TEMEL KAVRAMLAR

Daha önceki derslerde hipotez testinin NE olduğunu ve HANGİ testlerin kullanılması gerektiği anlatıldı. Şimdi NEDEN ve NASIL soruları derinlemesine incelenecek.

Bu bölümün sonunda şunları anlamış oalcaksınız:

- Bir p değerinin gerçekte ne anlama geldiği (sadece "$p < 0,05$ ise H_0'yi reddet" ifadesi değil)
- Neden $\alpha = 0,05$ kullanıyoruz (ve neden biraz keyfi)
- Güven aralıkları ve hipotez testleri aynı madalyonun iki yüzüdür
- Serbestlik dereceleri gerçekte neyi temsil eder (sadece formüller olarak değil, sezgisel olarak)
- Tüm bu vramlar birbirine nasıl bağlanır

Dedektif Hikayesi: İstatistiksel Kanıtları Anlamak

Diyelim ki bir madeni paranın adil mi yoksa taraflı mı olduğunu araştıran bir dedektifsiniz.

Setup: Biri size bir madeni para uzatıyor ve "Bu adil bir madeni para - %50 yazı, %50 tura." diyor. Ama siz şüpheleniyorsunuz. Bu yüzden denemeye karar veriyorsunuz.

Test: Madeni parayı 20 kere havaya atıyorsunuz ve 15 kez yazı 5 kez tura geliyor.

Sorunuz: "Eğer bu para gerçekten adil olsaydı, 20 atıştan 15'inin yazı gelme olasılığı sadece şans eseri ne kadar olurdu?"

Bu soru - bu TEK soru - TÜM hipotez testlerinin kalbidir. Her kavramı bu bakış açısıyla incelenmelidir.

P Değerini Anlamak: İstatistikte En Çok Yanlış Anlaşılan Kavram

İnsanların P-Değerinin Anlamını Düşünmesi (YANLIŞ!)

❌ "H_0'nin doğru olma olasılığı" ❌ "Sonucumun şans eseri ortaya çıkma olasılığı" ❌ "Hata yapma olasılığı" ❌ "H_1'nin doğru olma olasılığı

BUNLARIN HEPSİ YANLIŞ! İstatistiklerdeki 1 numaralı kafa karışıklığı kaynağıdır.

P-Değeri Aslında Ne Anlama Geliyor (DOĞRU!)

p-değeri:

"Eğer H_0 doğruysa, gerçekte gözlemlediğim kadar uç bir veri elde etme olasılığı nedir?"

Bunu madeni para örneğiyle dikkatlice açıklayalım.

Madeni Para Örneği: P-Değerinin Adım Adım Hesaplanması

Hipotezlerimiz:

- H_0: Madeni para adildir (yazı için p = 0,5)
- H_a: Madeni para yazıya doğru eğilimlidir (p > 0,5)

Gözlemimiz: 20 atıştan 15'i yazı.

P-Değeri Sorusu: "Eğer madeni para adil ise (H_0 doğruysa), 20 atıştan 15 veya daha fazlasının yazı gelme olasılığı nedir?"

Bunu bir düşünelim:

Eğer para gerçekten adilse, çoğu zaman 10 civarı yazı (20'nin yarısı) gelir. Bazen 9 veya 11, bazen 8 veya 12, nadiren 7 veya 13 ve çok nadiren 15 gibi uç bir sayı gelir.

Binom dağılımını kullanarak (hesaplama detaylarıyla ilgilenmeden) şunlar bulunabilir:

- P(exactly 15 heads) = 0.0148
- P(exactly 16 heads) = 0.0046
- P(exactly 17 heads) = 0.0011
- P(exactly 18 heads) = 0.0002
- P(exactly 19 heads) = 0.00002
- P(exactly 20 heads) = 0.000001

p değeri = tüm bu olasılıkların toplamı = 0,0207 (yaklaşık %2)

Bu P-Değerinin Yorumlanması

p = 0,0207'nin anlamı: "Madeni para gerçekten adil olsaydı, 20 atışta 15 veya daha fazla yazı gelme olasılığı yalnızca %2 olurdu."

Bunun bize gösterdiği şey şu: Adil bir para için 15 kez tura gelmesi oldukça sıra dışı bir durum. İmkansız değil, ama paranın taraflı olduğundan şüphelenmemize yetecek kadar sıra dışı.

Bunun bize söylemediği şey:

- ❌ Bize madeni paranın taraflı olma olasılığını söylemiyor (bu farklı bir soru!)
- ❌ Bize madeni paranın ne kadar taraflı olduğunu SÖYLEMEZ
- ❌ Bize %2 olasılıkla bir hata yaptığımızı söylemiyor

Resmi Tanım

p değeri:

P (elde ettiğimiz kadar uç gözlem verileri | H_0 doğrudur)

"|" sembolü "verildiği takdirde" veya "varsayıldığı takdirde" anlamına gelir.

Önemli Bilgi: p değeri, H_0'nin doğru olduğu varsayılarak hesaplanır. H_0'nin doğru olma olasılığı değildir!

