

İSTATİSTİKSEL TAHMİNLEME

Nokta Tahmini

Popülasyon parametresinin tek bir tahmin değerini verir

$$\bar{X} = \hat{\mu}$$

$$s = \hat{\sigma}$$

$$p = \hat{P}$$

Aralık Tahmini

Popülasyon parametresinin tahmin aralığını verir. Nokta tahmini kullanılarak hesaplanır.

$$20 \leq \mu \leq 60$$

$$2.5 \leq \sigma^2 \leq 3.4$$

$$0.25 \leq P \leq 0.35$$

Örnekleme Dağılımı

Toplumda sağa çarpık bir dağılım gösteren ve toplum ortalaması $\mu = 30.303$ ve toplum standart sapması $\sigma = 30.334$ olan bir değişken dizişinelim.

Toplum ortalaması $\mu = 30.303$

Toplum standart sapması $\sigma = 30.334$

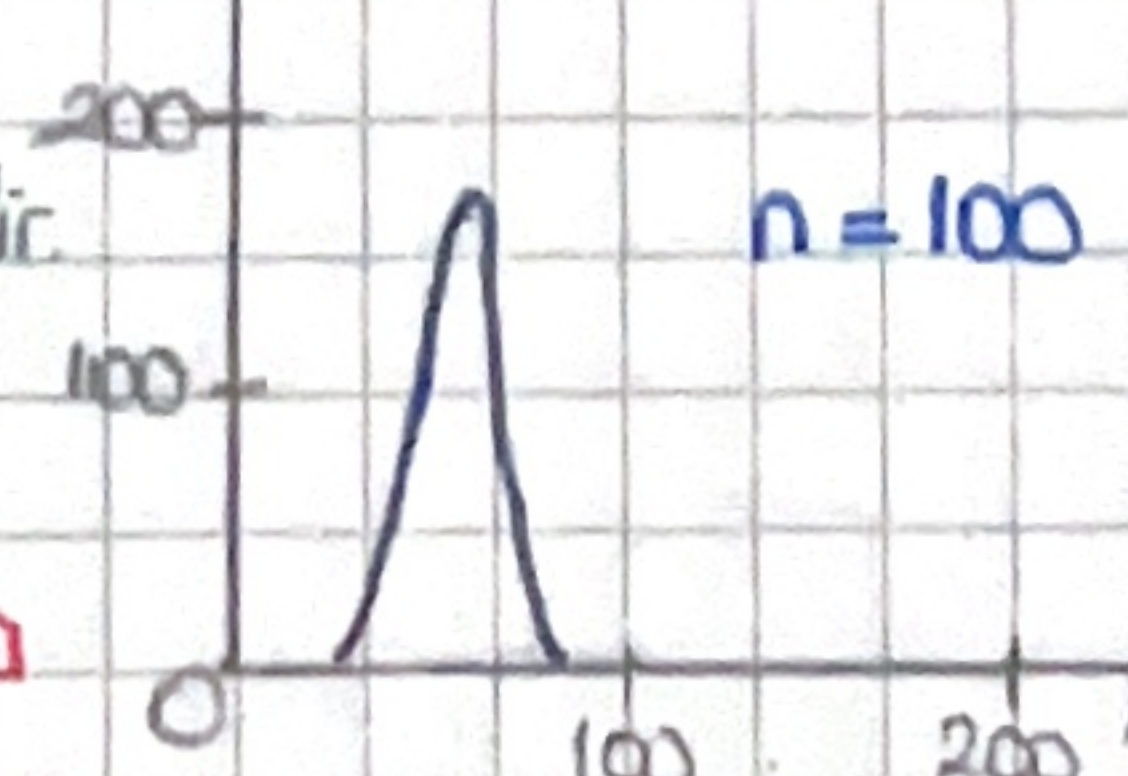
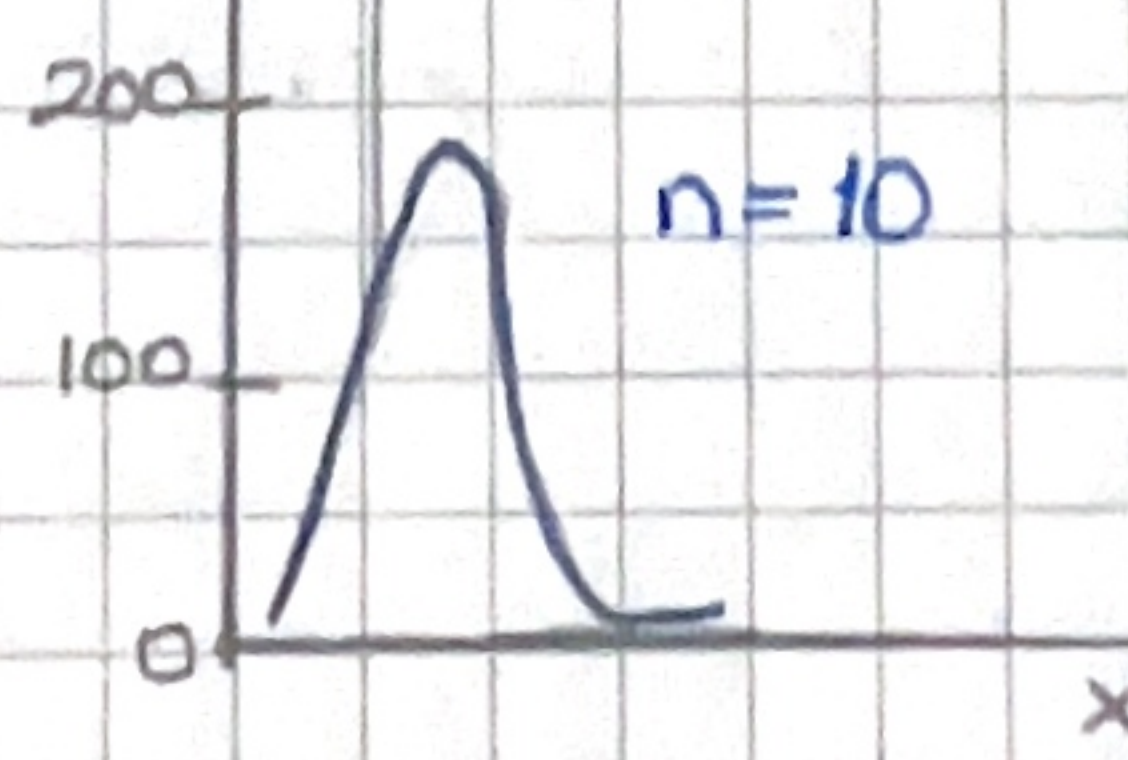
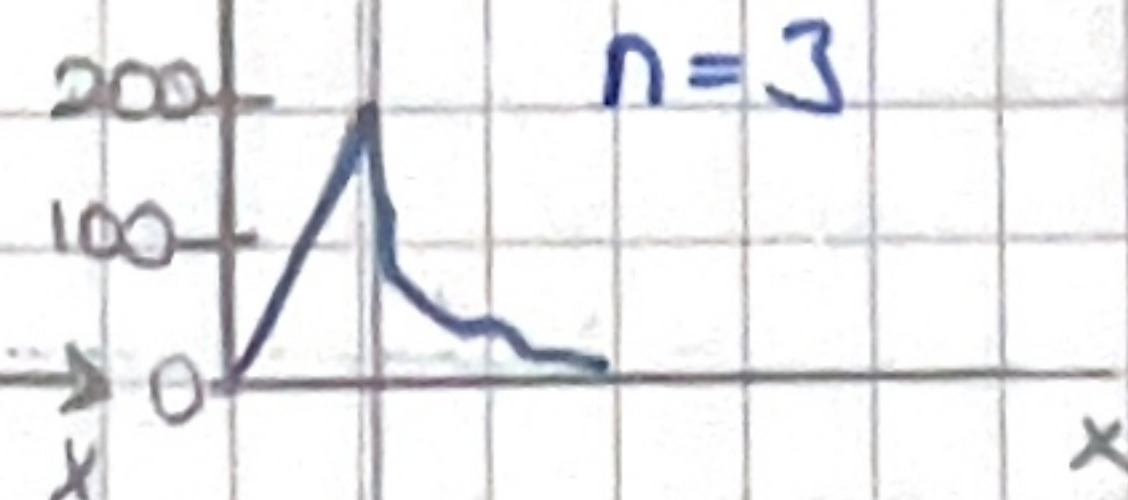
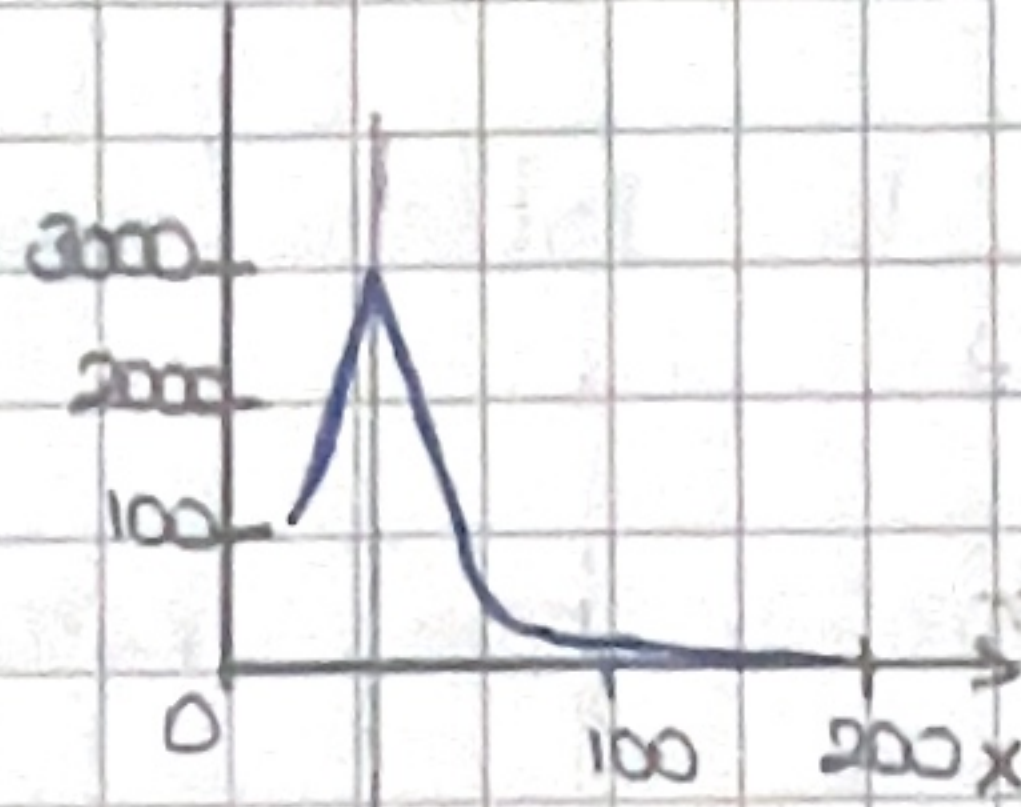
600 tane $\bar{X}$ için

