



T.C.
KARADENİZ TEKNİK ÜNİVERSİTESİ
MÜHENDİSLİK FAKÜLTESİ
ENDÜSTRİ MÜHENDİSLİĞİ BÖLÜMÜ



Mühendislik İstatistiği-II / İstatistiksel Tahmin

DR. ÖĞR. ÜYESİ FATMA BETÜL YENİ

İstatistiksel Tahmin

1. Nokta Tahmini

2. Aralık Tahmini

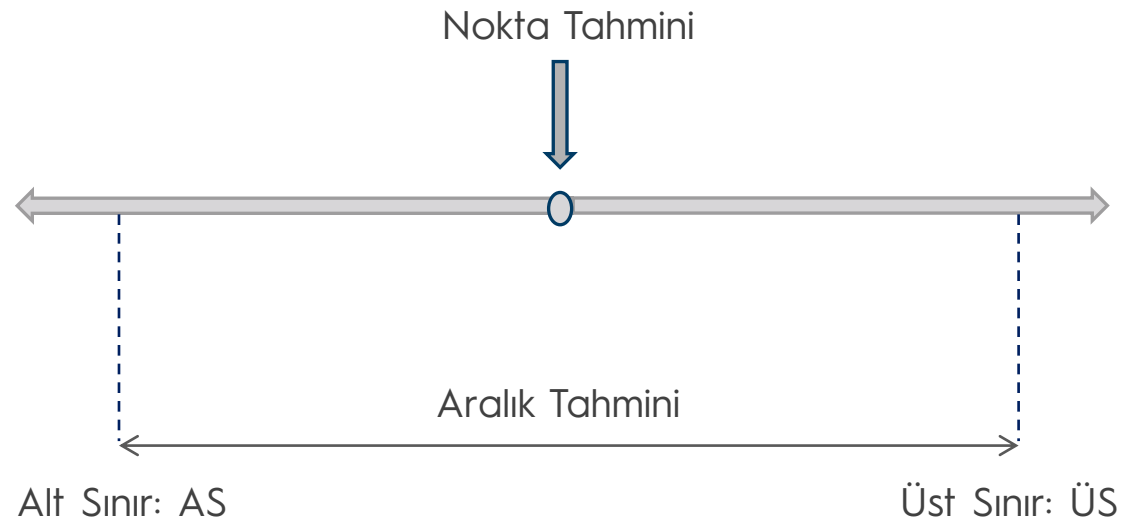
➤ Tek popülasyon için

- Yığın ortalaması μ 'nün aralık tahmini
- Yığın varyansı σ^2 'nin aralık tahmini
- Yığın (kitle) oranı π 'nin aralık tahmini

➤ İki popülasyon için

- Ortalamaların farkı $(\mu_1 - \mu_2)$ için aralık tahmini
- Varyansların oranı $(\frac{\sigma_1^2}{\sigma_2^2})$ için aralık tahmini

İstatistiksel Tahmin



Nokta Tahmini

Tahmin edici: Bilinmeyen bir θ kitle parametresini tahmin etmek için kullanılan T örneklem istatistiğine denir.

Nokta tahmini: Bir T tahmin edicisinin bir tek değerine denir. Bir rassal örneklemden hesaplanan T istatistiğinin değerini, ilgili yığın parametresi θ değerine eşit kabul eden tahminleme sürecidir. Nokta tahmini, tahmin işleminin güvenilirliği hakkında bilgi vermez.

Yığın Parametresi	Ortalama: μ	Oran: π	Varyans: σ^2
Nokta tahmincisi	$\bar{X}$	p	S_x^2

Nokta Tahmini

Örnek: $X_1=25$, $X_2=30$, $X_3=29$, $X_4=31$, olan X_i rassal değişkenleri, ortalaması μ olan normal dağılıma uysun. Bu durumda örneklem ortalaması, bilinmeyen popülasyon ortalaması μ 'nün nokta tahmin edicisidir.

- Bu durumda μ 'nün nokta tahmini şu şekildedir:

$$\bar{X} = \frac{25+30+29+31}{4} = 28,75$$

- Benzer şekilde, eğer popülasyon varyansı σ^2 bilinmiyorsa, tahmin edicisi örneklem varyansı S^2 'dir. Bu durumda örneklem verisinden hesaplanan $S^2 = 6,9$ sayısal değeri, popülasyon varyansının nokta tahminidir.

