İ.Ü.Siyasal Bilgiler Fakültesi Dergisi
No. 23-24 (Ekim 2000-Mart 2001)

DOĞRUSAL OLMAYAN REGRESYON MODELLERİ ÜZERİNE BİR NOT

Arş.Gör.Dr.Atıf EVREN*

Giriş

Çok değişkenli doğrusal regresyon modellerinde (ÇDDRM) tipik olarak değişkenlerden biri bağımlı değişken olarak ele alınır. Bu değişkenin kendi ortalaması etrafında gösterdiği değişim (variation) bir dizi bağımsız değişkenin etkisi aracılığı ile açıklanmaya çalışılır.

ÇDDRM'nin matematiksel işlemlere elverişli olmasının yanısıra, elde edilen parametrelerin genellikle fiziksel, ekonomik vb anlamlarının olması, parametre tahminlerinin yorumlanması aşamasında da araştırmacılara önemli kolaylıklar sunmaktadır.

Sabit bir sıcaklıkta, V gerilimine maruz kalan ve içinden I akımı geçen bir maddenin elektrik direnci; Ohm Yasası (V=IR) ile tahmin edilebilir. Değişik V değerlerine karşılık I değerleri ampermetreden okunarak R direncinin tahmini, bir doğrusal regresyon modeli ile gerçekleştirilebilir. Elektrik iletkenliği maddenin ayırdedici bir özelliği olduğu için bununla ilişkili bir kavram olan direncin parametre tahmininin yorumlanması kolaydır.

Yine $C = \beta_0 + \beta_1 Y$ gelir-tüketim modelinde (Y geliri, C de tüketimi göstermek üzere) β_0 parametresi Y=0 seviyesindeki tüketim miktarına, β_1 de marjinal tüketim eğilimine ($\frac{dC}{dY} = \beta_1$) eşit olacağı için parametre tahminlerinin yorumlanması zahmetsizce gerçekleştirilebilmektedir.

---

* İstanbul Üniversitesi, İktisat Fakültesi, Ekonometri Bölümü.

1.1.Doğrusal Modeller

ÇDDRM isminin çağrıştırdığından daha geniş bir pratik alanına sahiptir. Bunun önemli bir nedeni İstatistik'te ve Ekonometri'de doğrusal model kavramı ile parametrelerinde doğrusal olan modelin anlaşılmakta olmasıdır. Değişkenlerinde ve parametrelerinde doğrusal olan modeller ÇDDRM 'nin geleneksel araçları ile irdelenebilir. Yine parametrelerinde doğrusal, değişkenlerinde doğrusal olmayan modeller için de (değişkenlerinde uygun dönüşümlere gidildikten sonra) ÇDDRM 'nin araçlarından yararlanılabilir. Parametrelerinde doğrusal olmayan modellerden bazıları da uygun dönüşümlerle doğrusallaştırılabilmektedir[1].

1.2.ÇDDRM

$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon$ (1.2.1) denklemi ile ifade edilmektedir. Burada Y bağımlı değişkenin gözlem değerleri vektörünü, $X_1, X_2, \dots, X_k$ k tane bağımsız değişkenin gözlem değerleri vektörlerini, $\beta_0, \beta_1, \dots, \beta_k$ k tane parametreyi , ε ise modele rastlantısal olma özelliğini veren anakütle hata payı vektörünü simgelemektedir. Güven aralığı hesabı ya da parametrelerle ilgili hipotez testleri aşamasında, normal dağılım ya da onunla ilintili diğer olasılık dağılımlarını kullanabilmek için anakütle hata payının 0 ortalama ve σ^2 varyansı ile normal dağıldığı varsayılmaktadır.

Yukarıdaki model değişkenlerinde ve parametrelerinde doğrusal bir modeldir. Bununla birlikte sözgelimi,

$\sqrt{Y} = \beta_0 + \beta_1 X_1^2 + \beta_2 \log X_2 + \beta_3 Sin X_1 X_3 + \beta_4 e^{X_4} + \varepsilon$ gibi bir model de değişkenlerinde uygun dönüşümlere gidilerek doğrusallaştırılabilir.

$Y^* = \sqrt{Y}, X_1^* = X_1^2, X_2^* = \log X_2, X_3^* = Sin X_1 X_3, X_4^* = e^{X_4}$

denilecek olursa,

$Y^* = \beta_0 + \beta_1 X_1^* + \dots + \beta_4 X_4^* + \varepsilon$ dönüştürülmüş doğrusal modeli elde edilecektir.

---

[1] Bazı doğrusallaştırılabilir fonksiyonlar için bkz: Hocking, R.R., "Methods and Applications of Linear Models: Regression and the Analysis of Variance", John Wiley 1996, s383-385

Bazı modeller de parametrelerinde doğrusal değildir. Örneğin teorik kimyada sabit bir sıcaklıkta tutulan bir gazın hacmi (V) ile basıncı (P) arasında şu ilişki vardır:

$PV^{\gamma} = c$ (c bir sabit olmak üzere)

Burada γ her gaza göre farklı değerler alan, gazların ayırdedici bir özelliğini gösteren bir sabittir. Eğer Y=P ve $X = V^{-1}$ dönüşümleri gerçekleştirilecek olursa

$Y = c\,X^{\gamma}$ modeli elde edilecektir.

Eşitliğin her iki yanının logaritması alınacak olursa

$\log Y = \log c + \gamma \log X$ bağıntısı elde edilerek model doğrusallaştırılmış olur.

Öte yandan $Y = \beta_0 + \beta_1 e^{\beta_2 X}$ modeli β_0, β_1 parametrelerinde doğrusal olmasına karşılık, β_2 parametresinde doğrusal değildir. Uygun bir doğrusallaştırma yöntemi olmadığı için de doğrusallaştırılamamaktadır.

