MERKEZİ EĞİLİM ÖLÇÜTLERİ, ORTALAMALAR

Toplanan verileri özetlemek ve tek sayılı değerlere indirgemek için ortalamalar da denilen merkezi eğilim ölçütleri önemli kriterlerdir. Ortalamalar birçok halde incelenen olayın normal değerini gösterir ve istatistiksel analizlerde önemli bir dayanak noktasını oluşturur. Ayrıca ortalamalar kullanılarak seriler de karşılaştırılmaktadır.

1. Aritmetik Ortalama : Bir seride bulunan verilerin toplamının veri sayısına bölümünden elde edilen değerdir. $\overline{X}$ ile gösterilir. Aritmetik ortalamanın özellikleri;

- Bir serideki verilerle aritmetik ortalamanın farklarının kareleri toplamı minimumdur.

$$\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2=\text{min.}$$

- Bir serideki verilerle aritmetik ortalamanın farklarının toplamı sıfırdır.

$$\sum_{i=1}^{n}\left(X_i-\overline{X}\right)=0$$

- Bir serideki tüm verilere a sayısı eklenmesi, çıkarılması, çarpılması ve bölünmesiyle elde edilen yeni serinin aritmetik ortalaması ilk serinin aritmetik ortalamasına a sayısı eklenmesi, çıkarılması, çarpılması ve bölünmesiyle elde edilen değere eşittir.

$$Y_i = X_i + a, \quad \overline{Y} = \overline{X} + a, \quad Y_i = X_i - a, \quad \overline{Y} = \overline{X} - a$$
$$Y_i = X_i * a, \quad \overline{Y} = \overline{X} * a, \quad Y_i = X_i / a, \quad \overline{Y} = \overline{X}/a$$

n tane verisi olan bir örnek küme için aritmetik ortalama;

$$\overline{X} = \frac{X_1 + X_2 + ...+ X_n}{n} = \frac{\sum_{i=1}^{n} X}{n} = \frac{[X]}{n}$$ Deneysel Aritmetik ortalama,

Ölçü sayısının sonsuz olduğu durumdaki teorik aritmetik ortalama μ ile gösterilir ve

$$\mu = \frac{X_1 + X_2 + ...+ X_N}{N} = \frac{\sum_{i=1}^{n} X}{N} = \frac{[X]}{N}$$ eşitliğiyle hesaplanır.

2. Medyan (Ortanca Değer) : Bir veri grubundaki veriler küçükten büyüğe yada büyükten küçüğe sıralandığında veri grubunun tam ortasında kalan değere medyan (ortanca) denir. n veri sayısı olmak üzere medyan, n tek ise (n+1)/2 verisi, n çift ise n/2 ve (n/2)+1 verilerinin ortalamasıdır. Medyanın özellikleri;

- Medyan veri sayısının değişiminden etkilenir fakat uç değerlerin değişiminden etkilenmez.
- Veri değerleri ile medyanın farklarının yarısı pozitif yarı negatif olur.
- Cebirsel işlemlere uygun değildir.
- Medyan alt ve üst sınırları belli olmayan sınıfların varlığında önem kazanır.
- Verilerin medyandan farklarının mutlak değeri toplamı minimumdur.

Bir veri grubu için medyan hesaplanırken veriler sıralanır ve ortaya gelen değer n sayısının çift yada tek olması göz önünde bulundurularak hesaplanır.

3. Mod (Tepe Değeri) : Sıralanmış veri grubunda frekansı en büyük olan değer mod olarak alınır. Modun özellikleri;

- Veri grubundaki uç değerlerden etkilenmez,
- Cebirsel işlemlere elverişli değildir.
- Veri grubunda fazla ve tekrarlı ölçü olmayınca hesaplanamaz.
- Mod değeri veri grubundaki verilerin frekansı yüksekse en muhtemel değer olur.

Bir veri grubu için mod hesaplanırken veriler sıralanır ve en çok tekrar edilen veri mod olarak alınır.

DAĞILMA (YAYILMA), EĞİLİM VE BASIKLIK ÖLÇÜTLERİ

Bir veri grubunun merkezi eğilim ölçütlerinin bilmek veri grubunu tanımak için gereklidir fakat yeterli değildir. Aritmetik ortalaması eşit olan iki veri grubu çok farklı özellikler içerebilir. Bu nedenle merkezi eğilim ölçütlerinin yanında dağılma, eğilim ve basıklık ölçütleri de hesaplanmalıdır.

1. Dağılım Ölçütleri

1.1. Değişim Aralığı (Ranj): Veri grubundaki en büyük ve en küçük veri arasındaki fark değişim aralığı olarak adlandırılır. Değişim aralığı yalnızca maksimum ve minimum veriyle ilgili bir ölçüttür. Bu nedenle veri grubu hakkında çok bilgi vermez ve uygulamlarda çok fazla kullanılmaz.

