

Daha ileriye... En iyiye...

Hipotez Testlerine Giriş



www.hacettepe.edu.tr

Dr. Kübra Atalay
Kabasakal

Son dersimizde neler yaptık

Örneklem ortalamalarının dağılımı

- Belirli bir örneklem büyüklüğünde (N), seçilebilecek tüm olası örneklemlerden elde edilen ortalamaların dağılımı.

Örnekleme Hatası

- Örneklem istatistiği ile evrendeki karşılığı olan parametre arasındaki farka *örnekleme hatası* denir

Merkezi Limit Teoremi

- Ölçülen özellik evrende normal dağıldığında veya örneklem yeterince büyük olduğunda ($N \geq 30$) Merkezi limit teoremine göre örneklem ortalamalarının dağılımı normal dağılıma yaklaşır.
- Bu dağılımın ortalaması evren ortalamasının beklenen değeridir
- Bu dağılımın standart sapması (σ / \sqrt{N}) standart hatadır.

HİPOTEZ TESTLERİ

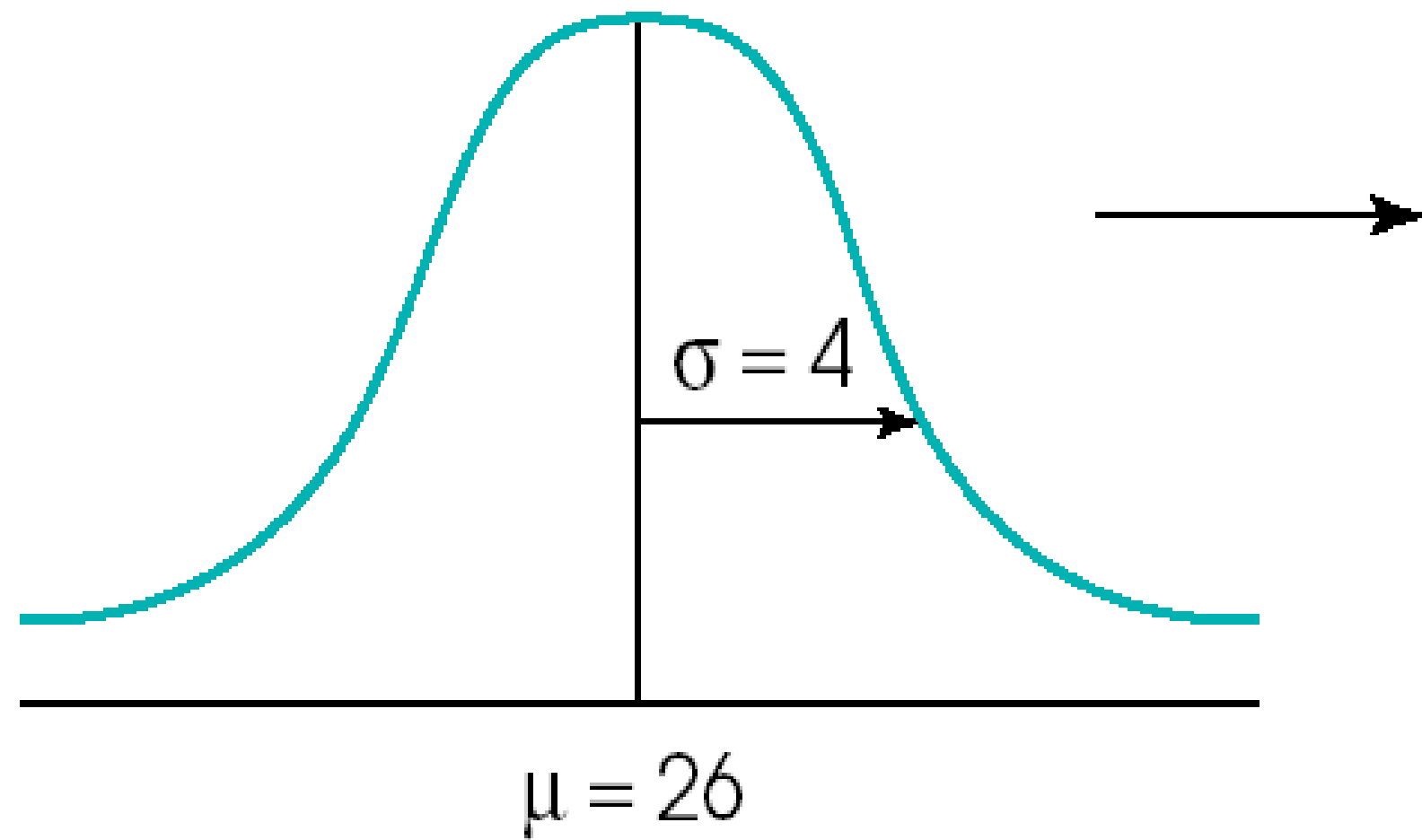
- Hipotez testi bir evren parametresi hakkında oluşturulan hipotezi örneklem verisi kullanarak test etmeye olanak veren bir istatistik yöntemidir.
- Gözlem ya da deneme sonucu elde edilmiş sonuçların, rastlantıya bağlı olup olmadığının incelenmesinde kullanılan istatistiksel yöntemlere HİPOTEZ TESTLERİ denir.
- Sonuçların rastlantıya bağlı olup olmadığı, evren parametreleri (ortalama, ortanca, varyans, vb.) üzerine kurulmuş hipotezlerin test edilmesi ile yapılır.

