**SİVAS CUMHURİYET ÜNİVERSİTESİ**

**Sosyal Bilimler Enstitüsü**

**Ekonometri Ana Bilim Dalı**

# DOĞRUSAL OLMAYAN KOVARYANS ANALİZİ MODELLERİ İLE SOSYAL GÜVENLİK VERİLERİ ÜZERİNE UYGULAMA

**Yüksek Lisans Tezi**

**Cansu MERCAN**

**Sivas**

**Temmuz  2019**

**SİVAS CUMHURİYET ÜNİVERSİTESİ**

**Sosyal Bilimler Enstitüsü**

**Ekonometri Ana Bilim Dalı**

# DOĞRUSAL OLMAYAN KOVARYANS ANALİZİ MODELLERİ İLE SOSYAL GÜVENLİK VERİLERİ ÜZERİNE UYGULAMA

**Yüksek Lisans Tezi**

**Cansu MERCAN**

**Tez Danışmanı**

**Doç. Dr. Necati Alp ERİLLİ**

**Sivas**

**Temmuz 2019**

## KABUL VE ONAY

**Üniversite** : **Sivas Cumhuriyet Üniversitesi**

**Enstitü** : **Sosyal Bilimler Enstitüsü**

**Ana Bilim Dalı** : **Ekonometri**

**Tezin Başlığı** : **Doğrusal Olmayan Kovaryans Analizi Modelleri ile Sosyal Güvenlik Verileri Üzerine Uygulama**

**Savunma Tarihi** : **12/07/2019**

**Danışmanı** : **Doç. Dr. Necati Alp ERİLLİ**

| | Unvanı - Adı Soyadı | İmza |
|---|---|---|
| **Jüri Başkanı:** | **Dr. Öğr. Üyesi Özge GÜNDOĞDU** | |
| **Üye:** | **Doç. Dr. Necati Alp ERİLLİ** | |
| **Üye:** | **Dr. Öğr. Üyesi Engin KARAKIŞ** | |

**Oy Birliği** [X]

**Oy Çokluğu** [ ]

**Cansu MERCAN tarafından hazırlanan "Doğrusal Olmayan Kovaryans Analizi Modelleri ile Sosyal Güvenlik Verileri Üzerine Uygulama" başlıklı tez, kabul edilmiştir. …/…/……**

**Prof. Dr. Ahmet ŞENGÖNÜL**

**Enstitü Müdürü**

# ETİK İLKELERE UYGUNLUK BEYANI

Sivas Cumhuriyet Üniversitesi Sosyal Bilimler Enstitüsü bünyesinde hazırladığım bu Yüksek Lisans tezinin bizzat tarafımdan ve kendi sözcüklerimle yazılmış orijinal bir çalışma olduğunu ve bu tezde;

1- Çeşitli yazarların çalışmalarından faydalandığımda bu çalışmaların ilgili bölümlerini doğru ve net biçimde göstererek yazarlara açık biçimde atıfta bulunduğumu;

2- Yazdığım metinlerin tamamı ya da sadece bir kısmı, daha önce herhangi bir yerde yayımlanmışsa bunu da açıkça ifade ederek gösterdiğimi;

3- Başkalarına ait alıntılanan tüm verileri (tablo, grafik, şekil vb. de dahil olmak üzere) atıflarla belirttiğimi;

4- Başka yazarların kendi kelimeleriyle alıntıladığım metinlerini, tırnak içerisinde veya farklı dizerek verdiğim yine başka yazarlara ait olup fakat kendi sözcüklerimle ifade ettiğim hususları da istisnasız olarak kaynak göstererek belirttiğimi,

beyan ve bu etik ilkeleri ihlal etmiş olmam halinde bütün sonuçlarına katlanacağımı kabul ederim.

Cansu MERCAN

16/07/2019

# ÖNSÖZ VE TEŞEKKÜR

Yüksek lisans ders ve tez aşaması süresince destek ve ilgilerini esirgemeyen, bilgilerinden faydalandığım danışman hocam Sayın Doç. Dr. N. Alp ERİLLİ'ye, her zaman teşvik ve destekleriyle yanımda bulunan ve beni büyüttüğü gibi kızlarımı da büyüten anneme, yine destekleri için babama, kardeşlerime ve eşime teşekkürlerimi sunarım.

Cansu MERCAN

# İÇİNDEKİLER

# KISALTMALAR

**SSK** : Sosyal Sigortalar Kurumu

**Bağkur** : Esnaf ve Sanatkarlar Ve Diğer Bağımsız Çalışanlar İçin Sosyal Sigortalar Kurumu

**SPSS** : (Statistical Package fort the Social Sciences) Sosyal Bilimler için İstatistik Paket Programı

**ANOVA** : (Analysis of Variance) Varyans Analizi

**ANCOVA** : (Analysis of Covariance) Kovaryans Analizi

**MANCOVA** : (Multivariate Analysis of Covariance) Çok Değişkenli Kovaryans Analizi

# TABLO LİSTESİ

# ŞEKİL LİSTESİ

# ÖZET

Çok değişkenli kovaryans analizi, birden fazla bağımlı değişkeni aynı anda inceleyerek grup ortalamaları arasındaki farkın istatistiksel olarak anlamlı olup olmadığını araştırmaktadır.

Kovaryans analizinin en temel varsayımlarından biri olan doğrusallık varsayımı, kovaryans analizi sonucunda oluşturulan modelin temsil gücünü etkileyen önemli bir etkendir. Doğrusal olmayan veriler üzerinde çok değişkenli kovaryans analizinin kullanılması, modelin temsil gücünü zayıflatacaktır.

Bu çalışmada, doğrusal olmayan kovaryans analizi tanıtılmıştır. Doğrusallık varsayımını sağlayan kovaryans analizi ile farklılıkları ortaya konmuş ve yapılan uygulama ile verinin doğrusal olup olmaması veya tam doğrusal olması durumlarının çok değişkenli kovaryans analizini nasıl etkilediği incelenmiştir.

Yapılan çalışma ile Sosyal Güvenlik Kapsamında Aktif Çalışan Kişi Sayısı ile Bakmakla Yükümlü Tutulanların Sayısının 4/a-4/b Çalışan Sigortalı Sayılarına Etkisi ile Zorunlu Sigortalı Sayısı ve İsteğe Bağlı Sigortalı Sayısının Cinsiyete Etkisi istatistiksel analizlerle incelenmiştir. İstatistiksel analizlerde; çok değişkenli doğrusal kovaryans analizi ve çok değişkenli doğrusal olmayan kovaryans analizi kullanılmış olup, doğrusal olmayan veriler üzerinde yapılacak analizin doğrusal olmayan kovaryans analizi ile yapılmasının daha uygun olduğunu ortaya koymuştur.

**Anahtar Kelimeler:** Çok Değişkenli Kovaryans Analizi, ANCOVA, Doğrusal ANCOVA, Doğrusal Olmayan ANCOVA, Kuadratik ANCOVA, Sosyal Güvenlik Verileri.

# ABSTRACT

By examining more than one dependent variable, multivariate analysis of covariance investigating whether differences between average of groups are statistically reasonable or unreasonable.

The basic assumption of linearity the assumption of covariance the analysis of covariance generated as a result of the analysis of representation is an important factor that affects the strength of the model.

Using multivariate analyse of covariance on the non-linear datas is going to weaken strenght of the model.

In this study is presented non-linear ANCOVA and is revealed differences with analyses of covariance providing hypothesis of linearity and is analyzed whether linear or non-linear and linear ANCOVA how effects of the MANCOVA.

In study is examined with statical analysis, within the scope of institution social security the number of active employees, the number of person, who is obliged to look after and 4/a-4/b the number of insured employees on the gender effects.

In statical analysis is used multi linear analysis of covariance and multi non linear ancona and is approved that non linear analysis of covariance more convenient than the analysis on non linear data. Performed analysis has revealed that analyze on non-linear datas should be practised with analyses of non-linear covarince.



**Keywords:** Multivariate Analysis of Covariance, ANCOVA, MANCOVA, Linear ANCOVA, Non-linear ANCOVA, Social Security’s Datas.

# GİRİŞ

Kovaryans analizinin temel varsayımlarından olan doğrusallık varsayımının sağlanmaması yapılan analizlerde problem teşkil etmektedir. Doğrusal olmayan veriler üzerinde yapılan doğrusal kovaryans analizi, kullanılacak olan modelin temsil etme gücünü zayıflatmaktadır. Bu nedenle de kovaryans analizi modellerinde doğrusallık şartının sağlanması önemlidir. Doğrusallık şartının sağlanmadığı veriler üzerinden yapılan analizlerde ise doğrusal olmayan kovaryans analizi modelinin kullanılması en sağlıklı sonuca ulaştıracaktır. Bu şekilde doğrusal olmayan veriler için doğrusal olmayan kovaryans analizi modeli kullanıldığında verilere en uygun ve en güçlü modele ulaşılır.

Yapılacak olan çalışmada çok değişkenli kovaryans modellerinden faydalanılacaktır. Kovaryans modellerinin doğrusal (lineer) olmaması durumunda izlenecek yollar incelenecektir. Doğrusal ve doğrusal olmayan modellerin karşılaştırması yapılarak, sosyal güvenlik verileri üzerinde uygulamaları yapılacaktır. Yapılan uygulama ile doğrusal olan veriye doğrusal ve doğrusal olmayan kovaryans analizi, aynı şekilde doğrusal olmayan veriye doğrusal ve doğrusal olmayan kovaryans analizi uygulanarak hangi yöntemin daha üstün olduğu tespit edilecek, uygulamada karşılaşılabilecek sorunlar tartışılacaktır.

## Literatür Taraması

**Adıgüzel (2010),** çalışmasında sosyal güvenlik sistemi ile ilgili olarak 5510 sayılı Kanunla yapılan reform ve daha çok kısa ve uzun vadeli sigorta kollarında yapılan değişiklikler ve sosyal güvenlik sisteminin sorunlarını anlatmış ve çözüm önerilerine yer vermiştir.

**Akçay vd. (2012),** çalışmalarında etlik piliçlerde cinsiyetin ve yemdeki protein kısıtlamasının abdominal yağ üzerine etkisini kovaryans analizi ile incelemişlerdir.

**Akıcı (2013),** çalışmasında söz konusu işletmenin yetiştirdiği Jersey ırkı ineklere ait süt verimi üzerinde durmuştur. Laktasyonun süt verimi üzerindeki üç çevresel faktör (Laktasyon Dönemi, Doğurma Dönemi, Doğurma Yılı) ile iki ortak değişken (Laktasyon Süresi, Buzağılama Yaşı) etkisi istatistiksel analizlerle incelenmiştir. İstatistiksel analizlerde; üç yönlü varyans analizi, çoklu doğrusal regresyon analizi ve iki ortak değişkenli üç yönlü kovaryans analizi kullanılmıştır.

**Burgazoğlu (2013),** çalışmasında 360 derece performans sistemini uygulayan bir şirketin farklı kaynaklardan alınan performans değerlendirme sonuçları ile çalışanların demografik özellikleri arasındaki ilişkiyi çok değişkenli kovaryans analizi kullanarak incelemiştir. Yapılan analizlerde yaş, eğitim seviyesi, unvan, kıdem yılı ve branş; çalışanların demografik özellikleri olarak bağımsız değişkenleri oluşturmuşlardır. Gerçekleştirilen çok değişkenli analizler ile elde edilen sonuçlara göre sayılan bu demografik özelliklerden sadece unvanın; alınan performans puanları üzerinde bir etkisi olduğu görülmüştür.

**Büyüköztürk (2001),** çalışmasında kovaryans analizini, bir örnek üzerinde varyans analizi ile karşılaştırmalı olarak tanıtmıştır. Çalışmada ayrıca kovaryans analizinin kullanımına ilişkin bir algoritma da tanıtılmıştır.

**Cantay (2005),** çalışmasında tesadüfi bloklar düzeninde kovaryans analizini tanıtmıştır. Çalışmada ilk önce kovaryans analizinin temel yaklaşımları ve teorik yapısı tanıtılmış, daha sonra da tesadüfi bloklar düzeninde kovaryans analizinin varyans analizine kıyaslaması yapılmıştır. Buna göre kovaryans analizinin deney hatasını azalttığı ve çalışmanın güvenilirliğini arttırdığı, örnekler ile ortaya konmuştur.

**Durmuş (2012),** çalışmasında sanal bilim ve teknoloji müzesinde eğitsel arayüz ajanı kullanımının öğrencilerin başarıları ve fen bilgisine yönelik ilgileri üzerindeki etkisini belirlemeyi amaçlamıştır. Sanal müze ortamının eğitsel arayüz ajanı içermesinin öğrencilerin fen bilgisine yönelik ilgileri üzerindeki etkisini belirlemeye yönelik olarak yapılan kovaryans analizi sonucunda, deney ve kontrol grupları arasında istatistiksel olarak anlamlı bir fark bulunmamıştır.

**Fisher (1932),** Kovaryans analizini ilk kez 1932 yılında tarımsal araştırmalarda uygulamıştır.

**Hinks ve Crassidis (2013),** çalışmalarında kovaryans analizi ile en çok olabilirlik tutum tahminini gerçekleştirmeye çalışmışlardır. Kalman filtresi ve Fischer bilgi matrisi yardımıyla önerilen yöntem, simülasyon çalışmaları ile desteklenmiştir.

**Huang ve Zhang (2014),** çalışmalarında uzay istasyonlarında randevu sistemleri için doğrusal olmayan kovaryans modeli önermişlerdir. Çalışmada uzay istasyonlarına buluşma zamanlaması ve yörünge manevra parametrelerinin hesaplanmalarında kullanılmak üzere kovaryans analizinden yararlanılmış ve önerilen yöntemin gücünü test etmek için Monte-Carlo simülasyon tekniklerini kullanılmıştır.

**Jamali vd. (2014),** çalışmalarında öğrenilmiş politika dayanıklılığını geliştirmek ve öğrenme takviyesi için bir ölçü olarak kovaryans analizini önermişlerdir. Deneysel sonuçların özellikle aykırı değerlerde daha başarılı olduğu görülmüştür.

**Jamiesson (2004),** çalışmasında son testi bağımlı değişken olarak almayıp, son test-ön test arasındaki farkı (difference skor) bağımlı değişken olarak kullanmış ve ön testi de ortak değişken olarak kabul ederek, tek bağımlı değişkenli kovaryans analizini uygulamıştır.

**Owen ve Froman (1981),** çalışmalarında kovaryans analizinin güçlü bir analiz yöntemi olduğunu, diğer yöntemlere göre varsayımlarının daha fazla olduğunu, varyans analizine regresyon eğimlerinin homojenlik varsayımının kovaryans analizi modeli için eklendiğini, kovaryans analizinde bağımlı değişken(ler) ve ortak değişken(ler) arasındaki ilişkinin doğrusal olduğunu, ilişkinin doğrusal olmaması durumunda istatistiksel gücün azalmasına sebep olacağını ileri sürmüşlerdir.

**Özer ve Sarı (2009),** çalışmalarında üniversite öğrencilerinin başarılarını etkileyen temel faktörleri tespit etmeyi ve bunların söz konusu öğrencilerin başarıları üzerindeki etkilerinin büyüklüğünü ve yönünü belirlemeyi amaçlamışlardır. Bu amaçla oluşturulan bir kovaryans analizi modeli, Atatürk Üniversitesi İktisadi ve İdari Bilimler Fakültesi öğrencilerine uygulanan ve değerlendirmeye alınan anketlerden elde edilen yatay kesit veriler kullanılarak tahmin edilmiştir.

**Özgören (2011),** çalışmasında Konya ilindeki Anadolu Lisesi, Normal Lise ve Teknik Liseden oluşan 9 farklı lisenin dokuzuncu, onuncu, on birinci ve on ikinci sınıflarına İnsani değerler ölçeğini belirlemek için hazırlanan anket soruları sorulmuştur. Öğrencilerin verdiği cevaplar SPSS ve diğer istatistiksel analiz programları kullanılarak değerlendirilmiş ve elde edilen sonuçlar incelenmiştir. Yapılan analizde kovaryans yöntemi kullanılarak farklı liselerdeki ve sınıflardaki öğrencilerin durumları anket sorularına göre değerlendirilmiştir.

**Şahin (2006),** çalışmasında çok sayıda grup, çok sayıda bağımlı değişken ve çok sayıda ortak değişken varken kovaryans analizinin nasıl işlediğini göstermeyi amaçlamıştır. Bu tez çalışmasında şeker pancarı ile ilgili olarak Konya ilinin toprağının yapısındaki kil, kum, mil, kireç değişkenlerinin etkisi giderildikten sonra yetişen şeker pancarı bitkisinin yaprakları ayasında bulunan azot, fosfor, potasyum, kalsiyum bağımlı değişkenleri bakımından Konya ilinin 5 farklı ilçesi arasında fark olup olmadığı incelenmiştir.

**Yazıcı (2001),** çalışmasında kategorik verilerde eş değişken varken kullanılan istatistiksel süreçler üzerinde durmuştur. Öncelikle nicel veriler için kullanılan Kovaryans analizi üzerinde kısaca durulmuş, ardından kategorik veriler için rassal model metotları, ağırlıklı en küçük kareler metodu, ağırlıklandırılmamış en küçük kareler metodu ve log-lineer model metotları anlatılmıştır. Bu metotlar bir grup romatizma hastasından alınan veriler üzerinde uygulanmış ve elde edilen sonuçlar yorumlanmıştır.

# 1. VARYANS ANALİZİ

İstatistik, geçmişteki ve şu andaki durumlarla ilgili verileri, geliştirilen bazı teknikler yardımıyla analiz edilerek gelecek hakkında sağlıklı karar vermemizi kolaylaştıran bir bilim dalıdır. İstatistik`te Varyans analizi (ANOVA), bazı değişkenlerin bir başka değişken üzerindeki etkisini incelemeye yarayan modelleme türüne verilen addır. Normal dağılmış iki değişken karşılaştırılırken t-testi, üç veya daha fazla değişken karşılaştırılırken ise varyans analiz tekniklerinden faydalanılır.

