

**GELİR VE HANEHALKI KİŞİ SAYISIYLA  
ET ve SEBZE TÜKETİMİ ARASINDAKİ İLİŞKİNİN  
ÇOK DEĞİŞKENLİ REGRESYON  
ANALİZİYLE BELİRLENMESİ**

**BERNA BALOĞLU**  
**Yüksek Lisans Tezi**

**İstatistik Anabilim Dalı**  
**EYLÜL-1996**

**ÖZET**  
*Yüksek Lisans Tezi*

**GELİR ve HANEHALKI KİŞİ SAYISIYLA ET ve SEBZE °  
TÜKETİMİ ARASINDAKİ İLİŞKİNİN ÇOK DEĞİŞKENLİ  
REGRESYON ANALİZİYLE İNCELENMESİ**

**BERNA BALOĞLU**

*Anadolu Üniversitesi  
Fen Bilimleri Enstitüsü  
Uygulamalı İstatistik Anabilim Dalı*

*Danışman: Doç. Dr. Embiya AĞAOĞLU  
1996*

*Bu çalışmada, iki veya daha fazla bağımsız değişkenin birden fazla sayıdaki bağımlı değişkeni etkilediği durumlarda kullanılan Çok Değişkenli Regresyon Analizi üzerinde durulmuştur. Analizde denklemlerin kurulması, katsayıların anlamlılığının sınanması, model uygunluğunun test edilmesi, güven sınırlarının tahmin edilmesi ele alınmıştır.*

*Araştırmanın uygulama bölümünde, hanehalkı kişi sayısı ve aylık geliri ile hanede bir ay boyunca tüketilen et ve sebze miktarı arasındaki ilişki ortaya konulmaya çalışılmıştır.*

*Anahtar Kelimeler: Çok değişkenli regresyon, Bağımlı, Bağımsız değişken, Hanehalkı harcamaları*

**ABSTRACT**  
*Master of Science Thesis*

**EXAMINING THE RELATIONSHIP BETWEEN INCOME, HOUSEHOLD  
POPULATION AND THE CONSUMPTION OF MEAT, VEGETABLE WITH  
MULTIVARIATE REGRESSION ANALYSIS**

**BERNA BALOĞLU**

*Anadolu University  
Graduate School of Natural and Applied Sciences  
Applied Statistics Program*

*Supervisor: Assoc. Prof. Embiya AĞAOĞLU  
1996*

*In this study, multivariate regression analysis has been studied for the situations in which two or more independent variables effect more than one dependent variable. During the analysis, evaluating the equations, testing the significance of constants, computing the confidence limits and testing the model fit have also been handled.*

*In the application part of this research, the relationship between household population, income and consumption of meat and vegetable during one month has been considered.*

**Key Words:** *Multivariate regression, Dependent, Independent variable, Household consumptions*

## **TEŞEKKÜR**

Bu çalışmanın her aşamasında ilgi ve yardımlarını esirgemeyen danışman hocam Anadolu Üniversitesi Fen Fakültesi öğretim üyesi Doç. Dr. Embiya AĞAOĞLU'na en içten teşekkürlerimi bir borç bilirim.

Çalışmalarım sırasında bana yardımcı olan D.İ.E. Yayın Haberleşme ve Halkla İlişkiler Bölümü çalışanlarına ve beni yönlendiren tüm hocalarıma en içten teşekkürlerimi sunarım.

## *İÇİNDEKİLER*

### Sayfa

ÖZET .....	i
ABSTRACT .....	ii
İÇİNDEKİLER .....	iv
ŞEKİLLER DİZİNİ .....	vii
TABLolar DİZİNİ .....	viii
SİMGELER VE KISALTMALAR DİZİNİ .....	ix
1. GİRİŞ .....	1
2. KONUYA İLİŞKİN TANIM VE KAVRAMLAR .....	3
2.1. Hanehalkı .....	3
2.2. İhtiyaç .....	3
2.3. Tüketim ve Tüketici Davranışları .....	4
2.4. Gelir .....	5
3. ÇOK DEĞİŞKENLİ REGRESYON ANALİZİ .....	8
3.1. Çok Değişkenli Regresyon Modelinin Matematiksel Gösterimi .....	9
3.2. Çok Değişkenli Regresyon Modelinin Varsayımları .....	12
3.3. Çok Değişkenli Regresyon Denkleminin Elde Edilmesinde Kullanılan Teknikler .....	13
3.3.1. En Küçük Kareler Tekniği .....	13
3.3.2. En Çok Olabilirlik Tekniği .....	15
3.4. Wishart Dağılımı .....	17
3.5. Regresyon Katsayılarının Anlamlılık Sınamaları .....	18
3.5.1. Wilks Lambda Olabilirlik Oran Sınaması .....	18

## İÇİNDEKİLER (DEVAM)

### Sayfa

3.5.2. Varyans Analizi Tekniđi .....	20
3.5.3. Parsimony Kavramına Dayalı Varyans Analizi Tekniđi...	21
3.6. Regresyon Katsayıları İin Gven Sınırlarının Belirlenmesi .....	23
3.7. Artıklarla Model Uygunluđunun İncelenmesi .....	24
3.7.1. Artık Deđerlerinin Tahmin Deđerlerine Gre izimi .....	25
3.7.2. Artık Deđerlerinin Bađımsız Deđeriskene Gre izimi .....	27
4. GELİR VE HANEHALKI KİŐİ SAYISIYLA ET VE SEBZE TKETİMİ ARASINDAKİ İLİŐKİNİN OK DEđerİŐKENLİ REGRESYON ANALİZİYLE BELİRLENMESİ .....	28
* 4.1. Problemin Tanımlanması, Veri Kaynađı ve İncelenecek Deđeriskenerin Belirlenmesi .....	28
4.1.1. Uygulama Alanı .....	28
4.1.2. Veri Kaynađı .....	28
4.1.3. Deđeriskenerin Belirlenmesi .....	30
* 4.2. Gelir ve Hanedeki KiŐi Sayısının Et ve Sebze Tketime Etkisinin ok DeđerİŐkenli Regresyon Tekniđi ile Belirlenmesi .....	31
✓ 4.2.1. Katsayıların Tahmini .....	31
✓ 4.2.2. Hesaplanan Regresyon Katsayılarının Anlamlılık Sınamaları .....	33
4.2.2.1. Hesaplanan Katsayıların Wilks Olabilirlik Oran Testi İle Sınanması .....	34
4.2.2.2. Hesaplanan Katsayıların Varyans Analizi Tekniđi İle Sınanması .....	35
4.2.2.3. Hesaplanan Katsayıların Parsimony Kavramına Dayalı Varyans Analizi Tekniđi İle Sınanması.....	37
✓ 4.2.3. Katsayılar İin Gven Sınırlarının Belirlenmesi .....	41

## *İÇİNDEKİLER (DEVAM)*

### Sayfa

4.2.4. Hesaplanan Artıklarla Model Uygunluğunun İncelenmesi ....	44
5. SONUÇ VE ÖNERİLER .....	49
6. KAYNAKLAR .....	50
7. EKLER .....	53
<b>EKLER</b>	
Ek-1 Uygulama Çalışmasında Kullanılan Veriler .....	53
Ek-2 X Veri Matrisi .....	54
Ek-3 Y Veri Matrisi .....	55
Ek-4 $\hat{Y}$ Tahmin Matrisi .....	56
Ek-5 $\hat{E}$ Hata Matrisi .....	57
Ek-6 İndirgenmiş Veri Matrisi .....	58
Ek-7 İndirgenmiş Modele İlişkin Tamin Matrisi .....	59
Ek-8 İndirgenmiş Model İçin Hesaplanmış Artıklar .....	60
Ek-9 İndirgenmiş Veri Matrisi .....	61
Ek-10 İndirgenmiş Modele İlişkin Tahmin Matrisi .....	62
Ek-11 İndirgenmiş Model İçin Hesaplanmış Artıklar .....	63

**ŞEKİLLER DİZİNİ**

3.1. Artıkların tahmin değerlerine göre çizimi	25
3.2. Artıkların tahmin değerlerine göre çizimi	25
3.3. Artıkların tahmin değerlerine göre çizimi	26
3.4. Artık değerlerinin bağımsız değişkene göre çizimi	27
4.1. Tahminlerin hata terimlerine göre çizimi	44
4.2. Tahminlerin hata terimlerine göre çizimi	45
4.3. Artık değerlerinin bağımsız değişken değerlerine karşı çizimi	46
4.4. Artık değerlerinin bağımsız değişken değerlerine karşı çizimi	46
4.5. Artık değerlerinin bağımsız değişken değerlerine karşı çizimi	47
4.6. Artık değerlerinin bağımsız değişken değerlerine karşı çizimi	48



***TABLULAR DİZİNİ***

3.1. Çok deęişkenli regresyon analizinde varyans analizi tablosu	20
3.2. İndirgenmiş model için varyans analizi tablosu	22
4.1. Çok deęişkenli regresyonda varyans analizi tablosu	36
4.2. İndirgenmiş model için varyans analizi tablosu	38
4.3. İndirgenmiş model için varyans analizi tablosu	40

***SİMGELER ve KISALTMALAR DİZİNİ***

$\Sigma$ : Varyans-kovaryans matrisi

$S$ : Örneklem için hesaplanmış varyans-kovaryans matrisi

$\Lambda$ : Wilks Lambda istatistiği

$X$ : Veri matrisi

$\hat{Y}$ : Bağımlı değişken için tahmin matrisi

$\hat{E}$ : Artık değerlerinin tahminleri

$\tilde{X}$ : İndirgenmiş model için veri matrisi

$\tilde{Y}$ : İndirgenmiş model için bağımlı değişken tahmin matrisi

$\tilde{E}$ : İndirgenmiş model için hata terimleri matrisi

D.İ.E. : Devlet İstatistik Enstitüsü

ANOVA: Analysis of Variance

## 1. GİRİŞ

Günümüzde dünya ekonomisinin seyri; istihdam, işsizlik, geçim sıkıntısı gibi pek çok sorunu ortaya çıkarmıştır. Ülkeler bu sorunlara çözüm bulmak amacıyla uygun ekonomi politikaları geliştirmek için konuya ilişkin araştırma çalışmalarına daha fazla yer vermektedir [1].

Son yıllarda ülkemizin karşı karşıya kaldığı sosyal sorunların başında hızlı fiyat artışları, bunun yol açtığı hayat pahalılığı ve geçim sıkıntısı gelmektedir. Özellikle sınırlı gelirle geçinmek zorunda kalan tüketiciler bu hızlı fiyat artışları sonucunda zor durumda kalmaktadır.

Eğer kişilerin gelirleri, zorunlu ihtiyaçlarını ancak karşılayabilecek kadarsa bu kişilerin kıt kanaat geçindiği söylenir. Kişilerin gelirleri zorunlu masraflarını karşılamadan yanısıra lüks mal almalarına yetiyor ve bir parça tasarruf yapmalarını sağlıyorsa kişilerin orta gelirli olduğu söylenebilir. Diğer taraftan kişiler ihtiyaç duydukları bir takım mallardan vazgeçmek zorunda kalıyorsa geçim sıkıntısı çektikleri söylenebilir [2].

Türkiye’de gözlenen fiyat artışları beraberinde sosyal sorunların yanısıra ekonomideki değişimi de getirmektedir. Bu durumdan etkilenen tüketici kitlesinin bazı ürünlerden vazgeçmesi, bazı ürünleri daha az tüketir hale gelmesi hem pazar hem tüketici açısından bu değişimin basit göstergelerindedir. Bu ekonomik değişim hanehalkının mutfak harcamalarını da etkileyebilmektedir. Harcamalar gelirle kısıtlı tutulur durumdadır. Ayrıca kişi sayısı yine mutfak harcamalarında etkili olabilmektedir.

Buradan hareketle, hanehalkının et ve sebze ürünlerini gelirlerine bağlı olarak ne miktarlarda tükettikleri, çalışmanın uygulama konusu olarak belirlenmiştir. Çalışmada, hem et hem de sebze tüketimiyle ilgileniliyor olması 2 tane bağımlı değişkenin varlığını göstermektedir. 2 bağımlı değişken en az 2 bağımsız değişkenle açıklanabilmektedir. Birden fazla bağımlı değişken bulunan regresyon analizi ise bu çalışmada kullanılan “Çok değişkenli regresyon analizi”dir.

Çok deęişkenli regresyon analizinin incelenmesi ve hanehalkının temel besinleri et ve sebze tüketimi ile gelir ve hanehalkı kiři sayısı arasındaki ilişkinin gösterilmesinin amaçlandığı bu çalışma 5 bölümden oluşmaktadır.

İlk bölüm giriş bölümüdür.

İkinci bölümde konuya ilişkin kavram ve tanımlara yer verilecektir. Ayrıca bu konularda daha önceden yapılmış bazı anket ve araştırma bulgularına da değinilecektir.

Üçüncü bölümde çok deęişkenli regresyon analizinin teorik yapısına yer verilecektir. Bu bölümde sözkonusu analizin tanımı, matematiksel yapısı, modelin varsayımları, analiz için kullanılacak teknikler ve model uygunluęunun test edilmesi üzerinde durulacaktır.

Dördüncü bölümde hanehalkı et ve sebze tüketimine gelir ve kiři sayısı deęişkenlerinin etkisinin belirlenmesi amaçlanmaktadır. Bu konuda gerçekleştirilen uygulama çalışması bu bölümde yer almaktadır.

Son bölüm olan beşinci bölüm sonuç ve önerilerin yer aldığı bölümdür.

## **2. KONUYA İLİŞKİN TANIM VE KAVRAMLAR**

### **2.1. Hanehalkı**

Bir ülkedeki gelir dağılımı, tüketim faaliyetleri, sosyo-ekonomik yapı genel olarak incelendiğinde gözlenen en küçük topluluk hanehalkı olmaktadır. Aralarında akrabalık bağı bulunsun ya da bulunmasın aynı evde oturan veya aynı evin bir kısmında yaşayan, gelir ve giderlerini ayırmayan, hanehalkı hizmet ve yönetimine katılan bir veya birkaç kişinin oluşturduğu topluluğa hanehalkı denilmektedir [3].

Buna göre hanehalkı bir veya daha fazla kişinin oluşturduğu, akrabalık bağı olan veya olmayan hatta bazı durumlarda sadece ekonomik bir bağı olan bir topluluktur [3].

### **2.2. İhtiyaç**

Hanehalkı bütün olarak sosyo-ekonomik açıdan incelenmek istenirse, ilk olarak hanehalkının ihtiyaçları gözönünde bulundurulur. Zaten tüketim olayının temelinde yatan ihtiyaç, insanların bazı temel doygunluklardan yoksun olduğunu ve bu eksikliği giderme arzusunu duymaları durumudur [4].

Bu noktada ihtiyaçların önem sırası akla gelmektedir. İnsanın yaşamını sürdürebilmesi için kesin olarak gidermek zorunda olduğu ihtiyaçlar beslenme ve barınma ihtiyaçlarıdır. Bu konuda en ileri teori Maslow'un ihtiyaç teorisidir. Bu teoriye göre, fizyolojik ihtiyaçlar olarak adlandırılan bu tür zorunlu ihtiyaçlar karşılanmadan önemi daha az olan diğer ihtiyaçların karşılanması düşünülmeyecektir. Bir başka ifadeyle ihtiyaçlar bir önem sırası izler ve bir üst basamaktaki ihtiyaç karşılanmadan diğerine geçilemez [5].

