

T.C.
MARMARA ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
EKONOMETRİ ANABİLİM DALI
İSTATİSTİK BİLİM DALI

KOVARYANS ANALİZİ VE BİR UYGULAMA

(YÜKSEK LİSANS TEZİ)

BELMA GÖREN

Prof.Dr.SELAHATTİN GÜRİŞ

İSTANBUL 1997



Marmara Üniversitesi
Kütüphane ve Dokümantasyon Daire Başkanlığı



T04360

İÇİNDEKİLER

GİRİŞ	1
-------------	---

BİRİNCİ BÖLÜM

ÇOK DEĞİŞKENLİ İSTATİSTİKSEL TEKNİKLER

1.1. ÇOK DEĞİŞKENLİ İSTATİSTİK	3
1.2. ÇOK DEĞİŞKENLİ İSTATİSTİK TEKNİKLER	6
1.2.1. BAĞIMLILIK ANALİZİNDE KULLANILAN YÖNTEMLER	6
1.2.1.1. REGRESYON ve KORELASYON ANALİZİ	6
1.2.1.2. ÇOK DEĞİŞKENLİ VARYANS ANALİZİ	7
1.2.1.3. KOVARYANS ANALİZİ	8
1.2.1.4. DİSKRİMİNANT ANALİZİ	8
1.2.1.5. EŞANLI DENKLEMLER SİSTEMİ	9
1.2.1.6. LOJİSTİK REGRESYON ANALİZİ	10
1.2.1.7. KANONİK REGRESYON ANALİZİ	10
1.2.2. KARŞILIKLI BAĞIMLILIK ANALİZLERİ	11
1.2.2.1. TEMEL BİLEŞENLER ANALİZİ	11
1.2.2.2. FAKTÖR ANALİZİ	12
1.2.2.3. KÜMELEME ANALİZİ	13
1.2.2.4. ÇOK BOYUTLU ÖLÇEKLEME	14
1.3. ÇOK DEĞİŞKENLİ İSTATİSTİK TEKNİKLERİNİN KARŞILAŞTIRILMASI	14

İKİNCİ BÖLÜM

KOVARYANS ANALİZİ

2.1. KOVARYANS KAVRAMI	18
2.2. KOVARYANS ANALİZİ	20
2.2.1. TAM RASTLANTISAL TASARIM İÇİN KOVARYANS ANALİZİ	24
2.2.1.1. GÖZLEM DEĞERLERİNİN DÜZENLENMESİ	24
2.2.1.2. TAM RASTLANTISAL TASARIM MODELİ	26
2.2.1.3. HİPOTEZ TESTLERİ	32

2.2.2. TESADÜFİ BLOK TASARIM İÇİN KOVARYANS ANALİZİ	35
2.2.2.1. GÖZLEM DEĞERLERİNİN DÜZENLENMESİ	36
2.2.2.2. TESADÜFİ BLOK TASARIM MODELİ	37
2.2.2.3. HİPOTEZ TESTLERİ	42
2.2.3. FAKTÖRİYEL TASARIM İÇİN KOVARYANS ANALİZİ	44
2.2.3.1. GÖZLEM DEĞERLERİNİN DÜZENLENMESİ	45
2.2.3.2. FAKTÖRİYEL TASARIM MODELİ	46
2.2.3.3. HİPOTEZ TESTLERİ	51

ÜÇÜNCÜ BÖLÜM
UYGULAMA

3.1. UYGULAMANIN AMACI	52
3.2. UYGULAMANIN VERİLERİ VE SONUÇLARI	53
3.3. UYGULAMANIN YORUMLANMASI	67
SONUÇ	69
KAYNAKLAR	71

GİRİŞ

Çok deęişkenli istatistik teknikleri ifadesinden de anlaşılacağı üzere deęişken sayısının çok olduęu tekniklerdir. Deęişken sayısının çok olması incelenen olayın kapsamını genişletmekte ve olayı bütün yönleriyle incelememize olanak tanımaktadır. Bu durumda çok deęişkenli istatistik teknikleri deęişken sayısının çok olmasına rağmen incelenen olayı daha iyi açıklayabilen ve uygulama teknikleri ile daha az sayıda deęişkenle ifade eden ve anlaşılabilir yorumlanmasını sağlayan teknikler olarak tanımlanabilir.

Çok deęişkenli istatistik teknikleri, Temel Bileşenler Analizi, Faktör Analizi, Kümeleme Analizi, Çok Boyutlu Ölçekleme, Regresyon Analizi, Eşanlı Denklemler Sistemi, Varyans Analizi, Kovaryans Analizi, Diskriminant Analizi, Lojistik Regresyon Analizi, Kanonik Korelasyon Analizidir. Bu analizler bağımlılık analizleri ve karşılıklı bağımlılık analizleri olarak sınıflandırılabilir.

Çalışmanın ilk bölümünde çok deęişkenli analiz tekniklerinin ne amaçla kullanıldıkları kısaca açıklanmaya çalışılmıştır.

Çok deęişkenli istatistik tekniklerinden biri olan Kovaryans Analizi ayrıntılı bir şekilde ele alınmıştır. Önce

kovaryansın ne anlama geldiđi sonra da Kovaryans analizi açıklanmaya çalışılmıştır.

Kovaryans analizi, varyans analizi ile regresyon analizini birleştiren bir tekniktir. Bağımlı deđişken deđişirken bağımlı deđişken ile birlikte deđişen deđişken veya deđişkenlerin etkisini ölçmek ve ortadan kaldırmak amacıyla uygulanır. Ayrıca, varyans analizinin kullanıldığı durumlarda kovaryans analizi de kullanılabilir.

Son bölümde Kovaryans analizi ile ilgili bir uygulama yapılmıştır. Marmara Üniversitesi öğrencilerine anket yapılmış ve Ekonometri, Maliye, İşletme ve İktisat bölümlerinden 10'ar öğrenci tesadüfen seçilmiştir. Amaç, bölümler arasında mezuniyet derecesine göre fark olup olmadığı, fark varsa sebebin dersaneye gitme faktörü olup olmadığı, dersane faktörünün etkisi giderildikten sonra bölümler arasında homojenliğin sağlanıp sağlanamayacağı sorularına Kovaryans Analizi uygulanarak cevap verebilmektir.

1.1. ÇOK DEĞİŞKENLİ İSTATİSTİK

Çok sayıdaki özelliğin bir bütün olarak ele alınması için, değişkenler arasındaki farkların veya benzerliklerin ölçülmesi, farklı gruplanması ve farklı yorumlanması gerekmektedir. İstatistik, kesin olmayışlığın ışığında karar verebilmek için kullanılan bir araç olarak tanımlanırsa¹, çok değişkenli istatistiksel analiz teknikleri de birbiriyle ilişkili ve özellikli değişkenleri analiz etmemize yardımcı olan istatistiki teknikler olarak tanımlanabilir.

Toplumda pek çok alanda bir takım bilgileri belli esaslara dayandırma gereği duyulmaktadır. Politika konusu ile ilgilenen bilim adamlarına göre oy oranları oylamanın yapıldığı sıklığa, yaşa, cinsiyete ve mesleklere göre değişim gösterecektir. sınıflandırabilirler. Bu durum yaşa, cinsiyete, toplumun farklı kesimlerine göre değişim gösterecektir. Bu değişim üç durumda incelenir. Birincisi, sayısal değişim gösterecektir ki burada ortalama değerler üzerindeki olası değerlerin değişimleri incelenecektir. İkincisi bağımlı değişkenler ve değişimler, oran ve regrasyon analizleri yapılacaktır. Üçüncüsü, bağımlılık ve bağımsızlık değişimleri içeren incelemedir. Bu durumu analiz etmek için çok değişkenli istatistik teknikleri kullanılacaktır.²

Çok değişkenli analizde deney birimlerinden gözlem ya da ölçüm yoluyla elde edilen özellikler göz önüne alınır. Değişken

¹ Charles R.Hicks, Deney Düzenlenmede İstatistiksel Yöntemler, (çev.) Zehra Muluk ve diğerleri, Ankara:Akademi Matbaası, 1985, s.1

² Albert R.Wildt, Olli Ahtola, Analysis of Covariance, Üniversty of Georgion and Üniversty of Florida, 1985, s.7

adı verilen bu özelliklerin çok sayıda olması, sorunun klasik istatistik teknikler ile çözümüne olanak tanımamaktadır.³ Çok değişkenli analiz teknikleri farklı amaçlarla kullanılmaktadırlar. Bu amaçlardan başlıcaları,

- a)Değişken sayısının aza indirgenmesi,
- b)Değişkenlerin aralarındaki ilişkilerin ölçümünden yararlanarak bağımlılığın incelenmesi,
- c)Sıralama ve ölçeklemedir.

Tek değişkenli istatistiksel tekniklerin her zaman yeterli olmamalarının nedeni, bu tekniklerin bazı kısıtlamalar altında geçerli olmasıdır. Bu kısıtlamalardan en önemlisi, bir çok faktörün deneysel olarak kontrol altında tutulması ve her defasında tek bir faktör etkisinin incelenmesidir.⁴ Çok değişkenli istatistiksel tekniklerin kullanımında böyle bir kısıtlama sözkonusu değildir. Çünkü çok değişkenli istatistiksel teknikler, daha önce de belirtildiği gibi incelenen olayı bir bütün olarak ele alır ve değişkenler arasındaki ilişkileri inceler. Bu araştırmalar bağımlılık ve karşılıklı bağımlılık analizleri ile yapılır. Bağımlılık analizinde değişkenler birbirleriyle ilişkili olup başka değişkenlerce açıklanabilmektedir. Bağımlılık analizinde kullanılan yöntemler, Regresyon Analizi, Eşanlı Denklemler Sistemi, Varyans Analizi, Kovaryans Analizi, Diskriminant Analizi, Lojistik Regresyon Analizi, Kanonik Korelasyon Analizi'dir.

³ Hüseyin Tatlıdil, Uygulamalı Çok Değişkenli İstatistiksel Analiz, Anaka; Cem Web Ofset Ltd.Şti., 1996, s.1

⁴ Hüseyin Tatlıdil, a.g.e., s.1

Karşılıklı bağımlılık analizinde bütün değişkenlerin arasındaki bağımlılık yerine karşılıklı bağımlılıklar söz konusudur. Bu analiz yöntemleri, değişkenler arasındaki karşılıklı bağımlılığın analiz edilerek, nesnelere veya değişkenlerin temel yapılarının anlaşılmasında ve onların yeni değişkenlerle birleştirilmesinde kullanılmaktadır.⁵ Karşılıklı bağımlılık analizinde kullanılan yöntemler, Temel Bileşenler Analizi, Faktör Analizi, Kümeleme Analizi ve Çok Boyutlu Ölçeklemedir.

Çok değişkenli analiz teknikleri adı verilen bu tekniklerin asıl amacı; istatistiğin öteki kollarında olduğu gibi, bilimsel çalışmaların sayı ile ifade edilebilen sonuçlarının özetlenmesi, yorumlanması ve karar verilirken kullanılmasının sağlanmasıdır. Bilimsel çalışmalarda ele alınan olaylar genellikle pek çok etkenin etkisi altındadır. Ayrıca gözleme konu olan nesnelere özellikleri de birbiriyle ilişkilidir. Bu nedenle, uygulamada çok sayıda değişkenle karşılaşmaktadır. Yapılan çalışmaların geçerli ve güvenilir sonuçlar verebilmesi için, inceleme konusu olayları bütün yönleriyle değerlendirilmesi gerekir. Bu zorunluluk sonucu araştırmacı çok değişkenli veri ve bunların analizleri ile karşı karşıya kalır.

⁵ David A.AKER, Multivariate Analysis in Marketing : Theory & Application, Wadsworth Publishing Company California, 1971, s.480

1.2. ÇOK DEĞİŞKENLİ İSTATİSTİKSEL TEKNİKLER

1.2.1. Bağımlılık analizinde kullanılan yöntemler;

Bağımlılık analizinde değişkenler birbiri ile ilişkili olup başka değişkenlerce açıklanabilmektedir. Bağımlılık analizinde kullanılan yöntemler aşağıda kısaca açıklanmıştır.

1.2.1.1. Regresyon Analizi

Değişkenler arasındaki ilişkinin matematiksel bir model yardımıyla incelenmesine “Regresyon analizi” denmektedir..⁶

Basit regresyon analizinin amacı iki değişken arasındaki ilişkiyi $Y = \alpha + \beta X$ denkleminin yardımıyla açıklamaktır. Burada, Y bağımlı değişken, X bağımlı değişkeni etkileyen bağımsız değişkendir. $Y = \alpha + \beta X$ basit doğrusal regresyon denklemdir. Eğer bağımlı değişken Y birden çok bağımsız değişken tarafından açıklanıyorsa, sözkonusu analiz çok değişkenli regresyon analizi olarak adlandırılır.

Bağımlı ve bağımsız değişkenler arasındaki ilişkinin yönü ve derecesi hakkında bilgi edinmek için korelasyon katsayısı hesaplanmaktadır. Korelasyon katsayısı +1 ile -1 arasında değer alır ve r ile gösterilir. Korelasyon katsayısı,

⁶ Necla Çömlekçi, Temel İstatistik İlke ve Teknikleri, Bilim Teknik Yayınevi, Eskişehir, 1989,s.426

$$r = (\Sigma XY - \Sigma X \Sigma Y / n) / [(\Sigma X^2 - \Sigma X^2 / n)(\Sigma Y^2 - \Sigma Y^2 / n)]^{1/2}$$

şeklinde formüle edilir.⁷

Eğer $r > 0$ ise değişkenler arasında aynı yönde, $r < 0$ ise ters yönde bir ilişki vardır, $r = 0$ ise değişkenler arasında ilişki yoktur denilir.

1.2.1.2. Çok Değişkenli Varyans Analizi (Manova)

İkiden fazla örneklem ortalamasının karşılaştırılması için geliştirilmiş olan yöntem, varyans analizi olarak isimlendirilir, sözkonusu yöntem değişkenlerin karşılaştırılmasına dayandırılmıştır.⁸

Varyans analizi aynı ana kütlede veya aynı varyansa sahip farklı ana kütlelerden seçilmiş iki veya daha fazla sayıda örneğin varyanslarının ortalaması alınarak ana kütle varyansının tahmininin yapılması prensibine dayanır.⁹

Varyans analizinin diğer bir amacı ise, bağımlı değişkendeki değişimi yorumlayabilmek ve bağımsız değişkende bir değişim olup olmadığını test etmektir.¹⁰

⁷ Şahin Akkaya, ve M.Vedat Pazarlıoğlu, Ekonometri I, Anadolu Matbaacılık, İzmir, 1995 s.78

⁸ Necla Çömlekçi, a.g.e. ,s.415

⁹ Bilge Aloba Köksal, İstatistik Analiz Metodları, İstanbul; Çağlayan Kitabevi, s.320

¹⁰ Michael R.Anderberg, Cluster Analysis for Applications, Academic Press, 1973, s.189

1.2.1.3. Kovaryans Analizi

Kovaryans analizi, varyans analizi ile regresyon analizinin bir bileşimidir. Analizin kullanılma amacı, bağımlı değişken değeri Y değişirken kontrol edilemeyen ve birlikte değişen bağımsız değişken X değerlerinin etkisini yok etmeye çalışmaktadır.

