TEZ YAZIMI

Dr.Önder Ergönül
Koç Üniversitesi Tıp Fakültesi
Enfeksiyon Hastalıkları ve Klinik Mikrobiyoloji Bölümü

7 Haziran 2014

AGUH, KLİMİK

Yazarlar neden yazar?

Umberto Eco;

“her türlü bilgi akışı anlatı üzerinden gerçekleşir”

“Üniversitede tezimi değerlendiren profesörlerimden biri beni ‘Sen tez değil tezi nasıl yazdığının romanını yazmışsın’ diye eleştirdi. O günden sonra felsefe kitabı gibi görünen romanlar yazmaya başladım. En sonunda da roman yazmayı öğrendim”

9 Nisan 2013, Boğaziçi Üniversitesi söyleşisi, Hürriyet

Yazarlar neden yazar?

Orhan Pamuk,

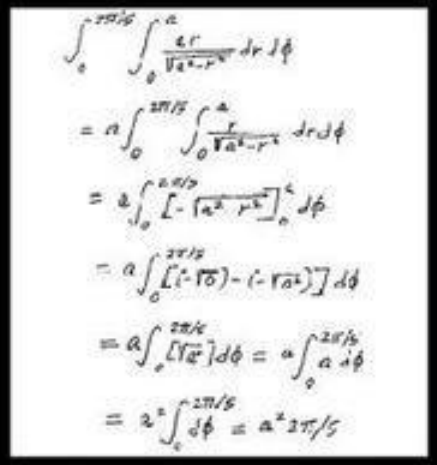
“Bir odada yalnız başıma yaşamak, aynı zamanda da kabul görmek istiyordum”

9 Nisan 2013, Boğaziçi Üniversitesi söyleşisi, Hürriyet

En büyük yazarlar, Homeros, Dante, Shakespeare, farklı olanla ilgilendikleri ve ancak kayda değer ayrıntıları aktardıkları için başarılıdır

William Strunk & E.B. White

STATISTICIAN



$$\int_0^{2\pi/5}\int_0^a \frac{ar}{\sqrt{a^2-r^2}}\,dr\,d\phi$$
$$= a\int_0^{2\pi/5}\int_0^a \frac{r}{\sqrt{a^2-r^2}}\,dr\,d\phi$$
$$= a\int_0^{2\pi/5}\left[-\sqrt{a^2-r^2}\right]_0^a d\phi$$
$$= a\int_0^{2\pi/5}\left[(-\sqrt{0})-(-\sqrt{a^2})\right]d\phi$$
$$= a\int_0^{2\pi/5}\left[\sqrt{a^2}\right]d\phi = a\int_0^{2\pi/5} a\,d\phi$$
$$= a^2\int_0^{2\pi/5} d\phi = a^2\,2\pi/5$$

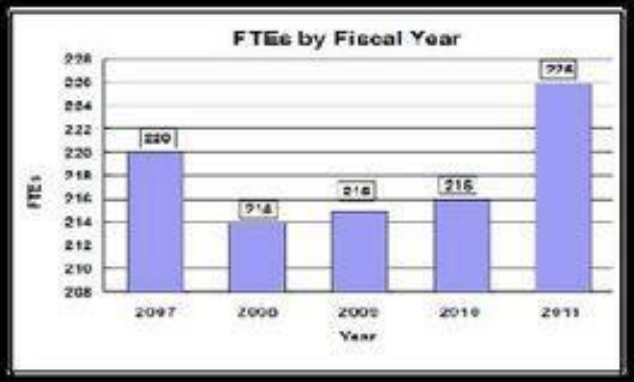
What My Friends Think I Do

What My Colleagues Think I Do

What My Parents Think I Do

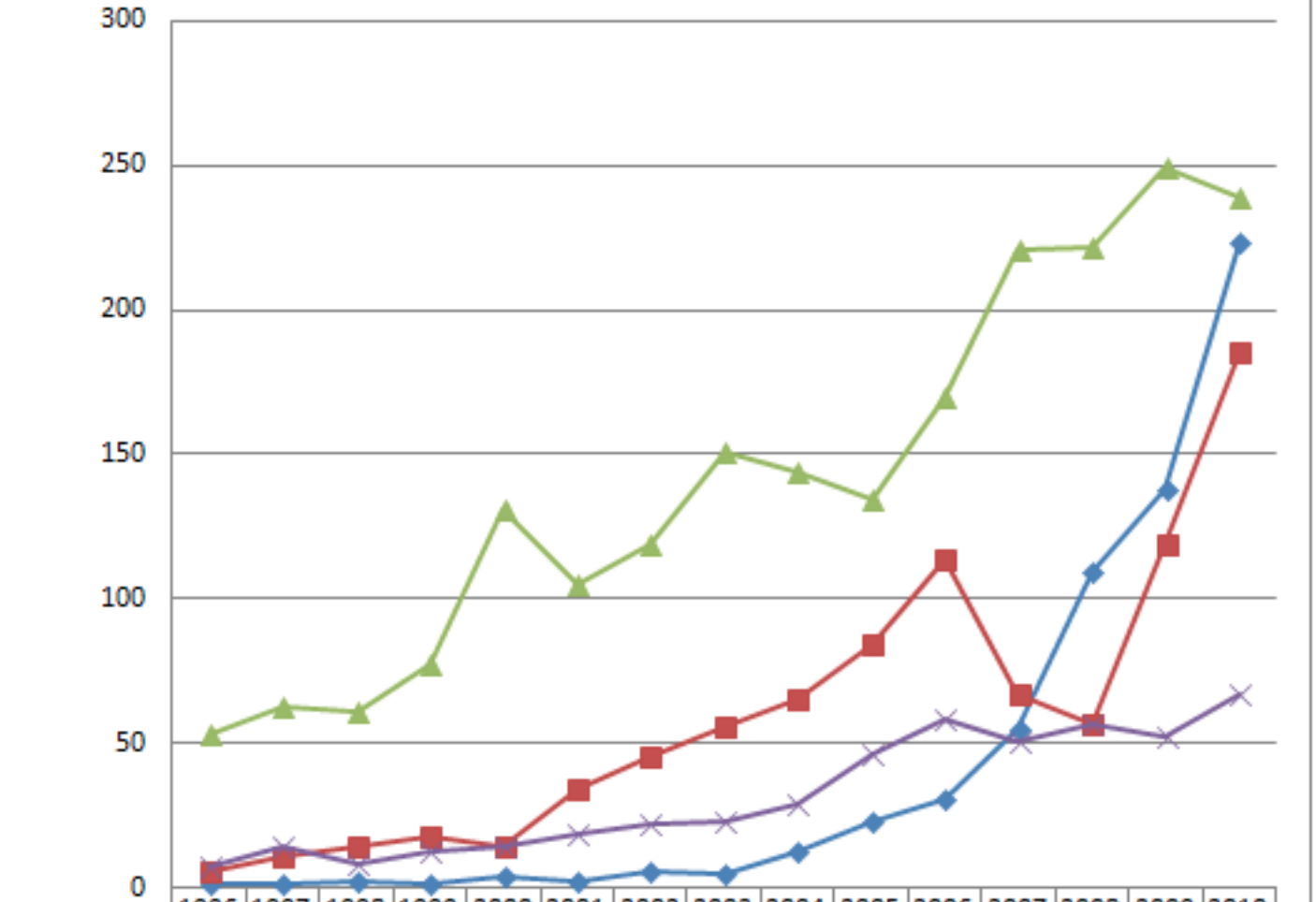
What Society Thinks I Do

What I Think I Do



What I Actually Do

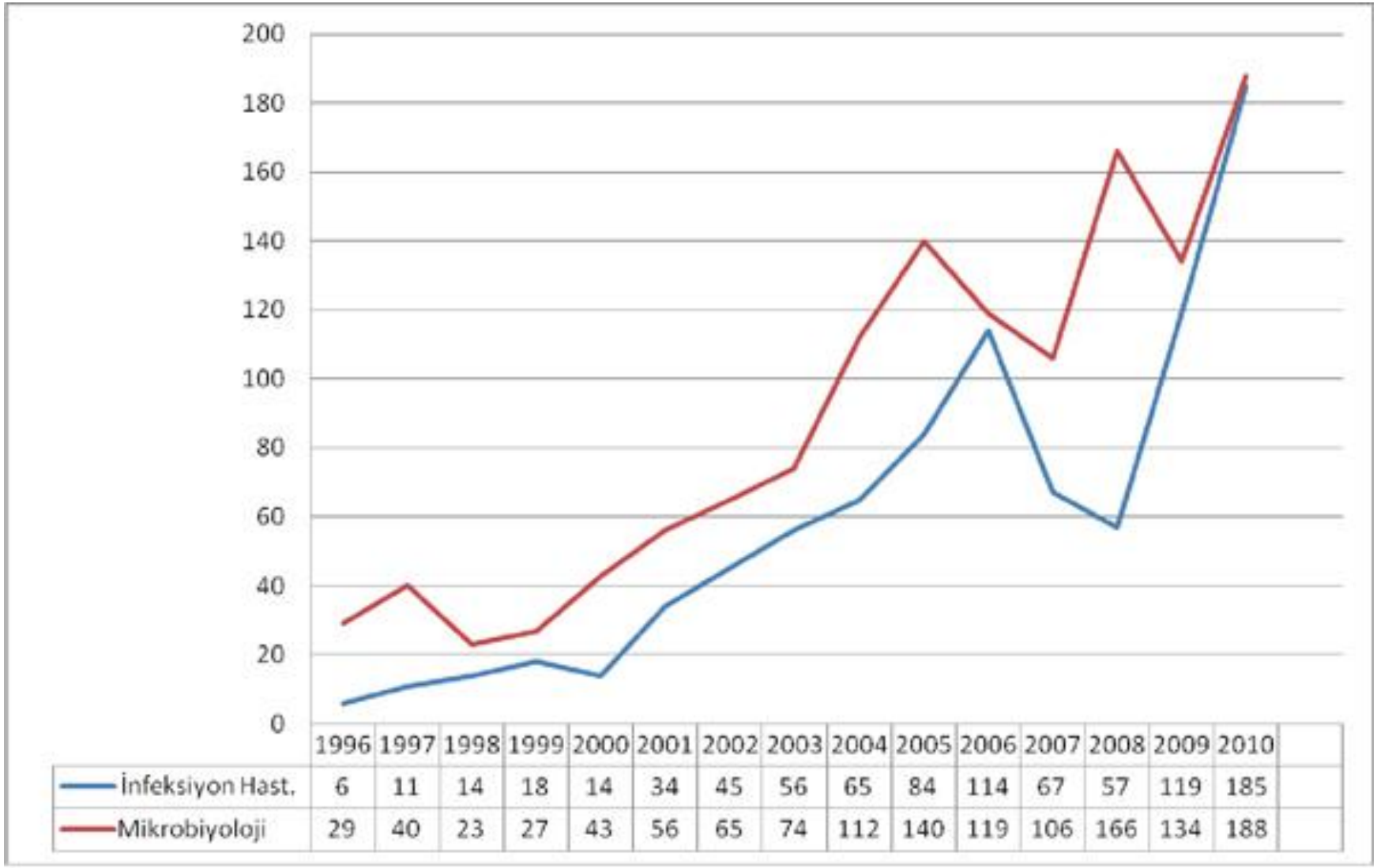
Enfeksiyon Hastalıklarında Yayınlar ve Türkiye'nin Durumu



	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Iran	1	1	2	1	4	2	6	5	13	23	31	55	109	138	223
Turkey	6	11	14	18	14	34	45	56	65	84	114	67	57	119	185
Italy	53	63	61	77	131	105	119	151	144	134	170	221	222	249	239
Greece	7	14	8	13	14	19	22	23	29	46	58	51	57	52	67

Klimik Bulteni, Ekim 2011

Yayın Üretimi 2006 Yılında Neden Düştü?



	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
İnfeksiyon Hast.	6	11	14	18	14	34	45	56	65	84	114	67	57	119	185
Mikrobiyoloji	29	40	23	27	43	56	65	74	112	140	119	106	166	134	188

Klimik Bulteni, Aralık 2011

Hekimlerin Yazma Nedenleri

- Entellektüel stimulasyon
- Düşüncelerimizi paylaşmak
- Araştırmalarımızı bildirmek
- Fikir ileri sürmek
- Tartışma başlatmak
- Disiplin oluşturmak
- Bir başlığa sahiplenmek
- Meslekte ilerlemek
- Olgu bildirmek
- Kişisel ün sağlamak
- Unutulmamak (ölümsüzlük)
- Para kazanmak

Taylor RB. The Clinician's Guide to Medical Writing Springer, 2005.

Hekimlerin Yazmama Nedenleri

- Zamansızlık
- Konusuzluk
- Çalışma arkadaşı yokluğu
- Sekreter yokluğu
- Yöntem bilmemek
- Mentor (danışman) yokluğu
- Motivasyon eksikliği
- Kendine güvensizlik
- Başlayamamak
- "Yazmaktan nefret etmek"

Taylor RB. The Clinician's Guide to Medical Writing Springer, 2005.

Çalışmalar Neden Yayınlanır?

- Sorunun “önemi”
- Yazarların ünü
- Denek sayısının çokluğu
- Beklenmedik girişim ve sonuçlar
- Günlük yayınların ilgisini çekecek mi?
- Sonuçlarda belirgin farklılıklar olması

Çalışmalar Neden Yayına Kabul Edilmez?

- “Önemli” bir bilimsel soruya yanıt vermemektedir
- Orijinal değildir
- Yazarın hipotezini araştırmamaktadır
- Tasarımı yanlıştır
- Örneklem sayısı azdır
- Kontrol kolu yoktur
- İstatistik analiz uygun değildir ya da yanlıştır
- Sonuçlar yanlış yorumlanmıştır
- Kötü yazılmıştır

Araştırma Sonuçları

1. Hata/Yanlılık (Bias)
2. Karıştırıcı etken
   (Confounder)

Tasarım

3. Şans faktörü → Analiz

Çalışma Evreleri

1. Ön hazırlık, hipotez kurma
2. Çalışma tasarımı
3. Veri toplama
4. Verilerin değerlendirilmesi ve analiz
5. Yazım

Occam' ın Usturası

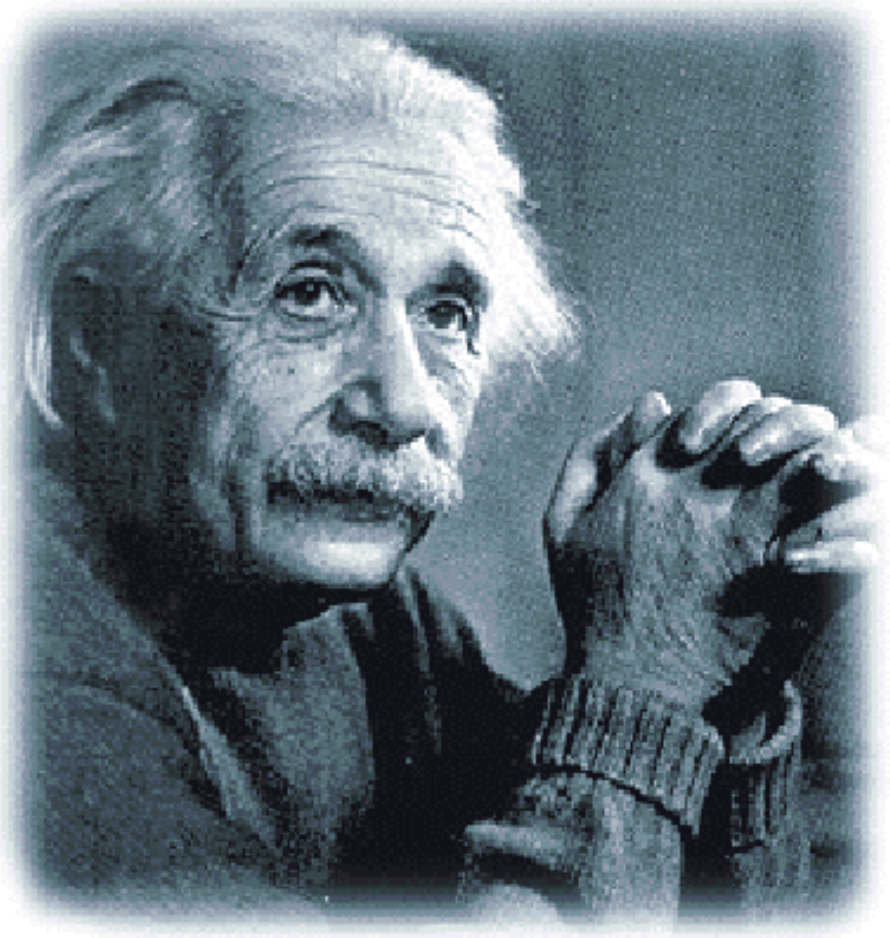
Occam: 14 yy, İngiliz felsefeci ve din bilimci

"Birşeyleri açıklamak için ortaya atılan varsayımlar, ihtiyacın ötesinde çoğaltılmamalıdır"

Taylor RB. The Clinician's Guide to Medical Writing Springer, 2005.

"Rakip varsayımlar arasında en doğru varsayım, en basit olanıdır"

Wynn CM, Wiggins AW. Yanlış Yönde Kuantum Sıçramalar. TÜBİTAK, 2005.



