

SOS207 SOSYAL BİLİMLERDE İSTATİSTİK I

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 çözümlyerek 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, 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 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 bilimizin çok sınırlı olduğu soruları güvenilir biçimde cevaplandırmanın en kolay yolu, bunlarla ilgili bilgi toplamak ve bilgileri çözümlyerek 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 kütleli 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, çözümlenme 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 anakütlede 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 geliřti. Eski Mısır'da bazı devlet görevlileri bütün aile reislerinin listesini tutuyorlardı. Yine Mısır'da M.Ö. 3000 yıllarında piramit inřaatına gerekli iř gücü talebini garanti etmek için ilk nüfus sayımı yapılmıřtır. Osmanlı devletinde de ilk dönemlerden itibaren istatistiksel bilgilerin toplanmasına önem verilerek, Orhan Bey zamanında çeřitli sayımlar yapılmıřtır.

İstatistiğin gerçek ilerlemesi Yeni Çağ' dan itibaren başlamıřtır. 17.yy.'da Fransa'da ilk defa maliye ve dıř ticaret istatistikleri düzenlenmesine başlanmıřtır. İlk bilimsel nüfus sayıma A.B.D.'de 1790 yılında yapılmıřtır.

İstatistik yöntem bilimi, istatistiksel iřlemlerin uygulamasından çok sonra ortaya çıkmıř ve çeřitli aşamalardan geçerek bugünkü durumunu almıřtır. 17.yy.'ın ilk yarısından itibaren bazı Alman Üniversitelerinde okutulan " Devletlerin Özellikleri "adlı yeni bir derste çeřitli 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 baėlı olarak meydana gelen olaylardır. Olayları meydana getiren gelen ve tesadüfi nedenler vardır. Genel yada temel nedenler her olaya eřit derecede ve tek yönde etki eden nedenlerdir. Tesadüfi nedenler ise her olaya farklı derecede ve çift yönde etki eden nedenlerdir. İřte 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 kiřinin kan grubunu belirlemek için o kiřiden defalarca ve vücudunun farklı bölgelerden kan alarak tahlil etmeye gerek yoktur. Aynı řekilde deniz seviyesinde suyun kaynama derecesini belirlemek için olayı defalarca tekrarlamaya gerek yoktur. Cisimlerin yer çekimi nedeni ile yere düřmelerini belirlemek için olayı defalarca tekrarlamaya veya farklı cisimler için bunu denemeye gerek yoktur. Bu tür olaylarda deėiřkenlik olmadığından benzerlikleri bulup bunları genellemeye çalıřmak söz konusu olmadığı için bunları istatistiğin konusu içine girmezler.

1.4. İstatistiğin Konusu

Yığın Olay:İstatistik diėer bilim dalları gibi olayları konu alır. Olay varsa istatistik vardır. Ancak her olay istatistiėe konu oluřturmaz.

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

Tipik Olay:Eėer bir olaylar kümesinde tek bir olay, tüm olaylar kümesini temsil edebiliyorsa, bu tür olaylara tipik olay denir. Örneğin ideal kořullar 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 anakütlerde 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 çözümlenmede "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 (Röleleler)

Bu aşama araştırmasının konusunun ve birimlerinin kesin tarifi ile başlar. Röleleinin 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 bağılı 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 inceleyen kolektif 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 sürdürdükleri 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 derecelerdir. 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 bekadır 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 anakütlerdeki 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 anakütlerdeki birey sayısını doğru olarak bulmak olanaksızdır.

Anakütlerdeki obje ya da birey sayısını tam bir doğrulukla saptamak, anakütlerdeki sınırlarını belirlemek ve anakütlenin kimlerden yada nelerden oluştuğunu belirtmek o kadar önemli olmayabilir, pek çok durumda anakütlerdeki 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 anakütlerdeki 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 vardamı istatistik denildiği bundan Önceki bölümde belirtilmişti.

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

Anakütleri farklı şekillerde gruplamak olanaklı ise de en önemli ayırım sonlu ve sonsuz anaküt ayırımıdır. Sonlu anakütlerde ilk ve son gözlem sonucunda bilinirken, sonsuz anakütlerde ilk gözlem sonucu bilinirken son gözlem sonucu bilinmez. Bu ayırım sayılabilir sayıda birim içeren ve sayılamayan sayıda birim içeren anakütler diye de yapılabilir. Örneğin bir sınıftaki öğrenciler bir sonlu anaküt iken, bir fabrikada üretilen ampuller yada Marmara denizindeki balıklar sonsuz anakütlerdir.

1.6.6. Örneklem (Sample)

Örneklem, anakütlerden 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 anaküt 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 anaküttelede 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.

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.

