

ANKARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

BAĞIMLI GÖZLEMLERLE BOOTSTRAP YÖNTEMİ

Bengül ARKANT

İSTATİSTİK ANABİLİM DALI

ANKARA

2008

Her hakkı saklıdır

Yrd. Doç. Dr. İhsan KARABULUT' un danışmanlığında, Bengül ARKANT tarafından hazırlanan bu çalışma 31/07/2008 tarihinde aşağıdaki jüri tarafından oy birliği ile Ankara Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı' nda Yüksek Lisans Tezi olarak kabul edilmiştir.

Başkan: Yrd. Doç. Dr. Bülent ALTUNKAYNAK

Gazi Üniversitesi İstatistik Anabilim Dalı

Üye : Prof. Dr. Yılmaz AKDİ

Ankara Üniversitesi İstatistik Anabilim Dalı

Üye : Yrd. Doç. Dr. İhsan KARABULUT

Ankara Üniversitesi İstatistik Anabilim Dalı

Yukarıdaki sonucu onaylarım.

Prof.Dr.Orhan ATAKOL

Enstitü Müdürü

ÖZET

Yüksek Lisans Tezi

BAĞIMLI GÖZLEMLERLE BOOTSTRAP YÖNTEMİ

Bengül ARKANT

Ankara Üniversitesi

Fen Bilimleri Enstitüsü

İstatistik Anabilim Dalı

Danışman: Yrd. Doç. Dr. İhsan KARABULUT

İstatistiksel çıkarımın önemli noktalarından birisi yığından alınan örneklemin yığını temsil etmesidir. Bootstrap, bu düşünceyi temel alarak geliştirilen bir yeniden örnekleme yöntemidir. Efron (1979)' un bootstrap yöntemi gözlemler birbirinden bağımsız ve aynı dağılımlı (b.b.a.d) olduğunda pek çok durumda kullanılabilen bir yeniden örnekleme yöntemidir. Gözlemlerin birbirinden bağımsız olmadığı durumlarda, bu yöntem gözlemlerin alındığı yığındaki bağımlılığı doğası gereği gözardı etmektedir. Gözlemlerin alındığı yığındaki bağımlılık yapısını yansıtabilecek yeniden örnekleme yöntemlerinden birisi de Blok Bootstrap yöntemleridir. Bu çalışmada, bu yöntemler tanıtılacak, Markov zincirlerinde yöntemlerin nasıl işlediği bir örnekle verilecektir. Ayrıca durağan otoregresif zaman serilerinde bootstrap yöntemi bir örnekle verilecektir.

Temmuz 2008, 76 sayfa

Anahtar Kelimeler: m-bağımlı, durağan, α -miksing, Hareketli Bloklarla Bootstrap, Örtüşmeyen Bloklarla Bootstrap, Çembersel Bloklarla Bootstrap, zaman serileri

ABSTRACT

Master Thesis

BOOTSTRAP METHOD WITH DEPENDENT OBSERVATIONS

Bengül ARKANT

Ankara University

Graduate School of Natural and Applied Sciences

Department of Statistics

Supervisor: Asst. Prof. Dr. İhsan KARABULUT

One of the important point of statistical inference is the representation ability of the sample of the population well enough. Bootstrap is a resampling method which depends on this idea. Efron (1979)' s bootstrap method works most of the situations when the observations are independent and identically distributed (iid). When the observations are dependent this method neglects the dependence structure of the population. There have been various works on the bootstrap method when the observations are dependent. One of the resampling methods which can reflect the dependence structure of the population is the Block Bootstrap methods. In this study these methods is going to be presented. Examples for the application of these methods are going to be given for the Markov Processes. And bootstrap method for stationary autoregressive time series is going to be given with an example.

July 2008, 76 pages

Key Words: m-dependent, stationary, α -mixing, Moving Block Bootstrap, Nonoverlapping Block Bootstrap, Circular Block Bootstrap, Time Series

TEŐEKKÜR

BaŐta alıŐmalarımı ynlemdirren, bana duyduėu gven ile alıŐmanın her safhasında bana destek olan, bilgi, neri ve yardımlarını esirgemeyerek akademik ortamda yetiŐme ve geliŐmeme katkıda bulunan danıŐman hocam, Sayın Yrd. Do. Dr. İhsan KARABULUT (Ankara niversitesi Fen Fakltesi İstatistik Blm)' a ve hibir zaman benden desteėini esirgemeyen, her zaman yanımda olan sevgili aileme teŐekkrlerimi sunarım.

Bengl ARKANT

Ankara, Temmuz 2008

İÇİNDEKİLER

ÖZET.....	i
ABSTRACT.....	ii
TEŞEKKÜR.....	iii
SİMGELER DİZİNİ.....	v
ŞEKİLLER DİZİNİ.....	vi
ÇİZELGELER DİZİNİ.....	vii
1. GİRİŞ.....	1
2. ÖNEMLİ TANIM ve KAVRAMLAR.....	4
3. BAĞIMLI GÖZLEMLERLE BOOTSTRAP YÖNTEMİ.....	15
3.1 Hareketli Bloklarla Bootstrap Yöntemi (HBB).....	19
3.2 Örtüşmeyen Bloklarla Bootstrap Yöntemi (ÖBB).....	25
3.3 Genellenmiş Blok Bootstrap Yöntemi (GBB).....	30
3.3.1 Çembersel Bloklarla bootstrap yöntemi (ÇBB).....	32
3.4 HBB, ÖBB ve ÇBB Yöntemleriyle Örneklem Varyans ve Dağılım Fonksiyonu Tahmin Edicilerinin Tutarlılığı.....	34
3.5 Varyans ve Yanlılık Bootstrap Tahmin Edicileri için Blok Bootstrap Yöntemlerinin Karşılaştırılması.....	38
3.6 Durağan Otoregresif Zaman Dizilerinde Bootstrap Yöntemi.....	44
4. UYGULAMA.....	47
4.1 Sonlu Durum Uzayına Sahip Bir Markov Zincirinin Geçiş Olasılıkları Matrisinin Blok Bootstrap Yöntemleriyle Tahmini.....	47
4.2 Birinci Dereceden Otoregresif Zaman Dizisinin α_1 Parametresinin Bootstrap Tahmini.....	54
5. TARTIŞMA ve SONUÇ.....	58
KAYNAKLAR.....	60
EKLER.....	63
EK 1 Bölüm 4.1 İçin Yazılan Matlab Bilgisayar Programları Kodları.....	64
EK 2 Bölüm 4.2 için Yazılan Matlab Bilgisayar Programı Kodları.....	72
EK 3 $X_i = 0.5X_{i-1} + \varepsilon_i, i = 2, \dots, 250$ zaman dizisi modeline uygun üretilen veriler.....	74
ÖZGEÇMİŞ.....	76

SİMGELER DİZİNİ

\mathbb{N}	Doğal sayılar
\mathbb{Z}	Tam sayılar
\mathbb{R}	Reel sayılar
$(X_n)_{n=1}^{\infty}$	Rasgele değişkenler dizisi
$\alpha(\cdot)$	Güçlü miksing katsayısı
p_{ij}	i. durumdan j. duruma geçiş olasılığı
\log	Doğal logaritma
\sup	Supremum
\xrightarrow{d}	Dağılımda yakınsama
\xrightarrow{p}	Olasılıkta yakınsama
F_n	X_1, X_2, \dots, X_n örnekleminin gözleme dayalı dağılımı
$F_{m,n}^{*(j)}$	$X_{j,1}^*, X_{j,2}^*, \dots, X_{j,m}^*$ bootstrap örnekleminin gözleme dayalı dağılımı
$F_{p,n}$	p -boyutlu gözleme dayalı dağılım
b.b.a.d	Birbirinden bağımsız aynı dağılımlı
\otimes	Kartezyen çarpım
$\llbracket \cdot \rrbracket$	Tam değer fonksiyonu
\equiv	Denk
HKO	Hata kareler ortalaması
Yan	Yanlılık

ŞEKİLLER DİZİNİ

Şekil 3.1 HBB için $\{B_1^{(1)}, B_2^{(1)}, \dots, B_N^{(1)}\}$ bloklarının elde edilişi.....	22
Şekil 3.2 ÖBB için $\{B_1^{(2)}, B_2^{(2)}, \dots, B_b^{(2)}\}$ bloklarının oluşturulması.....	27
Şekil 3.3 $Y_{n,i}$ değişkenlerinin oluşturulması.....	31
Şekil 4.1 Üretilen veriler için zaman serisi grafiği.....	55

ÇİZELGELER DİZİNİ

Çizelge 4.1 Markov zincirinin $n = 250$ adım sayısına sahip bir gerçekleşmesi.....	48
Çizelge 4.2 $l = 2, 5, 10, 25, 50$ blok uzunlukları için HBB, ÖBB ve ÇBB yöntemleri için oluşturulan blok sayıları ve seçilen blok sayıları.....	52
Çizelge 4.3 $l = 2, 5, 10, 25, 50$ blok uzunlukları için geçiş olasılıkları matrisinin en HBB, ÖBB ve ÇBB yöntemleri ile elde edilen bootstrap tahminleri....	53
Çizelge 4.4 $l = 2, 5, 10, 25, 50$ blok uzunlukları için HBB, ÖBB, ÇBB tahminleri...	57

1. GİRİŞ

İstatistiksel çıkarım yapmak amacıyla yığından alınan örneklemin yığını temsil etmesi (yığına ait dağılımı taklit ettiği) istatistiksel çıkarımın önemli dayanaklarından birisidir. Bu düşünceden hareketle yığından alınan örneklemin kendisi de yığındaki rasgeleliğin işleyişini kopya edecektir. Bootstrap bu düşünceleri temel alarak geliştirilen bir yeniden örnekleme tekniğidir. Bootstrapın işlevi ve istatistiksel çıkarımda kullanımı için Efron (1979), Hall (1992)' in çalışmaları önemli kaynaklar olarak gösterilebilir.

İlk tanımlanış haliyle bootstrap X_1, X_2, \dots, X_n örnekleminin birbirinden bağımsız ve aynı F dağılımından geldiği varsayımına dayanır ve yeniden örnekleme ile oluşturulan F_n gözleme dayalı dağılım fonksiyonu, $\frac{1}{n}$ olasılıkla yalnızca gözlenen değerlerin gözlenebildiği bir yığınun dağılım fonksiyonu gibi düşünülerek n çaplı $X_1^*, X_2^*, \dots, X_n^*$ bootstrap örneklemini her bir X_i^* rasgele değişkeni aynı F_n dağılımlı yığından birbirinden bağımsız olarak (yerine konularak) oluşturulur. Bu X_1, X_2, \dots, X_n ' nin F dağılımlı yığından örneklem alma işleminin taklit edilerek $X_1^*, X_2^*, \dots, X_n^*$ ' in F_n dağılımlı yığından elde edilmesidir. için Böyle bir yığından n^n tane değişik örneklem elde edilebileceği açıktır. Ele alınan bir istatistik n^n tane örneklemden hesaplanabilir ve dağılımı tanımlanabilir ki buna da bootstrap dağılımı denir. Pratikte bootstrap dağılımı da oluşturulamaz ancak n^n tane örneklem içinden B tanesi alınır ve bu işleme Bootstrap Monte Carlo' su denir. Yukarıdaki haliyle bootstrap birbirinden bağımsız aynı dağılımlı olma kurgusuna dayanır. Ancak üzerinden örneklem alınan yığın birbirinden bağımsız aynı dağılımlı örneklem alma kurgusu bozulduğunda bootstrap işleminin istatistiksel çıkarımda bu yöntemden beklenen tutarlı yaklaşımı sağlayamayacağı ilk olarak Singh (1981) tarafından saptanmıştır. Saptama (Remark 2.1) 2.1' de X_1, X_2, \dots, X_n örneklemini m-bağımlı olduğunda $X_1^*, X_2^*, \dots, X_n^*$ yukarıdaki haliyle bootstrap örneklemini, X_1, X_2, \dots, X_n koşulu altında birbirinden bağımsızdır. Yani,

$$P(X_1^* = x_{i_1}^*, X_2^* = x_{i_2}^*, \dots, X_n^* = x_{i_n}^* | X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) \\ = P(X_1^* = x_{i_1}^* | X_1 = x_1, X_2 = x_2, X_n = x_n) \dots P(X_n^* = x_{i_n}^* | X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$$

ve her bir $P(X_i^* = X_i) = \frac{1}{n}, i = 1, 2, \dots, n$ dir. Bu durumda \bar{X}_n^* bir bootstrap örnekleminde elde edilen örneklem ortalamasını, \bar{X}_n yığından alınan X_1, X_2, \dots, X_n örnekleminde ait örneklem ortalamasını göstermek üzere $n \rightarrow \infty$ iken Merkezi Limit Teoremi koşulları altında $\sqrt{n}(\bar{X}_n - \mu)$ rasgele değişkeninin bootstrap tahmin edicisi için,

$$\sqrt{n}(\bar{X}^* - \bar{X}_n) \xrightarrow{d} N(0, \sigma^2)$$

olacaktır. Halbuki m-bağımlılık yapısına sahip bir X_1, X_2, \dots, X_n örnekleminde için $n \rightarrow \infty$ iken gerekli koşullar sağlanması halinde Merkezi Limit Teoremi,

$$\sqrt{n}(\bar{X}^* - \bar{X}_n) \xrightarrow{d} N(0, \sigma^2 + 2 \sum_{i=1}^m Cov(X_1, X_{1+i}))$$

olduğunu ortaya koyar. $\sqrt{n}(\bar{X}^* - \bar{X}_n)$ için bağımlılık yapısı altında yine Merkezi Limit Teoremi'ne uygun yakınsama gerçekleşir fakat bu yakınsama farklı varyansla olacaktır. m- bağımlılık, ileride tanımlanacak olan zayıf bağımlılığın bir türüdür.

$(X_n)_{n=1}^{\infty}$ zayıf bağımlı ve durağan dağılımlı rasgele değişkenler dizisi olsun. Ayrıca X_1, X_2, \dots, X_n alınan örnekleminde gösterebiliriz. Bu örnekleme dayalı olarak tanımlanan bir T_n istatistiğinin bootstrap değerlendirmesi ile bu örneklem dağılımının bootstrap tahmini için önerilen bootstrap yöntemlerinden bloklar yönteminin tanıtımı çalışmanın konusu olacaktır.

Aşağıdaki ilk bölümde bazı tanım ve kavramlar tanıtılacak, izleyen bölümde ise bağımlı gözlemlerle bootstrap yönteminin tutarlılığına ilişkin sonuçlar aktarılacaktır.

Uygulama bölümünde ilk olarak sonlu durum uzayına sahip bir markov zincirinden üretilmiş veriler kullanılarak sonlu durum uzayına sahip bir markov zincirinin geçiş matrisi tahmininin blok bootstrap yöntemleriyle nasıl yapılabileceğine bir örnek verilecektir. İkinci uygulama ise durağan birinci dereceden otoregresif zaman serisi modeline uygun olarak üretilmiş veriler kullanılarak model parametresinin bootstrap tahmininin elde edilmesidir.

2. ÖNEMLİ TANIM VE KAVRAMLAR

İstatistik çıkarım ve yöntemlerinde genel olarak yapılan bağımsızlık varsayımı değişik biçimlerde gevşetilerek çeşitli özellik ve yapılaraya sahip bağımlı gözlemlerle de istatistiksel çıkarım yapılmaya çalışılır. Markov zincirleri ile bir sürecin modellenmesi ve zaman serilerinin analizi, bağımlılık yapılarının tanımlandığı uygulamalardır.

Tanım 2.1 $(X_n)_{n \in \mathbb{Z}}$ rasgele değişkenlerin bir dizisi olsun. Her $k \in \mathbb{N}$ ve $m \in \mathbb{Z}$ olmak üzere eğer her $i_1 < i_2 < \dots < i_k$ için $(X_{i_1}, \dots, X_{i_k})'$ rasgele vektörü ile $(X_{i_1+m}, \dots, X_{i_k+m})'$ rasgele vektörleri aynı ortak dağılımlara sahip iseler $(X_n)_{n \in \mathbb{Z}}$ “güçlü durağan” dır denir. Eğer $(X_n)_{n \in \mathbb{Z}}$ rasgele değişkenler dizisinin beklenen değeri ve varyansı n ’ den bağımsız, X_s ile X_t arasındaki kovaryans sadece $|t-s|$ ’ nin bir fonksiyonu ise $(X_n)_{n \in \mathbb{Z}}$ “zayıf durağan” dır denir (Akdi 2003).

Çalışmanın bundan sonraki kısmında, “zayıf durağan” diye belirtilmedikçe durağan terimi “güçlü durağan” anlamında kullanılmıştır.

Durağan rasgele değişken dizilerinin tipik bir örneği m-bağımlı dizilerdir. Eğer $(X_n)_{n \in \mathbb{Z}}$ rasgele değişken dizisi, her $k \in \mathbb{Z}$ için $m \in \mathbb{N}$ olmak üzere $(X_i)_{i \leq k}$ rasgele değişken dizisi ile $(X_i)_{i > k+m}$ rasgele değişken dizisi birbirlerinden bağımsız iseler $(X_n)_{n \in \mathbb{Z}}$ “m-bağımlı” dır denilir.

Örneğin, Y_1, Y_2, \dots birbirinden bağımsız aynı dağılımlı rasgele değişkenler olsun. f \mathbb{R}^{m+1} , de reel değerli bir fonksiyon olmak üzere $X_i = f(Y_i, \dots, Y_{i+m})$ olarak tanımlansın. Rasgele değişkenlerin dizisi $(X_n)_{n=1}^{\infty}$ hem durağan hem de m-bağımlıdır (Billingsley 1994).

Bağımsızlık varsayımının gevşetilmesi, yukarıdaki gibi durağanlık ve m-bağımlılık gibi kolayca tanımlanıp uygulanabilen bağımlılık yapıları ile söz konusudur. Bazı durumlarda da bağımlılık yapısı tanımlanmadan bağımsızlığın gevşetilmesi olanaklıdır. Bu yollardan birisi rasgele değişken dizisinin α -miksing (α -mixing) olarak tanımlanmasıyla sağlanır.

Tanım 2.2 X_1, X_2, \dots aynı olasılık uzayında tanımlı rasgele değişkenlerin bir dizisi ve $k \geq 1, n \geq 1$ tamsayıları için, $\sigma(X_1, X_2, \dots, X_k)$ X_1, X_2, \dots, X_k rasgele değişkenlerinin ürettiği σ -cebiri, $\sigma(X_{k+n}, X_{k+n+1}, \dots)$ de $X_{k+n}, X_{k+n+1}, \dots$ rasgele değişkenlerinin ürettiği σ -cebiri olsun. $A \in \sigma(X_1, X_2, \dots, X_k)$ ve $B \in \sigma(X_{k+n}, X_{k+n+1}, \dots)$ herhangi iki olay için

$$|P(A \cap B) - P(A)P(B)| \leq \alpha(n) \quad (2.1)$$

olacak biçimde bir $\alpha(n)$ sayısı var ve $\lim_{n \rightarrow \infty} \alpha(n) = 0$ oluyorsa, $(X_n)_{n=1}^{\infty}$ rasgele değişkenler dizisi α -miksing' dir denilir (Billingsley 1994).

