



İSTATİSTİK

DERS NOTLARI



Prof.Dr.YÜKSEL TERZİ

ONDOKUZ MAYIS ÜNİVERSİTESİ

FEN EDBEBİYAT FAKÜLTESİ

İSTATİSTİK BÖLÜMÜ

SAMSUN

2020

İÇİNDEKİLER

1. İSTATİSTİK KAVRAMI
 - 1.1. Giriş
 - 1.2. İstatistiğin Tanımı
 - 1.3. İstatistiğin Tarihsel Gelişimi
 - 1.4. İstatistiğin Konusu
 - 1.5. İstatistik Metodunun Aşamaları
 - 1.6. Temel Kavramlar
 - 1.6.1. Veri Ve Bilgi
 - 1.6.2. Birim
 - 1.6.3. Vasıf
 - 1.6.4. Gözlem Sonucu Ve Şık
 - 1.6.5. Anakütle
 - 1.6.6. Örneklem
 - 1.7. Değişkenler
 - 1.7.1. Nitel (Kalitatif) Değişkenler
 - 1.7.2. Nicel (Kantitatif) Değişkenler
 - 1.7.3. Kesikli Değişkenler
 - 1.7.4. Sürekli Değişkenler
 - 1.8. Ölçme Düzeyleri
 - 1.8.1. Adlandırma Ölçme Düzeyi
 - 1.8.2. Sıralama Ölçme Düzeyi
 - 1.8.3. Aralık Ölçme Düzeyi
 - 1.8.4. Oran Ölçme Düzeyi
 - 1.9. Oran ve Hız
2. SERİLER-FREKANSLAR-GRAFİKLER
 - 2.1. İstatistik Serileri
 - 2.1.1. Giriş
 - 2.1.2. Zaman ve Mekan Serileri
 - 2.1.3. Dağılım Serileri
 - 2.1.4. Birikimli Seriler
 - 2.1.5. Bileşik Seriler
 - 2.2. Frekans Tabloları
 - 2.3. Grafikler
 - 2.3.1. Basit Serilerinin Grafikle Gösterimi
 - 2.3.2. Frekans Serilerinin Grafikle Gösterimi
 - 2.3.3. Gruplanmış Serilerin Grafikle Gösterimi
 - 2.3.4. Birikimli Serilerin Grafikler Gösterimi
 - 2.3.5. Bileşik Serilerinin Grafikle Gösterilmesi
 - 2.4. Örnek Problemler
3. ORTALAMALAR
 - 3.1. Giriş
 - 3.2. Duyarlı Ortalamalar
 - 3.2.1. Aritmetik Ortalama
 - 3.2.2. Geometrik Ortalama
 - 3.2.3. Harmonik Ortalama
 - 3.2.4. Kareli Ortalama
 - 3.2.5. Tartılı Ortalama
 - 3.3. Duyarlı Olmayan Ortalamalar

