

YZM 511 - İstatistiksel Yapay Öğrenme

İstatistiksel Hipotez Testleri

6/11/2024

Hipotez Testinin Temel Kavramları

Hipotez Nedir?

Hipotez: İstatistiksel analizlerde doğruluğunu test etmek istediğimiz varsayımdır.

Sıfır Hipotezi (H_0): Genellikle "değişiklik yok" veya "etki yok" durumunu temsil eder. Hipotez testlerinin temel amacı, H_0 'ın geçerli olup olmadığını anlamaktır.

Alternatif Hipotez (H_1): H_0 'a karşı olan iddiadır; H_0 'ın yanlış olduğunu savunur.

Hipotez Testinin Temel Kavramları

Ornekler

- Bir ilaç şirketi, yeni geliştirdikleri bir ilacın migren ağrılarını azaltmada etkili olup olmadığını test etmek istiyor. Bu durumda hipotez şu olabilir:
 - **H0 (Sıfır Hipotezi):** Yeni ilaç migren ağrılarını azaltmada etkili değildir.
 - **H1 (Alternatif Hipotez):** Yeni ilaç migren ağrılarını azaltmada etkilidir.

- Bir şirket, müşteri memnuniyetinde yapılan yeni iyileştirmelerin etkisini ölçmek istiyor. Burada hipotez şu şekilde olabilir:
 - **H0:** Yapılan iyileştirmeler müşteri memnuniyetini artırmamıştır.
 - **H1:** Yapılan iyileştirmeler müşteri memnuniyetini artırmıştır.

- Bir e-ticaret şirketi, web sitesinin yeni tasarımının satışları artırıp artırmadığını test etmek istiyor. Hipotez şöyle kurulabilir:
 - **H0:** Yeni tasarım, satışları artırmamaktadır.
 - **H1:** Yeni tasarım, satışları artırmaktadır.

Hipotez Testinin Temel Kavramları

Hipotez Testi Adımları

➤ Araştırma Sorusu Belirleme

- Yeni bir eğitim programının öğrencilerin başarı oranını artırıp artırmadığını görmek istiyoruz.

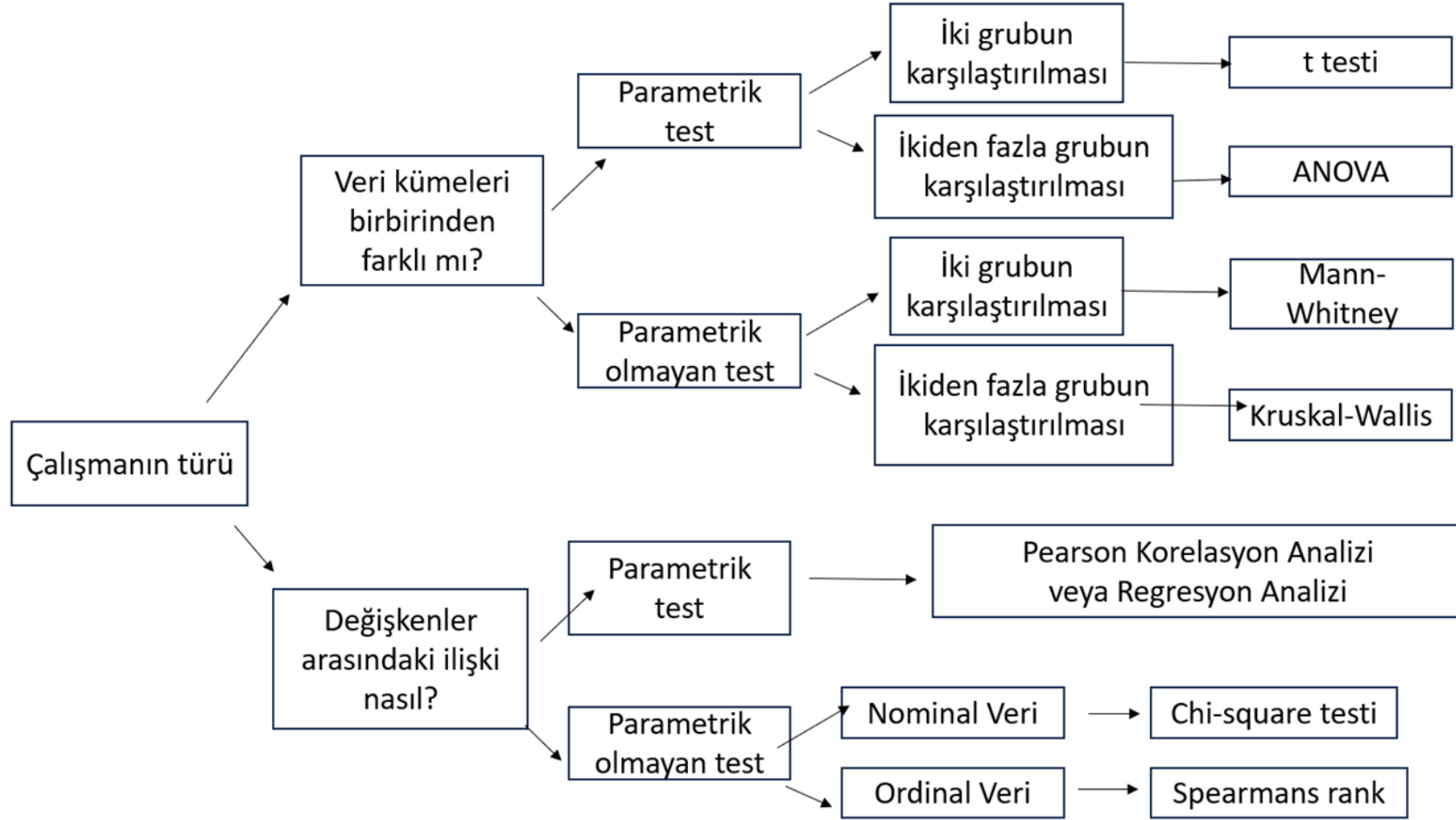
➤ Hipotezlerin Kurulması:

