

Çoklu Puanlanan Çok Boyutlu Madde Tepki Kuramında Kayıp Veri Yöntemlerinin İncelenmesi* **

Investigation of Missing Data Methods in Polytomous Multidimensional Item Response Theory

Mehmet Ali İŞİKOĞLU¹, Burcu ATAR²

¹Eskişehir Teknik Üniversitesi, Rektörlük. e-posta: mai@eskisehir.edu.tr

²Hacettepe Üniversitesi, Eğitim Bilimleri Bölümü, Eğitimde Ölçme ve Değerlendirme A. B. D. e-posta: burcuatar@hacettepe.edu.tr

Makale Türü/Article Types: Araştırma Makalesi/ Research Article

Makalenin Geliş Tarihi: 25.06.2023

Yayına Kabul Tarihi: 23.09.2023

ÖZ

Bu çalışmada, bazı kayıp veri ile baş etme yöntemlerinin çoklu puanlanan çok boyutlu madde tepki kuramı modellerinden olan çok boyutlu aşamalı tepki modelinden elde edilen madde ve yetenek parametrelerine etkisinin araştırılması amaçlanmıştır. İlk olarak belirlenen parametrelere uygun şekilde iki boyutlu bir referans veri seti üretilmiştir. Üretilen veri setinden tamamen rassal olarak kayıp mekanizmasında %5, %10 ve %20 oranında kayıplar oluşturulmuştur. Eksik veri setleri seri ortalaması atama (SO), regresyon atama (RA) ve beklenti maksimizasyonu (BM) yöntemleri ile tamamlanmış ve tamamlanan veri setlerinden madde ile yetenek parametreleri kestirilmiştir. Kestirilen parametreler, referans veri setinden kestirilen parametreler ile karşılaştırılmıştır. Elde edilen bulgulara göre madde ayırıcılık parametreleri için tüm kayıp oranlarında RA ve BM yöntemlerinin daha düşük RMSE değerleri ürettiği görülmüştür. Kategori sınır kesişim parametrelerinde ise d_1 ve d_3 için RA ve SO yöntemleri, d_2 için ise BM ve RA yöntemleri tüm kayıp oranlarında daha düşük hata üretmiştir. Genel olarak tüm madde parametreleri değerlendirildiğinde, RA yönteminin parametre değerlerinin geri kazanımında oldukça iyi sonuçlar ürettiği söylenebilir. Yetenek parametrelerinin geri kazanımı için ise tüm kayıp oranlarında en düşük RMSE değerleri üreten yöntemin BM olduğu görülmüştür.

Anahtar Sözcükler: Çoklu puanlanan çok boyutlu mtk, Kayıp veri, Çok boyutlu aşamalı tepki modeli

***Alıntılama:** İşikoğlu, M. A. ve Atar, B. (2023). Çoklu puanlanan çok boyutlu madde tepki kuramında kayıp veri yöntemlerinin incelenmesi. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, 43(3), 1645-1662.

** Bu makale birinci yazarın doktora tez çalışmasından üretilmiştir.

ABSTRACT

In this study, the effects of missing data handling methods on item and ability parameters obtained from the graded response model, which is one of the multidimensional item response theory models with polytomous scoring, are investigated. Firstly, a two-dimensional reference data set was generated. From the generated data set, missings of 5%, 10% and 20% were generated with missing completely at random mechanism. The data sets with missings were completed with series mean imputation (SM), regression imputation (RI) and expectation maximisation (EM) methods. Item and ability parameters were obtained from the completed data sets. The methods were evaluated by comparing the obtained parameters with the parameters obtained from the reference data set. According to the findings, it was observed that RI and EM methods have better results at all rates for item discrimination parameters. For the category boundary intersection parameters, RI and SM methods showed the best performance for d_1 and d_3 . For the d_2 parameter, it was observed that EM and RI produced results with very low error at all rates. In general, for all item parameters, it can be said that the RI method has very good results. For the ability parameters, it was observed that EM produced the lowest RMSE values at all rates.

Keywords: Polytomous MIRT, Missing data, Multidimensional GRM

GİRİŞ

Eğitim alanında gerçekleştirilen araştırmalar kapsamında veri toplanırken, ölçme aracını yanıtlayan bireylerden, ölçme aracından, puanlama aracından, uygulayıcıdan vb. kaynaklanan sebeplerle çoğu zaman veriler eksiksiz şekilde elde edilememektedir. Elde edilen verilerde değişken veya gözlem düzeylerinde kayıp veriler bulunabilmektedir. Veri setinde bulunan kayıplar araştırmalarda bazı sorunlara sebebiyet verebilmektedir. Bu sorunlar arasında bilgi kayıpları ve buna bağlı olarak kestirimlerin hata miktarlarının artması, yanlış sonuçların elde edilmesi ve analizlerin daha karmaşık hale gelmesi gibi sorunlar sayılabilmektedir (Little ve Schenker, 1995).

