

Eđitimle İlgili Sapan Deęer İeren Veri Kümelerinde En Küçük Kareler ve Robust M Tahmin Edicilerin Karşılaştırılması

Orkun COŞKUNTUNCEL

Özet – Eğitim arařtırmalarında regresyon katsayılarını tahmin etmek için en ok kullanılan yöntem klasik yöntem olarak da bilinen en küçük kareler (EKK) tahmin edicisidir. Ancak bu yöntem gözlenen verilerdeki sapan deęerlere karşı ok hassastır. Bu alıřmanın amacı, eğitim arařtırmalarında ortaya ıkabilen, sapan deęer(ler)e sahip verilerde adı geen problemlere karşı daha dayanıklı olan robust M regresyon tahmin edicisinin performansını incelemektir. Bunun için, Öğrenci Seçme Sınavı (ÖSS) puanları bağımlı deęişken, orta öğretim ders başarı puanları (OÖBP) ise bağımsız deęişken olarak ele alınarak, bağımsız deęişkenlerin bağımlı deęişkeni ne derece yordadığı iki farklı regresyon yöntemine dayalı olarak (En Küçük Kareler (EKK) ve robust M regresyon kestirimi) belirlenmiş ve sonuçlar karşılaştırılmıştır.

Anahtar kelimeler: Öğrenci seçme sınavı, en küçük kareler, robust tahmin, M tahmin edicisi, sapan deęer.

Abstract – *Comparison Least Squares Estimators with Robust M-Estimators Educational Data Contain Outliers* – Least square estimator which known classical method is most useable regression parameter estimator in educational research. But, when the educational data have outlier(s) the performance of the least square will be poor. The aim of this study is to propose robust M estimator for the model parameter of a regression models that can combat with the outlier(s) in educational research.

Key words: Student selection exam, least squares, robust estimation, M estimation, outliers.

Giriř

Eđitimde ve psikolojide ölçme araçlarından elde edilen sonuçlara dayalı ıkarımların doğruluęu ve buna dayalı verilecek kararların isabetli olması temelde ölçme araçlarının iki teknik özellięinin saęlanmasıya baęlıdır. Bunlar geerlik ve güvenirliliktir. Geerlik, ölçme aracıyla ölçülmek istenen özellięin amaca uygun olarak farklı özelliklerle karıştırılmadan ölçülmesidir. Başka bir deyiřle, bir ölçme aracının, geliřtirilmiş bulunduęu konuda amaca hizmet etmesidir. Güvenilirlik, bir ölçme aracının hatalardan arınık olma derecesidir. Ölçme aracının hatalardan arınıklığıının göstergelerinden biri de aynı birey üzerinde yapılan bir nitelięe ait ölçümlerin benzer şartlarda tekrar

Orkun Cořkuntuncel, Yrd. Do. Dr., Mersin Üniversitesi Eğitim Fakültesi. Bu alıřma Mersin Üniversitesi Bilimsel Arařtırma Projeleri birimi tarafından desteklenmiştir.

Mersin Üniversitesi Eğitim Fakültesi Dergisi, Cilt 5, Sayı 2, Aralık 2009, ss. 251-262.

Mersin University Journal of the Faculty of Education, Vol. 5, Issue 2, December 2009, pp. 251-262.

edilmesi halinde aynı sonuçları üretmesidir. Bu bir ölçme aracının hatadan arınıklığının veya hatasız ölçüm yapabilme gücünün bir göstergesidir (Erkuş, 2003).

Herhangi bir işe eleman alınırken, işin ölçütlerine uygun eleman seçme veya bir okula öğrenci alınırken okulun ya da bölümün gerektirdiği yetenek ve bilgi düzeyine sahip bireyleri seçme-yerleştirme için ölçme araçlarına ve testlere başvurulur. Bireylerin bir işi yapıp yapamayacağını ya da ölçülen özellikle ilgili geleceğe dönük performansının ölçüsünü veren seçme ve yerleştirme amaçlı kullanılan testlerin geçerliği, testi alan bireylerin testten elde ettiği puanlarla, ölçülen özellikle ilişkili geçerliği ve güvenilirliği kanıtlanmış ölçüt takımları arasındaki ilişkiyi hesaplamak suretiyle belirlenir. Bu tür kanıtlara dayalı geçerliğe, “ölçüt bağımlı geçerlik” adı verilir. Bir ölçüte dayalı geçerlik, bir grup bireyin ölçme aracından elde ettiği puanların, ölçüt durumundaki puanlar, sınıflamalar ya da diğer yetenek ve beceri ölçüleriyle karşılaştırmasına dayalı işlemleri içerir. Bir ölçüte dayalı geçerlik, ölçüt durumundaki puan ya da puan takımlarının elde edilmiş zamanına göre, zamandaş geçerlik ve yordama geçerliği olmak üzere iki gruba ayrılır. Ölçüt puanları yordayıcı puanlarla aynı zamanda veya daha önceden elde edilmişse, ölçüt puanları ile yordayıcı puanlar arasındaki ilişkiye dayalı geçerliğe zamandaş geçerlik, ölçüt puanların yordayıcı puanlardan sonra elde edilmesi durumunda ölçüt puanlar ile yordayıcı puanlar arasındaki ilişkiye dayalı geçerliğe yordama geçerliği adı verilir (Aiken 2000; Anastasi 1997; Crocker ve Algina 1986; Baykul 2000; Erkuş 2003; Baykul, Gelbal ve Kelecioğlu, 2001).

