BÖLÜM 12

VARYANS ANALİZİ

Hipotez testleri bölümünde ortalamaların, oranların ve varyansların testlerini açıkladık. Açıklanan testlerde bir veya iki anakütle ile ilgili karşılaştırmalar yapılabilmekteydi. Bu bölümde ikiden fazla anakütle ortalamasının karşılaştırılması için kullanılan varyans analizi (ANOVA=Analysis of Variance) ele alınacaktır.

12.1. İKİDEN FAZLA ORTALAMA ARASINDAKİ FARKLARIN TESTİ

Çeşitli araştırmalarda ikiden fazla anakütle ortalaması arasında fark olup olmadığının diğer bir ifade ile aralarındaki farkların istatistiksel olarak anlamlı olup olmadığının araştırılması gerekebilir. Hipotez testleri bölümünde açıklandığı gibi anakütleler ikişer ikişer eşlendirilerek aralarındaki farkların anlamlılığı test edilebilir.

Ortalamalar ikişer ikişer test edilirken test edilecek ortalama sayısı k ise, C_k^2 veya k(k-1)/2 tane test yapılması gerekecektir. Örneğin, üç ortalama için $C_3^2 = 3$, dört ortalama için $C_4^2 = 6$, beş ortalama için $C_5^2 = 10$ adet ikişerli ortalama testi yapılması gerekmektedir. Bu durumda tüm test işlemleri için 1.tip hata yapma olasılığı artacaktır. Örneğin, üç anakütle ortalaması ikişerli olarak karşılaştırılacak olsun. Hata payı yani 1.tip hata yapma olasılığı 0,05, güven olasılığı 0,95 ise üç testte de doğru sonuç elde etme olasılığı binom dağılımı ile,

$$P(X = 3) = C_3^3 (0{,}95)^3 (0{,}05)^0 = (0{,}95)^3 = 0{,}86$$

olacaktır. Bu durumda en az bir 1.tip hata yapma olasılığı (1-0,86=0,14)'tür. Aynı olasılık 0,05 hata payı ile 5 ortalamanın testi için,

$$P(X = 10) = C_{10}^{10} (0{,}95)^{10} (0{,}05)^0 = (0{,}95)^{10} = 0{,}60$$

ve 1.tip hata yapma olasılığı (1-0,60 =0,40) olacaktır. Görüldüğü gibi ortalama sayısı arttıkça en az bir 1.tip hata yapma olasılığı artmaktadır. Bu açıklamalardan da anlaşılacağı gibi ikiden fazla ortalamanın ikişer ikişer test edilmesi durumunda hata sözkonusudur, ikişerli testler birden fazla ortalamanın eşitliğinin testi için yetersizdir.

Varyans analizi ikiden fazla ortalamanın testi için kullanılan yeterli bir yöntemdir. Yöntemin temeli farklı kaynaklardan hesaplanan varyansların karşılaştırılmasına dayanmaktadır. k ortalama arasındaki fark test edilirken örneklerin varyansı σ^2 olan normal dağılmış anakütlelerden alındığı varsayılmaktadır. Diğer bir ifade ile, örneklerin aynı anakütleden alındığının varsayıldığını söyleyebiliriz. Temel hipotez ortalamaların eşitliğini ifade edecek şekilde oluşturulduğundan, örneklerin alındığı anakütlelerin normal dağıldığı ve varyanslarının eşit (homojen) olduğu kabul edilmektedir.

Varyans analizinin temeli daha önce açıklanan F dağılımına dayanmaktadır. k ortalama arasındaki farkların istatistiksel anlamlılığı varyans analizi ile şöyle yapılmaktadır.

1- Temel hipotez ortalamaların eşitliğini, karşıt hipotez ise temel hipotezin doğru olmadığını belirtecek şekilde hipotezler oluşturulur. Karşıt hipotez tüm ortalamaların eşit olmadığını gösterecektir.

$$H_0 : \mu_1 = \mu_2 = ... = \mu_k$$

$$H_1 : \mu_1 \neq \mu_2 \neq .... \neq \mu_k$$

Burada H_1 hipotezi ortalamalardan tümünün değil, en az birinin farklı olduğunu ifade etmektedir.

Örneklerin ortalamaları hesaplanarak, k örneğin ortalaması için ortalamaların standart hatası ($S_{\overline{X}}$) hesaplanır.

$$S_{\overline{X}} = \sqrt{\frac{\sum_{i=1}^{k} (\overline{X}_i - \overline{\overline{X}})^2}{k-1}}$$

Burada $\overline{\overline{X}}$ k örneğin ortalamalarının ($\overline{X}_i$) ortalamasıdır.

$$\overline{\overline{X}} = \frac{\sum_{i=1}^{k} \overline{X}_i}{k}$$

2- Hesaplanan ortalamaların standart hatası ($S_{\overline{X}}$) ile k örneğin alındığı anakütlenin varyansı tahmin edilir. Tahmin edilen anakütle varyansı S^2 ile ifade edilir ve gruplar arası kareler ortalaması ($KO_{G.Arası}$) olarak adlandırılır. Ortalamaların standart hatası,

$$S_{\overline{X}} = \frac{S}{\sqrt{n}}$$

olduğundan,

$$S = S_{\overline{X}} \sqrt{n}$$

$$S^2 = S_{\overline{X}}^2 . n = \text{KO}_{\text{G.Arası}}$$

olacaktır. Burada n örnek birim sayısıdır ve tüm örneklerin birim sayılarının eşit olduğu varsayılmaktadır.

3- Grupların ortalamalarına göre ve bağımsız olarak anakütle varyansı tahmin edilir. Tahmin edilen varyansların tartılı ortalamaları alınarak bunlar birleştirilir. Tartı olarak örnek birim sayılarının bir eksiği alınır. Örnek birim sayılarını n_i ile ifade edersek, tartı (n_i –1) olacaktır. Anakütle varyansının tahmin edilen bu değeri gruplar içi kareler ortalaması ($KO_{G.içi}$) olarak adlandırılır.

$$\text{KO}_{\text{G.İÇİ}} = \frac{\sum_{i=1}^{k} (n_i - 1) S_i^2}{\sum_{i=1}^{k} (n_i - 1)}$$

4- Elde edilen bu iki S^2 arasındaki fark F dağılımı ile tek taraflı olarak test edilir. Test istatistiği,

$$F = \frac{KO_{G.Arası}}{KO_{G.İçi}}$$

dir. F dağılımın serbestlik dereceleri,

SD_1= k-1

SD_2= nk-k

dır. Verilen serbestlik dereceleri ve belirlenen α hata payı ile F dağılımı tablosundan bulunan $F_{\alpha,SD1,SD2}$ değeri ile F test istatistiği karşılaştırılır.

$F < F_{\alpha,SD1,SD2}$	ise	H_0 kabul
$F > F_{\alpha,SD1,SD2}$	ise	H_1 kabul

edilecektir.