Eğitim alanında toplanan verilerin analizinde madde tepki kuramı (MTK) çerçevesindeki modellerin kullanımı sıklıkla tercih edilmektedir. MTK modelleri bir bireyin ölçme aracındaki maddeye verdiği yanıt ile bireyin yeteneği arasında ilişki kurmaktadır. Madde temelli olma özelliğinden dolayı maddeye verilen yanıtların eksiksiz şekilde elde edilmesi önemlidir. Ancak, ölçme araçlarından eksiksiz veri elde etmek her zaman kolaylıkla sağlanamamaktadır ve pratikte bu oldukça sık karşılaşılan bir durumdur. Kayıp veri çeşitli sebeplerde veri setinde eksik gözlemin bulunması

olarak tanımlanabilir. Ölçme aracındaki bazı maddelerin yanıtlanmaması, uygulama süresinin yeterli olmamasından kaynaklı olarak bazı maddelere erişememe (Goregebur, de Boeck ve Molenberghs, 2010), uygulama sırasında karşılaşılan teknik aksaklıklar, veri girişi sırasında yapılan hatalar (Field, 2005), boylamsal çalışmalarda deneklerin çalışmalardan çekilmeleri gibi nedenlerle veri setinde kayıplar ortaya çıkabilmektedir. Bu durumda kayıplar ile uygun yöntemleri kullanarak başa çıkılması gerekmektedir.

Veri setinde bulunan kayıp verilerin dağılımı, madde tepki kuramı modelleriyle gerçekleştirilecek olan parametre kestirimlerini etkileyen önemli durumlardan biridir. Kayıp verilerin sistematik olarak dağılımı yansız dağılmasına göre sonuçların genellenebilirliğini daha fazla etkileyebileceği için daha ciddi sorunlar oluşturmaktadır (Tabachnick ve Fidell, 2012). Little ve Rubin (2002) kayıp verilerde Tamamen Rassal Olarak Kayıp (TROK), Rassal Olarak Kayıp (ROK) ve İhmal Edilemez Kayıp (İEK) olarak üç çeşit mekanizma bulunduğunu belirtmiştir. TROK, kayıpların bulunduğu değişkene veya veri setindeki başka bir değişkene bağlı olmaması durumunda ortaya çıkmaktadır. Veri setinde bulunan kayıplar eksiksiz veri setinin rastgele bir örneklemini olarak düşünülebilir. Kayıp verilerin rastgele ortaya çıkmasından dolayı eksiksiz veri setinden elde edilen kestirimlerin eksiksiz veriden elde edilen kestirimlerden ortalama olarak farklı olmayacağı söylenebilir ancak bu durumda yine de standart hata artar ve güç etkilenir. ROK mekanizmasında ise verilerin kayıp olma durumu diğer gözlenen değişkenlerle ilişkilidir. ROK mekanizmasında kayıp veri ile ilişkili değişkenler kontrol edildikten sonra ilgili değişkende kayıp olma durumu rassal olarak düşünülebilir (Enders, 2013) çünkü kayıp olma durumu bulunduğu değişkenle ilişkili değildir. Aksine, kayıp verinin oluşma olasılığının, veri setinde kayıp verinin bulunduğu değişkene veya başka bir deyişle kayıp verinin gerçek değerine bağlı olması (Scheffer, 2002) ise İEK mekanizmasını göstermektedir.

Kayıp veri ile baş etme yöntemleri geleneksel yöntemler, en çok olabilirlik (Maximum Likelihood) ve çoklu atama yaklaşımlarına dayalı yöntemler olarak sınıflanmaktadır (Allison, 2001). Silme, kukla değişken atama ve ortalama atama gibi geleneksel yöntemlerin yanında beklenti maksimizasyonu ve çoklu atama yöntemleri de

bulunmaktadır. Bu çalışmada ortalama atama (SO), regresyon atama (RA) ve beklenti maksimizasyonu (BM) yöntemleri ele alınmaktadır. Tek değer atama yöntemlerinden yaygın olarak SO yöntemi kullanılmaktadır. Kayıp verilerin bulunduğu değişkene ait ortalama değeri o değişkendeki tüm kayıp veri noktalarına atanmaktadır. Anlaşılması ve uygulanmasının kolay olması ve veri setinde eksik bir gözlem bırakmaması gibi avantajlarının (Enders, 2010) yanı sıra bazı dezavantajları da bulunmaktadır. Tüm kayıp değerleri tek bir gözlem ile tamamlamak standart sapmayı ve varyansı azaltır (Enders, 2010; Little ve Rubin, 2002; Pigott, 2001). RA yönteminde kayıp verilerin tamamlanması için gözlenen değerler kullanılmaktadır. Gözlenen puanlardan elde edilen regresyon denklemi yardımıyla kayıp değerler atanmaktadır (Enders, 2010; Graham, 2012). Ancak bu durum, elde edilen veri setinde korelasyonların olduğundan fazla çıkmasına sebep olabilmektedir. Ayrıca, oluşturulan regresyon denklemi sonucunda belirlenen sınırlar dışında değerler de kayıp verilerin yerine atanabilmektedir (McKnight, McKnight, Sidani ve Figueredo, 2007). Beklenti ve maksimizasyon isimli iki adımdan oluşan BM yönteminde ise B adımında gözlenen değerler kullanılarak bir dizi regresyon denklemi elde edilir ve kayıp verilere atama yapılır. M adımında ise, atama yapılan değerler de kullanılarak yeni değer atamaları gerçekleştirilir. Bu adımlar çok sayıda tekrar edilerek B ve M basamaklarından her adımda yeni değerler hesaplanır. Ardışık M basamakları arasında değerler artık değişmediğinde, maksimum olasılık tahmini yapılmış olur (Enders, 2010). Parametre tahminleri, yineleme adımları arasında değerler artık değişmediğinde tamamlanır. Bu yöntemin dezavantajı olarak standart hataların gerçek standart hatalarla tutarlı olmayabileceği söylenebilir (Allison, 2003).

