JOHANSEN VE KISMİ LİNEER TEKİL DEĞER EŞBÜTÜNLEŞME TESTLERİNİN BİR UYGULAMA PROBLEMİ ÜZERİNDE KARŞILAŞTIRILMASI

Naciye PINARÖNÜ

YÜKSEK LİSANS TEZİ
İSTATİSTİK

GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

MAYIS 2007
ANKARA

Naciye PINARÖNÜ tarafından hazırlanan JOHANSEN VE KISMİ LİNEER TEKİL DEĞER EŞBÜTÜNLEŞME TESTLERİNİN BİR UYGULAMA PROBLEMİ ÜZERİNDE KARŞILAŞTIRILMASI adlı bu tezin Yüksek Lisans tezi olarak uygun olduğunu onaylarım.

Prof. Dr. Müslim EKNİ
Tez Yöneticisi

Bu çalışma, jürimiz tarafından oy birliği ile İstatistik Anabilim Dalında Yüksek Lisans tezi olarak kabul edilmiştir.

Başkan: : Prof. Dr. Hamza GAMGAM

Üye : Prof. Dr. Semra ORAL ERBAŞ

Üye : Doç. Dr. Yılmaz AKDİ

Üye : Yrd. Doç. Dr. Emel BAŞAR

Üye : Prof. Dr. Müslim EKNİ (Danışman)

Tarih : 25/05/2007

Bu tez, Gazi Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına uygundur.

TEZ BİLDİRİMİ

Tez içindeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edilerek sunulduğunu, ayrıca tez yazım kurallarına uygun olarak hazırlanan bu çalışmada orijinal olmayan her türlü kaynağa eksiksiz atıf yapıldığını bildiririm.

Naciye PINARÖNÜ

JOHANSEN VE KISMİ LİNEER TEKİL DEĞER EŞBÜTÜNLEŞME TESTLERİNİN BİR UYGULAMA PROBLEMİ ÜZERİNDE KARŞILAŞTIRILMASI

(Yüksek Lisans Tezi)

Naciye PINARÖNÜ

GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
Mayıs 2007

ÖZET

Bir zaman dizisi değişkeninin, başka bir ya da birden çok zaman dizisine göre regresyonu çoğu zaman yanıltıcı olabilir. Buna karşı korunmanın bir yolu değişkenler arasında eşbütünleşme (kointegrasyon) ilişkisine bakmaktır. Bu çalışmada, durağan olmayan zaman dizilerinden oluşan modellere ilişkin değişkenler arasındaki eşbütünleşme ilişkisi araştırılmıştır. Çok değişkenli eşbütünleşme testlerinden Johansen Testi ile Johansen yaklaşımından hareketle geliştirilmiş Kısmi Lineer Tekil Değer Testi tanıtılmış ve sayısal örnek üzerinde bir uygulama çalışması yapılmıştır. Amaç testlerde öne sürülen farklılıkları ortaya koymak ve bu yöntemleri ayrıntıları ile ele almaktır.

Bilim Kodu : 205 1.066
Anahtar Kelimeler : Durağanlık, Eşbütünleşme (Kointegrasyon), Johansen Testi, Kısmi Lineer Tekil Değer Testi
Sayfa Adedi : 83
Tez Yöneticisi : Prof. Dr. Müslim EKNİ

COMPARING THE COINTEGRATION TESTS OF SINGULAR PARTIAL LINEAR VALUE TEST WITH JOHANSEN TEST IN A PRACTICAL APPLICATIAN

(M.Sc. Thesis)

Naciye PINARÖNÜ

GAZİ UNIVERSITY
THE INSTITUTE OF NATURAL SCIENCES
May 2007

ABSTRACT

Most of the time, regression of a time series variable may be spurious when compared to that of one or more time series variable. To avoid this, it is recommended to examine the cointegration relationship between the variables. This study investigates the cointegration relationship between the variables in the models of non-stable time series and presents Johansen Test, being one of the multivariate cointegration tests, Singular Partial Linear Value Test, developed with reference to Johansen approach. Then, a sample numerical application is carried out. The aim is to establish the differences put forward in the tests and elaborate the methods.

Science Code : 205 1.066
Key Words : Stability, Cointegration, Johansen Test, Singular Partial Linear Value Test
Page Number: 83
Adviser : Prof. Dr. Müslim EKNİ

TEŞEKKÜR

Bu tezde çalışmalarım boyunca bilimsel tecrübeleri, değerli yardım ve katkılarıyla beni yönlendiren ve bundan sonraki çalışmalarımda bana destek olacağına inandığım Sayın Prof. Dr. Müslim EKNİ ’e teşekkürlerimi sunarım.

Ayrıca bu çalışmalarım süresince kıymetli tecrübelerinden yararlandığım, değerli fikirlerinden yardım aldığım Sayın Doç. Dr. M. Akif BAKIR ’a şükranlarımı sunarım.

Tüm yaşamım boyunca bana maddi ve manevi olanaklarının tamamını ellerinden geldiğince sunan ve başarılarımın altında her zaman imzalarının olduğuna inandığım aileme de şükranlarımı sunuyorum.

İÇİNDEKİLER

Sayfa

Sayfa

Sayfa

ÇİZELGELERİN LİSTESİ

Çizelge **Sayfa**

.

ŞEKİLLERİN LİSTESİ

Şekil **Sayfa**

SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış bazı simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

Simgeler	**Açıklama**
bütçe	Eğitim bütçesinin gsmh içinde oranı
d	Durbin Watson
$\mathbf{D_t}$	Mevsimsellik
h	Bant genişliği
ilk	İlköğretim okullaşma oranı
kbgsmh	Kişi başına gayri safi milli hasıla
orta	Ortaöğretim okullaşma oranı
$\mathbf{M_{ij}}$	Çarpım moment matrisi
P	İzdüşüm matrisi
r	Rank
$\mathbf{R^2}$	Belirleme katsayısı
t	t istatistiği
V	Özvektörler matrisi
yüksek	Yükseköğretim okullaşma oranı
$\mathbf{z_t}$	Zaman dizisi

Kısaltmalar	**Açıklama**
AR(1)	Bir Gecikmeli Otoregresyon Model
DF	Dickey-Fuller
DİE	Devlet İstatistik Enstitüsü
EG	Engle-Granger
EKK	En Küçük Kareler

Kısaltmalar	Açıklama
GEG	Genelleştirilmiş Engle-Granger
ERDW	Eşbütünleşme Regresyonu Durbin Watson
GDF	Genelleştirilmiş Dickey Fuller
GKVAR	Genelleştirilmiş Vektör Otoregresyon Testi
GSMH	Gayri Safi Milli Hasıla
HDM	Hata Düzeltme Modeli
K	Kernel
KBGSMH	Kişi Başına Gayri Safi Milli Hasıla
KHG	Kişisel Harcanabilir Gelir
KTH	Kişisel Tüketim Harcaması
KVAR	Kısıtlı Vektör Otoregresyon Testi
LR	Johansen En Çok Olabilirlik Testi
PLS	Kısmi Lineer Tekil Değer Testi
TÜİK	Türkiye İstatistik Kurumu
VAR	Vektör Otoregresyon Model

1. GİRİŞ

Çok değişkenli durağan olmayan zaman dizilerinin sözkonusu olduğu durumlarda regresyon parametrelerinin tahminleri uzun dönem denge modeli tahminlerini vermezler ve durağan olmama koşulu altında yanıltıcı regresyon sorunuyla karşılaşılır. Bu problemin üstesinden gelebilmek için zaman dizisi verilerinin durağan biçime getirilmesi gerekir. Ancak bu durum bilgi kaybına neden olacağından, eğer değişkenler I(1) dereceden durağansa bu durumda, eşbütünleşmenin (cointegration) varlığı araştırılır. I(0) dereceden durağansa, süreç durağandır ve eşbütünleşmenin varlığı halinde regresyon denklemlerine dayalı olarak hem kısa hem de uzun dönem model tahminlerini elde edebilmek mümkün olacaktır.

Eşbütünleşme testlerinden en yaygın olarak kullanılan testler Engle-Granger testi ve Johansen testidir. Engle-Granger testi eşbütünleşme vektörünün tahmini için EKK yöntemini kullanırken, Johansen testi ise; en çok olabilirlik yöntemini kullanır. Ancak parametrik modeller için geliştirilen bu testlere karşın bazı durumlarda modeldeki bir kısım değişkenler parametrik biçimde modellenemeyebilir. Bu çalışmada kısmi lineer model durumunda eşbütünleşme testi irdelenecek ve bu anlamda kısmi lineer tekil değer ayrışımına dayalı PLS olarak adlandırılan test ile Johansen testi, Türkiye için iktisadi büyüme ve eğitim ilişkisi ortaya koyan modelin değişkenleri üzerine uygulanacaktır.

Zaman dizileri yalnız kendi geçmiş değerleriyle ilişkili olarak değil aynı zamanda başka değişkenlerle de ilişkili olarak karşımıza çıkabilir. Değişkenlerin başka değişkenlere bağlı olarak da gözlenebilmesi o değişkene ilişkin birden fazla gözlem değeri elde etmemizi sağlar. Bu ilişkiye bir regresyon ilişkisi olarak bakmak hatalı olur. Nedeni ise; regresyonda bağımsız değişken değerlerinin sabit olma koşuludur. Burada değişkenler hem kendi geçmiş değerlerine ilişkin hem de başka değişkenlerle ilişkili tahminleri yapılabilmesi bu varsayımı sağlamamaktadır [1]. Bu nedenledir ki çok değişkenli zaman dizilerinde değişkenler arasındaki ilişkiyi ortaya koyabilecek ve bu ilişkiyi ifade eden regresyon denklemlerinin yanıltıcı olup olmadığını kanıtlayacak analiz yöntemi eşbütünleşme yöntemidir. Bu yöntem, aynı dereceden durağan olmayan değişkenlerin oluşturduğu lineer ilişkinin durağan olması temeline dayanır.

Hem çok değişkenli zaman dizilerinde hem de tek değişkenli zaman dizilerinde durağanlık önem arzeder [1].

Çalışmada, ilk olarak zaman dizilerinde durağanlık kavramının anlamı ve çözümlemelerde önemi üzerinde durularak, durağanlık testlerinden birim kök testi açıklanacaktır.

Tek değişkenli durağanlık testlerine değinildikten sonra asıl problemimiz olan çok değişkenli eşbütünleşme testleri açıklanacaktır. Çok değişkenli eşbütünleşme testlerinden Johansen Testi ve Johansen testinden hareketle yarı-parametrik modeller için Kısmi Lineer Tekil Değer Testinin (PLS) teorik alt yapısı açıklanacaktır.

PLS testinin kullanımında kısmi lineer model tahminleri gerekli olduğundan, bu türden modellere ilişkin kernel(çekirdek) tahminleri ve ilişkili kavramlar ele alınacaktır.

Son olarak Türkiye'de iktisadi büyüme ve eğitim arasındaki ilişkiyi tanımlayan model üzerinde yukarıda vurguladığımız çok değişkenli eşbütünleşme testleri uygulanarak, bu testlere ilişkin farklar ortaya konmaktadır.

Uygulamada birim kök testleri ve Johansen Eşbütünleşme testi Eviews 5.0 paket programı ile hesaplanmıştır. Kernel tahminlerine ilişkin hesaplamalarda Stata 10.0 kullanılmış Kısmi Lineer Tekil Değer Testi'ne ilişkin hesaplamalar ise; Matlab 7.04'te programlanmıştır.

2. ZAMAN DİZİLERİ ANALİZİNDE TEMEL KAVRAMLAR

İnsanoğlu, ilk çağlardan bu yana her zaman geleceğe karşı ilgi duymuştur. Medeniyetler ilerledikçe yaşamın karmaşası artmış ve artık gelecek hakkında belli bilgilere ulaşmak ihtiyaç haline gelmiştir. Dolayısıyla günümüzde kamu ve özel kuruluşlar, işletmeler ya da şirketler ve hatta şahıslar geleceğe yönelik araştırmalar ve tahminler yaparak bugünden önlemlerini alabilmektedirler [2].

Gelecek olaylar ya da koşulları tahmin etmeye *kestirim (öngörü)* denir. Kestirim, karar verme sürecinde vazgeçilmez bir unsurdur. Hükümet politikalarının oluşmasında işsizlik oranı, enflasyon oranı, vergi oranı, kişi başına düşen milli gelir gibi önemli ülke özelliklerini ortaya çıkarabilecek faktörlerle ilgili kestirimlerde bulunulması gerekmektedir. Böylece hükümetler sorunlar çıkabilecek alanlarda çözümler üretecek politikalar geliştirirler. Şirketlerde ise; daha çok pazarlama, finans ile ilgili bölümlerde, üretim planlama bölümlerinde ve süreç kontrol sektöründe kestirim işlemlerinden yararlanırlar.

Kestirim işlemi ile geçmişteki bilgilerden yararlanarak geleceğe ait tahminler yapılmaktadır. Zaman dizileri analizi de nicel yöntemlere dayanılarak elde edilen bilgilerin kullanılmasıyla gerçekleştirilir. Dolayısıyla zaman dizileri analizi, zaman içinde düzenli aralıklarla gözlenen verilerin istatistiksel olarak incelenmesini ve gelecek dönemlerde elde edilecek verilerin kestiriminin güvenilir bir şekilde sağlanmasını içermektedir.

3. ZAMAN DİZİSİ

Kronolojik sırayla elde edilen verilere sahip değişkenlere *zaman dizisi değişkeni* adı verilir. Zaman dizisi, n örnek hacmi olmak üzere, Z_t, $t = 1,2,...,n$ biçiminde gösterilir. Buna göre ilk gözlenen veri z_1, ikinci gözlenen veri z_2,… ve son gözlenen veri z_n ile ifade edilir. Olasılık teorisinde zaman dizisi $\{z_{t1}, z_{t2},..., z_{tn}\}$ sonlu bir rasgele değişkenler kümesi olup $\{z(w,t): t = 0, \pm1, \pm2,...\}$ stokastik sürecinden üretilmektedir. $z(w,t)$ stokastik süreci bir örnek uzayda tanımlı zamana ait rasgele değişkenleri içeren bir küme olmaktadır. Bu tanımda w değişkeni bir dönemde birden çok gözlem elde edildiğinde kullanılır.

Zaman içinde sürekli olarak kaydedilen verilere sahip dizilere *sürekli zaman dizileri* denir. Örneğin, elektrik sinyalleri, voltaj, ses titreşimleri…vb sürekli zaman dizileri tanımına uyar. Sadece belli aralıklarla elde edilen verilere sahip dizilere de *kesikli zaman dizileri* adı verilir. Faiz oranı, satış hacmi, üretim miktarı gibi diziler de kesikli zaman dizileridir [2].

Zaman dizileri tek değişkenli analize dayalı olabileceği gibi, birden fazla zaman dizisi birlikte düşünülerek modelleme ve analiz yapılması istenebilir. Literatürde vektörel ya da çok değişkenli zaman dizileri analizi olarak ifade edilen yöntemler ekonomi, işletme, mühendislik (meteoroloji) ve benzeri alanlarda kullanılmaktadır.

Çok değişkenli zaman dizileri analizi çalışmaları yalnız bireysel dizileri değil diziler arasındaki olası çapraz ilişkilerin belirlenmesi için de gereklidir. Dizileri modelleme ve analiz etmenin bir diğer ortak amacı diziler arasında zamana göre ilişkileri anlamak ve her bir dizinin kestirimlerinde ilişkili dizilerden mevcut bilgileri kullanarak tekli diziler için bulunan kestirimleri daha da iyileştirmeye çalışmaktır. Şimdi çok değişkenli zaman dizilerine bir kısa örnek verelim.

Çok değişkenli zaman dizilerine ilişkin herhangi bir VAR(1) modeli aşağıdaki gibi yazılmış olsun.

$$\begin{bmatrix} Y_t \\ Z_t \end{bmatrix} = \begin{bmatrix} 0{,}8 & 0{,}3 \\ 0{,}1 & 0{,}6 \end{bmatrix} \begin{bmatrix} Y_{t-1} \\ Z_{t-1} \end{bmatrix} + \begin{bmatrix} e_{Yt} \\ e_{Zt} \end{bmatrix}$$

Burada Y_t ve Z_t çok değişkenli zaman dizisini oluştururken katsayıların oluşturduğu matris ise; parametre matrisidir. Bu modelin durağanlık yapısını incelersek; A parametre matrisini ve λ özdeğer matrisini temsil etmek üzere, $|\lambda I - A| = 0$ kökleri birim çemberin içinde ise; sözkonusu model durağanlık koşulunu sağlayacaktır. Özdeğerleri bulmak için,

$$\left| \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0{,}8 & 0{,}3 \\ 0{,}1 & 0{,}6 \end{bmatrix} \right| = 0$$

denklem sisteminin çözülmesi gerekir. Buradan,

$$\left| \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} - \begin{bmatrix} 0{,}8 & 0{,}3 \\ 0{,}1 & 0{,}6 \end{bmatrix} \right| = 0$$

ve

$$\begin{vmatrix} \lambda - 0{,}8 & -0{,}3 \\ -0{,}1 & \lambda - 0{,}6 \end{vmatrix} = 0$$

şeklinde bulunur. Bu determinant alındığında,

$$\left[(\lambda - 0{,}8)(\lambda - 0{,}6) - (-0{,}1)(-0{,}3) \right] = 0$$

eşitliğinden $\lambda_1 = 0{,}9$ ve $\lambda_2 = 0{,}5$ elde edilir ve bu özdeğerler mutlak değer olarak 1'den küçük olduğu için, sözkonusu model durağanlık koşulunu sağlar.

4. ARDIŞIK BAĞIMLILIK

Ardışık bağımlılık terimi, zaman dizisi verilerinde zaman içinde sıralanan gözlem dizilerinin birimleri arasındaki ilişkidir. Klasik lineer regresyon modeli böyle bir ardışık bağımlılığın, u_t hata terimleri arasında olmadığını varsayar. Bu aşağıdaki gibi ifade edilir.

$$Kov(u_i, u_j) = 0 \qquad i \neq j$$

Başka bir deyişle, klasik bir modelde, herhangi bir gözleme ilişkin hata terimin, başka bir gözleme ilişkin hata teriminden etkilenmediğini ifade eder. Örneğin, üretimin, işgücü ve sermaye girdilerine göre regresyonuna ilişkin üçer aylık zaman dizisi verileriyle çalışılıyor ise; bir dönemde gerçekleşen grevin, üretimde neden olduğu kesintinin etkisinin bir sonraki üç aylık dönemde devam etmeyeceğini ifade eder [3].

Şayet bir bağımlılık varsa, bu ardışık bağımlılığı ifade eder ve şöyle simgelenir:

$$Kov(u_i, u_j) \neq 0 \qquad i \neq j$$

Bu durumda bir dönemde gerçekleşen grevin bu üç aylık dönemde neden olduğu kesintinin bir sonraki üç aylık dönemde de etkisinin devam edeceğini ifade eder.

5. DURAĞANLIK VE EŞBÜTÜNLEŞME (KOİNTEGRASYON)

Zaman dizilerine dayalı regresyon çalışmalarında değişkenlerin durağanlığı önemlidir.

