

T.C.
İSTANBUL ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
EKONOMETRİ ANABİLİM DALI

YÜKSEK LİSANS TEZİ

KESİRLİ BİRİM KÖK TESTLERİN
PERFORMANSLARININ MONTE CARLO
SİMÜLASYON TEKNİĞİ KULLANILARAK
İNCELENMESİ

HAMZA MUTLUAY
2501161105

TEZ DANIŞMANI
PROF. DR. NİLGÜN ÇİL

İSTANBUL, 2019



T.C.
İSTANBUL ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ



YÜKSEK LİSANS
TEZ ONAYI

ÖĞRENCİNİN;

Adı ve Soyadı : HAMZA MUTLUAY Numarası : 2501161105
Anabilim Dalı /
Anasanat Dalı / Programı : EKONOMETRİ Danışmanı : PROF. DR. NİLGÜN ÇİL
Tez Savunma Tarihi : 12.07.2019 Saati : 11.00
Tez Başlığı : KESİRLİ BİRİM KÖK TESTLERİNİN PERFORMANSLARININ MONTE CARLO
SİMÜLASYON TEKNİĞİ KULLANILARAK İNCELENMESİ

TEZ SAVUNMA SINAVI, İÜ Lisansüstü Eğitim-Öğretim Yönetmeliği'nin 36. Maddesi uyarınca yapılmış,
sorulan sorulara alınan cevaplar sonunda adayın tezinin KABULÜNE OYBİRLİĞİ / OYÇOKLUĞUYLA karar verilmiştir.

JÜRİ ÜYESİ	İMZA	KANAATİ (KABUL / RED / DÜZELTME)
1- PROF. DR. NİLGÜN ÇİL		Kabul
2- PROF. DR. FERDA YERDELEN TATOĞLU		Kabul
3- PROF. DR. MURAT GÜRAY KIRDAR		Kabul

YEDEK JÜRİ ÜYESİ	İMZA	KANAATİ (KABUL / RED / DÜZELTME)
1- PROF. DR. NURCAN METİN		
2- DOÇ. DR. BURCU KIRAN BAYGIN		

ÖZ

KESİRLİ BİRİM KÖK TESTLERİN PERFORMANSLARININ MONTE CARLO SİMÜLASYON TEKNİĞİ KULLANILARAK İNCELENMESİ

HAMZA MUTLUAY

Zaman serisi analizlerinde durağanlık sınaması birim kök testleriyle yapılmaktadır. Standart birim kök testlerinde serilerin bütünleşme derecelerinin tam sayı alması gibi bir kısıt bulunmaktadır. Fakat seriler kesirli bütünleşik sürece sahip olduğu durumda klasik birim kök testleri ile sınamalar yapılırsa testlerin gücünün düştüğü yapılan çalışmalarla ortaya konulmuştur. Bundan dolayı serilerin bütünleşme derecelerinin kesirli değerler almasına izin veren testler geliştirilmiştir. Bu tezde serilerin kesirli bütünleşik sürece sahip olduğu durum için Dolado, Gonzalo ve Mayoral (2002) tarafından geliştirilen kesirli Dickey Fuller ve kesirli genişletilmiş Dickey Fuller testlerinin performansları Monte Carlo simülasyon yöntemiyle araştırılmıştır. Simülasyonlar sonucunda modele deterministik bileşenler eklendiği durumda bu testlerin I. tip hata olasılıklarında bozulmalar meydana geldiği görülmüştür. Öte yandan kesirli genişletilmiş Dickey Fuller testinde hata teriminin otokorelasyonlu yapıya sahip olduğu durumda bozulan I. tip hata olasılığı sieve bootstrap yöntemi ile düzeltilmiştir. Bu tez çalışmasında önerilen yeni bootstrap birim kök testi ile Avrupa’da yer alan ve gelişmişlik düzeylerine göre sınıflandırılan 32 ülke grubu için hisse senedi piyasalarında zayıf formda etkinliğin geçerliliği sınanmıştır. Yapılan analizler sonucunda ülkelerin gelişmişlik seviyeleri arttıkça zayıf formda etkin olan piyasaların oranında da artışlar olduğu yönünde bulgular elde edilmiştir.

Anahtar Kelimeler: Kesirli Bütünleşme, Monte Carlo, Sieve Bootstrap, Etkin Piyasa Hipotezi, I.Tip Hata Bozulması

ABSTRACT

AN EXAMINATION OF PERFORMANCES OF FRACTIONAL UNIT-ROOT TESTS USING MONTE-CARLO SIMULATION TECHNIQUES

HAMZA MUTLUAY

In time series analysis, testing for stationarity is performed by unit root tests. However, in standard unit root tests, the degree of integration of the series typically takes integer numbers. However, it has been revealed by studies that the power of the classical unit root tests is reduced when the series has a fractional integrated process. Therefore, tests have been developed to allow the degree of integration of the series to take fractional values. In this study, we investigate the performance of the fractional Dickey-Fuller tests and the augmented fractional Dickey-Fuller tests developed by Dolado, Gonzalo and Mayoral (2002) using Monte Carlo simulation methods. The simulation results show that size distortions occur when the fractional Dickey-Fuller and fractional augmented Dickey-Fuller tests include a deterministic component. In addition, in the case of augmented fractional Dickey-Fuller test, we improve possible size distortions using the sieve bootstrap method when the error term is autocorrelated. As an example to these methods, we test the validity of the weak form market efficiency for the 32 countries in the European region, which are classified according to their level of development. We find that the fraction of markets which are efficient in the weak form rises as the development levels of the countries increase.

Key Words: Fractional Integration, Monte Carlo, Sieve Bootstrap, Efficient Market Hypothesis, Size Distortion

ÖNSÖZ

Ekonometrik analizlerde serilerin durağanlık seviyelerinin doğru olarak belirlenmesi analizlerin önemli bir aşamasıdır. Literatürde en çok bilinen ve kullanılan Dickey-Fuller, Phillips-Perron birim kök testlerinde serilerin eşbütülenleşme sıralarının tam değer aldığı varsayılmaktadır. Ancak gerçekte bütülenleşme sırası kesirli iken geleneksel birim kök testlerinin kullanılması durumunda testin gücünün düştüğü literatürdeki çalışmalarda gösterilmiştir.

Bu amaçla Dolado, Gonzalo ve Mayoral (2002) tarafından geliştirilen kesirli Dickey Fuller ve kesirli genişletilmiş Dickey Fuller testleri öncü çalışmalardan birisidir. Bu çalışmada ele alınan testin otokorelasyon ve deterministik bileşenler durumunda sınırlı veri kümesinde 1. tip hata olasılıklarında bozulmalar olduğu gözlemlenmiştir. 1. Tip hata olasılığında meydana gelen bu bozulmalar Sieve bootstrap yöntemi kullanılarak düzeltilmiştir.

Bu tezin yapılması aşamasında öneri ve yardımlarıyla her zaman yanımda olan başta danışman hocam Prof. Dr. Nilgün ÇİL olmak üzere, Prof. Dr. Ferda Yerdelen Tatoğlu'na, Prof. Dr. Murat Güray Kırdar'a, Doç.Dr. Furkan Emirmahmutoğlu'na, Arş. Gör.Halil İbrahim Gündüz'e teşekkürlerimi sunarım. Çalışmalarım süresince yanımda olan aileme de teşekkürü borç bilirim.

İSTANBUL 2019

HAMZA MUTLUAY

İÇİNDEKİLER

ÖZ.....	iii
ABSTRACT.....	iv
ÖNSÖZ.....	v
TABLolar LİSTESİ.....	viii
ŞEKİLLER LİSTESİ.....	ix
KISALTMALAR LİSTESİ	x
GİRİŞ	1

BİRİNCİ BÖLÜM

DOĞRUSAL ZAMAN SERİLERİ VE UZUN HAFIZA

1.1. Durağanlık	4
1.2. Bazı Stokastik Süreçler	5
1.2.1. Pür Rastgele Süreç.....	5
1.2.2. Hareketli Ortalama Süreci.....	6
1.2.3. Otoregresif Süreç	8
1.2.4. Otoregresif Hareketli Ortalama Süreci	10
1.2.5. Bütünleşik Otoregresif Hareketli Ortalama Süreci.....	12
1.3. Uzun Hafıza Süreci.....	13
1.3.1. Uzun Hafıza ve Yapısal Kırılma Arasındaki İlişkiler	18

İKİNCİ BÖLÜM

MONTE CARLO SİMÜLASYON TEKNİĞİ KULLANILARAK KESİRLİ BİRİM KÖK TESTLERİNİN PERFORMANSLARININ İNCELENMESİ

2.1. Dickey Fuller ve Genişletilmiş Dickey Fuller Testleri.....	21
2.2. Kesirli Dikey Fuller Testi.....	26

2.3. Deterministik Bileşenler Durumunda Dolado, Gonzalo ve Mayoral (2002)	
Birim Kök Testi: Monte Carlo Simülasyon Sonuçları.....	32
2.4. Zaman Serilerinde Bootstrap	43
2.4.1. Kesirli Birim Kök Testi İçin Sieve Bootstrap Algoritması	47
2.4.2. Monte Carlo Simülasyon Sonuçları	49
ÜÇÜNCÜ BÖLÜM	
FARKLI GELİŞMİŞLİK DÜZEYLERİNE SAHİP AVRUPA ÜLKELERİNDE	
ETKİN PİYASA HİPOTEZİNİN GEÇERLİĞİNİN SINANMASI	
3.1. Etkin Piyasa Hipotezi	53
3.2. Veri Seti ve Yöntem.....	56
3.3. Araştırma Bulguları	58
SONUÇ VE ÖNERİLER.....	61
KAYNAKLAR	63
ÖZGEÇMİŞ.....	68

TABLÖLER LİSTESİ

Tablo 2.1. Deterministik Bileşenler Durumunda FD-F Testinin Kritik Değerleri.....	33
Tablo 2.2. Deterministik Bileşensiz Durumda FD-F Testinin I. Tip Hata Olasılığı ve Testin Gücü	34
Tablo 2.3. Deterministik Bileşenli Durumda FD-F Testinin I. Tip Hata Olasılığı ve Testin Gücü	35
Tablo 2.4. Deterministik Bileşensiz Durumda AFD-F Testinin I. Tip Hata Olasılığı ve Testin Gücü	37
Tablo 2.5. Deterministik Bileşensiz Durumda AFD-F Testinin I. Tip Hata Olasılığı ve Testin Gücü	39
Tablo 2.6. Deterministik Bileşenli Durumda AFD-F Testinin I. Tip Hata Olasılığı ve Testin Gücü	41
Tablo 2.7. Deterministik Bileşenli Durumda AFD-F Testinin I. Tip Hata Olasılığı ve Testin Gücü	42
Tablo 2.8. AFD-F Sieve Bootstrap Testinin I. Tip Hata Olasılığı ve Testin Gücü....	49
Tablo 3.1. Avrupa'daki Gelişmemiş Ülkelerin Piyasaları İçin EPH'nin Sınanması.....	58
Tablo 3.2. Avrupa'daki Gelişmekte Olan Ülkelerin Piyasaları İçin EPH'nin Sınanması	59
Tablo 3.3. Avrupa'daki Gelişmiş Ülkelerin Piyasaları İçin EPH'nin Sınanması	60

ŞEKİLLER LİSTESİ

Şekil 1.1. AR(1) Sürecinin Otokorelasyon Fonksiyonu.....	17
Şekil 1.2 ARFIMA(0,0.4,0) Sürecinin Otokorelasyon Fonksiyonu.....	17



KISALTMALARA LİSTESİ

AD-F	: Genişletilmiş Dickey Fuller (Augmented Dickey Fuller)
D-F	: Dickey & Fuller
FD-F	: Kesirli Dickey & Fuller (Fractional Dickey & Fuller)
AFD-F	: Genişletilmiş Kesirli Dickey & Fuller (Augmented Fractional Dickey & Fuller)
AIC	: Akaike
AR	: Otoregresif (Autoregressive)
ARFIMA	: Otoregresif Kesirli Bütünleşik Hareketli Ortalama (Autoregressive Fractionally Integrated Moving Average)
ARIMA	: Otoregresif Bütünleşik Hareketli Ortalama (Autoregressive Integrated Moving Average)
ARMA	: Otoregresif Hareketli Ortalama (Outoregressive Moving Average)
BIC	: Schwarz Bilgi Kriteri
FI	: Kesirli Bütünleşme (Fractional Integration)
GMD	: Genelleştirilmiş En Küçük Uzaklık (Generalized Minimum Distance)
LM	: Lagrange Çarpanı (Lagrange Multiplier)
EPH	: Etkin Piyasa Hipotezi

GİRİŞ

Zaman serileri ile analizler yapılırken serilerin durağanlık özelliklerinin incelenmesi yapılan araştırmaların güvenilirliği açısından önem arz etmektedir. Literatürde serilerin durağan olup olmadıklarını araştırmak için birçok birim kök testi yer almakla beraber klasik testler serilerin bütünleşme derecelerini tam sayı olarak ele almaktadır. Fakat son zamanlarda, serilerin bütünleşme derecelerinin kesirli değerler aldığı uzun hafızalı süreçler araştırmalarda geniş yer bulmaktadır. Öte yandan uzun hafızalı sürece sahip zaman serilerinde durağanlık analizleri yapılırken klasik birim kök testlerinin kullanılmasının hatalı sonuçlar elde edilmesine neden olduğu belirtilmektedir. Çünkü klasik birim kök testleri aracılığıyla uzun hafızalı bir sürecin durağanlık özellikleri incelendiğinde testlerin güçleri azalmakta, dolayısıyla uzun hafızaya sahip durağan bir seri için durağan değilmiş gibi sonuçlar elde edilmektedir. Bu noktadan hareketle çalışmada Dolado, Gonzalo ve Mayoral (2002) tarafından geliştirilen kesirli Dickey Fuller (FD-F) ve genişletilmiş kesirli Dickey Fuller (AFD-F) testleri tüm yönleriyle ele alınacaktır. Bununla beraber Dolado, Gonzalo ve Mayoral (2002), modelde deterministik bileşenler olmadığı durum için testlerin sınırlı veri kümesinde ve asimptotik durumda I. tip hata olasılığı ve gücünü, diğer bir ifadeyle testlerin performanslarını incelemişlerdir. Bu çalışmada Dolado, Gonzalo ve Mayoral (2002)'den farklı olarak modelde deterministik bileşenler olduğu durum için de testlerin hem küçük hem de büyük örneklem durumlarında performansları analiz edilecektir. Yine Dolado, Gonzalo ve Mayoral (2002) çalışmalarında modelde deterministik bileşenlerin olmadığı ve otokorelasyon probleminin olduğu durum için AFD-F testinin performansını araştırmışlardır. Bu tez çalışmasında Dolado, Gonzalo ve Mayoral (2002)'den farklı olarak modelde deterministik bileşenlerin ve otokorelasyon probleminin olduğu durum için de AFD-F testinin performansı incelenecektir. Son olarak modelde otokorelasyon probleminin dolayı AFD-F testinin bozulan I. tip hata olasılığı sieve bootstrap yöntemi kullanılarak düzeltilecektir.

Çalışmanın ilk bölümünde zaman serisi analizlerinde sıklıkla kullanılan AR, MA, ARMA, ARIMA gibi modeller detaylı bir şekilde ele alınacak, daha sonra uzun hafıza kavramı ve uzun hafızalı modellerden ARFIMA modeli anlatılacaktır. Bölüm sonunda uzun hafıza ile yapısal kırılmalar arasındaki ilişkiden bahsedilecektir.

İkinci bölümün ilk kısmında kesirli birim kök testlerinden olan kesirli Dickey Fuller (FD-F) ve genişletilmiş kesirli Dickey Fuller (AFD-F) testinin aşamaları geniş bir çerçevede anlatılacaktır. Daha sonra bu testlerin sınırlı veri kümesinde ve asimptotik durumda performans değerlendirmesi yapmak için kullanılan Monte Carlo simülasyon yönteminin aşamaları verilerek simülasyon sonuçları tablolar halinde gösterilecektir. İkinci bölümün son kısmında ise AFD-F testinde modelin hata terimlerinde otokorelasyonlu bir yapı olduğu zaman bozulan I. tip hata olasılığını düzeltmek için başvurulmuş sieve bootstrap yönteminde bahsedilecektir. Son olarak sieve bootstrap yöntemi aracılığıyla elde edilen sonuçlar tablolar halinde verilecektir.

Son bölümde ise, Avrupa ülkelerinde hisse senedi piyasalarında etkinlik araştırması, bu tez kapsamında eksiklikleri düzeltilmiş AFD-F testi aracılığıyla yapılacaktır. Uygulamadan önce etkin piyasa hipotezinin tanımı yapılacak ve daha sonra arkasında yatan temel mantık ve tarihçesi anlatılacaktır. Sonraki aşamada gelişmişlik seviyelerine göre 3 gruba ayrılan Avrupa ülkeleri hisse senedi piyasası için zayıf formda etkinlik varsayımının geçerliliği sınanacaktır. Bölümün son kısmında analiz sonuçları tablolastırılarak yorumlanacaktır.

Sonuç ve öneriler kısmında ise, ilk olarak uygulamadan elde edilen sonuçlar yorumlandıktan sonra literatüre olan katkıyı arttırmak açısından çeşitli öneriler sunulacak çalışmada sonlandırılacaktır.

BİRİNCİ BÖLÜM

DOĞRUSAL ZAMAN SERİLERİ VE UZUN HAFIZA

Teorik açıdan bakıldığında, stokastik süreç rastgele değişkenler topluluğu $\{X_t\}$ olarak adlandırılmaktadır¹. Rastgele değişkenler dizisi X_t 'nin olasılık yapısı stokastik sürecin ortak dağılımı ile belirlenmektedir. Fakat stokastik sürecin ortak dağılım ile ifade edilen bu tanımı çok genel olduğu için stokastik süreç, rastgele değişken X_t 'nin birinci ve ikinci momentleriyle tanımlanacaktır. Belirli bir t zamanında X_t rastgele değişkeninin dağılım ve yoğunluk fonksiyonunu sırasıyla $F(X_t)$ ve $f(X_t)$ şeklinde gösterilir. Stokastik sürecin dağılım fonksiyonu birinci ve ikinci momentleriyle aşağıdaki gibi karakterize edilmektedir:

$$\text{ortalama} \quad \mu_t = E(X_t) \quad (1.1)$$

$$\text{varyans} \quad \sigma_t^2 = \text{var}(X_t) \quad (1.2)$$

$$\text{otokovaryans} \quad \gamma_{t_1, t_2} = \text{cov}(X_{t_1}, X_{t_2}) \quad (1.3)$$

Yukarıdaki eşitliklerden de anlaşıldığı üzere X_t 'nin birinci ve ikinci momentleri zamanın (t) bir fonksiyonudur. Eğer X_t normal dağılıma sahipse X_t için Gaussian sürece sahiptir denir. $\mu_t, \sigma_t^2, \gamma_{t_1, t_2}$ gibi bilinmeyen parametreler t 'nin bir fonksiyonu olduğu için t 'nin değerine bağlı olarak değişmektedir. Bundan dolayı tahmin edilmesi gereken çok fazla parametre vardır. Fakat gerçekte her bir rastgele değişken için bir tane gözlem bulunmaktadır. Örneğin rastgele değişken olarak haftalık işsizlik oranlarının incelediği düşünüldüğünde her hafta için bir tane gözlem vardır. Bu durum, literatürde tek bir gerçekleşme olarak ifade edilmektedir ve araştırmacıları stokastik süreçlerin istatistiksel yapılarını ortaya koymak için kısıtlayıcı modelleri kullanmak durumunda bırakmıştır. Dolayısıyla tek bir gerçekleşme için araştırmacılar parametreleri azaltmak durumundadırlar. Buradaki

¹G. S. Maddala and I. M. Kim, **Unit Roots, Cointegration, and Structural Change**, Cambridge, Cambridge University Press, 1998, s.8.

önemli nokta parametrelerin nasıl azaltılacağıdır. Tahmin edilecek parametreleri azaltma veya diğer adıyla indirgeme işlemleri bazı kısıtlamalar altında yapılabilmektedir². Bu kısıtlamalar;

- i. Durağanlık
- ii. Asimptotik bağımsızlık

olarak ifade edilmektedir.

1.1. Durağanlık

Durağanlık kavramı zaman serisi analizleri açısından büyük öneme sahiptir. Bir rastgele sürecin olasılık dağılımı zamandan bağımsız ise süreç durağandır denilmektedir³. Durağan stokastik süreçler, zayıf ve güçlü durağan olmak üzere ikiye ayrılmaktadır. Durağanlığı tanımlamak için X_t süreci ele alındığında t zamanı ifade etmektedir ve $T = \{1, 2, \dots\}$ olmak üzere $\{t \in T\}$ şeklindedir. Burada $\{X_{t_1}, X_{t_2}, \dots, X_{t_k}\}$ ve $\{X_{t_1+h}, X_{t_2+h}, \dots, X_{t_k+h}\}$ 'nin ortak dağılım fonksiyonu $t_1, t_2, \dots, t_k$ 'nın her bir gözlem değeri için $k \leq s$ olmak üzere X_t süreci s . mertebeden güçlü durağandır. Ortalaması ve varyansı zamandan bağımsız X_t sürecinin kovaryans fonksiyonu sadece iki zaman noktası arasındaki uzaklığa bağlı ise zayıf durağan, diğer bir ifadeyle kovaryans durağandır denir⁴. Kovaryans durağan bir sürecin kovaryans fonksiyonu $cov(X_t, X_{t-s}) = \gamma_s$ şeklinde tanımlanır. Öte yandan zaman serilerinin durağanlık analizleri serilerin otokorelasyon fonksiyonları incelenerek yapılmaktadır. Otokorelasyonları inceleme işlemlerine korelogram analizi adı da verilmektedir. Korelogram analizinde serinin gecikmeli değerleri ile anlık değeri arasındaki korelasyonlara bakılmaktadır. Matematiksel olarak;

$$\rho_s = corr(X_t, X_{t-s}) = \gamma_s / \gamma_0 \quad (1.4)$$

² a.e., s.9.