Nokta Tahmini

- Bir tahmin edici, bilinmeyen parametrenin gerçek değerine kısmen “yakın” olmalıdır. Eğer T ’nin beklenen değeri θ ’ya eşitse T ’nin θ ’nın yansız tahmin edicisi olduğunu söyleyebiliriz. Bu T ’nin olasılık dağılımının ortalamasının (ya da T ’nin örneklem dağılımının ortalamasının) θ ’ya eşit olduğunu söylemekle aynı anlama gelmektedir.
 - Eğer $E(T) = \theta$ ise T nokta tahmin edicisi θ parametresi için yansız tahmin edicidir.
 - Eğer tahmin edici yansız değilse (yanlı ise); $E(T) - \theta$ farkı T tahmin edicisinin hatası (bias) olarak bilinir.

Nokta Tahmini

Örnek: Örneklem ortalaması olan $\bar{X}$ istatistiği, ana kütle ortalaması olan μ 'nün yansız bir tahmin edicisidir. Çünkü $E(\bar{X}) = \mu$ 'dür.

Peki $\bar{X}^2$, μ^2 'nin yansız bir tahmin edicisi midir?

$$\frac{\sigma^2}{n} = E(\bar{X} - \mu)^2 \text{ olmak üzere;}$$

$$\frac{\sigma^2}{n} = E(\bar{X}^2) - \mu^2$$

$$\frac{\sigma^2}{n} + \mu^2 = E(\bar{X}^2) \text{ ise } E(\bar{X}^2) \neq \mu^2$$

Yani $\bar{X}^2$, μ^2 'nin **yanlı** bir tahmin edicisidir.

Nokta Tahmini

Örnek: Örneklem varyansı S^2 'nin popülasyon varyansı σ^2 'nin yansız bir tahmini olduğunu gösterelim.

$$\begin{aligned} E(S^2) &= E\left(\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}\right) = \frac{1}{n-1} E\left(\sum_{i=1}^n (X_i - \bar{X})^2\right) \\ &= \frac{1}{n-1} E\sum_{i=1}^n (X_i^2 + \bar{X}^2 - 2X_i\bar{X}) = \frac{1}{n-1} E\left(\sum_{i=1}^n X_i^2 - n\bar{X}^2\right) \\ &= \frac{1}{n-1} \left[\sum_{i=1}^n E(X_i^2) - nE(\bar{X}^2)\right] \end{aligned}$$

$E(X_i^2) = \mu^2 + \sigma^2$ ve $E(\bar{X}^2) = \mu^2 + \sigma^2/n$ olduğundan

$$E(S^2) = \frac{1}{n-1} (n\mu^2 + n\sigma^2 - n\mu^2 - \sigma^2) = \sigma^2$$

Nokta Tahmini

- Yanlı ve yansız tahmin ediciler söz konusu olduğunda yansız tahmin edici tercih edilir.
- İki ya da daha fazla yansız tahmin edici söz konusu olduğunda, parametreye yakın olma olasılığı yüksek olan bir tahmin edici tercih edilir. Yani en küçük varyanslı tercih hangisi ise o seçilir. Seçilen bu tahmin ediciye, minimum varyans yansız tahmin edicisi denir.

Örnek: $X_1, X_2, \dots, X_n$ örnekleminde $E(X_i) = \mu$ ve $E(\bar{X}) = \mu$ olduğundan X_i ve $\bar{X}$ 'nin her ikisi de μ 'nün yansız bir tahmin edicisidir. Bununla beraber;

$$\text{Var}(X_i) = \sigma^2 \quad \text{ve} \quad \text{Var}(\bar{X}) = \sigma^2/n \text{ 'dir.}$$

Bu durumda $\bar{X}$ daha küçük varyanslı olduğundan X_i 'e göre tercih edilir.

Aralık Tahmini

- Aralık tahmininde, belirlenen bir hata payı (anlam/önem düzeyi) α veya güven düzeyi $(1-\alpha)$ ile ana kütle parametresi θ için alt ve üst güven sınırları oluşturulur. Oluşturulacak alt sınır a ve üst sınır b ile ifade edilirse,

$$P(a < \theta < b) = 1 - \alpha \text{ olacaktır.}$$

- **Güven Düzeyi:** Aralık tahmininin popülasyon parametresini içermesi konusundaki kesinlik düzeyidir. $1 - \alpha$ ile gösterilir. En çok kullanılan güven düzeyleri: %90, %95, %99 'dur.
- **Önem Düzeyi:** Gerçek değerin tahmin aralığı dışında olma durumunu temsil eder. α ile gösterilir, anlamlılık düzeyi olarak da bilinir.
- **Güven Aralığı:** Tahminin güven düzeyini kullanarak bir parametre için belirlenen aralığa denir.

