

**ÇOK DEĞİŞKENLİ VARYANS ANALİZİ VE ÇOKLU
DOĞRUSAL REGRESYON ANALİZİNİN UYGULAMALI
OLARAK KARŞILAŞTIRILMASI**

Burcuhan BİÇKİCİ

**Yüksek lisans Tezi
Zootekni Anabilim Dalı
Prof. Dr. Necati YILDIZ
2007**

Her hakkı saklıdır

**ATATÜRK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

YÜKSEK LİSANS TEZİ

**ÇOK DEĞİŞKENLİ VARYANS ANALİZİ VE ÇOKLU DOĞRUSAL
REGRESYON ANALİZİNİN UYGULAMALI OLARAK
KARŞILAŞTIRILMASI**

Burcuhan BİÇKİCİ

ZOOTEKNİ ANABİLİM DALI

**ERZURUM
2007**

Her Hakkı Saklıdır

ÖZGEÇMİŞ

1983 yılında Erzurum'da doğdu. İlköğrenimini Erzurum İnönü ilkokulunda, orta öğrenimini Erzurum Sabancı orta okulunda, Lise öğrenimini Erzurum Lisesinde tamamladı. 2000 yılında girdiği Atatürk Üniversitesi Ziraat Fakültesi Hayvansal Üretim Zootekni Bölümü'nden 2004 yılında mezun oldu. 2004 yılında Atatürk Üniversitesi Ziraat Fakültesi Biyometri - Genetik Ana Bilim Dalı'nda Yüksek Lisans eğitimine başladı.

ÖZET

Yüksek Lisans Tezi

ÇOK DEĞİŞKENLİ VARYANS ANALİZİ VE ÇOKLU DOĞRUSAL REGRESYON ANALİZİNİN UYGULAMALI OLARAK KARŞILAŞTIRILMASI

Burcuhan BİÇKİCİ

Atatürk Üniversitesi
Fen Bilimleri Enstitüsü
Zootekni Anabilim Dalı

Danışman: Prof. Dr. Necati YILDIZ

Çok değişkenli varyans analizi, tek değişkenli tek faktörlü varyans analizi ve çok değişkenli regresyon analizinin uygulamalı olarak karşılaştırılması amacıyla yürütülen bu çalışmada, Atatürk Üniversitesi Ziraat Fakültesi Tarla Bitkileri Bölümünden alınan veriler kullanılmıştır. Bu verilerde, bağımlı değişken olarak, dane verimi ve bitki boyu bağımsız değişken olarak ta, sıra üzeri ve çeşit bağımlı değişken olarak kullanılmıştır. MANOVA’da elde edilen sonuçlara göre bağımsız değişkenlerin bağımlı değişkenler üzerindeki etkileri bazı test istatistiklerine göre önemli bazı istatistiklere göre önemsiz bulunmuştur. ANOVA’ya göre ise; bağımsız değişkenlerin bağımlı değişkenler üzerindeki etkileri önemsiz bulunmuştur. Çoklu Doğrusal Regresyon Analizi’nde, bağımsız değişkenlerin etkisinin; bitki boyu ve dane verimine olan etkisine önemli olduğu görülmektedir.

2007, 54 sayfa

Anahtar Kelimeler: MANOVA, ANOVA, Çoklu Regresyon Analizi

ABSTRACT

Master Thesis

PRACTICAL COMPARASION OF MULTI-LINEAR REGRESSION ANALYSIS AND MULTI- VARIABLE ANALYSIS OF VARIANCE

Burcuhan BİÇKİCİ

Ataturk University
Graduate School of Natural and Applied Sciences
Department of Zootechnology

Supervisor: Prof. Dr. Necati YILDIZ

In this study carried out with the aim of comparison of Multivariable Variance Analysis, single variable and single factored variance analysis and multiple linear regression analysis practically, The data obtained from Agriculture Faculty department of field plants were used. In this data, the yield of number and height of plant as dependent variable and over row and various dependent variables are used as independent variable. According to the results obtained from MANOVA, the effects of independents variables on dependent ones were significant accordingly to some tests, whereas it was insignificant according to some others. According to ANOVA, the effect of independent variables on dependent ones was found insignificant and according to Multiple Linear Regression Analysis, it is seen that independent variables had impact on plant height and number of yield.

2007, 54 pages

Keywords: MANOVA, ANOVA, Multi-linear regression

TEŐEKKÜR

Yüksek Lisans tezi olarak sunduđum bu alıŐmanın araştırma konusunun belirlenmesi, planlanıp yürütülmesi ve tez haline getirilmesinde bana yol gösteren danışmanım, hocam Sayın Prof. Dr. Necati YILDIZ'a,

alıŐmalarım esnasında gerekli yardımı ve yakın ilgiyi esirgemeyen Zootekni Bölüm Başkanı Sayın Prof. Dr.Hakkı EMSEN, Sayın Prof. Dr. Kemalettin Kara, Sayın Do.Dr. Ömer Cevdet BİLGİN, Sayın Yrd Do.Dr. Mehmet TOPAL'a, Sayın Yrd Do. Dr. MemiŐ ÖZDEMİR'e,

Zootekni Bölümü'nde ki diđer saygıdeđer hocalarıma,

alıŐmalarım esnasında görmüŐ olduđum maddi ve manevi destek ve teşviklerinden dolayı ailemin bütün bireylerine sonsuz teşekkürlerimi sunarım.

Burcuhan BİÇKİCİ

Ađustos 2007

İÇİNDEKİLER

ÖZET.....	i
ABSTRACT.....	ii
TEŞEKKÜR.....	iii
ÇİZELGELER DİZİNİ.....	v
1.GİRİŞ.....	1
2.KAYNAK ÖZETLERİ.....	7
3.MATERYAL ve YÖNTEM.....	10
3.1. Materyal.....	10
3.2. Yöntem	10
3.2.1. Tek değişkenli (ANOVA) ve çok değişkenli (MANOVA) varyans analizi... ..	10
3.2.2. Çok değişkenli analizlerde kullanılan test yöntemleri.....	18
3.2.2.a. Wilks lamda istatistiği.....	18
3.2.2.b. Hotelling-Lawley iz istatistiği.....	19
3.2.2.c. Roy'un en büyük karakteristik kök kriteri.....	20
3.2.2.d. Pillai iz kriteri test istatistiği.....	21
3.2.3. Regresyon analizi.....	23
3.2.3.a Basit doğrusal regresyon yöntemi.....	24
3.2.3.b. Çoklu doğrusal regresyon.....	24
3.2.3.c. Regresyon analizinin varsayımları.....	27
4. ARAŞTIRMA BULGULARI	31
4.2. Çok değişkenli varyans analizi.....	32
4.5. Çoklu doğrusal regresyon analizi	48
5. TARTIŞMA ve SONUÇ.....	50
KAYNAKLAR.....	53
ÖZGEÇMİŞ	

ÇİZELGELER DİZİNİ

Çizelge 1.1.	Tek deęişkenli tek yönlü varyans analiz tablosu.....	3
Çizelge 3.2.1.	Genel kareler ve çarpımlar toplamı matrisi.....	11
Çizelge 3.2.2.	Çok deęişkenli tek yönlü varyans analiz tablosu.....	12
Çizelge 3.2.3.	Çok deęişkenli iki yönlü varyans analiz tablosu.....	14
Çizelge 3.2.4.	Çoklu regresyon modelinde verilerin gösterimi.....	27
Çizelge 3.2.5.	Çoklu regresyonda varyans analiz tablosu.....	29
Çizelge 4.1.	Analizde kullanılan veriler.....	31
Çizelge 4.2.	Çok deęişkenli iki yönlü varyans analiz sonuçları.....	44
Çizelge 4.3.	SPSS’te MANOVA analizi sonuçları.....	46
Çizelge 4.4.	Tek Deęişkenli Varyans Analizi.....	47
Çizelge 4.5.	Çoklu regresyon analizi sonuçları.....	48

1. GİRİŞ

Araştırmalarda birçok problemle karşılaşılabilir, bundan dolayı problemleri ele alırken tüm yönüyle incelemek gerekir. İncelenen olayda çoğu kez bir sonuca tesir eden birden fazla faktör bulunabilmektedir. Bu nedenle incelemeye alınan olayı tüm faktörleri dikkate alarak incelemek, çözüm önerilerini ve sonuçlarını ortaya koymak gerekir (Özdamar 1999). Bu faktörler ve sonuçlar bağımlı ve bağımsız değişken olarak adlandırılır. Bağımlı değişken, bağımsız değişken tarafından açıklanmaya çalışılan değişkendir.

Varyans analizlerinde ele alınan bağımlı değişken sayısı bir tane olduğunda tek değişkenli varyans analizi, birden fazla olduğunda ise çok değişkenli varyans analiz teknikleri kullanılır (Terzi 1995). Genel olarak Varyans Analizi çalışmalarında amaç; sonuca tesir eden faktörlerin veya bağımlı değişkenlerin tesirlerinin büyüklüğünü ve önemlilik durumlarını araştırmaktır.

Varyans Analizi, bağımlı değişkenin gösterdiği değişkenliği bağımsız değişkenler yardımıyla açıklamaktadır. Varyans Analizi'nin doğrudan uygulanması için bazı faraziyelerin geçerli olması gereklidir (Yıldız vd. 2002).

Bu faraziyeler;

- a) Verilerin dağılışı normal dağılışa uygun olmalıdır.
- b) Muamele varyansları homojen olmalıdır.
- c) Analiz modelinde yer alan terimler toplanabilir olmalıdır.
- d) Veriler şans değişkeni olmalı ve örnek yeterli büyüklükte olmalıdır.

Bu şartlar sağlandığında varyans analizi güvenilir bir şekilde yapılabilir.

Uygulamalı araştırmalarda, araştırmaya konu olan faktörlerin incelendikten sonra tekrardan ele alınıp incelenmesi söz konusu olduğundan uygun denemelerle araştırılması yapılmaktadır.

Araştırmaya etki eden faktörlerle varyans analizi yapılır ve bunun sonucunda elde edilen hata ve muamele kareler toplamından yararlanarak F hesap değeri bulunur (Johnson ve Wichern 2002). Uygulanan deneme planı ve veriler ile testin önemliliğine bakıldığından, bu analizin yapımı için varyans analizi gereklidir. Araştırma sonunda incelenen faktörlerde belirlenen tekerrür sayıları eşitse Tek Yönlü Varyans Analizi'nin kullanılması gereklidir. ANOVA sonuçlarından elde edilen veriler homojense, normal dağılış gösteriyorsa varyans analizinin faraziyelerinin ve F istatistiğinin geçerli olduğunu, şayet elde edilen sonuçlar varyans analizinin faraziyelerini karşılamıyorsa F istatistiğinin geçerli ve güvenilir bir test olmadığını ortaya koyar (Arslan ve Mirtaghioğlu 2004).

Varyans analizinde tek ve çok değişkenli modeller kullanılabilir. Varyans Analizi; Tam Şansa Bağlı, Şansa Bağlı Bloklar ve Latinkare gibi deneme planlarına göre yapılmaktadır (Yıldız vd. 2002).

Varyans analizinde, F istatistiği muameleler arasında fark olup olmadığı hakkında bilgi verebilmektedir. Sabit modelde, farkın hangi gruptan geldiği, çoklu karşılaştırmalar yardımıyla bulunabilir. Bunun için geliştirilmiş birçok teknik vardır. Bunlardan bazıları; Scheffe, Tukey, Newman-Keuls, Brown- Forsythe, Duncan testleridir (Kirk 1982).

Varyans analizleri; tek değişkenli ve çok değişkenli varyans analizi olmak üzere ikiye ayrılır.

Tek değişkenli varyans analizi, bağımlı değişken sayısı bir tane olduğunda bu bağımlı değişkenin bağımsız değişken tarafından açıklanmasına denir. Tek değişkenli varyans analizinde amaç; faktör seviyelerinin bağımlı değişken üzerindeki etkilerinin önemliliğini araştırmaktır. Buradaki tek faktör ifadesi, grupları birbirinden ayıran tek özellik olduğu, ya da grupların tek bir değişkeninin değerleri ile açıklandığı anlamına gelir.

Tek değişkenli tek yönlü varyans analizi Çizelge 1.1' de verilen formüllerle yapılmaktadır (Terzi 1995).

Çizelge 1.1. Tek değişkenli tek yönlü varyans analiz tablosu

Varyans Kaynağı	Kareler Toplamı	Serbestlik Derecesi
Muameleler	$KT_{\text{muamele}} = \sum_{i=1}^g n_i (\bar{Y}_i - \bar{Y}_{..})^2$	g-1
Hata	$KT_{\text{hata}} = \sum_{i=1}^g \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2$	$\sum_{i=1}^g n_i - g$
Genel	$KT_{\text{genel}} = \sum_{i=1}^g \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_{..})^2$	$\sum_{i=1}^g n_i - 1$

KT : Kareler toplamı

i : 1,g

g : Toplam örnek sayısı

\bar{Y}_i : i. inci muameleye ait ortalama

$\bar{Y}_{..}$: Genel ortalama

n : Muamele sayısı

Çok Değişkenli Varyans Analizi genelde MANOVA (Multivariate Analysis Of Variance) adı ile bilinirler.

Bilimin ve teknolojinin gelişmesiyle birlikte karmaşık olan problemlerin çözümünde tek değişkenli analizler yeterli olmamaktadır. Tek değişkenli analizlerde araştırmadaki tüm değişkenlerin etkilerinin sabit kabul edilmekte ve her defasında sadece tek bir faktörün analizi yapılmaktadır. Fakat yapılan araştırmalar ve incelemeler sadece tek bir faktörün etkisiyle değil bir çok faktörün etkisi ile oluşmakta ve karmaşık bir yapı göstermektedir. Ancak herhangi bir araştırmada değişken sayıları birden fazla olduğu durumlarda tek

değişkenli varyans analizini kullanmak yeterli olmamakla birlikte işlemlerin uzun sürmesine ve daha çok hatanın olmasına sebep olur. Bundan dolayı tek değişkenli analiz yerine çok değişkenli analizin kullanılması daha doğru olmaktadır (Daşdemir ve Güngör 2002).

Çok değişkenli istatistiksel analiz, incelenen olay ve araştırmada çevresindeki bir çok sayıda iç ve dış faktörleri dikkate alarak, problemin yapısındaki bilgilere göre incelemek ve çözümlere ulaşmak için geliştirilmiş yöntemler bütünüdür (Özdamar 1999). Çok değişkenli teknikler kolayca görülmesi çok güç olan karmaşık ilişkileri açıklar (Anderson vd. 1998).

Çok Değişkenli analizler bilgisayar teknolojisinin gelişmesiyle yaygın bir kullanım alanına sahip olmuştur. Bunun sebebi; işlemlerin daha hızlı yapılarak zamandan tasarrufun sağlanması, yapılan işlemlerin daha doğru ve hassas yapılmasıdır. Çok Değişkenli Varyans Analizi başlangıçta eğitim, psikoloji, sosyoloji gibi alanlarda uygulanmış. Bunun nedeni; bu gibi alanlarda ele alınan konuların tek bir değişkenle ölçülmesinin olanaklı olmamasıdır. Örneğin; çocuk gelişiminin, tek bir faktöre bağlı olarak ölçülmesi ve belli bir değişkenle sınırlandırması doğru değildir. Çocuk gelişimine etki eden birçok faktör bulunmaktadır. Bu faktörleri birlikte incelemek ve bunların tümleşik sonucuna bakmak bize en doğru sonucu vermiş olur.