Deney, kontrol altındaki çeşitli durumların, deney birimlerinin bilinmeyen karakteristik özellikleri üzerindeki etkisini test etmek amacıyla uygulanan bir işlem veya süreç olarak tanımlanabilir. Deneyde birden fazla faktör olduğu durumlarda, faktör düzeylerinin kombinasyonları deneme olarak adlandırılır. Bu gibi birden fazla faktörün mevcut olduğu deneylerde temel amaç, faktörlerin ana etkileri ile beraber etkileşim etkilerini test etmektir. Deney birimlerinden elde edilen gözlem değerlerine Ronald Aylmer Fisher tarafından geliştirilen ve oldukça popüler bir teknik olan Varyans Analizi (**AN**alysis **O**f **V**Ariance - **ANOVA**) uygulanarak araştırma konusu olan faktör veya faktörlerin istatistiksel olarak anlamlı olup olmadığı belirlenmeye çalışılır (Şenoğlu, Acıtaş 2011:3).

Varyans analizi; ölçü ile belirtilen kitlelerde normal dağılıma uyan üç ya da daha fazla örneklem (gruplar) arasındaki farklılığın önemli olup olmadığını araştıran ve bu farklılığı meydana getiren sebepleri kontrol etmede kullanılan istatistiksel bir tekniktir (Cantay 2005:31). Başka bir ifade ile varyans analizi, ikiden fazla örneğin aynı ortalamaya sahip ana kitlelerden gelip gelmediği hakkında karar vermeye yarayan istatistiksel bir tekniktir (Hair vd. 1995).

Varyans analizinin genel işleyişi, farklı gözlem gruplarının varyanslarının karşılaştırılmasıyla bu grupların ortalamaları arasında bir fark olup olmadığının belirlenmesi şeklindedir (Lattin vd. 2003:389). Yapılan işleme, ortalamaların farklı olup olmadığının incelenmesine rağmen "Varyans Analizi" denmesinin altında yatan mantık bu şekilde özetlenebilir (Burgazoğlu 2013:10).

Varyans analizinin ikiden fazla ortalamanın toplu olarak kıyaslanmasında kullanılabilmesi için şu koşulların mevcut olması gerekir (Çömlekçi 1999):

i. Kıyaslanan farklı diziler, normal dağılım gösteren ana kitlelerden çekilmiş olmalıdır.

ii. Dizilerin ayrıldığı ana kitlelerin varyansları eşit olmalıdır.

Bu iki varsayım, F testinin sağlıklı sonuç verebilmesi için gereklidir. Bu iki varsayım kısaca $\varepsilon_{ij} \sim N(0,\sigma^2)$ şeklinde gösterilir. Buradaki $\varepsilon_{ij} \sim N(0,\sigma^2)$ ifadesi, hata terimlerinin bağımsız ve aynı normal dağılımlı oldukları anlamındadır (Akıcı 2013:9). Genel olarak araştırmaların birçoğunda bu varsayımlar sağlanmaktadır. Yapılan analizlerde de bu varsayımların sağlanmış olduğu varsayılarak işlem yapılır. Bu varsayımlar sağlanmıyor ise verilerin karekök alma, logaritmik dönüşüm, açı değiştirme veya araştırıcının belirlemiş olduğu dönüşümler yapılarak yeniden varsayımlar test edilir. Gerekli dönüşümlerden sonra varsayımlar sağlanmış ise varyans analizine geçilir (Efe vd. 2000). Örneklemlerin alındığı kitleler, normal dağılımlı ve eşit varyanslı olmak üzere, inceleme altındaki değişken veya değişkenler dış etkenlerden etkilenmiyorsa ortalamalar varyans analizi ile karşılaştırılır (Cantay 2005:31; Serper 2000:193).

Varyans analizinden deneysel çalışmalarda da yararlanılmaktadır. Deney, kontrol ve bekleme gruplarının deneysel işlem sonrası son test puanlarının farklı olup olmadığı varyans analizi ile test edilebilir. Varyans analizinde ise gruplar (örneklemler) topluca ele alınır ve sadece farklılığın anlamlı olup olmadığına bakılır. Varyans analizinde bu farklılığın kaynağı bilinemez. Farklılık kaynağının tespiti için çoklu karşılaştırma testleri yapılmalıdır (Akıcı 2013:9).

### 1.1. Tek Yönlü Varyans Analizi

Tek yönlü varyans analizi, deney birimleri homojen olduğunda kullanılması önerilen en uygun tasarımdır. Etkisi araştırılmak istenen yalnız bir tane faktör (bağımsız değişken) kullanıldığından dolayı, bir faktörlü varyans analizi olarak da adlandırılır (Şenoğlu, Acıtaş 2011:9).

Tek yönlü varyans analizi için matematiksel model Eşitlik (1.1) ile gösterildiği gibidir:

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}, \quad i = 1,2,...,a \qquad j = 1,2,...,n \tag{1.1}$$

Burada; $y_{ij}$; *i.* denemedeki *j.* gözlemi, $\mu$; genel ortalamayı, $\tau_i$; *i.* deneme etkisini ve $\varepsilon_{ij}$; rasgele hata terimini göstermektedir. (1.1) ile gösterilen model sabit etkili bir modeldir, bir başka deyişle Eşitlik (1.2)'deki gibi olduğu varsayılır.

$$\sum_{i=1}^{a} \tau_i = 0 \tag{1.2}$$

Tek yönlü ANOVA modelinin yapısı Şekil.1'de aşağıda gösterildiği gibidir;

| | *Gözlemler* | | | | | |
|---|---|---|---|---|---|---|
| *Denemeler* | 1 | 2 | $\cdots$ | *n* | *Toplam* | *Ortalama* |
| 1 | $y_{11}$ | $y_{12}$ | $\cdots$ | $y_{1n}$ | $y_{1.}$ | $\bar{y}_{1.}$ |
| 2 | $y_{21}$ | $y_{22}$ | $\cdots$ | $y_{2n}$ | $y_{2.}$ | $\bar{y}_{2.}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| *a* | $y_{a1}$ | $y_{a2}$ | $\cdots$ | $y_{an}$ | $y_{a.}$ | $\bar{y}_{a.}$ |
| | | | | | $y_{..}$ | $\bar{y}_{..}$ |

**Şekil 1:** Tek Yönlü ANOVA Modeli

Burada, $y_{i.} = \sum_{j=1}^{n} y_{ij}$ *ve* $\bar{y}_{i.} = \frac{y_{i.}}{n}$, $i = 1,2,...,a$ olmak üzere sırasıyla *i.* denemedeki gözlemlerin toplamını ve ortalamasını gösterir. Ayrıca, *N*=*an* toplam gözlem sayısını göstermek üzere;

$y_{..} = \sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}$ *ve* $\bar{y}_{..} = \frac{y_{..}}{N}$ eşitlikleri verilebilir. Bu eşitlikler sırasıyla tüm gözlemlerin toplamı ve tüm gözlemlerin ortalaması olarak tanımlanır (Şenoğlu, Acıtaş 2011:10).

Tek yönlü ANOVA modelinde amaç, denemelerin bağımlı değişken üzerindeki etkisinin incelenmesidir. Bir başka deyişle denemeler arasında istatistiksel olarak anlamlı bir farklılık olup olmadığını araştırmaktır. Bu durum için sıfır hipotezi üç farklı biçimde ifade edilebilir:

i. $\mathrm{H}_{0(1)}$ : Denemeler arası anlamlı bir fark yoktur.

ii. $\mathrm{H}_{0(1)}: \mu_1 = \mu_2 = \cdots = \mu_a = \mu$

iii. $\mathrm{H}_{0(1)}: \mathrm{A}_1 = A_2 = \cdots = A_a = 0$

Yukarıda belirtilen hipotezlerin üçü de birbirine denktir. Deneme ortalamalarının birbirine eşit olduğunu ifade eden *i, ii, iii* hipotezleri (1.3) eşitliği ile verilen genel kareler toplamının, deneme kareler toplamı ve hata kareler toplamı olarak bileşenlerine ayrılmasıyla elde edilen test istatistiği yardımıyla sınanır (Akıcı 2013:11).

$$\mathrm{KT}_{\mathrm{genel}} = \sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{..})^2 \tag{1.3}$$

Burada;

Deneme (Gruplar arası) Kareler Toplamı : $\mathrm{KT}_{\mathrm{deneme}} = \sum_{i=1}^{a}\sum_{j=1}^{n_i}(\bar{y}_{i.} - \bar{y}_{..})^2$ ,

Hata (Gruplar içi) Kareler Toplamı : $\mathrm{KT}_{\mathrm{hata}} = \sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i..})^2$ olup,

Eşitlik (1.3) kısaca; $\mathrm{KT}_{\mathrm{genel}} = KT_{deneme} + KT_{hata}$ şeklinde yazılır ve varyans analizinin temel denklemi olarak ifade edilir. Eşitlik (1.1) ile gösterilen modelde *i, ii, iii* hipotezlerini test edebilmek için Eşitlik (1.4)'deki test istatistiği kullanılır.

$$\mathrm{F}_{\mathrm{deneme}} = \frac{KT_{deneme}/(a-1)}{KT_{hata}/(N-a)} = \frac{KO_{deneme}}{KO_{hata}} \tag{1.4}$$

Eşitlik (1.1)'deki model, F test istatistiği $a-1$ ve $N-a$ serbestlik dereceli merkezi F dağılımına sahiptir. Eşitlik (1.4) ile elde edilen $F$ değeri için $F_{Hesap} > F_{\alpha;a-1;N-a}$ ise, denemeler arası istatistiksel olarak anlamlı bir farklılık olduğu, aksi takdirde denemeler arasında anlamlı bir farklılık bulunamadığı söylenir. Elde edilen sonuçlar Tablo 1'de olduğu gibi tek yönlü varyans analizi çizelgesinde özetlenebilir (Akıcı 2013:12).

**Tablo 1:** Tek yönlü varyans analizi çizelgesi.

| Değişim Kaynağı | Serbest-lik Derecesi | Kareler Toplamı | Kareler Ortalaması | Test İstatistiği |
|---|---|---|---|---|
| **Denemeler** | $a-1$ | $KT_{deneme}$ | $KO_{deneme} = KT_{deneme} \ / (a-1)$ | $F_{deneme} = \dfrac{KO_{deneme}}{KO_{hata}}$ |
| **Hata** | $N-a$ | $KT_{hata}$ | $KO_{hata} = KT_{hata} /(N-a)$ | |
| **Genel** | $N-1$ | $KT_{genel}$ | | |

## 1.2.İki Yönlü Varyans Analizi

İki yönlü varyans analizi, tek yönlü varyans analizine benzer şekilde, etkisi araştırılacak olan faktör sayısı iki olduğunda kullanılır. Bununla beraber, bir yönlü varyans analizinden farklı olarak, deney birimleri arasında sistematik farklılıklar söz konusudur. Bu sistematik farklılıkların etkisi kendi içinde homojen, kendi aralarında heterojen olan bloklar kullanılarak giderilmeye çalışılır (Şenoğlu, Acıtaş 2011:63).

İki yönlü varyans analizinin temel amacı, iki faktör için bir bağımlı değişken üzerindeki etkisinin ayrı ayrı test edilmesi değil, iki faktörün bağımlı değişken üzerindeki ortak etkisini bir bütün halinde ele alarak test etmektir.

İki faktörlü varyans analizinde üç ayrı test işlemi söz konusudur (Büyüköztürk 1997:141-185):

i. İki faktör birlikte ele alınarak istatistiksel olarak anlamlı bir etkiye sahip olup olmadıkları incelenebilir.

ii. Ortaya çıkmış olan ortak etkinin anlamlılığı araştırılabilir.

iii. Her bir faktörün etkisinin anlamlı olup olmadığı test edilebilir.

İki yönlü varyans analizi için matematiksel model ve veri yapısı tek yönlü varyans analizine benzer şekilde oluşturulur.

### 1.3. Latin Kare ve Greko- Latin Karesi Tasarımları

#### 1.3.1. Latin Karesi Tasarımı

Fisher (1925) ve Fisher (1926) çalışmalarında önerilen Latin Karesi Tasarımının (Latin square design-LSD) kullanımı, iki yönlü varyans analizinde olduğu gibi bloklama ilkesine dayanır. Ancak, iki yönlü varyans analizinden farklı olarak satır ve sütun adı verilen, deney birimleri arasındaki sistematik farklılıkların etkisini gidermek amacıyla yapılan iki tane bloklama değişkeni vardır. Deney birimleri arasındaki farklılık bir tane bloklama değişkeni kullanılarak giderilemeyecek kadar fazla ise iki farklı bloklama değişkeni kullanılarak homojenlik sağlanmaya çalışılır. Böylece deneysel hata azaltılmış olur. Latin karesi tasarımında, birincil derecede öneme sahip olan faktör sayısı, tek yönlü ve iki yönlü varyans analizinde olduğu gibi ‘bir’dir (Şenoğlu, Acıtaş 2011:97).

#### 1.3.2. Greko-Latin Karesi Tasarımı

Greko-Latin Karesi (Graeco-Latin square) tasarımında satır, sütun ve Yunan (Greek) harfleri olarak adlandırılan üç farklı bloklama değişkeni ile birincil derecede öneme sahip bir tane faktör vardır. Greko-Latin karesi tasarımında, bloklama değişkenleri ile faktöre ait düzey sayısının birbirine eşit olduğu kısıtı vardır. Bir başka kısıt ise faktör düzeylerini gösteren Latin harflerinin (A, B, C, … ) her satır ve her Yunan harfinde yalnız bir kez denenmesidir. Bu kısıtlardan dolayı, gerçek hayatta uygulamada nadiren kullanılan uygulamalardan biridir (Şenoğlu, Acıtaş 2011:104).

### 1.4. İç-İçe Tasarımlar

Aşamalı tasarımlar olarak da bilinen iç-içe tasarımlarda diğer tasarımlardan farklı olarak iki veya daha fazla faktör vardır.

Etkisi araştırılmak istenen A ve B gibi iki faktörden, A faktörünün düzeyi a, B faktörünün düzeyinin de b olduğu varsayıldığında ve B faktörünün b düzeyi, A faktörünün a düzeyinin her birinin içinde yuvalanmışsa bu tip tasarımlara iki aşamalı iç-içe tasarım denir. Bu durum, B faktörü, A faktörünün içinde yuvalanmıştır şeklinde ifade edilir ve B(A) sembolü ile gösterilir. Araştırmadaki önemine göre dıştaki faktör ana faktör, içteki faktör ise ikincil faktör olarak isimlendirilir. Üç faktör olduğunda ise üç faktöre sahip iç-içe tasarım ve n faktöre sahip olduğunda da n faktöre sahip iç-içe tasarım adını alır (Lindman 1992).

Yuvalanmış faktörün düzey sayısı, dıştaki faktörün her bir düzeyinde aynı ise ve her bir faktör kombinasyonundaki tekrar sayısı n ise, bu şekildeki iç-içe tasarımlara dengeli iç-içe tasarımlar adı verilir (Şenoğlu, Acıtaş 2011:125).

### 1.5. Faktöriyel Tasarımlar

Faktöriyel tasarımlar, iki ya da daha fazla faktörün ana etkilerini ve etkileşim etkilerini aynı anda araştırmak için kullanılan bir tasarımdır. Zaman ve para tasarrufu sağlamak bakımından, her seferinde bir tane faktörün etkisini araştıran geleneksel tasarımlara göre çok daha etkindirler. Bunun yanı sıra, faktörler arasındaki etkileşim veya etkileşimleri araştırmak bakımından da her seferinde bir tane faktörün etkisini araştıran tasarımlara göre avantaj sağlar (Lindman 1992).

# 2. KOVARYANS ANALİZİ

Kovaryans analizi (ANCOVA) deneme etkileri arasında istatistiksel olarak anlamlı bir farklılık olup olmadığını araştıran, kontrol edilebilen faktörlerle deney boyunca kontrol edilemeyen ortak değişken(ler)i de modelde birlikte değerlendirilen bir analiz yöntemidir. Kovaryans analizinde ortak değişken(ler)in bağımlı değişken üzerindeki etkisi arındırıldıktan sonra deneme etkileri hakkında bir sonuca varılır. Böylece deneyde hata oranı azaltılmış ve testin gücü artırılmış olur.

Kovaryans analizi ilk kez 1927 yılında Eden ve Fisher tarafından ortaya atılmıştır. 1930 yılında Sander kovaryans analizinin kullanımının etkinliği artırdığından bahsetmiştir. 1932 yılında Fisher kovaryans analizini bir istatistiksel araştırma yöntemi olarak kullanmıştır. 1934 yılında Wishard ve Wildson kovaryans analizinin uygulamalarını yapmışlardır. Pearson ise detaylı hesaplamaların nasıl yapıldığı konusunda önemli katkılarda bulunmuştur. 1957 yılında Cochran tarafından kovaryans analizi tekniğinin yararları ve uygulama alanları açıklanmıştır. 1967 yılında Quade tarafından ilk kez kategorik verilerle kovaryans analizi incelenmiştir. Bu konuda ileri çalışmalar Amara ve Koch tarafından 1980'de, Koch ve arkadaşları tarafından 1982'de yayınlanmıştır (Lindman 1992; Özer 2009; Şahin 2006:3; Şenoğlu, Acıtaş 2011:251).

Kovaryans analizi, iki veya daha fazla grupta bir bağımlı değişkenin ortalamalarının karşılaştırılması sırasında söz konusu değişkene etki eden başka bir sürekli değişkenin etkisinin ortadan kaldırılması veya bu etkinin artırılması amacıyla kullanılan istatistiksel bir süreçtir (Cantay 2005:1).

Temel olarak kovaryans analizi grup ortalamaları arasındaki farkın istatistiksel olarak anlamlı olup olmadığını incelemektedir. Kovaryans analizinde bağımlı değişken değerleri, ortak değişkene göre düzeltildikten sonra bağımsız değişkenin bağımlı değişken üzerindeki etkisi analiz edilir. Bu analizde gruplar arasındaki farklılık ölçülürken varyans analizinin yanı sıra regresyon analizi de kullanılmaktadır (Kalaycı 2010:396).

