

Güç ve Örneklem Büyüklüğü

Doç. Dr. Şeref Kerem Çorbacioğlu



@drserefkerem

 akamedika



istatistik@akamedika.com



@akamedika



Örneklem Büyüklüğü Hesaplama



1. Anlamı ve önemi nedir? İşin felsefesi yani...

2. Nasıl hesaplanır?



EN KOLAYI !!!

3. Hesaplama için gereken veriler nereden bulunur?

doqiii

 akamedika

Örneklem Büyüklüğü Hesaplama



Anlamı ve önemi

Örneklem Büyüklüğü Hesaplama

- **Tanım;** Çalışmanızın



- Kabul edilebilir hata boyutlarında
- Klinik anlamlılık ile uyumlu bir şekilde
- Gereken en az denek sayısını hesaplama

Tip 1 %5 ve Tip 2 için %20

Literatür, beklenti, hedef, maliyet-etkinlik

n: 324

işlemidir...

Belirli bir yargıya ulaşmak için **YETERLİ** delile ulaşmak gerekir...

Örneklem büyüklüğüdür

YARGI



Delil

Delil

Delil

Delil

Delil

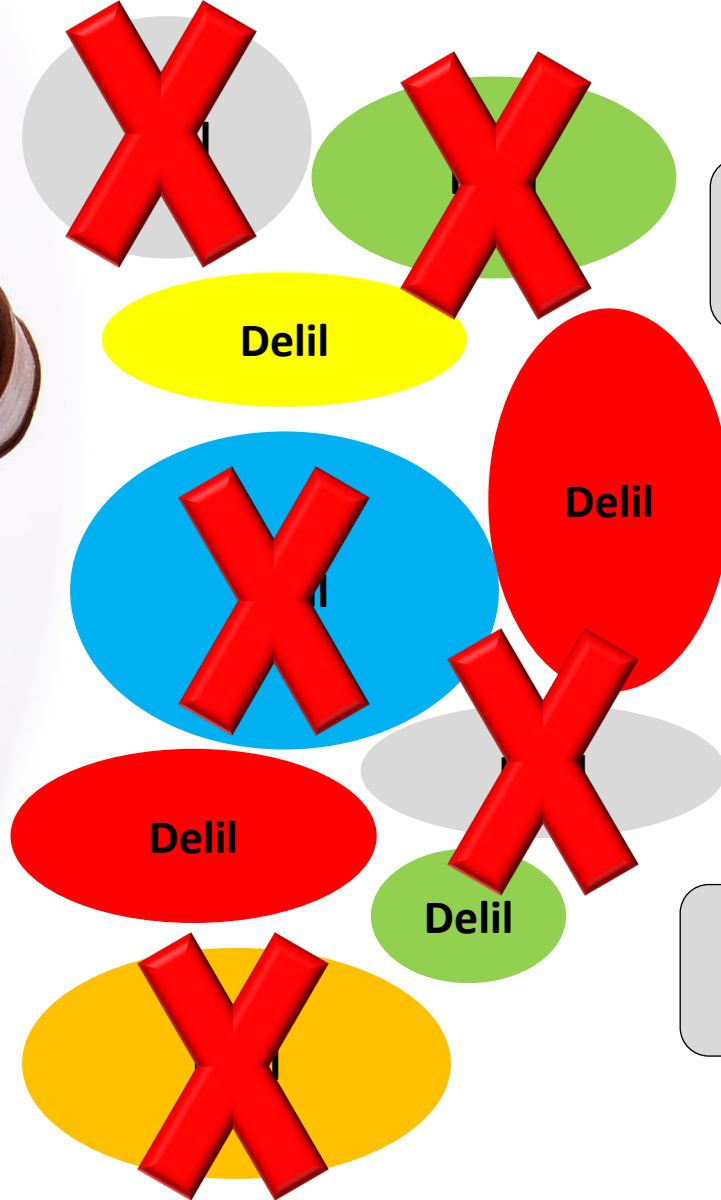
Delil

Delil

Delil

Delil

YETERLİ delil miktarı vereceğiniz kararın **önemine** göre değişebilir...



Hafif para cezası

Klinik anlamlılık – etki
büyüklüğü

YARGI

Müebbet



Yargılama sürecindeki doğal hata süreçlerine yaklaşımınız da delil miktarınızı etkileyebilir...

Tip-1 ve Tip-2 Hatalar

Hiçbir suçluyu
serbest bırakmak
istemezseniz

Tip-1 Hata

Hiçbir suçsuzu
hapse atmak
istemezseniz

Tip-2 Hata

YARGI



Tip-1 ve Tip-2 Hata ne?



- Çalışma koşullarını karşılayan gerçek hayattaki tüm bireyleri çalışmaya katmadan tip-1 ve tip-2 hata sıfırlanamaz.

Tüm evreni dahil etmeden Tip-1 ve Tip-2 Hatayı sıfırlayamayacağımız için belirli hata paylarını baştan belirler ve kabul ederiz.

Tip-1 Hata

Alfa değeri - α

%5



Tip-2 Hata

Beta değeri

%20



Bir örnek ile başlayalım...