“Basite indirgeyin, kolayına kaçmayın…”

Basit ≠ Kolay

Klinik Araştırmaların Sınıflandırılması

Kanıtın Kalitesi

I. En az bir uygun tasarlanmış randomize klinik çalışma Meta-analiz?

II-1. Randomizasyon yapılmamış iyi tasarlanmış kontrollü çalışma

II-2. İyi tasarlanmış, tercihen birden çok merkezde yapılmış kohort veya vaka-kontrol çalışma

II-3. Girişimli ya da girişimsiz zaman serileri

III. Saygın otoritelerin klinik deneyime dayanan görüşleri, tanımlayıcı çalışmalar, uzman komisyonların raporları

RKÇ veya
Meta-analiz
Randomizasyonsuz
kontrollü
Çok merkezli kohort veya
olgu-kontrol
Uzman görüşü, tanımlayıcı
çalışmalar, uzman raporları

Çalışma Tasarımları

Çalışma Tasarımları

Maruziyet (exposure) → **KOHORT** → Sonuç (outcome)

Maruziyet (exposure) ← **Vaka-Kontrol** ← Sonuç (outcome)

Kesitsel çalışma

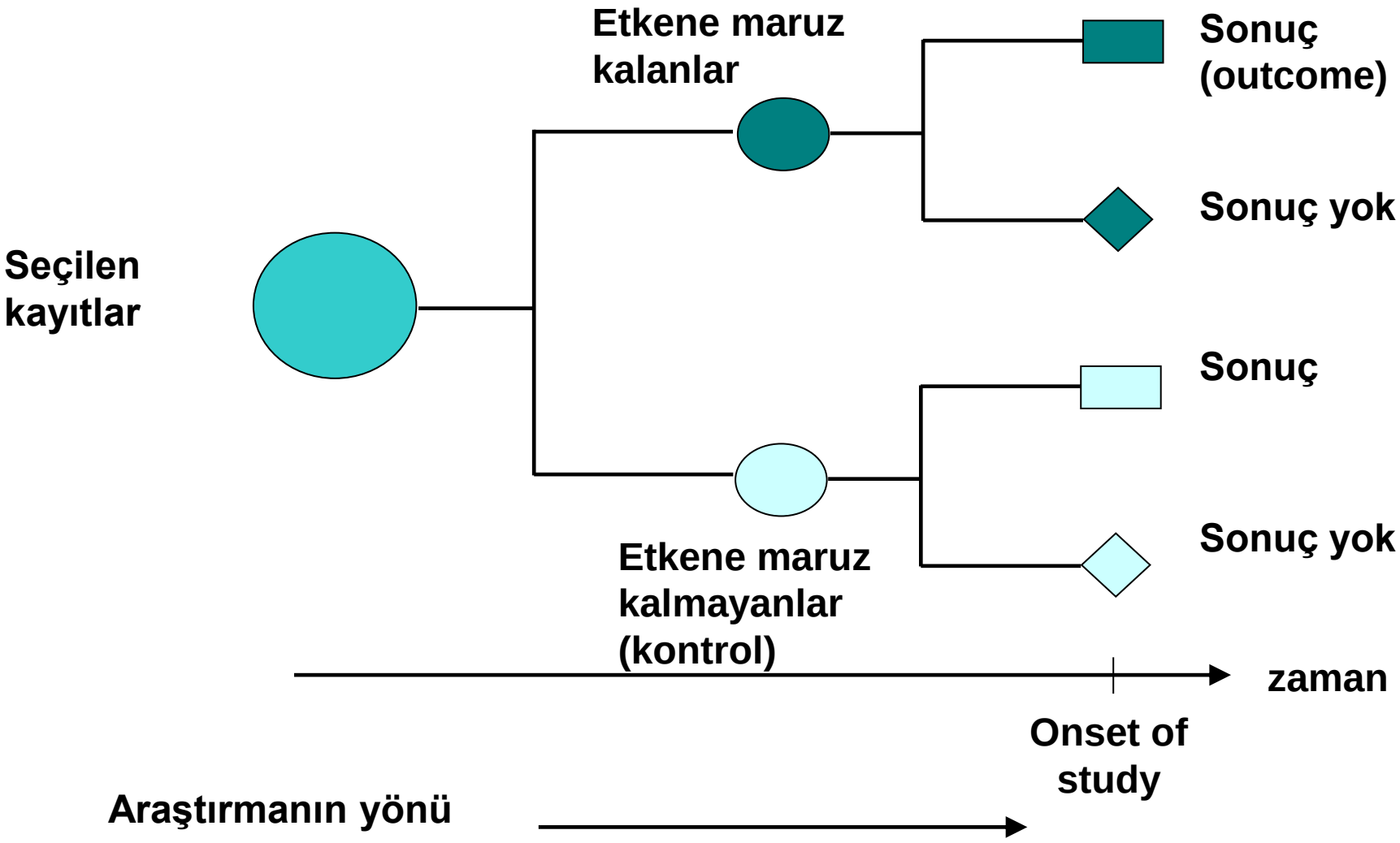
Maruziyet

↕

Sonuç

→

Retrospektif Kohort



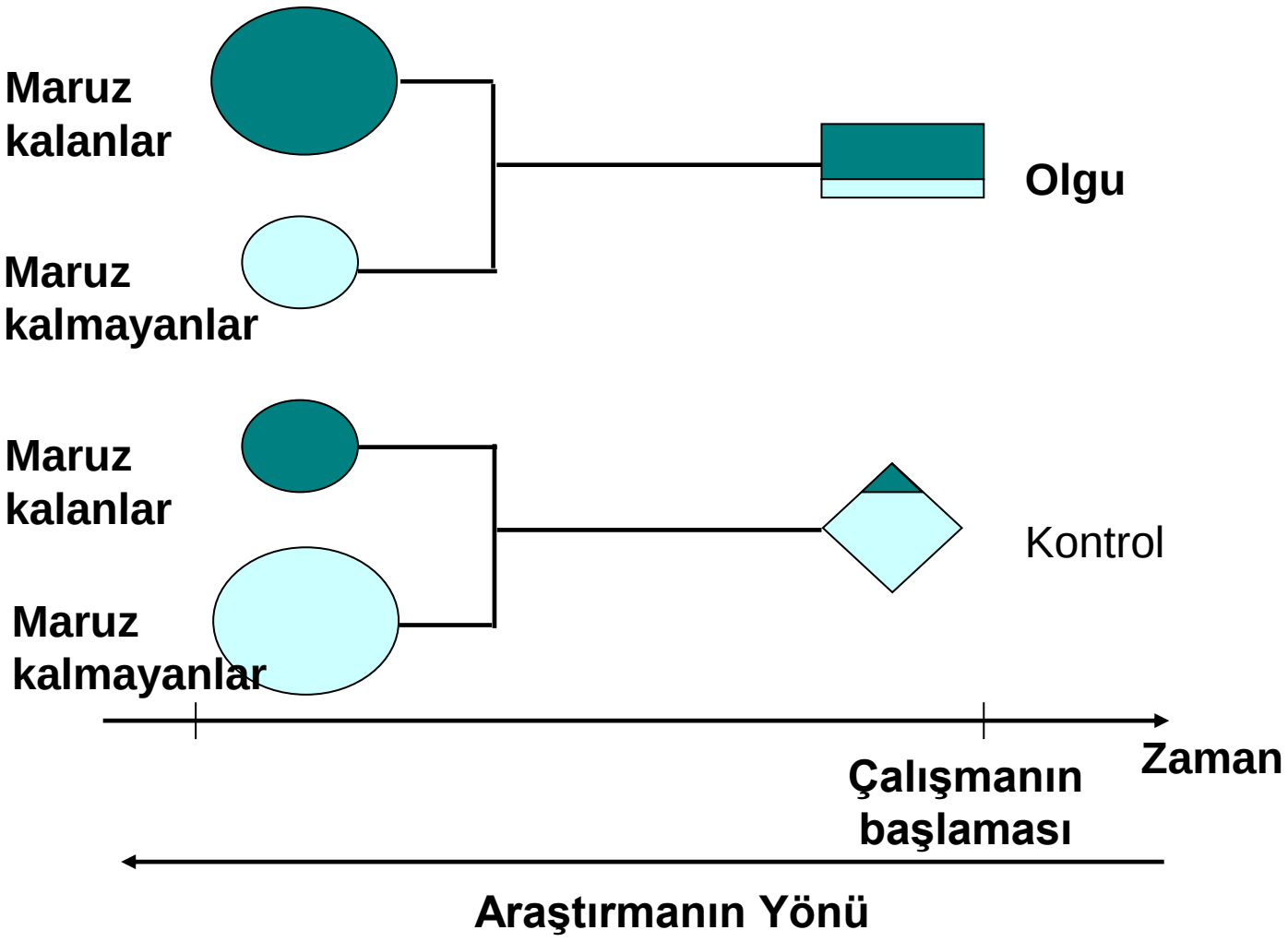
En İyi Bilinen Kohort Çalışmalar

- Framingham Kalp Çalışması
- Hemşirelerin Sağlığı Çalışması
- Hekimlerin Sağlığı Çalışması

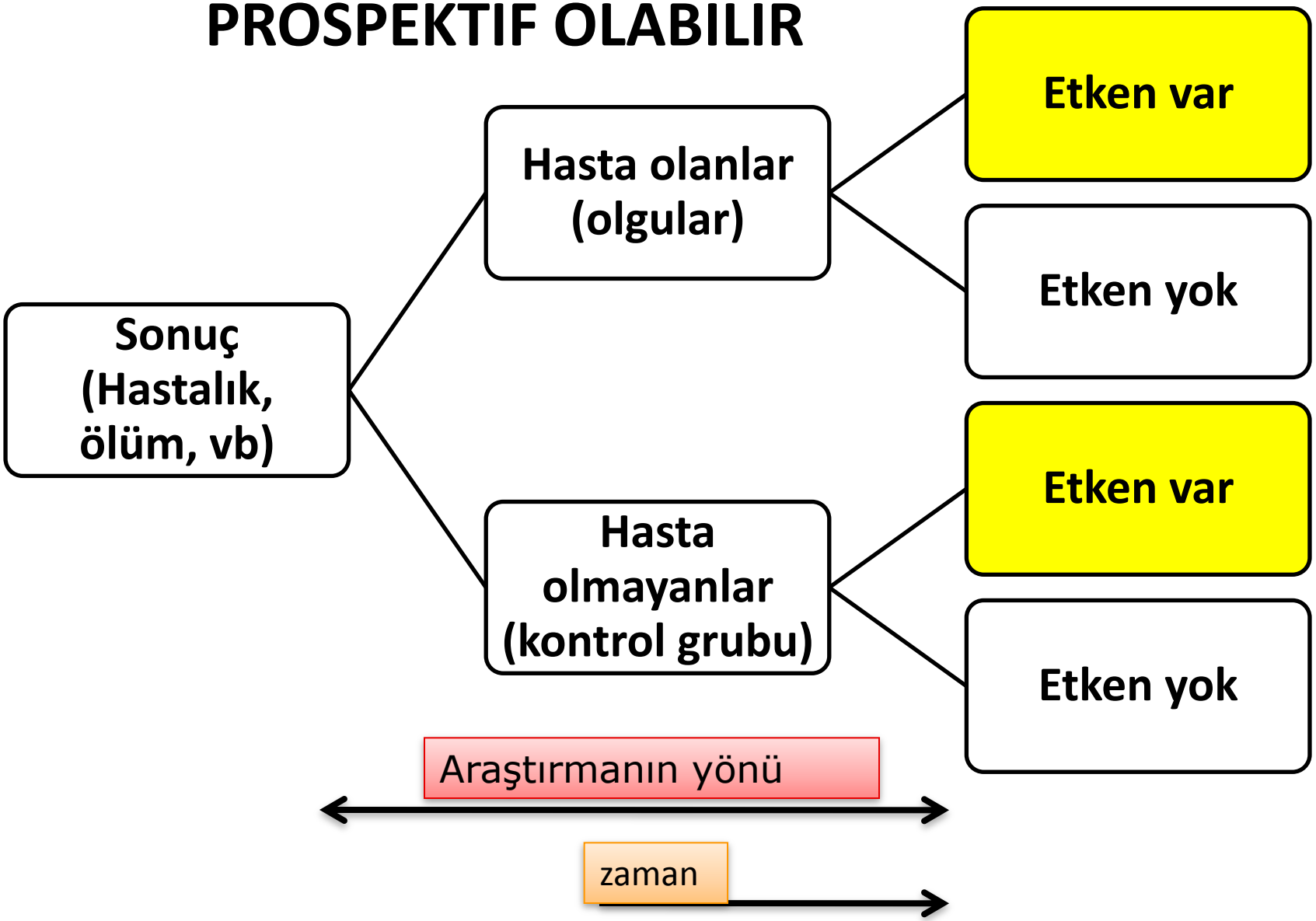
Kohort Çalışmalar: Framingham Heart Study

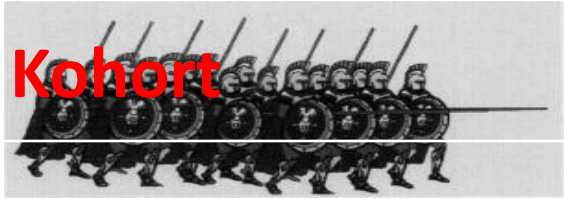
www.**framinghamheartstudy**.org bakılacak

Olgu Kontrol Çalışma Tasarımı



OLGU KONTROL ÇALIŞMALAR PROSPEKTİF OLABİLİR



	Kohort 	Vaka-Kontrol
1. Başlangıç noktası	Neden (maruz kalma)	Sonuç (vaka)
2. Çalışma süresi	Uzun	Kısa
3. Maliyet	Pahalı	Ucuz
4. Çalışmanın yönü	Etkenlerden sonuca ya da olgulara	Sonuçtan ya da olgulardan etkenlere
5. Gerçekleştirme	Zor	Kolay

Sıradışı bir durum!

Jaffe HWf, et al. National Case-Control Study of Kaposi's Sarcoma and Pneumocystis carinii Pneumonia in Homosexual Men. Ann Int Med 1983; 99: 145-151.

1981 Haziran/Temmuz
Kaliforniya ve New York'da,
genç, sağlıklı homoseksüellerde
pneumocytsis carinii pneumonia (PCP) ve
Kaposi sarcoma(KS) sıklığı artmaktadır.

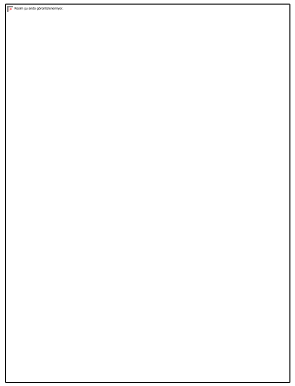
Sıradışı bir durum!

Amaç: Genç homoseksüellerde Kaposi sarkomu ve Pneumocystis carinii pnömonisi görülmesini belirleyen risk faktörlerinin araştırılması

Yöntem: Olgu-Kontrol / 50 olgu ve 150 kontrol

Sonuç:

- Saunalara gidenler
- Sifiliz öyküsü olanlar
- İlaç kullananlar
- İnhale nitrit kullananlar
- İtalyanlar (HLA-DR5 antijeni ile ilişkili olarak)



2008 Nobel Tıp Ödülü Sahipleri: HIV

Francoise Barre-Sinoussi **Luc Montagnier** **Harald zur Hausen**

Olgu Kontrol Çalışmalar

The NEW ENGLAND JOURNAL of MEDICINE

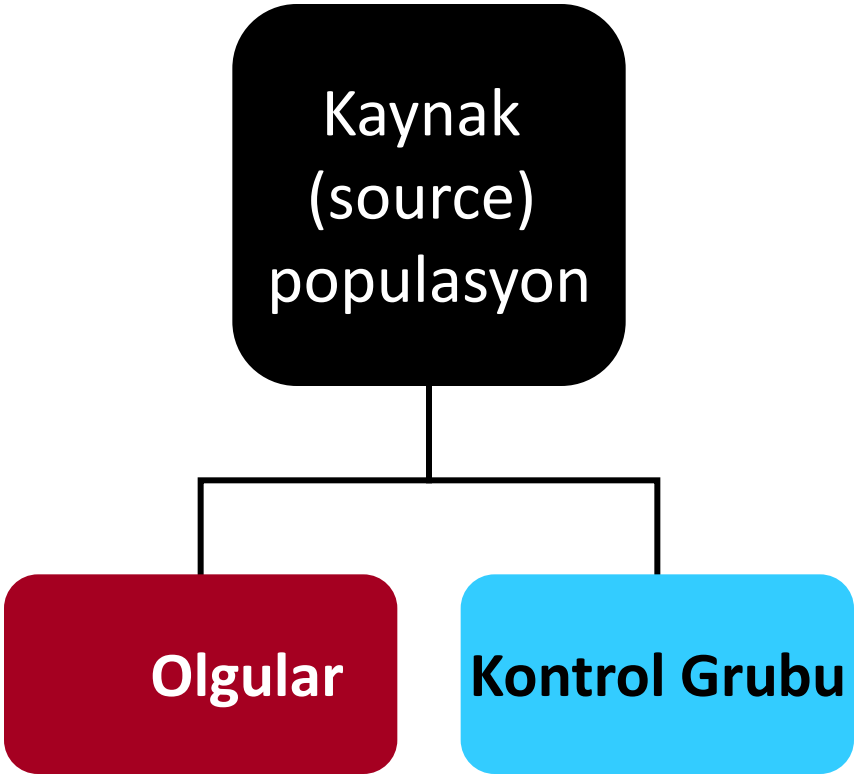
ORIGINAL ARTICLE

An Outbreak of Hepatitis A Associated with Green Onions

Variable	Patients (N=240)	Controls (N=134)	Odds Ratio (95% CI)
	no. (%)		
Age >12 yr	212 (88)	97 (72)	2.9 (1.7–5.0)
Menu items consumed			
Mild salsa†	218 (91)	45 (35)	19.6 (11.0–34.9)
Grilled chicken fajitas‡	29 (12)	8 (6)	2.2 (1.0–4.9)
Bar beverage§	49 (21)	16 (12)	1.9 (1.0–3.5)
Ground beef burrito	7 (3)	0	Indeterminate
Enchilada platter	7 (3)	0	Indeterminate
Ingredients consumed¶			
Green onions‖	234 (98)	76 (58)	33.3 (12.8–86.2)
Diced white onions**	237 (99)	95 (72)	30.8 (9.3–102.2)
Beans	113 (47)	35 (26)	2.5 (1.6–4.0)
Mixed cheese	181 (75)	76 (57)	2.3 (1.5–3.7)
Rice	124 (52)	46 (34)	2.0 (1.3–3.2)
Sour cream	92 (38)	33 (25)	1.9 (1.2–3.0)
Corn-cake mix	124 (52)	49 (37)	1.9 (1.2–2.9)

Çalışma Grubu

- Olgu kontrol çalışmalarında en büyük sorun, kontrol grubunun oluşturulmasıdır.
- Kontrol grubu, olgular ile aynı kaynak popülasyondan olmalıdır.



