Ahmet Burak Şişci

bu koınuyu hocanın dökümanınından çalışmak daha iyi olur. bunu özet gibi düşünün

hipotez testi 2 dokumanı

parametrik ve non parametrik testeler

önemli olan verinin normal deağılıp dağılmadığıdır. normal dağılıyorsa parametrik testler uygulanır.

Bu testler, istatistikte bir veya daha fazla grup arasında anlamlı bir fark olup olmadığını anlamak için kullanılan hipotez testleridir. Hangi test türünü seçeceğiniz, büyük ölçüde verilerinizin özelliklerine ve dağılımına bağlıdır.

Parametrik Testler

Bu testler, verilerinizin belirli varsayımları karşıladığını kabul eder. En temel varsayım, verilerin **normal dağılıma** (çan eğrisi şeklinde bir dağılım) sahip olmasıdır.

- **Varsayımları:**
    - Veriler normal dağılım göstermelidir.
    - Grupların varyansları (verilerin ortalama etrafındaki yayılımı) birbirine eşit (homojen) olmalıdır.
    - Veriler aralıklı veya oransal ölçekte ölçülmüş olmalıdır.
    - Gözlemler birbirinden bağımsız olmalıdır.
- **Ne Zaman Kullanılır?:** Veri setiniz yukarıdaki varsayımları karşılıyorsa ve genellikle örneklem büyüklüğünüz yeterince büyükse (genellikle n > 30 bir kılavuz olarak kabul edilir) kullanılır.
- **Özellikleri:** Varsayımlar karşılandığında non-parametrik testlere göre daha **güçlü** ve hassas sonuçlar verirler. Yani, gruplar arasında gerçek bir fark varsa, bu farkı tespit etme olasılıkları daha yüksektir. Analizlerde ortalama, standart sapma gibi değerler kullanılır.
- **Popüler Örnekler:** t-Testi, ANOVA (Varyans Analizi), Pearson Korelasyonu.

Non-Parametrik (Parametrik Olmayan) Testler

Bu testler, verinin belirli bir dağılıma sahip olduğu varsayımını yapmazlar. Bu nedenle "dağılımdan bağımsız" testler olarak da bilinirler.

- **Varsayımları:** Çok daha az ve esnek varsayımları vardır. Verinin normal dağılması veya varyansların eşit olması gibi katı kurallar gerektirmezler.
- **Ne Zaman Kullanılır?:**
  - Veri setiniz normal dağılım göstermiyorsa.
  - Örneklem büyüklüğünüz küçükse (genellikle n < 30).
  - Verileriniz sıralı (ordinal) veya nominal (kategorik) ise.
  - Veri setinizde aykırı değerler (outliers) varsa.
- **Özellikleri:** Parametrik testlerin varsayımları karşılanmadığında güvenilir bir alternatif sunarlar. Analizlerde ortalama yerine medyan (ortanca değer) ve sıralamalar kullanılır. Genellikle parametrik testlere göre daha az güçlü kabul edilirler.
- **Popüler Örnekler:** Mann-Whitney U Testi (Bağımsız t-Testi'nin alternatifi), Wilcoxon İşaretli Sıralar Testi (Bağımlı t-Testi'nin alternatifi), Kruskal-Wallis Testi (ANOVA'nın alternatifi), Spearman Korelasyonu.

Özetle

Özellik	Parametrik Testler	Non-Parametrik Testler
Ana Koşul	Veriler normal dağılmalıdır.	Verilerin dağılımı önemli değildir.
Veri Türü	Genellikle aralıklı veya oransal.	Nominal, sıralı veya sürekli olabilir.
Merkezi Eğilim	Ortalama (Mean) kullanılır.	Medyan (Ortanca) kullanılır.
Güç	Daha güçlüdür (farkı bulma olasılığı yüksek).	Daha az güçlüdür.
Kullanım Durumu	Büyük örneklem, normal dağılım.	Küçük örneklem, normal olmayan dağılım, aykırı değerler.
Örnek Testler	t-Testi, ANOVA	Mann-Whitney U, Kruskal-Wallis

Pratik bir kural olarak: Önce verilerinizin parametrik test varsayımlarını karşılayıp karşılamadığını kontrol edin. Eğer karşılıyorsa parametrik bir test kullanın çünkü daha güçlü sonuçlar verir. Eğer karşılamıyorsa, o zaman non-parametrik bir test kullanmak daha doğru ve güvenilir olacaktır.

