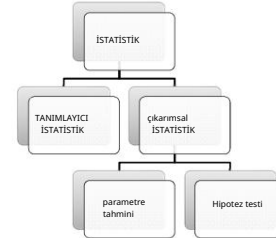


İSTATİSTİK

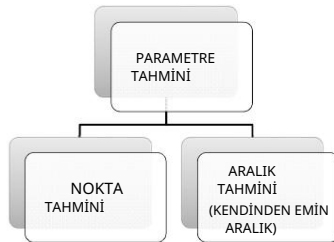
Güvenilirlik aralığı

Gözden geçirmek

İstatistik, karar vermek için veri toplama, düzenleme, analiz etme ve yorumlama bilimidir.



Gözden geçirmek

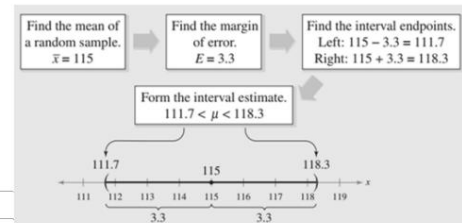


Motive edici örnek

Örneğin, bir satranç kulübü, üyelerinin ortalama IQ'sunu tahmin etmek istiyor. Rastgele bir üye örneğinin ortalaması 115'tir. Bu tahmin, sayı doğrusu üzerindeki bir noktayla temsil edilen tek bir sayıdan oluştuğu için buna nokta tahmini denir.

Bir nokta tahmini kullanmanın sorunu, nadiren popülasyonun tam parametresine (ortalama, standart sapma veya orantı) eşit olmasıdır.

Diyeelim ki kulüp, üyelerinin ortalama IQ'larına ilişkin tahmininden %90 emin olmak istiyor. Aralık Tahmini



Bölüm 6.1

Ortalama için Güven Aralıkları
(Bilinen)

Nüfus için Nokta Tahmini μ

Nokta tahmini

- Bir popülasyon parametresi için tek bir değer tahmini • Popülasyon ortalamasının en yansız nokta tahmini μ is the sample mean \bar{x}

Estimate Population Parameter...	with Sample Statistic
Mean: μ	\bar{x}

Örnek: için Puan Tahmini Nüfus μ (2'de 1)

A social networking website allows its users to add friends, send messages, and update their personal profiles. The following represents a random sample of the number of friends for 40 users of the website. Find a point estimate of the population mean μ . (Adapted from Facebook)

140 105 130 97 80 165 232 110 214 201 122
98 65 88 154 133 121 82 130 211 153 114
58 77 51 247 236 109 126 132 125 149 122
74 59 218 192 90 117 105

Örnek: için Puan Tahmini Nüfus μ (2/2)

The sample mean of the data is

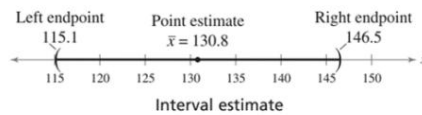
$$\bar{x} = \frac{\sum x}{n} = \frac{5232}{40} = 130.8.$$

web sitesinin tüm kullanıcıları için ortalama arkadaş sayısı için puan tahmini 130,8 arkadaşdır.

Aralık Tahmini

Aralık tahmini

- Bir popülasyon parametresini tahmin etmek için kullanılan bir aralık veya değer aralığı.

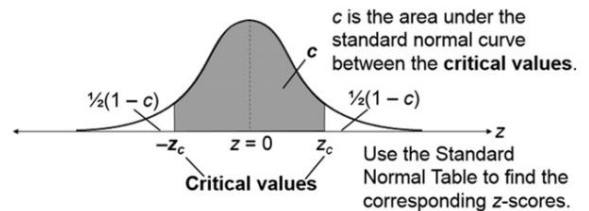


How confident do we want to be that the interval estimate contains the population mean μ ?

