

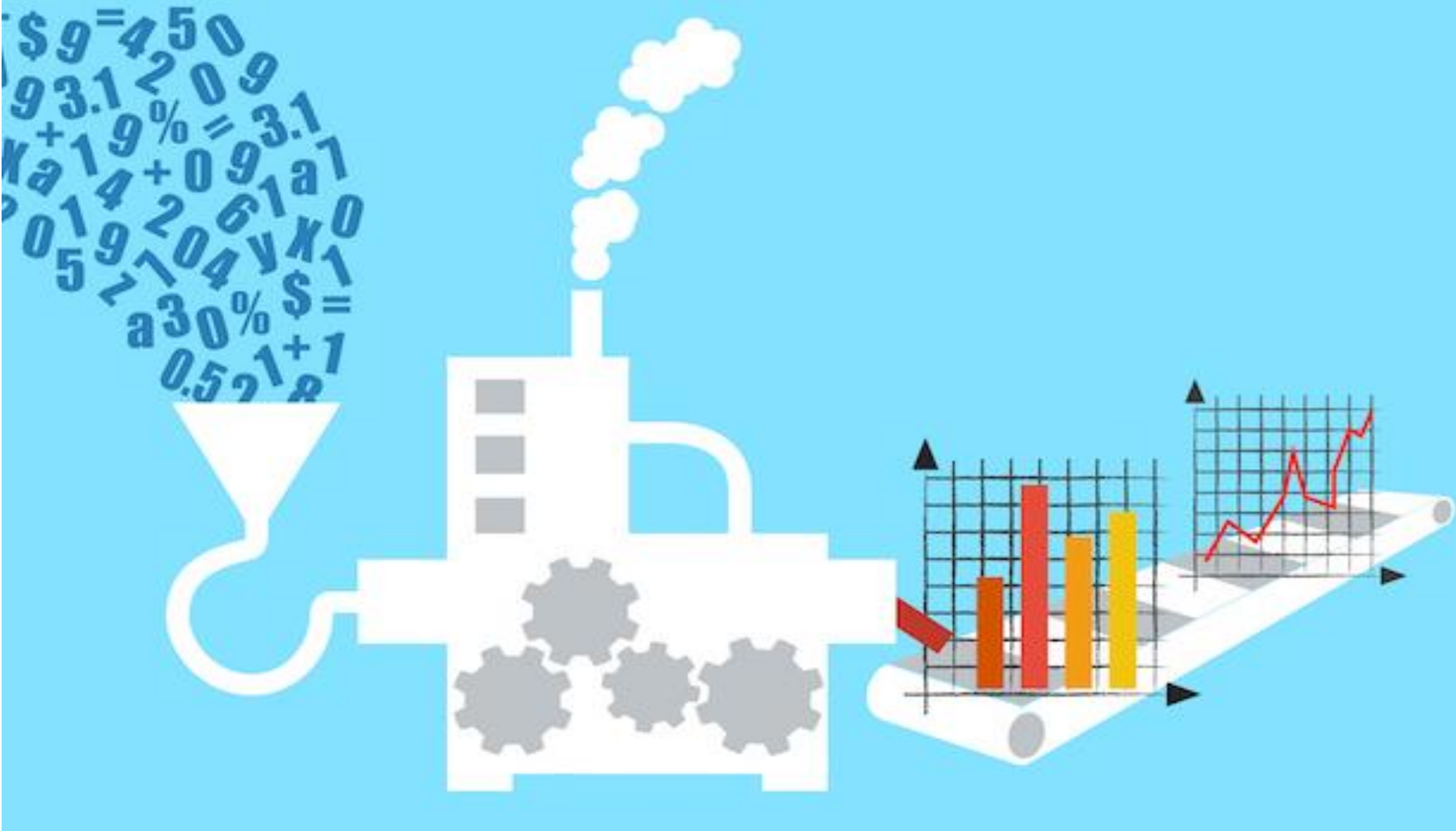
The word cloud features a variety of terms related to research, education, and business. Prominent words include "STATISTICS", "MARKETING", "COGNITION", "ANALYZE", "SEARCH", "DEVELOPMENT", "INNOVATION", "ARGUMENT", "DATA", "VALUE", "CORE", "KNOW", "CONCEPT", "CRITICISM", "DISPUTE", "PROOF", "SUMMARY", "REPORT", "PROCEDURE", "BOOK", "EXPERIMENT", "TEST", "CALCULATION", "REVIEW", "NET", "THESIS", "DEBATE", "CHECK", "DISCUSS", "CONCLUDE", "FINDING", "RESULT", "HYPOTHESIS", "METHOD", "DESIGN", "IMPLEMENTATION", "EVALUATION", "COMPARISON", "ANALYSIS", "SYNTHESIS", "GENERALIZATION", "SPECIFICATION", "DEFINITION", "CLASSIFICATION", "IDENTIFICATION", "DESCRIPTION", "EXPLANATION", "INTERPRETATION", "APPLICATION", "ADAPTATION", "MODIFICATION", "IMPROVEMENT", "ENHANCEMENT", "OPTIMIZATION", "MAXIMIZATION", "MINIMIZATION", "STABILIZATION", "FLUCTUATION", "OSCILLATION", "VARIATION", "TRANSFORMATION", "CONVERSION", "TRANSLATION", "TRANSFER", "TRANSITION", "TRANSFORMATION", "CONVERSION", "TRANSLATION", "TRANSFER", "TRANSITION". The words are interconnected by thin lines, suggesting relationships between different concepts. The overall layout is complex and visually engaging, representing a rich network of ideas.



burcinseyda.corba@omu.edu.tr
burcinseydazorlu@gmail.com

4. Hafta: Genel Kavramlar (Tekrar) ve Merkezi Eğilim Ölçüleri

İSTATİSTİK HAM VERİLERE ANLAM KAZANDIRIR



- Girdi çöp ise çıktı da çöp olacaktır.

İSTATİSTİĞİN KULLANILDIĞI ALANLAR



İSTATİSTİĞİN TANIMI

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

✓ **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, geçmişi ve şimdiki durumu çeşitli sayısal tekniklerle analiz ederek gelecek hakkında karar vermeyi sağlayan bir bilim dalıdır.***

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 evren hakkında tahminlerde bulunma, karşılaştırmalar yapma ve kararlara varma işlemleri çıkarımsal istatistiğin konularıdır.

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.

İSTATİSTİĞİN 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.

İSTATİSTİK METODUNUN AŞAMALARI

i). Bilgilerin Toplanması

Bu aşama araştırmasının konusunun ve birimlerinin kesin tarifi ile başlar. Bilgilerin 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.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.

VERİ VE BİLGİ

Veri (data) 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.

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.

VERİ VE BİLGİ

Veri (data) 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.

TEMEL KAVRAMLAR

BİRİM (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.

Birimler canlı yada cansız varlıklar olabileceği gibi, kurum, kuruluş da olabilir.

Bir olayın birim olabilmesi için kesinlikle ölçülmeye ve sayılmaya elverişli olması gerekir. İnsan, hayvan gibi canlı bir yaratık, bina, ağaç, araba gibi herhangi 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.

TEMEL KAVRAMLAR

Anakütle (Populasyon=Evren=Toplum) :

Üzerinde inceleme veya araştırma yapılacak olayın gözlenebileceği tüm birimlerin yer aldığı 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.

TEMEL KAVRAMLAR

Anakütle (Populasyon=Evren=Toplum) :

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.

TEMEL KAVRAMLAR

Örneklem (sample=örnek):

Örneklem, anakütleden seçilen ve anakütleye göre daha az sayıda birimden oluşan topluluktur. Örneklem, gözlem sonuçları açısından da anakütlede 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 anakütlede tüm birimlere ait gözlem sonuçlarına ulaşılamazken, ulaşılabilenlerle yetinip örneklem elde edilir. Örneklem konusunda dikkat edilmesi gereken en önemli nokta örneklemin tarafsız olması ve anakütleyi iyi temsil etmesidir.

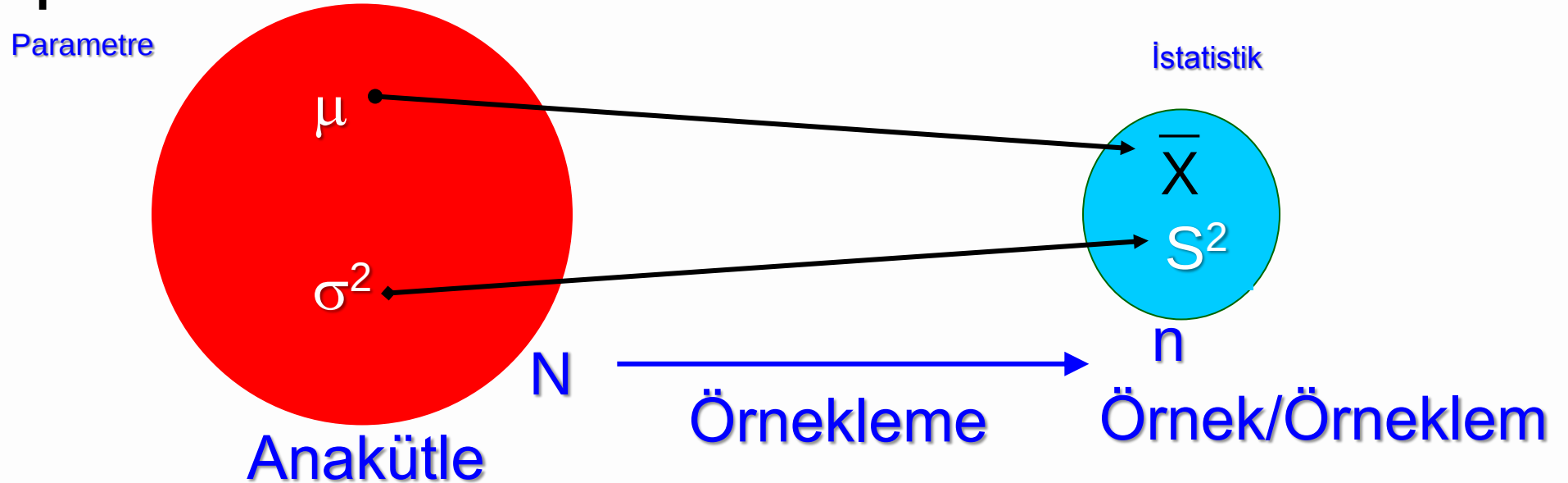
TEMEL KAVRAMLAR

Parametre: Anakütleyi tanımlamada kullanılabilen tipik değerlerdir. Anakütle ortalaması, anakütle varyansı gibi değerlerdir.

