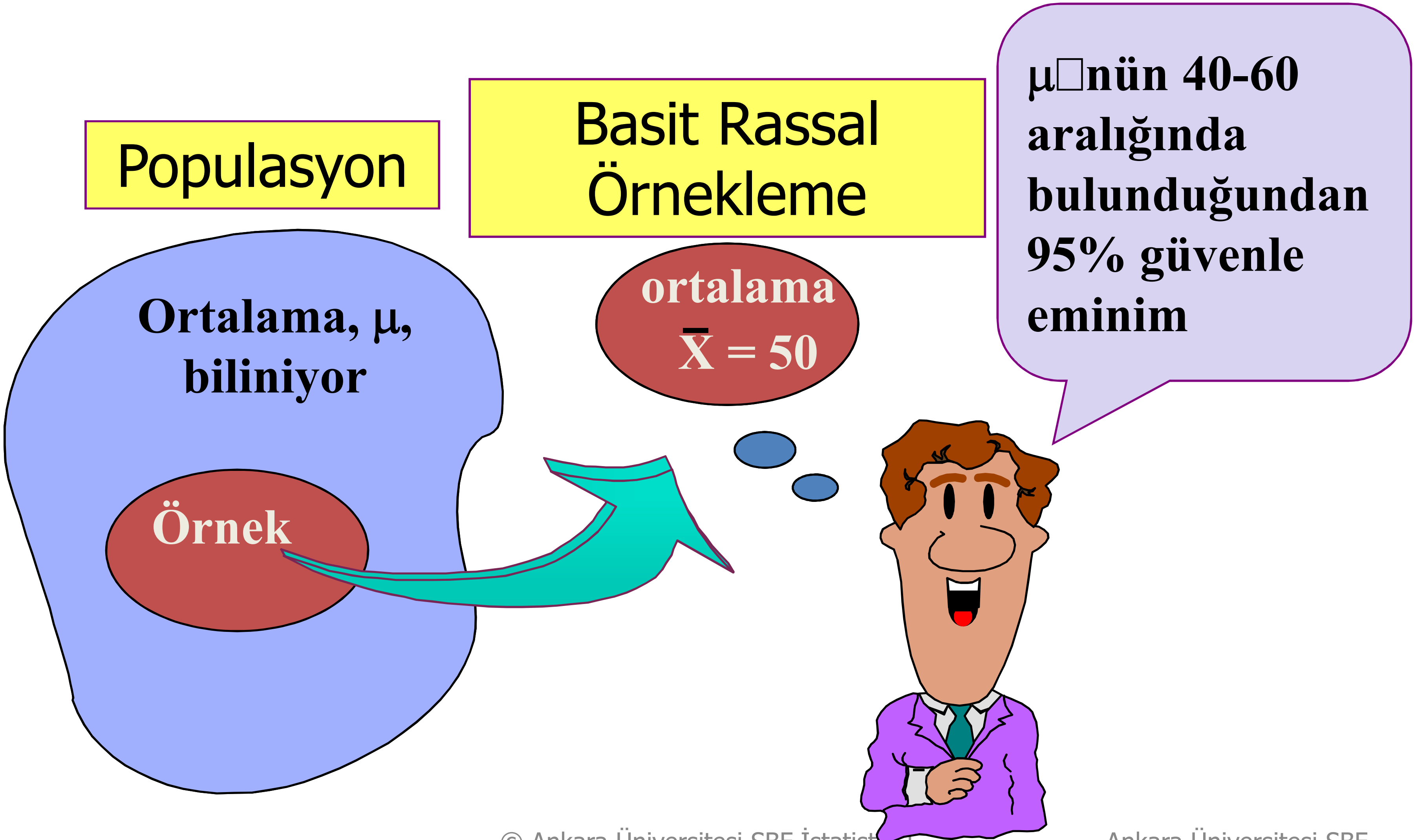
İstatistik 1
BÖLÜM 10
TAHMİN: GÜVEN ARALIKLARI

Bu Bölümde işlenecek Konular

- Tahmin süreci
- Nokta tahminleri
- Aralık tahminleri
- Güven aralıkları
- Örnek büyüklüğünün belirlenmesi

