İSTATİSTİK

Doç. Dr. Suat ŞAHİNLER

Tavsiye Edilen Kaynak Kitaplar

Her öğrencinin kendisinin tuttuğu düzenli notlar.

1. Akar, M. ve S. Şahinler, (1997). İstatistik. Ç.Ü. Ziraat Fakültesi Genel Yayın No: 74, Ders Kitapları Yayın No: 17, ADANA.
2. İkiz, F., H. Püskülcü, ve Ş. Eren, (1996). İstatistiğe Giriş. E.Ü. Basımevi, İZMİR.
3. Serper, Ö., (2000). Uygulamalı İstatistik. Ezgi Kitabevi, BURSA.

Bölüm 1

TEMEL KAVRAMLAR

İstatistik, değişik anlamlarda kullanılan ve farklı tanımları olan bir kavramdır.

Tanım1: Belirli konularda toplanan sayısal değerler ile ileri sürülen bir takım şekil, grafik ve tablolardır.

Örnek:

Bir ülkede bir yılda meydana gelen ölüm sayısı, kaza sayısı, ilkokula başlayan ve mezun olan öğrenci sayısı, Türkiye'de yıllar itibari ile tarım ve sanayi kesiminde çalışanların sayısı, okuma yazma oranı ile ihracat ve ithalattaki durumu gösteren bir takım şekil ve grafikler

Tanım2: Temelini matematikten alan bilim dalına verilen addır.

Bu yönden ele alındığında istatistik, pozitif bilimin esası olan deney veya denemeler planlama, gözlem yapma, verileri toplama ve toplanan verileri düzenleme, analiz etme, yorumlama , objektif ve doğru kararı verme ile ilgili bilimsel teknik ve metotlar geliştiren ve uygulayan bir bilim dalıdır. Geliştirilen bir istatistik teknik ve yöntem, ziraatte olduğu kadar ekonomi, tıp, eğitim veya biyolojide de uygulanabilir. Bu nedenle tüm araştırıcı veya bilim adamları istatistik teknik ve yöntemlerini en azından yakından tanımak ve belirli ölçüde bilmek zorundadır.

Tanım3: Şansa bağlı olarak elde edilen ve örneklerden hesaplanan değerlerin genel adıdır. Diğer bir ifadeyle, bir değer örnekten hesaplanıyorsa bu değere istatistik denir.

İstatistikte Kullanılan Temel Kavramlar

BİRİM(Ünite): Bir topluluğu oluşturan ve incelemeye konu olan obje yada bireye birim denir.

Nüfus olayında birim kişi, devlette çalışanlarda memur veya işçi, bir tarla denemesinde ya parsel ya da bir tek bitkinin kendisi, hastane olayında hasta ve üniversıte olayında ise ya öğretim elamanı yada öğrencidir.

KARAKTER(Vasıf): Birimin çeşitli özellikleri karakter olarak tanımlanabilir.

Örneğin, nüfus olayında kişinin cinsiyeti, yaşı, boyu, öğrenim durumu vb., tarımda buğdayın çeşidi, başaktaki dane sayısı, dekara verimi, bitkinin boyu, yaprak genişliği birer karakterdirler.

İstatistik açıdan tüm karakterler genel olarak ikiye ayrılır, bunlar;

a)Kalitatif (nitel) karakterler
b)Kantitatif (nicel) karakterler

POPULASYON(yığın, kitle): Belli karakterleri ortak olan birimlerin oluşturduğu topluluğa populasyon denir.

Örneğin: hayvan populasyonu, bitki populasyonu, öğrenci populasyonu

Populasyonu oluşturan birimlerin sayısı populasyon genişliği olarak bilinir (N) ve araştırıcının amacına bağlı olarak küçülebilir veya büyüyebilir.

Populasyonda üzerinde çalışılan karakter veya karakterler bakımından tüm obje veya bireyleri teker teker incelemek çoğu zaman mümkün değildir. Populasyondaki bireylerin tümünün incelenmesini sınırlayan faktörler;

Zaman, Maliyet,

İşçilik, Yasalar

Bundan dolayı, populasyonun tümünün üzerinde çalışılması yerine ondan belli yöntemlerle alınan örnek veya örnekler üzerinde çalışılır.

PARAMETRE: Populasyonun özelliklerini tanımlayan değerlere parametre denir. Populasyonun özelliklerini belirten parametrelerden en önemlileri populasyonun ortalaması (μ) ve populasyonun varyansı (σ^2)'dır.

