SPSS

Başlarken

SPSS

Bugünün iş dünyasında verilerden (sayılardan) kaçış yok.

- Bugünün dijital dünyasında ileri çalışmalar için herhangi bir olguya ait giderek artan miktarda veri toplanmakta, saklanmakta ve kullanılmaktadır.
- Her yerde veri kelimesini duymaktasınız.
- Veriler dünya hakkındaki gerçekleri göstererek raporlarlar.

SPSS

İstatistiğin Uygulanmasında Olası Hataların En Aza İndirilmesi İçin

TTDGA Süreci İzlenmelidir.

- Bir sorunu çözmek ya da bir hedefe ulaşmak için çalışmak istenilen verileri veya sorunu **Tanımlamak**
- Uygun kaynaklardan veri **Toplamak**
- Toplanan verileri **Düzenlemek**
- Grafiklerle verileri **Görselleştirmek**
- Sonuçları **Analiz** etmek ve yorumlamak

SPSS

TTDGA Sayesinde

- Sorunu özetler ve veriler görselleştirilebilir
- Bu verilerden sonuçlara ulaşılabilir
- Iş faaliyetleri hakkında güvenilir tahminler yapılabilir
- Iş süreçlerini iyileştirilebilir

Karar Vermede Sayılar Faydalıdır

- Yapılan araştırma sonucu gençlerin %54'ü A marka ürünü, %24'ü B marka ürünü ve %22'si C marka ürünü tercih etmektedir.
- 3000'den fazla öğrenciye yapılan bir araştırmada öğrencilerin %51'i boş zamanlarını sosyalleşme, eğlence ve diğer faaliyetlerle, % 19'u sınıf / laboratuvar çalışmalarına katılarak ve % 7'i de kitap okuyarak geçirmektedir.

Temel Kavramlar

DEĞİŞKEN

Bir ögenin ya da birimin özelliklerinin her biri.

VERİ

Bir değişken ile ilişkili öge ya da birim değerlerinin kümesi.

İSTATİSTİK

Karar vermede verileri yararlı bilgilere dönüştürmeye yardımcı yöntemler bütünü. Ham verilerden bilgi üretme.

İstatistik Olmadan Olmaz

- Araştırmadaki sayıların yararlı bilgiler olup olmadığını belirlemek
- Belirsizlik altında karar vermek
- Sınıflama ve kümeleme
- Nedensellik iddialarını doğrulamak (Değişimin ve farklılaşmanın sebebi)
- Büyük miktarda verilerin ortaya koyduğu kalıpları görmek
- Tahmin

İstatistik nedir?

- Üç çeşit yalan vardır:
  - Yalan
  - Kuyruklu yalan
  - İstatistik

Benjamin Disraeli

SPSS

Doğru Veri Toplama İstatistiksel Analiz İçin Çok Kritiktir

Doğru Veri

Doğru Yöntem

Doğru Analiz

Doğru Yorum

Doğru Bilgi

SPSS

İstatistiğin Konusu Olan Olaylar

İstatistik olaylarla ilgilenir.

Olayları ikiye ayırmak mümkündür.

Toplu olaylar

Bir çok faktör tarafından etkilenen olaylardır. İstatistiğin konusu kapsamındadır. (Enflasyon, başarıya etki eden faktörler…

Tekil olaylar

Tek bir faktör tarafından etkilenen olaylardır. İstatistiğin konusu kapsamında değildir. Belirli şartlar birleştiğinde daima aynı sonucu verir. (Kimyasal olaylar…)

İstatistiğin İki Dalı

İstatistik

Karar vermede verileri yararlı bilgilere dönüştürmeye yardımcı yöntemler bütünü. Ham verilerden bilgi üretme.

Tanımsal İstatistik

Verileri toplama, düzenleme, görselleştirme, analiz etme ve yorumlamadan oluşan süreci kapsar.

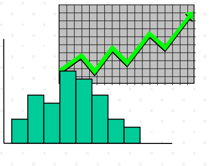
Çıkarımsal İstatistik

Küçük bir grubtan (örnekten) toplanan verileri kullanarak daha büyük bir grup (anakitle) hakkında sonuçlara varmak için kullanılır.

Tanımsal İstatistik

- Verilerin toplanması
  - Ör. Anket

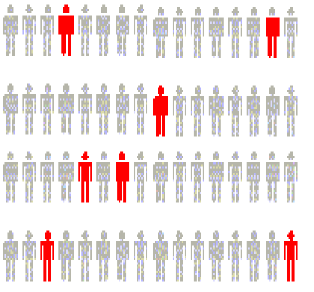
- Verilerin düzenlenmesi ve sunulması
  - Ör. Tablolar ve grafikler



- Karakteristik değerlerin hesaplanması
  - Ör. Örnek ortalaması = $\dfrac{\sum X_i}{n}$

Çıkarımsal İstatistik

- Tahmin
  - Ör. Anakitle ağırlık ortalamasının örnek ortalamasından yararlanarak tahmin edilmesi.
- Hipotez testleri
  - Ör. Anakitle ortalama ağırlığının 75 kg olduğu iddasının testi.



Örnekden hesaplanan sonuçlara göre anakitle hakkında karar verilir.

İstatistiğin Kullanım Alanları

- İşletmelerde; insan kaynakları, finansal analiz, Pazar araştırmaları, tedarik zinciri gibi…
- Psikloloji
- Sosyoloji
- Ekonomi
- Tıp
- Biyoloji
- Fizik
- Mühendislik
- vs.

İş analizi: İstatistik Değişen Yüzü

- Öngörülemeyen ilişkileri açığa çıkarmak için verileri analiz etmek ve keşfetmek için istatistiksel yöntemleri kullanın.
- Bir sistemin strateji, planlama ve operasyonlarını etkileyen optimizasyon modellerini geliştirmek için yönetim bilimi yöntemlerini kullanın.
- Verileri toplamak ve tüm yönleri ile verileri işlemek için bilgi sistemlerini kullanın. Aksi takdirde verileri etkin bir şekilde incelemek zor olur.

"Big Data" ve İş Analizi

- "Big Data" hala bulanık bir kavramdır.
- Yüksek hacimli verilerin otomatik olarak toplanabilmesi nedeniyle çok büyük veri setleri hızlı oranda artmaktadır.
- Eski istatistiksel tekniklerde büyük verilerin kullanımı pratik değildir.

İstatistikler, İşinizin Önemli Bir kısmıdır

- İşiniz giderek artan veri odaklı, analitik beceri gerektirir.
- Çalışmalar iş analitiğini uygulayan kuruluşların verimlilik, inovasyon ve rekabet gücünün arttığını göstermektedir.
- Hal Varian, Google Inc. Baş Ekonomist, "Önümüzdeki 10 yıl içinde en iyi meslek istatistik olacak. Ve ben şaka yapmıyorum."

Software (Bilg. Paket Programı) ve İstatistik

- Software, istatistiksel yöntemleri uygularken hesaplamalarda size yardımcı olacak programlardır.
- Microsoft Excel ile istatistiksel veri analizi yapabilirsiniz.
- Bir çok istatistik paket programı vardır. En bilinenleri;
  - SPSS
  - Minitab
  - R
  - Eviews
  - SAS

Verilerin Tanımlanması ve Toplanması

Temel Kavramlar

Değişken: Herhangi bir ögenin veya birimin herhangi bir özelliği

Veri (Data): Herhangi bir değişkenin birimlerine ilişkin değerler kümesi

İstatistik: Karar vermede verilerden yararlanarak, yararlı bilgiler üretmeye yardımcı yöntemler bütünü veya ham verilerden bilgi üretme süreci

Değişken Türleri

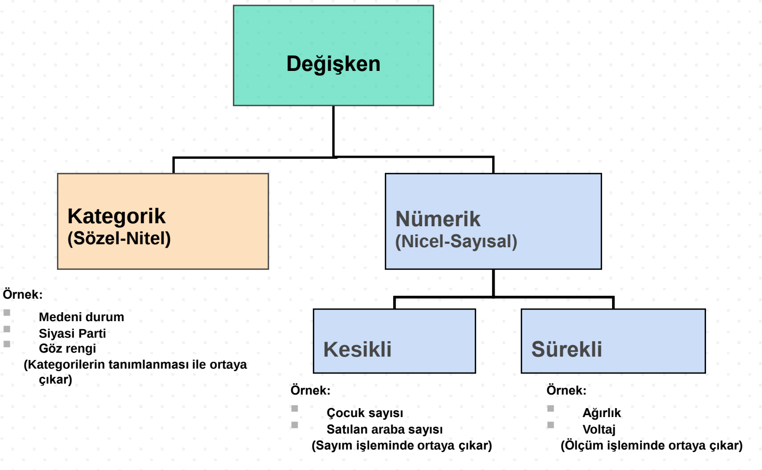
DCOVA

- **Kategorik** (***Kalitatif-Nesnel-Sözel***) variables have values that can only be placed into categories, such as "yes" and "no."
- **Sayısal** (***Kantitatif-Nicel***) variables have values that represent a counted or measured quantity.
  - **Kesikli değişken** sayım işleminde ortaya çıkar
  - **Continuous** değişken ölçüm işleminde ortaya çıkar

Tanımlar karışıklık ve hatalardan kaçınmak için çok önemlidir.

- Tanımlar ortak bir anlayışla açık ve kesin anlam sağlayan bir ifadedir.
- Tanımlar olmadığında iletişimsizlik ve hatalar oluşur.

Değişken Türleri



Değişken

Değişken : Gözlemden gözleme farklı değerler alabilen objelere, niteliklere ya da durumlara değişken denir. İstatistik birimlerinin sahip oldukları özellikler birer değişken olarak görülebilir.

- **Sürekli değişken :** Matematiksel olarak herhangi iki değeri arasında daima bir başka değeri bulunabilen değişken. (Örneğin: Uzunluk, ağırlık, yaş)
- **Süreksiz değişken :** Ölçüm birimleri daha küçük bölümlere bölünemediğinden ölçek üzerinde ayrı ayrı noktalar halinde yer alan değişken. (Örneğin Pekiyi 5, İyi 4 Orta 3 gibi)
- **Bağımsız değişken :** Başka bir değişkene bağlı olmadan değerler alabilen değiş-ken.
- **Bağımlı değişken :** Başka bir değişkene bağlı olarak değerler olabilen değişken.

Ölçüm Düzeyleri ve Ölçekler

		Örnek:
Oransal	Bu ölçme düzeyi, aralıklı ölçme düzeyinin bütün özelliklerine sahiptir. Aralıklı ölçme düzeyinden farklı olarak; oransal ölçekte sıfır gerçek yokluğu ifade eder ve iki sayı arasında oransal ilişki vardır.	Boy, yaş, haftalık tüketilen gıda miktarı…
Aralıklı	Bütün sıralı veri türlerini kapsar, değerler arasındaki uzaklık sabit büyüklüktedir, sayılar arasında oransal ilişki yoktur ve sıfırın gerçek bir yokluğu ifade etmez.	Hava sıcaklığı, standartlaştırılmış sınav skoru…
Sıralı	Veriler farklı sıralı kategorilere göre sınıflandırılır. Nominal ölçme düzeyi ile sıralı ölçme düzeyi arasındaki temel farklılık, sıralı ölçme düzeyi sınıfları arasında '… den daha iyi' ilişkisinin olmasıdır.	Hizmet kalite puanı, ürün memnuniyeti, akademik ünvan, S & P derecelendirmesi, Öğrenci bağıl notu (harf olarak)…
İsimsel	Nominal ölçekte veriler için hiçbir sıralama yoktur. Veriler farklı kategorilere göre sınıflandırılır.	Medeni durum, araba markası, facebook profili sahipliği, yatırım türü…

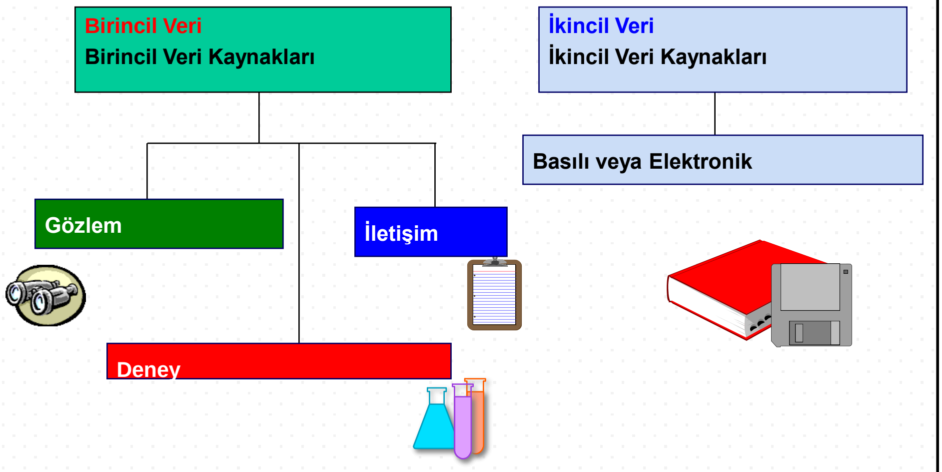
Doğru Veri Toplama İstatistiksel Analiz İçin Çok Kritiktir

- Yanlı(yanıltıcı), kusurlu veya hatalı verilerden kaçınmak
- Doğru olmayan verilerden elde edilen sonuçlar şüpheli ya da hatalı olacaktır.
- En iyi istatistiksel yöntemler bile, veri kusurlu olduğunda hatalı sonuçlara ve kararlara neden olacaktır.