HİPOTEZ TESTİNİN MANTIĞI

1. Evrene ilişkin bir hipotez kurulur
2. Araştırma yapılır, örneklem seçilir ve veri toplanır
3. Örneklemden elde edilen veri ile evrene yönelik kurulan hipotez karşılaştırılır.
4. Karşılaştırma sonucunda bir karara varılır. Eğer örneklemden elde ettiğimiz veri kurduğumuz hipotez ile tutarlıysa hipotezimizi kabul ederiz. Eğer veriden elde ettiğimiz sonuç ile hipotezin arasında farklılık çok ise o zaman hipotezimizin yanlış olduğuna karar veririz.

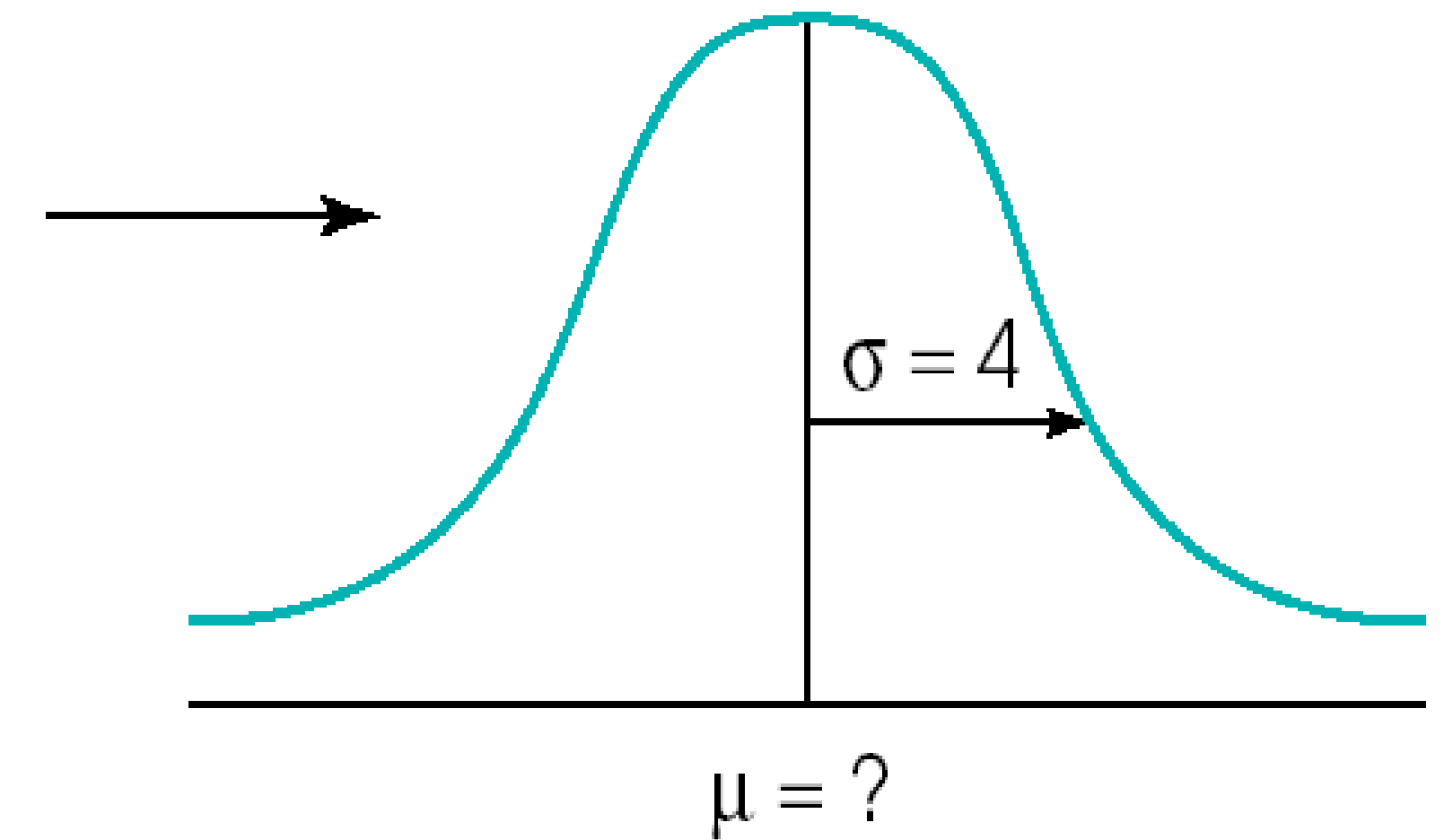
TEK ÖRNEKLEM Z TESTİ

- Deneysel bir araştırma deseni
Bilinen evren



Deneysel
İ etki/
Tedavi

Bilinmeyen evren



- Burada araştırmanın amacı deneysel etkinin evren ortalamasını değiştirip değiştirmediğini belirlemektir.
- Burada aslında şu soruyu sorarız: Eğer deney veya tedavi evrenin tamamına uygulansaydı ne olurdu?

- Z testinde başlangıçta evrenin ortalaması ve standart sapmasını biliriz

Z testinin varsayımları

- Örneklem rastgele seçilmiştir.
- Gözlemler bağımsızdır
- Dağılım normaldir (Not: örneklem 30 dan büyük olduğunda hipotez testleri normalliğin sağlanmamasına dayanıklıdır.)
- Uygulama evren varyansını değiştirmez

Araştırma Durumu:

Bir grup araştırmacı beyinin hafif düzeyde elektrik akımı ile uyarılmasının bireylerin matematik öğrenmelerini geliştirdiğini düşünmektedir (Kadosh et. al., 2010).

Araştırma kapsamında 25 öğrenci seçilmiştir. Bu öğrencilere soyut matematik, sayılar ve semboller konuları öğretilmiş ve öğrencilerin beyinlerine bu süreçte hafif düzeyde elektrik akımı verilmiştir.