Bir zaman dizisinin değişkeninin başka bir zaman dizisi değişkeni/değişkenlerine göre regresyonu hesaplanırken ikisi arasında anlamlı bir ilişki olmasa bile çoğunlukla yüksek bir belirleme katsayısı (R^2) bulunabilir. Bu durum *yanıltıcı regresyon problemi* olarak adlandırılır. Bu sorunun ortaya çıkış nedeni her iki zaman dizisinin de güçlü *genel eğilimler* (yukarı ya da aşağı doğru kalıcı hareketler-trend) taşımasıdır. Yüksek R^2 gerçek bir ilişkiden çok bu genel eğilimden kaynaklanabilir. Yani; yüksek bir R^2, regresyon kestirim denkleminin değerlendirilmesinde her zaman güvenilir bir sonuç olmayacaktır.

Dolayısıyla iki zaman dizisi değişkeninin arasındaki ilişkinin gerçek mi, yanıltıcı mı olduğu bu noktada büyük önem taşır.

Zaman dizileri içeren regresyon modelleri genellikle kestirim için kullanılır. Yukarıda ifade edildiği gibi eğer ilişkilendirilen zaman dizileri durağan değilse böyle kestirimlerin geçerli olup olmadığı bilinmek istenir.

5.1. Durağan Olasılık Süreçleri

Her zaman dizisinin bir olasılık ya da rassal süreç ile türediği düşünülür. Zaman dizisi analizcilerinin büyük ilgi gösterdiği ve incelenen bir olasılıklı süreç türü *durağan olasılıklı süreç*tir.

Genel olarak ifade edilecek olunursa, ortalaması ile varyansı zaman içinde değişmeyen ve iki dönem arasındaki kovaryansı, bu kovaryansın hesaplandığı döneme değil de yalnızca iki dönem arasındaki uzaklığa bağlı olan olasılıklı bir süreç için *durağandır (zayıf durağandır)* denir.

Y_t : olasılıklı bir zaman dizisini göstermek üzere,

$Ortalama: \ E(Y_t) = \mu$

$Varyans: \ Var(Y_t) = \sigma^2$

$Ko\,\mathrm{var}\,yans: \ \gamma_k = Kov(Y_t, Y_{t+k})$

olarak tanımlı olsun. Burada,

γ_k, k gecikme ile kovaryansı temsil etmektedir. Daha açık bir ifade ile aralarında k dönem fark olan Y_t ile Y_{t+k} arasındaki kovaryanstır. Aynı zamanda,

$k = 0 \ \ için \ \ \gamma_0 = Var(Y_t) = \sigma^2$

$k = 1$ *için* $\gamma_1 = Y$ nin ardışık iki değeri arasındaki kovaryanstır.

Y_t'nin sıfır noktası, Y_t'den Y_{t+m}'ye kaydırıldığı varsayılsın. Eğer Y_t durağansa Y_{t+m}'nin ortalaması, varyansı ve kovaryansı Y_t'ninkilerle aynı olmalıdır.

Kısaca; eğer bir zaman dizisi durağansa, ortalaması, varyansı ve (çeşitli gecikmelerde) kovaryansı, zamandan bağımsız aynı değere sahip olur.

Eğer zaman dizisi yukarıdaki anlamda durağan değilse, *durağan olmayan zaman dizisi* adını alır [3].

5.2. Durağanlığın Birim Kök İle Test Edilmesi

İki zaman dizisi arasındaki eşbütünleşmenin varlığını tespit edebilmek için öncelikle bu iki zaman dizisinin aynı dereceden durağan olup olmadığını test etmek gerekir. Şayet her iki dizide aynı dereceden durağan oluyorsa eşbütünleşme testine geçilir. Bu amaçla en çok yaygınlaşan testlerden biri *birim kök testi*dir.

Model aşağıdaki gibi verilmiş olsun.

$$Y_t = Y_{t-1} + u_t$$

Modelde, u_t, ortalaması 0, varyansı değişmeyen σ^2 olan, ardışık bağımlı olmayan olasılıklı hata terimidir (beyaz gürültü hata terimi).

Yukarıdaki model AR(1) modeli olup *t* dönemdeki *Y*'nin *(t-1)* dönemindeki kendi değerine göre regresyonudur.

Modelde, Y_{t-1}'nin katsayısı 1'e eşitse *birim kök sorunu* sözkonusudur. Model bir başka biçimde aşağıdaki gibi yazılmış olsun.

$$Y_t = \rho Y_{t-1} + u_t \tag{5.1}$$

Bu durumda, $\rho = 1$ bulunursa bu modelin birim kökü olduğu söylenir. Birim kökü olan bir zaman dizisi rassal yürüyüş olarakta bilinir ve değişkenin durağan olmayan bir zaman dizisine sahip olduğunu gösterir.

Eş.5.1'deki model,

$$\begin{aligned} \Delta Y_t &= (\rho - 1)Y_{t-1} + u_t \\ &= \delta Y_{t-1} + u_t \end{aligned} \tag{5.2}$$

biçiminde ifade edildiğinde, $\delta = (\rho - 1)$ ve Δ birinci fark işlemi göstermek üzere,

$$\Delta Y_t = (Y_t - Y_{t-1})$$

olur. δ gerçekte 0 ise;

$$\Delta Y_t = (Y_t - Y_{t-1}) = u_t \tag{5.3}$$

olur. Rassal bir yürüyüşün birinci farkları u_t olan durağan bir zaman dizisidir. Çünkü varsayım gereği u_t bütünüyle rassaldır [3].

Zaman dizisinin birinci farkları alındığında, dizi durağan çıkarsa, başlangıçtaki rassal yürüyüş dizisi 1. dereceden durağandır denir ve I(1) ile gösterilir. Durağan bir diziye ulaşmak için dizinin iki kez farkının alınması gerekiyorsa, dizi 2. dereceden durağandır ve I(2) gösterilir. Genel olarak ifade edilecek olursa, bir zaman dizisinin d kez farkının alınması gerekiyorsa o dizi d'inci dereceden durağandır denir ve I(d) ile gösterilir.

Bu nedenle 1. ya da daha yukarı dereceden durağan bir zaman dizisi varsa durağan olmayan bir zaman dizisi var demektir. $d = 0$ ise; I(0) ile gösterilir ve durağan süreç olarak ifade edilir.

Y_t gibi bir zaman dizisinin durağan olup olmadığını anlamak için Eş.5.1'deki regresyon hesaplanır ve $\hat{\rho}$'nin istatistik bakımından 1'e eşit olup olmadığına bakılır. Ya da buna eşdeğer olarak Eş.5.2 tahmin edilir, t istatistiğine göre $\delta = 0$ olup olmadığını bakılır. Ancak burada bir sorun, bu yolla bulunan t değerinin büyük örneklerde bile student t dağılımına uymamasıdır.

$H_0 : \rho = 1$, hipotezinin geleneksel yolla hesaplanan t istatistiği, τ(tau) istatistiği diye bilinir. Bunun kritik değeri Dickey ile Fuller (1979) tarafından Monte Carlo benzetimleriyle üretilerek tablolaştırılmıştır [3].

τ testi bulanların adı ile Dickey-Fuller (DF) testi diye tanınır. Eğer $H_0 : \rho = 1$ hipotezi reddedilirse (yani zaman dizisi durağansa) bilinen t testi kullanılabilir. Daha açık bir ifadeyle, Eş.5.1 regresyonun katsayıları tahmin edilir. Dickey-Fuller τ istatistiğinin hesaplanması için tahmin edilen $\hat{\rho}$ katsayısının kendi standart hatasına bölünmesi gerekir. $H_0 : \rho = 1$ hipotezinin reddedilip reddedilmediği görmek için Dickey-Fuller çizelgesine başvurulur. Ancak bu çizelge yeterli değildir ve Mackinnon tarafından Monte Carlo benzetimleriyle büyük ölçüde genişletilmiştir [4].

Eğer τ istatistiğinin mutlak değeri DF'nin ya da Mackinnon DF'nin mutlak kritik τ değerinden küçükse, zaman dizisinin durağan olmadığı biçiminde yorumlanır. Öte yandan, τ istatistik değeri mutlak değerce kritik değerden büyük ise; zaman dizisi durağandır denir [3].

Dickey-Fuller testi, Eş.5.2 ile aşağıda gösterilen biçimlerdeki regresyonlara uygulanır.

$$\Delta Y_t = \beta_1 + \delta Y_{t-1} + u_t \tag{5.4}$$

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + u_t \tag{5.5}$$

Bu eşitliklerde *t,* zaman ya da genel eğilim değişkenini göstermektedir.

Her bir durumda H_0 hipotezi $\delta = 0$, yani "birim kök var" biçiminde oluşturulur.

Eş.5.2'deki regresyon ile diğer iki regresyon arasındaki fark, sabit terimin ve eğilim değişkeninin denkleme katılmasıdır.

Eş.5.5 regresyonu aşağıda gösterildiği gibi yazılabilir.

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \alpha_i \sum_{i=2}^{m} \Delta Y_{t-i} + \varepsilon_t \tag{5.6}$$

Burada, $\Delta Y_{t-1} = (Y_{t-1} - Y_{t-2}), \Delta Y_{t-2} = (Y_{t-2} - Y_{t-3})$...vb. gecikmeli fark terimlerini göstermektedir.

Eş.5.6'daki modelde amaç, ε_t 'nin ardışık bağımsız olmasını sağlayacak kadar terimi modele katmaktır. Burada H_0 hipotezi,

$\delta = 0$, ya da $\rho = 1$ biçiminde kurulur. Daha açık bir ifadeyle, *Y'de birim kök vardır* ya da *Y durağan değildir* şeklinde ifade edilir.

Eş.5.6'daki modeldeki verilere DF uygulanırsa buna *Genişletilmiş Dickey Fuller (GDF) testi* denir.

5.3.Durağan Olmayan Çok Değişkenli Zaman Dizileri

Çok değişkenli zaman dizilerinde durağanlık tanımı, tek değişkenli zaman dizilerindeki gibidir. k-boyutlu bir $X = (X_1, X_2, ..., X_k)'$ rasgele vektörünün beklenen değer vektörü,

$$E(X) = (E(X_1), E(X_2), ..., E(X_k))'$$

şeklinde olur. Benzer şekilde varyans-kovaryans matrisi ise;

$$Var(X) = \begin{bmatrix} Var(X_1) & Kov(X_1, X_2) & . & . & Kov(X_1, X_k) \\ Kov(X_2, X_1) & Var(X_2) & . & . & Kov(X_2, X_k) \\ . & . & . & . & . \\ . & . & . & . & . \\ Kov(X_k, X_1) & Kov(X_k, X_2) & . & . & Var(X_k) \end{bmatrix}$$

olur. Çok değişkenli zaman dizilerinde durağanlık tanımı ise; herhangi bir k-değişkenli vektör zaman dizisi, $\{X : t \in n\}$ olsun. Şayet,

i. $E(X) = \mu$, t 'den bağımsız,

ii. $\Gamma(h) = E(X_t, X'_{t+h}) - E(X_t)E(X_{t+h})$ sadece h'nin bir fonksiyonu (t'den bağımsız) ise; $\{X_t : t = 1,2,3,...\}$ zaman dizisine durağandır denir. Bu özellikleri sağlamayan zaman dizileri de durağan olmayan çok değişkenli zaman dizileridir.

Burada, $E(X_t) = (E(X_{1,t}), E(X_{2,t}),..., E(X_{k,t}))'$ ve $\Gamma(h)$ matrisi simetrik değildir. Ama $\Gamma(-h) = [\Gamma(h)]'$ özelliğine sahiptir [1].

5.4. Yanıltıcı Regresyon (Spurious Regression)

Zaman dizisi kullanan regresyon sonuçları yüzeysel olarak gayet iyi görünmelerine karşın derinlemesine incelendiğinde kuşku uyandırmaları anlamında yanıltıcı ya da kuşkulu bulgular elde etme olasılığını içerebilir.

Yanıltıcı regresyon sorununu, Gujarati (2001)'nin çalışmasında güzel bir örnek problemle açıklanmıştır.

KTH: kişisel tüketim harcamasını
KHG: kişisel harcanabilir geliri

temsil etmek üzere,

$$K\hat{T}H_t = -171,4412 + 0,9672\ KHG_t$$

$$t = \ (-7,4809)\quad (119,8711) \qquad (5.7)$$

$$R^2 = 0.9940 \quad d = 0,5316$$

zaman dizilerini kullanan regresyon tahmin edilmiş olsun. Burada R^2 ve *KHG*'nin t oranı yüksek görünmektedir. Tek sorun Durbin-Watson d'nin düşük olmasıdır. $R^2 > d$ olması tahmin edilen regresyonun yanıltıcı olduğundan kuşkulanmanın pratik yollarından birisidir.

Yukarıdaki örneğe ilişkin farka dayalı regresyon tahminleri aşağıdaki gibi olsun.

$$\begin{aligned} \Delta K\hat{T}H_t &= 91{,}7110 + 0{,}7704t - 0{,}0432\,KTH_{t-1} \\ t &= (1{,}6358) \quad (1{,}2983) \quad (-1{,}3276) \end{aligned} \tag{5.8}$$

$$\begin{aligned} \Delta K\hat{H}G_t &= 326{,}2089 + 2{,}8834t - 0{,}1579\,KHG_{t-1} \\ t &= (2{,}7368) \quad (2{,}5243) \quad (-2{,}5751) \end{aligned} \tag{5.9}$$

Burada yalnızca gecikmeli *KTH* ile *KHG*'nin τ değerleriyle ilgilenilmektedir. MacKinnon'nun hesapladığı %1, %5, %10 kritik τ ya da DF değerleri sırasıyla -4,0673,-3,4620,-3,2447'dir. Gecikmeli *KTH* ile *KHG*'nin 1,3761 ve 2,5883 olan τ mutlak değerleri, %10 düzeyindeki kritik τ değerinden bile düşüktür. Sonuç olarak, *KTH* ile *KHG*'nin her ikisinde de birim kök sorunu vardır, yani ikisi de durağan değildir.

Eş.5.7'deki *KTH*'nin *KHG*'ye göre regresyonunu bulurken, durağan olmayan bir zaman dizisinin durağan olmayan bir zaman dizisine göre regresyonu sözkonusudur. Bu durumda bilinen t ve F testleri geçerli değildir. Eş.5.7'deki regresyon, bu anlamda yanıltıcı regresyondur. Burada dikkat edilmesi gereken bir nokta, Eş.5.9'daki KHG_{t-1}'in -2,5883 olan t değerine, bilinen t testi ile bakıldığında, %2 düzeyinde istatistiksel bakımdan anlamlı olduğunun görülmesidir. Oysaki, τ testi ile bu değer %10 düzeyinde bile anlamlı değildir. Bu ise, durağan olmayan zaman dizilerinde yalnızca t değerine neden güvenilmemesi gerektiğini göstermektedir.

Örnek problemdeki her bir dizi, rassal ilerler ama aralarında bir birliktelik bulunur. Aynı dereceden durağan zaman dizilerinin arkasındaki düşünce sezgisel olarak eşzamanlılıktır. Buna göre Eş.5.7'deki regresyon modeli yanıltıcı olmayabilir ve bilinen t ve F testleri kullanılabilir. Granger'in dediği gibi "Bir eşbütünleşme testi yanıltıcı regresyondan sakınmak için bir ön test olarak düşünülebilir" [3].

6. EŞBÜTÜNLEŞME VEKTÖRÜNÜN TAHMİNİ

Eşbütünleşme vektörünü tahmin etmek için çok değişkenli zaman dizisinin parametre matrisini tahmin etmeye gerek yoktur. Burada U_t durağan olmayan bir zaman dizisi, S_t ise durağan bir diziyi ifade etmek üzere iki değişkenli bir vektör zaman dizisinin bileşenleri,

$$Y_t = a_{11}U_t + a_{12}S_t$$
$$Z_t = a_{21}U_t + a_{22}S_t$$

olarak verilmiş olsun. Buradan haraketle,

$$X_t = Z_t - \frac{a_{21}}{a_{11}}Y_t = \left(a_{22} - \frac{a_{21}a_{12}}{a_{11}}\right)S_t$$

olur ve S_t durağan olduğu için herhangi bir sabit bir katsayı ile çarpılması durağanlığı etkilemez. Bu nedenle de X_t durağan bir zaman dizisi olur. Dolayısıyla eşbütünleşme vektörünü tahmin etmek için a_{21} / a_{11} oranını tahmin etmek yeterli olacaktır. Burada X_t dizisi, Z_t dizisinin Y_t üzerine regresyonu yapıldığında artıklar dizisine benzer. Yani,

$$Z_t = \beta Y_t + \varepsilon_t, \quad t = 1,2,3,...,n$$

regresyon denkleminden elde edilen artıklar dizisi durağan ise; dizi eşbütünleşiktir. Bu tahmin edicinin özellikleri Engle ve Granger (1987) tarafından incelenmiştir. Bu nedenle yöntem literatürde Engle-Granger yöntemi olarak bilinir. Z_t dizisinin Y_t üzerine regresyonunda β 'nın en çok olabilirlik tahmin edicisi,

$$\hat{\beta} = \frac{\sum_{t=1}^{n} Z_t Y_t}{\sum_{t=1}^{n} Y_t^2}$$

olmak üzere,

$$\frac{1}{n^2}\sum_{t=1}^{n} Z_t Y_t = a_{21}a_{11}\sum_{t=1}^{n} U_t^2$$

ve

$$\frac{1}{n^2}\sum_{t=1}^{n} Y_t^2 = a_{11}^2\sum_{t=1}^{n} U_t^2$$

olur. Burada,

$$\hat{\beta} = \frac{\sum_{t=1}^{n} Z_t Y_t}{\sum_{t=1}^{n} Y_t^2} = \frac{a_{21}}{a_{11}}$$

şeklinde elde edilir. Buradan,

$$\begin{aligned} X_t &= Z_t - \hat{\beta} Y_t \\ &= (a_{21}U_t + a_{22}S_t) - \frac{a_{21}}{a_{11}}(a_{11}U_t + a_{12}S_t) = CS_t \end{aligned}$$

olan durağan dizisi elde edilir. Yani, regresyon eşbütünleşme vektörünü bulunur ve bu vektör,

$$\beta' = \left(-\frac{a_{21}}{a_{11}}, 1\right)'$$

olur [1].