- **H0 (Sıfır Hipotezi):** Yeni eğitim programı ile başarı oranında değişiklik yoktur.
- **H1 (Alternatif Hipotez):** Yeni eğitim programı ile başarı oranı artmaktadır.

➤ Uygun Test Tipini Seçme: Verinin türüne ve araştırmanın amacına göre uygun testi belirlemek (t-testi, z-testi, ki-kare testi vb.).

➤ p-Değeri ve Karar Verme:

- **p-değeri:** Gözlemlenen sonuçların, H0 doğruysa, tesadüfen ortaya çıkma olasılığını verir.
- Eğer $p < 0.05$ ise, H0 reddedilir ve H1 kabul edilir.



Parametrik ve parametrik olmayan testler arasındaki ayrım

Parametrik Testler

- Parametrik testler, belirli varsayımlara dayanan testlerdir (normal dağılım, homojen varyans).
- Sürekli ve oranlı ölçeklerdeki verilerle çalışmak için uygundur.
 - ✓ t-Testleri, z-Testleri, ANOVA (Varyans Analizi)

Parametrik Olmayan Testler

- Belirli bir dağılım varsayımına dayanmaz. Bu testler, dağılımı bilinmeyen veya normal olmayan verilerle çalışırken tercih edilir ve genellikle medyan veya sıralama gibi daha az katı parametreler üzerinden analiz yaparlar
- Sürekli, sıralamalı ve kategorik verilerle çalışmak için uygundur.
 - ✓ Mann-Whitney U Testi, Kruskal-Wallis Testi, Ki-Kare Testi

Parametrik Testler

Uygun Test Tipini Seçme

t-Testleri

- iki grup arasındaki ortalamaları karşılaştırmak için kullanılır.
- Popülasyon standart sapması bilinmediğinde
- Örneklem küçük olduğunda (genellikle $n < 30$) tercih edilir.

