

**PARAMETRİK OLMAYAN REGRESYON  
MODELLERİNDEN ÇEKİRDEK REGRESYON  
VE TÜRKİYE'DE HANEHALKI HARCAMALARI  
ÜZERİNE BİR UYGULAMA**

**Elif GÖLVEREN**

**Yüksek Lisans Tezi  
İşletme Anabilim Dalı  
Doç. Dr. Mehmet Suphi ÖZÇOMAK  
2012  
Her hakkı saklıdır**

**ATATÜRK ÜNİVERSİTESİ**  
**SOSYAL BİLİMLER ENSTİTÜSÜ**  
**İŞLETME ANABİLİM DALI**

**Elif GÖLVEREN**

**PARAMETRİK OLMAYAN REGRESYON MODELLERİNDEN**  
**ÇEKİRDEK REGRESYON VE TÜRKİYE'DE HANEHALKI**  
**HARCAMALARI ÜZERİNE BİR UYGULAMA**

**YÜKSEK LİSANS TEZİ**

**TEZ YÖNETİCİSİ**  
**Doç. Dr. Mehmet Suphi ÖZÇOMAK**

**ERZURUM-2012**



T.C.  
ATATÜRK ÜNİVERSİTESİ  
SOSYAL BİLİMLERİ ENSTİTÜSÜ



**TEZ BEYAN FORMU**

05/07/2012

**SOSYAL BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜNE**

**BİLDİRİM**

Atatürk Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğine göre hazırlamış olduğum "Parametrik Olmayan Regresyon Modellerinden Çekirdek Regresyon Ve Türkiye'de Hanehalkı Harcamaları Üzerine Bir Uygulama " adlı tezin/raporun tamamen kendi çalışmam olduğunu ve her alıntıya kaynak gösterdiğimi taahhüt eder, tezimin/raporumun kağıt ve elektronik kopyalarının Atatürk Üniversitesi Sosyal Bilimler Enstitüsü arşivlerinde aşağıda belirttiğim koşullarda saklanmasına izin verdiğimi onaylarım:

Lisansüstü Eğitim-Öğretim yönetmeliğinin ilgili maddeleri uyarınca gereğinin yapılmasını arz ederim.

Tezimin/Raporumun tamamı her yerden erişime açılabilir.

Tezim/Raporum sadece Atatürk Üniversitesi yerleşkelerinden erişime açılabilir.

Tezimin/Raporumun ..... yıl süreyle erişime açılmasını istemiyorum. Bu sürenin sonunda uzatma için başvuruda bulunmadığım takdirde, tezimin/raporumun tamamı her yerden erişime açılabilir.

05/07/2012

Elif GÖLVEREN



T.C.  
ATATÜRK ÜNİVERSİTESİ  
SOSYAL BİLİMLER ENSTİTÜSÜ



TEZ KABUL TUTANAĞI

SOSYAL BİLİMLER ENSTİTÜSÜ MÜDÜRLÜĞÜNE

Doç. Dr. Mehmet Suphi ÖZÇOMAK danışmanlığında, Elif GÖLVEREN tarafından hazırlanan bu çalışma 05/07/2012 tarihinde aşağıdaki jüri tarafından İşletme Anabilim Dalı'nda Yüksek Lisans Tezi olarak kabul edilmiştir.

**Başkan** : Prof. Dr. Erkan OKTAY

**Jüri Üyesi** : Doç. Dr. Mehmet Suphi ÖZÇOMAK

**Jüri Üyesi** : Yrd. Doç. Dr. Emrah TALAŞ

İmza: .....

İmza: .....

İmza: .....

Yukarıdaki imzalar adı geçen öğretim üyelerine aittir. 05/07/2012

Prof. Dr. Mustafa YILDIRIM  
Enstitü Müdürü

F-85/00/22.02.2012

## İÇİNDEKİLER

ÖZET.....	III
ABSTRACT .....	IV
KISALTMALAR VE SİMGELER DİZİNİ .....	V
ŞEKİLLER DİZİNİ .....	VI
TABLOLAR DİZİNİ .....	VII
ÖNSÖZ.....	VIII
GİRİŞ .....	1

### BİRİNCİ BÖLÜM

#### REGRESYON ANALİZİ VE REGRESYON ANALİZİNE YAKLAŞIMLAR

1.1. REGRESYON ANALİZİ .....	3
1.1.1. Parametrik Regresyon Analizi .....	4
1.1.2. Parametrik Olmayan Regresyon Analizi .....	5

### İKİNCİ BÖLÜM

#### PARAMETRİK OLMAYAN REGRESYON KESTİRİCİLERİ

2.1. ÇEKİRDEK KESTİRİCİLERİ .....	8
2.1.1. Yerel Sabit Çekirdek Kestiricisi.....	8
2.1.2. Yerel Doğrusal Çekirdek Kestiricisi.....	9
2.1.3. Yerel Polinomial Çekirdek Kestiricisi .....	10
2.1.4. Gasser-Müller Çekirdek Kestiricisi.....	12
2.1.5. Priestley-Chao Çekirdek Kestiricisi .....	12
2.2. ÇEKİRDEK KESTİRİCİLERİNİN ÖZELLİKLERİ .....	13
2.2.1. Çekirdek Fonksiyonu .....	13
2.2.2. Pencere Genişliği .....	13
2.3. PENCERE GENİŞLİĞİ SEÇİCİLERİ.....	15
2.3.1. Çapraz Geçerlilik Yöntemi .....	16
2.3.1.1. En Küçük Kareler Çapraz Geçerlilik Yöntemi .....	20
2.3.1.2. Beklenen Kullback-Leibler Çapraz Geçerlilik Yöntemi .....	20
2.3.2. Cezalandırma Fonksiyonları .....	21
2.3.3. Plug-in Yöntemi .....	22
2.4. DİĞER PARAMETRİK OLMAYAN REGRESYON KESTİRİCİLERİ .....	22
2.4.1. En Yakın Komşuluk Kestiricisi (k-NN).....	22

2.4.2. Spline Modelleri.....	23
2.4.3. Ortogonal Seriler .....	25
2.4.4. Medyan Düzleştirme Tekniđi .....	26

## ÜÇÜNCÜ BÖLÜM

### TÜRKİYE'DE ZARURİ HARCAMA İLE GELİR ARASINDAKİ İLİŞKİNİN ÇEKİRDEK REGRESYON TEKNİĐİ İLE İNCELENMESİ

3.1. LİTERATÜR ÖZETİ .....	27
3.2. VERİ SETİ.....	30
3.3. TAHMİNLER.....	30
3.4. PARAMETRİK VE PARAMETRİK OLMAYAN TAHMİNLERİN KARŞILAŞTIRILMASI .....	40
TARTIŞMA VE SONUÇ.....	43
KAYNAKÇA .....	46
EKLER.....	51
ÖZGEÇMİŞ.....	57

**ÖZET**

**YÜKSEK LİSANS TEZİ**

**PARAMETRİK OLMAYAN REGRESYON MODELLERİNDEN ÇEKİRDEK  
REGRESYON VE TÜRKİYE'DE HANEHALKI HARCAMALARI ÜZERİNE  
BİR UYGULAMA**

**Elif GÖLVEREN**

**Tez Danışmanı: Doç. Dr. Mehmet Suphi ÖZÇOMAK**

**2012, 57 Sayfa**

**Jüri: Doç. Dr. Mehmet Suphi ÖZÇOMAK (Danışman)**

**Prof. Dr. Erkan OKTAY**

**Yrd. Doç. Dr. Emrah TALAŞ**

Bu çalışmada, parametrik olmayan regresyon kestiricilerinin incelenmesi ve Türkiye'de hanehalkı zaruri harcamaları (gıda ve alkolsüz içecekler, sağlık ve eğitim) ile hanehalkı kullanılabilir geliri, hanehalkı büyüklüğü (hanede yaşayan bireylerin sayısı), hanehalkı tipi, ikamet edilen konutun mülkiyet tipi ve hanehalkının coğrafi yerleşim birimi (kır/kent) arasındaki ilişkinin ortaya konulması amaçlanmıştır. Bu amaç kapsamında 2008 ve 2010 yıllarına ait TÜİK Hanehalkı Bütçe Anketi verileri üzerinde yerel doğrusal regresyon, en küçük kareler çapraz geçerlilik ve sabit pencere genişliği kullanılarak çeşitli tahminler yapılmıştır.

Yapılan analizler sonucunda ulaşılan temel sonuç gelir yükseldikçe zaruri harcamaların gelir esnekliğinin düştüğü, harcamaların gelirden bağımsız bileşenini baskılayan bir faktörün var olması durumunda ise (orta büyüklükte aile sahibi olmak, kentte yaşamak, kirada ikamet etmek) gelir esnekliğinin daha yüksek olduğudur. Öyle ki orta büyüklükteki hanelerde küçük hanelere kıyasla gelirden bağımsız olarak harcamalar düşük, harcamaların gelir esnekliği yüksek; aynı şekilde, kentsel hanehalkının kırsal hanehalkına göre; kirada oturan hanehalkının da kendi mülkiyetindeki bir evde yaşayan hanehalkına göre otonom zaruri harcamaları düşük, harcamalarının gelir esnekliği ise yüksek bulunmuştur.

**Anahtar Kelimeler:** Parametrik Olmayan Regresyon Analizi, Çekirdek Kestirim Tekniği, Harcama-Gelir İlişkisi

**ABSTRACT**

**MASTER'S THESIS**

**KERNEL REGRESSION WHICH IS NONPARAMETRIC REGRESSION  
MODEL AND AN APPLICATION IN REGARD TO HOUSEHOLD  
EXPENDITURE IN TURKEY**

**Elif GÖLVEREN**

**Advisor: Assoc. Prof. Dr. Mehmet Suphi ÖZÇOMAK**

**2012, Pages: 57**

**Jury: Assoc. Prof. Dr. Mehmet Suphi ÖZÇOMAK (Advisor)**

**Prof. Dr. Erkan OKTAY**

**Assist. Prof. Dr. Emrah TALAŞ**

This study investigates the relationship between the household expenditures and necessities (food and non-alcoholic beverages, health and education) and households' income, type, size, home ownership status and location. The relationships are estimated based on the Household Budget Surveys of 2008 and 2010 using the nonparametric kernel regression technique, the least squares cross-validation and the fixed bandwidth.

Our analyses suggest a decreasing pattern of income elasticity as income rise. Existence of a factor unrelated to income, such as having a medium size family, living urban, living as a tenant augments the income elasticity estimates. In medium size households, urban households and tenant households the expenditures on necessities are low yet income elasticity figures are high.

**Key Words:** Nonparametric Regression Analysis, Kernel Regression Model, Relation of Expenditure-Income



**KISALTMALAR VE SİMGELER DİZİNİ**

<b>ABK</b>	: Akaike Bilgi Kriteri
<b>ABK<sub>C</sub></b>	: Düzeltilmiş Akaike Bilgi Kriteri
<b>ATHKO</b>	: Asimptotik Toplanmış Hata Kareler Ortalaması
<b>CRAN</b>	: Comprehensive R Archive Network
<b>EKKCG</b>	: En Küçük Kareler Çapraz Geçerlilik
<b>GCG</b>	: Genelleştirilmiş Çapraz Geçerlilik
<b>GL</b>	: Hanehalkı Kullanılabilir Geliri
<b>HH</b>	: Hanehalkı Tipi
<b>HHB</b>	: Hanehalkı Büyüklüğü
<b>HKO</b>	: Hata Kareler Ortalaması
<b>KK</b>	: Kır-Kent Yerleşim Göstergesi
<b>k-NN</b>	: En Yakın Komşuluk Kestiricisi
<b>MLK</b>	: Hane Mülkiyet Tipi
<b>OECD</b>	: OECD Tanımlı Hanehalkı Büyüklüğü
<b>OHKO</b>	: Ortalanmış Hata Kareler Ortalaması
<b>OKH</b>	: Ortalanmış Kareli Hata
<b>THKO</b>	: Toplanmış Hata Kareler Ortalaması
<b>TKH</b>	: Toplanmış Kareli Hata
<b>TÜİK</b>	: Türkiye İstatistik Kurumu
<b>ZH</b>	: Zaruri Harcamalar
$\varepsilon_i$	: Gözlem hataları
$h$	: Pencere genişliği (düzleştirme parametresi)
$K(\cdot)$	: Çekirdek fonksiyonu
$m(\cdot)$	: Regresyon (ortalama) fonksiyonu
$w(\cdot)$	: Ağırlık fonksiyonu

**ŞEKİLLER DİZİNİ**

<b>Şekil 3.1.</b> Temel (Demografik Faktörleri İçermeyen) Modellerde Gelir Esneklikleri...	32
<b>Şekil 3.2.</b> Hanehalkı Tipinin Zaruri Harcamalara Etkisi .....	34
<b>Şekil 3.3.</b> Hanehalkı Tipi Kategorilerine Göre Gelir Esneklikleri.....	35
<b>Şekil 3.4.</b> Hanehalkı Büyüklüğünün (TÜİK Tanımı) Zaruri Harcamalara Etkisi.....	36
<b>Şekil 3.5.</b> Hanehalkı Büyüklüğüne (TÜİK Tanımı) Göre Gelir Esneklikleri .....	36
<b>Şekil 3.6.</b> Hanehalkı Büyüklüğünün (OECD Tanımı) Zaruri Harcamalara Etkisi .....	37
<b>Şekil 3.7.</b> Kır-Kent Ayrımının Zaruri Harcamalara Etkisi.....	37
<b>Şekil 3.8.</b> Kır-Kent Ayrımına Göre Gelir Esneklikleri .....	38
<b>Şekil 3.9.</b> Konut Mülkiyet Tipinin Zaruri Harcamalara Etkisi .....	38
<b>Şekil 3.10.</b> Konut Mülkiyet Tipine Göre Gelir Esneklikleri.....	39

**TABLolar DİZİNİ**

<b>Tablo 2.1.</b> Çekirdek Fonksiyonları .....	13
<b>Tablo 3.1.</b> Literatürde Ele Alınan Gıda Maddeleri ile Çeşitli Mal ve Hizmetler .....	28
<b>Tablo 3.2.</b> Seçilmiş Çalışmalarda Gelir Esneklikleri .....	29
<b>Tablo 3.3.</b> Literatürde Talebe Etkisi İncelenen Demografik Faktörler .....	29
<b>Tablo 3.4.</b> Betimleyici İstatistikler .....	31
<b>Tablo 3.5.</b> Tahmin Edilen Parametrik Olmayan Modellerin Özellikleri .....	31
<b>Tablo 3.6.</b> Parametrik Benchmark Modelleri (2008) .....	40
<b>Tablo 3.7.</b> Parametrik Benchmark Modelleri (2010) .....	41

**ÖNSÖZ**

Bu tez çalışmasının başlatılmasında, bir düzen içerisinde yürütülmesinde ve tamamlanmasında görüş ve desteklerinden yararlandığım danışmanım Sayın Doç. Dr. Mehmet Suphi Özçomak'a ve yine çalışmanın baştan sona tüm aşamalarının şekillenmesinde önemli katkıları olan hocam Sayın Doç. Dr. Eray Yücel'e sonsuz teşekkürlerimi sunarım. Ayrıca, ihtiyacım olduğunda desteklerini esirgemeyen diğer hocalarıma ve arkadaşlarıma ve çalışmada kullanılan veriye erişim konusunda göstermiş oldukları kolaylıktan ötürü Türkiye İstatistik Kurumu'na da teşekkürlerimi sunarım.

Çalışmaya doğrudan katkıları olmasa da, bu yolda beni yürekten destekleyen Sayın Prof. Dr. Adnan Ataç'a ve isminin anılması sona kalsa da şüphesiz en büyük teşekkürü hak eden sevgili annem Sayın Hülya Gölveren'e de minnettarlığımı dile getirmek isterim.

**Erzurum-2012**

**Elif GÖLVEREN**

## GİRİŞ

Parametrik regresyon analizinin bir alternatifi olarak gündeme gelen parametrik olmayan regresyon analizine özellikle deęişkenler arasındaki ilişkinin sahip olduęu fonksiyonel formun bilinmemesi durumunda başvurulmaktadır. Bu alıřmada da harcama-gelir ilişkisinin bilinen fonksiyonel biçimlerden bağımsız bir ekonometrik sunumunu yapmak için bu analize başvurulmuřtur. Söz konusu ilişkinin belirlenmesine yönelik yapılan uygulamada ise hanehalkı harcamalarının gelir esnekliğinin gelire göre deęişip deęişmedięi, deęiřtięi takdirde nasıl deęiřtięi incelenmiřtir.

Geleneksel parametrik ekonometrik teknikler nokta tahminlere odaklanmakta ve gözlem sayısı yüksek veri kümelerinde dahi sınırlı sayıda tahmin edilmiř parametre (katsayı) ile yetinmektedir. Bu ise, örneęin, harcamaların gelir esnekliğinin gelire göre deęişiminin incelenmesi gerektięinde, alt örneklemlerde tahminleri veya kukla deęişkenler üzerinden kategorizasyona gidilmesini zorunlu kılmaktadır. Fonksiyon biçimlerini dıřlayan bir tahmin yöntemi söz konusu zorunluluęu giderebilir. Nitekim alıřmada kullanılan parametrik olmayan ekirdek kestirim teknięi harcamaları gelire baęlayan tek bir tepki katsayısından söz etmek yerine, her gözlem noktasında o gözlem noktasına has bir katsayı tahmini üretmekte; böylelikle hem sürekli deęişkenler hem kategorik deęişkenler için daha zengin bulgulara ulařılmasını saęlamaktadır. alıřmada kullanılan yöntem sunduęu bu olanak nedeni ile tercih edilmiřtir.

Hanehalkı harcamalarının anlařılması ise hem ekonomik dinamikler hem politika üretimi açısından önem taşımaktadır. Toplam talebin nasıl biçimlendięinin anlařılmasında da harcamalar ile gelir arasındaki ilişkinin ölçülmesi başlangı noktasıdır. Bunların yanı sıra tüketimin ve harcamaların önemli bir belirleyicisi olduęu bilinen sosyo-demografik faktörlerin harcamalar üzerindeki etkisinin anlařılmasının da önemli olması, ele alınan uygulama alanının incelenmeye deęer olduęunu da göstermektedir.

alıřmada harcama-gelir ilişkisinin parametrik olmayan sunumuna yönelik tahminler R programında np paketi kullanılarak üretilmiřtir. Bu program, yaygın olarak kullanılan dięer programlara göre üstün sayılabilecek birok özellięe sahip olması nedeni ile tercih edilmiřtir. Programın önemli bir özellięi, hem bazı önde gelen

istatistikçiler hem de kullanıcılar tarafından sürekli geliştirilen bir konumda olmasıdır (R Development Core Team, 2012).

Yukarıda bahsedilen motivasyon kaynakları ile beslenen bu çalışmada, parametrik olmayan regresyon kestiricilerini çeşitli açılardan incelemek ve bu kestiricileri kullanarak Türkiye’de hanehalkı harcamaları ile hanehalkı kullanılabilir geliri ve sosyo-demografik özellikleri arasındaki ilişkiyi ortaya koymak amaçlanmıştır. Bu amaca dönük olarak hazırlanan içerik üç ana bölümden oluşmaktadır.

İlk bölümde regresyon analizinden ve bu analiz için bilinen iki yaklaşım olan parametrik ve parametrik olmayan regresyon analizlerinden bahsedilmiştir.

İkinci bölümde parametrik olmayan regresyon kestiricileri ele alınmış ve bu kestiricilerden çekirdek kestiricileri, en yakın komşuluk kestiricisi, spline modelleri, ortogonal seriler ve medyan düzleştirme tekniğine yer verilmiştir. Ayrıca bu bölümde ağırlıklı olarak çekirdek kestiricileri ile bu kestiriciler için belirleyici iki etmen olan çekirdek fonksiyonu ve pencere genişliği üzerinde durulmuş ve çekirdek kestiricileri için önerilen yerel sabit, yerel doğrusal, yerel polinomial, Gasser-Müller ve Priestley-Chao kestiricileri incelenmiştir.

Üçüncü bölümde ise Türkiye’de hanehalkı zaruri harcamaları ile hanehalkı kullanılabilir geliri ve sosyo-demografik özellikleri arasındaki ilişki incelenmiş ve yapılan analizler sonucu elde edilen bulgular sunulmuştur.

## BİRİNCİ BÖLÜM

### REGRESYON ANALİZİ VE REGRESYON ANALİZİNE YAKLAŞIMLAR

#### 1.1. REGRESYON ANALİZİ

İstatistiksel teknikler, bir deneysel verideki mümkün kalıpların tanımlanmasına olanak sağlayan araçlardır. Deneysel sorulara cevap arayan bilim adamları tarafından geliştirilen önemli istatistiksel analiz tekniklerinden biri *regresyon analizidir*. Regresyon analizi aracılığıyla bir veya daha fazla bağımsız değişkenin değerleri kullanılarak bağımlı değişkenin değeri tahmin edilmeye çalışılır (Allen, 1997).

Değişkenler arasındaki ilişkiyi incelemek bilimin uğraşlarından biri olagelmıştır. Gerek günlük hayatta gerek bilimsel araştırmalarda karşılaşılan sorunların çoğunluğu iki veya daha fazla değişken arasında ilişki olup olmadığının saptanması ile ilgilidir. Değişkenler arasındaki ilişkinin saptanması genellikle iki tür sorun için gerekli olur. Bunlardan birincisi, bir değişkene ilişkin gözlem değerleri yardımıyla diğer değişkenin alabileceği değerlerin ne kadar doğrulukla tahmin edilebileceği ile ilgilidir. Diğer bir sorun ise, değişken değerlerinde gözlenen farklılıkların ne dereceye kadar belirleyici bazı etmenlere bağlanabileceğiyle ilgilidir. İşte regresyon analizi, bu sorunları açıklayabilmek adına değişkenler arasındaki ilişkiyi inceler (Serper, 2004).