Bu tanımlama ile $(X_n)_{n=1}^{\infty}$ rasgele değişken dizisinin bağımlılık yapısını tanımlasak da $n \rightarrow \infty$ iken X_1, X_2, \dots, X_n rasgele değişkenlerinden tanımlayacağımız A olayları ile $X_{k+n}, X_{k+n+1}, \dots$ rasgele değişkenlerinden tanımlayacağımız B olaylarının yaklaşık olarak birbirinden bağımsız olacaklarını iddia edebiliriz.

Örnek 2.1 $(X_n)_{n=1}^{\infty}$ bir rasgele değişkenler dizisi olsun. Her $n > m$ için (X_1, X_2, \dots, X_k) ile $(X_{k+n}, X_{k+n+1}, \dots, X_{k+n+l})$ bağımsız iseler doğaldır ki $A \in \sigma(X_1, X_2, \dots, X_k)$ ve $B \in \sigma(X_{k+n}, X_{k+n+1}, \dots, X_{k+n+l})$ için,

$$|P(A \cap B) - P(A)P(B)| = 0$$

olacağından $n > m$ için $\alpha(n) = 0$ olmuş olacaktır. Böylelikle m-bağımlı bir rasgele değişkenler dizisi aynı zamanda α -miksing olacaktır (Billingsley 1994).

Örnek 2.2 (Markov Zinciri) (Y_n) sonlu durum uzayına sahip bir markov zinciri olsun. Geçiş olasılıkları matrisi P ve P ' nin elemanları p_{ij} ' ler sıfırdan büyük, f durum uzayı üzerinde tanımlı reel değerli bir fonksiyon ve $X_n = f(Y_n)$ olsun. Başlangıç olasılıkları p_i ' ler durağan olsunlar. Bu, i, j ' ler durum uzayında yer alan durumlar olmak üzere herhangi bir j -inci durum için bütün i ' ler üzerinden

$$\sum_i p_i p_{ij} = p_j$$

olmasıdır. Bu durumda (X_n) markov zinciri durağan olacaktır. Markov zinciri durağan ve sonlu durum uzayına sahip indirgenemez olup periyodik değilse i durumundan j durumuna n adımda geçiş olasılığı p_{ij}^n ile gösterilmek üzere $A \geq 0$ ve $0 \leq \rho < 1$ için,

$$|p_{ij}^{(n)} - p_j| \leq A\rho^n$$

dir (Billingsley 1994, Theorem 8.9).

Bu durumda

$$P(Y_1 = i_1, \dots, Y_k = i_k, Y_{k+n} = j_0, \dots, Y_{k+n+l} = j_l) = p_{i_1} p_{i_1 i_2} \dots p_{i_{k-1} i_k} p_{i_k j_0}^{(n)} p_{j_0 j_1} \dots p_{j_{l-1} j_l}$$

ile

$$P(Y_1 = i_1, \dots, Y_k = i_k).P(Y_{k+n} = j_0, \dots, Y_{k+n+l} = j_l) = p_{i_1} p_{i_1 i_2} \dots p_{i_{k-1} i_k} p_{j_0} p_{j_0 j_1} \dots p_{j_{l-1} j_l}$$

olasılıkları arasındaki fark en fazla $\rho^n p_{i_1} p_{i_2} \dots p_{i_{k-1} i_k} p_{j_0 j_1} \dots p_{j_{l-1} j_l}$ kadar olacaktır. Ek olarak eğer s durumların sayısı ise $A = [(Y_1, Y_2, \dots, Y_k) \in H]$ ve $B = [(Y_{k+n}, \dots, Y_{k+n+l}) \in H']$ kümeleri için (2.1) eşitsizliği $\alpha(n) = s\rho^n$ olmasıyla sağlanır. Bu kümeler k, n sabit olmak üzere sırasıyla $\sigma(X_1, \dots, X_k)$ ve $\sigma(X_{k+n}, \dots, X_{k+n+l})$ σ -cebiri üretirler (Billingsley 1994).

Rasgele değişken dizisinin durağan olmasının yanında α -miksing olması halinde $S_n = X_1 + X_2 + \dots + X_n$ için $n \rightarrow \infty$ iken $\frac{S_n}{\sigma\sqrt{n}} \xrightarrow{d} N(0,1)$ olduğu gösterilebilir. Çünkü bu varsayımlarla ve bazı ek koşullarla

$$\frac{Var(S_n)}{n} \rightarrow \sigma^2 = E(X_1^2) + 2 \sum_{k=1}^{\infty} E(X_1 X_{1+k})$$

olacak şekilde σ^2 'nin mutlak yakınsak olduğu gösterilebilir. Rasgele değişken dizisinin α -miksing olması, $E(X_1 X_{1+k})$ 'lerin sınırlandırılmasını ve söz konusu serinin mutlak olarak yakınsamasını sağlamaktadır. Bu nedenle $(X_n)_{n=1}^{\infty}$ rasgele değişkenler dizisinin zayıf bağımlılığının bir ifadesi olan α -miksing yerine zayıf bağımlılık aşağıdaki gibi de ifade edilebilir.

Tanım 2.3 \mathcal{F} reel değerli fonksiyonların bir sınıfı, $\theta = (\theta_r)_{r \in \mathbb{N}}$ $r \rightarrow \infty$ iken sıfıra doğru azalan bir dizi ve ψ , değişkenleri $(f, g, h, k) \in \mathcal{F}^2 \times \mathbb{N}^2$ olan bir fonksiyon olsun. $i_1 \leq \dots \leq i_h < i_h + r \leq j_1 \leq \dots \leq j_k$ için (i_1, \dots, i_h) h -boyutlu bir vektör ve (j_1, \dots, j_k) k -boyutlu bir vektör olmak üzere

$$|Cov(f(X_{i_1}, \dots, X_{i_h}), g(X_{j_1}, \dots, X_{j_k}))| \leq \psi(f, g, h, k) \theta_r$$

ise $(X_n)_{n \in \mathbb{N}}$ rasgele deęişkenler dizisi zayıf baęımlıdır (weak dependent) denir (Doukhan 1999).

α -miksing olma $(X_n)_{n=1}^\infty$ rasgele deęişkenler dizisinin zayıf baęımlı olmasını saęlayan bir özellik olarak karřımıza çıkmaktadır.

$(\varepsilon_t)_{t \in \mathbb{Z}}$ birbirinden baęımsız aynı daęılımlı rasgele deęişkenler dizisi olmak üzere $(X_t)_{t \in \mathbb{Z}}$,

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p} + \varepsilon_t \quad (2.2)$$

duraęan p . dereceden otoregresif zaman serisi (AR(p)) modeline uygun rasgele deęişkenler dizisi biçiminde tanımlansın. Bu durumda (2.2) eřitlięi $\sum_{j=0}^\infty |w_j| < \infty$ olmak üzere,

$$X_t = \sum_{j=0}^\infty w_j \varepsilon_{t-j} \quad (2.3)$$

řeklinde yazılabilir. Bu süreç α -miksingdir (Chanda 1974, Whithers 1981, Athreya and Pantula 1986).

Teorem 2.1 (Chanda 1974) $(\varepsilon_t)_{t \in \mathbb{Z}}$, $\delta > 0$ olmak üzere $E(|\varepsilon_1|^\delta) < \infty$ olan bir beyaz gürültü (white noise) dizisi olsun ve ϕ_0, ε_1 ' in integrallenebilen karakteristik fonksiyonu, $(2\pi)^{-1} \int_{-\infty}^\infty |\phi_0(u)| du \leq 1$ özellięini saęlasın. $(X_t)_{t \in \mathbb{Z}}$ rasgele deęişkenler dizisi

de (2.3) eřitlięindeki gibi tanımlanan bir dizi olsun. Eęer $\lambda = \frac{\delta}{1+\delta}$ için $\sum_{j=0}^\infty j |w_j|^\lambda < \infty$

ise $(X_t)_{t \in \mathbb{Z}}$ rasgele deęişkenler dizisi α -miksindir, yani M sadece ϕ_0 ' a baęlı pozitif bir sabit, $A \in \sigma(X_t : -\infty < t < 0)$ ve $B \in \sigma(X_t : k < t < \infty)$ için

$$|P(A \cap B) - P(A)P(B)| < M\alpha(k)$$

ve

$$\alpha(k) = \sum_{j=k}^{\infty} j |w_j|^\lambda$$

dır.

Örneęin, $(X_t)_{t \in \mathbb{Z}}$, $X_t = \alpha_1 X_{t-1} + \varepsilon_t$, $|\alpha_1| < 1$ zaman serisi modeline uygun rasgele deęişkenler dizisi olsun. Bu durumda, $w_j = \alpha_1^j$ için

$$\begin{aligned} \alpha(k) &= \sum_{j=k}^{\infty} j |w_j|^\lambda \\ &= \sum_{j=k}^{\infty} j (\alpha_1^j)^\lambda \\ &= \sum_{j=k}^{\infty} j (\alpha_1^\lambda)^j \\ &= k (\alpha_1^\lambda)^k \left[\frac{1}{(1 - \alpha_1^\lambda)} + \frac{1}{[k(1 - \alpha_1^\lambda)^2]} \alpha_1^\lambda \right] \end{aligned}$$

olacaktır.

$$\begin{aligned}\lim_{k \rightarrow \infty} \alpha(k) &= \lim_{k \rightarrow \infty} \frac{k(\alpha_1^\lambda)^k}{(1-\alpha_1^\lambda)} + \lim_{k \rightarrow \infty} \frac{(\alpha_1^\lambda)^{k+1}}{(1-\alpha_1^\lambda)^2} \\ &= \frac{1}{(1-\alpha_1^\lambda)} \lim_{k \rightarrow \infty} k(\alpha_1^\lambda)^k\end{aligned}$$

dır.

$$\begin{aligned}\frac{\partial}{\partial(\alpha_1^\lambda)} \lim_{k \rightarrow \infty} (\alpha_1^\lambda)^{k+1} &= \lim_{k \rightarrow \infty} \frac{\partial}{\partial(\alpha_1^\lambda)} (\alpha_1^\lambda)^{k+1} \\ &= \lim_{k \rightarrow \infty} (k+1)(\alpha_1^\lambda)^k \\ &= \lim_{k \rightarrow \infty} k(\alpha_1^\lambda)^k + \lim_{k \rightarrow \infty} (\alpha_1^\lambda)^k\end{aligned}$$

$\frac{\partial}{\partial(\alpha_1^\lambda)} \lim_{k \rightarrow \infty} (\alpha_1^\lambda)^{k+1} = 0$ ve $\lim_{k \rightarrow \infty} (\alpha_1^\lambda)^k = 0$ olduğundan

$$\lim_{k \rightarrow \infty} k(\alpha_1^\lambda)^k = 0$$

dır. Dolayısıyla $\lim_{k \rightarrow \infty} \alpha(k) = 0$ dır. O halde $(X_t)_{t \in \mathbb{Z}}$ zayıf durağan AR(1) serisi α -miksingdir.

Fakat M 'nin daha sonra Gorodetskii (1977) tarafından başka parametrelere de bağlı olduğunun gösterildiği, Whithers (1981) tarafından ifade edilmektedir. Whithers (1981) ise (2.3) eşitliğindeki gibi ifade edilen bir dizinin α -miksing olmasını başka koşullara bağlamıştır. Ayrıca dizideki rasgele değişkenlerin aynı dağılımlı olma koşulu da gevşetilmiştir.

Teorem 2.2 (Whithers 1981) (w_j) 'ler $\delta > 0$ için $t \rightarrow \infty$ iken $G_t = S_t(\min(1, \delta))^{\max(1, \delta)} \rightarrow 0$ olacak şekilde tanımlanan kompleks sayılar dizisi olsun. Burada $S_t(\cdot)$,

$$S_t(\delta) = \sum_{j=t}^{\infty} |w_j|^\delta$$

biçiminde tanımlıdır.

$$K = (2\pi)^{-1} \max_j \int_{-\infty}^{\infty} |\phi_j(t)| dt < \infty$$

ve

$$\gamma = \max_j E |\varepsilon_j|^\delta < \infty$$

için $(\varepsilon_j)_{j \in \mathbb{Z}}$ birbirinden bağımsız, karakteristik fonksiyonlar dizisi (ϕ_j) olan rasgele değişkenler dizisi olsun.

Bütün t ' ler için $X_{nt} = \sum_{j=0}^n w_j \varepsilon_{t-j}$, $n \rightarrow \infty$ iken olasılıkla X_t rasgele değişkenine yakınsar. Buna göre,

$$W_t = \sum_{j=0}^{t-1} w_j \varepsilon_{t-j}, \quad k \leq t \leq k+m-1$$

ve

$$\chi(v) = P\left(\bigcup_{j=1}^s \alpha_{jt} - v_t < W_t < \beta_{jt} - v_t, k \leq t \leq k+m-1\right)$$

olmak üzere

$$M_0 = \sup_{m,s,k \geq 1} \sup_{\alpha, \beta, \gamma} \max_t \left| \frac{\partial \chi(\mathbf{v})}{\partial v_t} \right| < \infty$$

ise $(X_t)_{t \in \mathbb{Z}}$ rasgele deęişkenler dizisi α -miksingdir ve $\alpha_0(k) = \sum_{t=k}^{\infty} G_t$ için

$$\alpha(k) \leq 2(4M_0 + \gamma)\alpha_0(k)$$

dır.

Bir dięer sonuç Athreya and Pantula (1986) tarafından ortaya konulmuştur. Athreya and Pantula (1986)'nın otoregresif bir zaman serisinin α -miksing olması için gösterdięi gerekli koşullar Teorem 2.3' de verilmiştir.

Teorem 2.3 $(X_t)_{t \in \mathbb{Z}}$, $X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p} + \varepsilon_t$ şeklinde ifade edilen p . dereceden bir otoregresif zaman serisi modeline uygun bir rasgele deęişkenler dizisi olsun.

(i) $\{\log |\varepsilon_1|\}^+$, $\log |\varepsilon_1|$ ifadesinin pozitif kısmını göstermek üzere $E \left[\{\log |\varepsilon_1|\}^+ \right] < \infty$,

(ii) ε_1 'in daęılım fonksiyonu aşikar olmayan (non-trivial) mutlak sürekli bileşene sahip,

(iii) $X_0 = (X_0, X_{-1}, \dots, X_{1-p})$, $(\varepsilon_t)_{t \in \mathbb{Z}}$ 'lerden baęımsız,

(iv) $(\varepsilon_t)_{t \in \mathbb{Z}}$ 'ler birbirinden baęımsız aynı daęılımlı,

(v) $z^p - \alpha_1 z^{p-1} - \dots - \alpha_p = 0$ karakteristik denkleminin kökleri mutlak deęerce 1' den küçük

olma koşulları sağlanıyorsa $(X_t)_{t \in \mathbb{Z}}$ dizisi α -miksingdir (Athreya and Pantula 1986).

Değişik zayıf bağımlılık ve bağımlılık koşulları altında merkezi limit teoremleri ifade edilir.

Teorem 2.4 $(X_i)_{i \in \mathbb{Z}}$ durağan rasgele değişkenler dizisi α -miksing olsun. $\delta \in (0, \infty)$

için $E|X_1|^{2+\delta} < \infty$ ve $\sum_{n=1}^{\infty} \alpha_n^{\delta/2+\delta} < \infty$ koşulları altında

$$\sigma_{\infty}^2 = \text{Var}(X_1) + 2 \sum_{k=1}^{\infty} \text{Cov}(X_1, X_{1+k})$$

olmak üzere $\sigma_{\infty}^2 > 0$ için

$$\frac{1}{\sqrt{n} \sigma_{\infty}} \sum_{i=1}^n (X_i - E(X_1)) \xrightarrow{d} N(0,1)$$

dir (Doukhan 1995).

Örnek 2.4 (Zayıf durağan birinci dereceden (AR(1)) zaman serisi)

$$(X_{t+1} - \mu) = \alpha(X_t - \mu) + \varepsilon_{t+1}, \quad |\alpha| < 1, t = 1, 2, \dots$$

$$\varepsilon_t \sim \text{b.b.a.d } N(0, \sigma_0^2)$$

olsun.

$$\text{Var}(X_{t+1}) = \alpha^2 \text{Var}(X_t) + \sigma_0^2 \tag{2.4}$$

olup söz konusu $(X_t)_{t \in \mathbb{N}}$ rasgele değişkenler dizisi zayıf durağan olduğu için

$$Var(X_{t+1}) = Var(X_t) = \sigma^2, t = 1, 2, \dots$$

dir. (2.4) ifadesinden

$$\sigma^2 = \frac{\sigma_0^2}{1 - \alpha^2}$$

dir. ε_t 'lerin normal dağılımlı olduğu göz önüne alınırsa (2.3) ifadesinden X_t 'lerin de normal dağılımlı olduğu söylenebilir. $h > 0$ için X_t ile X_{t+h} 'nin otokovaryans fonksiyonu,

$$\gamma(h) = Cov(X_t, X_{t+h}) = \alpha^h \sigma^2$$

olmak üzere

$$Var(\sqrt{n}(\bar{X} - \mu)) = \sigma^2 + \frac{2}{n} \sum_{k=1}^{n-1} (n-k)\gamma(k)$$

olup Teorem 2.4 uyarınca

$$\begin{aligned} \sigma_\infty^2 &= \lim_{n \rightarrow \infty} Var(\sqrt{n}(\bar{X} - \mu)) = \sigma^2 + 2 \sum_{k=1}^{\infty} \gamma(k) \\ &= \sigma^2 + 2(\alpha\sigma^2 + \alpha^2\sigma^2 + \dots) \\ &= \sigma^2 \left(1 + 2 \frac{\alpha}{1 - \alpha}\right) \\ &= \sigma^2 \frac{1 + \alpha}{1 - \alpha} \end{aligned}$$

dır (Lehmann 1999).

3. BAĞIMLI GÖZLEMLERLE BOOTSTRAP YÖNTEMİ

Bootstrap değerlendirmelerine yönelik teorik sonuçların birçoğu birbirinden bağımsız ve aynı dağılımlı (b.b.a.d) gözlemler için yapılmıştır. Ancak rasgele değişkenlerin her zaman birbirinden bağımsız olması söz konusu değildir. Bildik bootstrap yöntemi, gözlemler bağımlı olduğunda ele alınan probleme her zaman uygun yaklaşımlarda bulunamayabilir.

$(X_n)_{n=1}^{\infty}$, $m \in \mathbb{N}$ için beklenen değeri $E(X_1) = \mu$ ve $E(X_1^2) < \infty$ olan durağan ve m -bağımlı rasgele değişkenlerin bir dizisi olsun. $m \leq n$ için

$$\begin{aligned} n\text{Var}(\bar{X}_n) &= n^{-1}\text{Var}\left(\sum_{i=1}^n X_i\right) \\ &= n^{-1}\left[\sum_{i=1}^n \text{Var}(X_i) + 2\sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j)\right] \\ &= n^{-1}\left[n\text{Var}(X_1) + 2\sum_{i=1}^{n-1}\sum_{k=1}^{n-i} \text{Cov}(X_i, X_{i+k})\right] \\ &= n^{-1}\left[n\text{Var}(X_1) + 2\sum_{k=1}^m (n-k)\text{Cov}(X_1, X_{1+k})\right] \\ &= \text{Var}(X_1) + 2\sum_{k=1}^m \left(1 - \frac{k}{n}\right)\text{Cov}(X_1, X_{1+k}) \end{aligned}$$

olmak üzere

$$\sigma_m^2 \equiv \lim_{n \rightarrow \infty} n\text{Var}(\bar{X}_n) = \text{Var}(X_1) + 2\sum_{k=1}^m \text{Cov}(X_1, X_{1+k}), \quad \bar{X}_n = n^{-1}\sum_{i=1}^n X_i$$

dır. Eğer, $\sigma_m^2 \in (0, \infty)$ ise m-bağımlı rasgele değişkenler için merkezi limit teoreminden,

$$\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} N(0, \sigma_m^2)$$

dır (Athreya and Lahiri 2006).