Güven Aralığı

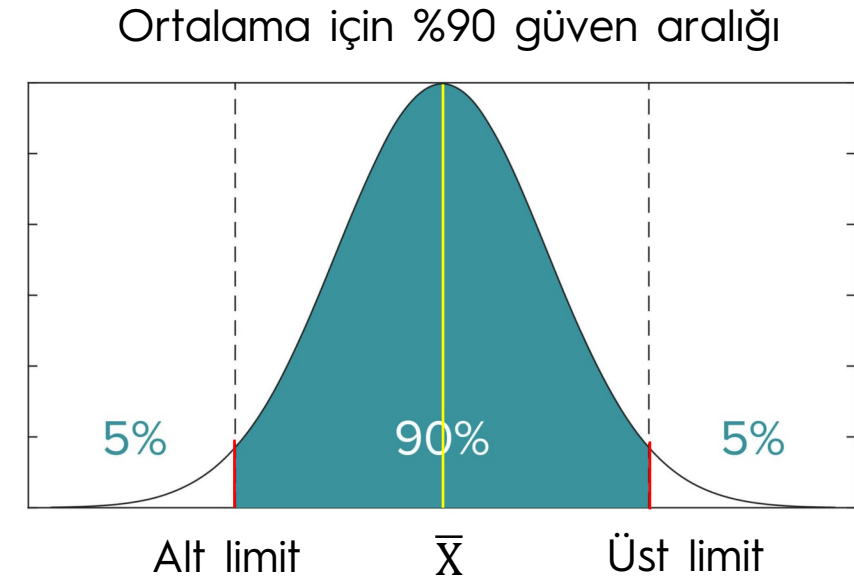
Örnek: Bir anakütlenin ortalaması tahmin edilirken nokta tahmini 5 olarak elde edilmiş olsun. 0.05 hata payı veya 0.95 güven olasılığı ile alt ve üst güven sınırları 4 ve 6 olarak hesaplanmış olsun.

Bu sonuç anakütle ortalaması değerinin 0.95 olasılıkla 4-6 arasında olduğunu; anakütle ortalaması değerinin 4'ten küçük veya 6'dan büyük olması olasılığının ise 0.05 olduğunu ifade etmektedir.

Güven Aralığı

Bir güven aralığı (interval):

- Örnekten örneğe, örnek istatistiklerindeki değişimi dikkate alan
- Bir örnekten elde edilen gözlemlere dayanan
- Bilinmeyen popülasyon parametrelerine yakınlık hakkında bilgi veren
- Güven seviyesi ifadesiyle belirtilen bir değer aralığı (range) verir.



Yığın Ortalaması μ 'nün Aralık Tahmini

- Varyans bilindiğinde yığın ortalaması μ 'nün aralık tahmini
- Varyans bilinmediğinde yığın ortalaması μ 'nün aralık tahmini
 - Örneklem hacmi $n \geq 30$ ise
 - Örneklem hacmi $n < 30$ ise

Tanım: Hata payı (e), güven aralığının kapsadığı nokta tahminden muhtemel en büyük uzaklık olmak üzere, μ için genel aralık tahmini:

$$\bar{X} - e < \mu < \bar{X} + e$$

σ^2 Bilindiğinde μ 'nın Aralık Tahmini

- Bir kitlenin μ ortalamasını tahmin etmek istediğimizi varsayalım. $\bar{X}$ arzu edilen özelliklere sahip olduğundan, onun üzerinden aralığı daha iyi kurabiliriz. $\bar{X}$, bilinen σ^2 varyanslı bir normal yoğunluktan alınan n rasgele değişkenden bulunmuş ise; $\bar{X}$ normal dağılıma ve $Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$ birim normal dağılıma sahiptir.

