



İSTATİSTİK

1. Giriş

Bilimsel Araştırmaların Amacı

1. Bir nesne veya olayı tanımlamak (Tanımlayıcı)
 - Mevcut durum nedir?
 - Tanımlama, sınıflandırma, kategorize etme
2. Çeşitli olaylar arasında ilişkiler bulmak (İlişkisel)
 - Bir veya daha fazla değişken arasında ilişki var mı yok mu? Varsa, ilişkinin derecesi ve yönü nedir?
3. Olaylar arasında neden sonuç ilişkilerini tespit etmek, böylece gelecekle ilgili tahminde bulunmak ve/veya bir takım değişiklikler yapabilmek (Deneyisel)

Bilimsel araştırmanın aşamaları

1. Problem ortaya koyma
2. Literatür taraması
3. Hipotez geliştirme
4. Veri toplama
5. Analizler
6. Yorumlama
7. Yazma

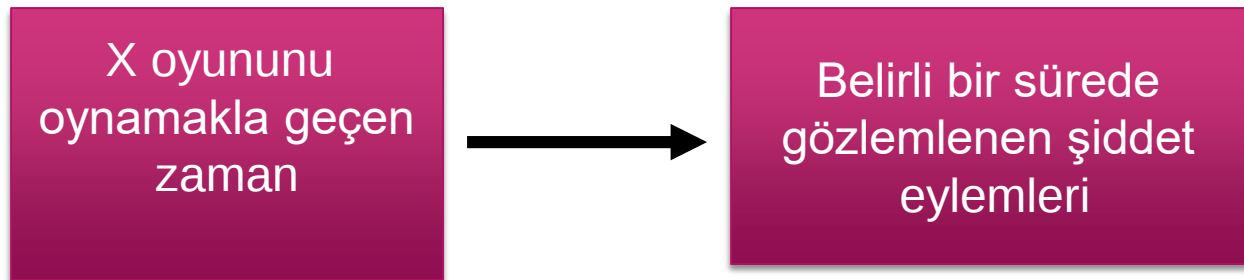
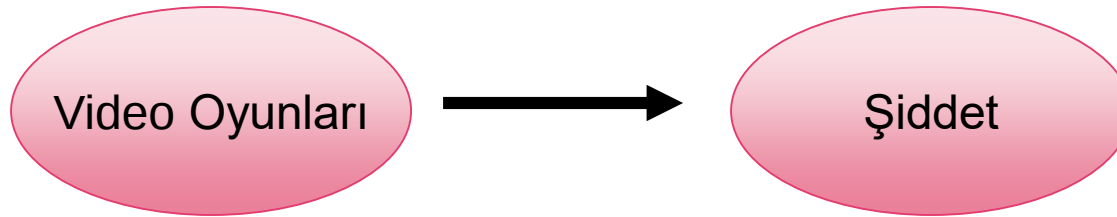
Hipotezler

- ▶ Araştırma problemi için önerdiğiniz cevaptır. Hipotezler araştırma sorusu için tahminde bulunur. Araştırma sonuçlarına göre bu tahminler ya desteklenir yada reddedilir.

Araştırma Problemi ve Hipotezler

- ▶ İyi bir araştırma problemi, bir yada daha fazla ölçülebilir hipotez üretir.
- ▶ Hipotezler bir tahmin ifadesidirler. Bazı araştırmalar keşifsel olup başlangıçta bir hipotez içermeyebilirler.
- ▶ Hipotezler farklı yapıları/değişkenleri belirli ilişkilerle birbirine bağlar
- ▶ Hipotez ifadeleri iki yada daha fazla ölçülebilen (en azından potansiyel olarak) değişken içerir ve bu değişkenlerin birbiriyle nasıl bir ilişki içinde olduğunu belirtir

