



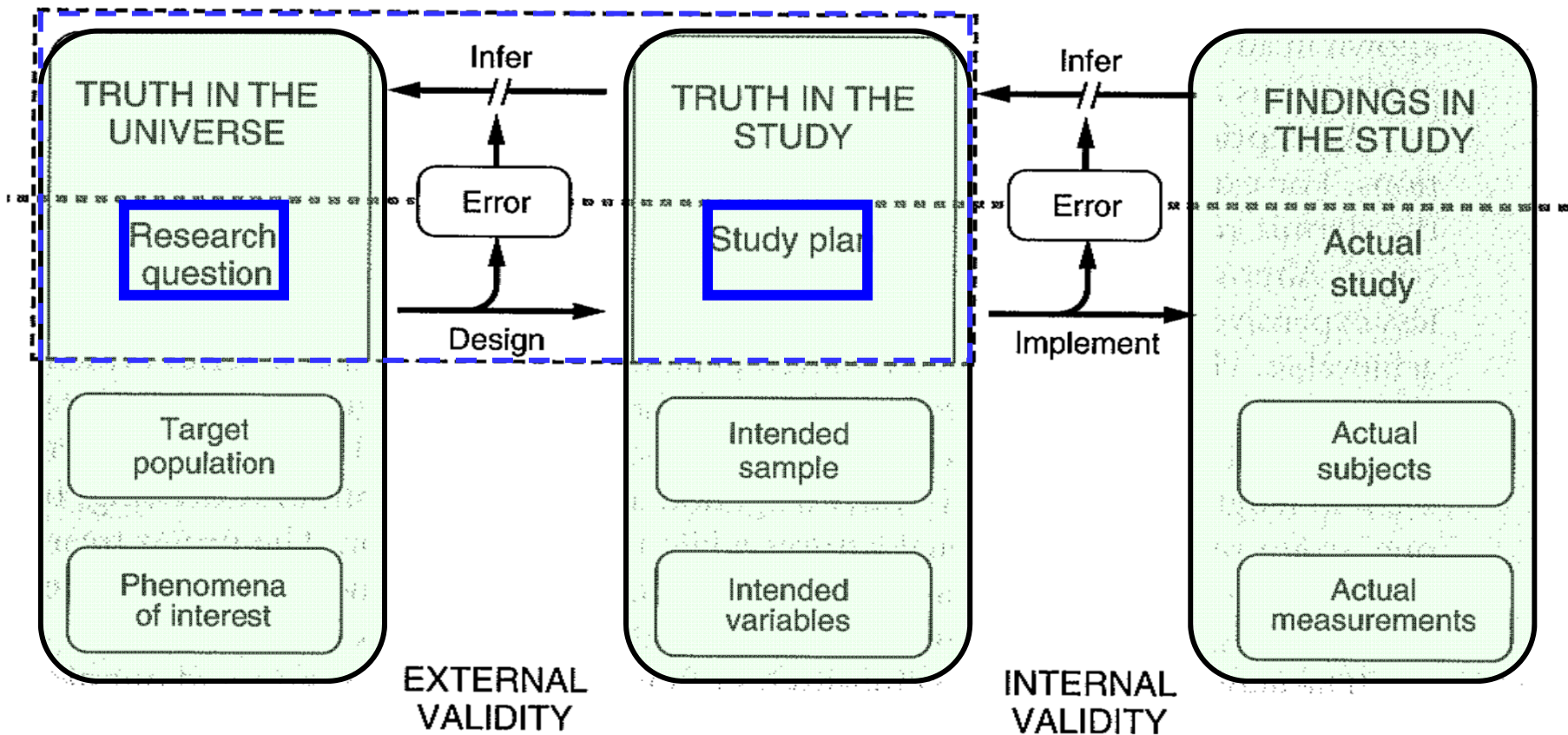
# Örneklem Seçimi ve Büyüklüğünün Hesaplanması

Türk Toraks Derneği  
Akademik Kurs, 10-11 Ekim 2015, İstanbul  
Ahmet Uğur Demir

# Sununun Amacı

- Bu sunuda aktarılanlar örneklem büyüklüğü gibi başlı başına bir disiplini anlatmak için yeterli değildir
- Bu sununun amaçları:
- Araştırma sonuçlarını değerlendirirken, okuduğunuz makaleyi yorumlarken, araştırma planlarken
  - Örneklem büyüklüğünün ne anlama geldiğini;
  - Örneklem büyüklüğü hesaplanırken –tahmin edilirken- dikkat edilmesi gereken temel noktaları;
  - Örneklem hesabı ile ilgili örneklerle birlikte özetlemektir

# Araştırmanın Anatomisi



# Arařtırmanın Anatomisi

- Evrendeki gereklik (hakikat, evren, hedef toplum)  
↓
- alıřmadaki gereklik (alıřma/arařtırma toplumu)  
↓
- alıřma bulguları (Örnekleme, ulařılan alıřmaya katılmayı kabul edenler)

# Örnek

- Araştırma Sorusu: Kronik obstrüktif akciğer hastalarında yaşam kalitesini etkileyen faktörler
- Hedef toplum: Kronik obstrüktif akciğer hastaları (**Tüm Türkiye**)
- Araştırma toplumu: **İstanbul'da yaşayan Kronik obstrüktif akciğer hastaları**
- Örneklem: **İstanbul'da eğitim araştırma hastanelerine son bir yılda başvuran Kronik obstrüktif akciğer hastalardan seçilen 1200 hasta**

# Neden Örneklem?

- Tüm toplumu arařtırmak zahmetli, çoğunlukla olanaksız
- Toplumunu temsil eden örneklem toplum hakkında geçerli bilgi verebilir
- Daha ucuz, kolay, uygulanabilir
- Toplumunu temsil etme özelliđi: *external validity* (genellenebilirlik, dışsal geçerlilik)
- Arařtırma toplumundaki gerçeklik: *internal validity* (dođruyu temsil etme, içsel geçerlilik, sonuçların güvenilirliđi)

# Arařtırma Toplumunu

Örneklem hedef toplumu temsil etmeli

- Her zaman mümkün deęilse de hatanın yönünü bilmek önemli (üçüncü basamaktaki bir hastaneden alınan örneklem toplumdaki genele göre daha ağır, kötü prognoza sahip hastaları temsil eder)

# Örnekleme Şekilleri

- Ardışık (Consecutive) Örnekleme (örneğin ikinci, beşinci hasta kuralı ile hasta alma... kural bilineceği için örnekleme/çalışmaya alınan kişiler tahmin edilebilir, randomizasyon kırılabilir)
- Olasılık (Probability) Örneklemleri: Basit rastgele örnekleme (tüm toplum üyeleri örnekleme alınmak için eşit şansa sahip, en ideal seçim)
- Yığınsal randomize seçim: mahalleler (bloklar) içinde hane ziyareti ile seçim (seçilen mahallede, ziyaret edilen hanede olmak örneklemede olmayı etkiliyor)
- Uygun (Convenience) Örnekleme: en pratik, kolay, ancak hataya en açık olanı (kongre katılımcıları arasında salonda bulunanlara uygulanan anket... geneli temsil etme olasılığı çok düşük)



# Örnekleme büyüklüğü

- Hesaplama değil tahmin etme
- Hipotezi test etmek için yeterli sayı... ???
- İstatistiksel (şansa bağlı) değişkenlik, dağılım
- Öngörülen/test edilen etki/ilişki
- Tahmin edilen yanılma payı: aralık
- Araştırmanın bütçesi, yapılabilirliği

# Örneklem büyüklüğü

- Araştırmanın sonuçta hipotezi kanıtlamak/sınamak için yeterli güce sahip olmaması: boşa harcama, boşa zaman
- Araştırmaya “gerekenden” fazla kişi, katılımcı alınması: maliyet, zaman kaybı, kişilerin sonucu iyi bilinmeyen madde/ilaç/tedavi vb ile karşılaştırılmaları, risk almaları

# İstatistiksel Güç

- Sıfır hipotezini doğru şekilde reddetme olasılığı
- Sıfır hipotez ( $H_0$ ): araştırmadaki grupların ölçülen etki göstergeleri (ortalama, sıklık, odds ratio) istatistiksel olarak farklı değil (iki grup arasında fark yok)
- Alternatif hipotez ( $H_A$ ): iki grup arasında fark var
- Güç için genel kabul: %80
- Daha kesin karar vermek için: %90

# P değeri

- Sıfır hipotezi doğru kabul edildiğinde bu araştırmada iki grup arasında gözlenen fark ne oranda şansa bağlı
- Aynı toplumda yapılan ölçümler arasında “...” kadar farklılık bulunabilir (istatistiksel dağılım)

Ne oranda/sıklıkla/olasılıkla?

