

## Çoklu Doğrusal Regresyon Modelinde Değişken Seçiminin Zootečniye Uygulanışı

G. Tamer KAYAALP<sup>(1)</sup>Melis ÇELİK GÜNEY<sup>(1)</sup>Zeynel CEBECİ<sup>(1)</sup>

### Özet

Çoklu doğrusal regresyon analizinde, en uygun modeli belirlemek için klasik yöntemler ve adımsal yöntemler kullanılabilir. Adımsal yöntemler; ileriye doğru seçim yöntemi, geriye doğru seçim yöntemi ve adımsal seçim yöntemidir. Klasik yöntemlerde ise değişken seçimi için geliştirilmiş ölçütlerden bazıları; belirtme katsayısı ( $R^2$ ), düzeltilmiş belirtme katsayısı, hata kareler ortalaması ve Cook istatistiğidir.

Bu çalışmanın amacı, yumurta kabuk ağırlığı, yumurta ak indeksi ve yumurta sarı indeksi verileri ele alınıp, çoklu regresyonda adımsal yöntemlerden en uygun modeli ortaya koyarak yumurta ağırlığı tahmini yapmaktır. Çalışmanın sonucunda üç adımsal yöntemde de aynı modele ulaşılmıştır. Kullanılan bağımsız değişkenlerin yumurta ağırlığı tahmininde açıklayıcı olmadığı sonucuna varılmıştır.

**Anahtar kelimeler:** Çoklu regresyon analizi, değişken seçimi, ileriye doğru seçim, geriye doğru seçim, adımsal seçim

## Variable Selection in Multiple Regression Model Application of Animal Science

### Abstract

Generally two groups of methods are used for variable selection in order to determine the best fit model in multiple linear regression analysis. Alternatively the stepwise such as backward, forward and stepwise methods can be applied as new/efficient algorithms. While the well known criteria such as  $R^2$ , adjusted  $R^2$ , MSE and Cook statistics are used for variable selection.

The purpose of this study, egg shell weight, taken up egg whites and yolk index, index data, stepwise multiple regression model is the most suitable method of making egg weight estimate putting out. Three stepwise method has also been achieved as a result of the work of the same model. Egg weight was concluded that the arguments used to illustrate the estimation.

**Key Words:** Multiple regression analysis, variable selection, forward, backward, stepwise.

### Giriş

Çoklu doğrusal regresyon modelinde amaç, bağımsız değişkenler (açıklayıcı değişkenler) ile bağımlı değişkendeki (cevap değişkeni) toplam değişimi açıklamaktır.

Çoklu doğrusal regresyon analizinde, model oluşturan bağımsız değişkenlerden bazılarının modele katkısı önemsiz olabilir. Bu nedenle, bağımlı değişkeni "en uygun" şekilde açıklayacak bağımsız değişkenlerin belirlenmesi ve önemsiz değişkenlerin modelden çıkarılması gerekir. Bu sürece "değişken seçimi" denir (Alpar, 2003).

Regresyonda değişken seçimi ya da en iyi modelin oluşturulmasıyla ilgili çalışmalar Cox ve Snell (1974), Hocking (1972), Myers (1990) ve Thompson (1978a, 1978b) tarafından yapılmıştır.

Değişken seçimi amacıyla değişik yöntemler geliştirilmiştir. Bu yöntemler hesaplama tekniklerine göre klasik yöntemler ve adımsal yöntemler olmak üzere iki grupta incelenir. Adımsal yöntemler ise;

- 1) İleriye doğru seçim yöntemi
- 2) Geriye doğru seçim yöntemi

3) Adımsal seçim yöntemi olarak bilinir (Alpar,2003).

Bu çalışmada adımsal yöntemler karşılaştırmalı olarak incelenmektedir. İleriye doğru seçim yönteminde, her defasında bir bağımsız değişken ilave edilerek en uygun regresyon modelinin bulunmasını arzu etmektedir. Geriye doğru seçim yönteminde, ileriye doğru seçim yönteminden tam tersi durum söz konusudur. Model tüm bağımsız değişkenler ile başlamaktadır. Adımsal seçim yönteminde ise hem ileriye doğru seçim yöntemi hem de geriye doğru seçim yöntemi aynı anda kullanılmaktadır (Çakır Zeytinoğlu, 2007).

Zootekniye bağımlı değişken olarak ele alınan bir faktöre birçok unsur etki etmektedir. Ancak bir çok araştırmacı ele alınan bir faktöre etki eden unsurlardan kaç tanesinin veya hangisinin regresyon modeline girip girmeyeceği konusunda tereddüt etmektedir. Bu nedenle regresyon analizinde ele alınan modelde, modele hangi bağımsız değişkenlerin gireceği konusu zootekniye büyük önem arz etmektedir (Berberoğlu, 2002).

Bu çalışmada yumurta ağırlığını en iyi tahmin eden regresyon modelini bulmak için çoklu doğrusal regresyon modelinde değişken seçimi yöntemleri karşılaştırmalı olarak incelenecektir.

### Materyal ve Yöntem

#### Materyal

Bu çalışmada, Çukurova Üniversitesi, Ziraat Fakültesi, Zootekni Bölümü, Tavukçuluk İşletmesinden alınan yumurtalar materyal olarak kullanılmıştır. 80 gün boyunca rastgele toplanan yumurtaların, iç kalite özelliklerinden ak indeksi ve sarı indeksi, dış kalite özelliklerinden ise kabuk ağırlığı değişkenleri incelenmiştir. Tüm analizlerde SPSS 17.0 V. paket programı kullanılmıştır.

Çalışmada kullanılan veri setinde yumurtanın aşağıdaki özellikleri bağımlı ve bağımsız değişken olarak kullanılmıştır.

Y: Yumurta ağırlığı

X<sub>1</sub>: Kabuk ağırlığı

X<sub>2</sub>: Yumurta ak indeksi (%)

X<sub>3</sub>: Yumurta sarı indeksi (%)

### Yöntem

Populasyon için çoklu doğrusal regresyon modeli,

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \varepsilon_{ij} ; \quad (1)$$

$i=1,2,\dots,n \quad j=1,2,\dots,k$

şeklinde tanımlanır.

Eşitlikte;

Y<sub>i</sub>:Bağımlı değişkeninin gözlenen i.inci değerini

X<sub>ij</sub>: j.inci bağımsız değişkenin i.inci düzeyindeki değerini

β<sub>j</sub>: j.inci regresyon katsayısını

ε<sub>ij</sub>: Hata terimini

k: Bağımsız değişken sayısını göstermektedir (Gunst ve Mason, 1980).

En küçük kareler yöntemi kullanılarak çoklu doğrusal regresyon modelindeki regresyon katsayısının tahminleri elde edilir ve  $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \dots + \hat{\beta}_k X_{ik} ; i=1,2,\dots,n$  (2) şeklinde model oluşturulur (Montgomery ve ark., 2001; İyi ve Erol, 2008).

### İleriye Doğru Seçim Yöntemi

İleriye doğru seçim yönteminde değişken seçme işlemine modelde sadece sabit terimin bulunduğu denklemlerle başlanır ve değişkenler modele teker teker eklenir. Modele alınması düşünülen ilk bağımsız değişken, bağımlı değişken ile en yüksek korelasyona sahip değişkendir. Aynı zamanda bu değişken Y bağımlı değişkeni ile en yüksek F istatistiğine sahip değişkendir. Hesapla bulunan F istatistiği α önem düzeyinde anlamlı ise, bu değişken modele alınır ve ileriye doğru seçim yöntemi

devam eder. Yapılan test sonucunda, bu değişken modele alınmazsa seçim süreci sona erer (Alpar,2003).

### **Geriye Doğru Seçim Yöntemi**

İlk aşamada model içine tüm değişkenler dahil edilir. Daha sonraki kademelerde her defasında bir tane olmak üzere en düşük kısmi F değerine sahip olan bağımsız değişken atılmak sureti ile işleme devam edilir. Atılan değişkenin katkısı her seferinde test edilir. Atılan değişkenin katkısı istatistiki olarak önemli ise atma işlemi gerçekleştirilmez ve işlem orada durdurulur (Efe ve ark.,2000).

### **Adımsal Seçim Yöntemi**

Adımsal seçim yöntemi, ileriye doğru seçim yönteminin düzenlenmesinden oluşur. Modele daha önce eklenen bağımsız değişken kısmi F istatistikleriyle yeniden değerlendirilir. Modele daha önceden eklenen bir bağımsız değişken daha sonraki adımlarda modelden çıkarılabilir (Montgomery ve ark., 2001).

Bu yöntemde amaç, Y bağımlı değişkenini etkileyebilecek bağımsız değişkenlerin neler olduğu teorik olarak belirlendikten sonra, bunlar arasından birbiriyle ilişkileri olmayan ve bağımlı değişkeni en çok etkileyen değişkenleri seçmektir. Adımsal seçim yönteminin en önemli yararı, çoklu doğrusal bağlantı sorununa çözüm getirmesidir (Işık, 2006).

Adımsal seçim yöntemine ilişkin sistematik yaklaşım aşağıdaki şekilde özetlenebilir.

Bağımsız değişkenlerin tek tek bağımlı değişkenle aralarındaki basit doğrusal korelasyon katsayılarıyla, bağımsız değişkenler arasındaki basit doğrusal korelasyon katsayıları hesaplanarak korelasyon matrisi oluşturulur. Bu korelasyon matrisi incelenerek korelasyon katsayısı en yüksek olan bağımsız değişken seçilir.

Birinci adımda, bağımlı değişkeni en fazla etkilediği düşünülen bağımsız değişken

modele dahil edilir. Bu bağımsız değişkenle bağımlı değişken arasındaki ilişkinin korelasyon katsayısının t veya F testi seçilen anlamlılık düzeyi için uygulanır. İlişkinin anlamlı olduğu kabul edilirse, yani  $H_0: \beta_1=0$  şeklinde kurulan hipotez reddedilirse o bağımsız değişken modelde kalır.

İkinci adımda, kalan (k-1) adet bağımsız değişken içinden birinin bağımlı değişken Y 'yi önemli ölçüde etkileyen ikinci bağımsız değişken olarak seçilir. Modele giren değişken dışındaki tüm bağımsız değişkenlerin kısmi korelasyonları hesaplanır ve test edilir. Bu kısmi korelasyon katsayılarının incelenmesinin nedeni modelde mutlaka kalması gereken en güçlü değişken sabit tutulduğunda Y'yi en fazla etkileyen değişkeni bulmaktır. Böylece ilk modele eklenen bağımsız değişken ile çoklu doğrusal bağlantısı olmayan ve aynı zamanda Y bağımlı değişkenini en fazla etkileyen bir değişken seçilmiş olacaktır. Yapılan test sonucunda ilk modele giren bağımsız değişken ile yeni modele giren bağımsız değişken arasında olası düşük ilişki nedeniyle test istatistiği değişecektir. Bu değişim kontrol edilip, değişkenin modelde kalıp kalmayacağına bakılmalıdır.

Üçüncü adımda, ikinci adımda yapılan işlemlere benzer işlemler yapılır. Modele giren değişkenler dışındaki değişkenlere ait kısmi korelasyon katsayıları hesaplanır ve test edilir.

Bu süreç, seçilen anlamlılık düzeyinde ilişkiler anlamlı olduğu sürece tüm değişkenler için tekrarlanarak modele girip girmeyeceğine karar verilir. İlişkiler anlamsız çıktığında model anlamlı ilişkiye sahip olan bağımsız değişkenler kadar değişkenle kalır (Keller ve ark., 1990; Orhunbilge, 1996; Mendenhall ve Sincich, 1996; Işık, 2006).

Bağımsız değişkenlerin sürekli modele girip sonra tekrar çıkışını engellemek için önem seviyesi farklı alınmaktadır. Bu çalışmada bağımsız değişkeni modele eklemeye  $\alpha =0.05$ ,

## Çoklu Doğrusal Regresyon Modelinde Değişken Seçiminin Zootekniye Uygulanışı

bağımsız değişkeni modelden çıkarmada ise  $\alpha=0.10$  alınmıştır. Böyle yapılmasının ana nedeni modeli daha esnek tutup hata payını genişleterek modelde daha fazla değişken kalmasını sağlamaktır.

### Araştırma Bulguları

#### İleriye Doğru Seçim Yöntemi

İlk olarak Y bağımlı değişkeni ile X bağımsız değişkenler arasındaki basit doğrusal

korelasyon matrisine göre korelasyon katsayısı en yüksek bağımsız değişkenin bulunması gerekir. Çizelge 1' de görüldüğü gibi yüksek korelasyon katsayısı  $r_{yx1}=0.658$ 'dir. Ardından modelde  $X_1$  değişkeni bulunuyorken regresyon analizi yapılır.

Çizelge 1. Yumurta verilerine ait korelasyon matrisi

	Y	$X_1$	$X_2$	$X_3$
Y	1.000			
$X_1$	0.658	1.000		
$X_2$	0.279	0.140	1.000	
$X_3$	0.359	0.302	0.687	1.000

Çizelge 2.  $X_1$  değişkeni modelde iken regresyon sonuçları

Model	Standartlaştırılmamış Katsayılar		Standartlaştırılmış Katsayılar	t	p
	B	Standart Hata	Beta		
1 Sabit	29.522	4.686		6.300	.000
$X_1$	5.070	.658	.658	7.710	.000

$X_1$  değişkenin anlamlılık düzeyi  $p<0.05$  olduğundan  $X_1$  değişkeni modele alınır. Daha sonra  $X_1$  değişkeni sabitken Y ile en yüksek kısmi korelasyona sahip bağımsız değişken  $X_2$ 'dir.

$$r_{y2.1} = \frac{r_{y2} - r_{21} * r_{y1}}{\sqrt{(1 - r_{21}^2) * (1 - r_{y1}^2)}} = \frac{0.279 - 0.140 * 0.658}{\sqrt{(1 - 0.140^2) * (1 - 0.658^2)}} = 0.251$$

$$r_{y3.1} = \frac{r_{y3} - r_{31} * r_{y1}}{\sqrt{(1 - r_{31}^2) * (1 - r_{y1}^2)}} = \frac{0.359 - 0.302 * 0.685}{\sqrt{(1 - 0.302^2) * (1 - 0.658^2)}} = 0.224 \quad (3)$$

Bu nedenle  $X_2$  değişkeni modele alınacaktır. Şimdi modelde  $X_1$  ve  $X_2$  değişkeni bulunuyorken regresyon analizi yapılır. Çizelge

3'te bulunan  $X_2$ 'nin anlamlılık düzeyi  $p<0.05$  olduğu için  $X_2$  değişkeni modele eklenecektir.

Geriye sadece bağımsız değişken olarak  $X_3$  kalmıştır. Bu nedenle tüm bağımsız değişkenler modelde iken regresyon analizi yapılır. Çizelge 4'e bakılarak  $X_3$  bağımsız değişkeninin anlamlılık düzeyi  $p>0.05$  olduğu için  $X_3$  değişkenin modele giremeyeceği anlaşılır. Sonuç olarak bulunan model  $X_1$  ve  $X_2$  değişkenlerinden oluşan modeldir.  $R^2=0.468$  bulunmuştur.

$$\hat{Y}_1 = 27.356 + 4.864 X_1 + 0.587 X_2 \text{ dir.}$$

Çizelge 3.  $X_1$  ve  $X_2$  değişkeni modelde iken regresyon sonuçları

Model	Standartlaştırılmamış Katsayılar		Standartlaştırılmış Katsayılar	t	p
	B	Standart Hata	Beta		
1 Sabit	27.356	4.664		5.865	.000
$X_1$	4.864	.647	.631	7.517	.000
$X_2$	.587	.258	.191	2.274	.026

Çizelge 4.  $X_1$ ,  $X_2$  ve  $X_3$  değişkeni modelde iken regresyon sonuçları

Model	Standartlaştırılmamış Katsayılar		Standartlaştırılmış Katsayılar	t	p
	B	Standart hata	Beta		
1 Sabit	25.710	5.323		4.830	.000
$X_1$	4.738	.678	.615	6.992	.000
$X_2$	.429	.355	.140	1.211	.230
$X_3$	.086	.132	.078	.650	.518

#### Geriye Doğru Seçim Yöntemi

İlk olarak modele katkısı en az olan değişkeni bulmak gerekir. Çizelge 4'te bütün bağımsız değişkenler modelde iken  $X_3$  bağımsız değişkeninin önem seviyesi  $P > 0.10$  olması modele katkısını anlamsız kılar. Ancak Çizelge 4'te anlamlılık düzeyi 0.10' dan büyük  $X_2$  değişkeni de yer almaktadır. Bu nedenle  $X_3$  değişkeni modelden atılmışken regresyon analizi uygulanır ve Çizelge 3'te görülen sonuçlar elde edilir. Hem  $X_1$  hem de  $X_2$

değişkeninin anlamlılık düzeyi 0.10'dan küçük olduğu için işlem sonlandırılır.

SPSS paket programında geriye doğru seçim yöntemi uygulanırsa Çizelge 5'deki sonuçlar elde edilir. Geriye doğru seçim yöntemiyle  $R^2=0.468$  bulunmuştur. Sonuç olarak bulunan model  $X_1$  ve  $X_2$  değişkenlerinden oluşan modeldir.

$$\hat{Y}_1 = 27.356 + 4.864 X_1 + 0.587 X_2 \text{ dir.}$$

Çizelge 5. SPSS'de Geriye doğru seçim yöntemi sonuçları

Model	R	$R^2$	Düzeltilmiş $R^2$	Standart Hatanın Tahmini
1	.686	.471	.450	1.75921
2	.684	.468	.454	1.75260

## Çoklu Doğrusal Regresyon Modelinde Değişken Seçiminin Zootekniye Uygulanışı

Çizelge 5. SPSS'de Geriye doğru seçim yöntemi sonuçları (devam)

Model		Standartlaştırılmamış Katsayılar		Standartlaştırılmış Katsayılar	t	p
		B	Standart Hata	Beta		
1	Sabit	25.710	5.323		4.830	.000
	X <sub>1</sub>	4.738	.678	.615	6.992	.000
	X <sub>2</sub>	.429	.355	.140	1.211	.230
	X <sub>3</sub>	.086	.132	.078	.650	.518
2	Sabit	27.356	4.664		5.865	.000
	X <sub>1</sub>	4.864	.647	.631	7.517	.000
	X <sub>2</sub>	.587	.258	.191	2.274	.026

### Adımsal Seçim Yöntemi

Modele alınacak değişken, bağımlı değişkenle en yüksek korelasyon katsayısına sahip olan değişkendir. Çizelge 1' de görüldüğü gibi yüksek korelasyon katsayısı  $r_{yx1}=0.658$ 'dir. Burada X<sub>1</sub> bağımsız değişken modele girmeye adaydır.

Modelde Y ve X<sub>1</sub> değişkeni bulunuyorken regresyon analizi yapılır. Çizelge 2'ye bakıldığında X<sub>1</sub> değişkenin anlamlılık düzeyi  $p<0.05$  olduğundan X<sub>1</sub> modele eklenecektir.

Daha sonra X<sub>1</sub> değişkeni sabitken Y ile en yüksek kısmi korelasyona sahip bağımsız değişkeni bulmak gerekir. (3) numaralı eşitliğe bakılırsa;  $r_{y2.1}=0.251$ 'dir. Bu nedenle X<sub>2</sub> değişkeni modele girmeye adaydır.

X<sub>1</sub> ve X<sub>2</sub> değişkeni modelde iken regresyon analizi yapılır. Çizelge 3'e göre

X<sub>2</sub>'nin anlamlılık düzeyi  $p<0.05$  olduğu için X<sub>2</sub> değişkeni modele eklenecektir. Ancak X<sub>1</sub>'in anlamlılık düzeyi değiştiği için modelden çıkarma durumu olup olmadığı için kontrol edilmesi gerekmektedir.  $p<0.10$  olduğu için X<sub>1</sub> değişkeni modelde kalmaya devam edecektir.

Geriye sadece bağımsız değişken olarak X<sub>3</sub> kalmıştır. Çizelge 4'e bakılarak X<sub>3</sub> bağımsız değişkeninin anlamlılık düzeyi  $p>0.05$  olduğu için X<sub>3</sub> değişkeni modele giremeyeceği anlaşılır. Sonuç model; X<sub>1</sub> ve X<sub>2</sub> değişkenlerinden oluşan modeldir.

SPSS paket programında adımsal seçim yöntemi uygulanırsa Çizelge 6'daki sonuçlar elde edilir. Adımsal seçim yöntemiyle  $R^2=0.468$  bulunmuştur. Sonuç olarak bulunan model X<sub>1</sub> ve X<sub>2</sub> değişkenlerinden oluşan modeldir.

$$\hat{Y}_1 = 27.356 + 4.864 X_1 + 0.587 X_2 \text{ 'dir.}$$

Çizelge 6. SPSS'de Adımsal seçim yöntemi sonuçları

Model	R	R <sup>2</sup>	Düzeltilmiş R <sup>2</sup>	Standart Hatanın Tahmini
1	.658	.432	.425	1.79886
2	.684	.468	.454	1.75260

Çizelge 6. SPSS'de Adımsal seçim yöntemi sonuçları (devam)

Model		Standartlaştırılmamış Katsayılar		Standartlaştırılmış Katsayılar	t	p
		B	Standart Hata	Beta		
1	Sabit	29.522	4.686		6.300	.000
	X <sub>1</sub>	5.070	.658	.658	7.710	.000
2	Sabit	27.356	4.664		5.865	.000
	X <sub>1</sub>	4.864	.647	.631	7.517	.000
	X <sub>2</sub>	.587	.258	.191	2.274	.026

### Tartışma ve Sonuç

Bu çalışmada incelenen veri setinde, ileriye doğru seçim yöntemi, geriye doğru seçim yöntemi ve adımsal seçim yöntemi incelenmiştir. Üç yöntemle de aynı modele ulaşılmıştır. Tahmin modeli  $\hat{Y}_1 = 27.356 + 4.864 X_1 + 0.587 X_2$  olarak bulunmuştur. Kabuk ağırlığı ve yumurta ak indeksi sıfır iken tahmini yumurta ağırlığı 27.356 gr olması beklenmektedir. Ancak gerçekte böyle bir durum olmayacağından 27.356 değeri, biyolojik olarak anlamlı bir değer değildir. Tahmin için gerekli bir kesme noktası parametresi olarak değerlendirilir. Modellerde  $R^2=0.468$  orta düzeyde bir açıklama değeridir. Bu da yumurta ağırlığını belirlemede başka değişkenlerin olacağını işaret etmektedir.

### Kaynaklar

Alpar, R. (2003) *Uygulamalı Çok Değişkenli İstatistiksel Yöntemlere Giriş 1. 2.* Baskı, Nobel Basımevi, Ankara.

Berberoğlu, E. (2002) Değişken Seçimi Yönteminin Kullanımı İle Model Belirlenmesinin Zootekniye Uygulanışı. Yüksek Lisans Tezi, Gaziosmanpaşa Üniversitesi.

Cox, D. R., Snell, E. J. (1974) The Choice of Variables in Observational Studies, *Appl. Statist.*, 23, 51-59.

Çakır Zeytinoğlu, F. (2007) İşletmelerin Dönen Varlıklarının Satışlar Üzerindeki Etkileri: En İyi Regresyon Denklemine Seçimi ve Sektörel Karşılaştırma. *Marmara Üniversitesi İ.İ.B.F. Dergisi*, 23 (2) : 331-349.

Efe, E., Bek, Y., Şahin, M. (2000) *SPSS'de Çözümleri ile İstatistik Yöntemler II*, Kahramanmaraş Sütçüimam Üniversitesi Rektörlüğü Yayın No:10, Kahramanmaraş.

Gunst, R.F., Mason, R.L. (1980) *Regression Analysis and Its Applications*, Marcel Dekker, New York.

Hocking, G. (1972) Criteria for Selection of a Subset Regression: Which one should be used. *Technometrics*, 14, 967-970.

Işık, A.(2006) *Uygulamalı İstatistik-II*. Beta Basım, İstanbul.

İyi, P., Erol, H. (2008) Çoklu Lineer Regresyonda En İyi Model Seçimi. *Ç.Ü. Fen Bilimleri Enstitüsü*, 17 (5) :48-56.

Keller, G., Warrack, B., Bartel, H. (1990) *Statistics for Management and Economics. A System Approach*. 2<sup>nd</sup>

## Çoklu Doğrusal Regresyon Modelinde Değişken Seçiminin Zootekniye Uygulanışı

- Ed. Wadsworth Publishing Company. Belmont, California.
- Mendenhall, W., Sincich, T. (1996) *A Second Course in Statistics: Regression Analysis*. 5th Edition. Simon and Schuster / A Viacom Company. USA.
- Myers, R. H. (1990) *Classical and Modern Regression with Applications*, 2nd ed., PWS-Kent Publishers, Boston.
- Montgomery, D.C., Peck, E.A., Vining, G.G. (2001) *Introduction to Linear Regression Analysis*, 3rd Edition, John Wiley & Sons, New York.
- Orhunbilge, N. (1996) *Uygulamalı Regresyon ve Korelasyon Analizi*. İ.Ü.İşletme Fakültesi, Yayın No: 267, İstanbul.
- Thompson, M. L. (1978a) Selection of variables in multiple regression: Part I. A review and evaluation, *Int. Statist. Rev.*, 46: 1-19.
- Thompson, M. L. (1978b) Selection of variables in multiple regression: Part II. Chosen procedures, computations and examples, *Int. Statist. Rev.*, 46: 129-146.