7. EŞBÜTÜNLEŞME TESTLERİ

Dizinin bileşenleri arasındaki lineer durağan ilişkiye *eşbütünleşme* denir. Eşbütünleşme testi için eşbütünleşme regresyonu aşağıdaki gibi yazılır.

$$Y_t = \beta_1 + \beta_2 X_t + u_t \qquad t = 1,2,...,n \tag{7.1}$$

Burada u_t, ortalaması 0, varyansı σ^2 olan, ardışık bağımsız hata terimidir. Bu regresyon denkleminde artıklar,

$$\hat{u}_t = Y_t - \hat{\beta}_1 - \hat{\beta}_2 X_t \tag{7.2}$$

olarak tanımlanır. Artıkların durağan olup olmadığın testi için,

$$\Delta u_t = \beta u_{t-1} + \varepsilon_t \tag{7.3}$$

denkleminde regresyon yapılır. Burada ε_t, ortalaması 0, varyansı σ^2 olan hata terimidir. Daha sonra $H_0 : \beta = 0$ hipotezi kurulur. Şayet hipotez kabul edilirse hata terimi u_t'nin durağan olmadığı anlaşılır. Bu ise; Eş.7.1'deki dizilerin eşbütünleşik olmadığını gösterir. Tam tersine, H_0 hipotezinin reddi söz konusu olduğunda ise, Y_t ve X_t dizileri arasında eşbütünleşmenin olduğu biçiminde yorumlanır.

Yukarıda verilen örnek problem için artıklar denklemi şöyle ifade edilir.

$$\hat{u}_t = KTH_t - \beta_1 - \beta_2 KHG_t \tag{7.4}$$

u_t'nin I(0) ya da durağan olduğu belirlenirse, *KTH* ve *KHG* değişkenlerinin *eşbütünleşik* oldukları söylenebilir.

Eğer her iki değişken de aynı dereceden durağan iseler, aynı yönlü eğilimler gösterecektir. Demek ki bir *Y* dizisi I(1), başka bir *X* dizisi de I(1) ise bunlar eşbütünleşik olabilir.

Kısacası, Eş.5.7'deki gibi regresyonlar da artıkların I(0) olup olmadığına bakmak koşuluyla bilinen regresyon yöntemleri (t ve F testi) zaman dizisi verilerine uygulanabilir.

Burada eşbütünleşme kuramı gereği Eş.5.7'deki gibi bir regresyon *eşbütünleşme regresyonu*, eğim katsayısı β_2'de *eşbütünleşme katsayısı* diye adlandırılır [3].[1]

Engle ve Granger (1987) çalışmasında yedi eşbütünleşme testine yer vermiştir. Bu testler için gerekli olan kritik değerler, 10000 tekrarlı bir benzetim ile elde edilmiştir. Bu kritik değerler 100 gözlem olmak üzere %1,%5 ve %10 yanılma düzeyleri için hesaplanmıştır.

7.1. Engle-Granger (EG) ya da Genişletilmiş Engle-Granger (GEG) Eşbütünleşme Testi

DF ya da GDF birim kök testini şu şekilde yapılır, Eş.5.7'deki gibi bir regresyonu tahmin etmek, artıklarını bulmak, DF ya da GDF testlerini uygulamaktır.[2] Ancak dikkat edilmesi gereken bir husus vardır. Tahmin edilen *u*, yığının tahmin edilen eşbütünleşme katsayısı β_2'ye bağlı olduğundan DF ve GDF kritik değerleri bu iş için pek de uygun sayılmaz. Engle ve Granger, kaynaklar arasında bulunabilecek bu değerleri hesaplamışlardır. Bu nedenle DF ve GDF testleri bu bağlamda *Engle-Granger (EG) testi* ve *Genişletilmiş Engle-Granger (GEG) testi* adlarını alır.

[1] Eşbütünleşim kavramı k tane açıklayıcı değişkeni olan bir regresyon modeline genişletilebilir. Bu durumda k tane eşbütünleşim katsayısı olur.

[2] KTH ya da KHG eşbütünleşik değilse, bunların herhangi bir lineer bileşimi, dolayısıyla u kalıntıları da lineer olmaz.

Şimdi Eş.5.7'deki KTH'in KHG'ye göre regresyonuna bakıp tahmin edilen DF testi uygulandığında şu bulgular elde edilir.

$$\Delta\hat{u}_t = -0{,}2716\,\hat{u}_{t-1}$$
$$t = (-3{,}6725) \tag{7.5}$$

Engle-Granger %1, %5, %10 kritik τ (yukarıdaki regresyonda t istatistiği) değerleri sırasıyla -2,5899, -1,9439, -1,6177'dir. 3,7791 olarak tahmin edilen τ değeri mutlak olarak bu kritik değerlerden büyük olduğu için, tahmin edilen u_t'nin durağan olduğu, dolayısıyla *KTH* ile *KHG*'nin de, tek başlarına durağan olmamakla birlikte, eşbütünleşik olduklarıdır. Bu test, bir gecikmeli denklemler için geçerlidir [3].[3]

$Y_t = \beta_0 + \beta_1 X_t + u_t$ eşbütünleşme regresyonundan elde edilen artıkların durağan olup olmadığının kontrolü için,

$$\Delta\hat{u}_t = \beta\,\hat{u}_{t-1} + b_1\Delta\hat{u}_{t-1} + \ldots + b_k\Delta\hat{u}_{t-k} + \varepsilon_t \tag{7.6}$$

biçiminde bir denkleme Engle-Granger Testi uygulanmasına ise, Genişletilmiş Engle-Granger Eşbütünleşme Testi adı verilir. Burada hata terimi durağan olarak bulunursa yani birim kök mevcut değil ise X_t ve Y_t dizileri eşbütünleşiktir denir. Bu test, k gecikmeli denklemler için geçerlidir [5].

7.2. Eşbütünleşme Regresyonu Durbin-Watson (ERDW) Testi

Daha çok ardışık bağımlılığı bulmak için kullanılan Durbin-Watson test istatistik değeri,

[3] Tahmin edilen eşbütünleşik "Eş.5.7" numaralı regresyonuna ilişkin teknik bir nokta ise;KTH ile KHG tek tek I(1) olduklarından, tahmin edilmiş standart hatalar ile yığın katsayılarına ilişkin çıkarsama yapma amacıyla bu regresyonda gösterilen tahmin edilmiş katsayılara ilişkin t değerleri de kullanılamaz.

$$d = \frac{\sum_{t=2}^{n} (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{n} \hat{u}_t^2}$$

şeklinde tanımlanır. Ardışık hata fark kareleri toplamının hata kareler toplamına oranıdır. Burada pay kısmında ardışık farklar alınması nedeniyle $(n-1)$ tane gözlem bulunmaktadır. Durbin-Watson testi,

$H_0 : d = 2$ (ardışık bağımlılık yoktur)

hipotezini test eder. Bu hipotezin reddi ya da kabulünü gösteren tek bir kritik değer yoktur. Durbin ile Watson öyle bir alt sınır d_L ile öyle bir üst sınır d_U türetmişlerdir ki hesaplanan d bu sınırlar dışındaysa aynı ya da ters yönlü ardışık bağımlılığın varlığına karar verilir. Bu d_L ve d_U değerleri n, anlamlılık düzeyleri ve bağımsız değişken sayısına (k) göre bulunan tablo değerleridir. Burada d 'nin sınırları,

$0 \leq d \leq 4$

olur ve her d değeri bu sınırlar içindedir. Burada hesaplanan d 'nin 2 civarlarında olması ardışık bağımlılığın olmadığını gösterir. Şayet,

$d < d_L$ ise; aynı yönlü ardışık bağımlılık,

$d_L \leq d \leq d_U$ ise; belirsiz durum,

$d_U < d < 4\text{-} d_U$ ise; ardışık bağımlılık yoktur,

$4 - d_U \leq d \leq 4 - d_L$ ise; belirsiz durum,

$4 - d_L < d$ ise; neğatif yönlü ardışık bağımlılık vardır denir [6].

Eşbütünleşme testinde ise; hata terimi tahmini,

$$\hat{u}_t = Y_t - \beta_1 - \beta_2 X_t$$

olan Eş.7.1'deki eşbütünleşme regresyonu kullanılır. Bu testin H_0 hipotezi;

$H_0 : d = 0$ (aynı yönlü ardışık bağımlı)

olarak ifade edilir. Bu hipotezin anlamı artıkların durağan olmadığıdır. Durbin-Watson istatistiği sıfıra yakın bir değer olduğunda, artıkların durağan olmadığı yani X_t ve Y_t dizileri arasında eşbütünleşmenin olmadığı anlamına gelir.

Şayet d istatistiği sıfırdan büyük bir değer ise; artıkların durağan olduğunu ve X_t ve Y_t dizilerinin eşbütünleşik olduğunu gösterir. Bu yöntem, bir gecikmeli dizi içeren denklemlerde kullanılır.

KTH ile *KHG*'nin eşbütünleşik olup olmadığını anlamanın başka ve daha basit bir yolu, kritik değerleri ilk kez Sargan ile Bhargava tarafından sağlanan ERDW testidir [7]. ERDW'de Eş.5.7'deki denklemde verilen $d = 0{,}5316$ gibi, eşbütünleşme regresyonla elde edilen Durdin-Watson d kullanılır. Her birinde 100 gözlem olan 10000 kere tekrarlanan benzetimle, gerçek $d = 0$ hipotezini test etmek için hesaplanan %1, %5, %10 kritik değerleri sırasıyla 0,511, 0,386, 0,322'dir. Öyleyse hesaplanan d değeri, diyelim ki 0,511'den küçükse, eşbütünleşme hipotezi %1 anlamlılık düzeyinde red edilir. Örneğimizde d değeri 0,5316 kritik değerinden büyüktür, bu da *KTH* ile *KHG*'nin eşbütünleşik olduğu anlamına gelir. Burada testin EG testi ile aynı sonucu verdiği görülür. ERDW'nin DF'ye üstünlüğü konusunda bir hayli tartışma vardır. Tartışma iki istatistiğin gücü, yani II.Tip hata (yokluk hipotezi yanlışken kabul etme) olasılığı çerçevesindedir [3].

Hem EG hem de ERDW testleri sonucunda *KTH* ile *KHG'*nin eşbütünleşik olduğu sonucuna varılır. Yani, bu iki zaman dizisinin uzun dönemde ilişkileri anlamlı olacaktır.

Şimdi Engle ve Granger (1987) çalışmasındaki diğer testlerden kısaca bahsetmek yararlı olacaktır.

7.3. Kısıtlı Vektör Otoregresyon Testi (KVAR) ya da Genişletilmiş Kısıtlı Vektör Otoregresyon Testi (GKVAR)

Eşbütünleşme regresyon denkleminden u_t artıkları elde edildikten sonra,

$$\Delta Y_t = \beta_1 \hat{u}_{t-1} + \varepsilon_{1t} \tag{7.7}$$

denklemi için,

$$H_0 : \beta_1 = 0$$

hipotezi test edilir. t testi $T_{\beta 1}$ şeklinde gösterilir. Eş.7.7'deki denklemden sonra,

$$\Delta X_t = \beta_2 \hat{u}_{t-1} + \gamma \Delta Y_t + \varepsilon_{2t} \tag{7.8}$$

denkleminde de,

$$H_0 : \beta_2 = 0$$

hipotezi test edilir. t testi de $T_{\beta 2}$ şeklinde gösterilir. $T_{\beta 1}$ ve $T_{\beta 2}$ test istatistikleri hesaplandıktan sonra bu istatistiklerin kareleri toplamı kısıtlı vektör otoregresyon test istatistiğini verir. Bu testin amacı, eşbütünleşme regresyonunda hata teriminin, Eş.7.7 ve Eş.7.8'deki denklemlerde önemli olup olmadığını test etmektir. Ayrıca bu test, bir gecikmeli dizi içeren denklemler için geçerli olmaktadır.

GKVAR testi k gecikmeli dizi içeren denklemlerde uygulanır. Bunun dışında KVAR ile GKVAR arasında hiçbir fark yoktur [5].

7.4. Kısıtsız Vektör Otoregresyon Testi ya da Genişletilmiş Kısıtsız Vektör Otoregresyon Testi

Bu teste ilişkin denklem,

$$Y_t = c_t + \beta_1 Y_{t-1} + \beta_2 X_{t-1} + \varepsilon_{1t} \tag{7.9}$$

şeklinde kurulur ve H_0 hipotezi,

$$H_0 : \beta_1 = \beta_2 = 0$$

olarak ifade edilip, F testi uygulanır ve F_1 ile gösterilir. Daha sonra,

$$\Delta X_t = c_2 + \beta_3 Y_{t-1} + \beta_4 X_{t-1} + \gamma \Delta Y_t + \varepsilon_{2t} \tag{7.10}$$

denklemine regresyon uygulanır ve

$$H_0 : \beta_3 = \beta_4$$

hipotezi test edilir. Test istatistiği olarak F testi kullanılır ve F_2 olarak gösterilir. Hesaplanan F_1 ve F_2 istatistik değerleri toplamının iki katı, Kısıtsız VAR test istatistik değerini verir. Bu test gecikmeli dizilerin denklemde anlamlı olup olmadığını gösterir. Bu test bir gecikmeli dizi içeren denklemlerde geçerli olmaktadır.

k gecikmeli denklemlerde, bu uygulamanın yapılması ile test, Genişletilmiş Kısıtsız VAR testini ifade eder [5].

8. ÇOK DEĞİŞKENLİ ZAMAN DİZİLERİNDE EŞBÜTÜNLEŞME

Eşbütünleşme vektörünün tahmin edilmesinde ve dizilerin eşbütünleşik olup olmadığının testi için literatürde çok sayıda çalışma vardır. Bunlardan biri yukarıda belirttiğimiz Engle-Granger testi ve Johansen (1988)'in önerdiği en çok olabilirlik yöntemi kullanılan standart yöntemlerdir.

Johansen'nin önerdiği yöntemin kullanılmasının iki nedeni vardır. Bunlardan ilki, ilgilenilen değişkenler için lineer bağımsız eşbütünleşmelerin sayısını bulmak, ikincisi ise, eşbütünleşme vektörünün ve ilgili parametrelerin en çok olabilirlik tahminlerini elde etmektir. Johansen yöntemi, dizinin eşbütünleşik olup olmadığının testi için parametre matrislerinin özdeğerlerinden yararlanır [1].

Johansen testinde ele alınan süreçlerin lineer modellerinde yer alan ε_t $(p \times 1)$ boyutlu vektör olmak üzere, ortalaması sıfır, varyans kovaryans matrisi V matrisi olan beyaz gürültü sürecini ifade etmektedir. $(p \times 1)$ boyutlu normal dağılıma sahip rasgele değişkenler olduğu varsayımı altında birinci dereceden vektör otoregresyon zaman dizisi,

$$X_t = AX_{t-1} + \varepsilon_t, \qquad (t = 1,2,....n) \tag{8.1}$$

şeklinde olur. Burada A matrisi $(p \times p)$ boyutlu parametre matrisidir. Eş.8.1'deki denklemde X_t birinci dereceden durağandır. Yani, her iki taraftan X_{t-1} çıkarılırsa, $\Pi = A - I$ olmak üzere,

$$\Delta X_t = \Pi X_{t-1} + \varepsilon_t \tag{8.2}$$

modeline ulaşılır. Burada I , $(p \times p)$ boyutlu birim matristir [8].

Burada Π matrisinin rankı eşbütünleşme vektörlerinin sayısını verir. Bu irdeleme genelleştirilecek olursa,

$$X_t = \pi_t X_{t-1} + ... + \pi_k X_{t-k} + \mu + \Phi D_t + \varepsilon_t, \quad (t = 1,2,...,n) \tag{8.3}$$

biçiminde ifade edilir. Burada D_t mevsimselliği gösteren değişkendir. Ekonomik veriler daha çok durağan olmayan veriler olduğundan dolayı Eş.8.3'deki VAR modeli çoğunlukla birinci farkları alınarak,

$$\Delta X_t = \Gamma_1 \Delta X_{t-1} + ... + \Gamma_{k-1} \Delta X_{t-k+1} + \Pi X_{t-k} + \mu + \Phi D_t + \varepsilon_t \tag{8.4}$$

şeklinde gösterilir(Johansen ve Juselius(1990)). Burada,

$$\Gamma_i = -(I - \pi_1 - ... - \pi_i), \quad (i = 1,...,k-1)$$

ve

$$\Pi = -(I - \pi_1 - ... - \pi_k)$$

olarak tanımlanır. Bu çalışmanın esas amacı, katsayılar matrisi Π'nin, veri vektöründeki değişkenler arasında uzun dönemli ilişki olup olmadığını araştırmaktır. Buradaki sonuçlar özetlenirse,

1. Rank(Π) = 0 ise; Π matrisinin sıfır matrisidir. Yani dizi durağan değildir ve dizinin hiç bir lineer birleşimi de durağan değildir. Eşbütünleşme yapısına uymaz.
2. Rank(Π) = p olması halinde Π matrisi tam ranklıdır ve X_t süreç vektörü durağandır.
3. 0 < Rank(Π) = r < p olması da uzun dönemli bir ilişkinin olduğunu ve eşbütünleşmenin varlığını gösterir.Yani $\Pi = \alpha\beta'$ olacak şekilde öyle bir α ve β' vektör veya matrisleri vardır ki $\beta' X_t$ durağandır [8].

9. ÇOK DEĞİŞKENLİ EŞBÜTÜNLEŞME TAHMİNİ VE EŞBÜTÜNLEŞME TESTİ

Johansen (1988) eşbütünleşme çalışmasında tahmini aşağıdaki gibi kurmaktadır.

$$\Pi = \alpha\beta' \tag{9.1}$$

Burada α ve β, her ikisi de $(p \times r)$ boyutlu iki matristir. X_t dizisi durağan değil iken, ΔX_t'nin durağan olması koşuluyla Eş.9.1'deki denklemin geçerliliği durumunda $\beta' X_t$ ile gösterilen lineer birleşimler durağandır. Durağan ilişkilere sahip olan $\beta' X_t$'ye *eşbütünleşik ilişkiler* adı verilir. Granger $\beta' X_t$'nin β eşbütünleme vektörü ile X_t vektör sürecinin eşbütünleşik olduğu söyler. Dolayısı ile Johansen testinde β vektörünün tahmini yerine Π matrisinin rankı üzerine testler oluşturulur [1].