- 3.3.1. Medyan
- 3.3.2. Mod
- 3.3.3. Kartiller
- 3.4. Ortalama Türünün Seçimi
- 3.5. Örnek Problemler
- 4. DEĞİŞİM ÖLÇÜLERİ
 - 4.1. Değişim Genişliği
 - 4.2. Kartiller Arası fark
 - 4.3. Ortalama Sapma
 - 4.4. Standart Sapma
 - 4.4.1. Varyans
 - 4.5. Standart Hata
 - 4.6. Değişim Katsayısı
- 5. MOMENTLER VE ASİMETRİ ÖLÇÜLERİ
 - 5.1. Momentler
 - 5.2. Çarpıklık
 - 5.3. Basıklık
 - 5.4. Örnek Problemler
- 6. OLASILIK
 - 6.1. Kümeler Kavramı
 - 6.2. Örnek Uzaylar, Örnek Noktalar ve Olaylar
 - 6.3. Sayma Kuralı
 - 6.4. Permütasyon - Kombinasyon
 - 6.5. Marjinal Olasılıklar
 - 6.6. Koşullu (Şartlı) Olasılık
 - 6.7. Bağımsız Olaylar
 - 6.8. Bayes Teoremi
 - 6.9. Olasılık Dağılımları
 - 6.9.1. Bernoulli Dağılımı
 - 6.9.2. Binom Dağılımı
 - 6.9.3. Normal Dağılım
- 7. HİPOTEZ TESTLERİ
 - 7.1. GİRİŞ
 - 7.2. Sıfır Hipotezi Ve Karşıt Hipotez
 - 7.3. I. ve II. Tip Hatalar
 - 7.4. Örnek Hacminin Belirlenmesi
 - 7.5. Tek Grup Ortalama Testi
 - 7.6. Tek Grup Oran Testi
 - 7.7. İki Bağımsız Grup Ortalama Testi
- 8. KORELASYON -BASİT DOĞRUSAL REGRESYON ANALİZİ
 - 8.1. Korelasyon
 - 8.2. Kısmi Korelasyon Katsayısı
 - 8.3. Spearman Sıra Korelasyon Katsayısı
 - 8.4. Basit Doğrusal Regresyon Analizi
 - 8.4.1. Parametrelerin En Küçük Kareler Tahmini (EKK)
- 9. ÇAPRAZ TABLOLAR-Kİ-KARE TESTLERİ
 - 9.1. Giriş
 - 9.2. Ki-Kare Uygunluk Testi
 - 9.3. Ki-Kare Bağımsızlık Testi

KAYNAKLAR

- ▶ Serper, Ö. (2000). *Uygulamalı İstatistik I,II*. Ezgi Kitabevi, Bursa.
- ▶ Tekin, V.N. (2006). *İstatistiğe Giriş*, Seçkin Yayıncılık, Ankara.
- ▶ Spigel, MR. and Stephens LJ. (1999). *İstatistik, Schaum's Outlines*, Çeviri editörleri: Alptekin Esin ve Salih Çelebioğlu, Nobel Dağıtım.
- ▶ ÇİL, B.(2004). *İstatistik*, Detay Yayıncılık, Ankara.
- ▶ Şenesen, Ü. (2002). *Matematiksel İstatistik, Literatür Yayınları*, İstanbul.
- ▶ Akdeniz, F.(1984). *Olasılık ve İstatistik*, Ankara Üniv. Basımevi, Ankara.
- ▶ Ünver, Ö. ve Gamgam, H.(2006). *Uygulamalı İstatistik Yöntemler*, Seçkin yayıncılık, Ankara.
- ▶ Çömlekçi, N.(1994). *Temel İstatistik*, Bilim Teknik Yayınevi, İstanbul.
- ▶ Işık A. (2006). *Uygulamalı İstatistik I-II*, Beta Yayınları, İstanbul.

İSTATİSTİĞE GİRİŞ

Günümüzde artan rekabet koşulları ve teknolojik gelişmeler bireyleri, işletmeleri ve ülkeleri çeşitli konularda verecekleri kararlarla ilgili olarak bilgi(veri) toplamaya ve bu verileri analiz etmeye zorunlu kılmaktadır. Gerek verilerin toplanması gerekse toplanan verilerin analiz edilmesi ve verilecek kararlarda kullanılması istatistiğin önemini ortaya koymaktadır

Günlük yaşamımızda pek çok olayla ilgilenir ve birçok sorular sorup cevaplar ararız. Örneğin; “Yaklaşmakta olan bayram tatilinde havalar nasıl olacak? Üniversite öğrencilerinin üniversiteye giriş sınavlarında aldıkları puanlarla üniversitedeki başarıları arasında ne-kadar ilişki vardır?” gibi sorular. Cevaplarından tam emin olmadığımız ve cevapları hakkında bilgimizin çok sınırlı olduğu soruları güvenilir biçimde cevaplandırmanın en kolay yolu, bunlarla ilgili bilgi toplamak ve bilgileri çözümleyerek anlamlı sonuçlar çıkarmaktır.

