**Marmara Üniversitesi**

**İstatistik Bölümü**

**Örnekleme Teorisi (9)**

**Doç. Dr. Atıf Evren**

**Örnekleme Yöntemleri**

## Basit Rassal Örnekleme

**Tanım:** **N** birimlik bir anakütleden **n** birimlik bir örneklem çekilmesi gerektiğini düşünelim. Basit rassal örnekleme süreci, n gözlemlik her örneklemin seçilme şansının eşit olduğu süreçtir.

## Anakütle Ortalamasının Tahmin Edilmesi

Ortalaması μ olan N birimli biir anakütleden çekilmiş n gözlemlik bir basit rassal örneklemin gözlenen değerleri $x_1, x_2, \dots, x_n$ olsun.

i) Örneklem ortalaması, anakütle ortalaması μ'nün yansız bir tahmin edicisidir. Nokta tahmini şöyledir:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

ii) Örneklem ortalamasının varyansını yansız tahmin eden süreç şu nokta tahminini verir:

$$\widehat{\sigma_{\bar{x}}^2} = \frac{s^2}{n}\frac{N-n}{N}$$

Burada $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$ örneklem varyansıdır.

iii) Örneklemin büyük olması koşuluyla anakütle ortalamasının %100(1-α) güven aralığı şöyle bulunur:

$\bar{x} - z_{\alpha/2}\widehat{\sigma_{\bar{x}}} < \mu < \bar{x} + z_{\alpha/2}\widehat{\sigma_{\bar{x}}}$ Burada $Z_{\alpha/2}$ , $P\left(Z > z_{\frac{\alpha}{2}}\right) = \frac{\alpha}{2}$ eşitliğini sağlayan bir sayı, Z de standart normal dağılıma uyan rassal bir değişkendir.

**Örnek**: Belli bir kentte geçen yıl 1118 ipotek işlemi yapılmıştır. Bunların 60 gözlemlik bir rassal örnekleminin ortalaması 87300 pb., standart sapması 19200 pb. dir. Bu kentte geçen yıl yapılan ipotek işlemlerinin ortalama değerini tahmin edip %95 güven aralığını bulun.

**Çözüm:** N=1,118 , n=60 $\bar{x} = 87300$ s=19200

Anakütle ortalamasının nokta tahmini $\bar{x} = 87300$ pb. dir.

Aralık tahminini elde etmek için

$$\widehat{\sigma^2_{\bar{x}}} = \frac{s^2}{n}\frac{N-n}{N} = \frac{19200^2}{60}\frac{1058}{1118} = 5814{,}268$$

$\widehat{\sigma_{\bar{x}}} = 2411\, pb.$

Bu kentte geçen yılki bütün ipotek işlemlerinin ortalama büyüklüğünün %95 güven aralığı aşağıdaki gibidir:

$\bar{x} - z_{\alpha/2}\widehat{\sigma_{\bar{x}}} < \mu < \bar{x} + z_{\alpha/2}\widehat{\sigma_{\bar{x}}}$

87300-1,96*(2411)<μ<87300+1,96*(2411)

82574<μ $< 92026$

**Anakütle ToplamınınTahmin Edilmesi**

N birimlik bir anakütleden n gözlemlik basit rassal bir örnek aldığımızı, tahmin edilecek büyüklüğün anakütle toplamı Nμ olduğunu düşünelim. Anakütle toplamı Nμ için yansız bir tahmin süreci $N\bar{x}$ nokta tahminini verir.

i) Anakütle toplamı tahmin edicisinin varyansı

$$N^2\widehat{\sigma_{\bar{x}}^2}=\frac{s^2}{n}(N-n)$$

ii) Örneklem büyük olmak koşulu ile anakütle toplamının %100(1-α) güven aralığı şöyle bulunur:

$$N\bar{x}-z_{\alpha/2}N\widehat{\sigma_{\bar{x}}}<N\mu<N\bar{x}+z_{\alpha/2}N\widehat{\sigma_{\bar{x}}}$$

**Örnek:** ABD'de 1395 fakülte arasından seçilen 400 gözlemlik bir basit rassal örneklemden, geçen yıl işletme istatistiği dersine kayıtlı öğrenci sayılarının ortalaması 320,3; standart sapması da 149,7 öğrencidir. O yıl işletme istatistiği dersine kayıtlı toplam öğrenci sayısını tahmin edip %99'luk güven aralığını bulun.

**Çözüm:** N=1395, n=400, $\bar{x}=320{,}3$ s=149,7

Toplam için nokta tahmini $N\bar{x}$ =1395*320,8=447516

Toplamın varyansı ise

$$N^2\widehat{\sigma_{\bar{x}}^2}=\frac{s^2}{n}(N-n)=\frac{149{,}7^2}{400}*(1395-400)=77764413$$

Standart hata: $N\widehat{\sigma_{\bar{x}}}=\sqrt{77764413}$ =8818,4

%99 güven aralığı

$N\bar{x} - z_{\alpha/2} N\widehat{\sigma_{\bar{x}}} < N\mu < N\bar{x} + z_{\alpha/2} N\widehat{\sigma_{\bar{x}}}$

$447516 - 2{,}575 * (8818{,}4) < N\mu < 447516 + {,}575 * (8818{,}4)$

$424809 < N\mu < 470223$

**Anakütle Oranının Tahmin Edilmesi**

Belirli bir özelliği taşıyan bireylerinin oranı p olan bir anakütleden çekilmiş n gözlemlik rassal bir örneklemde sözkonusu özelliği taşıyan birimlerin oranına $\hat{p}$ diyelim.

i) Örneklem oranı ($\hat{p}$) anakütle oranı p'nin yansız bir tahmin edicisidir.

ii) Anakütle oranı tahmincisinin varyansı

$$\widehat{\sigma_{\hat{p}}^2} = \frac{\hat{p}(1-\hat{p})}{n-1} \frac{N-n}{N}$$

iii) Örneklemin büyük olması koşuluyla anakütle oranının %100(1-α)'lık güven aralığı tahmini

$\hat{p} - z_{\frac{\alpha}{2}} \widehat{\sigma_{\hat{p}}} < p < \hat{p} + z_{\frac{\alpha}{2}} \widehat{\sigma_{\hat{p}}}$ ile bulunur.

