

AKT102 İSTATİSTİK

BÖLÜM 13
Kİ-KARE TESTİ VE F-DAĞILIMI

§ 13.1

Uyum İyiliđi Testi

Çokterimli Deneyler

Çok terimli bir deney, her bir bağımsız deneme için ikiden fazla olası sonucun (sadece iki olası sonucun olduğu binom deneyinin aksine) olduğu sabit sayıda denemeden oluşan bir olasılık deneyidir.

Örnek:

Bir araştırmacı, gençler arasında favori pizza malzemelerinin dağılımının aşağıda gösterildiği gibi olduğunu iddia ediyor.

Malzeme	Frekans, f
peynir	41%
biber	25%
sosis	15%
mantar	10%
soğan	9%

Her sonuç kategorilere ayrılır.

Her olası sonuç için olasılık sabittir.

Ki-Kare Uyum İyiliđi Testi

Bir frekans dađılımının beklenen dađılıma uyup uymadığını test etmek için Ki-Kare Uygunluk Testi kullanılır.

Ki-kare uyum iyiliđi testinde test istatistiđini hesaplamak için, gözlemlenen frekanslar ve beklenen frekanslar kullanılır.

Bir kategorideki gözlenen frekans (O), örneklem verilerinde gözlemlenen kategorinin frekansıdır.

Bir kategoriden beklenen frekansı (E), kategori için hesaplanan frekanstır. Beklenen frekanslar, belirli (veya varsayımsal) dađılım varsayımıyla elde edilir. i. kategori için beklenen sıklık

$$E_i = np_i$$

burada n denemelerin sayısı (örneklem büyüklüğü) ve p_i , i. kategorinin varsayılan olasılıđıdır.

Gözlenen ve Beklenen Sıklıklar

Örnek:

200 genç rastgele seçildi ve en sevdikleri pizza malzemesinin ne olduğunu soruldu. Sonuçlar aşağıda gösterilmiştir.

Gözlenen frekansları ve beklenen frekansları bulun.

Malzeme	sonuç ($n = 200$)	gençlerin yüzdesi
peynir	78	41%
biber	52	25%
sosis	30	15%
mantar	25	10%
soğan	15	9%

Gözlenen sıklık	Beklenen sıklık
78	$200(0.41) = 82$
52	$200(0.25) = 50$
30	$200(0.15) = 30$
25	$200(0.10) = 20$
15	$200(0.09) = 18$

Ki-Kare Uyum İyiliği Testi

Ki-kare uygun iyiliği testinin kullanılması için aşağıdakilerin olması gerekir..

1. Gözlemlenen frekanslar rasgele bir örneklem kullanılarak elde edilmelidir.
2. Beklenen her frekans 5'ten büyük veya ona eşit olmalıdır.

Ki-Kare Uyum İyiliği Testi

Yukarıda listelenen koşullar sağlanmışsa, uyum testi için örneklem dağılımı $k - 1$ serbestlik derecesine sahip bir ki-kare dağılımına yaklaşır, burada k kategori sayısıdır. Ki-kare uyum iyiliği testi için test istatistiği;

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

Test her zaman sağ kuyruklu bir testtir.

O her kategorinin gözlenen sıklığını, E ise her kategorinin beklenen frekansını temsil eder.

Ki-Kare Uyum İyiliği Testi

Ki-Kare Uyum İyiliği Testinin Yapılması

Açıklama

1. İddiayı tanımlayın. Sıfır ve alternatif hipotezleri belirtin.
2. Önem düzeyini belirtin.
3. Serbestlik derecelerini tanımlayın.
4. Kritik değeri belirleyin.
5. Reddetme bölgesini belirleyin.

Gösterim

$$H_0 \quad H_a$$

$$\alpha.$$

$$\text{s.d} = k - 1$$

Tabloyu kullanın

Ki-Kare Uyum İyiliđi Testi

Ki-Kare Uyum İyiliđi Testinin Yapılması

Açıklama

6. Test istatistiklerini hesaplayın.
7. Sıfır hipotezi reddetme veya reddetme konusunda bir karar verin.
8. Kararı yorumlayın.