ÖRNEK: Aşağıda öğrencilerin 50 üzerinden üç dersten aldıkları notlar verilmiştir.

Dersler	Notlar				
Matematik	29	16	36	15	24
İstatistik	41	28	33	11	37
Bilgisayar	49	29	38	25	19

0,05 hata payı ile ortalamalar arasındaki farkların anlamlılığını (örnek ortalamalarının aynı anakütleden alınıp alınmadığını) inceleyiniz.

Dersler	Notlar					Notların Toplamı	Xi
Matematik	29	16	36	15	14	110	22
İstatistik	41	28	33	11	37	150	30
Bilgisayar	49	29	38	25	19	160	32
							84

1- $H_0 : \mu_1 = \mu_2 = \mu_3$

$H_1 : \mu_1 \neq \mu_2 \neq \mu_3$

$$S_{\overline{X}} = \sqrt{\frac{\sum_{i=1}^{k}(\overline{X}_i - \overline{\overline{X}})^2}{k-1}}$$

$$\overline{\overline{X}} = \frac{\sum_{i=1}^{k}\overline{X}_i}{k} = \frac{84}{3} = 28$$

$$S_{\overline{X}} = \sqrt{\frac{(22-28)^2 + (30-28)^2 + (32-28)^2}{3-1}} = 5{,}29$$

2- $KO_{G.Arası} = S_{\overline{X}}^2 .n = (5{,}29)^2.(5) = 139{,}92$

3- $KO_{G.İçi} = \dfrac{\Sigma(X_i - \overline{X})^2}{n-1}$

$$S_1^2 = \frac{(29-22)^2 + (16-22)^2 + (36-22)^2 + (15-22)^2 + (14-22)^2}{5-1}$$

$$= 98{,}5$$

$$S_2^2 = \frac{(41-30)^2 + (28-30)^2 + (33-30)^2 + (11-30)^2 + (37-30)^2}{5-1}$$

$$= 136$$

$$S_3^2 = \frac{(49-32)^2 + (29-32)^2 + (38-32)^2 + (25-32)^2 + (19-32)^2}{5-1}$$

$$= 138$$

$$KO_{G.içi} = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2 + (n_3-1)S_3^2}{(n_1-1)+(n_2-1)+(n_3-1)}$$

$$= \frac{(5-1)(98{,}5) + (5-1)(136) + (5-1)(138)}{(5-1)+(5-1)+(5-1)} = 124{,}16$$

4- SD_1= k-1 = 3-1=2

SD_2= [(n.k)-k] = 15-3=12

$F_{\alpha,SD1,SD2} = F_{0,05,2,12} = 3,89$

$$F = \frac{KO_{G.arası}}{KO_{G.İçi}} = \frac{139,92}{124,16} = 1,12$$

1,12<3,89 olduğundan H_0 hipotezi kabul edilir. Ortalamalar arasındaki fark anlamsızdır.

12.2. TEK YÖNLÜ VARYANS ANALİZİ

Tek yönlü varyans analizinde k sayıdaki örnek ortalaması arasındaki fark test edilirken doğrusal model kullanılır. Ortalamaları test edilecek k sayıdaki anakütleyi $A_1, A_2, \ldots, A_k$ ile anakütle ortalamalarını μ_i ve varyanslarını σ^2 ile ifade edelim. Bu anakütleler deneme grubu olarak adlandırılmaktadır. Bu anakütlelerin biraraya gelmesi ile ortalaması μ olan bir anakütle oluşacaktır. Bu büyük anakütlenin ortalaması,

$$\mu = \frac{\sum_{i=1}^{k} \mu_i}{k}$$

olacaktır.

Tek yönlü varyans analizinde temel hipotez daha önce sözedildiği gibi,

$$H_0 : \mu_1 = \mu_2 = ... = \mu_k$$

olarak oluşturulacaktır. Temel hipotezin doğru olması durumunda tüm ortalamalar birbirlerine eşit olduğu gibi, bunlar tüm anakütlelerin ortalaması μ'ye de eşit olacaktır.

$$\mu = \mu_1 = \mu_2 = ... = \mu_k = \mu_i$$

Temel hipotezin doğru olmaması durumunda ise $\mu \neq \mu_i$ olacak ve aralarında a_i ile ifade edebileceğimiz farklar olacaktır.

$$\alpha_i = \mu_i - \mu$$

Burada α_i, A_i anakütlesinin işleyim etkisi olarak tanımlanmaktadır. Bu etki dikkate alınırsa, temel hipotez doğru ise,

$$\alpha_1 = \alpha_2 = ... = \alpha_k = 0$$

$$\sum_{i=1}^{k} \alpha_i = \sum_{i=1}^{k} (\mu_i - \mu) = 0$$

olacaktır. Yani deneme etkileri sıfırdır.

Normal dağılmış herhangi bir i. örneğin j. birimi için,

$$X_{ij} = \mu_i + \varepsilon_{ij}$$

yazılabilir. Burada ε_{ij} hata terimi veya artıktır. ε_{ij} 'lerin birbirinden bağımsız oldukları, sıfır ortalama ve σ^2 varyansı ile normal dağıldıkları kabul edilmektedir.

Daha önce $\alpha_i = \mu_i - \mu$ olarak tanımlanmıştı. Buradan,

$$\mu_i = \mu + \alpha_i$$

olarak elde edilir. Bu eşitlik X_{ij} eşitliğinde yerine konursa,

$$X_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

bulunur. Elde edilen bu model tek yönlü varyans analizi modelidir.

Tek yönlü varyans analizinde kullanılacak veri Tablo 12.1.'de görüldüğü gibi bir tablo oluşturulacaktır.

Tablo 12.1

Gözlemler	**Deneme Grupları**	**(j)**		
	1	2	......	k
1	X_{11}	X_{21}	......	X_{k1}
2	X_{12}	X_{22}	....	X_{k2}
...	...	....		
N	X_{1n1}	X_{2n2}	....	X_{knk}
Toplam	$\sum_{j=1}^{n_1} X_{1j}$	$\sum_{j=1}^{n_2} X_{2j}$	....	$\sum_{j=1}^{n_k} X_{kj}$
n_i	n_1	n_2	....	n_k

Bu verilerden örnek ortalamaları,

$$\overline{X}_i = \frac{X_{i1} + X_{i2} + ... + X_{in_i}}{n_i} = \frac{\sum_{j=1}^{n_i} X_{ij}}{n_i}$$

tüm örneklerin birleşmesi ile oluşacak büyük örneğin ortalaması,

$$\overline{\overline{X}} = \frac{\sum_{j=1}^{n_1} X_{1j} + \sum_{j=1}^{n_2} X_{2j} + ... + \sum_{j=1}^{n_k} X_{kj}}{N} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}}{N}$$

$$N = n_1 + n_2 + ... + n_k = \sum_{i=1}^{k} n_i$$

olarak hesaplanacaktır.