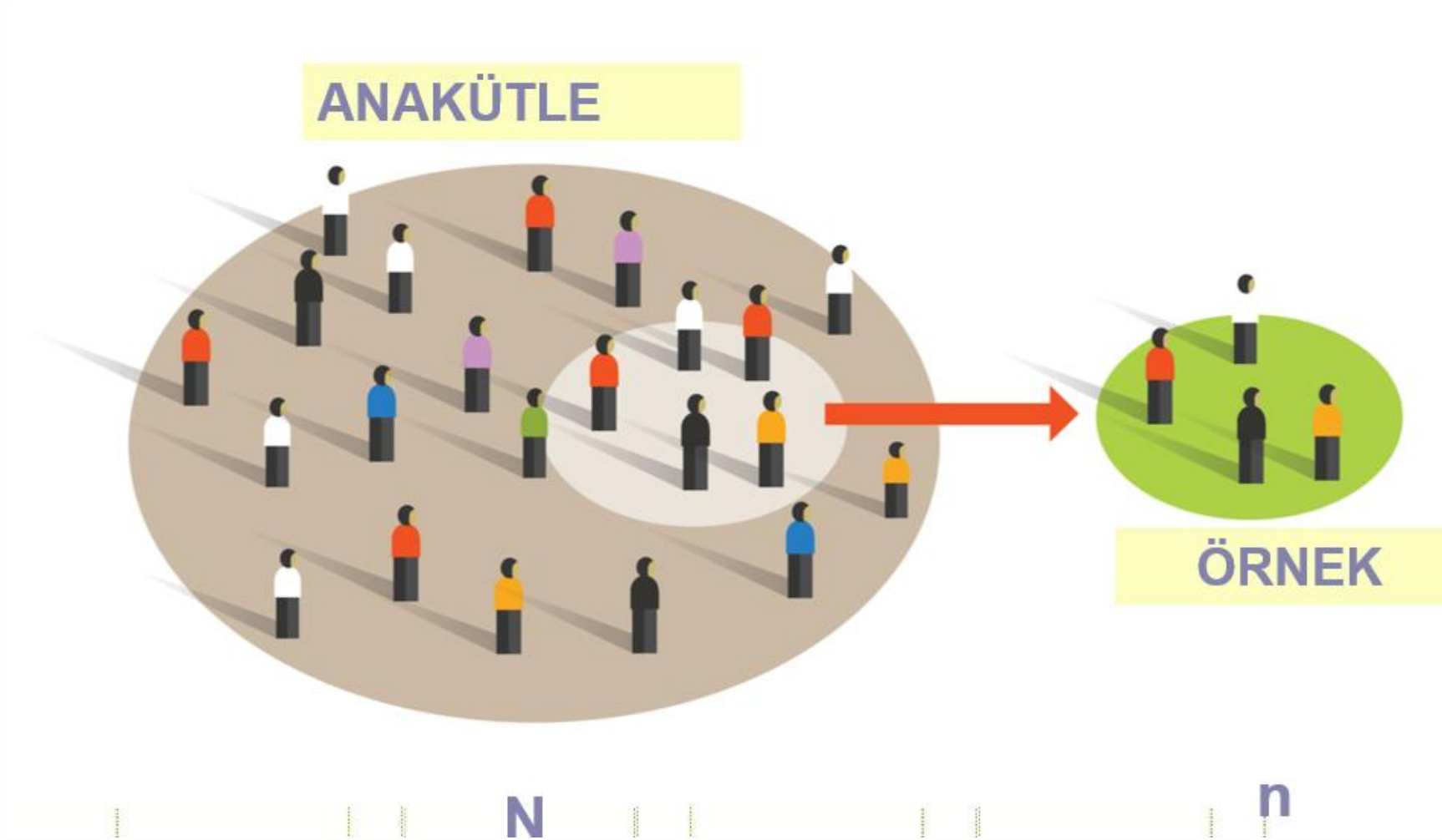
İstatistik: Örneği oluşturan birimlerden hesaplanmış, anakütleyi tanımlayan değerlere karşılık gelen değerlerdir.

Örnek ortalaması, örnek varyansı gibi.

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



TEMEL KAVRAMLAR



TEMEL KAVRAMLAR

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 yapılıp yapılmayacağını açıklayınız?

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

TEMEL KAVRAMLAR

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 kısa zamanda gerçeğe yakın bilgiler kısa zamanda elde edilebilir.
- Örneğe giren birimlerin fiziksel zarara uğramaması:** Birçok durumda gözlemlerin elde edilmesi deneklerin yok edilmesini gerektirebilir. Örneğin bir ilaç üzerinde deneme yapılıyorsa fabrikanın ürettiği tüm ilaçları denemeye almak ve yok etmek mümkün değildir.
- Doğru veri etme:** Küçük sayıda örneklerle çalışılırken daha hassas çalışma yapmak ve daha dikkatli ölçüm almak, daha hassas alet ve yöntemler kullanmak mümkündür. Yapılan işin denetlenmesi de daha kolay olur.

TEMEL KAVRAMLAR

Değişken (variable):

Birimlerin incelenen özelliklerine değişken denir.

Gözlemden gözleme değişik değerler alabilen objelere, özelliklere ya da durumlara değişken denir. Yani her birim için farklı değerler alabilen özelliklerdir.

Değişken kitleyi oluşturan birimlerin ölçülebilen yada sayılabilen özelliklerine denir.

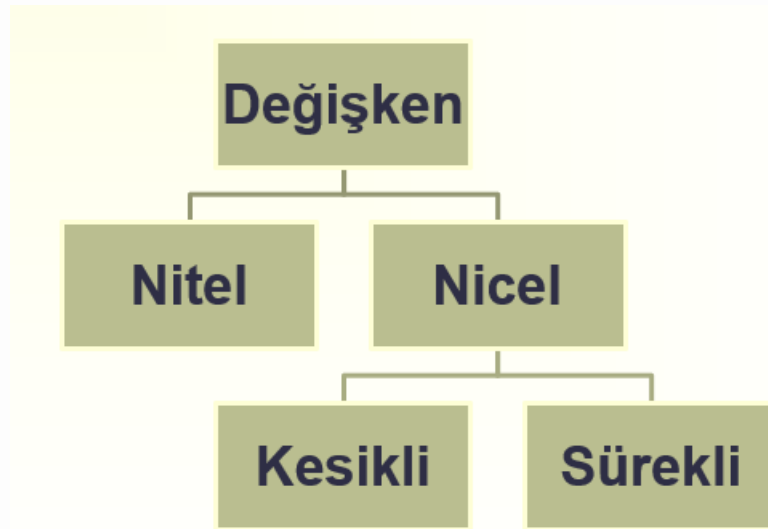
Genelde X,Y,Z gibi alfabenin son harfleri ile ifade edilir.

İstatistik birimlerinin sahip oldukları özelliklere değişken, değişkenlerin aldıkları değerler ise şık adı verilir.

TEMEL KAVRAMLAR

DEĞİŞKENLERİN ÖZELLİKLERİ

Değişkenler gözlenme biçimlerine Nitel ve Nicel değişken olarak ikiye ayrılır. Değişkenler aldıkları değerlere göre ise Kesikli ve Sürekli değişken olarak adlandırılırlar.



Nicel değişkenler aldıkları değerlerin türüne göre sürekli veya kesikli değişkenler olarak isimlendirilirler. Kesikli değişkenler sayılarak elde edilen veriler alırken, sürekli değişkenler ise bazı ölçümler sonunda elde edilen verileri alırlar.

TEMEL KAVRAMLAR

Nitel (Kalitatif) Değişkenler (Qualitative Variables)

Bu değişkenler gözlemden gözleme farklılık gösterirler, ancak bu farklılık derece yönünden değil kalite ve çeşit yönündendir. Bireyin nitelik olarak belirtilebilen özellikleridir. Her bireyin medeni durumu bekar, evli, boşanmış ya da dul kategorilerinden biri olabilir. Bu değişkenlere azlık ya da çokluk durumlarına göre sıraya koyamadığımız için "bir medeni durum, bir milliyet ya da bir din türü ötekilerden daha üstündür ya da daha azdır daha çoktur diyemeyiz. Çünkü bunları gerçek anlamıyla birimlerine ve derecelerine göre sıralamanın bir yolu yoktur. Ölçme düzeyi adlandırma veya sıralama ölçeklidir.