Matematiksel güçlüklerin yanısıra ya bazı ilişkilerin doğrusal olmaması nedeni ile ya da doğrusallaştırma halinde parametrelerin yorumlanmasında bazı sorunlar çıkabileceği gerekçesi ile doğrusallaştırma tercih edilmeyebilir. Bunun yanısıra anakütle hata payı ile ilgili varsayımlar bir dönüşüm sonrasında büyük bir olasılıkla geçerliliğini yitireceği için de doğrusallaştırma tercih edilmemektedir. Böyle durumlarda doğrusal olmayan optimizasyona gidilmektedir.

1.3.Anakütle Hata Payının Dağılımı

$Y = e^{\alpha+\beta X} + \varepsilon$ modeli için anakütle hata payının Y'nin ortalaması ile doğru orantılı olduğu (multiplicative) varsayılsın. Başka bir deyişle

$Y = e^{\alpha+\beta X}(1+\varepsilon_0) = e^{\alpha+\beta X} + \varepsilon_0 e^{\alpha+\beta X} = e^{\alpha+\beta X} + \varepsilon$ olsun.

$(\varepsilon = \varepsilon_0 e^{\alpha+\beta X})$

Yine $E(\varepsilon_0) = 0$ ve $Var(\varepsilon_0) = \sigma_0^2$ olarak varsayılsın. Bu durumda

$Var(\varepsilon) = Var(\varepsilon_0 e^{\alpha+\beta X}) = Var(\varepsilon_0)E(Y)) = \sigma_0^2(E(Y))^2$ olacaktır. Bu

eşitlikten hareketle anakütle hata payının varyansının X'in ve dolayısıyla Y'nin bir fonksiyonu olduğu söylenebilir.
Logaritma alınacak olursa

$$\log(Y) = \alpha + \beta X + \log(1 + \varepsilon_0)$$

$$= \alpha + \beta X + \varepsilon_0^*$$

Küçük ε_0^* değerleri için $E(\varepsilon_0^*) \cong E(\log(1 + \varepsilon_0) \cong 0$

Yine $Var(\varepsilon_0^*)$'ın X'den bağımsız olduğu düşünülebilir.[2]

Bütün bunlara ek olarak ε_0'ın (ve ε'un) normal dağılması halinde ε_0^*'ın normal dağılmayacağı söylenebilir. Öte yandan ana kütle hata payı sona eklenen (additive) biçiminde modelde yer alırsa

$$Y = e^{\alpha + \beta X} + \varepsilon_0$$

$$Y = e^{\alpha + \beta X}(1 + \frac{\varepsilon_0}{E(Y)})$$ ve bu durumda logaritma alınırsa

$$\log(Y) = \alpha + \beta X + \log(1 + \frac{\varepsilon_0}{E(Y)})$$

$$= \alpha + \beta X + \upsilon_0, \quad (\upsilon_0 = \log(1 + \frac{\varepsilon_0}{E(Y)}) \quad olmak \quad üzere)$$

Eğer $\varepsilon_0 \cong 0$ ise $E(\upsilon_0) = 0$ 'dır. Bununla birlikte $Var(\upsilon_0)$'ın E(Y)'nin bir fonksiyonu olacağı da söylenmelidir.

Özetlemek gerekirse ana kütle hata payının modele çarpan (multiplicative) olarak dahil edildiği birinci durumda logaritma almak hata payının varyansını stabilize ederken; hata payının modele eklenen (additive) olarak dahil edildiği ikinci durumda ise varyansı stabilize etmemektedir. Dolayısıyla dönüşüm yapılırken ana kütle hata payının dönüşümden nasıl etkileneceğini hesaba katmak gerekmektedir.[3]

---

[2] Seber, G.A.F., Wild, C.L., "Nonlinear Regression", John Wiley, 1989, s15-16
[3] a.g.y

Öte yandan dönüşüm yapma isteğinin doğrusallığı sağlamak, yaklaşık normal dağılan bir hata payı elde etmek ve sabit varyanslı bir hata payı elde etmek gibi üç temel nedeni olabilir. Bazen doğrusallığı sağlamak için gerçekleştirilen bir dönüşüm diğer hedeflere ulaşılmasını da sağlamaktadır. Bunun dışında sabıt varyanslı hata payı elde etmek için ağırlıklı en küçük kareler (AEKK) yönteminden ya da bazı dönüşümlerden (örneğin, Box-Cox Dönüşümleri) yararlanma yoluna gidilebilir.

2.1.Doğrusal Olmayan Modeller

Bazı modeller doğrusal değildir. Bunun bir nedeni bazı matematiksel modellerin diferansiyel denklemlerle ifade edilmesi, ve bir çok diferansiyel denklemin de kaçınılmaz olarak doğrusal olmayan modellere yol açmasıdır. Örneğin bir bakteri kolonisindeki bakteri sayısının zamana bağlı olarak üstel bir biçimde değişeceği düşünülebilir. Başka bir ifade ile sözkonusu koloni için zamana bağlı olarak gerçekleşen büyüme oranı (y koloni büyüklüğünü göstermek üzere) dy/dt ile gösterilecek olursa, bu oranın o anki büyüklük (bakteri sayısı) ile doğru orantılı olacağı öngörülebilir.

$\frac{dy}{dt} = \beta y$ (β bir sabit olmak üzere) diferansiyel denklemi düşünülebilir.

Bu denklemin çözümü ise

$y = \alpha \exp(\beta t)$ türünde bir modele yol açacaktır. Burada α, t=0 anındaki bakteri sayısını göstermektedir. β ise birim zamandaki bakteri sayısının yüzde ya da oransal olarak ne kadar arttığını gösterir. Dolayısıyla burada parametrelerin fiziksel anlamları oldukça yalındır. Eğer uygun bir doğrusallaştırma yöntemi kullanılsaydı; büyük bir olasılıkla dönüştürülmüş parametrelerin yorumlanması daha problemli olacaktı!