Bağımlı değişken üzerinde bir veya daha fazla etkenin etkisi araştırıldığında, bazen bağımlı değişken ile birlikte değişen başka bir değişken veya değişkenler vardır. Çoğunlukla bu diğer değişkeni (veya değişkenleri) deney süresince değişmez bir düzeyde denetim altında tutmak mümkün olmaz, ancak bu değişken (veya değişkenler) bağımlı değişkenler ile ölçülebilir. Bu değişkene, bağımlı değişken ile birlikte değiştiğinden "birlikte değişen değişken (ortak değişken)" adı verilir ve genellikle (X) ile gösterilir. Denemelerin bağımlı değişken üzerindeki etkisini ölçmek için öncelikle birlikte değişen değişkenin (veya değişkenlerin) bağımlı değişken üzerinde etkisinin giderildiği ve denemelerin bağımlı değişken üzerindeki kalan miktarının etkisinin çözümlendiği yönteme "Kovaryans Analizi" denir. Kovaryans analizi varyans analizi ile regresyon analizinin bir kombinasyonudur (Akgül 2005).

Kovaryans analizi eğitim, sağlık, ziraat gibi birçok alanda kullanılan istatistiksel bir yöntemdir. Bu yöntem gruplar arasında fark olup olmadığı sorusuna cevap vermektedir (Hicks 1994).

Ççok değişkenli kovaryans analizi (MANCOVA), varyans analizinin birden fazla bağımlı değişkeni aynı anda inceleyen çok değişkenli şeklidir. Ayrıca analize bağımlı değişkenlerle ilişkili bir ya da daha fazla sayıda ortak değişken eklenmektedir. Böylelikle çok değişkenli kovaryans analizi, ortak değişkenlerin etkisini bağımlı değişkenlerin üzerinden kaldırarak daha geçerli ve güvenilir sonuçlar elde edilmesini sağlamaktadır (Burgazoğlu 2013:1).

Yapılan analizlerde çok değişkenli kovaryans analizinin temel varsayımlarından olan doğrusallık varsayımının sağlanmaması problem teşkil etmektedir. Eğer doğrusal olmayan veriler üzerinde çok değişkenli kovaryans analizi kullanılıyorsa, kovaryans analizi modelinin temsil etme gücü zayıflamaktadır. Bu yüzden de çok değişkenli kovaryans analizi modellerinde doğrusallık şartının sağlanması önemlidir (Huitema 2011:285). Bu çalışmada doğrusal ve doğrusal olmayan modeller üzerinde çalışılacaktır.

Modelin doğrusal olmaması durumunda kovaryans analizi uygulanabilmesi için modelin doğrusal modele dönüştürülmesi gereklidir. Eğer bağımlı değişken ve bağımsız değişken arasındaki ilişki hem doğrusal olmayan hem de monoton değilse - yani bağımlı değişken artarken bağımsız değişken yalnızca bir noktaya kadar artıyor ve sonra bağımlı değişken artarken bağımsız değişken de artıyor ise- basit dönüşüm yaklaşımı da yeterli olmayacaktır. Bu durumda bağımsız değişkene polinomial yaklaşım uygulanmalıdır. Genellikle bu durumda, kuadratik veya kübik ANCOVA modellerinin kullanılması uygun olmaktadır (Huitema 2011:285).

Kovaryans analizinin iki temel amacı bulunmaktadır. Bunlardan ilki, sonuçları etkileyebilecek araştırmacının kontrolü dışında kalan sistematik hatayı ortadan kaldırmaktadır. Böylelikle hata terimi küçülecek ve F testinin duyarlılığı artacaktır. İkincisi ise grup üyelerinin belli karakteristikleri nedeniyle ortaya çıkan sonuçlar arasındaki farklılıklara açıklık getirmektir (Hair vd. 1995).

Kovaryans analizi, bağımlı değişken arasındaki ilişkinin doğrusal olup olmamasına ve ortak değişken sayısına göre aşağıdaki isimleri alır:

i. Bir tane ortak değişken varsa ve bu değişken ile bağımlı değişken arasındaki ilişki doğrusal ise basit (simple) kovaryans analizi,

ii. Bir tane ortak değişken varsa ve bu değişken ile bağımlı değişken arasındaki ilişki doğrusal değil ise eğrisel (curvilinear) kovaryans analizi,

iii. İki ya da daha fazla ortak değişken varsa ve bu değişkenler ile bağımlı değişkenler arasındaki ilişki doğrusal ise çoklu (multiple) kovaryans analizi olarak isimlendirilir (Weber, Skillings 2000).

Bağımlı değişken ile ortak değişken(ler) arasında yüksek bir korelasyon (ilişki) varsa, kovaryans analizi kullanarak analiz yapmak, varyans analizi kullanarak analiz yapmaktan daha avantajlıdır (Şenoğlu, Acıtaş 2011:252).

Kovaryans analizi tekniği kullanılarak yapılan analizler, varyans analizindeki gibi bazı temel varsayımlara dayanır. Bu beş temel varsayım aşağıda verilmiştir (Şenoğlu, Acıtaş 2011:253):

**1.Varsayım (Normallik) :** $\varepsilon_{ij}$ hata terimleri 0 ortalama ve $\sigma^2$ varyans ile normal dağılıma sahiptir.

**2.Varsayım (Eğimin Anlamlılığı) :** Ortak değişken ile bağımlı değişken arasındaki ilişkinin doğrusal olduğu varsayılır.

Varsayımın doğruluğu, $H_0 : \beta = 0$ hipotezinin sınanmasıyla veya görsel bir yöntem olan ve uygulamada yaygın olarak kullanılan saçılım (scatter) grafiğinin çizilmesiyle kontrol edilebilir.

$H_0 : \beta = 0$ hipotezi ret edilmezse deneme etkileri arasında anlamlı bir fark olup olmadığını belirlemek için varyans analizi kullanmak yeterlidir.

**3. Varsayım (Eğimlerin Homojenliği) :** Regresyon eğimlerinin homojen olduğu varsayılır.

Varsayımın doğruluğu, $H_0 : \beta_1 = \beta_2 = ... = \beta_\alpha$ veya $H_0$: "Denemelerin regresyon doğruları birbirine paraleldir" hipotezlerinin sınanmasıyla kontrol edilebilir. $H_0$ hipotezinin ret edilememesi, denemeler için bulunan regresyon doğrularının birbirine paralel olması demektir. Bu durumda, regresyon doğruları arasındaki mesafeler, ortak değişkenin tüm değerleri için aynı olacağından, önceden belirlenen herhangi bir ortak değişken değeri için, regresyon doğruları arasındaki mesafeleri karşılaştırmak, deneme etkilerini karşılaştırmaya denktir. $H_0$ hipotezinin ret edilmesi ise en az iki regresyon doğrusunun birbirine paralel olmadığını veya birbirini kestiğini gösterir. Bu durum, ortak değişken ile denemeler arasında etkileşimin varlığını ifade eder.

**4. Varsayım (Ölçüm Hatası) :** Ortak değişkenlerde ölçüm hatasının bulunması, tahmin değerlerinde ve testlerin gücünde olumsuz etkilere yol açar.

**5. Varsayım (Sabitlik) :** Ortak değişkenlerin sabit olduğu varsayılır. Bu varsayım pratikte çok geçerli değildir. Çünkü ortak değişkenlerin rasgele olduğu durumlar sabit olduğu durumlardan daha fazladır (Şenoğlu, Acıtaş 2011:255).

## 2.1. Tek Yönlü Kovaryans Analizi

### 2.1.1. Bir Ortak Değişkene Sahip Tek Yönlü Kovaryans Analizi

Bir ortak değişkene sahip tek yönlü ANCOVA modeli Eşitlik 2.1'de verildiği gibidir:

$$y_{ij} = \mu + \tau_i + \beta(x_{ij} - \bar{x}_{..}) + \varepsilon_{ij}, \qquad i=1,2,\ldots,a \quad j=1,2,\ldots,n \tag{2.1}$$

Burada, $y_{ij}$; faktörün i. düzeyinde j. deney birimine ait gözlem değerini, $\mu$; genel kitle ortalamasını, $\tau_i$; i. denemenin etkisini, $\beta$; tüm veri üzerinden ortak değişken ile bağımlı değişken arasındaki regresyon doğrusunun eğimini, $\varepsilon_{ij}$; rasgele hata terimini, $\bar{x}_{..}$; $x_{ij}$ gözlem değerlerinin ortalamasını, $a$; faktör düzey(grup) sayısını, $n_i$; i. gruptaki birim sayısını gösterir (Şenoğlu, Acıtaş 2011:257).

Tek yönlü kovaryans analizi modelinin veri yapısı Şekil.2'de gösterildiği gibidir:

| | *Denemeler* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | | 2 | | $\cdots$ | $a$ | |
| *Gözlemler* | $y$ | $x$ | $y$ | $x$ | $\cdots$ | $y$ | $x$ |
| 1 | $y_{11}$ | $x_{11}$ | $y_{21}$ | $x_{21}$ | $\cdots$ | $y_{a1}$ | $x_{a1}$ |
| 2 | $y_{12}$ | $x_{12}$ | $y_{22}$ | $x_{22}$ | $\cdots$ | $y_{a2}$ | $x_{a2}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ | $\vdots$ |
| $n$ | $y_{1n}$ | $x_{1n}$ | $y_{2n}$ | $x_{2n}$ | $\cdots$ | $y_{an}$ | $x_{an}$ |

**Şekil 2:** Tek Yönlü ANCOVA Modeli

Eşitlik (2.1) ile verilen tek yönlü kovaryans analizi modelinde parametrelerin tahmin edicileri en küçük kareler yöntemi ile elde edilir. Eşitlik (2.1) ile gösterilen modelin hata kareler toplamı;

$$\mathrm{KT}_{\mathrm{hata}} = \sum_{i=1}^{a}\sum_{j=1}^{n_i}\varepsilon_{ij}^{2} = \sum_{i=1}^{a}\sum_{j=1}^{n_i}\left\{y_{ij}-\mu-\tau_i-\beta\left(x_{ij}-\bar{x}_{..}\right)\right\}^2 \tag{2.2}$$

olmak üzere, model parametrelerinin hata kareler toplamından elde edilen denklem sistemlerinin çözümleri ile en küçük kareler tahmin edicileri, Eşitlik (2.3)'de verildiği gibidir:

$$\hat{\mu}=\bar{y}_{..} \qquad , \quad \hat{\tau}_i=\bar{y}_{i.}-\bar{y}_{..}-\hat{\beta}(\bar{x}_{i.}-\bar{x}_{..}), \quad \hat{\beta}=\frac{E_{xy}}{E_{xx}} \tag{2.3}$$

Burada,

$$E_{xx} = \sum_{i=1}^{a} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{i.})^2 \tag{2.4}$$

$$E_{xy} = \sum_{i=1}^{a} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{i.})(y_{ij} - \bar{y}_{i.}) \tag{2.5}$$

eşitlikleri ile gösterilir. Hata varyansı $\sigma^2$'nin en küçük kareler tahmin edicisi ise, Eşitlik (2.6)'da verildiği gibidir:

$$\hat{\sigma}^2 = \frac{E_{yy(düz)}}{N} \tag{2.6}$$

$E_{yy}$ ve $E_{yy(düz)}$ Eşitlik (2.7)'de verildiği gibidir:

$$E_{yy} = \sum_{i=1}^{a} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2 , \quad E_{yy(düz)} = \sum_{i=1}^{a} \sum_{j=1}^{n_i} \left\{ y_{ij} - \bar{y}_{i.} - \hat{\beta}(x_{ij} - \bar{x}_{i.}) \right\}^2 \tag{2.7}$$

Eşitlik (2.3), Eşitlik (2.6)'da yerine konulduktan sonra elde edilen ifade düzenlendiğinde Eşitlik (2.8)'deki ifade elde edilir.

$$E_{yy(düz)} = E_{yy} - \frac{E_{xy}^{\ 2}}{E_{xx}} \tag{2.8}$$

Eşitlik (2.1) ile verilen bir ortak değişkene sahip tek yönlü ANCOVA modelinde amaç, model parametrelerinin tahmin edilmesinin yanı sıra, deneme etkilerinin ve eğimin anlamlılığını test etmektir (Akıcı 2013:51).

Bu iki durum için sıfır hipotezleri,

$H_0 : \tau_1 = \tau_2 = \cdots = \tau_a = 0$ (veya $H_0$: Denemeler arasında anlamlı bir farklılık yoktur.) ve

$H_0 : \beta = 0$ (veya $H_0$: Eğim parametresi anlamlı değildir.)

şeklinde oluşturulur. Bu hipotezleri test etmek için gerekli olan test istatistikleri, bu hipotezler altında tanımlanacak olan indirgenmiş modellerin hata kareler toplamı ile tam modelin hata kareler toplamı arasındaki farktan yararlanılarak elde edilir.

Denemelerin anlamlılığını test etmek için hipotezi altında indirgenmiş model ise Eşitlik (2.9)'da verildiği gibidir.

$$y_{ij} = \mu + \beta(x_{ij} - \bar{x}_{..}) + \varepsilon_{ij}\ , \qquad i=1,2,\ldots,a \quad j=1,2,\ldots,n \tag{2.9}$$

Bu modele göre $\mu$ ve $\beta$ parametrelerinin en küçük kareler tahmin edicileri,

$\hat{\mu}_{(ind)} = \bar{y}_{..}\ ,\ \ \hat{\beta}_{(ind)} = \dfrac{S_{xy}}{S_{xx}}$ olmaktadır.

Burada, $S_{xx} = \sum_{i=1}^{a}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_{..})^2$ ; $S_{xy} = \sum_{i=1}^{a}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_{..})(y_{ij} - \bar{y}_{..})$ şeklinde ifade edilir.

Eşitlik (2.9) ile verilen indirgenmiş modelde $\mu$ ve $\beta$ parametreleri yerine tahmin edicileri kullanıldığında, indirgenmiş modelin hata kareler toplamı, Eşitlik (2.10)'daki gibi bulunur:

$$KT_{hata(ind)} = \sum_{i=1}^{a}\sum_{j=1}^{n_i}\left\{y_{ij} - \bar{y}_{..} - \hat{\beta}_{(ind)}(x_{ij} - \bar{x}_{..})\right\}^2$$

$$KT_{hata(ind)} = S_{yy} - \frac{S_{xy}{}^2}{S_{xx}} = S_{yy(düz)} \tag{2.10}$$

Eşitlik (2.10)'da yer alan $S_{yy}$ ifadesi, $S_{yy} = \sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{..})^2$ dir. Burada $T_{yy} = S_{yy} - E_{yy}$ olduğu dikkate alınırsa; $T_{yy} - \left\{\dfrac{S_{xy}{}^2}{S_{xx}} - \dfrac{E_{xy}{}^2}{E_{xx}}\right\} = T_{yy(düz)}$ şeklinde ifade edilir. Böylece hipotezini test etmek için, $F_{deneme} = \dfrac{T_{yy(düz)}/(a-1)}{\hat{\sigma}^2}$ eşitliği kullanılabilir. Eşitlikteki test istatistiği, $(a-1)$ ve $(N-a-1)$ serbestlik derecesi ile merkezi $F$ dağılımına sahiptir. Eğer, $\alpha$ önem seviyesinde $F_{deneme} > F_{\alpha;a-1;N-a-1}$ ise hipotezi ret edilir ve denemeler arası anlamlı bir farklılık vardır denir.

Bir ortak değişkene sahip tek yönlü kovaryans analizine ilişkin kovaryans analizi çizelgesi yukarıda elde edilen bilgiler ışığında Tablo 2'deki gibi oluşturulur.

**Tablo 2:** Bir ortak değişkene sahip tek yönlü kovaryans analizine ilişkin tablo.

| **Değişim Kaynağı** | **Serbestlik Derecesi** | **Kareler Toplamı** | **Kareler Ortalaması** | **Test İstatistiği** |
|---|---|---|---|---|
| **Denemeler** | $a-1$ | $T_{yy(düz)}$ | $KO_{deneme} = T_{yy(düz)} /(a-1)$ | $F_{deneme} = \dfrac{KO_{deneme}}{KO_{hata}}$ |
| **Hata** | $N-a-1$ | $E_{yy(düz)}$ | $KO_{hata} = E_{yy(düz)} /(N-a-1)$ | |
| **Genel** | $N-2$ | $S_{yy(düz)}$ | | |

### 2.1.2. İki Ortak Değişkene Sahip Tek Yönlü Kovaryans Analizi

İki ortak değişkene sahip tek yönlü ANCOVA modeli;

$$y_{ij} = \mu + \tau_i + \beta_1 (x_{1ij} - \bar{x}_1) + \beta_2 (x_{2ij} - \bar{x}_2) + \varepsilon_{ij} \,, \qquad i = 1,2,...,a \quad j = 1,2,...,n \tag{2.11}$$

şeklinde ifade edilir. Burada, $\beta_1$ : $x_1$ ile $y$ bağımlı değişkeni arasındaki regresyon doğrusunun eğimini, $\beta_2$ : $x_2$ ile $y$ bağımlı değişkeni arasındaki regresyon doğrusunun eğimini, $\bar{x}_1$ : $x_1$ ortak değişkeninin ortalamasını, $\bar{x}_2$ : $x_2$ ortak değişkeninin ortalamasını gösterir.

Eşitlik (2.11) ile gösterilen modelde parametrelerin tahmin edicileri bir ortak değişkene sahip tek yönlü kovaryans analizi modelinde olduğu gibi en küçük kareler yöntemi ile bulunabilir (Şenoğlu, Acıtaş 2011:266).