Amaç

Madde Tepki Kuramı (MTK) modelleri ile bir bireyin örtük özelliği ve maddeye verdiği yanıt arasında bir ilişki kurulmaktadır (Demars, 2010). Bu nedenle bireylerin maddelere verdikleri yanıtların her biri örtük özelliklerinin kestirilmesinde önemli bir rol oynamaktadır. Ancak, veri setinde bulunan kayıplar elde edilen madde parametrelerinde, yetenek parametrelerinde, model-veri uyumunda ve standart hata kestirimlerinde hatalara sebep olabilmektedir. Literatürde madde tepki kuramına ilişkin

kayıp veri çalışmalarının (Demir, 2013; Koçak ve Çokluk Bökeoğlu, 2017) genellikle tek boyutlu yapılarda incelendiği görülmektedir. Ancak, tek boyutlu MTK modellerinin en büyük sınırlılıklarından biri gerçekte çok boyutlu yapıya sahip olan verilerden elde edilen hata miktarlarını olduğundan fazla gösterebilmesi veya modelin veriye uyum sağlayamamasıdır. Ackerman (1996) çoğu psikolojik ve eğitimsel testin değişik derecelerdeki çoklu özelliği veya özellik bileşkesini ölçtüğünü belirtmiştir. Bu nedenle, gerçekte iki veya daha fazla boyut içeren karmaşık verilerin analizi için çok boyutlu MTK modelleri geliştirilmiştir. Çok boyutlu yapılarda gerçekleştirilen çalışmalar genellikle ikili puanlanan maddelere yönelik geliştirilen modeller üzerinde gerçekleştirilmiştir. Çok kategorili puanlanan maddelerin ikili hale dönüştürülerek iki kategorili maddeler için geliştirilmiş MTK modellerinin kullanılmasına ilişkin çalışmalar (Andreis ve Ferrari, 2012) bulunmasına rağmen, bu durumda çok kategorili tepkilerde bazı bilgiler kaybolmaktadır. Özellikle açık uçlu soruların cevaplanmasında ve likert tipi puanlanan psikolojik testlerde çok kategorili maddeler kullanılmaktadır. Dereceli ölçek verilerinde çoklu puanlanan MTK modellerine ilişkin kayıp veri etkisini inceleyen kısıtlı sayıda çalışmanın (Dai ve diğerleri, 2021) bulunması ile de araştırmancının literatüre katkı sağlayacağı düşünülmüştür.

Finch (2008) çalışmasında, üç parametrelili lojistik model için madde parametrelerinin tahmininde kayıp verilerin ele alınmasına yönelik çeşitli yaklaşımların etkisini değerlendirmiştir. Çalışmayı TROK ve ROK mekanizmalarında tek boyutlu MTK modeli çerçevesinde gerçekleştirmiştir. Dai ve diğerleri (2021) ise aşamalı tepki modeli ve genelleştirilmiş kısmi puan modellerinin performansını kayıp verilerin varlığı durumunda incelemiştir. Araştırma ROK mekanizmasına sahip, farklı oranlarında kayıp içeren örneklem büyüklüklerinde ve beş cevap kategorili maddeler kullanılarak ele alınmış ancak yalnızca dizin silme yöntemi ile sınırlı kalmıştır. Andreis ve Ferrari (2012) tarafından ise ÇB-MTK'da TROK, ROK ve İEK mekanizmalarında farklı kayıp veri atama yöntemlerinin performansları incelenmiştir. Ancak ÇB-MTK bağlamında incelenen bu çalışmada çoklu puanlanan maddeler ikili puanlanan maddelere dönüştürülerek iki parametrelili lojistik modelin çok boyutlu uzantısı kullanılmıştır.

Kayıp verilerle baş etme yöntemlerinin etkilerinin farklı koşullar altında birbirlerine göre farklılaşabildiği göz önünde bulundurularak literatürdeki çalışmaların bir adım ilerisi olarak, bu çalışmada çok kategorili maddeler için geliştirilmiş Çok Boyutlu Madde Tepki Kuramı (ÇB-MTK) modellerinde kayıp veri sorununun incelenmesi amaçlanmaktadır. Kayıp veriler ile başa çıkmada sıklıkla kullanılan SO, RA ve BM yöntemleri TROK koşulu altında değerlendirilmiştir. İlgili yöntemler kullanılması kolay ve yaygın olduğu için tercih edilmiştir. Farklı koşullar altında yöntemlerin etkililiği belirlenerek çok kategorili puanlanan ÇB-MTK çalışacak araştırmacılara kayıp veri yöntemlerinin hangisinin kullanılmasının daha etkili olacağı konusunda yol göstereceği düşünülmektedir.

YÖNTEM

Araştırmada kullanılan referans veri seti ÇBMTK yapısı, boyut sayısı, boyutlar arası korelasyon, örneklem büyüklüğü, boyut başına düşen madde sayısı ve maddelere ilişkin puan kategorileri sayısı göz önünde bulundurularak R programında “mirt” paketi (Chalmers, 2012) ile üretilmiştir. ÇBMTK modellerinden Çok Boyutlu Aşamalı Tepki Modeline (ÇBATM) uygun olarak veri üretimi gerçekleştirilmiştir. İlk boyut ile ikinci boyut arasında yüksek düzeyde korelasyon bulunması durumunda tek boyutlu modellerin değerlendirilebileceği (Zhang, 2008) göz önünde bulundurularak veri üretimi aşamasında boyutlar arasında 0,50 korelasyon olacak şekilde iki boyutlu, basit yapılu bir model kullanılmıştır. Boyut başına 5 madde ve her maddenin 0-3 arasında puanlanan dört puan kategorisi oluşturulmuştur. Çoklu puanlanan MTK modellerinde örneklem büyüklüğünün 250 ve üzerinde olmasının madde ve yetenek parametrelerinin geri kazanımında yeterli olduğu söylenmektedir (Valdivia ve Dai, 2023). Bu çalışmada, iki boyutlu bir ÇB-MTK yapısı incelendiğinden dolayı örneklem büyüklüğü 1000 olarak belirlenmiştir.

