesini arzu etmeleri sonucunda ortaya çıkmaktadır.
- Akıcı olması için bir önemlilik testleri aşamalı yöntem olarak sunulacaktır.

3/28/2013 102

ÖNEMLİLİK TESTLERİ

Bu aşamalar;

1. Veri
2. Varsayımlar
3. Hipotez
4. Test İstatistiği
3. Test istatistiğinin dağılım
4. Karar kuralı
5. Test istatististiğinin hesaplanması
6. İstatistiksel karar
7. Sonuç
8. p-değerleri

3/28/2013 103

ÖNEMLİLİK TESTLERİ

Bu aşamalar;

1. Veriler ve özellikleri; sürekli, kesikli, dağılımları
2. Varsayımlar

   Normal dağılım varsayımı,varyansların eşitliği, alınan örneklerin birbirinden bağımsızlığı
3. Hipotez

   H_0. sıfır hipotezi: farklılık yok

   H_A. Alternatif hipotez; farklılık var veya araştırma hipotezidir. Farklılık var, daha etkindir, daha büyüktür gibi.
4. Test istatistiği

   Verilerden elde edilir ve karar vericiye karar vermede yardımcı olur. Hipotezin kabul veya red edilmesi test istatistiğinin büyüklüğüne göre değişir.

3/28/2013 104

ÖNEMLİLİK TESTLERİ

5. Test istatistiğinin dağılım

Örneklem dağılımı istatistiksel kararın anahtarıdır.

6. Karar kuralı

Karar kuralı örnek verilerinden elde edilen test istatistiğinin değeri eğer red bölgesine düşüyorsa sıfır hipotezini red etmemizi eğer kabul bölgesine düşen bir değeri varsa etmememizi söyler.

Önemlilik seviyesi, 1. tip hata, red bölgesini belirler

7. Test istatistiğinin hesaplanması

Araştırmada elde edilen verilerden hesaplanır.

3/28/2013 105

ÖNEMLİLİK TESTLERİ

8. İstatistiksel karar

Null Hipothezi durumu

Alınan karar	Doğru	yanlış
Ret edememe, H_0	Doğru karar	Tip II hata
Ret etme, H_0	Tip I hata	Doğru karar

9. Sonuç

H_0 ret edilmesi durumunda, H_A hipotezinin doğru olduğu sonucu çıkarılır. H_0 ret edilemediğinde, H_0 hipotezinin doğru olabileceği kabul edilir.

10. p-değeri bize H_0 doğru olduğunda örneğe ait verilerin ne kadar ender olabileceğine ait bir rakamdır. $P < 0.05$ genellikle alternatif hipotezin ve $p > 0.05$ null hipotezin gerçek olabileceğini destekler.

3/28/2013 106

ÖNEMLİLİK TESTLERİ

- Örnek 1:

  H_0: Hastalık ve bir faktöre maruziyet arasında ilişki yok

  H_A: Hastalık ve bir faktöre maruziyet arasında ilişki var

  α, 1. tip hata:

  HASTALIKLA FAKTÖR ARASINDA İLİŞKİ OLMADIĞI HALDE İLİŞKİ VARDIR KARARININ VERİLMESİ

  β, II. TİP HATA:

  HASTALIKLA FAKTÖR ARASINDA İLİŞKİ OLDUĞU HALDE İLİŞKİ YOKTUR KARARININ VERİLMESİ

3/28/2013 107

ÖNEMLİLİK TESTLERİ

- *α*, (Tip I hata): Doğru olan null hipotezini red etme olasılığı
- *β*, (Tip II hata): Yanlış olan null hipotezini red etmeme olasılığı
- Örnek 1:

  H_0: Sanık suçsuz

  H_A: Sanık suçlu

  α (5%):

  sanık suçlu olmadığı halde sanık suçludur kararının verilmesi

  β, (20%):

  Sanık suçlu olduğu halde sanık suçlu değildir kararının verilmesi

3/28/2013 108

ÖNEMLİLİK TESTLERİ

- Example 3:
  H_0: Tedavi etkin değil
  H_A: Tedavi etkin

  α (5%):
  tedavi etkin olmadığı halde tedavi etkindir kararının alınması

  β, (20%): Tedavi etkin olduğu halde tedavi etkindeğildir kararının alınması

ÖNEMLİLİK TESTLERİ VE GÜVENLİK ARALIKLARI

- TEK BİR POPULASYON ORTALAMASI İÇİN ÖNEMLİLİK TESTİ
-
- **Adana da yaşayan erkeklerde ortalama kolesterol seviyelerinin gerçek değerini bilmek istiyoruz.**
  yapılan çalışmalardan;
  $\mu = 200$ mg/dL ve $\sigma = 50$ mg/dL.
- Fakat araştırıcı bu Adana populasyonundaki gerçek ortalama kolesterol değerinin 200 olabileceğine inanmamakta hatta geçmiş yıllarda meydana gelen diyet farklılıkları sonucu daha da yüksek olabileceğine inanmaktadır.