Ölçme aracının geçerliğini belirlemenin değişik yolları vardır. Bunlardan bazıları salt istatistiksel yöntemlerden yararlanmayı, bazıları mantıksal karşılaştırma ve yargılamayı, bazıları ise ikisini birden kullanmayı içerir. Bunlardan hangilerinin kullanılacağı ölçme aracının içeriğine, aracın ne gibi karar işlemleri için kullanılacağına ve var olan olanaklara bağlıdır. Öğrenci seçme ve yerleştirme işlemleri için hazırlanan testlerin ne ölçüde geçerli olduklarını, bu testler kullanılmadan önce belirlemek gereklidir. Ülkemizde, yüksek öğretim kurumlarına öğrenci seçme ve yerleştirme işlemlerinde bu gereklilik, testlerin gizlilik altında hazırlanması zorunluluğu yüzünden, mümkün olan her yol denenerek değil, olası bazı yöntemler uygulanarak yapılmaktadır. Örneğin, testler hazırlanmadan yapılan kapsam belirleme çalışmaları kapsam geçerliğini sağlama girişimlerinin yanında, testler uygulandıktan sonrada geçerliklerinin saptanması için ölçüt bağımlı geçerlik (yordama ve/veya zamandaş geçerliği) çalışmalarına başvurulmaktadır (Özçelik, 1982). Türkiye’de gerek lisansüstü düzeyde tez çalışması, gerekse bilimsel makale düzeyinde Öğrenci Seçme ve Yerleştirme Sınavının (ÖSS) geçerliğine ilişkin birçok çalışma yapılmıştır. Yapılan çalışmalarda, çeşitli istatistiksel teknikler kullanılarak, sınavın geçerliğine ilişkin çeşitli sonuçlar elde edilmiştir (Akhun, 1980; Aşkar, 1985; Demirok, 1990; Deniz ve Kelecioğlu, 2005; Erkuş, 1998; Gelbal, 1989; Kozan ve Tezer, 1979; Köse, 1990; Oral, 1985; Tezbaşaran, 1991).

Bir ölçüte dayalı geçerliliği belirlemede en çok kullanılan istatistiksel yöntemlerden biri regresyon analizidir. Regresyon analizinde değişkenler arasındaki ilişki matematiksel bir eşitlikle ifade edilir. Bu matematiksel eşitlik aracılığıyla bağımsız değişkenlerin, bağımlı değişken üzerindeki etkileri kestirilebilir. Değişkenler arasındaki ilişkiyi belirlemek amacıyla kullanılan en yaygın regresyon yöntemi klasik yöntem olarak da bilinen en küçük kareler (EKK) yöntemidir. EKK verideki sapan değerlere karşı çok hassas olan bir yöntemdir. Bu çalışmanın amacı regresyon analizine ait iki farklı kestirim yönteminin (EKK ve M-tahmin edicisi) geçerlilik türü üzerindeki etkilerini belirlemek ve sıkça karşılaşılan sapan değer problemine karşı daha etkili olan robust M tahmin edicisinin performansını incelemektir.

Yöntem

Bu çalışmada, ÖSS puanları bağımlı değişken, orta öğretim ders başarı puanları ise bağımsız değişken olarak ele alınmış ve bağımsız değişkenlerin bağımlı değişkeni ne derece yordadığı iki farklı tahmin yöntemine dayalı (en küçük kareler ve M tahmini) regresyon analizi ile belirlenmeye çalışılmıştır. Bu iki farklı kestirim yöntemin aynı ya da benzer sonuçlar üretip üretmediği araştırılmıştır. Böylece, farklı kestirim yöntemlerinin bir ölçüte dayalı geçerlik üzerindeki etkileri sorgulanmıştır.

Araştırma iki farklı yöntemin performansına dayalı olduğundan bir evren ve bu evreni temsil edecek bir örneklem üzerinde durulmamıştır. Bunun yerine bir orta öğretim kurumundan mezun olmuş 92 öğrenci üzerinde yürütülmüştür. Araştırmada, öğrencilerin öğrenim gördükleri dönemlerde elde ettikleri matematik, fizik, kimya, biyoloji, edebiyat, tarih derslerine ilişkin notlar ile ÖSS'den elde ettikleri, sözel, sayısal ve eşit ağırlıklı puanlar kullanılmıştır. Tüm bu veriler, öğrencilere ait oluşturulmuş formlar vasıtasıyla okul kayıtlarından elde edilmiştir. ÖSS sonuçları okul yönetimi ve ölçme ve değerlendirme servisi tarafından ÖSYM'den elde edilmiştir. Araştırmanın alt problemlerine çözüm bulmak amacıyla, öncelikle ders notları bağımsız değişken kümesi olarak, her bir ÖSS puanları ise bağımlı değişken olarak ele alınmıştır. Ayrıca, araştırmanın amacını daha iyi açıklayabilmek için, bir tek sapan değer EKK tahminine yaptığı olumsuz etkiyi gösterebilmek ve M-tahmin edicisinin bu olumsuz etkiye rağmen daha iyi bir tahmin üretebileceğini görebilmek amacıyla veriye 92 gözlemden farklı, sapan değer olan hayali bir gözlem eklenerek tahmin edicilerinin performansları incelenmiştir.