Örnek birimleri X_{ij}'lerin genel ortalama $\overline{\overline{X}}$'dan sapmaları, örnek birimlerinin kendi deneme ortalamalarından sapmaları ile deneme ortalamasının genel ortalamadan farklarının toplamına eşit olduğundan,

$$X_{ij} - \overline{\overline{X}} = (X_{ij} - \overline{X}_i) + (\overline{X}_i - \overline{\overline{X}})$$

olarak yazılabilir.

Hata teriminin tahmincisi $(X_{ij} - \overline{X}_i)$ ve deneme etkisinin tahmincisi $(\overline{X}_i - \overline{\overline{X}})$ olacağından örnek birimlerinin büyük ortalamadan farkı bunların toplamına eşittir.

Yukarıda verilen eşitlikten sapmaların kareleri toplamı hesaplanabilir.

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \overline{\overline{X}})^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left[(X_{ij} - \overline{X}_i)^2 + (\overline{X}_i - \overline{\overline{X}})\right] =$$

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \overline{X}_i)^2 + \sum_{i=1}^{k}\sum_{j=1}^{n_i}(\overline{X}_i - \overline{\overline{X}})^2 + 2\sum_{i=1}^{k}\sum_{j=1}^{n_1}(X_{ij} - \overline{X}_i)(\overline{X}_i - \overline{\overline{X}})$$

verilen eşitliğin son ifadesinde yer alan $(\overline{X}_i - \overline{\overline{X}})$, i. işlemdeki tüm gözlemler j.gözlemler için sabittir. Aynı ifade de yeralan $(X_{ij} - \overline{X}_i)$ toplamı ise, i. işlemdeki tüm j. gözlemler için sıfırdır. Bu nedenle,

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \overline{\overline{X}})^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \overline{X}_i)^2 + \sum_{i=1}^{k}\sum_{j=1}^{n_i}(\overline{X}_i - \overline{\overline{X}})^2$$

olacaktır. Burada eşitliğin sol tarafı genel sapma karelerin toplamların toplamıdır.

$$KT_G = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \overline{\overline{X}})^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}X_{ij}{}^2 - N(\overline{\overline{X}})^2$$

Eşitliğin sağ tarafındaki ilk ifade deneme içi sapma kareler toplamıdır.

$$KT_{D.içi} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \overline{X}_i)^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}X_{ij}{}^2 - \sum_{i=1}^{k}n_i(\overline{X})^2$$

Eşitliğin sağ tarafındaki ikinci ifade ise, denemeler arası sapma kareler toplamıdır.

$$KT_{D.arası} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(\overline{X}_i - \overline{\overline{X}})^2 = \sum_{i=1}^{k} n_i \overline{X}_i^2 - N\overline{\overline{X}}^2$$

Bu tanımlamalara göre kısaca,

$KT_G = KT_{D.içi} + KT_{D.arası}$

olacaktır. Bu açıklamalardan da anlaşılacağı gibi birimler arasındaki fark tesadüfi sapmalardan ve farklı anakütlelerden oluşmaktadır.

Tek yönlü varyans analizinde serbestlik dereceleri KT_G için (N-1), $K_{TD.arası}$ için (k-1) ve SSE için (N-k)'dır. Varyans analizinde σ^2'nin tahmincileri kareler ortalaması (KO) olarak adlandırılır. Kareler ortalaması ilgili sapma karelerin ilgili serbestlik derecelerine bölünmesi ile bulunacaktır.

$$KO_{D.arası} = \frac{KT_{D.arası}}{k-1}$$

$$KO_{D.içi} = \frac{KT_{D.İ.İ}}{N-k}$$

$\dfrac{KO_{D.arası}}{KO_{D.içi}}$ dağılımı (k-1) ve (N-k) serbestlik dereceli F

dağılımıdır. Test istatistiği,

$$F = \frac{KO_{D.arası}}{KO_{D.içi}}$$

olarak hesaplanır.

Bu açıklamalara göre varyans analizi tablosu Tablo 12.2. görüldüğü gibi oluşacaktır.

Tablo 12.2

Değişim Kaynağı	Kareler Toplamı	Serbestlik Derecesi	Kareler Ortalaması	F
Deneme Arası	$KT_{D.arası}$	k-1	$KO_{D.arası} = \frac{KT_{D.arası}}{k-1}$	
Deneme İçi	$KT_{D.İçi}=KT_G-KT_{D.arası}$	N-k	$KO_{D.içi} = \frac{KT_{D.İ.İ}}{N-k}$	$F = \frac{KO_{D.arası}}{KO_{D.içi}}$ $F = \frac{MSA}{MSE}$
TOPLAM	SST	N-1		

Tek yönlü varyans analizinde hipotezler daha önce açıklandığı gibi oluşturulur. Belirlenen α hata payı ve (k-1), (N-k) serbestlik dereceleri ile F tablosundan tablo değeri bulunur. Test istatistiği,

$$F=\frac{KO_{D.arası}}{KO_{D.içi}}$$

olarak hesaplanarak, $F < F_{\alpha,k-1,N-k}$ ise H_0 , $F > F_{\alpha,k-1,N-k}$ ise H_1 hipotezi kabul edilir.

ÖRNEK: Aşağıda öğrencilerin 50 üzerinden üç dersten aldıkları notlar verilmiştir.

Dersler	Notlar				
Matematik	29	16	36	15	14
İstatistik	41	28	33	11	37
Bilgisayar	49	29	38	25	19

0,05 hata payı ile ortalamalar arasındaki farkların anlamlılığını (örnek ortalamalarının aynı anakütleden alınıp alınmadığını) inceleyiniz.

Dersler	Notlar					Notların Toplamı	Toplamlar kareleri
Matematik	29	16	36	15	14	110	12100
İstatistik	41	28	33	11	37	150	22500
Bilgisayar	49	29	38	25	19	160	25600
						420	60200

1- $H_0 : \mu_1 = \mu_2 = \mu_3$

$H_1 : \mu_1 \neq \mu_2 \neq \mu_3$

$F_{\alpha,k-1,N-k}=3,89$

$\Sigma X_1 = 110$ $\qquad (\Sigma X_1)^2 = 12100$

$\Sigma X_2 = 150$ $\qquad (\Sigma X_2)^2 = 22500$

$\Sigma X_3 = 160$ $\qquad (\Sigma X_3)^2 = 25600$

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}(\Sigma X_1 + \Sigma X_2 + \Sigma X_3) = 420$$

$$\left(\sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}\right)^2 = 176400$$