ÖNEMLİLİK TESTLERİ

- İki hipotez tanımlayarak başlıyoruz

Sıfır hipotezi (H_0), inandığımız günümüzdeki bilgileri yansıtır. Bizim örneğimizde, literatür bilgilerinde verilen
$\mu = 200$ mg/dL.

H_0: Populasyon ortalama kolesterol (μ) = 200 mg/dL

Alternatife Hipotez (H_A), araştırıcının populasyonda var olduğuna inandığı değerlerdir.
Bu da $\mu > 200$ mg/dL.
H_A: Populasyon ortalama kolesterol (μ) $\neq$ 200 mg/Dl
H_0: $\mu = 200$
H_A: $\mu \neq 200$

3/28/2013 111

ÖRNEK VERİLERİNİ KULLANARAK ALINDIĞI POPULASYONLA İLGİLİ SONUÇ ÇIKARMAK

- Adana populasyonundan30 erkeğin rasgele alındığı ve örnek kolesterol ortalamasnında (x-bar) cholesterol 300 mg/dL olduğu bulunsun.

$\bar{x}$ = 300 mg/dL
n = 30

- Sorumuz ? İstatistiksel olarak 300 mg/dl 200 mg/dl den farklımıdır? Veya gerçek populasyon ortalamsı 200 değildir veya
- Null hipotezde belirtilen ($\mu = 200$) gerçekte yanlışmıdır?

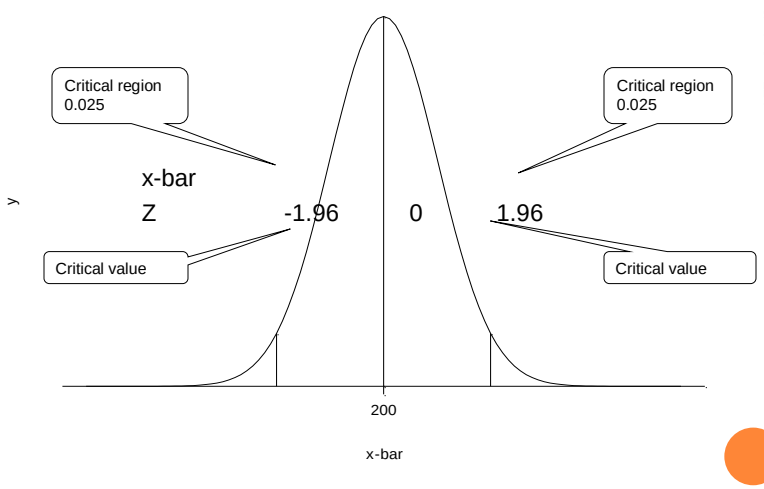
- Gerçek popülasyon ortalaması 200 olduğunda örnek ortalamasının 300 gözlenmesi olasılığı çok düşük. Çok düşük olasılıklı olduğu için popülasyon ortalaması 200 olmayabilir fakat ne olduğu konusunda şüphelerimiz vardır.
- Buna karar verirsek null hipotezi rededeyor ve alternatif hipotezin ($\mu \neq 200$) doğru olduğu sonucuna varıyoruz.

3/28/2013 112

ÖNEMLİLİK TESTİ

- Önce örnek büyüklüğü 30 olan tüm örneklerin ortalamalarının null hipotezi doğru olduğundaki grafiği elde edilir.

TEORIK ÖRNEKLEM DAĞILIMI, N=30



TEST UYGULAMASI

- 1. TİP HATA α = 0.05 OLSUN.