Daha sonra; özellikle tarım, tıp, biyoloji, ekonomi gibi alanlarda da kullanılmaya başlanmıştır. Çok değişkenli analizler istatistiğin temel dallarından biridir. Birimlerden birden fazla değişkenin alındığı durumlarda veya birimlere birden fazla etkinin bulunduğu durumlarda tek değişkenli varyans analiz teknikleri yeterli olmamaktadır. Bundan dolayı çok değişkenli analizler olayı daha doğru yaklaşımla incelemeye ve etkileri daha gerçekçi olarak ortaya çıkarmaya yardımcı olabilmektedir.

Hipotezlerin tek değişkenli ve çok değişkenli test sonuçları aynı olmayabilir. Örneğin; değişkenler tek tek test edildiğinde hipotez tüm değişkenler için kabul edildiği halde, birden fazla değişken birlikte MANOVA ile test edildiğinde bu hipotezler

reddedilebilir. Bunun sebebi; Tek deęişkenli varyans analizi yapıldığında hata terimi bir tanedir. Ancak MANOVA ile analiz yapıldığında deęişken sayısı ile birlikte analize dahil olan hata terimlerinin de sayısı artmış olacak ve analizdeki hata büyüyecektir. Bu durumdan dolayı; ANOVA ve MANOVA analizlerinin sonuçları arasında fark olabilir.

Çok deęişkenli varyans analizi karmaşık bir matematiksel modele sahip olup, matris cebirini çok geniş bir biçimde kullanılır. MANOVA'yı anlayabilmek için, araştırmaya alınan elemanların uygun bir şekilde tanımlanması yani hangilerinin bağımlı hangilerinin bağımsız deęişken olduğunu anlamamız gerekmektedir (Alpar 1997).

Çok deęişkenli istatistiklerde hipotez testleri çok deęişkenli normal dağılışa dayanılarak yapılmaktadır. Çok deęişkenli normal dağılış ilk kez 1898'de F.Galton adlı bir araştırmacı tarafından geliştirilmiştir. Yani iki deęişkenliden çok deęişkenliliğe geçiş yaparak çok deęişkenli istatistiğin kurucusu olmuştur (Balibeyođlu 1989).

Çok deęişkenli varyans analizi ile ilgili ilk çalışma Wilks'in Genelleştirilmiş Olabilirlik Oranı'dır (Wilks 1932). Daha sonraki yıllarda Tukey ve Bartlett bu konuda çalışmalar yapmışlardır (Bock ve Haggart 1963). Box, λ kriterini gözden geçirmiş Ki-kare dağılımına yaklaşımından yararlanarak ortalamaların testi için bir kriter geliştirmiştir. Bock MANOVA tekniklerini daha sonra davranış bilimlerine uygulamıştır (Box 1949). Birden fazla ölçümün yapıldığı tekrarlı ölçümlerde MANOVA teknikleri Moris ve Zeppa'nın yaptığı araştırmaya uygulanmıştır (Colej ve Grizzle 1966). Gabriel ve Krishnaiah eş zamanlı test istatistiklerinde çok deęişkenli varyans analizi tekniklerini kullanmıştır (Gabriel 1968, Gabriel 1969, Krishnaiah 1969). Pillai, Joyachandran ve Ito MANOVA test kriterlerinin güçleri üzerinde çalışmalar yapmışlardır (Ito 1962, Joyachandran 1967, Pillai 1967). Daha sonraki yıllarda, Hotelling T^2 istatistiğinin dağılımı, örnek sayısı az olduğu durumundaki dağılımı bulunmaya çalışılmıştır (Davis 1970). (MANOVA modelinde, yapılan araştırma ya da deneylerde toplam deney birimi sayısı n ile gösterilir. p ise cevap deęişkenlerinin sayısını gösterir. P cevaplarını her n deney, birimleri ile ölçüldüğünü farz edilerek $n \times p$ cevap matrisine sahip olunur.) Ayrıca

T^2 istatistiğinin deęişik yaklaşımlara göre dağılımı bulunmaya çalışılmıştır (Hugnes ve Saw 1972).

Bu çalışmada, Atatürk Üniversitesi Ziraat Fakültesi Tarla Bitkileri Bölümü'nün ayçiçeęi üzerine yapmış olduęu bir araştırmadan elde edilen veriler kullanılmıştır. Araştırmada bu veriler kullanılarak bitki boyu ve dane verimi üzerinde ayçiçeęi çeşidinin ve sıra üzerinin etkili olup olmadığı araştırılmıştır. Bu amaçla uygun MANOVA modeli kurulup, çok deęişkenli varyans analiz teknikleri yardımıyla tanımlanan bağımsız deęişkenlerin, bağımlı deęişkenler üzerindeki etkileri incelenmiştir. Daha sonra bu veriler dięer çok deęişkenli analiz yöntemi olan Çoklu Regresyon Analizi ile analiz edilmiştir. Çok Deęişkenli Varyans Analizi'nin, Çoklu Regresyon Analizi'nin ve Tek Deęişkenli Varyans Analizi'nin sonuçları karşılaştırılarak deęerlendirilmeleri yapılmıştır. Yapılan analizlerin sonuçları karşılaştırılarak, hangi metodun daha iyi sonuç verdięine bakılmıştır.

2. KAYNAK ÖZETLERİ

Çok deęişkenli analizler istatistięin temel dallarından biridir. Bu analizler; sosyal, saęlık ve fen bilimlerinde çokça kullanılmaktadır. Bu analizlerde baęımlı deęiřkene etki eden faktörler ve bu faktörlerin baęımlı deęiřkene ne kadar etki ettikleri üzerine çeřitli alanlarda farklı arařtırmalar ve çalıřmalar yapılmıřtır.

Balibeyoęlu (1989) çalıřmasında, birden fazla baęımlı deęiřken olduęunda baęımsız deęiřkenin etkisini incelemiřtir. MANOVA modelinin özel hali olan tekrarlı ölçmeler ve profil analizi için MANOVA ve ANOVA yaklařımlarına yer verilmiřtir. Burada yemeklik tanegillerden olan mercimek üretimiyle ilgili çalıřma ele alınmıř ve elde edilen verilerin çözümlemesi MANOVA ile yapılmıřtır.

Terzi'nin (1995) çalıřmasında iki baęımlı deęiřken ele alınmıřtır. Burada baęımlı deęiřkenler çift bakla sayısı ve bakla dal sayısı ve bu baęımlı deęiřkenleri açıklamak için kullanılan baęımsız deęiřkenler ise, tohum ekme biçimi ve m^2 'ye atılan tohum miktarı düşünölmüş. Bunlardan elde edilen verilere çok deęiřkenli varyans analiz teknikleri uygulanmıř. Sonuç olarak, Birinci faktörün baęımlı deęiřkenler üzerine etkisinin olmadığı, ikinci faktörün baęımlı deęiřkenler üzerine etkili olduęu, birinci ve ikinci faktörlerin etkileřiminin baęımlı deęiřkenler üzerinde bazı tekniklere göre etkili olduęu (Hotelling Lawley İz Kriteri) ve dięer tekniklere göre etkili olmadığı (Bartlett'in Ki-kare Yaklařımı, Pillai İz Kriteri, Roy'un En Büyük Özdeęer Kriteri) tespit edilmiş.

Vehid v.d. (2000) çalıřmasında, çok deęiřkenli varyans analizi yöntemini kullanarak allerjik olgularda etken olabilecek anne sütü alım süresi, total IgA ve IgE deęerleri gibi parametreleri birlikte incelemiř ve hastalıęın oluşumu üzerine etkilerini göstermesi amaçlanmıř. Bu amaçla, sistematik örneklemeyle elde edilen yařları 0 - 180 ay arasında deęiřen 1000 allerjik olguya ait verinin deęerlendirildięi çalıřmada, ortak deęiřken olarak anne sütü alım süresinin total IgE ve IgA düzeylerine etkisi, allerjik bronřiyal

astım, allerjik rinit ve atopik dermatit tanısı konmuş olgularda, cinsiyet ve yaş gruplarında çok deęişkenli varyans analizi uygulanmıştır.

Örsel vd. (2004) çalışmasında, 15-17 yaş arasında deęişen 531 lise öğrencisini rastgele örneklem yöntemiyle belirleyerek onlara bir anket uygulanmıştır. Uygulanan ankette, Ergenler İçin Kendilik Algısı Profili (EKAP), Beden Bölgesi ve Özelliklerinden Hoşnut Olma Anketi (BBÖHOA), Diyet Yapma Durumu ve ideal beden ağırlığı ve demografik bilgilere yer verilmiştir. EKAP ve BBÖHOA alt ölçeklerinin puanlarının bağımlı deęişken olarak yer aldığı cinsiyet, yaş, diyet yapma durumunu faktöryel düzende çok deęişkenli varyans analizi ve diyet yapanlarda etkenlerin saptanması amacıyla çoklu bileşik regresyon analizi uygulanmıştır.

Benli'nin (2005) çalışmasında 2002 yılında İMKB'de faaliyet gösteren sanayi işletmeleri bazında, sektörel farklılıkların oranlar üzerinde bir farklılık yaratıp yaratmadığı, eđer fark varsa bu farklılığın hangi oranlardan kaynaklandığı Çok Deęişkenli Varyans Analizi (MANOVA) ile saptanmaya çalışılmıştır. Buna göre sağlıklı verilerine ulaşılan yüz kırk sanayi işletmesinin bilanço ve gelir tablolarından yararlanılarak otuz altı oran hesaplanmıştır. Otuz altı oran kullanılarak, yedi alt sektöre uygulanan faktör analizi ile faktör grupları oluşturulmuştur. Daha sonra her bir faktör kendi altında yoğunlaşan deęişkenler ile ilişkilendirilmiş, negatif ve 0,5'den küçük katsayılara sahip olan deęişkenler analizden çıkartılmıştır. Geriye kalan yirmi yedi deęişken kullanılarak, Çok Deęişkenli Varyans Analizi (MANOVA) yapılmış ve sektörel farklılıkların oranlar üzerinde anlamlı derecede farklı olduğu sonucu çıkarılmıştır.

Çelik (2005) orta öğretim öğrencilerinin okula yabancılaşma düzeylerini, öfke ifade tarzları ve bazı sosyo-demografik deęişkenleri açısından incelenmiştir. Araştırmada cinsiyet, okulun içinde bulunduğu sosyo-ekonomik düzey ve öfke ifade tarzlarına baęlı yabancılaşma verilerini karşılaştırmak amacıyla çok deęişkenli varyans analizi uygulanmıştır.

Varinli v.d. (2005) çalışmasında, perakende sektöründe çalışan satış elemanlarının etik algılamaları ve bu algılamaların demografik özellikler itibariyle farklılık gösterip göstermediğini araştırmıştır. Araştırma, perakende sektöründe faaliyet gösteren departmanlı bir mağazanın iki farklı şehirde bulunan üç mağazasında çalışan satış elemanları üzerinde yapılmıştır. Toplanan veriler, SPSS istatistik programıyla analiz edilmiştir. Verilerin analizinde, frekanslar, yüzdeler, ortalamalar, standart sapmalar, faktör analizi ve çoklu varyans analizi (MANOVA) kullanılmıştır.

3. MATERYAL ve YÖNTEM

3.1. Materyal

Atatürk Üniversite'si Ziraat Fakültesi Tarla Bitkileri Bölümü'nde 1998 yılında yapılan bir ayçiçeği denemesinin verileri üzerinde uygulama yapılmıştır.

Uygulamada; iki faktör ve iki bağımlı değişken kullanılmıştır. Verilere göre bağımsız değişkenler; Çeşit ve Sıra üzeri olmak üzere iki değişken, bağımlı değişkenler ise; Dane Verimi ve Bitki Boyu olarak belirlenmiştir.

3.2. Yöntem

Çok değişkenli varyans analizi, tek değişkenli varyans analizi ve çoklu regresyon analizi SPSS paket programı kullanılarak yapılmıştır.

3.2.1. Tek değişkenli (ANOVA) ve çok değişkenli (MANOVA) varyans analizi

$$Y = A\theta + E$$

$$Y = \begin{pmatrix} Y_{11} & Y_{12} & \cdots & Y_{1n} \\ Y_{21} & Y_{22} & \cdots & Y_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ Y_{p11} & Y_{p2} & \cdots & Y_{pn} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ A_{m11} & A_{m2} & \cdots & A_{mn} \end{pmatrix} \begin{pmatrix} \theta_{11} & \theta_{12} & \cdots & \theta_{1m} \\ \theta_{21} & \theta_{22} & \cdots & \theta_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ \theta_{p11} & \theta_{p2} & \cdots & \theta_{pm} \end{pmatrix} + \begin{pmatrix} \epsilon_{11} & \epsilon_{12} & \cdots & \epsilon_{1n} \\ \epsilon_{21} & \epsilon_{22} & \cdots & \epsilon_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \epsilon_{p11} & \epsilon_{p2} & \cdots & \epsilon_{pn} \end{pmatrix}$$

Modelde;

Y: n x p boyutlu bağımlı değişken matrisi

A: n x m boyutlu katsayı (model) matrisi

θ : m x p boyutlu bilinmeyen parametre matrisi

E: n x p boyutlu hata matrisi

E : p değişkenli ortalaması ve varyansı $\Sigma_{p \times p}$ olan çok değişkenli normal dağılmış bir matristir (Smith vd. 1962).

Modelde verilen A matrisinin satırları birbirinden bağımsız değildir. Bazı satırlar birbirinin kombinasyonudur. θ matrisinin tüm elemanları tahmin edilemez. Örneğin; tek faktör MANOVA modelinde en fazla tahmin edilebilecek parametre sayısı $p \times (m-1)$ kadardır.

Tek yönlü varyans analizi, bir değişken açısından ikiden fazla grup arasında fark olup olmadığını test etmekte kullanılır. Çok değişkenli tek yönlü varyans analizi ise; her bir grupta iki veya daha fazla değişken olması durumunda kullanılır. Diğer bir deyişle ikiden çok grubun ortalama vektörleri karşılaştırılır. Çok değişkenli varyans analizinin genel varsayımları varyans-kovaryans matrislerinin homojen olması ve her bir grubun çok değişkenli normal dağılım göstermesidir (Balibeyoğlu 1989).

Çizelge 3.2.1. Genel kareler ve çarpımlar toplamı matrisi

T	=	B	+	W
Genel Kareler		Gruplar Arası		Hata Kareler Ve
Ve Çarpımlar	=	Kareler Ve Çarpımlar	+	Çarpımlar
Toplamı Matrisi		Toplamı Matrisi		Toplamı Matrisi

Kareler ve çarpımlar toplamı matrisleri şu şekilde ifade edilebilir:

$$W = \sum_{i=1}^g \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_{i.})(Y_{ij} - \bar{Y}_{i.})'$$

$$= (n_1 - 1)S_1 + (n_2 - 1)S_2 + \dots + (n_g - 1)S_g$$

S_1 , Birinci örneğin kovaryans matrisidir.

$$H_0 : a_1 = \dots = a_g$$

H_1 : En az iki a_i değeri birbirinden farklıdır.

