	<b>SAKARYA ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ DERGİSİ</b> <i>SAKARYA UNIVERSITY JOURNAL OF SCIENCE</i>		
	e-ISSN: 2147-835X Dergi sayfası: <a href="http://dergipark.gov.tr/saufenbilder">http://dergipark.gov.tr/saufenbilder</a>		
	<u>Geliş/Received</u> 05-05-2017 <u>Kabul/Accepted</u> 15-08-2017	<u>Doi</u> 10.16984/saufenbilder.310730	

## Çoklu doğrusal bağlantı durumunda doğrusal karma modelin kullanımı ve bir uygulama

Melike Bahçecitapar\*<sup>1</sup>, Serpil Aktaş<sup>2</sup>

### ÖZ

Doğrusal Karma Model (DKM), doğrusal regresyon modelinin bir uzantısıdır ve sabit etkilere ilave olarak rastgele etkilerin modele eklenmesi durumunda ortaya çıkan istatistiksel bir modeldir. Son yıllarda, DKM'ler özellikle, aynı birimlerin zaman boyunca gözlemlenmesi ile edilen uzun süreli verilerin analizi için sık kullanılmaktadır. DKM, bir bağımlı değişken ve birden fazla bağımsız değişken arasındaki ilişkiyi homojen olmayan birimlerin birim-özel etkilerini hesaba katarak inceleyebilmektedir. Bu çalışmanın amaçları, literatürdeki çalışmalardan farklı olarak, zaman noktasının bir olduğu durumda DKM'in kullanımını araştırmak ve bağımsız değişkenlerin birbiriyle ilişkili olması olarak tanımlanan çoklu doğrusal bağlantı sorununun olması durumunda DKM ile çoklu doğrusal regresyon modelini karşılaştırmaktır. Bu çalışmada, 2015 yılında Uluslararası Öğrenci Değerlendirme Programı (PISA)'na katılan ülkelerin fen bilimleri okuryazarlığı ortalama puanları ve bu puanlar üzerinde etkisi olduğu düşünülen bazı sosyoekonomik değişkenlerin verileri kullanılarak, DKM ile bir modelleme çalışması yapılmış ve çoklu doğrusal bağlantı sorununun DKM ve çoklu doğrusal regresyon modeli ile hesaplanan parametre kestirimleri üzerinde nasıl bir etkisi olduğu incelenmiştir. Çoklu doğrusal bağlantı sorunu ortadan kaldırıldığında, çoklu doğrusal regresyon modeli ile DKM'in oldukça benzer sonuçlar verdiği görülmüştür.

**Anahtar Kelimeler:** doğrusal karma model, çoklu doğrusal regresyon modeli, çoklu doğrusal bağlantı, PISA 2015 fen bilimleri puanları

## Use of linear mixed model in multicollinearity and an application

### ABSTRACT

A linear mixed model (LMM) is an extension of linear regression model to include both fixed effects and random effects. Recently, LMMs have been widely used especially for the analysis of longitudinal data taken from same subjects over time. The LMM can analyze the relationship between a dependent variable and two or more independent variables by taking into account subject-specific effects of nonhomogeneous subjects. Unlike previous studies, the aims of this study are to analyze the data taken at one time point by LMM and to compare LMM and multiple linear regression model in the presence of multicollinearity problem which occurs when two or more independent variables in the model are correlated with each other.

\* Sorumlu Yazar / Corresponding Author

<sup>1</sup> Hacettepe Üniversitesi, Fen Fakültesi, İstatistik Bölümü, Beytepe, Ankara, 06800 – [mlk@hacettepe.edu.tr](mailto:mlk@hacettepe.edu.tr)

<sup>2</sup> Hacettepe Üniversitesi, Fen Fakültesi, İstatistik Bölümü, Beytepe, Ankara, 06800 – [spxl@hacettepe.edu.tr](mailto:spxl@hacettepe.edu.tr)

We interpret the parameter estimates with application to the data set from Programme for International Student Assessment (PISA) in 2015. In this study, statistical analysis was performed for mean success scores of countries at science in PISA 2015 and data for some socioeconomic variables of countries which can be thought as factors affecting countries' scores at science. A modeling study was conducted by taking into account of effect of multicollinearity problem. According to the findings of the study, both models have similar results without multicollinearity problem.

**Keywords:** linear mixed model, multiple linear regression model, multicollinearity, PISA 2015 science scores

## 1. GİRİŞ (INTRODUCTION)

Doğrusal Karma Model (DKM) (Linear Mixed Model) [1], genel doğrusal modellerin özel bir türüdür ve bağımlı değişkenin ( $Y$ ) normal dağılım gösterdiği durumlar için kullanılan bir karma modeldir. Karma model olarak adlandırılmasının nedeni,  $k$  tane bağımsız değişkenin incelendiği bir doğrusal modelde sabit etkilere ( $\beta_j, j=1, \dots, k$ ) rastgele etkilerin ( $\gamma_i, i=1, \dots, m$ ) de eklenmesidir. Rastgele etkiler, her bir birimin ( $i=1, \dots, m$ ) birim-özel etkileridir. DKM, bazı yönleri bakımından klasik çoklu doğrusal regresyon modele benzerken, sabit etkilerin yanısıra rastgele etkileri de analizde dikkate alabilme özelliklerinden dolayı çoklu doğrusal regresyondan farklı olmaktadır. Son yıllarda, tekrarlı ölçümlü (repeated), uzun süreli (longitudinal) ya da çok aşamalı (multilevel) gibi çeşitli yapılarıdaki verilerin analizi için farklı DKM'ler kullanılmaktadır [2-4].

Uzun süreli veriler, aynı birimlerin farklı zaman noktalarında gözlemlenmesi ile elde edilen verilerdir. Uzun süreli verilerin DKM analizi, birimlerin zaman boyunca gösterdikleri bireysel değişimlerin detaylıca yorumlanmasını sağlar. Modele bir rastgele etki eklenmesi ile birimlerin farklı bir başlangıç değerine sahip olduğu varsayılır. Her bir bağımlı değişken için, farklı rastgele kesim noktaları varsayılarak, bireysel farklılıklar modellenir. DKM'de hata terimi ( $\varepsilon_i$ ) ile birim-içi değişim, rastgele etkiler ile birimden birime bireysel değişimin nasıl farklılık gösterdiğini açıklayan birim-arası değişim ayrı ifade edilebilmektedir [5]. DKM'de parametre tahmini en küçük kareler, en çok olabilirlik (maximum likelihood, ML) ve kısıtlandırılmış en çok olabilirlik (restricted maximum likelihood, REML) gibi yöntemlere dayanmaktadır [6,7]. DKM'de rastgele etkiler dikkate alınarak, sabit etkilerin genelleştirilmiş en küçük kareler tahminleri elde edilir. Rastgele etkilerin tahmini

için ampirik Bayes yöntemleri ve varyans-kovaryans parametrelerin tahmini için olabilirliğe dayalı yöntemlerden birisi tercih edilmektedir. Olabilirlik fonksiyonlarının çözümü karmaşıktır ve Beklenti Maksimizasyonu (Expectation Maximization EM), Newton-Raphson ve Fisher Skorlama (Fisher Scoring) gibi iteratif algoritmalar gerektirmektedir [1,8]. DKM'in uygulanabilmesi için rastgele etkilerin ve hataların birbirinden bağımsız olması ve normal dağılım göstermeleri, birimlerin birbirinden bağımsız olması, hataların sabit varyansa sahip olması ve bağımlı değişken ile bağımsız değişkenlerin doğrusal olarak ilişkili olması gibi temel varsayımların geçerli olması gerekir [1]. Normal dağılım göstermeyen, özellikle kesikli verilerin analizi için genelleştirilmiş karma modeller kullanılmaktadır [9].