Tahmin Süreci



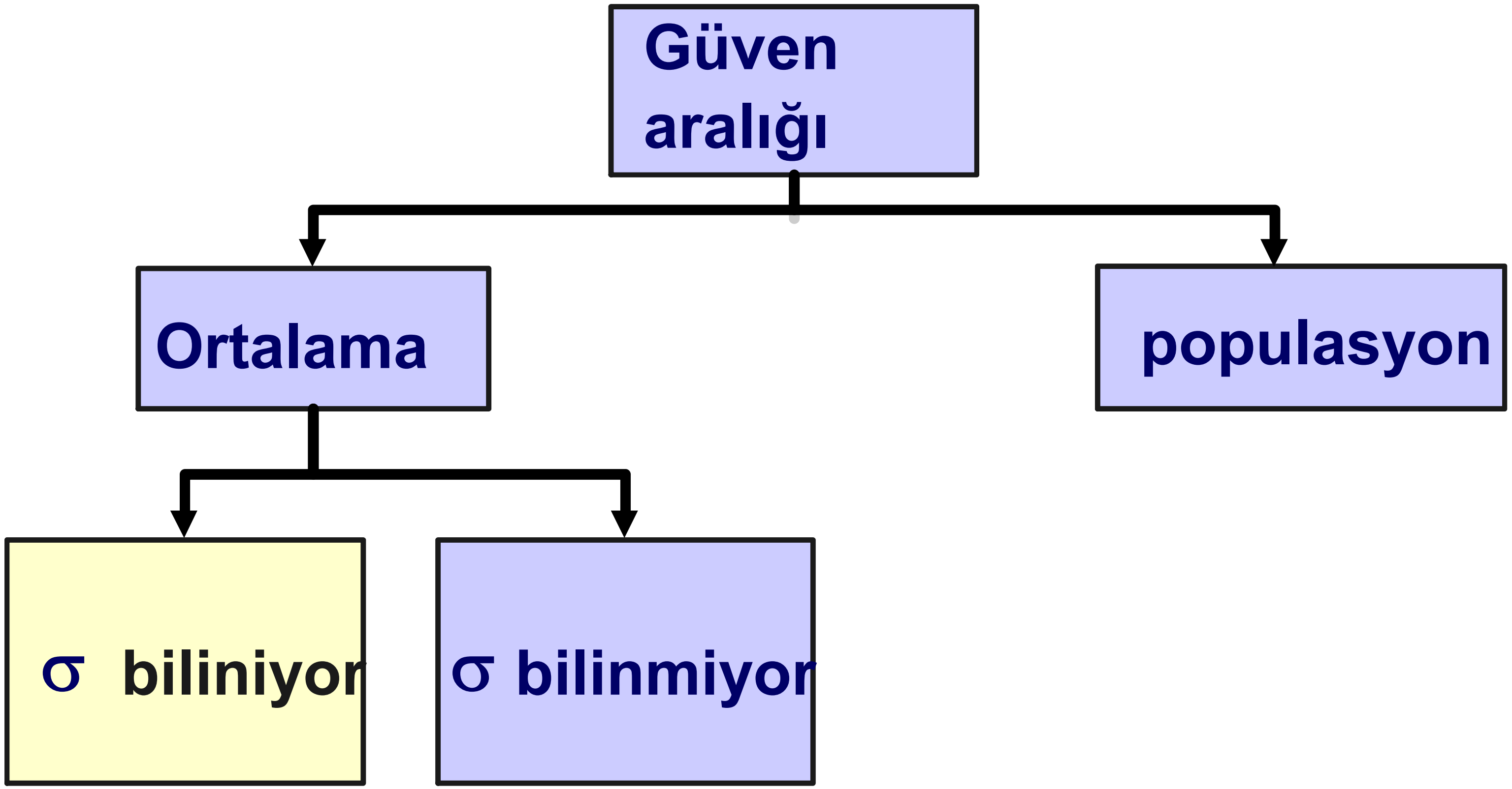
Nokta Tahmini

Tahmin edilen populasyon parametresi		Örnek istatistiği
Ortalama	μ	$\bar{X}$
Oran	p	P_S
Varyans	σ^2	S^2
Fark	$\mu_1 - \mu_2$	$\bar{X}_1 - \bar{X}_2$