$T_n = \sqrt{n}(\bar{X}_n - \mu)$ rasgele değişkeninin dağılımına bildik bootstrap yaklaşımı ele alınsın. X_1, X_2, \dots, X_n örnekleminin çapı ile bu örneklemden yeniden örnekleme ile elde edilen $X_1^*, X_2^*, \dots, X_n^*$ örnekleminin çapı eşit olsun. Bu durumda, T_n rasgele değişkeninin bootstrap uyarlaması $T_{n,n}^*$, $\bar{X}_n^* = n^{-1} \sum_{i=1}^n X_i^*$ olmak üzere

$$T_{n,n}^* = \sqrt{n}(\bar{X}_n^* - \bar{X})$$

olacaktır. $T_{n,n}^*$ ' in dağılımı σ_m^2 den başka bir varyans (σ^2) ile normal dağılıma yakınsayacaktır (Athreya and Lahiri 2006).

Teorem 3.1 (Athreya and Lahiri 2006) $(X_n)_{n=1}^\infty$, $E(X_1) = \mu$ ve $\sigma^2 = Var(X_1) \in (0, \infty)$ olan durağan m-bağımlı rasgele değişkenler dizisi olsun ($m \in \mathbb{Z}_+$). Bu durumda, $n \rightarrow \infty$ iken

$$\sup_x |P_*(T_{n,n}^* \leq x) - \Phi(x/\sigma)| = o(1)$$

dir. Ancak bu yakınsama σ_m^2 ' den farklı bir varyans ile olacaktır.

Bu teoremin koşulları altında eğer $\sigma_m^2 \neq 0$ ve $\sum_{k=1}^m Cov(X_1, X_{k+1}) \neq 0$ ise herhangi bir $x \neq 0$ için,

$$\lim_{n \rightarrow \infty} |P_*(T_{n,n}^* \leq x) - P(T_n \leq x)| = |\Phi(x/\sigma) - \Phi(x/\sigma_m)| \neq 0 \quad (3.1)$$

dır. Bu nedenle, bütün $x \neq 0$ için $P(T_n \leq x)$ 'in bildik bootstrap yöntemine dayanan bootstrap tahmin edicisi $P_*(T_{n,n}^* \leq x)$ hatalıdır, (3.1) eşitliğinde limitin sıfır olmaması $P(T_n \leq x)$ ' in bootstrap tahmin edicisinin tutarlı olmadığını göstermektedir. m-bağımlı X_i ' lerin yeniden örneklenmesinde, bildik bootstrap yöntemi, $(X_n)_{n=1}^{\infty}$ rasgele değişkenler dizisinin bağımlılık yapısını istatistiksel çıkarıma yansıtmaz. Bu sorunun aşılmasına yönelik bootstrap yöntemleri önerilmiştir. Örneğin, markov bağımlılığının görüldüğü sonlu durum uzayına sahip ergodik markov zincirleri için, Basawa *et al* (1990) makalelerinde geçiş olasılıkları matrisinin tahmini problemi için koşullu bootstrap yöntemini kullanmışlardır. Bu yöntemle geçiş sayılarının bilinmesi koşulu altında geçiş olasılıkları matrisinin bootstrap tahminleri ile tutarlı yaklaşımlarda bulunulduğunu göstermişlerdir. Değişik türlerde bağımlılık yapısını bünyesine alan ve bu bağımlılık yapılarını bootstrap tahminlerine yansıtan bir başka yeniden örnekleme biçimi ise, blok bootstrap yöntemleridir.

Blok bootstrap yöntemleri, $(X_i)_{i \in \mathbb{Z}}$ rasgele değişkenler dizisi durağan ve “kısa dönemli” bağımlılık yapısına sahip olduğunda gözlemler dizisi alındığı sıra ile dizideki bağımlılık yapısını yakalayabilen uygun uzunluklu bloklara ayrılırsa bu blokların “yaklaşık olarak” birbirinden bağımsız ve farklı bloklardaki gözlemlerin ortak dağılımının durağanlık koşulu altında aynı olması fikrine dayanır. Bağımlı gözlemlerle bootstrap uygulamalarında farklı blok bootstrap yöntemleri önerilmiştir. Bunlar, Künsch (1989), Liu and Singh (1992) tarafından önerilen Hareketli Bloklarla Bootstrap (HBB),

Carlstein (1986)' nin çalışmalarına dayanan Örtüşmeyen Bloklarla Bootstrap (ÖBB), Politis and Romano (1992) tarafından önerilen Çembersel Bloklarla Bootstrap (ÇBB) yöntemleridir.

Bu bölümde, adı geçen blok bootstrap yöntemleri tanıtıldıktan sonra bu yöntemler kullanılarak elde edilen istatistiğin varyans ve dağılım fonksiyonunun bootstrap tahmin edicilerinin tutarlı tahmin ediciler oldukları konusuna değinilecek ve bootstrap tahmin edicileri için blok bootstrap yöntemlerinin başarımları (yetkinlikleri) hakkında bilgiler verilecektir. Ayrıca otoregresif zaman dizilerinde bootstrap yönteminin nasıl işlediği hakkında bilgi verilecektir.

Bu bölümde kullanılan notasyonlar;

$l, 1 \leq l \leq n$ tamsayı değerli blok uzunluğu

N : HBB yöntemine göre oluşturulan blok sayısı

b : ÖBB yöntemine göre oluşturulan blok sayısı

k : HBB, ÖBB ve ÇBB yöntemlerine göre oluşturulan bloklar arasından rasgele seçilen blok sayısı

$B_i^{(j)}, j = 1, 2$: Sırasıyla HBB ve ÖBB yöntemlerine göre oluşturulan i -inci blok

$B_i^{*(j)}, j = 1, 2$: Sırasıyla HBB ve ÖBB yöntemlerine göre oluşturulan bloklar arasından rasgele seçilen i -inci blok

$B(i, j)$: GBB yöntemine göre oluşturulan j uzunluklu i -inci blok

$X_{j,1}^*, X_{j,2}^*, \dots, X_{j,m}^*, j = 1, 2, 3$: Sırasıyla HBB, ÖBB ve ÇBB yöntemlerine göre

oluşturulan bootstrap örnekleme

m : Bootstrap örnekleminin çapı

$\theta_{m,n}^{*(j)}, j = 1, 2$: $\hat{\theta}_n$ istatistiğinin HBB ve ÖBB yöntemleri ile elde edilen bootstrap uyarlamaları

$F_{m,n}^{*(j)}, j = 1, 2$: Bootstrap örnekleminin gözleme dayalı dağılımı

3.1 Hareketli Bloklarla Bootstrap Yöntemi (HBB)

Bir bootstrap örnekleme alınırken her defasında n kez tek bir gözlemin yeniden örneklenmesi yerine HBB yöntemi, her rasgele çekimde ardışık gözlemlerin bloklarını, herbir blokta orijinal gözlemlerin bağımlılığını muhafaza ederek yeniden örnekleme yapar. Ayrıca blok uzunluğu n ile birlikte sonsuza doğru büyütülürse, HBB bağımlılık yapısını yeniden üreticidir.

Yöntemin işleyişi hakkında bilgi verilmeden önce aşağıdaki gösterimlere değinilecektir.

$(X_i)_{i=1}^{\infty}$ durağan rasgele değişkenler dizisi olsun ve $\mathcal{X}_n = \{X_1, X_2, \dots, X_n\}$ bu diziden alınan bir örnekleme göstereyin. Öyle ki X_1, X_2, \dots, X_n örnekleminin bilinmeyen F gibi bir dağılımdan geldiği düşünölsün. F_n ise, X_1, X_2, \dots, X_n örnekleminin gözleme dayalı dağılımı ve $T(\cdot)$ de reel değeri bir fonksiyon olmak üzere ilgilenilen θ parametresi F 'nin bir fonksiyoneli olan $T(F)$ ile ifade edilir. Diğer taraftan $T(F_n)$ gözleme dayalı dağılım fonksiyonu F_n 'nin bir fonksiyoneli olup θ 'nin bir tahmin edicisi $\hat{\theta}_n$, $T(F_n)$ ile ifade edilir.

$h(x)$ herhangi bir fonksiyon olmak üzere

$$T(F) = \int_{-\infty}^{\infty} h(x) dF(x)$$

ve

$$T(F_n) = \int_{-\infty}^{\infty} h(x) dF_n(x) = \frac{1}{n} \sum_{i=1}^n h(X_i)$$

dir. Örneğin, ilgilenilen θ parametresi kitle ortalaması ise $h(x) = x$ olup

$$\theta = T(F) = \int_{-\infty}^{\infty} x dF(x)$$

biçimindedir. Kitle ortalamasının tahmin edicisi,

$$\hat{\theta}_n = T(F_n) = \int_{-\infty}^{\infty} x dF_n(x) = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}_n$$

dir. İlgilenilen θ parametresinin kitle varyansı (σ^2) olması durumunda ise

$$\sigma^2 = T(F) = \int_{-\infty}^{\infty} [x - \int_{-\infty}^{\infty} x dF(x)]^2 dF(x)$$

ve kitle varyansının tahmin edicisi,

$$\begin{aligned} \hat{\sigma}_n^2 &= T(F_n) = \int_{-\infty}^{\infty} [x - \int_{-\infty}^{\infty} x dF_n(x)]^2 dF_n(x) \\ &= \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2 \end{aligned}$$

dir (Serfling 1980).

Bu durumda,

$$\hat{\theta}_n = T(F_n) \quad (3.2)$$

şeklindeki bir istatistiğin HBB uyarlamasını elde etmek için aşağıdaki adımlar izlenir (Lahiri 2003).

Adım 1. l önceden belirlenen (rasgele olmayan) blok uzunluğu olmak üzere χ_n içinden i -inci blok,

$$B_i^{(1)} = (X_i, \dots, X_{i+l-1}) \quad 1 \leq i \leq N, \quad N = n - l + 1$$

olmak üzere N tane blok oluşturulur (bkz. Şekil 3.1.).

Adım 2. HBB örneklemini oluşturmak için $\{B_1^{(1)}, B_2^{(1)}, \dots, B_N^{(1)}\}$ blokları arasından yerine koyarak rasgele k tane $B_1^{*(1)}, B_2^{*(1)}, \dots, B_k^{*(1)}$ blokları seçilir. Burada,

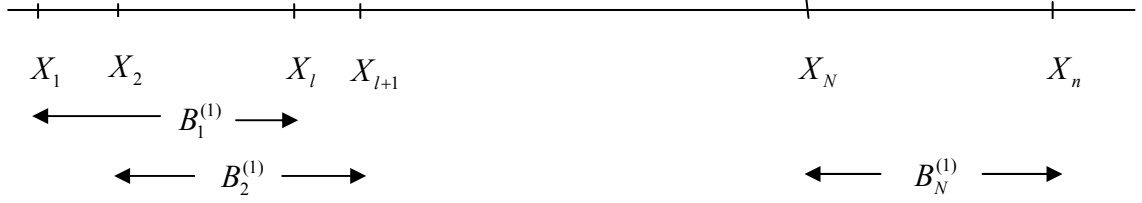
$$B_i^{*(1)} = (X_{1,(i-1)l+1}^*, \dots, X_{1,il}^*), \quad i = 1, \dots, k$$

dır. Sonra $m = k.l$ gözlemin bulunduğu m çaplı $X_{1,1}^*, \dots, X_{1,m}^*$ HBB örneklemini oluşturulur.

Adım 3. $F_{m,n}^{*(1)}, (X_{1,1}^*, \dots, X_{1,m}^*)$ m çaplı bootstrap örnekleminin gözleme dayalı dağılımı olmak üzere, $\hat{\theta}_n$ istatistiğinin bootstrap uyarlaması

$$\theta_{m,n}^{*(1)} = T(F_{m,n}^{*(1)})$$

hesaplanır.



Şekil 3.1 HBB için $\{B_1^{(1)}, B_2^{(1)}, \dots, B_N^{(1)}\}$ bloklarının elde edilişi (Lahiri 2003)

$\{B_1^{(1)}, B_2^{(1)}, \dots, B_N^{(1)}\}$ blokları arasından $B_i^{*(1)}$ bloklarının rasgele seçilmesi, $\{1, \dots, N\}$ kümesinden rasgele k tane indisin seçilmesine denktir. Buna göre, I_1, I_2, \dots, I_k birbirinden bağımsız ve olasılık dağılımı $P(I_i = j) = \frac{1}{N}, j = 1, \dots, N, i = 1, \dots, k$ olan kesikli düzgün dağılıma sahip rasgele değişkenlerdir. Eğer, $i = 1, \dots, k$ için $B_i^{*(1)}$ ' lar $B_i^{*(1)} = B_{I_i}^{(1)}$ şeklinde belirlenirse, $B_1^{*(1)}, B_2^{*(1)}, \dots, B_k^{*(1)}$ blokları $\{B_1^{(1)}, B_2^{(1)}, \dots, B_N^{(1)}\}$ kümesinden yerine koyarak rasgele seçilmiş m çaplı bir örneklem oluşturur (Lahiri 2003). $(X_{1,1}^*, \dots, X_{1,l}^*)', (X_{1,l+1}^*, \dots, X_{1,2l}^*)', \dots, (X_{1,(k-1)l+1}^*, \dots, X_{1,m}^*)'$ birbirinden bağımsız aynı dağılımlı l -boyutlu rasgele vektörler olup, her birinin oluşturulan $B_1^{(1)}, B_2^{(1)}, \dots, B_N^{(1)}$ blokları arasından seçilmesi olasılığı $1/N$ dir. P_*, χ_n verildiğinde koşullu olasılığı göstermek üzere

$$P_*((X_{1,1}^*, \dots, X_{1,l}^*)' = (X_j, \dots, X_{j+l-1})') = P_*(I_1 = j) = \frac{1}{N}, \quad j = 1, \dots, N \quad (3.3)$$

dır (Lahiri 2003). Eğer her bir blok sadece bir gözlem içerirse yani $l = 1$ ise (3.3) eşitliğinden $X_{1,1}^*, \dots, X_{1,m}^*$ bootstrap örnekleme birbirinden bağımsız aynı F_n dağılımına sahip rasgele değişkenler olurlar. Böylece HBB yöntemi, Efron (1979) tarafından önerilen bildik bootstrap yöntemine dönüşür.

Bildik bootstrap yönteminde olduğu gibi HBB yönteminde de bootstrap örneklem çapı χ_n ' nin örneklem çapı ile aynı seçilebilir. Eğer $b_1, b_1.l \geq n$ eşitsizliğini sağlayan en küçük tamsayı ise HBB örneklemini oluşturmak için $k = b_1$ tane blok seçilebilir ve $\hat{\theta}_n$ ' nin bootstrap uyarlamasını ifade etmek için HBB örnekleminin ilk n değeri kullanılabilir (Lahiri 2003).

Yukarıda (3.2) eşitliğiyle verilen $\hat{\theta}_n$ istatistiklerinin bir-boyutlu F_n gözleme dayalı dağılımına bağlı olduğundan, gözlemlerin ortak dağılımına dayanan örneğin kovaryans gibi istatistiklerin tahminlerinde yetersiz kalırlar. Dolayısıyla bu istatistiklerin bootstrap uyarlamalarında da aynı sorunla karşılaşılır. Bu sorunu ortadan kaldırmak amacıyla Künsch (1989), p-boyutlu gözleme dayalı dağılıma dayanan istatistiklerin bootstrap uyarlamaları için daha genel bir HBB yöntemi önermiştir.

χ_n gözlemleri boyunca $n - p + 1$ tane tanımlanacak olan $Y_j = (X_j, \dots, X_{j+p-1})$ p-boyutlu rasgele vektörleri için

$$\delta_{Y_j} = \begin{cases} (X_j \leq x_j, \dots, X_{j+p-1} \leq x_{j+p-1}), & 1 \\ \text{diğer durumlarda} & , 0 \end{cases}$$

olmak üzere,

$$F_{p,n} = (n - p + 1)^{-1} \sum_{j=1}^{n-p+1} \delta_{Y_j}$$

p-boyutlu gözleme dayalı dağılım fonksiyonu olsun. Bu durumda

$$\hat{\theta}_n = T(F_{p,n}) \tag{3.4}$$

biçimindeki istatistiklerin HBB uyarlamasını elde etmek için izlenen adımlar aşağıda verilmiştir (Künsch 1989).

Adım 1. $1 < l < n - p + 1$ için l önceden belirlenen (sabit tamsayı) blok uzunluğu olmak üzere

$$\tilde{B}_j = (Y_j, \dots, Y_{j+l-1}), 1 \leq j \leq n - p - l + 2$$

olacak şekilde bloklar oluşturulur.

Adım 2. HBB örneklemini oluşturmak için $\{\tilde{B}_i : 1 \leq i \leq n - p - l + 2\}$ blokları arasından yerine koyarak $k \geq 1$ olacak şekilde k tane blok rasgele seçilir. $m = k.l$ çaplı $Y_1^*, \dots, Y_l^*; Y_{l+1}^*, \dots, Y_{2l}^*, \dots, Y_m^*$ HBB örneklemini oluşturulur.

Adım 3. $\tilde{F}_{m,n}^* = m^{-1} \sum_{j=1}^m \delta_{Y_j^*}$, Y_1^*, \dots, Y_m^* bootstrap örnekleminin gözleme dayalı dağılımı olmak üzere (3.4) eşitliğinde ifade edilen $\hat{\theta}_n$ istatistiğinin bootstrap uyarlaması,

$$\theta_{m,n}^* = T(\tilde{F}_{m,n}^*)$$

kullanılarak hesaplanır.

Bu yöntemle doğrudan X - değerlerinin bloklanması yerine bu değerlerin oluşturduğu p -boyutlu rasgele vektörler yani Y -değerleri bloklanmıştır. Bu yaklaşım HBB' nin “ sıradan (ordinary)” yaklaşımıdır (Lahiri 2003).

Bir diğer yaklaşım ise HBB' nin “ sade (naive)” yaklaşımıdır. Bu yaklaşımda, $p \geq 2$ için (3.4) eşitliğinde verilen $\hat{\theta}_n$ istatistiği X_1, \dots, X_n gözlemlerinin bir fonksiyonu

biçiminde ifade edilebileceği için bu istatistiğin bootstrap uyarlaması için $B_i = \{X_i, \dots, X_{i+l-1}\}$, $i = 1, \dots, N$ l uzunluklu blokları arasından yerine koyarak k tane rasgele B_j^* , $j = 1, \dots, k$ blok seçilmesiyle oluşturulan $X_{1,1}^*, \dots, X_{1,m}^*$ bootstrap gözlemlerinden yararlanarak p -boyutlu $Y_i^{**} \equiv (X_{1,i}^*, \dots, X_{1,i+p-1}^*)'$, $i = 1, \dots, m - p + 1$ rasgele vektörleri oluşturulur. Buradan p -boyutlu bootstrap gözlemine dayalı dağılım fonksiyonu $\tilde{F}_{m,n}^{**} = \frac{1}{m - p + 1} \sum_{i=1}^{m-p+1} \delta_{Y_i^{**}}$ elde edilir. Bu yaklaşım altında $\hat{\theta}_n$ istatistiğinin bootstrap uyarlaması

$$\theta_{m,n}^* = T(\tilde{F}_{m,n}^{**})$$

olacaktır.

