

# Araştırma Yaparken İstatistiksel Yöntemlerin (Doğru) Kullanılması

Öğr.Gör. M.Sinan İyisoy

Tıp Eğitimi ve Bilişimi AD

*siyisoy@gmail.com*

# İstatistik Nedir?

İstatistiği basitçe, eldeki örneğe bakarak örneğin geldiği popülasyon hakkında konuşma sanatı olarak açıklayabiliriz.



# Neden İstatistik?

Elimizde kısıtlı sayıda laboratuvar imkanları, finansal ve insani kaynaklar vardır. İlgilendiğimiz popülasyondaki tüm bireyleri deneyimize dahil etmek mümkün değildir.

Bu yüzden popülasyondan seçilmiş belli sayıdaki kişi üzerinde çalışır, ölçümlemimizi onlar üzerinde yaparız.

Buradan elde ettiğimiz sonuçları tüm popülasyona genellemek isteriz. İşte bunu yapmamız istatistik sayesinde olur.

# İstatistik yapınca neler elde ederiz?

Temelde istatistik 4 farklı amaç için kullanılır.

- Tanımlamak (description)
- Çıkarımda bulunmak (inference)
- Kestirmek (estimation)
- Tahmin etmek (prediction)

# Tanımlamak (description)

300 hastanın kolesterol değerlerini ölçtüğümüzü düşünelim.

Bu bilgileri çalışma sonunda raporlamak için nasıl bir yöntem kullanmalıyız?

Bütün değerleri alt alta yazarak listeleyelim mi? Yoksa bir grafik mi çizmeliyiz?

İstatistik bize nasıl yapılması gerektiğini söyler.

# Çıkarımda bulunmak (inference)

Hiperkolesterolemi için yeni bir ilaç geliştirdiğimizi düşünelim.  
Halihazırda kullanılan standart ilaçtan daha etkili olduğunu göstermek için ne yapmalıyız?  
Deney tasarımı nasıl olmalıdır? Kaç kişiyle çalışılmalıdır?  
İstatistik bize nasıl yapılması gerektiğini söyler.

# Kestirmek (estimation)

Belirli bir popülasyonda kolesterolün normal sınırlarını bulmak istediğimizi düşünelim.

Nasıl bir yol izlemeliyiz?

Ya da, yeni bulduğumuz kolesterol ilacı standart ilaçtan **ne kadar** iyidir?  
İstatistik bize nasıl yapılması gerektiğini söyler.

# Tahmin etmek (prediction)

Belirli bir popülasyondaki kişilerin bazı temel özelliklerini (yaş,cinsiyet,aile geçmişi gibi) biliyoruz.

Bu özelliklere bakarak bu bireylerin gelecekte kolesterol hastası olma ihtimalini bulmak istediğimizi düşünelim.

Ne yapmalıyız?

İstatistik bize nasıl yapılması gerektiğini söyler.



Bir çalışma **planlarken** bazı istatistiksel süreçlerden geçmek lazımdır. Bunların nasıl yapılacağını **çalışmaya başlamadan** bir istatistikçi ile görüşmek iyi bir tercihtir. Sadece çalışma bittikten sonra istatistikçiye gitmek, ceset üzerinde ölüm sonrası araştırma yapmaya benzetilmiştir. Şimdi bu adımlardan bahsedelim.

# Bir çalışma planlayalım: Hipotez Belirlemek

Yeni bir ilacın hiperkolesterelomiye önlediğini iddia ettiğimizi varsayalım.

Çalışmaya (deneye) başlarken bir hipotezimiz olmalıdır.

"Yeni ilacımız standart ilaçtan daha iyidir" şeklindeki bir ifade uygun bir hipotez değildir.

Hipotezimiz istatistiksel terimler kullanılarak ifade edilebilmelidir.

Örneğin, yeni ilacı kullanan hastaların **ortalama** kolesterol değerleri, standart ilacı kullanan benzer hastaların **ortalama** kolesterol değerlerinden **20 birim** daha küçüktür.

# Bir çalışma planlayalım: Deney Tasarımını Belirlemek

Deney tasarımı önemli bir konudur. Belirlediğimiz hipotezi test edebileceğimiz uygun bir deney tasarım şeklini bulmamız gereklidir. Kolesterol çalışmamız için 2 grup hastaya ihtiyacımız olduğu açıktır.

1. gruptaki hastalar standart ilacı kullanan kontrol grubu hastaları,
2. gruptaki hastalar ise yeni ilacımızı kullanan hastalar olmalıdır.

Ayrıca her ilaç anlık olarak etki göstermeyeceği için belli bir süre kullanmak ve başlangıç kolesterol değeri ile son kolesterol değeri arasındaki değişimi de dikkate almak gerekecektir.