Bu tür bir büyüme denklemi başka büyüme problemlerinde de kullanılmaktadır. Örneğin, bankadaki bir tasarruf hesabının zamana göre değişimi $F = P\exp(rn)$ denklemine göre incelenebilir. (Burada F paranın gelecek değerini, P şimdiki değerini, r nominal faiz oranını, n de geçen periyod ya da dönem sayısını göstermektedir.) Yine bir kapasitörden geçen akımın zamana göre değişimi veya vücuttan atılan bir ilacın vücuttaki mıktarının zamana göre değişimi de benzer denklemlerle incelenmektedirler.[4]

2.2. Doğrusal Olmayan Modellerde Parametre Tahminleri

Doğrusal olmayan bir model

$Y_i = f(X_i;\theta) + \varepsilon_i$ (2.2.1) biçiminde ifade edilmektedir.[5] Burada f doğrusal haldeki gibi beklenen değer fonksiyonunu, X bağımsız değişkenler vektörünü; θ da parametreler vektörünü ifade etmektedir. ε da yine doğrusal halde olduğu gibi ana kütle hata payını simgelemektedir. Yine ε 'nun $E(\varepsilon) = 0$, *ve* $Var(\varepsilon) = \sigma^2 I$ olmak üzere normal dağıldığı varsayılmaktadır.

Bu model, beklenen değer fonksiyonunun, parametrelerin doğrusal olmayan bir fonksiyonu olması istisnası dışında, doğrusal model ile aynıdır. Böyle bir durumun sonucunda doğrusal olmayan regresyon modelinde beklenen değer fonksiyonunun parametrelere göre alınan türevlerinden en az bir tanesinin parametrelerden en az bir tanesine bağımlı olacağı söylenmelidir.[6]

Doğrusal olmayan durumda da parametrelerin EKK tahminleri kalıntı kareleri toplamı

$S(\theta) = \sum_{i=1}^{n} (Y_i - f(X_i;\theta))^2$ 'nın (2.2.2) minimize edilmesi ile bulunur.

Burada, doğrusal regresyondan farklı olarak $S(\theta)$ 'nın birden fazla göreli (veya yerel) minimuma sahip olabileceğinin belirtilmesi gerekir. S(θ*), $S(\theta)$'yı minimize eden değer, $\hat{\theta}$ da θ* 'nın herhangi bir tahmini değeri olsun. Yine ε_i değerlerinin; beklenen değerlerinin sıfır ve varyanslarının σ^2 olmak üzere; bağımsız ve özdeş dağıldıklarını

---

[4] Jennrich, R.I., "An Introduction to Computational Statistics&Regression Analysis", Prentice-Hall, 1995, s239

[5] Çok kullanılan bazı nonlineer modeller için bkz.,Ross,G.J.S., "Nonlinear Estimation";Springer Series in Statistics;1990, s1-10

[6] Bates, D. M.,Watts, D.G., "Nonlinear Regression Analysis and Its Applications", John Wiley,1988, s2

varsayarsak, $\hat{\theta}$ ve $S^2 = \frac{S(\hat{\theta})}{n-p}$ 'nın (2.2.3) θ^* ve σ^2 'nın tutarlı birer tahmincileri olduklarını söyleyebiliriz.

$n \to \infty$ olması halinde ise $\hat{\theta}$ 'nın asimptotik olarak normal dağıldığını ve bunlara ek olarak ana kütle hata payının normal dağılımı varsayımı altında $\hat{\theta}$ 'nın değerinin aynı zamanda maksimum olabilirlik tahmincilerine eşit olacağı söylenmelidir.

Matris notasyonu kullanılacak olursa

$S(\theta) = [Y - f(\theta)]' [Y - f(\theta)] = \|Y - f(\theta)\|^2$ (2.2.4) yazılabilir.

Burada Y nx1'lik bağımlı değişkenin gözlem değerleri vektörü $f(\theta)$ da nx1'lik bağımlı değişkenin beklenen değerleri vektörüdür.

$$f_i(\theta) = f(X_i;\theta) \quad ve \quad f(\theta) = (f_1(\theta); f_2(\theta); \ldots; f_n(\theta))'$$

$$F.(\theta) = \frac{\partial f}{\partial \theta'} = \left[\left(\frac{\partial f_i(\theta)}{\partial \theta_j}\right)\right] \quad olsun. \quad Yine$$

$$F. = F.(\theta^*) \quad ve \quad \hat{F} = F.(\hat{\theta})$$

biçiminde yazılsın. Bu durumda EKK tahmincilerini bulmak için kalıntı kareleri toplamı fonksiyonunun θ parametre vektörüne göre türevi alınıp sıfır vektörüne eşitlendiğinde

$\left.\frac{\partial S(\theta)}{\partial \theta_r}\right|_{\hat{\theta}} = 0$ (2.2.5) bulunur. Başka bir deyişle $\hat{\theta}$ vektörünün bir yerel minimumu verebilmesi için yukarıdaki eşitliği sağlaması gerekmektedir. Bu denklemden hareketle

$\left.\sum \{Y_i - f_i(\theta)\} \frac{\partial f_i(\theta)}{\partial \theta_r}\right|_{\theta=\hat{\theta}} = 0 \quad i = 1,2,3,\ldots,p$ (2.2.6) normal denklemleri elde edilir.

Matris notasyonu kullanılırsa

$$\hat{F}'\{Y - f(\hat{\theta})\} = 0 \qquad (2.2.7)$$

$$\hat{F}'\hat{\varepsilon} = 0$$

Bu denklemler doğrusal olmayan bir model için normal denklemler olarak adlandırılırlar ve pek çok doğrusal olmayan model için analitik olarak çözülemezler. Dolayısıyla tekrara dayalı (iterative) hesaplama yöntemleri kaçınılmaz olmaktadır.

2.3.Doğrusallaştırma

θ nın optimum değeri olan θ^* 'nın küçük bir komşuluğunda; beklenen değer fonksiyonunun birinci dereceden Taylor serisi açılımı yapılacak olursa

$$f_i(\theta) \cong f_i(\theta^*) + \sum_{r=1}^{p} \left.\frac{\partial f_i}{\partial \theta_r}\right|_{\theta^*} (\theta_r - \theta_r^*) \quad (2.3.1) \text{ ya da}$$

$$f(\theta) \cong f(\theta^*) + F.(\theta - \theta^*) \qquad (F. = F.(\theta - \theta^*) \text{ olmak üzere})$$

(2.3.2)

yazılabilir. Dolayısıyla

$$S(\theta) = \left\| Y - f(\theta) \right\|^2 \cong \left\| Y - f(\theta^*) - F.(\theta - \theta^*) \right\|^2 \quad (2.3.3) \text{ ve}$$

eğer

$Z = Y - f(\theta^*) = \varepsilon \quad \beta = \theta - \theta^*$ denilecek olursa

$S(\theta) = \left\| Z - F.\beta \right\|^2$ (2.3.4) eşitliği elde edilir. Doğrusal regresyon ile analoji kurularak $\hat{\beta}$ parametrelerinin EKK tahminleri

$\hat{\beta} = (F'.F.)^{-1} F'.Z$ (2.3.5) denklemi ile bulunmaktadır.

Belli koşullar altında $\hat{\theta}$ hemen hemen θ^* 'nın küçük bir komşuluğunda bulunduğundan

$\hat{\theta} - \theta^* \cong \quad \hat{\beta} = (F'F)^{-1} F' Z$ (2.3.6) denklemi ile minimizasyonun gerçekleşeceği söylenebilir. Dahası $\theta = \hat{\theta}$ eşitliğinden de yararlanılacak olursa

$$\begin{aligned} f(\hat{\theta}) - f(\theta^*) &\cong F(\hat{\theta} - \theta^*) \\ &\cong F(F'F)^{-1} F' Z \\ &\cong P_F Z \\ &\cong P_F \varepsilon \end{aligned} \qquad (2.3.7)$$

bulunur. Ve

$$\begin{aligned} Y - f(\hat{\theta}) &\cong Y - f(\theta^*) - F(\hat{\theta} - \theta^*) \\ &\cong \varepsilon - P_F \varepsilon = (I_n - P_F)\varepsilon \end{aligned} \qquad (2.3.8)$$ bulunur.

Burada $P_F = F(F'F)^{-1} F'$ ve $I_n - P_F$ simetrik ve idempotent projeksiyon matrisleridir. Yine doğrusal halde olduğu gibi P_F beklenen değer yüzeyine teğet olan düzleme projeksiyon yapılarak bağımlı değişkenin tahmin değerlerini bulmamıza ; $I_n - P_F$ de beklenen değer yüzeyine dik olan düzleme projeksiyon yapılarak kalıntı değerlerinin bulunmasına olanak vermektedir. Yine

$$(n-p)s^2 = S(\hat{\theta}) = \left\| Y - f(\hat{\theta}) \right\|^2 \cong \left\| (I_n - P_F)\varepsilon \right\|^2 = \varepsilon'(I_n - P_F)\varepsilon$$

(2.3.9) eşitliğinden hareketle ana kütle hata payının varyansına ilişkin tahmin gerçekleştirilmektedir.

3.1.Bir Doğrusal Olmayan Regresyon Problemi Olarak Genelleştirilmiş EKK

Doğrusal regresyonun doğrusal olmayan regresyonun özel bir hali olduğu düşüncesi şaşırtıcı gelebilir. Yine de genelleştirilmiş EKK ya da ağırlıklı EKK problemlerinin doğrusal olmayan EKK problemi olarak çözülebilmesi şüphesiz doğrusal olmayan regresyon tekniklerine önemli bir uygulama alanı sunmaktadır.

Genelleştirilmiş ya da ağırlıklı EKK'de minimize edilecek fonksiyon

$S(\theta) = [Y - f(\theta)]' V^{-1} [Y - f(\theta)]$ (3.1.1) biçimindedir.

Burada V pozitif belirli bir matristir ve değişen varyans (heteroscedasticity) durumuna bağlı olarak her bir gözlem noktasına farklı ağırlık verilmesini sağlayan matristir. Yine modelde ana kütle hata payına ilişkin olarak $E(\varepsilon) = 0$ ve $D(\varepsilon) = \sigma^2 V$ varsayımlarında bulunulmaktadır. Burada D ana kütle hata payına ilişkin varyans-kovaryans matrisini simgelemektedir. Yine $V = I_n$ olması halinde ağırlıklı EKK'in olağan EKK yöntemine dönüştüğünü belirtmek gerekir. Başka bir ifade ile EKK yötemi genelleştirilmiş EKK'in özel hali olmaktadır.

Öte yandan V matrisi U bir üst-üçgen matris olmak üzere $V = U'U$ Cholesky ayrıştırmasına tabi tutulsun. Yine

$R = (U')^{-1}$ olsun.

$Y = f(\theta) + \varepsilon$ denkleminin her iki yanı R matrisi ile soldan çarpıldığında

$RY = Rf(\theta) + R\varepsilon$; yine $Z = RY, k(\theta) = Rf(\theta), \eta = R\varepsilon$ olmak üzere

$Z = k(\theta) + \eta$ doğrusal olmayan modeli elde edilecektir. Bu durumda kalıntı kareleri toplamı

$$S(\theta) = [Y - f(\theta)]' V^{-1}[Y - f(\theta)]$$

$$= [Y - f(\theta)]' R'R[Y - f(\theta)] \qquad (3.1.2)$$

$$= [Z - R(\theta)]' [Z - R(\theta)]$$

olmaktadır. Dikkat edilirse yukarıdaki denklem doğrusal olmayan regresyon modelinden elde edilen kalıntı kareleri toplamı fonksiyonu ile tamamen aynıdır. Dolayısıyla genelleştirilmiş EKK problemlerini olağan doğrusal olmayan EKK problemi olarak formüle etmek olasıdır.[7]

3.2.Otokorelasyonlu Modeller

Bir dizi regresyon modelinde zaman içerisinde toplanan verilerle ilgili grafik gösterimler uzun dönemli pozitif ve uzun dönemli negatif kalıntıları sergilemektedir. Bu durum, modelin yetersiz formüle edilmesinden ya da ardışık hata terimleri arasındaki yüksek korelasyondan bulunmasından kaynaklanabilmektedir.

---

[7] Bu yöntemin algoritması için bkz. Seber, a.g.e., s30

Birinci dereceden otoregresif süreç AR(1)

$\varepsilon_i = \rho\varepsilon_{i-1} + a_i$ (3.2.1) biçiminde ifade edilmektedir. Burada a_i 'lerin korelasyonsuz, ve $E(a_i) = 0$ (3.2.2)

ve $Var(a_i) = \sigma_a^2$ (3.2.3) eşitliklerinin geçerli olduğu varsayılmaktadır. Yine $|\rho| < 1$ 'dir. Böyle bir durumda

$Corr[\varepsilon_i ; \varepsilon_j] = \rho^{|i-j|}$ 'dir.(3.2.4) Dolayısıyla ε_i ile ε_j arasındaki korelasyon Y_i ile Y_j arasındaki zaman aralığı büyüdüğünde üstel olarak azalmaktadır.

$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ i=1,2,3,....,n modeli için ana kütle hata payının AR(1) modeline uygun davrandığı ve $a_i = \varepsilon_i - \rho\varepsilon_{i-1}$ olduğu varsayılsın. Bu durumda

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

$$\rho Y_{i-1} = \rho\beta_0 + \rho\beta_1 X_{i-1} + \rho\varepsilon_{i-1}$$

$$Y_i - \rho Y_{i-1} = (1-\rho)\beta_0 + \beta_1 X_i - \rho\beta_1 X_{i-1} + a_i \quad (3.2.5)$$

Bu model ise

$Y_i = \beta_0^* + \rho X_{i1}^* + X_{i2}^* + \beta_1\rho X_{i3}^* + a_i$ (3.2.6) biçimine dönüştürülebilir. Dolayısıyla otokorelasyonlu hata terimine sahip bir doğrusal model; korelasyonsuz, sabit varyanslı hata terimine sahip bir doğrusal olmayan modele dönüştürülebilir. Sonuçta elde edilen ve doğrusal olmayan bu model de doğrusal olmayan EKK yöntemleri ile incelenebilir.

3.3.Bağımsız Değişkenlerde Dönüşüm

Ana kütle hata payı ile ilgili normal dağılma ve sabit varyans varsayımlarının yaklaşık olarak gerçekleştiği kabulü altında X ve Y arasındaki ilişkinin eğrisellik (curvature) derecesi araştırılsın. Yine burada bağımsız değişkende veya değişkenlerde dönüşüme gitmek sureti ile model biçimlendirmesinin gerçekleştirilmeye çalışıldığı varsayılsın.

Bu problemle ilgili olarak Box ve Tidwell (1962) X'in ne türde bir dönüşüme tabi tutulacağına ilişkin önerdikleri prosedür de doğrusal olmayan regresyona bir uygulama alanı sunması bakımından ilginçtir.[8] Bu prosedür çok değişkenli modeller için de uygulanabilir olmasına karşılık buradaki tartışma çerçevesinde iki değişkenli model ile kısıtlı olarak ele alınacaktır.

Bağımlı değişkenin bağımsız değişkenin bir kuvveti ile ilişkili olduğu varsayılsın. Başka bir deyişle

$\xi = X^{\alpha}$ ve yine

$$\xi = \begin{cases} X^{\alpha}, \alpha \neq 0 \\ \ln X, \alpha = 0 \end{cases} \quad \text{olmak üzere}$$

$$E(Y) = f(\xi, \beta_0, \beta_1) = \beta_0 + \beta_1 \xi \quad (3.3.1) \text{ kabul edilsin.}$$

β_0, β_1 ve α da değerleri tahmin edilecek parametreler olsun. Yine α_0'ın α'nın bir ilk tahmini olduğu varsayılsın. (Genellikle iterasyonlara $\alpha_0 = 1$ alınarak başlanmaktadır.) Bu ilk tahmin noktasında beklenen değer fonksiyonunun Taylor serisi açılımı yapılacak olursa

$$\begin{aligned} E(Y) &= f(\xi_0, \beta_0, \beta_1) + (\alpha - \alpha_0) \left\{ \frac{df(\xi_0, \beta_0, \beta_1)}{d\alpha} \right\}_{\substack{\xi = \xi_0 \\ \alpha = \alpha_0}} \\ &= \beta_0 + \beta_1 X + (\alpha - 1) \left\{ \frac{df(\xi_0, \beta_0, \beta_1)}{d\alpha} \right\}_{\substack{\xi = \xi_0 \\ \alpha = \alpha_0}} \end{aligned} \quad (3.3.2)$$

Burada eğer { } içindeki ifade bilinseydi, β_0, β_1 ve α'yı EKK ile bulmak mümkün olacaktı. Daha sonra da α , α_0 yerine konarak,

[8] Aktaran Montgomery, D.C.,Peck,E.A., "Introduction to Linear Regression Analysis", Second Edition, John Wiley, 1992, s105

daha iyi bir tahmin noktasına geçilmiş olabilirdi. Bu noktadan hareket etmek ve parantez içerisindeki ifadeyi şu şekilde yeniden ele almak mümkündür:

$$\left\{\frac{df(\xi,\beta_0,\beta_1}{d\alpha}\right\}_{\substack{\xi=\xi_0\\ \alpha=\alpha_0}} = \left\{\frac{df(\xi,\beta_0,\beta_1}{d\alpha}\right\}_{\xi=\xi_0} \left\{\frac{d\xi}{d\alpha}\right\}_{\alpha=\alpha_0} \quad 3.3.3)$$

Öte yandan

$$\left\{\frac{df(\xi,\beta_0,\beta_1)}{d\xi}\right\}_{\xi=\xi_0} = \beta_1 \text{ ve } \quad \left\{\frac{d\xi}{d\alpha}\right\}_{\alpha=\alpha_0} = X\ln X$$

yazılabilir.

