İstatistik I

3. Hafta

21 Ekim 2021

3. Merkezi Eğilim Ölçüleri

www.umitsarp.com

Ümit SARP, umit.sarp@ikcu.edu.tr

İZMİR KÂTİP ÇELEBİ ÜNİVERSİTESİ

Giriş

- Merkezi Eğilim Ölçüleri
  - Analitik (duyarlı) ortalamalar
  - Analitik olmayan (duyarsız) ortalamalar
  - Aritmetik ortalama
  - Geometrik ortalama
  - Tepe Değer (Mod)
  - Ortanca (Medyan)

Merkezi Eğilim Ölçüleri

Merkezi eğilim ölçüleri bir seri hakkında bilgi edinmemizi ve çeşitli kıyaslamalar yapmamıza olanak sağlayan ölçütlerdir. Verilerin daha çok hangi değerlere eğilimli olduklarını bulmak için merkezi eğilim ölçülerini kullanırız. Bu ölçüler analitik olan (duyarlı) ve analitik olmayan (duyarsız) ortalamalardır. Analitik ortalamalar; aritmetik ortalama, tartılı ortalama, kareli ortalama, harmonik ortalama ve geometrik ortalamadır. Analitik olmayan ortalamalar ise mod ve medyandır [3].

Merkezi Eğilim Ölçüleri

Analitik (duyarlı) ortalamalar

Analitik ortalamalara "duyarlı" denmesinin nedeni; **veri kümesindeki tüm gözlem birimlerini hesaba katmasıdır.** Bu hesaplama yöntemi, özellikle uç değerler (uç değer bir diğer ifade ile sapan değer, serideki değerlerden çok farklı olan değerlerdir) söz konusu olduğunda ortalamayı uç değerlere doğru çekerek doğru tahmin yapmayı zorlaştırırlar. Bu tip durumlarda ya uç değerler hesaplamadan dışlanır, ya da analitik olmayan ortalama türlerinden faydalanılır. Bu duruma en güzel örnek, bir sınıfın ortalama notunu hesaplarken en yüksek ve en düşük değerin dışlanarak hesaba alınmamasıdır [3].

Merkezi Eğilim Ölçüleri

Analitik ortalamalar;

- aritmetik ortalama*
- tartılı ortalama
- geometrik ortalama
- kareli ortalama
- harmonik ortalama

Analitik olmayan (duyarsız) ortalama

Analitik olmayan ortalamalara “duyarsız” denmesinin nedeni, **özellikle uç değerlerin olduğu durumlarda bu gibi değerleri göz ardı etmesi, tüm gözlemleri hesaba almaması ve sıralamaya dayanmasından kaynaklanmaktadır.** Bu sıralamanın ne anlama geldiğini ilerideki sayfalarda göreceğiz [3].

Merkezi Eğilim Ölçüleri

Analitik olmayan (duyarsız) ortalama;

- Tepe Değer (Mod)
- Ortanca (Medyan)

Merkezi Eğilim Ölçüleri

Ortalma hesaplanacak seride gözlem değerlerinin hepsi eşit önemde ise "Basit Aritmetik Ortalama" farklı önemde ise "Tartılı Aritmetik Ortalma" söz konusudur.

Basit Aritmetik Ortalama

Ortalamalar içinde en yaygın biçimde kullanılan ortalma biçimidir. Ortalma denilince ilk akla gelen ortalma türüdür. Aritmetik ortalma; gözlem değerlerinin cebirsel toplamının gözlem sayısına bölünmesi ile hesaplanan ortalamadır.
Örneklemde $\overline{X}$, kitlede (evrende) μ ile gösterilir [2].

Merkezi Eğilim Ölçüleri

Basit Aritmetik Ortalama

Ana kitlenin tamamı N hacimli ve sonlu olduğunda bu kitlenin sahip olduğu veriler $X_1, X_2, X_3, \ldots X_N$ ile ifade edilir. Bu durumda hesaplanacak aritmetik ortalama evrenin aritmetik ortalmasıdır ve μ ile gösterilir. Evrenin aritmetik ortalması;

$$\mu = \frac{\sum_{i=1}^{N} X_i}{N}$$

formülü ile hesaplanır [2].

