Olasılık ve İstatistik
BİL 3013

Fırat İsmailoğlu

Hafta 11

Tahmin

Çıkarımsal istatistikte örneğin (sample'ın) alındığı populasyonla ilgili çıkarım yaptığımızı söylemiştik. Bu çıkarımlar populasyon parametreleri (ortalama, varyans, populasyon oranı) ile ilgili tahminler (estimation) dir.

Populasyon parametreleri ile ilgili tahminlerimizi iki başlık altında inceleyeceğiz: nokta tahminler ve aralık tahminler.

Nokta Tahminler (Point Estimators)

1) Populasyon Ortalamasının Nokta Tahmini

$X_1, X_2, \dots, X_n$ değerleri populasyondan alınan n değer olsun. Bu değerlerin oluşturduğu örnek (sample)'ın ortalaması olan $\bar{X}$, populasyon ortalaması olan μ'nün bir tahmini olarak kullanılabilir. Zaten merkezi limit teoreminden hatırlarsak populasyondan rastgele eleman seçerek oluşturulan örneklerin ortalamalarının beklenen değeri μ idi. Bu yüzden herhangi bir örneğin ortalaması μ'için bir tahmin olabilir.

ör. Bir fabrikada üretilen 8 otomobilin ilk kez arızalanmaları (bin km cinsiden)

$175, 180, 185, 198, 148, 162, 191$ ve 205 km sonra gerçekleşmiş olsun. Buna göre bu fabrikada üretilen otomobiller ortalama kaç bin km'den sonra ilk kez arızalanırlar?

Çözüm.

Seçilen 8 otomobilin ortalaması olan $\bar{X} = \frac{175+180+185+198+148+162+191+205}{8} = 180.5$ fabrikanın ürettiği tüm otomobillerden oluşan populasyonun ortalamasının bir tahmini olabilir.

Bu tahmin sonuçta bir tahmin olduğundan bir hata payı vardır. Bu hatanın miktarı örneğin standart sapması olan $\frac{\sigma}{\sqrt{n}}$ kadardır. (σ populasyon standart sapması, n örnek büyüklüğü).

Örnek üzerinden populasyon parametresi nokta tahmini olarak yapılırken, tahminin hatasına standart hata (standard error) denir.

Standart hata populasyon ortalaması için $\frac{\sigma}{\sqrt{n}}$ kadardır.

Bir önceki soru için varsayalımki fabrikanın ürettiği tüm otomobillerin ilk kez arıza yapmalarının standart sapması 45 bin olsun. Bu durumda nokta tahminmiz olan 180.5 değerinin standart hatası = $\frac{45}{\sqrt{8}} = 15.95$ olur.

ör. Örnek büyüklüğü 1000 iken örneğin standart hatasını 4’te 1’e düşürmek için örneğe kaç eleman daha eklemek gerekir?

Çözüm.

Örnek büyüklüğü 1000 iken standart hata: $\frac{\sigma}{\sqrt{1000}}$ dir. Bunun 4’te 1’i $\frac{\sigma}{4\sqrt{1000}} = \frac{\sigma}{\sqrt{16000}}$ dir. O halde örnek büyüklüğü 16000 olamalıdır, yani 15000 eleman daha eklenmelidir.

2) Populasyon Oranının Nokta Tahmini

Varsayalımki bir populasyon elemanlarının belirli bir oranını merak ediyoruz. Örneğin Türkiye’deki seçmenlerin bir X partisine oy verenlerin oranı, bir fabrikanın ürettiği bütün mallardan bozuk olanlarının oranı, 18-25 yaş arası gençlerde sosyal medya bağımlılığı oranı..

Populasyon oranını tahmin etmek için populasyondan rastgele elemanlar seçerek bir örnek oluşturabilir; daha sonra bu örneğin oranını, populasyon oranının bir nokta tahmini olarak kullanabiliriz.

Populasyon oranının nokta tahmini: $\hat{p} = \frac{X}{n}$

Burada X örnekte, ilgilendiğimiz olaya katılan eleman sayısıdır: Örneğin önekteki bozuk ürün sayısı, belirli bir partiye oy verenlerin sayısı, sosyal medya bağımlılarının sayısı…

Bu tahminin standart hatası $\sqrt{\frac{p(1-p)}{n}}$ olur.

Burada p, populasyonun oranıdır. Fakat p'yı bilmedeğimizden p'nin yerine örnekten elde ettiğimiz tahmin olan : $\hat{p}$'yi kullanırız.

ör. Bir bölümde son sınıflar için mezuniyet yapılıp yapılmaması düşünülüyor.Bunun için son sınıf öğrencilerinden rastgele 50 kişi seçilip, fikirleri soruluyor. 30 öğrenci mezuniyetin olmasını istiyor.

Bu durumda tüm son sınıf öğrencileri arasında mezuniyet olmasını isteyen öğrencilerin oranı için bir tahmin yapınız. Bu yaptığınız tahminin standart hatası ne olur?

Çözüm.

Populasyon oranı için nokta tahmini $\hat{p} = \frac{30}{50} = 0.6$

Bu tahminin standart hatası: $\sqrt{\frac{0.6 \cdot 0.4}{50}} = 0.0693$

ör. Önümüzdeki seçimlerde seçmenlerin kime oy vereceğini tahmmin etmek için bir anket yapılıyor. 300 kişilik bu ankete 210 kişi X partisine oy vereceğini söylüyor. Buna göre tüm seçmenler içinde X partisine oy vereceklerin oranını tahmin ediniz. Bu tahminizin standart hatasını hesaplayınız.