Regresyon analizinin tarihi 19. yy sonları İngiltere'sine, bilim adamı Francis Galton'un araştırmalarına uzanır. Boy karakteristiğinin kalıtımını incelemek için bireylerin ve ebeveynlerinin boylarına ilişkin topladığı veriyi inceleyen Galton, uzun boylu ebeveynlere sahip bireylerin genellikle ortalamadan daha uzun ancak ebeveynleri kadar uzun olmadıklarını, kısa boylu ebeveynlere sahip bireylerin de ortalamadan daha kısa fakat genellikle ebeveynleri kadar kısa olmadıklarını ortaya koymuş, kalıtımın bu modelini *ortalamaya doğru regresyon (eğilim)* olarak adlandırmıştır. Bu, boy gibi herhangi bir karakteristiği ortalamadan uzak olan ebeveynlerin, karakteristikleri genellikle ortalamaya daha yakın çocuklara sahip oldukları anlamına gelmektedir.

Galton bu prensibi bulmakla, istatistik disiplininin gelişmesinde çok önemli olacak iki kavramı da tespit etmiştir. Ortalamaya doğru eğilim olsa bile bireylerin boyları genelde ebeveynlerinin boylarıyla ilişkilidir. Galton, iki nicel değişken arasında

böyle bir ilişki tanımlamak için matematiksel bir fonksiyon geliştirmenin mümkün olabileceğini bulmuş, ayrıca bazı ilişkilerin diğerlerinden daha tahmin edilebilir olduğunu fark etmiştir. Örneğin, bireylerin boylarının tamamen ebeveynlerinin boylarıyla tahmin edilemez olduğunu ileri sürmüştür. Galton, bir ilişkideki tahmin edilebilirliğin ölçüsünü *korelasyonun yakınlığı* olarak adlandırmıştır.

Galton tarafından geliştirilen regresyon ve korelasyon kavramlarına ait matematiksel formülleri geliştirme görevi ise Francis Edgeworth, Karl Pearson ve George Yule tarafından sonraki birkaç yılda tamamlanmıştır (Allen, 1997).

### 1.1.1. Parametrik Regresyon Analizi

Bir bağımlı değişkenin ( $Y$ ),  $k$  adet bağımsız değişkenin ( $X_1, X_2, \dots, X_k$ ) bir fonksiyonu olduğunu varsayan standart çok değişkenli regresyon modeli şu şekilde modellenebilir:

$$Y_j = \alpha + \left( \sum_{i=1}^k \beta_i X_{ij} \right) + \varepsilon_i$$

$Y_j$  ve  $X_{ij}$   $j$ . gözlem için  $Y$  ve  $X_i$ 'nin değerleridir.  $k$  adet bağımsız değişkenin her bir değer dizisi için ( $X_{1j}, X_{2j}, \dots, X_{kj}$ )  $Y$  değerlerinin koşullu olasılık dağılımı olduğu varsayılır. Bu dağılımlar sabit varyansa sahip normal dağılımdır (Berry, 1993).

Bir ilişkinin şekli, doğruluğu ve istatistiksel olarak anlamlı olup olmadığının sorgulanmasına ve kategorik değişkenlerin kullanılmasına olanak tanıyan regresyon analizi (Allen, 1997); modele dahil edilen değişkenlerin uygun olduğu, uygun fonksiyonel form kullanıldığı, değişkenlerin doğru bir şekilde ölçüldüğü ve verinin rastsal örneklemeden oluştuğu gibi birkaç temel varsayıma dayanır (Schroeder, Sjoquist ve Stephan, 1986). *Klasik doğrusal regresyon modeli* olarak adlandırılan temel regresyon modelinde bu varsayımlara ilaveten aşağıdaki hususların doğruluğu da kabul edilir:

1. Bağımlı değişken, bir dizi bağımsız değişkenin doğrusal bir fonksiyonudur.
2. Hata teriminin beklenen değeri sıfırdır.
3. Hata terimleri aynı varyansa sahip ve ilişkisizdir.
4. Tekrarlanan örneklemlerde bağımsız değişkenin değerleri sabittir.



5. Bağımsız değişkenler kendi aralarında doğrusal bir ilişki içinde değildir.

Bu varsayımların herhangi birinin ihlali söz konusuysen kullanılabilecek çeşitli modeller geliştirilmiştir (Kennedy, 2006).

Doğrusallık varsayımı ve bu varsayımın ihlali durumunda kullanılacak yaklaşımlarda, regresyon fonksiyonunun önceden bilinen bir fonksiyonel forma sahip olduğu varsayılır. Regresyon fonksiyonunu belirli bir forma dayandırmaksızın kestirmek için kullanılan teknik ise *parametrik olmayan regresyon analizi* olarak bilinir (Hardle, 1990). Parametrik regresyon analizinin aksine parametrik olmayan regresyon analizinde, tahmin edilen regresyon yüzeyinin düzgünlüğü dışında herhangi bir varsayım aranmaz (Kıroğlu, 1999).

### 1.1.2. Parametrik Olmayan Regresyon Analizi

Regresyon analizine parametrik yaklaşım gözlenen değerler kümesinin, parametrik bir ailenin üyesi olan bir dağılımdan; örneğin, ortalaması  $\mu$  ve varyansı  $\sigma^2$  olan bir normal dağılımdan alındığını varsayar. Parametrik olmayan yaklaşım ise, çoğu kez gözlemlerin dağılımından bağımsız bir süreç olarak tanımlanmıştır (Toktamış, 1995).

$\{(X_i, Y_i)\}_{i=1}^n$  şeklinde  $n$  tane veri noktası varsa, parametrik olmayan regresyon ilişkisi şu şekilde modellenebilir: ( $i = 1, \dots, n$ )

$$Y_i = m(X_i) + \varepsilon_i \quad (1.1)$$

$m$  bilinmeyen regresyon (ortalama) fonksiyonu,  $\varepsilon_i$  gözlem hatalarıdır. Burada,  $m$  fonksiyonunun varlığı ile hata terimi  $\varepsilon_i$ 'lerin bağımsız ve sıfır ortalamalı olduğu varsayılır ve amaç,  $m$  fonksiyonunun uygun bir kestirimini elde etmektir. Bu eğri kestirim prosedürü genellikle *düzgünleştirme (smoothing)* olarak adlandırılır (Hardle, 1990).

*Düzgünleştirici (smoother)*, bir veya daha fazla bağımsız değişkenin fonksiyonu olan bağımlı değişkenin sahip olduğu trendi ifade etmek için kullanılan bir araçtır. Bir düzgünleştiricinin en önemli özelliği değişkenler arasındaki ilişkinin biçimini kesin bir biçimde varsaymamasıdır ve bu özelliğinden dolayı parametrik olmayan regresyonda

sık kullanılır (Tezcan, 2009). Düzgünleştiriciler için iki genel görüş hâkimdir: Bunlardan, verideki genel yapıya dayalı olan ilk görüş *serpilme diyagramı düzgünleştiricileri* (*scatter plot smoothing*) olarak adlandırılır. Yoğun matematiksel analizler içeren metodolojik çalışmaların çoğunda tercih edilen ikinci görüş ise, rastsal verilere dayalı kestirimler yapmaktır. Bu görüş, parametrik olmayan regresyon kestiricilerini işaret eder (Chu ve Marron, 1991).

*Koşullu olasılık ve koşullu olasılık fonksiyonu* kavramları, hem parametrik hem de parametrik olmayan regresyon analizindeki en önemli ve en temel kavramlardır.  $X$  ve  $Y$ ,  $f(x, y)$  ortak olasılık yoğunluk fonksiyonuna sahip iki rastsal değişken olsun.  $X = x$  verilmişken  $Y$ 'nin koşullu olasılığı,

$$E(Y|X = x) = \int yf(y|x)dy = \int y \frac{f(x, y)}{f_X(x)} dy = m(x)$$

şeklindedir. Burada  $f(y|x)$ ,  $X = x$  verilmişken  $Y$ 'nin koşullu olasılık yoğunluk fonksiyonu;  $f_X(x)$  ise  $X$ 'in marjinal olasılık yoğunluk fonksiyonudur.  $E(Y|X = x)$ , yalnızca  $x$ 'in bir fonksiyonudur. Bu yüzden bu terim  $m(x)$  olarak ifade edilebilir.  $x$ 'e farklı değerler verilmesi ile koşullu olasılıklar, her bir  $x$  değerine karşılık gelen  $m(x)$  değerinin bulunması ile de  $E(Y|X)$  olarak gösterilen koşullu olasılık fonksiyonu elde edilir. Bu fonksiyon,  $X$  ve  $Y$ 'nin ortalama olarak ne kadar ilişkili olduğunu gösterir. Dolayısıyla buradaki asıl ilgi,  $m(\cdot)$  fonksiyonunun kestirilmesidir (Hardle, Müller, Sperlich ve Werwatz, 2004).

## İKİNCİ BÖLÜM

### PARAMETRİK OLMAYAN REGRESYON KESTİRİCİLERİ

Parametrik olmayan regresyon kestiricileri, parametrik modellerdeki varsayım ihlallerinin yanı sıra dağılım hakkında hiçbir bilgi olmadığı durumlarda kullanılır. Ayrıca parametrik model kurmaya yardımcı oluşu ve başta çekirdek regresyon olmak üzere diğer tüm kestiricilerinin tutarlı kestirimler sunmasından dolayı parametrik regresyon analizine alternatif olarak tercih edilmektedir (Kıroğlu, 1999). Parametrik olmayan regresyon kestiricilerinin temel amaçları; iki değişken arasındaki ilişkiyi açıklamada çok yönlü bir metot sağlamak, kayıp gözlemleri temsil etmede veya yakın  $X$  değerleri arasında bir değer bulmada esnek bir metot kurmak ve aykırı noktaların etkisini inceleyerek sahte gözlemleri ortaya çıkarmaktır (Hardle, 1990).

En basit ve en yaygın kullanılan parametrik olmayan regresyon kestiricileri *çekirdek metoduna* dayanır (Chu ve Marron, 1991). Çekirdek odaklı kestiriciler basit ve sezgisel olarak iyi anlaşıldığı için tercih edilmektedir. Diğer önemli parametrik olmayan eğri kestiricileri ise *k-en yakın komşuluk*, *spline modelleri* ve *ortogonal serilerdir* (Hardle, 1990).

Parametrik olmayan regresyon için önerilen bu modeller, *sabit* ve *rastgele tasarlanmış modeller* olmak üzere iki grupta toplanabilir. Sabit tasarım durumunda, tasarım değişkeninin belirli aralıkta olduğu varsayılır. Bu aralık genellikle  $[0,1]$  aralığıdır. (1.1) eşitliğindeki model sabit tasarım için tanımlandığında, eşit aralıkta sabit düzenlenmiş noktalar olan ve  $i/n$  şeklinde gösterilebilen  $X_i$ 'ler  $0 \leq X_1 \leq \dots \leq X_n \leq 1$  aralığında ifade edilir (Gökçe, 2008; Steinmetz ve Gasser, 1988). Rastgele tasarım durumunda ise,  $X$  ve  $Y$  değişkenlerinin rastgele olduğu varsayılır. Bu tasarımda oluşturulacak örneklem,  $f(x, y)$  şeklindeki iki değişkenli dağılımdan elde edilen rastsal örneklemdir. Rastgele seçilen  $n$  adet  $\{X_i, Y_i\}$  gözlem çiftleri ise, bağımsız ve özdeş dağılmış değişken çiftleridir. Çekirdek kestirim tekniklerinden Priestley-Chao ve Gasser-Müller çekirdek kestiricileri sabit tasarım durumu için, Nadaraya-Watson çekirdek kestiricisi ise rastsal tasarım durumu için ileri sürülmüştür (Steinmetz ve Gasser, 1988; Hardle vd., 2004).

## 2.1. ÇEKİRDEK KESTİRİCİLERİ

Çekirdek regresyonda bağımlı değişkenin yerel ağırlıklı ortalaması alınır. Ağırlıklar çekirdek olarak adlandırılan bir fonksiyona dayanır. Çekirdek adı da buradan gelmektedir (Chu ve Marron, 1991).

Çekirdek regresyona ilişkin geliştirilen yaklaşımlar arasında *yerel sabit* (Nadaraya-Watson), *yerel doğrusal* ve *yerel polinomial* çekirdek kestiricileri ile diğer alternatif kestiriciler (Priestley-Chao, Gasser-Müller) yer almaktadır.

### 2.1.1. Yerel Sabit Çekirdek Kestiricisi

Yerel sabit çekirdek kestirim tekniği Nadaraya (1965) ve Watson (1964) tarafından önerilmiş olup, *Nadaraya-Watson çekirdek kestiricisi* olarak da adlandırılır. Buradaki mantık, bağımlı değişkenin değerlerini yerel olarak ortalayarak koşullu ortalama fonksiyonunu kestirmektir.

$(X, Y)$  gözlem çiftleri için  $f(x, y)$  yoğunluk fonksiyonu kullanılarak elde edilen (Hardle ve Linton, 1994) yerel sabit çekirdek kestiricisi:

$$\hat{m}(x) = \frac{\sum_{i=1}^n Y_i K\left(\frac{x - X_i}{h}\right)}{\sum_{i=1}^n K\left(\frac{x - X_i}{h}\right)} \quad (2.1)$$

şeklindedir ve,

$$\min_a \sum_{i=1}^n (Y_i - a)^2 K\left(\frac{x - X_i}{h}\right)$$

minimizasyon problemine odaklıdır (Li ve Racine, 2007). Burada  $K$  çekirdek ya da *ağırlık fonksiyonu* olup, sıfır etrafında simetriktir ve  $\lim_{x \rightarrow \infty} |x|K(x) = 0$  özelliğini sağlamaktadır.  $h$ , *pencere genişliği* ya da *düzleştirme parametresi* adı verilen, kestirimin düzgünlüğünü kontrol eden ve genellikle pozitif değer alan parametredir (Schimek, 2000). Bir  $X_i$  gözlem noktasında  $h \rightarrow 0$  olduğunda  $\hat{m}_h(X_i) \rightarrow Y_i$ , keyfi bir  $x$  noktasında ise  $h \rightarrow \infty$  olduğunda  $\hat{m}_h(X_i) \rightarrow \bar{Y}$  olacaktır. Bu iki limit durumdan anlaşılacağı gibi, örnek büyüklüğü  $n$  ile ilişkili olan düzleştirme parametresi  $h$ , sıfıra ne çok hızlı ne de çok yavaş yakınsamalıdır (Hardle ve Linton, 1994).

$n \rightarrow \infty$  iken,  $h \rightarrow 0$  ve  $nh \rightarrow \infty$  olması şartıyla, (2.1)'deki formülde pay ve payda,

$$E \left[ \sum_{i=1}^n Y_i K \left( \frac{x - X_i}{h} \right) \right] \rightarrow m(x) f(x),$$

$$E \left[ \sum_{i=1}^n K \left( \frac{x - X_i}{h} \right) \right] \rightarrow f(x)$$

olacağından  $\hat{m}(x) \rightarrow m(x)$  olacaktır (Hardle ve Linton, 1994; He, 2009). Dolayısıyla bu kestirici ile elde edilen regresyon fonksiyonunun,  $m(x)$ 'in tutarlı bir kestiricisi olduğu söylenebilir (Demir, 2005).

Schuster (1972) tarafından gösterildiği gibi,

$$\sqrt{nh}[\hat{m}_h(x) - m(x) - h^2 \text{Yan}(x)] \Rightarrow N(0, \text{Var}(x))$$

$$\text{Yan}(x) = \frac{1}{2} \mu_2 K \left[ m''(x) + 2m'(x) \frac{f'(x)}{f(x)} \right]$$

$$\text{Var}(x) = v_2(K) \sigma^2(x) / f(x)$$

yerel sabit kestirici asimptotik olarak normaldir (Hardle ve Linton, 1994).

### 2.1.2. Yerel Doğrusal Çekirdek Kestiricisi

Veri kümesinin destek sınırına yakın regresyon fonksiyonlarında yerel sabit kestiricilerin yanlışlık göstermesi durumunda kullanılan en iyi yaklaşımlardan biri yerel doğrusal kestiricidir (Fan, 1993). Bu kestirici,

$$\hat{m}(x) = \frac{\sum_{i=1}^n Y_i K \left( \frac{x - X_i}{h} \right) (s_{n,2} - (x - X_i) s_{n,1})}{n^{-2} + \sum_{i=1}^n K \left( \frac{x - X_i}{h} \right) (s_{n,2} - (x - X_i) s_{n,1})}$$

$$s_{n,l} = \sum_{i=1}^n K \left( \frac{x - X_i}{h} \right) (x - X_i)^l, \quad l = 0, 1, 2.$$

şeklindedir ve,

$$\sum_{i=1}^n (Y_i - a - b(x - X_i))^2 K \left( \frac{x - X_i}{h} \right)$$

ifadesinin minimizasyonuna dayanır (Fan, 1993).

### 2.1.3. Yerel Polinomial Çekirdek Kestiricisi

Yerel doğrusal çekirdek kestiricisini genişletip veriye yüksek dereceden polinomların uydurmasıyla yerel polinomial çekirdek kestirici elde edilir (Wand ve Jones, 1995).  $p$ , polinomun derecesi olmak üzere bir  $x$  noktasındaki  $\hat{m}(x; p, h)$  kestiricisi,

$$\beta_0 + \beta_1(x_i - x) + \dots + \beta_p(x_i - x)^p$$

polinomunun  $K_h(x - x_i)$  ağırlıklı en küçük kareler tekniği kullanılarak  $(x_i, Y_i)$  gözlem çiftlerine uydurulmasıyla elde edilir (Wand ve Jones, 1995). Yerel polinomial çekirdek kestiricisi,

$$\min_{\beta} \sum_{i=1}^n \{Y_i - \beta_0 - \beta_1(x_i - x) - \dots - \beta_p(x_i - x)^p\}^2 K_h(x - x_i)$$

minimizasyon problemine odaklıdır. Burada  $\beta$ ,  $(\beta_0, \beta_1, \dots, \beta_p)^T$  katsayı vektörüdür.

$$X = \begin{bmatrix} 1 & x_1 - x & \dots & (x_1 - x)^p \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n - x & \dots & (x_n - x)^p \end{bmatrix}, \quad Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix},$$

$$W = \begin{bmatrix} K_h(x - x_1) & 0 & \dots & 0 \\ 0 & K_h(x - x_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & K_h(x - x_3) \end{bmatrix}$$

şeklindeki matrisler kullanılarak  $\beta$  katsayılarının kestirimi,

$$\hat{\beta}(x) = (X^T W X)^{-1} X^T W Y$$

olarak elde edilir. Parametrik en küçük karelerin aksine, bu kestirici  $x$  ile birlikte değişim gösterir. Bu da,  $x$  noktasında yerel bir regresyon olduğu anlamına gelir.  $m$  regresyon fonksiyonunun yerel polinomial kestiricisi,

$$\hat{m}_{p,h}(x) = \hat{\beta}_0(x)$$

şeklinindedir. Bu formülde,  $p$ 'nin bazı değerleri için diğer çekirdek kestiricileri elde edilir. Örneğin,  $p = 0$  için,

$$\hat{m}_{0,h}(x) = \hat{m}_h(x) = \frac{\sum_{i=1}^n K_h(x - X_i) Y_i}{\sum_{i=1}^n K_h(x - X_i)}$$

yerel sabit kestirici elde edilir.  $p = 1$  alındığında ise,

$$S_{h,j}(x) = \sum_{i=1}^n K_h(x - X_i) (X_i - x)^j,$$

$$T_{h,j}(x) = \sum_{i=1}^n K_h(x - X_i) (X_i - x)^j Y_i$$

ifadeleri kullanılarak  $\hat{\beta}(x)$ ,

$$\hat{\beta}(x) = \begin{pmatrix} S_{h,0}(x) & S_{h,1}(x) \\ S_{h,1}(x) & S_{h,2}(x) \end{pmatrix}^{-1} \begin{pmatrix} T_{h,0}(x) \\ T_{h,1}(x) \end{pmatrix}$$

olarak yazıldığında,

$$\hat{m}_{1,h}(x) = \hat{\beta}_0(x) = \frac{T_{h,0}(x)S_{h,2}(x) - T_{h,1}(x)S_{h,1}(x)}{S_{h,0}(x)S_{h,2}(x) - S_{h,1}^2(x)}$$

yerel doğrusal kestirici elde edilir (Hardle vd., 2004).

Yerel polinomial regresyon, locally weighted scatter plot smoothing ifadesindeki baş harflerin birleştirilmesiyle oluşan LOWESS adı ile bilinmektedir. LOWESS adı tek bağımsız değişkene sahip basit regresyon uygulamalarında kullanılmaktadır. Bağımsız değişken sayısı arttığında ise aynı yöntem LOESS olarak adlandırılmaktadır (Tezcan, 2009). LOWESS birinci dereceden bir fonksiyon olan doğrusal bir polinom fonksiyonu kullanırken, LOESS ikinci dereceden bir fonksiyon olan karesel polinom fonksiyonu kullanmaktadır (Çatalbaş, 2006).

#### 2.1.4. Gasser-Müller Çekirdek Kestiricisi

$[a, b]$  tanım aralığından elde edilen  $x_{(i)}$  ( $i = 1, \dots, n$ ) sıralı tasarım noktaları durumu için, Gasser-Müller tarafından regresyon fonksiyonunun çekirdek kestirimine yönelik önerilen ağırlık fonksiyonu,

$$W_{hi}^{GM}(x) = n \int_{s_{i-1}}^{s_i} K_h(x - u) du$$

şeklindedir. Burada  $s_i = (x_{(i)} + x_{(i+1)})/2$  olarak ifade edilir. Bu ağırlık fonksiyonuna sahip Gasser-Müller çekirdek kestiricisi ise,

$$m(x) = \frac{1}{n} \sum_{i=1}^n W_{hi}^{GM}(x) Y_i$$

olarak tanımlanmıştır (Hardle vd., 2004).

#### 2.1.5. Priestley-Chao Çekirdek Kestiricisi

Gasser-Müller kestiricisi gibi sabit tasarım durumu için önerilmiş olan Priestley-Chao çekirdek kestiricisinin temel aldığı ağırlık fonksiyonu,

$$W_{hi}^{PC}(x) = n(X_i - X_{i-1})K_h(x - X_i)$$

şeklindedir ( $X_0 = 0$ ) (Hardle, 1990). Bu fonksiyona sahip Priestley-Chao çekirdek kestiricisi ise,

$$m(x) = \frac{1}{n} \sum_{i=1}^n W_{hi}^{PC}(x) Y_i$$

olarak ifade edilir (Eubank, 1999).