ÖRNEK: İnceleme konusu olan populasyondan bir örnekleme yöntemiyle populasyonu temsil edebilecek büyüklükte alınan daha az sayıda birimlerin oluşturduğu topluluğa örnek adı verilir. Ayrıca örnek, araştırıcının çalıştığı konu ile ilgili olarak deney veya gözlemler sonucunda elde ettiği gözlemler topluluğudur şeklinde de tanımlanabilir.

Örneğin; bir koyun sürüsünden rastgele alınan 50 koyun, bir portakal bahçesinden rastgele seçilen 15 portakal ağacı birer örnektirler.

İSTATİSTİK: Örneğin özelliklerini tanımlayan değerlere istatistik adı verilir. Diğer bir ifadeyle, örnekten hesaplanan değerlere istatistik denir. En önemli İstatistikler örnek ortalaması($\bar{x}$) ve örnek varyansıdır (S^2). İstatistikler parametrelerin birer tahmini değeridirler. Yani $\bar{x}$, μ'nün, S^2 ise σ^2'nin tahmini değeridirler.

ÖRNEKLEME:Herhangi bir populasyondan örneği seçmek için uygulanan yöntemler topluluğuma örnekleme denir. Populasyonu temsil etmek amacıyla yapılan çeşitli örnekleme şekilleri vardır. Bunlardan en çok kullanılan ve en basit şans örneklemesi veya tesadüfi örneklemedir. Şans örneklemesinin esası , populasyonu oluşturan tüm birimlerin örneğe girme şanslarının eşit tutulmasıdır.

Populasyondan örnek alırken dikkat etmesi gereken hususlar.

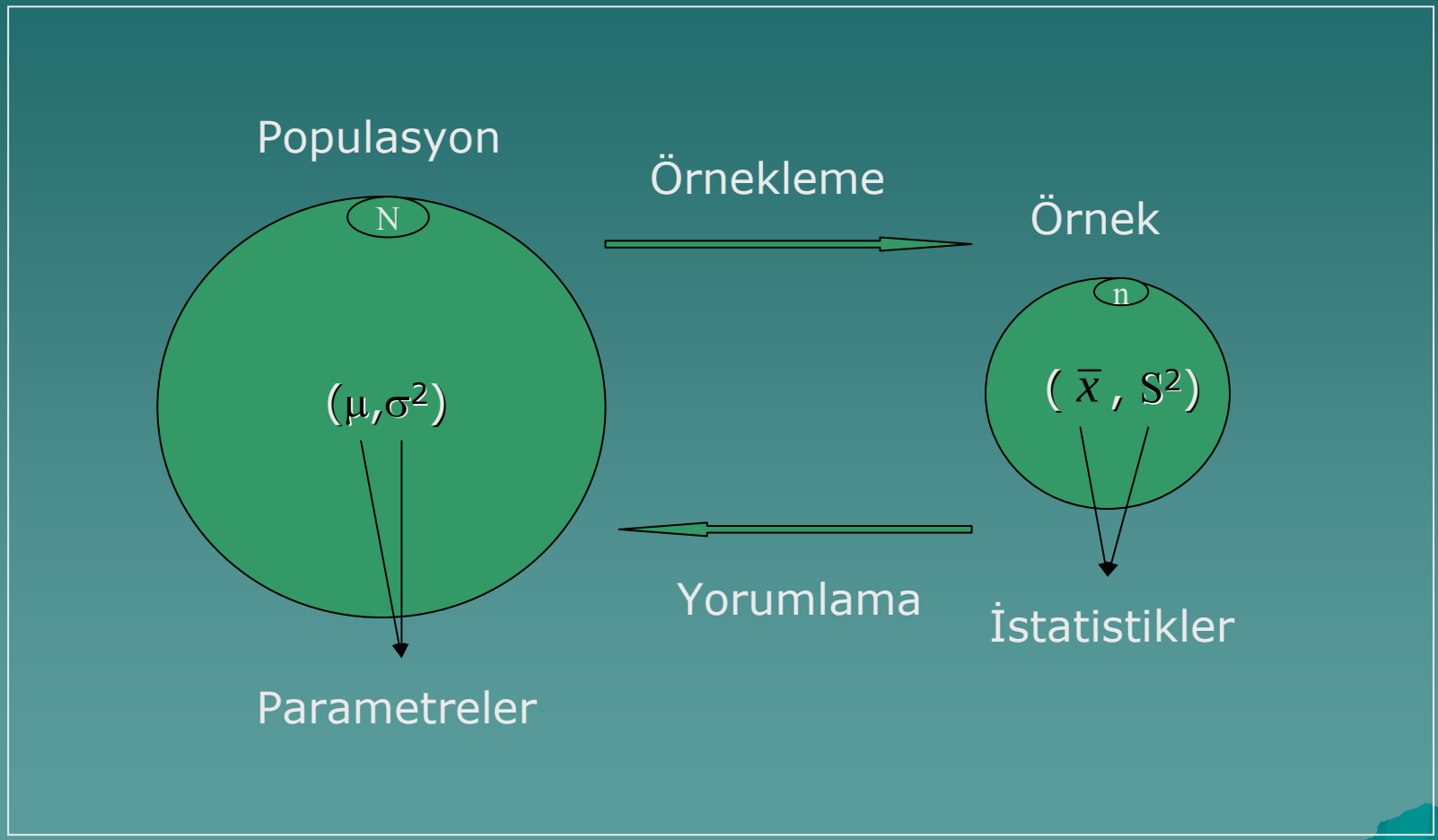
a) Birim sayısını geniş tutmak : Birim sayısı ne kadar geniş tutulursa veya örnekteki birim sayısı (n) populasyondaki birim sayısı (N) ne kadar yaklaşırsa , örneklerden hesaplanan istatistikler populasyon parametlerine o kadar yaklaşmış olur .Başka bir deyişle istatistiklerin güvenilirliği o oranda artar.Fakat bunu para, zaman ve eleman darlığı gibi etkenler sınırlar.

