



HACETTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Eğitim Bilimleri Ana Bilim Dalı
Eğitimde Ölçme ve Değerlendirme Programı

KAYIP VERİ İLE BAŞ ETME YÖNTEMLERİNİN GÜVENİRLİK KESTİRİMLERİ
ÜZERİNE ETKİSİ

İzel BAYRAM

Yüksek Lisans Tezi

Ankara, 2020

Liderlik, arařtırma, inovasyon, kaliteli eđitim ve deđiřim ile

Daha ileriye ... En İyiyeye ...



HACETTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Eğitim Bilimleri Ana Bilim Dalı
Eğitimde Ölçme ve Değerlendirme Programı

KAYIP VERİ İLE BAŞ ETME YÖNTEMLERİNİN GÜVENİRLİK KESTİRİMLERİ
ÜZERİNE ETKİSİ

COMPARISON OF INFLUENCE OF THE MISSING DATA HANDLING METHODS
ON RELIABILITY ESTIMATION

İzel BAYRAM

Yüksek Lisans Tezi

Ankara, 2020

Öz

Bu çalışmada farklı test uzunluklarında, farklı örneklem büyüklüklerine sahip veri setlerinden farklı oranlarda kayıp veri içeren veri setleri oluşturularak, bu veri setlerine uygulanan kayıp veri yöntemlerinin güvenilirlik kestirimi üzerine etkisi incelenmiştir. Bunun için 20 ve 40 soruluk 200, 500 ve 1000 kişilik örnekleme sahip yapay veri setleri oluşturulmuştur. Her bir veri setinin Cronbach alfa ve MTK temelli ampirik güvenilirlik katsayısı hesaplanıp referans değerler olarak kullanılmıştır. Sonrasında veri setleri, tamamen rassal olarak kayıp mekanizmasına sahip olacak şekilde %5, %10 ve %20 oranlarında eksiltiilip yeni veri setleri oluşturulmuştur. Oluşan veri setleri dizin silme (DS), sıfır atama, regresyon atama (RA), çoklu atama (ÇA) ve beklenti maksimizasyonu (BM) yöntemleriyle tamamlanmış ve güvenilirlik katsayıları kestirilmiştir.

Yapılan analizler sonucunda referans değerler ile farklı kayıp veri yöntemleriyle tamamlanmış veri setlerinin güvenilirlikleri arasında anlamlı bir farklılık bulunmadığı, ancak referans değerlere en yakın sonuçları regresyon ve çoklu atama yöntemlerinin verdiği görülmüştür. Sıfır atama yönteminin referans değerlere en uzak kestirimleri verdiği elde edilen sonuçlar arasındadır.

Anahtar sözcükler: kayıp veri, regresyon atama, çoklu atama, beklenti maksimizasyonu, sıfır atama, güvenilirlik

Abstract

The purpose of this study was to compare the missing data handling methods in terms of the influence on reliability coefficients under different simulation conditions. For this purpose, datasets with 200, 500, and 1000 sample sizes, 20 and 40 items were generated, and study was conducted with those simulated datasets. Cronbach alpha and IRT based reliability coefficients of each dataset was calculated and used as reference values. Afterwards, datasets were reduced by %5, %10 and %20 as completely random and generated new datasets. Using listwise deletion (LD), zero imputation, regression imputation (RI), multiple imputation (MI), and expectation maximization (EM) missing data handling methods were applied to all datasets. Cronbach alpha and IRT based empirical reliability coefficients of each dataset were calculated and compared with the reference values.

As a result of the study, there was no statistically significant difference between reliability coefficients. The methods which produced the closest values to the reference values were RI and MI methods.

Keywords: missing data, regression imputation, multiple imputation, expectation maximization, zero imputation, reliability

Teşekkür

Yüksek lisans ders ve tez döneminde değerli öğretileri ve fikirleriyle bana her aşamada yardımcı olan başta çok değerli tez danışmanım ve hocam Doç. Dr. Burcu ATAR'a, sonsuz teşekkürlerimi sunarım.

Değerli görüş ve önerileriyle tezimin son halini almasına katkıda bulunan Doç. Dr. Celal Deha DOĞAN ve Dr. Öğrt. Üyesi Sevda ÇETİN'e katkılarından dolayı teşekkür ederim.

Yüksek lisans öğrenimim süresinde bana kattıklarından dolayı başta Prof. Dr. Selahattin GELBAL olmak üzere tüm hocalarıma ve tez sürecimde her aşamada bana yardımını eksik etmeyen Arş. Gör. Mehmet Can DEMİR'e teşekkürlerimi sunarım.

Her zaman yanımda olan, bana olan güvenini ve desteğini her zaman hissettiğim ve bundan güç aldığım eşim Hasan BAYRAM'a ve tüm aileme sonsuz teşekkürlerimi sunarım.

İçindekiler

Öz.....	ii
Abstract.....	iii
Teşekkür.....	iv
Tablolar Dizini.....	vii
Şekiller Dizini.....	viii
Simgeler ve Kısaltmalar Dizini.....	ix
Bölüm 1 Giriş.....	1
Problem Durumu.....	1
Araştırmanın Amacı ve Önemi.....	3
Araştırma Problemi.....	4
Sınırlılıklar.....	5
Bölüm 2 Araştırmanın Kuramsal Temeli ve İlgili Araştırmalar.....	6
Araştırmanın Kuramsal Temeli.....	6
Güvenirlilik ile ilgili kuramsal çerçeve.....	6
Kayıp veri ile ilgili kuramsal çerçeve.....	12
İlgili çalışmalar.....	20
Bölüm 3 Yöntem.....	26
Simülasyon Verisi.....	26
Simülasyon Koşulları.....	26
Verilerin Analizi.....	27
Bölüm 4 Bulgular ve Yorumlar.....	33
Birinci Alt Probleme İlişkin Bulgular.....	34
İkinci Alt Probleme İlişkin Bulgular.....	38
Bölüm 5 Sonuç, Tartışma ve Öneriler.....	43
Sonuç.....	43

Tartışma.....	44
Öneriler	46
Araştırmanın sonuçlarına yönelik öneriler.....	46
Araştırmacılara yönelik öneriler.	47
Kaynaklar	48
EK-A: Etik Komisyonu Onay Bildirimi	53
EK-B: Etik Beyanı.....	54
EK-C: Yüksek Lisans/Doktora Tez Çalışması Orijinallik Raporu.....	55
EK-Ç: Thesis/ Dissertation Originality Report.....	56
EK-D: Yayımlama ve Fikrî Mülkiyet Hakları Beyanı.....	57

Tablolar Dizini

Tablo 1 Veri Setlerinin (Alfa) Güvenirlik Değerlerinin U-Testi Sonucu.....	28
Tablo 2 Veri Setlerinin (MTK Temelli Ampirik) Güvenirlik Değerlerinin U-Testi Sonucu	28
Tablo 3 DS Yöntemi Sonrası Veri Gruplarının Örneklem Sayıları	28
Tablo 4 Veri Setlerinin Betimsel İstatistik Sonuçları (Test Uzunluğu)	30
Tablo 5 Veri Setlerinin Betimsel İstatistik Sonuçları (Örneklem Büyüklüğü)	31
Tablo 6 Veri Setlerinin Betimsel İstatistik Sonuçları (Kayıp Oranı)	31
Tablo 7 Veri Setlerinin Kayıp Veri Oranlarına Göre Cronbach Alfa Güvenirlik Katsayıları	33
Tablo 8 Veri Setlerinin Kayıp Veri Oranlarına Göre MTK Temelli Ampirik Güvenirlik Katsayıları	34
Tablo 9 Tamamlanmış Veri Setlerinin Cronbach Alfa Güvenirlik Katsayıları	35
Tablo 10 Tamamlanmış Veri Setlerinin MTK temelli ampirik Güvenirlik Katsayıları	39

Şekiller Dizini

Şekil 1. Kayıp veri örüntüsü.....	15
-----------------------------------	----

Simgeler ve Kısaltmalar Dizini

- BM:** Beklenti Maksimizasyonu
ÇA: Çoklu Atama
DMF: Değişen Madde Fonksiyonu
DS: Dizin Silme
EM: Expectation Maximization
IRT: Item Response Theory
İEK: İhmal Edilemez Kayıp
KTK: Klasik Test Kuramı
LD: Listwise Deletion
MI: Multiple Imputation
MTK: Madde Tepki Kuramı
PISA: Uluslararası Öğrenci Değerlendirme Programı
RI: Regression Imputation
ROK: Rassal Olarak Kayıp
SIBTEST: Simultaneous Item Bias Test
SO: Seri Ortalamaları
TROK: Tamamen Rassal Olarak Kayıp

Bölüm 1

Giriş

Bu bölümde problem durumu, araştırmanın amacı ve önemi, araştırma problemi, sınırlılıklar ve tanımlar yer almaktadır.

Problem Durumu

Ölçme, genel anlamda nesnelerin veya bireylerin belli bir özelliğine ilişkin düzeylerini sayı ve sembollerle ifade etmektir (Turgut,1977). Eğitimdeki veya psikolojideki ölçmelerde ölçmek istediğimiz özellikler gizil özellikler olup bu özellikler, kendileri ile ilgili ölçülebilen özelliklerle belirlenir. Bireylerin gözlenebilen değişkenlere verdikleri tepkilerden yola çıkarak, gizil özelliklere yönelik çıkarımlar yapılır (Tavşancıl ve Akbaş, 2014).

Zeka, motivasyon, başarı, yetenek, kaygı gibi gizil özellikleri ölçecek ölçme aracı tasarlanırken, söz konusu ölçme aracından elde edilen puanların güvenilirliğini sağlamak ve puanlardan yapılacak çıkarımların geçerliğine ilişkin kanıt toplamak önemlidir. Çünkü test sonuçları öğrencileri başarılı olabilecekleri alana yöneltme, öğrenme güçlüklerini belirleme, öğretimde kullanılan metotların etkililik düzeyini saptama (Baykul, 2015) veya okula öğrenci kabul etme, klinik değerlendirme yapma, personel işe alımı gibi önemli kararların alınmasında rol oynarlar.