Dolayısıyla beklenen değer fonksiyonunun Taylor serisi açılımı

$E(Y) = \beta_0 + \beta_1 X + (\alpha - 1)\beta_1 X \ln X$ (3.3.4) biçiminde de ifade edilebilir.

$W = X\ln X$, $\gamma = (\alpha - 1)\beta_1$ denilecek olursa

$E(Y) = \beta_0 + \beta_1 X + \gamma W$ (3.3.5) denklemi elde edilecektir. Bu denklemden hareketle γ 'nın EKK tahmini gerçekleştirildikten sonra yeni α tahmini de

$\hat{\alpha} = 1 + \frac{\gamma}{\beta_1}$ (3.3.6) yardımıyla yapılmaktadır.

(Burada β_1'in $\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X$ den hareketle bulunduğu belirtilmelidir.)

Bulunan α 'nın tahmin değeri α_1 olarak adlandırılacak ve aynı hesaplama daha düşük bir kalıntı kareleri toplamı veren α_2'nin bulunması için tekrarlanacaktır. Box ve Tidwell buradaki yakınsamanın hızlı olduğunu ve genellikle birinci iterasyondan sonra α 'nın tatmin edici bir değerinin bulunduğunu belirtmektedirler.[9]

---

[9] a.g.y.

4.1.Uygulama

Bu aşamada doğrusal olmayan regresyon modellerinin , doğrusal modellerde ortaya çıkan bazı sorunların çözümünde kullanılabileceğine ilişkin bir uygulama yapılmıştır. Uygulamaların verileri Bates ve Watts'ın kitabının sonundaki verilerden alınmıştır.Bu veriler Newyork Cayuga Gölü'ndeki PCB (polichlorinated biphenyl) maddesinin kalıntıların yoğunluğu ile ilgilidir. Yine bu veriler kitabın doğrusal regresyon ile ilgili kısımlarında kullanılmaktadır. Orijinal veri seti şöyledir:

Y	X
0,6	1
1,6	1
0,5	1
1,2	1
2	1,25962998
1,3	1,25962998
2,5	1,25962998
2,2	1,441721509
2,4	1,441721509
1,2	1,441721509
3,5	1,586667686
4,1	1,586667686
5,1	1,586667686
5,7	1,709058826
3,4	1,816035636
9,7	1,816035636
8,6	1,816035636
4	1,911690788
5,5	1,911690788
10,5	1,911690788
17,5	1,998614186
13,4	1,998614186
4,5	1,998614186
30,5	2,07856091
12,4	2,222203177
13,4	2,287532932
26,2	2,287532932
7,4	2,287532932

Burada bağımsız değişken X yıl olarak geçen zamanın küpkökünü , bağımlı değişken Y ise PCB maddesinin göl içindeki yoğunluğunu ifade etmektedir. Bu iki değişkenli doğrusal regresyon modelinin oluşturulmasından elde edilen özetleyici istatistikler ise şöyledir:

Özetleyici İstatistikler

Regresyon İstatistikler	
Çoklu- R	0,70913742
R Kare	0,502875881
Ayarlanmış R Kare	0,483755723
Standard Hata	5,393871936
Gözlem Sayısı	28

Varyans Analizi

	Ser.Derecesi	*Kareler Toplamı*	*Ortalama Kareler*	*F*
Regresyon	1	765,1922841	765,1923	26,30082
Kalıntı	26	756,4402159	29,09385	
Toplam	27	1521,6325		

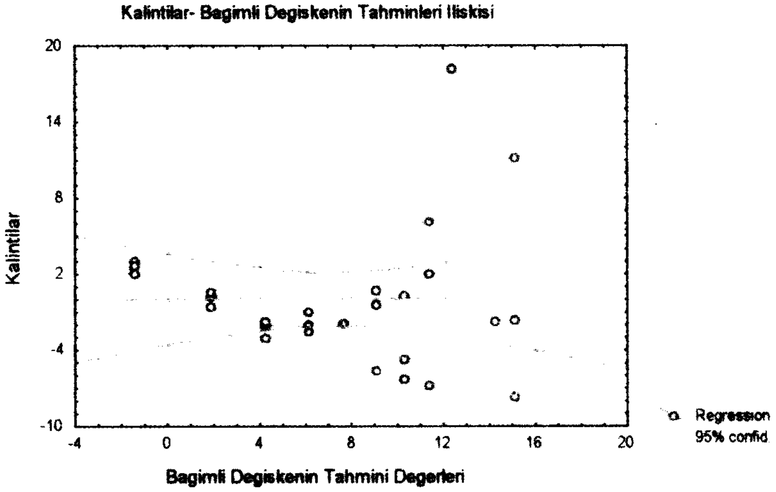
	Katsayılar	*Standard Hata*	*t -Değeri*	*P-Değeri*
Kesişim	-14,30118816	4,3099482	-3,31818	0,002684
Birinci Değişken	12,81736856	2,499275964	5,128433	2,4E-05

Uygulama ile ilgili diğer istatistikler de şöyledir:

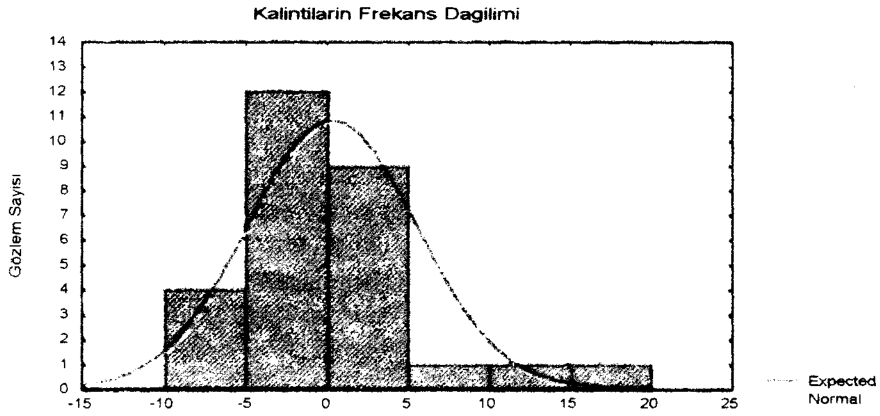
Durbin-Watson d ve Kalıntıların Serisel Korelasyonu

	Durbin- Watson d	Serisel Korelasyon
Tahmin	1,611332	-,255154

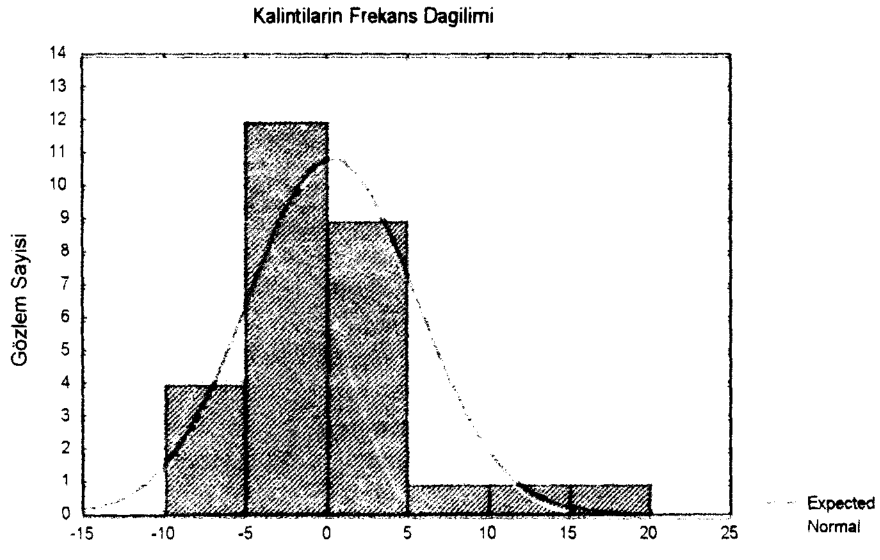
Yine aynı uygulama ile ilgili olarak kalıntıların bağımlı değişkenin tahmini değerleri karşısında gösterdiği değişimi ifade eden STATİSTİCA çıktısı şöyledir.



Yukarıdaki veri sonuçları için öncelikle hem EXCEL ve hem de STATİSTİCA çıktılarının elde edildiğini ve yine regresyon analizi sonuçlarına bakıldığında EXCEL sonuçları ile STATİSTİCA sonuçlarının birbirine yakın değerler verdiğini belirtmek gerekiyor. Özetleyici istatistiklerin elde edilmesinde EXCEL kullanılmasının nedeni STATİSTİCA'ya göre daha derli toplu tablolar elde edilebilmesidir. Bununla birlikte, hem Durbin-Watson ve serisel korelasyon rakamlarının elde edilmesinde ve hem de eldeki veriler ile modelin varsayımlarının (lineerlik, sabit varyans, anakütle hata payının normal dağılımı) uyumsuzluğunun tespit edilmesinde STATİSTİCA çıktılarından yararlanılmıtır. Bilindiği gibi EXCEL bu tür istatistikleri standard olarak sunmamaktadır.



Yukarıdaki verilerden hareketle modelin doğrusal olarak kabul edilmesi ile yapılan bir spesifikasyonun hatasının var olduğuna hükmedilebilir. Hem örneğin serisel korelasyonun -0,255 gibi bir değer alması ve hem de yukarıdaki grafikten hareketle kalıntı değerlerinin bağımlı değişkenin tahmini değerlerinin artışına bağlı olarak belirli bir artış göstermesi sabit varyans varsayımına ters düşmektedir. Yine kalıntıların dağılımına bakıldığında da şu STATİSTİCA çıktısı elde edilmektedir:



Burada da kalıntıların asimetrik bir dağılıma sahip olduğu görülmektedir. Bu noktadan hareketle ana kütle hata payının da normallikten uzak (asimetrik) dağılma özellikleri göstereceği düşünülebilir.

Doğrusal Olmayan Bazı Modellerin Denenmesi

Yukarıdaki yetersizlikten hareketle uygulamanın ikinci aşamasında doğrusal olmayan bir model (üstel büyüme modeli) için 10 civarında farklı parametre tahmin noktalarında denemeler yapılmış ve sonuçlar aşağıya aktarılmıştır.

Den .No	Denenen Model	Parametre Tahminleri	R	Açıklama
1	Y=a+exp(b+c*X)	a=-4,52; b=0,316,c=1,21	---	İyi uyum sağlanamadı. (İUS)
2	Y=a+b*exp(c*X)	a=-4,52; b=1,37;c=1,21	0,73	"İll-conditioning" problemi dolayısıyla İUS
3	Y=-4,52+b*exp(c*X)	b=1,36 c=1,21	0,73	"İll-conditioning" problemi dolayısıyla İUS
4	Y=a+1,37*exp(c*X)	a=-4,52 c=1.21 standard hatalar: s(a)=1,57; s(b)=0,055	0,73	Normale yakın dağılım özelliği gösteren kalıntılar
5	Y=a+exp(c*X)	a=-3,51; c=1,33 s(a)=1,52; s(c)=0,056	0,73	Normale yakın dağılım özelliği gösteren kalıntılar
6	Y=a+b*exp(X)	a=-6,80; b=2,41 s(a)=2,72; s(b)=0,44	0,73	Normale yakın dağılım özelliği gösteren kalıntılar
7	Y=b*exp(2,304*X)	b=0,091	0,71	Normale yakın dağılım özelliği gösteren kalıntılar
8	Y=-6.81+b*exp(c*X)	b=2,36; c=1,01	0,73	"İll-conditioning" problemi dolayısıyla İUS
9	Y= -6,81+2,41*exp(c*X)	c=1,001 s(c)=0,03	0,73	Normale yakın dağılım özelliği gösteren kalıntılar
10	Y=-6,80+b*exp(X)	b=2,41 s(b)=0,155	0,73	Normale yakın dağılım özelliği gösteren kalıntılar

Bilindiği gibi "ill-conditioning problemi" X'X projeksiyon matrisinin yapısına ilişkin bir problemdir ve sözkonusu matrisin determinantının sıfır ya da sıfıra yakın çıkması durumunda parametrelerin standard hataları çok büyük olacağı için parametrelerin güven aralığı tahminleri gerçekleştirilememektedir. Dolayısıyla ikinci, üçünce ve sekizinci modeller tatmin edici sonuçlar vermemektedir. Yine birinci modelin de tatmin edici olmadığı görülmektedir.

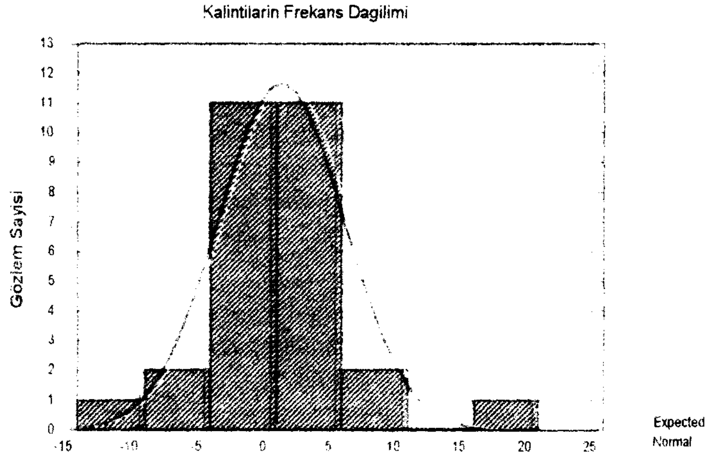
Diğer modeller ile elde edilen istatistiklerdeki bazı farklılıklara ise çok fazla takılmamak gerekir. Çünkü doğrusal olmayan modellerde birden fazla optimum olabilir. Bunun dışında parametrelere farklı başlangıç noktaları verilmesi de farklı optimum noktalarına ulaşılmasını sağlayabilir. O halde bu seçenekler içerisinde seçim yapılırken parametrele tahminlerinin fiziksel olarak anlamlılığı, daha yalın bir model seçilmiş olması, daha fazla serbestlik derecesine sahip olmak gibi "ikincil" ama hiç te önemsiz olmayan faktörlerden hareket etmek gerekir. Bu yüzden de yukarıdaki modellerden yedinci sırada olanı tercih edilmelidir. (Bu modelde bilinmeyen parametre sayısının diğerlerine göre daha azdır.Doğrusal olmayan modellerde doğrusal modellerden farklı olarak bir veri setinden hareketle birden çok çözüm bulmak mümkün olduğu için bazı özetleyici istatistiklerdeki küçük farklılıklar gerekçesi ile göz ardı edilerek daha basit olan bu model tercih edilmiştir.) Bu model için elde edilen sonuçlar ise şöyledir:

Model: Y=b*EXP(2.304*X)
BAĞIMLI DEĞİŞKEN: Y AMAÇ FONKSİYONU: EN KÜÇÜK KARELER(EKK)
EKK'NIN NİHAİ DEĞERİ:745,70851737 R=.71409
AÇIKLANAN DEĞİŞİM: 50,993%

b PARAMETRESİNİN TAHMİNİ: .098240

s(b)=0.01 ve t-değeri=8.96

Yine kalıntıların sıklık dağılımı ile ilgili görsel çıktı şu biçimdedir:



4.2.Sonuç:

Doğrusal olmayan regresyon doğrusal regresyonun genel hali olmasının yanısıra , doğrusal regresyonda karşılaşılan sorunlara da çözüm bulmak amacıyla gündeme getirilebilir. Birinci olarak genelleştirilmiş EKK problemlerinin, doğrusal olmayan EKK problemleri olarak formüle edilebildiğini belirtmek gerekir. İkinci olarak otokorelasyonlu hata payına sahip modeller yine doğrusal olmayan regresyon modellerine dönüştürülerek matematiksel bir biçimlendirme hatasının olup olmadığı sınanabilir. Son olarak model spesifikasyonuna ilişkin kimi problemlerin çözümünde (değişkenler arasındaki ilişkinin doğrusal olmadığının bilindiği ve bağımsız değişkende/değişkenlerde dönüşüme giderek doğrusal EKK modelinin varsayımları ile uyumlu bir yeni model yaratılmak istendiğinde) doğrusal olmayan EKK yöntemlerine başvurmak önerilebilir. Yine de doğrusal olmayan modellerin modelin varsayımları ile veriler arasındaki uyumsuzlukları eş zamanlı olarak düzeltmediğini, yukarıdaki örnekten hareketle söylenecek olursa bazı sapmaları düzeltirken (kalıntıların sıklık dağılımının nispeten simetrik hale getirilmesi) bazı sapmalara ise (DW, serisel korelasyon ile ifade edilen model biçimlendirme hatası gibi) çözüm getirememektedir.

Kaynaklar

1)Bates, D.M.,Watts, D.G., "Nonlinear Regression Analysis and Its Applications", John Wiley, 1988

2)Hocking, R.R., "Methods and Applications of Linear Models: regression and the Analysis of Variance", John Wiley, 1996

3)Jennrich, R.I., "An Introduction to Computational Statistics & Regression Analysis" , Prentice-Hall, 1995

4)Montgomery, D.C., Peck, E. A.; "An Introduction to Linear Regression Analysis" ;Second Edition, John Wiley, 1992

5)Ross, G.J.S., "Nonlinear Estimation" , Springer Series in Statistics, 1990

6)Seber, G.A.F.;Wild , C.L., "Nonlinear Regression" ; John Wiley, 1989