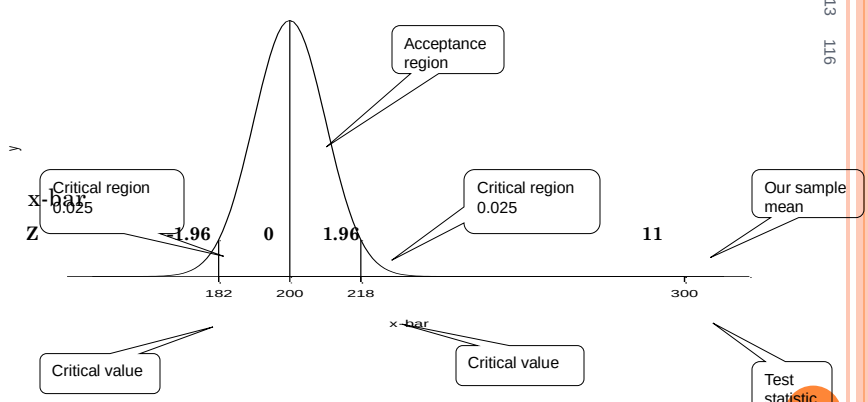
Test istatistiği: $z = \dfrac{\bar{x} - \mu}{\sigma / \sqrt{n}}$ **1 SE = 50 / sqrt(30)**

$$z = \frac{300 - 200}{50/\sqrt{30}} = 10.96 \approx 11$$

Bu da kritical alan içinde red bölgesine düşmektedir.

3/28/2013 115

TEORİK DAĞILIM



3/28/2013 116

SONUÇLAR

- Kritik değerler ± 1.96 örnek ortalamaları ; 182 ve 218 i temsil eder. Bu iki örnek ortalamaları yaklaşık ± 2 standard hatanın etrafında bulunurlar.
- (yani, ±1.96 standard hata).

$\overline{x}$

Böylece , 182(z = -1.96) altına düşen bir örnek ortalaması gözlersek veya 218 (z = 1.96), üzerine düşen ve kritik bölgenin red bölgesinde olacağımızdan null hipotezi red kararı veririz.

- 300, 218den büyük olduğundan veya (or z = 11 > z = 1.96), null hipotezi red edilir ve populasyon ortalamasının $\mu \neq 200$ olduğu sonucuna varılır. (yani, 11 standard hata > ~ 2 standard hatas)

3/28/2013 117

SONUÇLAR

a) Karar z-test istatistiğine bağlıdır.

- p< α- önemlilik seviyesi ise.
  bu olayda 1. tip hata 0.05.

p= Pr(sıfır hipotezi doğru olduğunda eldeki verilerden daha büyük test istatistiği elde etme olasılığıdır..

= Pr ($Z \geq 11 \mid \mu=200$ ($H_{0\ \text{doğru}}$) =0.00000 alternatif hipotez desteklenmektedir.

3/28/2013 118

GÜVENLİK ARALIĞI FORMÜLÜ

- Genel formül:

  %95 Güven aralığı= $\bar{x} \pm (z \times \frac{\sigma}{\sqrt{n}})$ (z değeri kritik z olarak kullanılmıştır

  $= 300 \pm (1.96 \times \frac{50}{\sqrt{30}})$

  = 300 mg/dL ± 18 mg/dL = (282 mg/dL, 318 mg/dL)

- μ = 200 ün bu aralığa düşmediğine dikkat et. Bu da sıfır hipotezinin reddi ile uyumludur. 200 gerçek ortalama kolesterol seviyesi değildir. Gerçek ortalama 282 ve 318 değerleri arasında olması çok muhtemeldir. %95 Güven aralığı testteki hipotezle uyumlu olmalıdır.

  Her bir durum için, her iki yaklaşım, önemlilik testi ve güven aralığı yaklaşımları sıfır hipotezinin ret edileceği sonucuna götürür. Yani gerçek popülasyon ortalaması 200 mg/dL.değildir.

3/28/2013 119

POPULASYON ORTALAMALARI FARKI İÇİN ÖNEMLİLİK TESTİ

Kolesterol seviyelerini azaltmada kullanılan Standard kontrol tedavi ile yeni tedavinin etkinliklerinin karşılaştırıldığı kontrollü randomize bir çalışma da hastalar da 6 ay takip sonunda kolesterol seviyeleri ölçülmüş ve bu iki hasta grubu karşılaştırılmıştır.

Çalışmaya alınan tüm hastalar yüksek kolesterole sahip hastalardır. Hastalar bu iki gruptan birine rastgele atanmışlardır.(Standard tedavi veya yeni tedavi)

Eğer yeni ilaç kolesterol azaltmada kontrol tedaviden iyi değilse, 6 ayın sonundaki ortalama kolesterol seviyelerindeki farklılık sıfır olacaktır.