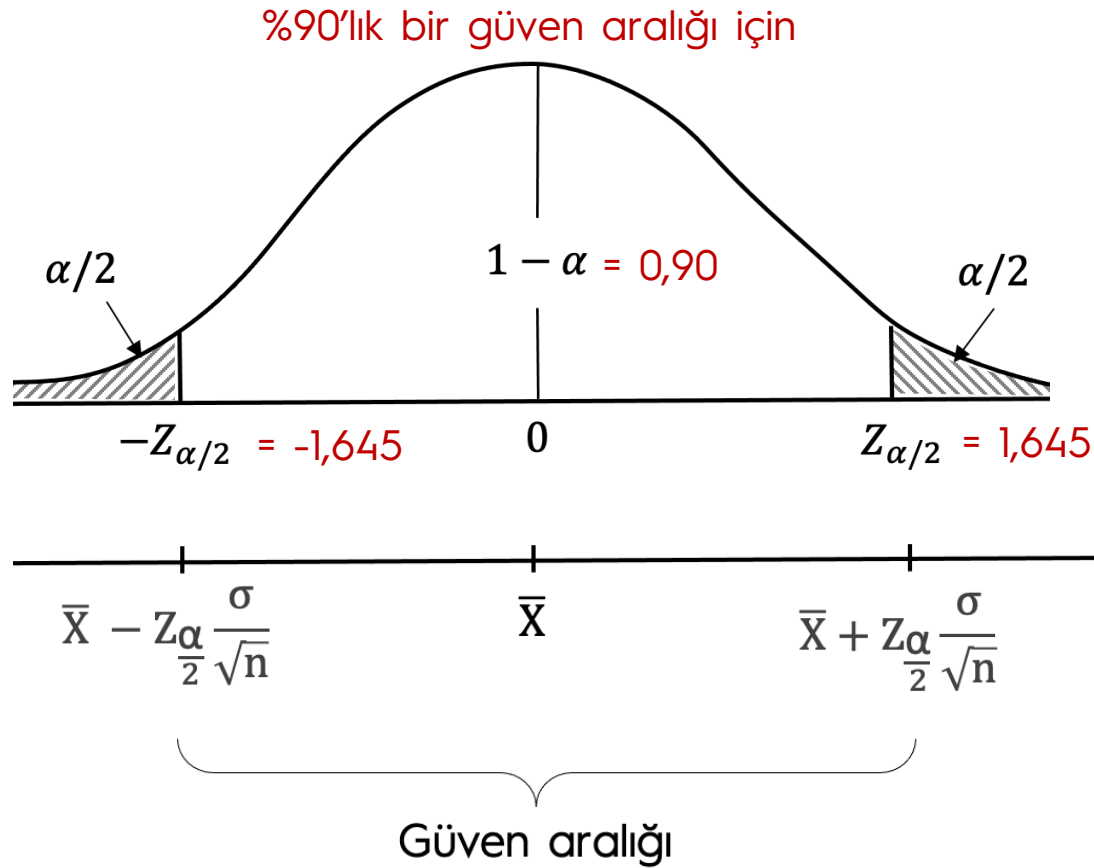
σ^2 Bilindiğinde μ 'nün Aralık Tahmini

Tanım: Anakütle ortalaması μ 'nün α önem düzeyi veya $1-\alpha$ güven düzeyi ile içinde bulunacağı aralığın alt ve üst sınırları hesaplanırken her iki kuyruk için de $Z_{\alpha/2}$ kritik değerine bakılacaktır. Bu durumda; $P(-Z_{\alpha/2} < Z < Z_{\alpha/2}) = 1-\alpha$ olmak üzere μ için aralık tahminimiz;

$$-Z_{\frac{\alpha}{2}} < \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} < Z_{\frac{\alpha}{2}} \implies \bar{X} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

Burada $\pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ bizim 'e' hata payımızdır.

σ^2 Bilindiğinde μ 'nün Aralık Tahmini



➤ $Z_{\alpha/2}$ değerinin hesaplanması

%90 güven düzeyini ele alalım.
Bu durumda $1 - \alpha = 0.90$ ve $\alpha = 0,1$ olur.

$Z_{0,1/2} = Z_{0,05}$ değerini bulmak istiyoruz. Bunun için Z tablosundan $1 - \alpha/2$ değerine bakarız. Bu örnek için tablodan $1 - 0,05 = 0,95$ değerini bulursak; **1,64** ile **1,65** arasında kaldığını görürüz.

Buna göre $Z_{0,05}$ değeri 1,645 olarak alınır.

σ^2 Bilindiğinde μ 'nün Aralık Tahmini

Örnek: Geniş normal dağılımlı bir popülasyondan alınan 11 elemanlı bir örneklem için ortalama direnç 2.20 ohm'dur. Geçmiş test verilerinden popülasyonun standart sapmasının 0.35 ohm olduğunu biliyoruz. Popülasyonun doğru ortalama direnci için %95 güven aralığı oluşturalım.

$n = 11$; $\bar{X} = 2,20$; $\sigma = 0,35$ olmak üzere

$$\bar{X} - Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$Z_{\alpha/2} = Z_{0,025}$ olduğuna göre tablodan 0,0975 değerini buluruz: 1,96

$$2,20 - 1,96 \frac{0,35}{\sqrt{11}} < \mu < 2,20 + 1,96 \frac{0,35}{\sqrt{11}} \implies 1,1932 < \mu < 2,4068$$

Buna göre; %95 güven aralığında ortalama direnç için sınırlarımız (1,1932 ; 2,4068)'dir.