³ M.H.Pesaran, **Time Series and Panel Data Econometrics**, Oxford, Oxford University Press, 2015, s.267.

⁴ Badi H. Baltagi, **Econometrics**, 4.bs., Berlin, Springer, 2008, s.356.

şeklinde ifade edilmektedir. Burada ρ_s , -1 ile $+1$ arasında değerler alan korelasyon katsayılarını ve γ_0 ise serinin varyansını göstermektedir. Uygulamada genellikle ana kütle otokorelasyon katsayıları bilinmediği için örnek otokorelasyonlar hesaplanmaktadır. Örnek otokorelasyon fonksiyonu $\hat{\rho}_s = \hat{\gamma}_s / \hat{\gamma}_0$ şeklinde ifade edilmektedir ve aşağıdaki eşitliklerin birbirine oranıdır:

$$\hat{\gamma}_0 = \frac{\sum_{t=1}^T (X_t - \bar{X})^2}{T} \quad (1.5)$$

$$\hat{\gamma}_s = \frac{\sum_{t=1}^{T-s} (X_t - \bar{X})(X_{t+s} - \bar{X})}{T} \quad (1.6)$$

Durağan bir süreç için gecikme sayısı s arttıkça otokorelasyon katsayıları hızlı bir şekilde azalmaktadır.

1.2. Bazı Stokastik Süreçler

Bu bölümde zaman serisi analizlerinde sıklıkla karşılaşılan bazı stokastik süreçler ele alınacaktır.

1.2.1. Pür Rastgele Süreç

Bir dizi bağımsız ve aynı dağılmış (i.i.d) rastgele değişkenler dizisinden oluşan X_t sürecine pür rastgele süreç diğer adıyla beyaz gürültü süreci denmektedir. Pür rastgele sürecin ortalaması ve varyansı zaman içinde değişmemektedir. Sürecin otokovaryans (γ_k) ve otokorelasyon fonksiyonları (ρ_k) sırasıyla aşağıdaki gibi ifade edilmektedir:

$$\gamma_k = cov(X_t, X_{t+k}) = 0 \quad k \neq 0 \quad (1.7)$$

$$\rho_k = \begin{cases} 1, & k = 0 \\ 0, & k \neq 0 \end{cases} \quad (1.8)$$

Beyaz gürültü süreci zayıf durağan bir süreçtir ve “hafızası” yoktur. Aynı zamanda X_t rastgele değişkeni normal dağılıma sahipse süreç güçlü durağandır⁵.

1.2.2. Hareketli Ortalama Süreci

Hareketli ortalama sürecine sahip X_t süreci aşağıdaki gibi ifade edilmektedir:

$$X_t = \beta_0 \varepsilon_t + \beta_1 \varepsilon_{t-1} + \cdots + \beta_q \varepsilon_{t-q} \quad (1.9)$$

Burada X_t serisi q . mertebeden hareketli ortalama sürecine sahiptir ve $MA(q)$ şeklinde gösterilmektedir. Burada $\beta_0 = 1$ ve $E(\varepsilon_t) = 0$ olduğu için $E(X_t) = 0$ ve $var(\varepsilon_t) = E(\varepsilon_t^2) = \sigma^2$ 'dir. X_t 'nin otokovaryans fonksiyonu aşağıdaki gibi yazılmaktadır:

$$\gamma_k = cov(X_t, X_{t-k}) = \begin{cases} \sigma^2 \sum_{i=0}^{q-k} \beta_i \beta_{i+k} & k = 0, 1, 2, \dots, q \\ 0 & k > q \end{cases} \quad (1.10)$$

Burada $\gamma_k = \gamma_{-k}$ şeklindedir. Diğer bir ifadeyle negatif ve pozitif otokovaryanslar birbirlerine simetriktir. X_t serisinin otokorelasyon fonksiyonu ise γ_k 'nin $var(X_t)$ 'ye bölünmesiyle elde edilmektedir. Kısacası otokorelasyon fonksiyonu ρ_k aşağıdaki gibi ifade edilmektedir:

$$\rho_k = \frac{\gamma_k}{\gamma_0} \quad (1.11)$$

Burada $\gamma_0 = var(X_t)$ 'dir. X_t serisinin otokovaryans fonksiyonu γ_k zamandan bağımsız olduğu için seri kovaryans durağandır. Hareketli ortalama sürecine sahip X_t süreci gecikme operatörü ile aşağıdaki gibi ifade edilmektedir:

⁵ Maddala and Kim, **a.g.e**, s.12.

$$X_t = (1 + \beta_1 L + \beta_2 L^2 + \dots + \beta_q L^q) \varepsilon_t = \beta(L) \varepsilon_t \quad (1.12)$$

Gecikme operatörü $L^j X_t = X_{t-j}$ şeklinde ifade edilmektedir. Böylelikle $LX_t = X_{t-1}$, $L^2 X_t = X_{t-2}$, $L^{-1} X_t = X_{t+1}$ şeklindedir. Yukarıda verilen (1.12) numaralı denklemin q tane kökü vardır ve aşağıdaki gibi yazılmaktadır:

$$X_t = (1 - \pi_1 L)(1 - \pi_2 L) \dots (1 - \pi_q L) \varepsilon_t \quad (1.13)$$

Burada $\pi_1, \pi_2, \dots, \pi_q$, $z^q + \beta_1 z^{q-1} + \dots + \beta_q = 0$ denkleminin kökleridir. (1.12) numaralı eşitlikteki modelinin parametrelerini tahmin ettikten sonra, $[\beta(L)]^{-1}$ 'nin yakınsak olması koşulu ile, modelin hataları $\varepsilon_t = [\beta(L)]^{-1} X_t$ eşitliğinden hesaplanabilir. Burada denklemin kökleri $|\pi_i| < 1$ sağladığı durumda MA(q) süreci sonsuz bir AR süreci şeklinde yazılabilmektedir. MA(2) süreci örnek olarak ele alındığında;

$$X_t = (1 + \beta_1 L + \beta_2 L^2) \varepsilon_t \quad (1.14)$$

π_1 ve π_2 , $z^2 + \beta_1 z + \beta_2 = 0$ denkleminin kökleridir. $|\pi_i| < 1$ eşitsizliğini aşağıdaki gibi gösterilebilmektedir:

$$\left| \frac{-\beta_1 \mp \sqrt{\beta_1^2 - 4\beta_2}}{2} \right| < 1 \quad (1.15)$$

Denklem (1.15)' ten hareketle β_1 ve β_2 'nin aşağıdaki koşulları

$$\beta_1 + \beta_2 > -1$$

$$\beta_2 - \beta_1 > -1$$

$$|\beta_2| < 1$$

sağlaması durumunda denklem (1.14)'deki MA(2) süreci sonsuz bir otoregresif süreç şeklinde yazılabilir.

1.2.3. Otoregresif Süreç

ε_t sürecinin 0 ortalama ve σ^2 varyansla pür rastgele sürece sahip olduğu düşünüldüğünde p . mertebeden otoregresif sürece sahip ve AR(p) şeklinde gösterilen X_t süreci aşağıdaki gibi ifade edilmektedir:

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \cdots + \alpha_p X_{t-p} + \varepsilon_t \quad (1.16)$$

Denklem (1.16)'daki AR(p) süreci gecikme operatörü ile aşağıdaki gibi yazılabilmektedir:

$$(1 - \alpha_1 L - \alpha_2 L^2 - \cdots - \alpha_p L^p) X_t = \varepsilon_t \quad (1.17)$$

Diğer bir ifadeyle;

$$\begin{aligned} X_t &= \frac{1}{(1 - \alpha_1 L - \alpha_2 L^2 - \cdots - \alpha_p L^p)} \varepsilon_t \\ &= \frac{1}{(1 - \pi_1 L)(1 - \pi_2 L^2) \cdots (1 - \pi_p L^p)} \varepsilon_t \end{aligned} \quad (1.18)$$

şeklinde yazılabilmektedir. Burada $\pi_1, \pi_2, \dots, \pi_p$, p . mertebeden $z^p - \alpha_1 z^{p-1} - \cdots - \alpha_p = 0$ denkleminin kökleridir. Denklemin kökleri, $|\pi_i| < 1$ olduğu durumda X_t serisi sonlu varyansa sahiptir. Aşağıda örnek olması açısından bir AR(2) sürecinin durağanlık koşulu ve otokorelasyon fonksiyonları incelenmiştir.

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \varepsilon_t \quad (1.19)$$

π_1 ve π_2 , $z^2 - \alpha_1 z - \alpha_2 = 0$ denkleminin karesel kökleridir ve aşağıdaki gibi elde edilebilmektedir.

$$|\pi_i| = \left| \frac{\alpha_1 \mp \sqrt{\alpha_1^2 + 4\alpha_2}}{2} \right| < 1 \quad (1.20)$$

Denklem (1.20)'den hareketle X_t serisinin durağan olabilmesi için aşağıdaki koşulları sağlaması gerekmektedir:

$$\alpha_1 + \alpha_2 < 1$$

$$\alpha_1 - \alpha_2 > -1$$

$$|\alpha_2| < 1$$

Bu koşullar altında seri durağandır denir. Öte yandan serinin otokorelasyon fonksiyonlarını bulmak için, denklem (1.16)'in her iki tarafı X_{t-k} ile çarpıldıktan sonra beklenen değerleri alınıp, γ_0 yani $var(X_t)$ ye bölüldüğünde aşağıdaki eşitlik elde edilmektedir⁶:

$$\rho_k = \alpha_1 \rho_{k-1} + \dots + \alpha_k \rho_{k-p}$$

AR(2) süreci için bu işlemler yapıldığında;

$$\rho_1 = \alpha_1 \rho_0 + \alpha_2 \rho_{-1}$$

$$\rho_2 = \alpha_1 \rho_1 + \alpha_2 \rho_0$$

⋮

$$\rho_k = \alpha_1 \rho_{k-1} + \alpha_2 \rho_{k-2}$$

Yukarıdaki eşitliklerde $\rho_1 = \rho_{-1}$ eşitliği kullanılarak işlemler yapılmıştır.

⁶ a.e, s.14.

Sonuç olarak, bir $MA(q)$ süreci gecikme sayısı (q) kadar otokorelasyonlara sahiptir ve gecikme sayısından daha büyük değerler için otokorelasyonlar sıfırdır denilebilmektedir. $AR(p)$ süreci için ise otokorelasyon katsayıları gecikme sayısı arttıkça üstel olarak sıfıra doğru azalmaktadır.

1.2.4. Otoregresif Hareketli Ortalama Süreci

AR ve MA süreçlerine sahip modeller ile analizler yapılırken iyi tahminler elde edebilmek için modelde çok fazla parametre yer alabilmekte ve bu durum yapılan tahminlerde etkinlik kaybına neden olmaktadır⁷. Dolayısıyla bu durumdan kaçınmak için AR ve MA modellerinin bir kombinasyonu olarak belirtilen otoregresif hareketli ortalama (ARMA) süreçleri kullanılabilmektedir. $ARMA(p, q)$ modeli aşağıdaki gibi ifade edilmektedir:

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \dots + \beta_q \varepsilon_{t-q} \quad (1.21)$$

Burada ε_t sıfır ortalama ve σ^2 sabit varyansla beyaz gürültü sürecine sahiptir. Gecikme operatörü kullanılarak süreç aşağıdaki gibi yazılabilmektedir:

$$\phi(L)X_t = \theta(L)\varepsilon_t$$

$$\phi(L) = 1 - \alpha_1 L - \dots - \alpha_p L^p$$

$$\theta(L) = 1 + \beta_1 L + \dots + \beta_q L^q$$

Burada durağanlık için ön koşul $\phi(L)$ 'nin köklerinin birim çember dışında olmasıdır. MA bileşenlerinin ters çevrilebilirliği için ön koşul ise $\theta(L)$ 'nin köklerinin birim çember dışında olmasıdır. ARMA modellerinin otokorelasyon fonksiyonlarının elde edilişi AR ve MA modellerine göre çok daha karmaşık olduğu için burada sadece $ARMA(1,1)$ sürecinin otokorelasyon fonksiyonları incelenecektir ve $ARMA(1,1)$ süreci;

⁷ Wei, William WS. , **Time series analysis**, 2.bs, Boston, Pearson Addison Wesley, 2006, s. 57

$$X_t = \alpha X_{t-1} + \varepsilon_t + \beta \varepsilon_{t-1} \quad (1.22)$$

olarak ifade edilmektedir. Süreç gecikme operatörü yardımıyla aşağıdaki gibi yazılmaktadır:

$$X_t = \alpha X_{t-1} + \varepsilon_t + \beta \varepsilon_{t-1} \quad (1.23)$$

$$(1 - \alpha L)X_t = (1 + \beta L)\varepsilon_t$$

veya;

$$X_t = \frac{1 + \beta L}{1 - \alpha L} \varepsilon_t = (1 + \beta L)(1 + \alpha L + \alpha^2 L^2 + \dots) \varepsilon_t \quad (1.24)$$

$$= [1 + (\alpha + \beta)L + \alpha(\alpha + \beta)L^2 + \alpha^2(\alpha + \beta)L^3 + \dots] \varepsilon_t$$

Burada ε_t sıfır ortalamalı ve σ^2 sabit varyanslı beyaz gürültü sürecidir. X_t serisinin varyansı ve kovaryansı aşağıdaki gibi elde edilmektedir;

$$var(X_t) = [1 + (\alpha + \beta)^2 + \alpha^2(\alpha + \beta)^2 + \dots] \sigma^2 \quad (1.25)$$

$$= \left[1 + \frac{(\alpha + \beta)^2}{1 - \alpha^2} \right] \sigma^2 = \frac{1 + \beta^2 + 2\alpha\beta}{1 - \alpha^2} \sigma^2$$

$$cov(X_t, X_{t-1}) = [(\alpha + \beta) + \alpha(\alpha + \beta)^2 + \alpha^3(\alpha + \beta)^2 + \dots] \sigma^2 \quad (1.26)$$

$$= \left[\alpha + \beta + \frac{(\alpha + \beta)^2 \alpha}{1 - \alpha^2} \right] \sigma^2$$

$$= \frac{(\alpha + \beta)(1 + \alpha\beta)}{1 - \alpha^2} \sigma^2$$

Birinci mertebe için otokorelasyon fonksiyonu:

$$\rho_1 = \frac{cov(X_t, X_{t-1})}{var(X_t)} = \frac{(\alpha + \beta)(1 + \alpha\beta)}{1 + \beta^2 + 2\alpha\beta} \quad (1.27)$$

şeklindedir.

1.2.5. Bütünleşik Otoregresif Hareketli Ortalama Süreci

Bütünleşik otoregresif hareketli ortalama (ARIMA) süreci otoregresif süreç, bütünleşik süreç ve hareketli ortalama süreçlerinden oluşmaktadır. Burada bütünleşik süreç kavramıyla ilgili anlatılmak istenen durağan olmayan serinin kaç kere farkı alınarak durağan hale getirildiğidir ve $I(d)$ şeklinde gösterilmektedir. Bu noktadan hareketle (d) parametresi negatif olmayan tamsayı değerleri almaktadır. Rastgele süreç durağan ise d parametresi 0 değerini almakta ve seri $I(0)$ şeklinde ifade edilmektedir. Diğer bir ifadeyle süreç $ARIMA(p, 0, q)$ sürecine sahiptir. Eğer seri durağan değilse ve birinci mertebe farkı alındığında durağan hale geliyorsa burada $d = 1$ ve süreç $ARIMA(p, 1, q)$ olarak gösterilmektedir. Süreç durağan hale getirildikten sonra sürecin $AR(p)$ ve $MA(q)$ bileşenlerini elde etmek için gerekli analizler yapılmaktadır. Otoregresif ve hareketli ortalama bileşenlerine sahip bir ARMA süreci önceki bölümlerde anlatıldığı gibi aşağıdaki şekilde ifade edilmektedir:

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \dots + \beta_q \varepsilon_{t-q} \quad (1.28)$$

(1.28) gecikme operatörü ile tekrardan aşağıdaki gibi yazılmaktadır:

$$\phi(L)X_t = \theta(L)\varepsilon_t$$

Burada;

$$\phi(L) = 1 - \alpha_1 L - \dots - \alpha_p L^p$$

$$\theta(L) = 1 + \beta_1 L + \dots + \beta_q L^q$$

olarak belirtilmektedir. Sürecin durağan olması için denklemin köklerinin (z) $|z| \leq 1$ ve $\phi(z) \neq 0$, ters çevrilebilir olması için ise $\theta(z) \neq 0$ ve $|z| \leq 1$ olması gerekli koşuldur. ARMA süreci durağan olduğunda $MA(\infty)$ süreci şeklinde, ters çevrilebilir olduğunda ise $AR(\infty)$ şeklinde ifade edilebilmektedir.⁸

Yukarıda ifade edilen ARMA sürecine sahip X_t , durağan değilse serinin d kez farkı alınarak durağan hale getirilmekte ve $ARIMA(p, d, q)$ şeklinde gösterilmektedir. Öte yandan seri $\Delta^d X_t \sim ARMA(p, q)$ şeklinde de gösterilmektedir. Bu noktada basitlik olması açısından $ARIMA(1,1,1)$ modeli ele alınarak aşağıdaki gibi ifade edilmektedir:

$$(1 - L)(X_t - \mu) = \phi_1(X_{t-1} - \mu) + \varepsilon_t - \theta_1\varepsilon_{t-1} \quad (1.29)$$

$$(1 - L)(1 - \phi_1 L)X_t = (1 - \phi_1)\mu + (1 - \theta_1 L)\varepsilon_t \quad (1.30)$$

Burada X_t ve ε_t sırasıyla t dönemindeki gözlem değeri ve serideki şoktur. Modelde μ serilerin ortalamasını, diğer bir ifadeyle modeldeki sabit terimi, θ_1 hareketli ortalama sürecinin parametresini, ϕ_1 otoregresif sürecin parametresini ifade etmektedir.

1.3. Uzun Hafıza Süreci

Basitçe ifade etmek gerekirse hafıza, zaman serilerinde birbirlerine uzak iki nokta arasındaki istatistiksel bağımlılığın seviyesini ifade etmektedir. Bu bağımlılığı incelemenin bir yolu serilerin otokorelasyon fonksiyonlarını analiz etmektir. Buradan hareketle birbirine uzak gözlemler arasındaki bağımlılık hızlı bir azalma gösterdiğinden dolayı, otokorelasyon fonksiyonları üstel bir azalma gösteren durağan zaman serileri kısa hafızalı olarak belirtilmektedir. Öte yandan kısa hafızaya sahip durağan zaman serilerinde bir şok meydana geldiğinde seri kısa bir süre sonra ortalamasına geri dönmektedir. Bununla birlikte, otokorelasyon fonksiyonları üstel

⁸ Lütkepohl, H., Krätzig, M., & Phillips, P. C. (Eds.), **Applied time series econometrics**, U.S.A. , Cambridge University Press, 2004, s.27.

bir azalmadan daha yavaş bir azalma gösteren, diğer bir ifadeyle hiperbolik olarak azalan seriler uzun hafızalıdır. Bu ilişkiye serilerdeki uzun dönemli bağımlılık adı da verilmektedir. Uzun hafızalı serilerde bir şok meydana geldiğinde ise serinin ortalamasına geri dönmesi, kısa hafızalı durağan süreçlere oranla daha geç olmaktadır. Serilerin bu otokorelasyon yapısını dikkate alan modeller, zaman serisi analizlerinde durağan olmayan serilerin durağanlaştırılması için fark parametresi d 'nin kesirli değerler alınmasına izin vererek sağlamaktadır. Bu noktada karşımıza literatürde genişçe yer bulan ARFIMA modelleri çıkmaktadır. Granger ve Joyeux (1980) ve Hosking (1981) tarafından geliştirilen kesirli otoregresif bütünleşik hareketli ortalama süreci (ARFIMA) modelleri d parametresinin hem kesirli değerler alabilmesine hem de serilerin uzun hafıza veya uzun dönemli bağımlılık özelliğini sergilemesine izin vermiştir. ARFIMA modelleri, ARIMA ve ARMA tipi modellerin daha kapsamlı bir versiyonu olarak düşünülebilir. ARFIMA modelleri kullanılarak hem d parametresi tahmin edilmekte hem de ARMA tipi modellerin sadece kısa dönemli bağımlılığı modellemesinin aksine, d parametresinin kesirli değerler almasına izin verilerek uzun dönem bağımlılıkla birlikte kısa dönemli bağımlılığı da modelleyebilme özelliğine sahip olmaktadır.

ARIMA tipi modeller gerçekte durağan olan fakat otokorelasyon fonksiyonları durağan serilere göre çok daha yavaş hızda azalan seriler için $d = 1, 2, \dots$ kez fark alarak durağan hale gelmektedir. Bu durum serilerin aşırı fark alınmasına neden olmakta ve uygulamalı çalışmalarda gerçeği yansıtmayan sonuçlara sebep olmaktadır. Bu problem ARFIMA modelleri kullanılarak ortadan kaldırılmaktadır.