İlk olarak, boyutlar arası korelasyon ($r=0,5$) olacak şekilde çok boyutlu normal dağılımdan 1000 birimlik yetenek parametreleri tanımlanmıştır. Madde parametreleri için ise Jiang, Wang ve Weiss (2016) tarafından önerilen parametreler kullanılmıştır.

Buna göre a_1 ve a_2 parametreleri minimum 1.1, maksimum 2.8 olacak şekilde tek biçimli (uniform) dağılımdan elde edilmiştir.

$$a_1, a_2 \sim U(1.1, 2.8)$$

b_1, b_2 ve b_3 parametreleri ise (-2,2) aralığında tek biçimli dağılım ile elde edilmiştir.

$$b_1 \sim U(-2, -0.67)$$

$$b_2 \sim U(-0.67, 0.67)$$

$$b_3 \sim U(0.67, 2)$$

Madde ve yetenek parametreleri belirlenen ölçütler bağlamında oluşturulduktan sonra ÇB-AT modeline göre madde yanıtları üretilmiştir. Daha sonra elde edilen veri seti kullanılarak, yöntemlerden elde edilen parametreler ile kıyaslanması amacıyla kullanılacak olan referans madde ve yetenek parametreleri kestirilmiştir.

Harwell, Stone, Hsu ve Kirisci (1996) MTK çalışmalarında en az 25 replikasyon önermelerine karşın MTK'da yöntem karşılaştırmalarında az sayıda replikasyonun yeterli olabileceğini belirtmişlerdir. Çalışmada replikasyon, üretilen parametrelerden kayıp veri içeren veri setlerinin elde edilmesi, elde edilen eksik veri setlerinin tamamlanması ve ÇB-AT modelinden madde ve yetenek parametrelerin elde edilmesi aşamasında uygulanmıştır. Çalışmanın karmaşık yapısı göz önünde bulundurularak replikasyon sayısı 10 olarak belirlenmiştir. Replikasyonlardan elde edilen değerlerin ortalaması kullanılarak sonuçlar yorumlanmıştır.

Kayıpların Elde Edilmesi Ve Tamamlanması

Üretilen veri setinden TROK mekanizmasına sahip olacak şekilde %5, %10 ve %20 kayıp içeren veri setleri elde edilmiştir. Kayıp verilerin elde edilmesinde R'da missMethods (Rockel, 2022) paketi kullanılmıştır.

Elde edilen veri setleri SPSS paket programında SO, RA ve BM yöntemleri ile tamamlanmıştır. Elde edilen ondalık değerler en yakın tamsayıya yuvarlanarak analizlere devam edilmiştir.

Kayıp Veri Yöntemlerinin Değerlendirilme Kriteri

Farklı kayıp veri yöntemleri ile tamamlanan veri setlerinden elde edilen madde ve yetenek parametreleri, üretilen eksiksiz veri setinden elde edilen parametreler ile karşılaştırılmıştır. Bu karşılaştırma için hataların ortalama karekökü (RMSE) değerleri kullanılmıştır.

$$RMSE = \sqrt{\frac{\sum(\hat{y}_i - y)^2}{N}}$$

RMSE, kestirilen değer ile gerçek değer arasındaki farkı ifade eden bir hata ölçüsüdür. Düşük RMSE değerleri kestirimlerin gerçek değere daha yakın olduğunu göstermektedir. Yöntemlerin tamamlanmasıyla elde edilen değerlerin karşılaştırılacağı eksiksiz veri setinden elde edilen madde parametrelerine ilişkin bilgiler Tablo 1’de verilmiştir.

Tablo 1. Eksiksiz Veri Setinden Elde Edilen Madde Parametreleri

Madde	Parametre Kestirimleri				
	a ₁	a ₂	d ₁	d ₂	d ₃
1	2.974	0	2.063	0.235	-1.407
2	1.435	0	1.149	-0.231	-1.680
3	1.948	0	1.513	0.076	-0.733
4	1.838	0	1.220	-0.347	-0.703
5	2.328	0	2.289	0.561	-1.323
6	0	2.463	1.841	-0.248	-1.869
7	0	1.428	1.105	-0.711	-1.691
8	0	0.970	1.784	-0.535	-1.732
9	0	2.631	0.955	0.581	-1.809
10	0	1.171	0.872	-0.416	-1.133