Farklı ekonomik düzeydeki fertlerin ihtiyalarının nem sırası da farklı olacaktır. Hatta gelir arttıka ya da azaldıka ihtiyalar da kiřiden kiřiye deęiřmeye başlayacaktır. Ekonomik düzey azaldıka ihtiyalar artık ihtiya deęil lks durumuna dnřrler. Bunun tersi olarak ekonomik düzey ykseldike lks diye nitelenen mallar ihtiya olarak dřnlr.

İhtiyalar yerleřim biimine gre deęiřmektedir. Byk řehirlerde yařayan fertlerle kırsal blgelerde yařayan fertlerin ihtiyaları ok farklı olacaktır. Kk yerleřim birimlerinde otomobile duyulan ihtiyacın byk řehirlerdekilere gre ok daha az olacaęı aıktır. Yine yerleřim birimi aısından dřnldęnde iklim ve doęa kořullarının da ihtiyaların ve nem sıralarının belirlenmesinde etkili olduęu gzlenir.

Farklı meslekler ya da iř gruplarında olanların ihtiyaları da farklı olacaktır. Bu durum zellikle konut ihtiyacında gzlenmektedir. Devlet memuru, zel sektr alıřanı, serbest meslek mensubu ve iřiler, tccar, sanayici ve dięer meslek gruplarındakilere gre daha ok konut ihtiyacı hissetmektedirler [6].

İhtiyalar, eęitim durumuna gre de farklılařır. Zaten eęitim durumunun hayata bakıř aısını etkiledięi bilinmektedir. Diplomasız ya da ilkokul mezunlarının iř amalı mallara olan ihtiyacı eęitim dzeyi daha yksek olanlara gre daha fazladır. Dięer taraftan hibir diploması bulunmayanlar iin konut en nemli ihtiyatır [6].

### ***2.3. Tketim ve Tketic Davranıřları***

Tketic davranıřının ve tketim olayının pazarlama dnyasındaki nemi byktr. Kiři yařamının byk blm tketim faaliyetlerini gerekleřtirmekle geirir. Kiři sadece tketim olayı deęil, tketicilecek mal ya da hizmetin seimi iin byk zaman ayırmaktadır. Pek ok tketic farkında olarak ya da olmayarak reklamları, ilanları izlemekte ve bunlardan etkilenmektedir. Yine pek ok kiři marka baęımlısıdır ve zamanının byk blmn maęazaları dolařarak, arkadař

çevresinden bilgiler alarak, tüketim sırasında ihtiyaçlarını ne derecede tatmin ettiğini değerlendirerek geçirir [5].

Bu nedenlerledir ki ilk zamanlarda tüketici davranışı pazarlamacılar tarafından sadece satın alma eylemi olarak ifade edilmekteydi. Oysaki günümüzde artık tüketici davranışı sadece bir satın alma faaliyeti olmaktan çıkmış, satın alma öncesi ve sonrası davranışların tümü olarak algılanır hale gelmiştir.

Tüketici davranışları hem sosyo-kültürel hem de psikolojik olguların etkisi altındadır. Sosyo-kültürel etkiler kişinin danışma grubu, ailesi, sosyal sınıf ve kültürle ilişkilidir. Psikolojik etkiler ise öğrenme ve bellek, güdülenme ve ilgilenme, algılama, tutumlar, kişilik ve yaşam biçiminden oluşmaktadır [5].

Tüketici davranışlarında etkin olan sosyo-kültürel yapı beslenme alışkanlıklarını etkilemektedir. Geleneksel yapıya sahip Türk ailesinde özellikle ev yapımı unlu gıda ürünleri oldukça fazla miktarlarda tüketilmektedir. Diğer taraftan vejeteryan bireylerin sayısı batı ülkelerine göre çok azdır. Ayrıca dışarıda beslenme alışkanlığı büyük şehirlerde son zamanlarda yaygınlaşmaya başlamıştır. Küçük yerleşim birimlerinde böyle bir alışkanlık henüz yerleşmemiştir.

Kişilerin tüketim davranışlarına gelirin etkisi çok büyüktür. Gelirleri farklı kişilerin ürünleri ve hizmetleri algılayışları da farklı olur. Mallara bakış açısı tüketicilerin davranışına doğrudan etki eder.

#### **2.4. Gelir**

Sosyo-ekonomik yapı, harcama yönü ve miktarı, sosyal sınıflar, yaşama biçimi hep gelirin etkisiyle şekillenir. Bu nedenle hanehalkı sosyo-ekonomik açıdan incelenirken gelir üzerinde durulmalıdır.

Hanehalkının parasal olarak elde ettiği gelir nakdi gelir olarak adlandırılır. Ancak hanehalkının gelirini sadece nakdi gelir belirlemez. Hanehalkı tarafından mal olarak elde edilen gelir aynı gelir olarak adlandırılır.

Gelirin belirlenmesine yönelik çalışmalarda aynı olarak elde edilen gelirin parasal karşılığı bulunur ve bu miktar gelirin hesaplanmasında kullanılır.

Geleneksel yapıya sahip Türk hanesindeki aynı gelirler hane dışından gelen yiyecekler, çalışılan birimin sağladığı avantajlar şeklindedir. Türkiye’de hanehalkı gelirlerinin 1/5 gibi büyük bir oranı aynı gelir olarak elde edilmektedir. Hanelerin yarısına yakın kısmı birden fazla gelir kaynağına sahiptir. Bu durum enflasyonun sabit gelirliler üzerindeki etkisini hafifletmektedir [7] .

Hanehalkının sosyo-ekonomik durumunun belirlenmesinde gelirin önemi çok büyüktür. Bununla birlikte hanenin gelirini doğru olarak saptamak çok zordur. Kişilere aylık gelirleri ve harcamaları sorulduğunda, genellikle cevap vermekten kaçınırlar. Cevap vermekte istekli olan kişiler de zaman zaman miktarları yuvarlarlar ya da yanlış hatırlayabilirler.

Bu sakıncaların giderilmesi için yapılacak en uygun çalışma haneye bir not defteri bırakarak hanenin gelirinin tam miktarları ile hergün satın aldıkları ürünleri ve tutarlarını not etmelerini sağlamaktır. Ancak bu tür çalışma anketi yürüten açısından zaman alıcı, yorucu ve maliyetli olmaktadır. Diğer taraftan çok uzun süren bu çalışma anketi cevaplayanlar için de sıkıcı olabilmektedir.

Türkiye’de gelir dağılımı açısından bir dengesizlik yaşandığı gözlenmektedir. Gelir konusunda araştırma yapmanın yukarıda sayılan güçlükleri nedeniyle bu konuda araştırmalar uzun aralıklarla yapılmaktadır. D.İ.E’nin yaptığı araştırmalar dışında bu konuda yayınlanmış son araştırma 1986 yılına ait TUSİAD’ın yaptığı araştırmadır.

Sözü edilen araştırma bulgularına göre yıllık en düşük hanehalkı kazancı 32.000 TL., en yüksek hanehalkı kazancı ise 50.000.000 TL. şeklindedir. Serbest meslekte çalışanlar, tüccar sanyici ve özel sektör çalışanları en çok kazananlar grubunu, esnaf ve zanaatkarlar orta kazançlılar grubunu çiftçi, devlet memuru, işçi ise düşük kazançlılar grubunu oluşturmaktadır [7].

Yine gelir beş ayrı düzeyde incelenirse birinci dilimde olanlar toplam gelirden % 3.9, ikinci düzeydekiler % 8.4, üçüncü düzeydekiler % 12.6, dördüncü düzeydekiler % 19.2 ve en çok kazanan beşinci düzeydekiler ise % 55.9’luk bir pay almaktadır.



1986 yılında TUSİAD tarafından yapılan araştırma sonuçlarına bağlı olarak, yerleşim birimlerine göre gelir dağılımına bakıldığında Türkiye’de en çok gelirli bireylerin İstanbul’da yaşadığı görülmektedir. Bunu diğer iki büyük şehir Ankara ve İzmir izlemektedir. Diğer şehirlere ve kırsal bölgelere gidildiğinde gelirin hızla azalmaya başladığı gözlenmektedir.[7]

### 3. ÇOK DEĞİŞKENLİ REGRESYON ANALİZİ

İstatistiğin amaçlarından biri değişkenler arasındaki ilişkileri incelemektir. İstatistiksel anlamda iki değişken arasındaki ilişki bunların değerlerinin karşılıklı değişimleri arasında bir bağıllık olarak anlaşılır. X değişkeninin değerleri değişirken buna bağlı olarak Y değişkeninin de değerleri değişiyorsa, bu ikisi arasında bir ilişki olduğu söylenebilir [8]. Bu ilişkinin fonksiyonel şekli istatistikte regresyonun konusudur.

Sözkonusu değişkenlerden X değişkeni bağımsız değişken Y değişkeni ise bağımlı değişkendir. Bağımlı değişkeni etkileyen ve bağımlı değişkendeki değişimleri açıklayan değişkene bağımsız değişken denilmektedir. Araştırmada özellikle üzerinde durulan değişken ise bağımlı değişkendir. Bu değişkenin aldığı değerlerin açıklanmasına çalışılır [9].

Regresyon analizinde X değişkeninin Y değişkenini ne şekilde etkilediği araştırılıyorsa basit doğrusal regresyon sözkonusudur. Bu durumda bağımlı değişkeni etkileyen tek bir bağımsız değişken vardır ve bu stokastik model aşağıdaki gibi ifade edilir:

$$Y_i = \alpha + \beta X_i + \varepsilon_i \quad (3-1)$$

Ancak çoğu zaman, özellikle ekonomik çalışmalarda, bağımlı değişken birden fazla bağımsız değişkenle açıklanır. Bu durumlarda uygulanan analiz çoklu regresyon adını alır ve şu doğrusal modelle ifade edilir:

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon_i \quad (i=1,2,\dots,n) \quad (3-2)$$

Çok değişkenli regresyon modeli, çoklu doğrusal regresyon modelinin, bağımlı değişkeni,  $m \times 1$  boyutlu rassal bir  $Y_m$  vektörüne genişletilmiş şeklidir [10]. Diğer bir deyişle birden fazla bağımsız değişken varken yine birden fazla bağımlı değişkenin bulunması durumudur. Bir başka ifadeyle aynı bağımsız değişkenler seti

ve aynı veri grubunun birden fazla bağımlı değişkeni nasıl etkilediğinin incelenmesidir.

### 3.1. Çok Değişkenli Regresyon Modelinin Matematiksel Gösterimi

Çok değişkenli regresyon analizinde model şu şekilde ifade edilir [11]:

$$Y = X \beta + \varepsilon \quad (3-3)$$

$(n \times m) \quad (n \times (r+1)) \quad ((r+1) \times m) \quad (n \times m)$

Model daha açık olarak şöyle yazılabilir:

$$Y_1 = \beta_{01} + \beta_{11}X_1 + \dots + \beta_{r1}X_r + \varepsilon_1$$

$$Y_2 = \beta_{02} + \beta_{12}X_1 + \dots + \beta_{r2}X_r + \varepsilon_2$$

$$\vdots \quad \quad \quad \ddots \quad \quad \quad \vdots$$

$$Y_m = \beta_{0m} + \beta_{1m}X_1 + \dots + \beta_{rm}X_r + \varepsilon_m$$

(3-4)

Matris gösteriminde veri matrisi şöyle olur:

$$X = \begin{bmatrix} X_{10} & X_{11} & \dots & X_{1r} \\ X_{20} & X_{21} & \dots & X_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n0} & X_{n1} & \dots & X_{nr} \end{bmatrix} \quad (3-5)$$

$(n \times (r+1))$

Dikkat edilirse yukarıda yer alan  $X$  veri matrisi  $(n \times (r+1))$  boyutludur. Normalde  $n \times r$  boyutlu olması gereken bu matris sabit terime de yer verilmesini

sağlamak amacıyla  $(n \times (r+1))$  boyutlu yazılır. Bu matris tek bağımlı değişkeni bulunan çoklu regresyon analizinde kullanılan veri matrisiyle aynıdır.

Bağımlı değişkeni gösteren  $n \times m$  boyutlu matris aşağıdaki gibi olur:

$$\mathbf{Y}_{(n \times m)} = \begin{bmatrix} Y_{11} & Y_{12} & \cdots & Y_{1m} \\ Y_{21} & Y_{22} & \cdots & Y_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{n1} & Y_{n2} & \cdots & Y_{nm} \end{bmatrix} = [Y_{(1)} : Y_{(2)} : \cdots : Y_{(m)}] \quad (3-6)$$

Yukarıda yer alan (3-6) matrisinin bileşenleri, herbiri ayrı bir bağımlı değişkeni ifade eden sütun vektörleridir. Pratikte bu vektörlerin herbir elemanını bağımlı değişkenin gözlem değerleri oluşturur.

Katsayılar matrisi  $(r+1) \times m$  boyutludur ve aşağıdaki gibidir:

$$\mathbf{\beta}_{((r+1) \times m)} = \begin{bmatrix} \beta_{01} & \beta_{02} & \cdots & \beta_{0m} \\ \beta_{11} & \beta_{12} & \cdots & \beta_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{r1} & \beta_{r2} & \cdots & \beta_{rm} \end{bmatrix} = [\beta_{(1)} : \beta_{(2)} : \cdots : \beta_{(m)}] \quad (3-7)$$

Yukarıdaki (3-7) ifadesinde aynı sütunda bulunan katsayılar aynı bağımlı değişkeni tahmin etmek için hesaplanan katsayılardır. (3-4) eşitliğinden de görülebileceği gibi aynı satırdaki katsayılar farklı bağımlı değişkenleri tahmin etmede kullanılırlar. Bir başka ifadeyle  $\beta_{(1)}$   $Y_1$ ,  $\beta_{(2)}$   $Y_2$ , ...,  $\beta_{(m)}$   $Y_m$  bağımlı değişkeninin tahmininde kullanılan katsayılar setini ifade eder.

$n \times m$  boyutlu artıklar matrisi ise şöyle yazılır:

$$\boldsymbol{\varepsilon}_{(n \times m)} = \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \cdots & \varepsilon_{1m} \\ \varepsilon_{21} & \varepsilon_{22} & \cdots & \varepsilon_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{n1} & \varepsilon_{n2} & \cdots & \varepsilon_{nm} \end{bmatrix} = [\boldsymbol{\varepsilon}_{(1)} \quad \boldsymbol{\varepsilon}_{(2)} \quad \cdots \quad \boldsymbol{\varepsilon}_{(m)}] \quad (3-8)$$

Katsayılar matrisinde (eşitlik (3-7)'de) gösterildiği gibi artıklara ilişkin matriste de her bir vektör aynı bağımlı değişken için hesaplanmış artık değerlerini gösterir.

Modelin tümü matrislerle ifade edildiğinde aşağıdaki gibi olur:

$$\begin{bmatrix} Y_{11} & Y_{12} & \cdots & Y_{1m} \\ Y_{21} & Y_{22} & \cdots & Y_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{n1} & Y_{n2} & \cdots & Y_{nm} \end{bmatrix} = \begin{bmatrix} X_{10} & X_{11} & \cdots & X_{1r} \\ X_{20} & X_{21} & \cdots & X_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n0} & X_{n1} & \cdots & X_{nr} \end{bmatrix} \begin{bmatrix} \beta_{01} & \beta_{02} & \cdots & \beta_{0m} \\ \beta_{11} & \beta_{12} & \cdots & \beta_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{r1} & \beta_{r2} & \cdots & \beta_{rm} \end{bmatrix} +$$

$$\begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \cdots & \varepsilon_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{21} & \varepsilon_{22} & \cdots & \varepsilon_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{n1} & \varepsilon_{n2} & \cdots & \varepsilon_{nm} \end{bmatrix} \quad (3-9)$$

(3-9)'daki gösterimin, (3-3) formülünün daha açık bir yazılımı olduğu açıktır.