Merkezi Eğilim Ölçüleri

Basit Aritmetik Ortalama

Burada $\sum_{i=1}^{N} X_i$: seri değerlerinin cebirsel toplamını,

N: gözlem sayısını (evrenin hacmini) gösterir.

Merkezi Eğilim Ölçüleri

Basit Aritmetik Ortalama

Aritmetik ortalma n-hacimlik **örneklemden** elde edilen $X_1, X_2, X_3, \ldots X_n$ veri kümesi için hesaplanacak olursa, $\overline{X}$ (X-ortalama veya X-bar diye okunur) simgesi ile gösterilir, örneklemin aritmetik ortalması olarak adlandırılır. Örneklemin aritmetik ortalması;

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

formülü ile hesaplanır [2].

Merkezi Eğilim Ölçüleri

En çok bilinen ve en sık kullanılan ortalamadır.

Sıfır içeren serilerde harmonik ortalama hesaplanamaz, geometrik ortalama hesaplanabilir ama anlamsız olur.

Seride negatif değer var ise, geometrik ortalama hesaplanamaz.

Ancak aritmetik ortalama her türlü veri seti için hesaplanabilir ve tek bir tanedir.

Merkezi Eğilim Ölçüleri

Aritmetik ortalama cebirsel işlemlere elverişli olması, kolay hesaplanması ve anlaşılması sebebiyle de yaygın kullanım alanına sahiptir. Nitekim ortalama dendiğinde akla sadece aritmetik ortalama gelmektedir.

Merkezi Eğilim Ölçüleri

Cebirsel işlemlere elverişli olması, aritmetik ortalamanın pek çok istatistik analiz yönteminin temelinde veya içeriğinde yer almasına sebep olmuştur.

Merkezi Eğilim Ölçüleri

Aritmetik ortalama, artış miktarı sabit aritmetik dizi özelliği gösteren serilerde daha temsili nitelik taşımaktadır. Başka bir deyişle, aritmetik ortalama bu tür serileri daha iyi temsil etmektedir.

Merkezi Eğilim Ölçüleri

Aritmetik ortalamanın kullanılabilmesi için verinin en az aralık ölçeğinde toplanması gerekir. Nominal ve ordinal ölçekte toplanmış veriler için aritmetik ortalama kullanılması uygun değildir.

Merkezi Eğilim Ölçüleri

Örnek:

Bir işyerinde çalışan 8 kişinin aylık kazançları aşağıda verilmiştir. Bu işçilerin aylık ortalma kazançlarını hesaplayalım.

$$\begin{array}{c} \underline{X_i} \\ X_1 = 640 \\ X_2 = 800 \\ X_3 = 860 \\ X_4 = 980 \\ X_5 = 1120 \\ X_6 = 1160 \\ X_7 = 1560 \\ X_8 = 1680 \\ \hline \end{array}$$

$$\sum_{i=1}^{8} X_i = 8800$$

Merkezi Eğilim Ölçüleri

Örnek:

İş yerinde çalışan toplam 8 kişinin aylık kazanç değişkeni incelenmiş ve tam sayım uygulanmıştır. Evrenin hacmi N=8 işçidir ve gözlem değerleri $\mu = \sum_{i=1}^{8} X_i = 8800$ olduğundan evrenin aritmetik ortalması:

$$\mu = \frac{\sum X}{N} = \frac{8800}{8} = 1100$$

olarak bulunur.

Hesaplanan aylık ortalma kazanç temsilidir.

Merkezi Eğilim Ölçüleri

Örnek:

Bir önceki örneği tekrar düşünelim bu kez dört kişiden oluşan bir örneklem alarak örneklemin aritmetik ortalamasını bulalım.