**P** oranında/sıklığında/olasılığında

- Güven aralığı (sıklıkla %95): bu (gerçek) fark %95 olasılıkla [...- ...] aralığındadır

# Alfa ve beta

**Gerçek durum**

**Test sonucu**

	$H_0$ doğru	$H_0$ yanlış (alternatif hipotez $H_A$ geçerli)
$H_0$ geçerli (kabul)	1-alfa	Beta (tip II hata)
$H_0$ geçerli değil (red)	Alfa (tip I hata) ( $\alpha$ )	İstatistiksel güç (1-beta)

# Dağılım (Sayısal değişken)

Dağılımın merkezi: central tendency

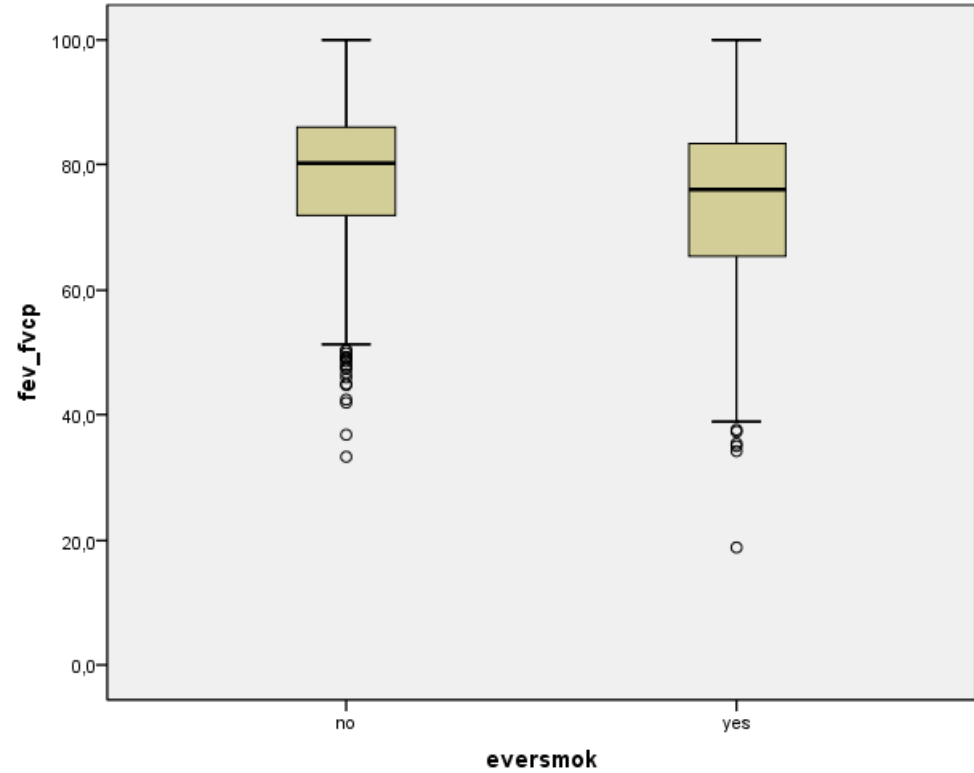
- Ortalama:  $(\sum x)/n$  (1,2,3,7,7,7,7,5,8 dizisinde 5.25)
- Medyan (ortanca): küçükten büyüğe sıralandığında ortadaki değer
- Mod: en sık tekrarlanan değer (1,2,3,7,7,7,7,5,8 dizisinde 7)

Dağılım (değişkenlik)

- Standart deviasyon (sapma):  $\sqrt{\sum (x - \bar{x})^2/(n-1)}$
- Aralık: en küçük-en büyük değerler
- Çeyrekler arası dilim (interquartile range): küçükten büyüğe sıralandığında ilk %25- son %25 (%75) değerler

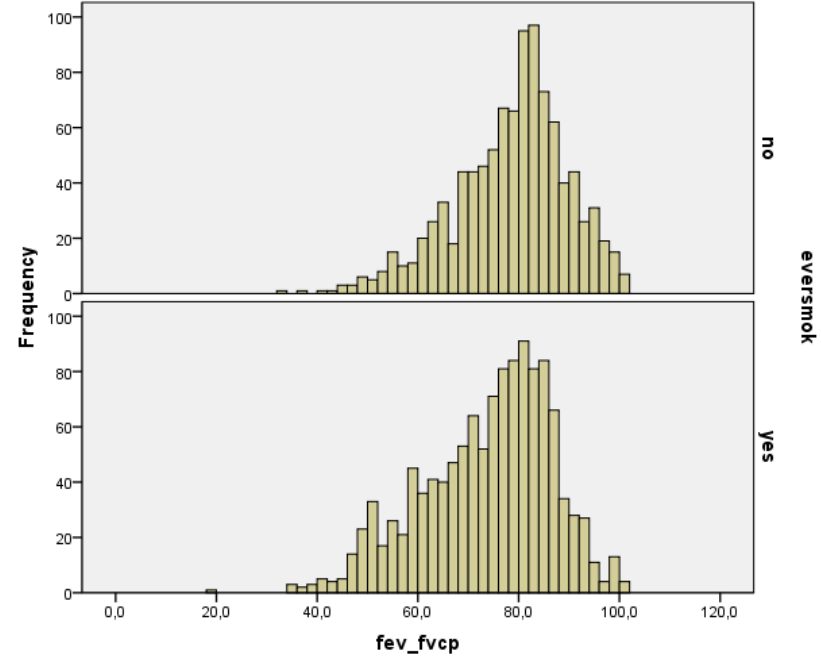
# Boxplot (kutu) Grafiđi

- Sigara içme durumuna göre FEV1/FVC oranı (%) dağılımı
- İçmemiş: 990, içmiş: 1214
- Orta çizgi: medyan
- Kenarlar: %25-%75 değerler
- Kuyruklar: %5-%95 değerler
- Halkalar: dışta kalanlar (outliers)