ARFIMA(p, d, q) sürecine sahip bir X_t değişkeni aşağıdaki gibi ifade edilmektedir:

$$\phi(L)(1-L)^d X_t = \theta(L)\varepsilon_t \quad \varepsilon_t \sim i.i.d(0, \sigma^2) \quad (1.31)$$

Burada;

$$\phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p$$

$$\theta(L) = 1 + \theta_1 L - \dots - \theta_p L^p$$

gösterilmektedir. Gerekli düzenlemeler yapıldığında denklem (1.31) aşağıdaki şekilde ifade edilmektedir:

$$X_t = (1 - L)^{-d} \phi(L)^{-1} \theta(L) \varepsilon_t \quad (1.32)$$

(1.32) numaralı eşitlikte $\phi(L)^{-1} \theta(L) \varepsilon_t$ kısmı X_t serisinin ARMA parametrelerini oluşturmakta ve bu parametrelerle serinin kısa hafıza özelliği modellenmektedir. Kesirli değerler alan fark parametresi d ile de serinin uzun hafıza özelliği modellenmektedir. ARFIMA modelleri kısa hafıza ve uzun hafıza özelliklerini ayrı ayrı modelleyebildiğinden dolayı ARIMA modellerine göre kullanımı çok daha basittir.

ARFIMA modelleri ARIMA modellerinin d parametresinin kesirli değerler almasına izin verilerek genişletilmiş halidir⁹. $\Gamma(\cdot)$ gamma fonksiyonu olmak üzere ARFIMA süreci için kesirli fark terimi $(1 - L)^d$ aşağıdaki şekilde yazılabilmektedir:

$$(1 - L)^d = \sum_{k=0}^{\infty} \frac{\Gamma(k - d) L^k}{\Gamma(-d) \Gamma(k + 1)} \quad (1.33)$$

Denklem (1.33)'teki kesirli fark parametresi (d), X_t serisinin durağanlığı ve hafıza özelliği üzerindeki etkisini göstermektedir. Bu noktadan hareketle ARFIMA sürecine sahip X_t serisinin durağanlık özelliklerinin incelenmesi noktasında sürecinin otokovaryans fonksiyonları γ_k ve otokorelasyon fonksiyonları ρ_k aşağıdaki gibi ifade edilmektedir:

⁹ Jan Beran, **Statistic For Long Memory Process**, Germany, Chapman & Hall, s.59.

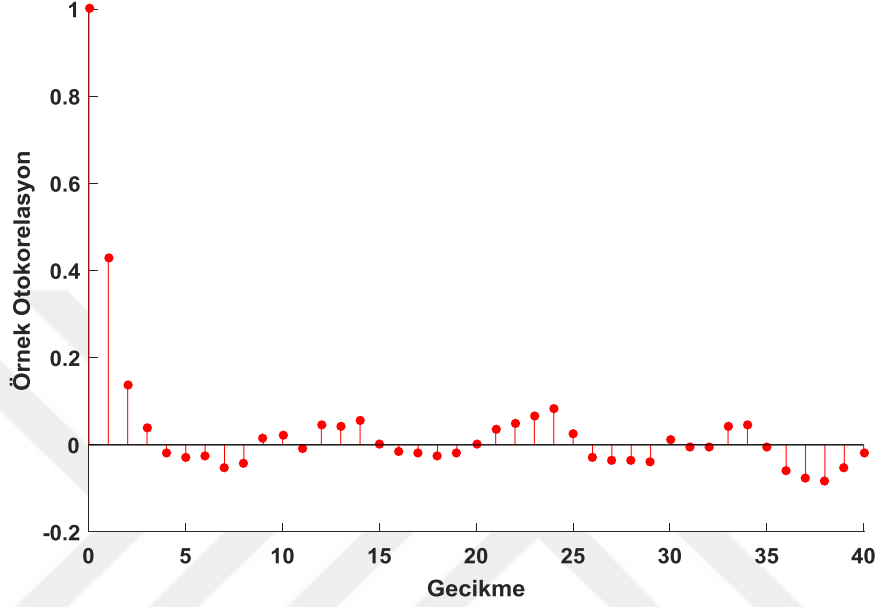
$$\gamma_0 = \frac{\sigma^2 \Gamma(1-2d)}{\Gamma(1-d)\Gamma(1-d)} \quad (1.34)$$

$$\gamma_k = \frac{\sigma^2 \Gamma(1-2d) \Gamma(k+d)}{\Gamma(d) \Gamma(1-d) \Gamma(k+1-d)} \quad k = \mp 1, \mp 2 \dots \quad (1.35)$$

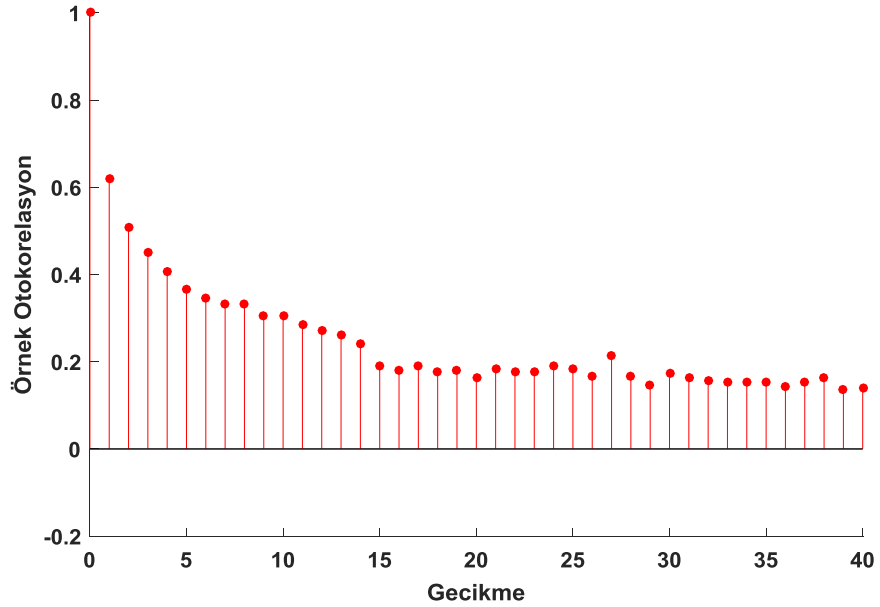
$$\rho_k = \frac{\Gamma(1-d) \Gamma(k+d)}{\Gamma(d) \Gamma(k+1-d)} \quad (1.36)$$

Burada kesirli fark parametresinin $0 < d < 0.5$ aralığında olduğu durumda seri durağan olmakla beraber ortalamasına geri dönme eğilimindedir. Fakat " d " parametresi $0.5 < d < 1$ aralığında ise X_t serisinin varyansı sonlu toplanabilir olmadığından ters çevrilebilir olmamakla beraber durağan olmayan bir süreçte sahiptir. Fakat seri durağan olmamasına rağmen serinin otokorelasyon fonksiyonları gecikme sayısı arttıkça sifıra doğru yaklaşmaktadır. Yani seri zayıf bağımlılık özelliğini korumaktadır. Bundan dolayı süreç sonlu fakat uzun hafızaya sahiptir, yani seride bir şok meydana geldiğinde seri uzun dönemde ortalamasına geri dönme eğilimindedir. $d > 1$ durumunda ise seri durağan olmamakla beraber ortalamasına geri dönmemektedir. Bu nedenle ARFIMA sürecinin hafıza özelliği önemli bir biçimde d parametresinin değerine bağlıdır ve sürecin otokorelasyon fonksiyonları durağan bir süreçte göre çok daha yavaş azalmaktadır. Uzun hafıza ile kısa hafıza süreçlerinin otokorelasyon fonksiyonları arasındaki bu farkı göstermek amacıyla parametre değeri $\emptyset = 0.4$ durağan AR(1) süreci ile fark parametresi $d = 0.4$ olan durağan ARFIMA(0,0.4,0) sürecinin otokorelasyon fonksiyonları Şekil 1.1 ve Şekil 1.2'de sırasıyla gösterilmektedir.

Şekil 1.1 AR(1) Sürecinin Otokorelasyon Fonksiyonu



Şekil 1.2 ARFIMA(0,0.4,0) Sürecinin Otokorelasyon Fonksiyonu



Yukarıdaki grafiklerde görüldüğü üzere durağan bir ARFIMA sürecinin otokorelasyon fonksiyonu durağan AR sürecine göre çok daha yavaş azalmaktadır.

Öte yandan serilerin gerçekten kesirli bütünleşik bir sürece sahip olup olmadıklarının incelenmesi gerekmektedir. Bu durum için kullanılan yöntemlerden birisi de birim kök testleridir. Klasik birim kök testlerinde serilerin bütünleşme derecelerinin (d) 0 ve 1 gibi tam sayı olduğu varsayılmaktadır. Bu bağlamda $d = 1$ durumu serinin durağan olmadığını ve bir kere fark alınarak durağan hale geldiğini belirtmektedir. Bir kere farkı alınarak durağan hale gelen serinin bütünleşme derecesi $I(1)$ şeklinde ifade edilmektedir. Klasik birim kök testleri, birim kökün varlığı araştırılırken serilerin bütünleşme derecelerinin 0 ya da 1 olması gibi kısıtlamalar getirmektedir. Baillie (1996), çalışmasında bu durumun çok kısıtlayıcı olabileceğini ifade etmektedir. Klasik birim kök testleri, seri aslında uzun dönemde dengeye gelmesine rağmen sanki durağan değilmiş gibi sonuçlar vermektedir. Dolayısıyla testlerin gücünün azalmaktadır. Bu nedenle serilerin bütünleşme derecelerinin tam sayı değerler yerine kesirli değerler almasına izin veren testler geliştirilmiştir.

1.3.1. Uzun Hafıza ve Yapısal Kırılma Arasındaki İlişkiler

Gerçekte birçok makroekonomik değişken durağan olmamasına rağmen bu durumun kaynağının belirlenmesi oldukça zordur. Öte yandan durağan olmayan, diğer adıyla bütünleşik süreçler ile yapısal kırılmaya maruz kalmış kısa hafızalı modellerin veri yapıları oldukça benzerlik göstermektedir. Bundan dolayı analizler yapılırken bu süreçleri birbirlerinden ayırmak zordur. Teknik olarak ifade etmek gerekirse süreçlerin otokorelasyon fonksiyonları benzerlik göstermektedir. Uzun hafızalı modeller kullanılarak yapılan analizlerde de bu duruma benzer bazı problemler ortaya çıkmaktadır. Granger ve Hyung (2004), yaptıkları çalışmada yapısal kırılmaya sahip kısa hafızalı bir sürecin otokorelasyon fonksiyonları ile uzun hafızalı bir sürecin otokorelasyon fonksiyonlarının birbirlerine benzerlik gösterdiğini ortaya koymaktadırlar. Diğer bir ifadeyle, yapısal kırılmalı kısa hafızalı bir sürecin otokorelasyon fonksiyonu, tıpkı uzun hafızalı modellerde olduğu gibi hiperbolik olarak, yani çok yavaş bir şekilde azalmaktadır. Bu durum literatürde sahte uzun hafıza olarak bilinmektedir. Bundan dolayı uzun hafızanın gerçekten verinin yapısından mı kaynaklandığı yoksa serideki yapısal kırılmaların mı uzun hafızaya neden olduğunun araştırması gerekmektedir. Bu iki olgu arasındaki bağlantıyı ortaya

koymak için Granger ve Hyung (2004), çalışmalarında yapısal kırılmalı kısa hafızalı süreci diğer adıyla rastgele yapısal kırılmalı süreci aşağıdaki gibi ifade etmektedirler;

$$Y_t = m_t + \varepsilon_t \quad t = 1, \dots, T \quad (1.37)$$

$$m_t = m_{t-1} + q_t \eta_t \quad (1.38)$$

Burada Y_t süreci yapısal kırılmaya sahip ARMA sürecidir¹⁰. Serinin yapısal kırılma kısmı, rastgele bir sürece sahiptir ve m_t ile ifade edilmektedir. q_t ve η_t , m_t sürecini oluşturan birbirinden bağımsız iki bileşendir. q_t binom dağılımına sahiptir ve kırılma tarihini ifade etmektedir. η_t ise birbirinden bağımsız ve aynı dağılmış sıfır ortalamalı σ_η^2 varyanslı bir dağılıma sahiptir ve kırılmanın şiddeti olarak ifade edilmektedir. Binom dağılımına sahip q_t rastgele değişkeni aşağıdaki gibi ifade edilmektedir:

$$q_t = \begin{cases} 0, & 1 - p \\ 1, & p \end{cases}$$

Burada p yapısal kırılmaların gerçekleşme olasılığıdır. Gerçekte bu olasılık değeri bilinmediği için tahmin edilmesi gerekmektedir. Olasılığın tahmini aşağıdaki formül yardımıyla yapılmaktadır:

$$\hat{p} = \frac{y}{T}$$

Burada $\hat{p}$, p ' nin sapmasız bir tahmin edicidir. 'y' binom dağılımındaki başarı sayısı olmakla birlikte burada verideki kırılma sayısıdır. Asimptotik olarak, diğer bir ifadeyle $T \rightarrow \infty$ durumunda kırılmaların olası değeri $p \rightarrow 0$ şeklinde ifade edilmektedir. Fakat $T \rightarrow \infty$ durumunda kırılmaların beklenen değeri Tp , sıfır olmayan sonlu bir değere yakınsamaktadır. Bu varsayım, yukarıdaki denklem (1.37)' de belirtilen sürecin hafıza özelliklerinin varlığını sürdürmesi için, örnekteki gözlem

¹⁰ Chen, C., & Tiao, G. C. , "Random level-shift time series models, ARIMA approximations, and level-shift detection", **Journal of Business & Economic Statistics**, 1990, 8(1), ss.85.

sayısı artarken yapısal kırılmaların sayısının sonlu olması gerektiğini ifade etmektedir. Denklem (1.37) ve denklem (1.38) birlikte;

$$Y_t = m_0 + q_1\eta_1 + \dots + q_t\eta_t + \varepsilon_t \quad (1.39)$$

şeklinde ifade edilebilmektedir. Burada ε_t , η_t ve q_t birbirlerinden bağımsız süreçlerdir ve dolayısıyla aralarında doğrusal herhangi bir ilişki yoktur. Kolaylık olması açısından $t \leq 0$ durumu için $m_0 = 0$, $q_t = 0$, $\varepsilon_t = 0$ ve $\eta_t = 0$ değerlerini almaktadır. Bu durumda $E(Y_t) = 0$ ve $var(Y_t) = Tp\sigma_\eta^2 + \sigma_\varepsilon^2$ şeklindedir. Y_t ve Y_{t+k} arasındaki kovaryans; $cov(Y_t, Y_{t+k}) = Tp\sigma_\eta^2$ gibi yazılmaktadır. Burada Y_t sürecinin k gecikmesi için otokorelasyon fonksiyonu aşağıdaki gibidir:

$$corr(Y_t, Y_{t+k}) = \frac{Tp\sigma_\eta^2}{\sqrt{Tp\sigma_\eta^2 + \sigma_\varepsilon^2} \sqrt{(T+k)p\sigma_\eta^2 + \sigma_\varepsilon^2}} \quad (1.40)$$

Matematiksel işlemler sonucunda $corr(Y_t, Y_{t+k}) = \left(1 + \frac{\sigma_\varepsilon^2}{Tp\sigma_\eta^2}\right)^{-1}$ şeklinde ifade edilebilmektedir. Açık bir şekilde de görüldüğü gibi $T \rightarrow \infty$ durumunda;

$$\left(1 + \frac{\sigma_\varepsilon^2}{Tp\sigma_\eta^2}\right)^{-1} > 0 \quad (1.41)$$

sıfırdan büyük pozitif bir değerdir. Dolayısıyla bu durum yapısal kırılmalı modelin otokorelasyon fonksiyonları ile uzun hafızalı modelin otokorelasyon fonksiyonları arasındaki benzerliği ortaya koymaktadır¹¹. Burada yapısal kırılmaların beklenen değeri Tp ve σ_η^2 , sürecin hafıza özelliğini ortaya koyan iki bileşendir. Tp ve σ_η^2 arttıkça sürecin otokorelasyon fonksiyonunun değerleri artarak bire doğru yaklaşmaktadır. Bu iki bileşen küçük olduğu durumda otokorelasyon fonksiyonu birden uzak bir görüntü sergilemesine rağmen otokorelasyonlar daima uzun hafızada olduğu gibi çok yavaş bir şekilde azalmaktadır.

¹¹ Granger and Hyung, **a.g.e**, s.403.

İKİNCİ BÖLÜM

MONTE CARLO SİMÜLASYON TEKNİĞİ KULLANILARAK KESİRLİ BİRİM KÖK TESTLERİNİN PERFORMANSLARININ İNCELENMESİ

Bu bölümde, durağanlık analizi yapılırken analizlerde yaygın bir şekilde kullanılan Dickey ve Fuller (D-F) testi ile Genişletilmiş Dickey Fuller (AD-F) testi ve uzun hafızanın varlığı araştırılırken kullanılan, Dolado, Gonzalo ve Mayoral (2002) (DGM) tarafından geliştirilen kesirli Dickey Fuller (FD-F) ve genişletilmiş kesirli Dickey Fuller (AFD-F) testi anlatılacaktır. Daha sonra kesirli birim kök testlerinin sınırlı veri kümesinde performansları Monte Carlo simülasyonları ile incelenecektir. Bu bölümün son aşamasında Sieve bootstrap yönteminden bahsedildikten sonra bu yöntem kullanılarak kesirli Dickey Fuller testinin bozulan 1. tip hata olasılığının iyileştirildiği Monte Carlo simülasyonları ile gösterilecektir. Tezin geri kalan kısmında kolaylık olması açısından testler anlatılırken uzun yazımlarının yerine kısaltmaları kullanılacaktır.

2.1. Dickey Fuller ve Genişletilmiş Dickey Fuller Testleri

Zaman serisi analizinde değişkenlerin durağan olup olmadığı araştırılırken, serilerde birim kökün varlığının test edilmesi gerekmektedir. Eğer seri yapısında birim kök barındırıyorsa durağan dışı olmakta ve durağan olmayan serilerle regresyon analizleri yapıldığında değişkenler arasında gerçekte olmayan ilişkiler meydana gelmektedir. Bu durumdan kaçınmak için serilerin durağan olup olmadığı, diğer bir ifadeyle serinin birim kök içerip içermediğinin araştırılması gerekmektedir. Öte yandan durağan olmayan serilerde bir şok meydana geldiğinde şoklar kalıcı olmakta ve seri ortalamasına geri dönmemektedir. Bunun sonucu olarak da serilerin şimdiki ve geçmiş dönemlerdeki gözlemleri kullanılarak ileriye dönük yapılan tahminler gerçeği yansıtmamaktadır. Dolayısıyla serilerde birim kök olduğu

durumda, fark alınarak durağan hale getirildikten sonra analizlerin yapılması gerekmektedir.

Buradan hareketle zaman serisi analizlerinde birim kökün önemini anlamak için deterministik bileşen içermeyen AR(1) sürecine sahip bir X_t serisi ele alındığında sürecin matematiksel ifadesi aşağıdaki gibidir:

$$X_t = \phi X_{t-1} + \varepsilon_t \quad (2.1)$$

$$X_{t-1} = \phi X_{t-2} + \varepsilon_{t-1} \quad (2.2)$$

$$X_{t-2} = \phi X_{t-3} + \varepsilon_{t-2} \quad (2.3)$$

⋮

Denklem (2.2), (2.1)' de yerine konularak yazıldığında aşağıdaki eşitlik elde edilir:

$$X_t = \phi(\phi X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \quad (2.4)$$

$$X_t = \phi^2 X_{t-2} + \phi \varepsilon_{t-1} + \varepsilon_t \quad (2.5)$$

Daha sonra, (2.3), (2.5)'de yerine konulduğu durumda aşağıdaki eşitlikler elde edilmektedir:

$$X_t = \phi^2(\phi X_{t-3} + \varepsilon_{t-2}) + \phi \varepsilon_{t-1} + \varepsilon_t \quad (2.6)$$

$$X_t = \phi^3 X_{t-3} + \phi^2 \varepsilon_{t-2} + \phi \varepsilon_{t-1} + \varepsilon_t \quad (2.7)$$

⋮

$$X_t = \phi^{T+1} X_{t-(T+1)} + \phi \varepsilon_{t-1} + \phi^2 \varepsilon_{t-2} + \phi^3 \varepsilon_{t-3} + \cdots + \phi^T \varepsilon_{t-T} + \varepsilon_t \quad (2.9)$$

Buradan hareketle üç tane sonuç çıkmaktadır:

- (i) $T \rightarrow \infty$ durumunda $\phi < 1$ ise $\phi^T \rightarrow 0$ olmaktadır. Seride bir şok meydana geldiğinde şokların etkisi zamanla ortadan kaybolmaktadır ve bu durum daha önce de bahsedildiği gibi durağan zaman serilerinin bir özelliğidir.
- (ii) $\phi = 1$ durumunda $\phi^T = 1$ 'dir ve $T \rightarrow \infty$ durumunda aşağıdaki gibi ifade edilmektedir:

$$X_t = X_0 + \sum_{t=0}^{\infty} \varepsilon_t$$

Burada X_t süreci X_0 gibi bir başlangıç noktası ve sonsuz bir geçmiş dönem şoklarının toplamından oluşmaktadır. $\phi = 1$ durumunda denklem (2.1)'in karakteristik kökü 1'dir ve durum karşımıza birim kök süreci olarak çıkmaktadır. Matematiksel olarak aşağıdaki gibi ifade edilmektedir:

$$X_t = X_{t-1} + \varepsilon_t \quad (2.10)$$

$$X_t - X_{t-1} = \varepsilon_t \quad (2.11)$$

(2.11) gecikme operatörü ile aşağıdaki gibi yazılabilmektedir:

$$(1 - L)X_t = \varepsilon_t \quad (2.12)$$

$$\Delta X_t = \varepsilon_t \quad (2.13)$$

Denklem (2.11)'in karakteristik kökü $(1 - z) = 0$ eşitliği ile elde edilir ve denklem çözüldüğünde $z = 1$ 'dir. Sonuç olarak X_t serisi birim köke sahiptir ve dolayısıyla durağan değildir.