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}^2 = (29)^2 + (16)^2 + (36)^2 + (15)^2 + (14)^2 + (41)^2 + (28)^2$$

$$+ (33)^2 + (11)^2 + (37)^2 + (49)^2 + (29)^2 + (38)^2 + (25)^2$$

$$+ (19)^2 = 13530$$

$$\overline{\overline{X}} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}}{N} = \frac{420}{15} = 28$$

$$KT_G = \sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}{}^2 - N\overline{\overline{X}}{}^2$$

$$= 13530 - 15(28)^2 = 13530 - 11760 = 1770$$

$$KT_{D.arası} = \sum_{i=1}^{k} n_i \overline{X}_i^2 - N\overline{\overline{X}}{}^2$$

$$= \left[5\left(\frac{110}{5}\right)^2 + 5\left(\frac{150}{5}\right)^2 + 5\left(\frac{160}{5}\right)^2\right] - 15(28)^2$$

$$= (2420 + 4500 + 5120) - 11760$$

$$= 12040 - 11760 = 280$$

$$KT_{D.İ.İ} = \sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}{}^2 - \sum_{i=1}^{k} n_i (\overline{X})^2$$

$$= 13530 - 12040 = 1490$$

veya

$KT_G = KT_{D.İÇİ} + KT_{D.arası}$

$1770 = 280 + 1490$

ilişkisinden yararlanılarak ikisi hesaplanıp üçüncü bu bağıntıdan bulunabilir.

$$KO_{D.arası} = \frac{KT_{D.arası}}{k-1} = \frac{280}{2} = 140$$

$$KO_{D.içi} = \frac{KT_{D.İ.İ}}{N-k} = \frac{1490}{12} = 124{,}16$$

$$F = \frac{KO_{D.arası}}{KO_{D.içi}} = \frac{140}{124{,}16} = 1{,}12$$

ANOVA Tablosu

Değişim Kaynağı	Kareler Toplamı	Serbestlik Derecesi	Kareler Ortalaması	F
Denemeler Arası	280	3-1=2	140	
Denemeler İçi	1490	15-3=12	124,66	F=1,12
TOPLAM	1770	14		

1.12 < 3,89 olduğundan H_0 hipotezi kabul edilir.

12.3. ÇİFT YÖNLÜ VARYANS ANALİZİ

Tek yönlü varyans analizinde incelenenin aksine bazı olaylarda değişkenlerdeki değişmeler iki faktörden kaynaklanır ve bunlar ile ilgili bilgiler bulunabilir. Bu durumda değişmeler iki bölümden oluşmaktadır. Bu nedenle tek yönlü varyans analizi yerine çift yönlü varyans analizi yapılması gerekmektedir. Örneğin daha önce öğrencilerin dersleri ile ilgili örnek çözmüştük. Bu örnekte sadece öğrencilerin derslerden aldıkları notlar yer almıştı. Oysa öğrencilerin orta eğitim yaptıkları okullar, o okulların eğitim şekilleri veya orta öğretimde mezun oldukları bölümler derslerdeki başarıyı etkiliyor olabilir. Bu bilgilerinde eklenmesi ile oluşacak veri tek yönlü analiz etkilemeyeceğinden, çift yönlü varyans analizi ile ortalamalar arasındaki farkların test edilmesi gerekecektir.

Farklı iki değişken sözkonusu olduğunda, değişkenlerin birbirlerine olan etkileri veya bağımsız olup olmadıkları da önemlidir. Bazı olaylarda değişkenler karşılıklı olarak birbirlerini etkilerler, yani bağımsız değillerdir. Bazı olaylarda ise karşılıklı etkileşim sözkonusu değildir. Bu nedenle çift yönlü varyans analizinin etkileşim olmaması ve etkileşim olması durumları için ayrı ayrı incelenmesi gerekmektedir.

Veriler iki değişken için düzenlendiğinde bir matris veya tablo oluşturulur. Çift yönlü varyans analizinde tablonun veya matrisin hücrelerinde birer gözlem yer alması durumu tesadüfi blok düzeni olarak tanımlanmaktadır. Bazı durumlarda ise birden fazla eşlenmiş gözlemler sözkonusudur. Bu gözlemlerin oluşturduğu gruplar blok olarak adlandırılır. Çift yönlü varyans analizinde blokların etkisi sözkonusu olmaktadır.

12.3.1. Etkileşimsiz Çift Yönlü Varyans Analizi

Çift yönlü varyans analizinde analize değişmeleri açıklayacak yeni kaynak eklenmesi ile hata kareler ortalaması (KO_{Hata}) azaltılabilir. Bu durumda bloklarında analize katılması ile varyans analizi modeli genişleyecektir. Etkileşim olmaması durumunda çift yönlü varyans analizi modeli,

$$X_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

olur. Burada,

X_{ij} : i. işleyimin j. gözlemini,

μ : genel anakütle ortalamasını

α_i : i. deneme etkisini

β_j : j. blok etkisini

ε_{ij} : hata terimini

ifade etmektedir. X_{ij}'nin dağılımı ortalaması μ ve varyansı σ^2 olan normal dağılım, ε_{ij}'nin dağılımı ise ortalaması sıfır, varyansı σ^2 olan normal dağılımdır. Ayrıca $\sum_{i=1}^{k} \alpha_i = 0$ ve $\sum_{j=1}^{n} \beta_j = 0$'dır.

Modelde α_i ve β_j ile ifade edilen etkiler yer aldığından temel hipotez bu etkilerin sıfır olacağını ifade edecek şekilde oluşturulacaktır.

$$H_0 : \alpha_i = 0 \qquad i = 1,2,3,.....,k$$
$$\beta_j = 0 \qquad j = 1,2,3,.....,n$$

Karşıt hipotez en az bir α_i veya en az bir β_j'nin sıfırdan farklı olduğunu ifade etmektedir.

Örnek birimlerinin değerlerinin (X_{ij} gözlemlerinin) büyük örnek ortalamasından farkları,

$$X_{ij} - \overline{\overline{X}} = \left(X_{ij} - \overline{X}_i - \overline{X}_j + \overline{\overline{X}}\right) + \left(\overline{X}_i - \overline{\overline{X}}\right) + \left(\overline{X}_j - \overline{\overline{X}}\right)$$

olacaktır. Sapmaları kareleri toplamı alınırsa,

$$\sum_{i=1}^{k}\sum_{j=1}^{n}\left(X_{ij} - \overline{\overline{X}}\right)^2 = \sum_{i=1}^{k}\sum_{j=1}^{n}\left(X_{ij} - \overline{X}_i + \overline{X}_j + \overline{\overline{X}}\right)^2 + \sum_{i=1}^{k}\sum_{j=1}^{n}\left(\overline{X}_i - \overline{\overline{X}}\right)^2$$

$$+ \sum_{i=1}^{k}\sum_{j=1}^{n}\left(\overline{X}_j^2 - \overline{\overline{X}}\right)^2$$

olacaktır. Burada $\overline{X}_i$, i. işleyimin gözlemlerinin ortalamasını, $\overline{X}_j$, j. işleyimin gözlemlerinin ortalamasını ve $\overline{\overline{X}}$ ise tüm gözlemlerin ortalamasını ifade etmektedir. Eşitliğin sol tarafı KT_G, sağ tarafı ise sırasıyla KT_H, $KT_{D.arası}$ ve $KT_{B.arası}$'dir. $KT_{B.arası}$ blokların kareleri toplamıdır. KT_G ve $KT_{D.arası}$ daha önce verilen formülerle,

$$KT_G = \sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}{}^2 - N\overline{\overline{X}}^2$$

$$KT_{D.arası} = \sum_{i=1}^{k} n_i \overline{X}_i^2 - N\overline{\overline{X}}^2$$