σ^2 Bilindiğinde μ 'nün Aralık Tahmini

Örnek: Bir tezgahta üretilen parçaların dış çapları normal dağılıma uymaktadır ve standart sapması 2.4 cm'dir. Tezgahın üretiminden rastgele seçilen 16 parçanın dış çap ortalaması 3.2 cm olarak bulunmuştur. %1 hata (%99 güven) seviyesinde anakütle ortalamasının güven aralığını tahmin edelim.

$n = 16$; $\bar{X} = 3,2$; $\sigma = 2,4$; $Z_{\alpha/2} = Z_{0,005} = 2,575$ olmak üzere

$$3,2 - 2,575 \frac{2,4}{\sqrt{16}} < \mu < 3,2 + 2,575 \frac{2,4}{\sqrt{16}} \quad \longrightarrow \quad 1,655 < \mu < 4,745$$

Alınan örneklerle göre sözü edilen tezgahta üretilen parçaların dış çapları ortalamasının %1 hata (%99 doğruluk) payı ile 1,655 cm ile 4.745 cm arasında olacağı söylenebilir.

σ^2 Bilindiğinde μ 'nün Aralık Tahmini

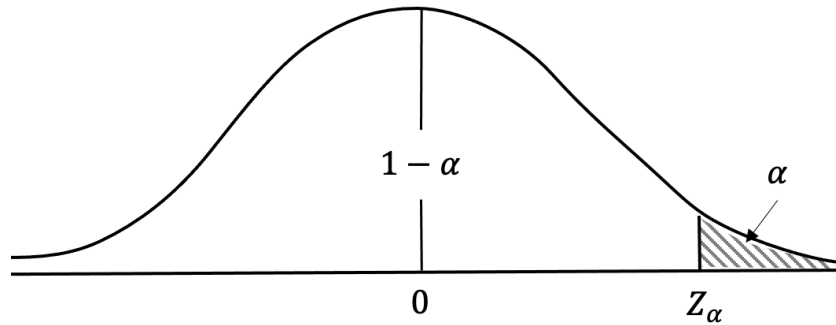
- Kitle sonlu elemanlı olduğunda ve iadesiz seçim söz konusu ise, μ için güven aralığı şu şekilde hesaplanır:

$$\bar{X} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} < \mu < \bar{X} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

- $\frac{n}{N} < 0,05$ olduğunda $\sqrt{\frac{N-n}{N-1}}$ terimi güven sınırlarını çok az etkilemektedir.

Tek Uçlu Güven Sınırı

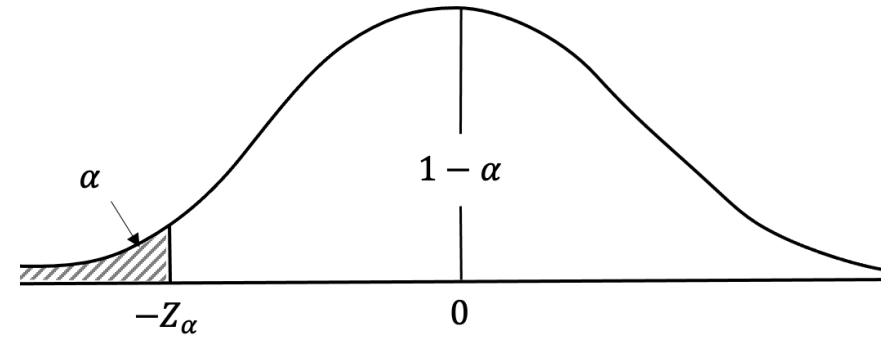
Üst güven sınırı



$$\mu < \bar{X} + Z_\alpha \frac{\sigma}{\sqrt{n}}$$

Ortalama için üst sınır

Alt güven sınırı



$$\bar{X} - Z_\alpha \frac{\sigma}{\sqrt{n}} < \mu$$

Ortalama için alt sınır

Örneklem büyüklüğü ne kadar olmalı?

➤ Kitle parametresinin iyi bir tahmini için yeterli büyüklükte örneklem seçilmelidir. Bu karar üç faktöre bağlıdır:

- Güven aralığı ne kadar dar olmalıdır?
- Güven aralığı ne kadar güvenle kitle parametresini içine almalıdır?
- Örneklemin seçileceği kitledeki değişkenlerin sayısı ne büyüklüktedir?