Priestley-Chao kestiricisinin, Gasser-Müller kestiricisinin bir Riemann toplam yaklaşımı olduğu belirtilmektedir. Bu nedenle, bu iki kestiricinin birbirine çok yakın olduğu söylenebilir (Demir, 2005).



## 2.2. ÇEKİRDEK KESTİRİCİLERİNİN ÖZELLİKLERİ

Her gözlem orta nokta olmak üzere her gözleme yerleştirilen tümseklerin toplamı olarak düşünülen çekirdek kestiricilerinde, hangi yaklaşım kullanılırsa kullanılsın, belirleyici iki etmen bulunmaktadır. Bunlardan biri tümseklerin biçimini ifade eden çekirdek fonksiyonu, diğeri ise tümseklerin yayılımını ifade eden pencere genişliğidir (Toktamış, 1995). Pencere genişliği seçimiyle karşılaştırıldığında çekirdek fonksiyonu seçimi daha az önem arz eden bir husustur. Çekirdek fonksiyonunun şeklindeki değişikliğin kestirime etkisi, pencere genişliğindeki değişikliğin neden olacağı etkiden çok daha azdır. Pencere genişliği seçimi ise tüm parametrik olmayan eğri kestiricilerinin uygulamasında ana sorundur (Marron, 1988).

### 2.2.1. Çekirdek Fonksiyonu

Olasılık fonksiyonu özelliklerini taşıyan herhangi bir fonksiyon olarak alınabilen çekirdek fonksiyonu, en iyi olduğu için değil, kolay uygulanırlığı ve diğer yöntemleri anlamada bir araç olması nedeni ile tercih edilmektedir (Kıroğlu, 1999). Yaygın olarak kullanılan çekirdek fonksiyonları Tablo 2.1’de sunulmaktadır (Hardle vd., 2004).

**Tablo 2.1.** Çekirdek Fonksiyonları

Çekirdek	K(u)
Tekbiçimli	$\frac{1}{2}I( u  \leq 1)$
Üçgensel	$(1 -  u )I( u  \leq 1)$
Epanechnikov	$\frac{3}{4}(1 - u^2)I( u  \leq 1)$
İki Ağırlıklı	$\frac{15}{16}(1 - u^2)^2I( u  \leq 1)$
Üç Ağırlıklı	$\frac{35}{32}(1 - u^2)^3I( u  \leq 1)$
Gaussian	$\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right)$
Kosinüs	$\frac{\pi}{4} \cos\left(\frac{\pi}{2}u\right)I( u  \leq 1)$

K ile gösterilen çekirdek fonksiyonu, her  $u$  için  $K(u) \geq 0$ ,  $K(u) = K(-u)$ ,  $\int_{-\infty}^{\infty} K(u)du = 1$ ,  $\int_{-\infty}^{\infty} K^2(u)du < \infty$  özelliklerine sahiptir (Benedetti, 1977).

### 2.2.2. Pencere Geniřliđi

Çekirdek kestirimlerinin elde edilmesinde önemli bir etken olan pencere genişliğinin çok küçük seçilmesi, yani çekirdeğin dar olması durumunda her noktadaki ortalama alma süreci daha az sayıda gözleme dayanacağından regresyon fonksiyonunun kestirimi kaba bir kestirim olacaktır (Toktamış, 1995). Başka bir deyişle, kestirim sonuçları örnek deđişkenliğine fazlaca bağımlı olacaktır (Marron, 1988). Bu durumda elde edilen kestirimlere *az düzleştirilmiş kestirimler* denir. Gereğinden büyük bir pencere genişliđi ise temel eğrinin özelliklerinin aşırı düzleştirilmesine (Marron, 1988), yani verilerin genel yapısının kaybolmasına neden olabilir. Bu durumda ise *fazla düzleştirilmiş kestirimler* elde edilir (Toktamış, 1995).

Bu kritik seçim için önerilen bir yöntem deneme yanılma yaklaşımıdır. Bunun için farklı düzgünlük miktarlarını temsil eden birkaç farklı grafiđe bakmak yeterlidir. Böyle bir strateji, geniş bir pencere genişliđi ile başlamaya ve dalgalanmaların yapısalardan daha rastgele olduđu görülene kadar genişliđi azaltmaya dayanır (Wand ve Jones, 1995). Bu yaklaşım şüphesiz veri seti hakkında fikir edinmeyi sağlar, fakat hipotez testini algılamada oluşan şüpheleri yok edemeyebilir (Marron, 1988). Pencere genişliğinin bu seçimi arařtırıcıların veride belirli yapılar olduđunu bildiđi durumlarda daha tutarlıdır. Ne yazık ki, çođu durumda böyle bir ön bilgiye sahip olunmaz. Ayrıca birden fazla parametrik olmayan regresyon yüzeyi olduđu durumlarda uygun genişliđin hangisi olduđuna karar vermek, özellikle gözlem sayısı fazlayken, hayli zaman alabilir. Tüm bu sebepler pencere genişliğini otomatik olarak seçmenin çok daha faydalı olduđunu göstermektedir. Çekirdek kestiricileri daha kapsamlı istatistiksel sürecin bir parçası olarak kullanıldıđında ise otomatik pencere genişliđi seçimi genellikle zorunludur.

Otomatik pencere genişliđi seçicileri kabaca iki sınıfa ayrılabilir. İlk sınıf, çok sayıda durum için uygun olan fakat optimal pencere genişliğine yakın olmanın hiçbir matematiksel garantisinin olmadığı bir pencere genişliđi bulmayı amaçlayan, basit ve kolayca hesaplanabilir formüllerden meydana gelir. Böyle pencere genişliđi seçicileri *hızlı ve basit* olarak adlandırılır. Hızlı ve basit pencere genişliđi seçicilerinin tercih edilme sebebi, pencere genişliğinin sübjektif seçimi için uygun bir başlangıç noktası sağlıyor olmalarıdır. *İleri teknoloji* olarak ifade edilen diđer pencere genişliđi seçicileri ise daha karmaşık matematiksel ispatlara dayanır ve oldukça fazla hesapsal çaba

gerektirir, fakat temel fonksiyon için iyi bir cevap vermeyi amaçlar (Wand ve Jones, 1995). Bu tür pencere genişliği seçicileri arasında *çapraz geçerlilik* ve *plug-in* yöntemi ile *cezalandırma fonksiyonları* yer alır.

Pencere genişliği seçimi için önerilen bu farklı yaklaşımlar kullanılırken, pencere genişliğini sabit veya değişken olarak ele almak gerekir. Sabit pencere genişliğini kullanan kestiriciler uzun kuyruklu dağılımlarda, çok tepeli dağılımlarda ve çok değişkenli dağılımlarda yetersiz kalmaktadır. Bu durumda değişen pencere genişlikli kestiriciler kullanılmaktadır (Var, 2005).

### 2.3. PENCERE GENİŞLİĞİ SEÇİCİLERİ

Bir pencere genişliği seçim yöntemi hem teorik olarak istenen özelliklere sahip hem de pratikte uygulanabilir olmalıdır. İlk şartla ilgili olarak kestirimin gerçek eğriye ne kadar uygun olduğunu ölçmek için önerilen bir dizi ölçüt bulunmaktadır. Bunlar:

#### Hata Kareler Ortalaması (HKO)

$$HKO(x, h) = HKO\{\hat{m}_h(x)\} = E[\{\hat{m}_h(x) - m(x)\}^2]$$

Burada  $m(x)$  bilinmeyen sabit değişken, kestirici  $\hat{m}_h(x)$  ise rastsal değişkendir. HKO, tek bir  $x$  noktasında  $\hat{m}_h$  kestiriminin  $m$ 'den kareli sapmasını ölçmektedir. Eğer  $m$ 'nin bir bütün olarak ne kadar iyi kestirildiğiyle ilgilenilecekse kestirimin gerçek eğriye yakınlığının genel bir ölçütü kullanılmalıdır.

#### Toplanmış Kareli Hata (TKH)

$$TKH(h) = TKH\{\hat{m}_h\} = \int_{-\infty}^{\infty} \{\hat{m}_h(x) - m(x)\}^2 w(x) fX(x) dx$$

Bu ölçüt, genel bir sapma ölçөгüdür. Fakat farklı örnekler  $\hat{m}_h(x)$ 'in dolayısıyla TKH'nin farklı değerler almasına neden olacağından  $\hat{m}_h(x)$ , hâlâ rastsal bir değişkendir. Ağırlık fonksiyonu  $w(\bullet)$ , verinin dağılık bölgelerindeki gözlemlere bu bölgelerdeki değişkenlik azalsın diye veya  $X$  dağılımının kuyruğundaki gözlemlere sınır etkisinden arınsın diye daha az ağırlık atamak için kullanılabilir.

TKH özünde HKO ile aynıdır. Tek farkı ağırlık fonksiyonu ve  $X$  değişkeninin yoğunluk fonksiyonu ile ağırlıklandırılarak hesaplanmasıdır. Dolayısıyla daha hassas bir tartım vermektedir.

### Toplanmış Hata Kareler Ortalaması (THKO)

$$THKO(h) = THKO\{\hat{m}_h\} = E\{TKH(h)\} = \int \dots \int \left[ \int_{-\infty}^{\infty} \{\hat{m}_h(x) - m(x)\}^2 w(x) f_X(x) dx \right] \cdot f(x_1, \dots, x_n, y_1, \dots, y_n) dx_1, \dots, dx_n, dy_1, \dots, dy_n$$

THKO, rastsal bir deęişken deęildir. Rastsal deęişken olan TKH'nin beklenen deęeridir.

### Ortalanmış Kareli Hata (OKH)

$$OKH(h) = OKH\{\hat{m}_h\} = \frac{1}{n} \sum_{j=1}^n \{\hat{m}_h(X_j) - m(X_j)\}^2 w(X_j)$$

TKH'ye farklı bir yaklaşımdır. TKH gibi hem rastsal bir deęişken hem de sapmanın genel bir ölçęüdür.

### Ortalanmış Hata Kareler Ortalaması (OHKO)

$$OHKO(h) = OHKO\{\hat{m}_h\} = E\{OKH(h)|X_1 = x_1, \dots, X_n = x_n\}$$

OKH'nin koşullu beklenen deęeridir.  $X_1, \dots, X_n$  rastsal deęişkenler ise OHKO da rastsal bir deęişkendir.

Hangi ölçęin kullanılacağı konusuna gelince, yoğunluk durumunda doğal bir tercih THKO veya asimptotik uyarlaması ATHKO ölçütünden yana olacaktır. Fakat regresyon durumundaki ATHKO yoğunluk durumundaki ATHKO ölçütünden daha fazla bilinmeyen nicelik gerektirir. Yerel sabit kestirici için düşünöldüğünde OKH, TKH ve THKO ölçütlerinin asimptotik olarak aynı düzleştirme seviyelerini verdikleri Marron ve Hardle (1986) tarafından ispatlanmıştır. Bu yüzden hesaplama ve kullanım açısından en kolay ölçüt olan OKH ölçütü kullanılabilir (Hardle vd., 2004).

#### 2.3.1. Çapraz Geçerlilik Yöntemi

Çapraz geçerlilik yöntemi, parametrik olmayan regresyon eğrisinin kestirimi için düşünöldüğünde, OHKO ölçütünü optimize etmeyi temel alır. Temel amaç,  $OKH(h)$  ölçütünü optimize eden pencere genişliğini bulmaktır. OKH ölçütü detaylı bir şekilde şöyle yazılabilir:

$$OKH(h) = \frac{1}{n} \sum_{i=1}^n m^2(X_i)w(X_i) + \frac{1}{n} \sum_{i=1}^n \hat{m}_h^2(X_i)w(X_i) - 2 \frac{1}{n} \sum_{i=1}^n m(X_i)\hat{m}_h(X_i)w(X_i)$$

OKH ölçütünün koşullu beklenen değeri OHKO ise şöyledir:

$$\begin{aligned} OHKO(h) &= E\{OKH(h)|X_1 = x_1, \dots, X_n = x_n\} \\ &= \frac{1}{n} \sum_{i=1}^n E[\{\hat{m}_h(X_i) - m(X_i)\}^2 | X_1 = x_1, \dots, X_n = x_n] w(X_i) \\ &= \frac{1}{n} \sum_{i=1}^n \left[ \frac{Var\{\hat{m}_h(X_i)\} | X_1 = x_1, \dots, X_n = x_n}{v(h)} \right. \\ &\quad \left. + \frac{Yan^2\{\hat{m}_h(X_i)\} | X_1 = x_1, \dots, X_n = x_n}{y^2(h)} \right] w(X_i) \end{aligned}$$

Burada yan kare:

$$y^2(h) = \frac{1}{n} \sum_{i=1}^n \left\{ \frac{1}{n} \sum_{j=1}^n \frac{K_h(X_i - X_j)}{\hat{f}_h(X_i)} m(X_j) - m(X_i) \right\}^2 w(X_i)$$

ve varyans:

$$v(h) = \frac{1}{n} \sum_{i=1}^n \left\{ \frac{1}{n^2} \sum_{j=1}^n \frac{K_h(X_i - X_j)}{\hat{f}_h(X_i)} \sigma(X_i)^2 \right\}^2 w(X_i) \quad (2.2)$$

şeklindedir. Pencere genişliği küçük olduğunda varyans büyük, yan kare küçük; pencere genişliği büyük olduğunda ise varyans küçük yan kare büyük olmaktadır. OKH ölçütünü minimuma indirmek için bu iki bileşen arasındaki en iyi denge elde edilmelidir. Bu denge sağlandığında, gerekenin altında düzleştirme yaparak veriden bilgi türetememe veya gereğinden fazla düzleştirme yaparak verinin içerdiği bilgiyi çarpıtma ihtimalleri bertaraf edilmektedir.  $OKH(h)$  dolayısıyla  $OHKO(h)$  ölçütü

kestirmek istenilen  $m(\bullet)$  fonksiyonuna bağlıdır. Bu nedenle veriden hesaplanabilen bir nicelikle  $OKH(h)$  ölçütünün yeniden yazılması gerekir. Bunun için,  $m(\bullet)$  yerine  $Y$  gözlemleri kullanılabilir:

$$p(h) = \frac{1}{n} \sum_{i=1}^n \{Y_i - \hat{m}_h(X_i)\}^2 w(X_i)$$

Bu kestirim, esasen ağırlıklı artık kareler toplamıdır.  $h \rightarrow 0$  iken  $p(h) \rightarrow 0$  olmaktadır. Fakat  $Y_i$  gözleminin kendi tahmini için  $\hat{m}_h(X_i)$ 'de kullanılması bir sorun teşkil eder. Bu sorunu göstermek için  $m(X_i)$  terimi eklenip çıkartılarak  $p(h)$  genişletilirse,

$$\begin{aligned} p(h) &= \frac{1}{n} \sum_{i=1}^n [\{Y_i - m(X_i)\} + \{m(X_i) - \hat{m}_h(X_i)\}]^2 w(X_i) \\ &= \frac{1}{n} \sum_{i=1}^n \varepsilon_i^2 w(X_i) + OKH(h) - \frac{2}{n} \sum_{i=1}^n \varepsilon_i \{\hat{m}_h(X_i) - m(X_i)\} w(X_i) \end{aligned}$$

ifadesi elde edilir. Burada  $\varepsilon_i = Y_i - m(X_i)$ 'dir. İlk terim  $\frac{1}{n} \sum_{i=1}^n \varepsilon_i^2 w(X_i)$   $h$ 'den bağımsızdır ve ikinci terim  $OKH(h)$ 'dir. Üçüncü terim  $\sum_{i=1}^n \varepsilon_i \{\hat{m}_h(X_i) - m(X_i)\} w(X_i)$  olmasaydı,  $p(h)$ 'yi optimize etmek  $OKH(h)$ 'yi optimize etmekle aynı sonucunu verir denilebilirdi. Fakat  $p(h)$ 'nin koşullu beklenen değeri hesaplanırsa,

$$E\{p(h)|X_1 = x_1, \dots, X_n = x_n\} \quad (2.3)$$

$$\begin{aligned} &= \frac{1}{n} \sum_{i=1}^n \sigma^2(x_i) w(x_i) + E\{OKH(h)|X_1 = x_1, \dots, X_n = x_n\} \\ &\quad - \frac{2}{n^2} \sum_{i=1}^n W_{hi}(x_i) \sigma^2(x_i) w(x_i) \end{aligned}$$

Burada  $W_{hi}(x)$ ,  $K_h(x - X_i)/\hat{f}_h(X_i)$  ( $\hat{f}_h(X_i) = n^{-1} \sum_{i=1}^n K_h(x - X_i)$ ) şeklindeki Nadaraya-Watson ağırlıklarındır. (2.3)'teki üçüncü terimin (2.2)'deki varyans  $v(h)$  ile aynı hızda sifıra yaklaştığı ve negatif işaretli olduğu görülür. Bu yüzden,  $p(h)$ 'yi

optimize eden pencere genişliği  $OKH(h)$ 'yi optimize eden pencere genişliğinin aşağıdan yanlı bir kestirimi olduğu gibi,  $p(h)$  de  $OKH(h)$ 'nin aşağıdan yanlı bir kestirimidir denir.

$OKH(h)$ 'nin kestirimindeki bu aşağıdan yanlılığı gidermek için önerilen bir yaklaşım,  $\hat{m}_h(X_i)$  teriminin *bir gözlem dışarı (leave-one-out)* kestirici  $\hat{m}_{h,-i}(X_i)$  ile değiştirilmesine dayanan çapraz geçerlilik yöntemidir.

Çapraz geçerlilik,  $Y_i$  gözleminin kendi tahmini için  $\hat{m}_h(X_i)$ 'de kullanılması sorununu bir gözlem dışarı kestiricisi ile çözmeye çalışır.

$$\hat{m}_{h,-i}(X_i) = \frac{\sum_{j \neq i} Y_j K_h(X_i - X_j)}{\sum_{j \neq i} K_h(X_i - X_j)}$$

Burada her  $X_i$  noktasındaki  $\hat{m}_h(\bullet)$ 'nin kestiriminde  $i$ . gözlem hariç tutulur. Böylece çapraz geçerlilik fonksiyonu elde edilir:

$$CG(h) = \frac{1}{n} \sum_{i=1}^n \{Y_i - \hat{m}_{h,-i}(X_i)\}^2 w(X_i)$$

$\hat{m}_h(X_i)$ 'nin yerine  $\hat{m}_{h,-i}(X_i)$ 'i kullanılırsa, (2.3)'deki üçüncü terimin beklenen değerinin sıfıra eşit olduğu görülür:

$$E \left[ -\frac{2}{n} \sum_{i=1}^n \varepsilon_i \{\hat{m}_{h,-i}(X_i) - m(X_i)\} w(X_i) \mid X_1 = x_1, \dots, X_n = x_n \right] = 0$$

Bu, ((2.3)'deki ilk terim  $h$ 'den bağımsız olduğundan)  $CG(h)$ 'yi optimize etmenin  $OKH$ 'yi optimize etmekle aynı şey olduğu anlamına gelir (Hardle vd., 2004). Çapraz geçerlilik yaklaşımının en büyük avantajı, kolay hesaplanabilir ve herhangi bir regresyon modeli için uygulanabilir olmasıdır (Demir, 2005).

Çapraz geçerlilik için önerilen çeşitli yaklaşımlar kullandıkları hata ölçütü bakımından farklılık gösterir. Bu yaklaşımlar aşağıda kısaca anlatılmaktadır.

### 2.3.1.1. En Küçük Kareler Çapraz Geçerlilik Yöntemi

Pencere genişliğinin seçimi için çapraz geçerlilik yönteminin uygulanmasında THKO ölçütü de kullanılmaktadır. THKO ölçütünü optimize eden çapraz geçerlilik yöntemi *en küçük kareler çapraz geçerlilik* olarak adlandırılır (Hepsağ, 2007).

$THKO(h) = E \int \{\hat{m}_h(x) - m(x)\}^2 dx$  olarak verilmişti. Gerekli işlemler yapıldığında:

$$THKO(h) - \int m(x)^2 dx = E \left[ \int \hat{m}_h(x)^2 dx - 2 \int \hat{m}_h(x)m(x) dx \right]$$

şeklinde elde edilen eşitliğin sağ taraftaki kısmının  $m(\bullet)$ 'e bağlı olduğu görülür. Bu ifade için kullanılan yansız bir kestirici,

$$EKKCG(h) = \int \hat{m}_h(x)^2 dx - 2n^{-1} \sum_{i=1}^n \hat{m}_{h,-i}(X_i)$$

şeklindedir. Burada,

$$\hat{m}_{h,-i}(X_i) = (n-1)^{-1} h^{-1} \sum_{j \neq i} K\left(\frac{X_i - X_j}{h}\right)$$

ifadesi bir gözlem dışarı kestiricisidir. Buradaki mantık, çapraz geçerlilikte olduğu gibi, örneklemin bir kısmını kullanarak diğer kısmı hakkında bilgi elde etmektir. EKKCG fonksiyonunu optimize eden pencere genişliği, THKO ölçütünü de optimize etmektedir (Wand ve Jones, 1995).

### 2.3.1.2. Beklenen Kullback-Leibler Çapraz Geçerlilik Yöntemi

Bu teknikte  $\log(\hat{\sigma}^2) + \psi(H)$  ifadesini minimize eden pencere genişliği tercih edilir. Bu ifadede  $\hat{\sigma}^2$  bileşeni,

$$\frac{1}{n} \sum_{i=1}^n \{Y_i - \hat{m}_h(X_i)\}^2 = \frac{Y'(I-H)'(I-H)Y}{n}$$

şeklindedir (Hurvich, Simonoff ve Tsai, 1998). Burada  $H$ ,  $n \times n$  ağırlık fonksiyonu yani çekirdek ağırlıkların matrisidir ve



$$H_{ij} = K_{h,ij} / \sum_{l=1}^n K_{h,il}$$

$$K_{h,ij} = \prod_{s=1}^q h_s^{-1} k((X_{is} - X_{js})/h_s)$$

olarak tanımlanır (Li ve Racine, 2007).