Güven Düzeyi (1/2)

Güven düzeyi c

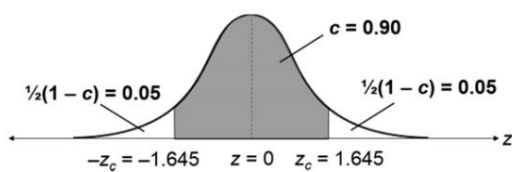
- Aralık tahmininin popülasyon parametresini içermesi olasılığı.



The remaining area in the tails is $1 - c$.

Güven Düzeyi (2/2)

Güven düzeyi %90 ise bu, aralığın nüfus ortalamasını μ içerdiğinden %90 emin olduğumuz anlamına gelir.



The corresponding z-scores are ± 1.645 .

Güven seviyesi

Level of Confidence	z_c
90%	1.645
95%	1.96
99%	2.575

Örnekleme Hatası

örnekleme hatası

- Nokta tahmini ile nokta tahmini arasındaki fark gerçek popülasyon parametre değeri.
- For μ :
 - the sampling error is the difference $\bar{x} - \mu$
 - μ is generally unknown
 - \bar{x} varies from sample to sample

Hata Marjı

Hata payı • Belirli

bir güven düzeyi için nokta tahmini ile tahmin ettiği parametrenin değeri arasındaki mümkün olan en büyük mesafe, c.

- Denoted by E .

$$E = z_c \sigma_{\bar{x}} = z_c \frac{\sigma}{\sqrt{n}}$$

1. The sample is random.
2. Population is normally distributed or $n \geq 30$. Sometimes called the maximum error of estimate or error tolerance.

Örnek: Marjını Bulma Hata (3'te 1)

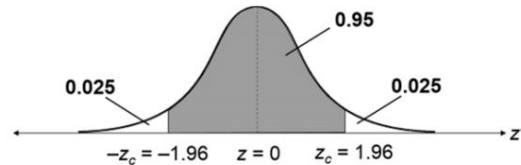
Web sitesinin tüm kullanıcıları için ortalama arkadaş sayısına ilişkin hata payını bulmak için örnekteki verileri ve %95'lik bir güven düzeyini kullanın.

Numune standart sapmasının yaklaşık 53.0 olduğunu varsayın

Örnek: Marjını Bulma Hata (2/3)

Solution

- First find the critical values



Standart normal eğri altındaki alanın %95'i, ortalamanın 1,96 standart sapması içinde yer alır. ($N = 40$ 30 olduğu için, örnek ortalamaların dağılımını Merkezi Limit Teoremi ile normal bir eğri ile yaklaşık olarak hesaplayabilirsiniz.)

Örnek: Marjını Bulma Hata (3/3)

Using the values $z_c = 1.96$, $\sigma \approx s \approx 53.0$, and $n = 40$,

$$\begin{aligned} E &= z_c \frac{\sigma}{\sqrt{n}} \\ &\approx 1.96 \cdot \frac{53.0}{\sqrt{40}} \\ &\approx 16.4. \end{aligned}$$

Nüfus ortalaması için hata payının yaklaşık 16,4 arkadaş olduğundan %95 eminsiniz.

için Güven Aralıkları

Nüfus Ortalaması

A c-confidence interval for the population mean μ

$$\bar{x} - E < \mu < \bar{x} + E \text{ where } E = z_c \frac{\sigma}{\sqrt{n}}$$

- Tahmin sürecinin çok sayıda tekrarlandığı varsayıldığında, güven aralığının μ içermesi olasılığı c'dir.

