

# Kayıp Veri ile Baş Etme Yöntemlerinin Madde Parametrelerine Etkisinin İncelenmesi\*

## Examination the Effect of Missing Data Techniques of Item Parameters

Ayfer SAYIN \*\*

Alperen YANDI \*\*\*

Esra OYAR \*\*\*\*

### Öz

Bu çalışmada madde bazında kayıp veri oranlarının bulunduğu farklı örneklem büyüklüğündeki verilere ait madde ve test parametrelerinin kayıp veri ile baş etme yöntemlerinden nasıl etkilendiğini belirlemek amaçlanmıştır. PISA 2015 uygulamasına katılan ve çalışma içerisinde yer alan “hırs algısı” ölçeğine cevap veren 5073 öğrenci içerisinde rastgele seçilen 500, 1000 ve 2500 öğrenci, araştırmanın çalışma grubunu oluşturmaktadır. Öncelikle her bir veri setinde normallik, tek boyutluluk, yerel bağımsızlık ve model-veri uyumu varsayımları incelenmiştir. Ölçekte 5 madde yer almaktadır ve kayıp veriler madde bazında oluşturulmuştur. Bu doğrultuda tamamen rastsal olacak şekilde her bir maddeden sırasıyla %5, %10, %15 ve %20’lik kayıp veriler oluşturulmuş, ölçek maddelerinden birinde ise hiç kayıp veri olmayacak şekilde analizler gerçekleştirilmiştir. Kayıp verilerin tamamen rastsal dağılım gösterdiği belirlendikten sonra öncelikle tam ve eksik verilerle; daha sonra silme, ortalama atama, yakın noktalar ortalama ataması, yakın noktalar medyan ataması, doğrusal değerlendirme, noktada doğrusal eğilim, regresyon atama ve beklenti maksimizasyonu yöntemleri sonucunda elde edilen tam veri setleri ile hesaplamalar gerçekleştirilmiştir. Hesaplama sürecinde betimsel istatistikler ve Cronbach-alfa güvenilirlik katsayısı; ardından Madde Tepki Kuramına dayalı Aşamalı Tepki Modeline göre ayırıcılık ve güçlük indeksleri ile marjinal güvenilirlik katsayısı hesaplanmıştır. Araştırma sonucunda madde ve test parametrelerinin eksik veriden ve kayıp veri ile baş etme yöntemlerinden etkilendiği; tam veri setine en yakın kestirimi sunan sonuçların doğrusal değerlendirme yöntemi ile elde edildiği belirlenmiştir.

*Anahtar Kelimeler:* Kayıp veri, tamamen rastsal kayıp veri, madde tepki kuramı, aşamalı tepki modeli

### Abstract

In this study, the aim is to determine how the item and test parameters affect the missing data techniques for different sample sizes and different items with different missing data rates. 500, 100 and 2500 students randomly selected from the 5073 students who participated in the PISA 2015 study and responded to the "ambition perception" scale included in the study constitute the study group of the research. First of all, the assumptions of normality, unidimensionality, local independence and model-data fit were examined for each data set. Afterwards, 5%, 10%, 15%, and 20% missing data were formed for four out of five items and there was no missing data in one item, then analyses were carried out. Once it is determined that the missing data are missing completely random, first with complete and incomplete data, then with serial mean, median of nearby points, mean of nearby points, linear interpolation, linear trend at point, regression, expectation maximization algorithm data item and test parameters were estimated. In the estimated process, descriptive statistics and

\* Bu çalışma 20-23 Nisan 2017 tarihlerinde gerçekleştirilen 26. Uluslararası Eğitim Bilimleri Kongresi’nde sözlü bildiri olarak sunulmuştur.

\*\* Dr., Gazi Üniversitesi, Gazi Eğitim Fakültesi, Ankara-Türkiye, ayfersayin@yahoo.com, ORCID ID: orcid.org/0000-0003-1357-5674.

\*\*\* Arş. Gör. Dr. Abant İzzet Baysal Üniversitesi, Eğitim Fakültesi, Bolu-Türkiye, e-posta:alperenyandi@gmail.com, ORCID ID: orcid.org/0000-0002-1612-4249

\*\*\*\* Arş. Gör. Gazi Üniversitesi, Gazi Eğitim Fakültesi, Ankara-Türkiye, esra.tas18@gmail.com, ORCID ID: orcid.org/0000-0002-4337-7815

cronbach alpha reliability coefficient and marginal reliability coefficient; the threshold parameters and the difficulty indices were estimated according to the graded response theory, which is one of the IRT models. The results of the study showed that the item and test parameters were influenced by incomplete and missing data techniques; it was determined that the best estimation results were obtained by linear interpolation method with different data.

*Keywords:* Missing data, missing data techniques, item response theory, graded response theory

## GİRİŞ

Eğitim ve psikoloji alanlarında araştırmacılar genellikle, belirledikleri psikolojik yapılara bireylerin ne düzeyde sahip olduğunu ortaya koymayı hedeflemektedir. Bu hedef doğrultusunda psikolojik yapılarla ilgili olarak elde edilen sonuçlara dayalı olarak birey veya gruplarla ilgili önemli kararlar alınmaktadır. Bireylerin psikolojik yapılara sahip olma düzeylerinin doğrudan ölçülmesi çoğu zaman mümkün değildir. Gizil değişken özelliğine sahip olan bu yapıların ölçülmesinde bireylerin ölçme araçlarındaki gözlenen değişkenlere (maddeler) vermiş olduğu tepkilerden yararlanılmaktadır (Hambleton, Swaminathan, Cook, Eignor ve Gifford, 1977).

Bireylerin psikolojik yapılarını belirlemeyi amaçlayan bu çalışmalarda, araştırmacılar bireylerin gözlenen değişkenlere eksiksiz şekilde yanıt vermelerini sağlamaya çalışmaktadır. Bu durumun nedeni olarak ölçme araçlarının psikometrik özelliklerinin incelenmesi ve belirlenen amaç doğrultusunda belirlenmeye çalışılan sonuçlar için yapılan analizlerde eksiksiz veri ile çalışılmasının gerekli olması durumu gösterilebilir. Ancak psikolojik yapılarla ilgili yapılan ölçme uygulamalarında bireylerin kendilerini ifade etmesi yoluyla veri toplama süreci yürütüldüğünden, çeşitli nedenlerden dolayı eksik verilerin ortaya çıkma durumu ile karşı karşıya kalılabilmektedir. Uygulama sonucunda elde edilen veri setlerinde ortaya çıkan eksiklikler kayıp veri olarak nitelendirilmektedir. Çok sayıda madde içeren ölçme araçlarının kullanılması, veri kayıt sürecinde ortaya çıkan teknik hatalar, bireylerin ölçülen konuya karşı hassas olmaları durumundan dolayı maddelere cevap vermek istememesi, fiziksel nedenler, konu ile ilgili duyarsızlık ve bilgi eksiklikleri, zamanı yetiştirememesi gibi nedenler kayıp verinin ortaya çıkmasına sebep olmaktadır (Field, 2005; Goregebeur, De Boeck ve Molenberghs, 2010).

Araştırmalar kapsamında elde edilen kayıp veriler, ölçülmek istenen psikolojik yapının tam olarak ölçülememesine yol açmaktadır. Kayıp verilerin varlığı nedeniyle, bireyler veya gruplar ilgili psikolojik özellik açısından yanlış kestirimlerde bulunulabilir. Bu durum elde edilen sonuçların doğruluğunu olumsuz yönde etkileyebilir. Leeuw, Hox ve Huisman (2003), eksik verinin en temelde bilgi eksikliğine yol açtığını, aynı zamanda kestirimlerin etkisini azalttığını ve istatistiksel test gücünü azalttığını belirtmektedir. Geniş ölçekli testlerde de öğrencilerin hem testlerde hem de ölçek ya da anket maddelerindeki eksik verilerinin olması durumunda bireyler hakkında eksik bilgiye ulaşıldığı için yanlışlık olabileceği ifade edilmektedir (Rose, Davier ve Xu, 2010). Benzer şekilde Doğanay Erdoğan (2012) da kayıp verinin bilgi kaybına neden olduğuna ve bu durumda ölçülmek istenen özelliğin doğru ve güvenilir olarak ölçülüp ölçülmediği sorununun ortaya çıktığını ifade etmektedir. Araştırma sürecinde gerçekleştirilen istatistiksel analizlerin işleyişinde kayıp verilerin yer aldığı veri setleri için problemlerle karşılaşılabilir (Bal, 2003) çünkü faktör analizi gibi bazı istatistikler tam veri seti matrisi gerektirmektedir (Peng, Harwell, Liou ve Ehman, 2007). Örneğin özellikle küçük örneklem büyüklüklerinde kayıp veri II. tip hatanın artmasına neden olabilir, pozitif olmayan kovaryans matrisi oluşturulmasına neden olabilir, varyansı azaltabilir (Acock, 2005). Kayıp verilerin olduğu durumlarda, özellikle kayıp veri göz ardı edilerek hesaplamalar yapılırsa sonuçlar yanlış hesaplanır ve bu da alınan kararların yanlış olmasına neden olmaktadır (Ambler ve Omar, 2007). Örneğin Hedeker, Mermelstein ve Demirtas (2007) tarafından sigara içmeye yönelik yürütülen bir çalışmada eksik verisi bulunan kişilerin gözlemlenen kişilere göre sigara içme olasılıklarının daha yüksek olduğu, başarısız bir sigara içmeyi bırakma süreçlerinin olduğu belirlenmiş, eksik verilerin göz ardı edilmesi durumunda bu bilgilerin kaybolacağına dikkat çekilmiştir. Kayıp değerler hem iki kategorili hem de çok kategorili puanlanan verilerde testin

parametre kestirimlerinde önemli bir yanlışlık sebebi olabilir (Demir ve Parlak, 2012; Demir, 2013). Finch (2008) de gerçekleştirdiği çalışmada kayıp verinin olduğu durumlarda maddelerin güçlük ve ayırıcılığının yanlış kestirildiğine işaret etmektedir. Parametrelerde olduğu gibi kayıp verinin bireylerin yeteneklerinin de beklenenin üstünde ya da altında kestirilmesine neden olduğu belirlenmiştir (Ayala, Plake ve Impara, 2001; Hohensinn ve Kubinger). Tüm bu süreçler dikkate alındığında kayıp veride oluşan sorunların dört başlıkta toplanabileceği (Peng ve diğerleri, 2007); (i.) kayıp verinin olması durumunda eksik veriler bilinemediği için yanlışlık oluşacağını, bunun da sonuçların temsil edilebilirlik yani ait olduğu grubu ya da evreni tanımlamada hatalara neden olabileceğini belirtmektedir. Ayrıca (ii.) istatistiksel analizlerin gücünün azalmasına, standart hatanın artmasına sebebiyet verebileceğini; (iii.) faktör analizi gibi bazı analizlerin tam veri ile gerçekleştirilmesi gerektiğine, aksi durumda hesaplama yapılamayacağına; (iv.) verilerin yeniden toplama sürecinin çaba, zaman ve maliyet açılarından da sıkıntı oluşturabileceğine dikkat çekmektedir (Peng ve diğerleri, 2007). Bütün bu nedenlerden dolayı kayıp veriler araştırma süreçlerinde üzerinde durulması gereken öncelikli konular arasında yer almaktadır.

Kayıp değerlere ilişkin olarak yapılan incelemelerin ilk adımı verinin örüntü varlığını incelenmek ile başlamaktadır. Veri setinde yer alan kayıp değerlerin oluşturduğu problemin ne derecede önemli olduğu, örüntüye sahip olup olmadığına bağlıdır (Çokluk ve Kayrı, 2011). Veri setleri içerisinde belirlenen kayıp verilerin rastsal olarak dağılım göstermesi, bu verilerle baş etme açısından daha düşük düzeyde bir problem teşkil etmektedir. Ancak kayıp değerlerin bir örüntüye bir başka ifadeyle rastsal olmayan bir dağılıma sahip olması; araştırmacıları belirlenen örüntü bağlamında bir yol izlemeye götürdüğünden daha önemli düzeyde bir problem oluşturmaktadır.

Rubin (1976) ve Little ve Rubin (1987) araştırmasında kayıp verilerin oluşum mekanizmasını üç başlıkta ele alınmıştır: tamamıyla rastsal kayıp-TRK (missing completely at random-MCAR), rastsal kayıp-RK (missing at random-MAR) ve rastsal olmayan kayıp-ROK (missing not at random-M-NAR). Allison (2002) ise kayıp veri mekanizmalarını iki varsayım altında incelemektedir. Bu varsayımlar tamamıyla rastsal kayıp-TRK (missing completely at random-MCAR) ve rastsal kayıp-RK (missing at random-MAR) şeklinde sıralanabilir. Acock (2005) da benzer doğrultuda üç tür kayıp veri olduğundan söz eder: tamamıyla rastsal kayıp-TRK, rastsal kayıp-RK ve ihmal edilemez kayıp (nonignorable-NI).