$\psi(\cdot)$  bileşeni ise  $\hat{m}_h$ 'nin düzgünlüğünü ayarlayan bir ceza fonksiyonudur.  $\psi$ 'nin bilinen tercihleri arasında,

*Genelleştirilmiş Çapraz Geçerlilik* ( $\psi(H) = -2 \log\{1 - \text{tr}(H)/n\}$ , GCG),

*Akaike Bilgi Kriteri* ( $\psi(H) = 2 \text{tr}(H)/n$ , ABK) ve

*T* (Rice, 1984) ( $\psi(H) = -\log\{1 - 2\text{tr}(H)/n\}$ ) yer alır.

Akaike bilgi kriteri ve genelleştirilmiş çapraz geçerlilik nispeten düşük cezalara sahip olduğundan az düzleştirmeye eğilimlidir. T ise bunun aksine  $\text{tr}(H)/n \geq 0,5$  için sonsuz olacağından çok güçlü cezaya sahip olup, pencere genişliği çok küçük alındığında aşırı düzleştirmeye neden olur. Akaike bilgi kriterinin yanlılığını gidermek için geliştirilmiş  $ABK_C$  ise bu iki uç arasında yer alır. Bu ölçüt,

$$ABK_C = \ln(\hat{\sigma}^2) + \frac{1 + \text{tr}(H)/n}{1 - \{\text{tr}(H) + 2\}/n}$$

olarak gösterilir (Hurvich vd., 1998).

### 2.3.2. Cezalandırma Fonksiyonları

Cezalandırma fonksiyonları, genellikle  $h$ 'nin çok küçük olması gibi istenmeyen çözümleri kontrol altında tutmak için kullanılmaktadır. Bu fonksiyonlar kullanılarak çok küçük pencere genişlikleri cezalandırılır ve pencere genişliğinin çok küçük olması engellenir (Demir, 2005). Cezalandırma fonksiyonları için birçok seçenek olmakla birlikte, pencere genişliği seçimi için genellikle birinci dereceden Taylor açılımına sahip bir cezalandırma fonksiyonu kullanılır. Bu fonksiyon,

$$\mathcal{E}(u) = 1 + 2u + O(u^2), u \rightarrow 0$$

şeklindedir. Bu fonksiyonun kullanılması ile elde edilecek pencere genişliği ise,

$$G(h) = n^{-1} \sum_{j=1}^n \left( Y_j - \hat{m}_h(X_j) \right)^2 \mathbb{E} \left( n^{-1} W_{hj}(X_j) \right) w(X_j)$$

olarak ifade edilir (Hardle, 1990).

### 2.3.3. Plug-in Yöntemi

Plug-in yaklaşımı, optimal pencere genişliğinin belirlenmesinde bilinmeyen fonksiyonlar yerine onların tutarlı kestiricilerini ( $\hat{m}_h''(x), \hat{\sigma}_h^2(x), \hat{f}_h(x), \hat{f}_h'(x)$ ) kullanır (Hardle ve Linton, 1994). Bu yaklaşım çapraz geçerlilik ve cezalandırma fonksiyonu yöntemleri kadar etkin sonuçlar vermesine rağmen, ilk adımda pencere genişliğinin seçiminde bir belirsizliğe sahiptir (Hepsağ, 2007).

## 2.4. DİĞER PARAMETRİK OLMAYAN REGRESYON KESTİRİCİLERİ

### 2.4.1. En Yakın Komşuluk Kestiricisi (k-NN)

Çekirdek kestiricisi  $x$  etrafında, genişliğinin  $h$  pencere genişliği tarafından belirlendiği, sabit bir komşulukta cevap değişkenlerinin ağırlıklı ortalamalarını hesaplayan bir metottur. k-en yakın komşuluk kestiricisi de  $x$  etrafında bir komşulukta cevap değişkenlerinin ağırlıklı ortalamasını hesaplamaktadır. Fakat buradaki önemli fark, komşuluğun genişliğinin sabit değil değişken olmasıdır.

k-NN kestiricisi şu şekilde gösterilir:

$$\hat{m}_k(x) = \frac{1}{n} \sum_{i=1}^n W_{ki}(x) Y_i$$

$\{W_{ki}(x)\}_{i=1}^n$  şeklindeki ağırlıklar ise,

$$W_{ki} = \begin{cases} \frac{n}{k}, & i \in J_x \text{ ise} \\ 0, & \text{diğer} \end{cases}$$

olarak tanımlanır.  $J_x = \{i: X_i, x$ 'e en yakın  $k$  gözlemlerden biridir. $\}$ .  $k$ , bu kestirimin pencere genişliğidir (Hardle vd., 2004).  $\{(X_i, Y_i)\}_{i=1}^5$  şeklindeki veri kümesine ait noktalar  $\{(1,5), (7,12), (3,1), (2,0), (5,4)\}$  olsun.  $x = 4$  ve  $k = 3$  için k-NN kestirimi

olan  $\hat{m}_k(x)$ 'in hesaplanmasında 4'e en yakın 3 adet gözlemin son üç veri noktası olduğu düşünülür. Dolayısıyla  $J_x = J_4 = \{3,4,5\}$ 'tir. Her bir veri noktası için ağırlıklar ise,

$$W_{k1}(4) = 0, W_{k2}(4) = 0, W_{k3}(4) = 1/3,$$

$$W_{k4}(4) = 1/3, W_{k5}(4) = 1/3$$

şeklinde olup,  $\hat{m}_3(4) = (1 + 0 + 4)/3 = 5/3$  olarak elde edilir (Hardle, 1990).

k-NN kestiricisi, tekbiçimli çekirdeğe ( $K(u) = \frac{1}{2}I(|u| \leq 1)$ ) ve değişken pencere genişliğine ( $h = R(k)$ ) sahip bir çekirdek kestiricisi olarak görülebilir.  $R(k)$ :  $x$  ve onun en uzak  $k$ -en yakın komşusu arasındaki mesafedir.

$$\hat{m}_k(x) = \frac{\sum_{i=1}^n K_R(x - X_i)Y_i}{\sum_{i=1}^n K_R(x - X_i)}$$

k-NN kestiricisi bu anlamda, tekbiçimli çekirdeklerden diğer çekirdeklere genelleştirilebilir. Genelleştirilmiş k-NN kestiricisinin yan ve varyansı aşağıdaki gibidir:

$k \rightarrow \infty, k/n \rightarrow 0$  ve  $n \rightarrow \infty$  iken,

$$E\{\hat{m}_k(x)\} - m(x) \approx \frac{\mu_2(K)}{8f_X(x)^2} \left\{ m''(x) + \frac{m'(x)f_X'(x)}{f_X(x)} \right\} \left( \frac{k}{n} \right)^2$$

$$Var\{\hat{m}_k(x)\} \approx 2\|K\|_2^2 \frac{\sigma^2(x)}{k}$$

Nadaraya-Watson çekirdek kestiricisinin varyansının aksine, k-NN kestiricisinin varyansı  $f_X(x)$ 'e bağlı değildir. Bu nedenle,  $Var\{\hat{m}_k(x)\} \sim \frac{1}{k}$  olur.  $k = 2nhf_X(x)$  olarak alınırsa, bir çekirdek kestiricisiyle yaklaşık olarak aynı olan bir k-NN kestiricisi elde edilir (Hardle vd., 2004).

#### 2.4.2. Spline Modelleri

Spline düzleştirme yöntemi, bir  $m(\cdot)$  fonksiyonunun veriye uygunluğu için artık kareler toplamı (RSS) ölçütünü kullanır. Bu ölçüt,

$$RSS = \sum_{i=1}^n \{Y_i - m(X_i)\}^2$$

şeklinde ifade edilir ve  $m(X_i) = Y_i$  ( $i = 1, \dots, n$ ) olarak tanımlandığında minimum olur. Fakat bu durumda, verideki mevcut yapı dikkate alınmaksızın sadece veri interpolate edilir (Hardle vd., 2004). Diğer bir deyişle, yerel değişimler olmadan bir eğri üretilir (Hepsağ, 2007). Spline metodu bu sorunu  $m(\cdot)$ 'nin düzgünsüzlüğünü cezalandıran  $\int \{m''(x)\}^2 dx$  şeklindeki bir dengeleyici ile çözmeye çalışır.

$$S_\lambda(m) = \sum_{i=1}^n \{Y_i - m(X_i)\}^2 + \lambda \int \{m''(x)\}^2 dx \quad (2.4)$$

$$\hat{m}_\lambda = \arg \min_m S_\lambda(m)$$

$\lambda$  parametresi, minimizasyondaki dengeleyiciye verilen ağırlıkları kontrol eder.  $\lambda$ 'nın büyük olması dengeleyiciye daha fazla ağırlık verildiği anlamına gelir.  $\lambda \rightarrow \infty$  olduğunda,  $\hat{m}_\lambda(\cdot)$  fonksiyonu doğrusal forma yaklaşır.  $\lambda \rightarrow 0$  olduğunda ise fonksiyon, sadece  $Y$  gözlemlerinin interpolasyonu olacaktır (Hardle vd., 2004). Diğer bir deyişle,  $Y$  değerleri interpolasyon fonksiyonuna yaklaşacak, dolayısıyla eşitlik 2.4'teki ilk terim herhangi bir  $m(x)$  fonksiyonu için sıfır olacaktır. Bundan dolayı, spline düzleştirme tekniği, uyum iyiliği ile yerel değişim arasındaki değiş tokuşu araştırma işlemi olarak kabul edilir (Çatalbaş, 2006).

$[a, b] = [X_{(1)}, X_{(n)}]$  aralığında iki kez türevi alınabilen tüm fonksiyonlar düşünüldüğünde,  $S_\lambda(\cdot)$  fonksiyonunun minimizasyon problemi *kübik spline* ( $\hat{m}_\lambda(x)$ ) olarak adlandırılan tek bir çözüme sahiptir.  $\hat{m}_\lambda(x)$ , ardışık  $X$  değerleri arasında kübik polinomialdır.  $X_i$  gözlem noktalarında  $\hat{m}_\lambda(x)$  eğrisi ve ilk iki türevi süreklidir, fakat üçüncü türevi süreksiz olabilir.  $X_{(1)}$  ve  $X_{(n)}$  sınır noktalarında ise  $\hat{m}_\lambda(x)$ 'in ikinci türevi sıfırdır.

Veri vektörü  $\{Y_i\}_{i=1}^n$ ,  $n$  tane koordinat vektörünün doğrusal bir bileşimi olarak yazıldığında  $W_{\lambda i}(x)$  ağırlıkları ortaya çıkmaktadır. Bu durumda  $\hat{m}_\lambda(x)$ ,

$$\hat{m}_\lambda(x) = n^{-1} \sum_{i=1}^n W_{\lambda i}(x) Y_i$$

şeklinde ifade edilir. Burada spline,  $Y$  gözlemlerinde doğrusaldır. Örnek büyüklüğüne eşit olan  $Y$  değeri hariç, sıfır olan  $Y$  değerlerine spline düzleştirme metodu uygulanarak ağırlık fonksiyonu çizilebilir. Ancak  $\{W_{\lambda i}(x)\}_{i=1}^n$ 'in fonksiyonel formunun açıkça yazılması zor, düzleştirme parametresi  $\lambda$  ve tasarı noktalarına olan bağımlılığı çok karmaşıktır.  $W_{\lambda i}(\cdot)$ ,  $(X_1, 0), \dots, (X_i, n), \dots, (X_n, 0)$  gözlemleri için bir spline'dır. Daha genel bir ifade ile,  $t$  noktasındaki spline  $W_\lambda(\cdot, t)$  olarak tanımlanır.

Silverman (1984), etkili bir ağırlık fonksiyonu  $W_\lambda(\cdot, t)$ 'nin

$$K_s(u) = \frac{1}{2} \exp(-|u|/\sqrt{2}) \sin(|u|/\sqrt{2} + \pi/4)$$

şeklinde tanımlanmış bir çekirdek fonksiyonuna benzediğini göstermiştir (Hardle, 1990).

### 2.4.3. Ortogonal Seriler

Uygun şartlar altında, bazı fonksiyonlar temel fonksiyon serileri kullanılarak tanımlanabilir. Örneğin, fourier serileri kullanılarak  $m(\cdot)$  fonksiyonu,

$$m(x) = \sum_{j=0}^{\infty} \beta_j \varphi_j(x) \quad (2.5)$$

şeklinde ifade edilebilir.  $\{\varphi_j(x)\}_{j=0}^{\infty}$  bilinen temel fonksiyonlar,  $\{\beta_j\}_{j=0}^{\infty}$  ise bilinmeyen fourier katsayılarıdır. Burada amaç, bilinmeyen fourier katsayılarını kestirmektir. Bu amaç doğrultusunda karşılaşılan bir sorun sonsuz sayıda, sıfırdan farklı  $\beta_j$  olduğunda (2.5)'teki toplamın sonsuz olacağı, diğer bir sorun ise sınırlı sayıda gözlemde sonsuz sayıda katsayı kestirilemeyeceğidir. Bu sorunlardan kurtulmak adına, örnek büyüklüğünün ( $n$ ) bir fonksiyonu olan ve  $N$  ile gösterilen bir sayı belirlemek gerekir.

Seri kestiricileri; temel fonksiyonun seçimi,  $n$ 'den küçük bir tamsayı olan  $N$ 'nin seçimi ve  $N$  tane bilinmeyen katsayının uygun bir metot ile kestirimi şeklinde üç adımda incelenir.  $N$ , bu kestiricilerin düzleştirme parametresidir.  $N$  büyük olduğunda fourier serilerindeki terim sayısı da artacak ve kestirim veriyi interpolate etme yönünde gelişecektir.  $N$ 'nin küçük değerleri için ise nispeten düzgün kestirimler üretilecektir. Buradaki  $\beta_j$  katsayıları *hızlı fourier dönüşümü* kullanılarak kestirilebilir (Hardle vd., 2004).

#### 2.4.4. Medyan Düzleştirme Tekniği

Medyan düzleştirme tekniği, koşullu olasılık fonksiyonu yerine koşullu medyan fonksiyonunun kestirimini ele alan en yakın komşuluk tekniği olarak tanımlanabilir. Medyan düzleştiricisi,

$$\hat{m}(x) = med\{Y_i: i \in J_x\}$$

şeklinde tanımlanır.  $J_x = \{i: X_i, x$ 'e en yakın  $k$  gözlemlerden biridir.}. Burada,  $x$ 'e en yakın  $k$  adet gözlemlerden biri olan  $X_i$  değerine karşılık gelen  $Y_i$  değerlerinin medyanı hesaplanır.

Koşullu medyan  $med(Y|X = x)$ , aykırı değerlere karşı koşullu olasılıktan  $E(Y|X = x)$  daha dayanıklıdır (Hardle vd., 2004). Ayrıca, regresyon eğrisinde beklenmeyen süresiz davranışların varlığında bile iyi modelleme gücüne sahiptir (Hepsağ, 2007).

## ÜÇÜNCÜ BÖLÜM

### TÜRKİYE'DE ZARURİ HARCAMA İLE GELİR ARASINDAKİ İLİŞKİNİN ÇEKİRDEK REGRESYON TEKNİĞİ İLE İNCELENMESİ

Hanehalkı harcamalarının anlaşılması hem ekonomik dinamikler hem de politika üretimi açısından önem taşımaktadır. Toplam talebin nasıl biçimlendiğinin anlaşılmasında harcamalar ile gelir arasındaki ilişkinin ölçülmesi başlangıç noktasıdır. Diğer taraftan, mikro iktisat teorisinde sağlam bir yer edinmiş olmasa da sosyo-demografik faktörlerin tüketimin ve harcamaların önemli bir belirleyicisi olduğu inkâr edilemez.

Çalışmanın bu bölümünde, Türkiye’de hanehalkı zaruri harcamaları (gıda ve alkolsüz içecekler, sağlık ve eğitim) ile hanehalkı kullanılabilir geliri ve sosyo-demografik özellikleri arasındaki ilişki 2008 ve 2010 yıllarına ait TÜİK Hanehalkı Bütçe Anketi verileri bazında incelenecektir.

İlerleyen alt bölümlerde çeşitli mal ve hizmetlere yönelik talebin incelenmesi kapsamında ele alınan bazı çalışmalar özetlenecek, çalışmada kullanılan veri seti tanıtılacak ve üretilen tahminler sunulacaktır.

#### 3.1. LİTERATÜR ÖZETİ

Literatürde gerek tek tek tüketim mallarına gerek belli mal sepetlerine yönelik talebi inceleyen çok sayıda çalışma bulunmaktadır. Konunun ilgi çekiciliğini kaybetmemesi sayesinde, bu çalışmalar zaman içinde süreklilik arz etmektedir. Bu çalışmanın arka planını oluşturmak için incelenen çalışmaların çoğunluğunda, literatürde yerleşik hale gelmiş talep modellerinin kullanıldığı söylenebilmektedir. Bunlar *Yaklaşık İdeal Talep Sistemi*, *Doğrusal Formda Yaklaşık İdeal Talep Sistemi*, *Working – Laser Talep Modeli* olarak sayılabilir.

Anılan talep modellerini tahmin etmeye yönelik yöntemler ise *En Küçük Kareler*, *Tobit Analizi* ve *Sıralı Logit Analizi* olarak sayılabilmektedir. Yöntemsel açıdan dikkat çeken önemli bir husus, parametrik yöntemlerin literatürdeki baskınlığıdır. Talep fonksiyonlarının teoride iyi parametrize edilmiş olması ve iyi tanımlı (sürekli ve türevlenebilir) fonksiyonlar halinde tanımlanması, ampirik düzlemde parametrik yöntemleri ön plana taşıyan temel neden olarak görülebilir. Uygulamada önem taşıyan

diğer bir husus da, parametrik modellerin kestirim sonrasında tahmin amaçlı kullanımının daha kolay olmasıdır. Araştırmacıların tahmin sürecinde kolayca senaryo üretebilmeyi ve ürettikleri senaryolara dayalı talep öngörülerini en az maliyetle hesaplayabilmeyi istemeleri anlaşılabilir bir davranış biçimidir.

Diğer taraftan, tüketim harcamalarının incelenmesinde yöntemsel açıdan farklılaşan çalışmalar da bulunmaktadır. Çolak, Öztürkler ve Tokatlıođlu (2008) *Dilim Regresyon* yöntemini, Bierens ve Pott-Buter (1990) *Parametrik Olmayan Çekirdek Regresyon* yöntemini, Gozalo (1997) *Düzgün Koşullu Moment Bootstrap* yöntemini ve Lee ve Chern (1989) *Lineer Olmayan Regresyon* yöntemini kullanmıştır. Bunlar arasında yalnız Bierens ve Pott-Buter (1990) çalışmasında bu çalışmada da kullanılan parametrik olmayan çekirdek regresyon metodu kullanılmıştır.

**Tablo 3.1.** Literatürde Ele Alınan Gıda Maddeleri ile Çeşitli Mal ve Hizmetler

1978	Atalay	Zeytinyađı, Ayçiçeđi yađı, Pamukyađı
1980	Yurdakul	Hayvansal Gıda Maddeleri
1985	Musgrove	Tahıl, Sebze, Meyve, Et, Süt Ürünleri, Yađ, Şekerleme, Alkolsüz İçecekler, Diğer
1987	Georgantelis vd.	Tahıl, Et, Yađ, Şekerleme, Süt Ürünleri, Meyve, Sebze, İçecek, Diğer
1987	Kaytaş	Alkollü İçki (Rakı, Bira, Şarap ve Diğerleri)
1991	Kasnakođlu	Et, Tahıl, Süt ve Ürünleri, Meyve, Sebze, Çay, Şeker, Zeytin, Alkol ve Sigara
1996	Ekinci	Ekmek, Şeker, Margarin, Et, Süt ve Ürünleri
1998	Baydemir	Tahıl, Sebze, Meyve, Et, Süt ve Ürünleri
2001	Koç, Tan	Süt, Yođurt, Peynir, Tereyađı
2006	De Meyrick, Yusuf	Sigara
2006	Tansel, Bircan	Özel Eğitim Harcamaları
2007	Djibuti vd.	Sigara
2007	Hatırlı vd.	Kırmızı Et, Beyaz Et, Tavuk Eti
2008	Emeç, Gülay	Alkol
2008	Öz	Cepten Sağlık Harcamaları
2008	Uraz	Eđlence ve Kültür Harcamaları
2011	Kim, Yang	Sađlık Harcamaları

Tablo 3.1’de belirtilen, literatürde analize tabi tutulan mal ve hizmetler incelendiđinde, hanehalkı tüketiminde merkezi ve çoklukla hayati önem taşıyan mal ve hizmetlerin ön plana çıktığı görülmektedir.

Çalışmalarda tüketimin gelire nasıl bađlı olduđu temel araştırma sorusu olarak karşımıza çıkmaktadır. Bu kapsamda seçilmiş çalışmalarda sunulan gelir esneklikleri Tablo 3.2’de verilmiştir.



**Tablo 3.2.** Seçilmiş Çalışmalarda Gelir Esneklikleri

		Gıda	Yakıt	Giyim	Sağlık	Ulaşım	Konut	Eğitim	Evde Yeme	Dışarıda Yeme	Eğlence	Kültür	Dayanıklı Mallar	Kişisel Bakım	Alkollü İçecekler	Tütün	Diğer
1961	Leser	Z	Z	L	.	Z	Z	.	.	.	.	.	L	.	Z	Z	L
1975	Mcelroy	.	B	B	L	.	.	B	.	L	B	.	L	L	L	.	.
1979	Izan, Clements	Z	.	B	.	.	B	.	.	.	.	.	.	.	.	.	L
1984	Barnes, Gillingham	.	.	L	.	.	Z	.	Z	L	.	.	.	.	.	.	.
1989	Lee, Chern	.	.	L	Z	Z	Z	.	Z	L	L	.	.	.	.	.	L
1995	Şenesen, Selim	Z	.	L	B	L	.	.	.	L	.	L	.	.	.	.	L
1998	Nişancı	Z	.	L	Z	L	Z	L	.	L	L	L	.	.	Z	Z	L
1999	Özer	Z	.	L	L	Z	Z	L	.	.	L	L	.	.	.	.	.
2002	Alpay, Koç	Z	.	L	Z	L	L	L	.	.	L	.	L	.	.	.	L
2005	Akbay	Z	.	.	Z	B	Z	L	.	.	B	B	L	.	Z	Z	B
2010	Fisunoğlu, Şengül	Z	.	Z	Z	.	Z	Z	.	.	.	.	.	.	.	.	.

