

HİPOTEZ TESTLERİNE GİRİŞ

UYGUN HİPOTEZİN KURULMASI

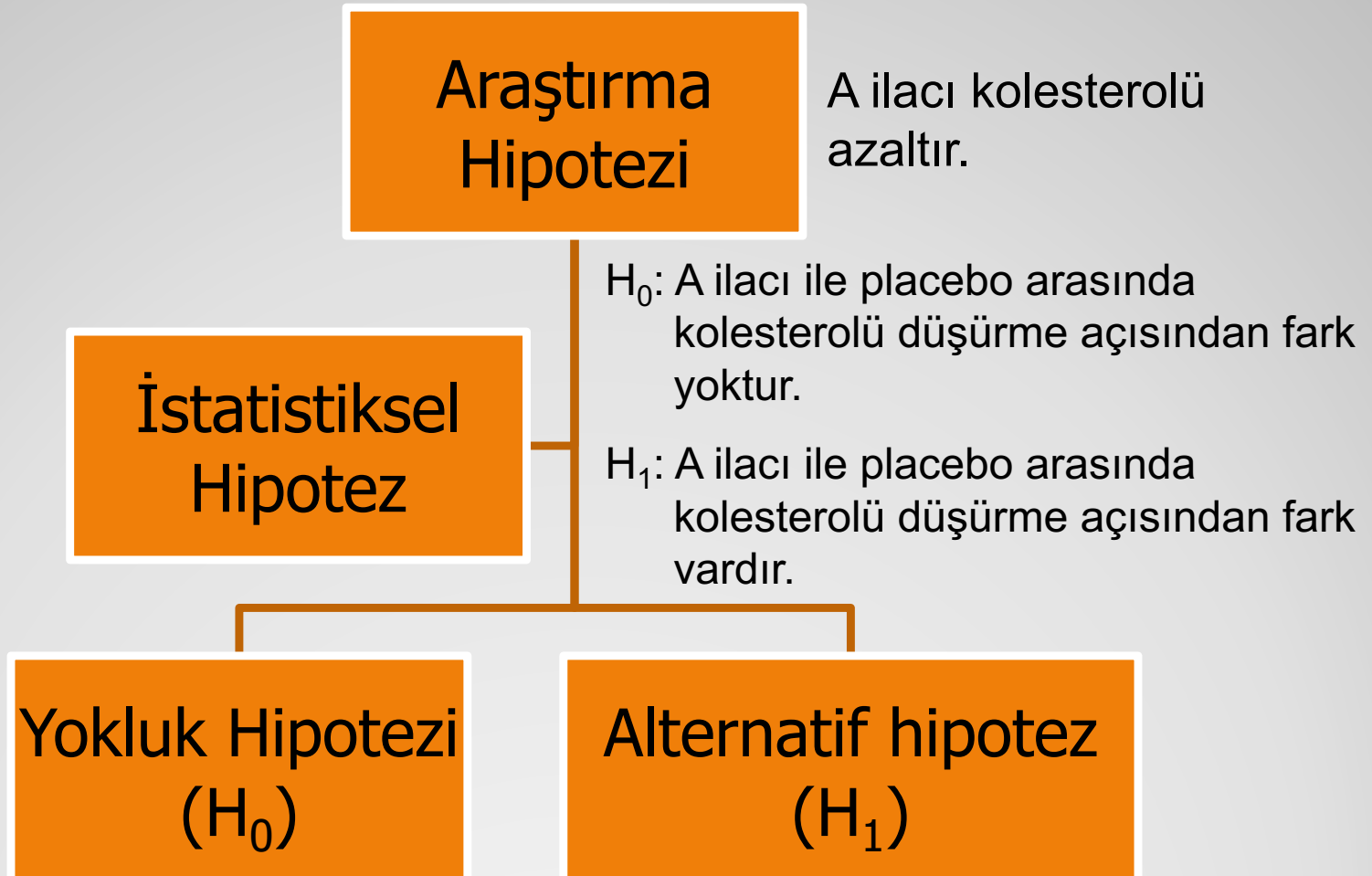
TİP I HATA VE TİP II HATA

- HİPOTEZ: Örneklemeye dayalı bir kitle parametresinin değeri hakkında ileri sunulan iddia olarak tanımlanabilir.

ÖRNEKLER:

- *Kandaki hemoglobin miktarı bakımından barsak paraziti olanlarla olmayanlar arasında fark yoktur.*
- *İki ilacın iyileştirme yüzdeleri aynıdır.*
- *Kızamık görülme sayısı bakımından mevsimler arasında fark yoktur.*

Hangi Hipotez?