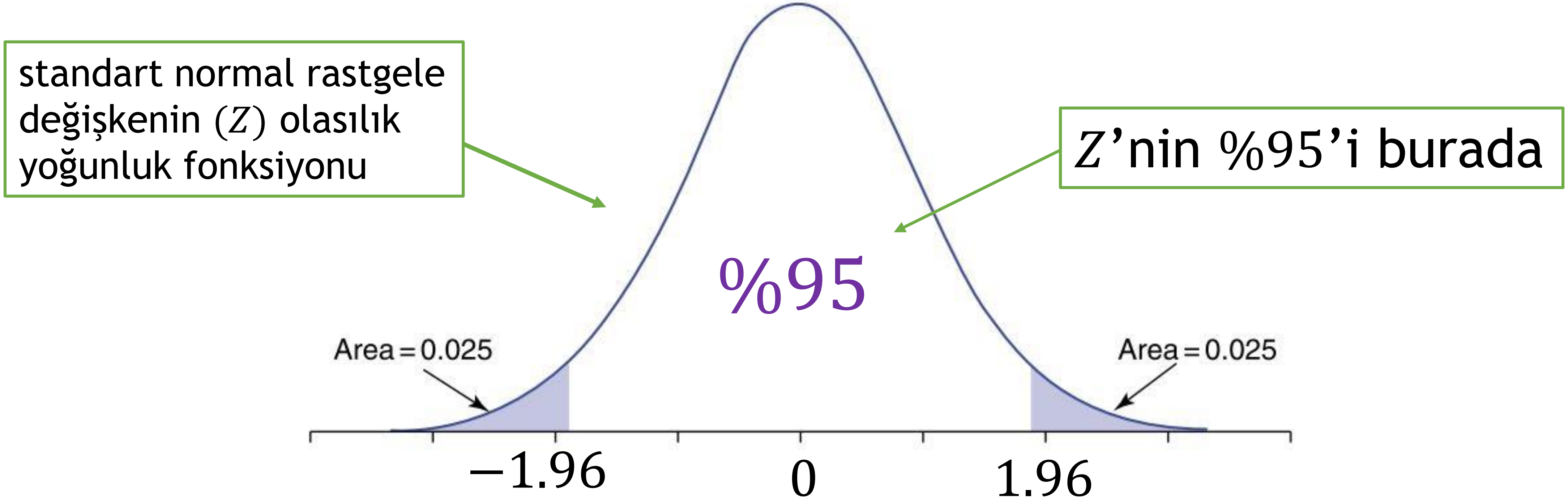
Çözüm.

X partisine oy vereceklerin oranının tahmini $\hat{p} = \frac{210}{300} = 0.7$

Bu tahminin standart hatası: $\sqrt{\frac{0.7 \cdot 0.3}{300}} = 0.02$

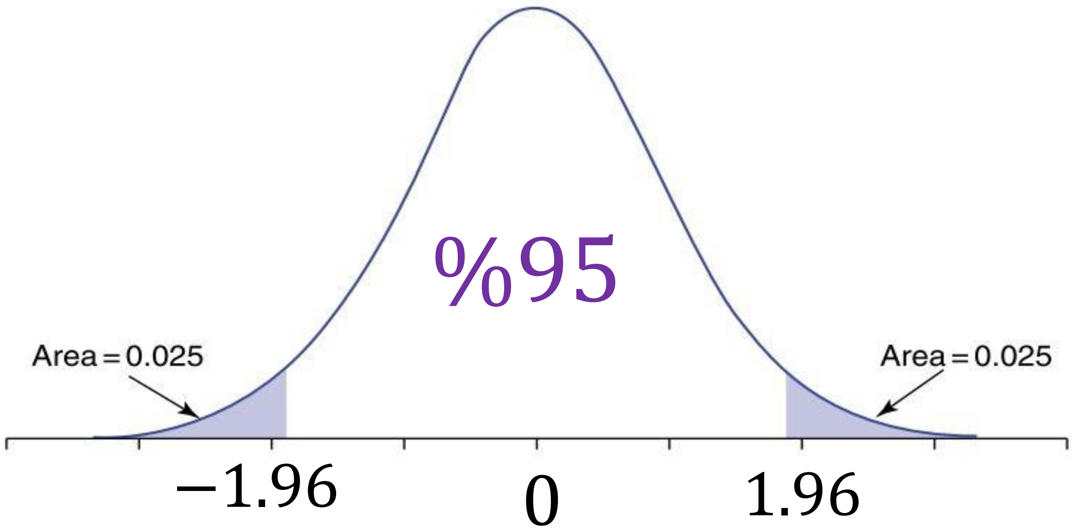
Aralık Tahminler (Interval Estimators)

Bir populasyon parametresinin aralık tahmini, bu parametreyi içerdiği düşünülen aralıktır. Bu aralığın güven değeri, aralığın paremetreyi içerme olasılığıdır.



Standart normal rastgele değişken olan Z’nin 1.96’dan büyük olma olasılığı: $P(Z \geq 1.96) = 0.025$

Z’nin 1.96’dan küçük olma olasılığı: $P(Z \leq -1.96) = 0.025$



$$\boxed{P(-1.96 \le Z \le 1.96) = 0.95}$$

$\bar{X}$ rastgele değişkeni populasyondan alınan bir n büyüklükteki örneğin ortalaması olsun. Merkezi limit teoremi gereği $\bar{X}$ normal dağılıma sahiptir, ortalaması (beklenen değeri) μ'dür ve standart sapması $\frac{\sigma}{\sqrt{n}}$'dir. (σ, populasyon standart sapması).

$\bar{X}$ normal değişkenini standart hale getirmek için (ortalamasını 0 yapmak ve standart sapmasını 1 yapmak icin):

$$Z \leftarrow \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \sqrt{n}\frac{\bar{X} - \mu}{\sigma}$$

$-1.96 \leq Z \leq 1.96$ aralığını $\bar{X}$ ve μ'yü içerecek şekilde düzenlersek:

$$-1.96 \leq \sqrt{n}\frac{\bar{X}-\mu}{\sigma} \leq 1.96$$

$$-1.96 \times \frac{\sigma}{\sqrt{n}} \leq \bar{X}-\mu \leq 1.96 \times \frac{\sigma}{\sqrt{n}}$$

$$\bar{X} - 1.96 \times \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + 1.96 \times \frac{\sigma}{\sqrt{n}}$$

olur. O halde populasyon ortalaması olan μ'nün $\left[\bar{X} - 1.96 \times \frac{\sigma}{\sqrt{n}}, \bar{X} + 1.96 \times \frac{\sigma}{\sqrt{n}}\right]$ aralığında olma olasılığı 0.95'tir. Bir başka deyişle biz %95 eminizki μ bu ralıktadır. Bu aralığa μ'nün %95 **güven aralığı** (confidence interval) denir.

$$\boxed{P\left(\bar{X} - 1.96 \times \frac{\sigma}{\sqrt{n}} \leq \ \mu \ \leq \bar{X} + 1.96 \times \frac{\sigma}{\sqrt{n}}\right) = 0.95}$$

Burada σ populasyonun standart sapmasıdır ve örnekten elde edilemez. Güven aralığını hesaplarken biz şimdilik σ'nın bilindiğini varsayalım.