Kovaryans analizi, değişken grupları arasındaki karşılıklı ilişkileri ölçmekte ve açıklamakta sıkça kullanılan bir tekniktir.¹¹

1.2.1.4. Diskriminant Analizi

Değişkenlerin bilinen özellikleri bir küme oluşturup karşılaştırıldığında, özelliği bilinen değişkenin hangi kümeye gireceğinin kestirilmesi için diskriminant analizi kullanılır. Diskriminant analizinde amaç, hatalı sınıflandırma olasılığını en aza indirgeyerek bireyleri ait oldukları gruplara ayırmak, çekilmiş oldukları kitleleri belirlemektir.¹²

Diskriminant analizinin sonucunda;¹³ Kümeler arasında anlamlı bir farklılık olup olmadığı denetlenir. Kümeler

¹¹ Ronald Gatty, "Multivariate Analysis For Marketing Search: An Evaluation", ss 167-172

¹² Hüseyin Tatlıdil, Uygulamalı Çok Değişkenli İstatistiksel Analiz, Anaka; Cem Web Ofset Ltd.Şti., 1996, s.256

¹³ Ahmet Dirican, N Değişkenli Araştırmalarda Diskriminant Çözümlemesinin Kümeler Arası Tamı Ayırımında Kullanımı ve Bilgisayar Uygulaması, İstanbul Üniversitesi, İstanbul, 1988. s.3

arasındaki farklılığa her bir değişkenin katkısı incelenebilir ve herhangi bir bireyin girebileceği küme saptanır.

Diskriminant analizi bu işlevi yerine getirebilmek için bireyi serbest değişkenlere bağlı ve kümeler arası farkı yansıtabilen doğrusal bir şekilde belirtir.

1.2.1.5. Eşanlı Denklem Sistemleri

Basit veya çoklu regresyon modellerinde değişkenler arasındaki ilişki, tek bir denklem ile ifade edilmektedir. Burada, bağımlı Y değişkeni ile bağımsız X değişkenleri arasındaki etkileşimin tek yönlü olduğu varsayılmaktadır. Yani Y değişkenini belirleyen X değişkenleri söz konusu iken, X değişkenlerini belirleyen diğer değişkenler dikkate alınmamaktadır. Eğer bir fonksiyonda iki yönlü etkileşim varsa, bu fonksiyon tek denklemlilikli model olarak ele alınmaz. Başka bir deyişle, Y ile X arasında karşılıklı belirleme söz konusu olduğunda, bu ilişki tek bir denklem ile gösterilemez. Aksine bütün değişkenler arasındaki ilişkileri açıklayan geniş bir denklem grubu oluşturulur. Değişkenlerin karşılıklı bağımlılığını gösteren bu denklem grubuna Eşanlı denklem sistemi denmektedir.¹⁴

¹⁴ A.KOUTSOYIANNIS , Ekonometri Yöntemlerinin Tanıtımına Giriş, çev. Ü.Şenesen, G.Şenesen, 1.Baskı, Ankara, s.333

1.2.1.6. Lojistik Regresyon Analizi

Gözlemlerin gruplara ayrılmasında kullanılan yöntemlerden üç tanesi; kümeleme, diskriminant ve lojistik regresyon analizleridir. Kümeleme analizinde gözlemlerin atanacağı küme sayısı tam bilinmezken, diskriminant ve lojistik regresyon analizinde grup sayısı bilinmekte, mevcut veriler kullanılarak bir ayırsama modeli elde edilmekte ve kurulan bu model yardımı ile veri kümesine eklenen yeni gözlemlerin gruplara atanması mümkün olabilmektedir.¹⁵

1.2.1.7. Kanonik Korelasyon Analizi

İki ayrı grupta toplanmış bağımlı Y ve bağımsız X değişkenler grubu arasındaki ilişki yapısı ile ilgilenen bu analizin esas amacı, iki grup arasındaki ilişkide en yüksek korelasyonu açıklamaktır.¹⁶

Kanonik korelasyon analizi, çok boyutlu kitleden çekilmiş iki yada daha çok değişken kümesi arasındaki ilişki ile ilgilenir. Raslantı değişkenler kümesinin doğrusal fonksiyonları arasındaki maksimum korelasyonları bulmaya çalışan kanonik korelasyon analizinde tüm formülasyonlar iki raslantı değişken kümesi için geliştirilmiş olup, küme sayısını ikiden çok olması durumlarında bu formüller geliştirilerek kullanılmaktadır.

¹⁵ Hüseyin Tatlıdıl, Uygulamalı Çok Değişkenli İstatistiksel Analiz, Akademi Matbaası, Ankara, 1996, s.289

¹⁶ David A. Aker, Multivariate Analysis in Marketing Analysis in Marketing Theory & Application, Wadsworth Publishing Company, California, 1971, s.156

Kanonik Korelasyon analizi bağımsız değişkenlerle ile bağımsız değişkenlerin arasındaki ilişkinin derecesini (korelasyonunu) ortaya koyan çok değişkenli istatistiksel analizlerden biridir ve Kanonik Korelasyon analizinin başlıca amaçları; aynı bireyden elde edilen değişkenler kümesinin birbirinden bağımsız olup olmadığının test edilmesi, kümeler arası korelasyona en fazla katkıda bulunan iki kümedeki değişkenlerin saptanması ve bağımlı ve bağımsız değişkenlere ait kümeler arasındaki korelasyonu maksimum yapan doğrusal bileşimlerin belirlenmesidir.¹⁷

1.2.2. Karşılıklı bağımlılık analizleri;

Karşılıklı bağımlılık analizinde, değişkenler arasında karşılıklı bağımlılık analiz edilerek değişkenlerin temel yapılarının anlaşılmasında ve yeni değişkenler ile birleştirilmesinde kullanılmaktadır.

1.2.2.1 Temel Bileşenler Analizi

Çok değişkenli analizde değişkenlerin sayısının çok olması, değişkenler arasında belli bir ilişkiyi beraberinde getirecektir. Bu durumda değişkenlerin birbirinden bağımsız olduğu varsayımı sağlanmamaktadır. Bilgisayar olanaklarının çok geliştiği günümüzde işlem yükü bir sorun olarak görünmese de çok sayıda değişkene ilişkin analiz sonuçlarının

¹⁷ Mustafa Tekin, Kanonik Korelasyon Analizi ve Bir Uygulama , İstanbul Üniversitesi, İstanbul,, 1993 , s.65

yorumlanması ve özetlenmesi gerçekten zor olabilmektedir. Böyle durumlarda başvurulan tekniklerin en önemlisi *Temel Bileşenler Analizi*'dir.¹⁸

Temel Bileşenler Analizinin asıl amacı, p tane değişkenin değerlerinin n gözlem birimi üzerinde ölçülmesiyle elde edilen verilere dayanarak, p 'ye göre daha küçük bir sayıda r kadar yeni bir değişken belirlemektir. Temel Bileşenler olarak adlandırılan bu yeni r tane değişken, p değişkendeki orjinal değişkenlerin büyük bir kısmını açıklayabilmektedir.¹⁹

1.2.2.2. Faktör Analizi

Faktör analizinin asıl amacı çok sayıdaki değişkeni bir kaç değişken altında toplamaktır. Sıkça kullanılan çok değişkenli analiz tekniklerinden biridir. Faktör analizi p değişkenli bir olayda (p boyutlu uzay) birbiriyle ilişkili değişkenleri bir araya getirerek, az sayıda yeni (ortak) ilişkisiz değişken bulmayı amaçlar. Yani temel bileşenler analizi gibi boyut indirgeme ve bağımlılık yapısını yok etme yöntemidir.²⁰

Faktör analizi, mevcut olduğu bilinmekle beraber direkt olarak gözlemlenemeyen gizli boyutları ortaya çıkarmada, gözlenebilen değerler arasında her zaman var olduğu bilinmekle beraber görülmesi kolay olmayan karşılıklı ilişkilerin

¹⁸ Hüseyin Tatlıdil, Uygulamalı Çok Değişkenli İstatistiksel Analiz, Akademi Matbaası, Ankara, 1996, s.138

¹⁹ Ümit Fırat, Marmara Üniversitesi Sosyal Bilim Öğrencilerinin Fakülte ve Bölümler Bazında Farklılıklarının Çok Değişkenli İstatistik Teknikleri ile Analizi, Doktora Tezi, M.Ü.Yayın No.573, İstanbul, 1996, s.23

²⁰ Hüseyin Tatlıdil, a.g.e., s.167

belirlenmesinde ve çok daha fazla sayıdaki veriler setinin azaltılması ve basitleştirilmesi amaçları ile kullanılmaktadır.²¹

Faktör analizi, yorumlanması güç olmakla beraber, çok sayıda ilişkili orjinal değişkenlerden bağımsız anlamlı az sayıda faktör elde ederek, bilgi kaybının minimum olmasını ve orjinal değişkenlerdeki bilgiyi özetlemek olanağını sağlamaktadır.²²

1.2.2.3. *Kümeleme Analizi*

Kümeleme terimi, homojenliğin derecesine göre bir takım varlıkları gruplamayı amaçlayan yöntemler için kullanılır.²³

Kümeleme analizinin asıl amacı, gruplandırılmış verileri benzerliklerine göre sınıflandırmak ve araştırmacıya uygun, işe yarar özetleyici bilgiler elde etmede yardımcı olmaktadır. Bireylerin gruplandırılmasında kullanılması nedeniyle kümeleme ve diskriminant analizleri arasında benzerlik olmakla birlikte, iki yöntem arasında önemli farklılıklar da bulunmaktadır. Herşeyden önce diskriminant analizinden elde edilen bilgiler gelecekte kullanılabilir. Oysa ki kümeleme analizinde küme sayısı bilinmemekte ve sadece verilerin mevcut durumuna ilişkin sonuçlar vermesi nedeniyle gelecekte kullanılabilmesi söz konusu olamamaktadır.

²¹ Suat Karagöz, Faktör Analizi Tekniği Kullanarak Üniversite Öğrencilerinin Gazete Tercihinde Etkin Faktörlerin Belirlenmesi, Yüksek Lisans Tezi, İ.Ü., 1991, s.s20-21

²² Hüseyin Tatlıdil, a.g.e., s.167

²³ S.James Press, Applied Multivariate Analysis: Using Bayesian and Frequentist Methods of Inference, Second Edition, Robert E.Krieger Publishing Company, Florida, 1982, s.435

1.2.2.4. Çok Boyutlu Ölçekleme

Çok Boyutlu Ölçekleme, n tane nesne arasındaki uzaklıklar ile bu uzaklıklar hakkında bilgileri açıklayan ve nesnelere arasında bir ilişki bulan yöntemlerden oluşmaktadır. Bu yöntemlerin tümü, çok boyutlu uzayda, nesnelere ilişkin yapılarını ve uzaydaki konumlarını en iyi şekilde ortaya koyan geometrik yerleşim düzeyini vermeyi amaçlamaktadır. Eğer nesnelere, bir çizgi halinde görülürse tek boyut, bir yüzey üzerine dağılmış görülürse iki boyut, uzayda noktalar tarafından temsil ediliyorsa üç boyuttan bahsedilebilir. Doğrudan doğruya geometrik bir şeklin mümkün olmadığı durumlarda ise çok boyutluluk söz konusudur.²⁴

1.3. ÇOK DEĞİŞKENLİ İSTATİSTİK TEKNİKLERİNİN KARŞILAŞTIRILMASI

Çok değişkenli istatistiksel teknikler iki grupta altında incelenmiştir. İstatistiksel tekniklerin bağımlılık ve karşılıklık bağımlılık analizleri olarak ikiye ayrılmaları istatistiksel teknikler arasındaki farkı ortaya koyan bir özelliktir.

Bağımlılık analizinde, değişkenler başka değişkenlerce açıklanmaktadır. Regresyon analizinde bağımlı değişken bağımsız değişkenlerce açıklanır. Korelasyon analizinde ise bağımlı değişken ile bağımsız değişkenler arasındaki ilişkinin yönü ve derecesi açıklanmaktadır.

²⁴ Filiz Çakır, Karşılıklı Bağımlılığın Ölçülmesinde Kümeleme Analizi ve Bir Uygulama, Yüksek Lisans Tezi, M.Ü., İstanbul, 1994, s.23

Korelasyon analizi ile Kanonik Korelasyon analizi arasındaki fark bağımlı değişken sayısının birden çok olmasıdır. Bağımlı değişkenler ile bağımsız değişkenler arasındaki maksimum ilişki belirlenmektedir.

Regresyon analizi modelindeki sabit katsayı gözardı edilirse ve bağımsız X değişkenleri rastgele olarak tanımlanırsa, regresyon modeline Diskriminant fonksiyonu da denilebilir.²⁵

Regresyon analizinde bağımlı değişken bağımsız değişkenlerce açıklanırken, Eşanlı denklemler bağımsız değişkenin başka değişkenlerce açıklanabileceğini göstermektedir.

Kümeleme analizi, Diskriminant analizi ve Lojistik regresyon analizi gözlemlerin gruplara ayrılmasında kullanılan yöntemlerdir. Kümeleme analizinde küme sayısı bilinmez ve gelecekle ilgili bir tahmin yapılamaz iken Diskriminant ve Lojistik regresyon analizinde küme sayısı bilinir ve gelecekle ilgili tahmin yapılabilmektedir. Kümeleme analizinde de diskriminant analizinde olduğu gibi verilerin normal dağılımlı olması gerektiği varsayımı olmakla birlikte normallik varsayımı prensipte kalmakta ve uzaklık değerlerinin normalliği yeterli görülmektedir.