### 3.2. Çok Değişkenli Regresyon Modelinin Varsayımları

Çok değişkenli regresyon analizindeki varsayımların bir kısmı basit doğrusal regresyon modelinin matris notasyonuna dönüştürülmüş şeklindedir. Çok değişkenli regresyon analizindeki varsayımlar aşağıdaki gibidir:

1- Tüm bağımlı değişkenler için hesaplanan artıkların beklenen değeri 0'dır [11].  $E(\varepsilon_{(i)}) = 0 \quad i = 1, 2, \dots, m$ . Diğer bir deyişle (3-10)'da gösterilen herbir sütun vektörünün beklenen değeri sıfırdır.

$$\varepsilon = [\varepsilon_1 : \varepsilon_2 : \dots : \varepsilon_m] \quad (3-10)$$

2-  $\hat{\beta}$ ,  $\varepsilon$  ve  $\Sigma$  kovaryans matrisinden bağımsızdır [12].

3- m gözlem sözkonusuyken j. olay için

$\text{Cov}(\varepsilon_{(i)}, \varepsilon_{(k)}) = \sigma_{ik} I \quad i, k = 1, 2, \dots, m$  ve  $\Sigma = \{\sigma_{ik}\}$  kovaryans matrisidir. Fakat farklı olaylardan olan gözlemler ilişkisizdirler [11].

4-  $\hat{\varepsilon}$  çok değişkenli normal dağılmaktadır.

Modele ilişkin özellikler ise aşağıdaki gibidir:

1-  $n \times r$  matrisi  $W_p$  ( $\Sigma$ ,  $n-r$ ) olmak üzere Wishart dağılımına sahiptir. Bu dağılıma izleyen bölümlerde değinilecektir.

2-  $\hat{\beta}$ ,  $\beta$ 'nin yansız kestiricisidir.

3-  $\hat{\beta}$  çok değişkenli normal dağılmaktadır.

### 3.3. Çok Değişkenli Regresyon Denkleminin Elde Edilmesinde Kullanılan Teknikler

Çok değişkenli regresyon modelindeki parametrelerin kestirimi için “En Çok Olabilirlik” ve “Sıradan En Küçük Kareler” teknikleri kullanılmaktadır. Bir sonraki kesimde bu tekniklere değinilecektir.

#### 3.3.1. En Küçük Kareler Tekniği

Çok değişkenli regresyon analizinde model daha önce gösterildiği gibi

$$Y = X \beta + \epsilon \quad \text{şeklindedir. Yine artık terimleri}$$

$(n \times m)$      $(n \times (r+1))$      $((r+1) \times m)$      $(n \times m)$

$$\epsilon = \begin{bmatrix} \epsilon_{11} & \epsilon_{12} & \dots & \epsilon_{1m} \\ \epsilon_{21} & \epsilon_{22} & \dots & \epsilon_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \epsilon_{n1} & \epsilon_{n2} & \dots & \epsilon_{nm} \end{bmatrix} \quad \text{olarak tanımlıdır.}$$

Bu eşitlikler gözönünde bulundurularak  $i$ .değişken için regresyon modeli

$$Y_{(i)} = X\beta_{(i)} + \epsilon_{(i)} \quad i=1,2, \dots, m \text{ şeklinde olur. Çoklu regresyon analizinde olduğu gibi}$$

$\beta_{(i)} = (X'X)^{-1} X'Y_{(i)}$  yazılabilir. Bu eşitlik aşağıdaki gibi genişletilirse çok değişkenli regresyon analizinin en küçük kareler tahminlerine ulaşılır.

$$\hat{\beta} = [\hat{\beta}_{(1)} : \hat{\beta}_{(2)} : \dots : \hat{\beta}_{(m)}] = (X'X)^{-1} X' [Y_{(1)} : Y_{(2)} : \dots : Y_{(m)}] \quad (3-11)$$

Buradan da kısaca

$$\hat{\beta} = (X'X)^{-1} X'Y \quad (3-12)$$

sonucuna ulaşılır. Diğer taraftan

$$n \hat{\Sigma} = (Y - X\hat{\beta})'(Y - X\hat{\beta}) \quad (3-13)$$

yazılabilir.

Bir başka ifadeyle model tanımından

$\hat{\epsilon} = Y - X\hat{\beta}$  yazılır.  $i$ .değişkenin seçildiğini düşünelim. Bu değişken için kareler toplamı

$(Y_{(i)} - X\beta_{(i)})'(Y_{(i)} - X\beta_{(i)})$  olur. Buradan

$\text{iz}[(Y - X\beta)(Y - X\beta)]$  ,  $\beta = \hat{\beta}$  seçilmesiyle en küçüklenir. Yine genelleştirilmiş varyans  $|(Y - X\hat{\beta})'(Y - X\hat{\beta})|$  en küçük kareler tahmini  $\hat{\beta}$  ile en küçüklenir.

$\hat{\beta}$  parametreleri en küçük kareler tekniğiyle elde edilmiş katsayı tahmin değerleridir. Bu nedenle  $\hat{\beta}$  katsayıları gözlemlenen  $y_j$  değerlerinden mümkün olduğunca küçük farklar kareler toplamı geliştiren tahmincilerdir. Yine bu teknikle elde edilen tahminciler yansız tahmincilerdir. Kısaca en iyi doğrusal yansız tahmincilerdir [13].



### 3.3.2. En Çok Olabilirlik Tekniđi

Bir regresyon denklemindeki özel deđişken deđerlerini elde etmeden önce  $\hat{\beta}$ 'nin ve ortak kareler toplamının  $\hat{\epsilon}\hat{\epsilon}'$  örnekleme dađılımını belirlemeliyiz. Bunu yapmak için  $\epsilon$  artık deđerlerini normal dađılan ve sabit varyansa sahip bađımsız deđerler olarak düşünüyoruz [11].

Bu varsayımın gerçekleştiđi durumlarda, çok deđişkenli regresyon analizinde parametre kestirimleri için kullandığımız teknik en çok olabilirlik tekniđidir.

Hata terimleri için yazılan varsayımların sağlanmasıyla, bađımlı deđişken deđerleri,  $N(X\beta, \sigma^2 I)$  olmak üzere normal dađılacaktır.

Normal dađılım için yoğunluk fonksiyonu;

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty \quad \text{şeklindedir.}$$

Matris gösteriminde benzerlik fonksiyonu şöyle yazılır:

$$L = \frac{1}{(\sigma\sqrt{2\pi})^{T/2}} e^{-\frac{(Y-X\beta)(Y-X\beta)}{2\sigma^2}} \quad (3-14)$$

Çok deđişkenli regresyon analizinde ise benzerlik fonksiyonu aşağıdaki gibi olur:

$$L(\beta, \Sigma) = \prod_{j=1}^n \frac{1}{(2\pi)^{m/2}} \frac{1}{|\Sigma|^{1/2}} e^{-\frac{1}{2}(y_j - \beta'x_j)' \Sigma^{-1} (y_j - \beta'x_j)} \quad (3-15)$$

Buradan kısaca

$$L(\beta, \Sigma) = -\frac{1}{2} n \log |2\pi \Sigma| - \frac{1}{2} (Y - X\beta)' \Sigma^{-1} (Y - X\beta) \quad (3-16)$$

elde edilir.<sup>1</sup>

Elde edilecek parametrelere göre türevler alınıp ifade 0'a eşitlendiğinde

$$\hat{\beta} = (X'X)^{-1} X'Y \quad (3-17)$$

$$n \hat{\Sigma} = (Y - X\hat{\beta})'(Y - X\hat{\beta}) \quad (3-18)$$

bulunur. Elde edilen parametreler en küçük kareler tahmincileriyle aynıdır.

Hesaplanan en çok olabilirlik kestiricilerinin özellikleri şunlardır [15]:

- Tahminciler tutarlıdır.  $\text{plim} \hat{\theta}_{ML} = \theta$
- Asimtotik olarak normal dağılırlar.  $\hat{\theta} \rightarrow N[\theta, \{I(\theta)\}^{-1}]$
- Asimtotik olarak etkindirler.

İkinci özellik hipotez testlerinin yapılması ve güven aralıklarının hesaplanmasında büyük kolaylık sağlar. Son özellik oldukça önemlidir. Tutarlı tahminci ile en çok olabilirlik tahmincisinin en küçük varyans özelliğine sahip olduğunu ifade eder.

<sup>1</sup> Ayrıntılı işlemler için bakınız [11].

### 3.4. Wishart Dağılımı

Bölüm 3.3.1. ve 3.3.2.'de tanımlanan  $n \Sigma$  matrisi Wishart dağılmaktadır. Bu nedenle bu bölümde Wishart dağılımına değinilecektir.

Ki-kare dağılımı Gamma dağılımının  $\alpha = k/2$  ve  $\beta=1/2$  olmak üzere özel bir şeklidir. Dağılımın yoğunluk fonksiyonu aşağıdaki gibidir [17]:

$$f(x) = \frac{1}{2^{k/2} \Gamma(k/2)} X^{k/2-1} e^{-X/2}, \quad X > 0 \quad (3-19)$$

Wishart dağılımı ise ki-kare dağılımının k tane parametreye genişletilmiş bir şeklidir.

$N(0, \Sigma)$  olmak üzere normal dağılan bir A parametresi düşünölsün. Bu matris p boyutlu N-1 ve N-k rassal vektörlerinin çarpımları olarak yazılabilir. Yine aynı şekilde matrisi  $\sum_{i=1}^n y_i y_i'$  toplamına dönüştürebiliriz.

Burada yazılan  $y_i$  vektörünün Wishart dağılımına sahip p değışkenli bağımsız dağılan bir vektör olduđu söylenir. Dağılımın yoğunluk fonksiyonu ise şöyle yazılır:

$$W(A; \Sigma, n) = \begin{cases} \frac{|A|^{1/2(n-p-1)} e^{(-1/2 \text{tr} A \Sigma^{-1})}}{2^{np/2} \pi^{1/4p(p-1)} |\Sigma|^{1/2n} \prod_{i=1}^p \Gamma[1/2(n+1-i)]} & , A > 0 \\ 0 & , \text{d. d.} \end{cases}$$

(3-20)

n serbestlik derecesi olmak üzere  $\Sigma$  ile beraber yoğunluk fonksiyonunun iki parametresi bulunmaktadır [16]. Eđer dağılımda p değışken yerine tek bir değışken

yer alırsa ve  $\Sigma$  skaler fomda ise Wishart yoğunluk fonksiyonu  $n$  serbestlik derecesine sahip bir ki-kare fonksiyonuna dönüşür ve bu fonksiyonun tüm özelliklerini taşır. Kısaca Wishart dağılımı  $\chi^2$  dağılımının çok değişkenli forma genelleştirilmiş şeklidir [18].

Wishart Dağılımı şu özelliklere sahiptir:

- Örneklem ortalamasını ifade eden  $\bar{X}$  vektörü ile çarpımlar ve kareler toplamı matrisi  $A$ , bağımsız olarak dağılan aynı örneklemden hesaplanmıştır.

-  $A_1, A_2, \dots, A_k$  Wishart dağılımına sahip matrisler gibi parametresi  $\Sigma$  matrisi ve serbestlik derecesi  $n_1, n_2, \dots, n_k$  olmak üzere bağımsızsa, toplamları da  $n=n_1+n_2+\dots+n_k$  olmak üzere Wishart dağılır.

- Eğer  $A, (W_m|\Sigma)$  olmak üzere Wishart dağılmışsa  $CAC'$  de  $W_m(CAC'|C\Sigma C')$  olmak üzere Wishart dağılır [11].

### 3.5. Regresyon Katsayılarının Anlamlılık Sınamaları

En küçük kareler ya da en çok olabilirlik teknikleriyle hesaplanan regresyon katsayıları çeşitli tekniklerle- sınanır.Bu sınamalara izleyen bölümlerde yer verilmiştir.

#### 3.5.1. Wilks Lambda Olabilirlik Oran Sınaması

En çok olabilirlik tekniği ya da en küçük kareler tekniği ile katsayıları hesaplanan regresyon denkleminin anlamlılığı olabilirlik oran testi ile sınanır. Adından da anlaşılacağı gibi olabilirlik oran testi, en çoklanmış olabilirlik değerlerinin bir oran kıyaslamasıdır [19]. Olabilirlik oran testinin dayandığı

hipotezler aynen tek bağımlı değişkene sahip regresyon analizinde olduğu gibi şöyle ifade edilir:

$$H_0: \beta_{ij}=0$$

$$H_1: \text{En az bir } \beta_{ij} \neq 0$$

Sınamanın dayandığı test istatistiği olan Wilks Lambda istatistiği aşağıdaki gibi hesaplanır:

$$\Lambda = \frac{|\hat{\Sigma}|}{|\tilde{\Sigma}|} \quad (3-21)$$

Burada pay aşağıdaki matrisin determinant değeriyle hesaplanır:

$$n \hat{\Sigma} = Y'Y - \hat{Y}'\hat{Y} \quad (3-22)$$

Payda ise

$$n \tilde{\Sigma} = Y'Y - n\bar{y}'\bar{y}' \quad (3-23)$$

matrisinin determinant değeridir.

İki değer oranlanmasıyla hesaplanan Wilks Lambda istatistiği aşağıda verilen Bartlett test istatistiğinde yerine konur.

$$\left[ n - r - 1 - \frac{1}{2}(m - r + 1) \right] \log \Lambda \quad (3-24)$$

Elde edilen değer ki-kare tablo değeriyle karşılaştırılır. Böylece denklemin anlamlılığına ya da anlamsızlığına karar verilir.

### 3.5.2. Varyans Analizi Tekniđi

Çok deđişkenli regresyon analizinde hesaplanan katsayıların anlamlılıđının sınanmasında kullanılan bir diđer teknik varyans analizi tekniđidir. Çok deđişkenli regresyonda varyans analizi tablosu řöyle oluşturulur:

Tablo 3.1. Çok deđişkenli regresyon analizinde varyans analizi tablosu

Deđişim Kaynađı	sd	Kareler Toplamı
Regresyon	r	$H = \hat{Y}'\hat{Y} - n\bar{y}\bar{y}'$
Hata	n-r-1	$\hat{E}'\hat{E} = Y'Y - \hat{Y}'\hat{Y}$
Genel	n-1	$T = Y'Y - n\bar{y}\bar{y}'$

Yukarıdaki varyans analizi tablosunda yer alan regresyon ve hata kareler matrislerinden yararlanarak lambda istatistiđi hesaplanabilir.

$$(\hat{E}'\hat{E})^{-1}H \quad (3-25)$$

matrisinin özdeđerleri  $\lambda_1, \lambda_2, \dots, \lambda_p$  deđerleri bulunur. Bu özdeđerlerden yararlanılarak lambda istatistiđi ařađıdaki formülle hesaplanır:

$$\Lambda = \prod_{i=1}^p (1 + \lambda_i)^{-1} \quad (3-26)$$

Hesaplanan deđer Bartlett istatistiđinde yerine konur:

$$\left[ n - r - 1 - \frac{1}{2}(m - r + 1) \right] \log \Lambda \quad (3-27)$$

Bölüm 3.5.1.'de de tanımlanan bu ifadede  $r$  ve  $m$  bilinmeyen  $\beta$  parametrelerine ilişkin matrisin boyutu ve  $n$  örneklem büyüklüğüdür. Bartlett istatistiği olarak bulunan değer  $m$  ve  $r$  serbestlik dereceleriyle ki-kare dağılımına sahiptir. Ki-kare tablo değeriyle karşılaştırılır ve katsayının anlamlılığına karar verilir.

