

TEMEL İSTATİSTİK I

DERS NOTLARI

Prof.Dr.YÜKSEL TERZİ

ONDOKUZ MAYIS ÜNİVERSİTESİ

FEN-EDEBİYAT FAKÜLTESİ

İSTATİSTİK BÖLÜMÜ

SAMSUN

2019

İÇİNDEKİLER

1. İSTATİSTİK KAVRAMLAR	4
1.1. Giriş	4
1.2. İstatistiğin Tanımı	4
1.3. İstatistiğin Tarihsel Gelişimi	7
1.4. İstatistiğin Konusu	8
1.5. İstatistik Metodunun Aşamaları	10
1.6. Temel Kavramlar	11
1.6.1. Veri ve Bilgi	11
1.6.2. Birim	11
1.6.3. Vasıf	12
1.6.4. Gözlem Sonucu ve Şık	13
1.6.5. Anakütle	13
1.6.6. Örneklem	14
1.7. Değişkenler	17
1.7.1. Nitel (Kalitatif) Değişkenler	19
1.7.2. Nicel (Kantitatif) Değişkenler	19
1.7.3. Kesikli Değişkenler	20
1.7.4. Sürekli Değişkenler	20
1.8. Ölçme Düzeyleri	21
1.8.1. Adlandırma Ölçme Düzeyi	22
1.8.2. Sıralama Ölçme Düzeyi	22
1.8.3. Aralık Ölçme Düzeyi	23
1.8.4. Oran Ölçme Düzeyi	23
1.9. Oran ve Hız	24
1.10. Örnek Problemler	27
2. SERİLER-FREKANSLAR-GRAFİKLER	29
2.1. İstatistik Serileri	29
2.1.1. Zaman ve Mekan Serileri	29
2.1.2. Dağılım Serileri	29
2.1.3. Birikimli Seriler	32
2.1.4. Bileşik Seriler	34
2.2. Frekans Tabloları	35
2.3. Grafikler	39
2.3.1. Basit Serilerinin Grafikle Gösterimi	39
2.3.2. Frekans Serilerinin Grafikle Gösterimi	46
2.3.3. Gruplanmış Serilerin Grafikle Gösterimi	50
2.3.4. Birikimli Serilerin Grafikler Gösterimi	51
2.3.5. Bileşik Serilerinin Grafikle Gösterilmesi	52
2.4. Örnek Problemler	55
3. ORTALAMALAR	56
3.1. Giriş	56
3.2. Duyarlı Ortalamalar	57
3.2.1. Aritmetik Ortalama	57
3.2.2. Geometrik Ortalama	61
3.2.3. Harmonik Ortalama	64
3.2.4. Kareli Ortalama	68
3.2.5. Tartılı Ortalama	69
3.3. Duyarlı Olmayan Ortalamalar	72

3.3.1. Medyan	72
3.3.2. Mod	77
3.3.3. Kartiller	81
3.4. Ortalama Türünün Seçimi	85
3.5. Excel ve SPSS'te Ortalama Hesabı	86
3.6. Örnek Problemler	89
4. DEĞİŞİM ÖLÇÜLERİ	91
4.1. Değişim Genişliği	92
4.2. Kartiller Arası fark	93
4.3. Ortalama Sapma	94
4.4. Standart Sapma	95
4.4.1. Varyans	98
4.5. Standart Hata	101
4.6. Değişim Katsayısı	104
5. MOMENTLER VE ASİMETRİ ÖLÇÜLERİ	107
5.1. Momentler	107
5.2. Çarpıklık	111
5.3. Basıklık	115
5.4. Excel ve SPSS Uygulaması	116
5.5. Örnek Problemler	118
6. NORMAL DAĞILIM	119
6.1. Puan Dönüşümleri	128
6.2. Binom Dağılımının Normal Dağılıma Yaklaşımı	131
6.3. Poisson Dağılımının Normal Dağılıma Yaklaşımı	134
Z TABLOSU	136
KAYNAKLAR	139

I.BÖLÜM

1. İSTATİSTİK KAVRAMLAR

1.1. Giriş

Günümüzde artan rekabet koşulları ve teknolojik gelişmeler bireyleri, işletmeleri ve ülkeleri çeşitli konularda verecekleri kararlarla ilgili olarak verileri toplamaya ve bunları analiz etmeye zorunlu kılmaktadır. İşte gerek verilerin toplanması gerekse toplanan verilerin analiz edilmesi ve verilecek kararlarda kullanılması istatistiğin önemini ortaya koymaktadır. Dar anlamda bir konuda toplanan verilerden o konu ile ilgili karar verinceye kadar yapılan tüm işlemleri kapsayan geniş anlamda ele alınan istatistik (**statistics**) günümüzde bu yelpazede çok farklı anlamlarda kullanılmaktadır.

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 sonuçlar çıkarmaktır. Topladığımız bilgiler genelde sayısaldir ya da anlam kolaylığı için biz sayılarla gösteririz. Mesela bir okula devam eden öğrencilerin boy uzunlukları, ağırlıkları, yaşları, zeka seviyeleri gibi bilgileri toplamaktır.

İstatistik sözcüğünün kökeni konusunda kesin bir görüş birliği yoktur. Bazı bilimciler sözcüğün Latince’de “devlet ve durum” anlamına gelen status kelimesinden, ya da Yunanca “gözlem” anlamına gelen “statizien” kelimesinden türemiştir.