Tanım: $\bar{X}$ ile μ arasındaki fark olarak tanımladığımız $e = Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ olmak üzere; seçilmesi gereken örneklem büyüklüğü şu şekilde hesaplanır:

$$n = \left[\frac{Z_{\frac{\alpha}{2}} \cdot \sigma}{e} \right]^2$$

Örneklem Büyüklüğü Hesabı

Örnek: Türkiye’de her bir şahsın yiyecek için yaptığı aylık harcamayı tahmin etmek amacıyla $n=49$ kişinin rasgele seçildiğini kabul edelim. Örneklem ortalaması 500 birim lira olsun. Geçmiş veriler göz önünde bulundurulduğunda kitle standart sapmasının 90 birim lira olduğu bilinmektedir. Buna göre aylık ortalama yiyecek harcaması için %95’lik güven aralığını bulalım. Güven aralığı genişliğinin 20 birim lira olabilmesi için gerekli örneklem büyüklüğünü hesaplayalım.

$$n = 49; \bar{X} = 500 ; \sigma = 90 ; Z_{\alpha/2} = Z_{0,025} = 1,96 \text{ olmak üzere}$$

$$500 - 1,96 \frac{90}{\sqrt{49}} < \mu < 500 + 1,96 \frac{90}{\sqrt{49}} \quad \Rightarrow \quad 479,8 < \mu < 525,2$$

Eğer güven sınırları arasındaki farkın 20 birim olmasını istiyorsak; $e = 20/2 = 10$ olur. Buna göre;

$$n = \left[\frac{1,96 * 90}{10} \right]^2 \cong 311 \text{ olmalıdır.}$$

Chebyshev Eşitsizliği ile n Hesabı

Tanım: Kitlenin dağılımı bilinmiyorsa ve örnek hacminin büyük olacağı kesin değilse örneklem büyüklüğünün hesaplanmasında Chebyshev (Çebişev) eşitsizliği kullanılır.

$$\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n} \quad \text{olmak üzere bu eşitsizlik;}$$

$$P \left[\frac{|\bar{X} - \mu|}{\sigma_{\bar{X}}} < k \right] \geq 1 - \frac{1}{k^2} \quad \text{olarak yazılır.}$$

Chebyshev Eşitsizliği ile n Hesabı

Örnek: Bir kitlenin standart sapması 2000 olsun. Bilinmeyen kitle ortalaması μ ile örneklem ortalaması arasındaki farkın 100'den az olduğunu %95 güvenle söyleyebilmemiz için örneklem büyüklüğü n ne olmalıdır?

➤ Bahsi geçen olasılık: $P(|\bar{X} - \mu| < 100)$

$$P\left[\frac{|\bar{X} - \mu|}{\sigma_{\bar{X}}} < \frac{100}{\sigma_{\bar{X}}}\right] \text{ Yazacak olursak;}$$

$$P\left[\frac{|\bar{X} - \mu|}{\sigma_{\bar{X}}} < k\right] \geq 1 - \frac{1}{k^2} \quad \text{için } k = \frac{100}{\sigma_{\bar{X}}} \text{ olur.}$$

Chebyshev Eşitsizliği ile n Hesabı

Buna göre; $P(|\bar{X} - \mu| < 100) \geq 1 - \frac{1}{\frac{100^2}{\sigma_{\bar{X}}^2}} = 0,95$ olmak üzere;

$$0,95 = 1 - \frac{\sigma_{\bar{X}}^2}{100^2} = 1 - \frac{\sigma^2}{n * 100^2} = 1 - \frac{2000^2}{n * 100^2}$$

➡ $n = 8000$ bulunur

!! Eğer n'nin bu kadar büyük olacağını bilseydik, aynı hesabı normal dağılım üzerinden yapardık.

O halde $n = \left[\frac{1,96 * 2000}{100} \right]^2 = 1600$ olarak bulunurdu.

σ^2 Bilinmediğinde μ 'nın Aralık Tahmini

- Anakütle varyansının bilinmediği, fakat örnek hacminin 30 veya daha büyük olduğu ($n \geq 30$) durumlarda örneklemin varyansı (S^2) kullanılarak Z dağılımı yardımıyla güven aralığı oluşturulur.

$$\bar{X} - Z_{\frac{\alpha}{2}} \frac{S_x}{\sqrt{n}} < \mu < \bar{X} + Z_{\frac{\alpha}{2}} \frac{S_x}{\sqrt{n}}$$

- Anakütle varyansının bilinmediği durumlarda örnek hacmi 30 dan küçük ($n < 30$) ise küçük örnek teorisine göre geliştirilen t dağılımı yardımıyla güven aralığı oluşturulur. Bu durumda