Bir veri grubu için değişim aralığı, $R = X_{mak} - X_{min}$ eşitliğiyle hesaplanır. Frekanslı veri grubu için de bu değişmez. Sınıflandırılmış veri grubu için ise,değişim aralığı ilk sınıfın alt sınır değeri ile son sınıfın üst sınır değeri arasındaki farkla yada ilk ve son sınıfın orta değeri arasındaki farkla bulunur.

1.2. Ortalama Sapma: Veri grubundaki veriler ile aritmetik ortalama aralarındaki farkların mutlak değeri toplamının veri grubundaki ölçü sayısına bölümüyle ortalama sapma hesaplanır. Ortalama sapmanın özellikleri;

- Veri grubundaki tüm değerlerden etkilenir,
- Verilerle aritmetik ortalama arasındaki farkların ortalaması olduğu için kesin bir anlamı vardır,
- Verilerin tümü hesaba katıldığı için güvenilirdir.

- Mutlak değerle hesaplandığı için cebirsel işleme uygun değildir.

Bir veri grubunda ortalama sapma,

$$OS = \frac{\sum_{i=1}^{n} \left| X_i - \overline{X} \right|}{n}$$ eşitliğiyle hesaplanır.

1.3. Standart Sapma ve Varyans : Ortalama sapma farkların mutlak değeriyle hesaplandığı için cebirsel işlemlere çok uygun olmayan bir ölçüttür. Standart sapma ise, aritmetik ortalamadan farkların karelerinin veri sayısına bölümünün karekökü alınarak hesaplanır ve cebirsel işlemlere uygun bir ölçüt olarak karşımıza çıkar. Standart sapma teorik olarak σ, uygulamada ise S ile gösterilir. Standart sapmanın karesi yani verilerin ortalamadan farklarının ortalaması varyansı verir ve V(X) yada σ^2 ile gösterilir.

Standart sapmanın ve varyansın özellikleri;

- Veri grubundaki tüm verilerden etkilenir,
- Cebirsel işlemler için uygundur,
- Ortalama sapmaya göre uç değerlerden daha az etkilenir,
- Veriler aritmetik ortalamaya yakınsa standart sapma küçük, uzaksa büyük olur. Veriler eşitse standart sapma sıfır olur.
- Aritmetik ortalaması eşit iki veri grubunun standart sapması küçük olanın verileri aritmetik ortalamaya daha yakındır,
- Bir veri grubunun sabit bir a sayısıyla çarpılması, bölünmesi ile oluşan yeni veri grubunun standart sapması ilk veri grubunun standart sapmasının a sayısıyla çarpımı ve bölünmesine eşit olur.
- Bir veri grubunun sabit bir a sayısıyla toplanması ve çıkarılması ile oluşan yeni veri grubunun standart sapması ilk veri grubunun standart sapmasına eşit olur.
- Bir X_i veri grubundan türetilen $Y_i = c_1+c_2X_i$ veri grubunun standart sapması $s_y^2 = c_2^2 * s_x^2$ olur.

İspat :

$$S_y^2=\frac{\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2}{n}=\frac{\sum_{i=1}^{n}\left(\left(c_1+c_2X_i\right)-\left(c_1+c_2\overline{X}\right)\right)^2}{n}=\frac{\sum_{i=1}^{n}\left(c_1+c_2X_i-c_1-c_2\overline{X}\right)^2}{n}=\frac{\sum_{i=1}^{n}c_2^2\left(X_i-\overline{X}\right)^2}{n}=c_2^2S_x^2$$

Bir veri grubu için standart sapma ve varyans,

$\sigma=\sqrt{\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{n}}$ teorik standart sapma, $S=\sqrt{\frac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{n}}$ standart sapma

$V(X)=\frac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{n}$ eşitlikleriyle hesaplanır. n < 30 ise $S=\sqrt{\frac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{n-1}}$ alınır.

Küme	**Aritmetik Ortalama**	**Standart Sapma**	**Varyans**
Ana Küme	μ	σ	σ^2
Örnek Küme	$\overline{X}$	S	S^2

1.4. Değişim Katsayısı (Relatif, Oransal Hata) : Ölçülmüş çeşitli veri gruplarının karşılaştırılabilmesi için kullanılan değişim katsayısı ölçütüdür ve şu şekilde hesaplanır.

$D = \frac{S}{\bar{X}} * 100$ örnek veri grubu için, $D = \frac{\sigma}{\mu} * 100$ ana veri grubu için

Bu şekilde bir hesaplamayla bilgiler oranlara dönüştürülür ve birim etkisinden kurtulur. İki veri grubu karşılaştırılırken ölçü aynı cins birimle yapılmış ortalamalar eşitse standart sapma, diğer durumlarda ise değişim katsayısı kullanılır.

2. Dağılımın Eğiklik Ölçütleri

Bir veri grubunun özelliklerini belirlerken merkezi eğilim (ortalamalar) ve dağılma (değişkenlik) ölçütlerinin yanında verilerin simetrik dağılımdan ne kadar uzaklaştığını gösteren ve dağılımın yüksekliğine ilişkin bilgileri veren eğiklik ve basıklık ölçütlerine de ihtiyaç duyulur.

