DEĞİŞKEN SAYISININ BELİRLENMESİ

Yrd. Doç. Dr. Osman Z. ÇETİNKAYA*

Özet

Regresyon analizinde açıklayıcı değişken sayısının belirlenmesi önemli bir sorundur. Bu amaçla, açıklama gücündeki kazançla serbestlik değerindeki azalma arasında bir tercih yapan uygunluk ölçütleri kullanılmaktadır. Modelin karmaşıklığı ve uyum arasında tercihe dayanan başka ölçütler de vardır. Bu yazı, bu ölçütlerin kısa bir özetini vermekte ve otoregresyonla ilgili bir ölçüt sunmaktadır.

Abstract

In regression analysis, to determine the number of explanatory variables is a serious problem. For this purpose, goodness of fit measures take into account the trade-off between gain in explanatory power and loss in degrees of freedom. There are other ways to measure this trade-off between model complexity and goodness of fit. This article gives a brief introduction to them and proposes a criteria which concerns about autoregression.

Tahmin edilen modelin veriye uygunluğunu belirlemek, ekonometrik çalışmaların önemli bir sorunu olmuştur. Uyumun ölçülmesinin nedeni, daha iyi uyumun daha iyi öngörü sağlayacağı beklentisidir. Uyumun iyiliği, öngörünün etkinliğini dolaysız olarak yansıtacaktır. Araştırmacı bu noktada bir dilemma ile karşı karşıya kalır; doğru model bilinmiş olsaydı, model seçme sorunu ortadan kalkardı. Farklı modeller denemek amacıyla bağımsız değişken sayısı artırıldıkça uygunluk artacak, ancak bu takdirde istenmeyen bir durum ortaya çıkacak ve modelin serbestlik derecesi azalacaktır. Uygunluk ve serbestlik dereceleri için optimum bir düzey bulabilmek model seçiminin temel sorunudur. Bu çalışmada açıklayıcı değişken sayısının belirlenmesi sorunu üzerinde durulacaktır.

---

* *İ.Ü. Siyasal Bilgiler Fakültesi.*

Adım-adım regresyon yöntemi ilk akla gelen bir yöntem olarak kullanılabilir. Bu yöntem her bağımsız değişkeni anlamlılık düzeyi temelinde değerlendirir. Her bağımsız değişken, bağımlı değişkenle yüksek bir korelasyona, dolayısıyla yüksek bir t istatistiğine sahip olabilir. Tek tek ele alındıklarında t istatistiklerinin yüksek olduğu a priori olarak söylenebilir. Ancak, bağımsız değişken sayısı arttıkça, bağımsız değişkenler arasındaki korelasyonlar (Çoklu Bağıntı) nedeniyle bu t istatistiğinin uygun bir yorumunun yapmak imkânsız hale gelir. Çoklu bağıntı nedeniyle, t istatistiği küçüldüğü ve/veya işaret değiştirdiği zaman, en son eklenen değişkenin modele alınmaması, yani modelin bir önceki adımda tamamlanmış olduğunu kabul etmek gibi bir çözüm önerilebilir. Uygulamada anlamlı olabilecek böyle bir yöntem, dışarıda bırakılan değişkenlerin etkisinin görülmesini imkânsız kılar ve değişkenlerin hangi sırayla modele eklenmesi gerektiği sorularına cevap vermekten de uzaktır.