- (iii) $\phi > 1$ durumunda ise zaman geçtikçe serideki şoklar çok daha büyük etkiye sahip olmaktadır. Bunun sebebi ise $\phi^T > \phi^{T-1} > \phi^{T-2} > \dots$, eşitsizliğinden dolayıdır. Bu duruma literatürde “*patlayan süreç*” denilmektedir.

Zaman serisinde birim kökün olup olmadığı araştırılırken başvuru yöntemlerinden birisi de serilerin otokorelasyon fonksiyonlarının incelenmesidir. Fakat, birim köke sahip diğer adıyla durağan olmayan serilerde, şoklar sisteme kalıcı olmakla beraber otokorelasyon fonksiyonları çok yavaş bir şekilde sıfıra doğru gidebilmektedir. Bu nedenle birim kök analizi için serilerin otokorelasyon fonksiyonları incelenirken dikkatli olunması gerekmektedir. Çünkü bazı durağan serilerin de otokorelasyon fonksiyonları çok yavaş bir şekilde azalarak sıfıra doğru gitmektedir¹². Bundan dolayı serilerde birim kökün varlığı araştırılırken otokorelasyon veya kısmi otokorelasyon fonksiyonlarının kullanılması yanıltıcı sonuçlara sebep olabilmektedir. Sonuç olarak yapılan analizlerin güvenilir olması açısından daha formel testlere ihtiyaç duyulmaktadır. Bu noktadan hareketle literatürde birim kök analizi için birçok test geliştirilmiştir. Bunlardan birisi D-F testidir. Testte temel alınan denklem aşağıdaki gibidir:

$$X_t = \phi X_{t-1} + \varepsilon_t \quad (2.14)$$

Durağanlık analizinde denklem (2.14)' de ϕ parametresi, temel hipotez altında $\phi = 1$, alternatif hipotez altında $\phi < 1$ e karşı sınanmaktadır. Burada boş hipotez seride birim kök olduğunu diğer adıyla serinin durağan olmadığını ifade etmektedir. Hesaplama kolaylığı açısından denklem (2.14) yerine durağanlık analizi için aşağıdaki model kullanılmaktadır:

$$\Delta X_t = \psi X_{t-1} + \varepsilon_t \quad (2.15)$$

Burada $\psi = \phi - 1$ dir ve boş hipotez altında $\psi = 0$ eşitliği sınanmaktadır. Model deterministik bileşenlerin de dahil edildiği en geniş haliyle aşağıdaki gibi ifade edilmektedir:

$$\Delta X_t = \psi X_{t-1} + \mu + \lambda t + \varepsilon_t \quad (2.16)$$

¹² Brooks, C. , **Introductory Econometrics For Finance**, 2.bs, UK, Cambridge University Press, 2008, s.327.

Burada μ ve t sırasıyla sabit terimi ve deterministik trendi ifade etmektedir. Test istatistiği $\hat{\psi}/sh(\hat{\psi})$ şeklinde elde edilmektedir ve Dickey ve Fuller (1979), yaptıkları çalışmada D-F test istatistiğinin sola çarpık bir dağılıma sahip olduğunu göstermişlerdir.

D-F testinde modelin hata teriminin otokorelasyonsuz olduğu varsayılmaktadır. Diğer bir ifadeyle hata terimi beyaz gürültü sürecine sahiptir. Fakat yapılan analizlerde bu durumun gerçeği yansıtmadığı görülmüştür. Modelin bağımlı değişkeninin (ΔX_t) gecikmeli değerlerinin modele dahil edilmediği durumda hata terimleri arasında otokorelasyon olabileceği yapılan çalışmalar sonucunda ortaya konulmuştur. Bu problemi çözmek için araştırmacılar bağımlı değişkenin gecikmeli değerlerini modele dahil etmişlerdir. Bu noktadan hareketle bağımlı değişkenin gecikmeli değerlerinin yer aldığı en geniş model aşağıdaki gibi yazılabilir:

$$\Delta X_t = \mu + \lambda t + \psi X_{t-1} + \sum_{i=1}^p \alpha_i \Delta X_{t-i} + \varepsilon_t \quad (2.17)$$

Burada ‘ p ’, modelde yer alan bağımlı değişkenin gecikmeli değerlerinin sayısıdır. Denklem (2.17), literatürde AD-F eşitliği olarak ifade edilmektedir. Burada bağımlı değişkenin gecikmeli değerleri modele ilave edilerek hata terimindeki otokorelasyon problemi çözülmeye çalışılmaktadır. D-F testinde olduğu gibi burada da ψ parametresinin anlamlılığı sınanmaktadır. AD-F testi için kullanılan kritik değerler, asimptotik olarak D-F testi için kullanılan kritik değerler ile aynıdır. Dolayısıyla, test istatistiğinin anlamlılığının sınanması için kurulan hipotezler de D-F testi ile aynıdır. Öte yandan uygun gecikme sayısının bulunması, yapılan analizlerin güvenilirliği açısından önemlidir. Çünkü AD-F testinin I. tip hata olasılığı ve testin gücü modele eklenecek gecikmelerin sayısına duyarlıdır¹³. Bundan dolayı uygun gecikme sayısının belirlenmesi durağanlık analizinin önemli bir aşaması olarak ifade edilebilir. Uygun gecikme sayısının belirlenmesi aşamasında literatürde çeşitli bilgi

¹³ Maddala and Kim, **a.g.e**, s.77.

kriterleri olmasıyla birlikte uygulamada Akaike ve Schwarz kriterleri sıklıkla kullanılmaktadır.

2.2. Kesirli Dikey Fuller Testi

Klasik birim kök testleriyle durağanlık analizleri yapılırken serilerin temel ve alternatif hipotezler altında bütünleşme derecelerinin 0,1,2,... şeklinde tamsayı olduğu varsayılmaktadır. Diğer bir ifadeyle seriler ya durağandır ya da durağan değildir şeklinde sınıflandırılabilmektedir. Ancak zaman serileri sadece ya durağan ya da durağan olmayan serilerden oluşmamaktadır, yani seriler durağandır ya da değildir gibi keskin bir ayırım çok kısıtlayıcı olmaktadır¹⁴. Bu noktadan hareketle klasik birim kök testleriyle analiz yapılırken serilerin bütünleşme derecelerinin (d) tamsayı olması varsayımı gevşetilerek yeni testler geliştirilmesine ihtiyaç duyulmuştur. Öte yandan Diebold ve Rudebusch (1991), Lee ve Schmidt (1996), alternatif hipotez altında kesirli bütünleşik bir süreç için klasik birim kök testleri kullanıldığında testlerinin tutarlı olduğu fakat güçlerinin azaldığını belirtmişlerdir. Bu noktadan hareketle literatürde bu durumu dikkate alan çeşitli testler geliştirilmiştir. Burada iki tür test grubu karşımıza çıkmaktadır. Bunlardan ilki Wald tipi testlerdir. Bu tür testler alternatif hipotez altında çalışmakta ve hafıza parametresi (d) için hem nokta tahmini hem de güven aralığı hesaplanmaktadır. İkinci grup testler ise temel hipotez altında çalışan Lagrange Multiplier (LM) tipi testlerdir. Birinci grupta d 'nin tahmini için parametrik ve yarı parametrik olmak üzere çeşitli testler yer almaktadır. Bunlara örnek olarak Geweke ve Porter-Hudak (1983), Fox ve Taqqu (1986), Sowell (1992), Robinson (1992) verilebilir. İkinci grup testlere ise Robinson (1994) ve Tanaka (1999) tarafından sırasıyla frekans tanım kümesinde (frequency domain) ve zaman tanım kümesinde (time domain) geliştirilen LM testi örnek olarak verilebilir. Bu her iki yaklaşımı klasik birim kök testlerinden ayıran en önemli özellik ise, klasik birim kök testlerinin aksine standart asimptotik dağılımlarının olmasıdır. Öte yandan LM tipi testler bu tip avantajlara sahip olmalarına rağmen boş hipotez altında çalıştığı için, boş hipotezin reddedilmesi durumunda d ile ilgili direk bir bilgi vermemektedir.

¹⁴ Kutlar, A., & Turgut, T., “ Türkiye’deki Başlıca Ekonomi Serilerinin ARFIMA Modelleri İle Tahmini ve Öngörülebilirliği” , **Kocaeli Üniversitesi Sosyal Bilimler Dergisi**, (11), 2006,s. 120-149.

Bu tez çalışmasında kullanılan Dolado, Gonzalo ve Mayoral (DGM, 2002) tarafından geliştirilen kesirli Dickey Fuller (FD-F) testi ise alternatif hipotez altında çalışan Wald tipi test grubundadır. Kolaylık olması açısından tezin geri kalan kısmında testin kısaltılmış biçimi FD-F kullanılacaktır. FD-F’de, D-F testinde serilerin bütünleşme derecelerinin tam sayı değerler alması durumu genişletilerek kesirli değerler alması sağlanmaktadır.

Dolado, Gonzalo ve Mayoral (2002)’de FD-F test süreci için kullanılan model aşağıdaki gibidir:

$$\Delta^{d_0} X_t = \psi \Delta^{d_1} X_{t-1} + \varepsilon_t \quad (2.18)$$

Burada $\varepsilon_t \sim i.i.d(0, \sigma_\varepsilon^2)$ şeklindedir, d_0 ve d_1 ise kesirli değerler alan fark parametreleridir. Denklem (2.18), bağımlı ve bağımsız değişkenler hem boş hem de alternatif hipotez altında farklı bütünleşme derecelerine sahip olduğundan dengeli olmayan bir regresyon yapısına sahiptir¹⁵. (2.18) nolu eşitlikte yer alan $\Delta^{d_1} X_{t-1}$ ifadesi aşağıdaki gibi hesaplanmaktadır:

$$\Delta^{d_1} X_t = \sum_{i=0}^{t-m-1} \pi_i(d_1) X_{t-i}$$

Burada m ($d_1 + 0.5$)’ın tam değer kısmını göstermektedir. $d_1 = 1$ olması durumunda bir gözlem kaybolacaktır.

D-F testinde olduğu gibi FD-F testinde de $\Delta^{d_1} X_{t-1}$ ’nin katsayısı olan ψ parametresinin anlamlılığı sınanmaktadır. Temel hipotez altında $\psi = 0$ olduğu durumunda süreç aşağıdaki gibi ifade edilmektedir:

$$\Delta^{d_0} X_t = \varepsilon_t \quad (2.19)$$

¹⁵ Dolado, J.J., Gonzalo, J., Mayoral, L., “A fractional Dickey–Fuller test for unit roots”, **Econometrica**, 2002, 70(5), s.1966

Burada X_t süreci kesirli bütünleşik sürece sahiptir ve $FI(d)$ şeklinde ifade edilmektedir. $\psi < 0$ durumunda X_t

$$(\Delta^{d_0-d_1} - \psi L)\Delta^{d_1}X_t = \varepsilon_t \quad (2.20)$$

şeklinde yazılabilir. Öte yandan FD-F testinde kolaylık olması ve literatürde yaygın olarak kullanılması açısından $d_0 = 1$ olarak alınmakta ve böylece X_t rastgele yürüyüş sürecine sahip olmaktadır. Buna karşın alternatif hipotez altında, $0 \leq d_1 < 1$ olmak üzere $X_t \sim FI(d_1)$ şeklinde ifade edilmektedir.

X_t 'nin birim köklü sürece sahip olduğunu belirten boş hipotez altında $\hat{\psi}$, ψ 'nin tutarlı bir tahmin edicisidir¹⁶. Aşağıdaki eşitliklerde d_1 parametresinin bulunduğu aralığa karşılık gelen bu tahmincinin ($\hat{\psi}$) gerçek değerine (ψ) yakınsama hızları verilmektedir:

$$0 < d_1 < 0.5 \quad T^{1-d_1} \quad (2.21)$$

$$d_1 = 0.5 \quad (T \log T)^{1/2} \quad (2.22)$$

$$0.5 < d_1 < 1 \quad T^{1/2} \quad (2.23)$$

Aynı zamanda $\hat{\psi}$ 'nin asimptotik dağılımı ise aşağıdaki gibidir:

$$0 \leq d_1 < 0.5 \quad T^{1-d_1}\hat{\psi} \xrightarrow{w} \frac{\int_0^1 W_{-d_1}(r)dB(r)}{\int_0^1 W_{-d_1}^2(r)dr} \quad (2.24)$$

$$d_1 = 0.5 \quad (T \log T)^{1/2}\hat{\psi} \xrightarrow{w} N(0, \pi) \quad (2.25)$$

$$0.5 < d_1 < 1 \quad T^{1/2}\hat{\psi} \xrightarrow{w} N\left(0, \frac{\Gamma^2(d_1)}{\Gamma(2d_1 - 1)}\right) \quad (2.26)$$

¹⁶ a.e, s.1969.

Burada, (Γ) gama fonksiyonunu göstermektedir. X_t 'nin rastgele yürüyüş sürecine sahip olduğu boş hipotez altında, $\hat{\psi}$ 'nin t test istatistiğine $(t_{\hat{\psi}})$ ait asimptotik dağılım aşağıdaki gibidir:

$$0 \leq d_1 < 0.5 \quad t_{\hat{\psi}} \xrightarrow{w} \frac{\int_0^1 W_{-d_1}(r) dB(r)}{\left(\int_0^1 W_{-d_1}^2(r) dr\right)^{1/2}} \quad (2.27)$$

$$0.5 \leq d_1 < 1 \quad t_{\hat{\psi}} \xrightarrow{w} N(0,1) \quad (2.28)$$

Burada, (B) standart Brownian hareketini, (W_d) standart kesirli Brownian hareketini, $\left(\xrightarrow{w}\right)$ olasılıktaki zayıf yakınsamayı ifade etmektedir. Hem $\hat{\psi}$ hem de $t_{\hat{\psi}}$ istatistiğinin asimptotik dağılımı, d_1 parametresinin değeri bilindiği varsayımı altında elde edilmiştir. İstatistiklerin asimptotik dağılımlarının normal veya normal olmayan süreçlere sahip olması, temel hipotez ve alternatif hipotez arasındaki uzaklığa bağlıdır¹⁷. Alternatif hipotez altında seri durağan olmayan bir sürece sahipse $(0.5 \leq d_1 < 1)$ $t_{\hat{\psi}}$ test istatistiği asimptotik olarak standart normal dağılıma yakınsamaktadır. Fakat alternatif hipotez altında seri durağan ise $(0 \leq d_1 < 0.5)$ $t_{\hat{\psi}}$ istatistiği asimptotik olarak standart olmayan bir dağılıma sahip olmaktadır. Standart olmayan dağılım (3.27)'de görüldüğü gibi kesikli Brownian hareketinin bir fonksiyonu olmaktadır.

Dolado, Gonzalo ve Mayoral (2002) yaptıkları çalışmada, öne sürdükleri bu testin (FD-F) I. tip hata olasılığını ve testin gücünü incelemekle birlikte literatürdeki alternatif testlerle karşılaştırma yapmışlardır. Bu alternatif testler, Geweke ve Porter-Hudak (1983), Tanaka (1999) ve Robinson (1994) tarafından önerilen LM tipi testlerdir. DGM (2002), 100 ve 400 gözlemlili 1000 Monte Carlo simülasyon deneyleri sonucunda testin I. tip hata olasılığı ve testin gücü açısından çok iyi sonuçlar verdiği belirtilmektedir. Öte yandan bu deneylerde d_1 parametresinin önceden bilindiği durum için sonuçlar elde edilmektedir.

¹⁷ a.e., s.1970.

Bununla birlikte, uygulamada genellikle kesirli fark parametresi d bilinmemektedir ve tahmin edilmesi gerekmektedir. Literatüre baktığımızda hem zaman tanım kümesinde (time domain) hem de frekans tanım kümesinde (frequency domain) kesirli fark parametresi (d) için çeşitli tahmin yöntemleri yer almaktadır. Bu tahmin yöntemlerine Velasco ve Robinson (2000) tarafından geliştirilen Whittle tahmincisi, Beran ve Tanaka (1999) tarafından geliştirilen en çok olabilirlik tahmincisi, Galbraith ve Zinde-Walsh (1997) tarafından öne sürülen en küçük uzaklık (minimum distance) tahmincisi verilebilir. Bu tahmincilere ek olarak Mayoral (2000) tarafından geliştirilen ve en küçük uzaklık tahmincisinin genişletilmiş hali olan GMD (generalized minimum distance) tahmincisi örnek verilebilir. Temelde d parametresinin tahmini için birçok yöntem karşımıza çıkmakla birlikte araştırmacılar, Dolado, Gonzalo ve Mayoral (2002) çalışmasında, Mayoral (2000) tarafından geliştirilen GMD tahmincisini kullanmışlardır. Öte yandan araştırmacılar çalışmada GMD tahmincisinin geçerliğini araştırmak için yaptıkları simülasyonlar sonucunda tahmincinin hem I. tip hata olasılığı hem de testin gücü açısından iyi sonuçlar verdiğini göstermişlerdir. (d) parametresi tahmin edildikten sonra testin gücü ile (d) parametresinin bilindiği durumda testin gücü karşılaştırıldığında aralarında çok küçük farklar olduğu yapılan simülasyonlarla Dolado, Gonzalo ve Mayoral (2002) tarafından ortaya konulmuştur.

Böylece (d) parametresinin tahmin edildiği durumda X_t 'nin rastgele yürüyüş sürecine sahip olduğu temel hipotez altında, aşağıdaki denklem yardımıyla $\hat{\psi}$ parametresine ait t istatistiği elde edilebilir:

$$\Delta X_t = \psi \Delta^{\hat{d}_1} X_{t-1} + \varepsilon_t \quad (2.29)$$

Burada ε_t beyaz gürültü sürecine sahip hata terimidir. Yukarıdaki test denkleminde (d) parametresi tahmin edildiği durumda $t_{\hat{\psi}}$ test istatistiği, d 'nin değerine bakılmaksızın standart normal dağılıma sahiptir¹⁸. Fakat test istatistiğinin asimptotik

¹⁸ A.e. , s.1977.

olarak standart normal dağılıma sahip olabilmesi için $(\hat{d})$ 'in 1'den küçük olması gerekir. Testin asimptotik dağılımı gereği $(\hat{d})$ aşağıdaki koşul ile belirlenmektedir:

$$\hat{d}_1 = \begin{cases} \hat{d}_T, & \hat{d}_T < 1 - c \\ 1 - c, & \hat{d}_T \geq 1 - c \end{cases} \quad (2.30)$$

Burada c sıfıra çok yakın pozitif bir sayıdır. Denklem (2.30)'da c 'nin 0.02 gibi bir değer aldığını farz edelim. Eğer $\hat{d}_1$, $1 - c$ farkından yani 0.98'den küçükse, tahmin edilen $\hat{d}_1$ 'in kendi değeri kullanılır. Fakat $\hat{d}_1$, $1 - c$ farkından büyükse $\hat{d}_1$ yerine $1 - c$ yani 0.98 kullanılır. Dolado, Gonzalo ve Mayoral (2002), yaptıkları çalışmada test istatistiğinin $(t_{\hat{\varphi}})$ standart normal dağılıma sahip olması için denklem (2.30)'da verilen kuralın gerekli olduğunu belirtmektedirler. Öte yandan Dolado, Gonzalo ve Mayoral (2002), kesirli fark parametresinin gerçek değeri (d_1) yerine tahmin edildiği durumda FD-F testinin performansının önemsenmeyecek derecede az bir şekilde etkilendiğini yaptıkları simülasyon çalışmalarıyla göstermişlerdir. Özetle ifade etmek gerekirse ilk adımda d parametresi tahmin edilmekte daha sonra $t_{\hat{\varphi}}$ test istatistiği farklı anlamlılık düzeyleri için normal dağılım tablosundan elde edilen kritik değerler ile karşılaştırılmaktadır. Test istatistiği verilen anlamlılık düzeyinde kritik değerden mutlak değerce büyükse serinin birim köklü sürece sahip olduğunu belirten boş hipotez reddedilerek alternatif hipotez altında serinin kesirli bütünleşik sürece sahip olduğu sonucuna varılmaktadır.

Daha önce de bahsedildiği gibi FD-F testinde modelin hata teriminin otokorelasyonsuz bir yapıya sahip olduğu varsayımı yapılmaktadır. Fakat hata teriminde otokorelasyonlu bir yapı gözlemlendiği durumunda FD-F testi geçersiz olmaktadır. Dolayısıyla otokorelasyon problemi bağımlı değişkenin gecikmeli değerlerinin de modele katılması ile düzeltilmektedir. Bu durumda genişletilmiş FD-F (AFD-F) testi Dolado, Gonzalo ve Mayoral (2002)'de aşağıdaki eşitlik ile ifade edilebilmektedir:

$$\Delta X_t = \psi \Delta^{\hat{d}_1} X_{t-1} + \sum_{i=1}^p \xi_i \Delta X_{t-i} + \varepsilon_t \quad (2.31)$$

FD-F testinde olduğu gibi AFD-F testinde de durağanlık analizi için test edilecek hipotezler aynı şekildedir. Daha açık bir şekilde ifade etmek gerekirse denklem (2.31)' den hareketle temel hipotez altında $\psi = 0$ durumu alternatif hipotez altında $\psi < 0$ durumuna karşı sınanmaktadır. Diğer bir ifadeyle X_t serisi temel hipotez altında birim köklü iken alternatif hipotez altında kesirli bütünleşik bir sürece sahiptir. Dolado, Gonzalo ve Mayoral (2002), yapılan analizler sonucunda AFD-F sürecinden elde edilen $t_{\hat{\psi}}$ test istatistiğinin asimptotik dağılımı ile FD-F testinden elde edilen test istatistiğinin asimptotik dağılımının aynı olduğunu ifade etmektedirler. Dolayısıyla daha önce de ifade edildiği gibi d_1 parametresinin değeri önceden bilindiğinde $0.5 \leq d_1 < 1$ durumu için $t_{\hat{\psi}}$ asimptotik normal dağılıma sahiptir. Öte yandan $0 \leq d_1 < 0.5$ durumunda $t_{\hat{\psi}}$ asimptotik olarak standart olmayan bir dağılıma sahip olmaktadır. d' nin tahmin edildiği durumda ise $t_{\hat{\psi}}$ asimptotik olarak standart normal dağılıma sahiptir.