Aralık Tahmini

- Bir değişkene ait değeri belli aralıklarda tahmin etme işi
  - Farklı örnekler için dağılımları dikkate al
  - Tek bir örnek için hesaplanan istatistikleri kullan
  - Parametrelere ne kadar yakın olduğunu belirt
  - Tahminin belli bir güven düzeyi ve hata payı için olduğunu belirt. Hiçbir zaman %100 emin olma

Tek Populasyona ait Aralık Tahminleri

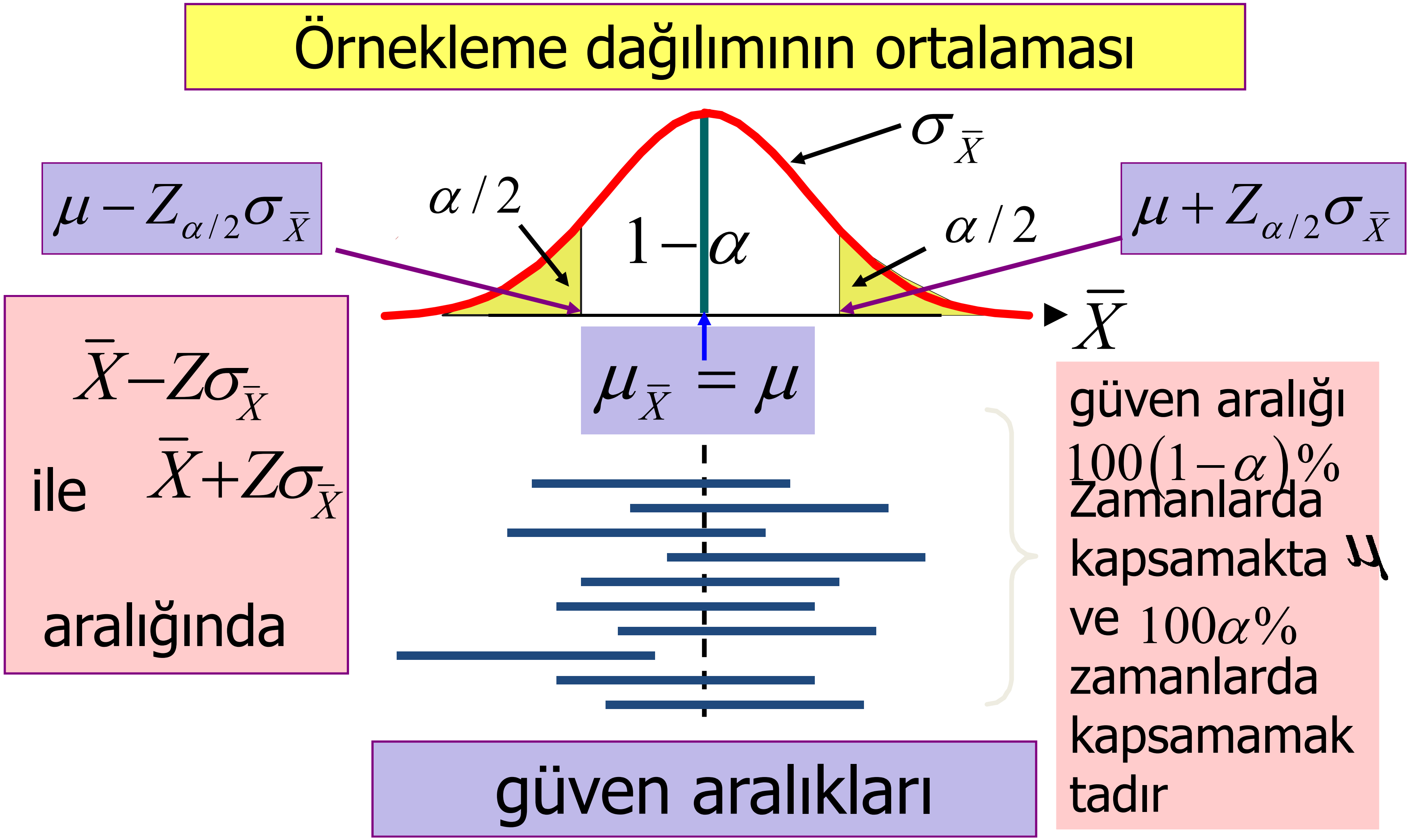


μ Nün tahmini
(σ biliniyor)