X_i
640
860
980
1160

$$\sum_{i=1}^{4} X_i = 3640$$

Merkezi Eğilim Ölçüleri

Örnek:

Çalışan işçilerden dört işçinin aylık kazançları gözlenmiştir. Bu nedenle $n = 4$'dir. Örneklemin aritmetik ortalaması:

$$\overline{X} = \frac{\sum X}{n} = \frac{3640}{4} = 910$$

olarak bulunur.

Hesaplanan aylık ortalma kazanç temsilidir.

Merkezi Eğilim Ölçüleri

Sıra sizde, Örnek:

Bir köyde yaşayan aileler arasında seçilen 6 ailenin aylık gelirleri aşağıdaki gibi olsun. Bu ailelerin aylık gelir ortalaması nedir?

X_i
200
300
400
500
600
8000

Toplamlarını bulalım;

Merkezi Eğilim Ölçüleri

Sıra sizde, Örnek:

Toplam;

X_i
200
300
400
500
600
8000
10000

Merkezi Eğilim Ölçüleri

Sıra sizde, Örnek:

Toplam;

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

- Neden $\overline{X}$ kullandık; μ kullanmadık?!
- n değeri kaçtır söylebilir misiniz?

Merkezi Eğilim Ölçüleri

Sıra sizde, Örnek:

Toplam;

$$X = \frac{\sum_{i=1}^{6} X_i}{6} = \frac{X_1 + X_2 + X_3 + X_4 + X_5 + X_6}{6} = \frac{10000}{6} = 1666,67$$

Merkezi Eğilim Ölçüleri

Sıra sizde, Örnek:

Sonuç olarak;

Hesaplanan ortalama değer, beşinci ve altıncı gözlem değerleri arasında yer alır. Gözlemlerin merkezinde değildir.

Bu sonuca göre bu özellikte uç değerler içeren seriler için aritmetik ortalama uygun bir method değildir. Uç değer içeren seriler için uygun methodlar önümüzdeki haftalarda anlatılacaktır.

Merkezi Eğilim Ölçüleri

Frekans Serilerinde Aritmetik Ortalama

Frekans serisi şeklinde düzenlenen serilerin aritmetik ortalması, gözlem değerlerinin frekanslarla çarpımları toplamının frekanslar toplamına oranı olarak tanımlanır. Frekans Serilerinde Aritmetik Ortalama;

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i \times f_i}{\sum_{i=1}^{n} f_i}$$

formülü ile hesaplanır [2].

Merkezi Eğilim Ölçüleri

Örnek:

Bir işletmenin kargo ile gönderdiği kolilerin dağılımı aşağıda verilmiştir. Kolilerin ağırlıklarının aritmetik ortlaması nedir?

X_i	f_i	$X_i \times f_i$
10	1	$10 \times 1 = 10$
20	2	$20 \times 2 = 40$
30	4	$30 \times 4 = 120$
40	2	$40 \times 2 = 80$
50	1	$50 \times 1 = 50$
	10	300

Merkezi Eğilim Ölçüleri

Örnek:

$$\bar{X} = \frac{\sum_{i=1}^{10} X_i \times f_i}{\sum_{i=1}^{10} f_i} = \frac{300}{10} = 30$$

Dikkat edersek düzgün artan teorik bir seri olduğu için ortalmanın tam ortayı ifade ettiğini görebiliriz.

Merkezi Eğilim Ölçüleri

Sınıflanmış Frekans Serilerinde Aritmetik Ortalama

Sınıflanmış frekans serisi şeklinde düzenlenen serilerin aritmetik ortalması, önce frekans serisine dönüştürülür ardından frekans serisi için kullanılan aritmetik ortlama formülü kullanıllır.

Merkezi Eğilim Ölçüleri

Örnek:

Aşağıdaki sınıflanmış frekans serisini düşünelim.

Gruplar	f_i		X_i	$X_i \times f_i$
100 – 200	7		150	1050
200 – 300	18		250	4500
300 – 400	25	$\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow$	350	8750
400 – 500	30		450	13500
500 – 600	20		550	11000
	100			38800

Dikkat ederseniz gruplarda yer alan verilerin orta noktaları kullanılarak dönüştürülmüştür. Sizce sebebi nedir?