1.2. İstatistiğin Tanımı

İstatistikle ilgili çeşitli tanımlar yapılmıştır. Bunlardan bazıları da bilim dışı, ciddiyeşsiz tanımlardır.

“Üç çeşit yalan vardır: yalan, kuyruklu yalan ve istatistik”

Disraeli

“İyi ambalajlanmış bir istatistik, Hitler’in büyük yalanından daha etkilidir”

Darell Huff

Bu tanımlar istatistiğin işe yaramadığını değil, her şeyin istatistikle kanıtlanamayacağını vurgulamaktadır. Örneğin Üsküdar-Eminönü arasında seyahat eden vapur seferleri ile İstanbul’daki boşanma oranları arasında bulunan yüksek bir

korelasyon (ilişki), bu iki olayın birbiri ile ilişkili olduğunu düşündürülebilir. Ancak bu reddedilmesi gereken bir durumdur. **Rakamlar yalan söylemez, fakat yalancılar rakam söyler.**

Ortalama sıcaklığın 40C° olduğu bir çöle üstündeki ince bir tişörtle giden bir kişi aldanacaktır. Çünkü çölde sıcaklık gündüz 50C°'ye çıkmakta, gece ise -10C°'ye düşmektedir.

İstatistik sözcüğü farklı yaklaşımlara göre değişik anlamlar taşır. Günlük dilde istatistik denildiği zaman, belirli bir olaya ilişkin derlenmiş sayısal bilgiler akla gelir. Örneğin dışalım, turizm, inşaat istatistikleri gibi.

İstatistiğin birçok tanımı mevcuttur. İstatistiğin değişik anlamlarda kullanılması doğaldır. Çünkü günümüzde bütün çalışma alanlarında kullanılmaktadır. 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 bilginin çok sınırlı olduğu soruları güvenilir biçimde cevaplandırmanın en kolay yolu, bunlarla ilgili bilgi toplamak ve bilgileri çözümleyerek sonuçlar çıkarmaktır.

Topladığımız bilgiler genelde sayısaldır ya da anlam kolaylığı için biz sayılarla gösteririz. Mesela bir okula devam eden öğrencilerin boy uzunlukları, ağırlıkları, yaşları, zeka seviyeleri gibi bilgileri toplamaktır.

Belirli amaçlarla toplanan sayısal bilgilere **veri** denir. Verileri inceleme işi ile uğraşan bilime **istatistik** denir. Başka bir deyişle istatistik; belirli amaçlar için veri toplama, toplanan verileri düzenleme, çözümleme yorumlama teknik ve yöntemleridir. İstatistik, bilimsel yöntemin en güçlü temel araçlarından biridir. Bir başka açıdan istatistik; hem teknik ve yöntemler geliştiren bir bilim, hem de yöntemler topluluğu sayılabilir.

İstatistikle ilgili çeşitli tanımlar yapılmıştır.

- ✓ Yığın olayları incelemek, olaylarla ilgili toplanan verileri analiz etmek, olayların sebep ve sonuçlarını açıklamak, aralarındaki ilişkileri ortaya koymak için kendine özgü yöntemleri olan bir bilim dalıdır.
- ✓ Olayların nicel ve nitel yönlerinin tablolar, grafikler veya sayısal değerler şeklindeki özet ifadelerdir. Ölüm istatistikleri, doğum istatistikleri, göç istatistikleri, kaza istatistikleri gibi özet bilgiler istatistikler olarak adlandırılır.
- ✓ Örneği oluşturan birimlerden hesaplanmış, ana kütleyi tanımlayan değerlere karşılık gelen değerlerdir. Örnek ortalaması, örnek varyansı gibi.
- ✓ ***İstatistik, geçmişi ve şimdiki durumu çeşitli sayısal tekniklerle analiz ederek gelecek hakkında karar vermeyi sağlayan bir bilim dalıdır.***
- ✓ İstatistik, çeşitli olaylarla ilgili toplanmış veriler demektir. Nüfus, istihdam, kaza istatistikleri, yıllık imalat sanayi anketleri veya Merkez Bankası'nın periyodik olarak yayınladığı bültenler örnek olarak verilebilir. Belirli bir konu ile ilgili sayısal

verilerin en uygun şekilde derlenerek açık ve anlaşılır bir şekilde ifade edilmesidir.

- ✓ İstatistik verilerin toplanması, organize edilmesi, özetlenmesi, sunulması, analiz edilmesi ve bu verilerden bir sonuca varılabilmesi ile ilgili olarak kullanılan bilimsel yöntemler topluluğudur.
- ✓ Belirli amaçlarla toplanan sayısal bilgilere veri denir. Verileri inceleme işi ile uğraşan bilime **istatistik** denir. Başka bir deyişle istatistik; belirli amaçlar için veri toplama, toplanan verileri düzenleme, çözümleme yorumlama teknik ve yöntemleridir. İstatistik, bilimsel yöntemin en güçlü temel araçlarından biridir. Bir başka açıdan istatistik; hem teknik ve yöntemler geliştiren bir bilim, hem de yöntemler topluluğu sayılabilir.