Crocker ve Algina (1986)'ya göre güvenilirlik, bir niteliğe ait ölçmelerin aynı kişilere uygulandığında benzer şartlarda benzer sonuçları vermesidir. Turgut (1977)'a göre ise güvenilirlik ölçme sonuçlarının tesadüfi hatalardan arınlık derecesidir. Tesadüfi hataların kaynakları tam olarak bilinemediğinden güvenilirlik duyarlılık, kararlılık, tutarlılık gibi anlamlarda da kullanılmaktadır (Baykul, 2015). Güvenirliğin kestirimine ilişkin birçok yöntem bulunmaktadır. Ancak bu yöntemler eksiksiz veri setleri üzerinde kullanılabilen ve genellikle de kayıp verilere yönelik bir çözüm mekanizması içermemektedir (Demir ve Parlak, 2012).

Kayıp veri, amaçlanan ölçmeye yönelik bir gözlemin bulunmaması durumunu tanımlamaktadır (Hohensinn & Kubinger, 2011). Bir başka deyişle kayıp veriler için bilgi eksikliği denilebilir ve bilgi eksikliği veri seti üzerinde yapılacak analizleri büyük ölçüde etkilemektedir (Bal, 2003). Kayıp veri, aynı zamanda veri setinin daralması ve dolayısıyla yapılacak kestirimlerin gücünün azalması anlamına gelmektedir

(Rubin, 1987). İstatistiksel analizler için hazırlanan programlar tam veri setlerine göre düzenlendiğinden, verideki eksikler analizlerde yanlı sonuçların ortaya çıkması, istatistiksel gücün abartılması veya azalması, örneklemin temsil edilebilirliğinin düşmesi, standart hatanın artması gibi sorunlara neden olmaktadır (Acock, 2005; Bal, 2003; Demir ve Parlak, 2012). Alanyazın incelendiğinde kayıp verilerin, sonuçların güvenilirliğini, geçerliliğini, genellenebilirliğini, ortalama ve standart sapma vb. gibi birçok istatistiksel değerlerini olumsuz yönde etkilediği görülmüştür (Little ve Rubin, 1987; Bal, 2003; Hohensinn & Kubinger, 2011).

Ölçme araçları tarafından elde edilen veriler, yanıtlayıcıdan, ölçme aracından veya uygulayıcıdan kaynaklanan sebeplerden dolayı her zaman eksiksiz şekilde elde edilememektedir (Işıkoğlu, 2017). Cheema (2012)'ya göre gözlemin dikkatsizliğinden, veri girişi yapılırken yapılan hatalardan ve yanlış işaretlemelerden dolayı kayıp veriler oluşabilir. Brown ve Kros (2003) kayıp verilerin oluşma sebeplerini üç grupta sınıflandırmıştır. Bunlardan ilki veri girişi sırasında karşılaşılan problemleri ele alan süreçle ilgili nedenlerdir. İkinci neden, cevaplamaı reddetme ile ilgili nedenlerdir. Üçüncü neden ise herhangi bir durum, grup veya konuyla ilgisi olmayan soruların sorulmasıdır. Maksimum düzeyde yansız ve güvenilir sonuçlar elde edebilmek için veri setlerinin eksiksiz ve yeterli büyüklükte olması gerektiği bilinmektedir. Maalesef, çoğu veri nadiren eksiksizdir (Cheema, 2012).

1970'lerin sonlarına kadar araştırmacılar kayıp verileri bir problem olarak görmemiş ve analizlere dahil etmemişlerdir. 1970'li yıllarda Afifi, Elashof, Rubin, Hartley, Hocking, Dempster, Orchard, Woodbury, Laird, ve Heckman gibi araştırmacıların yaptığı çalışmalar kayıp veriler ile ilgili yapılan ilk önemli çalışmalar olarak nitelendirilmektedir (Demir ve Parlak, 2012). Araştırmacılar silmeye ve atamaya dayalı birçok yöntem geliştirmişlerdir ve bu yöntemlerin her birinin analiz sonuçları üzerinde farklı etkileri olduğu bilinmektedir. Çıkarımları derinden etkileyen çeşitli kayıp veriyle baş etme yöntemi vardır. Bu nedenle araştırmacıların belli koşullarına uyacak uygun kayıp veri yöntemini seçmek önemlidir (Cheema, 2012).

Kayıp veriler yukarıda da belirtildiği gibi birçok farklı nedenden ötürü ortaya çıkabilmektedir. Kayıp veriye neyin sebep olduğunun ve kayıp veri mekanizmasının belirlenip ona uygun kayıp veriyle baş etme yönteminin seçilmesi sonuçların yanlı olmaması adına yapılması gereken bir işlemdir. Yine örneklem büyüklüğü ve kayıp veri miktarı da yöntem seçilirken göz önünde bulundurulması gereken özelliklerdir.

Kayıp veri yöntemleri seçilirken verinin yapısının da dikkate alınması gerekmektedir. İstatistiksel analizler için üretilen paket programlar daha çok sürekli veriler üzerinde çalışmaya olanak sağlamaktadır. İki kategorili puanlanan maddelerden oluşan testlerdeki veri setleri, ilgilenilen özelliğin yalnızca varlığını veya yokluğunu temsil ederler (Erkuş, 2012) ve bu veriler kesikli veri olarak adlandırılırlar. Kayıp veri sorunu istatistiksel programlara 1990'lı yıllardan sonra yansımaya başlamıştır ve çalışmalar büyük ölçüde çok kategorili veya sürekli veriler üzerinde yürütüldüğünden, analiz programları bu alanda gelişme göstermiştir (Demir, 2013). İki kategorili puanlanan diğer bir deyişle kesikli verilerde kayıp veri bulunması durumunda yaygın bir yaklaşım olarak verilerin yanlış kabul edilmesi veya 'uygulanmamış' olarak kodlanması yapılmakta ve bu nedenle veri setlerine yapılan analizlerde yanlış ve hatalı sonuçlara ulaşılmaktadır. Bunun sonucunda kesikli veri içeren veri setlerine daha karmaşık tamamlama yöntemlerinin kullanılmasının uygun olduğu belirtilmektedir (Demir, 2013; Hohensinn ve Kubinger, 2011).

Bu bilgilerden hareketle bu araştırmada farklı oranda kayıp veri içeren farklı örneklem büyüklüklerine sahip iki kategorili veri setlerine uygulanan kayıp veri yöntemlerinin güvenilirlik üzerine etkisi incelenerek karşılaştırılmıştır.

Araştırmanın Amacı ve Önemi

Bu araştırmada farklı test uzunluğunda, farklı büyüklükteki veri setleri ile bu veri setlerinden değişen oranlarda oluşturulan kayıp verilerin çeşitli kayıp veri ile baş etme yöntemleriyle tamamlanmasıyla elde edilen veri setlerinin güvenilirlik açısından karşılaştırılarak, hangi koşullarda hangi yöntemin gerçeğe daha yakın sonuçlar vereceğini belirlemek amaçlanmıştır. Bundan yola çıkarak $k=20$ ve $k=40$ test uzunluğunda $n=200$, $n=500$ ve $n=1000$ büyüklüklerinde iki kategorili veriler içeren yapay veri setleri türetilmiş ve her bir veri setinden %5, %10 ve %20 oranlarında kayıp veriler elde edilmiştir. Tam veri setleri ile kayıp veri içeren veri setleri çeşitli yöntemlerle tamamlanarak karşılaştırılmıştır.

Kayıp veri içeren araştırmalardaki analizlerin yanlış olduğu bilinmektedir. Bu da ölçme ve değerlendirmenin doğası gereği oldukça önemli ve giderilmesi gereken bir sorundur. Ancak uzun yıllar araştırmacılar bu sorunu göz ardı edip analiz dışı bırakmışlardır. Bu sorunu gidermek için araştırmacılar tarafından çeşitli yöntemler

geliştirilmiştir. Veri setlerinin büyüklüğü, kayıp veri miktarı, verilerin kesikli veya sürekli olması kullanılacak yöntemin seçilmesinde göz önünde bulundurulması gereken özelliklerdendir. Uygun kayıp veri yöntemini seçmek sonuçların güvenilir olması açısından büyük önem taşımaktadır.

Bir ölçmenin ölçülmek istenen özelliği hatalardan arınık olarak ölçmesi ve tekrarlanan ölçümlerde benzer sonuçları vermesi güvenilirlik kavramıyla ilgilidir. Yukarıda belirtildiği gibi güvenilirlik, testin standartlaştırılmasında yer alan önemli unsurlardan biridir. Veri setlerindeki kayıpların ölçmelerin güvenilirliğini tehlikeye attığı bilinmektedir. Bu bilgilerden hareketle bu araştırmanın sonuçlarının hem ilgili kişilere hem de alanyazına katkı sağlayacağı, güvenilirlik, kayıp veri veya kesikli veriler ile ilgili çalışma yapan araştırmacılara da yol gösterici olacağı düşünülmektedir.

Araştırma Problemi

Farklı test uzunluğundaki, farklı büyüklükteki tam veri setleri ile farklı oranlarda kayıp veri içeren veri setlerine uygulanan çeşitli kayıp veri yöntemleri ile elde edilen yeni veri setleri arasında güvenilirlik bakımından bir farklılık var mıdır?

Alt problemler. Araştırma problemi doğrultusunda aşağıdaki alt problemlere yanıt aranmıştır.