Johansen ve Juselius (1990), Eş.9.1'deki denklemde β'nin en çok olabilirlik tahminini bulmuşlardır. β'nin tahmininden önce bir takım tanımlamalar vermek gerekir. Johansen ve Juselius (1990) çalışmalarında Eş.8.4'deki VAR modelini kullanmışlardır. Bu modelden birtakım yeni gösterimler elde edilmiştir. Bunlar:

$$\mathrm{Z}_{0t} = \Delta X_t \qquad\qquad \mathrm{Z}_{1t} = (\Delta X_{t-1}, ..., \Delta X_{t-k+1}, D_t, 1)$$

değişkenlerini ifade ederken $Z_{kt} = X_{t-k}$ olur. Bu durumda Γ ise Z_{1t}'ye karşılık gelen parametreler matrisini tanımlar. Yani, $\Gamma' = (\Gamma_1, ..., \Gamma_k, \Phi, \mu)$ olur. Bu durumda Eş.8.4'deki denklem,

$$Z_{0t} = \Gamma Z_{1t} + \Pi Z_{kt} + \varepsilon_t, \qquad (t = 1,2,...,n) \tag{9.2}$$

olarak yeniden düzenlenir. Π 'nin sabit bir değeri için, en çok olabilirlik tahmini, $Z_{0t} - \Pi Z_{kt}$ bağımlı ve Z_{1t} bağımsız dizi olmak üzere bir regresyon içerir. Bu regresyon sonucunda,

$$\sum_{t=1}^{n} Z_{0t} Z_{1t}' = \Gamma \sum_{t=1}^{n} Z_{1t} Z_{1t}' + \Pi \sum_{t=1}^{n} Z_{kt} Z_{1t}' \qquad (9.3)$$

şeklinde bir denklem yazılabilir. Burada, n örnek büyüklüğünü ifade eder. Çarpım moment matrisi M_{ij} ise,

$$M_{ij} = n^{-1} \sum_{t=1}^{n} Z_{it} Z_{jt}', \qquad (i, j = 0,1,k) \qquad (9.4)$$

biçiminde tanımlanır. Bu durumda yukarıda verilen model,

$$M_{01} = \Gamma M_{11} + \Pi M_{k1}$$

veya

$$\Gamma = M_{01} M_{11}^{-1} - \Pi M_{k1} M_{11}^{-1} \qquad (9.5)$$

olarak yazılır. Buradan artıklar denklemi,

$$R_{0t} = Z_{0t} - M_{01} M_{11}^{-1} Z_{1t} \qquad (9.6)$$

$$R_{kt} = Z_{kt} - M_{k1} M_{11}^{-1} Z_{1t} \qquad (9.7)$$

biçiminde tanımlanır. Eş.9.6'daki denklemdeki artıklar ΔX_t bağımlı, $\Delta X_{t-1}, ..., \Delta X_{t-k+1}$, D_t ve 1 bağımsız diziler olmak üzere, regresyondan elde edilirken, Eş.9.7'deki denklemdeki artıklar da, X_{t-k} bağımlı $\Delta X_{t-1}, ..., \Delta X_{t-k+1}, D_t$ ve 1 bağımsız diziler olmak üzere regresyondan elde edilir. Artıkların çarpım moment matrisi de,

$$S_{ij} = n^{-1}\sum_{t=1}^{n} R_{it}R'_{jt} = M_{ij} - M_{i1}M_{11}^{-1}M_{1j}, \quad (i, j = 0, k) \tag{9.8}$$

olur. Bu işlemlerde matris boyutları önemlidir. Burada $(k-1)$ tane ΔX_t vektörü ve her ΔX_t vektöründe p tane değişken olduğundan, ayrıca üç tane mevsimsel değişken (D_t) ve bir tane de μ parametresinin katsayısı 1 olduğundan Z_{1t} vektörünün boyutu $(p(k-1)+3+1 \times 1)$ olur. Γ matrisinin boyutu ise $(p \times p(k-1)+3+1)$ olur. Aynı mantıkla Z_{0t} ve Z_{1t} vektörlerinin boyutunun $(p \times 1)$ olduğu kolayca görülür. Matrisyel çarpımlar yapıldığında M_{00}, M_{0k}, M_{k0} ve M_{kk} $(p \times p)$, M_{11} $((p(k-1)+3+1) \times (p(k-1)+3+1))$, M_{01} ve M_{k1} $(p \times (p(k-1)+3+1))$, M_{10} ve M_{1k} $((p(k-1)+3+1) \times p)$, R_{0t} ve R_{kt} vektörleri $(p \times 1)$ ve S_{00}, S_{0k}, S_{k0}, S_{kk} matrisleri de $(p \times p)$ boyutlu olur. Bu durumda,

$$H_1 : \Delta X_t = \sum_{j=1}^{k-1} \Gamma_j \Delta X_{t-j} + \Pi X_{t-k} + \mu + \Phi D_t + \varepsilon_t \tag{9.9}$$

hipotezini ele alınır ve EKK uygulanırsa,

$$\hat{\Pi} = S_{ok}S_{kk}^{-1} \tag{9.10}$$

ve

$$\hat{\Lambda} = S_{00} - S_{0k}S_{kk}^{-1}S_{k0} \tag{9.11}$$

$$L_{\max}^{-2/n}(H_1) = \left|\hat{\Lambda}\right| \tag{9.12}$$

bulunur. Π 'yi tahmin ettikten sonra Eş.9.5'deki denklemde yerine konularak Γ 'nın tahmini yapılır. Eş.9.9'daki hipotezin en çok olabilirlik tahmininden sonra Eş.9.1'deki denklemdeki β 'nın tahminine geçilir.

$$H_2 : \Pi = \alpha\beta' \tag{9.13}$$

olsun. Burada ilk olarak,

$$\left|\lambda S_{kk} - S_{k0} S'_{00} S_{0k}\right| = 0 \tag{9.14}$$

denklemi çözülür. Buradan özdeğerler $\hat{\lambda}_1 > \ldots > \hat{\lambda}_p$ olacak şekilde bulunur. Bu özdeğerlere karşılık gelen özvektörler, $\hat{V}' S_{kk} \hat{V} = I$ işlemi uygulanarak $\hat{V} = (\hat{v}_1, ..., \hat{v}_P)$ olarak bulunur. Bu özvektörler β 'nın tahmininin verir. Yani,

$$\hat{\beta} = (\hat{v}_1, ..., \hat{v}_p) \tag{9.15}$$

olur. H_2 eşbütünleşme hipotezi altında VAR modelinin en çok olabilirlik fonksiyonunun logaritması,

$$\log L = \frac{-n}{2} \log\left|S_{00}\right| - \frac{n}{2} \sum_{i=1}^{r} \log(1 - \hat{\lambda}_i)$$

şeklinde olur. Bir başka ifade ile en çok olabilirlik tahmini, Johansen ve Juselius (1990) tarafından,

$$L_{\max}^{-2/n}\ (H_2) = \left|S_{00}\right| \prod_{i=1}^{r} (1 - \hat{\lambda}_i)\ \prod_{i=1}^{r} (1 - \hat{\lambda}_i) \tag{9.16}$$

şeklinde verilmiştir. H_1 hipotezine göre H_2 hipotezinin olabilirlik oran test istatistiği ise;

$$-2\ln(Q;H_2/H_1) = -n\sum_{i=r+1}^{p}\ln(1-\hat{\lambda}_i) \qquad (9.17)$$

olarak ifade edilir.. Bu istatistiğe literatürde "iz (trace) istatistiği" denmektedir. H_1 hipotezi burada r'ye kadar iken H_2 hipotezi ise p'ye kadardır. Bu da H_1 hipotezinin H_2 hipotezinin özel bir hali olduğunu gösterir. Benzer biçimde $H_2(r+1)$ hipotezine göre $H_2(r)$ hipotezinin testi için olabilirlik oran test istatistiği,

$$-2\ln(Q;r/r+1) = -n\ln(1-\hat{\lambda}_{r+1}) \qquad (9.18)$$

biçiminde olur. Bu istatistiğe de literatürde *en büyük özdeğer (maximal eigenvalue)* denir. Eş.9.17 ve Eş.9.18'deki denklemlerin asimtotik dağılımları Johansen tarafından türetilmiştir. Bu dağılımlar bilinen χ^2 dağılımından farklı olup Dickey-Fuller'in bulmuş olduğu dağılımın çok değişkenli boyutudur. Bu dağılımların elde edilmesi için birtakım stokastik çalışmaların yapılması gerekir.

Johansen ve Juselius (1990), H_2 hipotezindeki β ve α terimlerine lineer kısıt getirmiştir. H, $(p\times s)$, φ ise, $(s \times r)$ boyutlu bir matris olmak üzere, ($r \le s \le p$)

$$\beta = H\varphi$$

şeklinde bir denklem tanımlamıştır. Böylece H_2 hipotezi,

$$H_3 : \Pi = \alpha\varphi' H'$$

veya

$$H_3 : \beta = H\varphi$$

biçiminde tanımlanır. Burada β 'nın en çok olabilirlik tahmin edicisini bulmak için öncelikle,

$$\left|\lambda H'S_{kk}H - H'S_{k0}S_{00}{}^{-1}S_{0k}H\right| = 0$$

denkleminden $\hat{\lambda}_{3.1} > \hat{\lambda}_{3.2} > \ldots > \hat{\lambda}_{3.S}$ olacak biçimde özdeğerler bulunur. Özvektörler $\hat{V}_3 = (\hat{v}_{3.1}, \hat{v}_{3.2}, ..., \hat{v}_{3.S})$ hesaplanıp $\hat{V}_3'(H'S_{kk}H)\hat{V}_3 = I$ formülüyle normalleştirilir.

Burada,

$$\hat{\varphi} = (\hat{v}_{3.1}, \hat{v}_{3.2}, ..., \hat{v}_{3.S})$$

ve β 'nın tahmini,

$$\hat{\beta} = H\hat{\varphi}$$

şeklinde hesaplanır. α 'nın tahmini,

$$\hat{\alpha} = S_{0k}\hat{\beta}(\hat{\beta}'S_{kk}\hat{\beta})^{-1}$$

Λ 'nın tahmini,

$$\hat{\Lambda} = S_{00} - \hat{\alpha}'(\hat{\beta}'S_{kk}\hat{\beta})\hat{\alpha}$$

ve

Γ 'nın tahmini ise,

$$\hat{\Gamma} = M_{01}M_{11}{}^{-1} - \hat{\pi}M_{k1}M_{11}^{-1}$$

ilişkileriyle elde edilir. H_3 hipotezinin en çok olabilirlik tahmini,

$$L_{\max}^{-2/n}(H_3) = |S_{00}| \prod_{i=1}^{r}(1-\hat{\lambda}_{3.i})$$

olarak tanımlanır. Ayrıca olabilirlik oran testi ise,

$$-2\ln(Q\ ; H_3 / H_2) = -n\sum_{i=1}^{r}\ln\left\{(1-\hat{\lambda}_{3.i})/(1-\hat{\lambda}_i)\right\}$$

biçiminde ifade edilir. Bu istatistiğin asimtotik olarak $r(p-s)$ serbestlik dereceli, χ^2 dağılımı olduğu Johansen (1988) tarafından gösterilmiştir. *H*$_2$ hipotezinde α, $\alpha = A\psi$ lineer kısıt ile aynı β tahminindeki gibi gerçekleştirilir.

Yukarıdaki Johansen yöntemi kısaca şöyle özetlemek istersek

1. VAR modelinde ΔX_t bağımlı ve $\Delta X_{t-1}, \Delta X_{t-2}, ..., \Delta X_{t-k+1}$ bağımsız diziler olmak üzere regresyon uygulanır. Bu regresyonda her X_t vektöründe p tane dizi olduğundan p tane regresyonu içerir. Yapılan bu regresyondan *t* zamanında $(p \times 1)$ boyutlu artık vektörü R_{0t} elde edilir. Bu regresyondan sonra yine VAR modelinde X_{t-k} bağımlı ve $\Delta X_{t-1}, \Delta X_{t-2}, ..., \Delta X_{t-k+1}$ bağımsız diziler olmak üzere bir regresyon daha hesaplanır. Bu ikinci regresyondan da $(p \times 1)$ boyutlu R_{kt} artık vektörü elde edilir.

2. Elde edilen bu artık vektörler kullanılarak $(p \times p)$ boyutlu S_{00}, S_{0k}, S_{k0} ve S_{kk} matrisleri Eş.9.8'deki denklemden bulunur.

3. $|\lambda S_{kk} - S_{k0}S'_{00}S_{0k}|$=0 eşitliğinden özdeğerler hesaplanır. Özdeğerler hesaplandıktan sonra $\hat{V} = \left[\hat{v}_1\hat{v}_2...\hat{v}_p\right]$ özdeğerler matrisindeki $\hat{v}_i\ (i = 1,2,...,p)$ özvektör bulunur. Bulunan bu özvektörler $\hat{V}'S'_{kk}\hat{V} = I$ formülü ile normalleştirilir. Daha sonra eşbütünleşme matrisi β'nın rankı $r < p$ olursa ilk r tane özvektöre ($\hat{v}_1, \hat{v}_2, ..., \hat{v}_r$) eşbütünleşme vektörü denir ve bu özvektörler β matrisinin sütunlarını oluşturur.

4. Son olarak her λ_i $(i = 1,2,..., p)$ değeri için olabilirlik oran istatistiği (LR) Eş.9.17'den hesaplanır [5].

Johansen ve Juselius (1990) çalışmasını kısaca özetlersek, λ_i'ler Π matrisinin özdeğerlerini göstermek üzere iki test istatistiği geliştirilmiştir. Bunlar,

$$\lambda_{trace} = -n \sum_{i=r+1}^{p} \ln(1 - \hat{\lambda}_i)$$

ve

$$\lambda_{\max} = -n \ln(1 - \hat{\lambda}_{r+1})$$

istatistik değerleridir.Burada λ_i, Π matrisinin özdeğerleri, n örnek çapıdır. Burada kurulan hipotez yapıları ise;

Çizelge 9.1. Johansen testinin hipotez yapıları

H_O HİPOTEZİ	H_1 HİPOTEZİ
λ_{trace} $r = 0$ $r \leq 1$ $r \leq 2$	$r \geq 1$ $r \geq 2$ $r \geq 3$
$\lambda_{\max}$ $r = 0$ $r = 1$ $r = 2$	$r = 1$ $r = 2$ $r = 3$

şeklinde olur. Bu testlerde kullanılan kritik değerlerde Monte Carlo yaklaşımı kullanılmıştır. $H_0 : r = 0$ eşbütünleşme ilişkisinin olmadığı anlamına gelir. Buna göre alternetif hipotez $H_1 : r \geq 1$ en az bir tane eşbütünleşme ilişkisi var şeklinde kurulur [8].

10. YARI PARAMETRİK YAKLAŞIM

Ekonometrideki birçok tahmin problemleri, *bilinmeyen sonlu boyutlu parametre* ve *bilinmeyen bir fonksiyonu* da içermesi durumunda yarı-parametrik olarak adlandırılır.

Y, gözlenen bağımlı değişken, *X*, gözlenen açıklayıcı değişkenler (satır) vektörü ve *u*, ortalaması 0 olan gözlenemeyen rasgele değişken olmak üzere, bunlardan bir tanesi lineer modellerdeki, β katsayılar vektörünün tahminidir. Model,

$$Y = X\beta + u$$

olur.

u 'nun dağılımının sonlu birçok parametreye göre bilinmesi durumunda, en çok olabilirlik yöntemi, β 'nın asimtotik olarak etkin tahmin edicilerini ve u 'nun dağılımının parametrelerini verir.

u 'nun dağılımının sonlu birçok parametreye bağlı olarak bilinmemesi durumunda, β 'nın tahmin problemi yarı parametriktir. *u*'nun dağılımına bakmaksızın tutarlı olan, en bilinen yarı parametrik tahmin edici EKK tahmin edicisidir.

Ekonometride bundan daha zor bir problem, iki düzeyli lojistik modeller için yarı parametrik yaklaşımlardır. *Y,* 0 ve 1 gibi sadece iki olası değer alan bir rasgele değişken, *X* de açıklayıcı değişken vektörünü göstersin. *X'*e bağlı olarak $Y = 1$ olasılığının tahmin problemi gözönüne alınsın.

Dağılım fonksiyonunu *G,* β, *X*'e uygun sabit parametreler vektörünü göstermek üzere, koşullu olasılık olarak aşağıdaki gibi ifade edilir.

$$P(Y = 1 / X = x) = E(Y / X = x) = G_{\varepsilon / x}\{v(x)\}$$

Bilindiği üzere standart parametrik modellerde, bilinen dağılım fonksiyonu $G_{\varepsilon/x} = G$ ile ε'nun da X ile bağımsız olarak dağıldığı ve indeks $v(.)$'nin aşağıdaki gibi basit formda olduğu varsayılmaktadır.

$$v(x) = \beta_0 + X'\beta$$

İki düzeyli probit modelde standart normal dağılım olarak alınan G'nin önceden bilindiği varsayılırsa, tek problem β'nın tahminidir. Bu tahminler en çok olabilirlik yöntemi ile elde edilebilir.

Yarı parametrik modellerin dezavantajları şunlardır,

i. Bu modeller etkin tahmin probleminin boyutunu azaltmaktadır.
ii. Bu, *E(Y | X=x)*'in biçimi ile ilgili varsayımlar yapılarak başarılır.

Sonuç olarak elde edilen modeller, daha güçlü varsayımlar ve daha büyük belirleme hatası riski pahasına "Boyutluluk Sorunu" probleminden uzaklaşmış olurlar[4].

Yarı parametrik modellerin avantajları:

i. Yorumlanması daha kolay tahminler içerir.
ii. *X*'in çok boyutlu olması durumunda daha yüksek tahmin doğruluğu sağlanır.

10.1. Kısmi Lineer Model (Partially Linear Model: PLM)

Analitik ya da ekonomik teoriden kaynaklanan nedenlerden dolayı ele alınan modelin sadece bir kısmının lineer olarak modellenmesi gerekebilir.

---

[4] Tamamen parametrik olmayan modellere göre daha fazla varsayım yapılırken, parametrik modellere göre daha az varsayım yapılmaktadır.

Daha ayrıntılı ifade etmek gerekirse, d-boyutlu değişken vektörü, $U = (U_1,...,U_P)'$ ve $T = (T_1,...,T_q)'$ olarak ayrılsın. Y'nin $X = (U,T)$ üzerine regresyonu, $m(.)$, T vektörünün bilinmeyen çok değişkenli fonksiyonu iken, aşağıdaki biçimde verilir.

$$E(Y \mid U,T) = U'\beta + m(T)$$

Böylece kısmi lineer model, tamamen parametrik kısım $U'\beta$ ile tamamen parametrik olmayan kısım ise, $m(T)$'nin bir toplamı olarak ifade edilir. Bundan dolayıdır ki, β ve $m(.)$'nin tahmini hem parametrik hem de parametre dışı regresyon tekniklerinin bir kombinasyonunu içerir.

11. PARAMETRE DIŞI TAHMİN YÖNTEMLERİNDEN KERNEL (ÇEKİRDEK, AĞIRLIK) TAHMİNİ

$X_1, X_2, \ldots, X_n$, sürekli, tek değişkenli bir dağılımdan alınan bağımsız, aynı dağılımlı gözlemlerin bir örneği olsun. Tahmin edilecek fonksiyon, bu gözlemlerin alındığı olasılık yoğunluk fonksiyonu f'dir. Rosenblatt (1956) tahmin edicilerinde kullanılan w(u) ağırlık fonksiyonu yerine,

$$\int_{\infty}^{\infty} K(x)dx = 1$$

koşulunu sağlayan bir K, kernel fonksiyonu olarak olasılık yoğunluk fonksiyonunun kernel(çekirdek) tahmin edicisinin aşağıdaki gibi ifade edilebileceğini belirtilmiştir:

$$\hat{f}(x,h) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right)$$

Olasılık yoğunluk fonksiyonu ve regresyon fonksiyonları tahminlerinde çok sık kullanılan parametre dışı tahmin edicilerden birisi kernel tahmin edicilerdir.

Kernel değeri $K_h(u) = h^{-1}K\left(\frac{u}{h}\right)$ olarak tanımlanıp tahmin edici,

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

biçiminde de yazılabilir. Yukarıda verilen tahmin ediciye *kernel tahmin edicisi* adı verilir. Burada K fonksiyonuna kernel fonksiyonu, h'ya da *pencere genişliği*, *düzleştirme parametresi* ya da *bant genişliği* adı verilir. K kernel fonksiyonu genellikle sıfır etrafında simetrik tek tepeli olasılık yoğunluk fonksiyonu olarak seçilir.