# Ortalama Standart Sapma Örnek

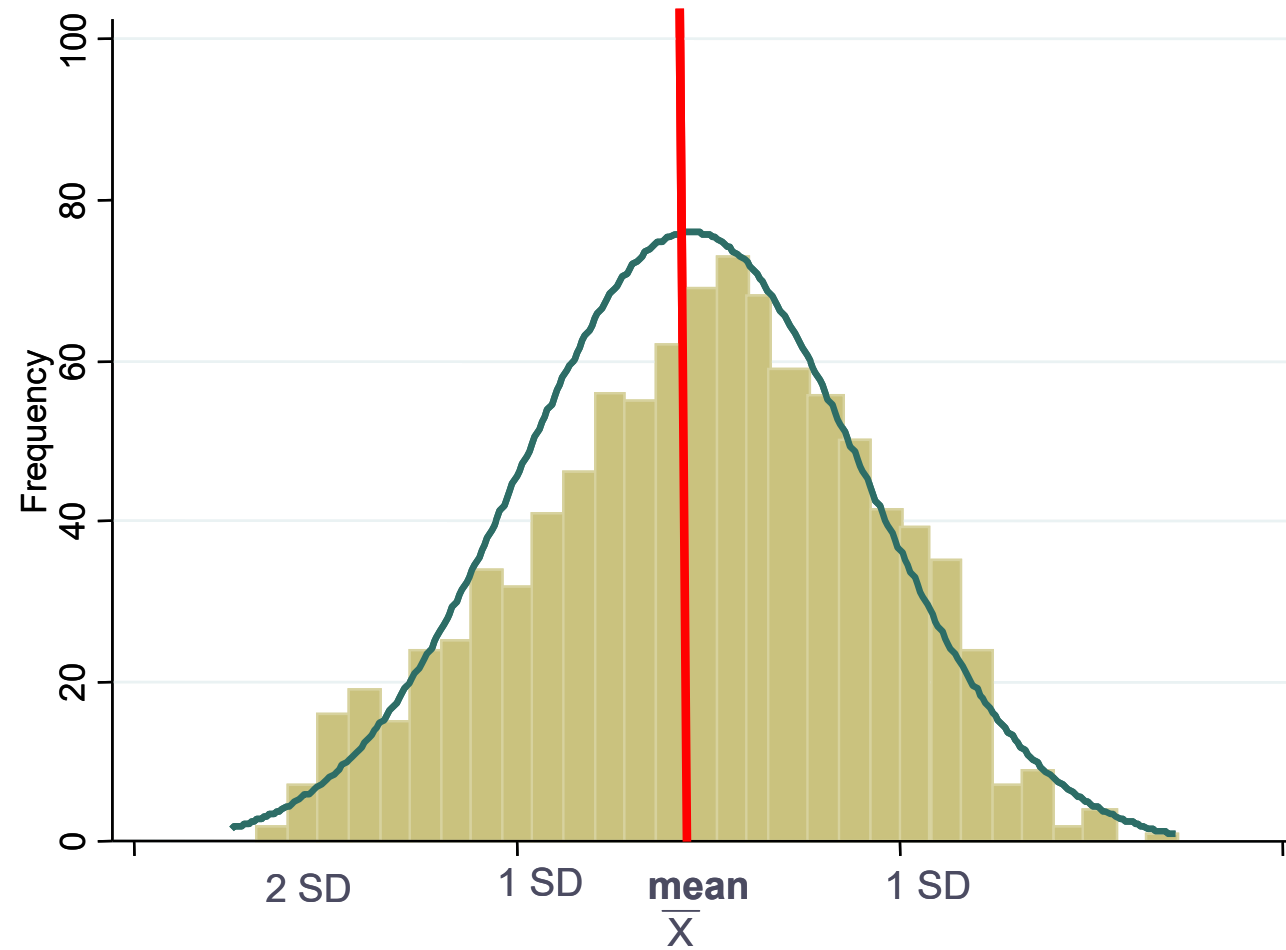
- Sigara içme durumuna göre FEV1/FVC oranı (%) dağılımı
- Karşılaştırma: bağımsız gruplar t testi,  $p < 0.001$



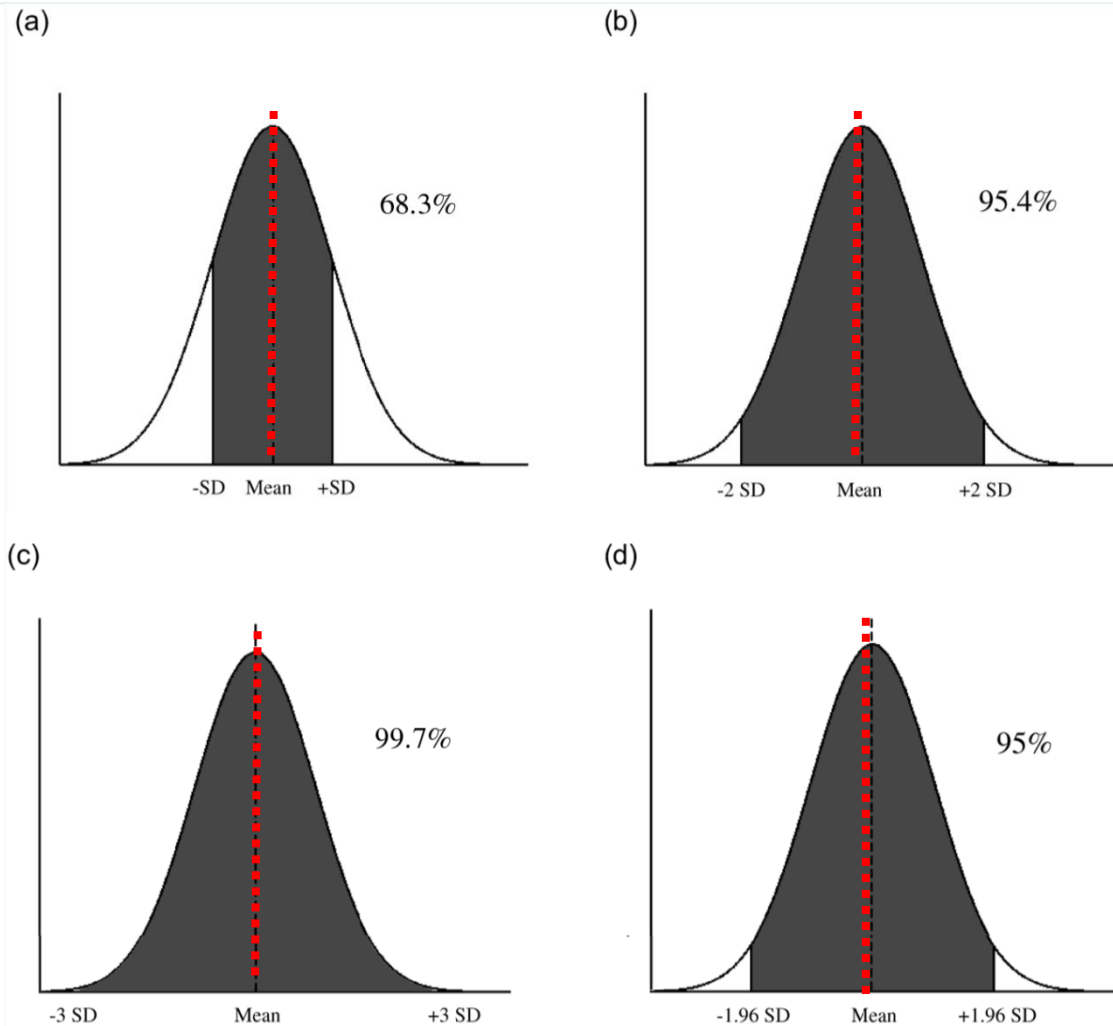
Sigara	FEV1/FVC (%) ortalama	Standart sapma (SD)	Standart hata (SE)
Hiç içmemiş	74.8	11.19	0.35
İçmiş	73.6	12.82	0.36



