

Normalite Testleri

Normalite Testleri

Biyoistatistiksel yöntemlerin temel amaçlarından birisi de n sayıdaki örneklerden elde edilmiş istatistikleri kullanarak topluma ilişkin parametre tahminleri yapmak, bu parametrelere ilişkin kurulan hipotezleri (varsayımları) test ederek doğru kararlara ulaşmaktır.

Normalite Testleri

Toplumda bir değişkenin parametrelerine, dağılım yapısına ya da ilişki düzeyine ilişkin kurulan hipotezlerin denetlenmesi için yararlanılan yöntemlere **Hipotez Testleri** denilmektedir.

Hipotez

Kuramsal olarak varsayılan ya da önceden yapılmış bir dizi gözleme dayanarak ortaya atılan gerçekleşmesi mümkün olabilen önermeye **Hipotez** denir.

Hipotez,

Araştırma Hipotezi

İstatistiksel Hipotez

olmak üzere iki tipte sınıflandırılabilir.

Araştırma Hipotezi

Bir araştırmada ortaya konulmaya çalışılan sonuçları, olumlu ya da olumsuz olarak cümlelerle ifade eden önermelere **Araştırma Hipotezi** denir.

Örneğin, yeni geliştirilen bir ilacın akciğer kanserinde olumlu bir etkiye sahip olup olmadığını saptamak üzere yapılan bir araştırmanın hipotezi

“Yeni geliştirilen ilaç akciğer kanserinde kanserli hücreleri yok etme özelliğine sahiptir.”

şeklinde kurulabilir.

İstatistiksel Hipotez

Bir ya da daha fazla toplumla ilgili olarak ileri sürülen ve belli bir dağılım varsayımı altında

- Parametrik değerlerin belli bir değere eşit olup olmadığını,
- İki ya da daha fazla toplumun parametrelerinin birbirlerine eşit olup olmadığını ya da farklı olduklarını,
- Değişkenlerin birbirlerinde bağımsız olduğunu vb.

gibi ifadelerle sembolik olarak gösterilen ve geçerliliği istatistiksel önemlilik testleri ile belirlenen önermelere

İstatistiksel Hipotez denir.

İstatistiksel Hipotez

İstatistiksel hipotez birbirlerine alternatif iki hipotez kullanılarak formüle edilir.

- 1- Sıfır Hipotezi (Null Hypothesis)
- 2- Karşıt Hipotez (Alternative Hypothesis)

İstatistiksel Hipotez

Sıfır Hipotezi (Null Hypothesis)

İncelenen değişkenin toplumdaki dağılımının, parametresinin değişmediği, belirli bir değere eşit olduğu, iki ya da daha fazla topluma ait parametrelerin birbirlerine eşit olduğu, aralarındaki farkların sıfır olduğu, ilişki düzeylerinin sıfır olduğu, değişkenlerin bağımsız olduğu vb. şeklinde formüle edilen istatistiksel hipoteze **Sıfır Hipotezi** denir ve H_0 ile gösterilir.

İstatistiksel Hipotez

Karşıt Hipotez (Alternative Hypothesis)

İncelenen değişkenin toplumdaki dağılımının, parametresinin değiştiği, belirli bir değere eşit olmadığı, iki ya da daha fazla topluma ait parametrelerin birbirlerine eşit olmadığı, aralarındaki farkların sıfır olmadığı, ilişki düzeylerinin sıfırdan farklı olduğu, değişkenlerin bağımlı olduğu vb. şeklinde formüle edilen istatistiksel hipoteze **Karşıt Hipotezi** denir **H_1** ile gösterilir.

İstatistiksel Hipotez

H_0 ve H_1 hipotezleri,

Parametre türüne (Ortalama, Oran, İlişki Katsayısı)

Örneklem sayısına (Tek, İki, ...)

Dağılım tipinin dikkate alınıp alınmamasına

göre farklı biçimlerde kurulur.

İstatistiksel Hipotez

Örnek 1:

Gebelikte sigara kullanımının düşük doğum ağırlığına neden olduğu araştırılmak isteniyor.

Araştırma öncesinde ise gebelikte sigara kullanan bayanların **2500 gr.dan daha düşük ağırlığa** sahip doğum yaptıkları ileri sürülüyor.