## 2.2. İki Yönlü Kovaryans Analizi

### 2.2.1. Etkileşimsiz İki Yönlü Kovaryans Analizi

Etkileşimsiz iki yönlü ANCOVA modeli Eşitlik (2.12)'de verilmiştir:

$$y_{ij} = \mu + \tau_i + \gamma_j + \beta (x_{ij} - \bar{x}..) + \varepsilon_{ij} \,, \qquad i = 1,2,...,a \quad j = 1,2,...,b \tag{2.12}$$

Burada $y_{ij}$; faktörün i. düzeyinde j. deney birimine ait gözlem değerini, $\mu$; genel kitle ortalamasını, $\tau_i$; birinci faktörün *i*. düzeydeki etkisini, $\gamma_i$; ikinci faktörün *j*. düzeydeki etkisini, $\beta$; tüm veri üzerinden ortak değişken ile bağımlı değişken arasındaki regresyon doğrusunun eğimini, $\varepsilon_{ij}$; rasgele hata terimini, $\bar{x}_{..}$; $x_{ij}$ gözlem değerlerinin ortalamasını, $a$ ve $b$; faktör düzey (grup) sayılarını gösterir (Şenoğlu, Acıtaş 2011:285).

Bu modele ilişkin veri yapısı Şekil 3'de gösterildiği gibidir;

| | *Denemeler* | | | | | | |
|---|---|---|---|---|---|---|---|
| | 1 | | 2 | | $\cdots$ | $a$ | |
| *Bloklar* | $y$ | $x$ | $y$ | $x$ | $\cdots$ | $y$ | $x$ |
| 1 | $y_{11}$ | $x_{11}$ | $y_{21}$ | $x_{21}$ | $\cdots$ | $y_{a1}$ | $x_{a1}$ |
| 2 | $y_{12}$ | $x_{12}$ | $y_{22}$ | $x_{22}$ | $\cdots$ | $y_{a2}$ | $x_{a2}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ | $\vdots$ |
| $b$ | $y_{1b}$ | $x_{1b}$ | $y_{2b}$ | $x_{2b}$ | $\cdots$ | $y_{ab}$ | $x_{ab}$ |

**Şekil 3:** Etkileşimsiz İki Yönlü Kovaryans Analizi Modeline İlişkin Veri Yapısı

Eşitlik (2.12) ile verilen modeldeki, $\varepsilon_{ij}$ hata terimlerinin kareleri toplamı; $S = \sum_{i=1}^{a}\sum_{j=1}^{b}\varepsilon_{ij}^2$ olmak üzere modeldeki bilinmeyen parametrelerin en küçük kareler tahmin edicileri Eşitlik (2.13)'de verilmiştir:

$$\hat{\mu} = \bar{y}_{..}, \quad \tau_i = \bar{y}_{i..} - \bar{y}_{..} - \hat{\beta}(\bar{x}_{i.} - \bar{x}_{..}), \tag{2.13}$$

$$\gamma_i = \bar{y}_{.j} - \bar{y}_{..} - \hat{\beta}(\bar{x}_{.j} - \bar{x}_{..}), \quad \hat{\beta} = \frac{E_{xy}}{E_{xx}}$$

Eşitlik (2.13)'de verilen formüldeki değişkenler aşağıda verildiği gibi tanımlanmıştır:

$$\bar{x}_{i.} = \frac{\sum_{j=1}^{b} x_{ij}}{b}, \quad \bar{x}_{.j} = \frac{\sum_{i=1}^{a} x_{ij}}{a}, \quad \bar{x}_{..} = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b} x_{ij}}{N},$$

$$\bar{y}_{i.} = \frac{\sum_{j=1}^{b} y_{ij}}{b}, \quad \bar{y}_{.j} = \frac{\sum_{j=1}^{a} y_{ij}}{a}, \quad \bar{y}_{..} = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij}}{N}, \ N=ab,$$

$$E_{xx} = \sum_{i=1}^{a}\sum_{j=1}^{b}(x_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x}_{..})^2,$$

$$E_{xy} = \sum_{i=1}^{a}\sum_{j=1}^{b}(x_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x}_{..})(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}),$$

$$E_{yy} = \sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2$$

Etkileşimsiz iki yönlü kovaryans analizine ilişkin ANCOVA çizelgesi yukarıda elde edilen bilgiler ışığında Tablo 3'deki gibi oluşturulur.

**Tablo 3:** Etkileşimsiz iki yönlü kovaryans analizine ilişkin tablo.

| **Değişim Kaynağı** | **Serbestlik Derecesi** | **Kareler Toplamı** | **Kareler Ortalaması** | **Test İstatistiği** |
|---|---|---|---|---|
| **Denemeler** | $a-1$ | $T_{yy(düz)}$ | $KO_{deneme} = T_{yy(düz)}/(a-1)$ | $F_{deneme} = \frac{KO_{deneme}}{KO_{hata}}$ |
| **Bloklar** | $b-1$ | $B_{yy(düz)}$ | $KO_{blok} = B_{yy(düz)}/(b-1)$ | $F_{blok} = \frac{KO_{blok}}{KO_{hata}}$ |
| **Hata** | $N-a-b$ | $E_{yy(düz)}$ | $KO_{hata} = E_{yy(düz)}/(N-a-b)$ | |
| **Genel** | $N-2$ | | | |

Tablo 3'de verilen formüller şu şekilde tanımlanmıştır:

$$E_{yy(düz)} = E_{yy} - \frac{E_{xy}^2}{E_{xx}}, \quad T_{yy(düz)} = T_{yy} - \frac{(T_{xy}+E_{xy})^2}{T_{xx}+E_{xx}} + \frac{E_{xy}^2}{E_{xx}},$$

$$B_{yy(düz)} = E_{yy(düz)} + T_{yy(düz)}, \quad F_{deneme} = \frac{T_{yy(düz)}/(a-1)}{\hat{\sigma}^2}, \quad \hat{\sigma}^2 = \frac{E_{yy(düz)}}{N-a-b}$$

### 2.2.2. Etkileşimli İki Yönlü Kovaryans Analizi

Etkileşimli iki yönlü ANCOVA modeli matematiksel olarak Eşitlik (2.14)'deki gibidir:

$$y_{ijk} = \mu + \tau_i + \gamma_j + \tau\gamma_{ij} + \beta(x_{ijk} - \bar{x}...) + \varepsilon_{ijk} \,, \qquad i = 1,2,...,a \quad j = 1,2,...,b \quad k = 1,2,...,n \tag{2.14}$$

Burada $\tau\gamma_{ij}$; faktörün i. deneme ile j. bloğun etkileşim etkisini gösteren parametredir (Lindman 1992). Eşitlik (2.14) ile gösterilen modelde parametrelerin tahmin edicileri etkileşimsiz iki yönlü ANCOVA modelinde olduğu gibi en küçük kareler yöntemi ile benzer şekilde bulunabilir.

## 2.3. Varyans Analizi ve Kovaryans Analizi Arasındaki Benzerlikler ve Farklar - ANOVA ve ANCOVA'nın karşılaştırılması

Deneysel çalışmalarda, deneylerde daha az hata bulunması ve yansız tahminler elde edilmesi amacıyla kontrol yöntemleri uygulanmaktadır. Burada kontrol, birimlerin rasgele olarak seçilmesi suretiyle homojen gruplar oluşturularak yapılır.

Kovaryans analizi bu amaçla kullanılan istatistiksel yöntemlerden biridir. Kovaryans analizi bir araştırmada etkisinin test edildiği bir faktör ya da faktörlerin dışında, bağımlı değişken ile ilişkisi bulunan bir ya da daha fazla değişkenin istatistiksel olarak kontrol edilmesini sağlayan bir teknik olarak bilinmektedir. Varyans analizi ile kovaryans analizi arasındaki en temel fark ANCOVA'nın analizde bağımlı değişken ile ilişkili olan ve ANOVA'da belirlenen bağımsız değişkenlerden farklı olarak bir ya da daha fazla değişkenin analize katılmasına olanak sağlamasıdır (Howitt, Cramer 1997; Özgören 2011:22).

Gerçekte doğru bir şekilde uygulandığında ANCOVA'nın, basit ANOVA'ya göre daha avantajlı olduğu söylenebilir. Bu iki temel avantaj; kovaryans analizinin hata varyansını azalttığından daha büyük bir istatistiksel güç sağlaması ve bir deneyin başlangıcında gruplar arası farkların olduğu durumlarda deneyin daha yansız olmasının sağlanmasıdır (Lindman 1992).

Kovaryans Analizi, sadece ortak bir değişkene ilişkin olarak gruplar arasında farkların anlamlı olması durumunda değil, bağımlı değişken ile ortak değişkene ait puanlar arasında doğrusal bir ilişki olması durumunda, başlangıçta grup ortalama puanlarının eşit olması koşulunda bile kullanılan güçlü bir istatistiktir. ANCOVA, varyans analizi ve regresyon analizi ile birlikte kullanıldığında deneyde ele alınan faktörlerin gerçek etkisi daha açık bir şekilde belirlenebilmektedir.

# 3. DOĞRUSAL MODELLER

## 3.1. Doğrusal Model Kavramı

$\underline{Y}$ vektörü bağımlı değişken, $X$ tasarım matrisi, $\underline{\beta}$ bilinmeyen parametrelerin vektörü ve $\underline{\varepsilon}$ da rasgele hataların vektörü olmak üzere; $Y$ , $X$ , $\beta$ ve $\varepsilon$ arasında

$$\underline{Y} = X\underline{\beta} + \underline{\varepsilon} \tag{3.1}$$

şeklinde ifade edilen bağıntıya doğrusal model denir. (3.1) modelinde $X\underline{\beta}$; modelin deterministik kısmı, $\underline{Y}$ ve $\underline{\varepsilon}$ ise modelin stokastik kısmı olarak adlandırılır. Modelde $I$ birim matris olmak üzere, $\underline{\varepsilon} \sim N(\underline{0}, \sigma^2 I)$ olduğu varsayılır (Akdeniz, Öztürk 1996:123).

Doğrusal modellerdeki "doğrusal" kelimesi, modelin $\underline{\beta}$ parametre vektörüne göre doğrusal olduğu anlamında kullanılmaktadır (Şenoğlu, Acıtaş 2011:321). Bu model pek çok özel hallere sahiptir. Bunlar, $\underline{\varepsilon}$'un dağılımına, $\Sigma$ kovaryans matrisine, $X$'in yapısına ve rankına bağlıdır. Bu çalışmada, aksi belirtilmedikçe, *rank( X ) = p* olduğunu kabul edilecektir yani modelimizdeki $X$ matrisi tam sütun ranklı olacaktır. $\underline{\varepsilon}$'nun dağılımı hakkında aşağıda verilen üç durumu göz önüne alınır (Akdeniz, Öztürk 1996:122).

**1. Durum:** $\underline{\varepsilon} \sim N(\underline{0}, \sigma^2 I)$

**2. Durum:** $\underline{\varepsilon}$ bilinmeyen bir dağılıma sahiptir ve $E(\underline{\varepsilon}) = \underline{0}$, $Cov(\underline{\varepsilon}) = \sigma^2 I$ dır. Bu durumu $\underline{\varepsilon} \sim (\underline{0}, \sigma^2 I)$ biçiminde gösterilir.

**3. Durum:** $Cov(\underline{\varepsilon}) = \sigma^2 V$, $V$ bilinen pozitif tanımlı bir matristir.

Birinci durumda her bir $\varepsilon_i$, 0 ortalamalı bilinmeyen $\sigma^2$ varyanslı normal dağılıma sahiptir ve $\varepsilon_i$ *(i=1,2,…,n)* 'ler bağımsızdır. İkinci durumda ise her bir $\varepsilon_i$'

nin beklenen değeri 0, $\varepsilon_i$ 'ler ilişkisiz ve $\varepsilon_i$ 'ler bilinmeyen ortak $\sigma^2$ varyansına sahiptirler (Akdeniz, Öztürk 1996:122).

### 3.2. Bir Açıklayıcı Değişkenli Doğrusal Modeller

*X* ve *Y* rasgele değişkenleri arasında bir ortak dağılım düşünmeden sadece *Y* bağımlı değişken ile ilgili gözlemlere dayanarak, Eşitlik (3.2)'deki gösterim altında, $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ , $\quad i = 1,2,...,n$ biçimindeki ifadeye basit doğrusal model denir.

$$\underline{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix}, \quad \underline{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \quad \underline{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \tag{3.2}$$

### 3.3. Birden Çok Açıklayıcı Değişkenli Doğrusal Modeller

Bir doğrusal modelde açıklayıcı değişken sayısı birden çok olduğunda bu modele çoklu doğrusal model denir. Bu modelde bağımlı değişken birden çok olduğunda modele çok değişkenli model denir (Akdeniz, Öztürk 1996:126).

Genel olarak çoklu doğrusal modeller

$\underline{Y} : n \times 1,\ X : n \times p,\ \underline{\beta} : p \times 1,\ \underline{\varepsilon} : n \times 1,\ E(\underline{\varepsilon}) = 0,\ Cov(\underline{\varepsilon}) = \sigma^2 I$ olmak üzere, Eşitlik (3.3)'de gösterildiği gibidir:

$$Y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + \underline{\varepsilon}_i, \qquad i = 1,2,...,n \tag{3.3}$$

### 3.4. Kovaryans Analizi İçin Doğrusal Modeller

En genel haliyle kovaryans analizi modeli Eşitlik (3.4)'de verildiği gibidir:

$$\mathrm{y} = \mathrm{Z}\alpha + \mathrm{X}\beta + \varepsilon \tag{3.4}$$

Burada Z ; 0 ve 1 leri, $\alpha$ ; $\mu$ ve $\alpha_i$ , $\beta_i$ , $\gamma_{ij}$ gibi faktör ve parametrelerini veya diğer etkilerini, X; ortak değişken değerlerini, $\beta$ ; ortak değişkenlerin katsayılarını içerir. Böylece ortak değişkenler Eşitlik (3.4)'ün sağ tarafında bağımsız değişken gibidir (Rencher 2001).

Tek yönlü bir ortak değişkenli model Eşitlik (3.5)'deki gibi ifade edilir.

$$\mathrm{y_{ij}} = \mu + \alpha_i + \beta x_{ij} + \varepsilon_{ij} \qquad (i = 1,2,\ldots \quad k_j = 1,2,\ldots,n) \tag{3.5}$$

Burada $\alpha_i$; muamele etkisi, $x_{\mathrm{ij}}$; $y_{\mathrm{ij}}$ ile aynı örneklemde gözlenen ortak değişken değerleri ve $\beta$; $x_{\mathrm{ij}}$ 'nin $y_{\mathrm{ij}}$ ile ilişkisinin eğimi olarak tanımlanmıştır.

Eğer (3.5) modeli, bir regresyon modeli olarak göz önüne alınırsa *k* grup için $\mu + \alpha_i$ regresyon sabitleri olarak düşünülür. *kn* adet gözlem Eşitlik (3.5)'de verilen model;

$\mathrm{y} = \mathrm{Z}\alpha + \mathrm{X}\beta + \varepsilon$ formunda Eşitlik (3.4)'de olduğu gibi yazılabilir.

Burada;

$$Z = \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 1 \end{bmatrix}, \qquad \alpha = \begin{bmatrix} \mu \\ \alpha_1 \\ \vdots \\ \alpha_k \end{bmatrix}, \qquad X = x = \begin{bmatrix} x_{11} \\ \vdots \\ x_{1n} \\ x_{2n} \\ \vdots \\ x_{kn} \end{bmatrix} \tag{3.6}$$

ve $\beta = \beta$ 'dır. Bu durumda $\mathrm{Z}$ , (3.7) deki $\mathrm{X}$ ile aynıdır.

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix} = \begin{bmatrix} j & j & 0 & \cdots & 0 \\ j & 0 & j & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ j & 0 & 0 & \cdots & 0 \end{bmatrix} . \begin{bmatrix} \mu \\ \alpha_1 \\ \vdots \\ \alpha_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_k \end{bmatrix} \tag{3.7}$$

Dolayısıyla $\mathrm{y} = \mathrm{X}\beta + \varepsilon$ olarak yazılacaktır. Tek yönlü *q* ortak değişkenli model ise (3.8)'de olduğu gibi verilir.

$$\mathrm{y_{ij}} = \mu + \alpha_{\mathrm{i}} + \beta_1 \mathrm{x_{ij1}} + \cdots + \beta_{\mathrm{q}} x_{ijq} + \varepsilon_{ij}, \qquad \mathrm{i} = 1,2,\ldots,\mathrm{k} \quad \mathrm{j} = 1,2,\ldots,\mathrm{n} \tag{3.8}$$

Bu durumda (3.6) ‘de verilen; Z , $\alpha$ ve $X\beta$ aşağıda verilen formda yazıldığı gibi olacaktır.

$$X\beta = \begin{bmatrix} x_{111} & x_{112} & \cdots & x_{11q} \\ x_{121} & x_{122} & \cdots & x_{12q} \\ \vdots & \vdots & \ddots & \vdots \\ x_{kn1} & x_{kn2} & \cdots & x_{knq} \end{bmatrix} . \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_q \end{bmatrix} \tag{3.9}$$

Tek ortak değişkenli iki yönlü model ise (3.10) eşitliğinde verildiği gibi yazılabilir.

$$y_{ijk} = \mu + \alpha_i + \gamma_i + \delta_{ij} + \beta x_{ijk} + \varepsilon_{ijk} \quad \begin{matrix} i = 1,2,\ldots,a \\ j = 1,2,\ldots,c \\ k = 1,2,\ldots,n \end{matrix} \tag{3.10}$$

Bu şekilde verilen iki yönlü model bir kaç ortak değişken içeren modele genişletilebilir (Rencher 2001).

# 4. DOĞRUSAL OLMAYAN KOVARYANS ANALİZİ

Ortak değişken ve bağımlı değişken puanları arasındaki ilişki her zaman doğrusal olamamaktadır. ANCOVA modelinin altında yatan bir düşünceye göre, *X* ve *Y* arasındaki gruplar arası ilişki doğrusal olduğu için, araştırmacılar doğrusal olmama konusunun farkında olamayabilirler. ANCOVA analizinde; veriler doğrusal olmadığında kullanıldığı zaman, F-testinin gücü azalacak ve düzeltilmiş ortalamaların deneme etkileri zayıf gözükebilecektir (Huitema 2011:285).

