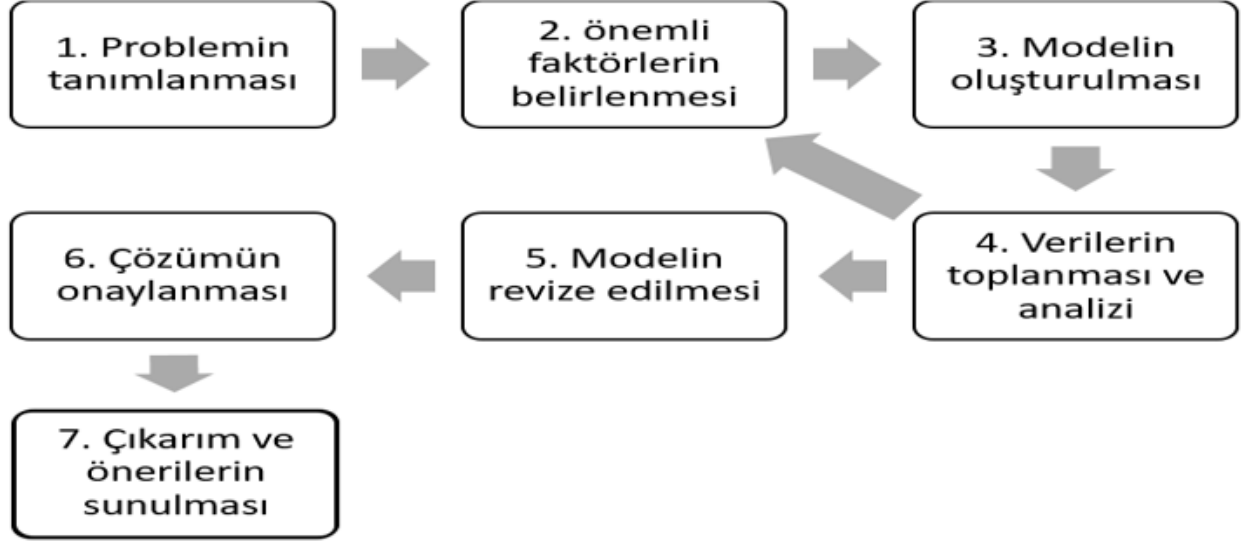


# OLASILIK VE İSTATİSTİK DERS NOTU

## 1. TEMEL İSTATİSTİKSEL KAVRAMLAR

Problem çözümünde istatistiğin rolü:



### 1.1. Tanımlamalar:

**İstatistik bilimi;** verilerin toplanması, düzenlenmesi, özetlenmesi, takdimi, analizi ve bu analizler vasıtasıyla elde edilen sonuçların yorumlanması ve bir karara bağlanması ile ilgilenir. Buna göre istatistiği ikiye ayırabiliriz: Betimsel istatistik ve çıkarımsal istatistik.

Frekans dağılımları, merkezi ölçüleri (aritmetik ortalama, geometrik ortalama, mod, medyan, ...), dağılış ölçüleri (standart sapma, varyans, değişim aralığı, ...), çarpıklık ve basıklık gibi konular verilerin özetlenmesi ve betimlenmesi ile ilgili olup betimsel istatistiğin konusudur. Örneklem teorisi, tahmin, hipotez testleri, ilişki katsayıları ve regresyon analizi gibi konular ise çıkarımsal istatistiğin konusudur.

**Kitle:** Üzerinde ölçüm ve araştırmaların yapılacağı, belirli bir ya da birkaç özelliği incelenecek olan birimler topluluğudur. Örneğin; Denizli'deki üniversite öğrencilerinin aylık giderleri için kitle PAÜ öğrencileridir.

**Tamsayım:** Kitle hakkında tüm bilgilerin elde edilmesi kitleyi oluşturan tüm birimlerin incelenmesi demektir. Kitlenin tüm birimlerinin incelenmesi işlemine tam sayım denir.

**Örneklem:** Kitleye ulaşmanın mümkün olmadığı ya da yüksek maliyetler gerektirdiği durumlarda kitleden belirli yöntemlerle seçilen gözlemlerin oluşturduğu kümeye denir.

**Örnekleme:** Kitleden, kitle birim sayısından daha az sayıda birimin seçilmesi ile kütleye tahmin yapma işlemlerine denir.