2.3. Deterministik Bileşenler Durumunda DGM (2002) Birim Kök Testi: Monte Carlo Simülasyon Sonuçları

Dolado, Gonzalo ve Mayoral (2002) çalışmasında (2.29) ve (2.31) nolu eşitliklere herhangi bir deterministik bileşen (sabit terim ve trend) dahil etmeden FD-F ve AFD-F birim kök testlerinin sınırlı veri kümesindeki performansını incelemişlerdir. Bu tez çalışmasında Dolado, Gonzalo ve Mayoral (2002)'den farklı olarak deterministik bileşenlerin regresyon eşitliğine dahil edildiği durumlar da ele alınarak her iki testin sınırlı veri kümesinde I. tip hata olasılığı ve gücü hesaplanacaktır.

Bu kısımda ilk olarak deterministik bileşenler durumunda FD-F birim kök testinde (d) parametresinin tahmin edildiğinde test istatistiğinin ($t_{\hat{\psi}}$) kritik değerleri

elde edilecektir. X_t serisi temel hipotez altında pür rastgele yürüyüş sürecine sahip olmak üzere test istatistiğinin $(t_{\hat{\varphi}})$ %1, %5 ve %10 anlamlılık düzeylerinde farklı örnek büyüklüklerine göre kritik değerleri Tablo 2.1’de verilmiştir. Kritik değerlerin elde edilmesinde 50000 Monte Carlo simülasyonu yapılmıştır.

Tablo 2.1. Deterministik Bileşenler Durumunda FD-F Testinin Kritik Değerleri

T	Sabit Terim			Sabit Terim ve Trend		
	1%	5%	10%	1%	5%	10%
100	-2.49	-1.77	-1.38	-2.65	-1.89	-1.51
200	-2.42	-1.71	-1.34	-2.51	-1.79	-1.42
500	-2.38	-1.69	-1.32	-2.43	-1.74	-1.36
1000	-2.36	-1.66	-1.30	-2.40	-1.69	-1.33
2000	-2.35	-1.66	-1.29	-2.36	-1.68	-1.31
5000	-2.32	-1.64	-1.28	-2.33	-1.65	-1.30
10000	-2.32	-1.64	-1.29	-2.34	-1.65	-1.30

Daha önce de ifade edildiği gibi deterministik bileşenlerin modele dahil edilmediği ve (d) parametresinin tahmin edildiği durumda test istatistiği $(t_{\hat{\varphi}})$ asimptotik olarak standart normal dağılıma sahip olmaktadır. Benzer bir durum deterministik bileşenlerin modele dahil edildiği her iki durumda da Tablo 2.1’den görülmektedir. Tablo 2.1’e göre örnek büyüklüğü T arttıkça $(t_{\hat{\varphi}})$ ’nın dağılımının standart normale yaklaştığı, hatta $T = 10000$ durumunda standart normal dağılım kritik değerlerine yakınsadığı görülmektedir.

Kritik değerler elde edildikten sonra FD-F testinin sınırlı veri kümesinde I. tip hata olasılığı ve gücünü hesaplamak amacıyla aşağıda verilen veri üretim süreci kullanılmıştır:

$$\Delta^{d_1} X_t = \varepsilon_t \quad (2.32)$$

$$\varepsilon_t \sim i.i.d N(0,1) \quad (2.33)$$

Burada I.tip hata olasılığı (α) , gerçekte H_0 hipotezi doğruyken reddedilme olasılığıdır. II. tip hata olasılığı ise gerçekte H_0 yanlışken reddedilememe olasılığıdır

ve β ile ifade edilmektedir. Bu noktada I.tip hatanın, II. tip hataya oranla çok daha tehlikeli olduğu ifade edilmektedir¹⁹. Öte yandan testin I.tip hata olasılığı ile II. tip hata olasılığı arasında ters yönlü bir ilişki söz konusudur. Diğer bir ifadeyle I.tip hata olasılığı (α)’nın azaltılması durumunda testin II. tip hata olasılığı (β) artacaktır. Testin gücü ise $(1 - \beta)$ şeklinde hesaplanmakta ve II. tip hata yapmama olasılığı olarak ifade edilmektedir. Diğer bir ifadeyle testin gücü, gerçekte alternatif hipotez doğru iken temel hipotezin (H_0) reddedilme olasılığıdır²⁰.

(2.32)’deki veri üretim sürecinde d_1 parametresinin 0’dan başlayarak ve 0.1 arttırılarak 1’e kadarki değerleri ele alınmıştır. $d_1 = 1$ parametre değeri testin I. tip hata olasılığını gösteriyorken diğer değerler testin gücünü ifade etmektedir. Aynı zamanda test aşamasında d_1 parametresini tahmin etmek için, Dolado, Gonzalo, Mayoral (2002)’ çalışmasında kullandığı genelleştirilmiş en küçük uzaklık (GMD) tahmincisi yerine Shimotsu ve Phillips (2005) tarafından geliştirilmiş olan “Exact Local Whittle” tahmincisi kullanılmıştır. Hem deterministik bileşen içermeyen hem de içeren durumlar için $T = 100, 200, 400$ ve 1000 örnek büyüklüklerinde 5000 Monte Carlo deneyinden FD-F birim kök testinin I. tip hata olasılığı ve gücü hesaplanmıştır. Nominal anlamlılık düzeyi %5 olarak seçilmiştir. İlk olarak deterministik bileşensiz model için elde edilen I. tip hata olasılığı ve testin gücü olasılığı Monte Carlo sonuçları Tablo 2.2’de gösterilmiştir:

Tablo 2.2. Deterministik Bileşensiz Durumda FD-F Testinin I. Tip Hata Olasılığı ve Testin Gücü

T/d_1	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
100	100	100	100	100	99.98	99.84	98.84	89.68	63.12	26.16	5.82
200	100	100	100	100	100	100	100	99.64	88.52	41.58	5.52
400	100	100	100	100	100	100	100	100	99.42	64.76	5.82
1000	100	100	100	100	100	100	100	100	100	94.54	4.82

¹⁹ Baron, M. , **Probability and statistics for computer scientists** , 2.bs, USA, Chapman and Hall/CRC , 2013 , s.267.

²⁰ Montgomery, Douglas C., and George C. Runger. , **Applied Statistics And Probability For Engineers** , 3.bs , USA, John Wiley & Sons, 2007, s.285.

Tablo 2.2'deki sonuçlara göre deterministik bileşenlerin dahil edilmediği model için küçük örneklerde bile testin I. tip olasılığı nominal anlamlılık düzeyi %5 civarındadır. Örneğin $T = 100$ için testin I. tip hata olasılığı %5.82 ile %5 anlamlılık düzeyine oldukça yakındır. Örnek büyüklüğü arttıkça da FD-F birim kök testinin I. tip hata olasılığının %5 anlamlılık düzeyine yakın olduğu görülmektedir. Dolayısıyla kesirli fark parametresinin tahmin edildiği ve modele herhangi bir deterministik bileşen ilave edilmediği durumda testin I. tip hata olasılığında herhangi bir bozulma görülmemektedir. FD-F testinin gücü incelendiğinde ise farklı d parametrelerinde örnek çapı T arttıkça %100'e doğru gitmektedir. Örneğin $d_1 = 0.9$ parametresi için $T = 100$ durumunda testin gücü %26.16 iken bu değer $T = 1000$ durumunda %94.54 olarak elde edilmiştir. Dolayısıyla deterministik bileşensiz FD-F testi için elde edilen Monte Carlo sonuçları Dolado, Gonzalo ve Mayoral (2002) çalışmasındaki sonuçlarla hemen hemen aynıdır.

İkinci olarak FD-F modeline deterministik bileşenler eklenerek testlerin I. tip hata olasılıkları ve güçleri analiz edilmiştir. Tablo 2.3'de sabit terimli ve sabit terim ve trendli modellerde testin I. tip hata olasılıkları ve güçleri sunulmuştur:

Tablo 2.3. Deterministik Bileşenli Durumda FD-F Testinin I. Tip Hata Olasılığı ve Testin Gücü

Sabit Terimli											
T/d_1	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
100	100	100	100	100	99.98	99.86	98.84	90.16	64.18	27.10	6.54
200	100	100	100	100	100	100	99.98	99.62	89.20	41.96	5.76
400	100	100	100	100	100	100	100	100	99.28	64.42	5.50
1000	100	100	100	100	100	100	100	100	100	94.52	4.84
Sabit Terimli ve Trendli											
T/d_1	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
100	100	100	100	100	100	99.86	99.04	90.58	65.54	29.42	7.96
200	100	100	100	100	100	100	100	99.72	89.52	43.90	6.54
400	100	100	100	100	100	100	100	100	99.46	65.96	6.52
1000	100	100	100	100	100	100	100	100	100	94.76	5.24

Tablo 2.3 incelendiğinde sabit terimli model için küçük örneklerde testin I. tip hata olasılığında ciddi olmamakla beraber bozulmalar meydana gelmektedir. Örneğin

$T = 100$ için testin I. tip hata olasılığı %6.54'tür ve %5 anlamlılık düzeyinden deterministik bileşenlerin olmadığı modele kıyasla kısmen uzaktır. Sabit terimli ve trendli model için ise bozulmalar daha fazladır. Örneğin $T = 100$ için testin I. tip hata olasılığı %7.96'dır ve %5 anlamlılık düzeyinden bir hayli uzaktır. Sonuç olarak modele sabit terim ve trend eklendiği durumda testin küçük örneklerde I. tip hata olasılığında bozulmalar görülmektedir. Bununla birlikte, deterministik bileşenli modellerde testin I. tip hata olasılığındaki bu bozulmalar örnek büyüklüğü arttıkça anlamlılık düzeyi %5'e yaklaşmaktadır. Bu durum $T = 1000$ örnek büyüklüğünde sabit terimli modelde %4.84 ve sabit terim ve trendli modelde %5.24 ile net olarak görülmektedir.

Tablo 2.3'e göre modelde deterministik bileşenler yer aldığı durumda testin gücü sınırlı veri kümesinde d_1 parametresinin değeri sıfıra doğru yaklaştıkça %100'e doğru artmaktadır. Öte yandan örnek büyüklüğü arttıkça $d_1 \rightarrow 1$ yaklaşırsa bile testin gücü %100'e doğru gitmektedir.

Şimdi AFD-F birim kök testine deterministik bileşenler ilave edildiğinde testin I. tip hata olasılığı ve gücü incelenecektir. Bu amaçla kullanılacak veri üretim süreci aşağıdaki gibidir:

$$\Delta^{d_1} X_t = \varepsilon_t \quad (2.34)$$

$$\varepsilon_t = \rho \varepsilon_{t-1} + u_t \quad (2.35)$$

$$u_t \sim i.i.d N(0,1) \quad (2.36)$$

(2.34)'deki veri üretim sürecinde FD-F simülasyonlarında olduğu gibi $d_1 = 1$ parametre değeri testin I. tip hata olasılığını, (0.0 – 0.9) arasındaki değerler ise testin gücünü ifade etmektedir. (2.35)'de verilen veri üretim sürecinde hata terimi ε_t 'nin durağan AR(1) yapısında olduğu varsayılmış ve otoregressif katsayının $\rho = 0.2$ ve $\rho = 0.6$ değerleri için ayrı ayrı analizler yapılmıştır. (2.31)'de verilen AFD-F test denkleminde gecikme uzunluğu p , veri üretim sürecinde hata terimi birinci mertebeden otokorelasyon yapısında olduğundan dolayı “1” olarak

seçilmiştir. Buna ek olarak otokorelasyon mertebesinin önceden bilindiğini varsaymak yerine test aşamasında belirlemek için alternatif olarak Akaike ve Schwarz bilgi kriterlerinden yararlanılmıştır. Aynı zamanda (2.31)’de $\hat{d}_1$ FD-F simülasyon çalışmasında olduğu gibi “Exact Local Whittle” tahmincisi ile bulunmuştur. AFD-F birim kök testinin I. Tip hata olasılığı ve gücünü hesaplamak için standart normal dağılım kritik değerleri kullanılmıştır. Tüm simülasyon deneylerinde Monte Carlo deney sayısı 5000 olarak seçilmiş olup anlamlılık düzeyi %5 olarak atanmıştır.

Dolado, Gonzalo ve Mayoral (2002) çalışmalarında FD-F testinde olduğu gibi burada da sadece sabit terim ve trend içermeyen model için Monte Carlo simülasyonları yapılarak testin I. tip hata olasılığı ve gücünü analiz etmişlerdir. Tablo 2.4’de $\rho = 0.2$ için deterministik bileşensiz AFD-F modelinin farklı örnek büyüklüklerinde hesaplanan testin I. tip hata olasılığı ve gücü olasılıkları yer almaktadır:

Tablo 2.4. Deterministik Bileşensiz Durumda AFD-F Testinin I. Tip Hata Olasılığı ve Testin Gücü

$\rho = 0.2$												
T/d_1	p	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
100	1	100	100	100	99.72	98.44	92.88	80.20	57.70	32.96	14.92	6.34
	bic	100	100	100	99.84	98.78	94.28	82.16	60.58	35.18	17.00	7.74
	aic	99.72	99.26	98.90	97.80	96.04	92.84	82.60	64.32	39.68	22.34	12.32
200	1	100	100	100	100	100	99.78	97.84	85.58	53.60	22.08	6.14
	bic	100	100	100	100	99.98	99.92	97.86	87.36	56.88	22.98	7.16
	aic	100	99.98	99.90	99.76	99.08	98.48	95.42	87.28	60.80	28.44	11.30
400	1	100	100	100	100	100	100	100	98.98	82.20	33.80	5.50
	bic	100	100	100	100	100	100	99.98	99.38	82.50	34.46	6.68
	aic	100	100	100	100	99.94	99.74	99.24	96.50	83.84	42.02	10.24
1000	1	100	100	100	100	100	100	100	100	99.54	65.34	5.52
	bic	100	100	100	100	100	100	100	100	99.62	66.08	5.86
	aic	100	100	100	100	100	100	100	99.46	97.12	67.94	10.36
10000	1	100	100	100	100	100	100	100	100	100	100	5.24
	bic	100	100	100	100	100	100	100	100	100	100	5.32
	aic	100	100	100	100	100	100	100	100	100	99.72	9.42

Tablo 2.4’e göre sabit terim ve trendsiz AFD-F modeli için küçük örneklerde testin I. tip hata olasılığının %5 anlamlılık düzeyinden sapmalar görülmektedir.

Örneğin $T = 100$ için testin I. tip hata olasılığı %6.34 ve %5 anlamlılık düzeyinin yukarısındadır. Bununla beraber hata teriminin yapısı bilgi kriterleri aracılığıyla belirlendiği durumda testin I. tip hata olasılığındaki sapmanın boyutu artmaktadır. Örneğin Schwarz bilgi kriteri kullanıldığında $T = 100$ için testin I. tip hata olasılığı %7.74 iken Akaike bilgi kriteri kullanıldığında testin I. tip hata olasılığı %12.32 olmaktadır. Diğer yandan örnek büyüklüğü arttıkça testin I. tip hata olasılığı p 'nin 1 olduğu ve Schwarz bilgi kriteri ile seçildiği durumlarda anlamlılık düzeyine oldukça yakındır. Dolayısıyla testin küçük örneklerde gösterdiği I. tip hata bozulması ortadan kalkmaktadır. Örneğin $T = 10000$ için testin I. tip hata olasılığı sırasıyla $p = 1$ için %5.24 ve BIC için %5.32 dir. Ancak bu durum Akaike bilgi kriteri için geçerli değildir. Burada Akaike bilgi kriteri temel alınarak hata teriminin yapısı modellendiğinde testin I. tip hata yapma olasılığının Schwarz bilgi kriterine göre çok daha fazla olmasının sebebi, Akaike'in otokorelasyon mertebesini olduğundan daha yüksek seçmesinden kaynaklanmaktadır. Tablo (2.4)' e göre AFD-F testinin gücü incelendiğinde, FD-F testinde modelde deterministik bileşen olmadığı durumda benzer şekilde gözlem sayısı arttıkça testin gücü %100' e doğru artmaktadır.

Deterministik bileşenlerin dahil edilmediği AFD-F birim kök testinde $\rho = 0.6$ durumunda Monte Carlo sonuçları Tablo 2.5'te sunulmuştur. p 'nin hem önceden bilindiği hem de bilgi kriterlerine göre seçildiği durumlarda AFD-F testinin I. tip hata olasılığı Tablo 2.4'de olduğu gibi benzer bir kalıp sergilemektedir. Örnek büyüklüğü arttıkça AIC durumunda I. tip hata olasılığındaki bozulmalar azalmakla birlikte hala sapmanın yüksek olduğu görülmektedir. Örneğin $T = 100$ olduğunda AIC'de I. tip hata olasılığı %12.22 iken bu olasılık $T = 10000$ durumunda %8.98 olarak bulunmuştur. Ancak p 'nin 1 alındığı durumla BIC'le seçildiği durumlarda $\rho = 0.2$ sonuçlarında olduğu gibi I. tip hata olasılığındaki bozulmalar örnek büyüklüğü arttıkça azalmaktadır. Öte yandan testin gücü gözlem sayısı arttıkça %100'e doğru gitmektedir. Örneğin $d_1 = 0.9$ ve $T = 100$ p 'nin 1 alındığı, AC ve BIC'te testin gücü sırasıyla %8.02 , %10.22 , %16.12 olmakla beraber $T = 10000$ olduğunda sırasıyla %79.60, %88.18 , %97.88'dir.

Tablo 2.5. Deterministik Bileşensiz Durumda AFD-F Testinin I. Tip Hata Olasılığı ve Testin Gücü

$\rho = 0.6$												
T/d_1	p	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
100	1	98.34	94.86	85.50	72.14	54.56	39.50	27.10	18.02	11.72	8.02	6.80
	bic	99.36	97.10	90.82	79.42	63.00	48.40	33.94	23.38	15.36	10.22	7.82
	aic	99.24	97.68	95.00	88.10	77.22	63.56	48.24	35.02	23.22	16.12	12.22
200	1	100	99.92	99.16	95.44	84.90	66.10	43.40	28.86	16.00	8.90	6.10
	bic	100	100	99.64	97.86	90.42	73.94	52.56	35.50	20.12	10.80	6.82
	aic	100	99.92	99.40	99.18	96.00	87.96	71.38	53.90	32.40	18.38	10.26
400	1	100	100	100	100	99.12	91.34	71.68	46.78	24.64	11.60	5.80
	bic	100	100	100	100	99.74	95.94	79.92	54.18	29.58	14.00	6.62
	aic	100	100	99.98	99.98	99.84	98.8	93.98	76.86	49.66	23.12	10.12
1000	1	100	100	100	100	100	100	98.38	82.06	47.84	17.50	5.12
	bic	100	100	100	100	100	100	99.46	88.80	55.46	20.86	5.28
	aic	100	100	100	100	100	100	99.84	97.50	80.10	38.00	9.42
10000	1	100	100	100	100	100	100	100	100	100	79.60	5.54
	bic	100	100	100	100	100	100	100	100	100	88.18	5.48
	aic	100	100	100	100	100	100	100	100	100	97.88	8.98

Tablo 2.6 ve Tablo 2.7’de deterministik bileşenlerin AFD-F modeline dahil edildiği durumlarda birim kök testinin I. tip hata olasılığı ve gücü değerleri $\rho = 0.2$ ve $\rho = 0.6$ için sırasıyla verilmiştir. AFD-F modeline deterministik bileşenlerin dahil edildiği her iki durumda da tablolar incelendiğinde I. tip hata olasılıklarında oldukça yüksek oranda bozulmalar ortaya çıkmaktadır. Örneğin $\rho = 0.2$ ve sabit terimli modelde $T = 100$ için $p = 1$ iken I. tip hata olasılığı %10.22’dir. Yine hata teriminin yapısı Schwarz ve Akaike bilgi kriterleri aracılığıyla modellendiğinde $T = 100$ durumunda olasılıklar sırasıyla %12.10 ve %20.16 bulunmuştur. Örnek büyüklüğü arttıkça p ’nin 1 alındığı ve BIC ile seçildiği durumlarda I. Tip hata olasılıkları %5 anlamlılık düzeyine yaklaşarak bozulmaların düzeldiği görülmektedir. Ancak AIC için aynı sonuçlar söz konusu değildir. $T = 10000$ durumunda bile testin I. tip hata olasılığı AIC’e göre %5 anlamlılık düzeyinin oldukça üzerinde %10.48 bulunmuştur. Tablo 2.7’de sabit terimli modelde $\rho = 0.6$ için benzer sonuçlar elde edilmiştir.

AFD-F modeline hem sabit terim hem de trend terimin dahil edildiği durumda ise bozulmanın şiddeti küçük örneklerde çok daha büyüktür. Örneğin $T = 100$ ve $\rho = 0.2$ durumunda testin I. tip hata olasılığı $p = 1$ için %15.68'dir. Benzer şekilde BIC ve AIC durumları için sırasıyla I. tip hata olasılıkları %19.50 ve %34.66 şeklindedir. Sadece sabit terimli AFD-F testinde olduğu gibi $T = 10000$ iken p 'nin 1 alındığı ve BIC ile seçildiği durumlarda I. tip hata olasılıkları sırasıyla %6.86 ve %6.76 ile anlamlılık düzeyine oldukça yakındır. Örnek büyüklüğü sonsuza yaklaştıkça testin I. Tip hata olasılığındaki bozulmalar AIC ile seçilen durum hariç diğer iki durumda düzelecektir. Ancak AIC durumunda bu bozulmalar örnek çapı büyüdüğünde bile hala devam edecektir. Tablo 2.7'de sabit terim, sabit terim ve trendli AFD-F modelinde $\rho = 0.6$ için benzer sonuçlar bulunmuştur.