z-Testleri

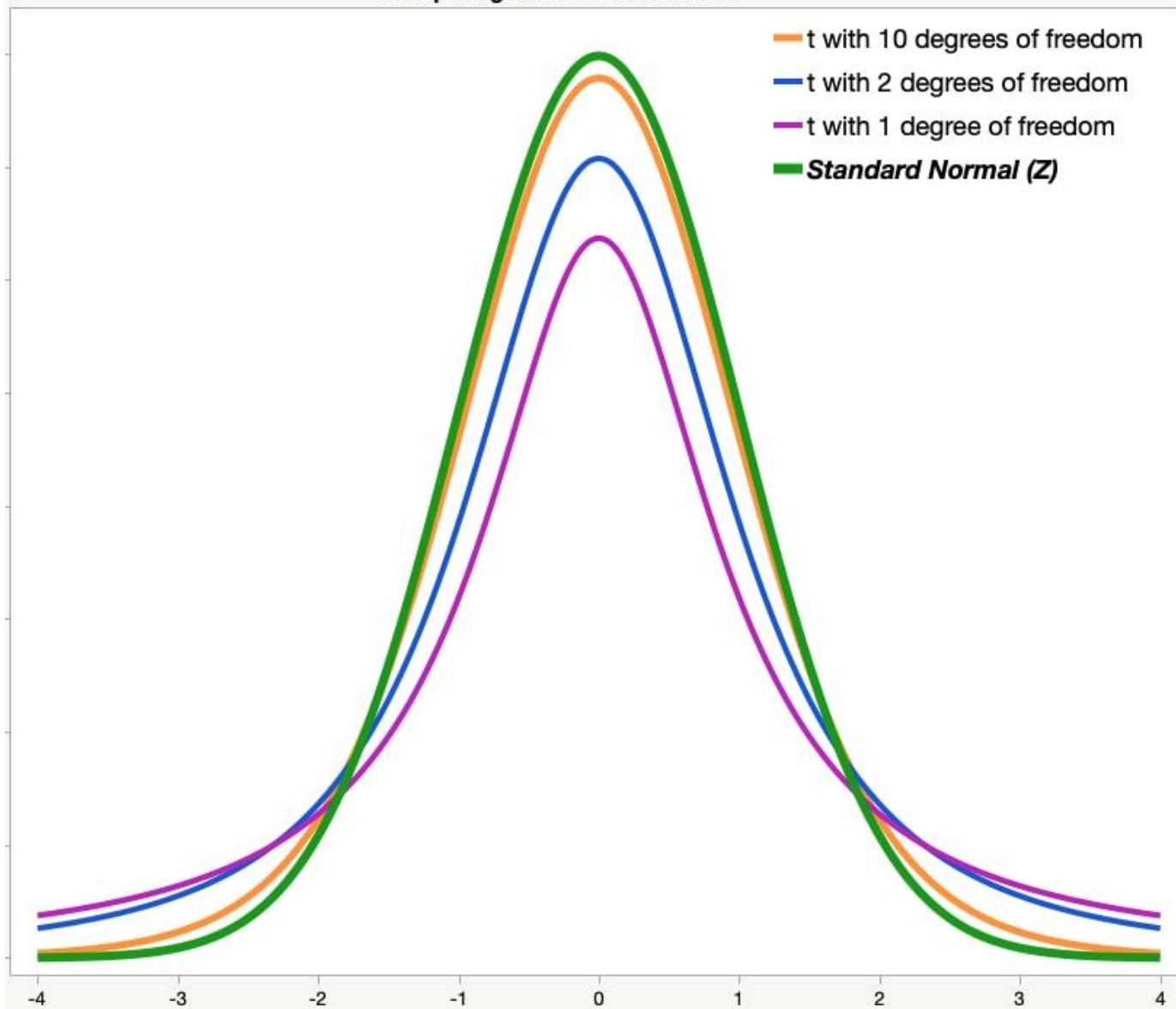
- iki grup arasındaki ortalamaları/oranları karşılaştırmak için kullanılır.
- Popülasyon standart sapması **bilindiğinde**
- Büyük örneklerle (genellikle $n > 30$) çalışırken

ANOVA (Varyans Analizi)

- Birden fazla grubun ortalamalarını karşılaştırmak için kullanılır.

t-dağılımı, örneklem belirsizlik veya değişkenlik payını da hesaba katmak için biraz daha yayvandır (kuyrukları daha geniştir), bu da küçük örneklemle çalışırken daha doğru bir analiz sağlar.