Veri Kaynakları

- **Birincil Veriler**: Veri analizi yapacak kişi/kişiler tarafından toplanmış veriler
  - Siyasetle ilgili anketlerden elde edilen veriler
  - Deneylerden elde edilen veriler
  - Gözlemlerden elde edilen veriler
- **İkincil Veriler**: Veri analizi yapacak kişi(ler)den farklı kişiler tarafından toplanmış veriler
  - Nüfus sayımı verileri
  - İnternet veya basılı yayınlardaki yer alan veriler

Veri Toplama



SPSS

Birincil Veri Kaynakları

- Deney
- Gözlem
  - Doğal yolla
  - Doğal olmayan yolla
- İletişim
  - Anket
    - Mektup
    - Telefon, e-mail, Web
    - Kişisel
  - Mülakat

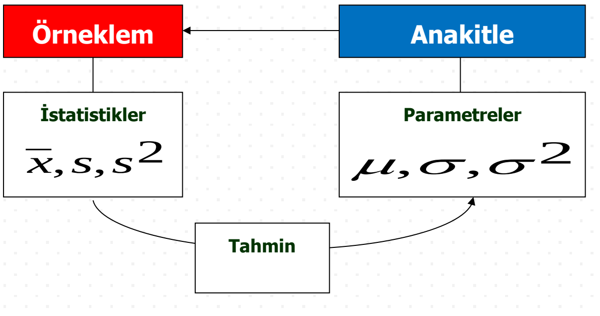
SPSS

Anakitle-Örneklem

Tanımlar

- **Anakitle**: Araştırmaya konu olan birimlerin oluşturduğu kümeye denir.
- **Örneklem**: Belli kurallara göre, belli bir anakitleden seçilmiş ve seçildiği anakitleyi temsil yeterliliği olan alt kümedir. (temsil gücü ve yeterlilik)
- **Parametre**: Anakitleyi tanımlamak için hesaplanan karakteristik değerler
- **İstatistik**: Örnekten hesaplanan karekteristik değerler
- **Tamsayım**: Anakitleyi oluşturan birimlerin tamamının sayılması
- **Örnekleme**: Bir araştırmanın konusunu oluşturan anakitlenin bütün özelliklerini yansıtan bir parçasının seçilmesi ve seçilen bu örneklemden yararlanarak hesaplanan karekteristik değerlerden (istatistik) yararlanarak anakitle karekteristik değerlerinin (parametre) tahmin edilmesi
- **Birim**: Anakitleyi oluşturan en küçük parça. Birim tekil olmak zorunda değildir.
- **Karekteristik Değer**: Herhangi bir verinin veya değişkenin özelliklerini tanımlamak için hesaplanan değerlerdir (aritmetik ortalama, mod, medyan, standart sapma vb...)

Parametre ve İstatistik



SPSS

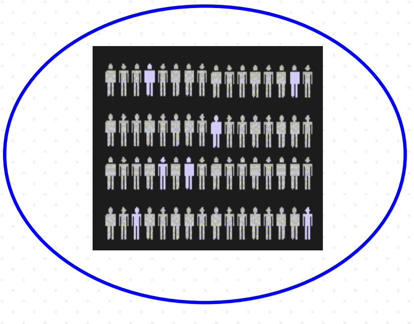
Parametre Ve İstatistik Simgeleri

DEĞER	PARAMETRE	İSTATİSTİK
Birim Sayısı	N	n
Aritmetik Ortalama	μ	$\bar{x}$
Standart Sapma	σ	s
Varyans	σ^2	s^2
Standart Hata	σ_μ	$s_{\bar{x}}$
Oran	π	p

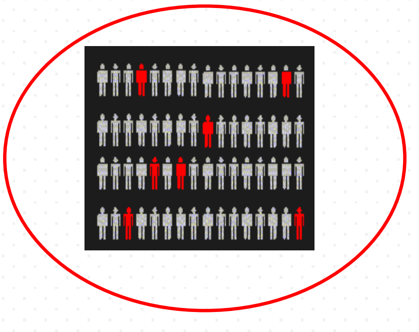
SPSS

Anakitle - Örneklem

Anakitle



Örneklem

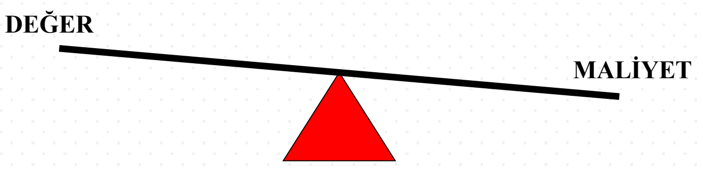


Araştırmaya konu olan birimlerin oluşturduğu kümedir

Bir anakitleden seçilmiş ve seçildiği anakitleyi temsil yeterliliği olan alt kümedir.

Niçin TAMSAYIM?

- Kesin sonuç



Niçin Örnekleme?

- Anakütleye ulaşılamaması
- Zaman
- Maliyet
- Kolaylık

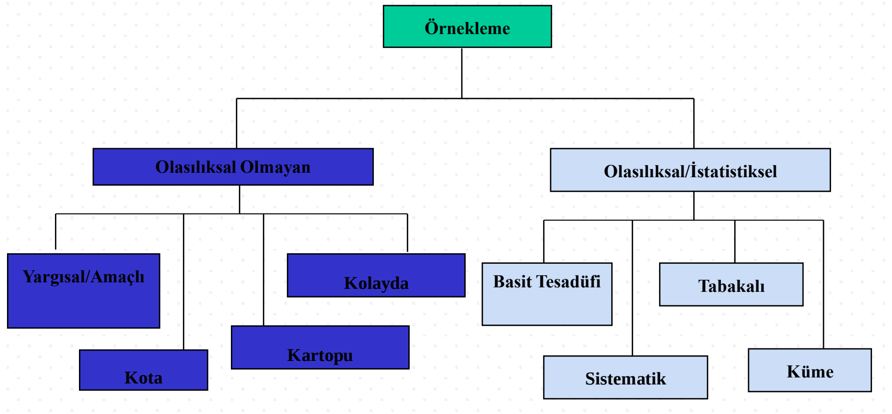
SPSS

Veri Kaynakları ve Veriler ile İlgili Bilinmesi Gerekenler

- Veri kaynağı yapılandırılmış veya yapılandırılmamış mı?
  - Yapılandırılmamış/Düzenlenmiş
  - Yapılandırılmış/Düzenlenmemiş
- Elektronik veriler hangi formatta yer almaktadır?
- Veriler nasıl kodlanmış?
  - Veriler kodlanmış mı?
  - Kodlanmış verilerin tekrar orijinal hale dönüştürümesi gerekir mi?
- Veri temizlemesi yapılmış mı?
  - Veri yanlışlıklar, Kayıp veriler, Uç değerler...
  - Tanımlanamayan veriler vs.

SPSS

Örnekleme Yöntemleri



Olasılıksal Olmayan Örnekleme Yöntemleri

Olasılıksal olmayan örnekleme, birimlerin seçiminde keyfi seçim yönteminin uygulandığı örnekleme yöntemleridir.

Kolayda (Gelişigüzel) Örnekleme: Kolayca ulaşılabilir birimleri seçmek suretiyle bir örnek oluşturulmaya çalışılır. Örneklemede birimlerinin seçimi görüşmeci tarafından doğru zamanda doğru yerde bulunan birimler, gönüllü katılımcılar arasından yapılır. Herhangi bir fakülteye gidip saptanacak sayıda rastlanan öğrenciyi örnekleme alma

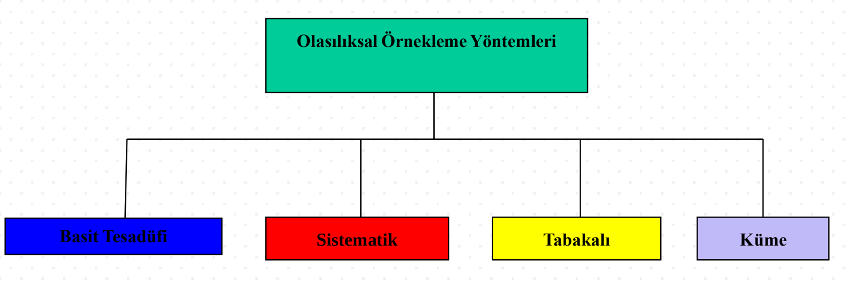
Yargısal Örnekleme: Birimlerin seçiminin araştırmacının amacına, arzu, düşünce ve deneyimlerine dayanarak yapılmasıdır. Meslek hastalıklarıyla ilgili yapılacak bir araştırmada örneklemin, meslek hastalıklarının tüm anakitle içinden değil, özellikle belli bir hizmet süresini aşmış ya da belli bir yaş sınırının üstündekiler arasından seçmesi gibi.

Kota Örneklemesi: Bu yöntemde tabakalı örnekleme yönteminde olduğu gibi anakitle alt tabakalara ayrılır. Her alt tabakanın temsili için kota konulur. Bu kota belirlenen tabakanın anakütleye oranına göre belirlenir. Kota örneklemede örneğe girecek elemanlar tesadüfen değil araştırmacını kendi isteğine göre belirlenir.

Kartopu Örneklemesi: Anakitleye ulaşmak mümkün olmadığında, ulaşabilen ilk birim belirlenir. Bu birimden elde edilen bilgilerle diğer birimlere ve bu şekilde zincirleme olarak anakitleyi temsil eden örneğe ulaşılmaya çalışır.

Olasılıksal Örnekleme Yöntemleri

- Olasılık örnekleme yöntemlerinde, birimler bilinen olasılıklara bağlı olarak seçilir.



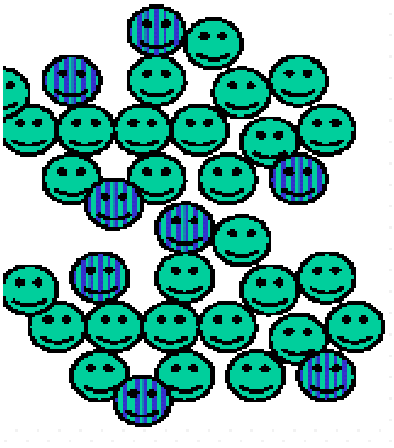
SPSS

Basit Tesadüfi Örnekleme

- Anakitlede yer alan her bir birimin örneklem kümesine girme şansı var ve bu şanslar eşit
- Seçimler iadeli olarak yapılabilir.
- Birimler tesadüfi sayılar tablosu veya bilgisayar yardımı ile çekilebilir.
- Anakütle incelenen konu açısından HOMOJEN yapıda olduğunda iyi sonuç verir
- Anakitleyi oluşturan birimlere birer numara verilir ve rasgele bu numaralar çekilir.

SPSS

Basit Tesadüfi Örnekleme



Rasgele Sayılar Tablosu

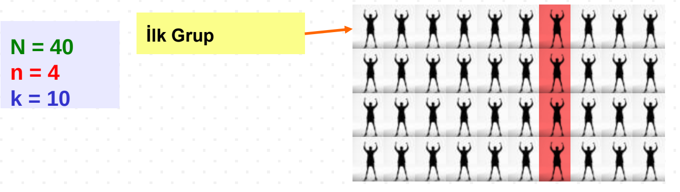
49280	88924	35779	00283	81163	07275
11100	02340	12860	74697	96644	89439
09893	23997	20048	49420	88872	08401

Örnekleme seçilen ilk 5 birim

Item # 492
Item # 808
Item # 892 -- iptal böyle bir gözlem yok
Item # 435
Item # 779
Item # 002

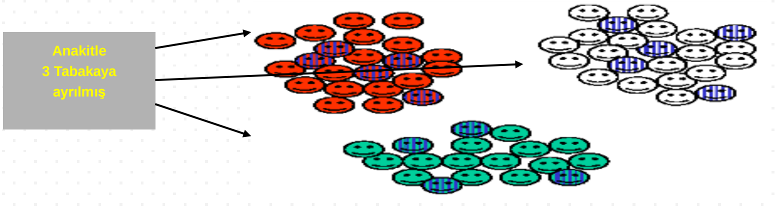
Sistematik Örnekleme

- Anakitle birimlerini kurala göre numaralandırılır (1...N) ve örneklem büyüklüğünü (n) belirlenir
- Örnekleme oranı k'yı (k=N/n) hesaplanır ve anakitle sıra numarasına göre her biri k birimden oluşan n gruba ayırılır.
- 1 ile k arasında rasgele bir rakam (s) seçilir.
- Her gruptaki s'inci sıradaki birim örneklem kümesine dahil edilir.



Tabakalı Örnekleme

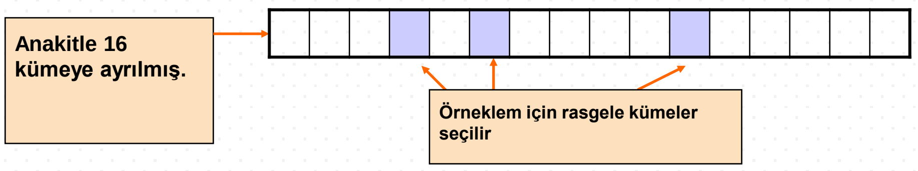
- Homojen olmayan anakitle birimleri, karakteristik özelliklerine göre tabaka denilen homojen alt gruplara ayrıştırılır
- Her tabakadan anakitle içindeki oranına bağlı olarak basit tesadüfi örnekleme yöntemi ile birimler seçilir
- Bu tabakalardan seçilen birimler birleştirilerek örneklem oluşturulur
- Çok yaygın kullanılan bu teknikte tabakalar kendi içinde homojen birbirleri arasında heterojendir.



SPSS

Küme Örnekleme

- Anakitle, anakitleyi temsil eden birden fazla “küme”ye bölünür
- Kümeler arasından basit tesadüfi örnekleme ile rasgele seçim yapılır
- Seçilen küme içindeki tüm birimler örneklem içinde yer alır veya seçilen kümelerdeki birimler başka bir örnekleme tekniğinde kullanılabilir
- Kümeler kendi içinde heterojen, kümeler arasında homojendir.



SPSS

Örnekleme Yöntemlerinin Karşılaştırılması

- Basit tesadüfi örnekleme ve sistematik örneklem
  - Kullanımı kolay
  - Anakitle özelliklerini için temsil sorunu yaşanabilir.
- Tabakalı örnekleme
  - Anakitleyi oluşturan ve farklı karakteristiklere sahip tüm birimlerin temsil edilmesini sağlar.
- Küme örnekleme
  - Daha düşük maliyetlidir.
  - Daha az etkindir. Etkinliğin ve temsiliyetin diğerleri kadar olabilmesi için daha yüksek örneklem büyüklüğüne ihtiyaç vardır.

SPSS

Hata Türleri

- Kapsam hatası

Listeden dışlananlar

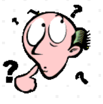
- Tepki hatası

Cevaplamayanları takip

- Örnekleme hatası

Tesadüfilik var

- Ölçme hatası



Kötü veya yönlendirici soru

SPSS

Hata Türleri

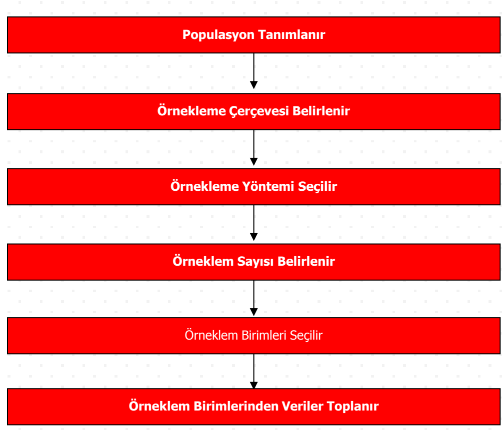
- Kapsama hatası veya seçim yanlılığı
  - Bazı gruplar çerçeve dışında kalması nedeniyle seçilme şansları yoktur
- Cevaplamama hatası
  - Cevap vermeyen insanlar cevap verenlerden farklı kanaate sahip olabilir
- Örnekleme hatası
  - Herzaman var olur ve örneklemden örnekleme değişkenlik gösterir.
- Ölçme hatası
  - Yanlış ve yönlendirici hazırlanmış sorular nedeniyle yanlı cevaplar olacaktır.

