GÖREVDE YÜKSELME YAZILI SINAVI

8. GRUP ŞUBE MÜDÜRÜ (İSTATİSTİK) DERS NOTU

DEVLET HAVA MEYDANLARI İŞLETMESİ
GENEL MÜDÜRLÜĞÜ

STRATEJİ VE GELİŞTİRME - İSTATİSTİK DERS NOTLARI

İçindekiler

1. TEMEL İSTATİSTİK 5
1.1. İSTATİSTİĞİN TARİHÇESİ 5
1.2. TARİHTEN GÜNÜMÜZE ÖNE ÇIKAN İSTATİSTİKSEL ÇALIŞMALAR....................7
1.3. İSTATİSTİK TANIMLARI....................8
1.4. İSTATİSTİĞİN TEMEL KONULARI 10
1.4.1. Betimsel İstatistik 10
1.4.2. Çıkarımsal İstatistik....................10
1.5. İSTATİSTİĞİN KONUSU 11
1.6. İSTATİSTİK METODUNUN AŞAMALARI....................12
1.7. TEMEL KAVRAMLAR 13
1.7.1. Gözlem 13
1.7.2. Veri 13
1.7.3. Faktör 14
1.7.4. Risk Faktörü 14
1.7.5. Birim 14
1.7.6. Anakütle (Evren-Toplum-Popülasyon)....................14
1.7.7. Örneklem....................16
1.7.8. Parametre....................18
1.7.9. İstatistik (statistics)....................18
1.7.10. Örnekleme (sampling) 18
1.7.11. Tamsayım 19
1.7.12. Oran (Ratio) 20
1.7.13. Yüzde 21
1.7.14. Hız (Rate) 21
1.8. İSTATİSTİKTE DEĞİŞKENLER 23
1.8.1. Nitel Değişkenler 25
1.8.2. Nicel Değişkenler 26
1.8.2.1. Kesikli Değişkenler 26
1.8.2.2. Sürekli Değişkenler 27
1.9. ÖLÇME DÜZEYLERİ 27
1.9.1. Adlandırma Ölçme Düzeyi (Nominal) 28
1.9.2. Sıralama Ölçme Düzeyi (Ordinal)....................29
1.9.3. Aralık Ölçme Düzeyi (İnterval) 29
1.9.4. Oran Ölçme Düzeyi (Ratio) 30
1.10. ORTALAMALAR 31

1.10.1. Duyarlı Ortalamalar 32
1.10.1.1. Aritmetik Ortalama 32
1.10.1.2. Geometrik Ortalama 37
1.10.1.3. Harmonik Ortalama 38
1.10.1.4. Tartılı (Ağırlıklı) Ortalama 41
1.10.2. Duyarlı Olmayan Ortalamalar 42
1.10.2.1. Medyan (Ortanca) 42
1.10.2.2. Mod (Tepe Değeri) 44
1.11. DEĞİŞİM ÖLÇÜLERİ 47
1.11.1. Değişim Genişliği (Range) 48
1.11.2. Çeyrek Sapma 49
1.11.3. Ortalama Sapma 50
1.11.4. Standart Sapma (Standard Deviation) 50
1.11.5. Varyans 53
1.11.6. Standart Hata 54
1.11.7. Değişim (Varyasyon) Katsayısı 54
1.11.8. Momentler 55
1.12. İSTATİSTİKSEL DAĞILIMLAR 57
1.12.1. Kesikli Dağılımlar 57
1.12.2. Sürekli Dağılımlar 61
1.13. KORELASYON 69
1.14. REGRESYON ANALİZİ 70
1.15. ZAMAN SERİLERİ 72
1.15.1. Zaman Serisi Verileri 72
1.15.2. Zaman Serisi Verilerinin Türleri 72
1.15.3. Zaman Serisi Verilerinin Bileşenleri 73
1.15.4. Zaman Serisi Analizinin Amaçları 78
1.15.5. Durağanlık Kavramının Açıklanması 79
2. **İSTATİSTİK ŞUBESİ İLE İLGİLİ BİLGİLER** 81
2.1. Görev ve Sorumluluklar 81
2.1.1. İSTATİSTİK ŞUBE MÜDÜRLÜĞÜ **ŞUBE MÜDÜRÜ** YETERLİLİK KRİTERLERİ, GÖREV, YETKİ VE SORUMLULUKLARI: 81
2.1.2. İSTATİSTİK ŞUBE MÜDÜRLÜĞÜ **ŞEF** YETERLİLİK KRİTERLERİ, GÖREV, YETKİ VE SORUMLULUKLARI: 83
2.1.3. İSTATİSTİK ŞUBE MÜDÜRLÜĞÜ **İSTATİSTİKÇİ/ENDÜSTRİ MÜHENDİSİ** YETERLİLİK KRİTERLERİ, GÖREV, YETKİ VE SORUMLULUKLARI: 84

2.3.1. İSTATİSTİK ŞUBE MÜDÜRLÜĞÜ **BİLGİSAYAR İŞLETMENİ/MEMUR** YETERLİLİK KRİTERLERİ, GÖREV, YETKİ VE SORUMLULUKLARI: 86
2.4. HAVAYOLU UÇAK VE YOLCU TRAFİĞİ İSTATİSTİKLERİ VE TAHMİNLERİ ÜRETİM SÜRECİ 87
2.5. Uçuş Amaçları 88
2.6. Havalimanları ICAO ve IATA Kodları 90
2.7. Anatablo Değişkenleri 91

1. TEMEL İSTATİSTİK

Bu bölümde temel istatistiksel yöntemler ele alınacaktır.

1.1. İSTATİSTİĞİN TARİHÇESİ

İstatistik kelimesi Almanya'da Achenwall tarafından 1748'de kullanılmaya başlanmışsa da, istatistik çalışmaları çok eski çağlara kadar uzanır. Tacitus, imparatorluk zenginlikleri üzerinde Augustus'un geniş bir anket yaptığını yazar. Böylece askerler, gemiler her çeşit gelir kaynakları sayılmış ve kamu gelirleri tespit edilmiştir.

Bütün Ortaçağ boyunca ve 17. Yy'e kadar istatistik sadece tasviri olarak kalmıştır. Bu alanda iki okul ortaya çıktı: Göttingen Üniversitesi profesörü ve Almanlar tarafından istatistiğin babası olarak bilinen Achenwall'in temsil ettiği tasviri okul ve bazı sosyal olayların yaklaşık düzenliliğinden öngörüler ve kanunlar çıkarmayı deneyen siyasi matematikçiler okulu. 18. Yy'de Fransız Desparcieux ve İsviçreli Wargentin toplum olaylarını öngörmenin pratik önemini gösteren ve çok gelişmekte olan sigorta sanayinin hareket noktasını teşkil eden ilk ölüm tablolarını (mortalite tablosu) düzenlemişlerdi.

Fransa'da ilk resmi istatistik bürosu 17. Yy'in sonlarında kurulmuş ve içişleri bakanı Chaptal 1801'de genel nüfus sayımını gerçekleştirmişti. Türkiye'de ilk resmi istatistik bürosu 1933'te kuruldu.

Jacques Bernoulli ve özellikle Laplace, tasviri sayısal bilgilere matematik bilgileri ilave ederek sonuçların değerlendirilmesinde olasılık hesap imkânlarının kullanılmasını araştırdılar. Birincisi ünlü "büyük sayılar kanunu"nu ortaya attı, ikincisi ise"analitik olasılık teorisi" adlı eserinde karmaşık sebepli doğal olayların incelenmesinde olasılık hesaplarından sağlanacak faydaları açıkça belirtti. Quotelet, çevresinde bütün diğer insanların yer alacağı hayali bir "ortalama insan" bulmak maksadıyla bu yöntemin uygulanışını insanların fiziki, ahlaki ve düşünsel özelliklerine uyguladı.

"Quotelet bütün sosyal olayları birey etrafında toplamıştır. Daha ileri giderek istatistiği bütün sosyal alanlara hatta ahlaki olaylara da yaymıştır. İlk "sosyal kriminolog" olarak adlandırılmaktadır. 1836 yılında Quetelet, "Toplum suçu hazırlar, Suçlu ise ancak bir araçtır" demiştir. Bugün Quotelet modern istatistiğin kurucusu olarak kabul edilmekle birlikte, ileri sürdüğü sistem ve vardığı sonuçlar reddedilmektedir.

Quotelet'ten sonraki gelişmeler istatistik yöntembiliminde matematiğin daha yaygın bir şekilde uygulanmasını sağlamış ve matematiksel istatistik disiplininin meydana çıkmasına yol açmıştır.

İstatistik Türklere de yabancı değildir. Selçuklular döneminde İran'da İlhanlılar döneminde Hindistan'da nüfus sayımları düzenlenmişti. Öte yandan Osmanlı İmparatorluğu'nun yönetim sistemi nüfus ve arazi hakkında düzenli biçimde bilgi toplamasını gerektiriyordu.

Bu nedenle imparatorluk zamanında 30-40 yıl gibi aralıklarla nüfus sayımları ve arazi yazımları yapılmaktaydı. Bu istatistiksel işlemler sırf vergi toplanması amacıyla yapıldığı için sadece vergi mükellefleri göz önünde bulunduruluyordu. İmparatorluğun gerileme devrinde bu sayımların arkası kesilmiş, modern tekniklere dayanan istatistikler de 20. Yy'e kadar ülkeye girememiştir. Gerçi bu gibi istatistiklerin düzenlenmesi yolunda 19. Yy'de bazı çalışmalar yapılmıştır.

Örneğin orduyu modernleştirmek amacıyla askere alınabilecek erkeklerin sayısı öğrenilmek istenmiş ve böylece ilk sayım 1831'de gerçekleştirilmiştir. Ancak bu sayım yalnız erkekleri dikkate aldığı gibi ülkenin bütünü de kapsamamıştır. Nitekim sayım yapılan yerler Rumeli ve Anadolu sancakları ve kasabalarıdır. Dolayısıyla bu sayım bugünkü anlamıyla gerçek bir nüfus sayımı olmaktan uzak kalmıştır. 1844'te yapılan ikinci nüfus sayımının amacı ise kimlik belgesi verilecek vatandaşların belirlenmesi idi. Bu nedenle kadınlar da erkeklerle birlikte sayılmıştır. Ne var ki askerlik korkusuyla birçok erkek sayım dışında kalmıştır. 19. Yy'de ülke nüfusu hakkında bilgi verecek sonucu sayım 1874'te başlamış ve araya giren Rus savaşı nedeniyle 1884'e kadar sürmüştür. Nüfus kütüklerini düzenleme ve halka nüfus tezkeresi verme amaçlarıyla başlatılan bu sayımın uygulanmasında da bazı aksaklıklar görülmüştür. Birçok teknik hata içeren ve kısmi nitelikteki bu sayımlar yanında ölüm ve doğumlar gibi nüfus hareketlerini suçları ithalat ve ihracatı belirlemek amacıyla çalışmalar yapılmıştır. Ne var ki bu çabalar da sonuçsuz kalmıştır."

1.2. TARİHTEN GÜNÜMÜZE ÖNE ÇIKAN İSTATİSTİKSEL ÇALIŞMALAR

- Antikçağda Çin'de, Mısır'da, Yunanistan'da nüfus ve bu nüfusun maddi yaşam koşullarına ilişkin sayısal verilere sahip olma gereksinimi izlerine rastlanmıştır.

- Roma'da düzenli nüfus sayımlarına rastlanır. Roma'da önce 5 yıl, sonra da 10 yıl ve 15 yıl gibi düzenli aralıklarla yapılan ve "census" adı verilen bu sayımlar 600 yıl kadar sürmüştür. Her Romalı kendisinin ve babasının adını, servetini, arazisini, kölelerinin sayısını bildirmek zorundaydı.

- 1662'de Graunt erkek çocuklarının doğumlarının kız çocuklarının doğumlarına oranı gibi değişmezlikleri ortaya çıkardı.

- 1742’de Edmond Halley çağdaş aktüer (hayat sigortacılığı) çalışmalarının temelini oluşturan bir hayat tablosu yaptı.

- 1767’de Süssmilch erkek doğum oranı ve 20 yaşına dek erkek çocuk oranındaki gelişme üzerine önemli çalışmalar yayınladı.

- Pierre Simon De Laplace (1749-1827) “Çözümsel Olasılık Kuramı” (1812) adlı eserinde, nedenlerinin tümünü bilip tek tek çözümlenemeyecek kadar karmaşık olan doğal olayların incelenmesinde bu kuramdan sağlanacak yararları ortaya koydu. Önceleri betimsel amaçlar taşıyan istatistik, 17-18 ve 19. yy’lar boyunca Bernoulli, Gauss, Laplace, Moivre gibi ünlü matematikçilerin olasılık teorisini geliştirmeleriyle çıkarımsal istatistiğin temellerini atmışlardır.

- Adolphe Quotelet (1796-1874) yöntemin uygulama alanını, canlı varlıkların antropometrik, psikolojik ve toplumsal açıdan incelenmelerini sağlamıştır. Onun girişimiyle 1885’te Londra’da kurulan Uluslararası İstatistik Enstitüsü’nün öncüsü olan ilk uluslararası istatistik kongresi 1853’te Brüksel’de toplanmıştır.

- Karl Pearson (1857-1936) Biyoistatistik’i buldu. Bu yıllarda istatistik ile iktisadın arasındaki ilişkilerin fark edilmesiyle ekonometri doğmuş oldu.

1.3. İSTATİSTİK TANIMLARI

Tanım 1: İstatistik, belirli bir amaç için veri toplama, tablo ve grafiklerle özetleme, sonuçları yorumlama, sonuçların güven düzeylerini açıklama, örneklemlerden elde edilen (hesaplanan) sonuçları kitle için genelleme (çıkarsama), karakteristikler arasındaki ilişkileri araştırma, çeşitli konularda geleceğe ilişkin projeksiyon yapma, deney tasarlama ve düzenleme vb. kapsayan bir yöntemler topluluğudur.

Tanım 2: İstatistik verileri işler ve özetler, araştırmacılara yol gösterir. Değişkenler arasındaki ilişkileri incelememize, tahminler ve öngörüler yapmamıza, doğru karar vermemize yarar.

Tanım 3: Belirli amaç ya da amaçlar doğrultusunda gözlenen yığın olaylardan sayısal verilerin işlenerek, ilgili olayların oluşturduğu yığınların bilimsel olarak incelenmesinde kullanılan teknik ve yöntemler bilimi olarak tanımlanabilir.

Tanım 4: İstatistik; gözlem ve ölçme sonucunda elde edilen sayısal (rakamsal) verilerin tablolar veya grafikler halinde sunulması ve bunların karar alma sürecinde nasıl kullanılacağına ilişkin yöntemler sunar. İstatistiğe konu olacak olaylar gözlemlenebilir ve ölçülebilir olmalıdır. İstatistikler belirli ölçümler yapılarak elde edilmişlerdir yani istatistikler rastgele bulunmuş sayılar değildirler. Belirli bir kitleyi oluşturan birimlerin gözlemlenmesiyle elde edilen, ölçülebilen veya sayılabilen nicel bilgileri ifade etmektedirler. Bu nedenle; logaritma cetveli, telefon rehberi, milli piyango sonuçları birer istatistik değildirler.

Tanım 5: İstatistik; olasılığa dayalı yani sonucun baştan kesin olarak bilinmediği rastlantısal (tesadüfî) durumlar üzerinde çalışırken, belirsizlik altında karar vermeyi sağlar.

Tanım 6: Verilerin derlenmesi, işlenmesi ve yorumlanmasına yönelik yöntemlerin tümü. (İşleme ve yorum verileri temsil edebileceği düşünülen kuramsal dağılım ve modellere geçişi de içermek üzere matematiksel istatistik kapsamına girer.)

1.4. İSTATİSTİĞİN TEMEL KONULARI

İstatistik konuları bakımından Betimsel (tasviri) istatistik ve Çıkarımsal istatistik olarak ikiye ayrılır.

1.4.1. Betimsel İstatistik

Verilerin toplanması, yorumlanıp özetlenmesi gibi konularla ilgilenir. Tanımsal istatistik, verilerin bir takım sayısal ve(ya) grafiksel yöntemlerin kullanımı ile bilgiye dönüştürülmesini ve karar verme sürecinde kullanılmasını sağlar. İstatistik, geçmişi ve içinde bulunulan durumu tanımlayarak bir veri kümesine ilişkin özet değerler ve grafikler ortaya koyduğunda betimsel istatistik (descriptive statistics) adını alır.

Özetleme ve verilerin yoğunlaştırılması diğer bir deyişle verilerin hacimleri azaltılarak kullanım değerlerinin arttırılması betimsel istatistiğin konusuna girer. Tablolar ve grafikler yardımı ile verilerin özetlenmesi ve çok sayıda sayıdan oluşan bir veri grubunun“ortalama” gibi tek bir sayıya indirgenmesi yine bu alan için geçerlidir. Kısaca betimsel istatistik bir veri kümesinde bulunan bilgiyi sayısal ve grafiksel yöntemleri kullanarak özetler ve sunar. Betimsel amaçlı istatistik kitledeki tüm birimlerden ilgili değişken ya da değişkenler için veri toplandığında bunları kullanarak kitleyi özetlemeyi (betimlemeyi) amaçlar. Bu ise frekans dağılımı oluşturularak, grafikler çizilerek ya da parametreler (kitle ortalaması ve varyansı) hesaplanarak yapılır.