Gösterim

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

Eđer χ^2 ret bölgesindeyse H_0 reddedilir.

Ki-Kare Uyum İyiliđi Testi

Örnek:

Bir arařtırmacı, gençler arasında favori pizza malzemelerinin dađılımının ařađıda gösterildiđi gibi olduđunu iddia ediyor. 200 rastgele seilen genç ankete katıldı.

Malzeme	Frekans, f
peynir	39%
biber	26%
sosis	15%
mantar	12.5%
sođan	7.5%

$\alpha = 0.01$ ve daha önce hesaplanan gözlenen ve beklenen deđerleri kullanarak, anketörün ki kare uyum iyiliđi testi ile iddiasını test edin.

Ki-Kare Uyum İyiliği Testi

Örneğin devamı:

H_0 : Pizza malzemelerinin dağılımı % 39 peynir, % 26 biberli, % 15 sosis, % 12,5 mantar ve % 7,5 soğandır.
(İddia)

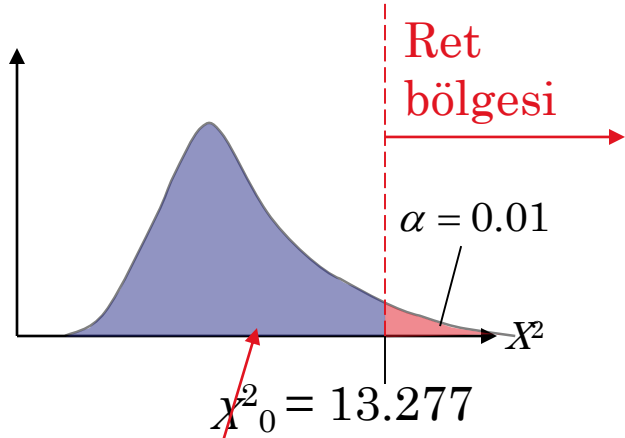
H_a : Pizza malzemelerinin dağılımı iddia edilen veya beklenen dağılımdan farklıdır.

5 kategori olduğundan, ki-kare dağılımı $k - 1 = 5 - 1 = 4$ serbestlik derecesine sahiptir.

s.d= 4 ve $\alpha = 0.01$ ile kritik değer $\chi^2_0 = 13.277$.

Ki-Kare Uyum İyiliği Testi

Örneğin devamı:



Malzeme	Gözlenen sıklık	Beklenen sıklık
peynir	78	82
biber	52	50
sisis	30	30
mantar	25	20
soğan	15	18

$$X^2 = \sum \frac{(O - E)^2}{E} = \frac{(78 - 82)^2}{82} + \frac{(52 - 50)^2}{50} + \frac{(30 - 30)^2}{30} + \frac{(25 - 20)^2}{20} + \frac{(15 - 18)^2}{18}$$

≈ 2.025

H_0 reddedilemez

Araştırmacının iddiasını reddetmek için % 1 düzeyinde yeterli kanıt yoktur.

§ 13.2

Bağımsızlık

Çapraz Tablo

Bir $r \times c$ çapraz tablosu, iki deęişken için gözlenen frekansları gösterir. Gözlenen frekanslar r satır ve c sütunlarında düzenlenmiştir. Bir satır ve bir sütunun kesişimine hücre adı verilir.

Aşağıdaki çapraz tablo, rasgele seçilen 321 ölümcül yaralı araç sürücüsünün yaş ve cinsiyete göre durumunu göstermektedir.

	yaş					
cinsiyet	16 – 20	21 – 30	31 – 40	41 – 50	51 – 60	61 +
erkek	32	51	52	43	28	10
kadın	13	22	33	21	10	6

Beklenen Sıklıklar

İki değişkenin bağımsız olduğu varsayılarak, her hücre için beklenen sıklığı bulmak için çapraz tablo kullanılabilir.

Çapraz Tablo Hücrelerinde Beklenen Frekansı Bulma

Çapraz tablodaki $E_{r,c}$ hücresi için beklenen frekans:

$$\text{Expected frequency } E_{r,c} = \frac{(\text{Sum of row } r) \times (\text{Sum of column } c)}{\text{Sample size}}.$$

Beklenen Sıklıklar

Örnek:

321 ölümcül yaralı sürücü örneklemini için çapraz tabloda her "Erkek" hücre için beklenen frekansı bulun. Değişkenlerin yaş ve cinsiyetin bağımsız olduğunu varsayalım.