Örneklem Hataları

- Örneklem yöntemlerine göre yapılan tahminlerde iki çeşit hata vardır. Tesadüfi hatalar, örnek sayısı artırılarak giderilirken, sistematik hatalar örnekleme sürecinde ortay çıkar ve sonradan giderilmesi zordur. Bu hatalar:

1. Örnekleme yönteminin yanlış seçilmesi
2. Populasyonun yanlış tanımlanması
3. Örnek çerçevesinin yanlış belirlenmesi
4. Örnek birimlerinin doğru alınmayışından
5. Örnek büyüklüğünün yanlış belirlenmesinden kaynakalnır.

Örnekleme Süreci



SPSS

Örneklem Büyüklüğünün Saptanması

%95 güven aralığında %3, %5, %10 örnekleme hataları için karşılık gelen örneklem büyüklükleri yanda verilmiştir.

	$\alpha = 0.05$ için örneklem büyüklükleri					
	± %3 örnekleme hatası (d)		± %5 örnekleme hatası (d)		± %10 örnekleme hatası (d)	
Hedef Kitle Büyüklüğü (N)	p=0.5 q=0.5	p=0.8 q=0.2	p=0.5 q=0.5	p=0.8 q=0.2	p=0.5 q=0.5	p=0.8 q=0.2
100	92	87	80	71	49	38
250	203	183	152	124	70	49
500	341	289	217	165	81	55
750	441	358	254	185	85	57
1.000	516	406	278	198	88	58
2.500	748	537	333	224	93	60
5.000	880	601	357	234	94	61
10.000	964	639	370	240	95	61
25.000	1023	665	378	244	96	61
50.000	1045	674	381	245	96	61
100.000	1056	678	383	245	96	61
1.000.000	1066	682	384	246	96	61
100.000.000	1067	683	384	246	96	61

SPSS

Araştırma Sürecinin Aşamaları

SPSS

Araştırma Süreci

- Araştırma Konusunun Belirlenmesi
- Problemin Ortaya Konması
- Konuyla İlişkili Kaynakların Taranması
- Hipotezlerin Yazılması
- Araştırma Yönteminin Belirlenmesi
- Süre ve Olanakların Belirlenmesi
- Araştırmanın Sonuçlandırılması

SPSS

Araştırmanın Raporlaştırılması

- Araştırma planlanan şekilde gerçekleştirildikten sonra, araştırmanın verilerinin analizi sonucunda elde edilen bulgular yazılır ve bu bulguların yorumları yapılır.
- Bilimsel araştırma sürecinin son aşamasında ise araştırma raporu hazırlanır. Sosyal bilim araştırmaları genellikle dört ana bölümden ve çeşitli alt bölümlerden oluş-maktadır. Son yıllarda en yaygın kullanılan raporlaştırma biçimi şöyledir:
- I. GİRİŞ
Problem
- Kaynak Taraması Önem Hipotezler
- II. YÖNTEM
Evren ve Örneklem
Araştırma Modeli
- Verilerin Toplanması ve Analizi
- III. BULGULAR
- IV. SONUÇ (TARTIŞMA) Bulguların Yorumu Sınırlılıklar Öneriler

Verilerin Düzenlenmesi

- Veriler hangi yöntemle toplanırsa toplan-sın, elde edilen veriler genellikle istatistik-sel analize hazır değildir (bu veriler ham veri olarak adlandırılır)
- Bu verilerin analize uygun hale getirilmesi için düzenlenmeleri gerekir

Basit Seriler

- **Tanım**

  Elde edilecek ham verilerin küçükten büyüğe doğru sıralanması ile elde edilen serilere basit seri denir
- Basit bir seri birden fazla birimden oluşur ve birim sayısı *n* ile gösterilir
- Serinin i'ninci elemanı X_i değişkeni ile gösterilir

Basit Seriler

- Basit serinin toplam değeri X_i'lerin toplamına eşit olacağından,

Serinin toplam değeri= $X_1 + X_2 + ... + X_n = \sum_{i=1}^{n} X_i$

ÖRNEK:

Bir işletmede 25 işçiye verilecek çocuk paraları ile ilgili bir araştırma yapılmaktadır. İşçilerin çocuk sayıları aşağıda verilmiştir.

1,3,2,2,3,1,4,5,3,6,0,5,2,3,2,4,8,0,1,2,3,3,1,0,4

17

Basit Seriler

- Verilen değerleri basit seri şeklinde düzenleyelim:

  0,0,0,1,1,1,1,2,2,2,2,2,3,3,3,3,3,3,4,4,4,5,5,6,8

18

Frekans Serileri (Tasnif Edilmiş Seriler)

- İncelenecek birim sayısı arttıkça basit seriler çok uzun olacaktır
- Bu durumda çalışma kolaylığı sağladığı için frekans serileri kullanılır
- Frekans serilerinin değişkenin çok sayıda farklı değer almadığı durumlarda kullanılması daha uygun olur

19

Frekans Serileri (Tasnif Edilmiş Seriler)

- Düzenlenen bir frekans serisi iki sütundan oluşur:
  - Birinci Sütun: Değişkenin aldığı farklı değerler yer alır
  - İkinci Sütun: f_i ile gösterilir ve değişkenin aldığı değerlerin tekrar sayısı gösterilir

X_i	f_i
X_1	f_1
X_2	f_2
X_3	f_3
.	.
.	.
.	.
X_k	f_k

Değişkenin k sayıda farklı değer aldığı bir frekans serisi

20

Frekans Serileri (Tasnif Edilmiş Seriler)

- Frekans serisinin toplam değerini aşağıdaki gibi hesaplayabiliriz:

$$f_1X_1 + f_2X_2 + \ldots + f_kX_k = \sum_{i=1}^{k} f_iX_i$$

ÖRNEK: Daha önce basit seri olarak düzenlenen seriyi frekans serisi olarak düzenleyiniz

31

Frekans Serileri (Tasnif Edilmiş Seriler)

- Örnekteki basit serinin frekans serisi olarak düzenlenmiş hali:

X_i	f_i
0	3
1	4
2	5
3	6
4	3
5	2
6	1
8	1
	25

32

Grupandırılmış Seriler (Sınıflı Seriler)

- İncelenecek birim sayısının çok fazla olması durumunda frekans serileri de uzun sayı dizilerine dönüşür
- Bu durumda gruplandırılmış seriler düzenlenir
- Gruplandırılmış seriler iki sütundan oluşur:
  - Birinci Sütun: Sınıflar sütunudur
  - İkinci Sütun: frekans sütunudur

Gruplandırılmış Seriler (Sınıflı Seriler)

- Gruplandırılmış serilerde en önemli nokta sınıf sayısının kaç tane veya sınıf aralığının ne olacağının belirlenmesidir.
- Sınıf aralığının ne olması gerektiği konusunda bazı yazarlar çeşitli formüller önermektedir.
- Ancak bu formüller sadece birer öneridir, yani kesin değildir.

Gruplandırılmış Seriler (Sınıflı Seriler)

- Sınıf sayısının az olması serinin verdiği bilgilerin kaybına yol açacağından sınıf sayısının dörtten az olmaması
- Diğer yandan, çok fazla sınıf sayısının ise işlem zorluğu ve serinin yorumlanmasını zorlaştıracağı için sekizden fazla olmaması tavsiye edilir
- Önerilen kurallardan biri, sınıf sayısının birim serideki sayısının kare kökü olarak seçilmesidir

Gruplandırılmış Seriler (Sınıflı Seriler)

- Bir diğer kural ise Sturges kuralıdır:

$$\boxed{k = 1 + 3{,}322 \log\left(\sum f_i\right)}$$

k: Minimum Sınıf Sayısı

- Sınıf Genişliğinin hesaplanması

$$S = \frac{X_{max} - X_{min}}{k}$$ **(Eşit sınıf aralığı)**

S: Sınıf Genişliği

Gruplandırılmış Seriler (Sınıflı Seriler)

- Sınıf aralığı belirlendikten sonra
  1. İlk sınıfın alt sınır değeri sınıflar sütununa yazılır
  2. Alt sınır değere sınıf aralığı eklenerek üst sınır değeri elde edilir
  3. Her sınıfın üst sınıf değeri bir sonraki sınıfın alt sınır değerini oluşturur

Sınıflar değişkenin tüm değerlerini kapsayana dek işleme devam edilir

27

Gruplandırılmış Seriler (Sınıflı Seriler)

- **ÖRNEK:** Daha önce frekans serisi olarak düzenlenen örneği 2 eşit aralıklı sınıflı seri olarak düzenleyiniz

Frekans Serisi

X_i	f_i
0	3
1	4
2	5
3	6
4	3
5	2
6	1
8	1
	25

Sınıflandırılmış Seri

Sınıflar	f_i
0-2	7
2-4	11
4-6	5
6-8	1
8-10	1

*** Sınıflardaki frekanslar belirlenirken alt Sınıf değeri dahil üst sınıf değeri hariç tutulur

Gruplandırılmış Seriler (Sınıflı Seriler)

- Gruplandırılmış seriler ile işlem yapılırken frekansların her bir sınıf içerisinde eşit biçimde dağıldığı düşünülür
- Bu nedenle sınıf ortaları X_i olarak düşünülerek serinin toplamı frekans serilerinde olduğu gibi hesaplanır

Serinin toplam değeri= $f_1 m_1 + f_2 m_2 + ... + f_k m_k = \sum_{i=1}^{k} f_i m_i$

$m_i = \dfrac{X'_{alt} + X'_{üst}}{2}$ i'ninci sınıf ortası m_i sınıf alt sınır ve üst sınır toplamının ikiye bölünmesi ile bulunur

Gruplandırılmış Seriler (Sınıflı Seriler)

ÖRNEK: Bir önceki örnekteki gruplandırılmış serinin toplamını hesaplayınız

$f_1 m_1 + f_2 m_2 + ... + f_k m_k = \sum_{i=1}^{k} f_i m_i$

Sınıflar	f_i	m_i	$f_i m_i$
0-2	7	(0+2)/2=1	7
2-4	11	(2+4)/2=3	33
4-6	5	(4+6)/2=5	25
6-8	1	(6+8)/2=7	7
8-10	1	(8+10)/2=9	9

$\sum_{i=1}^{k} f_i m_i = 81$

Gruplandırılmış Seriler (Sınıflı Seriler)

- Gruplandırılmış serilerde sınıflardaki verilerin sınıf aralığında düzgün dağıldığının varsayılması hesaplanan ölçülerde sapmalara neden olacaktır.
- Basit seride, $\sum_{i=1}^{n} X_i = X_1 + X_2 + \ldots + X_n$

$$= 0 + 0 + 0 + \ldots + 5 + 6 + 8 = 68$$

- Frekans serisinde,

$$\sum_{i=1}^{n} f_i X_i = f_1 X_1 + f_2 X_2 + \ldots + f_k X_k$$

$$= 0(3) + 1(4) + 2(5) + \ldots + 6(1) + 8(1) = 68$$

Birikimli Frekanslar

- Bazı istatistiksel çalışmalarda bir frekans serisinde veya gruplandırılmış seride belirli bir değerden daha küçük veya daha büyük değer alan alan birim sayısının belirlenmesi gerekebilir
- Bu durumlarda birikimli frekanslar hesaplanır

Artan Birikimli Frekanslar

- Artan birikimli frekanslar hesaplanırken frekanslar sütunun ilk frekans değeri aynen alınır
- İkinci frekans değeri ilk frekans değeri ile toplanarak artan birikimli frekanslar sütunun ikinci değeri olarak yazılır
- Bu işleme frekanslar bitene dek devam edilir

33

Artan Birikimli Frekanslar

ÖRNEK: Daha önce frekans serisi olarak düzenlenen serinin artan birikimli frekansını hesaplayınız

X_i	f_i	Artan Birikimli Frekanslar
0	3	3
1	4	3+4=7
2	5	7+5=12
3	6	12+6=18
4	3	18+3=21
5	2	21+2=23
6	1	23+1=24
8	1	24+1=25
	25	

Seride 4 ve daha küçük birimlerin sayısı 21'dir

34

Azalan Birikimli Frekanslar

- Artan birikimli frekanslar için yaptığımız işlemleri bu kez aşağıdan yukarıya doğru yaparız
- Diğer bir yol ise frekanslar toplamından başlanarak çıkarma işlemi yapılmasıdır:
  - Azalan frekanslar sütununa ilk olarak frekanslar toplamı yazılır
  - İkinci frekans değeri frekanslar toplamından çıkarılarak azalan frekanslar sütununa yazılır
  - Bu işleme serinin son değerine dek devam edilir

Azalan Birikimli Frekanslar

ÖRNEK: Daha önce artan birikimli frekansı hesaplanan frekans serisinin azalan birikimli frekanslarını açıklandığı gibi iki şekilde hesaplayınız

X_i	f_i	1. Yol Azalan Birikimli Frekanslar	2. Yol Azalan Birikimli Frekanslar
0	3	25	25
1	4	22	25-3=22
2	5	18	22-4=18
3	6	13	18-5=13
4	3	7	13-6=7
5	2	2+2=4	7-3=4
6	1	1+1=2	4-2=2
8	1	1	2-1=1
	25		

Azalan Birikimli Frekanslar

ÖRNEK: Daha önce gruplandırılmış seri olarak düzenlediğimiz serinin azalan birikimli frekanslarını hesaplayınız

Sınıflar	f_i	Azalan Birikimli Frekanslar
0-2	7	
2-4	11	
4-6	5	
6-8	1	
8-10	1	

37

Oransal Frekanslar

- Bazı durumlarda birimlerin sayısı yerine birimlerin toplam birim sayısına oranının hesaplanması gerekir
- Değişkenin aldığı değerlerin veya sınıf frekanslarının toplam frekansa bölünmesi ile elde edilen frekanslara **oransal frekans** denir

38

Oransal Frekanslar

ÖRNEK: Daha önce ele aldığımız frekans serisi ve sınıfsal serinin oransal frekanslarını hesaplayınız

X_i	f_i	Oransal Frekanslar
0	3	3/25= 0,12
1	4	4/25= 0,16
2	5	5/25= 0,20
3	6	6/25= 0,24
4	3	3/25= 0,12
5	2	2/25= 0,08
6	1	1/25= 0,04
8	1	1/25= 0,04
	25	1,00

Sınıflar	f_i	Oransal Frekanslar
0-2	7	
2-4	11	
4-6	5	
6-8	1	
8-10	1	

Grafikler

- **Tanım**

  Araştırma sonucunda elde edilen ve düzenlenen verilerin daha kolay anlaşılabilmesi için gösterildiği şekillere grafik denir.