b) Materyal ve araştırmanın konusu hakkındabilgi sahibi olmak:

Araştırıcı üzerinde çalıştığı konu hakkında tam bir bilgiye sahip olmalıdır.Aksi taktirde,uygulayacağı istatistik yöntemi yanlış belirleyebilir.

c)Birimlerin örneğe girmesini rastgele yapmak: Araştırıcının örneği seçerken tamamıyla objektif ve yansız olabilmesi ancak örneğin seçimini rastgele yapmasıyla mümkündür.

Buraya kadar anlatılanlar Şekil 1.1'de özetlenmiştir.



Değişken: Tekrarlanan bir olayda her defasında farklı değerler alabilen sembollere değişken denir. Matematikte Değişken; denklem ve eşitsizliklerde belli koşulları sağlayan sayı ve olaylar topluluğunu gösteren sembollerdir.

Değişkenler genellikle alfabenin X,Y,V,W,Z gibi son harfleriyle gösterilir.Değişkenin aldığı değerler ise o değişkenin küçük harfi ile sembolize edilirler.Örneğin ,

$X=\{x_1,x_2,x_3,\ldots,x_n\}$

$Y=\{y_1,y_2,y_3,\ldots,y_n\}$

$Z=\{z_1,z_2,z_3\ldots\ldots,z_n\}$

$W=\{w_1,w_2,w_3\ldots,w_n\}$

Bir değişken iki tip değer alabilir;

Teorik (mümkün) Değer: Değişkenin alabileceği değerlerin tümüne "mümkün değer" denir.

X≡"Not" ise $X=\{0,1,2,\ldots,100\}$

Gözlenen Değer: Değişkenin deneme, deney veya gözlem sonunda almış olduğu değerlere denir.

X≡"5 öğrencinin notu" ise $X=\{77,65,44,99,15\}$

Veri Tipleri:

Değişkenin aldığı değerlerin bir ölçü birimiyle ifade edilmiş olan haline veri denir ve gözlemden gözleme farklılık gösterir. Esas olarak istatistikçi veya araştırıcı şans değişkenleri üzerinde çalışır. Araştırıcı herşeyden önce üzerinde çalıştığı şans değişkeninin tipini bilmek zorundadır. Çünkü uygulanacak istatistik yöntem değişkenin tipine göre farklılık gösterir. Genel olarak iki tip şans değişkeni vardır;

Kalitatif veriler; sınıflandırılabilen veriler olup her sınıfa düşen gözlem sayısı şeklinde gösterilirler.

Kantitatif veriler ise; kesikli ve sürekli olmak üzere kendi çerisinde ikiye ayrılır. Kesikli veriler, bir bireyin sayma sayıları ile ifade edilebilen özelliklerini, sürekli veriler ise, bir ölçüm veya tartım ile ölçülebilen özelliklerini yansıtan verilerdir.

Tablo 1. Değişken tipleri ve bazı özellikleri.