Normal dağılım gösteren bağımlı değişkenlerin analizinde yaygın olarak kullanılan istatistiksel yöntemlerden bir diğeri klasik doğrusal regresyon analizidir. Çoklu doğrusal regresyon analizinde, bir bağımlı değişken ve bağımlı değişkeni açıkladığı varsayılan iki veya daha fazla bağımsız değişken arasındaki ilişkiyi açıklamak için bir doğrusal regresyon modeli kullanılır. Çoklu doğrusal regresyonda parametreler genellikle en küçük kareler (EKK) ve en çok olabilirlik yöntemleri ile tahmin edilir. Çoklu doğrusal regresyon modelinin temel varsayımları, hata teriminin sıfır ortalama ve sabit varyans ile normal dağılım göstermesi, hata terimleri arasında otokorelasyonun olmaması, hata terimi ile bağımsız değişkenler arasında bir ilişkinin olmaması, bağımsız değişkenler arasında çoklu doğrusal bağlantı durumunun olmaması; yani, bağımsız değişkenler arasında doğrusal bir ilişkinin bulunmaması şeklinde ifade edilebilir [10].

DKM ve çoklu doğrusal regresyon modeli ile analizlerin amacı, aralarında sebep-sonuç ilişkisi bulunan birden fazla değişken arasındaki ilişkiyi ortaya çıkararak tahminler yapmak ve bu

tahminler yardımıyla değişkenler arasındaki fonksiyonel ilişkiyi matematiksel olarak modellemektir. Tahmin edilen modelin doğru modele yaklaşması için tüm varsayımların sağlanıp sağlanmadığı kontrol edilmelidir. Eğer varsayımlarda bozulmalar söz konusu olursa, modelden yapılacak yorumlar güvenilir olmayacaktır.

Çoklu doğrusal bağlantı sorunu, çok sayıda bağımsız değişkenin incelendiği çalışmalarda en çok karşılaşılan sorunlardan birisidir. İncelenen değişkenlerden bazıları arasında gerçekten ilişki olması ve araştırmacıdan kaynaklanan hatalar çoklu bağlantıya neden olabilir. Çoklu bağlantı durumunda, regresyon katsayıları tahminlerinin varyans ve kovaryans değerleri olması gerekenden büyük olacak, dolayısıyla güven aralıkları genişleyecek ve önem kontrollerinde hesaplanan test istatistik değerleri olması gerekenden küçük çıkacaktır. Bunun yanı sıra, bir veya daha fazla regresyon katsayısının tahminleri yanlış işaretli olabilecektir. Bu durum, modeldeki parametre tahminlerinin yanlış işaretli olmasına bağlı olarak, yapılan yorumların hatalı olmasına neden olabilecektir [11,12]. Çoklu doğrusal bağlantı sorunun ortaya çıkması durumunda, ilgili bağımsız değişken veya değişkenler modelden çıkartılabilir, temel bileşenler regresyonu ya da Ridge regresyon yöntemi kullanılabilir [11,13,14].

Bu çalışmada, DKM'in genel kullanımından farklı olarak tekrar sayısının olmadığı; yani, zaman noktasının bir olduğu durum için DKM'in bir uygulaması yapılmış ve çoklu doğrusal bağlantı durumunda DKM ile çoklu doğrusal regresyon analizi karşılaştırılmıştır. Uygulama için, 2015 yılında Uluslararası Öğrenci Başarılarını Değerlendirme Programı (Programme for International Student Assessment: PISA)'na katılan ülkelerin fen bilimleri okuryazarlığındaki ortalama puanları ve bu puanları etkilediği düşünülen bazı sosyoekonomik değişkenlerin 2014 yılına ait verileri kullanılmıştır [15]. PISA'da araştırmaya katılan ülkelerdeki 15 yaş grubu öğrencilerin matematik okuryazarlığı, fen bilimleri okuryazarlığı ve okuma becerileri alanlarında aldığı puanlar değerlendirilerek ülkelerin eğitim sistemleri karşılaştırılmaktadır. Literatürde, ülkelerin ortalama puanları üzerinde öğrencilerin kendisi, aile ve okul ortamları ile ilgili değişkenlerin etkisi ki-kare analizi, lojistik regresyon modeli, aşamalı doğrusal modeller, yapısal eşitlik modeli, çoklu regresyon analizi vb. çok sayıda istatistiksel yöntem ile incelenmiştir

[16-21]. Buna karşın, ülkelerin sosyoekonomik özelliklerinin puanlar üzerindeki etkisini analiz etmeye yönelik yürütülen çalışma sayısının az olduğu görülmüştür.

Bu çalışmanın amaçları, zaman noktasının bir olduğu durumda çoklu doğrusal bağlantı sorununu dikkate alarak DKM ile bir modelleme çalışması yapmak, DKM ile çoklu doğrusal regresyon modelini karşılaştırmak ve literatürdeki çalışmalardan farklı olarak, ülkelerin PISA araştırmalarında özellikle dikkate alınan sosyoekonomik değişkenlerinin ortalama fen bilimleri başarı puanlarındaki etkisini DKM ile incelemektir.

## 2. MATERYAL VE YÖNTEMLER (MATERIAL AND METHODS)

Bu çalışmada, 2015 yılında PISA'ya katılan ülkelerin “kişi başına düşen gayri safi milli hasılası”, “6-15 yaş grubundaki öğrenci başına kümülatif harcaması”, “yüksek öğrenimli 35-44 yaş arasındaki bireylerin yüzdesi”, “dezavantajlı sosyoekonomik kökenli öğrencilerin payı” ve “15 yaş grubundaki göçmen öğrencilerin oranı” gibi beş sosyoekonomik değişkenin ortalama fen bilimleri başarı puanları üzerindeki etkisi çoklu doğrusal regresyon modeli ve DKM ile incelenmiştir. İncelenebilecek başka birçok sosyoekonomik değişken bulunmasına rağmen, bu çalışmanın kapsamı 2009 yılından itibaren PISA çalışmalarında özellikle dikkate alınan değişkenlerle sınırlandırılmıştır [15].

### 2.1. Çoklu Doğrusal Regresyon Modeli (Multiple Linear Regression Model)

Çoklu doğrusal regresyon modeli  $i$ . ülke için,

$$Y_i = \beta_0 + \beta_1 X_{i1} + \dots + \beta_5 X_{i5} + \varepsilon_i, \quad (1)$$

şeklinde ifade edilmiştir. Eşitlik (1)'de,  $Y_i$ ,  $i=1, \dots, m$ ,  $i$ . ülkenin fen bilimleri puanı,  $X_{ij}$ ,  $j=1, \dots, 5$ ,  $i$ . ülkenin  $j$ . bağımsız değişken değeri ve  $\varepsilon_i$ , rastgele hata terimidir. Eşitlik (1)'de ifade edilen regresyon denkleminin sabit terimi  $\beta_0$ ,  $X_{ij}=0$  iken  $Y_i$ 'nin alacağı değeri ve  $j$ . regresyon katsayısı  $\beta_j$ ,  $X_{ij}$ 'deki bir birim değişimin  $Y_i$ 'de yaratacağı değişmeyi gösterir. Eşitlik (1) matris formunda,

$$Y = X\beta + \varepsilon$$

şeklinde ifade edilir. Burada,  $Y$ ,  $m \times 1$  boyutlu bağımlı değişken vektörü,  $X$ ,  $m \times 6$  boyutlu bağımsız değişkenler matrisi,  $\beta$ ,  $6 \times 1$  boyutlu

bilinmeyen regresyon katsayıları vektörü ve  $\varepsilon$ ,  $m \times 1$  boyutlu hata terimleri vektörüdür.