Toplanan bilgiler genelde sayısaldır ya da anlam kolaylığı için sayılarla gösterilir. Örneğin bir okula devam eden öğrencilerin boy uzunlukları, kiloları, yaşları, cinsiyetleri, zeka seviyeleri gibi bilgiler.

İstatistik; belirli amaçlar için veri toplama, toplanan verileri düzenleme, analiz etme ve elde edilen sonuçların yorumlaması ile ilgilenen bilim dalıdır.

İstatistiğin bir diğer tanımı; örneklem verileri üzerinden hesaplanan ortalama, varyans ve oran gibi tahmin edicilere istatistik denir.

İstatistik teknik ve yöntemlerini kullanılış amaçlarına göre iki grupta toplanabilir.

a) Tanımlayıcı İstatistik (descriptive statistics):

Elde edilen verilerin sınıflandırılması, frekans dağılımlarının yapılması, bu dağılımların ortalamalar, yüzdelikler ve standart sapma gibi ölçülerle tanımlanarak tablo ya da grafiklerle okuyuculara sunulması tanımlayıcı istatistiğin konularıdır.

b) Çıkarımsal İstatistik (inferential statistics):

Örneklemeden elde edilen bulgularla, örneklemin seçildiği(çekildiği) anakütle (kitle, popülasyon, evren, yığın) hakkında tahminlerde bulunma, karşılaştırmalar yapma ve kararlara varma işlemleri çıkarımsal istatistiğin konularıdır.

Günümüzde istatistik, deney yada gözlemlere dayalı tüm bilim dallarında geniş bir uygulama alanına sahiptir. Fen bilimleri, sağlık bilimleri, eğitim bilimleri ve sosyal bilimler gibi çok geniş bir alanda istatistik kullanılmaktadır.

İstatistiğin Tarihsel Gelişimi

İstatistiğin uygulamada kullanılışı çok uzun bir gelişime sahip olmasına rağmen bilim olarak 19.yy. ve 20.yy. da ortaya çıkmıştır. İnsanlar topluluklar halinde yaşamaya başlayıp devletler kurunca, yönetenler işleri daha düzenli yürütülebilmek için bir takım bilgilere ihtiyaç duydular. Bunlar başlangıçta toplumdaki birey sayısı, asker sayısı hayvan sayısı vb. hususları kapsıyordu.

Zamanla bu bilgiler yenilendi ve gelişti. Eski Mısır'da bazı devlet görevlileri bütün aile reislerinin listesini tutuyorlardı. Yine Mısır'da M.Ö. 3000 yıllarında piramit inşaatına gerekli iş gücü talebini garanti etmek için ilk nüfus sayımı yapılmıştır. Osmanlı devletinde de ilk dönemlerden itibaren istatistiksel bilgilerin toplanmasına önem verilerek, Orhan Bey zamanında çeşitli sayımlar yapılmıştır.

İstatistiğin gerçek ilerlemesi Yeni Çağ' dan itibaren başlamıştır. 17.yy.'da Fransa'da ilk defa maliye ve dış ticaret istatistikleri düzenlenmesine başlanmıştır. İlk bilimsel nüfus sayıma A.B.D.'de 1790 yılında yapılmıştır.

İstatistik yöntem bilimi, istatistiksel işlemlerin uygulamasından çok sonra ortaya çıkmış ve çeşitli aşamalardan geçerek bugünkü durumunu almıştır. 17.yy.'ın ilk yarısından itibaren bazı Alman Üniversitelerinde okutulan " Devletlerin Özellikleri "adlı yeni bir derste çeşitli ülkelerin tarihi, mali. askeri ve idari özellikleri hakkında bilgi veriliyordu. Bir müddet sonra bu konuya, statüs (devlet) 'den gelme statistik (istatistik) denilmeye başlandı.