### 3.5.3. Parsimony Kavramına Dayalı Varyans Analizi Tekniği

İstatistiksel analizlerin hepsi gözlemlere bağlı bir mekanizmanın geliştirilmesiyle oluşan bir matematiksel modele dayanır. Bu matematiksel model analizin işleyişini belirler. Gereksiz değişkenler fonksiyonu daha karmaşık bir hale getirir. Bununla beraber analiz de gittikçe daha zor olacaktır. Ancak çoğu zaman ilk anda hangi değişkenlerin analize alınacağına karar vermek oldukça zordur. Bu nedenle olaya etkisi olduğu düşünülen tüm değişkenler analize alınır. Daha sonra bağımsız değişkenlerin tamamı için varyans analizi yapılarak olayı açıklamadaki etkileri araştırılır.

Değişken seçiminde izlenen bir diğer yol da parsimony kavramı üzerine kuruludur. Parsimony fonksiyonel formun basit ve değişken sayısının minimum olmasıdır [16]. Buna göre bağımsız değişkenlerin bir kısmının anlamlı ve diğer kısmının anlamsız olduğu varsayılarak regresyon katsayıları hesaplanır. Bulunan bu katsayılar varyans analiziyle sınanır.

Varyans analizinin gerçekleştirilmesi için  $r$  tane bağımsız değişken  $r_1$  ve  $r_2$  olmak üzere iki alt gruba ayrılır. İlk grubun regresyon katsayılarının anlamlı, diğer grubun katsayılarının anlamsız olduğu şeklindeki hipotezler kurulur ve sınanır. Duruma ilişkin varyans analizi tablosu aşağıdaki gibidir [12]:

Tablo 3.2. İndirgenmiş model için varyans analizi tablosu

Değişim Kaynağı	Kareler Toplamı	sd
Regresyon (q değişken)	$H = \hat{Y}'\hat{Y} - n\bar{y}\bar{y}'$	r
r <sub>1</sub> değişken	$H_{(r_1)} = T - \tilde{E}'\tilde{E}$	r <sub>1</sub>
r <sub>2</sub> değişken	$H_{(r_2)} = H - H_{(r_1)}$	r <sub>2</sub>
Hata	$\hat{E}'\hat{E} = Y'Y - \hat{Y}'\hat{Y}$	n-r-1
Genel	$T = Y'Y - n\bar{y}\bar{y}'$	n-1

İndirgenmiş modele ilişkin denklemin sınanması da Wilks lambda istatistiğinin elde edilmesini gerektirir. Bu istatistiği

$(\hat{E}'\hat{E})^{-1}H_{(r_2)}$  matrisinin özdeğerlerine  $(\lambda_1, \lambda_2, \dots, \lambda_p)$  bağlı olarak aşağıdaki gibi elde ederiz:

$$\Lambda = \prod_{i=1}^p (1 + \lambda_i)^{-1} \quad (3-28)$$

Wilks istatistiği yine aşağıda verilen Bartlett istatistiğinde yerine konur:

$$\left[ n - r - 1 - \frac{1}{2}(m - r + 1) \right] \log \Lambda \quad (3-29)$$

Bulunan değer ki-kare tablo değer yine ki-kare dağılımına sahiptir. Ki-kare tablo değeriyle karşılaştırılır ve modelin seçilen kısmının istatistiksel olarak anlamlı ya da anlamsız olduğuna karar verilir. Seçilen bağımsız değişkenler istatistiksel olarak anlamlı ise bu değişkenlerin olayı açıklamada yeterli olduğuna karar verilir ve diğer değişkenler denklemden çıkarılır. Diğer taraftan seçilen değişkenlerin anlamsız olduğu gözlenirse bağımlı değişkenleri açıklamak için diğer değişkenlere de ihtiyaç duyulduğuna karar verilir.



### 3.6. Regresyon Katsayıları İçin Güven Sınırlarının Belirlenmesi

Çok değişkenli regresyon analizinde hesaplanan katsayılara ilişkin güven sınırlarının belirlenmesi mümkündür. Bu amaçla  $X$  veri matrisi iken  $X_0$  bu matrisin sütun ortalamalarının oluşturduğu bir sütun vektörü olsun. 2 bağımlı 2 bağımsız değişken içeren bir regresyon analizi için  $\hat{\beta}_1$ ,  $Y_1$  bağımlı değişkenini hesaplamada kullanılan katsayılar vektörü iken,  $\hat{\beta}_2$ ,  $Y_2$  bağımlı değişkenini hesaplamada kullanılan katsayılar vektörüdür. Dolayısıyla

$$\hat{Y}_1 = X\hat{\beta}_1 \quad \text{ve} \quad \hat{Y}_2 = X\hat{\beta}_2$$

olur. Elde edilen tahmin vektörlerine göre  $n\hat{\Sigma}$  matrisi aşağıdaki gibi hesaplanmaktadır [11]:

$$n\hat{\Sigma} = \begin{bmatrix} (\hat{Y}_2 - X\hat{\beta}_2)'(\hat{Y}_1 - X\hat{\beta}_1) & (\hat{Y}_1 - X\hat{\beta}_1)'(\hat{Y}_2 - X\hat{\beta}_2) \\ (\hat{Y}_2 - X\hat{\beta}_2)'(\hat{Y}_1 - X\hat{\beta}_1) & (\hat{Y}_2 - X\hat{\beta}_2)'(\hat{Y}_2 - X\hat{\beta}_2) \end{bmatrix} \quad (3-30)$$

Matris elemanlarının  $n$ 'e oranlanmasıyla  $\hat{\Sigma}$  matrisi elde edilir. Bu matrisin köşegen elemanları  $\hat{\sigma}_{ii}$  değerlerini verir. Başlangıçta verilen  $X_0$  ortalamalar vektörü de kullanılarak güven sınırları aşağıdaki formülle hesaplanır:

$$X_0' \hat{\beta}_k \pm \sqrt{\left( \frac{p(n-q-1)}{n-p-q} \right) F_{p;n-q-p;(\alpha)}} \sqrt{X_0' (X'X)^{-1} X_0 \left( \frac{n}{n-q-1} \right) \hat{\sigma}_{ii}} \quad (3-31)$$

Yukarıdaki ifadede  $k$  tahminci sayısını göstermektedir. Bu hesaplamalardan sonra her iki bağımlı değişken için güven sınırları hesaplanmış olur. Bu güven sınırları, ortalamaları merkez noktasının koordinatlarını belirleyen bir elipsi oluşturur.

### 3.7. Artıklarla Model Uygunluğunun İncelenmesi

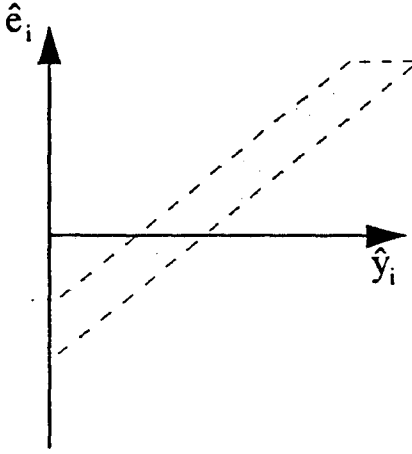
Modele göre çıkarsamalar yapmadan önce model uygunluğunun incelenmesi gerekmektedir. Modeldeki uyum eksikliği doğrudan artıklara yansiyacaktır. Bu nedenledir ki yanlış model hem artıkların büyümesine hem de yanlış çıkarsamaların yapılmasına neden olur.

Çok değişkenli regresyon analizinde model uygunluğunun incelenmesi tek bağımlı değişkeni bulunan regresyon analizinde çizilen grafiklerin aynılarının çizilmesiyle gerçekleştirilir. Bununla birlikte model uygunluğunu belirlemede kullanılan artıklara göre grafiklerin çizim yöntemi artıkların çok değişkenli normal dağılıma uygunluğunu kesin olarak belirtmede yeterli olmamaktadır. Bu konuda şüpheye düşüldüğünde ya çok karmaşık çizimler yapılmalı ya da daha önce tanımlanan artıklara ilişkin matrise temel bileşenler analizi uygulanmalıdır [12].

Eğer model doğru belirlenmişse artık değerleri için ortalama 0 ve varyans  $\sigma^2$  olmak üzere normal dağıldığı şeklindeki varsayım sağlanmış olur. Varsayımların sağlanıp sağlanmadığını görebilmek amacıyla artıklara göre iki farklı çizim yapılmaktadır. Bunlardan ilki artıkların tahmin değerlerine göre çizimi diğeri ise artıkların bağımsız değişkene göre çizimidir.

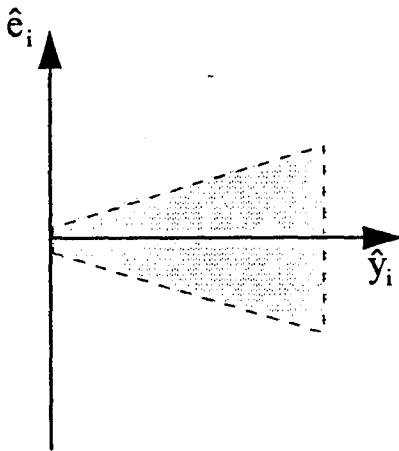
Bölüm 3.7.1. ve 3.7.2.'de çok değişkenli regresyon analizinde model uygunluğunun gözlenebilmesi amacıyla artıklara göre çizilmiş çeşitli grafikler ve yorumları yer almaktadır [11].

### 3.7.1. Artık Değerlerinin, Tahmin Değerlerine Göre Çizimi



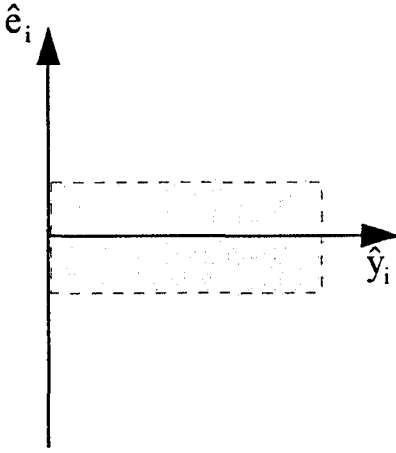
Şekil 3.1. Artıkların tahmin değerlerine göre çizimi

Artıkların tahmin değerlerine göre çizimi yukarıdaki gibi bir şekil sergilerse, tahmin değerleri ile artıklar arasında bir ilişkinin varlığı sözkonusudur. Yapılması gereken öncelikle işlemlerde bir hatanın olup olmadığının kontrol edilmesidir. Bir başka olasılık ise  $\beta_0$  katsayısının denkleme ilave edilmemiş olmasıdır. Bu katsayı ilave edildikten sonra grafik yeniden çizilmeli ve ilişkinin yer alıp almadığı tekrar gözden geçirilmelidir.



Şekil 3.2. Artıkların tahmin değerlerine göre çizimi

Yukarıda olduğu gibi artıkların tahmin değerlerine göre çizimi sağa megafon şeklinde ise sabit varyans varsayımı çığnenmiş demektir. Bu durum  $y_j$  değerleri için az olması beklenen değışkenliğın çok olduğunun bir göstergesidir. Sabit varyans varsayımının çığnendiğı gözlemlendiğinde yapılması gereken, tartılı en küçük kareler tekniğinin kullanılması ya da bazı dönüşümlere yer verilmesidir. Bazı durumlarda ise her iki yola birden başvurulması mümkündür.

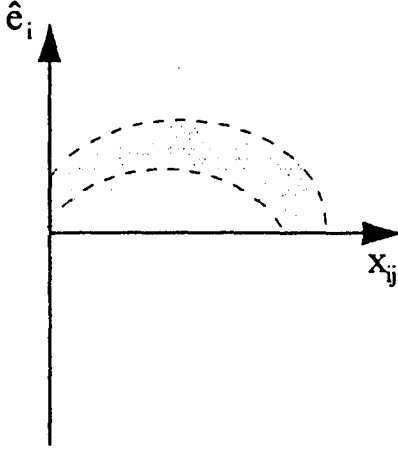


Şekil 3.3. Artıkların tahmin değerlerine göre çizimi

Artıkların tahmin değerlerine göre çizimi Şekli 3.3'te olduğu gibi bir dikdörtgeni andırıyorsa hem bağımsızlık varsayımının hem de sabit varyans varsayımının gerçekleştiğı söylenir.

Artıkların tahminlere göre grafikleri çizilirken herbir tahmin edilen değışken için ayrı ayrı çizim yapılmaktadır.

### 3.7.2. Artık değerlerinin, Bağımsız Değişkene Göre Çizimi



Şekil 3.4. Artık değerlerinin bağımsız değişkene göre çizimi

Artık değerlerinin bir bağımsız değişkene göre çizimi yukarıda olduğu gibi u veya ters u şeklindeyse bu durum modele bir başka açıklayıcı değişkenin ilave edilmesi gerektiğini ya da karesel olması gereken terimlerin doğrusal olarak modelde yer aldığını gösterir. Böyle bir durumda dışlanmış açıklayıcı değişken belirlenmeli ya da değişkenler karesel forma dönüştürüldüğünde aynı grafik bir kez daha çizilip nasıl bir şekil sergilediği incelenmelidir.

Artıkların bağımsız değişkene göre grafiği çizilirken, herbir bağımsız değişken için ayrı ayrı çizimler yapılmaktadır.

## **4. GELİR VE HANEHALKI KİŞİ SAYISIYLA ET VE SEBZE TÜKETİMİ ARASINDAKİ İLİŞKİNİN ÇOK DEĞİŞKENLİ REGRESYON ANALİZİYLE İNCELENMESİ**

### **4.1. Problemin Tanımlanması, Veri Kaynağı ve İncelencek Değişkenlerin Belirlenmesi**

Çok değişkenli regresyon analizinin ekonomi alanında uygulandığı bu çalışmanın üçüncü bölümünde çok değişkenli regresyon tekniğinin matematiksel yapısı üzerinde durulmuştur. Bu bölümde yer verilecek uygulama aşamasında Ankara’da bulunan 35 hanenin oluşturduğu örneklem üzerinde çok değişkenli regresyon analizi uygulanarak gelir ve hanede yaşayan kişi sayısının et ve sebze miktarlarına olan etkisi araştırılacaktır.

#### **4.1.1. Uygulama Alanı**

Alım gücü ve dolayısıyla tüketilen mal miktarını gelir belirlemektedir. Diğer taraftan tüketilen besin miktarının doğrudan bu besini tüketen kişi sayısı ile ilişkili olduğu düşünülmektedir.

Buradan hareketle “Çok Değişkenli Regresyon Analizi”nin uygulandığı bu çalışmada, hanehalkının geliri ve kişi sayısı ile temel besin maddelerinden et ve sebzeyi tüketme miktarları arasındaki ilişkinin ortaya konması amaçlanmaktadır.

#### **4.1.2. Veri Kaynağı**

İkinci Bölümde değinildiği gibi gelir ve harcamalara ilişkin araştırma yapmak oldukça zordur. Ayrıca bu tür bir çalışmanın güvenilirliğinden de şüphe duyulabilir. Bu nedenle bu çalışmada Devlet İstatistik Enstitüsü’nün (D.İ.E.) 1994

yılında yaptığı “Hanehalkı Gelir ve Tüketim Harcamaları” adlı anket sonucunda elde ettiği hazır verilerden yararlanılmıştır.