$$\bar{X} = 30.68, s = 17.60, n = 3$$

$$\bar{X} = 30.23, s = 9.13, n = 10$$

$$\bar{X} = 30.331, s = 3.05, n = 100$$

> Grafikler incelendiğinde, örnek hacmi arttıkça ortalamaların örneklem dağılımının ortalaması toplum ortalamasına yakınlasmakta, standart sapması ise küçülmektedir. Ortalamaların örneklem dağılımının standart sapmasına **standart hata** adı verilir.



$\mu = 30.303$ Örnek Ortalamaları

3. Hafta

Hipotezlerin Oluşturulması (Sınav ?)

(n) sayıdaki, örneklemden elde edilmiş istatistikleri kullanarak topluma örneklem ilişkin parametre tahminleri yapmak, bu parametrelere ilişkin kurulan hipotezleri test ederek doğru kararlara ulaşmaktır.

H_0, H_1

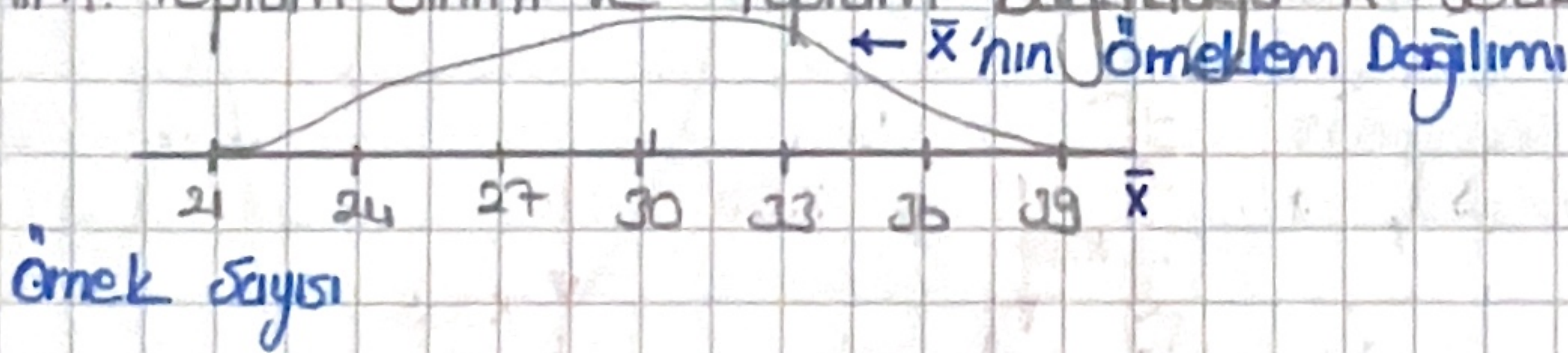
$H_0 : \mu = 45 \rightarrow$ çift kuyruk

$H_1 : \mu > 45$
 $< 45 \rightarrow$ tek kuyruk
 $\neq 45$

2. Hafta Devamı:

GÜVEN ARALIĞI NASIL YORUMLANIR?

Örnek: Toplum ortalaması 30.303 için $n=100$ birimlik 600 adet örnek seçelim ve her bir örnek için %99 güven sayısına sahip güven aralıklarını hesaplayalım. Toplum sınırlı ve toplum büyüklüğü $N=8042$ olsun.

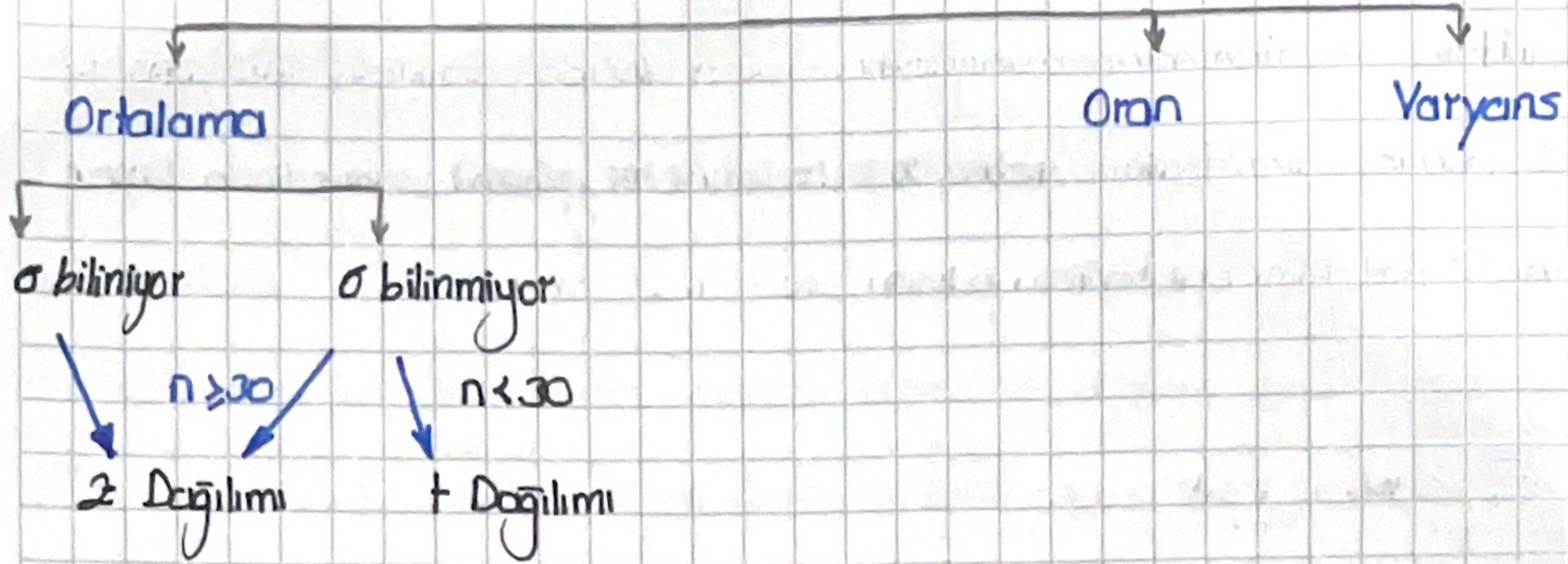


Elde edilen sonuçlara göre sadece 39. örneğe ait hesaplanan güven aralığı toplum ortalaması $\mu = 30.303$ 'ü kapsamamaktadır. Bu durumda hesaplanan 600 tane güven aralığından sadece 1 tanesi toplum ortalamasını kapsamamaktadır. Kapsayın güven aralığı oranı ise $599/600 = 0.9983$ olarak elde edilir.

Bu durumda hesaplanan güven aralığı; örneğin %99 güven katsayısı ile hesaplanmış ise şu şekilde yorumlanır: Elde edilen bu güven aralığı, toplum ortalamasını %99 olasılıkla kapsayan bir aralıktır. Toplum ortalaması %99 olasılıkla bu aralıktadır demek YANLIŞ bir yorumdur.