$\mu_C - \mu_T = 0$

3/28/2013 120

HİPOTEZ

- .Bu sıfır(null) hipotezini temsil eder 6. ayda her iki grupta ki ortalama kolesterol seviyeleri arasında bir farklılığın olmadığının ifadesidir:

  H_0: $\mu_C - \mu_T = 0$

- Alternatif hipotez de:

  H_A: $\mu_C - \mu_T \neq 0$

- Diğer bir deyişle ,6. ayda her iki grupta ölçülen ortalama kolesterol seviyelerinde farklılıktan şüphelenilmektedir.

- Hangi grupta ortalama kolesterol seviyesinin daha yüksek olduğu tanımlanmadığından buradaki test iki yönlü test olarak bilinir.

ÖRNEKLER

- Birisi yeni tedaviyi diğeri ise standart-kontrol tedavisini almış iki hasta grubumuz olacaktır. Sonra da 6. aydaki örnek ortalamaları elde edilecektir.

  KONTROL: $n = 30$
  $S = 40$ mg/dL
  $\bar{x} = 300$ mg/dL

  YENİ TEDAVİ: $n = 35$
  $S = 45$ mg/dL
  $\bar{x} = 220$ mg/dL

- $\alpha = 0.05$ Test

- İki grubun ortalamaları arasındaki farkın örneklem dağılımı çizeriz.

Table B Probability Between $\pm t$ Value (Two-Sided)

Area = 0.20	0.40	0.60	0.80	0.90	0.95	0.98	0.99	0.999
Probability Below t Value (One-Sided)								
df area = 0.60	0.70	0.80	0.90	0.95	0.975	0.99	0.995	0.9995
1 0.3250	0.7270	1.376	3.078	6.3138	12.706	31.821	63.657	636.619
2 0.2885	0.6172	1.061	1.886	2.9200	4.3027	6.965	9.9248	31.598
3 0.2766	0.5840	0.978	1.638	2.3534	3.1825	4.541	5.8409	12.924
4 0.2707	0.5692	0.941	1.533	2.1318	2.7764	3.747	4.6041	8.610
5 0.2672	0.5598	0.920	1.476	2.0150	2.5706	3.365	4.0321	6.869
6 0.2648	0.5536	0.906	1.440	1.9432	2.4469	3.143	3.7074	5.959
7 0.2632	0.5493	0.896	1.415	1.8946	2.3646	2.998	3.4995	5.408
8 0.2619	0.5461	0.889	1.397	1.8595	2.3060	2.896	3.3554	5.041
9 0.2610	0.5436	0.883	1.383	1.8331	2.2622	2.821	3.2498	4.781
10 0.2602	0.5416	0.879	1.372	1.8125	2.2281	2.764	3.1693	4.587
11 0.2596	0.5400	0.876	1.363	1.7939	2.2010	2.718	3.1058	4.437
12 0.2590	0.5387	0.873	1.356	1.7823	2.1788	2.681	3.0545	4.318
13 0.2586	0.5375	0.870	1.350	1.7709	2.1604	2.650	3.0123	4.221
14 0.2582	0.5366	0.868	1.345	1.7613	2.1448	2.624	2.9768	4.140
15 0.2579	0.5358	0.866	1.341	1.7530	2.1315	2.602	2.9467	4.073
16 0.2576	0.5358	0.865	1.337	1.7459	2.1199	2.583	2.9208	4.015
17 0.2574	0.5344	0.863	1.333	1.7396	2.1098	2.567	2.8982	3.965
18 0.2571	0.5338	0.862	1.330	1.7341	2.1009	2.552	2.8784	3.922
19 0.2569	0.5333	0.861	1.328	1.7291	2.0930	2.539	2.8609	3.883
20 0.2567	0.5329	0.860	1.325	1.7247	2.0860	2.528	2.8453	3.850
21 0.2566	0.5325	0.859	1.323	1.7207	2.0796	2.518	2.8314	3.819
22 0.2564	0.5321	0.858	1.321	1.7171	2.0739	2.508	2.8188	3.792
23 0.2563	0.5318	0.858	1.319	1.7139	2.0687	2.500	2.8073	3.767
24 0.2562	0.5315	0.857	1.318	1.7109	2.0639	2.492	2.7969	3.745
25 0.2561	0.5312	0.856	1.316	1.7081	2.0595	2.485	2.7874	3.725
26 0.2560	0.5309	0.856	1.315	1.7056	2.0555	2.479	2.7787	3.707
27 0.2559	0.5307	0.855	1.314	1.7033	2.0518	2.473	2.7707	3.690
28 0.2558	0.5304	0.855	1.313	1.7011	2.0484	2.467	2.7633	3.674
29 0.2557	0.5302	0.854	1.311	1.6991	2.0452	2.462	2.7564	3.659
30 0.2556	0.5300	0.854	1.310	1.6973	2.0423	2.457	2.7500	3.616
35 0.2553	0.5292	0.8521	1.3062	1.6896	2.0301	2.438	2.7239	3.5919
40 0.2550	0.5286	0.8507	1.3031	1.6839	2.0211	2.423	2.7045	3.5511
45 0.2549	0.5281	0.8497	1.3007	1.6794	2.0141	2.412	2.6896	3.5207
50 0.2547	0.5278	0.8489	1.2987	1.6759	2.0086	2.403	2.6778	3.4965
60 0.2545	0.5272	0.8477	1.2959	1.6707	2.0003	2.390	2.6603	3.4606
70 0.2543	0.5268	0.8468	1.2938	1.6669	1.9945	2.381	2.6480	3.4355
80 0.2542	0.5265	0.8462	1.2922	1.6641	1.9901	2.374	2.6388	3.4169
90 0.2541	0.5263	0.8457	1.2910	1.6620	1.9867	2.368	2.6316	3.4022
100 0.2540	0.5261	0.8452	1.2901	1.6602	1.9840	2.364	2.6260	3.3909
120 0.2539	0.5258	0.8446	1.2887	1.6577	1.9799	2.358	2.6175	3.3736
140 0.2538	0.5256	0.8442	1.2876	1.6558	1.9771	2.353	2.6114	3.3615
160 0.2538	0.5255	0.8439	1.2869	1.6545	1.9749	2.350	2.6070	3.3527
180 0.2537	0.5253	0.8436	1.2863	1.6534	1.9733	2.347	2.6035	3.3456
200 0.2537	0.5252	0.8434	1.2858	1.6525	1.9719	2.345	2.6006	3.3400
∞ 0.2533	0.5244	0.8416	1.2816	1.6449	1.9600	2.326	2.5758	3.2905