$\beta$ 'nın en küçük kareler (EKK) tahmin edicisi

$$\hat{\beta}=(X'X)^{-1}X'Y$$

şeklinde bulunur ve  $E(\hat{\beta})=\beta$ 'dir. EKK'in uygun bir tahmin edici olabilmesi için aşağıdaki temel varsayımların geçerli olması gerekmektedir [10]:

- Bağımlı ve bağımsız değişkenler arasındaki ilişki doğrusaldır.
- Hata terimlerinin beklenen değeri sıfırdır ( $E(\varepsilon|X)=\mathbf{0}_{m \times 1}$ ).
- Hata terimlerinin varyansı sabittir ve aralarında otokorelasyon yoktur ( $E(\varepsilon\varepsilon')=\sigma^2 I_{m \times 1}$ ).
- Bağımsız değişkenler ile hata terimleri arasında bir ilişki yoktur ( $E(X'\varepsilon)=\mathbf{0}$ ).
- $X$  açıklayıcı değişken matrisinin sütunları birbirinden bağımsızdır, yani, çoklu doğrusal bağlantı sorunu yoktur.
- Hata terimleri normal dağılıma sahiptir ( $\varepsilon \sim N(\mathbf{0}, \sigma^2 I)$ ).

Yukarıdaki varsayımlardan birisinin sağlanmaması durumunda EKK tahmin edicisi yanlış, tutarsız ve etkisiz olmaktadır [22]. Hataların normal dağılıp dağılmadığı çarpıklık ve basıklık ölçüleri, grafik yöntemi ve normallik testleri ile araştırılabilir. Değişen varyanslılık sorununun araştırılmasında grafik yöntemi, Glejser testi ve Spearman sıra korelasyon testi kullanılabilir. Otokorelasyon sorunun ortaya çıkarılmasında grafik yöntemi ve Durbin-Watson (d) ile Von-Neumann testleri sık kullanılan yöntemlerdir. Değişen varyanslılık durumunda regresyon katsayıları yansız olmasına karşın büyük standart hatalara sahip olacağından, parametrelere ilişkin güven aralıkları genişleyecek ve katsayılara ilişkin testler yanlış yorumlanacaktır. Hatalar arasında otokorelasyon varlığı regresyon katsayılarının standart hatalarının olması gerekenden düşük çıkmasına neden olabilir [11]. Çok sayıda bağımsız değişkenin incelendiği çalışmalarda bağımsız değişkenlerin ilişkili olması; yani, çoklu doğrusal bağlantı sorunu sık karşılaşılan bir durumdur. Çoklu doğrusal bağlantı sorununun tespit edilmesi için basit korelasyon katsayısı, varyans artış değişkenü (Variance Inflation Factors-VIF), tolerans değeri (T), bağımsız değişkenlere ilişkin korelasyon matrisinin özdeğerleri ( $\lambda_j$ ) ve bu özdeğerlere ilişkin koşul endeksleri (Condition Index=CI) hesaplanabilir. Çoklu bağlantıya neden olduğu düşünülen

bağımsız değişkenlerin modele alınmaması, gerekli başka bir değişkenin modele eklenmesi, gözlem sayısının artırılması, modelin matematiksel kalıbının değiştirilmesi veya Ridge regresyon analizi çoklu doğrusal bağlantı durumunda tercih edilen bazı yöntemlerdir [11, 22-25].

## 2.2. Doğrusal Karma Model (Linear Mixed Model)

Doğrusal karma model (DKM),  $i$ . ülke için Eşitlik (2)'deki gibi ifade edilmiştir:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \dots + \beta_5 X_{i5} + \gamma_i + \varepsilon_i. \quad (2)$$

Eşitlik (2)'de,  $Y_i$ ,  $i=1, \dots, m$ ,  $i$ . ülkenin ortalama fen bilimleri puanı,  $X_{ij}$ ,  $i$ . ülkeye ilişkin  $j$ . bağımsız değişkenin değeri,  $\beta_j$ ,  $j=1, \dots, 5$ ,  $j$ . bağımsız değişkenin bilinmeyen regresyon katsayısı,  $\gamma_i$ ,  $i$ . ülkenin rastgele etkisi ve  $\varepsilon_i$  rastgele hata terimidir. Eşitlik (2)'de verilen  $i$ . ülkeye ilişkin DKM'de,  $\gamma_i$  rastgele etkilerin ve  $\varepsilon_i$  rastgele hataların 0 ortalama ve sırasıyla,  $\sigma_\gamma^2$  ve  $\sigma^2$  varyansları ile normal dağıldıkları ( $\gamma_i \sim N(0, \sigma_\gamma^2)$ ,  $\varepsilon_i \sim N(0, \sigma^2)$ ) ve birbirinden bağımsız olduğu varsayılmaktadır.

Eşitlik (2)'de ifade edilen DKM matris formunda,

$$Y = X\beta + Z\gamma + \varepsilon \quad (3)$$

şeklinde ifade edilir. Eşitlik (3)'te,  $Y$ ,  $m \times 1$  boyutlu bağımlı değişken vektörü,  $X$ ,  $m \times 6$  boyutlu modeldeki sabit etkilere ilişkin tasarım matrisi ( $\text{rank}(X) \leq m$ ),  $\beta$ ,  $6 \times 1$  boyutlu bilinmeyen sabit etkiler vektörü,  $Z$ , elemanları 1 olan  $m \times 1$  boyutlu modeldeki rastgele etkilere ilişkin bir tasarım vektörü,  $\gamma$ ,  $1 \times m$  boyutlu rastgele etkiler vektörü ve  $\varepsilon$ ,  $m \times 1$  boyutlu hata vektörüdür.

Tek bir zaman noktasında elde edilen verilerin analizi için kullanılan DKM'in varsayımları çoklu doğrusal regresyon modeli varsayımlarına benzerdir. DKM'de  $\gamma$  ve  $\varepsilon$  terimleri,  $\gamma \sim N(0, \Omega)$  ve  $\varepsilon \sim N(0, A)$  olan normal dağılıma sahip rastlantı değişkenleridir.  $\Omega$  ve  $A$ , sırasıyla, rastgele etkilere ve rastgele hatalara ilişkin varyans-covaryans matrisidir ve DKM için  $\text{Cov}(\varepsilon, \gamma) = \mathbf{0}_{m \times 1}$  'dir. DKM'de,

$$E(Y) = X\beta, \quad (4)$$

$$\text{Cov}(Y) = Z\Omega Z' + A = \Sigma \quad (5)$$

'dir. Eşitlik (4) ve Eşitlik (5)'ten,  $Y \sim N(X\beta, \Sigma)$  olarak ifade edilir.  $l/\beta$ ,  $\beta$ 'nin tahmin edilebilir bir fonksiyonu olmak üzere ve  $\Omega$  ve  $A$  bilindiğinde,

$l'\beta$  genelleştirilmiş en küçük kareler tahmin edicisi (GEKK) ile tahmin edilebilir.

$$\hat{\beta}=(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Y$$

olmak üzere,  $l'\beta$  'nın GEKK tahmin edicisi  $l'\hat{\beta}$  'dır. DKM varsayımları altında,  $l'\hat{\beta}\sim N(l'\beta, l'(X'\Sigma^{-1}X)l)$  'dir. Genellikle,  $\Omega$  ve  $\Lambda$  bilinmez.  $\Omega$  ve  $\Lambda$  'nin en çok olabilirlik veya kısıtlandırılmış en çok olabilirlik tahminleri altında,  $\hat{\Sigma}=\mathbf{Z}\hat{\Omega}\mathbf{Z}'+\hat{\Lambda}$  ve

$$\hat{\beta}_{\hat{\Sigma}}=(X'\hat{\Sigma}^{-1}X)^{-1}X'\hat{\Sigma}^{-1}Y$$

olmak üzere,  $l'\beta$  'nın GEKK tahmin edicisi  $l'\hat{\beta}_{\hat{\Sigma}}$  ile tahmin edilir [2,26]. Bu çalışmada, DKM 'de kısıtlandırılmış en çok olabilirlik fonksiyonlarının çözümü için Fisher Skorlama algoritması kullanılmıştır.