K kernel fonksiyonu için birçok fonksiyon benimsenebilir. EK-1'de bu fonksiyonlardan bazıları verilmiştir. K kernel fonksiyonunun, geniş seçilebilme şansına sahip olmanın getirebileceği yarar, bazı K fonksiyonlarında daha güvenilir tahminler elde edebilecek daha büyük düzleştirme etkisine sahip olmalarıdır.

x noktasındaki kernel tahmin edicisinin değeri şu şekilde elde edilir: x gözlem değeri ortada olmak üzere kernel fonksiyonu x gözleminin üzerine yerleştirilir. Veriler sınıf genişliği h ve x gözlem değeri ortada olmak üzere sınıflara ayrılır. Bu kernel fonksiyonundan yararlanılarak her bir sınıftaki olasılıklar hesaplanır. Bu olasılıklar her bir gözleme, içinde bulunduğu sınıfa göre ağırlık olarak verilerek ağırlıklı ortalama elde edilir. Bu ağırlıklı ortalama x gözlem değerinin kernel tahmin edicisidir. Başka bir ifade ile kernel tahmin edicisi, komşu gözlemlere de belli ağırlıklar vererek işin içine katmaktadır. Kernel fonksiyonu ağırlıkların hesaplanmasında, h bant genişliği ise, ölçekleme etkeni rolünü üslenmektedir. h bant genişliği ve K kernel fonksiyonunun farklı seçimleri ile elde edilen yoğunluk fonksiyonu tahmin edicileri de değişmektedir. h bant genişliğinin seçimi, çok önemlidir. K kernel fonksiyonu bir olasılık yoğunluk fonksiyonu olduğundan $\hat{f}$ de olasılık yoğunluk fonksiyonu olacaktır. Ayrıca $\hat{f}$, K kernel fonksiyonunun süreklilik ve türevlenebilme özelliklerini de taşıyacaktır. $\hat{f}$ kernel tahmin edicisi, K kernel fonksiyonuna ve h bant genişliğine bağlı olmasının yanı sıra gözlem değerlerine de bağlıdır. Her gözlem için $\hat{f}$, bir rasgele değişken tanımlar [9].

11.1. Regresyon Fonksiyonunun Kernel Tahmini

n gözlemli bir örnekde, $x_1, x_2, ..., x_n$ değerlerine karşılık gelen $y_1, y_2, ..., y_n$ gözlem değerleri,

$$y_i = f(x_i) + \varepsilon_i, \quad (i = 1, ..., n)$$

model denklemi ile tanımlansın. Burada $\{x_i\}$'ler bilinen değerlere sahip ve $0 \le x_1 \le x_2 \le ... \le x_n \le 1$ olmak üzere, $f(x)$, $0 \le x \le 1$ için tanımlanmış bilinmeyen regresyon fonksiyonudur. Ayrıca ε_i'ler ilişkisiz rasgele değişkenler olmak üzere ,

$$E(\varepsilon_i) = 0,\ Var(\varepsilon_i) = \sigma_i^2 \qquad (i = 1,...,n)$$

özelliklerine sahip olsun.

Klasik regresyon problemlerinde, $f(x)$ özel olarak belirtilmiş bir fonksiyon olup, belli varsayımlar altında parametre tahminleri EKK tahmin yöntemiyle tahmin edilmektedir. Modelin uygunluğu artıkların normal dağıldığı varsayımı altında test edilmektedir. Burada regresyon fonksiyonunun tahmini kernel fonksiyonları kullanılarak yapılacaktır. Gözlem değerleri, Rozenblatt tarafından önerilen bir *kernel* lar sınıfına dayanan ağırlıklara sahiptirler [10]. Bu ağırlıklar parametre dışı yoğunluk tahmininde kullanılanlara benzemektedir. Şimdi x_i'lerin eşit aralıklı olması varsayımı altında tahminler ele alalım. Birinci farkları;

$$x_i - x_{i-1} = \delta \qquad (i = 2,...,n)$$

ifadesi kullanılarak tanımlansın. Aynı yoğunluk fonksiyonunda olduğu gibi regresyon fonksiyonunun tahmini için de kernel fonksiyonları kullanılarak $\hat{f}_n^{\,1}(x)$ ve $\hat{f}^{\,2}{}_n(x)$ aşağıdaki biçimde yazılabilir [11].

$$K^{(1)}(z) = \begin{cases} 1 & -1 \le z \le 0 \\ 0 & d.h \end{cases} \tag{11.1}$$

Bu kernel fonksiyonu kullanılarak aşağıdaki regresyon fonksiyonu tahmini,

$$\hat{f}^{\,1}{}_n(x) = \delta \sum_{i=1}^{n} y_i \left\{ \delta^{-1} K^{(1)}\left(\frac{x - x_i}{\delta} \right) \right\}$$

şeklinde yazılır. Yine aynı şekilde,

$$K^{(2)}(z) = \begin{cases} 1-|z| & 0 \le |z| \le 1 \\ 0 & d.h \end{cases} \tag{11.2}$$

fonksiyonu kullanılarak da regresyon fonksiyonu tahmini,

$$\hat{f}^2{}_n(x) = \delta \sum_{i=1}^{n} y_i \left\{ \delta^{-1} K^{(2)}\left(\frac{x - x_i}{\delta} \right) \right\}$$

olarak bulunur. Eş.11.1'deki ifade "*dikdörtgensel kernel fonksiyonu*", Eş.11.2'deki ifade ise "*üçgensel kernel fonksiyonu*"dur. $f(x)$ 'in genel bir tahmin sınıfının,

$$\hat{f}(x) = \delta \sum_{i=1}^{n} y_i K_h(x - x_i)$$

biçiminde olacağı düşünülür. Burada her *h* için,

$$K_h(x) = h^{-1} K_0(x/h)$$

olup, K_0 ağırlık fonksiyonunun genel bir tipidir. Sonuç olarak ifade edilirse, $h = \delta$ olmak üzere tahmin ediciler,

$\hat{f}^1{}_n(x)$ için $K_0(x) = K^{(1)}(u)$

$\hat{f}^2{}_n(x)$ için $K_0(x) = K^{(2)}(u)$

şeklinde tanımlanır [11].

11.2. Optimum Bant Genişliği ve Kernel (Çekirdek) Fonksiyonunun Seçimi

h bant genişliğinin seçimi çok önemlidir. h'nin çok küçük seçilmesi durumunda, her noktadaki kernel tahmin edicisini elde etmek için uygulanan ağırlık alma süreci daha az sayıda gözleme dayanacağından, f 'in tahmini kaba bir tahmin olacaktır. Bu tahminlere az düzleştirilmiş tahmin denir. h çok büyük ise ağırlık ortalama alma süreci çok sayıda gözleme dayanacağından elde edilen tahmin değeri gerçek değerden uzaklaşabilir. Bu tahmine de aşırı düzleştirilmiş tahmin denir. Tahmin performansı açısından çok önemli olan h bant genişliğinin seçimi için çalışmalar devam etmekle birlikte herkes tarafından benimsenen genel bir yöntem henüz bulunamamıştır [9].

K kernel fonksiyonunu simetrik bir olasılık yoğunluk fonksiyonu olarak seçmek, sağında ve solunda gözlemlere, verilen noktaya olan uzaklıklarına göre eşit ağırlıklar vermek mantığa en uygun olanıdır. K kernel fonksiyonunun, f olasılık yoğunluk fonksiyonundan farkı kullanıcının kontrolü altında olmasıdır.

12. TEKİL DEĞER AYRIŞIMI

$\Pi_{n\times p}$, π_{ij} ve $rank((\Pi_{n\times p}) = r \le p$) olsun.

$$\pi_{ij} = \theta_1 u_{1i} v_{1j} + \theta_2 u_{2i} v_{2j} + ... + \theta_r u_{ri} v_{rj}$$

veya;

$$\pi_{ij} = \sum_{k=1}^{r} \theta_k u_{ki} v_{kj}$$

şeklinde ifade edilsin. Burada $\theta_1 \ge \theta_2 \ge ... \ge \theta_r$, u ve v ise, r satırdan oluşan vektörlerdir. Bu ifade *tekil değer ayrışımı* olarak bilinir. (Π matrisinin tekil değer ayrışımı)

Ayrıca, u ve v ’nin her biri ortagonaldir.

$$\sum_i u_{ki}^2 = \sum_j v_{kj}^2 = 1 \quad (tüm\ \ k\ \ için)$$

Her bir vektör uzunluğu ℓ ’dir. En genel hali ile tekil değer ayrışımı,

$$\Pi_{nxp} = U_{nxr} \theta_{r\times r} V'_{rxp}$$

biçiminde gösterilir.

Burada, θ köşegen matris tanımlar ve tüm θ ’lar pozitiftir.

$U : u$ vektörlerinden, $V : v$ vektörlerinden oluşur. Bu durumda,

$U'U = \mathrm{I}_{r \times r}$
$VV' = \mathrm{I}_{r \times r}$

olur. θ_k , $\Pi'\Pi$ ve $\Pi\Pi'$ matrislerinin özdeğerlerinin kareköküdür.

U : $\Pi\Pi'$ matrisinin,

V : $\Pi'\Pi$ matrisinin,

özvektörleridir [12].

13. KISMİ LİNEER MODELLERDE KERNEL YÖNTEMİNİN KULLANILMASI VE KISMİ LİNEER TEKİL DEĞER TESTİ (PLS)

13.1. Toplanabilir Kısmi Lineer Modellerin Tahmini

Bu kısımda PLS testine geçmeden önce toplanabilir kısmi lineer model yapısı açıklanacaktır.

$$Y_i = \beta_0 + X_i'\beta + g_1(Z_{1i}) + g_2(Z_{2i}) + ... + g_p(Z_{pi}) + U_i \tag{13.1}$$

modeli toplanabilir kısmi lineer model yapısını tanımlar. Burada X_i'lerin her biri $(q \times 1)$ boyutlu rasgele değişkenler vektörüdür ve $\beta = (\beta_1, \beta_2, ..., \beta_q)'$ ise $(q \times 1)$ boyutlu bilinmeyen parametreler vektörüdür. Aynı zamanda modelde yer alan β_0, modelin sabit parametresidir. $Z_{\alpha i}$'ler ise tek değişkenli sürekli rasgele değişkenler olmakla birlikte $g_\alpha(.)$, $\alpha = 1,...,p$ tane bilinmeyen düzleştirme fonksiyonudur. Burada gözlenen $\{Y_i, X_i', Z_{1i}, ..., Z_{pi}\}$, $i = 1,...,n$ bağımsız, aynı dağılımlı rasgele değişkenlerdir. Tüm α'lar için $E(g_\alpha(Z_\alpha)) = 0$ dır.

$Z_{\underline{\alpha} i} = (Z_{1i}, ..., Z_{\alpha-1,i}, Z_{\alpha+1,i}, ..., Z_{pi})$ vektörü, $(Z_{1i}, Z_{2i}, ..., Z_{pi})$'den $Z_{\alpha i}$'nin kaldırılmış halidir. $G_{\underline{\alpha}}(z_\alpha) = g_1(z_1) + ... + g_{\alpha-1}(z_{\alpha-1}), g_{\alpha+1}(z_{\alpha+1}) + ... + g_p(z_p)$ de olarak tanımlanır ve Eş.13.1'deki model şu şekilde ifade edilir.

$$Y_i = \beta_0 + X_i'\beta + g_\alpha(Z_{\alpha i}) + G_{\underline{\alpha}}(Z_{\underline{\alpha i}}) + U_i$$

Tahmin yönteminde,

$$a(z_\alpha, z_{\underline{\alpha}}, x) = E(Y_i \mid Z_{\alpha i} = z_\alpha, Z_{\underline{\alpha i}} = z_{\underline{\alpha}}, X_i = x)$$

eşitliği kullanılır [13].

13.2. Eşbütünleşme Testlerinden Kısmi Lineer Tekil Değer Testi (PLS)

Bu kısımda sabit ortak bir değişkenin parametre dışı olarak girdiği eşbütünleşmeye ait kısmi lineer modellerle ilgilenilecektir. Johansen ve Juselius (1992) çalışmasında eşbütünleşme denkleminde durağan olmayan değişkenlere yine lineer olarak ve diğer değişkenlerle ilişkili yeni durağan bir değişkeni modele katıp testin gücünü artırmaya yönelik bir çalışma yapmışlardır. Bu çalışmada konu edilecek Ted Juhl ve Zhijie Xiao (2005) çalışmasının, Johansen ve Juselius (1992) çalışmasından farkı, eklenen ortak değişkenin lineer olmayan bir formda (nonlinear) ve durağanlık şartını kaldırarak modeli kısmi lineer model olarak kurmalarıdır. Modelde Π 'nin özdeğerlerinin tekil değer ayrışımıyla tahmin edilmesine dayanır. Ted Juhl ve Zhijie Xiao (2005) bu testi *Kısmi Lineer Tekil Değer Testi(PLS)* olarak adlandırmışlardır.

$$\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \mu_0 + \mu_{1t} + V_t \qquad (13.2)$$

biçiminde tanımlanan model p boyutlu *vektörel otoregresif model* (VAR)dir. Burada, V_t ortalaması sıfır, aynı dağılımlı rasgele değişkendir. V_t'yi tanımlayan q sayıda ortak değişken Z_t olduğunu varsayalım. V_t, Z_t'lerin bir fonksiyonudur ve şu şekilde ifade edilir:

$$V_t = \lfloor g(Z_t) - \mu_g \rfloor + \varepsilon_t \qquad g : \Re^q \longrightarrow \Re^r$$

$$\mu_g = E(g(Z_t)) = 0$$

Bu açıklamaların ışığı altında yeni model,

$$\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \mu_0 + \mu_{1t} + g(Z_t) + \varepsilon_t \qquad (13.3)$$

olarak düzenlenir [14].

Burada X_{t-1} 'li kısım modelin lineer kısmını, $g(Z_t)$ 'li kısım ise lineer olmayan kısmı ifade eder. Böylece model, yarı parametrik bir model olmuş olur. Modelin özellikleri şunlardır,

- Model asimetriktir.
- Model lineer kısım ve lineer olmayan kısım olarak iki kısıma ayrılır.
- Modele sabit bir değişken Z_t eklenmesi ile genişletilir.
- Amaç artıkların varyansını azalmaktır.
- Z_t durağan olmayan bir dizi olabilir.

Burada yarı parametrik modelin tahmin problemi ile karşı karşıya kalınır. Kısmi lineer modeller için parametrik kısmın parametre tahminleri, parametrik olarak tahmin edilirken, lineer olmayan kısmın parametre tahminleri, parametre dışı regresyon yöntemleriyle elde edilir.

Eğer X_t, vektör zaman dizileri r kadar sırada eşbütünleşirse, $rank(\Pi) = r$ olur ve $\Pi = \alpha\beta'$ olur. Burada ki α ve β, $p \times r$ boyutlu r ranklı matrislerdir.

Eş.13.3'deki model aşağıdaki gibi yazılabilir.

$$\Delta X_t = \Pi X_{t-1} + \Gamma_i Y_t + \mu_0 + \mu_{1t} + g(Z_t) + \varepsilon_t$$

Burada,

$$\Gamma = (\Gamma_1 \vdots \Gamma_2 \vdots \ldots \vdots \Gamma_{k-1}), \qquad Y_t = (\Delta X'_{t-1} \vdots \Delta X'_{t-2} \vdots \ldots \vdots \Delta X'_{t-k-1})'$$

olmak üzere $g(Z_t)$ 'lerin varlığı hesaplanarak Π 'nin tahmini yapılır [14].

$\mu_0 = \mu_1 = 0$ olduğu koşulu altında $E(\varepsilon_t \mid Z_t) = 0$ olarak alınır. Eş.13.3'de verilen modelin her iki tarafından $E(\Delta X_t \mid Z_t)$ çıkarılırsa model,

$$\Delta X_t - E(\Delta X_t \mid Z_t) = \Pi(X_t - E(\Delta X_t \mid Z_t)) + \Gamma(Y_t - E(\Delta X_t \mid Z_t)) + \varepsilon_t$$

şeklinde yazılarak lineer hale gelir.

PLS testine geçmeden önce kernel değerlerinin hesaplamasına gidilerek kernel fonksiyonu elde edilir. Kernel fonksiyonun,

$$K(u) = \prod_{j=1}^{q} k(u_j) \qquad u : q \text{ boyutlu,}$$

ve

$$K_{ts} = K\left(\frac{(z_t - z_s)}{h}\right)$$

formülünden hesaplanır. Burada h, bant genişliğini ifade eder. Buradan elde edilen olasılık yoğunluk kernel tahmin fonksiyonu ise,

$$\hat{f}_t = \frac{1}{nh^q} \sum_{s=1, s\neq t}^{n} K_{ts}$$

biçiminde tahmin edilir. Bu fonksiyonda n, örnek hacmi, $\hat{f}_t$, olasılık yoğunluk kernel tahmin edicisidir ve bu fonksiyon parametre dışı regresyon fonksiyonu olarak tanımlanır. Bu kernel tahmin edicisi kullanılarak aşağıdaki tahminler bulunur [14].

$$\hat{X}_{t-1} = \frac{1}{nh^q} \sum_{s=1, s\neq t}^{n} K_{ts} X_{s-1} / \hat{f}_t$$

$$\Delta\hat{X}_t = \frac{1}{nh^q} \sum_{s=1, s\neq t}^{n} K_{ts} \Delta X_s / \hat{f}_t$$

$$\hat{Y}_t = \frac{1}{nh^q} \sum_{s=1, s\neq t}^{n} K_{ts} Y_s / \hat{f}_t$$

Yukarıdaki denklemlerden elde edilen tahminlerle $\Delta X_t - \Delta\hat{X}_t = f(X_{t-1} - \hat{X}_{t-1}, Y_t - \hat{Y}_t)$ EKK yöntemi regresyonu tahmin edilir. Burada böylelikle tahminin tahmini yapılmış olur.

Bu tahmin değerler kullanılarak $t'nci$ satır değerleri;

$$\Delta\widetilde{X} \rightarrow (\Delta X_t - \Delta\hat{X}_t)'\hat{f}_t$$
$$\widetilde{X} \rightarrow (X_{t-1} - \hat{X}_{t-1})'\hat{f}_t$$
$$\widetilde{Y} \rightarrow (Y_t - \hat{Y}_t)'\hat{f}_t$$

bulunur. Burada $\Delta\widetilde{X}$, $(n \times p)$ boyutlu bir matristir. Elde edilen bu değerlerden izdüşüm matrisi,

$$P = \widetilde{Y}(\widetilde{Y}'\hat{Y})^{-1}\widetilde{Y}'$$

olarak hesaplanır. Buradaki izdüşüm matrisinin boyutu, $(n \times n)$ dir.