- Varsayımlar
  - μ biliniyor
  - Populasyon normal dağılmıştır
  - Populasyon normal dağılıma sahip değilse örnek büyük olmalı
- Güven aralığı: (1-α) olasılıkla aşağıdaki aralıktadır

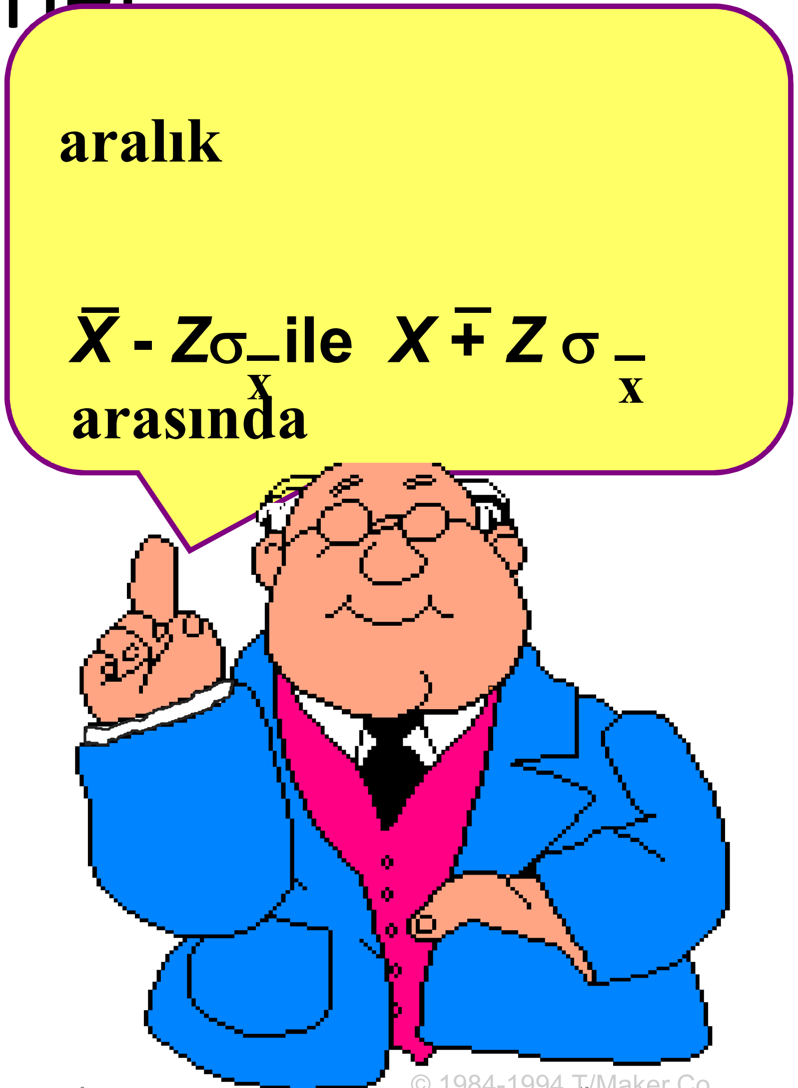
$$\bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Aralık ve Güven Düzeyi