- Akciğer Kanseri için kullanılabilecek yeni bir KT ilacını standart KT uygulaması ile üç yıllık mortalite açısından karşılaştırmak istiyorsunuz.
- Bunun için her gruba 25'şer hasta alıyorsunuz...

morta * Kemoterapi Crosstabulation

morta	surv	Count	Kemoterapi		Total
			plasebo	KT	
		Count	15	20	35
		Expected Count	17.5	17.5	35.0
		% within morta	42.9%	57.1%	100.0%
		% within Kemoterapi	60.0%	80.0%	70.0%
		% of Total	30.0%	40.0%	70.0%
mortalite		Count	10	5	15
		Expected Count	7.5	7.5	15.0
		% within morta	66.7%	33.3%	100.0%
		% within Kemoterapi	40.0%	20.0%	30.0%
		% of Total	20.0%	10.0%	30.0%
Total		Count	25	25	50
		Expected Count	25.0	25.0	50.0
		% within morta	50.0%	50.0%	100.0%
		% within Kemoterapi	100.0%	100.0%	100.0%
		% of Total	50.0%	50.0%	100.0%

%33'lük bir mortalite azalması var

AMA anlamlı değil...

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	2.381 ^a	1	.123		
Continuity Correction ^b	1.524	1	.217		
Likelihood Ratio	2.416	1	.120		
Fisher's Exact Test				.217	.108
Linear-by-Linear Association	2.333	1	.127		
N of Valid Cases	50				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 7.50.

b. Computed only for a 2x2 table

	Kemoterapi	uc_yillik_mortalite	var
10	KT	surv	
11	KT	surv	
12	KT	surv	
13	KT	surv	
14	KT	surv	
15	KT	surv	
	KT	surv	
	KT	surv	
	KT	surv	
	KT	surv	
	KT	surv	
	KT	mortalite	
	KT	mortalite	
	KT	surv	
	KT	surv	
25	KT	surv	
26	standart	surv	
27	standart	surv	
28	standart	surv	
29	standart	surv	
30	standart	surv	
31	standart	surv	
32	standart	surv	

Bir örnek ile başlayalım...

- Akciğer Kanseri için kullanılabilecek yeni bir KT ilacını standart KT uygulaması ile üç yıllık mortalite açısından karşılaştırmak istiyorsunuz.
- Bunun için her gruba 50'şer hasta alıyorsunuz...

		Kemoterapi			
		standart	KT	Total	
uc_yillik_mortalite	surv	Count	30	40	70
		Expected Count	35.0	35.0	70.0
		% within uc_yillik_mortalite	42.9%	57.1%	100.0%
		% within Kemoterapi	60.0%	80.0%	70.0%
		% of Total	30.0%	40.0%	70.0%
	mortalite	Count	20	10	30
		Expected Count	15.0	15.0	30.0
		% within uc_yillik_mortalite	66.7%	33.3%	100.0%
		% within Kemoterapi	40.0%	20.0%	30.0%
		% of Total	20.0%	10.0%	30.0%
Total	Count	50	50	100	
	Expected Count	50.0	50.0	100.0	
	% within uc_yillik_mortalite	50.0%	50.0%	100.0%	
	% within Kemoterapi	100.0%	100.0%	100.0%	
	% of Total	50.0%	50.0%	100.0%	

%33'lük bir mortalite azalması var

İstatistiksel anlamlı

Chi-Square Tests				
	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	4.762 ^a	1	.029	
Continuity Correction ^b	3.857	1	.050	
Likelihood Ratio	4.831	1	.028	
Fisher's Exact Test				.049
Linear-by-Linear Association	4.714	1	.030	
N of Valid Cases	100			

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 15.00.

b. Computed only for a 2x2 table

	Kemoterapi	uc_yillik_mortalite	var
10	KT	surv	
11	KT	surv	
12	KT	surv	
13	KT	surv	
14	KT	surv	
15	KT	surv	
	KT	surv	
	KT	surv	
	KT	surv	
	KT	surv	
	KT	surv	
	KT	mortalite	
	KT	mortalite	
	KT	surv	
	KT	surv	
25	KT	surv	
26	standart	surv	
27	standart	surv	
28	standart	surv	
29	standart	surv	
30	standart	surv	
31	standart	surv	
32	standart	surv	

Pragmatist ama gerekli bir tanım...

- Bizim için klinik açıdan önemsenecek bir fark çıktığında istatistiksel olarak ta anlamlı çıkması için gereken en az örneklem sayısının belirlenmesi aslında örneklem büyüklüğü hesabı...

Etki büyüklüğü kavramı ile
açıklayacağız...

Bir başka örneğe bakalım...