# Normal (Gauss) Dağılım



a: 1 SD, b: 1.96 SD, c: 2 SD, d: 3 SD

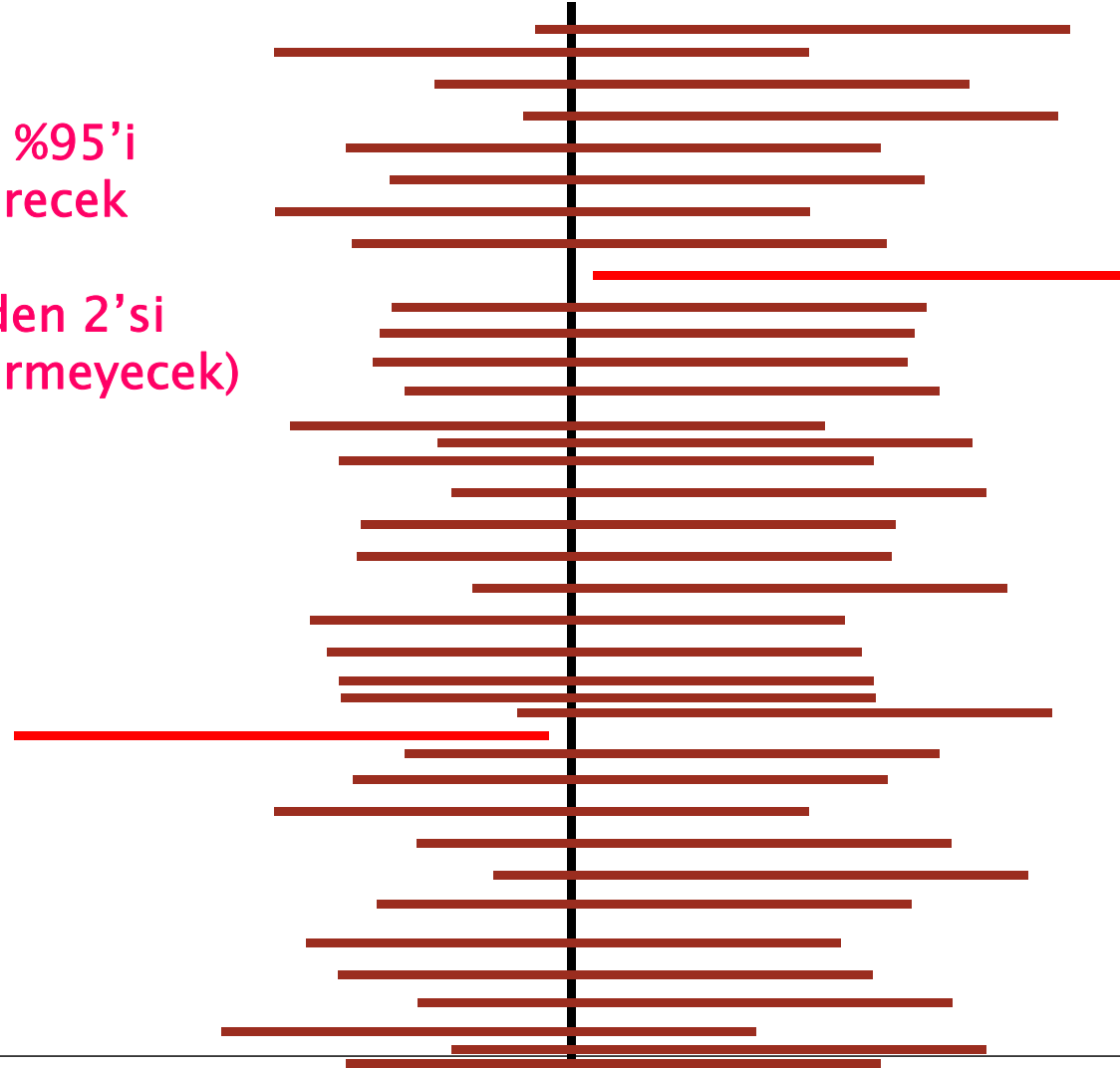


# Örneklemlerin Dağılımı

Toplumun ortalaması

Örneklemlerin %95'i  
Ortalamayı içerecek

(40 örneklemden 2'si  
Ortalamayı içermeyecek)



# %95 Güven Aralığı

- %95 Güven Aralığı (GA) = Ortalama  $\pm \sqrt{SE}$   
= Ortalama  $\pm SD/\sqrt{n}$
- Örnek: plazma sodyumu (Na): 138 (133, 143)
- N artarsa: GA nasıl değişir?

- Örneklemin ortalama değeri: genel ortalamaya ne kadar yakın?
- Ne kadar büyük örneklem alınırsa genele o kadar yakın olması beklenir
- Örneklem ortalaması da normal dağılım gösterir
- Örneklem ortalaması dağılımının standart sapması: standart hata:  $SD/\sqrt{n}$

# Örneklem büyüklüğünü etkileyen faktörler

	Değer	Yorum	Örneklem Büyüklüğü
P (alfa)	↓	Katı ölçüt, istatistiksel anlamı bulmak zor (H0 reddetmek daha zor)	↑
	↑	Gevşek ölçüt, istatistiksel anlamı bulmak daha kolay (H0 reddetmek daha kolay)	↓
Güç	↓	Farkı göstermek oldukça zor	↓
	↑	Farkı göstermek daha kolay	↑
Etki /fark	↓	Etki varsa göstermek zor	↑
	↑	Etki varsa göstermek daha kolay	↓
Alternatif Hipotez	Tek taraflı	Farkı göstermek daha kolay (çift taraflı olan alfa artıyor)	↓
	Çift taraflı	Genel kabul, Farkı göstermek daha zor	↑

# Tasarımın örneklem büyüklüğüne etkisi

- Randomize klinik çalışmalar: aynı etkiyi göstermek gözlemsel çalışmalardan daha ↓
- Karıştırıcı etkenleri kontrol etmek için: %20 ↑
- Pre post: kontrol grubu ile olan araştırmanın yarısı kadar
- Cross-over: kontrol grubu ile olan araştırmanın ¼'ü kadar
- Tek taraflı test: %20 ↓

# Sonuçlar

Tür	Örnek	Ölçüt
Kategorik	Evet/hayır, var/yok Örneğin: astım, obezite	Sıklık, oran, yüzde
Kategorik	Fazla sayıda kategori Örneğin: ağırlık dereceleri (hafif, orta, ağır),	Sıklık, oran, yüzde
Sürekli	Sayılarla ölçülen değer Örneğin: FEV1, Kan basıncı	Ortalama, standart sapma



- Sürekli deęişkenler için gerekli örneklem büyüklüęü daha azdır
- Deęişkenlięin artması (standart sapma) örneklem büyüklüęünü arttırır

# Minimum saptanabilecek fark / Minimum önemli fark

- İki grup arasındaki fark veya ilişki için gereken minimum değer: klinik olarak önemli, anlamlı minimum değer
- Literatür
- Deneyim
- Pilot çalışma
- Değerlendirme, klinik anlamlılık
- İstatistiksel olarak anlamlı olan –her zaman- klinik olarak anlamlı değildir

# Minimum saptanabilecek fark / Minimum önemli fark

- Araştırma planlarken örneklem büyüklüğü tahmininde klinik anlamlılık göz önüne alınmalı
- Araştırmalarda klinik anlamlı fark bilinmeyebilir
- O konuda yapılmış araştırmaların sonuçlarından yararlanılarak, diğer ölçeklerle karşılaştırma yapılarak istatistiksel yöntemler yardımıyla minimum önemli fark tahmin edilebilir

# Minimum önemli fark: Örnek

Pulmoner hipertansiyon tedavi çalışmalarında kullanılan 6 dakika yürüme testi için minimum önemli fark tahmin edilmiş. SF-36 yaşam kalitesi ölçeğinden yararlanılmış

TABLE 2. ANCHOR- AND DISTRIBUTIONAL-BASED ESTIMATES OF THE MINIMAL IMPORTANT DIFFERENCE FOR THE 6-MINUTE WALK TEST

Method	MID Estimate (m)
Anchor method (MID PCS = 5)	38.6
ES [Eos – baseline/SD(bl)]	38.4
SRM [Eos – baseline/SD(Eos – baseline)]	28.8
SEMeas [SD(baseline) × $\sqrt{1-ICC}$ ]	25.1
0.5 SD [0.5 × SD of baseline 6MWT]	38.5

*Definition of abbreviations:* 6MWT = 6-minute walk test; ICC = intraclass correlation coefficient; Eos = end of study; ES = effect size; MID = minimal important difference; PCS = physical component summary score; SEMeas = standard error of the measurement; SRM = standardized response mean.