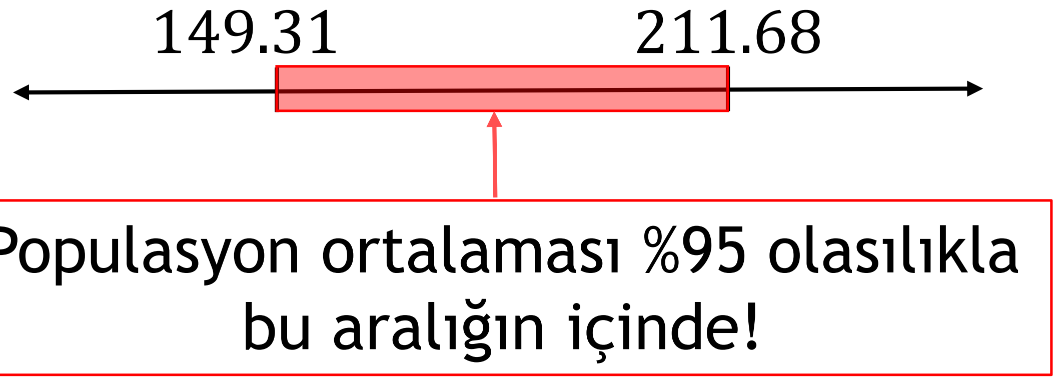
ör. Daha önceki soruda bir fabrikada üretilen 8 otomobilin ilk kez arızalanmaların kaç bin km sonra gerçekleştiğini vermistik: 175, 180, 185, 198, 148, 162, 191 ve 205. Diyelimki bu fabrikada üretilen tüm otomobillerin (populasyonun) kaç bin km sonra arızalanmalarının standart sapması 45 olsun. Bu durumda üretilen otomobillerin ortalama kaç bin km den sonra arızalancağının %95 güven aralığını bulunuz.

Çözüm.

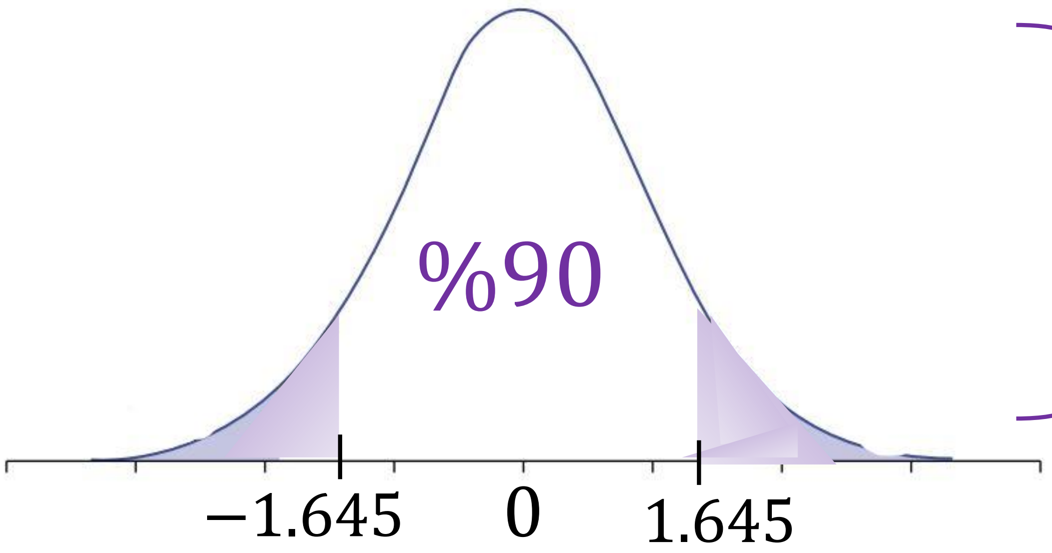
$$\bar{X} = \frac{175+180+185+198+148+162+191+205}{8} = 180.5$$

%95 güven aralığı $\left[180.5 - 1.96 \times \frac{45}{\sqrt{8}}, 180.5 + 1.96 \times \frac{45}{\sqrt{8}}\right] = [149.31, 211.68].$

Yani biz %95 eminizki populasyonun ortalaması $[149.31, 211.68]$ aralığındadır.



Bir başka önemli güven aralığı $\%90$ güven aralığıdır. Z standart normal rastgele değişkeninin $[-1.645, 1645]$ aralığında olma olasılığı 0.9 dur:



$$\boxed{P(-1.645 \leq Z \leq 1.645) = 0.9}$$

Populasyon ortalaması olan μ'nün $\left[\bar{X} - 1.645 \times \frac{\sigma}{\sqrt{n}}\ ,\ \bar{X} + 1.645 \times \frac{\sigma}{\sqrt{n}}\right]$ aralığında olma olasılığı 0.9'dur.

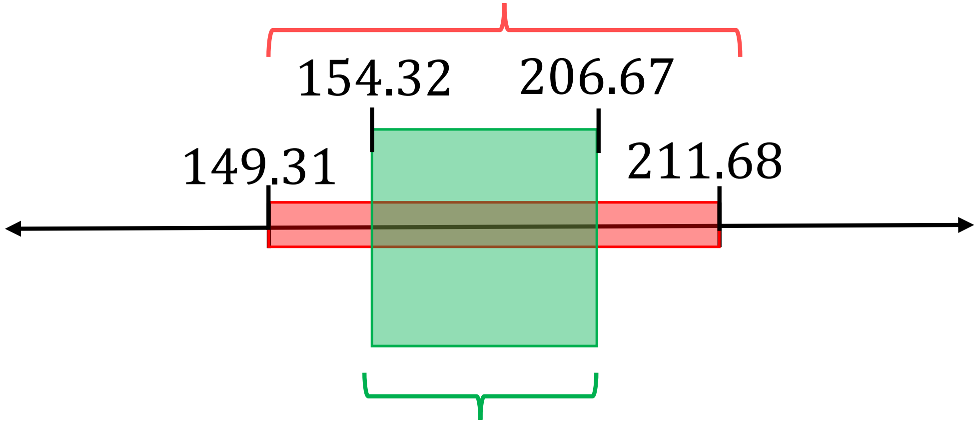
Dikkat edilirse bu aralık, $\%95$ aralığına göre daha dar bir aralıktır. O halde aralık darlaştıkça (genişledikçe), μ'nün bu aralığa düşme şansı azalır (artar).

Bir önceki soru için %90 güven aralığını hesaplayalım.

$\bar{X} = 180.5$ ve $\sigma = 45$ idi. O halde μ'nün %90 güven aralığı:

$$\left[180.5 - 1.645 \times \frac{45}{\sqrt{8}}, 180.5 + 1.645 \times \frac{45}{\sqrt{8}}\right] = [154.32, 206.67]$$