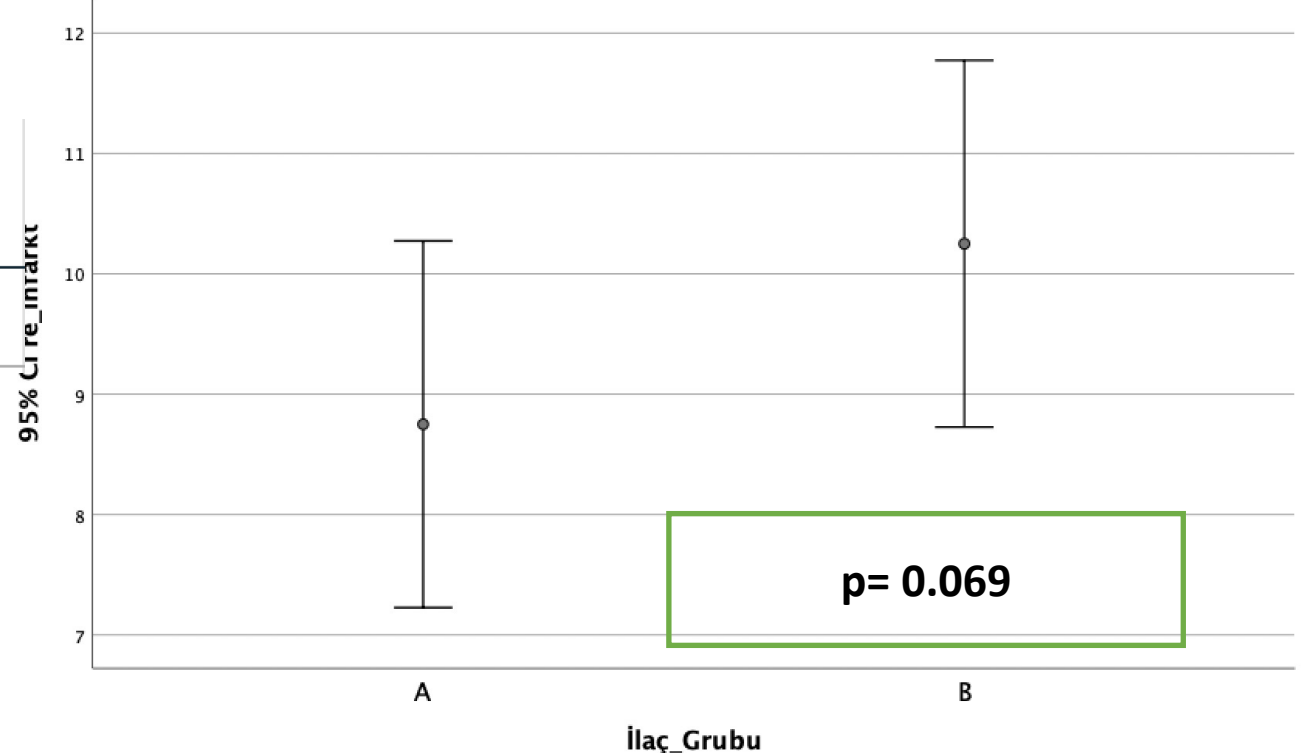
- STEMI sonrası başlanan iki farklı anti-platelet ajanın re-infarkt süreleri (ay cinsinden) açısından bir araştırma planlıyorsunuz. Toplam 8 hasta (her iki gruba 4'er) ile yaptığınız araştırmanın sonuçları;

Group Statistics

İlaç_Grubu	N	Mean	Std. Deviation	Std. Error Mean
re_infarkt A	4	8.75	.957	.479
re_infarkt B	4	10.25	.957	.479

Neredeyse 2 ay var!!!

Sig. (2-tailed)
.069



p= 0.069

8 :

	İlaç_Grubu	re_infarkt
1	1	8
2	1	9
3	1	8
4	1	10
5	2	10
6	2	11
7	2	11
8	2	9

Bir başka örneğe bakalım...

- Sayıyı iki katına çıkarıyoruz...

Group Statistics

	İlaç_Grubu	N	Mean	Std. Deviation	Std. Error Mean
re_infarkt	A	4	8.75	.957	.479
	B	4	10.25	.957	.479

Sig. (2-tailed)

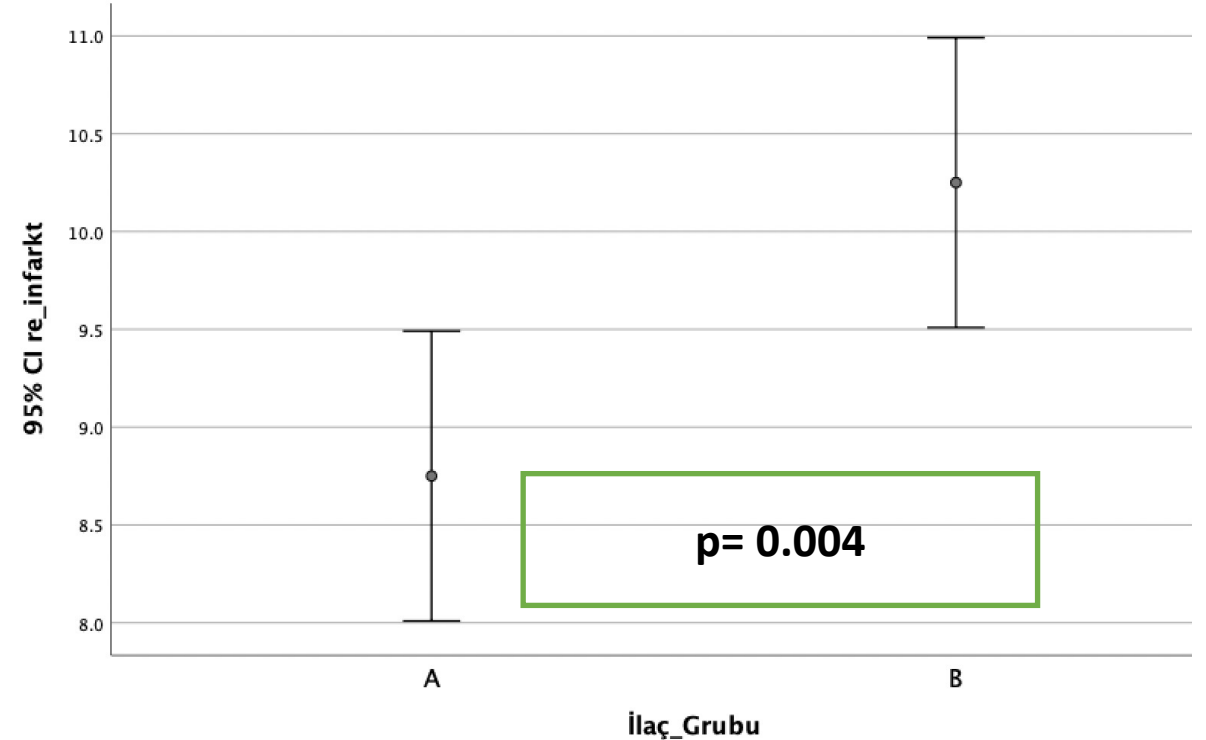
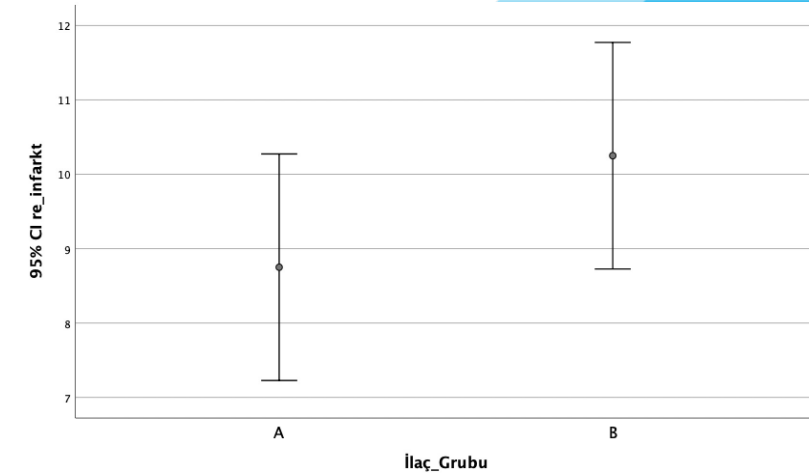
.069