Anket Türkiye’de tüm yerleşim yerleri örnek seçimi için kapsama dahil edilmiş; 1994 yılında Türkiye sınırları içinde yaşayan T.C. uyruklu tüm hanehalkı fertleri kapsamıştır. Okul, yurt, otel, çocuk yuvası, huzurevi, hastane, hapisane, kışla ve orduevinde bulunanlar ile göçer nüfus ise kapsam dışında tutulmuştur.

Anket için öncelikle coğrafi bölgeler ile kırsal ve kentsel yerleşim yerleri ve nüfus grupları düzeyinde tabakalama yapılmıştır. Örneklem yöntemi gereği, örnek iller tesadüfi olarak seçildikten sonra, bu örnek illere bağlı yerleşim yerleri belirlenen tabakalar ayırımında ele alınarak, yine tesadüfi olarak seçilmiştir. Son aşamada ise tekrar tesadüfi seçim yapılarak örnek haneler belirlenmiştir.

1994 Hanehalkı Gelir ve Tüketim Harcamaları Anketi, 7 coğrafi bölge içinde kır-kent, kır-kent içinde nüfus grupları dikkate alınarak 6 tabakada toplam 26.256 hanehalkında uygulanmıştır.

	<u>Nüfus grubu</u>	<u>Tabaka</u>
Kır	0 – 2001	1
	2001 – 5000	2
	5001 – 20000	3
Kent	20001 – 50000	4
	50001 – 200000	5
	200001 +	6

Ankara ilinde toplam 690 hanehalkına anket uygulanmıştır.

1994 yılında sözkonusu anket bir yıl boyunca her ay farklı 2188 örnek hanehalkına uygulanmıştır.<sup>1</sup>

<sup>1</sup> Yazılı Görüşme, 1996. KABADAYI, G., D.İ.E. Yayın Haberleşme ve Halkla İlişkiler Şubesi Müdürü

Anket formu 7 bölümden oluşmaktadır [20]:

1. Hanehalkı tanım ve soru kağıdı işletim bilgileri (demografik bilgiler)
2. Konut ve konut kolaylıkları
3. Ekonomik durum
4. Hanehalkı tüketim harcamaları
5. Hanehalkı bileşimi, istihdam ve gelir durumu
6. Tarımsal işletme ile ilgili bilgiler
7. Gelir ve harcama dengeleri

Çalışmada, anketin Ankara'da yaşamakta olan ailelerden, D.İ.E. tarafından rassal olarak seçilen 35 haneyi kapsayan kısmı kullanılmıştır.

1994 yılı Ocak ayından Aralık ayı sonuna kadar 12 aya ilişkin ayrı ayrı yapılan anketlerin Ocak ayına ait olanı kullanılmıştır. Anket, D.İ.E. çalışanın haneye bir anket kitapçığı bırakarak, her günkü harcamalarını not etmelerini sağlanmasıyla gerçekleştirilmiştir. Bu nedenle verilerin güvenilirliğinden şüphe duyulmamaktadır.

Uygulama çalışmasında kullanılan et ve sebze tüketim miktarlarına ilişkin veriler anket kitapçığının dördüncü bölümü olan hanehalkı tüketim harcamaları bölümünden alınmıştır. Diğer değişken verileri, kişi sayısı ve gelir, anketin hanehalkı bileşimi istihdam ve gelir durumu adlı beşinci bölümünden alınmıştır. Veriler Ek-1'de yer almaktadır.

#### **4.1.3. Değişkenlerin Belirlenmesi**

Çok değişkenli regresyon tekniğinde aynı anda birden fazla sayıda bağımlı değişken aynı bağımsız değişkenlerle tahmin edilebilmektedir. Burada dikkat edilmesi gereken, bağımsız değişken sayısı, bağımlı değişken sayısına eşit ya da daha fazla olmalıdır. Bu uygulama çalışmasında 2 bağımlı, 2 bağımsız değişken kullanılmıştır.



Hanehalkının ekonomik durumu incelenirken ilk düşünülmesi gereken hanehalkının geliridir. Diğer taraftan hanede yaşayan bireylerin tükettikleri et ve sebze miktarının hanede yaşayan birey sayısı ile ilişkili olduğu düşünülmektedir. Dolayısıyla et ve sebze tüketim miktarı bağımlı değişkenler, gelir ve hanede yaşayan kişi sayısı bağımsız değişkenler olarak tanımlanmıştır:

$Y_1$ : Hanede 1 ayda tüketilen et miktarı

$Y_2$ : Hanede 1 ayda tüketilen sebze miktarı

$X_1$ : Hanede yaşayan kişi sayısı

$X_2$ : Hanehalkı toplam aylık geliri

Bağımlı değişkenler, 1 ay boyunca hanede tüketilen et ve sebze miktarlarıdır. Her iki bağımlı değişken kg. cinsinden verilmektedir.

İlk bağımsız değişken hanede yaşayan kişi sayısı olarak belirlenmiştir. Hanedeki kişi sayısı doğrudan alınmıştır.

Son değişken hanehalkının geliri olarak tanımlanmıştır. Diğer değişken değerleri doğrudan kullanılmasına rağmen, gelir değişkeni çok büyük değerler aldığından 1.000.000'a oranlanmış ve sonuçlar  $X_2$  bağımsız değişkeni olarak kaydedilmiştir.

## ***4.2. Gelir ve Hanedeki Kişi Sayısıyla Et ve Sebze Tüketimi Arasındaki İlişkinin Çok Değişkenli Regresyon Tekniği İle Belirlenmesi***

### ***4.2.1. Katsayıların Tahmini***

Çok değişkenli regresyon analizinde bağımlı ve bağımsız değişken değerleri matris formunda yazılmaktadır. Sabit terime bağımlı değişkenlerin tahmin denklemlerinde yer verebilmek amacıyla  $X$  veri matrisinin ilk sütunu 1 değerlerinden oluşmaktadır.  $X$  veri matrisi Ek-2'de yer almaktadır. Bağımlı değişken veri setinin bulunduğu  $Y$  veri matrisi Ek-3'te gösterilmiştir.

Verilere ilişkin varyans kovaryans matrisi aşağıdaki gibidir:

$$S = \begin{bmatrix} 15,8521 & 597,5619 & 26,5424 & 169,8771 \\ 597,5619 & 1587,5933 & 318,1521 & 1683,7243 \\ 26,5424 & 318,1521 & 1,4319 & 80,4058 \\ 169,8771 & 1683,7243 & 40,8058 & 195,5541 \end{bmatrix} \quad (4-1)$$

Katsayıların tahmininde MapleV 2 paket programından yararlanılmış ve sonuçlar aşağıdaki gibi bulunmuştur:

$$\hat{\beta} = (X'X)^{-1} X'Y$$

$$\hat{\beta} = \begin{bmatrix} 2,23 & 17,65 \\ 0,48 & 16,27 \\ 0,13 & 0,01 \end{bmatrix} \quad (4-2)$$

Çalışmada iki bağımlı, iki bağımsız değişkene yer verildiğinden çok değişkenli regresyon modeli şöyle olacaktır:

$$\hat{Y}_1 = \hat{\beta}_{01} + \hat{\beta}_{11}X_1 + \hat{\beta}_{21}X_2$$

$$\hat{Y}_2 = \hat{\beta}_{02} + \hat{\beta}_{12}X_1 + \hat{\beta}_{22}X_2$$

Model için katsayılar hesaplanmış ve sonuçta aşağıdaki denklemler elde edilmiştir:

$$\hat{Y}_1 = 2,23 + 0,48X_1 + 0,13X_2 \quad (4-3)$$

$$\hat{Y}_2 = 17,65 + 16,27X_1 + 0,01X_2$$

Elde edilen denklemler doğrusal olduğundan çok değişkenli regresyon tekniği ile elde edilen katsayıların yorumu çoklu doğrusal regresyon tekniğindeki katsayıların yorumuna benzemektedir. Denklemlerdeki katsayılar, diğer bağımsız değişken sabit tutulurken bir bağımsız değişkendeki 1 birim değişimden dolayı bağımlı değişkendeki değişimi ifade etmektedir.

Her iki bağımsız değişken iki denklemde de yer almaktadır. Bir başka deyişle, gelirdeki değişim hem etin hem de sebzenin tüketim miktarını etkileyecektir. Aynı şekilde kişi sayısındaki bir değişim hem et hem de sebze tüketim miktarına yansiyacaktır. Buna göre, örneğin hanenin geliri 1.000.000TL. arttığında (küçültülmüş değerlerle çalışıldığından) et tüketimi 0,13 birim artmaktadır. Diğer taraftan sebze tüketimindeki artış sadece 0,01'e oranlı olmaktadır. Buradan artan gelirin et tüketimini daha fazla etkilediği sonucuna varılabilir. Kişi sayısındaki 1 birimlik artış et tüketiminde 0,48 birimlik artışa neden olmaktadır. Oysa aynı artış sebze tüketimini 16,27 birim etkilemektedir. Yine buradan kişi sayısındaki artışın sebze tüketiminde daha büyük bir etkiye neden olduğu söylenebilir.

#### ***4.2.2. Hesaplanan Regresyon Katsayılarının Anlamlılık Sınamaları***

Bu bölümde gelir ve hanehalkı kişi sayısı ile et ve sebze tüketimi arasındaki ilişkinin belirlenmesi amacıyla hesaplanan regresyon katsayıları için anlamlılık sınamalarına yer verilecektir. Bu amaçla ilk olarak Wilks Lambda olabilirlik oran testi, daha sonra varyans analizi tekniği ve Parsimony kavramına dayalı varyans analizi tekniğine değinilecektir.

Tahmin değerleri  $X$  veri matrisiyle katsayı tahminleri matrisinin çarpımından elde edilmiştir. Sonuçlar Ek-4'te yer almaktadır. Gözlem değerlerinden tahmin değerlerinin çıkarılmasıyla elde edilen hata matrisi Ek-5'te verilmiştir. Hesaplamalarda Minitab ve MapleV 2 paket programları kullanılmıştır.

#### 4.2.2.1. Katsayıların Wilks Lambda Olabilirlik Oran Testi İle Sınanması

En küçük kareler ya da en çok olabilirlik tekniği ile hesaplanan katsayıların sınanması için hipotezler aşağıdaki gibi kurulmuştur:

$$H_0 : \beta_{ij} = 0$$

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

Uygulama çalışmasının her aşamasında  $\alpha = 0.05$  alınmıştır.

Sınama için gerekli matris değerleri hesaplanmış ve sonuçlar Wilks lambda olabilirlik oran test istatistiğinin bulunmasında kullanılmıştır:

$$\Lambda = \frac{|\hat{\Sigma}|}{|\Sigma|}$$

$$n \hat{\Sigma} = Y'Y - \hat{Y}'\hat{Y}$$

$$n \hat{\Sigma} = \begin{bmatrix} 399,1 & 1474,6 \\ 1474,6 & 4062,2 \end{bmatrix} \quad (4-4)$$

$$n \tilde{\Sigma} = Y'Y - n\bar{y}\bar{y}'$$

$$n \tilde{\Sigma} = \begin{bmatrix} 538,6 & 2029,0 \\ 2024,7 & 53982,1 \end{bmatrix} \quad (4-5)$$

$$\Lambda = \frac{14213478,86}{24966642,76} = 0,5693 \quad (4-6)$$

Hesaplanan Wilks lambda istatistiği aşağıdaki Bartlett test istatistiğinde yazılmış ve şu sonuçlar elde edilmiştir:

$$\begin{aligned}\chi_{\text{hes}}^2 &= - \left[ n - q - 1 - \frac{1}{2}(p - q + 1) \right] \ln \Lambda \\ &= - \left[ 35 - 2 - 1 - \frac{1}{2}(2 - 2 + 1) \right] - 0,5693 \\ &= 17,744 \quad (4-7)\end{aligned}$$

$\chi_{\text{hes}}^2 = 17,744 > \chi_{(4,0,05)}^2 = 9,49$  olduğundan  $H_0$  hipotezi reddedilir. Buna göre et ve sebze tüketim miktarlarını tahmin etmek amacıyla hesaplanan regresyon denklemlerinin istatistiksel olarak anlamlı olduğuna karar verilir.

#### 4.2.2.2. Hesaplanan Katsayıların Varyans Analizi Tekniği İle Sınanması

Bu bölümde hesaplanan  $\hat{\beta}$  katsayıları için varyans analizi yapılacaktır. Böylelikle her iki denklemde bulunan bütün katsayıların istatistiksel olarak anlamlı olup olmadığına karar verilecektir.

Varyans analizi için kurulacak hipotezler bölüm 4.2.2.1.'de sınanan hipotezlerle aynıdır.

Varyans analizi için gerekli hesaplamalar yapılmış ve elde edilen sonuçlar Tablo 4.1'deki varyans analizi tablosunda verilmiştir. Burada kareler toplamları skaler formda değil matris olarak elde edilmiştir:

Tablo 4.1. Çok deęişkenli regresyonda varyans analizi tablosu

Deęişim Kaynaęı	Kareler Toplamı Matrisi	Serbestlik Derecesi
Regresyon	$H = \begin{bmatrix} 139,5 & 554,4 \\ 550,1 & 12919,9 \end{bmatrix}$	2
Hata	$\hat{E}'\hat{E} = \begin{bmatrix} 399,1 & 1474,6 \\ 1474,6 & 41062,2 \end{bmatrix}$	32
Genel	$T = \begin{bmatrix} 538,6 & 2029,0 \\ 2024,7 & 53982,1 \end{bmatrix}$	34

Yukarıdaki matrisler kullanılarak  $(\hat{E}'\hat{E})^{-1}H$  matrisi şöyle bulunmuştur:

$$(\hat{E}'\hat{E})^{-1}H = \begin{bmatrix} -0,345916 & -0,261256 \\ -0,000975 & -0,305261 \end{bmatrix} \quad (4-8)$$

Wilks istatistięi bu matrisin özdeęerleri yardımıyla bulunmaktadır. Özdeęerler aşığıdaki gibidir:

$$\lambda_1 = -0,3514328777 \quad \lambda_2 = -0,2997441223 \quad (4-9)$$

Özdeęerler  $\Lambda$  istatistięinin elde edilmesi için aşığıdaki eşitlikte yerine konur:

$$\Lambda = \prod_{i=1}^P (1 + \lambda_i)^{-1} = (-1,351432877 \times -1,2997441223)^{-1}$$

$$\Lambda = 0,569308486 \quad (4-10)$$

Elde edilen katsayılar bu aşamada aşağıdaki ki-kare istatistiği ile test edilecektir:

$$\begin{aligned}\chi_{\text{hes}}^2 &= \left[ n - q - 1 - \frac{1}{2}(p - q + 1) \right] \ln \Lambda \\ &= \left[ 35 - 2 - 1 - \frac{1}{2}(2 - 2 + 1) \right] - 0,56333283 \\ &= 17,74498433 \quad (4-11)\end{aligned}$$

$$\chi_{\text{hes}}^2 = 17,74498433 > \chi_{(4;0,05)}^2 = 9,49 \quad \text{olduğundan } H_0 \text{ hipotezi reddedilir.}$$

Buna göre et ve sebze tüketim miktarlarını tahmin etmek amacıyla hesaplanan tüm  $\hat{\beta}$  katsayılarının istatistiksel olarak anlamlı olduğuna karar verilir.