1.4.2. Çıkarımsal İstatistik

Tümevarımsal amaçlı istatistik kitle(ler)den rastgele seçilen örneklem(ler)den toplanan verileri kullanarak kitle(ler) nin parametrelerini tahmin etmeyi veya parametrelerle ilgili olan savların doğru olup olmadığının araştırılmasını amaçlar. Yanlı örneklem gözlendiğinde belirsizlikler içeren kitle hakkındaki önermelere geçmek için geliştirilen süreçle istatistiksel sonuç çıkarır.

Çıkarımsal istatistik, araştırma sürecinin betimsel istatistiğin bıraktığı yerden devam ettirilmesi suretiyle, çalışmanın amacına uygun bir şekilde daha ileri tekniklerle (olasılık başta olmak üzere) karar verme ve kitleye dair çıkarsamalar yapma sürecidir. Bu çıkarsamalar örnekleme dayalı olarak yapıldıkları için belirli bir hata düzeyinde ifade edilirler. (Kitle N adet birimden oluşan teorik topluluktur, aslında tüm birimlere ulaşmak olasıdır ancak zaman ve maliyet kısıdından dolayı genellikle istatistiksel çalışmalar kitleyi iyi bir biçimde temsil eden daha küçük alt gruplar yani örneklemler üzerinden yürütülür. Örneklemin büyüklüğü n ile gösterilir. O halde N büyüklüğündeki bir kitleden n büyüklüğünde tane mümkün örneklem çekilebilir.

Bu örneklemler büyüklük yani hacim olarak birbirinin aynı fakat içerdikleri birimler bakımından birbirinden farklıdırlar, kombinasyonun doğası gereği seçilen birim yerine konmamaktadır, bu nedenle mümkün örneklem sayısı hesaplanırken permütasyon yerine kombinasyon kullanılmaktadır.

Örnekleme dayanan çıkarımlar mutlak doğru olamayacağından "olasılık" kelimesi kullanılarak ifade edilir. Çıkarımsal istatistiğin genel olarak kullandığı yöntem tümevarımdır. Önceleri betimsel amaçlar taşıyan istatistik, 17-18 ve 19. yy'lar boyunca Bernoulli, Gauss, Laplace, Moivre gibi ünlü matematikçilerin olasılık teorisini geliştirmeleriyle çıkarımsal istatistiğin temellerini atmışlardır.

1.5. 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.

1.6. İSTATİSTİK METODUNUN AŞAMALARI

İstatistik metodu dört aşamada uygulanır. Bu aşamalar:

1. **Bilgilerin Toplanması (Röleveler)**

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

2. **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.

3. **Verilerin Sunulması**

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

4. **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.

1.7. TEMEL KAVRAMLAR

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

1.7.1. Gözlem

Birimlerde incelenen özelliğin gözlenmesi veya ölçülmesi suretiyle elde edilen değerlerdir.

1.7.2. 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 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.

1.7.3. Faktör

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

1.7.4. 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.

1.7.5. Birim

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

Birimler canlı ya da 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.

1.7.6. 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 ya da inceleme kapsamına giren obje ya da bireylerin tümüne anakütle ya da kütle denir. Gözlemin amacına bağlı olarak, anakütle küçülebilir ya da 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 ya da nelerden oluştuğunu ve sayısını saptamak bazen kolay bazen de çok zor hatta olanaksız olabilir. Bu zorluklar özellikle anakütledeki obje ya da deney sayısını saptamada ortaya çıkar. Çünkü çoğu kez belirli bir anakütleye girmesi gereken obje ya da bireyleri teker teker bulup ortaya karmak ve saymak olanaksızdır. Örneğin Türkiye'deki kanserli hastalar Üzerinde inceleme yapmak isteyen bir araştırmacıyı düşünelim. Bu araştırmacı için anakütlenin sınırlarını çizmek, kimlerin anakütleye girip kimlerin anakütleye girmeyeceğini saptamak oldukça kolaydır. Çünkü belirli bir zamanda Türkiye'de yaşayan ve kanserli olan herkes anakütleye dahildir, bunun dışında kalanlar dahil değildir. Fakat Türkiye'deki kanserli hepsini teker teker saptamak ve toplam sayılarını bulmak olanaksızdır. Çünkü kanserli olduğu bilinenlerin yanında kanserli olup ta bilinmeyen daha birçok kimsenin de bulunduğu bir gerçektir. Türkiye'deki bütün insanları kısa bir zaman içinde muayene edip kanserli olanları doğru bir şekilde saptama olanağı olmadığına göre bu anakütledeki birey sayısını doğru olarak bulmak olanaksızdır.

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

Pek çok araştırma amaçları için anakütleyi oluşturan obje ya da 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 sonuçlan anakütleye genelleme olanağı vardır. İstatistik teknik ve yöntemlerinin birçoğu da bu amaçla geliştirilmiştir. Bu tür istatistik teknik ve yöntemlerinin oluşturduğu bu kısma vardamdı istatistik denildiği bundan Önceki bölümde belirtilmişti.

Nüfus, yığın, anakütle (*population universe*) gibi adlarla da ifade edilen anakütle (*population*) incelenen konudaki olası tüm birimlerin oluşturduğu topluluktur. Anakütle birimler topluluğu yerine gözlem sonuçlarının oluşturduğu topluluk diye de tanımlanabilir. Anakütle N sembolü ile gösterilir ve anakütle hacmi ya da büyüklüğü (*population size*) şeklinde ifade edilir. Anakütleleri farklı şekillerde gruplamak olanaklı ise de en önemli ayrım sonlu ve sonsuz anakütle ayrımıdır. Sonlu anakütlede ilk ve son gözlem sonucunda bilinirken, sonsuz anakütlede 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 ya da Marmara denizindeki balıklar sonsuz anakütledir.

1.7.7. Örneklem

Örneklem, anakütleden seçilen ve anakütleye göre daha az sayıda birimden oluşan topluluktur. Örneklem, gözlem sonuçları açısından da anakütlede ulaşılabilen ya da 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 ya da olanaklı gözlem sonuçlarını

kapsarken, örneklem ona göre daha az sayıda birim ya da 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.

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 örneklemelere 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 örneklemelere de yanlı örneklemler denir. Yanlı örneklemelerden 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.

1.7.8. Parametre

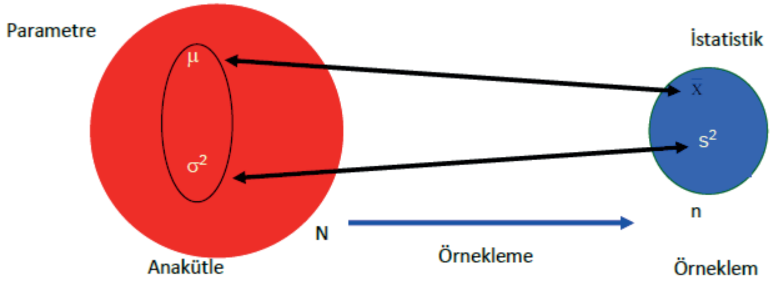
Anakütlenin özelliklerini belirleyen sayısal karakteristiklere parametre adı verilir. Anakütleyi tanımlamada kullanılabilen tipik değerlerdir. Anakütle aritmetik ortalaması, anakütle varyansı gibi değerlerdir.

1.7.9. İstatistik (statistics)

Örneği oluşturan birimlerden hesaplanmış, anakütleyi tanımlayan değerlere karşılık gelen değerlerdir. Örnek ortalaması, örnek varyansı gibi.

1.7.10. 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 (σ2) parametredir. Anakütledeki toplam gözlem sayısı N ile gösterilir.

Örneklemin aritmetik ortalaması ($\overline{X}$) ve örneklemin varyansı (S^2) ise istatistiktir. Anakütleden seçilen örneklem sayısı ise **n** kadardır.

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. Hâlbuki 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.

1.7.11. Tamsayım

Sonlu bir anakütlenin bütün birimlerinin incelenmesi ya da 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. 25.000 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=25.000, örneklem ise n=400 işçiden oluşmaktadır. 25.000 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

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

Oran= 20/80= 0,25 olur.

1.7.13. Yüzde

Oranın özel bir halidir. Oran olarak elde edilen değerin 100 ile çarpımının ifadesidir.

$$\text{Yüzde} = \frac{n}{N} \times 100$$

Örnek: Yukarıda ki örnekte sınıfta mavi gözlü öğrencileri yüzdesi;

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

1.7.14. 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 kat eden arabanın bir saat sonra nereye varabileceğini tahmin edebilmemiz gibi, hızı bilinen bir hastalığın gelecek yıl kaç kişiyi etkileyebileceği kestirilebilir.

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

Hız formülünde *k* 100, 1000, 100.000 gibi alınabilir.

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

a-) Erkeklerin bayanlara oranını

b-) 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: 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.

1.8. İSTATİSTİKTE DEĞİŞKENLER

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 ya da 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 gözlemde 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.

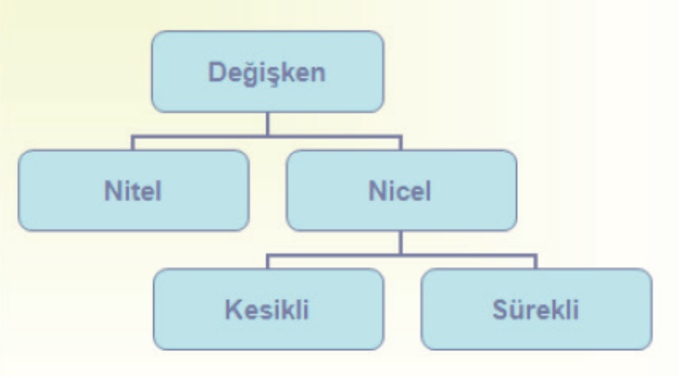
İkinci durumda ise tek bir nitelikle ilgili gözlem işini yaklaşık olarak aynı zamanda ve başka obje ya da bireyler üzerinde ya da ortamlarda yapı farklı sonuçlar alabiliriz. Örneğin günün belirli bir zamanında Türkiye'de 25 şehirde hava sıcaklığını ölçüp farklı sonuçlar gözleyebiliriz ya da aynı yaştaki 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 mekâna göre oluşuyorsa bu tür değişkenlere **mekân değişkeni** adı verilir. Örneğin doğum yeri ve üniversitelerin bulundukları şehirler mekân 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.

Değişkenler gözlenme biçimlerine Nitel ve Nicel değişken olarak ikiye ayrılır.



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.

1.8.1. Nitel Değişkenler

Bu değişkenler gözlemden gözleme farklılık gösterirler, ancak bu farklılık derece yönünden değil kalite ve çeşit yönündendir. Cinsiyet, medeni durum, göz rengi, din vb. Nitel değişkenler; birimlerin kalite, kategorik, ya da isimsel olarak belirtilebilen karakteristik özelliklerini, durumlarını ve pozisyonlarını belirten değişkenlerdir. Bu değişkenlerin verileri isimsel ya da sıralı ölçekle elde edilmişlerdir ve iki ya da daha fazla kategoriye (alt seçenek, sınıf, grup) ayrılarak sayımla elde edilir. Bu değişkenlerin verilerine nitel veriler adı verilir. Örneğin birimlerin, cinsiyeti, kan grubu, medeni durum, göz rengi, mesleği, yerleşim yeri, tuttuğu futbol takımı (fanatikler için) gibi nitelik bildiren durumları açıklayan değişkenlerdir.

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 değer 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.

1.8.2. Nicel Değişkenler

Birimlerin ölçüm ve tartım sonucu değerleri saptanan sayısal özelliklerini belirten değişkenlerdir. Bu değişkenler değerleri, mekanik ve elektronik araçlara sayısal olarak aralıklı ölçekli ya da orantılı ölçekli verileridir. Nicel değişkenlerin verilerine nicel veri adı verilir. Değişik derecelerde az ya da çok değerler alabilen değişkendir. Örneğin birimlerin, boy uzunluğu, vücut ağırlığı, kilosu, kan basıncı gibi özellikler nicel değişkenlerdir.

Bu tür değişkenler farklı derecelerde az ya da çok değerler alan değişkenlerdir. Yaş, ağırlık, boy uzunluğu, yıllık ya da aylık gelir, zekâ düzeyi, matematik ya da tarih bilgisi, havanın sıcaklığı, hava basıncı, hız, nüfus yoğunluğu vb. nicel değişkenlerdir. Bütün insanlar boy uzunluğu ve ağırlığa sahiptir ancak bunun miktarı kişiden kişiye ya da bir kişi için zamandan zamana değişebilir. Bütün şehirlerde insan yaşar sayısı şehirden şehre değişir. Bu gibi değişkenleri sayabildiğimiz gibi ölçerek derece sırasına koymak ve bir ölçek üzerinde işaretleme olanağı vardır. Bu tür değişkenlerin çoğu genellikle normal adını verdiğimiz türden bir dağılım gösterir. Bazen de bazı gerekçelerle bunların normal bir dağılım gösterdiği ya da göstereceği kabul edilir.

1.8.2.1. Kesikli Değişkenler

Kesikli değişkenler sayılarak elde edilen verileri alırlar. Bu değişkenler miktar yönünden değişiklik yerine tür yönünden değişiklik gösterir. Dolayısıyla bir obje ya da birey bir özelliğe sahiptir ya da değildir. Yani kesin değerler alırlar. Nitel değişkenlerin hemen hepsi süreksiz değişkendir. Cinsiyet, medeni durum, göz rengi gibi... Birinin diğerine göre daha çok veya az olması mümkün değildir.

1.8.2.2. Sürekli Değişkenler

Sürekli değişkenler bazı ölçümler sonunda elde edilen verilerin değerlerini alırlar. 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. Örneğin; balıkların ağırlıkları, tohum ağırlıkları, tohum çapı, bitki boyu, ortamdan bakterilerin tüketmiş oldukları şeker miktarı, ineklerin yıllık süt verimleri gibi. Noktadan sonraki hane sürekli veri olduğunu gösterir. Örneğin pamuk tohumu: 1.3; 1.32; 1.32076 gr. Bir ineğin bir senelik süt miktarı 3762 kg tartmaya dayandığı için süreklidir.

Kalitatif özellikler için sayma yoluyla veri elde edilir. Dolayısıyla bu veriler kesiklidir. Örneğin; doğan yavruların erkek ya da dişi oldukları (cinsiyet özellikleri) incelendiğinde erkek ve dişilerin sayısı ancak elde edilebilir. Kantitatif özellik için ise hem sayma hem ölçme, hem de tartma yoluyla veri elde edilebilir. Örneğin kök sayısı, balık ağırlığı, tohum çapı sürekli değişkenlerdir.

Sürekli değişkenleri, sınıflayarak kesikli değişkenlere dönüştürebiliriz. Örneğin kalsiyum düzeyi sürekli değişkendir. Eğer kalsiyum düzeyinin rakam olarak değerinden 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 değer alabilen bir kesikli değişkene dönüştürebiliriz. Hastaların yaşlarını da "20'den küçük", "20-34", "35-50", "50'den büyük" gibi sınıflandırarak, kesikli değişkene dönüştürebiliriz.

1.9. ÖLÇME DÜZEYLERİ

Ölçme 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.

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

İstatistik analize başlamadan önce ilk yapılacak şey, değişkenlerin nasıl ölçüldüğünün belirlenmesidir. İstatistikte ölçüm denildiği zaman anlaşılan, değişkenin alabileceği değerlerle 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. Değişkenlerin alabileceği değerlerin neler olabileceği, neler olamayacağı, yani nasıl ölçüldüğünü belirlemek, yapılacak istatistik analizin seçimi için çok önemlidir. Dört farklı ölçme düzeyi vardır.

1.9.1. Adlandırma Ölçme Düzeyi (Nominal)

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.

1.9.2. Sıralama Ölçme Düzeyi (Ordinal)

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 arsındaki 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.

1.9.3. Aralık Ölçme Düzeyi (İnterval)

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.

Isıyı ölçen santigrad ölçeği bir aralık ölçeğidir. 0 $_0$C başlangıç noktasıdır. Ölçümler arasındaki farklar birbirinin katı olarak ifade edilebilir. 0 $_0$C-18 $_0$C arasındaki fark 9 $_0$C-18 $_0$C arasındaki farkın iki katıdır. Ancak 18 $_0$C ısı 9 $_0$C ısının iki katıdır denilemez.

Aralık ölçeğinde sıfır noktası mutlak sıfır noktası değildir. İtibari bir sıfır noktasıdır. Sıcaklığın 0 0C 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 ölçeğinden birisiyle elde edilen ölçüm diğerine çevrilebilir. $_0$C derece Fahranayta, hicri takvim miladi takvime çevrilebilir.

1.9.4. Oran Ölçme Düzeyi (Ratio)

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.

1.10. ORTALAMALAR

Sık Kullanılan Bazı Matematiksel İfadeler :

$\{X_i: X_1 + X_2 + \ldots + X_n\}$ ifadesi X değişken kümesini göstermektedir.

$$\sum_{i=1}^{n} X_i = X_1 + X_2 + \ldots + X_n \quad ,$$

$$\sum_{i=1}^{n} c X_i = c(X_1 + X_2 + \ldots + X_n) = c\sum_{i=1}^{n} X_i$$

$$\sum_{i=1}^{n} (X_i + Y_i) = (X_1 + Y_1) + (X_2 + Y_2) + \ldots + (X_n + Y_n) = \sum_{i=1}^{n} X_i + \sum_{i=1}^{n} Y_i$$

$$\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{c} = \mathrm{c} + \mathrm{c} + \ldots + \mathrm{c} = \mathrm{nc}$$

$$\sum_{i=1}^{n} X_i Y_i = (X_1 Y_1) + (X_2 Y_2) + \ldots + (X_n Y_n)$$

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.