Hipotez Testleri

Hipotez testi; kitleden çekilen örneklem yardımıyla kitle hakkında bir karara varmak için örneklemeyle dayalı sistematik izlenen bir seri işleminden oluşur.

Örneklem İstatistiği

$\bar{x}$: Örneklem ortalaması

SS: Standart sapma

Kitle parametresi

μ : Aritmetik ortalama

σ : Standart sapma

1. Hipotezlerin Belirlenmesi

İstatistiksel hipotezler, örneklemelerin çekildiği kitle parametreleri ile ilgili önermelerdir. İstatistikte kullanılan iki tür hipotez vardır:

- H_0 hipotezi
- H_1 hipotezi

H_0 hipotezine, yokluk hipotezi denir. Bir testte öne sürülen ve test edilmek istenen hipotezdir.

H_1 hipotezine, alternatif hipotez denir ve yokluk hipotezine karşı kurulan hipotezdir.

- İstatistikte kullanılan hipotezler tek ya da çift yönlü olabilir.
- Bir hipotezin tek ya da çift yönlü olduğunu H_1 hipotezi (alternatif hipotez) belirler.
- Yokluk hipotezine karşılık kurulabilecek 3 farklı alternatif hipotez vardır.

1. ÇİFT YÖNLÜ HİPOTEZ

H_0 : Hemoglobin ortalamaları bakımından gebe olan kadınlarla, gebe olmayanlar arasında fark yoktur

H_1 : Hemoglobin ortalamaları bakımından gebe olan kadınlarla, gebe olmayanlar arasında fark vardır.

şeklindeki hipotez **çift yönlüdür**. Çünkü hemoglobin ortalamaları arasında fark olması önemli olup, gebe olan ya da olmayan kadınlara ait hemoglobin ortalamasının diğerinden düşük ya da yüksek olması önemli değildir.

2. TEK YÖNLÜ HİPOTEZ

H_0 : Hemoglobin ortalamaları bakımından gebe olan kadınlarla, gebe olmayanlar arasında fark yoktur

H_1 : Gebe olmayan kadınların hemoglobin ortalaması, gebe olanların hemoglobin ortalamasından daha yüksektir.

şeklindeki hipotez **tek yönlüdür**. Çünkü gebe olmayan kadınların hemoglobin ortalamasının, gebe olanların hemoglobin ortalamasından yüksek olduğu iddia edilmektedir.

3. TEK YÖNLÜ HİPOTEZ

H_0 : Hemoglobin ortalamaları bakımından gebe olan kadınlarla, gebe olmayanlar arasında fark yoktur

H_1 : Gebe olmayan kadınların hemoglobin ortalaması, gebe olanların hemoglobin ortalamasından daha düşüktür.

şeklindeki hipotez **tek yönlüdür**. Çünkü gebe olmayan kadınların hemoglobin ortalamasının, gebe olan kadınların hemoglobin ortalamasından düşük olduğu iddia edilmektedir.

Hipotezlerde, kitleyi tanımlamada kullanılan parametrelerden yararlanılır.

$$\begin{aligned} 1. \quad H_0 : \mu_1 &= \mu_2 \\ H_1 : \mu_1 &\neq \mu_2 \end{aligned}$$

$$\begin{aligned} 2. \quad H_0 : \mu_1 &= \mu_2 \\ H_1 : \mu_1 &> \mu_2 \end{aligned}$$

$$\begin{aligned} 3. \quad H_0 : \mu_1 &= \mu_2 \\ H_1 : \mu_1 &< \mu_2 \end{aligned}$$

2. Hipotezler İçin Uygun Test ve Test İstatistiğine Karar Verilmesi

Değişik hipotez testleri için değişik test istatistiklerinden yararlanılır.

Örneğin:

İki grubun ortalamasının karşılaştırılması durumunda kullanılacak hipotez testi, üç grubun ortalamasının karşılaştırılması durumunda kullanılacak hipotez testinden farklıdır.