#### 4.2.2.3. Hesaplanan Katsayıların Parsimony Kavramına Dayalı Varyans Analizi Tekniği İle Sınanması

Regresyon denkleminin 2 ayrı teknikle anlamlılığı sınıandıktan sonra şimdi indirgenmiş model yardımıyla sınamaya yapılacaktır.

İlk olarak hanehalkı kişi sayısı değişkeninin önemsiz, gelir değişkeninin ise önemli olduğu şeklindeki hipotezler aşağıdaki gibi kurulmuştur:

Simgelerle hipotezler aşağıdaki gibi olur:

$$H_0: \beta_{1j} = 0$$

$$H_1: \beta_{1j} \neq 0$$

İndirgenmiş modele ilişkin yeni  $X_1$  veri matrisinin ilk sütununda sabit terimin de hesaplanmasını sağlamak amacıyla 1 değerlerine yer verilmiştir.

Sözkonusu matris Ek-6'da olduğu gibidir. Buradan hesaplanan tahmin değerleri ise Ek-7'de verilmiştir.

İndirgenmiş denklemin katsayılarının sınanması amacıyla gerekli hesaplamalar yapılmış ve sonuçlar Tablo 4.2'de verilmiştir. Kareler toplamı sütunlarını  $2 \times 2$  boyutlu matrisler oluşturmaktadır.

Tablo 4.2. İndirgenmiş model için varyans analizi tablosu

Değişim Kaynağı	Kareler Toplamı Matrisi	Serbestlik Derecesi
Regresyon		
Tüm model için	$H = \begin{bmatrix} 139,5 & 554,4 \\ 550,1 & 12919,9 \end{bmatrix}$	2
$r_1$ değişken için	$H_{r_1} = \begin{bmatrix} 128,083 & 178,510 \\ 174,229 & 253,250 \end{bmatrix}$	1
$r_2$ değişken için	$H_{r_2/r_1} = \begin{bmatrix} 11,42 & 375,89 \\ 375,87 & 12666,65 \end{bmatrix}$	1
Hata	$\hat{E}'\hat{E} = \begin{bmatrix} 399,1 & 1474,6 \\ 1474,6 & 41062,2 \end{bmatrix}$	32
Genel	$T = \begin{bmatrix} 538,6 & 2029,0 \\ 2024,7 & 53982,1 \end{bmatrix}$	34

İndirgenmiş yeni model için hata terimleri matrisi  $\tilde{E}_1$  Ek.8'de verilmiştir. Özdeğerleri aranan yeni matris aşağıdaki gibidir:

$$(\tilde{E}\tilde{E})^{-1}H_{r_2/r_1} = \begin{bmatrix} -0,004408 & -0,174036 \\ 0,007148 & 0,241745 \end{bmatrix} \quad (4-12)$$

Matrisin özdeğerleri şöyle hesaplanmıştır:



$$\lambda_1 = 0,0007540585 \quad \lambda_2 = 0,2365829415 \quad (4-13)$$

Özdeğerler yardımıyla  $\Lambda$  istatistiği aşağıdaki gibi bulunmuştur:

$$\Lambda = \prod_{i=1}^P (1 + \lambda_i)^{-1} = (1,0007540585 \times 1,2365829415)^{-1}$$

$$\Lambda = 0,808070753 \quad (4-14)$$

İndirgenmiş modelin katsayılarının sınanması amacıyla ki-kare istatistiği aşağıdaki gibi hesaplanmıştır:

$$\begin{aligned} \chi_{\text{hes}}^2 &= - \left[ n - q - 1 - \frac{1}{2}(p - q_1 + 1) \right] \ln 0,808070753 \\ &= - \left[ 35 - 2 - 1 - \frac{1}{2}(2 - 1 + 1) \right] - 0,21310565 \\ &= 6,60627515 \end{aligned} \quad (4-15)$$

$\chi_{\text{hes}}^2 = 6,60627515 > \chi_{(2,0,05)}^2 = 5,99$  olduğundan  $H_0$  reddedilir. Buna göre hanehalkı kişi sayısı değişkeni bağımlı değişkenleri açıklamada etkilidir. Denklemden çıkarılmamalıdır.

İkinci olarak gelir değişkeninin et ve sebze tüketimini açıklamada etkisi olmadığı varsayımıyla ANOVA yapılacaktır. Olaya ilişkin hipotezler aşağıdaki gibidir:

$$H_0: \beta_{2j} = 0$$

$$H_1: \beta_{2j} \neq 0$$

İndirgenmiş model için yeni veri matrisinde ilk sütunda yine sabit terimin de hesaplanması için 1 değerlerine yer verilmiştir. Bu matris Ek-9'da yer almaktadır. Buradan elde edilen tahmin matrisi ise Ek-10'da verilmiştir.

Varyans analizi için hesaplamalar yapılmış ve Tablo 4.3'te yer alan sonuçlar elde edilmiştir. Kareler toplamı sütunu yine matris formundadır.

Tablo 4.3. İndirgenmiş model için varyans analizi tablosu

Değişim Kaynağı	Kareler Toplamı Matrisi	Serbestlik Derecesi
Regresyon		
Tüm model için	$H = \begin{bmatrix} 139,5 & 554,4 \\ 550,1 & 12919,9 \end{bmatrix}$	2
$r_1$ değişken için	$H_{r_1} = \begin{bmatrix} 22,64 & 544,07 \\ 539,79 & 12520,16 \end{bmatrix}$	1
$r_2$ değişken için	$H_{r_2/r_1} = \begin{bmatrix} 116,857 & 10,333 \\ 10,314 & 399,740 \end{bmatrix}$	1
Hata	$\hat{E}'\hat{E} = \begin{bmatrix} 399,1 & 1474,6 \\ 1474,6 & 41062,2 \end{bmatrix}$	32
Genel	$T = \begin{bmatrix} 538,6 & 2029,0 \\ 2024,7 & 53982,1 \end{bmatrix}$	34

Yeni hata matrisi  $\tilde{E}_{(2)}$  Ek-11'de verilmiştir. Özdeğerleri aranan matris ise şöyledir:

$$(\tilde{E}'\tilde{E})^{-1} H_{r_1/r_2} = \begin{bmatrix} 0,251697 & -0,008607 \\ -0,008765 & 0,009949 \end{bmatrix} \quad (4-16)$$

$$\lambda_1 = 0,2520086602 \quad \lambda_2 = 0,009637339829 \quad (4-17)$$

Özdeğerler kullanılarak  $\Lambda$  Wilks Lambda istatistiği şöyle bulunmuştur:

$$\Lambda = \prod_{i=1}^p (1 + \lambda_i)^{-1} = (1,2520086602 \times 1,009637339829)^{-1}$$

$$\Lambda = 0,791092493 \quad (4-18)$$

İndirgenmiş denklemin katsayılarının sınanması için ki-kare istatistiği hesaplanmış ve aşağıdaki sonuca ulaşılmıştır:

$$\begin{aligned} \chi_{\text{hes}}^2 &= - \left[ n - q - 1 - \frac{1}{2}(p - q_1 + 1) \right] \ln 0,791092493 \\ &= - \left[ 35 - 2 - 1 - \frac{1}{2}(2 - 1 + 1) \right] - 0,23434038 \\ &= 7,26455195 \quad (4-19) \end{aligned}$$

$\chi_{\text{hes}}^2 = 7,26455195 > \chi_{(2;0,05)}^2 = 5,99$  olarak hesaplanmıştır. Buna göre hanehalkı geliri değişkeninin olayı açıklayamadığı şeklindeki  $H_0$  hipotezi reddedilir.

#### 4.2.3. Katsayılar İçin Güven Sınırlarının Belirlenmesi

Katsayı tahminleri hesaplandıktan ve denklem için anlamlılık testleri yapıldıktan sonra şimdi katsayılar için güven sınırları belirlenecektir. Hesaplanan noktalar 2 boyutlu bir düzlemde bir elipsi ifade eder.

$$n \hat{\Sigma} = \begin{bmatrix} 399,2468 & 1474,6377 \\ 1474,6377 & 41062,8281 \end{bmatrix} \quad (4-20)$$

$n\hat{\Sigma}$  matrisinin tüm elemanlarının  $n$ 'e oranlanmasıyla  $\hat{\Sigma}$  matrisi hesaplanmıştır.

$$\hat{\Sigma} = \begin{bmatrix} 11,41 & 42,13 \\ 42,13 & 1173,22 \end{bmatrix} \quad (4-21)$$

Matrisin köşegen elemanlarından  $\hat{\sigma}_{11} = 11,41$  ve  $\hat{\sigma}_{22} = 1173,22$  elde edilir. Diğer taraftan  $X$  veri matrisinin ortalama değerleri olan 3,74 ve 20,89 kullanılarak aşağıdaki sonuçlar elde edilmiştir:

$$X_0'(XX)X_0 = 0,0286 \quad (4-22)$$

Diğer taraftan güven sınırlarının belirlenmesinde aşağıdaki matrisler kullanılmıştır:

$$X_0 = \begin{bmatrix} 1 \\ 3,74 \\ 20,89 \end{bmatrix} \quad (4-23)$$

$$\hat{\beta}_{(1)} = \begin{bmatrix} 2,2328 \\ 0,4817 \\ 0,1337 \end{bmatrix} \quad (4-24)$$

$$\hat{\beta}_{(2)} = \begin{bmatrix} 17,6543 \\ 16,2709 \\ 0,0104 \end{bmatrix} \quad (4-25)$$

$$\begin{aligned}
& X_0' \hat{\beta}_{(1)} \pm \sqrt{\left(\frac{p(n-q-1)}{n-q-p}\right) F_{q,n-q-p}(\alpha)} \sqrt{(X_0' (X'X) X_0) \left(\frac{n}{n-q-1}\right) \hat{\sigma}_{ii}} \\
& = 6,83 \pm \sqrt{\left(\frac{2(35-3)}{31}\right) F_{2,31;(0,05)}} \sqrt{0,0286 \left(\frac{35}{32}\right) 11,41} \\
& = 6,83 \pm 2,62 \times 0,59 \tag{4-26}
\end{aligned}$$

Buradan  $P(5,27 \leq \mu \leq 8,39) = 0,95$  olarak hesaplanır.

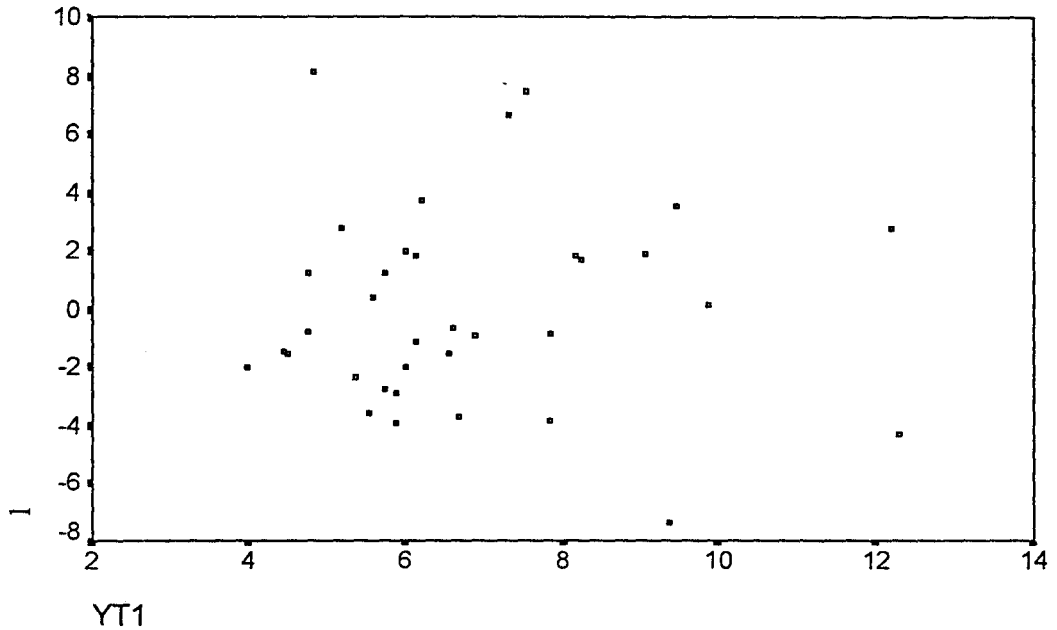
$$\begin{aligned}
& X_0' \hat{\beta}_{(2)} \pm \sqrt{\left(\frac{p(n-q-1)}{n-q-p}\right) F_{q,n-q-p}(\alpha)} \sqrt{(X_0' (X'X) X_0) \left(\frac{n}{n-q-1}\right) \hat{\sigma}_{ii}} \\
& = 78,77 \pm \sqrt{\left(\frac{2(35-3)}{31}\right) F_{2,31;(0,05)}} \sqrt{0,0286 \left(\frac{35}{32}\right) 1173,22} \\
& = 78,77 \pm 2,62 \times 6,06 \tag{4-27}
\end{aligned}$$

Buradan da  $P(62,89 \leq \mu \leq 94,64) = 0,95$  olarak hesaplanır. Tek bağımlı değişken için güven aralığı hesaplandığında bulunan aralık merkezi ortalama olan tek bir doğru gibi düşünülebilir. Oysa çok değişkenli regresyon analizinde bağımlı değişkenin tahmin değerleri için hesaplanan bu değerler merkezi ortalama değerleri 6,83 ve 78,77 olan bir elipsi ifade eder. Bu nedenle artık bulunan değerler güven aralığı olarak adlandırılmaz, güven elipsoidi olarak adlandırılır.

#### 4.2.4. Hesaplanan Artıklarla Model Uygunluğunun İncelenmesi

Model uygunluğunun incelenmesi amacıyla aşağıdaki şekiller çizilmiş ve yorumlanmıştır. Bu bölümdeki tüm şekiller SPSS paket programının 5.0 versiyonunda çizdirilmiştir.

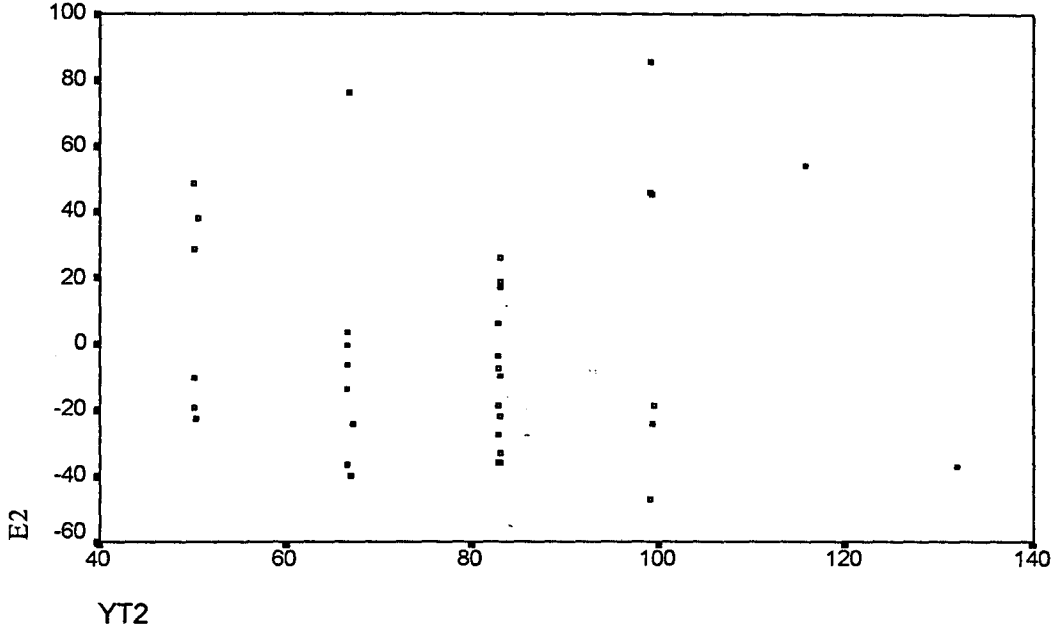
Aşağıda yer alan Şekil 4.1, hanehalkı kişi sayısı ve gelir bağımsız değişkenleri ile tahmin edilen et tüketim miktarı ve bu tahminlere karşı gelen artık değerleri için çizilmiştir.