Hipotezler



Hipotezler

- ▶ H_0 : Sıfır Hipotezi : İlişki yok
- ▶ H_a : Alternatif Hipotez: İlişki var
- ▶ Tek kuyruklu: İlişki yönü belirtiliyor
- ▶ Çift kuyruklu: İlişki yönü belirtilmiyor
 - ▶ $H_0: \mu_1 = \mu_c$
 - ▶ $H_A: \mu_1 < \mu_c$
 - ▶ $H_A: \mu_1 > \mu_c$
 - ▶ $H_A: \mu_1 \neq \mu_c$

Analizler

- ▶ Genellikle istatistiki teknikler kullanılarak yapılır.
- ▶ Uygulanacak analiz türü veriye ve araştırma problemine bağlıdır.
- ▶ Analiz sonucunda yapılabilecek hata - yanılma olasılığı olabildiğince düşük olmalıdır (genellikle %5 veya %1)
- ▶ I. Ve II. tip hata

Hipotez testlerinde hata türleri

Gerçek Durum	Araştırma Sonucu	
	H0 doğru (fark yok)	H0 yanlış (fark var)
H0 doğru (fark yok)	Doğru	I. Tip hata
H0 yanlış (fark var)	II. Tip hata	Doğru

I. Tip Hata

- ▶ Ana kütlede H_0 geçerli olduğu halde, örnekleme dayalı analiz sonucunda H_0 reddedilirse, I. Tip hata yapılmış olur.
- ▶ I. tip hata yapma olasılığı α (alfa) ile gösterilir ve genellikle α değeri 0,05 veya 0,01 olarak belirlenir.
- ▶ α değeri anlam düzeyi (significance level) olarak adlandırılır. $1 - \alpha$ değeri ise güven düzeyi (confidence level) olarak adlandırılır.