*X* ve *Y* arasındaki doğrusal olmayan ilişkinin iki sebebi vardır: Değişkenlerin özellikleri ve değişkenlerdeki ölçekleme hataları. Burada, değişkenlerin temel özelliklerinin ölçülebilmesinde doğrusal ilişki olmaması durumu, en önemli sorun olarak karşımıza çıkmaktadır.

Bu konuya bir örnek olarak motivasyon (*X*) ve performans (*Y*) ölçümlerini verebiliriz. Motivasyon alanında çalışan psikologlar bazen belirli bir işle ilgili çalışan bir birey için optimal motivasyon veya uyarılma seviyesi olduğunu varsayarlar. Çok düşük veya çok yüksek uyarılma seviyelerinde performans, optimal uyarılma seviyesinden daha düşük olabilmektedir. Burada; *X* ve *Y* skorları arasındaki ilişkinin doğrusal olmaması beklenir zira gözlemlenen (ölçülen) skorların altında yatan temel özellikler doğrusal olmayabilir. Bu durum, ölçekleme hatası sorununu ortaya çıkarmaktadır. Birçok ölçekleme hatası doğrusal olmama sorununa yol açsa da en çok karşılaşılan tür olarak *tavan* veya *taban* etkisini gösterebiliriz.

Her iki durumda da sorun; *X* veya *Y* değişkeninin (ya da her ikisinin de) ölçümünde kullanılan ölçme aleti veya ölçünün, ölçülen özelliklerdeki gerçek farkı yansıtmak için yeterli olmamasıdır. Örneğin, eğer bir çalışmada kullanılan deneklerin çoğu bir ölçümde mümkün olan neredeyse en yüksek sonucu alıyorsa, aynı yüksek sonucu alanlar arasında da ölçülmemiş farklılıklar olması muhtemeldir.

Ölçüm prosedürü; ölçülen özelliklerde, denekler arasındaki farklılıkları yansıtmak için yeterli "tavan" değerine sahip olmayabilir. Ortak değişken olarak kullanılan 50 puanlık bir testten deneklerin birçoğunun 50 aldığını varsaydığımızda; testin, araştırmaya katılan denekler için çok kolay olduğu görülmektedir. Eğer bu ölçümdeki puanlar uygun zorluk seviyesinde bir son test puanlarına karşı planlanırsa, doğrusal olmama durumu görülecektir. Burada, *X* ve *Y* özellikleri arasındaki doğal ilişki doğrusaldır, ama gözlemlenen ölçümler arasında elde edilen ilişki doğrusal değildir. Bu yüzden, *XY* ilişkisindeki doğrusal olmama durumunun bir sebebi ölçme hatası veya uygun olmayan ölçümler olarak karşımıza çıkacaktır.

Doğrusal olmama durumunun sebebine bakılmaksızın, doğrusal ANCOVA modelinin doğrusal olmama durumu çok şiddetli ise doğrusal yöntemlerin uygulanması uygun olmamaktadır (Huitema 2011:286).

### 4.1. Doğrusal Olmama Durumu

Verinin doğrusal olup olmadığını kontrol etmek için yapılması gereken ilk iş verinin grafiğini çizmektir. Bu ilk aşama, her gruptaki *X* puanlarına karşılık gelen *Y* puanlarının birlikte grafiğini çizme ile gerçekleştirilir. Sorun yaratacak doğrusal olmama durumu genellikle serpme diyagramında gözlemlenen trend veya marjinal dağılımlar şeklinde açıkça görülebilmektedir. Doğrusal olmama durumunu belirlemek için daha hassas yaklaşımlar içerisinde ANCOVA modelinin görsel olarak incelenmesi ve verilere çeşitli alternatif modellerin uydurulması tavsiye edilmektedir.

Doğrusal olmama durumunun problem olduğuna karar verildiğinde, bir sonraki adımda şu iki yol izlenmelidir (Huitema 2011:286):

i. Değiştirilmiş verilere, doğrusal bir ilişki ile sonuçlanacak orijinal *X* ve/veya *Y* puanlarının değişmesi için bir yol aramak.

ii. Orijinal verilere uygun bir polinominal ANCOVA modeli uydurmak.

#### 4.1.1. Veri Dönüştürme

Eğer *X* ve *Y* arasındaki ilişki, doğrusal olmayan bir ilişki olarak tanımlanmış ama monoton ise (yani, *X* arttıkça *Y*'nin de artması ama fonksiyonun doğrusal olmaması durumu var ise) *X* değişkenine dönüşüm uygulanmalıdır. İstenilen doğrusallığı sağladıkları için en sık kullanılan dönüştürme yöntemleri; logaritmik, karekök alma ve ters dönüştürmelerdir.

Dönüştürme yöntemi seçildikten sonra, dönüştürülmüş veri üzerinde her zaman olduğu gibi ANCOVA uygulanır. Örneğin, eğer $\ln X$ ve $Y$ arasındaki ilişkinin doğrusal olduğuna inanmak için sebep varsa, $\ln X$ ortak değişken olarak kullanılır ve ANCOVA uygulanır. Bununla birlikte, analizin yorumlanmasında $X$ değişkeninden ziyade $\ln X$ 'in ortak değişken olduğunu belirtmek gerekir. Dönüştürülmüş ve dönüştürülmemiş veriler için hesaplanan ANCOVA sonuçlarının doğruluğunu tespit etmek için sonuçların birlikte grafikleri çizilerek karşılaştırılmalıdır (Huitema 2011:287).

### 4.1.2. Polinominal ANCOVA Modelleri

Eğer $X$ ve $Y$ arasındaki ilişki monoton değilse, basit bir dönüştürme bile doğrusallığı sağlamayabilir. Doğrusal olmayan ve monoton durumda, $X$ değişkeninin değeri arttıkça $Y$ değişkeninin değeri de artacaktır (Huitema 2011:287).

Doğrusal olmayan ve monoton olmayan durumda ise $X$ değişkeninin değeri arttıkça $Y$ değişkeninin değeri sadece bir noktaya kadar artacak ve sonra $X$ değişkeninin değeri arttıkça $Y$ değişkeninin değeri azalacaktır. Monoton olmayan durumda $X$ değişkenini $\ln X$ 'e dönüştürdüğümüzde, $X$ değişkeninin değeri arttıkça $\ln X$ değerleri de artacaktır. $\ln X$ ve $Y$ değişkenlerinin birlikte grafiği çizildiğinde doğrusal olmama durumu görülebilecektir. Bu durumda uygulanacak alternatif yöntem; ikinci derece polinominal (kuadratik) bir ANCOVA modelidir. Bu model Eşitlik (4.1)'de verildiği gibidir:

$$\overline{Y}_{ij} = \mu + \alpha_j + \beta_1(X_{ij} - \overline{X}_{..}) + \beta_2(X_{ij}^2 - \overline{X}_{..}^2) + \varepsilon_{ij} \tag{4.1}$$

Burada;

$Y_{ij}$ : $j$. gruba ait $i$. bireyin bağımlı değişkene etkisi,

$\mu$ : Bağımlı değişkenin anakitle ortalaması,

$\alpha_j$ : $j$. deneme (işlem) etkisi,

$\beta_1$ : doğrusal regresyon katsayısı,

$X_{ij}$ : $j$. gruba ait $i$. bireyin ortak değişkene etkisi,

$\overline{X}_{..}$ : ortak değişkenlerdeki tüm gözlemlerin ortalaması,

$\beta_2$ : eğim katsayısı, $X_{ij}^2$ : j. gruba ait i. bireyin ortak değişkenin karesi,

$\overline{X}_{..}^2$ : ortak değişkenlerde gözlemlerin ortalamalarının karesi,

$\varepsilon_{ij}$ : j. gruba ait i. bireyin hata terimi olarak tanımlanmıştır.

Bu model, eğim etkisi $\beta_2(X_{ij}^2 - (\overline{X}_{..}^2))$ değerini içerdiği için doğrusal modelden farklılık göstermektedir.

Eğer bağımlı değişken puanları, ortak değişkenlerin doğrusal fonksiyonu değil de kuadratik bir fonksiyonu ise bu modelin daha uygun bir model olduğu söylenebilir. Böylece uygulanacak metotları test etmek açısından daha başarılı olacaktır. Çoklu ortak değişken analizinde iki eş değişken varsa kuadratik ANCOVA, $X$ ve $X^2$, kullanılarak hesaplanır.

Temel ANCOVA testi, regresyon testinin homojenliği, düzeltilmiş ortalamaların hesaplanması ve çoklu karşılaştırma testlerinin hepsi sıradan iki ortak değişkenli ANCOVA ile uygulanır. Eğer $X$ ve $Y$ değişkenleri arasındaki ilişki kuadratik fonksiyondan daha karmaşık yapıda ise yüksek derecede bir polinominal yapı daha kullanışlı olabilir. Üçüncü dereceden polinom (kübik) ANCOVA modeli Eşitlik (4.2)'deki gibi yazılır:

$$\overline{Y}_{ij} = \mu + \alpha_j + \beta_1(X_{ij} - \overline{X}_{..}) + \beta_2(X_{ij}^2 - \overline{X}_{..}^2) + \beta_3(X_{ij}^3 - \overline{X}_{..}^3) + \varepsilon_{ij} \qquad (4.2)$$

Ortak değişken ve bağımlı değişken arasındaki ilişki kübik fonksiyon olduğunda bu model daha iyi bir uyum sağlayacaktır. Kübik ANCOVA'da $X$, $X^2$, $X^3$ ortak değişken olarak kullanılır. Daha karmaşık fonksiyonlar için daha yüksek derecede polinomialler kullanılabilir, fakat bu çeşit durumlarla karşılaşmak son derece olağan dışıdır.

Daha yüksek derecede polinomial modeller neredeyse her zaman örnek verilere daha basit polinomial modellerden daha iyi uyum sağlarlar ama bu durum daha karmaşık modellerin daha basitlere tercih edildiği anlamına gelmemektedir.

Gerektiğinden daha karmaşık bir model kullanmamaya dikkat edilmelidir. Modeli mümkün olduğu kadar basit tutmak için temelde iki neden vardır (Huitema 2011:288).

İlk olarak, ANCOVA modele eklenen her yeni tanım için ANCOVA hata kare ortalamasından bir serbestlik derecesi kaybolur. Eğer örneklem sayısı fazla değil ise, serbestlik derecesi kaybı kolayca dengelenebilir. Hata Kareler toplamı daha karmaşık modellerde daha küçük olmasına rağmen, hata kare ortalaması karmaşık modellerde daha büyük olabilir.

Daha büyük hata teriminin sonuçları uygulanan yolların daha az kesin tahminleridir ve bu da uygulanan yollar arasındaki farklılıklar üzerinde daha az kesin testler anlamına gelir. Gerekli olandan daha karmaşık bir model uygulamamanın ikinci sebebi ise tutumluluk ilkesidir.

Eğer doğrusal bir model bir veriye neredeyse kuadratik bir model kadar uyuyorsa, daha basit model seçilmelidir çünkü sonuçların yorumlanması ve genellenmesi daha kolay olacaktır. Polinomial regresyon modellerin kullanımı ile ilişkili olarak burada bahsedilen ANCOVA ile ilişkili iki ek nokta vardır.

İlk olarak, ortak değişkenin sabit bir değişken olması gerekmez. Bazen polinamial regresyonun sadece X sabitine uygun olduğuna dair yanlış bir inanış vardır. İkinci olarak, polinomial regresyon parametrelerini bazı çoklu regresyon yöntemleri ve bilgisayar programları ile hesaplamak bazen zordur. Çünkü bu programlar belirli veri setleri ile gerekli matrisin tersini vermezler.

$X$, $X^2$, $X^3$ ve bunun gibi değişkenler, birbirleriyle yüksek derecede ilişkili olduğundan bu problem genişletilebilir. Bu hesaplama güçlükleri genellikle regresyon analizi yapılmadan önce ham $X$ puanlarını sapma puanlarına (yani ortalanmış puanlara) dönüştürerek azaltılabilir. Böylece, kuadratik ANCOVA'da $X$ ve $X^2$, yerine $(X-\overline{X})$ ve $(X-\overline{X})^2$ ortak değişken olarak kullanılmalıdır.

Düzeltilmiş deneme etkisi, çoklu belirlilik katsayıları $R^2_{YX}$ ve $R^2_{YD,X}$ 'e bağlıdır. $R^2_{YX}$ ; kuadratik regresyon ile açıklanan toplam değişkenlik oranını temsil etmektedir. (Bağımlı değişken $Y$, bağımsız değişkenler $X$ ve $X^2$, olmak üzere) $R^2_{YD,X}$ ise kuadratik regresyon ile açıklanan toplam değişkenlik oranı ve denemeleri temsil etmektedir. $Y$ 'nin, $X$ ve $X^2$ üzerindeki regresyonu $R^2_{YX}$, kuadratik regresyonu ise $R^2_{YD,X}$ dir. Bu nedenle iki katsayı arasındaki fark, bağımsız kuadratik regresyon ile hesaplanmış olan denemelerin değişkenlik oranını temsil eder. Bu yüzden iki katsayı temsilcisi arasındaki farkın oranı kuadratik regresyon tarafından hesaplanan çözümdür. Açıklanamayan değişkenlik oranı; $1-R^2_{YD,X}$ olarak tanımlanır.

Kuadratik ANCOVA'nın genel hali Tablo 4'de verilmiştir (Huitema 2011:290):

**Tablo 4:** Kuadratik ANCOVA'nın Genel Haline İlişkin Tablo.

| **Değişkenlik Kaynağı** | **Serbestlik Derecesi** | **Kareler Toplamı** | **Kareler Ortalaması** | **F Hesap Değeri** |
|---|---|---|---|---|
| *Düzeltilmiş Değerlendirme* | $J-1$ | $\left(R^2_{yD,X}-R^2_{yX}\right)KT$ | $\dfrac{\left(R^2_{yD,X}-R^2_{yX}\right)}{(J-1)}$ | $\dfrac{\left(R^2_{yD,X}-R^2_{yX}\right)/(J-1)}{\left(1-R^2_{yD,X}\right)/(N-J-2)}$ |
| *Kuadratik* $Hata_w$ | $N-J-2$ | $\left(1-R^2_{yD,X}\right)KT$ | $\dfrac{\left(1-R^2_{yD,X}\right)}{(N-J-2)}$ | |
| *Kuadratik* $Hata_t$ | $N-1-2$ | $\left(1-R^2_{yX}\right)KT$ | | |

Tabloda yer alan KT; kareler toplamını, N; toplam gözlem sayısını, J ise değişken sayısını ifade etmektedir.

Çoklu ANCOVA'da regresyonun homojenlik testi önemli olduğu gibi, kuadratik ANCOVA'da da ayrı gruplar için kuadratik regresyonun homojenlik testi önemlidir. Bu test regresyon yüzeylerinin homojenlik testi ile aynı şekilde hesaplanır. Herhangi bir derecedeki kuadratik regresyon modeli için Homojenlik testinin genel formu Tablo 5'de verildiği gibidir:

**Tablo 5:** Kuadratik Regresyon Modeli için Homojenlik Testinin Genel Formu

| Değişkenlik Kaynağı | Serbestlik Derecesi | Kareler Toplamı | Kareler Ortalaması | F Hesap Değeri |
|---|---|---|---|---|
| *Kuadratik Regresyonun Heterojenliği* | $2(J-1)$ | $\left(R^2_{yD,X,DX} - R^2_{yD,X}\right)KT$ | $\dfrac{\left(R^2_{yD,X,DX} - R^2_{yD,X}\right)}{\left(2(J-1)\right)}$ | $\dfrac{\dfrac{\left(R^2_{yD,X,DX} - R^2_{yD,X}\right)}{\left(2(J-1)\right)}}{\dfrac{\left(1-R^2_{yD,X,DX}\right)}{\left(N-J-3\right)}}$ |
| *Kuadratik* $Hata_i$ | $N-(J-3)$ | $\left(1-R^2_{yD,X,DX}\right)KT$ | $\dfrac{\left(1-R^2_{yD,X,DX}\right)}{\left(N-J-3\right)}$ | |
| *Kuadratik* $Hata_w$ | $N-J-2$ | $\left(1-R^2_{yD,X}\right)KT$ | | |

Herhangi bir derecedeki polinomial regresyon modeli için Homojenlik testinin genel formu ise Tablo 6'da verildiği gibidir:

**Tablo 6:** Herhangi bir derecedeki Polinomial Regresyon Modeli için Homojenlik Testinin Genel Formu

| Değişkenlik Kaynağı | Serbestlik Derecesi | Kareler Toplamı | Kareler Ortalaması | F Hesap Değeri |
|---|---|---|---|---|
| *Polinomial Regresyonun Heterojenliği* | $C(J-1)$ | $\left(R^2_{yD,X,DX} - R^2_{yD,X}\right)KT$ | $\dfrac{\left(R^2_{yD,X,DX} - R^2_{yD,X}\right)}{C(J-1)}$ | $\dfrac{\dfrac{\left(R^2_{yD,X,DX} - R^2_{yD,X}\right)}{\left(C(J-1)\right)}}{\dfrac{\left(1-R^2_{yD,X,DX}\right)}{\left(N-J(C+1)\right)}}$ |
| *Polinomial* $Hata_i$ | $N-J(C+1)$ | $\left(1-R^2_{yD,X,DX}\right)KT$ | $\dfrac{1-R^2_{yD,X,DX}}{\left(N-J(C+1)\right)}$ | |
| *Polinomial* $Hata_w$ | $N-J-C$ | $\left(1-R^2_{yD,X}\right)KT$ | | |

Tablo 6'da yer alan C değeri; homojenlik derecesini ifade etmektedir.