Güven Aralığının Genişliğini Etkileyen Faktörler

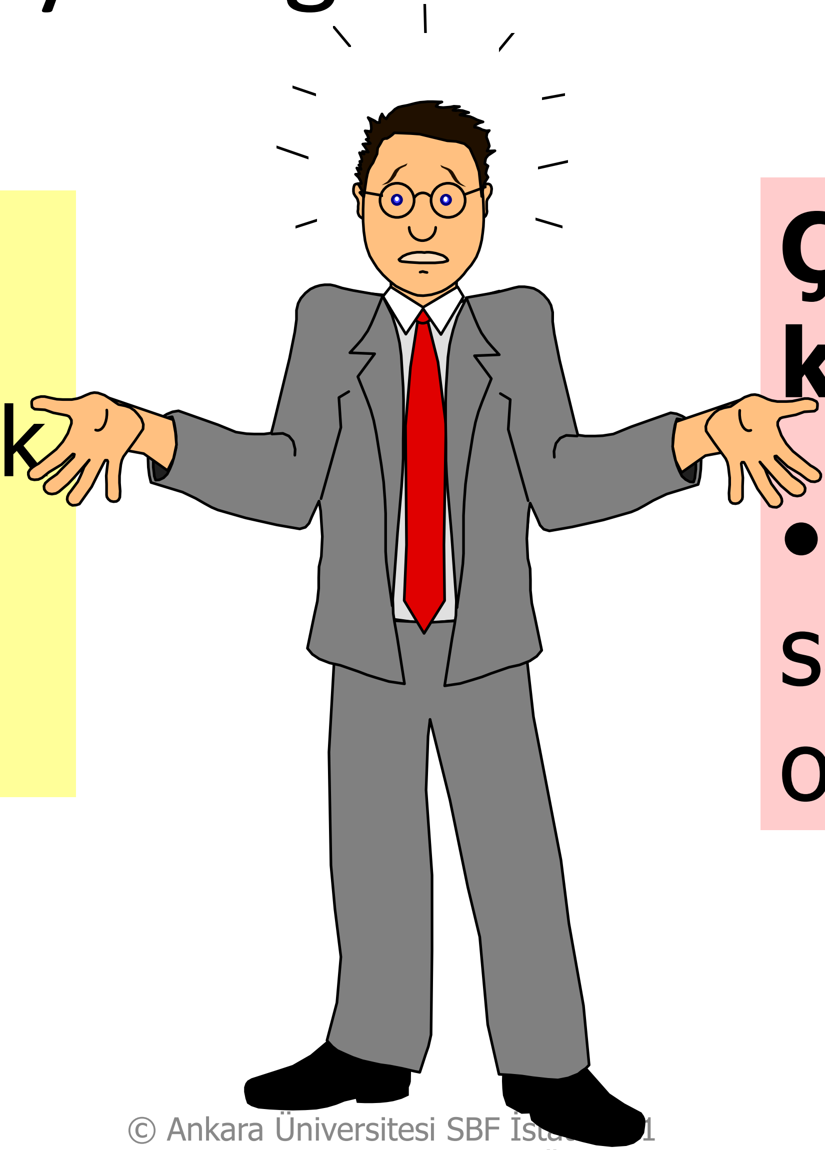
- Verilerin yayılımı
  - σ ile ölçümlenmekte
- Örnek büyüklüğü
  - $$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$
- Güven düzeyi
  - $100(1-\alpha)\%$



Örnek Büyüklüğünün Belirlenmesi

Çok büyük:

- fazla kaynak kullanılması gerekir



Çok küçük:

- tahmin sağlıklı olmayabilir

Ortalama İçin Örnek Sayısının Tespiti

Oluşturulan güven aralığının %90 güven düzeyinde ± 5 hata payı ile doğru olabilmesi için örnek büyüklüğü ne olmalıdır? Not: Standart sapma 45.

$$n = \frac{Z^2\sigma^2}{\text{Error}^2} = \frac{1.645^2\left(45^2\right)}{5^2} = 219.2 \cong 220$$

$$\mu - Z_{\alpha/2}\sigma_{\bar{X}}$$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