Şöyle düşünün: H_0'yi yargılıyorsunuz. p-değeri, H_0'nin masum olması durumunda kanıtınızın ne kadar şaşırtıcı olacağını ölçer. Küçük p-değeri = çok şaşırtıcı = H_0 suçlu görünüyor.

Alfa'yı (α) Anlamak: Önem Düzeyi

Artık p-değeri anlaşıldığına göre, şimdi α (alfa) hakkında konuşalım.

Alfa Nedir? Alfa (α), H_0'yi reddetmeden önce p değerinin ne kadar küçük olması gerektiğine dair **önceden belirlenen eşiktir.**

Yaygın tercihler:

- $\alpha = 0,05$ (%5) - en yaygın
- $\alpha = 0,01$ (%1) - daha risksiz
- $\alpha = 0,10$ (%10) - daha liberal

Karar Kuralı

p-değeri ≤ α ise: H_0'yi reddedin (kanıt yeterince güçlüdür) **p-değeri > α ise:** H_0'yi reddetmeyin (kanıt yeterince güçlü değildir)

Madeni Para Örneğimize Devam Edelim

p değeri 0,0207 (2,07%) olarak hesaplandı.

α = 0,05 olarak belirlenmişse:

- p-değeri (0.0207) < α (0.05)
- Karar: H_0'yi Reddet
- Sonuç: "Madeni paranın yazı gelme eğiliminde olduğu sonucuna varmak için yeterli kanıt var"

α = 0,01 olarak belirlenmişse:

- p-değeri (0.0207) > α (0.01)
- Karar: H_0 reddedilemedi
- Sonuç: "Madeni paranın taraflı olduğu sonucuna varmak için yeterli kanıt yok"

Dikkat: Aynı veriler, aynı p değeri, α'ya bağlı olarak farklı sonuçlar!

Neden α = 0,05?

İşte bir sır: **Bu biraz keyfi!**

α = 0,05 kullanma geleneği 1920'lerde Ronald Fisher'dan geliyor. %5'in makul bir eşik gibi göründüğünü söylemiş. Hepsi bu. Yani, %5'te büyülü bir şey yok.

α'nın Gerçek Anlamı:

Alfa, yanlış alarmlara (Tip I hataları - bu konu hakkında daha fazla bilgi bir sonraki derste anlatılacaktır) karşı toleransınızdır.

- $\alpha = 0,05$ şu anlama gelir: "Yanılma olasılığım %5 olduğunda H_0'yi (bir şeyin doğru olduğunu iddia etmeyi) reddetmeye hazırım"
- $\alpha = 0,01$ şu anlama gelir: "Daha temkinliyim - sadece %1 yanılma olasılığı istiyorum"
- $\alpha = 0,10$ şu anlama gelir: "Daha hoşgörülüyüm - %10 yanılma olasılığını kabul ederim"

**Alfa Seçimi"

Aşağıdaki durumlarda $\alpha = 0,01$ (daha katı) değerini kullanın:

- Yanlış pozitifler çok maliyetliyse
- Güçlü kanıtlara ihtiyaç duyan iddialı bir iddiada bulunuluyorsa
- Örnek: Bir ilacın kanseri iyileştirdiğini iddia etmek - çok emin olunması gerekir!

Aşağıdaki durumlarda $\alpha = 0,05$ (standart) değerini kullanın:

- Orta düzeyde sonuçlar
- Akademik araştırmaların çoğu

Aşağıdaki durumlarda $\alpha = 0,10$ (daha esnek) değerini kullanın:

- Keşifsel araştırma
- Ön çalışmalar
- Gerçek bir etkiyi gözden kaçırmak, yanlış pozitif bir sonuçtan daha maliyetlidir

Kritik Bir Nokta: Veri Toplamadan ÖNCE Alfa'yı Ayarlayın

Kötü uygulama (p-hacking):

1. Veri topla
2. p-değerini hesapla
3. Sonucu anlamlı kılmak için α'yı seç
4. "Bak, $p = 0,06$... yayınlayabilmek için $\alpha = 0,10$ kullanalım!"

İyi uygulama :

1. Alanınızın standartlarına ve sonuçlarına göre α'yı seçin
2. Veri toplayın
3. P-değerini hesaplayın
4. "Önceden belirlediğiniz α ile karşılaştırın
5. Kararınızı verin

P-Değeri ile Alfa Arasındaki İlişki:

"H_0'ye Karşı Kanıt" Sayı Doğrusu varsayalım

No description has been provided for this image

p-değeriniz bu çizginin bir yerindedir.

p-değeriniz α'nın SOL'undaysa (daha küçükse): Kanıt yeterince güçlüdür → H_0'yi reddedin
p-değeriniz α'nın SAĞ'ındaysa (daha büyükse): Kanıt yeterince güçlü değildir → H_0'yi reddedin

Örnek Senaryolar

Senaryo 1: p = 0,001, α = 0,05

No description has been provided for this image

p değeri α'nın çok solunda → Çok güçlü kanıt → H_0'yi kesinlikle reddet

Senaryo 2: p = 0.04, α = 0.05

No description has been provided for this image

p değeri α'nın hemen solunda → Çok az önemli → H_0'yi reddet (ama dikkatli ol)

Senaryo 3: p = 0.06, α = 0.05

No description has been provided for this image

p-değeri α'nın tam sağında → Tam olarak anlamlı değil → H_0'yi reddedemedi (ama yakın!)