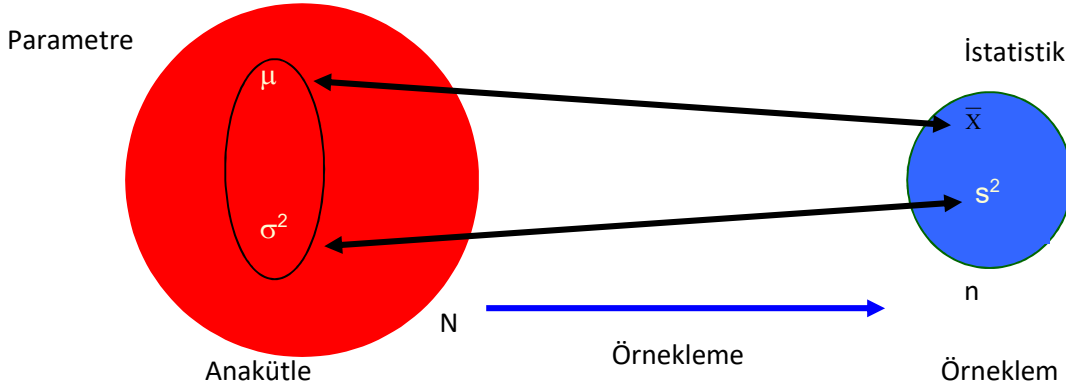
Örneklem grubu üzerinde gözlem sonuçlarını genellerden en az hata ile tahminde bulunmak için örneklemin anakütleyi temsil etmesi temel nitelikleri yansıtması gerekir. Örneklemin anakütleyi temsil etmesi içinde en başta yansız olması gerekir. Herhangi bir örneklem grubu seçildiği anakütleyi belirli bir 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 anakütledeki durumu tam yansıtmayacağından yanıltıcı olur.

Bu anakütleden amaca uygun örneklem seçme işine örnekleme denir. Anakütlelerden 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 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 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.

1.7. Değişken (Variable)

Değişkenin kelime anlamı; değişme özelliği gösteren, çok değişen, değişebilir, kararsız, değişici şeklindedir. Matematiksel tanımı ise; gözlemden gözleme değişik değerler alabilen objelere, özelliklere ya da durumlara "Değişken" denir. Değişken kitleyi oluşturan birimlerin ölçülebilen yada sayılabilen özelliklerine denir. Değişken birimden birime değişen bir sayı ile ifade edilebilen bir bilgidir.

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

Gözlemden gözleme farklı değerler alabilen objelere niteliklere ya da durumlara değişken denir. Gözlemden gözleme farklı değerler alabilme iki durumda ortaya çıkar birinci şekilde bir obje ya da bireyin belirli bir niteliği üzerinde iki ayrı zamanda gözlemlenip bulunup ayrı ayrı değerler gözleyebiliriz. Bir şehirde sıcaklığın sabah ve öğleyin ayrı ayrı ölçülüp farklı sonuçlar bulunması ya da bir çocuğun ağırlığının 6 ay arayla iki kez ölçülüp farklı sonuçlar alınabilmesi gibi. Birinci örnekte değişkenimiz havanın sıcaklığı, ikincisinde ise ağırlıktır. Her iki durumda da gözlemler arasındaki farklar gözlemlenebileceğinden değişken tanımımıza uymaktadır.

İkinci durumda ise tek bir nitelikle ilgili gözlem işini yaklaşık olarak aynı zamanda ve başka obje ya da bireyler üzerinde ya da ortamlarda yapı farklı sonuçlar

alabiliriz. Örneğin günün belirli bir zamanında Türkiye’de 25 şehirde hava sıcaklığını ölçüp farklı sonuçlar gözleyebiliriz ya da aynı yaştaki 15 çocuğu tarttığımız zaman ağırlıklarının az çok değiştiğini görürüz. İnceleme konusu yaptığımız sıcaklık ve ağırlık bu durumda da değişken tanımımıza uymaktadır.

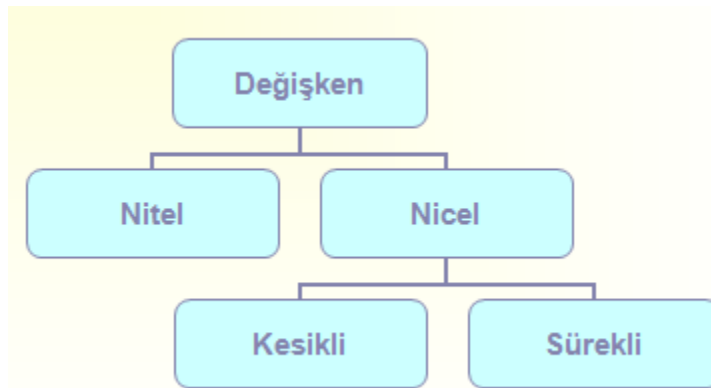
Yukarıda verdiğimiz tanım ve açıklamadan değişkenlerin türleri olabileceği sonucu çıkarılabilir. Aslında değişkenlerin türleri yerine sınıflamasından bahsetmek daha doğrudur. Değişkenler farklı şekillerde sınıflanabilir. Ancak burada her durum için geçerli sayılabilecek bir sınıflama yapma yerine uygulamada sıkça kullanılan bazı deyimler örneklerle açıklanacaktır. Değişkenler X, Y, Z, W gibi harflerle genelde gösterilirler.