²⁵ Hüseyin Tatlıdil, Uygulamalı Çok Değişkenli İstatistiksel Analiz, Akademi Matbaası, Aknara, 1996, s.256

Diskriminant analizindeki varsayımların, Lojistik regresyon analizinde geçerli olmaması, kullanım rahatlığı ve çözümlenmeden elde edilen modelin matematiksel olarak çok esnek olması, kolay yorumlanabilir olması Lojistik regresyon analizine olan ilgiyi artırmıştır.²⁶

Çok Boyutlu ölçekleme, nesnelere arasındaki uzaklıkları kullanarak, bunlar arasındaki ilişki yapısını vermek amacıyla geliştirilmiştir. Gözlem çiftleri arasındaki uzaklıkların bulunmasının mümkün olması, çok boyutlu ölçeklemenin kullanım şansını artırmaktadır.²⁷



²⁶ Hüseyin Tatlıdil a.g.e., s.289

²⁷ Filiz Çakır, Karşılıklı Bağımlılığın Ölçülmesinde Kümeleme Analizi ve Bir Uygulama, Yüksek Lisans Tezi, M.Ü., İstanbul, 1994, s.30

İKİNCİ BÖLÜM

KOVARYANS ANALİZİ

2.1. KOVARYANS KAVRAMI

Varyans değişim ölçüsü iken, kovaryans iki değişkenin birlikte değişim ölçüsü olarak tanımlanabilir.²⁸

Kovaryans iki tesadüfi değişken arasındaki doğrusal ilişkinin yönünü belirleyen bir ölçüdür.²⁹ Matematiksel ifadesi aşağıdaki gibidir.

$$\text{Kov}(X, Y) = E (X - E(X)) (Y - E(Y))$$

Kovaryansın değeri (+) pozitif ise iki değişken değerleri aynı yönde değişim gösteriyor, (-) negatif ise değişken değerlerinden biri artarken diğeri azalıyor demektir.

Kovaryans değerinin 0 olması yani $\text{Kov}(X, Y) = 0$ iki nedenden kaynaklanmaktadır. X ile Y değişkenleri birbirinden bağımsız değişkenler ise kovaryans değeri 0 çıkacaktır. Ayrıca,

²⁸ Orhan Düzgüneş ve diğerleri, Araştırma ve Deneme Metodları, (İstatistik Metodları II), Ankara Üniversitesi Ziraat Fakültesi Yayınları, Ankara, 1987, s.229

²⁹ Selahattin Gürüş ve Şahamet Bülbül, Olasılık, Marmara Üniversitesi Nihad Sayar Eğitim Vakfı, İstanbul, 1995, s.356

X ile Y deęişkenleri doğrusal olmayan bir ilişki içindeyseler kovaryansın değeri yine 0 olarak bulunmaktadır.³⁰

Kovaryansın başlıca özellikleri;³¹

1- $Kov(X, Y)$ ile $Kov(Y, X)$ birbirine eşittir.

2- Bir tesadüfi deęişkenin kendisi ile kovaryansı varyansına eşittir.

$$Kov(X, X) = Var(X)$$

3- Tesadüfi deęişkenlerin sabit sayı ile çarpımlarının kovaryansı, sabit sayı ile deęişkenler arasındaki kovaryansın çarpımlarına eşittir.

$$Kov(aX, bY) = abKov(X, Y)$$

4- Tesadüfi deęişkenlerin herhangi bir sabit ile toplanmalarının kovaryansı, deęişkenler arasındaki kovaryansa eşittir.

$$Kov(a+X, b+Y) = Kov(X, Y)$$

5- $Kov(X+Z, Y) = Kov(X, Y) + Kov(Z, Y)$ dir.

³⁰ Mehmet Genceli, Ekonometride İstatistik İlkeler

³¹ Selahattin Güriş ve Şahamet Bülbül, a.g.e, ss.359-361

2.2. KOVARYANS ANALİZİ

Kovaryans analizi bağımlı değişken üzerindeki değişimlere etkiyen bağımsız değişkenin etkilerinin giderilmesi için uygulanan bir tekniktir. Bağımlı değişkenin üzerinde bir ya da birden çok değişkenin etkisi araştırılır ve bağımlı değişken değişirken bağımlı değişkeni etkileyen diğer değişkenleri kontrol altında tutmak mümkün değildir. Birlikte değişen bağımsız X değişkenin etkisinin giderildiği ve denemelerin bağımlı değişken üzerinden yapıldığı yönteme kovaryans çözümlemesi denmektedir.³²

Kovaryans analizi varyans ve regresyon analizlerinin özelliklerini birleştiren bir tekniktir. Bu teknik sınıflandırılmış ve sıradan regresyon analizinden elde edilen bilgi içeriğine imkan sağlayan bir veri analizini ortaya koyar.³³

Kovaryans analizi, sınıflar arasında uygun bir şekilde standartlaştırılmayan kontrol dışı değişkenlerin etkilerinin istatistiksel olarak düzeltilmesi için geliştirilmiştir³⁴. Kovaryans analizinin amacı, esas değişkene ait grup ortalamalarından birlikte değişen değişkenin etkisinin çıkartılmasıdır. Bir ya da daha çok kontrol dışı bağımsız değişkenin Y bağımlı değişkeni ile birlikte değişebileceği düşünüldüğünde önce X değişkeninin etkilerinin giderilmesine çalışılmalıdır.

³² Charles R.Hicks, Deneysel Düzenlemede İstatistiksel Yöntemler, çev. Zehra Muluk ve diğerleri, Ankara, Akademi Matbaası, 1985, s.227

³³ Ben W.Bolch and Cliff J.Huang, Multivariate Statistical Methods for Business and Economics, Prentice-Hall, Inc., New Jersey, 1989, s.185

³⁴ Yüksel İşyar, Ekonometrik Modeller, Uludağ Üniversitesi Basımevi, Bursa, 1994, s. 206

Kovaryans analizi, Varyans analizinin uygulandığı tam rastlantısal tasarım, tesadüfi blok tasarımı ve faktöryel tasarım gibi bütün deneysel düzenler için uygulanabilmektedir.

Örneğin; çeşitli yemlerin, hayvanlar üzerindeki şişmanlatma etkileri konusunda elde edilen son ağırlık değerleri kullanılarak kovaryans analizi yapılabilir. Fakat son ağırlık değerleri, deneyin başındaki ilk ağırlık değerlerinin farklılıklarının yanısıra yemlerdeki farklılıklara da dayanır. Kovaryans analizi bu baştaki farkları ayarlamamızı ya da düzeltmemizi sağlayan çok değişkenli istatistiksel tekniktir.

Bir bağımlı değişkene karşılık tek bağımsız değişkenin etkisinin giderildiği istatistiksel tekniğe Anacova denir. Kovaryans analizinin tek etkenli modeli;

Bağımlı Değişken = Sabit + Yöntemlerin Veya Grubun Etkisi + Ortak Değişim Etkisi + Hata Etkisi

Bu modelin matematiksel ifadesi;

$$Y_{ij} = \mu + \beta_j + \gamma(x_{ij} - \bar{x}_{..}) + e_{ij} \quad \begin{matrix} j=1,2,\dots,r \\ i=1,2,\dots,n_j \end{matrix}$$

şeklinde ifade edilir.³⁵

³⁵ William C. Guenter, Analysis of Variance, Universty of Wyoming, Prentice-Hall, Englewood Cliffs, s.145

e_{ij} normal dağılımlı 0 ortalamalı σ^2 varyanslı hata terimidir.

$$e_{ij} \sim N(0, \sigma^2)$$

γ Regresyon katsayısı, tek etkende bağımlı veri üzerine ortak değişkenin seviye etkisini gösterir.

$$\sum_{j=1}^r n_j \beta_j = 0$$

Bu değerleri farklı bir yolla aşağıdaki şekilde yazmak mümkündür.³⁶

$$Y_{ij} \sim N(\mu_{ij}, \sigma^2)$$

$$\mu_{ij} = \mu + \beta_j + \gamma (x_{ij} - \bar{x}_{.j}) \quad \begin{matrix} i=1,2,\dots,n_j \\ j=1,2,\dots,r \end{matrix}$$

$$\sum_{j=1}^r n_j \beta_j = 0$$

β_j 'lerin 0 olması bağımlı ve bağımsız değişkenlerin arasında bir etkileşimin olmadığı anlamına gelir. β_j 'lerin 0 dan farklı olmaları ise bir etkinin olduğunu gösterir. Her bir grup için doğruların eğiminin aynı olması yani homojen ve r grup

³⁶ William C. Guenter, a.g.e., s. 145

için ya da yöntemler için doğruların birbirine paralel olması gereklidir.

MANCOVA analizinde ise, bağımlı değişken birden çok bağımsız değişkene karşılık gelir. Değişkenler arasında bir ilişkinin olduğu ve bu ilişkilerin Y bağımlı değişkeni üzerindeki etkilerinin giderildiği yöntemdir. İki bağımsız değişken olması durumunda model,

$$Y_{ij} = \mu + \beta_j + \gamma_1(x_{ij} - \bar{x}_{..}) + \gamma_2(z_{ij} - \bar{z}_{..}) + e_{ij} \quad \begin{matrix} j=1,2,\dots,r \\ i=1,2,\dots,n_j \end{matrix}$$

şeklinde ifade edilebilir.³⁷

Varyans analizinde yapılan, e_{ij} 'nin tesadüfi değişken olduğu, normal dağılıma sahip olduğu ve gruplar için varyansın homojen olduğu varsayımları kovaryans analizi için de geçerlidir. Ancak regresyon modelinin doğrusal olduğu, eğimin 0 dan farklı olduğu, her bir grup için regresyon katsayılarının homojen olduğu, bağımsız değişken X'in gruplara uygulanan denemeden etkilenmediği varsayımları, kovaryans analizini varyans analizinden ayıran özelliklerdir.

³⁷ Ben W. Bolch and Cliff J. Huang, *Multivariate Statistical Methods for Business and Economics*, Vanderbilt University, New Jersey, 1985, s.192

2.2.1. TAM RASTLANTISAL TASARIM İÇİN KOVARYANS ANALİZİ

Tam rastlantısal tasarım, bağımlı değişkenin gruplar arasındaki farklılıklarının saptanmasına ilişkin bir tasarımdır. Örnek birimleri tesadüfi olarak meydana gelen gözlem değerlerini ifade eder.³⁸

Uygulamada bağımsız değişken, bağımlı değişkenin herbir gözlemi için ayrı bir değer alacaktır.

Kovaryans analizi bağımlı değişkenin her bir gözlemi için birlikte değişen bağımsız değişken değerine göre ayarlanan bir uygulamadır. Böylece tam rastlantısal tasarımda grup etkileri hakkında daha doğru bilgi edinilmesi mümkün olur. Daha iyi bir sonuç için hata varyansının azaltılması yani birlikte değişen bağımsız değişken etkisinin giderilmesi, bağımlı değişken üzerinde grup ortalamalarını ilgili bağımsız değişken gruplarının ortalama değerlerindeki değişikliklere göre ayarlamaları ile mümkün olur.

2.2.1.1. Gözlem Değerlerinin Düzenlenmesi

Tam rastlantısal tasarımda deney sonunda elde edilen gözlemler aşağıdaki şekilde düzenlenmektedir.

³⁸ Albert R. Wildth and Olli Ahtola, Analysis of Covariance, Universty George and Universty of Florida, 1985, s.19

	Gruplar										
	1		2		r						
	x	y	x	y	x	y					
1	x ₁₁	y ₁₁	x ₁₂	y ₁₂	·	·	·	x _{1r}	y _{1r}		
2	x ₂₁	y ₂₁	x ₂₂	y ₂₂	·	·	·	x _{2r}	y _{2r}		
·	·	·	·	·	·	·	·	·	·		
·	·	·	·	·	·	·	·	·	·		
n _j	x _{nj1}	y _{nj1}	x _{nj2}	y _{nj2}	·	·	·	x _{njr}	y _{nj1}		
Toplam	T _{x,1}	T _{y,1}	T _{x,2}	T _{y,2}	·	·	·	T _{x,r}	T _{y,r}	T _{x..}	T _{y..}
Ortalama	$\bar{x}_{.1}$	$\bar{y}_{.1}$	$\bar{x}_{.2}$	$\bar{y}_{.2}$	·	·	·	$\bar{x}_{.r}$	$\bar{y}_{.r}$	$\bar{x}_{..}$	$\bar{y}_{..}$

$$T_{x,j} = \sum_{i=1}^{n_j} x_{ij} \quad \text{x'in j. sütun toplamları}$$

$$\bar{x}_{.j} = T_{x,j} / n_j \quad \text{x'in sütun ortalaması}$$

$$T_{y,j} = \sum_{i=1}^{n_j} y_{ij} \quad \text{y'nin j. sütun toplamları}$$

$$\bar{y}_{.j} = T_{y,j} / n_j \quad \text{y'nin sütun ortalaması}$$

$$T_{x..} = \sum_{j=1}^r \sum_{i=1}^{n_j} x_{ij} \quad \text{Tüm x'lerin toplamları}$$

$$\bar{x}_{..} = T_{x..} / N \quad \text{Tüm x'lerin ortalaması}$$

$$N = \sum_{j=1}^r n_j$$

$$T_{y..} = \sum_{j=1}^r \sum_{i=1}^{n_j} y_{ij} \quad \text{Tüm } y\text{'lerin toplamları}$$

$$\bar{y}_{..} = T_{y..} / N \quad \text{Tüm } y\text{'lerin ortalaması}$$

2.2.1.2 Tam Rastlantısal Tasarım Modeli

Kovaryans analizinde her bir Y ölçümü yapılırken Y ölçümü ile birlikte her birimde X değişkeninin de ölçümünün yapılabilmesi ve Y değişkeninin birlikte değişken bağımsız X değişkenlerinden etkilenmesi söz konusudur.

Varyans analizinde tam rastlantısal tasarım için model denklemi,

$$Y_{ij} = \mu + \beta_j + e_{ij}$$

şeklinde ifade edilir.³⁹

μ ortak etkiyi, β_j grup etkisini ve e_{ij} hatayı belirtmektedir. Varyans analizin de grup etkileri test edilir. Test sonucunda grup etkileri söz konusu ise bu grupların birbirlerinden farklı olmalarına sebep bir bağımsız X değişkeni olabilir. Eğer bağımlı ve bağımsız değişkenler arasında kuvvetli bir ilişki var ise yani X ile Y değişkenleri arasında pozitif korelasyon mevcut ise regresyon denklemi,

³⁹ Richard A. Johnson and Dean W. Wichern. Applied Multivariate Statistical Analysis, University of Wisconsin-Madison, s.252

$$Y_{ij} = \mu + \gamma (X_{ij} - \bar{X}) + e_{ij}$$

şeklinde ifade edilir.⁴⁰

her iki denklem birleştirildiğinde,

Bağımlı Değişkenin= (Sabit) + UygulamaSeviyesi + Kovaryans +Artık
Gözlem Değeri veya Grup Etkisi Etkisi Etkisi

$$Y_{ij} = \mu + \beta_i + \gamma (X_{ij} - \bar{X}) + e_{ij}$$

sonucu elde edilen model kovaryans modelidir. Varyans Analizi ve Regresyon analizine ait model denklemlerinden çıkarılan sonuç olan Kovaryans analizi, Varyans analizi ve Regresyon analizinin bir bileşimidir.