Potansiyel Yanlılık (bias, hata) Tipleri

- Örnek seçerken *(selection bias)*
  - Kendi kendine seçim
  - Tanısal

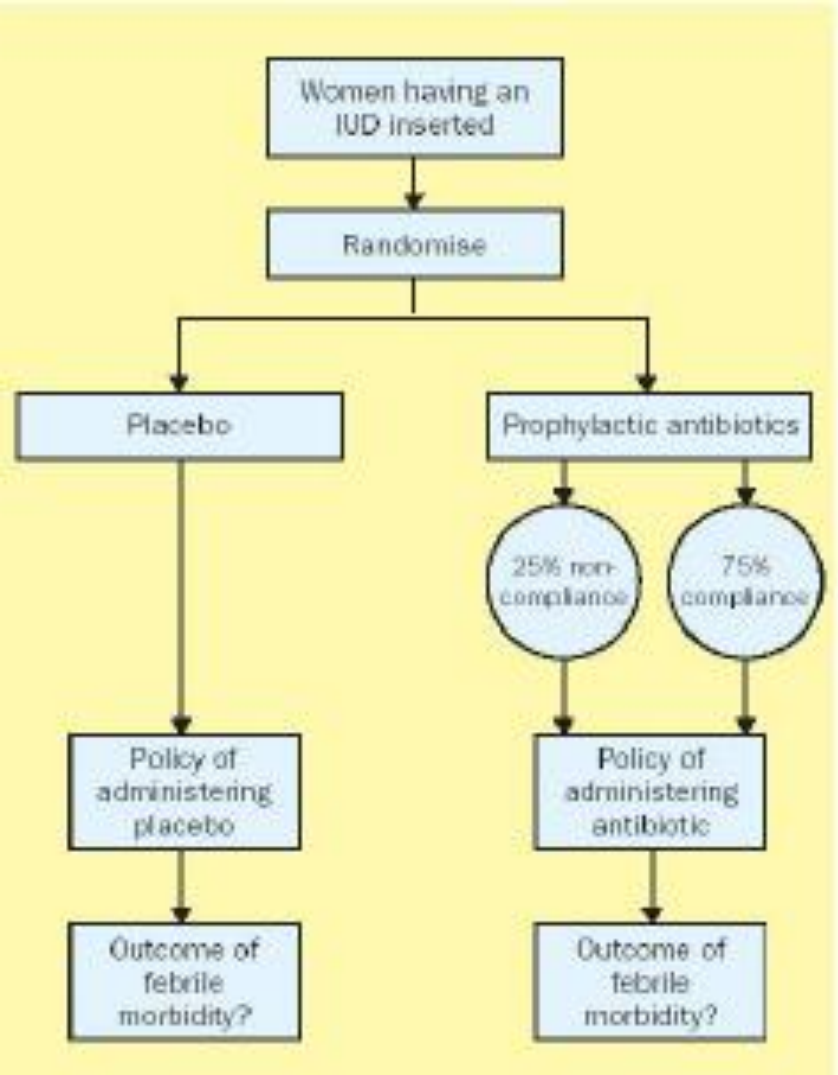
- Bilgi toplarken
  - Yanlış gruplama *(misclassification bias)*
  - Hatırlayamama *(recall bias)*

Akış Diyagramı

Tasarlarken

Analiz yaparken

Okuyucu için



Tanımlayıcı İstatistikte Temel İstatistik Kavramlar

Bilimsel Yayının Bulgular Bölümünde 3 Ana Tablo Vardır

1. Demografik veriler
2. Tek değişkenli analiz bulguları
3. Çok değişkenli analiz bulguları

Sunum Planı

1. Veri Türleri
2. Merkezi Eğilim Ölçütleri
3. Merkezi Yayılım Ölçütleri
4. Örneklem Büyüklüğü

Değişken (*variable*)

- İstatistikte değişken (variable) terimi, verilere ait özellikler anlamında kullanılır.
- Cinsiyet, boy, ağırlık, kan basıncı, ilaç grubu, tedavi öncesi kan basıncı vb. hepsi birer değişkendir.
- Bilgisayar veritabanlarında;
  - Değişkenler, kolonda
  - Veriler, satırlarda gösterilir.

Veri Toplama

- Kağıt form
  - Her durumda gerekli, en azından şimdilik...
- Excell
- Stata
  - Veri editörü
- Access
  - Elektronik form oluşturma; birden çok sonuç alınmak istendiğinde

Veri Toplama

Veri Toplama Tez Yazımında En Önemli Adımlardan Biridir!

- Veriler kesişmemeli
- Seçenekler birbirini dışlamalı
- Her bir kolonun amacı ve anlamı olmalı
- Tüm veriler tek bir veritabanında toplanmalı

Stata/SE 8.0

File Edit Prefs Data Graphics Statistics User Window Help

Stata Editor

Preserve Restore Sort << >> Hide Delete...

death[1] = 0

	death	incubation	prehospsym~s	incless8	gender	age	fever	lowext	upperext	h
1	0	3	1	0	0	29	0	1	0	
2	0	2	1	0	0	62	0	0	1	
3	1	2	2	0	0	68	1	0	0	
4	1	2	3	0	1	70	0	1	0	
5	0	3	1	0	0	25	0	1	0	
6	0	2	1	0	0	48	0	1	0	
7	0	4	1	0	0	20	0	1	0	
8	0	2	2	0	0	66	0	1	0	
9	1	2	1	0	0	18	1	0	1	
10	1	2	1	0	0	68	0	1	0	
11	1	2	2	0	0	68	1	1	0	
12	1	2	2	0	0	57	1	1	0	
13	1	2	2	0	0	51	1	0	1	
14	0	2	1	0	0	22	0	1	0	
15	0	2	1	0	0	51	0	1	0	
16	1	2	1	0	1	45	1	0	0	
17	1	2	2	0	0	52	0	1	0	
18	0	2	1	0	0	38	0	0	1	
19	0	3	1	0	0	54	0	1	0	
20	0	3	1	0	0	30	1	1	0	
21	1	1	1	1	1	54	0	1	0	

C:\Documents and Settings\Onder Ergonul

Microsoft Excel

File Edit View Insert Format Tools Data Window Help

F516 =

keskinercalismasonuclariSON

	E	F	G	H	I	J	K	L	M	N
1	SKAR	SON	35 yaş altı Doktor				TBC	AİLEDE	PPD	BOOSTE
2	SAYISI	BCG TAR	MESLEK	M.YILI	BÖLÜMÜ	B.YILI	ÖYKÜ	TBC	mm	mm
3	3	0	D	7	9	3	0	0	9mm	10mm
4	2	0	D	1	9	1	0	0	10mm	
5	2	0	D	3	1	1 ay	0	0	11mm	
6	2	0	D	2	1	3 ay	0	0	5mm	10mm
7	2	0	D	3	1	3	0	0	7mm	10mm
8	1	3	D	3	1	3	0	1Ağabey	14mm	
9	2	0	D	2	1	2	0	0	8mm	11mm
10	1	0	D	4.5	4	4	0	0	12mm	
11	2	0	D	3	1	2	0	0	12mm	
12	1	3	D	9	5	3	0	1Ağabey	0mm	0mm
13	3	0	D	2.5	1	1.5	0	0	11	
14	1	0	D	4	1	3	0	1A.anne	18	
15	3	0	D	5	1	5	0	0	17	

Sayfa1 Sayfa2 Sayfa3

Ready

Microsoft Access

File Edit View Insert Format Records Tools Window Help

Demographics

reviewed ☐

MRN:

Account: 0

Age: 0

Gender:

InitialInfectionDate:

Admit Date:

Discharge Date:

Location at time of Infectio

Admitting Service:

primary diagnosis:

☐ IDconsultation

ID notes:

discharge status:

disposition:

☐ Readmission for same target diagnosis

☐ Readmission within 60 days (Yes/No)

Comments

ABXhospital | micro | Mod | dischargeABX | radiology | clinical findings

datestamp	ID	account	AbxName	StartDate	EndDate	Route	Frequency	Dose	s
)03 15:10:56	(AutoNumber)	0						0	

Record: 1 of 1

Record: 1 of 1

Form View

Microsoft Access

File Edit View Insert Format Records Tools Window Help

Demographics

IDconsultation

Record:

Demographics : Table

	reviewed	ID	MRN	Account	Age	Gender	Location at time	Admitting Service	Admit Date	Dis
▶	☐	1793		0	0					
*	☐	(AutoNumber)		0	0					

Record: 1 of 1

Datasheet View

Çalışma Birimi Nedir?

- Hasta günü
- İlaç günü

Her bir satır neyi gösterir?

Veriler Hangi Tipte?

Veri Türleri

- Kategorik veriler
  - Nominal
    - Dikotom: kadın veya erkek
    - Kan grupları

Characteristic	Peginterferon Alfa-2a plus Placebo (N=271)	Peginterferon Alfa-2a plus Lamivudine (N=271)	Lamivudine (N=272)
Male sex — no. (%)	214 (79)	208 (77)	215 (79)
Race or ethnic group — no. (%)†			
White	24 (9)	23 (8)	32 (12)
Asian	237 (87)	236 (87)	232 (85)
Black	4 (1)	4 (1)	3 (1)
Other	6 (2)	8 (3)	5 (2)

Veri Türleri

- Kategorik veriler
  - Ordinal
    - Ciddiyet skalası
    - Derecelendirme

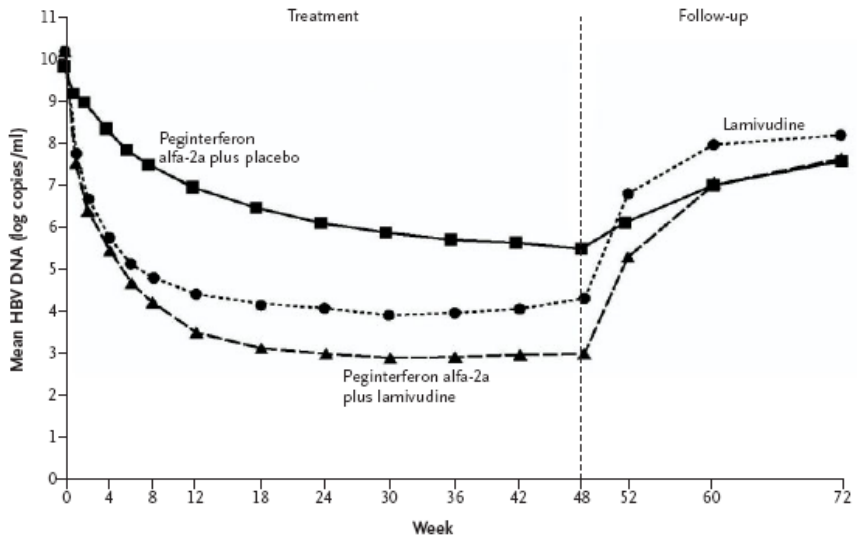
Table 1. (Continued.)

Variable	1 Yr of Trastuzumab (N=1694)	Observation (N=1693)
Histologic grade of tumor — no. (%)		
3 (Poorly differentiated)	1015 (59.9)	1012 (59.8)
2 (Moderately differentiated)	546 (32.2)	554 (32.7)
1 (Well differentiated)	45 (2.7)	42 (2.5)
Not assessed	75 (4.4)	76 (4.5)
Missing	13 (0.8)	9 (0.5)

Veri Türleri

- Devamlı: ölçülebilir değerler
  - Yaş
  - Kolesterol düzeyi

B



No. of Patients													
Peginterferon alfa-2a plus placebo	271	269	265	265	262	263	259	258	255	248	254	241	248
Peginterferon alfa-2a plus lamivudine	271	268	268	267	268	263	259	254	255	249	254	244	254
Lamivudine	272	267	267	267	263	263	259	260	244	249	248	228	241

Figure 1. Rates of HBeAg Seroconversion (Panel A) and HBV DNA Levels (Panel B), from Baseline to Week 72.
HBeAg seroconversion was defined by the loss of HBeAg and the presence of anti-HBe antibody. Log to the base 10 was used. The information about the number of patients refers only to Panel B.

Veri Türleri

	Dikotom	Ordinal	Devamlı
	0	1	11
	1	2	24
	0	3	224
	1	5	45
	1	2	56
	0	4	57
	0	3	866
	1	3	34
	0	1	23

Veri Türleri

	Devamlı, Sürekli	**Kategorik, Nominal**
	17	1
	19	1
	21	1
	24	1
	27	2
	32	2
	33	2
	36	3
	41	3

Veri Türleri

	String, Sözel	**Nümerik, Sayısal**
	Hemşire	1
	Hemşire	1
	Hekim	2
	hemşire	1
	Hekim	2
	teknisyen	3
	hemşire	1
	hekim	2
	teknisyen	3

Tanımlayıcı İstatistik
merkezi eğilim ölçütleri

- Aritmetik Ortalama
- Ortanca
- Mod
- Geometrik Ortalama

Tanımlayıcı İstatistik
merkezi eğilim ölçütleri

Aritmetik ortalama (mean, kısaca ortalama)

- Değerlerin toplamının denek sayısına bölünmesiyle elde edilir
- X üzerinde çizgiyle gösterilir.
  - Sürekli değişkenler için kullanılır
    - Trombosit sayısı, karaciğer enzimleri, diğer biyokimya testleri

UYARILAR

- Aşırı değerlerden etkilenir.
- Kategorik (nominal veya ordinal) değişkenler için kullanılamaz.

Tanımlayıcı İstatistik
merkezi eğilim ölçütleri

Ortanca *(median)*

- Değerler küçükten büyüğe doğru sıralandığında tam ortadaki deneğin değeridir. Denek sayısı çiftse, ortada yer alan iki deneğin değerlerinin ortalaması olarak hesaplanır.
- Ordinal veriler için en iyi merkezi dağılım ölçütü
- Aşırı değerlerden etkilenmez. Bu nedenle aşırı uç değerler varsa, sayısal veriler için de ortanca tercih edilmelidir.

Tanımlayıcı İstatistik
merkezi eğilim ölçütleri

Mod

- Değişken değerleri içinde en sık görülen değerdir.
- Örneğin 100 kişilik bir grupta:
  - 12 kişinin Hb'i 14.5 g/dL
  - 10 kişinin 14.0 g/dL
  - 5 kişinin 13.5 g/dL
  - 3 kişinin 15.5 g/dL ve
  - 70 kişinin Hb değerleri farklı değerler alıyorsa,
    - Hb değerlerinin modu 14.5 g/dL'dir.
- Mod, tıpta nadiren kullanılan bir merkezi eğilim ölçütüdür.

Tanımlayıcı İstatistik
merkezi eğilim ölçütleri

Geometrik ortalama

- Asıl olarak logaritmik dağılım gösteren veriler için kullanılır.
- Aşırı değerlerin olduğu sayısal verilerin merkezi eğilim ölçütü olarak da kullanılabilir.

$$\log(GM) = \sum \frac{\log X}{n}$$

Tanımlayıcı İstatistik
merkezi eğilim ölçütleri

İki grup veri:

1, 2, 3, 4, 1000

1, 2, 3, 4, 5

Stata/SE 9.2

File Edit Prefs Data Graphics Statistics User Window Help

Results

```
. sum  var2, det

                              var2
-------------------------------------------------------------
      Percentiles      Smallest
 1%            1              1
 5%            1              2
10%            1              3       Obs                   5
25%            2              4       Sum of Wgt.           5

50%            3                      Mean                  3
                        Largest       Std. Dev.      1.581139
75%            4              2
90%            5              3       Variance            2.5
95%            5              4       Skewness              0
99%            5              5       Kurtosis            1.7

.
```

D:\data

Başlat 2 Win... (0 unre... Ribaviri... EndNot... Microsof... Stata/S... TR 06:38

Stata/SE 9.2

File Edit Prefs Data Graphics Statistics User Window Help

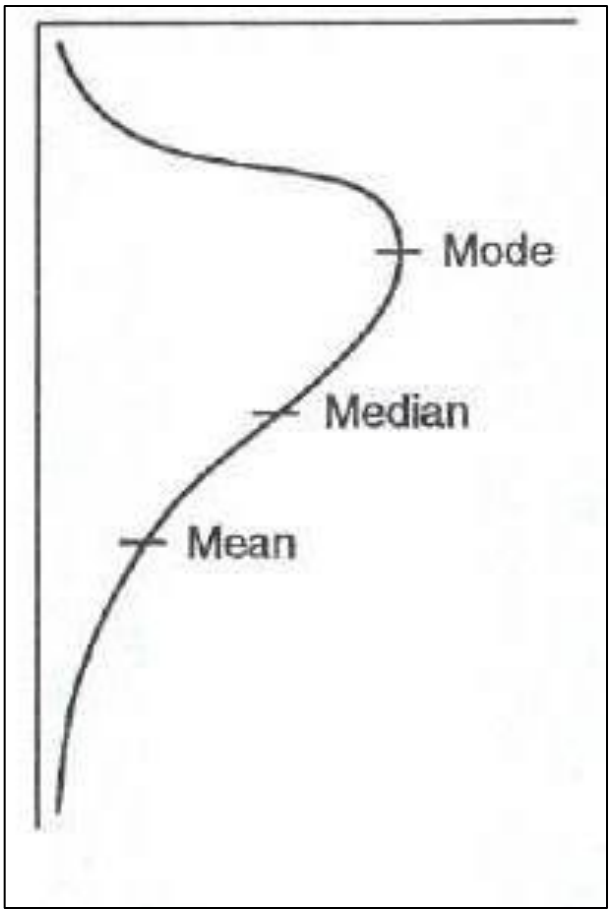
Results

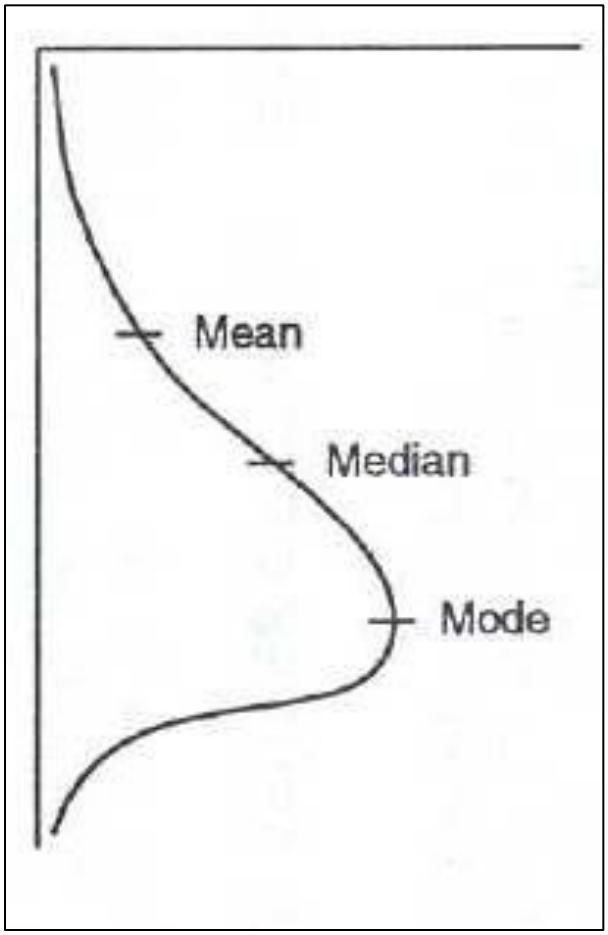
```
. sum  var1,det

                              var1
-------------------------------------------------------------
      Percentiles      Smallest
 1%            1              1
 5%            1              2
10%            1              3       Obs                   5
25%            2              4       Sum of Wgt.           5

50%            3                      Mean                202
                        Largest       Std. Dev.       446.097
75%            4              2
90%         1000              3       Variance       199002.5
95%         1000              4       Skewness       1.499976
99%         1000           1000       Kurtosis       3.249969

.
```