$(X_n)_{n=1}^{\infty}$ durağan rasgele değişkenlerin bir dizisi olduğundan her i için $Y_i = (X_i, \dots, X_{i+p-1})'$ vektörü ile $(X_1, \dots, X_p)'$ vektörü aynı dağılımlı olacaktır. Böylece “ sıradan” yaklaşım kullanılarak oluşturulan Y_i^* HBB gözlemleri $(X_1, \dots, X_p)'$ ’ nin bağımlılık yapısını muhafaza edecektir. Fakat $l < p$ olması durumunda eğer bloklar “ sade” yaklaşım kullanılarak oluşturulursa, ard arda gelen B_j^* ve B_{j+1}^* bloklarının sınırlarında bulunan $X_{1,i}^*$ bootstrap gözlemleri birbirinden bağımsız olacaktır dolayısıyla bağımlılık yapısı bozulacaktır (Lahiri 2003).

3.2 Örtüşmeyen Bloklarla Bootstrap Yöntemi (ÖBB)

Bu bölümde, Carlstein (1986)’ nin bloklama yönteminden bahsedilecektir. Kolaylık açısından $p = 1$ olması durumunda (3.4) eşitliğinde verilen $\hat{\theta}_n$ biçimindeki istatistiklerin bootstrap uyarlamasının nasıl ifade edileceğine yer verilecektir.

Carlstein (1986) tarafından önerilen blok bootstrap yöntemi, HBB yönteminden farklı olarak gözlemlerin iç içe örtüşmeyecek şekilde bloklanmasına dayanır. Örtüşmeyen bloklarla bootstrap yönteminin işleyişi aşağıdaki gibidir.

Adım 1. $l \equiv l_n \in [1, n]$ önceden belirlenmiş (sabit tamsayı) blok uzunluğu ve $l \cdot b \leq n$ eşitsizliğini sağlayan en büyük tamsayı b , oluşturulan blok sayısı olmak üzere

$$B_i^{(2)} = (X_{(i-1)l+1}, \dots, X_{il})', \quad i = 1, 2, \dots, b$$

olacak şekilde bloklar tanımlanır (bkz. Şekil 3.2.).

Adım 2. ÖBB örneklemini oluşturmak için $\{B_1^{(2)}, \dots, B_b^{(2)}\}$ blokları arasından yerine koyarak $k \geq 1$ olacak şekilde rasgele k tane $B_1^{*(2)}, \dots, B_k^{*(2)}$ blokları seçilir. $m = k \cdot l$, ÖBB örnekleminin çapı olmak üzere

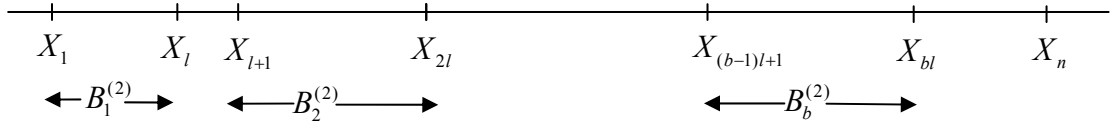
$$(X_{2,1}^*, \dots, X_{2,l}^*; \dots; X_{2,(k-1)l+1}^*, \dots, X_{2,m}^*)$$

ÖBB örneklemini oluşturulur.

Adım 3. $F_{m,n}^{*(2)}$, $(X_{2,1}^*, \dots, X_{2,m}^*)$ ' in gözleme dayalı dağılımı olmak üzere $\hat{\theta}_n$ istatistiğinin bootstrap uyarlaması

$$\theta_{m,n}^{*(2)} = T(F_{m,n}^{*(2)})$$

dir.



Şekil 3.2 ÖBB için $\{B_1^{(2)}, B_2^{(2)}, \dots, B_b^{(2)}\}$ bloklarının oluşturulması (Lahiri 2003)

ÖBB yöntemiyle elde edilen $(X_{2,1}^*, \dots, X_{2,l}^*)', \dots, (X_{2,\{(k-1)l+1\}}^*, \dots, X_{2,m}^*)'$ bootstrap değişkenleri birbirinden bağımsız aynı dağılımlı l -boyutlu rasgele vektörler olup P_* , χ_n verildiğinde koşullu olasılığı göstermek üzere,

$$P_*((X_{2,1}^*, \dots, X_{2,l}^*)' = (X_{(j-1)l+1}, \dots, X_{jl})') = 1/b, \quad j = 1, 2, \dots, b \quad (3.5)$$

dır (Lahiri 2003).

ÖBB yönteminde gözlemlerin örtüşmeyecek şekilde bloklanması dolaylı ÖBB yönteminde oluşturulan blok sayısı, HBB yönteminde oluşturulan blok sayısından daha azdır. Ayrıca $\theta_{m,n}^{*(1)}$ ile $\theta_{m,n}^{*(2)}$ bootstrap tahmin edicileri farklı dağılım özelliklerine sahiptir.

$\hat{\theta}_n = \frac{1}{n} \sum_{j=1}^n X_j$, örneklem ortalaması olmak üzere $\hat{\theta}_n$ ' nin HBB ve ÖBB yöntemleri

kullanılarak elde edilen bootstrap uyarlamaları sırasıyla

$$\theta_{m,n}^{*(1)} = \frac{1}{m} \sum_{j=1}^m X_{1,j}^* = \bar{X}_m^{*(1)} \quad \text{ve} \quad \theta_{m,n}^{*(2)} = \frac{1}{m} \sum_{j=1}^m X_{2,j}^* = \bar{X}_m^{*(2)}$$

olacaktır. $U_i^{(1)} = (X_i + \dots + X_{i+l-1})/l$ HBB yöntemi ile oluşturulan her bir blok ortalaması ve bu bloklar arasından yerine koyarak rasgele seçilen k tane bloğun her birinin ortalaması $U_i^{*(1)} = (X_{1,(i-1)l+1}^* + \dots + X_{1,li}^*)/l$ olmak üzere HBB örneklem ortalamasının beklenen değeri

$$E_*(\bar{X}_m^{*(1)}) = E_*(k^{-1} \sum_{i=1}^k U_i^{*(1)})$$

biçimindedir. $U_1^{*(1)}, \dots, U_k^{*(1)}$ rasgele değişkenlerinin aynı dağılımlı olmasından dolayı yukarıdaki eşitlik,

$$\begin{aligned} E_*(\bar{X}_m^{*(1)}) &= \frac{1}{k} k E_*(U_1^{*(1)}) \\ &= E_*(U_1^{*(1)}) \end{aligned}$$

olacaktır. (3.3) eşitliğinden

$$P_*(U_1^{*(1)} = U_i^{(1)}) = 1/N, 1 \leq i \leq N$$

olup, $E_*(U_1^{*(1)}) = \frac{1}{N} \sum_{i=1}^N U_i^{(1)}$ dir ve $U_i^{(1)} = (X_i + \dots + X_{i+l-1})/l$ ifadesi kullanılarak

$$\begin{aligned} E_*(\bar{X}_m^{*(1)}) &= \frac{1}{N} \sum_{i=1}^N (X_i + \dots + X_{i+l-1})/l \\ &= N^{-1} \left[\sum_{i=l}^N X_i + \sum_{i=1}^{l-1} i/l (X_i + X_{n-i+l}) \right] \end{aligned}$$

biçiminde elde edilir (Athreya and Lahiri 2006). $U_i^{(2)} = (X_{(i-1)l+1} + \dots + X_{il})/l$ ÖBB yöntemi ile oluşturulan her bir bloğun ortalaması ve bu bloklar arasından yerine koyarak rasgele seçilen k tane bloğun her birinin ortalaması $U_i^{*(2)} = (X_{2,(i-1)l+1}^* + \dots + X_{2,li}^*)/l$ olmak üzere, ÖBB örneklem ortalamasının beklenen değeri

$$E_*(\bar{X}_m^{*(2)}) = E_* \left(k^{-1} \sum_{i=1}^k U_i^{*(2)} \right)$$

biçimindedir. $U_1^{*(2)}, \dots, U_k^{*(2)}$ rasgele değişkenlerinin aynı dağılımlı olmasından dolayı yukarıdaki eşitlik,

$$\begin{aligned} E_*(\bar{X}_m^{*(2)}) &= \frac{1}{k} k E_*(U_1^{*(2)}) \\ &= E_*(U_1^{*(2)}) \end{aligned}$$

olacaktır. (3.5) eşitliğinden

$$P_*(U_1^{*(2)} = U_i^{(2)}) = 1/b, 1 \leq i \leq b$$

olup, $E_*(U_1^{*(2)}) = \frac{1}{b} \sum_{i=1}^b U_i^{(2)}$ dır ve $U_i^{(2)} = (X_{(i-1)l+1} + \dots + X_{il})/l$ ifadesi kullanılarak

$$\begin{aligned} E_*(\bar{X}_m^{*(2)}) &= \frac{1}{b} \sum_{i=1}^b l^{-1} \sum_{j=1}^l X_{(i-1)l+j} \\ &= (bl)^{-1} \left\{ n \bar{X}_n - \sum_{i=bl+1}^n X_i \right\} \end{aligned} \quad (3.6)$$

biçiminde elde edilir (Lahiri 2003). Eğer l, n ' i tam bölen bir sayı ise (3.6) eşitliği \bar{X}_n ' e eşittir.

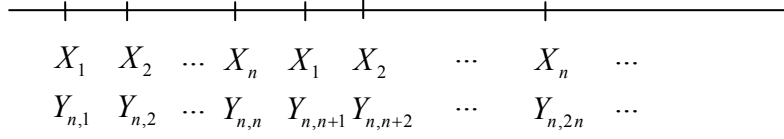
Her iki yöntem altında bootstrap tahmin edicilerinin χ_n verildiğinde koşullu beklenen değerleri farklıdır. Ancak $(X_n)_{n=1}^{\infty}$ rasgele değişkenler dizisi bazı beklenen değer ve α - miksing olma koşullarını sağlarsa $E \left\{ E_*(\theta_{m,n}^{*(1)}) - E_*(\theta_{m,n}^{*(2)}) \right\}^2 = o(l/n^2)$ olacaktır. Bundan dolayı, büyük örneklem çapları için, iki değer arasındaki fark önemsizdir (Lahiri 2003).

HBB yöntemi, gözlemlerin baş ve son kısımlarından daha çok orta kısımlarına daha fazla ağırlık vererek istenilmeyen sınır (uç) etkilerine maruz kalmaktadır. $l \leq j \leq n-l$ olmak üzere j -inci gözlem $X_j \{B_1^{(l)}, B_2^{(l)}, \dots, B_N^{(l)}\}$ bloklarından tam olarak l tanesinde gözlemlenirken, $1 \leq j \leq l-1$ için X_j ve X_{n-j+1} gözlemleri sadece j tane blokta gözlemlenir. X_n gözleminde sonra veya X_1 gözleminde önce herhangi bir gözlem bulunmadığından bu etkiyi ortadan kaldırmak için yeni bloklar da tanımlanamaz. Benzer problem n , blok uzunluğu l 'nin tam katı olmadığına ÖBB yönteminde de ortaya çıkmaktadır (Lahiri 2003). Politis and Romano (1992) bu sınır problemini ortadan kaldırmaya yönelik bir yöntem önermişlerdir. Bu yöntem blokları, gözlemleri bir çemberin etrafına sararak oluşturmaya dayanır ve bu yönteme “Çembersel (circular) Bloklarla Bootstrap (ÇBB) yöntemi” adı verilmiştir. Aşağıdaki bölümde literatürde bilinen blok bootstrap yöntemlerinin ana fikrini oluşturan Genelleştirilmiş Blok Bootstrap (GBB) yöntemi tanıtıldıktan sonra izleyen bölümde bu yöntemin özel bir hali ÇBB yönteminden bahsedilecektir.

3.3 Genelleştirilmiş Blok Bootstrap (GBB) Yöntemi

$\mathcal{X}_n = \{X_1, \dots, X_n\}$ gözlemleri gösterecek ve $\Gamma_n, \mathbb{R}^n \times \bigotimes_{t=1}^{\infty} (\{1, \dots, n\} \times \mathbb{N})$ kümesinde tanımlı geçiş olasılık fonksiyonu olsun. Örneğin, $\forall x \in \mathbb{R}^n$ için $\Gamma_n(x; \cdot)$, $\bigotimes_{t=1}^{\infty} (\{1, \dots, n\} \times \mathbb{N}) \equiv \left\{ \{i_t, l_t\}_{t=1}^{\infty} : 1 \leq i_t \leq n, 1 \leq l_t < \infty, \forall t \geq 1 \right\}$ kümesinde tanımlı bir olasılık ölçüsüdür. Burada $\{i_t, l_t\}$ elemanları ile olabilecek tüm sıralı ikililer kümesi ifade edilmektedir. GBB yöntemi için izlenen adımlar aşağıda verilmiştir (Lahiri 2003).

Adım 1. Herhangi bir $i \geq 1$ ve $j_i \in [1, n]$ için $i = j_i \text{ mod}(n)$ olmak üzere $Y_{n,i} = X_{j_i}$ olacak şekilde periyodik olarak ilerleyen $Y_{n,i}$ değişkenleri tanımlanır. Aslında bu tanımlama X_1, X_2, \dots, X_n değişkenlerini arka arkaya defalarca yazarak $i \geq 1$ için her bir X_i değişkenini $Y_{n,i}$ değişkenleri ile indirmekle aynıdır (bkz. Şekil 3.3.).



Şekil 3.3 $Y_{n,i}$ değişkenlerinin oluşturulması (Lahiri 2003)

Adım 2. $i \geq 1$ ve $j \geq 1$ için

$$B(i, j) = (Y_{n,i}, \dots, Y_{n,(i+j-1)})$$

olacak şekilde bloklar tanımlanır.

Adım 3. $(I_1, J_1), (I_2, J_2), \dots$ ikilileri χ_n verilmişken ortak koşullu dağılımları $\Gamma_n(\chi_n; \cdot)$ olan rasgele vektörler dizisi olmak üzere, GBB örneklemini oluşturmak için $\{B(i, j) : i \geq 1, j \geq 1\}$ blokları arasından $B(I_1, J_1), B(I_2, J_2), \dots$ blokları seçilir.

Adım 4. $X_{G,1}^*, X_{G,2}^*, \dots, X_{G,m}^*$ m çaplı GBB örnekleme ve $m \geq 1$ için $F_{m,n}^{*(G)}$ bu örneklemin gözleme dayalı dağılımı olmak üzere (3.2) eşitliğinde verilen $\hat{\theta}_n$ tahmin edicisinin bu yöntem kullanılarak elde edilen bootstrap uyarlaması

$$\theta_{m,n}^{*(G)} = T(F_{m,n}^{*(G)})$$

dır.

Bilinen tüm blok bootstrap yöntemleri GBB yönteminin özel hali olarak gösterilebilir. Örneğin, $1 \leq l \leq n$ için blok uzunluğu l olan HBB yöntemi düşünülün. $N = n - l + 1$

olmak üzere geçiş olasılık fonksiyonu, $\Gamma_n(x;.)$ olsun. Bu durumda $(I_1, J_1), (I_2, J_2), \dots$ rasgele indisleri

$$P_*(I_1 = j, I_2 = k) = \begin{cases} N^{-1}, & 1 \leq j \leq N \text{ ve } k = l \\ 0, & \text{d.y.} \end{cases}$$

olasılıkları ile b.b.a.d rasgele vektörler olacaktır. Böylece Bölüm 3.1' de bahsedilen $\{B_1, \dots, B_N\}$ kümesi ile $\{B(i, j) : 1 \leq i \leq N, j = l\}$ kümesi aynı olup $B(I_1, J_1), B(I_2, J_2), \dots$ blokları bu kümeden seçilmiş olur (Lahiri 2003).

3.3.1 Çembersel Bloklarla Bootstrap (ÇBB) Yöntemi

Politis and Romano (1992) tarafından önerilen bu yöntem $\mathcal{X}_n = \{X_1, \dots, X_n\}$ kümesindeki gözlemleri bir çemberin etrafına sardıktan sonra her bir gözlemi Bölüm 3.3' de tanımlanan $Y_{n,i}$ değişkenleri ile indisleyerek $1 \leq i \leq n$ için $B(i, l) = (Y_{n,i}, \dots, Y_{n,(i+l-1)})$ olacak şekilde oluşturulan n tane örtüşen bloktan k tanesini yeniden örneklemeğe dayanır. Böylelikle her bir X_i gözlemi $\{B(1, l), \dots, B(n, l)\}$ kümesindeki bloklardan l tanesinde yer alır.

ÇBB yöntemi için geçiş fonksiyonu, $\Gamma_n(x;.)$ olmak üzere $I_{3,1}, I_{3,2}, \dots$ rasgele değişkenleri $\{B(1, l), \dots, B(n, l)\}$ kümesinden rasgele seçilen blokların indisleri olarak ifade edilirse bu rasgele değişkenler birbirinden bağımsız ve

$$P_*(I_{3,1} = i) = \frac{1}{n}$$

olasılık fonksiyonu ile aynı dağılıma sahip olacaktır. Her bir X_i gözleminin $\{B(1, l), \dots, B(n, l)\}$ kümesindeki bloklardan l tanesinde yer alması ve bu kümeden eşit

olasılıkla bloklar seçilip yeniden örnekleme yapıldığı için her bir X_i gözlemi ÇBB yöntemi altında eşit ağırlığa sahiptir.

$m = k.l$ için $X_{3,1}^*, X_{3,2}^*, \dots, X_{3,m}^*$ m çaplı ÇBB örneklemini ve $\overline{X}_m^{*(3)}$ ÇBB örneklem ortalamasını gösterebiliriz. $U_i^{(3)} = (Y_{n,i} + \dots + Y_{n,(i+l-1)})/l$ ÇBB yöntemi ile oluşturulan her bir bloğun ortalaması ve bu bloklar arasından yerine koyarak rasgele seçilen k tane bloğun her birinin ortalaması $U_i^{*(3)} = (Y_{n,(i-1)l+1}^* + \dots + Y_{n,li}^*)/l$ olmak üzere $U_1^{*(3)}, \dots, U_k^{*(3)}$ rasgele değişkenleri birbirinden bağımsız aynı dağılımlı olup

$$P_*(U_1^{*(3)} = U_i^{(3)}) = \frac{1}{n}$$

dir. Bu durumda ÇBB örneklemini sonucunda $\overline{X}_m^{*(3)}$ bootstrap örneklem ortalamasının beklenen değeri,

$$E_*(\overline{X}_m^{*(3)}) = E_*(k^{-1} \sum_{i=1}^k U_i^{*(3)})$$

biçiminde verilir. $U_1^{*(3)}, \dots, U_k^{*(3)}$ rasgele değişkenlerinin aynı dağılımlı olmasından,

$$\begin{aligned} E_*(\overline{X}_m^{*(3)}) &= E_*(U_1^{*(3)}) \\ &= \frac{1}{n} \sum_{i=1}^n U_i^{(3)} \end{aligned}$$

dır. $U_i^{(3)} = (Y_{n,i} + \dots + Y_{n,(i+l-1)})/l$ ifadesi yardımıyla

$$\begin{aligned} E_*(\overline{X}_m^{*(3)}) &= \frac{1}{n} \sum_{i=1}^n l^{-1} \left(\sum_{j=1}^l Y_{n,(i+j-1)} \right) \\ &= l^{-1} (l \overline{X}_n) = \overline{X}_n \end{aligned}$$

elde edilir (Lahiri 2003).