Comparing t and Z Distributions



Parametrik Olmayan Testler

Uygun Test Tipini Seçme

Ozellikleri

- **Dağılım Varsayımı Olmaması**
- **Sıralamalı ve Medyan Üzerine Çalışır**
- **Küçük ve Büyük Örneklemeler** Hem küçük hem de büyük örneklemelerle güvenle uygulanabilir.
- **Veri Tipi:** Sürekli, sıralamalı ve kategorik verilerle çalışmak için uygundur.

Ornekler

Mann-Whitney U Testi: İki bağımsız grubun medyanlarını karşılaştırmak için kullanılır.

Wilcoxon İşaretli Sıra Testi: Aynı grubun iki farklı durumda ölçülen medyanlarını karşılaştırmak için kullanılır.

Kruskal-Wallis Testi: Üç veya daha fazla grubun medyanlarını karşılaştırmak için kullanılan parametrik olmayan bir alternatiftir.

Ki-Kare Testi: Kategorik verilerle çalışarak beklenen ve gözlenen frekanslar arasındaki farkı inceler.

p-Degeri ve Karar Verme Sureci

Anlamlilik Duzeyi

Anlamlılık Düzeyi (Alpha)

- Anlamlılık düzeyi (α - alpha) bir tür hata oranıdır ve yapılan testte sıfır hipotezini (H_0) yanlışlıkla reddetme olasılığını gösterir.
- Örneğin, $\alpha=0.05$ olarak belirlendiğinde, araştırmacı sıfır hipotezi doğru olsa bile onu %5 oranında yanlışlıkla reddetmeyi göze almış demektir. Bu, yapılan testte %5'lik bir hata payı bırakıldığı anlamına gelir.

- ❖ **$p < 0.05$ olduğunda:** Sonuçlar, sıfır hipotez doğru olsa bile yalnızca %5'lik bir olasılıkla şansa bağlı olarak ortaya çıkabilir. Bu düşük olasılık, sıfır hipotezini reddetmek için yeterli kanıt olduğuna işaret eder.
- ❖ **$p \geq 0.05$ olduğunda:** Bu durumda sıfır hipotezin doğru olma olasılığı daha yüksek kabul edilir. Yani, elde edilen sonuçların şansa bağlı olma ihtimali göz ardı edilemeyecek kadar yüksektir.