μ için Güven Aralıkları Oluşturma (1/2)

Nüfus Ortalaması için Güven Aralığı Bulma (σ Biliniyor)

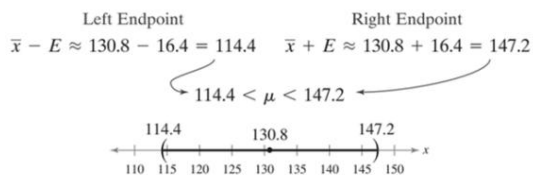
In Words	In Symbols
1. Verify that σ known, sample is random, and either the population is normally distributed or $n \geq 30$.	
2. Find the sample statistics n and \bar{x} .	$\bar{x} = \frac{\sum x}{n}$

μ için Güven Aralıkları Oluşturma (2/2)

In Words	In Symbols
3. Find the critical value z_c that corresponds to the given level of confidence.	Use Table 4, Appendix B.
4. Find the margin of error E .	$E = z_c \frac{\sigma}{\sqrt{n}}$
5. Find the left and right endpoints and form the confidence interval.	Left endpoint: $\bar{x} - E$ Right endpoint: $\bar{x} + E$ Interval: $\bar{x} - E < \mu < \bar{x} + E$

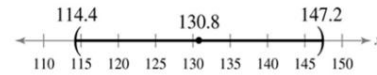
Örnek 1: Güven Oluşturma Aralık (1/2)

Web sitesinin tüm kullanıcıları için ortalama arkadaş sayısı için %95'lik bir güven aralığı oluşturun.



Örnek 1: Güven Oluşturma Aralık (2/2)

$$114.4 < \mu < 147.2$$



Nüfusun ortalama arkadaş sayısının 114.4 ile 147.2 arasında olduğunu %95 güvenle söyleyebiliriz.

Örnek 2: Güven Oluşturma Aralık (1/4)

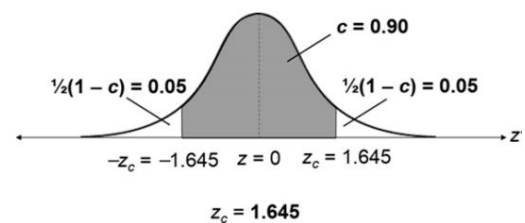
Bir kolej kabul müdürü, şu anda kayıtlı olan tüm öğrencilerin ortalama yaşını tahmin etmek istemektedir. 20 öğrenciden oluşan rastgele bir örnekte ortalama yaş 22,9 olarak bulunmuştur. Geçmiş çalışmalardan, standart sapmanın 1,5 yıl olduğu ve popülasyonun normal dağıldığı bilinmektedir. Nüfus ortalama yaşının %90 güven aralığını oluşturun.



Örnek 2: Güven Oluşturma Aralık (2/4)

Çözüm

• Önce kritik değerleri bulun



Örnek 2: Güven Oluşturma Aralık (3/4)

- Margin of error:

$$E = z_c \frac{\sigma}{\sqrt{n}} = 1.645 \cdot \frac{1.5}{\sqrt{20}} \approx 0.6$$

- Confidence interval:

Left Endpoint:

$$\bar{x} - E$$

$$\approx 22.9 - 0.6$$

$$= 22.3$$

Right Endpoint:

$$\bar{x} + E$$

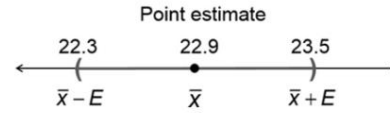
$$\approx 22.9 + 0.6$$

$$= 23.5$$

$$22.3 < \mu < 23.5$$

Örnek 2: Güven Oluşturma Aralık (4/4)

$$22.3 < \mu < 23.5$$



%90 güvenle tüm öğrencilerin yaş ortalamasının 22,3 ile 23,5 arasında olduğunu söyleyebilirsiniz.

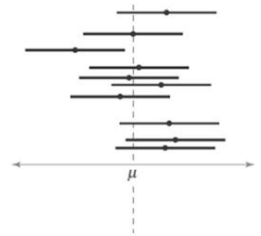
Sonuçları Yorumlama (1/2)

- μ sabit bir sayıdır. Ya güven aralığındadır ya da değildir.
- Yanlış: "Gerçek ortalamanın (22.3, 23.5) aralığında olma olasılığı %90'dır."
- Düzeltilen: "Eğer çok sayıda numune toplanırsa ve her numune için bir güven aralığı oluşturulur, bu aralıkların yaklaşık %90'ı μ içerecektir."