- Grafikler göze hitap ettikleri için, toplanan verilerin daha açık bir şekilde görülmesine ve yorumlanmasına yardımcı olur. Buradaki en önemli nokta grafiklerin açık ve anlaşılır biçimde çizilmeleridir.

Zaman Serilerinin Grafikleri

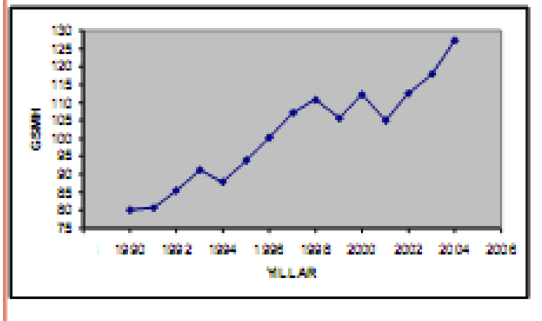
- Zaman serileri, değişkenin aldığı değerlerin zaman birimlerine göre dağılımını gösterir; bu nedenle bunların grafikleri bir koordinat sistemi üzerine çizilebilir.
- Bu koordinat sisteminde yatay eksende zaman birimleri ve düşey eksende ise değişkenin aldığı değerler yer alır.
- Eksenler, zaman birimleri ve değişkenin aldığı değerler dikkate alınarak ölçeklendirilir.

3

Zaman Serilerinin Grafikleri

ÖRNEK: Aşağıda verilen serinin grafiğini çiziniz.

Yıllar	Sabit Fiyatlar ile GSMH (1987) (YTL)
1990	80,1247
1991	80,7768
1992	85,4189
1993	91,3018
1994	87,8921
1995	94,0664
1996	100,1083
1997	107,1451
1998	110,6840
1999	105,5264
2000	112,2314
2001	104,9710
2002	112,4851
2003	117,9757
2004	127,2194



4

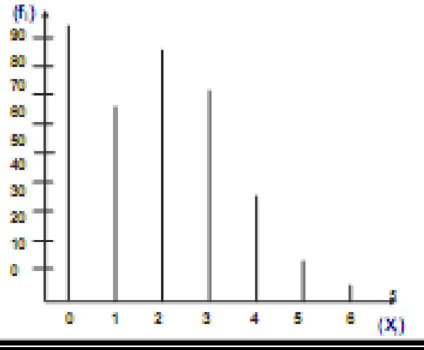
Dağılım Serilerinin Grafikleri

Tanım: Değişkenlerin aldıkları değerlerin dağılımını gösteren ve serilerin türlerine göre çizilen grafiklerdir.

- Basit serilerde birim sayısı az olduğundan ve frekanslar bulunmadığından bu serilerin grafikleri, değişkenin aldığı değerlerin büyüklüklerini gösterecek şekilde çizilir.

ÖRNEK:

Çocuk Sayısı (X_i)	Aile Sayısı (f_i)
0	94
1	65
2	87
3	71
4	34
5	12
6	5

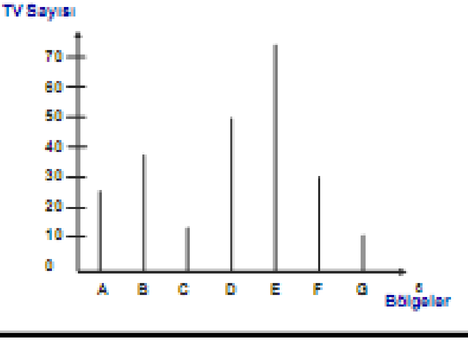


Dağılım Serilerinin Grafikleri (Frekans Serisi)

- Frekans serilerinin grafikleri, değişkenin aldığı değerlere göre frekansların dağılımlarını gösterir ve bir koordinat sistemi üzerine çizilebilir.

ÖRNEK: Bölgelere göre televizyon sayılarının dağılımı aşağıda verilmiştir. Bu dağılımın grafiğini çiziniz.

Bölgeler	Televizyon Sayısı(1=10.000)
A	25
B	36
C	12
D	49
E	71
F	28
G	12



Dağılım Serilerinin Grafikleri (Sınıflı Seri)

Tanım: Sınıf aralıklarının alt ve üst sınırlarından frekans değerlerine kadar çizilen dikmelerin, yatay eksene paralel çizgiler ile birleştirilmesi ile elde edilen dikdörtgenlerin tümüne histogram denir.

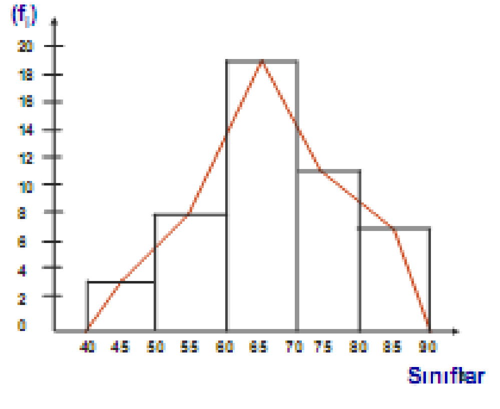
Tanım: Histogramı oluşturan bu dikdörtgenlerin üst orta noktalarının birleştirilmesi ile elde edilen grafiğe de frekans poligonu denir.

7

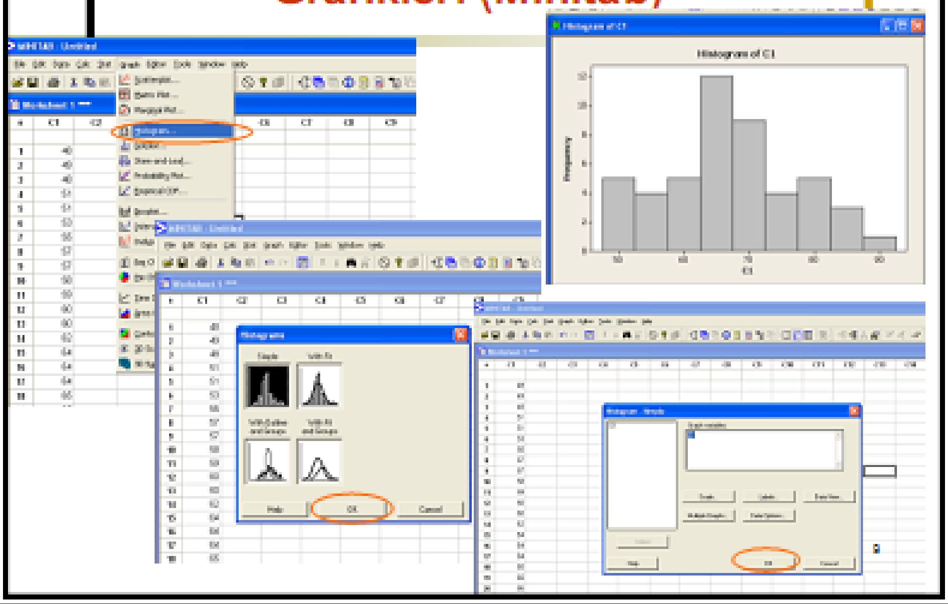
Gruplandırılmış Serilerin Grafikleri

Örnek: Bir sınıftaki 48 öğrencinin ağırlıklarına göre dağılımı aşağıdadır. Bu serinin histogram ve frekans poligonunu çiziniz.

Sınıflar	(f_i)
40-50	3
50-60	8
60-70	19
70-80	11
80-90	7
	48



Gruplandırılmış Serilerin Grafikleri (Minitab)

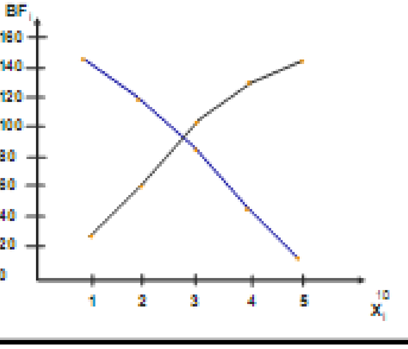


Dağılım Serilerinin Grafikleri (Artan-Azalan Birikimli Frekans Serileri)

- Artan ve azalan birikimli frekans serilerinin grafikleri de koordinat sistemi üzerine çizilebilir. Değişkenin aldığı değerler ile birikimli frekansların kesiştiği noktaların birleştirilmesiyle birikimli serilerin grafiği çizilir.

ÖRNEK: Aşağıda verilen bir bölgedeki binaların kat sayısına göre dağılımı gösteren serinin artan ve azalan frekanslarının grafiğini çiziniz.

Xi	(f_i)	Artan Birikimli Frekanslar	Azalan Birikimli Frekanslar
1	26	26	145
2	34	60	119
3	43	103	85
4	29	132	42
5	13	145	13

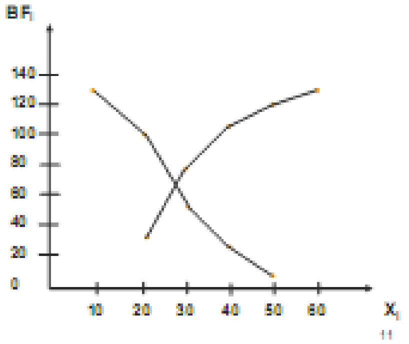


(Artan-Azalan Birikimli Sınıflı Frekans Serileri)

ÖRNEK: Dayanıklı tüketim malları satan bir mağazadaki aylık kredili satış miktarlarının kişilere göre dağılımı yansıda verilmiştir. Serinin artan ve azalan birikimli frekanslarının grafiğini çiziniz.

Sınıflı serilerde artan birikimli frekansların grafikleri çizilirken **sınıfın üst sınır değeri**, azalan birikimli frekanslar çizilirken ise **alt sınır değeri** alınır.

Sınıflar (1=1000TL)	(f_i)	Artan Birikimli Frekanslar	Azalan Birikimli Frekanslar
10-20	32	32	128
20-30	44	76	96
30-40	29	105	52
40-50	15	120	23
50-60	8	128	8



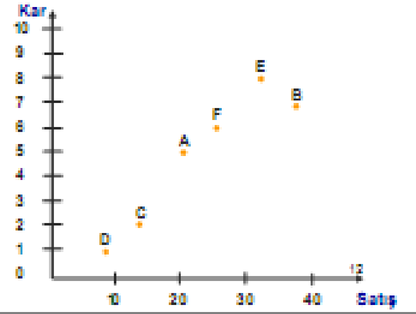
11

Bileşik Serilerin Grafikleri

- **Tanım:** Birimlerin birden fazla karaktere göre dağılmalarını gösteren serilere bileşik seriler denir.
- Bileşik serilerin grafikleri iki karakter söz konusu olduğunda bir karakter yatay eksende, diğer karakter düşey eksende gösterilerek çizilebilir. Karakter sayısı ikiden fazla ise çizim güçleşir.

ÖRNEK: Aynı iş kolunda faaliyet gösteren 6 işletmenin yıllık satış miktarları ve kararları aşağıda verilmiştir. Verilen serinin grafiğini çiziniz.

İşletmeler	Satış Miktarı (1=10000TL)	Kar Miktarı (1=10000TL)
A	22	5
B	38	7
C	14	2
D	8	1
E	33	8
F	27	6



12

SPSS

Merkezi Eğilim Ölçüleri (Ortalamalar)

Prof. Dr. Ünal H. ÖZDEN 91 ©12 Kasım 2015 Perşembe

Ortalamalar

- **Tanım:** Bir serideki tüm gözlem değerlerini temsil eden tek bir rakama ortalama denir.
- Ortalamalar özellikle tek maksimumlu serilerde gözlemlerin hangi değer etrafında toplandığını gösterir.
- Ortalama değer daima serinin minimum ve maksimum değerleri arasında yer alır.

$$X_{min} \leq Ortalama \leq X_{max}$$

Ortalamaların Faydaları:

1. Ortalamalar çoğu zaman serinin normal değerini gösterir. Tabi bunun için serinin dağılımının da aşırı çarpık olmaması gerekir.
2. İstatistik analiz işleminin temel elemanlarından biridir.
3. Aynı birimle ölçmek kaydıyla farklı serileri karşılaştırmaya imkan tanır.
4. Tek bir rakam olması sebebiyle hatırda tutulması kolaydır.

13

Ortalamalar

Ortalamalar hesaplanmalarında kullanılan birimlere göre iki ana gruba ayrılabilir.

1.Grup (Analitik)
- Aritmetik Ortalama
- Kareli Ortalama
- Geometrik Ortalama
- Harmonik Ortalama

2.Grup(Analitik Olmayan)
- Mod
- Medyan

- 1.Gruptaki ortalamaların değeri, serinin herhangi bir biriminin değeri değiştiğinde değişir.
- 2. gruptaki ortalamaların değerinin değişmesi için bu ortalamaların hesabında kullanılan birimlerin değerinin değişmesi gerekir.

14

Aritmetik Ortalama

- **Tanım:** Serideki gözlem değerleri toplamının, toplam gözlem sayısına oranı şeklinde hesaplanır.

Basit seride
$$\bar{X} = \frac{X_1 + X_2 + \dots + X_N}{N} = \frac{\sum_{i=1}^{N} X_i}{N}$$

Frekans serisinde
$$\bar{X} = \frac{f_1 X_1 + f_2 X_2 + \dots + f_k X_k}{f_1 + f_2 + \dots + f_k} = \frac{\sum_{i=1}^{k} f_i X_i}{\sum_{i=1}^{k} f_i}$$

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

X_i : i. gözlem değeri,
f_i : i. gözlemin tekrar sayısı (frekansı)
m_i : i. sınıfın sınıf orta noktası

15

Aritmetik Ortalama

Örnek 1: Bir işletmede aynı parçayı üreten işçilerin bu parçayı üretim sürelerinin dağılımı aşağıdaki gibi gözlenmiştir. Parça üretim süresinin aritmetik ortalamasını bulunuz.