İleri sürülen bu hipotezin test edilmesi için kurulacak sıfır ve karşıt hipotezler nelerdir, nasıl formüle edilir?

İstatistiksel Hipotez

Bu araştırmada;

Toplum “gebeliği esnasında sigara kullanan tüm hamile bayanlar”

olarak tanımlanabilir.

Bu topluma ait

Parametre “sigara kullanan hamile bayanlara ait ortalama doğum ağırlığı”

şeklinde ifade edilir.

İstatistiksel Hipotez

H_0 : Sigara kullanımı düşük doğum ağırlığına neden olmaz.

H_1 : Sigara kullanımı düşük doğum ağırlığına neden olur.

$H_0 : \mu = 2500 \text{ gr.}$

$H_1 : \mu < 2500 \text{ gr.}$

İstatistiksel Hipotez

Örnek 2:

Bir yerleşim yerinin su ihtiyacını karşılamak üzere doğal bir akarsudan su sağlanarak baraj kurulmak ve bu barajdan su ihtiyacının karşılanması planlanıyor.

Ancak akarsuyun **pH seviyesinin nötr olmadığı yani 7'ye eşit olmadığı** iddia ediliyor.

İleri sürülen bu hipotezin test edilmesi için kurulacak sıfır ve karşıt hipotezler nelerdir, nasıl formüle edilir?

İstatistiksel Hipotez

Bu araştırmada;

Toplum “akarsuyun kaynağından kurulacak baraja kadar geçtiği tüm noktalardaki suyun pH değerleri” olarak tanımlanabilir.

Bu topluma ait

Parametre “tüm noktalardaki pH değerlerinin ortalaması” şeklinde ifade edilir.

İstatistiksel Hipotez

H_0 : Akarsuyun pH seviyesi nötrdür.

H_1 : Akarsuyun pH seviyesi nötr değildir.

$$H_0 : \mu = 7$$

$$H_1 : \mu \neq 7$$

İstatistiksel Hipotez

Örnek 3:

Sigara ve alkol kullananların Sistolik Kan Basınçlarının normal seviye olarak kabul edilen **120 mm/Hg'den daha yüksek olduğu**, sigara ve alkol kullanımının hipertansiyon hastalığının en önemli etkenleri arasında yer aldığı iddia ediliyor.

Bu amaçla yapılacak olan bir araştırmada ileri sürülen bu hipotezin test edilmesinde kullanılacak sıfır ve karşıt hipotezler nelerdir, nasıl formüle edilirler?

İstatistiksel Hipotez

Bu araştırmada;

Toplum “sigara ve alkol kullanan tüm yetişkin bireyler”

olarak tanımlanabilir.

Bu topluma ait

Parametre “sigara ve alkol kullanan tüm yetişkin bireylerin ortalama sistolik kan basıncı”

şeklinde ifade edilir.

İstatistiksel Hipotez

H_0 : Sigara ve alkol kullananların sistolik kan basıncı ortalaması normaldir.

H_1 : Sigara ve alkol kullananların sistolik kan basıncı ortalaması yüksektir.

$$H_0 : \mu = 120 \text{ mm/Hg}$$

$$H_1 : \mu > 120 \text{ mm/Hg}$$

İstatistiksel Hipotez

Örnek 1

$$H_0: \mu = 2500$$

$$H_1: \mu < 2500$$

Örnek 2

$$H_0: \mu = 7$$

$$H_1: \mu = 7$$

Örnek 3

$$H_0: \mu = 120$$

$$H_1: \mu > 120$$

Herhangi bir değişimi, farklılığı, eşitsizliği, bağımlılığı vb. ifadeleri içeren önermeler her zaman karşıt hipotezde belirtilir.

Yanılma Payı, I. ve II. Tip Hata

Sıfır ve karşıt hipotezden hangisini seçmeliyiz?

Hangi hipotezin doğru olduğuna nasıl karar vermeliyiz?

Yanılma Payı, I. ve II. Tip Hata

Örnek 1'deki araştırmada sigara kullanan 18 gebe doğuma kadar takip edilmiş ve doğum sonrası bebeklerin doğum ağırlıkları elde edilmiş ve ortalama olarak $\bar{X}=2395$ gr. bulunmuş olsun.

Bu değer H_1 hipotezini kabul etmemiz için yeterli midir?