1. Uygunluk ölçütleri

Tahmin edilen modelin veriye uygunluğu için bir ölçüt, regresyon denklemi tarafından açıklanan kısmın bağımlı değişkendeki toplam değişkenliğe oranı olan R^2 değeridir: $R^2 = SSR / SST$. Modele açıklayıcı değişken olarak eklenen her yeni değişken R^2 değerini artırır fakat serbestlik derecesinin düşmesine sebep olur. R^2 değeri her zaman artacaktır, çünkü az da olsa eklenen yeni değişkenle bağımlı değişken arasında bir korelasyon vardır. Bu nedenle R^2 değerinin artması, modelin daha uygun hale geldiğinin bir ölçütü olamaz. Hatta, çok yüksek R^2 değerleri, modelin stokastik olduğu varsayımına da aykırı düşer. Eklenen yeni değişken aynı zamanda yeni bir katsayı tahmini demektir ve modelin tahmin veya açıklama gücünü artırdığı da söylenemez. Bu sorun, zaman serileri analizinde, ARIMA modellerinin p ve q parametrelerinin belirlenmesi sorunuyla da özdeştir. Başka bir ifadeyle, kalıntıların kareler toplamında (SSE) bir azalma ile, serbestlik derecesindeki kayıp arasında bir seçim yapma sorunu ortaya çıkar. Serbestlik derecesinin yüksek olması, bağımlı değişkenlerdeki değişmelerin daha az değişkenle açıklanması demektir ki, bu da "parsimony" ilkesine daha uygundur. Bütün bilimsel disiplinlerde, modelin daha az varsayım ve daha az değişkenle açıklanması istenilen bir ilkedir.

Sorun açıklayıcı değişken sayısının belirlenmesi olarak da ifade edilebilir. Bu amaçla geliştirilen çeşitli ölçütlerden ilki, Theil'in önerdiği düzeltilmiş R_d^2 değeridir. Bilindiği gibi bu ölçüt,

$$R_d^2 = 1 - S_e^2 / S_y^2$$

olarak tanımlanmıştır ve modelin açıklanamayan kısmını, kalıntıların varyansının, bağımlı değişkenin varyansına oranı, yani toplam değişkenliğin yüzdesi olarak, S_e^2 / S_y^2 olarak ifade eder. Se^2 ve S_y^2 ifadeleri yerine konursa,

$$R_d^2 = 1 - \left(1 - R^2\right)(n-1)/(n-k)$$

eşitliği elde edilir[1]. Modele yeni açıklayıcı değişkenler eklendiği takdirde k artacak, R_d^2 değeri de, k'nın belli bir değerinden sonra düşmeye başlayacaktır. k'daki bir artışın, her zaman R^2 değerini artırıcı, R_d^2 değerini ise, k'nın belli bir değerinden sonra azaltıcı bir etkisi olacaktır. Bu ölçüte göre model, R_d^2'nin maksimum olduğu k değeri için belirlenir. Ancak bu ölçüt, açıklayıcı değişkenler arasındaki korelasyon (Çoklu Bağıntı) ve kalıntılar arasındaki korelasyon (Ardışık Bağımlılık) sorunları için yön gösterecek bir ölçü vermez, sadece SSE'deki değişmelere dayanır.

R_d^2'ye benzer bir ölçüt Goldberger[2] tarafından önerilen "uyarlanmış R_u^2" ölçütüdür:

$$R_u^2 = (1 - k/n)\, R^2$$

Amemiya[3], 1985'de önerdiği "Öngörü ölçütü"nü "Prediction Criterion (PC)";

$$PC = S_e^2(1 + k/n) \tag{1}$$

olarak ifade etmiştir. Amemiya, Theil'in önerdiği düzeltilmiş R_d^2 değerinin teorik gerekçelerini yetersiz bulmuş, bu değerin k'nın artmasına eğilimli olduğunu iddia etmiştir. Ayrıca standart olarak kullanılan *minimax* ilkesinin yerine *minimini* (minimum riski minimum yapma) ilkesini tercih etmiştir. "Koşulsuz tahmin hataları kareleri ortalaması" olarak,

$$E\left(\hat{y}_i - y_i\right)^2 = S^2(1 + k/n)$$

ifadesini tanımlar[4], ve buradan hareket ederek (1) eşitliğindeki tanımı elde eder.