3/28/2013 123

TEST İSTATİSTİĞİ

- Test istatistiğinin hesaplanması.
- İki tane örnek standard sapma değerleri olduğundan test istatistiği formülünde kullanmak üzere 'pooled Standard sapma' Sp değerini elde etmemiz gerekmektedir. :

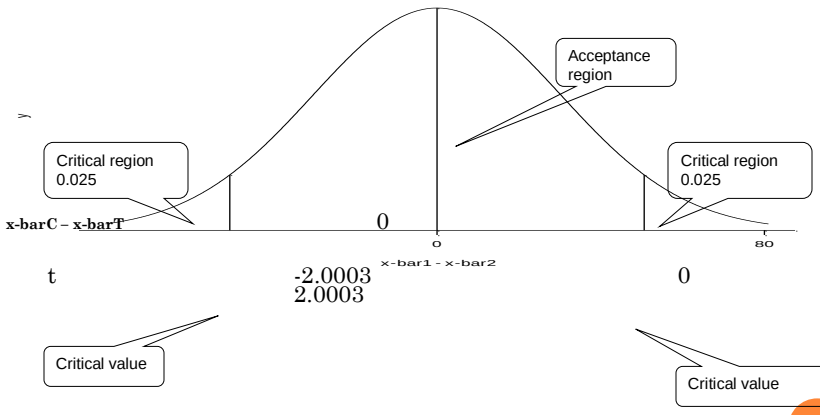
$$s_p = \sqrt{\frac{s_1^2(n_1-1)+s_2^2(n_2-1)}{n_1+n_2-2}} = \sqrt{\frac{46400+68850}{63}} = 42.8 \quad \text{mg/dL}$$

- test istatistiği: 1 SE_{Diff} = 10.7 mg/dL

$$t = \frac{\bar{x}_1-\bar{x}_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{300-220}{42.8\times\sqrt{\frac{1}{30}+\frac{1}{35}}} = \frac{80}{10.7} = 7.48$$

3/28/2013 124

İKI GRUBUN ORTALAMALARI ARASINDAKI FARKIN ÖRNEKLEM DAĞILIMI



SONUÇ

- 7.48 OLAN T-DEĞERİ GERÇEKTE 80 mg/Dl OLAN FARKLLILIĞI TEMSİL EDER.
- 7.48 değeri kritik bölgeye düştüğünden sıfır hipotezi ret edilir. Sonuçta 6 ay sonunda ortalama kolesterol seviyeleri iki grupta farklıdır (H_A).