Hipotezi, muameleler ve hata kareler ve çarpımlar toplamı matrisleri dikkate alınarak test edilir.

Çizelge 3.2.2. Çok değişkenli tek yönlü varyans analiz tablosu

Varyans Kaynağı	Kareler ve Çarpımlar Toplamı Matrisleri	Serbestlik Derecesi
Gruplararası kareler toplamı ve çapraz çarpımlar	$B = \sum_i \sum_j np (Y_{ij} - \bar{Y}_{...})(Y_{ij} - \bar{Y}_{...})'$	$g-1$
Grup içi kareler toplamı ve çapraz çarpımlar	$W = \sum_i \sum_j \sum_k (Y_{ijk} - \bar{Y}_{ij.})(Y_{ijk} - \bar{Y}_{ij.})'$	$\sum_{i=1}^g n_i - g$
Genel	$T = B + W = \sum_i \sum_j \sum_k (Y_{ijk} - \bar{Y}_{...})'$	$\sum_{i=1}^g n_i - 1$

Çok değişkenli varyans oranı:

$$\lambda^x = \frac{|W|}{|B + W|} = \frac{\left| \sum_{i=1}^g \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)(Y_{ij} - \bar{Y}_i)' \right|}{\left| \sum_{i=1}^g \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y})(Y_{ij} - \bar{Y})' \right|} \quad \text{değeri çok küçük ise}$$

$H_0 : a_1 = a_2 = \dots = a_g$ hipotezi reddedilir. Büyük örnekler için Bartlett'in ki-kare yaklaşımı kullanılır (Terzi 1995).

Çok değişkenli iki yönlü varyans analizi de iki bağımsız değişkenin etkisi altında iki veya daha fazla bağımlı değişkenin olması durumunda kullanılır. Tek değişkenli iki yönlü model genelleştirilirse,

$$Y_{ikr} = \mu + a_i + B_k + \alpha_{ik} + e_{ikr}$$

$$i = 1, 2, \dots, g$$

$$k = 1, 2, \dots, b$$

$$r = 1, 2, \dots, n$$

a_i : i. inci faktörün deneme etkisi

B_k : k.ıncı faktörün deneme etkisi

α_{ik} : i.inci ve k.ıncı faktör arasındaki etkidir.

e_{ikr} : hatanın etkisi. Burada,

$$\sum_{i=1}^g a_i = \sum_{k=1}^b B_k = \sum_{i=1}^g \alpha_{ik} = \sum_{k=1}^b \alpha_{ik} = 0$$

Modeldeki parametre değerleri yerine tahmin değerleri yazılırsa,

$$Y_{ikr} = \bar{Y}_{...} + (\bar{Y}_{i..} - \bar{Y}_{...}) + (\bar{Y}_{.k.} - \bar{Y}_{...}) + (\bar{Y}_{ik.} - \bar{Y}_{i..} - \bar{Y}_{.k.} + \bar{Y}_{...}) + (Y_{ikr} - \bar{Y}_{ik.}) \text{ olur.}$$

Tek değişkenli iki yönlü model genelleştirilirse kareler ve çarpımlar toplamı matrisleri ile serbestlik dereceleri şu şekilde olur:

$$\begin{aligned} & \sum_{i=1}^g \sum_{k=1}^b \sum_{r=1}^n (Y_{ikr} - \bar{Y}_{...})(Y_{ikr} - \bar{Y}_{...})' = \sum_{i=1}^g bn (\bar{Y}_{i..} - \bar{Y}_{...})(\bar{Y}_{i..} - \bar{Y}_{...})' \\ & + \sum_{k=1}^b gn (\bar{Y}_{.k.} - \bar{Y}_{...})(\bar{Y}_{.k.} - \bar{Y}_{...})' \\ & + \sum_{i=1}^g \sum_{k=1}^b n (\bar{Y}_{ik.} - \bar{Y}_{i..} - \bar{Y}_{.k.} + \bar{Y}_{...}) + (\bar{Y}_{ik.} - \bar{Y}_{i..} - \bar{Y}_{.k.} + \bar{Y}_{...}), \\ & + \sum_{i=1}^g \sum_{k=1}^b \sum_{r=1}^n (Y_{ikr} - \bar{Y}_{ik.})(Y_{ikr} - \bar{Y}_{ik.})' \end{aligned}$$

Serbestlik dereceleri ise;

$$gbn-1 = (g-1) + (b-1) + (g-1)(b-1) + gb(n-1) \text{ şeklinde olur.}$$

Çizelge 3.2.3. Çok değişkenli iki yönlü varyans analiz tablosu

Varyans Kaynağı	Kareler ve Çarpımlar Toplamı Matrisleri	Serbestlik Derecesi
Faktör 1	$KTC_{\text{fak.1}} = \sum_{i=1}^g bn (\bar{Y}_{i..} - \bar{Y}_{...}) (\bar{Y}_{i..} - \bar{Y}_{...})'$	$g-1$
Faktör 2	$KTC_{\text{fak.2}} = \sum_{k=1}^b gn (\bar{Y}_{..k} - \bar{Y}_{...}) (\bar{Y}_{..k} - \bar{Y}_{...})'$	$b-1$
Karş. Et.	$KTC_{\text{etk.}} = \sum_{i=1}^g \sum_{k=1}^b n (\bar{Y}_{ik.} - \bar{Y}_{i..} - \bar{Y}_{..k} + \bar{Y}_{...}) + (\bar{Y}_{ik.} - \bar{Y}_{i..} - \bar{Y}_{..k} + \bar{Y}_{...})'$	$(g-1)(b-1)$
Hata	$KTC_{\text{hata}} = \sum_{i=1}^g \sum_{k=1}^b \sum_{r=1}^n (Y_{ikr} - \bar{Y}_{ik.}) (Y_{ikr} - \bar{Y}_{ik.})'$	$gb(n-1)$
Genel	$KTC_{\text{genel}} = \sum_{i=1}^g \sum_{k=1}^b \sum_{r=1}^n (Y_{ikr} - \bar{Y}_{...}) (Y_{ikr} - \bar{Y}_{...})'$	$gbn-1$

KTC : Kareler toplamlar ve çarpımlar matrisi

g : Birinci muamele sayısı

b : İkinci muamele sayısı

i : 1, ..., g

k : 1, ..., b

$\bar{Y}_{i..}$: i . inci muameleye ait ortalama

$\bar{Y}_{...}$: Genel muamele ortalaması

Buna göre hipotez testi şu şekilde olur;

$H_0 = \alpha_{11} = \alpha_{12} = \dots = \alpha_{gb}$ (Karşılıklı etkileşim yoktur.)

$H_1 =$ En az bir tane karşılıklı etkileşim vardır.

$$\lambda_1^* = \frac{|SSP_{hata}|}{|SSP_{tek} + SSP_{hata}|}$$

Buradan da,

$F_1 = [(1 - \lambda_1^*) / \lambda_1^*] \frac{[gb(n-1) - p + 1]/2}{[(g-1)(b-1) - p + 1]/2}$ elde edilir. Bu test istatistiği hesaplanır,

serbestlik derecesi ile F tablo değeri ile mukayese edilir (Alpar 1997).

Tek Değişkenli Varyans Analizi (ANOVA) ile Çok Değişkenli Varyans Analizi (MANOVA) arasında belirgin benzerlikler vardır. Tek Değişkenli Varyans Analizi'nde denemeye tesir eden bileşenler ile hata bileşenleri incelenir.

$$KT_{genel} = KT_{muamele} + KT_{hata}$$

Test olarak hata kısmıyla muamele kısımlarının karşılaştırıldığı F oranı kullanılır.

$$F = \frac{KT_{muamele} / S.d._{muamele}}{KT_{hata} / S.d._{hata}}$$

Bağımlı değişkenler arasındaki ilişkinin derecesini belirtmek için muameleleri hesaba katmalıyız. ANOVA'da iki veya daha fazla bağımlı değişkenin kullanılması istenebilir. Bu durumda bağımlı değişkenler arasındaki kareler ve çarpımlar toplamı matrislerinin hesaplanması gerekir. Kareler ve çarpımlar toplamı matrislerini KTC ile gösterilir.

ANOVA'da olduğu gibi MANOVA'da da muamele ve hata etkileri için ayrı ayrı KTC matrisleri hesaplanır (Terzi 1995).

$$GKTC_{genel} = KTC_{muamele} + KTC_{hata}$$

MANOVA modellerinde bağımsız değişken değerleri etkisi birden fazla bağımlı değişken üzerinde gözlenmektedir. Yani Y_1, Y_2, \dots, Y_p tane bağımlı değişken bulunmaktadır. Burada amaç, işlemlerin etkisini p tane değişken üzerinde eş zamanlı olarak incelemektir.

Tek deęişken varyans analizinde kurulabilecek hipotez ANOVA modelinde

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

Buna karşılık MANOVA modeli,

$$H_0 : \begin{bmatrix} \mu_{11} \\ \mu_{21} \\ \cdot \\ \cdot \\ \cdot \\ \mu_{p1} \end{bmatrix} = \begin{bmatrix} \mu_{12} \\ \mu_{22} \\ \cdot \\ \cdot \\ \cdot \\ \mu_{p2} \end{bmatrix} = \dots = \begin{bmatrix} \mu_{1k} \\ \mu_{2k} \\ \cdot \\ \cdot \\ \cdot \\ \mu_{pk} \end{bmatrix} \quad \text{şeklinde ifade edilmektedir.}$$

Hipotezi genel olarak yazılırsa;

$$H_0 : \theta_h = 0$$

$$H_1 : \theta_h \neq 0$$

Örneęin $H_0 : \theta_h = 0$ olabilir. Daha açık biçimde,

$$H_0 : \alpha_1 - \alpha_2 = 0 \text{ yazılabilir.}$$

θ : parametre vektörü

$$\theta = \begin{bmatrix} \theta_0 \\ \theta_h \end{bmatrix} \quad \text{burada } \theta_0, \text{ k-n}_b \text{ ve } \theta_h, \text{ n}_b \text{ dir.}$$

n_b : θ_h hipotezine ilişkin serbestlik derecesidir.

θ nın ortogonal tahmin edicisi U da aynı biçimde ikiye ayrılabilir.

$$U = \begin{bmatrix} U_0 \\ U_h \end{bmatrix} \quad \text{burada da } U_0, \text{ k-n}_b \text{ ve } U_h, \text{ n}_b \text{ dir.}$$

Test etmek istedięimiz hipoteze ait kareler toplamı ve çarpımları toplamı matrisi

$$S_B = \sum U_i U_i' = U_h U_h \text{ idi.}$$

θ ve θ_0 tahmin edicilerinin varyans-kovaryans matrisleri sırasıyla $\hat{\Sigma}$ ve $\hat{\Sigma}_0$ olsun.

$|\hat{\Sigma}| > 0$, $|\hat{\Sigma}_0| > 0$ ve inversi alınabilir matrislerdir.

H_0 hipotezinin doęruluęu altında model

$Y_i = K_0 \theta_0 + E$ yazılabiliyordu. Likelihood olabilirlik oranı yazılırsa,

$$L = \frac{|\Sigma|}{|\Sigma_0|}$$

Eğer H_0 doğru ise $|\Sigma_0|$, $|\Sigma|$ dan daha büyük olacaktır.

$$S_E = (N-k) \hat{\Sigma}_0$$

$$S_E = Y'Y - U'U$$

$$(N-k) \hat{\Sigma}_0 = Y'Y - U_0'U_0 = Y'Y - U_0'U_0 + U_h'U_h - U_h'U_h$$

$$(N-k) \hat{\Sigma}_0 = S_E + S_H$$

$$L = \frac{|S_E|}{|S_E + S_H|} \text{ şeklinde olacaktır (Anderson 1958).}$$

Bu oran genelleştirilmiş likelihood oranı olan λ dır. Bu oran Wilks tarafından geliştirilmiş Box tarafından 1945'de kullanılır hale getirilmiştir (Box 1949, Wilks 1932). λ oranı tek değişkenli varyans analizinde F oranının rolünü oynamaktadır. Box ve daha sonra Anderson tarafından Wilks Lamda kriteri için asimptotik yaklaşımlar verilmiştir (Anderson 1958, Box 1949). Daha sonraları Pillai ve Gupta tarafından Wilks Lamda Oranı'nın yüzdelerik değerlerini veren tablolar geliştirilmiştir. Pearson bu tabloları "Tables For Multivariate Analysis" adlı kitabında toplamıştır. Bu tablolar lamda dağılımının yüzdelerik değerlerini doğrudan vermezler. Bu tablolar p, q, M = n_e-p-q+1 ve önem seviyelerine bağlı olarak C_α(p,q,m) dönüşüm faktörünün değerini vermektedir.

$$W_{\alpha}(n,p,q) = C_{\alpha}(p,q,M) x_{pq}^2(\alpha)$$

$$W = -m \log_e \lambda(n,p,q)$$

Pearson'ın hazırladığı tablolar $\alpha = 0,01$ ve $0,25$ içindir. Fakat orjinal tablolar $\alpha = 0.025$; 0.05 ; $0,10$ için hazırlanmıştır (Balibeyoğlu 1989).

3.2.2. Çok deęişkenli analizlerde kullanılan test yöntemleri

Çok deęişkenli analizlerde ortalama vektörleri arasında fark olup olmadığını incelemekte kullanılan birçok test yöntemi vardır. Bu test yöntemlerine çoklu karşılaştırma testleri denir. Bu testlere ait istatistikleri farklı dağılımlara sahiptir. Ancak bu istatistikler F istatistiğine dönüştürülebilir. İstatistik yazılımlarının çoęu bu istatistiklere ilişkin sonuçları vermektedir. Bu test istatistiklerinin bazıları aşağıda verilmiştir (Alpar 1997).

3.2.2.a. Wilks lamda istatistięi

Bu yöntem Genelleştirilmiş Olabilirlik Oranına dayalıdır. Bu test istatistięi 0–1 arasında deęer alır. Lamda deęeri sıfıra yaklaştıkça H_0 hipotezi reddedilir, 1'e yaklaştıkça H_0 hipotezinin kabul edilir (Terzi 1995). Araştırmaya konu olan deęişkenler için gruplar arasında fark olup olmadığı Wilks Lamda istatistięi ile belirlenir. Bunun için yapılacak olan MANOVA analizinin sonucunda gruplar arasında bir fark varsa grup ortalamalarını eşit olmadığı, şayet anlamlı bir fark bulunmazsa tüm grup ortalamalarının eşit olduğu yani gruplar arasında fark olmadığı söylenebilir (Ünsal ve Güler 2005).