İstatistik Metodunun Aşamaları

İstatistik metodu dört aşamada uygulanır. Bu aşamalar:

i). Bilgilerin Toplanması

Bu aşama araştırmasının konusunun ve birimlerinin kesin tarifi ile başlar. Bilgi toplamanın ne zaman yapılacağına ve kapsamının ne olacağına bu aşamada karar verilir.

ii). Bilgilerin Organize Edilmesi

Bu aşamada toplanmış olan ham veriler matematik ve istatistik analizlere elverişli, düzenli bir hale getirilir. Verilerin tasnif edilmesi ve gruplandırılması bu aşamada yapılması gereken işlerdir.

iii). Verilerin Sunulması

Düzenli ve gruplanmış verilerin tablo ve grafik halinde sunulması ve bu işlemlerle ilgili metotlar bu aşamada uygulanır.

iv). İstatistik Tahsil

Çeşitli metotlar kullanarak düzenli verilerin derinlemesine analizini yapmak, olaylarla ilgili eğilimleri ortaya çıkarmak, istatistik testler yardımıyla sonuca varmak ve karar vermek bu aşamanın incelediği konulardır. Bu metotlar, istatistik metodolojisinin önemli bir kısmını meydana getirir.

Temel Kavramlar

İstatistiğin iyi anlaşılması için istatistikte çok sık kullanılan bazı kavramların anlamlarının, birbirleri ile olan ilişkilerinin ve farklılıklarının iyi bilinmesi gerekir.

Veri: Belirli amaçlar için toplanan bilgilere veri denir

Birim: İncelenmesi amaçlanan konudaki en küçük parçaya birim denir. Birimler canlı ya da cansız varlıklar, kurumlar, kuruluşlar olabilir.

(Üzerinde gözlem ve ölçüm yapılan ve anakütleyi oluşturan en küçük öğeye birim adı verilir)

Anakütle (Evren, Kitle, Yığın, Popülasyon): Hakkında bilgi edinmek istenen bütüne anakütle denir.

Bir başka ifadeyle, üzerinde inceleme ve araştırma yapılacak olayın gözlenebileceği tüm birimlerin yer aldığı topluluktur.

Kitledeki toplam birim sayısı N ile gösterilir ve “kitle büyüklüğü” yada “kitle hacmi” olarak ifade edilir.

Sayılabılır sayıda birim içeren anakütlelere “sonlu anakütle”, sayılamayan sayıda birim içeren anakütlelere “sonsuz anakütle” denir.

Sonlu anakütlerde ilk ve son birim bilinirken, sonsuz anakütlerde ilk birim bilinir fakat son birim bilinmez.

“Türkiye’de çalışma çağındaki insanlar” kitlesi sonlu kitleye, “Karadeniz’deki balıklar” kitlesi sonsuz kitleye örnek olarak verilebilir.

Örneklem(Örnek): Anakütleden seçilen ve anakütleye göre daha az sayıda birimden oluşan topluluktur.

Örneklemdeki birim sayısı n ile gösterilir ve “örneklem büyüklüğü” ya da “örneklem hacmi” olarak ifade edilir ve $n < N$ dir.

Örneklemden edindiğimiz bilgilere dayanarak anakütle hakkında genellemeler yaparız. Bu sebeple örneklemin anakütleyi iyi temsil etmesi gerekir.

Örnekleme: Kitleden, amaca uygun örneklem seçme(çekme) işine örnekleme denir.

Anakütleden yansız örneklem seçebilmek için geliştirilmiş çeşitli örnekleme yöntemleri vardır. Basit rastgele örnekleme, Küme örnekleme, Sistemantik örnekleme, Kota örnekleme gibi.