Yanılma Payı, I. ve II. Tip Hata

Örnek 2'deki araştırmada akarsu yatağı boyunca 40 farklı noktadan alınan su örneklerinin laboratuvar ortamında pH değerleri hesaplanıp ortalaması $\bar{X}=7.4$ olarak bulunsun.

Acaba bu değer H_1 hipotezini kabul etmemiz için tek başına yeterli midir?

Yanılma Payı, I. ve II. Tip Hata

Örnek 3'deki araştırmada sigara ve alkol kullanan 20 bireyin sistolik kan basınçlarının ortalaması $\bar{X}=149 \text{ mm/Hg}$ olarak elde edilsin.

Acaba bu değer H_1 hipotezini kabul etmemiz için tek başına yeterli midir?

Yanılma Payı, I. ve II. Tip Hata

Örnekten elde edilen değerleri kullanarak sıfır ya da karşıt hipotezden birinin doğru olduğuna karar verirken hata yapma olasılığımız her zaman vardır.

Yanılma Payı, I. ve II. Tip Hata

Çünkü deneyi tekrarladığımızda farklı örnekten farklı ortalama sonuçları bulmamız olasıdır.

Bu durumda gerçekte sıfır hipotezi doğru iken örneklem hatası nedeniyle karşıt hipotezi doğru kabul edebiliriz ya da tam tersini yapabiliriz.

Yanılma Payı, I. ve II. Tip Hata

Gerçekte doğru olan hipotezi kabul etmeyip yanlış olan hipotezi doğru olarak kabul etme olasılığına **Yanılma Payı** denir.

İki tip yanılma payı vardır, bunlar

I. Tip Hata

II. Tip Hata

olarak adlandırılır.

Yanılma Payı, I. ve II. Tip Hata

Gerçekte doğru olan H_0 hipotezinin kabul edilmeyerek red edilmesine ve H_1 hipotezinin kabul edilmesine **I. Tip Hata** denir ve α ile gösterilir.

Gerçekte yanlış olan H_0 hipotezinin red edilmeyerek kabul edilmesine **II. Tip Hata** denir ve β ile gösterilir.

Yanılma Payı, I. ve II. Tip Hata

Hipotez Testinin Sonucu	Gerçek Durum	
	H_0 Doğru	H_0 Yanlış
H_0 Kabul	$1 - \alpha$ (Kabul Olasılığı)	β (II. Tip Hata)
H_0 Red	α (I. Tip Hata)	$1 - \beta$ (Testin Gücü)