D:\data

Command

Başlat 2 Window... (0 unread) ... Ribavirin_s... EndNote X ... Microsoft P... Stata/SE 9.2 TR 06:44

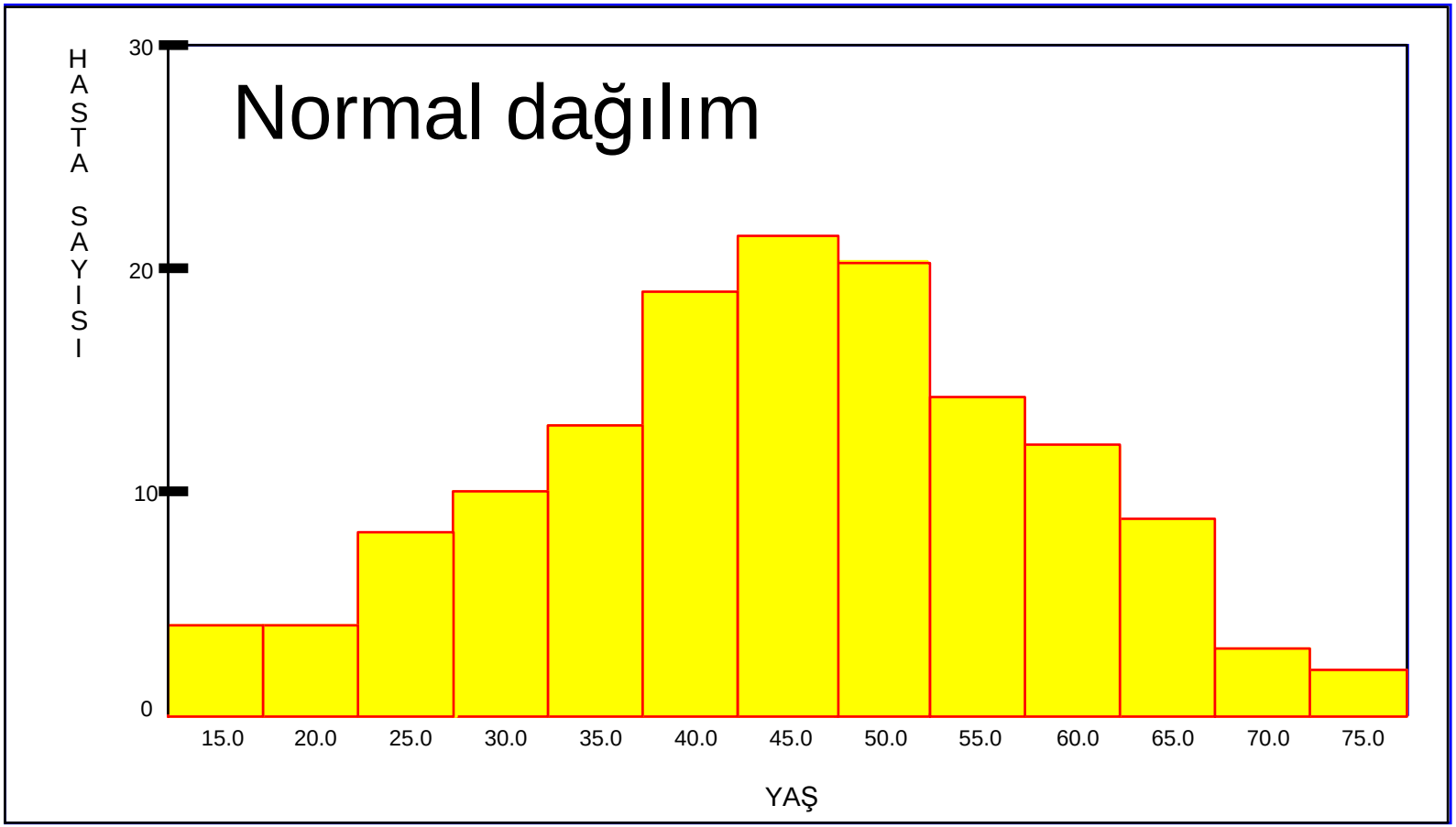

Mode
Median
Mean

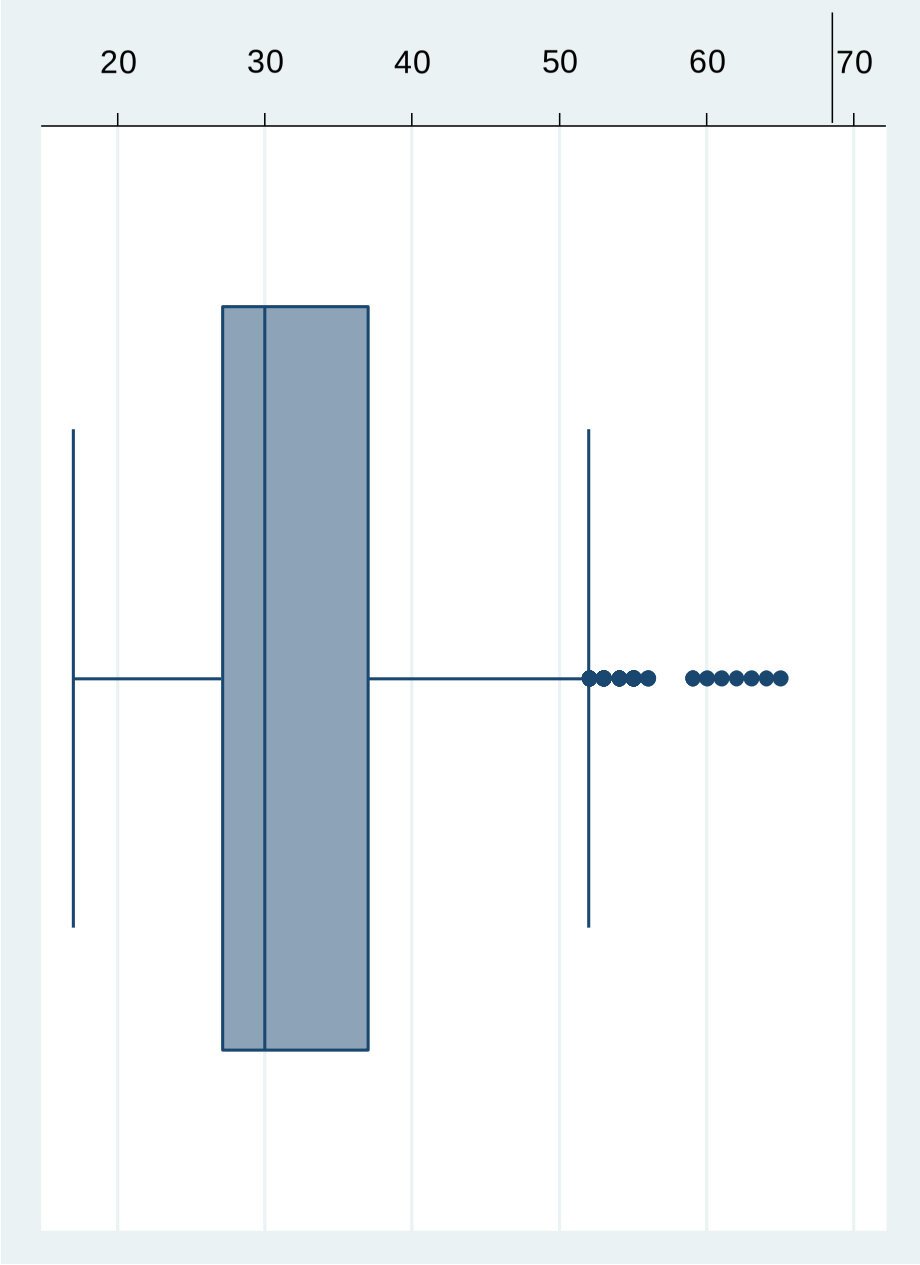

Mean
Median
Mode

Tanımlayıcı İstatistik merkezi eğilim ölçütleri

- Görsel yöntem
  - Histogram
  - Gövde-yaprak grafiği
  - Boksör torbası grafiği
- İstatistik testler
  - Kolmogorov-Smirnov
  - Lilliefors
  - Shapiro wilk
- Varyasyon katsayısı (SD/ortalama)
  - SD/ortalama ≤%30 ise, dağılım ≈ normaldir.

Tanımlayıcı İstatistik
merkezi eğilim ölçütleri




20
30
40
50
60
70

Tanımlayıcı İstatistik yayılım ölçütleri

Değer aralığı *(range) : min-max*

Standart sapma *(standard deviation)*

Persentil *(percentile)*

Çeyreklerarası aralık *(interquartile range)*

Tanımlayıcı İstatistik yayılım ölçütleri

Sedimentasyon değerleri

- Grup 1: 11, 11, 12, 12, 12, 12, 13
- Grup 2: 4, 5, 6, 8, 19, 20, 21

sedim	Obs	Mean	Std. Dev.	Min	Max
grup 1	7	11.85714	.6900656	11	13

sedim	Obs	Mean	Std. Dev.	Min	Max
grup 2	7	11.85714	7.733662	4	21

Tanımlayıcı İstatistik yayılım ölçütleri

- Standart sapma
  - SD, sd, Σ, σ (sigma)
  - Varyans' ın karekökü
  - Varyans' dan daha sık kullanılır
  - Tüm değerler eşitse SD sıfıra eşittir.
  - Değerler arasında farklar arttıkça SD büyür.

Tanımlayıcı İstatistik yayılım ölçütleri

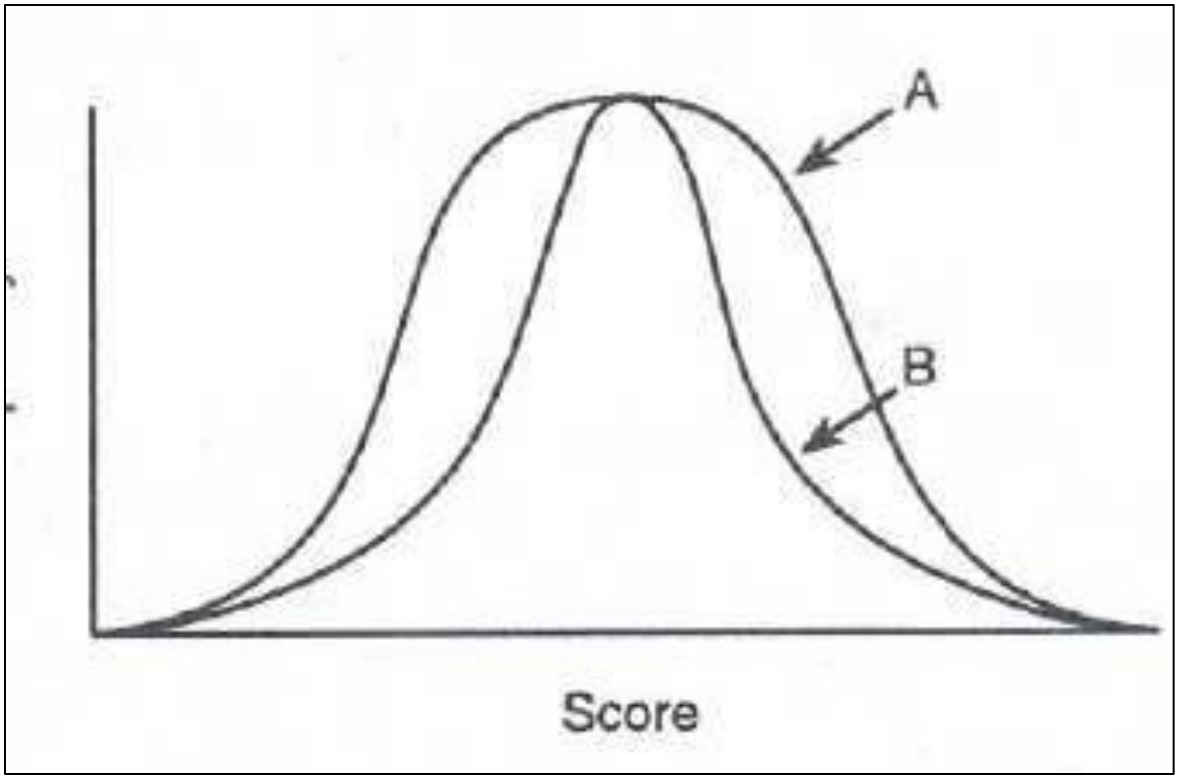
Standart sapma

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

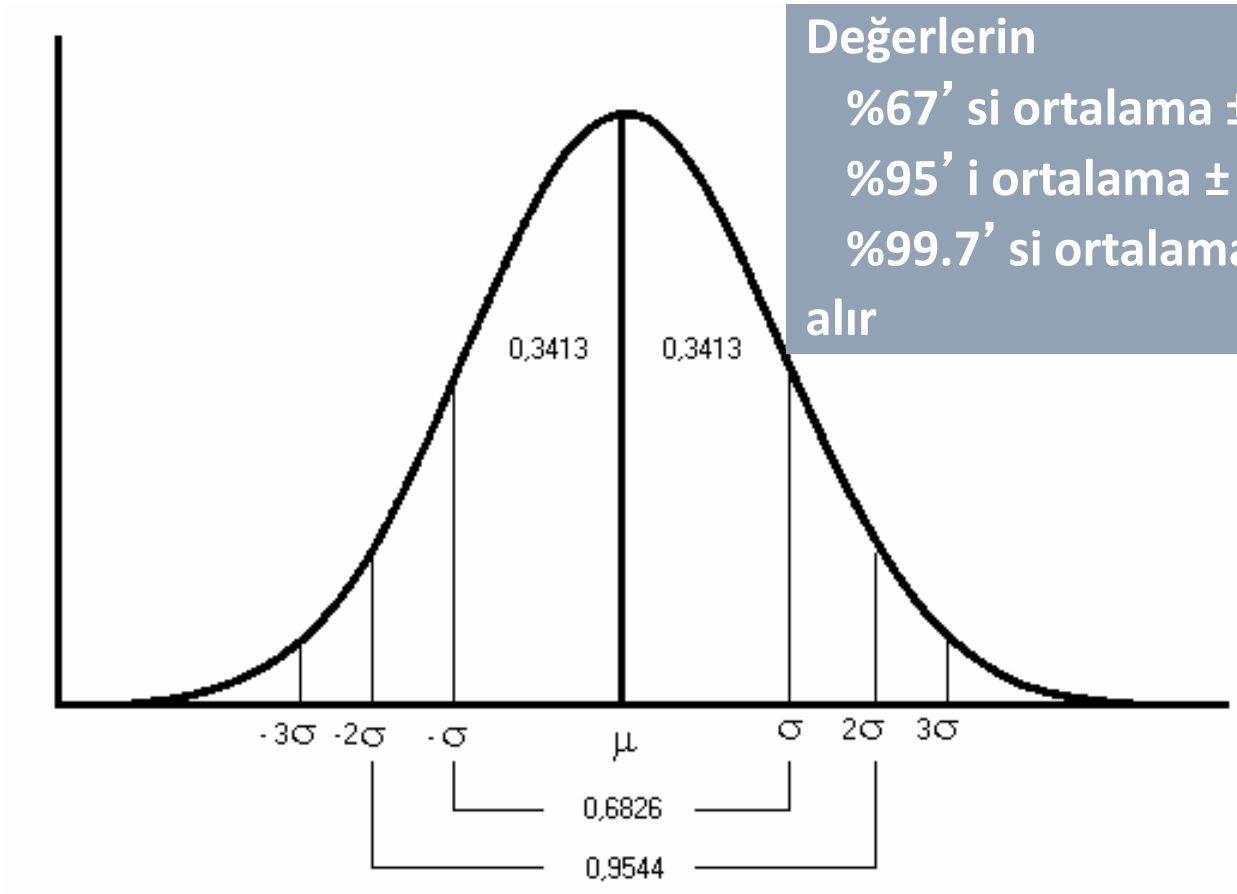
X=ortalama

n=denek sayısı

Hangi verinin standart sapması daha büyük?