	Age						
Gender	16 – 20	21 – 30	31 – 40	41 – 50	51 – 60	61 and older	Total
Male	32	51	52	43	28	10	216
Female	13	22	33	21	10	6	105
Total	45	73	85	64	38	16	321

Beklenen Sıklıklar

Örneğin devamı:

	Age						
Gender	16 – 20	21 – 30	31 – 40	41 – 50	51 – 60	61 and older	Total
Male	32	51	52	43	28	10	216
Female	13	22	33	21	10	6	105
Total	45	73	85	64	38	16	321

Expected frequency $E_{r,c} = \frac{(\text{Sum of row } r) \times (\text{Sum of column } c)}{\text{Sample size}}$

$$E_{1,1} = \frac{216 \cdot 45}{321} \approx 30.28 \quad E_{1,2} = \frac{216 \cdot 73}{321} \approx 49.12 \quad E_{1,3} = \frac{216 \cdot 85}{321} \approx 57.20$$

$$E_{1,4} = \frac{216 \cdot 64}{321} \approx 43.07 \quad E_{1,5} = \frac{216 \cdot 38}{321} \approx 25.57 \quad E_{1,6} = \frac{216 \cdot 16}{321} \approx 10.77$$

Ki-Kare Bağımsızlık Testi

İki değişkenin bağımsızlığını test etmek için ki-kare bağımsızlık testi kullanılır. Ki-kare testi kullanılarak, bir değişkenin ortaya çıkmasının diğer değişkenin oluşma olasılığını etkileyip etkilemediğini belirleyebilirsiniz.

Kullanılacak ki-kare bağımsızlık testi için, aşağıdakilerin olması gerekir.

1. Gözlemlenen frekanslar rasgele bir örneklem kullanılarak elde edilmelidir.
2. Beklenen her frekans 5'ten büyük veya ona eşit olmalıdır.

Ki-Kare Bağımsızlık Testi

Ki-Kare Bağımsızlık Testi

Listelenen koşullar yerine getirilirse, ki-kare bağımsızlık testi için örneklem dağılımı,

$$(r - 1) (c - 1)$$

serbestlik derecesi ile (r ve c çapraz tablodaki sırasıyla satır ve sütun sayısıdır) ki-kare dağılımına yaklaşır. Ki-kare bağımsızlık testi için test istatistiği;

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

Test her zaman sağ kuyruklu bir testtir.

O, gözlemlenen frekansları ve E, beklenen frekansları temsil eder.

Ki-Kare Bağımsızlık Testi

Ki-Kare Bağımsızlık Testi

Açıklama

1. İddiayı tanımlayın. Sıfır ve alternatif hipotezleri belirtin.
2. Önem düzeyini belirtin.
3. Serbestlik derecelerini tanımlayın.
4. Kritik değeri belirleyin.
5. Reddetme bölgesini belirleyin

Gösterim

$$H_0 \quad H_a$$

$$\alpha.$$

$$\text{s.d} = (r - 1)(c - 1)$$

Tabloyu kullanın

Ki-Kare Bağımsızlık Testi

Ki-Kare Bağımsızlık Testi

Açıklama

6. Test istatistiklerini hesaplayın.
7. Sıfır hipotezi reddetme veya reddetme konusunda bir karar verin.
8. Kararı yorumlayın.

Gösterim

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

Eğer χ^2 ret bölgesindeyse H_0 reddedilir.

Ki-Kare Bağımsızlık Testi

Örnek:

Aşağıdaki acil durum tablosu, 321 ölümcül yaralı binek araç sürücüsünün yaş ve cinsiyete göre rastgele bir örneğini göstermektedir. Beklenen frekanslar parantez içinde gösterilir. $\alpha = 0,05$ 'te, sürücülerin yaşlarının bu tür kazalarda cinsiyetle ilgili olduğu sonucuna varabilir misiniz?