FD-F ve AFD-F testlerinin güçleriyle ilgili genel bir yorum yapılacak olursa, FD-F testi ile karşılaştırıldığında AFD-F testinin gücü d_1 parametresinin değeri 1'e doğru yaklaştıkça FD-F testine oranla daha çok azaldığı görülmektedir. Dolayısıyla modelde otokorelasyon probleminin olması halinde, testin hem I. tip hata olasılığı ve gücündeki bozulmalar otokorelasyon probleminin olmadığı duruma göre çok daha ciddidir. Öte yandan bu bozulmalar modeldeki otokorelasyon probleminin şiddetine paralel olarak artmaktadır.

Sonuç olarak AFD-F birim kök testine deterministik bileşenler eklendiğinde testin I. tip hata olasılığı nominal anlamlılık düzeyine göre özellikle T küçük iken ciddi oranda artmaktadır. Öte yandan hata terimlerinin *i.i.d* özelliği bozulduğu zaman otokorelasyonlu bir yapıya sahip olduğu durumda testin I. tip hata olasılığı ciddi derecede artmaktadır. Bu durum testin güvenilirliği açısından büyük sorunlar yaratmaktadır.

Tablo 2.6. Deterministik Bileşenli Durumda AFD-F Testinin I. Tip Hata Olasılığı ve Testin Gücü

Sabit Terim												
$\rho = 0.2$												
T/d_1	p	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
100	1	100	100	99.98	99.78	98.52	93.28	81.82	61.12	38.46	20.24	10.22
	bic	100	100	100	99.86	98.82	94.74	84.04	64.54	41.80	23.14	12.10
	aic	99.90	99.68	99.32	98.60	97.28	94.38	85.90	70.62	49.34	32.22	20.16
200	1	100	100	100	100	100	99.78	97.92	86.66	57.32	26.70	9.04
	bic	100	100	100	100	99.96	99.86	98.14	88.10	60.50	28.22	10.08
	aic	100	100	99.92	99.86	99.38	98.92	96.58	90.16	66.76	36.56	17.76
400	1	100	100	100	100	100	100	100	99.02	83.44	37.26	7.88
	bic	100	100	100	100	100	100	99.98	99.00	84.56	39.84	8.58
	aic	100	100	100	99.96	100	99.94	99.18	97.62	85.96	48.54	15.14
1000	1	100	100	100	100	100	100	100	100	99.52	66.34	7.26
	bic	100	100	100	100	100	100	100	100	99.44	67.26	8.22
	aic	100	100	100	100	100	100	99.94	99.54	97.84	71.16	13.52
10000	1	100	100	100	100	100	100	100	100	100	100	6.12
	bic	100	100	100	100	100	100	100	100	100	100	6.38
	aic	100	100	100	100	100	100	100	100	100	99.32	10.48
Sabit Terim ve Trend												
T/d_1	p	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
100	1	100	100	99.98	99.80	98.62	93.88	83.66	65.08	43.90	26.92	15.68
	bic	100	100	100	99.90	99.00	95.40	86.04	69.10	48.36	31.36	19.50
	aic	99.94	99.90	99.62	99.18	98.22	96.38	89.78	77.14	60.62	45.98	34.66
200	1	100	100	100	100	100	99.78	98.20	88.30	61.66	33.04	12.96
	bic	100	100	100	100	100	99.94	98.26	89.04	65.38	34.86	14.52
	aic	100	100	99.96	99.90	99.68	99.20	97.64	92.26	74.04	47.54	28.20
400	1	100	100	100	100	100	100	100	99.06	85.26	42.66	10.62
	bic	100	100	100	100	100	100	100	99.20	86.42	44.48	12.24
	aic	100	100	100	99.98	100	99.94	99.4	98.44	89.14	56.98	23.72
1000	1	100	100	100	100	100	100	100	100	99.58	69.74	9.56
	bic	100	100	100	100	100	100	100	100	99.72	70.24	9.74
	aic	100	100	100	100	100	100	99.94	99.74	98.30	76.60	18.80
10000	1	100	100	100	100	100	100	100	100	100	92.02	6.86
	bic	100	100	100	100	100	100	100	100	100	100	6.76
	aic	100	100	100	100	100	100	100	100	100	99.68	12.44

Tablo 2.7. Deterministik Bileşenli Durumda AFD-F Testinin I. Tip Hata Olasılığı ve Testin Gücü

Sabit Terim												
$\rho = 0.6$												
T/d_1	p	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
100	1	98.40	94.94	85.64	72.74	55.76	41.30	29.86	21.92	15.88	13.14	11.86
	bic	99.38	97.20	90.96	80.04	64.20	50.70	37.58	28.24	20.16	16.08	13.42
	aic	99.38	98.08	95.4	88.76	78.58	67.64	54.56	43.06	31.92	26.10	21.14
200	1	100	99.94	99.16	95.52	85.38	66.98	45.06	31.60	19.64	12.66	9.94
	bic	100	100	99.64	97.92	90.88	74.94	54.58	38.78	24.38	15.22	11.16
	aic	100	99.96	99.48	99.30	96.42	88.84	74.56	58.78	40.32	26	17.02
400	1	100	100	100	100	99.16	91.72	72.26	48.16	27.52	14.68	9.20
	bic	100	100	100	100	99.74	96.00	80.88	56.18	32.44	17.66	10.36
	aic	100	100	99.98	99.98	99.84	98.98	94.66	78.82	54.56	29.66	15.32
1000	1	100	100	100	100	100	100	98.44	82.54	49.34	19.76	7.38
	bic	100	100	100	100	100	100	99.46	89.50	56.56	22.36	7.72
	aic	100	100	100	100	100	99.98	99.82	97.76	81.34	41.44	13.28
10000	1	100	100	100	100	100	100	100	100	100	78.04	6.08
	bic	100	100	100	100	100	100	100	100	100	87.02	6.08
	aic	100	100	100	100	100	100	100	100	100	97.62	10.14
Sabit Terim ve Trend												
$\rho = 0.6$												
T/d	p	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
100	1	98.44	95.22	86.42	73.82	57.94	43.78	33.28	25.66	20.52	18.86	18.78
	bic	99.44	97.36	91.54	81.24	66.60	53.78	41.82	33.86	26.62	24.14	22.44
	aic	99.46	98.50	95.96	90.56	82.48	72.84	62.50	52.94	45.34	40.86	37.24
200	1	100	99.92	99.18	95.56	85.64	68.38	47.48	34.54	23.30	17.20	15.04
	bic	100	100	99.64	97.96	91.04	76.50	57.12	42.58	29.18	20.80	17.18
	aic	100	99.96	99.60	99.38	96.94	90.54	78.26	65.02	49.48	37.28	28.56
400	1	100	100	100	100	99.16	91.98	73.24	50.36	31.06	18.52	12.96
	bic	100	100	100	100	99.72	96.30	81.42	58.50	36.60	20.84	14.40
	aic	100	100	99.98	99.98	99.90	99.08	95.22	81.72	61.76	39.26	24.82
1000	1	100	100	100	100	100	100	98.50	83.14	51.44	22.62	9.78
	bic	100	100	100	100	100	100	99.46	89.98	59.38	27.40	10.10
	aic	100	100	100	100	100	100	99.88	98.12	84.32	47.90	19.10
10000	1	100	100	100	100	100	100	100	100	100	79.82	7.14
	bic	100	100	100	100	100	100	100	100	100	88.96	7.14
	aic	100	100	100	100	100	100	100	100	100	98.18	12.98

2.4. Zaman Serilerinde Bootstrap

Bootstrap yöntemi genel hatlarıyla simülasyona benzese de bir noktada ayrılmaktadır²¹. Simülasyon yapılırken veriler tamamen yapay olarak oluşturulmakta iken bootstrap yönteminde ise gerçek bir veri seti kullanılmaktadır. Eldeki tek bir veri seti popülasyon olarak değerlendirilmekte ve yinelemeli örnekler çekilmektedir. Her bir çekilen örneklemden istatistikler hesaplanmakta ve bunların dağılımları hakkında çıkarımlar yapılmaktadır. Bootstrap yönteminin mantığını kavramak için elde $x_1, \dots, x_n$ şeklinde bir örnek olduğu ve bu örneğin $F(x) = P(x_i \leq x)$ birikimli olasılık fonksiyonuna sahip bir olasılık dağılımından rastgele olarak çekildiğini farz edildiğinde ilgilenilen istatistikler $T_n(x_1, \dots, x_n)$ şeklinde gösterilebilmektedir. Burada T_n olarak ifade edilen istatistikler genellikle bilinmeyen popülasyon parametrelerinin tahmincilerini içerisinde barındırmakta ve verinin rastgele olarak çekildiği popülasyonun dağılımı ile ilişkilidir. Daha açık ifade etmek gerekirse popülasyon ortalaması μ 'nın tahmincisi olarak örnek ortalamasını ele alındığında T_n aşağıdaki eşitlikle ifade edilmektedir:

$$T_n = \frac{1}{n} \sum_{i=1}^n x_i \quad (2.37)$$

Burada T_n örnek ortalamasıdır ve çekilen her bir farklı örnekten farklı örnek ortalamaları elde edileceği için yani istatistik olduğu için kendisi de rastgele bir değişkendir ve bundan dolayı bir olasılık dağılımına sahiptir. Bu örnek ortalamasının dağılımını matematiksel olarak aşağıdaki gibi ifade edilebilmektedir:

$$G_n(x) = P(T_n \leq x) \quad (2.38)$$

Burada istatistiğin dağılımı $G_n(x)$ kullanılarak μ için istatistiksel çıkarımlar yapılabilmektedir. Öte yandan $G_n(x)$, veri setinin rastgele olarak çekildiği popülasyonun dağılımına bağlıdır. Bu bağımlılık yapısı;

²¹ Brooks, a.g.e , s.553

$$G_n(x, F) \quad (2.39)$$

şeklinde yazılabilmektedir. (2.39) incelendiğinde örnek ortalamasının olasılık dağılımının bu örneğin çekildiği popülasyonun dağılımına bağlı olduğu görülmektedir. Fakat çoğu zaman veri setinin çekildiği popülasyonun dağılımı bilinmemekte ve hesaplanamamaktadır. Dolayısıyla istatistiklerin gerçek dağılımları bilinmeyen popülasyon dağılımlarına bağımlı olmaktadır²². Bundan dolayı istatistiğin olasılık dağılımının (G_n) tahmin edilmesi diğer bir ifadeyle elde edilmesi gerekmektedir. İstatistiğin olasılık dağılımını tahmin etmenin bir yolu gözlem sayısının sonsuza doğru arttığı durumda ($n \rightarrow \infty$) asimptotik teoriyi kullanmaktır. Bu duruma örnek olarak merkezi limit teoremi gösterilebilir. Merkezi limit teoremi kullanılarak istatistiğin örnekleme dağılımı tahmin edilebilmektedir. Birçok istatistiğin asimptotik dağılımı ($G_\infty(x)$) popülasyonun dağılımına (F) bağlı değildir. Öte yandan bu tarz yaklaşımlar gözlem sayısının büyük olduğu durumlar için geçerli olmaktadır. Fakat uygulamada genellikle gözlem sayısı (n) çok küçüktür ve bu sebeple asimptotik teoriyi kullanarak istatistiğin dağılımını tahmin etmek çok zordur. Bu problemin çözümü için bootstrap yöntemine başvurularak istatistiğin olasılık dağılımı tahmin edilebilmektedir. Bootstrap yönteminde istatistiğin asimptotik dağılımını kullanmak yerine popülasyonun dağılımı tahmin edilmektedir ($\hat{F}$). Popülasyonun dağılımını tahmin etmek için iki türlü yaklaşım bulunmaktadır. Bu yaklaşımlardan ilki parametrik bootstrap yöntemidir. Örneğin, popülasyonun dağılımının bilindiği fakat ortalama ve varyansı bilinmediği durumda popülasyonun dağılımının tahmini ($\hat{F}$), tahmin edilen ortalama ve varyansla normal dağılıma sahiptir denilebilmektedir. Diğer yaklaşım ise parametrik olmayan bootstrap yöntemidir. Parametrik olmayan bootstrap yöntemi popülasyonun dağılımı ile ilgili hiçbir bilgi olmadığı durumda başvurulacak bir yöntemdir. Bu durumda popülasyonun dağılımı (F), ampirik dağılım fonksiyonu (EDF) elde edilerek tahmin edilmektedir. Ampirik dağılım fonksiyonu aşağıdaki şekilde ifade edilmektedir:

²² Stephan Smeekes, “Bootstrapping Nonstationary Time Series”, Maastricht University, **Yayınlanmış Doktora Tezi**, 2009, s.3-4.

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n I(x_i \leq x) \quad (2.40)$$

Burada A doğru ise $I(A) = 1$ değilse 0 değerini almaktadır²³. $\hat{F}_n$ elde edildiği zaman ilgilenilen istatistiğin olasılık dağılımı $G_n(x, \hat{F}_n)$ elde edilebilmektedir. Elde edilen istatistiğin olasılık dağılımı bootstrap dağılımı olarak ifade edilmektedir ve $G^*(x)$ şeklinde ifade edilmektedir. Burada ‘*’ işareti bootstrap yöntemi kullanılarak elde edilen çıktıları göstermektedir.

Genellikle uygulamada istatistiğin olasılık dağılımının analitik olarak ifade edilmesi çok karmaşıktır. Çünkü gerçekte elde tek bir örnek seti vardır ve bu örnek seti kullanılarak çeşitli istatistikler hesaplanmaktadır. Bu problemin üstesinden gelmek için simülasyon yöntemlerine başvurulmaktadır. Simülasyon sayesinde popülasyonun tahmin edilen dağılımından ($\hat{F}_n$) yeni örnekler elde edilmektedir. Bu elde edilen yeni örnekler bootstrap örnekleri adı verilmektedir. Bootstrap örnek seti $x_1^*, \dots, x_n^*$ şeklinde ifade edilmektedir. Tekrarlı örnekler yardımıyla istatistiğin olasılık dağılımı $G_n(x, \hat{F}_n)$ elde edilmektedir. Burada ampirik olasılık dağılımı $\hat{F}_n$ olarak düşünülürse, $\hat{F}_n$ ’den örnek çekme işlemi, eldeki gerçek örnekten $(x_1, \dots, x_n)$ tekrarlı örnek setleri çekme anlamına gelmektedir. Burada her bir gözlem noktası $\frac{1}{n}$ olasılıkla çekilmektedir. Dolayısıyla bu yöntemle *i.i.d.* bootstrap yöntemi adı verilmektedir. *i.i.d* bootstrap algoritması aşağıdaki gibidir:

- i. Eldeki veri setinden $(x_1, \dots, x_n)$ tekrarlı olarak örnekler çekilir ve bootstrap veri setleri oluşturulur $(x_1^*, \dots, x_n^*)$.
- ii. Bootstrap veri setleri yardımıyla istatistikler hesaplanır $(T_n^* = T_n(x_1^*, \dots, x_n^*))$.
- iii. 1. ve 2. adımlar yapılacak bootstrap sayısı (B) kadar tekrar edilir ve istatistiklerin olasılık dağılımları $G_n(x, \hat{F}_n) = P^*(T_n^* \leq x)$ aşağıdaki eşitlik yardımıyla tahmin edilir:

²³ Stephan Smeekes, a.g.e. s.3.

$$\frac{1}{B} \sum_{l=1}^B I(T_{n,b}^* \leq x) \quad (2.41)$$

Zaman serisi analizlerinde veriler rastgele bir örnekleme süreci sonucunda elde edilmediği için *i.i.d* bootstrap yöntemi geçerli olmamaktadır. Zaman serisi verilerinde gözlem değerleri rastgele olarak elde edilmemekte ve farklı zaman noktaları arasında bağımlılık gözlenmektedir.

Bu noktadan hareketle zaman serisi analizlerinde yaygın olarak kullanılan iki farklı bootstrap yöntemi vardır. Birincisi, Carlstein (1986) ve Künsch (1989) tarafından geliştirilen Block bootstrap yöntemidir. Block bootstrap yöntemi parametrik olmayan bir yöntemdir. Bu yöntemde veri seti birkaç alt gruba ayrılır, alt gruplar tekrarlı olarak çekilir ve gerekli analizler yapılır. Burada amaç grupların içindeki gözlem verilerinin bağımlılık yapısını korumaktır. Diğer yöntem ise Kreiss (1992) ve Bühlmann (1997) tarafından geliştirilen yarı parametrik Sieve bootstrap yöntemidir. Bu yöntemde verinin bağımlılık yapısı modellenmekte ve bu modelden elde edilen kalıntılar *i.i.d.* olarak değerlendirilmektedir. Daha sonra *i.i.d.* bootstrap yöntemi uygulanarak tahmin edilen model, gerçek veri seti ile aynı bağımlılık yapısına sahip bootstrap örneklerini elde etmek için kullanılmaktadır. Durağan zaman serisi analizlerinde kullanılan sieve bootstrap yönteminin aşamaları genel hatlarıyla aşağıdaki gibidir²⁴:

- i. Otoregresif model p gecikme için tahmin edilir:

$$x_t = \sum_{j=1}^p \hat{\phi}_j x_{t-j} + \hat{\varepsilon}_t \quad (2.42)$$

- ii. İlk aşamada elde edilen kalıntılardan $(\hat{\varepsilon}_{p+1}, \dots, \hat{\varepsilon}_n)$ yerine koyma yöntemiyle tekrarlı olarak örnekler çekilir ve bootstrap kalıntıları oluşturulur (ε_t^*) .

²⁴ A.e. , s.9.

iii. Bu aşamada x_t^* bootstrap serisi oluşturulur:

$$x_t^* = \sum_{j=1}^p \hat{\phi}_j x_{t-j}^* + \varepsilon_t^* \quad (2.43)$$

iv. Üçüncü aşamada elde edilen regresyon modeli kullanılarak çeşitli test istatistikleri ($T_n^* = T_n(x_1^*, \dots, x_n^*)$) hesaplanır.

iv. Son olarak (ii–iv) adımları B kez tekrarlanır ve B kadar bootstrap test istatistikleri ($T_{n,b}^*$) elde edilir. Son aşamada istatistiklerin olasılık dağılımları $G_n(x, \hat{F}_n) = P^*(T_n^* \leq x)$;

$$\frac{1}{B} \sum_{i=1}^B I(T_{n,b}^* \leq x) \quad (2.44)$$

şeklinde elde edilebilmektedir.

2.4.1. Kesirli Birim Kök Testi İçin Sieve Bootstrap Algoritması

Daha önce de bahsedildiği gibi modelin hata terimlerinin otokorelasyonlu bir yapıya sahip olduğu durumda AFD-F testinin I. tip hata olasılığı bozulmaktadır. Diğer bir ifadeyle gerçekte doğru olan temel hipotezinin reddedilme olasılığı artmaktadır. Bunun sonucu olarak da testin güvenilirliği azalmaktadır. Bahsi geçen problemin üstesinden gelebilmek için bu tez çalışmasında hata terimlerindeki otokorelasyonu modelleyebilen sieve bootstrap yöntemi kullanılmıştır.

Sieve bootstrap algoritması aşağıdaki gibidir:

- i. X_t serisinin birim kök temel hipotezi altında birinci mertebeden farkı alınır ve bu seri $\hat{u}_t$ ile gösterilirse daha sonra $\hat{u}_t$ kendi gecikmeleri üzerine regresyona sokulup bu regresyondan $\hat{\varepsilon}_t$ elde edilir:

$$\hat{u}_t = \Delta X_t \quad (2.45)$$

$$\hat{\varepsilon}_t = \hat{u}_t - \sum_{j=1}^p \hat{\phi}_j \hat{u}_{t-j} \quad (2.46)$$

- ii. *i.i.d* ($\hat{\varepsilon}_t$) kalıntılarının ortalamasının 0 olmasını garanti altına almak için ortalamasından farkı alınarak $\tilde{\varepsilon}_t$ serisi oluşturulur:

$$\tilde{\varepsilon}_t = \hat{\varepsilon}_t - \frac{1}{n} \sum_{t=1}^n \hat{\varepsilon}_t \quad (2.47)$$

- iii. $\tilde{\varepsilon}_t$ 'dan yerine koyma yöntemiyle tekrarlı olarak örnekler çekilir ve bootstrap kalıntıları oluşturulur (ε_t^*). Daha sonra, ε_t^* ve (i). adımdan elde edilen $\hat{\phi}_j$ 'ler kullanılarak bootstrap örnekleri u_t^* 'ler bulunur:

$$u_t^* = \sum_{j=1}^p \hat{\phi}_j u_{t-j}^* + \varepsilon_t^* \quad t = p + 1, \dots, T \quad (2.48)$$

Burada u_t^* 'nin ilk p değeri $\hat{u}_t$ 'nin ilk p değeri olarak alınabilir.

- iv. Durağan olmayan bootstrap serisi x_t^* , u_t^* 'den hareketle aşağıdaki gibi oluşturulur:

$$X_t^* = X_{t-1}^* + u_t^* \quad (2.49)$$

- v. X_t^* serisi kullanılarak AFD-F test istatistiği hesaplanır.

- vi. (ii) ve (v). adımlar B kadar tekrarlanır ve $t_{\hat{\psi}}^* = (t_{\hat{\psi}_1}^*, t_{\hat{\psi}_2}^*, \dots, t_{\hat{\psi}_B}^*)'$ bootstrap test istatistikleri elde edilir.