Değişken Tipleri	**Soru Tipi**		**Cevap**	**Dağılış Şekli**
Kalitatif	Parazit Var mı?		Evet Hayır	Ki-Kare(χ^2)
Kantitatif	Kesikli	Kaç Yumurta?	... adet	Binom, Poisson
	Sürekli	Yumurta ağırlığı ne kadar?	… gr	Normal

İstatistikte Kullanılan İşaretler:

" Σ " grek alfabesinde büyük harf sigmadır. Matematik ve istatistikte sayıların toplamı anlamına gelir.

$$\sum_{i=1}^{i=n} x_i \qquad \sum_{i=1}^{n} x_i \qquad \sum x_i \qquad \sum x$$

X ={x_1,x_2,x_3,...,x_n} gibi n tane değer alıyorsa bunların toplamı,

$$\sum_{i=1}^{n} x_i = x_1+x_2+x_3+\ldots+x_n$$

şeklinde gösterilir.

Örneğin; X={5,7,3,2,8) ise, n=5

$$\sum_{i=1}^{5} x_i = x_1 + x_2 + x_3 + x_4 + x_5 = 5+7+3+2+8 = 25$$

Olur.

$$\sum_{i=2}^{4} x_i = x_2 + x_3 + x_4 = 7 + 3 + 2 = 12$$

Bir değişkenin kareler toplamının bulunması; Bu, değişkenin aldığı değerlerin ayrı ayrı karelerini alıp toplamak suretiyle bulunur ve istatistikte buna kısaca "kareler toplamı" denir.

$$\sum_{i=1}^{n} x_i^{\,2} = x_1^{\,2} + x_2^{\,2} + x_3^{\,2} + \ldots + x_n^{\,2}$$

Örnek: Yukarıdaki X değişkenine ait kareler toplamını bulalım,

$$\sum_{i=1}^{5} x_i^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 = 5^2 + 7^2 + 3^2 + 2^2 + 8^2$$

$$= 25 + 49 + 9 + 4 + 64 = 151$$

Olur.

Bir değişkenin kareler toplamı(KT) , o değişkenin toplamının karesine eşit değildir. Yani,

$$\sum_{i=1}^{n} x_i^2 \neq (\sum_{i=1}^{n} x_i)^2$$

Yukarıdaki X değişkeni değerleri için bu durum; $\sum_{i=1}^{5} x_i = 25$ ve $\sum_{i=1}^{5} x_i^2 = 151$
İdi ve böylece;

$25^2 \neq 151$

Olur.

İki değişkenin çarpımlar toplamı:
X ={x_1,x_2,x_3,...,x_n} ve Y={y_1,y_2,y_3,...,y_n} olsun. Bu durumda çarpımlar toplamı;

$$\sum_{i=1}^{n} x_i * y_i = x_1 * y_1 + x_2 * y_2 + \ldots + x_n * y_n$$

Örneğin; X={5,7,3,2,8) ve Y={4,2,5,6,3} ise, (n=5) çarpımlar toplamı

$$\sum_{i=1}^{5} x_i * y_i = x_1 * y_1 + x_2 * y_2 + x_3 * y_3 + x_4 * y_4 + x_5 * y_5 = 5*4+7*2+3*5+2*6+8*3$$

$$= 20+14+15+12+24 = 85$$

İki değişkenin çarpımlar toplamı,değişkenlerin ayrı ayrı toplamlarının çarpımına eşit değildir. Yani,

$$\sum_{i=1}^{n} x_i * y_i \neq (\sum_{i=1}^{n} x_i) * (\sum_{i=1}^{n} y_i)$$

Yukarıdaki örnek için bu durum $\sum_{i=1}^{5} x_i = 25,\ \sum_{i=1}^{5} y_i = 20$ ve $\sum_{i=1}^{5} x_i y_i = 85$

Olduğuna göre;

85≠500 olduğu açıkça görülür.

Toplama İşaretinin (Σ) Özellikleri:

a)X ve Y gibi iki değişkenin değerlerinin ikişer ikişer toplamlarının toplamı (farkı) bu iki değişkene ait değerlerin ayrı ayrı toplamlarının toplamına (farkına) eşittir.Yani,

$$\sum_{i=1}^{n}(x_i \pm y_i) = \sum_{i=1}^{n} x_i \pm \sum_{i=1}^{n} y_i$$

b)Bir değişkene ait değerlerin bir sabit sayı ile çarpımının toplamı, değişkenin aldığı değerlerin toplamının o sabit ile çarpımına eşittir.

$$\sum_{i=1}^{n} kx_i = k\sum_{i=1}^{n} x_i$$

c) Bir sabit sayının 1'den n'e kadar toplamı, o sabitin n ile çarpımına eşittir.

$$\sum_{i=1}^{n} k = nk$$