Senaryo: Yeni Bir Gübre Etkili mi?

Bir ziraat mühendisi, domates verimini artırdığı iddia edilen yeni bir organik gübre geliştirdi. Bu gübrenin gerçekten işe yarayıp yaramadığını test etmek istiyor.

Deney Tasarımı:

Mühendis, özellik ve büyüklük olarak birbirine çok benzer 20 adet domates fidesi alıyor. Bu fideleri rastgele iki gruba ayırıyor:

- **A Grubu (Kontrol Grubu - 10 fide):** Bu fideler standart gübre ile yetiştirilecek.

- **B Grubu (Deney Grubu - 10 fide):** Bu fideler yeni geliştirilen organik gübre ile yetiştirilecek.

Hasat zamanı geldiğinde, her bir fideden elde edilen domateslerin toplam ağırlığını (kilogram cinsinden) kaydediyor. Amacı, iki grup arasında verim açısından istatistiksel olarak anlamlı bir fark olup olmadığını bulmak.

Veri Analizi Adımı:

Mühendis, verileri analiz etmek için bir istatistik programı kullanır. İlk yapması gereken, hangi istatistiksel testi kullanacağına karar vermektir.

1. Parametrik Yaklaşım (İdeal Durum):

- **Varsayım Kontrolü:** Mühendis önce her iki grubun verilerinin de **normal dağılıp dağılmadığını** test eder (örneğin, Shapiro-Wilk testi ile). Ayrıca, grupların varyanslarının benzer olup olmadığını kontrol eder.
- **Sonuç:** Test sonuçları, her iki grubun verilerinin de normal dağılıma uyduğunu ve varyansların benzer olduğunu gösterir. Örneklem boyutu küçük (n=10) olmasına rağmen, normal dağılım varsayımı sağlandığı için parametrik test kullanmaya karar verir.
- **Kullanılacak Test: Bağımsız Örneklem t-Testi (Independent Samples t-Test)**. Bu test, iki bağımsız grubun ortalamaları arasında anlamlı bir fark olup olmadığını ölçer.
- **Hedef:** Yeni gübre kullanılan B grubunun ortalama veriminin, standart gübre kullanılan A grubunun ortalama veriminden anlamlı derecede yüksek olup olmadığını bulmak.

2. Non-Parametrik Yaklaşım (Gerçekçi / Sorunlu Durum):

- **Varsayım Kontrolü:** Mühendis, normal dağılım testini yapar. Ancak bir sorunla karşılaşır: Deney grubundaki fidelerden biri hastalanmış ve neredeyse hiç verim vermemiştir. Bu durum, verilerde bir **aykırı değer (outlier)** yaratır ve verilerin normal dağılımını bozar.
- **Sonuç:** Normal dağılım testi, verilerin normal dağılmadığını gösterir. Bu durumda parametrik test varsayımları karşılanmamaktadır.
- **Kullanılacak Test: Mann-Whitney U Testi**. Bu test, t-Testi'nin non-parametrik alternatifidir. Verilerin kendisi yerine, verilerin sıralamasını (en düşük verimden en yükseğe doğru) karşılaştırır. Bu sayede aykırı değerin ezici etkisinden kurtulmuş olur.
- **Hedef:** Yeni gübre kullanılan B grubunun verim sıralamalarının medyanının (ortanca değer), A grubunun verim sıralamalarının medyanından anlamlı derecede yüksek olup olmadığını bulmak.