TRK varsayımı, içerisindeki kayıp veri barındıran bir Y değişkeninde bulunan kayıp verilerin olasılığının, bu Y değişkeninin kendi değerine ve veri setindeki diğer değişkenlerle ilişkisiz olmasını ifade etmektedir. TRK varsayımı, Y değişkenindeki kayıp verilerin, veri setindeki bir başka değişkendeki kayıp verilerle ilişkili olmasına izin vermemektedir. Ancak bu durumda dahi verilerin tamamen rastsal olarak kayıp veri olması mümkündür. Veriler büyük bir matris olarak düşünüldüğünde TRK varsayımında kayıp veriler, matrise rastsal bir şekilde dağılmaktadır (Acock, 2005). RK varsayımı ise Y değişkenindeki kayıp verilerin, Y değişkeninin kendi değeri ile ilişkisiz olmasını ifade etmektedir. Bir başka ifadeyle bu varsayım hem Y hem de veri setindeki bir başka değişken olan X değişkeni birlikte ele alınırken, Y değişkeninde kayıp veri görülme olasılığı ile sadece X değişkeni ele alındığında Y değişkeninde kayıp veri görülme olasılığının eşit olması durumudur. Araştırmacılar Little'ın MCAR testi ile TRK varsayımını test etme imkanına sahipken RK varsayımının test edilmesi daha güçtür ve henüz önerilen kesin bir test bulunmamaktadır. Tam veri seti Y'de kayıp verilerin kendisi, gözlenmeyen cevapların yani kayıp verilerin oluşmasına neden oluyorsa bu durumda kayıp veri mekanizması rastsal olmayan kayıp-ROK (M-NAR) olacaktır. ROK veri yapısı kayıp verilerin kendisinin yanı sıra gözlenen diğer verilere de bağlı olabilmektedir (Rubin, 1976). Üniversite öğrencilerinin akademik performanslarının incelendiği bir panel çalışmada üniversiteden ayrılan öğrencilerin akademik performanslarının daha düşük olma ihtimali bulunmaktadır ve bu durumda oluşan kayıp veri de "ihmal edilemez kayıp" olarak adlandırılmaktadır (Acock, 2005).

Veri setinde yer alan kayıp verilerin belli bir örüntü oluşturmadığı durumlarda kayıp verilerin veri setinden çıkarılması ve tamamlanması gibi farklı çözüm yolları önerilmektedir (Allison, 2002; Carpita ve Manisera, 2011; Demir ve Parlak, 2012; Şahin Kürşad ve Nartgün, 2015). Bu yöntemler

silme, yaklaşık değer atama ve yeni yaklaşımlar (Demir ve Parlak, 2012) ya da silme-basit atama ve model tabanlı atama yöntemleri (Schafer ve Graham, 2009) olarak sınıflandırılabilir.

- a) Silme Yöntemi: Bu yöntemde kayıp veri bulunan satırların veri seti dışında bırakılması söz konusudur. Ancak örneklem büyüklüğündeki azalma, çalışmanın geçerlik ve güvenilirliğinde yanlı kestirimler elde edilmesine sebep olmaktadır (Akbaş ve Tavşancıl, 2015; Baygül, 2007; Çüm ve Gelbal, 2015; Şahin Kürşad ve Nartgün, 2015; Yılmaz, 2014).
- b) Yaklaşık değer atama yöntemleri (Çokluk ve Kayri, 2011): Bu yöntemde geçmiş bilgileri kullanmak, ortalama değer atamak ve regresyon işlemleri gerçekleştirilebilir.
  - i) Seri ortalaması (SO): Bu yöntemde kayıp veri içeren değışkende mevcut olan diğer değerlerin tüm katılımcılar için ortalaması alınarak atama işlemi gerçekleştirilir.
  - ii) Yakın noktaların ortalaması (YNO): Kayıp verinin bulunduğu hücrenin yakınındaki değerlerin aritmetik ortalaması üzerinden atama işlemleri gerçekleştirilmektedir. Atama işlemi esnasında kayıp verinin bulunduğu hücrenin altındaki ve üstündeki tam değerlerden yararlanılmaktadır.
  - iii) Yakın noktaların medyanı (YNM): Bir önceki yöntemle benzer şekilde atama işlemi gerçekleştirilen bu yöntemde kayıp veri hücrenin altında ve üstünde yer alan tam veriler kullanılarak hesaplanan medyan değeri atanarak işlem yapılmaktadır.
  - iv) Doğrusal değerleme (DD): Bu yöntemde ise kayıp veri hücrenden önceki ve sonraki ilk tam verinin ortalaması, kayıp değeri yerine atanarak işlem yapılmaktadır.
  - v) Noktanın doğrusal eğimi (NDE): Bu yöntemde kayıp veri dışında kalan yöntemlerden yararlanılması söz konusudur. Mevcut tam verilerin sahip olduğu yükseliş veya düşüş eğilimi doğrultusunda atama işlemi gerçekleştirilir.
  - vi) Regresyon ataması (RA): Kayıp veriler dışında kalan tam veriler kullanılarak elde edilen regresyon modeli aracılığı ile kayıp veriler yerine atama yapılır.
- c) Yeni yaklaşımlar (Yılmaz, 2014):
  - i) Beklenti maksimizasyonu (BM): Bu yöntem iteratif şekilde tekrar eden iki aşama üzerinden atama işlemi yapılmaktadır. İlk aşama olan beklenti aşamasında, kayıp verilere başlangıç değerleri atanır. İkinci aşama olan maksimizasyon aşamasında ise, bu başlangıç değerleriyle oluşan beklentiler maksimize edilir. Bu beklenti-maksimize etme döngüsü, bundan sonra, atanan değerler, önceden belirlenmiş bir yakınsama kriterine dayalı olarak benzer hale gelene kadar tekrarlanmaktadır.

Kayıp veri ile baş etme yöntemlerinin araştırma sonuçları üzerine etkileri ile ilgili olarak alanyazında farklı birçok çalışma mevcuttur. Çokluk ve Kayri (2011), Ankara Üniversitesi Eğitim Bilimleri Fakültesi Sınıf Öğretmenliği bölümünde öğrenimine devam eden 200 öğrencinin Fatalizm Ölçeği maddelerine verdikleri yanıtlardan oluşan verileri %15 - %20 ve %0 - %50 oranında kayıp veri içerecek şekilde düzenlemiştir. Çalışma kapsamında elde edilen bu veri setlerindeki kayıp verilerin tamamlama işlemi sonrasında faktör yapıları, düzeltilmiş madde-toplam korelasyonları ve Cronbach-alfa iç tutarlık katsayıları karşılaştırılmıştır. Sonuç olarak çalışma kapsamında tam veri için elde edilen faktör yapılarının benzer olduğu ancak tamamlanmış veriler üzerinden elde edilen açıklanan varyans, öz değeri ve iç tutarlılık katsayısında bir düşüş olduğu gözlenmiştir. Köse ve Öztumur (2014) ise örneklem büyüklüğü ve kayıp veri oranının t testi ve ANOVA olmak üzere test istatistikleri üzerindeki etkisini incelemiş, kayıp veri yöntemlerinin fark testlerine etki ettiğini ortaya koymuşlardır. Akbaş ve Tavşancıl (2015) da araştırmalarında liste bazında silme tekniğinin test istatistiklerinde yanlı sonuçlara neden olduğunu ve beklenti maksimizasyonu ve çoklu değeri regresyon atama tekniklerinin ise genel olarak daha yüksek kestirimler gerçekleşmesine neden olduğunu belirtmişlerdir. Çüm ve Gelbal (2015), yapmış oldukları çalışmada PISA 2012 Türkiye örneklemini, tam veri seti üzerinden tamamen rastsal ve tamamen rastsal olmayacak şekilde %20 ve %30 oranında kayıp veri olacak şekilde veri setlerini düzenlemiştir. Düzenlenen bu veri setlerinde kayıp veriler

yerine 10 farklı yöntemle eksik verilerle baş etme yöntemlerine dayalı tam hâle getirilmiş ve bu durumun model veri uyumu değerlerine etkisi ele alınmıştır. Araştırma sonucunda kayıp verilerin tamamıyla rastsal olarak dağıldığı durumlarda regresyonla atama yöntemi sonrası elde edilen veri üzerinden kestirilen model veri uyum değerlerinin tam veri setinin model uyum değerlerine en yakın sonuçları verdiği tespit edilmiştir. Çalışmada ayrıca, yaklaşık değer atamalarının veri seti için belirlenen dağılımları önemli düzeyde etkilediği, bu nedenle araştırmacıların uygun atama yöntemlerini kullanarak süreci gerçekleştirmeleri önerilmiştir. Şahin Kürşad ve Nartgün (2015) ise yapmış oldukları çalışmada PISA 2012 “Matematik Çalışma Etiği” ölçeği Türkiye örnekleme verileri içerisinde 200 kişilik bir alt örneklem seçmiştir. Seçilen bu veri seti üzerinde %5, %10 ve %20 oranında, tamamen rastsal olacak şekilde veri silme işlemi yapılmıştır. Oluşturulan kayıp verili setleri farklı yöntemlerle tamamlandıktan sonra geçerlik ve güvenilirliğe ilişkin kestirimler gerçekleştirilmiştir. Araştırma sonucunda değer atama yöntemleri ile oluşturulan veri setlerinden elde edilen parametre değerlerinin kayıp veri oranının düşük olduğu durumlarda genel olarak tam veri setinden elde edilen değerlere yakın veya aynı değerler verdiği rapor edilmiştir. Bunun yanı sıra ele alınan tüm durumlar için çoklu atama, beklenti maksimizasyonu ve regresyon ataması yöntemlerinin tam veri setinden elde edilen değerlere en yakın değer veren yöntemler olduğu belirtilmiştir. Soysal ve Akın Arıkan (2017) da kayıp veri yöntemleri ile örneklem büyüklüğünün faktörleşmeye olan etkisini araştırdıkları çalışmalarında, faktörleştirme tekniklerinin hemen hemen her koşulda benzer performans gösterdiği ve atama yöntemleri açısından farklılaşmadığı sonucuna ulaşmışlardır. Uluslararası alanyazında da benzer şekilde kayıp veri yöntemlerinin açılımlayıcı faktör analizi gibi genel test istatistikleri üzerindeki etkilerinin incelendiği çalışmalar mevcuttur (Josse ve Husson, 2012; McNeish, 2016; Weaver ve Maxwell, 2014).

Ele alınan bu çalışmalar kapsamında kayıp veri ile baş etme yöntemlerinin test parametrelerini etkilediği görülmektedir çünkü alanyazındaki çalışmalarda veri setlerindeki kayıp, testin geneli için oluşturulmuş, veriler madde bazına indirgenmemiştir. Bir başka ifadeyle maddelerde bulunan kayıp veri oranının değişiklik göstermesinin diğer maddeler üzerinde farklı etkileme durumuna yönelik inceleme yapılmamıştır. Ayrıca bu yöntemlerin Madde Tepki Kuramı (MTK) temelli parametreler üzerinde etkisi sınırlı çalışmada ele alınmış (Koçak ve Çokluk Bökeoğlu, 2017), çalışmalarda genel olarak Klasik Test Kuramı temelinde incelemeler yapılmıştır. Madde bazında ortaya çıkan kayıp veri miktarının diğer maddelerin parametreleri üzerindeki etkisinin araştırılması, araştırmacılara kullandıkları kayıp veri ile baş etme yönteminin seçiminde bilgi sağlayabilir. Ayrıca MTK'nın temel varsayımlarından biri olan madde parametresi kestirimlerinin maddelerin birbirlerinden bağımsız gerçekleştirilip gerçekleştirilmediği hakkında da bilgi verebilir. Araştırmacılar çalışma kapsamındaki veri setinde bulunan maddelerdeki kayıp veri oranlarına göre uygun olan yöntemi seçmeleri ile birlikte elde edilen sonuçların doğruluğunu olumlu yönde etkileyebilir. Ayrıca MTK kapsamında ele alınan parametrelerin kayıp veri baş etme yöntemlerinden ne şekilde etkilendiğinin belirlenmesi, bu kuram kapsamında temellendirilen çalışmaların sürecinde araştırmacılara katkı sağlayabilir.