Z: Birden küçük - zorunlu, L: Birden büyük - lüks, B: Yaklaşık olarak bire eşit - birim

**Tablo 3.3.** Literatürde Talebe Etkisi İncelenen Demografik Faktörler

1963	Leser	Ülke
1965	Houthakker	Ülke
1973	Parks, Barten	Ülke
1981	Pollak, Wales	Hanehalkı Büyüklüğü*, Çocuk Sayısı
1984	Barnes, Gillingham	Ev Sahipliği*, Hanehalkı Tipi*, Çocuk Sayısı
1993	Ersoy	Gelir*, Eğitim Düzeyi, Çocuk Yaşı, Hayat Devresi, Evlilik Yılı
1993	Sawtelle	Hanehalkı Büyüklüğü*, 18 Yaş Altı Çocuk Sayısı, Bölge, Barınma Durumu*, Yaş, Cinsiyet, Irk, Eğitim Düzeyi
1997	Driel, Nadall, Zeelenberg	Ülke
1997	Gozalo	Gelir*, Hanehalkı Büyüklüğü*, Yaş, Cinsiyet
1998	Jensen, Manrique	Gelir*, Bölge, Hanehalkı Yaş Bileşeni
1999	Akın	Gelir*, Hanehalkı Büyüklüğü*, Eğitim Düzeyi
1999	Özer	Bölge, Kır-Kent*
1999	Sarımışeli	Kır-Kent*
2000	Case	İrk
2000	Selim	Kır-Kent*
2001	Emeç	Bölge
2007	Beyaz	Kır-Kent*, Eğitim Düzeyi
2008	Choi, Lee	Gelir*, Hanehalkı Tipi*
2008	Çolak, Öztürkler	Yaş, Sosyal Güvenlik, Servet
2009	Zortuk	Cinsiyet
2010	Şahinli	Kır-Kent*

\* Bu çalışmada incelenen faktörler

Tablo 3.3'te görüldüğü üzere, literatürde pek çok sosyal ve demografik değişkenin talep üzerindeki etkileri incelenmiştir. Bu çalışmada ele alınan faktörler Tablo 3.3'te verilen literatürle büyük ölçüde uyum göstermektedir.

### 3.2. VERİ SETİ

Hanehalkı Bütçe Anketi; hanelerin sosyo-ekonomik yapıları, yaşam düzeyleri ve tüketim kalıpları hakkında bilgi veren, uygulanan sosyo-ekonomik politikaların test edilmesi amacıyla kullanılan en önemli kaynaklardan biridir. Bu çalışmada 1 Ocak-31 Aralık 2008 tarihleri arasında toplam 8640, 1 Ocak-31 Aralık 2010 tarihleri arasında ise toplam 13248 örnek hanehalkına uygulanan Hanehalkı Bütçe Anketleri'nden elde edilen veriler kullanılmıştır.

Çalışmada bağımlı değişken olarak *zaruri harcamalar* diye nitelendirilen gıda ve alkolsüz içecekler, sağlık ve eğitim harcamaları ele alınmış; zaruri harcamaların açıklanmasında ise *yıllık kullanılabilir gelir*, *hanehalkı büyüklüğü*, *OECD tanımlı hanehalkı büyüklüğü*, *hanehalkı tipi*, *kır-kent yerleşim göstergesi* ve *hane mülkiyet tipi* isimli değişkenler kullanılmıştır. Söz konusu değişkenlere ilişkin tanımlar Ek 3'te sunulmuştur.

### 3.3. TAHMİNLER

Çalışmanın bu kısmında, ikinci bölümde tanıtılan parametrik olmayan çekirdek kestirim tekniğinin 2008 ve 2010 yıllarına ait TÜİK Hanehalkı Bütçe Anketi verilerine uygulanması ile elde edilen sonuçlar sunulmuştur. Başlangıçta zaruri harcamaların sadece kullanılabilir gelir ile açıklandığı temel model tahmin edilmekte olup, sonraki tahminler sosyal ve demografik değişkenleri içermeyen temel model yapısına her seferinde yalnız bir sosyal/demografik değişken eklenerek oluşturulmuştur. Modellere dahil edilen açıklayıcı değişken sayısının belirlenmesinde, zaman maliyetleri önemli bir belirleyici olarak karşımıza çıkmaktadır. Bunun başlıca nedeni, parametrik olmayan çekirdek kestirim modellerinde en iyi pencere genişliği tespitinin ele alınan veri kümesinin büyüklüğünün karesi mertebesinde bir süre gerektirmesidir.

Çalışmada sunulan tahminler R ortamında np paketi kullanılarak üretilmiştir (Hayfield ve Racine, 2007). Tahminlerin üretilmesinde kullanılan kodların bir örneği Ek 2'de sunulmuştur.

**Tablo 3.4.** Betimleyici İstatistikler

2008	ZH	GL	HH	HHB	OECD	KK	MLK
<b>Ortalama</b>	8.335	9.660	3.2	3.8	2.2	1.6	1.5
<b>Medyan</b>	8.351	9.669	3.0	4.0	2.1	2.0	1.0
<b>En Yüksek</b>	11.910	12.633	7.0	23.0	9.6	2.0	4.0
<b>En Düşük</b>	3.178	6.915	1.0	1.0	1.0	1.0	1.0
<b>Standart Sapma</b>	0.623	0.705	1.657	1.837	0.766	0.459	0.935
<b>Çarpıklık</b>	-0.380	-0.007	0.233	1.503	1.539	-0.857	1.726
<b>Basıklık</b>	5.596	3.668	1.873	8.425	8.507	1.734	4.845
<b>Gözlem Sayısı</b>	8540	8540	8540	8540	8540	8540	8540
2010	ZH	GL	HH	HHB	OECD	KK	MLK
<b>Ortalama</b>	8.445	9.811	3.3	3.7	2.2	1.6	1.6
<b>Medyan</b>	8.459	9.811	3.0	4.0	2.1	2.0	1.0
<b>En Yüksek</b>	11.310	13.343	7.0	20.0	8.9	2.0	4.0
<b>En Düşük</b>	3.178	6.927	1.0	1.0	1.0	1.0	1.0
<b>Standart Sapma</b>	0.661	0.683	1.779	1.901	0.790	0.464	1.041
<b>Çarpıklık</b>	-0.400	0.020	0.256	1.498	1.477	-0.802	1.409
<b>Basıklık</b>	5.292	3.874	1.926	7.637	7.535	1.643	3.577
<b>Gözlem Sayısı</b>	10062	10062	10062	10062	10062	10062	10062

**Tablo 3.5.** Tahmin Edilen Parametrik Olmayan Modellerin Özellikleri

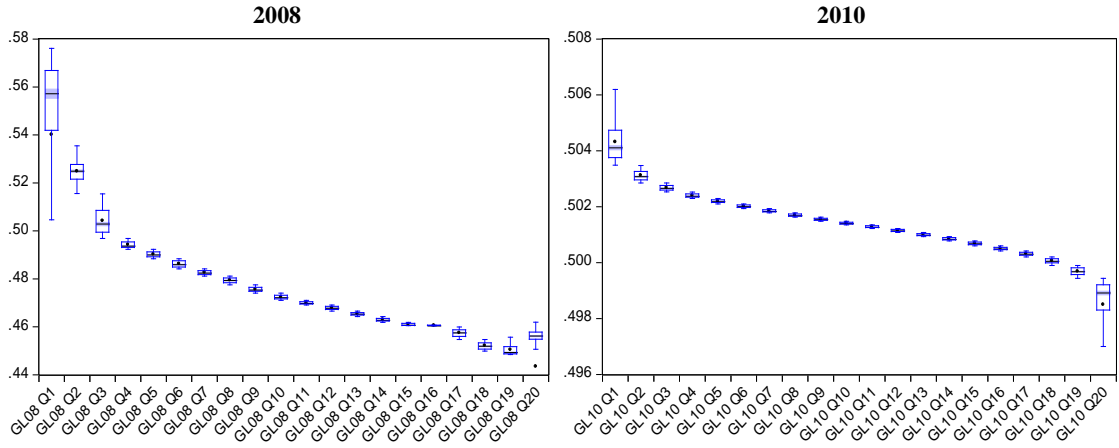
Kategorik Değişken	Çekirdek Tipi ve Bant Genişliği (Gelir)	Çekirdek Tipi ve Bant Genişliği (Kategorik)	KSH	R <sup>2</sup>	İlgili Şekiller
<b>2008</b>					
<b>Yok</b>	Epanechnikov(2) 0,50909	-	0,27332	0,29589	3.1
<b>Hanehalkı tipi</b>	Epanechnikov(2) 0,52436	Aitchison ve Aitken 0,02551	0,25724	0,33731	3.2, 3.3, Ek 1
<b>Hanehalkı büyüklüğü (TÜİK)</b>	Epanechnikov(2) 1,00932	Wang ve Van Ryzin 0,04407	0,25021	0,35542	3.4, 3.5, Ek 2
<b>Hanehalkı büyüklüğü (OECD)</b>	Epanechnikov(2) 0,60606	Epanechnikov(2) 1,26658	0,25177	0,35140	3.6, Ek 3
<b>Kır-Kent</b>	Epanechnikov(2) 0,60395	Aitchison ve Aitken 0,00000	0,27021	0,30390	3.7, 3.8, Ek 4
<b>Mülkiyet Tipi</b>	Epanechnikov(2) 0,32439	Aitchison ve Aitken 0,10334	0,26832	0,30875	3.9, 3.10, Ek 5
<b>2010</b>					
<b>Yok</b>	Epanechnikov(2) 2,08098	-	0,31997	0,26791	3.1
<b>Hanehalkı tipi</b>	Epanechnikov(2) 1,17199	Aitchison ve Aitken 0,01751	0,29629	0,32208	3.2, 3.3, Ek 1
<b>Hanehalkı büyüklüğü (TÜİK)</b>	Epanechnikov(2) 2,09637	Wang ve Van Ryzin 0,02330	0,29155	0,33292	3.4, 3.5, Ek 2
<b>Hanehalkı büyüklüğü (OECD)</b>	Epanechnikov(2) 57720,7	Epanechnikov(2) 0,86322	0,29369	0,32802	3.6, Ek 3
<b>Kır-Kent</b>	Epanechnikov(2) 1,15765	Aitchison ve Aitken 0,00000	0,31519	0,27883	3.7, 3.8, Ek 4
<b>Mülkiyet Tipi</b>	Epanechnikov(2) 1,30614	Aitchison ve Aitken 0,03106	0,31331	0,28313	3.9, 3.10, Ek 5

**KSH:** Kalıntı terimleri standart hatası. İlgili şekiller, tabloda özellikleri sunulan modellerin sonuçlarının çalışmadaki hangi şekillerde sunulduğunu göstermektedir.

Tahminlerde kullanılan zaruri harcama (ZH) ve kullanılabilir gelir (GL) serileri, orijinal serilerin doğal logaritması alınarak oluşturulmuştur. Hanehalkı büyüklüğü (HHB) serisi, orijinal veride 7'nin üzerinde olan değerlerin 7 değerine kesilmesi ile oluşturulmuştur. Söz konusu dönüşüm, örneklem büyüklüğüne kıyasla az sayıda

gözlemede rastlanan 7'den büyük HHB değerlerinin optimal pencere seçimi üzerindeki olumsuz etkilerini sınırlamaya dönük olarak yapılmıştır. OECD tanımlı hanehalkı büyüklüğü (OECD), hanehalkı tipi (HH), kır-kent yerleşim göstergesi (KK) ve hane mülkiyet tipiyle (MLK) ilgili olarak herhangi bir dönüşüm kullanılmamıştır. Veriye ilişkin tanımlayıcı istatistikler Tablo 3.4'te sunulmuştur.

Sunulan tahminlerin tümünde sürekli değişkenler için Epanechnikov ikinci derece çekirdek fonksiyonu, nominal kategori değişkenleri için Aitchison ve Aitken çekirdek fonksiyonu, ordinal kategori değişkenleri içinse Wang ve Van Ryzin çekirdek fonksiyonu kullanılmıştır. Tüm tahminlerde yerel doğrusal regresyon, en küçük kareler çapraz geçerlilik ve sabit pencere genişliği kullanılmıştır. Tahmin edilen parametrik olmayan kestirim modellerinin özellikleri Tablo 3.5'te sunulmuştur.



**Şekil 3.1.** Temel (Demografik Faktörleri İçermeyen) Modellerde Gelir Esneklikleri

2008 ve 2010 veri kümeleri için ayrı ayrı tahmin edilen temel modellerde zaruri harcamalar, hanehalkı kullanılabilir geliri ile açıklanmaktadır. Söz konusu modellerden 2008 verisindeki 8.540 ve 2010 verisindeki 10.062 gözlem noktası için elde edilen eğim tahminleri Şekil 3.1'de harcanabilir gelirin 20 sıklık dilimi için box-whisker grafikleri ile özetlenmektedir. Dikey eksen zaruri harcamaların gelir esnekliği tahminlerini, yatay eksen gelir sıklık derecelerini göstermektedir. Tahminlerin özetlenmesinde her yıl için o yıla ait 20 sıklık derecesi kullanılmıştır. Box-Whisker grafiğinde, dikdörtgen kutu grafiği çizilen verinin 1. ve 3. kartil arasındaki kısmını temsil etmekte, kutuyu enine kesen çizgi medyan değerini, bağımsız çizilen nokta ortalama değeri, medyanı çevreleyen taralı alan %95 güven aralığını, kutudan yukarı ve aşağı uzanan çizgiler ise

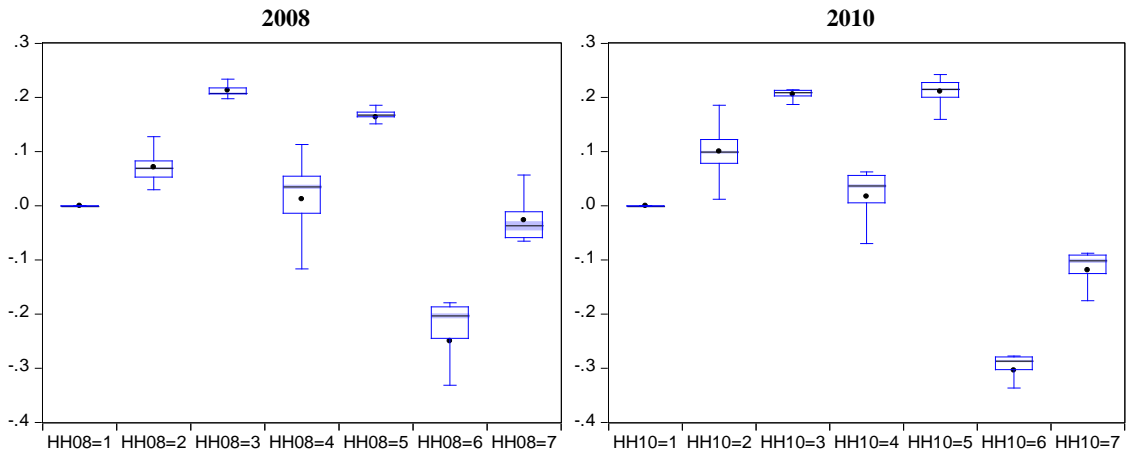
verinin dağılım aralığını göstermektedir. Grafiğin görsel basitliğini sağlamak için yakın ve uzak uç değerler grafiğe alınmamıştır. Sürekli değişkenlerin logaritmik olması nedeniyle eğim tahminleri her bir gözlem noktası için, o noktada geçerli esneklik değerine karşılık gelmektedir. Şekil 3.1 incelendiğinde gelir yükseldikçe zaruri harcamaların gelir esnekliğinin düştüğü gözlemlenmektedir. Daha üst gelir gruplarında haneye giren fazladan yüzde birlik gelirin zaruri harcamaya dönüşme oranı daha düşüktür. 2010 yılında söz konusu düşüşün daha sınırlı kaldığı ayrıca ifade edilebilir.

Şekil 3.1’de sunulan temel model sonuçları istatistiksel açıdan anlam taşımakla birlikte, oldukça sınırlı bir bilgi sağlamaktadır. Yapılan analizde gözlem birimi hanehalkı olduğundan, hanehalkının çeşitli sosyal ve demografik özelliklerinin zaruri tüketim fonksiyonunda yer alması beklenebilir. Bu çerçevede hanehalkı büyüklüğü, hanehalkının tipi, ikamet edilen konutun mülkiyet tipi ve hanehalkının coğrafi yerleşimi (kır/kent) gibi pek çok etmen dikkate alınmalıdır. Bu çalışmada sadece anılan değişkenler modellerde kullanılmıştır. Bu değişkenlerin analizine geçmeden önce, değişkenlere ilişkin hipotezlere kısaca değinmek yerinde olacaktır. Hanehalkı büyüklüğü hane olgunluğunun bir ölçüsü olarak ele alınmaktadır. Tipik bir aile esas alındığında, hanede yaşayan kişi sayısı eşlerin daha fazla çocuğa sahip olması veya yaşlıların bakımını üstlenmesi ile artabilir. Ailenin eline geçen parasal kazancı ne şekilde tüketime dönüştüreceğinin hane büyüklüğünden etkileniyor olması beklenmekle birlikte, beklenen etkinin yönü belirsizdir. Benzer bir durum hanehalkı tipi için de geçerlidir. Ailenin çekirdek aile olup olmaması, çocuk sayısı, çocukların yaşı ve benzeri etmenler zaruri harcamaların nasıl şekilleneceğini belirliyor olmalıdır. Hanehalkının kırdan mı yoksa kentte mi yerleştiğinin de zaruri harcamaları etkilemesi beklenmektedir. Son olarak, hanehalkının ikamet ettiği konuttaki mülkiyet biçimi konuta yönelik harcamaların başlıca belirleyicisi olarak zaruri harcamaları da etkiliyor olmalıdır.

Diğer tahmin sonuçlarının sunumundan önce metinde yer alan terminolojik bir noktanın altının çizilmesinde fayda görülmektedir. Başlangıçta da belirtildiği gibi çalışmada kullanılan ankete dayalı verinin gözlem birimi hanehalkıdır ve hanehalkı terimi tanımsal olarak aile terimiyle eşdeğerlik taşımamaktadır. Bununla birlikte, izleyen kısımlarda görüleceği gibi, TÜİK Hanehalkı Bütçe Anketi’nde hanehalkı tipi tanımları aileye referansla yapıldığı için izleyen tartışmalarda hane, hanehalkı ve aile terimleri birbirinin yerine kullanılabilir.

İzleyen egzersizlerde, yukarıda ele alınan sosyal ve demografik değişkenlerin zaruri harcamalara etkileri tahmin edilmekte ve sonuçlar iki şekilde özetlenmektedir. İlk olarak, hanehalkının belli bir kategoride yer almasının zaruri harcamaları nasıl etkilediği ortaya konulmakta olup, bu analiz parametrik regresyonlarda kukla değişkenler ile yapılan analizin parametrik olmayan karşılığını oluşturmaktadır. İkinci olarak ise, örneklemdaki her bir gözlem için elde edilen gelir esnekliği tahminlerinin hanehalkının farklı kategorilerinde nasıl olduğu özetlenmektedir. Esneklik bulgularının bu özeti, parametrik regresyondaki eğim değişkeni-kukla değişken etkileşiminin katsayısının parametrik olmayan karşılığıdır.

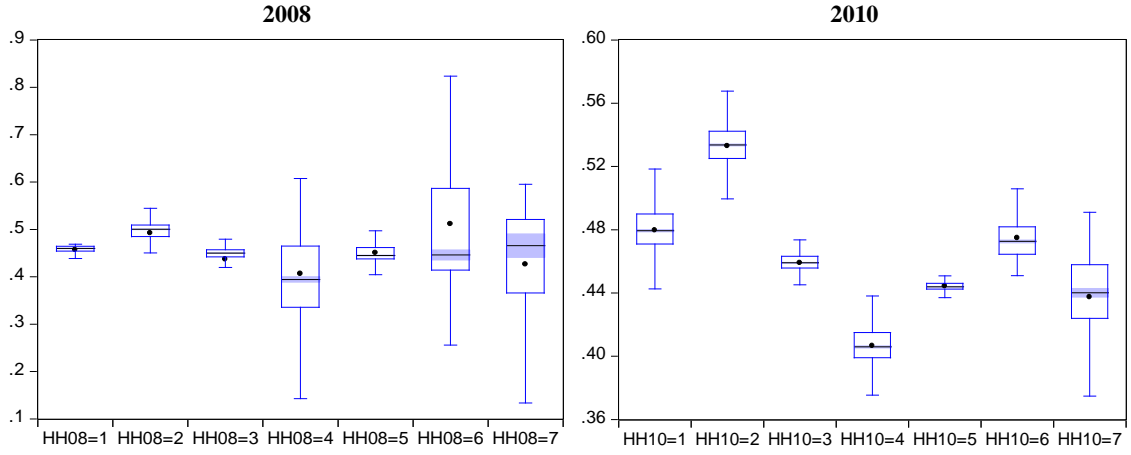
Zaruri harcamaların hanehalkı tiplerine göre nasıl şekillendiğine ilişkin tahmin sonuçları Şekil 3.2’de, hanehalkı tiplerine göre gelir esneklikleri ise Şekil 3.3’te sunulmuştur. Söz konusu grafiklerde hanehalkı tipi değişkeninin 1 olması 18 yaşından küçük tek çocuklu çekirdek aileyi, 2 olması 18 yaşından büyük tek çocuklu çekirdek aileyi, 3 olması 18 yaşından küçük iki çocuklu çekirdek aileyi, 4 olması en az biri 18 yaşından büyük iki çocuklu çekirdek aileyi, 5 olması 18 yaşından küçük üç veya daha fazla çocuklu çekirdek aileyi, 6 olması en az biri 18 yaşından büyük üç veya daha fazla çocuklu çekirdek aileyi ve 7 olması ise çocuksuz çekirdek aileyi ifade etmektedir.