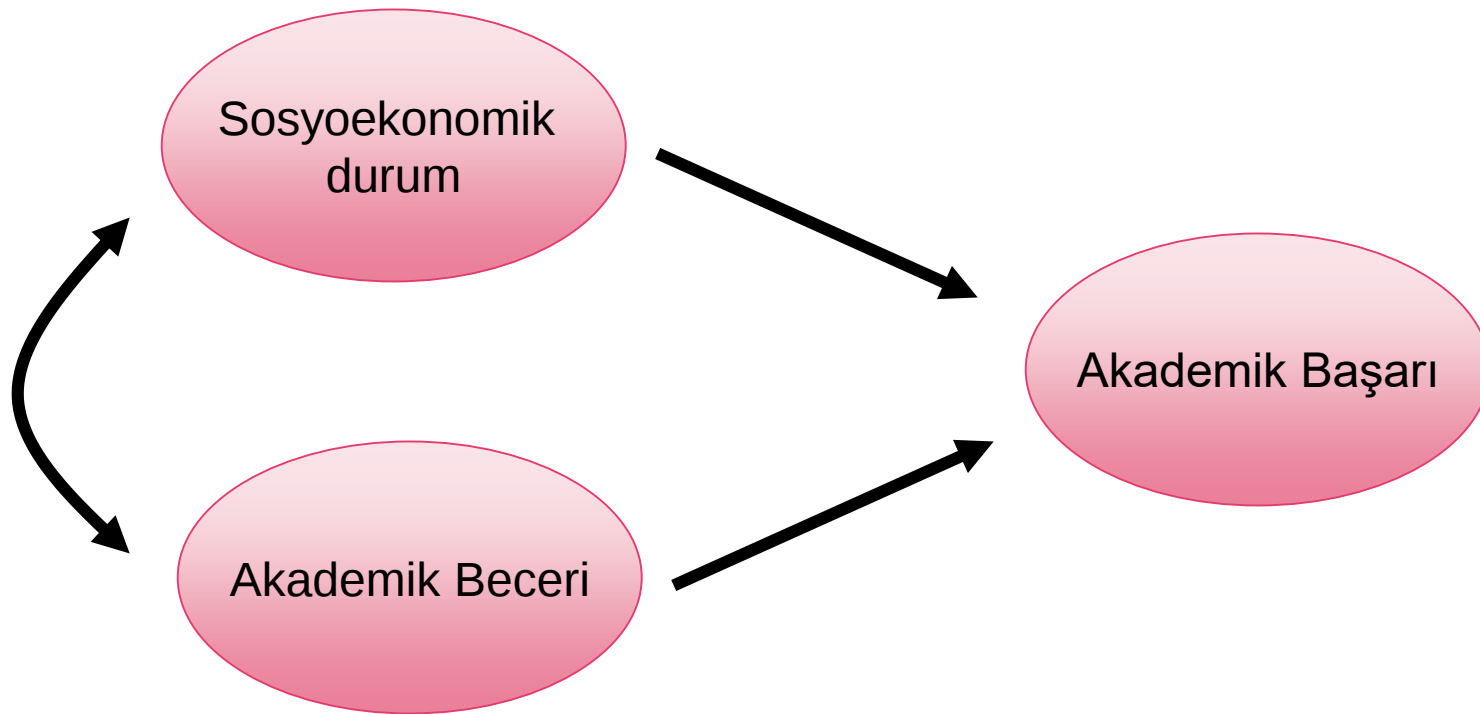
II. Tip Hata

- Ana kütlede H_0 geçerli olmadığı halde, örnekleme dayalı analiz sonucunda H_0 kabul edilirse, II. Tip hata yapılmış olur.
- II. tip hata yapma olasılığı, β (beta) ile gösterilir.
- $1 - \beta$, gerçekte bir etki varken bunu bulma olasılığı olarak tanımlanır ve **testin gücü** olarak adlandırılır.
- İstatistik testlerinde I. ve II. tip hata yapma olasılıklarının her ikisini de olabildiğince küçük tutmak istenir ancak bu ikisi arasında ters bir ilişki olduğundan örneğin birinci tip hata yapma olasılığını düşürmek ikinci tip hata yapma olasılığını artırır

İstatistik Testlerinde Güç Kavramı

- İstatistiksel bir testin gücünü etkileyen faktörler:
 1. Etki büyüklüğü: Eğer ana kütlede var olan ilişkinin derecesi büyükse o zaman bu etkiyi örneklemden de bulmak kolay olacaktır yani testin gücünü artıracaktır.
 2. α : α ne kadar küçük alınırsa, güç o kadar azalacaktır.
 3. Örneklem büyüklüğü: Örneklem büyüklüğü arttıkça testin gücü de artacaktır. Düşük hacimli örneklemelerde test duyarsız (ilişkileri bulamaz) olurken, büyük hacimli örneklemelerde ise aşırı duyarlı (en küçük ilişkileri bile bulur) olacaktır.

Araştırma modeli – teorik model

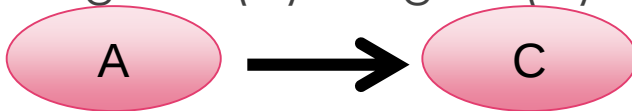


Değişken kavramı

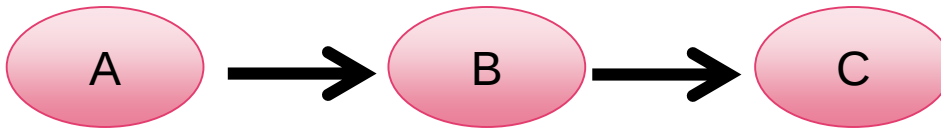
- ▶ Değişken farklı değerler alabilen birimlerdir.Ör: yaş, boy, kilo
- ▶ Sayısal olmayabilir. Örneğin, cinsiyet. Ancak çoğu zaman analiz yapılabilmesi için sayısallaştırılmaları gerekir.
- ▶ Bağımsız değişken (Independent Variable): manipüle edip üzerinde değişiklik yaptığınız değişken
- ▶ Bağımlı değişken (Dependent Variable): Bağımsız değişkene bağlı olarak değişen değişken
- ▶ Aracı Değişken (Mediating Variable): Bağımsız değişkenin etkisini bağımlı değişkene aktaran değişken
- ▶ İlimlaştırmacı Değişken (Moderating Variable): Bağımlı ve bağımsız değişken arasındaki ilişkiyi etkileyen değişken