Senaryo 4: p = 0.42, α = 0.05

No description has been provided for this image

p değeri α'nın en sağında → Çok zayıf kanıt → H_0'yi reddetmede açıkça başarısız

Önemli Bir Bakış Açısı: Parlak Çizgi Sorunu

Yukarıdaki 2. ve 3. senaryolara dikkat edin: p = 0,04 - p = 0,06.

Biri "önemli", diğeri "önemsiz". Ama pratikte neredeyse hiç fark yok! Her ikisi de benzer kanıt seviyeleri gösteriyor.

Ders: α'yı "gerçek" ile "yalan" arasındaki ince bir çizgi olarak görmeyin. Bu, sürekli bir kanıt ölçeğidir.

Daha iyi yorumlama:

- p = 0,001: H_0'ye karşı çok güçlü kanıt
- p = 0,02: H_0'ye karşı güçlü kanıt
- p = 0,06: H_0'ye karşı orta düzeyde kanıt
- p = 0,15: H_0'ye karşı zayıf kanıt
- p = 0,40: H_0'ye karşı çok zayıf kanıt

Sadece "önemli" ve "önemli değil" demekten daha iyi.

Test İstatistikleri ve Kritik Değerler: Z-puanlarını, T-puanlarını vb. Anlamak

Şimdi bu test istatistiklerinin (z, t, χ^2, F) aslında neyi temsil ettiğini ve p-değerleri ve α ile nasıl ilişkili olduğunu anlamaya çalışalım.

Test İstatistiği Nedir?

Bir test istatistiği, örnek verilerinizin H_0 altında beklediğinizden ne kadar uzakta olduğunu özetleyen tek bir sayıdır.

Bunu şöyle düşünün: "Sonucum beklenenden kaç 'standart birim' uzakta?"

Z-Puanını Derinlemesine Anlamak

Senaryo: IQ testleri şu şekilde tasarlanmıştır:

- Ortalama (μ) = 100
- Standart sapma (σ) = 15

Bir kişiyi test ediyorsunuz ve 130 puan alıyor. Bu ne kadar sıra dışı?

Z-puanının hesaplanması:

$z = (X - \mu) / \sigma$ $z = (130 - 100) / 15$ $z = 30 / 15$ $z=2$

z = 2'nin anlamı: "Bu kişinin IQ'su ortalamanın 2 standart sapma üstündedir."

Z Dağılımı (Standart Normal Dağılım)

Z dağılımı şu şekilde bir çan eğrisidir:

- Ortalama = 0
- Standart sapma = 1

Her z-puanı bu standart eğri üzerinde nerede olduğunuzu söyler.

- Verilerin yaklaşık %68'i z = -1 ile +1 arasında değişmektedir.
- Verilerin yaklaşık %95'i z = -2 ile +2 arasında değişmektedir.
- Verilerin yaklaşık %99,7'si z = -3 ile +3 arasında değişmektedir.

Yani z = 2 ise, tipik %95 aralığının sınırındasınız. İnsanların yalnızca yaklaşık %2,5'i bu kadar yüksek veya daha yüksek puan alıyor.

Z-Puanından P-Değerine

Sağ kuyruklu bir test için (H_a: $\mu > \mu_0$):

- Z-puanını hesapla
- p-değeri = z'yi bu kadar büyük veya daha büyük alma olasılığı
- Z-tablosuna bakın veya yazılım kullanın

Örnek: Sağ kuyruklu bir testte z = 2.0 ise:

- p-value ≈ 0.0228 (2.28%)
- Bu, z = 2'nin sağındaki eğrinin altındaki alandır

Alfa'dan Kritik Değere

Kritik değer, seçilen α değerine karşılık gelen test istatistiği değeridir.

α = 0,05 için sağ kuyruklu test:

- Kritik z değeri = 1.645
- Bu, sağında alanın tam %5'i olan z puanıdır

α = 0,05 için iki kuyruklu test:

- Kritik z değerleri = ±1,96
- Bunlar her kuyrukta %2,5'e sahip olan z puanlarıdır (toplam %5)