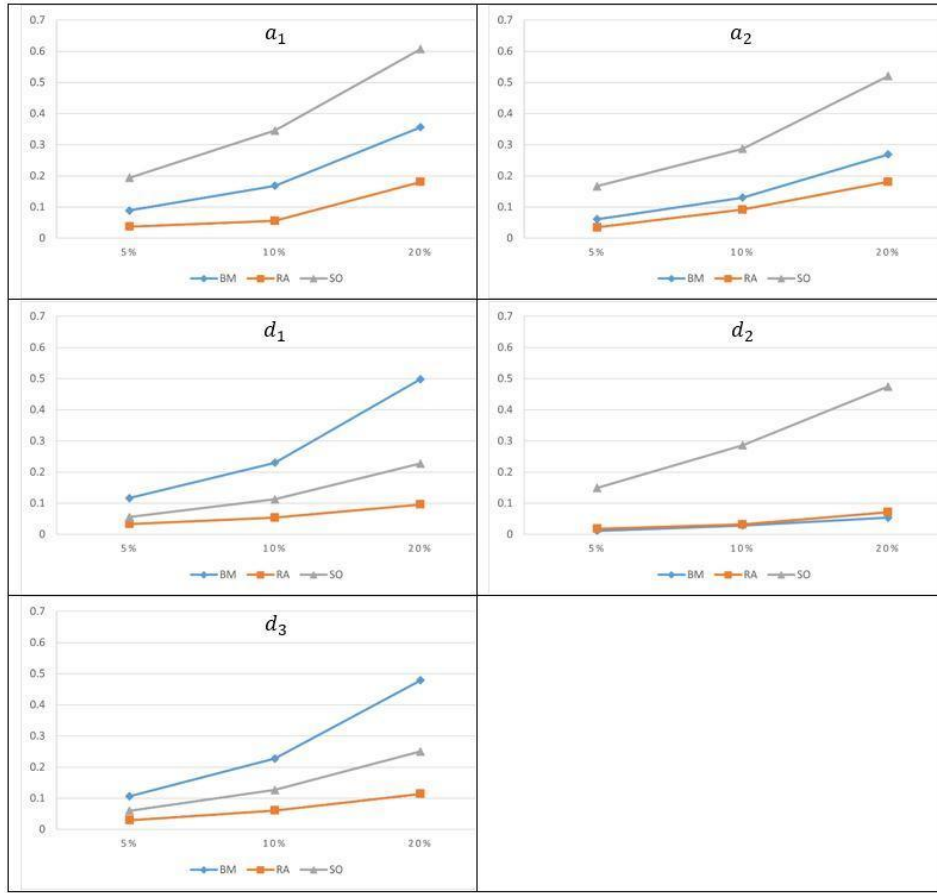
Tek boyutlu AT modelinde kategori eřik deđerleri artan řekilde sıralanmasına karřın B-AT modelinde kategori keřiřim parametreleri (d_1, d_2, d_3) iin bu durum ters řekilde gerekleřmektedir. ünkü B-AT modelinde pozitif keřiřim formu kullanılmaktadır (Paek & Cole, 2020). Bu nedenle alıřmada ayırıcılık parametreleri ile birlikte kategori keřiřim parametreleri kullanılmıřtır.

BULGULAR

Arařtırmadan elde edilen madde ve yetenek parametrelerine iliřkin bulgular bu blmde yer almaktadır.

Madde Parametrelerine İliřkin Bulgular

B-AT modelinden elde edilen a_1 ve a_2 ayırıcılık parametresi ile her bir kategoriye iliřkin d_1, d_2 ve d_3 keřiřim parametrelerine iliřkin ortalama RMSE deđerleri elde edilmiřtir. Farklı kayıp veri oranlarında kayıp veri yntemlerinden elde edilen ortalama RMSE deđerleri řekil 1’de verilmiřtir.



Şekil 1.: Madde Parametrelerine İlişkin Ortalama RMSE Değerleri

Şekil 1 incelendiğinde, veri setinde bulunan kayıp veri miktarı arttığında tüm parametrelerde RMSE'nin de arttığı görülmektedir. Bu durum kayıp gözlem miktarının artması ile kaybedilen bilginin de artmasına bağlı olarak gerçekleşmektedir (Kaya Kalkan, Kara ve Kelecioğlu, 2018). Böylelikle kestirilen parametrelerdeki hata miktarı da artmaktadır. %5 kayıp veri durumunda, parametre kestirimlerinin hata miktarları arasındaki farkın en az düzeyde olduğu görülmektedir.

SO yönteminin a_1 ve a_2 parametrelerinde en yüksek düzeyde ortalama RMSE'ye sahip olduğu görülmüştür. Madde ayırıcılık parametreleri için tüm kayıp veri oranlarında en

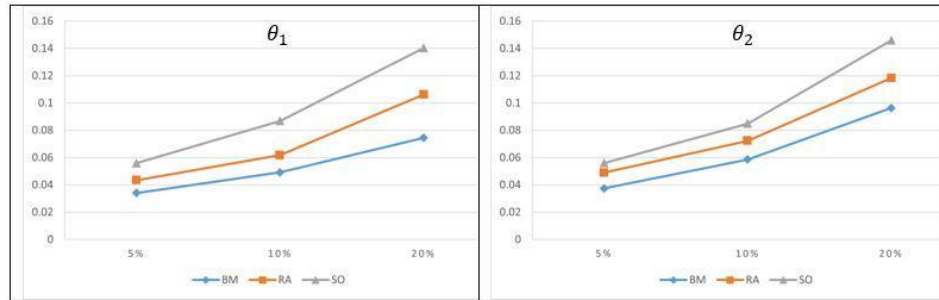
iyi performans gösteren yöntem RA yöntemi olmuştur. Bu yöntemde özellikle %5 ve %10 kayıp veri durumlarında düşük RMSE değerleri elde edilmiştir.