Sonuçları Yorumlama (2/2)

Yatay bölümler, farklı uygulamalar için %90 güven aralıklarını temsil eder.

Aynı büyüklükteki örnekler. Uzun vadede, bu tür her 10 aralığın 9'u μ içerecektir.



Örnek boyut

- Bir c-güven düzeyi ve bir hata payı E verildiğinde, popülasyon ortalamasını tahmin etmek için gereken minimum örneklem büyüklüğü n

$$n = \left(\frac{z_c \sigma}{E} \right)^2$$

- If σ is unknown, you can estimate it using s provided you have a preliminary sample with at least 30 members.

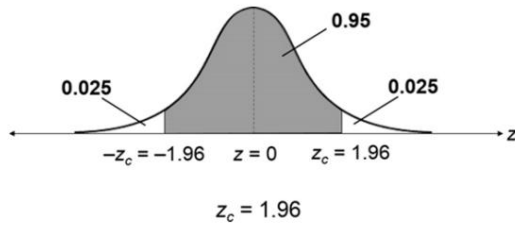
Örnek: Bir Minimum Belirleme Örnek Büyüklüğü (1/3)

Herkes için ortalama arkadaş sayısını tahmin etmek istiyorsunuz. web sitesinin kullanıcıları. Örneklem ortalamasının popülasyon ortalamasından yedi arkadaş içinde olduğundan %95 emin olmak istiyorsanız, örneğe kaç kullanıcı dahil edilmelidir? Numune standart sapmasının yaklaşık 53.0 olduğunu varsayalım.

Örnek: Bir Minimum Belirleme Örnek Büyüklüğü (2/3)

Çözüm

- Önce kritik değerleri bulun



Örnek: Bir Minimum Belirleme Örnek Büyüklüğü (3/3)

$$z_c = 1.96 \quad \sigma \approx s = 53.0 \quad E = 7$$

$$n = \left(\frac{z_c \sigma}{E} \right)^2 \approx \left(\frac{1.96 \cdot 53.0}{7} \right)^2 \approx 220.23$$

Gerektiğinde, bir tam sayı elde etmek için yuvarlayın .

Örnekleminize en az 221 kullanıcı dahil etmelisiniz .

Bölüm 6.2

Ortalama için Güven Aralıkları

(σ Unknown)

Slide 35

t-Dağılımı

- Kitle standart sapması bilinmediğinde, örneklem büyüklüğü 30'dan küçük olduğunda ve x rasgele değişkeni yaklaşık olarak normal dağıldığında, bir t dağılımı izler.

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

- t'nin kritik değerleri ile gösterilir.

t_c

t-Dağılımının Özellikleri (1/3)

1. t dağılımının ortalaması, ortancası ve modu 0'a eşittir.
2. t-dağılımı çan şeklinde ve simetrik bir ortalama hakkındadır.
3. Bir t-eğrisinin altındaki toplam alan 1'dir.
4. t dağılımındaki kuyruklar "daha kalın" standart normal dağılımdadır.
5. t dağılımının standart sapması değişir örneklem büyüklüğü ile, ancak 1'den büyüktür.

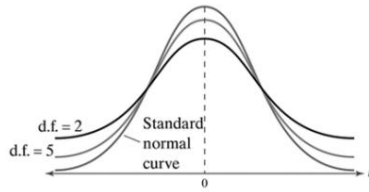
t-Dağılımının Özellikleri (2/3)

6. t-dağılımı, her biri bir eğri ailesidir. serbestlik derecesi adı verilen bir parametre tarafından belirlenir. Serbestlik dereceleri, gibi bir örnek istatistik hesaplandıktan sonra kalan serbest seçeneklerin sayısıdır. Bir t- \bar{x} dağılımı kullandığınızda bir popülasyon ortalamasını tahmin edin, serbestlik dereceleri örneklem büyüklüğünden bir eksik olur.