Group Statistics

	İlaç_Grubu	N	Mean	Std. Deviation	Std. Error Mean
re_infarkt	A	8	8.75	.886	.313
	B	8	10.25	.886	.313

Sig. (2-tailed)

.004



Bir başka pragmatist ama gerekli bir tespit...

- İki örnekte de vaka sayısını artırarak aslında Tip-2 Hata endişesini ortadan kaldırmaya çalıştık.

İkinci çalışmada 8 hasta aldığımızda re-infarkt açısından iki ay fark çıkmıştı

Ama anlamsızdı...

Gerçek hayatta fark olsaydı ve sırf biz az vaka aldık diye anlamsız çıkmışsa;
TİP-2 HATA
Yaptık demektir

Gerçek hayatta da fark yoksa sorun yok aslında. AMA bu gerçeği henüz bilmediğimiz için sırf biz az vaka aldık;
TİP-2 HATA
riskini anlatamayız.

16 hastaya çıktığımızda ise aynı farkı (2 aylık) bulduk ve bu istatistiksel olarak anlamlı çıkmış ise Tip-2 Hata riskini egale ettik demektir.

Gerçek hayatta fark yok ise bu durumda biz 16 vakaya çıktığımızda arada **bulduğumuz farkın 1 ay gibi** süreye düşmesi ve yine istatistiksel olarak anlamsız çıkması lazımdır.

Devam edelim...

- İki örnekte de vaka sayısını artırarak aslında Tip-2 Hata endişesini ortadan kaldırmaya çalıştık...

Biz 16 vakaya çıktığımızda arada **bulduğumuz farkın 1 ay gibi** süreye düştü...
Matematiksel fark: 1 ay

Ama anlamsızdı...

Gerçek hayatta da 1 aylık fark yoksa sorun yok aslında. AMA bu gerçeği henüz bilmediğimiz için sırf biz az vaka (16 vaka) aldık;
TİP-2 HATA
riskini anlatamayız.

Gerçek hayatta 1 aylık ta fark yok ise bu durumda biz 100 vakaya çıktığımızda da arada **bulduğumuz farkın 10 günlük bir** süreye düşmesi ve yine istatistiksel olarak anlamsız çıkması lazımdır.

Bu böyle tüm evreni alana kadar sürebilir... 😊

Nerede bitmeli...

Artık aradaki fark çıksa bile çıkan farkın bizim için çok ta önemli olmadığı bir klinik durum.
Örnek: 10 günlük fark veya daha az varsa da bunun klinik pratikte bir karşılığı olmayabilir.

ETKİ BÜYÜKLÜĞÜ!!!

- Örneklem sayısı artıkça sonucunuzun kesinlik düzeyi artıyor. Yani güven aralığı daralıyor.
- Çünkü evrenin gerçek sayısına örneklem büyüdükçe yaklaşıyorsunuz.



Nedir bu etki büyüklüğü? (Effect Size)

- Etki boyutu, deney ve kontrol grubu arasında gözlemlenen farkın (etkinin) büyüklüğünün nesnel ve (genellikle) standartlaştırılmış bir ölçüsüdür.



Londra Kulučka ve
Şartlandırma Merkezi



Boy Ortalaması: 170cm



Boy Ortalaması: 175cm

Aradaki fark: **ETKİ BÜYÜKLÜĞÜ: 5cm**

Erkek olmanın boyunuzun kaç olacağı üzerine etkisi 5cm'dir.