Varyans analizinde Genel Kareler Toplamı = Gruplar Arası KT + Grup İçi KT şeklinde verilen eşitlik, çok deęişkenli varyans analizinde matrislerle aşağıdaki gibi ifade edilir.

$$\Lambda = W / (W+B)$$

Λ : Lamda test istatistięi deęeri

B: Gruplar arası kareler ve çarpımlar toplamı matrisi

W: Grup içi kareler ve çarpımlar toplamı matrisi

Bu yöntemde kritik deęer olarak;

$$F = \frac{(1-\lambda)(\sum n_i - g)}{\lambda(g-1)} \text{ alınır. Büyük örnekler için bu test istatistięinin yerine Bartlett}$$

yaklaşımı tercih edilir. Bartlett yaklaşımı için test istatistięi olarak;

$$L = -[n-1-(p+k)/2] \ln \lambda \text{ eşitlięi kullanılır. Bu ise } p(g-1) \text{ serbestlik dereceli bir ki-kare dağılımı göstermektedir (Tatlídil 1992).}$$

Çok değişkenli çok faktör varyans analizi için test istatistiği ise;

$$\Lambda_k = \frac{|W|}{|W + B_k|}, \quad k = 1, 2, \dots, m \quad \text{olarak ifade edilir.}$$

$m = 2$ alındığında iki faktör varyans analizine ait değerler elde edilir.

$$\Lambda_k = \frac{|W|}{|W + B_k|} \quad \text{ifadesi faktör interaksiyonlarının test edilmesinde kullanılmak}$$

istenirse, bu durumda k inci ve t inci faktör interaksiyonları için

$$\Lambda_{kt} = \frac{|W|}{|W + B_{kt}|}, \quad k < t = 1, 2, \dots, m \quad \text{olarak alınır.}$$

Birinci faktör, ikinci faktör ve interaksiyon için test istatistikleri sırasıyla;

$$F_1 = \frac{(1 - \lambda)(gb(n-1) - p + 1)/2}{\lambda((g-1) - p + 1)/2}$$

$$F_2 = \frac{(1 - \lambda)(gb(n-1) - p + 1)/2}{\lambda((b-1) - p + 1)/2}$$

$$F_{12} = \frac{(1 - \lambda)(gb(n-1) - p + 1)/2}{\lambda((g-1)(b-1) - p + 1)/2}$$

şeklinde olacaktır. Bu eşitlikler ile hesaplanacak olan F istatistik değerleri;

$$1. \text{ faktör için } V_1 = ((g-1) - p + 1) \text{ ve } V_2 = (gb(n-1) - p + 1)$$

$$2. \text{ faktör için } V_1 = ((b-1) - p + 1) \text{ ve } V_2 = (gb(n-1) - p + 1)$$

$$\text{etkileşim için } V_1 = ((g-1)(b-1) - p + 1) \text{ ve } V_2 = (gb(n-1) - p + 1) \text{ serbestlik dereceli}$$

F tablo değerleri ile karşılaştırılır. İstatistik değeri tablo değerinden büyük olduğunda

$H_0 : a_1 = a_2 = \dots = a_g$ şeklinde kurulan H_0 hipotezi reddedilir (Terzi 1995).

3.2.2.b. Hotelling-Lawley iz istatistiği

Hotelling T^2 testi çok değişkenli normal dağılım varsayımına göre kurulan çok değişkenli hipotezlerin test edilmesini amaçlayan bir yöntemdir. Hotelling Student t 'nin çok değişkenli genellemesi olan T^2 istatistiğinin önemliliğini değerlendirmek için bir dağılım ortaya koymuştur ve bu dağılım çok değişkenli hipotezlerin test edilmesinde

kullanılmaktadır. T^2 testi tek deęişkenli hipotezlerin test edilmesinde yararlanılan t testinin çok deęişkenli hipotezler için genellenmiş biçimidir (Özdamar 1999).

Çok deęişkenli tek faktör varyans analizinde λ_i ler BW^{-1} matrisinin öz deęeri olmak üzere (yani, BW matrisinden elde edilen i.karakteristik kök) $T^2 = \sum_{i=1}^s \lambda_i$ şeklinde verilir.

T^2 deęerinin büyüklüğü H_0 hipotezinin reddedileceğini gösterir (Alpar 1997). Örnek sayısı yeterli iken, $nT > \chi_{[p(k-1),\alpha]}^2$ ise ortalama vektörleri arasında fark olduğu söylenir ve bu durumda, $HL = n T^2$ istatistięi $p(k-1)$ serbestlik dereceli ki-kare dağılımı gösterir.

T^2 istatistięini test etmek için F dağılımından da yararlanılır. Buna göre;

$$F = \frac{p(n-p-1)+2}{(k-1)p^2} T^2$$
 dönüşümü yapılması durumunda $(k-1)p^2$ ve $p(n-p-1)+2$ serbestlik dereceli F tablo deęeri kritik deęer olarak kullanılır.

Çok deęişkenli iki faktör analizinde ise, birinci faktör, ikinci faktör ve interaksiyon için λ_i ler $|B_1W^{-1} - \lambda_i| = 0$, $|B_2W^{-1} - \lambda_i| = 0$, $|B_{12}W^{-1} - \lambda_i| = 0$ eşitliklerinden elde edilir. Serbestlik dereceleri ise, 1. faktör için $p(g-1)$, ikinci faktör için $p(b-1)$ ve interaksiyon için $p(g-1)(b-1)$ olur (Terzi 1995).

3.2.2.c. Roy'un en büyük karakteristik kök kriteri

En büyük özdeęer λ_{\max} ile gösterilirse, Roy'un En Büyük Öz Deęer'e dayalı test istatistięi, $\theta = \lambda_{\max} / (1 + \lambda_{\max})$ ile gösterilir. θ de 0 ile 1 arasında deęer alır. Çok Deęişkenli Tek Faktör Varyans Analizi'nde $|B W^{-1} - \lambda_i| = 0$ determinant deęerinden elde edilen p tane öz deęerden en büyüğü seçilir ve bu deęerden test istatistięi θ elde edilir. Burada B gruplar arası kareler ve çarpımlar toplamı matrisi, W ise gruplar içi kareler ve çarpımlar toplamı matrisidir. θ nın 1'e yaklaşması H_0 hipotezinin reddedileceğini gösterir.

Bulunan θ istatistięi,

$$s = \min(k-1, p)$$

$$m = [(k-p)-1] / 2$$

$y = (N-k-p-1) / 2$ olmak üzere S, M, Y parametrelili Heck Grafik (Heck's Chart) değeri ile karşılaştırılır. θ istatistiğinin Heck grafik değerinden büyük olması durumunda ortalama vektörleri arasında fark olduğu söylenir.

k = karşılaştırılacak grup sayısı

p = her gruptaki değişken sayısı

n = toplam denek sayısı olarak belirtilir.

Çok değişkenli iki faktör varyans analizinde ise; birinci faktör, ikinci faktör ve interaksiyon için öz değerler;

$$|B_1 W^{-1} - \lambda_i| = 0 \quad , \quad |B_2 W^{-1} - \lambda_i| = 0 \quad , \quad |B_{12} W^{-1} - \lambda_i| = 0 \text{ şeklindedir.}$$

Heck tablo parametreleri;

1. faktör için;

$$S = \min(g-1, p) \quad , \quad M = (|g - p| - 1) / 2 \text{ ve}$$

$$Y = (N-g-p-1) / 2$$

2. faktör için;

$$S = \min(b-1, p) \quad , \quad M = (|b - p| - 1) / 2 \text{ ve } Y = (N-b-p-1) / 2$$

interaksiyon için;

$$S = \min((g-1)(b-1), p) \quad , \quad M = (|(g-1)(b-1) - p| - 1) / 2 \text{ ve}$$

$$Y = (N-(g-1)(b-1)-p-1) / 2 \text{ olacaktır (Terzi 1995).}$$

3.2.2.d. Pillai iz kriteri test istatistiği

$$T = \sum_{i=1}^s \frac{\lambda_i}{1 + \lambda_i} \text{ ile verilir. Testin incelenmesinde Roy'un}$$

Kök Yöntemi'ndeki gibi S, M ve Y parametreleri bulunur. Daha sonra test istatistiği Heck grafiğinden yararlanarak bulunan tablo değerleriyle karşılaştırılır (Terzi 1995).

Bu test istatistikleri içinde Roy tarafından geliştirilen En Büyük Karakteristik Kök istatistiğinin ayrıcalıklı yeri vardır. En Büyük Kök istatistiği basit kombinasyonların

fonksiyonunun hipotezinde daha güçlüdür ve en büyük karakteristik kök en dar güven aralığını vermektedir (Gabriel 1968).

Roy'un En Büyük Kök Kriteri, Hotelling-Lawley İz İstatistiği, Pillai İz İstatistiği, Wilks Lamda Oranı istatistikleri ile H_0 hipotezinin red edilmesi ile ortalamaların birbirinden farklı olduğunu diğer bir değişle birimlerin farklı gruplardan oluştuğu söylenebilir. En az bir ortalama vektörü farklı bile olsa H_0 hipotezi red edilecektir. Farklılığın hangi değişkenden geldiği güven aralığı ile bulunabilir. Buna göre güven aralığı ile farklı değişkenin belirlenmesi şu şekilde olur;

$H_0 : \mu_1 = \mu_2$ (iki ortalama vektörü arasında fark yoktur.)

$H_1: \mu_1 \neq \mu_2$ şeklinde hipotezler kurulur ve buna göre;

T^2 istatistiği; $T^2 = \frac{n_1 n_2}{n_1 + n_2} (\bar{x}_1 - \bar{x}_2)' S^{-1} (\bar{x}_1 - \bar{x}_2)$ ile verilir.

$n_1 = 1.$ grubun denek sayısı

$n_2 = 2.$ grubun denek sayısı

$\bar{x}_1 = 1.$ grubun ortalama vektörü

$\bar{x}_2 = 2.$ grubun ortalama vektörü

$S =$ Ortak varyans - kovaryans matrisi

Yukarıda verilen T^2 istatistiği;

$F = \frac{n_1 + n_2 - p - 1}{(n_1 + n_2 - 2)p} T^2$ ile p ve $n_1 + n_2 - p - 1$ serbestlik dereceli bir F dağılımı gösterir

ve bunun yardımıyla elde edilen $T^2_{(p, n_1 + n_2 - p - 1; \alpha)} = \frac{(n_1 + n_2 - 2)p}{n_1 + n_2 - p - 1} F_{(p, n_1 + n_2 - p - 1, \alpha)}$ olmak

üzere her bir değişkenin ortalamalarının farkına ilişkin eşanlı güven aralıkları;

$$a' (\bar{x}_1 - \bar{x}_2) - \sqrt{\frac{n_1 + n_2}{n_1 n_2} a' S a T^2_{(p, n_1 + n_2 - p - 1; \alpha)}} \leq a' S \leq a' (\bar{x}_1 - \bar{x}_2) +$$

$$\sqrt{\frac{n_1 + n_2}{n_1 n_2} a' S a T^2_{(p, n_1 + n_2 - p - 1; \alpha)}} \quad \text{ile verilir.}$$

Buradan elde edilen aralıkların sıfırı içermesi durumunda ilgili değişken açısından gruplar arasında fark yoktur denilebilir (Alpar 1997).

3.2.3. Regresyon analizi

Regresyon analizi aralarında sebep sonuç ilişkisi olan iki veya daha fazla değişken arasındaki ilişkiyi incelemek ve o konuyla ilgili tahmin yapabilmek amacıyla oluşturulan ve matematiksel bir model ile belirtilen istatistiksel bir analizdir. Bağımlı değişkenler ile bağımsız değişkenler arasında kurulan istatistiksel modelle bağımsız değişkenlerin belirli değerleri için bağımlı değişkenlerin alacağı değeri tahmin etme yöntemidir (Şahinler 2000). Burada bir bağımlı değişken (Y) ile bir bağımsız değişken (X_1) arasındaki bağıntıyı inceleyen yönteme basit regresyon, bir bağımlı değişken (Y) ile iki ya da daha fazla bağımsız değişken ($X_1, X_2, X_3, \dots, X_p$) arasındaki bağıntıları modeller aracılığı ile inceleyen yönteme ise çoklu regresyon adı verilmektedir (Özdamar 1999). Burada verilen bağımlı değişkenlerin, deney esnasında ardışık olarak eklenerek çoklu regresyon analizini oluşturur. Böylece analizin sonunda gözlemlerin sayısının her eklenen değişkenle azaldığı görülür (Raats v.d. 2005). Kurulan modellerin geçerliliğini analiz için yararlanılan yönteme ise regresyon analizi adı verilmektedir. Örneğin, yaş (x) ile kan basıncı (y), boy (x) ile ağırlık (y) ilacın dozu (x) ile kalp atım hızı (y) arasında bir ilişki var mıdır? Sorularının cevabı regresyon analizi ile bulunabilmektedir (Alpar 1997).

Basit ve çoklu doğrusal regresyon modelleri aşağıdaki gibi ifade edilmektedir.

$$Y = \beta_0 + \beta_1 X_1 + e_{ij} \quad \text{Basit Doğrusal Model}$$

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + e_{ij} \quad \text{Çoklu Doğrusal Model}$$

$$Y_1, Y_2 = f(X_1, X_2, \dots, X_p) \quad \text{Çok Değişkenli Doğrusal Model}$$

Çok Değişkenli Doğrusal Regresyon Analizi cevap değişkeninin iki veya daha fazla olduğu, bağımsız değişkenin ise bir ya da daha fazla olduğu regresyon modellerini ifade etmektedir. Regresyon uygulamaları için verilerin uyması gerekli bazı koşullar

bulunmaktadır. Bu koşullardan en önemlileri; Y ve X'in normal dağılım göstermesi, X'in hatasız ölçümler içermesi, e'nin ise $N(0, \sigma^2)$ parametrelili normal dağılım göstermesi sayılabilir.

3.2.3.a Basit doğrusal regresyon yöntemi

Bir bağımlı değişken (Y) ve birden fazla bağımsız değişken (X) arasındaki fonksiyonel ilişkiyi inceleme yöntemidir.

$Y = a + bX$ biçiminde ifade edilen modele basit doğrusal regresyon adı verilir.

Basit regresyon modelinde bağımlı değişken ile bağımsız değişken arasındaki fonksiyonel bağıntı matris formu ile aşağıdaki gibi gösterilir.

$$Y = X\beta + \varepsilon$$

Bu modelde; Y, (n x 1) boyutlu bağımlı değişken gözlem vektörü; X, (n x (p + 1)) boyutlu bağımsız değişken gözlem matrisi; β , ((p + 1) x 1) boyutlu katsayılar vektörü; ε , (n x 1) boyutlu hata vektörüdür.

Basit Doğrusal Regresyon modeli matris formunda;

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \cdot & \cdot \\ \cdot & \cdot \\ 1 & X_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \cdot \\ \cdot \\ e_n \end{bmatrix} \text{ şeklinde ya da}$$

$Y = X\beta + \varepsilon$ şeklinde yazılır (Özdamar 1999).