- d.f. = $n - 1$ Özgürlük derecesi

t-Dağılımının Özellikleri (3/3)

7. Serbestlik derecesi arttıkça t dağılım normal dağılıma yaklaşır. 30 veya daha fazla serbestlik derecesi için, t-dağılımı standart normal dağılıma yakındır.



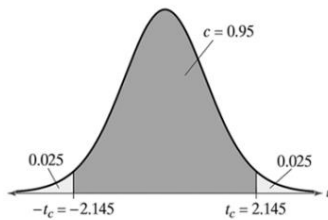
Örnek: t'nin Kritik Değerlerini Bulma

Bir örneklem büyüklüğü için kritik t_c 95% güven ne zaman değeri 15 olarak bulun.

	Level of confidence, c	0.80	0.90	0.95	0.98	0.99
	One tail, α	0.10	0.05	0.025	0.01	0.005
d.f.	Two tails, α	0.20	0.10	0.05	0.02	0.01
1		3.078	6.314	12.706	31.821	63.657
2		1.886	2.920	4.303	6.965	9.925
3		1.638	2.353	3.182	4.541	5.841
12		1.356	1.782	2.179	2.681	3.055
13		1.350	1.771	2.160	2.650	3.012
14		1.345	1.761	2.145	2.624	2.977
15		1.341	1.753	2.131	2.602	2.947
16		1.337	1.746	2.120	2.583	2.921

Çözüm: t'nin Kritik Değerleri (1/2)

Tablodan bunu anlayabilirsiniz. Şekil, 14. serbestlik derecesi için t dağılımını göstermektedir ve $c = 0.95$, $t_c = 2.145$.



için bir Güven Aralığı Oluşturma Nüfus Ortalaması (Sigma Bilinmiyor) (1/2)

kelimelerle

Sembollerde

1. Bilinmediğini, örneğin rastgele olduğunu ve popülasyonun normal dağıldığını veya $n \geq 30$.

2. Örneği ve s'yı bulun. istatistik n , \bar{x}

$$\bar{x} = \frac{\sum x}{n} \quad s = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}}$$

için bir Güven Aralığı Oluşturma Nüfus Ortalaması (Sigma Bilinmiyor) (2/2)

kelimelerle

Sembollerde

3. Serbestlik derecelerini, güven düzeyini c ve kritik değeri tanımlayın.

t_c

d.f. = $n - 1$;
Ek B'deki Tablo
5'i kullanın.

4. E hata payını bulun.

$$E = t_c \frac{s}{\sqrt{n}}$$

5. Solu ve sağı bulun uç noktaları ve güven aralığını oluşturun.

Sol uç nokta: $\bar{x} - E$
Sağ uç nokta: $\bar{x} + E$
Aralık: $\bar{x} - E < \mu < \bar{x} + E$

Örnek: Güven Oluşturma Aralık (1/2)

Rastgele 16 kahve dükkanı seçiyorsunuz ve her birinde satılan kahvenin sıcaklığını ölçüyorsunuz. Numune ortalama sıcaklığı, numune standart sapması ile . Satılan kahvenin popülasyon ortalama sıcaklığı için bir güven aralığı oluşturun.

Sıcaklıkların yaklaşık olarak normal dağıldığını varsayalım.

Çözüm: Güven Oluşturmak Aralık (6'da 1)

Bilinmediği için, numune rastgele olduğundan ve sıcaklıklar yaklaşık olarak normal dağıldığından, t dağılımını kullanın. kullanarak, bunu bulmak için Tablo 5'i kullanabilirsiniz. $\bar{x} = 162.0$, $s = 10.0$, $c = 0.95$, Ve $d.f. = 15$
 $t_c = 2.131$.

