D.E.Ü. İ.İ.B.F. Dergisi
Cilt: 9, Sayı: I, 1994, ss.195-205.

DOĞRUSAL OLMAYAN EKONOMETRİK MODELLERDE HETEROSKEDASTİSİTE

Murat KARAGÖZ (*)

ÖZET

Doğrusal olmayan regresyonda değişken varyans ve diğer sorunlar doğrusal modellere nazaran daha kompleks bir yapıda ortaya çıkmaktadır. Bu çalışmada, veri dönüştürme teknikleri, özellikle "İki Yanlı Dönüştürme" yöntemi incelenmiştir. Verilerin çarpık ve değişken varyans özelliğine sahip olduğu durumlarda bu yöntem yararlı olmaktadır.

GİRİŞ

Regresyon analizlerinde genelleştirilmiş En Küçük Kareler (EKK) yöntemine alternatif olarak iki yanlı dönüştürme yöntemleri de değişken varyans problemine karşı kullanılabilir. Veriler modellenirken tesadüfi hata teriminin yokluğunda y bağımlı değişkeni x tahmin değişkenlerinden

$$y = f(x, \beta) \qquad (1)$$

deterministik ilişkisi ile tahmin edilebilir. ancak gerçekte değişkenlerin ölçümü tesadüfi bir hata unsuru içerdiğinden böyle bir deterministik ilişki kullanılamaz. Bunun için şu varsayımları yapıyoruz.

$$E(y) = f(x, \beta) \qquad (2)$$

$$y - f(x, \beta) = u \;, \quad \text{var}(y) = \text{var}(u) = \sigma^2 \qquad (3)$$

u hata terimi herbir x için özdeş olarak dağılmaktadır. (4)

u hata terimi herbir x için birbirinden bağımsız olarak dağılmaktadır . (5)

Kuşkusuz, (4) varsayımı (3) 'ü ima etmektedir. Ancak bazen (3) varsayımı ile yetinilmekte olduğundan bunları ayrı ayrı gösterdik. Hatalar normal dağıldığından β nın EKK ile tahmini etkin olmaktadır. Buradan beşinci varsayım

Hatalar normal olarak dağılmaktadır. (6)

Doğrusal olmayan regresyon uygulamalarında (3) ve (4) nolu varsayımla-

---

(*) Yrd.Doç.Dr. İ.Ü. İ.İ.B.F. Ekonometri Bölümü.

rın ciddi olarak ihlal edildiği görülmektedir. Bu çalışmanın başlıca amacı, (3) ve (4) nolu varsayımların ihlali yani değişken varyans durumunun analizidir.

Değişken varyansın gözlenmesi durumunda β ların tahmini için iki temel strateji kullanılabilir. Birincisi, hataların normal dağıldığı varsayımı altında ağırlıklı EKK (iterasyonlu) yöntemini kullanmaktadır. İkinci strateji ise, hata teriminin normal dağılım varsayımı yerine sabit varyanslı başka bir dağılım varsaymaktadır. Bu starateji içerisindede iki önemli seçim mevcuttur. İlki hataların gamma dağılımına sahip olduğunu varsaymak. Gamma dağılımı, bir üstel dağılım ailesi olduğundan bu bizi genelleştirilmiş doğrusal (olmayan) modellere götürecektir. Bu konu titiz bir şekilde McCullagh ve Nelder (Mc Cullagh, 1983) tarafından tartışılmaktadır. Burada tahmin metodu normal dağılımlı gözlemlerde olduğu gibi iteratif ağırlıklı EKK yöntemidir. İkinci yaygın seçim alanı ise, verilerin log-normal dağılıma sahip olduğunu varsaymaktadır. Yani y'nin logaritmaları normal dağılıma sahiptir. İşte bu sonuncu iki yanlı dönüştürmeye bir örnek teşkil etmektedir, şöyle ki,

$$\ln y = \ln [f(x,\beta)] + u \tag{7}$$

olmakta ve u klasik varsayımları, en azından yaklaşık olarak, sağlamaktadır.

Logaritma dışında başka dönüştürmeler de sabit varyans ve normal dağılımı sağlamak için kullanılabilir. Bütün bu dönüştürme seçenekleri tek bir model halinde

$$h(y, \lambda) = h[f(x, \beta), \lambda] + u$$

ile temsil edilebilir. h (y, λ) fonksiyonu Box-Cox (Box- Cox, 1964, 211-46) üstel dönüştürme ailesi olup

$$h(y, \lambda) = \begin{cases} (y - 1)/\lambda & \text{eğer } \lambda \neq 0 \text{ ise} \\ \ln(y) & \text{eğer } \lambda = 0 \text{ ise} \end{cases} \tag{8}$$

ile verilmektedir. λ = 0 olması logaritmik dönüştürme anlamına gelirken λ = 0.5 karekök dönüştürmesi, λ = 1 karekök dönüştürme olmaması anlamına gelmektedir. Aşağıda bu metodu ayrıntılı olarak tartışacağız.

2.İKİ YANLI DÖNÜŞTÜRME METODOLOJİSİ

(1) ile verilen genel bir doğrusal olmayan regresyon modelinde toplumsal normal dağılımlı bir hata terimi varsayımı yaygın bir uygulamadır. Yani u'lar 0 ortalama ve σ^2 sabit varyansı ile normal dağılıma sahiptir. Ancak çoğu kez kalıntılar heteroskedastik (değişken varyanslı) ve/veya çarpık olmakta ve aşırı uç değerler içermektedir. Tek başına heteroskedastisite genelleştirilmiş EKK ile giderilebilmektedir. Fakat aşırı normal-dışı dağılımlı u'lar durumunda varyansın tersi ile ağırlıklandırmak yeterli olmamaktadır.

Çarpık verilerin modellemesi için iki iyi gelişmiş yöntem mevcuttur. Bir yaklaşım, u'nun yoğunluğunun parametrik bir çarpık yoğunluk sınıfı, sözgelimi gamma yoğunluğu, içerisinde olmasına izin vermektir. Bu yaklaşım genelleştirilmiş doğrusal modeller olarak bilinmektedir. Bu çalışmada biz, y'nin simetrik bir dağılıma dönüştürülebileceğini varsaymaktayız. Uygulamada genelleştirilmiş doğrusal modeller ve dönüştürme modelleri başarılı bir şekilde uygulanmaktadır. Ancak her iki yöntem de esasen doğrusal olmayan modelleri de kapsayacak şekilde genelleştirilmelidir. Örneğin genelleştirilmiş doğrusal modeller genelleştirilmiş doğrusal olmayan modellere genişletilebilir. h (y, λ), y için λ parametresi ile indekslenmiş bir dönüştürme ailesi olsun. Bir çok örnekte λ skaler olarak verilmekle birlikte bu şart değildir. Genel olarak y için h (y, λ) nin monoton bir fonksiyon olması gereklidir. Aksi halde h (y, λ) için bir model, y için bir modeli dönüştürmenin tersi ile yeteri kadar genelleştiremez.

En yaygın kullanılan dönüştürme (8) ile verilen Box-Cox dönüştürmesidir. Bir başka muhtemel durum ise kayan üstel dönüştürmedir. Yani,

$$h\,(y,\,\lambda,\mu) = [y - \mu]^{(\lambda)} \tag{9}$$

(8)'in kullanımı y'nin pozitif olduğunu varsaymaktadır. Eğer y muhtemel bir minimum değere sahip ise o zaman, μ 'nün uygun seçimi ile (y - μ) de pozitif yapılabilir. Kayan üstel dönüştürme Atkinson (Atkinson, 1985) tarafından etraflıca tartışılmaktadır.

Dönüştürme ailesi seçildikten sonra (1) ile verilen model

$$h\,(y,\,\lambda,\mu) = h\,[f\;(x,\beta),\,\lambda] + u \tag{10}$$

ile ikame edilmektedir. Bu model "iki yanlı dönüştürme" olarak anılacak ve İYD ile kısaltılacaktır. Bu model şunu ifade etmektedir; y ve f (x,β)'e aynı dönüştürme işlemi uygulandığında kalıntılar sabit bir varyans ile normal dağılıma sahip olmaktadır (Carroll ve Ruppert, 1984, 321-8).

Buradaki düşünce şöyledir. Hiçbir varyans kaynağı olmadığında, yani u = 0 ise, model (1) ile model (10) eşdeğerdir. Sabit varyanslı simetrik bir hata terimi elde etmek için y'yi dönüştürmeye tabi tutuyoruz. Daha sonra aynı dönüştürmeyi f (x,β)'ya uygulayarak y ile x arasındaki ilişkiyi muhafaza etmiş oluyoruz. (10) nolu denklem tesadüfi değişmenin modele çeşitli şekillerde girmesini sağlamaktadır. Hatta eğer h (y, λ) bazı λ'lar için y bakımından doğrusal ise bu takdirde (10) nolu model, (1) nolu modeli özel bir durum olarak içermektedir. Gerektiğinde dönüştürme yapmama seçeneği de mevcuttur. Örneğin, H_0: λ =1 ile verilen " dönüştürme gereksizdir" hipotezini test edebiliriz.

Eğer h (y, λ) yeteri kadar zengin bir aile oluşturuyorsa, örneğin modifiye edilmemiş üstel dönüştürme; o zaman model (10) aşırı heteroskedastisite ve normal-dışılık sergileyen verilere başarılı bir şekilde uyarlanabilir. λ parametresi β ve σ ile beraber tahmin edilebilir. Böylece veriler aracılığı ile u'ların y bağımlı değişkenini nasıl etkilediğini modellemiş oluyoruz. Burada (10) modelinin Box-Cox modelinden çok farklı olduğu vurgulanmalıdır. Box-Cox modelinde, bağımlı değişken y dönüştürmeye tabi olurken f (x, β) regresyon fonksiyonuna dokunulmamaktadır. f (x, β)'nın basit bir formda örneğin doğrusal, seçildiğinde bunun dönüştürülmemiş bağımsız değişken y'ye uymadığı durumlarda Box-Cox dönüştürmesi çok uygun olmaktadır. h dönüştürmesi, f (x, β) fonksiyonu h (y, λ)'ya uyacak şekilde seçilir. Böylelikle h'nin normalite ve sabit varyansı sağlayacağı beklenmektedir.

Bizim dönüştürme modelimiz halihazırda y = f (x,β) regresyonunun yeterli bir uyuma sahip olduğunu varsaymakla beraber kalıntıların değişimi normal-dışı ya da heteroskedastiktir. Bu özellikle iktisadi modellerde makul bir varsayım olmaktadır. Zira fonksiyonel biçim çoğu kez a priori olarak iktisat teorisi tarafından sınırlanmaktadır. Değişken varyans veya normal-dışı dağılım durumunda regresyon fonksiyonunu yeniden tek yanlı olarak modellemek yerine iki yanlı dönüştürme yapmak fonksiyonel biçimi bozmayacaktır.

Kalıntıların değişimi çarpıklık veya heteroskedastisite sergiliyorsa o zaman basit olarak ağırlıksız EKK ile y'nin f (x,β)'ya regresyonu yerine, İYD yönteminin kullanılması için birkaç neden mevcuttur. İlk olarak, normalite ve homoskedastisiteye dönüştürme β parametre vektörünün daha etkin tahmin edilmesini sağlamaktadır. Dönüştürmesiz durumda β'nın basit EKK tahmin edicisi gereksiz yere büyük bir varyansa sahip olacaktır. Bundan başka, İYD modeli veri bir x değişkeni için tam bir şartlı dağılım modeli sağlamaktadır.

Regresyon modellerinde iki yanlı dönüştürme düşüncesi uzunca bir zamandır gündemde olmakla beraber geleneksel olarak, esasen doğrusal olmayan modellerin doğrusallaştırılması alanında kullanılmakta idi. Literatürde f (x,β) regresyon fonksiyonunun basit bir dönüştürme ile, sözgelimi çarpmaya göre tersi veya logaritması alınarak, parametreleri açısından doğrusal hale getirildiği bir çok örnek mevcuttur. Ancak modern doğrusal olmayan paket programlarında artık bu tür dönüştürmeler gereksiz hale gelmiştir. Hatta doğrusallaştırma dönüştürmeleri asimetri veya heteroskedastisiteye yol açarak EKK tahminlerinde önemli ölçüde etkinlik kaybına sebep olabilir. Box ve Hill (Box - Hill, Technometric, 385-9) doğrusallaştırmanın heteroskedastisiteye ve fiziksel olarak tahmini mümkün olmayan parametrelere yol açtığı bir örnek vermektedir. Box ve Hill bu durumda, dönüştürmeden vazgeçmeyip ortaya çıkan heteroskedastisiteyi ağırlıklandırma ile giderilmesini önermektedir. Carroll ve Ruppert (Carroll - Ruppert , 1984a) homoskedastisite için bir dönüştürme ile aynı ölçüde iyi bir uyumun elde edilebileceğini gösteriyorlar.

İki yanlı dönüştürme tekniğini kuşkusuz veri analizcileri normalite ve kararlı varyanslara ulaşmak için zaman zaman kullanmışlardır. Snee ve Irr (Snee, Irr, 1981, 77 - 93) bu fikri analizlerinde kullanmışlardır. Ancak İYD'yi istatistiksel bir teknik olarak ilk önerenlerin Carroll ve Ruppert (Carroll - Ruppert , 1984a) olduğu görülmektedir. İYD'nin özel bir hal olarak bir başka erken uygulaması Leech (Leech, 1975, 719-25) de görülmektedir. Bu araştırmacı, İYD'yi, hata teriminin çarpan olarak mı (üstel dönüştürmede $\lambda = 0$ 'a karşı geliyor) yoksa toplumsal olarak mı ($\lambda = 1$ 'a karşı geliyor) modele gireceğine karar vermede kullanmaktadır. Leech λ 'nın diğer değerlerine de bazı anlamlar verilebileceğini önermişse de daha fazla konu ile ilgilenmemiştir.

Bir başka kullanışlı model ise dönüştürülmüş $y^{(\lambda)}$ bağımlı değişkeninin dönüştürülmüş açıklayıcı değişken $x^{(\lambda)}$ üzerine regresyonudur. Yani,

$$y^{(\lambda)} = a_0 + a_1 x^{(\lambda)} + u \qquad (11)$$

Bu model özellikle x değişkeni y'nin gecikmeli değerleri olduğunda uygun olmakta ve bir ARIMA (Autoregressive Integrated Moving Avarage) modelini uyarlamadan önce zaman serisinin dönüştürülmesinde yaygın bir pratiktir. Örnek için bakınız Hopwood ve diğ. (1984, 57-61) ve Box ve Jenkins (1976). Egy ve Lahiri (1979) nolu modeli heteroskedastik hatalarda kullanmaktadır (1979). Bu model aynı zamanda y ve x değişkenlerinin aynı miktarlarının farklı ölçüm birimleri sözkonusu olduğunda kullanılabilir. Örneğin Leurgans (1980) iki ölçüm ile ilgili testlerin karşılaştırılması için log-dönüştürme yapmaktadır. Model (11),

$$y^{(\lambda)} = [\, B_0 + B_1 x \,] \qquad (12)$$

iki yanlı dönüştürmesine benzemektedir. a_0 ve β_0 değerleri sıfır olduğunda bu iki model özdeş olmaktadır. (12) nolu dönüştürmenin bir avantajı, x değerleri veri iken $\beta_0 + \beta_1$ x değeri y'nin şartlı beklenen değeri olduğundan λ bilinmese dahi

β_1 İktisadi olarak anlamlı bir parametre olmaktadır. (11) nolu ifadeden, y'nin şartlı medyanı

$$h^{-1}\,[\alpha + \alpha_1 x^{(\lambda)},\ \lambda]$$

olup, bilinmeyen λ değerine bağlıdır. Burada $h^{-1}(y + \lambda)$, sabit λ ile y nin fonksiyonu olan $h(y, \lambda)$ nın tersidir. Yani,

$$h^{-1}\,[h\ (y, \lambda), \lambda] = y$$

olmaktadır.

İki yanlı dönüştürmenin "sadece bağımlı değişkenin dönüştürülmesi" yaklaşımı ile çakıştığı örnekleme durumları da mevcuttur. Örneğin k-örneklem durumunda i = 1 ....k olmak üzere i inci yığından n_i boyutlu örnekleme sahip olduğunu düşünelim. Bu durumda İYD modeli,

$$h\ (y_{ij}\,,\ \lambda) = h\ (\mu_i\,,\ \lambda) + u_{ij} \qquad (13)$$

burada y_{ij}, i inci yığındaki j inci gözlem, ve u ise sabit varyanslı bağımsız normal bir değişkendir. Bağımsız değişkenin dönüştürülmesi modeli,

$$h\ (y_{ij}\,,\ \lambda) = \theta_i + u_{ij} \qquad (14)$$

olur. (13) ile (14) arasındaki tek fark μ_i, ile θ_i 'nin anlamındadır. (13) de μ_i, y_{ij} 'nin medyanıdır. Oysa (14) de θ_i, $h(y_{ij}, \lambda)$ nın hem medyanı hem ortalamasıdır. Burada dikkat edilecek husus $\theta_i = h(\mu_i, \lambda)$ dır ve yeniden parametrelemede en çok olabilirlik tahminleri değişmekle aynı ilişkiyi sağlamaktadır. Buna göre, iki yanlı dönüştürmeden bağımlı değişkenin dönüştürülmesine ya da tersi olarak, yeniden tahmin yapmaksızın sonuçların çevrilmesi mümkündür. Aynı zamanda, x'ler veri iken, her iki model de y'nin aynı dağılım tahminlerini vermektedir.

$H_0 : \mu_1 = ... = \mu_k$ sıfır hipotezini ancak ve ancak $H_0 : \theta_1 = ... = \mu_k$ sıfır hipotezinin doğruluğu halinde doğru olduğundan, bir grup etkisini test ederken modeller eşdeğer olmaktadır.

Tahmin amaçları için iki model farklı farklı davranışa sahiptir. (14) modelindeki $\hat{\theta}$ büyük varyanslara sahip olacaktır ve $\hat{\lambda}$i le de, bir hayli korelasyonlu olacaktır. bakınız Bickel ve Doksum (1981, 296-311, Box ve Cox (1982) ve kesim 3. (14) modelini Hinkley ve Runger (1984, 302-8), üzerine şartlı olmasını önermektedir. veri iken $\hat{\lambda}$'nin şartlı varyansı küçüktür fakat tahmin edilen parametre E [$\hat{\theta}_i$ $\hat{\lambda}$] olup stokastiktir. Model (13) de, $\hat{\mu}$ 'nün örneklem değişkenliği düşük olmakta ve $\hat{\lambda}$ ile yüksek bir korelasyon içerisinde olmaktadır. Bu durum, λ bilinmese dahi μ_i 'nin iktisadi bir anlamı olduğu gerçeği ile ilgilidir. Eğer (13) modeli kullanılıyorsa, Hinkley ve Runger'in şartlı yaklaşımının gereksiz olduğu görülüyor.

İYD modelinin bir başka takdimi Snee (1986) tarafından yapılmaktadır. Snee'nin bu mükemmel makalesinde bir dizi ilginç uygulamalara yer verilmektedir.

Box ve Cox'un "bağımlı değişkenin dönüştürülmesi" yöntemi ile İYD yöntemi arasındaki ilişkiyi araştıran bir kaç makale mevcuttur. Carroll (1982, 149-52) aşırı uçların tahmin edilen dönüştürme parametreleri üzerinde önemli etkilerinin olduğu bir örnek vermektedir. Carroll bu makalesinde, diyagnostik ve robast prosedürlere olan ihtiyaca işaret etmektedir. Robast Box - Cox tahminini içeren makaleler arasında Andrews (1971, 249-54) , Carroll (1983, 411-16), ve Bickel and Doksun (1981) sayılabilir. Carroll, $\hat{\lambda}$ 'nın belirli bir kümesi örneğin (-1, -1/2, 0, 1/2, 1) 'e kısıtlandığı durumların muhtemel etkilerini tartışmaktadır ki bu yaygın bir uygulamadır. Carroll (1983, 217-22) , ve Doksum and Wang (1980,1224-33) dönüştürme çerçevesinde regresyon parametreleri hakkında yapılan hipotezlerin testini tartışmaktadır. Carroll ve Ruppert (1981, 609-15; 1984b, 312-13) , x'ler veri iiken, dönüştürülmemiş y'nin ön tahmini ile ilgilenmekte özellikle y'nin şartlı medyanın tahmini üzerinde durmaktadır.

3.HETEROSKEDASTİSİTE, ÇARPIKLIK VE DÖNÜŞTÜRME

Dönüştürme yönteminin akıllıca bir kullanımı, bunların normaldışılık varyans heterojenliği üzerindeki etkisinin anlaşılmasını gerektirir. Bu kesimde, varyans ortalamanın bir fonksiyonu olduğunda dönüştürmenin heteroskedastisiteyi nasıl gidermekte olduğu ve dönüştürmenin içbükey ya da dışbükey oluşunun çarpıklık üzerindeki etkileri araştırılmaktadır.

Varsayalım ki, bazı g fonksiyonları ve bir σ sabiti için, y_i tesadüfi değişkeni μ_i ortalama ve $\sigma_i{}^2$ varyansına sahiptir öyleki, $\sigma_i = \sigma_g(\mu_i)$ gibi fonksiyonel bir ilişki mevcuttur. Bartlett (1947, 39-52) de gösterildiği gibi eğer y_i değişkeni $h(y_i)$' ye dönüştürülürse o zaman $h(y_i)$'nin varyansına basit bir Taylor açılımıyla şöyle yaklaşılabilir.

$$
\begin{aligned}
\mathrm{Var}(h(y_i)) &\cong E[h(y_i) - h(\mu_i)]^2 \\
&\cong (h'(\mu_i))^2 \quad E[y_i - \mu_i]^2 \\
&= [(\mu_i)]^2 \quad [\sigma_g(\mu_i)]^2
\end{aligned}
$$

burada h', h nin türevini göstermektedir. Bundan dolayı eğer, h (y) ile $[g(y)]^{-1}$ doğru orantılı ise $h(y_i)$'nin varyansı yaklaşık olarak sabit olacaktır.

Tipik olarak g pozitif ve h ise monoton artan olacaktır. h nin monoton oluşu dönüştürmenin yorumunu kolaylaştırmaktadır çünkü, y ile bağımsız değişkenler arasındaki bazı niteliksel ilişkiler örneğin y'nin herhangi özel bağımsız bir değişkenin artan bir fonksiyonu olması, monoton bir dönüştürme durumunda muhafaza edilmektedir. Hatta eğer y için bir öngörü modeli de amaçlanıyorsa bu durumda, h'nin tersi alınabilir bir dönüştürme olması esastır. Genel olarak y_i daha küçük bir varyansa sahip olduğunda veya h nin ikinci türevi birinciye nazaran küçük olduğunda dönüştürme doğrudur denir.

g 'nin üssel bir fonksiyon olduğu birçok önemli çalışma mevcuttur. Öte yandan, h de üssel ya da değiştirilmiş üssel bir fonksiyon olabilir. Örneğin, poisson dağılımlı veriler için $g(\mu) = \mu^{1/2}$ ve buradan $h(\mu)$, $\mu^{1/2}$ nin doğrusal bir fonksiyonu olmakta ve $\lambda = 1/2$ değeri de Box - Cox dönüştürme ailesi içerisinde kalmaktadır. Eğer $g(\mu)$, μ ' ye eşit ise o zaman değişim katsayısı sabittir ve $h'(\mu)$ de $\mu - 1$ ile orantılı olmalıdır. Bu şekilde, log-dönüştürme varyansı kararlı hale getirmektedir. Genel olarak $g(\mu) = \mu^{(1-\lambda)}$ olduğunda o zaman h', $\mu^{(\lambda-1)}$ ile orantılı olmalıdır.

Van Zwet (1964) nisbi olarak (sağa doğru) çarpıklığın bir tanımını vermektedir. Eğer $R(t) = G^{-1}[F(t)]$ konveks (dışbükey) bir konfeksiyon ise G'nin dağılımının F'ye nazaran (sağa doğru) daha çarpık olacağını tanımlamaktadır. Bu düşüncedeki mantığı görmek için, eğer bir x değişkeni F-dağılımlı olacağına dikkat edilmelidir. Konveks bir R fonksiyonu artan birinci türeve sahip olacağından, x'ten y yaklaştıran bir özellikte olacaktır. x'e nazaran y'nin sağ tarafı daha uzun ve sol tarafı daha kısa olacaktır. Bu sebepten van Zwet'in tanımı, çarpıklık kavramının hassas bir matematiksel kavram haline dönüştürülmesidir.

4.SONUÇ

Doğrusal olmayan regresyonda değişken varyans problemi bu çalışmanın eksenini oluşturmaktadır. Bunun için çeşitli dönüştürme yöntemleri incelenmiştir.

Dönüştürmeler heteroskedastisite ve normal-dışılığı giderebilirse de ikisini aynı anda gidermesi garantili değildir. Eşanlı olarak çarpıklığı gidermek için dönüştürme ve varyansı kararlı hale getirmek için ağırlıklandırma gerekebilir. Aşırı uç değerler parametre tahminlerinde önemli bir etkiye sahip olabilirler. Örneğin bir aşırı uç büyük bir çarpıklığa yol açabilir ya da haddizatında normal ve homoskedastik olan bir veri kümesini heteroskedastik bir şekilde gösterebilir. Tersi durumda, sağa doğru bir çarpıklık sol taraftaki aşırı bir uç ile maskelenebilir. Küçük varyans değerlerinde aşırı bir uç değer heteroskedastisiteyi saklayabilir.

Yukarda özetlenen iki yanlı dönüştürme yöntemleri ile parametrelerin tahmini için çeşitli paket programlar geliştirilmiştir. SAS ve SHAZAM yaygın olarak kullanılmaktadır.

SUMMARY

Nonlinear regression models often exhibit combinations of nonconstant variance, outlier and skewness. In this study,we discuss data transformations, particularly the technique called "Transform Both Sides". This technique is usually found very useful when fitting nonlinear regression models to data that exhibit skewness and nonconstant variance.

KAYNAKÇA

Andrews, D.F.(1971) A note on the selection of data transformation. **Biometrika 58,** 249-54.

Atkinson, A.C. (1985) **Plots, Tranasformations and Regression**. Clarendon Press. Oxford.

Barlett, M.S. (1947) The use of transformations, **Biometrics 3, 39-52.**

Bickel, P.J. and Doksum, K.A. (1981) An analysis of transformations revisited.**J.Am.Stat.Assoc**.76, 296 - 311.

Box, G.E.P. and Cox.D.R. (1964) An analysis of transformations **J.R.Stat.Soc.**B26, 211 - 46.

Box, G.E.P. and Cox.D.R. (1982) An analysis of transformations revisited. **J.Am.Stat.Assoc**. 77, 209 -10.

Box, G.E.P. Hill, W.J. (1974) Correcting inhomogenity of variance with power transformation weighting. **Technometrics** 16, 385-9.

Box, G.E.P. and Jenkins, G.M. (1976) **Time Series Analysis: Forecasting and Control**. Holden - Day, San Fransisco.

Carroll, R.J. (1980) A robust method for testing transformations to achieve approximate normality. **J.R.Stat.Soc.**B42, 1224 Carrol, R.J.(1982) Two examples of transformations where there are possible outliers. **Appl. Stat.**31, 149-52.

Carroll, R.J. (1983) Tests for regression panameters in power transformation models. Scand. **J.Stat**. 9, 217-22.

Carroll, R.J. and Ruppert, D. (1981) On prediction and the power transformation family. **Biometrika** 68, 609-15.

Carroll, R.J. and Ruppert, D. (1984a) Power transformations when fitting theoretical models to data, **J.Am.Stat.** Assoc.79,321-8.

Carroll, R.J. and Ruppert, D. (1984b) Discussion of Hinkley and Runger's paper The analysis of transformed data'. **J.Am.Stat.Assoc**.79, 312-13.

Doksum, K.A.and Wong, C.W. (1983) Statistical tests based on transformed data. **J.Am.Stat.Assoc**.78, 411-16.

Egy, D.and Lahiri K. (1979) On maximum likelihood estimation of functional forms and heteroskedasticity. **Economics Lett.**2, 155 - 40.

Harvey, A.C. (1976) **Estimating regression models with multiplicative heteroscedasticity.**

Hinkley, D.V. and Runger, G. (1984) Analysis of transformed data, **J.Am.Stat.Assoc.** 79, 302-8.

Hopwood, W.S.et al (1984) Time series forcasting models involving power transformations,**J.Forcasting** 3,57-61.

Just, R.E. and R.D. (1978) Stocastic Specification of production functions and economic implications, **J.Econometrics** 7, 67-86.

Leech, D. (1975) Tosting the error specification in nonlinear regression. **Econometrica** 43, 719 - 25.

Leurgans, S. (1980) Evaluating laboratory measurement techniques, **Biostatics Casebook** eds. R.G.Miller Jr, B.Efron, B.W.Brown Jr.and L.E.Moses. Wiley, New York.

McCullagh, P. and Nelder, J.A. (1983) **Generalized Linear Model**, Chapman and Hall, New York.

Snee, R.D. (1986) An alternative approach to fitting models when reexpression of response is useful. **J. Qual. Technol.**18, 211-25.

Snee, R.D. and Irr, J.D. (1981) Design of statistical method for the analysis of mutagenesis at the HGPRT locus of cultured chinese hamster ovary cells, Mutation Res.85, 77-93.

van Zwet, W.R. (1964) Convexs Transformations of Random Variables, Mathematisch Centrum, Amsterdam.