Değişken Türleri

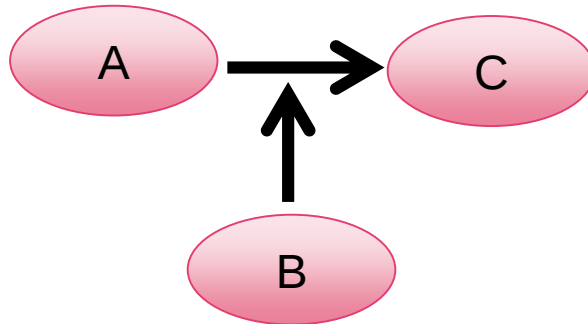
- Bağımsız (A) –Bağımlı (C) değişkenler



- Bağımsız (A), Aracı (B) ve Bağımlı (C) değişkenler



- Bağımsız(A), İlimlaştırmacı (B) ve Bağımlı (C) değişkenler



Veri Türleri (Veri Ölçüm Düzeyleri)

- ▶ Değişkenler 4 farklı ölçüm seviyesinde ölçülebilirler
- ▶ Veri ölçüm düzeyleri, veriler üzerinde yapabileceğimiz istatistiksel yöntemler konusunda bir kısıt oluşturur. Bir başka ifadeyle hangi istatistiksel yöntemi uygulayacağımız, verinin türüne bağlı olarak değişir.

1. Nominal - Kategorik

- ▶ Veriler bir yada daha fazla kategoriye göre sınıflandırılarak ölçüldükleri zaman, veri **nominal** seviyede ölçülmüş demektir.
- ▶ Ör: Renk, Cinsiyet, Dini ve politik tercihler, vs.

2. Sıralı (Ordinal)

- ▶ Ölçümün farklı kategorileri arasında bir sıra (öncelik) durumu var ancak sıralar arasındaki mesafe eşit veya ölçülebilir değilse, ölçüm seviyesi **sıralıdır**.
- ▶ Ör: Eğitim durumu, askeri rütbe, elbise büyüklüğü, vs.
- ▶ Kategoriler arasında bir sıra farkı vardır ancak bu farkın büyüklüğü belirli değildir.

3. Aralık (Interval)

- ▶ Ölçüm değerleri arasında büyüklük olarak ölçülen bir farklılık var ancak mutlak bir sıfır noktası yoksa, ölçüm seviyesi **aralık** seviyesindedir.
- ▶ Hava sıcaklığı en iyi örnektir.

4. Oran (Ratio)

- ▶ Veriler arasındaki fark sayısal olarak belirli ve mutlak bir sıfır noktası var ise, ölçüm seviyesi **oran** olarak adlandırılır.
- ▶ Ör: Ağırlık, yaş, boy, vs.

İstatistik Türleri

- ▶ **Tanımlayıcı (descriptive) İstatistik**
 - ▶ Verileri toplama, sunma ve açıklama
- ▶ **Çıkarımsal (inferential) İstatistik**
 - ▶ Örneklem verilerine dayalı olarak ana kütle hakkında çıkarımlar yapma veya karar verme

