EKONOMETRİK MODEL SEÇİM KRİTERLERİ ÜZERİNE KISA BIR İNCELEME

Meltem Şengün UCAL*

Özet

"Hangi değişkenler önemli? , Bir model nasıl seçilir? gibi sorular modellemede çok önemlidir. İyi araştırma (ekonometrik) teknikleri altında iyi bir model kesinlikle verileri uygun tahmin eder. Birden fazla uygun model tanımlamasının bulunduğu durumlarda ekonometrisyen ve istatistikçi mevcut olan veri setinden uygun modeli seçmek ister. Model seçim kriterleride en uygun model kararının verilmesi için bir yoldur.

Bu çalışma farklı model seçim kriterlerinin kısa incelemesini ve birbirleriyle karşılaştırmasını içermektedir. Analiz edilen model seçim kriterleri geleneksel hipotez testine bağlı metodlardan farklı olarak bilgi teorisine dayanmaktadır. Kullback-Leibler uyumsuzluğuna dayanan Akaike Bilgi Kriteri ile bilgi teorisi yaklaşımı 1970'lerde popülerdi. Daha sonraları bu yaklaşım Bayes Bilgi Kriteri(BIC), Schwartz Bilgi Kriteri (SIC), Mallow'un Cp kriteri gibi örneklerle çeşitlenerek gelişmiştir.

Bu çalışmada ayrıca yeniden örnekleme methodlarından bootstrap ve çapraz-geçerlilik.te model seçim kriterleri içinde anlatılmıştır.

Anahtar Kelime: Bilgi Teorisi, Model Seçim Kriterleri, Bootstrap, Çapraz Geçerlilik.

A Brief Survey Of Econometrics Model Selection Criteria

Abstract

"Which variables are important?, How to select a model?" kind of questions are very important for the modeling. A good model certainly fits the well in to the data under investigation (econometrics). The econometrician and statistician would like to select most appropriate model from data sets, where there may be more than one definition of "appropriate". Model selection criteria are one way to decide on the most appropriate model.

This paper surveys briefly the different model selection criteria and compares them with each other. The analyzed model selection criteria are based on the information theory and are quite different from the usual methods based on null hypothesis testing. Information theory approaches were popular in the 1970s with the land mark Akaike Information Criteria based on the Kullback-Leibler discrepancy. Later, those approaches were diversified and such criteria as Bayes Information Criterion (BIC), Schwartz Information Criterion (SCI), and Mallow's Cp were developed.

* Dr., Kadir Has Üniversitesi, İİBF, İktisat Bölümü, Istanbul

In the paper, the resample methods (bootstrap and cross validation) were also explained in the contents of model selection criteria.

Key Words: Information Theory**,** Model Selection Criteria, Bootstrap, Cross-Validation.

I. GİRİŞ

Tüm çalışmalarda ekonometrik modelin doğru kurulduğu varsayımı altında ilerlenir. Klasik doğrusal regresyon modelinin bilindik önsav sınamaları gözden geçirilerek model desteklenmeye çalışılır. Fakat modelin istatistik testlerindeki sorunlar modelin yeniden kurulmasına yönlendirebilir (model kurma hatası yapıldığı düşünülerek). Ekonometristler **doğru** model kurma ve seçme kavramları üzerine uzun yıllardan beri tartışırlar. Tartışmalarda en çok karşılaşılan sorular şöyledir: Hangi değişkenler modelde önemlidir[1]?, İyi bir model nasıl tanımlanır?. Bir model nasıl seçilir?. Tüm bu soruların başlangıç cevapı aslında cimrilik prensibi[2] (The principle of parsimony, Box ve Jenkins, 1977) ile ortaya çıkar. Prensibe gore "verinin en iyi şekilde yansıtılabilmesi için gerekli en az sayıda parametre yani değişken kullanılmalıdır". Böylece baştan bazı sapmalar engellemiş olacaktır. Aslında doğru modelin kurulma aşamasının dikkatli ve en az eğilimle tamamlanması istenmektedir. Bu süreç verilerin toplanması, doğru veri[3]/örnek büyüklüğünün oluşturulması ile başlar. Modelin yapısının doğru kurulup kurulmadığının sınanmasında araştırmacıya yardımcı olacak teknikler ise model yapı testleridir Bunlar, R^2 ve Durbin - Watson d , Ramsey's RESET (Ramsey, 1969), değişken eklemede Lagrange Çarpanı (Engel, 1982), Wald test, White test, vb.sayılabilir.. Modelin yapısının irdelenmesi dışında bizi ilgilendiren diğer bir soru da araştırma için kurulan modeler arasından en iyi performansa sahip olanın nasıl seçilmesi gerektiğidir. Söz konusu soru araştırılırken belirli model kalıpları için düşünülmüş model seçim kriterlerinden faydalanılmalıdır (Mcquarrie, Tsai, 1998). Ancak o zaman daha etkin bir çalışma platformu yaratılabilir. Bu kapsamda model seçiminde en göze çarpan yaklaşımlar Leamer (Leamer, 1978) ve Hendry (Hendry, 1980-1983; SSS:sına,sına,sına ;TTT:test test test mantığına dayanan) yaklaşımlarıdır. Bilindik bu yaklaşımların haricinde Gaver ve Geiels

---

1 Seçilen değişkenlerin ekonomik ve istatistik olarak en etkin olması ile model en iyi performansı sağlar. Ekonomik olarak alt örnek kütlenin bulunmasıyla değişken seçim süreci başlar. Değişkenlerin seçilmesindeki ekonomik kriterler aşağıdaki kriterlerle yakından ilgilidir (Onishi, 1983):

- İşaret kriteri
- Önem kriteri . Bunlar tamamlandıktan sonra istatistiksel kriterler işlemeye başlar.