Özetle

- Eğer domates verimleri her iki grupta da beklendiği gibi, aşırı uç değerler olmadan dağılsaydı, mühendis **t-Testi (parametrik)** kullanarak daha güçlü bir analiz yapabilirdi.

- Ancak hastalanan fide gibi beklenmedik bir durum veriyi bozduğu için, varsayımları ihlal etmemek ve güvenilir bir sonuç almak adına **Mann-Whitney U Testi (non-parametrik)** kullanması daha doğru bir yaklaşım oldu.

Bu senaryo, veri analizinde neden tek bir doğru olmadığını ve verinin özelliklerine göre doğru testin seçilmesinin ne kadar kritik olduğunu göstermektedir.

kategorik verilerde sayılar sadecee sırayı temsil ettiği için nonparametrik uygulama yapmak daha matıklıdır. (önemli)

Test Seçim Çerçevesi

Z-Testi

- popülasyon standart sapması bilindiğinde kullanılır. (gerçek hayatta bunu bilemeyiz.)
- elimizdeki veri büyüdükçe standart sapmayı daha iyi tahmin ederiz.
- elimizdeki veri 30 dan büyük ise z-testi kullanılabilir

Z-Testi Nedir?

Z-Testi, elinizdeki bir grubun (örneklem) ortalamasının, hakkında her şeyi bildiğiniz büyük bir grubun (popülasyon) ortalamasından **anlamlı derecede farklı olup olmadığını** anlamaya yarayan bir istatistik testidir.

En kritik kuralı: Bu testi kullanabilmek için büyük grubun (popülasyonun) **standart sapmasını kesin olarak bilmeniz gerekir.** Bu, gerçek hayatta nadir bir durumdur.

T-Testi: Popülasyon Standart Sapması Bilinmediğinde

one sample z test

One Sample Z Test, **bilinen standart sapma** ile bir örneklemin ortalamasının belirli bir değere (μ_0) eşit olup olmadığını test eder.

Formül: $Z = (\bar{x} - \mu_0) / (\sigma/\sqrt{n})$

Kullanım şartları:

- Popülasyon standart sapması (σ) biliniyor
- $n \geq 30$ veya popülasyon normal dağılımlı

Hipotezler:

- H_0: $\mu = \mu_0$
- H_1: $\mu \neq \mu_0$ (iki yönlü) veya $\mu > \mu_0$ / $\mu < \mu_0$ (tek yönlü)

Çift Kuyruk (Two-tailed):

- H_1: $\mu \neq \mu_0$ (eşit değil)
- Red bölgesi: Her iki uçta ($\alpha/2$)
- Örnek: "Ortalama 100'den **farklı** mı?"

Sağ Kuyruk (Right-tailed):

- H_1: $\mu > \mu_0$ (büyüktür)
- Red bölgesi: Sağ uçta (α)
- Örnek: "Ortalama 100'den **büyük** mü?"

Sol Kuyruk (Left-tailed):

- H_1: $\mu < \mu_0$ (küçüktür)
- Red bölgesi: Sol uçta (α)
- Örnek: "Ortalama 100'den **küçük** mü?"

Karar: Z hesaplanan > Z kritik ise H_0 red edilir.

anova testi

ANOVA, **ikiden fazla** grubun ortalamaları arasında anlamlı bir fark olup olmadığını anlamak için kullanılan bir istatistik testidir.

Basitçe: "Bu gruplar birbirinden gerçekten farklı mı, yoksa aralarındaki farklar tesadüf mü?" sorusuna cevap verir.

Örnek: Üç farklı diyet programının kilo verdirme ortalamaları arasında bir fark var mı?

two sample z test

2. Two-Sample T-Test (Independent Samples)

İşte asıl karışıklık burada başlıyor!

Cevapladığı soru: "İki bağımsız grubun ortalamaları anlamlı derecede farklı mıdır?" **Ne zaman kullanılır:** İki ayrı, ilgisiz grubunuz olduğunda ve bunların ortalamalarını karşılaştırmak istediğinizde.