Şekil 4.1. Tahminlerin hata terimlerine göre çizimi

Şekil 4.1'de uç değerler bulunmasına rağmen artıklar sağa megafon, birlikte artan ya da azalan bir eğilim göstermemektedirler.

Şekil 4.2 hanehalkı kişi sayısı ve gelir bağımsız değişkenleri ile tahmin edilen aylık sebze tüketim miktarı ve bu tahminlere karşı gelen artık değerlerini göstermektedir.



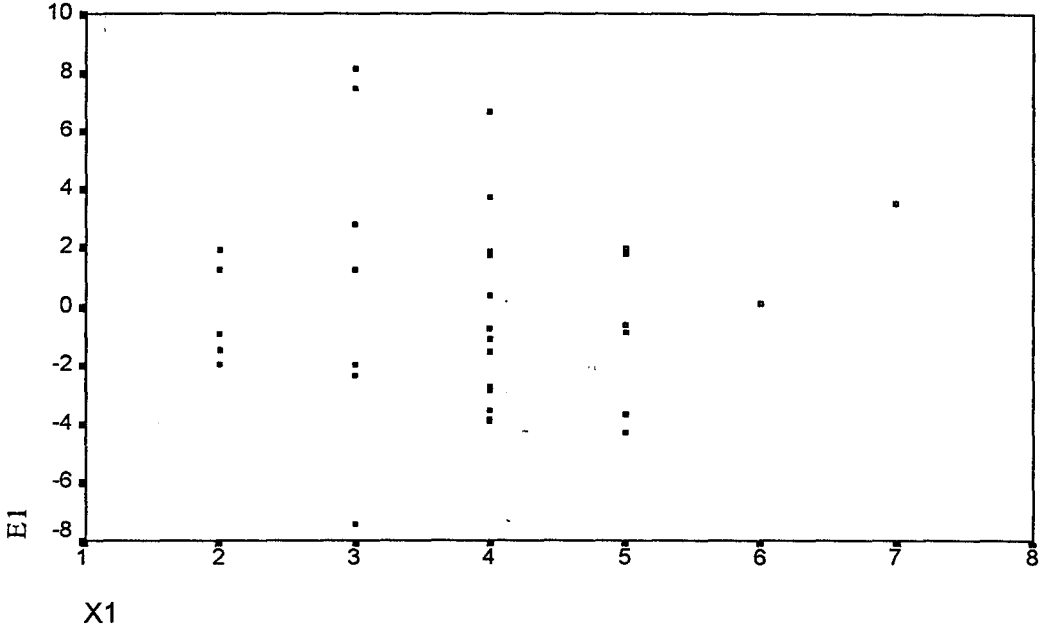
Şekil 4.2. Tahminlerin hata terimlerine göre çizimi

Yukarıdaki Şekil 4.2’de artıkların dağılımı, Şekil 4.1’de olduğundan daha düzgündür.

Tahmin değerlerinin artıklara karşı çizildiği her iki şekilde de artıklar sağa megafon, sürekli artan bir seyir sergilememektedir. Dolayısıyla Şekil 4.1 ve Şekil 4.2’ye göre modelin uygun olduğu söylenebilir.

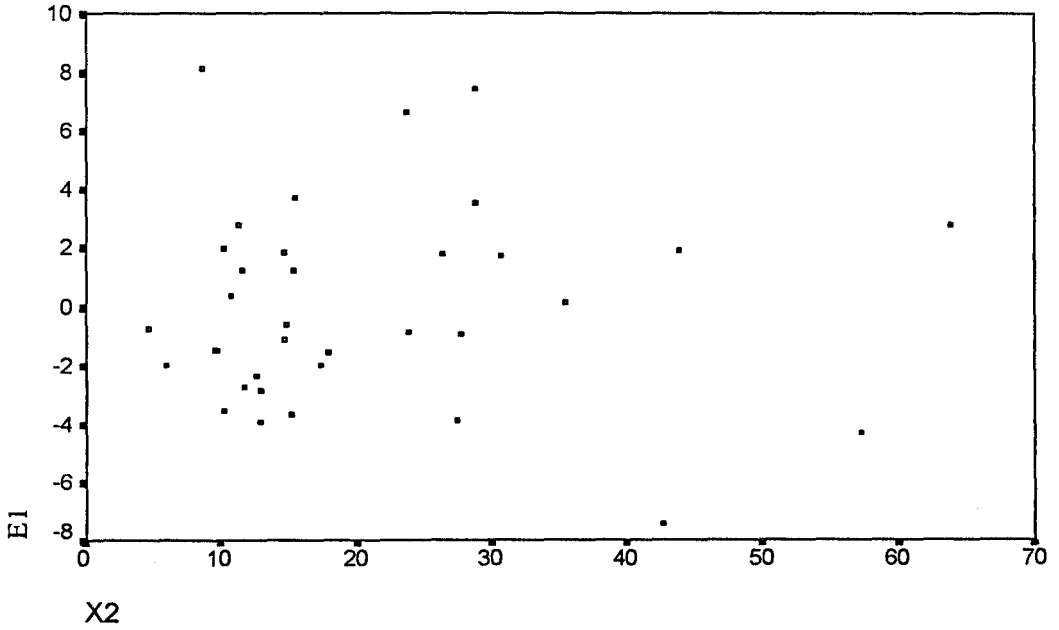
Bundan sonra yer verilen şekiller bağımsız değişken gözlem değerleri ile bunlara karşı gelen artık değerleri için çizilmiştir.

Şekil 4.3 hanehalkı kişi sayısı bağımsız değişkeni ile buna karşı gelen aylık et tüketim miktarı bağımlı değişkeni artık değerlerini göstermektedir.



Şekil 4.3. Artık değerlerinin bağımsız değişken değerlerine karşı çizimi

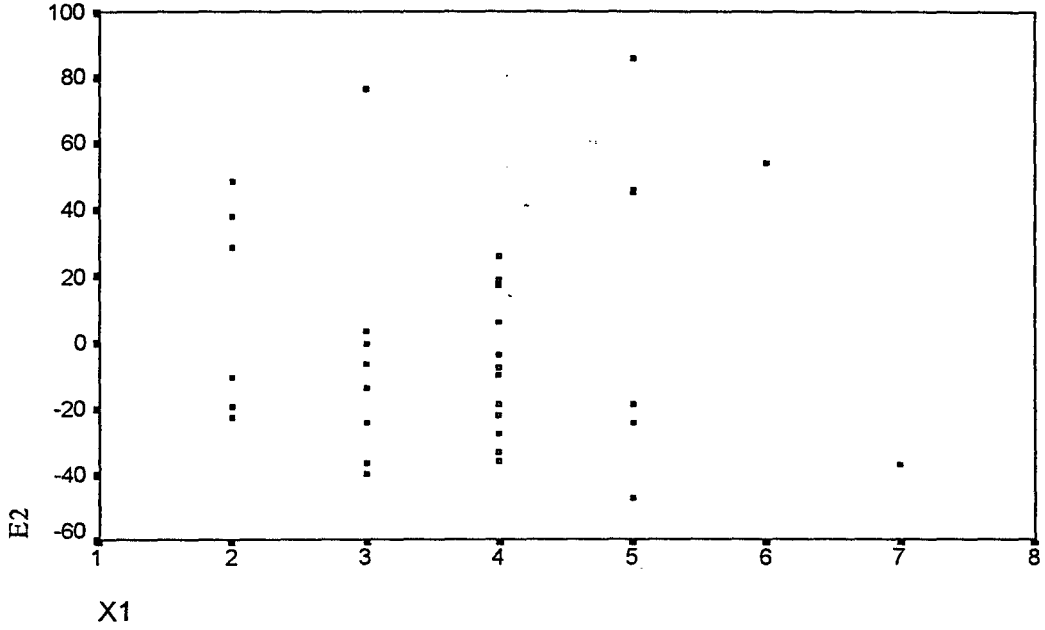
Şekil 4.4'de gelir bağımsız değişken değerleri yatay, aylık et tüketim miktarı artık değerleri dikey eksenini göstermektedir.



Şekil 4.4. Artık değerlerinin bağımsız değişken değerlerine karşı çizimi

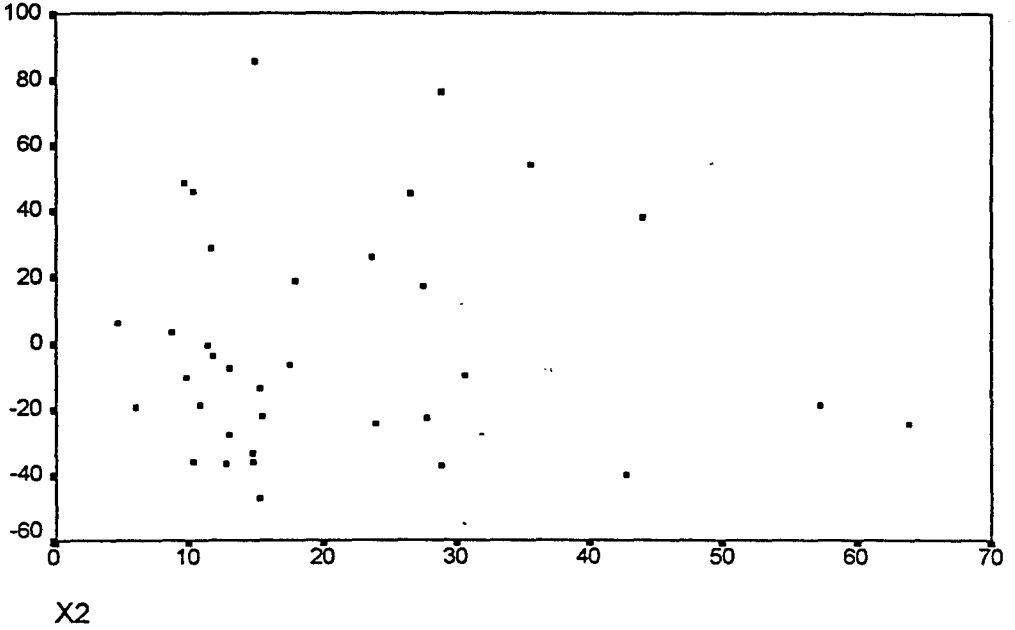


Aşağıdaki Şekil 4.5'te yatay ekseni hanehalkı kişi sayısı bağımsız değişkeni gözlem değerleri oluşturmaktadır. Dikey eksen ise sebze tüketimi bağımlı değişkeni için hesaplanmış artık değerlerini göstermektedir.



Şekil 4.5. Artık değerlerinin bağımsız değişken değerlerine karşı çizimi

Şekil 4.6, gelir bağımsız değişkeni değerleri yatay, tüketilen sebze miktarı tahmin değerlerine göre hesaplanmış artıklar ise dikey ekseninde yer almak üzere çizilmiştir.



Şekil 4.6. Artık değerlerinin bağımsız değişken değerlerine karşı çizimi

4 şekilde de artıklar sürekli artan yada azalan bir formda görünmemektedirler. Diğer taraftan u ya da ters u şeklinde değildirler. Beraber dalgalanma da sözkonusu değildir.

Sonuç olarak gerek artıkların tahmin değerleri için yapılmış çizimlerinde gerekse yine artıkların gözlem değerleri için yapılmış çizimlerinde modelin uygun olduğuna karar verilebilir.

## 5. SONUÇ VE ÖNERİLER

Çok değişkenli regresyon analizinin uygulandığı bu çalışmada aylık et tüketim miktarı ( $Y_1$ ), aylık sebze tüketim miktarı ( $Y_2$ ) ile bağımsız değişkenler hanehalkı kişi sayısı ( $X_1$ ) ve gelir ( $X_2$ ) arasındaki ilişki ortaya konulmaya çalışılmıştır. Bu amaçla hesaplanan doğrusal regresyon denklemleri varyans analiziyle test edilmiş ve denklemlerin anlamlı olduğuna karar verilmiştir. Ayrıca modelin uygun olup olmadığı gözlenilmek istenmiş ve bu amaçla bazı grafikler çizilmiştir. Sonuçta modelin uygun olduğuna karar verilmiştir.

Bu çalışmada olduğu gibi iki veya daha fazla sayıda bağımsız değişkenin aynı anda birden fazla sayıda bağımlı değişkeni etkilediğinin düşünüldüğü durumlarda, araştırmacılar bağımlı değişkenler için denklemleri ayrı ayrı tahmin etmek yerine çok değişkenli regresyon analizini kullanarak tüm katsayıları aynı anda hesaplamalıdır. Böylece değişmeyen veri setiyle aynı anda çok sayıda bağımlı değişken tahmin edilebilir. Yine ANOVA ile denklemlerin tümü sınanabilir.

Bu çalışmada D.İ.E.'den alınan 1994 yılının tek bir ayına ait veriler kullanılmıştır. Dolayısıyla bir aya ilişkin yatay kesit verilerden yararlanılmıştır. Çalışmada gelir ve temel besin ürünlerinden et ve sebze tüketimi ile ilgilenildiğinden çalışmanın ekonomiyle yakından ilişkisi bulunmaktadır.

Türkiye'de yaşanan enflasyon gerçeğini de gözlemlemek amacıyla araştırmacılar daha uzun süreyi kapsayan veri setini kullanarak (en az 1 yıl) çok değişkenli regresyon analiziyle birlikte zamanın ve enflasyonun etkisini de gözlemleyebilirler. Araştırmacılara başka tüketim malları ve bağımsız değişkenler için bu tür bir çalışma önerilebilir. Böylece araştırmacılar ilişki analizi yapmanın yanı sıra enflasyonun etkisiyle bu malların satın alınan miktarlarında bir değişim olup olmadığını da inceleyebilirler. Aynı zaman sürecinde kişilerin gelirleri de farklılaşacaktır. Yine gelirden gözlemlenen artışların mal alımına nasıl yansıdığı da aynı çalışmayla gözlemlenebilir.