Tanımlayıcı İstatistik yayılım ölçütleri



Değerlerin
%67' si ortalama ± 1 SD
%95' i ortalama ± 2 SD,
%99.7' si ortalama ± 3 SD içinde yer alır

Stata/SE 9.2

File Edit Prefs Data Graphics Statistics User Window Help

Results

```
. sum yas, det

                              1.soru
-------------------------------------------------------------
      Percentiles      Smallest
 1%           19             17
 5%           23             17
10%           24             18       Obs                2372
25%           27             18       Sum of Wgt.        2372

50%           30                      Mean           32.06661
                        Largest       Std. Dev.      7.468858
75%           37             62
90%           43             63       Variance       55.78384
95%           46             64       Skewness       .9172034
99%           54             65       Kurtosis       3.664814

.
```

D:\data

Başlat medikomarketing (0 unread) Yahoo! M... Microsoft PowerPoint ... Stata/SE 9.2 TR 06:49

Ortalamanın Standart Hatası

- Standart sapmadan farklı bir ölçüt
- Standart sapma ve örneklem sayısına bağlıdır
- Standart sapma / √ örneklem sayısı
- Örneklem sayısı artıkça standart hata küçülür
- SS: verilerin yayılımını,
  SH: ortalamanın belirsizliğini gösterir

Altman DG. Standard deviations and standard errros. BMJ 2005; 331: 903

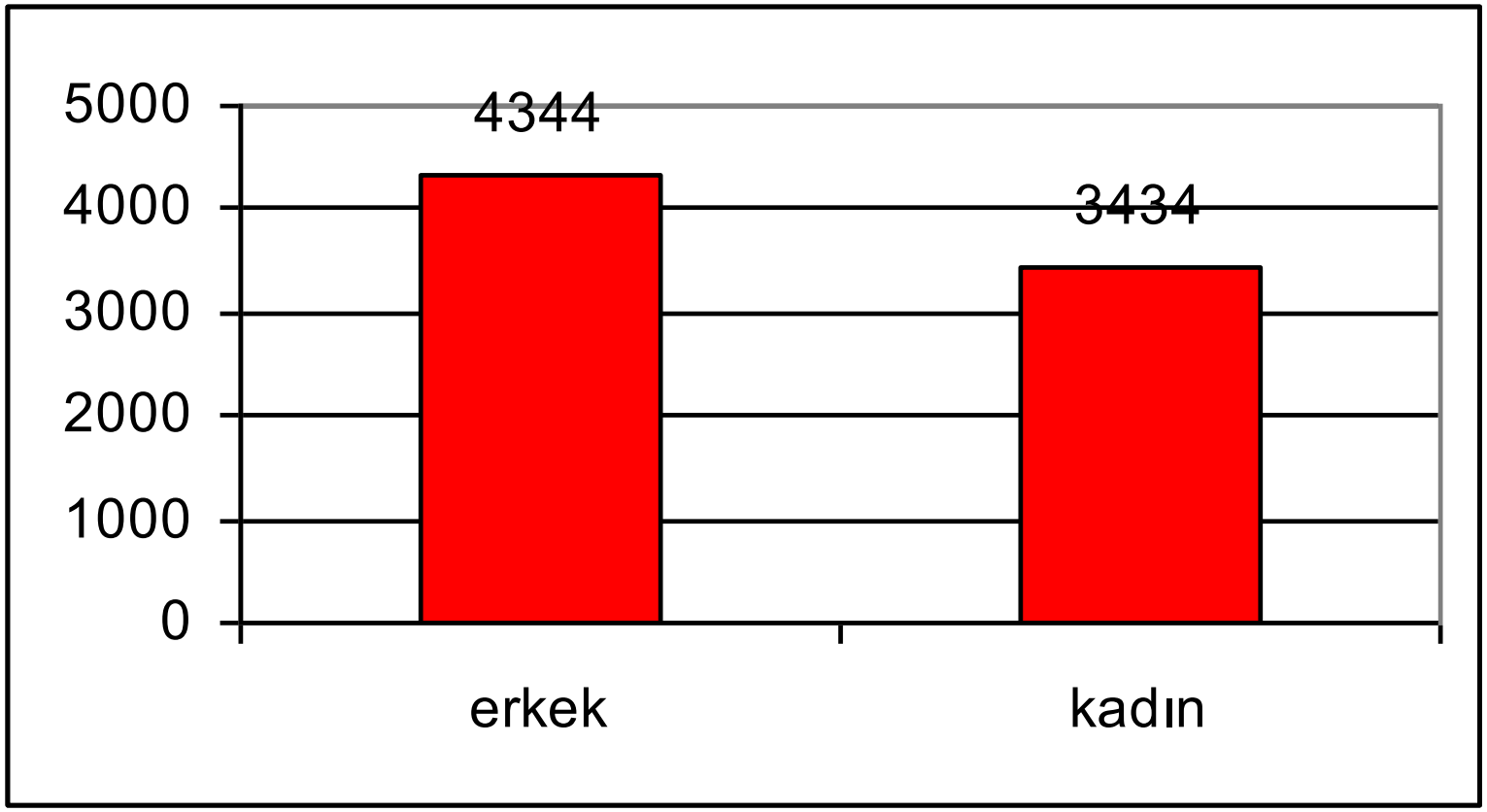
Tanımlayıcı İstatistik

Frekans dağılımı ve kümülatif frekans

Yaş	n=	%	Kümülatif sıklık
5-14	15	17.6	17.6
15-24	19	22.3	39.9
25-34	21	24.8	64.7
35-44	30	35.3	100

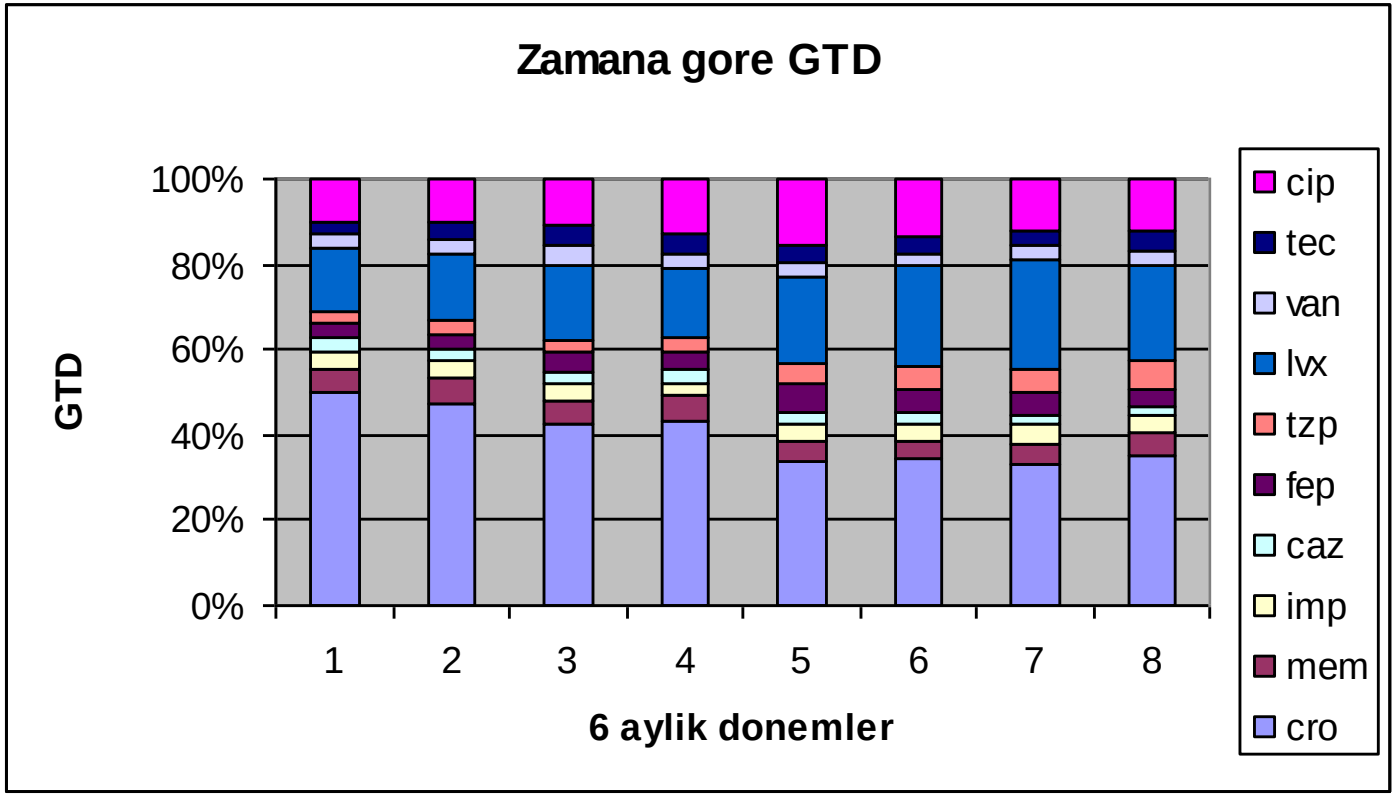
Tanımlayıcı İstatistik

Veri türleri ve grafik tipleri



Tanımlayıcı İstatistik

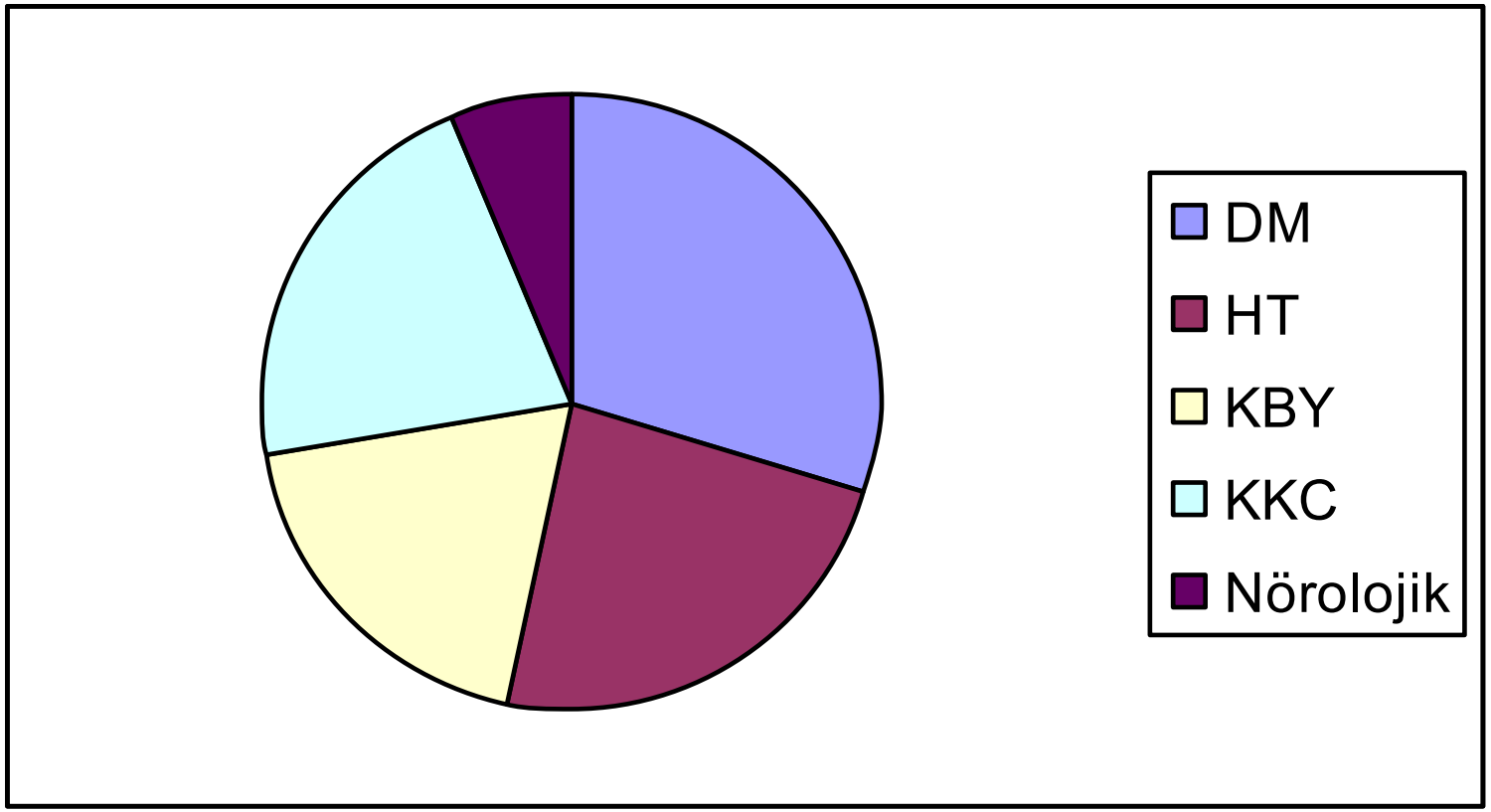
Veri türleri ve grafik tipleri



Yıllar içinde DDD

Tanımlayıcı İstatistik

Veri türleri ve grafik tipleri



Uyarı: Gruplar arasında kesişim olmamalı

Hipotez Testleri
ve
Örneklem Büyüklüğü

İstatistik yöntemlerin evrimi

%	1978-79	1989	2004-05
Sadece tanımlayıcı	27	12	13
İstatistik tabloları (contingency tables)	27	36	53
Epidemiyolojik istatistikler (relatif risk, odds)	10	22	35
Sağkalım analizi	11	32	61
Çok değişkenli regresyon	5	14	51
Güç analizi	3	3	39

Horton NJ, Switzer SS. NEJM 2005; 353.

Hipotezlerin tanımlanması

Araştırmanın amacı belirlendikten sonra bu amaca uygun bir hipotez cümlesi kurulmalıdır.

- **Gruplar arasında fark:**
  - “Erkekler ve kadınlarda ilaca yanıt oranları farklıdır.
  - “Tutulan eklem sayısının 4’ten az olduğu grupla 4 ya da daha fazla olduğu grup arasında ilaca yanıt oranları farklıdır.”
- **Kestirim:**
  - “Yaş, cinsiyet, hastalık şiddeti vb. veriler bilindiğinde hastanın ilaca yanıt verip vermeyeceği kestirilebilir.”

Hipotezlerin tanımlanması

Hipotezinizin önceden belirlenmesi,
uygun istatistik analiz yönteminin seçilmesi için gereklidir.

- Yukarıdaki örneklerde hipotez **“gruplar arasında fark olduğu”** şeklindeyse, ortalamalar/oranlar arası farkların test edilmesi gerekirken, **“kestirim”** yapmak yönünde hipotez kurulmuşsa, regresyon modellerinden birinin seçilmesi uygun olacaktır.

Hipotezlerin tanımlanması

Bunu nasıl ortaya çıkarıyoruz?

- **“0”** Hipotezinin (yani farksızlık hipotezinin) doğruluk olasılığını test ederiz.
- Olasılık önceden belirlenmiş bir düzeyin altındaysa **“H0”** reddedilir.
- Bu düzeye **Alpha Düzeyi** denir.

Hipotez testi

- “Null” hipotezi (H0): “Karşılaştırılan iki değer arasında fark yoktur” önermesi
- Alternatif hipotez (H1): “Karşılaştırılan iki değer arasında fark vardır” önermesi

Hipotez testi

		Gerçek durum	
		Fark var (H1)	Fark yok (H0)
Test sonucu	Fark var (H0 red)	**OK** **Güç veya 1- β**	**Tip I hata (α)**
	Fark yok (H0 kabul)	**Tip II hata (β)**	**OK**

Tip I hata

- Gerçekte tedavi ve kontrol gruplarının **EŞİT** olmasına rağmen istatistik anlamlılık saptanması, **yanlış pozitiflik**, p değeri
- Genellikle 0.05,
  - Araştırmacılar, **%5**'den daha az yanlış pozitiflik hedeflemek isterler.

Tip II hata

- Gerçekte tedavi ve kontrol gruplarının **FARKLI** olmasına rağmen istatistik anlamlılık saptanmaması, **yanlış negatiflik**, p değeri
- Eğer 0.20 ise,
  - Araştırmacılar, **%20**'den daha düşük bir yanlış negatiflik isterler.

“P” değeri nedir?

P *(probability)*, alfa, tip I hata ile ilişkilidir; gruplar arasında gözlenen farkın şans eseri olma olasılığıdır.

- P=0.05 ise, farklılığın şans eseri olma olasılığı %5, yani 20’de 1’dir.
- Bu rakam küçüldükçe şans eseri olma olasılığı azalır.

“P” değeri nedir?

Örnek:

- **Sıfır hipotezi (H0):**

  Çalışma ilacı ile plasebo eşdeğerdir.

- **Analiz sonucu:**

  (istatistiksel olarak anlamlı fark var),

  H0 reddedilir; arada anlamlı bir fark vardır sonucuna varılır.

  ilaç plasebodan farksızdır.

“P” değeri nedir?

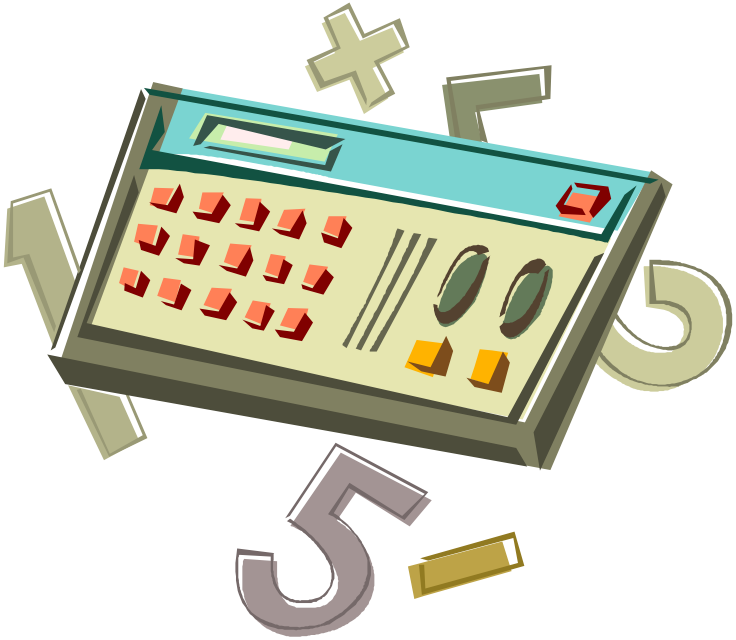
Örnek:

- **H0:** Erkek ve kadınların gelir ortalamaları arasında bir fark yoktur.
  - Testin verdiği değer: **P=0.001 ise:**
    - Bu kabul edilmiş **P=0.05** seviyesinin altındadır.
    - **0** Hipotezini destekleyecek ölçüde yeterli kanıt yok.
    - **0** Hipotezi dolayısıyla reddediliyor.
    - Binde bir olasılıkla **H0** hipotezi gerçekte doğrudur, ancak gösterilememiştir.