3.4 HBB, ÖBB ve ÇBB Yöntemleriyle Örneklem Varyansı ve Dağılım Fonksiyonu Tahmin Edicilerinin Tutarlılığı

$E(X_1) = \mu$ ve $E(X_1^2) < \infty$ olmak üzere $(X_i)_{i \in \mathbb{Z}}$, \mathbb{R} ' de değerler alan $\alpha(\cdot)$ güçlü miksing katsayısına sahip durağan rasgele değişkenlerin bir dizisi olsun. Basitlik sağlaması nedeniyle her bir blok bootstrap yöntemi için yeniden örneklenen blok sayısı $k = n/l$ olacak şekilde belirlensin. Bu durumda her bir blok bootstrap yöntemi ile elde edilen örneklem, ana örneklem ile aynı sayıda gözlem içerecektir. $\bar{X}_n^{*(1)}$, $\bar{X}_n^{*(2)}$, $\bar{X}_n^{*(3)}$ sırasıyla HBB, ÖBB ve ÇBB yöntemleri ile elde edilen bootstrap örneklem ortalamaları olmak üzere $T_n = \sqrt{n}(\bar{X}_n - \mu)$ istatistiğinin blok bootstrap tahmin edicileri sırasıyla,

$$T_n^{*(j)} = \sqrt{n}(\bar{X}_n^{*(j)} - E_*(\bar{X}_n^{*(j)})), j = 1, 2, 3$$

olacaktır.

$Var(T_n)$ ' in blok bootstrap varyans tahmin edicileri $Var_*(T_n^{*(j)}), j = 1, 2, 3$, χ_n gözlemlerini içeren basit formüllerle hesaplanabilir. $U_i^{(1)} = (X_i + \dots + X_{i+l-1})/l$ HBB yöntemi ile oluşturulan her bir bloğun ortalaması, $U_i^{(2)} = (X_{(i-1)l+1} + \dots + X_{il})/l$ ÖBB yöntemi ile oluşturulan her bir bloğun ortalaması ve $U_i^{(3)} = (Y_{n,i} + \dots + Y_{n,i+l-1})/l$ ÇBB yöntemi ile oluşturulan her bir bloğun ortalaması olsun. HBB ve ÖBB yöntemleri için bloklar arasından yerine koyarak rasgele seçilen k tane bloğun her birinin ortalaması $U_i^{*(j)} = (X_{j,(i-1)l+1}^* + \dots + X_{j,il}^*)/l, j = 1, 2$ ve ÇBB yöntemi için bloklar arasından yerine koyarak rasgele seçilen k tane bloğun her birinin ortalaması $U_i^{*(j)} = (Y_{n,(i-1)l+1}^* + \dots + Y_{n,il}^*)/l, j = 3$ olmak üzere $U_1^{*(j)}, \dots, U_k^{*(j)}, j = 1, 2, 3$ rasgele değişkenlerinin bağımsızlığı kullanılarak bootstrap varyans tahmin edicileri,

$$\begin{aligned}
Var_*(T_n^{*(j)}) &= nVar_*(\bar{X}_n^{*(j)}) \\
&= nVar_*(\frac{1}{k} \sum_{i=1}^k U_i^{*(j)}) \\
&= n \frac{1}{k} Var_*(U_1^{*(j)})
\end{aligned}$$

olacaktır (Athreya and Lahiri 2006).

Böylece $\hat{\mu}_{n,1} = E_*(U_1^{*(1)}) = \frac{1}{N} \sum_{i=1}^N U_i^{(1)}$, $\hat{\mu}_{n,2} = E_*(U_1^{*(2)}) = \frac{1}{b} \sum_{i=1}^b U_i^{(2)}$ olmak üzere

$$Var_*(U_1^{*(1)}) = \frac{1}{N} \sum_{i=1}^N (U_i^{(1)})^2 - \hat{\mu}_{n,1}^2$$

$$Var_*(U_1^{*(2)}) = \frac{1}{b} \sum_{i=1}^b (U_i^{(2)})^2 - \hat{\mu}_{n,2}^2$$

$$Var_*(U_1^{*(3)}) = \frac{1}{n} \sum_{i=1}^n (U_i^{(3)})^2 - \bar{X}_n^2$$

dır. Buradan,

$$Var_*(T_n^{*(1)}) = l \left[\frac{1}{N} \sum_{i=1}^N (U_i^{(1)})^2 - \hat{\mu}_{n,1}^2 \right]$$

$$Var_*(T_n^{*(2)}) = l \left[\frac{1}{b} \sum_{i=1}^b (U_i^{(2)})^2 - \hat{\mu}_{n,2}^2 \right]$$

$$Var_*(T_n^{*(3)}) = l \left[\frac{1}{n} \sum_{i=1}^n (U_i^{(3)})^2 - \bar{X}_n^2 \right]$$

olacaktır (Athreya and Lahiri 2006).

$\delta > 0$ için $E|X_1|^{2+\delta} < \infty$ ve $\sum_{n=1}^{\infty} \alpha(n)^{\delta/2+\delta} < \infty$ koşulları sağlandığında $Z_i = X_i - \mu, i \in \mathbb{Z}$

olmak üzere T_n istatistiğinin asimptotik varyansı,

$$\sigma_{\infty}^2 = \lim_{n \rightarrow \infty} \text{Var}(T_n) = \sum_{i=-\infty}^{\infty} E(Z_1 Z_{1+i})$$

sonsuz serisi olup yukarıdaki koşullara ek olarak eğer $n \rightarrow \infty$ iken $l^{-1} + n^{-1}l = o(1)$ ise

$$\text{Var}_*(T_n^{*(j)}) \xrightarrow{p} \sigma_{\infty}^2, n \rightarrow \infty, j = 1, 2, 3 \quad (3.7)$$

olacaktır (Athreya and Lahiri 2006, Theorem 17.4.1).

T_n istatistiğinin örneklem dağılım fonksiyonu,

$$G_n(x) = P(T_n \leq x), x \in \mathbb{R}$$

ve $\hat{G}_n^{(j)}(x) = P_*(T_n^{*(j)} \leq x), j = 1, 2, 3$ sırasıyla HBB, ÖBB ve ÇBB yöntemleri altında $T_n^{*(j)}$ blok bootstrap tahmin edicisinin dağılım fonksiyonu olsun. Klasik olarak G_n dağılımında G_{∞} limit dağılımına yakınsadığında yaklaşılmaya çalışılırken, blok bootstrap yöntemlerinde örneklem çapı n değiştikçe değişen, rasgele olan $\hat{G}_n^{(j)}$ ' leri üreterek G_n dağılımına yaklaşılmaya çalışılır (Lahiri 2003). T_n istatistiğinin asimptotik dağılımı

$$T_n \xrightarrow{d} N(0, \sigma_{\infty}^2)$$

dır. $\delta > 0$ için $E|X_1|^{2+\delta} < \infty$ ve $\sum_{n=1}^{\infty} \alpha(n)^{\delta/2+\delta} < \infty$ koşulları sağlansın,

$\sigma_{\infty}^2 = \sum_{i \in \mathbb{Z}} \text{Cov}(X_1, X_{1+i})$ ve $n \rightarrow \infty$ iken $l^{-1} + n^{-1}l = o(1)$ ise

$$\sup_{x \in \mathbb{R}} \left| P_*(T_n^{*(j)} \leq x) - P(T_n \leq x) \right| \xrightarrow{p} 0, n \rightarrow \infty \quad (3.8)$$

olacaktır (Athreya and Lahiri 2006, Theorem 17.4.3).

Sırasıyla (3.7) ve (3.8) ifadeleri $(X_i)_{i \in \mathbb{Z}}$ rasgele değişkenler dizisi üzerine konulan bazı moment ve ifade edilen güçlü miksing koşulları altında bootstrap varyans tahmin edicileri $\text{Var}_*(T_n^{*(j)}), j=1,2,3$ ve G_n örneklem dağılımının blok bootstrap tahmin edicileri $\hat{G}_n^{(j)}, j=1,2,3$ ' lerin, $n \rightarrow \infty$ iken l blok uzunluğunun n ' e göre daha yavaş hızda sonsuza gitmesi durumunda tutarlı olduğunu göstermektedir. Bu nedenle, $\text{Var}_*(T_n^{*(j)}), j=1,2,3$ blok bootstrap tahmin edicilerinin tutarlı olmasını sağlayan $l = \log \log n$ ve $0 < \varepsilon < 1$ için $l = n^{1-\varepsilon}$ blok uzunlukları kabul edilebilir blok uzunluklarıdır (Lahiri 2003). Eğer l sınırlı olursa blok bootstrap yöntemleri ana örneklemin bağımlılık yapısını muhafaza edemezler ve Singh (1981)' in örneğinde olduğu gibi σ_{∞}^2 ' dan farklı bir varyans ile normal dağılıma yakınsama olur (Athreya and Lahiri 2006). Diğer taraftan eğer l , gözlem sayısı n ile aynı hızda sonsuza giderse ($n^{-1}l = o(1)$ şartı bozulursa) sonsuz bir kitleyi temsil edecek bir örnekleme yaratmak için yeterli sayıda farklı blok yoktur ve bu nedenle örneğin $\hat{G}_n^{(j)}$ rasgele bir dağılım fonksiyonuna yakınsar (Lahiri 2003).

3.5 Varyans ve Yanlılık Bootstrap Tahmin Edicileri için Blok Bootstrap Yöntemlerinin Karşılaştırılması

χ_n örnekleme bağılı olarak θ parametresinin bir tahmin edicisi $\hat{\theta}_n$ olsun. θ ilgi duyduğumuz yığının parametresi olup düzey-1 parametresi olarak bilinir. $\hat{\theta}_n$ ' nin örnekleme dağılımına ait örneğin $HKO(\hat{\theta}_n)$ gibi bir parametresi, düzey-2 parametresi olarak bilinir. Diğer yeniden örnekleme yöntemlerinde olduğu gibi bootstrap yöntemi de asli olarak düzey-2 parametreleri ile ilgili istatistiki çıkarıma yönelir.

Bağımlı gözlemler durumunda özel olarak, $HKO(\hat{\theta}_n)$ üzerine yapılmış çalışmalardan Hall, Horowitz and Jing (1995), Lahiri (1999) ve Politis and White (2004) olarak kaynak verilebilir.

Bağımlı gözlemlerle bootstrap yöntemleri için $HKO(\hat{\theta}_n)$ hakkında bilgi sahibi olmanın diğer bir yararı da bootstrapün başarımı için belirleyici role sahip blok uzunluğunun belirlenmesidir. Bu amaçla Hall, Horowitz and Jing (1995) blok bootstrap yöntemleri için n gözlem çapı ve C pozitif bir sabit olmak üzere eğer varyans ve yan tahmini yapılacaksa blok uzunluğunun $l \sim Cn^{1/3}$, eğer $P(X \leq x)$ dağılım fonksiyonu tahmin edilecekse $l \sim Cn^{1/4}$ ve $P(a < X \leq b)$ tahmin edilecekse $l \sim Cn^{1/5}$ olması gerektiğini açıklamıştır.

$(X_i)_{i \in \mathbb{Z}}, \mathbb{R}^d$ de değerler alan $E(X_1) = \mu$ ortalamalı, $\alpha(\cdot)$ güçlü miksing katsayısına sahip durağan rasgele değişkenlerin bir dizisi olsun. $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ ve $H: \mathbb{R}^d \rightarrow \mathbb{R}$ düzgün (smooth) bir fonksiyon olmak üzere ilgili düzey-1 parametresi θ ve onun tahmin edicisi $\hat{\theta}_n$ olmak üzere $\theta = H(\mu)$ ve $\hat{\theta}_n = H(\bar{X}_n)$ şeklinde ifade edilebilir.

$X_{j,1}^*, X_{j,2}^*, \dots, X_{j,n}^*$ $j=1,2,3$ sırasıyla HBB, ÖBB ve ÇBB yöntemleri ile elde edilen bootstrap örneklemini gösterebilir ve her üç yöntem için l blok uzunluğu olmak üzere

$k = \lceil n/l \rceil$ tane blok yeniden örneklensin. Bu durumda $\overline{X}_{n,l}^{*(j)}$, $j = 1, 2, 3$ her üç yöntem için bootstrap örneklem ortalamasını göstermek üzere $T_n = \hat{\theta}_n - \theta$ istatistiğinin her üç yöntem altında blok bootstrap tahmin edicisi

$$T_{n,l}^{*(j)} = H(\overline{X}_{n,l}^{*(j)}) - H(E_*(\overline{X}_{n,l}^{*(j)})), j = 1, 2, 3$$

olacaktır (Lahiri 2003).

φ_n , $\hat{\theta}_n$ tahmin edicisinin dağılımının bir fonksiyonu olan ilgili düzey-2 parametresi olmak üzere

$$\varphi_{1n} \equiv \text{Yan}(\hat{\theta}_n) = E(\hat{\theta}_n) - \theta = E(T_n) \quad (3.9)$$

$$\varphi_{2n} \equiv \text{Var}(\hat{\theta}_n) = E(\hat{\theta}_n - E(\hat{\theta}_n))^2 = \text{Var}(T_n) \quad (3.10)$$

şeklinde tanımlansın. $\hat{\varphi}_n(j;l)$, $j = 1, 2, 3$ j -inci blok bootstrap yöntemi için φ_n ' nin l blok uzunluklu blok bootstrap tahmin edicisi olmak üzere (3.9) ve (3.10) ifadelerinde verilen düzey-2 parametrelerinin blok bootstrap tahmin edicileri sırasıyla,

$$\hat{\varphi}_{1n}(j;l) \equiv \widehat{\text{Yan}}_j(l) = E_*(T_{n,l}^{*(j)}), j = 1, 2, 3$$

$$\hat{\varphi}_{2n}(j;l) \equiv \widehat{\text{Var}}_j(l) = \text{Var}_*(T_{n,l}^{*(j)}), j = 1, 2, 3$$

dır. Lahiri (1999) bu tahmin edicileri elde etmede HBB, ÖBB ve ÇBB yöntemlerinin başarımlarını karşılaştırırken bir düzey-2 parametresi olan hata kareler ortalaması (HKO) ölçütünü kullanmıştır bu amaçla aşağıda blok bootstrap tahmin edicileri $\widehat{\text{Yan}}_j(l)$ ve $\widehat{\text{Var}}_j(l)$ ' nin hata kareler ortalamaları için bazı koşullar altında Lahiri (1999) tarafından elde edilen açılımları verilecektir.

Y rasgele deęişkeni için $HKO(Y) = [Yan(Y)]^2 + Var(Y)$ olmak üzere Lahiri (1999) HKO' nun bir açılımını elde etmek için bootstrap tahmin edicilerinin ayrı ayrı yan ve varyanslarını inceleyerek bunları HKO ölçütü bakımından tahmin edicilerin iyiliklerinin basit bir ölçüsünü elde etmek için birleştirmiştir. $v = (v_1, \dots, v_d)' \in \mathbb{Z}_+^d$ için $|v| = v_1 + v_2 + \dots + v_d$ ve $D^v = \frac{\partial^{|v|}}{\partial x_1^{v_1} \dots \partial x_d^{v_d}}$ olmak üzere açılımları elde etmek için aşağıdaki koşullar kullanılmıştır (Lahiri 1999).

Koşul D_r : $H: \mathbb{R}^d \rightarrow \mathbb{R}$ r kez diferansiyellenebilir bir fonksiyon ve $a_0 \geq 1$ tamsayısı için $\max \left\{ |D^v H(x)| : |v| = r \right\} \leq C(1 + \|x\|^{a_0}), x \in \mathbb{R}^d$ olsun.

Koşul M_r : $\delta > 0$ için $E\|X_1\|^{2r+\delta} < \infty$ ve $\Delta(r; \delta) = 1 + \sum_{n=1}^{\infty} n^{2r-1} \alpha(n)^{\delta/(2r+\delta)} < \infty$ olsun.

$\hat{\phi}_{1n}(j; l)$ ve $\hat{\phi}_{2n}(j; l)$ blok bootstrap tahmin edicilerinin yan ve varyanslarına ilişkin sonuçlar sırasıyla Teorem 3.2 ve Teorem 3.3' de verilmiştir.

Teorem 3.2 $l, n \rightarrow \infty$ iken $l^{-1} + n^{-1/2}l = o(1)$ şartını sağlayan bir blok uzunluğu olsun.

(a) $r = 3$ için Koşul D_r ve $r = 3 + a_0$ için Koşul M_r sağlansın ve $c_1 = c_1(f)$ için

$$Yan(\widehat{Yan}_j(l)) = n^{-1}l^{-1}c_1 + o(n^{-1}l^{-1}), j = 1, 2, 3, n \rightarrow \infty$$

(b) $r = 2$ için Koşul D_r ve $r = 4 + 2a_0$ için Koşul M_r sağlansın. Bu durumda $c_2 = c_2(f)$ için

$$Yan(\widehat{Var}_j(l)) = n^{-1}l^{-1}c_2 + o(n^{-1}l^{-1}), j = 1, 2, 3, n \rightarrow \infty$$

dır (Lahiri 1999).

Teorem 3.2, ister örtüşen bloklar isterse örtüşmeyen bloklar kullanılsın φ_{1n} ve φ_{2n} ' nin blok bootstrap tahmin edicilerinin yanlarının her üç blok bootstrap yöntemi için aynı olduğunu göstermektedir.

Teorem 3.3 $l, n \rightarrow \infty$ iken $l^{-1} + n^{-1/2}l = o(1)$ olan bir blok uzunluğu olsun. Teorem 3.2' nin (a) ve (b) kısımları için r tamsayısı üzerine konulan şartlar sağlansın. Hesaplamaları oldukça karmaşık olan $g_1 = g_1(f)$ ve $g_2 = g_2(f)$ fonksiyonları için

$$(a) \text{Var}(\widehat{Yan}_j(l)) = \frac{4\pi^2 g_1 l}{3n^3} + o(n^{-3}l), j = 1, 3, n \rightarrow \infty$$

$$\text{Var}(\widehat{Yan}_j(l)) = \frac{2\pi^2 g_1 l}{n^3} + o(n^{-3}l), j = 2, n \rightarrow \infty$$

$$(b) \text{Var}(\widehat{Var}_j(l)) = \frac{4\pi^2 g_2 l}{n^3} + o(n^{-3}l), j = 1, 3, n \rightarrow \infty$$

$$\text{Var}(\widehat{Var}_j(l)) = \frac{2\pi^2 g_2 l}{n^3} + o(n^{-3}l), j = 2, n \rightarrow \infty$$

dır (Lahiri 1999).