GÜVEN ARALIĞI TAHMİNLERİ

Güven Aralıkları

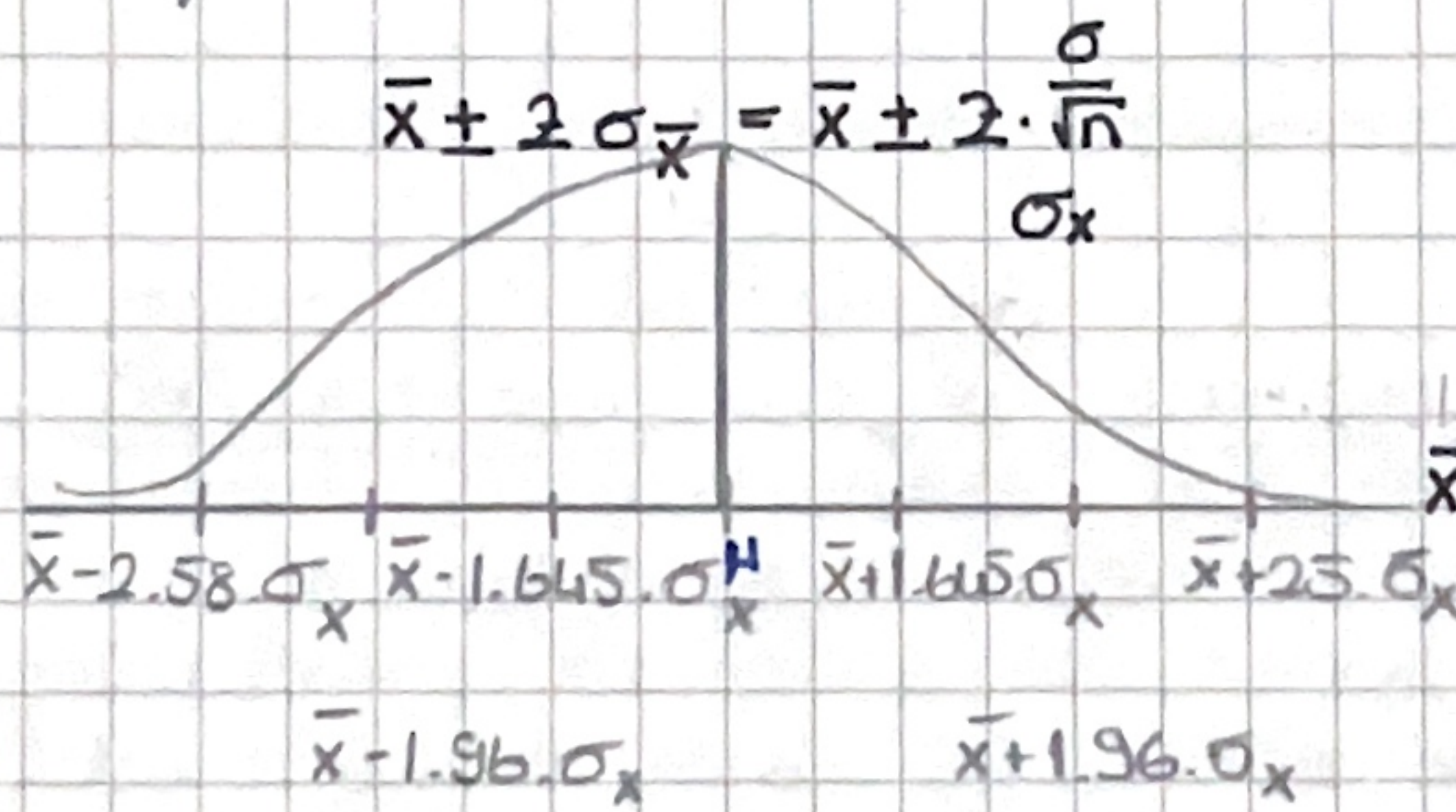


ORTALAMALAR İÇİN GÜVEN ARALIĞI

Gerçek kitle ortalaması, $1-\alpha$ güven seviyesinde

$$P\left(\bar{X} - z_{\alpha/2} \frac{\sigma_x}{\sqrt{n}} \leq \mu_x \leq \bar{X} + z_{\alpha/2} \frac{\sigma_x}{\sqrt{n}}\right) = 1-\alpha$$

aralığında yer alır.



$$\bar{x} \pm (2 \sigma_{1/2}) \cdot \frac{\sigma}{\sqrt{2}}$$

Örneklerin %90 $\rightarrow 1.645$

Örneklerin %95 $\rightarrow 1.96$

Örneklerin %99 $\rightarrow 2.58$

ÖRNEK:

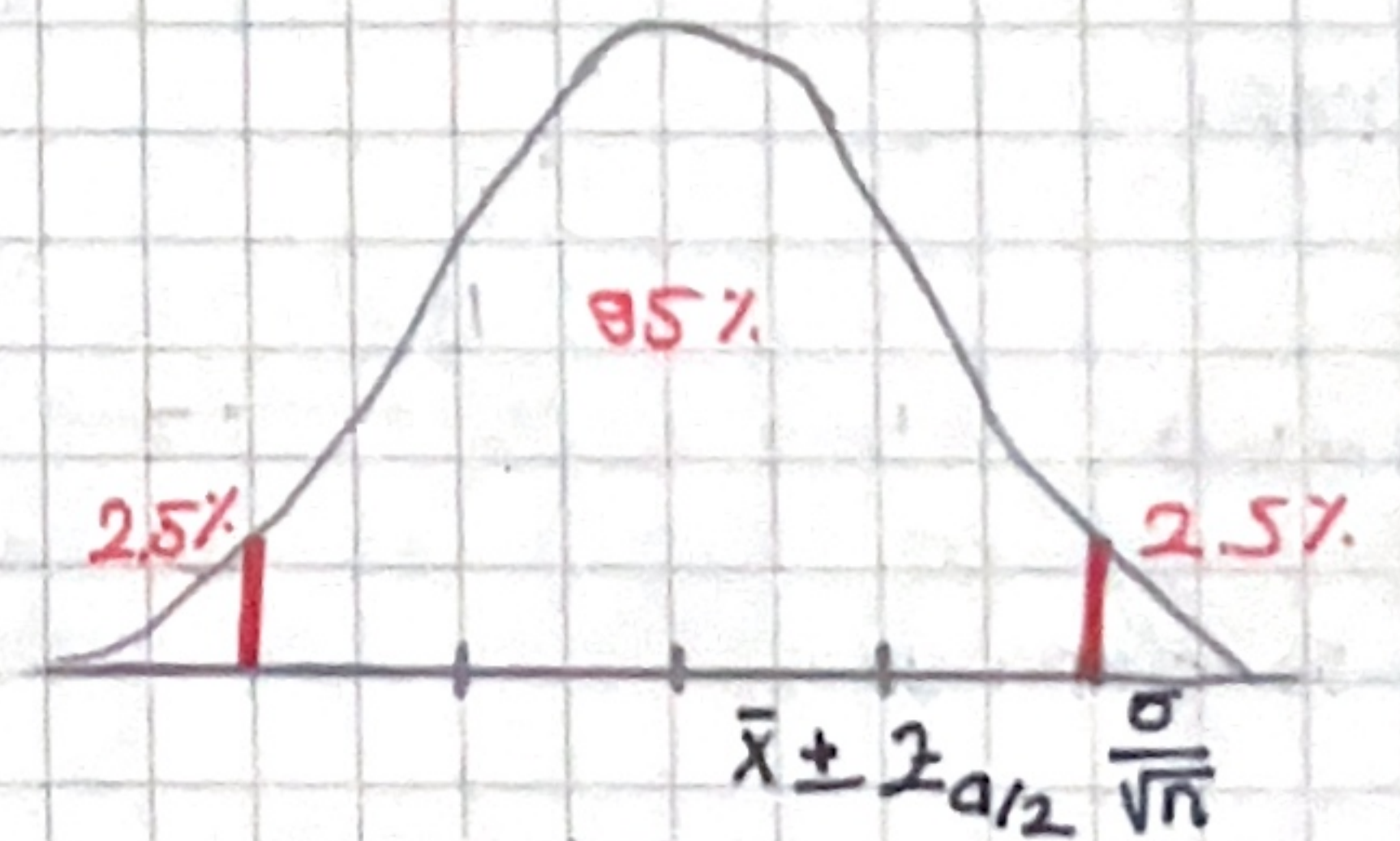
Standart sapması 100 olarak bilinen 25 diabet hastasının örneklem kan şekeri ortalaması 520 olarak bulunmuştur. %95 güven aralığını göre ortalamayı ifade ederiz.