Merkezi Eğilim Ölçüleri

Tartılı Aritmetik Ortalama

Bir istatistik serisindeki bazı gözlem değerlerinin önem derecesi farklılık gösterebilir. Bu önem derecesinin hesaba katılmasına tartılı aritmetik ortalma denir. t_i i'ninci gözlem değerini gösterir. Tartılı Aritmetik Ortalama;

$$\bar{X}_t = \frac{\sum\limits_{i=1}^{n} X_i \times t_i}{\sum\limits_{i=1}^{n} t_i}$$

formülü ile hesaplanır [2].

Merkezi Eğilim Ölçüleri

Örnek:

Kredili sistemde öğrenim gören bir coğrafya bölümü örencisinin almış olduğu dersler, derslerin kredi değerleri (tartı) ve örencinin başarı puanları aşağıda verilmiştir. Öğrencinin ortalma başarı puanı nedir?

Ders	Başarı Notu	Kredi Değeri (tartı)
Çevre ve Ekoloji	2 (CC)	3
TDE	4 (AA)	4
ATA	3 (BB)	3
İstatistik	1 (DD)	2
Bilim Felsefesi	4 (AA)	4

Merkezi Eğilim Ölçüleri

Örnek:

X_i	t_i	$X_i \times t_i$
2	3	6
4	4	16
3	3	9
1	2	2
4	4	16
14	16	49

Merkezi Eğilim Ölçüleri

Örnek:

Basit aritmetik ortalama;

$$\bar{X} = \frac{\sum X_i}{n} = \frac{14}{5} = 2,8$$

Tartılı aritmetik ortalama;

$$\bar{X}_t = \frac{\sum_{i=1}^{n} X_i \times t_i}{\sum_{i=1}^{n} t_i} = \frac{49}{16} = 3,06$$

Merkezi Eğilim Ölçüleri

Geometrik Ortalama

Gözlem değerleri çarpımının gözlem sayısına eşit derecede kökü alınarak hesaplanan ortalmadır. Geometrik Ortalama;

$$GO = \sqrt[n]{X_1 \times X_2 \times X_3 \times \ldots \times X_n}$$

formülü ile hesaplanır [2].

Merkezi Eğilim Ölçüleri

Örnek:

Aynı bölgede bulunan dört kömür madenindeki rezervler 2001 yılından bugüne kadar sırayla 3,2,4 ve 6 kat (misli) artış göstermiştir. Buna göre GO kaçtır?

$$GO = \sqrt[4]{3 \times 2 \times 4 \times 6} = \sqrt[4]{144} = (144)^{1/4} = 3,46$$

Merkezi Eğilim Ölçüleri

Kareli Ortalama

Gözlem değerleri karelerinin aritmetik ortalmasının kareköküne eşit olan ortalmadır. Sıfır değerli ya da negatif işaretli değişkenler olduğunda hesaplanamaz. Kareli Ortalama;

$$KO = \sqrt{\frac{\sum X_i^2}{n}}$$

formülü ile hesaplanır [2].

Merkezi Eğilim Ölçüleri

Örnek:

Merkeze olan birim cinsinden belirlendiği bir harita örnekleminde negatif değerler batıyı pozitif değerler doğu göstermektedir. Buna göre merkeze uzaklıkları $-3, -2, 1, 3, 3$ olan noktaların kareli ortalaması nedir?

X_i	X_i^2
-3	9
-2	4
1	1
3	9
3	9

$$\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow KO = \sqrt{\frac{32}{6}} = 2,53$$

Bu dizi için kareli ortalama uygun değildir. Çünkü dizide negatif terimler vardır.

Merkezi Eğilim Ölçüleri

Harmonik Ortalama

Oransal olarak belirtilebilen değişkenlerin ortalamalarının hesaplanmasında harmonik ortalama kullanılır. Sıfır değerli ya da negatif işaretli değişkenler olduğunda hesaplanamaz. Harmonik Ortalama;

$$HO = \frac{n}{\sum\limits_{i=1}^{n} \frac{1}{X_i}}$$

formülü ile hesaplanır [2].

