ANADOLU ÜNİVERSİTESİ BİLİM VE TEKNOLOJİ DERGİSİ
ANADOLU UNIVERSITY JOURNAL OF SCIENCE AND TECHNOLOGY
Cilt/Vol.:7-Sayı/No: 2 : 387-391 (2006)

ARAŞTIRMA MAKALESİ /**RESEARCH ARTICLE**

UZUN DÖNEM BAĞIMLI NORMAL AKGÜRÜLTÜ SÜRECİNDE KESİRLİ FARK PARAMETRESİNİN BAYES TAHMİNİ

Erol EĞRİOĞLU[1], Süleyman GÜNAY

ÖZ

Uzun dönem bağımlı zaman serilerinde kesirli fark parametresinin tahmin edilmesi önemli bir sorun olarak düşünülür. Kesirli fark parametesinin tahmin edilmesinde bazı parametrik ve yarı parametrik yöntemler vardır. Bu çalışmada uzun dönem bağımlı normal akgürültü sürecinde kesirli fark parametresinin tahmin edilmesinde bir Bayesci yaklaşım önerildi. Önerilen yaklaşımda normal önsel dağılım ve tam olabilirlik fonksiyonu kullanıldı. Bir benzetim çalışması ile önerilen yaklaşım ile bulunan tahmin edici logaritmik periodogram tahmin edicisi ile karşılaştırıldı. Önerilen yaklaşımın periodogram tahmin edicisine göre daha iyi sonuçlar verdiği görüldü.

Anahtar Kelimeler : Uzun dönem bağımlılık, Bayes tahmini, Zaman serileri

BAYESIAN ESTIMATION of THE FRACTIONALLY DIFFERENCING PARAMETER in THE LONG RANGE DEPENDENCE NORMAL WHITE NOISE PROCESSES

ABSTRACT

The estimation of the fractionally differencing parameter is an important problem in the long memory time series. There are many parametric and semi parametric methods for the estimation of this parameter. In this study, a Bayesian approximation for estimation of the fractionally differencing parameter in long range dependence white noise processes is proposed. Normal prior and exact likelihood functions are used in the proposed Bayesian method. This method is compared with the logarithmic periodogram regression method by a simulation study. The proposed method is better than logarithmic periodogram regression.

Keywords: Long range dependence, Bayesian estimation, Time series

[1,] Hacettepe Üniversitesi, Fen Fakültesi, İstatistik Bölümü, Beytepe-ANKARA.
E-posta: erole@hacettepe.edu.tr

Geliş: 20 Ekim 2005; **Düzeltme:** 03 Nisan 2006; **Kabul:** 01 Ağustos 2006

1. GİRİŞ

Literatürde uzun dönem bağımlı zaman serilerinde kesirli fark parametresinin tahmin edildiği çalışmalar yapılmıştır. Bu yöntemleri parametrik ve yarı parametrik yöntemler olarak gruplandırmak mümkündür. R/S istatistiği (Hurst, 1951) , logaritmik periodogram regresyonu yöntemi (Geweke ve Porter-Hudak , 1983) ve otokorelasyon regresyonu yöntemi (Eğrioğlu ve Günay , 2005) yarı parametrik tahmin yöntemlerine örnek olarak verilebilir. Parametrik tahmin yöntemlerine örnek olarak kesilmiş tahmin yöntemi (Li ve McLeod, 1986), Hosiking'in yaklaşık olabilirlik yöntemi (Hosking, 1984), Whitle'ın yaklaşık olabilirlik yöntemi (Beran ,1994) ve Sowell'ın en çok olabilirlik yöntemi (Sowell, 1992) verilebilir. Kesirli fark parametresinin tahmin edildiği parametrik tahmin yöntemlerinde yaklaşık olabilirlik fonksiyonu kullanılmaktadır. Nan-Jung Hsu ve F. Jay Breidt (2003) de ARFIMA(0,d,0) sürecinde kesirli fark parametresini Bayesci olarak tahmin etmiştir. Çalışmalarında ARFIMA(0,d,0) sürecini ARMA(2,2) ile yaklaştırarak Markov Zinciri Monte Carlo yöntemlerini kullanmışlardır.