- Kullanılan istatistiksel teste bağlı olarak, kullanılacak test istatistikleri de farklılık gösterir.
- *Uygun testi, dolayısıyla test istatistiğini seçmek, hipotez testlerinin en önemli aşamasıdır.*

3. İstatistiksel Test için Yanılma Düzeyinin Belirlenmesi

Yanılma düzeyi çoğunlukla, istatistiksel test uygulanmadan önce araştırmacı tarafından belirlenir ve α (alfa) olarak tanımlanır. α , gerçekte H_0 hipotezi doğru iken, test sonucunda H_0 hipotezini reddetme olasılığını verir. Yanılma düzeyi olarak $\alpha = 0.05$, $\alpha = 0.01$, $\alpha = 0.001$ gibi küçük değerler alınır. Çünkü, H_0 hipotezi gerçekten doğru iken reddedilmek istenmez.

3. İstatistiksel Test için Yanılma Düzeyinin Belirlenmesi (Devam...)

Yanılma düzeyi ile ilgili diğer bir kavram, *p değeri*dir. İstatistik paket programlarının tümü, sonuç aşamasında elde edilen test istatistiklerine ilişkin p değerini çıktı olarak verirler.

- p değeri de bir olasılık değeri olup, gerçekte H_0 hipotezi doğru iken, gözlenen değerlere bağlı olarak bu hipotezin reddedilme olasılığını verir.
- Örneğin $p=0.197$ ise, reddetme olasılığımız 0.197'dir. p değeri istatistiksel testin uygulanmasından sonra elde edilir.

4. İstatistiksel Karar

Yapılan test sonucunda elde edilen test istatistiği değeri, her bir test yöntemi için farklı şekilde bulunan teorik tablo değerine eşit ya da tablo değerinden büyük ise H_0 hipotezi reddedilir. Aksi durumda H_0 hipotezi reddedilemez.

Test istatistiği değeri $\geq$ Tablo değeri $\rightarrow H_0$ red

Test istatistiği değeri $<$ Tablo değeri $\rightarrow H_0$ reddedilemez

4. İstatistiksel Karar (Devam...)

H_0 hipotezi, p değeri yardımıyla da reddedilebilir ya reddedilemez. Hesaplanan test istatistiğine bağlı olarak bulunan p değeri, daha önce seçtiğimiz yanılma düzeyinden küçük ise H_0 hipotezi reddedilir, küçük ise reddedilemez. Örneğin, $\alpha = 0.05$ olarak belirlendiğinde, test istatistiğine ilişkin p değeri 0.002 bulunmuş ise H_0 hipotezi reddedilir.

$$p < \alpha \Rightarrow$$

H_0 red

$$p \geq \alpha \Rightarrow$$

H_0 reddedilemez

Hipotez Testleri Sırasında Ortaya Çıkabilecek Hatalar:

Bir hipotez testinde H_0 hipotezini kabul ya da reddederken iki tür hata yapılabilir.

Tip I Hata (α): H_0 hipotezi gerçekte doğru iken, bu hipotezin yanlışlıkla reddedilmesidir.

Tip II Hata (β): H_0 hipotezi gerçekte yanlış iken, bu hipotezin yanlışlıkla kabul edilmesidir.

Tip I ve Tip II Hata

		Hipotez Testi Sonucundaki Karar	
		H_0 red	H_0 reddedilemez
Gerçek Durum	H_0 doğru	Yanlış Karar Tip I hata (α)	Doğru karar ($1 - \alpha$)
	H_0 yanlış	Doğru karar (Testin Gücü) $1 - \beta$	Yanlış Karar Tip II hata (β)

- İstatistiksel yöntemler kullanılırken α türü hata olasılığı araştırmacı tarafından belirlenir ve genellikle 0.05, 0.01 ya da 0.001 olarak alınır.
- β türü hatalar konusunda genellikle bir denetimimiz yoktur, ancak 0.20 civarında olması istenir.
- α ve β türü hatalar arasında ilişki vardır, α azalırken β artar. Ancak α sabit tutulurken denek sayısının artması β türü hata olasılığını küçültür.

ÖRNEKLEM BÜYÜKLÜĞÜ

ÖRNEKLEM BÜYÜKLÜĞÜ

Bir araştırma planlanırken ilk sorulan soru, örneklem büyüklüğünün nasıl belirleneceği ve ne kadar olacağıdır. Bu soruya kolayca cevap vermek her zaman mümkün olmamaktadır.

Genellikle bilimsel çalışmaların çoğunda birden fazla değişken incelenmektedir. Ancak, çalışmanın ana amacı ile ilişkili olan değişken için örneklem büyüklüğünün hesaplanması yeterlidir.

Örneklem büyüklüğünü hesaplamadan önce popülasyon ya da örneklem hakkında bir takım bilgilere ihtiyacımız olmaktadır.