İstatistikte **değişken**(*variable*) terimi, deneklere ait özellikler anlamında kullanılır. Örneğin iki gruptan birine A ilacı, diğerine B ilacının verildiği ve deneklerin ilaca verdikleri yanıtın karşılaştırıldığı bir klinik çalışma düşünelim. Bu çalışmada hastaların yaşı, bir değişkendir. Aynı şekilde cinsiyet, boy, ağırlık, kan basıncı, hangi ilaç grubuna girdikleri, tedavi öncesi kan basıncı vb. hepsi birer değişkendir. Pratik anlamda düşünürsek, her hastaya ait veri giriş formundaki her bilgi bir değişkendir.

Eğer bir değişkenin şıkları mekana göre oluşuyorsa bu tür değişkenlere **mekan değişkeni** adı verilir. Örneğin doğum yeri ve üniversitelerin buldukları şehirler mekan değişkenidir. Değişkenin şıkları zamana göre oluşuyorsa bu tür değişkenlere **zaman değişkeni** denir. Örneğin doğum yılları, üniversitelerin kuruluş yılları, aylara göre enflasyon rakamları zaman değişkenidir. Zaman ve mekan değişkenleri dışındaki tüm değişkenlere **maddesel değişken** adı verilir. Örneğin insanların medeni durumu, öğrenci notları, işletmenin maliyetleri maddesel değişkendirler.

Değişkenlerin Özellikleri

Değişkenler gözlenme biçimlerine Nitel ve Nicel değişken olarak ikiye ayrılır.



Şekil 1.2. Değişkenlerin özellikleri

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

1.7.1.Nitel Deęişkenler (Kalitatif-Qualitative Variables)

Bu deęişkenler gözlemden gözleme farklılık gösterirler, ancak bu farklılık derece yönünden deęil kalite ve çeşit yönündendir. Cinsiyet, medeni durum, göz rengi, din vb.

Nitel deęişkenler; birimlerin kalite, kategorik, yada isimsel olarak belirtilebilen karakteristik özelliklerini, durumlarını ve pozisyonlarını belirten deęişkenlerdir. Bu deęişkenlerin verileri isimsel ya da sıralı ölçekle elde edilmişlerdir ve iki yada daha fazla kategoriye (alt seçenek, sınıf, grup) ayrılarak sayımla elde edilir. Bu deęişkenlerin verilerine nitel veriler adı verilir. Örneğin birimlerin, cinsiyeti, kan grubu, medeni durum, göz rengi, mesleęi, yerleşim yeri, tuttuęu futbol takımı (fanatikler için) gibi nitelik bildiren durumları açıklayan deęişkenlerdir.

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

Nitel verileri el ile sınıflandırırken iki sütunlu bir tablo düzenlenir. Bu tabloda birinci sütuna deęişkenin kategorileri yazılır. İkinci sütuna ise veri setindeki her deęer ele alınarak bu deęerlerin hangi kategoriye girdięi çeteleme yoluyla ait olduęu sınıfa yazılır. Tüm deęerler tarandıktan sonra çeteleme deęerleri sayılır ve her bir kategoride gözlenen birim sayısını gösteren frekanslar belirlenir.

1.7.2. Nicel Deęişkenler (Kantitatif-Quantitative Variables)

Birimlerin ölçüm ve tartım sonucu deęerleri saptanan sayısal özelliklerini belirten deęişkenlerdir. Bu deęişkenler deęerleri, mekanik ve elektronik araçlara sayısal olarak aralıklı ölçekli yada orantılı ölçekli verileridir. Nicel deęişkenlerin verilerine nicel veri adı verilir. Deęişik derecelerde az ya da çok deęerler alabilen deęişkendir. Örneğin birimlerin, boy uzunluęu, vücut aęırlıęı, kilosunu, kan basıncı gibi özellikler nicel deęişkenlerdir.