Sosyal Bilimler açısından bakıldığında istatistik; çeşitli anlamlarda kullanılan bir sözcüktür. Günlük yaşamda istatistik deyişi ile çeşitli olaylara ilişkin olarak toplanan rakamlar yani verileri örneğin tarım istatistikleri, turizm istatistikleri gibi rakamlarla istatistik anlatılmak istenir. İstatistik, bu anlamda daima çoğul olarak kullanılır. Ancak istatistik, herhangi bir rakam değildir. Belirli bir olayın gözlenmesiyle onun hacmi, bölünüşü, büyüklüğü vs. hakkında elde edilen rakamların ifade eder. Bu yüzden sözcüğü milli piyangodan ikramiye kazananların listesi ve logaritma cetvellerine istatistik denilemez.

Daha geniş anlamda, istatistik sözcüğü belirli olaylar hakkında nicel bilgilerin toplanmasında, işlenmesinde, analiz ve yorumunda kullanılan bütün yöntemi ifade eder. Ancak bu anlamda istatistik yöntemi şeklinde belirtilmelidir.

İstatistik için yapılabilecek bir başka tanımda örneklem değeri olması ile ilgilidir. İstatistik hakkında bilgi edinilmek istenen bütüne anakütle, ondan alınan parçalara da örneklem adı verilir. Dolayısı ile bir özellik hem anakütlerde hem de örneklem de söz konusu olabilmektedir. İşte örnekleme ait olan ve örnekleme karakterize eden değerlere istatistik adı verilir. İstatistikte genellikle örneklem incelenerek anakütle hakkında bilgi sahibi olmaya çalışıldığından, yani kısaca tümevarım yöntemi kullanıldığından istatistik için buna uygun bir tanımda yapılabilmektedir. Buna göre istatistik, örneklem istatistiğinden hareketle anakütle parametresini tahmin etmeye çalışan bir bilim dalıdır şeklinde tanımlanabilir.

İstatistik teknik ve yöntemlerini kullanılış amaçlarına göre, iki genel grupta toplayabiliriz:

Tanımlayıcı İstatistik (descriptive statistics):

Tüm istatistikleri elde etmek için kullanılan ve sonuçları yorum yapmadan veren istatistiklere verilen addır. Tanımlayıcı istatistik verilerin kapsamlı açıklaması şeklinde de tanımlanabilir.

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

Çıkarımsal İstatistik (inferential statistics):

Kitleye ilişkin genellemelerin yapılmasını sağlayan yöntemlerdir. Örneklemden elde edilen bulgularla, örneklemin seçildiği anakütle hakkında tahminlerde bulunma, karşılaştırmalar yapma ve kararlara varma işlemleri çıkarımsal istatistiğin konularıdır.

Bir bilim olarak istatistik, uygulamalı matematiğin bir dalı olup bütün alanlarla ilişkilidir. İstatistiğin genel bir kuramlar, teknikler ve yöntemler bütünü vardır. Bunlar gözlem yapılabilen her alana uygulanır. Ancak, bazı teknik ve yöntemler bazı alanlarda karşılaşılan özel durumlara daha uygun düşer ve daha fazla kullanılır.

Dar alanda istatistik, çeşitli olaylarla ilgili toplanmış veriler demektir. Nüfus, istihdam, kaza istatistikleri, yıllık imalat sanayii anketleri veya Merkez Bankası'nın periyodik olarak yayınladığı bültenler örnek olarak verilebilir. Belirli bir konu ile ilgili sayısal verilerin en uygun şekilde derlenerek açık ve anlaşılır bir şekilde ifade edilmesi büyük önem taşısa da, bu konu genel anlamdaki istatistiğin sadece ilk safhasını meydana getirmektedir.

Günümüzde istatistik, deney yada gözlemlere dayalı tüm bilim dallarında geniş bir uygulama alanına sahiptir. Fen bilimleri (biyoloji, fizik, kimya gibi), sağlık bilimleri (tıp, diş, eczacılık gibi), eğitim bilimleri ve sosyal bilimleri (sosyoloji, psikoloji, coğrafya gibi) çok geniş bir alanda istatistik kullanılmaktadır.

1.3. İSTATİSTİĞİN TARİHSEL GELİŞİMİ

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.

Tipik olaylar, birbirinin tam benzeri olan ve tek nedenle bağlı olarak meydana gelen olaylardır. Olayları meydana getiren gelen ve tesadüfi nedenler vardır. Genel yada temel nedenler her olaya eşit derecede ve tek yönde etki eden nedenlerdir. Tesadüfi nedenler ise her olaya farklı derecede ve çift yönde etki eden nedenlerdir. İşte tipik olaylar sadece genel nedenlerin etkisi ile ortaya çıkan ve birbirleri ile tamamen aynı olan dolayısı ile birbirinin incelenmesi ile hepsi hakkında kesin bilgi elde edilen ve tekrarı gerektirmeyen olaylardır. Örneğin; bir kişinin kan grubunu belirlemek için o kişiden defalarca ve vücudunun farklı bölgelerden kan alarak tahlil etmeye gerek yoktur. Aynı şekilde deniz seviyesinde suyun kaynama derecesini belirlemek için olayı defalarca tekrarlamaya gerek yoktur. Cisimlerin yer çekimi nedeni ile yere düşmelerini belirlemek için olayı defalarca tekrarlamaya veya farklı cisimler için bunu denemeye gerek yoktur. Bu tür olaylarda değişkenlik olmadığından benzerlikleri bulup bunları genelemeye çalışmak söz konusu olmadığı için bunları istatistiğin konusu içine girmezler. 