Kategori sınır kesişim parametreleri incelendiğinde ise d_1 ve d_3 parametresinde yöntemlerin benzer sonuçlar ürettiği görülmektedir. d_1 ve d_3 parametreleri için en küçük RMSE üreten yöntem ayırıcılık parametrelerinde olduğu gibi RA olmuştur. d_2 parametresinde ise diğer parametrelerden farklı olarak oldukça düşük RMSE değerleri elde edilmiştir. BM ve RA yöntemlerinin neredeyse aynı ve oldukça düşük sonuçlar ürettiği ancak SO yönteminin diğer yöntemlere kıyasla oldukça fazla RMSE ürettiği görülmüştür.

Madde parametreleri için tüm kayıp oranlarında ve tüm parametrelerde en az hata ile kestirim yapan yöntemin RA olduğu söylenebilir.

Yetenek Parametresine İlişkin Bulgular

ÇB-AT modelinden elde edilen yetenek parametrelerine ilişkin ortalama RMSE değerleri elde edilmiş ve farklı kayıp veri oranlarında yöntemlerden elde edilen ortalama RMSE değerleri Şekil 2’de verilmiştir.



Şekil 2. Yetenek Parametrelerine İlişkin Ortalama RMSE Değerleri

Şekil 2 incelendiğinde, θ_1 ve θ_2 yetenek parametreleri için tüm yöntemlerde düşük RMSE değerleri elde edildiği görülmektedir. Veri setindeki kayıp oranı arttığında tüm yöntemler için de RMSE değerinin artış gösterdiği görülmektedir. Kayıp bilginin her iki yetenek parametresinde de oran arttıkça arttığı söylenebilir.

Her iki yetenek parametresi için de kayıp veri yöntemlerinin benzer şekilde çalıştığı ve en düşük RMSE değerlerini üreten yöntemin BM olduğu görülmüştür. SO yöntemi için ise yetenek parametrelerinde diğer yöntemlere kıyasla daha az etkili sonuçlar elde edildiği söylenebilir.

TARTIŞMA VE SONUÇ

Çalışmada, basit yapılı çok boyutlu yapılar için geliştirilmiş olan ÇB-MTK modellerinden ÇB-AT modelinde kayıp veri ile baş etme yöntemlerinin madde ve yetenek parametrelerine etkisi araştırılmaktadır. İki boyutlu ve 0-3 şeklinde çoklu puanlanan veri setinde TROK mekanizmasına sahip olacak şekilde %5, %10 ve %20 oranında kayıplar oluşturulmuş ve SO, RA, BM yöntemleri ile tamamlanmıştır. Tamamlanan veri setinden elde edilen parametreler eksiksiz veri setinden elde edilen referans parametreler ile karşılaştırılmış ve elde edilen bulgular yorumlanmıştır.

Kalkan, Kara ve Kelecioğlu'nun (2018) ikili puanlanan MTK modelinden elde ettikleri sonuçlara benzer olarak, tüm madde ve yetenek parametreleri için %5 kayıp içeren veri setlerinde RMSE değerlerinin diğer kayıp oranlarından daha düşük olduğu görülmüştür. Bu durum, TROK mekanizmasına sahip %5 ve daha altında kayıp içeren veri setlerinde kayıp verilerin etkisinin az olacağını göstermektedir. Dai ve diğerlerinin (2021) çoklu puanlanan MTK modellerinde kayıp veri etkisini inceledikleri çalışmalarında da buna benzer sonuçlar ortaya çıkmıştır. Kayıp veri oranının %10 veya daha az olduğu durumlarda çoklu puanlanan MTK modellerinde kayıp veri etkisinin kabul edilebilir düzeyde olduğunu belirtmişlerdir. Ancak, yüksek oranlarda kayıp içeren veri setlerinde kayıp verilerin uygun yöntemlerle başa çıkılması gerektiği önerilmiştir.

Madde ayırıcılık parametreleri için tüm kayıp oranlarında RA ve BM yöntemlerinin SO yöntemine göre daha iyi performans sergilediği görülmüştür. Kategori sınır kesişim parametrelerinde ise d_1 ve d_3 için RA ve SO yöntemleri en düşük RMSE değerini üretirken d_2 parametresi için ise BM ve RA'nın tüm kayıp oranlarında oldukça düşük sonuçlar ürettiği görülmüştür. SO yönteminin d_2 parametresinde çok yüksek hata oranları ürettiğinden bu yöntemin madde parametrelerinin geri kazanımında dikkatli

şekilde kullanılması önerilir. Genel olarak tüm madde parametreleri değerlendirildiğinde RA yönteminin parametre değerlerinin geri kazanımında diğer yöntemlere kıyasla oldukça iyi sonuçlar elde ettiği söylenebilir. Çoklu puanlanan ölçekler üzerinde kayıp veri etkisinin incelendiği çalışmalarda RA yönteminin sonuçları yeniden üretmede oldukça iyi olduğu sonucuna ulaşan çalışmalar (Roth, Switzer ve Switzer, 1999) bulunmaktadır. MTK modelleri bağlamında ise RA yönteminin ROK mekanizmasında model veri uyumu üzerinde iyi sonuçlar ürettiği görülmektedir (Koçak ve Çokluk Bökeođlu, 2017). Bernaards ve Sijtsma (2010) ise çoklu puanlanan çok boyutlu madde tepki verilerinde kayıp veri yöntemlerinin performansını faktör analizi bağlamında incelemiştir. BM yönteminin diğer yöntemlere kıyasla daha iyi sonuçlar ürettiğini ve BM yöntemine alternatif olarak SO yönteminin de kullanılabileceğini belirtmişlerdir.