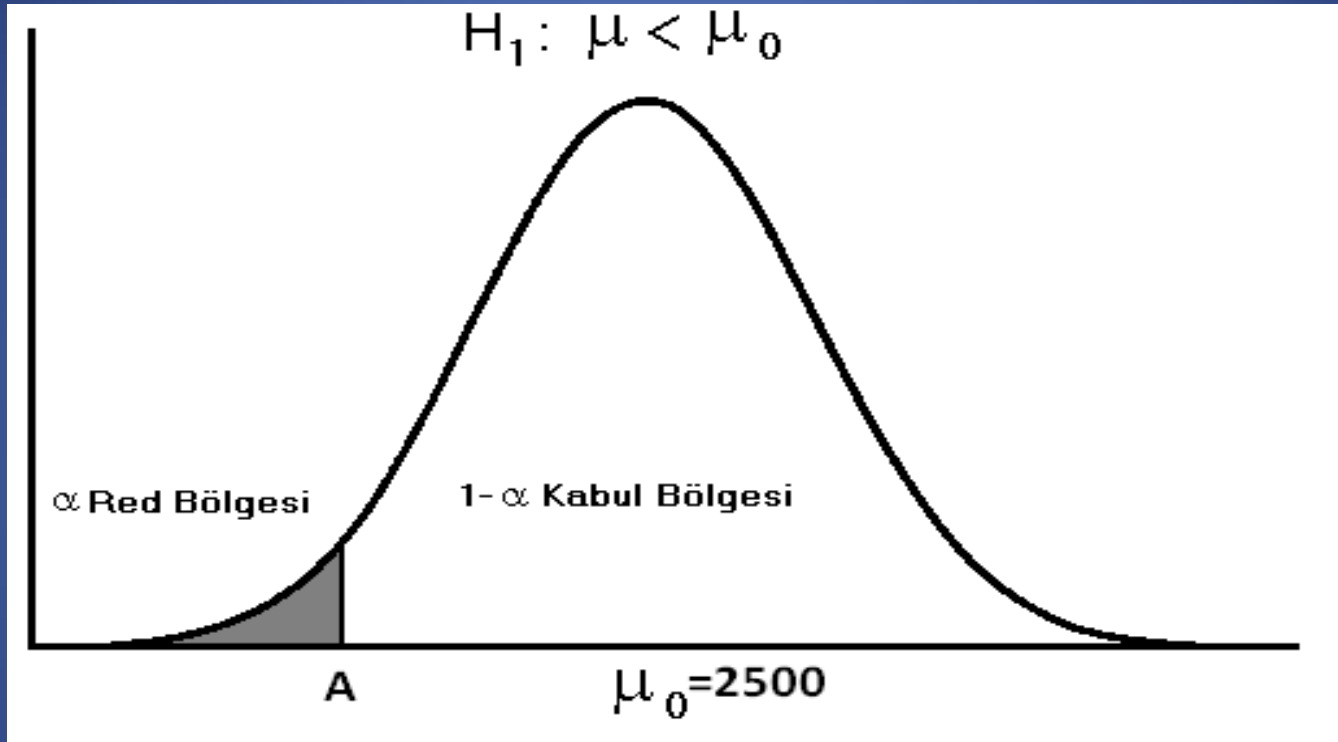
Yanılma Payı, I. ve II. Tip Hata

Örnek 1'de acaba örnekten hesapladığımız ortalama değer kaç olsun ki biz H_0 hipotezini kabul ya da red edelim.

Bu kritik değeri bulmak için bir dağılım varsayımı yapılır ve **I. Tip Hata** değeri belirlenir.

Bu örnekte normal dağılım varsayımı kullanılarak bir **A** kritik değeri elde edilir.

Yanılma Payı, I. ve II. Tip Hata



Eğer örnekten elde edilen ortalama değer bu **A** kritik değerinden;

küçük ise **H_0 red** edilir,

büyük ya da eşitse **H_0 kabul** edilir.