Π 'nin tahmini,

$$\hat{\Pi} = \Delta\widetilde{X}'(\mathrm{I} - P)\widetilde{X}(\widetilde{X}'(\mathrm{I} - P)\widetilde{X})^{-1}$$

denkleminden bulunur. Burada I birim matristir. $\hat{\Pi}$ 'nin boyutu ise, $(p \times p)$ dir. Bu tahminden faydalanılarak $\widetilde{\varepsilon}$ 'nin $t'nci$ satır değerleri;

$$\widetilde{\varepsilon}_{n\times p} \rightarrow \left[\left(\Delta X_t - \Delta\hat{X}_t\right) - \hat{\Pi}\left(X_{t-1} - \hat{X}_{t-1}\right)\right]'\hat{f}_t$$

formülünden bulunur.

Bu tahmin standartlaştırılır ise;

$$\widetilde{\Pi} = (n^{-1}\widetilde{\varepsilon}'(\mathrm{I}-P)\widetilde{\varepsilon})^{-1/2}\hat{\Pi}(n^{-1}\widetilde{X}'(\mathrm{I}-P)\widetilde{X})^{1/2}$$

elde edilir. Burada $\widetilde{\Pi}$ 'nin boyutu ise; $(p \times p)$ dir.

$\widetilde{\Pi}'\widetilde{\Pi}$ matrisinden tekil değer ayrışımı yöntemi kullanılarak, $\widetilde{\lambda}_1 > \widetilde{\lambda}_2 > \ldots > \widetilde{\lambda}_p$ olmak üzere $\widetilde{\lambda}_i$ değerleri bulunur. Bu değerlerin hesaplamaları yapıldıktan sonra elde edilen değerler test istatistiği,

$$PLS = n \sum_{i=r+1}^{p} \widetilde{\lambda}_i$$

denkleminde yerine konularak hesaplanır [14][5].

Burada asimtotik değerlerin tahmininde Gamma yaklaşımı kullanılmıştır.

Çizelge 13.1. PLS test istatistik yapısı

r	*p-r*	*Test istatistiği PLS*
0	5	$n \times (\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5)$
1	4	$n \times (\lambda_2 + \lambda_3 + \lambda_4 + \lambda_5)$
2	3	$n \times (\lambda_3 + \lambda_4 + \lambda_5)$
3	2	$n \times (\lambda_4 + \lambda_5)$
4	1	$n \times (\lambda_5)$

Çizelge 13.1'den faydalanılarak H_0 ve H_1 hipotezleri şöyle kurulur:

[5] Şayet $g(Z_t)$ mevcut değilse ve $\widetilde{\lambda}_i$ yerine $\lambda_i = \widetilde{\lambda}_i / (1 - \widetilde{\lambda}_i)$ yazılabilirse PLS testi, Johansen testine benzerlik gösterir.

$H_0 : r = 0$ (hiç eşbütünleşme denklemi yok)

$H_1 : r \geq 1$ (bir tane eşbütünleşme denklemi var)

H_0 red edilir ise, bir sonraki hipoteze geçilir.

$H_0 : r = 1$ (bir tane eşbütünleşme denklemi var)

$H_1 : r \geq 2$ (iki tane eşbütünleşme denklemi var)

Bu hipotezler $H_0 : r = p$ değerinin testine kadar devam eder. H_0'ın red edilmediği denklem sayısı eşbütünleşme denklem sayısını verir [15].

14. VERİ ANALİZLERİ

Bu bölümde, Türkiye'deki eğitim ve iktisadi büyüme arasındaki ilişkiyi ortaya koyan model üzerinde, Engle ve Granger tarafından geliştirilen ve zaman dizileri modellerinde uzun dönemli ilişkilerin varlığını araştırmaya yönelik olarak kullanılan eşbütünleşme analizleri uygulanmıştır. Çalışmada kısmi lineer model durumunda eşbütünleşme testi irdelenmiş ve bu anlamda kısmi lineer modellerde tekil değer ayrışımına dayalı Kısmi Lineer Tekil Değer Testi (PLS) testi ile Johansen en çok olabilirlik testi (LR) gerçekleştirilecektir.

Dickey-Fuller çalışmasında, zaman dizilerinin birçoğunun varyans ve ortak varyanslarının zaman içinde artma eğiliminde olduğunu vurgulamaktadır. Bir de değişkenler arasında eşbütünleşme ilişkisini araştırabilmek için koşul, değişkenlerin aynı dereceden durağan olmasıdır. Bu durumu ortaya koyabilmek için, öncelikle durağanlık testleri yapılmıştır. Bu çalışmada durağanlık kavramı, Genişletilmiş Dickey-Fuller (GDF) testi ile incelenmiştir.

Modeldeki değişkenlerin aynı dereceden durağan olduklarına karar verilirse, eşbütünleşme testi ile değişkenler arasında uzun dönemli ilişkinin varlığı araştırılabilecektir [16].

Verilerin sayısal analizlerine geçmeden önce eğitim ve iktisadi büyüme arasındaki ilişkiyi, bu ilişkiyi ortaya koyabilecek değişkenleri ve ortaya konulan modelin tanımı, içeriğini ve amaçlarını açıklayalım.

14.1. Eğitim ve İktisadi Büyüme Arasındaki İlişkinin Tanımı

Ekonomik büyümenin temelinde fiziksel yatırımlardan ziyade teknolojik yatırımlar ve ARGE faaliyetleri için ayrılan kaynaklar gibi insan kaynaklarının geliştirilmesine yönelik yatırımlar olduğu pek çok çalışma sonucu ortaya konmuştur. Birçok iktisatçıya göre ekonomik kalkınmanın temek noktası eğitimdir. Birçok ülkenin eğitim harcamalarında yaptıkları artışla, kaydettikleri ekonomik gelişme konusunda sıkı ilişkiler olduğu bilimsel çalışmalarla desteklenmiştir.

Eğitim türü ile ekonomik yapı arasındaki ilişkiyi ele alan araştırmalar; kişi başına düşen gelir ile eğitim türü arasında güçlü bir ilişki olduğunu gösterir. Genellikle bu çalışmalar sonucunda gelir düzeyinin ilk ve orta öğretim ile ilişkisinde belirli bir eşbütünleşim saptanamasa da yüksek öğrenim ile oldukça güçlü ilişkisi olduğu saptanmıştır [16].

14.2. Eğitim ve İktisadi Büyümeye İlişkin Model

Çalışmamızda Türkiye'de eğitim ve iktisadi büyüme arasındaki ilişkiyi açıklamak üzere ele aldığımız değişkenler, ekonomik büyümeyi temsil etmek üzere kişi başına gayri safi milli hasıla "*kbgsmh*", eğitimi temsil etmek üzere, ilköğretim okullaşma oranı "*ilk*", ortaöğretim okullaşma oranı "*orta*", yükseköğretim okullaşma oranı "*yüksek*" ve eşbütünleşme analizine ortak değişken olarak katılacak değişken de eğitim bütçesinin GSMH içindeki oranı ,"*bütçe*" şeklinde ele alınmıştır. Bu durumda modelimiz,

$$kbgsmh_t = \beta_0 + \beta_1 (ilk)_t + \beta_2 (orta)_t + \beta_3 (yüksek)_t + u_t$$

şeklinde olur.

14.3. Türk Milli Eğitim Sistemi

Türk milli eğitimi 1973 yılında yürürlüğe konulan 1739 sayılı Milli Eğitim Temel Kanunu ile belirlendiği üzere örgün eğitim ve yaygın eğitim olmak üzere iki ana bölümden meydana gelmektedir.

14.3.1. Örgün eğitim

Örgün eğitim, okul öncesi eğitim, ilköğretim, ortaöğretim ve yükseköğretimi kapsamaktadır. Örgün öğretim belirli yaş grubundaki ve aynı seviyedeki bireylere,amaca göre hazırlanmış programlarla okul çatısı altında yapılan düzenli eğitimdir.

Okulöncesi eğitim

Okulöncesi eğitim, isteğe bağlı olarak zorunlu ilköğretim çağına gelmemiş 3-5 yaş grubundaki çocukların eğitimini kapsar. Okulöncesi eğitim kurumları bağımsız anaokulları olarak kurulabildiği gibi, gerekli görülen yerlerde ilköğretim kurumlarına bağlı anasınıflar olarak da kurulabilirler. Okulöncesi eğitimin amacı, çocukların bedensel, zihinsel, duygusal gelişimini sağlamak ve iyi alışkanlıklar kazandırarak onları ilköğretime hazırlanmalarını ve ayrıca Türkçe'nin doğru ve güzel kullanılmasını sağlamaktır.

İlköğretim

İlköğretim, 6-13 yaş grubundaki çocukların eğitim ve öğretimini kapsar. Resmi ve özel eğitim veren ilköğretim okulları ile açık ilköğretimi içerir. İlköğretimin amacı, her Türk çocuğunun iyi birer yurttaş olabilmesi için, gerekli temel bilgi, beceri, davranış ve alışkanlıklar kazandırmak, milli ahlak anlayışına uygun olarak yetiştirilerek ilgi ve becerileri doğrultusunda bir üst öğrenime hazırlamaya çalışmaktır. Türkiye'de bu eğitim sekiz yıllık ve zorunludur. Resmi eğitim veren okullar tarafından parasız sağlanır.

Ortaöğretim

Ortaöğretim, ilköğretime dayalı en az üç yıllık genel, mesleki ve teknik öğretim kurumlarının tümünü kapsar. Ortaöğretimin amacı, ortak bir genel kültür vermek, birey ve toplum sorunlarını tanıtmak, ülkenin sosyo-ekonomik ve kültürel gelişimine katkıda bulunacak bilinci kazanmış bireyler yetiştirerek ilgi ve becerileri doğrultusunda yükseköğretime, hayata veya iş alanlarına hazırlamaktır.

Yükseköğretim

Yükseköğretim, ortaöğretime dayalı, en az iki yıllık yüksek öğretim veren resmi ve özel (vakıf üniversiteleri) eğitim kurumlarının tümünü kapsar. Yükseköğretim önlisans, lisans (4-5-6 yıllık fakülteler, açıköğretim fakültelerini kapsar), yüksek lisans ve doktora olmak üzere dört bölüm olarak hesaplanır. Yükseköğrenimin amacı, ülkenin bilim politikasına, toplumun yüksek düzeyde ve çeşitli kademelerinde insan gücü gereksinimlerine göre öğrencileri ilgi ve becerileri doğrultusunda yetiştirmek, bilimsel alanlarda araştırma yapmak, araştırma ve inceleme sonuçlarını gösteren ve bilim tekniğin ilerlemesini sağlayan her türlü yayını yapmak, hükümet tarafından istenecek incelemeler ve araştırmalar yaparak düşüncelerini bildirmek, Türk toplumunun genel seviyesini yükseltici ve yararına olan bilimsel çalışmaları kamuoyuna sözle veya yazı ile bildirmek ve yaygın eğitim hizmetinde bulunmaktır [17].

14.3.2. Yaygın eğitim

Yaygın eğitimin amacı, milli eğitimin genel amaçlarına ve temel ilkelerine uygun olarak, örgün eğitim sistemine hiç girmemiş olan veya herhangi bir kademesinde bulunan ya da bu kademelerden çıkmış yurttaşlara örgün eğitimin yanında veya dışında;

- Okuma yazma öğretmek, eksik eğitimlerini tamamlamak için sürekli eğitim imkanı sağlamak,
- Bilimsel, teknolojik, ekonomik, sosyal ve kültürel gelişimi uyum sağlamalarına yardımcı olmak,
- Milli kültür değerlerini koruyucu, geliştirici, tanıtıcı nitelikte eğitimler vermek,
- Toplu yaşama, dayanışma, birlikte çalışma, örgütlenme anlayış ve alışkanlıklar kazandırma,

- Ekonomik gelişme doğrultusunda ve istihdam politikasına uygun olarak meslek edinmelerini sağlayıcı olanaklar hazırlamak,
- Beslenme ve sağlıklı yaşam hakkında bilinçlendirme çalışmaları yapmak,
- Çeşitli mesleklerde çalışanlara, gelişmelerini sağlayıcı bilgi ve beceriler kazandırmak,
- Boş zamanlarını yararlı bir biçimde değerlendirme ve kullanma alışkanlıkları kazandırmaktadır.

14.4. Okullaşma Oranı

14.4.1. Brüt okullaşma oranı

İlgili öğrenim türündeki tüm öğrencilerin, ait olduğu öğrenim türündeki teorik yaş grubunda bulunan toplam nüfusa bölünmesi ile elde edilir.

A: Toplam öğrenci sayısı
B: Teorik yaş grubundaki toplam nüfus

olmak üzere,

$$\text{Brüt okullaşma oranı} = \frac{A}{B}$$

olur.

14.4.2. Net okullaşma oranı

İlgili öğrenim türündeki teorik yaş grubunda bulunan öğrencilerin, ait olduğu öğrenim türündeki teorik yaş grubunda bulunan toplam nüfusa bölünmesi ile elde edilir.

A: Teorik yaş grubundaki öğrenci sayısı

B: Teorik yaş grubundaki toplam nüfus

olmak üzere,

$$\text{Net okullaşma oranı} = \frac{A}{B}$$

olur. Öğrencilerin bitirdiği yaş baz alınarak, ilköğretimde teorik yaş 6-13, ortaöğretimde teorik yaş 14-16, yüksek öğretimde teorik yaş 17-21 olarak kabul edilmiştir [18].

Bizim buradaki çalışmamızda net okullaşma oranlarını kullanılacaktır.

14.5. Kişi Başına Gayri Safi Milli Hasıla

Bir ekonomide yerleşik olan üretici birimlerin belli bir dönemde yurtiçi faaliyetleri sonucu yaratmış oldukları tüm mal ve hizmetlerin üretim değerleri toplamından bu mal ve hizmetlerin üretimde kullanılan girdiler toplamının düşülmesi sonucu, Gayri Safi Yurtiçi Hasıla (*GSYİH*) değerine , dış alım net faktör gelirleri değerinin eklenmesiyle Gayri Safi Milli Hasıla (*GSMH*) değeri elde edilir. Kişi Başına Gayri Safi Milli Hasıla (*KBGSMH*) değeri;

$$KBGSMH = \frac{GSMH}{nüfus}$$

olarak elde edilir.

14.6. Eğitim Bütçesinin GSMH İçindeki Oranı

Ülkelerin bütçesi içerisindeki eğitime ayrılan paylar *GSMH* içinde belli bir oranı içermektedir. Türkiye büyüme sürecini devam ettirebilmek için yapısal reformlar vasıtasıyla eğitim alanında ve diğer sosyal nitelikli alanlarda yatırım ortalamasını iyileştirmek zorundadır. Bu eğitim ve sosyal nitelikli harcamaların *GSMH* içindeki payı artırılarak, toplumun yaşam kalitesi yükseltilebilir.

14.7. Analiz

Bu bölümde öncelikle değişkenlerin durağanlık durumlarını inceleyen birim kök testleri gerçekleştirilecek daha sonra da Johansen ve Kısmi Lineer Tekil Değer eşbütünleşme testleri uygulanacaktır.

14.7.1. Durağanlık testi

EK-2'de verilen uygulama verileri 1983 yılından başlayıp 2005 yılına kadar devam eden 23 tane veri içeren zaman dizileri bulunmaktadır. Her bir değişkene ilişkin birim kök testi sonuçları Çizelge 14.1'de izlenebilir.

Çizelge 14.1. GDF birim kök testi

Açıklayıcı değişken	Test istatistik değeri	Kritik değeri %5	Birinci dereceden farkı	Kritik değer %5	Sonuç
kbgsmh	-2,3890	-3,6329	-3,9055	-3,6908	I(1)
ilk	0,5638	-1,9581	-2.5507	-1,9580	I(1)
orta	-1,5969	-3,6329	-8,9949	-3,7104	I(1)
yüksek	-2,7143	-3,6736	-4,3213	-3,6449	I(1)
bütçe	-2,3669	-3,0124	-3,6078	-3,0123	I(1)

Tahmin edilen ilköğretim ve bütçe değişkenli modeller haricindeki modeller, sabit ve trend içermektedir. İlköğretim değişkenine ilişkin model sabitsiz ve trendsiz bir model iken, bütçe değişkenine ilişkin model sadece sabit içermektedir. Bu değişkenlere ilişkin farklı gecikme değerlerinde ve farklı sabit ve trend durumlarındaki test sonuçları EK-3'de verilmiştir.

Dickey-Fuller (DF) ve Genelleştirilmiş Dickey-Fuller (GDF) birim kök testlerine ilişkin kritik değerler, MacKinnon tarafından üretilen değerlere dayanmaktadır.

Çizelge 14.1 incelendiğinde ikinci sütundaki GDF test istatistiği değerleri, tablo kritik değerlerinden mutlak değer olarak küçük olduğu için “H_0: Birim kök vardır” yokluk hipotezi red edilememiştir ve dizinin orjinalinde durağan olmadığı sonucuna varılır. Dördüncü sütundaki istatistik değerleri ise, değişkenin birinci farkları alındıktan sonra elde edilen istatistik değerleridir ve tablo kritik değerlerinden mutlak değerce büyük olduğu için yokluk hipotezi red edilmiştir.

Değişkenlerin seviyelerine uygulanan GDF test sonuçları değişkenlerin durağan olmadığını göstermektedir. Aynı testlerin, değişkenlerin birinci farklarına uygulanmasıyla elde edilen sonuçlar, değişkenlerin birinci dereceden farklarının durağan olduğunu gösterir.

14.7.2. Johansen eşbütünleşme testi

Değişkenlerin zaman dizisi özellikleri incelendikten sonraki adımı, söz konusu değişkenler arasındaki uzun dönemli ilişki olup olmadığının incelenmesi oluşturacaktır. Çalışmamızın bu bölümünde değişkenler arasındaki uzun dönemli ilişkilerin varlığı Johansen eşbütünleşme testi ile araştırılmıştır. Johansen ve Juselius (1990), değişkenler arasında uzun dönemli bir ilişkinin olup olmadığını test etmek için maksimum özdeğer (Maximum Eigen) ve iz (Trace) istatistikleri kullanılmaktadır. Johansen eşbütünleşme testi ile değişkenler arasındaki uzun dönemli ilişki araştırılırken kurulan VAR’da önemli bir rolü olan gecikme sayısı uygun testlerin kullanımı (Akaike Bilgi Kriteri) ile belirlenmiştir.

Eşbütünleşme test sonuçları Çizelge 14.2’de görülmektedir. Maximum Eigen ve Trace istatistikleri elde edilmiştir.