Cinsiyet

1=(Kadın)
2= (Erkek)

Kan Grubu

1- 0 2-A 3-B 4-AB

TEMEL KAVRAMLAR

Sosyal bilimlerde incelenen ölçümlerin büyük bir kısmı nitel türden değişkenler üzerinde yapılan gözlemlerden elde edilir. Nitel veriler sınıflandırırken değişkenin kaç kategorisi (sınıf, alt seçenek) varsa belirlenir ve her bir kategoride kaç birim bulunduğu sayılarak belirlenir.

Nitel verileri el ile sınıflandırırken iki sütunlu bir tablo düzenlenir. Bu tabloda birinci sütuna değişkenin kategorileri yazılır. İkinci sütuna ise veri setindeki her değe ele alınarak bu değerlerin hangi kategoriye girdiği çeteleme yoluyla ait olduğu sınıfa yazılır. Tüm değerler tarandıktan sonra çeteleme değerleri sayılır ve her bir kategoride gözlenen birim sayısını gösteren frekanslar belirlenir.

TEMEL KAVRAMLAR

Nicel (Kantitatif) Değişkenler:

Birimlerin ölçülerek ya da tartılarak değeri saptanan özellikleridir. Bu değişkenlere sayısal değişken adı da veriliyor. Bu tür değişkenler farklı derecelerde az yada çok değerler alan değişkenlerdir. Yaş, ağırlık, boy uzunluğu, yıllık yada aylık gelir, zeka düzeyi, matematik yada tarih bilgisi, havanın sıcaklığı, hava basıncı, hız, nüfus yoğunluğu vb. nicel değişkenlerdir. Ölçme düzeyleri oran veya aralık ölçeğindedir.

TEMEL KAVRAMLAR

Değişkenlerin aldıkları değerlere göre sınıflandırılması :

Kesikli(süreksiz) değişkenler:

Bu değişkenler genelde sayılarak elde edilen değişkenlerdir. Aldıkları değerler sadece sonlu ve tamsayı değerlerdir. Örneğin bir şahsın ağızındaki çürük dişlerin sayısı, mikroskopla bir alanda ölçülen bakteri sayısı, çocuk sayısı, hastanede kalma gün sayısı, galerideki araç sayısı gibi

Sürekli değişkenler:

İki ayrı ölçüm arası kuramsal olarak sonsuz parçaya bölünebilir. Ölçüm söz konusu olduğu için sürekli değişken değerleri her zaman tam değeri vermez.rasyonel sayılar kümesinin elemanları ile belirtilirler. Yaş, uzunluk ve ağırlık gibi. Ölçme tartma yoluyla elde edilen dolayısıyla nokta içermesi mümkün olan verilerdir. Sayı ıskalası üzerindeki her değeri (sonsuz) alabilen değişkenlerdir. Tamsayı veya kesirli sayı olabilir. Boy, ağırlık, sistolik kan basıncı gibi değişkenler bu gruba girer.

TEMEL KAVRAMLAR

Sürekli ve Kesikli Veriler

Özellik (birim)	Değişken	Açıklama
Doğum ağırlığı (kg)	X	Kantitatif özelliktir, sürekli veridir, tartma yoluyla elde edilebilir. 4.72, 5.18, 5.02
100 tohum ağırlığı (g)	Y	Kantitatif özelliktir, sürekli veridir, tartma yoluyla elde edilebilir. 10.8, 11.2, 11.3
Laktasyon (yıllık) süt verimi (kg)	Z	Kantitatif özelliktir, sürekli veridir, tartma yoluyla elde edilebilir. 3572, 4080, ..kg
Başaktaki dane veya tohum sayısı (adet)	T	Kantitatif özelliktir, kesikli veridir, sayma yoluyla elde edilir. 38, 62, 78
Cinsiyet	W	Kalitatif özelliktir, kesikli veridir. 26 erkek, 18 dişi
Puldaki halka sayısı (adet)	Q	Kantitatif özelliktir, kesikli veridir, sayma yoluyla elde edilir. 8, 6, 10, 12

TEMEL KAVRAMLAR

ÖLÇME DÜZEYLERİ

- Ölçme düzeyi bir değişkene atanan değerler arasındaki ilişki ile ilgili bir durumdur. Objelere ve ya bireylere, belirli bir özelliğe sahip oluş derecelerini belirtmek için, belirli kurallara uyarak sembolik değerler verme işlemidir.
- Ölçme gözlem sonuçlarının sayısal sembollerle ifade edilmesidir. Cinsiyet, kilo, sıcaklı birer ölçmedir.
- Ölçmede ölçme konusu olan şey özelliiktir. Bu özellik nesneden nesneye, zamandan zamana değişebilir. Yani bir farklılık mevcuttur. Farklılık ölçme için temeldir.

TEMEL KAVRAMLAR

ÖLÇME DÜZEYLERİ

- Ölçme düzeyi bir değişkene atanan değerler arasındaki ilişki ile ilgili bir durumdur. Objelere ve ya bireylere, belirli bir özelliğe sahip oluş derecelerini belirtmek için, belirli kurallara uyarak sembolik değerler verme işlemidir.
- Ölçme gözlem sonuçlarının sayısal sembollerle ifade edilmesidir. Cinsiyet, kilo, sıcaklı birer ölçmedir.
- Ölçmede ölçme konusu olan şey özelliktir. Bu özellik nesneden nesneye, zamandan zamana değişebilir. Yani bir farklılık mevcuttur. Farklılık ölçme için temeldir.

DÖRT Temel Ölçme Düzeyi Vardır:

1. Adlandırma-Sınıflama (Nominal)
2. Sıralama (Ordinal)
3. Aralık (Interval)
4. Oran (Proportion)

TEMEL KAVRAMLAR

ÖLÇME DÜZEYLERİ

Adlandırma-Sınıflama Ölçeği (Nominal Scale)

Bir sıralaması olmayan kategorileri temsil eden değişkenler sınıflama ölçeğinde tarif edilir. Rakamlar ise farklı bir sınıfı temsil eder. Veriler kodlanabilir fakat sıralanamaz. Yani ölçüm düzeyleri arasında bir sıralama ya da uzaklık-yakınlık gibi belirli bir mesafe yoktur. Aritmetik işlem yapılamaz. Sadece = (eşittir) vardır. Aynı ismi taşıyan fertlerde belirli özelliği taşıma bakımından eşitlik vardır denir.

Bu tip verilerde özellikle parametrik olmayan istatistikler kullanılır. Adlandırma ölçeğine örnek olarak nitel değişkenlerden cinsiyet, bölüm, fakülte, kan grubu, saç rengi verilebilir.

Adlandırma ölçeğinde 4 işlem yapılmaz, çünkü bir anlam taşımaz

TEMEL KAVRAMLAR

ÖLÇME DÜZEYLERİ

Sıralama ölçeği (Ordinal Scale):

Belli bir sıralamaya göre gözlem sonuçlarının sıralanmasıdır. Sınıflar belli bir özelliğe sahip olma bakımından sıralanabiliyorsa, ölçek sıralama ölçeğindedir. Kategorilerin birinin diğerinden ne kadar büyük, ne kadar önemli olduğu bilinemez.

Bu ölçekte bir büyüklük veya önemlilik söz konusudur. Kişilerin eğitim durumunu gösterirken, Eđitimsiz=1, İlkokul=2, Ortaokul=3, Lise=4, Üniversite=5, Yüksek lisans=6 kodları verildiğinde sayı büyüdükçe eğitim düzeyinin arttığı anlaşılır. Ancak 1 ile 2 arasındaki mesafe ile 5 ile 6 arasındaki mesafe aynı değildir. Yani **sıralama ölçeğinde kod olarak kullanılan sayılar arasındaki mesafe önemli değildir. Aynı şekilde oranlarda bir anlam ifade etmez.**

Sıralamada en iyiye büyük sayı verilebileceği gibi küçük sayıda verilebilir, bu kullanılacak analiz yöntemini deđiştirmez. Bu ölçekle elde edilmiş verilerde ortalama yerine medyan kullanılır, daha çok parametrik olmayan testler kullanılır.

TEMEL KAVRAMLAR

ÖLÇME DÜZEYLERİ

Aralık Ölçeği (Interval Scale):

Nesnelerin belli bir başlangıç noktasına göre ve belli bir özelliğe sahip oluş derecesine göre eşit aralıklarla sıralandığı ölçektir. **Aralık ölçeğinde başlangıç noktası keyfi seçilir** ve bu noktadan itibaren belli bir ölçü birimiyle bölümlenerek genişletilir.

Aralık ölçeğinde sıfır noktası mutlak sıfır noktası değildir. İtibari bir sıfır noktasıdır. Sıcaklığın 0°C olması sıcaklığın hiç bulunmaması anlamına gelmez. Bu ölçme düzeyinde toplama, çıkarma yapılabilir. Sayılar arası mesafelerin anlamı önemlidir. **Bu tip sayılar toplanarak ortalama alınabilir. Ancak çarpma ve bölme işlemleri yapılamaz, oranlar bir anlam taşımaz.**

Isıyı ölçen santigrad ölçeği bir aralık ölçeğidir. 0°C başlangıç noktasıdır. Ölçümler arasındaki farklar birbirinin katı olarak ifade edilebilir. 0°C - 18°C arasındaki fark 9°C - 18°C arasındaki farkın iki katıdır. Ancak 18°C ısı 9°C ısının iki katıdır denilemez.