$$R_{PC}^2 = (n + k)/(n - k)\left(1 - R^2\right)$$

1 Gujarati, Damodar N., "Temel Ekonometri", Literatür yay. Çev: Ümit Şenesen-Gülay Günlük Şenesen, 1999.

2 Goldberger, Arthur S., "A Course in Econometrics", Harvard University Press, 1991.

3 Amemiya, Takeski, "Advanced Econometrics", Basil Blackwell Ltd., 1985, s. 49.

4 Amemiya, Takeski, age. s. 40.

değeri de model seçiminde kullanılır. PC değerini minimum yapmakla, R_{PC}^2 değerini maksimum yapmak aynı anlamdadır. (1) tanımından hareketle, PC değerindeki artışı (p), SSE'deki artışla (e) ifade etmek mümkündür;

$$1+p=(1+e)\ \{(n+k+1)/(n+k)\cdot(n-k)/(n-k-1)\} \qquad (1a)$$

Görüldüğü gibi, PC değerindeki artış, n ile birlikte k'ya da bağlıdır. k'ya bağlı olması, PC ölçütünü, aşağıda ele alınacak diğer ölçütlerden ayıran önemli bir özelliktir.

2. Model seçme ölçütleri

R^2 ve R_u^2 daha çok modelin verilere uygunluğunu gösteren ölçütler olarak kullanılır. Aşağıda ele alacağımız ölçütler ise model seçme ölçütleri olarak adlandırılmaktadır. Uygunluk ölçütleri açıklama gücündeki artışla serbestlik derecesindeki azalma arasında bir tercihe dayanırken, model seçme ölçütlerinin, modelin karmaşıklığı (complexity) ve uyumun iyiliği arasındaki tercihe dayandığı söylenebilir. Bu ölçütler hata terimlerinin kareleri toplamını (SSE) veya gözlenen olabilirlik değerlerini minimum kılmak gibi optimallik ilkelerine dayanırlar. Bütün bu ölçütler, SSE'nin ortalamasının bir "ceza" faktörü ile çarpılmasına dayanır. "Ceza" faktörü de, tahmin edilecek regresyon katsayılarının sayısı (k) ile ölçülen modelin karmaşıklığı ile ifade edilir. Hepsinde de, ölçütün daha küçük değer aldığı model tercih edilir.

Bu ölçütler tarih sırası ile aşağıdaki gibi özetlenebilir[5]:

Akaike FPE (Finite Prediction Error) ölçütü (1970):

$FPE = \ln(SSE / n) + \ln(n+k / n-k)$

AIC (Akaike Information Criterion) ölçütü (1973):

$AIC = \ln(SSE / n) + 2k / n$

Schwarz Ölçütü (1978):

$SC = \ln(SSE / n) + (k / n) \ln n$

Hannan ve Quinn ölçütü (1979):

$HQ = SSE / n \quad (\ln n)^{2k / n}$

5 Ramanathan, Ramu, "Statistical Methods in Econometrics", Academic Press Inc., 1993, s.280-28.

Craven ve Wakka'nın önerdiği GCV (Generalized Cross Validation) ölçütü (1979):

$$GCV = SSE / n \quad (1 - k / n)^{-2}$$

Shibata ölçütü (1981):

$$Shibata = \ln (SSE / n) + \ln (n+k / n)$$

Rice ölçütü (1984) :

$$Rice = \ln(SSE / n) - \ln(1 - 2k / n)$$

H. Akaike'nin[6] önerdiği "Akaike Information Criterion (AIC)" bu konuda en çok kullanılan ölçütlerden birisidir. Akaike, $L(\alpha) = S^{-n}$ olarak tanımladığı $L(\alpha)$ fonksiyonunu, α_1, α_2 ... modellerine uygulamıştır. $L(\alpha_1)$, $L(\alpha_2)$ ... "Olabilirlik Oranı – Likelihood Ratio" testinden hareket ederek, aşağıdaki kayıp fonksiyonunu önermektedir:

$$\omega(\theta,\widehat{\alpha})=-\frac{2}{n} \int \left[\ln \frac{L(\widehat{\alpha})}{L(\theta)} \right] L(\theta)\, dx$$