! Hesaplamayı sonucu diyebiliriz ki %95

güvenle ve %5 hata payı ile hastaların

kan şekeri değeri 480.8 ile 559.2 arasında

değişmektedir.



$$520 - (1.96) \frac{100}{\sqrt{25}} = 480.8$$

$$520 + (1.96) \frac{100}{\sqrt{25}} = 559.2$$

3. HAFTA

ÖRNEK: Bir tesgahta üretilen parçaların dış çaplarının standart sapması $\sigma = 2.4$ cm'dir. Tesgahtın üretiminde rastgele seçilen 16 parçanın dış çap ortalaması 3.2 cm olarak bulunmuştur. %5 hata (%95 güven) seviyesinde kitle ortalamasının güven aralığını tahmin ediniz.

CEVAP:

$$3.2 - 1.96 \cdot \frac{2.4}{\sqrt{16}} < \mu < 3.2 + 1.96 \cdot \frac{2.4}{\sqrt{16}}$$

$$2.024 < \mu < 4.376$$

3.2

ÖRNEK: Bir fabrikada üretilen 100 ürünün ortalama ağırlığı 1040 gr, standart sapması 25 gr bulunmuştur. Bu imalat sürecinde üretilen ürünlerin ortalama ağırlığı %95 güvenle hangi aralıktadır?

CEVAP:

$n = 100$ %95 güven aralığı $1040 - 1.96 \cdot \frac{25}{\sqrt{100}} < \mu < 1040 + 1.96 \cdot \frac{25}{\sqrt{100}}$
 $\bar{x} \rightarrow 1040$
 $\sigma = 25$ gr $1035.1 < \mu < 1044.9$

ÖRNEK: Standart sapması 4 olarak bilinen 100 kişinin Hb değeri ortalaması 13.6 olarak bulunmuştur. %95 güven aralığını göre ortalamayı ifade ediniz.

CEVAP:

$\sigma = 4$ $13.6 - 1.96 \cdot \frac{4}{\sqrt{100}} < \mu < 13.6 + 1.96 \cdot \frac{4}{\sqrt{100}}$
 $n = 100$
 $\bar{x} \rightarrow 13.6$ $12.816 < \mu < 14.384$
%95 güven aralığı

HİPOTEZLERİN OLUŞTURULMASI

→ Biyostatistiksel yöntemlerin temel amaçlarından birisi de n sayıdaki örneklerden elde edilmiş istatistikleri kullanarak topluma ilişkin parametre tahminleri yapmak, bu parametrelere ilişkin kurulan hipotezleri test ederek doğru kararlara ulaşmaktır.

→ Toplumda bir değişkenin parametrelerine, dağılım yapmasına ya da ilişki düzeyine ilişkin kurulan hipotezlerin denetlenmesi için yararlanılan yöntemlere **hipotez testleri** denir.

Hipotez: Kuramsal olarak varsayılan ya da önceden yapılmış bir dizi gözleme dayanarak ortaya atılan gerçekleşmesi mümkün olan önermedir.

→ Örnek veriler, toplum ortalaması ya da başka bir parametre hakkında sonuç çıkarmak için hipotez testlerinde kullanılır.

→ Hipotez testlerinde genel olarak parametrenin;

- Belli bir değere eşit olup olmadığı,
- Belli bir değerden az ya da yüksek olup olmadığı,
- Birden fazla toplum σ^2 konusu ise incelenen değişkene ait parametrenin toplumlara göre değişip değişmediği vb. şekilde kurulan hipotezler örnek veriler ile test edilir.

HİPOTEZ TESTLERİ

Hipotezler

H_0 : Test hipotezi, kontrol hipotezi, sıfır hipotezi veya yokluk hipotezi olarak adlandırılır. Bu hipotezde araştırılan konuya bağlı olarak ele alınan örneğin popülasyona ait olduğu, yani belirtilen popülasyonu temsil ettiği ileri sürülür.

H_1 : Karşıt hipotez veya alternatif hipotez olarak adlandırılır. Bu hipotezde ise H_0 hipotezinin aksine örneğin popülasyona ait olmadığı, bu popülasyondan başka bir popülasyonu temsil ettiği ifade edilir.

★ Bu iki durumun dışındaki başka bir durum σ^2 konusu değişebilir.

→ Hipotezler genel olarak, yığınların karakteristikleri ile ilgili önermelerdir. Bu karakteristikler genel olarak parametrelerdir.

→ **HİPOTEZ**: parametre hakkındaki bir inanıştır. Parametre hakkındaki inancı test etmek için hipotez testi yapılır.

→ Hipotez testleri sayesinde örneklerden elde edilen istatistikler aracılığıyla yığın parametreleri hakkında karar verilir. Yığın parametreleri hakkında karar verirken doğru ya da yanlış olması muhtemel yargılardan hareket edilir.

→ **İSTATİSTİKSEL HİPOTEZ**: Bir ya da daha fazla kitleye ilişkin olarak ileri sürülen ve teorik bir dağılım varsayımı altında, parametrik değerin belirli bir değere eşit olduğunu veya iki ya da daha fazla ana kütle parametresinin birbirlerine eşit olduğunu belirten ve geçerliliği istatistiksel testlerle denetlenen bir önermedir.