### ***Araştırmanın Amacı***

Bu araştırmada diğer araştırmalardan farklı olarak likert tipi derecelendirilmiş bir ölçekte yer alan kayıp veriyle baş etme yöntemlerinin Aşamalı Madde Tepki Kuramı (AMTK) ile kestirilen parametrelere etkisi incelenmiştir. Ayrıca yine diğer araştırmalardan farklı olarak kayıp veri oranı testin geneli için değil, madde bazında değiştirilmiş ve ölçekteki bir madde hiç eksik veri olmayacak şekilde düzenlenmiştir. Bu doğrultuda da bir ölçme aracındaki herhangi bir maddede bulunan kayıp veri miktarının eksik verisi olmayan diğer madde parametrelerine etkisi incelenmiştir. Alanyazında gerçekleştirilen çalışmalarda kayıp veri, örneklem setinin tamamı üzerinden yapılmışken bu çalışmada madde bazlı yapılmış olması ve kestirimlerin aşamalı tepki modeli ile gerçekleştirilmiş olmasının araştırmayı önemli kıldırdığı düşünülmektedir.

## YÖNTEM

Bu araştırmada, farklı örneklem büyüklüklerinde ve kayıp veri oranlarına sahip maddeleri bulunan veri setlerinde madde parametrelerinin kayıp veri baş etme yöntemlerinden ne düzeyde etkilendiğinin belirlenmesi amaçlanmaktadır. Bu doğrultuda araştırmanın temel bir araştırma niteliğinde olduğu söylenebilir.

### *Örneklem*

PISA 2015 çalışmasında yer alan “hırs algısı” ölçeğine eksiksiz cevap veren 5073 öğrenci içerisinde rastgele seçilen 500, 1000 ve 2500 öğrenci, araştırmanın çalışma grubunu oluşturmaktadır. Hesaplamalar MTK’ya göre gerçekleştirildiği için örneklem büyüklüklerinin yeterli sayıda olmasına dikkat edilmiştir. MTK modelleri için alanyazında farklı örneklem büyüklükleri önerilmektedir. Tsutakawa ve Johnson (1990) parametrelerin doğru kestirilmesi için en az 500 örneklem büyüklüğüne ihtiyaç olduğunu belirtmişlerdir.

### *Verilerin Elde Edilmesi*

Araştırmada PISA 2015 uygulamasında yer alan ve öğrencilerin hırs algısını belirlemeyi amaçlayan 5 maddelik “hırs algısı” ölçeği kullanılmıştır. Bu ölçeğin seçilmesinde maddelerin tek boyutta toplanması göz önünde bulundurulmuştur. Ölçekte yer alan maddeler şu şekildedir:

- Derslerimin çoğunda veya tamamında en yüksek notu almak istiyorum.
- Mezun olduğumda bana uygun en iyi fırsatlardan birini seçmek istiyorum.
- Ne yaparsam yapayım en iyisi olmak istiyorum.
- Kendimi hırslı bir insan olarak görüyorum.
- Sınıftaki en iyi öğrencilerden biri olmak istiyorum.

4’lü likert tipinde derecelendirilen ölçek maddeleri olumsuz ifade içermemektedir.

### *İşlem*

Araştırma verilerine OECD’nin internet adresinden ([www.oecd.org/pisa/](http://www.oecd.org/pisa/)) ulaşılmıştır. Veri seti içerisinde Türkiye örneklemini alınmıştır. Örneklem, OECD sekreteryasında Westat (ABD) liderliğinde seçilmiş ve örneklemin evreni temsil ettiği belirtilmiştir.

### *Verilerin Analizi*

Verilerin analizi sürecinde öncelikle MTK varsayımlarının karşılanıp karşılanmadığının incelenmesi, ardından veri seti içerisinde kayıp verilerin oluşturulması ve kayıp verilerin tamamen rastsal olup olmama durumunun test edilmesi, son olarak da MTK’ya dayalı olarak madde parametrelerinin hesaplanması işlemleri gerçekleştirilmiştir.

### *Varsayımların İncelenmesi*

Araştırma kapsamında hesaplamalar MTK’ya dayalı olarak gerçekleştirildiği için MTK’nın temel varsayımları olan normallik, tek boyutluluk, yerel bağımsızlık ve değişmezlik varsayımları seçilen 500, 1000 ve 2500 kişilik örneklem grupları için incelenmiştir.

Örneklemden elde edilen verilerin evrene ait uzayda normal dağılım gösterip göstermediğinin incelenmesi MTK'nın temel varsayımlarından biridir. Tek değişkenli normalliğin test edilmesinde, çarpıklık ve basıklık katsayıları hesaplanmış ve sonuçlar Tablo 1'de gösterilmiştir.

Tablo 1. Farklı Örneklem Büyüklüğündeki Setlerin Çarpıklık ve Basıklık Katsayıları

Değerler	Örneklem büyüklükleri		
	n=500	n=1000	n=2500
Çarpıklık	-0,975	-0,951	0,929
Basıklık	0,677	0,720	0,691

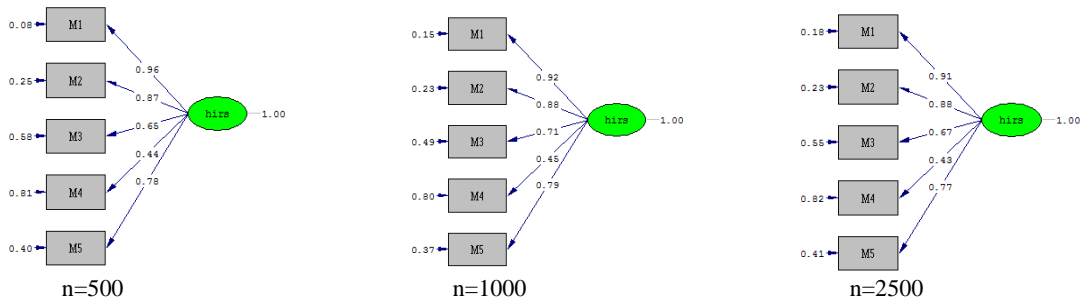
Çarpıklık ve basıklık katsayılarının  $\pm 1$  arasında hesaplanması, verilerin normal dağılımdan aşırı bir sapma göstermediğini belirtmektedir (Büyüköztürk, 2017). Tablo 1'de yer alan bilgiler incelendiğinde veri setlerinin genel olarak normal bir dağılım gösterdiği görülmektedir. Çarpıklık ve basıklık katsayılarına ek olarak verilere ilişkin histogram grafikleri de incelenmiş, veri setlerinin normal bir dağılım gösterdiği tespit edilmiştir.

MTK'nın diğer önemli varsayımlarından biri tek boyutluluktur. Tek boyutluluğun test edilmesi için açımlayıcı ve doğrulayıcı faktör analizi teknikleri kullanılmıştır. Açımlayıcı faktör analizi sonuçlarına Tablo 2'de yer verilmiştir.

Tablo 2. Farklı Örneklem Büyüklüğündeki Setlerin Açımlayıcı Faktör Analizi Sonuçları

Değerler	Örneklem büyüklükleri		
	n=500	n=1000	n=2500
Özdeğer	2,760	2,813	2,704
Açıklanan varyans (%)	55,204	56,251	54,084

Tablo 2'de yer alan değerler incelendiğinde tüm örneklem setlerinde 5 maddeden oluşan ölçek maddelerinin tek bir boyutta toplandığı ve boyutun özdeğerinin 2'den, açıklanan varyans oranının ise %30'dan büyük (Çokluk, Şekercioğlu ve Büyüköztürk, 2014) olduğu belirlenmiştir. Ardından doğrulayıcı faktör analizine başvurulmuş ve analizin sonucunda tüm maddelerin anlamlı t değerlerine sahip oldukları, başka bir ifadeyle anlamlı bir açıklayıcılıklarının bulunduğu belirlenmiştir. Bu analiz sonucunda oluşturulan yol (path) diyagramları Şekil 1'de gösterilmiştir.



Şekil 1. Farklı Örneklem Büyüklüğündeki Setlerin Doğrulayıcı Faktör Analizi Yol Diyagramları

Açımlayıcı ve doğrulayıcı faktör analizi sonuçlarına göre ölçek maddelerinin 500, 1000 ve 2500 kişilik örneklem setlerinde tek boyutta toplandığı tespit edilmiştir. Yerel bağımsızlık tek boyutlulukla ilişkili bir özelliktir. Eğer bir ölçek tek boyutluluk özelliğini gösteriyorsa, bu ölçekte yer alan maddelerin yerel bağımsızlık özelliğine sahip olduğu söylenebilir (Hambleton ve Swaminathan, 1997).

MTK'nın bir diğer varsayımı madde parametrelerinin kestirimin yapıldığı gruptan; yetenek parametrelerinin ise kestirimin yapıldığı maddelerden bağımsız olmasıdır (Wright, 1997). Bu varsayımın test edilmesi amacıyla öncelikle madde parametrelerini kestirmek için grup rastgele ikiye bölünmüş, sonrasında her iki grup için de madde parametreleri kestirilmiştir. Kestirilen parametreler arasında Pearson korelasyon katsayısı hesaplanmış ve korelasyon değerlerinin kabul edilebilir aralıkta olduğu ( $>.70$ ) belirlenmiştir. Parametre değişmezliğinin ikinci kısmında ölçekte yer alan maddeler tek ve çift maddeler olarak ayrılmış sonrasında ölçeceği alan bireyler için iki ayrı yetenek parametresi kestirilmiştir. Kestirilen yetenek parametreleri arasında Pearson korelasyon katsayısı hesaplanmış ve korelasyon değerlerinin anlamlı olduğu bulunmuştur.

Hırs algısı ölçeğinde yer alan beş maddeye ilişkin verilerin model veri uyumu için “-2likelihood (negative twice the loglikelihood)” değeri hesaplanmıştır. Bu değer 500 örneklem büyüklüğünde 2215,3 hesaplanmış, 1000 örneklem büyüklüğünde -5972.8 olarak; 2500 örneklem büyüklüğünde -19003.9 olarak hesaplanmıştır; -2likelihood değerinin yüksek olması modelin uyumlu olduğunun bir göstergesidir.

#### *Kayıp Verilerin Oluşturulması ve Test Edilmesi*

Varsayımların incelenmesinin ardından tamamen rastsal kayıp veri-TRK (MCAR) setleri oluşturulabilmesi için tüm gözlemlerin kayıp olma ihtimallerinin birbirine eşit olması gerekmektedir. Bu nedenle öncelikle gözlemler ile değişkenler birbiri ile çarpılarak toplam hücre sayısı bulunmuştur. Her bir hücre içinde eksiltilecek veri sayısı belirlenmiştir. Ardından MS Excell dosyasında oluşturulan bir betikle (random seçim; aynı casede dört veri silme vb.) silme işlemi gerçekleştirilmiştir. Bu kısımda yapılan veri silme işlemi tamamen rastsaldır ve TRK mekahnizmasının mantığını uygun yapılmaktadır. Öncelikli olarak kayıp veri oluşturulmuş ve sonrasında kayıp verinin rastsal olup olmadığı Little'in MCAR testi uygulanarak test edilmiştir. Daha önce de belirtildiği gibi eksiltme işlemleri maddeler bazında gerçekleştirilmiştir. Her bir veri setinde birinci maddenin %20'si; ikinci maddenin %15'i; üçüncü maddenin %10'u; dördüncü maddenin %5'i kayıp veri durumuna getirilmiştir. Beşinci maddede ise herhangi bir kayıp veri oluşturulmamıştır. Bu doğrultuda her bir veri setinde %12 kayıp veri bulunmuştur.

Kayıp veriler oluşturulduktan sonra kayıp verilerin, veri setlerinde tamamen rastsal bir şekilde dağılıp dağılmadığının belirlenmesi amacıyla Little'in MCAR testi hesaplanmıştır. Hesaplama sonuçları Tablo 3'te gösterilmiştir.

Tablo 3. Farklı Örneklem Büyüklüğündeki Setlerdeki Eksik Verilerin Tamamen Rastsallığına İlişkin Hesaplanan Little'in MCAR Testi Sonuçları

Değerler	Örneklem büyüklükleri		
	n=500	n=1000	n=2500
X <sup>2</sup>	38,001	27,381	48,309
sd	36	40	40
p	0,378	0,935	0,172

Little'in MCAR testi eksik verilerin TRK olarak dağılıp dağılmadığının belirlenmesinde kullanılan en yaygın testtir. Little'in MCAR testi için hesaplanan p değerinin anlamlı olmaması yani 0,05'ten büyük olması, eksik verilerde herhangi bir örüntü bulunmadığını içeren yokluk hipotezinin kabul edildiğini, kayıp verinin TRK olduğunu göstermektedir (Garson, 2015). Bu doğrultuda Tablo 3 incelendiğinde her bir veri setinde yer alan eksik verilerin tamamen rastsal bir dağılım gösterdiği belirlenmiştir ( $p>0,05$ ).



*Verilerin Çözümlemesi*

Verilerin çözümleme sürecinde öncelikle eksik veriler için silme, ortalama atama, yakın noktalar ortalama ataması, yakın noktalar medyan ataması, doğrusal değerlendirme, noktada doğrusal değerlendirme, regresyon atama ve beklenti maksimizasyonu atama olmak üzere 8 kayıp veri ile baş etme yöntemine dayalı olarak eksik veri setleri tamamlanmıştır. Bu doğrultuda her bir örneklem büyüklüğü için tam veri seti, eksik (eksiltilmiş, kayıplar oluşturulmuş) veri seti, 8 farklı kayıp veri başa çıkma yöntemine göre tam hâle getirilmiş veri setleri olmak üzere 10'ar veri seti oluşturulmuştur. Bu veri setleri üzerinde test istatistikleri için ortalama, standart sapma, Cronbach alfa güvenilirlik ve marjinal güvenilirlik katsayısı hesaplanmıştır. Ardından da Aşamalı Tepki Modeli'ne göre ayırıcılık ve güçlük indeksleri hesaplanarak sonuçlar karşılaştırılmıştır. Araştırmada kullanılan ölçek maddeleri çok kategorili puanlandığı için Samejima'nın Aşamalı Tepki Modeli olarak adlandırılan model kullanılmıştır. Bu kuram cevaplayıcının yetenek düzeyi ile belli bir kategoriye tepki vermesi arasında doğrusal olmayan ilişkilere dayanmaktadır ve bireylerin yetenekleri hakkında bilgi elde edilmektedir. Eşik parametresi, bir maddenin her bir kategoriye kadar olan sınıflarının 0,50 olasılıkla yanıtlanması için gerekli olan düzeyi tanımlar ve kategori sayısının bir eksiği kadar eşik parametresi hesaplanır. Modelde bir madde için her bir eşik değeri için bir işlem karakteristik eğrisi, her eşik için güçlük indeksi ve maddenin tümü için bir ayırıcılık parametresi kestirilmektedir (Emretson ve Reise, 2000). Madde parametrelerinin hesaplanmasında MULTILOG programı kullanılmıştır.

**BULGULAR****1. 500 örneklem büyüklüğünde kayıp veri ile baş etme yöntemlerine dayalı olarak test ve madde istatistikleri nasıl bir değişim göstermektedir?**

Araştırma kapsamında 500 örneklem büyüklüğü için oluşturulan 10 veri için betimsel istatistikler ve güvenilirlik katsayıları hesaplanmış, sonuçlar Tablo 4'te gösterilmiştir.

Tablo 4. Betimsel İstatistikler ve Güvenirlik Katsayıları (n=500)

Veri setleri	Ortalama	SS	Cronbach alfa güvenilirlik	Marginal güvenilirlik
Tam veri	17,080	2,580	0,775	0,738
Eksik veri	15,370	3,250	0,772*	0,731
Silme	17,160	2,540*	0,764	0,735
Ortalama atama	17,120	2,350	0,731	0,711
Yakın noktalar ortalama ataması	17,100*	2,510	0,770	0,735
Yakın noktalar medyan ataması	17,110	2,530	0,771	0,734
Doğrusal değerlendirme	17,100*	2,530	0,770	0,736*
Noktada doğrusal eğilim	17,005	2,400	0,735	0,727
Regresyon atama	17,110	2,540	0,770	0,732
Beklenti maksimizasyonu	17,110	2,520	0,782	0,734

\* tam veri parametrelerine en yakın olan değerler

Tablo 4'te görüldüğü gibi tam veri setinden elde edilen değerler, yani veri setinin hiçbir eksik verisi bulunmayan hâlinde hesaplanan değerler, (ilk satır) hem sonradan oluşturulan eksik veri seti için hem de eksik verilerle baş etme yöntemlerine dayalı tam hâle getirilmiş veri setleri için bir referans oluşturmaktadır. Bu doğrultuda eksik veriler ile baş etme yöntemlerine dayalı tam hâle getirilmiş veri setlerinde hesaplanan değerler; veri setlerinin hesaplama yanlılığından ne düzeyde etkilendiğini belirlemek amacıyla tam veri setinden elde edilen değerler ile karşılaştırılmıştır.

Tablo 4'te yer alan bilgiler incelendiğinde, 500 örneklem büyüklüğüne tam veri seti ile eksik veri setinin ortalama ve standart sapma betimsel istatistikleri arasında farklılıklar olduğu görülmektedir. Bununla birlikte kayıp veri ile baş etme yöntemlerine dayalı olarak oluşturulan veri setlerinin ortalama, standart sapma ve güvenilirlik katsayılarının tam veri ile büyük ölçüde benzerlik gösterdiği belirlenmiştir. Noktada doğrusal atama yöntemi (-0,075 fark) dışındaki eksik veri ile baş etme

yöntemleri sonucunda ortalama değerin tam veriye yakın olmakla birlikte biraz daha yüksek kestirildiği belirlenmiştir. Güvenirlik katsayılarının yüksek kestirilmiş olmakla birlikte gerçek değere oldukça yakın olduğu görülmektedir. Başka bir ifadeyle 500 örneklem büyüklüğünde eksik veri için kullanılan kayıp veri ile baş etme yöntemlerinin test istatistiklerine olan etkisi benzerlik göstermektedir.

500 örneklem büyüklüğünde kayıp veri ile baş etme yöntemlerinin madde parametreleri üzerindeki etkisini belirlemek amacıyla aşamalı tepki modeli ile madde parametreleri hesaplanmış ve sonuçlar Tablo 5'te gösterilmiştir.

Tablo 5. Aşamalı Madde Tepki Modeline Dayalı Kestirilen Ayırıcılık ve Güçlük İndeksleri (n=500)

Madde no	Madde parametreleri	Veri setleri									
		1	2	3	4	5	6	7	8	9	10
M1 (%20 eksik veri)	a	5,86	7,51	7,58	3,63	7,79	8,69	6,81*	4,44	7,05	11,5
	b <sub>1</sub>	-2,26	-2,21	-2,23	-2,51	-2,24	-2,24	-2,25*	-2,45	-2,3	-2,15
	b <sub>2</sub>	-1,81	-1,84	-1,82*	-2,1	-1,78	-1,77	-1,8*	-2,06	-1,82*	-1,72
	b <sub>3</sub>	-0,52	-0,57	-0,61	-0,81	-0,54	-0,54	-0,53	-0,55	-0,55	-0,52*
M2 (%15 eksik veri)	a	3,2	2,25	3,39	2,01	3,02	3,41	3,17*	2,27	2,99	4,64
	b <sub>1</sub>	-2,44	-2,41*	-2,32	-2,83	-2,51	-2,49	-2,55	-2,86	-2,55	-2,35
	b <sub>2</sub>	-1,95	-2	-1,98*	-2,36	-2,1	-1,98*	-2,08	-2,39	-1,99	-1,89
	b <sub>3</sub>	-0,61	-0,64	-0,59*	-0,85	-0,67	-0,66	-0,63*	-0,75	-0,63*	-0,58
M3 (%10 eksik veri)	a	1,69	1,75	1,81	1,96	1,65	1,65	1,69*	2,0	1,96	1,75
	b <sub>1</sub>	-3,08	-2,97	-2,79	-3,01	-3,08*	-3,09	-3,07	-3,08*	-2,96	-2,98
	b <sub>2</sub>	-1,64	-1,62	-1,55	-1,68	-1,7	-1,71	-1,68	-1,71	-1,64*	-1,66
	b <sub>3</sub>	-0,03	-0,07	-0,13	0,11	-0,05	-0,05	-0,04*	0,01	-0,08	-0,02*
M4 (%5 eksik veri)	a	0,95	1,01	0,88	0,98	0,99	0,99	0,95*	1,11	1,01	0,95*
	b <sub>1</sub>	-3,47	-3,41	-3,54	-3,43	-3,39	-3,39	-3,52	-3,12	-3,3	-3,47*
	b <sub>2</sub>	-1,43	-1,4	-1,57	-1,45	-1,41	-1,44*	-1,44*	-1,05	-1,39	-1,42*
	b <sub>3</sub>	0,97	0,94	0,9	1,03	0,91	0,9	0,96	0,95	0,97*	1,09
M5 (%0 eksik veri)	a	2,3	2,1	2,12	2,05	2,24	2,23	2,28*	1,94	2,08	2,03
	b <sub>1</sub>	-2,5	-2,6	-2,65	-2,62	-2,51	-2,52	-2,5*	-2,72	-2,6	-2,6
	b <sub>2</sub>	-1,75	-1,82	-1,92	-1,84	-1,76	-1,77	-1,75*	-1,89	-1,82	-1,81
	b <sub>3</sub>	-0,1	-0,12	-0,14	-0,12	-0,11*	-0,11*	-0,11*	-0,12	-0,11*	-0,09*

1: tam veri, 2: eksik veri, 3: silme, 4: ortalama atama, 5: yakın noktalar ortalama ataması, 6: yakın noktalar medyan ataması, 7: doğrusal değerlendirme, 8: noktada doğrusal eğilim, 9: regresyon atama, 10: beklenti maksimizasyonu

\* tam veri parametrelerine en yakın olan değerler

Tablo 5'te görüldüğü gibi birinci maddede %20 eksik veri bulunmaktadır. Ortalama atama (-2,23 fark), doğrusal değerlendirme (-0,95 fark) ve noktada doğrusal eğilim (-1,42 fark) yöntemleri ile tamamlanan veri setlerindeki ayırıcılık indekslerinin tam veridekinden daha düşük; diğer yöntemlerle kestirilen verilerin de daha yüksek olduğu belirlenmiştir. Ayırıcılığın en yakın hesaplandığı yöntemin 0,95 farkla doğrusal değerlendirme olduğu tespit edilmiştir. Benzer şekilde %15 eksik verisi bulunan ikinci maddede ortalama atama (-1,19 fark), yakın noktalar ortalama ataması (-0,18 fark), noktada doğrusal eğilim (-0,93 fark), regresyon atama (-0,21 fark) başa çıkma yöntemleriyle oluşan setlerindeki ve eksik veri setindeki (-0,95 fark) ayırıcılık indekslerinin daha düşük; diğer yöntemlerle olanlarda daha yüksek olduğu belirlenmiştir. Tam veri ile hesaplanan ayırıcılık indeksinin en yakın olduğu yöntemin doğrusal değerlendirme yöntemiyle (0,03 fark) yapılan kestirim olduğu belirlenmiştir. Üçüncü maddede %10 eksik veri bulunmaktadır ve doğrusal değerlendirme yöntemi kullanılarak kayıp verilerin atandığı veri seti ile tam veri setinin ayırıcılık indeksleri bire bir aynı (0,00 fark) hesaplanmıştır. Diğer eksik verilerle baş etme yöntemlerine dayalı tam hâle getirilmiş veri setleri ile gerçekleştirilen hesaplamaların da genel olarak benzerlik gösterdiği belirlenmiştir. %5 eksik verisi bulunan dördüncü maddenin ayırıcılık indekslerinin noktada doğrusal eğilim yöntemi (0,16 fark) ile oluşturulan veri setinde farklılaştığı, diğer veri setlerinde tam veri seti ile büyük ölçüde benzerlik gösterdiği belirlenmiştir. Benzer şekilde doğrusal değerlendirme (0,00 fark) ve beklenti maksimizasyonu (0,00 fark) yöntemleriyle oluşturulan veri setlerinin madde ayırıcılık indekslerinin tam veri seti ile aynı çıktığı görülmektedir. Ölçeğin beşinci ve son maddesinde eksik veri bulunmamaktadır. Ancak

diğer maddelerdeki eksik verilerle baş etme yöntemlerinden madde ayırıcılık indeksinin etkilendiği belirlenmiştir. Doğrusal değerlendirme yöntemi (0,02 fark) ile gerçekleştirilen atama sonucunda beşinci maddenin ayırıcılık indeksinin diğer veri setlerinden daha yakın olduğu tespit edilmiştir. Madde güçlük indekslerinin veri setleri bazında büyük farklılıklar göstermediği; tam veri seti ile en fazla benzerliğin yine doğrusal değerlendirme yöntemiyle atanan veri setinde olduğu belirlenmiştir.

500 örneklem büyüklüğünde maddedeki kayıp veri oranı azalma gösterdikçe; kayıp veri ile baş etme yöntemleri sonucunda oluşturulan veri setlerinin ayırıcılıklarının tam veride hesaplanan madde ayırıcılık indeksine yaklaştığı belirlenmiştir. Aynı zamanda maddenin ayırıcılık indeksi artış gösterdikçe de farkın arttığı tespit edilmiştir. Araştırma kapsamında beşinci maddede hiç kayıp veri bulunmamasına ve MTK'nın yerel bağımsızlık varsayımı karşılanmasına karşın beşinci maddenin parametre kestiriminin diğer maddelerdeki kayıp verilerden etkilendiği saptanmıştır. Madde güçlük indekslerinin ise madde ayırıcılık indeksine göre kayıp veri ile baş etme yöntemlerinden daha az etkilendiği belirlenmiştir. Tam veri seti üzerinden hesaplanan madde ayırıcılık ve madde güçlük indeksinin en çok doğrusal değerlendirme yönteminin kullanıldığı veri seti sonuçları ile benzerlik gösterdiği belirlenmiştir.