# 5. UYGULAMA

## 5.1. Veri Seti

Uygulamada 2018 yılı Aralık ayına ait Sosyal Güvenlik verileri kullanılmıştır. Yukarıdaki bölümlerde bahsedilen yöntemler 2018 yılı Aralık ayına ait Sosyal Güvenlik verilerine uygulanmış ve analiz sonuçları değerlendirilmiştir.

Bu bölümde Sosyal Güvenlik Kurumu ve sigortalılık hakkında kısa bir bilgi verilmiştir.

20.05.2006 tarihinde yürürlüğe giren 5502 sayılı Kanunla Sosyal Güvenlik Kurumu kurulmuştur. Sosyal Güvenlik Kurumu, vatandaşların doğumundan ölümüne kadar, ölümünden sonra da ölen kişinin hak sahiplerine sağlık, sigorta ve sosyal yardım alanlarında, "Değişen sosyal güvenlik ihtiyaç ve risklerine karşı toplumu güvence altına alarak güvenilir, kaliteli ve yenilikçi bir anlayışla sürdürülebilir sosyal güvenlik hizmeti sunmak" misyonu doğrultusunda hizmetler sunmaktadır.

Sosyal Güvenlik Kurumu, Türkiye'de daha önceleri sosyal güvenlik sistemini yürüten Sosyal Sigortalar Kurumu (SSK), Bağ-Kur ve Emekli Sandığı'nın görevlerini devralmıştır. 20.05.2006 tarihinden itibaren sosyal güvenlik sistemi tek bir kurum ile yürütülmeye başlanmıştır. Birleşmenin ardından çıkartılan 5510 sayılı Sosyal Sigortalar ve Genel Sağlık Sigortası Kanunu ile birlikte, daha önce beş farklı kanun çerçevesinde sosyal güvenceye sahip olan kişilerin tümü 5510 sayılı kanun kapsamına alınarak sosyal güvenlik sitemi uygulamalarında norm ve standart bütünlüğü sağlanmıştır (Adıgüzel 2010:4).

5510 sayılı Sosyal Sigortalar ve Genel Sağlık Sigortası Kanununa göre sigortalı; kısa (iş kazası, meslek hastalığı, hastalık ve analık sigortası kolları) veya uzun vadeli (malullük, yaşlılık ve ölüm sigortası kolları) sigorta kolları bakımından adına prim ödenmesi gereken veya kendi adına prim ödemesi gereken kişiyi temsil etmektedir. (https://www.sgk.gov.tr, Erişim Tarihi: 11.07.2019).

5510 sayılı Kanununun 4. Maddesi 1. Fıkrasının (a) bendinde sayılanlar kısaca eski SSK statüsü, aynı Kanunun 4. Maddesi 1. Fıkrasının (b) bendinde sayılanlar kısaca eski Bağ-Kur statüsü, yine aynı Kanunun 4. Maddesi 1. Fıkrası (c) bendinde sayılanlar kısaca eski Emekli Sandığı statüsü ile aynı konumdadır.

Çalışmada anlatılan 4/a'lı çalışan sigortalı sayısı eski SSK çalışanlarını, 4/b'li çalışan sigortalı sayısı eski Bağ-Kur çalışanlarını, 4/c'li çalışan sigortalı sayısı eski Emekli Sandığı çalışanlarını ifade etmektedir.

Yapılan uygulamalarda SPSS 21 paket programı kullanılmış, tüm uygulama boyunca istatistiksel anlamlılık düzeyi olan alfa değeri 0,05 olarak alınmış olup, Sosyal Güvenlik verileri ile üç farklı yapıdaki veri seti için Doğrusal ve Doğrusal olmayan kovaryans analizleri kullanılmıştır.

Birinci uygulamada; Sosyal Güvenlik Kapsamında Kişi Sayısı ve Türkiye Nüfusuna Oranı Tablosu adı altında verilerin doğrusal olmadığı durumda Sosyal Güvenlik Kapsamında Aktif Çalışan Kişi Sayısı ile Bakmakla Yükümlü Tutulanların Sayısının 4/a ve 4/b li Çalışan Sigortalı Sayısına Etkisi araştırılmış, doğrusal ve doğrusal olmayan ANCOVA yöntemlerinin uygulanmasının sonuçları verilmiş ve tartışılmıştır.

Uygulama 1(a) ile Sosyal Güvenlik Kapsamında Aktif Çalışan Kişi Sayısı ile Bakmakla Yükümlü Tutulanların Sayısının 4/a ve 4/b li Sigortalı Sayısına Çalışan Sigortalı Sayısına Etkisi veriler doğrusal değilken doğrusal olmayan ANCOVA analizi uygulanmıştır. Uygulama 1(b) ile aynı veriye verinin %3,7'si kadar daha veri eklenerek doğrusal yapıdan uzaklaştırılmış, yine doğrusal olmayan veriler için doğrusal olmayan ANCOVA analizi uygulanmıştır. Uygulama 1(c) ile aynı verinin doğrusal olduğu varsayılarak doğrusal ANCOVA analizi uygulanmıştır. Burada, $X_1$: 4/a'lı aktif çalışan kişi sayısı, $X_2$: 4/b'lı aktif çalışan kişi sayısı, $X_3$: 4/a'lı bakmakla yükümlü tutulanların sayısı, $X_4$: 4/a'lı bakmakla yükümlü tutulanların sayısı olarak tanımlanmıştır.

İkinci uygulamada; 4/b Kapsamındaki Aktif Sigortalıların İl ve Cinsiyet Tablosu adı altında verinin doğrusal olduğu durumda 4/b Kapsamındaki Aktif Sigortalılardan Zorunlu Sigortalı Sayısı ve İsteğe Bağlı Sigortalı Sayısının Cinsiyete Etkisi araştırılmış, doğrusal ve doğrusal olmayan ANCOVA yöntemlerinin uygulanmasının sonuçları verilmiş ve tartışılmıştır. Burada, $X_1$: Zorunlu sigortalı erkeklerin sayısı, $X_2$: Zorunlu sigortalı kadınların sayısı, $X_3$: İsteğe bağlı sigortalı erkeklerin sayısı, $X_4$: İsteğe bağlı sigortalı kadınların sayısı olarak tanımlanmıştır.

Uygulama 2(a) ile 4/b Kapsamındaki Aktif Sigortalılardan Zorunlu Sigortalı Sayısı ve İsteğe Bağlı Sigortalı Sayısının Cinsiyete Etkisi veriler doğrusal iken doğrusal ANCOVA analizi uygulanmıştır. Uygulama 2(b) ile aynı verinin doğrusal olmadığı varsayılarak doğrusal olmayan ANCOVA analizi uygulanmıştır.

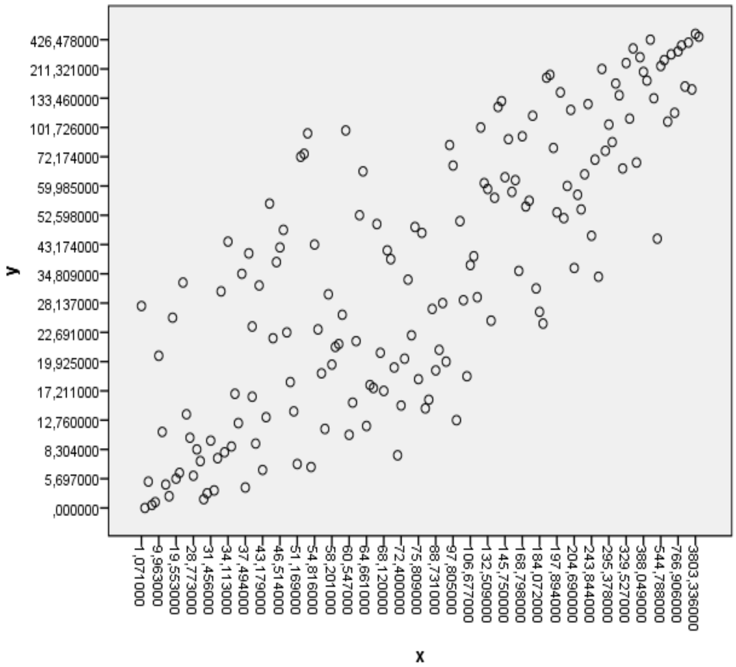
Üçüncü uygulamada ise ikinci uygulamada kullanılan veri, dönüşüm yöntemleri ile tam doğrusala yakın hale getirilmiş, uygulanan doğrusal ve doğrusal olmayan ANCOVA yöntemlerinin sonuçları yapılan analizlerle açıklanmıştır.

Uygulama 3(a) ile Uygulama 2'de kullanılan logaritmik dönüşüm işlemi ile verilerin doğrusallığı artırılarak doğrusal ANCOVA analizi uygulanmıştır. Uygulama 3(b) ile aynı veri tam doğrusal yapıya yakınlaşacak şekilde verinin %4,9'unda değişiklik yapılarak, yine doğrusal veriler için doğrusal ANCOVA analizi uygulanmıştır. Uygulama 3(c) ile verinin doğrusal olmadığı varsayılarak doğrusal olmayan ANCOVA analizi uygulanmıştır.

## 5.2. Doğrusal Olmayan Veri Üzerinden Yapılan Uygulamalar

### 5.2.1. Uygulama 1(a): Veriler doğrusal değilken Aktif Çalışan Sayısı ile Bakmakla Yükümlü Tutulanların Sayısının 4/a ve 4/b'li Sayısına Etkisi

Kuadratik (doğrusal olmayan) ANCOVA, çoklu ANCOVA ile aynı özelliklere sahip olduğundan, hesaplamalar benzer şekilde olmaktadır. Aşağıda verilen yayılım grafiğinde uygulamada kullanılan verinin doğrusal olmadığı görülmektedir:




Önceki araştırmalar veya teorik karşılaştırmalarda, X ve Y arasındaki ilişki en iyi kuadratik fonksiyon olarak önerilmiştir. Verinin yayılım grafiği kuadratik fonksiyon fikrini de desteklemektedir. Böylece deney tasarımını gerçekleştiren araştırmacının kuadratik ANCOVA modelini seçmesi için geçerli sebepleri bulunmaktadır.

Çoklu kovaryans analizinde $X$ ve $X^2$ ’yi ortak değişken olarak kullanarak hesaplamalara devam edilir. Daha önce belirtildiği gibi; düzeltilmiş deneme etkisi, çoklu belirtme katsayıları $R^2_{YX}$ ve $R^2_{YD,X}$ ’e bağlıdır. $R^2_{YX}$ ; kuadratik regresyon ile açıklanan toplam değişkenlik oranını temsil etmektedir. (Bağımlı değişken $Y$, bağımsız değişkenler $X$ ve $X^2$ olmak üzere) $R^2_{YD,X}$ ise kuadratik regresyon ile açıklanan toplam değişkenlik oranı ve denemeleri temsil etmektedir. $Y$ ’nin, $X$ ve $X^2$ üzerindeki regresyonu $R^2_{YX}$ , kuadratik regresyonu ise $R^2_{YD,X}$ dir.

Bu nedenle iki katsayı arasındaki fark, bağımsız kuadratik regresyon ile hesaplanmış olan denemelerin değişkenlik oranını temsil eder. Bu yüzden iki katsayı temsilcisi arasındaki farkın oranı kuadratik regresyon tarafından hesaplanan çözümdür. Açıklanamayan değişkenlik oranı ise $1 - R^2_{YD,X}$ tir.

Araştırma konusu olan ve modelde test edilecek hipotezler ise şunlardır:

$H_0 =$ Grup ortalamaları arasında fark yoktur.

$H_1 =$ En az bir grup ortalaması diğerlerinden farklıdır.

Modele regresyon analizi uygulandığında $R^2_{YD,X} = R^2_{Y123} = 0{,}752$ ve $R^2_{YX} = R^2_{Y23} = 0{,}696$ sonuçları elde edilir. Kukla değişkenlerin fark katkısı üzerinden kuadratik regresyon katsayısı ise 0,056 olarak hesaplanmıştır. Genel kareler toplamı da *KT* = 3764269,011 olarak hesaplanmıştır.

Kuadratik ANCOVA'nın genel hali Tablo 7'de verilmiştir:

**Tablo 7:** Kuadratik ANCOVA'nın Genel Hali

| **Değişkenlik Kaynağı** | **Serbestlik Derecesi** | **Kareler Toplamı** | **Kareler Ortalaması** | **F Hesap Değeri** |
|---|---|---|---|---|
| *Düzeltilmiş Değerlendirme* | $J-1$ | $\left(R^2_{yD,X} - R^2_{yX}\right)KT$ | $\dfrac{\left(R^2_{yD,X} - R^2_{yX}\right)}{(J-1)}$ | $\dfrac{\left(R^2_{yD,X} - R^2_{yX}\right)/(J-1)}{\left(1-R^2_{yD,X}\right)/(N-J-2)}$ |
| *Kuadratik* $Hata_w$ | $N-J-2$ | $\left(1-R^2_{yD,X}\right)KT$ | $\dfrac{\left(1-R^2_{yD,X}\right)}{(N-J-2)}$ | |
| *Kuadratik* $Hata_t$ | $N-1-2$ | $\left(1-R^2_{yX}\right)KT$ | | |

Uygulama verisi için kuadratik ANCOVA sonuçları Tablo 8'de verilmiştir:

**Tablo 8:** Uygulama 1(a) Verisi için Kuadratik ANCOVA Sonuçları

| **Değişkenlik Kaynağı** | **Serbestlik Derecesi** | **Kareler Toplamı** | **Kareler Ortalaması** | **F Hesap Değeri** |
|---|---|---|---|---|
| *Düzeltilmiş Değerlendirme* | 2 | 229620,4 | 114810,2 | 25,443 |
| *Kuadratik* $Hata_w$ | 156 | 703918,3 | 4512,296 | |
| *Kuadratik* $Hata_t$ | 158 | 933538,7 | | |

$F_{tablo} = F_{0,05;2;156}$ değerinin 3,054 olduğu göz önüne alındığında; hesaplanan F değerinin anlamlı olduğu, deney ve kontrol grupları arasında farkın olduğu görülmektedir. Bu kuadratik yani doğrusal olmayan ANCOVA modeli, verileri temsil edecek şekilde kabul edilir.

Düzeltilmiş ortalama ve çoklu karşılaştırma işlemleri çoklu ANCOVA modeli gibi işlem görürler. Düzeltilmiş ortalamalar $R^2_{Y123}$ üzerinden hesaplanan regresyon eşitliği ile elde edilir.

Modele regresyon analizi uygulandığında regresyon katsayıları $\beta_0 = 46{,}548;\ \beta_1 = \text{-}72{,}761;\ \beta_2 = 0{,}363;\ \ \beta_3 = \text{-}0{,}000$ olarak bulunmuştur.

Ayrıca 1.gruba ait ortalama gölge değişken skoru 1, ortalama ortak değişken skoru 213,333 ve ortalama kareler ortak değişken skoru 280341,187 olarak hesaplanmıştır. Buna göre 1. gruba ait düzeltilmiş deneme etkisi;

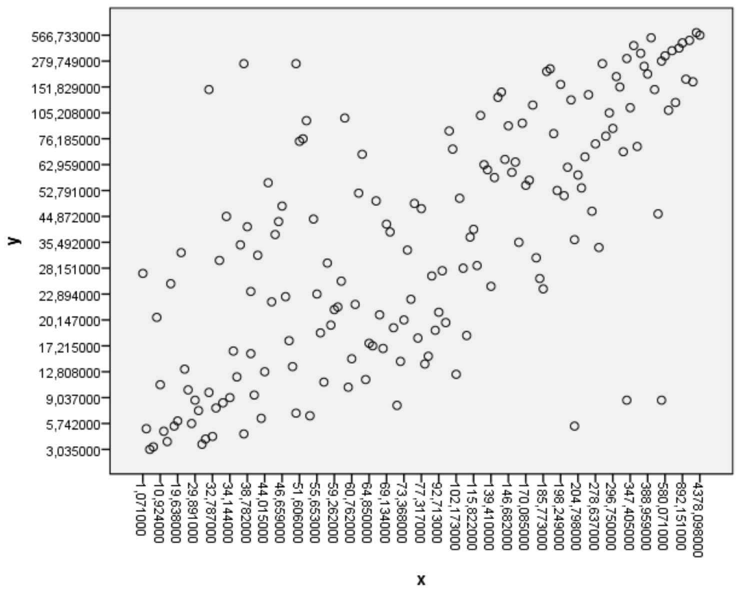
$\overline{Y}_1(düz) = 46{,}548 - 72{,}761(1) + 0{,}363(213{,}333) - 0(280341{,}187) = 51{,}226$ olmaktadır.

Benzer şekilde 2. gruba ait ortalama gölge değişken skoru 0, ortalama ortak değişken skoru 213,333 ve ortalama kareler ortak değişken skoru 280341,187 olarak hesaplanmıştır. Buna göre 2. gruba ait düzeltilmiş deneme etkisi;

$\overline{Y}_2(düz) = 46{,}548 - 72{,}761(0) + 0{,}363(213{,}333) - 0(280341{,}187) = 123{,}987$ olmaktadır.

### 5.2.2. Uygulama 1(b): Veriler Doğrusal Yapıdan Uzaklaştıkça Aktif Çalışan Kişi Sayısı ile Bakmakla Yükümlü Tutulanların Sayısının 4/a ve 4/b li Sayısına Etkisi

Uygulama 1(b) ile verilere doğrusallığı bozacak şekilde %3,7 oranında yeni gözlem eklenerek veriler doğrusal yapıdan uzaklaştırılmış olup kuadratik ANCOVA analizi uygulanmıştır. Aşağıda verilen yayılım grafiğinde uygulamada kullanılan verinin bir önceki uygulamadaki yayılım grafiğindekine göre doğrusallıktan uzaklaştığı görülmektedir.