TEMEL KAVRAMLAR

ÖLÇME DÜZEYLERİ

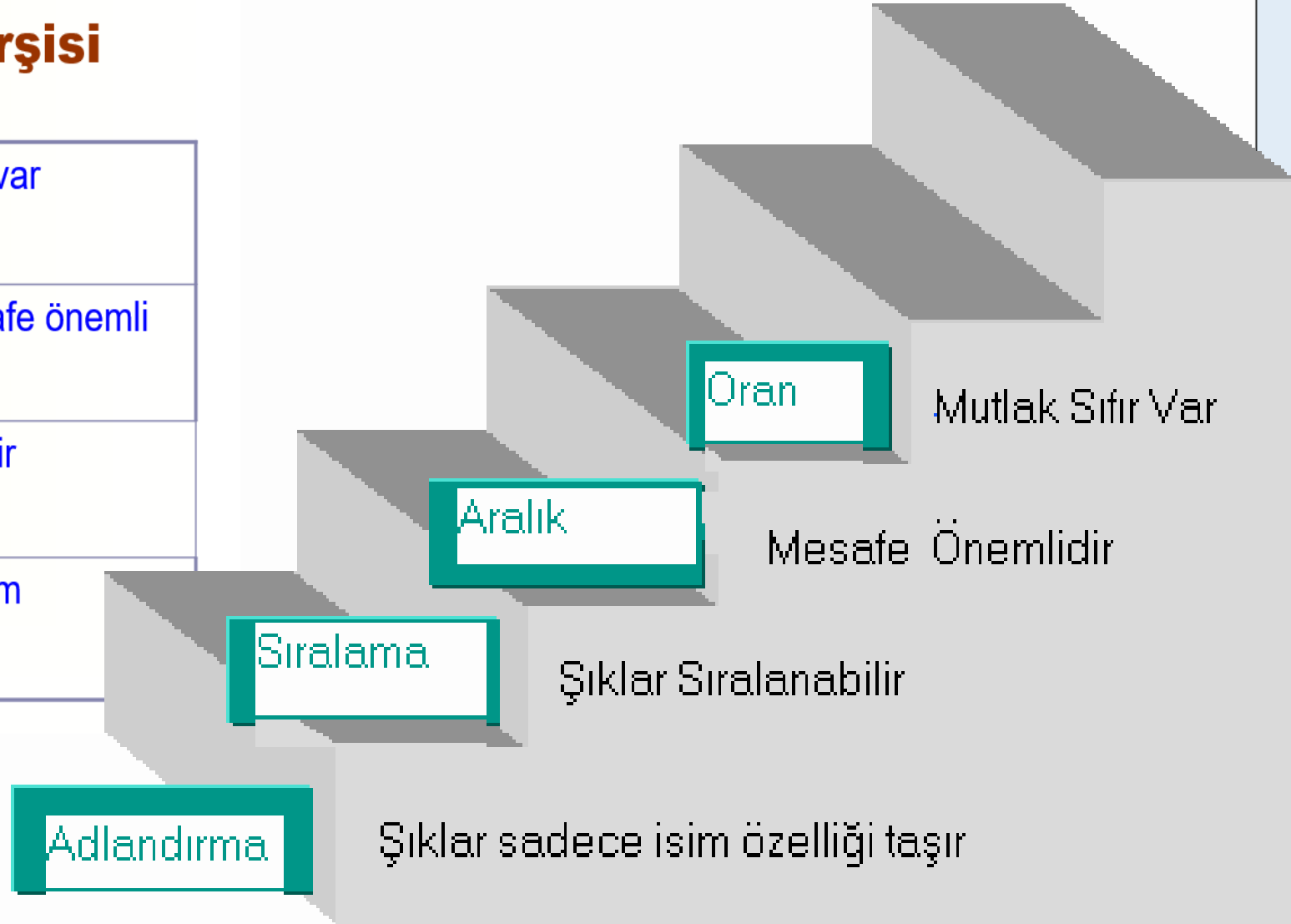
Oran Ölçeği (Ratio Scale):

- Mutlak sıfır noktası olan bir ölçektir. Ölçek üzerindeki noktalar birbirinin katı olarak ifade edilebilir. Tüm matematiksel işlemler kullanılabilir. Metre, kg oran ölçeği için uygundur. Bir şeyin uzunluğu sıfır metre demek ölçülecek uzunluk yok demektir.
- Sayılarak elde edilen değişkenlerin çoğu oran ölçeğindedir. Örneğin; geçen 6 aydaki hasta sayısı nedir dendiğinde, bu sıfır olabilir. Nicel değişkenlerin ölçme düzeyleri oran ölçeklidir. Örneğin boy uzunluğu, kilo, sınavdan alınan notlar, hava sıcaklığı gibi.
- **Tüm parametrik tekniklerin uygulanabildiği ölçek oran ölçeğidir.**

TEMEL KAVRAMLAR

Ölçeklerin Güç Hiyerarşisi

Oran		Mutlak sıfır noktası var (en güçlü ölçek)
Aralık		Değerler arası mesafe önemli
Sıralama		Sonuçlar sıralanabilir
Adlandırma		Sonuçlar sadece isim (en zayıf ölçek)



2.İSTATİSTİK SERİLERİ VE FREKANSLAR

İstatistik serileri gözlem değerlerinin büyüklüklerine göre sıralanmasıyla oluşturulur. İstatistik serileri ilgili yığın olayın kavranması açısından etkin bir araç olup, istatistik analizlere temel oluşturur. Değişik ölçütler temel alınarak istatistik seriyle ilgili farklı sınıflandırmalar yapılabilir.

- **ZAMAN VE MEKAN SERİLERİ**

Eğer gözlem sonuçları yıl, ay, hafta, gün yada saat gibi bir zaman değişkeninin şıklarına göre sıralanırsa oluşturulan seriye **zaman serisi** adı verilir. Yıllara göre ülke nüfusları, günün belli saatlerindeki trafik yoğunluğu, aylara göre ortalama sıcaklık değerleri zaman serisine örnek verilebilir.

- Eğer gözlem sonuçları ülke, bölge, şehir yada köy gibi mekan değişkeninin şıklarına göre sıralanırsa, elde edilen seriye **mekan serisi** adı verilir. Şehirlere göre elektrik tüketimi, bölgelere göre tahıl üretimi bu tür serilere örnek verilebilir.

İSTATİSTİK SERİLERİ VE FREKANSLAR

DAĞILMA SERİLERİ

Gözlem sonuçlarının maddesel bir değişkenin (zaman ve mekan değişkenleri dışındaki değişkenler) şıklarına göre sıralanmasıyla oluşturulan serilere dağılma serileri adı verilir. Dağılma serileri nitel ve nicel dağılma serileri olarak ikiye ayrılır. Nicel dağılma serileri de üçe ayrılır.

- a) Basit Seriler
- b) Frekans Serileri
- c) Gruplandırılmış (Sınıflandırılmış) Seriler

Eğer derlenen veriler ilgilenilen konunun dışında başka bir yönde örneğin gözlem sırasına göre sıralanmış ise, bu sıralamaya **liste** adı verilir.

İSTATİSTİK SERİLERİ VE FREKANSLAR

Örnek 20 öğrencinin istatistik notları aşağıdaki gibidir.

70,60,45,85,72,61,38,90,75,78,38,45,45,72,72,90,72,90,45,70

Eğer liste belirtilen amaçlar doğrultusunda düzenlenirse sonuçlara daha kısa sürede ulaşılır. Böyle bir sıralama (küçükten büyüğe doğru sıralama gibi) sonucu elde edilen seriye **Basit Seri** adı verilir. Verilen notları daha kolay anlaşılması için liste halinde küçükten büyüğe doğru yazalım.

38	38	45	45	45	45	60	61	70	70
72	72	72	72	75	78	85	90	90	90

Verilerin daha kolay anlaşılması için gözlem değerlerinin yanına kaç kez tekrarlandığı kaydedilerek oluşturulan seriye **frekans serisi**, tekrarlara da frekans denir.

Örnek . Örnekdeki basit seriyi frekans serisi biçiminde gösterelim.

Notlar	:38	45	60	61	70	72	75	78	85	90	Toplam
Frekans:	2	4	1	1	2	4	1	1	1	3	20

İSTATİSTİK SERİLERİ VE FREKANSLAR

- Deney yada gözlem sonuçları çok iken, deney yada gözlem sonuçlarının belirli aralıklar (sınıflar) içinde kalan sıklara göre düzenlenmesiyle oluşturulan istatistik serilerine **gruplandırılmış (sınıflandırılmış) seri** denir.

Not Sınıfı	Frekans
31-45	6
46-60	1
61-75	8
76-90	5

- Bir sınıfın alt ve üst sınırları arasındaki farka sınıf aralığı yada sınıf büyüklüğü adı verilir ve h ile gösterilir. Yukarıdaki tabloda 31 alt sınıf, 45 üst sınıf, 31-45 aralığı 15 olup bu değer sınıf aralığıdır.
- Not:** Bir sınıfın üst değeri diğer sınıfın alt değerine eşit verilmişse, alt sınır değeri ilgili sınıfa dahil, üst sınır değeri ise hariç tutulur.

İSTATİSTİK SERİLERİ VE FREKANSLAR

- **Örnek .** Aşağıda verilen frekans dağılımında ilgili sınıflara karşı gelen değişken değerlerini belirleyiniz?