3.2.3.b. Çoklu doğrusal regresyon

Diğer bir regresyon analiz yöntemi ise çoklu doğrusal regresyondur. Araştırma ve çalışmaların çoğu çeşitli faktörlere bağlı olarak meydana gelmektedir. Basit regresyonda analizinde bir olayı etkileyen sadece bir faktöre yer verilir. Oysa gerek sosyal, gerek ekonomik, gerek psikolojik v.b. olaylar çok sayıda faktöre bağlı olan olaylarla birlikte değişim göstermektedirler. Bu nedenle basit regresyon yetersiz kalmaktadır. İşte bir

olayı etkileyen birden fazla faktör olması durumunda sebep-sonuç ilişkisinin araştırılması çoklu regresyon analiziyle mümkün olacaktır.

Genelde çoklu regresyon analizi:

- 1- Bağımlı değişkendeki değişimi açıklayabilmek için kullanılır.
- 2- Faktörlerin, diğer faktörlerin etkisi olmaksızın bağımlı değişkene etkilerinin kestirimi için yapılır. Bu kestirim her değişken için bulunur. Bunlardan hangilerinin açıklamada daha önemli oldukları da aynı amaç içinde incelenir.
- 3- Bağımlı değişkene ilişkin ortalama Y ya da \hat{Y} değerlerinin bulunması amacıyla kullanılır (Şıklar 2000).

Bağımlı değişken ile bağımsız değişkenler arasında doğrusal bir bağıntı varsa bu tip regresyona Basit Doğrusal ya da Çoklu Doğrusal Regresyon adı verilmektedir. Çoklu doğrusal regresyonda, bağımlı değişkeni etkileyen birden fazla bağımsız değişken söz konusudur.

Çalışmaların çoğunda bağımlı değişkeni etkileyen birden çok bağımsız değişken söz konusudur. Bu tür çalışmalarda amaç; bağımlı değişkene etki eden birden çok bağımsız değişkenin etkisini incelemek ya da sadece aralarındaki karmaşık yapıyı tanımlamak olabildiği gibi, bağımsız değişkenlerden hangisi ya da hangilerinin bağımlı değişkeni daha çok etkilediğini bulmak ya da bağımsız değişkenler yardımıyla bağımlı değişkeni kestirmek şeklinde de olabilmektedir.

Çoklu doğrusal regresyon modeli; Basit doğrusal regresyon modelinde bir bağımlı bir bağımsız değişken yer almaktaydı. Çoklu regresyon modelinde ise; bir bağımlı ve bağımlı değişkeni etkileyen birden fazla bağımsız değişken yer almaktadır. Örneğin, bir bağımlı değişkeni etkileyen üç bağımsız değişken söz konusu olduğu zaman kestirimlerde kullanılacak regresyon denklemi:

$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$ şeklinde ifade edilmektedir. Örneklem regresyon denklemi ise,

$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + e$ yazılır. Bu eşitlikte,

Y: Bağımlı değişken

X: Bağımsız değişken

ε : Hata terimi olarak tanımlanır. Modelde β 'lar bilinmeyen parametrelerdir. Örneklem regresyon denklemiyle kestirimler elde edilir. b_0 ; sabit terimdir.

Şöyle ki; $X_1 = X_2 = X_3 = 0$ iken Y'nin aldığı ortalama değeri gösterir. b_1, b_2, b_3 ise regresyon katsayılarıdır, örneğin b_2 , X_1 ve X_3 değişkeni aynı kaldığında X_2 değişkeninin bağımlı değişken Y'ye etkisini gösterir. Bu kısmi regresyon katsayısı olarak da isimlendirilir. Böylece b_2 diğer değişkenlerin etkisi artırıldıktan sonra X_2 bağımsız değişkenindeki bir birimlik değişimin bağımlı değişkende yaratacağı değişiklik miktarını verir (Şıklar 2000).

Sonuçta, verilerin genel durumundan ilişkinin şekli tahmin edilmeye çalışılır. Şayet verilere göre diyagram çizilebilirse diyagram üzerinde her gözlem bir nokta ile temsil edildiğinden her bir değişken için noktalar bir doğru etrafında toplanmış gibi görünüyorsa en iyi fonksiyonun doğrusal bir fonksiyon olduğu söylenebilir (Balibeyoğlu 1989).

Çoklu regresyon analizinde; bağımlı değişken y, bağımsız değişkenler x_1, x_2, \dots, x_p ile gösterilişi Çizelge 3.2.4.'de verilmiştir.

Çizelge 3.2.4. Çoklu regresyon modelinde verilerin gösterimi

Gözlem	y	x ₁	x ₂	x ₃ ... x _p
1	y ₁	x ₁₁	x ₁₂	x ₁₃ ... x _{1p}
2	y ₂	x ₂₁	x ₂₂	x ₂₃ ... x _{2p}
3	y ₃	x _{31x}	x ₃₂	x ₃₃ ... x _{3p}
.
.
.
n	y _n	x _{n1}	x _{n2}	x _{n3} ... x _{np}

Çoklu doğrusal regresyonda y ile x_1, x_2, \dots, x_p arasındaki ilişki;

$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon_i$ ile verilir.

Burada; $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ bilinmeyenlerin kısmi regresyon katsayıları ya da kısaca regresyon katsayıları denir. Örneklemeyle ilişkin regresyon denklemi;

$y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \dots + b_p x_{ip} + e_i$ ile verilirken, beklenen regresyon denklemi;

$\hat{y}_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \dots + b_p x_{ip}$ ile verilir.

$e_i = y_i - \hat{y}_i$ olarak yazılır. e_i 'ler çoğunlukla modelin yeterliliğini değerlendirmekte kullanılır.

3.2.3.c. Regresyon analizinin varsayımları

Basit doğrusal regresyonda olduğu gibi çoklu doğrusal regresyonun da varsayımları vardır. Bu varsayımlar:

1- x_i değişkenlerine ilişkin değerler sabittir. (Rastlantı değişkeni değildir.)

- 2- Her bir x_i deęişken kümesi için birden çok y deęeri vardır ve y alt kümeleri normal dağılım gösterir.
- 3- y alt kümelerinin varyansları eşittir.
- 4- y deęerleri bağımsızdır.
- 5- ϵ_i ' nin ortalaması sıfıra eşittir.
- 6- ϵ_i ' nin dağılımı normaldir.
- 7- Hata terimleri arasında ilişki yoktur.
- 8- Hata terimi eşit varyanslıdır.
- 9- Hata terimleri ile bağımsız deęişken arasında ilişki yoktur.
- 10- Modelde belirleme hatası yoktur.
- 11- Bağımsız deęişkenler arasında ilişki yoktur.

Bu varsayımlardan, özellikle son varsayım çoklu modellerde yani birden fazla bağımsız deęişken olan modellerde söz konusu olur (Şıklar 2000).

Bu gerekçeler, özellikle regresyon katsayıları için hipotezin test edilmesi için gereklidir. Ayrıca, bağımsız deęişkenler arasında yüksek derecede ilişki olmaması istenir. Bazı durumlarda ise, bağımsız deęişkenler arasındaki ilişki çok belirgindir. Eğer bir bağımsız deęişken dięer bağımsız deęişken ya da deęişkenlerin bir doğrusal fonksiyonu olarak yazılabiliyorsa, deęişkenler arasında doğrusal bağımlılık söz konusu olacak ve regresyon katsayıları bulunamayacaktır.

$$y_i = b_0 + b_1x_{i1} + b_2x_{i2} + \dots + b_px_{ip} + \epsilon_i$$

$Y = X\hat{\beta} + e$ ile verilir. Burada,

Y : $n \times 1$ boyutlu bağımlı deęişken vektörü

X : $n \times (p + 1)$ boyutlu girdi matrisi olup bu matrisin ilk kolonu 1'lerden, dięer kolonları deęişken deęerlerden oluşur.

$\hat{\beta}$: $(p + 1) \times 1$ boyutlu katsayılar vektörü

e : $n \times 1$ boyutlu hata vektörü olup, matris eşitlięi aşıęıdaki gibidir:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_p \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

Çoklu regresyon denklemi elde edildikten sonra, çeşitli hipotezler test edilebilir. Öncelikle, varyans analizi yapılarak, bağımlı değişkenin bağımsız değişkenler tarafından açıklanıp açıklanamadığı, diğer bir değişle, bağımlı değişkenle bağımsız değişkenler kümesi arasında doğrusal bir ilişki olup olmadığı test edilir. F dağılışı yardımıyla yapılan bu testte H_0 hipotezi, tüm regresyon katsayılarının sıfıra eşit olduğu ($H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$) şeklinde kurulurken, H_1 hipotezi, en az bir β_i nin sıfırdan farklı olduğu şeklinde kurulur. Bilindiği gibi, F değerini bulabilmek için çeşitli kareler toplamları, serbestlik dereceleri ve kareler ortalamalarına gereksinim duyulmakta ve bu sonuçlar varyans analizi tablosu ile özetlenmektedir.

Çizelge 3.2.5. Çoklu regresyonda varyans analiz tablosu

VK	KT	Sd	KO	F
YOA	$Y'Y - n\bar{y}^2$	n-1		RAKO / RAKO
R	$\beta'X'Y - n\bar{y}^2$	p	RKO	
Hata (RA)	$Y'Y - \hat{\beta}'X'Y$	n-p-1	RAKO	

Yukarıdaki tablodaki gibi elde edilen F istatistiği p, n-p-1 serbestlik dereceli F tablo değeri ile karşılaştırılır. H_0 hipotezinin kabul edilmesi durumunda, seçilen α yanılma

düzeyi için, bağımlı değişkenle bağımsız değişkenler arasında doğrusal bir ilişki olmadığı ya da bağımsız değişkenlerin bağımlı değişkendeki değişimi açıklayamadığı sonucuna varılabilir (Alpar 1997).

4. ARAŞTIRMA BULGULARI

Çizelge 4.1. Analizde kullanılan veriler

Çeşit	Sıra Üzeri					
	30		40		50	
	DV	BB	DV	BB	DV	BB
Yağlık	270	135	237	118	179	117
	454	193	364	142	427	143
	384	138	350	148	297	149
	269	136	236	135	223	136
	410	163	337	149	338	174
	400	147	276	123	269	136
	274	141	206	145	194	136
	310	141	258	138	319	147
	177	116	308	142	164	113
	245	142	266	142	169	119
	423	157	230	145	239	146
	303	130	216	130	190	131
	$\sum x$	3919	1739	3284	1657	3008
\bar{x}	326.6	144.9	273.7	138.1	250.7	137.3
$\sum x^2$	1360201	256123	930782	229869	826988	229119
$\sum x_1x_2$	585745		458247		423685	
Çerezlik	471	178	319	150	333	157
	600	169	504	168	327	164
	390	144	302	134	399	166
	423	162	397	158	259	130
	602	179	481	163	454	175
	459	153	378	164	367	162
	349	149	269	140	371	155
	503	176	317	150	394	171
	496	169	212	131	345	157
	235	124	323	168	274	163
	429	167	409	165	315	166
	266	151	417	164	324	168
	$\sum x$	5223	1921	4328	1855	4162
\bar{x}	435.3	160.1	360.7	154.6	346.8	161.2
$\sum x^2$	2170787	310479	1642128	288715	1476084	313134
$\sum x_1x_2$	853211		679230		675145	

Çizelge 4.1’de, sıra üzeri ve çeşide göre bitki boyu ve dane verimine ilişkin veriler verilmiştir. Yukarıdaki tabloda gösterilmiş olan veriler, ayçiçeği denemesinde ele alınan iki ayrı çeşit (yağlık – çerezlik) ve üç ayrı sıra arası (30 - 40 - 50) kullanılmıştır. Bitki boyu ve dane verimine olan veriler belirtilmiştir. Ayrıca her bir sıra arası ve çeşit için

verilen dane verimi ve bitki boyuna ait toplam değerleri, ortalama değerleri, kareler toplamları ve çarpımlar toplamları hesaplanmış olup, bu veriler ile Çok Değişkenli Varyans Analizi, Tek Değişkenli Varyans analizi ve Çoklu Doğrusal Regresyon Analizi yapılmış, elde edilen sonuçların karşılaştırılması yapılmıştır.

DV = Dane Verimi

BB = Bitki Boyu

$\sum x$ = Faktörlere ait toplamlar

\bar{x} = Faktörlere ait ortalama değerler

$\sum x^2$ = Faktörlere ait toplamların karesi

$\sum x_1x_2$ = birinci ve ikinci değişkene ait değerlerin çarpımlar toplamı

Bu çalışmada, Atatürk Üniversitesi Ziraat Fakültesi Tarla Bitkileri Bölümü'nün ayçiçeği üzerine yapmış olduğu bir araştırmadan elde edilen veriler kullanılmıştır. Araştırmada bu veriler kullanılarak bitki boyu ve dane verimi üzerinde ayçiçeği çeşidinin ve sıra üzerinin etkili olup olmadığı araştırılmıştır. Bu amaçla uygun MANOVA modeli kurulup, çok değişkenli varyans analiz teknikleri yardımıyla tanımlanan bağımsız değişkenlerin, bağımlı değişkenler üzerindeki etkileri incelenmiştir.