1. Farklı örneklem büyüklüğünde ve farklı test uzunluğundaki eksiksiz veri setleri ile bu veri setlerinin farklı oranlarda eksiltilmesiyle oluşan veri setlerine uygulanan dizin silme, sıfır atama, regresyon atama, beklenti maksimizasyonu ve çoklu atama yöntemleri ile tamamlanarak elde edilen yeni veri setleri arasında Cronbach alfa güvenilirlik kestirimleri açısından bir farklılık var mıdır?

a. Test uzunluğu dikkate alınarak karşılaştırıldığında, Cronbach alfa güvenilirlik kestirimleri arasında bir farklılık var mıdır?

b. Örneklem büyüklüğü dikkate alınarak karşılaştırıldığında, Cronbach alfa güvenilirlik kestirimleri arasında bir farklılık var mıdır?

c. Kayıp veri oranı dikkate alınarak karşılaştırıldığında, Cronbach alfa güvenilirlik kestirimleri arasında bir farklılık var mıdır?

2. Farklı örneklem büyüklüğünde ve farklı test uzunluğundaki eksiksiz veri setleri ile bu veri setlerinin farklı oranlarda eksiltilmesiyle oluşan veri setlerine uygulanan dizin silme, sıfır atama, regresyon atama, beklenti maksimizasyonu ve çoklu atama yöntemleri ile tamamlanarak elde edilen yeni veri setleri arasında MTK temelli ampirik güvenilirlik kestirimleri açısından bir farklılık var mıdır?

a. Test uzunluğu dikkate alınarak karşılaştırıldığında, arasında MTK temelli ampirik güvenilirlik kestirimleri arasında bir farklılık var mıdır?

b. Örneklem büyüklüğü dikkate alınarak karşılaştırıldığında, arasında MTK temelli ampirik güvenilirlik kestirimleri arasında bir farklılık var mıdır?

c. Kayıp veri oranı dikkate alınarak karşılaştırıldığında arasında, MTK temelli ampirik güvenilirlik kestirimleri arasında bir farklılık var mıdır?

Sınırlılıklar

1. Araştırmadaki veri setleri kayıp veri mekanizmalarından tamamen rassal olarak kayıp (TROC) mekanizmasıyla sınırlıdır.
2. Bu çalışma kayıp veri ile baş etme yöntemlerinden dizin silme (DS), sıfır atama, regresyon atama (RA), beklenti maksimizasyonu (BM) ve çoklu atama (ÇA) yöntemleriyle sınırlıdır.
3. Güvenirlik kestirimlerinde Cronbach alfa ve MTK temelli ampirik güvenilirlik katsayılarına yer verilmiştir.
4. Bu çalışmada her bir simülasyon koşulu için 25 replikasyon yapılmıştır.

Bölüm 2

Araştırmanın Kuramsal Temeli ve İlgili Araştırmalar

Araştırmanın Kuramsal Temeli

Bu bölümde araştırmanın kuramsal temeli incelenmiş, güvenilirlik ve kayıp veri olmak üzere iki başlık ile ilgili detaylı bilgiye yer verilmiştir.

Güvenirlik ile ilgili kuramsal çerçeve. Ölçme, belirli bir özellik bakımından bireyler, olaylar veya nesnelere hakkında değerlendirme yapmak veya karar vermek için yapılır. Bu değerlendirme ve kararların doğru ve uygun olması için ölçme sonuçlarının ve ölçüt koşulunun standardize olması beklenir. Güvenirlik ve geçerlilik kavramlarının birlikte kullanılmasıyla standardizasyon kavramı ortaya çıkmaktadır. Ölçme aracının standardize olabilmesi için maddeler analiz edilerek incelenir ve uygun geçerlik ve güvenilirlik katsayılarına ulaşılmaya çalışılır. Ölçüm standartlaştırılmış olur (Çakmur, 2012; Ercan ve Kan, 2004).

Güvenirlik, test veya ölçüğün istenilen özelliği doğru bir şekilde ortaya çıkarması ve farklı uygulamalarda benzer sonuçları vermesidir. Turgut (1977)'un güvenilirliğin 'Ölçme sonuçlarının tesadüfi hatalardan arınlık derecesi' tanımı da yaygın kullanılan tanımları arasındadır. Güvenirliğin birden çok anlamda kullanıldığı da görülmektedir. Örneğin, gram bölmeli terazi ile yapılan ölçmeler kilogram bölmeli terazilerle yapılan ölçmelerden daha duyarlı dolayısıyla da daha güvenilirdir. Burada güvenilirliğin duyarlılık anlamından bahsedilmiştir. Bir özellik aynı araçla birden çok defa ölçüldüğünde her seferinde aynı sonuçlar elde ediliyorsa, bu ölçme sonuçları kararlıdır ve dolayısıyla güvenilirdir. Burada ise güvenilirliğin kararlılık anlamından bahsedilmiştir. Testi oluşturan maddelerin testin bütünüyle tutarlı olması da güvenilirliğin tutarlılık anlamına değinmektedir (Baykul, 2015).

Klasik test kuramı (KTK)'na göre güvenilir bir ölçek, ölçme sonuçlarına karışan hatanın en az olduğu ölçek olarak tanımlanabilir. Ancak hatasız ölçme mümkün olmadığından güvenilirliği arttırmak adına hatanın en aza indirilmesi beklenir. Bundan yola çıkılarak herhangi bir ölçme sürecinde elde edilen ölçümlerde gerçek değer ve ölçme hatası bulunmaktadır. Gözlenen ölçüm değerini (x_i), iki ögesi olan gerçek değer (t_i) ve ölçüm hatası (e_i) unsurlarını da dikkate alarak aşağıdaki şekilde ifade edebiliriz:

$$x_i = t_i + e_i$$

Hataların tesadüfi olarak dağıldığı varsayılarak ve hataların tesadüfi olarak pozitif ve negatif yönde gelişip birbirlerini elemesi sonucunda ölçüm hataları ortalamasının sıfır olduğu sonucuna ulaşılır ($E(e_i)=0$). Bu varsayımdan hareketle ölçme puanlarının varyansı şöyle formüle edilebilir:

$$\sigma_x^2 = \sigma_t^2 + \sigma_e^2$$

Burada

σ_x^2 = Gözlenen değerlerin varyansı

σ_t^2 = Gerçek değerlerin varyansı

σ_e^2 = Ölçüm hatalarının varyansı

Bu bilgilerden yola çıkarak ölçümlerin ne derece güvenilir olduğunu belirten güvenilirlik katsayısına, gerçek değerlerin varyansını gözlenen değerlerin varyansına oranlayarak ulaşılabilir (Turgut, 1993; Gürsakal, 2001).

$$\rho_x = \frac{\sigma_t^2}{\sigma_x^2}$$

Burada

ρ_x = güvenilirlik katsayısı

Ancak gerçek değerlerin uygulamada bilinmesinin mümkün olmadığından bu hesaplama formül üzerinde kalmaktadır. Bunun için de güvenilirlik katsayısını belirlemek için birtakım yöntemler geliştirilmiştir. Bunlar;

1. Birden çok uygulama gerektiren yöntemlerdir.
2. Tek uygulama gerektiren yöntemler,

Birden çok uygulama gerektiren yöntemler.

Test-tekrar test yöntemi. Test-tekrar test yönteminde test aynı gruba, hatırlamalarını engelleyecek ölçüde uzun fakat ölçülecek özellikte bir değişme olmayacak kadar kısa bir süre sonra tekrar uygulanır ve iki uygulamadan elde edilen puanlar arasındaki korelasyon, güvenilirlik kestirimi olarak kullanılır. Bu katsayıya kararlılık katsayısı da denir.

Paralel formlar yöntemi. Paralel formlar yönteminde aynı özelliği ölçen iki test geliştirilir. Bu geliştirilen testlerin yapısının, zorluk derecesinin, madde sayısının,

puanlamasının aynı olmasına dikkat edilmelidir. Bu iki test gruba aynı anda veya ölçülmek istenilen özelliğin niteliğinde değişme olmayacak kadar kısa bir süre sonra uygulanır. Bu iki uygulamadan elde edilen puanlar arasındaki korelasyon, eşdeğerlik katsayısı veya eşdeğer formlar güvenilirliği olarak adlandırılır.

Bu yöntemde ölçmelerin eşdeğer olabilmesi için puanlar arasındaki korelasyonun yanında ortalama ve standart sapma istatistiklerine de bakılmalıdır. İki uygulamadan elde edilen puanların ortalama ve standart sapmaları arasında anlamlı bir farklılık olup olmadığı hipotez testleriyle incelenmelidir. Bu istatistikler arasında anlamlı farklılık yok ise puanlar arasındaki korelasyonun büyüklüğü ölçüsünde testin güvenilir olduğu söylenebilir (Baykul, 2015).

Tek uygulama gerektiren yöntemler.

Eşdeğer yarılar yöntemi. Eşdeğer yarılar yönteminde bir test, iki alt teste ayrılır. Alt testlerden her biri testin bütününe göre yarı uzunluktadır. Bu alt testlerden her birine 'yarı' denir. Testi iki yarıya ayırmada, maddeleri tesadüfi olarak çekmek, çift ve tek numaralı maddeleri ayırmak, maddelerin güçlük sırasına göre sıralanmasından sonra tek ve çift numaraları ayırmak, maddelerin kapsamlara göre eşlenip eşlerden her birini farklı yarılarına koymak gibi yöntemler uygulanabilir. Test iki yarıya bölündükten sonra bir gruba bir defa uygulanır ve testlerden elde edilen puanlar arasındaki korelasyon hesaplanarak güvenilirlik kestirimi olarak kullanılır.

İç tutarlık yöntemi. İç tutarlık yönteminde tek bir test ve tek bir ölçüm yapılarak maddelerin belirli bir kavramsal yapıyı tutarlı bir şekilde ölçüp ölçmediği tartışılır (Çakmur,2012). İç tutarlığı belirlemek için birtakım yöntemler geliştirilmiştir.