	Age						
Gender	16 – 20	21 – 30	31 – 40	41 – 50	51 – 60	61 and older	Total
Male	32 (30.28)	51 (49.12)	52 (57.20)	43 (43.07)	28 (25.57)	10 (10.77)	216
Female	13 (14.72)	22 (23.88)	33 (27.80)	21 (20.93)	10 (12.43)	6 (5.23)	105
	45	73	85	64	38	16	321

Ki-Kare Bağımsızlık Testi

Örneğin devamı:

Beklenen her frekans en az 5 olduğundan ve sürücüler rasgele seçildiğinden, değişkenlerin bağımsız olup olmadığını test etmek için ki-kare bağımsızlık testi kullanılabilir.

H_0 : Sürücülerin yaşları cinsiyetten bağımsız.

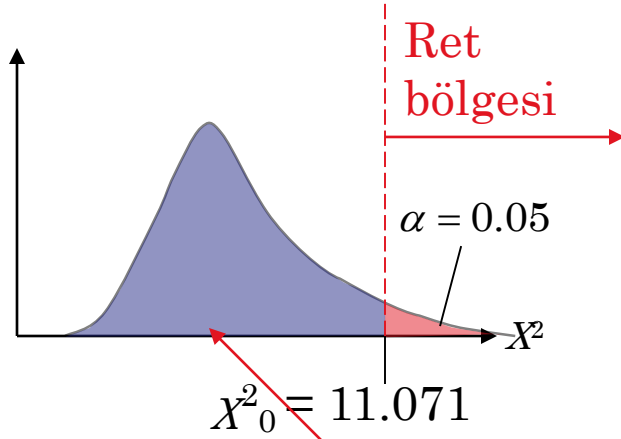
H_a : Sürücülerin yaşları ile cinsiyetleri bağımlı. (iddia)

$$\text{s.d} = (r - 1)(c - 1) = (2 - 1)(6 - 1) = (1)(5) = 5$$

s.d = 5 ve $\alpha = 0.05$ ile kritik değer $\chi^2_0 = 11.071$.

Ki-Kare Bağımsızlık Testi

Örneğin devamı:



$$X^2 = \sum \frac{(O - E)^2}{E} = 2.84$$

H_0 reddedilemez

O	E	$O - E$	$(O - E)^2$	$\frac{(O - E)^2}{E}$
32	30.28	1.72	2.9584	0.0977
51	49.12	1.88	3.5344	0.072
52	57.20	-5.2	27.04	0.4727
43	43.07	-0.07	0.0049	0.0001
28	25.57	2.43	5.9049	0.2309
10	10.77	-0.77	0.5929	0.0551
13	14.72	-1.72	2.9584	0.201
22	23.88	-1.88	3.5344	0.148
33	27.80	5.2	27.04	0.9727
21	20.93	0.07	0.0049	0.0002
10	12.43	-2.43	5.9049	0.4751
6	5.23	0.77	0.5929	0.1134

Bu kazalarda yaşın cinsiyete bağlı olduğu sonucuna varmak için % 5 anlam düzeyinde yeterli kanıt yoktur.

§ 13.3

**İki Varyansın
Karşılaştırılması**

F-Dağılımı

s_1^2 ve s_2^2 iki farklı popülasyonun örneklem varyanslarını temsil etsin. Her iki popülasyon normal ve popülasyon varyansları σ_1^2 ve σ_2^2 eşitse, örneklem dağılımı