Hata payı 95% güven seviyesi

$$E = t_c \cdot \frac{s}{\sqrt{n}} = 2.131 \cdot \frac{10}{\sqrt{16}} \approx 5.3$$

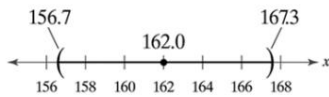
Çözüm: Güven Oluşturmak Aralık (2/6)

• Güven aralığı:

Sol Bitiş Noktası:	Sağ Uç Nokta:
$\bar{x} - E$	$\bar{x} + E$
$\approx 162 - 5.3$	$\approx 162 + 5.3$
$= 156.7$	$= 167.3$
$156.7 < \mu < 167.3$	

Çözüm: Güven Oluşturmak Aralık (3/6)

• $156.7 < \mu < 167.3$



Güvenli, satılan kahvenin ortalama sıcaklığının ve arasında olduğunu söyleyebilirsiniz. 156.7°F 167.3°F .

Örnek: Güven Oluşturma Aralık (2/2)

Bir araba galerisinde satılan aynı modelden 36 arabayı rastgele seçersiniz ve her arabanın satılmadan önce bayinin arsasında durduğu gün sayısını belirlersiniz. Örnek ortalaması 9,75 gündür ve örnek standart sapması 2,39 gündür.

Araba modelinin bayii arsasında bulunduğu ortalama gün sayısı için bir güven aralığı oluşturun.

Çözüm: Güven Oluşturmak Aralık (4/6)

Çözüm

Bilinmediği için, örnek rastgeledir ve t dağılımını kullanır. ve $n = 36 \geq 30$, kullanarak, kullanabilirsiniz. Hata $n = 36$, $\bar{x} = 9.75$, $s = 2.39$, $c = 0.99$, payı $d.f. = 35$
Güven düzeyini bulmak $t_c = 2.724$
 99% için Tablo 5

$$E = t_c \cdot \frac{s}{\sqrt{n}} = 2.724 \cdot \frac{2.39}{\sqrt{36}} \approx 1.09.$$

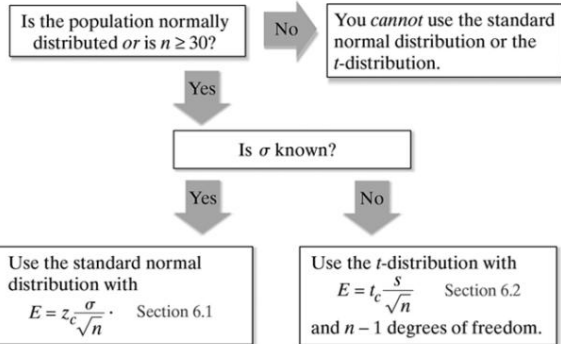
Çözüm: Güven Oluşturmak Aralık (5/6)

Çözüm

Güven aralığı gösterildiği gibi oluşturulur.

Sol Bitiş Noktası	Sağ Uç Nokta
$\bar{x} - E \approx 9.75 - 1.09$	$\bar{x} + E \approx 9.75 + 1.09$
$= 8.66$	$= 10.84$
$8.66 < \mu < 10.84$	

Normal mi yoksa t-Dağılımı mı?



Örnek: Normal mi yoksa t-Dağılımı mı?

Yeni inşa edilmiş 25 evi rastgele seçiyorsunuz. Örnek ortalama inşaat maliyeti ve popülasyon standart sapması . İnşaat maliyetlerinin normal dağılımı varsayarsak, normal dağılım, t-dağılımı mı kullanmalısınız, yoksa nüfus ortalama inşaat maliyeti oluşturmak için ikisini birden mi kullanmalısınız? 95% için güven aralığı Nedenini açıkla.

Çözüm: Normal mi yoksa t-Dağılımı mı?

Çözüm

Popülasyon normal dağılıyor mu yoksa Evet, $n \geq 30$?
popülasyon normal dağılıyor. Popülasyon normal olarak dağıldığı için standart normal dağılımı veya t-dağılımı kullanmaya devam edebileceğinizi unutmayın.

Bilgin?

Evet.

Karar:

Standart normal dağılımı kullanın.

Sonraki Ders

- TEK ÖRNEK İLE HİPOTEZ TESTİ