Nedir bu etki büyüklüğü? (Effect Size)

- Etki boyutu, deney ve kontrol grubu arasında gözlemlenen farkın (etkinin) büyüklüğünün nesnel ve (genellikle) standartlaştırılmış bir ölçüsüdür.

Gerçek Hayatta iki sorun;

1. Standart sapmalar
2. Elde ettiğimiz farkın önemine nasıl karar vereceğiz. Başka çalışmalarla nasıl karşılaştıracacağız.



Boy Ortalaması: 170cm

+SS

Group Statistics

	cinsiyet	N	Mean	Std. Deviation	Std. Error Mean
boy	kadın	5	170.00	7.906	3.536
	erkek	5	175.00	14.036	6.277



Boy Ortalaması: 175cm

+SS

Cohen d

$$\frac{\mu_1 - \mu_2}{SS}$$

Etki büyüklüğünü hesaplayalım...

$$\frac{\mu_1 - \mu_2}{SS} = \frac{175 - 170}{SS}$$

	cinsiyet	N	Mean	Std. Deviation	Std. Error Mean
boy	kadın	5	170.00	7.906	3.536
	erkek	5	175.00	14.036	6.277

1.

Deney ve kontrol grubu olan çalışmalarda «Kontrol Grubunun» SS'si kullanılması önerilir. Veya bir grubun bazal grup kabul edilip onun SS'sini kullanılması önerilir.

2.

İkinci opsiyon iki ayrı gruba ait iki ayrı SS'yi kullanarak «Havuzlanmış SS»'yi kullanmaktır.

Hangisini Kullanalım?

$$s_p = \sqrt{\frac{(N_1 - 1)s_1^2 + (N_2 - 1)s_2^2}{N_1 + N_2 - 2}}$$

Etki büyüklüğünü hesaplayalım...

$$\frac{\mu_1 - \mu_2}{SS} = \frac{175 - 170}{SS}$$

		N	Mean	Std. Deviation	Std. Error Mean
boy	kadın	5	170.00	7.906	3.536
	erkek	5	175.00	14.036	6.277

$$s_p = \sqrt{\frac{(N_1 - 1)s_1^2 + (N_2 - 1)s_2^2}{N_1 + N_2 - 2}}$$

$$\sqrt{\frac{(5 - 1) 14,03^2 + (5 - 1) 7,9^2}{5 + 5 - 2}}$$

$$= 11.3$$

$$\frac{5}{11,3} = 0,44$$

Ee b

Cohen d:



Aradaki farkı bulmak isteseydiniz hangisini kullanırdınız???

1988 yılında Cohen	Effect size	Small	Medium	Large
Difference between two means	d	0.20	0.50	0.80



Ee bu ne demek? Ne işe yarıyor?

Cohen d: effect size:0,44

1988 yılında Cohen	Effect size	Small	Medium	Large
Difference between two means	d	0.20	0.50	0.80

İster santim ile ölç ister arşın ile bana etki büyüklüğü ile gel.

İki farklı çalışmada kullanılan iki farklı ölçüm yöntemine bağlı çıkan sonuçları karşılaştırılabilmesine olanak veriyor...



- Deprem yaşamış kişilerle yaşamamış kişilerin arasında anksiyete düzeyleri açısından fark olup olmadığını araştıran iki çalışma düşünün...

Yani deprem yaşamış olmanın anksiyete üzerinde ne kadar BÜYÜK ETKİ'si olduğunu araştıran iki çalışma...

A Çalışması

Beck Ölçeği Kullanılmış;

- Deprem grubu ort: 26 SS: 10
- Kontrol grubu ort: 10 SS: 7

16 puan
fark

EB: 0.7

B Çalışması

Hamilton Ölçeği Kullanılmış;

- Deprem grubu ort: 18 SS: 9
- Kontrol grubu ort: 8 SS: 5

10 puan
fark

EB: 0.63



Çalışmanın Gücü (Power)

Güç: 1-Tip2 Hata

- Tanım olarak güç yani «power»; yapılan çalışmada fark olduğu varsayılan değişken açısından gerçekten de fark varken, var olan bu farkı doğru olarak tespit edebilme olasılığıdır.



- Araştırılan değişken açısından gerçek hayatta fark varken sizin çalışmanızda fark bulamama olasılığına TİP-2 hata deniliyordu.