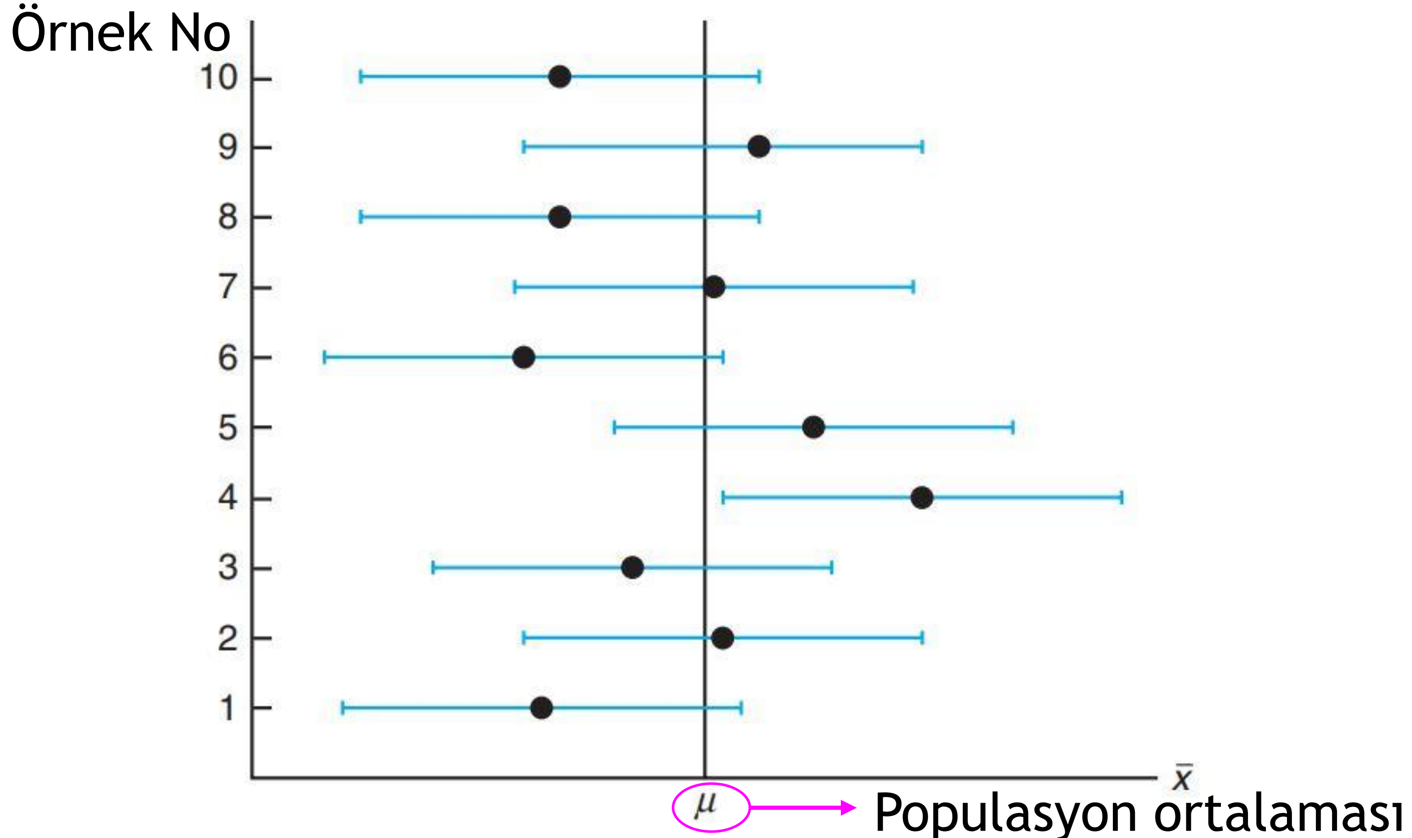
Ortalama %95 olasılıkla bu aralığın içinde!



Ortalama %90 olasılıkla bu aralığın içinde!

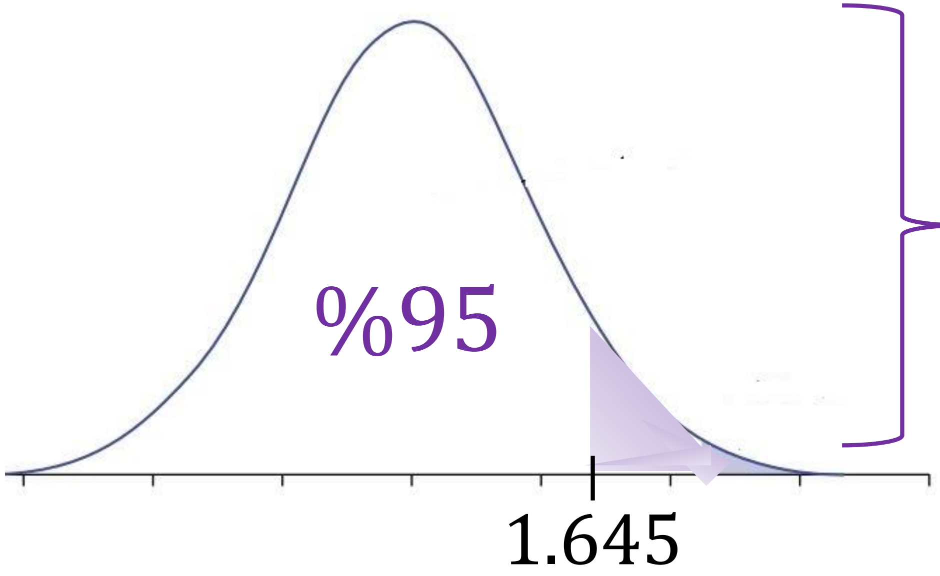
Unutulmaması gerekirki güven aralıkları eldeki örnekten elde edilir. O yüzden aynı populayondan alınan farklı örneklerden elde edilen güven aralıkları birbirinden farklı olabilir.

Örneğin aşağıda aynı populasyondan alınan 10 farklı örnekten elde edilen güven aralıkları verilmiştir (mavi aralıklar). Bu örneklerin kendi ortalamaları ise nokta (•) ile verilmiştir. Görülüyorki 10 farklı örnekten elde edilen 10 farklı güven aralığının 9'u populasyon ortalaması olan μ'yü içeriyor.



Alt ve Üst Güven Aralıkları

Bazen populasyon ortalamasının yalnızca bır üst sınırdan daha küçük olması, yada yalnızca bir alt sınırdan daha büyük olması ile ilgileniriz. Yani populasyon ortalamasının alabileceği en büyük değer yada bu ortalamanın alabilceği en küçük değer ile ilgileniriz.