2.1. Aritmetik ortalama, Medyan ve Mod Karşılaştırması: Aritmetik ortalama, Medyan ve mod karşılaştırılarak dağılımın eğikliği hakkında bilgi edinilebilir.

$\bar{X}$ =Medyan = Mod ise dağılım simetrik,

$\bar{X}$ > Medyan > Mod ise dağılım sağa eğik,

$\bar{X}$ < Medyan < Mod ise dağılım sola eğik olacaktır.

2.2. Pearson Eğiklik Katsayısı: Ayrıca eğikliğin derecesinin de belirlenmesi dağılım açısından önemli bir kriterdir. Mod ve aritmetik ortalama arasındaki kabaca fark eğikliğin derecesini gösterir. Eğikliğin fazla olmadığı veri grubunda aritmetik ortalama ile mod arasındaki fark aritmetik ortalama ile medyan arasındaki farkın yaklaşık 3 katına eşittir.

$\bar{X}$-Mod = 3 ($\bar{X}$-Medyan) olur. Bu eşitliğin her iki yanı standart sapma S'e bölünürse eğiklik ölçütü eşitliği birimsiz olur. Bu şekilde α_1 ve α_2 olarak gösterilen pearson eğiklik katsayıları elde edilmiş olur.

$$\frac{\bar{X}-\text{Mod}}{S} = 3\left(\frac{\bar{X}-\text{Medyan}}{S}\right) \Rightarrow \alpha_1 = 3\alpha_2$$

Bu durumda;

$\overline{X}$ =Medyan = Mod ise α_1=0, α_2=0 olur ve dağılım simetrik,

$\overline{X}$ > Medyan > Mod ise α_1>0, α_2>0 olur ve dağılım sağa eğik,

$\overline{X}$ < Medyan < Mod ise α_1<0, α_2<0 olur ve dağılım sola eğik olur.

2.3. Momentler ve Momentlerle Eğiklik Ölçütleri:

Bir veri grubunun dağılımının duyarlı olarak belirlenmesinde moment ölçütü kullanılır. Moment, verilerin sıfıra, aritmetik ortalamaya ya da herhangi bir a sayısına göre farklarının çeşitli derecelerden kuvvetlerinin aritmetik ortalaması alınarak hesaplanır ve istatistikte önemli bir yer tutar. Momentler bir veri grubu için,

$m = \frac{\sum_{i=1}^{n}(X_i - a)^r}{n}$ eşitliğiyle, frekanslı ve sınıflandırılmış veri grupları içinse,

$m = \dfrac{\sum_{i=1}^{n} f_i (X_i - a)^r}{\sum_{i=1}^{n} f_i}$ eşitliğiyle hesaplanır. Moment eşitliğinde kullanılan a sayısı herhangi bir sayı olabileceği gibi a=0 olduğu durumda sıfıra göre moment, a= $\overline{X}$ olduğu durumda ise aritmetik ortalamaya göre moment olarak adlandılır.

- Sıfıra Göre Moment çeşitli r derecelerine göre aşağıdaki gibi alınır.

	Basit Seri
1. derece	$m = \dfrac{\sum_{i=1}^{n} X_i}{n} = \overline{X}$
2. derece	$m = \dfrac{\sum_{i=1}^{n} X_i^2}{n} = K^2$
3. derece	$m = \dfrac{\sum_{i=1}^{n} X_i^3}{n}$
4. derece	$m = \dfrac{\sum_{i=1}^{n} X_i^4}{n}$

- Aritmetik ortalamaya göre moment çeşitli r derecelerine göre aşağıdaki gibi alınır.

	Basit Seri
1. derece	$m = \dfrac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)}{n} = 0$
2. derece	$m = \dfrac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{n} = S^2$
3. derece	$m = \dfrac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^3}{n}$
4. derece	$m = \dfrac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^4}{n}$

Aritmetik ortalamaya göre momentin ilk dört derecesi önemlidir. 1. derecede moment sıfıra, ikinci derecede momentse varyansa eşittir.3. moment eğiklik, 4. momentse basıklık için kullanılan kriterlerdir.

3. Basıklık Ölçütleri

3.1. Momente göre Basıklık Ölçütü: Basıklık bir dağılımın diklik ölçütüdür. Basıklık bir eğri altındaki toplam alanın ne kadarının ortalamadan ±1,2,3S (standart sapma) uzaklıkta yer aldığını gösterir. Normal dağılım eğrisi simetrik bir özellik taşır ve %99'u ortalamadan ±3S'lık aralık içinde yer alır. Basıklık ölçütü aritmetik ortalamaya göre 4. momentin, aritmetik ortalamaya göre 2. momentin karesine bölünmesiyle elde edilir ve α_4 ile gösterilir.

$\alpha_4 = \frac{m_4}{m_2^2}$ yazılır. Eşitliklerde $m_2 = S^2$ olduğu için eşitlik $\alpha_4 = \frac{m_4}{S^4}$ şekline gelir.

Eğer dağılım normal ise $\alpha_4 = 3$

dağılım dik ise $\alpha_3 > 3$

dağılım basık ise $\alpha_3 < 3$ olur.