American Journal of Respiratory and Critical Care Medicine. 2012;186(5):428-433.  
doi:10.1164/rccm.201203-0480OC.

# FEV1 deęiřimi- Minimum klinik önemli fark

“A minimum clinically important difference (MCID) has not been defined, although improvement of about 100–120 mL in trough FEV1 has been suggested as a possible benchmark for bronchodilators and anti-inflammatory agents, but this is not based on quantitative data.”\*

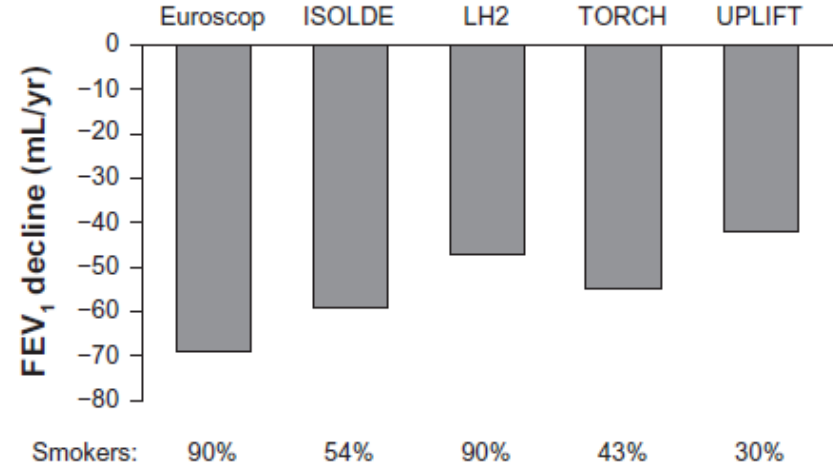


Figure 2 Rate of FEV<sub>1</sub> decline in the placebo arm of long-term COPD studies.<sup>4,36,55-57</sup>  
Abbreviations: COPD, chronic obstructive pulmonary disease; FEV<sub>1</sub>, forced expiratory volume in 1 second.

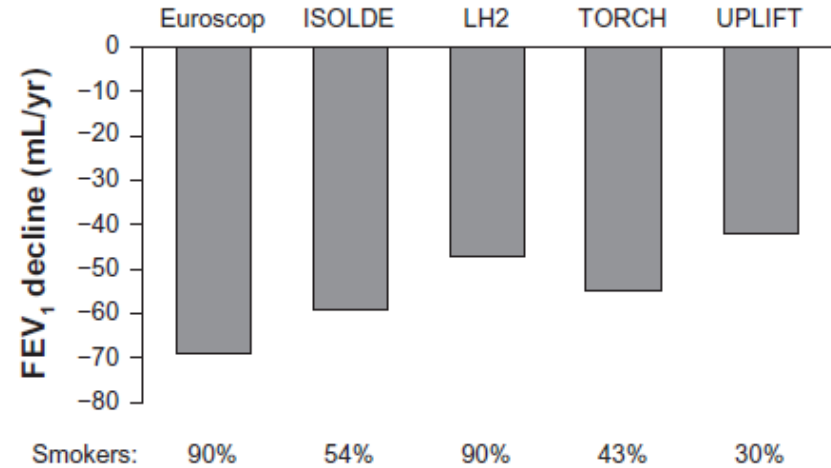
Farklı arařtırmalarda plasebol kolundaki yıllık FEV1 deęiřimi

\*: “Minimum klinik önemli fark tanımlanmamıřtır, FEV1 bazalinde 100-120 ml iyileřme bronkodilatörler ve anti-inflamatuar bir hedef olarak önerilmiřtir ancak sayısal verilerle dayanmamaktadır.”

# FEV1 deęiřimi- Minimum klinik önemli fark

“A minimum clinically important difference (MCID) has not been defined, although improvement of about 100–120 mL in trough FEV1 has been suggested as a possible benchmark for bronchodilators and anti-inflammatory agents, but this is not based on quantitative data.”

\*: “Minimum klinik önemli fark tanımlanmamıştır, FEV1 bazalinde 100-120 ml iyileřme bronkodilatörler ve anti-inflamatuar bir hedef olarak önerilmiştir ancak sayısal verilerle dayanmamaktadır.”



**Figure 2** Rate of FEV<sub>1</sub> decline in the placebo arm of long-term COPD studies.<sup>4,36,55-57</sup>  
**Abbreviations:** COPD, chronic obstructive pulmonary disease; FEV<sub>1</sub>, forced expiratory volume in 1 second.

Farklı arařtırmalarda  
plasebo kolundaki  
yıllık FEV1 deęiřimi

# İstatistiksel güç

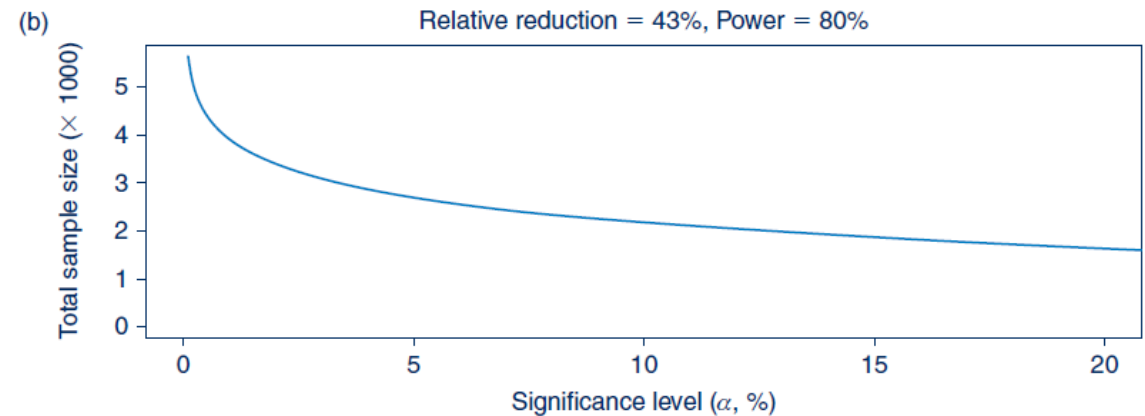
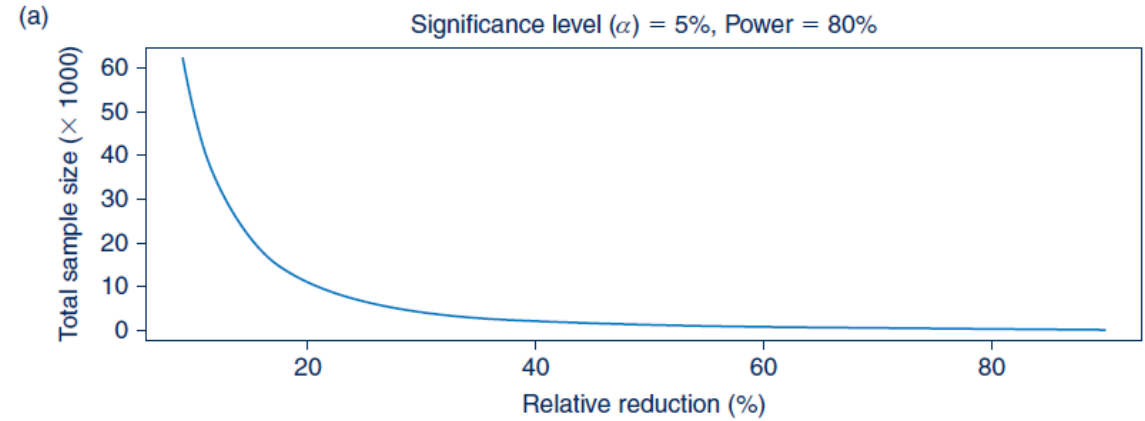
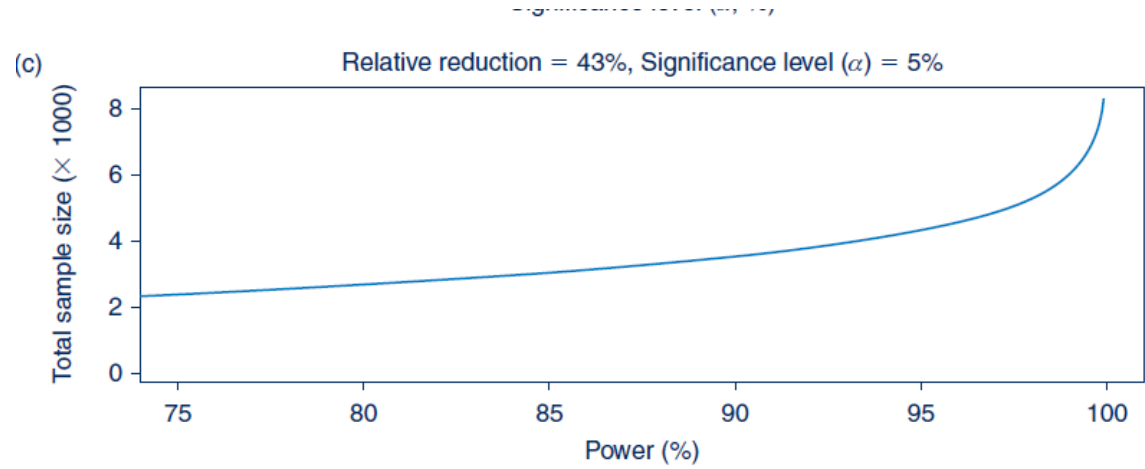
- **Araştırma planlama aşamasında:** veriler toplanmadan önce aranan farkı bulmak için gerekli örneklem büyüklüğünü hesaplamak için kullanılır
- **Veri analizinden sonra:** Fark istatistiksel anlamlı bulunamadığında, bu farkı göstermek için yeterli istatistiksel güce ulaşıldı mı, sorusunu yanıtlamak için kullanılır(örneğin, işyeri çalışmaları, örneklem büyüklüğü sınırlı)