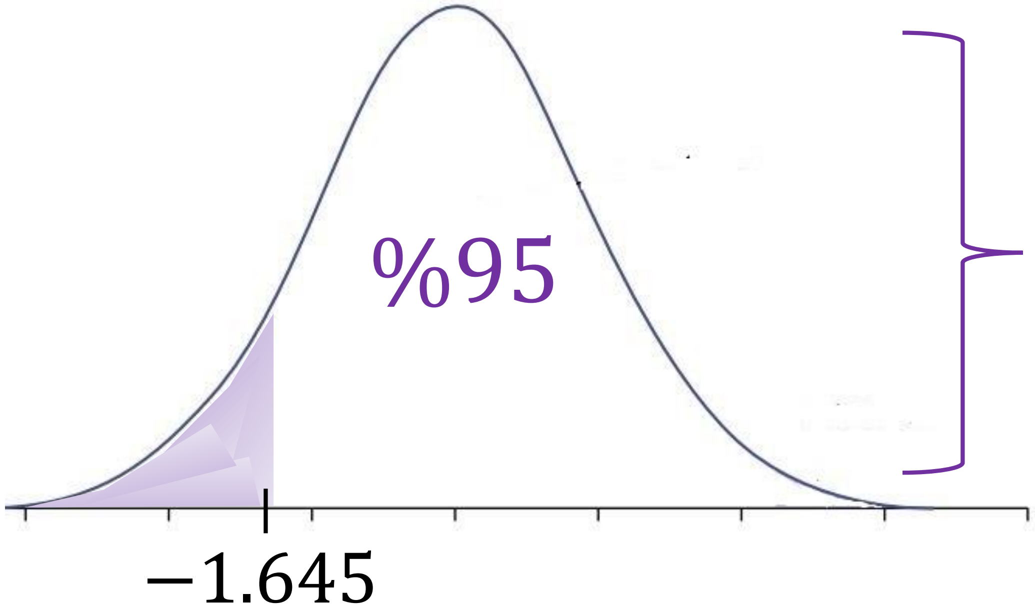


$$P(Z \le 1.645) = 0.95$$

$$P(Z \le 1.645) = P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \le 1.645\right) = P\left(\mu \ge \bar{X} - 1.645 \times \frac{\sigma}{\sqrt{n}}\right) = 0.95$$

$$P\left(\mu \geq \bar{X} - 1.645 \times \frac{\sigma}{\sqrt{n}}\right) = 0.95$$

Populasyon ortalaması μ, %95 kesinlikle en az $\bar{X} - 1.645 \times \frac{\sigma}{\sqrt{n}}$ ‘dir. Başka bir deyişle %95 eminizki populasyon ortalaması $\bar{X} - 1.645 \times \frac{\sigma}{\sqrt{n}}$ değerinden büyüktür.



$$\boxed{P(Z \geq -1.645) = 0.95}$$

$$P(Z \geq -1.645) = P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \geq 1.645\right) = P\left(\mu \leq \bar{X} + 1.645 \times \frac{\sigma}{\sqrt{n}}\right) = 0.95$$

$$P\left(\mu \leq \bar{X} + 1.645 \times \frac{\sigma}{\sqrt{n}}\right) = 0.95$$

Populasyon ortalaması μ, %95 kesinlikle en fazla $\bar{X} + 1.645 \times \frac{\sigma}{\sqrt{n}}$ ‘dir. Başka bir deyişle %95 eminizki populasyon ortalaması $\bar{X} + 1.645 \times \frac{\sigma}{\sqrt{n}}$ değerinden küçüktür.

ör. 16 kişilik bir sınıfta öğrencilerin boyları cm cinsinden

$$188, 181, 165, 172, 181, 168, 165, 190, 177, 176, 170, 161, 182, 194, 157, 173$$

olsun. Bu sınıfın dahil olduğu okuldaki tüm kişilerin boy ortalamalarının standart sapmasının bilindiği ve bu standart sapmanın 14 olduğu varsayılsın. Buna göre

i) okuldaki kişilerin boy ortalamasını %90 ve %95 güven aralığı içersinde veriniz.

ii) okuldaki kişilerin boy ortalaması %95 kesinlikle en fazla kaç olur?

iii) okuldaki kişilerin boy ortalaması %95 kesinlikle en az kaç olur?

Çözüm. Örneğin ortalaması:

$$\bar{X} = \frac{188 + 181 + 165 + 172 + 181 + 168 + 165 + 190 + 177 + 176 + 170, +161 + 1 + 194, +157 + 173}{16}$$

$$\bar{X} = 175$$

Örnek büyüklüğü: $n = 16$

Populasyon standart sapması: $\sigma = 14$

Populasyon ortalaması μ'nün %90 güven aralığı: $\left[\bar{X} - 1.645 \times \frac{\sigma}{\sqrt{n}} \text{ , } \bar{X} + 1.645 \times \frac{\sigma}{\sqrt{n}}\right]$

(yani bir başka deyişle biz %90 eminizki μ bu ralığın içindedir)

O halde %90 güven aralığı: $\left[175 - 1.645 \times \frac{14}{\sqrt{16}}, 175 + 1.645 \times \frac{14}{\sqrt{16}}\right] = [169.24, 180.75]$

Populasyon ortalaması μ'nün %95 güven aralığı: $\left[\bar{X} - 1.96 \times \frac{\sigma}{\sqrt{n}} \text{ , } \bar{X} + 1.96 \times \frac{\sigma}{\sqrt{n}}\right]$

O halde %90 güven aralığı: $\left[175 - 1.96 \times \frac{14}{\sqrt{16}}, 175 + 1.96 \times \frac{14}{\sqrt{16}}\right] = [168.14, 181.86]$

Çözüm.

ii) $P\left(\mu \geq \bar{X} - 1.645 \times \frac{\sigma}{\sqrt{n}}\right) = 0.95$ idi. Şu halde $\bar{X}$, σ ve $\sqrt{n}$ yerlerine konulursa:

$$P\left(\mu \geq 175 - 1.645 \times \frac{14}{\sqrt{16}}\right) = P(\mu \geq 169.24) = 0.95$$

olur. Yani %95 eminizki populasyon ortalaması 169.24'ten büyüktür.

iii) Üst sınır için $P\left(\mu \leq \bar{X} + 1.645 \times \frac{\sigma}{\sqrt{n}}\right) = 0.95$ idi.

$$P\left(\mu \leq 175 + 1.645 \times \frac{14}{\sqrt{16}}\right) = 0.95$$

$$P(\mu \leq 180.75) = 0.95$$

%95 eminizki populasyon ortalaması 180.75'ten küçüktür.

Populasyon Standart Sapması Bilinmezken Ortalama İçin Güven Aralığı

Şu ana kadar populasyon ortalaması (μ) için bir güven aralığı oluştururken populasyon standart sapması olan σ'yı bildiğimizi varsayıyorduk. Şimdi ise σ'yı bilmediğimizi varsayalım; ki bu durumla karşılaşmak çok daha olasıdır.

Merkezi limit teoreminden gördükki

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

standart normal dağılıma sahipti (yani ortalaması 0, standart sapması 1 idi).