$KT_{Blok\ arası}$ ise,

$$KT_{B.arası} = \sum_{i=1}^{k}\sum_{j=1}^{n} (\overline{X}_j^2 - \overline{\overline{X}})^2 = \sum_{j=1}^{n} \overline{X}_j^2 - nk\overline{\overline{X}}^2$$

olarak bulunacaktır. Burada KT_{Hata} daha önce verilen toplamlar yardımı ile,

$$KT_{Hata} = KT_G - KT_{D.arası} - KT_{B.arası}$$

olarak belirlenebilir. Hesaplanan bu kareler toplamları ile ortalama kareler (MS) şöyle hesaplanacaktır.

$$KO_{D.arası} = \frac{KT_{D.arası}}{k-1}$$

$$KO_{B.arası} = \frac{KT_{B.arası}}{n-1}$$

$$KO_{hata} = \frac{KT_{Hata}}{(k-1)(j-1)}$$

$KO_{D.arası}$ ve $KO_{B.arası}$ için iki F test istatistiği ise,

$$F = \frac{KO_{D.arası}}{KO_{hata}}$$

$$F = \frac{KO_{B.arası}}{KO_{Hata}}$$

olacaktır. Bu test istatistiklerinin birincisi için S_{D1}=(k-1) ve S_{D2}=(k-1)(n-1); ikincisi için S_{D1}=(j-1) ve S_{D2}=(k-1)(n-1) serbestlik dereceleri ve belirlenen α hata payı ile F dağılımı tablosundan bulunacak değerler ile karşılaştırılarak karar verilir. Test hipotezlerde yeralan α_i ve β_j'ler için ayrı ayrı yapılmış olacaktır.

Bu durumda oluşacak varyans analizi tablosu Tablo 12.3.'te görülmektedir.

Tablo 12.3

Değişim Kaynağı	Kareler Toplamı	Serbestlik Derecesi	Kare Ortalaması	F
Gruplar Arası (A)	$KT_{D.arası}$	k-1	$KO_{D.arası}$	$F = \frac{KO_{D.arası}}{KO_{hata}}$
Bloklar Arası (B)	$KT_{B.arası}$	n-1	$KO_{B.arası}$	$F = \frac{KO_{B.arası}}{KO_{Hata}}$
Örnekleme Hatası(E)	KT_H	(k-1)(j-1)	KO_H	
TOPLAM (T)		N-1		

Yukarıdaki tablodan da görüldüğü gibi tek yönlü varyans analizi tablosuna SSB'nin eklenmesi ile çift yönlü varyans analizi tablosu oluşturulmuştur. Bu tablo etkileşim olmaması durumunda çift yönlü varyans analizi tablosudur. Etkileşim olması durumunda tablo değişecektir.

ÖRNEK: Liselerin A,B,C,D ve E bölümlerinden mezun üç öğrencinin 100 üzerinden 3 dersten aldıkları final notları aşağıda verilmiştir.

	Dersler		
Bölümler	**Matematik**	**İstatistik**	**Bilgisayar**
A	77	64	87
B	43	51	62
C	86	79	57
D	27	36	48
E	59	42	76

0,05 hata payı ile öğrencilerin final notları ve liselerden mezun oldukları bölümler arasında fark olup olmadığını test ediniz.

		Dersler(i)			
Bölümler (j)	**Matematik**	**İstatistik**	**Bilgisayar**	**Toplam**	**Ortalama ($\overline{X}_j$)**
A	77	64	87	228	76
B	43	51	62	156	52
C	86	79	57	222	74
D	27	36	48	111	37
E	59	42	76	177	59
Toplam	292	272	330	894	
Ortalama ($\overline{X}_i$)	58,4	54,4	66		$\overline{\overline{X}} = 59,6$

1-

$H_1 : \alpha_i \neq 0$

$\beta_j \neq 0$

$\alpha = 0{,}05$

$F_{\alpha,k-1,(k-1)(n-1)} = F_{0,05,2,8} = 4{,}46$

$F_{\alpha,j-1,(k-1)(n-1)} = F_{0,05,4,8} = 3{,}84$

2- $$KT_G = \sum_{i=1}^{3}\sum_{j=1}^{5} X_{ij}^{\ 2} - N\overline{\overline{X}}^2$$

$$\sum_{i=1}^{3}\sum_{j=1}^{5} X_{ij}^{\ 2} = (77)^2 + (43)^2 + (86)^2 + (27)^2 + (59)^2 + (64)^2 + (51)^2 + (79)^2 + (36)^2 + (42)^2 + (87)^2 + (62)^2 + (57)^2 + (48)^2 + (76)^2 = 58124$$

$KT_G = 58124 - 15(59{,}6)^2 = 4841{,}6$

$$KT_{D.arası} = \sum_{i=1}^{3} n\overline{X}_i^2 - N(\overline{\overline{X}})^2$$

$$= \left[\left(\frac{292}{5}\right)^2 + \left(\frac{272}{5}\right)^2 + \left(\frac{330}{5}\right)^2 5\right] - 15(59{,}6)^2$$

$$= 53629{,}6 - 53282{,}4 = 347{,}2$$

$$KT_{B.arası} = \sum_{j=1}^{5} \overline{X}_j^2 - nk(\overline{\overline{X}})^2$$

$$= \left[\frac{(228)^2}{3} + \frac{(156)^2}{3} + \frac{(222)^2}{3} + \frac{(111)^2}{3} + \frac{(177)^2}{3}\right] - 15(59{,}6)^2$$

$$= 56418 - 53282{,}4 = 3135{,}6$$

$KT_{Hata} = KT_G - KT_{D.arası} - KT_{B.arası}$

$= 4841{,}6 - 347{,}2 - 3135{,}6 = 1358{,}8$

$$KO_{D.arası} = \frac{KT_{D.arası}}{k-1} = \frac{347{,}2}{3-1} = 173{,}6$$

$$KO_{B.arası} = \frac{KT_{B.arası}}{n-1} = \frac{3135,6}{5-1} = 783,9$$

$$KO_{hata} = \frac{KT_{Hata}}{(k-1)(j-1)} = \frac{1358,8}{(3-1)(5-1)} = 169,85$$

$$F = \frac{KO_{D.arası}}{KO_{hata}} = \frac{173,6}{169,85} = 1,02$$

$$F = \frac{KO_{B.arası}}{KO_{Hata}} = \frac{783,9}{169,85} = 4,61$$

1,02 < 4,46 olduğundan H_0 hipotezi kabul edilir. ($\alpha_i = 0$)

4,61 < 3,84 olduğundan H_1 hipotezi kabul edilir. ($\beta j \neq 0$)