$$F = \frac{s_1^2}{s_2^2}$$

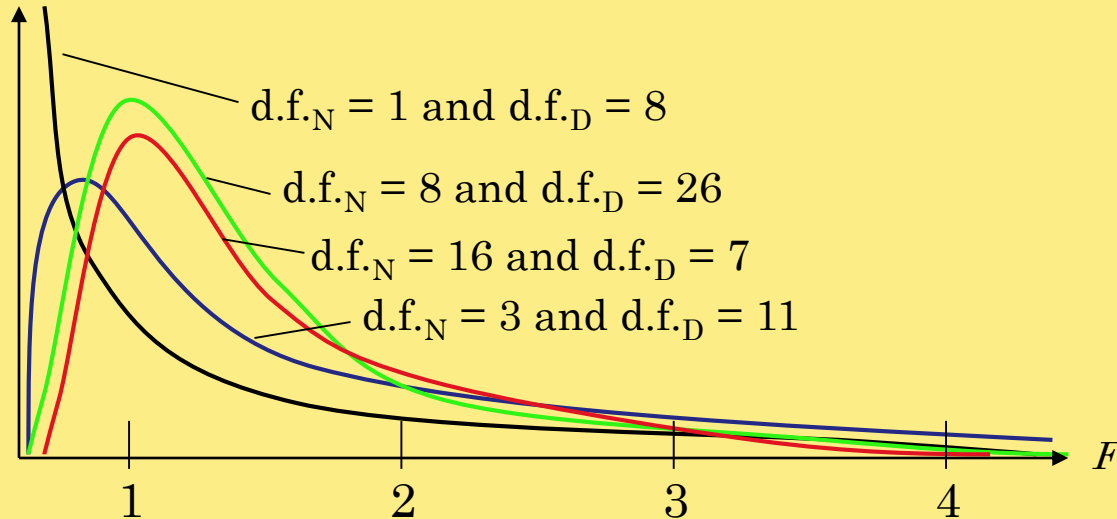
F-dağılımı adını alır.

Bu dağılımın birkaç özelliği vardır.

1. F-dağılımı, her biri iki serbestlik derecesi ile belirlenen bir eğri ailesidir: paydaki varyansa karşılık gelen serbestlik derecesi (d.f._N) ve paydadaki varyansa karşılık gelen serbestlik derecesi (d.f._D).

F-Dağılımı

2. *F*-dağılımı pozitif çarpık bir dağılımdır.
3. Bir *F* dağılımının her eğrisinin altındaki toplam alan 1'e eşittir.
4. *F* değerleri her zaman 0'dan büyük veya 0'a eşittir.
5. Tüm *F* dağılımları için, *F*'nin ortalama değeri yaklaşık 1'e eşittir.



F-dağılımı için Kritik Değerler

F-dağılımı için Kritik Değer Bulma

1. α anlamlılık seviyesini belirtin α .
2. Pay için serbestlik derecelerini belirleyin, d.f._N
3. Payda için serbestlik derecelerini belirlemek, d.f._D
4. Kritik değeri bulmak tabloyu kullanın. Eğer hipotez testi
 - a. Tek yönlü ise, F-tabloda α olan kısma bakılır.
 - b. İki yönlü ise F-tabloda $\frac{1}{2} \alpha$ olan kısma bakılır.

F-dağılımı için Kritik Değerler

Örnek:

$\alpha = 0.05$, $d.f._N = 5$ ve $d.f._D = 28$ iken sağ kuyruklu F kritik değerini bulun.

F-Dağılımı

d.f. _D : Paydanın serbestlik derecesi	$\alpha = 0.05$					
	d.f. _N : Payın serbestlik derecesi					
	1	2	3	4	5	6
1	161.4	199.5	215.7	224.6	230.2	234.0
2	18.51	19.00	19.16	19.25	19.30	19.33
27	4.21	3.35	2.96	2.73	2.57	2.46
28	4.20	3.34	2.95	2.71	2.56	2.45
29	4.18	3.33	2.93	2.70	2.55	2.43

Kritik değer $F_0 = 2.56$.

F-dağılımı için Kritik Değerler

Example:

$\alpha = 0.10$, d.f._N = 4 ve d.f._D = 6 iken iki yönlü F kritik değerini bulun.

$$\frac{1}{2}\alpha = \frac{1}{2}(0.10) = 0.05$$

F-Dağılımı

d.f. _D : Paydanın serbestlik derecesi	$\alpha = 0.05$					
	d.f. _N : Payın serbestlik derecesi					
	1	2	3	4	5	6
1	161.4	199.5	215.7	224.6	230.2	234.0
2	18.51	19.00	19.16	19.25	19.30	19.33
3	10.13	9.55	9.28	9.12	9.01	8.94
4	7.71	6.94	6.59	6.39	6.26	6.16
5	6.61	5.79	5.41	5.19	5.05	4.95
6	5.99	5.14	4.76	4.53	4.39	4.28
7	5.59	4.74	4.35	4.12	3.97	3.87

Kritik değer $F_0 = 4.53$.