Cronbach alfa katsayısı. Cronbach (1951) tarafından geliştirilen α katsayısı, K bileşenden oluşan bir bileşik ölçmeye ait gözlenen puanların gerçek puanlarla korelasyonunu verir. Bu korelasyon, bileşenlerin varyanslarının bileşik ölçmeye ait toplam puanların varyanslarına bölünerek elde edildiğinden, bileşenlerin ölçmenin bütünüyle ne kadar tutarlı olduğunu göstermektedir (Baykul, 2015). Cronbach alfa çoklu ve ikili kategorili puanlanan testlerde testin bütününe ait güvenilirlik katsayısının hesaplanmasında kullanılır.

Cronbach alfa şu formülle hesaplanmaktadır:

$$\alpha = \left(\frac{K}{K-1} \right) \cdot \left[1 - \frac{\sum_{j=1}^K \sigma^2(X_j)}{\sigma^2(X)} \right]$$

K= madde sayısı

$\sigma^2(X_j)$ = j maddesinin varyansı

$\sigma^2(X)$ = toplam puan varyansı

Kuder-Richardson güvenilirlik katsayısı. İç tutarlığı belirlemek için kullanılan yöntemlerden biri de KR-20 ve KR-21 güvenilirlik katsayılarıdır. KR-20 eşitliği maddeleri 0 ve 1 şeklinde puanlanan testlerde, testin bütününe ait güvenilirlik katsayısının hesaplanmasında kullanılır.

Kuder ve Richardson tarafından oldukça karmaşık varsayımlar altında elde edilmiş bu formüller α katsayısının özel bir hali olarak aşağıdaki gibi yazılabilir.

$$\text{KR-20} = \left(\frac{K}{K-1} \right) \cdot \left[1 - \frac{\sum_{j=1}^K \pi_j \cdot (1-\pi_j)}{\sigma^2(X)} \right]$$

Burada

K= madde sayısı

$\sigma^2(X)$ = toplam puan varyansı

π_j = madde güçlük indeksi

Formülde görüldüğü üzere KR-20 güvenilirlik katsayısı madde sayısına bağlıdır, madde sayısı arttıkça güvenilirlik de artacaktır.

$$\text{KR-21} = \left(\frac{K}{K-1} \right) \cdot \left[1 - \frac{K \cdot \mu(X) - [\mu(X)]^2}{K \cdot \sigma^2(X)} \right]$$

Burada

K= madde sayısı

$\sigma^2(X)$ = toplam puan varyansı

$\mu(X)$ = toplam puanların ortalaması

Formülden görüldüğü üzere KR-21 güvenilirlik katsayısı madde sayısı, test ortalaması ve test varyansına bağlıdır. KR-21 hesaplanırken tüm maddelerin güçlükleri eşit kabul edilmiştir, bu da pratikte pek mümkün olan bir şey değildir. Bu

nedenle KR-21 güçlükler eşit olmadığında KR-20'den düşük çıkacaktır (Baykul, 2015).

KTK'da yukarıda da bahsedildiği üzere $x_i = t_i + e_i$ sayılısını temel alan sınıf içi korelasyon olarak da adlandırılan güvenilirlik hesaplamaları yapılmaktadır. Bunun ikili veriye uygulanmasıyla gizil korelasyon denilen korelasyon ortaya çıkar ve gizil korelasyon kullanılarak hesaplanan güvenilirliğin gizil ölçekte olacağı sonucuna varılmıştır (Molenberghs ve Verbeke, 2005). Ancak araştırmacılar gözlenen puanlar arasındaki korelasyonu baz alan ve daha anlamlı güvenilirlik ölçümleri yapan belirgin korelasyonu kullanmaya eğilimlidirler. Normal dağılıma sahip veri setlerinde bu iki korelasyon birbirleriyle çakışmaktadır ancak normal dağılıma sahip olmayan veya ikili veriye sahip veri setlerinde bu mümkün olmamaktadır. Bunun için Milanzi, Molenberghs, Alonso, Verbeke ve Boeck (2015)'e göre anlamlı güvenilirlik ölçümleri içeren yeni formül türetilmelidir. Ancak sürekli veriler için uygun olan belirgin korelasyonu içeren bu formülü üretmek iki kategorili veriler için uygun değildir. Bunun sebebi ise ikili veriler için kullanılacak bu formülün, kapalı formları olmadığı için yaygın olarak kullanılmayan integral hesapları içermesidir. Bu sebeplerden dolayı iki kategorili verilerde KTK temelli yaklaşımların yanında MTK temelli ampirik yaklaşımların kullanıldığı da görülmektedir.

KTK'da birey özelliklerinin test özelliklerinden ayrılamaması, güvenilirlik katsayısının hesaplanmasında kullanılan ölçmenin standart hatası kavramının testi alan bireyler için aynı kabul edilmesinin yalnızca teoride olabilmesi, paralel formlardan elde edilen korelasyon olarak tanımlanan güvenilirlik katsayısı hesabında paralel formlar tanımını sağlamanın uygulamada zor olması vb. gibi nedenlerden dolayı psikometristler tarafından alternatif ölçme kuramları arayışı sonucunda, 1950'li yıllarda bir maddenin doğru yanıtlandırılma olasılığı ile bireyin ölçülmek istenen yetenek düzeyi arasında ilişki kuran bir model olan Madde Tepki Kuramı (MTK) ortaya çıkmıştır (Crocker ve Algina, 1986; Akyıldız ve Şahin, 2017). MTK'da KTK'nın aksine madde parametreleri örneklemden bağımsızdır ve her bir birey için ayrı standart hata kestirimleri yapılabilir. MTK, KTK'dan farklı olarak yetenek kestirimlerini dikkate almaktadır. Bireylerin testteki bazı maddelere yanıt vermeleri yetenek kestirimleri için yeterli olduğundan, bu sayede kayıp verilere yönelik alternatiflerin üretilmesini sağlamıştır (Lord, 1980). Bir bireyin yeteneği teoride

negatif sonsuz ve pozitif sonsuz arasındadır ancak çalışmalarda genellikle +3 ve -3 ranjı kullanılmaktadır.

MTK'da veri yapısının kestirilecek parametre sayısı üzerinde etkisi vardır. İki kategorili verilerde 1, 2 ve 3 parametrelili lojistik model kullanılmaktadır. 1-parametrelili modelde madde güçlüğü (b), 2-parametrelili modelde madde ayırt ediciliği (a) ve madde güçlüğü (b), 3-parametrelili modelde madde ayırt ediciliği (a), madde güçlüğü (b) ve şans faktörü (c) ile çalışılmaktadır (Crocker ve Algina, 1986).

2-parametrelili modelde güvenilirlik hesabı şu formüllerle yapılmaktadır.

$$Y_{ji} = \mu_{ji} + \varepsilon_{ji}$$

Burada

Y_{ji} = j kişinin i maddesinde gözlenen puanı

μ_{ji} = j kişinin i maddesinde gerçek puanı

ε_{ji} = hata puanı

2-parametrelili modelde gözlenen puanın formülü şu şekilde olmaktadır:

$$Y_{ji} = \mu_{ji} + \varepsilon_{ji} = \frac{\exp[\alpha_i(\theta_i - \beta_i)]}{1 + \exp[\alpha_i(\theta_i - \beta_i)]} + \varepsilon_{ji}$$

Burada

θ_i = kişinin gizil özellik puanı

α_i = maddenin ayırt ediciliği

β_i = maddenin güçlüğü

Aşağıdaki formüllerden elde edilen değerler gözlenen madde puanı ve gözlenen toplam puanlardır.

$$Y_i = \int \left\{ \frac{\exp[\alpha_i(\theta_i - \beta_i)]}{1 + \exp[\alpha_i(\theta_i - \beta_i)]} + \varepsilon_{ji} \right\} \phi(\theta_i | 0, \sigma_\theta^2) d\theta = \mu_{ji} + \varepsilon_{ji}$$

$$S_T = \int \sum_{i=1}^I \left\{ \frac{\exp[\alpha_i(\theta_i - \beta_i)]}{1 + \exp[\alpha_i(\theta_i - \beta_i)]} + \varepsilon_{ji} \right\} \phi(\theta_i | 0, \sigma_\theta^2) d\theta = \mu_{ji} + \varepsilon_{ji}$$

Y_i = gözlenen madde puanı

S_T = gözlenen toplam puan

Dimitrov (2003) varyansları şu şekilde tanımlamıştır (akt. Milanzi ve diğerleri, 2015):

$$\text{Var}(\varepsilon_i) = \int \mu_{ji}(1 - \mu_{ji}) \phi(\theta_i | 0, \sigma_\theta^2) d\theta$$

$$\text{Var}(\mu_i) = \mu_{ji}(1 - \mu_{ji}) - \text{Var}(\varepsilon_{ji})$$

$$\text{Var}(\varepsilon) = \sum_{i=1}^I \text{Var}(\varepsilon_{ji})$$

$$\begin{aligned} \text{Var}(\mu) = \int \left\{ \sum_{i=1}^I \frac{\exp[\alpha_i(\theta_i - \beta_i)]}{1 + \exp[\alpha_i(\theta_i - \beta_i)]} \right\}^2 \phi(\theta_i | 0, \sigma_\theta^2) d\theta \\ + \left\{ \int \sum_{i=1}^I \frac{\exp[\alpha_i(\theta_i - \beta_i)]}{1 + \exp[\alpha_i(\theta_i - \beta_i)]} (\theta_i | 0, \sigma_\theta^2) d\theta \right\}^2 \end{aligned}$$

$\text{Var}(\mu)$ = gerçek puanların varyansı

Buradan yola çıkarak, gerçek varyansın gözlenen varyansa oranı olarak tanımlanan güvenilirliği şu şekilde yazabiliriz:

$$\rho_i = \frac{\text{Var}(\mu_i)}{\text{Var}(\mu_i) + \text{Var}(\varepsilon_i)}$$

$$\rho_S = \frac{\text{Var}(\mu)}{\text{Var}(\mu) + \text{Var}(\varepsilon)}$$

ρ_S = güvenilirlik katsayısı

Gözlenen madde puanı ve gözlenen toplam puan, güvenilirliği ölçebilir. Ancak bu ölçümler için gereken bileşenlerin her birini elde etmek için iki kategorili veri dağılımındaki, kapalı formlarının bulunmadığı bilinen normal rastgele etkilerin integralini almayı gerektirmektedir (Milanzi ve diğerleri, 2015).