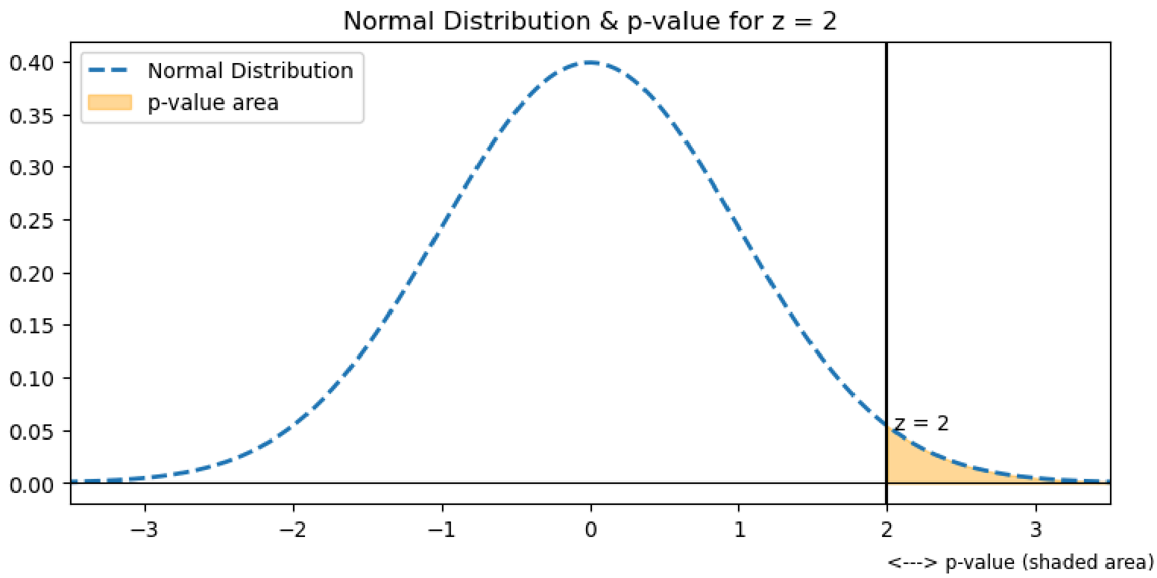
Kritik değerleri kullanan karar kuralı:

- Hesaplanan z > kritik z ise: H_0'yi reddet
- Hesaplanan z < kritik z ise: H_0'ı reddedemezsin

In [5]:

```
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import norm

# x ekseni
x = np.linspace(-4, 4, 400)

# Normal dağılım
y = norm.pdf(x)

# z-değeri
z_value = 2

plt.figure(figsize=(8, 4))

# Normal dağılım eğrisi
plt.plot(x, y, linestyle="--", linewidth=2, label="Normal Distribution")

# p-value alanını renklendir (z'nin sağ tarafı)
x_fill = np.linspace(z_value, 4, 200)
plt.fill_between(x_fill, norm.pdf(x_fill), color="orange", alpha=0.4, label="p-v

# z = 2 çizgisi
plt.axvline(z_value, color="black", linestyle="-", linewidth=1.5)
plt.text(z_value + 0.05, 0.05, "z = 2", fontsize=10)

# Eksen ve etiketler
plt.axhline(0, color="black", linewidth=0.8)
plt.xlim(-3.5, 3.5)
plt.xticks([-3, -2, -1, 0, 1, 2, 3])

plt.title("Normal Distribution & p-value for z = 2")

plt.legend()

# Alt açıklama
plt.text(2, -0.08, "<---> p-value (shaded area)", fontsize=9)
```

```
plt.tight_layout()
plt.show()
```



Aynı Kararı Vermenin İki Yolu

Yöntem 1: p değerini α ile karşılaştırın

- Test istatistiğini hesapla → p değerini al → α ile karşılaştır
- Modern yaklaşım, yorumlaması daha kolay

Yöntem 2: Test istatistiğini kritik değerle karşılaştırın

- Test istatistiğini hesaplayın → Kritik değerle karşılaştırın
- Geleneksel yaklaşım, hala yaygın

İkisi de aynı! İkisi de aynı sonuca varıyor.

Örnek:

Veriler: ● z = 2,0 ● α = 0,05 ● Sağ kuyruk testi

Yöntem 1 (p-değeri yaklaşımı): ● p-değeri = 0,0228 ● 0,0228 < 0,05 ● H_0'yi reddet

Yöntem 2 (kritik değer yaklaşımı): ● Kritik değer = 1,645 ● 2,0 > 1,645 ● H_0'yi reddet