**Gözlem:** Belirli bir özellik hakkında bilgi ve veri toplamak için kayıt altına alınan her bir ölçüme denir.

**Değişken:** Gözlemden gözleme değişik değerler alabilen objelere, özelliklere ya da durumlara denir. Örneğin; cinsiyet, eğitim düzeyi, boy, kilo, aylık harcama.

**Parametre ve Örneklem İstatistiği:** Kitlenin tanımlayıcı sayısal değerlerine parametre, örneklemin tanımlayıcı sayısal değerlerine de örneklem istatistiği denir. İstatistiksel çıkarımda amaç örneklem istatistiği kullanılarak kitle parametreleri hakkında çıkarımda bulunmaktır.

	Parametre	Örneklem istatistiği
Ortalama	$\mu$	$\bar{x}$
Varyans	$\sigma^2$	$s^2$
Standart sapma	$\sigma$	$s$
Oran	$\Pi$	$p$
Gözlem sayısı	$N$	$n$

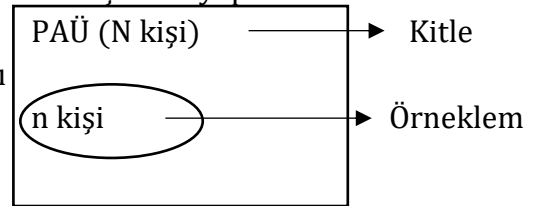
**Örnek:** PAÜ'deki öğrencilerin aylık harcamaları hakkında araştırma yapılsın.

Kitle: PAÜ'de kayıtlı öğrenci sayısı  $N=44\ 733$

Parametre ( $\mu$ ): öğrencilerin ortalama aylık harcaması

Örneklem: Rasgele seçilen  $n=250$  öğrenci

Örneklem istatistiği ( $\bar{x}$ ): örneklemin ortalaması



## 1.2. Verilerin Toplanması

- **Geçmişe Dönük Araştırma:** Geçmişteki bir zaman diliminde toplamış verilerin bir kısmının ya da tamamının kullanılması durumudur. Maliyet ve zaman tasarrufu sağlamaktadır. Fakat verinin geçerliliği ve güvenilirliği şüphelidir. Zaman içerisinde gözlem kaybı yaşanmış olabilir. Veri toplanırken hatalar yapılmış olabilir. Tam amacımıza hitap etmiyor olabilir. Örnek: Eski laboratuvar çalışmasında elde edilen veriler.
- **Gözleme Dayalı Araştırma:** Bir süreç veya kitlenin belirli bir periyotta gerçekleştirilen rutin işlemler sırasında verilerin gözlemlenip kayıt altına alınmasıdır. Örnek: Yeni kurulan elektrik santralinin bölgedeki sıcaklık değişimine etkisi.

- **Deney Tasarımı:** Sistemde yer alan kontrol edilebilir değişkenler (faktörler) üzerinde bilinçli değişiklikler yapılarak sistemin bu değişikliklere verdiği tepkiyi gözlemleyerek değişkenlerin çıktı performansının etkilerini belirleyebilmesidir. Örnek: Araç yakıt tüketimini etkileyen faktörler örneği.

### 1.3. Bazı Olasılığa Dayalı Örneklem Yöntemleri:

- **Basit rasgele örneklem:** Bir kitleden her bir elemanın seçilme şansının eşit olduğu örneklem yöntemidir. Örneğin, PAÜ aylık ortalama harcama örneği gibi rasgele 250 öğrenci seçilmesi.
- **Sistematiik örneklem:** Örneklemi oluşturacak ilk eleman kitleden rasgele seçilir, diğer elemanların ise belirli ve aynı kural çerçevesinde seçimi ile belirlenen yöntemdir. Genel olarak (örneklem aralığı=Kitle büyüklüğü/örneklem büyüklüğü) formülü ile rasgele seçilen ilk elemana örneklem aralığı eklenerek oluşturulur. Örneğin sınıftaki 80 öğrenciden 10 kişilik örneklem seçelim. Örneklem aralığı=(80/10) dan 8 olup sınıf listesinden rasgele 2 öğrenci ilk eleman olarak seçilsin. O halde örneklemi 2, 10, 18, 26, 34, 42, 50, 58, 66 ve 74 sıra numarasına sahip öğrenciler oluşturacaktır.
- **Tabakalı Örneklem:** Öncelikle kitle belirli bir çerçevede gruplandırılır ve her bir gruptan rasgele örneklem çekilerek birleştirilmesiyle elde edilen yöntemdir. Burada önemli olan gruplar kendi içinde homojen, gruplar arası ise heterojen olmalıdır. Örneğin, PAÜ deki öğrencilerin aylık ortalama harcamaları için kitleyi gruplandıralım. Tıp, Mühendislik, Fen Edebiyat, İktisadi ve İdari Bilimler, Mimarlık ve Tasarım gibi fakültelerden yoğunluklarına göre her birinden rasgele öğrenciler seçip örneklem oluşturmak.
- **Küme örneklem:** Kitleyi ayrı ayrı temsil edebilen kendi içerisinde heterojen, aralarında homojen yapıya sahip kümelerin seçimine bağlı yöntemdir. Örneğin, sanayi bölgesinde tekstil fabrikalarında çalışan işçilerin çalışma koşulları üzerine araştırma yapılmak istensin. Aynı koşullara sahip her bir fabrika birer küme olmuş olup rasgele fabrikalar seçilerek örneklemi oluşturma yöntemidir.