p-Degeri ve Karar Verme Sureci

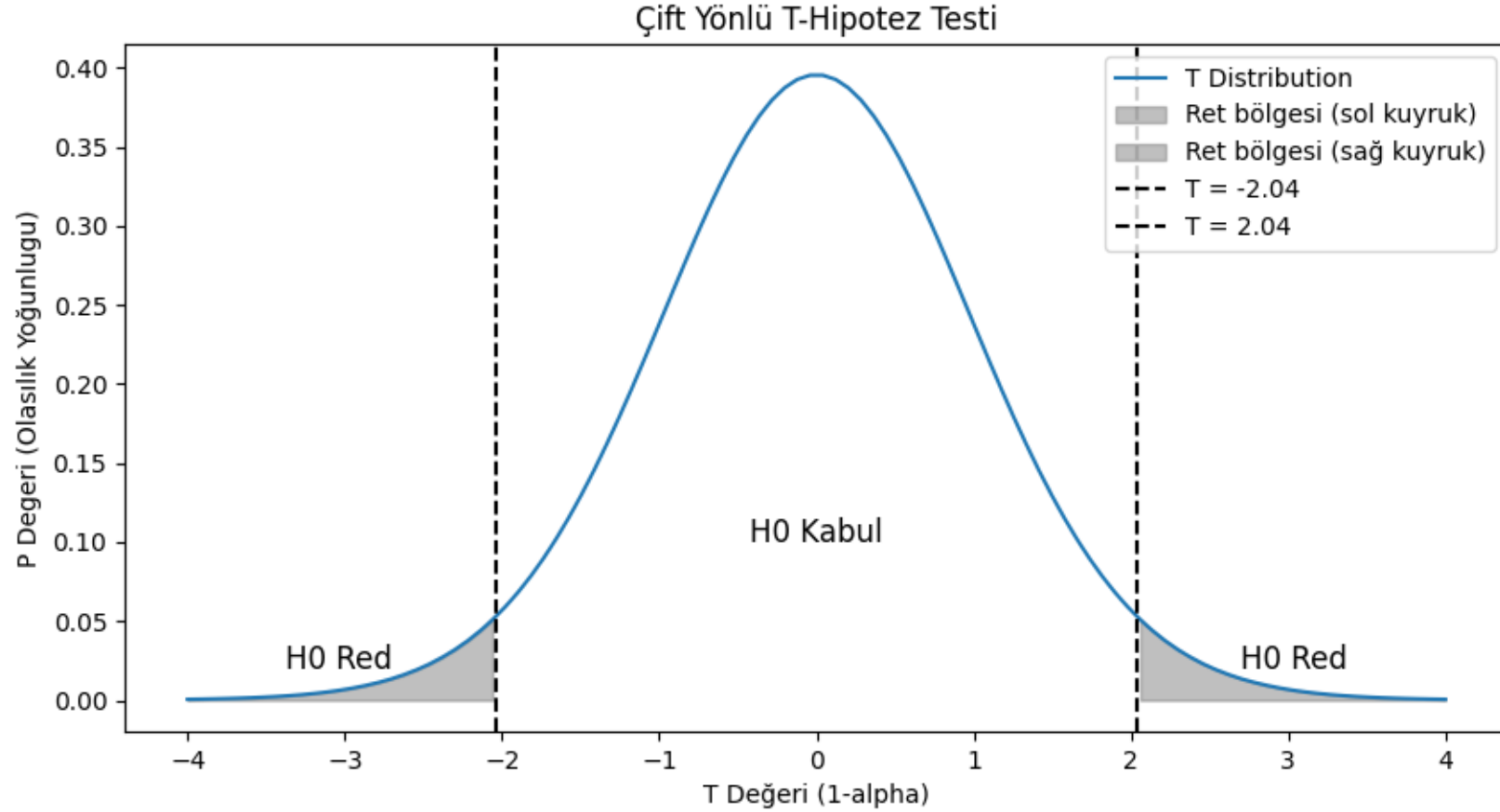
Anlamlilik Duzeyi

Alternatif Anlamlılık Düzeyleri

0.05 yaygın bir standart olsa da, bazı durumlarda farklı anlamlılık düzeyleri kullanılır. Örneğin:

- **0.01 (1%):** Daha katı bir anlamlılık düzeyidir. Özellikle, hata riskinin çok düşük tutulması gerektiği (örneğin, tıbbi araştırmalar gibi) çalışmalarda tercih edilir.
- **0.10 (10%):** Daha esnek bir anlamlılık düzeyidir. Sosyal bilimler gibi alanlarda bazen kullanılır, çünkü bazı çalışmalarda hata payının biraz daha yüksek olmasına izin verilebilir.

α = Anlamlılık düzeyini ifade eder. (ret bölgeleri)
(1- α) = Güven aralığı



$$H_0 : \mu_1 = \mu_2$$

$$H_A : \mu_1 \neq \mu_2$$

Hipotez Testlerinde Hatalar

Hipotez testinde iki tür hata yapılabilir: **Tip I Hata** ve **Tip II Hata**.

Tip I Hata (α): H_0 doğruyken yanlışlıkla reddetmek.

❖ Yeni eğitim programı aslında etkisizken, onu etkili gibi kabul etmek.

Tip II Hata (β): H_0 yanlışken yanlışlıkla kabul etmek.

❖ Yeni eğitim programı etkiliyken, etkisiz olduğunu düşünmek.

Hipotez Testlerinde Hatalar

Tip I Hata (α) – Yanlış Pozitif

Anlamı: Tip I hata, elde edilen sonucun yalnızca tesadüf nedeniyle ortaya çıkmasına rağmen anlamlı bir fark varmış gibi yorumlanmasıdır. Bu hata türü, "yanlış pozitif" bir sonuca yol açar, yani gerçekte olmayan bir etki veya fark varmış gibi kabul edilir.