T-İstatistik ve T-Dağılımını Anlamak

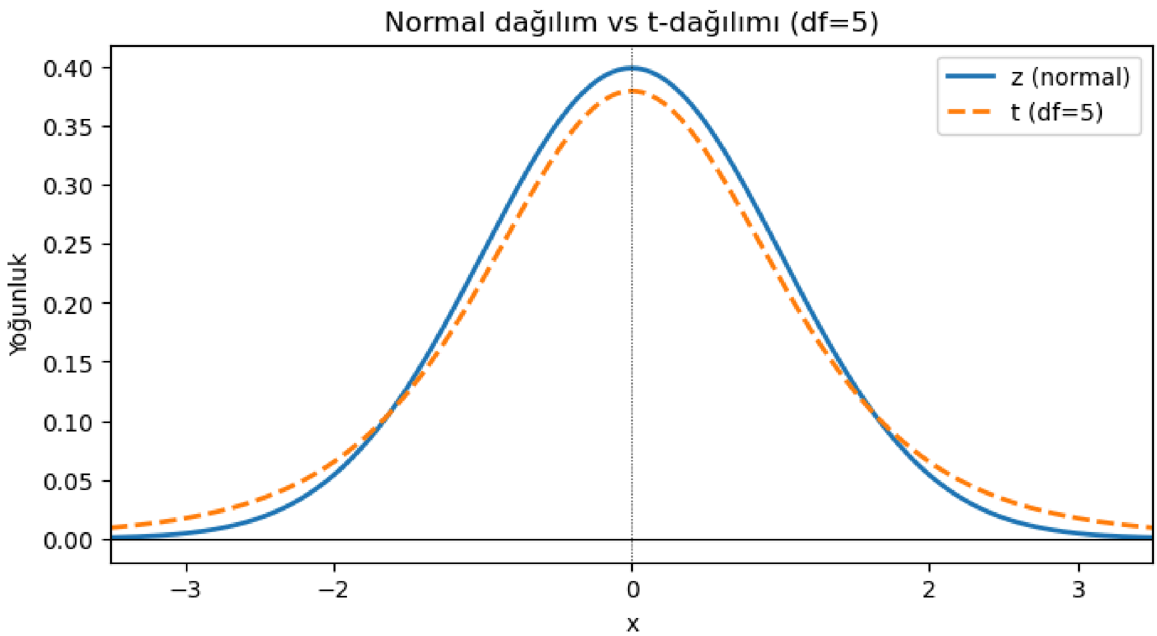
t istatistiği z ile aynı şekilde çalışır, ancak σ'yu bilmediğinizde ek belirsizlik ortaya çıkar.

T-Dağılım Ailesi

Z-dağılımının aksine (ki her zaman aynıdır), her serbestlik derecesi için farklı bir t-dağılımı vardır.

In [6]:
```
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import norm, t
```

```python
# x ekseni
x = np.linspace(-4, 4, 400)

# Dağılımlar
y_norm = norm.pdf(x, loc=0, scale=1)   # z ~ N(0,1)
df = 5
y_t = t.pdf(x, df=df)                  # t(df=5)

plt.figure(figsize=(8, 4))

# Eğriler
plt.plot(x, y_norm, label="z (normal)", linewidth=2)
plt.plot(x, y_t, linestyle="--", label=f"t (df={df})", linewidth=2)

# Eksen ve arka plan ayarları
plt.axhline(0, color="black", linewidth=0.8)
plt.axvline(0, color="black", linewidth=0.5, linestyle=":")

plt.xlim(-3.5, 3.5)
plt.xticks([-3, -2, 0, 2, 3])

plt.xlabel("x")
plt.ylabel("Yoğunluk")
plt.title("Normal dağılım vs t-dağılımı (df=5)")

plt.legend()

plt.show()
```



T dağılımı normal dağılıma göre daha kalın kuyruklara sahiptir. df arttıkça, t dağılımı → normal dağılım.

T-Puanı Yorumlanması Z-puanı gibi t-puanı da "beklenenden kaç standart hata uzakta olduğunuzu" söyler.

Örnek: t = 2.5 ve df = 10

Bunun anlamı: "Örneklem ortalaması, varsayılan ortalamanın 2,5 standart hata üzerindedir ve bu, standart sapma tahminimizdeki belirsizliği hesaba katar."

p-değerini elde etmek: t-tablosunda df = 10'a bakılır veya bir yazılım kullanılır.

df = 10 için sağ kuyruklu test:

● α = 0,05'teki kritik t değeri yaklaşık 1,812'dir ● 2,5 > 1,812 olduğundan → H_0'yi reddet

Serbestlik Dereceleri (df):

Bu, yeni başlayanlar için en gizemli kavramlardan biridir.

Serbestlik derecesi = Bir istatistiğin hesaplanmasında kullanılan bağımsız bilgi parçalarının sayısı.

Ya da daha basiti: **Kaç değer "serbestçe değişebilir"?**

Örnekler

Örnek 1: Örnek Ortalaması (Tek Örneklem T-Testi)

5 sayınız var. Size şunu söyleyeyim: ● Ortalama 10 ● Sayılardan dördü: 8, 9, 11, 12

Soru: Beşinci sayı nedir?

Çözüm: Ortalama 10 ise, toplam 5 × 10 = 50 olmalıdır. Mevcut toplam = 8 + 9 + 11 + 12 = 40 Beşinci sayı = 50 - 40 = 10 olmalıdır.

Önemli nokta: Ortalama ve 4 sayı bilindiğin, 5. sayının değişmesi mümkün değildir - belirlenmiştir!

Serbestlik Derecesi = n - 1 = 5 - 1 = 4

Sadece 4 değer değişebilir. 5.si kısıtlanmıştır.

Örnek 2: İki Örneklemli T Testi

Grup 1: n_1 = 20 kişi Grup 2: n_2 = 25 kişi

İki örneklem ortalaması hesaplarsınız (her grup için bir tane).

df = (n_1 - 1) + (n_2 - 1) = 19 + 24 = 43 Veya eşdeğer olarak: df = n_1 + n_2 - 2 = 20 + 25 - 2 = 43

Neden 2 çıkarılır? Çünkü 2 parametre (iki ortalama) tahmin edildi ve her biri 1 serbestlik derecesine mal oldu.