Çizelge 14.2. Johansen eşbütünleşim test sonuçları

H_0 Hipotezi	H_1 Hipotezi	Trace Test İstatistiği	Kritik Değer %5	Max. Eigen Test İstatistiği	Kritik Değer %5
r=0	r$\geq$1	77,7229*	60,0614	33,9965*	30,4396
r=1	r$\geq$2	43,7263*	40,1749	22,5139	24,1592
r=2	r$\geq$3	21,2124	24,2759	11,4265	17,7973

Eşbütünleşme analizi sonucuna göre diziler arası uzun dönemli bir ilişki vardır. Bu sonuca olabilirlik oran istatistiği değerleri ile tablo değerlerinin karşılaştırılması sonucu elde edilmiştir. $H_0 : r = 0$ (Eşbütünleşme denklemi oluşturulamaz) şeklindeki yokluk hipotezi Trace istatistik değeri (77,7229) ve Max. Eigen test istatistik değeri (33,9965), tablo değerinden büyük değerler olduğu için yokluk hipotezi red edilmektedir. Bundan sonra gelen $H_0 : r = 1$ (bir tane eşbütünleşme denklemi vardır) şeklindeki hipotez Trace test istatistiğine göre red edilirken Max. Eigen test istatistiğine göre kabul edilmiştir. Dolayısıyla Trace test istatistiğine göre iki tane, Max. Eigen test istatistiğine göre bir tane eşbütünleşme denklemi yazılabilmektedir. Yani diziler arasında uzun dönemli bir ilişki mevcuttur. Örneğin birinci eşbütünleşme denkleminde ilköğretim okullaşma oranına ilişkin dizinin katsayısı (-1426007), ortaöğretim okullaşma oranına ilişkin katsayı (13320,09) yükseköğrenim okullaşma oranına ilişkin katsayı (-62204,89) ve bütçe değişkenine ilişkin katsayıda (-165561,5) olmaktadır. Diğer eşbütünleşim denklemleri de aynı şekilde yorumlanır. Bu değerler Eviews programını kullanılarak yapılan Johansen eşbütünleşme test sonuçlarından elde edilmiştir ve ayrıntılı bilgileri EK-4'de yer almaktadır.

14.7.3. Kısmi lineer tekil değer eşbütünleşme testi

Bu bölümde bu teste ilişkin bilgisayar programı bulunmadığından hesapları tek tek ele alınmıştır. Buradaki hesaplamalar kısmi lineer modele göre yapılacaktır.

$\mu_0 = 0$ *ve* $\mu_1 = 0$ olma koşulu altında yani, sabitin ve trend değişkeninin olmadığı *VAR* modelini,

$$\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + g(Z_t) + \varepsilon_t$$

gibi ifade edilmektedir. Burada,

$$\Gamma = (\Gamma_1 \vdots \Gamma_2 \vdots \ldots \vdots \Gamma_{k-1})$$

ve

$$Y_t = (\Delta X_{t-1} \vdots \Delta X_{t-2} \vdots \ldots \vdots \Delta X_{t-k-1})$$

olmak üzere yeni modelimiz,

$$\Delta X_t = \Delta X_{t-1} + \Gamma Y_t + g(Z_t) + \varepsilon_t$$

olur.

Yukarıda verilen modelin her iki tarafından $E(\Delta X_t \mid Z_t)$ çıkarılırsa model,

$$\Delta X_t - E(\Delta X_t \mid Z_t) = \Pi(X_t - E(\Delta X_t \mid Z_t)) + \Gamma(Y_t - E(\Delta X_t \mid Z_t)) + \varepsilon_t$$

lineer hale gelir .

Öncelikle EK-5'de verildiği gibi optimal bant genişliğine ilişkin farklı kernel ağırlıklandırmalar hesaplanmıştır. Burada parametre dışı regresyon tahminlerinden Nadaraya-Watson kernel tahmini kullanılmıştır. Bunlara ilişkin grafikler Şekil 14.1-14.2-14.3'de gösterilmiştir. Daha sonra optimal bant genişliği konusunda tavsiye edilen $n^{-1/(q+4)}$ formülüne ilişkin ve bu optimal genişliğin iki katı ve yarısı oranında bant genişliklerinde tahmin değerleri elde edilmiştir. Tüm bu hesaplamalarda Stata programından yararlanılmıştır. Kernel fonksiyonlarını kullanarak değişkenlerin parametre dışı birinci fark tahmin değerleri, (Robinson (1988))

$$\Delta\hat{X}_t = \frac{1}{nh^q}\sum_{s=1,s\neq t}^{n} K_{ts}\Delta X_s / \hat{f}_t$$

formülü kullanılarak,

$$\Delta(kb\hat{g}smh)_t \rightarrow g(bütçe)_t$$
$$\Delta(i\hat{l}k)_t \rightarrow g(bütçe)_t$$
$$\Delta(o\hat{r}ta)_t \rightarrow g(bütçe)_t$$
$$\Delta(yü\hat{k}sek)_t \rightarrow g(bütçe)_t$$

bulunur. Aynı şekilde değişkenlerin gecikmeli ve birinci fark gecikmeli tahmin değerleri,

$$\hat{X}_{t-1} = \frac{1}{nh^q}\sum_{s=1,s\neq t}^{n} K_{ts}X_{s-1} / \hat{f}_t \quad \text{ve} \quad \hat{Y}_t = \frac{1}{nh^q}\sum_{s=1,s\neq t}^{n} K_{ts}Y_s / \hat{f}_t$$

formülü kullanılarak,

$$(kb\hat{g}smh)_{t-1} \rightarrow g(bütçe)_{t-1}$$
$$(i\hat{l}k)_{t-1} \rightarrow g(bütçe)_{t-1}$$
$$(o\hat{r}ta)_{t-1} \rightarrow g(bütçe)_{t-1}$$
$$(yü\hat{k}sek)_{t-1} \rightarrow g(bütçe)_{t-1}$$

ve

$$\Delta(kb\hat{g}smh)_{t-1} \rightarrow g(bütçe)_{t-1}$$
$$\Delta(i\hat{l}k)_{t-1} \rightarrow g(bütçe)_{t-1}$$
$$\Delta(o\hat{r}ta)_{t-1} \rightarrow g(bütçe)_{t-1}$$
$$\Delta(yü\hat{k}sek) \rightarrow g(bütçe)_{t-1}$$

bulunur. İkinci adımda bu tahmin değerlerinin kullanıldığı $\Delta X_t - \Delta\hat{X}_t = f(X_{t-1} - \hat{X}_{t-1}, Y_t - \hat{Y}_t)$ formülünden,

$$\begin{aligned}\Delta(kbgsmh)_t - \Delta(kb\hat{g}smh)_t = f[&((kbgsmh)_{t-1} - (kb\hat{g}smh)_{t-1}), ((ilk)_{t-1} - (i\hat{l}k)_{t-1}),\\ &((orta)_{t-1} - (o\hat{r}ta)_{t-1}), ((yüksek)_{t-1} - (yü\hat{k}sek)_{t-1}),\\ &(\Delta(kbgsmh)_{t-1} - \Delta(kb\hat{g}smh)_{t-1}),\\ &(\Delta(ilk)_{t-1)} - \Delta(i\hat{l}k)_{t-1}), (\Delta(orta)_{t-1} - \Delta(o\hat{r}ta)_{t-1}),\\ &(\Delta(yüksek)_{t-1} - (\Delta(yü\hat{k}sek)_{t-1})]\end{aligned}$$

bulunur.

Aynı şekilde,

$(\Delta(ilk)_t - \Delta(i\hat{l}k)_t)$, $(\Delta(orta)_t - \Delta(o\hat{r}ta)_t)$ ve $(\Delta(yüksek)_t - \Delta(yü\hat{k}sek)_t)$ değerleri de EKK regresyon tahmininden elde edilir.

Π 'nin tahminine geçmeden önce,

$$\Delta\widetilde{X} \rightarrow (\Delta X_t - \Delta\hat{X}_t)'\hat{f}_t$$
$$\widetilde{X} \rightarrow (X_{t-1} - \hat{X}_{t-1})'\hat{f}_t$$
$$\widetilde{Y} \rightarrow (Y_t - \hat{Y}_t)'\hat{f}_t$$

formülleri kullanılarak,

$$\Delta(kb\tilde{g}smh)_t \rightarrow [\Delta(kbgsmh)_t - \Delta(kb\hat{g}smh)_t]'\hat{f}_t$$
$$\Delta(i\tilde{l}k)_t \rightarrow [\Delta(ilk)_t - \Delta(i\hat{l}k)_t]'\hat{f}_t$$
$$\Delta(o\tilde{r}ta)_t \rightarrow [\Delta(orta)_t - \Delta(o\hat{r}ta)]'\hat{f}_t$$
$$\Delta(y\ddot{u}\tilde{k}sek)_t \rightarrow [\Delta(y\ddot{u}ksek)_t - \Delta(y\ddot{u}\hat{k}sek)_t]'\hat{f}_t$$

$$(kb\tilde{g}smh)_{t-1} \rightarrow [(kbgsmh)_{t-1} - (kb\hat{g}smh)_{t-1}]'\hat{f}_t$$
$$(i\tilde{l}k)_{t-1} \rightarrow [(ilk)_{t-1} - (i\hat{l}k)_{t-1}]'\hat{f}_t$$
$$(o\tilde{r}ta)_{t-1} \rightarrow [(orta)_{t-1} - (o\hat{r}ta)_{t-1}]'\hat{f}_t$$
$$(y\ddot{u}\tilde{k}sek)_{t-1} \rightarrow [(y\ddot{u}ksek)_{t-1} - (y\ddot{u}\hat{k}sek)_{t-1}]'\hat{f}_t$$

ve

$$\Delta(kb\tilde{g}smh)_{t-1} \rightarrow [\Delta(kbgsmh)_{t-1} - \Delta(kb\hat{g}smh)_{t-1}]'\hat{f}_t$$
$$\Delta(i\tilde{l}k)_{t-1} \rightarrow [\Delta(ilk)_{t-1} - \Delta(i\hat{l}k)_{t-1}]'\hat{f}_t$$
$$\Delta(o\tilde{r}ta)_{t-1} \rightarrow [\Delta(orta)_{t-1} - \Delta(o\hat{r}ta)_{t-1}]'\hat{f}_t$$
$$\Delta(y\ddot{u}\tilde{k}sek)_{t-1} \rightarrow [\Delta(y\ddot{u}ksek)_{t-1} - \Delta(y\ddot{u}\hat{k}sek)_{t-1}]'\hat{f}_t$$

vektörleri bulunur. Buradaki her bir vektör (21×1) boyutludur.

Yukarıda elde edilen değerlerle izdüşüm matrisi, $P = \tilde{Y}(\tilde{Y}'\hat{Y})^{-1}\tilde{Y}'$ formülünde ve $\tilde{\varepsilon}_{n\times p} \rightarrow \left[\left(\Delta X_t - \Delta\hat{X}_t\right) - \hat{\Pi}\left(X_{t-1} - \hat{X}_{t-1}\right)\right]'\hat{f}_t$ formülünde Matlab programında yerlerine konularak hesaplanmıştır.

Son olarak bu izdüşüm ve $\tilde{\varepsilon}$ matrisleri, $\tilde{\Pi} = (n^{-1}\tilde{\varepsilon}'(\mathrm{I}-P)\tilde{\varepsilon})^{-1/2}\hat{\Pi}(n^{-1}\tilde{X}'(\mathrm{I}-P)\tilde{X})^{1/2}$ formülünde yerine yazılarak $\tilde{\Pi}$ matrisi elde edilir.

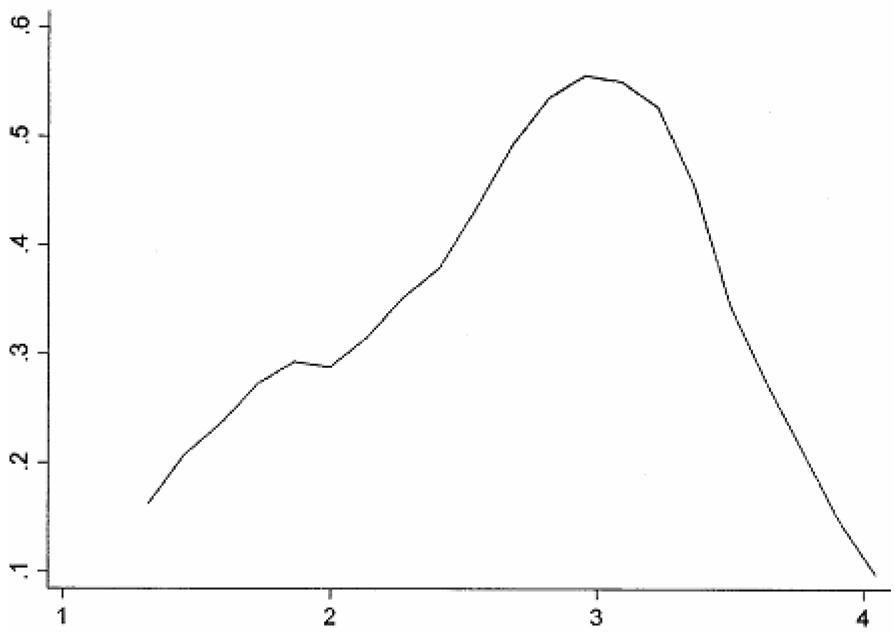
Kernel değerleri ve parametre dışı regresyondan hesaplanan tahmin değerleri Stata pogramından elde edilmiştir. Matlab programında yazılan program ile $\tilde{\Pi}'\tilde{\Pi}$

matrisinin özdeğerleri tekil değer ayrışımı ile elde edilmiştir. Bu özdeğerlere ilişkin bilgiler Çizelge 14.3'de verilmiştir.

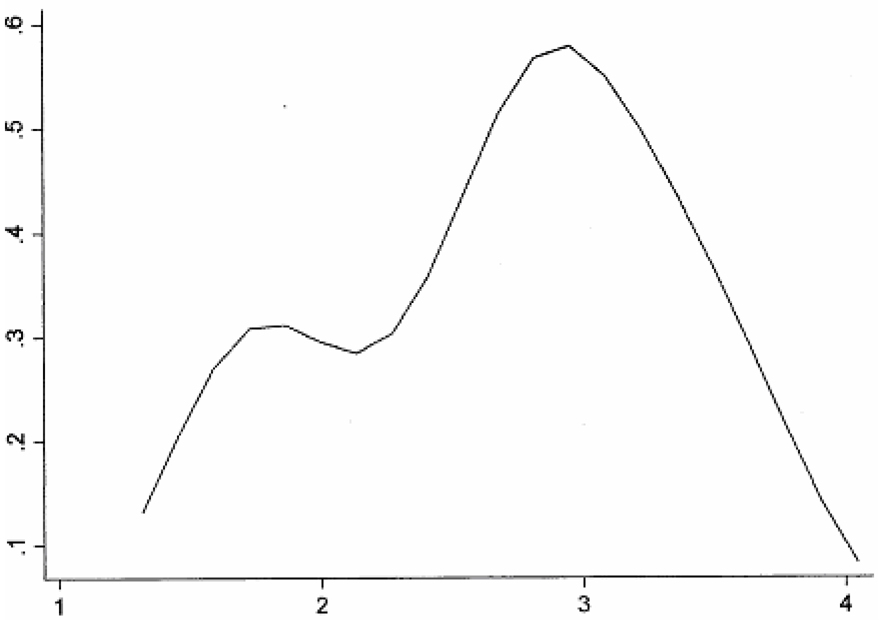
Tahminler için optimal bant genişliği;

$n^{-1/(q+4)} = 21^{-1/(1+4)} = 0{,}54$ olur.

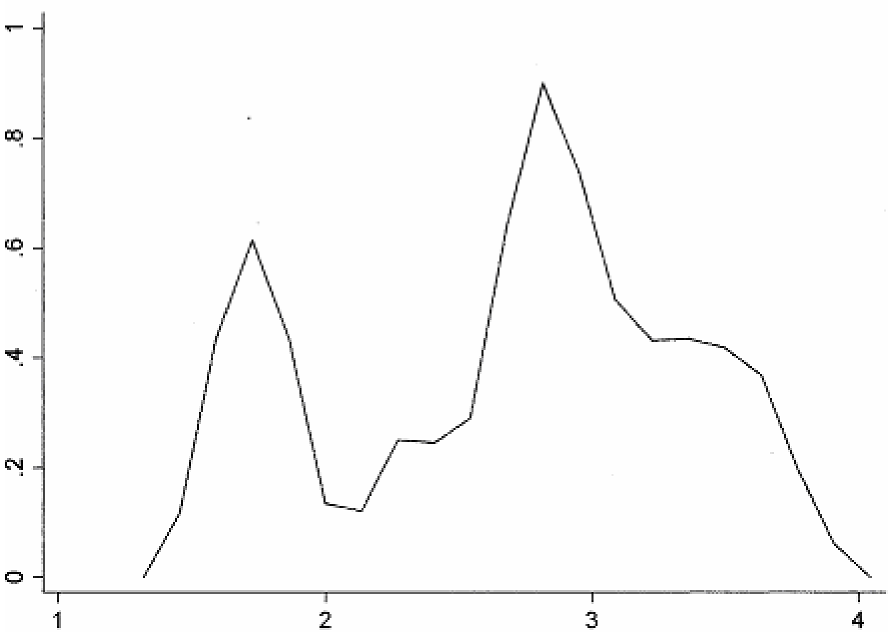
Burada n örnek hacmi, q ise modele lineer olmayan biçimde eklenen değişken sayısıdır. Örnek hacminin 21 olma nedeni PLS testindeki hesaplamalara ilişkin fark işlemleri ve gecikme değerleridir. Bu optimal bant genişliğinin iki katı $2n^{-1/(q+4)} = 1{,}08$ ve bu optimal bant genişliğinin yarısı $0{,}5n^{-1/(q+4)} = 0{,}27$ dir.



Şekil 14.1. Optimal bant genişliğinde Epanechnikov kernel ağırlıklandırma eğrisi



Şekil 14.2. Optimal bant genişliğinde Gaussian kernel ağırlıklandırma eğrisi



Şekil 14.3. Optimal bant genişliğinde Quadratic (Biweight) kernel ağırlıklandırma eğrisi

Çizelge 14.3. Farklı bant genişliklerinde ve farklı kernel ağırlıklandırmaları kullanılarak tekil değer ayrışımı ile elde edilen özdeğerler

	h= 0,27	h=0,54	h=1,08
Epanechnikov	14,9558 7,7129 1,9979 0,6537	405,3865 8,8946 0,8277 0,1450	350,1994 27,2317 0,1355 0,0059
Gaussian	238,8164 3,4484 0,9979 0,1017	530,7673 221,7562 0,6049 0,0054	244,3341 90,7732 1,4292 0,0181
Quadratic(Biweight)	6,2526 3,1081 1,3020 0,0488	73,3744 6,8563 1,8217 0,0398	26,7709 10,1413 3,4221 0,0668

Burada *h*, bant genişliğini, Epanechnikov, Gaussian ve Quadratic ise kernel ağırlıklandırmaları ifade etmektedir.

Bu özdeğerlerle elde edilen PLS değerleri Çizelge 14.4'de yer almaktadır.