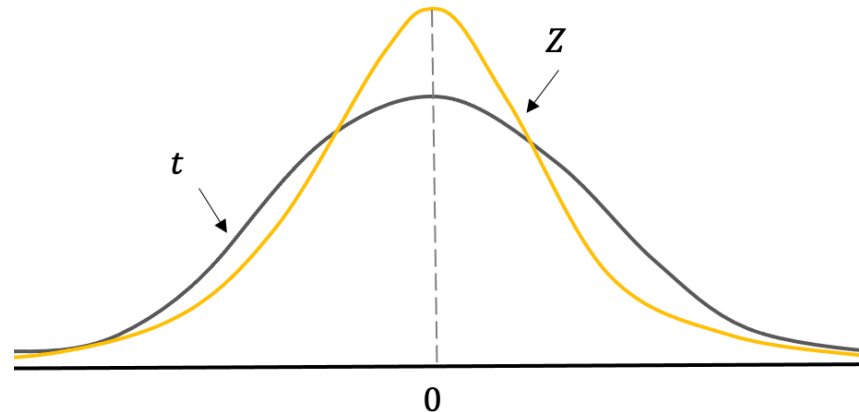
$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \text{ istatistiği yerine } t_{n-1} = \frac{\bar{X} - \mu}{\frac{S_x}{\sqrt{n}}} \text{ istatistiği kullanılır.}$$

σ^2 Bilinmediğinde μ 'nın Aralık Tahmini

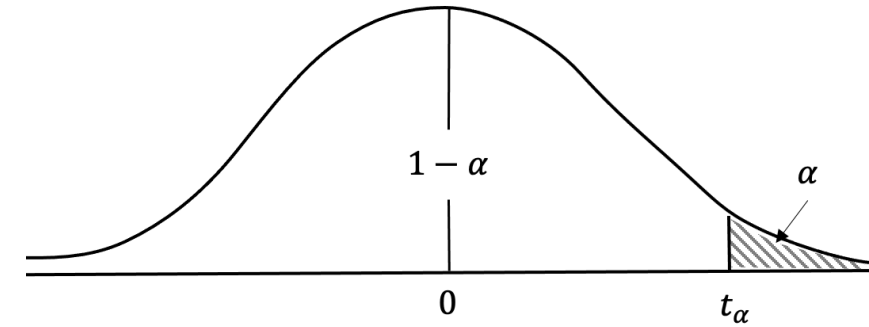
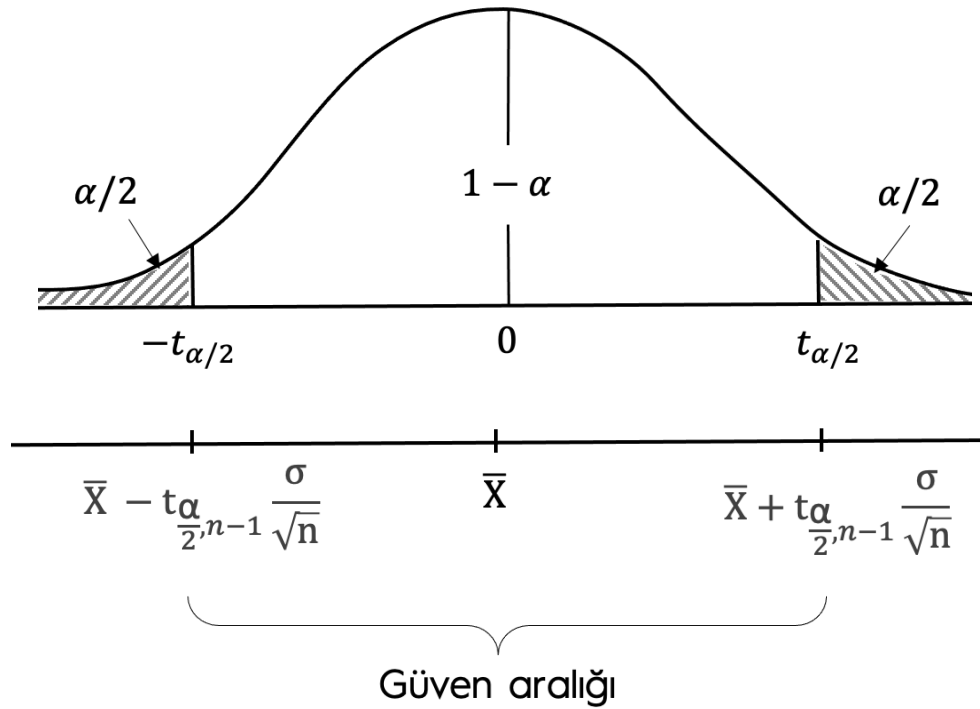
Tanım: Eğer $\bar{X}$ ve S , varyansı bilinmeyen normal dağılımdan alınmış rassal örneklemin ortalaması ve standart sapması ise μ için $(1 - \alpha)$ düzeyinde oluşturulan güven aralığı;

$$\bar{X} - t_{\frac{\alpha}{2}, n-1} \frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{\frac{\alpha}{2}, n-1} \frac{S}{\sqrt{n}}$$