Parça üretim süresi (dakika) (X_i)	İşçi sayısı (f_i)	f_iX_i
12	2	24
13	4	52
14	7	98
15	6	90
16	1	16
Toplam	$\sum f_i = 20$	$\sum f_ix_i = 280$

$$\bar{X} = \frac{\sum_{i=1}^{5} f_i X_i}{\sum_{i=1}^{5} f_i} = \frac{280}{20} = 14 \text{ dk/parça}$$

Aritmetik Ortalama

Örnek 2: Bir işyerinde yapılan telefon görüşmelerinin süresinin dağılımı için aşağıdaki gruplanmış seri verilmiştir. Buna göre görüşme süresinin aritmetik ortalamasını bulunuz.

Görüşme süresi (dakika)	Görüşme sayısı (f_i)	m_i	f_im_i
0 - 2 den az	5	1 [(0+2)/2]	5
2 - 4 " "	10	3 [(2+4)/2]	30
4 - 6 " "	40	5 [(4+6)/2]	200
6 - 10 " "	30	8	240
10 - 20 " "	25	15	375
Toplam	$\Sigma f_i = 110$		$\Sigma f_im_i = 850$

$$\bar{X} = \frac{\sum_{i=1}^{k} f_i m_i}{\sum_{i=1}^{k} f_i} = \frac{850}{110} \cong 7.73 \text{ dk./görüşme}$$

Gruplanmış serilerde ortalamayı hesaplayabilmek için her sınıfı temsil eden tek bir değere ihtiyaç vardır. Bu değer o sınıfı en iyi temsil etmesi muhtemel olan sınıf orta noktasıdır.

$$m_i = \frac{l_{i1} + l_{i2}}{2}$$

Tartılı Aritmetik Ortalama

Tanım: Bir serideki değerler arasında önem derecesi farklı oluyorsa, bu tür serilerin aritmetik ortalaması tartılı olarak hesaplanır. Bunun için önem düzeyini gösteren katsayılar (tartılar) kullanılır.

Basit seride	$X_T = \dfrac{\sum_{i=1}^{n} t_i X_i}{\Sigma t_i}$
Frekans serilerinde	$X_T = \dfrac{\sum_{i=1}^{k} t_i f_i X_i}{\sum_{i=1}^{k} t_i f_i}$
Gruplanmış seride	$X_T = \dfrac{\sum_{i=1}^{k} t_i f_i m_i}{\sum_{i=1}^{k} t_i f_i}$

18

Tartılı Aritmetik Ortalama

Örnek 1: Aşağıda bir öğrencinin almış olduğu dersler, notları ve kredileri verilmiştir. Not ortalamasını tartılı aritmetik ortalama cinsinden hesaplayınız.

Dersler	Notlar (X_i)	Kredi (t_i)	$t_i X_i$
İstatistik	70	3	210
Matematik	60	4	240
Fizik	50	3	150
Kimya	80	2	160
Toplam	260	$\Sigma t_i = 12$	$\Sigma t_i X_i = 760$

Tartılı Ortalama

$$\bar{X}_T = \frac{\sum_{i=1}^{n} t_i X_i}{N} = \frac{760}{12} = 63.33$$

Basit Ortalama

$$\bar{X} = \frac{\sum_{i=1}^{N} X_i}{N} = \frac{260}{4} = 65$$

19

Tartılı Aritmetik Ortalama

Örnek 2: Bir işletmede işçilerin saat ücretleri çalıştıkları süre (kıdem) dikkate alınarak belirlenmektedir. Veriler aşağıdaki gibi olduğuna bu işletmede ortalama saat ücretini tartılı aritmetik ortalama cinsinden hesaplayınız.

Saat ücreti (milyon TL)	İşçi sayısı (f_i)	Ortalama kıdem (t_i)	m_i	f_it_i	$f_it_im_i$	f_im_i
1.00 – 1.40	10	2.5	1.20	25	30.0	12.00
1.40 – 1.60	30	5.0	1.50	150	225.0	45.00
1.60 – 1.80	50	9.5	1.70	475	807.5	85.00
1.80 – 2.00	15	13.0	1.90	195	370.5	16.90
2.00 – 2.50	5	18.0	2.25	90	202.5	11.25
Toplam	110			935	1635.5	170.15

Tartılı Ortalama

$$X_T = \frac{\sum f_it_im_i}{\sum f_it_i} = \frac{1635.5}{935} = 1.75 \text{ milyon TL/saat}$$

Basit Ortalama

$$\bar{X} = \frac{\sum f_im_i}{N} = \frac{170.15}{110} = 1.547 \text{ milyon TL/saat}$$

20

Tartılı Aritmetik Ortalama

Tartılı aritmetik ortalamanın kullanıldığı yerler:

- Veriler arasında önem farkı bulunması halinde kullanılır.
- Oranların ve ortalamaların ortalaması hesaplanırken kullanılır.

Örnek: Bir işletmede bulunan üç tezgahın belli bir günde ürettikleri malların sayısı ve üretimlerindeki kusurlu oranları aşağıdaki tabloda verilmiştir. Buna göre bu tezgahlarının ürettiği mamul kütlesinin ortalama kusurlu oranını bulunuz.

Tezgahlar	Üretim miktarı (t_i)	Kusurlu oranı (X_i)	t_iX_i
A	100	0.03	3
B	200	0.05	10
C	50	0.01	0.5
	$\Sigma t_i = 350$	$\Sigma X_i = 0.09$	$\Sigma t_iX_i = 13.5$

21

Tartılı Aritmetik Ortalama

Basit aritmetik ortalama ile kütlenin kusurlu oranı;

$$\bar{X} = \frac{\sum X_i}{N} = \frac{0.09}{3} = 0.03$$

Tartılı aritmetik ortalama ile kütlenin kusurlu oranı;

$$X_T = \frac{\sum_{i=1}^{n} t_i X_i}{\sum_{i=1}^{n} t_i} = \frac{13.5}{350} = 0.03857$$

Kütlenin gerçek kusurlu oranı 0.03857 dir

22

Aritmetik Ortalamanın Özellikleri

1. Aritmetik ortalama hassas (diğer ortalamalara nazaran) bir ortalama olup serideki aşırı değerlerden etkilenir ve aşırı değere doğru kayma gösterir.

2. Serinin gözlem sayısı ile aritmetik ortalaması çarpılırsa serinin toplam değeri elde edilir.

$$N.\ \bar{X} = \sum Xi$$

3. Serideki gözlem değerlerinin aritmetik ortalamadan sapmaları toplamı sıfır olur.

$$\sum(Xi - \bar{X}) = 0$$

23

Aritmetik Ortalamanın Özellikleri

4. Serideki değerlerin aritmetik ortalamadan sapmalarının kareleri toplamı minimum olur.

$$\sum(X_i - \bar{X})^2 \rightarrow \text{minimum olur.}$$

5. Bir serinin değerleri, diğer iki serinin değerleri toplamından oluşuyorsa bu serinin aritmetik ortalaması da diğer iki serinin aritmetik ortalamaları toplamına eşit olur.

$$\bar{X} = \bar{Y} + \bar{Z}$$

24

Geometrik Ortalama(G)

Tanım: Bir serideki gözlem değerlerinin birbirleri ile çarpımlarının, gözlem sayısı derecesinde kökünün alınması ile elde edilir.

$$G = \sqrt[N]{X_1 \cdot X_2 \cdot X_3 \cdot \dots \cdot X_N}$$

$$G = \sqrt[N]{\prod_{i=1}^{N} X_i}$$

25

Geometrik Ortalama(G)

- Bu yoldan geometrik ortalamayı bulmak için gözlem sayısının az olması gerekir.
- Gözlem sayısı arttıkça bu yoldan geometrik ortalamayı hesaplamak güçleşmektedir.
- Bunun yerine logaritmik dönüşüm uygulanarak geometrik ortalama hesaplanır.

Basit seride

$$G = \sqrt[N]{X_1 \cdot X_2 \cdot X_3 \cdot \cdots X_N} \qquad \log G = \frac{\sum_{i=1}^{N} \log X_i}{N}$$

Geometrik Ortalama(G)

Frekans serileri için

$$G = \sqrt[\sum f_i]{X_1^{f_1} \cdot X_2^{f_2} \cdot X_3^{f_3} \cdots X_k^{f_k}}$$

$$\log G = \frac{f_1 \log X_1 + f_2 \log X_2 + f_3 \log X_3 + \cdots + f_k \log X_k}{f_1 + f_2 + f_3 + \cdots f_k} = \frac{\sum_{i=1}^{k} f_i \log X_i}{\sum_{i=1}^{k} f_i}$$

Gruplanmış seri için

$$G = \sqrt[\sum f_i]{m_1^{f_1} \cdot m_2^{f_2} \cdot m_3^{f_3} \cdots m_k^{f_k}}$$

$$\log G = \frac{f_1 \log m_1 + f_2 \log m_2 + f_3 \log m_3 + \cdots + f_k \log m_k}{f_1 + f_2 + f_3 + \cdots f_k} = \frac{\sum_{i=1}^{k} f_i \log m_i}{\sum_{i=1}^{k} f_i}$$

logG'yi G'ye çevirmek için; G =10logG dönüşümü yapılır

Harmonik Ortalama

Tanım: Harmonik ortalama bir serideki gözlem değerlerinin terslerinin aritmetik ortalamasının tersine eşittir.

Basit Seri için;

$$H = \frac{1}{\frac{\frac{1}{X_1}+\frac{1}{X_2}+\frac{1}{X_3}+\dots+\frac{1}{X_N}}{N}} = \frac{N}{\frac{1}{X_1}+\frac{1}{X_2}+\frac{1}{X_3}+\dots+\frac{1}{X_N}} \Rightarrow H = \frac{N}{\sum_{i=1}^{N}\frac{1}{X_i}}$$

Harmonik Ortalama

Frekans serisi için

$$H = \frac{\sum f_i}{\sum \frac{f_i}{X_i}}$$

Gruplanmış seride

$$H = \frac{\sum f_i}{\sum \frac{f_i}{m_i}}$$

Kareli Ortalama (K)

Tanım: Kareli ortalama serideki değerlerin karelerinin aritmetik ortalamasının kareköküdür

Basit seride

$$K = \sqrt{\frac{X_1^2 + X_2^2 + X_3^2 + \cdots + X_N^2}{N}} \Rightarrow K = \sqrt{\frac{\sum_{i=1}^{N} X_i}{N}}$$

Tasnif edilmiş seride

$$K = \sqrt{\frac{f_1 X_1^2 + f_2 X_2^2 + f_3 X_3^2 + \cdots f_k X_k^2}{f_1 + f_2 + f_3 + \cdots + f_k}} \Rightarrow K = \sqrt{\frac{\sum_{i=1}^{k} f_i X_i^2}{\sum_{i=1}^{k} f_i}}$$

Gruplanmış seride

$$K = \sqrt{\frac{f_1 m_1^2 + f_2 m_2^2 + f_3 m_3^2 + \cdots f_k m_k^2}{f_1 + f_2 + f_3 + \cdots + f_k}} \Rightarrow K = \sqrt{\frac{\sum_{i=1}^{k} f_i m_i^2}{\sum_{i=1}^{k} f_i}}$$

Kareli Ortalama (K)

Örnek: Bir şehirdeki konutlarda elektrik enerjisi tüketimi üzerine yapılan araştırmada, 200 konut rasgele seçilmiş ve aşağıdaki sonuçlar elde edilmiştir.

Aylık elektrik tüketimi (Kwh)	Konut sayısı (f_i)	m_i	$f_i m_i$
0 – 60	10	30	9000
60 – 100	20	80	128000
100 – 120	40	110	484000
120 – 140	50	130	845000
140 – 180	45	160	1152000
180 – 250	35	215	1617875
Toplam	200		4235875

$$K = \sqrt{\frac{\sum_{i=1}^{k} f_i m_i^2}{\sum_{i=1}^{k} f_i}} = \sqrt{\frac{4235875}{200}} = \sqrt{21179{,}375} \Rightarrow K \cong 145{,}53 \text{ Kwh/ay}$$

Analitik Ortalamalar Arasındaki İlişkiler

Normal bir seride ortalamalar arasında aşağıdaki gibi bir büyüklük ilişkisi vardır.

$$K > \bar{X} > G > H$$

Mod

Tanım: Bir seride en çok tekrarlanan değere mod adı verilir.

İstatistikte nispeten az kullanılan bu ölçü özellikle verilerin simetrik bir dağılış göstermediği durumlarda iyi bir ölçü olarak düşünülebilir.

Tanım: Eğer seride en çok tekrarlanan birden fazla eleman varsa bu tür seriler çok modlu seriler olarak isimlendirilir.

Böyle serilerde modun tek bir değerle ifade edilmesi istenirse seri gruplanmış hale dönüştürülerek modu hesaplanabilir. Gruplama sonrasında da en yüksek frekansa sahip tek bir sınıf bulunamazsa sınıflar birleştirilerek mod hesaplanabilir.

Mod

Örnek: Aşağıda basit ve tasnif edilmiş iki seri verilmiştir. Bu serilerin modlarını bulunuz

X_i
10
12
14
15
15
15
16
16
18
20

Mod=15

X_i	(f_i)
10	1
12	1
14	1
15 ←	3
16	2
18	1
20	1

Mod=15

Gruplanmış Seride Modun Bulunması

- Gruplanmış seride modu bulmak için serinin frekanslar sütunundan hareketle ençok tekrarlanan sınıf belirlenir.
- Belirlenen sınıf içerisinde mod'a karşılık gelen değeri elde etmek için aşağıdaki formül kullanılır

l_1 : mod sınıfı alt sınırı
Δ_1 : mod sınıfı ile bir önceki sınıf frekansı arasındaki fark
Δ_2 : mod sınıfı ile bir sonraki sınıf frekansı arasındaki fark
s : serinin sabit sınıf aralığı

$$Mod = l_1 + \frac{\Delta_1}{\Delta_1 + \Delta_2} \cdot s$$

- Gruplanmış serilerde modu bulabilmek serinin sınıf aralıklarının eşit olmasına dikkat edilmelidir. Eğer sınıf aralıkları eşit verilmemişse, aralıkları eşit hale getirebilmek için sınıfları birleştirme yoluna gidilebilir. Buna rağmen sınıf aralıkları eşit hale getirilemiyorsa mod hesaplanamaz.
- Bazen seri iki, üç modlu olabilir. Böyle serilerde tek modlu hale getirmek için tasnif edilmiş seriyi sınıflandırmak, sınıflandırılmış serinin de sınıflarını birleştirmek gerekebilir.