Vol. 23(7–8) 2012 NSW  
Public Health Bulletin

i. Relatif azalma ve istatistiksel anlamlılık verili: güç – örneklem büyüklüğü ilişkisi

ii. istatistiksel anlamlılık ve güç verili: Relatif azalma - örneklem büyüklüğü ilişkisi

iii. Relatif azalma ve güç verili: istatistiksel anlamlılık - örneklem büyüklüğü ilişkisi





# Katılım Oranı

- Örnekleme seçilenler içinde araştırmaya katılmayı kabul etmeyen, bilgi vermeyenler olabilir
- Temsiliyeti etkileyebilir: katılanlardan ne oranda farklı (örneğin araştırılan hastalığa sahip olanlar, daha ağır durumda olanlar araştırmaya katılmaya daha istekli olabilir...)
- Ulaşılması istenen sayı hesaplanırken tahmin edilen katılım oranına göre düzeltim yapılması önerilir (örneğin: hesaplanan örneklem büyüklüğü 480, tahmin edilen katılım oranı %75, örneklem büyüklüğüne ulaşmak için çalışmaya davet edilecek sayı:  $480/0.75 = 600$  olarak hesaplanır. )

# Özet

1. Araştırmada amaç hedef toplum hakkında geçerli bilgi elde etmektir (çıkarım)
2. Örneklem alınması hedef toplumda geçerli bilgiyi daha kolay elde etme amacıyla (yapılabilirlik) uygulanır
3. Örneklem büyüklüğü hesabı araştırılan sorunun niteliği, tahmini dağılım bilgileri, minimum anlamlı fark ile belirlenen tahminlerden oluşur

# Özet

4. Daha küçük bir farkı göstermeyi hedeflemek, daha büyük bir değişkenlik, tip I hatanın küçük olması, tip II hatanın küçük olması örneklem büyüklüğünü artırır
5. Sürekli değişken için gerekli örneklem büyüklüğü kategorik değişken için gerekli olandan daha azdır

# Örneklem Büyüklüğü Hesaplarken

1.  $H_0$  sapta
2. Bu hipotezi sınamak için uygun istatistiksel yöntemi seç
3. Mantıklı (gerçekçi) bir etki büyüklüğü (effect size) ve değişkenlik (gerekirse) belirle
4.  $\alpha$  ve  $\beta$ 'yi seç (tek veya çift taraflı  $\alpha$  için karar ver)
5. Yazılım programına bu değerleri gir (tablo veya formül de olabilir)
6. Örneklem için seçilen kabulleri gözden geçir, tekrar dene

Hesaplamalar İin Web Sitelerinden Yararlanılabilir  
Örneđin: <http://www.stat.ubc.ca/~rollin/stats/ssize/>

---

# Formüller

# $\alpha/\beta$ değerleri

	Alpha		Beta			
	0.05	0.01	0.20	0.10	0.05	0.01
<b>Multiplier</b>	1.96	2.58	0.842	1.28	1.64	2.33

# Prevalans araştırması

$$N = \frac{Z_{\alpha/2}^2 * P * (1-p) * D}{E^2}$$

P: araştırılan toplumda tahmin edilen oran,  
prevalans

E: hata sınırı (genelde P'nin %5-10'u kabul edilir)

$Z_{\alpha/2}$ : normal dağılımda çift taraflı test için saptanan  
 $\alpha$  değeri