Bu tür deęişkenler farklı derecelerde az yada çok deęerler alan deęişkenlerdir. Yaş, aęırlık, boy uzunluęu, yıllık yada aylık gelir, zeka düzeyi, matematik yada tarih bilgisi, havanın sıcaklıęı, hava basıncı, hız, nüfus yoğunluęu vb. nicel deęişkenlerdir. Bütün insanlar boy uzunluęu ve aęırlıęa sahiptir ancak bunun miktarı kişiden kişiye

ya da bir kiři için zamandan zamana deęiřebilir. Bütün řehirlerde insan yařar sayısı řehirden řehre deęiřir. Bu gibi deęiřkenleri sayabildiđimiz gibi ölçerek derece sırasına koymak ve bir ölçek üzerinde iřaretleme olanađı vardır. Bu tür deęiřkenlerin çođu genellikle normal adını verdiđimiz türden bir dađılım gösterir. Bazen de bazı gerekçelerle bunların normal bir dađılım gösterdiđi ya da göstereceđi kabul edilir.

1.7.3. Kesikli Deęiřkenler

Kesikli deęiřkenler sayılarak elde edilen verileri alırlar. Bu deęiřkenler sonlu ve tamsayı deđerler alırlar. Sınıftaki öğrenci sayısı, acil servisteki hasta sayısı, galerideki araba sayısı vb. kesikli deęiřkenlere örnek verilebilir.

1.7.4. Sürekli Deęiřkenler

Sürekli deęiřkenler bazı ölçümler sonunda elde edilen verilerin deđerlerini alırlar. İki ayrı ölçüm arası kuramsal olarak sonsuz parçaya bölünebilir. Ölçüm söz konusu olduđu için sürekli deęiřken deđerleri her zaman tam deđer vermez, rasyonel sayılar kümesinin elemanları ile belirtilirler. Yař, uzunluk ve ađırlık gibi

Ölçme tartma yoluyla elde edilen dolayısıyla nokta içermesi mümkün olan verilerdir. Örneđin; balıkların ađırlıkları, tohum ađırlıkları, tohum çapı, bitki boyu, ortamdan bakterilerin tüketmiř oldukları řeker miktarı, ineklerin yıllık süt verimleri gibi. Noktadan sonraki hane sürekli veri olduđunu gösterir. Örneđin pamuk tohumu: 1.3; 1.32; 1.32076 gr. Bir ineđin bir senelik süt miktarı 3762 kg tartmaya dayandıđı için süreklidir.

Kalitatif özellikler için sayma yoluyla veri elde edilir. Dolayısıyla bu veriler kesiklidir. Örneđin; dođan yavruların erkek yada diři oldukları (cinsiyet özellikleri) incelendiđinde erkek ve diřilerin sayısı ancak elde edilebilir. Kantitatif özellik için ise hem sayma hem ölçme, hem de tartma yoluyla veri elde edilebilir. Örneđin kök sayısı, balık ađırlıđı, tohum çapı sürekli deęiřkenlerdir.

Sürekli deęiřkenleri, sınıflayarak kesikli deęiřkenlere dönüřtürebiliriz. Örneđin kalsiyum düzeyi sürekli deęiřkendir. Eđer kalsiyum düzeyinin rakam olarak deđerinden çok, "normalden düşük", "normal" ya da "normalden yüksek" olması önemliyse, "8.9'dan düşük", "8.9-10.1 arasında" ve "10.1'den yüksek" olarak yalnız üç deđer alabilen bir kesikli deęiřkene dönüřtürebiliriz. Hastaların yařlarını da "20'den küçük", "20-34", "35-50", "50'den büyük" gibi sınıflandırarak, kesikli deęiřkene dönüřtürebiliriz.

Sürekli deęiřkenleri çok gerekmedikçe, kesikli deęiřkenlere dönüřtürmek uygun deđildir. İstatistik analiz sırasında hataya yol açmaz, ama daha az bilgi veren yöntemler kullanılmasını gerekli kılabilir. Bu nedenle çalıřmalar sırasında verileri

toplarken, deneklere ait sayısal deęişkenleri sınıflandırmadan, gerçek deęerleri ile kaydediniz.

Çizelge 1.1. Sürekli ve Kesikli Veriler

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

1.8. Ölçme Düzeyleri

Ölçme objelere ve ya bireylere, belirli bir özellięe sahip oluş derecelerini belirtmek için, belirli kurallara uyarak sembolik deęerler verme işlemdir. Ölçme gözlem sonuçlarının sayısal sembollerle ifade edilmesidir. Cinsiyet, kilo, sıcaklı birer ölçmedir. Ölçmede ölçme konusu olan şey özelliştir. Bu özellik nesneden nesneye, zamandan zamana deęişebilir. Yani bir farklılık mevcuttur. Farklılık ölçme için temeldir.

Deęişken için elde edilen deęerlerin nasıl yorumlanacağına karar vermede yardımcı olur. Bir ölçüm adlandırma ölçeğinde yapılmışsa buna atfen verilen rakamsal deęerler, sadece kısa gösterim içindir. Yanlış seçilen ölçme düzeyleri istatistik analizi etkiler.