Hosking(1981) de yarı parametrik tahmin yöntemleri ile Box- Jenkins algoritmasının birlikte kullanıldığı iki aşamalı bir tahmin yöntemi önermiştir. Hosking'in önerdiği yaklaşımın daha sonra üstün ve zayıf yönleri Smith vd. (1997) ve Reinsen vd. (2001) de bir benzetim çalışması ile ortaya konmuştur. Reinsen vd. (2001) de Hosking'in yaklaşımı ile yeteri kadar iyi sonuçlar elde edilebileceğini vurgulamaktadır. Bu durumda logaritmik periodogram regresyonu gibi yarı parametrik tahmin yöntemlerinin kullanımındaki önemi artırmaktadır. Yarı parametrik tahmin yöntemlerinde sürecin varyansının tahmini ihmal edilerek sadece kesirli fark parametresinin tahmini üzerinde yoğunlaşılmaktadır.

Bu çalışmada uzun dönem bağımlı normal akgürültü sürecinde kesirli fark parametresinin tam olabilirlik fonksiyonu ve önsel dağılım olarak normal dağılım kullanılarak Bayesci tahmini elde edilmiştir. Literatürde kesirli fark parametresinin tahmini üzerine yapılan yaklaşımlarda yaklaşık olabilirik fonksiyonu üzerinden tahmin yapılmaktadır. Bu nedenle bu çalışmada verilen tam olabilirliğe dayalı bir yöntem ilk defa önerilmektedir. Önerilen yöntem en yaygın kulanılan yarı parametrik tahminm yöntemi olan logaritmik periodogram yöntemi ile çeşitli örnek büyüklükleri ve farklı parametre değerleri açısından benzetim çalışması ile karşılaştırılarak sonuçlar tablolar halinde verilmiştir.

Çalışmamızın yazım planı su şekildedir. Çalışmanın ikinci bölümünde uzun dönem bağımlı normal akgürültü sürecine ilişkin tanımlar verilmiştir. Üçüncü bölümde Geweke ve Porter-Hudak'ın logaritmik periodogram regresyonu, dördüncü bölümde önerilen Bayes yaklaşımı sunulmuştur. Beşinci bölümde benzetim çalışmasının sonuç tabloları, son bölümde ise sonuçlar ve tartışmalar verilmiştir.

2. UZUN DÖNEM BAĞIMLI NORMAL AKGÜRÜLTÜ SÜRECİ

X_t uzun dönem bağımlı normal akgürültü süreci için d uzun dönem bağımlılık parametresi olmak üzere model denklemi şöyle yazılabilir:

$$X_t = (1-B)^{-d}\varepsilon_t \quad , \ \varepsilon_t \sim N(0{,}1)\ , \ 0 < d < 0.5 \qquad \textbf{(1)}$$

Kesirli fark parametresi $d \geq 0.5$ olduğunda (1) modelinin durağan olmadığı Hosking (1981) de belirtilmiştir. Uzun dönem bağımlı normal akgürültü süreci için kuramsal otokovaryans, otokorelasyon katsayısı Beran (1994) ve Baillie (1996) da

$$\gamma(k) = \frac{(-1)^k \Gamma(1-2d)}{\Gamma(k-d+1)\Gamma(1-k-d)} \qquad \textbf{(2)}$$

$$\rho(k) = \frac{\Gamma(1-d)\Gamma(k+d)}{\Gamma(d)\Gamma(k+1-d)} \qquad \textbf{(3)}$$

formülleri ile verilmektedir. X_t zaman serisinin n birimlik gözlemi $X_1,...,X_n$ olsun. Ω, $X_1,...,X_n$ lerin otokovaryans matrisini göstersin. Buna göre $X_1,...,X_n$ lerin olabilirlik fonksiyonu $\varepsilon_t \sim N(0{,}1)$ olduğundan,

$$L(d;X) = (2\pi)^{-n/2}|\Omega|^{-1/2}\, exp\left\{-\frac{1}{2}X'\Omega^{-1}X\right\} \qquad \textbf{(4)}$$

eşitliği ile verilir.