2 Cimrilik prensibi ek. 'de kısaca anlatılmıştır.

3 (Sample bias, errors of measurement) Örnek büyüklüğünü yalnış seçme veya toplanan verilerden/değişkenlerden kaynaklanan ölçme hataları model kurmada dikkat edilmesi gereken noktalardandır.

(1974), Jeffreys (1961), Kiefer ve Richard (1979), Leamer (1978), Lindley (1957), Poirrier (1981) ve Zellner (1971, 1978, 1979) model seçiminde ön bilginin minimizasyonuna dayalı posterior odds(evvelki şarta bağlı ve veri-örnek bilgisi) kriterini kullanarak model seçimine destek vermeye çalışmışlardır. Bu çalışmaların alternatifi olarak da ekonometriciler rakip modeler arasındaki seçimi kolaylaştırma amaçlı model seçim sınamaları geliştirmişlerdir. Genel olarak sınamalar iki grupta toplanır:

- Yuvalanmış(nested) model sınamaları[4]
- Yuvalanmamış(nonnested) model sınamaları[5]
  - Ölçüte gore ayrıt edici yaklaşım
  - Öteki model bilgisiyle ayırt edici yaklaşım olmak üzere iki türlü incelenir (Harvey, 1990)

Çalışmada yuvalanmış, yuvalanmamış model sınamalarının ayrıntılarına girilmeden, araştırmacı tarafından oluşturulan modellerin karşılaştırılmasını sağlayan seçim teknikleri irdelenecektir. Bu günkü çalışmaların çoğunda da görülen ve en çok kullanılan model seçim kriteri R^2, $\overline{R}^2$ (düzeltilmiş)'dir. Fakat modele alınan değişken sayısının artmasıyla değişen R^2, $\overline{R}^2$ (düzeltilmiş)'nin yanında diğer model seçim kriterlerini kullanmak daha doğru olacaktır. Model seçim kriterleri ilk olarak 60'ların başı ve 70'lerin sonlarında Akaike'nin FPE (Akaike, 1969) ve Mallow'un C_p testleri ile ortaya çıktı. Kullback.-Leibler (Kullback ve Leibler, 1951) uyumsuzluğuna dayanan ve daha çok zaman serilerinde kullanılan Akaike Bilgi kriterinin (Akaike, 1973, 1974) ortaya çıkmasıyla başlayan bu dönüm noktasında 70'lerde bilgi teorileri yaklaşımı görülmeye başladı. 70'lerin sonlarına gelindiğinde bilgi teorileri alanında çalışmalar artarak gelişti ve yeni kriterler doğdu. Bunlardan bazıları: Bayes bilgi kriteri (BIC, Akaike, 1978), Schwarz bilgi kriteri (SIC, Schwarz, 1978), Hannan ve Quinn Kriteri (HQ, Hannan ve Quinn,1979), FPE$\propto$ (Bhansali ve Downham, 1977), GM (Geweke ve Meese, 1981), yeniden örneklemeye dayanan cross-validation ve bootstrap (Efron, 1979) sayılabilir. Ayrıca 1980'lerin sonunda küçük örnek için Kullback.-Leibler uyumsuzluğunun eğilimsiz tahmin edicisi, Hurvich ve Tsai (1989) tarafından geliştirilerek model seçim kriterleri alanına sokulmuştur.

[4] Benzerlik oranı, Wald ve Lagrange çarpanı yuvalanmış modellerin karşılaştırmalarında kullanılan en bilindik testlerdir.

[5] Pesaran and Deaton (1978), Davidson ve Mackinnon(1981) yuvalanmamış modeller için seçim kriterleri oluşturmuşlardır.

II. MODEL SEÇİM KRİTERLERİ

1. R^2, Kriteri

Herhangi bir regresyonun uygunluk derecesini ölçmede R^2, 'nin kullanıldığı bilinmektedir. 0 ve 1 aralığında değerler alabilen ve 1'e yaklaştığında en iyi açıklanabilirliği veren istatistiği aşağıdaki şekliyle elde edebiliriz.

$$R^2 = {ESS}/{TSS} = 1 - {RSS}/{TSS} \quad \textbf{(1.1)}$$

Fakat R^2 model seçim kriteri olarak alındığında bazı sorunlarla kaşılaşılır (Gujarati, 2003).

- Verilen örnek büyüklüğü içinde tahmini değerler gerçek değerlere ne kadar yakın bulunsalarda gelecek tahmininde bu garantiyi sağlamak mümkün olmayabilir.
- R^2'lerin karşılaştırılabilmeleri için modellerin fonksiyonel yapısının ve tahmin edicilerinin aynı olması gerekmektedir. Farklı model yapıları için bir çok R^2 örneği verilebilir. Bunlardan bazıları: Maddala R^2, Gragg-Uhler R^2, McFadden R^2, Estrella R^2, Pseudo R^2(McFadden'a benzer fakat probit modellemelerinde kullanılır)
- Modele alınan açıklayıcı değişken sayısı arttıkça R^2 değeride artmaktadır. Bu yöntemle maksimum R^2'ye ulaşılabilinir. Fakat bu durum gelecek tahmin hata varyansınında yükselmesine neden olacaktır. Bu nedenle model seçim kriteri olarak alınması her zaman yeterli değildir.