Sınıflar	f	Sınıf Orta Noktası
0-4	4	$(0+4)/2=2$
4-8	10	$(4+8)/2=6$
8-12	17	$(8+12)/2=10$
12-16	25	$(12+16)/2=14$
16-20	14	$(16+20)/2=18$
20-24	6	$(20+24)/2=22$
24-28	4	$(24+28)/2=26$

İSTATİSTİK SERİLERİ VE FREKANSLAR

- **BİRİKİMLİ SERİLER**
- Bir frekans dağılımında, her sınıfın frekansına bir önceki sınıfın frekansı eklenerek oluşturulan seriye birikimli seri, bu tür oluşturulan frekanslara da birikimli frekanslar adı verilir. Eğer birikimli seriler küçükten büyüğe doğru oluşturulmuşsa **–den az**, büyükten küçüğe doğru oluşturulmuşsa **–den çok** olarak isimlendirilir.
- **Örnek** Bir doğum evinde doğan 100 bebeğe ilişkin sınıflandırılmış seriyi ele alarak –den az ve –den çok birikimli serilerini bulunuz?

İSTATİSTİK SERİLERİ VE FREKANSLAR

Ağırlık (kg)	Frekans	-den az	-den çok ↑
1,50-1,75	5	5	95+5=100
1,75-2,00	6	6+5=11	89+6=95
2,00-2,25	10	10+11=21	79+10=89
2,25-2,50	10	10+21=31	69+10=79
2,50-2,75	35	35+31=66	34+35=69
2,75-3,00	15	15+66=81	19+15=34
3,00-3,25	13	13+81=94	6+13=19
3,25-3,50	2	2+94=96	4+2=6
3,50-3,75	4	4+96=100 ↓	4

100

Yukarıdaki tablodan –den az serisi yardımıyla 66 bebeğin ağırlıklarının 2,75kg az olduğu, –den çok serisi yardımıyla da 34 bebeğin 2,75 kg dan fazla olduğu görülebilir.

İSTATİSTİK SERİLERİ VE FREKANSLAR

Örnek Aşağıda verilen seri için –den az ve –den çok serilerini oluşturunuz?

X	5	10	15	20	25	30	Toplam
f	3	5	8	6	3	5	30

X	5	10	15	20	25	30
f	3	5	8	6	3	5
-den az	3	8	16	22	25	30
-den çok	30	27	22	14	8	5

İSTATİSTİK SERİLERİ VE FREKANSLAR

BİLEŞİK SERİLER

Birimlerin birden fazla değişkene göre dağılımlarını bir arada gösteren serilerdir.

Örnek Bir sınıftan rastgele seçilen 5 öğrencinin boy uzunlukları ve ağırlıkları aşağıdaki gibidir.

Öğrenci No	Uzunluk (m) X	Ağırlık (kg) Y
1	1,72	68
2	1,68	70
3	1,80	76
4	1,74	73
5	1,76	71

FREKANS (SIKLIK) TABLOLARI

Verilerin sınıflara göre ve ilgili sınıfın frekansları ile birlikte bir tablo halinde düzenlenmesine frekans dağılımı veya frekans tablosu adı verilir.

Frekans tabloları ele alınan bir değişkene ilişkin ölçüm değerlerinden istenen sınıflara düşen değerlerin sayısını veya frekansını, oran veya % cinsinden belirlenmiş dağılımlarını gösteren tablolarıdır. Frekans tablolarının hazırlanmasında aşağıdaki esaslar dikkate alınmalıdır.

- i. Herhangi bir gözlem değeri ancak bir sınıfta yer almalı, sınıflar çakışmamalıdır.
- ii. Hiçbir gözlem değeri belirlenen sınıfların dışında kalmamalı, tüm gözlemler sınıflarda yer almalıdır.
- iii. Frekans dağılımında kullanılan sınıf aralıkları mümkün olduğunca eşit olmalıdır. Eşit olmayan sınıf aralıkları dağılımların grafiksel olarak gösterilmesinde sorunlar yaratır. Ancak bazen çok sayıda boş sınıfın ayıklanması için eşit olmayan aralıkların kullanılması gerekebilir.

Araştırmacılar kimi durumlarda ele alınan değişkenin sonuçları için frekansları kullanmak yerine frekansların toplam frekansa göre oranlarını kullanma eğilimi taşıyabilirler. Nicel değişkenlerde olduğu gibi nitel değişkenler için oluşturulan serilerde de bu oranlar hesaplanabilir. Oranlara ihtiyaç duyulduğunda, nitel frekans serisine oran sütunu eklenir

Pasta lezzeti	Frekans	Oran
Mükemmel	10	$10 / 50 = 0,20$
İyi	18	$18 / 50 = 0,36$
Orta	12	$12 / 50 = 0,24$
Kötü	6	$6 / 50 = 0,12$
Çok kötü	4	$4 / 50 = 0,08$
Toplam	50	

Ödeme Miktarı	Frekans	Oran
25 - 35	4	$4 / 50 = 0,08$
35 - 45	10	$10 / 50 = 0,20$
45 - 55	11	$11 / 50 = 0,22$
55 - 65	18	$18 / 50 = 0,36$
65 - 75	4	$4 / 50 = 0,08$
75 - 85	3	$3 / 50 = 0,06$
Toplam	50	

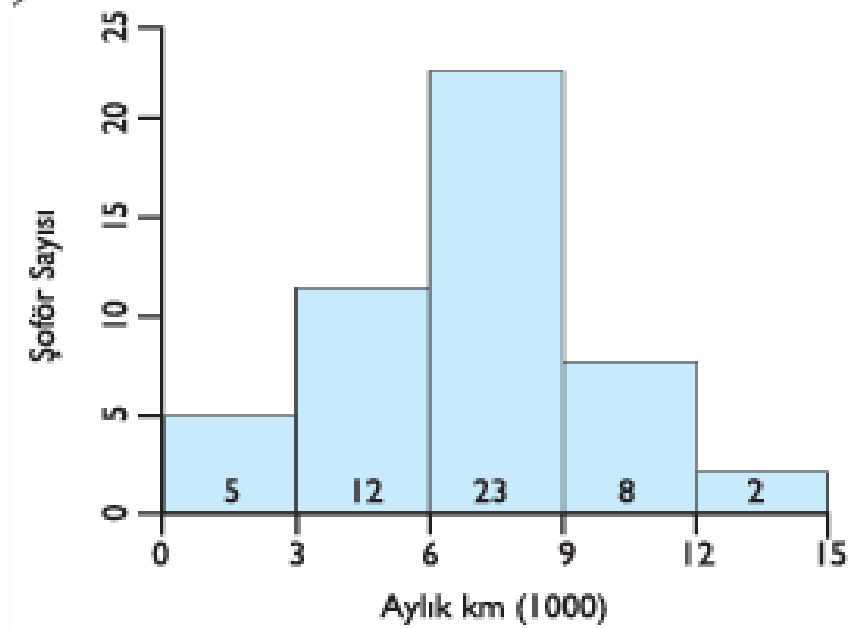
GRAFİKLER

Grafikler gözlem sonuçlarının anlaşılması kolaylaştırır. Çünkü ilk bakışta görsel kolaylık sağlarlar. Ancak grafikler rakamlar derecesinde önemli değildir. Çünkü grafiksiz istatistik olabilir ancak rakamsız grafik asla olamaz. Ayrıca grafikler büyük farkları ve değişimleri gösterebilir. Ancak küçük farklar ve değişimleri gösteremez. Çeşitli grafik türleri vardır.

Nicel Değişkenlerin Grafikselsel Gösterimi Nicel veri çalışmalarında çeşitli deney ya da ölçümler ile elde edilmiş rakamlar ile çalışılır. Nicel verinin grafikselsel gösterimi için sıklıkla kullanılan grafik teknikleri; histogram, frekans poligonu ve birikimli frekans poligonu'dur.

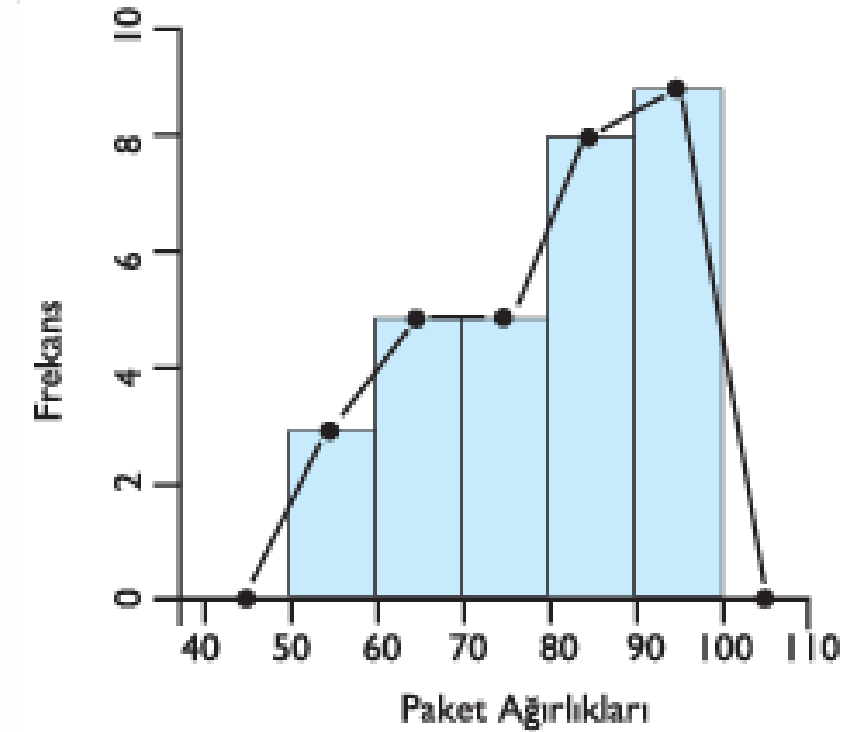
HİSTOGRAM

Aylık km (1000)	Şoför sayısı	Sınıf Aralığı (h)
0 - 3	5	3
3 - 6	12	3
6 - 9	23	3
9 - 12	8	3
12 - 15	2	3
Toplam	50	



