

AR(1) MODELİ İÇİN AYIRIM FONKSİYONU; NORMAL VE NORMAL OLMAYAN DAĞILIMLAR ÜZERİNE BİR SİMULASYON ÇALIŞMASI

Reşat KASAP, İhsan ALP

Gazi Üniversitesi, Fen-Edebiyat Fakültesi, İstatistik Bölümü, 06500 / Ankara

ÖZET

Bu çalışmada AR(1) için $\{Z_t\}$ sürecinin p_1 ve p_2 hipotezleri ile verilen iki kategoriden birine ait olduğunu gözönüne almaktayız. Bu hipotezler sırasıyla p_1 ve p_2 altında $\{Z_t\}$, $f(w)$ ve $g(w)$ spectral yoğunluk fonksiyonlara sahip olduğunu belirtir. Eğer $f(w)$ ve $g(w)$ biliniyorsa $I(f:g)$ 'yi ayırım fonksiyonu olarak kullanabiliriz. Çalışmada Normal ve normal olmayan durumlar için simülasyon çalışmasıyla bir sayısal örnek verilecektir.

Anahtar Kelimeler : Otoregressif süreç, Ayırım fonksiyonu, İstatistiksel dağılımlar, Simülasyon

A DISCRIMINATION FUNCTION FOR THE AR (1) MODEL : A SIMULATION STUDY ON THE NORMAL AND ABNORMAL DISTRIBUTIONS

ABSTRACT

We shall consider the case where a process $\{Z_t\}$ for an AR(1) belongs to one of two categories given by two hypotheses p_1 and p_2 in this paper. These hypotheses specify that $\{Z_t\}$ has spectral density function $f(w)$ and $g(w)$ under p_1 and p_2 respectively. We can use $I(f:g)$ as a discrimination function, if $f(w)$ and $g(w)$ are known. A numerical example will be given by using simulation for the normal and abnormal distribution cases in this study.

Key Words : Autoregressive process, Discrimination function, Statistical distributions, Simulation

1. GİRİŞ

Çeşitli bilim dallarında özellikle mühendislik ve fen bilimlerinde sinyallerin birbirinden ayırımını ortaya koymak oldukça önemlidir. Bu ise literatürde, zaman zaman istatistiksel çalışmaların konusu olmaktadır. Özellikle zaman ekseninde meydana gelen hareketlerin hangi temel yapıya sahip olduğunun ayırımı bu çalışmanın konusudur. Veri yapısının istatistiksel dağılımlardan Normal, Düzgün ve Üstel'e uyması durumu incelenmiştir.

Shumway ve Unger (1974), doğrusal süzgeçleme ile ilgili iki Gaussian sürecin ayırımının probleminde, Kullback-Leibler bilgi ve J-ayırma'sının bilinen

spektral yaklaşımlarını verdi. Adı geçen çalışmada bu yaklaşımların maksimizasyonu için doğrusal ayırım süzgeçleri tanımlanarak sismik kayıtlar için uygulandı. Shumway (1982) ise, zaman dizilerinde çeşitli ayırım problemlerini gözden geçirdi. Bu konudaki daha sonraki çalışmalardan en önemlilerinden biri Zhang ve Taniguchi (1994) tarafından verilmiştir. Söz konusu çalışmada durağan zaman dizileri için ayırım analizi yapıldı.

Yukarıda verilenler, bu konudaki çalışmalarını özetlemekle beraber, benzer çalışmalar günümüzde de devam etmektedir. Bu çalışmada, AR(1) modeli için belirlenmiş olan ayırım fonksiyonu kullanılarak, normal ve normal olmayan süreçlerden hareketle,

ayırım fonksiyonun işlevini ortaya koymak amacıyla bir simulasyon çalışması yapılmıştır. Bunun için öncelikle 2. Bölümde AR(1) modeli için ayırım fonksiyonu verilmiştir. Simulasyon çalışmasında kullanılmak üzere oluşturulan algoritma bölüm 3'dedir.

Bölüm 4'de ise elde edilen sonuçlar verilmiştir.

2. MATERYAL VE METOT

2. 1. Ayırım Fonksiyonu

Z_t , ($t = 1, 2, \dots, T$) durağan bir zaman dizisi olmak üzere, ayırım fonksiyonu

$$I(f;g) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left[\log \left\{ \frac{g(w)}{f(w)} \right\} + I(w) \left\{ \frac{1}{g(w)} - \frac{1}{f(w)} \right\} \right] dw \quad (1)$$

şeklinde tanımlanmıştır (Zhang ve Taniguchi, 1994). Burada $f(w)$ ve $g(w)$ spektral yoğunluk fonksiyonları olup, $I(w)$, Z_t 'nin periodogramıdır (Priestley, 1981).