Değişim Kaynağı	Kareler Toplamı	Serbestlik Derecesi	Kare Ortalaması	F
Deneme Arası (A)	347,2	3-1=2	173,6	1,02
Bloklar Arası (B)	3135,6	5-1=4	783,9	4,61
Örnekleme Hatası(E)	1358,8	(3-1)(5-1)=8	169,85	
TOPLAM	4841,6			

12.3.2. Etkileşimli Çift Yönlü Varyans Analizi

Çift yönlü varyans analizinde, iki faktör arasında karşılıklı etkileşim de olabilir. Deneme grupları ile bloklar arasındaki etkiyi de modele dahil edersek, doğrusal çift yönlü varyans analizi modeli,

$$X_{ijh} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijh} \qquad i = 1,2,3,.....,k$$

$$j = 1,2,3,.....,n$$

$$h = 1,2,3,.....,r$$

olacaktır. Burada,

- r : bloklarla deneme grupları birleşiminin sayısını
- Xijh : gözlemleri
- μ : büyük anakütle ortalamasını
- α_i : i. deneme etkisini
- β_j : j. blok etkisini
- γ_{ij} : i. işleyim ile j. deneme arasındaki etkiyi (etkileşimi)
- ε_{ijh} : hata terimini

ifade etmektedir. Gözlemler ve hata teriminin dağılımı normal dağılımdır.

Temel hipotez oluşturulurken etkileşim olmaması durumundan farklı olarak faktörler arası etkileşiminde dikkate alınması gerekmektedir. Burada temel hipotez deneme etkilerinin, blok etkilerinin ve karşılıklı etkileşim sıfır olduğunu belirtecektir.

$$H_0 : \alpha_i = 0 \qquad i = 1,2,3,......,k$$

$$H_0 : \beta_j = 0 \qquad j = 1,2,3,.....,n$$

$$H_0 : \gamma_{ij} = 0 \qquad i = 1,2,3,....,k \quad ; \quad j = 1,2,3,.....,n$$

Karşıt hipotez ise temel hipotezin doğru olmadığını belirtecek şekilde oluşturulacaktır.

$$H_1 : \alpha_i \neq 0$$

$$H_1 : \beta_j \neq 0$$

$$H_1 : \gamma_{ij} \neq 0$$

Karşıt hipotez α_i ve β_j için en az bir değerin, γ_{ij} için en az bir çiftin değerinin sıfırdan farklı olduğunu ifade etmektedir.

Örnek gözlemlerinin genel örnek ortalamasından farkları,

$$(X_{ijh} - \overline{\overline{X}}) = \left(X_{ijh} - \overline{X}_{ij}\right) + (\overline{X}_i + \overline{\overline{X}}\Big) + \left(\overline{X}_j - \overline{\overline{X}}\right) + \left(\overline{X}_{ij} - \overline{X}_i - \overline{X}_j + \overline{\overline{X}}\right)$$

olacaktır. Sapmaların kareleri toplamı alınırsa,

$$\sum_{i=1}^{k}\sum_{j=1}^{n}\sum_{h=1}^{r}\left(X_{ijh} - \overline{\overline{X}}\right)^2 = \sum_{i=1}^{k}\sum_{j=1}^{n}\sum_{h=1}^{r}(X_{ijh} - \overline{X}_{ij})^2 + nr\sum_{i=1}^{k}\left(\overline{X}_i - \overline{\overline{X}}\right)^2$$

$$+ kr\sum_{j=1}^{n}\left(\overline{X}_j - \overline{\overline{X}}\right)^2 + r\sum_{i=1}^{k}\sum_{j=1}^{n}\left(\overline{X}_{ij} - \overline{X}_i - \overline{X}_j + \overline{\overline{X}}\right)^2$$

olacaktır. Burada $\overline{X}_i$, i. işleyim grubunun; $\overline{X}_j$, j. bloğun gözlemlerinin ortalamasıdır. $\overline{X}_{ij}$ ise i. işleyim grubu ile j. bloğun r sayıda gözleminin ortalamasıdır. $\overline{\overline{X}}$ büyük örnek ortalaması, yani (knr) gözlemin ortalamasıdır. Eşitliğin sol tarafı KT_G, sağ tarafı ise sırasıyla KT_H, $KT_{D.arası}$, $KT_{B.arası}$ ve etkileşim için tanımlanan $KT_{etkileşim}$'dır.

Etkileşim olması durumunda çift yönlü varyans analizinde de KT_G, $KT_{D.arası}$, $KT_{B.arası}$ daha önce açıklandığı gibi hesaplanır. $KT_{etkileşim}$ ise,

$$KT_{etkileş} = \frac{1}{r}\sum_{i=1}^{k}\sum_{j=1}^{n}\left(\sum_{h=1}^{r} X_{ijh}\right)^2 - \frac{\left(\sum_{i=1}^{k}\sum_{j=1}^{n} X_{ij}\right)^2}{N} - SSA - SSB$$

olacaktır. KT_H yine toplamlardan hareketle,

$KT_H = KT_G - KT_{D.arası} - KT_{B.arası} - KT_{etkileşim}$

olarak bulunabilir.

Hesaplanan bu kareler toplamı ile ortalama kareler (MS) şöyle hesaplanacaktır.

$$KO_D = \frac{KT_D}{k-1}$$

$$KO_B = \frac{KT_B}{n-1}$$

$$KO_E = \frac{KT_E}{(k-1)(n-1)}$$

$$KO_H = \frac{KT_H}{kn(r-1)}$$

KO_D, KO_B ve KO_E için üç F test istatistiği ise,

$$F = \frac{KO_D}{KO_H}$$

$$F = \frac{KO_B}{KO_H}$$

$$F = \frac{KO_E}{KO_H}$$

olacaktır. Bu test istatistiklerinin birincisi için S_{D1}=(k-1) ve S_{D2}=kn(r-1); ikincisi için S_{D1}=(n-1) ve S_{D2}=kn(r-1); üçüncüsü için S_{D1}=(k-1)(n-1) ve S_{D2}=kn(r-1) serbestlik dereceleri ve belirlenen α hata payı ile F dağılımı tablosundan bulunacak değerler ile karşılaştırılarak karar verilir. Test, hipotezlerde yeralan α_i, β_j ve γ_{ij}'ler için ayrı ayrı yapılmış olacaktır. Bu durumda oluşacak varyans analizi tablosu Tablo 11.4.'te görülmektedir.

Tablo 12.4.

Değişim Kaynağı	**Kareler Toplamı**	**Serbestlik Derecesi**	**Kareler Ortalaması**	**F**
Gruplar Arası (A)	KT_D	k-1	KO_D	$F = \frac{KO_D}{KO_H}$
Bloklar Arası (B)	KT_B	n-1	KO_B	$F = \frac{KO_B}{KO_H}$
Gruplar Arası ve Bloklararası Etkileşim(I)	KT_E	(k-1)(n-1)	KO_E	$F = \frac{KO_E}{KO_H}$
Örnekleme Hatası(E)	KT_H	Kn(r-1)	KO_H	
TOPLAM	KT_G	N-1		

Görüldüğü gibi etkileşim olmaması durumu ile etkileşim olması durumunda çift yönlü varyans analizi arasındaki fark, faktörler arası etkileşimin üçüncü F istatistiği ile test edilmesidir.