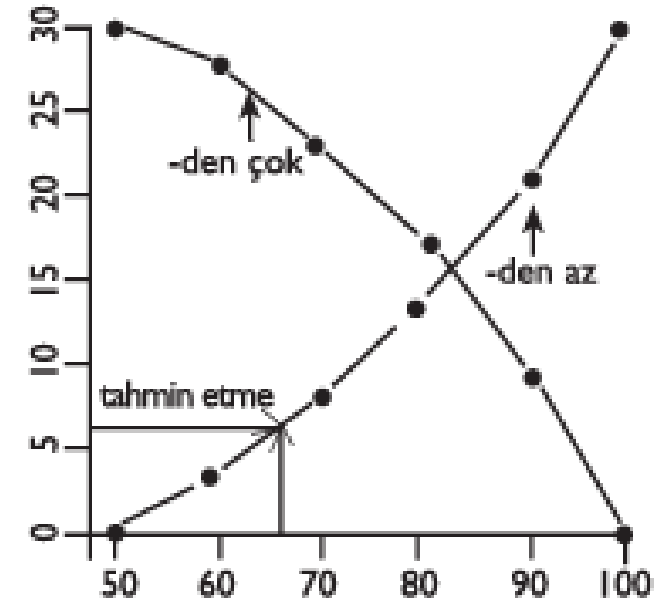
Ağırlık Sınıfları	Frekans	Sınıf Ortası	Koordinat İkili
40 - 50	0	45	(45,0)
50 - 60	3	55	(55, 3)
60 - 70	5	65	(65, 5)
70 - 80	5	75	(75, 5)
80 - 90	8	85	(85, 8)
90 - 100	9	95	(95, 9)
100 - 110	0	105	(105,0)
Toplam	30		

Frekans poligonu



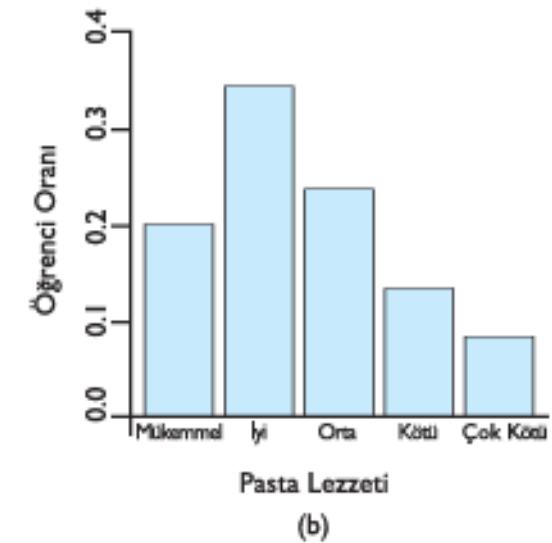
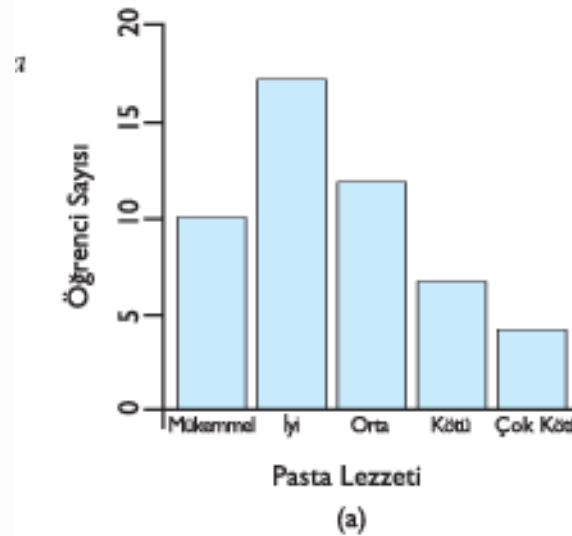
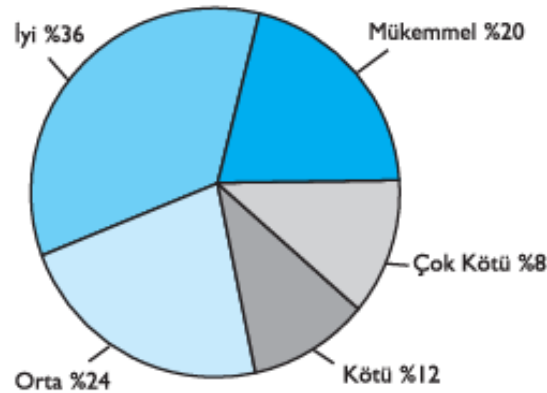
Ağırlık Sınıfları	Frekans	-den az	-den az ikilileri	-den çok	-den çok ikilileri
			(50, 0)		
50 - 60	3	3	(60, 3)	$3 + 27 = 30$	(50, 30)
60 - 70	5	$5 + 3 = 8$	(70, 8)	$5 + 22 = 27$	(60, 27)
70 - 80	5	$5 + 8 = 13$	(80, 13)	$5 + 17 = 22$	(70, 22)
80 - 90	8	$8 + 13 = 21$	(90, 21)	$8 + 9 = 17$	(80, 17)
90 - 100	9	$9 + 21 = 30$	(100, 30)	9	(90, 9)
					(100, 0)
Toplam	30				

Birikimli Frekans poligonu



Nitel Değişkenlerin Grafikselle Gösterimi Sütun ve pasta sıklıkla kullanılmaktadır.

Pasta Lezzeti	Frekans	Oran
Mükemmel	10	$10 / 50 = 0,20$
İyi	18	$18 / 50 = 0,36$
Orta	12	$12 / 50 = 0,24$
Kötü	6	$6 / 50 = 0,12$
Çok Kötü	4	$4 / 50 = 0,08$
Toplam	50	



3MERKEZİ EĞİLİM ÖLÇÜLERİ (ORTALAMALAR)

3.1. Giriş

İstatistikte birçok terimden oluşan bir sayıyı temsil ve ifadeye yeterli olan tek bir rakama ortalama denir. Ortalama aynı zamanda serinin özelliklerini de belirler. Gözlemlerin hangi nokta etrafında toplanmış olduğunu göstermesi gerektiğinden ortalama adı verilmektedir.

“Başı fırında ayakları buzdolabında olan bir adam ortalama olarak rahattır”

Malcolm E.Lines

Bir mahalledeki ortalama gelir düzeyi bir araştırmaya göre yılda 15000€, diğerine göre ise 5000€ çıkmıştır. İki sonuçta aynı kişi ve yerden bulunmuştur. Burada ki hile “ortalama” kelimesidir. Çünkü hangi ortalama olduğu belirtilmemiştir. Eğer büyük bir sayıya ulaşmak istiyorsanız aritmetik ortalama, küçük bir sayıya ulaşmak için en çok tekrarlanan değer olan mod kullanılabilir. Bunun için hangi ortalamanın kullanıldığı ve araştırmanın kimleri kapsadığı sorulmalıdır.

Sayı yığınlarının kolayca anlaşılması için sayı yığınlarının en fazla yığıldığı bölgeyi tarif eden tipik değerlerin verilmesi gerekir. Bu değerler dağılışın merkezini gösterdikleri için merkezi eğilim ölçüleri olarak da bilinir. İstatistikte bir seriyi temsil etmeye yarayan tek bir rakama **ortalama** denir.

Ortalamalar serinin bütün terimlerinin hesaba katıldığı ortalamalar (**duyarlı ortalamalar**) ve serinin bütün terimlerinin hesaba katılmadığı ortalamalar (**duyarlı olmayan ortalamalar**) diye ikiye ayrılır.

3.2 Duyarlı Ortalamalar

Aritmetik ortalama, geometrik ortalama, harmonik ortalama, kareli ortalama

3.2.1. Aritmetik Ortalama

Aritmetik ortalama deneklerin aldıkları değerlerin toplanıp denek sayısına bölünmesiyle elde edilen değerdir. Tablo 3.1.'de $r=1$ alındığında aritmetik ortalama formülleri elde edilir.

Aritmetik Ortalama	Basit Serilerde	$\bar{X} = \frac{\sum X_i}{n}$
	Frekans Serilerde	$\bar{X} = \frac{\sum f_i X_i}{\sum f_i}$
	Gruplanmış Serilerde	$\bar{X} = \frac{\sum f_i m_i}{\sum f_i}$

Örnek 3.1. 20 pnömoni (zatürre) hastası için hastalık süreleri (gün) aşağıdaki şekilde bulunmuştur.

Xi: { 6,7,8,8,10,11,11,11,8,10,10,10,12,12,14,
14,12,7,10,11}

$$\overline{X} = \frac{\sum_{i=1}^N X_i}{n}$$

N=20

Aritmetik Ortalama= 202/20=10.1gün

Örnek 3.2. Aşağıdaki serinin aritmetik ortalamasını bulunuz?

Notlar (X _i)	:	40	60	70	80	100
Frekans(f _i)	:	5	4	5	4	2

$$\overline{X} = \frac{\sum f_i X_i}{\sum f_i} = \frac{(5 \times 40) + (4 \times 60) + (5 \times 70) + (4 \times 80) + (2 \times 100)}{5 + 4 + 5 + 4 + 2} = \frac{1310}{20} = 65,5$$

Örnek 3.3. Aşağıda verilen serinin aritmetik ortalamasını bulunuz?