Kovaryans analizi bu durumda bağımsız değişkenlerin, bağımlı değişken üzerindeki etkisini saptamak amacıyla bir veya daha çok bağımsız değişkenin etkisini yok etmek için yapılır.⁴¹

Y_{ij} : i. gruptaki j. gözlemin bağımlı değişkenin değerini ifade eder.

X_{ij} : i. gruptaki gözlemin birlikte değişen değeridir.

⁴⁰ Charles R.Hicks, Deney Düzenlemede İstatistiksel Yöntemler,çev. Zehra Muluk ve diğerleri, Ankara, Akademi Matbaası, 1985, s.229

⁴¹ Charles R.Hicks, a.g.e., s.230

γ : Regresyon katsayısıdır. Birlikte değişen değişkendeki bir birimlik değişimin bağımlı değişken üzerindeki ortalama etkisini ifade eder. γ en küçük kareler yöntemi ile elde edilir ve genel regresyon katsayısı olarak kabul edilir.

e_{ij} 'ler bağımsızdır, normal dağılımlı 0 ortalamalı σ^2 varyanslıdır.

$$e_{ij} \sim N(0, \sigma^2)$$

Kısaca kovaryans analizinin tam rastlantısal tasarımı için varsayımlar aşağıdaki gibidir,

- a) Her bir örnek N birimlik anakütlesinden tesadüfi olarak seçilmiştir.
- b) N birimli anakütle normal dağılımlıdır.
- c) N birimli anakütle aynı varyansa sahiptir.
- d) Her bir grup içindeki anakütle ortalaması bir doğru üzerinde yer alır.
- e) Doğruların eğimi her bir grup için aynıdır.

Sonuç olarak d ve e varsayımları varyans analizini kovaryans analizinden ayıran özelliklerdir.⁴²

Kovaryans analizinin genel amacı, grup ortalamalarının eşit olduğu hipotezini test etmek, birlikte değişimin kaynaklarını ve bunların etki ve derecelerini tesbit etmek, X ile

⁴² William C. Guenther, Analysis of Variance, University of Wyoming, Prentice-Hall, Englewood Cliffs, s.146

Y deęişkenleri için elde edilen verilerle kareler toplamları ve çarpımlar toplamlarını varyans analizinde olduğu gibi hesaplamaktır. Kovaryans analizinde birlikte deęişen deęişkenin etkisi giderilmekte ve düzeltilmiş kareler toplamları elde edilmektedir.

Kareler, çarpımlar toplamları ve düzeltilmiş kareler toplamları için aşağıdaki hesaplamalar yapılmalıdır.

Birlikte deęişen deęişken için kareler toplamı,

$$SS_{XT} = SS_{XT_r} + SS_{X_E}$$

$$SS_{XT} = \sum_{j=1}^r \sum_{i=1}^{n_j} x_{ij}^2 - Tx^2 \dots / N$$

Etkilenen bağımlı deęişken için kareler toplamı,

$$SS_{YT} = SS_{YT_r} + SS_{Y_E}$$

$$SS_{YT} = \sum_{j=1}^r \sum_{i=1}^{n_j} y_{ij}^2 - Ty^2 \dots / N$$

Birlikte deęişen deęişkenin grup etkilerinin kareler ortalaması,

$$SS_{XT} = \sum_{j=1}^r Tx^2 \dots / n_j - Tx^2 \dots / N$$

Birlikte deęişen deęişkenin hata kareler toplamı,

$$SS_{XE} = SS_{XT} - SS_{XT_r}$$

Etkilenen baęımlı deęişkenin grup etkişlerinin kareler ortalaması,

$$SS_{YT} = \sum_{j=1}^r T y^2_{.j} / n_j - T y^2_{..} / N$$

Baęımlı deęişken için hata kareler toplamı,

$$SS_{YE} = SS_{YT} - SS_{YT_r}$$

Baęımlı ve baęımsız deęişkenler için çarpımlar toplamı,

$$SP_T = \sum \sum (x_{ij} - \bar{x}_{..})(y_{ij} - \bar{y}_{..})$$

$$SP_T = SP_{T_r} + SP_E$$

$$SP_E = SP_T - SP_{T_r}$$

Düzeltilmiş kareler toplamları;

$$SS'_{YE} = SS_{YE} - SP_E^2 / SS_{XE}$$

$$SS'_{YT} = SS_{YT} - SP_T^2 / SS_{XT}$$

$$SS'_{T_r} = SS'_{YT_r} - SS'_{YE}$$

Elde edilen hesaplamalar sonucunda oluşturulacak Kovaryans Analiz tablosu aşağıdaki gibidir.

Tam Rastlantısal Tasarım için KOVARYANS ANALİZİ

Kaynak	Kareler ve Çarpımlar Toplamı			Düzeltilmiş Kareler	Serbestlik Derecesi	Düzeltilmiş Kareler	F Oranı
				Toplamı		Ortalaması	
	XX	XY	YY	SS'		MS' _{YY}	
Gruplar Arası	SS _{XTr}	SP _{Tr}	SS _{YTr}	SS' _{YTr}	r-1	MS' _{YTr}	MS' _{YTr}
Hata	SS _{XE}	SP _E	SS _{YE}	SS' _{YE}	N-r-1	MS' _{YE}	MS' _E
Toplam	SS _{XT}	SP _T	SS _{YT}	SS' _{YT}			

$$SS_{XE} = \sum \sum (X_{ij} - \bar{X}_i)^2$$

$$SP_E = \sum \sum (X_{ij} - \bar{X}_i) (Y_{ij} - \bar{Y}_i)$$

$$SS_{YE} = \sum \sum (Y_{ij} - \bar{Y}_i)^2$$

$$SS_{XT} = \sum \sum (X_{ij} - \bar{X})^2$$

$$SP_T = \sum \sum (X_{ij} - \bar{X}) (Y_{ij} - \bar{Y})$$

$$SS_{YT} = \sum \sum (Y_{ij} - \bar{Y})^2$$

$$SS'_{YE} = SS_{YE} - SP_E^2 / SS_{XE}$$

$$SS'_{YT} = SS_{YT} - SP_T^2 / SS_{XT}$$

$$SS'_{Tr} = SS'_{YT} - SS'_{YE}$$

2.2.1.3. Hipotez Testi

Varyans analizi için geçerli varsayımların kovaryans analizi için de geçerli olduğunu daha önce belirtmiştik, Bunlar, normal dağılımlı, bağımsız dağılmış hata grupları için varyansın homojenliği gibi varsayımlardı. Kovaryans çözümlemesinde buna ek olarak regresyonun doğrusal olduğu, eğimin sıfır olmadığı, her bir grup için regresyon katsayılarının homojen olduğunu ve bağımsız değişken X'in gruplara uygulanan denemeden etkilenmediği varsayımları da yapılmakta ve bu varsayımların doğruluğu hipotez testleri yardımı ile test edilmektedir.

Eşitliğin her bir kolonu için Y_{ij} doğrusaldır ve her bir r grubu için doğrunun eğimi aynıdır. Doğrusallık varsayımı için X ve Y değişkenlerine ait gözlem değerlerinin yayılım çizgisine bakmak yeterli olacaktır.

$$H_0: \beta_j=0 \quad (j=1,2,\dots,r)$$

$$H_1: \beta_j \text{ 'ler } 0 \text{ 'dan farklıdır.}$$

Regresyonun homojen olup olmadığına karar verebilmek için $\beta_1=\beta_2$ yani regresyon katsayıları arasında fark olmadığı şeklindeki sıfır hipotezinin test edilmesi gerekir.

Eğim β nın sıfırdan farklı olduğuna dair hipotezi test etmek için, hata kareler toplamlarındaki azalmanın önemli olup olmadığına hesaplaması yapılır. Bunun için aşağıdaki hesaplama yapılır.

$$F_{1, N-r-1} = (SP_E)^2 / S_{XE} / MS'_{YE}$$

1, (N-r-1) serbestlik dereceli F dağılımına sahiptir ve α hata ile 1, (N-r-1) serbestlik dereceli F tablo değeri ile karşılaştırılır. Hesaplanan F değeri tablo F değerinden büyükse H_0 red edilecek ve eğimin 0'dan farklı olduğu hipotezi test edilmiş olacaktır.

$$H_0: \mu_{.1} - \gamma \bar{x}_{.1} = \mu_{.2} - \gamma \bar{x}_{.2} = \dots = \mu_{.r} - \gamma \bar{x}_{.r}$$

$$H_1: \mu_{.j} - \gamma \bar{x}_{.j} \text{ birbirine eşit değildir.}$$

Daha sonra $\mu_1 = \mu_2$ yani, X değişkenine göre düzeltilmiş ve denk duruma getirilmiş Y değerleri ortalamaları arasında fark olmadığı şeklindeki sıfır hipotezi test edilir. Birinci hipotezde kovaryans analizinin çözümünde eğimin 0 olmaması ancak her bir grup içindeki regresyon katsayılarının eşit (homojen) olduğu ve X birlikte değişen değişkenin Y değişkenini etkilemediği varsayılır.

Gruplar için regresyon katsayısının eşitliği hipotezinin test etmek için herbir örneklem katsayısının hesaplamak ve her bir grup içi kareler toplamlarını kendi regresyon katsayıları ile düzeltmek gerekmektedir. Bunun için yapılacak hesaplamalar aşağıdaki gibidir.

$$b = \Sigma xy / \Sigma x^2$$

$$b = (\Sigma xy - \Sigma x \Sigma y) / (\Sigma x^2 - \Sigma x^2 / n)$$

$$\text{düzeltmiş } \Sigma y^2 = \Sigma y^2 - \Sigma xy / \Sigma x^2$$

$$F_{k-1, N-2k} = \left[\text{düzeltmiş } \Sigma y^2 \text{ toplam grup içi regresyona dayalı} - \text{düzeltmiş } \Sigma y^2 \text{ her bir grup içi regresyona dayalı} \right] / (k-1)$$

$$\text{düzeltmiş } \Sigma y^2 \text{ her bir grup içi reg dayalı } / N-2k$$

k-1, N-2k serbestlik dereceli F dağılımına sahiptir ve α hata ile F tablo değeri ile karşılaştırıldığında F tablo değeri hesap değerinden büyükse H_0 kabul edilecek ve regresyon katsayılarının her bir grup için homojen olduğu test edilmiş olacaktır.

$$H_0 : \gamma = 0$$

$$H_1 : \gamma \neq 0$$

Daha sonraki test istatistiği, bağımsız değişken X'in gruplara uygulanan denemelerden etkilenip etkilenmediğinin incelenmesi için yapılan testtir. Bunun için,

$$F_{r-1, n-r-1} = (SS_{XE}/r-1) / (SS_{XT} / N-r-1)$$

F değeri hesaplanır. r-1 , N-r-1 serbestlik dereceli F dağılımına sahiptir. α hata ile F tablo değeri ile karşılaştırılır. Sonuçta F tablo değeri hesap değerinden büyükse sıfır hipotezi kabul edilir ve X'in gruplara uygulanan denemelerden etkilenmediği söylenir.

2.2.2. TESADÜFİ BLOK TASARIMI İÇİN KOVARYANS ANALİZİ

Bazı çalışmalarda tek etkenli düzenlemeler için, bir kategoride bağımlı değişkenin değerini etkileyen birden çok faktör olabilir. Bu faktörden dolayı bağımlı değişkeni bir kategoride değil farklı iki kategoride nitelendirmek gerekir. Tesadüfi blok tasarımı için düzenlenen bağımlı değişkenin iki kategorisinden biri grup diğeri de blok etkisini gösterir. Dolayısıyla blok etkisinin hata varyansında bir azalma yapması söz konusu olur.

Sonuçta deneyin güvenilirliğini artırmak için, deney tasarımı iki kategoride yani grup ve blok etkisi olarak düşünülür. Bu etkenlerden herhangi birinin olmaması durumunda tasarım tam rastlantısal tasarım olur.

Herhangi bir grubun içindeki değişkenlik sonucunda iki kategorinin olması gerekir ve bu değişkenlik gruplar arasındaki değişkenlikten daha az olabilir. Bu nedenle blok tasarımının homojenlikten kaynaklanan değişimleri kontrol ederek deneyin hata payının azaltılmasını sağlayan bir kontrol şekli olduğu görülür.⁴³

Tesadüfi blok tasarımı için gerekli ek bir şart, uygulama seviyelerinin her bir blok içindeki test ünitelerine tesadüfi olarak dağıtılabilmesidir. Bu tesadüfi dağıtım işleminin bir

⁴³ Albert R. Wildth and Olli Ahtola, Analysis of Covariance, Universty George and Universty of Florida, 1985, s.57

istisnası, bloğun bütün seviyelerine tabi tutulan tek bir birimden oluşmasıdır. Her durumda, deneyin niteliği izin verdiği ölçüde, uygulama seviyelerinin sıralaması her bir test ünitesi itibariyle rastlantısal olarak belirlenir.

Tesadüfi blok tasarımı için gözlenen verilerin düzenlenmesi aşağıdaki gibidir.