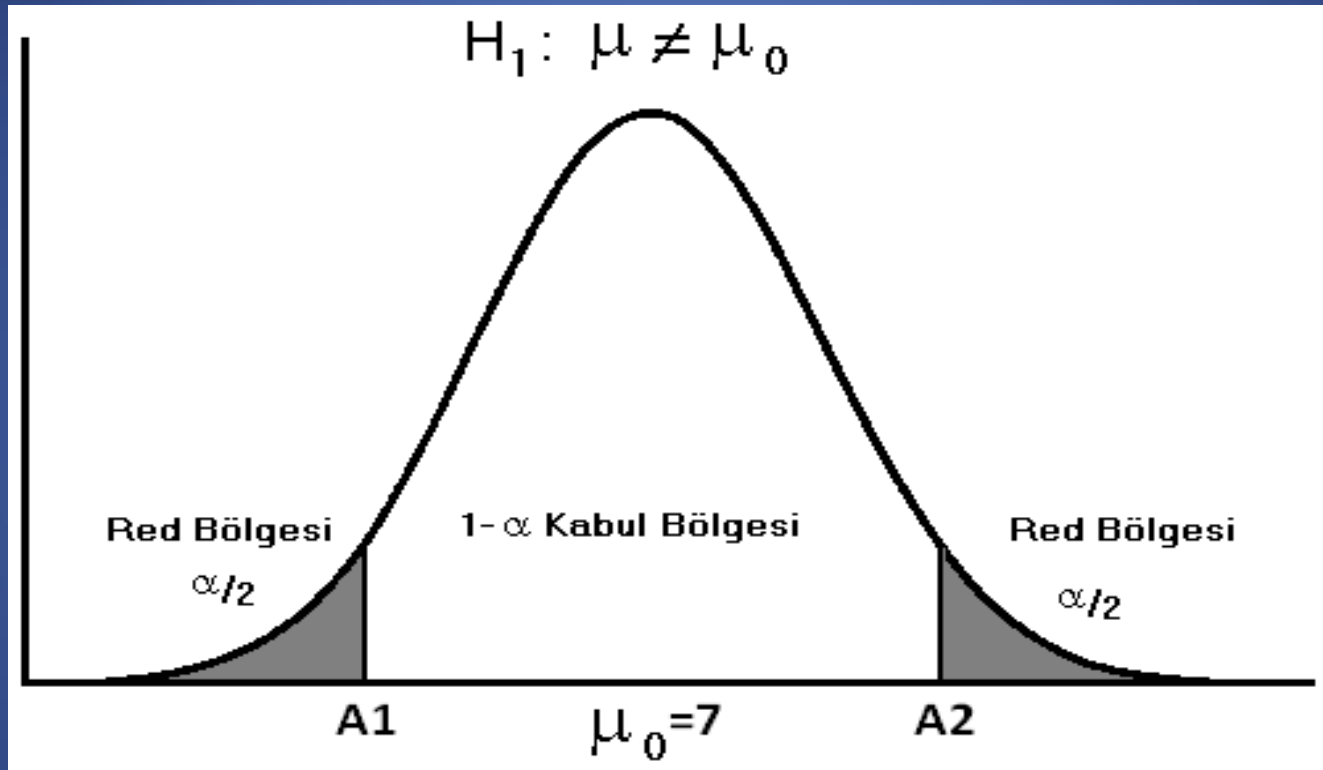
Yanılma Payı, I. ve II. Tip Hata

Örnek 2'de acaba örnekten hesapladığımız ortalama değer kaç olsun ki biz H_0 hipotezini kabul ya da red edelim.

Bu kritik değeri bulmak için bir dağılım varsayımı yapılır ve I. Tip Hata değeri belirlenir.

Bu örnekte normal dağılım varsayımı kullanılarak iki kritik değer $A1$ ve $A2$ elde edilir.

Yanılma Payı, I. ve II. Tip Hata



Eğer örnekten elde ettiğimiz ortalama bu iki değer arasındaysa;

H_0 kabul edilir,

A1'den küçük ya da A2'den büyük ise H_0 red edilir.