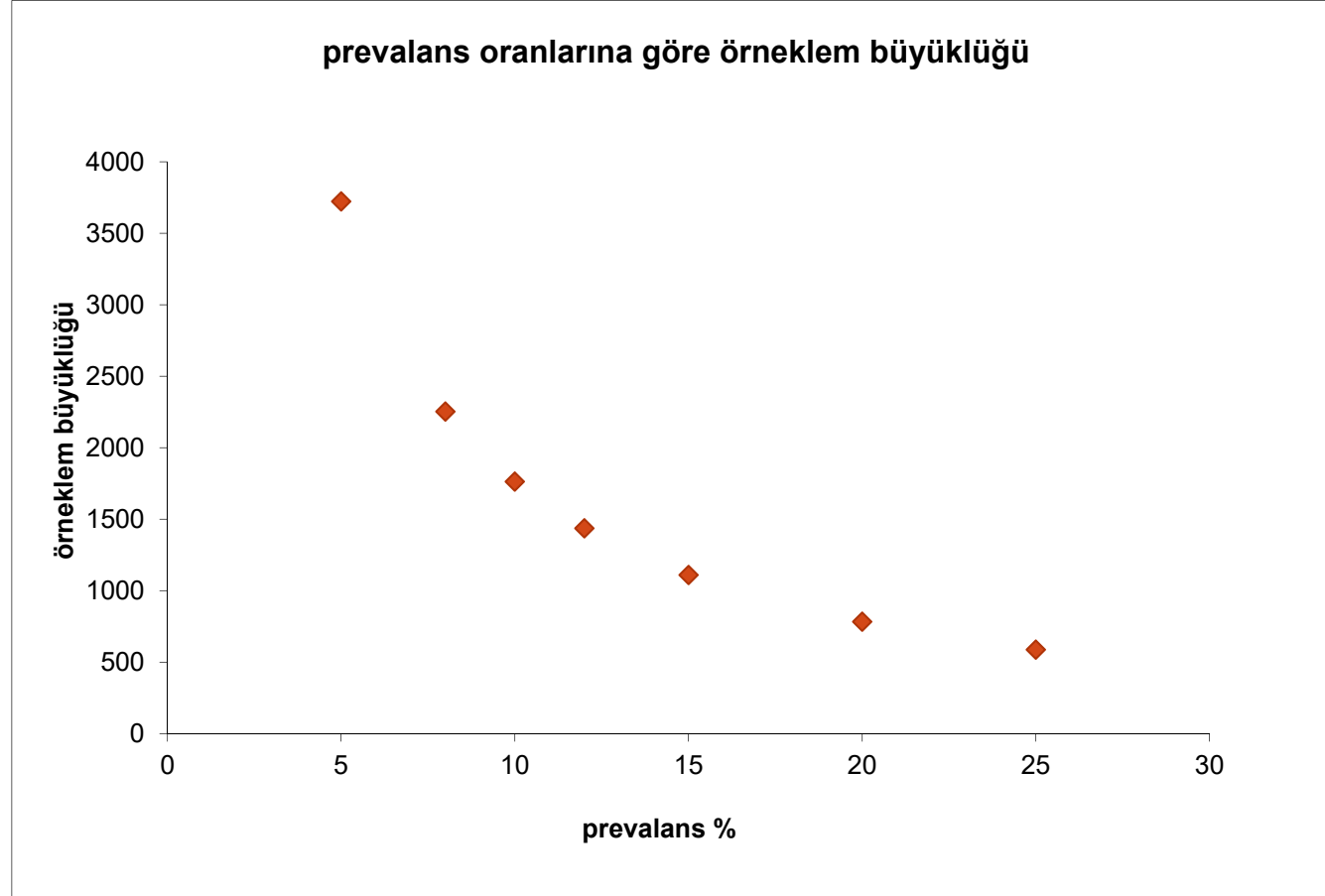
D: tasarım etkisi (genelde 1-2 arası, basit rastgele örneklem seçiminden sapmayı düzeltmek için, sistematik cluster örneklemede 1.5-2, amaçlı, uygun olanı alan örneklem için 10)

# Prevalans araştırması örneklem tahmini örnekler

	senaryo 1	senaryo 2	senaryo 3	senaryo 4	senaryo 5	senaryo 6	senaryo 7
z	1,96	1,96	1,96	1,96	1,96	1,96	1,96
p	0,08	0,05	0,1	0,12	0,15	0,2	0,25
1-p	0,92	0,95	0,9	0,88	0,85	0,8	0,75
hata (p/10)	0,008	0,005	0,01	0,012	0,015	0,02	0,025
%95 GA	0,06	0,04	0,08	0,10	0,12	0,16	0,20
	0,10	0,06	0,12	0,14	0,18	0,24	0,30
örneklem büyüklüğü (n)	2254	3724	1764	1437	1110	784	588



# Prevalans arařtırması rneklem tahmini rnekler



# Grup ortalaması

$$N = Z_{\alpha/2}^2 s^2 / d^2$$

s: standart sapma (önceki, pilot çalışma)

$Z_{\alpha/2}$ : normal dağılımda çift taraflı test için saptanan  $\alpha$  değeri

d: tahminin doğruluğu, ortalamaya ne kadar yakın olduğu

# İki ortalamanın karşılaştırılması

$$N = \frac{(r + 1) (Z_{\alpha/2} + Z_{1-\beta})^2 \sigma^2}{rd^2}$$

r:  $n_1/n_2$  (iki grubun sayılarının oranı)

$Z_{\alpha/2}$ : normal dağılımda çift taraflı test için saptanan  $\alpha$  değeri

$Z_{1-\beta}$ : normal dağılımda çift taraflı test için saptanan  $1 - \beta$  değeri ( $\beta$ : istatistiksel güç,  $1 - \beta$ : tip II hata)

$\sigma$ : standart sapma

d: ortalamalar arasındaki fark

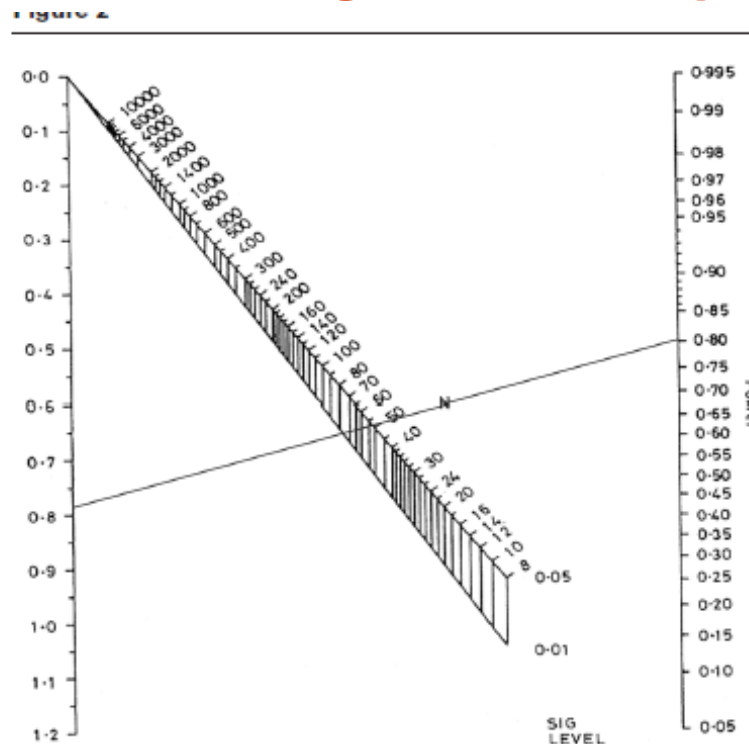
- Standardize fark: hedeflenen fark/ standart sapma
- Oranlar için:  $p_1 - p_2 / (\sqrt{p_{ort} (1 - p_{ort})})$
- Standart sapma: literatür, pilot çalışma

Critical Care 2002; 6:335-341

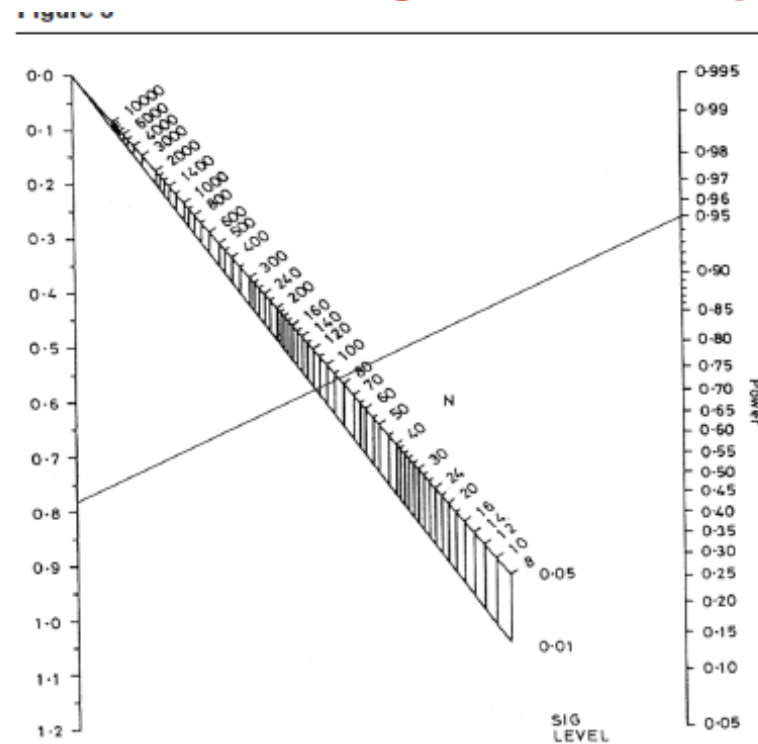
Örnek: beklenen kan basıncı farkı 18 mmHg, standart sapma: 18mmHg, Standardize fark: 14/18: 0.78

**Güç: 0.80, p: 0.05 >>> n: 52, her iki grupta 26 kişi**

**Güç: 0.95, p: 0.05 >>> n: 86, her iki grupta 43 kişi**



Nomogram showing sample size calculation for a standardized difference of 0.78 and 80% power.