### 3. BULGULAR (RESULTS)

PISA 2015 araştırmasına katılan ülkelerin fen bilimleri okuryazarlığı alanında aldıkları ortalama puanların ve bu puanları etkilediği düşünülen beş sosyoekonomik değişkenin istatistiksel analizi IBM SPSS 20 programı yardımıyla yapılmıştır.

Bu çalışmada kullanılan bağımlı değişken;

Y: Ortalama fen bilimleri puanı (FEN),

ve bağımsız değişkenler;

$X_1$ : Kişi başı gayri safi milli hasıla (PPP kullanarak dönüştürülmüş bin \$) (GSMH),

$X_2$ : 6-15 yaş grubundaki öğrenci başına kümülatif harcama (PPP kullanarak dönüştürülmüş bin \$) (KHarcama),

$X_3$ : 35-44 yaş arasındaki yüksek öğrenimli bireylerin yüzdesi (%) (YÖBY),

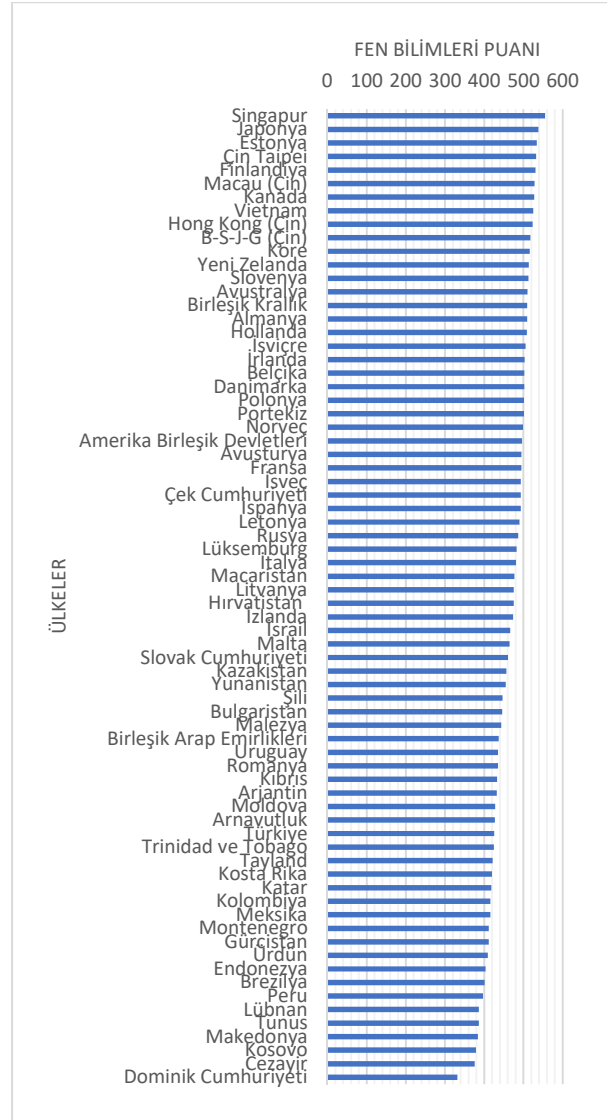
$X_4$ : Dezavantajlı sosyoekonomik kökenli öğrencilerin payı (%) (DSKÖP),

$X_5$ : 15 yaş grubundaki göçmen öğrencilerin oranı (%) (GÖO),

olarak tanımlanmıştır.

Şekil 1 'de, PISA 2015 araştırmasına katılan toplam 72 ülkenin fen bilimleri okuryazarlığındaki ortalama başarı puanları gösterilmiştir. Singapur, Japonya ve Estonya en başarılı ülkelerdir. Bunun yanı sıra, Uzakdoğu ülkelerinin üst sıralarda yer aldığı görülmektedir. Örneğin, Singapur 556, Japonya 538, Çin (Tapei) 532, Macau (Çin) 529 ve Vietnam 525 puanlarına sahiptir. Türkiye'nin

PISA 2015 fen bilimleri puanı 425 olup birçok ülkenin gerisinde kalmıştır.



Şekil 1. PISA 2015 araştırmasına katılan 72 ülkenin fen bilimleri okuryazarlığı ortalama puanları (Mean scores at science of 72 countries in PISA 2015)

Türkiye'nin bağımlı ve bağımsız değişkenlere ilişkin değerleri Tablo 1 'de ve tüm ülkelerin bağımlı ve bağımsız değişkenlerinin değerleri üzerinden hesaplanan bazı tanımlayıcı istatistikler Tablo 2 'de gösterilmektedir.

Tablo 1. Türkiye için bağımlı ve bağımsız değişkenlerin değerleri (Values of dependent and independent variables for Turkey)

Değişken	Değerler
FEN	425
GSMH	19788
KHarcama	32752
YÖBY	15.90
DSKOP	64.40
GÖO	0.30

Tablo 2. Tanımlayıcı istatistikler (Descriptive statistics)

Değişken	<i>m</i>	En Küçük Değer	En Büyük Değer	Ortalama $\pm$ St. Sapma
FEN	72	332 (Dominik Cum.)	556 (Singapur)	465.19 $\pm$ 49.455
GSMH	71	4983 (Moldova)	138050 (Katar)	33555.17 $\pm$ 24674.629
KHarcama	53	11704 (Gürcistan)	187459 (Lüksemburg)	76976.68 $\pm$ 39794.364
YÖBY	41	8.50 (Endonezya)	60.60 (Kanada)	35.8268 $\pm$ 13.297
DSKÖP	72	1.60 (İzlanda)	80.40 (Vietnam)	25.4139 $\pm$ 19.470
GÖO	72	0.00 (Vietnam ve Cezayir)	40.00 (Katar)	4.8264 $\pm$ 7.397

*m*: Ülke sayısı

Tablo 3. Değişkenler arasındaki Pearson korelasyon katsayıları (*r*) ve önem kontrolleri (Pearson correlation coefficients (*r*) between variables and statistical significance tests)

	Y (FEN)	X <sub>1</sub> (GSMH)	X <sub>2</sub> (KHarcama)	X <sub>3</sub> (YÖBY)	X <sub>4</sub> (DSKÖP)	X <sub>5</sub> (GÖO)
Y (FEN)	1	0.478**	0.562**	0.612**	-0.792**	0.321*
X <sub>1</sub> (GSMH)		1	0.923**	0.616**	-0.585**	0.729**
X <sub>2</sub> (KHarcama)			1	0.582**	-0.625**	0.669**
X <sub>3</sub> (YÖBY)				1	-0.695**	0.569**
X <sub>4</sub> (DSKÖP)					1	-0.363**
X <sub>5</sub> (GÖO)						1

\*\* :  $p < 0.01$

\* :  $p < 0.05$

Tablo 1 ve Tablo 2’den, kişi başına düşen gayri safi milli hasıla (bin \$) (GSMH), 6-15 yaş grubundaki öğrenci başına yapılan kümülatif harcama (bin \$) (KHarcama) ve 35-44 yaş arasındaki yüksek öğrenimli bireylerin yüzdesi (%) (YÖBY) bağımsız değişkenleri ve FEN bağımlı değişkeni bakımından Türkiye’nin genel ortalamanın çok altında değerlere sahip olduğu görülmektedir. Buna karşın, Türkiye’deki dezavantajlı sosyoekonomik kökenli öğrencilerin payı (DSKÖP) %64.4’tür ve genel ortalamadan (%25.41) oldukça yüksektir. PISA 2015 araştırmasında en düşük fen puanına sahip olan ülke Dominik Cumhuriyeti’dir. Vietnam’da PISA 2015 araştırmasına katılan göçmen öğrencilerin oranı (GÖO) sıfır olmasına karşın, DSKÖP %80 olup, en yüksek orana sahip olan ülkedir.