1.4. İstatistiğin Konusu

Yığın Olay: İstatistik diğer bilim dalları gibi olayları konu alır. Olay varsa istatistik vardır. Ancak her olay istatistiğe konu oluşturmaz.

Bir olaylar kümesindeki tek bir olay kümedeki diğer olayları temsil edemiyorsa, bu tür olaylara yığın olay denir. İstatistik yığın olaylarla ilgilenir. Örneğin firmaların yıllık ciroları, trafik kazaları, evlenmeler, boşanmalar, doğumlar, ölümler gibi her gün karşılaşılan olaylar yığın olaydır.

Tipik Olay: Eğer bir olaylar kümesinde tek bir olay, tüm olaylar kümesini temsil edebiliyorsa, bu tür olaylara tipik olay denir. Örneğin ideal koşullar altında ve uygun bir laboratuvar ortamında iki hidrojen ve bir oksijen atomu bir araya gelirse su elde edilir. Deneyin her tekrarında aynı sonuç elde edileceğinden tek bir deney ilgili olaylar kümesini temsil eder. İstatistik tipik olaylarla ilgilenmez.

Neden İstatistik?

İçinde yaşadığımız Dünya hızla değişmektedir. Günlük gazetelere bir göz atar ve radyo Tv yayınlarını izlerse ele alınan konu ne olursa olsun, sayısal ifadelerin sık sık kullanıldığını görürüz. Ayrıca, günlük yaşantımız süresince karşılaştığımız "ne kadar", "ne zaman", "nerede" 7 "nasıl" ve "kaç tane" gibi soruları çoğu kez sayısal ifadelerle cevaplayabiliriz. İstatistik sayısal bilgileri inceleyen bilim olduğuna göre, istatistik bilgisi en azından anakütlede olup bitenleri anlama ve bunları başkalarına anlatmada yardımcı olur.

Öte yandan, çeşitli alanlarda karşılaşılan sorunlara çözüm yolları bulma ihtiyacı gün geçtikçe artmaktadır. Bunun yanı sıra, sorunları çözümlemede "pratik muhakeme" ya da geleneklere dayanma yerine gözlemlerde bulunarak sonuçları bilimsel yollarla inceleme ihtiyaç ve eğilimini de gün geçtikçe güçlendirmektedir. Böyle bir ortamda farklı alanlarda günlük meselelerin ötesinde işlem ve sorunlarla uğraşan kişilerin, doğal bir parçası olan istatistiğin yöntemlerini bilmesi bir zorunluluktur.

Bir araştırmanın düzeni teknik yönden hatalıysa, hiçbir istatistik teknik ve yöntemi böyle bir araştırmadan geçer ve güvenilir sonuçlar çıkarma olanağı sağlayamaz; başı ve sonu belirsiz verileri anlamlı hale getiremez. Bu nedenle istatistik bilgi ve anlayışı yalnız veriler toplandıktan sonra değil, araştırmanın düzenlenmesi ve yürütülmesi aşamalarında da gereklidir. Bunu sağlamanın en uygun yolu da istatistik teknik ve yöntemlerini gereğince öğrenmektir.

Bir bölgede veya ülkede sağlık durumunun saptanması, başka bölgeler veya ülkelerle karşılaştırılması, buna paralel olarak sağlık tedbirlerinin alınması ve uygulamaların kontrolü, toplumun değişik özelliklerinin sağlık sorunları üzerine etkilerinin saptanması hep istatistik yöntemler kullanılarak yapılır. Koruyucu hekimlikte istatistik yöntemlerin bilinmesi ve uygulanması bu açıdan oldukça önemlidir.

Tedavi edici hekimlikte hastaların yaş, cinsiyet, ırk, ekonomik durumu, mesleği, yaşadığı çevre gibi değişik özelliklerinin hastalığın gelişimi üzerine etkileri, tedavi yöntemlerinin etkinliği v.b durumlar için istatistik analizlerden yararlanmak, sonuçların yorumuna objektiflik kazandırmak, çalışmanın inandırıcılığı yönünden çok önemlidir.

Karşımıza çıkan bir istatistik sonucu aşağıdaki gibi sorgulanmalıdır.

- Önce bilinçli sapma olup olmadığına bakılır. İşe yarayan veriler seçilmiş, işe yaramayan veriler örtbas edilmiş olabilir.
- Uygun olmayan ölçü kullanılmış olabilir. Medyan yerine aritmetik ortalama gibi.
- Örneklem güvenilir bir sonuç verecek kadar büyük mü? Korelasyon bir anlam verecek kadar büyük mü?
- Güvenilirlik ölçüsü (olası hata, standart hata) verilmeden önünüze konan bir korelasyon ciddiye alınmaz.
- Ortalamanın türü muhakkak belirtilmelidir.