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 *kareli ortalama*, küçük bir sayıya ulaşmak için *harmonik ortalama* 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, duyarlı (analitik) ortalamalar ve duyarlı olmayan (analitik olmayan) ortalamalar şeklinde iki gruba ayrılmaktadır.

1.10.1. Duyarlı Ortalamalar

Duyarlı ortalamalar, serinin bütün terimlerinin hesaba katıldığı ortalamadır. Duyarlı ortalamalar, aritmetik ortalama, geometrik ortalama, harmonik ve kareli ortalamaları içerir.

1.10.1.1. Aritmetik Ortalama

Aritmetik ortalama deneklerin aldıkları değerlerin toplanıp denek sayısına bölünmesiyle elde edilen değerdir.

Basit serilerin aritmetik ortalaması, terimlerin toplamının terim sayısına bölünmesine eşittir.

Biri gözlem değerleri, diğeri frekansları gösteren iki sütundan oluşan sınıflanmış serilerin aritmetik ortalaması hesaplanırken gözlem değerleri ile frekanslar çarpımlarının toplamı frekanslar toplamına bölünmektedir.

Gruplanmış serilerde aritmetik ortalama hesabının yapılabilmesi için öncelikle sınıf orta noktalarının (m) bulunması gerekmektedir. Sınıf orta noktalarının hesaba katılmasında gruba dâhil birimlerin tamamının sınıf orta noktasında toplanmış olduğu varsayımından hareket edilmektedir.

Aritmetik Ortalama Formülleri

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

Örnek: 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}

N=20

$$\overline{X} = \frac{\sum_{i=1}^{N} X_i}{n} = \frac{202}{20} = 10{,}1$$

Örnek: Aşağıdaki sınıflanmış serinin aritmetik ortalamasını bulunuz?

Notlar (X_i)	:	40	60	70	80	100
Frekans(f_i)	:	5	4	5	4	2

$$\bar{X} = \frac{\sum f_i X_i}{f_i} = \frac{1310}{20} = 65{,}5$$

Örnek: Aşağıda gruplanmış olarak verilen serinin aritmetik ortalamasını bulunuz?

Sınıf sınırları	Sıklık (frekans= fi)	Sınıf Değeri (mi)	fi*mi
1.45 - 1.95	2	1.7	3,4
1.95 - 2,45	18	2,2	39,6
2,45 – 2,95	24	2,7	64,8
2,95 – 3,45	19	3,2	60,8
3,45 – 3,95	18	3,7	66,6
3,95 – 4,45	9	4,2	37,8
4,45 – 4,95	6	4,7	28,2
4,95 – 5,45	4	5,2	20,8
	100	TOPLAM	322

$$\bar{X} = \frac{\sum f_i m_i}{f_i} = \frac{322}{100} = 3{,}22$$

Örnek: 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

$$\bar{X} = \frac{\sum f_i m_i}{f_i} = \frac{898}{56} = 16\ gün$$

Aritmetik Ortalamanın Fayda ve Sakıncaları

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

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

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

Aritmetik Ortalamanın Özellikleri

i) Aritmetik ortalamanın terim sayısı ile çarpımı seri toplamına eşittir.

Terimlerin aritm $\overline{X} = \dfrac{\sum_{i=1}^{N} X_i}{n} \Rightarrow \overline{X}n = \sum_{i=1}^{N} X_i$ nalarının toplamı sıfırdır.

$$\sum (X_i - \overline{X}) = \sum X_i - nX = \sum X_i - n\frac{\sum X_i}{n} = 0$$

ii) Terimlerin aritmetik ortalamadan sapmalarının kareleri toplamı minimumdur.

$$\sum (X_i - a)^2 = \sum [(X_i - \overline{X}) + (\overline{X} - a)]^2$$
$$= \sum (X_i - \overline{X})^2 + 2(\overline{X} - a)\underbrace{\sum (X_i - \overline{X})}_{0} + N(\overline{X} - a)^2 = \sum (X_i - \overline{X})^2 + n(\overline{X} - a)^2$$

Son eşitlikte $n(\overline{X} - a)^2 > 0$ olacağından aşağıdaki sonuç elde edilir.

$$\sum (X_i - a)^2 > \sum (X_i - \bar{X})^2$$

iii) Bir serinin bütün terimlerine aynı sayıyı eklersek (çıkarırsak) aritmetik ortalama eklenen (çıkarılan) sayı kadar arta (azalır).

$$\frac{\sum (X_i + k)}{n} = \frac{\sum X_i + nk}{n} = \overline{X} + k \quad , \quad \frac{\sum (X_i - k)}{n} = \frac{\sum X_i - nk}{n} = \overline{X} - k$$

iv) Bir serinin bütün terimlerinin aynı sayıyla çarptığımızda (böldüğümüzde) aritmetik ortalama çarptığımız (böldüğümüz) sayıyla orantılı olarak büyür (küçülür).

$$\frac{\sum kX_i}{n} = \frac{k\sum X_i}{n} = k\overline{X} \quad , \quad \frac{\sum X_i / k}{n} = \frac{1/k \sum X_i}{n} = \frac{\overline{X}}{k}$$

v) Aritmetik ortalama çok duyarlı bir ortalamadır. Çünkü serinin bütün terimleri aritmetik ortalamayı etkiler. Özellikle de **aşırı uç değerlerden** çok etkilenir ve dolayısıyla temsili olma özelliğini kaybeder.

vi) İki serinin bütün terimleri karşılıklı olarak toplanarak (çıkartılarak) elde edilen serinin aritmetik ortalaması bu serilerin aritmetik ortalamalarının toplamına (farkına) eşittir.

$$\frac{\sum (X_i + Y_i)}{n} = \frac{\sum X_i + \sum Y_i}{n} = \frac{\sum X_i}{n} + \frac{\sum Y_i}{n} = \overline{X} + \overline{Y}$$

$$\frac{\sum (X_i - Y_i)}{n} = \frac{\sum X_i - \sum Y_i}{n} = \frac{\sum X_i}{n} - \frac{\sum Y_i}{n} = \overline{X} - \overline{Y}$$

1.10.1.2. Geometrik Ortalama

X_i: {X_1,X_2,…, X_n} pozitif sayılar kümesinin geometrik ortalaması, sayıların çarpımlarının n. dereceden köküdür.

$$GO = \sqrt[n]{X_1 X_2 \dots X_n}$$

Örnek: X_i: { 8, 12, 25, 6 }

$$GO = \sqrt[4]{8.12.15.6} = 10{,}95$$

Geometrik Ortalamanın Özellikleri

I. Geometrik ortalama özellikle aynı oranda artma yada azalma eğilimi gösteren olaylara ilişkin serilere uygulanır. Örneğin nüfus çoğalması, bakteri üremesi gibi geometrik dizilerde birim zamandaki artışı bulmak için GO kullanılır.

II. Simetrik olmayan ancak logaritmaları alındığında simetrik hale dönüşen serilere geometrik uygulama uygulanabilir.

III. Serideki terimler arasında bazı değerler sıfır veya negatifse GO hesaplanamaz.

IV. Geometrik uygulama aşırı uç değerlerden aritmetik ortalamaya göre daha az etkilenir.

V. GO ≤ AO ilişkisi vardır. Bütün X_i ler eşitse GO=AO olur.

Geometrik ortalama özellikle aynı oranda artma veya azalma eğilimi gösteren olaylara ilişkin serilere uygulanır. Bu olaylar arasında öncelikle nüfus belirtilebilir. Öte yandan, aslında simetrik olmadığı halde logaritmaları alındığında simetrik hale dönüşen serilere de geometrik ortalamayı uygulamak gerekir.

Not:

Başlangıçta A kadar birey varsa, bu bireyler birim zamanda r kadar bir hızla artıyorsa, n birim zaman sonra sayıları B kadar olmuş ise **$B = A(1+r)^n$** olur. Bu formül ***bileşik faiz formülü*** olarak adlandırılır.

Ortalama artış (r) buradan hesaplanır.

$$r = \sqrt[n]{\frac{B}{A}} - 1$$

Örnek: Bir bakteri kültürü 3 günde 1000'den 4000'e çıkmış ise ortalama günlük artış hızı(r) nedir?

$$r = \sqrt[3]{\frac{4000}{1000}} - 1 = \sqrt[3]{4} - 1 = 0.587$$

Örnek: Bir bölgenin nüfusu 2000 yılında 500.000 ölçülmüştür. Bu bölgenin yıllık nüfus artışı binde 15 ise ise 2005 yılında bu bölgenin nüfusu kaç olur?

$$B = A(1+r)^n = 500000(1+0,015)^5 = 538642$$

1.10.1.3. Harmonik Ortalama

Harmonik ortalama terimlerin terslerinin aritmetik ortalamasının tersidir.

X_i:{ X_1,X_2,…, X_n } değerlerinin harmonik ortalaması:

$$\text{Harmonik Ortalama} = \frac{1}{\frac{1}{\text{n}}\sum_{i=1}^{n}\frac{1}{X_i}}$$

Harmonik Ortalama	Basit Serilerde	$HO = \frac{n}{\sum \frac{1}{X_i}}$
	Sınıflanmış Serilerde	$HO = \frac{\sum f_i}{\sum \frac{f_i}{X_i}}$
	Gruplanmış Serilerde	$HO = \frac{\sum f_i}{\sum \frac{f_i}{m_i}}$

Örnek: Bir fabrikada üretim 4 makineyle yapılmaktadır. Bu makinelerin bir mamul için harcadıkları süreden (dk.) yararlanarak bir mamulün bu fabrikada ortalama kaç dkda üretildiğini bulunuz?

Makineler	X=Üretim Süresi (dk./parça)	1/X
I	2,5	0,4
II	2	0,5
III	1,6	0,625
IV	4	0,25
		1,775

$$\mathrm{HO} = \frac{n}{\sum \frac{1}{X_i}} = \frac{4}{1.775} = 2,254\, dk / parca$$

Örnek: X_i: { 6,8,3,5,4 } veri setinin harmonik ortalaması?

$$HO = \frac{1}{\frac{1}{5}\left(\frac{1}{6}+\frac{1}{8}+\frac{1}{3}+\frac{1}{5}+\frac{1}{4}\right)} = 4{,}65$$

Örnek: Bir otobüs firması iki şehir arasında (600 km.) 37 otobüsle seferler düzenlemektedir. Bu otobüslerin hızlarına göre dağılımı aşağıdaki gibidir. Otobüslerin ortalama hızını bulunuz?

Hız (km/saat)	Otobüs sayısı (f_i)	f_i/X_i
60	3	3/60=0,05
75	6	6/75=0,08
80	10	10/80=0,125
90	18	18/90=0,200
	37	0,455

$$\text{HO} = \frac{\sum f_i}{\sum \frac{f_i}{X_i}} = \frac{37}{0.455} = 81{,}32\ km/saat$$

Harmonik Ortalamanın Özellikleri

I. Serideki terimlerden biri sıfır ise harmonik ortalama sıfır çıkar.

II. Seri terimleri farklı işaretli olursa harmonik ortalamanın sonucu anlam taşımaz. Mesela verilerimiz -4, -2, 1,2,5 olsun. Buna göre HO=5,05 çıkar. Bu sonuç, bir ortalama maksimum değerden daha büyük bir değere sahip olamayacağı için, ortalama olarak kabul edilmez.

$$HO = \frac{1}{\frac{1}{5}\left(\frac{1}{-4}+\frac{1}{-2}+\frac{1}{1}+\frac{1}{2}+\frac{1}{5}\right)} = 5{,}05$$

III. HO ≤ GO ≤ AO ilişkisi vardır.

IV. HO sınırlı hallerde kullanılır. Tersine çevrildiğinde taşıyacağı anlama önem verilen oran türündeki niceliklerin ortalamasını bulmak için kullanılır. Bu niceliklere örnek olarak fiyat=para/mal, prodüktivite=iş/emek, verim=ürün/ekim alanı, hız=uzaklık/zaman verilebilir.

1.10.1.4. Tartılı (Ağırlıklı) Ortalama

Seri terimleri veya sınıfları arasında önem farkını dikkate almak için her terime veya sınıfa önemi ile orantılı bir tartı verilerek tartılı ortalama hesaplanır.

Örnek: Bir ildeki 5 hastanede acile gelen hastaların ortalama yaşları aşağıdaki şekilde bulunmuştur.

Hastane	Hasta (ti)	Ortalama Yaş(Xi)
1	10	25
2	15	30
3	20	40
4	5	20
5	30	15

Bu ilde acile gelen hastaların ortalama yaşı nedir? Bunun için tartılı ortalama bulunur.

T_i= i nci grubun tartısı X_i= i nci grubun değeri

$$TO = \frac{\sum_{i=1}^{k} t_i X_i}{\sum_{i=1}^{k} t_i} = \frac{2050}{80} = 25,625$$

1.10.2. 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.

1.10.2.1. Medyan (Ortanca)

Terimlerin küçükten büyüğe doğru (ya da büyükten küçüğe doğru) sıralanmış bir seride tam ortaya düşen ve seriyi iki eşit kısma bölen değere medyan (ortanca) denir. Medyanın hesabı basit, sınıflandırılmış ve gruplanmış serilerde farklıdır.

Basit serilerde, terimlerin sayısı tek ise tam ortadaki değer, çift ise ortadaki iki terimin aritmetik ortalaması medyanı verir. (N+1)/2 terime karşılık gelen terim medyandır.

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

Sayılar büyüklük sırasına dizilirse, { 1, 3, 6, 6, 8, 13, 27 } olur.

Ortada kalan sayı 6 olduğundan Ortanca=6 olur.

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

X_i : { 3, 7, 8, 9, 9, 21 } olur.

Ortada kalan iki değerin ortalaması ortancadır:

Ortanca= (8+9)/2= 8,5

Diğer bir ifade ile (N+1)/2. değer ortanca değerdir. Gözlem değeri çift ise sonuç şöyle bulunur. (N+1)/2 = (6+1)/2 = 3,5 yani 3. ve 4. gözlemlerin ortalaması ortancadır.

Gruplanmış Verilerden Ortanca Hesabı

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

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

Medyanın Özellikleri

I. Terimlerin medyandan mutlak sapmalarının toplamı minimumdur.

$$\sum|X_i - Medyan| = min$$

II. Basit bir sıralama ile bulunması mümkün olduğundan, medyan bir çok durumda pratiktir. Örneğin bir grup öğrencinin boy uzunluğunu teker teker ölçmeye gerek yoktur. Öğrenciler küçükten büyüğe doğru sıralanıp ortadaki öğrenci (ler) ölçülerek ortanca boy uzunluğu bulunabilir.

III. Seride açık (alt sınırı veya üst sınırı belli olmayan) sınıfların varlığı halinde medyan hesabı önem kazanır. Medyan sınıfı serinin ilk sınıfı olduğunda, sınıfın alt sınırı tahminsel olarak ele alınır.

IV. Diğer ortalamaların aksine, gruplanmış serinin medyan hesabında sınıf genişliklerinin tamamının eşit olması gerekmez.

V. Medyan serdeki anormal terimlerden etkilenmez.

Medyanı Kullanmanın Sakıncaları

I. Ortancanın standart hatası aritmetik ortalamadan daha büyüktür.

II. Ortanca üzerinde cebirsel işlemler yapılamaz.

III. Farklı alt grupların ortancaları biliniyorsa bu gruplar birleştiğinde ortanca nedir sorusu hesaplama ile bulunamaz.

1.10.2.2. Mod (Tepe Değeri)

Bir seride en çok tekrarlanan terime mod denir. En yüksek frekansa karşılık gelen X değeri Modu verir. Basit serilerde mod hesabı yapılmaz. Çünkü basit serilerde X'e karşılık gelen tüm frekanslar 1 olduğu için frekans sütunu bulunmaz.

Sınıflanmış serilerde modun belirlenmesi için frekans sütununda en yüksek frekans değerini veren X değeri bulunur.

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

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

Burada en fazla tekrarlanan değer 6 olduğu için Mod= 6 olur.

Örnek: Aşağıdaki sınıflandırılmış 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.

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

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

Örnek 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	3	
33.5-37.5	7	
37.5-41.5	14	
41.5-45.5	**17**	**Mod Sınıfı**
45.5-49.5	7	
49.5-53.5	2	
Toplam	50	

Mod=41.5+(17 - 14)*4/[(17-14)+(17-7)] =42.42

1.11. DEĞİŞİM ÖLÇÜLERİ

Bir anakütleyi (popülasyonu) tanıtmak, başka ana kütlelerle karşılaştırabilmek için merkezi eğilim ölçüleri yanında dağılışın genişliğini, değişkenliğin büyüklüğünü gösteren bir başka tipik değerin verilmesi gerekmektedir. Bu tipik değer değişim ölçüsüdür.

Ortalamaları eşit olan seriler, değişkenlikleri veya bölüne şekilleri farklı olduğunda birbirine benzemez. Bu nedenle serileri tam olarak tanımlayabilmek için ortalamayı, değişimlerini ve bölünme şekillerini incelemek gerekir.