Sınıf sınırları	Sıklık (frekans= f_i)	Sınıf Değeri (m_i)	$f_i \cdot m_i$
1.45 - 1.95 den az	2	1.7	3,4
1.95 - 2,45 den az	18	2,2	39,6
2,45 – 2,95 den az	24	2,7	64,8
2,95 – 3,45 den az	19	3,2	60,8
3,45 – 3,95 den az	18	3,7	66,6
3,95 – 4,45 den az	9	4,2	37,8
4,45 – 4,95 den az	6	4,7	28,2
4,95 – 5,45 den az	4	5,2	20,8
	n=100	TOPLAM	322

- Frekans tablosunda sınıf orta değerleri (X_i) bulunur. Frekanslarla (f_i) birlikte bu değerler kullanılarak ortalama hesaplanır. Şöyleki:
- k = sınıf sayısı, yani örnek için $k=8$ olur.

$$\bar{X} = \frac{\sum_{i=1}^k f_i m_i}{\sum_{i=1}^k f_i} = \frac{f_1 m_1 + f_1 m_2 + \dots + f_k m_k}{f_1 + f_2 + \dots + f_k}$$

Aritmetik Ortalama= $322/100= 3.22$

Olur.

Soru. Aşağıda hastaların hastanede kalış süreleri verilmiştir. Buna göre ortalama hastanede kalış süresini bulunuz?

Kalış Süresi (Gün)	Frekans (f_i)	Sınıf Orta Noktası (m_i)	$f_i m_i$
1-5	4	3	12
6-10	10	8	80
11-15	17	13	221
16-20	8	18	144
21-25	10	23	230
26-30	4	28	112
31-35	3	33	99
Toplam	56		898

$$\bar{X} = \frac{\sum f_i m_i}{\sum f_i} = \frac{898}{56} = 16.036 \text{ gün}$$

Aritmetik Ortalamanın Fayda ve Sakıncaları

Aritmetik ortalama kavram olarak basittir, hesaplanması kolay olduđu gibi cebirsel işlemlere de elverişlidir. Bu bakımdan en çok kullanılan ortalamadır.

Aritmetik ortalama dağılımdaki terimlerden herhangi birinde meydana gelen kıymet değışikliğinden etkilenir; bu özellik aritmetik ortalama için bir üstünlük olduđu kadar, sakıncalıdır aynı zamanda. Dağılımda terim sayısının az olması durumunda olađanüstü küçük veya büyük terimler aritmetik ortalamanın değeriini etkiler ve simgeleyici olmasını engeller.

Diđer taraftan dağılımın alt ve/veya üst sınırının belirsiz olması durumunda aritmetik ortalamayı hesaplamak olanaksızdır; belirsiz olan sınırlar için yapılacak kestirimler, ortalamanın kesin değeriinin hesaplanılmasına olanak vermeyecektir. Bu bakımdan sözü edilen iki durumda dağılım terimlerini normal büyüklüğünün belirlenmesinde aritmetik ortalama kullanılmamalıdır.

Soru : Aşağıda hastaların hastanede kalış süreleri verilmiştir. Buna göre ortalama hastanede kalış süresini bulunuz?

Kalış Süresi (Gün)	Frekans (f_i)	Sınıf Orta Noktası (m_i)	$f_i m_i$
1-5	4	3	12
6-10	10	8	80
11-15	17	13	221
16-20	8	18	144
21-25	10	23	230
26-30	4	28	112
31-35	3	33	99
Toplam	56		898

$$\bar{X} = \frac{\sum f_i m_i}{\sum f_i} = \frac{898}{56} = 16.036 \text{ gün}$$

3.3. Duyarlı Olmayan Ortalamalar

Duyarlı ortalamalarda bütün terimler veya sınıflar dikkate alınır. Hesaplamalarda bazen serinin bütün terimleri veya sınıfları dikkate alınmayabilir. Bu durumda duyarlı olmayan ortalamalar ortaya çıkar.

3.3.1. Ortanca (Medyan)

Sayılar kümesini iki eşit yarıya bölen değer ortanca değerdir $(n+1)/2$. Gözlem değerleri büyüklük sırasına dizildiğinde tam ortada kalan değer yada iki orta değer aritmetik ortalamasıdır.

Örnek 3.19. $X_i : \{ 3, 1, 13, 27, 6, 8, 6 \}$ gözlem değerlerinin ortancası nedir?

Sayılar büyüklük sırasına dizilirse, $\{ \underline{1}, \underline{3}, \underline{6}, \underline{6}, \underline{8}, \underline{13}, \underline{27} \}$ olur.
Ortada kalan sayı 6 olduğundan
Ortanca=6 olur.

Örnek 3.20. $X_i : \{ 21, 9, 8, 3, 7, 9 \}$ olsun. Gözlemler büyüklük sırasına dizilirse,

$$X_i : \{ \underline{3}, 7, 8, 9, \underline{9}, 21 \}$$

olur. Ortada kalan iki değer 8 ve 9 ortancadır:

$$\text{Ortanca} = (8+9)/2 = 8.5$$

Diğer bir ifade ile $(n+1)/2$ nci değer ortanca değerdir. Gözlem değeri çift ise sonuç şöyle bulunur. $(n+1)/2 = (6+1)/2 = 3.5$, yani 3 ncü ve 4 ncü gözlemlerin ortalaması ortancadır.

Sınıflandırılmış verilerden ortanca hesaplamak için önce medyan sınıfının bulunması gerekir. Bunun içinden az eklemeli frekans (... den az F_i) bulunur ve bu kullanılarak $(n+1)/2$ nci gözlemin düştüğü sınıf ortanca sınıfı olarak tanımlanır.

Örnek 3.21. Aşağıda olan serilerin meydanlarını bulunuz?

A Serisi			B Serisi		
X_i	f_i	$\sum f_i$	X_i	f_i	$\sum f_i$
11	2	2	13	3	3
22	3	5	24	6	9
34	4	9	37	4	13
45	2	11	48	5	18

A serisinde frekans toplamı 11 olup $(n+1)/2=(11+1)/2=6$. terim medyandır. Kümülatif frekans toplamlarında 6. terim (9. terim dahil) 34 değerine sahiptir. O halde A serisi için medyan 34 olur.

B serisinde frekans toplamı 18 olup $(n+1)/2=(18+1)/2=9.5$ terim medyandır. Ancak seride 9.5 terimi olmadığından 9 ve 10. terimlerin ortalaması medyayı verecektir. Seride 9. terim 24 ve 10. terim 37 değerine sahiptir. Dolayısıyla medyan $=(24+37)/2=30.5$ olur.

Gruplanmış Verilerden Ortanca Hesabı

$$\text{Ortanca} = L + \frac{\frac{n}{2} - F_{(i-1)}}{f_i} \cdot C$$

Burada

L: Medyan sınıfının alt sınır değeri

n= Toplam gözlem sayısı

Fi-1 = Medyan sınıfından önceki sınıfların frekans toplamı

fi= Medyan sınıfının kendi sınıf frekansı

C= sınıf genişliği

Kümülatif frekanslarda $n/2$ 'nci terimi içeren sınıf medyan sınıfı kabul edilir. Medyan değeri, medyan sınıfının alt sınırından küçük ve üst sınırından büyük olamaz.

Örnek 3.22. Aşağıda verilen serinin meydanını bulunuz?

Sınıf sınırları	Sıklık (frekans= f_i)	den Az (F_i)
1.45 - 1.95	2	2
1.95 - 2,45	18	20
2.45 – 2.95	24	44
2.95 – 3.45	19	63
3.45 – 3.95	18	81
3.95 – 4.45	9	90
4.45 – 4.95	6	96
4.95 – 5.45	4	100
	100	

Sınıf Sınırı	2.95			3.45
Fi	45.....	50	51.....	63

Medyan Sınıfı
Başlangıcı

Medyan Sınıfı
Sonu

$$Medyan = L + \frac{\frac{n}{2} - F_{(i-1)}}{f_i} \cdot C$$

.

$$\text{Medyan} = L + \frac{\frac{n}{2} - F_{(i-1)}}{f_i} \cdot C$$

L= 2.95, n=100, Fi-1=44, fi=19 olur.

Ortanca= 2.95+ [(100/2) – 44]*0.5/19

= 3.108

Örnek 3.23. Aşağıda serinin meydanını hesaplayınız?

Sınıflar	:4-7 den az	7-10 dan az	10-13 den az
f_i	: 8	5	2
Toplam f_i	: 8	13	15

Toplam frekans $n/2=15/2=7.5$ terim medyandır. Medyan sınıfı 4-7 den az sınıfıdır. Bu sınıfın alt sınırı $L=4$, genişliği $C=3$, medya sınıfının frekansı $n=8$ dir. Medyan sınıfından önceki sınıf seride bulunmadığından bunun kümülatif frekansı $F_{i-1}=0$ kabul edilir. Buna göre medyan aşağıdaki gibi bulunur.

$$Medyan = 4 + \frac{7.5 - 0}{8} \times 3 = 6.81$$

3.3.2. Tepe Değer(Mod)

Bir sayı kümesi içinde en fazla tekrarlanan değer o kümenin tepe değerini oluşturur.

Örnek 3.25. X_i : { 1,2, 6, 3, 7, 3,5, 6,6, 8,9 } serisinin modu nedir?

Burada en fazla tekrarlanan değer 6 olduğu için $\text{Mod} = 6$ olur.

Örnek 3.26. Aşağıdaki serinin modu nedir?