## KAYNAKLAR

1. D.İ.E. YAYINI, *Kentsel Yerler Hanehalkı İşgücü Anket Sonuçları*, Yayın No:1159, 1983.
2. YÜKSEL, B., *Hayat Pahalılığı Sorununun Çözümünde Tüketim Kooperatiflerinin Rolü*, Afyon İktisadi İdari Bilimler Fakültesi Yayınları No:5, Anadolu Üniversitesi Yayınları No:365;1989.
3. D.İ.E. YAYINI, *Türkiye İstatistik Yıllığı*, Yayın No:1845, 1995.
4. CEMALCILAR, İ., *Pazarlama Kavramlar - Kararlar*, Beta Basım Yayın A.Ş., 1988.
5. ODABAŞI, Y., *Tüketici Davranışları ve Pazarlama Stratejisi*, Anadolu Üniversitesi Yayınları No:908, İşletme Fakültesi Yayınları No:2, 1996.
6. TUSİAD ARAŞTIRMALARI SONUCU, *Türkiye 'de Sosyo-ekonomik Öncelikler Hane Gelirleri, Harcamaları ve Sosyo-ekonomik İhtiyaçlar Üzerine Araştırma Dizisi, Türk Halkının Sosyo-ekonomik İhtiyaçları*, Yayın No: TUSİAD-T/86.92, 1986.
7. TUSİAD ARAŞTIRMALARI SONUCU, *Türkiye 'de Sosyo-ekonomik Öncelikler Hane Gelirleri, Harcamaları ve Sosyo-ekonomik İhtiyaçlar Üzerine Araştırma Dizisi, Türkiye 'de Hane Geliri Hanehalkı Harcamaları ve Hayat Standardı*, Yayın No: TUSİAD-T/86.92, 1986.
8. SERPER, Ö., *Uygulamalı İstatistik 2*, Filiz Kitabevi, İstanbul, 1993.

**KAYNAKLAR (devam)**

9. ÇÖMLEKÇİ, N., *Temel İstatistik İlke ve Teknikleri*, Bilim Teknik Yayınevi, Eskişehir, 1989.
10. SPANOS, A., *Statistical Foundations of Econometric Modelling*, Cambridge University Press, 1986.
11. JOHNSON, R.A., WICHERN, D.W., *Applied Multivariate Statistical Analysis*, Prentice Hall International, Inc. 2nd Edition, 1988.
12. TATLIDİL, H., *Uygulamalı Çok Değişkenli İstatistiksel Analiz*, Hacettepe Üniversitesi, 1992.
13. GUJARATI, D.N., *Basic Econometrics*, 3th. Edition, Mc Graw Hill International Editions, 1988.
14. AMEMIYA, T., *Advanced Econometrics*, Basil Blackwell Ltd., 1985.
15. GREENE, W.H., *Econometric Analysis*, New York University, Macmillan Publishing Company, 1990.
16. MORRISON, D.F., *Multivariate Statistical Methods*, Mc Graw Hill Book Company, 1, 1978.
17. KARA, İ., *Olasılık*, T.C. Anadolu Üniversitesi, Eğitim, Sağlık ve Bilimsel Araştırma Çalışmaları Vakıf Yayınları, No:69, Eskişehir, 1989.
18. KSHIRSAGAR, A.M., *Multivariate Analysis*, Marcel Dekker, Inc., New York, 1972.

**KAYNAKLAR (devam)**

19. KLEINBAUM, D.G., KUPPER, L.L., MULLER, K.E., *Applied Regression Analysis And Other Multivariate Methods*, 493, PWS-KENT Publishing Company, Boston, 1987.
20. D.İ.E. YAYINI, *1994 Hanehalkı Gelir ve Tüketim Harcamaları Soru Kağıdı (Aylık)*, 1994.

## EK-1 Uygulama çalışmasında kullanılan veriler

Hane No	Kişi Sayısı	Gelir	Et Tüketimi	Sebze Tüketimi
1	4	30.676.000	10	73
2	4	27.464.000	4	100
3	2	6.000.000	2	31
4	3	8.666.000	13	70
5	5	10.250.000	8	145
6	4	14.800.000	5	50
7	4	15.541.000	10	61
8	6	35.540.000	10	170
9	3	11.428.000	8	66
10	3	15.433.000	7	53
11	4	10.250.000	2	47
12	2	9.800.000	3	40
13	4	13.025.000	3	55
14	5	15.340.000	3	52
15	4	12.995.000	2	75
16	4	11.800.000	3	79
17	4	18.036.000	5	102
18	3	42.750.000	2	27
19	5	26.536.000	10	145
20	2	27.775.000	6	28
21	4	4.600.000	4	89
22	3	12.700.000	3	30
23	5	23.950.000	7	75
24	4	14.750.000	8	47
25	4	23.700.000	14	109
26	4	10.835.000	6	64
27	5	14.975.000	6	185
28	5	57.275.000	8	81
29	2	11.723.000	6	79
30	7	28.892.000	13	95
31	3	17.498.000	4	60
32	2	44.000.000	11	89
33	3	63.875.000	15	43
34	3	28.947.000	15	143
35	2	9.625.000	3	99

*EK-2 X Veri Matrisi*

$$X = \begin{bmatrix} 1 & 4 & 30.7 \\ 1 & 4 & 27.5 \\ 1 & 2 & 6 \\ 1 & 3 & 8.7 \\ 1 & 5 & 10.3 \\ 1 & 4 & 14.8 \\ 1 & 4 & 15.5 \\ 1 & 6 & 35.5 \\ 1 & 3 & 11.4 \\ 1 & 3 & 15.4 \\ 1 & 4 & 10.3 \\ 1 & 2 & 9.8 \\ 1 & 4 & 13 \\ 1 & 5 & 15.3 \\ 1 & 4 & 13 \\ 1 & 4 & 11.8 \\ 1 & 4 & 18 \\ 1 & 3 & 42.8 \\ 1 & 5 & 26.5 \\ 1 & 2 & 27.8 \\ 1 & 4 & 4.6 \\ 1 & 3 & 12.7 \\ 1 & 5 & 23.9 \\ 1 & 4 & 14.8 \\ 1 & 4 & 23.7 \\ 1 & 4 & 10.8 \\ 1 & 5 & 14.9 \\ 1 & 5 & 57.3 \\ 1 & 2 & 11.7 \\ 1 & 7 & 28.9 \\ 1 & 3 & 17.5 \\ 1 & 2 & 44 \\ 1 & 3 & 63.9 \\ 1 & 3 & 28.9 \\ 1 & 2 & 9.6 \end{bmatrix}$$



*EK-3 Y* Veri Matrisi

$$Y = \begin{bmatrix} 10 & 73 \\ 4 & 100 \\ 2 & 31 \\ 13 & 70 \\ 8 & 145 \\ 5 & 50 \\ 10 & 61 \\ 10 & 170 \\ 8 & 66 \\ 7 & 53 \\ 2 & 47 \\ 3 & 40 \\ 3 & 55 \\ 3 & 52 \\ 2 & 75 \\ 3 & 79 \\ 5 & 102 \\ 2 & 27 \\ 10 & 145 \\ 6 & 28 \\ 4 & 89 \\ 3 & 30 \\ 7 & 75 \\ 8 & 47 \\ 14 & 109 \\ 6 & 64 \\ 6 & 185 \\ 8 & 81 \\ 6 & 79 \\ 13 & 95 \\ 4 & 60 \\ 11 & 89 \\ 15 & 43 \\ 15 & 143 \\ 3 & 99 \end{bmatrix}$$

**EK-4**  $\hat{Y}$  Tahmin Matrisi
$$\hat{Y} = \begin{bmatrix} 8.26 & 83.06 \\ 7.84 & 83.02 \\ 3.99 & 50.26 \\ 4.84 & 66.56 \\ 6.02 & 99.12 \\ 6.14 & 82.89 \\ 6.23 & 82.89 \\ 9.87 & 115.65 \\ 5.20 & 66.59 \\ 5.74 & 66.63 \\ 5.54 & 82.85 \\ 4.51 & 50.29 \\ 5.89 & 82.87 \\ 6.69 & 99.17 \\ 5.89 & 82.87 \\ 5.74 & 82.86 \\ 6.57 & 82.93 \\ 9.39 & 66.91 \\ 8.18 & 99.29 \\ 6.91 & 50.49 \\ 4.77 & 82.79 \\ 5.38 & 66.59 \\ 7.84 & 99.26 \\ 6.14 & 82.89 \\ 7.33 & 82.98 \\ 5.60 & 82.85 \\ 6.63 & 99.16 \\ 12.3 & 99.61 \\ 4.76 & 50.32 \\ 9.47 & 131.85 \\ 6.02 & 66.65 \\ 9.08 & 50.65 \\ 12.2 & 67.13 \\ 7.54 & 66.77 \\ 4.48 & 50.29 \end{bmatrix}$$

**EK-5** Hata Matrisi

$$\hat{E} = \begin{bmatrix} 1.74 & -10.06 \\ -3.84 & 16.98 \\ -1.99 & -9.26 \\ 8.16 & 3.44 \\ 1.98 & 45.88 \\ -1.14 & -32.89 \\ 3.77 & -21.89 \\ 0.13 & 54.35 \\ 2.80 & -0.59 \\ 1.26 & -13.63 \\ -3.54 & -35.85 \\ -1.51 & -10.29 \\ -2.89 & -27.87 \\ -3.69 & -47.17 \\ -3.89 & -7.87 \\ -2.74 & -3.86 \\ -1.57 & 19.07 \\ -7.39 & -39.91 \\ 1.82 & 45.71 \\ -0.91 & -22.49 \\ -0.77 & 6.21 \\ -2.38 & -36.59 \\ -0.84 & -24.26 \\ 1.86 & -35.89 \\ 6.67 & 26.02 \\ 0.40 & -18.85 \\ -0.63 & 85.84 \\ -4.30 & -18.61 \\ 1.24 & 28.68 \\ 3.53 & -36.85 \\ -2.02 & -6.65 \\ 1.92 & 38.35 \\ 2.78 & -24.13 \\ 7.46 & 76.23 \\ -1.48 & 48.71 \end{bmatrix}$$

**EK-6** İndirgenmiş Veri Matrisi

$$\tilde{X}_{(1)} = \begin{bmatrix} 1 & 30.7 \\ 1 & 27.5 \\ 1 & 6 \\ 1 & 8.7 \\ 1 & 10.3 \\ 1 & 14.8 \\ 1 & 15.5 \\ 1 & 35.5 \\ 1 & 11.4 \\ 1 & 15.4 \\ 1 & 10.3 \\ 1 & 9.8 \\ 1 & 13 \\ 1 & 15.3 \\ 1 & 13 \\ 1 & 11.8 \\ 1 & 18 \\ 1 & 42.8 \\ 1 & 26.5 \\ 1 & 27.8 \\ 1 & 4.6 \\ 1 & 12.7 \\ 1 & 23.9 \\ 1 & 14.8 \\ 1 & 23.7 \\ 1 & 10.8 \\ 1 & 14.9 \\ 1 & 57.3 \\ 1 & 11.7 \\ 1 & 28.9 \\ 1 & 17.5 \\ 1 & 44 \\ 1 & 63.9 \\ 1 & 28.9 \\ 1 & 9.6 \end{bmatrix}$$

**EK-7 İndirgenmiş Modele İlişkin Tahmin Matrisi**

$$\tilde{Y}_{(1)} = \begin{bmatrix} 8.19 & 80.67 \\ 7.75 & 80.05 \\ 4.76 & 75.89 \\ 5.13 & 76.41 \\ 5.36 & 76.72 \\ 5.98 & 77.59 \\ 6.08 & 77.73 \\ 8.86 & 81.58 \\ 5.51 & 76.93 \\ 6.06 & 77.71 \\ 5.34 & 76.72 \\ 5.29 & 76.62 \\ 5.73 & 77.24 \\ 6.05 & 77.69 \\ 5.73 & 77.24 \\ 5.56 & 77.01 \\ 6.43 & 78.21 \\ 9.88 & 83.01 \\ 7.61 & 79.86 \\ 7.79 & 80.11 \\ 4.56 & 75.62 \\ 5.69 & 77.19 \\ 7.25 & 79.35 \\ 5.98 & 77.59 \\ 7.22 & 79.31 \\ 5.43 & 76.82 \\ 5.99 & 77.61 \\ 11.9 & 85.82 \\ 5.55 & 76.99 \\ 7.94 & 80.32 \\ 6.36 & 78.11 \\ 10.1 & 83.24 \\ 12.8 & 87.09 \\ 7.94 & 80.32 \\ 5.26 & 76.59 \end{bmatrix}$$

**EK-8 İndirgenmiş Model İçin Hesaplanmış Artıklar**

1.81	- 7.67
-3.75	19.95
-2.76	- 44.89
7.87	- 6.41
2.64	68.28
-0.98	- 27.59
3.92	- 16.73
1.14	88.40
2.49	- 10.93
0.94	- 24.71
-3.36	- 29.72
-2.29	- 36.62
-2.73	- 22.24
-3.05	- 25.69
-3.73	- 2.24
-2.56	1.99
-1.43	23.79
-7.88	- 56.01
2.39	65.14
-1.79	- 52.11
-0.56	13.38
-2.69	- 47.19
-0.25	- 4.35
2.02	- 30.59
6.78	29.69
0.57	- 12.82
0.01	107.39
-3.89	- 4.82
0.45	2.01
5.06	14.68
-2.36	- 18.11
0.96	5.76
2.19	- 44.09
7.06	62.68
-2.26	22.42

 $\tilde{E}_{(1)} =$

**EK-9 İndirgenmiş Veri Matrisi**

$$\tilde{X}_{(2)} = \begin{bmatrix} 1 & 4 \\ 1 & 4 \\ 1 & 2 \\ 1 & 3 \\ 1 & 5 \\ 1 & 4 \\ 1 & 4 \\ 1 & 6 \\ 1 & 3 \\ 1 & 3 \\ 1 & 4 \\ 1 & 2 \\ 1 & 4 \\ 1 & 5 \\ 1 & 4 \\ 1 & 4 \\ 1 & 4 \\ 1 & 3 \\ 1 & 5 \\ 1 & 2 \\ 1 & 4 \\ 1 & 3 \\ 1 & 5 \\ 1 & 4 \\ 1 & 4 \\ 1 & 4 \\ 1 & 4 \\ 1 & 5 \\ 1 & 5 \\ 1 & 2 \\ 1 & 7 \\ 1 & 3 \\ 1 & 2 \\ 1 & 3 \\ 1 & 3 \\ 1 & 2 \end{bmatrix}$$

**EK-10 İndirgenmiş Modele İlişkin Tahmin Matrisi**

$$\tilde{Y}_{(2)} =$$

7.02	85.42
7.02	85.72
5.64	55.90
6.33	70.66
7.71	100.18
7.02	85.42
7.02	85.42
8.40	114.94
6.33	70.66
6.33	70.66
7.02	85.42
5.64	55.90
7.02	85.42
7.71	100.18
7.02	85.42
7.02	85.42
7.02	85.42
6.33	70.66
7.71	100.18
5.64	55.90
7.02	85.42
6.33	70.66
7.71	100.18
7.02	85.42
7.02	85.42
7.02	85.42
7.71	100.18
7.71	100.18
5.64	55.90
9.09	129.70
6.33	70.66
5.64	55.90
6.33	70.66
6.33	70.66
5.64	55.90



*EK-11* İndirgenmiş Model İçin Hesaplanmış Artıklar

	2.98	-12.42
	-3.02	14.58
	-3.64	-24.90
	6.67	-0.66
	0.29	44.82
	-2.02	-35.42
	2.98	-24.42
	1.60	55.06
	1.67	-4.66
	0.67	-17.66
	-5.02	-38.42
	-2.64	-15.90
	-4.02	-30.42
	-4.71	-48.18
	-5.02	-10.42
	-4.02	-6.42
	-2.02	16.58
$\tilde{E}_{(2)} =$	-4.33	-43.66
	2.29	44.82
	0.36	-27.90
	-3.02	3.58
	-3.33	-40.66
	-0.71	-25.18
	0.98	-38.42
	6.98	23.58
	-1.02	-21.42
	-1.71	84.82
	0.29	-19.18
	0.36	23.10
	3.91	-34.70
	-2.33	-10.66
	5.36	33.10
	8.67	-27.66
	8.67	72.34
	-2.64	43.10