Merkezi Eğilim Ölçüleri

Örnek:

Aşağıdaki veriler için harmonik ortalmayı hesaplayalım?

X_i	$^1/_{X_i}$
4	0,25
5	0,2
7	0,14
8	0,12
16	0,06

$$\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow HO = \frac{n}{\sum_{i=1}^{n} \frac{1}{X_i}} = \frac{5}{0,77} = 6,49$$

Merkezi Eğilim Ölçüleri

Analitik olmayan (duyarsız) ortalamalar;

Tepe Değer (Mod)

Küçükten büyüğe doğru sıralanmış gözlem serisinde en fazla tekrarlanan (frekansı en çok olan) verisine tepe değeri veya mod ismi verilir [2].

Mod, verilerin en çok tekrar eden değer etrafında toplandığını ifade eden bir ölçüdür. Veri grubunu betimlemede, tüm verilerden ziyade en çok tekrar eden verinin kullanılmasından dolayı mod, diğer merkezi eğilim ölçülerine kıyasla veri hakkında en az bilgi veren ölçüdür. Hiçbir aritmetik işlem gerektirmez.

Merkezi Eğilim Ölçüleri

Örnek:

Aşağıdaki veriler için mod değerini bulalım?

$$1, 1, 1, 2, 2, 3, 3, 4, 4, 4, 4, 4, 4, 4, 4, 5, 5, 7, 8, 8, 15, 45, 72$$

Cevap?

Merkezi Eğilim Ölçüleri

Örnek:

Aşağıdaki veriler için mod değerini bulalım?

$$1, 1, 1, 2, 2, 3, 3, 4, 4, 4, 4, 4, 4, 4, 4, 5, 5, 7, 8, 8, 15, 45, 72$$

Cevap=4

Merkezi Eğilim Ölçüleri

Mod, duyarsız ortalamalar içinde en temsili ortalamadır. Hesabının kolay ve anlaşılır olması sebebiyle duyarsız bir ortalama kullanılması gerekiyorsa tercih edilebilecek bir ortalamadır.

Merkezi Eğilim Ölçüleri

Mod, seride yer alan bütün birimleri dikkate almaksızın hesaplanır. Nitekim en çok tekrar eden değer olarak tanımladığımız modun değerinin ne olacağı üzerinde frekansı daha az olan diğer gözlem değerlerinin hiçbir etkisi ve katkısı bulunmamaktadır. Bu özelliği sebebiyle mod, seride yer alan aşırı küçük ya da aşırı büyük değerlerden etkilenmemesi nedeniyle en temsili ortalama olma özelliği kazanır. Ancak öte yandan, seride yer alan tüm gözlem değerlerinden hareketle hesaplanmaması modu, cebirsel işlemler açısından elverişsiz kılar ve daha ileri istatistiksel analizler açısından uygun bir ortalama olmamasına yol açar.

Merkezi Eğilim Ölçüleri

Dolayısıyla mod, bir seriyi en iyi temsil eden ortalama olarak, amaç sadece bir seriyi özetlemek veya seriyi tek bir değer ile ifade etmek olduğunda rahatlıkla kullanılabilir bir ortalama olma özelliği taşır. Fakat, amaç seriye yönelik bir takım istatistiksel analizler yapmaksa o takdirde mod yetersiz bir ortalama olacaktır.

Merkezi Eğilim Ölçüleri

Mod, nicel vasıflara uygulanabilirliği yanında nitel vasıflara da uygulanabilen bir özellik taşır. Nicel, yani sayısal verilere yönelik örnekleri bu başlık altında gördük. Nitel vasıflara yönelik olarak da örneğin, bir toplulukta kadın ve erkek kişi sayısının ortalaması hesaplanmak istendiğinde, ya da eğitim düzeyleri açısından insanlar tasnif edildiğinde en yüksek frekansı belirlemek suretiyle modu tespit etmek mümkündür.