Evrende ortalaması 80 ve standart sapması 20 olan standart bir matematik testi bu öğrencilere deneyden 4 hafta sonra uygulanmıştır. Araştırmanın sonucunda 25 öğrencinin bu testten aldıkları puanların ortalaması 90 olarak hesaplanmıştır.

Araştırmanın amacı: Araştırmanın amacı hafif düzeyde beyne verilen elektrik akımının öğrencilerin matematik performanslarına etkisinin incelenmesidir

Araştırma sorusu: Beyne verilen hafif elektrik akımı öğrencilerin matematik performanslarını etkilemekte midir?

HİPOTEZ TESTİNİN AŞAMALARI

1. *Evren dağılım özelliklerinin belirlenmesi*

- Evrende matematik testi ortalaması 80 standart sapması 20 olan normal dağılım göstermektedir.
- Evren ortalamasını ve standart sapmasını bildiğimiz için Z testi kullanmak için uygundur.

2. *Hipotezlerin kurulması*

- Sıfır hipotezi: Test edilen hipotez

$H_0 : \mu_{\text{deneysonrası}} = 80$ ---- H_0 : Elektrik akımının matematik performansını etkilemez

$H_1 : \mu_{\text{deneysonrası}} \neq 80$ ---- H_1 : Elektrik akımı matematik performansını etkiler

3. Karar Kuralının Belirlenmesi

Red bölgesi: Örneklem ortalamasının kurduğumuz hipotezleri sorgulamamıza neden olacak değerleridir.

Anlamlılık/Manidarlık Düzeyi: Hipotez testlerinde «olma ihtimali çok düşük» olarak tanımladığımız olasılık değeri. Bu değer birinci tip hataya eşittir. Alfa düzeyi olarak da isimlendirilir.