Çizelge 14.4. Özdeğerlerle elde edilen PLS test istatistikleri

	h=0,27				h=0,54			
Hipotezler	r=0 $r \geq 1$	r=1 $r \geq 2$	r=2 $r \geq 3$	r=3 $r \geq 4$	r=0 $r \geq 1$	r=1 $r \geq 2$	r=2 $r \geq 3$	r=3 $r \geq 4$
Epanechnikov	531,73	217,65	55,68	13,73	8720,33	207,21	20,43	3,05*
Gaussian	5110,65	95,51	23,09	2,14*	15815,81	4669,70	12,82*	
Quadratic (Biweight)	224,94	93,64	28,37	1,02*	1723,94	183,07	39,09	0,84*

Çizelge 14.4. (Devam) Özdeğerlerle elde edilen PLS test istatistikleri

	h=1,08			
Hipotezler	r=0 r≥1	r=1 r≥2	r=2 r≥3	r=3 r≥4
Epanechnikov	7929,02	574,84	2,97*	
Gaussian	7067,65	1936,63	30,39	0,38*
Quadratic (Biweight)	848,42	286,23	73,27	1,40*

Çizelge 14.4'de elde edilen sonuçlara göre yıldızlı değerler, bunlara ilişkin tablo değerleri ile karşılaştırıldığında yokluk hipotezleri red edilememiştir [19]. Yokluk hipotezlerine karşılık gelen rank değerleri eşbütünleşme denklem sayılarını vermektedir.

Buna göre Epanechnikov kernel ağırlıklandırmasında 0,27 bant genişliğinde dört, 0,54 bant genişliğinde üç, 1,08 bant genişliğinde iki tane eşbütünleşme denklem sayısı bulunmuştur.

Gaussian kernel ağırlıklandırmasında 0,27 bant genişliğinde üç, 0,54 bant genişliğinde iki, 1,08 bant genişliğinde ise üç tane eşbütünleşme denklem sayısı bulunmuştur.

Quadratic (Biweight) kernel ağırlıklandırmasında ise, 0,27 bant genişliğinde üç, 0,54 bant genişliğinde üç, 1,08 bant genişliğinde de üç tane eşbütünleşme denklem sayısı bulunmuştur.

15. SONUÇ

Zaman dizilerine dayalı regresyon çözümlemeleri genelde, kullanılan zaman dizilerinin durağan olması koşulu altında yapılmakta ve klasik F, t…vb testler yapılmaktadır. Ancak uygulamaların çoğunda zaman dizileri durağan olmayacaktır.

Bu çalışmada, durağan olmayan zaman dizilerinden oluşan modellere ilişkin değişkenler arasındaki eşbütünleşme ilişkinin araştırılmasına yönelik kullanılan bilinen Johansen Testi ile Johansen yaklaşımından hareketle oluşturulan Kısmi Lineer Tekil Değer Testi uygulanmıştır.

Bu çalışmanın amacı, yukarıda belirtilen testler üzerinde, öne sürülen farklılıkları ortaya koymak ve bu yöntemleri ayrıntılarıyla ele almaktır. Elde edilen sonuçlar testleri birbirlerinden ayrılan farkları ortaya koyacak niteliktendir.

Öncelikle durağanlık ne demektir ve bir dizinin durağan olmaması ne anlama geldiği vurgulanmıştır. Durağanlık, zaman dizisinin birim kökü olup olmadığına bakılarak anlaşılır. Genişletilmiş Dickey-Fuller testi bu amaçla yapılmıştır.

Bir zaman dizisi değişkeninin, başka bir ya da birden çok zaman dizisine göre regresyonu çoğu zaman yanıltıcı olabilir. Buna karşı korunmanın bir yolu zaman dizilerinin eşbütünleşik olup olmadıklarına bakmaktır.

Eşbütünleşme tek tek aynı dereceden durağan olmayan iki ya da daha çok zaman dizisinin lineer birleşimlerinin durağan olması, bunların uzun dönemdeki ilişkilerinin anlamlı olduğu anlamına gelmektedir. Yani, değişkenleri durağanlaştırmadan uzun dönem model tahmini yapılabilir.

Tek değişkenli eşbütünleşme testleri, Engle Granger ve Genişletilmiş Engle Granger Testi ile Eşbütünleşme Regresyonu Durbin-Watson Testi gibi testlerdir.

Çok değişkenli eşbütünleşme testlerinden en yaygın kullanılan ise, Johansen en çok olabilirlik testidir. Bu test, parametrik testdir ve modelin lineer olduğunu varsayar. Johansen ve Juselius (1992) çalışmasında bu testi biraz daha geliştirerek bu durağan olmayan değişkenlerden oluşan eşbütünleşme denklemine yine lineer olarak durağan ortak bir değişken katmıştır ve böylece "eşbütünleşme yoktur" şeklindeki testin gücünü artırdığını göstermiştir.

Diğer çok değişkenli eşbütünleşme testi ise, Kısmi Lineer Tekil Değer Testi'dir.Bu test, Johansen testindeki eklenen durağan değişkene ilişkin lineer olma koşulunu rahatlatarak, bu değişkenin herhangi bir formda eklenebileceğini göstererek bu duruma ilişkin PLS testini geliştirmiştir. Bu eklenen ortak değişkenin aynı zamanda durağan olma koşulu da ortadan kaldırılmıştır. Bu değişken durağan olmayabilir. Böylece varyansın daha da küçültülebileceğini göstermiştir.

Bu durumda eşbütünleşme denklemi lineer ve lineer olmayan kısım olmak üzere iki kısımdan oluşmaktadır. Modelimiz böylece yarı parametrik bir model halini alır. Bunun sonunda yarı parametrik model tahmin problemiyle karşı karşıya kalınır.

Kısmi Lineer Model için parametrik kısmın parametre tahminleri parametrik olarak tahmin edilirken, lineer olmayan kısmın parametre tahminleri parametre dışı regresyon yöntemleriyle tahmin edilir. Kernel, Spline Smoothing…vb. yöntemlerle bu tahminler yapılır.

Çalışmada Türkiye'de eğitim ve iktisadi büyüme ilişkisini tanımlayan modele ilişkin yukarıda belirtilen eşbütünleşme testleri yapılmıştır.

Johansen testi sonuçunda Trace test istatistiğine göre iki, Max. Eigen test istatistiğine göre ise, bir tane eşbütünleşme denklemi bulunmuştur.

Kısmi Lineer Tekil Değer testinin sonuçunda ise, farklı bant genişliklerinde ve farklı kernel ağırlıklandırmalara ilişkin eşbütünleşme denklem sayıları elde edilmiştir.

Bu çalışma sonucunda eşbütünleşme testlerine ilişkin denklemlerin farklı kurulabileceği görülmüştür. Ayrıca farklı bant genişlikleri ve farklı kernel tahminlerinden elde edilen eşbütünleşme denklem sayıları da farklılık göstermektedir. İktisadi modeller her zaman lineer model yapıları oluşturmayabilirler. Bu nedenle kısmi lineer model yapıları ekonomik çalışmalar yönünde avantaj sağlayacaktır.

Şunu belirtmek gerekir ki, testlerin tam karşılaştırılmasına yönelik çalışma, testlerin gücüne bakılarak gerçekleştirilmelidir.

KAYNAKLAR

1. Akdi, Y., “Çok Değişkenli Zaman Serileri Analizi”,Zaman Serileri Analizi (Birim Kökler ve Kointegresyon), ***Bıçaklar Kitabevi,*** Ankara, 249-296 (2003)

2. Kadılar, C., “Zaman Serileri Analizinde Temel Kavramalar ve Önemli İşlemler”, SPSS Uygulamalı Zaman Serileri Analizine Giriş, ***Bizim Büro Basımevi***, Ankara, 1-2 (2005)

3. Damodar, N. G., “Zaman Serileri Ekonometrisi 1”, Temel Ekonometri, Ümit Şenesen,Gülay Günlük Şenesen, ***Literatür Yayıncılık,*** 709-730 (2001)

4. Mackinnon, J. G., “Critical Values of Cointegration Tests”,Engle , R. F., Granger, C. W. J., “Long-Run Economic Relationships:Readings in Cointegration”, 13. Bölüm, ***Oxford University Pres***, New York, (1991)

5. Kadılar, C., “Birim Kök Testleri”, “Eşbütünleşme Analizi”,Uygulamalı Çok Değişkenli Zaman Serileri, ***Bizim Büro Basımevi,*** 15-21, 112-149 (2000)

6. Brocklebank, J. C., Dickey D. A., “Simple Models: Regression” , Box 8000, SAS Circle, ***SAS Institute Inc.***, 6-9 (1986)

7. Sargan, J. D., Bhargava, A. S., “Testing Residuals from Least Squares Regression for Being Generated by the Gaussian Random Walk”, ***Econometrica,*** 153-174 (1983)

8. Enders, W., “Cointegration and Error Correction Models”, Applied Econometric Time Series, Second Edition , David J. Balding, Peter Bloomfield, Noel A. C. Cressie, Nicholas I. Fisher, lain M. Johnstone, J. B. Kadane, Louise M. Ryan, David W. Scott, Adrian F. M. Smith, Jozef L. Teugels, ***University of Alabama***, United States of America, 335-386 (2003)

9. Cula, S., G., “Çok Değişkenli Olasılık Yoğunluk Fonksiyonunun Çekirdek Fonksiyonlarıyla Kestirimi” ***Hacettepe Üniversitesi Fen Bilimleri Enstitüsü,*** Ankara, 2-36 (1998)

10. Rosenblatt, M., “Remarks on Some Nonparametric Estimate of a Density Function”, ***Ann. Math. Stat,*** 832-837 (1956)

11. Arıkan, E., “Parametrik Olmayan Regresyon ve Bir Uygulama”, ***Hacettepe Üniversitesi Fen Bilimleri Enstitüsü,*** Ankara, 6-19 (1993)

12. Mandel, J., “Use of the Singular Value Decomposition in Regression Analysis” ***The American Statistician*** , 15-17 (1982)

13. Yangin F.and Qi Li "A Kernel-Based Method For Estimation Additive Partially Linear Models", ***Statistica Sinica,*** 739-762 (2003)

14. Ted, J., Zhijie, X. "Testing for cointegration using partially linear models", ***Journal of Econometrics,*** 365-394 (2005)

15. Boswijk, H. P., Doornik, J. A., "Distribution Appoximations for Cointegration Test with Stationary Exogenous Regressors",***Universiteit van Amsterdam,*** 1-10 (1999)

16. Doğan, S., Bozkurt, H. Y., "Eğitim-İktisadi Büyüme İlişkisi ve Türkiye İçin Kointegresyon Analizi", ***www.bilgiyonetimi.org/cm/pages/yazArk.php?page*** , 1-8 (2006)

17. TÜİK, "Türkiye Eğitim Harcamaları Araştırması",(2002)

18. DİE , "Türkiye Eğitim İstatistikleri",(2005-2006)

19. Seo, B., "Statistical inference on cointegration rank in error correction models with statistionary covariates",***Journal of Econometrics,*** 339-385 (1998)

20. İnternet :Türkiye İstatistik Kurumu "Eğitim İstatistikleri" http://www.tuik.gov.tr/PreIstatistikTablo.do?istab_id=102 (2006)

EKLER

EK-1 Bazı Kernel Fonksiyonları [11]

Kernel	$K(u)$
Uniform	$1/2 \qquad \lvert u\rvert < 1$ $0 \qquad \lvert u\rvert > 1$
Triangular	$1-\lvert u\rvert \qquad \lvert u\rvert \le 1$ $0 \qquad \lvert u\rvert > 1$
Normal	$(\sqrt{2\pi})^{-1}\ e^{-u/2}$
Double Exponential	$1/2\ \exp(-\lvert u\rvert)$
Quadratic	$(3/4\sqrt{5})\ (1-u^2/5) \qquad \lvert u\rvert \le \sqrt{5}$ $0 \qquad \lvert u\rvert > \sqrt{5}$
Fejer	$(2\pi)^{-1}\left\{\frac{\sin(u/2)}{u/2}\right\}^2$
Caucy	$\{\pi(1+u^2)\}^{-1}$

EK-2 Eğitim ve İktisadi Veriler [20]

Yıl	***kbgsmh***	***ilk***	***orta***	***yüksek***	***bütçe***
1983	1196703	0,8368162	27,9	8	2,75
1984	1250251	0,831799	29,7	9,7	2,07
1985	1271997	0,8311897	31,7	10,7	1,74
1986	1328231	0,8290612	33,3	11,3	1,64
1987	1427282	0,8447743	34,3	11,7	1,67
1988	1416888	0,8366882	35,2	12,8	1,87
1989	1409056	0,8335614	36,6	14,5	1,74
1990	1505110	0,8295347	38,5	15,7	2,77
1991	1481321	0,8346436	41,7	16,4	2,9
1992	1546592	0,8347982	44,9	18,1	3,57
1993	1641872	0,8276188	47,7	22,2	3,72
1994	1514346	0,8276306	53	22,1	3,21
1995	1606454	0,8211217	55	22,4	2,3
1996	1691943	0,8196418	54,7	23,2	2,37
1997	1838576	0,8951	52,79	19,52	2,81
1998	1880016	0,9431	57,15	21,67	3,39
1999	1741293	0,9752	58,84	21,05	3,5
2000	1766124	1,0093	60,97	22,25	2,783
2001	1570770	0,9945	67,89	23,37	2,716
2002	1670893	0,9649	80,76	27,12	2,751
2003	1741783	0,963	80,97	28,15	2,946
2004	1884802	0,9574	80,9	30,48	3,041
2005	2021100	0,9559	85,18	34,46	3,197

EK-3 Durağanlık Test Sonuçları

Açıklayıcı değişken	Gecikme değerleri	Test istatistik değeri	Kritik değer %5	Model yapısı
kbgsmh	0	1,7894	-1,9572	Sabit yok Trend yok
kbgsmh	0	-0,7595	-3,0048	Sabit var Trend yok
kbgsmh	0	-2,3890	-3,6328	Sabit var Trend var
kbgsmh	3	-3,9055	-3,6908	Sabit var Trend var

Açıklayıcı değişken	Gecikme değerleri	Test istatistik değeri	Kritik değer %5	Model yapısı
ilk	0	-2,5507	-1,9580	Sabit yok Trend yok
ilk	0	-2,6010	-3,0123	Sabit var Trend yok
ilk	0	-2,5120	-3,6449	Sabit var Trend var
ilk	1	0,5638	-1,9580	Sabit var Trend var
ilk	1	-1,2952	-3,0123	Sabit var Trend yok
ilk	1	-2,4028	-3,6449	Sabit yok Trend yok

EK-3 (Devam) Durağanlık Test Sonuçları

Açıklayıcı değişken	Gecikme değerleri	Test istatistik değeri	Kritik Değer %5	Model yapısı
orta	0	4,1628	-1,9572	Sabit yok Trend yok
orta	0	0,8298	-3,0048	Sabit var Trend yok
orta	0	-1,5969	-3,6328	Sabit var Trend var
orta	4	-8,9949	-3,7104	Sabit var Trend var

Açıklayıcı değişken	Gecikme değerleri	Test istatistik değeri	Kritik değer %5	Model yapısı
yüksek	0	3,3857	-1,9572	Sabit yok Trend yok
yüksek	0	0,4431	-3,0048	Sabit var Trend yok
yüksek	0	-4,3213	-3,6449	Sabit var Trend var
yüksek	3	-2,7142	-3,6736	Sabit var Trend var

EK-3 (Devam) Durağanlık Test Sonuçları

Açıklayıcı değişken	Gecikme değerleri	Test istatistik değeri	Kritik Değer %5	Model yapısı
bütçe	0	-0,1500	-1,9572	Sabit yok Trend yok
bütçe	0	-3,6078	-3,0123	Sabit var Trend yok
bütçe	1	-2,3669	-3,0123	Sabit var Trend yok
bütçe	1	-2,9845	-3,6449	Sabit var Trend var
bütçe	2	-3,6617	-3,6736	Sabit var Trend var

EK-4 Johansen Eşbütünleşme Denklem Yapıları

1 Eşbütünleşme denklemi, Log likelihood $= -265,2795$

kbgsmh	*ilk*	*orta*	*yüksek*	*bütçe*
1,000000	-1426007	13320,09	$-62204,89$	$-165561,5$
	(82718,5)	(3762,81)	(9995,75)	(50567,4)

2 Eşbütünleşme denklemi, Log likelihood $= -254,0225$

kbgsmh	*ilk*	*orta*	*yüksek*	*bütçe*
1,000000	0,000000	2758395	10914963	$2,92E+08$
		(6354597)	$(1,7E+07)$	$(8,6E+07)$
0,000000	1,000000	1,925008	7,697835	204,5585
		(4,45862)	(11,7002)	(60,1962)

3 Eşbütünleşme denklemi, Log likelihood $= -248,3092$

kbgsmh	*ilk*	*orta*	*yüksek*	*bütçe*
1,000000	0,000000	0,000000	-12110465	$-1,50E+08$
			(1605111)	$(5,6E+07)$
0,000000	1,000000	0,000000	$-8,370976$	$-103,6125$
			(1,11116)	(38,6670)
0,000000	0,000000	1,000000	8,347400	160,0881
			(1,46100)	(50,8411)

EK-4 (Devam) Johansen Eşbütünleşme Denklem Yapıları

4 Eşbütünleşme denklemi, Log likelihood = −243,4377

kbgsmh	*ilk*	*orta*	*yüksek*	*bütçe*
1,000000	0,000000	0,000000	0,000000	− 4237606
				(1559758)
0,000000	1,000000	0,000000	0,000000	− 2,824209
				(1,01378)
0,000000	0,000000	1,000000	0,000000	59,58375
				(13,6584)
0,000000	0,000000	0,000000	1,000000	12,04020
				(4,49543)

EK-5 Yıllara Göre Epanechnikov, Gaussian ve Quadratic Kernel Ağırlıklandırma Değerleri

Yıl	Epanechnikov	Gaussian	Quadratic
1985	0,16302	0,13269	2,74E-32
1986	0,20811	0,20532	0,11628
1987	0,23686	0,27074	0,43337
1988	0,27315	0,30829	0,61503
1989	0,29315	0,31160	0,43563
1990	0,28868	0,29508	0,13438
1991	0,31569	0,28493	0,12024
1992	0,3523	0,30331	0,25037
1993	0,37937	0,35728	0,24526
1994	0,43416	0,43612	0,29036
1995	0,49328	0,51556	0,64153
1996	0,53617	0,56855	0,90211
1997	0,55719	0,57937	0,73816
1998	0,55109	0,55085	0,50542
1999	0,52771	0,49798	0,43168
2000	0,4548	0,43447	0,43505
2001	0,34567	0,36547	0,41891
2002	0,2742	0,29110	0,36807
2003	0,21151	0,21391	0,19715
2004	0,14791	0,14152	0,06171
2005	0,09839	0,08270	2,74E-32

ÖZGEÇMİŞ

Kişisel Bilgiler

Soyadı, adı	: PINARÖNÜ, Naciye
Uyruğu	: T.C.
Doğum tarihi ve yeri	: 26.01.1981 Ankara
Medeni hali	: Bekar
Telefon	: 0 (312) 480 72 59
e-mail	: npinaronu@hotmail.com

Eğitim

Derecesi	Eğitim Birimi	Mezuniyet Tarihi
Yüksek lisans	Gazi Üniversitesi /İstatistik Bölümü	2007
Lisans	Gazi Üniversitesi/ İstatistik Bölümü	2004
Lise	Sokullu Mehmet Paşa Lisesi(YDA)	1999

Yabancı Dil

İngilizce

Hobiler

Şiir, sinema, dış politika, edebiyat