4.2. Çok Değişkenli Varyans Analizi

Birinci faktör olarak çeşit, ikinci faktör olarak sıra üzeri belirlendi. Çok değişkenli varyans analizini yapabilmek için faktörlere ait ortalama vektörleri ve örnek sayıları hesaplanır. Buna göre,

1. Faktörün 1. düzeyi için ortalama vektörü ve örnek sayısı;

$$\bar{x}_{11} = \begin{bmatrix} (326.6 + 273.7 + 250.7)/3 \\ (144.9 + 138.1 + 137.3)/3 \end{bmatrix} = \begin{bmatrix} 283.6667 \\ 140.1000 \end{bmatrix} \quad n = 36$$

1. Faktörün 2. düzeyi için ortalama vektörü ve örnek sayısı;

$$\bar{x}_{21} = \begin{bmatrix} (435.3 + 360.7 + 346.8)/3 \\ (160.1 + 154.6 + 161.2)/3 \end{bmatrix} = \begin{bmatrix} 380.9333 \\ 158.6333 \end{bmatrix} \quad n = 36$$

2. Faktörün 1. düzeyi için ortalama vektörü ve örnek sayısı;

$$\bar{x}_{12} = \begin{bmatrix} (326.6 + 435.3)/2 \\ (144.9 + 160.1)/2 \end{bmatrix} = \begin{bmatrix} 380.95 \\ 152.50 \end{bmatrix} \quad n = 24$$

2. Faktörün 2. düzeyi için ortalama vektörü ve örnek sayısı;

$$\bar{x}_{22} = \begin{bmatrix} (273.7 + 360.7)/2 \\ (138.1 + 154.6)/2 \end{bmatrix} = \begin{bmatrix} 317.20 \\ 146.35 \end{bmatrix} \quad n = 24$$

2. Faktörün 3. düzeyi için ortalama vektörü ve örnek sayısı;

$$\bar{x}_{32} = \begin{bmatrix} (250.7 + 346.8)/2 \\ (137.3 + 161.2)/2 \end{bmatrix} = \begin{bmatrix} 298.75 \\ 149.25 \end{bmatrix} \quad n = 24$$

Genel ortalama vektörü ve örnek sayısı;

$$\bar{x}_{32} = \begin{bmatrix} 332.3000 \\ 149.3667 \end{bmatrix} \quad n = 72$$

Birinci faktör için kareler ve çarpımlar toplamı matrisi;

$$KT\check{C} \text{ Fak 1} = \sum_{i=1}^g bn (\bar{Y}_{i..} - \bar{Y}_{...}) (\bar{Y}_{i..} - \bar{Y}_{...})'$$

b = Birinci faktörün sayısı

n = Birinci faktörün birinci sütununa karşılık gelen dane verimindeki toplam veri sayısı

$i = 1, \dots, g$

$\overline{Y_{i..}}$ = i . İnci faktöre ait ortalama değer

$\overline{Y_{...}}$ = Genel ortalama

$$\text{Fak1KÇT} = (3)(12) \left\{ \begin{array}{l} \left[\begin{array}{l} 383.6667 - 332.3000 \\ 140.1000 - 149.3667 \end{array} \right] \left[\begin{array}{l} 383.6667 - 332.3000 \\ 140.1000 - 149.3667 \end{array} \right] + \left[\begin{array}{l} 380.9333 - 332.3000 \\ 158.6333 - 149.3667 \end{array} \right] \left[\begin{array}{l} 380.9333 - 332.3000 \\ 158.6333 - 149.3667 \end{array} \right] \end{array} \right\}$$

$$\text{Fak1KÇT} = (3)(12) \left\{ \begin{array}{l} \left[\begin{array}{ll} 4730.3958 & 901.3404 \\ 901.3404 & 171.7434 \end{array} \right] \end{array} \right\}$$

$$\text{Fak1KÇT} = \left[\begin{array}{ll} 340588.4976 & 64896.5088 \\ 64896.5088 & 12365.5248 \end{array} \right]$$

İkinci faktör için kareler çarpımlar ve toplamlar matrisi;

$$\text{KTÇ Fak2} = \sum_{k=1}^b gn (\overline{Y_{..k}} - \overline{Y_{...}})(\overline{Y_{..k}} - \overline{Y_{...}})'$$

$\overline{Y_{..k}}$ = k . İnci faktöre ait ortalama değer

$k = 1, \dots, b$

g = İkinci faktörün sayısı

$$\text{Fak2KÇT} = (2)(12) \left\{ \left[\begin{array}{l} 380.95 - 332.3000 \\ 152.50 - 149.3667 \end{array} \right] \left[\begin{array}{ll} 380.95 - 332.3000 & 152.50 - 149.3667 \end{array} \right] + \right\}$$

$$\begin{bmatrix} 317.20 - 332.3000 \\ 146.35 - 149.3667 \end{bmatrix} \begin{bmatrix} 317.20 - 332.3000 & 146.35 - 149.3667 \end{bmatrix} +$$

$$\begin{bmatrix} 298.75 - 332.3000 \\ 149.25 - 149.3667 \end{bmatrix} \begin{bmatrix} 298.75 - 332.3000 & 149.25 - 149.3667 \end{bmatrix}$$

$$\text{Fak2K}\check{\text{C}}\text{T} = 24 \left\{ \begin{bmatrix} 2366.8225 & 152.4205 \\ 152.4205 & 9.8157 \end{bmatrix} \right\} + \begin{bmatrix} 228.01 & 45.5522 \\ 45.5222 & 9.1005 \end{bmatrix} +$$

$$\begin{bmatrix} 1125.6025 & 3.9153 \\ 3.9153 & 0.0136 \end{bmatrix}$$

$$\text{Fak2K}\check{\text{C}}\text{T} = \begin{bmatrix} 89290.44 & 4845.312 \\ 4845.312 & 454.3152 \end{bmatrix}$$

Hata kareler çarpımlar ve toplamlar matrisi;

$$\text{KTC}_{\check{\text{C}}\text{hata}} = \sum_{i=1}^g \sum_{k=1}^b \sum_{r=1}^n (Y_{ikr} - \overline{Y_{ik.}})(Y_{ikr} - \overline{Y_{i.k}})'$$

$r = 1, \dots, n$

$\overline{Y_{ik.}}$ = ik. inci faktöre ait ortalama değer

$$\text{Hata K}\check{\text{C}}\text{T} = \begin{bmatrix} 270 - 326.6 \\ 135 - 144.9 \end{bmatrix} \begin{bmatrix} 270 - 326.6 & 135 - 144.9 \end{bmatrix} +$$

$$\begin{bmatrix} 454 - 326.6 \\ 193 - 144.9 \end{bmatrix} \begin{bmatrix} 454 - 326.6 & 193 - 144.9 \end{bmatrix} + \begin{bmatrix} 384 - 326.6 \\ 138 - 144.9 \end{bmatrix} \begin{bmatrix} 384 - 326.6 & 138 - 144.9 \end{bmatrix}$$

$$+ \dots \begin{bmatrix} 303 - 326.6 \\ 130 - 144.9 \end{bmatrix} [303 - 326.6 \quad 130 - 144.9] +$$

$$\begin{bmatrix} 237 - 273.7 \\ 118 - 138.1 \end{bmatrix} [237 - 273.7 \quad 118 - 138.1] + \begin{bmatrix} 364 - 273.7 \\ 142 - 138.1 \end{bmatrix} [364 - 273.7 \quad 142 - 138.1]$$

$$+ \dots + \begin{bmatrix} 216 - 273.7 \\ 130 - 138.1 \end{bmatrix} [216 - 273.7 \quad 130 - 138.1] +$$

$$\begin{bmatrix} 179 - 250.7 \\ 117 - 137.3 \end{bmatrix} [179 - 250.7 \quad 117 - 137.3] + \begin{bmatrix} 427 - 250.7 \\ 143 - 137.3 \end{bmatrix} [427 - 250.7 \quad 143 - 137.3]$$

$$+ \dots + \begin{bmatrix} 190 - 250.7 \\ 131 - 137.3 \end{bmatrix} [190 - 250.7 \quad 131 - 137.3] +$$

$$\begin{bmatrix} 471 - 435.3 \\ 178 - 160.1 \end{bmatrix} [471 - 435.3 \quad 178 - 160.1] + \begin{bmatrix} 600 - 435.3 \\ 169 - 160.1 \end{bmatrix} [600 - 435.3 \quad 169 - 160.1]$$

$$+ \dots + \begin{bmatrix} 266 - 435.3 \\ 151 - 160.1 \end{bmatrix} [266 - 435.3 \quad 151 - 160.1] +$$

$$\begin{bmatrix} 319 - 360.7 \\ 150 - 154.6 \end{bmatrix} [319 - 360.7 \quad 150 - 154.6] + \begin{bmatrix} 504 - 360.7 \\ 168 - 154.6 \end{bmatrix} [504 - 360.7 \quad 168 - 154.6]$$

$$+ \dots + \begin{bmatrix} 417 - 360.7 \\ 164 - 154.6 \end{bmatrix} [417 - 360.7 \quad 164 - 154.6] +$$

$$\begin{bmatrix} 333 - 346.8 \\ 157 - 161.2 \end{bmatrix} [333 - 346.8 \quad 157 - 161.2] + \begin{bmatrix} 327 - 346.8 \\ 164 - 161.2 \end{bmatrix} [327 - 346.8 \quad 164 - 161.2]$$

$$+ \dots + \begin{bmatrix} 324 - 346.8 \\ 168 - 161.2 \end{bmatrix} \begin{bmatrix} 324 - 346.8 & 168 - 161.2 \end{bmatrix}$$

$$\text{Hata K}\check{\text{C}}\text{T} = \begin{bmatrix} 80320.92 & 14817.854 \\ 14817.854 & 4113.8821 \end{bmatrix} + \begin{bmatrix} 32060.68 & 1717 \\ 1717 & 1064.92 \end{bmatrix} +$$

$$\begin{bmatrix} 72982.68 & 11088.18 \\ 11088.18 & 3068.28 \end{bmatrix} + \begin{bmatrix} 143492.28 & 17725.76 \\ 17725.76 & 2958.92 \end{bmatrix} + \begin{bmatrix} 61236.88 & 11203.7 \\ 11203.7 & 1869.32 \end{bmatrix} +$$

$$\begin{bmatrix} 32563.68 & 3739.32 \\ 3739.32 & 1437.68 \end{bmatrix}$$

$$\text{Hata K}\check{\text{C}}\text{T} = \begin{bmatrix} 419657.56 & 60291.814 \\ 60291.814 & 22648.4221 \end{bmatrix}$$

Toplam kareler ve çarpımlar toplamı matrisi;

$$\text{KT}\check{\text{C}} \text{ Toplam} = \sum_{i=1}^g \sum_{k=1}^b \sum_{r=1}^n (Y_{ikr} - \bar{Y}_{...})(Y_{ikr} - \bar{Y}_{...})'$$

$$\text{Toplam K}\check{\text{C}}\text{T} = \begin{bmatrix} 270 - 332.3000 \\ 135 - 149.3667 \end{bmatrix} \begin{bmatrix} 270 - 332.3000 & 135 - 149.3667 \end{bmatrix}$$

$$+ \begin{bmatrix} 454 - 332.3000 \\ 193 - 149.3667 \end{bmatrix} \begin{bmatrix} 454 - 332.3000 & 193 - 149.3667 \end{bmatrix} + \dots + \begin{bmatrix} 303 - 332.3000 \\ 130 - 149.3667 \end{bmatrix}$$

$$\begin{bmatrix} 303 - 332.3000 & 130 - 149.3667 \end{bmatrix} + \begin{bmatrix} 237 - 332.3000 \\ 118 - 149.3667 \end{bmatrix} \begin{bmatrix} 237 - 332.3000 & 118 - 149.3667 \end{bmatrix}$$

$$\begin{aligned}
& + \dots + \begin{bmatrix} 216 - 332.3000 \\ 130 - 149.3667 \end{bmatrix} [216 - 332.3000 \quad 130 - 149.3667] + \\
& \begin{bmatrix} 179 - 332.3000 \\ 117 - 149.3667 \end{bmatrix} [179 - 332.3000 \quad 117 - 149.3667] + \dots + \\
& \begin{bmatrix} 190 - 332.3000 \\ 131 - 149.3667 \end{bmatrix} [190 - 332.3000 \quad 131 - 149.3667] \\
& + \begin{bmatrix} 471 - 332.3000 \\ 178 - 149.3667 \end{bmatrix} [471 - 332.3000 \quad 178 - 149.3667] + \dots + \\
& \begin{bmatrix} 266 - 332.3000 \\ 151 - 149.3667 \end{bmatrix} [266 - 332.3000 \quad 151 - 149.3667] + \\
& \begin{bmatrix} 319 - 332.3000 \\ 150 - 149.3667 \end{bmatrix} [319 - 332.3000 \quad 150 - 149.3667] + \\
& \begin{bmatrix} 504 - 332.3000 \\ 168 - 149.3667 \end{bmatrix} [504 - 332.3000 \quad 168 - 149.3667] + \dots + \\
& \begin{bmatrix} 417 - 332.3000 \\ 164 - 149.3667 \end{bmatrix} [417 - 332.3000 \quad 164 - 149.3667] + \begin{bmatrix} 333 - 332.3000 \\ 157 - 149.3667 \end{bmatrix} \\
& [333 - 332.3000 \quad 157 - 149.3667] + \begin{bmatrix} 327 - 332.3000 \\ 164 - 149.3667 \end{bmatrix} [327 - 332.3000 \quad 164 - 149.3667] \\
& + \dots + \begin{bmatrix} 324 - 332.3000 \\ 168 - 149.3667 \end{bmatrix} [324 - 332.3000 \quad 168 - 149.3667]
\end{aligned}$$

$$\text{Toplam K}\check{\text{C}}\text{T} = \begin{bmatrix} 1678630.24 & 361750.884 \\ 361750.884 & 181187.3768 \end{bmatrix}$$

Etkileşim kareler ve çarpımlar toplamı matrisi (Etkileşim KÇT);

Toplam KÇT = Fak1KÇT + Fak2KÇT + Etkileşim KÇT + Hata KÇT

$$\begin{bmatrix} 1678630.24 & 361750.884 \\ 361750.884 & 181187.3768 \end{bmatrix} = \begin{bmatrix} 340588.4976 & 64896.5088 \\ 64896.5088 & 12365.5248 \end{bmatrix} +$$

$$\begin{bmatrix} 89290.44 & 4845.312 \\ 4845.312 & 454.3152 \end{bmatrix} + \text{Etkileşim K}\check{\text{C}}\text{T} + \begin{bmatrix} 419657.56 & 60291.814 \\ 60291.814 & 22648.4221 \end{bmatrix}$$

$$\text{Etkileşim K}\check{\text{C}}\text{T} = \begin{bmatrix} 829093.742 & 231717.249 \\ 231717.249 & 145719.639 \end{bmatrix}$$

Wilks Lamda istatistiklerini bulabilmek için; Fak₁, Fak₂ ve etkileşime ilişkin kareler ve çarpımlar toplamı matrisleri, hata kareler ve çarpımlar toplamı matrisi ile toplanır ve determinantları alınır.