Tablo 9'da ise Uygulama 1(b) için kuadratik ANCOVA sonuçları verilmiştir:

**Tablo 9:** Uygulama 1(b) Verisi için Kuadratik ANCOVA Sonuçları

| Değişkenlik Kaynağı | Serbestlik Derecesi | Kareler Toplamı | Kareler Ortalaması | F Hesap Değeri |
|---|---|---|---|---|
| *Denemeler* | 2 | 255771 | 127885,72 | 27,409 |
| *Hata* | 162 | 755862 | 4665,814 | |
| *Genel* | 164 | 1011633 | | |

$F_{tablo} = F_{0,05;2;162}$ değerinin 3,051 olduğu göz önüne alındığında; hesaplanan F değerinin anlamlı olduğu, deney ve kontrol grupları arasında farkın olduğu görülmektedir. Bu doğrusal ANCOVA modeli, verileri temsil edecek şekilde kabul edilir.

### 5.2.3. Uygulama 1(c): Veriler Doğrusal İken Aktif Çalışan Sayısı ile Bakmakla Yükümlü Tutulanların Sayısının 4/a ve 4/b li Sayısına Etkisi

Uygulama 1(a) ve Uygulama 1(b) ile yapılan analizlerde veriler doğrusal olmadığından kuadratik ANCOVA yöntemi uygulanmıştır. Uygulama 1(c) ile de verileri doğrusal kabul edip doğrusal ANCOVA analizi uygulanmıştır. Tablo 10'da ise Uygulama 1(c) için ANCOVA sonuçları verilmiştir:

**Tablo 10:** Uygulama 1(c) Verisi için Doğrusal ANCOVA Sonuçları

| Değişkenlik Kaynağı | Serbestlik Derecesi | Kareler Toplamı | Kareler Ortalaması | F Hesap Değeri |
|---|---|---|---|---|
| *Denemeler* | 1 | 241809 | 241809,33 | 38,673 |
| *Hata* | 159 | 994167 | 6252,622 | |
| *Genel* | 160 | | | |

$F_{tablo} = F_{0,05;1;159}$ değerinin 3,900 olduğu göz önüne alındığında; hesaplanan F değerinin anlamlı olduğu, deney ve kontrol grupları arasında farkın olduğu görülmektedir. Bu doğrusal ANCOVA modeli, verileri temsil edecek şekilde kabul edilir.

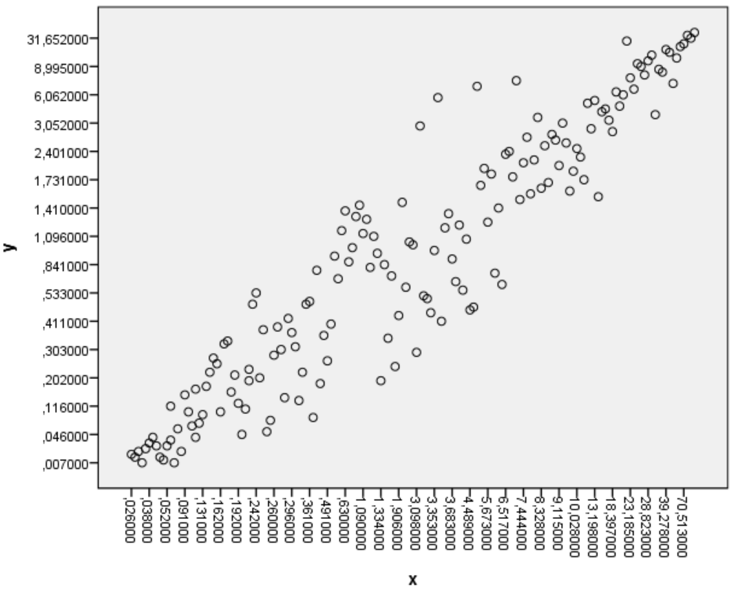
Aynı veriler kullanılarak yapılan iki farklı analizden verinin daha fazla doğrusal olmayan durumunun sonuçlar üzerindeki etkisi yapılan uygulamalarla desteklenmiştir. Veriler doğrusal değilken yapılan kuadratik ANCOVA analizlerinde ve verilerin doğrusal kabul edildiği durumda uygulanan doğrusal ANCOVA analizinde oluşturulan her üç model de anlamlı çıkmıştır. Ancak ilk modelde hesaplanan F değeri diğer iki modeldeki hesaplanan F değerlerinden daha küçük olduğundan ilk modelin yani kuadratik ANCOVA modelinin kullanılması daha uygun olduğu söylenebilir.

Böylece $\overline{Y}\left(düz\right) = 46{,}548 - 72{,}761d + 0{,}363x - 0x^2$ modeli verileri temsil edecek en uygun model olarak belirlenmiştir.

### 5.3. Doğrusal Veri Üzerinden Yapılan Uygulamalar

#### 5.3.1. Uygulama 2(a): Veriler doğrusal iken Zorunlu Sigortalı Sayısı ve İsteğe Bağlı Sigortalı Sayısının Cinsiyete Etkisi

Aşağıda verilen yayılım grafiğinde uygulamada kullanılan verinin doğrusal olduğu görülmektedir.




Tablo 11'de Uygulama 2(a) verisi için doğrusal ANCOVA sonuçları verilmiştir:

**Tablo 11:** Uygulama 2(a) Verisi için Doğrusal ANCOVA Sonuçları

| **Değişkenlik Kaynağı** | **Serbestlik Derecesi** | **Kareler Toplamı** | **Kareler Ortalaması** | **F Hesap Değeri** |
|---|---|---|---|---|
| *Denemeler* | 1 | 42,574 | 42,574 | 8,648 |
| *Hata* | 159 | 782,793 | 4,923 | |
| *Genel* | 160 | 825,367 | | |

$F_{tablo} = F_{0,05;1;159}$ değerinin 3,900 olduğu göz önüne alındığında; hesaplanan F değerinin anlamlı olduğu, deney ve kontrol grupları arasında farkın olduğu görülmektedir. Bu doğrusal ANCOVA modeli, verileri temsil edecek şekilde kabul edilir.

Modele regresyon analizi uygulandığında regresyon katsayıları $\beta_0 = 6,903$; $\beta_1 = 42,574$; $\beta_2 = 14589,475$ olarak bulunmuştur.

1.gruba ait ortalama gölge değişken skoru 1, ortalama ortak değişken skoru 10,668 olarak hesaplanmıştır. Böylece 1. gruba ait düzeltilmiş deneme etkisi

$$\overline{Y}_1(d\ddot{u}z) = 6,903 + 42,574\,(1) + 14589,475\,(10,668) = 155694,499$$

olarak bulunmuştur.

Benzer şekilde 2.gruba ait ortalama gölge skoru 0, ortalama ortak değişken skoru 10,668 olarak hesaplanmıştır. Böylece 2. gruba ait düzeltilmiş deneme etkisi

$$\overline{Y}_2(d\ddot{u}z) = 6,903 + 42,574\,(0) + 14589,475\,(10,668) = 155651,925$$

olarak bulunmuştur.

### 5.3.2. Uygulama 2(b): Verilerin Doğrusal Olmadığı Varsayıldığında Zorunlu Sigortalı Sayısı ve İsteğe Bağlı Sigortalı Sayısının Cinsiyete Etkisi

Uygulama 2(a) ile yapılan analizde veriler doğrusal olduğundan doğrusal ANCOVA yöntemini kullanılmıştır. Uygulama 2(b) ile de verilerin doğrusal olmadığı varsayılarak kuadratik yani doğrusal olmayan ANCOVA analizi uygulanmıştır.

Regresyon analizi ile $R^2_{YD,X} = R^2_{Y123} = 0,954$ ve $R^2_{YX} = R^2_{Y23} = 0,950$ sonuçları elde edilir. Kukla değişkenlerin fark katkısı üzerinden kuadratik regresyon katsayısı 0,004 olarak hesaplanmıştır. Genel kareler toplamı 16238,639 olarak bulunmuştur. Hesaplanan kuadratik ANCOVA sonuçları Tablo 12’de verilmiştir:

**Tablo 12:** Uygulama 2(b) Verisi için Kuadratik ANCOVA Sonuçları

| **Değişkenlik Kaynağı** | **Serbestlik Derecesi** | **Kareler Toplamı** | **Kareler Ortalaması** | **F Hesap Değeri** |
|---|---|---|---|---|
| *Düzeltilmiş Değerlendirme* | 2 | 16,238 | 8,119 | 1,733 |
| *Kuadratik* $Hata_w$ | 156 | 730,738 | 4,684 | |
| *Kuadratik* $Hata_t$ | 158 | 746,977 | | |

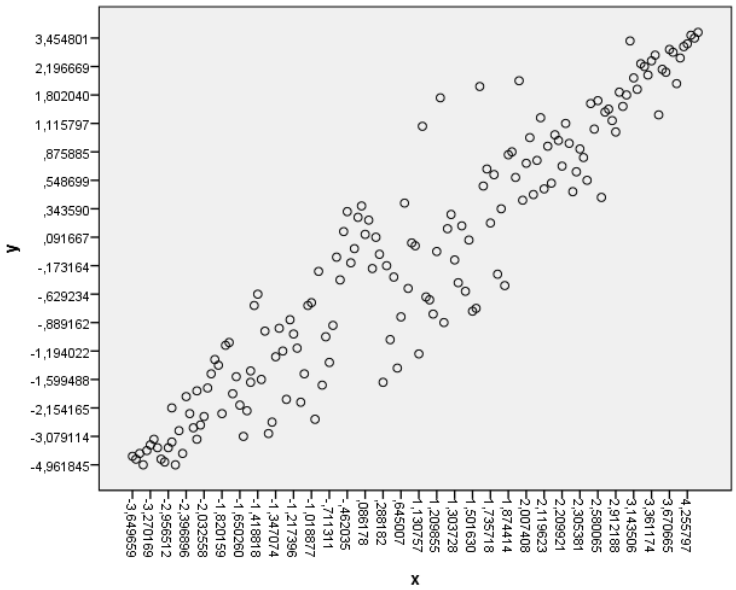
$F_{tablo} = F_{0,05;2;162}$ değerinin 3,054 olduğu göz önüne alındığında; hesaplanan F değerinin anlamlı olmadığı, deney ve kontrol grupları arasında farkın olmadığı görülmektedir. Bu doğrusal ANCOVA modeli, verileri temsil edecek şekilde kabul edilemez.

Aynı veriler kullanılarak yapılan iki farklı analizden verilerin doğrusal olup olmama durumlarının sonuçlar üzerindeki etkisi yapılan uygulamalarla desteklenmiştir. Veriler doğrusal iken yapılan doğrusal ANCOVA analizinde oluşturulan model anlamlı çıkarken, verilerin doğrusal olmadığı varsayılarak aynı veri ile yapılan kuadratik ANCOVA analizinde model anlamsız çıkmaktadır.

Sonuç olarak verileri temsil edecek en uygun modelin; $\overline{Y}\left(d\ddot{u}z\right) = 6{,}903 + 42{,}574\,d + 14589{,}475\,x$ olduğu belirlenmiştir.

### 5.3.3. Uygulama 3(a): Veriler Tam Doğrusal Yapıya Yakınlaştıkça Zorunlu Sigortalı Sayısı ve İsteğe Bağlı Sigortalı Sayısının Cinsiyete Etkisi

Bir önceki uygulamada kullanılan verilere logaritmik dönüşüm işlemi uygulanmıştır. Aşağıda verilen yayılım grafiğinde uygulamada kullanılan verinin bir önceki uygulamadaki yayılım grafiğindekine göre daha doğrusal olduğu görülmektedir.




Uygulama 3(a) hesaplanan verisi için doğrusal ANCOVA sonuçları Tablo 13'de verilmiştir:

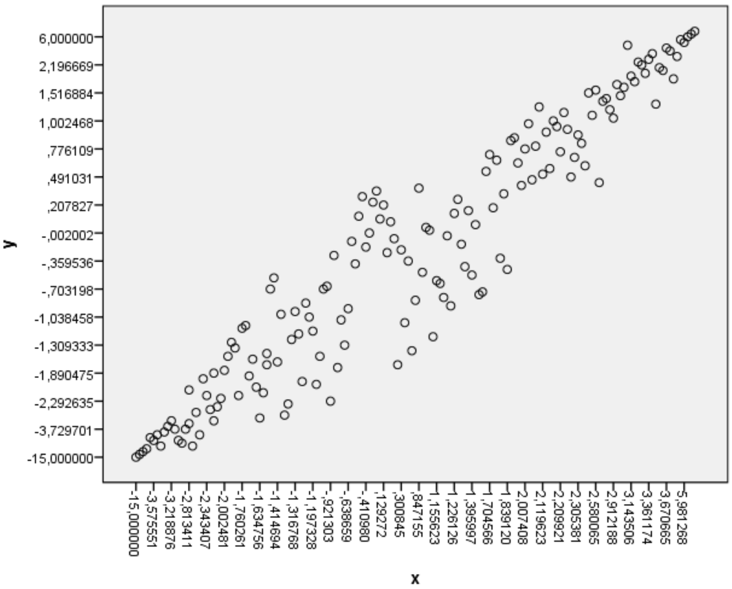
**Tablo 13:** Uygulama 3(a) Verisi için Doğrusal ANCOVA Sonuçları

| Değişkenlik Kaynağı | Serbestlik Derecesi | Kareler Toplamı | Kareler Ortalaması | F Hesap Değeri |
|---|---|---|---|---|
| *Denemeler* | 1 | 38,446 | 38,446 | 142,454 |
| *Hata* | 159 | 42,911 | 0,27 | |
| *Genel* | 160 | 81,357 | | |

$F_{tablo} = F_{0,05;1;159}$ değerinin 3,90061 olduğu göz önüne alındığında; hesaplanan F değerinin anlamlı olduğu, deney ve kontrol grupları arasında farkın olduğu görülmektedir. Bu doğrusal ANCOVA modeli, verileri temsil edecek şekilde kabul edilir.

### 5.3.4. Uygulama 3(b): Veriler Tam Doğrusal İken Zorunlu Sigortalı Sayısı ve İsteğe Bağlı Sigortalı Sayısının Cinsiyete Etkisi

Bir önceki uygulamada kullanılan verilere tam doğrusallığın elde edilebilmesi için logaritmik dönüşüm işlemi uygulanmıştır. Veri grafiğinde tam doğrusallığı bozan aykırı gözlemler çıkarılmış yerine tam doğrusallığı sağlayan verilerin %4,9'u kadar yeni gözlemler eklenmiştir. Aşağıda verilen yayılım grafiğinde verinin tam doğrusal olduğu görülmektedir.




Uygulama 3(b) hesaplanan verisi için doğrusal ANCOVA sonuçları Tablo 14'de verilmiştir:

**Tablo 14:** Uygulama 3(b) Verisi için Doğrusal ANCOVA Sonuçları

| **Değişkenlik Kaynağı** | **Serbestlik Derecesi** | **Kareler Toplamı** | **Kareler Ortalaması** | **F Hesap Değeri** |
|---|---|---|---|---|
| *Denemeler* | 1 | 43,608 | 43,608 | 137,643 |
| *Hata* | 159 | 50,374 | 0,317 | |
| *Genel* | 160 | 93,982 | | |

$F_{tablo} = F_{0,05;1;159}$ değerinin 3,90061 olduğu göz önüne alındığında; hesaplanan F değerinin anlamlı olduğu, deney ve kontrol grupları arasında farkın olduğu görülmektedir. Bu doğrusal ANCOVA modeli, verileri temsil edecek şekilde kabul edilir.

Modele regresyon analizi uygulandığında regresyon katsayıları $\beta_0 = 108{,}331;\ \beta_1 = 43{,}608;\ \beta_2 = 751{,}989$ olarak bulunmuştur.

1.gruba ait ortalama gölge değişken skoru 1, ortalama ortak değişken skoru 0,316 olarak hesaplanmıştır. Böylece 1. gruba ait düzeltilmiş deneme etkisi

$$\overline{Y}_1(d\ddot{u}z) = 108{,}331 + 43{,}608\,(1) + 751{,}989\,(0{,}486) = 389{,}778$$

olarak bulunmuştur.

Benzer şekilde 2.gruba ait ortalama gölge skoru 0, ortalama ortak değişken skoru 0,316 olarak hesaplanmıştır. Böylece 2. gruba ait düzeltilmiş deneme etkisi

$$\overline{Y}_2(d\ddot{u}z) = 108{,}331 + 43{,}608\,(0) + 751{,}989\,(0{,}486) = 346{,}170$$

olarak bulunmuştur.

### 5.3.5. Uygulama 3(c): Verilerin Doğrusal Olmadığı Varsayıldığında Zorunlu Sigortalı Sayısı ve İsteğe Bağlı Sigortalı Sayısının Cinsiyete Etkisi

Uygulama 3(a) ve Uygulama 3(b) ile yapılan analizlerde veriler doğrusal olduğundan doğrusal ANCOVA yöntemini kullanılmıştır. Uygulama 3(c) ile de verilerin doğrusal olmadığı varsayılarak kuadratik yani doğrusal olmayan ANCOVA analizi uygulanmıştır.

Uygulama 3(b) hesaplanan verisi için doğrusal ANCOVA sonuçları Tablo 15'de verilmiştir:

**Tablo 15:** Uygulama 3(c) Verisi için Kuadratik ANCOVA Sonuçları

| **Değişkenlik Kaynağı** | **Serbestlik Derecesi** | **Kareler Toplamı** | **Kareler Ortalaması** | **F Hesap Değeri** |
|---|---|---|---|---|
| *Düzeltilmiş Değerlendirme* | 2 | 0,577 | 0,288 | 1,083 |
| *Kuadratik Hata*$_w$ | 156 | 41,602 | 0,266 | |
| *Kuadratik Hata*$_t$ | 158 | 42,18 | | |

$F_{tablo} = F_{0{,}05;2;156}$ değerinin 3,054 olduğu göz önüne alındığında; hesaplanan F değerinin anlamlı olmadığı, deney ve kontrol grupları arasında farkın olmadığı görülmektedir. Bu doğrusal ANCOVA modeli, verileri temsil edecek şekilde kabul edilemez.