İstatistikte kullanılan bazı simgeler:

	Örneklem Parametresi	Anakütle Parametresi
Aritmetik ortalama	$\overline{X}$	μ
Standart sapma	s	σ
Varyans	s^2	σ^2
Gözlem sayısı	n	N
Korelasyon	r	ρ

Örnek: Üç farklı bölümde okuyan öğrencilere ait notlar aşağıdaki gibidir:

X_i: { 49,49,49,50,51,52 } ; $\overline{X}$ =AO(X)= 50

Y_i : { 35,41,50,55,58,61 } ; $\overline{Y}$ =AO(Y)= 50

Z_i: { 15,21,33,49,90,92 } ; $\overline{Z}$ =AO(Z)= 50

Bu üç değişkenin ortalaması aynı olduğu halde X, Y ve Z değişkenlerinin aldığı değerlerin en küçük ve en büyük değerlerine bakıldığında birbirlerinden çok farklıdır. X değerleri 50'nin etrafında çok yakın kümelendiği halde Z değerleri 50'den çok uzakta yer almakta, yani daha büyük değişim göstermektedir.

Bir popülasyonun bireylerinin değerleri arasındaki bu değişimin bir ölçü ile ifade edilmesi gerekir. Bu ölçülere yayılım veya değişim ölçüleri denir.

1.11.1. Değişim Genişliği (Range)

En basit değişim ölçüsü Değişim Genişliğidir.

Değişim Genişliği = (En Büyük Gözlem – En Küçük Gözlem)

DG(X)= 52-49= 3

DG(Y)= 61-35= 26

DG(Z)= 92-15= 77

Değişim genişliği aslında çok kullanılan bir ölçüdür. Genelde şeker hastasının en küçük ve en yüksek değerleri nedir, ateşli bir hastanın en düşük ve en yüksek ateş derecesi nedir, bunlarla ilgilenilir.

Ancak değişim genişliğinde sadece iki aşırı uç değer kullanıldığı için istikrarlı bir ölçü değildir, bilgi kaybı çoktur. Diğer bir sakıncası değişim genişliği örnek büyüklüğüne çok bağlıdır. Örnek büyüdükçe aşırı değerleri kapsama olasılığı artar. Ölçü birimleri farklı serilerin değişkenliğinin kıyaslanmasında değişim genişliğinden yararlanmak yanlış olur.

1.11.2. Çeyrek Sapma

Değişim genişliğinin serinin iki ucunda yer ala aşırı değerlerden etkilenmesi sakıncasını gidermek için *çeyrekler arası fark* kullanılır. Bu ölçü erinin 3. ve 1. çeyrekleri arasındaki farkı esas alır.

Çeyrekler arası fark (IQR)= Q3-Q1

Çeyrekler bütün verileri hesaba katmadığından ve Q1'den küçük, Q3'den büyük değerlerin değişkenliğini ihmal etmesi sakınca meydana getirir.

Bir veri kümesinin ***çeyrek sapması*** Q ile gösterilir ve aşağıdaki gibi hesaplanır.

$$Q = \frac{Q_3 - Q_1}{2}$$

Örnek: {1,3,4,6,10,12} serisinin çeyrek sapmasını bulunuz?

Q_1=(n+1)/4=(6+1)/4=1.75 1.terim =1

(3-1)*0,75=1,5 Q_1=1+1,5=2,5

Q_3=3(n+1)/4=3(6+1)/4=5.25 5. terim 10

(12-10)*0,25=0,5 Q_3=10+0,5=10,5

Çeyrek sapma = Q3- Q1 = 10.5-2,5 = 8

1.11.3. Ortalama Sapma

Ortalama mutlak sapma olarak da anılan bu ölçü, terimlerin aritmetik ortalamadan mutlak sapmalarının aritmetik ortalamasıdır. Aritmetik ortalamaya göre ortalama sapma formülleri aşağıdaki gibidir:

Ortalama Sapma	**Basit Serilerde**	$O.S.=\frac{\sum\lvert X_i-\bar{X}\rvert}{n}$
	Sınıflanmış Serilerde	$O.S.=\frac{\sum f_i\lvert X_i-\bar{X}\rvert}{\sum f_i}$
	Gruplanmış Serilerde	$O.S.=\frac{\sum f_i\lvert m_i-\bar{X}\rvert}{\sum f_i}$

Ortalama sapma hesabı medyana göre yapıldığında $\bar{X}$ yerine Medyan konur.

Örnek: Aşağıdaki basit serinin ortalama sapmasını bulunuz?

X : {1,4,5,5,7,14}

$\bar{X}$ =36/6=6 $\lvert X_i-\bar{X}\rvert$ ={5,2,1,1,1,8} O.S.=18/6=3

1.11.4. Standart Sapma (Standard Deviation)

Kareli ortalama sapma adı verilen bu ölçü, terimlerin aritmetik ortalamadan cebirsel sapmalarının kareli ortalamasıdır. Standart sapma gözlemlerin ortalamadan ne kadar uzaklaştığını gösterir. Yani gözlemler arasında ne kadar yaygınlık olduğunu ifade eder.

Bir çalışma grubundaki her bir verinin ortalamaya göre ne kadar uzaklıkta olduğunu, bir diğer deyişle dağılımın ne yaygınlıkta olduğunu gösteren bir ölçüdür. Varyans ise belirli bir popülasyonda incelenen özelliğin ya da ölçümlerin ne genişlikte yani nasıl bir aralıkta (dar veya geniş) dağıldığının göstergesidir.

Standart sapma *varyansın* kare köküdür. *Varyans birimsiz ifade edildiği halde standart sapma ölçülen özelliğin birimi ile ifade edilir.* Yani milimetre yerine milimetre kare gibi veri biriminin karesinin kullanılması uygun olmadığından, varyansın yerine standart sapma kullanılır.

Ortalama etrafındaki saçılma fazla ise varyans büyük olacak, dolayısıyla standart sapma da büyük çıkacaktır. Veri seti içinde aşırı değer fazla ise bunların varyansı etkilemesi olağan karşılanabilir, ancak birkaç aşırı değerin varyansı büyütmesi olağan karşılanmaz.

Dağılımdaki maksimum ve minimum değerler arasındaki farkı dörde bölmek suretiyle, standart sapma kabaca tahmin edilebilir.

Moses'e göre standart sapma değeri örnekteki gözlem değerleri arasındaki varyasyonun derecesini tanımlayan bir istatistiktir. Değişkenin dağılışı normal dağılış gösteriyorsa, standart sapma verilerin ortalama etrafındaki dağılışını iyi bir şekilde açıklar.

$$s = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n}}$$

Standart Sapma	**Basit Serilerde**	$s = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n}}$
	Sınıflanmış Serilerde	$s = \sqrt{\frac{\sum f_i (X_i - \bar{X})^2}{\sum f_i}}$
	Gruplanmış Serilerde	$s = \sqrt{\frac{\sum f_i (m_i - \bar{X})^2}{\sum f_i}}$

Gözlem sayısı n<30 ise örnek verisine ilişkin standart sapma hesaplanırken paydada n yerine n-1 alınır. Çünkü elde edilen değer örneğin alındığı yığının standart sapmasını daha iyi verir. n'in büyük değerleri için (n≥30) iki tanım arasında bir fark yoktur.

$$s = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n-1}}$$

Standart sapma daima ortalama sapmadan büyük çıkar. Çünkü standart sapma kareli ortalama cinsinden, ortalama sapma ise aritmetik ortalama cinsinden hesaplanan değişkenlik ölçüdür.

Gözlem değerlerinin tamamı birbirine eşit olamayacağından standart sapmanın değerleri pozitiftir.

Örnek: Xi:{49,49,49,50,51,52} veri setinin standart sapmasını bulunuz?

A.O.=50 , s^2= [(49-50)2+ (49-50)2+ (49-50)2+ (50-50)2+(51-50)2+ (52-50)2]/(6-1)

= (1+1+1+0+1+4)/5= 8/5=1,6

Standart sapma s veya s_x şeklinde değişken ismi ile birlikte yazılabilir.

$$s = \sqrt{Varyans} = \sqrt{s^2} = \sqrt{1.6} = 1.26$$

1.11.5. Varyans

Dağılış ölçüsü olarak kullanılan en uygun ölçü standart ayrılış ölçüsüdür. Standart ayrılış ölçüsü varyansın kareköküdür. Bu nedenle önce varyans hesaplanmasını görelim.

Xi: { X1,X2,…, Xn} değerler kümesinin varyansı:

$$Varyans = s^2 = \frac{1}{(n-1)}\sum_{i=1}^{N}(X_i - \overline{X})^2 = \frac{1}{(n-1)}\left\{\sum_{i=1}^{N} X_i^2 - \frac{(\sum X_i)^2}{N}\right\}$$

Örnek: Xi:{49,49,49,50,51,52} değişkeninin varyansı nedir?

$\overline{X}$ = 300/6=50

s^2= [(49-50)2+(49-50)2+ (49-50)2+ (50-50)2+(51-50)2+ (52-50)2]/(6-1)

= (1+1+1+0+1+4)/5 = 8/5=1,6

1.11.6. Standart Hata

Standart hata örnekleme dağılımındaki ortalamaların standart sapmasıdır.

Seçilecek örneklerde ortalamalar arasındaki yaygınlığı gösterir. Standart hata örnek büyüklüğünün fonksiyonudur. Böylece n arttıkça hata küçülür. *Standart hata* Aritmetik ortalamada oluşan hatanın belirlenmesi için bulunur.

Standart hata ortalamalarla ilgilidir, deneklerle ilgili değildir. Standart hata, standart sapma gibi değişkenin dağılışı hakkında bilgi vermez.

$$S.H. = s_{\overline{X}} = \sqrt{\frac{Varyans}{n}} = \frac{s}{\sqrt{n}}$$

1.11.7. Değişim (Varyasyon) Katsayısı

Bir serinin standart sapmasının aritmetik ortalamasına bölünüp 100 ile çarpılması sonucu elde edilen değere değişim katsayısı adı verilir. Değişim katsayısının ölçü birimi yoktur.

$$DK = \frac{s}{\overline{X}} \times 100$$

Değişim katsayısı küçük olan serilerin diğerlerine göre aha az değişken olduğu söylenir. Bunun anlamı ise seri terimlerinin aritmetik ortalama etrafında daha homojen olarak dağıldığıdır.

Örnek: Xi:{49,49,49,50,51,52}serisinin değişim katsayısını bulunuz?

D.K= 1.26 *100 /50=%2,52

1.11.8. Momentler

Terimlerin sıfırdan sapmalarının veya aritmetik ortalamadan sapmalarının çeşitli kuvvetlerinin aritmetik ortalamalarının tamamına moment adı verilir. Momentler sıfır etrafındaki momentler ve aritmetik ortalama etrafındaki momentler olmak üzere ikiye ayrılır. Sapmaların derecesi (r) ile gösterilir ve r momentin derecesini belirler.

Sıfıra Göre Momentler	Basit Serilerde	$m_r = \dfrac{\sum X_i^r}{n}$
	Sınıflanmış Serilerde	$m_r = \dfrac{\sum f_i X_i^r}{\sum f_i = n}$
	Gruplanmış Serilerde	$m_r = \dfrac{\sum f_i m_i^r}{\sum f_i = n}$

Burada r=1 ise birinci moment $m_1 = \overline{X}$ yani aritmetik ortalamayı verir.

Aritmetik Ortalamaya Göre Momentler	Basit Serilerde	$\mu_r = \dfrac{\sum (X_i - \bar{X})^r}{n}$
	Sınıflanmış Serilerde	$\mu_r = \dfrac{\sum f_i (X_i - \bar{X})^r}{\sum f_i}$
	Gruplanmış Serilerde	$\mu_r = \dfrac{\sum f_i (m_i - \bar{X})^r}{\sum f_i}$

Bu serilerde r=1 için $\mu_1 = 0$, r=2 için $\mu_2 = \sigma^2$ yani varyanstır.

Sıfıra göre momentler bilindiğinde, aritmetik ortalamaya göre momentler bulunabilir. Bu şekilde elde edilen bağıntılara **König Teoremi** adı verilir.

$$\mu_2 = m_2 - m_1^2$$

$$\mu_3 = m_3 - 3m_1m_2 + 2m_1^3$$

$$\mu_4 = m_4 - 4m_1m_3 + 6m_1^2m_2 - 3m_1^4$$

Örnek: Aşağıdaki basit serinin sıfıra ve aritmetik ortalamaya göre momentlerini bulunuz?

X_i	X_i^2	X_i^3	X_i^4	$X_i - \overline{X}$	$(X_i - \overline{X})^2$	$(X_i - \overline{X})^3$	$(X_i - \overline{X})^4$
1	1	1	1	-5	25	-125	625
3	9	27	81	-3	9	-27	81
4	16	64	256	-2	4	-8	16
6	36	216	1296	0	0	0	0
10	100	1000	10000	4	16	64	256
12	144	1728	20736	6	36	216	1296
$\sum$ 36	306	3036	32370	0	90	120	2274

Sıfıra göre momentler:

m_1=36/6=6 m_2=306/6=51 m_3=3036/6=506 m_4=32370/6=5395

Aritmetik Ortalamaya Göre Momentler:

μ_1=0/6=0 μ_2=90/6=15 μ_3=120/6=20 μ_4=2274/6=379

1.12. İSTATİSTİKSEL DAĞILIMLAR

1.12.1. Kesikli Dağılımlar

Bernoulli Dağılımı

Bir deneyde başarı ve başarısızlık diye nitelendirilen iki sonuçla ilgilenildiğinde bu deneye (iki sonuçlu) Bernoulli deneyi ya da Bernoulli denemesi denir.

başarı olasılığı $\rightarrow p$, $(0 < p < 1)$

başarısızlık olasılığı $\rightarrow 1 - p = q$

başarı-başarısız/ sağlam-bozuk/ olumlu-olumsuz/ ölü-canlı

Bernoulli dağılımının olasılık fonksiyonu

$$f(x) = P(X = x) = p^x(1-p)^{1-x}, \quad x = 0,1$$

şeklinde verilir.

x	0	1
$P(X = x)$	$1-p$	p

Bernoulli dağılımın beklenen değer ve varyansı aşağıdaki gibidir:

$$E(X) = \sum_x x\, f(x) = 0(1-p) + 1p = p$$

$$E(X^2) = \sum_x x^2\, f(x) = 0^2(1-p) + 1^2 p = p$$

$$Var(X) = E(X^2) - [E(X)]^2 = p - p^2 = p(1-p) = pq$$

Binom Dağılımı

Başarı olasılığı p olan bir Bernoulli denemesinin aynı şartlar altında (bağımsız olarak) n kez tekrarlanması ile oluşan deneye binom deneyi denir.

Binom deneyinin aşağıdaki koşulları sağlaması gerekir:

- Deney süresince örneklemde denek sayısı ya da deneme sayısı değişmez olmalıdır.
- Denemeler birbirinden bağımsızdır.
- Her denemede iki olası sonuç vardır (istenen ve istenmeyen olay).
- Her denemede ilgilenilen olay olasılığı p değişmezdir. Dolayısıyla istenmeyen olay olasılığı $q = 1 - p$ de değişmezdir.

Binom dağılımı kesikli bir olasılık dağılımıdır. X rasgele değişkeni binom dağılımına sahip olduğunda $X \sim b(n, p)$ ile gösterilir.

Binom dağılımının olasılık fonksiyonu,

$$f(x) = P(X = x) = \binom{n}{x} p^x q^{n-x}, \quad x = 0,1, \dots, n$$

şeklinde verilir.

Binom dağılımının beklenen değer ve varyansı aşağıdaki gibidir:

$\mu = E(X) = np$

$\sigma^2 = Var(X) = npq$

Çarpıklık katsayısı $ÇK = \frac{q-p}{\sqrt{npq}}$, basıklık katsayısı $BK = 3 + \frac{1-6pq}{npq}$

Örnek Bir kutuda bulunan 10 tabletten 5 tanesi aspirindir. Bu kutudan yerine koyarak 3 tablet çekildiğinde 2 tanesinin aspirin olması olasılığı nedir?

X: Çekilen tabletin aspirin olması

$X \sim b(n = 3, p = \frac{1}{2})$

$$P(X = 2) = \binom{3}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^{3-2} = \frac{3!}{2!\,(3-2)!}\left(\frac{1}{2}\right)^3 = 0.375$$

Poisson Dağılımı

Bu dağılım, belirli bir aralıkta gerçekleşme olasılığının çok küçük olduğu durumlarda kullanılır. Örneğin Ankara'da Beşevler kavşağında bir gün içerisinde meydana gelen trafik kazaları, belli bir yılda meydana gelen doğal afetler, az rastlanan hastalıklar gibi.

Denek sayısı olan n büyük iken p de çok küçük ise binom dağılımı poisson dağılımına yaklaşır. Genel olarak $np \leq 5$ olduğu zaman binom dağılımı yerine poisson dağılımı kullanılabilir. Ayrıca n' nin 20 den büyük olması koşulu vardır.

X rasgele değişkeni Poisson dağılımına sahipse, bu değişkenin olasılık fonksiyonu aşağıdaki gibidir:

$$f(x) = P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0,1,2, \ldots$$

λ gerçekleşen ortalama olay sayısı olup $\lambda = np$ dir.

Poisson dağılımının beklenen değer ve varyansı aşağıdaki gibidir.

$\mu = E(X) = \lambda$

$\sigma^2 = Var(X) = \lambda$

Çarpıklık katsayısı $ÇK = \frac{1}{\sqrt{\lambda}}$, basıklık katsayısı $BK = 3 + \frac{1}{\lambda}$

Örnek 3.18. Bir şehirde ender rastlanan bir hastalıktan, bir hafta içinde ortalama ölen kişi sayısı 4' dür. Belli bir hafta içinde bu hastalıktan,

a) Hiç kimsenin ölmemesi
b) En az 2 kişinin ölmesi
c) 3 kişinin ölmesi

olasılıklarını hesaplayınız.