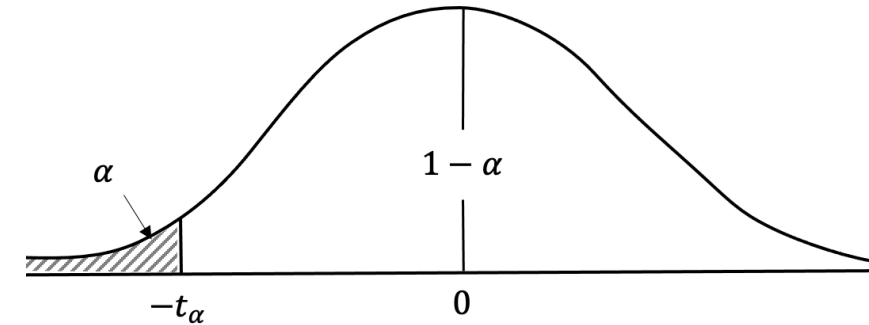
Şeklinde hesaplanır. $t_{\frac{\alpha}{2}, n-1}$ 'deki $(n-1)$ ifadesi t dağılımının serbestlik derecesini ifade eder.



σ^2 Bilinmediğinde μ 'nın Aralık Tahmini



Tek yönlü güven sınırları



σ^2 Bilinmediğinde μ 'nın Aralık Tahmini

Örnek: Normal dağılıma uyduğu bilinen bir veri setinden ortalaması $\bar{X} = 50$ ve standart sapması $S = 8$ olan $n = 25$ büyüklüğünde bir örnek alalım. μ için %95 güven aralığını oluşturalım

➤ Serbestlik derecesi $n - 1 = 24$

➤ $\alpha = 0,05$ ve $t_{\frac{\alpha}{2}} = t_{0,025 ; 24} = 2,064$ olmak üzere μ için güven aralığı;

$$50 - 2,064 \frac{8}{\sqrt{25}} < \mu < 50 + 2,064 \frac{8}{\sqrt{25}}$$

$$46,698 < \mu < 53,302$$

σ^2 Bilinmediğinde μ 'nın Aralık Tahmini

Örnek: Otomobil lastiği üreten bir firmanın yöneticisi, üretilen lastiklerin ortalama ömürlerinin km olarak tahmin etmek istemektedir. Bu amaçla rassal olarak 25 lastik seçilmiş ve bu lastiklerin ortalama ömürlerinin 30.000 km ve standart sapmasının da 1.500 km olduğu tespit edilmiştir. %99 güven düzeyinde otomobil lastiklerinin en az ve en fazla dayanma ömürlerini tahmin edelim.

➤ Serbestlik derecesi $n - 1 = 24$

➤ $\alpha = 0,01$ ve $t_{\frac{\alpha}{2}} = t_{0,005} = 2,797$ olmak üzere μ için güven aralığı şu şekildedir;

$$30000 - 2,797 \frac{1500}{\sqrt{25}} < \mu < 30000 + 2,797 \frac{1500}{\sqrt{25}}$$

$$29.160,9 < \mu < 30.839,1$$

σ^2 Bilinmediğinde μ 'nın Aralık Tahmini

Örnek: 10 hastaya A ve B olarak nitelendirilen iki çeşit uyku ilaç verilmiştir. Her bir hasta için B ilacının etkisinin A ilacına eşit ya da daha fazla olduğu saptanmıştır. Fazla uyku saatlerinin ortalaması için %95'lik güven aralığını bulalım.

Hasta	B ilacı ile kazanılan fazla uyku saati (X)
1	1,2
2	2,4
3	1,3
4	1,3
5	0
6	1
7	1,8
8	0,8
9	4,6
10	1,4

σ^2 Bilinmediğinde μ 'nün Aralık Tahmini

Çözüm:

- $\bar{X} = 1,58$ 'dir
- $S_x^2 = \frac{1}{10} \sum_{i=1}^{10} (x_i - 1,58)^2 = 1,513$
- $S_x = 1,23$
- $n = 10$, serbestlik derecesi $n - 1 = 9$
- $t_{\frac{0,05}{2}}$ için tablo değeri 2,262 olmak üzere

$$1,58 - 2,262 \frac{1,23}{\sqrt{10}} < \mu < 1,58 + 2,262 \frac{1,23}{\sqrt{10}}$$

$$0,70 < \mu < 2,46$$

Sorularınız?