Örnek 3: Eşleştirilmiş T-Testi

15 çift ölçümü var (15 kişide öncesi/sonrası). 15 fark hesaplanıyor. Ardından bu farkların ortalamasını hesaplanıyor.

df = n - 1 = 15 - 1 = 14

Tıpkı tek örneklemli bir test gibi! Çünkü farklılıklar "örneklem" olarak ele alınıyor.

Örnek 4: Ki-Kare Testi

Uyum iyiliği testi: 6 kategoriniz var (zar yüzleri gibi).

df = (number of categories) - 1 = 6 - 1 = 5

Neden? Toplam sayı ve 5 kategori sayısı biliniyorsa, 6. belirlenir.

Bağımsızlık testi (2×3 tablo): 2 satır, 3 sütun

df = (rows - 1) × (columns - 1) = (2-1) × (3-1) = 1 × 2 = 2

Neden? Satır toplamları, sütun toplamları ve (satır-1)×(sütun-1) hücre verildiğinde diğer tüm hücreler belirlenir.

Örnek 5: ANOVA

Toplam N = 40 gözlem içeren k = 4 grup karşılaştırıldığında.

Gruplar arası df: df_1 = k - 1 = 4 - 1 = 3 (4 grup ortalamasını tahmin ediyorsunuz, ancak bunlar genel ortalama ile sınırlandırılmıştır)

Grup içi df: df_2 = N - k = 40 - 4 = 36 (Her grup ortalamasından 1 df kaybeder)

Serbestlik Derecesinin Önemi

1. df değeri, istatistiksel testte kullandığın dağılımın şeklini belirler.
   A. t-dağılımı df'ye bağlı olarak şekil değiştirir.
   B. df küçükse (ör. 5), dağılım daha yayvan ve kuyrukları daha şişmandır.
   C. df büyüdükçe (örn. 30), dağılım normal dağılıma yaklaşır. ⇒ df → ∞ → t-dağılımı = normal dağılım gibi olur.
2. df, kritik değerleri değiştirir → bu da H_0'ı reddetmeyi etkiler.
   Örneğin α = 0.05, iki yönlü test için:
   A. z-kritik (df = ∞) = 1,96
   B. t-kritik (df = 10) = 2,228
   C. t-kritik (df = 5) = 2,571
   D. t-kritik (df = 2) = 4,303
   👉 df azaldıkça kritik değer büyür

   👉 Kritik değer büyüdükçe H_0'ı reddetmek daha zorlaşır

   NOT:Daha küçük bir df → daha fazla belirsizlik → daha geniş/şişman kuyruk → daha ekstrem bir test istatistiği gerektirir.
3. df örneklem büyüklüğünü ve model karmaşıklığını yansıtır
   A. Daha büyük sample size → daha büyük df → daha hassas (precise) tahmin
   B. Modelde daha fazla parametre tahmin edildikçe → df azalır → tahminler daha az güvenilir
   Yani df hem kaç gözlemin var hem de modelde kaç parametre kullandığın konusunda bilgi taşır.

Felsefi Anlam Serbestlik dereceleri bilgiyi temsil eder.

Verilerinizden bir parametreyi (örneğin bir ortalama) her tahmin ettiğinizde, bilgilerinizin bir kısmını "tüketirsiniz". Geriye kalanlar ise hipotezleri test etmek için kullanılabilir.

Bunu para gibi düşünün: ● n dolarla (n veri noktası) başlıyorsunuz. ● Tahmin ettiğiniz her parametrenin maliyeti 1 dolardır. ● Geriye kalan, çıkarımlar yapmak için gereken "satın alma gücünüz" olan df'nizdir.

Güven Aralıkları: Madalyonun Diğer Yüzü

Şimdi güven aralıklarından ve bunların hipotez testleriyle ilişkisinden bahsedelim.

Güven Aralığı Nedir?

Bir güven aralığı, bir popülasyon parametresi için makul değerler aralığı verir.

Genel biçimi:

CI = (point estimate) ± (critical value) × (standard error)

Ortalama İçin:

$$CI = \bar{x} \pm t(\alpha/2,\, df) \times \left(\frac{s}{\sqrt{n}}\right)$$

"%95 Güven"i Anlamak

Yaygın Yanlış Yorumlama: ❌ "Gerçek ortalamanın bu aralıkta olma olasılığı %95'tir."

Doğru Yorumlama: ✓ "Bu işlemi birçok kez tekrarlarsak, oluşturduğumuz güven aralıklarının yaklaşık %95'i gerçek popülasyon ortalamasını içerecektir."

Frekansçı Yorum

Diyelim ki:

1. Örnek alın
2. %95 GA hesaplayın
3. Bunu 100 kez tekrarlayın

Sonuç: Aralıklarınızın yaklaşık %95'i gerçek μ'yi içerecek, yaklaşık %5'i ise içermeyecektir.