İstatistik analize başlamadan önce ilk yapılacak şey, deęişkenlerin nasıl ölçüldüğünün belirlenmesidir. İstatistikte ölçüm denildięi zaman anlaşılan, deęişkenin alabileceęi deęerlerle ilgili kısıtlamalardır. Örneğin bir kadının gebelik sayısı 5.5 olamaz, ama yaşı 25.64 yıl olabilir. Ya da yeni doğan bir bebeğin APGAR skoru 0 ile 10 arasındaki tüm tam sayılar olabilir, ama 8.4 ya da 15 olamaz. Deęişkenlerin alabileceęi deęerlerin neler olabileceęi, neler olamayacağı, yani nasıl ölçüldüğünü

belirlemek, yapılacak istatistik analizin seçimi için çok önemlidir. Dört farklı ölçme düzeyi vardır.

1.8.1. Adlandırma Ölçme Düzeyi (nominal level)

Bir sıralaması olmayan kategorileri temsil eden değişkenler sınıflama ölçeğinde tarif edilir. Rakamlar ise farklı bir sınıfı temsil eder. Verile 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.

Örneğin, psikiyatride hastalar şizofrenik, paranoid, manyak depresif, psikonörotik gibi isimlendirilir. Bu isimler hastalığın tipine ait sembollerdir. Bu isimlendirme A,B,C,D, veya 1,2,3,4,5 diye de yapılabilirdi. Bu tip verilerde özellikle parametrik olmayan istatistikler kullanılır. Adlandırma ölçeğine örnek olarak nitel değişkenlerden cinsiyet, bölüm, fakülte, kan grubu, saç rengi verilebilir.

1.8.2. Sıralama Ölçme Düzeyi (ordinal level)

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.

Ölçme sonucunda verilen sayısal değerler büyükten küçüğe sıralanabilir. Bir özelliğe sahip oluş derecesidir. Örneğin, yarışma 1.'si 2.'si 3.'sü, birinci tercih, ikinci tercih vb. Bu ölçekte bir büyüklük veya önemlilik söz konusudur. Kişilerin eğitim durumunu gösterirken, Eğitimsiz=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. Sıralamada en iyiye büyük sayı verilebileceği gibi küçük sayıda verilebilir, bu kullanılacak analiz yöntemini değiştirmez.

Örneğin tümörlü hastaların evresinin girildiği "EVRE" adını verdiğimiz bir değişken olsun. Bu değişkene evre 1'den 4'e karşılık gelmek üzere, 1, 2, 3 ve 4 değerleri girilebilir. Bu değerler belirli bir sıra ifade etmektedir. Örneğin "evre 3, evre 2'den daha ileri evredir", "evre 1 en iyi, evre 4 en kötü evredir" vb. Ama değişkenin düzeyleri arasındaki mesafeler belirli değildir. Örneğin matematik işlem yapıldığında $2 - 1 = 4 - 3 = 1$ doğrudur, ama "evre 2, evre 1'den ne kadar ileriye, evre 4 de evre 3'den o kadar ileridir" denemez.

Sıralama ölçme düzeyinde değişken değerleri yalnızca ">" ve "<" işlemleri için sayı gibi değerlendirilir; bunlar dışındaki matematik işlemler uygulanamaz. Sıralama ölçmeyle elde edilmiş verilere genellikle parametrik olmayan teknikler uygulanır. Ayrıca parametrik test varsayımları yerine getirilemiyorsa, hangi ölçmeyle toplanmış olursa olsun parametrik olmayan teknikler tercih edilmelidir. Nominal ve ordinal ölçümle belirtilen değişkenlere **kategorik değişkenler** adı da verilmektedir.

1.8.3. Aralık Ölçme Düzeyi (interval level)

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

Isıyı ölçen santigrad ölçeği bir aralık ölçme düzeyidir. 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 ısıya iki katıdır denilemez.

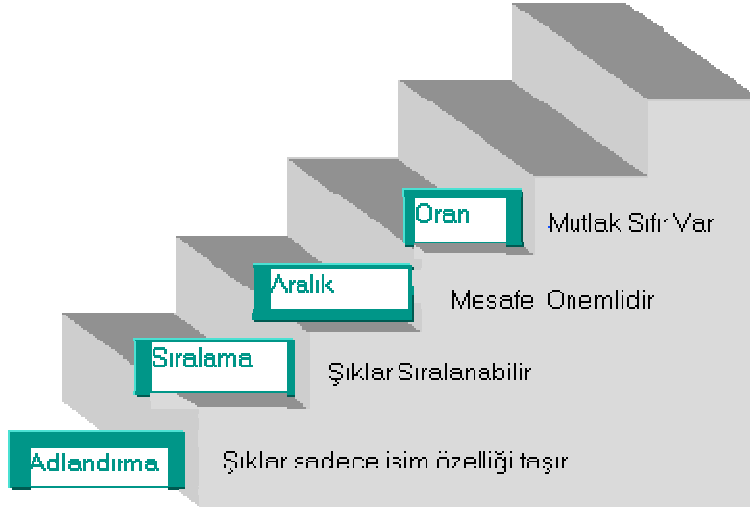
Aralık ölçme düzeyinde sıfır noktası mutlak sıfır noktası değildir. İtibari bir sıfır noktasıdır. Sıcaklığın 0 °C olması sıcaklığın hiç bulunmaması anlamına gelmez. Bu ölçme düzeyinde toplama, çıkarma yapılabilir. Sayılar arası mesafelerin anlamı önemlidir. **Bu tip sayılar toplanarak ortalama alınabilir.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