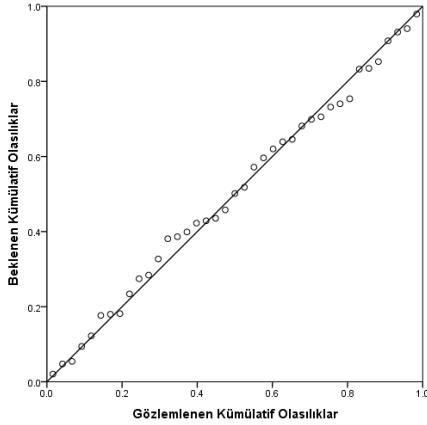
Çoklu doğrusal regresyon ve DKM kayıp gözleme duyarlı olduğundan, SPSS programı, tüm bağımsız değişkenler bakımından kayıp gözlem içermeyen 39 ülkenin verilerini analiz etmiştir. Analizlere başlamadan önce, Pearson korelasyon katsayısı (*r*) ile değişkenler arasındaki ilişkilerin yönü ve derecesi incelenmiştir.

Tablo 3’te korelasyon katsayıları ve önem kontrollerinden görüldüğü gibi, ülkelerin fen bilimleri okuryazarlığındaki başarı puanları ve tüm bağımsız değişkenler arasındaki ilişkiler önemlidir ( $p < 0.01$  ve  $p < 0.05$ ). Bağımlı değişken FEN ile en yüksek ilişkiye sahip olan bağımsız

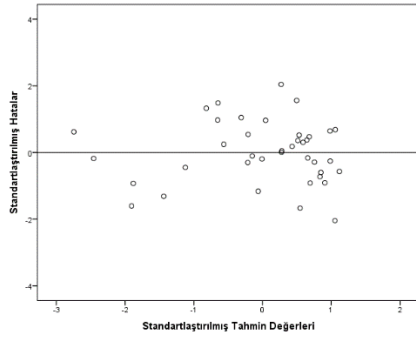
değişkenin DSKÖP olduğu bulunmuştur. FEN bağımlı değişkeni ve DSKÖP bağımsız değişkeni arasında negatif yönde anlamlı bir ilişki söz konusudur ( $r = -0.792$ ,  $p < 0.01$ ). DSKÖP arttıkça, fen puanları azalmaktadır. Bunun yanı sıra, GSMH, KHarcama, YÖBY ve GÖO bağımsız değişkenleri fen puanıyla pozitif ilişkilidir. Bağımsız değişkenler arasındaki ilişkiler incelendiğinde, GSMH ve KHarcama bağımsız değişkenleri arasında 0.90’ın üzerinde önemli derecede pozitif yönlü bir ilişki bulunmaktadır ( $r = 0.923$ ,  $p < 0.01$ ). Bu durum, çoklu bağlantı sorununun olabileceğini göstermektedir. Diğer bağımsız değişkenler arasında ise düşük ve orta düzeyde fakat önemli bir ilişki vardır.

Analizler için gerekli olan regresyon modeli varsayımlarının sağlanıp sağlanmadığı kontrol edilmiştir. Normallik ve sabit varyans varsayımları standartlaştırılmış hatalar kullanılarak test edilmiştir. Şekil 2’de Normal P-P grafiğinde hataların doğru etrafında toplanması ve Şekil 3’te saçılım grafiğinde tahmin edilen puanlar ile standartlaştırılmış hatalar arasındaki ilişkinin düz bir doğru şeklinde olması hataların normal dağıldığını ve hataların sabit varyansa sahip olduğunu göstermektedir. Otokorelasyon varsayımı Durbin-Watson test istatistiği ile kontrol edilmiştir. Durbin-Watson istatistiği  $d = 1.899$  olup, Durbin-Watson tablosunda anlamlılık düzeyi  $\alpha = 0.05$ , gözlem sayısı  $m = 39$  ve bağımsız değişken sayısı  $k = 5$  için kritik değerler olan  $d_L = 1.033$  ve

$d_U=1.583$  ile karşılaştırıldığında, hataların birbirinden bağımsız olduğu söylenebilir.



Şekil 2.  $m=39$  ve  $k=5$  için standartlaştırılmış artıkların Normal P-P grafiği (Normal P-P graphics for standardized residuals, when  $m=39$  and  $k=5$ )



Şekil 3.  $m=39$  and  $k=5$  için standartlaştırılmış tahmin değerleri ve hataların saçılım grafiği (Scatter plot for standardized predicted values and residuals, when  $m=39$  and  $k=5$ )

Çoklu doğrusal bağlantı sorununun olup olmadığını tespit etmek için ilk olarak VIF ve T değerleri hesaplanmıştır. Tablo 4’te modelde yer alan bağımsız değişkenlerin VIF değerleri 10’dan büyük ve T değerleri 0.10’dan küçük olmadığı için değişkenler arasında güçlü bir çoklu bağlantının olmadığına karar verilmiştir [26]. Ancak, GSMH ve KHarcama değişkenlerinin tolerans değerleri 0.10’a çok yakın bir değerlerdir (GSMH için

$T=0.121$ , KHarcama için  $T=0.135$ ). Bu nedenle, VIF ve T değerlerine ilaveten, çoklu bağlantının olup olmadığını tespit edilmesinde kullanılan diğer yöntemlerden bağımsız değişkenlere ait korelasyon matrisinin özdeğerleri ve koşul endeksleri (Condition Index:CI) hesaplanmıştır.

Tablo 4’te görüldüğü gibi, son iki özdeğer sıfıra çok yakındır. Bunun yanısıra, özdeğerlerin terslerinin toplamı  $\sum_{j=1}^5 1/\lambda_j=187.333$  olup, değişken sayısı 5’ten oldukça büyüktür. Bu iki durum çoklu bağlantının varlığını güçlendirmektedir. Koşul endeksleri değerlendirildiğinde, son iki koşul endeksinin 15 ile 30 arasında olması, modelde orta derecede çoklu bağlantı sorununun varlığını göstermektedir [23, 24, 26]. En son özdeğere ilişkin varyans oranlarına bakıldığında, GSMH ve KHarcama değişkenlerinin çoklu doğrusal bağlantı sorununa neden olduğu söylenebilir.

PISA 2015 fen puanlarının çoklu doğrusal regresyon modeli ve DKM ile yapılan analiz sonuçları Tablo 5’te gösterilmektedir. Analizlerde tüm bağımsız değişkenler aynı anda regresyon denkleminde alınmıştır. Anlamlılık düzeyi olarak 0.05 belirlenmiş ve katsayılar buna göre yorumlanmıştır. Tablo 5’te, çoklu doğrusal regresyon analizinde çoklu korelasyon katsayısı  $R=0.812$  ve çoklu belirlilik katsayısı  $R^2=0.660$  olarak bulunmuştur. Fen puanı ile tüm bağımsız değişkenler arasında %80’lik bir ilişki söz konusudur.  $R^2$  değerine göre, fen puanlarındaki değişimin %66’sı GSMH, KHarcama, YÖBY, DSKÖP ve GÖO bağımsız değişkenleri ile açıklanmaktadır. Tablo 5’teki varyans analizi sonuçlarından görüldüğü gibi, tahmin edilen regresyon modeli istatistiksel olarak anlamlıdır ( $F$ -istatistiği=12.812,  $p=0.000<0.05$ ). Bağımsız değişkenler ile FEN bağımlı değişkeni arasında doğrusal bir ilişki olduğuna karar verilmiştir.