Varyans için İki Örneklemli F-testi

Varyans için İki Örneklemli F-testi

Bir örneklem her popülasyondan rasgele seçildiğinde iki popülasyon varyansını σ_1^2 ve σ_2^2 karşılaştırmak için iki örneklemli F testi kullanılır. Popülasyonlar bağımsız olmalı ve normal dağılmalıdır. Test istatistiği

$$F = \frac{s_1^2}{s_2^2}$$

s_1^2 ve s_2^2 , $s_1^2 \geq s_2^2$ olacak şekilde örneklem varyansını temsil eder. Payın s.d. si d.f._N = $n_1 - 1$ ve paydanın s.d.si d.f._D = $n_2 - 1$ dir. Burada n_1 , s_1^2 varyansına sahip örneklem boyutu ve n_2 , s_2^2 varyansına sahip örneklem boyutunu gösterir.

Varyans için İki Örneklemli F-testi

İki Örneklemli F-Testi Kullanılarak
Karşılaştırılması

σ_1^2 ve σ_2^2

Açıklama

Gösterim

1. İddiayı tanımlayın. Sıfır ve alternatif hipotezleri belirtin.

H_0 H_a

2. Önem düzeyini belirtin.

α .

3. Serbestlik derecelerini tanımlayın.

d.f._N = $n_1 - 1$

d.f._D = $n_2 - 1$

Tabloyu kullanın.

4. Kritik değeri belirleyin

Varyans için İki Örneklemli F-testi

İki Örneklemli F-Testi Kullanılarak σ_1^2 and σ_2^2
Karşılaştırılması

Açıklama

5. Reddetme bölgesini belirleyin.
6. Test istatistiklerini hesaplayın.
7. Boş hipotezi reddetme veya reddetme konusunda bir karar verin.
8. Kararı yorumlayın.

Gösterim

$$F = \frac{s_1^2}{s_2^2}$$

Eğer F ret bölgesindeyse H_0 reddedilir.

İki Örneklemli F-Testi

Örnek:

Bir seyahat acentesinin pazarlama broşürü, iki şehir için otel oda fiyatlarının standart sapmalarının aynı olduğunu gösterir. Bir şehirde rasgele seçilen 13 otel odası 27,50 \$ standart sapmaya, diğer şehirden rasgele seçilen 16 otel odası ise 29,75 \$ standart sapmaya sahiptir. Ajansın $\alpha = 0.01$ 'deki iddiasını reddedebilir misiniz?

29.75 > 27.50 olduğundan, $s_1^2=885.06$ ve $s_2^2 = 756.25$.

$$H_0: \sigma_1^2 = \sigma_2^2 \quad (\text{iddia})$$

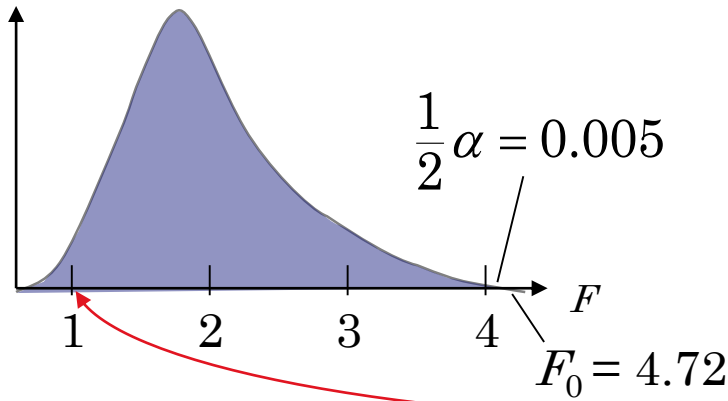
$$H_a: \sigma_1^2 \neq \sigma_2^2$$

İki Örneklemli F-Testi

Örneğin devamı:

Bu soru iki yönlü bir test ve $d.f._D = 12$.

$$\frac{1}{2} \alpha = \frac{1}{2} (0.01) = 0.005, d.f._N = 15$$



Kritik değer $F_0 = 4.72$.