Anahtar kelime: BAĞIMSIZ

- A Grubu insanları B Grubu insanlarından tamamen farklıdır
- Bir gruba olan şey diğerini etkilemez
- Gözlemler arasında doğal bir eşleşme yok

Formül (eşit varyanslar varsayılarak): $t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$

$s_p^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$ -> (birleştirilmiş varyans)pooled variance

Burada neler oluyor:

- $(\bar{x}_1 - \bar{x}_2)$: İki örnek ortalaması arasındaki fark
- s_p^2 : Daha iyi bir tahmin elde edebilmek için her iki gruptan elde edilen varyanslar birleştirilir.
- Payda: Ortalamalar arasındaki farkın standart hatası

Yaygın Karışıklık #1: "Farklılıkları neden birleştiriyoruz?"

Diyelim ki 10 erkek ve 10 kadının boyları karşılaştırılıyor. Her grup size boylardaki değişkenlik hakkında bilgi verir. Sadece bir grubun değişkenliğini kullanmak yerine, daha istikrarlı bir tahmin elde etmek için her iki grubun bilgileri birleştirilir. Bu da daha fazla istatistiksel güç sağlar.

Not: Eşit varyansları varsaymayan (Welch t-testi olarak adlandırılır) ve pratikte kullanımı daha güvenli olan bir versiyonu da mevcuttur. Çoğu yazılım bunu otomatik olarak yapar.

Örnek : Yeni bir öğretim yönteminin (Yöntem A) geleneksel yöntemden (Yöntem B) daha iyi olup olmadığını öğrenmek istiyorsunuz. 30 öğrenciyi rastgele Yöntem A'ya, 28 öğrenciyi ise Yöntem B'ye atıyorsunuz. Bir dönem sonra, sınav puanlarını karşılaştırıyorsunuz.

Veriler

- Yöntem A : n_1 = 30, $(\bar{x}_1$ = 78, s_1 = 8
- Yöntem B : n_2 = 28, $(\bar{x}_2$ = 72, s_2 = 10

- **Bağımsızlığı doğrulayın ✓** A Yöntemindeki öğrenciler, B Yöntemindeki öğrencilerden farklı kişilerdir ✓ Rastgele atanırlar ✓ Bir öğrencinin puanı diğerinin puanını etkilemez. Bu, bağımsız iki örneklemli bir durumdur.
- Hipotezler: H_0: $\mu_1 = \mu_2$ (Yöntemler arasında fark yok) H_a: $\mu_1 \neq \mu_2$ (Yöntemler farklı sonuçlar üretir)

İki taraflı test (herhangi bir fark olup olmadığını kontrol etmek için).
Birleştirilmiş varyansı hesapla: $s\_p^2 = ((30-1)\times 8^2 + (28-1)\times 10^2) / (30 + 28 - 2)$ $s\_p^2 = (29\times 64 + 27\times 100) / 56$ $s\_p^2 = (1,856 + 2,700) / 56$ $s\_p^2 = 4,556 / 56$ $s\_p^2 = 81.36$
t-istatistiğini hesapla: $t = (78 - 72) / \sqrt{(81.36 \times (1/30 + 1/28))}$ $t = 6 / \sqrt{(81.36 \times 0.0690)}$ $t = 6 / \sqrt{5.61}$ $t = 6 / 2.37$ $t = 2.53$
Degrees of freedom $df = n_1 + n_2 - 2 = 30 + 28 - 2 = 56$
df = 56 için, α = 0,05'te iki kuyruklu test: Kritik değerler = ±2,003
2,53 > 2,003 olduğundan H_0'yi reddediyoruz
Sonuç: "Yöntem A ve Yöntem B'nin farklı ortalama puanlar ürettiğine dair önemli kanıtlar mevcuttur. Yöntem A'nın daha etkili olduğu (daha yüksek ortalama puan) görülmektedir."

bu kısımda önemli olan p değerini bulmaktır. karar mekanızması buna nağlıdır

Proportion Z Test

Proportion Z Test, bir örneklemdeki **oran ($\hat{p}$)** ile popülasyon oranının (p_0) eşit olup olmadığını test eder.