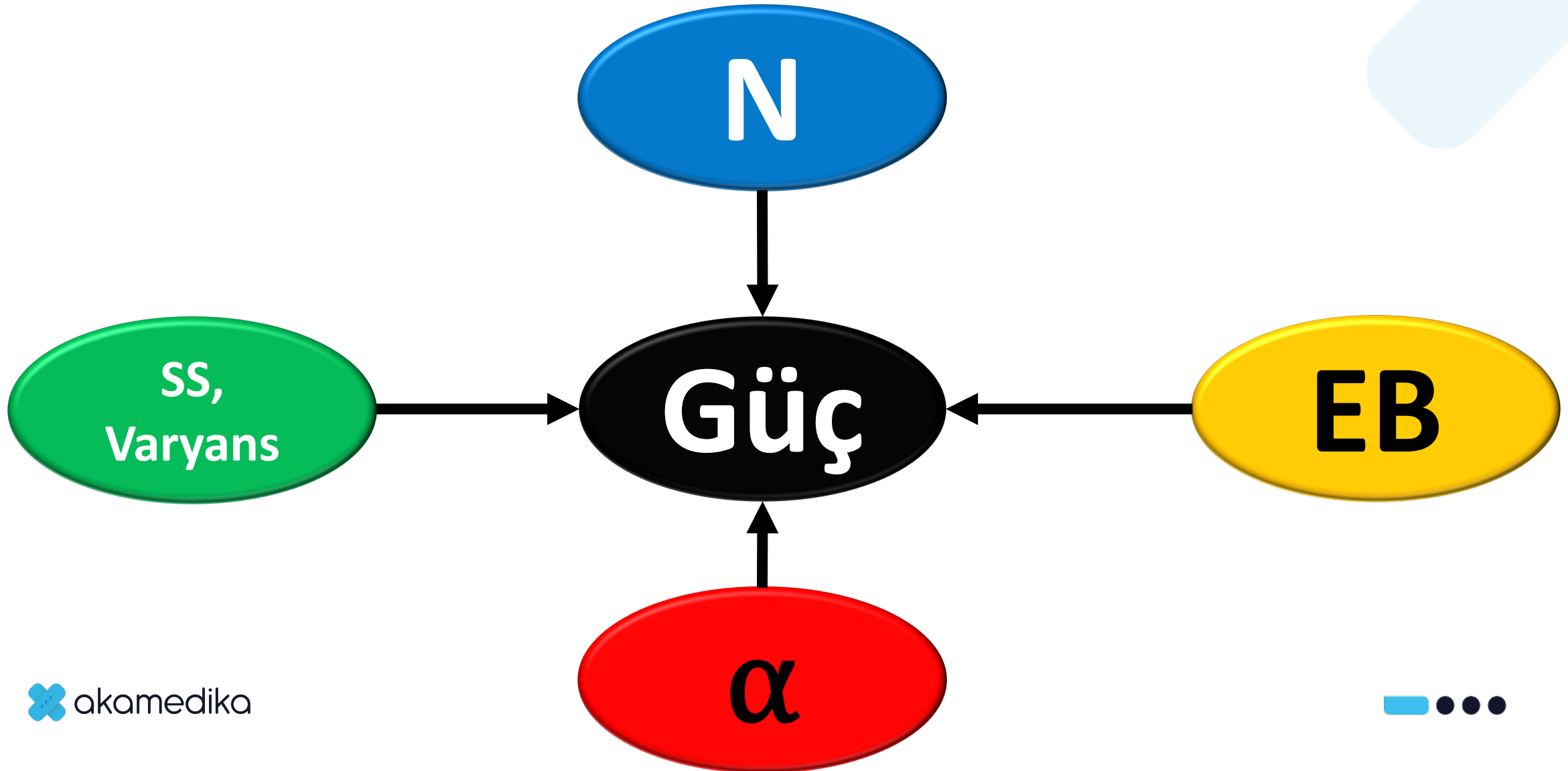
Tip-2 Hata

Beta değeri

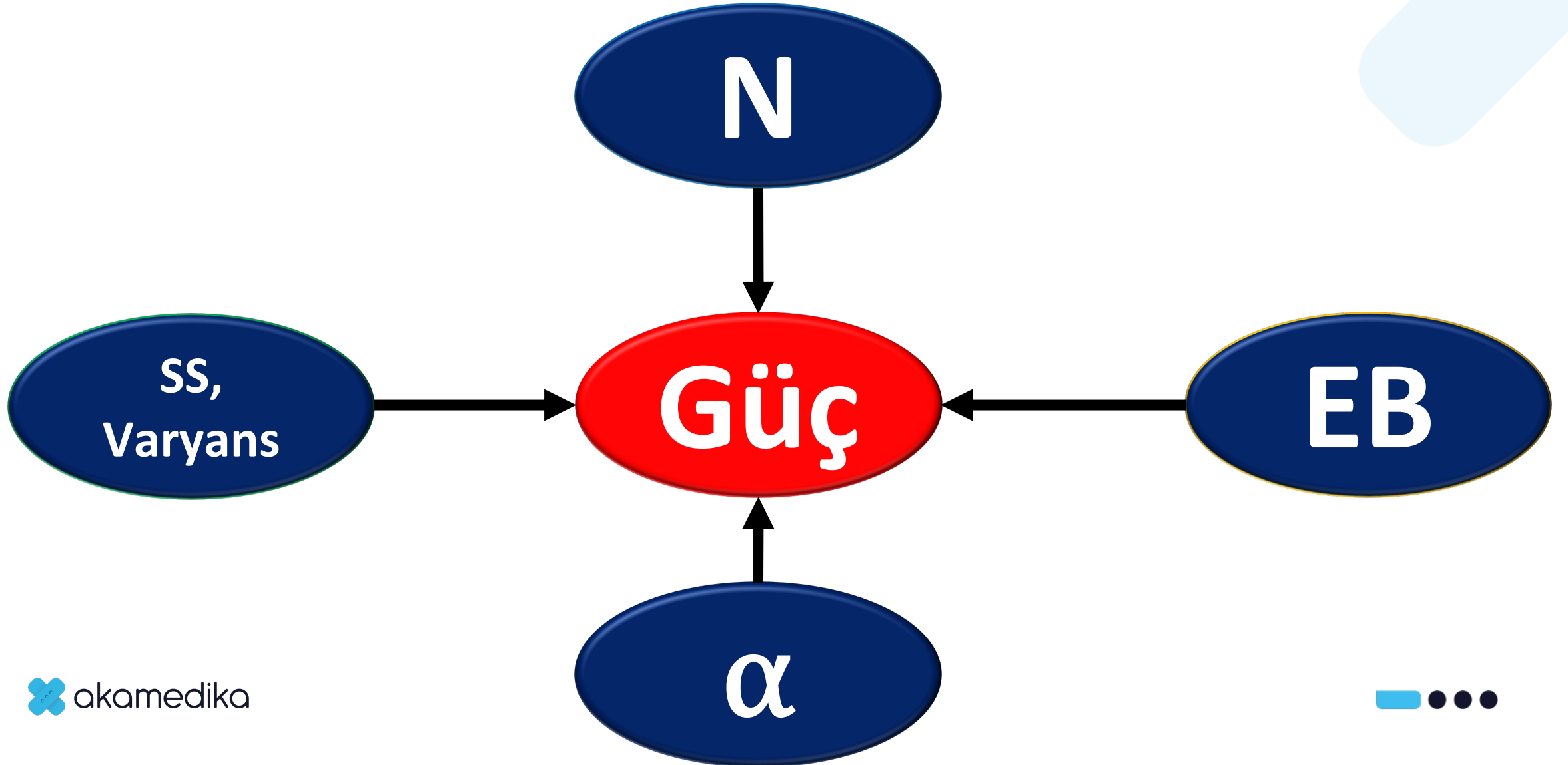
%20



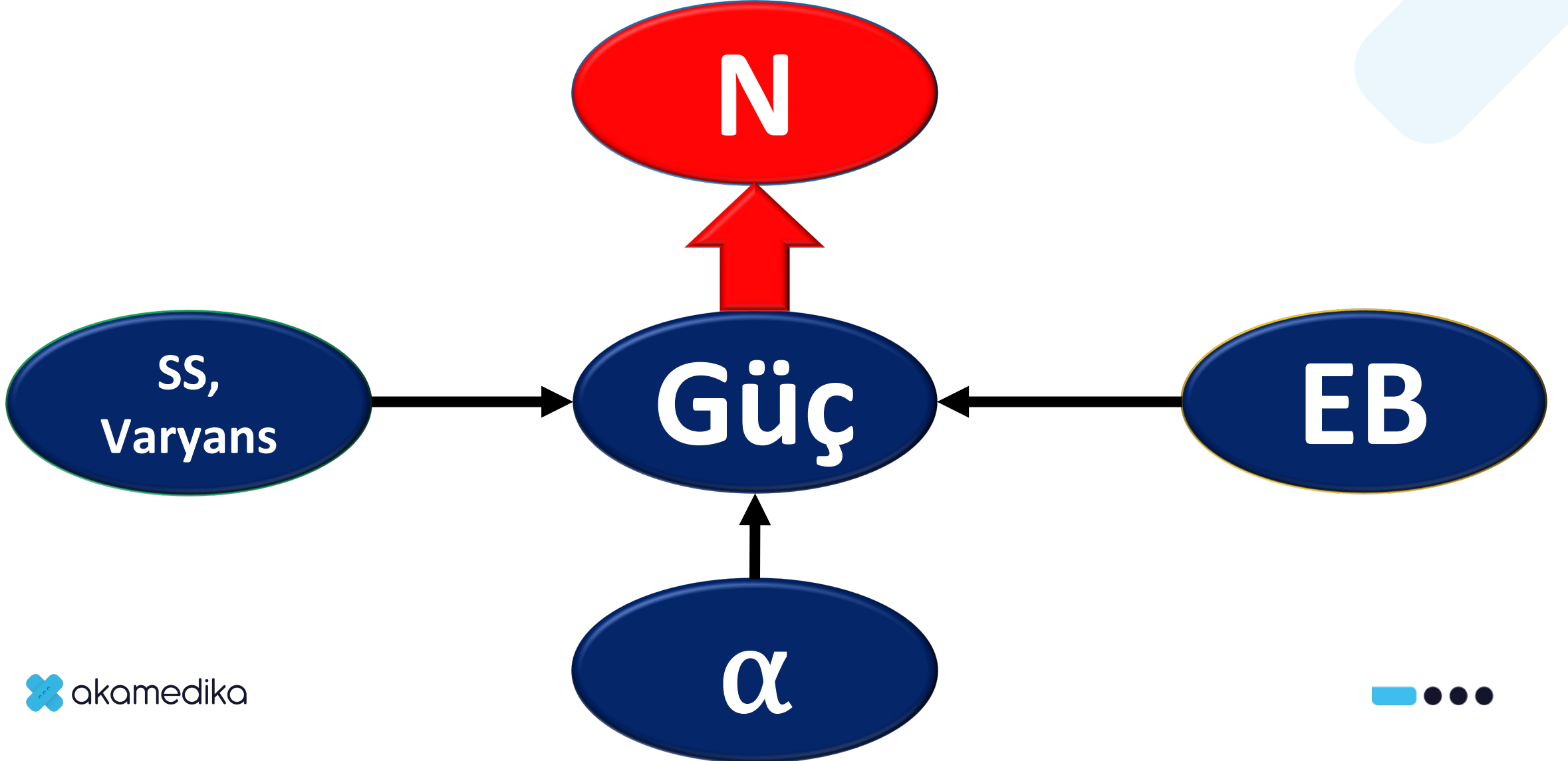
Çalışmanın Gücünü Belirleyen Parametreler



Power Analizi: Çalışmanızı yapıp elde ettiğiniz değerler ile çalışmanın gücünü hesaplayabilirsiniz...