Örneklem büyüklüğü hesaplamaları

Kaç hasta olmalı??

50 hasta yeter mi???



Örneklem büyüklüğünün önemi

P değeri her şeyin çaresi mi?

"null" hipotezi (H0) reddedilir,
artık elimizde ilginç bir sonuç vardır.

– İlaç ve plasebo arasındaki karşılaştırmada $p<0.05$ ise, ilaç plasebodan farklıdır.

- Michael Jordan vs siz: 7/7 vs 4/7 basket
- $p=0.07$ (Fisher kesinleme testi)
  - **En az kaç atışta Michael Jordan farkını gösterebilir?**

Örneklem büyüklüğünün hesaplanması

Gerekli bileşenler:

1. Anlamlılık kriteri: Tip I hata **(α)**
2. Çalışmanın gücü: 1-Tip II hata **(β)**
3. Beklenen minimum etki **(olay sıklığı)** farkı
   a. İki grup arasında beklenen minimum etki farkı, veya
   b. Kontrol grubunda (veya popülasyonda) ve tedavi grubunda beklenen etki büyüklüğü / olay görülme sıklığı
4. Öngörülen ölçüm değişkenliği

Örneklem büyüklüğünün hesaplanması

- Gerçekten var olan bir farklılığın istatistiksel olarak anlamlı olarak saptanma olasılığı
- Bir çalışmanın var olan gerçek bir farkı saptayabilme gücü
- Gerçekte fark varken, H0'ı reddetme ve alternatif hipotezi doğrulama olasılığı, yani fark olduğunu ortaya koyma olasılığı
  - **β=0.20 ise, güç %80**

Hipotez testi

GÜÇ
testin, anlamlı farkı bulma olasılığı

		Gerçek durum	
		Fark var (H1)	Fark yok (H0)
Test sonucu	Fark var (H0 red)	**OK** **Güç = 1- β**	**Tip I hata (α)**
	Fark yok (H0 kabul)	**Tip II hata (β)**	**OK**

testin anlamlılık düzeyi

Örneklem büyüklüğünün önemi

- Örneklem, **gruplar arasındaki gerçek farkı gösterebilecek kadar büyük olmalı**
- Örneklem, **gereksiz maliyet ve etik olmayan uygulamaları önleyebilecek kadar küçük olmalı**

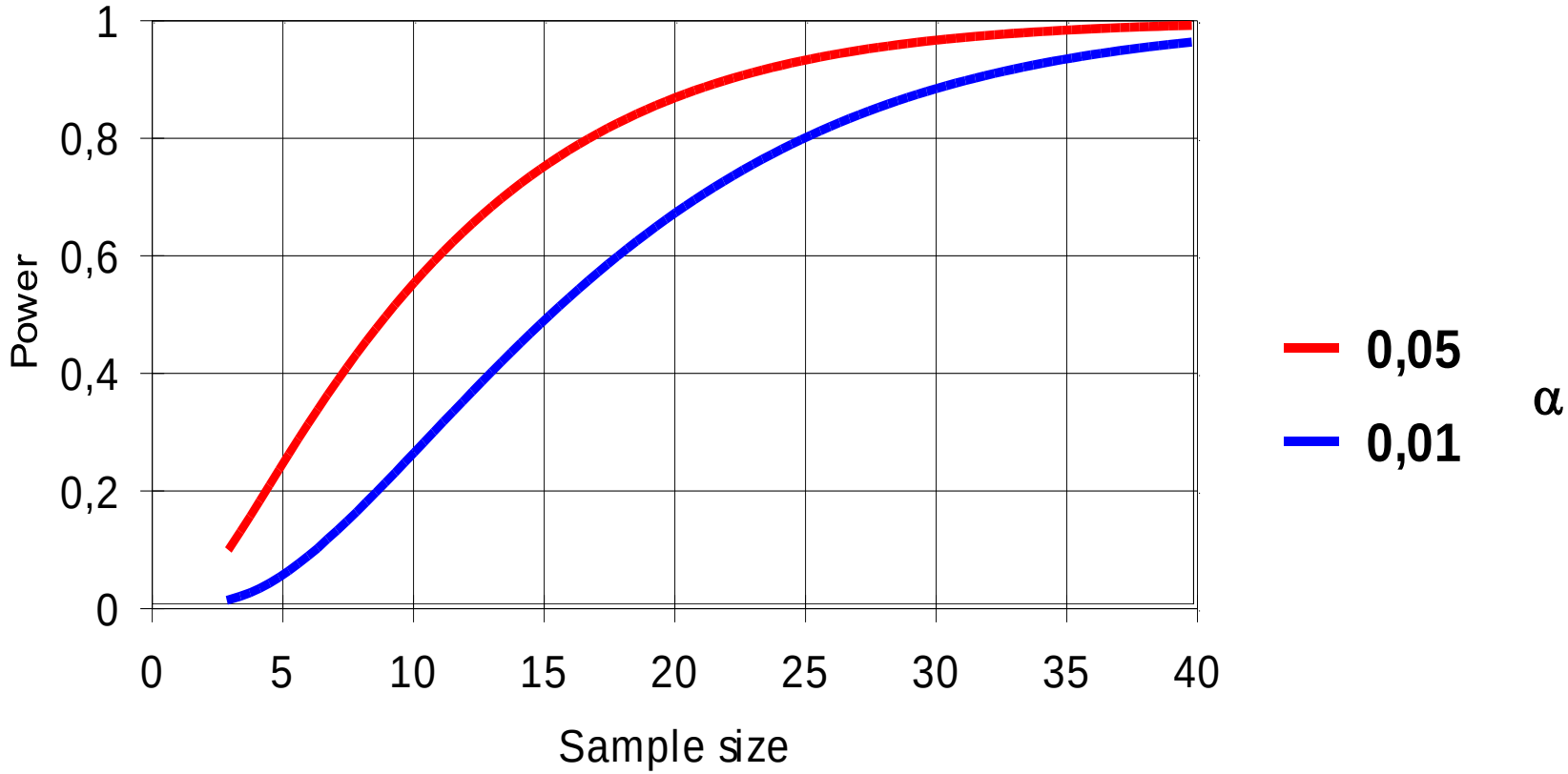
Örneklem büyüklüğünün hesaplanması

- **Bu durumda β büyür ve güç azalır**
  - Örneğin yeni bir ilacın yan etkisi araştırılırken α küçük olsun istenir.

- **Bu durumda β küçülür, α'nın büyümesinde sakınca yoktur.**
  - Örneğin, vitamin vs profilaktik antibiyotik

Örneklem büyüklüğü ve tip I hata

Güç, örneklem büyüklüğü ve tip I hata arasındaki ilişki



Beklenen minimum etki farkı

- Karşılaştırılan gruplar arasında ölçülen ve klinik açıdan önemli olabilecek en küçük fark
- Literatüre bakılarak veya klinik deneyimle belirlenebilir

Öngörülen ölçüm değişkenliği

- Her bir grupta yapılacak ölçümlerin beklenen **standart sapması**
- Ön sonuçlara, literatüre veya sübjektif deneyime göre belirlenir

Örneklem büyüklüğü

- $\alpha \downarrow$
- Güç $\uparrow$
- Minimum beklenen (etki / olay sıklığı) farkı $\downarrow$
- Ölçüm değişkenliği (SS) $\uparrow$

Örneklem büyüklüğü $\uparrow$

Grupsal parametreler nasıl tahmin edilecek?

- Tedavi grubunun parametrelerini tahmin etmeye çalışmıyoruz !
- “Saptamaya değer tedavi etkisi”ni tahmin ediyoruz.
- Genellikle kontrol grubuna dair veriler başlangıç noktası olur.

Yayınlardan örnek

Yayınlardan örnek:

"Güç=0.90 ve α = 0.05 kabul ederek her bir grup için 120 kişilik örneklem büyüklüğü hesapladık."

Çünkü, **beklenen minimum etki /olay sıklığı farkı tahmini** ve **değişkenliği** belirtilmemiştir.

Makale okurken “örneklem büyüklüğü”

Örneklem büyüklüğü hesabı belirtmemek çalışmanın naifliğine işaret eder. Bu durumda başka eksikliklere de *(bias)* bakılmalıdır;

- **Randomizasyon**
- **Körleme**
- **Karıştırıcı faktörler**

Örneklem hesabının yapılmamış olması büyük bir eksiklik değildir, çünkü örneklem büyüklüğü “**p**” değeri ve **güven aralıkları** ile anlaşılır.

İzlemden çıkanların sayısı ve bunun gruplar arasındaki farkına dikkat etmelidir.

Örneklem büyüklüğü hesabı

Rosuvastatin to Prevent Vascular Events in Men and Women with Elevated C-Reactive Protein

Paul M Ridker, M.D., Eleanor Danielson, M.I.A., Francisco A.H. Fonseca, M.D., Jacques Genest, M.D., Antonio M. Gotto, Jr., M.D., John J.P. Kastelein, M.D., Wolfgang Koenig, M.D., Peter Libby, M.D., Alberto J. Lorenzatti, M.D., Jean G. MacFadyen, B.A., Børge G. Nordestgaard, M.D., James Shepherd, M.D., James T. Willerson, M.D., and Robert J. Glynn, Sc.D., for the JUPITER Study Group*

STATISTICAL ANALYSIS

JUPITER was an event-driven trial designed to continue until 520 confirmed primary end points had been documented, to provide a statistical power of 90% to detect a 25% reduction in the rate of the primary end point, with a two-sided significance level of 0.05. Pretrial estimates of the duration of follow-up and number of participants were based on event rates in earlier prevention trials[19,20] and were modified to take into account plans to include low-risk groups, including women.

Yeterli örneklem büyüklüğü sağlanamazsa!

- **Çalışma tasarımı daha özenli yapılmalı**
  - Randomizasyona dikkat edilmeli
  - Karıştırıcı faktörlere dikkat etmeli
  - Güven aralıkları bildirilmeli
- **Çalışma sonuçları yine de yayınlanmalı**
  - Sonuçlar hiç değilse meta-analize dahil edilebilir.

Tek Değişkenli Analiz

Risk (Şans) Nedir?

Hastalanma riski

Hastaların ölüm riski

İstanbul' da deprem riski

Kumarda kaybetme riski

Maç sayısı

Analitik İstatistik

- Çalışma örnekleminden elde edilen değerlerden, popülasyon değerlerinin kestirilmesi ve bunun doğru olma olasılığının belirlenmesidir
- P-değeri hesaplanması (Hipotez testleri)
- Güven aralıkları

Sonuçların Yorumlanması

- İlişki var mı?
- İstatistiksel olarak anlamlı mı?
- İlişkinin yönü nasıl?
- Neden sonuç ilişkisi kurulabiliyor mu?
- Sonuçlar diğer çalışmalarla tutarlı mı?

İstatistik Test Seçimi

Parametrik teknikler

- Araştırma düzeni, deneklerin sayısı, verilerin dağılımı, verilerin yayılım ölçütleri gibi çeşitli özellliklerle ilgili koşullar sağlanabildiği takdirde uygulanabilecek testler parametrik testlerdir.
- Verilerin normal dağılıma uyması gerekir.

İstatistik Test Seçimi

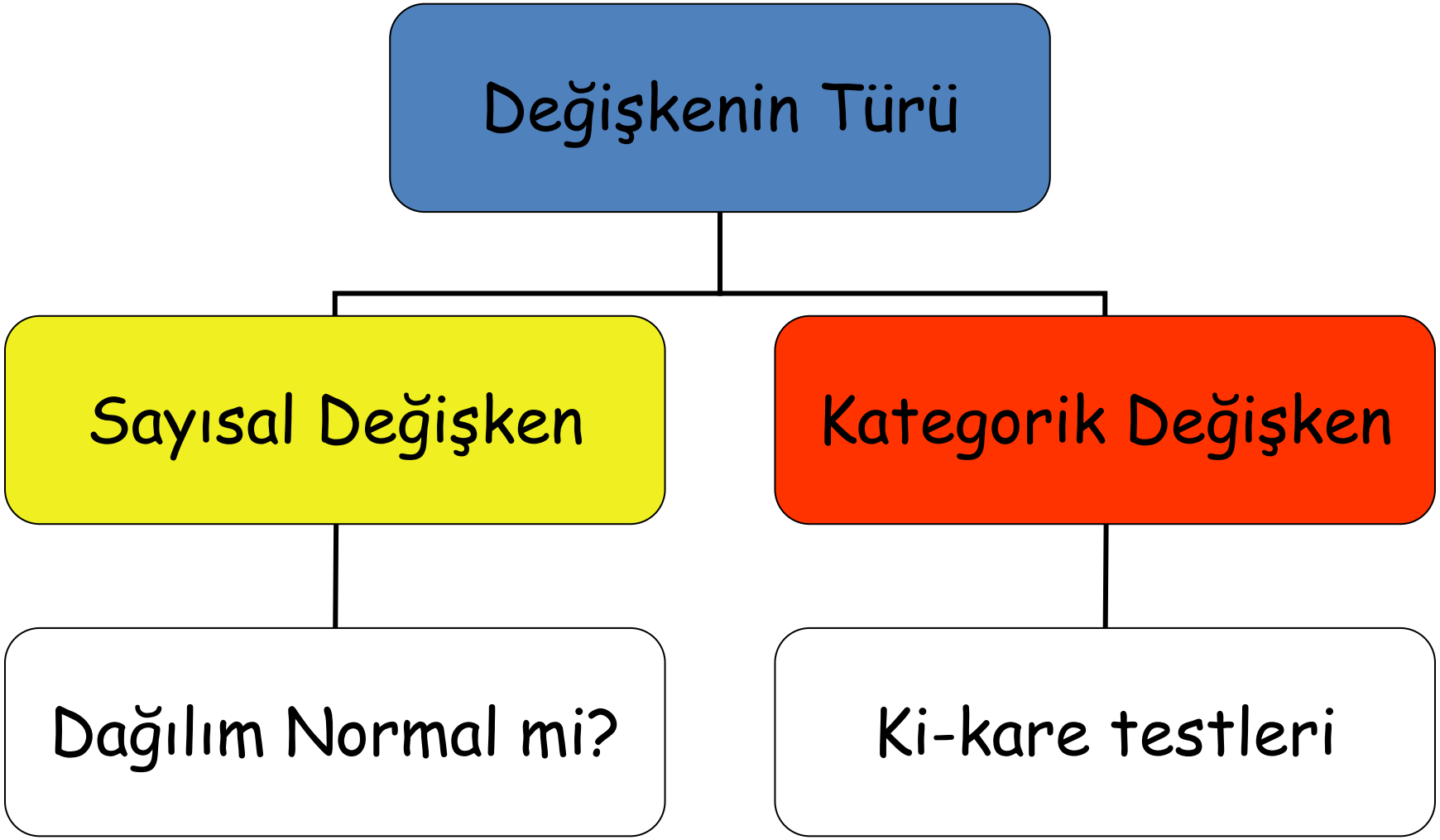
Nonparametrik teknikler

- Verilerin normal dağılım göstermediği, denek sayısının az olduğu ya da diğer parametrik test uygulayabilme koşullarından herhangi birinin sağlanamadığı durumlarda, parametrik testlerin eşdeğeri ya da alternatifi olan nonparametrik testler uygulanabilir.