X: *bir haftada bu hastalıktan ölenlerin sayısı*

$$f(x) = P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0,1,2, \dots, \lambda = 4$$

a) $P(X = 0) = \frac{e^{-4}4^0}{0!} = 0.0183$

b) $P(X \geq 2) = 1 - P(X < 2) = 1 - \left(P(X = 0) + P(X = 1)\right) = 1 - \left(\frac{e^{-4}4^0}{0!} + \frac{e^{-4}4^1}{1!}\right) =$
$= 1 - (0.0183 + 0.0733) = 1 - 0.0916 = 0.9084$

c) $P(X = 3) = \frac{e^{-4}4^3}{3!} = 0.195$

Geometrik Dağılım

Arka arkaya n kez tekrarlanan bir Bernoulli deneyinde ilk istenen sonucun (başarı ya da başarısızlık) elde edilmesi için yapılan deney sayısı olan X' e geometrik rasgele değişken denir. Bu değişkenin dağılımı geometrik dağılım adını alır.

X rasgele değişkeni geometrik dağılıma sahipse, $X \sim Geo(p)$ biçiminde gösterilir.

X rasgele değişkeninin olasılık fonksiyonu,

$$f(x) = P(X = x) = p(1 - p)^{x-1}, \quad x = 1,2,3, \dots \quad 0 < p < 1$$

biçimindedir.

Geometrik dağılımın beklenen değer ve varyansı aşağıdaki gibidir.

$\mu = E(X) = \frac{1}{p}$

$\sigma^2 = Var(X) = \frac{1 - p}{p^2}$

Örnek . Bir torbada 8 beyaz, 4 siyah top bulunmaktadır. Her defasında yerine konularak bir top çekiliyor.

a) Beyaz topun ilk defa 5'inci çekilişte çıkma olasılığı nedir?
b) X rasgele değişkeni beyaz bir top çekmek için yapılan deney sayısı ise X rasgele değişkenin beklenen değer ve varyansı nedir?

X: İlk başarıya ulaşıncaya kadar yapılan deney sayısı

$$X \sim Geo(p = \frac{8}{12} = \frac{2}{3})$$

$$f(x) = P(X = x) = \frac{2}{3}(1 - \frac{2}{3})^{x-1}, \quad x = 1,2,3, \ldots$$

a) $P(X = 5) = \frac{2}{3}(1 - \frac{2}{3})^{5-1} = \frac{2}{3}(\frac{1}{3})^4 = \frac{2}{243}$

b) $E(X) = \frac{1}{p} = \frac{1}{2/3} = \frac{3}{2}$

$$Var(X) = \frac{1-p}{p^2} = \frac{1 - \frac{2}{3}}{\left(2/3\right)^2} = \frac{1}{3} \cdot \frac{9}{4} = \frac{3}{4}$$

1.12.2. Sürekli Dağılımlar

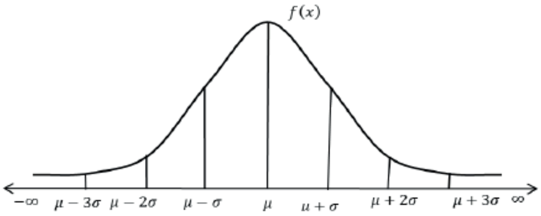
Normal Dağılım

Normal dağılım ya da Gauss dağılımı pratikte çok sık karşılaşılan sürekli bir dağılımdır. İnsanların boy uzunlukları, zekâ seviyeleri gibi değişkenler normal dağılmış tesadüfi değişkenlere örnek olarak verilebilir.

Tanım : X normal dağılmış bir tesadüfi değişken iken bunun olasılık yoğunluk fonksiyonu

$$f_X(x) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} & , \quad -\infty < x < \infty \\ 0 & , \quad d.d. \end{cases}$$

biçimindedir. Bu dağılım iki parametrelidir, bunlar $-\infty < \mu < \infty$ ve $\sigma^2 > 0$ dır. Normal dağılım için $X \sim N(\mu, \sigma^2)$ gösterimi kullanılır. Normal dağılımın olasılık yoğunluk fonksiyonu aşağı şekildeki gibidir.



Şekil $X \sim N(\mu, \sigma^2)$'nin olasılık yoğunluk fonksiyonu.

Şekilde görüldüğü gibi normal dağılım $\mu -$ ye göre simetriktir. μ nün sağındaki ve solundaki iki alan eşittir.

Standart Normal Dağılım

$X \sim N(\mu, \sigma^2)$ iken $F(x)$'e ait integralin kapalı hali elde edilememektedir. Dolayısıyla $F(x)$ değerlerinin hesaplanması için tablolar geliştirilmiştir. Bu bağlamda X tesadüfi değişkeni standartlaştırılarak Z değişkenine dönüştürülür.

$$Z = \frac{X-\mu}{\sigma}$$

alındığında

$$E(Z) = E\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma}(E(X)-\mu) = \frac{1}{\sigma}(\mu-\mu) = 0$$

$$Var(Z) = Var\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2}Var(X-\mu) = \frac{1}{\sigma^2}Var(X) = 1$$

bulunur. Ayrıca Normal dağılıma sahip bir X tesadüfi değişkenin herhangi bir lineer fonksiyonu da normal dağılıma sahip olduğundan dolayı $Z \sim N(0,1)$ yazılır.

Tanım $Z \sim N(0,1)$ tesadüfi değişkeninin olasılık yoğunluk fonksiyonu

$$f_Z(z) = \begin{cases} \dfrac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} & , \quad -\infty < z < \infty \\ 0 & , \quad d.d. \end{cases}$$

yazılır. $f_Z(z)'$ nin grafiği aşağıdaki gibidir.

Örnek: Bir konserve fabrikasındaki konserve kutularının ağırlıkları $\mu = 500$ gr ve $\sigma^2 = 100$ gr olarak normal dağılıma sahiptir. Buna göre tesadüfi olarak seçilen bir kutunun,

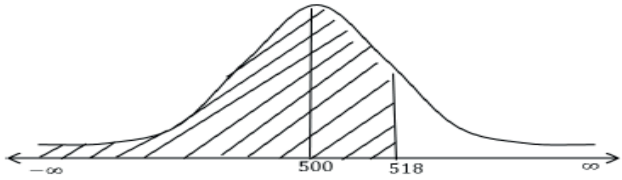
a) 518 gr dan az olma olasılığı nedir?

b) 480 gr dan çok olma olasılığı nedir?

c) 485 ile 515 gr arasında olma olasılığı nedir?

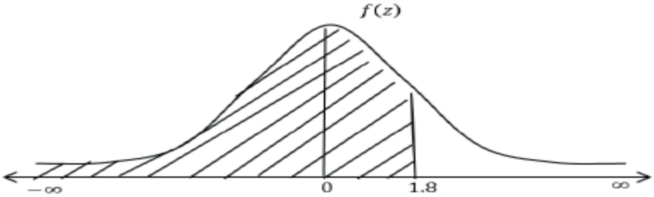
Çözüm.

a) $P(X < 518) = P\left(Z < \frac{518-500}{10}\right) = P(Z < 1{,}8)$



Şekil $X \sim N(500{,}100)$'nin olasılık yoğunluk fonksiyonu ve $P(X < 518)$ olasılığı

X tesadüfi değişkeni standartlaştırıldığında istenen olasılık aşağıdaki şekilde gösterilir.



Şekil $Z \sim N(0{,}1)$'nin olasılık yoğunluk fonksiyonu ve $P(Z < 1.8)$ olasılığı

$$P(Z < 1.8) = \Phi(1.8) = 0{,}9641$$

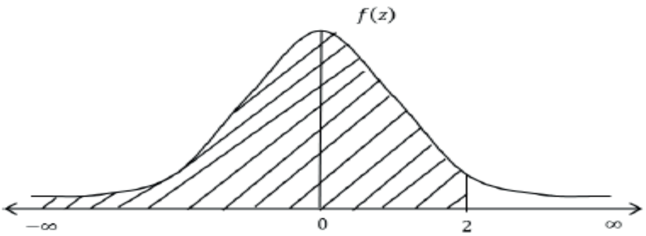
olarak bulunur.

b)

$$P(X > 480) = P(Z > -2) = P(Z < 2) = \Phi(2) = 0{,}9772$$

veya şöyle bulunur.

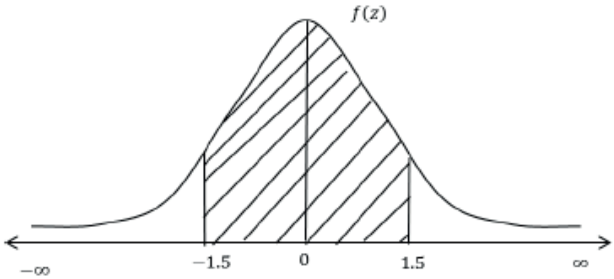
$$P(X > 480) = P(Z > -2) = 1 - P(Z \leq -2)$$
$$= 1 - \Phi(-2) = 1 - 0{,}0228 = 0{,}9772$$



Şekil $Z \sim N(0{,}1)$'nin olasılık yoğunluk fonksiyonu ve $P(Z < 2)$

c)

$$\begin{aligned} P(485 < X < 515) &= P\left(\frac{485-500}{10} < Z < \frac{515-500}{10}\right) \\ &= P(-1.5 < Z < 1.5) \\ &= \Phi(1.5) - \Phi(-1.5) = 0{,}8664 \end{aligned}$$



Şekil $Z \sim N(0,1)$**'nin olasılık yoğunluk fonksiyonu ve** $P(-1.5 < Z < 1.5)$

Düzgün Dağılım

Tanım X tesadüfi değişkeninin olasılık yoğunluk fonksiyonu

$$f_X(x) = \begin{cases} \dfrac{1}{b-a} & , \quad a \le x \le b \\ 0 & , \quad \text{d.d.} \end{cases} \qquad (5.26)$$

ise, X tesadüfi değişkeni $[a,b]$ kapalı aralığında düzgün dağılıma sahiptir denir ve $X \sim U(a,b)$ biçiminde gösterilir.
$f_X(x)$ ' in grafiği aşağıdaki gibidir.

$X \sim U(a,b)$'nin dağılım fonksiyonu

$$F_X(x) = \begin{cases} 0 & , \quad x < a \\ \dfrac{x-a}{b-a} & , \quad a \le x \le b \\ 1 & , \quad x > b \end{cases}$$

Beklenen Değer ve Varyans

$$E(X) = \int_{D_X} x f_X(x) dx$$

$$\begin{aligned} E(X) &= \int_a^b x \frac{1}{b-a} dx = \left.\frac{x^2}{2(b-a)}\right|_a^b = \frac{b^2-a^2}{2(b-a)} \\ &= \frac{(b+a)(b-a)}{2(b-a)} \\ &= \frac{b+a}{2} \end{aligned}$$

bulunur.

$$E(X^2) = \int_{D_X} x^2 f_X(x)dx$$

$$E(X^2) = \int_a^b x^2 \frac{1}{b-a}dx = \frac{1}{(b-a)}\frac{x^3}{3}\Big|_a^b = \frac{b^3-a^3}{3(b-a)}$$

$$= \frac{(b-a)(b^2+ab+a^2)}{3(b-a)}$$

$$= \frac{a^2+ab+b^2}{3}$$

$$Var(X) = E(X^2) - \left(E(X)\right)^2$$

$$= \frac{(b-a)^2}{12}$$

Moment Çıkaran Fonksiyon

$$M_X(t) = \int_{D_X} e^{tx} f_X(x)dx$$

$$M_X(t) = \int_a^b e^{tx}\frac{1}{b-a}dx = \frac{1}{t(b-a)}e^{tx}\Big|_a^b$$

$$= \frac{e^{tb}-e^{ta}}{t(b-a)}$$

Örnek : Bir otobüs durağına saat 10.00'da geldiğimizi varsayalım. Bu durağa otobüs, saat 10.00 ile 10.30 arasında düzgün dağılıma uygun herhangi bir zamanda gelmektedir. Buna göre,

a) Otobüsün gelmesi için 10 dakikadan fazla bekleme olasılığını hesaplayınız.

b) Eğer saat 10.15 ve otobüs hala gelmemiş ise, en az 10 dakika daha bekleme olasılığı nedir?

***Çözüm*.a)**

$$f_X(x) = \begin{cases} \frac{1}{30} & , \quad 10.00 < x < 10.30 \\ 0 & , \quad \text{d.d.} \end{cases}$$

$$F_X(x) = \begin{cases} 0 & , \quad x < 10.00 \\ \frac{x}{30} & , \quad 10.00 < x < 10.30 \\ 1 & , \quad x \geq 10.30 \end{cases}$$

X: "Otobüsün gelmesi için beklediğimiz süre" olsun.

$$P(10 < X) = 1 - P(X \leq 10) = 1 - \frac{10}{30} = 0{,}67$$

b)

$$P(X > 25 | X > 15) = \frac{P(X > 25)}{P(X > 15)} = \frac{1 - P(X \leq 25)}{1 - P(X \leq 15)}$$

$$=\frac{1-\frac{25}{30}}{1-\frac{15}{30}}=\frac{1}{3}$$

Üstel Dağılım

Tanım . : Negatif olmayan değerler alan sürekli X tesadüfi değişkenin olasılık yoğunluk fonksiyonu $\lambda > 0$ için,

$$f_X(x)=\begin{cases}\lambda e^{-\lambda x} & , \quad x>0\\ 0 & , \quad x\le 0\end{cases}$$

ise X, λ parametreli üstel dağılıma sahiptir ve $X \sim Üstel(\lambda)$ ile gösterilir.

$X \sim Üstel(\lambda)$'nın dağılım faonksiyonu da

$$F_X(x)=\begin{cases}0 & , \quad x<0\\ 1-e^{-\lambda x} & , \quad x\ge 0\\ 1 & , \quad x\to\infty\end{cases}$$

biçimindedir. Üstel dağılım özellikle bekleme hattı, güvenilirlik, sağ kalım analizi ve sürekli parametreli Markov zincirlerinde sıkça kullanılır.

Beklenen Değer ve Varyans

$$E(X) \quad = \int_{D_X} x f_X(x)dx$$

$$E(X) \quad = \int_0^\infty x\lambda e^{-\lambda x}dx$$

kısmi integralden,

$$x=u \text{ ve } \lambda e^{-\lambda x}dx \Rightarrow dx=du, u=-e^{-\lambda x}$$

olur. Böylece

$$E(X) \quad = \int_0^\infty x\lambda e^{-\lambda x}dx = (-xe^{-\lambda x})\Big|_0^\infty + \int_0^\infty e^{-\lambda x}dx$$

$$=\frac{1}{\lambda}$$

$$E(X^2) \quad = \int_{D_X} x^2 f(x)dx$$

$$= \int_0^\infty x^2\lambda e^{-\lambda x}dx$$

iki kez kısmi integral alındığında,

$$E(X^2) = \frac{2}{\lambda^2}$$

bulunur.

$$Var(X) = E(X^2) - [E(X)]^2$$
$$= \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

elde edilir.

Moment Çıkaran Fonksiyon

$$M_X(t) = E(e^{tX}) = \int_{D_X} e^{tx} f_X(x)dx$$

$$M_X(t) = \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx$$
$$= \lambda \int_0^\infty e^{x(t-\lambda)} dx = \frac{\lambda}{\lambda - t}$$

Örnek Bir elektronik parçanın ömrü, yıl olarak aşağıdaki olasılık yoğunluk fonksiyonu ile verilmiştir.

$$f_X(x) = \begin{cases} \frac{1}{3} e^{-\frac{1}{3}x} & , \quad x > 0 \\ 0 & , \quad d.d \end{cases}$$

Bu elektronik parçanın ömrünün

a) En çok 4 yıl olma olasılığı nedir?

b) En az 3 yıl olma olasılığı nedir?

c) Beklenen ömrü kaç yıldır?

d) Ömrünün standart sapması nedir ?

Çözüm.

a)

$$F_X(x) = \begin{cases} 0 & , \quad x < 0 \\ 1 - e^{-\frac{1}{3}x} & , \quad x \geq 0 \\ 1 & , \quad x \to \infty \end{cases}$$

$$P(X \leq 4) = F_X(4) = 1 - e^{-\frac{4}{3}} = 0{,}7364$$

a) $P(X \geq 3) = 1 - P(X < 3) = 1 - F_X(3) = e^{-1} = 0{,}3678$

c)

$$E(X) = \frac{1}{\frac{1}{3}} = 3$$

d)

$$Var(X) = \frac{1}{\left(\frac{1}{3}\right)^2} = 9$$

ise $\sigma_X = \sqrt{V(X)} = \sqrt{9} = 3$ olarak elde edilir.

$X \sim \ddot{U}stel(\lambda)$ tesadüfi değişkeni için aşağıdaki eşitlikler mevcuttur.

$$P(X \geq a) = P(X > a) = e^{-\lambda a}$$

$$P(X \leq a) = 1 - P(X > a) = 1 - e^{-\lambda a}$$

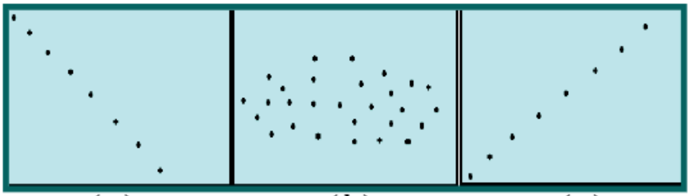
$$P(a \leq X \leq b) = F(b) - F(a) = e^{-\lambda a} - e^{-\lambda b}$$

- Gamma Dağılımı ve Beta Dağılımı da sürekli dağılımlardandır.