Visual:

Gerçek μ = 50

Örnek 1: [48 ══════════ 52] ✓ μ İçerir Örnek 2: [47 ══════════ 51] ✓ μ İçerir Örnek 3: [49 ══════════ 53] ✓ μ İçerir Örnek 4: [51 ══════════ 55] ✗ μ'yi kaçırır (zamanın yaklaşık %5'inde gerçekleşir) Örnek 5: [46 ══════════ 50] ✓ μ İçerir

Örnek: Güven Aralığının Hesaplanması

Senaryo: 25 kişinin boyunu ölçüyorsunuz.

● Sample mean: $\bar{x}$ = 170 cm ● Sample SD: s = 10 cm ● Sample size: n = 25

İstenen: Gerçek ortalama yükseklik μ için %95 güven aralığı.

Adım 1: Kritik t değerini bulun

● df = n - 1 = 24 ● %95 CI için α = 0,05, dolayısıyla α/2 = 0,025 (iki kuyruklu) ● t-tablosundan: t(0,025, 24) = 2,064

Adım 2: Standart hatayı hesaplayın

SE = s / √n = 10 / √25 = 10 / 5 = 2

Adım 3: Hata payını hesaplayın ME = t × SE = 2.064 × 2 = 4.128

Adım 4: Aralığı oluşturun

CI = $\bar{x}$ ± ME CI = 170 ± 4.128 CI = (165.87, 174.13)

Yorum: "Nüfusun gerçek ortalama boyunun 165,87 cm ile 174,13 cm arasında olduğundan %95 eminiz."

Derin Bağlantı: Güven Aralıkları ve Hipotez Testleri

İşte asıl ilginç nokta burada başlıyor: Güven aralıkları ve hipotez testleri aynı şeyi söylemenin iki farklı yoludur!

Eşdeğerlik

α önem düzeyinde iki taraflı bir test için:

Hipotez Testi Yaklaşımı:

- H_0: $\mu = \mu_0$
- Test istatistiğini hesapla
- p-değerini al
- Eğer $p < \alpha$ ise, H_0'yi reddedin

Güven Aralığı Yaklaşımı:

- (1-α)% güven aralığını hesaplayın
- μ_0 aralıkta DEĞİLSE, H_0'yi reddedin
- μ_0 aralıktaysa, H_0'yi reddetmeyi başaramaz.

Aynı cevabı veriyorlar!

Örnek: Eşdeğerliği Görmek

Senaryo: μ = 165 cm olup olmadığını test edin. Daha önce, %95 CI'mız = (165,87, 174,13) idi.

Hipotez Testi:/br>

- H_0: $\mu = 165$
- H_a: $\mu \neq 165$
- $\alpha = 0.05$

t'yi hesaplayın:

$t = (\bar{x} - \mu_0) / (s/\sqrt{n})$

$t = (170 - 165) / 2$ $t = 2.5$

$\alpha = 0{,}05$, df = 24'teki kritik t değeri: ±2,064 2,5 > 2,064 olduğundan → H_0'yi reddet

Güven Aralığı Yöntemi: %95 GA'mız (165.87, 174.13)'tür. Bu aralıkta $\mu_0 = 165$ mi? HAYIR! (165 < 165.87) → H_0'yi Reddet

Aynı sonuç!

Görsel Anlama

No description has been provided for this image

Güven Aralıkları Neden Kullanılır?

CI'lerin hipotez testlerine göre avantajları:

1. **Daha bilgilendirici**
   A. Hipotez testi: "Evet, bir fark var" (ikili cevap)
   B. CI: "Fark muhtemelen 5 ile 15 birim arasındadır" (nicel cevap)
2. **Hassasiyet gösterir**
   A. Dar GA: Kesin tahmin
   B. Kesin olmayan tahmin (daha fazla veriye ihtiyaç var)
3. **Pratik önemi**
   A. Etki büyüklüğünün anlamlı olup olmadığını görebilirsiniz
   B. Örnek: Maaş farkı için CI = $(1, 5)$ → istatistiksel olarak anlamlı ancak pratikte küçük bir fark

Örnek Karşılaştırma:

Çalışma 1:

● Örneklem büyüklüğü: n = 1000 ● Ortalama iyileşme için %95 GA: (2.1, 2.3) gün ● Yorum: Çok kesin bir tahmin. İyileşme kesinlikle 2.2 gün civarındadır.

Çalışma 2:

● Örneklem büyüklüğü: n = 20 ● Ortalama iyileşme için %95 GA: (0.5, 8.7) gün ● Yorum: Çok belirsiz. Küçük veya büyük bir iyileşme olabilir. Daha fazla veriye ihtiyacımız var!

Hepsini Bir Araya Getirelim:

Tam Hipotez Test Süreci

Senaryo: Bir kahve şirketi, poşetlerinin 500 gr kahve içerdiğini iddia ediyor. Siz de poşetlerin yetersiz olduğundan şüpheleniyorsunuz. 16 poşet seçtiniz.