3. LOGARİTMİK PERİODOGRAM REGRESYONU

Uzun dönem bağımlılık parametresinin tahmini için farklı bir yöntem de Geweke ve Porter-Hudak (1983) tarafından geliştirilmiştir. Bu çalışmada sıfır yöresinde spektral yoğunluk fonksiyonu incelenmiştir. Sıfıra yakın frekanslar kullanılarak, frekansların logaritması üzerine logaritmik periodogramın regresyonu yapılmış ve eğimin tahmini uzun dönem bağımlılık parametresinin tahmini olarak alınmıştır. Geweke ve Porter-Hudak (1983) çalışmasından bazı önemli noktalar şöyle verilebilir.

$\{X_t\}$'nin T büyüklüğünde bir örneklemin verildiği varsayılsın. Buna göre

$$\lambda_{j,T} = \frac{2\pi j}{T}\ , \ j = 0,...,T-1$$

yazılabilir. Burada $\lambda_{j,T}$ harmonik ordinatları ve $I(\lambda_{j,T})$ bu ordinatlardaki periodogramı göstersin. Buna göre logaritmik periodogram ARFIMA$(0,d,0)$'ın spektral yoğunluğun bir tahmin edicisi olarak Geweke ve Porter-Hudak (1983) de ,

$$ln\{I(\lambda_{j,T})\} = ln\{\sigma^2 f_\varepsilon(0)/2\pi\} - d\,ln\{4\,sin^2(\lambda_{j,T}/2)\}$$
$$+ ln\{f_\varepsilon(\lambda_{j,T})/f_\varepsilon(0)\} + ln\{I(\lambda_{j,T})/f(\lambda_{j,T})\}$$

şeklinde verilmiştir. Bu eşitliğin nasıl elde edildiği detaylı bir şekilde Geweke ve Porter-Hudak (1983) de verilmektedir. Logaritmik periodogramın yukarıda verilen eşitliğine doğrusal regresyon denklemi gözüyle bakılarak d'nin tahmini elde edilebilir. Burada $ln\{I(\lambda_{j,T})\}$ bağımlı değişken , $ln\{4sin^2(\lambda_{j,T})\}$ açıklayıcı değişken, -d eğim parametresidir. Öte yandan $ln\{f_\varepsilon(\lambda_{j,T})/f_\varepsilon(0)\}$ terimi harmonik frekanslar sıfıra yakın olduğundan ihmal edilebilir. Bu eşitlikte $ln\{\sigma^2 f_\varepsilon(0)/2\pi\}$ ve $ln\{I(\lambda_{j,T})/f(\lambda_{j,T})\}$'nin ortalamasına sabit terim gözüyle bakılabilir.

Geweke ve Porter-Hudak (1983) de regresyon tahmin edicisi ile ilgili aşağıdaki teorem verilmiştir. Teoremin tanıtı Geweke ve Porter-Hudak (1983) den incelenebilir.

Teorem: $\{X_t\}$, $d < 0$ uzun dönem bağımlı süreç ve $I(\lambda_{j,T})$, $\lambda_{j,T} = \pi j/T$ harmonik frekanslarda $\{X_t\}$'nin periodogramını göstersin. $b_{1,T}$, $V\hat{a}r(b_1)$, aşağıda verilen (5) regresyon denkleminden β_1'in ve $var(b_1)$ en küçük kareler tahmini olmak üzere $\lim_{T\to\infty} g(T) = \infty$ ve $\lim_{T\to\infty} g(T)/T = 0$ özelliğine sahip bir $g(T)$ fonksiyonu için, $n = g(T)$, $p\lim_{T\to\infty} b_1 = -d$ ve $\lim_{T\to\infty}(ln\,T)^2/g(T) = 0$ ise;
$(b_1 + d)/(V\hat{a}r(b_1))^{1/2} \xrightarrow{D} N(0,1)$ olur.