P<0.05, istatistiksel önemliliği gösterir.Gerçek populasyonlar ortalamaları arasındaki farklılık olmadığı verildiğinde veya bilindiğinde, örnek ortalamaları arasında gözlenen farkın 80 mg/dL daha yüksek olması olasılığı %5 ten küçüktür. (Gerçekte< 0.00000 t-dağılımlar tablosu)

GERÇEK POPULASYON ORTALAMALARI FARKI İÇİN %95 GÜVEN ARALIĞI

- Yine %95 güven aralığı elde etmek için aşağıdaki formül kullanılmaktadır:

$\mu_C - \mu_T$ için %95 güven aralığı:

$$(\bar{x}_1 - \bar{x}_2) \pm (t_{n_1+n_2-2} \times s_p \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}})$$

t t-kritik değeridir.

$$= (300 - 220) \pm (2.0003 \times 42.8 \times \sqrt{\frac{1}{30} + \frac{1}{35}})$$

= 80mg/dL ± 21mg/dL = (59mg/dL, 101mg/dL)

- ($H_{0\ \text{doğru iken}}$) o güven aralığına düşmemektedir. Burada %95 güvenilirlikle gerçek populasyon ortalamaları farkı 59 ile 101 arasındadır.Her iki önemlilik testi ve güven aralığı yaklaşımında sıfır hipotezinin yanlış oluğu gösterilmektedir. P<0.05, istatistiksel olarak önemli farklılık.

3/28/2013 127

BAĞIMLI GRUPLARDA ÖNEMLİLİK TESTLERİ

- Sadece bir hasta grubu araştırıldığında ve ölçümler defalarca tekrarlandığında.
- Bu verilere tekrarlanmış ölçümler adı verilir.(takip çalışmalarından, kontrollü kliniksel denemelerden)
- Ayrı çiftlerde veya aynı kişinin çift olan organlarında ölçümler.Sol kulak- sağ kulak,sol kol-sağ kol, akciğerde sağ lob ve sol lob
- İkiz ve üçüzlerde aynı verilerin ölçülmesi
- Vaka kontrol çalışmalarında hastaların eşleştirilmesi.

3/28/2013 128

ÖRNEK PROBLEMLER

- 1000 kişinin yaşadığı bir belde de ilk fiziksel muayenede 250 inin hipertansif hasta (sistolik kan basıncının 140 mm Hg üzeri ve diyastolik kan basıncının ise >90 mm Hg) olduğu görülmüştür. Bunlar 5 yıl daha takip edildiklerinde 50 kişinin daha hipertansif oldukları gözlenmiştir. Bunlardan 5 inde kooner kalp hastalığı tanısı konulmuş ve tanıdan 2 sene sonra ölmüşlerdir.
- Bu bilgiler ışığında bu belde de ilk muayenede hipertansiyon prevalansı aşağıdakilerden hangisidir?
- A-%25, b-%75, c-%05, d-%99.5 –e-%25.5
- Cevap: a: araştırmanın başlangıcında var olan vakaların sayısı (250) /beldedeki kişi sayısı(1000)
- Bu belde de 5 yılın sonundaki kumulatif hipertansiyon insidansı aşağıdakilerden hangisidir?
- 50/750, 50/1000,300/1000, 300/750 ve 305/1000
- Cevap: a: yeni hipertansiyon vakası(50)/risk altındaki kişiler(1000-250=750)

129

ÖRNEK PROBLEMLER

- Bir hastalık için Endemi tanımı;
  - A-hastalık beklenenin üzerinde ortaya çıkar
  - B-toplumlarda doğal olarak mevcuttur
  - C- mevsimsel bir değişim modeli gösterir
  - D-aynı zamanda bir çok ülkeyi etkilemesi durumudur.
  - E-Hastalığın o sene çok düşük hızlarda olmasıdır.
  - Cevap:b
- Bir hastalık için epidemi tanımı
  - A-Hastalığın bir bölge veya beldede sürekli fakat düşük hızlarda görülmesi durumudur
  - B-Atak hızının 1000 de 10 dan daha yüksek olması durumu
  - C-Belirli zamanda toplumda hastalığın varlığı normal beklenen sayının üzerinde olması durumudur.
  - D- 100000 de Yıllık vaka hızı
  - E-Hastalığın o sene çok yüksek hızlarda olmasıdır.
  - Cevap:c