Tablo 4. Bağımsız değişkenlerin VIF ve Tolerans (T) değerleri, korelasyon matrisi özdeğerleri ( $\lambda_j$ ), koşul endeksleri (CI) ve varyans oranları (VIF, Tolerance (T) values, eigenvalues of correlation matrix, conditional index and variance proportions for independent variables)

Değişken	VIF	T	$\lambda_j$	CI	Varyans Oranları					
					Sabit Etki	GSMH	KHarcama	YÖBY	DSKOP	GÖO
GSMH	8.259	0.121	0.738	2.588	0.00	0.00	0.00	0.00	0.16	0.06
KHarcama	7.425	0.135	0.228	4.652	0.01	0.00	0.00	0.02	0.11	0.56
YÖBY	2.536	0.394	0.067	8.567	0.01	0.05	0.08	0.29	0.02	0.08
DSKÖP	2.481	0.403	0.015	18.006	0.82	0.17	0.01	0.55	0.55	0.30
GÖO	2.438	0.410	0.010	21.758	0.16	0.77	0.90	0.14	0.15	0.00

Tablo 5. Çoklu doğrusal regresyon modeli ve DKM ile hesaplanan katsayı tahminleri, önem kontrolleri ve bazı istatistikler (Coefficient estimates, significance tests and some statistics obtained by multiple linear regression model and LMM)

Parametre	Çoklu Doğrusal Regresyon Modeli				DKM			
	Tahmin	St. Hata	<i>t</i> -istatistiği	<i>p</i>	Tahmin	St. Hata	<i>t</i> -istatistiği	<i>p</i>
Sabit terim	494.090	22.974	21.506	0.000*	531.884	12.3898	42.929	0.000*
GSMH	-0.001	0.001	-1.324	0.195	3.341E-5	0.0003	0.102	0.919
KHarcama	0.000	0.000	1.545	0.132	4.622E-5	0.0001	0.338	0.737
YÖBY	0.414	0.428	0.967	0.340	0.1163	0.2208	0.527	0.602
DSKÖP	-1.330	0.330	-4.027	0.000*	-0.6085	0.1847	-3.293	0.002*
GÖÖ	-0.029	1.061	-0.027	0.979	-0.3434	0.5439	-0.631	0.532
<i>m</i> =39, <i>R</i> =0.812, <i>R</i> <sup>2</sup> =0.66, <i>F</i> =12.812 ( <i>p</i> =0.000)					$\hat{\sigma}_Y^2=4.2007$ , St. Hata=0.600, Wald Z=0.70 ( <i>p</i> =0.484)			

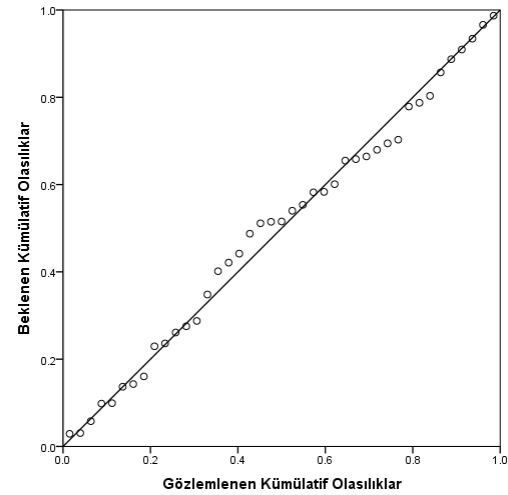
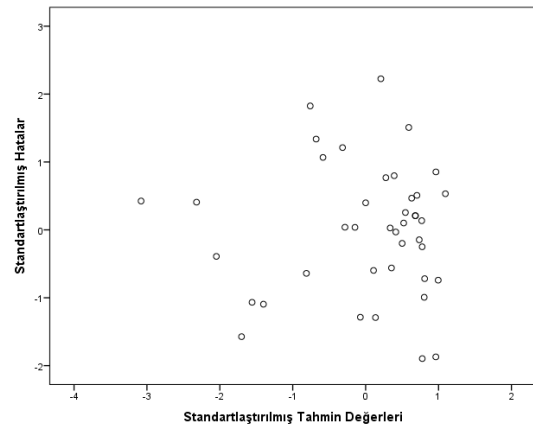
\* : *p*<0.05

Tablo 5'te, çoklu doğrusal regresyon modeli ve DKM için, sırasıyla, standartlaştırılmamış regresyon katsayıları tahminleri ( $\hat{\beta}_j$ ),  $\hat{\beta}_j$ 'lerin standart hatası, *t*-istatistiği ve önem kontrolleri gösterilmektedir. Her iki modelde fen puanları ve değişkenler arasındaki ilişkilerin benzer olduğu görülmektedir. GSMH, DSKÖP ve GÖÖ değişkenleri negatif regresyon katsayılarına sahip iken, KHarcama ve YÖBY değişkenleri fen puanları için pozitif regresyon katsayılarına sahiptir. GSMH, DSKÖP ve GÖÖ bağımsız değişkenleri ile fen puanları arasındaki ilişki ters orantılıdır. Ülkelerin DSKÖP ve GÖÖ arttıkça, fen bilimleri okuryazarlığındaki başarıları azalmaktadır. Buna karşın, KHarcama ve YÖBY bağımsız değişkenleri ile fen puanları arasındaki ilişki doğru orantılıdır ve bu değişkenlerdeki artış fen puanlarındaki başarıyı artırmaktadır.

Tablo 5'te verilen sonuçlara göre, sosyo-ekonomik değişkenlerin PISA 2015 fen bilimleri puanları üzerindeki etkisinin incelendiği her iki modelde, DSKÖP değişkeninin *t* testine göre 0.05 anlamlılık düzeyinde istatistiksel olarak anlamlı olduğu görülmektedir. Fen puanlarını etkileyen en önemli değişken DSKÖP olarak bulunmuştur. Sadece DSKÖP bağımsız değişkeninin regresyona önemli bir katkısı vardır ve DSKÖP fen puanları ile negatif yönde anlamlı ilişkiye sahiptir. DSKÖP'daki artış ülkelerin fen bilimleri okuryazarlığındaki başarısını önemli derecede azaltmaktadır. Bunun yanısıra, KHarcama ve YÖBY değişkenleri fen puanları ile pozitif yönde ilişkiye sahiptir, ancak bu değişkenlere ilişkin regresyon katsayı tahminleri istatistiksel olarak anlamlı bulunmamıştır (*p*>0.05).

Tablo 5'te görüldüğü gibi, DKM sonuçlarının çoklu doğrusal regresyon modeli sonuçlarıyla uyumlu olduğu söylenebilir. Ancak, DKM ile hesaplanan standart hatalar çoklu doğrusal regresyon modeli ile hesaplanan standart

hatalardan daha küçüktür ve tahmin değerleri birbirinden oldukça farklıdır. Tablo 5'te, birim-arası değişimi ifade eden,  $\gamma_i$  rastgele etkilerin varyans tahmini  $\hat{\sigma}_Y^2=4.2007$  (*p*=0.484>0.05) olarak elde edilmiştir.