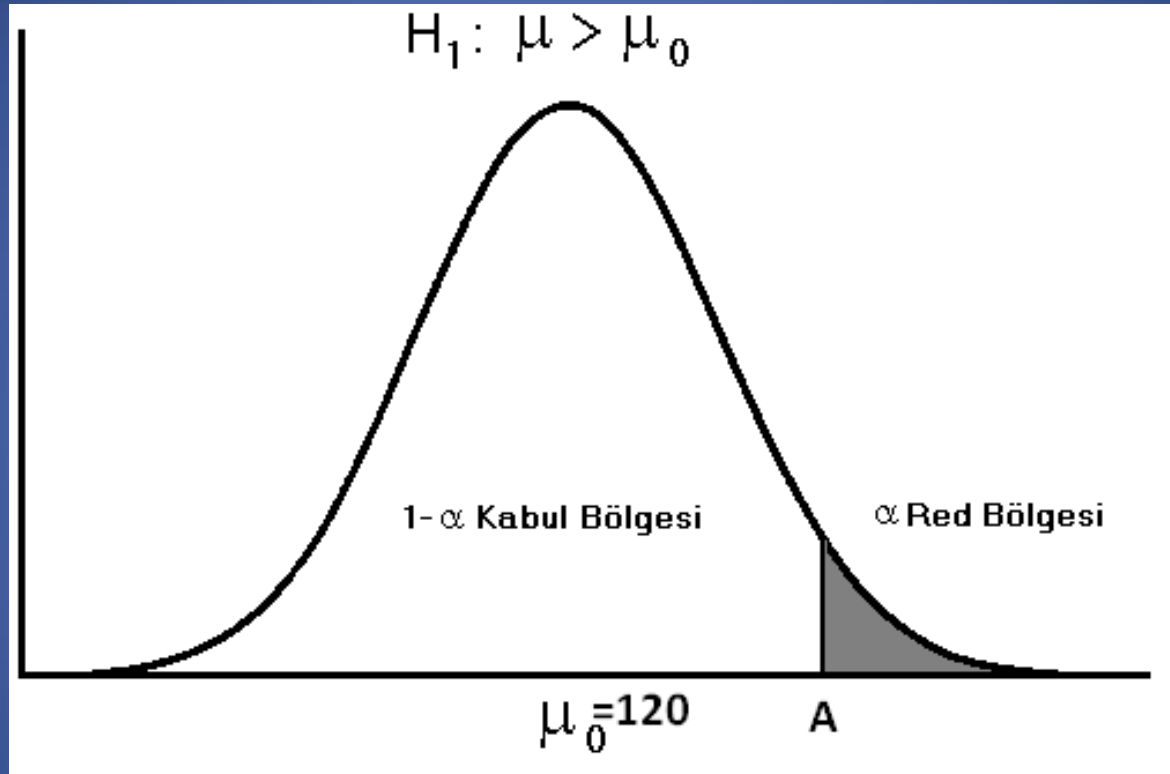
Yanılma Payı, I. ve II. Tip Hata

Örnek 3'de acaba örnekten hesapladığımız ortalama değer kaç olsun ki biz H_0 hipotezini kabul ya da red edelim.

Bu kritik değeri bulmak için bir dağılım varsayımı yapılır ve I. Tip Hata değeri belirlenir.

Bu örnekte normal dağılım varsayımı kullanarak aşağıdaki gibi bir A kritik değer elde edilir.