- Bazen oranlar verilir, ancak sayılar ortada görünmez. Bu yanıltıcı olabilir. Mesela bir büroda çalışan bayanların %33'ü (1/3) bankacılar ile evlenmiştir. Ancak burada sayı belirtilmemiştir. Bu işyerinde 3 bayan vardır ve biri bankacı ile evlenmiştir.
- Bazen olgulardaki değişikliğe neden olan faktör görünmez. Mesela Ekim ayı perakende satışları geçen yılın Ekim ayına göre artış göstermiştir. Ancak burada özel olan durum, bu seneki ekim ayının ramazan ayına denk gelmesidir. Bu ayrıntı açıklamada yer almamıştır.
- İstatistiği değerlendirirken elde edilmiş sayılardan sonuca giderken bir saptırma yapıp yapılmadığına bakılır. Mesela Çin'de bir bölgenin nüfusu 28 milyon bulunmuştur. Aynı bölgenin 5 yıl sonraki nüfus sayımı ise 105 milyon çıkmıştır. Ancak burada hangi sayımın yapıldığı belirtilmemiştir. Birinci sayım vergi ve askerlik için, ikinci sayım ise gıda yardımı üzerine yapılmıştır.

1.5. İstatistik Metodunun Aşamaları

İstatistik metodu dört aşamada uygulanır. Bu aşamalar:

i). Bilgilerin Toplanması (Röleleveler)

Bu aşama araştırmasının konusunun ve birimlerinin kesin tarifi ile başlar. Rölevenin 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 Tahlil

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.

1.6. TEMEL KAVRAMLAR

İstatistiğin iyi anlaşılması için istatistikte çok sık kullanılan bazı kavramların anlamlarının, birbirleri ile olan ilişkilerinin ve farklılıklarının iyi bilinmesi gerekir. **İstatistik yığın olaylarla** ilgilenir. Yığın olay, bir olaylar kümesinde tek bir olayın diğerlerini bağlı olarak da ait olduğu kümeyi temsil edemeyen olaylardır.

1.6.1. Veri ve Bilgi

Gözlem: Birimlerde incelenen özelliğin gözlenmesi veya ölçülmesi suretiyle elde edilen değerlerdir.

Veri: İki veya daha fazla denek üzerinden elde edilen bir veya daha fazla değişkene ait sayısal değerler kümesi veridir. Yani bir gözlem veya deney sonucunda ölçümlerle elde edilmiş olan bilgilerdir. Bir klinikte muayene edilen şahıslara ait tansiyon değerleri veridir.

İstatistikte en çok eşanlamlı kullanılan kavramlar veri ve bilgi kavramıdır. **Veri** (data-gözlem) incelenen birimlerin çeşitli özelliklerine ait sembolik değerlerdir. Semboller yerine çoğunlukla rakamlar kullanılır. **İstatistiksel olarak veri, analiz ve yorumlama için kullanılan bilgidir.** **Bilgi** (information) ise birimlerden elde edilen verilerin işlenerek anlamlı hale getirilmiş halidir. Bir başka ifade ile ham verilerin işlenmiş halidir. Elde edilen veriler birtakım işlemlere tabi tutulduktan sonra yani süzgeçten geçirildikten sonra bilgiye dönüştüklerine göre hacimce küçülürken değerce büyümektedirler. Dolayısı ile veriler hacim olarak büyük değer olarak küçük iken, bilgi aksine hacim olarak küçük değer olarak büyüktür.

Faktör: Birimlerin incelemeye alınan özellikleri üzerinde etkileri olduğu kabul edilen dış etmenlerdir. Birimin incelenen özelliği (diyelimki tansiyon) dışında birimin yaşı, cinsiyeti, sosyo-ekonomik durumu, diğer etmenler faktör olarak alınır.

Risk Faktörü: Bir olayın ortaya çıkmasında kesin etkisi olup olmadığı bilinmeyen, ancak varlığında olayın ortaya çıkmasını etkilediğinden şüphelenilen faktörlere risk faktörü denir. Örneğin sigara akciğer kanseri için bir risk faktörüdür.

1.6.2. Birim (olgu=denek=case)

Üzerinde gözlem ve ölçüm yapılan ve anakütleyi oluşturan en küçük öğeye birim adı verilir. Yığın olay niteliğindeki her bir olaya birim denir. Kütleyi oluşturan ve sayısal olarak incelebilen kollektif olaylardan her birine birim denir.

Birimler canlı yada cansız varlıklar olabileceği gibi, kurum, kuruluş da olabilir. Her ilaç(veya her kutu) bir birimi oluşturabilir. Bir olayın birim olabilmesi için kesinlikle ölçülmeye ve sayılmaya elverişli olması gerekir. İnsan, hayvan gibi canlı bir yaratık, bina, ağaç, araba gibi her hangi bir şey, aile, banka, şirket gibi sosyal bir kuruluş veya

doğum, ölüm, evlenme, boşanma, suç işleme gibi bir olay birime örnek olarak gösterilebilir. Renkler, kokular ve rüya birim olmazlar.

i) Maddesel Bir Varlığa Sahip Olan ve Olmayan Birimler:

Eğer birimler insan, araba, ve benzeri gibi canlı yada cansız maddesel bir varlığa sahipse, bu tür birimlere maddesel varlığa sahip birimler denir. Eğer birimler doğum, ölüm, trafik kazası gibi olay niteliğindeyse bu tür birimlere maddesel varlığa sahip olmayan birimler adı verilir.

ii) Doğal-Doğal Olmayan Birimler:

Nitelikleri açısından bir bütün oluşturan, parçalanmaları yada birleştirilmeleri halinde niteliklerini kaybeden birimlere doğal birim adı verilir. Örneğin bir canlı parçalandığında canlı olma niteliğini kaybeder.