$$\text{Fak1 K}\check{\text{C}}\text{T} + \text{Hata K}\check{\text{C}}\text{T} = \begin{bmatrix} 340588.4976 & 64896.5088 \\ 64896.5088 & 12365.5248 \end{bmatrix} + \begin{bmatrix} 419657.56 & 60291.814 \\ 60291.814 & 22648.4221 \end{bmatrix}$$

$$\text{Fak1 K}\check{\text{C}}\text{T} + \text{Hata K}\check{\text{C}}\text{T} = \begin{bmatrix} 760246.0576 & 125188.3228 \\ 125188.3228 & 35013.9469 \end{bmatrix}$$

$$\text{Fak2 K}\check{\text{C}}\text{T} + \text{Hata K}\check{\text{C}}\text{T} = \begin{bmatrix} 89290.44 & 4845.312 \\ 4845.312 & 454.3152 \end{bmatrix} + \begin{bmatrix} 419657.56 & 60291.814 \\ 60291.814 & 22648.4221 \end{bmatrix}$$

$$\text{Fak2 K\c{C}T + Hata K\c{C}T} = \begin{bmatrix} 508948 & 65137.126 \\ 65137.126 & 23102.7373 \end{bmatrix}$$

$$\text{Etkileşim K\c{C}T + Hata K\c{C}T} = \begin{bmatrix} 829093.742 & 231717.249 \\ 231717.249 & 145719.639 \end{bmatrix} + \begin{bmatrix} 419657.56 & 60291.814 \\ 60291.814 & 22648.4221 \end{bmatrix}$$

$$\text{Etkileşim K\c{C}T + Hata K\c{C}T} = \begin{bmatrix} 1248751.302 & 292009.063 \\ 292009.063 & 168368.0611 \end{bmatrix}$$

$$\begin{aligned} \text{Determinant Fak1} &= (760246.0576 \times 35013.9469) - (125188.3228 \times 125188.3228) \\ &= 2661921509 - 1567211617 \end{aligned}$$

$$\text{Determinant Fak1} = 1094709892$$

$$\text{Determinant Fak2} = (508948 \times 23102.7373) - (65137.126 \times 65137.126)$$

$$\text{Determinant Fak2} = 3067035990$$

$$\text{Determinant Etkileşim} = (1248751.302 \times 168368.0611) - (292009.063 \times 292009.063)$$

$$\text{Determinant Etkileşim} = 1249805426.30$$

$$\text{Determinant Hata} = (419657.56 \times 22648.4221) - (60291.814 \times 60291.814)$$

$$\text{Determinant Hata} = 9504581556 - 3635102835$$

$$\text{Determinant Hata} = 5869478721$$

Daha sonra,

$\lambda_i = \frac{|HataK\check{C}T|}{Fak_i K\check{C}T + HataK\check{C}T}$ Formülü ile Wilks Lamda istatistikleri bulunur. Buna göre,

$$\lambda_1 = \frac{5869478721}{5869478721 + 1094709892} = 0.8428$$

$$\lambda_2 = \frac{5869478721}{5869478721 + 3067035990} = 0.6568$$

$$\lambda_3 = \frac{5869478721}{5869478721 + 1249805426} = 0.2361$$

Bu istatistiklere ilişkin F istatistikleri ise aşağıda hesaplanmıştır. Birinci faktörün düzey sayısı 2 olduğu için F istatistiği ve serbestlik dereceleri;

$$F_{(sd_1, sd_2)} = \left(\frac{1 - \lambda}{\lambda} \right) \left(\frac{Sd_2}{Sd_1} \right)$$

$$Sd_2 = 2(\text{Diğer faktörün düzey sayısı})(n-1) - p + 1$$

$Sd_1 = |1 - p| + 1$ ile hesaplanacak olursa,

$$Sd_2 = 2(3)(12-1) - 2 + 1 = 65$$

$$F_1 = \left(\frac{1 - 0.8428}{0.8428} \right) \frac{(2 * 3)(12-1) - 2 + 1}{|1-2| + 1} = 0.18652112 * \frac{65}{2}$$

$$F_1 = 6.0619$$

$$F_{c(2, 65)} \cong 3.15$$

$F_1 > F_c$ olduğu için, bitki boyunun ve dane veriminin çeşide göre değiştiği yani çeşidin önemli bir faktör olduğu söylenir.

İkinci faktörün düzey sayısı üç olduğu için aşağıdaki formüllerden yararlanılır.

$$F_{(sd_1, sd_2)} = \left(\frac{1-y}{y} \right) \left(\frac{Sd_2}{Sd_1} \right)$$

$$y = \lambda^{1/s} \text{ olup, } S = \sqrt{\frac{p^2(\text{İncelenenfak.Sd})^2 - 4}{p^2(\text{İncelenenfak.Sd})^2 - 5}} \text{ ve}$$

$$Sd_2 = S \left[\text{HataSd} - \frac{p - \text{İncelenenfak.Sd} + 1}{2} \right] - \left[\frac{p(\text{İncelenenfak.Sd}) - 2}{2} \right]$$

$Sd_1 = p$ (İncelenen fak. Sd) ile verilir. Buna göre;

İkinci Faktörün,

$$S = \sqrt{\frac{2^2(2^2) - 4}{2^2 + (2^2) - 5}} = 2 \quad y = \lambda^{1/s} = \sqrt{0.6568} = 0.810431983$$

$$Sd_2 = 2 \left[66 - \frac{2 - (2) + 1}{2} \right] - \left[\frac{2(2) - 2}{2} \right] = 129$$

$$Sd_1 = 2 (2) = 4$$

$$F_2 = \left(\frac{1 - 0.810431983}{0.810431983} \right) \left(\frac{129}{4} \right) \quad F_2 = 7.5435 \quad F_{c(4, 129)} \cong 2.37$$

$F_2 > F_c$ olduğu için bitki boyu ve bitki veriminin sıra üzerine göre değiştiği yani sıra üzerinin bu iki bağımlı değişken üzerindeki etkisinin önemli olduğu söylenir.

Etkileşimin F hesap değeri ve serbestlik derecelerine yukarıdaki formüller kullanılarak hesaplanır,

$$S = \sqrt{\frac{2^2(2^2) - 4}{2^2 + (2^2) - 5}} = 2 \quad Sd_2 = 2 \left[66 - \frac{2 - (2) + 1}{2} \right] - \left[\frac{2(2) - 2}{2} \right] = 129$$

$$y = \lambda^{1/s} = \sqrt{0.236} = 0.4859 \quad F_{\text{etk.}} = \left(\frac{1 - 0.4859}{0.4859} \right) \left(\frac{129}{4} \right) = 34.12$$

$F_{c(4, 129)} \cong 2.37$ $F_{\text{etk.}} > F_c$ olduğu için bitki boyu ve bitki veriminde etkileşimin önemli olduğu söylenir.

Bu durumda, her bir bağımlı değişken için çift yönlü varyans analizi ile etkileşimin hangi değişken için var olduğu bulunur. Uygulama için bu hesaplamalar yapılmış olup, her iki değişken için de etkileşimin anlamlı olduğu görülmüştür.

Çizelge 4.2. Çok deęişkenli iki yönlü varyans analiz sonuçları

VK	KÇT	Sd	KÇT+Hata	Det.	γ	F
Fak.1	$\begin{bmatrix} 340588.4976 & 64896.5088 \\ 64896.5088 & 12365.5248 \end{bmatrix}$	1	$\begin{bmatrix} 760246.0576 & 125188.3228 \\ 125188.3228 & 35013.9469 \end{bmatrix}$	1094709892	0.8428	6.0619
Fak.2	$\begin{bmatrix} 89290.44 & 4845.312 \\ 4845.312 & 454.3152 \end{bmatrix}$	2	$\begin{bmatrix} 508948 & 65137.126 \\ 65137.126 & 23102.7373 \end{bmatrix}$	3067035990	0.6568	7.5435
Etk.	$\begin{bmatrix} 829093.742 & 231717.249 \\ 231717.249 & 145719.639 \end{bmatrix}$	2	$\begin{bmatrix} 1248751.302 & 292009.063 \\ 292009.063 & 168368.0611 \end{bmatrix}$	1249805426.30	0.2361	34.12
Hata	$\begin{bmatrix} 419657.56 & 60291.814 \\ 60291.814 & 22648.4221 \end{bmatrix}$	67	$\begin{bmatrix} 419657.56 & 60291.814 \\ 60291.814 & 22648.4221 \end{bmatrix}$	5869478721		
Top.	$\begin{bmatrix} 1678630.24 & 361750.884 \\ 361750.884 & 181187.3768 \end{bmatrix}$	72				

VK: Varyasyon Kaynağı

Fak1: Bağımsız deęişkenlerden çeşit

Fak1: Bağımsız deęişkenlerden sıra üzeri

Etk. : Çeşit ve sıra üzeri interaksyonu

KÇT: Kareler ve çarpımlar toplamı matrisi

Sd: Serbestlik Derecesi

KÇT+Hata: Kareler ve çarpımlar toplamı matrisi + Hata Kareler ve çarpımlar toplamı matrisi

Det. : Faktörlere ait determinant deęerleri

γ : Wilks hesap deęeri

F: Faktörlere ait hesap deęerleri

Burada Lamda istatistiklerine ilişkin F istatistikleri,

Birinci faktör olan çeşit için $F_1 = 6,0619$ olarak hesaplanmıştır. Buna ait F cetvel değeri ise $F_{c(2, 65)} \cong 3.15$ olarak bulunmuştur. Elde edilen bu sonuçlara göre, $F_1 > F_c$ olduğu için, bitki boyunun ve dane veriminin çeşide göre değiştiği söylenebilir.

İkinci faktör olan Sıra üzeri için $F_2 = 7,5435$ Buna ait F cetvel değeri ise $F_{c(4, 129)} \cong 2.37$ olarak bulunmuş ve elde edilen bu sonuçlara göre, $F_2 > F_c$ olduğu için bitki boyu ve bitki veriminin sıra üzerine göre değiştiği söylenebilir.

Etkileşim, yani çeşit-sıra üzeri interaksyonu için F hesap değeri= 34.12 olarak hesaplanmıştır. $F_{c(4, 129)} \cong 2.37$ olarak bulunmuştur. $F_{\text{etk}} > F_c$ olduğu için bitki boyu ve bitki veriminde etkileşimin önemli olduğu söylenir.

Çizelge 4.3. SPSS’te MANOVA analizi sonuçları

Etki		Değer	F	Muamele S.D.	Hata S.D.	P
Kesişim	Pillai İz İst.	.993	4393.609	2.000	65.000	.000
	Wilks Lamda	.007	4393.609	2.000	65.000	.000
	Hotelling T ²	135.188	4393.609	2.000	65.000	.000
	Roy İz İst.	135.188	4393.609	2.000	65.000	.000
Sıraüzeri	Pillai İz İst.	.238	4.449	4.000	132.000	.002
	Wilks Lamda	.766	4.624	4.000	130.000	.002
	Hotelling T ²	.300	4.793	4.000	128.000	.001
	Roy İz İst.	.281	9.274	2.000	66.000	.000
Çeşit	Pillai İz İst.	.313	14.834	2.000	65.000	.000
	Wilks Lamda	.687	14.834	2.000	65.000	.000
	Hotelling T ²	.456	14.834	2.000	65.000	.000
	Roy İz İst.	.456	14.834	2.000	65.000	.000
Sıra üz. X Çeşit	Pillai İz İst.	.044	.734	4.000	132.000	.570
	Wilks Lamda	.957	.728	4.000	130.000	.574
	Hotelling T ²	.045	.722	4.000	128.000	.578
	Roy İz İst.	.041	1.346	2.000	66.000	.267

Değer = Verilere göre hesaplanmış olan Pillai İz İstatistiği, Wilks Lamda İstatistiği, Hotelling T² İstatistiği, Roy İz İstatistiği değerleridir.

P = Önemlilik

F = Hesap değeri

Muamele S.D. = Test istatistiklerine ait faktörlerin serbestlik dereceleri

Hata S.D. = Test istatistiklerine ait hata serbestlik dereceleri

Çizelge 4.4. Tek Değişkenli Varyans Analizi

Kaynak Bağımlı Değişken	Kareler Toplamı	S.D.	Kareler Ort.	F	P
Karışık Model					
Dane Verimi	256.258.878	5	51251.776	7.393	.000
Bitki Boyu	6713.891	5	1342.778	6.170	.000
Kesişim					
Dane Verimi	7891371.281	1	7891371.281	1138.260	.000
Bitki Boyu	1608799.427	1	1608799.427	7392.586	.000
Sıra Üzeri					
Dane Verimi	92539.451	2	46269.725	6.674	.002
Bitki Boyu	433.919	2	216.959	.997	.374
Çeşit					
Dane Verimi	161170.031	1	161170.031	23.247	.000
Bitki Boyu	6053.667	1	6053.667	27.817	.000
Sıra Üzeri					
Dane Verimi	2549.396	2	1274.698	.184	.832
X Çeşit	226.305	2	113.153	.520	.597
Hata					
Dane Verimi	457567.371	66	6932.839		
Bitki Boyu	14363.142	66	217.623		
Toplam					
Dane Verimi	8605197.530	72			
Bitki Boyu	1629876.460	72			
Gen. Top.					
Dane Verimi	713826.249	71			
Bitki Boyu	21077.033	71			

Kaynak: Bağımsız değişkenler

S.D. : Serbestlik derecesi

F: Hesap değeri

P: Önemlilik derecesi

Tek Değişkenli Varyans Analizi'nde; Sıra üzerinin; dane verimine ve bitki boyuna olan etkisi önemsiz olduğu, Çeşidin de; dane verimine ve bitki boyuna etkisi önemsiz olduğu, Sıra üzeri-Çeşit interaksyonunun; dane verimine ve bitki boyuna etkisi önemsiz olduğu bulunmuştur ($p > 0.05$).

4.5. Çoklu Doğrusal Regresyon Analizi

Çoklu Doğrusal Regresyon Analizi'ne göre; Regresyon Analizleri $R^2 = 0.4-0.6$ arasında değer alırsa orta uyum iyiliğine, $R^2 > 0.6$ ise çok iyi uyum iyiliğine sahiptir denir.

Çizelge 4.5. Çoklu regresyon analizi sonuçları

Bağımlı Değişken	Bağımsız Değişken	R^2	b	Std. Hata	Beta	t	P	F	S.D.
Dane Verimi	a	.340	454.363	61.809	-.337	-3.447	.001	5.724	4
	Sıra Üzeri		-4.113	1.193					
	Çeşit		94.625	19.483					
Bitki Boyu	a	.292	454.363	.212	-.070	-.688	.494	7.110	4
	Sıra Üzeri		-.146	3.466					
	Çeşit		18.339						

a = Regresyon doğrusunun Y eksenini kestiği nokta

b = Regresyon katsayısı

Beta = Standardize edilmiş regresyon katsayısı

Std. Hata = Regresyon katsayılarının standart hatası

P = önemlilik derecesi

R^2 = Çoklu belirtme katsayısı

F = Regresyon hesap değeri

S.D. = Serbestlik derecesi

Çoklu belirtme katsayısı R^2 birden fazla bağımsız değişkenin bağımlı değişkeni açıklama miktarıdır. Regresyon denkleminin uygunluk ölçüsü olarak ta kullanılan R^2 , regresyon modeli ile açıklanan değişimin toplam değişimdeki yüzdesi olarak bilinir. R^2 değeri 0-1 arasında değer alır. R^2 değeri 1'e yaklaştıkça bağımsız değişkenlerin bağımlı değişkeni daha iyi açıkladığı, R^2 değeri sıfıra yaklaştıkça modelde yer alan bağımsız değişkenlerin bağımlı değişkeni iyi açıklamadığı şeklinde ifade edilir. Buna göre, bağımsız değişkenler; çeşit, sıra üzeri olduğunda dane verimine etki eden R^2 değeri 0.340 olarak hesaplanmış olup, bunun iyi uyum iyiliğine sahip olmadığı, Bağımlı

değişken, bitki boyu alındığında R^2 değeri 0.292 olarak hesaplanmış ve bununda iyi uyum iyiliğine sahip olmadığı yani modeldeki gerçek verileri yansıtmadığı görülmüştür.

Regresyon modelinde hangi değişkenin modele en fazla katkıda bulunduğunu tespit etmek için standardize edilmiş regresyon katsayılarından faydalanılır. Buna göre dane verimini açıklayan modelde çeşide ait standardize edilmiş regresyon katsayısı (0.475) ve sıra üzerine ait standardize edilmiş regresyon katsayısı (-0.337) bulunmuş olup, çeşidin dane verimini açıklayan modelde pozitif yönde olumlu, sıra üzerinin ise negatif yönde etki yaptığı gözlenmiştir. Bitki boyunu oluşturan modelde ise çeşide ait standardize edilmiş regresyon katsayısı (0.536) ve sıra üzerine ait standardize edilmiş regresyon katsayısı (-0.070) bulunmuş olup, çeşidin bitki boyunu açıklayan modelde pozitif yönde olumlu, sıra üzerinin ise negatif yönde etki yaptığı gözlenmiştir.