Anlamlılık Düzeyi (α): Tip I hata olasılığı, anlamlılık düzeyi (α) ile belirlenir. Genellikle %5 (0.05) olarak seçilir, yani Tip I hata yapma olasılığı %5 kabul edilir.

Sonuç: Tip I hata, araştırmacıyı yanlış bir pozitif sonuç kabul etmeye yönlendirir ve bu yanlış kabul, etkisiz bir yöntemin etkili gibi görünmesine yol açar.

Hipotez Testlerinde Hatalar

Tip II Hata (β) – Yanlış Negatif

Tanım: Sıfır hipotezi (H_0) yanlış olduğu halde, H_0 'ı reddetmeyip kabul etmek anlamına gelir. Bu hata, bir etki veya fark olmasına rağmen, testin bu farkı anlamlı bir sonuç olarak gösterememesi durumunda ortaya çıkar.

Anlamı: Tip II hata, "yanlış negatif" bir sonuca yol açar, yani var olan bir etki veya fark göz ardı edilir. Test, aslında var olan bir değişimi veya etkili bir durumu ortaya çıkaramaz.

Sonuç: Tip II hata, gerçek bir etkinin göz ardı edilmesine neden olur, bu da etkili bir yöntemin etkisiz olduğu sonucuna yol açar.

Hipotez Testlerinde Hatalar

Tip I ve Tip II Hataların Karşılaştırılması

	Tip I Hata (α)	Tip II Hata (β)
Tanım	H0 doğruyken onu yanlışlıkla reddetmek	H0 yanlışken onu yanlışlıkla kabul etmek
Sonuç	Yanlış pozitif: Gerçekte olmayan bir etki var gibi kabul edilir	Yanlış negatif: Gerçekte olan bir etki göz ardı edilir
Kontrol Yolu	Anlamlılık düzeyini (α \alpha) düşük tutmak	Örneklem büyüklüğünü artırmak veya testin gücünü yükseltmek
Örnek Durumu	Etkisiz bir ilacın etkili olduğunu düşünmek	Etkili bir ilacın etkisiz olduğunu düşünmek

Özet Olarak Hipotez Testi Aşamaları

- 1.Hipotezleri Belirleme (H_0 , H_1)
- 2.Anlamlılık Düzeyini Belirleme (α)
- 3.Uygun Testi Seçme
- 4.Test İstatistiğini Hesaplama
- 5.p-Değerini Bulma ve Karar Verme
- 6.Sonuçların Yorumlanması ve Raporlanması

Aşama 1: Hipotezlerin kurulması

Web Tasarımının Gücü: Yeni Arayüz, Okuyucu Etkileşimini Artırıyor mu?