Formül: $Z = (\hat{p} - p_0) / \sqrt{[p_0(1-p_0)/n]}$

Kullanım şartları:

- $np_0 \geq 5$ ve $n(1-p_0) \geq 5$
- Büyük örneklem

Hipotezler:

- H_0: $p = p_0$
- H_1: $p \neq p_0$ (veya >, <)

Örnek: "Seçmenlerin %60'ı destekliyor mu?" gibi oran testleri.

Fark: One sample Z test ortalama için, Proportion Z test **oran** için kullanılır.

Müşteri Memnuniyet Oranı Örneği:

Bir restoran zinciri müşteri memnuniyet oranının %90'dan yüksek olduğunu iddia ediyor. Bunu test etmek istiyoruz.

H_0: $p = 0.90$ (Müşteri memnuniyet oranı %90'dır)

H_1: $p > 0.90$ (Müşteri memnuniyet oranı %90'dan **büyüktür**)

Test tipi: Sağ kuyruk (right-tailed)

Senaryo: 500 müşteriye anket yapıldı, 465'i memnun olduğunu belirtti.

- $\hat{p} = 465/500 = 0.93$
- $Z = (0.93 - 0.90) / \sqrt{[0.90(0.10)/500]}$ = **2.24**

$\alpha = 0.05$ için Z kritik = 1.645

Sonuç: $Z = 2.24 > 1.645 \rightarrow H_0$ red. Memnuniyet oranı gerçekten %90'dan yüksek! ✓

popülasyon stanstart sapmasını bilmiyorken z testi kullanılır.

gerçek hayatta stanstart sapma bilinmewmdiği için genelllikle t testi kullanılır.

Z ve T Arasındaki Temel Farklar:

Bakış Açısı	Z-Test	T-Test
Popülasyon Standart Sapması	Bilinen (σ)	Bilinmeyen(s kullanın)
Gerekli örneklem büyüklüğü	Büyük ($n \geq 30$) genellikle	Her boyutta çalışır

Dağılım Şekli	Standart normal (her zaman aynı)	Örneklem büyüklüğüne bağlı değişiklikler
Ne zaman kullamnılır?	Pratikte nadirdir	Çok yaygın

t testi türlerei

one sample t test

One Sample T Test, **bilinmeyen standart sapma** ile bir örneklemin ortalamasının belirli bir değere (μ_0) eşit olup olmadığını test eder.

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

Kullanım şartları:

- Popülasyon standart sapması (σ) **bilinmiyor** (s kullanılır)
- Küçük örneklem ($n < 30$) için popülasyon normal dağılımlı olmalı
- Serbestlik derecesi: $df = n - 1$

Hipotezler:

- H_0: $\mu = \mu_0$
- H_1: $\mu \neq \mu_0$ (veya >, <)

Z Test ile farkı:

- **Z Test:** σ biliniyor, büyük örneklem
- **T Test:** σ bilinmiyor (s kullan), küçük örneklem için uygun

Örnek: 20 öğrencinin sınav ortalaması 75'ten farklı mı? (σ bilinmiyor)

ONE SAMPLE vs TWO SAMPLE Testleri:

ONE SAMPLE (Tek Örneklem):

- **1 grup** var, **sabit bir değerle** (μ_0) karşılaştırılır
- **Soru:** "Bu grubun ortalaması belirli bir değere eşit mi?"
- **Örnek:** "Öğrencilerin boy ortalaması 170 cm'den farklı mı?"
- **Formül:** $t = (\bar{x} - \mu_0) / (s/\sqrt{n})$