ω kayıp fonksiyonu, farklı modeller arasında minimum yapılacak bir fonksiyon olarak kullanılmıştır. Akaike'nin çözümü;

$$AIC = (-2 / n) \ \ln L(\alpha) + 2 k / n$$

olarak bulmuştur. Bu ifadede k, α vektörünün boyutuna karşılık gelmektedir. $L(\alpha) = S^{-n}$ değeri yerine konursa,

$$AIC = \ln (SSE / n) + 2 k / n \tag{2}$$

eşitliğine ulaşılır. Bu ölçüt, modele yeni bir değişken eklemenin marjinal katkısını, yani SSE'deki azalmayı, k'daki artma, yani serbestlik derecesindeki azalma ile kıyaslamaktadır. k artıkça ilk terim sürekli olarak azalacak, ikinci terim ise artacaktır. AIC ölçütü artışa mutlak, azalmaya ise logaritmik bir ağırlık vermek suretiyle artışı, yani k'nın etkisini daha önemli kabul etmiş olmaktadır. Değişken sayısı k_0'dan $k_1 = k_0 + 1$'e arttığı zaman, (2) denklemi;

$$e^{AIC_0} = SSE_0 / n \quad e^{2k / n}$$

$$e^{AIC_1} = SSE_1 / n \quad e^{2 (k+1) / n}$$

şeklinde ifade edilebilir. Bu eşitlikler taraf tarafa bölünürse;

6 Akaike, H., "Information Theory and an Extension of the Maximum Likelihood Principle", 1973.

$$e^{AIC_1 - AIC_0} = SSE_1 / n \quad n / SSE_0 \; e^{2(k+1) - 2k / n} = \frac{SSE_1}{SSE_0} e^{2/n}$$

$$AIC_1 - AIC_0 = \ln \left(\frac{SSE_1}{SSE_0} \right) + 2 / n \tag{2a}$$

AIC ölçütündeki artış, yani $\Delta AIC = AIC_1 - AIC_0$ elde edilir. SSE_1 / SSE_0 oranı, SSE'deki artış oranıdır. Bu oran e ile gösterilirse, (2a) eşitliği de;

$$\Delta AIC = \ln (1 + e) + 2 / n \tag{2b}$$

olarak yazılabilir. Her zaman $SSE_1 < SSE_0$ eşitsizliği geçerli olduğundan, $e < 1$'dir. n sabit olduğundan, "ceza"nın artışı da sabittir: 2/n.

Genel kabul gören bir diğer ölçüt, Schwarz ölçütüdür "Schwarz Criterion (SC)"[7]. Bu ölçüt de;

$$SC = \ln SSE / n + k (\ln n) / n \tag{3}$$

olarak tanımlanmıştır. AIC ölçütüne benzer şekilde, burada da k'daki bir artışın SC değerinde meydana getireceği artış;

$$\Delta SC = \ln (1 + e) + (\ln n) / n \tag{3a}$$

olarak ifade edilebilir.

AIC ve SC ölçütlerinde, gözlem sayısı n değişmeden, her ikisinin de küçük değer aldığı modelin tercih edilmesi gerekir. Başka bir deyişle, bu ölçütlerin daha küçük değer aldığı modelin daha iyi uyum sağladığı kabul edilir[8]. Eklenen yeni değişkenin açıklama gücünün azalması, AIC ve SC değerlerinin artmasına sebep olacaktır. Her iki ölçüt de, bağımsız değişkenlerin sayısı arttıkça SSE'de ortaya çıkacak azalmaya bir "ceza" verir.

İki ölçütü karşılaştırmak amacıyla, (2b) ve (3a) eşitliklerinin farkı alınırsa,

$$\Delta SC - \Delta AIC = (\ln n - 2) / n \tag{4}$$

bulunur. Bu eşitlik ise, $\ln n > 2$, yani $n > 8$ olduğu sürece SC ölçütünün, AIC ölçütünden daha fazla değişeceğini gösterir. Başka bir ifadeyle, SC her zaman (genellikle $n > 8$ olacağından) daha "parsimonious" modeli seçecektir[9]. Çünkü, (4)