→ Araştırmacıların yığın parametresi ya da yığınların parametreleri hakkında ortaya attığı iddialar (önermeler) genellikle alternatif hipotez veya karşıt hipotez olarak bilinir ve H_1 notasyonu ile gösterilir. Karşıt hipotez, incelenen değişkenin toplumdaki parametresinin değiştiği, belirli bir değere eşit olmadığı vb. şekilde formüle edilir.

→ **Örneğin**, Konya'da KTO Karatay Üniversitesi'nde okuyan öğrencilerin okula gidiş dönüşte harcadığı ortalama süre 45 dakikadan fazladır, iddiası için alternatif hipotez $H_1: \mu > 45$ olarak ifade edilir.

örnek: Gebelikte sigara kullanımının düşük doğum ağırlığına neden olduğu araştırılmak isteniyor. Araştırma öncesinde ise gebelikte sigara kullanan kadınların 2500 gr'dan daha düşük ağırlığı sahip doğum yaptıkları ileri sürülüyor.

H_0 : Sigara kullanımı düşük doğum ağırlığına neden olmaz.

H_1 : Sigara kullanımı düşük doğum ağırlığına neden olur.

$$H_0: \mu = 2500$$

$$H_1: \mu < 2500$$

örnek: Sigara ve alkol kullananların sistolik kan basıncalarının normal seviye olarak kabul edilen 120 mm/Hg'den daha yüksek olduğu, sigara ve alkol kullanımının hipertansiyon hastalığının en önemli etkenleri arasında yer aldığı iddia ediliyor.

H_0 : Sigara ve alkol kullananların sistolik kan basıncı ortalaması normaldir.

H_1 : Sigara ve alkol kullananların sistolik kan basıncı ortalaması yüksektir.

$$H_0: \mu = 120$$

$$H_1: \mu > 120$$

→ Hipotez kontrolü iki şekilde yapılır.

1. Tek Taraflı Hipotez Kontrolü: Bu hipotez kontrolünde popülasyonun sadece sağ tarafı veya sadece sol tarafıyla ilgilenilir. Hipotezin tek taraflı mı yoksa çift taraflı mı olduğu, ilgilenilen amacın son cümlesinden anlaşılır. Eğer amacın sonunda "artmış mıdır, artırmış mıdır, azalmış mıdır, azaltmış mıdır, önerir misiniz" soruları varsa test tek taraflı yapılabilmek demektir.

2. Çift Taraflı Hipotez Kontrolü: Bu tip hipotez kontrolünde dağılımın her iki tarafıyla ilgilenilir. Eğer amacın "değiştirmiş midir, değişmiş midir, farklılık tesadüften mi ileri gelmektedir" şeklinde sorular varsa hipotez kontrolü çift taraflı yapılacaktır.

→ Uygulamalarda daha çok çift yönlü hipotezlere rastlanmaktadır.

H_0 : İki grubun ortalaması arasında fark yoktur.

Yaş ve boy değişkenleri ilişkisizdir.

A ilacı ile B ilacının iyileştirme yüzdeleri arasında fark yoktur.

H_1 : İki grubun ortalaması farklıdır.

Yaş ve boy değişkenleri ilişkilidir.

A ilacı ile B ilacının iyileştirme yüzdeleri arasında fark vardır.

Örnek 1

Örnek 2

Örnek 3

$H_0: \mu = 2500$

$H_0: \mu = 7$

$H_0: \mu = 120$

$H_1: \mu < 2500$

$H_1: \mu \neq 7$

$H_1: \mu > 120$

! Herhangi bir değişimi, farklılığı, eşitsizliği, bağımlılığı vb. ifadeleri içeren önermeler her zaman karşıt hipotezde belirtilir.

! Örnekten elde edilen değerleri kullanarak sıfır ya da karşıt hipotezden birinin doğru olduğuna karar verirken hata yapma olasılığımız her zaman vardır.

NEDEN?

→ Çünkü deneyi tekrarladığımızda farklı örnekten farklı ortalamaya sonuçları bulmamız olasıdır. Bu durumda gerçekte sıfır hipotezi doğru iken örnekleme hatası nedeniyle karşıt hipotezi doğru kabul edebiliriz ya da tam tersini yapabiliriz.

4. HAFTA

P Değerinin α ile Kıyaslanması

Eğer;

$P < \alpha$ ise H_0 reddedilir

$P > \alpha$ ise H_0 reddedilemez.

P Değeri Yaklaşımı

H_0 ile karşıt hipotez H_1 belirlenir α ve n seçilir. Uygun test istatistiği ve örneklem değeri belirlenir. Veri toplanır ve test istatistiği değeri ile P değeri hesaplanır istatistiksel karar verilir. Yöntemsel akışım belirlenir.

$$Z_{\text{test}} = \frac{\bar{x} - M}{\frac{\sigma}{\sqrt{n}}}$$

! H_0 hipotezi her zaman verilen değere eşit olmalıdır.

Soru: İstanbul'da Beşiktaş ilçesinde ikamet eden öğrencilerin haftalık ortalama kültürel amaçlı harcamalarının 60 TL'den farklı olduğu iddia edilmektedir. Rastgele seçilen 9 öğrencinin örneklem değerleri aşağıda verilmiştir.

35, 60, 70, 95, 30, 110, 80, 95, 130

$$\sigma^2 = 300$$

$$\alpha = 0,05$$

$$Z_t = \frac{\bar{x} - M}{\frac{\sigma}{\sqrt{n}}} \Rightarrow \sigma \text{ biliniyorsa } Z \text{ tablosu kullanılır.}$$