Üzerinde sıfır noktası 10. noktaya kaydırılan bir metre ile ölçüm yapılırsa, 2m uzunluk 210 cm, 1m uzunluk 110cm olur. Normal bir metrede 2m uzunluk 1m uzunluğun 2 katı olmasına rağmen, sıfır noktası kaydırılarak yapılan ölçümde 210 cm, 110cm'nin 2 katı değildir. İki aralık ölçme düzeyinden birisiyle elde edilen ölçüm diğerine çevrilebilir. °C derece Fahrenheit'a, hicri takvim miladi takvime çevrilebilir.

1.8.4. Oran Ölçme Düzeyi (ratio level)

Mutlak sıfır noktası olan bir ölçektir. Ölçek üzerindeki noktalar birbirinin katı olarak ifade edilebilir. Tüm matematiksel işlemler kullanılabilir. Metre, kg oran ölçme düzeyi için uygundur. Bir şeyin uzunluğu sıfır metre demek ölçülecek uzunluk yok demektir.

Sayılar elde edilen değişkenlerin çoğu oran ölçme düzeyindedir. Örneğin; geçen 6 aydaki hasta sayısı nedir dendiğinde, bu sıfır olabilir. **Nicel değişkenlerin ölçme düzeyleri oran ölçme düzeyindedir.** Örneğin boy uzunluğu, kilo, sınavdan alınan notlar, hava sıcaklığı gibi.



Şekil 1.3. Ölçeklerin Güç Hiyerarşisi

1.9. Oran (Ratio)

İki değişken arasındaki ilişkiyi, birinin diğerine bölünmesiyle ifade eden değişkendir. İncelenen birimler (N) içerisinde istenen özelliğe sahip olanların (n) payını gösterir. Oran aşağıdaki gibi bulunur:

$$Oran = \frac{n}{N}$$

Örnek 1.4. Bir sınıfta bulunan 80 kişiden 20 tanesi mavi gözlü ise bu sınıftaki mavi gözlü öğrencilerin oranı;

$$Oran = 20/80 = 0,25 \text{ olur.}$$

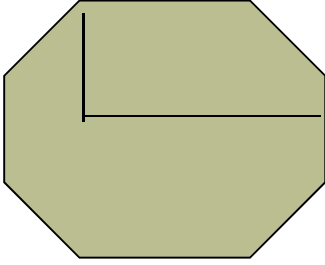
Yüzde : Oranın özel bir halidir. Oran olarak elde edilen değer 100 ile çarpımının ifadesidir.

$$Yüzde = \frac{n}{N} \times 100$$

Örnek 1.5. Yukarıda ki örnekte sınıfta mavi gözlü öğrencileri yüzdesi;

$$Yüzde = \frac{20}{80} \times 100 = \%25$$

Hız (Rate): Olay sayısının maruz kalınan insan-yıla (toplam maruziyet süresi) oranıdır. Hız bir olayın oluş olasılığı olarak tanımlanır. Toplumda hız hesabı yapılırken kişi başına risk altındaki zaman izlenemez. Yıl ortası nüfus, yaklaşık olarak insan-yıl olarak toplam maruziyet süresini verir. Orandan farkı, zaman boyutunu kullanmasıdır. Kesrin payında belirli bir zaman içinde belirli bir toplumda saptanan toplam olgu sayısı, paydada ise bu vakaların içinden geldiği toplumda aynı zaman dilimi içinde hastalık ile karşılaşma ve yakalanma riski olan kişilerin (risk altındaki popülasyon) toplam sayısı verilir. Hız ifadesi geleceğe dair bir kestirim imkanı sunar. Saatte belirli bir kilometre mesafe kateden arabanın bir saat sonra nereye varabileceğini tahmin edebilmemiz gibi, hızı bilinen bir hastalığın gelecek yıl kaç kişiyi etkileyebileceği kestirilebilir.



$$HIZ = \frac{a}{a + b} \times k$$

Hız formülünde k 100, 1000, 100.000 gibi alınabilir.

Örnek 1.6 150 kişilik bir toplumda 70 erkek ve 80 bayan varsa;

Erkeklerin bayanlara oranını

Bu toplumda erkek bulunma hızını bulunuz?