#### 1.4. Değişken ve değişkenlerin ölçülme düzeyleri

**Değişken:** Birimlerin farklı değerler alabildikleri nitelik veya niceliklerdir. Örneğin müşterilerin oluşturduğu bir kitle de; meslek, eğitim düzeyi, cinsiyet, yaş, aylık gelir gibi.

- **Nitel (Kategorik) değişkenler:** Gözlemlerin sınıflara ayrılarak açıklandığı ölçümlerdir. Cinsiyet, eğitim durumu, Kan grubu gibi.
- **Nicel (Sayısal) değişkenler:** Sayısal değerlerden oluşan ölçümlerdir. Boy, kilo, yaş, nüfus gibi. Sayısal değişkenler *kesikli* ve *sürekli* olmak üzere ikiye ayrılır. *Kesikli değişkenler*, belirli aralıkta tam değerler alabilen değişkenlerdir. Sınıftaki öğrenci sayısı, otoparktaki araba sayısı gibi. Sürekli değişkenler ise belirli aralıkta tüm değerleri alır. Ağırlık, uzunluk gibi.

Değişkenlerin ölçümleri genel olarak dört farklı şekilde yapılabilir.

- **Sınıflama Ölçeği:** Gözlemlerin sadece sınıflanması veya etiketlenmesiyle sınırlıdır. Sınıflar arasında büyüklük ya da sıra açısından bir ilişki bulunmaz. Kategorik değişkenlerdir. Örneğin, Cinsiyet, kan grubu, meslek grubu, Gözlük kullanımı (0: Gözlüksüz, 1: Gözlüklü) gibi.
- **Sıralama Ölçeği:** Gözlemlerin belirli bir kritere göre sıralayarak etiketlenmesidir. Değişken değerleri bakımından önem sırası bulunur. Kategorik değişkenlerdir. Örneğin, Eğitim durumu (ilkokul, ilköğretim, lise, üniversite), Üniversite öğretim üyesi unvanları ve askeri rütbeler, ekonomik durum (1: çok kötü, 2: kötü, 3: orta, 4: İyi, 5: çok iyi ).
- **Eşit aralıklı Ölçek:** Bu ölçekte üzerinde durulan değişken belirli iki değer arasında sonsuz değer alabilir. Bu ölçekteki 0 değeri, ölçülen karakteristiğin olmadığını göstermez. Aynı şekilde ölçüm karakteristiklerinden biri diğerinin katlarıyla ifade edilemez. Sayısal değişkenlerdir. Örneğin, Sıcaklık ölçümleri. 0 derece, sıcaklığın olmadığını göstermez ve 4 derece, 2 derecenin iki katı değildir.
- **Oran Ölçeği:** Zayıftan kuvvetliye doğru sıraladığımızda yukarıdaki ölçeklerin en hassas olanıdır. Ölçülen karakteristiğin 0 olması o karakteristiğin olmadığını gösterir. Aynı şekilde ölçülen bir karakteristik diğerinin katları ile ifade edilebilir. Sayısal değişkenlerdir. Örneğin; ağırlık, uzunluk, hız, gelir gibi. Sıfır ağırlığa sahip olmak yokluğu ifade ettiği gibi 50 kg mın iki katı 100 kg ifade eder. Tüm matematiksel işlemler yapılabilir.