TWO SAMPLE (İki Örneklem):

- **2 grup** var, **birbirleriyle** karşılaştırılır
- **Soru:** "İki grubun ortalamaları birbirine eşit mi?"
- **Örnek:** "Erkeklerin boy ortalaması kadınlardan farklı mı?"
- **Formül:** $t = (\bar{x}_1 - \bar{x}_2) / SE$ (standart hata)

Özetle:

- **One Sample:** Grup vs Sabit değer
- **Two Sample:** Grup 1 vs Grup 2

t-Testte Varyansın Önemi (Ders Notu)

t-testi, örneklem verilerinden yola çıkarak **iki ortalama arasında istatistiksel olarak anlamlı bir fark** olup olmadığını test eder. Bu testte **varyans** (ya da onun karekökü olan standart sapma), hem testin formülünde hem de türünün seçiminde kritik rol oynar.

1. **Varyans, Hata Payını Belirler**

t-test istatistiğinde paydada **standart hata** yer alır; bu da örneklem varyansına bağlıdır:
$$ t = \frac{\bar{x}_1 - \bar{x}_2}{\text{Standart Hata}} $$
→ Varyans ne kadar büyükse, standart hata o kadar büyük olur → t değeri küçülür → farkın anlamlı çıkma olasılığı azalır.

2. **Eşit Varyans Varsayımı (Homoskedastisite)**

Bağımsız iki örneklem t-testinde iki durum vardır:

- **Eşit varyans varsayılır** (Student’s t-test): İki grubun popülasyon varyansları aynı kabul edilir.
- **Eşit varyans varsayılmaz** (Welch’s t-test): Varyanslar farklıysa kullanılır; daha güvenlidir.

  > **Not:** Varyans homojenliği Levene testiyle kontrol edilir. $p < 0.05$ ise varyanslar **eşit değildir** → Welch t-testi tercih edilir.

3. **Eşleştirilmiş t-test**

Aynı bireylerin iki farklı durumdaki ölçümleri karşılaştırılır. Burada **farkların varyansı** önemlidir: Farkların varyansı küçükse, test daha güçlü (anlamlı sonuç çıkma şansı daha yüksek) olur.

Özet:
Varyans, t-testin **duyarlılığını**, **güvenilirliğini** ve **uygun türünün seçimini** doğrudan etkiler. Varyans ne kadar küçükse, ortalama farklarını tespit etmek o kadar kolay olur.

Ki-Kare (χ^2) Testleri: Kategorik Veriler İçin

Şimdi tamamen farklı bir veri türüne geçiyoruz: sayılar yerine kategoriler.

Ki-Kare Testlerini Ne Zaman Kullanırız?

Şimdiye kadar ortalamaları karşılaştırdık. Peki ya verileriniz sayısal değilse? Ya farklı kategorilere kaç kişinin girdiğini sayıyorsanız?

Örnekler

- Kırmızı arabalar, mavi arabalar, beyaz arabalar
- Ürün A'yı ve Ürün B'yi Ürün C'ye tercih eden kişiler
- Hastalar ve sağlıklılar
- Başarılılar ve başarısızlar

Ki-kare testleri bunlar için mükemmeldir!

Üç Tür Ki-Kare Testi

Bu durum insanları şaşırtıyor çünkü testlerin isimleri benzer ama cevapları farklı. Her birini açıklayalım.

1. Chi-Square Goodness of Fit Test (Ki-Kare Uyum İyiliği Testi)

Cevapladığı soru: "Gözlemlediğim kategorilerin dağılımı beklenen dağılımla uyuşuyor mu?"

Ne zaman kullanılır: BİR kategorik değişkeniniz olduğunda ve frekansların varsayılan bir desenle eşleşip eşleşmediğini test etmek istediğinizde.