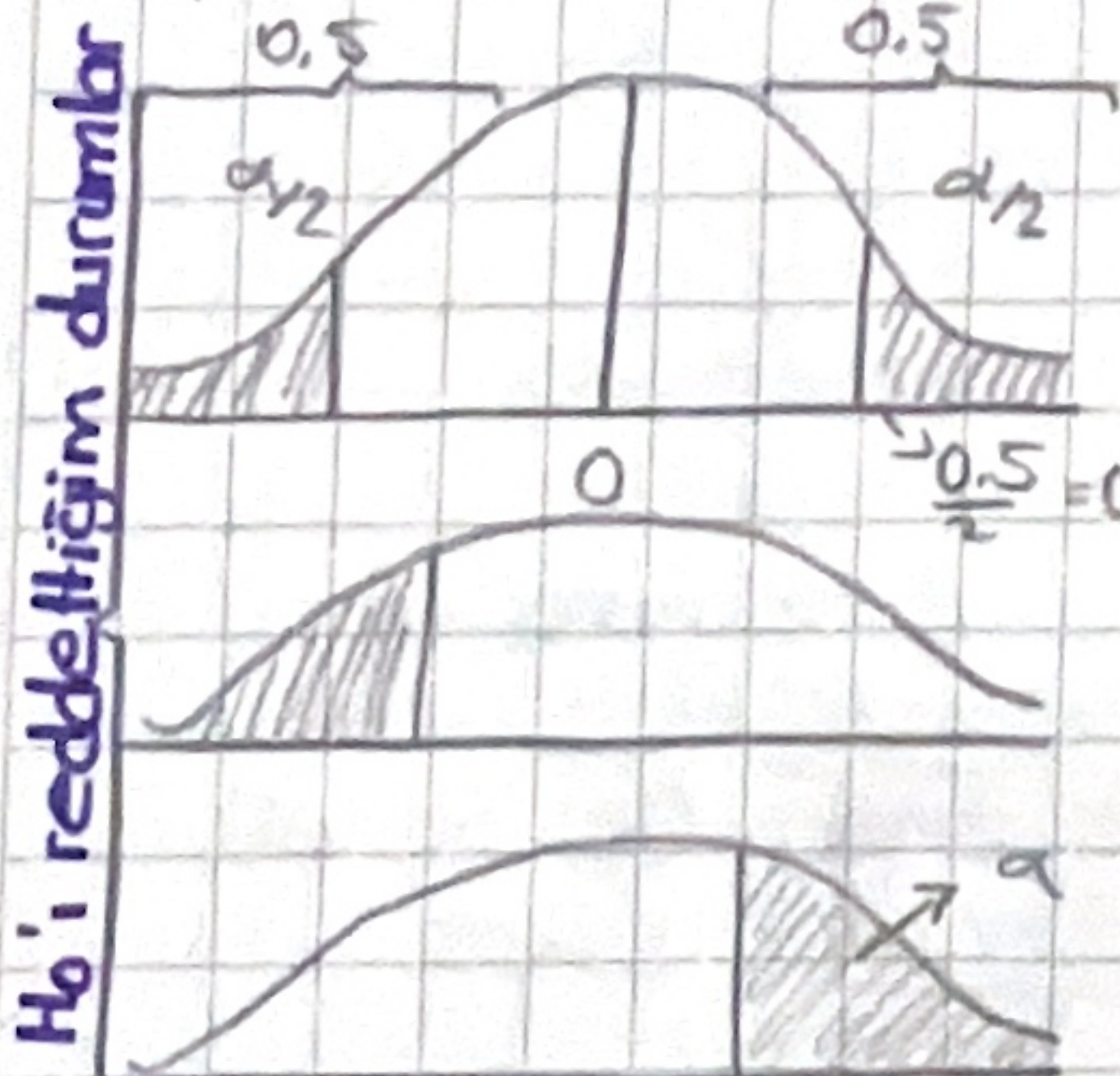
$$H_0: M = 60$$

$$\bar{x} = \frac{35 + 60 + 70 + 95 + 30 + 110 + 80 + 95 + 130}{9} = 78,3$$

$$H_1: M \neq 60$$

$$Z_t = \frac{78 - 60}{\frac{30}{\sqrt{9}}} = \frac{18,3}{10} = 1,83$$

$$M = 60$$



$$H_0: M = 60$$

$$H_1: M \neq 60$$

$$0,5 - 0,025 = 0,475$$

$$1,96$$

Z tablosu

$$H_0: M = 60$$

$$H_1: M < 60$$

$$H_0: M = 60$$

$$H_1: M > 60$$

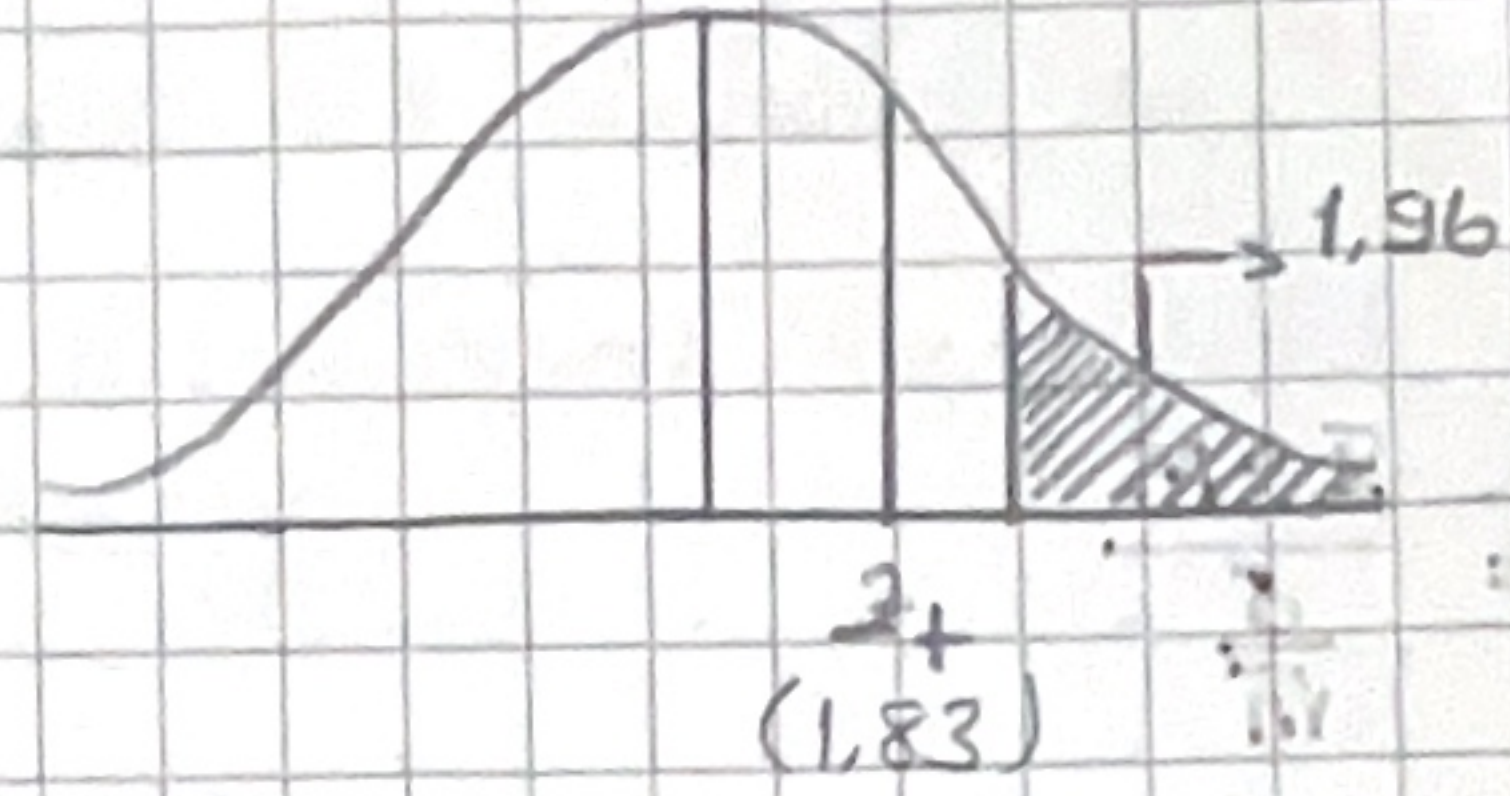
⇒ Kabul olan bir şeyin gerçekte reddedilmesi I. Tip Hata