Belirli bir dersle ilgili başarı notu, o dersle ilgili meydana gelen hedeflerle tutarlı öğrenme düzeyini gösterir. O halde, normal şartlar altında araştırmacı, derslerindeki başarı notu yüksek olan bir öğrencinin daha iyi sınav notuna sahip olabileceğini düşünebilir. Ancak gözlemler arasında düşük başarı notuna sahip bir öğrencinin diğerlerinden daha yüksek not alması EKK tahminini ve dolayısıyla bu tahminde yola çıkılarak yapılacak çalışmaları olumsuz yönde etkileyebilir. Bu sebeple ek olarak; bir tek sapan değer bile EKK tahminine yaptığı olumsuz etkiyi gösterebilmek ve M

tahmin edicisinin bu sapan değer in olumsuz etkilerini ortadan kaldırma performansını daha iyi görebilmek amacıyla 92 gözlem içerisinde en düşük başarı notu ile en yüksek sınav notu belirlenerek hayali bir gözlem eklenmiştir. Şüphesiz bu eklenen hayali gözlem geriye kalanlardan farklı, sapan değer, olan bir gözlem olacaktır.

EKK ve M tahminlerinin hesaplanmasında R ve S-Plus istatistik paket programlarından yararlanılmıştır.

İşlem

Matris gösteriminde,

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}, \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

şeklinde olmak üzere

$$y = X\beta + \varepsilon \quad (1)$$

çoklu doğrusal regresyon modelini ele alalım. Genel olarak, y , $n \times 1$ tipinde gözlemlerin vektörü, X matrisi, $n \times p$ tipinde bağımlı değişken matrisi, β , $p \times 1$ tipinde regresyon katsayılarının vektörü ve ε , $n \times 1$ tipinde rastgele hataların vektörüdür. Burada k bağımlı değişken sayısı olmak üzere $p = k + 1$ 'dir ve 1'ler sabit terim içindir. Ayrıca $k = 1$ alınırsa basit doğrusal regresyon modeli elde edilir. En küçük kareler tahmin edicileri $\varepsilon = y - X\beta$ olmak üzere $\sum \varepsilon_i^2$ hata kareleri toplamını minimize eder.

Böylece β 'nin en küçük kareler tahmini $\hat{\beta}$,

$$\hat{\beta} = (X'X)^{-1}X'y \quad (2)$$

ve varyans-kovaryans matrisi,

$$\text{Var}(\hat{\beta}) = \sigma^2(X'X)^{-1} \quad (3)$$

şeklinde elde edilir. En küçük kareler için standartlaştırılmış rezidüel $e_i = y_i - \hat{y}_i$ olmak üzere,

$$r_i = \frac{e_i}{\hat{\sigma}} \quad (4)$$

ile verilir. Burada,

$$\hat{\sigma}^2 = \frac{1}{n-p} \sum_{i=1}^n e_i^2 \quad (5)$$

dir ve hatalar bağımsız, sıfır ortalamalı, σ standart sapmalı, özdeş dağılıma sahip olduğunda $\hat{\sigma}^2$, σ^2 'nin yansız tahmin edicisidir. Bazen standartlaştırılmış rezidüleri yerine

$$t_i = \frac{e_i}{\hat{\sigma} \sqrt{1-h_{ii}}} \quad (6)$$

studentized rezidüleri kullanılır. Bu iki çeşit rezidüye genel olarak standartlaştırılmış rezidüleri denir.

En küçük kareler yönteminin klasikleşmiş olmasını temel nedeni hesaplanmasının kolaylığıdır. Tahmin veriler yardımıyla herhangi bir iteratif yöntem ihtiyacı olmaksızın direk ve kolaylıkla hesaplanmaktadır. Ayrıca diğer tüm yansız tahmin ediciler içinde en iyi lineer yansız tahmin edicidir ve eğer hatalar normal dağılıyorsa, Maksimum Likelihood tahmin edicisine benzediği gibi bu durumda diğer yansız tahmin edicilere göre minimum varyanslı tahmin verir.

Robust istatistik, istatistiksel yöntemlerin sağlamlık teorisidir. Robust istatistik verilerin modelden sapmalarının klasik yöntemlere etkilerini inceler ve gerekiyorsa daha uygun bir yöntem geliştirir. En küçük kareler, regresyon modeli normal dağılıma sahipken çoğu zaman uygun yöntemdir ve iyi istatistiksel özelliklere sahip tahminler verir. Ancak normallikten sapmalar olduğu durumlarda uygun yöntem olmaktan çok uzaktır. Bu tip durumlarda alternatif tahmin edici olarak robust tahmin edicileri düşünebiliriz. İstatistiksel çalışmalarda, regresyon tahmin edicisi için başlangıçta kabul edilen istatistiksel model doğru olmazsa bile iyi sonuçlar verebilen yöntemler elde edebilme uğruna hatırı sayılır çabalar sarf edilmiştir. En küçük karelerin normallik varsayımı altında bile en uygun tahmin edici olduğu düşüncesi Tukey'in (1960) "A survey of sampling from contaminated distributions" adlı çalışması ile son bulmuştur. Daha sonra bu çalışmadan esinlenerek birbirine paralel dört önemli robust teori, Huber (1964), Huber (1965), Hampel (1968), Rousseeuw (1987) tarafından ortaya atılmıştır.