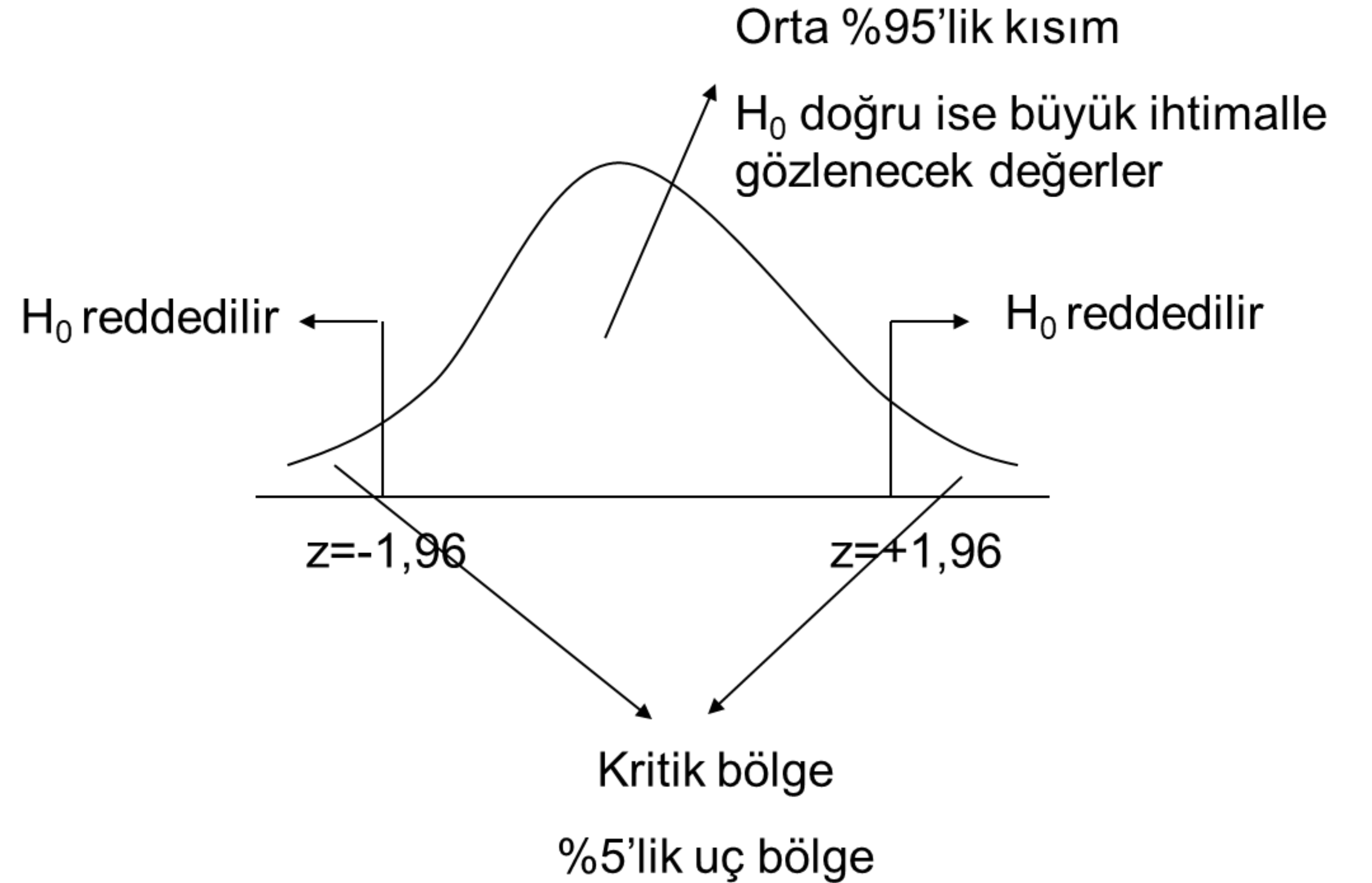
- H_0 doğru ise büyük ihtimalle elde edilebilecek örneklem ortalamaları



3. Karar Kuralının Belirlenmesi

- Alfa düzeyini $\alpha = 0.05$ olarak belirleyelim
- Yüksek olasılıklı ve düşük olasılıklı örneklem ortalamalarının sınırlarını belirleyelim

1.04	.9071	.0329	.4071
1.85	.9678	.0322	.4678
1.86	.9686	.0314	.4686
1.87	.9693	.0307	.4693
1.88	.9699	.0301	.4699
1.89	.9706	.0294	.4706
1.90	.9713	.0287	.4713
1.91	.9719	.0281	.4719
1.92	.9726	.0274	.4726
1.93	.9732	.0268	.4732
1.94	.9738	.0262	.4738
1.95	.9744	.0256	.4744
1.96	.9750	.0250	.4750
1.97	.9756	.0244	.4756
1.98	.9761	.0239	.4761
1.99	.9767	.0233	.4767



4. Örneklem İstatistiğini Hesaplama

Araştırmada N=25 öğrenciden veri toplandı ve elde edilen örneklem ortalaması $\bar{X} = 90$

Örneklem ortalaması için z istatistiği: $z = (\bar{X} - \mu) / SH$

- Z istatistiğini hesaplayabilmek için öncelikle standart hata hesaplanmalıdır.
- Merkezi limit teoreminden örneklem ortalamalarının standart hatası:

$$SE = \sigma / \sqrt{N} = 20 / \sqrt{25} = 4$$

$$z = (\bar{X} - \mu) / SH = (90 - 80) / 4 = 2.5 \text{ olur}$$

5. Karar verme

İki olası sonuç vardır:

1. Örneklem ortalamasına yönelik hesaplanan z değeri eğer kritik bölgedeyse H_0 reddedilir.
 - Alfa 0.05 anlamlılık düzeyi kullanıldığında sınır z değerleri -1.96 ve 1.96 olarak elde edilir. Eğer örneklemde elde edilen z istatistiği 1.96'dan büyük veya -1.96'dan küçükse H_0 reddedilir.
2. Örneklem ortalamasına yönelik hesaplanan z değeri eğer kritik bölgede değilse H_0 kabul edilir.
 - Alfa 0.05 anlamlılık düzeyi kullanıldığında sınır z değerleri -1.96 ve 1.96 olarak elde edilir. Eğer örneklemde elde edilen $-1.96 < z$ istatistiği < 1.96 aralığında ise H_0 kabul edilir.