Bu çalışmada kullanılacak zaman dizileri süreci, birinci dereceden otoregressif (AR)'dir. Buna göre

$$(1 - \phi B) Z_t = A_t \quad (2 a)$$

veya,

$$Z_t = \phi Z_{t-1} + A_t \quad (2 b)$$

olarak yazılır. Burada $-1 < \phi < 1$ olarak tanımlıdır (Pandit ve Wu, 1993). Anderson (1984)'den

$$I(f;g) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left[\log \left\{ \frac{1 + \phi^2 - 2\phi \cos w}{1 + \phi^{*2} - 2\phi^* \cos w} \right\} + T \left(\frac{1}{T} \sum_{t=1}^T Z_t \cos wt \right)^2 \right]$$

$$\frac{2\pi}{\sigma_A^2} \left\{ (1 + \phi^{*2} - 2\phi^* \cos w) - (1 + \phi^2 - 2\phi \cos w) \right\} dw \quad (7 a)$$

olur. İşlemler yapıldıktan sonra ayırım fonksiyonu

$$I(f;g) = \frac{1}{2T} \left\{ (\phi^{*2} - \phi^2) \sum_{t=1}^T Z_t^2 + 2(\phi - \phi^*) \sum_{t=1}^{T-1} Z_t Z_{t+1} \right\} \quad (7 b)$$

olarak yazılır.

2. 2. Algoritma

hareketle ayırım fonksiyonu kullanılarak Z_t sürecinin sınıflandırılması için

$$p_1 : f(w | \phi) \quad (3 a)$$

$$p_2 : g(w | \phi + \frac{h}{\sqrt{T}}) \quad (3 b)$$

katogorisinden birine dahil edilecektir; eğer $D(f,g) > 0$ ise p_1 'e, değilse p_2 'ye gönderilecektir.

Ayrıca ;

$h(j) = \sqrt{1 - \phi^2}$, $j = 1, \dots, 4$ dir. Yukarıda verilen $D(f,g)$ 'yi, $f(w)$ ve $g(w)$ değerlerini bularak yerine konulduğunda aşağıdaki gibi yazılabilir (Zhang ve Taniguchi, 1994). Öyleyse, öncelikle $\phi + h / \sqrt{T} = \phi^*$ alındığında

$$f(w) = \frac{\sigma_A^2}{2\pi} \frac{1}{(1 + \phi^2 - 2\phi \cos w)} \quad (4)$$

$$g(w) = \frac{\sigma_A^2}{2\pi} \frac{1}{(1 + \phi^{*2} - 2\phi^* \cos w)} \quad (5)$$

ve

$$I(w) = \frac{1}{T} \sum_{t=1}^T Z_t \cos wt \quad (6)$$

dır. Buradan

Bu simulasyon çalışmasında kullanılan algoritma aşağıda verildiği gibidir:

Begin

for I : = 1 to number of simulation

for J : = 1 to number of data { T }

ET (J) : = number from related distribution

End ;

sn % = .9/ stepsize ; { DT }

Q = 0 ;

```

for O1 : = 1 to sn % do
for J : = 1 to number of J do
XT(J) := Q*XT(J-1) + ET(J);
H(J) := J* (3- O^2)^.5;
MU := Q + H(J)/T^ . 5 ;
TOP1 := 0 ;
TOP2 := 0 ;
for L := 1 to T do
TOP1 := TOP1 + XT(L) ^ 2 ;
End ;
for L := 1 to T-1 do
TOP2 := TOP2 + XT(L) * XT(L + 1) ;
End ;
I := (1/(2*T))* ((MU^2-O^2)* TOP1 + 2*
(O - MU)*TOP2 ;
iF I > 0 then say (O1, j) = say (O1,j) +1 ; end ;
end;

```

end ;

print say (O1, J)

İstenildiği takdirde çalışır durumdaki programlar yazarlardan elde edilebilir.

3. SİMULASYON SONUÇLARI

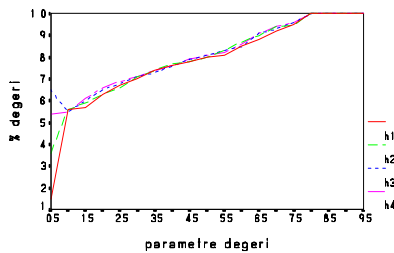
AR(1) süreci için ayırım fonksiyonunun Normal ve Normal olmayan durumlar için incelenmesi yukarıda verilen algoritma yardımıyla bir simulasyon çalışmasıyla yapılmıştır. Bunun için Normal (N), Üstel (Ü) ve Düzgün (D) dağılımlardan hareketle üretilen veriler kullanılmıştır. Her bir çalışma; 1500 gözlemlik, 1000 tekrar, $h = 4$ alınmıştır. ϕ parametresi ise, 0.05 den 0.95'e kadar 0.05 artırılarak yapılmıştır. Bunlara ilişkin sonuçlar yüzde (%) olarak Tablo 1'de verilmiştir.