Elde edilen bootstrap test istatistiklerinin belirlenen anlamlılık düzeyinde kuantilleri alınarak kritik değerleri bulunur. Bunu göstermek için $t_{\hat{\varphi}}^*$ 'nın ampirik dağılımı $D_T^*(m)$ olmak üzere α . örnek kantili

$$q_T(\alpha) = \inf\{m: D_T^*(m) \geq \alpha\} \quad \alpha \in [0,1] \text{ için} \quad (2.50)$$

ile elde edilir. Şayet modelde deterministik bileşenler varsa;

$$\tilde{X}_t = \begin{cases} X_t - \bar{X}, & \text{sabit terimli model için} \\ X_t - \hat{\beta}_0 - \hat{\beta}_1 t, & \text{sabit terim ve trenli model için} \end{cases}$$

şeklinde X_t serisi sabit terim ve/veya deterministik trendden arındırılır. Burada, $\hat{\beta}_0$ ve $\hat{\beta}_1$ en küçük kareler tahmincileridir. Yukarıdaki işlemler hem orjinal veriye hem de bootstrap örnekleri X_t^* 'ye uygulanır.

2.4.2. Monte Carlo Simülasyon Sonuçları

Bu kısımda, daha önce de bahsedildiği gibi deterministik bileşenlerin varlığında ve modelin hata teriminin otokorelasyonlu yapıya sahip olduğu durumda AFD-F birim kök testinin bozulan I.tip hata olasılığı sieve bootstrap yöntemiyle düzeltilerek tablolastırılmıştır. AFD-F birim kök testinin bootstrap versiyonun sınırlı veri kümesindeki performansını incelemek için kullanılan veri üretim süreci (2.34), (2.35) ve (2.36) verilen ile aynıdır. Veri üretim sürecindeki tüm parametre değerleri de aynı alınmıştır. Sieve bootstrap algoritmasında p gecikme uzunluğu schwarz bilgi kriteri ile seçilmiştir. Chang ve Park (2003), veri üretim sürecinde otoregresif gecikme uzunluğu bilindiğinden dolayı schwarz bilgi kriteri akaike bilgi kriterine göre daha tutarlı sonuçlar verdiğini belirtmektedir. Örnek büyüklüğünün $T = 100, 200$ ve 400 değerlerinde 5000 Monte Carlo denemesi yapılmıştır. Aynı zamanda bootstrap sayısı ise Monte Carlo aşamasında Giacomini, Politis ve White (2013) tarafından geliştirilen “Warp Speed” yöntemine göre 1 olarak alınmıştır.

Tablo 2.8. AFD-F Sieve Bootstrap Testinin I. Tip Hata Olasılığı ve Testin Gücü

Deterministik Bileşensiz											
$\rho = 0.2$											
T/d_1	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
100	90.48	84.26	74.88	60.82	48.56	43.26	40.56	26.32	17.64	9.06	4.02
200	97.18	93.96	87.04	76.78	70.42	57.66	50.62	42.34	32.16	15.04	4.28
400	99.28	97.58	94.10	88.36	83.38	73.48	66.00	56.10	42.88	22.76	4.90
$\rho = 0.6$											
T/d_1	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
100	68.38	64.48	57.78	46.60	39.06	26.06	19.66	13.10	9.44	5.38	4.90
200	84.04	79.74	79.78	71.84	68.06	52.52	39.44	25.18	13.26	8.00	5.28
400	95.00	93.34	90.66	87.68	84.90	82.84	70.02	45.20	23.56	10.54	5.22
Sabit Terimli											
$\rho = 0.2$											
T/d_1	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
100	61.78	53.04	46.52	33.08	29.00	29.44	24.92	20.26	15.42	7.50	4.08
200	95.00	93.34	90.66	87.68	84.90	82.84	70.02	45.20	23.56	10.54	5.22
400	85.86	81.28	70.02	60.54	52.16	43.96	39.52	34.12	26.56	12.90	4.38
$\rho = 0.6$											
T/d_1	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
100	44.52	46.16	43.34	34.26	27.34	19.42	15.26	10.76	8.30	5.04	5.22
200	72.40	68.68	68.90	62.00	60.72	45.86	33.12	20.32	11.86	7.20	4.96
400	90.90	88.60	84.66	84.02	80.76	79.18	63.28	39.60	19.94	9.04	4.90
Sabit Terim ve Trendli											
$\rho = 0.2$											
T/d_1	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
100	38.04	31.28	26.00	25.60	17.64	14.08	16.02	14.28	8.94	6.16	4.12
200	69.80	63.90	51.84	45.16	37.42	36.40	28.94	24.60	19.50	11.34	4.10
400	90.62	84.34	77.20	67.74	52.80	51.90	45.94	33.10	32.22	17.94	4.10
$\rho = 0.6$											
T/d_1	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
100	28.56	25.54	20.92	21.20	18.24	14.54	10.22	8.06	5.98	4.40	3.30
200	54.70	51.88	45.96	47.12	45.72	34.34	22.02	16.30	9.04	5.48	4.64
400	78.70	75.82	67.84	72.80	71.18	68.20	56.66	32.80	17.62	7.66	4.52

Tablo 2.8'deki sonuçlara göre üç durum için de Sieve bootstrap'a dayalı AFD-F testinin I. tip hata olasılığındaki bozulmaların düzelerek %5 anlamlılık düzeyine yaklaştığı görülmektedir. Örneğin $T = 100$, $\rho = 0.2$ olduğunda deterministik bileşensiz AFD-F testinin bootstrap versiyonunun I. tip hata olasılığı %4.02'dir ve anlamlılık düzeyi %5'e oldukça yakındır. Hata terimindeki otokorelasyonun şiddeti arttığında, diğer bir ifadeyle $\rho = 0.6$ ve $T = 100$ olduğunda

deterministik bileşensiz AFD-F testinin bootstrap versiyonunun I. tip hata olasılığı %4.90'dir. Testin gücü, kesirli fark parametresinin değeri 1'e yaklaştıkça önemli ölçüde azalmaktadır. Örnek olması açısından $d_1 = 0.9$ için $T = 100$, $\rho = 0.2$ durumunda testin gücü %9.06'dır. Gözlem değeri arttıkça testin gücünde iyileşmeler görülse de yine de çok azdır. Örneğin $d_1 = 0.9$, $T = 400$, $\rho = 0.2$ olduğunda testin gücü %22.76'dır. Modeldeki otokorelasyonun şiddeti arttıkça testin gücündeki azalışlar da paralel olarak artmaktadır. Örneğin $\rho = 0.6$ ve $T = 100$ durumunda testin gücü %5.38'dir. Gözlem değeri arttıkça testin gücünde kayda değer olmamakla beraber çok az artışlar görülmektedir. Örneğin $d_1 = 0.9$, $\rho = 0.6$ ve $T = 400$ testin gücü %10.54'tür.

Öte yandan $T = 100$, $\rho = 0.2$ ve $\rho = 0.6$ olduğunda sabit terimli AFD-F testinin bootstrap versiyonunun I. tip olasılığı ise sırasıyla %4.08 ve %5.22 ile nominal anlamlılık düzeyine oldukça yakın bulunmuştur. Modelde deterministik bileşenlerin yer almadığı durumda ki gibi sabit terimli modelde testin gücü $d_1 \rightarrow 1$ durumunda önemli ölçüde azalmaktadır. Örneğin $d_1 = 0.9$, $T = 100$, $\rho = 0.2$ için testin gücü %7.50 iken $\rho = 0.6$ olduğunda bu değer %5.04'tür. Yine $d_1 = 0.9$, $T = 400$, $\rho = 0.2$ durumunda testin gücü %12.90, $\rho = 0.6$ için bu değer %9.04'tür.

Sabit terimli ve deterministik trendli model incelediğinde ise $T = 100$, $\rho = 0.2$ ve $\rho = 0.6$ olduğunda testin I. tip olasılığı sırasıyla %4.12 ve %3.30'dur. Diğer durumlarda olduğu gibi burada da testin gücü kesirli fark parametresinin değeri 1'e doğru artarken oldukça düşmektedir. Gözlem sayısının büyük olduğu durumda bile testin gücü oldukça düşüktür. Örneğin $d_1 = 0.9$, $T = 100$, $\rho = 0.2$ durumunda testin gücü %6.16 iken $\rho = 0.6$ için %4.40'tır. Ayrıca $d_1 = 0.9$, $T = 400$, $\rho = 0.2$ ve $\rho = 0.6$, durumunda testin gücü sırasıyla %17.94 ve %7.66'dır.

Tablo 2.8'de ki sonuçlara göre AFD-F Sieve bootstrap testinin I. Tip Hata olasılığındaki bozulmaların düzeldiği fakat testin gücünün önemli ölçüde düştüğü görülmektedir. Daha önce de bahsedildiği gibi testin I. Tip ve II. Tip hata olasılığı arasında ters yönlü bir ilişki mevcut olduğu için AFD-F Sieve bootstrap testinde I. Tip hata olasılığındaki bozulmalar azalırken II. Tip hata olasılığında artışlar meydana

gelmektedir. Sonuç olarak da testin gücü azalmaktadır. Dolayısıyla bu tez çalışmasında ADF-F testinde I. Tip hata olasılığındaki bozulmalar ele alınmış ve Sieve bootstrap yöntemiyle düzeltilerek AFD-F Sieve bootstrap testi olarak literatüre kazandırılmıştır.



ÜÇÜNCÜ BÖLÜM

FARKLI GELİŞMİŞLİK DÜZEYLERİNE SAHİP AVRUPA ÜLKELERİNDE ETKİN PİYASA HİPOTEZİNİN GEÇERLİLİĞİNİN SINANMASI

Tezin bu bölümü üç kısımdan oluşmaktadır. İlk kısımda etkin piyasa hipotezinin anlamı ve türleri açıklanacaktır ve sonrasında piyasaların etkinliği ile ilgili literatürde yer alan çalışmalardan bahsedilecektir. İkinci kısımda tezde kullanılan veri seti ve yöntem anlatılacaktır. Son kısımda ise ikinci bölümde önerilen Sieve bootstrap kesirli bütünleşme testi aracılığıyla çeşitli ülke gruplarının sermaye piyasaları için yapılan etkinlik analizlerinin sonuçları verilerek elde edilen sonuçlar yorumlanacaktır.

3.1. Etkin Piyasa Hipotezi

Etkin piyasa hipotezine göre, sermaye piyasalarında işlem gören finansal varlık fiyatları, piyasadaki tüm bilgileri yansıtmaktadır. Nath (2001), yaptıkları çalışmada piyasanın etkin olabilmesi için, sermaye piyasalarında kar amacı güden çok fazla katılımcı olması ve bu katılımcıların her birinin bir diğerinden bağımsız olması gerektiği, finansal varlık fiyatlarıyla ilgili piyasalara gelen bilgilerin tamamen rastgele olduğu ve yatırımcıların piyasaya gelen bu yeni bilgilerin etkisini yansıtmak için varlık fiyatlarında hızlı bir şekilde düzenlemeler yaptığı şeklinde varsayımları sağlaması gerektiğini belirtmektedir.

Fama (1970) çalışmasında piyasadaki bilginin derecesine bağlı olarak etkin piyasa hipotezini üç farklı formda aşağıdaki gibi ifade etmektedir:

- i. Zayıf formda etkinlik
- ii. Orta kuvvetli formda etkinlik
- iii. Güçlü formda etkinlik

Zayıf formda etkinlikte piyasadaki varlık fiyatları geçmişteki tüm bilgileri içermektedir. Öte yandan orta kuvvetli formda etkinlikte hisse senedi fiyatları halka açık bilgileri anında yansıtmaktadır. Güçlü formda etkinlikte ise, finansal değerlerin fiyatları, halka açıklanmayan bilgiler de dahil olmak üzere tüm piyasadaki tüm bilgileri içerisinde barındırmaktadır. Bu noktadan hareketle menkul kıymetler piyasası zayıf formda etkin olduğu durumda yatırımcıların, yatırım araçlarının geçmiş fiyat bilgileri kullanılarak gelecekteki piyasa değerlerinin tahmin edemeyeceği ifade edilmektedir. Piyasa etkin olduğu durumda yeni bilgilerin elde edilmesine karşılık fiyatlarda dalgalanmalar meydana gelmektedir²⁵. Piyasaya giren bilgiler rassal olduğu için hisse senedi fiyatlarındaki değişimlerin tahmin edilemez olması gerektiği belirtilmektedir.

Finansal varlık fiyatlarının rassal yürüyüş sürecine sahip olduğu durumda varlıkların geçmiş fiyat bilgileri kullanılarak gelecekteki değeri tahmin edilememekte ve böylelikle etkin piyasa hipotezi geçerli olmaktadır. Öte yandan hisse senedi fiyatlarının rassal yürüyüş sürecine sahip olmadığı ve ortalamasına geri dönme eğiliminde olduğu durumda, varlığın geçmişteki fiyat bilgileri gelecekteki değerini tahmin etmede kullanılabilmektedir.

Ekonometrik çerçeveden bakıldığında bir serinin rastgele yürüyüş sürecine sahip olup olmamasının araştırılması ile serilerin bütünleşme dereceleri arasında doğrudan bir ilişki söz konusudur. Öte yandan durağan olmayan serilerin rastgele yürüyüş sürecine sahip olduğu ifade edilmektedir. Bundan dolayı literatüre bakıldığında hisse senedi fiyatlarının durağanlık analizleri yapılarak piyasaların etkin olup olmadıkları incelenmiştir. Başka bir ifadeyle hisse senedi fiyatlarının durağanlık analizleri birim kök testleri yardımıyla ortaya konulmuştur.

Literatüre bakıldığında etkin piyasa hipotezi, farklı yöntem ve veri setleri kullanılarak araştırmacılar tarafından incelenmiştir. Bu incelemelerin geçmişi 1960'lı yıllara kadar uzanmaktadır. Bu yıllarda yapılan çalışmalara örnek olarak Fama

²⁵ Balcilar, M., Cakan, E., & Ozdemir, Z. A. ,”Structural breaks, long memory, or unit roots in stock prices: Evidence from emerging market”. International Econometric Review, 2015 , s.13

(1965) ve Samuelson (1965) örnek olarak verilebilir. Fama (1965), çalışmada kullandığı geniş veri setinin dağılımını analiz etmiştir ve rassal yürüyüş sürecinin geçerli olduğuna dair bulgular elde etmiştir. Diğer bir ifadeyle Fama (1965) çalışmasında hisse senedi piyasasının etkin olduğuna dair kanıtlar elde etmiştir. Sharma ve Kennedy (1977) yaptıkları çalışmada 1963 ile 1973 yıl aralığında aylık endeks verilerini kullanarak Hindistan'daki Bombay Borsasını incelemişlerdir. Analiz sonucunda piyasanın zayıf formda etkin olduğuna dair bulgular elde etmişlerdir.

Chan vd. (1997), yaptığı çalışmasında 1961 ile 1992 aralığında aylık veriler kullanarak 18 farklı ülke için zayıf formda piyasa etkinliğini araştırmışlardır. Çalışmanın sonucunda her bir piyasa için zayıf formda etkinliğin geçerli olduğuna dair sonuçlar elde edilmiştir. Andor vd. (1999), korelasyon ve koşu testleri aracılığıyla Macaristan piyasası için etkin piyasa hipotezinin geçerliliğini sınamışlardır. Yapılan analizler sonucunda piyasanın zayıf formda etkin olduğuna dair kanıtlar elde edilmiştir.

Literatürde Avrupa ülkeleri için yapılan çalışmalara bakıldığında farklı veri aralığı ve testler için farklı sonuçlar elde edildiği görülmektedir. Harvey (1993), gelişmekte olan ülkeler için yapmış olduğu piyasa etkinliği analizi sonucunda piyasaların etkin olmadığı bulgusuna rastlamıştır. Öte yandan Aktan, Iren ve Omay (2019), yaptıkları çalışmada 32 Avrupa ülkesinde hisse senedi piyasaları için etkinlik analizi yapmışlardır. Analiz sonucunda gelişmekte ve gelişmiş ülkelerde etkin piyasa hipotezinin geçerli olmadığı bulgusunu rastlamışlardır. Worthington ve Higgs (2004), gelişmekte olan ülkelerin çoğunluğunda piyasaların etkin olmadığı fakat gelişmiş ülkelerde bu durumun tersi olduğu sonucuna ulaşmıştır. Chaudhuri ve Wu (2003), gelişmekte olan ülke piyasalarının zayıf formda etkin olup olmadıklarını araştırmış ve sonuç olarak ülkelerde etkinlik hipotezinin geçerli olmadığına dair bulgular elde etmişlerdir.

Son zamanlarda etkin piyasa hipotezi uzun hafızalı modeller aracılığıyla da incelenmeye çalışılmıştır. Balcılar, Cakan ve Özdemir (2015) yaptıkları çalışmada 14

gelişmekte olan ülkelerin hisse senedi piyasalarında etkin piyasa hipotezinin geçerliliğini sınamışlardır. Araştırmacılar serilerin temel hipotez altında birim köklü olması durumunu alternatif hipotez altında kesirli bütünleşik sürece sahip olmasına diğer adıyla serinin uzun hafıza özelliği göstermesi duruma karşı sınamışlardır. Analiz sonucunda 4 ülkede hisse senedi fiyatlarının kesirli bütünleşik sürece sahip olduğu yani ortalamasına geri dönme eğiliminde olduğu bulgularına rastlanmıştır. Bu ülkelerde etkin piyasa hipotezinin geçerli olmadığı sonucuna ulaşılmıştır. Geri kalan 10 ülkede ise, etkin piyasa hipotezinin geçerli olduğuna dair bulgular elde edilmiştir. Aye, Gil-Alana, Gupta ve Wohar (2017), 15 farklı fiyat endeksini doğrusal ve doğrusal olmayan uzun hafıza modelleri ile incelemişlerdir. Çalışmanın sonucunda kimi serilerde uzun hafıza özelliği görülmekle beraber etkin piyasa hipotezinin geçerli olmadığı sonucuna ulaşmışlardır. Caporale, Gil-Alana, Plastun ve Makarenko (2016), Ukrayna hisse senedi piyasasında market etkinliğini araştırmışlardır. Analizler farklı kesirli birim kök testleri aracılığı ile yapılmıştır. Çalışmanın sonucunda Ukrayna hisse senedi piyasasında zayıf formda etkinliğin geçerli olmadığı yönünde sonuçlar elde edilmiştir.

Genel olarak literatür incelendiğinde farklı ülke grupları, farklı veri setleri ve yöntemler kullanılarak bir çok çalışma yapılmıştır. Farklı teknikler ve ülke grupları için farklı sonuçlar elde edilmiştir. Bazı araştırmacılar etkin piyasa hipotezinin geçerli olduğu yönünde bulgular elde ederken kimileri hipotezin geçerli olmadığı yönünde bulgular elde etmiştir. Dolayısıyla veri setleri değiştikçe, yeni yöntemler geliştikçe piyasa etkinliği analiz edilmeye devam edecektir.

3.2. Veri Seti ve Yöntem

Bu tez çalışmasında Avrupa'da yer alan 32 ülke için hisse senedi piyasalarında etkin piyasa hipotezinin geçerliliği araştırılmıştır. Tezin önceki bölümlerinde açıklandığı gibi etkin piyasa hipotezinin farklı türde çeşitleri olmasıyla beraber bu tez çalışmasında piyasaların zayıf formda etkin olup olmadıkları incelenmiştir. Çalışmada yer alan ülkelerin hisse senedi endekslerinin kapanış fiyatları kullanılmıştır. Çalışmada veriler aylık olarak değerlendirilmiştir. Tezde

kullanılan veri aralığı Ocak 2009 ile Nisan 2019 aralığını kapsamaktadır. Analizde verilerin 2009 yılından başlamasının sebebi 2008 yılında meydana gelen küresel krizin serilerin durağanlık özelliklerini etkilemesidir. Öte yandan tezin önceki bölümlerinde de bahsedildiği gibi yapısal kırılmalara maruz kalan serilerin otokorelasyon fonksiyonları ile gerçek bir uzun hafıza özelliğine sahip serilerin otokorelasyon fonksiyonları birbirlerine benzerlik gösterdiğinden dolayı verilerin 2008 krizinin etkisine maruz kalmaması açısından 2009 yılından başlaması uygun bulunmuştur. Literatürde bu durumu birbirinden ayırmak için geliştirilmiş testler olsa da, bu tez çalışmasının kapsamı dışında kaldığı için yapısal kırılmalara neden olan aralıklar veri setinden çıkarılmıştır.

Tez çalışmasında kullanılan veriler “investing.org” sitesinden elde edilmiştir. Verisinde eksik olmayan 32 Avrupa ülkesi analize dahil edilmiştir. Analizde kullanılan verilerin endeksleri, kodları ve zaman aralıkları Ek A’da yer alan Tablo A.1’de gösterilmiştir. Tez çalışmasına konu olan ülkeler az gelişmiş, gelişmekte ve gelişmiş olmak üzere gruplandırılmıştır. Gelişmişlik dereceleri birçok faktöre dayanmakta olduğundan bu ayrımı yapabilmek için Financial Times Stock Exchange (FTSE) Mart 2017 “Pazar Kalitesi Kriterleri (Quality of Markets Criteria)” raporu temel alınmıştır²⁶.

Çalışmada verilerin logaritması alınarak analize dahil edilmiştir. Analizde I. tip hata olasılığı düzeltilmiş Dolado,Gonzalo ve Mayoral (2002) testi kullanılarak serilerde zayıf formda piyasa etkinliği hipotezinin geçerli olup olmadığı araştırılmıştır. Analizler yapılırken Matlab programı kullanılmıştır. Sieve bootstrap ile kritik değerlerin elde edilmesinde 50000 bootstrap tekrarı yapılmıştır. Sieve bootstrap AFD-F birim kök testinde hem test aşamasında hem de bootstrap algoritmasında gecikme uzunluğunun seçiminde maksimum uzunluk 12 olmak üzere Akaike bilgi kriteri kullanılmıştır. Analizlerde AFD-F eşitliğine deterministik bileşen olarak sabit terim ilave edilmiştir.