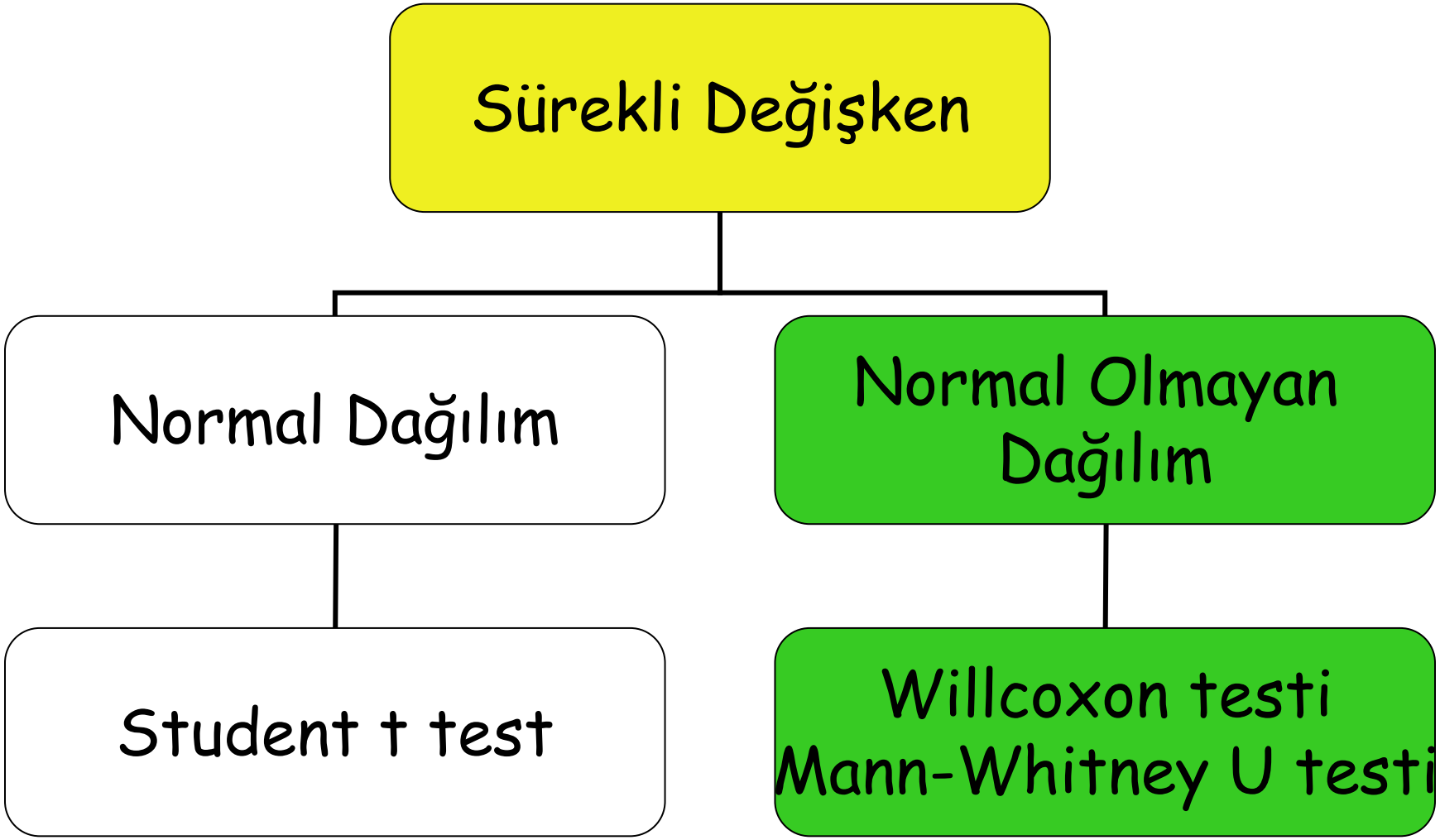
Sunumun Tabloları

Tablolar	Test	Örnek
Tablo 1 Tanımlayıcı istatistik	Ortalama, ortanca, standart sapma, vb.	Hastaların %51'i kadındır
Tablo 2 Farkları inceleyen analitik istatistik	İki grup karşılaştırması için t-testleri ve ki-kare (çapraz tablo) testleri	Ciddi seyirli hastalarda AST düzeyi daha yüksektir
Tablo 3 İlişki ve kestirim inceleyen analitik istatistik	Bağıntı analizi ve regresyon analizi	Çok değişkenli analiz sonucunda, AST düzeyi 300'ün üzerinde olanlarda olmayanlara göre ölüm oranı 3 kat fazladır (odds oranı)

Tek Değişkenli Analiz
iki grubun karşılaştırılması



Tek Değişkenli Analiz
iki grubun karşılaştırılması



Stata/SE 9.2

Results

0	1,654	183	1,837
1	442	79	521
Total	2,096	262	2,358

Pearson chi2(1) = 11.1177 Pr = 0.001

```
. tab  asisonrasigrip  asi, chi
```

asisonrasi grip	asi oldum=1 olmadim=0 0	1	Total
0	1,654	183	1,837
1	442	79	521
Total	2,096	262	2,358

Pearson chi2(1) = 11.1177 Pr = 0.001

.

C:\Documents and Settings\ONDER ERGONUL\Belgelerim\influenza

Başlat | 2 Microsof... | influenza | epidemiologi... | Stata/SE 9.2 | http://bus.... | multivariabl... | TR | 00:25

Etken - Sonuç İlişkisi Saptanmasında Temel Hücre

Dört gözlü tablo=2x2 tablosu

	Sonuç VAR	Sonuç YOK
Etken VAR	a	b
Etken YOK	c	d

Etken – Sonuç İlişkisi

	Şizofreni VAR	Şizofreni YOK
Çocukken SSF enf. VAR	a	b
YOK	c	d

İlişki *(association)*

Relatif risk (kohort)

a/a+b : c/c+d

Odds oranı (olgu-kontrol)

a/a+c : b/b+d

Relatif Risk, Risk Oranı

$$\frac{\text{Maruz kalan bireylerin riski}}{\text{Maruz kalmayan bireylerin riski}}$$

Odds Oranı

Olgular için odds= $\dfrac{\text{Maruz kalan olguların sayısı}}{\text{Maruz kalmayan olguların sayısı}}$

Kontrol için odds= $\dfrac{\text{Maruz kalan kontrollerin sayısı}}{\text{Maruz kalmayan kontrollerin sayısı}}$

Odds oranı= $\dfrac{\text{Olgular için odds}}{\text{Kontroller için odds}}$

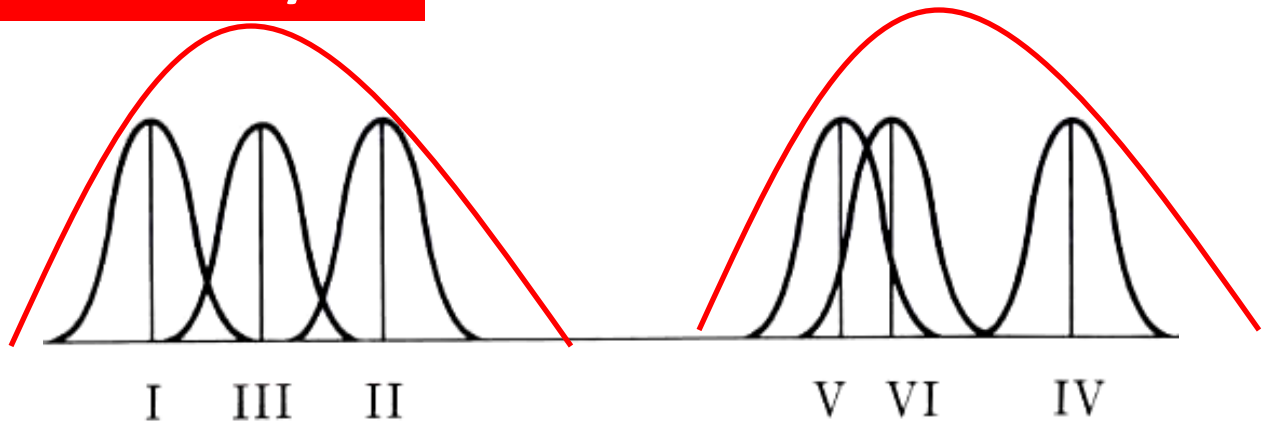
Odds ve Odds Oranı Nedir?

- Bir ilacın yan etki gelişme oranı 1/100 ise, odds=1/99= 0.0101
- Doğumda erkek çocuk olma ihtimali ½, odds=1/1=1
- İki odds' un birbirlerine oranı=odds oranı

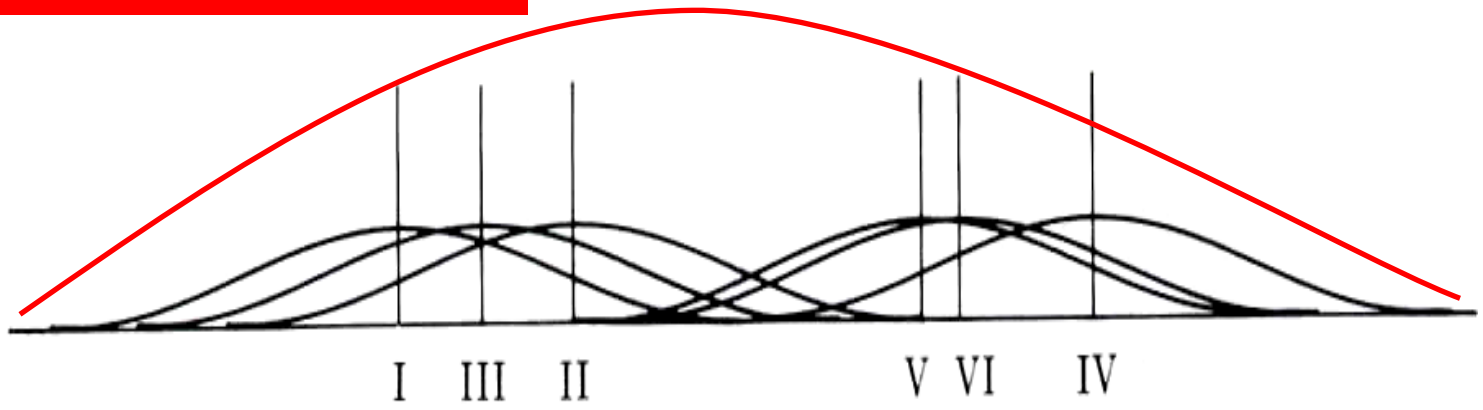
ANOVA

(ikiden çok sayıda grubun karşılaştırılması)

Gruplararası varyans



Grup içinde varyans

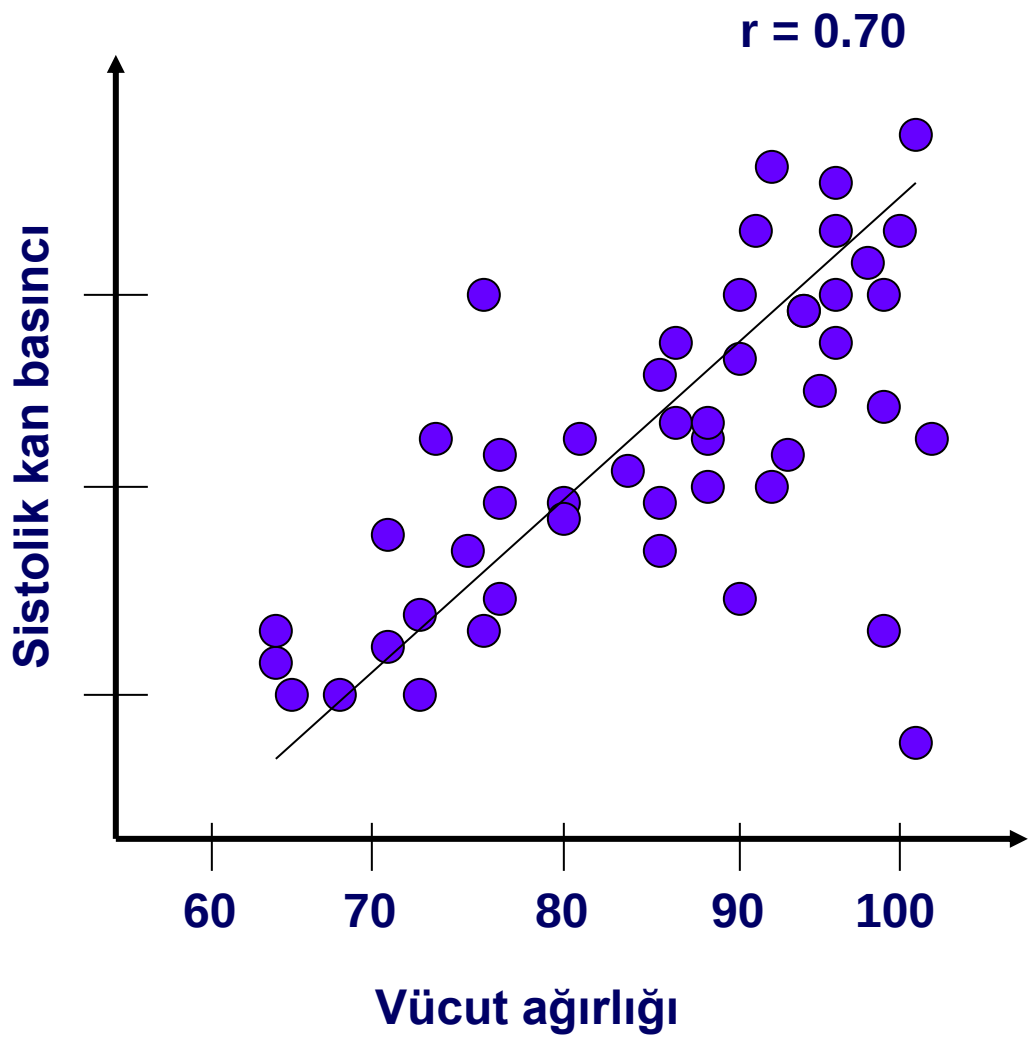


ANOVA

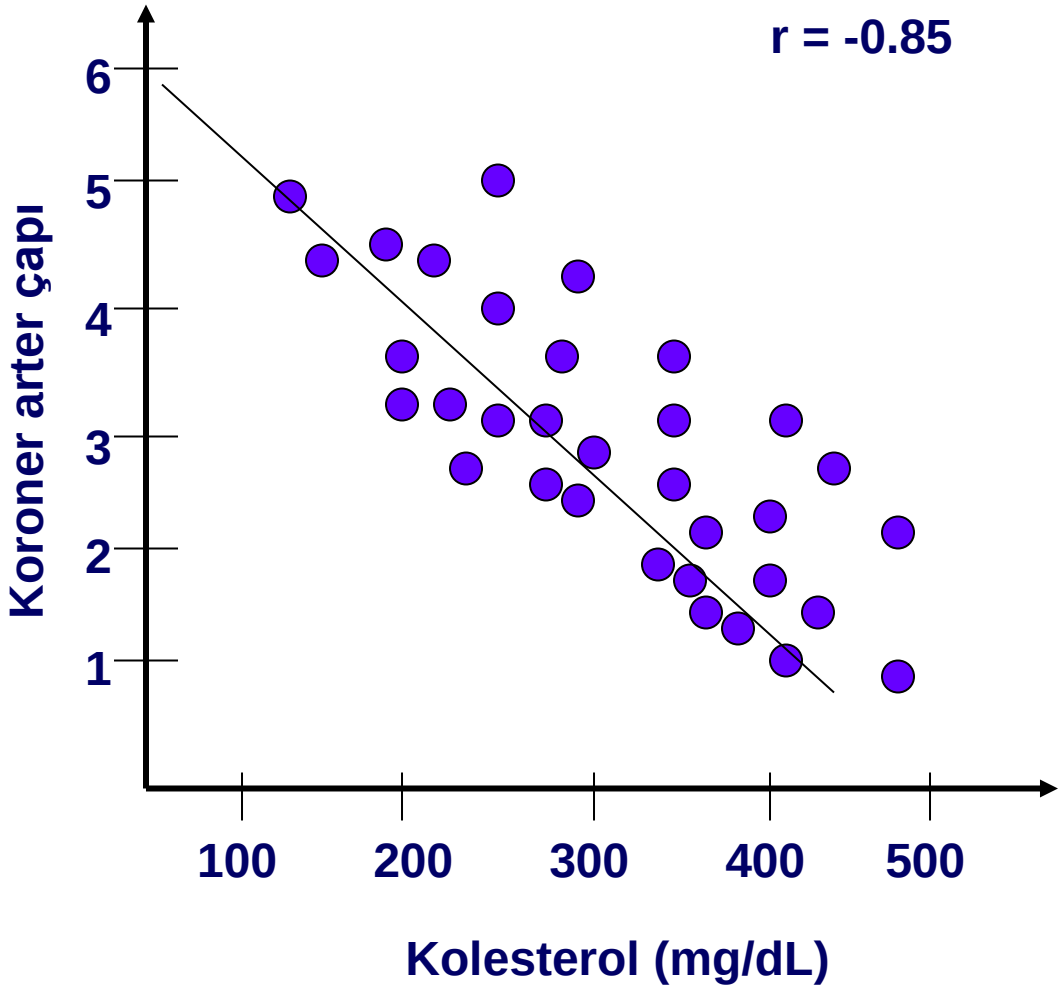
- Birden çok grup varsa
- F testi
- ANOVA, çok değişkenli analiz testi değildir