Şimdi ise populasyon standart sapması σ'yı bilmediğimizden, bunun yerine örneğin standart sapması olan s'yi yazarsak:

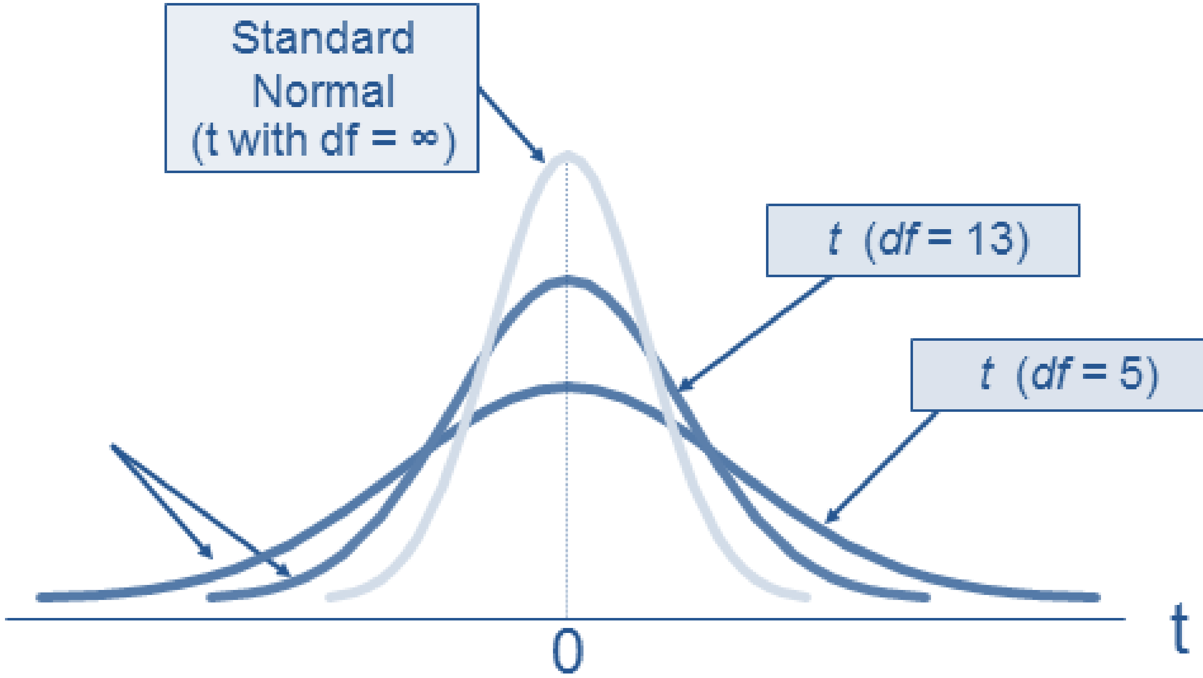
$$T = \frac{\bar{X} - \mu}{s/\sqrt{n}}$$

t-dağılımına (t-distribution); tam olarak t_{n-1} dağılımına sahip olur. Burada $n-1$ serbestlik derecesidir(degree of freedom).

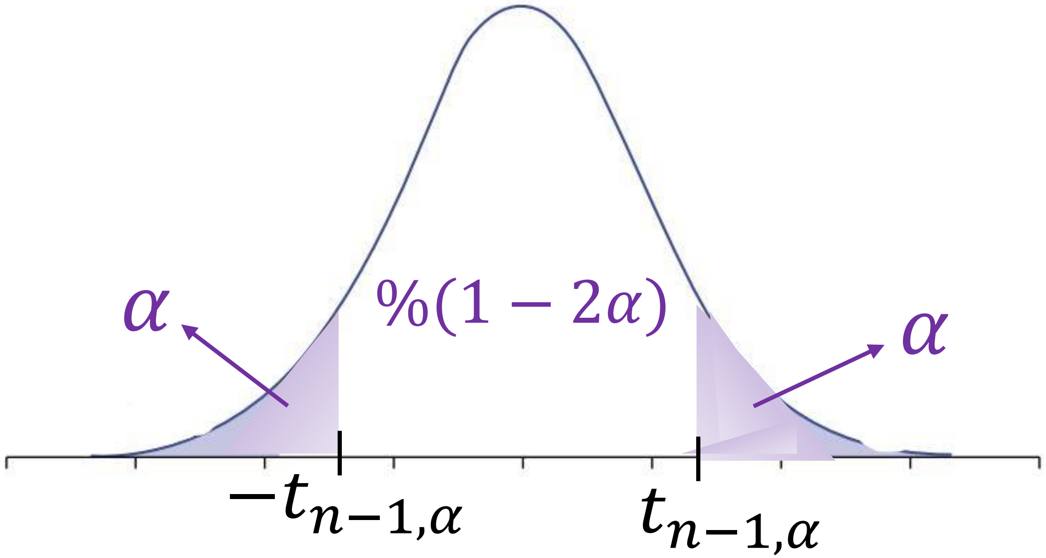
t-dağılımları

t-dağılımları genel olarak standart normal dağılıma benzer: çan eğrileri şeklindedir fakat biraz daha düzdür, pek sivri değildir.

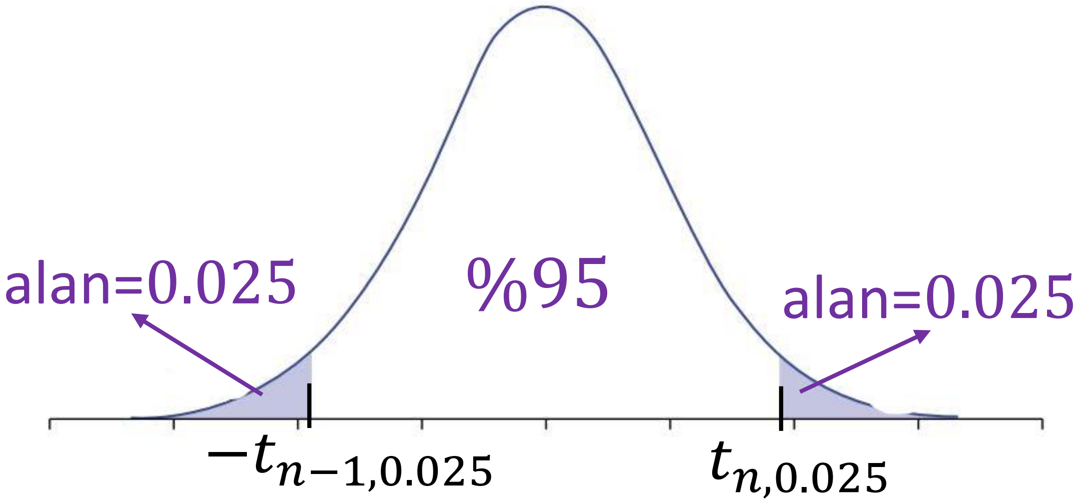
t-dağılımları tek bir dağılım değildir. Serbestlik derecesi $(n-1)$ değiştikçe farklılık gösterirler. Genel olarak n (örnek büyüklüğü) arttıkça standart normal dağılıma yaklaşırlar.