Regresyon analizinde karşılaşılan en büyük problem verideki bir veya daha fazla gözlemin diğer gözlemlerden farklı olması, yani sapan değer problemidir. Robust istatistiksel yöntemlerin esas hedefi bu tip sapan değer içeren veriler için kullanılabilir, tutarlı sonuçlar veren yöntemler geliştirmektir. Robust kelimesi ilk olarak 1953'te G.E.P. Box tarafından kullanılmıştır.

Robust regresyon tahmin yöntemleri, genellikle en küçük kareler yönteminden daha iyi istatistiksel sonuçlar üretmelerine rağmen, literatür incelendiğinde, istatistiksel analizlere gereksinim duyan ve istatistiksel yöntemler kullanan bilim dallarından eğitim bilimlerinde kullanılmamaktadır. Bunun başlıca nedeni robust yöntemlerin en küçük karelerin aksine iteratif yöntemlere ihtiyaç duymalarından dolayı hesaplanmasının zor

olmasıdır. Birçok robust tahmin yöntemi vardır. Bunlar arasında en çok kullanılanlar klasik robust tahmin yöntemi olarak da bilinen M-tahmin edicileri (*maximum likelihood* tipi tahmin edici), L-tahmin edicileri (sıra istatistiklerinin lineer kombinasyonları), R-tahmin edicileri (ranka dayalı veya rank dönüşümüne dayalı tahmin edici), RM-tahmin edicileri (*repeated median*-tekrarlı medyan tahmin edicileri), LMS-tahmin edicileri (medyan karelerinin en küçüğünü kullanan tahmin edici)'dir.

(1)'de verilen model için en küçük kareler tahmini (2)'de verilmişti. En genel halde (1)'deki β katsayısı için M tahmin edicisi, $\rho(e)$,

$$i) \rho(e) \geq 0 \quad ii) \rho(0) = 0 \quad iii) \rho(e) = \rho(-e) \quad iv) |e_i| > |e_j|, i \neq j \text{ iken } \rho(e_i) \geq \rho(e_j), e_i = y_i - x_i' \beta$$

koşullarını sağlayan bilinen bir fonksiyon olmak üzere

$$\sum_{i=1}^n \rho(e_i) = \sum_{i=1}^n \rho(y_i - x_i' \beta) \quad (7)$$

fonksiyonunu minimum yapar. Fonksiyonu minimum yapacak β değerini elde etmek için (12)'deki fonksiyonun β 'ya göre türevi sıfıra eşitlenirse,

$$\sum_{i=1}^n \rho'(y_i - x_i' \beta) x_i' = 0 \quad (8)$$

elde edilir. (13)'teki fonksiyonu

$$\sum_{i=1}^n \frac{\rho'(e_i)}{e_i} \cdot e_i x_i = 0, e_i \neq 0 \quad (9)$$

olarak yazabiliriz. (14)'teki fonksiyonda $w_i = \rho'(e_i)/e_i$ dersek fonksiyon,

$$\sum_{i=1}^n w_i x_i e_i = \sum_{i=1}^n w_i x_i (y_i - x_i' \beta) = 0 \quad (10)$$

şeklinde yazılır ve buradan

$$\sum_{i=1}^n w_i x_i y_i = \sum_{i=1}^n w_i x_i x_i' \beta \quad (11)$$

normal denklemleri elde edilir. Buna göre β için M tahmin edicisi,

$$\hat{\beta}_M = \left(\sum_{i=1}^n w_i x_i x_i' \right)^{-1} \sum_{i=1}^n w_i x_i y_i \quad (12)$$

dir. Matris formunda, (12)'deki normal denklemler, W matrisi köşegeninde w_i ağırlıkları bulunan köşegen bir matris olmak üzere,

$$X'WX\beta = X'Wy \quad (13)$$

ve $\hat{\beta}_M$ tahmini

$$\hat{\beta}_M = (X'WX)^{-1}X'Wy \quad (14)$$

olarak elde edilir. Buradaki w_i ağırlıkları her bir gözlem için ele edilecek ve eğer gözlem veya gözlemlerin bir gurubu geriye kalanlardan farklı olan, sapan değerlerse, sıfıra yakın ağırlıklara sahip olacaktır. Buna karşılık kalan gözlemler bire yakın ağırlıklara sahip olacaktır. Böylece sapan değerlerin modele yaptıkları olumsuz katkı minimuma indirilmiş olacaktır. Gerek bağımlı, gerekse bağımsız değişkenler sapan değerlere sahip olabilirler. Bu çalışmada sadece bağımlı değişkenlerin sapan değer veya değerlere sahip olması durumu ele alınacaktır.

Bulgular

Öncelikle her puan türü için EKK ile elde edilen sonuçlar 92 ve 93 gözlem için karşılaştırılmıştır. Katsayı tahminlerinin daha iyi karşılaştırılabilmesi için sonuçlar tablolarda verilmiştir. Tablo 1'de her puan türü için EKK ile elde edilen katsayı tahminleri yer almaktadır.