Veriler:

- n = 16 poşet
- $\bar{x}$ = 495g (sample mean)
- s = 12g (örnek SD)

Adım 1: Hipotezleri belirtin ve α'yı seçin

- H_0: μ ≥ 500 g (eksik doldurma değil)
- H_a: μ < 500 g (eksik doldurma)
- α = 0,05 (önceden seçilmiştir)
- Bu sol kuyruklu bir testtir

Adım 2: Test istatistiğini hesaplayın

t = ($\bar{x}$ - μ_0) / (s/√n) t = (495 - 500) / (12/√16) t = -5 / (12/4) t = -5 / 3 t = -1.667

Bu t-puanı ne anlama geliyor: "Örneklem ortalaması, iddia edilen ortalamanın 1,667 standart hata altındadır."

Adım 3: Serbestlik derecelerini belirleyin

df = n - 1 = 16 - 1 = 15

Adım 4: Kritik değeri bulun (geleneksel yaklaşım)

Sol kuyruklu test için α = 0,05, df = 15: Kritik t değeri = -1,753

Kritik değer kullanılarak karar: t = -1,667, -1,753'ten daha mı uç değerde? Hayır! -1,667 > -1,753 (sıfıra daha yakın) → **H_0 reddedilemedi**

Adım 5: p-değerini bulun (modern yaklaşım)

t = -1,667, df = 15 için t-tablosu veya yazılım kullanıldığında: p-değeri ≈ 0,058

p-değeri kullanılarak karar: p = 0,058 < α = 0,05 mi? Hayır! 0,058 > 0,05 → **H_0 reddedilemedi**

Her iki şekilde de aynı sonuç!

Adım 6: Güven aralığını hesaplayın

df = 15 olan %95 CI için: t-kritik (iki kuyruklu) = 2,131

$$CI = \bar{x} \pm t \times \left(\frac{s}{\sqrt{n}}\right)$$

CI = 495 ± 2.131 × 3 CI = 495 ± 6.393 CI = (488.61, 501.39)

CI'nin Yorumlanması: "Gerçek ortalama ağırlığın 488,61 g ile 501,39 g arasında olduğundan %95 eminiz."

Dikkat: Bu aralık 500g'ı kapsamaktadır! Bu, H_0'yi reddetmeme kararımızı doğrulamaktadır.

Adım 7: Sonucu tamamlayın

"16 poşetlik örneğimize dayanarak (ortalama = 495 g, SD = 12 g), poşetlerin 500 g'ın altında yetersiz doldurulduğu sonucuna varmak için α = 0,05 düzeyinde yeterli kanıtımız yok (t = -1,667, df = 15, p = 0,058). %95 güven aralığı (488,61 g, 501,39 g), iddia edilen 500 g değerini de içermektedir. Örneklem ortalaması 500 g'ın altında olsa da, gerçek ortalama 500 g ise bu farkın tesadüfen ortaya çıkması makul bir olasılıktır."

Tüm Kavramlar Nasıl İlişkilendirilir: Birleşik Bir Çerçeve

```
HİPOTEZ TESTİ
                               |
               --------------------------------
               |                              |
             SORU                           CEVAP
     "Bir etki/fark var mı?"      Kanıtlara göre, muhtemelen
                                         evet veya hayır
               |                              |
       -----------------            -----------------------
       |               |            |                     |
 HİPOTEZLER        ALFA (α)     TEST İSTATİSTİĞİ     GÜVEN ARALIĞI
 H0 vs H1        Karar eşiği    Kanıtı özetler   Olası değerlerin
aralığı
                       |               |                     |
                       |       ----------------        Hipotez
testine
                       |       |              |         eşdeğer bilgi
                       |      Z, T, χ², F  SERBESTLİK
                       |  "Kaç sd uzakta?"   DERECESİ
                       |                     |
                       |              Hassasiyeti etkiler
                       |                     |
                       |                 P-DEĞERİ
                       |       "H0 doğruysa bu ne kadar sıra dışı?"
                       |                     |
                       |----------------------|
                               |
                             KARAR
                  H0 reddet veya H0 reddedilemez
```

Mantık Akışı

1. Bir popülasyon parametresi hakkında bir soruyla başlayın
2. Hipotezler olarak çerçeveleyin (H_0 ve H_a)
3. Ne kadar muhafazakar olmak istediğinize göre α'yı seçin
4. Veri toplayın ve örneklem istatistiklerini hesaplayın

5. Test istatistiğini hesaplayın (z, t vb.) - verilerinizin H_0'den ne kadar uzakta olduğunu özetler
6. Serbestlik derecelerini kullanarak belirsizliği hesaba katın
7. p-değerini hesaplayın - H_0 doğruysa verileriniz ne kadar sıra dışıdır?
8. p-değerini α ile karşılaştırın VEYA test istatistiğini kritik değerle karşılaştırın
9. Karar verin: H_0'yi reddedin veya reddetmeyin
10. Güven aralığı olarak da ifade edilebilir - makul değerler aralığı verir

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]: