

# Kayıp Veriyle Baş Etme Yöntemlerinin Model Veri Uyumu Ve Madde Model Uyumuna Etkisi\*

## The Effect Of Missing Data Tecniques On Model Fit And Item Model Fit

Duygu KOÇAK\*\* Ömay ÇOKLUK BÖKEOĞLU\*\*\*

### Öz

Bu çalışmanın amacı, kayıp veri baş etme yöntemlerinin Madde Tepki Kuramı 1 parametrelili lojistik modelinde model veri uyumuna ve madde model uyumuna olan etkisini incelemektir. Bu amaç doğrultusunda örneklem büyüklüğünün 500, 1000 ve 1500 olarak manipüle edildiği, madde sayısının 20 olarak sabitlendiği IPLM'e uyumlu veri setleri üretilmiştir. Üretilen verilerde madde güçlüğü -2 ile +2 arasında sınırlandırılmış, madde ayırt ediciliği 1.5 olarak sabitlenmiştir. Üretilen eksiksiz veri setleri üzerinde tamamen rastgele kayıp ve rastgele kayıp koşulları altında %5, %10 ve %15 oranlarında silme işlemi gerçekleştirilmiştir. Tamamen rastgele kayıp mekanizması, veri setindeki toplam hücre sayısı arasından rastgele değerler silinerek oluşturulmuştur. Liste bazında silme yöntemi için belirlenen oranda rastgele birim (kişi) silinmiştir. Rastgele kayıp mekanizması, veri setine üç düzeyi olan bir değişken tanımlanıp 1. düzeyden %20, 2. düzeyden %30 ve 3. düzeyden %50 oranında olacak şekilde hücrelerin rastgele silinmesi ile oluşturulmuştur. Oluşturulan kayıp veriler liste bazında silme, regresyonla atama ve beklenti maksimizasyon algoritması yöntemleriyle giderilmiştir. Model veri uyumunun kestirilmesinde  $-2 \log \lambda$ , AIC ve BIC değerlendirme kriterlerinden, madde model uyumunun kestirilmesinde  $\chi^2$  istatistiğinden faydalanılmıştır. Eksiksiz veri setlerinden elde edilen değerler, kayıp veri baş etme yöntemleriyle tamamlanan veri setlerinden elde edilen kestirimler için referans olarak kullanılmıştır. İncelemeler sonucunda, beklenti maksimizasyon algoritması yönteminin rastgele kayıp mekanizmasında iyi, tamamen rastgele kayıp mekanizmasında kısmen iyi performans sergilediği sonucuna ulaşılmıştır. Regresyonla atama yönteminin de belirli koşullar altında iyi performans sergilediği ancak liste bazında silme yönteminin performansının düşük olduğu sonucuna ulaşılmıştır. Tüm kayıp veri mekanizmalarında kayıp veri oranı arttıkça, kayıp veri baş etme yöntemlerinin performansı da düşmektedir. Tüm mekanizmalarda ve koşullarda kullanılabilecek, en iyi sonuçları veren tek bir yöntemin varlığından söz etmenin mümkün olmadığı belirlenmiştir.

*Anahtar Kelimeler:* kayıp veri, kayıp veri baş etme yöntemleri, model veri uyumu, madde model uyumu, madde tepki kuramı.

### Abstract

The purpose of this study was to examine the effects of missing data handling techniques on model data fit and item model fit in the one parameter logistic Item Response Theory Model. For this purpose, data sets with sample sizes of 500, 1000, and 1500 and with 20 items that fit to one parameter logistic model were generated. Item difficulty values of the items in the generated data sets ranged from -2 to +2 and item discrimination was fixed as 1.5. The generated complete data sets were exposed to deletion at %5, %10, and %15 under missing at complete random (MCR) and missing at random (MR) conditions. Missing at complete random mechanism was obtained as a result of random values deleted among the total number of cells in the data set. A particular percentage of random units (individuals) were deleted for listwise deletion method. Missing at random mechanism was reached as a result of random deletion of cells pursuant to defining a three level variable in the data set at the following percentages: 20% at Level 1, 30% at Level 2 and 50% at Level 3. The generated missing data were resolved using listwise deletion method (LM), regression imputation, and expectation maximization algorithm (EMA).  $-2 \log \lambda$ , AIC, and BIC evaluation criteria were used for model data fit

\*Bu çalışma, birinci yazarın Doç. Dr. Ömay ÇOKLUK BÖKEOĞLU danışmanlığında tamamlanan doktora tezinden türetilmiştir.

\*\*Yrd. Doç. Dr., Alanya Alaaddin Keykubat Üniversitesi, Eğitim Fakültesi, Antalya-Türkiye, e-posta: duygu.kocak@alanya.edu.tr

\*\*\*Doç. Dr., Ankara Üniversitesi, Eğitim Bilimleri Fakültesi, Ankara-Türkiye, e-posta: cokluk@aducation.ankara.edu.tr

estimation and  $\chi^2$  statistics were used for item model fit estimation. Values obtained from the complete data sets were taken as reference for predictions in the data sets that were completed with the effect of missing data handling techniques. As a result of the examinations, it was concluded that expectation maximization algorithm had good performance in missing at random mechanism but partially good in missing at complete random mechanism. It was also seen that regression imputation had good performance under certain conditions but the performance of listwise deletion method was poor. In all missing data mechanisms, the performance of the effect of missing data handling techniques declines as missing data increase. It is certain that a single method to give best results in all mechanisms and under any conditions is unlikely to be assumed.

*Keywords:* missing data, missing data techniques, model fit, item model fit, item response theory.

## GİRİŞ

Eğitim ve psikoloji alanında ölçme konusu olan özellikler çoğunlukla doğrudan gözlenemeyen, gizil (örtük) değişkenlerdir. Bireylerin ölçülmek istenen bu değişkenlerle ilgili olduğu düşünülen ve gözlenebilen değişkenlere verdikleri tepkiler aracılığıyla, gizil özelliklere dair çıkarımlar yapılmaktadır. Bireylerin gizil özelliklerini gözlenen değişkenler ile açıklamak için farklı test kuramları geliştirilmiştir ve söz konusu özellikler, bu kuramlara dayalı olarak geliştirilen ölçme araçları ile ölçülmektedir.

Psikolojik ölçme tarihindeki kuramlar ve bunların gelişimleri incelendiğinde test geliştirmede, geliştirilen testlerin puanlanmasında ve analizinde yaygın olarak Klasik Test Kuramı'nın (KTK) kullanıldığı görülmektedir. KTK'nın çeşitli sınırlılıkları nedeniyle Madde Tepki Kuramı (MTK), Genellenabilirlik Kuramı (GK), Konjenerik Test Kuramı (KonTK) gibi farklı kuramlar da giderek daha fazla tercih edilmeye ve kullanılmaya başlanmıştır (Reise, Ainsworth ve Haviland, 2005).

Klasik Test Kuramı'nda doğrudan gözlenemeyen iki farklı puan tanımlanmaktadır. Bunlar, gerçek puan ve hata puanıdır. Ölçülmesi hedeflenen gizil özelliğe ilişkin gerçek değeri ifade eden gerçek puan tanımlanır ve temel olarak ölçülen özelliğin gerçek değerine ulaşmak amaçlanır. Fakat ölçme sürecini etkileyen hatalar nedeniyle bu değer doğrudan kestirilememektedir. Bu nedenle kuramda gerçek puan, çeşitli varsayımlar ışığında gözlenen puanlar aracılığıyla kestirilir. Kuramda gerçek ile gözlenen puan farkı hata puanı olarak tanımlanır (Baykul, 2000; Crocker ve Algina, 1986; Embretson, 1999; Lord, 1980; Turgut, 1992; Wainer ve Thissen, 2001; Wilson, 2005).

KTK'da parametreler madde ve test olmak üzere iki başlıkta ele alınmaktadır. Madde güçlük indeksi, madde ayırt edicilik indeksi, madde standart sapması, madde varyansı ve madde güvenilirliği madde parametreleri olarak gruplandırılır (Cronbach, 1990). KTK'da madde güçlüğü ve madde ayırt ediciliği gibi parametreler verilerin toplandığı gruptan gruba farklılık göstermektedir bir diğer ifadeyle gruba bağlı olarak farklı değerler alabilmektedir (Lord ve Novick, 1968). Uygulama yapılan grup değiştiğinde, madde parametrelerinin de değişmesi, farklı gruplara ilişkin test puanlarının karşılaştırılmasında önemli bir sınırlılıktır. Bu durum KTK'da farklı testleri alan kişilerin karşılaştırılmasını güçleştirmektedir (Hambleton, Swaminathan ve Rogers, 1991). Bir diğer önemli sınırlılık ise testin bütünü için tek bir hata değerinin kestirilmesi ve testi alan bireylerin tümü için hata varyansının eşit kabul edilmesidir (Hambleton, Swaminathan ve Rogers 1991). Madde parametrelerinde olduğu gibi güvenilirlik gibi test parametreleri de uygulama yapılan gruba bağlı olarak farklı değerler alabilmektedir. Dolayısıyla bir grupta güvenilir olan test, başka bir grupta aynı durumu gösteremeyebilir yani güvenilir olmayabilir (Crocker ve Algina, 1986). Diğer bir sınırlılığı da testi alan kişilerin testte yer alan bir maddeye nasıl tepki vereceğine ilişkin kestirim imkânı sunamamasıdır (Hambleton ve Swaminathan, 1989).

Klasik Test Kuramı'nın yukarıda değinilen temel bazı sınırlılıklarına alternatif olarak geliştirilen önemli kuramlardan biri de MTK'dır. Kuram testi alan kişinin yetenek düzeyi ile maddelere verdiği yanıtlar arasındaki ilişkinin matematiksel fonksiyonlar ile açıklanabileceği iddiasındadır (Embretson ve Reise, 2000; Hambleton ve Swaminathan, 1989). Buna bağlı olarak da bireylerin yetenek düzeyleri testten bağımsız olarak maddelere verilen yanıtlardan hareketle kestirilmektedir. Çünkü MTK, farklı yetenek düzeylerindeki bireylerin her bir maddede nasıl tepki vereceklerine dair matematiksel modelleme sunmaktadır (Crocker ve Algina, 1986).

Madde Tepki Kuramı, test ve madde parametrelerinin gruptan, yetenek düzeyi kestiriminin ise testten bağımsız olduğu iddiasındadır. Bununla birlikte büyük örneklem aracılığıyla gruptan bağımsız şekilde kestirilen madde parametrelerinin değişmezlik özelliği ile farklı gruplar için ölçme sonuçlarının karşılaştırılmasına olanak sağlamaktadır. MTK'nın KTK'dan bir diğer üstünlüğü ise, KTK'da tek bir güvenilirlik katsayısı hesaplanırken MTK'da her bir yetenek düzeyi için test ve madde bilgi fonksiyonları aracılığıyla ayrı standart hata kestirimi yapılabilmesidir (Adams, 2005; Çelen, 2008; Embretson ve Reise, 2000; Yurdugül, 2006).

Madde Tepki Kuramı'nda yetenek ( $\theta$ ), maddelere verilen yanıtlar arasındaki kovaryansa karşılık gelen kesiksiz ve başat yapı yani özellik olarak ifade edilmektedir (Reeve, 2002).  $\theta$ , lojit birimi ile ifade edilir ve kuramsal olarak  $-\infty$  ile  $+\infty$  arasında değer alır. Madde puanının,  $\theta$  yetenek düzeyi üzerindeki regresyonu "Madde Karakteristik Fonksiyonu (MKF)", fonksiyonun grafiği ise "Madde Karakteristik Eğrisi (MKE)" olarak adlandırılır (Embretson ve Reise, 2000; Hambleton ve Swaminathan, 1989; Lord ve Novick, 1968). Madde Karakteristik Eğrisi'nin şeklini ise, maddeye ait üç parametre (güçlük, ayırt edicilik ve şans) belirlemektedir.

Madde Tepki Kuramı'nda testte yer alan her bir maddenin taşıdığı bilgi "Madde Bilgi Fonksiyonu (MBF)" ile hesaplanır (Yu, 2013). Madde bilgi fonksiyonları ve madde bilgi fonksiyonlarından elde edilen test bilgi fonksiyonları, madde ve testleri tanımlamada, test maddelerini seçmede ve testleri karşılaştırmada kullanılan önemli göstergelerdir. Test geliştirme sürecinde MBF'den faydalanılması MKE'nin veriye uyumlu olmasına bağlıdır (Hambleton ve diğ., 1991). MKE ile test verisi yeterince uyumlu değilse, bu durum bilgi fonksiyonundan yapılacak kestirimlerin hatalı olabileceğini gösterir. Bunun önüne geçebilmek adına varsayımlar test edilmeli ve sağlandığından emin olunmalıdır.

Madde Bilgi Fonksiyonu, maddenin ölçülen özelliğe ilişkin ne kadar bilgi verdiğini ortaya koyan bir fonksiyondur. Bir madde tarafından sağlanan bilgi, maddenin bireyin yetenek düzeyinin tahmin edilmesine sağladığı katkı olarak yorumlanmaktadır. Madde bilgisi, maddeye ilişkin hata varyansı ile ters orantılıdır (Reid, Kolakowsky-Hayner, Lewis ve Armstrong, 2007). TBF, uygulanan testin her bir yetenek düzeyinde verdiği bilgiyi gösterir ve MTK'da güvenilirliğin kestirilmesinde, standart hatanın hesaplanmasında temel alınır (DeMars, 2010). Bu nedenle testin geliştirilmesi sürecinde önemli bir yere sahiptir. Standart hata daha öncede ifade edildiği gibi her madde için, her yetenek düzeyi için ve her birey için kestirilebilmektedir ve özellikle testin uygulandığı kişilerin yetenek düzeyine uyumlu maddelerin testte yer almasıyla ilişkilidir (Hambleton ve diğ., 1991). MTK'nın, KTK'nın sınırlılıklarına karşılık olarak sağladığı birçok avantaj bulunmaktadır; ancak MTK'nın yukarıda bahsedilen avantajları yalnızca model veri uyumu kabul edilebilir derecede sağlandığında geçerlidir. Model-veri uyumu düşük olduğunda madde ve yetenek parametrelerinin değişmezliği sağlanamayacağından, avantajlarından da faydalanılamayacaktır (Hambleton ve diğ., 1991; Orlando ve Thissen, 2000). Bu nedenle parametrelerin kestiriminde hangi modelin kullanılacağına belirlenmesi gerekir. Bu belirleme model veri uyumunun değerlendirilmesi, dolayısıyla uyum indekslerinin incelenmesi ile mümkündür (Orlando ve Thissen, 2000).

Model veri uyumunun değerlendirilmesinde belirli bir MTK modelinin veriye ve maddelerin modele ne kadar uyumlu olduğunu ortaya koymak amaçlanır. Bir modelin bir veri setine kesin olarak uyumlu veya uygun olup olmadığı yargısına varmayı sağlayan mutlak bir yöntem henüz geliştirilememiştir. Model veri uyumunda ayrı ayrı maddelerin model ile uyumu ve modelin veri ile uyumu incelenebilir. Genellikle MTK uygulayıcıları, tüm maddelere karşılık gelen model uyumu yerine, her bir maddenin model ile uyumuna odaklanmıştır (Demars, 2010).

Bir modelde, model veri uyumu zayıf ise, madde istatistikleri ve madde bilgi fonksiyonları da yanıltıcı olacaktır (Ackerman, 2005; Chang, 1992; Hambleton ve diğ., 1991). Başka bir deyişle, MBF ve dolayısıyla TBF'nin ve standart hatanın yanıltıcı olması, model veri uyumunun sağlanamaması ile açıklanabilir. Bu nedenle güvenilirliğin sağlanması ve elde edilecek bilgilerin doğruluğu için, model ve madde uyumunun sağlandığı en doğru modelin seçilmesi gerekmektedir. Görüldüğü üzere model veri uyumu ölçmenin geçerliği ve güvenilirliği açısından büyük önem taşımaktadır.

Madde Tepki Kuramı'nda model veri uyumu aracılığıyla kullanılacak modele karar verirken göz önünde bulundurulmuş temel ölçüt birey hakkında en çok bilgiyi en az hata ile kestirebilecek modeli

seçmektir. Böylece bireyin verdiği yanıtlardan hareketle, ölçülmeye çalışılan gizil özelliği hakkında daha yüksek doğrulukla kestirimler yapılabilir. Testlerle ölçülmeye çalışılan gizil değişkenlere ulaşılabilmesi, bireylerin kendilerine testle yöneltilen maddelere yanıt vermesi ile mümkündür. Bu bağlamda yanıtlayıcıların kendilerine yöneltilen maddeleri herhangi bir nedenle yanıtlamaktan kaçınması, boş bırakması ya da atlaması, aslında bu tür bir çıkarımın yapılabilmesinin önündeki en önemli engeli oluşturur (Hohensinn ve Kubinger, 2011).

Araştırmalarda kayıp veri sorununun bulunması durumunda, tüm veriler ile uyumlu en doğru modelin seçildiğinin düşünülmesi doğru olmayacaktır ya da bir başka deyişle, veriler için doğru modelin seçiminde sorunlar yaşanacaktır. Yanıt örüntüsündeki eksiklikler, yeteneğin kestirilmesinin önünde de önemli bir engel oluşturacaktır. Ancak araştırmalarda her zaman eksiksiz veri setleri elde edebilmek ya da başka bir deyişle kayıp verilerin önüne geçebilmek çok da mümkün değildir. Hohensinn ve Kubinger (2011), kayıp değerlerin birçok çalışma kapsamında karşılaşılan ve çalışma uzman kişilerce ne kadar dikkatli planlanırsa planlansın tam olarak önüne geçilmesi mümkün olmayan bir problem olduğunu belirtmektedirler.

Kayıp veriler araştırma sürecinde önemli bir sorundur; çünkü standart istatistiksel yöntemlerin uygulandığı veri setleri, satırları gözlemlerden, sütunları ise değişkenlerden oluşan matrislerdir ve bir değişkene ilişkin bir gözlemin eksik olması, söz konusu gözlemi temsil eden hücrenin boş kalmasına neden olur. Bu da analiz sürecini engelleyecek / etkileyecek bir problem ile karşılaşılması anlamına gelir (Little ve Rubin, 1987). Bununla birlikte yanıtlamama ya da boş bırakmaya bağlı olarak ortaya çıkan kayıp veriler, veri setinin daralmasına ve yapılacak kestirimlerin zayıflamasına yol açar. Kayıp veriler içeren veri seti üzerinde istatistiksel analizler yapılamadığı gibi, yanıt veren ve vermeyen bireyler arasında da, çoğunlukla sistematik olan farklılıklardan dolayı, bir yanlılık oluşmasına neden olabilir (Rubin, 1987).

Rubin (1987), “yanıtlamama oranının” dikkat edilmesi gereken bir nokta olduğunu vurgulamaktadır. Çünkü yanıtlamama oranı arttıkça, sistematik hatanın yani yanlılığın ortaya çıkma olasılığının da artacağı yönünde görüşler bulunmaktadır (Baraldi ve Enders, 2010; Enders ve Bandalos, 2001; Hohensinn ve Kubinger, 2011). Buna karşın yanıtlamama oranı ile yanıtlamama yanlılığı arasında doğrudan bir ilişki bulunmadığı yönünde görüşler de mevcuttur. Bu farklı görüşler kayıp verilere ilişkin tartışmaları, kayıp verilerin ihmal edilebilir olup olmadığının yoklanması noktasına taşımıştır. Kayıp verilerin ihmal edilebilir olması, kayıp verilerin seçkisiz olarak oluştuğu, yani herhangi bir örüntü/yanlılık içermediği, dolayısıyla veri dağılımında bir sapma ya da farklılık oluşmayacağı, verinin çok değişkenli olması durumunda, çok değişkenli normallik sağlanırsa, her bir değişkene ilişkin kayıp veri oluşma durumunun eşit olacağı gibi anlamlara gelmektedir. Rubin (1976), kayıp verilerin herhangi bir örüntüye sahip olmadığını, rastgele dağıldığının kanıtlanması durumunda, eksiksiz veri seti ile kayıp veri içeren veri setinden yapılacak kestirimler arasında manidar bir fark olmayacağını ifade etmektedir.

Alanyazında, kayıp verileri ve kayıp veri baş etme yöntemlerini konu alan çok sayıda araştırma mevcuttur (Chen, Wang ve Chen, 2011; Çokluk ve Kayri, 2011; Demir, 2013; Demir ve Parlak, 2012; Doğanay Erdoğan, 2012; Enders ve Bandalos, 2001; Enders, 2004; Fiona ve diğ., 2006; Furlow ve diğ., 2007; Köse, 2014; Sijtsma ve van der Ark., 2003; Van Ginkel, 2007). Kayıp verilerin neredeyse her çalışmada karşılaşılan bir sorun olması ve önceden kestirilemediği için önlenemiyor olması, konuya ilişkin araştırmaların her geçen gün artmasına ve farklı koşullar altında etkilerinin neler olacağını incelenmesine yol açmaktadır.

### ***Araştırmanın Amacı***

Bu çalışmada kayıp verilerle baş etmede kullanılan yöntemlerin MTK modellerinden 1PLM’de model veri uyumu ve madde model uyumu üzerindeki etkilerinin farklı örneklem büyüklüğü ve kayıp veri koşulları dikkate alınarak incelenmesi amaçlanmaktadır. Bu genel amaç doğrultusunda aşağıdaki sorulara yanıt aranmıştır:

1. Farklı örneklem büyüklüğü ve kayıp veri oranlarında tamamen rastgele kayıp mekanizmasında liste bazında silme, regresyonla atama ve beklenti maksimizasyon algoritması ile atama yöntemlerinin

a. model veri uyumu

b. madde model uyumu üzerindeki etkileri nelerdir?

2. Farklı örneklem büyüklüğü ve kayıp veri oranlarında rastgele kayıp mekanizmasında liste bazında silme, regresyonla atama ve beklenti maksimizasyon algoritması ile atama yöntemlerinin

a. model veri uyumu

b. madde model uyumu üzerindeki etkileri nelerdir?

## YÖNTEM

Bu bölümde araştırmanın modeli, verilerin üretilmesi, istenen koşullara uygun kayıplar içeren veri setlerinin elde edilmesi, kayıp veri içeren veri setlerinin farklı kayıp veri baş etme yöntemleri ile tamamlanması ve verilerin analizi ile ilgili açıklamalara yer verilmiştir.

### *Araştırmanın Modeli*

Araştırma, farklı kayıp veri baş etme yöntemlerinin Madde Tepki Kuramı bir parametrelili lojistik modelinde model veri uyumuna ve standart hataya olan etkisini yapay (simulatif) veri setleri kullanarak ortaya koymayı amaçlayan bir simülasyon çalışmasıdır. Aynı zamanda var olan kuramsal bilgiye yenilerini katmayı amaçladığından, (Karasar, 2007) temel araştırma niteliği taşıdığı ifade edilebilir.

### **Verilerin Üretilmesi**

Bu çalışmada verilerin üretilmesinde, üretilen eksiksiz veri setlerinde kayıp veri mekanizmalarına uygun ve istenen oranlarda kayıp veri oluşturulmasında R programı kullanılmıştır (R Development Core Team, 2011). Araştırmada veri üretimi, Madde Tepki Kuramı tek boyutlu lojistik modellerinden 1PLM temel alınarak gerçekleştirilmiştir. Üretilen veri seti için madde sayısı 20, kişi sayısı 500, 1000 ve 1500 ve her maddeye ait yanıt kategori sayısı iki (1-0 puanlama) olarak belirlenmiştir. Schafer (1997), tekrar ve kestirim sayısının 100'den az olduğu durumlarda kestirimlerin birleştirilmesi ile elde edilecek değerlerin güvenilir olmayacağını belirtmiştir. Bu nedenle üretilen verilerde, yapılacak atama sayısında ve parametre kestirimlerinde 100 Monte Carlo (MC) tekrarı yapılmıştır.

### *Madde sayısının belirlenmesi*

Aiken'in (1995), iki kategorili puanlanan maddelerden oluşan testlerde 20 maddeden daha az madde bulunması durumunda, testle ölçülmek istenen özelliğin süreklilik gösterme olasılığının düşeceği yönündeki görüşü dikkate alınarak, üretilen veri setlerinin 20 madde ile sınırlandırılmasına karar verilmiştir.

### *Örneklem büyüklüğünün belirlenmesi*

Alanyazında örneklem büyüklüğünü (n) dikkate alan araştırmalar incelendiğinde genellikle, 250 (Goldman ve Raju, 1986), 500 (Baker, 1998; Gao ve Chen, 2005; Goldman ve Raju, 1986; Hulin, Lissak ve Drasgow, 1982; Thissen ve Wainer, 1982), 1000 (Goldman ve Raju, 1986; Lord, 1968; Hulin, Lissak ve Drasgow, 1982; Thissen ve Wainer, 1982; Yen, 1987) ve 2000 (Gao ve Chen, 2005;

Hulin, Lissak ve Drasgow, 1982) kişilik örneklemelerin kullanıldığı belirlenmiştir. Bu araştırmaların bulguları ve Madde Tepki Kuramı'nın varsayımları göz önünde bulundurularak, bu çalışmada örneklem büyüklüklerinin 500,1000 ve 1500 kişi ile sınırlandırılmasına karar verilmiştir.

#### *Yetenek düzeyinin belirlenmesi*

Veriler üretilirken örneklem büyüklüğü koşulunda belirlenen kişi sayıları göz önünde bulundurularak, kişilere ait yetenek ( $\theta$ ) parametrelerini içeren vektör minimum -3 logit ve maksimum +3 logit olacak şekilde sınırlandırılmıştır. Veriler, belirlenen bu değerler arasında  $(\theta_{maksimum}-\theta_{minimum})/(kişi\ sayısı-1)$  birimlik artışlarla normal dağılım gösterecek şekilde oluşturulmuştur.

#### *Madde parametrelerinin belirlenmesi*

Veri üretiminde temel alınan modelde olması gereken parametreler güçlük (b) ve ayırt edicilik (a) parametreleridir. Madde güçlükleri (b), minimum -2 logit, maksimum +2 logit olacak şekilde sınırlandırılmıştır. Madde ayırt edicilik indeksinin (a) tanımlanan normal aralığı (0, 2)'dir (Hambleton, Swaminathan ve Rogers, 1991). Hambleton'a (1989) göre, 1.35 ile 1.69 aralığı yüksek ayırt ediciliğe işaret etmektedir buna göre üretilen verilerde madde ayırt ediciliği 1.5 olarak sabitlenmiştir.

#### *Kayıp Verilerin Oluşturulması*

Alanyazın incelendiğinde kayıp verilerle ilgili yapılan çalışmalarda en çok %5, %10 ve %15 (Chen, Wang, Chen, 2005; Doğanay Erdoğan 2012; Enders ve Bandalos 2001; Fiona ve diğ., 2006; Furlow ve diğ., 2007; Sijtsma ve van der Ark, 2003) kayıp veri oranlarının kullanıldığı belirlenmiştir. Bu çalışmaların bulguları göz önünde bulundurularak, üretilen verilerde Tamamen Rastgele Kayıp (TRK) ve Rastgele Kayıp (RK) mekanizmalarını sağlayan %5, %10 ve %15 oranlarında kayıp veri oluşturulmuştur.

Veri silme işlemi için her bir kayıp veri mekanizmasının belirlenen koşullarını sağlayacak kodlar R programında yazılmış ve kayıp verilerin oluşturulması sağlanmıştır. TRK mekanizmasında kayıplar oluşturulurken 500, 1000 ve 1500 kişilik örneklemelerde herhangi bir değişkene bağlı olmaksızın, örneklem büyüklüğünden belirlenen oranlarda gözenek silinmiştir. RK, kayıp veri mekanizmasını oluşturabilmek için öncelikle ölçülen özellikten farklı ama ölçülen bireylere ait, üç düzeyi olan sıralama ölçeği düzeyinde bir başka değişken rastgele olarak tanımlanmıştır. Daha sonra 1. 2. ve 3. düzeylerde sırasıyla %20, %30 ve %50 oranlarında kayıp veri oluşturulmuştur.

Elde edilen eksikli veri setleri, kayıp veri baş etme yöntemlerinden atamaya dayalı yöntemler ile veri tamamlama aşamasında tamamlanmıştır. Bu yöntemler basit atama yöntemlerinden olan Regresyonla Atama (RA), model tabanlı atama yöntemlerinden Beklenti Maksimizasyon Algoritması ile Atama (BMA) ve silmeye dayalı yöntemlerden biri olan Liste bazında Silme (LBS) yöntemleridir. LBS yönteminde veri seti, veri atama ile değil, eksik gözeneği olan kişileri silme yoluyla eksiksiz hale getirilir. Bu nedenle LBS yöntemi için farklı bir kayıp veri mekanizması kodu yazılmıştır. TRK'da 500, 1000 ve 1500 kişilik örneklemelerde herhangi bir değişkene bağlı olmadan örneklem büyüklüğünden belirlenen oranlarda kişi silinmiştir. RK mekanizmasını oluşturmak için ise kişilere rastgele olarak, üç düzeyi olan sıralama ölçeği düzeyinde bir başka değişken tanımlanmıştır. Daha sonra 1. 2. ve 3. düzeylerden sırasıyla %20, %30 ve %50 oranlarında kayıp veri oluşturulmuştur.

#### *Eksik Verilerin Tamamlanması*

Beklenti maksimizasyon algoritması ve regresyonla atama yöntemleri ile veri ataması IBM SPSS 21.00 paket programı ile yapılmıştır. Liste bazında silme yöntemi için R programı kullanılmıştır.

Liste bazında silme yönteminde belirlenen senaryolara (koşullara) uygun olacak şekilde kişi silme işlemi gerçekleştirilmiştir.

### Verilerin Analizi

Araştırmada öncelikle üretilen eksiksiz veri setlerinden model veri uyumu kestirilmiştir. Ardından tamamlanmış veri setlerinde aynı kestirimler yapılmış ve eksiksiz veri setinden elde edilen değerler referans kabul edilerek, bu değerlerle karşılaştırılmıştır. Kestirimlerin hesaplanmasında R programı içerisinde yer alan {lrm} ve {dplyr} paketlerinden yararlanılmıştır.

### Model veri uyumu

En iyi modeli seçmek için kullanılan yöntemlerden biri  $-2 \log \lambda$  fark istatistiğine yani  $G^2$ 'ye göre karar vermektir (Dais, 2006; Morren, Gelisen ve Vermut, 2011).  $G^2$ 'nin istatistiksel olarak manidar olup olmadığına, söz konusu iki modelin serbestlik dereceleri arasındaki farkın kritik değer ile karşılaştırılması yoluyla karar verilir.

Serbestlik derecesinin yüksek olduğu koşullarda  $G^2$  istatistiklerinin  $\chi^2$  dağılımı göstermeyebileceği belirtilmektedir (Collins ve Lanza, 2010). Bu durumlar için alternatif yol olarak AIC ve BIC bilgi kriterlerinin kullanılması genel olarak kabul görmüştür (Bauer ve Curran, 2003; Dias, 2006; Lin, 2006;2012; Nylund ve diğ., 2007; Vrieze, 2012; Yang ve Yang, 2007). Buna bağlı olarak bu çalışmada model veri uyumunun belirlenmesinde  $-2 \log \lambda$  ile birlikte AIC ve BIC yöntemleri de kullanılmıştır.

Model veri uyumunun kestirilmesinde R programında yer alan {lrm} paketindeki fonksiyonlardan faydalanılarak  $-2 \log \lambda$ , AIC ve BIC değerleri kestirilmiştir. Aşağıda eksiksiz veri setlerinden kestirilen model veri uyumu değerleri sunulmuştur.

Tablo 1. Eksiksiz Veri Setinden Kestirilen Bir Parametrelili Lojistik Modelde  $-2 \log \lambda$ , AIC ve BIC Kriterlerine Göre Model Veri Uyumu

Bilgi Kriterleri	Örneklem Büyüklüğü		
	500	1000	1500
$-2 \log \lambda$	-4575.22	-8982.30	-13779.42
AIC	9192.45	18006.61	27600.85
BIC	9280.96	18109.68	27712.43

Eksiksiz veri setinin ve kayıp veri baş etme yöntemleri ile tamamlanmış veri setlerinin IPLM'e uyumlarını gösteren  $-2 \log \lambda$  değerleri hesaplanmış ve eksiksiz veri seti ile aralarındaki farkın manidarlığın belirlenmiştir. Böylece hangi kayıp veri koşulunda, hangi kayıp veri baş etme yönteminin eksiksiz veri setine daha benzer sonuçlar ürettiği belirlenmeye çalışılmıştır.

Tamamlanmış veri setlerinden elde edilen AIC ve BIC değerleri bulgular bölümünde eksiksiz veri setinden elde edilen değerler referans alınarak yorumlanmıştır. AIC değeri doğrudan yorumlanamaz; çünkü bir modele ilişkin model veri uyumunun iyi ya da kötü olduğuna ilişkin bilgi sunmaz, karşılaştırılan modellerden hangisinin veriye daha uygun olduğuna dair karşılaştırma yapma olanağı sunar. Buna göre, daha küçük AIC ve BIC değerine sahip modelin veriye ya da daha küçük AIC değerine sahip veri setinin modele daha uyumlu olduğu sonucuna varılır (Blozis ve diğ. 2007, Duncan ve diğ., 2006 ). Bu doğrultuda eksiksiz veri seti ve kayıp veri baş etme yöntemleri ile tamamlanmış veri setlerinden elde edilen AIC ve BIC değerleri karşılaştırılırken, kayıp veri baş etme yöntemlerinin performansı betimsel olarak yorumlanmıştır.

### Madde model uyumu

Embretson ve Reise (2000), yalnız model veri uyumunun belirlenmesinin yeterli olmayacağını bunun yanı sıra madde model uyumunun da belirlenmesi gerektiğini belirtmiştir. Orlando ve Thissen (2000), bir maddenin model ile uyumunun belirlenmesinde bir grafik ya da  $\chi^2$  benzeri bir istatistik kullanılarak, gözlenen değerler ile beklenen (model yardımıyla kestirilen) değerlerin karşılaştırılmasını önermektedir. Bu nedenle bu çalışmada da madde model uyumunun belirlenmesinde  $\chi^2$  madde-uyum yaklaşımı kullanılmıştır. Bu yaklaşıma göre, modelde her bir madde için gözlenen ve beklenen değerler, farklı yetenek düzeylerindeki bireyler ve gruplar için karşılaştırılarak  $\chi^2$  dağılımı gösteren olabirlik (likelihood) oran istatistiği elde edilir ve bu istatistik için serbestlik derecesi, cevaplayıcılar için oluşturulan yetenek aralıkları sayısına eşittir (Mislevy ve Bock, 1990). Eksiksiz veri seti ve kayıp veri baş etme yöntemleri ile tamamlanan veri setlerinde madde model uyumu hesaplanmış ve eksiksiz veri seti referans alınarak tamamlanan veri setlerindeki uyumlu madde sayıları karşılaştırılmıştır.

Tablo 2. Eksiksiz Veri Setlerinde Kestirilen Madde Model Uyumu

Madde No	Örneklem Büyüklüğü					
	500		1000		1500	
	$\chi^2$	P	$\chi^2$	p	$\chi^2$	p
M1	14.602	0.406	10.094	0.814	18.738	0.095
M2	18.925	0.168	15.252	0.433	9.387	0.670
M3	16.549	0.281	23.728	0.070	16.208	0.182
M4	8.373	0.869	17.199	0.307	12.024	0.444
M5	11.115	0.677	18.830	0.222	23.400	0.065
M6	12.606	0.558	21.085	0.134	23.241	0.126
M7	17.422	0.234	33.536	0.004	33.687	0.011
M8	11.605	0.638	26.549	0.033	22.789	0.070
M9	16.845	0.265	14.092	0.519	20.796	0.053
M10	23.987	0.046	14.500	0.488	19.967	0.068
M11	7.647	0.907	16.586	0.344	28.788	0.004
M12	10.611	0.716	24.414	0.058	15.794	0.201
M13	20.327	0.120	31.305	0.008	24.366	0.018
M14	18.154	0.200	18.884	0.219	39.570	0.000
M15	8.518	0.861	30.995	0.009	15.356	0.223
M16	6.161	0.962	19.946	0.174	23.006	0.028
M17	13.799	0.465	21.659	0.117	9.846	0.629
M18	20.957	0.103	28.457	0.019	24.337	0.058
M19	16.180	0.303	28.957	0.016	13.688	0.321
M20	13.608	0.479	24.459	0.058	20.566	0.089

Tablo 2’de yer alan değerler tamamlanmış veri setlerinden kestirilecek madde model uyumu için referans olarak kabul edilmiş ve yorumlamalar da bu değerlerle karşılaştırılarak yapılmıştır.

### BULGULAR

Bu bölümde araştırmanın amaçları doğrultusunda elde edilen bulgulara yer verilmiştir. Araştırmada 500, 1000 ve 1500 kişilik 1PLM model ile uyumlu verilerde, %5, %10 ve %15 oranında TRK ve RK veri mekanizmalarında LBS, RA ve BMA yöntemlerinin model veri uyumuna, madde model uyumuna etkisi araştırılmıştır. Elde edilen bulgular TRK ve RK mekanizmaları için sırasıyla sunulmuştur. Tablo 3’te Tamamen Rastgele Kayıp mekanizmasında liste bazında silme yöntemi ile tamamlanan veri setlerinden elde edilen  $G^2$  olabirlik oranlarına ilişkin bulgular sunulmuştur.



Tablo 3. Tamamen Rastgele Kayıp Mekanizmasında Liste Bazında Silme Yöntemi ile Tamamlanan Veri Setlerinden Elde Edilen  $G^2$  Olabilirlik Oranı

Eksiksiz Veri Seti	n=500		n=1000		n=1500	
	$-2 \log \lambda$	$G^2$	$-2 \log \lambda$	$G^2$	$-2 \log \lambda$	$G^2$
	-4575.27		-8982.32		-13779.54	
%5	-4347.43	227.83*	-8524.2	458.12*	-13059.80	719.74*
%10	-4123.78	451.48*	-8065.191	917.12*	-12418.96	1360.58*
%15	-3897.29	677.97*	-7650.436	1331.88*	-11716.71	2062.83*

\* $p(\chi^2_{sd=19} > 30,144) < 0.05$

Tablo 3 incelendiğinde veri setinde TRK olduğu durumda LBS yöntemi ile tamamlanan veri setinde model veri uyumu  $-2 \log \lambda$  ile sınılandığında tüm örneklem büyüklüklerinde (500, 1000 ve 1500) kayıp veri oranı arttıkça bu değerin düştüğü görülmektedir. Her bir  $-2 \log \lambda$  sütununda ilk olarak eksiksiz veri setinden elde edilen  $-2 \log \lambda$  değeri verilmiş olup sütunda aşağı doğru inildikçe kayıp veri miktarlarında da artış görülmektedir. Ancak buna karşılık  $-2 \log \lambda$  değeri ise azalmaktadır. Dolayısıyla eksiksiz veri setinden elde edilen  $-2 \log \lambda$  değeri ile, tamamlanan veri setlerinden elde edilen  $-2 \log \lambda$  arasındaki fark yani  $G^2$  değeri de giderek artmaktadır.  $\chi^2$  dağılımının tablo değeriyle karşılaştırıldığında eksiksiz veri seti ve LBS yöntemi ile tamamlanan veri setlerinden kestirilen model veri uyumları arasında manidar farklılık olduğu görülmektedir. TRK mekanizmasında LBS yöntemi ile tüm örneklem büyüklüklerinde, kayıp veri oranı arttıkça eksiksiz veri setinden sapmanın giderek artması, LBS yönteminden kaynaklı olarak örneklemin küçülmesi ile açıklanabilir. Çünkü LBS yönteminin TRK mekanizmasında kullanılması durumunda, kayıp verili kişiler silindiğinden, veri setinde kalan gözlemler eldeki verinin rastgele bir alt örneklemini oluşturmaktadır (Allison, 2003; Enders, 2010; McKnight ve diğ., 2007; Rosenthal ve Rosnow, 2008) ve örneklemin küçülmesi  $\chi^2$  istatistiğine dayalı uyum iyiliğini artırmaktadır (Drasgow, Levine, Tsien, Williams ve Mead, 1995; Bock, 1997). Madde Tepki Kuram'ında model parametrelerinin kestiriminde ve model veri uyumu üzerinde örneklem büyüklüğü önemli bir etkidir (Bolt ve Lall, 2003; Seungho-Yang, 2007). Hambleton ve diğerleri (1991),  $-2 \log \lambda$ 'in  $\chi^2$  dağılımı gösterdiğini ve bu istatistiğin örneklem büyüklüğüne karşı oldukça hassas olduğunu, büyük örneklem söz konusu olduğunda model veri uyumunun neredeyse tüm modellerde sağlanmadığını vurgulamaktadır. Tablo 4'te TRK mekanizmasında LBS silme yöntemi ile tamamlanan veri setlerinden elde edilen AIC değerleri sunulmuştur.

Tablo 4. Tamamen Rastgele Kayıp Mekanizmasında Liste Bazında Silme Yöntemi ile Tamamlanan Veri Setlerinden Elde Edilen AIC Değerleri

Eksiksiz Veri Seti	n=500		n=1000		n=1500	
	AIC	$\Delta$ AIC	AIC	$\Delta$ AIC	AIC	$\Delta$ AIC
	9190.54		18004.66		27599.07	
%5	8736.86	453.67	17090.40	914.26	26161.59	1437.48
%10	8289.57	900.96	16172.38	1832.22	24879.91	2719.16
%15	7836.58	1353.95	15342.87	2661.79	23475.42	4123.65

Tablo 4 incelendiğinde, tüm örneklem büyüklüklerinde kayıp veri oranı arttıkça, AIC değerinde düşüş yaşandığı görülmektedir. Dolayısıyla kayıp veri oranı arttıkça, LBS yöntemi, AIC kriteri açısından eksiksiz veri setinden uzaklaşmış ve veriyi modele daha uyumlu hale getirmiştir. Tablo 5'te TRK mekanizmasında LBS yöntemi ile tamamlanan veri setlerinden elde edilen BIC değerleri sunulmuştur.

Tablo 5. Tamamen Rastgele Kayıp Mekanizmasında Liste Bazında Silme Yöntemi ile Tamamlanan Veri Setlerinden Elde Edilen BIC Değerleri

Eksiksiz Veri Seti	n=500		n=1000		n=1500	
	BIC	ΔBIC	BIC	ΔBIC	BIC	ΔBIC
	9274.84		18102.81		27705.34	
%5	8824.29	450.54	17192.39	910.42	26272.09	1433.25
%10	8375.86	898.97	16273.23	1829.58	24989.28	2716.06
%15	7921.67	1353.16	15442.52	2660.29	23583.58	4121.76

Tablo 5 incelendiğinde, eksiksiz veri seti ve tamamlanan veri setlerinden elde edilen değerler karşılaştırıldığında, tüm örneklem büyüklüklerinde kayıp veri oranı arttıkça, BIC değerinin düştüğü ve eksiksiz veri setinden uzaklaştığı görülmektedir. Bu uzaklaşma veri setinin 1PLM'e daha uyumlu hale gelmesine yol açmıştır. Model veri uyumunun değerlendirilmesinde kullanılan AIC ve BIC kriterlerinin formülleri incelendiğinde ( $AIC = -2 \log \lambda + 2k$ ,  $BIC = -2 \log \lambda + k \log(n)$ ) k sabit terim dahil parametre sayısı ve n gözlem sayısı (Akaike, 1973;1974; Hurvich ve Tsai, 1989; Siungiura, 1978) olmak üzere, her iki kriterin de  $-2 \log \lambda$ 'den türetildiği görülmektedir. Dolayısıyla örneklemin küçülmesinin  $\chi^2$  istatistiğine dayalı uyum iyiliğini artırdığı (Bock, 1997; Drasgow ve diğ., 1995) ve  $-2 \log \lambda$ 'in  $\chi^2$  dağılımı gösterdiği göz önünde bulundurulduğunda, AIC ve BIC kriterlerinin de örneklemden etkileneceği, buna bağlı olarak LBS yönteminin örneklemini daraltmasıyla model veri uyumunu artırdığı sonucuna varılabilir. Kankaras ve diğerleri (2011) AIC ve BIC değerlerinin örneklem büyüklüğünden etkilendiğini belirtmektedir. Dolayısıyla LBS yönteminin kullanılması örneklemin küçültmesi sonucu AIC, BIC ve  $-2 \log \lambda$  değerlerinde bir küçülmeye ve buna bağlı olarak da model veri uyumunda artışa neden olmaktadır. Tablo 6'da TRK mekanizmasında RA ile tamamlanan veri setlerinden elde edilen  $G^2$  olabilirlik oranları sunulmuştur.

Tablo 6. Tamamen Rastgele Kayıp Mekanizmasında Regresyonla Atama Yöntemi ile Tamamlanan Veri Setlerinden Elde Edilen  $G^2$  Olabilirlik Oranı

Eksiksiz Veri Seti	n=500		n=1000		n=1500	
	$-2 \log \lambda$	$G^2$	$-2 \log \lambda$	$G^2$	$-2 \log \lambda$	$G^2$
	-4575.27		-8982.32		-13779.54	
%5	-4505.32	69.945*	-8818.56	163.76*	-13490.39	289.15*
%10	-4392.78	182.48*	-8626.59	355.27*	-13259.96	519.58*
%15	-4290.10	285.16*	-8441.24	541.07*	-12929.28	850.26*

\* $p(\chi^2_{sd=19} > 30,144) < 0.05$

Tablo 6 incelendiğinde, RA yöntemi ile tamamlanan veri setlerinin TRK mekanizması olduğu durumda tüm örneklem büyüklüklerinde kayıp veri oranı arttıkça, eksiksiz veri setinden kestirilen  $-2 \log \lambda$  değerinden uzaklaştığı görülmektedir.  $G^2$  oranının manidarlığı test edildiğinde ise, RA yöntemi ile tamamlanan veri setlerinin 1PLM'e eksiksiz veri setinden daha uyumlu olduğu görülmektedir. RA yöntemiyle tamamlanan veri setlerinde TRK olması durumunda, tüm örneklem büyüklüklerinde kayıp veri oranı arttıkça, eksiksiz veri setinden kestirilen  $-2 \log \lambda$  değerinden daha da uzaklaştığı ifade edilebilir. Tablo 7'de TRK mekanizmasında RA yöntemi ile tamamlanan veri setlerinden elde edilen AIC değerleri sunulmuştur.

Tablo 7. Tamamen Rastgele Kayıp Mekanizmasında Regresyonla Atama Yöntemi ile Tamamlanan Veri Setlerinden Elde Edilen AIC Değerleri

Eksiksiz Veri Seti	n=500		n=1000		n=1500	
	AIC	$\Delta$ AIC	AIC	$\Delta$ AIC	AIC	$\Delta$ AIC
	9190.54		18004.66		27599.07	
%5	9050.64	139.89	17677.12	327.54	27020.78	578.29
%10	8825.57	364.97	17293.19	711.47	26559.91	1039.16
%15	8620.21	570.32	16922.48	1082.12	25898.56	1700.51

Tablo 7 incelendiğinde, tüm örneklem büyüklüklerinde kayıp veri oranı arttıkça, AIC değerinin azaldığı görülmektedir. Buna bağlı olarak RA yönteminin TRK mekanizmasına sahip kayıp veri olduğu durumda, tüm örneklem büyüklüklerinde kayıp veri oranı arttıkça eksiksiz veri setinden uzaklaştığı ve IPLM'e daha uyumlu bir hale dönüştürdüğü belirtilebilir. Tablo 8'de TRK mekanizmasında RA yöntemi ile tamamlanan veri setlerinden elde edilen BIC değerleri sunulmuştur.

Tablo 8. Tamamen Rastgele Kayıp Mekanizmasında Regresyonla Atama Yöntemi ile Tamamlanan Veri Setlerinden Elde Edilen BIC Değerleri

Eksiksiz Veri Seti	n=500		n=1000		n=1500	
	BIC	$\Delta$ BIC	BIC	$\Delta$ BIC	BIC	$\Delta$ BIC
	9274.84		18102.81		27705.34	
%5	9134.94	139.89	17775.28	327.53	27127.05	578.29
%10	8909.86	364.97	17391.35	711.46	26666.18	1039.16
%15	8704.50	570.33	17020.64	1082.17	26004.83	1700.51

Tablo 8 incelendiğinde, tüm örneklem büyüklüklerinde kayıp veri oranı arttıkça, BIC değerinin düştüğü görülmektedir. Dolayısıyla  $\Delta$ BIC değerleri de göz önünde bulundurulduğunda, TRK mekanizmasında kayıp veri oranı arttıkça, RA yönteminin BIC değeri bakımından eksiksiz veri setinden uzaklaştığı ve IPLM için model veri uyumunu artırdığı görülmektedir. TRK durumunda model veri uyumunda RA yöntemi kullanılarak tamamlanan veri setlerinden elde edilen değerlerin eksiksiz veri setinden küçük sapmalar gösterdiği görülmektedir.  $-2 \log \lambda$ , AIC ve BIC değerleri bakımından sapmaların olması, alanyazın ile de uyum göstermektedir. TRK mekanizması olduğu durumda RA yöntemi, bazı durumlarda tam veri setlerinden elde edilen değerlerden farklı değerler alabilmektedir (Roth ve Switzer, 1995; Suraphee, Raksmanee, Busaba, Chaisorn ve Nakornthai, 2006). Tablo 9'da TRK mekanizmasında BMA yöntemi ile tamamlanan veri setlerinden kestirilen model veri uyumu kriteri olan  $-2 \log \lambda$  değerleri sunulmuştur.

Tablo 9. Tamamen Rastgele Kayıp Mekanizmasında Beklenti Maksimizasyon Algoritması Yöntemi ile Tamamlanan Veri Setlerinden Elde Edilen  $G^2$  Olabilirlik Oranı

Eksiksiz Veri Seti	n=500		n=1000		n=1500	
	$-2 \log \lambda$	$G^2$	$-2 \log \lambda$	$G^2$	$-2 \log \lambda$	$G^2$
	-4575.27		-8982.32		-13779.54	
%5	-4498.01	77.26*	-8806.507	175.82*	-13485.65	293.89*
%10	-4391.44	183.83*	-8613.895	184.13*	-1368.43	239.74*
%15	-4269.46	305.81*	-8410.51	571.81*	-12894.70	884.84*

\* $p(\chi^2_{sd=19} > 30.144) < 0.05$

Tablo 9'da sunulan  $G^2$  değerlerinin manidarlık testleri edildiğinde tüm örneklem büyüklüklerinde kayıp veri oranı arttıkça, model uyumunun iyileştiği gözlenmiştir. BMA yönteminin TRK mekanizmasında model veri uyumunda  $-2 \log \lambda$  kriteri açısından uyumu artırıcı bir performans

sergilediği görülmektedir. Tablo 10'da TRK mekanizmasında BMA yöntemi ile tamamlanan veri setlerinden kestirilen AIC değerleri sunulmuştur.

Tablo 10. Tamamen Rastgele Kayıp Mekanizmasında Beklenti Maksimizasyon Algoritması Yöntemi ile Tamamlanan Veri Setlerinden Elde Edilen AIC Değerleri

Eksiksiz Veri Seti	n=500		n=1000		n=1500	
	AIC	$\Delta$ AIC	AIC	$\Delta$ AIC	AIC	$\Delta$ AIC
	9190.54		18004.66		27599.07	
%5	9036.11	154.42	17654.61	350.05	27015.34	583.73
%10	8822.88	367.65	17273.01	731.65	26528.14	1070.93
%15	8582.22	608.31	16871.29	1133.37	25849.71	1749.36

Tablo 10 incelendiğinde, BMA yöntemiyle tamamlanan veri setlerinin daha düşük AIC değeri verdiği görülmektedir. Tüm örneklem büyüklüklerinde kayıp veri oranı arttıkça, AIC değerinin eksiksiz veri setinden kestirilen değerden daha da saptığı ve model veri uyumunun arttığı görülmektedir. TRK mekanizması olduğu durumda BMA yönteminin AIC kriteri bakımından uyumu artırıcı atamalar yaptığı sonucuna ulaşılmaktadır. Tablo 11'de TRK mekanizmasında BMA yöntemi ile tamamlanan veri setlerinden elde edilen BIC değerleri sunulmuştur.

Tablo 11. Tamamen Rastgele Kayıp Mekanizmasında Beklenti Maksimizasyon Algoritması Yöntemi ile Tamamlanan Veri Setlerinden Elde Edilen BIC Değerleri

Eksiksiz Veri Seti	n=500		n=1000		n=1500	
	BIC	$\Delta$ BIC	BIC	$\Delta$ BIC	BIC	$\Delta$ BIC
	9274.84		18102.81		27705.34	
%5	9120.41	154.42	17752.77	350.04	27121.61	583.73
%10	8907.17	367.66	17371.17	731.64	26634.41	1070.93
%15	8666.51	608.32	16969.44	1133.37	25955.98	1749.36

Tablo 11 incelendiğinde, tüm örneklem büyüklüklerinde kayıp veri oranı arttıkça, BIC değerinin düştüğü, dolayısıyla da model veri uyumunun BIC kriteri açısından arttığı görülmektedir. Kayıp veri oranı sabit tutulduğunda, örneklem büyüklüğü arttıkça eksiksiz veri setinden kestirilen BIC değeri ile BMA yöntemi ile tamamlanmış veri setlerinden elde edilen BIC değerleri arasındaki farkın giderek arttığı gözlenmektedir. TRK koşulunda BMA yöntemi ile yapılan atamalar model veri uyumu bakımından eksiksiz veri setine yakın kestirimler sunmaktadır. Enders (2010) ve Schafer ve Graham (2002) TRK mekanizmasında BMA yöntemi ile yansız parametre tahminleri elde edilebildiğini belirtmektedir. Bu çalışmada elde edilen bulgular da bunu desteklemektedir. Tablo 12'de TRK mekanizmasında 1PLM ile uyum gösteren madde sayıları sunulmuştur.

Tablo 12. Tamamen Rastgele Kayıp Mekanizmasında 1PLM ile Uyum Gösteren Madde Sayısı

	n=500			n=1000				n=1500				
	Eksiksiz Veri Seti	%5	%10	%15	Eksiksiz Veri Seti	%5	%10	%15	Eksiksiz Veri Seti	%5	%10	%15
LBS		20	19	20		15	14	17		17	16	17
RA	19	16	15	16	14	14	13	12	15	13	10	9
BMA		17	17	12		15	11	9		8	6	6

Tablo 12 incelendiğinde, örneklem büyüklüğü  $n=500$  olan eksiksiz veri setinde 1PLM'e uyum gösteren madde sayısı 19,  $n=1000$  örneklem büyüklüğünde 14 ve  $n=1500$  örneklem büyüklüğünde 15'tir. LBS yöntemi ile tamamlanan veri setlerinde tüm örneklem büyüklüklerinde ve kayıp veri oranlarında LBS yöntemi ile eksiksiz hale getirilen verilerle uyumlu olan madde sayısının, eksiksiz veri setinden daha fazla olduğu görülmektedir. Madde Tepki Kuramı'nda model parametrelerinin kestiriminde ve model veri uyumu üzerinde "örneklem büyüklüğü" önemli bir etkidir (Bolt ve Lall, 2003; Seungho-Yang, 2007). Madde model uyumunun belirlenmesinde kullanılan  $\chi^2$  istatistiğinin de örneklem büyüklüğünden etkilendiği düşünülürse (Hambleton ve diğ.,1991), LBS yönteminin kullanılmasının örnekleme küçültürken  $\chi^2$ 'nin manidar çıkmasına yol açtığı ve dolayısıyla uyumlu madde sayısını artırdığı düşünülebilir.

Regresyonla atama yöntemi ile tamamlanan veri setinde model ile uyum sağlayan madde sayısı incelendiğinde, tüm örneklem büyüklüklerinde, genel olarak eksiksiz veri setinden daha az sayıda maddenin uyumu olduğu görülmektedir. Örneklem büyüdükçe ve kayıp veri oranı arttıkça, uyumlu madde sayısı giderek azalmaktadır. Alanyazında RA yönteminin, kayıp veri mekanizmasının TRK olduğu durumlarda yansız parametre tahminleri sunduğu vurgulanmaktadır (Baraldi ve Enders, 2010; McKnight ve diğ., 2007). Ancak elde edilen bulgular bunu desteklememektedir. Buna karşılık bu bulgular, alanyazında RA yönteminin, bazı durumlarda tam veri setlerinden elde edilen değerlerden farklı değerler alabileceği bilgisiyle örtüşmektedir (Roth ve Switzer, 1995; Suraphee ve diğ., 2006).

Beklenti maksimizasyon algoritması yöntemi ile tamamlanan veri setlerindeki model ile uyumlu madde sayısı incelendiğinde, tüm örneklem büyüklüklerinde eksiksiz veri setinden daha az sayıda maddenin uyumlu olduğu ve kayıp veri oranı arttıkça, uyumlu madde sayısının da azaldığı görülmektedir. Enders (2010) ve Schafer ve Graham (2002), TRK koşullarında, BMA yönteminin yansız parametre tahminleri sağlayabildiğini belirtmektedir. Ancak elde edilen sonuçlar bunu desteklememektedir. En çok olabilirlik temelli atama yöntemleri, özellikle RK varsayımının sağlandığı durumlarda daha iyi tahmin değerleri vermektedir (Allison, 2003; Baraldi ve Enders, 2010). Buna bağlı olarak BMA yönteminin TRK mekanizmasında daha düşük bir performans sergilemesi olması bir sonuç gibi görünmektedir. Tablo 13'te RK mekanizmasında LBS yöntemi kullanılarak eksiksiz hale getirilen veri setlerinden kestirilen  $-2 \log \lambda$  değerleri sunulmuştur.

Tablo 13. Rastgele Kayıp Mekanizmasında Liste Bazında Silme Yöntemi ile Tamamlanan Veri Setlerinden Elde Edilen  $G^2$  Olabilirlik Oranı

Eksiksiz Veri Seti	n=500		n=1000		n=1500	
	$-2 \log \lambda$	$G^2$	$-2 \log \lambda$	$G^2$	$-2 \log \lambda$	$G^2$
	-4575.27		-8982.32		-13779.54	
%5	-4361.74	213.52*	-8579.01	403.30*	-13164.56	614.98*
%10	-4141.48	433.78*	-8122.06	860.26*	-12500.73	1278.81*
%15	-3915.42	659.84*	-7685.31	1297.00*	-11812.62	1966.92*

\* $p(\chi^2_{sd=19} > 30.144) < 0.05$

Tablo 13 incelendiğinde,  $G^2$  değerleri  $\chi^2$  dağılımının tablo değeri ile karşılaştırıldığında, tüm örneklem büyüklüklerinde ve kayıp veri oranlarında LBS yöntemi ile eksiksiz hale getirilen veri setlerinin 1PLM ile daha uyumlu olduğu görülmektedir. Tablo 14'te RK mekanizmasında LBS yöntemi ile tamamlanan veri setlerinden elde edilen AIC değerleri sunulmuştur.

Tablo 14. Rastgele Kayıp Mekanizmasında Liste Bazında Silme Yöntemi ile Tamamlanan Veri Setlerinden Elde Edilen AIC Değerleri

Eksiksiz Veri Seti	n=500		n=1000		n=1500	
	AIC	ΔAIC	AIC	ΔAIC	AIC	ΔAIC
	9190.54		18004.66		27599.07	
%5	8765.48	425.05	17200.03	804.63	26371.12	1227.95
%10	8324.97	865.65	16286.12	1718.54	25043.45	2555.62
%15	7872.84	1317.69	15412.63	2592.03	23667.24	3931.83

Tablo 14 incelendiğinde, tüm örneklem büyüklüklerinde kayıp veri oranı arttıkça AIC değerinin eksiksiz veri setinden uzaklaşarak düştüğü, dolayısıyla model veri uyumunun arttığı görülmektedir. RK koşulu altında LBS yönteminin kullanılması IPLM’de model veri uyumunda AIC kriteri açısından model veri uyumunu artırmaktadır. Tablo 15’te RK mekanizmasında LBS yöntemi ile tamamlanan veri setlerinden elde edilen BIC değerleri sunulmuştur.

Tablo 15. Rastgele Kayıp Mekanizmasında Liste Bazında Silme Yöntemi ile Tamamlanan Veri Setlerinden Elde Edilen BIC Değerleri

Eksiksiz Veri Seti	n=500		n=1000		n=1500	
	BIC	ΔBIC	BIC	ΔBIC	BIC	ΔBIC
	9274.84		18102.81		27705.34	
%5	8852.91	421.92	17302.01	800.80	26481.62	1223.72
%10	8411.26	863.57	16386.97	1715.84	25152.82	2552.52
%15	7957.94	1316.89	15512.31	2590.50	23775.40	3929.94

Tablo 15 incelendiğinde tüm örneklem büyüklüklerinde ve kayıp veri oranlarında, LBS yöntemi ile tamamlanan veri setlerinin, eksiksiz veri setinden daha düşük BIC değerlerine sahip oldukları görülmektedir. Tüm örneklem büyüklüklerinde LBS yöntemi ile tamamlanan veri setlerindeki kayıp veri oranı arttıkça, eksiksiz veri setinden uzaklaştıkları ve IPLM’e uyumun arttığı görülmektedir. Bu durum da LBS yöntemine bağlı olarak örneklemin küçülmesi ve böylelikle  $\chi^2$  istatistiğine dayalı uyum iyiliğinin artması ile ilişkilendirilebilir (Bock, 1997; Drasgow ve diğ., 1995). Tablo 16’da RA yöntemi ile tamamlanan veri setlerinden elde edilen  $-2 \log \lambda$  değerleri arasındaki farklara ilişkin  $G^2$  olabilirlik oranları sunulmuştur.

Tablo 16. Rastgele Kayıp Mekanizmasında Regresyonla Atama Yöntemi ile Tamamlanan Veri Setlerinden Elde Edilen  $G^2$  Olabilirlik Oranları

Eksiksiz Veri Seti	n=500		n=1000		n=1500	
	$-2 \log \lambda$	$G^2$	$-2 \log \lambda$	$G^2$	$-2 \log \lambda$	$G^2$
	-4575.27		-8982.32		-13779.54	
%5	-4574.09	1.17	-8965.79	16.52	-13761.04	18.50
%10	-4553.29	21.97	-8946.36	29.95	-13716.70	62.84
%15	-4562.16	13.10	-8932.99	49.32	-13695.28	84.26

\* $p(\chi^2_{df=19} > 30,144) < 0.05$

Tablo 16 incelendiğinde,  $G^2$  oranının manidarlığı test edildiğinde örneklem büyüklüğü n=500 iken RA yönteminin tüm kayıp veri oranlarında eksiksiz veri seti ile benzer sonuçlar verdiği görülmektedir. RA yöntemi küçük örneklerde tüm kayıp veri oranlarında  $-2 \log \lambda$  kriteri açısından eksiksiz veri seti ile karşılaştırıldığında, aradaki farkların manidar olmadığı görülmektedir. Örneklem büyüklüğü n=1000 olduğunda %5 ve %10 kayıp veri oranlarında, örneklem büyüklüğü

n=1500 iken, %5 kayıp veri oranında RA yöntemi eksiksiz veri seti ile benzer sonuçlar üretmektedir. Örneklem büyük ve kayıp veri miktarı yüksek olduğunda, RA yöntemi ile tamamlanan veri setlerinde eksiksiz veri setinden önemli sapmalar meydana gelmektedir. Tablo 17’de RA yöntemi ile tamamlanan veri setlerinden kestirilen 1PLM’e uyumu gösteren AIC değerleri sunulmuştur.

Tablo 17. Rastgele Kayıp Mekanizmasında Regresyonla Atama Yöntemi ile Tamamlanan Veri Setlerinden Elde Edilen AIC Değerleri

Eksiksiz Veri Seti	n=500		n=1000		n=1500	
	AIC	$\Delta$ AIC	AIC	$\Delta$ AIC	AIC	$\Delta$ AIC
	9190.54		18004.66		27599.07	
%5	9188.19	2.34	17971.59	33.07	27562.08	36.99
%10	9146.59	43.94	17938.72	65.94	27473.39	125.68
%15	9132.45	58.08	17905.99	98.67	27430.57	168.5

Tablo 17 incelendiğinde, tüm örneklem büyüklüklerinde kayıp veri oranı arttıkça AIC değerlerinin düştüğü ve eksiksiz veri setinden uzaklaşarak 1PLM’e daha uyumlu hale geldiği görülmektedir. Tablo 18’de RA yöntemi ile tamamlanan veri setlerinden kestirilen BIC değerleri sunulmuştur.

Tablo 18. Rastgele Kayıp Mekanizmasında Regresyonla Atama Yöntemi ile Tamamlanan Veri Setlerinden Elde Edilen BIC Değerleri

Eksiksiz Veri Seti	n=500		n=1000		n=1500	
	BIC	$\Delta$ BIC	BIC	$\Delta$ BIC	BIC	$\Delta$ BIC
	9274.84		18102.81		27705.34	
%5	9272.48	2.35	18069.75	33.06	27668.34	37.00
%10	9230.88	43.95	18036.88	65.93	27579.66	125.68
%15	9248.62	26.21	18004.14	98.67	27536.83	168.51

Tablo 18 incelendiğinde, tüm örneklem büyüklüklerinde, RA yöntemi ile tamamlanan veri setlerinde kayıp veri oranı arttıkça, BIC kriterinin düştüğü ve dolayısıyla model veri uyumunun arttığı görülmektedir. %5 kayıp oranında eksiksiz veri setine en yakın kestirimlerin elde edilmiş olması, RA yönteminin iyi performans sergilediğini göstermektedir. Allison (2002) ve Van Ginkel (2007) de, kayıp veri oranı %5’in üzerine çıktığında kayıp veri baş etme yöntemlerinin performansının düşeceğini belirtmektedir. Burada da elde edilen bulgu, bu durumu desteklemektedir. RK kayıp veri mekanizmasında RA yöntemi yansız parametre tahminleri sunmaktadır (Baraldi ve Enders, 2010; McKnight ve diğ., 2007; Roth ve Switzer, 1995; Suraphee ve diğ., 2006). Elde edilen sonuç alanyazın ile uyumludur. Tablo 19’da BMA ile tamamlanan veri setlerinden kestirilen  $-2 \log \lambda$  değerleri sunulmuştur.

Tablo 19. Rastgele Kayıp Mekanizmasında Beklenti Maksimizasyon Algoritması Yöntemi ile Tamamlanan Veri Setlerinden Elde Edilen  $G^2$  Olabilirlik Oranı

Eksiksiz Veri Seti	n=500		n=1000		n=1500	
	$-2 \log \lambda$	$G^2$	$-2 \log \lambda$	$G^2$	$-2 \log \lambda$	$G^2$
	-4575.27		-8982.32		-13779.54	
%5	-4573.19	2.07	-8962.41	19.90	-13760.38	19.16
%10	-4557.04	22.22	-8947.24	35.07*	-13714.20	65.34*
%15	-4553.94	21.32	-8929.07	53.24*	-13690.30	89.24*

\* $p(\chi^2_{sd=19} > 30.144) < 0.05$

Tablo 19 incelendiğinde, BMA ile tamamlanan veri setleri ve eksiksiz veri seti arasındaki  $-2 \log \lambda$  değeri farkını gösteren  $G^2$  olabilirlik oranının manidarlığı test edildiğinde, kayıp veri oranı %5 iken,  $-2 \log \lambda$  farkının manidar olmadığı görülmektedir. Örneklem büyüklüğü  $n=500$  olduğunda tüm kayıp veri oranlarında eksiksiz veri seti ile farkın manidar olmadığı, dolayısıyla iyi performans sergilediği görülmektedir. Örneklem büyüklüğü  $n=1000$  ve  $n=1500$  olduğunda kayıp veri oranı %10 ve üzerine çıkınca, tamamlanan veri setleri ile eksiksiz veri seti arasında  $-2 \log \lambda$  farklarının manidar olduğu görülmektedir. BMA yöntemi kayıp veri oranı düşük olduğunda iyi performans sergilemektedir. Allison (2002) ve Van Ginkel (2007) kayıp veri oranı %5'in üzerine çıktığında kayıp veri baş etme yöntemlerinin performansının düşebileceğini belirtmektedir. BMA yönteminin %5 veri kayıplarının üzerinde düşük performans sergilemesi alanyazın ile uyumlu görünmektedir. Tablo 20'de BMA yöntemi ile tamamlanan veri setlerinden kestirilen AIC değerleri sunulmuştur.

Tablo 20. Rastgele Kayıp Mekanizmasında Beklenti Maksimizasyon Algoritması Yöntemi ile Tamamlanan Veri Setlerinden Elde Edilen AIC Değerleri

Eksiksiz Veri Seti	n=500		n=1000		n=1500	
	AIC	$\Delta$ AIC	AIC	$\Delta$ AIC	AIC	$\Delta$ AIC
	9190.54		18004.66		27599.07	
%5	9186.39	4.14	17964.84	39.82	27560.76	38.31
%10	9154.09	36.44	17934.49	70.17	27468.40	130.67
%15	9147.89	42.64	17898.15	106.51	27420.61	178.46

Tablo 20 incelendiğinde, tüm örneklem büyüklüklerinde kayıp veri oranı arttıkça, AIC değerinin düştüğü görülmektedir. Kayıp veri oranı arttıkça BMA yöntemi, AIC kriteri bakımından model veri uyumunu artırıcı bir performans sergilemektedir. Kayıp veri oranı sabit tutulup örneklem büyüklüğü artırıldığında ise, eksiksiz veri setinden kestirilen AIC değerinden uzaklaşmanın arttığı görülmektedir. Tablo 21'de BMA yöntemi ile tamamlanan veri setlerinden kestirilen BIC değerleri sunulmuştur.

Tablo 21. Rastgele Kayıp Mekanizmasında Beklenti Maksimizasyon Algoritması Yöntemi ile Tamamlanan Veri Setlerinden Elde Edilen BIC Değerleri

Eksiksiz Veri Seti	n=500		n=1000		n=1500	
	BIC	$\Delta$ BIC	BIC	$\Delta$ BIC	BIC	$\Delta$ BIC
	9274.84		18102.81		27705.34	
%5	9270.68	4.15	18062.99	39.82	27667.03	38.31
%10	9238.42	36.41	18032.64	70.17	27574.67	130.67
%15	9232.19	42.64	17996.31	106.5	27526.87	178.47

Tablo 21 incelendiğinde, tüm örneklerde kayıp veri oranı arttıkça, BIC değerinin azaldığı görülmektedir. Eksiksiz veri setinden elde edilen BIC değeri ile kıyaslandığında, BMA yönteminin veri setini IPLM'e daha uyumlu bir forma dönüştürdüğü görülmektedir. Kayıp veri oranı sabit tutulup örneklem büyüklüğü artırıldığında, eksiksiz veri setinden kestirilen BIC değeri ile olan farkın da arttığı görülmektedir. Bir diğer ifade ile örneklem büyüklüğü arttıkça, BMA yöntemi BIC değerinde sapma miktarını artırma eğilimi göstermektedir. Bu durum BMA'nın küçük örneklerde daha iyi performans sergilediği sonucuna götürmektedir. Tablo 22'de RK mekanizmasında model ile uyum gösteren madde sayısı sunulmuştur.



Tablo 22. Rastgele Kayıp Mekanizmasında Model ile Uyum Gösteren Madde Sayısı

	n=500			n=1000				n=1500				
	Eksiksiz Veri Seti	%5	%10	%15	Eksiksiz Veri Seti	%5	%10	%15	Eksiksiz Veri Seti	%5	%10	%15
LBS		19	20	18		17	16	19		8	12	9
RA		18	18	14		16	12	14		15	16	7
BMA	19	19	17	12	14	15	12	12	15	15	14	12

Tablo 22 incelendiğinde, eksiksiz veri setinde IPLM ile uyum gösteren madde sayılarının örneklem büyüklüğü n=500 olduğunda 19, n=1000 olduğunda 14 ve n=1500 olduğunda ise 15 olduğu görülmektedir. Kayıp veriler LBS yöntemi ile tamamlandığında, örneklem büyüklüğü n=500 olduğunda %5 kayıpta 19, %20 kayıpta 20 ve %15 kayıpta 18 madde uyumlu olacak şekilde etki gösterdiği görülmektedir. Örneklem büyüklüğü n=1000 olduğunda eksiksiz veri setinden daha fazla sayıda madde uyumlu görünürken, örneklem büyüklüğü n=1500 olduğunda, eksiksiz veri setinden daha az sayıda maddeyi uyumlu görmektedir. Bu sonuçlara bağlı olarak RK mekanizmasında LBS yönteminin, eksiksiz veri seti ile uyumlu bir performans göstermediği ve tutarlı olarak “düşük” ya da “yüksek” kestirimlere neden olduğu şeklinde bir genellemeye varılamayacağı görülmektedir. Madde Tepki Kuramı’nda model parametrelerinin kestiriminde ve model veri uyumunda örneklem büyüklüğü önemli bir etkidir (Bolt ve Lall, 2003; Seungho-Yang, 2007). Madde model uyumunun belirlenmesinde kullanılan  $\chi^2$  istatistiğinin örneklem büyüklüğünden etkilendiği göz önünde bulundurulduğunda (Hambleton ve diğ.,1991), LBS yönteminin kullanılmasının örneklemini küçülterek  $\chi^2$ ’nin manidar hale gelmesine yol açması beklenir. Ancak buradaki temel sorun kaybın TRK olmamasıdır. LBS yönteminin TRK koşulunda kullanılması durumunda, veri setinde kalan gözlemler eldeki verinin rastgele bir alt örneklemini olacağından, parametre kestirimlerinde yanlılığa yol açmayacaktır. Ancak TRK şartının sağlanmadığı durumlarda, LBS yönteminin kullanılmasıyla elde edilen alt örneklem, tüm gözlemlerin rastgele bir alt örneklemini olmayacağından, parametre kestirimleri yanlılık içerecektir (Allison, 2003; Enders, 2010; McKnight ve diğ., 2007; Rosenthal ve Rosnow, 2008;). LBS yönteminin kullanılması, yanıtlarında kayıp bulunan ve yanıtlarında hiç kayıp bulunmayan kişiler arasında sistematik farklılığın olduğu ölçüde yanlılığa da neden olabilecek ve buna bağlı olarak da kayıp veri içeren kişilerin veri setinden çıkarılmasıyla veri setinde kalan kişilerin evreni temsiliyeti düşecektir (Montalto ve Sung, 1996; Schafer, 1997). Dolayısıyla veri setinde RK mekanizmasında kayıp olduğu durumda, LBS yöntemi eksiksiz veri setini yansıtır bir performans sergilememektedir.

Regresyonla atama yöntemi ile kestirilen madde model uyumunda, uyumlu madde sayıları incelendiğinde bu yöntemin genel olarak eksiksiz veri setinden daha az sayıda maddeyi uyumlu gösterecek atamalar yaptığı görülmektedir. tüm örneklem büyüklüklerinde kayıp veri oranı arttıkça kestirilen değerlerin eksiksiz veri setinden uzaklaştığı bir diğer ifadeyle yöntemin performansı düşmektedir. Alanyazında RA yöntemi, bazı durumlarda eksiksiz veri setlerinden elde edilen değerlerden farklı değerler olsa da, bu değerlerin genel olarak eksiksiz veri setlerinden elde edilen değerlere çok benzer ya da aynı olacağı belirtilmektedir (Roth ve Switzer, 1995; Suraphee ve diğ., 2006). Elde edilen sonuç da bu bilgi ile tutarlıdır.

Beklenti maksimizasyon algoritması ile atama yapıldığında, tüm örneklem büyüklüklerinde kayıp veri oranı arttıkça, uyumlu madde sayısının azaldığı görülmektedir. Özellikle kayıp veri oranı %15 olduğunda model ile uyumlu madde sayısında önemli azalışların olduğu göze çarpmaktadır. Bu nedenle rastgele kayıp mekanizmasına sahip kayıp veri ile karşılaşıldığı durumda kayıp veri oranı yüksek (%15) ise bu yöntemi kullanmaktan kaçınılması önerilebilir. BMA yöntemi özellikle RK varsayımının sağlandığı durumlarda daha iyi tahmin değerleri üretmektedir (Allison, 2003; Baraldi ve Enders, 2010). Ancak elde edilen bulgular madde model açısından bunu desteklememektedir.

## SONUÇLAR ve TARTIŞMA

Bu çalışmadan elde edilen bulgular,  $-2 \log \lambda$  kriterine bakılarak model veri uyumuna karar vermek istendiğinde, kayıp veri baş etme yöntemlerinden LBS yönteminin diğer yöntemlere göre eksiksiz veri setinden daha uzak kestirime yol açtığı, BMA yönteminin özellikle örneklem büyüklüğü  $n=500$  olduğunda tüm kayıp veri oranlarında eksiksiz veri setinden kestirilen  $-2 \log \lambda$  değerine istatistiksel olarak en yakın sonuçları verecek atamalar yaptığı sonucuna götürmektedir.

Model veri uyumunun belirlenmesinde kullanılan AIC ve BIC kriterleri bakımından kayıp veri baş etme yöntemlerinin performansları karşılaştırıldığında; LBS yönteminin diğer yöntemlere göre, eksiksiz veri setinden kestirilen AIC ve BIC değerlerinden daha büyük sapmalara neden olduğu göze çarpmaktadır. AIC ve BIC kriterleri bakımından eksiksiz veri setinden yakın kestirimlerin RA ve BMA yöntemleri ile tamamlanan veri setlerinden elde edildiği görülmektedir. Ancak örneklem büyüklüğü ve kayıp veri oranı arttıkça, bu yöntemlerdeki sapmanın da artış gösterdiği söylenebilir. Bu durumda LBS, RA ve BMA yöntemlerinin tümünün örnekleme heterojenleştirecek performans sergiledikleri belirlenmiştir.  $\chi^2$ 'nin hesaplanmasında, evren ortalaması ile bağımsız değişken ortalaması arasındaki farkın, evren varyansına oranlanması söz konusudur (Chao, 1968). Buradaki oranlama nedeniyle varyans ile  $\chi^2$  ters orantılıdır buna bağlı olarak da varyansın büyümesi  $\chi^2$ 'yi düşürürken, varyansın azalması  $\chi^2$ 'yi artıracaktır. Dolayısıyla LBS, BMA ve RA yöntemleri, varyansı artırıcı, bir başka deyişle grubu heterojenleştirici performanslar sergiledikleri saptanmıştır.

Tamamen rastgele kayıp koşulunda, RA ve BMA yöntemleri ile tamamlanan veri setlerinde ise birbirine yakın sonuçlar elde edilmiştir. Model veri uyumu AIC kriteri açısından ele alındığında,  $-2 \log \lambda$ 'ya benzer olarak, eksiksiz veri setine en uzak kestirimlerin LBS yöntemi ile tamamlanan veri setlerinden elde edildiği belirlenmiştir. BMA ve RA yöntemleri ile tamamlanan veri setlerinden ise birbirine yakın sonuçlar elde edilmiştir. Model veri uyumu BIC kriteri açısından ele alındığında ise AIC ve  $-2 \log \lambda$ 'de olduğu gibi, eksiksiz veri setine en uzak kestirim yine LBS yöntemi ile tamamlanan veri setlerinden elde edildiği, RA, ÇA ve BMA yöntemleri ile tamamlanan veri setlerinde ise, birbirine yakın kestirimler elde edildiği belirlenmiştir.

Liste bazında silme yöntemi, tüm kayıp veri mekanizmalarında  $-2 \log \lambda$  kriteri açısından benzer performanslar göstermiştir. Eksiksiz veri seti ile tamamlanan veri seti arasındaki  $-2 \log \lambda$  farkını gösteren  $G^2$  değerleri incelendiğinde, bu oranın tüm kayıp veri mekanizmalarında birbirine çok yakın değerlerde olduğu görülmektedir. Örneğin; örneklem büyüklüğü  $n=1000$ , kayıp veri oranı %10 olduğunda  $G^2$  olabilirlik oranı TRK kayıp koşulunda yaklaşık 917, RK koşulunda yaklaşık 860, ROK koşulunda ise yaklaşık 899'dur. Bu değerlerin manidarlığı test edildiğinde, tüm kayıp veri mekanizmalarında eksiksiz veri seti ile arasındaki farklılığın manidar olduğu görülmektedir. Dolayısıyla LBS yöntemi, 1PLM'de model veri uyumu için  $-2 \log \lambda$  bakımından isabetsiz kestirimler yapılmasına yol açmaktadır. TRK ve RK koşullarında diğer yöntemlerle kıyaslandığında ise eksiksiz veri setinden en uzak kestirimleri sunduğu ve bu nedenle de en düşük performansı sergilediği yönünde karara varıldığı söylenebilir.

Liste bazında silme yönteminin model veri uyumu üzerindeki etkisini AIC ve BIC kriterleri bakımından değerlendirildiğinde,  $-2 \log \lambda$ 'de olduğu gibi, tüm kayıp veri mekanizmalarında benzer performansı sergilediği görülmektedir. Eksiksiz veri setinden uzaklaşma miktarının diğer kayıp veri baş etme yöntemlerine göre fazla olmasının, LBS yönteminin model veri uyumunda da AIC ve BIC kriterleri bakımından düşük performans sergilemesine neden olduğu belirtilebilir. LBS yönteminin performansı diğer kayıp veri baş etme yöntemleri ile karşılaştırıldığında ise, en düşük performansı TRK ve RK mekanizmalarında sergilediği ifade edilebilir.

Regresyonla atama yönteminin 1PLM'de model veri uyumu üzerindeki performansı incelendiğinde,  $-2 \log \lambda$  kriteri açısından en iyi performansı RK mekanizmasında gösterdiği saptanmıştır. Eksiksiz veri seti ile tamamlanan veri seti arasındaki  $-2 \log \lambda$  farkını gösteren  $G^2$  değerleri incelendiğinde, bu oranın tüm kayıp veri mekanizmalarında birbirinden farklı değerlerde olduğu görülmektedir. RA yöntemi 1PLM'de model veri uyumu için  $-2 \log \lambda$  bakımından RK mekanizmasında isabetli kestirimler yapılmasını sağlamaktadır. RK ve TRK mekanizması olduğu durumda, RA ve diğer yöntemlerin  $-2 \log \lambda$  üzerindeki performansları karşılaştırıldığında, RK durumunda RA yönteminin,

TRK koşulunda ise BMA yönteminin daha iyi performans sunduğu sonucuna ulaşılmıştır. Burada performansın iyi olmasından kastedilen, eksiksiz veri setine en yakın kestirimlerin elde edilmiş olmasıdır. RA yönteminin model veri uyumu üzerindeki etkisi, AIC ve BIC kriterleri bakımından değerlendirildiğinde,  $-2 \log \lambda$  ile benzer bir tablo ile karşılaşılmıştır. Eksiksiz veri setinden uzaklaşma miktarının kayıp veri baş etme yöntemlerinin performansını gösterdiği göz önünde bulundurulduğunda, RA yöntemi IPLM’de model veri uyumu için AIC ve BIC bakımından RK mekanizmasında oldukça isabetli kestirimler yaparken, TRK mekanizmasında performansının düşük olduğu belirlenmiştir. RK ve TRK mekanizmalarının olduğu durumlarda, RA ve diğer yöntemlerin AIC ve BIC üzerindeki performansları karşılaştırıldığında, RK durumunda RA yönteminin oldukça iyi ve BMA ile benzer performans sergilediği, TRK koşulunda ise BMA yöntemleri ile benzer sonuçlar sunduğu söylenebilir.

Beklenti maksimizasyon algoritması ile atama yönteminin IPLM’de model veri uyumu üzerindeki performansı incelendiğinde,  $-2 \log \lambda$  kriteri açısından en iyi performansı RK mekanizmasında gösterdiği sonucuna ulaşılmıştır. Eksiksiz veri seti ile tamamlanan veri seti arasındaki  $-2 \log \lambda$  farkını gösteren  $G^2$  değerleri incelendiğinde, bu oranın RK mekanizmasında düşük, TRK mekanizmasında ise daha yüksek olduğu görülmüştür. BMA yöntemi IPLM’de model veri uyumu için  $-2 \log \lambda$  bakımından RK mekanizmasında bazı koşullarda isabetli kestirimler sunmaktadır. Eksiksiz veri setinden sapmaların yönüne bakıldığında ise TRK ve RK mekanizmalarında model veri uyumunu artırma yönünde performans gösterirken ROK mekanizmasında model veri uyumunu düşürme yönünde performans sergilediği belirlenmiştir. RK mekanizması olduğu durumda BMA ve diğer yöntemlerin  $-2 \log \lambda$  üzerindeki performansları incelendiğinde, RA ile benzer kestirime yol açtığı görülmüştür. RK mekanizmasında özellikle örneklem büyüklüğü  $n=500$  ve kayıp veri oranı %5 olduğunda en iyi performansı gösteren yöntem olarak ön plana çıkmaktadır. TRK durumunda da RA yöntemleri ile benzer ve iyi sonuçlar sunduğu belirlenmiştir. BMA ile atama yönteminin model veri uyumu üzerindeki etkisi AIC ve BIC kriterleri bakımından değerlendirildiğinde, diğer yöntemlerde olduğu gibi  $-2 \log \lambda$ ’de karşılaşılan sonuç ile karşılaşılmaktadır. BMA yöntemi IPLM’de model veri uyumu için AIC ve BIC bakımından RK mekanizmasında oldukça isabetli kestirimler sunmuştur. TRK mekanizmasında eksiksiz veri setine en yakın kestirimleri veren yöntem olmamakla birlikte, eksiksiz veri setine yakın kestirimler üretmiştir. RK ve TRK mekanizması olduğu durumda BMA ve diğer yöntemlerin AIC ve BIC üzerindeki performansları incelendiğinde, her iki durumda da BMA yönteminin oldukça iyi ve RA yöntemine benzer performans sunduğu ifade edilebilir.

Model veri uyumu söz konusu olduğunda TRK mekanizmasında LBS, BMA yöntemlerinin iyi, RA yönteminin kısmen iyi, RK mekanizmasında tüm yöntemlerin iyi performans sergilediği saptanmıştır. Buna ek olarak tüm kayıp veri mekanizmalarında, tüm kayıp veri baş etme yöntemlerinin, kayıp veri oranı arttıkça, model veri uyumu kestirimleri bakımından eksiksiz veri setinden uzaklaştığı ve performanslarının düştüğü vurgulanabilir.

Liste bazında silme yöntem ile tamamlanan kayıp verilerin ardından TRK ve RK mekanizmalarında IPLM ile uyum gösteren madde sayıları, eksiksiz veri seti ile uyumu referans alınarak değerlendirildiğinde, TRK mekanizmasında olması gerekenden fazla sayıda maddeyi model ile uyumlu gösterme eğilimi taşıdığı görülmüştür. LBS yönteminin hiçbir kayıp veri mekanizmasında eksiksiz veri seti ile uyumlu sonuçlar üretmediği, dolayısıyla LBS yönteminin IPLM’de madde model uyumu kestirimi söz konusu olduğunda kayıp veri baş etme yöntemi olarak tercih edilmemesi gerektiği sonucuna ulaşılmıştır.

Regresyonla atama yöntemi, madde model uyumu söz konusu olduğunda TRK mekanizması altında genel olarak daha az sayıda maddeyi uyumlu gösterecek yönde performans sergilediği belirtilebilir. TRK mekanizmasında madde sayısındaki azalma kayıp veri oranı ile ilişkilidir. Bir diğer deyişle kayıp veri miktarı arttıkça, uyumlu madde sayısı azalmaktadır. Bu durum RA yönteminin kayıp veri oranından etkilendiğini göstermektedir. RK mekanizması genel olarak daha az sayıda maddeyi uyumlu gösterme eğiliminde olmakla birlikte, kayıp veri oranı %15 olduğunda, tüm örneklem büyüklüklerinde uyumlu madde sayısında önemli düşümlere yol açacak atamalar yaptığı belirtilebilir.

Beklenti maksimizasyon algoritması yönteminin TRK ve RK mekanizmalarında IPLM’de madde model uyumuna etkisi değerlendirildiğinde, RK mekanizmasında tüm örneklem büyüklüklerinde ve

%5 kayıp oranında, eksiksiz veri setine en yakın kestirimleri sunmuştur. Dolayısıyla veri setinde RK mekanizması olduğunda, kayıp veri oranı düşük ise kullanılabilir bir yöntemdir. Kayıp veri oranı arttıkça, uyumlu madde sayısı azalmakta ve dolayısıyla BMA yönteminin RK mekanizmasında performansı düşmektedir. TRK mekanizmasında performansı incelediğinde ise, genel olarak tüm örneklem büyüklüklerinde eksiksiz veri setinden daha az sayıda maddeyi uyumlu kılmaktadır. Kayıp veri oranı %15 olduğunda performansında önemli düşüşler yaşanmaktadır. Dolayısıyla TRK mekanizmasında kayıp veri oranı yüksek olduğunda, BMA'nın kullanılmaması gereken bir kayıp veri baş etme yöntemi olduğu belirtilebilir.

Veri setinde TRK mekanizmasında kayıp veriler olduğunda MTK 1 PLM'e göre modelleme yapılmak istendiğinde, model veri uyumunun en iyi şekilde tahminlenebilmesi için; LBS, RA ve BMA yöntemlerinin kullanılması önerilebilir. Veri setinde RK mekanizmasında kayıp veriler olduğunda ve MTK 1PLM'e göre modelleme yapılmak istendiğinde ise, model veri uyumunun en iyi şekilde tahmin edilebilmesi için, BMA ve RA yöntemlerinin kullanımı önerilebilir. Veri setinde TRK mekanizmasına sahip kayıp veri olması durumunda 1PLM'de kayıp veri baş etme yöntemlerinin performansları göz önünde bulundurulduğunda, madde model uyumu için eğer örneklem küçükse (örneğin n=500) BMA yöntemi tercih edilebilir. Veri setinde RK mekanizmasına sahip kayıp veri olması durumunda 1PLM'de kayıp veri baş etme yöntemlerinin performansları göz önünde bulundurulduğunda madde model uyumu söz konusu ise, BMA yöntemi düşük kayıp veri oranları için kullanılabilir bir yöntemdir.

#### KAYNAKÇA

- Ackerman, T.A. (2005). Multidimensional item response theory modeling. In A. Maydeu-Olivares & J.J. McArdle (Eds.). *Contemporary psychometrics* (3–26) Mahwah, NJ: Lawrence Erlbaum Associates.
- Ackerman, T.A., Gierl, M.J., & Walker, C.M. (2003). *Using multidimensional item response theory to evaluate educational and psychological tests*. *Educational Measurement: Issues and Practice*, 22(3), 37-51.
- Adams, R. (2005). *Reliability as a measurement design effect*. *Studies in Educational Evaluation*, 31, 162–172. DOI: 10.1016/j.stueduc.2005.05.008
- Aiken, L. R. (1995). *Psychological testing and assessment*. Boston: Allyn and Bacon
- Allison, P. (2002). *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
- Baker, F. B. (1998). An investigation of the item parameter recovery of a Gibbs sampling procedure. *Applied Psychological Measurement*, 22, 153-169.
- Baraldi, A. N., & Enders, C. K. (2010). An introduction to modern missing data analysis. *Journal of School Psychology*, 48, 5–37.
- Bauer, D. J., & Curran, P. J. (2003). Distributional assumptions of growth mixture models: Implications for overextraction of latent trajectory classes. *Psychological Methods*, 8, 338-363.
- Baykul, Y. (2000). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulanması*. Ankara: ÖSYM.
- Blozis, S. A., Conger K. J., & Harring, J. R. (2007). Nonlinear latent curve models for multivariate longitudinal data. *International Journal of Behavioral Development: Special Issue on Longitudinal Modeling of Developmental Processes*, 31, 340-346.
- Bock, R. D. (1997). A brief history of Item Response Theory. *Educational Measurement: Issues and Practice*, 16, 21-33.
- Bolt, D. M., & Lall, V. F. (2003). Estimation of compensatory and non-compensatory multidimensional Item Response models using Markov Chain Monte Carlo. *Applied Psychological Measurement*, 27(6), 395-414. Web: <http://apm.sagepub.com> adresinden 02.10.2015'de alınmıştır.
- Cao, Y., & Poh, K. L. (2006). An accurate and robust missing value estimation for microarray data: Least absolute deviation imputation, *Proceedings of the 5th International Conference on Machine Learning and Applications (ICMLA'06)*, 157-161.
- Chang, Y. W. (1992). A comparison of unidimensional and multidimensional IRT approaches to test information in a test battery (Unpublished doctoral dissertation, University of Minnesota). <http://files.eric.ed.gov/fulltext/ED344940.pdf> adresinden 20.01.2015'de alınmıştır.
- Chen, S. F., Wang, S., & Chen, Y. C. (2011). A simulation study using EFA and CFA programs based the impact of missing data on test dimensionality. *Expert Systems with Applications*, 39, 4026–4031.