Burada

$$ln\{I(\lambda_{j,T})\} = \beta_0 + \beta_1 ln\{4sin^2(\lambda_{j,T}/2)\} + \varepsilon_{j,T} \quad , j = 1,2,...,n \qquad \textbf{(5)}$$

4. BAYESCİ YAKLAŞIM

(1) eşitliği ile verilen uzun dönem bağımlı normal akgürültü sürecinde kesirli fark parametresinin önsel dağılımı olarak normal dağılım (d~N(μ,σ^2)) ve (4) eşitliği ile verilen tam olabilirlik fonksiyonu kullanılarak sonsal dağılım aşağıdaki gibi elde edilebilir.

$$P(d) = (2\pi\sigma^2)^{-1/2} exp\left\{-\frac{1}{2\sigma^2}(d-\mu)^2\right\}$$

$$P(X/d) = L(d;X)$$

$$P(d/X) \propto P(d)L(d;X)$$

$$P(d/X) \propto (2\pi\sigma^2)^{-1/2} exp\left\{-\frac{1}{2\sigma^2}(d-\mu)^2\right\} exp(2\pi)^{-n/2}|\Omega|^{-1/2} exp\left\{-\frac{1}{2}X'\Omega^{-1}X\right\} \qquad \textbf{(6)}$$

Sonsal dağılımın hesaplanmasında Ω matrisinin tersinin ve determinantının hesaplanması gerekmektedir. Bu işlemleri kolaylaştırmak için Cholesky ayrıştırması kullanılabilir. Cholosky ayrıştırması Ω matrisinin alt ve üst üçgen matrislerin çarpımı şeklinde yazılmasıdır. Bu ayrıştırma ile $n \times n$ boyutlu Ω matrisinin tersini hesaplamak yerine alt ya da üst üçgen matrislerin tersi hesaplanmaktadır. Bu durumda hesaplamalarda büyük kolaylık ve zaman tasarrufu sağlanır. Ansley (1979) da otoregresif hareketli ortalamalar modelinin olabilirlik fonksiyonunun hesaplanması için bu yöntemi kullanmıştır.

$$\Omega = LL'$$

(6) sonsal dağılımında,

$$W = L^{-1}X$$

$$X'\Omega^{-1}X = X'(LL')^{-1}X = (L^{-1}X)'(L^{-1}X) = W'W$$

$$|\Omega|^{-1/2} = |LL'|^{-1/2} = |L|^{-1}$$

eşitliklerini yerine yazarak aşağıdaki kolay hesaplanabilir sonsal dağılım elde edilebilir:

$$P(d/X) = (2\pi)^{-1/2} exp\left\{-\frac{1}{2\sigma^2}(d-\mu)\right\} exp(2\pi)^{-n/2}|L|^{-1} exp\left\{-\frac{1}{2}W'W\right\} \qquad \textbf{(7)}$$

Bu çalışmada (7) sonsal dağılımının kesirli fark parametresine (d) göre maksimize edilmesi işleminde Brent (1973) de verilen polinomsal interpolasyona dayalı tek değikenli optimizasyon algoritması kullanılmıştır. Bu algoritma S-PLUS 6.2 paket programı içinde bulunmaktadır. Burada önerilen yeni yöntem S-PLUS 6.2 paket programında kodlanmıştır.