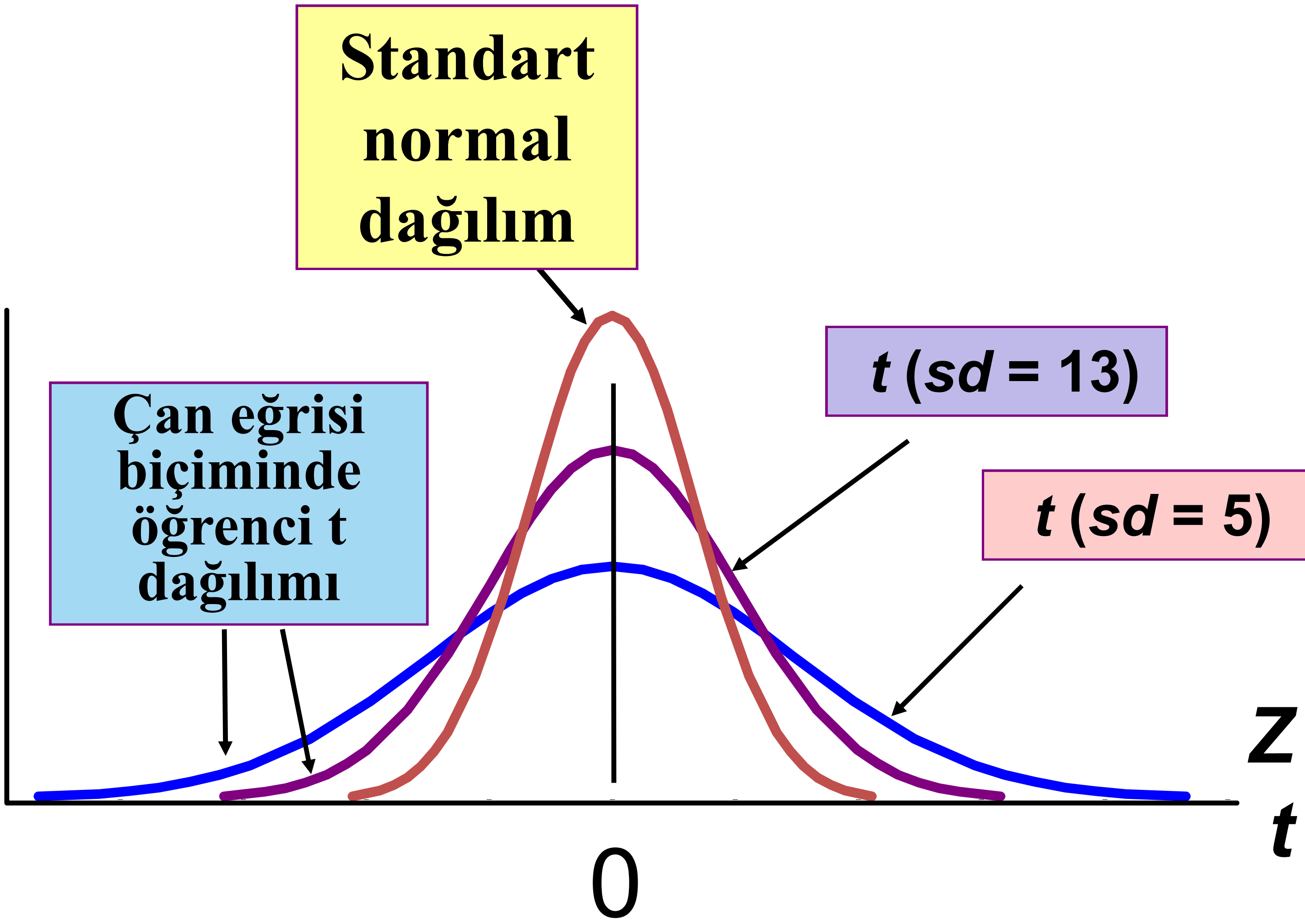
μ İçin Güven Aralığı (σ bilinmiyor)

1. Populasyon varyansı bilinmemektedir.
2. populasyon normal dağılıma sahiptir.
3. örnek sayısı n<30 olmalıdır.
4. Populasyon normal dağılıma sahip değilse büyük ölçekli örnek kullanılmalıdır.

t- dağılımı için güven aralığı şöyle oluşur:

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

Öğrenci t dağılımı



Serbestlik derecesi (sd)

– Sd örnek ortalaması hesaplandıktan sonra dağılma serbestisine sahip gözlem sayısıdır

– örnek

- n-5 ise sd= 4
- 1, 2, 3, 4 nolu gözlemler
- Dağılım gösterebilir
- 5 nolu gözlem dağılamaz

sd

$$= n - 1$$
$$= 5 - 1$$
$$= 4$$

t tablosu

	Upper Tail Area		
df	.25	.10	**.05**
1	1.000	3.078	6.314
2	0.817	1.886	**2.920**
3	0.765	1.638	2.353