ÖRNEK: Bir önceki örnekte 3 öğrenci için verilen notlar, 6 öğrenci için aşağıda verilmiştir.

Bölümler (j)	Matematik	Dersler(i) İstatistik	Bilgisayar
A	77	64	87
	81	70	89
B	43	51	62
	55	40	70
C	86	79	57
	91	80	87
D	27	36	48
	51	42	63
E	59	42	76
	67	54	62

0,05 hata payı ile tüm farkları test ediniz.

Bölümler (j)	Matematik	Dersler(i) İstatistik	Bilgisayar	Toplam	Ortalama
A	77	64	87	468	78
	81	70	89		
B	43	51	62	321	53,5
	55	40	70		
C	86	79	57	480	80
	91	80	87		
D	27	36	48	267	44,5
	51	42	63		
E	59	42	76	360	60
	67	54	62		
Toplam	637	558	701	1896	
Ortalama	63,7	55,8	70,1		$\bar{\bar{X}} = 63,2$

1-

$$H_1 : \alpha_i \neq 0$$
$$: \beta_j \neq 0$$
$$: \gamma_{ij} \neq 0$$

2- $F_{\alpha,k-1,nk(r-1)} = F_{0,05,2,15} = 3,68$

$F_{\alpha,n-1,nk(r-1)} = F_{0,05,4,15} = 3,06$

$F_{\alpha,(k-1),(n-1),nk(r-1)} = F_{0,05,8,15} = 2,64$

3-

$$KT_G = \sum_{i=1}^{k}\sum_{j=1}^{n}\sum_{h=1}^{r} X_{ijh}^2 - N(\overline{\overline{X}})^2$$

$$= (77)^2 + (81)^2 + \dots\dots\dots + (62)^2 - 30(63,2)^2$$

$$= 128864 - 30(63,2)^2 = 9036,8$$

$$KT_D = \sum_{i=1}^{3} n\overline{X}_i^2 - N\left(\overline{\overline{X}}\right)^2$$

$$= \frac{(637)^2}{10} + \frac{(558)^2}{10} + \frac{(701)^2}{10} - 119827,2$$

$$= 120853,4 - 119827,2 = 1026,2$$

$$KT_B = \sum_{j=1}^{n} \overline{X}_j^2 - nk\left(\overline{\overline{X}}\right)^2$$

$$= \frac{(468)^2}{6} + \frac{(321)^2}{6} + \frac{(480)^2}{6} + \frac{(267)^2}{6} + \frac{(360)^2}{6} - 119827,2$$

$$= 125559 - 119827,2 = 5731,8$$

$$KT_E = \frac{1}{r}\sum_{i=1}^{k}\sum_{j=1}^{n}\left(\sum_{h=1}^{r} X_{ijh}\right) - \frac{\left(\sum_{i=1}^{k}\sum_{j=1}^{n} X_{ij}\right)^2}{N} - SSA - SSB$$

$$= \frac{1}{2}[(77+81)^2 + (43+55)^2 + (86+91)^2 + (27+51)^2 + (59+67)^2 +$$

$$(64+70)^2 + (51+40)^2 + (79+80)^2 + (36+42)^2 + (42+54)^2 +$$

$$(87+89)^2 + (62+70)^2 + (57+87)^2 + (48+63)^2 + (76+62)^2] -$$

$$119827,2 - 2154,2 - 5731,8$$

$$= 127588 - 119827,2 - 1026,2 - 5731,8 = 1002,8$$

$$KT_H = KT_G - KT_D - KT_B - KT_E$$
$$= 9036{,}8 - 1026{,}2 - 5731{,}8 - 1002{,}8$$
$$= 1276$$

$$KO_D = \frac{KT_D}{n-1} = \frac{1026{,}2}{2} = 513{,}1$$

$$KO_B = \frac{KT_B}{n-1} = \frac{5731{,}8}{4} = 1432{,}95$$

$$KO_E = \frac{KT_E}{(k-1)(n-1)} = \frac{1002{,}8}{8} = 125{,}35$$

$$KO_H = \frac{KT_H}{kn(r-1)} = \frac{1276}{15} = 85{,}06$$

$$F = \frac{KO_D}{KO_H} = \frac{513{,}1}{85{,}06} = 6{,}03$$

$$F = \frac{KO_B}{KO_H} = \frac{1432{,}95}{85{,}06} = 16{,}84$$

$$F = \frac{KO_E}{KO_H} = \frac{125{,}35}{85{,}06} = 1{,}47$$

6,03 > 3,68 olduğundan H_1 hipotezi kabul edilir.α_i'ler anlamlı.($\alpha_i \neq 0$)
16,84 > 3,06 olduğundan H_1 hipotezi kabul edilir.β_j'ler anlamlı.($\beta_j \neq 0$)
1,47 <2,64 olduğundan H_0 hipotezi kabul edilir.γ_{ij}'ler anlamsız.($\gamma_{ij} = 0$)

Değişim Kaynağı	**Kareler Toplamı**	**Serbestlik Derecesi**	**Kareler Ortalaması**	**F**
Deneme Arası (A)	1026,2	3-1=2	513,1	6,03
Deneme Arası (B)	5731,8	5-1=4	1432,95	16,84
Gruplar Arası Ve Bloklararası Etkileşim(I)	1002,8	(3-1)(5-1)=8	125,35	1,47
Örnekleme Hatası(E)	1276	(3)(5)(2-1)=15	85,06	
TOPLAM	9036,8			

12.4. ÇÖZÜMLÜ ÖRNEKLER

ÖRNEK 1: Bir paketleme fabrikası meyva sularının paketlenmesi için üç ayrı teknolojiye sahip dolum makinesi kullanmaktadır. Yapılacak yeni yatırım için teknoloji seçimi yapılacaktır. Kullanılan teknolojinin dolum ağırlıkları üzerinde etkinliği incelenmek istenmektedir. Standart net 1 litre olması gereken paketlerden her makineden 5 adet seçilmiştir ve ağırlık ml olarak ölçülmüştür. %95 güven düzeyinde test ediniz.