5. SİMÜLASYON ÇALIŞMASI SONUÇLARI

Önerilen Bayesci yaklaşımın üstün ve zayıf yönlerini ölçmek için bir benzetim çalışması yapıldı. Örneklem büyüklüğü 100,150,200,500 ve kesirli fark parametresinin 0.10, 0.3, 0.4 olduğu herbir durum için 200 tekrar, toplam 2400 tane uzun dönem bağımlı normal akgürültü süreci üretildi. Benzetim çalışmasında yapılan tekrar sayısının 200 den fazla alınması durumunda sonuçlar fazla değişmediğinden daha fazla tekrar yapılmamıştır. Uzun dönem bağımlılık parametresi $0 < d < 0.5$ aralığında incelendiğinden 0.1 ve 0.4 parametre değeri durağanlık sınırına yakın 0.3 ise durğanlık sınırından uzak parametre değeri olarak düşüldü. Kesirli fark parametresi tahmin edicilerinin büyük örneklemler için daha başarılı sonuçlar verdiği bilinmektedir. Bu nedenle örneklem büyüklüğü en fazla 500 alınmıştır. Benzetim çalışmasında 100 ve üzeri örneklem büyüklüğü için ARFIMA$(0,d,0)$ modellerinin doğru bir şekilde üretilebildiği görülmektedir. Bu nedenle en küçük örneklem büyüklüğü ise 100 olarak alınmıştır. Üretilen serilerde logaritmik periodogram regresyonu ve önerilen Bayesci yaklaşım ile sadece kesirli fark parametresi tahmin edicileri yanlılık bakımından karşılaştırıldı. Benzer şekilde Reinsen (2001)

Tablo 1. n=100 için HKT

Yöntem	**d=0.10**	**d=0.30**	**d=0.40**
L. Periodogram Regresyonu	20,92	19,19	16,81
Bayesci Yaklaşım	3,26	0,61	1,90

Tablo 2. n=150 için HKT

Yöntem	**d=0.10**	**d=0.30**	**d=0.40**
L. Periodogram Regresyonu	17,62	18,74	16,48
Bayesci Yaklaşım	1,66	0,60	1,78

Tablo 3. n=200 için HKT

Yöntem	**d=0.10**	**d=0.30**	**d=0.40**
L. Periodogram Regresyonu	17,56	20,69	15,38
Bayesci Yaklaşım	0,43	0,38	0,92

Tablo 4. n=500 için HKT

Yöntem	**d=0.10**	**d=0.30**	**d=0.40**
L. Periodogram Regresyonu	5,02	5,85	4,84
Bayesci Yaklaşım	0,39	0,37	0,54

de de çeşitli tahmin ediciler yanları bakımından karşılaştırılmıştır. Bayesci yaklaşımda önsel dağılımın parametreleri $\mu = 0.3$, $\sigma^2 = 1$ olarak alındı. Sonuçlar tablolar halinde yukarıda verildi. Tabloyu oluşturan değerler elde edilen tahmin sonuçları için yanlılığın bir ölçüsü olan hata kareler toplamlarıdır.

$$HKT = \sum_{i=1}^{200} (\hat{d} - d_i)^2$$

6. SONUÇ VE TARTIŞMA

Tablo1,2,3 ve 4 incelendiğinde aşağıda verilen sonuçlara ulaşmak mümkündür. Önerilen Bayesci yaklaşım tüm durumlarda logaritmik periodogram regresyonundan iyi sonuç vermektedir. Örneklem büyüklüğü arttıkça önerilen Bayesci yaklaşım daha başarılı tahmin sonuçları vermektedir. Örneğin d=0.1 durumunda örnek büyüklüğü 100 iken HKT=3,26 , örnek büyüklüğü 150 iken HKT=1,66 , örnek büyüklüğü 200 iken HKT=0.43 ve örnek büyüklüğü 500 iken HKT=0,38 olmaktadır. Benzer durumlar d'nin diğer değerleri içinde görülmektedir. Önerilen Bayesci yaklaşım durağanlık sınırından uzak olan d=0.3 durumu için daha başarılı sonuçlar vermektedir. Örneği n=100 örnek genişliğinde d=0.1 için HKT=3,26 , d=0.4 için HKT=1,90 , d=0.3 durumu için ise HKT=0,61 değerini almaktadır. Önerilen yaklaşımın zayıf yönü (1) eşitliğindeki gibi hata teriminin ε_t 'nin $N(0,1)$ kabul edilmesidir. ε_t 'nin dağılımı $N(0,\sigma_\varepsilon^2)$ kabul edilerek σ_ε^2 parametreside tahmin edilebilir. Bu durumda sonsal dağılım çok değişkenli olacaktır. Yine sonsal dağılımın optimizasyonunda çok değişkenli optimizasyon teknikleri kullanılabilir. Ancak Hosking (1981) de önerilen iki aşamalı yaklaşımın ilk aşamasında logaritmik periodogram regresyonu yöntemi gibi sadece kesirli fark parametresi d'nin tahmini veren yöntemler kullanılmaktadır. Bu durum düşünülürse burada önerilen yaklaşım Hosking (1981) de önerilen iki aşamalı yaklaşımın ilk aşamasında kullanılabilir.