Çizelge 4.5'te verilen değerlere göre, her bir bağımlı değişkene karşılık bağımsız değişkenlerin R^2 değerlerine bakılarak ne kadar uyum iyiliği gösterdiği belirtilmiş olup; bağımlı değişken dane verimi ve bağımsız değişkenler; çeşit, sıra üzeri olduğunda dane verimine etki eden R^2 değeri 0.340 olarak hesaplanmıştır ve bunun iyi uyum iyiliğine sahip olmadığı yani toplam değişimin % 34'ünün ele alınan iki bağımsız değişken tarafından açıklanabilmekte olduğu görülmektedir, bitki boyu bağımlı değişken alındığında ise $R^2 = 0.292$ olarak hesaplanmış olup, iyi uyum iyiliğine sahip olmadığı ve toplam değişimin % 34'ünün ele alınan iki bağımsız değişken tarafından açıklanabilmekte olduğu görülmektedir.

Önemlilik derecelerine bakıldığında; Sıra üzeri ve çeşidin dane verimi üzerine etkisinin önemli olduğu ve bitki boyuna ise sadece çeşidin etkili olduğu, sıra üzerinin ise önemli olmadığı görülmüştür.

Dane verimi ve bitki boyuna göre model aşağıdaki gibi olur,

$$Y = a + b_1X_1 + b_2X_2 + e_{ij}$$

$$Y_{\text{Dane verimi}} = 454.363 + (-4.113) * \text{sıra üzeri} + 94.625 * \text{çeşit} + 61.809$$

$$Y_{\text{Bitki boyu}} = 454.363 + (-0.146) * \text{sıra üzeri} + 18.339 * \text{çeşit} + 61.809$$

5. SONUÇ VE TARTIŞMA

Bilimin ve teknolojinin gelişmesine paralel olarak karmaşık yapıdaki problemlerin çözümünde ve herhangi bir araştırmada değişken sayıları birden fazla olduğu durumlarda tek değişkenli varyans analizi kullanmak yeterli olmamakla birlikte işlemlerin uzun sürmesi ve daha çok hatanın olmasına sebep olmaktadır. Bundan ötürü Çok Değişkenli Analiz'ler bilgisayar teknolojisinin gelişmesiyle yaygın bir kullanım alanına sahip olmuştur. Bunun sebebi; işlemlerin daha hızlı yapılarak zamandan tasarrufun sağlanması, yapılan işlemlerin daha doğru ve hassas yapılması, çözümü zor karmaşık modellerin çözümünde kullanıcıya yardım etmesinden, elde edilen sonuçların yorumunun daha kolay yapılmasından dolayı yaygın bir kullanım alanına sahiptir.

Tek değişkenli analizlerde, en önemli varsayım, araştırmadaki diğer boyutların etkilerinin sabit kabul edilmesi ve her defasında sadece bir faktörün inceleme konusu yapılmasıdır. Halbuki yapılan araştırmalar ve çalışmalar sadece tek bir faktörün etkisiyle değil çok sayıda faktörün ve bu faktörlerin karşılıklı interaksiyonu ile oluşmakta ve karmaşık bir yapı göstermektedir.

Çok Değişkenli Varyans Analizi'nden elde edilen sonuçlara göre, Çizelge 4.2'de birinci faktör olan çeşit için $F_1 = 6,0619$ ve serbestlik derecesi: $2-1 = 1$ olarak hesaplanmıştır. Buna ait F cetvel değeri ise $F_{c(2, 65)} \cong 3.15$ olarak bulunmuştur. Elde edilen bu sonuçlara göre, $F_1 > F_c$ olduğu için, bitki boyunun ve dane veriminin çeşide göre değiştiği yani çeşidin bitki boyu ve dane verimi üzerine olan etkisinin önemli olduğu söylenir.

İkinci faktör olan Sıra üzeri için $F_2 = 7,5435$ ve serbestlik derecesi: $3-1 = 2$ olarak hesaplanmıştır. Buna ait F cetvel değeri ise $F_{c(4, 129)} \cong 2.37$ olarak bulunmuş ve elde edilen bu sonuçlara göre, $F_2 > F_c$ olduğu için bitki boyu ve bitki veriminin sıra üzerine göre değiştiği yani sıra üzerinin bu iki bağımlı değişken üzerindeki etkisinin önemli olduğu söylenebilir.

Etkileşim, yani çeşit-sıra üzeri interaksyonu için F hesap değeri= 34.12 olarak hesaplanmıştır. Serbestlik derecesi= $2 \times 1 = 2$ dir. F cetvel değeri ise, $F_{c(4, 129)} \cong 2.37$ olarak bulunmuştur. $F_{\text{etk}} > F_c$ olduğu için bitki boyu ve bitki veriminde etkileşimin önemli olduğu söylenir.

Tek Değişkenli Varyans Analizi'nin sonuçlarına göre; Sıra üzerinin; dane verimine ve bitki boyuna etkisi 0.05 lik hata payı ile önemsiz olduğu ($p > 0.05$), Çeşidin de; dane verimine ve bitki boyuna etkisi 0.05'lik hata payı ile önemsiz ($p > 0.05$) olduğu, Sıra üzeri-Çeşit interaksyonunun; dane verimine ve bitki boyuna etkisi 0.05'e göre önemsiz ($p > 0.05$) olduğu bulunmuştur.

Çoklu Doğrusal Regresyon Analizi sonuçlarına göre; Çizelge 4,5.'de Sıra üzeri ve çeşidin dane verimi üzerine etkisinin önemli olduğu ve bitki boyuna ise sadece çeşidin etkili olduğu, sıra üzerinin ise önemli olmadığı görülmüştür.

Çoklu Regresyon'un uyum iyiliklerine, yani verilerin gerçek verilere uyumlu olup olmadığına bakıldığında;

Bağımlı değişken dane verimi ve bağımsız değişkenler; çeşit, sıra üzeri olduğunda dane verimine etki eden R^2 değeri 0.340 olarak hesaplanmış olup, bunun iyi uyum iyiliğine sahip olmadığı, yani ilgili modelin gerçek verilere yakın olmadığı, buda toplam değişimin % 34'ünün ele alınan iki bağımsız değişken tarafından açıklanabilmekte olduğu görülmektedir.

Bağımlı değişken, bitki boyu alındığında $R^2 = 0.292$ olarak hesaplanmış olup, iyi uyum iyiliğine sahip olmadığı yani modeldeki gerçek verileri yansıtmadığı, yani toplam değişimin % 29'unun ele alınan iki bağımsız değişken tarafından açıklanabilmekte olduğu görülmüştür.

Tek Değişkenli Regresyonun hesap değerlerine bakıldığında, Birinci bağımlı değişken yani dane verimi için $F_{\text{cetvel}(4,67;\alpha=0.05)} \cong 2.53$ bulunmuştur. $F_{\text{hesap}} = 5.724$ olarak

hesaplanmıştır. $F_{\text{hesap}} > F_{\text{cetvel}}$ olduğu için doğrusal ilişkinin anlamlı olduğu ya da dane verimindeki değişimin sıra üzeri ve çeşit ile açıklanabileceği söylenir.

İkinci bağımlı değişken olan bitki boyu için $F_{\text{cetvel}(4,67;\alpha=0.05)} \cong 2.53$ bulunmuştur. $F_{\text{hesap}} = 7.110$ olarak hesaplanmıştır. $F_{\text{hesap}} > F_{\text{cetvel}}$ olduğu için doğrusal ilişkinin anlamlı olduğu ya da bitki boyundaki değişimin sıra üzeri ve çeşit ile açıklanabileceği söylenebilir.

Çok Değişkenli Regresyonda, Çizelge 4,5'e bakıldığında, sıra üzerinin dane verimine olan etkisi negatif yönde yani dane verimiyle ters orantılı olarak artmaktadır. Çeşitin dane verimi üzerine olan etkisi pozitif yönde önemli bulunmuştur. Yani çeşit dane verimiyle doğru orantılı olarak artmaktadır. Sıra üzerinin bitki boyuna olan etkisi önemsiz, çeşitin dane verimine olan etkisi pozitif yönde önemli bulunmuştur.

Sonuçta, Hem Çok Değişkenli Varyans Analizi'nden elde edilen sonuçlarda hem de Çoklu Regresyon Analizi'nden elde sonuçlarda bağımsız değişkenlerin bağımlı değişkenler üzerinde olan etkileri önemli bulunmuştur. Bunun nedeni her iki analizde de değişkenler tek bir değişken gibi ele alınıp, birlikte olan etkileri incelenmiş olmasıdır. Tek Değişkenli Varyans Analizi'ndeki sonuçlar değişkenlerin tek tek ele alınması ile elde edilmiş ve sonuçlar önemsiz bulunmuştur.

KAYNAKLAR

- Alpar, R., 1997. Uygulamalı Çok Değişkenli İstatistiksel Yöntemlere Giriş I, Ankara, 120-210.
- Anderson, T.W., 1958. An Introduction to Multivariate Statistical Analysis, John Willey and Sons, Newyork.
- Anderson, Black, Hair, Joseph F., Rolf E., Ronald L., Tatham, William C., 1998. Multivariate Data Analysis, prentice-Hall international, Inc., New Jersey, s.12.
- Arslan, S., Mirtaghioglu, H., 2004. Tek Yönlü Sınıflanmış Şansa Bağlı Modelde Desen Dengesizliğinin F Testinin Gücü Üzerine Etkisi. 4. Ulusal Zootekni Bilim Kongresi Sözlü Bildiriler Formu, Ankara.
- Balibeyoğlu, J., 1989. Çok Değişkenli Varyans Analizi ve Yazlık Mercimek Kalitesi Üzerine Uygulaması (Doktora Tezi). Gazi Üniversitesi Fen Bilimleri Enst. Ankara.
- Benli, Y.K., 2005. Sektörel Farklılıkların Oranlar Üzerinde Etkisi. Gazi Üniversitesi Endüstriyel Sanatlar Eğitim Fakültesi Dergisi, Sayı: 16, s.14-30.
- Bock, R.D., Haggart E., 1963. The Use of Multivariate Analysis of Variance in Behavioral Sciences ch. 3.
- Box, G.E.P., 1949. A General Distribution Theory for a Classes of Likelihood Criteria Biometrika, 36, 317–346.
- Colej, W., Grizzle J, 1966. Application of Multivariate analysis of variance to repeated measurements experiments, Biometrics, 22, 810–828.
- Çelik, F., 2005. Orta Öğretim Öğrencilerinin Okula Yabancılaşma Düzeylerinin Bazı Değişkenler Açısından İncelenmesi (Yüksek Lisans Tezi).Çukurova Üniversitesi Sosyal Bilimleri Enst., Adana.
- Daşdemir, İ., Güngör, E., 2002. Çok Boyutlu Karar Verme Metotları Ve Ormancılıkta Uygulama Alanları. ZKÜ, Bartın Orman Fak. Der. Vol 1-2.
- Davis, A.W., 1970. Exact Distribution of Hotelling's Generalized T_0^2 , Biometrika, 57, 187–191.
- Gabriel, K.R, 1968. Simultaneous test procedures in multivariate analysis of variance Biometrika, 55, 489–504.
- Gabriel, K.R, 1969. Comparasion of Some Methods of Simultaneous Inference in MANOVA Multivariate Analysis II (P.R. Krishnaiah ed) Academic Newyork.
- Hugnes, D., Saw J., 1972. Approximating The Percentage Points Of Hotelling Generalized T_0^2 Statistic, Biometrika 59, 224–226.
- Ito, K., 1962. Comparison Of The Powers Of Two Multivariate Analysis Of Variance Test, Biometrika, 49, 455-462.
- Johnson, R.A., Wichern, D.W., 2002. Applied Multivariate Statistical Analysis. Bıçaklar Kitapevi Basım Yayın ve Dağıtım Ltd. Şti., Ankara, 298-299.
- Kirk, R., 1982. Experimental Desing, Brooks/Cole Publishing Company California.
- Krishnaiah, P.R., 1969. "Simultaneous test procedures under general MANOVA models" Multivariate Analysis 2. Academics Pres, Newyork, 121–143.
- Örsel, S., Canpolat, B., Akdemir, A., Özbay, M., 2004. Diyet Yapan ve Yapmayan Ergenlerin Kendilik Algısı, Beden İmajı ve Beden Kitle İndeksi Açısından Karşılaştırılması. Türk Psikiyatrisi Dergisi, 15(1), 5-15.
- Özdamar, K.,1999. Paket Programlar ile İstatistiksel Veri Analizi. Kaan Kitap Evi, Eskişehir, 137–210.

- Pillai, K.C.S and Joyachandran, K., 1967. Power comparison of test of two Multivariate hypotheses based of four criteria *Biometrika* 54, 195–210.
- Raats, V.M., Genugten B. B., Moors J. J. A., 2005. Multivariate Regression With Consecutively Added Dependent Variables, 410, 170–197.
- Smith, H., Gnanadesikan,R., Hugnes J.B., 1962. Multivariate Analysis of Variance (MANOVA), *Biometrics*, 18, 22–41.
- Şahinler, S., 2000. En Küçük Kareler Yöntemi İle Doğrusal Regresyon Modeli Oluşturmanın Temel Prensipleri. Mustafa Kemal Üniversitesi Ziraat Fakültesi Dergisi, 5(1-2), 57-73.
- Şıklar, E., 2000. Regresyon Analizine Giriş. T.C. Anadolu Üniversitesi Yayınları, s.1-51, Eskişehir.
- Şirvancı, M., 1997. Kalite İçin Deney Tasarımı. Literatür Yayıncılık, İstanbul.
- Tatlıdil, H., 1996. Uygulamalı Çok Değişkenli İstatistiksel Analiz. Cem Ofset, Ankara.
- Terzi, Y., 1995. Değişkenli Varyans Analizi ve Bir Uygulama (Yüksek lisans Tezi). Ondokuz Mayıs Üniversitesi Fen Bilimleri Enst. Samsun.
- Ünsal, A., Güler H., 2005. Türk Bankacılık Sektörünün Lojistik Regresyon Ve Diskriminant Analizi İle İncelenmesi. VII. Ulusal Ekonometri ve İstatistik Sempozyumu. 26-27 Mayıs. İstanbul.
- Varinli, İ., Kurtoğlu, R., 2005. Satış Elemanlarının Etik Algılamaları ve Perakende Sektöründe Bir Uygulama. C.Ü. İktisadi ve İdari Bilimler Dergisi, 6 (2).
- Vehid, E.H., Özden, Y.I., Dirican, A., Öneş, Ü., 2000. Çok Değişkenli Varyans Analizi ve Uygulaması. İstanbul Tıp Fakültesi Dergisi, 63 (4).
- Yıldız, N., Akbulut, Ö., Bircan, H., 2002. Uygulamalı İstatistik. Atatürk Üniversitesi Ziraat Fakültesi Ofset Tesisi, 60, 283-288.
- Wilks, S.S., 1932. Certain genalizations in the analysis of variance *Biometrika* 24 , 471-494.