Yetenek parametrelerinin geri kazanımı için ise tüm kayıp oranlarında en düşük RMSE değerleri üreten yöntemin BM olduğu görülmüştür. Dolayısıyla yetenek parametrelerinin değerlendirilmesi için gerçekleştirilecek çalışmalarda BM yönteminin kullanılması daha uygun olacaktır. Bu sonuç, Bernaards ve Sijtsma'nın (1999) çok boyutlu ölçeklerde kayıp veri yöntemlerinin faktör puanlarına etkisini inceledikleri çalışmasında BM algoritmasının, faktör puanlarını tahmin etmek için kullanılabileceğini belirttikleri sonuç ile tutarlıdır.

KAYNAKLAR

- Ackerman, T. A. (1996). Graphical representation of multidimensional item response theory analyses. *Applied Psychological Measurement*, 20, 311-329. doi: 10.1177/014662169602000402
- Andreis, F., & Ferrari, P. A. (2012). Missing data and parameters estimates in multidimensional item response models. *Electronic Journal of Applied Statistical Analysis*, 5(3), 431-437. doi: 10.1285/i20705948v5n3p431
- Allison, P. D. (2001). *Missing data*. Thousand Oaks, CA: Sage.
- Allison, P. D. (2003). Missing data techniques for structural equation modeling. *Journal of Abnormal Psychology*, 112(4), 545-557. doi: 10.1037/0021-843X.112.4.545
- Bernaards, C., & Sijtsma, K. (1999). Factor Analysis of Multidimensional Polytomous Item Response Data Suffering From Ignorable Item Nonresponse. *Multivariate Behavioral Research*, 34(3), 277-313. doi: 10.1207/S15327906MBR3403_1
- Bernaards, C., & Sijtsma, K. (2010). Influence of Imputation and EM Methods on Factor Analysis When Item Nonresponse in Questionnaire Data Is Nonignorable. *Multivariate Behavioral Research*, 5(3), 321-364. doi:10.1207/S15327906MBR3503_03
- Chalmers, R. P. (2012). mirt: A Multidimensional Item Response Theory Package for the R Environment. *Journal of Statistical Software*, 48(6), 1-29. doi: 10.18637/jss.v048.i06
- Dai, S., Vo, T. T., Kehinde, O. J., He, H., Xue, Y., Demir, C., & Wang, X. (2021). Performance of Polytomous IRT Models With Rating Scale Data: An Investigation Over Sample Size, Instrument Length, and Missing Data. *Front. Educ.*, 6. doi: 10.3389/educ.2021.721963
- Demars, C. (2010). *Item Response Theory: Understanding Statistics Measurement*. Oxford University Press, Oxford.
- Demir, E. (2013). *Kayıp Verilerin Varlığında İki Kategorili Puanlanan Maddelerden Oluşan Testlerin Psikometrik Özelliklerinin İncelenmesi* (Doktora tezi). Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Enders, C. K. (2010). *Applied missing data analysis*. New York: The Guilford Press.
- Enders, C. K. (2013). Dealing With Missing Data in Developmental Research. *Child Development Perspectives*, 7(1), 27-31. doi: 10.1111/cdep.12008
- Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London: Sage.
- Finch, H. (2008). Estimation of item response theory parameters in the presence of missing data. *Journal of Educational Measurement*, 45(3), 225-245. doi: 10.1111/j.1745-3984.2008.00062.x

- Goegebeur, Y., De Boeck, P. & Molenberghs, G. (2010). Person fit for test speededness: normal curvatures, likelihood ratio tests and empirical bayes estimates. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 6(1), 3–16. doi: 10.1027/1614-2241/a000002
- Graham, J. W. (2012). *Missing data: analysis and design*. New York: Springer
- Harwell, M., Stone, C. A., Hsu, T. C., & Kirisci, L. (1996). Monte Carlo studies in item response theory. *Applied psychological measurement*, 20(2), 101-125. doi: 10.1177/014662169602000
- Jiang, S., Wang, C., & Weiss, D. J. (2016). Sample Size Requirements for Estimation of Item Parameters in the Multidimensional Graded Response Model. *Front. Psychol.*, 7, doi: 10.3389/fpsyg.2016.00109
- Kaya-Kalkan, Ö., Kara, Y. & Kelecioğlu, H. (2018). Evaluating Performance of Missing Data Imputation Methods in IRT Analyses. *International Journal of Assessment Tools in Education*, 5(3), 403–416. doi: 10.21449/ijate.430720
- Koçak, D. ve Çokluk-Bökeoğlu, Ö. (2017). Kayıp veriyle baş etme yöntemlerinin model veri uyumu ve madde model uyumuna etkisi. *Journal of Measurement and Evaluation in Education and Psychology*, 8(2), 200-223. doi: 10.21031/epod.303753
- Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data*. (2nd edition). New York: Wiley
- Little, R. J. A., & Schenker, N. (1995). Missing data. In Arminger, G., Clogg, C. C., & Sobel, M. E. (Eds), *Handbook of statistical modeling for the social and behavioral sciences*. Boston, MA: Springer. doi: 10.1007/978-1-4899-1292-3_2
- McKnight, P. E., McKnight, K. M., Sidani, S. & Figueredo, A. J. (2007). *Missing Data: A Gentle Introduction*. New York: The Guilford Publications, Inc.
- Paek, I., & Cole, K. (2019). *Using R for item response theory model applications*. London: Routledge.
- Pigott, T. D. (2001). A review of methods for missing data. *Educational Research and Evaluation: An International Journal of Theory and Practice*, 7(4), 353-383. doi: 10.1076/edre.7.4.353.8937
- Rockel, T. (2022). *missMethods: Methods for missing data*. R package version 0.4.0.
- Roth, P. L., Switzer, F. S., & Switzer, D. M. (1999). Missing Data in Multiple Item Scales: A Monte Carlo Analysis of Missing Data Techniques. *Organizational Research Methods*, 2(3), 211-232. doi: 10.1177/109442819923001
- Scheffer, J. (2002). Dealing with missing data. *Research Letters in the Information and Mathematical Sciences*, 3, 153-160.
- Tabachnick, B. G., & Fidell, L. S. (2012). *Using Multivariate Statistics* (6th Edition). Boston: Person Education.