**Şekil 3.2.** Hanehalkı Tipinin Zaruri Harcamalara Etkisi

Gerek 2008 gerek 2010 örnekleminde, çocukların ikisinin de 18 yaşından küçük olduğu iki çocuklu çekirdek ailelerin ve çocukların hepsinin 18 yaşından küçük olduğu üç veya daha fazla çocuklu çekirdek ailelerin zaruri harcamalarının örneklem geneline kıyasla daha yüksek olduğu söylenebilmektedir. Bunu sırasıyla, çocuğu 18 yaşından

büyük tek çocuklu çekirdek aile ve en az bir çocuğu 18 yaşından büyük iki çocuklu çekirdek aile izlemektedir. Çocuğu 18 yaşından küçük tek çocuklu çekirdek aile örneklemdaki ortalama zaruri harcama seviyesini temsil etmekte, en az bir çocuğun 18 yaşından büyük olduđu üç veya daha fazla çocuklu çekirdek aile ile çocuksuz çekirdek ailenin örneklem ortalamasına kıyasla zaruri harcama eğiliminin daha düşük olduđu görülmektedir.



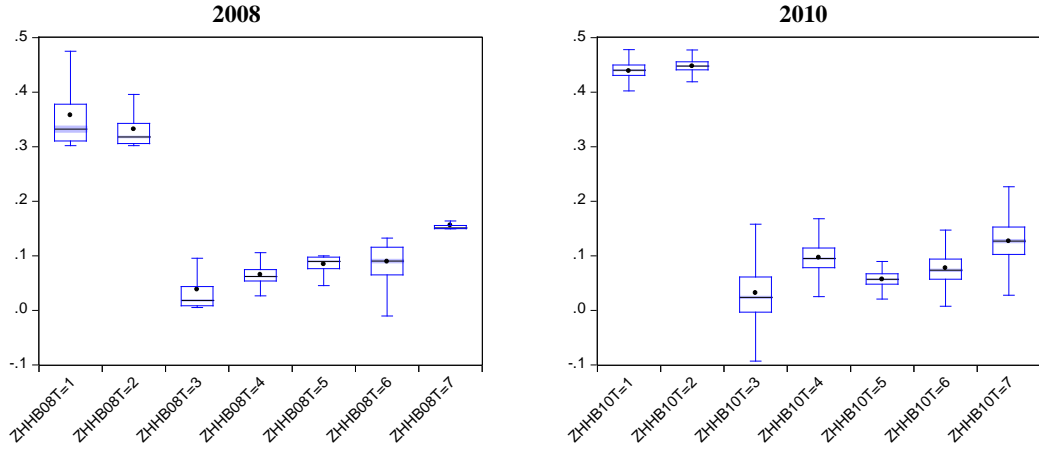
**Şekil 3.3.** Hanehalkı Tipi Kategorilerine Göre Gelir Esneklikleri

Şekil 3.3'te sunulan hanehalkı tipi kategorilerine göre gelir esneklikleri incelendiğinde; esnekliklerin çocuğun 18 yaşından büyük olduđu tek çocuklu aile, çocuğun 18 yaşından küçük olduđu tek çocuklu aile, en az bir çocuğun 18 yaşından büyük olduđu üç veya daha fazla çocuklu çekirdek aile, çocukların hepsinin 18 yaşından küçük olduđu iki çocuklu çekirdek aile, çocukların hepsinin 18 yaşından küçük olduđu üç veya daha fazla çocuklu çekirdek aile, çocuksuz çekirdek aile ve en az bir çocuğun 18 yaşından büyük olduđu iki çocuklu çekirdek aile sırasıyla azaldığı görülmektedir.<sup>1</sup>

Hanehalkı tipleri bazında yürütölen analiz zaruri harcama davranış kalıplarına ilişkin bazı fikirler vermekle birlikte, bulguların netliğinin büyük ölçüde sınırlı kaldığı görülmektedir. Bunun ardındaki birinci nedenin hanehalkı tipi analizinde çok sayıda ve nominal kategorilerin ele alınması olduđu söylenebilir. Bunun yanında, hanehalkı tiplerinin tanımında hanedeki çocuk sayısına ve çocukların yaşlarına yer veriliyor olmasının bulguların yeterince düzen arz etmemesine neden olduđu düşünülmektedir. Söz konusu potansiyel sorunları gidermek amacıyla, hanehalkı tipi yerine hanehalkı

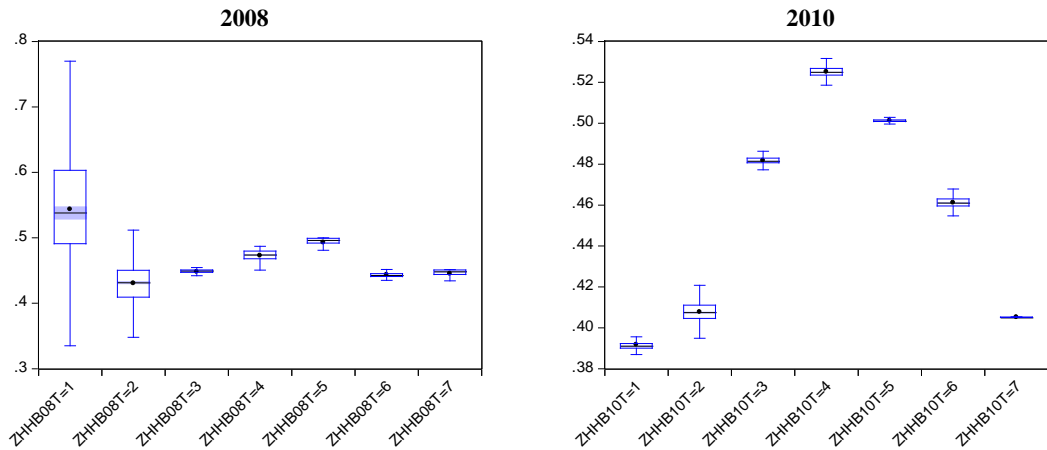
<sup>1</sup> Kategori değişkenlerini içeren modellerde elde edilen gelir esnekliklerinin gelir dilimlerine göre değişimi Ek 1'de sunulan grafiklerde yer almaktadır.

büyükliğini dikkate alan analizler yapılmıştır. Sayısal olarak kesikli değerler aldığı için ordinal kategori değişkeni olarak tanımlanan hanehalkı büyüklüğüne ilişkin sonuçlar Şekil 3.4 ve Şekil 3.5'te sunulmuştur. Grafiklerde, 7'nin üzerindeki hanehalkı büyüklükleri 7 değeri ile temsil edilmiştir.



**Şekil 3.4.** Hanehalkı Büyüklüğünün (TÜİK Tanımı) Zaruri Harcamalara Etkisi

Şekil 3.4'e göre hanehalkı büyüklüğünün zaruri harcamalara etkisi, HHB 1 veya 2 iken en yüksek olup, HHB 3'e yükseldiğinde en düşük değerini almaktadır. Hanede yaşayan birey sayısı 3'ten 7'ye doğru yükselirken etki de artmakta ancak HHB=1 ve HHB=2 iken aldığı değere ulaşmamaktadır.

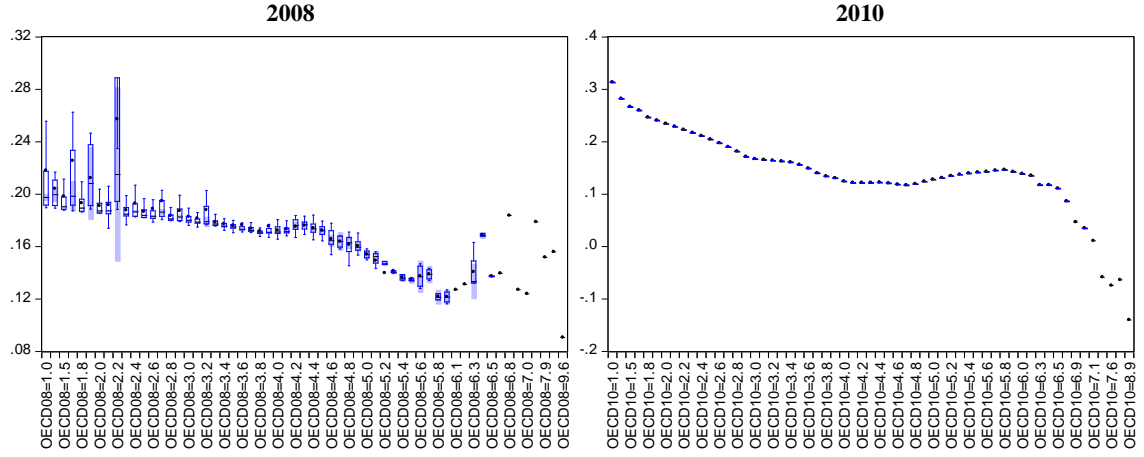


**Şekil 3.5.** Hanehalkı Büyüklüğüne (TÜİK Tanımı) Göre Gelir Esneklikleri

Şekil 3.5'te sunulan hanehalkı büyüklükleri bazında özetlenen gelir esnekliği tahminlerinin daha ilgi çekici olduğu düşünülmektedir. 2008 yılı örneklemindeki HHB=1 kategorisi göz ardı edildiğinde, hanehalkı büyüklüğü 5 bireye (2010 için 4)

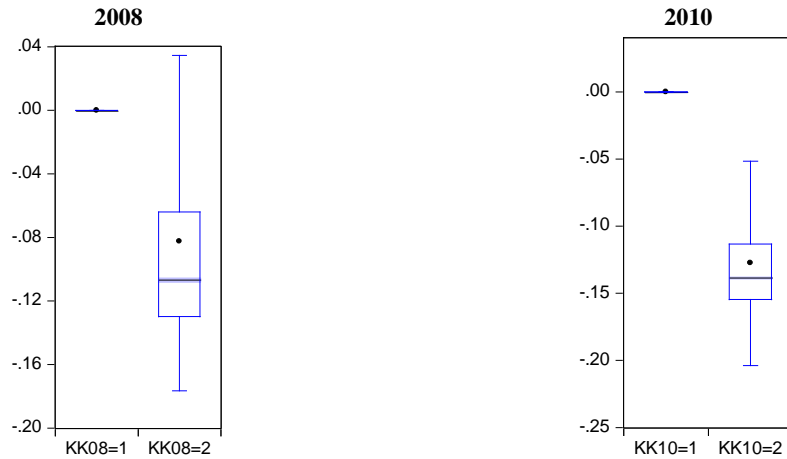


ulaşana kadar zaruri harcamaların gelir esnekliğinin arttığı sonrasında ise düştüğü gözlenmektedir.



**Şekil 3.6.** Hanehalkı Büyüklüğünün (OECD Tanımı) Zaruri Harcamalara Etkisi

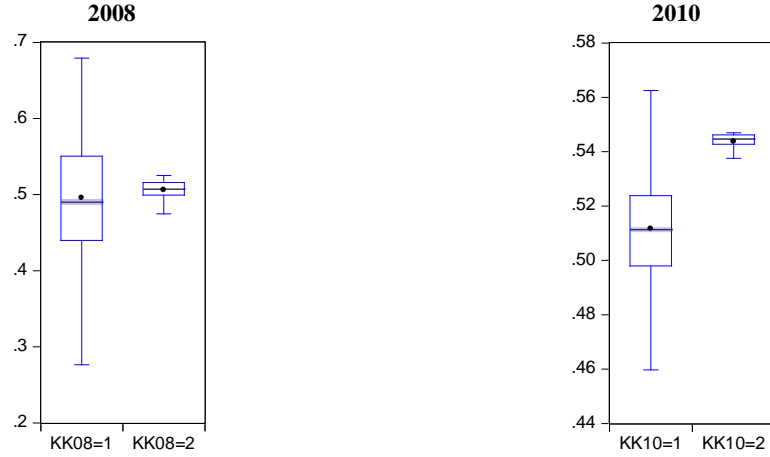
TÜİK tarafından kullanılan ve hanedeki birey sayısını esas alan hanehalkı büyüklüğü (HHB) tanımı yerine OECD tarafından geliştirilen efektif hanehalkı büyüklüğü (OECD) kullanıldığında da Şekil 3.4'te olduğu gibi daha az nüfuslu hanelerde zaruri harcamalar üzerindeki etkinin daha yüksek olduğu Şekil 3.6'dan anlaşılmaktadır. Bununla birlikte Şekil 3.4 için dile getirilen “keskin düşüş ve sonrasında tedrici artış” gözlemi OECD değişkeni için söz konusu olmamakta, zaruri harcamalar üzerindeki etki hanehalkı büyüklüğü arttıkça tedrici olarak azalmaktadır.



**Şekil 3.7.** Kır-Kent Ayrımının Zaruri Harcamalara Etkisi

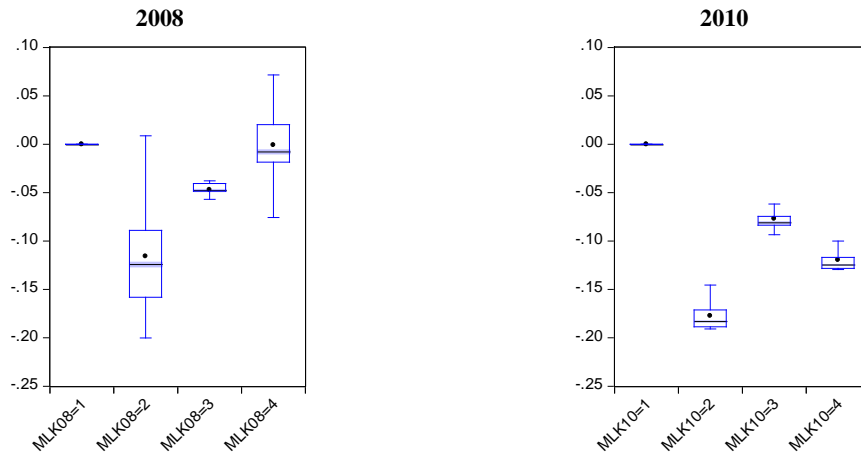
Hanehalkının yaşadığı yerleşim biriminin kır veya kent olmasının zaruri harcamalar üzerindeki etkisi Şekil 3.7 ve Şekil 3.8'de özetlenmiştir. Söz konusu

grafiklerde yer alan kır kent değişkeninde 1 kır, 2 ise kenti ifade etmektedir. Kentte yaşamın zaruri harcamalar üzerindeki tahmin edilmiş etkisi negatiftir. Diğer taraftan, kentte yaşayan hanehalkı için zaruri harcamaların gelir esnekliği kırdaki yaşayan hanehalkınıninkine kıyasla daha yüksektir.



**Şekil 3.8.** Kır-Kent Ayrımına Göre Gelir Esneklikleri

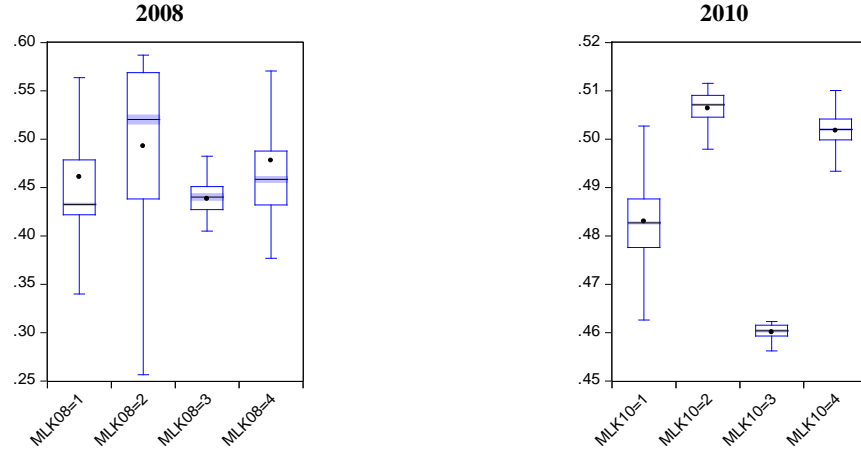
Analizlerde son olarak, hanehalkının ikamet ettiği konuttaki mülkiyet tipinin zaruri harcamalara etkileri ele alınmış, elde edilen sonuçlar Şekil 3.9 ve Şekil 3.10'da sunulmuştur. Söz konusu grafiklerde yer alan konut mülkiyet tipi değişkeninde 1 ev sahibi, 2 kiracı, 3 lojman ve 4 ise diğeri ifade etmektedir.



**Şekil 3.9.** Konut Mülkiyet Tipinin Zaruri Harcamalara Etkisi

Konut mülkiyet durumunun “Diğeri” kategorisi dışarıda bırakıldığında, hanehalkı kendi mülkiyetindeki bir evde yaşıyorsa zaruri harcamalar üzerinde bir etki yok iken, kirada oturma durumunda zaruri harcamalar üzerinde negatif bir etki oluşmakta, lojman

imkanlarından faydalananlar için de kıyasla sınırlı ancak negatif bir etki görülmektedir. “Diğer” kategorisi için etki 2008 örnekleminde yaklaşık 0 iken, 2010 örnekleminde negatiftir. Gelir esneklikleri ise kiracılar, ev sahibi, lojmanda ikamet eden sırasıyla azalmaktadır.



**Şekil 3.10.** Konut Mülkiyet Tipine Göre Gelir Esneklikleri

### 3.4. PARAMETRİK VE PARAMETRİK OLMAYAN TAHMİNLERİN KARŞILAŞTIRILMASI

Harcamaların gelir ve sosyo-demografik faktörlerle nasıl ilişkili olduğunun parametrik olmayan çekirdek kestirim tekniği ile izahından sonra, bu kısımda da söz konusu ilişkinin parametrik Benchmark modeli ile izahına yer verilmiş ve bu iki farklı yöntemin kullanılmasıyla elde edilen tahmin sonuçları karşılaştırılmıştır. Tablo 3.6 ve Tablo 3.7’de sırasıyla 2008 ve 2010 yıllarına ait parametrik model sonuçları sunulmuştur.

**Tablo 3.6.** Parametrik Benchmark Modelleri (2008)

Bağımlı Değişken: zh						
	1	2	3	4	5	6
<b>Sabit</b>	3,802 (0,000)	3,592 (0,000)	3,569 (0,000)	3,576 (0,000)	3,597 (0,000)	3,784 (0,000)
<b>gl</b>	0,453 (0,000)	0,460 (0,000)	0,458 (0,000)	0,449 (0,000)	0,498 (0,000)	0,474 (0,000)
<b>hh=1</b>	0,081 (0,155)					
<b>hh=2</b>	0,154 (0,006)					
<b>hh=3</b>	0,299 (0,000)					
<b>hh=4</b>	0,107 (0,064)					
<b>hh=5</b>	0,251 (0,000)					
<b>hh=6</b>	-0,154 (0,012)					
<b>hhb</b>		0,075 (0,000)				
<b>hhb*</b>			0,088 (0,000)			
<b>kk=2</b>					-0,105 (0,000)	
<b>mlk=1</b>						-0,005 (0,787)
<b>mlk=2</b>						-0,136 (0,000)
<b>mlk=3</b>						-0,093 (0,058)
<b>oecd</b>				0,187 (0,000)		
<b><math>\bar{R}^2</math></b>	0,329	0,343	0,342	0,345	0,299	0,301
<b>F</b>	599,6	2233,3	2222,1	2259,5	1827,3	923,5

Parantez içindeki değerler p-değerleri olup, White heteroskedastisite düzeltmesi kullanılmıştır. Hanehalkı tipi için HH=7, kır-kent için KK=1 ve mülkiyet tipi için MLK=4 referans kategorileri olarak alındığından, bu kategorilere ilişkin etki değerleri sabit terim tahmini üzerinden okunabilir.

2008 ve 2010 yıllarına ait parametrik Benchmark modeli sonuçları bir önceki kısımda yer alan parametrik olmayan model sonuçları ile karşılaştırıldığında, harcamaların gelir ve bir sosyo-demografik faktörle açıklandığı 6 modele ilişkin uyum iyiliği değerlerinin parametrik olmayan çekirdek kestirim tekniğinde daha yüksek

olduğu görülmektedir. Bu, söz konusu ilişkiyi açıklamak için oluşturulan parametrik olmayan modellerin daha uygun olduğu anlamına gelmektedir.

**Tablo 3.7.** Parametrik Benchmark Modelleri (2010)

Bağımlı Değişken: zh						
	1	2	3	4	5	6
<b>Sabit</b>	3,728 (0,000)	3,416 (0,000)	3,403 (0,000)	3,380 (0,000)	3,368 (0,000)	3,577 (0,000)
<b>gl</b>	0,462 (0,000)	0,482 (0,000)	0,478 (0,000)	0,472 (0,000)	0,527 (0,000)	0,489 (0,000)
<b>hh=1</b>	0,118 (0,000)					
<b>hh=2</b>	0,219 (0,000)					
<b>hh=3</b>	0,328 (0,000)					
<b>hh=4</b>	0,146 (0,000)					
<b>hh=5</b>	0,336 (0,000)					
<b>hh=6</b>	-0,179 (0,000)					
<b>hhb</b>		0,076 (0,000)				
<b>hhb*</b>			0,092 (0,000)			
<b>kk=2</b>					-0,143 (0,000)	
<b>mlk=1</b>						0,127 (0,000)
<b>mlk=2</b>						-0,059 (0,002)
<b>mlk=3</b>						0,041 (0,261)
<b>oecd</b>				0,193 (0,000)		
$\bar{R}^2$	0,317	0,316	0,316	0,320	0,277	0,281
<b>F</b>	670,4	2326,1	2332,1	2370,7	1928,9	988,8

Parantez içindeki değerler p-değerleri olup, White heteroskedastisite düzeltmesi kullanılmıştır. Hanehalkı tipi için HH=7, kır-kent için KK=1 ve mülkiyet tipi için MLK=4 referans kategorileri olarak alındığından, bu kategorilere ilişkin etki değerleri sabit terim tahmini üzerinden okunabilir.

Zaruri harcamaların açıklayıcısı olarak tüm modellere dahil edilen gelir değişkeninin etkisi, parametrik yöntem kullanıldığında tek bir katsayı ile ölçülebilmektedir. Oysa parametrik olmayan yöntem kullanıldığında, modele ilave edilen her bir değişkenin farklı kategorileri için gelir esneklikleri ayrı ayrı hesaplanmakta, böylelikle her bir kategorinin gelir esnekliklerini nasıl etkilediği görülebilmektedir. Ayrıca parametrik olmayan yöntem ile, gelirin farklı sıklık dilimleri için esneklik tahminlerinin izlediği trendi görebilmek de mümkündür. Bununla birlikte her iki yöntemden elde edilen sonuçlar, her bir değişkenin zaruri harcamalar üzerindeki etkisini farklı kategoriler bazında özetlemektedir.