7 Schwarz, G., "Estimating the Dimension of a Model", Annals of Statistics, 6, 1978.
8 Greene, William H., "Econometric Analysis", Prentice-Hall, Inc., 1997, s. 400.
9 Johnston, Jack – Dinardo, John; "Econometric Methods", McGraw-Hill, forth ed., 1997.

eşitliğinden açıkça görüldüğü gibi, $n > 8$ için, iki ölçütteki değişme arasındaki fark pozitif olacağından, yeni bir açıklayıcı değişken eklemenin marjinal etkisi, SC'de, AIC ölçütüne göre daha büyük olacaktır. Ayrıca, "SC ölçütü büyük örnek özelliklerinde daha üstündür. SC asimptotik olarak tutarlıdır. AIC ölçütü ise parametresi fazla olan modellere doğru sapmalıdır." [10]

Klasik süreci kullanmak, araştırmacı için metodolojik bir dilemmaya yol açar. Bir yandan a priori olarak bilinen doğru bir model varsayımı vardır. Model doğru ise, araştırmacı sadece bu basit modeli hesaplamalı ve sonucu yorumlamalıdır. Diğer yandan uygulamada her zaman araştırma niteliğinde çalışmalar yapmak zorunlu olur. Tamamen teorik bir şekilde kurulan modeli irdelemek ve başka değişkenleri de denemek her zaman mümkündür.

Bu ölçütlerin çok sağlam teorik temelleri olduğu iddia edilememektedir. Her birinin doğruluk ve eksiklikleri vardır. Mesela AIC ölçütünün, n'in büyük değerleri için aşırı uygunluğa doğru sapmalı bir eğilimi vardır. Buna rağmen ardışık yapılan tahminlerde çıkarsama sorunundan etkilenmez. F testleri de uygun anlamlılık düzeyinin ardışık olarak gözden geçirilmesini gerektirir; fakat istatistiksel olarak uygun değişken sayısını olması gerekenin altında belirlemeye eğilimlidir.

3. Ardışık bağımlılık sorununun ölçüte eklenmesi

Düzeltilmiş R^2 için söylediğimiz sakıncaları bir kez daha tekrar edersek, yukarıda ele alınan bütün ölçütler, açıklayıcı değişkenler arasındaki korelasyon (Çoklu Bağıntı) ve kalıntılar arasındaki korelasyon (Ardışık Bağımlılık) sorunları için yön gösterecek bir ölçü vermezler, sadece SSE'deki değişmelere dayanırlar. Oysa, özellikle ardışık bağımlılık sorunu model seçiminde SSE'deki düşüş veya R^2'deki artış kadar öneme sahiptir.

Ardışık bağımlılığın, buna yol açan en önemli nedenin, açıklayıcı bir değişkenin modelde yer almaması (model kurma hatası) olduğu hatırlanırsa, değişken sayısının belirlenmesiyle çok yakından ilgili olduğu açıktır. Bu amaçla, aşağıda,ardışık bağımlılık katsayısını, değişken sayısının belirlenmesinde dikkate alan bir ölçüt denenecektir. Ardışık bağımlılık katsayısı;

$$\rho = \Sigma e_t e_{t-1} / \Sigma e_t^2 \tag{5}$$

veya; basitlik amacıya $E = | \Sigma e_t e_{t-1} |$ denirse,

$$\rho = E / SSE \tag{5a}$$

10 Endres, W., "Applied Econometric Time Series", John Wiley and Sons, Inc., 1995, s.88.

şeklinde ifade edilebilir. E terimini mutlak değer olarak tanımlamakta, ardışık bağımlılığın yönü ile değil, sadece büyüklüğü ile ilgilenildiği sürece bir sakınca yoktur. Değişken sayısı arttıkça SSE'nin azalacağı bilinmektedir. SSE'nin azalması da ρ'nun eğilim olarak artmasına yol açar. Ancak E hakkında sadece, E'nin başlangıçta büyük olacağı, değişken sayısı arttıkça, ardışık bağımlılığa sebep olan açıklayıcı değişkenin dışarıda bırakılması sorunu ortadan kalkacağı için giderek azalacağı, fakat bir minimum değerden sonra tekrar artmaya başlayacağını a priori olarak söyleyebiliriz. Dolayısıyla E terimi bağımsız bir değişken gibi ele alınabilir. E'nin artışı ρ üzerinde belirleyici bir pozitif etkiye sahip olacaktır. ρ'daki artış oranına ς, E'deki artış oranına da ε denirse,