Teorem 3.3' den $\varphi_{1n} = Yan(\hat{\theta})$ ve $\varphi_{2n} = Var(\hat{\theta})$ ' in HBB ve ÇBB yöntemleri ile elde edilen tahmin edicilerinin varyansları,

$$\text{Var}(\widehat{Yan}_j(l)) = \frac{2}{3} \text{Var}(\widehat{Yan}_2(l)), j = 1, 3$$

$$\text{Var}(\widehat{Var}_j(l)) = \frac{2}{3} \text{Var}(\widehat{Var}_2(l)), j = 1, 3$$

eşitliklerinden dolayı ÖBB yöntemi ile elde edilen tahmin edicilerin varyanslarına göre daha küçüktür (Lahiri 1999). HBB ve ÇBB yöntemleriyle elde edilen tahmin edicilerin varyanslarının küçük olmasının sebebi bu yöntemlerde örtüşen blokların kullanılmasıdır.

Teorem 3.2 ve Teorem 3.3' den her üç blok bootstrap yöntemi için eğer l blok uzunluğu arttırılırsa, blok bootstrap tahmin edicilerinin yanları azalırken varyanslarının arttığı görülmektedir. Sonuç olarak her bir blok bootstrap tahmin edicisi için hata kareler ortalamasını en küçük yapan bir l blok uzunluğu vardır ve bu blok uzunluğuna HKO-optimal blok uzunluğu (MSE-optimal block length) denir (Lahiri 1999).

$0 < \varepsilon < \frac{1}{2}$ için

$$l_{1j}^o = \arg \min \left\{ HKO(\widehat{Yan}_j(l)) : n^\varepsilon \leq l \leq n^{(1-\varepsilon)/2} \right\}, j = 1, 2, 3$$

ve

$$l_{2j}^o = \arg \min \left\{ HKO(\widehat{Var}_j(l)) : n^\varepsilon \leq l \leq n^{(1-\varepsilon)/2} \right\}, j = 1, 2, 3$$

olmak üzere aşağıdaki sonuç her üç blok bootstrap yöntemi için blok bootstrap tahmin edicilerini elde etmek için kullanılan $l_{kj}^o, k=1,2, j=1,2,3$ optimal blok uzunluklarını göstermektedir (Lahiri 1999).

Sonuç 3.1 $c_k \neq 0, g_k \neq 0, k=1,2$ olmak üzere Teorem 3.2 ve Teorem 3.3' ün koşulları sağlandığında

$$l_{kj}^o = \left(\frac{3c_k^2}{2\pi^2 g_k} \right)^{1/3} n^{1/3} (1 + o(1)), j = 1, 3 \quad (3.11)$$

$$l_{kj}^o = \left(\frac{c_k^2}{\pi^2 g_k} \right)^{1/3} n^{1/3} (1 + o(1)), j = 2 \quad (3.12)$$

dır.

Bootstrap tahmin edicilerini elde etmek için HBB ve ÇBB yöntemleri kullanılırsa bu yöntemlerdeki optimal blok uzunlukları (3.11) ve (3.12) eşitliklerinden ÖBB yönteminde kullanılan optimal blok uzunluğunun $\left(\frac{3}{2}\right)^{1/3}$ katıdır.

Sonuç 3.1 in koşulları altında, optimal blok uzunlukları kullanılarak $\varphi_{1n} = \widehat{Yan}(\hat{\theta})$ ve $\varphi_{2n} = \widehat{Var}(\hat{\theta})$ ' in her üç yöntemle elde edilen blok bootstrap tahmin edicilerinin hata kareler ortalamaları

$$(a) \quad HKO(\widehat{Yan}_j(l_{1j}^o)) = 3^{1/3} [2\pi^2 g_1 c_1]^{2/3} n^{-8/3} + o(n^{-8/3}), j = 1, 3$$

$$HKO(\widehat{Yan}_j(l_{1j}^o)) = 3[\pi^2 g_1 c_1]^{2/3} n^{-8/3} + o(n^{-8/3}), j = 2$$

$$(b) \quad HKO(\widehat{Var}_j(l_{2j}^o)) = 3^{1/3} [2\pi^2 g_2 c_2]^{2/3} n^{-8/3} + o(n^{-8/3}), j = 1, 3$$

$$HKO(\widehat{Var}_j(l_{2j}^o)) = 3[\pi^2 g_2 c_2]^{2/3} n^{-8/3} + o(n^{-8/3}), j = 2$$

dır (Lahiri 1999). Her üç yöntem için ilgili HKO-optimal blok uzunlukları kullanıldığında,

$$HKO(\widehat{Yan}_j(l_{1j}^o)) = (2/3)^{2/3} HKO(\widehat{Yan}_2(l_{12}^o)), j = 1, 3$$

$$HKO(\widehat{Var}_j(l_{2j}^o)) = (2/3)^{2/3} HKO(\widehat{Var}_2(l_{22}^o)), j = 1, 3$$

olduğundan dolayı HBB ve ÇBB yöntemleriyle elde edilen tahmin edicilerin hata kareler ortalaması ÖBB yöntemiyle elde edilen tahmin edicilerin hata kareler ortalamasından daha küçüktür ve hata kareler ortalaması ölçütü bakımından HBB ve ÇBB yöntemleri ile elde edilen tahmin ediciler, ÖBB yöntemi ile elde edilen tahmin edicilere göre daha iyidir (Lahiri 1999).

3.6 Durağan Otoregresif Zaman Dizilerinde Bootstrap Yöntemi (ARB)

$(\varepsilon_i)_{i \in \mathbb{Z}}$ sıfır ortalamalı, birbirinden bağımsız aynı dağılımlı rasgele değişkenler dizisi olmak üzere $AR(p)$ serisi

$$X_i = \alpha_1 X_{i-1} + \alpha_2 X_{i-2} + \dots + \alpha_p X_{i-p} + \varepsilon_i \quad (3.13)$$

biçiminde verilmiş olsun. Bu model altında X_i rasgele değişkenleri bağımlı değişkenlerdir ve bu bölümde blok bootstrap yöntemlerinden farklı Bose (1988)' in bootstrap yönteminden bahsedilecektir.

$(X_i)_{i \in \mathbb{Z}}$ rasgele değişkenler dizisinden alınan bir örneklem X_1, \dots, X_n olsun. X_1, \dots, X_n örneklemini kullanarak $\alpha_1, \dots, \alpha_p$ parametrelerinin en küçük kareler tahmin edicileri $\hat{\alpha}_{1n}, \dots, \hat{\alpha}_{pn}$ olmak üzere

$$\left(\hat{\alpha}_{1n}, \dots, \hat{\alpha}_{pn} \right)' = (V_n' V_n)^{-1} V_n' (X_{p+1}, \dots, X_n)'$$

dır (Lahiri 2003). Burada V_n i -inci satırı (X_{i+p-1}, \dots, X_i) , $i = 1, \dots, n-p$ olan $(n-p) \times p$ boyutlu bir matristir.

$$e_i = X_i - \hat{\alpha}_{1n}X_{i-1} - \dots - \hat{\alpha}_{pn}X_{i-p}, \quad i = p+1, \dots, n$$

artıkları göstermek üzere (3.13) eşitliğinden artıklar,

$$e_i = \varepsilon_i - \sum_{j=1}^p (\hat{\alpha}_{jn} - \alpha_j) X_{i-j}, \quad p+1 \leq i \leq n \quad (3.14)$$

şeklinde ifade edilebilir. $j = 1, \dots, p$ için $\hat{\alpha}_{jn} \xrightarrow{p} \alpha_j$ yani $\hat{\alpha}_{jn}$, en küçük kareler tahmin edicileri α_j parametrelerinin tutarlı tahmin edicileri olduğu için (3.14) eşitliğinde bulunan $\sum_{j=1}^p (\hat{\alpha}_{jn} - \alpha_j) X_{i-j}$ terimi $n \rightarrow \infty$ iken küçük bir değere eşit olacağından artıklar yaklaşık olarak (approximately) bağımsızdır. Böylece bu yöntem bir bootstrap örnekleme oluşturulurken $n-p$ kez tek bir artığın yeniden örneklenmesine dayandığı için Efron (1979) tarafından önerilen bildik bootstrap yöntemine benzemektedir. Fakat geçerli bir yaklaşımda bulunabilmek için öncelikle artıkların merkezileştirilmesi gerekmektedir. Bunun için $\bar{e}_n = \frac{1}{n-p} \sum_{i=p+1}^n e_i$ artıkların ortalaması olmak üzere merkezileştirilmiş artıklar,

$$\tilde{e}_i = e_i - \bar{e}_n, \quad i = p+1, \dots, n \quad (3.15)$$

olacaktır. Bootstrap hata değişkenleri $\varepsilon_i^*, i \in \mathbb{Z}$ ' ları elde etmek için $\{\tilde{e}_{p+1}, \dots, \tilde{e}_n\}$ kümesinden yerine koyarak rasgele örnekleme yapılır. Böylece $\varepsilon_i^*, i \in \mathbb{Z}$ rasgele değişkenleri, X_1, \dots, X_n verildiğinde koşullu olarak birbirinden bağımsız ve

$$P_*(\varepsilon_i^* = \tilde{e}_i) = \frac{1}{n-p}, \quad p+1 \leq i \leq n \quad (3.16)$$

olasılık dağılımı ile aynı dağılıma sahip rasgele değişkenler olacaktır.

(3.15) ve (3.16) eşitliklerinden $E_*(\varepsilon_1^*) = \frac{1}{n-p} \sum_{i=p+1}^n \tilde{\varepsilon}_i = 0$ olup (3.13) eşitliğindeki model

için verilen $E(\varepsilon_1) = 0$ koşulu bu modelin bootstrap uyarlaması için de sağlanmış olur.

Böylece (3.13) eşitliğinde verilen modelin bootstrap uyarlaması

$$X_i^* = \hat{\alpha}_{1n} X_{i-1}^* + \dots + \hat{\alpha}_{pn} X_{i-p}^* + \varepsilon_i^*, \quad p+1 \leq i \leq n \quad (3.17)$$

olacaktır. Burada $(X_1^*, X_2^*, \dots, X_p^*) \equiv (X_1, X_2, \dots, X_p)$ olarak alınabilir (DasGupta 2008).

$\alpha_{1n}^*, \dots, \alpha_{pn}^*, \alpha_1, \dots, \alpha_p$ parametrelerinin (3.17) eşitliğindeki model kullanılarak elde edilen en küçük kareler tahmin edicileridir. Bu değerler $\hat{\alpha}_{1n}, \dots, \hat{\alpha}_{pn}$ değerlerinin bootstrap uyarlamasıdır.

4. UYGULAMA

Bu bölümde ilk olarak geçiş matrisi bilinen bir markov zinciri tanımlanmış ve bu markov zincirinden simülasyonla $n = 250$ adım sayısına sahip bir örneklem elde edilerek bu örnekleme dayalı geçiş matrisinin en çok olabilirlik tahmini ve farklı blok uzunlukları için HBB, ÖBB ve ÇBB tahminleri elde edilmiştir. Elde edilen sonuçlar Çizelge 4.1’ de verilmiştir. Daha sonra $\alpha_1 = 0.5$ için $X_i = \alpha_1 X_{i-1} + \varepsilon_i, i = 1, \dots, n$ zaman serisi modeline uygun olarak simülasyonla $n = 250$ adet veri üretilmiş ve α_1 parametresinin en küçük kareler tahmini ve bootstrap tahmini elde edilmiştir.

4.1 Sonlu Durum Uzayına Sahip Bir Markov Zincirinin Geçiş Olasılıkları

Matrisinin Blok Bootstrap Yöntemleriyle Tahmini

$\{X_n, n \geq 1\}$ stokastik süreci durum uzayı $S = \{1, 2, \dots, k\}, k \geq 2$ ve geçiş olasılıkları matrisi $P = (p_{ij})$ olan ergodik markov zinciri olsun. $\sum_{j=1}^k P(X_t = j | X_{t+1} = i) = 1, i, j \in S$ olması durumunda $P = (p_{ij})$ ’ nin en çok olabilirlik tahmin edicisi $\hat{P} = (\hat{p}_{ij}), i.$ durumdan $j.$ duruma geçişlerin sayısı n_{ij} ve $i.$ durumdan geçişlerin sayısı $n_i = \sum_j n_{ij}$ olmak üzere

$$\hat{p}_{ij} = \frac{n_{ij}}{n_i}$$

olacaktır (Basawa and Rao 1980).

$\{X_n; n \geq 1\}$ stokastik süreci $S = \{1, 2, 3\}$ durum uzayına sahip ve geçiş olasılıkları matrisi,

$$P = \begin{bmatrix} 0.25 & 0.40 & 0.35 \\ 0.30 & 0.50 & 0.20 \\ 0.10 & 0.30 & 0.60 \end{bmatrix}$$

olan bir markov zinciri olsun (Duman 2006). Bu markov zincirinden, yapılan simülasyon çalışması (EK 1) sonucunda elde edilen $n = 250$ adım sayısına sahip bir örneklem Çizelge 4.1' de verilmiştir .

Çizelge 4.1 Markov zincirinin $n = 250$ adım sayısına sahip bir gerçekleşmesi

2 2 1 2 2 1 2 1 2 1 3 2 3 2 3 3 3 2 1 2 2 2 2 2 3 2 3 3 1 2 2 3 1 3 3 1 2 2 1 2 1 2 1 3 3 3 3
3 1 1 2 2 2 2 3 3 3 3 3 2 2 1 2 2 2 3 3 2 2 1 3 3 3 3 2 2 2 3 3 3 3 2 2 2 2 2 3 3 3 3 3 3
3 2 2 2 3 3 1 3 3 3 2 2 2 2 1 1 3 2 3 3 3 3 2 1 3 2 2 2 2 1 2 3 3 2 2 3 3 3 3 3 3 3 3 3 3
2 2 1 3 2 2 2 1 1 1 3 3 3 3 1 3 2 2 1 1 2 1 2 2 1 1 2 2 3 2 2 2 1 2 2 1 1 3 3 2 3 3 3 2 2 2 3
3 1 1 3 2 1 3 3 3 2 2 2 2 1 2 1 3 2 1 1 2 3 3 3 2 1 1 3 2 2 1 3 3 3 2 2 2 2 3 2 3 3 1 2 3 3 2
3 3 1 1 1 3 3 2 2 1 2 1 2 2 2

Yukarıda elde edilen veri setine ait P geçiş olasılıkları matrisinin en çok olabilirlik tahmini, \hat{P}_n ,

$$\hat{P}_n = \begin{bmatrix} 0.2449 & 0.4082 & 0.3469 \\ 0.2745 & 0.5098 & 0.2157 \\ 0.0909 & 0.2929 & 0.6162 \end{bmatrix}$$

olacaktır.

Geçiş olasılıkları matrisinin blok bootstrap yöntemleri ile elde edilen tahminlerini bulmak için izlenen adımlar aşağıdaki gibidir. Kolaylık açısından l blok uzunluğu n örneklem çapını tam bölen bir sayı ve seçilen blok sayısı k , $k.l = n$ olacak şekilde belirlensin.

Hareketli bloklarla bootstrap yöntemi için,

Adım 1. $B_i^{(1)} = (X_i, \dots, X_{i+l-1})$, $1 \leq i \leq N$, $N = n - l + 1$ olacak şekilde N tane blok oluşturulur.

Adım 2. Oluşturulan bloklar arasından yerine koyarak rasgele k tane blok seçilir ve $k.l = n$ çaplı $X_{1,1}^*, \dots, X_{1,n}^*$ HBB örnekleme oluşturulur.

Adım 3. B yineleme sayısı olmak üzere Adım 2 B defa tekrarlanır ve B tane n çaplı HBB örnekleme elde edilir. Her bir \underline{X}_1^{*s} , $s = 1, \dots, B$ HBB örnekleme için

$$p_{ij}^{*s(1)} = \hat{p}_{ij}(\underline{X}_1^{*s})$$

değeri hesaplanır.

Adım 4. \hat{p}_{ij} 'nin HBB tahmini,

$$p_{ij}^{*(1)} = \frac{1}{B} \sum_{s=1}^B p_{ij}^{*s(1)}$$

olacaktır.

Örtüşmeyen bloklarla bootstrap yöntemi için,

Adım 1. $B_i^{(2)} = (X_{(i-1)l+1}, \dots, X_{il}), 1 \leq i \leq b$ olacak şekilde $b = \frac{n}{l}$ tane blok oluşturulur.

Adım 2. Oluşturulan bloklar arasından yerine koyarak rasgele k tane blok seçilir ve $k.l = n$ çaplı $X_{2,1}^*, \dots, X_{2,n}^*$ ÖBB örnekleme oluşturulur.

Adım 3. B yinleme sayısı olmak üzere Adım 2 B defa tekrarlanır ve B tane n çaplı ÖBB örnekleme elde edilir. Her bir $\underline{X}_2^{*s}, s = 1, \dots, B$ ÖBB örnekleme için

$$p_{ij}^{*s(2)} = \hat{p}_{ij}(\underline{X}_2^{*s})$$

değeri hesaplanır.

Adım 4. \hat{p}_{ij} 'nin ÖBB tahmini

$$p_{ij}^{*(2)} = \frac{1}{B} \sum_{s=1}^B p_{ij}^{*s(2)}$$

olacaktır.

Çembersel bloklarla bootstrap yöntemi için,

Adım 1. Her bir X_i gözlemi $Y_{n,i}$ değişkeni ile etiketlenilerek $1 \leq i \leq n$ için $B(i, l) = (Y_{n,i}, \dots, Y_{n,i+l-1})$ olacak şekilde n tane blok oluşturulur.

Adım 2. Oluşturulan bloklar arasından yerine koyarak rasgele k tane blok seçilir ve $k.l = n$ çaplı $X_{3,1}^*, \dots, X_{3,n}^*$ ÇBB örnekleme oluşturulur.

Adım 3. B yineleme sayısı olmak üzere Adım 2 B defa tekrarlanır ve B tane n çaplı DBB örnekleme elde edilir. Her bir $\underline{X}_3^{*s}, s = 1, \dots, B$ ÇBB örnekleme için

$$p_{ij}^{*s(3)} = \hat{p}_{ij}(\underline{X}_3^{*s})$$

değeri hesaplanır.

Adım 4. \hat{p}_{ij} 'nin ÖBB tahmini

$$p_{ij}^{*(3)} = \frac{1}{B} \sum_{s=1}^B p_{ij}^{*s(3)}$$

olacaktır.

Yukarıdaki algoritmalara uygun olarak yazılan Matlab bilgisayar programı kodları (EK 1) ve Çizelge 4.2' de verilen N, k, l, b değerleri kullanılarak yapılan simülasyon çalışması sonucunda $B = 200$ yineleme ve $n = 250$ örneklem çapı için geçiş olasılıkları matrisinin blok bootstrap yöntemleri ile elde edilen tahminleri Çizelge 4.3' de verilmiştir.