Tablo 1: 92 ve 93 Gözlemlili Veriler İçin EKK Sonuçları

Puan	Gözlem	Sabit	Matematik	Fizik	Kimya	Biyoloji	Edebiyat	Tarih	R ²
Sözel (y ₁)	92	72,45	-0,12	0,14	-0,02	0,14	0,13	0,63	0,56
	93	88,84	-0,15	0,03	-0,01	0,17	0,09	0,55	0,43
Sayısal (y ₂)	92	34,18	-0,01	0,59	0,22	0,33	0,38	-0,03	0,70
	93	62,51	-0,06	0,39	0,23	0,38	0,30	-0,15	0,48
Eşit Ağr.(y₃)	92	53,31	-0,07	0,36	0,10	0,24	0,25	0,30	0,76
	93	74,78	-0,10	0,21	0,10	0,27	0,19	0,20	0,51

Tablo 1'deki sonuçlara dikkat edilirse, eklenen yeni gözlem EKK sonuçlarını oldukça değiştirmiş ve özellikle sabit terim (kesim noktası) bundan çok fazla etkilenmiştir. Ayrıca R² değerleri de dikkate alınacak oranda azalmıştır. Bu durumda, robust M regresyon tahmin edicisinden beklentimiz 93 gözlemlilik veriye uygulandığında 92 gözlem için elde edilen EKK sonuçlarına yakın katsayı tahminleri üreterek eklenen sapan değerlerin modele yaptığı olumsuz katkıyı kısmen veya tamamen ortadan kaldıracak tahminler üretmesidir. Tablo 2'de 93 gözleme M tahmin edicisinin uygulanması ile elde edilen sonuçlar verilmiştir (Robust M regresyon tahmini için R² değerini hesaplamak teorik olarak mümkün değildir).

Tablo 2. 93 Gözlemlili Veriler için M Tahmin Edicisi ile Elde Edilen Sonuçlar

Puan	Sabit	Matematik	Fizik	Kimya	Biyoloji	Edebiyat	Tarih
Sözel (y_1)	75,74	-0,11	0,14	-0,03	0,14	0,14	0,60
Sayısal (y_2)	40,67	-0,04	0,66	0,17	0,35	0,33	-0,08
Eşit Ağr. (y_3)	56,49	-0,07	0,39	0,07	0,26	0,24	0,27

93 gözlemlilik veriye uygulanan robust M tahmin edicisi ile elde edilen sonuçlar beklenildiği gibi 92 gözlemlilik veri için elde edilen EKK tahminine yakın elde edilmiştir. Tablo 3'te M tahmininin her gözlem için kullandığı ağırlıklar verilmiştir.

Tablo 3. 93 Gözlemlili Veriler için M Tahmin Edicisi ile Elde Edilen Ağırlıklar

Sözel (y_1)		Sayısal (y_2)				Eşit Ağırlık (y_3)											
1	0,69	32	1,00	63	1,00	1	0,99	32	1,00	63	1,00	1	0,57	32	1,00	63	1,00
2	1,00	33	1,00	64	1,00	2	1,00	33	0,52	64	0,63	2	1,00	33	1,00	64	1,00
3	1,00	34	1,00	65	1,00	3	0,66	34	1,00	65	1,00	3	1,00	34	1,00	65	1,00
4	1,00	35	1,00	66	1,00	4	1,00	35	1,00	66	1,00	4	1,00	35	1,00	66	1,00
5	0,64	36	0,87	67	1,00	5	1,00	36	1,00	67	1,00	5	1,00	36	1,00	67	1,00
6	1,00	37	1,00	68	1,00	6	1,00	37	1,00	68	1,00	6	1,00	37	1,00	68	1,00
7	1,00	38	1,00	69	1,00	7	1,00	38	1,00	69	0,68	7	1,00	38	1,00	69	1,00
8	1,00	39	1,00	70	0,67	8	1,00	39	1,00	70	1,00	8	1,00	39	1,00	70	0,70
9	1,00	40	1,00	71	1,00	9	1,00	40	1,00	71	1,00	9	1,00	40	1,00	71	1,00
10	1,00	41	1,00	72	1,00	10	1,00	41	1,00	72	1,00	10	1,00	41	1,00	72	1,00
11	1,00	42	1,00	73	1,00	11	0,94	42	1,00	73	1,00	11	0,88	42	1,00	73	1,00
12	0,95	43	1,00	74	1,00	12	1,00	43	1,00	74	1,00	12	1,00	43	0,93	74	1,00
13	1,00	44	1,00	75	1,00	13	<u>0,49</u>	44	1,00	75	0,79	13	0,56	44	1,00	75	0,76
14	1,00	45	1,00	76	1,00	14	0,85	45	1,00	76	1,00	14	1,00	45	1,00	76	1,00
15	0,91	46	0,64	77	1,00	15	1,00	46	0,64	77	0,73	15	0,95	46	<u>0,44</u>	77	0,80
16	1,00	47	1,00	78	1,00	16	0,76	47	1,00	78	1,00	16	0,85	47	1,00	78	1,00
17	1,00	48	1,00	79	1,00	17	1,00	48	1,00	79	0,80	17	1,00	48	1,00	79	1,00
18	1,00	49	1,00	80	1,00	18	1,00	49	1,00	80	1,00	18	1,00	49	0,91	80	1,00
19	0,50	50	1,00	81	1,00	19	1,00	50	0,77	81	1,00	19	<u>0,46</u>	50	0,86	81	1,00
20	1,00	51	1,00	82	1,00	20	0,82	51	0,76	82	0,93	20	1,00	51	1,00	82	0,85
21	0,87	52	0,63	83	1,00	21	1,00	52	0,61	83	1,00	21	1,00	52	1,00	83	1,00
22	1,00	53	0,92	84	1,00	22	1,00	53	1,00	84	<u>0,40</u>	22	1,00	53	1,00	84	0,82
23	1,00	54	1,00	85	1,00	23	1,00	54	1,00	85	1,00	23	1,00	54	1,00	85	1,00
24	1,00	55	1,00	86	1,00	24	1,00	55	1,00	86	1,00	24	1,00	55	1,00	86	1,00
25	1,00	56	1,00	87	1,00	25	1,00	56	1,00	87	1,00	25	1,00	56	1,00	87	1,00
26	1,00	57	1,00	88	1,00	26	1,00	57	1,00	88	1,00	26	1,00	57	1,00	88	1,00
27	0,85	58	1,00	89	1,00	27	1,00	58	1,00	89	<u>0,33</u>	27	1,00	58	0,78	89	<u>0,40</u>
28	1,00	59	1,00	90	1,00	28	1,00	59	1,00	90	1,00	28	1,00	59	1,00	90	1,00
29	1,00	60	1,00	91	0,88	29	1,00	60	1,00	91	1,00	29	1,00	60	0,80	91	1,00
30	0,57	61	1,00	92	1,00	30	1,00	61	1,00	92	1,00	30	0,68	61	1,00	92	1,00
31	1,00	62	1,00	93	<u>0,21</u>	31	0,94	62	1,00	93	<u>0,12</u>	31	1,00	62	1,00	93	<u>0,11</u>