Kayıp veri ile ilgili kuramsal çerçeve. Eğitim veya sosyal bilimler alanlarında yapılan birçok araştırmalarda, belirli birtakım sebeplerden ötürü veri setlerinde eksikler bulunabilmektedir. Standart istatistiksel yöntemler, dikdörtgensel veri setlerinin analizine yönelik olarak geliştirilmiştir. Matris şeklindeki bu veri setlerinde satırlar gözlemleri, sütunlar ise değişkenleri göstermekte ve bu matrislerin 'hücre' olarak isimlendirilen her bir birimine gerçek sayı değerleri girilmektedir. Bir değişkene ait gözlemin bulunmaması durumunda, söz konusu gözlemi temsil eden hücre boş kalır ve bu da kayıp veri anlamına gelmektedir (Little & Rubin, 1987). İstatistiksel programlar eksiksiz veri setlerine yönelik tasarlandığından matristeki hücrelerde eksiklerin olması bir diğer deyişle veri setlerinde kayıp veri bulunması analizleri önemli ölçüde etkilemektedir.

İstatistiksel analizlerde kayıp veri her zaman sorun teşkil etmiştir. Bunun sebepleri, çoğu yirminci yüzyılın başlarında geliştirilen ve araştırmacılar tarafından kullanılan analiz programlarının eksiksiz veri setleri üzerinden geliştirilmesi (Graham, 2009) ve araştırmacıların kayıp verileri analizlere dahil etmek yerine analiz dışı bırakıp ilgisiz kalması (Demir ve Parlak, 2012) şeklinde açıklanabilir. İstatistiksel analizlerde kayıp veri sorunu 1930'lu yıllara dayanmakla beraber bu sorunun istatistiksel programlara yansımaları 1970'li yılları bulmuş ve asıl ilerleme 1980'li yıllarda kaydedilmiştir. İstatistiksel programlarda kayıp veri ile ilgili analizler 1990'lı yıllarda oldukça yaygınlaşmıştır (Kürşad ve Nartgün, 2015; Bayhan, 2018). Bu geliştirilen istatistiksel yöntemler genellikle sürekli veriler üzerinde kullanılabilir. İki kategorili puanlanan maddelerin analizleri çok daha karmaşık olmaktadır. İki kategorili puanlanan maddelerden oluşan veri setlerindeki kayıp veriler yaygın olarak 'yanlış' veya 'uygulanmamış' olarak kabul edilirler. Ancak bunun yanı sıra kestirime yol açtığı bilinmektedir.

'Yanıtlanmama (nonresponse)', 'boş bırakma (omitted)' ve 'kayıp veri (missing value)' tanımlamaları, literatürde oldukça yakın anlamda ve sıklıkla birbirine yerine kullanılmaktadır. Esas itibarıyla boş bırakma yanıtlanmama durumuna, yanıtlanmama durumu ise kayıp veriye yol açmaktadır.

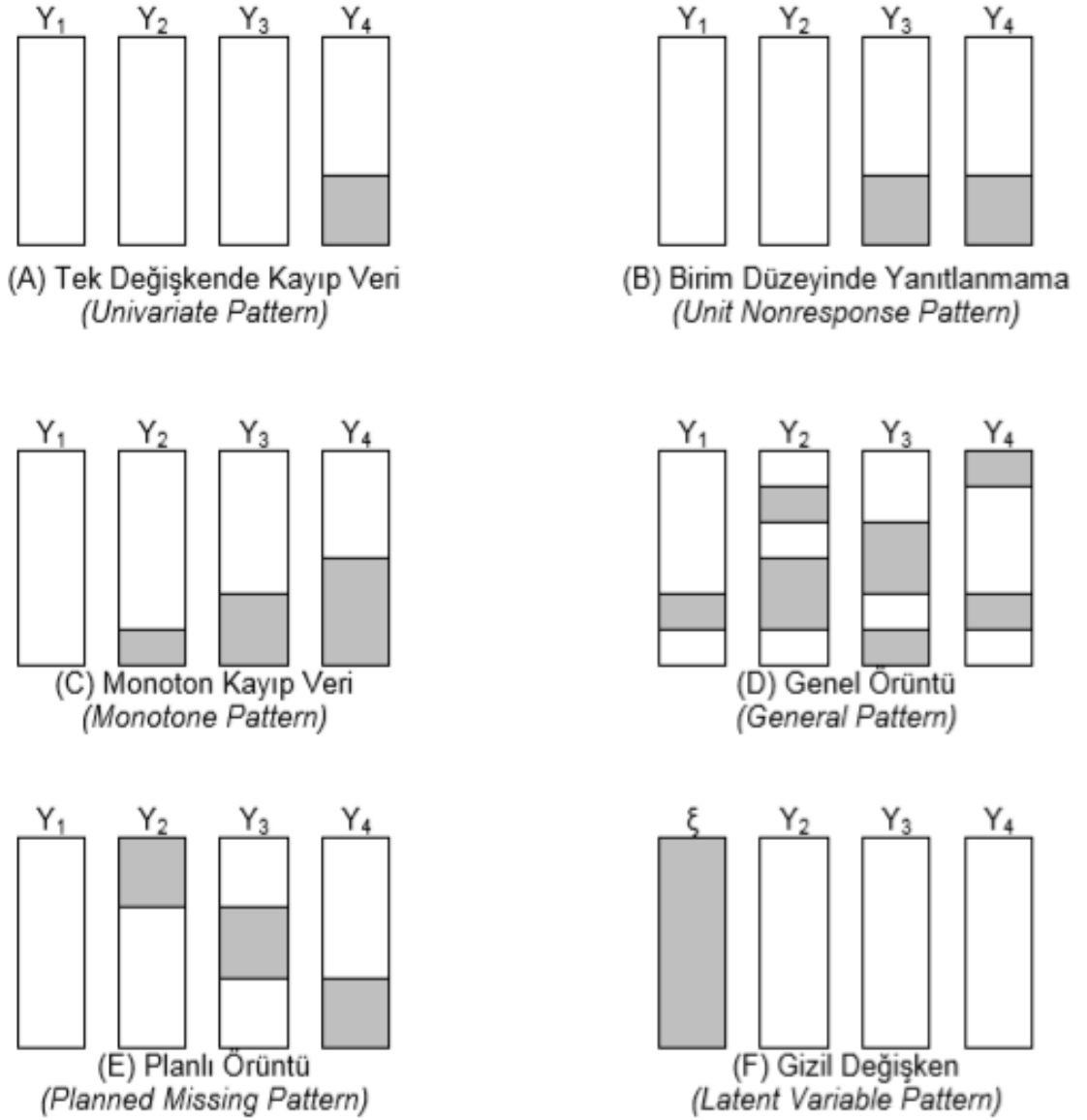
Kayıp verilerin ortaya çıkmasında birçok etken rol oynamaktadır. McKnight, McKnight, Sidani ve Figueredo (2007) bu sebepleri çalışmaya katılanlarla ilgili, çalışmanın tasarımıyla ilgili ve katılımcılarla çalışmanın tasarımı arasındaki etkileşimle ilgili olmak üzere üç gruba ayırmışlardır. Örneğin; bazı katılımcıların bir ankette kendilerine rahatsız edici gelen bazı maddeleri cevaplamaması (katılımcılarla ilgili), çalışmanın katılımcının vaktini çok fazla alıyor olması (tasarımla ilgili) ya da çalışmanın külfetli/ağır kısımları nedeniyle yanıtlayıcının bıkkınlık ve isteksizliğe kapılıp cevaplamayı tamamlamaması (katılımcı ve tasarım özellikleriyle ilgili) gibi durumlar ortaya çıkabilmektedir. Rubin (1987), yanıtlanmama durumunu ikiye ayırmıştır. Bir araştırmada gözlem birimi tarafından kendisine yöneltilen tüm maddelerin yanıtlanmaması durumu 'birim düzeyinde yanıtlanmama (unit nonresponse)', bazı maddelerin yanıtlanmaması durumu 'madde düzeyinde yanıtlanmama (item nonresponse)' olarak tanımlanmaktadır. Yanıtlanmama durumlarında bağlı olarak ise araştırmalarda 'yanıtlanmama yanlılığı' ve 'boş bırakma/atlama yanlılığı' oluşabilmektedir.

Arařtırmalarda kayıp verilerin bulunması bilgi eksikliđine yol amaktadır ve bu da yapılan analizlerin etkililiđinin azalmasına sebep olmaktadır. Peng, Harwell, Liou ve Ehman (2007) kayıp verilere bađlı olarak ortaya ıkabilecek drt problemi řu řekilde zetlemiřtir:

1. İstatistiksel bir modelden hareketle yapılacak kestirimlerde kayıp verilerin yol aacađı yanlılık, en nemli sorun olarak ifade edilmektedir.
2. Kayıp veriler bilgi eksikliđine ve dolayısıyla analizlerin gcnn azalmasına sebep olabilmektedir.
3. Yaygın olarak kullanılan istatistiksel yntemlerin, veri setlerinde kayıpların olduđu durumlarda kullanımı zor olabilmektedir.
4. Kayıp veriler, deđerlendirilebilir kaynakların bořa harcanmasına yol aabilmektedir.