1.13. KORELASYON

Saçılım(scatterplot) grafikleri iki değişken arasındaki ilişki hakkında genel bir bilgi edinmemizi sağlar. Ancak, ilişkinin miktarı konusunda yorum yapabilmek için korelasyon katsayısının hesaplanması gerekmektedir.

Korelasyon katsayısı (r), iki değişken arasındaki ilişkinin ölçüsüdür ve -1 ve +1 arasında değişim gösterir.



(a)	(b)	(c)
r= -1	r= 0	r= 1
Mükemmel negatif İlişki	İlişki Yok	Mükemmel pozitif İlişki

Yukarıdaki saçılım grafikleri ;

(a) değişkenlerden birisinin artışına bağlı olarak diğerinde azalma olan doğrusal ilişki olduğu,

(b) iki değişken arasında ilişki olmadığı

(c) değişkenlerden birisindeki artışa bağlı olarak diğerinde de artış olan doğrusal ilişki olduğu

şeklinde açıklanır.

Korelasyonun katsayısının gücü ile ilgili olarak aşağıdaki tanımlamalar yapılmıştır:

0.00 - 0.25 Çok zayıf ilişki
0.26 - 0.49 Zayıf ilişki
0.50 - 0.69 Orta ilişki
0.70 - 0.89 Yüksek ilişki
0.90 - 1.0 Çok yüksek ilişki

Korelasyon katsayısı, örneklem büyüklüğünden etkilenmektedir. Küçük hacimli örneklerde, elde edilen korelasyon katsayısı büyük bile olsa istatistiksel olarak önemli bir değer olmayabilir. Dolayısıyla, elde edilen değerin hipotez testinin yapılması gereklidir.

1.14. REGRESYON ANALİZİ

Regresyon analizi, bilinen bulgulardan, bilinmeyen gelecekteki olaylarla ilgili tahminler yapılmasına izin verir. Regresyon, bağımlı ve bağımsız değişken(ler) arasındaki ilişkiyi ve doğrusal eğri kavramını kullanarak, bir tahmin eşitliği geliştirir. Değişkenler arasındaki ilişki belirlendikten sonra, bağımsız değişken(ler)in skoru bilindiğinde bağımlı değişkenin skoru tahmin edilebilir.

Bağımlı Değişken (y)

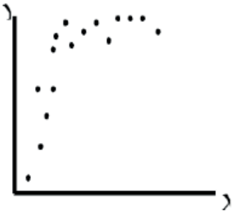
Bağımlı değişken, regresyon modelinde açıklanan ya da tahmin edilen değişkendir. Bu değişkenin bağımsız değişken ile ilişkili olduğu varsayılır.

Bağımsız Değişken (x)

Bağımsız değişken, regresyon modelinde açıklayıcı değişken olup; bağımlı değişkenin değerini tahmin etmek için kullanılır.

- Değişkenler arasında doğrusal ilişki olabileceği gibi, doğrusal olmayan bir ilişki de olabilir. Bu nedenle, saçılım grafiği yapılmadan (ilişki yok/doğrusal ilişki var/doğrusal olmayan ilişki var) ve değişkenler arasında korelasyon varlığına rastlanmadan regresyon analizine karar verilmemesi gerekir.
- Bu bilgiler doğrultusunda, tek/çok değişkenli doğrusal regresyon analizlerinin yanı sıra, tek/çok değişkenli doğrusal olmayan regresyon analizleri de mevcuttur.

(a) (+) yönlü Doğrusal İlişki **(b) (-) yönlü Doğrusal İlişki**



(c) Doğrusal Olmayan İlişki **(d) İlişki Yok**

Bağımlı değişken ile bağımsız değişken arasındaki doğrusal ilişkiyi açıklayan tek değişkenli regresyon modeli aşağıdaki gibidir:

$$y = a + bx$$

Burada;

y = Bağımlı değişkenin değeri

a = Regresyon doğrusunun kesişim değeri (Sabit değer)

b = Regresyon doğrusunun eğimi

x = Bağımsız değişkenin değerini göstermektedir.

1.15. ZAMAN SERİLERİ

1.15.1. Zaman Serisi Verileri

Eşit zaman aralıklarında, bir değişkene ait gözlemlerin oluşturduğu seriye "zaman serisi" denir. Bu gözlem sonuçlarının yıl, hafta ve gün gibi bir zaman vasfının şıklarına göre sıralanmasıyla elde edilen zaman serilerinde, zaman vasfının karşısında gözlem değerleri yer almaktadır ve bu şekilde istatistik araştırmasına konu olan olayın zaman içinde sergilediği değişkenlik gözlenmektedir. Zaman serisi verileri genellikle günlük, haftalık, aylık, üç aylık, altı aylık, yıllık ve daha uzun dönemli aralıklarla derlenir ve toplanır. Genel olarak zaman serisi, *T* örneklem büyüklüğü olmak üzere, Z_t, t= 1, 2, …, T biçiminde gösterilir. Buna göre ilk gözlemlenen veri Z_1; ikinci gözlemlenen veri Z_2; son gözlemlenen veri Z_T ile ifade edilir.

1.15.2. Zaman Serisi Verilerinin Türleri

Zaman içinde sürekli olarak kaydedilebilen verilere sahip serilere "sürekli zaman serileri", sadece belli aralıklarda elde edilebilen verilere sahip serilere de "kesikli zaman serileri" adı verilmektedir. Elektrik sinyalleri, voltaj, ses titreşimleri gibi mühendislik alanlarına ait seriler sürekli zaman serileri iken; faiz oranı, satış miktarı ve üretim gibi iktisadi seriler kesikli zaman serilerine örnektirler. Anlaşılacağı gibi zaman serileri farklı şekillerde karşımıza çıkmaktadırlar. Bunlardan bazıları aşağıdaki gibi sıralanabilmektedir:

- İktisadi ve finansal zaman serileri: İktisadi ve finansal verilerin büyük bir kısmı zaman serilerinden oluşmaktadır. Bunlara örnek olarak, günlük döviz kuru, hisse senedi getirisi, yıllık faiz oranı, enflasyon oranı gibi seriler verilebilir.

- Fiziksel zaman serileri: Özellikle coğrafya, meteoroloji, denizcilik bilimleri gibi alanlarda sürekli olarak gözlemlenen serilere örnek olarak günlük nem oranı, ortalama hava sıcaklığı, aylık yağış miktarı gibi ölçümler verilebilir.

- İşletme zaman serileri: Farklı dönemlerde gözlemlenen, işletmelerin satış analizleri, karlılık oranları, maliyet hesaplamaları gibi veriler işletme politikalarının belirlenmesi, yönlendirilmesi veya değiştirilmesinde etkin bir şekilde kullanılmaktadır.

- Demografik zaman serileri: Genellikle nüfus çalışmalarında ortaya çıkan ve hükümetlerin orta ve uzun vadeli planlamalarında kullanılan yıllık ortalama nüfus artışı, ölüm oranı, doğum oranı gibi veriler, ilgili zaman serilerine örnek olarak gösterilebilir. - Süreç kontrol verileri: Herhangi bir üretim sürecinin kalitesini ölçmek ve belirlenen bir hedeften ne kadar ve hangi yönde sapma gösterdiğini incelemek amacıyla zamana karşı grafik çizme yoluyla gözlemlenen veriler bu sınıfa girmektedir.

- İkili süreç verileri: Genellikle iletişim teknolojisinde ortaya çıkan ve gözlemlerinin 0 ve 1 gibi iki değerden sadece birisini aldığı zaman serisi verileridir. Örnek olarak, elektronik bir cihazın açma / kapama düğmesinin açık ve ya kapalı olma durumuna göre ölçeklendirilmesi verilebilir.

- Nokta süreç verileri: Belli bir dönem içinde rassal olarak ortaya çıkan bir olaylar dizisi biçiminde oluşan zaman serileridir. Bunlara örnek olarak, havayolu ulaşımında bir yolcu uçağının belli bir dönem içerisinde arızalanma, bakım ve onarım alma zamanlarının nokta süreç olarak ifade edilmesi verilebilir.

1.15.3. Zaman Serisi Verilerinin Bileşenleri

Belli bir zaman biriminde bir gözleme ait veriler incelendiğinde bunların bir takım dalgalanmaların etkisi altında kaldığı gözlemlenmiştir. Zaman serilerinin bileşenleri olarak da tanımlayabileceğimiz bu etkiler sırayla;

- Trend (T)

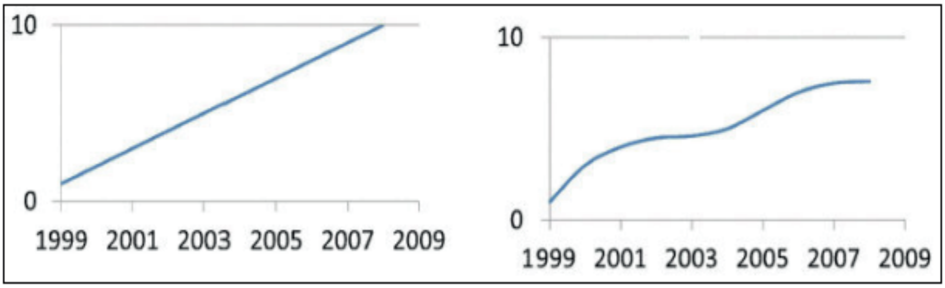
- Mevsimsel dalgalanmalar (M)

- Konjonktürel dalgalanmalar (K)

- Düzensiz hareketler (D)

olmak üzere dört temel faktörden kaynaklanmaktadırlar. Bu faktörlerden her birinin olay üzerindeki etkileri farklı yön ve şiddette olabileceği gibi, aynı yön ve şiddette de olabilmektedir. Dolayısıyla zaman serileri ile analiz yapıldığında bu faktörlerin etkilerinin araştırılması zorunludur. Bu bileşenler ayrıntılı şekilde aşağıda sunulmaktadır:

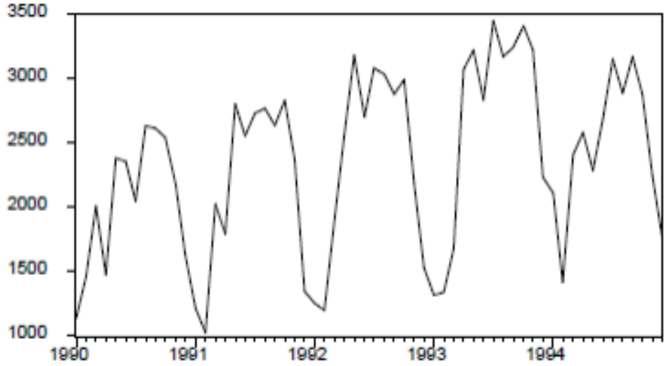
-**Trend (T):** Bir zaman serisinin uzun bir dönem içerisinde gösterdiği ana eğilime trend adı verilmektedir. Zaman serilerinde trend yapan kalıplar genellikle seride uzun süreli artışları veya azalışları yansıtmaktadır. Trend etkisine örnek olarak, demografik özelliklerdeki, coğrafi dağılımdaki, teknolojik gelişmelerdeki, tüketici zevk ve alışkanlıklarındaki ve fiyatlardaki değişmeleri vermek mümkündür. Etkilerin şiddetine bağlı olarak, artış ve azalış yönündeki değişmeler bazen artabilir, bazen de azalabilir. Yani, trend aynı kalmaz. Trend doğrusal olabileceği gibi eğrisel de olabilmektedir. Zaman içinde artış ve azalış göstermeyen, hemen hemen aynı düzeyde kararlılık gösteren serilerin trendi yoktur. İlgili trend çeşitleri Şekil 1'de görülmektedir.



***Şekil 1:** Doğrusal ve eğrisel trend sergileyen zaman serisi örneği*

Değişkenler arasında ekonometrik olarak anlamlı ilişkiler elde edebilmek için analizi yapılan serilerin güçlü bir trend taşımaması gerekir. Eğer değişkenlere ait zaman serilerinde trend bulunuyorsa, ilişki gerçek olmaktan çok sahte regresyon şeklinde ortaya çıkmaktadır. Birçok ekonometrik analizde ele alınan iki serinin de güçlü genel eğilimler (trend) taşıması nedeniyle değişkenler arasında anlamlı bir ilişki olmasa dahi yüksek bir R2 bulunmaktadır. Gözlenen yüksek R2 iki değişken arasındaki gerçek ilişkiden ziyade bu eğilimden kaynaklanmaktadır. Bu nedenle regresyonun gerçek bir ilişkiyi mi yoksa sahte bir ilişkiyi mi ifade ettiği, zaman serilerinin durağan olup olmamasıyla yakından ilgilidir.

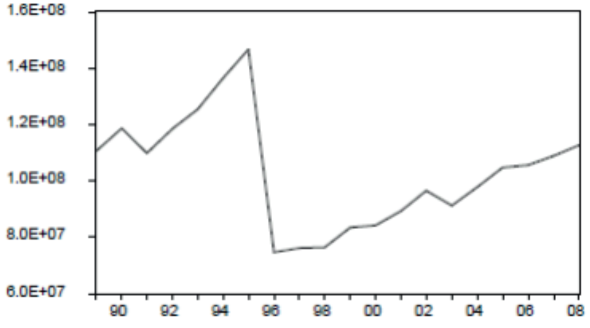
-Mevsimsel Dalgalanmalar (M): Mevsimlik dalgalanmalar, zaman serilerinde kolayca izlenebilen ve sık rastlanan bir etkidir. Periyodik hareketlerle kendini gösterir. Bir yıl ve daha az süre içinde gerçekleşen tam dairesel süreçte mevsim hareketlerinin verilere etkisini ifade eder. Genelde mevsimsel etkiler aylık dönemler itibariyle ortaya çıkar. Mevsimin etkisinde olan değişkenler yılın bazı dönemlerinde diğerlerine oranla daha yüksek veya daha düşük değerlere ulaşırlar. Satış rakamları, sıcaklık göstergeleri, turizm istatistikleri gibi değişkenlere ait verilerde mevsim etkisini görmek mümkündür. Mevsim etkisi değişimleri genellikle yaz aylarından kış aylarına geçişlerde kendilerini gösterirler. Kış aylarında enerji tüketiminin artması, yazın yağışların azalması, okulların tatil olduğu yaz aylarında işsizlik oranının artması da mevsimsel dalgalanmalara örnek olarak verilebilir.



***Şekil 2:** 1990:01 – 1994: 12 yıllarına ait çimento üretim miktarı*

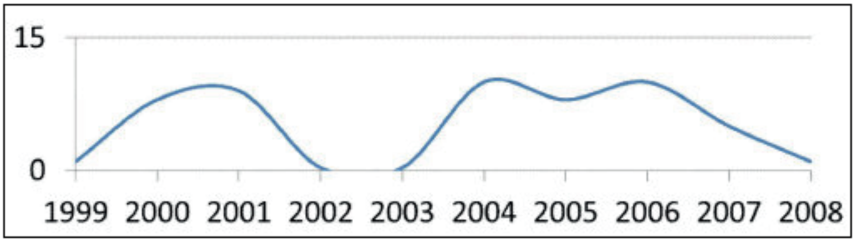
Mevsimsel dalgalanmalar sistematik etkileri olan faktörlerdir ve bu etkinin seriden arındırılması gerekir. Zaman serilerinde mevsimlik etkinin varlığını araştırmak için çeşitli yöntemler geliştirilmiştir. Bunlar, hareketli ortalamalar yöntemi, ortalama yüzde yöntemi, mevsim indeksi gibi yöntemlerdir. Bunların yanı sıra regresyon ve varyans analizi yöntemleri aracılığıyla da, zaman serilerinde mevsimsel etkinin olup olmadığı ortaya çıkarılabilir.

-Konjonktürel Dalgalanmalar (K): Konjonktürel dalgalanmalar, 2-10 yıl veya daha uzun bir dönemde serinin seyrinde oluşan değişmelerdir ya da zaman serisindeki dalgalanmalar bir yıldan daha uzun dönemi kapsar şekilde seyir izliyorsa bu gidişat konjonktürel dalgalanma olarak adlandırılır. Konjonktürel hareketler daha çok ekonominin veya sektörlerin refah ya da durgunluk (ekonomik kriz) dönemlerini içeren değişmelerdir. Refah dönemlerinde yatırımlar, üretimler, gelirler ve satışlar gibi ekonomik göstergeler bir süre için artış gösterir ve durgunluk dönemlerinde ise düşmeler baş gösterir. Genelde konjonktürel hareketler periyodik olmayan fakat 5 ile 8 yıllık dalgalanmalar ile tekrarlanır. Mevsimsel hareketlerde dönemler düzenli ve periyodik bir salınım gösterirken, konjonktürel hareketlerde dönemler düzensiz ve periyodik olmayan bir yapıdadır. Ayrıca konjonktürel hareketlerin ortalama uzunlukları mevsimsel dalgalanmalardan daha uzundur ve konjonktürün hacmi (genişliği) mevsimselliğe göre daha fazla bir değişkenliğe sahiptir.



***Şekil 3:** 1989 – 2008 yıllarına ait GSYİH verisi*

-Düzensiz Hareketler (D): En son faktör olan düzensiz hareketler, periyodik olmayan değişimleri gösterir. Varlığı önceden tahmin edilemeyen tesadüfi olayların ortaya çıkardığı dalgalanmalardır ve hata terimi ile ifade edilebilecek değişmelerdir. Örnek olarak, doğal felaketlerin etkisi ile verilerde oluşan artışlar ya da azalmalar verilebilir.