Altı çizgili olarak verilen ağırlıklar diğerlerine göre kısmen düşük olup dikkat edilmesi gereken gözlemlerdir. Genel olarak 0,5 değerinden daha düşük olan ağırlığa sahip gözlemler incelenmesi gereken gözlemlerdir. Burada özellikle 93. gözleme ait ağırlıklar diğerlerinin aksine sıfıra daha yakın elde edilmiş ve böylece modelde oluşturduğu hasar oldukça düzeltilmiştir.

92 gözlemden oluşan veri için robust M tahmin edicisi ile elde edilen ağırlıklar Tablo 4'te verilmiştir ve M tahmin edicisi sonuçları Tablo 5'te verilmiştir.

Tablo 4. 92 Gözlemlili Veriler için M Tahmin Edicisi ile Elde Edilen Ağırlıklar

Sözel (y_1)			Sayısal (y_2)			Eşit Ağırlık (y_3)											
1	0,70	32	1,00	63	1,00	1	0,95	32	1,00	63	1,00	1	0,57	32	1,00	63	1,00
2	1,00	33	1,00	64	1,00	2	1,00	33	0,50	64	0,61	2	1,00	33	1,00	64	1,00
3	1,00	34	1,00	65	1,00	3	0,61	34	1,00	65	1,00	3	1,00	34	1,00	65	1,00
4	1,00	35	1,00	66	1,00	4	1,00	35	1,00	66	1,00	4	1,00	35	1,00	66	1,00
5	0,67	36	0,95	67	1,00	5	1,00	36	1,00	67	1,00	5	1,00	36	1,00	67	1,00
6	1,00	37	1,00	68	1,00	6	1,00	37	1,00	68	1,00	6	1,00	37	1,00	68	1,00
7	1,00	38	1,00	69	1,00	7	1,00	38	1,00	69	0,64	7	1,00	38	1,00	69	1,00
8	1,00	39	1,00	70	0,68	8	1,00	39	1,00	70	1,00	8	1,00	39	1,00	70	0,69
9	1,00	40	1,00	71	1,00	9	1,00	40	1,00	71	1,00	9	1,00	40	1,00	71	1,00
10	1,00	41	1,00	72	1,00	10	1,00	41	1,00	72	1,00	10	1,00	41	1,00	72	1,00
11	1,00	42	1,00	73	0,96	11	0,98	42	1,00	73	1,00	11	0,95	42	1,00	73	1,00
12	0,94	43	1,00	74	1,00	12	1,00	43	1,00	74	1,00	12	1,00	43	0,94	74	1,00
13	1,00	44	1,00	75	1,00	13	<u>0,46</u>	44	1,00	75	0,76	13	0,54	44	1,00	75	0,77
14	1,00	45	1,00	76	1,00	14	<u>0,79</u>	45	1,00	76	1,00	14	1,00	45	1,00	76	1,00
15	0,92	46	0,64	77	1,00	15	1,00	46	0,60	77	0,69	15	0,95	46	<u>0,43</u>	77	0,79
16	1,00	47	1,00	78	1,00	16	0,75	47	1,00	78	1,00	16	0,87	47	1,00	78	1,00
17	1,00	48	1,00	79	1,00	17	1,00	48	1,00	79	0,79	17	1,00	48	1,00	79	1,00
18	1,00	49	1,00	80	1,00	18	1,00	49	1,00	80	1,00	18	1,00	49	0,88	80	1,00
19	0,51	50	1,00	81	1,00	19	1,00	50	0,74	81	1,00	19	<u>0,46</u>	50	0,86	81	1,00
20	1,00	51	1,00	82	1,00	20	0,79	51	0,72	82	0,87	20	1,00	51	1,00	82	0,84
21	0,89	52	0,64	83	1,00	21	1,00	52	0,59	83	1,00	21	1,00	52	1,00	83	1,00
22	1,00	53	1,00	84	1,00	22	1,00	53	1,00	84	<u>0,40</u>	22	1,00	53	1,00	84	0,88
23	1,00	54	1,00	85	1,00	23	1,00	54	1,00	85	1,00	23	1,00	54	1,00	85	1,00
24	1,00	55	1,00	86	1,00	24	1,00	55	1,00	86	1,00	24	1,00	55	1,00	86	1,00
25	1,00	56	1,00	87	1,00	25	1,00	56	1,00	87	1,00	25	1,00	56	1,00	87	1,00
26	1,00	57	1,00	88	1,00	26	1,00	57	1,00	88	1,00	26	1,00	57	1,00	88	1,00
27	0,85	58	1,00	89	1,00	27	1,00	58	1,00	89	<u>0,32</u>	27	1,00	58	0,81	89	<u>0,41</u>
28	1,00	59	1,00	90	1,00	28	1,00	59	1,00	90	1,00	28	1,00	59	1,00	90	1,00
29	1,00	60	1,00	91	0,94	29	1,00	60	0,96	91	1,00	29	1,00	60	0,78	91	1,00
30	0,57	61	1,00	92	1,00	30	1,00	61	1,00	92	1,00	30	0,66	61	1,00	92	1,00
31	1,00	62	1,00			31	0,91	62	1,00			31	1,00	62	1,00		