Yanılma Payı, I. ve II. Tip Hata

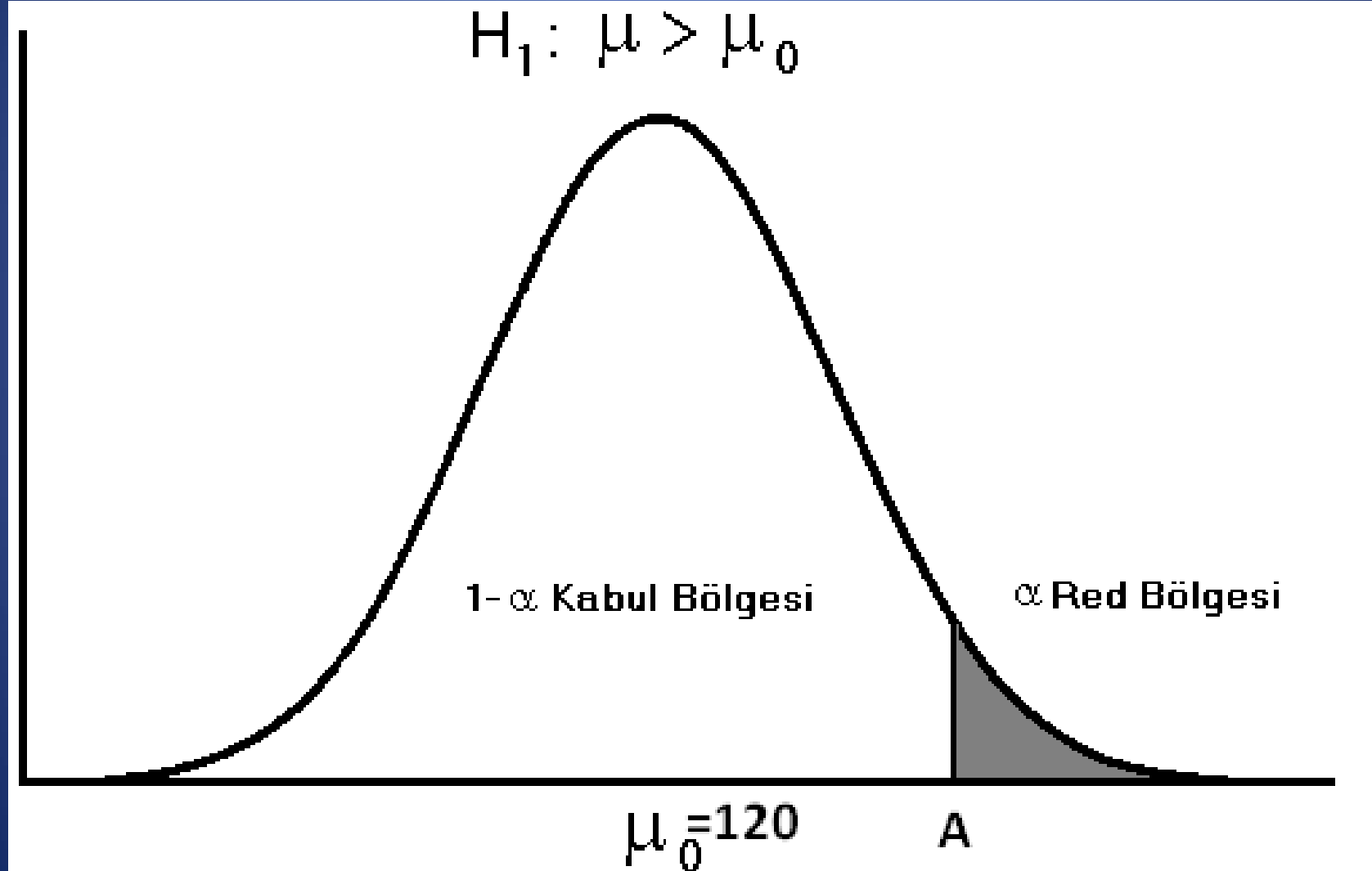


Eğer örnekten elde edilen ortalama değer bu **A** kritik değerden büyükse;

H₀ red edilir,

küçük ya da eşitse **H₀** kabul edilir.

Yanılma Payı, I. ve II. Tip Hata



Önemlilik Düzeyleri

Pratikte hipotez testlerinden elde edilen sonuçları belirli bir yanılma payına göre değerlendirmek yerine, seçilen

$\alpha=0.05$, $\alpha=0.01$ ve $\alpha=0.001$

düzeylerinin üçünü de birlikte göz önüne alarak önemliliği belirlemek daha uygun olmaktadır.

Önemlilik Düzeyleri

α değerlerine bağlı olarak kritik A noktalarının belirlenmesi ve karara varılması **İstatistiksel Önemlilik Testleri** ile yapılmaktadır.

Önemlilik Düzeyleri

Z, T, F gibi herhangi bir test istatistiğinin gözlenme olasılığına bakarak önemliliği belirlemek için dört koşul vardır.

Test İstatistiği ve Yanılma payı	P(T) Olasılık	Karar Olasılığı	Karar	Önemlilik Düzeyi
$ T < \alpha_{0.05}$	$P > 0.05$	$P > 0.05$	Önemli değil	ns
$\alpha_{0.05} \leq T < \alpha_{0.01}$	$0.05 \leq P < 0.01$	$P < 0.05$	Önemli	*
$\alpha_{0.01} \leq T < \alpha_{0.001}$	$0.01 \leq P < 0.001$	$P < 0.01$	Çok önemli	**
$ T \geq \alpha_{0.001}$	$P \leq 0.001$	$P < 0.001$	İleri düzeyde önemli	***

Tek ve İki Yönlü Test

Bir hipotez testinde, test istatistiğinin sonucu H_0 ve H_1 varsayımlarının formülasyonuna göre değerlendirilir.

$$H_0 : \mu = \mu_0 \text{ iken}$$

$$H_1 : \mu > \mu_0 \text{ ya da } H_1 : \mu < \mu_0$$

gibi yön belirtiyorsa test istatistiğinin önemliliği kabul ve red bölgelerinin yerine göre farklılık gösterecektir.

Tek ve İki Yönlü Test

Eğer **red** bölgesini belirten α olasılığı dağılımın sağ ya da sol ucunda yer alıyorsa bu teste **Tek Yönlü Test** adı verilir.

Eğer $H_0:\mu=\mu_0$ iken $H_1:\mu\neq\mu_0$ formüle edilmiş ise test istatistiğinin değerlendirilmesi farklılık gösterir.

Bu durumda **red** bölgesi dağılımın her iki ucunda $\alpha/2$ olarak ele alınır. Bu teste de **İki Yönlü Test** adı verilir.

Tek ve İki Yönlü Test

Sağlık alanında değişkenlerin ortalama değerden çok küçük ya da çok büyük değerler alması genel eğilimden önemli farklılaşmaları belirttiğinden dolayı, sonuçların her iki yönde farklı sonuçları daha iyi yansıtacağı kabul edilmektedir.

Bu nedenle tıbbi denemelerde ya da araştırmalarda **İki Yönlü Hipotezlerin** kurulması ve **İki Yönlü Testlerin** yapılması daha yaygındır.