## 2. 1000 örneklem büyüklüğünde kayıp veri ile baş etme yöntemlerine dayalı olarak test ve madde istatistikleri nasıl bir değişim göstermektedir?

1000 öğrenciden oluşan veri setinin de benzer şekilde birinci maddesinin %20'si, ikinci maddesinin %15'i, üçüncü maddesinin %10'u, dördüncü maddesinin %5'i eksiktir, beşinci maddede ise eksik veri bulunmamaktadır. 1000 örneklem büyüklüğünde de benzer şekilde tam veri, eksik veri ve 8 farklı kayıp veri baş etme yöntemi ile tam hale getirilmiş veri setleri üzerinde betimsel istatistikler ve güvenilirlik katsayıları hesaplanmış, sonuçlar Tablo 6'da gösterilmiştir.

Tablo 6. Betimsel İstatistikler ve Güvenirlik Katsayıları (n=1000)

Veri setleri	Ortalama	SS	Cronbach güvenirlilik	alfaMarginal güvenirlilik
Tam veri	17,180	2,570	0,779	0,741
Eksik veri	15,400	3,300	0,784	0,738
Silme	17,220	2,560	0,784	0,743
Ortalama atama	17,180*	2,370	0,743	0,723
Yakın noktalar ortalama ataması	17,180*	2,550	0,779*	0,740
Yakın noktalar medyan ataması	17,190	2,550	0,778	0,738*
Doğrusal değerlendirme	17,180*	2,550	0,777	0,738
Noktada doğrusal eğilim	17,180*	2,380	0,738	0,722
Regresyon atama	17,160	2,570*	0,778	0,736
Beklenti maksimizasyonu	17,170	2,540	0,792	0,743

\* tam veri parametrelerine en yakın olan değerler

Tablo 6 incelendiğinde, 1000 örneklem büyüklüğünde tam veri ile hesaplanan ortalama değer ile kayıp veri ile baş etme yöntemleri sonucunda oluşturulan veri seti sonuçlarının büyük oranda benzerlik gösterdiği görülmektedir. Ancak eksik veri ile hesaplanan betimsel istatistiklerin tam veri setinden uzaklaştığı belirlenmiştir. Güvenirlik katsayılarının da benzer şekilde 9 veri setinde tam veri seti ile hesaplanan sonuçlara oldukça yakın olduğu tespit edilmiştir.

1000 örneklem büyüklüğünde hırs algısı ölçeğinde yer alan maddelerin ayırıcılık ve güçlük indeksleri AMTK'ya dayalı olarak hesaplanmış; sonuçlar Tablo 7'de gösterilmiştir.

Tablo 7. Aşamalı Madde Tepki Modeline Dayalı Kestirilen Ayırıcılık ve Güçlük İndeksleri (n=1000)

Madde no	Madde parametreleri	Veri setleri									
		1	2	3	4	5	6	7	8	9	10
M1 (%20 eksik veri)	a	4,45	4,39	5,79	3,26	4,61	4,58	4,47*	3,18	4,39	6,23
	b <sub>1</sub>	-2,43	-2,45	-2,29	-2,67	-2,4	-2,4	-2,42*	-2,78	-2,45	-2,32
	b <sub>2</sub>	-1,8	-1,83	-1,77	-2,02	-1,79*	-1,79*	-1,79*	-2,1	-1,82	-1,76
	b <sub>3</sub>	-0,66	-0,69	-0,67	-0,91	-0,66*	-0,66*	-0,67	-0,61	-0,64	-0,61
M2 (%15 eksik veri)	a	3,41	3,19	3,52	2,3	3,22	3,18	3,4*	3,13	3,47	4,69
	b <sub>1</sub>	-2,52	-2,61	-2,52*	-3,01	-2,68	-2,69	-2,65	-2,71	-2,62*	-2,41
	b <sub>2</sub>	-2,01	-2,08	-2,05*	-2,4	-2,1	-2,13	-2,11	-2,2	-1,98	-1,88
	b <sub>3</sub>	-0,57	-0,58	-0,59	-0,8	-0,62	-0,63	-0,62	-0,57*	-0,56	-0,55
M3 (%10 eksik veri)	a	2,12	2,16	2,24	1,93	2,15	2,1*	2,1*	1,68	1,99	2,07
	b <sub>1</sub>	-2,91	-2,88	-3,04	-3,11	-2,93*	-2,96	-2,96	-3,42	-2,96	-2,95
	b <sub>2</sub>	-1,49	-1,51*	-1,53	-1,66	-1,52	-1,53	-1,53	-1,79	-1,56	-1,55
	b <sub>3</sub>	-0,11	-0,14	-0,16	0,04	-0,11*	-0,12	-0,12	0,03	-0,11*	-0,09
M4 (%5 eksik veri)	a	1,08	1,07	1,09	1,07	1,07	1,07	1,08*	0,87	0,99	1,03
	b <sub>1</sub>	-3	-2,99*	-3,01*	-3,07	-3,02	-3,03	-3,01*	-3,59	-3,13	-3,11
	b <sub>2</sub>	-1,17	-1,15	-1,09	-1,23	-1,18*	-1,19	-1,18*	-1,41	-1,21	-1,2
	b <sub>3</sub>	0,81	0,84	0,74	0,94	0,81*	0,81*	0,81*	1,08	0,9	0,98
M5 (%0 eksik veri)	a	2,59	2,61*	2,73	2,73	2,61*	2,67	2,61*	2,38	2,5	2,49
	b <sub>1</sub>	-2,42	-2,4	-2,27	-2,37	-2,39	-2,39	-2,39	-2,55	-2,45	-2,41*
	b <sub>2</sub>	-1,6	-1,6*	-1,53	-1,59	-1,59	-1,59	-1,59	-1,68	-1,63	-1,6*
	b <sub>3</sub>	-0,22	-0,23	-0,19	-0,23	-0,22*	-0,22*	-0,22*	-0,22*	-0,22*	-0,21

1: tam veri, 2: eksik veri, 3: silme, 4: ortalama atama, 5: yakın noktalar ortalama ataması, 6: yakın noktalar medyan ataması, 7: doğrusal değerlendirme, 8: noktada doğrusal eğilim, 9: regresyon atama, 10: beklenti maksimizasyonu

\* tam veri parametrelerine en yakın olan değerler

Tablo 7’de yer alan bilgiler doğrultusunda maddelerin ayırıcılık parametrelerinin güçlük parametrelerine göre kayıp veriden ve kayıp veri yöntemlerinden daha fazla etkilenmiş olduğu görülmektedir. Ölçekte yer alan tüm maddelerin tam veri ile hesaplanan ayırıcılık indekslerinin en yakın kayıp verileri doğrusal değerlendirme yöntemiyle atanan veri setinde hesaplandığı belirlenmiştir. %20 kayıp veriye sahip birinci maddenin ayırıcılık indeksinin doğrusal değerlendirme yöntemi ile verilerin atandığı veri seti ile farkının 0,02; ikinci maddenin (%15 kayıp verili) farkının 0,01; üçüncü maddenin (%10 kayıp verili) ayırıcılık indeksleri arasındaki farkın 0,02; dördüncü maddenin (%5 eksik verili) ayırıcılık indeksi arasındaki farkın 0,00; beşinci maddenin (%0 kayıp verili) ayırıcılık indeksleri arasındaki farkın da 0,01 olduğu belirlenmiştir.

1000 örneklem büyüklüğünde de maddelerdeki kayıp veri oranının ve madde ayırıcılık gücünün kayıp veri ile baş etme yöntemlerine etkisi olduğu tespit edilmiştir. Aynı zamanda maddede kayıp veri olmasa da diğer maddelerdeki kayıp verinin madde parametreleri üzerinde etkisi olduğu belirlenmiştir.

### 3. 2500 örneklem büyüklüğünde kayıp veri ile baş etme yöntemlerine dayalı olarak test ve madde istatistikleri nasıl bir değişim göstermektedir?

Araştırma kapsamında son olarak 2500 kişilik örneklem büyüklüğünde tam veri seti ile birlikte 10 veri seti oluşturulmuştur. Veri setlerine yönelik önce betimsel istatistikler ve güvenilirlik katsayıları hesaplanmış, sonuçlar Tablo 8’de gösterilmiştir.

Tablo 8. Betimsel İstatistikler ve Güvenirlik Katsayıları (n=2500)

Veri setleri	Ortalama	SS	Cronbach alfa güvenirlik	Marginal güvenirlik
Tam veri	17,140	2,550	0,766	0,735
Eksik veri	15,380	3,300	0,769	0,734
Silme	17,200	2,560	0,769	0,734
Ortalama atama	17,150	2,350	0,729	0,712
Yakın noktalar ortalama ataması	17,130	2,550*	0,769	0,735
Yakın noktalar medyan ataması	17,140*	2,550*	0,768	0,734
Doğrusal değerlendirme	17,140*	2,550*	0,767	0,734
Noktada doğrusal eğilim	17,150	2,370	0,730	0,725
Regresyon atama	17,130	2,540	0,766*	0,736*
Beklenti maksimizasyonu	17,140*	2,530	0,782	0,7412

\* tam veri parametrelerine en yakın olan değerler

Tablo 8’de yer alan bilgiler doğrultusunda 2500 örneklem büyüklüğünde tam veri setinde ölçek maddelerine verilen cevapların ortalaması 17,14; standart sapması 2,55 olarak hesaplanmıştır. Örneklemin Cronbach alfa güvenirlik katsayısı 0,766; marginal güvenirlik katsayısı da 0,7351 olarak hesaplanmıştır. Kayıp verilere herhangi bir müdahalede bulunulmadan veri setinin %12’si kayıp veri durumundayken yapılan hesaplama sonucunda ortalama değerler gerçek değerden daha düşük kestirildiği (-1,76 fark) belirlenmiştir. Bununla birlikte eksik verilerle farklı baş etme yöntemlerine dayalı tam hâle getirilmiş veri setleri gerçekleştirilen hesaplamalar sonucunda ortalama değer, standart sapma ve güvenirlik katsayılarının tam veri seti ile büyük ölçüde benzerlik gösterdiği tespit edilmiştir.

Araştırmada 2500 kişilik örneklem setindeki maddelerin ayırıcılık ve güçlük parametreleri aşamalı tepki modeli ile kestirilmiş, sonuçlar Tablo 9’da gösterilmiştir.

Tablo 9. Aşamalı Madde Tepki Modeline Dayalı Kestirilen Ayırıcılık ve Güçlük İndeksleri (n=2500)

Madde no	Madde parametreleri	Veri setleri									
		1	2	3	4	5	6	7	8	9	10
M1 (%20 eksik veri)	a	4,2	4,09	4,34	2,82	4,18	4,17	4,19*	3,59	4,18	5,58
	b <sub>1</sub>	-2,22	-2,23*	-2,24	-2,54	-2,23*	-2,23*	-2,23*	-2,39	-2,27	-2,16
	b <sub>2</sub>	-1,74	-1,78	-1,71	-2,07	-1,74	-1,74*	-1,75	-1,96	-1,8	-1,7
	b <sub>3</sub>	-0,62	-0,63	-0,66	-0,88	-0,62*	-0,62*	-0,62*	-0,57	-0,58	-0,57
M2 (%15 eksik veri)	a	3,55	3,23	3,39	2,42	3,58	3,49	3,57*	3,37	3,4	4,57
	b <sub>1</sub>	-2,4	-2,47*	-2,42	-2,81	-2,43	-2,42*	-2,42*	-2,53	-2,49	-2,33
	b <sub>2</sub>	-1,91	-1,97	-1,9*	-2,25	-1,92*	-1,92*	-1,93	-2,05	-1,89	-1,81
	b <sub>3</sub>	-0,57	-0,59	-0,58	-0,8	-0,58	-0,59	-0,58	-0,57*	-0,54	-0,53
M3 (%10 eksik veri)	a	1,8	1,8*	1,8*	1,66	1,81	1,8*	1,8*	1,51	1,7	1,79
	b <sub>1</sub>	-2,93	-2,94*	-2,95	-3,15	-2,98	-3	-3,01	-3,35	-3,04	-2,98
	b <sub>2</sub>	-1,62	-1,62*	-1,6	-1,78	-1,63	-1,64	-1,64	-1,88	-1,67	-1,66
	b <sub>3</sub>	-0,09	-0,09*	-0,14	0,1	-0,09*	-0,09*	-0,09*	0,11	-0,07	-0,04
M4 (%5 eksik veri)	a	0,99	1,03	1,07	1,03	0,99*	0,99*	0,99*	0,91	1,01	0,98
	b <sub>1</sub>	-3,44	-3,34	-3,34	-3,39	-3,45*	-3,45*	-3,47	-3,73	-3,41	-3,49
	b <sub>2</sub>	-1,41	-1,39	-1,36	-1,46	-1,41*	-1,41*	-1,41*	-1,59	-1,39	-1,46
	b <sub>3</sub>	0,84	0,79	0,71	0,89	0,84*	0,84*	0,84*	0,97	0,81	0,95
M5 (%0 eksik veri)	a	2,32	2,35	2,4	2,39	2,31	2,32*	2,32*	2,13	2,23	2,21
	b <sub>1</sub>	-2,54	-2,55	-2,48	-2,52	-2,54*	-2,54*	-2,54*	-2,66	-2,59	-2,57
	b <sub>2</sub>	-1,63	-1,64	-1,61	-1,63*	-1,64	-1,64	-1,63*	-1,71	-1,66	-1,64
	b <sub>3</sub>	-0,15	-0,16	-0,18	-0,16	-0,15*	-0,15*	-0,15*	-0,15*	-0,15*	-0,13

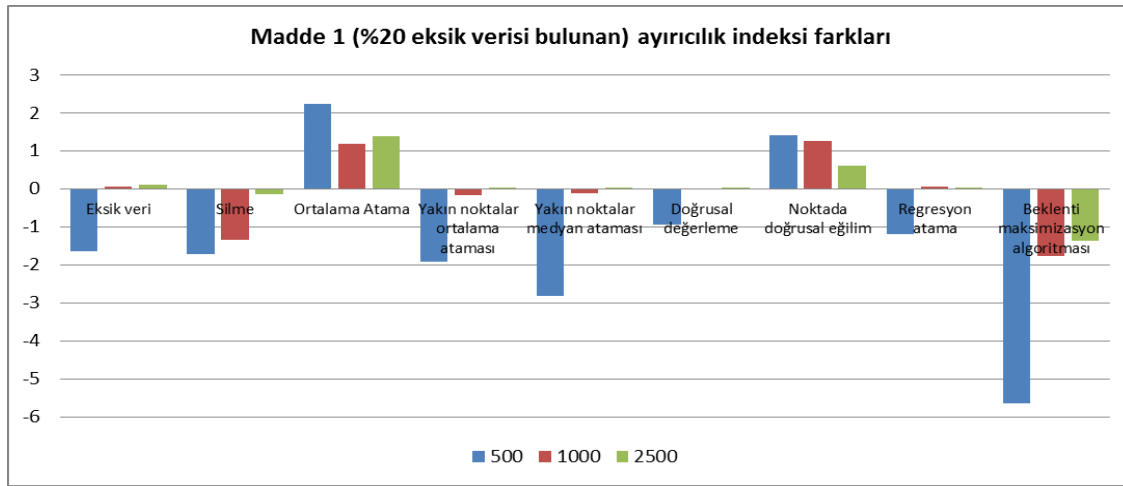
1: tam veri, 2: eksik veri, 3: silme, 4: ortalama atama, 5: yakın noktalar ortalama ataması, 6: yakın noktalar medyan ataması, 7: doğrusal değerlendirme, 8: noktada doğrusal eğilim, 9: regresyon atama, 10: beklenti maksimizasyonu

\* tam veri parametrelerine en yakın olan değerler

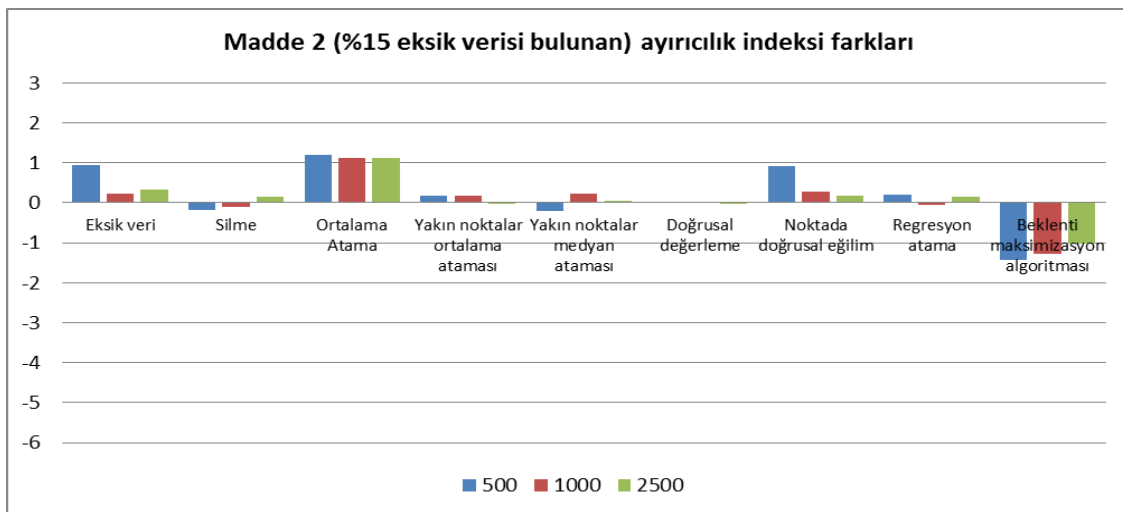
Tablo 9’da yer alan bilgiler doğrultusunda 2500 örneklem büyüklüğünde tam veri setinde hesaplanan madde ayırıcılık ve madde güçlük indekslerinin eksik veri setinde ve 8 kayıp veri ile baş etme yöntemleri ile elde edilen veri setlerinde büyük ölçüde benzerlik gösterdiği; farkların oldukça düşük olduğu tespit edilmiştir. Diğer örneklem büyüklüklerine benzer şekilde, tam veri setinden hesaplanan ayırıcılık indekslerine en yakın olan değerlerin doğrusal değerlendirme kayıp veri atama yöntemi ile elde edilen veri setlerine ait olduğu belirlenmiştir. Birinci maddede (%20 eksik verili) farkın 0,01; ikinci maddede (%15 eksik verili) farkın 0,02; üçüncü maddede (%10 eksik verili), dördüncü maddede (%5 eksik verili) ve beşinci maddede (%0 eksik verili) ise farkın olmadığı (0,00), madde ayırıcılık indeksinin bire bir aynı hesaplandığı belirlenmiştir.

#### 4. Farklı örneklem büyüklüklerinde kayıp veri ile baş etme yöntemlerine dayalı olarak hesaplanan madde ayırıcılık indeksi ve tam veri setinden hesaplanan madde ayırıcılık indeksi farkları nedir?

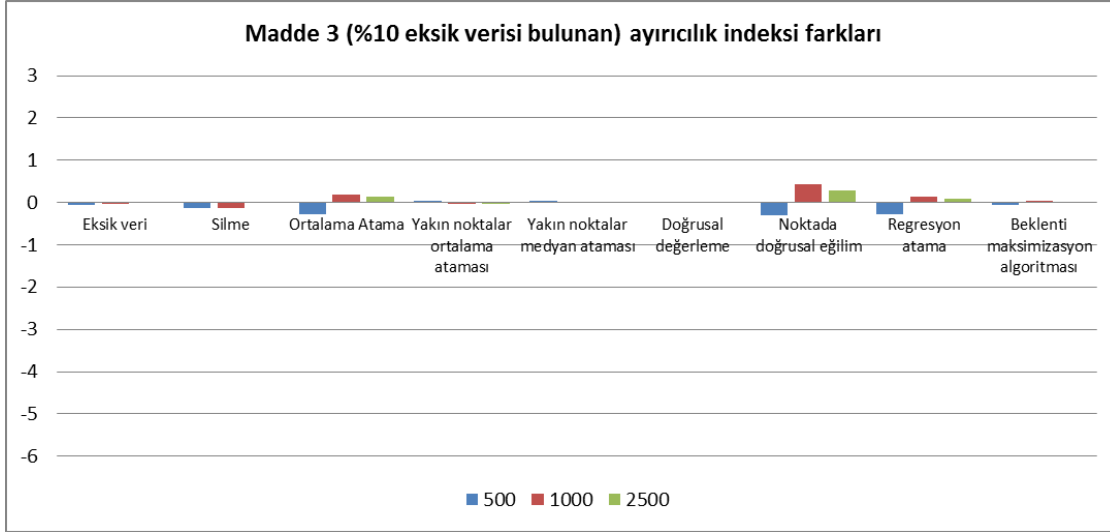
Araştırmada kayıp veriden ve kayıp veri ile baş etme yöntemlerinden en çok etkilenen değerlerin madde ayırıcılık parametresi olduğu belirlenmiştir. Bu doğrultuda her bir madde için eksik veri seti ve kayıp veri baş etme yöntemlerine dayalı oluşturulmuş veri setlerinin ayırıcılık indekslerinin tam veri setinde hesaplanan değerle arasındaki fark hesaplanmıştır; sonuçlar Şekil 2-Şekil 6’da gösterilmiştir.



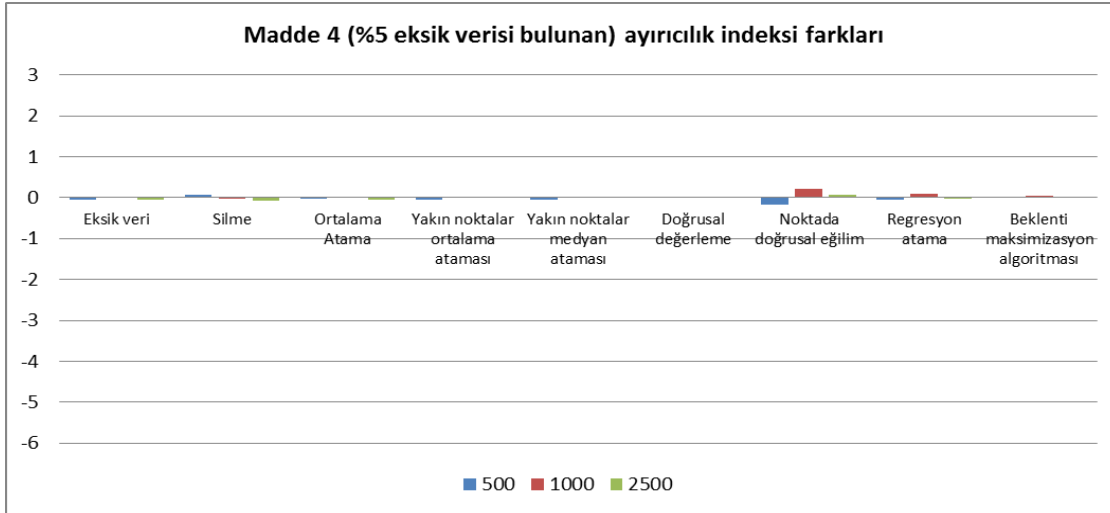
Şekil 2. Madde 1 (%20 eksik verili) için Hesaplanan Ayırıcılık İndeksleri Arasındaki Farklar



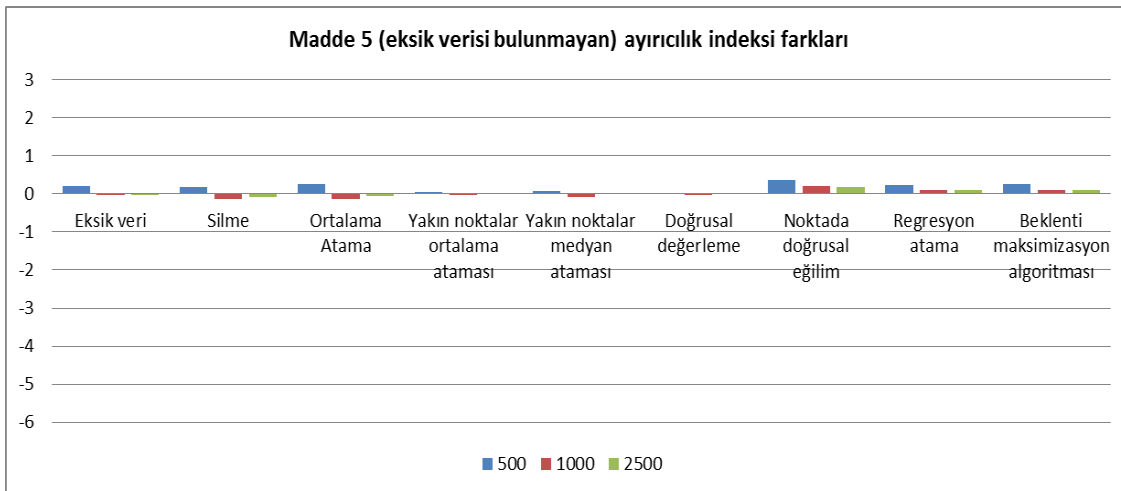
Şekil 3. Madde 2 (%15 eksik verili) için Hesaplanan Ayırıcılık İndeksleri Arasındaki Farklar



Şekil 4. Madde 3 (%10 eksik verisi) için hesaplanan ayırıcılık indeksleri arasındaki farklar



Şekil 5. Madde 4 (%5 eksik verili) için hesaplanan ayırıcılık indeksleri arasındaki farklar



Şekil 6. Madde 5 (eksik verisi bulunmayan) için hesaplanan ayırıcılık indeksleri arasındaki farklar

Şekil 2-6 incelendiğinde tüm maddelerde örneklem büyüklüğüne bağlı olarak tam veri ile eksik veri ve kayıp veri ile baş etme yöntemlerine dayalı olarak oluşturulan veri setlerinde hesaplanan ayırıcılık indeksleri arasındaki farkların azalma gösterdiği belirlenmiştir. Maddelerdeki kayıp veri oranları azalma gösterdikçe ayırıcılık indeksleri arasındaki farkın da azalma gösterdiği tespit edilmiştir. Bununla birlikte ölçek maddeleri içerisinde en yüksek ayırıcılığa birinci madde; en düşük ayırıcılığa ise dördüncü madde sahiptir. Düşük ayırıcılıktaki maddeler arasındaki fark da ranjı daha küçük olduğu için daha az hesaplanmıştır. Beşinci maddede hiç kayıp veri olmamasına rağmen madde ayırıcılık indekslerinde ufak da olsa farklılaşmalar saptanmıştır.