$$1 + \varsigma = (1 + \varepsilon) / (1 + e)$$

yazılabilir. Buradan;

$$1 + e = (1 + \varepsilon) / (1 + \varsigma) \qquad (6)$$

eşitliğine ulaşılır. AIC ölçütünü için yazılan (2b) eşitliğinde bu değer yerine konursa,

$$\Delta AIC = \ln (1 + \varepsilon) - \ln (1 + \varsigma) + 2 / n \qquad (7)$$

eşitliğine ulaşılır. Bu eşitlik bize, AIC ölçütünün ardışık bağımlılığı da ifade edeceğini göstermektedir.

Eğer AIC'e benzer, yalnızca ardışık bağımlılıktaki değişmeyi ifade eden yeni bir ölçüt aranırsa (A), (2) denklemindeki AIC ölçütünün tanımına uygun olarak,

$$A = \ln (E / SSE) + 2k / n \qquad (8)$$

ölçütü tanımlanabilir. Bu tanımdaki ilk terim $|\rho|$'ya eşittir ve ardışık bağımlılığı yansıtmaktadır. k arttıkça ilk terim azalacak, bir minimum değerden geçtikten sonra "dalgalı" olarak artmaya başlayacaktır. k artmaya devam ettikçe, ilk terimin tekrar azaldığı gözlenebilir. Bu nedenle, sürekli bir artış seyrinden ayırt etmek için, "dalgalı" bir artış deyimi tercih edilmiştir. İlk terimdeki bu dalgalı seyir, E'nin etkisinden kaynaklanmaktadır. AIC ölçütüne benzer şekilde, bu ölçütün minimum değerini aldığı ilk k değeri, ardışık bağımlılığın minimum olduğu optimum açıklayıcı değişken sayısını verecektir. İkinci terim ise her zaman k'nın artan bir fonksiyonudur.

(2) ve (8) eşitlikleri taraf tarafa toplanırsa;

$$AIC + A = \ln\left(\frac{E}{SSE}\right) + \ln\left(\frac{SSE}{n}\right) + 4k/n$$

$$= \ln(E/n) + (4k/n)$$

eşitliği geçerlidir. k'daki bir artışın, yani eklenen yeni değişkenin bu ölçüte etkisi de, (2a) eşitliğinin elde edilmesine benzer şekilde;

$$\Delta A = \ln(1+\varepsilon) - \ln(1+e) + 2/n \tag{8a}$$

ifadesiyle hesaplanabilir. (2b) eşitliğine dayanarak,

$$\Delta A + \Delta AIC = \ln(1+\varepsilon) + (4/n)$$

eşitliği elde edilir. Bu son eşitlik, E terimindeki (veya değişkenindeki) artış oranının logaritmasının, iki farklı ölçütteki değişmelerin toplamı ile doğru orantılı olduğu şeklinde de yorumlanabilir. E'deki değişkenlik fazla olduğu için, ρ ve A ölçütlerinde, diğer ölçütlere nispetle değişkenlik fazla olacaktır.

4. Uygulama

Türkiye'de enflasyonu açıklamak amacıyla, aşağıdaki beş denklem denenmiştir. Enflasyon DİE TEFE (Toptan Eşya Fiyatları Endeksi)'ndeki artışlarla ile ifade edilmiş, açıklayıcı değişkenler olarak da TL/USD kuru (Kur) ve para arzının (M2) gecikmeli değerlerindeki artışlar kullanılmıştır. 1994 yılı başından 2002 yılı ilk çeyreğine kadar 33 dönemi kapsayan ekteki veriler, www.die.gov.tr adresinden elde edilmiştir.