***Şekil 4:** Düzensiz hareketler sergileyen zaman serisi örneği*

Kısaca özetlemek gerekirse, zaman serileri trend mevsimsel dalgalanmalar, konjonktürel dalgalanmalar ve düzensiz hareketler gibi faktörlerin etkisi altında kalmaktadırlar. Bu bileşenler t sürecinde gözlenen bir değişkenin zaman serisi modelini tanımlamada kullanılır. Söz konusu tanımlamalar iki türlüdür:

1- **Çarpımsal Model: T x M x K x D**
2- **2- Toplamsal Model: T + M + K + D**

Zaman serileri sıraladığımız faktörlerden birini ya da hepsini içerebilir ve bir seride bu hareketlerden birinin etkisi, bir diğerinin içinde kendisini gösterebilir. Örneğin; mevsimsel dalgalanmalar yıllık zaman serilerinde bulunmaz. Sadece aylık ve 3 aylık gibi zaman serileri söz konusu olduğunda mevsim etkisinin araştırılması gerekmektedir. Bu, ayrıntılı şekilde aşağıdaki gibi ifade edilebilmektedir.

Yıllık zaman serisi: T x K x D veya T + K + D

Aylık veya 3 aylık zaman serisi: T x M x K x D veya T + M + K + D

Zaman serileri üzerinde gerçekçi bir analiz yapmak ve ileriye dönük doğru tahminlerde bulunmak için serilerin bu etkilerden arındırılması gerekmektedir.

1.15.4. Zaman Serisi Analizinin Amaçları

Zaman serilerini analiz etmenin amaçları aşağıdaki gibi sıralanabilmektedir:

-Zaman serisini bileşenlerine ayırarak özelliklerini ortaya çıkarmak: Amacımız ne olursa olsun öncelikle serinin özelliklerini araştırmamız gerekmektedir. Serinin grafiğini çizerek ilk bakışta hangi etkiler altında olduğu gözlemlenebilir. Daha sonra ise bazı analiz ve teknikler kullanarak gözlemlerimizden vardığımız sonucu kontrol etmeliyiz. Zaman serilerini birtakım etkilerden arındırmak için bileşenlerine ayırmak ve hangi etkiler altında olduğunu bulmak önemli bir aşamadır.

-Zaman serileri arasındaki ilişkiyi belirlemek: Elimizde birden fazla seri varsa bazı gelişmiş yöntem ve teknikler kullanarak ilgili zaman serileri arasındaki ilişkileri inceleyebiliriz. Birinde meydana gelen değişmelere bakarak, diğerleri üzerindeki etkileri açıklayabilir, aralarındaki kısa ve uzun dönem ilişkilerini analiz edebilir hatta aralarındaki ilişkinin yönünü ifade eden nedensel ilişkileri inceleyebiliriz. Bunları yapabilmemize olanak veren yöntem ve teknikler ilerleyen derslerde ayrıntısıyla incelenecektir.

-Geleceğe yönelik tahminlerde bulunmak: Zaman serileri ile ilgili analizlerin yapılma amacından biri de, gözlem kümesince temsil edilen gerçeğin anlaşılması ve zaman serisindeki değişkenlerin gelecekteki değerlerinin doğru bir şekilde tahmin edilmesidir. Bu amaçla da bazı gelişmiş tahmin teknikleri kullanmamız gerekmektedir.

-Kontrol: Seriyi oluşturan olayın mekanizmasını araştırarak, sistemi kontrol etmek ve sistemin istenen yönde gelişip gelişmediğini anlamaya çalışmak zaman serileri analizinin bir diğer amacıdır.

1.15.5. Durağanlık Kavramının Açıklanması

Zaman serileri ile yapılan ampirik çalışmalarda verilerin "durağan" olduğu varsayılmaktadır. Fakat zaman serilerinin önemli bir kısmı durağan değildir. Değişkenler arasındaki ilişkilerin anlamlı olabilmesi için kullandığımız zaman serilerinin durağan özellikler göstermesi gerekmektedir. İki değişken arasında anlamlı ilişkiler olmamasına rağmen aralarında ilişki varmış gibi görünebilir. Birini diğerine regres ettiğimizde aralarında ilişki olmasa bile yüksek bir R^2 değeri elde edilebilir. Bu durumda "sahte regresyon" problemi ortaya çıkmış olacaktır.

Bu sorunun kaynağı şudur: Her iki zaman serisi de güçlü bir trende sahipse aralarında gözlenen yüksek R^2 'nin sebebi; söz konusu iki değişken arasındaki doğrusal ilişki değil, bu güçlü trend ilişkisidir. Dolayısıyla durağan olmayan seriler ile analiz yapıldığında geleneksel *t* ve *F* testleri ile R^2 değeri yanıltıcı sonuçlar vermektedir. Zaman serisi yöntemlerini kullanarak durağan olmayan zaman serilerinden kaynaklanan bu problemin ne şekilde ortaya çıktığı analiz edilebilecektir. Bu kısımda öncelikle durağanlık kavramının açıklanmasında fayda vardır. Zaman serilerinin zaman içinde belli bir değere yaklaşmasına ve ya beklenen değeri etrafında dalgalanmasına "durağanlık" denir. Zaman serileri üzerinde etkili olan faktörlerin etkileri geçici olabileceği gibi kalıcı da olabilmektedir. Söz konusu kalıcı etkiler serilerin durağanlıktan sapmasına sebep olmakta ve serinin belli bir değere yaklaşmasını engellemektedir. Bunun yanı sıra bazı etkiler kısa dönemde etkili olmakla birlikte etkilerini belli bir dönem sonra kaybetmektedirler. Literatürde zamandan etkilenmeyen, ortalaması, varyansı ve kovaryansı sabit olan serilere zayıf durağan (ya da kovaryans durağan) seriler adı verilir ve bu, geniş anlamda durağanlık olarak bilinir. Güçlü durağanlıkta sonlu ortalama ve varyansa gerek yoktur. Zayıf durağanlık güçlü durağanlığa göre daha kısıtlı şartlar taşımaktadır. Herhangi bir Y_t serisinin durağan olması için gerekli şartlar aşağıdaki gibi sıralanabilir.

Sabit ortalama: $E(Y_t) = \mu$ (tüm t'ler için)

Sabit varyans: $Var(Y_t) = E(Y_t - \mu)^2 = \sigma^2$ (tüm t'ler için)

Gecikme mesafesine bağlı kovaryans: $\gamma_k = E[(Y_t - \mu)(Y_{t+k} - \mu)]$ (tüm t'ler ve tüm $k \neq 0$ için)

Burada *k*; gecikme mesafesidir. γ_k ise; aralarında k dönem fark bulunan iki Y değeri arasındaki kovaryanstır. Buradan hareketle durağan bir sürecin, ortalaması ve varyansı zaman içinde değişmeyen ve iki dönem arasındaki kovaryansın bakılan döneme değil de dönemler arasındaki uzaklığa bağlı olan bir süreç olduğu söylenebilir. Bu durağanlık türü, zayıf durağanlık olarak bilinmektedir. Böyle bir seri, kendi ortalaması çevresinde sabit genişlikte salınımlar gösterir. Bu özelliğe "ortalamaya dönüş" de denmektedir. Aynı zamanda bir sürecin ortak ve koşullu olasılık dağılımı da zaman içinde değişmiyorsa, bu seri güçlü durağan olarak nitelendirilmektedir. Uygulamalarda genellikle zayıf durağanlık kavramının araştırılması yeterli görülmektedir.

Bir serinin ortalama, varyans ve kovaryansı zamana bağlı olarak değişiyorsa durağan değildir. Serinin davranışı sadece ele alınan tahmin dönemi için geçerli olacaktır. Seri hakkında diğer dönemler için bir genelleme yapılamayacak ve değişken üzerindeki etkiler kalıcı olacaktır. Analizlerden anlamlı sonuçlar elde edebilmek için durağan olmayan serilerin durağan hale getirilmesi gerekmektedir. Bunu yapmak için kullanılan başlıca yöntemler; fark alma, logaritmasını alma veya logaritmasının farkını alma şeklinde sıralanabilir.

2. İSTATİSTİK ŞUBESİ İLE İLGİLİ BİLGİLER

Bu bölümde DHMİ Genel Müdürlüğü İstatistik Şubesi ile ilgili temel bilgiler ele alınacaktır.

2.1. Görev ve Sorumluluklar

2.1.1. İSTATİSTİK ŞUBE MÜDÜRLÜĞÜ **ŞUBE MÜDÜRÜ** YETERLİLİK KRİTERLERİ, GÖREV, YETKİ VE SORUMLULUKLARI:

1. AMAÇ:

Bu dokümanın amacı İstatistik Şube Müdürlüğü hizmetleri kapsamında yürütülen hizmetlerde yönetim ve organizasyon yönünden görev yapan personelin yeterlilik kriterleri, görev tanımlarının, yetkilerinin ve sorumluluklarının belirlenmesidir.

2. YETERLİLİK KRİTERLERİ:

2.1. 657 sayılı Devlet Memurları Kanunu hizmet şartlarına uygun olmak,

2.2. Atanacağı görevin liyakat şartlarını taşımak,

2.3. DHMİ Personeli Görevde Yükselme ve Unvan Değişikliği Yönetmeliği şartlarını taşımak,

3. GÖREV VE SORUMLULUKLAR:

3.1. 5429 Sayılı "Türkiye İstatistik Kanunu" ve 23.12.2011 tarihli ve 2011/2630 Sayılı Bakanlar Kurulu Kararı ile kabul edilen "Resmi İstatistik Programlarının öngördüğü gereklerin ve sorumlulukların yerine getirilmesini sağlar, 3.2. Verilerin, belirlenmiş kontrol kriterlerine göre kontrol ve geçerliliğinin sağlanmasını, analiz edilmesini, değerlendirilmesini ve arşivlenmesini sağlar,

3.3. Kısa-Orta ve Uzun dönemli uçak, yolcu ve yük trafiğinin istatistiksel tahminlerinin üretilmesi, üretilen tahminlerin raporlanma aşamasında metodolojik özet ve varsayımlarının yayımlanmasını sağlar,

3.4. İstatistiki bilgilerin yurtiçi ve yurtdışındaki ilgili mercilere gönderilmesi ve belirli sürelerle yayımlanmasını sağlar,

3.5. Kullanılan veri toplama, değerlendirme sistem ve metotlarının, ihtiyaçlar ve yenilikler doğrultusunda geliştirilmesini sağlar,

3.6. ICAO İstatistik Programına uyum çerçevesinde üretilen istatistiklerin diğer uluslararası standartlara da uygunluğunu sağlamak yönünde gerekli çalışmaların yapılmasını sağlar,

3.7. Benzer istatistikleri üreten ulusal ve uluslararası kuruluşların yayınlarının takip ve analiz edilmesini sağlar,

3.8. Uluslararası sivil havacılık alanında uygulanan istatistiki veri çeşitleri ile istatistik teknik ve yöntemlerinin takip edilmesi, incelenmesi ve bunlara göre gerekli düzenlemelerin yapılmasını sağlar,

3.9. İstatistiki göstergeler ile sektörel değerlendirme raporlarının hazırlanmasını sağlar,

3.10. Kurum içi istatistiki veri taleplerinin karşılanması, Makam’a karar destek amaçlı raporların hazırlanmasını sağlar,

3.11. İstatistik Yıllığı hazırlık çalışmalarının yürütülmesinde gerekli koordinasyonu sağlar, 3.12. Müdürlük personelinin sicil işlemlerinin yürütülmesi, performans durumu ve disiplin işlemlerinin Başkanlığa arz edilmesini yapar,

3.13. Müdürlük personelinin günlük yoklamasını idari işlere bildirir,

3.14. Başkanlıkça verilen diğer görevleri yerine getirir,

2.1.2. İSTATİSTİK ŞUBE MÜDÜRLÜĞÜ **ŞEF** YETERLİLİK KRİTERLERİ, GÖREV, YETKİ VE SORUMLULUKLARI:

1. AMAÇ:

Bu dokümanın amacı İstatistik Şube Müdürlüğü hizmetleri kapsamında yürütülen hizmetlerde yönetim ve organizasyon yönünden görev yapan personelin yeterlilik kriterleri, görev tanımlarının, yetkilerinin ve sorumluluklarının belirlenmesidir.

2. YETERLİLİK KRİTERLERİ:

2.1. 657 sayılı Devlet Memurları Kanunu hizmet şartlarına uygun olmak,

2.2. DHMİ Personeli Görevde Yükselme ve Unvan Değişikliği Yönetmeliğinde belirtilen şartları taşımak.

2.3. DHMİ Personel Sicil Yönetmeliğinde belirtilen şartları taşımak.

3. GÖREV VE SORUMLULUKLARI

3.1. Müdürlük bünyesinde birden fazla birimi ilgilendiren konularda birimler arası koordineyi sağlamak, 3.2. Şube Müdürünün bulunmadığı zamanlarda şube içindeki iş ve işlemlerin koordinasyonunu sağlamak,

3.3. İstatistik Yıllığı hazırlık çalışmalarını yürütmek,

3.4. ALO DHMİ başvurularında İstatistik Şubesi adına koordineyi sağlamak,

3.5. Resmi yazışma kurallarını takip etmek, gerektiğinde yeniden düzenlenmesini sağlamak,

3.6. Başkanlıkça/Müdürlükçe verilen diğer görevleri yerine getirmek,

2.1.3. İSTATİSTİK ŞUBE MÜDÜRLÜĞÜ **İSTATİSTİKÇİ/ENDÜSTRİ MÜHENDİSİ** YETERLİLİK KRİTERLERİ, GÖREV, YETKİ VE SORUMLULUKLARI:

1. AMAÇ:

Bu dokümanın amacı İstatistik Şube Müdürlüğü hizmetleri kapsamında yürütülen hizmetlerde yönetim ve organizasyon yönünden görev yapan personelin yeterlilik kriterleri, görev tanımlarının, yetkilerinin ve sorumluluklarının belirlenmesidir.

2. YETERLİLİK KRİTERLERİ:

2.1. 657 sayılı Devlet Memurları Kanunu hizmet şartlarına uygun olmak,

2.2. DHMİ Personeli Görevde Yükselme ve Unvan Değişikliği Yönetmeliğinde belirtilen şartları taşımak,

2.3. DHMİ Personel Sicil Yönetmeliğinde belirtilen şartları taşımak.

3. GÖREV VE SORUMLULUKLARI

3.1. Verilerin, belirlenmiş kontrol kriterlerine göre kontrol ve geçerliliğinin sağlanması, analiz edilmesi, değerlendirilmesi çalışmalarını yapmak,

3.2. İstatistik Yıllığı hazırlık çalışmalarına katılım sağlar,

3.3. Kısa-Orta ve Uzun dönemli uçak, yolcu ve yük trafiğinin istatistiksel tahminlerinin üretilmesi, üretilen tahminlerin raporlanma aşamasında metodolojik özet ve varsayımlarının yayımlanması çalışmalarına katılım sağlar, İstatistiki bilgilerin yurtiçi ve yurtdışındaki ilgili mercilere gönderilmesi ve belirli sürelerle yayımlanması çalışmalarına katılım sağlar,

3.4. Kullanılan veri toplama, değerlendirme sistem ve metotlarının, ihtiyaçlar ve yenilikler doğrultusunda geliştirilmesi çalışmalarına katılım sağlar,

3.5. ICAO İstatistik Programına uyum çerçevesinde üretilen istatistiklerin diğer uluslararası standartlara da uygunluğunu sağlamak yönünde gerekli çalışmaların yapılmasına katılım sağlar,

3.6. Benzer istatistikleri üreten ulusal ve uluslararası kuruluşların yayınlarının takip ve analiz edilmesini sağlar,

3.7. Uluslararası sivil havacılık alanında uygulanan istatistiki veri çeşitleri ile istatistik teknik ve yöntemlerinin takip edilmesi, incelenmesi ve bunlara göre gerekli düzenlemelerin yapılmasını sağlar,

3.8. İstatistiki göstergeler ile sektörel değerlendirme raporlarının hazırlanmasını sağlar,

3.9. Kurum içi istatistiki veri taleplerinin karşılanması, Makam’a karar destek amaçlı raporların hazırlanmasını sağlar, 3.10. ALO DHMİ başvurularında İstatistik Şubesi adına koordineyi sağlamak,

3.11. Başkanlıkça/Müdürlükçe verilen diğer görevleri yerine getirmek

2.3.1. İSTATİSTİK ŞUBE MÜDÜRLÜĞÜ **BİLGİSAYAR İŞLETMENİ/MEMUR** YETERLİLİK KRİTERLERİ, GÖREV, YETKİ VE SORUMLULUKLARI:

1. AMAÇ:

Bu dokümanın amacı İstatistik Şube Müdürlüğü hizmetleri kapsamında yürütülen hizmetlerde yönetim ve organizasyon yönünden görev yapan personelin yeterlilik kriterleri, görev tanımlarının, yetkilerinin ve sorumluluklarının belirlenmesidir.