Örneklem Büyüklüğü Hesabı: Çalışmanızı yapmadan önce ilgili değerleri önceden belirleyip çalışmanızın belirli bir güce ulaşması için gereken en az örneklem sayısını hesaplayabilirsiniz...



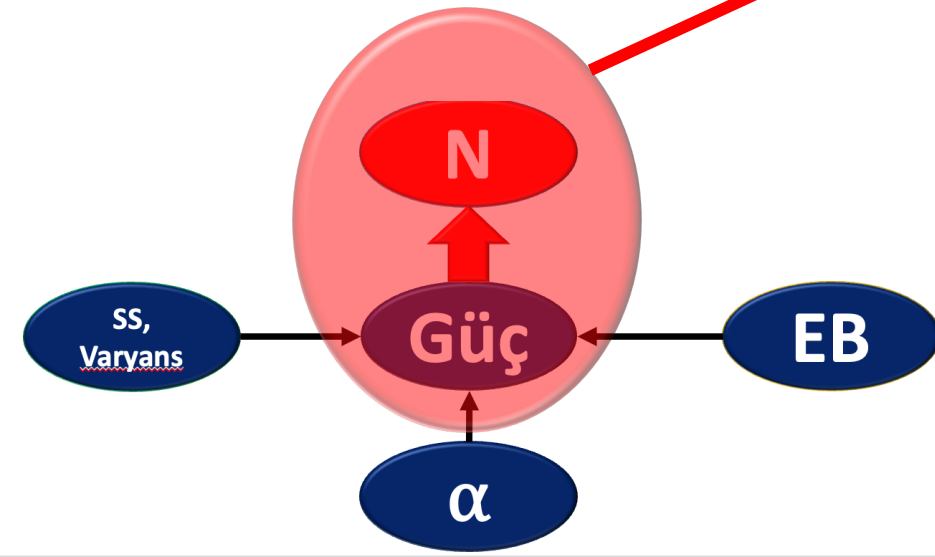
Çalışmanın Gücü (Power)

Güç: 1-Tip2 Hata

- Tanım olarak güç yani «power»; yapılan çalışmada fark olduğu varsayılan değişken açısından gerçekten de fark varken, var olan bu farkı doğru olarak tespit edebilme olasılığıdır.

Örneklem büyüklüğü hesaplamadan yapılan çalışmalarda asıl problem çalışmanızda araştırılan değişken açısından fark çıkmadığı zaman yaşanır.

Çünkü gerçekten mi fark yok yoksa beklenen değerin üzerinde TİP-2 Hata mı yaptınız emin olunamaz.



Sample size calculation and data analysis

A previous study reported a reduction in mortality from 15% to 8.6% with total-body CT scanning as the single diagnostic procedure during trauma evaluation as compared to historical control data [29]. Analysis on the large German polytrauma registration database performed by Huber-Wagner et al. showed a significant reduction in the 24-h mortality in patient who underwent immediate total-body CT compared to the conventional group (10% vs. 12%, $P = 0.038$) [25]. Historical AMC data show a mortality rate of 12% for trauma patients matching the current trial inclusion criteria. Based on the combination of the AMC data and the participation of the other trauma centers with comparable trauma populations, it is expected to find a reduction in mortality from 12% to 7%. The detection of such a difference requires 539 patients per group using a power of 80% and a two-sided alpha of 5%. Based on the historical and estimated data of the participating centers the inclusion period will take about 1,5 years and the follow-up period will take an additional year.

REACT-2 Çalışması

Conventional
with severe



G W Dijkgraaf, Jan S K Luitse,

Outcomes

The primary endpoint was **in-hospital mortality**, defined as mortality during the index hospital admission after trauma, including in patients who were transferred to

	Total-body CT		Standard work-up	
	Number of patients	Data	Number of patients	Data
Age (years)	541	42 (27-59)	542	45 (26-59)

Total-body CT		Standard work-up		p value
Number of patients	Data	Number of patients	Data	
86 (16%)		542	85 (16%)	0.92*

İki yaklaşım arasında fark bulamadık!!!

İki yaklaşım arasında fark olmadığını bulduk !!!

Sonucu nasıl yazalım?



Bir örnek; REACT-2 Çalışması; Okumaya devam!!!

Sample size calculation and data analysis

A previous study reported a reduction in mortality from 15% to 8.6% with total-body CT scanning as a diagnostic procedure during trauma care compared to historical control data from a large German polytrauma registry. The study was performed by Huber-Wagner et al. (2010) and showed a reduction in the 24-h mortality rate from 12% to 8.6% with immediate total-body CT compared to historical control data (10% vs. 12%, $P = 0.001$). The study also showed a mortality rate of 12% in the historical control group (10% vs. 12%, $P = 0.001$). In the current trial including the combination of the AMC data with data from other trauma centers with similar characteristics, it is expected to find a reduction in mortality from 12% to 7%. The detection of such a difference requires 539 patients per group using a two-sided alpha of 5%. Based on the historical data of the participating centers, the investigators take about 15% as the expected mortality rate. In addition,

da fark olmadığını bulduk !!!

ark olmadığını bulduk

**Peki ama nasıl
hesapladın?**

Örneklem büyüklüğü mortalitenin %5'ini hesaplanmış!!!

→ 881
→ 1636
%2 → 3837
%1 → 15957



Teşekkürler

Soru ve Katkılarınızı Bekliyoruz.



@drserefkerem