Basit Bir Benzetme: Bir zarı 60 kez atıyorsunuz. Zar adilse, her sayının (1-6) yaklaşık 10 kez gelmesini beklersiniz. Ama aslında şunu elde edersiniz:

- 1: 8 kere
- 2: 12 kere
- 3: 9 kere
- 4: 11 kere
- 5: 13 kere
- 6: 7 kere

"Her biri 10"dan sapmalar sadece rastgele bir değişim mi, yoksa zarın bir etkisi mi var?

Formül:

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

O = Gözlenen frekans E = Beklenen frekans Σ = Tüm kategorilerdeki toplam

Bu formülün anlamı:

- Her kategori için şunları hesaplayın: (Elde ettiğiniz sonuç - Beklediğiniz sonuç)^2 / Beklediğiniz sonuç
- Tüm bu değerleri toplayın
- Büyük χ^2, gözlemlenen verilerinizin beklenenden çok farklı olduğu anlamına gelir
- Küçük χ^2, gözlemlenen verilerinizin beklenene yakın olduğu anlamına gelir

Örnek Senaryo: Bir şirket M&M şekerlemeleri üretiyor. Dağıtımın şu şekilde olduğunu iddia ediyorlar:

- %30 Kahverengi
- %20 Sarı
- %20 Kırmızı
- %10 Turuncu
- %10 Yeşil
- %10 Mavi

İçinde 500 adet M&M bulunan büyük bir çanta satın alıyorsunuz ve iddialarının doğru olup olmadığını test etmek istiyorsunuz.

Gözlemlediğiniz sayılar:

Renk	Gözlemlenen(O)	Beklenen(E)
Kahverengi	140	150 (30% of 500)
Sarı	110	100 (20% of 500)
Kırmızı	95	100 (20% of 500)
Turuncu	60	50 (10% of 500)
Yeşil	50	50 (10% of 500)
Mavi	45	50 (10% of 500)
Toplam	500	500

Adım 1: Hipotezler

● H_0: Dağılım iddia edilen oranlarla eşleşiyor ● H_a: Dağılım iddia edilen oranlarla eşleşmiyor

Adım 2: χ^2'yi hesaplayın

Kahverengi: $(140 - 150)^2 / 150 = 100 / 150 = 0.67$ Sarı: $(110 - 100)^2 / 100 = 100 / 100 = 1.00$ Kırmızı: $(95 - 100)^2 / 100 = 25 / 100 = 0.25$ Turuncu: $(60 - 50)^2 / 50 = 100 / 50 = 2.00$ Yeşil: $(50 - 50)^2 / 50 = 0 / 50 = 0.00$ Mavi: $(45 - 50)^2 / 50 = 25 / 50 = 0.50$

$\chi^2 = 0.67 + 1.00 + 0.25 + 2.00 + 0.00 + 0.50 = 4.42$

Adım 3: Serbestlik derecesi df = (kategori sayısı) - 1 = 6 - 1 = 5

Adım 4: Karar df = 5 ve $\alpha = 0,05$ için, χ^2 tablosundan kritik değer = 11,07

$4.42 < 11.07$ olduğundan H_0'yi reddedemiyoruz.

Sonuç: "Gözlemlenen dağılım, şirketin iddia ettiği oranlarla tutarlıdır. Herhangi bir sapma, rastgele şansa bağlanabilir."

Serbestlik Derecesi Nedir?

Basit tanım: Serbestlik derecesi, bir hesaplamada "özgürce değişebilen" değerlerin sayısıdır.

Sezgisel Açıklama

Şöyle düşünün: 4 sayınız var ve bunların toplamının 100 olması gerekiyor.

- İlk 3 sayıyı **özgürce** seçebilirsiniz: mesela 20, 30, 15
- Ama 4. sayıyı **seçemezsiniz** - otomatik olarak 35 olmak zorunda (çünkü toplam 100 olmalı)

Bu durumda **serbestlik dereceniz = 3** (4 değer - 1 kısıt = 3)

Ki-Kare Testlerinde Neden Önemli?