Nitelikleri açısından bir bütün oluşturmayan, parçalanmaları yada birleştirilmeleri halinde niteliklerini kaybetmeyen birimlere doğal olmayan birim adı verilir. Bir arsa kaç parçaya bölünürse bölünsün, o parçaları yine arsadır.

iii) Sürekli-Ani Birimler:

Belirli bir zaman aralığı içinde herhangi bir anda gözlenebilen birimlere sürekli birimler adı verilir. Örneğin insan, bina, firma gibi birimler sürekli birimlerdir. Bu tür birimler varlıklarını sürdükleri sürece gözlenebilirler. Maddesel bir varlığa sahip olan birimler aynı zamanda sürekli birimlerdir.

Evlenme, boşanma, trafik kazası gibi ömürleri kısa olan ve aniden ortaya çıkan birimlere de ani birimler denir. Ani birimler maddesel bir varlığa sahip olmayan birimlerdir.

1.6.3. Vasıf

Birimlerin birbirlerinden ayırt edilmesini sağlayan özelliklerine (*characteristic*) **vasıf** (*quality attribute*) denir. Örneğin; bir sınıftaki öğrencilerin her biri bir birimdir. Öğrencilerin boy uzunlukları, yaşları cinsiyetleri, babalarının meslekleri, üniversiteye girişte aldıkları puanlar, bitirdikleri liseler, bir konudaki düşünceleri, kan grupları, IQ puanları, tuttıkları futbol takımı, sevdikleri renkler, okudukları gazete, beğendikleri film türü v.s. birer vasıftır. Birimlerin tanımında da olduğu gibi vasıfların tanımında da çok dikkatli davranmak gerekir. İncelenecek vasfın tanımında tereddütlerin ortaya çıkması yapılan çalışmanın güvenilirliğini zedeleyeceğinden bu konuda gerekli hassasiyetin gösterilmesi gerekmektedir.

Vasıfları farklı şekillerde gruplandırmak olanaklıdır. Ancak bu konudaki en önemli ayırım nitel ve nicel vasıf ayrımıdır. Buradaki esas nokta da vasıfların sayılarla ifade edilip edilememesidir. Sayılarla ifade edilebilen nicel, edilemeyenler niteldir. Ancak vasıfları aldıkları değerler açısından cinsiyet vasfında olduğu gibi bekar, evli, dul, boşanmış şeklinde çok sonuçlu vasıflar diye de sınıflandırmak olanaklıdır.

1.6.4. Gözlem Sonucu ve Şık

Gözlem sonucu yada görünüm ve şık kavramları da istatistikte çok karşılaşılan ve eş anlamlı olarak kullanılan iki kavramdır. Görünüm bir bireyin bir özelliğine ait veri iken, şık çeşitli vasıfların bireylerden bağımsız olarak ortaya çıkış şekilleri yada dereceleridir. Yani görünüm bireye ait bir değer iken, şık özelliğe aittir. Örneğin bir sınıftaki bir öğrencinin boy uzunluğu o öğrenciye ait bir görünümdür. Boy uzunluğu vasfına ait değerler de şık tır. Ali'nin boyu 170 cm dir dendiğinde bu sonuç Ali'nin boyunun görünümüdür. Genel olarak boy uzunluğu vasfına ait 170 cm. ise şıktır.

Bir başka örnek medeni durumu şıkları bekar, evli, dul ve boşanmış olup herkes için geçerli olan değerlerdir. Ali'nin medeni durumu bekardır dendiğinde bu medeni durumun Ali'deki görünümüdür.

1.6.5. Anakütle (Evren-Toplum-Popülasyon)

Üzerinde inceleme veya araştırma yapılacak olayın gözlenebileceği tüm birimlerin yer aldığı topluluktur. Anakütle yığın olay niteliğinde ve aynı cins birimlerin oluşturduğu topluluktur. Bir fabrikanın ürettiği aynı türden ilaçlar anakütleyi oluşturur.

Herhangi bir gözlem yada inceleme kapsamına giren obje yada bireylerin tümüne anakütle ya da kütle denir. Gözlemin amacına bağlı olarak, anakütle küçülebilir yada büyüyebilir. Örneğin bir araştırmacı Türkiye'de 5 yaşında çocukların boy uzunlukları üzerinde bir inceleme yapmak isterse, araştırmacının anakütleyi Türkiye'de 5 yaşındaki çocukların tümünün oluşturduğu gruptur. Öte yandan başka bir araştırmacı Ankara şehrinde yaşamakta olan 60 yaşından büyük kişilerde bir inceleme yapmak ister ve elde edeceği sonuçları Ankara şehrinin dışında kalan 60 yaşındaki kişilere genelleme amacı taşımazsa, bu araştırmacının anakütleyi Ankara şehrinde yaşayan ve 60 yaşından büyük olanların oluşturduğu grup olur.

Belirli bir amaç için anakütle kabul edilen grup başka bir amaç için anakütle olmayabilir. Anakütlenin sınırlarını anakütleyi kimlerin ya da nelerin oluşturduğunu gözlemin amacı ve gözlem sonuçlarının kimlere genelleneceğini belirler. Anakütlenin sınırlarını belirlemek ve anakütlenin kimlerden yada nelerden oluştuğunu ve sayısını saptamak bazen kolay bazen de çok zor hatta olanaksız olabilir. Bu zorluklar özellikle anakütledeki obje ya da deney sayısını saptamada ortaya çıkar. Çünkü çoğu kez belirli bir anakütleye girmesi gereken obje ya da bireyleri teker teker bulup ortaya karmak ve saymak olanaksızdır. Örneğin Türkiye'deki kanserli hastalar Üzerinde inceleme yapmak isteyen bir araştırmacıyı düşünelim. Bu araştırmacı için anakütlenin sınırlarını çizmek, kimlerin anakütleye girip kimlerin anakütleye girmeyeceğini saptamak oldukça kolaydır. Çünkü belirli bir zamanda Türkiye'de yaşayan ve kanserli olan herkes anakütleye dahildir, bunun dışında kalanlar dahil değildir. Fakat Türkiye'deki kanserli hepsini teker teker saptamak ve toplam sayılarını bulmak olanaksızdır. Çünkü kanserli

olduğu bilinenlerin yanında kanserli olup ta bilinmeyen daha birçok kimsenin de bulunduğu bir gerçektir. Türkiye'deki bütün insanları kısa bir zaman içinde muayene edip kanserli olanları doğru bir şekilde saptama olanağı olmadığına göre bu anakütledeki birey sayısını doğru olarak bulmak olanaksızdır.

Anakütledeki obje ya da birey sayısını tam bir doğrulukla saptamak, anakütledeki sınırlarını belirlemek ve anakütlenin kimlerden yada nelerden oluştuğunu belirtmek o kadar önemli olmayabilir, pek çok durumda anakütledeki obje ya da birey sayısını saptama yerine onu belirli yollarla tahmin etmeye çalışırız, incelemelerde de tahmin edilen bu sayıyı kullanırız. Böyle yapmak bir bakıma zorunluluktur. Ayrıca istatistikte bazı nitelikleri bilinen bir anakütledeki obje yada birey sayısını oldukça güvenilir ve geçerli bir şekilde tahminde yararlı olabilecek bazı teknikler geliştirilmiştir.

Pek çok araştırma amaçları için anakütleyi oluşturan obje yada bireylerin tümünü ayrı ayrı gözlemlemek olanaksız olduğu gibi zorunlu da değildir. Geliştirilmiş olan bazı teknik yöntemlerden yararlanarak anakütleden seçilecek daha küçük sayıda bir grubu gözleyip elde ettiğimiz sonuçları anakütleye genelleme olanağı vardır. İstatistik teknik ve yöntemlerinin birçoğu da bu amaçla geliştirilmiştir. Bu tür istatistik teknik ve yöntemlerinin oluşturduğu bu kısma vardamdı istatistik denildiği bundan Önceki bölümde belirtilmişti.

Nüfus, yığın, anakütle (*population universe*) gibi adlarla da ifade edilen anakütle (*population*) incelenen konudaki olası tüm birimlerin oluşturduğu topluluktur. Anakütle birimler topluluğu yerine gözlem sonuçlarının oluşturduğu topluluk diye de tanımlanabilir. Anakütle N sembolü ile gösterilir ve anakütle hacmi yada büyüklüğü (*population size*) şeklinde ifade edilir.

Anakütleleri farklı şekillerde gruplamak olanaklı ise de en önemli ayrım sonlu ve sonsuz anakütle ayrımıdır. Sonlu anakütlerde ilk ve son gözlem sonucunda bilinirken, sonsuz anakütlerde ilk gözlem sonucu bilinirken son gözlem sonucu bilinmez. Bu ayrım sayılabilir sayıda birim içeren ve sayılamayan sayıda birim içeren anakütleler diye de yapılabilir. Örneğin bir sınıftaki öğrenciler bir sonlu anakütle iken, bir fabrikada üretilen ampuller yada Marmara denizindeki balıklar sonsuz anakütledir.

1.6.6. Örneklem (Sample)

Örneklem, anakütleden seçilen ve anakütleye göre daha az sayıda birimden oluşan topluluktur. Örneklem, gözlem sonuçları açısından da anakütlerde ulaşılabilen yada elde edilebilen gözlem sonuçlarını oluşturduğu topluluk şeklinde tanımlanabilir. Örneklem istatistikte n sembolü ile gösterilir ve örneklem hacmi (*sample size*) veya örneklem büyüklüğü diye de ifade edilir. Örneklem ve anakütle hacimleri arasında $n < N$ durumu geçerli olup $n = N$ durumunda örneklem kavramı önemini kaybeder.

Dikkat edilecek olursa anakütle bir konudaki tüm birimleri yada olanaklı gözlem sonuçlarını kapsarken, örneklem ona göre daha az sayıda birim yada gözlem sonucundan meydana gelmektedir. Çünkü bir konudaki olanaklı tüm gözlem

sonuçlarına ulaşmak her zaman söz konusu olamaz. Bir başka ifade ile sonsuz bir anakütlerde tüm birimlere ait gözlem sonuçlarına ulaşamazken, ulaşılabilenlerle yetinip örneklem elde edilir. Örneklem konusunda dikkat edilmesi gereken en önemli nokta örneklemin tarafsız olması ve anakütleyi iyi temsil etmesidir.