Makine1	Makine2	Makine3
1007	995	1002
1001	900	996
1009	1011	1000
996	1000	998
1002	1019	1004
5015	4925	5000

Çözüm:

$$\sum_{j=1}^{5} X_{1j} = 5015 \quad , \quad \sum_{j=1}^{5} X_{2j} = 4925 \quad , \quad \sum_{j=1}^{5} X_{3j} = 5000$$

$$\overline{X}_1 = \frac{5015}{5} = 1003 \quad , \quad \overline{X}_2 = \frac{4925}{5} = 985 \quad , \quad \overline{X}_3 = \frac{5000}{5} = 1000$$

şeklinde örnek ortalamaları hesaplanır. Genel ortalama ise ,

$$\sum_{i=1}^{3}\sum_{j=1}^{5} X_{ij} = \sum_{i=1}^{3} \overline{X}_i n_i = 14940 \quad , \quad \overline{\overline{X}} = \frac{14940}{15} = 996$$

Toplam Varyans : $$S_t^2 = \frac{\sum_{i=1}^{3}\sum_{j=1}^{5}\left(X_{ij} - \overline{\overline{X}}\right)^2}{n-1} = \frac{10458}{14} = 747$$

Gruplar Arası Varyans: $$S_b^2 = \frac{\sum_{i=1}^{3} n_i\left(\overline{X}_i - \overline{\overline{X}}\right)^2}{k-1} = \frac{930}{2} = 465$$

Gruplar İçi Varyans : $$S_w^2 = \frac{\sum_{i=1}^{3}\sum_{j=1}^{5}\left(X_{ij} - \overline{X}_i\right)^2}{n-k} = \frac{9528}{12} = 794$$

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

H_1 : Ortalamaların en az biri diğerlerinden farklıdır.

Red bölgesi $F > F_{v1,v2,\alpha}$

Test istatistiği $F = \dfrac{S_b^2}{S_w^2} = 0{,}586 < F_{2,12,0,05} = 3{,}88$

Ortalamaların farklılığını gösterecek yeterli kanıt bulunamamıştır.

ÖRNEK 2: Ekonometri, Maliye ve İktisat bölümlerinde verilen "İstatistiğe Giriş" derslerindeki başarının okunulan bölüme göre farklılık gösterip göstermediği incelenecektir. Bölümlerden öğrenci mevcuduna göre %5 oranında örneklemler oluşturulmuş ve başarı notları listelenmiştir. %1 anlamlılık düzeyinde test ediniz.

Ekonometri	**Maliye**	**İktisat**
20	10	10
70	40	20
80	45	25
90	65	45
	90	50
		80
		85

Çözüm:

Grup(i)	$\sum_{j=1}^{n_i} X_{ij}$	n_i	$\overline{X}_i$	$\sum_{j=1}^{n_i} (X_{ij})^2$	$n_i(\overline{X}_i - \overline{\overline{X}})^2$
1	260	4	65	19800	722,27
2	250	5	50	16050	12,21
3	315	7	45	19275	301,46
Toplam	825	16	51,5625	55125	1035,94

$$S_w^2 = \frac{\sum_{i=1}^{3}\left[\sum_{j=1}^{n_i}(X_{ij})^2 - n_i(\overline{X}_i)^2\right]}{n-k} = \frac{11550}{16-3} = 888{,}46$$

$$\overline{\overline{X}} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}}{n} = \frac{825}{16} = 51{,}5625$$

$$S_b^2 = \frac{\sum_{i=1}^{k} n_i(\overline{X}_i - \overline{\overline{X}})^2}{k-1} = \frac{1035{,}94}{3-1} = 517{,}97$$

$$S_t^2 = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(X_{ij}\right)^2 - \overline{\overline{X}}\sum_{i=1}^{k}\sum_{j=1}^{n_i}X_{ij}}{n-1} = \frac{12585{,}94}{15} = 839{,}06$$

Test istatistiği

$$F = \frac{S_b^2}{S_w^2} = 0{,}583 < F_{2,13,0,01} = 6{,}70$$

Sıfır hipotezinin reddi için yeterli kanıt yoktur. Bölümlere göre başarı notları ortalamalarında anlamlı bir farklılık yoktur.

ÖRNEK 3: Bir ihracat firması 3 ayrı ülkeye yaptığı ihracatlar arasında miktar açısından farklılık olup olmadığını incelemek istiyor. Bu amaçla üç ayrı ülkeye yapılan ihracat miktarları aşağıdaki gibi toparlanmıştır. Farklılık olup olmadığını 0,01 anlam düzeyinde inceleyiniz.

	A Ülkesi	B Ülkesi	C Ülkesi
	10	50	10
	12	51	15
	17	40	14
	18	30	17
	15		18
			16
$\sum X_{ij}$	72	171	90
$\overline{X}_{.j}$	14,4	42,75	15

Çözüm:

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

H_1 : Ortalamaların en az biri diğerlerinden farklıdır

$\overline{X}_{..} = \frac{333}{15} = 22{,}2$

$KT_{GENEL} = (10-21{,}53)^2 + \ldots\ldots\ldots + (16-21{,}53)^2 = 2680{,}4$

$KT_{G.ARASI} = 5(12{,}4-21{,}53)^2 + 4(42{,}75-21{,}53)^2 + 6(15-21{,}53)^2$

$= 2304{,}45$

$KT_{G.İÇİ} = KT_{GENEL} - KT_{G.ARASI} = 2680{,}4-2304{,}45 = 375{,}95$

Kaynaklar	sd.	KT	KO
Gruplar Arası	2	2304,45	
Grup İçi	12	375,95	1,305
Genel	14	2680,4	

$F_{0,01,2,12} = 6,93 > F_{hesap} = 1,305$ H_0 kabul

ÖRNEK 4: Üç ayrı eğitim sisteminin etkileri incelenmek için 5'er öğrenci bu sistemler ile çalışarak bir test sınavına alınmışlardır. Bu sınavdan aldıkları notlar aşağıdaki gibidir. Eğitim sistemleri arasında farklılık olup olmadığını 0,01 anlam düzeyinde inceleyiniz.

	A Sistem	B Sistemi	C Sistemi
	70	50	60
	75	65	70
	80	70	80
	90	40	85
	60	55	90
$\sum X_{ij}$	375	280	385
$\overline{X}_{.j}$	75	56	77

Çözüm:

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

H_1 : Ortalamaların en az biri diğerlerinden farklıdır.

$\overline{\overline{X}} = 69,33$

$$KT_{G.ARASI} = r.\sum_{j=1}^{c}\left(\overline{X}_{.j} - \overline{\overline{X}}\right)^2$$

$$= 5.[(75-69,33)^2 + (56-69,33)^2 + (77-69,33)^2]$$

$$= 1343,3$$

$$KT_{GENEL} = \sum\sum\left(X_{ij} - \overline{X}_{.j}\right)^2$$

$$= (70-75)^2 + (75-75)^2 + ... + (90-77)^2$$

$$= 1650$$

$KT_{G.İÇİ} = KT_{GENEL} - KT_{G.ARASI}$

$= 1650 - 1343,3 = 306,7$

Kaynaklar	**sd.**	**KT**	**KO**
Gruplar Arası	3-1=2	1343,3	671,65
Grup İçi	14-2=12	306,7	25,55
Genel	15-1=14	1650	

$F_{hesap} = 29{,}66$

$F_{0,01,2,12} = 6{,}93 < F_{hesap} = 29{,}66$ H_0 red

En az bir grup diğerlerinden farklıdır.