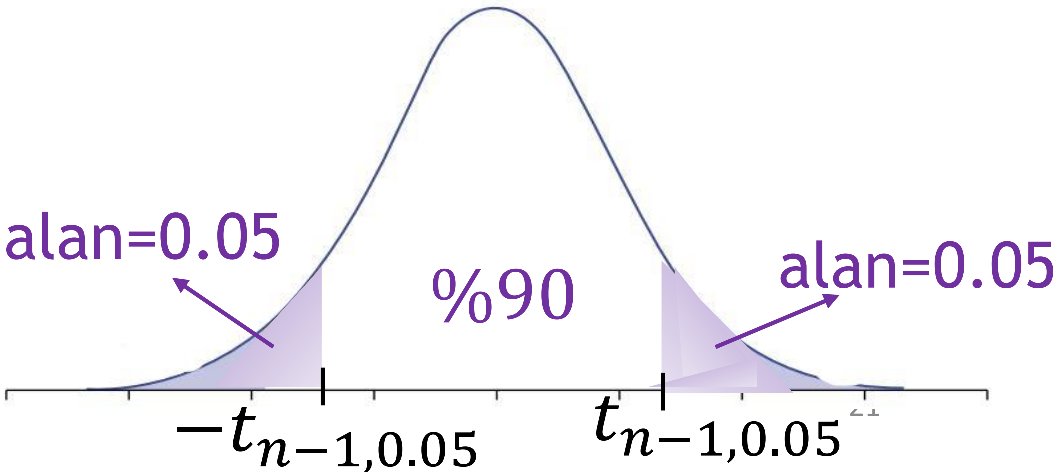


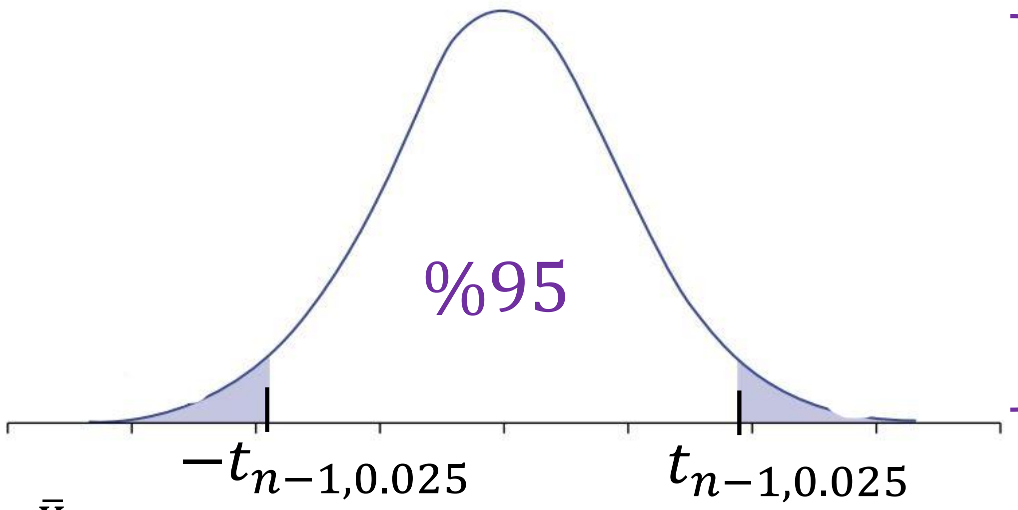
t-dağılımları



$t_{n,\alpha}$ ‘nın sağındaki alanın büyüklüğü α kadardır. $-t_{n,\alpha}$ ‘nın solundaki alanın büyüklüğü α kadardır. Biz genelde %95 ve %90 güven aralıkları ile ilgileniriz. %95 güven aralığını veren t değerleri $-t_{n,0.025}$ ve $-t_{n,0.025}$; %90 güven aralığını veren t değerleri $-t_{n,0.05}$ ve $-t_{n,0.05}$ dir.







$$\boxed{P\left(-t_{n-1,0.025} \leq T \leq t_{n-1,0.025}\right) = 0.95}$$

$T = \frac{\bar{X}-\mu}{s/\sqrt{n}}$ idi. $-t_{n-1,0.025} \leq T \leq t_{n-1,0.025}$ eşitsizliğinde yerine yazılırsa, populasyon ortalaması için %95 güven aralığı:

$$\left[\bar{X} - t_{n-1,0.025} \times \frac{s}{\sqrt{n}}, \bar{X} + t_{n-1,0.025} \times \frac{s}{\sqrt{n}}\right]$$

olur. Burada $t_{n,0.025}$ değerleri tablodan bulunur.

ör. Bir önceki soru için, okuldaki kişilerin boylarının standart sapmalarını bilmediğimiz varsayarak yalnızca elimizdeki veriye dayanarak okul boy ortalaması için %95 ve %90 güven aralıklarını bulalım.

Çözüm. Örneğin standart sapması:

$$s = \sqrt{\frac{1}{15} \times \left((188-175)^2 + (181-175)^2 + \cdots + (173-175)^2\right)} = 10.6$$

Table D.2 Percentiles $t_{n,\alpha}$ of t Distributions

n	α									
	0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657	127.32	318.31	636.62
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925	14.089	23.326	31.598
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841	7.453	10.213	12.924
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	0.265	0.718	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	0.263	0.711	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	0.262	0.706	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	0.261	0.703	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	0.260	0.697	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	0.259	0.695	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	0.259	0.694	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	0.258	0.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	0.258	0.691	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	0.258	0.690	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	0.257	0.689	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	0.257	0.688	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922

$t_{15,0.025} = 2.131$

Populasyon ortalaması için %95 güven aralığı:

$$\left[\bar{X} - t_{n-1,0.025} \times \frac{s}{\sqrt{n}}, \bar{X} + t_{n-1,0.025} \times \frac{s}{\sqrt{n}}\right]$$

$$\left[175 - 2.131 \times \frac{10.6}{\sqrt{16}}, 175 + 2.131 \times \frac{10.6}{\sqrt{16}}\right] = [169.35\,, 180.64]$$

Ikinci olarak %90 güve aralğını bulalım. Bunun için $t_{15,0.05}$ değerini tablodan bulmamız gerekir. Bir önceki slayttaki tabloya bakılırsa, bu değer 1.753'tür. O halde %90 güven aralığı:

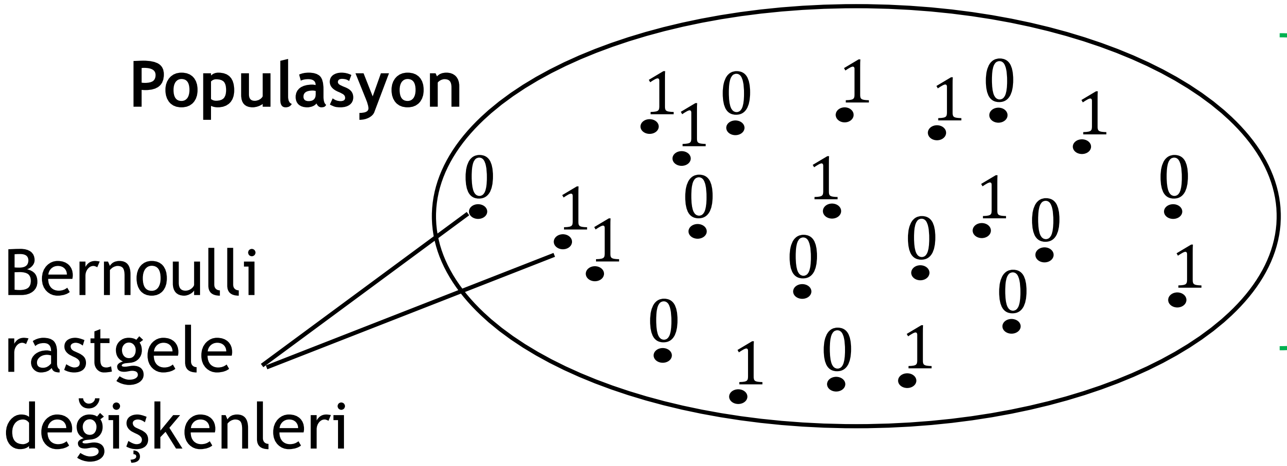
$$\left[175 - 1.753 \times \frac{10.6}{\sqrt{16}}, 175 + 1.753 \times \frac{10.6}{\sqrt{16}}\right] = [170.35\,, 179.64]$$

Dikkat edilirse t-dağılımı yardımıyla bulunan %90 güven aralığı, yine t-dağılımı yardımıyla bulunan %95 güven aralığından daha dardır; tıpkı populasyon standart sapmasını bildiğimizi varsaydığımız durumda standart normal dağılım yarımıyla bulduğumuz %90 ve %95 güven aralıkları gibi.

Populasyon Oranı Için Güven Aralığı

Tıpkı populasyon ortalamasında olduğu gibi, populasyon oranı içinde güven aralığı oluşturabiliriz.

Daha önce normal dağılım yardımıyla binomial dağılıma yaklaşmıştık:



Populasyon ortalaması (aynı zamanda oranı): p
Populasyon standart sapması: $\sqrt{p(1-p)}$

Örnek

$X_1 = 0$
$X_2 = 1$
$X_n = 1$

Bu örneğin ortalaması:

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n} = \frac{X}{n} = \hat{p}$$

Örneğin ortalaması aynı zamanda oranı olan $\hat{p} = \frac{X_1+X_2+\cdots+X_n}{n}$, merkezi limit teoremi gereği normal dağılıma sahiptir. Örneklerden alınan bu oranların ortalaması p (populasyon oranı), standart sapması $\frac{\sqrt{p(1-p)}}{\sqrt{n}} = \sqrt{\frac{p(1-p)}{n}}$ dir.

Normal dağılıma sahip $\hat{p}$'yi standart hale getirisek:

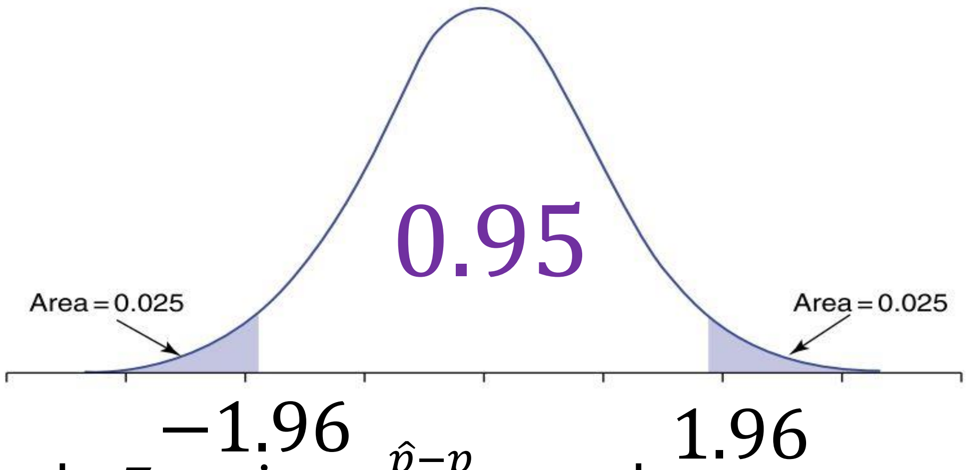
$$Z \leftarrow \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

olur. Populasyon oranı olan p yerine onun bir tahmini olan örnekten elde ettiğimiz $\hat{p}$'yi yazabiliriz. Bu durumda:

$$Z \leftarrow \frac{\hat{p} - p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}$$

yazabiliriz. Bu değişken standart normal dağılıma sahip olur!

Daha önceden de görmüştük ki:



$$\boxed{P(-1.96 \ \leq Z \leq 1.96) = 0.95}$$

Burada Z yerine $\frac{\hat{p}-p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}$ yazılırsa:

$$P\left(-1.96 \ \leq \frac{\hat{p}-p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} \leq 1.96\right) = 0.95$$

$$P\left(\hat{p} - 1.96 \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq \ p \ \leq \hat{p} + 1.96 \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right) = 0.95$$

Sonuç olarak populasyon oranı p için %95 güven aralığı:

$$\left[\hat{p} - 1.96 \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + 1.96 \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

olur. Eğer 1.96 değeri yerine 1.645 yazılırsa p için %90 güven aralığı elde edilir:

$$\left[\hat{p} - 1.645 \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + 1.645 \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

ör. Daha önceki örnekte önümüzdeki seçimlerde seçmenlerin kime oy vereceğini tahmmin etmek için bir anket yapıldığını, 300 kişilik bu anketdee 210 kişi X partisine oy vereceğini söyledğini belirtmiştik. Bu tahminin %95 ve %90 güven aralığı ne olur?

Çözüm.

$$\hat{p} = \frac{210}{300} = 0.7$$

%95 güven aralığı:

$$\left[0.7 - 1.96 \times \sqrt{\frac{0.7 \times 0.3}{300}}, 0.7 + 1.96 \times \sqrt{\frac{0.7 \times 0.3}{300}}\right] = [0.64, 0.75]$$

%90 güven aralığı:

$$\left[0.7 - 1.645 \times \sqrt{\frac{0.7 \times 0.3}{300}}, 0.7 + 1.645 \times \sqrt{\frac{0.7 \times 0.3}{300}}\right] = [0.65, 0.74]$$

olur. Buna göre biz %95 eminizki tüm seçmeler içersinde X partisine oy verme oranı $[0.64, 0.75]$ arasındadır; %90 eminizki tüm seçmeler içersinde X partisine oy verme oranı $[0.65, 0.74]$ arasındadır.