130

ÖRNEK PROBLEMLER

- Kronik airflow sınırlaması olan 6 hastada , Fernandez ve arkş bu gibi hastalarda gaz değişimini iyileştirmek amacı ile uyguladıkları tedavinin etkilerini vücut yüzey alanlarını(metre-kare) olarak aşağıdaki ölçmüşlerdir.
  - 2.2, 1.8,1.6,1.9, 1.4, 2.2
- Verileri özetlemede aşağıdaki ölçütler hesaplanmıştır. Yanlış değer alan ölçütü bulunuz?
  a) Ortalama:1.85
  b) Medyan=1.85
  c) Mode:2.2
  d) Range:0.80
  e) Minimum değer=1.4 bulmuşlardır.

  Cevap:b
- Standard sapma verilerde;
  a) Farklılığın boyutunu ölçer
  b) Uç değerleri standardize eder
  c) Yaygınlığı ölçer
  d) Pozitif ve negatif değerlerdeki kargaşayı önler
  e) Standardize edilmiş değerleri ölçer
  - Cevap:c

131

ÖRNEK PROBLEMLER

- Detaylı bir araştırmadan sonra bir gastroenterolog intestenal hastalıklarda başlangıç yaşının 25 yaş, yine medyan başlangıç yaşını 17.6 yaş ve modal yaşın 15.1 yaş olduğunu rapor eder. Hastalık 8 -45 yaş grubu hastalarda görülmektedir. Aşağıdakilerden hangi eğilim ölçütü tercih edilen eğilim ölçütüdür?
  a) Aritmetik ortalama
  b) Medyan
  c) Mod(tepe değeri)
  d) Minimum
  e) Maksimum

  Cevap:b

132

ÖRNEK PROBLEMLER

- Aşağıdaki problemi okuduktan sonra alternatif hipotez hangisidir işaretleyiniz?
- Hamileliği sırasında günde bir paketten fazla sigara için 23 kadının ilk doğan çocuklarının ağırlıkları hiç sigara içmeyen 16 kadının ilk çocuklarının doğum ağırlıklarından ortalama 200 gr daha düşüktür. Farklılık önemlilik 0.05 seviyesinde istatistiksel olarak önemli bulunmuştur ($p<0.05$). Bu aşağıdakilerden hangisidir?
  a) Gebelikte sigara kullanımı fetal (cenin in) gelişimini geciktirir.
  b) Ortalama doğum ağırlıklarında gözlenen farklılık şansa bağlı ortaya çıkabilmesi için çok yüksektir.
  c) Ortalama doğum ağırlıklarında gözlenen farklılık kolaylıkla sadece şansa bağlı olarak ortaya çıkabilir.
  d) Bu çalışmaya alınan hasta sayıları bu sonucu bulmada yeterli değildir.
  e) Gebelikte sigara kullanımı fetal (cenin in) gelişimini etkilemez..

  Cevap:b

133

ÖRNEK PROBLEMLER

- Üç ay boyunca 25 erlik 50 gönüllü kolesterol düşürücü iki farklı diyete tabi tutulmuşlardır. İki diyet kolesterolü düşürmede farklı etkiye sahipmidirler? Aşağıdaki veriler toplanmıştır.
- 1. grupta: 180,225,243,150,212,218,235,299,310,177,275,288,220,199,204,266,…,200 mg/dL
- 2. grupta:182,220,241,140,222,216,232,291,305,176,175,282,218,195,260,184,…,199 mg/Dl

Kolesterol seviyelerinin Normal dağılım gösterdiği varsayıldığında, bu iki ayrı diyet uygulanan gruplar ortalama kolesterol seviyeleri yönünden farklı olup olmadıklarını test etmede hangi uygun istatistiksel test kullanılmalıdır?