- Zhang, B. (2008). Application of unidimensional item response models to tests with items sensitive to secondary dimensions. *The Journal of Experimental Education*, 77 (2), 147-166. doi: 10.3200/JEXE.77.2.147-166
- Valdivia, D. S., & Dai, S. (2023). Number of Response Categories and Sample Size Requirements in Polytomous IRT Models. *The Journal of Experimental Education*, doi: 10.1080/00220973.2022.2153783

SUMMARY

In this study, the effect of missing data methods on item and ability parameters obtained from the multidimensional graded response model, which is one of the multidimensional item response theory models with polytomous items, is investigated.

Method

In the study, data generation was carried out in accordance with the Multidimensional Graded Response Model (M-GRM). Two dimensions, five items per dimension and four response categories where each item is scored between 0-3 were formed. The sample size was determined as 1000. Firstly, ability parameters were defined from a multidimensional normal distribution with an inter-dimensional correlation of $r=0.5$. For item parameters, the parameters suggested by Jiang, Wang, and Weiss (2016) were used. Discrimination parameters were obtained from a uniform distribution with a minimum of 1.1 and a maximum of 2.8. Category boundary intercept parameters were obtained from a uniform distribution in the range of (-2,2). Item responses were generated using item and ability parameters. Reference item and ability parameters were estimated from the generated data set. The number of replications was set as 10. From the complete data set, data sets containing 5%, 10% and 20% missingness with MCAR mechanism were obtained. The obtained data sets were completed with SMI, RI and EM methods. The item and ability parameters obtained from the data sets completed with different methods were compared with the reference parameters. RMSE were used for this comparison.

Results

Average RMSE values were obtained for $a1$ and $a2$ discrimination parameters and $d1$, $d2$ and $d3$ intercept parameters for each category. It is seen that RMSE increases in all parameters when the amount of missing data increases. In the case of 5% missing data, the difference between the error amounts of the parameter estimates is minimal. SMI method has the highest average RMSE for $a1$ and $a2$. For the item discrimination parameters, RI was the best performing method in all ratios. In this method, low RMSE values were obtained especially in 5% and 10% missingness. When the category boundary intersection parameters are analysed, the methods have similar results for $d1$ and $d3$ parameters. As in the discrimination parameters, the method which had the smallest RMSE for $d1$ and $d3$ was RI. For parameter $d2$, unlike the other parameters, very low RMSE values were obtained. EM and RI methods produced almost the same and quite low results, but SMI method produced higher RMSE. For the item parameters, it can be said that RI is the method that predicts with the least error at all missing rates and for all parameters.

When the average RMSE values for the ability parameters are analysed, low RMSEs were obtained from all methods for the ability parameters. As the missing rate increases, the RMSE increases for all methods. It was observed that the methods worked similarly for both ability parameters and the method that produced the lowest RMSE was BM. For the SO method, it can be said that less effective results are obtained for the ability parameters compared to the other methods.


Discussion and Conclusion

For all parameters, RMSEs are quite low in data sets with 5% missingness. This indicates that the effect of missing data will be less in data sets with 5% or less missing with MCAR mechanism. Similar results were obtained in the study of Dai et al.(2021).

It was observed that RI and EM methods performed better than SMI method at all missing rates for item discrimination parameters. For the category boundary intersection parameters, RI and SMI methods produced the lowest RMSE values for d1 and d3, while for the d2 parameter, EM and RI produced very low results at all missing rates. Since the SMI method produces very high error rates for the d2 parameter, it is recommended that this method be used with caution in the recovery of item parameters. In general, when all item parameters are evaluated, it can be said that the RI method has very good results in the recovery of parameters. For the recovery of ability parameters, it was observed that the EM method produced the lowest RMSE values at all missing rates. Therefore, it would be more appropriate to use the EM method in the studies to be carried out for the evaluation of ability parameters.

ORCID

Mehmet Ali İşikoğlu  ORCID 0000-0001-5104-5661

Burcu Atar  ORCID 0000-0003-3527-686X

Araştırmacıların Katkı Oranı Beyanı

Bu çalışmanın planlanması, yürütülmesi ve yazılı hale getirilmesinde araştırmacılar eşit oranda katkı sağlamıştır.

Destek ve Teşekkür Beyanı

Bu araştırmada herhangi bir kurum, kuruluş ya da kişiden destek alınmamıştır.

Çatışma Beyanı

Araştırmacıların, araştırma ile ilgili diğer kişi ve kurumlarla herhangi bir kişisel ve finansal çıkar çatışması yoktur.

Etik Kurul Beyanı

Bu araştırma, simülatif verilere dayalı bir çalışma olduğu için etik kurul izni gerektirmemektedir.