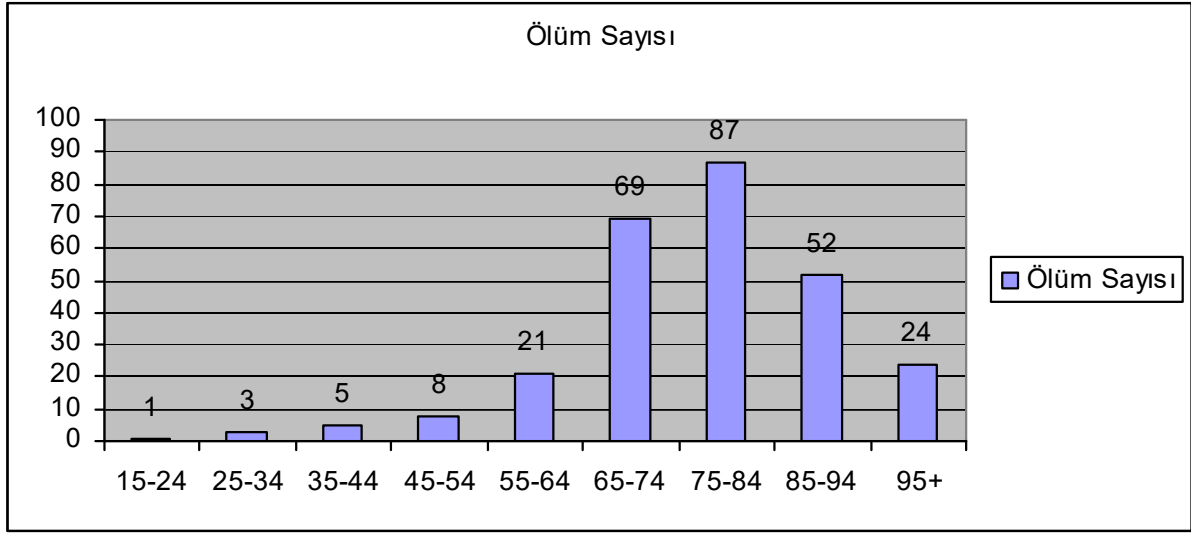
a- Oran= $a/b=70/80=0,875$

b- Hız= $a/(a+b)=70/(70+80)=0,4667$ %46,67 binde 466,7

Örnek 2.7. Bir toplumda 15 yaş üzerindeki gruplarda görülen ölüm sayıları aşağıdaki gibidir.

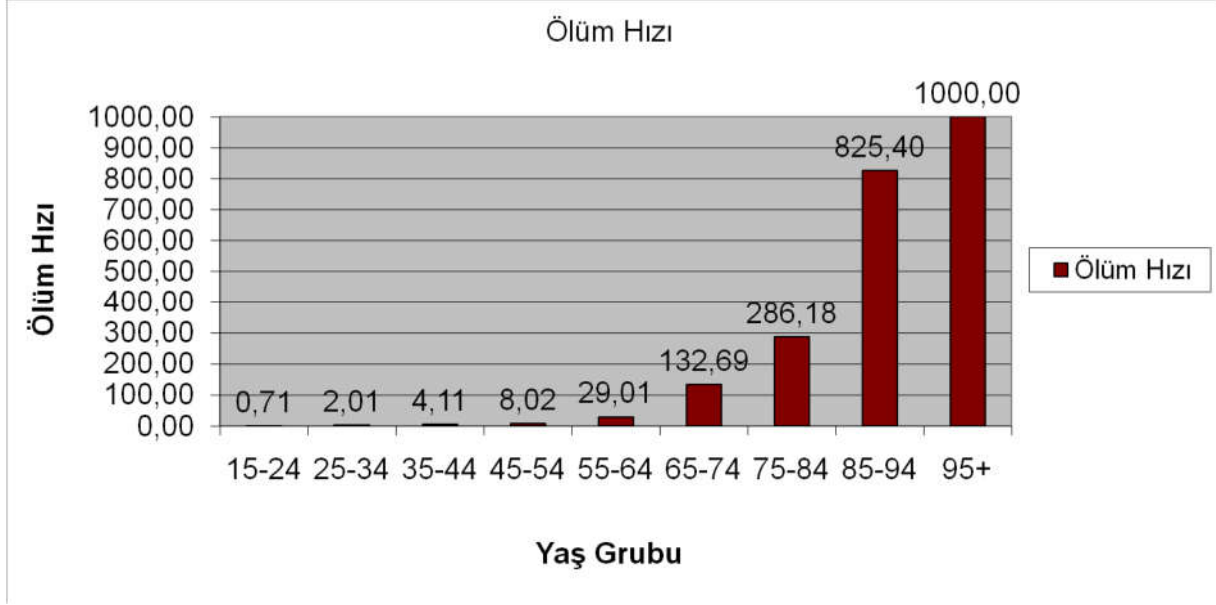
Yaş Grubu	15-24	25-34	35-44	45-54	55-64	65-74	75-84	85-94	95+
Ölüm Sayısı	1	3	5	8	21	69	87	52	24

Tabloda genç yaş gruplarında düşük olan ölüm sayıları yaş ilerledikçe artmakta, belirli bir yaştan sonra ise düşüşe geçmektedir. Veriye göre ölümün genç yaşta az görüldüğü, ileri yaşta fazla görüldüğü ve çok ileri yaşta ise azaldığı söylenebilir.