Test istatistiği

$$F = \frac{s_1^2}{s_2^2} = \frac{885.06}{756.25} \approx 1.17.$$

H_0 reddedilemez

İki şehir için otel oda fiyatlarının standart sapmasının aynı olduğu iddiasını reddetmek için % 1 düzeyinde yeterli kanıt yoktur.

§ 13.4

Varyans Analizi

Tek Yönlü ANOVA

Tek yönlü varyans analizi, üç veya daha fazla popülasyondan gelen ortalamaları karşılaştırmak için kullanılan bir hipotez test tekniğidir. Varyans analizi genellikle ANOVA olarak kısaltılır.

Tek yönlü bir ANOVA testinde aşağıdakilerin olması gerekir.

1. Her örneklem normal veya yaklaşık olarak normal bir popülasyondan rasgele seçilmelidir.
2. Örneklemeler birbirinden bağımsız olmalıdır.
3. Her popülasyon aynı varyansa sahip olmalıdır.

Tek Yönlü ANOVA

$$\text{Test istatistiği} = \frac{\text{Örneklem arası varyans}}{\text{Örneklem içi varyans}}$$

1. Örneklem arası varyans (MS_B) her örnekleme verilen işlemle ilgili farklılıkları ölçer ve bazen örneklem arası ortalama kare olarak adlandırılır.
2. Örneklem içi varyans (MS_W) aynı örneklem içindeki girdilerle ilgili farklılıkları ölçer. Bazen örneklem içi ortalama kare olarak da adlandırılan bu sapma, genellikle örnekleme hatasından kaynaklanır.

Tek Yönlü ANOVA

Tek Yönlü Varyans Testi

Listelenen koşullar sağlanmışsa, test için örnekleme dağılımına F dağılımı yaklaşır. Test istatistiği

$$F = \frac{MS_B}{MS_W}.$$

F-testinin serbestlik derecesi

$$\text{d.f.}_N = k - 1$$

ve

$$\text{d.f.}_D = N - k$$

k örneklem sayısı ve N örneklem büyüklüklerinin toplamıdır.

Tek Yönlü ANOVA İçin Test İstatistikleri

Tek Yönlü ANOVA İçin Test İstatistiklerinin Bulunması

Açıklama

1. Her örneklemin ortalama ve varyansını bulun.
2. Tüm örneklerdeki tüm girdilerin ortalamasını bulun (genel ortalama).
3. Örnekler arası karelerin toplamını bulun.
4. Örnekler içi karelerin toplamını bulun.

Gösterim

$$\bar{x} = \frac{\sum x}{n} \quad s^2 = \frac{\sum (x - \bar{x})^2}{n - 1}$$

$$\bar{\bar{x}} = \frac{\sum x}{N}$$

$$SS_B = \sum n_i (\bar{x}_i - \bar{\bar{x}})^2$$

$$SS_W = \sum (n_i - 1) s_i^2$$

Tek Yönlü ANOVA İçin Test İstatistikleri

Tek Yönlü ANOVA İçin Test İstatistiklerinin Bulunması

Açıklama

5. Örneklem arası varyansı bulun.
6. Örneklem içi varyansı bulun
7. Test istatistiklerini bulun.

Gösterim

$$MS_B = \frac{SS_B}{k-1} = \frac{SS_B}{d.f._N}$$

$$MS_W = \frac{SS_W}{N-k} = \frac{SS_W}{d.f._D}$$

$$F = \frac{MS_B}{MS_W}$$

Tek Yönlü ANOVA Testinin Gerçekleştirilmesi

Tek Yönlü ANOVA Testinin Gerçekleştirilmesi

Açıklama

1. İddiayı tanımlayın. Sıfır ve alternatif hipotezleri belirtin.
2. Önem düzeyini belirtin.
3. Serbestlik derecelerini tanımlayın.
4. Kritik değeri belirleyin.