Denenen denklemler aşağıdaki gibidir:

I $TEFE_t = \beta_0 + \beta_1 Kur_t + e_t$

II $TEFE_t = \beta_0 + \beta_1 Kur_t + \beta_2 Kur_{t-1} + e_t$

III $TEFE_t = \beta_0 + \beta_1 Kur_t + \beta_2 Kur_{t-1} + \beta_3 Kur_{t-2} + e_t$

IV $TEFE_t = \beta_0 + \beta_1 Kur_t + \beta_2 Kur_{t-1} + \beta_3 Kur_{t-2} + \beta_4 M2_t + e_t$

V $TEFE_t = \beta_0 + \beta_1 Kur_t + \beta_2 Kur_{t-1} + \beta_3 Kur_{t-2} + \beta_4 M2_t + \beta_5 M2_{t-1} + e_t$

Denklemlerin özet çözümleri aşağıdaki tabloda verilmiştir:

Denklem	1 R^2_d	2 S_e	3 sd	4 n	5 k	6 SSE	7 F	8 E	9 ρ	10 AIC	11 A	12 Dw
I	0,110	8,849	30	32	2	2349,3	4,8	2010,1	0,856	4,421	-2,231	0,289
II	0,246	4,622	28	31	3	598,1	5,9	60,3	**0,101**	3,116	-2,107	1,798
III	**0,305**	**4,481**	26	30	4	522,1	5,2	67,5	0,129	**3,042**	-1,795	1,741
IV	0,277	4,569	25	30	5	522,0	3,8	65,3	0,125	3,104	**-1,766**	1,750
V	0,269	4,596	24	30	6	507,1	3,1	63,3	0,125	3,138	-1,769	1,750

Birinci sütunda yer alan düzeltilmiş belirlilik katsayısı R^2_d ölçü alınırsa, bu değerin en yüksek olduğu III. denklem tercih edilmelidir. Bu denklemde, doğal olarak S_e değeri de en düşük değerini almaktadır. 10. sütunda yer alan AIC değerinin en düşük olduğu denklem de III. denklemdir. Bu ölçütlere göre III. denklemin seçilmesi beklenir.

Ardışık bağımlılık sorunu ele alınırsa, 9. sütundaki ardışık bağımlılık katsayısı ρ'nun birinci denklemde yüksek bir değer aldıktan sonra hızla düştüğü, en düşük değerini de II. denklemde aldığı görülmektedir. Ancak bu değerlerin genellikle düşük olması, I. Denklemden sonra önemli bir ardışık bağımlılık sorunu olmadığını göstermektedir. Nitekim 12. sütunda görülen Durbin-Watson (dw) değerleri, I. denklemde pozitif otokorelasyon olduğunu, diğer denklemlerde ise otokorelasyon olmadığını göstermektedir.

Amaç para arzının enflasyon üzerindeki etkisini de görmek ise, IV. denklem de kullanılabilir. ρ'daki azalışla, bu azalışı sağlayan yeni değişken eklenmesi arasındaki tercihin bir ölçüsü olan A ölçütünün en düşük değerini bu denklemde aldığı görülmektedir. Eğer IV. denklem seçilirse hem para arzındaki değişmelerin etkileri ölçülmüş olacak, hem de dw testi nedeniyle ardışık bağımlılıkta bir sorun çıkmayacaktır.

5. Sonuç

Farklı sorulara cevap vermek amacıyla, ekonometrik çalışmalarda farklı modeller denemek zorunluluğu ortaya çıkmaktadır. Bu amaçla bağımsız değişken sayısı artırıldığı takdirde, modelin uygunluğu artacak, ancak serbestlik derecesi azalacaktır. Uygunluk ve serbestlik dereceleri için optimum bir düzey bulabilmek model seçiminin temel sorunudur. Ardışık bağımlılık gibi ekonometrik modellerin temel varsayımlarını çiğneyen durumlarda da bir tercih yapılması gerekmektedir. Mesela, yukarıdaki örnekte olduğu gibi, M2 değişkeninin etkisini görmek amacıyla AIC ölçütünün önerdiği III. denklem yerine IV. denklemin seçilmesi mümkündür. R^2_d ölçütüne göre ise bir başka denklem seçilmesi gerekebilirdi. Görüldüğü gibi model seçiminde, araştırmacının amacı ile tamamıyla çakışan tek bir ölçüt bulmak mümkün değildir. Bu yazıda önerilen A ölçütü de, ardışık bağımlılıktaki azalmayı bir ölçüt olarak ileri sürmektedir.