²⁶ Ceyda Aktan, “Avrupa Hisse Senedi Piyasalarında Zayıf Formda Piyasa Etkinliğinin Test Edilmesi”, Türk Hava Kurumu Üniversitesi, İşletme Anabilim Dalı, **Yayınlanmamış Doktora Tezi**, 2018, s.59

3.3. Araştırma Bulguları

Daha önce de belirtildiği gibi, zayıf formda etkin olan piyasalarda finansal varlık fiyatları rassal yürüyüş sürecine sahip olmakta ve böylelikle yatırımcıların geçmişteki fiyat bilgilerini kullanarak gelecek fiyatların geleceği hakkında anlamlı tahminler yürütemeyecekleri ifade edilmektedir. Bundan dolayı piyasada hiçbir yatırımcı normalin üzerinde getiri elde edemeyecektir.

Bu çalışmada 32 tane Avrupa ülkesi için hisse senedi piyasalarının durağanlık analizleri, bu tez kapsamında önerilen Sieve bootstrap'a dayalı kesirli birim kök testi aracılığı ile sınanmıştır. Analiz sonuçları Tablo 3.1 – Tablo 3.3'de gösterilmiştir.

Tablo 3.1. Avrupa'daki Gelişmemiş Ülkelerin Piyasaları İçin Etkin Piyasa Hipotezi'nin Sınanması

Ülke	$t_{\hat{\rho}}$	Bootstrap p-value
Bulgaristan	2.90	0.64
Hırvatistan	-3.58	0.08*
Kıbrıs Rum Kesimi	-1.24	0.13
Estonya	-0.47	0.33
Letonya	1.09	0.61
Litvanya	-2.16	0.07*
Romanya	-3.52	0.02**
Sırbistan	2.96	0.65
Slovakya	-0.75	0.56
Slovenya	0.64	0.62

Not: ** ve * sırasıyla %5 ve %10 anlamlılık düzeyini ifade etmektedir.

Tablo 3.1' deki sonuçlara göre Bulgaristan, Kıbrıs Rum Kesimi, Estonya, Letonya, Sırbistan, Slovakya, Slovenya ülkeleri için yapılan durağanlık analizi sonucuna göre piyasaların birim kök içermediği yönünde bulgular elde edilmiştir. Dolayısıyla bu ülke piyasalarının zayıf formda etkin olduğu anlaşılmaktadır. Hırvatistan ve Litvanya ülke piyasaları için yapılan analizler sonucunda %10 anlamlılık düzeyinde serilerin rassal yürüyüş sürecine sahip olduğunu belirten temel hipotez reddedilebilmektedir. Dolayısıyla Hırvatistan ve Litvanya ülkelerinin hisse senedi piyasaları %10 anlamlılık düzeyi için zayıf formda etkin değildir sonucuna

ulařılabilir. Romanya hisse senedi piyasası iin hisse senedi piyasasının durađan olmadıđını ifade eden temel hipotez %5 anlamlılık dzeyinde reddedilerek zayıf formda etkinliđin geerli olmadıđı sonucuna varılmıřtır.

Litvanya ve Hırvatistan piyasaları iin test istatistiđi asimptotik kritik deđerle yani -1.645 ile karřılařtırıldıđında serilerin durađan olmadıđını belirten temel hipotez %5 anlamlılık dzeyinde reddedilebilmektedir. Fakat bootstrap kritik deđerleri ile karřılařtırıldıđında serilerin birim kkl bir srece sahip olduđunu ifade eden temel hipotez %5 anlamlılık dzeyinde reddedilememektedir. Dolayısıyla AFD-F testinde I. tip hata olasılıđında bozulmalardan dolayı AFD-F testinin serilerin durađan olmadıđını ifade eden temel hipotezi, olması gerekenden daha ok reddettiđi anlařılmaktadır.

Tablo 3.2. Avrupa'daki Geliřmekte Olan lkelerin Piyasaları iin Etkin Piyasa Hipotezi'nin Sınanması

lke	$t_{\hat{\phi}}$	Bootstrap p-value
ek Cumhuriyeti	0.20	0.58
Yunanistan	0.15	0.31
Macaristan	1.67	0.62
İzlanda	-0.17	0.65
Polonya	-0.51	0.55
Rusya	-1.27	0.23
Trkiye	-1.08	0.32
Ukrayna	-2.67	0.02*

Not: ** ve * sırasıyla %5 ve %10 anlamlılık dzeyini ifade etmektedir.

Tablo 3.2'ye gre ek Cumhuriyeti, Yunanistan, İzlanda, Macaristan, Trkiye, Polonya, Rusya, lkelerinin hisse senedi piyasaları iin yapılan durađanlık sınaması sonularına gre piyasalarda zayıf formda etkinliđin geerli olduđu grlmektedir. Birim kk testi sonucuna gre Ukrayna iin zayıf formda etkinliđin geerli olmadıđı sonucuna varılmıřtır. Bu sonulara dayanarak Ukrayna hisse senedi piyasasında yatırımcılar gemiřteki fiyat hareketlerinden hareketle gelecekteki fiyatlar hakkında bilgi sahibi olabileceđi ifade edilebilmektedir. Dolayısıyla yatırımcıların Ukrayna hisse senedi piyasasında normalin zerinde getiriler elde edilebileceđi sylenebilmektedir.

Tablo 3.3. Avrupa'daki Gelişmiş Ülkelerin Piyasaları İçin Etkin Piyasa Hipotezi'nin Sınanması

Ülke	$t_{\hat{\psi}}$	Bootstrap p-value
Belçika	0.63	0.60
Danimarka	1.68	0.62
Finlandiya	0.96	0.46
Fransa	-0.46	0.41
Almanya	-0.68	0.11
İrlanda	0.20	0.43
İtalya	-0.22	0.55
Hollanda	-0.51	0.47
Norveç	0.51	0.44
Portekiz	1.20	0.63
İspanya	-0.16	0.58
İsveç	0.50	0.58
İsviçre	0.36	0.59
İngiltere	-1.56	0.41

Not: ** ve * sırasıyla %5 ve %10 anlamlılık düzeyini ifade etmektedir.

Tablo 3.3'e göre gelişmiş ülkelerdeki hisse senedi piyasası için yapılan durağanlık analizleri sonucunda hiçbir ülke için seride birim kökün olduğunu belirten boş hipotez reddedilememiş ve böylelikle bu ülkelerdeki piyasaların zayıf formda etkin olduğu sonucuna ulaşılmıştır.

SONUÇ VE ÖNERİLER

Bu tez çalışmasında, Avrupa ülkelerinin hisse senedi piyasalarında zayıf formda etkinlik kavramının geçerliliği kesirli birim kök testi bu tez kapsamında önerile Sieve bootstrap'a dayalı kesirli birim kök testi aracılığıyla araştırılmıştır. Tezin uygulama kısmında elde edilen sonuçların daha açık bir şekilde ifade edilebilmesi için ilk bölümde, bazı zaman serilerinin birbirlerine uzak gözlemleri arasında mesafe arttıkça istatistiksel bağımlılığının üstel azalma yerine hiperbolik olarak azalan bir yapı sergilediğini ifade eden uzun hafıza kavramı açıklanmıştır. Bu noktadan hareketle serilerin uzun ve kısa hafıza özelliklerini birlikte modelleyebilen ve literatürde sıklıkla kullanılan ARFIMA modelinden bahsedilmiştir.

İkinci bölümde serilerde kesirli bütünleşme kavramı açıklanmış, serilerde kesirli bütünleşme sürecinin araştırılmasıyla ilgili kullanılan testlerden birisi olan FD-F ve AFD-F testleri anlatılmıştır. İlk olarak FD-F testinin I. tip hata olasılığı ve gücü modelde deterministik bileşenler olmadığı durum için incelenmiştir. Yapılan analizler sonucunda modelde deterministik bileşenler olmadığı durumda FD-F testinin I. tip hata olasılığı ve gücünde herhangi bir bozulma olmadığı tespit edilmiştir. Öte yandan modelde deterministik bileşenler yer aldığı zaman testin I. tip hata olasılığı ve gücünde bozulmalar olduğu tespit edilmiştir. AFD-F testi için yapılan analizler sonucunda modelde deterministik bileşenler yer aldığı ve hata terimindeki otokorelasyonun şiddeti arttığı zaman testin I. tip hata olasılığı ve gücündeki bozulmaların çok daha ciddi olduğu sonucuna ulaşılmıştır. Bölümün son kısmında FD-F ve AFD-F testindeki bu bozulmaları düzeltmek için kullanılan sieve bootstrap yöntemi aşamalarıyla anlatılmıştır.

Çalışmanın son bölümünde 32 Avrupa ülkesi için hisse senedi piyasalarında zayıf formda etkinlik varsayımı incelenmiştir. Analizler yapılmadan önce etkin piyasa hipotezinin temeli, tarihçesi ve çeşitleri kısaca anlatılmıştır. Sonrasında uygulamada kullanılan veri seti ve yöntemden bahsedilmiştir. Çalışmada Ocak 2009 ve Ocak 2019 yıl aralığında doğal logaritması alınmış aylık veriler için bu tez çalışmasında önerilen Sieve bootstrap'e dayalı kesirli birim kök testi uygulanarak

serilerde zayıf formda etkin piyasa hipotezinin geçerliliği araştırılmıştır. Uygulama sonucunda gelişmemiş ülkeler grubunda Hırvatistan, Litvanya ve Romanya piyasaları için zayıf formda etkin piyasa hipotezinin geçerli olmadığı bulgusuna rastlanılmıştır. Öte yandan gelişmekte olan ülkeler için yapılan analizler sonucunda Ukrayna hariç tüm piyasalarda zayıf formda etkinliğin geçerli olduğu yönünde sonuçlar elde edilmiştir. Son aşamada gelişmiş ülke piyasaları için analizler yapılmış tüm ülkelerde zayıf formda etkinlik hipotezinin geçerli olduğu yönünde bulgular elde edilmiştir. Literatüre bakıldığında etkin piyasa hipotezi ile ilgili yapılan birçok çalışma birbirleri ile örtüşmeyen sonuçlar verse de genel anlamda gelişmiş ülke piyasaların gelişmekte olan ülke piyasalarından daha etkin olduğu ifade edilmektedir. Tez kapsamında yapılan uygulamanın sonucunda elde edilen bulgular da bu durumla örtüşmektedir. Dolayısıyla uygulama sonucunda, gelişmişlik düzeyleri gelişmiş ülkelere oranla daha az olan piyasalarda yatırımcılar geçmiş fiyat bilgilerinden hareketle gelecekteki fiyatlarla ilgili tahminler yapabilmekte ve normalin üzerinde kazançlar elde edebilmektedir sonucuna ulaşılmıştır.

Çalışmanın literatüre yaptığı katkı anlamında, DGM (2002) çalışmasından farklı olarak FD-F ve AFD-F testlerinin modelde deterministik bileşenler olduğu durum için testin I. tip hata olasılığı ve gücü incelenmiştir, hata terimlerindeki otokorelasyon probleminden dolayı AFD-F testinde bozulan I. tip hata olasılığı düzeltilmiştir. Bundan sonraki aşamalarda literatüre yapılan katkıyı arttırmak için modelde yapısal kırılmaları dikkate alan kesirli bütünleşme testlerinin kullanılması ve kesirli eş bütünleşme testleri kullanılarak piyasaların birbirleriyle olan etkileşimlerinin incelenmesi, çalışmaların detaylandırılması açısından olanak sağlayacaktır. Bununla beraber testlerin performansları değerlendirilirken modelde otokorelasyon problemine ek olarak değişen varyans probleminin de araştırılması, problemin olması durumunda wild bootstrap gibi çeşitli yöntemlerle testlerin güvenilirliğinin artırılması analizlerden daha doğru sonuçlar elde edilmesine olanak sağlayacaktır.

KAYNAKLAR

- Akaike, Hirotugu : "A new look at the statistical model identification", **Selected Papers of Hirotugu Akaike**. Springer, New York, NY, s.215-222, 1974.
- Aktan, Ceyda; Iren, Perihan; Omay, Tolga : "Market development and market efficiency: evidence based on nonlinear panel unit root tests", **The European Journal of Finance**, 25.11 , s.979-993 , 2019.
- Andor, György; Ormos, Mihály; Szabó, Balázs "Return Predictability in the Hungarian Capital Market", **Periodica Polytechnica Social and Management Sciences**, 7.1 , s. 29-46 , 1999.
- Aye, G. C. ; Gil-Alana, L. A. ; Gupta, R. ; & Wohar, M. E. : "The efficiency of the art market: Evidence from variance ratio tests, linear and nonlinear fractional integration approaches", **International Review of Economics & Finance**, 51, s.283-294 , 2017.
- Baillie , Richard : "Long memory processes and fractional integration in econometrics", **Journal of Econometrics**, 73(1), s.3-30, 1996.
- Balcılar, Mehmet; Cakan, Esin; Ozdemir, Zeynel Abidin : "Structural breaks, long memory, or unit roots in stock prices: Evidence from emerging markets", **International Econometric Review**, 7.1 , s.13-33 , 2015.
- Baltagi, Badi H. : **Econometrics**, 4. Bs. , Berlin , Springer, 2008.
- Baron, Michael : **Probability and statistics for computer scientists**, 2.bs, USA , Chapman and Hall/CRC, 2013.
- Beran, Jan : **Statistic For Long Memory Process**, Germany, Chapman & Hall, 1994
- Beran, Jan : "Maximum likelihood estimation of the differencing parameter for invertible short and long memory autoregressive integrated moving average models", **Journal of the Royal Statistical Society**, 57.4, s.659-672, 1995.
- Chan, Kam C. ; Gup, Benton E. ; Pan, Ming-Shiun : "International stock market efficiency and integration: A study of eighteen nations", **Journal of Business Finance**

- Chaudhuri, Kausik; Wu, Yangru : **& Accounting**, 24.6 , s.803-813 , 1997.
- Brooks, Chris : "Random walk versus breaking trend in stock prices: Evidence from emerging markets", **Journal of Banking & Finance** , 27.4 , s.575-592 , 2003.
- Bühlmann, Peter : **Introductory Econometrics for Finance**, 2.bs. Cambridge, Cambridge University Press, 2008.
- Caporale, G. M. ; Gil-Alana, L. ; Plastun, A. ; Makarenko, I. : "Sieve bootstrap for time series", **Bernoulli**, 3.2 , s.123-14 , 1997.
- Carlstein, Edward : "Long memory in the Ukrainian stock market and financial crises" **Journal of Economics and Finance**, 40.2 , s.235-257 , 2016.
- Yoosoon, Chang ; Joon, Park, : "The use of subseries values for estimating the variance of a general statistic from a stationary sequence", **The Annals of Statistics**, 14.3 ,s.1171-1179 , 1986.
- Chen, Chung ; TIAO, George C. : "A sieve bootstrap for the test of a unit root.", **Journal of Time Series Analysis**, s.379-400, 2003.
- Dickey, David A. ; Fuller, Wayne A. : "Random level-shift time series models, ARIMA approximations, and level-shift detection" **Journal of Business & Economic Statistics**, 8.1 , s. 83-97. , 1990
- Dickey, David A. ; Fuller, Wayne A. : "Distribution of the estimators for autoregressive time series with a unit root" **Journal Of The American Statistical Association**, 74.366, s.427-431, 1979.
- Diebold, Francis : "Distribution of the estimators for autoregressive time series with a unit root" **Econometrica**, 49, s.1057-1072, 1981.
- Dolado, Juan J.; Gonzalo, Jesús; Mayoral, Laura : "On the Power of the Dickey-Fuller Tests against Fractional Alternatives" **Economic Letters** , 35, s.155-160, 1991.
- Fama, Eugene F. : "A fractional Dickey-Fuller test for unit roots" , **Econometrica** , 70.5 ,s.1963-2006 , 2002.
- Fama, Eugene F. : "Efficient capital markets: A review of theory and empirical work", **The**

- Fama, Eugene F. : **Journal of Finance** , 25.2 , s.383-417 , 1970.
"The behavior of stock-market prices" , **The Journal of Business** 38.1 , s.34-105 , 1965.
- Fox, Robert ; Taqqu, Murad S. : Fox, Robert, and Murad S. Taqqu. "Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series." **The Annals of Statistics** , 14, s.517-532, 1986.
- Galbraith, J. W. ; Zinde-Walsh, Victoria : "Time domain methods for the estimation of fractionally-integrated time series models", **Forthcoming in Biometrika** , 1997.
- Geweke, John ; Porter-Hudak, Susan : "The estimation and application of long memory time series models." **Journal Of Time Series Analysis** , 4.4 , s. 221-238, 1983.
- Giacomini, Raffaella ; Politis, Dimitris N. ; White, Halbert : "A warp-speed method for conducting Monte Carlo experiments involving bootstrap estimators", **Econometric Theory** , 29.3 s.567-589 , 2013.
- Granger, Clive Wj; Joyeux, Roselyne : "An introduction to long-memory time series models and fractional differencing", **Journal of Time Series Analysis** , 1.1 , s.15-29 , 1980.
- Granger, Clive WJ ; Hyung. Namwon : "Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns." **Journal of empirical finance** 11.3, s.399-421, 2004
- Hosking, Jonathan R.M. ; "Fractional Differencing", **Biometrika**, 68, s.165-176, 1981.
- Kreiss, Jens-Peter : Kreiss, J. P. (1992). **Bootstrap procedures for AR (∞)—processes**. In Bootstrapping and Related Techniques , Springer, Berlin, Heidelberg , 1992.
- Kutlar, Aziz, ; Turgut, Tuba : "Türkiye'deki Başlıca Ekonomi Serilerinin ARFIMA Modelleri İle Tahmini ve Öngörülebilirliği." **Kocaeli Üniversitesi Sosyal Bilimler Dergisi** ,11, s.120-149 , 2006
- Kunsch, Hans R. : "The jackknife and the bootstrap for general stationary observations", **The Annals of Statistics**, s.1217-1241 ,

- Lee, Dongin ; Schmidt, Peter : 1989.
"On the power of the KPSS test of stationarity against fractionally-integrated alternatives" , **Journal of Econometrics** , 73.1 , s.285-302, 1996.
- Lütkepohl, Helmut; Kratzig, Markus; Peter CB, Phillips : **Applied time series econometrics**, U.S.A. , Cambridge university press, 2004.
- Mayoral, Laura : "A new minimum distance estimation procedure of ARFIMA processes" , **Department of Economics and Business Universitat Pompeu Fabra** , 2004.
- Maddala, G.S. ; Kim, In-Moo: **Unit Roots, Cointegration, and Structural Change**, Cambridge, Cambridge University Press, 1998.
- Montgomery, Douglas C.; Runger, George C. **Applied statistics and probability for engineers**, John Wiley & Sons, 2007.
- Nath, Golaka C. : "Long memory and Indian stock market—an empirical evidence" , **National Stock Exchange of India Ltd. , Mumbai** , 2001.
- Robinson, Peter M. : "Semiparametric analysis of long-memory time series" , **The Annals of Statistics**, 22.1 , s.515-539 , 1992.
- Robinson, Peter M. : "Efficient tests of nonstationary hypotheses" , **Journal of the American Statistical Association** , 89.428 , s.1420-1437 , 1994.
- Samuelson, Paul A. : "Proof that properly anticipated prices fluctuate randomly" , **The World Scientific Handbook of Futures Markets**, s.25-38 , 2016
- Schwarz, Gideon : "Estimating the dimension of a model" , **The Annals of Statistics**, 6.2 , s.461-464 , 1978.
- Sharma, J. L.; Kennedy, Robert E. : "A comparative analysis of stock price behavior on the Bombay, London, and New York stock exchanges" , **Journal of Financial and Quantitative Analysis**, 12.3 s.391-413 , 1977.
- Shimotsu, Katsumi ; Peter CB Phillips : "Exact local Whittle estimation of fractional integration" , **The Annals of Statistics**, 33.4 , s.1890-1933, 2005.
- Smeeke, Stephan : "Bootstrapping nonstationary time series" Maastricht University, Doktora Tezi, Hollanda , 2009.

- Sowell, Fallaw : "Maximum likelihood estimation of stationary univariate fractionally integrated time series models", **Journal of Econometrics**, 53.1-3 ,s. 165-188, 1992.
- Tanaka, Katsuto : "The nonstationary fractional unit root", **Econometric Theory**, 15.4 , s.549-582, 1999.
- Worthington, Andrew C. ; Helen Higgs : "Art as an investment: Risk, return and portfolio diversification in major painting markets.", **Accounting & Finance** , 44.2 s.257-271 , 2004.
- Velasco, Carlos; Robinson, Peter M. : "Whittle pseudo-maximum likelihood estimation for nonstationary time series", **Journal of the American Statistical Association** , 95.452, s.1229-1243 , 2000.

ÖZGEÇMİŞ

Soyadı, adı :MUTLUAY Hamza
Uyruğu : T.C.
Doğum tarihi ve yeri :1990 - Sakarya
Medeni hali : Bekar
e-mail : hamzanutluay@gmail.com

Eğitim

Derece	Eğitim Birimi	Mezuniyet tarihi
Yüksek lisans	İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, Ekonometri ABD	2019
Lisans	Gazi Üniversitesi İktisadi ve İdari Bilimler Fakültesi Ekonometri Bölümü	2015

Mesleki Deneyim

- Araştırma Görevlisi
BETAM (Bahçeşehir Üniversitesi Ekonomik Ve Toplumsal Araştırmalar
Merkezi)
Haziran 2019-Devem ediyor
- 107K042 Kodlu TÜBİTAK Projesi'nde Bursiyer Asistan
Proje Yöneticisi: Prof. Dr. Murat G. Kırdar
Aralık 2018-Mayıs 2019

Yabancı Dil

İngilizce