Karar:

Deney sonrasında grubun test ortalaması 90 olarak elde edilmiştir. Bu puana karşılık gelen z değeri 2.5 olarak hesaplanmıştır. $z=2.5 > 1.96$ olduğundan z değeri red bölgesindedir. Dolayısıyla H_0 reddedilir.

6. *Sonuçların yorumlanması*

İstatistiksel anlamlılık: H_0 hipotezi doğru olduğunda bir sonucun elde edilme olasılığı çok düşükse, bu sonuç istatistiksel olarak anlamlıdır.

- Raporlama: Beyine hafif düzeyde elektrik akımı verilmesinin bireylerin matematik performansları üzerinde anlamlı bir etkisi vardır ($z=2.5$, $p < .05$).

6. Sonuçların yorumlanması

Pratik Anlamlılık / Etki büyüklüğü:

Hipotez testlerinde sonuca şu şekilde varılır

1. Standart hata hesaplanır. Standart hata aslında bize evren ortalaması (μ) ile örneklem ortalaması (\bar{X}) arasında kabul edilebilir farkın miktarını söyler.
2. Elde ettiğimiz farkın standart hatadan daha yüksek olduğunu göstermeye çalışırız. Yani elde edilen farkın hata ile açıklanamayacak kadar büyük olduğunu söyleriz. Dolayısıyla aslında bize sadece görel bir karşılaştırma verir.

Hipotez testleri bize asıl farkın ne kadar olduğuna ilişkin bir bilgi vermez

Örnek

Birinci Durum

Aynı çalışmayı 25 kişiyle yaptığımızı ve örneklem ortalamasını 81 olarak bulduğumuzu varsayalım

Bu durumda standart hatamız:

$$SE = \sigma / \sqrt{N} = 20 / \sqrt{25} = 4 \text{ olur}$$

Z değerimiz

$$z = (\bar{X} - \mu) / SH = (81 - 80) / 4 = 0.25 \text{ olur.}$$

Bu durumda $z=0.25$ kritik değer olan 1.96 dan küçük olduğundan H_0 kabul edilir

Deneyin bir etkisi olmamıştır.

İkinci Durum

Aynı çalışmayı 2000 kişiyle yaptığımızı ve örneklem ortalamasını 81 olarak bulduğumuzu varsayalım

Bu durumda standart hatamız:

$$SE = \sigma / \sqrt{N} = 20 / \sqrt{2000} = 0.45 \text{ olur}$$

Z değerimiz

$$z = (\bar{X} - \mu) / SH = (81 - 80) / 0.45 = 2.2 \text{ olur.}$$

Bu durumda $z=2.2$ kritik değer olan 1.96 dan büyük olduğundan H_0 reddedilir

Deneyin anlamlı bir etkisi olmuştur.

- İstatistiksel anlamlılık bize yalnızca H_0 doğru olduğunda elde ettiğimiz sonucun ortaya çıkmasının olasılığının çok düşük olduğunu söyler
- Farkın pratikte anlamlılığı hakkında bir bilgi vermez
- İstatistiksel olarak anlamlı olan sonuçlar her zaman pratikte anlamlı olmayabilir.
- Bu nedenle hipotez testleri yapıldığında her zaman etki büyüklüğü de raporlanmalıdır.
- Etki büyüklüğü bize farkın pratik anlamlılığına ilişkin bilgi veren ölçümlerdir.

Cohen d

- Yaygın olarak kullanılan bir etki büyüklüğü ölçümüdür
- Hesaplaması ve yorumlaması kolaydır

$$d = \frac{\text{Ortalamalar arası fark}}{\text{standart sapma}} = \frac{\mu_{\text{deneysonrası}} - \mu_{\text{deneyoncesi}}}{\sigma}$$

Z testi için etki büyüklüğü hesaplanması:

Deney sonrası evrenin ortalamasını bilmiyoruz. Bu nedenle evren ortalamasının bir kestirimi olarak örneklem ortalamasını kullanıyoruz.

$$\text{Kestirilen } d = \frac{\bar{X}_{\text{deneysonrası}} - \mu_{\text{deneyoncesi}}}{\sigma}$$

- Bizim orijinal örneğimiz için hesaplanan etki büyüklüğü:

$$d = \frac{90 - 80}{20} = 0.50$$

Etki büyüklüğünün yorumlanması	
d= 0.2	Küçük etki
d=0.5	Orta etki
d=0.8	Büyük etki

- Dikkat ederseniz etki büyüklüğü örneklem büyüklüğünden bağımsızdır.