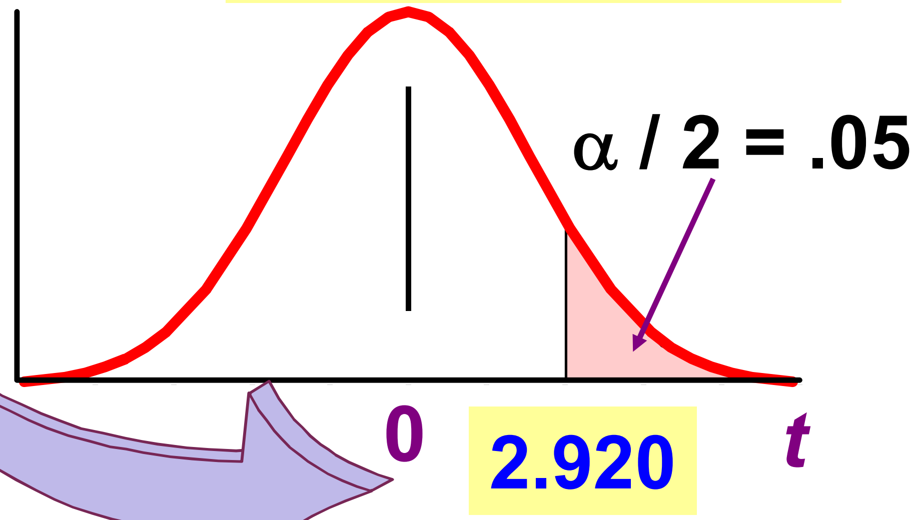
t değerleri

$n = 3$

$sd = n - 1 = 2$

$\alpha = .10$

$\alpha/2 = .05$



Örnek

Ortalaması 50 standart sapması 8 olan ve n=25 gözlemden oluşan bir örnek veri setine sahip olunsun. Populasyon için %95'lik güven aralığı oluşturunuz

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

$$50 - 2.0639\frac{8}{\sqrt{25}} \le \mu \le 50 + 2.0639\frac{8}{\sqrt{25}}$$

$$46.69 \le \mu \le 53.30$$

Popülasyon Oranı İçin Güven Aralığı

– Varsayımlar:

1. Populasyon binom olasılık dağılımına sahiptir.
2. np ≥ 5 ve n(1-p) ≥ 5 olması durumunda populasyon düzeltme faktörü kullanılır.

– Güven aralığı:

$$1-\alpha = P(\overline{p} - z_{\alpha/2}\sqrt{\frac{\overline{p}(1-\overline{p})}{n}} < p < p + z_{\alpha/2}\sqrt{\frac{\overline{p}(1-\overline{p})}{n}})$$

Örnek

n=400 gözlemden oluşan bir örnek veri setine için yapılan bir anket sonucunda 32 kişinin öğrenci temsilciliği için Gözde'yi tercih ettikleri saptanmıştır. P için %95'lik güven aralığı oluşturunuz

$$p_s - Z_{\alpha/2}\sqrt{\frac{p_s\left(1-p_s\right)}{n}} \le p \le p_s + Z_{\alpha/2}\sqrt{\frac{p_s\left(1-p_s\right)}{n}}$$

$$.08 - 1.96\sqrt{\frac{.08\left(1-.08\right)}{400}} \le p \le .08 + 1.96\sqrt{\frac{.08\left(1-.08\right)}{400}}$$

$$.053 \le p \le .107$$

Örnek Oranı İçin Oluşturulan Güven Aralığı İçin Örnek Sayısının Belirlenmesi

1,000, üyeden oluşan bir populasyondan rassal yöntemle 100 gözlem tespit edildi. 30 adet defolu ürün bulunmakta. Oluşturulan güven aralığının %90 güven düzeyinde ± 5 hata payı ile doğru olabilmesi için örnek büyüklüğü ne olmalıdır?

$$n = \frac{Z^2 p(1-p)}{\text{Error}^2} = \frac{1.645^2(0.3)(0.7)}{0.05^2}$$

$$= 227.3 \cong 228$$

Varyans İçin Güven Aralığı

- Varyansı σ^2 olan normal dağılıma sahip bir populasyondan elde edilmiş olan örneklerden oluşturulan rassal değişken ($\chi^2 = (n-1)\, S^2/\sigma^2$) (n-1) serbestlik derecesi ile ki kare dağılımına sahiptir. $(n-1)\, S^2/\sigma^2$ değişkeni ki kare istatistiği olarak isimlendirilmekte ve $\chi 2$ (ki) harfi ile gösterilmektedir.
- k inci serbestlik derecesinden, ki kare dağılımına sahip bir rassal değişken (k= n – 1) şeklinde gösterilir. Belirli bir olasılık (örneğin α) için ki kare olasılık dağılım tablosundan bu olasılığa karşılık gelen değer bulunur. Bu olasılık , biçiminde gösterilmektedir.