Aynı veriler kullanılarak yapılan üç farklı analizden verilerin doğrusal olup olmama durumları ile veriler tam doğrusal hale getirildikten sonraki durumlarının sonuçlar üzerindeki etkisi yapılan uygulamalarla desteklenmiştir. Uygulama 2 ve 3'de kullanılan verilerin doğrusal ve tam doğrusal olduğu durumlarda yapılan analizlerin ikisinde de doğrusal ANCOVA analizleri ile oluşturulan modeller anlamlı çıkmış, Kuadratik ANCOVA modelleri anlamsız bulunmuştur. Yapılan üç doğrusal ANCOVA analizlerinden F değeri daha küçük olarak bulunan tam doğrusal veri üzerinden oluşturulan model daha anlamlı bulunmuştur.

Sonuç olarak Uygulama 3 için yani tam doğrusala yakın hale getirilmiş olan verileri temsil edecek en uygun modelin; $\overline{Y}(düz) = 108{,}331 + 43{,}608\,d + 751{,}989\,x$ olduğu, Uygulama 2 ve Uygulama 3 birlikte değerlendirildiğinde ise verileri temsil edecek en uygun modelin yine; $\overline{Y}(düz) = 108{,}331 + 43{,}608\,d + 751{,}989\,x$ olduğu belirlenmiştir.

# SONUÇ VE TARTIŞMA

Kovaryans analizi, ancak varsayımların sağlanması durumunda güçlü ve yararlı bir analiz olmaktadır. Hata varyansını en aza indirgemek modelin gücünü artırır. Ayrıca kovaryans analizi, küçük örneklemlere ya da etki büyüklüğünün küçük olduğu durumlarda uygulandığında daha anlamlı sonuçlar vermektedir.

Geleneksel ANCOVA modeli varsayımlarındaki bağımlı değişken ve bağımsız değişkenler arasındaki ilişki her zaman doğrusal olmamaktadır. Yapılan analizlerde çok değişkenli kovaryans analizinin temel varsayımlarından olan doğrusallık varsayımının sağlanmaması problem teşkil etmektedir. Eğer doğrusal olmayan veriler üzerinde çok değişkenli kovaryans analizi kullanılıyorsa, kovaryans analizi modelinin temsil etme gücü zayıflamaktadır. Bu yüzden de çok değişkenli kovaryans analizi modellerinde doğrusallık şartının sağlanması önemlidir. Şiddetli doğrusal olmama durumu, grup içi *XY* yayılım grafiği ile kolaylıkla ortaya çıkarılabilecektir.

Eğer ilişki doğrusal değil fakat monoton ise basit dönüşümlere (genellikle *X*'e uygulanacak dönüşümler) *Y* ve dönüştürülmüş *X* arasında doğrusal ilişki bulunabilecektir. Değişkenlere dönüşüm uygulandıktan sonra verilere kovaryans analizi uygulanabilecektir.

Modelin doğrusal olmaması durumunda kovaryans analizi uygulanabilmesi için modelin doğrusal modele dönüştürülmesi gereklidir. Eğer bağımlı değişken ve bağımsız değişken arasındaki ilişki hem doğrusal olmayan hem de monoton değilse - yani bağımlı değişken artarken bağımsız değişken yalnızca bir noktaya kadar artıyor ve sonra bağımlı değişken artarken bağımsız değişken de artıyor ise- basit dönüşüm yaklaşımı da yeterli olmayacaktır. Bu durumda bağımsız değişkene polinomial yaklaşım uygulanacaktır. Genellikle bu durumda, kuadratik veya kübik Ancova modelleri kullanılmalıdır. Karmaşık polinom modellerin kullanımı sadece daha basit modellerse eğer yeterli olabilir. Aksi halde polinom modeller de yetersiz olacaktır. Daha basit olan modeller tercih edilir çünkü karmaşık modellere dayalı sonuçlar daha zor yorumlanır ve genel olarak daha az kararlıdır. Polinom ANCOVA modelleri açıkça uygulanabildiğinde hesaplamalar çoklu ANCOVA'nın bir uzantısını içerir.

Yapılan uygulamalarda Sosyal Güvenlik Kapsamında Kişi Sayısı ve Türkiye Nüfusuna Oranı Tablosu adı altında Sosyal Güvenlik Kapsamında Aktif Çalışan Kişi Sayısı ile Bakmakla Yükümlü Tutulanların Sayısının 4/a ve 4/b li Çalışan Sigortalı Sayıları; 4/b Kapsamındaki Aktif Sigortalıların İl ve Cinsiyet Tablosu adı altında 4/b Kapsamındaki Aktif Sigortalılardan Zorunlu Sigortalı Sayısı ve İsteğe Bağlı Sigortalı Sayısının Cinsiyet Dağılımı verileri kullanılmıştır.

Uygulamalardaki amaç Sosyal Güvenlik Kurumundan alınan verilerle doğrusal ve doğrusal olmayan veriler için en uygun kovaryans analizi yöntemini belirlemektir.

Üç farklı veri yapısı için yapılan uygulama sonuçları Tablo 16'da verilmiştir:

**Tablo 16:** Uygulama Verileri İçin Karşılaştırmalı ANCOVA Sonuçları

| Uygulama | Verinin Doğrusallık Durumu | Uygulanan Model | Hesaplanan F Değeri | F Tablo Değeri | Serbestlik Derecesi |
|---|---|---|---|---|---|
| Uygulama 1(a) | Veri doğrusal değil iken | Kuadratik Ancova | 25,443 | 3,054 | ( 2 , 156 ) |
| Uygulama 1(b) | Veri doğrusallıktan uzaklaştıkça | Kuadratik Ancova | 27,409 | 3,051 | ( 2 , 162 ) |
| Uygulama 1(c) | Verinin doğrusal olduğu varsayıldığında | Doğrusal Ancova | 38,673 | 3,9 | ( 1 , 159 ) |
| Uygulama 2(a) | Veri doğrusal iken | Doğrusal Ancova | 8,648 | 3,9 | ( 1 , 159 ) |
| Uygulama 2(b) | Verinin Doğrusal olmadığı varsayıldığında | Kuadratik Ancova | 1,733 | 3,054 | ( 2 , 156 ) |
| Uygulama 3(a) | Veri doğrusal yapıya yaklaştıkça | Doğrusal Ancova | 142,454 | 3,9 | ( 1 , 159 ) |
| Uygulama 3(b) | Veri tam doğrusal hale getirildiğinde | Doğrusal Ancova | 137,643 | 3,9 | ( 1 , 159 ) |
| Uygulama 3(c) | Verinin doğrusal olmadığı varsayıldığında | Kuadratik Ancova | 1,083 | 3,054 | ( 2 , 156 ) |

Uygulama 1(a), Uygulama 1(b) ve Uygulama 1(c) ile hesaplanan F değerleri karşılaştırıldığında, ortak değişken ile bağımlı değişken arasındaki ilişki düşük olduğunda, kuadratik ANCOVA modelinin F değeri daha güçlü sonuçlar vermiştir. Yani doğrusal olmayan veriye uygulanacak en uygun modelin kuadratik ANCOVA modeli olduğu yapılan uygulamalarla gösterilmiştir.

Uygulama 2(a) ve Uygulama 2(b) ile hesaplanan F değerleri karşılaştırıldığında, doğrusal veriye kuadratik ANCOVA uygulanması modeli anlamsız hale getirdiğinden doğrusal veriye uygulanacak en uygun modelin doğrusal ANCOVA modeli olduğu yapılan uygulamalarla gösterilmiştir.

Uygulama 3(a), Uygulama 3(b) ve Uygulama 3(c) ile hesaplanan F değerleri karşılaştırıldığında, tam doğrusal veriye kuadratik ANCOVA uygulanması modeli anlamsız hale getirirken aynı veriye doğrusal ANCOVA uygulanması modeli anlamlı hale getirmektedir. Bu durumda tam doğrusal veriye uygulanacak en uygun modelin yine doğrusal ANCOVA modeli olduğu yapılan uygulamalarla gösterilmiştir.

Sonuç olarak, doğrusal olmayan veriler için uygulanacak en uygun modelin kuadratik ANCOVA modeli olduğu, doğrusal olduğu kabul edilen veriler için de doğrusal ANCOVA modeli olduğu değerlendirilmiştir. Ayrıca veri ne kadar doğrusal ise doğrusal ANCOVA modelinin F değerinin daha güçlü sonuçlar verdiği, verinin doğrusal olmadığı durumlarda da kuadratik ANCOVA modelinin F değerinin daha güçlü sonuçlar verdiği değerlendirilmiştir.

Her ne kadar doğrusal olmayan veriler için kuadratik ANCOVA modelinin daha uygun olsa da, analiz verileri için daha az karmaşık modelin tercih edilmesi daha sağlıklı sonuçlar verecektir. Eğer doğrusal olmayan bir veri için doğrusal model de uygunluk sağlıyorsa, daha az karmaşık model olan doğrusal modelin tercih edilmesi gerekmektedir.

Yapılacak olan çalışmalarda doğrusal olmayan veriler için doğrusal olmayan ANCOVA analizinin doğruluğunun tespiti için daha büyük gözlem verileri kullanılarak simülasyon teknikleri kullanılabilir.

# KAYNAKÇA


Adıgüzel Ahmet (2010). *5510 Sayılı Sosyal Sigortalar ve Genel Sağlık Sigortası Kanununda Kısa ve Uzun Vadeli Sigorta Kollarında Yapılan Yenilikler, Sorunlar ve Çözüm Önerileri.* Tokat: Sosyal Bilimler Enstitüsü, Yayınlanmamış Yüksek Lisans Tezi.

Akçay Aytaç, Uğurlu Mustafa, Yakan Akın ve Atasoy Fatih (2012) "*Etçil Piliçlerde Cinsiyetin ve Protein Kısıtlanmasının Abdominal Yağ Birikimi Üzerine Etkisinin Kovaryans Analizi ile İncelenmesi*", Veterinerlik Fakültesi Dergisi, 9/2: (2012), 107-112.

Akdeniz Fikri ve Öztürk Fikri (1996). *Lineer Modeller,* Ankara: A.Ü.F.F. Döner Sermaye İşletmesi Yayınları.

Akgül Aziz (2005). *Tıbbi Araştırmalarda İstatistiksel Analiz Teknikleri*. Ankara: Emek Ofset.

Akıcı Ceren (2013). *Jersey Irkı Süt İneklerinde Laktasyon Süt Verimi Üzerinde Etkili Olan Faktörlerin Kovaryans Analizi İle İncelenmesi: Samsun Karaköy Tarım İşletmesi Müdürlüğü Örneği*. Samsun: Fen Bilimleri Enstitüsü, Yayınlanmamış Yüksek Lisans Tezi.

Burgazoğlu Hüseyin (2013). *Çok Değişkenli Kovaryans Analizi ve 360 Derece Performans Değerlendirmesi Üzerine Bir Uygulama*. İstanbul: Sosyal Bilimler Enstitüsü, Yayınlanmamış Yüksek Lisans Tezi.

Büyüköztürk Şükrü (1997). İki Faktörlü Varyans Analizi, *Eğitim Bilimleri Fakültesi Dergisi*. Ankara: 30, 141-185.

Büyüköztürk Şükrü (2001). *Kovaryans Analizi (Varyans Analizi ile Karşılaştırmalı Bir İnceleme)*. Ankara: Sosyal Bilimler Enstitüsü, Yayınlanmamış Yüksek Lisans Tezi.

Cantay Tülin (2005). *Tesadüfi Bloklar Düzeninde Kovaryans Analizi ve Uygulaması.* Konya: Fen Bilimleri Enstitüsü, Yayınlanmamış Yüksek Lisans Tezi.

Cochran William G. (1957). *Analysis of Covariance: Its Nature and Uses*, Biometrics. New York.

Çömlekçi Necla (1994). *Temel İstatistik*. İstanbul: Bilim Teknik Yayınevi.

Durmuş Alpaslan. (2012). *Sanal Bilim ve Teknoloji Müzesinde Eğitsel Arayüz Ajanı Kullanımının Öğrencilerin İlgi ve Başarılarına Etkisi.* Ankara: Eğitim Bilimleri Enstitüsü, Yayınlanmamış Doktora Tezi.

Eden Thomas ve Fisher Ronald Aylmer (1927). "Studies in Crop Variation. Iv. The Experimental Determination of The Value of Top Dressings With Cereals", *Journal of Agricultural Science*. 17/4.

Efe Ercan, Bek Yüksel, Şahin Mustafa (2000) *Spss'te Çözümleri ile İstatistiksel Yöntemler-2*, Kahramanmaraş: Sütçü İmam Üniversitesi, Yayın No:10.

Fisher Ronald Aylmer (1932). *Statistical Methods for Research Workers*. (4th Edition), Edinburgh: Oliver & Boyd.

Fisher Ronald Aylmer (1925). *Statistical Methods for Research Workers*. Edinburgh Oliver & Boyd.

Fisher Ronald Aylmer (1926). *The Arrangement of Field Experiments*. England: J. Ministry Agric.

Hair Joseph F., Anderson Rolph E., Tatham R. L., Black William C. (1995). *Multivariate Data Analysis with Readings*. Amerika: Prentice Hall.

Hicks C.R. (1994). *Deney Düzenlemede İstatistiksel Yöntemler*, Çev.: Muluk Zehra, Toktamış Öniz, Kurt Serdar, Karaagaoglu Ergun, İzmir: Ege Üniversitesi Basımevi.

Hinks Joanna C. ve Crassidis John (2013). "Covariance Analysis of Maximum Likelihood Attitude Estimation", *Journal of The Astronautical Sciences* 60/2

Howitt Dennis ve Cramer Duncan (1997) *An Introduction to Statistics in Phychology: A Complete Guide for Students*. London: Prentice Hall.

Huang Haibing ve Zhang Jin (2014). A Nonlinear Covariance Analysis Method for Space Station Rendezvous Phasing Based on Unscented Transformation, *Journal of National University of Defense Technology* 36/2: (2014), 18-23.

Huitema Bradley E. (2011). *The Analysis of Covariance and Alternatives.* New Jersey: A. John Wiley and Sons, Inc., Publication.

Jamali Nawid, Kormushev Petar, Ahmadzadeh Reza ve Caldwell Darwin G. (2014). *Covariance Analysis as a Measure of Policy Robustness*, Ieee-Xplore Oceans.

Jamiesson John (2004). "Analysis of Covariance (Ancova) With Difference Scores", *International Journal of Psychophysiology,* 52: (2004), 277-283.

Koch Gary G., Amara Ingrid A., Davis Gordon W. ve Gillings Dennis B. (1982). A *Review of Some Statistical Methods for Covariance Analysis of Categorical Data*.

Lattin James M., Carroll Douglas, Green Paul (2003). *Analyzing Multivariate Data*, Kanada.

Lindman Harold R. (1992) *Analysis of Variance in Experimental Design,* New York: Springer-Verlag.

Owen Steven V., Froman Robin D. (1981). "Uses and Abuses of The Analysis of Covariance", *Bureau of Educational Research,* John Wiley&Sons., Connecticut, 21: (1981), 557-562.

Özer Hüseyin ve Sarı Ahmet (2009) "Kovaryans Analizi Modelleriyle Üniversite Öğrencilerinin Başarılarını Etkileyen Faktörlerin Belirlenmesi: Atatürk Üniversitesi İİBF Öğrencileri İçin Bir Uygulama", *Dokuz Eylül Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 24/2: (2009), 105-126.

Özgören Berrin (2011) *Değerler Ölçeği ve Kovaryans Analizi.* Konya: Fen Bilimleri Enstitüsü, Yayınlanmamış Yüksek Lisans Tezi.

Quade Dana (1967) *Rank Analysis of Covariance. Journal of The American Statistical Association 62.*

Rencher Alvin (2001). *Lineer Models in Statistics.* New York: John Wıley & Sons.

Serper Özer (2000) *Uygulamalı İstatistik 2,* Bursa: Ezgi Kitabevi.

Şahin Hilal (2006). *Kovaryans Analizi ve Bir Uygulama*. Ankara: Fen Bilimleri Enstitüsü, Yayınlanmamış Yüksek Lisans Tezi.

Şenoğlu Birdal ve Acıtaş Şenol (2010). *İstatistiksel Deneysel Tasarımı: Sabit Etkili Modeller*, Ankara: Nobel Yayın Dağıtım.

Şenoğlu Birdal ve Acıtaş Şenol (2011) *İstatistiksel Deney Tasarımı*, Ankara: Nobel Yayınları (2. Basım).

Weber Donald. C. ve Skıllıngs John H. (2000). *A First Course in The Design of Experiments: A Linear Models Approach.* New York: Crc Press Llc.

Yazıcı Berna (2001). *Kategorik Veri Analizinde Eş Değişken Bulunması Durumunda Genelleştirilmiş Tahmin Denklemleri Yaklaşımı ve Bir Uygulama.* Eskişehir: Fen Bilimleri Enstitüsü, Yayınlanmamış Doktora Tezi.

**İnternet Kaynakları**

Sosyal Güvenlik Kurumu resmi web sitesi https://www.sgk.gov.tr

# ÖZGEÇMİŞ

**KİŞİSEL BİLGİLER**

| | | |
|---|---|---|
| Adı Soyadı | : | Cansu MERCAN |
| Uyruğu | : | T.C. |
| Doğum Tarihi ve Yeri | : | 05.01.1991 / Sivas |
| e-posta | : | cansuozkan08@hotmail.com |

**EĞİTİM**

| **Derece** | **Kurum** | **Mezuniyet Yılı** |
|---|---|---|
| Lisans | Selçuk Üniversitesi / İstatistik Bölümü | 2012 |
| Yüksek Lisans | Cumhuriyet Üniversitesi | |

**İŞ TECRÜBESİ**

| **Tarih** | **Kurum** | **Görev** |
|---|---|---|
| 10.02.2014 | Sosyal Güvenlik Kurumu | Sosyal Güvenlik Denetmeni |

**YABANCI DİL BİLGİSİ**

| **Yabancı Dilin Adı** | **KPDS (   )** | **ÜDS (   )** | **TOEFL (   )** | **EILTS (   )** |
|---|---|---|---|---|