KAYNAKLAR

Ansley,C.F. (1979). An Algorithm for the Exact Likelihood of a Mixed Autoregressive Moving Avarage Process , *Biometrica* 66, 59-65.

Baillie R. T. (1996). Long Memory Processes and Fractionally Integration in Econometrics, *Journal of Econometrics* 73, 5-59.

Beran,J. (1994). *Statistics for Long-Memory Processes*, Chapman&Hall CRC.

Brent, R. (1973). *Algorithms for Minimization without Derivatives*. Prentice-Hall, Englewood Cliffs, NJ, USA.

E.Eğrioğlu ve S. Günay (2005). Uzun Dönem Bağımlı Normal Akgürültü Sürecinde Otokorelasyon Regresyonu ile Parametre Tahmini. *Anadolu Üniversitesi Bilim ve Teknoloji Dergisi* 6, 61-66.

Geweke, J. and Porter-Hudak, S. (1983). The Estimation and Application of Long-Memory Time Series Models. *Journal of Time Series Analysis* 4, 221-237.

Hosking J.R.M. (1981). Fractionally Differencing, *Biometrica* 68, 1, 165-176.

Hosking J.R.M. (1984). Modeling Persistence in Hydrological Time Series Using Fractionally Differencing. *Water Resources Research* 20(12), 1898-1908.

Hurst H.E. (1951). Long-Term Storage Capacity of Reservoirs. *Trans. Am. Soc. Civil Engineers* 116, 770-799.

Li, W.K. ve Mcleaod, A.I. (1986). Fractional Time Series Modelling. *Biometrica* 73, 217-221.

Nan-Jung Hsu ve F. Jay Breidt (2003). Bayesian Analysis of Fractionally Integrated ARMA with Additive Noise. *Journal of Forecasting* 22 , 491-514.

Reinsen V., Abraham B. ve Lopes S. (2001). Estimation of Parameters in ARFIMA Processes: A simulation Study. *Communications Statistics and Simulation* 30(4),787-803.

Smith J., Taylo N. ve Yadav S (1997). Comparing The Bias And Misspecificiation in ARFIMA Models. *Journal of Time Series Analysis* 18, 507-527.

Sowell F. (1992). Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models. *Journal of Econometrics* 53,165-188.

Erol Eğrioğlu, lisans eğitimini Ondokuz Mayıs Üniversitesi, Fen-Edebiyat Fakültesi, İstatistik Bölümünde, 1998 yılında tamamladı. Bilim uzmanlığı derecesini 2002 yılında Ondokuz Mayıs Üniversitesi Fen-Edebiyat Fakültesi, İstatistik Bölümünde aldı. 2006 yılında doktora derecesini aldı

Süleyman Günay, Ankara Üniversitesi, Fen Fakültesi, Matematik Bölümünde, 1968 yılında lisans eğitimini tamamladı. Hacettepe Üniversitesi, Fen Fakültesi, İstatistik Bölümünde 1972 yılında doktora derecesini aldı. 1978 yılında Doçent, 1988 yılında Profesör oldu. Halen Hacettepe Üniversitesi, Fen Fakültesi, İstatistik Bölümünde Bölüm Başkanı olarak çalışmalarına devam etmektedir.