2.2.2.1. Gözlem Değerlerini Düzenlenmesi

	Gruplar						Toplam	Ortalama
	1		2		r			
Bloklar	x	y	x	y	x	y		
1	x ₁₁	y ₁₁	x ₁₂	y ₁₂	x _{1r}	y _{1r}	T _{x1.} T _{y1.}	$\bar{x}_1.$ $\bar{y}_1.$
2	x ₂₁	y ₂₁	x ₂₂	y ₂₂	x _{2r}	y _{2r}	T _{x2.} T _{y2.}	$\bar{x}_2.$ $\bar{y}_2.$
.
.
n	x _{n1}	y _{n1}	x _{n2}	y _{n2}	x _{nr}	y _{nr}	T _{xn.} T _{yn.}	$\bar{x}_{xn.}$ $\bar{y}_{n.}$
Toplam	T _{x.1}	T _{y.1}	T _{x.2}	T _{y.2}	T _{x.r}	T _{y.r}	T _{x..} T _{y..}	
Ortalama	$\bar{x}_{.1}$	$\bar{y}_{.1}$	$\bar{x}_{.2}$	$\bar{y}_{.2}$	$\bar{x}_{.r}$	$\bar{y}_{.r}$		$\bar{x}_{..}$ $\bar{y}_{..}$

$$Tx_i = \sum_{j=1}^r x_{ij} \quad x' \text{in } i. \text{ blok toplamları}$$

$$x_{.j} = Tx_{.j} / r \quad x' \text{in blok ortalaması}$$

$$Ty_i = \sum_{j=1}^r y_{ij} \quad y\text{'nin } i. \text{ blok toplamları}$$

$$\bar{y}_{.j} = Ty_{.j} / r \quad y\text{'nin } i. \text{ blok ortalaması}$$

2.2.2.2. Tesadüfi Blok Tasarımı Modeli

Bağımlı Değişkenin Gözlem Değeri = (Sabit) + UygulamaSeviyesi veya Grup Etkisi + **Blok Etkisi** + Kovaryans Etkisi + Artık Etkisi

$$Y_{ij} = \mu + \alpha_i + \beta_j + \gamma (x_{ij} - \bar{x}_{..}) + e_{ij} \quad \begin{matrix} i=1,2,\dots,n_i \\ j=1,2,\dots,r \end{matrix}$$

şeklinde ifade edilir.⁴⁴

e_{ij} 'ler bağımsızdır, normal dağılımlı, 0 ortalamalı σ^2 varyanslıdır.

$$e_{ij} \sim n(0, \sigma^2)$$

$$\sum_{i=1}^n \alpha_i = \sum_{j=1}^r \beta_j = 0$$

⁴⁴ William C.Guenther , Analysis of Variance , Universty of Wyoming, Prentice-Hall, Englewood Cliffs,s.152

Y_{ij} 'ler bağımsızdır, normal dağılımlı, μ_{ij} ortalamalı σ^2 varyanslıdır.

$$Y_{ij} \sim n(\mu_{ij}, \sigma^2)$$

$$\mu_{ij} = \mu + \alpha_i + \beta_j + \gamma(x_{ij} - \bar{x} \dots) \quad j=1,2,\dots,r$$
$$i=1,2,\dots,n_i$$

şeklinde ifade edilebilir.⁴⁵

$$\sum_{i=1}^n \alpha_i = \sum_{j=1}^r \beta_j = 0$$

Kovaryans analizinde tesadüfi blok tasarımı için gerekli varsayımlar aşağıdaki gibidir,

- a) Herbir örneklem rn birimli anakütleden tesadüfi olarak seçilmiştir.
- b) rn birimli anakütle normal dağılımlıdır.
- c) Bloklar ve gruplar birbirlerini etkilemez.
- d) Anakütle ortalamaları aynı doğru üzerindedir.
- e) Eğimler her bir grup için aynıdır.
- f) rn anakütle için grupların eğimi aynıdır.⁴⁶

⁴⁵ William C.Guenther , a.g.e., s.152

⁴⁶ William C.Guenther , a.g.e, s.153

Varyans analizinde olduğu gibi blok etkilerinin testi için kullanılan çözüm yöntemleri, kovaryans analizi için de geçerlidir.

Tesadüfi blok tasarımının çözümü için;

Birlikte değişen değişken için kareler toplamı,

$$SS_{XT} = SS_{XB} + SS_{XT_r} + SS_{XE}$$

Etkilenen bağımlı değişken için kareler toplamı,

$$SS_{YT} = SS_{YB} + SS_{YT_r} + SS_{YE}$$

Çarpımların kareler toplamı,

$$SP_T = \sum_{j=1}^r \sum_{i=1}^n (x_{ij} - \bar{x}_{.j}) (y_{ij} - \bar{y}_{.j})$$

$$SP_T = \sum_{j=1}^r \sum_{i=1}^n X_{ij} Y_{ij}$$

$$X_{ij} = (\bar{x}_{i.} - \bar{x}_{..}) + (\bar{x}_{.j} - \bar{x}_{..}) + (x_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x}_{..})$$

$$Y_{ij} = (\bar{y}_{i.} - \bar{y}_{..}) + (\bar{y}_{.j} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})$$

$$SP_T = \sum_{i=1}^r \sum_{j=1}^n (\bar{x}_{i.} - \bar{x}_{..})(\bar{y}_{.j} - \bar{y}_{..}) + \sum_{j=1}^r \sum_{i=1}^n (x_{.j} - \bar{x}_{..})(\bar{y}_{.j} + \bar{y}_{..}) +$$

$$\sum_{i=1}^r \sum_{j=1}^n (x_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x}_{..}) (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}) + 0$$

$$SP_T = SP_B + SP_{Tr} + SP_E$$

Bloklara ait kareler toplamı,

$$SP_B = \sum_r \frac{T_{X_i.} T_{Y_{.j}}}{r} - \frac{T_{X_{i.}} T_{Y_{.j}}}{rn}$$

Blok etkisi düşüldükten sonraki hata kareler toplamı,

$$SP_E = SP_T - SP_B - SP_{Tr}$$

Düzeltilmiş kareler toplamları,

$$SS'_{YE} = SS_{YE} - (SP_E)^2 / SS_{YE}$$

$$SS'_{Y(Tr + E)} = SS_{YTr} + SS_E - (SP_{Tr} + SP_E)^2 / (SS_{XTr} + SS_{XE})$$

$$SS'_{YTr} = SS'_{Y(Tr + E)} - SS'_{YE}$$

Elde edilen hesaplamalar sonunda tesadüfi blok tasarımı için oluşturulacak Kovaryans analizi tablosu aşağıdaki gibidir.

Tesadüfî Blok Tasarımı için KOVARYANS ANALİZİ

Kaynak	Kareler ve Çarpımlar Toplamı			Düzeltilmiş Kareler	Serbestlik Derecesi	Düzeltilmiş Kareler Ortalaması	F Oranı
				Toplamı		MS'Y	
	XX	XY	YY	Y'			
Blok Etkisi	SS_{XB}	SP_B	SS_{YB}	SS'_{YB}	r-1	MS'_{YB}	MS'_{YB}/MS'_{YE}
Gruplar Arası	SS_{XTr}	SP_{Tr}	SS_{YTr}	SS'_{YTr}	n-1	MS'_{YTr}	MS'_{YTr}/MS'_{YE}
Hata	SS_{XE}	SP_E	SS_{YE}	SS'_{YE}	N-r-1	MS'_{YE}	MS'_{YE}
Toplam	SS_{XT}	SP_T	SS_{YT}	SS'_{YT}			

$$SS_{XB} = r \sum (X_{ij} - \bar{X}_i)^2$$

$$SP_B = r \sum (X_{ij} - \bar{X}_i)(Y_{ij} - \bar{Y}_i)$$

$$SS_{YB} = r \sum (Y_{ij} - \bar{Y}_i)^2$$

$$SS_{XTr} = n \sum (X_{ij} - \bar{X}_i)^2$$

$$SP_{Tr} = n \sum (X_{ij} - \bar{X}_i)(Y_{ij} - \bar{Y}_i)$$

$$SS_{YTr} = n \sum \sum (Y_{ij} - \bar{Y}_i)^2$$

$$SS'_{YE} = SS_{YE} - SP_E^2 / SS_{XE}$$

$$SS'_{YT} = SS_{YT} - SP_T^2 / SS_{YT}$$

$$SS'_{Tr} = SS'_{YT} - SS'_{YE}$$

2.2.2.4. Hipotez Testi

Tesadüfi blok tasarımındaki varsayımların doğruluğunu test etmek için yapılması gereken hesaplamalar,

- Hata kareler toplamları ve çarpımlar toplamı
 SS_{XE}, SP_E, SS_Y
- Gruplar arası kareler toplamları ve çarpımlar toplamları
 $SS_{XTr}, SS_{PTr}, SS_{YTr}$
- Düzeltilmiş hata ve gruplar arası kareler toplamları
 SS'_{YTr}, SS'_E ve
- F oranı'dır.

Tesadüfi bloklar için yapılan kovaryans analizinde, regresyon katsayısına ilişkin bazı varsayımlar yapılır. Bunlar, regresyonun bütün gruplar için homojen olduğu ve regresyon katsayısının sıfır olmaması gibi, $\beta=0$ sıfır hipotezinin testi için F değeri,

$$F = SP^2_{E (N-r-1)} / (SS_{YE} - SS_{XE} - SP^2_E)$$

şeklinde ifade edilir.⁴⁷ Burada hata karelerinin terimleri SP_E, SS_{YE}, SS_{XE} dir. 1, (N- r - 1) serbestlik dereceli F dağılımına sahiptir.

⁴⁷ Albert R. Wildthand Olli Ahtola, Analysis of Covariance, Universty George and Universty of Florida, 1985, s.63

Bütün gruplar için regresyonun homojenliği testini tesadüfi bloklar için yapmak oldukça zordur. Regresyon katsayılarının bütün uygulama-blok kombinasyonlarında eşit olduğu hipotezi test edilir, ancak uygulama-blok kombinasyonu başına sadece bir gözlem olduğunda böyle bir testi yapmak mümkün değildir. Regresyonun homojenliğini, bloklarda regresyon katsayılarının eşit olduğu varsayımı altında test etmektir. Bu testi yapmak için, tamamen rastlantısal tasarımda kullanılan yöntemle benzer bir işlem yapılabilir.⁴⁸

Hipotez testi;

$$H_0: \beta_j=0 \quad (j=1,2,\dots,r)$$

$H_1: \beta_j$ 'ler 0'dan farklıdır.

$$H_0: \alpha_i=0 \quad (i=1,2,\dots,r)$$

$H_1: \alpha_j$ 'ler 0'dan farklıdır.

$$H_0: \mu_{.1} - \gamma \bar{x}_{.1} = \mu_{.2} - \gamma \bar{x}_{.2} = \dots = \mu_{.r} - \gamma \bar{x}_{.r}$$

$H_1: \mu_{.j} - \gamma \bar{x}_{.j}$ 'ler birbirinden farklıdır.

⁴⁸ Albert R. Wildth and Olli Ahtola, a.g.e., s.63

2.2.3. FAKTÖRYEL TASARIM İÇİN KOVARYANS ANALİZİ

Faktörel tasarımda, iki veya daha çok kategorik bağımsız değişken eş zamanlı değerlendirilmektedir. Buna benzer pek çok olay faktörel tasarım sınıfına sokulan değişik tasarımlarla ele alınır. Bu tasarımda araştırmacı bütün kategorik bağımsız değişkenlerin etkisiyle eş zamanlı olarak ilgilenir. Test üniteleri, araştırmacının bütün kategorik bağımsız değişkenlerin çeşitli seviye kombinasyonlarına tesadüfi olarak ayrılır. Bu özellikle tesadüfi blok tasarımından farklıdır. Bu tasarımda test ünitesi karakteristikleri araştırmacının kontrolü altında değildir.

Faktörel tasarımda ikiden fazla bağımsız değişken olabilir. Kategorik bağımsız değişkenlere ek olarak niceliksel bir bağımsız değişken ilave edilirse, kullanılan analiz prosedürü *Kovaryans Faktörel Analizi* olur. Bu tasarımda iki kategorik bağımsız değişken ele alınır; birincisi p seviyesinde tasarım faktörü A ve ikincisi q seviyesinde tasarım faktörü B dir. pq uygulama kombinasyonlarında, birbirine rastlantısal olarak ayrılmış eşit sayıda test ünitesi ve birlikte değişen bağımsız değişken vardır.⁴⁹

⁴⁹ Albert R. Wildth and Olli Ahtola, a.g.e., s.69

2.2.3.2. Gözlem Değerlerinin Düzenlenmesi

Faktörel tasarım için gözlemlerin düzenlenmesi aşağıdaki gibidir.

		Gruplar A Etkeni için							
		A1		A2		Aa		Toplam	Ortalama
		x	y	x	y	x	y		
B1		X ₁₁₁	Y ₁₁₁	X ₁₂₁	Y ₁₂₁	X _{1a1}	Y _{1a1}	T _{x1} · T _{y1} ·	\bar{x}_1, \bar{y}_1
		X ₁₁₂	Y ₁₁₂	X ₁₂₂	Y ₁₂₂	X _{1a2}	Y _{1a2}		
		X ₁₁₃	Y ₁₁₃	X ₁₂₃	Y ₁₂₃	X _{1a3}	Y _{1a3}		
			
		X _{11k}	Y _{11k}	X _{12k}	Y _{12k}	X _{1ak}	Y _{1ak}		
B2		X ₂₁₁	Y ₂₁₁	X ₂₂₁	Y ₂₂₁	X _{2a1}	Y _{2a1}	T _{x2} · T _{y2} ·	\bar{x}_2, \bar{y}_2
		X ₂₁₂	Y ₂₁₂	X ₂₂₂	Y ₂₂₂	X _{2a2}	Y _{2a2}		
		X ₂₁₃	Y ₂₁₃	X ₂₂₃	Y ₂₂₃	X _{2a3}	Y _{2a3}		
			
		X _{21k}	Y _{21k}	X _{22k}	Y _{2k}	X _{2ak}	Y _{2ak}		
Bb		X _{b11}	Y _{b11}	X _{b21}	Y _{b21}	X _{ba1}	Y _{ba1}	T _{xb} · T _{yb} ·	x_b, \bar{y}_b
		X _{b12}	Y _{b12}	X _{b22}	Y _{b22}	X _{ba2}	Y _{ba2}		
		X _{b13}	Y _{b13}	X _{b23}	Y _{b23}	X _{ba3}	Y _{ba3}		
			
		X _{b1k}	Y _{b1k}	X _{b2k}	Y _{b2k}	X _{bak}	Y _{bak}		
Toplam		T _{x,1}	T _{y,1}	T _{x,2}	T _{y,2}	T _{x,r}	T _{y,r}	T _{x...}	T _{y...}
Ortalama		$\bar{x}_{.1}$	$\bar{y}_{.1}$	$\bar{x}_{.2}$	$\bar{y}_{.2}$	$\bar{x}_{.r}$	$\bar{y}_{.r}$	$\bar{x}_{..}$	$\bar{y}_{..}$

2.2.3.2. Faktöriyel Tasarımı Modeli

Faktöriyel tasarım için kovaryans modeli;

Bağımlı Değişkenin= (Sabit) + A Etkeni+ B Etkeni + AB Etkeni + Artık
Gözlem Değeri Etkisi Etkisi Etkisi Etkisi

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \gamma(x_{ijk} - \bar{x}_{...}) + e_{ijk} \quad \begin{array}{l} j=1,2,\dots,b \\ i=1,2,\dots,a \\ k=1,2,\dots,n \end{array}$$

şeklinde ifade edilir.⁵⁰

e_{ijk} 'ler bağımsızdır, normal dağılımlı, 0 ortalamalı, σ^2 varyanslıdır.

$$e_{ijk} \sim (0, \sigma^2) \quad (j=1,2,\dots,b \quad i=1,2,\dots,a \quad k=1,2,\dots,n)$$

Y_{ij} 'ler bağımsızdır, normal dağılımlı, μ_{ij} ortalamalı, σ^2 varyanslıdır.

$$Y_{ij} \sim (\mu_{ij}, \sigma^2) \quad (j=1,2,\dots,r \quad i=1,2,\dots,n_j)$$

$$\sum_{i=1}^a \alpha_i = \sum_{j=1}^b \beta_j = \sum_{i=1}^a \alpha\beta_{ij} = \sum_{j=1}^b \alpha\beta_{ij} = 0$$

$$N=abn$$

⁵⁰ William C. Guenther, Analysis of Variance, University of Wyoming, Prentice-Hall, Englewood Cliffs, s.156

Kovaryans analizinde, faktöriyel tasarım için gerekli varsayımlar,

- a) $N=abn$ birimli anakütleden tesadüfi olarak seçilmiştir.
- b) $N=abn$ birimli anakütle normal dağılımlıdır.
- c) Her bir $N=abn$ birimli anakütlenin varyansı aynıdır.
- d) Anakütle ortalamaları eşit ve doğrusaldır.
- e) Her bir $r=ab$ doğrularının eğimi aynıdır.⁵¹

Tam rastlantısal tasarım için tesadüfi bloklar için geçerli olan çözüm yöntemi faktöriyel analiz içinde geçerlidir.