a) Ki-kare testi
b) İki bağımsız grupta t-testi
c) Eşleştirilmiş t-testi
d) İşaret testi
e) Yukarıdakilerin hiç birisi

f) Cevap: b

134

ÖRNEK PROBLEMLER

- Üç ay boyunca 25 erlik 50 gönüllü kolesterol düşürücü iki farklı diyete tabi tutulmuşlardır. İki diyet kolesterolü düşürmede farklı etkiye sahipmidirler? Aşağıdaki veriler toplanmıştır.
- 1. grupta: 180,225,243,150,212,218,235,299,310,177,275,288,220,199,204,266,…,200 mg/dL
- 2. grupta:182,220,241,140,222,216,232,291,305,176,175,282,218,195,260,184,…,199 mg/Dl

Gerçekte diyet kolesterol seviyelerini düşürdüğü halde araştırma sonucunda , diyet kolesterol seviyelerini düşürmemektedir kararı verildiğinde aşağıdaki lerden hangisini düşünürsünüz?

a) Tip I hata
b) Sıfır hipotezi
c) Tip II hata
d) Alternatif hipotezi
e) Doğru kabul kararını

135

ÖRNEK PROBLEMLER

- Sığınmacı kamplarında yaşayan çocuklar arasında beslenme durumunun araştırıldığı bir çalışmada aşağıda verilen özet ölçütler elde edilmiştir.
  - Ortalama beslenme indeksi:89.5
  - Standard sapma:9.9
  - Ortalamanın standard hatası:0.7
  - Bu verilen bilgiler e dayanarak ve dağılımın normal dağılım olduğu varsayımı ile ortalama için %95 lik güven aralığı yaklaşık
    - 70.1 ve 108.9
    - 79.6 ve 99.4
    - 88.1 ve 90.9
    - 88.8 ve 90.2
    - 71.8 ve 98.9
    - Cevap: c
    - %95 lik güven aralığı formülü
    - %95 Güven aralığı= (z değeri kritik z olarak kullanılmıştır

136

ÖRNEK PROBLEMLER

Cevap: c %95 lik güven aralığı formülü
%95 Güven aralığı hesaplamada (z =1.86değeri kritik z olarak kullanılmıştır)

Formül:

$$\overline{x} \pm \left(z \times \frac{\sigma}{\sqrt{n}}\right)$$

89.5± 1.96 x (0.7)= 88.1,90.9)
Beslenme indeksi %95 güvenilirlikle alt sınırı 88.1 ve üst sınırı 90.9 olan aralığa düşecektir.

137

ÖRNEK PROBLEMLER

- Standard Hata aşağıdakilerden hangisini temsil eder?
  a) Gerçek populasyon ve örnek ortalaması arasındaki farkı
  b) Ortalamayı ölçmede sistematik hatayı
  c) Ortalamada gözlemlerdeki yaygınlığı
  d) Gerçek populasyon ortalaması hakkındaki örnek ortalamalarındaki yaygınlığı
  e) Verilerdeki yaygınlığı
  f) Cevap: d

İstatistiksel önemlilik:

a) p-değeri , 1.tip hatadan büyük olduğunda (p>α)
b) P-değeri II.tip hatadan büyük olduğunda (p>β)
c) P<α
d) β= α
e) 1- β>0.90
f) Cevap:c

138

•467 üniversite öğrencisi Human Papilloma Virus
•infeksiyonu yönünden PCR testi ile taranmışlardır. Gerçek durum belirlemede Southern blot teknikği gold standard olarak kabul edilmektedir. Buna göre

Test sonuçları	Gerçek enfeksiyon durumu +	-	Toplam
PCR +	190	23	213
PCR -	10	244	254
Toplam	200	267	467

Buna göre :Duyarlılık=a/(a+c), seçicilik=d/(b+d)

Pozitif prediktif değeri= a/ (a+b)

Negatif prediktif değeri= d/ (c+d)

139

ÖRNEK PROBLEMLER

Bu çalışma grubundaki Gerçek HPV prevalansı

a) 190/467=0.41
b) 200/467=0.43
c) 213/467=0.46
d) 190/213=0.89
e) 190/200=0.95 cevap:b

Bu çalışma grubundaki PCR duyarlılığı nedir?

a) 190/467=0.41
b) 200/467=0.43
c) 213/467=0.46
d) 190/213=0.89
e) 190/200=0.95 cevap:e

140

ÖRNEK PROBLEMLER

Gerçekte enfeksiyon olmayan bir kadının PCR test sonucu pozitif olduğunda test sonuçları aşağıdakilerden hangisine girer?

a) Yanlış negatif
b) Yanlış pozitif
c) Gerçek negatif
d) Gerçek pozitif
e) Hiç birisi

Cevap :b

141