Gruplanmış Seride Modun Bulunması

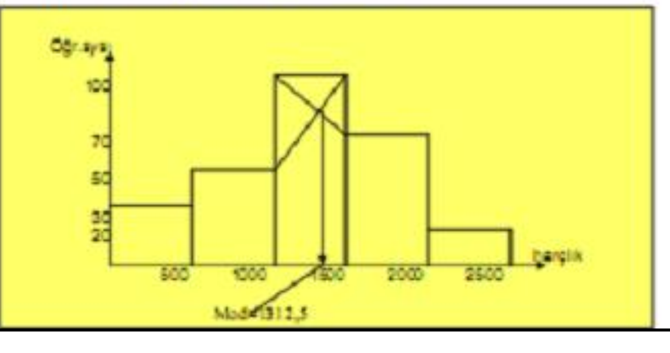
Örnek: Bir ilköğretim okulunda öğrencilerin günlük olarak aldıkları harçlıkların dağılımı aşağıda verilmiştir. Öğrencilerin aldıkları günlük harçlık miktarının ortalamasını mod ile belirleyiniz.

Harçlık (10^3)TL/gün	*Öğrenci sayısı*
0 – 500	*30*
500 – 1000	*50*
1000 – 1500	*100*
1500 – 2000	*70*
2000 – 2500	*20*

$$Mod = l_1 + \frac{\Delta_1}{\Delta_1 + \Delta_2} \cdot s \Rightarrow Mod = 1000 + \frac{50}{50+30} \cdot 500 \Rightarrow Mod = 1312{,}5 \Rightarrow Mod = 1312500TL / gün$$

Modun Grafikle Gösterilmesi

- Modun grafikle gösterilebilmesi için serinin histogramı çizilir.
- Histogramda en yüksek sütun mod sınıfına karşılık gelir.
- Modun yerini tayin etmek için en yüksek sütunun üst köşegenleri ile komşu sütunların bitişik üst köşeleri çapraz olarak birleştirilir.
- İki doğrunun kesim noktasından yatay eksene çizilen doğrunun ekseni kestiği nokta mod olarak tespit edilir.



Modun Özellikleri

- Ortalamalar oranında en temsili alanıdır.
- Pratik hayatta çok kullanılan ortalamalardandır
- Özellikle kalitatif (niteliksel) serilerin ortalaması mod ile ifade edilir.
- Mod serideki aşırı değerlere karşı hassas değildir.
- Yukarıdaki avantajlarının yanında analitik olmaması sebebi ile matematik işlemlere elverişli değildir.
- J, ters J ve U tipi serilerde mod temsili alma özelliğini kaybeder. Böyle serilerde mod ya en küçük veya en büyük değere karşılık gelir

Medyan (Ortanca)

Tanım: Serideki değerler küçükten büyüğe sıralandığında tam ortaya düşen ve seriyi iki eşit parçaya bölen değere medyan adı verilir.

Basit ve tasnif edilmiş seride medyanın bulunuşu:

- Bunun için serideki değerler küçükten büyüğe sıralanır daha sonra medyana karşılık gelen değerin sıra değeri belirlenir.
- $\frac{N+1}{2}$ işlemi ile medyanın hangi sıradaki eleman olduğu belirlenir.
- Eğer bu işlemin sonucu tam sayı ise bu sıradaki eleman medyan olarak belirlenmiş olur. Eğer bu işlemin sonucu kesirli çıkarsa medyan iki değerin tam ortasına düşeceğinden bu iki değerin ortalaması alınarak medyan bulunur.

Medyan (Ortanca)

Örnek: *X_i:15,8,12,23,45,32,5,18,16,28,39,51*

Yukarıdaki serinin medyanını bulunuz.

Önce seri büyüklük sırasına göre dizilir

X_i : 5,8,12,15,16,18,23,28,32,39,45,51

$$\frac{N+1}{2}=\frac{12+1}{2}=6.5.$$ sıradaki değer medyandır.

Bu değer 18-23 arasında olduğundan

$$Medyan=\frac{18+23}{2}\Rightarrow Medyan=20.5$$ olarak bulunur

Gruplanmış Seride Medyanın Hesaplanması

- Gruplanmış seriler sürekli karekterde seriler olduğu için medyanın sıra değeri N/2 şeklinde bulunur.
- Bu değer toplam frekansın yarısına eşit olup serinin orta noktasını gösterir.
- Ortaya düşen sınıf yani medyan sınıfı kolayca belirlenebilir.
- Belirlenmiş olan medyan sınıfından hareketle aşağıdaki formül yardımı ile medyan değeri hesaplanır.

$$Medyan=l_1+\frac{\frac{N}{2}-\sum_{i=1}^{m-1}N_i}{N_m}\cdot s_m$$

l_1 : Medyan sınıfının alt sınırı

N/2 : Medyanın sıra değeri

$\sum_{i=1}^{m-1}N_i$: Medyandan önceki frekanslar toplamı

N_m : Medyan sınıfının frekansı

S_m : Medyan sınıfının sınıf aralığı

Gruplanmış Seride Medyanın Hesaplanması

Örnek: Bir işletmede işçilere ödenen saat ücretlerinin dağılımı aşağıda verilmiştir. Bu verilere göre medyan saat ücretini hesaplayınız.

Saat ücreti (Bin TL)	İşçi sayısı
500 – 600	10
600 – 700	50
700 – 800 ⇐ Medyan sınıfı	40 ←
800 – 1000	30
1000 – 1500	20
Toplam	150

$\frac{N}{2}=\frac{150}{2}=75.$ değer medyan dır. Bu değer 800 – 1000 ücret sınıfında yer almaktadır.

$$Medyan = l_1 + \frac{\frac{N}{2} - \sum_{i=1}^{m-1} N_i}{N_m} \cdot s_m \Rightarrow Medyan = 700 + \frac{75-60}{40} \cdot 100 \Rightarrow Medyan = 737.5\ TL/saat$$

Medyanın Grafikle Belirlenmesi

- Medyanın grafik üzerinde gösterilebilmesi için kümülatif ve ters kümülatif frekans serilerin oluşturulması gerekir.
- Bu serilerin grafiği birlikte çizildiğinde iki eğrinin birbirini kestiği noktadan yatay eksene çizilen doğrunun ekseni kestiği nokta medyan olarak tespit edilir. Bu işlem sadece eğrilerden birinin çizilmesiyle de yapabiliriz.
- Eğrilerden biri çizildiğinde Y ekseninde N/2 değerine karşılık gelen noktadan X eksenine paralel çizildiğinde, bu doğrunun kümülatif eğriye temas ettiği noktadan X eksenine çizilen doğrunun ekseni kestiği noktada medyanı gösterecektir.

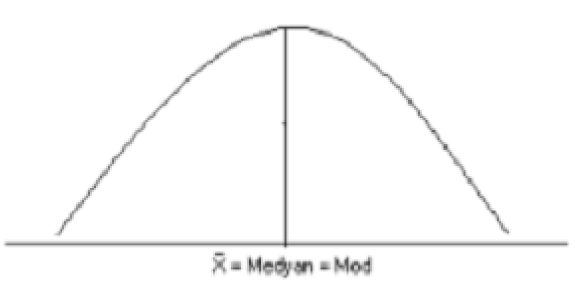
Medyanın Özellikleri

1. Pratik bir ortalamadır, sadece basit bir sıralama işlemi gerektirir.
2. Özellikle açık sınıflı seriler için medyan daha bir önem kazanır. Medyan böyle serilerin ortalamasında problemsiz olarak hesaplanabilir.
3. Serideki aşırı değerlere karşı hassas değildir. Çünkü medyan serinin ortasına rastladığında, uçlarda oluşan aşırı değerler medyanı etkilemez.
4. Serideki değerlerin medyandan mutlak farkları toplamı minimum olur. $\sum |X_i - \text{medyan}| \rightarrow \text{minimum}$
5. Medyanın zayıf tarafı serideki bütün değerleri dikkate almaması sebebi ile matematik işlemlere elverişli değildir.

Mod, Medyan ve Aritmetik Ortalama Arasındaki İlişkiler

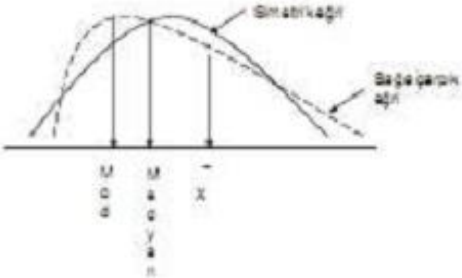
1- Simetrik serilerde her üç ortalama birbirine eşit olur.

X=medyan=mod



Mod, Medyan ve Aritmetik Ortalama Arasındaki İlişkiler

2- Sağa çarpık serilerde $\bar{X}$ > Medyan > Mod olur.



DEĞİŞKENLİK ÖLÇÜLERİ

Giriş

Ortalamalar, serilerin karşılaştırılmasında her zaman yeterli ölçüler değildir. Aynı ortalamayı sahip seriler farklı dağılım gösterebilirler. Bu nedenle serilerin karşılaştırılmasında, değişkenlik ve asimetri ölçülerine bakılır.

2

Değişkenlik

- X_i ve Y_i aşağıdaki gibi iki seri verilmiş olsun:

Xi	Yi
0	2
1	2
2	3
3	4
9	4

Serilerin aritmetik ortalamaları

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{15}{5} = 3$$

$$\bar{Y} = \frac{\sum_{i=1}^{N} Y_i}{N} = \frac{15}{5} = 3$$

3

Değişkenlik Ölçüleri

Değişkenliği az olan serilerin ortalamaları daha temsili oldukları halde, değişkenliği fazla olanların ortalamaları seriyi daha az temsil eder .

Değişkenlik Ölçüleri

Mutlak Değişkenlik Ölçüleri

- Değişim Aralığı
- Kartil ve Desil Aralığı
- *Standart Sapma*

Nisbi Değişkenlik Ölçüleri

- Değişim Katsayısı

Değişim Aralığı

Tanım: Gözlem değerlerinin maksimum ve minimumu arasındaki fark olup, verilerin ne kadarlık bir aralıkta değiştiğini gösterir.

Değişim aralığı frekans serilerinde X_i sütununun maksimum ve minimum değerleri arasındaki farka, gruplandırılmış serilerde ise ilk grubun alt sınır değeri ile, son sınıfın üst sınır değeri arasındaki farka eşittir.

$$R = X_{max} - X_{min}$$

Örnek: X_i : 12,15,20,30,50,52,58,70,90 olan bir serinin değişim aralığı

R=90-12 =78

Yani gözlem değerleri 78 birimlik bir aralıkta değişim göstermektedir .

Standart Sapma

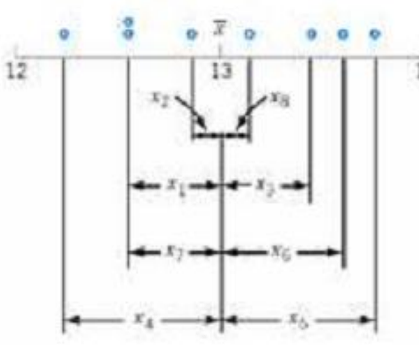
Tanım: Seri değerlerinin aritmetik ortalamadan farklarının kareli ortalamasına standart sapma denir. Standart sapma σ ile gösterilir.

Tanım: Standart sapmanın karesine varyans adı verilir. Varyans V(X) yada σ^2 ile ifade edilir.

6

Standart Sapma

- *Varyans serideki değişimi nasıl ölçer?*



$X_i - \overline{X}$ her zaman sıfır olacağından farkların karelerini alınız

7

Standart Sapma

Basit Seride

$$\sigma = \sqrt{\frac{\sum (Xi - \overline{X})^2}{N}}$$

Frekans Serisinde

$$\sigma = \sqrt{\frac{\sum fi(Xi - \overline{X})^2}{\sum fi}}$$

Gruplanmış Seride

$$\sigma = \sqrt{\frac{\sum fi(mi - \overline{X})^2}{\sum fi}}$$

8

Standart Sapma

Örnek: Doğru , yanlış şeklinde cevap şıkları olan 10 soruya öğrencilerin verdikleri doğru cevap sayılarının dağılımı aşağıda verilmiştir . Bu serinin standart sapmasını ve varyansını bulunuz .

Doğru Cev.Say.	Öğr.Sayısı
2	2
3	4
4	5
5	10
6	20
7	30
8	20
9	10
10	3

$$\sigma = \sqrt{\frac{\sum fi(Xi - \overline{X})^2}{\sum fi}}$$

$\overline{X} = \frac{696}{104} \cong 6{,}69$ Standart Sapma $\sigma = \sqrt{\frac{\sum fi(Xi - \overline{X})^2}{\sum fi}} = \sqrt{\frac{296{,}15}{104}} \Rightarrow \sigma \cong 1{,}69$

Varyans $\sigma^2 = 2{,}847$ olarak bulunur.

9

Standart Sapma

Örnek: Bir liseden mezun olan ve ÖSS sınavına giren öğrencilerin puanlarının dağılımı aşağıda verilmiştir. Buna göre öğrenci puanlarının standart sapmasını bulunuz.

ÖSS Puanları	Öğr.Sayısı
90-110	10
110-130	30
130-150	50
150-170	25
170-210	5
Toplam	120

$$\overline{X} = \frac{16550}{120} \cong 137,9 \ puan$$

$$\sigma = \sqrt{\frac{\sum fi(mi - \overline{X})^2}{\sum fi}} = \sqrt{\frac{49979,2}{120}} \Rightarrow \sigma \cong 20,4 \ puan$$

10

Değişim Katsayısı

Tanım: Standart sapmanın ortalamanın bir yüzdesi olarak ifade edilmesine değişim katsayısı adı verilir . Bu tanıma göre standart sapmanın büyüklüğü ortalamaya göre ifade edilmektedir .

$$\boxed{D.K = \frac{\sigma}{\overline{X}}.100}$$

- Bu hesaplama ile ölçü birimlerinin etkisi giderilmiş olmaktadır. Bu nedenle bu karşılaştırılmak istenen serilerin değerleri farklı ölçü birimleri ile ifade edildiği durumlarda değişim katsayısı kullanılabilir.

- Değişim katsayısı küçük olan serilerde, birimlerin ortalama etrafında daha uygun dağıldıkları sonucuna varılır.