Faktöriyel tasarım için gerekli hesaplamalar,

$$SS_{XT} = SS_{XA} + SS_{XB} + SS_{XAB} + SS_{XE}$$

$$SS_{YT} = SS_{YA} + SS_{YB} + SS_{YAB} + SS_{YE}$$

$$SP_T = SP_A + SP_B + SP_{AB} + SP_E$$

$$SP_T = \sum_i \sum_j \sum_k (x_{ijk} - \bar{x} \dots) (y_{ijk} - \bar{y} \dots)$$

$$SP_A = \sum_i \sum_j \bar{\sum}_k (x_{i..} - \bar{x} \dots) (y_{i..} - \bar{y} \dots)$$

⁵¹ William C.Guenthet, a.g.e., s.156

$$SP_B = \sum_i \sum_j \sum_k^n (\bar{x}_{.j} - \bar{x}_{...}) (\bar{y}_{.j} - \bar{y}_{...})$$

$$SP_{AB} = \sum_i \sum_j \sum_k^n (\bar{x}_{ij} - \bar{x}_{i..} - \bar{x}_{.j} + \bar{x}_{...}) (\bar{y}_{ij} - \bar{y}_{i..} - \bar{y}_{.j} + \bar{y}_{...})$$

$$SS_{Tr} = \sum_i \sum_j \sum_k^n (\bar{x}_{ij} - \bar{x}_{...}) (\bar{y}_{ij} - \bar{y}_{...})$$

$$SS_{Tr} = \sum_i \sum_j \sum_k^n X_{ijk} Y_{ijk}$$

$$X_{ijk} = (\bar{x}_{i..} - \bar{x}_{...}) + (\bar{x}_{.j} - \bar{x}_{...}) + (\bar{x}_{ij} - \bar{x}_{i..} - \bar{x}_{.j} + \bar{x}_{...})$$

$$Y_{ijk} = (\bar{y}_{i..} - \bar{y}_{...}) + (\bar{y}_{.j} - \bar{y}_{...}) + (\bar{y}_{ij} - \bar{y}_{i..} - \bar{y}_{.j} + \bar{y}_{...})$$

$$SP_T = \sum_i \sum_j \sum_k^n X_{ijk} Y_{ijk} - T_{x...} T_{y...}/abn$$

$$SP_A = \sum_i \sum_j \sum_k^n T_{x_{i..}} T_{y_{i..}}/bn - T_{x...} T_{y...}/abn$$

$$SP_B = \sum_i \sum_j \sum_k^n T_{x_{.j}} T_{y_{.j}}/an - T_{x...} T_{y...}/abn$$

$$SP_{AB} = \sum_i \sum_j \sum_k \bar{T}_{x_{ij.}} T_{y_{ij.}} / n - T_{x...} T_{y...} / abn$$

$$SP_E = SP_T - SP_{Tr}$$

$$SP_E = SP_T - SP_A - SP_B$$

$$SS'_{YE} = SS_{YE} - (SP_E)^2 / SS_{XE}$$

$$SS'_{Y(A+E)} = SS_{YA} + SS_{YE} - (SP_A + SP_E)^2 / (SS_{XA} + SS_{XE})$$

$$SS'_A = SS'_{Y(A+E)} - SS'_{YE}$$

$$SS'_{Y(B+E)} = SS_{YB} + SS_{YE} - (SP_B + SP_E)^2 / (SS_{XB} + SS_{XE})$$

$$SS'_B = SS'_{Y(B+E)} - SS'_{YE}$$

$$SS'_{Y(AB+E)} = SS_{YAB} + SS_{YE} - (SP_{AB} + SP_E)^2 / (SS_{XAB} + SS_{XE})$$

$$SS'_{AB} = SS'_{Y(AB+E)} - SS'_{YE}$$

Yapılan hesaplamalar sonucunda faktöriyel tasarım için oluşturulacak Kovaryans analizi tablosu aşağıdaki gibidir.

Faktöriyel Tasarım için KOVARYANS ANALİZİ

Kaynak	Kareler ve Çarpımlar Toplamı			Düzeltilmiş Kareler	Serbestlik Derecesi	Düzeltilmiş Kareler Ortalaması	F Oranı
	XX	XY	YY	Toplamı		MS' _Y	
	SS _{XA}	SP _A	SS _{YA}	SS' _A		MS' _{YA}	
A Etkisi	SS _{XA}	SP _A	SS _{YA}	SS' _A	a-1	MS' _{YA}	MS' _{YA} /MS' _{YE}
B Etkisi	SS _{XB}	SP _B	SS _{YB}	SS' _{YB}	b-1	MS' _{YB}	MS' _{YB} /MS' _{YE}
AB Etkisi	SS _{XAB}	SP _{AB}	SS _{YAB}	SS' _{YAB}	(a-1)(b-1)	MS' _{YAB}	MS' _{AB} /MS' _{YE}
Hata	SS _{XE}	SP _E	SS _{YE}	SS' _{YE}	ab(n-1)-1	MS' _{YE}	MS' _{YE}
Toplam	SS _{XT}	SP _T	SS _{YT}	SS' _{YT}	abn-2		

$$SS_{XA} = bn \sum (\bar{X}_{.i} - \bar{X})^2$$

$$SS_{YA} = bn \sum (\bar{Y}_{.j} - \bar{Y})^2$$

$$SS_{PA} = bn \sum (\bar{X}_{.i} - \bar{X})(\bar{Y}_{.j} - \bar{Y})$$

$$SS_{XB} = an \sum (\bar{X}_{.i} - \bar{X})^2$$

$$SS_{YB} = an \sum (\bar{Y}_{.j} - \bar{Y})^2$$

$$SS_{PB} = an \sum (\bar{X}_{.i} - \bar{X})(\bar{Y}_{.j} - \bar{Y})$$

$$SS_{XE} = \sum \sum \sum (\bar{X}_{ijk} - \bar{X}_{ij})^2$$

$$SS_{YE} = \sum \sum \sum (\bar{Y}_{ijk} - \bar{Y}_{ij})^2$$

$$SS_{PE} = \sum \sum \sum (\bar{X}_{ijk} - \bar{X}_{ij})(\bar{Y}_{.j} - \bar{Y}_{ij})$$

$$SS_{XAB} = n \sum (\bar{X}_{ij} - \bar{X}_{.i} - \bar{X}_{.j} + \bar{X})^2$$

$$SS_{YAB} = n \sum (\bar{Y}_{ij} - \bar{Y}_{.i} - \bar{Y}_{.j} + \bar{Y})^2$$

$$SS_{PAB} = n \sum (\bar{X}_{ij} - \bar{X}_{.i} - \bar{X}_{.j} + \bar{X})(\bar{Y}_{ij} - \bar{Y}_{.i} - \bar{Y}_{.j} + \bar{Y})$$

$$MS'_{YE} = SS_{YE} - SP_E^2 / SS_{XE}$$

$$MS'_{YA} = (SS_{YA} + SS_{YE}) - (SP_A + SP_E)^2 / (SS_{XA} + SS_{XE}) - MS'_{YE}$$

$$MS'_{YB} = (SS_{YB} + SS_{YE}) - (SP_B + SP_E)^2 / (SS_{XB} + SS_{XE}) - MS'_{YE}$$

$$MS'_{YAB} = (SS_{YAB} + SS_{YE}) - (SP_{AB} + SP_E)^2 / (SS_{XAB} + SS_{XE}) - MS'_{YE}$$

2.2.3.3. *Hipotez Testi*

İki faktörlü kovaryans analizi, deneysel ve gözlemsel araştırmalara eşit ölçüde uygulanabilir. Ancak her hücre için (iki kategorik değişkenin seviye kombinasyonları) eşit sayıda gözlemin meydana gelmesi, gözlemsel araştırma sayısı sınırlıdır. Gözlemsel araştırmalarda eşit olmayan hücre büyüklüğü eğilimi vardır ve buna göre analiz edilmeleri gerekir.

Tamamen rastlantısal faktörel tasarımda ağırlık kategorik bağımsız değişkenlerin (faktörlerin) çeşitli seviyelerindeki farklılıkların ve bunların karşılıklı etkileşimlerinin incelenmesine olanak sağlar. Daha önce ele alınan tasarımlarda olduğu gibi, faktörlere ait hipotezler için, düzeltilmiş ortalama karelerinin F oranları kullanılarak test edilecektir.

Kovaryans modelinde hücre içi regresyon katsayısının A ve B faktörlerinin bütün kombinasyonları için eşit olduğu varsayılır. Regresyonun homojenliği hipotezi, tamamen rastlantısal tasarımdaki gibi olup,

$$F = \frac{SP_E^2}{(N-ab-1)} / (SS_{YE} - SP_E^2)$$

şeklindeki F testi ile test edilir.⁵² Serbeslik derecesi 1,(N-2ab-1) olan F dağılımına sahiptir.

⁵² Albert R. Wildth and Olli Ahtola, Analysis of Covariance, Universty George and Universty of Florida, 1985, s.78

Faktörel tasarım için test edilecek hipotezler aşağıdaki gibidir.

$$H_0: \beta_j=0 \quad (j=1,2,\dots,b)$$

$H_1: \beta_j$ 'ler 0'dan farklıdır.

$$H_0: \alpha_i=0 \quad (i=1,2,\dots,a)$$

$H_1: \alpha_j$ 'ler 0'dan farklıdır.

$$H_0: \alpha\beta_{ij}=0 \quad (i=1,2,\dots,a; j=1,2,\dots,b)$$

$H_1: \alpha\beta_{ij}$ 'ler 0'dan farklıdır.

$$\alpha_i = (\mu_{i..} - \mu) - \gamma (\bar{x}_{i..} - \bar{x}_{...})$$

$$\beta_j = (\mu_{.j.} - \mu) - \gamma (\bar{x}_{.j.} - \bar{x}_{...})$$

$$\beta_{ij} = (\mu_{ij} - \mu_{i..} - \mu_{.j.} + \mu) - \gamma (\bar{x}_{ij} - \bar{x}_{i..} - \bar{x}_{.j.} + \bar{x}_{...})$$

$$H_0: \mu_{1..} - \gamma \bar{x}_{1..} = \mu_{2..} - \gamma \bar{x}_{2..} = \dots = \mu_{a..} - \gamma \bar{x}_{a..}$$

$H_1: \mu_{i..} - \gamma \bar{x}_{i..}$ 'ler birbirinden farklıdır.

$$H_0: \mu_{.1.} - \gamma \bar{x}_{.1.} = \mu_{.2.} - \gamma \bar{x}_{.2.} = \dots = \mu_{.b.} - \gamma \bar{x}_{.b.}$$

$H_1: \mu_{.j.} - \gamma \bar{x}_{.j.}$ 'ler birbirinden farklıdır.

ÜÇÜNCÜ BÖLÜM

UYGULAMA

3.1. UYGULAMANIN AMACI

Kovaryans analizi uygulaması tek etkenli ve tam rastlantısal tasarım için düzenlenmiştir. Öğrencilerin mezuniyet dereceleri bağımlı değişken ve mezuniyet derecelerini etkileyen faktörler; dersaneye gitmek, kolej mezunu olmak, ailelerin gelir durumları (ortalama), özel ders almak gibi faktörler de bağımsız değişkenler olarak düşünülerek anket yapılmıştır.

Yapılan anket Marmara Üniversitesi Bahçelievler Kampüsündeki öğrencilere uygulanmıştır. Ekonometri, Maliye, İşletme, İktisat bölümlerinden 10'ar öğrenci tesadüfi yöntemle seçilerek anket yapılmıştır.

Uygulamanın amacı bu öğrencilerin mezuniyet derecelerine göre bölümler arasında fark var mıdır, fark varsa bunu etkileyen faktörler nelerdir, etkileyeceğini düşündüğümüz faktörlerin etkisi gerçekte var mıdır sorularına cevap bulmaktır. Bu çalışma için kullanılacak yöntem çok değişkenli istatistik tekniklerinden biri olan kovaryans analizidir. Daha önce de belirtildiği gibi kovaryans analizi ortak değişken değerlerinin etkisini kaldırmak amacıyla uygulanan istatistiksel bir analizdir.

3.2. UYGULAMANIN VERİLERİ VE SONUÇLARI

İlk aşamada mezuniyet derecesine göre bölümler arasında fark var mıdır? sorusunu cevaplamak gerekir. Bunun için varyans analizinin uygulanması gerekir. Varyans analizine göre sıfır hipotezi; bölümler arasında fark yoktur, karşıt hipotez ise bölümler arasında fark vardır şeklinde olacaktır.

Ankete göre mezuniyet derecelerine ait veriler aşağıdaki gibidir.