Herhangi bir anakütleden belirli bir yolla seçilmiş daha küçük sayıdaki obje ve bireylerin oluşturduğu gruba örneklem denir. Örneklemden edindiğimiz bilgilere dayanarak anakütle hakkında tahminde bulunuruz. Çünkü pek çok durumda asıl amacımız örneklem grubunu tanımlamak değil, anakütleyi tanımak, onunla ilgili sonuçlar çıkararak karar vermektir.

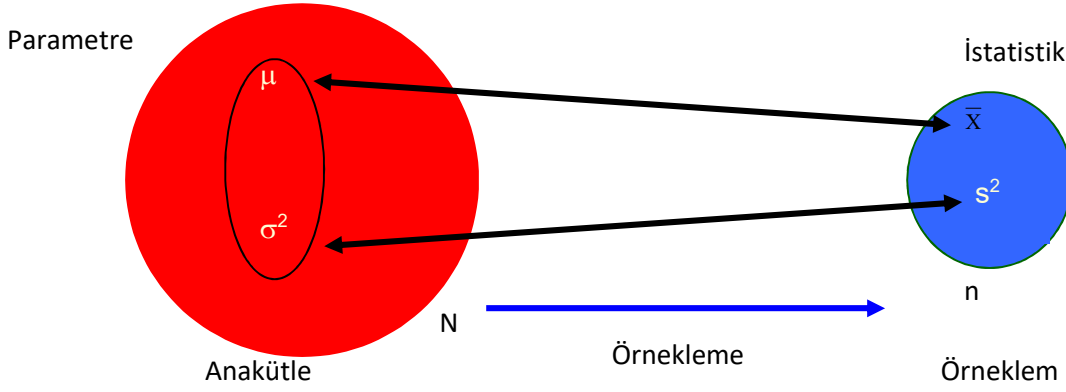
Örneklem grubu üzerinde gözlem sonuçlarını genellerden en az hata ile tahminde bulunmak için örneklemin anakütleyi temsil etmesi temel nitelikleri yansıtması gerekir. Örneklemin anakütleyi temsil etmesi içinde en başta yansız olması gerekir. Herhangi bir örneklem grubu seçildiği anakütleyi belirli bir alt gruba (alt anakütleye) ya da bazı niteliklere sahip olanlara gerçekte olduğundan daha çok ya da daha az ağırlık vermeden temel nitelikleriyle yansıtıyorsa, ya da temsil ediyorsa bu gibi örneklemlere yansız örneklemler denir. Öte yandan seçildiği anakütleyi temel nitelikleriyle tam yansıtmaya, bazı alt gruplara ya da belirli niteliklere taşımaya gerçekte olduğundan daha çok ya da daha az ağırlık veren örneklemlere de yanlı örneklemler denir. Yanlı örneklemlerden elde edilecek bilgiler anakütlerdeki durumu tam yansıtmayacağından yanıltıcı olur.

Bu anakütleden amaca uygun örneklem seçme işine örnekleme denir. Anakütlerden yansız örneklemler seçebilmek için geliştirilmiş çeşitli örnekleme yöntemleri vardır. Bu yöntemler ve uygulaması istatistiğin çok ilginç bir o kadar da karmaşık çalışma alanlarından biridir. Bir örnekleme işleminde araştırmacının amacına anakütlenin yapısına ve olanaklara bağlı olarak bu yöntemlerden bir ya da birkaçı birlikte kullanılabilir.

Parametre: Anakütlenin özelliklerini belirleyen sayısal karakteristiklere parametre adı verilir. Anakütleyi tanımlamada kullanılabilen tipik değerlerdir. Anakütle aritmetik ortalaması, anakütle varyansı gibi değerlerdir.

İstatistik (statistics): Örneği oluşturan birimlerden hesaplanmış, anakütleyi tanımlayan değerlere karşılık gelen değerlerdir. Örnek ortalaması, örnek varyansı gibi.

Örnekleme (sampling): Bir örneklem yardımıyla ilgilenilen anakütleye ilişkin genelleme yapma sürecine örnekleme denir.



Şekil 1.1. Anakütle-örneklem ilişkisi

Şekil 1.1.'de görüldüğü gibi anakütlenin aritmetik ortalaması (μ) ve anakütlenin varyansı (σ^2) parametredir. Anakütledeki toplam gözlem sayısı N ile gösterilir. Örneklem aritmetik ortalaması ($\bar{X}$) ve örneklem varyansı (S^2) ise istatistiktir. Anakütleden seçilen örneklem sayısı ise n kadardır.

Tamsayım : Sonlu bir anakütlenin bütün birimlerinin incelenmesi yada sayılması işlemidir.

Örnek 1.1. 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 1.2. 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 kaçtır? Neden tamsayım yapılmamıştır?

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

Örnek 1.3. Bir fabrikada üretilen bisküvi paketlerinin, planlanan ağırlıkta üretilip üretilmediğinin araştırılması amacıyla, üretilen paketler arasından 200 paket seçilmiştir. Bu araştırmada anakütle ve örneklem nedir? Tamsayımın yapıp yapılmayacağını açıklayınız?

KAYNAKLAR

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