(TEK) YÖNLÜ HİPOTEZ TESTLERİ

- Yönlü hipotez testinde (tek yönlü test) istatistiksel hipotezler (H_0 ve H_1) evren ortalamasının arttığını veya azaldığını belirtir. Yani, hipotezler uygulamanın etkisinin yönünü belirtir.
- Bazen bir deneye başlarken deneyin etkisine ilişkin bir öngörümüz olur.

Örneğin

- Performansın artması
- Şiddet eğilimli davranışların azalması
- Hipotezlerimizi de öngörümüze göre kurabiliriz.

- Bir önceki örneğimiz üzerinden gidersek eğer, deneyimizin sonucunda öğrencilerin matematik performanslarının artmasını bekleyebiliriz.

Bu durumda araştırma sorumuz:

Beyne uygulanan hafif elektrik akımı öğrencilerin matematik performanslarını **artırır** mı?

<u>Sıfır hipotezi:</u>	$H_0: \mu_{\text{deney sonrası}} \leq 80$ H_0 : Elektrik akımı öğrencilerin performanslarını iyileştirmez
<u>Alternatif hipotez:</u>	$H_1: \mu_{\text{deney sonrası}} > 80$ H_1 : Elektrik akımı öğrencilerin performanslarını iyileştirir.

Karar kuralının belirlenmesi

- Kritik bölge sıfır hipotezinin doğru olduğu durumlarda olasılığı düşük örneklem ortalamalarından oluşur. Bu sebeple, yönlü hipotezde kritik bölge dağılımın tek bir tarafındaki uç bölgedir.
- Alfa değeri olarak 0.05 belirleyelim
 - Dağılımın bir ucunda % 5 lik kısmı ayıran kritik z değeri belirlenir
- Dikkat ederseniz dağılımının sağ tarafında kritik bölge belirledik. Çünkü bizim için önemli olan ortalamanın artması durumu
- Eğer öngörümüz ortalamanın azalması yönünde olsaydı bu kez dağılımın sol kuyruğunda kritik bölge

(A) z	(B) Proportion in Body	(C) Proportion in Tail	(D) Proportion Between Mean and z
1.50	.9332	.0668	.4332
1.51	.9345	.0655	.4345
1.52	.9357	.0643	.4357
1.53	.9370	.0630	.4370
1.54	.9382	.0618	.4382
1.55	.9394	.0606	.4394
1.56	.9406	.0594	.4406
1.57	.9418	.0582	.4418
1.58	.9429	.0571	.4429
1.59	.9441	.0559	.4441
1.60	.9452	.0548	.4452
1.61	.9463	.0537	.4463
1.62	.9474	.0526	.4474
1.63	.9484	.0516	.4484
1.64	.9495	.0505	.4495
1.65	.9505	.0495	.4505
1.66	.9515	.0485	.4515
1.67	.9525	.0475	.4525
1.68	.9535	.0465	.4535
1.69	.9545	.0455	.4545
1.70	.9554	.0446	.4554

- Burada kritik z değerimiz 1.65.
- Örneklemden hesaplayacağımız z istatistiği eğer 1.65'den büyük olursa H_0 reddedilir.
- Örneklemden hesaplayacağımız z istatistiği eğer 1.65'den küçük olursa H_0 Kabul edilir

Örneklem İstatistiğini Hesaplama

Bu kısım yönsüz testler ile aynı şekilde hesaplanır

Araştırmada N=25 öğrenciden veri toplandı ve elde edilen örneklem ortalaması $\bar{X} = 90$

Örneklem ortalaması için z istatistiği: $z = (\bar{X} - \mu) / SH$

- Z istatistiğini hesaplayabilmek için öncelikle standart hata hesaplanmalıdır.
- Merkezi limit teoreminden örneklem ortalamalarının standart hatası:

$$SE = \sigma / \sqrt{N} = 20 / \sqrt{25} = 4$$

$$z = (\bar{X} - \mu) / SH = (90 - 80) / 4 = 2.5 \text{ olur}$$

Karar:

Deney sonrasında grubun test ortalaması 90 olarak elde edilmiştir. Bu puana karşılık gelen z değeri 2.5 olarak hesaplanmıştır. $z=2.5 > 1.65$ olduğundan z değeri red bölgesindedir. Dolayısıyla H_0 reddedilir.

Yorum:

Beyine hafif düzeyde elektrik akımı verilmesinin bireylerin matematik performanslarını anlamlı düzeyde artırmıştır($z=2.5, p < .05$).