BÖLÜMLER

	EKONOMETRİ	MALİYE	İŞLETME	İKTİSAT
	Y	Y	Y	Y
1	6	9.08	7.68	6.23
2	5.56	5.78	6.5	7
3	6	5.08	5.7	7.81
4	5.19	9.36	6.3	6.4
5	7	6.74	7	8.55
6	6.4	6	9	8.94
7	6.55	6.35	9.8	8.90
8	6.4	6	8	8.25
9	6	6.4	6	7.5
10	5.6	6	5.33	6
TOPLAM	60.07	66.79	71.33	75.65

$$\Sigma Y = 274.45$$

Varyans analizi için gerekli hesaplamalar,

$$\begin{aligned} \text{Genel KT} &= 6^2 + 5.56^2 + \dots + 7.5^2 + 6^2 - 274.45^2 / 40 \\ &= 1946.29 - 1883.07 = 63.22 \end{aligned}$$

$$\begin{aligned} \text{Gruplar Arası KT} &= 60.70^2 / 10 + \dots + 75.65^2 / 10 - 274.45^2 / 40 \\ &= 1895.34 - 1883.07 = 12.273 \end{aligned}$$

$$\begin{aligned} \text{Hata KT} &= \text{Genel KT} - \text{Gruplar Arası KT} \\ &= 63.22 - 12.273 = 50.947 \end{aligned}$$

Varyans Analizi Tablosu

Kaynak	Serbestlik Derecesi	Kareler Toplamı	Kareler Ortalaması	F ORANI
Gruplar Arası	4-1=3	12.273	4.0914	2.88
Gruplar İçi (Hata)	40-4=36	50.947	1.415	
Genel	40-1=39	63.22		

Serbestlik derecesi 3, 36 ve $\alpha=0.05$ hata payı ile F tablo değeri 2.86 dır. Hesaplanan F değeri ile tablo değeri karşılaştırıldığında hesap değeri tablo değerinden büyük olduğundan bölümler arasında bir fark yoktur hipotezi reddedilir.

Tablo değeri ile hesaplanan değer arasındaki kabul edilebilirlik farkının oldukça küçük olmasına rağmen sıfır hipotezi reddedilmiştir. Çünkü seçilen örnek birim sayısı küçük olduğundan güven derecesini düşürdüğümüzde tablo değeri küçülecek ve hipotez red edilecektir. Yani mezuniyet derecelerine göre bölümler arasında fark vardır.

Varyans analizi sonucuna göre bölümler arasında fark varsa, bu fark hangi sebeplerden kaynaklanır? Yapılan ankete göre ailenin gelir durumunun, öğrencinin dersaneye gitme süresinin (yıl olarak), özel ders almasının, kolej mezunu olup olmamasının mezuniyet derecelerine etkilerinin olabileceği düşünülmüştür.

Yapılan hesaplamalarla mezuniyet derecesiyle bağımsız değişkenler arasında bir ilişkinin var olduğu söylenebilmektedir ancak ilişkinin derecesi oldukça düşüktür. Bütün bağımsız değişkenler ele alındığında ilişki katsayısı daha yüksek bir değer çıkmaktadır.

Kovaryans analizi için yapılan hesaplamalar bilgisayarla yapılmadığından ortak değişkenlerin etkileri çok etkenli olmasına rağmen her etken tek faktörlü olarak test edilecektir.

Kovaryans analizinin, varyans analizi ile regresyon analizinin bir bileşimi olduğunu daha önce belirtmiştik. Varyans analizi bölümler arasındaki farkı bulmamız için kullanılmıştır. Mezuniyet derecesi olarak tanımladığımız bağımlı Y değişkeni ve Y değişkenini etkileyen ailenin gelir durumu bağımsız X değişkeni olarak tanımlanırsa, X ile Y arasındaki ilişkinin ölçüsü de regresyon analizi ile hesaplanacaktır.

Sonuçta ailenin gelir durumunun mezuniyet derecesine etkisi varsa, var olan bu etki kovaryans analizi ile test edilmiş olacaktır.

Ailenin gelir durumlarına ait veriler aşağıdaki gibidir.

Gözlenen değerler;

BÖLÜMLER

	EKONOMETRİ		MALİYE		İŞLETME		İKTİSAT	
	Y	X	Y	X	Y	X	Y	X
1	6	35	9.08	30	7.68	60	6.23	80
2	5.56	100	5.78	25	6.5	120	7	80
3	6	40	5.08	100	5.7	150	7.88	40
4	5.19	80	9.36	27	6.3	50	6.4	55
5	7	100	6.74	60	7	50	8.55	50
6	6.4	1000	6	80	9	27	8.94	100
7	6.55	40	6.35	100	9.8	41	8.90	80
8	6.4	80	6	60	8	80	8.25	80
9	6.	50	6.4	30	6	200	7.5	100
10	5.6	70	6	38	5.33	120	6	80
TOP.	60.07	1595	66.79	550	71.33	898	75.65	745

$$\Sigma Y = 274.45$$

$$\Sigma X = 3838$$

Genel Kareler Toplamları;

$$T_{YY} = 6^2 + 5.56^2 + \dots + 7.5^2 + 6^2 - 274.45^2 / 40 = 63.22$$

$$T_{XX} = 35^2 + 100^2 + \dots + 100^2 + 80^2 - 3838^2 / 40 = 887392$$

$$T_{XY} = 6 \times 35 + 5.56 \times 100 + \dots + 7.5 \times 100 + 6 \times 80 - 274.45 \times 3838 / 40 \\ = -1045.96$$

Gruplar Arası Kareler Toplamları;

$$B_{YY} = 60.70^2 / 10 + \dots + 75.65^2 / 10 - 274.45^2 / 40 = 12.27$$

$$B_{XX} = 1545^2 / 10 + \dots + 745^2 / 10 - 3838^2 / 40 = 58289$$

$$B_{XY} = 6.70 \times 1545 / 10 + \dots + 75.65 \times 745 / 10 - 274.45 \times 3838 / 40 \\ = -604.87$$

Hata Kareler Toplamları;

$$E_{YY} = T_{YY} - B_{YY} = 63.22 - 12.27 = 50.95$$

$$E_{XX} = T_{XX} - B_{XX} = 887.392 - 58289 = 828103$$

$$E_{XY} = T_{XY} - B_{XY} = -1045.96 - (-604.87) = -441.09$$

Düzeltilmiş Y değerleri;

$$\text{Genel KT} \Rightarrow T_{YY} - T_{XY}^2 / T_{XX} \\ \Rightarrow 63.22 - (-1045.96)^2 / 887392 \\ \Rightarrow 61.98$$

$$\text{Gruplar Arası KT} \Rightarrow B_{YY} - B_{XY}^2 / B_{XX} \\ \Rightarrow 12.27 - (-604.87)^2 / 58289 \\ \Rightarrow 5.99$$

$$\begin{aligned}
\text{Hata} &\Rightarrow E_{YY} - E_{XY}^2 / E_{XX} \\
&\Rightarrow 50.95 - (-441.09)^2 / 829103 \\
&\Rightarrow 50.71
\end{aligned}$$

Kovaryans Analizi Tablosu

Kaynak	KARELER TOPLAMLARI			Düzeltilmi ş Y	sd	KARELER ORTALAMASI	
	sd	XX	XY				YY
Gruplar Arası	3	58289	-604.87	12.27	5.99	2	2.995
Hata	36	829103	-441.09	50.95	50.71	35	1.44
Genel	39	887392	-1.045.96	63.22	61.98	37	

$$F_{2,35} = 2.995 / 1.448 = 149.12 = 2.079$$

Serbestlik derecesi 2, 35 ve $\alpha=0,05$ hata ile tablo değeri 3.23 ile karşılaştırıldığında hata kareler toplamındaki azalma önemli değildir hipotezi KABUL EDİLİR. Yani gelir durumunun etkisinin hata kareler toplamındaki azalmada etkisi önemsizdir.

$$\begin{aligned}
b &= \Sigma XY / \Sigma X^2 = E_{XY} / E_{XX} \\
&= -441.09 / 829103 = -0.00053
\end{aligned}$$

düzeltilmiş Y değerleri;

$$\begin{aligned}\bar{Y}_{.1} &= 60.70 / 10 - 0.0005x(1595 / 10 - 3838 / 40) \\ &= 6.070 - 0.0005 x (-63.55) = 6.070 - 0.03 = 6.03\end{aligned}$$

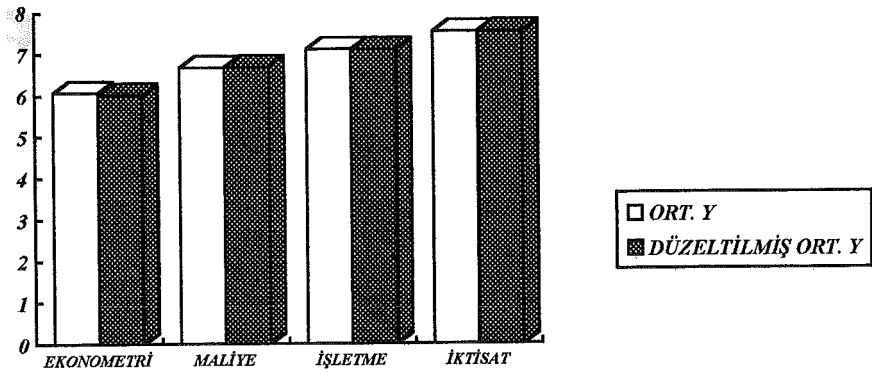
$$\begin{aligned}\bar{Y}_{.2} &= 66.79 / 10 - 0.0005x(600 / 10 - 3838 / 40) \\ &= 6.679 - 0.0005 x (-35.95) = 6.679 + 0.02 = 6.696\end{aligned}$$

$$\begin{aligned}\bar{Y}_{.3} &= 71.31 / 10 - 0.0005x(898 / 10 - 3838 / 40) \\ &= 7.131 - 0.0005 x (-6.15) = 7.131 + 0.003 = 7.134\end{aligned}$$

$$\begin{aligned}\bar{Y}_{.4} &= 75.65 / 10 - 0.0005x(745 / 10 - 3838 / 40) \\ &= 7.565 - 0.0005 x (-21.45) = 7.565 + 0.011 = 7.575\end{aligned}$$

Sonuç olarak ailenin gelir durumunun mezuniyet derecesine etkisi önemsizdir. Düzeltilmiş Y değerleri ile ortalama Y değerleri arasındaki fark oldukça küçüktür. Bir grafik ile gösterirsek,

	\bar{Y}	\bar{Y} DÜZELTİLMİŞ
EKONOMETRİ	6.070	6.030
MALİYE	6.679	6.696
İŞLETME	7.131	7.134
İKTİSAT	7.565	7.575



Anket sonucunda elde edilen ve diğer etken olarak düşünülen dersaneye gitme faktörünün mezuniyet derecesine etkisinin incelenmesi için bu etkinin bölümler arasındaki farkı açıklayacağı hipotezi test edilecektir.

Anket sonucunda öğrencilerin dersaneye kaç yıl gittiklerine dair veriler aşağıdaki gibidir.

Gözlemlenen değerler;

BÖLÜMLER

	EKONOMETRİ		MALİYE		İŞLETME		İKTİSAT	
	Y	X	Y	X	Y	X	Y	X
1	6	2	9.08	2	7.68	2	6.23	2
2	5.56	1	5.78	3	6.5	2	7	2
3	6	1	5.08	1	5.7	1	7.88	1
4	5.19	1	9.36	0	6.3	1	6.4	1
5	7	1	6.74	2	27	1	8.55	1
6	6.4	2	6	2.5	9.0	0	8.94	1
7	6.55	1	6.35	0	9.8	0	8.90	1
8	6.4	1	6	0	8	1	8.25	1
9	6.	1	6.4	2	6	1	7.5	1
10	5.6	1	6	2	5.33	2	6	1
TOP.	60.07	12	66.79	14.5	71.33	11	75.65	12

$$\Sigma Y = 274.45$$

$$\Sigma X = 49.5$$

Genel Kareler Toplamları;

$$T_{YY} = 6^2 + 5.56^2 + \dots + 7.5^2 + 6^2 - 274.45^2 / 40 = 63.22$$

$$T_{XX} = 2^2 + 1^2 + \dots + 1^2 + 1^2 - 49.5^2 / 40 = 20.03$$

$$T_{XY} = 6 \times 1 + 5.56 \times 1 + \dots + 7.5 \times 1 + 6 \times 1 - 274.45 \times 49.5 / 40 \\ = -11.77$$

Gruplar Arası Kareler Toplamları;

$$B_{YY} = 60.70^2 / 10 + \dots + 75.65^2 / 10 - 274.45^2 / 40 = 12.27$$

$$B_{XX} = 12^2 / 10 + \dots + 12^2 / 10 - 49.5^2 / 40 = 0.67$$

$$B_{XY} = 6.70 \times 12 / 10 + \dots + 75.65 \times 12 / 10 - 274.45 \times 49.5 / 40 \\ = -0.723$$

Hata Kareler Toplamları;

$$E_{YY} = T_{YY} - B_{YY} = 63.22 - 12.27 = 50.95$$

$$E_{XX} = T_{XX} - B_{XX} = 20.03 - 0.67 = 19.36$$

$$E_{XY} = T_{XY} - B_{XY} = -11.77 - (-0.723) = 11.047$$

Düzeltilmiş Y değerleri;

$$\text{Genel KT} \Rightarrow T_{YY} - T_{XY}^2 / T_{XX}$$

$$\Rightarrow 63.22 - (-11.77)^2 / 20.03$$

$$\Rightarrow 70.136$$

$$\begin{aligned} \text{Gruplar Arası KT} &\Rightarrow B_{YY} - B_{XY}^2 / B_{XX} \\ &\Rightarrow 12.27 - (-0.723)^2 / 0.67 \\ &\Rightarrow 11.49 \end{aligned}$$

$$\begin{aligned} \text{Hata} &\Rightarrow E_{YY} - E_{XY}^2 / E_{XX} \\ &\Rightarrow 50.95 - (-11.047)^2 / 19.36 \\ &\Rightarrow 44.64 \end{aligned}$$

Kovaryans Analizi Tablosu

Kaynak	KARELER TOPLAMLARI				Düzeltilmiş Y	sd	KARELER ORTALAMASI
	sd	XX	XY	YY			
Gruplar Arası	3	0.67	-0.723	12.27	11.49	2	5.74
Hata	36	19.36	-11.047	50.95	44.64	35	1.275
Genel	39	20.03	-11.77	63.22	70.136	37	

$$F_{2,35} = 5.74 / 1.275 = 4.501$$

Serbestlik derecesi 2,35 olan ve $\alpha=0,05$ için tablo değeri 3.23 ile karşılaştırıldığında hata kareler toplamındaki azalma

önemli değildir hipotezi REDDEDİLİR. Yani hata kareler toplamındaki azalma önemlidir. Dersaneye gitmenin mezuniyet derecesine etkisi vardır ve bu etkinin hata kareler toplamındaki azalmada etkilidir.

$$b = \frac{\sum XY}{\sum X^2} = \frac{E_{XY}}{E_{XX}} = -11.047 / 19.36 = -0.57$$

düzeltilmiş ortalama Y değerleri;

$$\begin{aligned}\bar{Y}_{.1} &= 60.70 / 10 - (-0.57)x(12/10 - 49.5 / 40) \\ &= 6.070 - 0.0213 \\ &= 6.0487\end{aligned}$$

$$\begin{aligned}\bar{Y}_{.2} &= 66.79 / 10 - (-0.57)x(14.5 / 10 - 49.5 / 40) \\ &= 6.679 - 0.121 \\ &= 6.557\end{aligned}$$

$$\begin{aligned}\bar{Y}_{.3} &= 71.31 / 10 - (-0.57)x(11 / 10 - 49.5 / 40) \\ &= 7.131 - 0.0783 \\ &= 7.05\end{aligned}$$

$$\begin{aligned}\bar{Y}_{.4} &= 75.65 / 10 - (-0.57)x(12 / 10 - 49.5 / 40) \\ &= 7.565 - 0.0213 \\ &= 7.54\end{aligned}$$

Dersane faktörünün etkisinin önemli olduğu test edildikten sonraki aşamada oluşturulacak hipotez, eğimin 0

olmadığı, regresyonun doğrusal olduğu hipotezidir ki, bu faktör kovaryans analizi için gereklidir.

$$H_0 : \beta=0$$

$$H_1 : \beta \neq 0$$

$$F_{1,N-k-1} = (E_{XY} / E_{XX}) / (\text{düzeltilmiş } E_{YY} / N-k-1)$$

$$F_{1,N-k-1} = (-11.0472 / (-0.723)) / (44.69 / 36)$$

$$F_{1,N-k-1} = 136.121$$

Serbestlik derecesi 1, 35 olan $\alpha=0,05$ hata ile tablo değeri 4.08 dir. Bu durumda hipotez reddedilir. Kovaryans analizi için gerekli bir koşul kabul edilir.