Şekil 4. *m*=41 ve *k*= 3 için standartlaştırılmış artıkların normal P-P grafiği (Normal P-P graphics for standardized residuals, when *m*=41 and *k*=3)Şekil 5. *m*=41 and *k*=3 için standartlaştırılmış tahmin değerleri ve hataların saçılım grafiği (Scatter plot for standardized predicted values and residuals, when *m*=41 and *k*=3)

Tablo 5 ve Tablo 1 karşılaştırıldığında, Tablo 1'de çoklu korelasyon matrisinde GSMH ile FEN ve GÖÖ ile FEN değişkenleri arasında pozitif yönlü



bir ilişki görülmesine rağmen, Tablo 5’te çoklu doğrusal regresyon analizinde GSMH ve GÖO değişkenlerinin fen puanı üzerinde negatif bir etkiye sahip olduğu bulunmuştur. GSMH’nın fen puanları üzerinde negatif etkisi olması en dikkat çekici sonuçtur. Bu durumun değişkenler arasındaki orta derecede çoklu doğrusal bağlantıdan kaynaklandığı söylenebilir. Çoklu doğrusal bağlantı sorununu ortadan kaldırmak için gerekli analizler yapılmış ve bu soruna neden olduğu görülen GSMH ve KHarcama değişkenleri her iki modele dahil edilmemiştir. YÖBY, DSKÖP ve GÖO değişkenleri bakımından kayıp gözlem içermeyen 41 ülkenin verileri çoklu doğrusal regresyon ve DKM ile tekrar analiz edilmiştir.

Şekil 4 ve Şekil 5’te, 41 ülkenin verilerinin normallik ve sabit varyans varsayımları sağlanmaktadır. Hatalar arasında otokorelasyon olmaması varsayımını araştırmak için Durbin-Watson testi hesaplanmış ve modelde otokorelasyon olmadığına karar verilmiştir ( $d=1.892>d_U=1.659$ ,  $\alpha=0.05$ ,  $m=41$ ,  $k=3$ ). FEN bağımlı değişkeni ve YÖBY, DSKÖP ve GÖO bağımsız değişkenlerine ilişkin çoklu korelasyon matrisi Tablo 6’da gösterilmiştir.

Tablo 6. FEN bağımlı değişkeni ve YÖBY, DSKÖP ve GÖO bağımsız değişkenleri arasındaki Pearson korelasyon katsayıları ve önem kontrolleri (Pearson correlation coefficients between dependent variable FEN and independent variables YÖBY, DSKÖP and GÖO and their significance tests)

	Y (FEN)	X <sub>3</sub> (YÖBY)	X <sub>4</sub> (DSKÖP)	X <sub>5</sub> (GÖO)
Y (FEN)	1	0.664**	-0.814**	0.351*
X <sub>3</sub> (YÖBY)		1	-0.729**	0.580**
X <sub>4</sub> (DSKÖP)			1	-0.386**
X <sub>5</sub> (GÖO)				1

\*\* :  $p < 0.01$

\* :  $p < 0.05$

Tablo 6’da fen puanları ile en yüksek ilişkiye sahip olan bağımsız değişkenin DSKÖP olduğu görülmektedir. FEN ve DSKÖP değişkenleri arasında negatif yönlü ve önemli bir ilişki söz konusudur ( $r=-0.814$ ,  $p<0.05$ ). Bağımsız değişkenler arasında orta derecede önemli bir ilişki vardır. Aralarında en yüksek ilişkiye sahip olan bağımsız değişkenler YÖBY ve DSKÖP olarak bulunmuştur ( $r=-0.729$ ,  $p<0.05$ ).

Tablo 7. YÖBY, DSKÖP ve GÖO bağımsız değişkenlerinin VIF ve Tolerans (T) değerleri, korelasyon matrisi özdeğerleri ( $\lambda_j$ ) ve koşul endeksleri (CI) (VIF, Tolerance (T) values, eigenvalues of correlation matrix, conditional index)

Bağımsız Değişkenler	VIF	T	$\lambda_j$	CI
X <sub>3</sub> (YÖBY)	2.748	0.364	0.717	2.064
X <sub>4</sub> (DSKÖP)	2.142	0.467	0.212	3.797
X <sub>5</sub> (GÖO)	1.513	0.661	0.018	12.904

Ancak, Tablo 7’de VIF, Tolerans ve koşul endeks değerleri ile YÖBY, DSKÖP ve GÖO değişkenlerine ait korelasyon matrisinin özdeğerleri birlikte değerlendirildiğinde, önemsenmeyecek derecede çoklu doğrusal bağlantı sorununun olduğu görülmektedir ( $VIF_j \neq 1$ ,  $VIF_j < 5$ ,  $T_j > 0.10$ , CI değerleri  $< 15$ ,  $\sum_{j=1}^3 1/\lambda_j = 56.95$ ) [11].

Çoklu doğrusal bağlantı sorununun olmadığına karar verilerek incelenen doğrusal regresyon modeli ve DKM’ne göre regresyon katsayı tahminleri Tablo 8’de verilmiştir. Tablo 8’de doğrusal regresyon denkleminin belirlilik katsayısı  $R^2=0.67$  olarak verilmiştir.

Fen puanlarındaki değişimin %67’si YÖBY, DSKÖP ve GÖO bağımsız değişkenleri ile açıklanmaktadır. Bu model için  $F$ -test istatistiği anlamlıdır ve FEN bağımlı değişkeni ile YÖBY, DSKÖP ve GÖO bağımsız değişkenleri arasında doğrusal ilişki olduğuna karar verilmiştir ( $F=25.437$ ,  $p<0.001$ ).

Tablo 8’de görüldüğü gibi, parametre tahminleri ve standart hatalar bakımından, çoklu doğrusal bağlantı sorunu olmadığına DKM ve çoklu doğrusal regresyon modeli oldukça benzer sonuçlar vermektedir.

Seçilen 0.05 anlamlılık düzeyinde, DSKÖP’nın fen puanları üzerinde önemli bir etkisi olduğu bulunmuştur ( $p<0.05$ ). YÖBY ve GÖO bağımsız değişkenlerine ait regresyon katsayıları anlamsız çıkmıştır ( $p>0.05$ ).

Tablo 8 ve Tablo 5 karşılaştırılmak istenirse, Tablo 5’te çoklu doğrusal bağlantı durumunda elde edilen sonuçlara benzer olarak, her iki modelde, sadece DSKÖP bağımsız değişkenine ait regresyon katsayısı anlamlı bulunmuştur. Tablo 8’de fen puanlarını DSKÖP ve GÖO değişkenleri negatif yönde etkilerken, YÖBY değişkeni pozitif yönde etkilediği görülmektedir.

**Tablo 8.** Çoklu doğrusal bağlantı olmaksızın çoklu doğrusal regresyon modeli ve DKM ile hesaplanan katsayı tahminleri, önem kontrolleri ve bazı istatistikler (Coefficient estimates, significance tests and some statistics for multiple linear regression and LMM without multicollinearity)

Parametre	Çoklu Doğrusal Regresyon Modeli				DKM			
	Tahmin	St. Hata	t-istatistiği	p	Tahmin	St. Hata	t-istatistiği	p
Sabit terim	497.238	18.225	27.283	0.000*	497.292	18.193	27.333	0.000*
YÖBY	0.428	0.419	1.021	0.314	0.442	0.420	1.052	0.300
DSKÖP	-1.362	0.266	-5.111	0.000*	-1.344	0.270	-4.970	0.000*
GÖO	-0.086	0.831	-0.104	0.918	-0.099	0.830	-0.119	0.906
$m=41, R=0.821, R^2=0.673, F=25.437 (p=0.000)$					Var( $\gamma_i$ )=0.1390, St. Hata=0.1715, Wald Z=0.081( $p=0.935$ )			

\*:  $p < 0.05$

#### 4. TARTIŞMA VE SONUÇ (DISCUSSION AND CONCLUSION)

Bu çalışmada, önceki çalışmalardan farklı olarak, zaman noktasının bir olduğu durum için DKM'in kullanımı araştırılmış ve çoklu doğrusal bağlantı durumunda DKM ve çoklu doğrusal regresyon modeli karşılaştırılmıştır. Parametre tahminleri ve standart hataları bakımından çoklu doğrusal bağlantı sorunu varlığında, çoklu doğrusal regresyon modeli ile DKM birbirinden oldukça farklı sonuçlar vermektedir. DKM'de hesaplanan standart hatalar çoklu doğrusal regresyon modeline göre daha küçüktür. Buna karşın, çoklu doğrusal bağlantı sorunu olmadığında her iki modelden elde edilen parametre tahminleri ve standart hatalar birbirine oldukça benzerdir.