Arařtırmacılar genellikle eksiksiz veri setleri zerinde alıřma veya kayıp verileri analiz dıřı bırakmanın manidar bir farka yol amadıđını kanıtlama eđilimindedirler. Rubin (1987)'e gre kayıp verilerin analiz dıřı bırakılmasının sebebi birtakım varsayımlara bađlı olarak kayıp verilerin 'ihmal edilebilir' olarak grlmesidir. Kayıp verilerin ihmal edilebilirliđi kayıp verilerin karakteriřtiđi ve kayıp verilerin oluřmasına yol aan srelerle ilgilidir ve bu kayıp verilerin analiz dıřı bırakılması iin ihmal edilebilir olduđunun kanıtlanması gerekir. Uygun analiz yntemlerinin belirlenmesi ve sonuların yorumlanması, kayıp veri rnts ve mekanizmasıyla yakından ilgilidir (Little ve Rubin, 1987).

Kayıp veri rnts. Kayıp veri rnts, gzlenen ve kayıp verilerin bir veri setindeki konumlanma biimlerini tanımlanmaktadır. Kayıp veri rnts, verideki kaybın konumunu ifade eder ve bu kaybın neden oluřtuđunu aıklamaz (Enders, 2010). řekil 1'de kayıp verilerin veri setindeki konumlanma biimine gre sınıflandırılmasını betimleyen grafıklere yer verilmiřtir.



Kaynak: Enders, C.K. (2010), *Applied Missing Data Analysis*. New York: The Guilford Press, sy.4

Şekil 1. Kayıp veri örüntüsü

Tek değişken örüntüsünde veri setindeki değişkenlerden yalnızca birinde kayıp veri gözlenir ve bu durum daha çok deneysel desenlerde karşılaşılan bir durumdur. Birim yanıtlanmama örüntüsü, genellikle anket araştırmalarında ortaya çıkıp, kişilerin bazı değişkenlere cevap vermemesi durumudur. Monoton örüntü, boylamsal çalışmalarda zamanın geçmesiyle katılımcıların çalışmalara geri dönmemeleri sonucu ortaya çıkan kayıp verilerde görülmektedir. Kayıp verilere yönelik en yaygın örüntü olan genel örüntüde, kayıp veriler veri grubundaki değişkenlerde düzensiz şekilde dağılmaktadır. Planlanmış örüntü, katılımcıların

yükünü azaltmak için anketlerin bölünüp farklı formlara dağıtılmasıyla oluşmaktadır. Gizil değişken örüntüsünde, yapısal eşitlik modeli gibi gizil değişken içeren analizlerde gözlenen bir veri olmamasından kaynaklanarak bu gizil değişken başlı başına bir kayıp veri olarak nitelendirilir (Enders, 2010).

Kayıp veri mekanizmaları. Kayıp veri mekanizmaları, ölçülen değişkenlerle kayıp veri olasılığı arasındaki olası ilişkileri tanımlamaktadır (Enders, 2010). Bu nedenle kayıp veri mekanizmaları, kayıp verilerin karakteristiğini belirlemede ve uygun kayıp veri yöntemlerini kullanma ve dolayısıyla analizlerde önemli rol oynamaktadır. Kayıp veri mekanizmalarına yönelik ilk sınıflandırmalar Rubin (1976) tarafından yapılmıştır. Bu sınıflandırmaya göre kayıp veri mekanizmaları tamamen rassal olarak kayıp (TROK), rassal olarak kayıp (ROK) ve ihmal edilemez kayıp (İEK) olarak 3'e ayrılmıştır.

Tamamen rassal olarak kayıp. Herhangi bir verinin kayıp olma olasılığının, kaybın bulunduğu değişkenin kendisinden ve veri grubundaki diğer değişkenlerden bağımsız olmasıdır (Little ve Rubin, 1987). Tüm değişkenleri için bu varsayımı sağlayan bir veri seti sadece gözlenen verilerin oluşturduğu bir örneklem, üzerinde çalışılması planlanan tam veri setinin rasgele seçilmiş bir örnekleme olarak düşünülebilir (Allison, 2001; Enders, 2010).

X gözlenen, Y kayıp veri içeren değişken olmak üzere tamamen rassal olarak kayıp (TROK) mekanizmasının formül olarak gösterimi aşağıdaki gibidir:

$$P(Y \text{ kayıp} | X, Y) = P(Y \text{ kayıp})$$

Örneğin, X değişkeni tutum Y değişkeni başarı durumunu gösteriyor olsun. Y değişkeninde kayıp veri olması ne X değişkenine ne de Y değişkenine bağlı değilse Y değişkenindeki kayıp veriler tamamen rassal olarak kayıptır. Yani başarı değişkenindeki kayıp veriler ne tutuma ne de başarıya bağlı değildir (Allison, 2003).

Rassal olarak kayıp. Herhangi bir verinin kayıp olmasının, kaybın bulunduğu değişkenin kendisinden bağımsız, ancak araştırmada ele alınan diğer bir değişkene veya değişkenlere bağımlı olmasıdır (Little ve Rubin, 1987).

X gözlenen, Y kayıp veri içeren değişken olmak üzere RK mekanizmasının formül olarak gösterimi aşağıdaki gibidir;

$$P(Y \text{ kayıp} | X, Y) = P(Y \text{ kayıp} | X)$$

Örneğin, X ve Y gibi iki tane test olsun. X sınavından düşük alan bireyler Y sınavına girmeyi reddetmektedirler. Bu durumda Y sınavını tamamlama durumu X sınavından alınan sonuçlara bağlıdır ve Y sınavının kendisi ile ilişkisi yoktur. Bu durumda Y sınav sonuçlarındaki kayıp veriler rassal olarak kayıptır (Enders, 2001 (b)).

İhmal edilemez kayıp. İhmal edilemez kayıp mekanizmasında verilerin kayıp olma olasılığı gözlenemeyen yani kayıp verinin yer aldığı değişkenle ilgilidir (Little ve Rubin ,1987). Eğer, diğer değişkenlerin etkileri kontrol altına alındığında, bir değişkendeki kayıp veri içeren durumlar, aynı değişkendeki tam veri içeren durumlardan önemli derecede farklıysa ihmal edilemez kayıp vardır (Cheema, 2012). Bu durumda kayıp veri mekanizmasının, iyi bir parametre tahmini yapabilmek için modellenmesi gerekir (Allison, 2001).

X daima gözlenen, Y ise kayıp veri içeren değişken olmak üzere bu yapının formüsel olarak gösterimi şu şekildedir;

$$P(Y \text{ kayıp} / Y, X) = P(Y \text{ kayıp} / Y)$$

Örneğin, öğrencilere verilen bir okuma testinde, okuma güçlüğü çeken öğrencilerin bazı test sorularını boş bıraktığı görülebilir. Bu durumda okuma testindeki kayıp veriler öğrencilerin okuma güçlüğü ile ilişkilendirilebilir (Baraldi ve Enders, 2010).

Kayıp verilerin rassallığının incelenmesinde 3 yöntem kullanılır. Bunlardan ilki, kayıp veri içeren ve içermeyen veri setleri arasındaki ortalama farkının incelendiği t-testidir. Anlamlı bir fark çıkmaması kayıp verilerin tamamen rassal olarak dağıldığını göstermektedir. İkincisi, kayıp veri içeren ve içermeyen veri grupları arasındaki Pearson korelasyon katsayısını karşılaştırmaktadır. Düşük korelasyonlar rassallığı göstermektedir. Üçüncüsü ve en yaygın kullanılanı Little'ın TROK testidir. Bu test, aynı kayıp veri örüntüsüne sahip gözlemlerin alt gruplarının ortalama farklılıklarını karşılaştırmaktadır. Alt grupların ortalamaları ve genel ortalama arasındaki farkların ağırlıklı toplamı olarak da ifade edilir. Bu testte anlamlı fark çıkmaması kayıp verilerin tamamen rassal olarak dağıldığını göstermektedir (Baygül, 2007; Enders, 2010).

Kayıp veri ile baş etme yöntemleri. Kayıp veriler hemen her çalışmada karşılaşılan bir sorundur ve ilgili kişilere ulaşılarak eksik verilerin her zaman tamamlanmamasının yarattığı sorun, bazı istatistikler yöntemlerin geliştirilmesine neden olmuştur. Bu yöntemlerin kullanılması analizlerin sonucunu doğrudan etkileyeceğinden, uygun yöntemi seçmek önem kazanmaktadır.

Allison (2009)'a göre bu yöntemleri seçerken bazı kriterleri göz önünde bulundurmak gerekmektedir. Bu kriterler şu şekildedir:

1. Seçilecek yöntem, parametre tahminlerindeki yanlılığı mümkün olduğunca düşük tutmalıdır.
2. Seçilecek yöntem, kayıp verileri analiz dışı bırakmak yerine elde olan bilgiyi en iyi şekilde kullanmalıdır.
3. Standart hata, güven aralığı ve p değerleri için yapılan tahminler hassas ve doğru olmalıdır.

Literatür incelendiğinde silmeye ve atamaya yönelik olarak pek çok kayıp veri ile baş etme yönteminin olduğu görülmektedir. Bu bölümde araştırmada kullanılan dizin silme (DS), sıfır atama, regresyon atama (RA), beklenti maksimizasyonu (BM) ve çoklu atama (ÇA) yöntemleri ile ilgili bilgilere yer verilmiştir.

