

TEMEL İSTATİSTİK KAVRAMLARI

Günümüzde bir konu ile ilgili bilgi elde etmek için veri toplamak gerekir. İstatistik veriden öğrenme sanatıdır. İstatistik verilerin toplanması, daha sonra verilerin tanımlanması, verilerin analizi ve tahmin problemleri ile ilgilenir.

Verilerin Toplanması ve Tanımlayıcı İstatistik

İstatistiksel analiz bazen bir veri kümesi ile başlar. Örneğin düzenli olarak alınan yıllık yağış toplamları, ortaya çıkan depremler, işsizlik oranı, enflasyon verileri gibi veriler düzenli olarak toplanır ve kamuoyuna duyurulur. İstatistik bu verileri betimlemek, özetlemek, analiz etmek ve gelecek tahmini yapmak için kullanılır.

Bazı durumlarda da veri henüz mevcut değildir, böyle durumlarda veri üretmek için istatistik uygun bir deney hazırlamada kullanılabilir. Seçilen deney veriden yararlanmak isteyen kişinin kullanımına bağlı olmalıdır. Örneğin, bir öğretim görevlisinin bilgisayar programlama öğrenmeye yeni başlayanlara iki farklı öğretim yönteminden hangisinin daha etkili olduğu araştırılmak istenebilir.

Çıkarımsal İstatistik ve Olasılık Modelleri

İstatistiğin sonuç çıkarma ile ilgilenen bu kısmına çıkarımsal istatistik denir. Veriler yardımıyla bir sonuç çıkarabilmek için

MERKEZİ EĞİLİM ÖLÇÜLERİ

Aritmetik Ortalama

En çok kullanılan merkezi eğilim ölçüsüdür. Örneklem değerlerinin toplanarak örneklem sayısına bölünmesiyle, yani

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

formülü ile elde edilir. Eğer kitle değerleri kullanılarak aritmetik ortalama hesaplanırsa, kitleye ilişkin aritmetik ortalama μ ile gösterilir.

Mod (Tepe Değer)

Bir veri setinde en çok tekrarlanan değere mod (tepe değer) denir.

Medyan (Ortanca)

Bir veri setinde değerler küçükten büyüğe sıralandığında ortaya düşen değer medyan değeridir. Veri sayısının tek yada çift olmasına göre medyan değerinin hesaplanması değişir. Medyan değeri hesaplanırken ilk olarak veriler küçükten büyüğe doğru sıralanır. n veri sayısını göstermek üzere medyan

$$Medyan = \begin{cases} x_j & , j = \frac{n+1}{2}, n \text{ tek ise} \\ \frac{x_j + x_{j+1}}{2}, j = \frac{n}{2}, n \text{ çift ise} \end{cases}$$

formülü ile hesaplanır.

Örnek: Koronavirüs hastalığı nedeniyle yoğun bakım servisinde yatan hastalar arasından rastgele 10 hasta seçilmiş olup, yoğun bakımda kalma süreleri aşağıdaki gibi verilmektedir:

$$x_i: 9 \ 8 \ 5 \ 5 \ 10 \ 11 \ 9 \ 8 \ 7 \ 9$$

- Hastaların yoğun bakımda kalma sürelerinin aritmetik ortalamasını bulunuz.
- Hastaların yoğun bakımda kalma sürelerinin mod değerini bulunuz.
- Hastaların yoğun bakımda kalma sürelerinin medyan değerini bulunuz.