İşte bu gibi noktaları göz önüne alarak uygun bir deney tasarım şekli belirlemek ilk aşamada yapılacak şeylerden biridir.

# Bir çalışma planlayalım: İstatistiksel Testi Belirlemek

Her hipotezi test için farklı bir istatistiksel yöntem vardır.

Deney tasarımı belirlendikten sonra bir bakıma hangi istatistiksel testin kullanılacağı da belirlenmiş olur.

Kolesterol çalışmamızda 2 grubumuz, her grubun 2 ölçümü (başlangıç ve son) olduğu ve biz bunların ortalamalarını karşılaştıracığımız için **Mixed ANOVA** adlı istatistiksel testi (yöntemi) kullanarak başta belirlediğimiz hipotezimiz test edebiliriz.

# Bir çalışma planlayalım: Karıştırıcı Değişkenleri Düşünmek

Deneyimize dahil edeceğimiz kişilerin bazı özellikleri elde edeceğimiz sonuçların doğruluğunu etkileyebilir.

Standart ilacı alan gruptaki hastalar daha yaşlı hastalar olabilir. Daha çok sigara içen hastalar olabilir.

Gruplar arasındaki bu farklılık ve hiperkolesterolemiye etkisi olan sigara içme gibi faktörleri göz önüne almadan yapacağımız çalışma bizi doğru olmayan sonuçlara götürecektir. Bu faktörlerin etkisini nasıl engelleyeceğimizi çalışmaya başlamadan düşünmemiz gerekir.

# Bir çalışma planlayalım: Örnek Büyüklüğünü Belirlemek

Deneyimizin temel özelliklerinden bir tanesi de kaç hastayı deneye alacağımızdır.

Bunun için daha önce hipotezimizde belirttiğimiz yeni ilacı kullanan hastalardaki 20 birimlik düşme bilgisini kullanacağız.

Kullanacağımız istatistik yöntemi (Mixed ANOVA) ve yeni ilacın etki büyüklüğünü (20 birim) kullanarak her 2 grubumuza kaçar hasta alacağımızı belirleyebiliriz.

# Bir çalışma planlayalım: Örneklem Yöntemini Belirlemek

Deneyimizin **dahil etme** ve **dışlama** kriterlerini belirlediğimizi düşünelim. Deneyimize komorbid hastalığı olmayan 18-65 yaş arası hastalarda yapmak istediğimizi varsayalım.

Deneye alacağımız kişileri popülasyondan nasıl seçmemiz gereklidir?  
Deneydeki bir kişiyi hangi ilaç grubuna alacağımızı nasıl belirleyeceğiz?  
Deney gruplarındaki hastaların özelliklerinin mümkün olduğunca aynı olmasını nasıl sağlayacağız?

Çalışma sırasında bazı kişiler deneyden çıkmak isteyecek, örneğin ilacı kullanmayı bırakacak, ya da ölçümlere gelmeyecek. Bu durumların üstesinden nasıl geleceğiz?

# Bir çalışma planlayalım: Sonuçları Raporlama

Deneyimizi bitirdiğimizi ve her hastadan gerekli ölçümleri aldığımızı düşünelim.

Alınan ölçümleri uygun bir şekilde bilgisayar ortamına nasıl girmek gereklidir?

Elimizdeki verinin türü nedir? Bu türe hangi istatistiksel yöntemler uygundur?

Bulunan sonuçlar nasıl değerlendirilecek ve raporlanacak? Hangi grafikler kullanılacak?

Tablolar nasıl yapılacak?



# p Değerinin "Anlamı"

## İstatistiksel

analiz sonucunda bazı p değerleri bulduk ve bunlar 0.05'ten küçük çıktı.

Artık yeni ilacımızın standart ilaçtan daha iyi olduğunu söyleyebilir miyiz?

Acaba beklediğimiz etki büyüklüğü olan kolesterolde en az 20 birimlik bir düşme yeni ilacı kullananlarda görüldü mü? Yoksa

bulduğumuz istatistiksel anlamlılık **klินิก olarak anlamlı bir farkı** göstermiyor mu?

<u>P-VALUE</u>	<u>INTERPRETATION</u>
0.001	HIGHLY SIGNIFICANT
0.01	
0.02	
0.03	
0.04	SIGNIFICANT
0.049	
0.050	OH CRAP. REDO CALCULATIONS.
0.051	ON THE EDGE OF SIGNIFICANCE
0.06	
0.07	HIGHLY SUGGESTIVE, SIGNIFICANT AT THE $P < 0.10$ LEVEL
0.08	
0.09	
0.099	HEY, LOOK AT THIS INTERESTING SUBGROUP ANALYSIS
$\geq 0.1$	

# Teşekkürler