Dizin silme. DS, herhangi bir değişkene ait gözlemin bulunmadığı durumlarda kayıp veri içeren bireye ait tüm gözlemlerin veri setinden çıkarılıp analizlere dahil edilmediği yöntemdir. Tam hücre analizi (complete case analysis) de denilen bu yöntemde kayıp veri bulunan gözlemcilerin verileri analiz dışı bırakılarak eksiksiz bir veri seti oluşturulur ve bu veri grubu üzerinde her türlü istatistiksel işlemler uygulanabilir (Işıkoğlu, 2017). Bu yöntem çok pratik gibi gözükmemektedir ancak gözlemcilerin analiz dışı bırakılması veri kaybı anlamına gelmektedir ve veri kaybı standart hatanın artmasına, güvenirliliğin azalmasına ve hipotez testinin gücünün düşmesine neden olmaktadır (Demir, 2013). Ancak kayıp verilerin TROK mekanizmasında olması durumunda, verilerin çıkarılarak oluşturulduğu eksiksiz veri seti, orijinal veri setinin seçkisiz bir örnekleme olacağından standart hata kestirimleri, gerçek verinin standart hata kestirimlerine yakın olacaktır (Allison, 2003). Kayıp veri mekanizmasının ROK veya IEK olması durumunda oluşturulan eksiksiz veri seti, orijinal veri setinin seçkisiz bir örnekleme olmayacaktır ve bu nedenle de analizler sonucu elde edilecek parametrelerde yanlılık gözlemlenecektir (Allison, 2003;

McKnight ve diğeri, 2007). DS yönteminin örneklem büyüklüğünün fazla ve kayıp verinin az olduğu durumlarda kullanılması önerilmektedir çünkü bu durumlarda sonuçlar arasında önemli ölçüde fark olmamaktadır ancak örneklem büyüklüğünün az olması durumunda kayıp veriler tamamen rassal olarak dağılsalar bile, oluşacak veri kaybı sonuçları önemli ölçüde etkileyecektir.

Sıfır atama. Bu yöntemde kayıp verilerin yerine 0 değeri atanmaktadır. Bu yöntemin kullanılabilmesi için 0'ın gözlemlenebilir değerler arasında olması gerekmektedir (McKnight ve diğeri, 2007). Bu yöntem genellikle başarı testlerinde ve iki kategorili cevapların olduğu durumlarda yaygın kullanılmaktadır (Tamcı, 2018).

Regresyon atama. Regresyon atama yönteminde kayıp veriler, kayıp veri içermeyen diğer değişkenler kullanılarak oluşturulan regresyon denklemine göre tahmin edilmektedir (Enders, 2010). Kayıp veriler bağımlı değişken, diğer değişkenler de bağımsız değişkenler kabul edilerek bir regresyon denklemi kurulur ve kayıp veriler için tahminlerde bulunulur. Bu yöntem kullanılırken bağımlı ve bağımsız değişkenler arasındaki ilişkinin yeterli düzeyde olmasına dikkat edilmelidir. Tabachnick & Fidell, (2013)'e göre, kayıp veriler diğer değişkenlerden tahmin edildiği için gerçekte olması gerekenden daha iyi uyum gösterebilir veya elde edilen değer gözlenen verilerin değer aralığında olamayabilir. Bunlar RA yönteminin dezavantajlarıdır.

Beklenti maksimizasyonu. Dempster, Laird & Rubin (1977) tarafından önerilen, Beklenti (B) ve maksimizasyon (M) olmak üzere iki aşamadan oluşan bu yöntem bir en çok olabilirlik yöntemidir (McKnight ve diğeri, 2007). B basamağında gözlenen değerler kullanılarak bir dizi regresyon denklemi elde edilir ve kayıp verilere atama yapılır. M basamağında ise, atama yapılan değerler de kullanılarak yeni değer atamaları yapılır (Enders, 2010). Art arda yapılan her tekrarlı atama sonucunda daha fazla bilgi edinilir çünkü her tekrar, önceki tekrarlarda yer alan bilgileri kullanmaktadır (Acock, 2005). Bu tekrarlama işlemi, önce ve sonra yapılan atamaların parametrelerinde anlamlı farklılık oluşmayana kadar devam eder. Beklenti maksimizasyonunda temel çıktılar ortalama, varyans ve kovaryansların maksimum olabilirlik kestirimleridir (Howell, 2007). Bu belirtilen çıktı tahminlerinin neredeyse yansız yapılması bu yöntemin avantajı olarak gösterilebilir (Howell, 2007), ancak atama sürecinde regresyon denklemi kullanılmasından ve

kovaryans matrisindeki bileşenlerin hata bileşeni olmadan atanmasından dolayı değişkenliğin korunamaması (Enders,2001(b)) da yöntemin dezavantajıdır.

Çoklu atama. Rubin (1987) tarafından geliştirilen bu yöntemde kayıp veriler yerine birden fazla değer atanmakta ve böylelikle birden fazla tam veri seti oluşmaktadır. Bu yöntem BM yönteminin tekrarlı bir uzantısıdır (Olinsky, Chen ve Harlow, 2003). BM yönteminde kayıp verilere tek bir değer atanırken ÇA yönteminde m sayıda veri atanmaktadır.

Bu yöntem atama (imputation), analiz etme (analysis) ve bir araya getirme (pooling) olmak üzere üç adımdan oluşmaktadır. Atama aşamasında kayıp veriler yerine m tane değer atanarak m tane veri seti elde edilir. Buradaki m sayısının 2 ile 10 arasında olması idealdir (Cheema, 2012). Analiz aşamasında veri seti tam halinde olması halinde yapılacak analizler m tane veri setinin her birine ayrı uygulanır. Bir araya getirme aşamasında elde edilen parametrelerin genellikle ortalamaları alınarak, parametre kestirimleri tek bir değerde toplanır (Enders, 2010).

Bu yöntem her türlü veri ve modele uygulanabilmektedir. Diğer yöntemler veri mekanizmasına duyarlı iken bu yöntemin rasgele ve rasgele olmayan kayıp veri mekanizmalarında dahi başarılı olduğu ve oldukça isabetli kestirimler yaptığı görülmüştür (Tamcı, 2018). Bu yöntem analizler esnasında kayıp veriyle ilgili belirsizliği göz önüne aldığı için standart hatalardaki yanlılığı azaltmaktadır (Van Ginkel ve diğerleri, 2010). Bunlar ÇA yönteminin avantajları olarak gösterilebilir. Ancak kayıp veriler tekrarlı regresyon işlemlerine dayanarak yordandığı için korelasyon katsayısı olduğundan büyük çıkabilir veya bu yöntemde rasgele varyans eklendiğinden aynı veri setinden farklı sonuçlar çıkabilir (Allison, 2003). Bunlar ise ÇA yönteminin dezavantajlarıdır.

İlgili çalışmalar

Bu bölümde, alanyazında kayıp veri ile baş etme yöntemlerinin ve güvenirliliğinin incelendiği yurt içi ve yurt dışındaki araştırmalar incelenmiş ve tarihlerine göre sıralanarak verilmiştir.

Enders (2003), kayıp veri ile baş etme yöntemlerini Cronbach alfa katsayısı temelinde incelemiştir. Bunun için farklı kayıp veri mekanizmaları altında farklı örneklem büyüklüklerine sahip çok kategorili olarak puanlanan veriler içeren veri

setleri, liste bazında silme, çiftler bazında silme, bireysel ortalama atama, ortalama atama ve beklenti maksimizasyonu yöntemleri kullanılarak tamamlanmış ve yapılan analizler sonucunda beklenti maksimizasyonu ve bireysel ortalama atama yöntemlerinin etkili sonuçlar verdiği, liste bazında silme yönteminin yanlı sonuçlara yol açtığı bilgilerine ulaşılmıştır.

Enders (2004), kayıp verinin en çok olabilirlik yaklaşımı kullanılarak güvenilirlik üzerine etkisini araştırmak için yaptığı çalışmasında iki farklı kayıp veri mekanizması altında 1000 kişilik örnekleme sahip yapay veri kullanmıştır. Çok kategorili puanlanan ve sürekli veri içeren veri setlerindeki kayıplar beklenti maksimizasyonu, ortalama atama, liste bazında silme ve çiftler bazında silme yöntemleri kullanılarak tamamlanmıştır ve kovaryans matrisleri oluşturularak alfa katsayıları kestirilmiştir. Yapılan analizler sonucunda beklenti maksimizasyonu yaklaşımının, daha az yanlı güvenilirlik kestirimleri sağladığı, tahminler arası örneklem dalgalanmalarını önemli ölçüde düşürdüğü, daha kesin güven aralığı sonuçları verdiği ulaşılmıştır.

Acock (2005), kayıp veri ile baş etme yöntemlerini karşılaştırmak için 2002 yılında 818 kişiye uygulanan sosyal bir araştırmaya ait ölçekten elde ettiği tam veri setlerinden, TROK ve ROK mekanizmalarına sahip farklı oranlarda kayıp veri içeren veri setleri oluşturmuştur. Bu veri setlerinin liste bazında silme, çiftler bazında silme, ortalama atama, beklenti maksimizasyonu ve çoklu değer atama yöntemleriyle tamamladıktan sonra regresyon kestirimleri ve hata oranlarını incelemiştir. Yapılan analizler sonucunda ortalama atama yönteminin TROK ve ROK varsayımı altında en kötü sonuçları verdiği, çağdaş yöntemlerin geleneksel yöntemlerden daha iyi sonuçları verdiği ulaşılmıştır.

Baygül (2007), kayıp veri ile baş etme yöntemlerini incelemek amacıyla sürekli veri içeren simülasyon veri setleri üzerinde çalışmıştır. Kayıp verileri tamamlamak için yerine ortalama koyma, son gözlemi ileri taşıma, beklenti maksimizasyonu, regresyon ataması yöntemler kullanılmış ve yapılan analizler sonucunda silme ve atama yöntemlerinin anlamlı sonuçlar verdiği ulaşılmıştır.

Ambley, Omar ve Royston (2007), yaptıkları çalışmada iki kategorili veri setlerinde yer alan kayıp verilere uygulanan kayıp veri ile baş etme yöntemlerini karşılaştırmak ve güvenilirlik kestirimlerini karşılaştırmak için ulusal kalp ameliyatı veri tabanını kullanarak farklı örneklem büyüklüklerinde ve farklı kayıp veri

oranlarına sahip veri setleri oluşturmuşlardır. Bu veri setlerini ortalama atama, mod atama, hotdeck atama, çoklu atama yöntemlerini kullanarak tamamlamış ve yapılan analizler sonucunda çoklu atama yönteminin en güçlü kestirimleri veren yöntem olduğuna ulaşmışlardır.