⇒ Gerçekte hata olan bir şeyin kabul edilmesi II. Tip Hata

➢ Eğer test istatistiğim z_+ , tabanlı alana düşerse H_0 'i reddedeceğim, H_1 'i kabul edeceğim.

➢ Beyaz olma düşerse H_0 'i kabul edeceğim.

		Gerçek Durum	
		H_0 Doğru	H_0 Reft
Test Sonucu	H_0 Ret	I. Tip Hata (α)	Doğru Karar
	H_0 Kabul	Doğru Karar	II. Tip Hata (β)



z_+ Beyaz olma düştü

$H_0 = 60 \rightarrow H_0$ kabul ✓

soru: Bebek maması üreten bir işletme, ürettiği bebek mamalarının bir paketinin 350 gr, standart sapmasının $\sigma = 15$ gr olduğunu belirtmiştir. Bir tüketici bir bebek maması paketinin ortalama ağırlığının 350 gramdan az olduğunu iddia etmektedir. Bu iddiayı ispatlamak için 36 paketlik bir örneklem alınmış ve bu örneklem ortalaması 346 gram olarak bulunmuştur. Buna göre bu iddianın doğruluğunu %5 güven aralığında test ediniz.

çözüm:

$$\bar{x} = 346$$

$$H_0: \mu = 350$$

σ biliniyor $\rightarrow$ z tablosu

$$n = 36$$

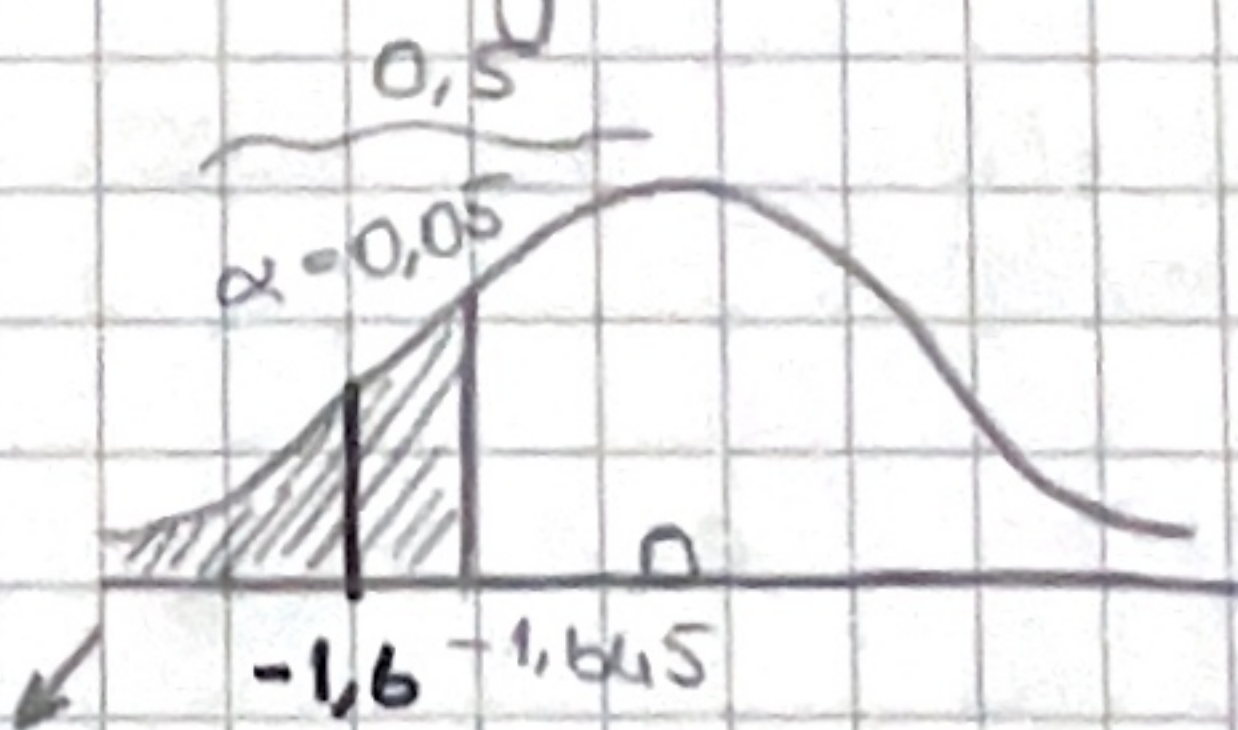
$$H_1: \mu < 350$$

$$\sigma = 15$$

$$z_+ = \frac{346 - 350}{\frac{15}{\sqrt{36}}} = \frac{-4}{15/6}$$

$$z_+ = -1,6$$

H_0 'i reddettik,
iddiayı kabul ettik.



$$0,5 - 0,05 = 0,450$$

z tablosuna git

$$\frac{1,64 - 1,65}{\text{arasında (ise ort. al)}}$$

$$\frac{1,64 + 1,65}{2} = 1,645$$