## SONUÇLAR ve TARTIŞMA

Bu araştırma kapsamında PISA 2015 uygulamasına katılan ve çalışma içerisinde yer alan “hırs algısı ölçek maddelerine” cevap veren 5073 öğrenci arasından rastgele 500, 1000 ve 2500 büyüklüğünde örneklem seçilmiştir. Hırs algısı ölçeğinde tek boyutta toplandığı belirlenen 4’lü likert tipinde derecelendirilmiş 5 madde yer almaktadır. Her bir örneklem büyüklüğünde birinci maddenin %20’si, ikinci maddenin %15’i, üçüncü maddenin %10’u ve dördüncü maddenin %5’ine denk gelen veriler silinmiştir. Beşinci madde ise eksik veri içermeyecek şekilde aynen bırakılmıştır. Allison (2002) ile Tabachnick ve Fidell (2014) kayıp verinin TRK olduğu durumda, veri setindeki kayıp veri oranının %5 ve altında olması hâlinde test ve madde istatistiklerinde önemli sorunlar yaşanmayacağını, ancak kayıp veri oranının %5’in üzerine çıkması durumunda sonuçların yanıltıcı olarak kestirilebileceğini belirtmişlerdir. Bu doğrultuda bu çalışmada ele alınan veri setlerinde %12’lik kayıp veri oluşturulmuştur. Veri setlerindeki eksik verilerin mekanizmasının tamamen rastsal olduğu Little’in MCAR testi ile test edilmiştir. Her bir örneklem büyüklüğünde tam veri, eksik veri ve silme, ortalama atama, yakın noktalar ortalama ataması, yakın noktalar medyan ataması, doğrusal değerlendirme, noktada doğrusal değerlendirme, regresyon atama ve beklenti maksimasyonu algoritması atama yöntemleriyle eksik verilerin tamamlanması sonucunda 10’ar olmak üzere toplam 30 veri seti oluşturulmuştur. Veri setlerinde önce test istatistikleri (ortalama, standart sapma, Cronbach alfa ve marginal güvenilirlik katsayıları) ardından madde ayırıcılık ve madde güçlük parametreleri Aşamalı Tepki Modeline dayalı olarak hesaplanmıştır.

Hesaplamalar sonucunda eksik veri ve eksik verilerle baş etme yöntemlerine dayalı tam hâle getirilmiş veri setlerine ait madde istatistiklerinin test istatistiklerine göre daha fazla etkilendiği belirlenmiştir. Tam veri setleri ölçüt (referans) alınarak karşılaştırmalar yapılmış; eksik verilerle baş etme yöntemlerine dayalı tam hâle getirilmiş veri setlerinde test istatistiklerinin tam veri seti ile büyük ölçüde benzerlik gösterdiği saptanmıştır. Ancak eksik veri seti ile hesaplanan ortalama değerlerin tam veri setinden elde edilen ortalama değerinden düşük olduğu belirlenmiştir. Araştırma sonuçları ve bu bilgiler doğrultusunda kayıp veri ile baş etme yöntemlerinin testin betimsel istatistiklere olan etkisinin yordayıcı istatistiklere olan etkisinden daha az olduğu söylenebilir. Eksik veri ve kayıp veri ile baş etme yöntemlerinin kullanılmasıyla elde edilen veri setlerine ait madde parametreleri incelendiğinde, madde ayırıcılık indeksinin madde güçlük indekslerinden daha fazla etkilendiği tespit edilmiştir. Maddedeki kayıp veri oranının ve maddenin ayırıcılık indeksinin ayrıca kayıp veri ile baş etme yöntemleri üzerinde etkisi olduğu belirlenmiştir. Koçak ve Çokluk Bökeoğlu (2017) çalışmalarında MTK’ya dayalı test istatistiklerini karşılaştırmış, tüm setlerde kayıp veri oranı artış gösterdikçe kayıp veri ile baş etme yöntemleri sonuçları arasındaki farkın artış gösterdiğini tespit etmiştir.

Bu çalışmada madde parametreleri Aşamalı Tepki Modeline dayalı olarak kestirilmiştir. Ölçeğin beşinci maddesinde hiç kayıp veri bulunmamasına karşın madde ayırıcılık ve güçlük indeksinin veri setleri bazında değişiklik gösterdiği saptanmıştır. Bu durum, eksik verisi bulunan maddelerin diğer maddelerin parametrelerine de etki ettiğini göstermektedir. Bu nedenle araştırmacıların kayıp veri ile karşılaştığı durumda veri setini inceleyerek kayıp verisi için en uygun baş etme yöntemini kullanması gerekmektedir. Araştırma sonucu dikkate alınarak eksik verisi olmayan maddelerin de süreçten etkilendiği göz önünde bulundurulmalıdır. Garrett (2009) yapmış olduğu çalışmada



değişen madde fonksiyonun kayıp veri olduğu durumda referans ve odak grup kestirimlerinin benzerlik gösterdiğini belirlemiştir. Ergün (2013) 0-1 veri setiyle çalıştığı araştırmasında kayıp veri yöntemlerinin madde ayırıcılık ve güvenilirlik katsayılarına etki ettiğini; en çok olabilirlik ve çoklu veri atama yöntemlerinde madde parametre kestirimlerinin benzer sonuçlar ürettiğini tespit etmiştir.

Örneklem büyüklüğü artış gösterdikçe kayıp veri ile baş etme yöntemlerinin madde ve test parametreleri üzerindeki etkisinin azaldığı, gerçek değere oldukça yaklaştığı belirlenmiştir. Soysal ve Akın Arıkan (2017) tamamen rastsal kayıp verilerle gerçekleştirdikleri açıklayıcı faktör analizi sonucunda örneklem büyüklüğünün artış gösterdikçe hata oranının da azalma gösterdiğini tespit etmişlerdir.

Araştırma sonucunda her bir örneklem büyüklüğünde tam veri setine en yakın madde ayırt edicilik indeksi sunan sonuçların doğrusal değerlendirme yöntemi ile elde edildiği belirlenmiştir. Kayıp verinin kendisinden bir önceki ve bir sonraki veri ortalaması alınarak hesaplandığı bu yöntemin genellikle boylamsal veriler ya da ardışık kategorilerin olduğu veri setlerinde kullanıldığı belirlenmiştir. Bu araştırmalarda kayıp verilerle baş etme amacıyla doğrusal değerlendirme yöntemi kullanılmış ve tam veri setine yakın sonuçlar elde edilmiştir. (Juninen, Niska, Tuppurainen, Ruuskanen ve Kolehmainen, 2004; Kohn ve Ansley, 1986; Twisk ve de Vente, 2002). Yapılan bu çalışmada da likert tipinde derecelendirilmiş bir ölçek kullanılmıştır ve ölçek dereceleri ardışıklık özelliği göstermektedir. Aşamalı Tepki Modeli ile kestirilen eşik parametresi (bij), bireylerin örtük özellik boyunca 0,50 olasılıkla bir kategori eşliğinin üzerinde tepki verme olasılığını tanımlar. bij, aynı zamanda bireylerin örtük özellik boyutunda o kategoride ya da daha üst kategorilerde 0,50 tepki verme olasılığının da bir tanımıdır ve bu nedenle eşik parametreleri sıralı özellik göstermektedir. Eğitim parametresi olan  $a_i$  de sıralı kategorilerin tamamı için kestirilmektedir (Emretson ve Reise, 2000; Ostini ve Nering, 2006). Bu bağlamda araştırmada kayıp gözlemlere doğrusal değerlendirme yöntemi ile atanan veriler sonucunda hesaplanan ayırıcılık ve güçlük değerlerinin tam veri setine yakın hesaplanması, cevap kategorilerinin kendi içinde ardışıklık göstermesinden kaynaklanmış olabilir. Bu doğrultuda başka çalışmalarda bu yöntemin yinelenmesi, iki kategorili veri setlerinde de uygulanarak sonuçlarının karşılaştırılması önerilmektedir.

Araştırma kapsamında doğrusal değerlendirme dışında da kayıp veri ile baş etme yöntemleri kullanılmış ve tam veri seti ile, özellikle küçük örneklem büyüklüklerinde, benzer olmayan ayırıcılık indeksleri hesaplanmıştır. Çokluk ve Kayri (2011) de yaklaşık değer atama yöntemlerinin test maddelerinin açıklanan varyans oranında ve güvenilirlikte düşüş meydana getirdiğini belirlemiştir. Bernaards ve Sijtsma (2000) kayıp veri yöntemleri içerisinde beklenti maksimizasyonu yönteminin diğer yöntemlerden daha iyi sonuç verdiği bilgisine ulaşmıştır. Sezgin ve Çelik (2013) araştırmalarında kayıp veri ile baş etme sürecinde durum düzeyinde silme yönteminin kullanılmasının basit olmasına rağmen veri kaybı yaşattığı için varyansı değiştirdiğini; regresyon atamasının da özellikle ilişkisizlik durumunda hata oluşumuna neden olduğunu belirtmektedir. Garlett (2009) de silme yöntemi ile değişen madde fonksiyonunun farklılık gösterdiğini ortaya koymuştur. Köse ve Öztumur (2014) da araştırmalarında silme yönteminin özellikle t testi sonuçlarına etki ettiğini, sonuçların yanlış kestirilmesine neden olduğunu ortaya koymuşlardır. Acock (2005), silme yönteminin TRK dışında özellik gösteren kayıp veriler için hiç uygun olmadığını; TRK özellik gösteren kayıp veri setlerinde de özellikle küçük örneklem büyüklüklerinde testin istatistiksel gücünde bir azalma meydana getirdiğini ifade etmektedir. Nartgün (2015) araştırmasında silme yönteminin, tüm koşullarda, tam veri setinden elde edilen değere en uzak değerleri verdiğini, en yakın değerlerin ise çoklu veri atama ve regresyon ataması yöntemleriyle sağlandığını belirlemiştir. Bu çalışmada örneklem büyüklüğü artış gösterdikçe kayıp veri ile baş etme yöntemleri arasındaki farklar benzerlik gösterse de tüm örneklem büyüklüğünde ve kayıp veri oranlarında tam veri setine en yakın sonuçların doğrusal değerlendirme yönteminin kullanıldığı koşullarda ulaşıldığı tespit edilmiştir. Diğer araştırma bulgularına benzer şekilde diğer eksik verilerle baş etme yöntemlerinin tam veri seti ile benzerlik gösterdiği; ancak söz konusu çalışmaların çoğunda doğrusal değerlendirme yöntemi kullanılmadığı için başka çalışmalarda da doğrusal değerlendirme yönteminin incelenmesi ve özellikle bu yöntemin küçük örneklem gruplarından elde edilen veri setleri üzerinde kullanılması önerilmektedir.

Bu araştırmada sekiz farklı kayıp veri ile baş etme yönteminin madde ve test parametreleri kestirimindeki performansları karşılaştırılmıştır. Başka çalışmalarda eksik veriler yerine atama ya da silme yöntemleri yerine kayıp verilerin de analizin içine dahil edildiği tam bilgi en çok olabilirlik (full information maximum likelihood) gibi yöntemler ile de hesaplamalar yapıp elde edilen sonuçlar karşılaştırılabilir.