2. Düzeltilmiş R^2

Henry Theil, açıklayıcı değişkenlerin katkılasıyla ortaya çıkan R^2 değerindeki yükselmeyi önlemek için düzeltilmiş R^2'yi geliştirmiştir.

$$\bar{R}^2 = 1 - {RSS/(n-k)}/{TSS/(n-1)} = 1 - (1 - R^2)\, {n-1}/{n-k} \quad \textbf{(1.2)}$$

Düzeltilmiş R^2, R^2'den farklı olarak eklenen değişkenin sadece mutlak t değerinin 1'den büyük olduğu durumlarda yükselir ve daima $\overline{R}^2 \le R^2$'dir. Fakat unutulmaması gereken model karşılaştırmalarında ister $\overline{R}^2$ veya R^2 kullanılsın herzaman modelin fonksiyonel yapısın ve tahmin edicilerinin aynı olmasıdır.

Modeller seçim analizinde karşılaştırıldığında her zaman maximum $\overline{R}^2$ değerini veren model tercih edilir. Maximum düzeltilmiş R^2 minimum kalıntı-varyans seçim kriterine[6] eş değer bir seçim kriteridir (Ebbeler, 1975). Düzeltilmiş R^2 en küçük kareler ile tahmin edilmiş regresyonlarda daha çok kullanılmaktadır. Bayes yaklaşımlarında zayıf kaldığı görülmektedir (Burham ve Anderson, 1998).

3. Akaike Bilgi Kriteri (AIC: Akaike Information Criterion)

Akaike Bilgi Kriteri (Akaike, 1973,1974 ve düzeltimiş AICc, Siungiura, 1978-Hurvich ve Tsai, 1989) modele eklenen değişkenlerin yarattığı yükselmeye sınırlama getirerek düzenlenmiştir.

$$AIC = e^{2k/n} \sum \hat{u}_i^2 \Big/ n = e^{2k/n} RSS \Big/ N \quad \textbf{(1.3)} \text{ veya}$$

$$AIC = -2\log(L) + 2k$$ [7]

(1.3) ve (1.4) eşitliklerinde k sabit terim dahil parameter sayısı ve n gözlem sayısını, L=benzer (likelihood)'liği vermektedir. (1.3) eşitliğinin matematiksel olarak her iki tarafının logaritması alınarak uygunlaştırılırsa,

---

6 $(A)\ \gamma = X\beta + \varepsilon$ (B) $\gamma = X\nu + \tau\ ,\ \gamma \approx N(\mu, \sigma^2 I_N)$

$M_X = I_N - X(X'X)^{-1}X'$

$M_X = I_N - Z(Z'Z)^{-1}Z'$

A ve B modelleri arasında seçim yaparken yukarıdaki minimum kalıntı-varyans kriteri dahilinde $\hat{\sigma}^2 = \gamma M_i k_i /(N - k_i) \quad i = x, z$ minimum yapan model seçilir.

7 Kullback-Leibler Uyumsuzluğu (Discrepancy)

$$\Delta_{K-L}(f, g_\vartheta) = -E_F \log g_\vartheta(x) = -\int \log g_\vartheta(x) f(x) dx$$

$$\ln AIC = \left(2k/n\right) + \ln\left(RSS/n\right) \quad \text{veya}$$

$$AIC = -2\log(L) + 2k \qquad \textbf{(1.4)},$$

(1.4) eşitliğinde lnAIC=AIC'nin doğal logaritması ve 2k/n ise *sınırlama (penalty)* faktörüdür. Yukarıdaki formüldende görüldüğü üzere AIC, $\overline{R}^2$'den eklenen yeni açıklayıcı değişkenler hususunda dahada serttir.

Daha sonraki dönemlerde Hurvich ve Tsai (1989)'nin küçük örnek zaman serisi regresyon modeleri için kullanılan eğilimsiz AIC'den türetmiş oldukları AICc aşağıdaki gibidir (Zucchini, 2000).

$$AIC_C = AIC + 2k(k+1)/(n-k-1) \quad \textbf{(1.5)}$$

Shibata (1981) AIC ve FPE'nin asimtotik olarak yeterli kriterler olduğunu söylerken, Nishii (1984) ve diğerleri de AICc , Cp'nin asimtotik olarak yeterli olmakla beraber; AIC, FPE, Cp'nin de asimtotik olarak eşit oldukları sonucuna varmışlardır.

AIC'nin bazı özellikleri söyle sıralanır:

- Model karşılaştırmalarında her zaman en düşük AIC değerini veren model tercih edilir.
- AIC sadece seçili örnek büyüklüğü içinde değil aynı zamanda seçili örnek büyüklüğü dışındaki gelecek tahmini içinde geçerlidir.
- Yuvalanmış, yuvalanmamış ve geçikmeli modellerde rahatlıkla kullanılabilir.

4. Bayes Bilgi Kriteri (BIC: Bayes Information Criterion)

Akaike (1978) ve Schwarz (1978) bayes perspektifinden birbirine yakın tutarlı iki model seçim kriteri tasarlamışlardır. Schwarz Koopman-Darmois türünde seçme modeler için SIC (Schwarz Information Criteria) kriterini türetirken buna karşın Akaike doğrusal regresyonda seçilmiş model problemleri için BIC (Bayesian Information Criterion) model seçim kriterini türetmiştir (McQuarrie, Tsai, 1998). Eşitlik (1.6)'daki gibidir.

$BIC = -2\log(L) + k\log(n)$ **(1.6)**

BIC eşitliğin sağ tarafındaki örnek büyüklüğüne bağlı olan ikinci kısım itibariyle AIC'den farklılık gösterir. Fakat AIC ve BIC arasındaki yüzeysel benzerliğe rağmen, daha sonraları bayes[8] yapısı içinde farklılıklar gösterdiği ortaya çıkmıştır (Rafley, 1995 ve Wasserman, 2000).

Literatür içindeki çalışmalara bakacak olursak BIC bayes factörü[9]nden daha fazla kullanılmaktadır. Bunun da sebeplerinden biri analiz sonrasında büyük hesaplamalara ihtiyaç duyulmasıdır (Zucchini, 2000).