Yönlü ve Yönsüz Hipotez Testlerinin Karşılaştırılması

- Yönlü ve yönsüz hipotez testi arasındaki en büyük fark H_0 'ı reddetmek için kullanılan kriterdedir.
- Örneğimizde hatırlarsanız yönsüz hipotez testi için kritik değerimiz 1.96 iken yönlü testte 1.65 'dir.
- **Yönlü hipotezde** H_0 'ı reddetmek için evren ile örneklem arasında belirtilen yönde ve görelî olarak **daha küçük** bir farkın bulunması yeterlidir.
- **Yönsüz hipotez** testinde ise H_0 'ı reddetmek için etkinin yönüne bağılı olmayan nispeten **daha büyük** bir fark gerekmektedir.

- Bazı araştırmacılar yönsüz testin daha kuvvetli ve ikna edici olduğunu savunurlar çünkü H_0 'ı reddetmek için daha çok kanıta ihtiyaç vardır.
- Bazı araştırmacılar ise yönlü testi savunurlar çünkü bu test daha duyarlıdır. Uygulamanın etkisi küçük olsa bile bu saptanabilir.
- Yönlü test uygulamanın etkisi hakkında kuvvetli bir beklenti olduğu durumlarda kullanılmalıdır.
- Uygulamanın etkisinin yönü ile ilgili farklı teoriler olduğu zaman yönsüz testler kullanılmalıdır.
- Yönlü test hiçbir zaman sonucu anlamlı çıkmayan yönsüz bir testten sonra ikinci bir deneme olarak kullanılmamalıdır.

İSTATİSTİKSEL GÜÇ

- Hipotez testinin doğru bir şekilde gerçekten var olan uygulama etkisini bulma olasılığı kaçtır?

Bir hipotez testinde olası kararlar:

		H0	
		Doğru	Yanlış
Karar	H0 reddedilir	Tip I hata(α)	
	H0 kabul edilir		Tip II hata(β)

- Bir istatistik testinin gücü, yanlış olan sıfır hipotezini doğru bir şekilde reddetme olasılığıdır. Testin gücü, gerçekten var olan uygulama etkisini ortaya çıkarma olasılığıdır.

$$\text{Güç} = p(\text{yanlış } H_0 \text{'ı reddetme}) = 1 - \beta$$

- Testin gücü uygulamanın etkisinin büyüklüğüne bağlıdır.
- Uygulamanın büyük bir etkisi varsa, bu etkiyi ortaya çıkarmak kolaydır, böylece testin gücü de yüksektir.
- Uygulamanın küçük bir etkisi varsa, bu etkiyi ortaya çıkarmak zordur, bu sebeple testin gücü de düşüktür.

- İstatistiksel güç genel olarak bir araştırmayı yapmadan önce hesaplanır.
- Bu yolla sonuçların anlamlı farklılık gösterme olasılığı hesaplanır.
- İstatistiksel gücün hesaplanması için aşağıdaki özellikleri bilmemiz gerekir.
 - Örneklem büyüklüğü
 - Gerçek deneysel etkinin büyüklüğü
 - Alfa düzeyi
 - Hipotez testinin türü (yönlü veya yönsüz)

Örnek:

Evrende ortalaması 80 standart sapması 10 olan ve normal dağılan bir değişken olsun. Bu evrenden $N=25$ kişilik bir örneklem seçelim ve bu örnekleme bir deney uygulayalım. Deneyin sonucunda kişilerin puanlarında 8 puanlık bir artış öngörelim.

- Örneklem büyüklüğü: 25
- Alfa= 0.05
- Hipotez testi yönsüz
- Gerçek etki: 8 puan

- Hipotezlerimiz
 - $H_0: \mu_{deneysonrası} = 80$ (*Deneyin bir etkisi yoktur*)
 - $H_1: \mu_{deneysonrası} \neq 80$ (*Deneyin bir etkisi vardır.*)
- Olası sonuçlar:
 - Sıfır hipotezi doğrudur. Deneyin puanlar üzerinde etkisi yoktur. Deney yapıldıktan sonra da evren ortalaması 80'dir.
 - Sıfır hipotezi yanlıştır. Deneyin puanlar üzerinde etkisi vardır. Deney yapıldıktan sonra evren ortalaması artık 88'dir. (Hatırlayalım 8 puanlık bir etki öngörmüştük)

- Alfa değeri 0.05 olarak tanımlanmıştı. Dolayısıyla kritik z değerleri 1.96 ve -1.96 olur.
- Standart hata (Örneklem dağılımının standart sapması) ise $\sigma/\sqrt{N} = 10/\sqrt{25} = 2$ dir.