$$\mu_a = 800$$



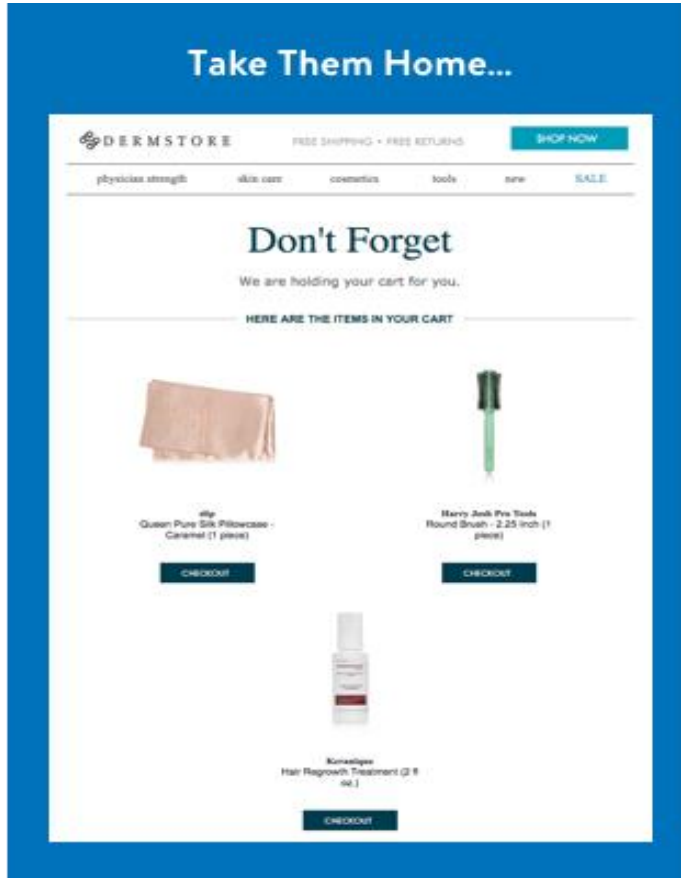
$$\mu_b = 1100$$



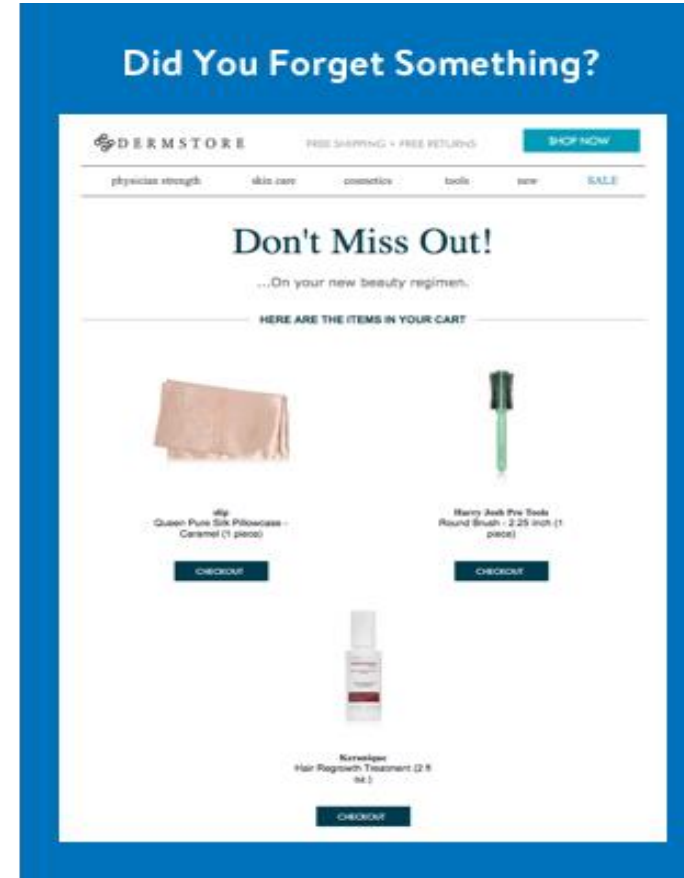
$$H_0 : \mu_a = \mu_b$$

$$H_A : \mu_a \neq \mu_b$$

Unutmayın,
sizin için
tutuyoruz!



Kaçırmayın!
Sepetinizdeki
güzellik
rejiminizi
tamamlayın.