15

Değişim Katsayısı

Örnek: **Konutlarda tüketilen aylık elektrik ve su miktarları için aşağıdaki veriler elde edilmiştir . Değişim katsayılarını bularak hangi grupta değişkenliğin daha fazla olduğunu araştırın.**

Elekt. Tük.(kwh)	Konut Say.
50-100	10
100-150	20
150-200	30
200-300	15
300-500	5
	80

Su Tük.(ton/h)	Konut Say.
5-15	10
15-25	30
25-35	40
35-45	20
45-65	10
	110

16

Değişim Katsayısı

Elektrik Tüketimi İçin

$$\overline{X} = \frac{14250}{80} = 178,125 \qquad K^2 = \frac{3025000}{80} = 37812,5$$

$$\sigma = \sqrt{K^2 - \overline{X}^2} = \sqrt{37812,5 - 178,125^2} = \sqrt{6083,98} \cong 78$$

$$DK = \frac{\sigma}{\overline{X}}.100 \Rightarrow DK = \frac{78}{178,125}.100 \cong 44,8$$

Su Tüketimi İçin

$$\overline{X} = \frac{3250}{110} = 29,55 \qquad K^2 = \frac{111250}{110} = 1011,4$$

$$\sigma = \sqrt{K^2 - \overline{X}^2} = \sqrt{1011,4 - 29,55^2} = \sqrt{138,16} \cong 11,75$$

$$DK = \frac{\sigma}{\overline{X}}.100 \Rightarrow DK = \frac{11,75}{29,55}.100 \cong 39,76$$

17

Asimetri Ölçüleri

Tanım: Serilerin dağılma şekillerini belirlemek için hesaplanan ölçülere asimetri ölçüleri denir.

- Değişkenlik ölçüleri serilerin değişkenliğini ölçebilmesine rağmen, onların dağılma şekilleri hakkında bilgi vermez. Oysa ortalamaları ve değişkenlik ölçüleri birbirine eşit veya yakın olan serilerin dağılımları birbirinden farklı olabilir. Bu nedenle serilerin sadece değişkenli ölçüleri değil, dağılma şekilleri de araştırılmalıdır.

18

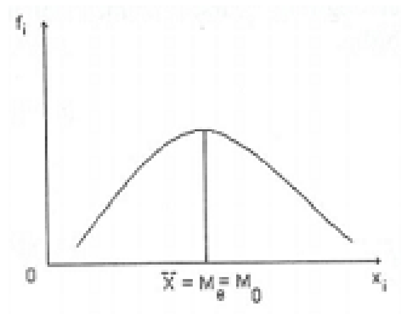
Asimetri Ölçüleri

- Dağılımı çan eğrisine benzeyen, tek maksimumlu ve simetrik olan dağılım istatistikte önemli bir yere sahiptir ve bazı yöntemlerin uygulanması dağılımların normal dağılım olduğu varsayımına dayanmaktadır.
- Bir dağılımın normal dağılım olarak kabul edilebilmesi için simetri durumunun bulunması ve basıklık ölçüsünün hesaplanması gerekir. Basıklık ölçüsü ile bir dağılımın basıklığının, normal dağılımın basıklığından farklı olup olmadığı belirlenmektedir.

19

Ortalamalar Arası Farklar ile Hesaplanan Asimetri Ölçüleri

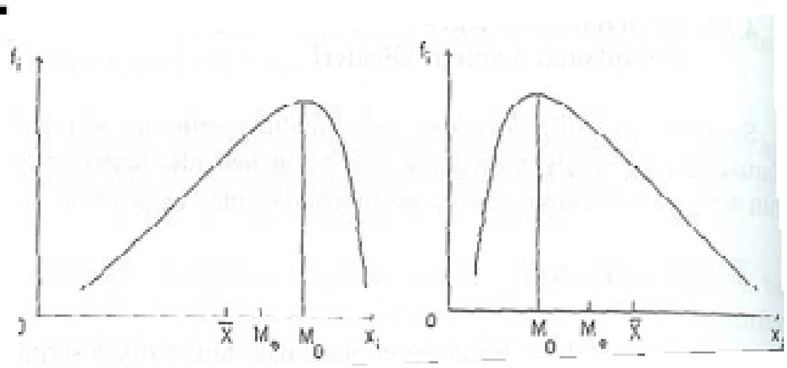
Simetrik dağılım gösteren tek modlu serilerde aritmetik ortalama, mod ve medyan birbirine eşittir.



$\bar{X} = M_e = M_o$ ise seri simetriktir. Bu değerler arasındaki fark arttıkça serinin eğikliği artar.

20

Ortalamalar Arası Farklar ile Hesaplanan Asimetri Ölçüleri

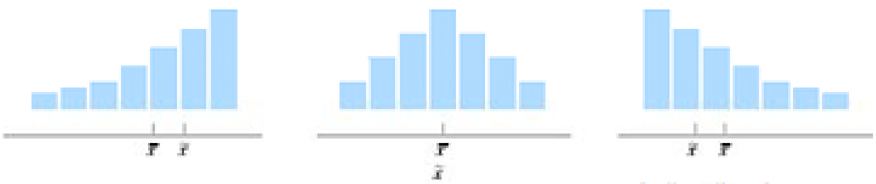


$\bar{X} < M_e < M_o$ ise seri sola çarpıktır.

$\bar{X} > M_e > M_o$ ise seri sağa çarpıktır.

21

Asimetri Ölçüleri



Sola Eğik Seri | Simetrik Seri | Sağa Eğik Seri

Ortalamalar Arası Farklar ile Hesaplanan Asimetri Ölçüleri

Sağa ve sola eğikliği fazla olmayan serilerde

$$\boxed{\bar{X} - M_o \cong 3(\bar{X} - M_e)}$$

- Bu eşitlikten yararlanarak Pearson asimetri ölçüleri ($\alpha_1 ve \alpha_2$) hesaplanabilir.
- Bu ölçülerin hesaplanmasında farklı ölçü birimlerinin etkisini gidermek ve oransal bir ölçü etmek için aritmetik ortalama ile mod yada medyan arasındaki farklar standart sapmaya bölünür.

$$\alpha_1 = \frac{\bar{X} - M_o}{\sigma} \qquad \alpha_2 = \frac{3(\bar{X} - M_e)}{\sigma}$$

Ortalamalar Arası Farklar ile Hesaplanan Asimetri Ölçüleri

Seri simetrik ise,

$\overline{X} = M_e = M_o$ olacağından $\alpha_1 = 0$ $\alpha_2 = 0$

Seri sola eğikse,

$\overline{X} < M_e < M_o$ olacağından $\alpha_1 < 0$ $\alpha_2 < 0$

Seri sağa eğikse,

$\overline{X} > M_e > M_o$ olacağından $\alpha_1 > 0$ $\alpha_2 > 0$

28

Kartiller Arası Farklar ile Hesaplanan Asimetri Ölçüleri

Asimetri ölçüleri ortalamalar arası farka dayanarak hesaplanabildikleri gibi, kartiller arası farka dayanarak da hesaplanabilir.

a) Simetrik serilerde

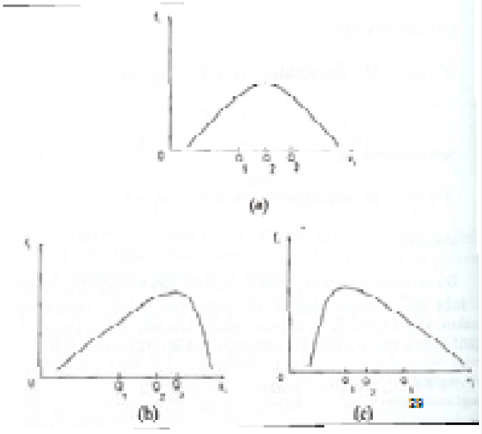
$Q_3 - Q_2 = Q_2 - Q_1$

b)Seri sola eğikse,

$Q_3 - Q_2 < Q_2 - Q_1$

c) Seri sağa eğikse,

$Q_3 - Q_2 > Q_2 - Q_1$



29

Kartiller Arası Farklar ile Hesaplanan Asimetri Ölçüleri

Tanım: Bu ölçülerin hesaplanmasında farklı ölçü birimlerinin etkisini gidermek ve oransal bir ölçü etmek için kartiller arası farklar, kartiller arası farklar toplamına bölünür. Bu şekilde elde edilen ölçü Bowley asimetri ölçüsü olarak adlandırılır.

$$A_s = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_2) + (Q_2 - Q_1)} \qquad A_s = \frac{(Q_3 - 2Q_2 + Q_1)}{Q_3 - Q_1}$$

Bowley simetri ölçüsüne göre elde edilen sonuç (+1) ile (-1) arasında değişmektedir. Sonuç:

(0) ise seri simetrik,
(-) ise sola asimetrik,
(+) ise sağa asimetriktir.

30

SPSS

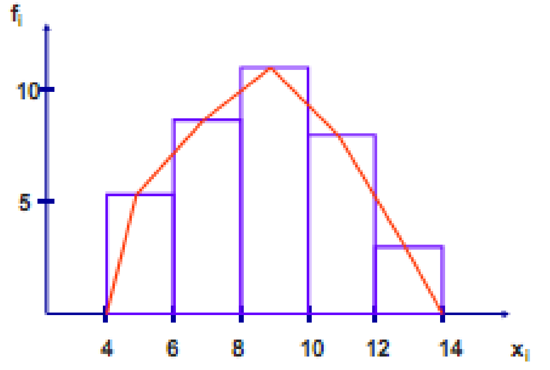
Örnek: Aşağıda verilen gruplandırılmış serinin simetrik olup olmadığını

a) Grafiğini çizerek,
b) Mod, medyan ve aritmetik ortalamasını karşılaştırarak,
c) Pearson katsayılarını ve Bowley asimetri ölçüsünü hesaplayarak inceleyiniz.

Sınıflar	f_i	Artan Birikimli Frekanslar	X_i	$f_i X_i$	$f_i X_i^2$
4-6	5	5	5	25	125
6-8	8	13	7	56	392
8-10	12	25	9	108	972
10-12	7	32	11	77	847
12-14	3	35	13	39	507
	35			305	2843

31

Prof. Dr. Ünal H. ÖZDEN 148 ©12 Kasım 2015 Perşembe a)



32

b) $M_o = l_a + \frac{\nabla_1}{\nabla_1 + \nabla_2}.c = 8 + \frac{4}{4+5}.2 = 8,88$

$$M_e = l_a + \frac{\frac{n}{2} - \sum_{i=1}^{k-1} f_b}{f_m}.c$$

$$\frac{n}{2} = \frac{35}{2} = 17,5$$

$$M_e = 8 + \frac{17,5 - 13}{12}.2 = 8,75$$

$$\bar{X} = \frac{\sum f_i X_i}{\sum f_i} = \frac{305}{35} = 8,71$$

$\bar{X} < M_e < M_o$ olduğundan seri sola asimetriktir. 33

c) $\sigma = \sqrt{\frac{\sum fi(Xi - \bar{X})^2}{\sum fi}}$

$$\sigma = \sqrt{\frac{2843}{35} - (8,71)^2} = 2,31$$

$$\alpha_1 = \frac{\bar{X} - M_o}{\sigma}$$

$$\alpha_1 = \frac{8,71 - 8,88}{2,31} = -0,07$$

$$\alpha_2 = \frac{3(\bar{X} - M_e)}{\sigma}$$

$$\alpha_2 = \frac{3(8,71 - 8,75)}{2,31} = -0,05$$

$\alpha_1 < 0$ $\alpha_2 < 0$ olduğundan seri sola asimetriktir. 34

c) $Q_{h/r} = I_a + \frac{n\frac{h}{r} - \sum_{i=1}^{k-1} f_i}{f_b} c$

$$\frac{n}{4} = \frac{35}{4} = 8,75$$

$$\frac{3n}{4} = \frac{3(35)}{4} = 26,25$$

$$Q_1 = 6 + \frac{8,75 - 5}{8} 2 = 6,93$$

$$Q_2 = M_e = 8,75$$

$$Q_3 = 10 + \frac{26,25 - 25}{7} 2 = 10,35$$

$$A_s = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_2) + (Q_2 - Q_1)}$$

$$A_s = \frac{(10,35 - 8,75) - (8,75 - 6,93)}{(10,35 - 8,75) + (8,75 - 6,93)} = \frac{1,6 - 1,82}{1,6 + 1,82} = -0,06$$

Sonuç negatif olduğundan seri sola asimetriktir. Ancak asimetri katsayısının değeri sıfıra yakın olduğundan seri simetrik kabul edilebilir. 35

Momentler ile Asimetri ve Basıklık Ölçüleri

Tanım: Seri değerlerinin belirli bir orijine göre çeşitli derecelerden kuvvetlerinin ortalamalarına moment adı verilir.

Basit Seriler için

$$_a m_q = \frac{\sum_{i=1}^{n} (X_i - a)^q}{n}$$

Frekans ve Gruplanmış Seriler için

$$_a m_q = \frac{\sum_{i=1}^{k} f_i (X_i - a)^q}{\sum_{i=1}^{k} f_i}$$

a=0 ise orijine yada 0'a göre momentler (m)

a= $\overline{X}$ ise ortalamaya göre momentler (μ)

38

Momentler

Örnek: Aşağıda verilen frekans serisinin orijine ve ortalamaya göre momentlerini hesaplayınız.