**Örnek:** ABD'de 1395 fakülte arasından seçilen 400 gözlemlik bir basit örneklemden 141 tanesinde işletme istatistiği dersinin iki dönem okutulduğu anlaşılmıştır. Bu dersi iki dönem boyunca okutan bütün fakültelerin oranını tahmin edip %90'lık güven aralığını oluşturunuz.

**Çözüm:** N=1395 n=400 $\hat{p} = \frac{141}{400} = 0{,}3525$

$$\widehat{\sigma_{\hat{p}}^2} = \frac{\hat{p}(1-\hat{p})}{n-1}\frac{N-n}{N} = \frac{0{,}3525*0{,}6475}{399}\frac{995}{1395} = 0{,}0004079$$

$$\hat{\sigma}_{\hat{p}} = 0{,}0202$$

%90'lık güven aralığı

$$\hat{p} - z_{\frac{\alpha}{2}}\widehat{\sigma_{\hat{p}}} < p < \hat{p} + z_{\frac{\alpha}{2}}\widehat{\sigma_{\hat{p}}}$$

$$0{,}3525 - 1{,}645 * 0{,}0202 < p < 0{,}3525 + 1{,}645 * 0{,}0202$$

0,3193<p<0,3857

**Örneklem Büyüklüğünün Belirlenmesi**

**1)Basit Rassal Örneklemede Aracılığıyla Anakütle Ortalaması ya da Toplamını Tahmin Ederken Örneklem Büyüklüğünün Belirlenmesi**

Anakütle ortalamasını n gözlemlik basit rassal örnekleme ile tahmin edelim. Burada örneklem ortalamasının varyansı daha önceden tartışıldığı gibi

$Var(\bar{X}) = \frac{\sigma^2}{n}\frac{N-n}{N-1}$ idi. Buradan n çekilecek olursa

$$n = \frac{N\sigma^2}{(N-1)\sigma_{\bar{X}}{}^2 + \sigma^2}$$

**Örnek:** Bir kentte geçen yıl 1118 ipotek işlemi yapıldığını ve bunların ortalama tutarını tahmin edebilmek için basit rassal bir örnek alındığını düşünelim. Geçmiş deneylerden hareketle anakütle standart sapmasının yaklaşık 20000 pb. olduğu bilinmektedir. Anakütle ortalamasının %95 güven aralığı örneklem ortalamasının her iki yanında 4000 pb. kadar olmalıdır. Bu amaca ulaşabilmek için örneklemde kaç gözlem olmaldır?

**Çözüm**: N=1118 , $\sigma = 20000$ , 1,96$\sigma_{\bar{X}}$=4000 $\Rightarrow$ $\sigma_{\bar{X}}$=2041

Bu durumda aranan örneklem büyüklüğü

$$n = \frac{N\sigma^2}{(N-1)\sigma_{\bar{X}}{}^2 + \sigma^2} = \frac{1118 * 20000^2}{(1117 * 2041^2 + 20000^2)} = 88{,}5$$

Demek ki 89 gözlemli bir basit rassal örneklem amacımızı karşılamaya yetmektedir.

**2) Basit Rassal Örneklemede Aracılığıyla Anakütle Oranını Tahmin Ederken Örneklem Büyüklüğünün Belirlenmesi**

$Var(\hat{p}) = \frac{p(1-p)}{n}\frac{N-n}{N-1}$ idi. Buradan n çekilecek olursa

$n = \frac{Np(1-p)}{(N-1)\sigma_{\hat{p}}^2 + p(1-p)}$ olur.

P'nin değeri ne olursa olsun örneklem hacminin en büyük değeri p=1/2 için bulunur. Bu değer yukarıdaki formülde yerine konacak olursa

$n_{maks} = \frac{0{,}25N}{(N-1)\sigma_{\hat{p}}^2 + 0{,}25}$ elde edilir. Anakütle oranının %95'lik güven aralığı örneklem oranının her iki yanında yaklaşık 1,96*$\sigma_{\hat{p}}$ kadar uzanır.

**Örnek:** İşletme istatistiği dersinin iki dönem okutulduğu okulların oranını tahmin etmek için 1395 ABD fakültesinden basit bir rassal örnek alınsın. Gerçek oran ne olursa olsun, bunun %95'lik güven aralığının, örneklem oranının iki yanında 0,04'ten daha fazla uzanmaması için örneklem içindeki gözlem sayısı ne olmalıdır?

**Çözüm:** 1,96*$\sigma_{\hat{p}} = 0{,}04 \Longrightarrow \sigma_{\hat{p}} = 0{,}0204$

$$n_{maks} = \frac{0{,}25N}{(N-1)\sigma_{\hat{p}}^2 + 0{,}25} = \frac{0{,}25 * 1395}{(1394 * 0{,}0204^2 + 0{,}25)} = 420{,}1$$

421 gözlemli bir basit rassal örneklem yeterlidir.