1. Doğru Kritik Değeri Bulmak İçin

Her serbestlik derecesi için farklı bir χ^2 dağılımı vardır:

```
df = 1  →  Kritik değer (α=0.05): 3.84
df = 5  →  Kritik değer (α=0.05): 11.07
df = 10 →  Kritik değer (α=0.05): 18.31
```

Yanlış df kullanırsanız, yanlış sonuca varırsınız!

2. Kategori Sayısı Arttıkça

- Daha fazla kategori = Daha yüksek df
- Daha yüksek df = Daha yüksek kritik değer
- Yani H_0'ı reddetmek daha zor olur

Mantığı: Daha fazla kategori varsa, rastgele sapmaların olması daha normaldir.

Farklı Ki-Kare Testlerinde df

1. Uyum İyiliği Testi

```
df = (kategori sayısı) - 1
```

Neden -1? Çünkü toplam frekans sabittir (örneğinizde 500). Son kategoriyi bilindiğinde, diğer kategorileri biliyorsanız, son kategori otomatik olarak bellidir.

Örnek: 6 renk → df = 6 - 1 = 5

2. Bağımsızlık Testi (İki değişken arasındaki ilişki)

```
df = (satır sayısı - 1) × (sütun sayısı - 1)
```

Örnek: 3×4'lük bir tablo → df = (3-1) × (4-1) = 2 × 3 = 6

3. Homojenlik Testi (Gruplar arasında dağılım karşılaştırma)

```
df = (satır sayısı - 1) × (sütun sayısı - 1)
```

Pratik Önem

Yanlış df kullanırsanız:

M&M örneğinizde df = 5 yerine yanlışlıkla df = 4 kullansaydınız:

- Kritik değer 9.49 olurdu (11.07 yerine)
- χ^2 = 4.42 hala reddedemezsiniz ama...
- Başka bir örnekte bu fark yanlış sonuca götürebilir!

df'nin Etkisi Görsel:

```
df küçük  → Dağılım sağa çarpık, kritik değer düşük
df büyük  → Dağılım daha simetrik, kritik değer yüksek
```

Özet

Serbestlik derecesi:

1. ✅ Hangi χ^2 dağılımını kullanacağınızı söyler
2. ✅ Kritik değeri belirler
3. ✅ Test sonucunu doğrudan etkiler
4. ✅ Her testte farklı hesaplanır
5. ✅ Mutlaka doğru hesaplanmalıdır

Pratik kural: df hesaplamayı formüle güvenin ve her zaman kontrol edin. Tabloda veya yazılımda doğru df değerini kullandığınızdan emin olun!

Başka bir soru var mı? Ki-kare testlerinin diğer türleri hakkında da konuşalım mı?

Chi-square (χ^2) test of independence

🎯 Amaç:

Bir olayın meydana gelme sıklığının, diğer bir değişkene **bağımlı mı yoksa bağımsız mı** olduğunu anlamak.

🔹 Örnek:

Bir araştırmacı, **cinsiyet** (erkek/kadın) ile **sigara kullanımı** (evet/hayır) arasında ilişki var mı diye merak ediyor.

- Cinsiyet ve sigara kullanımı **iki kategorik değişken** olduğu için → **Chi-square bağımsızlık testi** yapılır. Eğer test sonucu anlamlı çıkarsa, “**cinsiyet ile sigara kullanımı arasında ilişki vardır**” denir.

💡 Gerçek Hayat Örnekleri:

- Eğitim düzeyi ile internet kullanım alışkanlığı arasında ilişki var mı?
- Müşteri cinsiyeti ile ürün tercihi arasında bağlantı var mı?
- Yaş grubu ile oy verme eğilimi arasında ilişki var mı?

Kısaca:

> **Chi-square bağımsızlık testi**, “Bu iki kategori arasında bağlantı var mı?” sorusuna yanıt verir.

haftaya hoca soru soracak ve hangi test olacağını bilmeni isteyecek.

```
In [ ]:
```