X_i	f_i
1	12
2	17
3	20
4	21
5	10
	80

Orijine göre momentler:

$$m_1 = \frac{\sum_{i=1}^{5} f_i X_i}{\sum_{i=1}^{5} f_i} = \frac{240}{80} = 3 = \overline{X}$$

$$m_2 = \frac{\sum_{i=1}^{5} f_i X_i^2}{\sum_{i=1}^{5} f_i} = \frac{846}{80} = 10,575 = K^2$$

$$m_3 = \frac{\sum_{i=1}^{5} f_i X_i^3}{\sum_{i=1}^{5} f_i} = \frac{3282}{80} = 41,025$$

$$m_4 = \frac{\sum_{i=1}^{5} f_i X_i^4}{\sum_{i=1}^{5} f_i} = \frac{13530}{80} = 169,125$$

38

Momentler

Ortalamaya göre momentler:

X_i	f_i
1	12
2	17
3	20
4	21
5	10
	80

$$\mu_1 = \frac{\sum_{i=1}^{5} f_i (X_i - \overline{X})}{\sum_{i=1}^{5} f_i} = \frac{0}{80} = 0$$

$$\mu_2 = \frac{\sum_{i=1}^{5} f_i (X_i - \overline{X})^2}{\sum_{i=1}^{5} f_i} = \frac{126}{80} = 1{,}575 = \sigma^2$$

$$\mu_3 = \frac{\sum_{i=1}^{5} f_i (X_i - \overline{X})^3}{\sum_{i=1}^{5} f_i} = \frac{-12}{80} = -0{,}15$$

$$\mu_4 = \frac{\sum_{i=1}^{5} f_i (X_i - \overline{X})^4}{\sum_{i=1}^{5} f_i} = \frac{390}{80} = 4{,}875$$

39

Momentler Yardımıyla Hesaplanan Asimetri Ölçüsü

Momentler ile asimetri ölçüsü α_3 aşağıdaki şekilde hesaplanır.

$$\alpha_3 = \frac{\mu_3}{\mu_2^{3/2}} \qquad \mu_2 = \sigma^2 \qquad \alpha_3 = \frac{\mu_3}{\sigma^3}$$

$\alpha_3 = 0$ ise seri simetrik

$\alpha_3 < 0$ ise seri sola çarpık

$\alpha_3 > 0$ ise seri sağa çarpıktır

41

Momentler Yardımıyla Hesaplanan Asimetri Ölçüsü

Örnek:

Sınıf	X_i	f
0-4	2	15
4-8	6	44
8-12	10	30
12-16	14	18
16-20	18	13
		120

$$\overline{X} = \frac{\sum f_i X_i}{\sum f_i} = \frac{1080}{120} = 9$$

$$\mu_2 = \sigma^2 = \frac{\sum f_i (X_i - \overline{X})^2}{\sum f_i} = \frac{2664}{120} = 22,2$$

$$\sigma = \sqrt{22,2} = 4,71$$

$$\mu_3 = \frac{\sum f_i (X_i - \overline{X})^3}{\sum f_i} = \frac{5424}{120} = 45,2$$

$$\alpha_3 = \frac{\mu_3}{\mu_2^{3/2}} = \frac{\mu_3}{\sigma^3} = \frac{45,2}{(4,71)^3} = 0,43$$

α3 >0
seri sağa çarpıktır

42

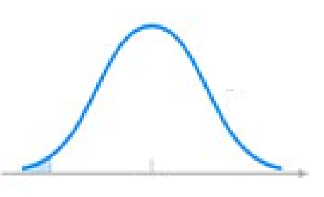
Momentler ile Hesaplanan Basıklık Ölçüsü

Momentler ile basıklık ölçüsü α_4 aşağıdaki şekilde hesaplanır.

$$\alpha_4 = \frac{\mu_4}{\mu_2^2} \qquad \mu_2 = \sigma^2 \qquad \alpha_4 = \frac{\mu_4}{\sigma^4}$$

Normal dağılım için α_4=3 olduğundan

- α_4 =3 ise seri normal
- α_4 <3 ise seri basık
- α_4 >3 ise seri sivri yada yüksektir.



Normal dağılım

43

Basıklık Ölçüsü

Örnek: Verilen serinin basıklığını inceleyiniz.

Sınıf	X_i	f
0-4	2	15
4-8	6	44
8-12	10	30
12-16	14	18
16-20	18	13
		120

$$\overline{X}=\frac{\sum f_i X_i}{\sum f_i}=\frac{1080}{120}=9$$

$$\mu_2=\sigma_2=\frac{\sum f_i (X_i-\overline{X})^2}{\sum f_i}=\frac{2664}{120}=22{,}2$$

$$\mu_4=\frac{\sum f_i (X_i-\overline{X})^4}{\sum f_i}=\frac{136152}{120}=1134{,}6$$

$$\alpha_4=\frac{\mu_4}{\mu_2^{\,2}}=\frac{\mu_4}{\sigma^4}=\frac{1134{,}6}{(22{,}2)^4}=\frac{1134{,}6}{492{,}84}=2{,}230$$

α4 <3 seri basık

44

SPSS

İNDEKSLER

- *İndeks Kavramı*
- *Mekan İndeksleri*
- *Zaman İndeksleri*
- *Bileşik İndeksler*

İNDEKS KAVRAMI

İstatistik oranlarının bir türü olan indeksler kısaca, sayısal değerlerin birbirine oranlanması ile elde edilir ve hesaplanabilmeleri için en az iki değerin bilinmesi gerekmektedir.

Bir olaya ait bir veya birden fazla değişkenin, farklı yerlerdeki veya zaman içindeki oransal değişimini belirlemek için hesaplanan sayılar, "indeks sayılar" olarak tanımlanır. İndeks sayılar, kısaca "indeksler" diye adlandırılmaktadır. İndeksler, fiyat, üretim, yatırım, ücret, satış değişmelerinin belirlenmesi gibi farklı olaylarda kullanılabilir.

İndeksler, hesaplandıkları seri türüne göre zaman ve mekan indeksleri; hesaplandıkları madde sayısına göre basit ve bileşik indeksler olmak üzere ikiye ayrılırlar. Zaman ve mekan indeksleri aynı zamanda basit veya bileşik indeksler olabilir.

Zaman indeksleri, hesaplanmalarında temel alınan devreye göre sabit ve değişken esaslı indeksler olarak, bileşik indeksleri ise hesaplanmalarına konu olan maddelerin önem derecelerine göre tartılı ve tartısız bileşik indeksler olarak ikiye ayrılırlar.

Mekan İndeksleri

Mekan serilerinin mekan içindeki oransal değişimini belirlemek için hesaplanan indekslere mekan indeksleri denir. Mekan indeksleri, nüfus, üretim, tüketim, fiyat gibi çeşitli değişkenlerin bölgeler, iller gibi mekan birimlerine göre oransal değişimlerinin belirlenmesi için kullanılırlar.

Mekan indeksleri seriyi oluşturan değerlerden her birinin serinin aritmetik ortalamasına oranlanması ile hesaplanır.

$$I_i = \frac{X_i}{\overline{X}_i} .100$$

I_i = Hesaplanan indeks

X_i = Seri değeri

$\overline{X}$ = Seri değerinin artimetik ortalaması

Örnek: **Beş ayrı bölgenin 1992 yılı buğday üretim miktarları aşağıda verilmiştir. (1=10.000 ton) Bu bölgelerin mekan indekslerini hesaplayınız.**

Bölgeler	Buğ.Üret.(X_i)
A	15
B	30
C	9
D	27
E	39
	120

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{120}{5} = 24$$

Mekan İndeksleri;

$$I_A = \frac{X_A}{\overline{X}}.100 = \frac{15}{24}.100 = 62{,}5$$

$$I_B = \frac{X_B}{\overline{X}}.100 = \frac{30}{24}.100 = 125$$

$$I_C = \frac{X_C}{\overline{X}}.100 = \frac{9}{24}.100 = 37{,}5$$

$$I_D = \frac{X_D}{\overline{X}}.100 = \frac{27}{24}.100 = 112{,}5$$

$$I_E = \frac{X_E}{\overline{X}}.100 = \frac{39}{24}.100 = 162{,}5$$

Bu sonuçlara göre, 1992 yılı buğday üretim miktarı, A, B, C, D, E bölgeleri ortalamasına göre, A bölgesinde %37,5 (=100-62,5), C bölgesinde %62,5 (=100-37,5) daha az, B bölgesinde %12,5 (=112,5-100) ve E bölgesinde %62,5 (=162,5-100) daha fazladır.

Zaman İndeksleri

Zaman serilerinin oransal değişimini belirlemek için hesaplanan indekslere zaman indeksleri denir.

Herhangi bir zaman birimine ait indeksin hesaplanmasında, zaman serisinin herhangi bir zaman biriminin değeri temel alınarak, diğerleri bu değere oranlanabileceği gibi, her zaman biriminin değeri bir önceki zaman biriminin değerine de oranlanabilir. Buna dayanarak zaman indeksleri "sabit esaslı" ve "değişken esaslı" olarak ikiye ayrılır.

SPSS

o *Sabit esaslı indeksler*

Zaman indekslerinde hesaplanacak zaman birimlerini devre olarak adlandırırsak; belirli bir devrenin değerini temel alıp, diğer devrelerdeki oransal değişmeleri bu değere göre belirlemek için hesaplanan indekslere, "sabit esaslı indeksler" denir.

Sabit esaslı indekslerde hesaplanmaya temel alınan devre, "temel devre" veya "esas devre" olarak adlandırılır ve değeri 100'e eşittir.

$$I_i = \frac{X_i}{X_0}.100$$

I_i = i'nci devrenin indeks sayısı

X_i = indeksi hesaplanacak devrenin değeri

X_0 = esas devre değeri

Not: İndekslerde değişken genellikle x_i ile gösterilse de, fiyat indekslerinde p_i, miktar indekslerinde ise q_i değişkenleri kullanılmaktadır.

Örnek: Aşağıda bir malın 6 yıllık satış fiyatları verilmiştir. Bu verilerden yararlanarak 1994 yılı temel yıl ve 1997 yılı temel yıl olarak

Yıllar	p_i (A malı fiyatı)	I_i (1994=100)	I_i (1997=100)
1994	35	100,0	125,0
1995	40	114,2	142,8
1996	32	91,4	114,2
1997	28	80,0	100,0
1998	50	142,8	178,5
1999	73	208,5	260,7

1994=100 (1994 temel yıl ise)

$$I_i = \frac{p_i}{p_0}.100$$

$$I_{94} = \frac{p_{94}}{p_{94}}.100 = \frac{35}{35}.100 = 100,0$$

$$I_{95} = \frac{p_{95}}{p_{94}}.100 = \frac{40}{35}.100 = 114,2$$

$$I_{96} = \frac{p_{96}}{p_{94}}.100 = \frac{32}{35}.100 = 91,4$$

$$I_{97} = \frac{p_{97}}{p_{94}}.100 = \frac{28}{35}.100 = 80,0$$

$$I_{98} = \frac{p_{98}}{p_{94}}.100 = \frac{50}{35}.100 = 142,8$$

$$I_{99} = \frac{p_{99}}{p_{94}}.100 = \frac{73}{35}.100 = 208,5$$

1997=100 (1997 temel yıl ise)

$$I_i = \frac{p_i}{p_0}.100$$

$$I_{94} = \frac{p_{94}}{p_{97}}.100 = \frac{35}{28}.100 = 125,0$$

$$I_{95} = \frac{p_{95}}{p_{97}}.100 = \frac{40}{28}.100 = 142,8$$

$$I_{96} = \frac{p_{96}}{p_{97}}.100 = \frac{32}{28}.100 = 114,2$$

$$I_{97} = \frac{p_{97}}{p_{97}}.100 = \frac{28}{28}.100 = 100,0$$

$$I_{98} = \frac{p_{98}}{p_{97}}.100 = \frac{50}{28}.100 = 178,5$$

$$I_{99} = \frac{p_{99}}{p_{97}}.100 = \frac{73}{28}.100 = 260,7$$

1994 yılı temel yıl olarak alındığında yapılabilecek yorum:

Malın fiyatları 1994 yılına göre 1995 yılında %14,2, 1998 yılında %42,8, 1999 yılında %108,5 artarken; 1996 yılında %8,6, 1997 yılında %20 azalmıştır.

1995 yılı temel yıl olarak alındığında yapılabilecek yorum:

Malın fiyatları 1997 yılına göre 1994 yılında %25, 1995 yılında %42,8, 1996 yılında %14,2, 1998 yılında %78,5 ve 1999 yılında %160,7 oranında yükselmiştir.

Sabit esaslı indekslerin hesaplanmasında, bir devre yerine, birkaç devrelik bir dönem de temel alınabilir. Bu durumda birkaç devrelik dönemin aritmetik ortalaması alınarak, bu ortalama esas devre değeri kabul edilir ve diğer devre değerleri hesaplanan bu ortalamaya oranlanır.

Örnek: Aşağıda bir bölgenin yıllık kömür üretim miktarları verilmiştir. Bölgenin yıllık kömür üretimine ait sabit esaslı indeksleri 1993-1995 dönemi temel alarak hesaplayınız.

Yıllar	Kömür üretimi (1=10000ton)
1993	11
1994	18
1995	16
1996	19
1997	13
1998	21

$$\bar{q}_0 = \frac{q_{93}+q_{94}+q_{95}}{3} = \frac{11+18+16}{3} = \frac{45}{3} = 15$$

$$I_i = \frac{p_i}{p_0}.100$$

$$I_{93} = \frac{q_{93}}{q_{93-95}}.100 = \frac{11}{15}100 = 73{,}3$$

$$I_{94} = \frac{q_{94}}{q_{93-95}}.100 = \frac{18}{15}100 = 120{,}0$$

$$I_{95} = \frac{q_{95}}{q_{93-95}}.100 = \frac{16}{15}100 = 106{,}7$$

$$I_{96} = \frac{q_{96}}{q_{93-95}}.100 = \frac{19}{15}100 = 126{,}6$$

$$I_{97} = \frac{q_{97}}{q_{93-95}}.100 = \frac{13}{15}100 = 86{,}66$$

$$I_{98} = \frac{q_{98}}{q_{93-95}}.100 = \frac{21}{15}100 = 140{,}0$$

Yorum: Bölgenin yıllık kömür üretim miktarı 1993-1995 dönemine (1993-1995 yılları ortalamasına) göre; 1993'de %16,7 daha az, 1994'de %20, 1995'de %6,7, 1996'da %26,6, 1998'de %40 daha az olmuştur.

o Değişken esaslı indeksler

Zaman indekslerinde, indeksi hesaplanacak devre değerinin bir önceki devre değerine oranlanması ile hesaplanacak indeksler “değişken esaslı indeksler” olarak adlandırılır.

Değişken esaslı indeksler her devre indeksinin bir önceki devreye göre oransal artışını veya azalışını gösterecektir. Bu indekslerde hesaplanacak her devre için bir önceki devre 100 kabul edilmektedir.

I_i = i’nci devrenin indeks sayısı

X_i = indeksi hesaplanacak devrenin değeri

X_{i-1} = bir önceki devre değeri

$$I_i = \frac{X_i}{X_{i-1}}.100$$

SPSS

Örnek: Aşağıda bir fabrikada üretilen yıllık buzdolabı miktarları verilmiştir. Bu verilerden yararlanarak değişken esaslı indeksleri hesaplayınız.

Yıllar	Üretim mik (1=1000ad)	I_i
1995	5	-
1996	4	80
1997	3	75
1998	6	200
1999	8	133

$$I_i = \frac{q_i}{q_{i-1}} 100$$

$$I_{96} = \frac{4}{5} 100 = 80$$

$$I_{97} = \frac{3}{4} 100 = 75$$

$$I_{98} = \frac{6}{3} 100 = 200$$

$$I_{99} = \frac{8}{6} 100 = 133$$

SPSS

Yıllar	Üretim mik	I_i
1995	5	-
1996	4	80
1997	3	75
1998	6	200
1999	8	133

Yorum: Fabrikanın yıllık üretim miktarı, 1996 yılında 1995 yılına göre %20 azalmış; 1997 yılında 1996 yılına göre %25 azalmış; 1998 yılında 1997 yılına göre %100 artmış; 1999 yılında 1998 yılına göre %33 artmıştır.