Uygulama için 2015 yılında PISA araştırmasına katılan ülkelerin fen bilimleri okuryazarlığı puanları ve kişi başına düşen gayri safi milli hasıla, 6-15 yaş grubundaki öğrenci başına yapılan kümülatif harcama, 35-44 yaş arasındaki yüksek öğrenimli bireylerin yüzdesi, dezavantajlı sosyoekonomik kökenli öğrencilerin payı ve 15 yaş grubundaki göçmen öğrencilerin oranı gibi beş sosyoekonomik değişken incelenmiştir. Bu değişkenlerin fen bilimleri puanları üzerindeki etkisi DKM modeli ile açıklanmaya çalışılmıştır.

DKM ile çoklu doğrusal regresyon modelinden elde edilen benzer sonuçlara göre, dezavantajlı sosyoekonomik kökenli öğrencilerin payının PISA 2015 fen puanları üzerinde istatistiksel olarak anlamlı ve negatif yönlü bir etkiye sahip olduğu bulunmuştur. Ülkelerin fen puanları üzerinde 35-44 yaş arasındaki yüksek öğrenimli bireylerin yüzdesi pozitif ve göçmen öğrencilerin oranı ise negatif yönlü bir etkiye sahiptir, ancak istatistiksel olarak bu etkiler anlamlı bulunmamıştır.

Türkiye, PISA kapsamında gerçekleştirilen fen puanları değerlendirilmelerinde birçok ülkenin gerisindedir. Uzakdoğu ülkelerinden Singapur, Japonya, Çin (Tapai), Çin (Macao) ve Vietnam en üst sıralarda yer almaktadır. Dezavantajlı sosyoekonomik kökenli öğrencilerin payının fen puanları üzerinde oldukça etkili olduğu düşünülürse, eğitimde fırsat eşitliği her gelişmekte olan ülke için sağlanması gereken en temel koşullardan birisidir. Türkiye fen bilimleri alanında öğrencilerin başarılarını artırmaya yönelik strateji ve politikalar geliştirmelidir.

#### KAYNAKÇA (REFERENCES)

- [1] N. M. Laird ve J. H. Ware, "Random-effects models for longitudinal data," *Biometrics*, cilt 38, no. 4, pp. 963-974, 1982.
- [2] G. Verbeke ve G. Molenberghs, *Linear Mixed Models for Longitudinal Data*, New York: Springer, 2000.
- [3] S. Raudenbush ve A. S. Bryk, *Hierarchical Linear Models: Applications and Data Analysis Methods*, Thousand Oaks, CA: SAGE Publications, Inc., 2002.
- [4] J. W. R. Twisk, *Applied Multilevel Analysis: A Practical Guide for Medical Researchers*, Cambridge, UK: Cambridge University Press, 2006.
- [5] M. Bahçecitapar, *Uzun Süreli Verilerin Analizinde Kullanılan İstatistiksel Modeller*, Ankara: Hacettepe Üniversitesi Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, 2006.
- [6] H. T. R. Patterson, "Recovery of interblock information when block sizes are unequal," *Biometrika*, pp. 545-554, 58 1971.
- [7] D. A. Harville, "Maximum likelihood approaches to variance component

- estimation and to related problems," *Journal of the American Statistical Association*, pp. 320-338, 358 72 1977.
- [8] F. N. Gumedze ve T. T. Dunne, "Parameter estimation and inference in the linear mixed model," *Linear Algebra Applications*, cilt 435, pp. 1920-1944, 435 2011.
- [9] J. W. Hardin ve J. M. Hilbe, *Generalized Linear Models and Extensions*, 3rd edition,, College Station: TX: Stata Press, 2012.
- [10] T. Yamane, *Statistics: An Introductory Analysis*, New York: Harper&Row, 1969.
- [11] R. Alpar, *Uygulamalı Çok Değişkenli İstatistiksel Yöntemler*, Ankara: Detay Yayıncılık, 2011.
- [12] E. Bulut ve A. Alın, "Kısmi en küçük kareler regresyon yöntemi algoritmalarından nıpls ve pls - kernel algoritmalarının karşılaştırılması ve bir uygulama," *Dokuz Eylül Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, cilt 24, no. 2, pp. 127-138, 2009.
- [13] K. Özkan, "Toprağın tarla kapasitesi değişiminin toprak türüne göre temel bileşenler regresyon analizi ile modellenmesi," *Süleyman Demirel Üniversitesi Orman Fakültesi Dergisi*, cilt A, no. 2, pp. 1-9, 2009.
- [14] A. S. Albayrak, *Uygulamalı Çok Değişkenli İstatistik Teknikleri*, Ankara: Asil Yayın Dağıtım, 2006.
- [15] OECD, *PISA 2015 Results (Volume I): Excellence and Equity in Education*, Paris: OECD Publishing, 2016.
- [16] P. Aşkar ve S. Olkun, "PISA 2003 sonuçları açısından bilgi ve iletişim teknolojileri kullanımı," *Eurasian Journal of Educational Research*, cilt 19, pp. 15-34, 2005.
- [17] F. K. Çelen, A. Çelik ve S. S. Seferoğlu, "Türk Eğitim Sistemi ve PISA Sonuçları," *Akademik Bilişim Konferansı, İnönü Üniversitesi*, Malatya, 2011.
- [18] S. Gürsakal, "PISA 2009 öğrenci başarı düzeylerini etkileyen faktörlerin değerlendirilmesi," *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, cilt 17, no. 1, pp. 441-452, 2012.
- [19] D. Anıl, "Uluslararası öğrenci başarılarını değerlendirme programı(PISA)'da Türkiye'deki öğrencilerin fen bilimleri başarılarını etkileyen faktörler," *Eğitim ve Bilim*, cilt 34, no. 152, pp. 87-100, 2009.
- [20] Y. Özer ve D. Anıl, "Öğrencilerin fen ve matematik başarılarını etkileyen faktörlerin yapısal eşitlik modeli ile incelenmesi," *H.Ü. Eğitim Fakültesi Dergisi*, cilt 41, pp. 313-324, 2011.
- [21] M. Özmuşul ve A. Kaya, "Türkiye'nin PISA 2009 ve 2012 sonuçlarına ilişkin karşılaştırmalı bir analiz," *Journal of European Education*, cilt 4, no. 1, pp. 23-40, 2014.
- [22] A. S. Albayrak, "Çoklu doğrusal bağlantı halinde en küçük kareler tekniğinin alternatifi yanlı tahmin teknikleri ve bir uygulama," *ZKÜ Sosyal Bilimler Dergisi*, cilt 1, no. 1, pp. 105-126, 2005.
- [23] D. Gujarati, *Basic Econometrics*, 3rd Edition, New York: McGraw-Hill International Editions, 1995.
- [24] N. Orhunbile, *Uygulamalı Regresyon ve Korelasyon Analizi*, 2. Baskı, İstanbul: İ.Ü. Basım Yayın, 2002.
- [25] M. Topal, E. Eyduran, A. M. Yağanoğlu, A. Y. Sönmez ve S. Keskin, "Çoklu doğrusal bağlantı durumunda ridge ve temel bileşenler regresyon analiz yöntemlerinin kullanımı," *Atatürk Üniversitesi Ziraat Fakültesi Dergisi*, cilt 41, no. 1, pp. 53-57, 2010.
- [26] D. E. Johnson, "An Introduction to the analysis of mixed models," *SUGI 28 Proceedings*, Seattle, Washington, 2003.