## KAYNAKÇA

- Acar, T. ve Kelecioğlu, H. (2010). Maddenin farklı fonksiyonlaşmasını belirleme tekniklerinin karşılaştırılması: GADM, LR ve MTK-OO. *Kuram ve Uygulamada Eğitim Bilimleri*, 10(2), 639-649.
- Acock, A. (2005). Working with missing values. *Journal of Marriage and Family*, 67(4), 1012-1028.
- Akbaş, U. ve Tavşancıl, E. (2015). Farklı örneklem büyüklüklerinde ve kayıp veri örüntülerinde ölçeklerin psikometrik özelliklerinin kayıp veri baş etme teknikleri ile incelenmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 6(1), 38-57.
- Allison, P. D. (2002). *Missing data*. California: Sage Publication, Inc.
- Ambler, G., Omar, R. Z., & Royston, P. (2007). A comparison of imputation techniques for handling missing predictor values in a risk model with a binary outcome. *Statistical Methods in Medical Research*, 16(3), 277-298.
- Bal, C. (2003). *Çok gruplu veri setlerinde eksik gözlem sorununun çözümlenmesi ve sağlık alanında bir uygulama* (Yayımlanmamış Yüksek Lisans Tezi, İstanbul Üniversitesi, Sağlık Bilimleri Enstitüsü, İstanbul). Erişim adresi: <http://tez2.yok.gov.tr>
- Baygül, A. (2007). *Kayıp veri analizinde sıklıkla kullanılan etkin yöntemlerin değerlendirilmesi*. (Yayımlanmamış Doktora Tezi, Osmangazi Üniversitesi, Sağlık Bilimleri Enstitüsü, Eskişehir). Erişim adresi: <http://tez2.yok.gov.tr>
- Bernaards, C. A., & Sijtsma, K. (2000). Influence of imputation and EM methods on factor analysis when item nonresponse in questionnaire data is nonignorable. *Multivariate Behavioral Research*, 35(3), 321 – 364.
- Carpita, M., & Manisera, M. (2011). On the imputation of missing data in surveys with Likert-type scales. *Journal of Classification*, 28(1), 93-112.
- Çokluk, Ö., Şekercioğlu, G. ve Büyüköztürk, Ş. (2014). *Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL uygulamaları*. Ankara: Pegem Akademi.
- Çokluk, Ö. ve Kayri, M. (2011). Kayıp değerlere yaklaşık değer atama yöntemlerinin ölçme araçlarının geçerlik ve güvenilirliği üzerindeki etkisi. *Kuram ve Uygulamada Eğitim Bilimleri*, 11(1), 289-309.
- Çüm, S. ve Gelbal, S. (2015). Kayıp veriler yerine yaklaşık değer atamada kullanılan farklı yöntemlerin model veri uyumu üzerindeki etkisi. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi*, 35, 87-111.
- De Ayala, R. J., Plake, B. S., & Impara, J. C. (2011). The impact of omitted responses on the accuracy of ability estimation in item response theory. *Journal of Educational Measurement*, 38(3), 213-234.
- De Leeuw, E. D., Hox, J., & Huisman, M. (2003). Prevention and treatment of item nonresponse. *Journal of Official Statistics*, 19(2), 153-176.
- Demir, E. (2013). Kayıp verilerin varlığında çoktan seçmeli testlerde madde ve test parametrelerinin kestirilmesi: SBS örneği. *Eğitim Bilimleri Araştırmaları Dergisi*, 3(2), 48-68.
- Demir, E. (2013). *Kayıp verilerin varlığında iki kategorili puanlanan maddelerden oluşan testlerin psikometrik özelliklerinin incelenmesi*. (Yayımlanmamış Doktora Tezi, Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara). Erişim adresi: <http://tez2.yok.gov.tr>
- Demir, E. ve Parlak, B. (2012). Türkiye’de eğitim araştırmalarında kayıp veri sorunu. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 3(1), 230-241.
- Doğanay-Erdoğan, B. (2012). *Çoklu atama yöntemlerinin Rasch modelleri için performansının benzetim çalışması ile incelenmesi*. (Yayımlanmamış Doktora Tezi, Ankara Üniversitesi, Sağlık Bilimleri Enstitüsü, Ankara). Erişim adresi: <http://tez2.yok.gov.tr>
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Field, A. (2005). *Discovering statistics with SPSS*. California: Sage Publication, Inc.
- Finch, H. (2008). Estimation of item response theory parameters in the presence of missing data. *Journal of Educational Measurement*, 45(3), 225-245.
- Garrett, P. L. (2009). *A Monte Carlo study investing ating missing data, differential item functioning and effect size* (Yayımlanmamış doktora tezi, Georgia State University, Atlanta). Erişim adresi: <https://scholarworks.gsu.edu>
- Garson, D. (2015). *Missing data analysis & data imputation*. Asheboro: Statistical Publishing Associates.

- 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.
- Hambleton, R. K., Swaminathan, H., Cook, L. L., Eignor, D. R., & Gifford, J. A. (1977). Developments in latent trait theory: models, technical issues, and applications. *Review of Educational Research*, 48(4), 467-510.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. California: Sage Publication, Inc.
- Hedeker, D., Mermelstein, R. J., & Demirtas, H. (2007). Analysis of binary outcomes with missing data: Missing = smoking, last observation carried forward, and a little multiple imputation. *Addiction*, 102(10), 1564-1573.
- Hohensinn, C., & Kubinger K. D. (2011). On the impact of missing values on item fit and the model validness of the Rasch model. *Psychological Test and Assessment Modeling*, 53(3), 380-393.
- Josse, J., & Husson, F. (2012). Handling missing values in exploratory multivariate data analysis methods. *Journal de la Société Française de Statistique*, 153(2), 79-99.
- Junninen, H., Niska, H., Tuppurainen, K., Ruuskanen, J., & Kolehmainen, M. (2004). Methods for imputation of missing values in air quality data sets. *Atmospheric Environment*, 38, 2895-2907.
- Koçak, D. ve Çokluk Bökeoğlu, Ö. (2017). Kayıp veriyle baş etme yöntemlerinin model veri uyumu ve madde model uyumuna etkisi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 8(2), 200-223.
- Kohn, R., & Ansley, C. F. (1986). Estimation, prediction, and interpolation for ARIMA models with missing data. *Journal of the American Statistical Association*, 81, 751-761.
- Köse, A. ve Öztumur, B. (2014). Kayıp veri ele alma yöntemlerinin t-testi ve ANOVA parametreleri üzerine etkisinin incelenmesi. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 14(1), 400-412.
- McNeish, D. (2016). Exploratory factor analysis with small samples and missing data. *Journal of Personality Assessment*, 99(6), 637-652.
- Nartgün, Z. (2015). Kayıp veri sorununun çözümünde kullanılan farklı yöntemlerin farklı kayıp veri koşulları altında ölçeklerin psikometrik nitelikleri ve ölçme sonuçları bağlamında karşılaştırılması. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 7(4), 252-265.
- Ostini, R., & Nering, M. L. (Eds.). (2006). *Polytomous item response theory models*. U.S.A: Sage Publication.
- Peng, C. Y., Harwell, M. R., Liou, S. M., & Ehman, L. H. (2007). Advances in missing data methods and implications for educational research. In S. S. Sawilowsky (Ed.), *Real data analysis* (pp. 31-78). Charlotte, NC: New Information Age.
- Rose N., von Davier, M., & Xu, X. (2010). *Modeling nonignorable missing data with item response theory (IRT)* (Report No: RR-10-11). Princeton: ETS Research Report Series.
- Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7(2), 147 - 177
- Sezgin E. ve Çelik Y. (2013). *Veri madenciliğinde kayıp veriler için kullanılan yöntemlerin karşılaştırılması*, Akademik Bilişim Konferansı, Akdeniz Üniversitesi, Antalya.
- Soysal, S. ve Akın Arıkan, Ç. (2017). Kayıp veri atama yöntemlerinin faktörleşme teknikleri üzerindeki etkisi. PegemA Yayıncılık (Ed.), *Küreselleşen Dünyada Eğitim*, (syf. 316-332) Ankara: PegemA Yayıncılık.
- Şahin Kürşad, M. ve Nartgün, Z. (2015). Kayıp veri sorununun çözümünde kullanılan farklı yöntemlerin ölçeklerin geçerlik ve güvenilirliği bağlamında karşılaştırılması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 6(2), 254-267.
- Tabachnick, B. G., & Fidell, L. S. (2014). *Using multivariate statistics*. USA: Pearson Education Limited.
- Tsutakawa, R. K., & Johnson, J. C. (1990). The effect of uncertainty of item parameter estimation on ability estimates. *Psychometrika*, 55, 371-390.
- Twisk, J., & de Vente, W. (2002). Attrition in longitudinal studies: How to deal with missing data. *Journal of Clinical Epidemiology*, 55, 329-337.
- Weaver, B., & Maxwell, H. (2014). Exploratory factor analysis and reliability analysis with missing data: A simple method for SPSS users. *The Quantitative Methods for Psychology*, 10(2), 143-152.
- Witta, E. L. (2000). *Comparison of missing data treatments in producing factor scores* (Report No:1). Honolulu: Annual Meeting of the American Educational Research Association.
- Wright, B. D. (1997). A History of social science measurement. *Educational measurement: Issues and practice*, 16, 33-45.
- Yılmaz, H. (2014). *Random forests yönteminde kayıp veri probleminin incelenmesi ve sağlık alanında bir uygulama*. (Yayımlanmamış Yüksek Lisans Tezi, Eskişehir Osmangazi Üniversitesi, Sağlık Bilimleri Enstitüsü, Eskişehir). Erişim adresi: <http://tez2.yok.gov.tr>

## EXTENDED ABSTRACT

### *Introduction*

One of the major problems in the literature in social and behavioural sciences is missing data. Since large samples are needed especially in items response theories, missing data can be seen in the data sets. In this study, item and test parameters were estimated according to the graded response theory, which is one of the item response theory (IRT) models. In the cases of missing data, researchers generally prefer to listwise or pairwise deletion techniques for the missing data cases from the analysis. However, this situation effects the validity and reliability of the results (Akbaş ve Tavşancıl, 2015; Baygül, 2007; Çüm ve Gelbal, 2015; Kaspar, 2011; Kürşad, 2014). And also different methods have been developed in order to solve the missing data problem in the studies so there are a lot of techniques of handling missing data.

Unlike other studies, in this study the missing data rate is not based on the test, but it is item-based. In this direction, the effect of the amount of missing data on any item to other non-missing item parameters is examined. The other importance of the study is that this study is made on the basis of items and the estimations are done by using the graded response model.

So in this study, the aim is to determine how the item and test parameters are affected by the missing data techniques for different sample sizes and different items with different the missing data rates.

### *Method*

PISA data was used to achieve the purpose of the study. In the PISA study, there is an ambition perception scale with 5 items. All items are rated in 4-point likert type and there isn't any reverse item. In the study in 2015, there are a total of 5073 Turkish students responding to all of the scale items. 500, 100 and 2500 students randomly selected from 5073 students. First of all, the assumptions of normality, uni-dimensionality, local independence and model-data fit are examined for each data set. Afterwards, 5%, 10%, 15% and 20% missing data were formed from for the items of the scale and there is no missing data in one (fifth) item. Tabachnick and Fidell (2014) indicate that the missing data above 5% of the test will lead to biased estimates. In the scope of the research, 12% of data sets have missing data.

First, whether the consistency of data had missing completely at random (MCAR) mechanism was also controlled with Little's MCAR test. Then 10 data sets were obtained for each sample. Complete data, incomplete data and 8 data sets which were obtained by missing data techniques including series mean, median of nearby points, mean of nearby points, linear interpolation, linear trend at point, regression.

For the analysis, descriptive statistics and Cronbach alpha reliability coefficient and marginal reliability coefficient were estimated for each sample. And then the threshold parameters and the difficulty indices were estimated according to the graded response theory from the item response theory models. Estimates were done primarily for each sample size, the parameters of the items with different missing data rates were compared. The results were then compared in terms of sample sizes.

### *Results and Discussion*

As a result of the study, it was determined that the item statistics were affected from missing data and missing data techniques, which was more than the test statistics. When only the listwise deletion method was used, the mean value was estimated lower than the actual value.

It was found that the item threshold parameters were affected from missing data and missing data techniques for each sample size, which was more than the item difficulty indices. Sample size, missing data techniques and value of the threshold index had an effect on the threshold index.

In this study, item parameters were estimated based on the item response theory. Item response theory assumes that the item parameters are independent of each other. But this study showed that last items which belong to complete data (no missing data) were affected by other items and missing data techniques.

In the study, it was determined that as the sample size increased, the difference between item parameters in terms of the missing data techniques decreased. For this reason, the missing data technique is not important in large samples (ex.  $n=2500$ ).

The results of the study showed that the item and test parameters were influenced by incomplete and missing data and missing data techniques; it was determined that the best estimation results were obtained by linear interpolation method with different data.

In this study, the results of 8 different missing data techniques were compared. For the future research, the full information maximum likelihood method can also be used and the results can be compared. By using full information maximum likelihood method, missing values are not replaced or imputed, but the missing data is handled within the model.