Tablo 5. 92 Gözlemlili Veriler için M Tahmin Edicisi ile Elde Edilen Sonuçlar

<i>Puan</i>	<i>Sabit</i>	<i>Matematik</i>	<i>Fizik</i>	<i>Kimya</i>	<i>Biyoloji</i>	<i>Edebiyat</i>	<i>Tarih</i>
Sözel (y₁)	71,41	-0,10	0,17	-0,03	0,13	0,15	0,61
Sayısal (y₂)	36,21	-0,03	0,69	0,17	0,34	0,35	-0,06
Eşit Ağr. (y₃)	53,40	-0,07	0,41	0,06	0,26	0,25	0,28

Dikkat edilirse sözel puan için elde edilen sonuç, EKK ile hemen hemen aynı olmakla beraber M tahmininin ürettiği ağırlıklar incelendiğinde tümünün 1 veya 1'e çok yakın olduğu görülmektedir. Sayısal puan için elde edilen sonuçlar ise dikkate değer şekilde değişmiştir. Ağırlıklar incelendiğinde 13, 84 ve 89. gözlemlere ait ağırlıkların sifira yakın diğerlerinin ise 1'e yakın olduğu görülmektedir. Benzer şekilde eşit ağırlık puanı için elde edilen sonuçların EKK sonuçlarından gösterdiği farklılığın sayısal puandaki kadar olmadığı görülmektedir. Ancak eşit ağırlık için M tahmininin ürettiği ağırlıklarda 19, 46 ve 89. gözlemlere ait ağırlıkların diğerlerinin aksine 0'a yakın olduğu görülmektedir.

Sapan değerlerin modele olan etkilerini araştırmak için en pratik yöntem bu gözlemleri veriden çıkararak tekrar EKK uygulamaktır. Bu durumda sayısal ve eşit ağırlık için elde edilen M tahmini sonuçlarında düşük ağırlık alan gözlemleri çıkararak EKK uyguladığımızda elde ettiğimiz sonuçlar Tablo 6'da verilmiştir.

Tablo 6. Düşük Ağırlığa Sahip Gözlemlerler Çıkarıldıktan Sonra Elde Edilen Sonuçlar

<i>Puan</i>	<i>Sabit</i>	<i>Matematik</i>	<i>Fizik</i>	<i>Kimya</i>	<i>Biyoloji</i>	<i>Edebiyat</i>	<i>Tarih</i>	<i>R²</i>
Sayısal (y₂)	41,75	-0,03	0,77	0,04	0,42	0,24	-0,06	0,78
Eşit Ağr. (y₃)	56,50	-0,08	0,47	0,01	0,28	0,20	0,27	0,79

Elde edilen sonuçlar 92 gözleme uygulanan EKK sonuçlarıyla karşılaştırıldığında çıkarılan gözlemlerin modele etkileri açıkça görülebilmektedir. Ancak burada bir noktayı tekrar vurgulamakta yarar vardır. Son yapılan veriden gözlem çıkarma işlemi sadece şüphelenilen gözlemlerin regresyon katsayıları üzerindeki etkilerini incelemek için yapılmaktadır. Çünkü bu gözlemler atıldığında zaman geriye kalan gözlemler içinde sapan değerler olabilir ve bunlardan kaynaklanan katsayı tahmin hataları oluşabilir; hatta bağımsız değişkenlerin oluşturduğu matriste olması muhtemel kötü koşulluluk daha da vahim bir düzeye ulaşabilir, belki de ortadan kalkabilir. Bu yüzden sadece katsayılar bakarak gözlemler hakkında oluşan sapan değer yargısının ölçmek için gözlem çıkartılır. Bunun dışında veriyi ilk hali ile ele almak gerekir.