Merkezi Eğilim Ölçüleri

Ortanca (Medyan)

Basit seri özelliği gereği verilerin küçükten büyüğe ya da büyükten küçüğe sıralandığı seridir. Böyle bir seride sıralamanın tam ortasında yer alan değer medyandır.

Basit seride medyanın hangi sırada yer alan değer olduğunu belirlemek üzere aşağıdaki basit formül kullanılmaktadır:

$$\text{Medyanın Sıra Değeri} = \frac{n+1}{2}$$

Merkezi Eğilim Ölçüleri

Örnek:

Aşağıdaki veriler için medyan değerini bulalım?

$$1, 1, 1, 2, 2, 3, 3, 4, 4, 4, 4, 4, 4, 4, 4, 5, 5, 7, 8, 8, 15, 45, 72$$

Cevap?

Merkezi Eğilim Ölçüleri

Örnek:

Aşağıdaki veriler için medyan değerini bulalım?

$$1, 1, 1, 2, 2, 3, 3, 4, 4, 4, 4, 4, 4, 4, 4, 5, 5, 7, 8, 8, 15, 45, 72$$

Cevap$=$ 4

Merkezi Eğilim Ölçüleri

Örnek:

Aşağıdaki veriler için medyan değerini bulalım?

$$1, 1, 1, 2, 2, 7, 8, 15, 45, 50, 60, 65, 72$$

Cevap?

Merkezi Eğilim Ölçüleri

Örnek:

Aşağıdaki veriler için medyan değerini bulalım?

$$1, 1, 1, 2, 2, 7, 8, 15, 45, 50, 60, 65, 72$$

Cevap$=$ 7

Ödev

1- Aşağıda bir sınıftaki öğrencilerin aldıkları notların dağılımı mevcuttur. Sınıfın ortalama başarısı nedir?

Sınıflar	Öğrenci Sayısı
$0-20$	10
$20-40$	15
$40-60$	20
$60-80$	5
$80-100$	2

Ödev

2- Bir öğrencinin vizeden aldığı not 50 ve finalden aldığı not 52 ‘dir. Bu sınıfta geçme notu 60 ise öğrenci ilgili dersten geçebilir mi? (Vize ağırlığı %30)

Ödev

3- 13, 13, 12, 19, 15, 2, 17, 12, 10, 10, 4, 5, 10, 13, 19 değerlerini, bir de sıklık serisi oluşturarak bu serilerde aritmetik ortalama, kareli ortalama, geometrik ortalama, mod ve medyan değerlerini hesaplayarak bu değerler arasındaki sıralamayı belirtin?

4. Hafta: Merkezi Dağılma Ölçüleri
İpucu : Standart Sapma

Kaynaklar I

[1] K. Mert Çubukçu,
"Planlamada ve Coğrafyada Temel İstatistik ve Mekansal İstatistik",
Ankara: Nobel Akademik Yayıncılık Eği. Dan. Tic. Ltd. Şti., (2015).

[2] A. Özmen, F. Er, M. Atlas, E. Şıklar,
"İstatistik (AÖF)",
Eskişehir: Anadolu Üniversitesi Yayınları, (2012).

[3] L. İşbilen Yücel,
"İstatistik Maliye Uzaktan Eğitim",
İstanbul Üniversitesi Açık ve Uzaktan Eğitim Fakültesi Ders Notu.

[4] Ö. Serper,
"Uygulamalı İstatistik 1",
Bursa: Ezgi Kitapevi, (2004).

Kaynaklar II

[5] Murat Komisyon,
"İstatistik",
Murat Açıköğretim Yayınları, (2004).

[6] N. Gürsakal, A. Oğuzlar,
"Betimsel İstatistik",
Dora Yayıncılık, (2019).

[7] Y. Baykul, C. O. Güzeller,
"Sosyal Bilimler için İstatistik Uygulamaları",
Ankara: Pegem Akademi, (2014).