Oluşturulan her bir model için parametrik ve parametrik olmayan yöntemlerden elde edilen sonuçlar karşılaştırıldığında, sonuçların büyük ölçüde paralellik gösterdiği

görülmektedir. Fakat parametrik olmayan yöntemin veri kümelerine ilişkin daha kapsamlı sonuçlar sunması, bu yöntemin tercih edilmesi için bir gerekçe olarak gösterilebilir.

## TARTIŞMA VE SONUÇ

Hanehalkı zaruri harcamalarının hanehalkı gelirinin bir fonksiyonu olduğu bilgisi iktisat teorisince ortaya konmakta ve önceki ampirik literatürce teyit edilmektedir. Yapılan parametrik olmayan tahminler de literatüre paralellik arz etmektedir. 2008 ve 2010 yıllarına ait Hanehalkı Bütçe Anketi verileri kullanılarak yapılan parametrik olmayan tahminlerin işaret ettiği dikkate değer ilk bulgu ise, zaruri harcamaların gelir esnekliğinin artan gelire düştüğü tespitidir. Buna göre, haneye giren fazladan %1'lik gelirin zaruri harcama olarak değerlendirilen kısmı hanenin gelir düzeyi yükseldikçe azalmaktadır. Çalışmada sunulan spesifikasyonların hemen hepsi bu tespiti içermektedir.

Ancak daha önce de belirtildiği gibi, harcamaların gelire nasıl bağlı olduğunun anlaşılması ne yeterlidir, ne de ampirik açıdan yenilik içermektedir. Farklı hanehalkı özelliklerinin harcamaları nasıl etkilediği incelenmeden harcama kalıplarına ilişkin bütüncül bir bilgi türetmek zordur. Buradan hareketle çalışmada zaruri harcamaların hanehalkı demografik özelliklerine nasıl bağlı olduğu araştırılmış olup, spesifik olarak hanehalkı büyüklüğü, hanehalkının tipi, ikamet edilen konutun mülkiyet tipi ve hanehalkının coğrafi yerleşimi (kır/kent) dikkate alınmıştır.

Bulguların kapsamlı bir tartışmasını yapmadan önce parametrik çekirdek tahmin sürecinde parametrik ekonometrideki sabit terimin var olmadığı ifade edilmelidir. Diğer bir deyişle parametrik olmayan teknik, “gelirden bağımsız harcama” olarak yorumlanabilecek tahminler üretmemektedir. Ancak kategori değişkenleri için elde edilen katsayı tahminleri düzleştirilmiş parametrik olmayan tahmin yüzeyine kıyasla harcamaların her bir kategoride nasıl farklılaştığını göstermektedir. Böylelikle gelir değişkeni için her gözlem noktasında elde edilen katsayıların uyarılmış harcamaların esnekliğini, kategori değişkenleri için elde edilen katsayıların ise otonom harcamalardaki farklılaşmayı ölçtüğü söylenebilir.

Çocuk sayısı ve çocukların anketin gerçekleştirildiği esnada hanede yaşayıp yaşamadığı üzerinden tanımlanan hanehalkı tipi değişkeninin zaruri harcamaları istatistiksel açıdan anlamlı biçimde etkilediği görülmekte, ancak söz konusu etkiler net biçimde sınıflandırılmamaktadır. Fazla sayıda nominal kategorinin varlığı bu zorluğun bir nedeni olarak görülebilir. Net bulgulara ulaşmak için hanehalkı büyüklüğünün zaruri

harcamalara etkisi incelenmiş olup hanehalkı büyüklüğü 1 veya 2 iken harcamaların daha yüksek olduğu görülmüştür. Daha önemlisi, hanehalkı büyüklüğü 3 olduğunda zaruri harcamalar düşmekte, birey sayısı 3'ten 7'ye yükselirken etki de yükselmektedir. Gelir esnekliği ise hanehalkı büyüklüğü 4-5 civarında iken tepe noktasına ulaşan bir parabolü andırmaktadır. Özetle, küçük hanelerde gelirden bağımsız olarak harcamalar yüksek, harcamaların gelir esnekliği düşüktür. Haneye ilk çocuğun katılımıyla birlikte, harcamaların otonom kısmı azalmakta ancak gelir esnekliği yükselmektedir. Veri kümemizde her bir hanehalkının kaç yıllık olduğuna dair bilgi bulunmamaktadır. Ancak hanedeki birey sayısının hanehalkının birlikte yaşama süresi arttıkça arttığı ve gelirin bireylerin yaşı arttıkça arttığı varsayımları bir arada yapıldığında, orta büyüklükteki haneler için bu bulgular daha rahat anlaşılabilir. Orta büyüklükteki hanelerde harcamaların gelirden bağımsız bileşeninin düşük ancak gelir esnekliğinin yüksek olması, yaşam döngüsünün orta aşamasındaki bu hanelerin zaruri dahi olsa harcamalarını ancak gelir artışı ile destekleyebildikleri ölçüde gerçekleştirdikleri biçiminde yorumlanabilir.

Küçük hanelerde harcamaların gelirden bağımsız bileşeninin yüksek olması, yaşam döngüsünün erken aşamasında olduğunu varsaydığımız bu hanelerde riskten kaçınmanın daha sınırlı olduğunu ima ediyor olabilir. Benzer biçimde, birey sayısı orta büyüklüğün (4-5) üzerine çıktığında harcamaların otonom bileşeninin artması genişleyen ailenin ihtiyaçlarının karmaşıklaşması ile açıklanabilir. Küçük ve büyük hanelerde gelir esnekliğinin düşük olması, bu hanelerin örneklem ortalamasının üzerinde otonom harcamada bulunmasına bağlanmaktadır. Zaruri harcamaları gelirlerinden bağımsız olarak yüksek olan bu hanelerin gelirlerindeki artışlar zaruri harcamaları ancak sınırlı ölçüde uyarmaktadır. TÜİK tanımlı hanehalkı büyüklüğü için özetlenen bu bulgular OECD tanımı kullanıldığında kısmen değişmekte, hanehalkı büyüklüğünün zaruri harcamalar üzerindeki etkisi büyüklük arttıkça devamlı düşüş göstermektedir.

Kentsel hanehalkının otonom zaruri harcamalarının kırsal hanehalkınıninkine göre daha sınırlı, harcamalarının gelir esnekliğinin ise daha yüksek olması, yukarıda da kullandığımız risk algısı kanalı ile açıklanabilir. Öngörümüz, kentsel hanehalkının tasarruflarını yüksek tutabilmek için otonom harcamalarını sınırladığı yönündedir.



Bununla birlikte kentsel hanehalklarında haneye giren fazladan %1'lik gelirin zaruri harcamaya dönüşme oranı kırsal hanehalkıninkine kıyasla daha yüksektir.

Kirada veya lojmanda oturan hanehalkının zaruri harcamalarının kendi mülkiyetindeki bir evde yaşayan hanehalkıninkine kıyasla daha düşük olduğu bulgusu oldukça akla uygun bir yapı arz etmektedir. Gelirin muhtemelen dikkate değer bir kısmının konut kirasına ayrılması ile zaruri harcamalara (gıda ve alkolsüz içecekler, sağlık ve eğitim) hanehalkı gelirinden ayrılan pay baskılanmaktadır. Kirada yaşayan hanehalkının fazladan gelirini zaruri harcamalara yönlendirme oranı daha yüksektir.

Buraya kadar özetlediğimiz bulguların genelinde, harcamaların gelirden bağımsız bileşenini baskılayan bir faktörün var olması durumunda (orta büyüklükte aile sahibi olmak, kentte yaşamak, kirada ikamet etmek) gelir esnekliğinin daha yüksek olduğu ifade edilmelidir. Zaruri harcamaların otonom ve uyarılmış bileşenleri arasında bir al-ver (trade-off) söz konusudur.

Yapılan çalışmada parametrik olmayan çekirdek kestirim tekniği ile hanehalkı zaruri harcamalarının gelire ve sosyo-demografik değişkenlere nasıl bağlı olduğu ortaya konulmuştur. Oluşturulan modellerden elde edilen sonuçlar parametrik model sonuçları ile karşılaştırıldığında, sonuçların büyük ölçüde paralellik gösterdiği görülmektedir. Fakat parametrik olmayan yöntemin veri kümelerine ilişkin daha kapsamlı sonuçlar sunması, bu yöntemin tercih edilmesi için bir gerekçe olarak gösterilebilir.

## KAYNAKÇA

- Akbay, C. (2005). "Kahramanmaraş'ta Hanehalklarının Gıda Tüketim Talebi Ekonometrik Analizi". *KSÜ Fen ve Mühendislik Dergisi*, 8 (1), 114-121.
- Akın, F. (Mayıs 1999). "Hanehalkı Gelir ve Tüketim Harcamaları Üzerine Quantile Regresyon" [Bildiri]. *IV. Ulusal Ekonometri ve İstatistik Sempozyumu*, Antalya.
- Allen, M. P. (1997). *Understanding Regression Analysis*. New York: Plenum Press.
- Alpay, S. ve Koç, A. (2002). "Household Demand in Turkey: An Application of Almost Ideal Demand System with Spatial Cost Index". *Economic Research Forum*. (Working Paper).
- Atalay, M. (1978). "Türkiye'de Bitkisel Yağ Talebi Üzerine Bir Deneme". *ODTÜ Gelişme Dergisi*, 19, 1-8.
- Barnes, R. ve Gillingham, R. (1984). "Demographic Effects in Demand Analysis: Estimation of the Quadratic Expenditure System Using Microdata". *The Review of Economics and Statistics*, 66 (4), 591-601.
- Baydemir, M. (1998). *Doğrusala Yakın İdeal Talep Sisteminin (LA/AIDS) Bir Uygulaması: Erzurum İli Verileri*. (Yüksek Lisans Tezi). Erzurum: Atatürk Üniversitesi Fen Bilimleri Enstitüsü.
- Benedetti, J. K. (1977). "On the Nonparametric Estimation of Regression Functions". *Journal of the Royal Statistical Society*, 39 (2), 248-253.
- Berry, W. D. (1993). *Understanding Regression Assumptions*. London: Sage Publications.
- Beyaz, F. B. (2007). *Türkiye'de Hanehalkı Tüketim Harcamaları ve Talep Tahmini*. (Yüksek Lisans Tezi). Antalya: Akdeniz Üniversitesi Sosyal Bilimler Enstitüsü.
- Bierens, H. J. ve Pott-Buter, H. A. (1990). "Specification of Household Engel Curves by Nonparametric Regression". *Econometric Reviews*, 9 (2), 123-184.
- Case, A. (2000). "Income Distribution and Expenditure Patterns in South Africa" [Bildiri]. Princeton University, USA.
- Choi, M. ve Lee, S. (2008). "Expenditure Patterns by Household Type and Income Class: The Korean Experience". *The Gerontologist*, 48, 699-700.
- Chu, C. K. ve Marron, J. S. (1991). "Choosing a Kernel Regression Estimator". *Statistical Science*, 6 (4), 404-419.
- Çatalbaş, K. G. (2006). *Parametrik Olmayan Tahmin Teknikleri: Türkiye Verileri ile Bir Uygulama*. (Doktora Tezi). Ankara: Gazi Üniversitesi Sosyal Bilimler Enstitüsü.
- Çolak, Ö. F., Öztürkler, H. ve Tokatlıoğlu, İ. (2008). "Türkiye'de Tüketim Fonksiyonunun Dilim Regresyon Yöntemi ile Tahmini". *İktisat İşletme ve Finans*, 23 (268), 62-93.
- De Meyrick, J. ve Yusuf, F. (2006). "The Application of Household Expenditure Data in the Development of Anti-Smoking Campaigns". *Health Education*, 106 (3), 227-237.

- Demir, S. (2005). *Regresyon Fonksiyonlarının Uyarlanabilir Nadaraya-Watson Çekirdek Kestirimleri*. (Doktora Tezi). Ankara: Hacettepe Üniversitesi Fen Bilimleri Enstitüsü.
- Djibuti, M., Gotsadze, G., Mataradze, G. ve Zoidze, A. (2007). "Influence of Household Demographic and Socio-Economic Factors on Household Expenditure on Tobacco in Six New Independent States". *BMC Public Health*, 7 (1).
- Driel, H. V., Nadall, V. ve Zeelenberg, K. (1997). "The Demand for Food in the United States and the Netherlands: A Systems Approach with the CBS Model". *Journal of Applied Econometrics*, 12, 509-532.
- Ekinci, S. (1996). *Türkiye’de Bazı Gıda Maddelerinin Talep Analizi*. (Yüksek Lisans Tezi). Adana: Çukurova Üniversitesi Fen Bilimleri Enstitüsü.
- Emeç, H. (2001). "Türkiye’de Bölgelerarası Tüketim Harcamaları Tobit Model Yaklaşımı". *DEÜ İİBF Dergisi*, 16 (2), 61-73.
- Emeç, H. ve Gülay, E. (2008). "Alkol Tüketimi ve Sosyo-Demografik Değişkenlerin Alkol Tüketimi Üzerine Etkisi". *Ekonomik Yaklaşım*, 19 (68), 115-134.
- Ersoy, A. F. (1993). *Düzenli Gelire Sahip Ailelerin Tüketim Harcamaları ve Bunu Etkileyen Faktörler*. (Doktora Tezi). İstanbul: İstanbul Üniversitesi Sosyal Bilimler Enstitüsü.
- Eubank, R. L. (1999). *Nonparametric Regression and Spline Smoothing*. New York: Marcel Dekker.
- Fan, J. (1993). "Local Linear Regression Smoothers and Their Minimax Efficiencies". *The Annals of Statistics*, 21 (1), 196-216.
- Fisunoğlu, H. M. ve Şengül, S. (2010). *Adana İli Seyhan ve Yüreğir İlçe Merkezlerinde Hanehalkı Gelir, Tasarruf Eğilimleri ve Yoksulluğun Ölçülmesi (Proje No:107K065)*. Ankara: TÜBİTAK.
- Georgantelis, S., Phillips, G. D. A. ve Zhang, W. (1987). "Estimating and Testing an Almost Ideal Demand System". *The Practice Econometrics*.
- Gozalo, P. L. (1997). "Nonparametric Bootstrap Analysis with Applications to Demographic Effects in Demand Functions". *Journal of Econometrics*, 81, 357-393.
- Gökçe, U. (2008). *Finansal Varlıkların Fiyatlamasında Parametrik Olmayan Regresyon Modelleri*. (Yüksek Lisans Tezi). İstanbul: Marmara Üniversitesi Sosyal Bilimler Enstitüsü.
- Hardle, W. (1990). *Applied Nonparametric Regression*. Cambridge: Cambridge University Press.
- Hardle, W. ve Linton, O. (1994). "Applied Nonparametric Methods". *Handbook of Econometrics*, 4 (38), 2295-2339.
- Hardle, W., Müller, M., Sperlich, S. ve Werwatz, A. (2004). *Nonparametric and Semiparametric Models*. Berlin: Springer.

- Hatırlı, S. A., Öztürk, E. ve Aktaş, A. Z. (2007). “Kırmızı, Tavuk ve Beyaz Et Talebinin Tam Talep Sistemi Yaklaşımıyla Analizi”. *SDÜ Sosyal Bilimler Enstitüsü Dergisi*, 2 (6), 211-221.
- Hayfield, T. ve Racine, J. S. (2007). *Package ‘np’: Nonparametric Kernel Smoothing Methods for Mixed Data Types*. R package version 0.40-4.
- He, B. (2009). *Smoothing Parameter Selection for a New Regression Estimator for Nonnegative Data*. (Yüksek Lisans Tezi). Canada: Concordia University.
- Hepsağ, A. (2007). *Parametrik Olmayan Regresyon ve Tahmin Yöntemleri: İMKB’de İşlem Gören Hisse Senetlerine Ait Piyasa Riskinin Tahmini Üzerine Bir Uygulama*. (Yüksek Lisans Tezi). İstanbul: İstanbul Üniversitesi Sosyal Bilimler Enstitüsü.
- Houthakker, H. S. (1965). “New Evidence on Demand Elasticities”. *Econometrica*, 33 (2), 277-288.
- Hurvich, C. M., Simonoff, J. S. ve Tsai, C. L. (1998). “Smoothing Parameter Selection in Nonparametric Regression Using a Improved Akaike Information Criterion”. *Journal of the Royal Statistical Society*, 60 (2), 271-293.
- Izan, H. Y. ve Clements, K. W. (1979). “A Cross-Cross-Section Analysis of Consumption Patterns”. *Economics Letters*, 4 (1), 83-86.
- Jensen, H. H. ve Manrique, J. (1998). “Demand for Food Commodities by Income Groups in Indonesia”. *Applied Economics*, 30, 491-501.
- Kasnakoğlu, Z. (1991). “Income and Expenditure Elasticities in Turkey for Selected Products in the Provinces of Ankara and Erzurum”. *The State Institute of Statistics*, EB. OB./91-4 (Working Paper).
- Kaytaş, M. (1987). “Koşullu Talep Denklemleri ve Bir Uygulama”. *İktisat Fakültesi Mecmuası*, 45 (1-4), 136-158.
- Kennedy, P. (2006). *Ekonometri Klavuzu*. (Çev. Muzaffer Sarımeşeli ve Şenay Açıkgöz). Ankara: Gazi Kitabevi. (2003).
- Kıroğlu, B. G. (1999). “Parametrik Olmayan Regresyonda Kestirim Yöntemleri: Kernel Kestirim Yöntemi” [Bildiri]. *İstatistik Araştırma Sempozyumu*, DİE, Ankara.
- Kim, Y. ve Yang, B. (2011). “Relationship Between Catastrophic Health Expenditures and Household Incomes and Expenditure Patterns in South Korea”. *Health Policy*, 100 (2-3), 239-246.
- Koç, A. A. ve Tan, S. (2001). “The Demand of Dairy Products in Turkey: The Impact of Household Composition on Consumption”. *METU Studies in Development*, 28, 169-181.
- Lee, H. J. ve Chern, W. S. (1989). “Nonlinear Regression for Estimating a Quadratic Expenditure System” [Bildiri]. The 1989 Proceedings of the Business and Economic Statistics Section, American Statistical Association.
- Leser, C. E. V. (1961). “Commodity Group Expenditure Functions for the United Kingdom, 1948-1957”. *Econometrica*, 29 (1), 24-32.
- Leser, C. E. V. (1963). “Forms of Engel Functions”. *Econometrica*, 31 (4), 694-703.

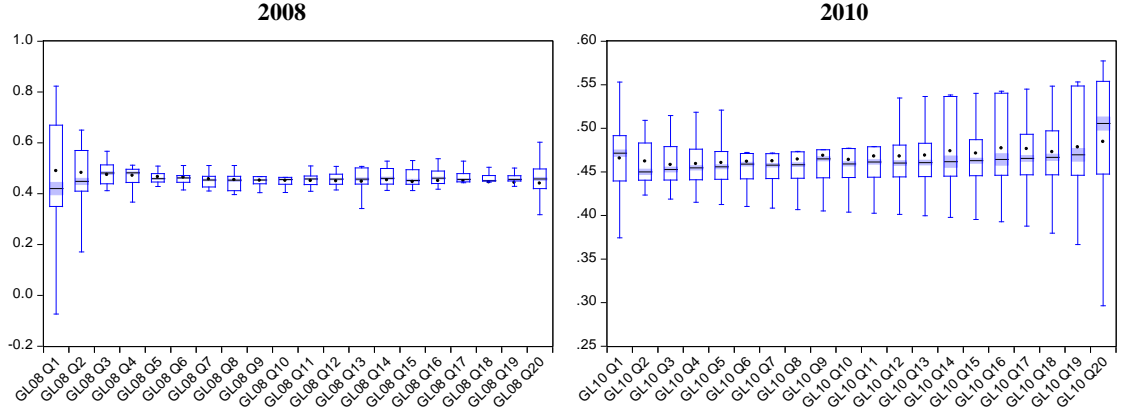
- Li, O. ve Racine, J. S. (2007). *Nonparametric Econometrics: Theory and Practice*. Princeton: Princeton University Press.
- Marron, J. S. (1988). "Automatic Smoothing Parameter Selection: A Survey". *Empirical Economics*, 13 (3/4), 187-208.
- Marron, J. S. ve Hardle, W. (1986). "Random Approximations to Some Measures of Accuracy in Nonparametric Curve Estimation". *Journal of Multivariate Analysis*, 20, 91-113.
- Mcelroy, M. B. (1975). "A Spliced CES Expenditure System". *International Economic Review*, 16 (3), 765-780.
- Musgrove, P. (1985). "Household Food Consumption in the Dominican Republic: Effects of Income, Price and Family Size". *Economic Development and Cultural Change*, 34 (1), 83-101.
- Nişancı, M. (1998). *Türkiye’de Tüketici Harcamalarının Analizi –İdeale Yakın Talep Sistemi Uygulaması–*. (Doktora Tezi). Erzurum: Atatürk Üniversitesi Sosyal Bilimler Enstitüsü.
- Öz, E. (2008). *Ulusal Hanehalkı Cepten Sağlık Harcamalarının Finansmanda Dikey Hakkaniyet Açısından Değerlendirilmesi*. (Doktora Tezi). Ankara: Hacettepe Üniversitesi Sağlık Bilimleri Enstitüsü.
- Özer, H. (1999). *Türkiye’de Hanehalkı Tüketim Harcamalarının Doğrusal Harcama Sistemi Yaklaşımıyla Analizi*. (Doktora Tezi). Erzurum: Atatürk Üniversitesi Sosyal Bilimler Enstitüsü.
- Parks, R. W. ve Barten, A. P. (1973). "A Cross-Country Comparison of the Effects of Prices, Income and Population Composition on Consumption Patterns". *The Economic Journal*, 83 (331), 834-852.
- Pollak, R. A. ve Wales, T. J. (1981). "Demographic Variables in Demand Analysis". *Econometrica*, 49 (6), 1533-1551.
- R Development Core Team (2012). *R: A Language and Environment for Statistical Computing*. Austria: R Foundation for Statistical Computing. Erişim tarihi: 11 Temmuz 2012. <http://www.R-project.org/>
- Sarimeşeli, M. (1999). "Hanehalkları Harcama Eğilimleri". *GÜ İİBF Dergisi*, 2, 41-50.
- Sawtelle, B. A. (1993). "Income Elasticities of Household Expenditures: A US Cross-Section Perspective". *Applied Economics*, 25, 635-644.
- Schimek, M. G. (2000). *Smoothing and Regression*. Canada: A Wiley-Interscience Publication.
- Schuster, E. F. (1972). "Joint Asymptotic Distribution of the Estimated Regression Function at a Finite Number of Distinct Points". *Annals of Mathematical Statistics*, 43 (1), 84-88.
- Schroeder, L. D., Sjoquist D. L. ve Stephan, P. E. (1986). *Understanding Regression Analysis (An Introductory Guide)*. London: Sage Publications.
- Selim, R. (Hazian 2000). "The Changes in the Consumption Expenditure Patterns in Turkey: 1987-1994" [Bildiri]. *First International Joint Symposium on Business*

*Administration: Challenges for Business Administrators in the New Millennium*, Çanakkale.