$$\chi^2_{k,\alpha}$$

Populasyon Varyanslarının Oranları İçin Güven Aralığı Oluşturulması

- Varyansları σ_1 ve σ_2 olan iki adet normal dağılıma sahip populasyondan rassal örnekleme yöntemi ile n_1 ve n_2 sayıdan oluşan gözlemler elde edilmiş ve elde edilen bu örnekler için varyanslar S_1^2 ve S_2^2 olarak hesaplanmış olsun bu durumda rassal değişken F, şöyle bir F dağılımına sahiptir

$$\frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2}$$

Populasyon Varyanslarının Oranları İçin Güven Aralığı Oluşturulması

- Varyansların oranları için güven aralığı şu formülle oluşturulur

$$\left(\left(\frac{S_1^2}{S_2^2}\right)\frac{1}{F_{\alpha/2,V1,V2}} < \frac{\sigma_1^2}{\sigma_2^2} < \left(\frac{S_1^2}{S_2^2}\right)F_{\alpha/2,V2,V1} \right) = 1-\alpha$$

Normal Dağılıma Sahip İki Populasyonun Ortalamalarının Farkları İçin Güven Aralığı Oluşturulması

- Uygun Örnek Çiftlerine İlişkin Güven Aralığı Oluşturulması:
- Ortalamaları μ_1 ve μ_2 ve varyansları $\sigma 1$ ve $\sigma 2$ olan normal dağılıma sahip iki populasyonun her birinden rassal yöntemle n (eşit) sayıda gözlemden oluşan iki ayrı örnek seti elde edilmiş olsun. Oluşturulan bu iki örnek setinde yer alan gözlemlerin birbirleri ile uyumlu olmaları durumunda eşleştirilebilmeleri ve farkları için ortalamaları hesaplanabilmektedir.

Normal Dağılıma Sahip İki Populasyonun Ortalamalarının Farkları İçin Güven Aralığı Oluşturulması

- Güven aralığı şu formülle belirlenir

$$\bar{d} - \frac{t_{n-1,\alpha/2} S_d}{\sqrt{n}} < \mu_1 - \mu_2 < \bar{d} + \frac{t_{n-1,\alpha/2}\, S_d}{\sqrt{n}}$$

Normal Dağılıma Sahip İki Bağımsız Populasyonun Ortalamalarının Farkları İçin Güven Aralığı Oluşturulması

- Güven aralığı: populasyon varyansı biliniyor

$$(\bar{x}_1 - \bar{x}_2) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Normal Dağılıma Sahip İki Bağımsız Populasyonun Ortalamalarının Farkları İçin Güven Aralığı Oluşturulması

- Güven aralığı: örnek varyansı biliniyor

$$(\bar{x}_1 - \bar{x}_2) - z_{\alpha/2}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + z_{\alpha/2}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

Normal Dağılıma Sahip İki Bağımsız Populasyonun Ortalamalarının Farkları İçin Güven Aralığı Oluşturulması

- Güven aralığı: populasyon varyansı bilinmiyor ve n<30

$$(\bar{x}_1-\bar{x}_2)-t_{n_1+n_2-2,\alpha/2}S\sqrt{\frac{n_1+n_2}{(n_1)(n_2)}}<\mu_1-\mu_2<(\bar{x}_1-\bar{x}_2)+t_{n_1+n_2-2,\alpha/2}S\sqrt{\frac{n_1+n_2}{(n_1)(n_2)}}$$

Dikkat Edilecek Noktalar

- Nokta tahmini ile birlikte aralık tahmini de belirtilmeli
- Güven düzeyi belirtilmeli
- Örnek sayısı belirtilmeli
- Sonuçla ilgili yorum yapılmalı