$$F = \frac{\text{Between Group Variability}}{\text{Within Group Variability}}$$

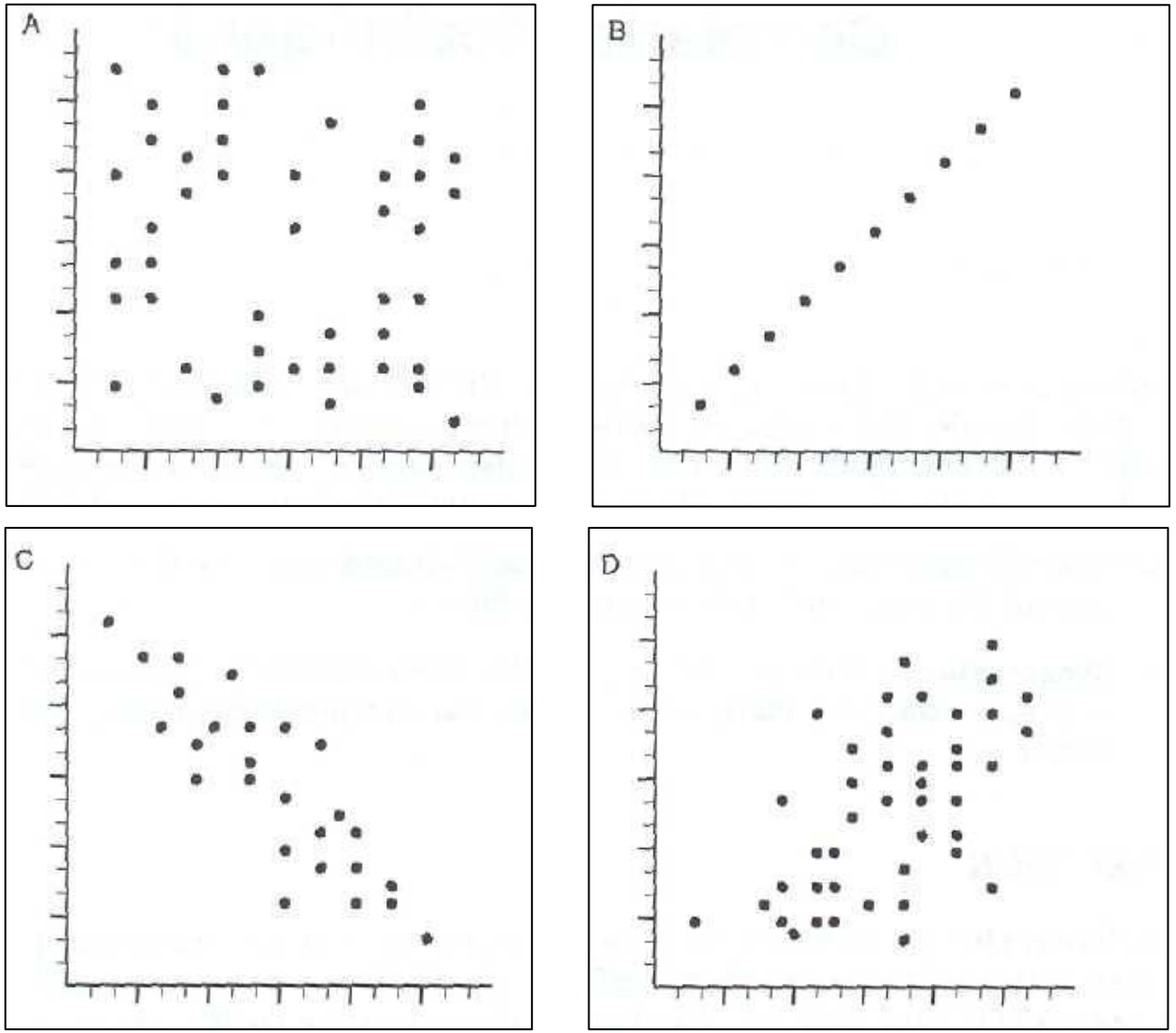
Bağıntı analizi: *Korelasyon*



Bağıntı analizi: *Korelasyon*

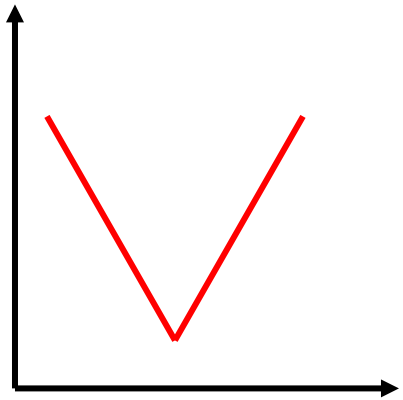


Bağıntı analizi: *Korelasyon*

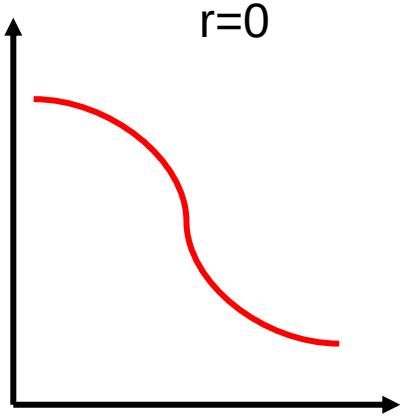


Bağıntı analizi: *Korelasyon*

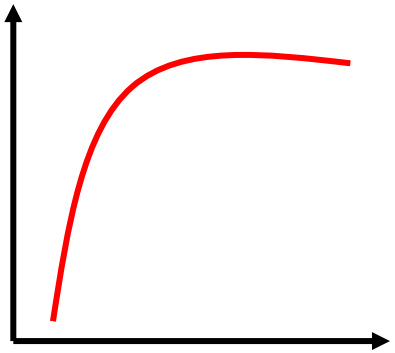
r=0



r=0



r=0



Bağıntı analizi: *Korelasyon*

- **Korelasyon kat sayısı -1 ve +1 arasında değişir**
  - 0-0.25 Hiç ilişki yok/zayıf ilişki
  - 0.25-0.50 Zayıf – orta derece ilişki
  - 0.50-0.75 İyi derece ilişki
  - 0.75-1.00 Çok iyi derecede ilişki
- **Spearman vs Pearson KK**
- **P değeri !**
- **Aşırı değerler problemi**

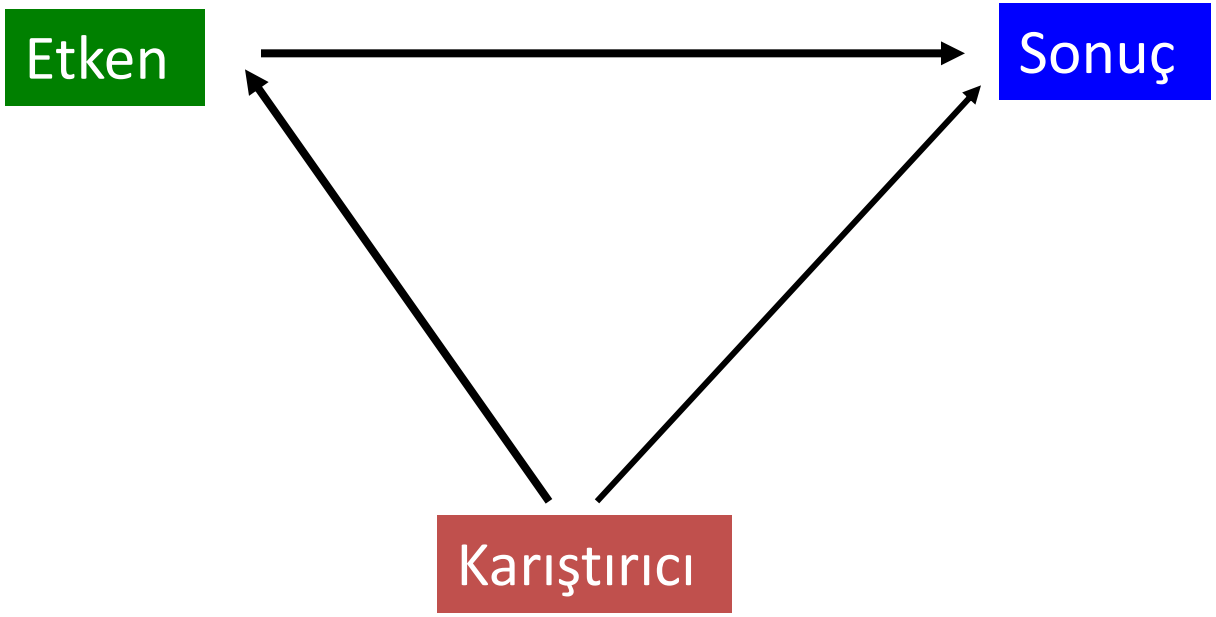
Örnek

Table 5: Correlations between changes in blood pressure and changes in metabolic parameters

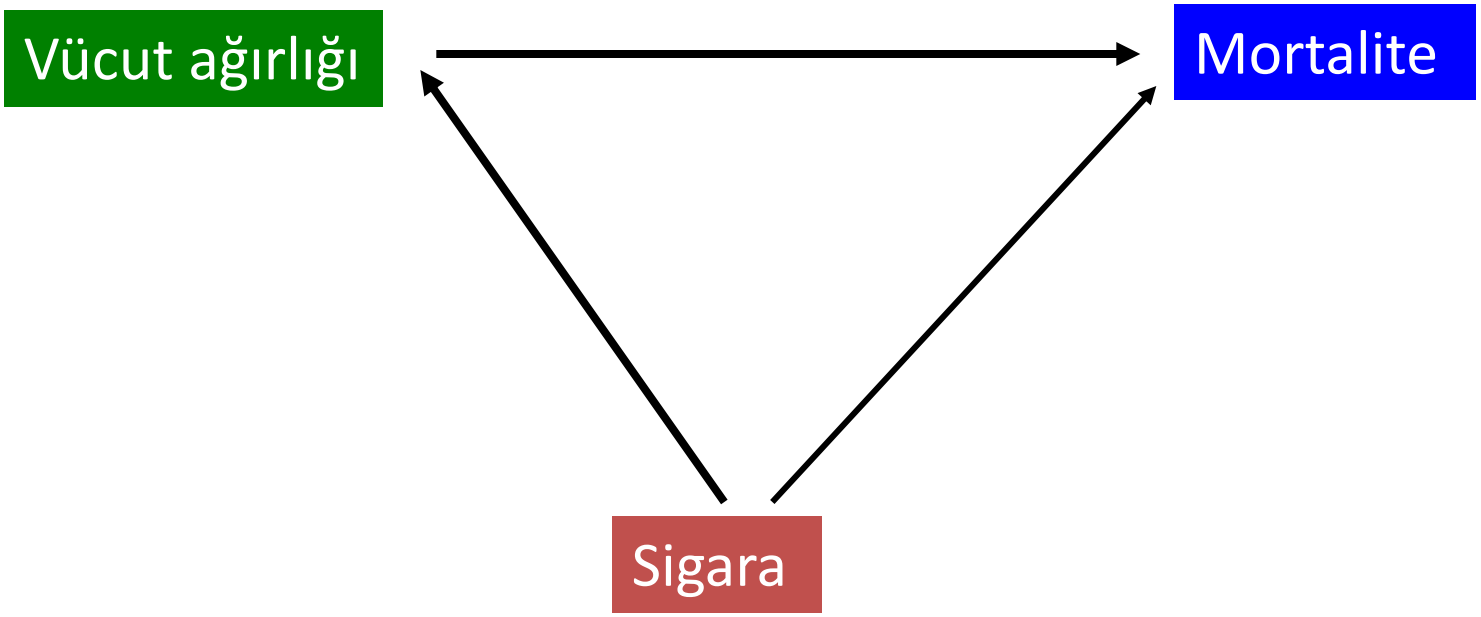
Parameter	Systolic blood pressure		Diastolic blood pressure	
	r	p	r	p
HbA1c [%]	0.114	<0.0001	0.064	0.002
Glucose [mg/dl]	0.078	<0.0001	0.045	0.016
LDL-cholesterol [mg/dL]	0.097	<0.0001	0.132	<0.0001
HDL-cholesterol [mg/dL]	-0.025	ns	-0.064	0.001
Triglycerides [mg/dl]	0.085	<0.0001	0.115	<0.0001

Karıştırıcı Etken

- Baktığınız etkenle sonuç arasında, başka bir etkenden dolayı bir ilişki varmış gibi sonuç çıkması



Karıştırıcı Etken



Karıştırıcı Etkenleri Kontrol Etmek İçin

- Randomizasyon
- Tabakalama
- Düzeltme (adjustment)
  - Standardizasyon
  - Lojistik regresyon
  - Çoklu doğrusal
  - Regresyon

Tabakalama yapılması gereken birden çok değişken varsa,

ÇOK DEĞİŞKENLİ ANALİZ

yapılır.

Çok Değişkenli Analiz Yöntemleri

- Çoklu Regresyon *(multiple/multivariate regression)*
  - bağımlı değişken – sürekli
  - bağımsız değişken – sürekli / kesikli
- Lojistik Regresyon
  - bağımlı değişken – dikotom
  - bağımsız değişken – sürekli / kesikli
- Cox Regresyon
  - bağımlı değişken – dikotom
  - bağımsız değişken - sürekli / kesikli
  - zaman değişkeni

- Kolesterol düzeyini belirleyen değişkenler hangileridir ?

  - Yaş
  - Cinsiyet
  - BMI

- Her değişken için tek tek lineer regresyon analizi yapılabilir.

Lojistik Regresyon

- Birden fazla X bağımsız değişkeninin Y bağımlı değişkeni ile ilişkisini göstermek için kullanılan matematiksel model yaklaşımıdır.

Lojistik Regresyon

- Kohort çalışmalarda kullanmak için geliştirilmiştir.
- Olgu – kontrollerde ve kesitsel çalışmalarda da sık kullanılır.
- Olgu- kontrolden OR, kohorttan RR elde edilir.
- Olgu – kontrolde örneklem oranı bilinmediği için ά bilinemez, onun için bireysel risk hesaplanamaz.

Çok Değişkenli Analiz

Regresyon türü	Bağımlı değişken (outcome, sonuç)
Linear	Sürekli
Logistic	Dikotom
Cox	Dikotom
Poisson	Nadir

Bilmemiz Gerekenler

1. Model oluşturulması
2. Değişkenlerin seçimi
3. Modelin uygunluğu
4. Sonuçların yorumu

Model Kurma Stratejileri

- Modele alınacak değişkenlerin belirlenmesi
- Tekli analizlerde önemli olan değişkenler
- Literatürde önemli olan değişkenler
- Birbiriyle yüksek korelasyonu olan değişkenler alınmamalı
- Bir değişken için 10 olgu

Değişkenlerin Modele Hazırlanması

- Kategorik değişkenlerin frekansı alınır ve kategorilere dağılımına bakılır.
- Kategorilerin dengeli dağılımı olması sağlanır.
- Referans kategorisi belirlenir.
- Dummy değişken kavramının anlaşılması.

Değişkenlerin Modele Hazırlanması

- Bağımlı değişkeni 1 ve 0 olarak kodlamak gerekli
  - 1: hasta 0: sağlam
- Dikotom değişkenleri de 1 ve 0 olarak kodlamak gerekli
  - 1: risk var 0: risk yok

Meslek grubu	Nominal	Dummy Çiftçi	Dummy Hemşire
Çiftçi	1	1	0
Ev hanımı	2	0	0
Hekim	3	0	0
Hemşire	4	0	1
Polis	5	0	0

Model Oluşturmak

İki bağımsız değişken > 0.90 korele ise (multicollinearite), birini tercih et.

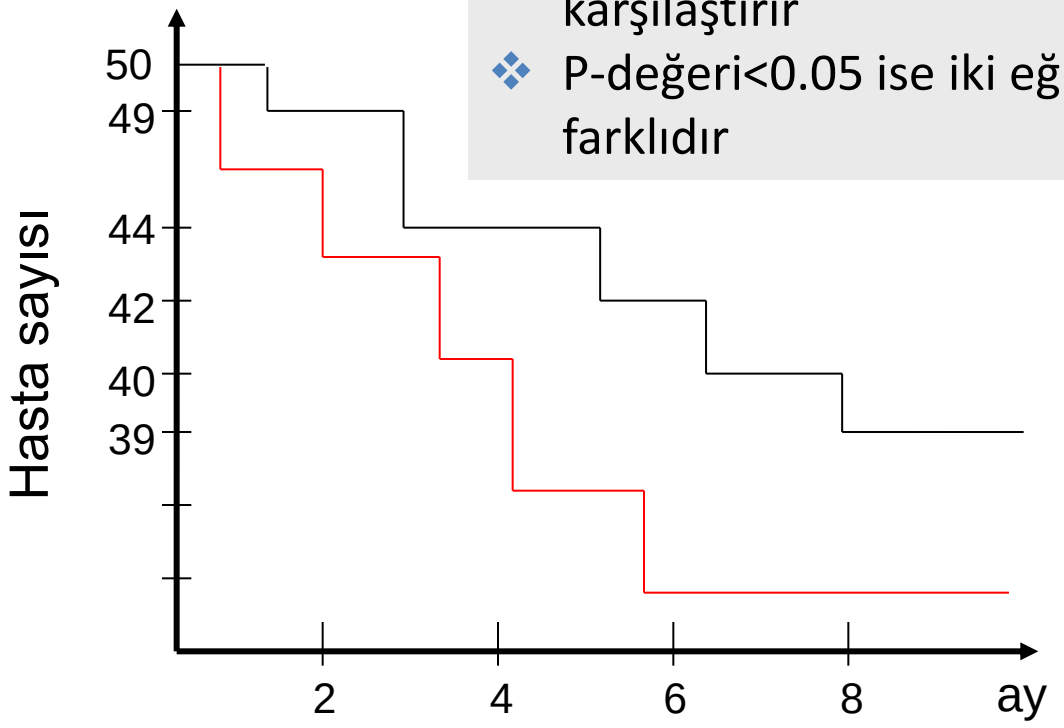
Daha pratik bir yol, “biyolojik ilişki” (sigara, kibrit taşıma ve kanser)

Sağkalım Analizi

Kaplan-Meier Grafikleri

- Sonuç değişkeni = Bir "olaya" kadar geçen süre *(time until an event occurs)*
- "Olay" = ölüm, iyileşme, relaps

- Log-rank testi iki eğriyi istatistiksel olarak karşılaştırır
- P-değeri<0.05 ise iki eğri istatiksel olarak farklıdır



Yazmak için öneriler

1. Yazmaya erken aşamada başlamak
2. Günlük iş listesi yapmak
3. Hergün bir şey yazmak
4. ısınma süresini azaltmak
5. Uygun bir ortam hazırlamak
6. Çok yazdıktan sonra kısaltabilmek
7. Uygun referans programları kullanmak

Yazar Blokunun Farkına Var, Üzerine Git

Kaygılanma. Daha önce hep yazdın, şimdi de yazacaksın. Tüm yapman gereken gerçek bir cümle yazmak. Bildiğin en hakiki cümleyi yaz.

Ernest Hemingway

Hepimiz, kimsenin asla usta olamayacağı bir zanaatta çırağız

Ernest Hemingway

Orhan Pamuk, 25.05.2014, Hürriyet, Çınar Oskay ile söyleşi

“Depresyon zamanları, romanınızdan memnun olmadığınız zamanlar da olur. Şu an memnunum, çünkü gül bahçesindeki fazlalıkları takır tukur kesip atıyorum. Bu en zorudur. Çünkü attığınız her kısım hayatınızın 15 günüdür, bir ayıdır. Roman onu mükemmel yapma azmi ve korkusuyla başlar.

40 yıldır roman yazıyorum. Hala ilk kitabını yayınlayan romancı gibi titriyorum. Ustalık diye bir şey yok. Her seferinde aynı korku... Yıllarca çalışarak kazandığınız okurlarınız vardır ama romanınız kötü olursa bir sonrakini almazlar.”

Her ne yapabiliyor ya da yapabileceğini hayal ediyorsan yapmaya başla. Cürette deha, güç ve büyü vardır

Johann Wolfgang Von Goethe

Gerçekten de hiçbir şey bu kadar eğlenceli değildi. Tam da bu nedenle yazdın. Bunu daha önce fark etmemiştin. Bilinçli yaptığın bir şey değildi. En büyük zevki orada bulduğundandı. Hiç bir şey bu kadar sürükleyici değildi.

Ernest Hemingway

Süresini geciktirenler için teselli

Sevgili Herr Doktor,

Bizim için yazmayı kabul ettiğiniz Das Kapital'in tesliminde daha şimdiden on ay geciktiniz. El yazmasını önümüzdeki altı ay içinde alamazsak bu işi başkasına vermek zorunda kalacağız.

Leipzig'deki yayıncısından Karl Marx'a mektup