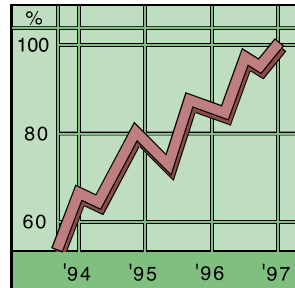
Tanımlayıcı İstatistik

- Veri kümesini anlamaya vve özetlemeye yönelik istatistikler

- Sayılarla

- Ör: Ortalama =
$$\frac{\sum x_i}{n}$$

- Grafiklerle



Çıkarımsal İstatistik

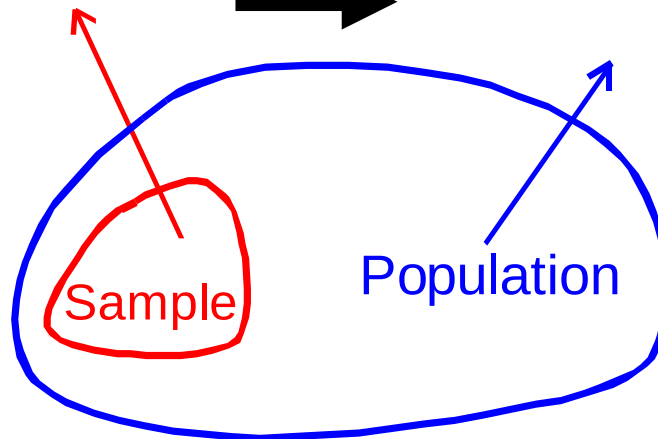
- Örnekleme (sample) bakarak ana kütle (population) hakkında yorumlarda bulunmak

Örneklem istatistikleri
(biliniyor)

Çıkarım



Ana kütle parametreleri
(bilinmiyor, ancak
örnekleme bakılarak
tahmin edilebilir)



Örneklem Alma

- Bilimsel araştırmalarda çoğu zaman bir örneklem üzerinden toplanan veriler kullanılarak analizler yapar ve çıkarımlarda bulunuruz. Araştırma sonuçlarının geçerliliği bu örneklemin ana kütleyi temsil gücüne bağlıdır. Bu nedenle örneklemin ana kütleyi temsil etmesi oldukça önemlidir.
- Bir örnekleme dayalı olarak yapılan çıkarımların kalitesi hem örnekleme yöntemi hem de örneklemin büyüklüğü ile ilişkilidir. Dolayısıyla uygun yöntem ve büyüklükte örneklem oluşturmak oldukça önemlidir.

Örnekleme Alma...

- ▶ **Ana kütle** araştırmanın ilgilendiği kişi, nesne veya olayların bütününe verilen isimdir. Eğer bir çalışmada ana kütlelerin tümünden veri toplanırsa buna **tamsayım** denir.
- ▶ **Örnekleme** bir ana kütlede seçilen elemanların oluşturduğu kümedir.
- ▶ Örnekleme almamıza neden olan şey ana kütlelerin tümü hakkında veri toplamamın gerektirdiği çaba, zaman ve maliyetin yüksek oluşu veya bazen imkansız oluşudur.
- ▶ Örnekleme almanın amacı ise yukarıdaki sayılan nedenlerle tüm verileri ölçülemeyen ana kütle hakkında ondan seçilen örneklemden edinilen bilgilere dayalı olarak çıkarımlarda bulunmaktır.

Örneklem Alma...

- **Ana kütle parametresi:** Ana kütlenin herhangi bir özelliği. Yunan harfleri ve büyük harfler ile gösterilir (N, μ)
- **Örneklem istatistiği:** Örneklemün ölçülen bir özelliği. Latin harfleri ve küçük harfler ile gösterilir (n, s)
- **Örneklem Hatası:** Genellikle araştırmalarda tamsayım yapıp ana kütle parametresi hesaplanmaz. Ancak ana kütleden seçilen bir örneklemün istatistiği hesaplanır ve bu istatistiğe dayalı olarak ana kütle parametresi tahmin edilmeye çalışılır. Tamsayım yerine örneklem alındığı için oluşan hata örneklem hatası olarak adlandırılır.

Örneklem alma yöntemleri

- Örneklem alma yöntemleri tesadüfi olan ve olmayan olmak üzere iki ana grupta toplanabilir. Tesadüfi olan yöntemlerde bütün birimlerin örnekleme girme şansı varken tesadüfi olmayan yöntemlerde bu durum söz konusu değildir. Aşağıda her iki gruba giren temel yöntemler verilmiştir

1. Tesadüfi (rassal, random) örnekleme yöntemleri
 - a. Basit rassal
 - b. Sistematik
 - c. Tabakalı
 - d. Küme
2. Tesadüfi olmayan (non-random) örnekleme yöntemleri
 - a. Kolayda
 - b. Yargısal
 - c. Kota
 - d. Kartopu