Gösterim

$$H_0 \quad H_a$$

$$\alpha$$

$$\text{d.f.}_N = k - 1$$

$$\text{d.f.}_D = N - k$$

Tabloyu kullanın.

Tek Yönlü ANOVA Testinin Gerçekleştirilmesi

Tek Yönlü ANOVA Testinin Gerçekleştirilmesi

Açıklama

5. Reddetme bölgesini belirleyin.
6. Test istatistiklerini hesaplayın.
7. Sıfır hipotezi reddetme veya reddetme konusunda bir karar verin.
8. Kararı yorumlayın.

Gösterim

$$F = \frac{MS_B}{MS_W}$$

Eğer F ret bölgesindeyse H_0 reddedilir..

ANOVA Özet Tablosu

Tablo, sonuçları tek yönlü ANOVA testinde özetlemenin uygun bir yoludur.

Değişim	Kareler toplamı	Serbestlik derecesi	Ortalama kareler	F
Örneklem arası	SS_B	d.f. _N	$MS_B = \frac{SS_B}{d.f._N}$	$MS_B \div MS_W$
Örneklem için	SS_W	d.f. _D	$MS_W = \frac{SS_W}{d.f._D}$	

Tek Yönlü ANOVA Testinin Gerçekleştirilmesi

Örnek:

Aşağıdaki tablo, dört büyükşehir bölgesinden rasgele seçilen kişilerin maaşlarını göstermektedir. $\alpha = 0,05$ anlam düzeyinde ortalama maaşın bölgelerden en az birinde farklı olduğu sonucuna varabilir misiniz?

Pittsburgh	Dallas	Chicago	Minneapolis
27,800	30,000	32,000	30,000
28,000	33,900	35,800	40,000
25,500	29,750	28,000	35,000
29,150	25,000	38,900	33,000
30,295	34,055	27,245	29,805

Tek Yönlü ANOVA Testinin Gerçekleştirilmesi

Örneğin devamı:

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$$

H_a : En az bir ortalama diğerlerinden farklı. (iddia)

$k = 4$ örneklem, $d.f._N = k - 1 = 4 - 1 = 3$.

Örneklem büyüklüklerinin toplamı

$$N = n_1 + n_2 + n_3 + n_4 = 5 + 5 + 5 + 5 = 20.$$

$$d.f._D = N - k = 20 - 4 = 16$$

$\alpha = 0.05$, $d.f._N = 3$, ve $d.f._D = 16$ kullanılarak, kritik değer $F_0 = 3.24$.

Tek Yönlü ANOVA Testinin Gerçekleştirilmesi

Örneğin devamı:

Test istatistiğini bulmak için aşağıdakiler hesaplanmalıdır.

$$\bar{\bar{X}} = \frac{\sum X}{N} = \frac{140745 + 152705 + 161945 + 167805}{20} = 31160$$

$$\begin{aligned} MS_B &= \frac{SS_B}{\text{d.f.}_N} = \frac{\sum n_i (\bar{X}_i - \bar{\bar{X}})^2}{k - 1} \\ &= \frac{5(28149 - 31160)^2 + 5(30541 - 31160)^2}{4 - 1} + \\ &\quad \frac{5(32389 - 31160)^2 + 5(33561 - 31160)^2}{4 - 1} \\ &\approx 27874206.67 \end{aligned}$$

Tek Yönlü ANOVA Testinin Gerçekleştirilmesi

Örneğin devamı:

$$MS_W = \frac{SS_W}{\text{d.f.}_D} = \frac{\sum(n_i - 1)s_i^2}{N - k}$$
$$\approx \frac{(5 - 1)(3192128.94) + (5 - 1)(13813030.08)}{20 - 4} +$$
$$\frac{(5 - 1)(24975855.83) + (5 - 1)(17658605.02)}{20 - 4}$$
$$= 14909904.97$$

$$F = \frac{MS_B}{MS_W} = \frac{27874206.67}{14909904.34} \approx 1.870$$

Test
istatistiği

Kritik
değer

1.870 < 3.24.

H₀ reddedilemez

Ortalama maaşın bölgelerden en az birinde farklı olduğu sonucuna varmak için % 5 anlam düzeyinde yeterli kanıt yoktur.