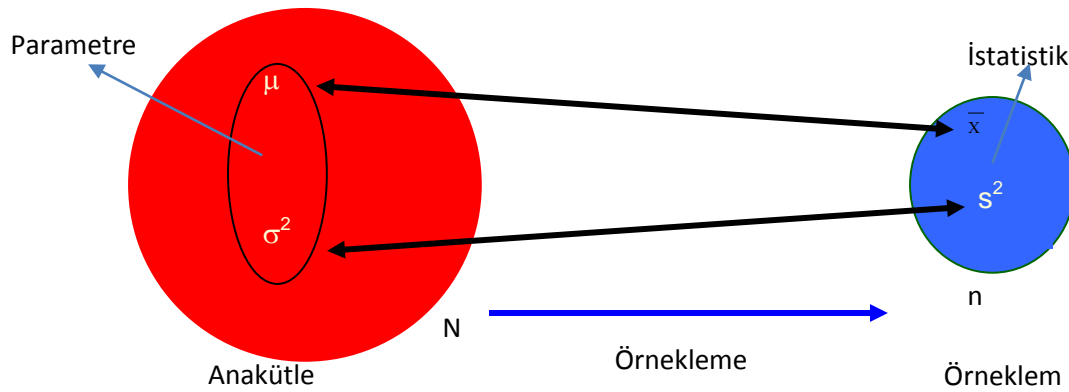
Parametre: Anakütlenin özelliklerini belirleyen sayısal karakteristiklere parametre adı verilir.

Parametreler, anakütleyi tanımlamada kullanılabilen tipik değerlerdir. Kitle ortalaması (μ), kitle varyansı (σ^2), kitle oranı (π) gibi değerlerdir.

İstatistik: Örneklem verileri üzerinden hesaplanan, anakütleyi tanımlayan değerlere karşılık gelen tahmin edicilere istatistik denir. Örneklem ortalaması (\bar{x}), örneklem varyansı (s^2), örneklem oranı (p) gibi.

Tamsayım: Sonlu bir anakütlenin bütün birimlerinin incelenmesi ya da sayılması işlemidir.

Yukarıda verilen tanımların birbiriyle olan ilişkisi aşağıda toplu olarak gösterilmiştir.



Şekil: Anakütle-örneklem ilişkisi

Örnek 1. Yeni bir ücret sisteminin uygulandığı 30 işçisi olan bir işletmede, işçilerin yeni ücret sisteminden memnuniyetleri araştırılmak istenmektedir. Burada tamsayım yapılabilir mi?

Çözüm: Burada anakütle $N=30$ işçiden oluşmaktadır ve küçük hacimli bir anakütledir. İşçilerin her birine ulaşmak ve bunlardan veri elde etmek kolaydır. Bu yüzden tamsayım yapılabilir.

Örnek 2. 25000 öğrencisi bulunan bir üniversitede, öğrencilerin kendilerine sunulan hizmetleri yeterli bulup bulmadıklarını belirlemek amacıyla, bir araştırma planlanmış ve rasgele seçilen 400 öğrenciden görüşleri alınmıştır. Bu araştırmada anakütle hacmi ve örneklem hacmi kaçtır? Neden tamsayım yapılmamıştır?

Çözüm: Anakütle büyük hacimli sonlu bir anakütledir ve $N=25000$, örneklem ise $n=400$ öğrenciden oluşmaktadır. 25000 öğrencinin tümünün görüşüne başvurmak, onlara ulaşmak zordur. Bu yüzden tamsayım oldukça zordur.

Örnek 3. Bir fabrikada üretilen bisküvi paketlerinin, planlanan ağırlıkta üretilip üretilmediğinin araştırılması amacıyla, üretilen paketler arasından 200 paket seçilmiştir. Bu araştırmada anakütle ve örneklem nedir? Tamsayımın yapılıp yapılmayacağını açıklayınız?

Çözüm: Anakütle sonsuzdur. Açıktır ki bu tür anaküteller üzerinde tam sayım yapılamaz, örnekleme zorunludur. Örneklem 250 bisküvi paketinden oluşan topluluktur.

Örnekleme Yapmayı Gerekli Kılan Sebepler:

Maliyet: Popülasyonun hepsini incelemek çok masraflı olabilir. Popülasyondan alınacak küçük örnekler yardımı ile gerçeğe yakın bilgiler elde edilebilir.

Zaman: Popülasyonla yapılacak bir çalışma çok uzun zamana ihtiyaç gösterebilir. Halbuki örnekle çalışılırsa gerçeğe yakın bilgiler kısa zamanda elde edilebilir.