ÖRNEKLEME TEORİSİ

- Bir popülasyonu istatistiksel açıdan incelemek ve işlemler yapabilmek için popülasyon içerisinden seçilen örneklemlerden yararlandığımızı söylemiştik.
- Peki popülasyonun istatistiksel parametrelerini örneklemlerle belirlemenin nedenleri neler olabilir?

....

- Popülasyonu tümü üzerinde çoğu zaman işlem yapmanın imkansız ve/veya oldukça maliyetli olması
- Örneklem ile çalışmanın vakit ve maliyet açısından tasarruflu olması
- İyi seçilmiş bir örneklemin popülasyonu en iyi şekilde temsil edebilmesi

Örnekleme Çeşitleri

ÖRNEKLEME

Rasgele	*Karara dayalı (iradi)*
**basit*	
**sistematik*	
**cluster*	

1. RASGELE ÖRNEKLEME :
   ***Basit Rasgele Örnekleme:** Popülasyondaki her bir üyenin seçilme olasılığı birbirine eşittir.*
   ***Sistematik Rasgele Örnekleme:** Rastgele belirlenen başlangıç noktasından itibaren her n. üye örneklem içerisine dahil edilir.*
   ***Cluster Örnekleme:** Gruplara ayrılan popülasyondan rasgele örneklem ayarlanır.*
2. KARARA DAYALI ÖRNEKLEME:örneklemeyi yapan kişi kendi isteğine göre üye seçimi yapar. Bu tip örneklem popülasyonu iyi yansıtmayacağı için hata payı büyüktür.

ÖRNEKLEME HATASI

Popülasyonun gerçek parametresi ile örnekleme istatistiği arasındaki farka **örnekleme hatası** denir. Ortalamadaki standart hata olarak da bilinir. Buna göre;

s:örneklemedeki gözlemlerin standart sapması

n:örneklemedeki gözlem sayısı

Bir örneklemede tahmin edilen ortalamadaki standart hata ➡ $\sigma_{\ddot{x}} = \frac{s}{\sqrt{n}}$

- **MERKEZİ LİMİT TEOREMİ:** Dağılım ne olursa olsun ve dağılımın bilinmediği durumlarda da örneklem hacmi (n) yeteri kadar büyük olduğunda örneklemi normal dağılıma çevirebilen bir teoremdir. Buna göre;
  - Popülasyon normal dağılıma sahip ise örneklemlerin aritmetik ortalamaları da normal dağılım gösterir.
  - Popülasyon normal dağılıma sahip değil iken örneklemlerin aritmetik ortalamaları normale yaklaşan bir dağılım sergiler. Örnek sayısı arttıkça normale daha çok yakınsanır.

Popülasyon parametreleri iki şekilde tahmin edilebilir:

- ***NOKTA TAHMİNİ (point estimate):*** Tek bir değer kullanılarak parametre tahmin edilir.
- ***ARALIK TAHMİNİ (invertal estimate):*** Popülasyon parametresi için bir aralık tespiti yapılır. Bu aralık tespiti yapılırken popülasyon için yapılan bir hesaplama ile güven aralığı bulunur.

GÜVEN ARALIĞI

n:gözlem sayısı

s:örneklemenin standart sapması

z:standart değer

$$\left(\bar{x} - z\frac{s}{\sqrt{n}}, \bar{x} + z\frac{s}{\sqrt{n}}\right)$$

n≥30 için;

- %95 güven aralığında z=1,96
- %99 güven aralığında z=2,58 alınır.

❖ 1,96 ve 2,58 değerleri gözlemlerin sırasıyla %95 ve %99 una karşılık gelen

standard değerlerdir. Bu güven aralıklarına karşılık gelen değerler standart normal

dağılım tablolarından hesaplanır.

❖ Örneğin yarım normal dağılıma göre 0,95/2 =0,475 . Bu değere karşılık gelen standart değer normal dağılım tablosundan 1,96 olarak kolaylıkla okunabilir.

- Bir örneklemede popülasyonun ortalaması için gerekli olan gözlem sayısı yani n;

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

E:izin verilebilir max. hata oranı

z:seçilen güven aralığı

s:verilerin standart sapması

- Eğer örneklemedeki veri sayısı (n) popülasyonun (N) %5inden büyük ise hem popülasyon ortalaması hem de standart hataya bir düzeltme katsayısı uygulanmalıdır. Bu katsayı

(N-n) / (N-1)

O halde;

- n/N>0.05 iken popülasyon ortalamasının standart hatası:

$$\sigma_{\overline{X}} = \frac{s}{\sqrt{n}}\left(\frac{N-n}{N-1}\right)$$

- Ortalamanın güven aralığı:

$$\overline{X} \pm z\frac{s}{\sqrt{n}}\left(\frac{N-n}{N-1}\right)$$

……

- Bu durumda (n/N>0.05 iken) düzeltme katsayısı standart hatayı azalttığı için popülasyon ortalamasının aralığı daralır. Yani örneklem sayısı arttıkça ortalamanın standart hatası azalır.