$$z = (\bar{X} - \mu) / SH$$

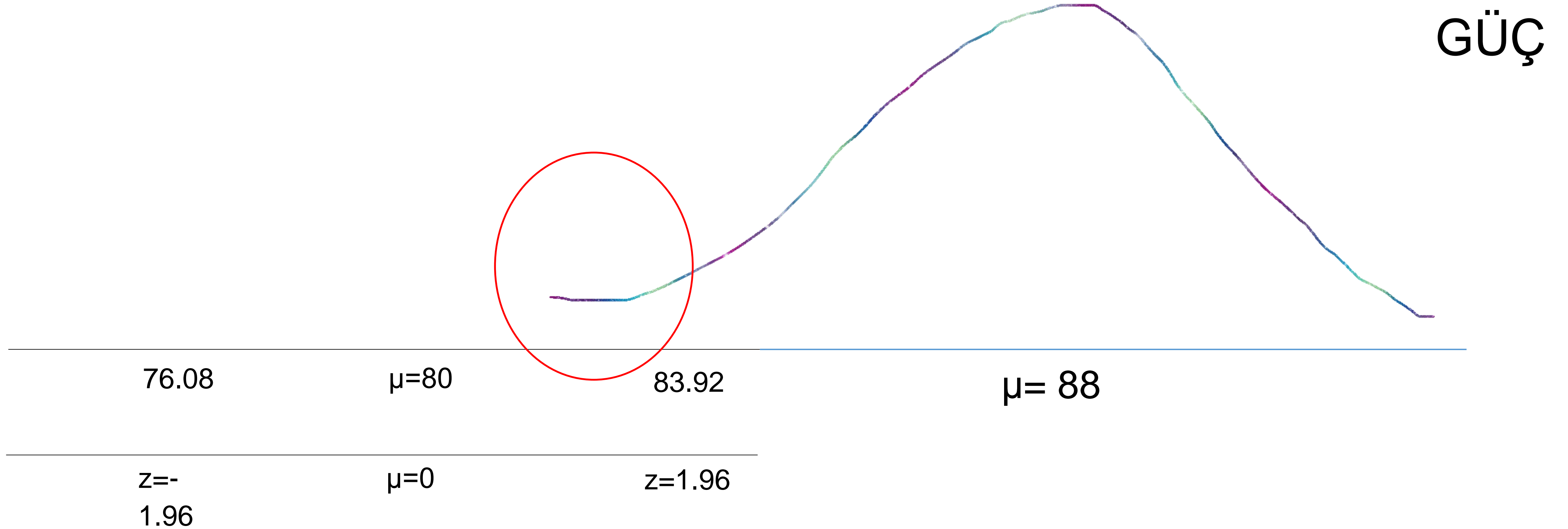
$$1.96 = (\bar{X} - 80) / 4$$

$$\bar{X} = 83.92$$

z=-	$\mu=0$	z=1.96
1.96		
76.08	$\mu=80$	83.92

H₀ doğru ise örneklem ortalamalarının dağılımı

H₀ yanlış ise, gerekte 8 puanlık bir etki varsa örneklem ortalamalarının dağılımı



- Ortalaması 88 standart sapması 2 olan örneklem ortalamaları dağılımında (deneye maruz kalmış hipotetik olan evren) $\bar{X} = 83.92$ ye karşılık gelen z değerini hesaplayalım.

$$z = (\bar{X} - \mu) / SH = (83.92 - 88) / 2 = -2.04$$

- Birim normal dağılım tablosundan bu bölgenin alanını bulalım

Z=-
2.04

(A) z	(B) Proportion in Body	(C) Proportion in Tail	(D) Proportion Between Mean and z
2.00	.9772	.0228	.4772
2.01	.9778	.0222	.4778
2.02	.9783	.0217	.4783
2.03	.9788	.0212	.4788
2.04	.9793	.0207	.4793
2.05	.9798	.0202	.4798
2.06	.9803	.0197	.4803
2.07	.9808	.0192	.4808
2.08	.9812	.0188	.4812
2.09	.9817	.0183	.4817
2.10	.9821	.0179	.4821
2.11	.9826	.0174	.4826
2.12	.9830	.0170	.4830
2.13	.9834	.0166	.4834
2.14	.9838	.0162	.4838
2.15	.9842	.0158	.4842
2.16	.9846	.0154	.4846
2.17	.9850	.0150	.4850
2.18	.9854	.0146	.4854
2.19	.9857	.0143	.4857
2.20	.9861	.0139	.4861

- Bu tabloda yer alan B sütunu bize taralı bölgenin alanını verir: 0.9793
- Testin gücü %97.93'tür
- Eğer araştırmacı 25 kişilik bir örneklem seçerse ve deneyin gerçekten 8 puanlık bir etkisi varsa, yapılacak olan hipotez testlerinin %97.93' ünde anlamlı bir farklılık bulunacaktır.

- İstatistiksel güce ilişkin görsel anlatım için aşağıdaki linki inceleyebilirsiniz. (Link İngilizcedir.)

<http://rpsychologist.com/d3/NHST/>

- Ayrıca Türkçe kaynak olarak Çapık (2014)'e bakabilirsiniz.