Çizelge 4.2 $l = 2, 5, 10, 25, 50$ blok uzunlukları için geçiş olasılıkları matrisinin HBB, ÖBB ve ÇBB yöntemleri için oluşturulan ve seçilen blok sayıları

	<u>$l = 2$</u>		
	<u>HBB</u>	<u>ÖBB</u>	<u>ÇBB</u>
Oluşturulan blok sayısı:	$N = 249$	$b = 125$	$n = 250$
Seçilen blok sayısı:	$k = 125$	$k = 125$	$k = 125$
	<u>$l = 5$</u>		
	<u>HBB</u>	<u>ÖBB</u>	<u>ÇBB</u>
Oluşturulan blok sayısı:	$N = 246$	$b = 50$	$n = 250$
Seçilen blok sayısı:	$k = 50$	$k = 50$	$k = 50$
	<u>$l = 10$</u>		
	<u>HBB</u>	<u>ÖBB</u>	<u>ÇBB</u>
Oluşturulan blok sayısı:	$N = 241$	$b = 25$	$n = 250$
Seçilen blok sayısı:	$k = 25$	$k = 25$	$k = 25$
	<u>$l = 25$</u>		
	<u>HBB</u>	<u>ÖBB</u>	<u>ÇBB</u>
Oluşturulan blok sayısı:	$N = 225$	$b = 10$	$n = 250$
Seçilen blok sayısı:	$k = 10$	$k = 10$	$k = 10$
	<u>$l = 50$</u>		
	<u>HBB</u>	<u>ÖBB</u>	<u>ÇBB</u>
Oluşturulan blok sayısı:	$N = 201$	$b = 5$	$n = 250$
Seçilen blok sayısı:	$k = 5$	$k = 5$	$k = 5$

Çizelge 4.3 $l = 2, 5, 10, 25, 50$ blok uzunlukları için geçiş olasılıkları matrisinin HBB, ÖBB ve ÇBB yöntemleri ile elde edilen tahminleri

	<u>$l = 2$</u>	<u>$l = 5$</u>
$HBB : P^{*(1)} =$	$\begin{bmatrix} 0.2201 & 0.4099 & 0.3644 \\ 0.2324 & 0.4528 & 0.3095 \\ 0.1474 & 0.3460 & 0.5027 \end{bmatrix}$	$P^{*(1)} =$ $\begin{bmatrix} 0.2380 & 0.3984 & 0.3587 \\ 0.2627 & 0.4835 & 0.2500 \\ 0.1123 & 0.3155 & 0.5683 \end{bmatrix}$
$ÖBB : P^{*(2)} =$	$\begin{bmatrix} 0.2146 & 0.4366 & 0.3452 \\ 0.2454 & 0.4389 & 0.3120 \\ 0.1351 & 0.3658 & 0.4945 \end{bmatrix}$	$P^{*(2)} =$ $\begin{bmatrix} 0.2159 & 0.4505 & 0.3293 \\ 0.2850 & 0.4672 & 0.2444 \\ 0.0960 & 0.3246 & 0.5752 \end{bmatrix}$
$ÇBB : P^{*(3)} =$	$\begin{bmatrix} 0.2201 & 0.4016 & 0.3735 \\ 0.2324 & 0.4596 & 0.3040 \\ 0.1442 & 0.3562 & 0.4958 \end{bmatrix}$	$P^{*(3)} =$ $\begin{bmatrix} 0.2365 & 0.4071 & 0.3515 \\ 0.2540 & 0.4989 & 0.2422 \\ 0.1139 & 0.3137 & 0.5689 \end{bmatrix}$
	<u>$l = 10$</u>	<u>$l = 25$</u>
$HBB : P^{*(1)} =$	$\begin{bmatrix} 0.2445 & 0.3806 & 0.3707 \\ 0.2619 & 0.4985 & 0.2361 \\ 0.1022 & 0.3038 & 0.5897 \end{bmatrix}$	$P^{*(1)} =$ $\begin{bmatrix} 0.2504 & 0.3863 & 0.3592 \\ 0.2533 & 0.5203 & 0.2230 \\ 0.0918 & 0.2903 & 0.6123 \end{bmatrix}$
$ÖBB : P^{*(2)} =$	$\begin{bmatrix} 0.2168 & 0.4447 & 0.3331 \\ 0.2784 & 0.4811 & 0.2380 \\ 0.0926 & 0.3140 & 0.5887 \end{bmatrix}$	$P^{*(2)} =$ $\begin{bmatrix} 0.2359 & 0.4129 & 0.3479 \\ 0.2880 & 0.5051 & 0.2010 \\ 0.0871 & 0.2944 & 0.6161 \end{bmatrix}$
$ÇBB : P^{*(3)} =$	$\begin{bmatrix} 0.2339 & 0.4089 & 0.3539 \\ 0.2611 & 0.5102 & 0.2244 \\ 0.1014 & 0.3058 & 0.5889 \end{bmatrix}$	$P^{*(3)} =$ $\begin{bmatrix} 0.2364 & 0.4036 & 0.3553 \\ 0.2648 & 0.5160 & 0.2157 \\ 0.0943 & 0.2968 & 0.6047 \end{bmatrix}$
	<u>$l = 50$</u>	
$HBB : P^{*(1)} =$	$\begin{bmatrix} 0.2634 & 0.3773 & 0.3554 \\ 0.2710 & 0.5248 & 0.3854 \\ 0.0821 & 0.2986 & 0.6154 \end{bmatrix}$	
$ÖBB : P^{*(2)} =$	$\begin{bmatrix} 0.2424 & 0.3973 & 0.3520 \\ 0.2891 & 0.5024 & 0.2046 \\ 0.0898 & 0.3050 & 0.6034 \end{bmatrix}$	
$ÇBB : P^{*(3)} =$	$\begin{bmatrix} 0.2424 & 0.4056 & 0.3473 \\ 0.2744 & 0.5163 & 0.2053 \\ 0.0924 & 0.2929 & 0.6112 \end{bmatrix}$	

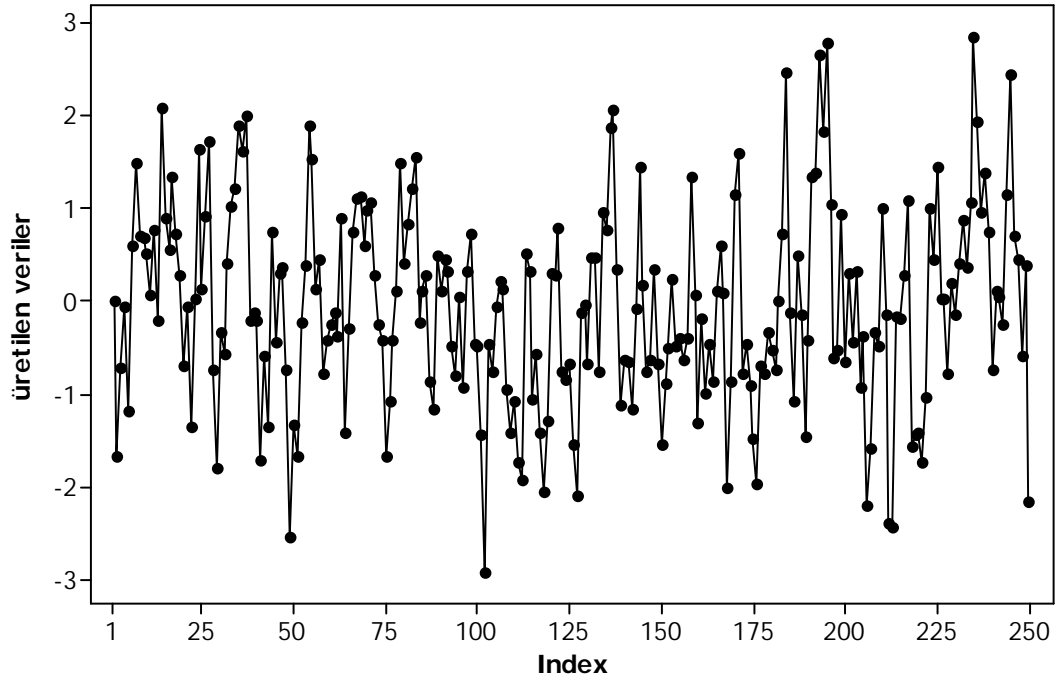
4.2 Birinci Dereceden Otoregresif Zaman Serisinin α_1 Parametresinin Bootstrap Tahmini

$(X_i)_{i \in \mathbb{Z}}$, $X_i = \alpha_1 X_{i-1} + \varepsilon_i, i = 1, 2, \dots, n, |\alpha_1| < 1$ durağan zaman serisi modeline (AR(1)) uygun rasgele değişkenlerin bir dizisi olsun. α_1 parametresinin en küçük kareler tahmin edicisi,

$$\hat{\alpha}_{1n} = \frac{\sum_{i=1}^n X_i X_{i-1}}{\sum_{i=1}^n X_{i-1}^2}$$

olup $\hat{\alpha}_{1n} \xrightarrow{p} \alpha_1$ dir (Akdi 2003). Yani $\hat{\alpha}_{1n}$, en küçük kareler tahmin edicisi, α_1 ' in tutarlı bir tahmin edicisidir.

$\varepsilon_i \sim b.b.a.d N(0,1)$ olmak üzere, $\alpha_1 = 0.5$ için $X_i = \alpha_1 X_{i-1} + \varepsilon_i, i = 1, \dots, n$ zaman serisi modeline uygun olarak $X_0 = 0$ başlangıç koşulu altında simülasyon çalışması sonucunda üretilen $n = 250$ tane veri EK 3' de verilmiştir. Bu veriler kullanılarak çizdirilen zaman serisi grafiği Şekil 4.1. deki gibidir.



Şekil 4.1 Üretilen veriler için zaman serisi grafiği

α_1 parametresinin ARB yöntemi ile bootstrap tahminini elde etmek için aşağıdaki adımlar izlenir.

Adım 1. $\hat{\alpha}_{1n}$, en küçük kareler tahmin edicisi hesaplanır.

Adım 2. $e_i = X_i - \hat{\alpha}_{1n}X_{i-1}$, $i = 2, \dots, n$ AR(1) modelinin artıklarını göstermek üzere

$$\tilde{e}_i = e_i - \frac{1}{n-1} \sum_{i=2}^n e_i, i = 2, \dots, n \text{ merkezileştirilmiş artıklar hesaplanır.}$$

Adım 3. Bootstrap hata değişkenleri ε_i^* ' ları elde etmek için $\{\tilde{e}_2, \dots, \tilde{e}_n\}$ kümesinden yerine koyarak rasgele örnekleme yapılır. $X_1^* = X_1$ koşulu altında

$$X_i^* = \hat{\alpha}_{1n}X_{i-1}^* + \varepsilon_i^*, 2 \leq i \leq n$$

modeline uygun $X_1^*, X_2^*, \dots, X_n^*$ bootstrap örnekleme oluşturulur.

Adım 4. B yineleme sayısı olmak üzere Adım 3, B defa tekrarlanır ve B tane n çaplı bootstrap örnekleme elde edilir. Her bir $\underline{X}^{*b}, b=1, \dots, B$ bootstrap örnekleme için $\alpha_{1n}^{*b} = \hat{\alpha}_1(\underline{X}^{*b})$ değeri hesaplanır.

Adım 5. α_1 parametresinin bootstrap tahmini,

$$\alpha_{1n}^* = \frac{1}{B} \sum_{b=1}^B \alpha_{1n}^{*b}$$

olarak hesaplanır.

Yukarıdaki algoritmaya uygun olarak yazılan Matlab bilgisayar programı kodları (EK 2) sonucunda $B = 200$ yineleme ve $n = 250$ örneklem çapı için α_1 parametresinin, en küçük kareler tahmini ve bootstrap tahmini sırasıyla,

$$\hat{\alpha}_{1n} = 0.498953185434769$$

$$\alpha_{1n}^* = 0.499228537939652$$

olarak bulunmuştur.

α_1 parametresinin $B=200$ yineleme ve $l = 2, 5, 10, 25, 50$ blok uzunlukları için HBB, ÖBB ve ÇBB yöntemleriyle elde edilen tahminleri Çizelge 4.4' de verilmiştir.

Çizelge 4.4 $l = 2, 5, 10, 25, 50$ blok uzunlukları için HBB, ÖBB, ÇBB tahminleri

	$l = 2$	$l = 5$	$l = 10$	$l = 25$	$l = 50$
HBB:	0.22812	0.37953	0.41953	0.44924	0.47593
ÖBB:	0.25794	0.39748	0.44130	0.46639	0.46873
ÇBB:	0.23294	0.37790	0.42959	0.45205	0.45383

Çizelge 4.4 incelendiğinde, blok bootstrap yöntemleriyle elde edilen tahmin değerlerinin blok uzunluğu arttıkça $\alpha_1 = 0.5$ değerine yaklaştığı görülmektedir. ARB yönteminden elde edilen tahmin değeri ile blok bootstrap yöntemleri kullanılarak elde edilen tahmin değerleri karşılaştırıldığında, ARB tahmin değeri gerçek parametre değerine daha yakın sonuç vermektedir. Bunun sebebi ARB yönteminin asıl modelin yapısını bozmadan model parametresini tahmin etmesidir. Blok bootstrap yöntemlerinde ise bootstrap örneklemeleri oluşturulurken gözlemler l uzunluklu bloklara ayrılıp, bu bloklar arasından rasgele bloklar seçildiğinden bootstrap örneklemeleri asıl modelin yapısını yansıtmamaktadır.

5. TARTIŞMA ve SONUÇ

Gözlemlerin “kısa dönemli” bağımlılık yapısına sahip olması durumunda bootstrap metodunun geçerliliğinin araştırılması bu çalışmanın esas amacıdır, bu nedenle önce bazı önemli tanım ve kavramlar irdelenmiştir. Daha sonra üçüncü bölümde gözlemlerin bağımlı olması durumunda kullanılan blok bootstrap yöntemlerinden, hareketli bloklarla bootstrap, örtüşmeyen bloklarla bootstrap ve çembersel bloklarla bootstrap yöntemleri ve bu yöntemlerle bulunan tahmin edicilerin tutarlılığı konularına değinilmiştir.

Blok bootstrap yöntemleri, $(X_i)_{i \in \mathbb{Z}}$ rasgele değişkenler dizisi durağan ve “kısa dönemli” bağımlılık yapısına sahip olduğundan gözlemler dizisi alındığı sıra ile dizideki bağımlılık yapısını yakalayabilen uygun uzunluklu bloklara ayrılırsa bu blokların “yaklaşık olarak” birbirinden bağımsız ve farklı bloklardaki gözlemlerin ortak dağılımının durağanlık koşulu altında aynı olması fikrine dayanır. Bu nedenle blok bootstrap yöntemleri, Efron (1979) tarafından önerilen bildik bootstrap yöntemiyle aynı temel düşünceye sahiptir.

Uygulama bölümünde ise ilk olarak geçiş matrisi bilinen bir markov zinciri tanımlanmış ve bu markov zincirinden simülasyonla $n = 250$ adım sayısına sahip bir örneklem elde edilerek bu örnekleme dayalı geçiş matrisinin en çok olabilirlik tahmini ve $l = 2, 5, 10, 25, 50$ blok uzunlukları için HBB, ÖBB ve ÇBB tahminleri elde edilmiştir. Elde edilen sonuçlar Çizelge 4.1’ de verilmiştir. Elde edilen sonuçlara bakıldığında blok uzunluğu arttırıldıkça özellikle blok uzunluğu $l > 5$ için daha iyi sonuçlar elde edildiği görülmektedir. Bunun sebebi blok uzunluğu arttırıldıkça elde edilen blok bootstrap örnekleme ile başlangıçta oluşturulan markov zincirinin bağımlılık yapısının daha iyi yansıtılabilmesidir. Böylelikle geçiş matrisinin bootstrap tahmini için, Basawa *et al* (1990)’ da bahsedilen karmaşık algoritmaya kıyasla daha basit bir algoritmaya sahip olan blok bootstrap yöntemlerinin de kullanılabilirdiği gösterilmiştir.

Uygulama bölümünün ikinci kısmında $\varepsilon_i \sim b.b.a.d N(0,1)$ olmak üzere $X_i = \alpha_1 X_{i-1} + \varepsilon_i, i = 1, \dots, n$ zaman serisi modeli için α_1 parametresinin bootstrap tahmininin nasıl yapılacağına ilişkin bir örnek verilmiştir. Buna göre $\alpha_1 = 0.5$ için modele uygun olarak simülasyonla $n = 250$ adet veri üretilmiş ve α_1 parametresinin ARB yöntemi ile bootstrap tahmini elde edilmiştir. Aynı model için blok bootstrap yöntemleri ile model parametresi tahmin edilmiş sonuçlar Çizelge 4.4' de verilmiştir.

Blok bootstrap yöntemleri kullanılarak oluşturulan bootstrap örneklemelerinin durağan olmaması sorunu ile karşılaşılabilir. Politis and Romano (1994) bu sorunu ortadan kaldırmak amacı ile durağan bootstrap yöntemini önermişlerdir.

KAYNAKLAR

- Akdi, Y. 2003. Zaman serileri analizi (birim kökler ve kointegrasyon). Bıçaklar, 300 s. Ankara.
- Athreya, K.B. and Pantula, S. G. 1986. Mixing properties of haris chains and autoregressive processes. Applied Probability Trust 23 (4); 880-892.
- Athreya, K. B. and Lahiri, S. N. 2006. Measure theory and probability theory. Springer, 612p., New York.
- Basawa, V. I. and Prakasa-Rao, B. L. S. 1980. Statistical Inference for Stochastic Processes. Academic Press London.
- Bilingsley, P. 1994. Probability and measure, third edition. Wiley, 593 p., New York.
- Bose, A. 1988. Edgeworth correction by bootstrap in autoregressions. Annals of Statistics 16 (4);1709-1722.
- Carlstein, E. 1986. The use of subseries methods for estimating the varience of a general statistic from a stationary time series. Annals of Statistics 14 (3); 1171-1179
- Chanda, K. C. 1974. Strong mixing properties of linear stochastic processes. Journal of Applied Probability 11 (2); 401-408.
- DasGupta, A. 2008. Asymptotic theory of statistics and probability. Springer, 722p., New York.
- Doukhan, P. 1995. Mixing properties and examples. Springer, 146 p., New York.
- Doukhan, P. 1999. A new weak dependence condition and applications to moment inequalities. Stochastic Processes and their Applications 84, 313-342.
- Duman, S. 2006. Markov zincirlerinde bootstrap. Yüksek lisans tezi. Ankara Üniversitesi, 54 s., Ankara.

- Efron, B. 1979. Bootstrap methods; another look at the jackknife. *Annals of Statistics*, 7 (1); 1-26.
- Gilat, A. 2008. *An introduction with applications*. Third edition. Wiley, 384p., New York.
- Gorodetskii, V. V. 1977. On the strong mixing property for linear sequences. *Theory Probab. Appl.* 22; 411-413.
- Hall, P. 1992. The Bootstrap and Edgeworth Expansion. *Journal of the Royal Statistical Society*, 156 (3); 504-505.
- Hall, P., Horowitz, J. L. and Jing, B.-Y. 1995. On blocking rules for the bootstrap with dependent data. *Biometrika*, 82 (3); 561-574.
- Künsch, H. R. 1989. The jackknife and the bootstrap for general stationary observations. *Annals of Statistics*, 17 (3); 1217-1261.
- Lahiri, S. N. 1999. Theoretical comparisons of block bootstrap methods. *Annals of Statistics*. 27 (1); 386-404.
- Lahiri, S. N. 2003. *Resampling methods for dependent data*. Springer, 374p., New York.
- Lehmann, E. L. 1999. *Elements of large-sample theory*. Springer, 631p., New York.
- Liu, R. Y. and Singh, K. 1992. Moving blocks jackknife and bootstrap capture weak dependence, in *Exploring the Limits of Bootstrap*, (Lepage, R. and Billard, L. eds.). Wiley, pp. 225-248. New York.
- Politis, D. N. and Romano, J. P. 1992. A circular block resampling procedure for stationary data, in *Exploring the Limits of Bootstrap*, (Lepage, R. and Billard, L. eds.). Wiley, pp. 263-270. New York.
- Politis, D. N. and Romano, J. P. 1994. The stationary bootstrap. *Journal of the American Statistical Association*. 89; 1303-1313.