β katsayısı 0 dan farklıdır, regresyon doğrusaldır. Bu hipotez de doğrulandıktan sonra, bölümler için regresyon katsayılarının eşit olduğu hipotezini test etmek gerekir. Bu durumda herbir bölüm için regresyon katsayılarını hesaplamak ve her bir bölüm için kareler toplamlarını kendi regresyon katsayıları ile düzeltmek gerekir.

$$b = \Sigma XY / \Sigma X^2 = (\Sigma XY - \Sigma X \Sigma Y / n) / (\Sigma X^2 - (\Sigma X)^2 / n)$$

$$\text{düzeltilmiş } \Sigma Y^2 = \Sigma Y^2 - (\Sigma Y)^2 / n$$

Ekonometri için;

$$b = (73.10 - 12 \times 60.70) / (16 - 122 / 10) = 0.26 / 1.6 = 0.1625$$

$$\begin{aligned} \text{düzeltilmiş } \Sigma Y^2 &= (371.03 - 60.702/10) - (0.26)^2 / 1.6 \\ &= 2.53 \end{aligned}$$

Maliye için;

$$\begin{aligned} b &= (93.86 - 14.50 \times 66.79) / (32.25 - 14.52 / 10) \\ &= -2.9855 / 11.225 = -0.2659 \end{aligned}$$

$$\begin{aligned} \text{düzeltilmiş } \Sigma Y^2 &= (463.98 - 66.792/10) - (-2.9855)^2 / 11.225 \\ &= 17.089 \end{aligned}$$

İşletme için;

$$b = (72.02 - 1 \times 71.31) / (17 - 112 / 10) = -6.421 / 4.9 = -1.31$$

$$\begin{aligned} \text{düzeltilmiş } \Sigma Y^2 &= (527.86 - 71.312/10) - (-6.421)^2 / 11.225 \\ &= 10.934 \end{aligned}$$

İktisat için;

$$b = (88.88 - 12 \times 75.65) / (16 - 122 / 10) = -1.9 / 1.6 = -1.1875$$

$$\begin{aligned} \text{düzeltilmiş } \Sigma Y^2 &= (583.42 - 75.652/10) - (-1.9)^2 / 1.6 \\ &= 8.8715 \end{aligned}$$

Herbir Grup İçi Düzeltilmiş Kareler Toplamı

	sd	ΣY^2	b	düzeltilmiş ΣY^2	sd
EKONOMETRİ	9	2.581	0.1625	2.53	8
MALİYE	9	17.889	-0.2659	17.089	8
İŞLETME	9	19.348	-1.3100	10.934	8
İKTİSAT	9	11.127	-1.1875	8.8715	8
TOPLAM	36	50.945		39.4245	32

$$F_{k-1, N-2k} = \frac{[\text{düzeltilmiş } \Sigma Y^2 \text{ (Hata Kareler Ortalaması)} - \text{düzeltilmiş } \Sigma Y^2 \text{ (Her bir bölüm için Hata Kareler Ortalaması)}] / k-1}{\text{düzeltilmiş } \Sigma Y^2 \text{ (Her bir bölüm için Hata Kareler Ortalaması)} / N-2k}$$

$$F_{8, 32} = \frac{(44.64 - 39.4245)/8}{39.425/32}$$

$$F_{8, 32} = 0.65/1.232 = 0.527$$

$$F_{\text{tablo}} > F_{\text{hesap}}$$

Serbestlik derecesi 8, 32 olan ve $\alpha=0,05$ hata ile tablo değeri 2.33 olduğundan sıfır hipotezi kabul edilir. Yani her bir bölüm için regresyon katsayıları homojendir.

Enson aşamada oluşturulacak hipotez dersaneye gitme faktörünün bölümlerden etkilenmediği yani bağımsız değişken olan dersaneye gitme faktörünün bölümlere uygulanan denemelerden etkilenmediği şeklindeki hipotezi test etmek olacaktır.

$$F_{3, 36} = (0.67 / 3) / 19.36/36 = 0.41$$

Elde edilen değer tablo değerinden küçük olduğundan sıfır hipotezi kabul edilir. Yani öğrencinin okuduğu bölümün dersaneye gitme faktörünü etkilemediği söylenir.

3.3. UYGULAMANIN YORUMLANMASI

Kovaryans analizi için yapılan anket sonucunda bağımlı değişken olarak düşünülen öğrencilerin mezuniyet derecelerinin, bölümler (Ekonometri, Maliye, İktisat, İşletme) arasında bir fark yarattığı ortaya çıkmıştır. Bu farkı açıklamak için ailelerin gelir durumlarının mezuniyet derecelerine bir etkisinin olduğu düşünülmüş ancak hesaplamalar sonucunda, ailelerin gelir durumlarının mezuniyet dereceleri ile ilişkisinin olmadığı, bölümler arasındaki farkı düzeltmek için yetersiz olduğu sonucu ortaya çıkmıştır.

Etkisi olacağını düşündüğümüz diğer faktörlerden dersaneye gitme faktörünün, öğrencilerin mezuniyet dereceleri üzerinde etkili olup olmadığı incelendiğinde ise dersaneye gitme faktörünün, mezuniyet dereceleri üzerinde etkili olduğu ve bu

etkinin bölümler arasındaki fark için önemli olduğu sonucu ortaya çıkmıştır.

Kovaryans analizi için gerekli olan $\beta \neq 0$ ve regresyon doğrusaldır koşulu sağlanmış, her bir bölüm için regresyon katsayısının homojen olduğu test edilmiştir.

Son aşama olarak nitelendirdiğimiz bölümlerin dersaneye gitme faktörü üzerinde etkilerinin olmadığı sonucu da anlamlı çıkmıştır.

Mezuniyet derecelerine göre bölümler arasında bir fark olduğu görülmüş ve öğrencilerin dersaneye gitme süreleri ile mezuniyet dereceleri arasında bir ilişki belirlenmiştir. Bölümler arasındaki farkı, dersaneye gitme faktörü açıklamaktadır ve bu faktörün etkisi önemlidir. Bu etki düzeltildiğinde Kovaryans analizi ile bölümler arasındaki homojenlik sağlanmış olur.

SONUÇ:

Çok Değişkenli İstatistiksel Analizlerde değişken sayısı birden fazladır ve değişkenler başka değişkenlerce açıklanmaktadır. Değişkenler arasındaki bağımlılık analiz edilerek değişkenlerin temel yapılarını anlamak, hangi değişkenin hangi grup içerisinde yer alacağına karar vermek çok değişkenli analiz uygulamalarının gerekliliğini ortaya koyar.

Çok Değişkenli İstatistiksel Analiz bağımlılık ve karşılıklı bağımlılık analizleri olarak iki grupta incelenmektedir. Bağımlılık analizinde, değişkenler birbiriyle ilişkili olup başka değişkenlerce açıklanmaktadır. Karşılıklı bağımlılık analizinde, bütün değişkenlerin arasındaki bağımlılık yerine karşılıklı bağımlılıklar sözkonusudur.

Bağımlı değişken değişirken, bağımsız değişken değerlerini kontrol altında tutmak mümkün değildir. **Kovaryans Analizi**, bağımlı değişken üzerindeki, bağımsız değişkenin etkilerinin giderildiği çok değişkenli istatistiksel bir analizdir. Bağımlı değişkeni etkileyen birden çok bağımsız değişken olabilir.

Kovaryans analizi için, Marmara Üniversitesi, İktisadi ve İdari Bilim Fakültesinin Ekonometri, Maliye, İşletme ve İktisat bölümlerinden onar öğrenci teadüfen seçilerek anket yapılmıştır. Ankette öğrencilere mezuniyet dereceleri ve mezuniyet dereceleriyle etkili olduğunu düşündüğümüz ailelerin gelir durumları ve dersaneye gitme durumlarına ait bilgiler alınmıştır.

Mezuniyet derecelerine göre bölümler arasında bir fark olduğu, varyans analizi ile test edilmiştir. Bölümler arasındaki bu fark ile ailelerin gelir durumlarının bir ilişkisinin olduğu düşünülmüş, ancak kovaryans analizi sonucunda mezuniyet dereceleri ile ailelerin gelir durumları arasındaki ilişkinin bölümler arasındaki fark için anlamlı olmadığı sonucu ortaya çıkmıştır.

Diğer bir etken olarak düşündüğümüz dersaneye gitme faktörünün ise mezuniyet dereceleri ile bir ilişkisinin olduğu ortaya çıkmış ve Kovaryans analizi ile bu faktörün etkisi giderildiğinde, Ekonometri, Maliye, İktisat ve İşletme bölümleri arasında mezuniyet derecelerine bir fark olmadığı belirlenmiştir.

KAYNAKLAR

Kitaplar

- A.Aker David, (1971). Multivariate Analysis in Marketing Theory & Application, Wadsworth Publishing Company, California
- A.Johnson Richard, (1985). Dean W. Wichern, Applied Multivariate Statistical Analysis, University of Wisconsin-Madison
- A. Koutsoyiannis, (1989). Ekonometri Kuramı-Ekonometri Yöntemlerinin Tanıtımına Giriş, çev. Ü. Şenesen, G. Şenesen, 1. Baskı, Ankara
- Chatfield Christopher, (1980) Alexander J. Collins Introduction to Multivariate Analysis, Chapman and Hall
- C.Guenther William, Analysis of Variance, Prentice-Hall Inc., Englewood Cliffs, N.J.
- Çömlekçi Necla, (1989). Temel İstatistik İlke ve Teknikleri , Bilim Teknik Yayınevi, Eskişehir
- Düzgüneş Orhan, (1987). Tahsin Kesicic, Orhan Kavuncu, Fikret Gürbüz, Araştırma ve Deneme Metodları, Ankara Üniversitesi Ziraat Fakültesi Yayınları, Ankara
- Gatty Ronald, Multivariate Analysis for Marketing Search: An Evaluation
- Genceli Mehmet,(1991). Ekonometride İstatistiksel İlkeler, Filiz Kitabevi, İstanbul,
- Giriş Selahattin, (1995). Şahamet Bülbül, Olasılık, Marmara Üniversitesi Nihad Sayar Eğitim Vakfı, İstanbul
- İşyar Yüksel, (1995). Ekonometriye Giriş, Uludağ Üniversitesi Yayınları, Bursa,
- Köksal Bilge Aloba, (1985). İstatistik Analiz Metodları, Çağlayan Kitabevi, İstanbul, 3. Baskı
- Muluk Zehra ve diğ.,(1985). Deney Düzenlemede İstatistik Yöntemler, Akademi Matbaası, Ankara, (çev.) Charles R.Hicks

- Press S.James, (1982) Applied Multivariate Analysis: Using Bayesian and Frequentist Methods of Inference, Second Edition, Robert E. Krieger Publishing Company Florida
- R.Anderberg Michael, (1973). Cluster Analysis for Applications, Academic Press,
- R.Wildt Albert, (1985). Olli Ahtola, Analysis of Covariance, University of Georgia and University of Florida
- Şahin Akkaya, (1995). Vedat Pazarlıoğlu, Ekonometri I, Anadolu Matbaacılık, İzmir
- Tatlıdil Hüseyin, (1996). Uygulamalı Çok Değişkenli İstatistiksel Analiz , Akademi matbaası, Ankara
- W.Bolch Ben, (1985). Cliff J. Huang, Multivariate Statistical Methods for Business and Economics, Vanderbilt University

Diğer Kaynaklar

- Çakır Filiz, (1994). Karşılıklı bağımlılığın Ölçülmesinde Kümeleme Analizi ve Bir Uygulama , Yüksek Lisans Tezi, M.Ü., İstanbul
- Dirican Ahmet, (1988). n Değişkenli Araştırmalarda Diskriminant Çözümlemesinin Kümeler Arası Tanı Ayrımında Kullanımı ve Bilgisayar Uygulaması, Yüksek Lisans Tezi, İ.Ü., İstanbul
- Karagöz Suat, (1991). Faktör Analizi Tekniğini Kullanarak Üniversite Öğrencilerinin Gazete Tercihinde Etkin Faktörlerin Belirlenmesi, Yüksek Lisans Tezi, İ.Ü., İstanbul
- Oktay Ümit Fırat, (1996). Marmara Üniversitesi Sosyal Bilim Öğrencilerinin Fakülte ve Bölümler Bazında Farklılıklarının Çok Değişkenli İstatistik Teknikleri ile Analizi, Doktora Tezi, M.Ü., Yayın No:573, İstanbul
- Taşkın Veyis, (1991). Lojistik Regresyon Yöntemi ile Angina Pektoris Tanısına Çeşitli Ölçütlerin Katkısının İrdelenmesi, Yüksek Lisans Tezi, İ.Ü., İstanbul
- Tekin Mustafa, (1993). Kanonik Korelasyon Analizi ve Bir Uygulama, Doktora Tezi, İ.Ü., İstanbul