2. YETERLİLİK KRİTERLERİ:

2.1. 657 sayılı Devlet Memurları Kanunu hizmet şartlarına uygun olmak,

2.2. DHMİ Personeli Görevde Yükselme ve Unvan Değişikliği Yönetmeliğinde belirtilen şartları taşımak,

2.3. DHMİ Personel Sicil Yönetmeliğinde belirtilen şartları taşımak.

3. GÖREV VE SORUMLULUKLARI

3.1. Verilerin, belirlenmiş kontrol kriterlerine göre kontrol ve geçerliliğinin sağlanması, analiz edilmesi, değerlendirilmesi çalışmalarını yapmak,

3.2. İstatistik Yıllığı hazırlık çalışmalarına katılım sağlar, 3.3. Kullanılan veri toplama, değerlendirme sistem ve metotlarının, ihtiyaçlar ve yenilikler doğrultusunda geliştirilmesi çalışmalarına katılım sağlar,

3.4. ICAO İstatistik Programına uyum çerçevesinde üretilen istatistiklerin diğer uluslararası standartlara da uygunluğunu sağlamak yönünde gerekli çalışmaların yapılmasına katılım sağlar,

3.5. Benzer istatistikleri üreten ulusal ve uluslararası kuruluşların yayınlarının takip ve analiz edilmesini sağlar,

3.6. İstatistiki göstergeler ile sektörel değerlendirme raporlarının hazırlanmasını sağlar,

3.7. Kurum içi istatistiki veri taleplerinin karşılanması, Makam'a karar destek amaçlı raporların hazırlanmasını sağlar,

3.8. ALO DHMİ başvurularında İstatistik Şubesi adına koordineyi sağlamak,

3.9. Başkanlıkça/Müdürlükçe verilen diğer görevleri yerine getirmek

2.4. HAVAYOLU UÇAK VE YOLCU TRAFİĞİ İSTATİSTİKLERİ VE TAHMİNLERİ ÜRETİM SÜRECİ

Veri kaynakları: Kuruluş idari kayıtları ve kurum dışı veri kaynakları olmak üzere iki türlüdür. Kuruluş idari kayıtları, Devlet Hava Meydanları İşletmesi Genel Müdürlüğü uçuş verileri, kurum dışı veri kaynakları, Yer Hizmet Kuruluşları ve Havayolu Şirketleri yolcu, yük (Kargo + Posta + Bagaj) verileridir.

Yöntem: Kuruluş idari kayıtları ile kurum dışı veri kaynaklarının birleştirilmesi ile elde edilen veriler, tam sayım esasına göre kâğıt/manyetik ortamda sistematiğine uygun olarak derlenmektedir. Elde edilen toplulaştırılmış veriler, muhtelif kontrol kriterleri ile geçerli kılınmakta, kesinleştirilerek yayımlanmaktadır.

Açıklama: İlk verinin ve revize edilmiş verinin tanımlanması: "ön bilgilere göre" ve "kesinleşmiş" başlıkları altında tanımlanmaktadır.

Ön Bilgilere Göre İstatistikler: Havalimanlarından her ayın ilk 3 günü içinde gönderilmesi talep edilen ve kesin olmayan bir önceki aya ait uçak, yolcu ve yük verisidir.

Kesinleşmiş Değerler: Belirli periyotlarla (aylık,3 aylık, yıllık) veri tabanında izlenen tüm değişkenleri içeren, çeşitli iç kontrolleri yapılmış, tutarlılığı ve geçerliliği sağlanmış verinin nihai halidir.

Kısa Dönemli Tahmin Verileri: Nisan ve Eylül aylarında olmak üzere yılda iki kez tahmin çalışması yapılmaktadır. Tahmin çalışmasında Ön Bilgilere Göre İstatistiklerinin o aya kadar gerçekleşen aylık verileri kullanılarak yılsonu verileri tahmin edilmektedir. Yılsonu tahminlerinin oluşmasını müteakiben gelecek 3 yılda gerçekleşmesi öngörülen uçak ve yolcu trafiği tahmini yapılmaktadır. Tahmin çalışmasında Zaman Serileri Analizi yöntemi kullanılmaktadır.

2.5. Uçuş Amaçları

Havayolu istatistikleri üretim sürecinde kullanılan uçuş amaçları, uçuş amacı tipleri, uçuş amaçlarına göre tarife durumlarının yer aldığı tablo aşağıda yer almaktadır:

UÇUŞ AMACI LİSTESİ

UÇUŞ KODU	UÇUŞ AMACI ADI	TARİFE BİLGİSİ	UÇUŞ KATEGORİSİ	AÇIKLAMA
110	YOLCU	Tarifeli	T	
120	KARGO-GENEL KARGO	Tarifeli	T	
121	KARGO-CANLI HAYVAN NAKLİ	Tarifeli	T	
122	KARGO-BOZULABİLİR KARGO	Tarifeli	T	
123	KARGO-TEHLİKELİ MADDE NAKLİ	Tarifeli	T	
124	KARGO-POSTA	Tarifeli	T	Dünya Posta Birliği (UPU) kurallarına göre her türlü yazışma ve diğer malzemelerin bütününün taşındığı kargo hava araçlarına
125	KARGO-YOLCU/KABİNDE TAŞINAN KARGO	Tarifeli	T	Yolcu uçağı içerisinde yolcu yerine kargo (yük,valiz vb.) taşınan hava araçlarına
210	YOLCU-TURİSTİK CHARTER	Tarifesiz	T	
212	TIBBİ-AMBULANS (ÜCRET ALINAN)	Tarifesiz	G	Muafiyet kapsamı dışında hasta veya organ nakli yapan yerli ve yabancı hava araçlarına
214	YOLCU-HAC ve UMRE UÇUŞU	Tarifesiz	T	
215	YOLCU-ÖZEL YOLCU (TUTUKLU,GÖÇMEN..vb)	Tarifesiz	T	
217	YOLCU-TAHLİYE	Tarifesiz	T	Kurtarma/Tahliye uçuşlarına
220	KARGO-GENEL KARGO	Tarifesiz	T	Uçuş amacı kod listesinde özel olarak tanımlanmış kargo içerikleri (örn: 221,222,223,224,225, 226 uçuş amacı kodlarında belirtilen kargolar dışındaki) dışında kalan kargoları taşıyan hava araçlarına
221	KARGO-CANLI HAYVAN NAKLİ	Tarifesiz	T	Permisinde canlı hayvan nakli yazan hava araçlarına
222	KARGO-BOZULABİLİR KARGO	Tarifesiz	T	Permisinde bozulabilir ürün (Kuru ve Yaş Gıda, Sebze, Bitki ve Bitki türleri) yazan hava araçlarına
223	KARGO-TEHLİKELİ MADDE NAKLİ	Tarifesiz	T	Permisinde Sağlığa, güvenliğe, eşyaya ve çevreye zarar verebilme riskine sahip ve teknik talimatlarda bu konu kapsamında olan ya da bu kapsamda sınıflandırılmış olan maddeleri taşıyan hava araçlarına
224	KARGO-POSTA	Tarifesiz	T	Sadece kargo hava aracında Dünya Posta Birliği (UPU) kurallarına göre her türlü yazışma ve diğer malzemelerin bütününü taşıyan hava araçlarına
225	KARGO-YOLCU/KABİNDE TAŞINAN KARGO	Tarifesiz	T	Yolcu uçağı içerisinde yolcu yerine kargo (yük,valiz vb.) taşınan hava araçlarına
226	KARGO-BAVUL TİCARETİ	Tarifesiz	T	Permisinde Charter Bagaj Uçuşu yazan hava araçlarına
230	TEKNİK İNİŞ-KALKIŞ	Tarifesiz	S	Herhangi fiziksel bir yük değişimi olmayan, ticari amaç dışında veya yakıt almak gibi teknik sebeplerle uçuş yapan hava araçlarına (Teknik iniş yapan bütün uçaklar teknik iniş yaptıkları havalimanından kesinlikle yolcu, yük ve posta alamazlar.)
240	İNTİKAL UÇUŞ (BOŞ UÇUŞ)	Tarifesiz	S	
250	İŞ UÇUŞU	Tarifesiz	S	
260	DIVERT-KALKTIĞI MEYDANA DÖNÜŞ	Tarifesiz	S	
270	DIVERT-YEDEK MEYDANA İNİŞ	Tarifesiz	S	
310	KAMU TESCİLLİ	Tarifesiz	S	Kamu, Kurum ve Kuruluşlarına ait Hava Araçlarıl ve OR-0177 Tecilli Uçak
320	VIP-DEVLET UÇUŞU (ÜCRETSİZ)	Tarifesiz	S	Sadece Devlet Başkanını taşıyan Hava Araçlarına
321	RESMİ OLMAYAN VIP (ÜCRET ALINAN)	Tarifesiz	G	Devlet Hava Araçları dışında kalan ve Sivil Havacılık Genel Müdürlüğü tarafından VIP uçuşu olarak permisi bulunmayan hava araçlarına (yolculuk esnasında içinde halihazırda görev yürüten Bakan veya üst düzey bürokrat bulunsa dahi)
330	ASKERİ- YABANCI ÜCRETLİ	Tarifesiz	A	
331	ASKERİ-T.C. (ÜCRETSİZ)	Tarifesiz	A	
332	ASKERİ- YABANCI ÜCRETSİZ	Tarifesiz	A	ABD ve PAKİSTAN
411	ORMAN YANGIN SÖNDÜRME	Tarifesiz	S	Yangın Söndürme (Özel Şirketler)
413	EMNİYET GENEL MÜDÜRLÜĞÜ (ÜCRETSİZ)	Tarifesiz	S	
415	ÖZEL UÇUŞ (TİCARİ OLMAYAN)	Tarifesiz	S	
420	HAVA TAKSİ (TİCARİ)	Tarifesiz	G	
421	KKTC (Yurtdışı bağlantılı yolcu/yük almayan T.C. Tescilli)	Tarifesiz	T	
440	EĞİTİM UÇUŞU (TİCARİ AMAÇLI UÇUŞ EĞİTİMİ)	Tarifesiz	S	
450	TEST,TECRÜBE UÇUŞU (ÜCRETSİZ)	Tarifesiz	S	Bakım sonrası aynı havalimanı kullanılarak yapılan uçuşlar
460	UÇUŞ KONTROL (FLIGHT CHECK) (DHMİ) (ÜCRETSİZ)	Tarifesiz	S	Sadece DHMİ Hava Araçları
470	İNSANİ YARDIM (ÜCRETSİZ)	Tarifesiz	S	
471	SİVİL UÇAKLA ASKER SEVKİYATI	Tarifesiz	T	
480	ARAMA KURTARMA (ÜCRETSİZ)	Tarifesiz	S	
491	PANO ÇEKİMİ/DÖVİZ ATIMI	Tarifesiz	G	
492	FOTOĞRAF-FİLM ÇEKİMİ	Tarifesiz	G	
493	PARAŞÜT ATLAYIŞI	Tarifesiz	S	
494	MADEN ARAMA	Tarifesiz	S	
495	AKROBASİ UÇUŞU	Tarifesiz	S	
496	MOTORLU YELKEN KANAT UÇUŞU	Tarifesiz	S	
497	SU UÇUŞU-ZEPLİN	Tarifesiz	S	
498	BALON UÇUŞU	Tarifesiz	S	
499	SPORTİF FAALİYETLER	Tarifesiz	S	
500	IHA (DRONE UÇUŞLARI)	Tarifesiz	S	
	T - Ticari			
	G - Hava Taksi			
	G - Diğer Ticari Genel Havacılık			
	S - Ticari Olmayan Genel Havacılık			

88

2.6. Havalimanları ICAO ve IATA Kodları

Ülkemizde yer alan 56 havalimanının dörtlü kodları(ICAO) ve üçlü kodları(IATA) aşağıda yer almaktadır.

ICAO Kod	IATA Kod	Meydan Adı	ICAO Kod	IATA Kod	Meydan Adı
LTBA	ISL	İstanbul Atatürk	LTBY	AOE	Eskişehir Hasan Polatkan
LTFM	IST	İstanbul	LTCW	YKO	Hakkari Yüksekova Selahaddin Eyyubi
LTFJ	SAW	İstanbul Sabiha Gökçen	LTDA	HTY	Hatay
LTAC	ESB	Ankara Esenboğa	LTCT	IGD	Iğdır Şehit Bülent Aydın
LTBJ	ADB	İzmir Adnan Menderes	LTFC	ISE	Isparta Süleyman Demirel
LTAI	AYT	Antalya	LTCN	KCM	Kahramanmaraş
LTFG	GZP	Gazipaşa Alanya	LTCF	KSY	Kars Harakani
LTBS	DLM	Muğla Dalaman	LTAL	KFS	Kastamonu
LTFE	BJV	Muğla Milas-Bodrum	LTAU	ASR	Kayseri
LTAF	ADA	Adana	LTBQ	KCO	Kocaeli Cengiz Topel
LTCG	TZX	Trabzon	LTAN	KYA	Konya
LTCE	ERZ	Erzurum	LTAT	MLX	Malatya
LTAJ	GZT	Gaziantep	LTCR	MQM	Mardin
LTCP	ADF	Adıyaman	LTCK	MSR	Muş Sultan Alparslan
LTCO	AJI	Ağrı Ahmed-i Hani	LTAZ	NAV	Kapadokya
LTAP	MZH	Amasya Merzifon	LTCB	OGU	Ordu-Giresun
LTBD	CII	Aydın Çıldır	LTFH	SZF	Samsun Çarşamba
LTFD	EDO	Balıkesir Koca Seyit	LTCL	SXZ	Siirt
LTBF	BZI	Balıkesir Merkez	LTCM	NOP	Sinop
LTCJ	BAL	Batman	LTAR	VAS	Sivas Nuri Demirağ
LTCU	BGG	Bingöl	LTCS	GNY	Şanlıurfa GAP
LTBR	YEI	Bursa Yenişehir	LTCV	NKT	Şırnak Şerafettin Elçi
LTBH	CKZ	Çanakkale	LTBU	TEQ	Tekirdağ Çorlu Atatürk
LTFK	GKD	Çanakkale Gökçeada	LTAW	TJK	Tokat
LTAY	DNZ	Denizli Çardak	LTBO	USQ	Uşak
LTCC	DIY	Diyarbakır	LTCI	VAN	Van Ferit Melen
LTCA	EZS	Elazığ	LTBZ	KZR	Zafer
LTCD	ERC	Erzincan Yıldırım Akbulut	LTAS	ONQ	Zonguldak Çaycuma

2.7. Anatablo Değişkenleri

Kurumumuz veri tabanında yer alan değişkenler ve açıklamaları aşağıdaki tabloda yer almaktadır.

Anatablo Değişkenlerinin Özellikleri		
Değişken	**Açıklama**	**type**
meydan	Trafiği raporlayan havalimanı	nvarchar(4)
tarih	Tarih (g/a/y)	datetime
cagri	Hava-yer muhaberesinde hava aracının tanımlanması için kullanılır.	nvarchar(8)
inis	Trafiğin kalkış veya iniş olduğunu gösterir	nvarchar(1)
hat	Trafiğin iç hat veya dış hat olduğunu gösterir	nvarchar(1)
pistno	Uçağın kullandığı pisti belirtir	nvarchar(3)
alcalma	İniş prosedürünü belirtir	nvarchar(5)
ifrvfr	Trafiğin IFR veya VFR olduğunu gösterir	nvarchar(1)
saat	Gerçek kalkış saati	nvarchar(5)
tahminisaat	Planlanan kalkış saati	nvarchar(5)
motorc	Motor çalışma saati	nvarchar(5)
motorctalep	Motor çalıştırma için izin verilen saat	nvarchar(5)
sirketicao	Havayolu şirketi ICAO kodu	nvarchar(6)
sirketiata	Havayolu şirketi IATA kodu	nvarchar(2)
permino	Trafiğin permi numarası	nvarchar(10)
uamaci	Uçuş amacı kodu Örn. 110 - yolcu	nvarchar(3)
uamacitipi	Uçuş amacı tipi Örn. T - ticari	nvarchar(1)
tarifelimi	Tariğin tarifeli olup olmadığını gösterir	nvarchar(1)
milliyetkodu	Havayolunun milliyet kodunu gösterir Örn. THY için T.C.	nvarchar(5)
milliyetadi	Havayolunun milliyetini gösterir Örn. THY için Türkiye	nvarchar(50)
kmeykod	Kalkış meydanının ICAO kodu	nvarchar(4)
vmeykod	Varış meydanının ICAO kodu	nvarchar(4)
dmeykod	Divert meydanın ICAO kodu	nvarchar(4)
tescil	Uçak tescil numarası	nvarchar(11)
tonaj	Uçak tonajı	int
koltukarz	Uçak koltuk sayısı	int
tesciltipi	Uçak tescil tipi	nchar(1)
utipiicao	Uçak tipi ICAO kodu	nvarchar(4)
utipiiata	Uçak tipi IATA kodu	nvarchar(3)
fgyeri	Fır sahası giriş noktası	nvarchar(5)
fcyeri	Fır sahası çıkış noktası	nvarchar(5)
fdevyer	Fır sahası transfer noktası	nvarchar(5)
fgzam	Fır sahasına giriş saati	nvarchar(5)
fczam	Fır sahasından çıkış saati	nvarchar(5)
yolcutrafik	Yer hizmetleri kodu	nvarchar(6)
transitdr	Direkt transit yolcu sayısı	int
yolcuucretli	Ücretli yolcu sayısı	int
yolcumuaf	Muaf yolcu sayısı	int
transfer	Transfer yolcu sayısı	int
transit	Transit yolcu sayısı	int
yolcutoplam	Toplam yolcu sayısı	int
kargoterm	Term. kargo miktarı	float
kargotransfer	Transfer kargo miktarı	float
kargotransit	Transit kargo miktarı	float
kargotoplam	Toplam kargo miktarı	float
postaterm	Term. mail miktarı	float
postatransfer	Transfer mail miktarı	float
postatransit	Transit mail miktarı	float
postatoplam	Toplam posta miktarı	float
bagajtoplam	Toplam bagaj miktarı	float
askeri	Tariğin askeri olup olmadığını gösterir	int