Sonuçlar ve Öneriler

Yukarıda verilen örnekte de görüldüğü gibi regresyon analizinde sapan değerlerin regresyon katsayıları üzerinde çok olumsuz etkileri vardır. Bu olumsuz etkileri ham veri ile oynamadan azaltmak veya tamamen ortadan kaldırmak için robust M regresyon tahmin edicileri uygun tahmin yöntemi olarak ele alınabilir. Bilindiği gibi veri toplama her aşaması çok dikkat, zaman ve emek gerektirmektedir; ayrıca ekonomik şartlar da buna eklendiğinde, önemi bir kat daha artmaktadır. Özellikle eğitim bilimleri alanında yapılan çalışmalar örneklemin tutarlılığı ve genellenebilirliği açısından çoğunlukla yüksek gözlem sayısına sahiptir. Bundan dolayı zaman zaman verinin normalliğini bozan gözlemlerin atılması veya toplanan verinin toplandığı örneklemden normal dağılacak şekilde yeni bir örneklem çekilmesi sıkça karşılaşılan bir durumdur. Ancak bu durum bazen çalışmanın geçerliği ve genellenebilirliği açısından sorunlar doğurabilir. Ayrıca gözlem sayısı fazla olan verilerde gruplaşmaların olması ihtimalinin yüksek olması araştırmacıyı küçük olan grubun veriden çıkarılmasına sevk edebilecektir. İşte oluşabilecek bu ve benzeri sorunlarla mücadele edebilmek için eğitim araştırmalarında robust tahmin ediciler EKK tahminine alternatif olarak düşünülebilirler.

Kaynakça

- Akhun, İ. (1980). *Akademik başarının kestirilmesi*. Ankara Üniversitesi Eğitim Fakültesi Yayınları.
- Anastasi, A. (1997). *Psychological testing*. New York, USA: Prentice Hall Inc.
- Arslan, O. (2004). Convergence behavior of an iterative reweighting algorithm to compute multivariate M-estimates for location and scatter. *Journal of Statistical Planning and Inference*, 118, 115-128.
- Aşkar, P. (1985). *Yükseköğretime öğrenci seçme ve yerleştirme sisteminin geçerliği*. Yayımlanmamış doktora tezi, Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Baykul, Y. (2000). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması*. Ankara: ÖSYM.
- Baykul, Y., Gelbal, S., ve Kelecioğlu, H. (2001). *Eğitimde ölçme ve değerlendirme*, Ankara: MEB.
- Crocker, L. ve Algina, J. (1986). *Introduction to classical and modern test theory*. Florida, USA: Harcourt Brace Jovanovich College Publishers.
- Demirok, S. (1990). *ÖSS ve ÖYS puanları ile lise ve dengi okullardaki başarının yüksek öğretimdeki başarıyla ilişkisi*. Yayımlanmamış yüksek lisans tezi, Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Deniz, Z.K. ve Kelecioğlu, H. (2005). İlköğretim başarı ölçüleri ile Ortaöğretim Kurumları Öğrenci Seçme ve Yerleştirme Sınavı arasındaki ilişkiler, *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 38, 127-143.
- Doğan, N. (1999). *Dershane deneme sınavları ile ÖSS ve ÖYS arasındaki ilişki*. Yayımlanmamış yüksek lisans tezi, Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.

- Erkuş, A. (1998). *1991-1994 ÖYS tercihlerinin test-tekrar test güvenilirlikleri ile yordama geçerliklerinin zaman içinde gösterdikleri değişim*. 10. Ulusal Psikoloji Kongresi (Serbest Bildiri), Ankara Üniversitesi.
- Erkuş A (2003) Psikometri üzerine yazılar: ölçme ve psikometrinin tarihsel kökenleri, güvenilirlik, geçerlik, madde analizi, tutumlar; bileşenleri ve ölçülmesi. 1. baskı, Ankara. Türk Psikologlar Derneği Yayınları No:24. s. 34-148.
- Gelbal, S. (1989). *Öğrenci seçme sınavı ile ÖSS testlerinin öğrenci başarıları yönünden ilişkileri, güvenilirlikleri ve ÖYS'yi yordama güçleri*. Yayınlanmamış yüksek lisans tezi, Hacettepe Üniversitesi, Ankara.
- Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J., ve Stahel, W.A. (1986). *Robust statistics: The approach based on influential functions*. New York: Wiley.
- Huber, P.J. (1964). Robust estimation of a location parameters. *The Annals of Mathematical Statistics*, 35, 73-101.
- Huber, P.J. (1981). *Robust statistics*. New York: Wiley.
- Kozan, K. ve Tezer, E. (1979). *Üniversitelerarası seçme sınavı geçerlik araştırması*, ÜSYM, AB-20-77-0020.
- Köse, M.R. (1990). Üniversiteye giriş ve liselerimiz. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 15, 51-60.
- Oral, T. (1985). *Lise başarı ölçüleri ile ÖSYS puanları arasındaki uyum*. Yayınlanmamış doktora tezi, Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Özçelik, D.A. (1982). *Öğrenci seçme ve yerleştirme sınavı geçerlik araştırması. Öğrenci seçme ve yerleştirmede kullanılan yöntemlere ilişkin bazı sorunlar*. ÖSYM Araştırma-Geliştirme Birimi, Ankara.
- Rousseeuw, P.J. ve Leroy, A.M. (1987). *Robust regression and outlier detection*. New York: Wiley.
- Tezbaşaran, A. (1991). *Yüksek öğretime öğrenci seçme ve yerleştirme sisteminde 1987 yılında yapılan değişiklikler üzerine bir araştırma*. Yayınlanmamış doktora tezi, Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Tokat, E. ve Demirtaşlı, N. (2004). Lisansüstü eğitimi giriş sınavı (LES) ve diğer kabul ölçülerinin yordama geçerliliğine ilişkin bir çalışma. *Eğitim Bilimleri ve Uygulama*, 3(5), 35-55.