Örneklem Büyüklüğünü Etkileyen Başlıca Faktörler

- 1) Tip I Hata (α):** Doğru bir yokluk hipotezinin (H_0), yanlışlıkla reddedilmesi olasılığıdır. $(1 - \alpha)$ ise testin güvenirlik düzeyidir.
- 2) Güç ($1 - \beta$):** H_0 hipotezi yanlış olduğu zaman, H_0 hipotezini reddetme olasılığıdır. Diğer bir ifadeyle güç, gerçekte gruplar arasında fark varken, test sonucunda H_0 'ın reddedilerek fark vardır denilmesi olasılığıdır. β ise Tip II hatadır.
- 3) Etki Büyüklüğü (δ):** Çalışma sonucunda elde edilen tahmin değerinin, popülasyon değerinden ne kadar sapabileceğini gösterir ve araştırmacı tarafından belirlenir. Gruplar arasında ne kadarlık bir fark, klinik olarak anlamlı kabul edilmelidir? Bu sorunun cevabı etki büyüklüğünü verecektir. Eğer araştırmacı bu soruyu cevaplandıramıyorsa, daha önce yapılmış çalışmalardan ya da pilot çalışmalardan faydalanır.

TEST SONUCUNDAKİ KARAR

GERÇEK

H_0 Hipotezini
Kabul Etme
(fark yok)

H_0 Hipotezini
Ret etme
(fark var)

H_0 Hipotezi Doğru
Olduğu Zaman
(fark yok)

Doğru karar

Tip I hata
(α)

H_0 Hipotezi Yanlış
Olduğu Zaman
(fark var)

Tip II hata
(β)

Doğru karar
($1 - \beta$)

Çalışmaya başlamadan önce Tip I ve Tip II hata düzeyleri araştırmacı tarafından belirlenir.

Genellikle Tip I hata 0.05'ten küçük, Tip II hata ise 0.20'den daha küçük olarak belirlenir.

Araştırmacı, önemlilik düzeyini seçmek suretiyle Tip I hatayı kontrol eder.

Tip II hatayı kontrol altına almak sadece denek sayısını değiştirmekle mümkündür. Denek sayısı arttıkça Tip II hata yapma olasılığı azalmaktadır.

İstatistikte kullanılan birçok hipotez testi olduğu bilinmektedir. Bu testlerin her biri için farklı örneklem büyüklüğü formülleri olduğu düşünülürse, bunlar için örneklem büyüklüğü hesaplamanın çokta kolay bir işlem olmadığı söylenebilir. Bu nedenle, örneklem büyüklüğü hesaplamak amacıyla kullanılan birçok paket program geliştirilmiştir. Sonuç olarak, araştırmalar gerçekleştirilmeden önce bu programlar kullanılarak uygun örneklem büyüklüğü hesaplanmalıdır.

Örnek olarak, gebelerde hemoglobin ortalamasını tahmin etmek istediğimizi varsayalım. Ancak ne kadarlık bir örneklem büyüklüğüne ihtiyaç olduğunu hesaplamak için aşağıdaki sorulara cevap verilmesi gerekmektedir.

1) Popülasyon ortalaması (μ) ile tahmin değeri arasında en fazla ne kadarlık bir fark kabul edilebilir? ($d = ?$)

2) Popülasyon ortalaması ile tahmin değeri arasındaki farkın, d değerinden büyük olma olasılığı en fazla yüzde kaç olacak? ($\alpha = ?$)

3) İncelenecek değişkene ait popülasyon varyansı ya da onun tahmini ($\sigma^2 = ?$ ya da $s^2 = ?$)

Belirli bir bölgede yaşayan gebelerde hemoglobin ortalamasını tahmin etmek istediğimizi varsayalım. Örneklem büyüklüğünü hesaplamak için gerekli olan istatistikleri elde etmek amacıyla 100 kişilik bir örneklem seçilerek bir pilot çalışma düzenlenmiştir. Bu çalışma sonunda hemoglobin ortalaması 11, standart sapması (s) 4 olarak bulunmuştur. Bilinmeyen popülasyon ortalamasını %95 güvenirlilikle ± 0.5 'lik bir hata (d) ile bulabilmek için aşağıdaki formül kullanılarak yaklaşık 251 gebenin seçilmesi gereklidir.

$$n = \frac{z_{1-\alpha/2}^2 s^2}{d^2} = \frac{1.98^2 \times 4^2}{0.5^2} \cong 251$$