Yaş Grubu	15-24	25-34	35-44	45-54	55-64	65-74	75-84	85-94	95+
Ölüm Sayısı	1	3	5	8	21	69	87	52	24
Nüfus	1.417	1.496	1.218	998	724	520	304	63	24
Ölüm Hızı (binde)	0,7	2,0	4,1	8,0	29,0	132,7	286,2	825,4	1000,0

Yaş gruplarındaki kişi sayıları farklı verinin hız olarak gösterilmesi gerekir. Yani yaş gruplarındaki ölüm sayıları toplam kişi sayısına bölünerek her yaş grubu için ölüm hızları bulunur. Bulunan değerlere bakıldığında görüldüğü gibi yaş ilerledikçe ölümün arttığı görülmektedir. Yani ölümün yaşla birlikte düzenli olarak arttığı söylenebilir.



1.10. Örnek Problemler

1. Aşağıdakilerden hangisi doğal olmayan birimdir?

- a)Araba b) Öğrenci c) Arsa d) Doğum e)Banknot

2. Aşağıdakilerden hangisi ani birimdir?

- a) Banknot b) Otomobil c) Kitap d) Boykot e) Öğrenci

3. Aşağıdakilerden hangisi kesikli rassal değişkendir?

- a)Sınavda bir sorunun çözülme süresi b) Bir evin fiyatı c)Bir şişe sütun ağırlığı
d) Bir kişinin boy uzunluğu e)Bir ailenin çocuk sayısı

4. Ana kütlelin özelliklerini belirleyen sayısal karakteristiklere ne ad verilir?

- a)Parametre b) İstatistik c) Örnekler d)Tamsayım e)Gözlem birimi

5. Aşağıdaki tabloda verilen seri hangi istatistik serisine örnektir?

İller	Yüzölçümleri(km ²)
Afyon	14.718,63
Ankara	25.401,94
Bursa	10.886,38

- a) Zaman serisi b) Birikimli seri c)Bileşik seri
d) Dağılma serisi e)Mekan serisi

6. Eşit aralıklı zaman noktalarında bir değişken ile ilgili elde edilen gözlem değerlerini zamana göre sıralanmış olarak gösteren serilere ne ad verilir?

- a) Çapraz tablo b) Mekan serisi c) Bölünme serisi
d) Bileşik bölünme serisi e) Zaman serisi

7. Ağıdakilerden hangisi ani birimdir?

- a) Renkler b) okul c) Deprem d) Evlenmelere) Kan grubu

8. Sonlu bir kütleinin tüm birimlerinin sayılması işlemine ne ad verilir?

- a) Parametre b) Kontrol grubu c) Örnekleme d) İstatistik e) Tam sayım

9. Belirlenen amaçlar doğrultusunda hakkında bilgi edinilmek istenen yığının tümüne ne ad verilir?

- a) Topluluk b) Örneklem c) Anakütle d) Örnekleme e) Örnek

10. Aşağıdakilerden hangisi sürekli rassal değişkendir?

- a) Medeni durum b) Bir maddenin ağırlığı c) Fakülteler
d) Bir evdeki çocuk sayısı e) Kan grubu

11. Aşağıdakilerden hangisi kesikli rassal değişkendir?

- a) Sınavda bir sorunun çözülme süresi b) Bir otomobilin fiyatı
c) Bir madenin ağırlığı d) Bir kişinin boy uzunluğu e) Bir galerideki araç sayısı

12. Anakütleden uygun tekniklerle seçilen alt birimlerin topluluğuna ne denir?

- a) Değişken b) Kütle c) Örneklem d) İstatistik e) Birim

13. Cinsiyet (Erkek, Bayan) değişkeninin ölçme düzeyi nedir?

- a) Adlandırma b) Sıralama c) Aralık d) Oran e) Hiçbiri

14. Eğitim düzeyi (ilkokul, Orta, Lise, Üniversite) değişkeninin ölçme düzeyi nedir?

- a) Adlandırma b) Sıralama c) Aralık d) Oran e) Hiçbiri

15. Bir sınava giren 50 öğrenciye ait notların ölçme düzeyi ne olur?

- a) Adlandırma b) Sıralama c) Aralık d) Oran e) Hiçbiri