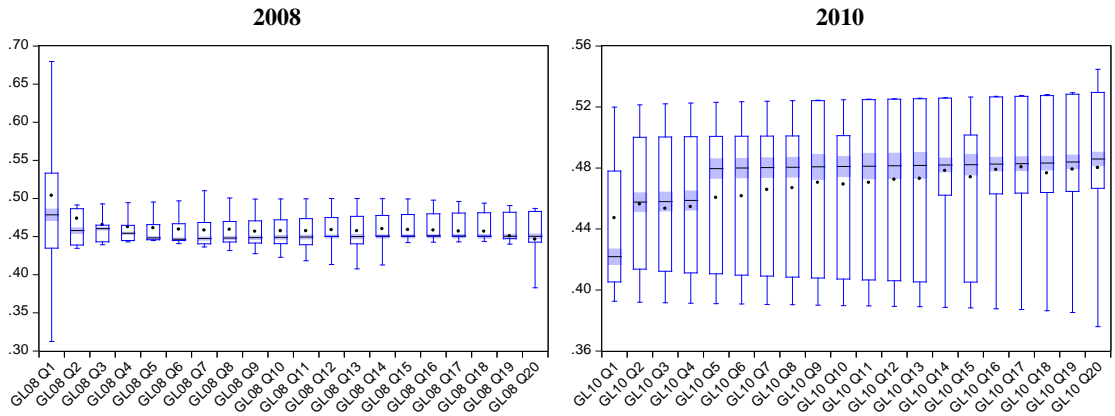
- Serper, Ö. (2004). *Uygulamalı İstatistik 2*. Bursa: Ezgi Kitabevi.
- Steinmetz, J. C. ve Gasser, T. (1988). "A Unifying Approach to Nonparametric Regression Estimation". *Journal of the American Statistical Association*, 83 (404), 1084-1089.
- Şahinli, M. A. (2010). "Yaklaşık İdeal Talep Analizi Yöntemi ile Harcama ve Fiyat Esnekliklerinin Tahmini". *Eskişehir Osmangazi Üniversitesi İİBF Dergisi*, 5 (2), 147-159.
- Şenesen, Ü. ve Selim, R. (1995). "Consumption Patterns of Turkish Urban and Rural Household in 1987". *METU Studies in Development*, 22 (2), 207-220.
- Tansel, A. ve Bircan, F. (2006). "Demand for Education in Turkey: A Tobit Analysis of Private Tutoring Expenditures". *Economics of Education Review*, 25, 303-313.
- Tezcan, N. (2009). "Parametrik Olmayan Regresyon Analizi" [Bildiri]. *10. Ekonometri ve İstatistik Sempozyumu*, Atatürk Üniversitesi, Erzurum.
- Toktamış, Ö. (1995). "Olasılık Yoğunluk Fonksiyonunun Çekirdek Kestirimi Üzerine Bir Çalışma". *Hacettepe Fen ve Mühendislik Bilimleri Dergisi*, 16, 145-163.
- TÜİK. (2008/2010). *Hanehalkı Bütçe Anketlerinin Uygulama Yöntemi, Tanım ve Kavramlar*. Ankara: TÜİK.
- Uraz, A. (2008). *Assessing Patterns of Household Expenditures on Recreation and Culture in Turkey in 2003*. (Yüksek Lisans Tezi). Ankara: ODTÜ Sosyal Bilimler Enstitüsü.
- Var, A. (2005). *Olasılık Yoğunluk Fonksiyonu İçin Uyarlanabilir Çekirdek Kestirim Yöntemi*. (Yüksek Lisans Tezi). Ankara: Hacettepe Üniversitesi Fen Bilimleri Enstitüsü.
- Wand, P. M. ve Jones, M. C. (1995). *Kernel Smoothing*. New York: Chapman and Hall.
- Yurdakul, O. (1980). "Adana'da Hayvansal Gıda Maddeleri Tüketimi ve Gelir-Harcama Esneklikleri". *Çukurova Üniversitesi Ziraat Fakültesi Yıllığı*, 11 (1-4).
- Zortuk, M. (2009). "Cinsiyet Değişkeni Bağlamında Harcama Alt Grupları ve Gelir İlişkisi: Dumlupınar Üniversitesi Öğrencileri Üzerine Bir Uygulama". *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*, 23.

## EKLER

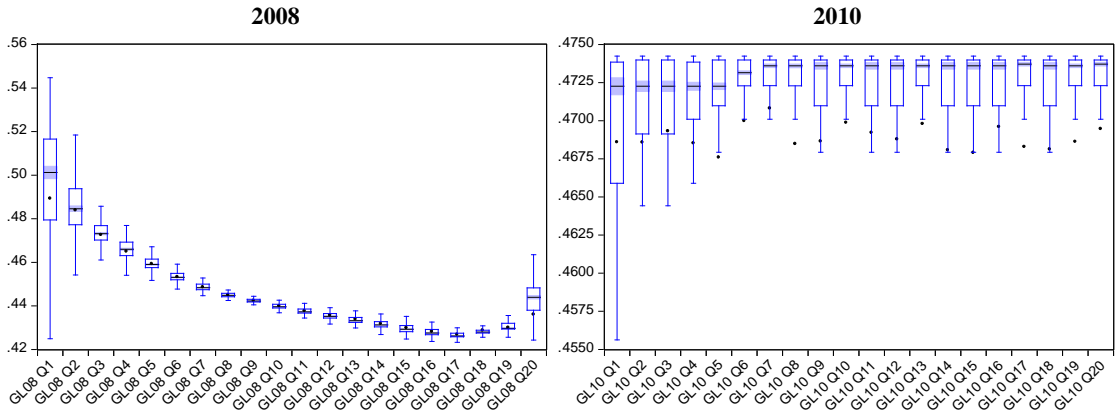
## Ek 1. Alternatif Modellere Göre Gelir Esneklikleri



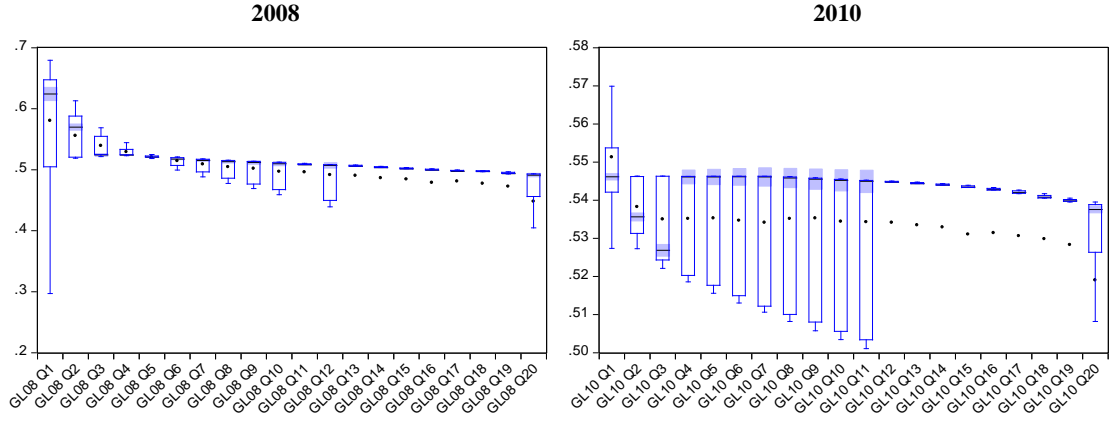
Şekil Ek 1. Hanehalkı Tipi Modelinde Gelir Dilimlerine Göre Gelir Esneklikleri



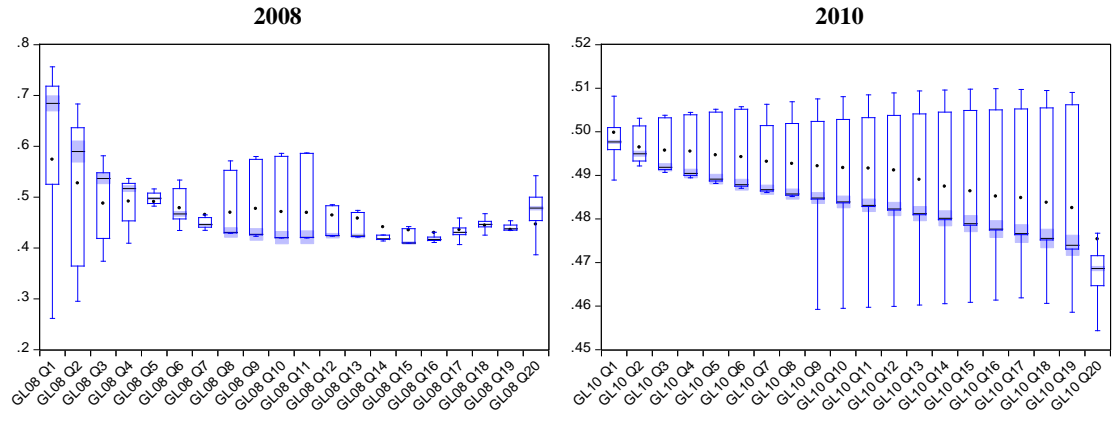
Şekil Ek 2. Hanehalkı Büyüklüğü (TÜİK Tanımı) Modelinde Gelir Dilimlerine Göre Gelir Esneklikleri



Şekil Ek 3. Hanehalkı Büyüklüğü (OECD Tanımı) Modelinde Gelir Dilimlerine Göre Gelir Esneklikleri



Şekil Ek 4. Kır-Kent Ayrımı Modelinde Gelir Dilimlerine Göre Gelir Esneklikleri



Şekil Ek 5. Konut Mülkiyet Tipi Modelinde Gelir Dilimlerine Göre Gelir Esneklikleri



**Ek 2. Tahminlerin Üretilmesinde Kullanılan Kodlar****Tablo Ek 1.** Tahminlerin Üretilmesinde Kullanılan Kodlar

---

```
setwd("C:/calc/makale")
rm(list=ls(all=TRUE))
hhdt1<-read.csv(file="hh_10.csv",header=TRUE)
attach(hhdt1)
library(np)

sink("results_factor(hh)_10_hh_ls_epanechnikov_ll.txt")

zh<-zh
x<-data.frame(gl,factor(hh))

bwhh<-
npregbw(xdat=x,ydat=zh,ckertype="epanechnikov",ckerorder=2,regtype="ll",bwmet
hod="cv.ls",bwscaling=FALSE)

reshh<-npreg(bws=bwhh,gradients=TRUE,residuals=TRUE)

npplot(xdat=x,ydat=zh,bws=bwhh,plot.errors.method=c("none"),view="rotate",theta=
310,phi=30,gradients=TRUE)

reshh

summary(reshh)

m1<-reshh$mean
m2<-reshh$merr
g1<-reshh$grad
g2<-reshh$gerr

cbind(m1,m2)
cbind(g1,g2)

sink()
```

---

### **Ek 3. Veri Kümesi**

#### **Hanehalkı Değişkeni 1: BULTEN (Bülten numarası)**

#### **Hanehalkı Değişkeni 2: HHTIPI (Hanehalkı tipi)**

**Aldığı Değerler:** (1) Tek çocuklu çekirdek aile, çocukların hepsi 18 yaşından küçük (2) Tek çocuklu çekirdek aile, en az bir çocuk 18 yaşından büyük (3) İki çocuklu çekirdek aile, çocukların hepsi 18 yaşından küçük (4) İki çocuklu çekirdek aile, en az bir çocuk 18 yaşından büyük (5) Üç ve daha fazla çocuklu çekirdek aile, çocukların hepsi 18 yaşından küçük (6) Üç ve daha fazla çocuklu çekirdek aile, en az bir çocuk 18 yaşından büyük (7) Çekirdek aile, çocuksuz (REFERANS) (8) Ataerkil veya geniş aile, çocuksuz (9) Ataerkil veya geniş aile, çocukların hepsi 18 yaşından küçük (10) Ataerkil veya geniş aile, en az bir çocuk 18 yaşından büyük (11) Tek yetişkinli aile, çocuksuz (12) Tek yetişkinli aile, çocukların hepsi 18 yaşından küçük (13) Tek yetişkinli aile, en az bir çocuk 18 yaşından büyük (14) Bir arada yaşayan öğrenci, işçi vb. kişiler (15) Bir arada yaşayan akrabalar

**Açıklama:** Çekirdek aile: Bir arada yaşayan anne, baba ve/veya çocuklardan oluşan ailedir. Ataerkil veya geniş aile: Anne, baba ve/veya çocuklar ile büyükanne, büyükbaba, hala, amca vb. Akrabalardan oluşan, en az iki neslin bir arada yaşadığı kalabalık ailelerdir. Tek yetişkinli aile: Tek başına yaşayan kişiler ile ebeveynlerinden birinin bulunmadığı (başka bir yere, çalışma amacıyla gidenler, boşanma veya ölüm gibi nedenlerden dolayı) ailelerdir. Bir arada yaşayan öğrenci, işçi vb. Kişiler: Aralarında kan bağı bulunmayan ancak bir arada yaşayan kişilerin oluşturduğu topluluktur. Bir arada yaşayan akrabalar: Aralarında kan bağı olan (aile birlikteliği olmaksızın) ve bir arada yaşayan kişilerin oluşturduğu topluluktur.

#### **Hanehalkı Değişkeni 4: MULKIYET (Konuta mülkiyet durumu)**

**Aldığı Değerler:** (1) Ev sahibi (2) Kiracı (3) Lojman (4) Diğer (REFERANS)

**Açıklama:** Ev sahibi: Hanehalkı reisinin ya da hanehalkı fertlerinden birinin sahip olduğu (mülkiyetinde olduğu) konutta oturan hanelerin konuta tasarruf şeklidir. Kiracı: Mülkiyeti kendilerine ait olmayan bir konutta aynı veya nakdi bir bedel ödeyerek oturan hanelerin konuta tasarruf şeklidir. Lojman: Mülkiyeti devlete veya çalıştığı işyerine ait olan bir konutta oturan hanelerin konuta tasarruf şeklidir. Diğer: Hanehalkının hiç bir bedel ödemedi ya da piyasa fiyatının altında çok az değerde bedel ödeyerek baba evi, akraba evi vb. şeklinde oturduğu konuta tasarruf şeklidir. Hanehalkı birden fazla konut biriminde ikamet ediyorsa, sadece hanehalkı reisinin ikamet ettiği konutun mülkiyet durumu bu soruda belirtilmiştir.

#### **Hanehalkı Değişkeni 133: Yıllık hanehalkı kullanılabilir geliri**

#### **Hanehalkı Değişkeni 134: EFB\_OECD (Yenilenmiş OECD'ye göre eşdeğer hane büyüklüğü)**

**Aldığı Değerler:** 0..99,9

**Açıklama:** Hanedeki ilk yetişkin için 1, 14 ve daha yukarı yaştaki fertler için 0.5, 14 yaşından küçük fertler için 0.3 değerleri dikkate alınarak hesaplanan hanehalkı büyüklüğüdür.

#### **Hanehalkı Değişkeni 136: HHB (Hanenin büyüklüğü)**

**Aldığı Değerler:** 1..25, Açıklama: Hanede yaşayan toplam fert sayısı alınmıştır.

**Hanehalkı Değişkeni 137: KIRKNTKD (Kır ve Kent yerleşim yeri kodu)**

**Aldığı Değerler:** (1) Kır (REFERANS), (2). Kent

**Açıklama:** Nüfusu 20001 ve daha fazla olan yerleşim yerleri “Kent”, 20000 ve daha az olan yerleşim yerleri ise “Kır” olarak belirlenmiştir.

**Tüketim Değişkeni 1: BULTEN (Bülten numarası)**

**Tüketim Değişkeni 3: HBS\_KOD (COICOP-HBS Madde Kodu (5 Basamak))**

**Aldığı Değerler:** COICOP kod listesindeki değerleri alır.

**Açıklama:** COICOP sınıflamasına göre 3 farklı kodlama mevcuttur. Bu kodlamalar, satın alma gücü, fiyat indeksi ve hanehalkı bütçe anketine göre yapılmaktadır. COICOP-HBS sınıflaması ise hanehalkı bütçe anketine göre 5 basamaklı olarak kodlanmıştır.

**Tüketim Değişkeni 4: DEGERD (Harcama değeri)**

**Aldığı Değerler:** 0-999999999999

**Çalışmada Dikkate Alınan Maddeler:** 1111 Pirinç, 1112 Ekmek, 1113 Makarna, 1114 Hamur Ürünleri, 1115 Diğer tahıl ürünleri, 1121 Büyükbaş (sığır) hayvanların taze, soğutulmuş ya da dondurulmuş eti, 1122 Taze, soğutulmuş ya da dondurulmuş domuz eti, 1123 Taze, soğutulmuş ya da dondurulmuş koyun ve keçi eti, 1124 Taze, soğutulmuş ya da dondurulmuş tavuk eti (kümes hayvanları), 1125 Kurutulmuş, tuzlanmış ya da tütsülenmiş et ve yenilebilir sakatat, 1126 Diğer konserve ya da işlenmiş et ve et preparatları, 1127 Diğer taze, soğutulmuş ya da dondurulmuş yenilebilir et, 1131 Taze, soğutulmuş ya da dondurulmuş balık, 1132 Taze, soğutulmuş ya da dondurulmuş deniz ürünleri, 1133 Kurutulmuş, tütsülenmiş ya da tuzlanmış balık ve deniz ürünleri, 1134 Konserve ya da işlenmiş balık ve deniz ürünleri; balık ve deniz ürünleri preparatları, 1141 Süt, 1142 Az yağlı süt, 1143 Konserve (şişe) süt, 1144 Yoğurt, 1145 Peynir ve Lor, 1146 Diğer süt ürünleri, 1147 Yumurtalar, 1151 Tereyağı, 1152 Margarin ve diğer bitkisel yağlar, 1153 Zeytin Yağı, 1154 Yenilebilir Yağlar, 1155 Diğer yenilebilir hayvansal yağlar, 1161 Turunçgiller (taze, soğutulmuş ya da dondurulmuş), 1162 Muzlar (taze, soğutulmuş ya da dondurulmuş), 1163 Elmalar (taze, soğutulmuş ya da dondurulmuş), 1164 Armutlar (taze, soğutulmuş ya da dondurulmuş), 1165 Çekirdekli meyveler (taze, soğutulmuş ya da dondurulmuş), 1166 Kabuksuz yumuşak meyveler (etli ve zarlı) (taze, soğutulmuş ya da dondurulmuş), 1167 Diğer taze, soğutulmuş ya da dondurulmuş meyveler, 1168 Kuru meyveler, 1169 Konserve meyveler ve meyve özlü ürünler, 1171 Yapraklı ve Gövdeli sebzeler (taze, soğutulmuş ya da dondurulmuş), 1172 Lahanalar (taze ya da soğutulmuş), 1173 Meyveler için yetiştirilen sebzeler (taze, soğutulmuş ya da dondurulmuş), 1174 Köklü bitkiler, nişastasız (bitki) soğanlar ve mantarlar (taze, soğutulmuş ya da dondurulmuş), 1175 Kurutulmuş Sebzeler, 1176 Diğer konserve ya da işlenmiş sebzeler, 1177 Patatesler, 1178 Diğer yumru köklü ve yumrulu sebzelerden yapılan ürünler, 1181 Şeker, 1182 Reçeller, marmelatlar, 1183 Çikolata, 1184 Şekerleme Ürünleri, 1185 Yenilebilir buz ve dondurma, 1186 Diğer şekerli ürünler, 1191 Soslar, çeşniler, 1192 Tuz, baharatlar, 1193 Ekmek mayası, tatlı preparatları, çorbalar, 1194 Diğer Gıda Ürünleri, 1211 Kahve, 1212 Çay, 1213 Kakao ve çikolata tozu, 1221 Maden ya da kaynak suları, 1222 Alkolsüz içecekler, 1223 Meyve suları, 1224 Sebze suları, 6111 Eczacılıkla (ilaçla) ilgili ürünler, 6121 Diğer tıbbi ürünler, 6131 Tedavide kullanılan alet ve ekipmanlar, 6211 Tıbbi

hizmetler (doktor), 6221 Diřçilik hizmetleri, 6231 Medikal analiz laboratuvarlarının ve röntgen merkezlerinin hizmetleri, 6232 Tıbbi yardımcılarının hizmetleri, 6233 Hastaneyle ilgili olmayan diđer hizmetler, 6311 Hastane hizmetleri, 10111 Okul öncesi eğitim ve ilköğretim, 10211 Orta Öğretim, 10311 Orta öğretim sonrası ve üniversite öncesi eğitim, 10411 Yüksek Eğitim, 10511 Seviyesi belirlenemeyen eğitim

**ÖZGEÇMİŞ****Kişisel Bilgiler:**

Adı Soyadı Elif Gölveren

Doğum Yeri ve Tarihi Aşkale 1987

**Eğitim Durumu:**

Lisans Öğrenimi Atatürk Üniversitesi İİBF İşletme Bölümü

Yüksek Lisans Öğrenimi Atatürk Üniversitesi Sosyal Bilimler Enstitüsü İşletme  
Bilim Dalı Sayısal Yöntemler Anabilim Dalı

Bildiği Yabancı Diller İngilizce

**İş Deneyimi:**

Çalıştığı Kurum Atatürk Üniversitesi İİBF Ekonometri Bölümü  
İstatistik Anabilim Dalı, Arş. Gör.

**İletişim:**

E-posta Adresi elifgolveren@gmail.com

**Tarih:** 11 Temmuz 2012