- Collins, L. M., & Lanza, S. T. (2010). *Latent class and latent transition analysis with applications in the social, behavioral, and health sciences*. Hoboken, New Jersey: John Wiley & Sons, Inc.
- Crocker, L., & Algina, J. (1986). *Introduction to classical & modern test theory*. Newyork: Holt, Rinehart and Winston
- Cronbach, J. L. (1990). *Essentials of psychological testing* (5. Ed). Harper Collins Publishers, Inc
- Çelen, Ü. (2008). Klasik Test Kuramı ve Madde Tepki Kuramına dayalı olarak geliştirilen iki testin psikometrik özelliklerinin karşılaştırılması. *İlköğretim Online*, 7(3), 758-768.
- Çokluk, Ö. ve Kayrı, M. (2011). Kayıp değerlere yaklaşık değer atama yöntemlerinin ölçme araçlarının geçerlik ve güvenilirliği üzerindeki etkisi. *Kuram ve Uygulamada Eğitim Bilimleri*, 11(1), 289-309.
- DeMars, C. (2010). *Item response theory: Understanding statistics measurement*. London: Oxford Press.
- Demir, E. (2013). *Kayıp verilerin varlığında iki kategorili puanlanan maddelerden oluşan testlerin psikometrik özelliklerinin incelenmesi* (Yayımlanmamış doktora tezi, Ankara Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara). <http://acikarsiv.ankara.edu.tr/> adresinden 08.10.2016 tarihinde alınmıştır.
- Demir, E. ve Parlak, B. (2012). Türkiye’de eğitim araştırmalarında kayıp veri sorunu. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 3(1), 230-241
- Dias, J. G. (2006). *Latent class analysis and model selection*. In M. R. Kruse, C. Borgelt, A. Nürberger, & W. Gaul, *From data and information analysis to knowledge engineering* (pp. 95-102). Berlin: Springer-Verlag.
- Drasgow, F., Levine, M., Tsien, S., Williams, B., & Mead, A. (1995). Fitting polytomous Item Response Theory models to multiple-choice tests. *Applied Psychological Measurement*, 19(2), 143-165. doi:10.1177/014662169501900203
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, 43(6), 1428-1446. Doi:10.1037/0012-1649.43.6.1428
- Doğanay Erdoğan, B., (2012). *Çoklu atama yöntemlerinin Rasch modelleri için performansının benzetim çalışması ile incelenmesi* (Yayımlanmamış doktora tezi, Ankara Üniversitesi Sağlık Bilimleri Enstitüsü, Ankara). <http://acikarsiv.ankara.edu.tr> adresinden 13.01.2015 tarihinde alınmıştır.
- Embretson, S. E. (1999). Generating items during testing: Psychometric issues and models. *Psychometrika*, 64, 407-433.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Enders, C. K. (2010). *Applied missing data analysis*. (1st ed.). New York: The Guilford Publications, Inc.
- Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*. 8(3), 430-457.
- Enders, C. K. (2004). The impact of missing data on sample reliability estimates: Implications for reliability reporting practices. *Educational and Psychological Measurement*. 64(3), 419-436.
- Fiona, M. S., Heather, S., Hude, G., & William G. (2006). Dealing with missing data in a multi-question depression scale: A comparison of imputation methods. *BMC Medical Research Methodology* 6:57. DOI: 10.1186/1471-2288-6-57
- Furlow C. F., Fouladi R. T., Gagne P., & Whittaker T.vA. (2007). A Monte Carlo study of the impact of missing data and differential item functioning on theta estimates from two polytomous Rasch family models. *Journal of Applied Measurement*. 7(8),388-403
- Gao, F., & Chen, L. (2005). Bayesian or Non-Bayesian: A comparison study of item parameter estimation in the three-parameter logistic model. *Applied Measurement in Education*, 18(4), 351-380.
- Goldman, S. H., & Raju, N. S. (1986). Recovery of one- and two-parameter logistic item parameters: An empirical study. *Educational and Psychological Measurement*, 46(1),11-21.
- Hambleton R. K., Swaminathan H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: SAGE Publications, Inc.
- Hambleton, R. K. (1989). Principles and salected applications of Item Response Theory. R. L. Linn (Ed), *Educational Measurement* (pp.147-200). Washington, DC: American Council of Education.
- Hohensinn, C., & Kubinger, K. D. (2011). On the impact of missing values on the item fit and the model validness of the Rasch model. *Psychological Test and Assessment Modeling*, 53(3), 380-393.
- Hulin, C. L., Lissak, R. I., & Drasgow, F. (1982). Recovery of two and three-parameter logistic item characteristic curves: A Monte Carlo study. *Applied Psychological Measurement*, 6, 249-260.
- Hurvich, C., & Tsai, C. (1989). Regression and time series model selection in small samples. *Biometrika*, 76, 297-307

- Kankaras, M., Vermunt, J. K., & Moors, G. (2011). Measurement equivalence of ordinal items. A comparison of factor analytic, Item Response Theory and Latent Class Approaches. *Sociological Methods & Research*, 40(2), 279-310.
- Karasar, N. (2007). *Bilimsel araştırma yöntemi: Kavramlar, ilkeler, teknikler* (17. Baskı). Ankara: Nobel.
- Köse, İ. A. (2014). The effect of missing data handling methods on goodness of fit indices in confirmatory factor analysis. *Educational Research and Reviews*, 9(8), 208-215. Doi:10.5897/ERR2014.1709
- Lin, T. H. (2006). A comparison of model selection indices for Nested Latent Class Models. *Monte Carlo Methods and Applications*, 12(3-4), 239-259.
- Lin, T. H. (2012). Model selection information criteria in Latent Class Models with missing data and contingency question. *Journal of Statistical Computation and Simulation*, 84(1),159-170 doi:10.1080/00949655.2012.698621.
- Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: Wiley.
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading MA: Addison-Wesley.
- Lord, F. M. (1968). An Analysis of the Verbal Scholastic Aptitude Test Using Birnbaum's Three-parameter Logistic Model. *Educational and Psychological Measurement*, 28, 989-1020.
- Lord, F. (1980). *Applications of Item Response Theory to practical testing problems*. New Jersey: Lawrence Erlbaum Associates Publishers.
- McKnight, P. E., McKnight, K. M., Sidani, S., & Figueredo, A. J. (2007). *Missing data: A gentle introduction*. New York: The Guilford Publications, Inc.
- Mislevy, R. J., & Bock, R. D. (1990). *BILOG 3- item analysis and test scoring with binary logistic models* (2nd Ed). Mooresville: Scientific Software Inc.
- Montalto, C. P., & Sung, J. (1996). Multiple imputation in the 1992 survey of consumer finances. *Financial Counseling and Planning*, 7, 133 – 141.
- Morren, M., Gelissen, J., & Vermunt, J. K. (2011). Dealing with extreme response style in cross-cultural research: A restricted Latent Class factor analysis. *Sociological Methodology*, 1(1), 13-47. Doi: 10.1111/j.1467-9531.2011.01238
- Nylund, K. L., Asparouhov, T., & Muthen, B. O. (2007). Deciding on the number of classes in Latent Class analysis and Growth Mixture Modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, 14(4), 535-569.
- Orlando, M., & Thissen, D. (2000). Likelihood-based item-fit indices for dichotomous Item Response Theory models. *Applied Psychological Measurement*, 24(1), 24-50
- R Development Core Team (2011). R: A language and environment for statistical computing, a foundation for statistical computing. Vienna, Austria, ISBN 3-900051-0-70, Erişim:[<http://www.R-project.org>].
- Reeve, B. B. (2002). *An introduction to modern measurement theory*. <http://appliedresearch.cancer.gov/areas/cognitive/immmt.pdf> adresinden 21.12.2015 tarihinde alınmıştır.
- Reid, C. A., Kolakowsky-Hayner, S. A., Lewis, A. N., & Armstrong, A. J. (2007). Modern psychometric methodology: Applications of Item Response Theory. *Rehabilitation Counselling Bulletin*, 50(3), 177-178.
- Reise, S. P., Ainsworth, A. T., & Haviland, M. G. (2005). Item Response Theory. Fundamentals, applications, and promise in psychological research. *Current Directions in Psychological Science*, 14(2), 95-101.
- Rosenthal, R., & Rosnow, R. (2008). *Essentials of behavioral research: Methods and data analysis* (3rd Ed). Boston : McGraw-Hill.
- Roth, P. L., & Switzer, F. S. (1995). A Monte Carlo analysis of missing data techniques in a HRM setting. *Journal of Management*, 21(5), 1003-1023.
- Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63, 581-592.
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: John Wiley & Sons, Inc.
- Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. New York: Chapman & Hall/Crc.
- Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7(2), 147 – 177.
- Seungho Yang, M. A. (2007). *A comparison of unidimensional and multidimensional Rasch models using parameter estimates and fit indices when assumption of unidimensionality is violated* (Unpublished Doctoral Dissertation, The Ohio State University). <https://etd.ohiolink.edu> adresinden alınmıştır.
- Sijtsma, K. & van der Ark, L. (2003). Investigation and treatment of missing item scores in test and questionnaire data. *Multivariate Behavioral Research*, 38(4), 505-528.
- Suraphee, S., Raksmanee, C., Busaba, J., Chaisorn, C., & Nakornthai, W. (2006). A comparison of estimation methods for missing data in Multiple Linear Regression with two independent variables. *Thailand Statistician*, 4, 13-26.
- Thissen, D., & Wainer, H. (1982). Some standart errors in Item Response Theory. *Psychometrika*, 47, 397-412.

- Turgut, F. (1992). *Eğitimde ölçme ve değerlendirme metotları*. Ankara: Yargı.
- Van Ginkel, J. R. (2007). *Multiple imputation for incomplete test, questionnaire, and survey data*. (Unpublished Doctoral Dissertation, Tilburg University). <https://pure.uvt.nl> adresinden 21.09.2016 tarihinde alınmıştır.
- Vrieze, S. I. (2012). Model selection and psychological theory: A discussion of the differences between the Akaike Information Criterion and the Bayesian Information Criterion. *Psychological Methods*, 17(2), 228-243.
- Wainer, H., & Thissen, D. (2001). True score theory: The traditional method. D. Thissen, & H. Wainer (Ed). *Test scoring*. New Jersey: Lawrence Erlbaum Associates Publishers.
- Wilson, M. (2005). *Constructing measures: An item response modeling approach*. New Jersey: Lawrence Erlbaum Associates.
- Yang, C., & Yang, C. (2007). Separating Latent Classes by information criteria. *Journal of Classification*, 24, 183-203.
- Yen, W. M. (1987). A comparison of the efficiency and accuracy of BILOG and LOGIST. *Psychometrika*, 52, 275-291.
- Yu, H. C. (2013). A simple guide to the Item Response Theory (IRT) and Rasch Modeling. <http://www.creative-wisdom.com/computer/sas/IRT.pdf> adresinden 18.09.2016 tarihinde alınmıştır.
- Yurdugül, H. (2006). Paralel, eşdeğer ve konjenetik ölçmelerde güvenilirlik katsayılarının karşılaştırılması. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 39(1), 15-37.

## EXTENDED ABSTRACT

### Introduction

The measured properties in education and psychology are mostly latent variables that do not allow direct observation. Various test theories have been developed to account for latent individual properties with the help of observable variables and those properties have been measured with instruments that have been developed based on the theories. One of the most significant theories developed is Item Response Theory (IRT). IRT is based on responses of test takers to items. It grounds on expressing the relationship between ability levels of individuals associated with the measured property and responses of test takers with the help of a mathematical function (Embretson and Reise, 2000; Hambleton and Swaminathan, 1989).

Item Response Theory claims the following: tests and item parameters must be independent from the group under application and individual levels of ability ( $\theta$ ) must be independent from the measurement instrument applied. Nevertheless, when the chosen model does not fit to the data or when there is low model data fit, the benefits offered by the theory will not be utilized in terms of items and ability parameters (Hambleton et al., 1991; Orlando and Thissen, 2000). If model data fit in a model is poor, item statistics and item information functions will be misleading (Ackerman, 2005; Chang, 1992; Hambleton et al., 1991).

In IRT, the main criterion to be considered when deciding a model to be employed through model data fit is choosing the model to present maximum individual data with minimum error in estimation. Thus, most probable estimations of latent features under measurement are likely with the help of individual responses. However, reaching latent variables that are attempted to be measured with tests is likely with individual responses to test items. In this context, when individuals avoid responses to items, skip items or give no response for any reason, the greatest handicap in implications will appear (Hohensinn and Kubinger, 2011).

In case of missing data issues in research, it is not appropriate to think that the best fitting model to all data has been decided or in other words, problems will arise in the process of choosing the accurate model for data. Deficiencies in response patterns will considerably hinder ability estimations. The study aims at examining the effects of missing data handling techniques on model data fit and item model fit of IRT one parameter logistic model, taking different sample sizes and missing data conditions.

### **Method**

The study is a simulation that attempts to reveal the effect of various missing data handling techniques on model data fit in Item Response Theory one parameter logistic model, using artificial (simulative) data sets. At the same time, it could be suggested that it has the quality of fundamental research as it aims at contributing new findings to the available theoretical knowledge (Karasar, 2007).

In data generation, missing data generation R program was employed at a desired level in accordance with missing data mechanisms in the generated complete data sets (R Development Core Team, 2011). The number of items for the generated data set was 20 and the number of response category for each item was two (1-0 scoring). The number of individuals was 500, 1000 and 1500 and the data were generated based on IRT one parameter logistic model. Item difficulty in the generated data ( $b$ ) was restricted from minimum -2 logit to maximum +2 logit and item discrimination was fixed as 1.5.

In the generated data sets, missing data at 5%, 10% and 15% were obtained to provide missing at complete random (MCR) and missing at random (MR) mechanisms. Missing data were generated in MCR mechanism whereas pores at pre-defined ratios were deleted in the samples of 500, 1000 and 1500 individuals, independently from any variables. In order to generate MR mechanism, another variable was randomly defined at a three level sequencing scale which was different from the measured feature, but still belonged to the measured individuals. Then, missing data were generated respectively at Level 1 (20%), Level 2 (30%), and Level 3 (50%). The generated missing data were resolved, using expectation maximization algorithm, regression imputation, and listwise deletion method (LM).

In the study, model data fit was estimated first in the generated complete data sets. Then, the same predictions were applied in the complete data sets and resulting values obtained from the complete data set were taken as reference and compared to the above mentioned values.  $-2 \log \lambda$ , AIC, and BIC values were used to evaluate the model data fit.  $\chi^2$  item fit approach was used to decide item model fit.

### **Results and Discussion**

It is seen that data sets which are completed with MR and EMA methods give the closest estimations to the complete data set in terms of  $-2 \log \lambda$ , AIC, and BIC criteria that are employed to decide model data fit. However, as sample sizes and missing data amount increase, deviation in these methods could be suggested as increasing. In this case, all LM, MR, and EMA methods were found to have performance to heterogenize the sample. Under missing at complete random (MCR) condition, similar results were obtained in data sets completed with missing at random (MR) and expectation maximization algorithm (EMA) methods. When compared to other methods, listwise deletion method under MCR and MR conditions could be decided to present the most improbable estimations from the complete data set and thus to have the lowest performance.

When there is missing data in MCR mechanism of the data set, and modelling according to IRT one parameter logistic method is desired, LM, MR and EMA methods were found eligible for the best model data fit estimation; when there are missing values in MR mechanism, and modelling according to IRT one parameter logistic method is desired, EMA and MR were found eligible for the best model data fit estimation. In case of missing data in data sets with MCR and MR mechanisms, EMA method could be preferred for item model fit if sample is small.