İstatistiksel Önemlilik Testleri

İstatistiksel Önemlilik Testleri, kurulan sıfır ve karşıt hipotezlerin test edilerek bu hipotezlerden hangisinin doğru olarak seçileceğini gösteren testlerdir.

Başka bir anlatımla H_0 ya da H_1 hipotezlerinden hangisinin doğru olduğuna karar vermede kullanılan yöntemlere **İstatistiksel Önemlilik Testleri** denir.

İstatistiksel Önemlilik Testleri

İstatistiksel önemlilik testleri; kurulan hipotezde parametre kullanılıp kullanılmamasına, parametre sayısına, grup sayısına, örneklemin bağımlı olup olmadığına ve test edilecek değişken sayısına göre farklı biçimlerde sınıflandırılır.

İstatistiksel Önemlilik Testleri

Kullanılacak değişken sayısına göre

- 1- Tek değişkenli (univariate) önemlilik testleri
- 2- Çok değişkenli (multivariate) önemlilik testleri

Kurulan hipotezin parametreye dayalı olup olmamasına göre

- 1- Parametrik önemlilik testleri
- 2- Parametrik olmayan önemlilik testleri

İstatistiksel Önemlilik Testleri

Parametre sayısına göre

- 1- Tek örneklem testler
- 2- İki örneklem testler
- 3- k örneklem ($k > 2$) testler

Örneklemin bağımlı olup olmadığına göre

- 1- Bağımlı örneklem testleri

Bağımlı iki örneklem, bağımlı k örneklem

- 2- Bağımsız örneklem testleri

Bağımsız iki örneklem, bağımsız k örneklem

İstatistiksel Önemlilik Testleri

Test edilen değişkenin parametresinin toplumda ele alınışına göre ya da testte hipotezin kuruluş biçimine göre

- 1- Ortalamaya dayalı testler
- 2- Oransal değerlere dayalı testler
- 3- Gözlem sayılarına dayalı testler
- 4- Ortanca değere dayalı testler
- 5- Dağılım varsayımına dayalı testler
- 6- İlişkiye dayalı testler
- 7- Ard arda gelişlerin rastgeleliğine dayalı testler
- 8- Uyuma, uygunluğa, uyuşuma dayalı testler

İstatistiksel Önemlilik Testleri

Bağımlı Örneklem

Aynı birimden bir değişkene ilişkin en az iki kere ölçüm sonunda elde edilen veri yapısına Bağımlı Örneklem adı verilir.

Aynı birimden elde edilen ölçüm sayısı,

iki ise Bağımlı İki Örneklem,

ikiden fazla ise Bağımlı k Örneklem

olarak adlandırılır.

İstatistiksel Önemlilik Testleri

Batın ameliyatı geçirmiş kişilerde ameliyat anında ve ameliyattan 1 saat sonra ölçülen kalp atım hızları bağımlı iki örneklem veri yapısına uymaktadır.

Çünkü aynı kişilere ait kalp atım hızları hem ameliyat anında hem de ameliyattan 1 saat sonra ölçülmüştür.

İstatistiksel Önemlilik Testleri

Bağımsız Örneklem

Bir değişkene ilişkin ölçümlerin birbiriden farklı en az iki grupta yapılmasıyla elde edilen veri yapısına Bağımsız Örneklem adı verilir.

İstatistiksel Önemlilik Testleri

Hipertansiyonu düşürücü etkisi olan iki farklı ilaçtan hangisinin daha etkin olduğunu araştırmak üzere bir çalışma planlansın.

Birinci ilaç hipertansiyon hastası 15 kişide, ikinci ilaç ise yine hipertansiyon hastası 17 kişide denensin.

Bu çalışmadan elde edilen veriler bağımsız iki örneklem veri yapısına uymaktadır. Çünkü gruplardaki kişiler aynı bireyler değildir.

Konu Başlıkları

Parametrik Hipotez Testleri

Bağımsız Gruplarda t testi

Bağımlı Gruplarda t testi

Tek Yönlü Varyans Analizi

İki Yönlü Anova

Parametrik Olmayan Hipotez Testleri

MannWhitney U Testi

Wilcoxon T testi

Kruskal Wallis Testi

Friedman Testi

2*2 Kikare Testi

Pearson Kikare

Yates Kikare

Fisher Kikare

R*C Kikare testi