$$H_0 : \mu_1 = \mu_2$$

$$H_A : \mu_1 \neq \mu_2$$

$$\mu_1 = 30.000$$



$$\mu_2 = 35.000$$

Kadıköy

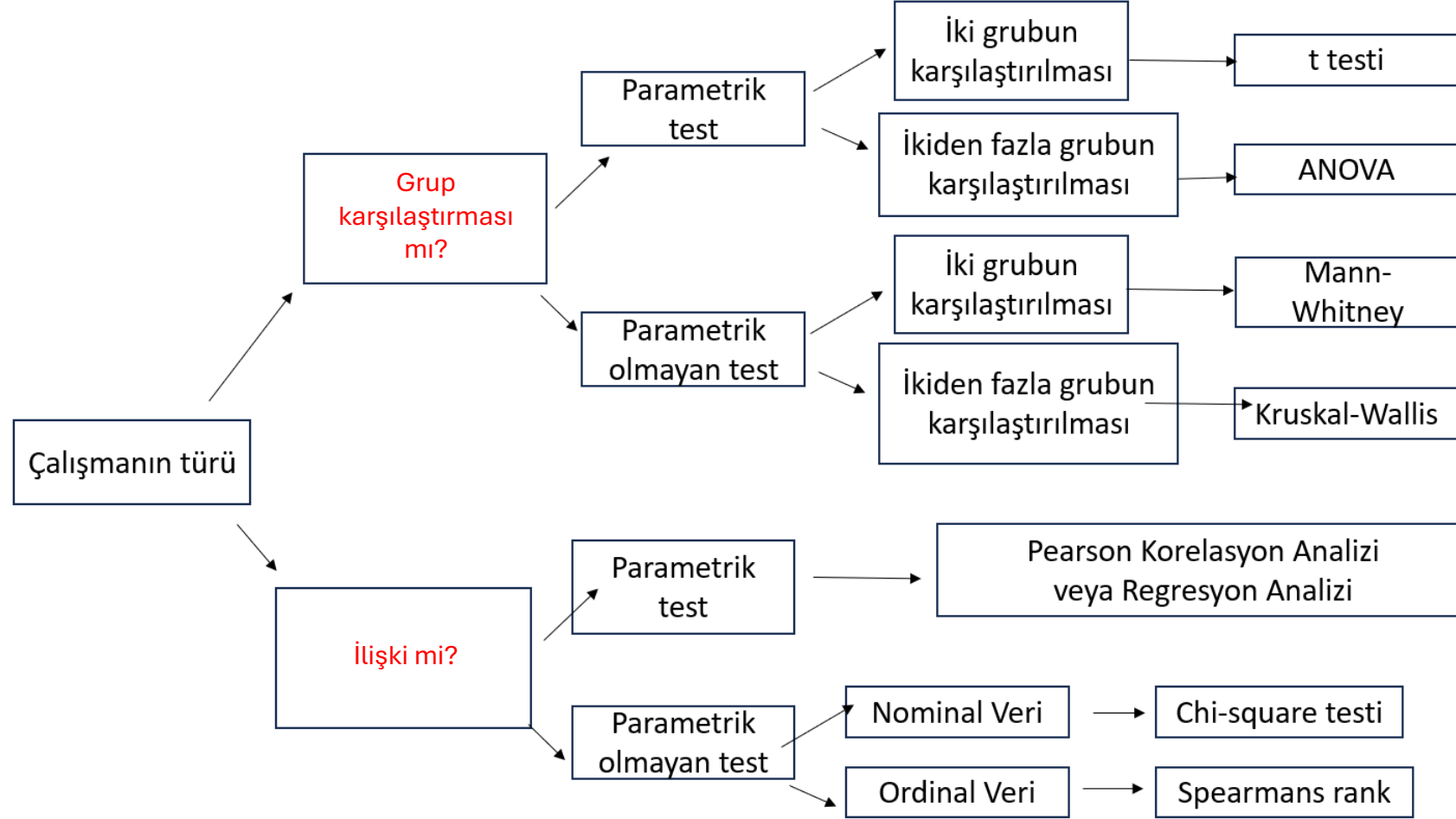


$$H_0 : \mu_1 = \mu_2$$

$$H_A : \mu_1 \neq \mu_2$$

Aşama 2: Uygun İstatistiksel Teste Karar Verilmesi

Hipotezler İçin Uygun İstatistiksel Teste Karar Verilmesi



Parametrik Testler: Veri dağılımları ile ilgili belirli varsayımları karşılayan veriler için kullanılır. Bu testler genellikle aşağıdaki varsayımlara dayanır:

Normal Dağılım: Verinin normal (Gauss) dağılımına uyduğu varsayılır.

Varyans Homojenliği: Test edilen grupların varyanslarının (homoscedasticity) aynı olduğu varsayılır.

Bağımsızlık: Gözlemlerin birbirinden bağımsız olduğu varsayılır.

Parametrik olmayan test: Non-parametrik testler, verinin dağılımı hakkında spesifik varsayımlar yapmadan kullanılan testlerdir. (Yukarıdaki varsayımlar aranmaz.)

Aşama 3: Test İstatistiği ve P-Değeri (Olasılık Değeri) Hesaplama

Aşama 4: Karar Verme

A/B Testi

(bağımsız iki örneklem t-testi)