- Politis, D. N. and White, H. 2004. Automatic block-length selection for the dependent bootstrap. *Econ. Rev.* 23 (1); 53-70.
- Serfling, R. J. 1980. *Approximation theorems of mathematical statistics*. Wiley, 371p., New York.
- Singh, K. 1981. On the asymptotic accuracy of Efron' s bootstrap. *Annals of Statistics*, 9 (6); 1187-1195.
- Whithers, C. S. 1981. Conditions for linear processes to be strong-mixing. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*. 57 (4); 477-480

EKLER

EK 1 Bölüm 4.1 İçin Yazılan Matlab Bilgisayar Programları Kodları.....	64
EK 2 Bölüm 4.2 için Yazılan Matlab Bilgisayar Programı Kodları.....	72
EK 3 $X_i = 0.5X_{i-1} + \varepsilon_i, i = 2, \dots, 250$ zaman serisi modeline uygun üretilen veriler.....	74

EK 1 Bölüm 4.1 İçin Yazılan Matlab Bilgisayar Programı Kodları

Bu bölümde kullanılan kodlar için Gilat (2008) kaynak olarak gösterilebilir.

```
clc
clear all
close all
%markov zinciri oluşturma
P=[0.25 0.40 0.35;0.30 0.50 0.20;0.10 0.30 0.60]
nn=size(P,1);durum=1
for n=1:250
    i=durum;a=rand(1,1);durum=1;
    for j=1:(nn-1)
        if a>sum(P(i,1:j))
            durum=j+1;
        end
    end
    zincir(n)=durum;
    zincir
y=zincir'
end

% markov zincirinin en çok olabilirlik tahmin edicisini bulma

function gecis=mrkov(zincir)
zincir=zincir';
[n,mmmm]=size(zincir);
y=zincir;
%geçiş olasılıkları matrisinin en çok olabilirlik tahmin edicisi
for i=1:n
    for j=1:3
        if zincir(i)==1
            ni(i,1)=1;
        else
            ni(i,1)=0;
        end
        if zincir(i)==2
            ni(i,2)=1;
        else
            ni(i,2)=0;
        end
        if zincir(i)==3
            ni(i,3)=1;
        else
            ni(i,3)=0;
        end
    end
end
end
```

EK 1 Bölüm 4.1 İçin Yazılan Matlab Bilgisayar Programı Kodları (Devam)

```
i=0;
for j=1:n-1
i=j+1;

    if y(i)==1 & y(j)==1
        yy1(i)=1;
    else
        yy1(i)=0
    end
end
i=0;
for j=1:n-1
i=j+1;

    if y(i)==2 & y(j)==1
        yy12(i)=1;

    else
        yy12(i)=0;
    end
end
i=0;
for j=1:n-1
i=j+1;

    if y(i)==3 & y(j)==1
        yy13(i)=1;
    else
        yy13(i)=0;

    end
end

yy1dengecis=[yy1' yy12' yy13'];
yy1dengecis(1,1)=0;
yy1dengecis(1,2)=0;
yy1dengecis(1,3)=0;
%=====

i=0;
for j=1:n-1
i=j+1;

    if y(i)==1 & y(j)==2
```

EK 1 Bölüm 4.1 İçin Yazılan Matlab Bilgisayar Programı Kodları (Devam)

```
yy2(i)=1;
else
    yy2(i)=0;
end
end
i=0;
for j=1:n-1
i=j+1;

    if y(i)==2 & y(j)==2
        yy22(i)=1;

    else
        yy22(i)=0;
    end
end
i=0;
for j=1:n-1
i=j+1
if y(i)==3 & y(j)==2
    yy23(i)=1;
else
    yy23(i)=0;

end
end

yy2dengecis=[yy2' yy22' yy23'];
yy2dengecis(1,1)=0;
yy2dengecis(1,2)=0;
yy2dengecis(1,3)=0;

%=====

i=0;
for j=1:n-1
i=j+1;

    if y(i)==1 & y(j)==3
        yy3(i)=1;
    else
        yy3(i)=0;
    end
end
end
```

EK 1 Bölüm 4.1 İçin Yazılan Matlab Bilgisayar Programı Kodları (Devam)

```
i=0;
for j=1:n-1
i=j+1;

    if y(i)==2 & y(j)==3
        yy32(i)=1;

    else
        yy32(i)=0;
    end
end
i=0;
for j=1:n-1
i=j+1;

    if y(i)==3 & y(j)==3
        yy33(i)=1;
    else
        yy33(i)=0;
    end
end

yy3dengecis=[yy3' yy32' yy33'];
yy3dengecis(1,1)=0;
yy3dengecis(1,2)=0;
yy3dengecis(1,3)=0;

toplamni=sum(ni);
toplamy1dengec=sum(yy1dengecis);
toplamy2dengec=sum(yy2dengecis);
toplamy3dengec=sum(yy3dengecis);

oran11=toplamy1dengec/toplamni(1);
oran12=toplamy2dengec/toplamni(2);
oran13=toplamy3dengec/toplamni(3);

gecis=[oran11; oran12; oran13];
end
```

EK 1 Bölüm 4.1 İçin Yazılan Matlab Bilgisayar Programı Kodları (Devam)

% en çok olabilirlik tahmin edicisinin blok bootstrap yöntemleri ile elde edilen tahminlerini veren program

```
clc
clear all
close all
% HBB Yöntemi
l=input('blok uzunluğunu giriniz');
%c=y';
zincir=[2 2 1 2 2 1 2 1 2 1 3 2 3 2 3 3 3 2 1 2 2 2 2 3 2 3 3 1 2 2 3 1 3 3 1 2 2 1 2 1 2 1
3 3 3 3 3 1 1 2 2 2 2 3 3 3 3 3 2 2 1 2 2 2 3 3 2 2 1 3 3 3 3 2 2 2 3 3 3 3 2 2 2 2 2 2 3 3 3
3 3 3 3 3 2 2 2 3 3 1 3 3 3 2 2 2 2 1 1 3 2 3 3 3 3 2 1 3 2 2 2 2 1 2 3 3 2 2 3 3 3 3 3 3 3 3
3 3 3 3 2 2 1 3 2 2 2 1 1 1 3 3 3 3 1 3 2 2 1 1 2 1 2 2 1 1 2 2 3 2 2 2 1 2 2 1 1 3 3 2 3 3 3
2 2 2 3 3 1 1 3 2 1 3 3 3 2 2 2 2 1 2 1 3 2 1 1 2 3 3 3 2 1 1 3 2 2 1 3 3 3 2 2 2 2 3 2 3 3 1
2 3 3 2 3 3 1 1 1 3 3 2 2 1 2 1 2 2 2];
c=zincir;
%n=length(c);
%l=5;
%for i=1:l-1
% c(i+n)=c(i)
%end

N=250-l+1;%n-l+1
k=input('kaç tane blok seçeceksiniz');
m=k*l;
for i=1:N
    for j=i:i+l-1
        blok(i,j)=c(j);
    end
end
% sadece bloklardaki elemanları göster
artim=0;
for i=1:N
    for j=1:l%j=1:l
        bb(i,j)=blok(i,artim+j);
    end
    artim=artim+l;
end
% elimizde N tane blok var bunların arasından rasgele k tane blok seçilecek
% bu işlem yapılırken 1 ile N arasındaki kesikli düzgün dağılımdan sayı
% üretilerek üretien sayıya karşılık gelen bloklar alınabilir bu işlem
% simülasyon sayısı (200) kadar tekrarlanacak

for simulasyon=1:200
    secilenbloklar(simulasyon,:)=unidrnd(N,k,1);
    % böylelikle k tane blok seçildi
```

EK 1 Bölüm 4.1 İçin Yazılan Matlab Bilgisayar Programı Kodları (Devam)

```
% bu bloklardaki gözlemler yerine yazılacak
for i=1:k% satırları göstereceğiz

orneklem(simulasyon,i,:)=bb((secilenbloklar(simulasyon,i)),:);
    %end
end

end
v=1;j=1;
for i=1:simulasyon
    %for j=1 : k*1
        for t=1:k
            for z=1:l
                sonornek(i,v)=orneklem(i,t,z);

                v=v+1;
            end

            %end
        end
        zzincir=sonornek(i,:);
        %zzincir=zzincir';
        mle(i,,:)=mrkov(zzincir);
        v=1;
    end

    %ÖBB için
    cc=c;
    %l=5;

    k2=input('kac tane bloko sececeksiniz');
    b=k2;
    for i=1:b
        for j=(i-1)*l+1:i*l
            bloko(i,j)=cc(j);
        end
    end
    %sadece bloklardaki elemanları göster
    artimm=0;
    for i=1:b
        for j=1:l%j=1:l
            bbo(i,j)=bloko(i,artimm+j);
        end
        artimm=artimm+l;
    end
end
```

EK 1 Bölüm 4.1 İçin Yazılan Matlab Bilgisayar Programı Kodları (Devam)

```
% elimizde N tane blok var bunların arasından rasgele k tane blok seçilecek
% bu işlem yapılırken 1 ile N arasındaki kesikli düzgün dağılımdan sayı
% üretilerek üretien sayıya karşılık gelen bloklar alınabilir bu işlem
% simülasyon sayısı (200) kadar tekrarlanacak
for simulasyono=1:200
    secilenbloklaro(simulasyono,:)=unidrnd(b,k2,1);
    %böylelikle k2 tane blok seçildi bu bloklardaki gözlemler yerine yazılacak
    for i=1:k2% satırları göstereceksin

        orneklemo(simulasyono,i,:)=bbo((secilenbloklaro(simulasyono,i)),:);
        %end
    end
end

v=1;j=1;
for i=1:simulasyono
    %for j=1 : k*1
        for t=1:k2
            for z=1:l
                sonorneko(i,v)=orneklemo(i,t,z);

                v=v+1;
            end

            %end
            end
            zzinciro=sonorneko(i,:);
            %zzincir=zzincir';
            mleo(i,,:)=mrkov(zzinciro);
        v=1;
    end
    %Geçiş matrisinin en çok olabilirlik tahmin edicisinin DBB şekli
    ccc=cc;
    n=length(ccc);
    %l=5;
    k3=input('kac tane db bloğu sececeksiniz');
    for i=1:l-1
        ccc(i+n)=ccc(i);
    end
    for i=1:n
        for j=i:i+1-1
            blokd(i,j)=ccc(j);
        end
    end
end
%sadece blokları göster
```

EK 1 Bölüm 4.1 İçin Yazılan Matlab Bilgisayar Programı Kodları (Devam)

```
artimmm=0;
for i=1:n
    for j=1:l%j=1:l
        bbd(i,j)=blokd(i,artimmm+j);
    end
    artimmm=artimmm+1;
end
% elimizde N tane blok var bunların arasından rasgele k tane blok seçilecek
% bu işlem yapılırken 1 ile N arasındaki kesikli düzgün dağılımdan sayı
% üretilerek üretien sayıya karşılık gelen bloklar alınabilir bu işlem
% simülasyon sayısı (200)kadar tekrarlanacak

for simulasyond=1:200
    secilenbloklard(simulasyond,:)=unidrnd(n,k3,1);
    %böylelikle k tane blok seçildi
    % bu bloklardaki gözlemler yerine yazılacak
    for i=1:k% satırları göstereceksin

    orneklemd(simulasyond,i,:)=bbd((secilenbloklard(simulasyond,i),:));
        %end
    end
end
end
v=1;j=1;
for i=1:simulasyond
    %for j=1 : k*1
        for t=1:k3
            for z=1:l
                sonornekd(i,v)=orneklemd(i,t,z);

                v=v+1;
            end

            %end
            end
            zzincird=sonornekd(i,:);
            %zzincir=zzincir';
            mled(i,,:)=mrkov(zzincird);
        v=1;
    end
```

EK 2 Bölüm 4.2 için Yazılan Matlab Bilgisayar Programı Kodları

```
clear all
clc
close all
alfa=0.5;
en=normrnd(0,1,250,1);
toplaml=0;
%modelden sayı üretme
x(1)=0;
for i=2:250
    x(i)=alfa*x(i-1)+en(i);
end
x;
b=x';
%a1 katsayısının ekk tahmin edicisi
for i=2:250
    carpim(i)=x(i)*x(i-1);
toplaml=toplaml+carpim(i);
end
toplam2=0;
for i=2:250
    toplam2=toplam2+x(i-1)^2;
end
ekk=toplam1/toplam2;
for i=2:250
    ensapka(i)=x(i)-(ekk*x(i-1));
end
for i=2:250
    enstd(i)=ensapka(i)-mean(ensapka);
end
secilenenyildiz=fix(unifrnd(2,251,250,200));
```

EK 2 Bölüm 4.2 için Yazılan Matlab Bilgisayar Programı Kodları (Devam)

```
for i=2:250
wnyildiz(i,:)=enstd(secilenwnyildiz(i,:));
end
%bootstrap örnekleminin oluşturulması ve ekk tahmin edicisinin bootstrap
%şekli
a2=ekk;

xboot=zeros(250,200);
xboot(1,:)=0;
for i=2:250
    xboot(i,:)=a2*xboot(i-1,:)+enyildiz(i,:);
end
c=xboot';

toplam3=zeros(1,200);

for i=2:250
    for j=1:200
        toplamt(i,j)=xboot(i,j)*xboot(i-1,j);
    end
end
toplamt=sum(toplamt);
tttt=sum(xboot.^2)-(xboot(250,:)).^2;

ekkboot=toplamtt./tttt;
```

EK 3 $X_i = 0.5X_{i-1} + \varepsilon_i, i = 2, \dots, 250$ zaman serisi modeline uygun üretilen veriler

Sıra No	Üretilen Veri	Sıra No	Üretilen Veri	Sıra No	Üretilen Veri	Sıra No	Üretilen Veri
1	-1,12256	38	0,654379	75	0,100585	112	-0,35287
2	-1,95003	39	-0,42119	76	0,391416	113	0,694028
3	-1,0594	40	0,97174	77	-1,43503	114	1,004698
4	-1,06261	41	0,131704	78	-0,16299	115	0,233551
5	-1,54879	42	-1,68073	79	-0,28621	116	0,19691
6	-0,52426	43	-1,09166	80	-0,31845	117	0,04514
7	-0,79558	44	0,927082	81	-0,48585	118	0,04383
8	0,618073	45	2,700928	82	-0,51349	119	-0,21198
9	0,078495	46	1,091079	83	-0,04859	120	0,520941
10	0,638042	47	1,737083	84	-1,31746	121	1,452232
11	2,183093	48	-0,00977	85	-0,79018	122	0,394929
12	-0,82609	49	0,053702	86	-1,18761	123	0,998603
13	-1,62856	50	-0,87867	87	0,156873	124	-1,07642
14	0,477679	51	0,442346	88	-0,13814	125	0,501186
15	-0,93999	52	0,740269	89	0,261614	126	0,466597
16	-0,53864	53	2,920016	90	0,211211	127	-0,68465
17	-0,03326	54	2,298934	91	-0,10999	128	0,540616
18	0,826208	55	1,484524	92	0,58645	129	-1,01233
19	-1,19865	56	0,287995	93	1,423306	130	-0,8869
20	-0,87104	57	-0,07544	94	-0,60329	131	0,089656
21	-1,23136	58	0,117565	95	0,021601	132	0,047711
22	-0,30701	59	-0,84608	96	-0,77942	133	1,616131
23	-0,30956	60	-1,44407	97	-0,72271	134	0,124083
24	0,26883	61	-1,15798	98	-0,03747	135	0,127432
25	0,9024	62	0,759225	99	1,59252	136	-1,02186
26	0,52625	63	1,449226	100	1,11381	137	-2,3907
27	-0,52569	64	1,598318	101	-1,44525	138	-0,73806
28	0,894344	65	0,845648	102	1,653391	139	-1,04409
29	-1,22101	66	1,230957	103	0,894407	140	-2,17065
30	-0,18398	67	2,007049	104	0,117585	141	-0,99045
31	0,353985	68	1,390305	105	0,716297	142	0,44511
32	0,574773	69	1,093494	106	1,916896	143	-0,24154
33	0,291925	70	2,844846	107	0,927272	144	2,092694
34	-0,20483	71	2,143421	108	1,868031	145	0,893469
35	-0,15228	72	0,243174	109	1,288446	146	1,261827
36	-0,88604	73	0,053533	110	-0,39388	147	0,82403
37	-0,36764	74	0,082458	111	-0,53819	148	2,195817

EK 3 $X_i = 0.5X_{i-1} + \varepsilon_i, i = 2, \dots, 250$ zaman serisi modeline uygun üretilen veriler
(Devam)

Sıra No	Üretilen Veri	Sıra No	Üretilen Veri	Sıra No	Üretilen Veri	Sıra No	Üretilen Veri
149	0,877435	175	0,385163	201	0,996005	227	-2,25344
150	0,606655	176	0,118383	202	-0,31435	228	-0,3572
151	1,111333	177	0,37992	203	0,794723	229	-0,13325
152	1,952357	178	0,499928	204	-0,78582	230	-0,57874
153	0,306221	179	-0,65569	205	0,700757	231	-1,11071
154	-0,78505	180	0,40771	206	0,414249	232	-0,72108
155	0,203502	181	-0,44705	207	-0,6819	233	0,385932
156	0,254439	182	0,581631	208	1,444571	234	-0,77648
157	-1,32249	183	0,885752	209	0,779458	235	0,96208
158	-2,39704	184	0,68854	210	1,374415	236	-0,323
159	-1,40962	185	0,436153	211	-0,88715	237	-1,20811
160	0,481045	186	2,115882	212	-1,11732	238	-0,71647
161	0,435808	187	1,727323	213	0,556398	239	-0,70867
162	0,458577	188	1,188597	214	0,767298	240	0,023395
163	0,832372	189	0,155793	215	2,523867	241	-2,78164
164	1,016277	190	0,55114	216	1,538469	242	-2,08536
165	1,288629	191	0,049461	217	-0,19508	243	-1,50162
166	-0,58619	192	-2,25977	218	-0,97223	244	-1,10548
167	-1,02635	193	-1,75578	219	0,777477	245	-0,44153
168	-0,29064	194	-0,7105	220	-1,10636	246	-0,10476
169	0,4391	195	-0,37578	221	0,040032	247	-0,9579
170	-0,61435	196	-1,19665	222	-0,87122	248	-1,92819
171	-0,85367	197	-0,67198	223	-1,06342	249	0,240521
172	-1,16233	198	0,278396	224	-0,94839		
173	-1,80613	199	0,766438	225	-0,58198		
174	-1,98929	200	0,486013	226	-0,24303		

ÖZGEÇMİŞ

Adı Soyadı : Bengül ARKANT

Doğum Yeri : Ödemiş

Doğum Tarihi : 17.06.1983

Medeni Hali : Bekar

Yabancı Dili : İngilizce

Eğitim Durumu

Lise : Hulusi Uçaçelik Anadolu Lisesi (1994-2001)

Lisans : Ankara Üniversitesi Fen Fakültesi İstatistik Bölümü (2001-2005)

Yüksek Lisans : Ankara Üniversitesi Fen Bilimleri Enstitüsü İstatistik

Anabilim Dalı (Eylül 2005-Ağustos 2008)

Çalıştığı Kurum/ Kurumlar ve Yıl

Ankara Üniversitesi Fen Fakültesi İstatistik Bölümü Araştırma Görevlisi (2007-)