EK: VERİLER

		TEFE	Artış %	TL / USD	Artış %	M2	Artış %
1994	1	71		20.584		284.905	
	2	102	43,1	31.725	54,1	469.332	64,7
	3	113	10,4	33.895	6,8	563.013	20,0
	4	135	19,4	37.407	10,4	642.490	14,1
1995	1	168	24,6	41.646	11,3	765.592	19,2
	2	182	8,5	43.055	3,4	979.810	28,0
	3	199	9,4	47.632	10,6	1.126.714	15,0
	4	223	12,1	56.615	18,9	1.270.423	12,8
1996	1	277	24,3	68.150	20,4	1.435.626	13,0
	2	321	15,6	79.497	16,6	1.825.307	27,1
	3	358	11,7	88.671	11,5	2.253.409	23,5
	4	413	15,2	104.392	17,7	2.801.675	24,3
1997	1	491	19,0	124.521	19,3	3.232.451	15,4
	2	563	14,8	143.806	15,5	3.768.239	16,6
	3	664	17,8	169.790	18,1	4.484.355	19,0
	4	788	18,7	199.079	17,3	5.264.529	17,4
1998	1	913	15,9	234.827	18,0	5.985.260	13,7
	2	996	9,1	260.196	10,8	7.781.176	30,0
	3	1.101	10,6	274.473	5,5	9.014.456	15,8
	4	1.215	10,3	306.123	11,5	10.856.763	20,4
1999	1	1.353	11,3	360.513	17,8	13.254.674	22,1
	2	1.497	10,6	411.466	14,1	14.922.280	12,6
	3	1.701	13,7	452.063	9,9	18.343.392	22,9
	4	1.980	16,4	525.491	16,2	22.596.061	23,2
2000	1	2.247	13,5	579.195	10,2	22.491.499	-0,5
	2	2.346	4,4	614.999	6,2	24.680.949	9,7
	3	2.448	4,3	662.536	7,7	26.625.880	7,9
	4	2.626	7,3	676.000	2,0	32.812.563	23,2
2001	1	3.035	15,6	953.108	41,0	38.289.265	16,7
	2	3.796	25,1	1.208.651	26,8	40.365.605	5,4
	3	4.277	12,7	1.463.304	21,1	42.291.251	4,8
	4	4.952	15,8	1.448.212	-1,0	46.985.987	11,1
2002	1	5.388	8,8	1.352.496	-6,6	50.104.303	6,6

Kaynak: www.die.gov.tr

KAYNAKÇA

Akaike, H., (1973), "Information Theory and an Extension of the Maximum Likelihood Principle".

Amemiya,Takeski, (1985), "Advanced Econometrics", Basil Blackwell Ltd.

Endres, W., (1995), "Applied Econometric Time Series", John Wiley and Sons, Inc.

Goldberger, Arthur S., (1991), "A Course in Econometrics", Harvard University Press.

Greene, William H., (1997), "Econometric Analysis", Prentice-Hall, Inc.

Gujarati, Damodar N., (1999), "Temel Ekonometri", Literatür yay. Çev: Ümit Şenesen-Gülay Günlük Şenesen.

Johnston, Jack – Dinardo, John; (1997) "Econometric Methods", McGraw-Hill, forth ed.

Ramanathan, Ramu, (1993), "Statistical Methods in Econometrics", Academic Press Inc.

Schwarz, G., (1978), "Estimating the Dimension of a Model", Annals of Statistics, 6.