5. Schwarz Bilgi Kriteri (SIC: Schwarz Information Criteria)

SIC kriteride AIC'ye benzemektedir. Formülü aşağıdaki gibidir:

$$SCI = n^{k/n}\frac{\sum \hat{u}^2}{n} = n^{k/n} RSS / n \qquad \textbf{(1.7)}$$

Logaritmik form ise:

$$\ln SCI = \frac{k}{n}\ln n + \ln(RSS/n) \qquad \textbf{(1.8)}$$

k/n ln*n sınırlama (penalty)* faktördür.

SIC'nin bazı özellikleri söyle sıralanır:

- SIC, AIC'ye göre yeni değişkenlerin modele eklendiğinde ortaya çıkacak durumu değerlendirme hususunda daha dikkatli düzenlenmiştir.
- SIC herzaman AIC'den daha düşük çıkar

[8] Bayes kuramı, önsezi olabilecek bilgilerin somut çıkarımlar olarak değerlendirilmesini sağlar ("Essay Towards Solving A Problem in the Doctine of Chance", Journal of Royal Society, 1763.) Bayes elde bir bulgu yokken, önsel olasılığın 0 ile 1 arasında herhangi bir değer alabileceğini iddia ederek, önsel olasılık ile baş etmeye çalışır.

[9] BAYES FAKTÖRÜ: $\frac{P(G_2 \mathrm{I} D)}{P(G_2 \mathrm{I} D)} = \frac{P(D \mathrm{I} G_2)P(G_2)}{P(D \mathrm{I} G_1)P(G_1)}$ eşitliğin sağ tarafındaki ilk bölüm bayes faktörü , ikinci bölüm ise prior odd (öncekinden artan) olarak bilinir.

- AIC'de olduğu gibi sadece seçili örnek büyüklüğü içinde değil aynı zamanda seçili örnek büyüklüğü dışındaki gelecek tahmini içinde geçerlidir.

6. Mallow'un C_p Kriteri

Sabit terimide içeren k açıklayıcı değişkenin bulunduğu model varsayımı altında p (p≤k) kadar açıklayıcı değişkenin seçildiği ve buradan kalıntı kareler toplamı hesaplanırsa C.P. Mallow'un model seçim kriteri (Mallows, 1973, 1995) (1.9) 'daki gibi olur:

$$C_p = \frac{RSS_p}{\hat{\sigma}^2} - (n-2p) \quad \mathbf{(1.9)}$$

$E(\hat{\sigma}^2)$ anakütle varyansının sapmasız tahmin edicisidir. Eğer etkin bir tahmin için p açıklayıcı değişken yeterli ise, $E(RSS_p) = (n-p)\sigma^2$ olduğu görülür. Sonuç olarak,

$$E(C_p) \approx \frac{(n-p)\sigma^2}{\sigma^2} - (n-2p) \approx p \quad \mathbf{(1.10)} \text{ bulunur}$$

Mallow'un C_p Kriterinin bazı özellikleri söyle sıralanır:

- *Cimrilik (parsimony)* prensibi dahilinde p<k yaklaşımı altında modeler karşılaştırılırken en düşük C_p değerini veren model seçilir (Gujarati,2003).
- Mallow'un C_p Kriteri, diğerleri gibi sabit varyans ve normal dağılımlı en küçük kareler tahmin regreyonunda kullanılırken bayes yaklaşımlarında .bu methodun göze çarpmadığı dikkat çekicidir (Seber, 1977. Burham ve Anderson, 1998).

7. Bootstrap Model Seçim Kriteri

Efron (1979) tarafından literature katılan bootstrap methodu öngörü veya tahmin hatasına dayanan model seçim kriteri olarak diğer kriterler içinde yer almaktadır. Linhart ve Zucchini (1986) istatistiksel özelliklerinin tam olarak bahsedilmediği bootstrap model seçim algoritmaları geliştirmişlerdir. Schall ve Zucchini (1990) bu algoritmaları kullanarak bootstrap model seçim kriterini tekrar yenilemişlerdir. Çalışmada, kriterin modifikasyonları ile birlikte parametrik ve nonparametrik açıdan ayrı ayrı incelemek gerektiği unutulmayarak temel anlayış yansıtılmaya çalışılmıştır.

X ve y arasındaki ilişkinin doğrusal olduğu bir durumda bootstrap gözlemlerinin oluşturulması için iki farklı yol olan bootstrap kalıntıları ve bootstrap çifti (x,y) model ve değişken seçim için kullanılabilir (Draper, 1998). Fakat bootstrap çifti bootstrap kalıntılarına nazaran varsayımlara daha az duyarlıdır. Ayrıca bootstrap kalıntıları x'in deterministik olduğu bir durum için daha uygundur. Eğer x rastlantısal ise bootstrap çifti daha uygun bir durum sergiler, fakat boostrap çifti deterministik x içinde kullanılabilmektedir. Her iki durumda da amaç öngörü hatasının bootstrap metodu kullanılarak tahmin edilmesidir (Efron-ve Tibshirani, 1993).

Öncelikle ayrı ayrı model seçimine geçmeden önce bootstrap ve çapraz-geçerlilik (cross-validation) için açıklamalara ışık tutacak bazı kavramlar ele alınmalıdır. 2^p-1(model seçimi için kullanılacak alternatif model sayısı)'e dayanan doğrusal modellerden biri için V ifadesini matris mantığı içinde kullanalım. V bir alt karakter olarak kullandığında ise söz konusu modelin elemanları kastedilmektedir. $n \times p_V$ tanımlı X_V matrisi, model V içindeki beraber değişkenleri ihtiva eden X matrisinin p_V tane sütununu içerir; X_V'in j. satırı ise X_{VJ}^T, V'deki en küçük kareler katsayı tahmincileri $\hat{\beta}_V$, ve H_V ise modelin tahmin edilmiş değerlerini tanımlayan şapka matrisidir (hat matris $X_V(X_V^T X_V)^{-1} X_V^T$). Model V içindeki regresyon katsayılarının toplam adedi $q_V = p_V + 1$'dir ve burada her zaman sabit katsayının olduğu varsayılır. Aynı ifadeler basit doğrusal bir model içinde söz konusudur. Tüm bu bilgiler ışığında çapraz geçerlilik ve bootstrap metodlarının model seçimi için nasıl kullanılacağına geçilebilir.

Kütle öngörü hatasının (aggregate prediction errror) çapraz geçerlilik ve bootstrap metodu kullanılarak minimize edilmesiyle en uygun V modeline ulaşılacağı düşünülmektedir (Hinkley ve Davidson, 1997). Buna göre V modeli kullanılarak ortalama öngörü karesi (average square prediction) şöyledir.

$$n^{=1} \sum_{j=1}^{n} (y_{+j} - x_{Vj}^T \hat{\beta}_V)^2$$ **(1.11)** buradan kütle öngörü hatası,

D(V)= $$\sigma^2 + n^{-1} \sum_{j=1}^{n} (\mu_j - x_{Vj}^T \hat{\beta}_V)^2$$ **(1.12)** de yer alan $\mu^{\mathrm{T}} = (\mu_1, ... \mu_n)$ gerçek çoklu regresyon modelinin ortalama (responses) tepkilerinin vektörüdür. Yukarıdaki ifadenin veri dağılımı üzerinden beklentisi alınırsa aşağıdaki (1.13) elde edilir.

$$\Delta(V) = E\{D(V)\} = (1 + n^{-1} q_V)\,\sigma^2 + \mu^{\mathrm{T}} (I - H_V)\,\mu \quad \textbf{(1.13)}$$

bu ifade de eğer model V doğruysa $\mu^{\mathrm{T}}(I - H_V)\mu$ sıfıra eşit olacaktır.

Prensipte, en iyi model D(V)'i minimize eden modeldir, fakat modelin parametreleri bilinmediği için D(V) ve Δ(V) için iyi birer tahminci bulunmalıdır. Böylece en iyi yaklaşım, iyi bir yöntem seçerek bunu bütün olabilecek modellere uygulamaktır

Bootstrap metod amaca göre uygulanırken bootstrap örneklemi B'nin belirlenmesi önem taşımaktadır. Buna göre konunun yazarlarının edinmiş oldukları tecrübelerden biraraya getirilen iki kural göze çarpar (Efron, 1982). .

1- Küçük sayıda bootstrap örneklemleri örneğin, B=25 konu için genellikle aydınlatıcı iken B=50'nin ancak $se_F(\hat{\theta})$ 'nın iyi tahminini vermekte yeterli olduğu ifade edilmektedir.

2- Bir standart hatanın tahmin edilmesi için B=200'den fazla örnekleme çok fazla ihtiyaç duyulmazken, güven aralıklarının belirlenmesinde B'nin daha büyük değerleri istenmektedir.

Bootstrap metodun model seçimi için diğer metodlara ve özelliklede aynı grupta yer almaları açısından çapraz geçerliliğe rağmen kullanılmasının en az iki nedeni sayılabilir (Shao,1996).

1. Doğrusal regresyon durumunda bootstrap farklı çıkarsamalar sağlamaktadır (örneğin, güven setlerinin oluşturulabilmesi..) ki bu diğer metodlar ile elde edilenden asimtotik olarak daha doğrudur. Kullanıcı açısından seçilmiş modele dayanan ardışık çıkarsama ve model seçimi, her ikisi de çalışmada tercih edilebilir. Eğer ardışık çıkarsama ve model seçimi için bootstrap kullanılıyorsa, model seçimi için oluşturulan bootstrap gözlemleri aynı zamanda ardışık çıkarsama içinde kullanılır. Bootstrap'in model seçimi için kullanılırken aynı zamanda ardışık çıkarsama için de kullanılması araştırma açısından ekstra bir maliyet getirmemektedir. Fakat ardışık çıkarsama için bootstrap ve model seçimi için çapraz geçerlilik metodunun kullanılması ekstra bir yeniden örneklem oluşturulmasını gerektirecektir. Bu nedenle model seçimi için çapraz geçerlilik metodundan kaçınılmalıdır.

2. Bootstrap model seçim prosedürleri, doğrusal regresyon için oluşturulmasına rağmen daha zor problemlerde de örneğin doğrusal olmayan modeller, genelleştirilmiş doğrusal modeller ve otoregresyon modellerinde de diğer kriterlerde (AIC, BIC, ….) olduğu gibi kullanım için ortam sağlayabilir. Fakat çapraz geçerlilik metodu bir yeniden örnekleme metodu olmasına rağmen otoregresyon modelleri için düşünülemez.

Bootstrap model seçim kriterinin performansı uygulamada diğer bilinen bazı seçim kriterlerine gore parametrik olarak sınandığında daha fazla üstünlük sergilediği görülmektedir (Ucal, 2005). Bu konuda da öne sürülen bazı tartışmalar mevcuttur. Çapraz-geçerlilik ve bootstrap metod birlikte veya karşılaştırmalı olarak model seçiminde kullanılırken yukarıdaki ayrımlara kesinlikle dikkat edilmelidir.

8. Çapraz-Gerçerlilik (Cross-validation) Model Seçim Kriteri

Eğer kare hatalar kullanılacaksa iyi bir tahmin edici olarak birini-dışarda-bırak çapraz geçerlilik tahmincisi düşünülebilir. Bu metod modellerin öngörü kabiliyetine göre model seçimi yapılmasını benimsemektedir (Shao, 1993).

Öncelikle basit bir $y = \beta_0 + \sum_{i=1}^{p-1} \beta_j x_j + \varepsilon$ **(1.14)** modeli için

$Q_{CV} = \frac{1}{n} \sum_{j=1}^{n} L(y_j, \hat{y}_j) = \frac{PRESS}{n}$ **(1.15)** olarak alınabilir

(PRESS=prediction Error Sum of Square, yani öngörü hata karesi). $L(y_j, \hat{y}_j)$ ifadesi kayıp fonksiyonunu (loss function) verirken, $L(y_j, \hat{y}_j) = (y_j - \hat{y}_j)$ ile de açıklanabilir (Banke ve Drage, 1984). Regresyon modelinde yer alan katsayılar yalnız bir kez tahmin edileceğine, birini-dışarda-bırak çapraz geçerlilik ile bir çok kez tekrarlanmaktadır. Ayrıca $Q_{CV} = \frac{1}{n} \sum_{j=1}^{n} L(y_j, \hat{y}_j) = \frac{PRESS}{n}$, $s^2 = RSS/(n-p)$ ile de karşılaştırılabilir (n=gözlem sayısı, p=sabit terim dahil parametre sayısı, RSS=Residual sum of square-Kalıntı kareler toplamı). Fakat $\mathrm{Q_{CV}}$, s^2 'ye göre her zaman daha büyük değerler verecektir. Ayrıca, PRESS genellikle $\mathrm{s}^2(1+\mathrm{p/n})$ 'e hayli yakın sonuçlar vermektedir (Hjorth, 1994).

Çapraz-geçerlilk geniş bir model ve matris mantığı içinde şöyle olacaktır,

$$\hat{\Delta}_{CV}(V) = \frac{1}{n}\sum_{j=1}^{n}\frac{(y_j - \hat{y}_{VJ})^2}{(I - h_{VJ})^2} \qquad \textbf{(1.16)}$$

(1.16)'da $\hat{y}_{VJ}$ bütün verilere dayanarak model V'nin tahmin edilmiş değeri, h_{VJ} ise j vakası için model V'nin kaldıracıdır.

$\hat{\Delta}_{CV}(V)$'in sapması küçüktür, fakat bu iyi bir seçim yapmak için yeterli değildir. Bunun niçin olduğunu görmek için herşeyden önce bir geniş formülasyona ihtiyaç vardır.

Formülasyon, $n\,\hat{\Delta}_{CV}(V) = \varepsilon^T(I - H_V)\varepsilon + 2p_V + \mu^T(I - H_V)\mu$ **(1.17)**

şeklindedir. Buradan eğer V doğru model ve daha büyük bir modeli ifade ediyorsa, bu durumda büyük n için şunlar geçerlidir: $\Pr\{\hat{\Delta}_{CV}(V) < \hat{\Delta}_{CV}(V')\} = \Pr(\chi^2_d < 2d)$ eşitliği sağlanmalıdır ki burada $d = P_{V'} - P_V$'dir. Bu olasılık eğer d büyük değilse, 1'in altında sağlanmaktadır. Dolayısıyla, $n\,\hat{\Delta}_{CV}(V)$ modelindeki $\mu^T(I - H_V)\mu$ teriminin doğru olmayan modellerin seçilemeyeceğini garanti etmesine rağmen, $\hat{\Delta}_{CV}(V)$'nin minimizasyonu tutarlı seçimi sağlamaz.

Daha yeterli ve doğru bir $\Delta(V)$'i tahmin etmek için hem model V'yi oluşturacak çok büyük miktarda veriye hem de büyük sayıda bağımlı öngörüye ihtiyaç vardır. Bundan dolayı $\Delta(V)$'i daha da genelleştirerek şöyle ifade edebiliriz.

$$\hat{\Delta}_{CV}(V) = R^{-1}\sum_{r=1}^{R} V^{-1} \sum_{J \in S_{a,r}} \{y_j - \hat{y}_{VJ}(S_{t,r})\}^2$$ **(1.18)**'de

$n_h = n - m$ ve $n_d = m$ büyüklüğünde hazırlama ve değerlendirme kümeleri için R farklı parçayı kullanan bir ifade vardır. Aynı zamanda, ifade de $\hat{y}_{VJ}(S_{t,r}) = X^T_{VJ}\,\hat{\beta}_V(S_{t,r})$ ve $\hat{\beta}_V(S_{t,r})$ ise V'nin en küçük kareler kullanılarak r'inci hazırlanma kümesinin tahmin edildiği katsayıları ve ($S_{t,r}$) ise hazırlanma kümelerini ifade eder. Burada dikkat edilmesi gereken R parçanın (split) tüm modeller için kullanılmasıdır.

$n \to \infty$ iken $n \to m \to \infty$ ve $m/n \to 1$ olacak şekilde m'nin seçildiği varsayılırsa, bu $\hat{\Delta}_{CV}(V)$'nin minimizasyonu ile $n \to \infty$ iken $R \to \infty$ iken doğru model için tutarlı bir seçim sağlanacağı mümkündür..

III. SONUÇ

Ekonomik bir olguyu incelerken modelin sağlam ve yeterli olduğundan emin olmak gerekir. Diğer bölümlerde de bahsedildiği gibi, kurulan modelin veri aşamasından başlayarak bir süreç içinde sınanması gerekmektedir. Ekonomik araştırmalarda yön verici olması dolayısıyla uygun modelin seçilmesine büyük önem gösterilmelidir. Amaç sadece model kurmak değil ekonomik olguyu ve seçili kütleyi doğru yansıtmasını sağlamak olmalıdır.

Model Hendry ve Richard'ın önerdiği üzere, yapılan kestirimler mantığa uymalı, kuramla uyumlu olmalı, açıklayıcı değişkenleri hata terimiyle ilişkisiz olmalı, katsayıları durağan olmalı, tahmin edilen kalıntılar beyaz gürültülü olmalı, son olarakta model tüm diğer rakip modelleri kapsamalı ve onların bulgularınıda açıklayabilmelidir.

Modellerin birbirleri arasındaki performansını belirlemede sadece klasik regresyon testlerinin yeterli kalmadığı bunların yanında daha önceki bölümlerde bahsedilen model seçim kriterlerine ihtiyaç duyulduğu açıktır. Model seçim kriterlerinin birbirlerine göre üstünlükleri ve nasıl kullanılacaklarına dikkat edilmelidir. Aslında McQuarrie ve Tsai'nin tasarladığı üzere tüm kriterlerin modellerin yapılarına (tek değişkenli, çok değişkenli, vektör otoregresif, robust, nonparametrik ve semiparametrik modeller gibi) göre ayrı ayrı incelenerek karşılaştırılmasıdır. Çünkü her kriterin modellin yapısına göre göstereceği performans farklıdır.

Son olarak, araştırmacı Hendry'nin TTT: test, test, test (bahsedilen model seçim kriterleri düşünülerek) anlayışını ve cimrilik prensibini, kullanılacak doğru modele ulaşmak için hedef olarak kabul etmelidir.

KAYNAKLAR

Akaike, H. (1970),"Statistical Predictor Identification", *Annals of the Institute of Statistical Mathematics*, 22, pp. 203-217.

------------ (1973), Information Theory and an Extension of the Maximum Likelihood Principle. InB.N. Petro and F. Csaki ed. 2nd International Symposium on Information Theory, pp. 267-281.

------------ (1978), “A Bayesian analysis of the Minimum AIC Procedure,” *Annals of the Institute of Statistical Mathematics*, 20, pp. 9-14.

Banke, O. and Drage, B.(1984), “Bootstrap and Cross-Validation Estimates of the Prediction Error for Linear Regression Models”, *The Annals of Statistics*, 12, pp. 1400-1424.

Bhansalli, R. J. Ve Downham, D. Y. (1977), “Some Properties of the Order of an Autoregressive Model Seected by a Generalization of Akaike’s EPF Criterion,” *Biometrika*, 64, pp. 547-551.

Box, G.E.P. ve Jenkins, G.M. (1970), *Time Series Analysis: Forcasting and Control,* Holden-Day, London.

Burham, K.P. ve Anderson, D.R.(1998), *Model Selection an Inference: A Practical Information Theoric Approach*, Springer-Verlag, New York.

Mallows, C.L.(1973), ”Some Comments on C_p”, *Technometrics*, 15, pp. 661-675.

---------------- (1995), “More Comments on Cp,” *Technometrics*, 37, pp. 362-372.

Davidson, A. C. ve Hinkley, D.V. (1997), *Bootstrap Methods and Their Application,* New York, Cambridge University Press.

Draper, N. R., *Applied Regression Analysis,* John Wıley & Sons, New York, 1998.

Ebbler, D. H.(1975), “On the Probability of Correct Model Selecting Using the Maximum R-Square Choıce Criterion,” *International Economic Review*, 16-2, pp. 516-520.

Efron, B. (1979), “Bootstrap methods: Another look at the jackknife”, *Ann. Statist*., 1, pp.446- 453.

----------- (1982), *The Jacknife, the Bootstrap and other Resampling Plans,* Philadelphia, Society for Industrial and Applied Mathematics.

----------- (1983), “Estimating the Error Rate of Prediction Rule Improvement on Cross-Validation”, *Journal of the American Statistical Association*, 78, 316-331.

Efron, B. ve Tibshirani, R. J. (1993), *An Introduction to the Bootstrap,* New York: Champman & Hall.

Geweke, J. Ve Meese, R. (1981), “Estimating Regression Models of Finite but Unknown Order,” International Economic Review, 22, pp. 55-70.

Gujarati, D. (2003), *Basic Econometrics*, McGraw Hill, Singapure.

Hanmam E. J. ve Quinn, B. G. (1979), “The Determination of the Order of an Auturegression,” *Jounal of the Royal Statistical Society*, 41, pp. 190-195.

Harvey, A. (1990), *The Econometric Analysis of Time Series,* The MIT Press, 2.b., Cambridge Mass.

Hendry, D.F. (1980),"Econometrics-Alchemy or Science", *Economica,* c.47.

Hjorth, J.S. (1994), *Computer Intensive Statistical Methods Validation Model Selection and Boostrap,* New York: Champman & Hall.

Hurvich, C.M. ve Tsai, C. (1989), " Regression and Time Series Model Selection in Small Samples", *Biometrika,* 76, pp.297-307.

Kullback, S. ve Leibler, R. A.(1951), "On Information and Sufficiency," *Annals of Mathematical Statistics*, 22, pp. 79-86.

Leamer, E. E.(1979), "Information Criteria for Choice of Regression Models: A Comment", *Econometrica,* 47, pp.507-510.

Lindley, D.(1956), " On A Measures of the Information Provided By An Experiment." *Annals of Mathematics and Statistics,* 27, pp.986-1003.

Linhart, H. ve Zucchini, W (1986), *Model Selection*, John Wıley & Sons, New York.

McQuarrie, A.D. ve Tsai, C. L., *Regression and Time Series Model Selection*, World Sciencetific, 1998.

Nishi, R. (19849, "Asymptotic Properties of Criteria for Selection of Variable in Multible Regression," *Annals of Statistics***,** 12, pp.758-765.

Onish, H. (1983),"A Variable Selection Procedure for Econometrics Models," *Computational Statistics & Data Analysis,* pp. 85-95.

Poirier, D. (1991)," Posterior Odds Analysis When All Competing Models are False,"*Econometric Society Meeting*, Amsterdam, The Netherlands.

Raftery, A. E. (1995), Bayesian Model Selection in Social Research(with Discussion by Andrew Gelman, Donald B. Rubin and Robert M. Hauser). In P.V. Marsdn (Ed.), *Socialogical Methodology*, pp. 111-196.

Saho, J. (1993), "Linear Model Selection by Cross-Validation", *Journal of the American Assosiational,* Vol. 88.

---------- (1996),"Bootstrap Model Selection", *Journal of the American Statistical Association,* 91, 434, pp.655-665.

Shao, J.ve. Tu, D (1996), *The Jacknife and Bootstrap,* Newyork, Springer.

Schwarz, G. (1978), "Estimating the Dimensions of a Model," *The Annals of Statistical***,** 6, pp.461-464.

Seber, G.A.F.(1977), *Linear Regression Analysis*, John Wıley & Sons, New York.

Sengun, M. (1999), "A Nonparametric Approach to Resampling (Bootstrap) Methods", *Fourth National Symposium of Econometrics and Statistics,* Antalya.

Shibata, R. (1981), "An Optimal Selection of Regression Variables," *Biometrika***,** 68, pp.45-54.

Sugiuna, N. (1978), "Further Analysis of the Data by Akaike's Information Criterion and the Finite Corrections," *Communication in Statistics-Theory and Methods,* 57, pp. 13-26..

Ucal, Ş. M. (2005), "Parametric Bootstrap Model Selection Criteria with in Linear Model Compared to Other Criteria", *The 9th World Multi-Conference on Systemics, Cybernetics and Informatics,* July 10-13, 2005-Orlando, Florida.

Weasserman, L. (2000), " Bayesian Model selection and Model averaging," *Jounal of Mathematical Psychology*, 44, 92-107.

Wu, C.F.J. (1986), "Jackknife, bootstrap and other resampling methods in regression analysis(with discussions), *Ann. of Statist.,* 14, pp. 1261-1350.

Zellner, A. (1971), *An Introduction Bayesian Inference in Econometrics,* John Wiley and Sons. Inc., NewYork.

------------ (1977), "Maximal Data Information Prior Distribution," in New Developments in the *Applications of Bayesian Methods* ed. By A. Ayka and C. Brumat, Amesterdam, North-Holland, pp.211-232.

------------ (1977), "Posterior Odds Ratio for Regression Hypothesis: General Considerations and Some Specific Results," *Econometric Society Meeting.*

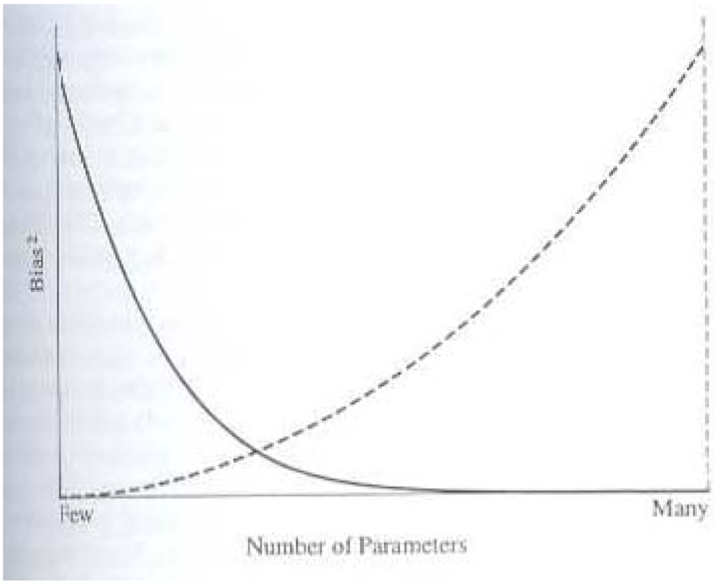
------------ (1978), "Jeffreys Bayes Posterior Odds Ratio and the Akaike Information Criterion for Discriminating Between Models," *Economics Letters,* pp.337-342.

Zucchini, W.(2000),"An Introduction to Model Selection,"*Journal of Mathematical Psychology,* 44, pp.41-61.

EK

Cimrilik Prensibi:

Box ve Jenkins (1970:17) "verinin en iyi şekilde yansıtılabilmesi için en az sayıda parametre yani değişken kullanılmalıdır" ifadesiyle ortaya attıkları cimrilik prensibi, tüm model seçim kriterlerinde genişletilerek kullanılmıştır. Breiman (1992) ve Zhang (1994) yazdıkları makalelerinde bu konuya ayrıntılı olarak değinmişlerdir. Prensibi aşağıdaki şekil ile açıklayabiliriz..



Yukarıdaki şekle dikkatlice bakıldığında, eğilim ve varyans arasındaki ilişkinin parametre sayısıylada yakından ilişkili olduğu görülür. Parametre sayısının artırılmasıyla sağlanan modelde eğilimin azaldığı buna karşın varyansın büyüdüğü görülür. Fakat seçilmiş bir modelde çıkarsamalar açısından varyansın artması düşündürücü olmaktadır. Ayrıca, en iyi modelin eğilim ve varyansın kesiştiği noktaya ihtiyacı yoktur. Cimrilik prensibi eğilim ve varyans arasındaki uygun dengeyi başarmaya çalışmaktadır. Zaten tüm model seçim kriterlerinin temelinde de bu yaklaşım vardır. Daha sonraları prensip genişletilerek kriterlerde yer almaya devam etmiştir.