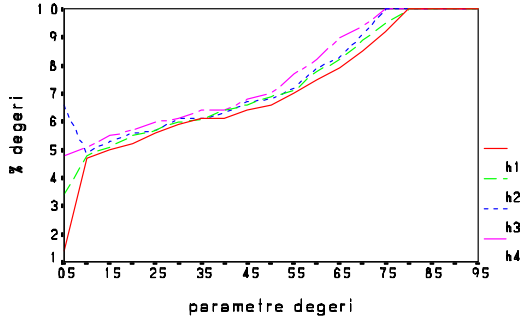
Tablo 1. Simulasyon Sonucunda Elde Edilen Değerler (%)

ϕ	N				Ü				D			
	h(1)	h(2)	h(3)	h(4)	h(1)	h(2)	h(3)	h(4)	h(1)	h(2)	h(3)	h(4)
0.05	15	36	65	54	14	34	66	48	24	26	61	54
0.10	56	56	55	55	47	48	49	51	57	57	58	58
0.15	57	59	60	61	50	51	53	55	59	59	60	59
0.20	63	63	65	66	52	55	56	57	60	62	62	64
0.25	67	66	68	69	56	57	57	60	66	68	72	70
0.30	70	71	71	71	59	60	61	61	72	72	72	71
0.35	74	74	73	74	61	61	61	64	73	73	73	73
0.40	76	77	76	76	61	64	63	64	75	75	74	75
0.45	78	78	79	79	64	66	67	68	76	76	75	76
0.50	80	80	81	81	66	69	68	70	78	78	77	78
0.55	81	83	83	82	70	71	72	77	78	79	80	82
0.60	85	87	85	86	75	78	79	82	83	83	83	84
0.65	88	90	91	91	79	82	83	90	85	85	86	87
0.70	92	94	93	94	85	89	91	94	88	88	90	89
0.75	95	95	96	96	92	95	100	100	91	94	97	98
0.80	100	100	100	100	100	100	100	100	100	100	100	100
0.85	100	100	100	100	100	100	100	100	100	100	100	100
0.90	100	100	100	100	100	100	100	100	100	100	100	100
0.95	100	100	100	100	100	100	100	100	100	100	100	100

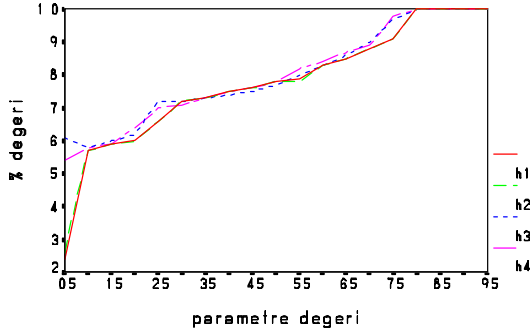
Bunlara karşılık gelen her bir dağılımlı durum için h değerleri açısından grafikleri Şekil 1, 2 ve 3'de sunulmuştur.

Şekil 1. Normal dağılımlı durum için ayırım yüzdeleri grafiği



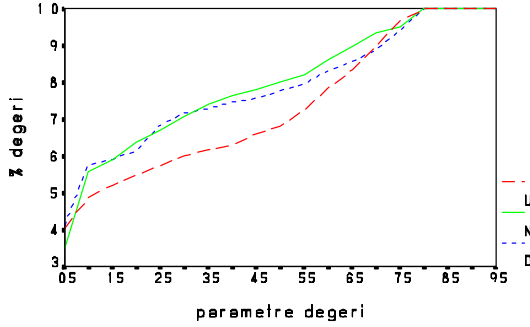


Şeki 2. Üstel dağılımlı durum için ayırım yüzdeleri grafiği



Şeki 3. Düzgün dağılımlı durum için ayırım yüzdeleri grafiği

Ayrıca, herbir dağılımın ortalama ayırım yüzdelerini karşılaştıran grafik ise Şeki 4' de görülmektedir.



Şeki 4. U, N ve D dağılımlı durumlar için ortalama ayırım yüzdeleri grafiği

Yukarıda verilen Tablo 1 ve şekiller göz önüne alındığında, AR (1)'in ayırım fonksiyonu bütün

dağılımlar için, ϕ ve h değerleri arttıkça daha doğru seçim yaptığı görülmektedir. Özellikle ϕ 'nin artışı ile bu daha da belirgin bir görüntü içerisinde. Normal dağılım için ayırımların diğerlerine göre daha iyi olduğu, Üstel dağılım değerleri ise diğerlerine göre daha kötü olduğu söylenebilir. Ayrıca, ϕ 'nin 0.80 den 0.95'e kadar değişen değerleri için % 100'nün doğru ayırma sahip olduğu görülmüştür.

KAYNAKLAR

Anderson, T. W. 1984. An Intoduction to Multivariate Statistical Analysis, JW, New York.

Pandit, S. M., Wu, S. M. 1993. Time Series and System Analysis With Applications. Krieger Publishing Company, Florida

Priestley, M. B. 1981. Spectral Analysis and Time Series, Volume 1-2, Academic Pres, London.

Shumway, R. H. 1982. Discriminant Analysis for time Series, In Handbook of Statistics, Vol. 2, ed. P. R. Krishnaiah and L.N. Kanal, Amsterdam; Nort-Holland, 1-4

Shumway, R. H., Unger, A. N. 1974. Linear Discriminant Functions for Stationarytime Series, J. Statist. Assoc., (69), 948-956.

Zhang, G., Masanobu, T. 1994. Discriminant Analysis for Stationary Vector Time Series, Journal of Time Series Analysis, 15 (1), 117-126.