X_i :	2	3	6	7
f_i :	3	6	4	5

Seride en yüksek frekans 6 olduğuna göre, buna karşı gelen değer yani 3 moddur.

Gruplanmış seride mod hesabı için aşağıdaki formül kullanılır.

$$Mod = L + \frac{d_1}{d_1 + d_2} \cdot C$$

d_1 = Mod sınıfı frekansı – bir önceki sınıf frekansı,

d_2 = Mod sınıfı frekansı – bir sonraki sınıf frekansı,

C = Mod sınıfının genişliği

L = Mod sınıfının alt sınırı

Buradan bulunacak mod yaklaşık bir değere sahip olur. Gruplanmış serilerde mod sınıfının belirlemek için, frekans sütunundaki en yüksek frekansa bakılır. En yüksek frekansa sahip sınıf mod sınıfı kabul edilir. Mod değeri, mod sınıfının alt sınırından küçük ve üst sınırından büyük olamaz.

Örnek 3.27. Bir taramada 50 kadının kanındaki (gr/lt) serum albümin değerleri aşağıdaki gibi bulunmuştur.

41 41 42 44 44 36 38 41 42 44 42 39 49 40 45 **32** 34 43
37 39 41 39 48 42 43 33 43 35 32 34 39 35 43 44 47 40
39 42 41 46 37 49 41 39 43 42 47 48 51 **52**

Bu verilere ait sıklık tablosu 6 sınıf olacak şekilde yapılsın.

EnK.=32, EnB.= 52, sınıf aralığı= 4 olsun.

$$N=50, C=4, d_1=(17-14)=3, d_2=(17-7)=10, L=41.5$$

Sınıf Sınırları **Frekans (fi)**

29.5-33.5 den az	3	
33.5-37.5 den az	7	
37.5-41.5 den az	14	
41.5-45.5 den az	17	Mod Sınıfı
45.5-49.5 den az	7	
49.5-53.5 den az	2	
Toplam	50	

$$\text{Mod}=41.5+(17 - 14)*4/[(17-14)+(17-7)] =42.42$$

Tepe değerin (mod) kullanışlı olabilmesi için gözlem sayısının çok fazla olması gerekir. Bazı durumlarda dağılışın birden fazla modu olabilir, çok modlu dağılışlar olabilir. Modların aynı yükseklikte olması da gerekmez. Ancak bu modların sınıf aralarının küçük değışikliğı ile kaybolmayacak ayrıklıkta olması gerekir. Bu durumlarda örneğın farklı gruplardan oluştuğı anlamı çıkar.

Bazen serinin iki maksimum değeri olabilir. Bunun nedeni incelenen kütlenin homojen olmamasından ileri gelir. Örneğın kadın ve erkeklerin boy uzunluklarını gösteren seride iki maksimum nokta vardır. Biri kadınların boy uzunluğu, diğeri de erkeklerin boy uzunluğudur. Burada yapılması gereken kütleyi homojen gruplara ayırmak ve her grup için ayrı mod hesaplamaktır.

Örnek 3.28. Aşağıdaki serinin modunu hesaplayınız?

Sınıflar:	1-3 den az	3-5 den az	5-7 dan az	7-9 den az	9-11 den az	11-13 den az
f_i	:2	7	18	12	18	5

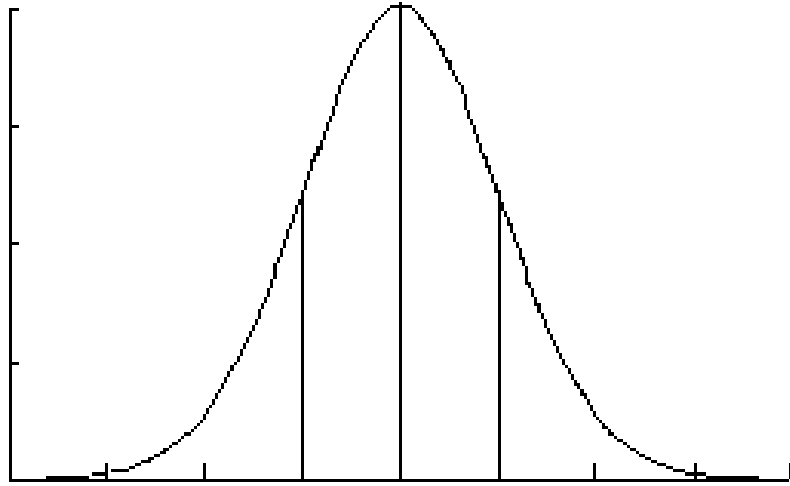
Serinin en yüksek frekansı olan 18 hem 5-7 den az hem de 9-11 den az sınıfına aittir. Bu yüzden çift tepeli serinin sınıfları birleştirilir.

Sınıflar	:1-5 den az	5-9 dan az	9-13 den az
f_i	:9	30	23

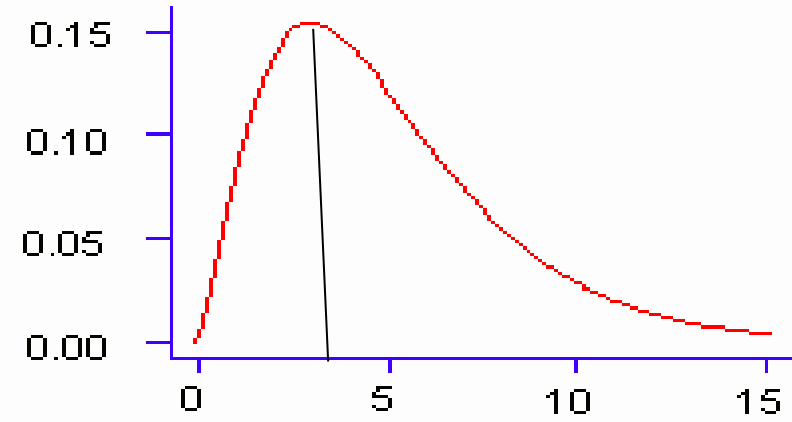
$$C=4 \quad d_1=30-9=21 \quad d_2=30-23=7 \quad L=5$$

$$\text{Mod} = L + \frac{d_1}{d_1 + d_2} \cdot C = 5 + \frac{21}{21 + 7} \times 4 = 8$$

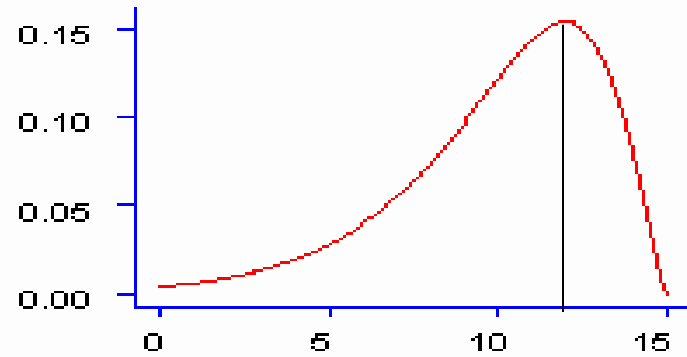
NOT:



AO=Mod=Medyan
Simetrik eğri



Mod<Medyan<AO
Sağa eğik eğri



AO<Medyan<Mod
Sola eğik eğri

3.4. ORTALAMA TİPİNİN SEÇİMİ

- ✓ Ortalama kıyaslama amacıyla hesaplandığında *Aritmetik Ortalama* tercih edilir. Çünkü Aritmetik Ortalama bütün terimler yada sınıflar üzerinden hesaplanan en duyarlı ortalamadır.
- ✓ Araştırmanın amacı seriyi kıyaslamayıp, seriyi temsil etmek ise yerine göre *Mod* yada *Medyan* tercih edilir.
- ✓ Terimlerin kendileri yerine oranları bizi ilgilendiriyorsa *Geometrik Ortalama* tercih edilir.
- ✓ Terimlerin tersleri ile ilgileniliyorsa *Harmonik ortalama* kullanılır.
- ✓ Sıfır veya negatif işaretli değerlere sahip serilerde *Harmonik ve Geometrik Ortalama* hesaplanamaz.
- ✓ Sınıf genişlikleri eşit olamayan gruplanmış serilerde *Medyanın* hesaplanması uygundur.
- ✓ Seri terimleri arasında önem farkı bulunduğunda *Tartılı Ortalama* uygulanır.
- ✓ Ortalama, ortanca ve mod arasında aşağıdaki genel ilişki vardır.
$$\text{Ortalama} - \text{Mod} = 3 * (\text{Ortalama} - \text{Ortanca})$$
- ✓ Sıralamalı ölçümlü özelliklerde veya bütün değerlerin elde edilmesinin uzun zaman aldığı bazı durumlarda *Medyanın* kullanılması uygundur. Örneğin öğrenme davranışının incelendiği bir araştırmada bazı bireyler çok geç öğrenebilir, ortalama için bunu beklemek gerekir, *Medyan* için bunu beklemeye gerek kalmaz.

KAYNAKLAR

- ▶ Prof. Dr. Yüksel Terzi, Temel İstatistik Ders Notları
- ▶ 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 istatistik Yöntemler, Seçkin yayıncılık, Ankara.
- ▶ Çömlekçi, N.(1994). Temel İstatistik, Bilim Teknik Yayınevi, İstanbul.
- ▶ Işık A. (2006). *Uygulamalı İstatistik I-II*, Beta Yayınları, İstanbul.
- ▶ Ağaoğlu, E., Atlas, M. (2013). İstatistik I, Anadolu Üniv. Yayınları, Eskişehir.