Nomogram showing sample size calculation for a standardized difference of 0.78 and 95% power.

- Standardize fark: hedeflenen fark/ standart sapma
- Oranlar için:  $p_1 - p_2 / (\sqrt{p_{ort} (1 - p_{ort})})$
- Standart sapma: literatür, pilot çalışma

Critical Care 2002; 6:335-341

# Her iki grupta eşit sayı

- $n = (2/d^2) \times C_{p,power}$
- $n$ : her grup için gereken sayı
- $d$ : Standardize fark

Table 2

Commonly used values for  $C_{p,power}$

$P$	Power (%)			
	50	80	90	95
0.05	3.8	7.9	10.5	13.0
0.01	6.6	11.7	14.9	17.8

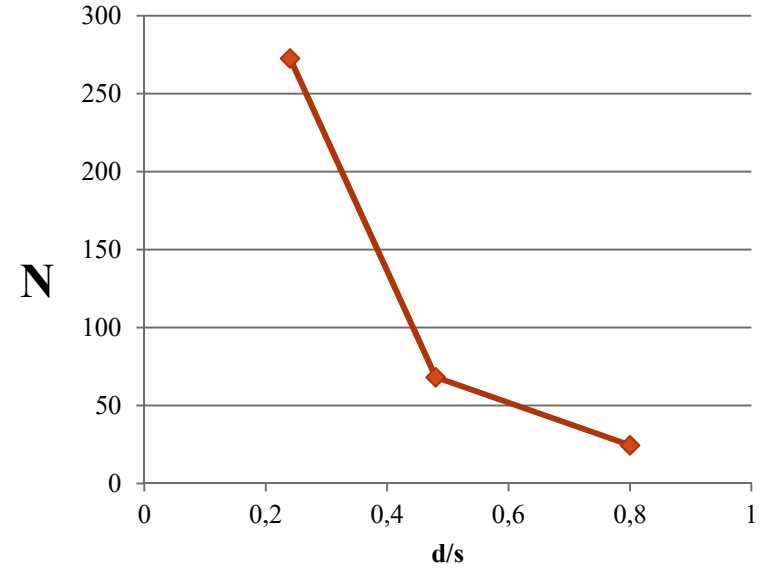
# Standardize fark –örneklem büyüklüğü

p	Güç	C-power	d (standardize fark)	N (her gruptaki sayı)
0,05	0,8	7,9	0,78	25
0,01	0,8	11,7	0,78	38
0,05	0,9	10,5	0,78	34
0,01	0,9	14,9	0,78	48
0,05	0,8	7,9	0,35	128
0,01	0,8	11,7	0,35	191
0,05	0,9	10,5	0,35	171



# Örnekler

z-a	1,96	1,96	1,96
z-b, 0.80	0,842	0,842	0,842
s	15	25	50
r	1	1	1
d	12	12	12
d/s	0,8	0,48	0,24
N	25	68	273



Z-a: alfa (tip I hata:  $p$ ) için z değeri

Z-b, %80: istatistiksel güç için z değeri

(tip iki hata: beta: 1-güç: 1-0.8: 0.2)

# İki oranın karşılaştırılması

$$N = \frac{(Z_{\alpha/2} \sqrt{2p(1-p)} + Z_{\beta/2} \sqrt{p_1(1-p_1) + p_2(1-p_2)})^2}{(p_1 - p_2)^2}$$

$p_1 - p_2$ : iki gruptaki oran tahminleri

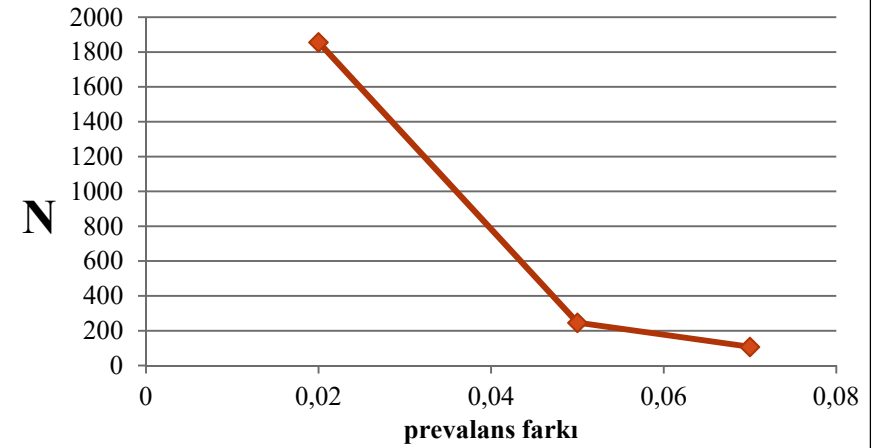
$p$ :  $(p_1 + p_2)/2$

$Z_{\alpha/2}$ : normal dağılımda çift taraflı test için saptanan  $\alpha$  değeri

$Z_{1-\beta}$ : normal dağılımda çift taraflı test için saptanan  $1 - \beta$  değeri ( $\beta$ : istatistiksel güç,  $1 - \beta$ : tip II hata)

# Örnekler

z-a	1.96	1.96	1.96
z-b, 80	0.842	0.842	0.842
p1	0.1	0.1	0.1
p2	0.03	0.05	0.08
p	0.065	0.075	0.09
N	108	247	1857



Z-a: alfa (tip I hata: p) için z değeri

Z-b, %80: istatistiksel güç için z değeri

(tip iki hata: beta: 1-güç: 1-0.8: 0.2)

# Korelasyon katsayısı

$$N = \frac{(Z_{\alpha/2} + Z_{1-\beta})^2}{4 \cdot \frac{1}{1-r} [\ln(1+r)]}$$

r: korelasyon katsayısı ( $H_0: r=0$ )

$Z_{\alpha/2}$ : normal dağılımda çift taraflı test için saptanan  $\alpha$  değeri

$Z_{1-\beta}$ : normal dağılımda çift taraflı test için saptanan  $1 - \beta$  değeri ( $\beta$ : istatistiksel güç,  $1-\beta$ : tip II hata)

# Odds ratio (olasılıklar oranı)

$$N = \frac{(1+r)^2 (Z_{\alpha/2} + Z_{1-\beta})^2}{r (\ln OR)^2 [p(1-p)]}$$

$$OR = \frac{p_1(1-p_2)}{p_2(1-p_1)}$$

p: durumun (hastalığın) toplumdaki prevalansı

r: n1/n2 (iki grubun sayılarının oranı)

$Z_{\alpha/2}$ : normal dağılımda çift taraflı test için saptanan  $\alpha$  değeri

$Z_{1-\beta}$ : normal dağılımda çift taraflı test için saptanan  $1-\beta$  değeri ( $\beta$ : istatistiksel güç,  $1-\beta$ : tip II hata)

# Kaynaklar

1. Journal of Human Reproductive Sciences. 2012;5(1):7-13. doi:10.4103/0974-1208.97779.
2. Critical Care. 2002;6(4):335-341.
3. The Diabetes Educator September/October 2010 36: 701-707.  
doi: 10.1177/0145721710380791
4. Nephron Clin Pract 2011;118:c319–c323.  
doi:10.1159/000322830
5. MECOR programındaki sunular