Finch (2010), çalışmasında kategorik değişkenlerden oluşan veri setlerinde kayıp verinin bulunması durumunda kategorik model ve normal model yaklaşımlarını karşılaştırmak amacıyla yapay olarak elde edilmiş 5'li Likert tipi ölçek verisiyle çalışmıştır. Bunun için TROK ve ROK mekanizmasına ait kayıp veri içeren veri setleri oluşturmuştur. Yapılan analizler sonucunda regresyona ait eğitim kestirimlerini, standart hataları ve hipotez testi sonuçlarını karşılaştırmış ve kategorik modelin daha iyi sonuçlar verdiği bilgisine ulaşmıştır.

Çokluk ve Kayrı (2011), çok kategorili puanlanan veri setlerinde kayıp verilerin bulunduğu ve bulunmadığı durumlarda kayıp veri ile baş etme yöntemlerinin faktör yapısına, düzeltilmiş madde toplam korelasyonuna, Cronbach alfa iç tutarlık katsayısına etkisini incelemiştir. Bu çalışma için Ankara Üniversitesi Eğitim Fakültesi İlköğretim Bölümündeki 200 öğretmen adayının ölçek verileri kullanılmıştır. Kayıp verileri tamamlamada seri ortalamaları, yakın noktaların ortalaması, yakın noktaların medyanı, linear interpolasyon yöntemleri kullanılmıştır. Yapılan analizler sonucunda tek faktör yapısının korunduğu, fakat varyans, özdeğerler ve Cronbach alfa katsayısında düşüş olduğu görülmüştür.

Cheema (2012), kayıp verileri tamamlama işlemlerinin tahminlerin doğruluğuna etkisini incelemek için analitik metodun seçimi, kayıp veri ile baş etme yöntemi, örneklem büyüklüğü, eksik verilerin oranı gibi faktörleri incelemiştir. Bu çalışma için simülasyon ve sürekli veri içeren veri kullanılmıştır. Yapılan analizler sonucunda eğitim araştırmaları yapan kişiler için uygulaması kolay bir klavuz ortaya çıkmıştır.

Demir ve Parlak (2012), kayıp veri sorununun Türkiye'deki eğitim araştırmalarında ne ölçüde ele alındığını ve hangi kayıp veri yönteminin daha yaygın kullanıldığını incelemiştir. Bunun için birtakım kriterler doğrultusunda eğitim dergilerindeki makaleleri incelemiştir. Bunun sonucunda Türkiye'deki eğitim araştırmalarında kayıp veri sorununa tamamen ilgisiz kalındığını belirtmişlerdir.

Misztal (2012), R paket programında yer alan kayıp veri atama yöntemlerini karşılaştırmak amacıyla kredi verilen 467 kişilik eksiksiz veri grubunu belirli oranda eksilterek kayıp veri içeren veri seti oluşturmuştur. Lojistik regresyon temelinde yapılan analizler sonucunda en kötü sonucun liste bazında silme yönteminde, en iyi sonucun çoklu atama yöntemleri ile ortaya çıktığı görülmüştür.

Demir (2013), kayıp verilerin bulunması durumunda iki kategorili puanlanan maddelerden oluşan testlerin psikometrik özelliklerini incelemiştir. Bunun için SBS 2011 verileri kullanılmıştır. Kayıp verileri tamamlamak için dizin silme, seri ortalamaları ataması, yakın noktalar ortalama ataması, gözlem birimi ortalaması ataması, yakın noktalar medyan ataması, doğrusal interpolasyon, doğrusal eğilim noktası ataması, regresyon, sıfır atama, beklenti maksimizasyonu, MarkovChain Monte Carlo yöntemleri kullanılmıştır. Yapılan analizlere bakılarak dizin silme ve sıfır atama yöntemlerinin iki kategorili puanlanan veriler için uygun olmadığı, basit atama yöntemlerinin yanlış sonuçlar verdiği, en çok olabilirlik ve çoklu atama yöntemlerinin en uygun yöntemler olduğu belirtilmiştir.

Tavşancıl ve Akbaş (2014), kayıp veri ile baş etme yöntemlerinin ölçeklerin psikometrik özelliklerine etkisini incelemiştir. Bu çalışma için ürettiği yapay verilerden farklı oranlarda kayıp veri içeren veri setleri oluşturmuş ve bu veri setlerini liste bazında silme, Öklid uzaklığına dayalı benzer tepki örüntüsüne dayalı atama, stokastik regresyonla değer atama, beklenti – maksimizasyon algoritması ve çoklu değer atama yöntemleri ile tamamlamıştır. Yapılan analizlere göre silme yönteminin önemli sorunlar yaratabileceği, beklenti maksimizasyonu ve çoklu değer atama yönteminin genel olarak uygun yöntemler olduğu sonucuna varılmıştır.

Köse ve Öztumur (2014), kayıp veri ile baş etme yöntemlerinin varyans analizi parametreleri üzerine etkisini incelemiştir. Farklı örneklem büyüklükleri ve farklı oranlarda kayıp veriye sahip yapay ve sürekli veri içeren veri setleri oluşturulmuş ve bu veri setleri silme yöntemi, yerine ortalama koyma yöntemi, regresyon yöntemi ve beklenti maksimizasyonu yöntemleri ile tamamlanmıştır. Yapılan ANOVA testlerinin sonucunda düşük birimli veri setlerinde regresyon ve BM yöntemlerinin; yüksek birimli veri setlerinde ise regresyon ve yerine ortalama koyma yönteminin daha etkili sonuçlar verdiğine ulaşılmıştır.

Zhu (2014), boylamsal alıřmalarda karřılařılan kayıp veri sorununu incelemek iin yaptığı simülasyon alıřmasında farklı kořullar altında veriler üreterek yanlılık, RMSE, %95 kapsama olasılığı gibi deęerleri karřılařtırmıřtır. Kayıp veriyle bař etme yöntemlerinden ortalama atama, liste bazında silme, gözlenen son deęeri atama ve oklu atama yöntemlerini kullanarak analizlerini yapmıřtır. Yapılan analizler sonucunda TROK varsayımı altında liste bazında silme ve oklu atama yöntemleri en iyi sonuları verirken ROK varsayımı altında oklu atama yöntemi en iyi sonuları vermiřtir.

Kürřad ve Nartgün (2015), kayıp veri ile bař etme yöntemlerini betimsel istatistik, güvenilirlik ve geerlik aısından karřılařtırmıřtır. Bu alıřma PISA 2012 Türkiye uygulamasına katılan öęrencilerin verileri ile yapılmıřtır. Farklı kořullar ieren veri setleri oluřturulmuř ve bu veri setleri liste bazında silme, seri ortalaması, yakın noktaların ortalaması, yakın noktaların medyanı, noktanın doęrusal eęimi, beklenti maksimizasyonu, doęrusal deęer kestirimi, regresyon ataması ve oklu atama yöntemleri ile tamamlanmıřtır. Yapılan analizler sonucunda liste bazında silme yöntemi referans deęerlere en az benzerlik gösterirken; BM, regresyon ve oklu atama yöntemleri referans deęerlere en yakın sonuları vermiřtir.

Iřıkoęlu (2017), kayıp veri ile bař etme yöntemlerinin ölçme deęiřmezlięi üzerine etkisini incelemiřtir. Bu alıřma iin PISA 2015 Türkiye uygulamasına katılan 15 yař grubu öęrencilerin verilerini kullanmıřtır. Farklı örneklem büyüklüklerine ve farklı oranlarda kayıp veri ieren veri setleri oluřturulmuř ve bu veri setleri dizin silme, seri ortalaması atama, regresyon atama, beklenti maksimizasyonu ve oklu atama yöntemleri ile tamamlanmıřtır. Analizler sonucunda beklenti maksimizasyonu ve oklu atama yöntemlerinin daha etkili sonular verdięine ulařılmıřtır.

Arıkan ve Soysal (2018), kayıp verilerin güvenilirlik üzerine etkisini arařtırmak iin farklı örneklem büyüklüklerine sahip ok kategorili puanları ieren yapay veri seti kullanmıřlardır. Farklı oranlarda kayıp veri ieren veri setleri sıfır atama, regresyon atama, ortalama atama ve oklu atama yöntemleri kullanılarak tamamlanmıř ve yapılan analizler sonucunda kayıp veriye dayalı olarak tek bir güvenilirlięin öne ıkmadıęını söylemiřler ve oklu atama ve regresyon yöntemlerinin kullanılmasını önermiřlerdir.

Tamcı (2018), kayıp veri ile baş etme yöntemlerinin değişen madde fonksiyonu (DMF) üzerine etkisini incelemiştir. Bu çalışma için PISA 2012 uygulamasına katılan öğrencilerin verileri kullanılmıştır. Verileri kesikli veri haline dönüştürüp analizlere bu veri setleri üzerinden devam etmiştir. Farklı örneklem büyüklüklerinde ve farklı oranlarda kayıp veri içeren veri setleri oluşturulmuş ve bu veri setleri sıfır atama (SA), çoklu atama (ÇA) ve beklenti maksimizasyonu (BM) yöntemleri ile tamamlanmıştır. Veri setlerine SIBTEST yöntemiyle DMF analizi yapılmıştır. Analizler sonucunda I. tip hata açısından SA yöntemi %30 kayıp veri durumunda diğer kayıp veri oranlarına göre, A ve B düzeyinde DMF için C düzeyindeki DMF'ye göre daha iyi performans sergilemiştir. İncelenen tüm koşullar açısından ÇDA yöntemi I. tip hata açısından iyi sonuçlar vermiştir. BM yöntemi A düzeyinde DMF koşulunda, B ve C düzeyinde DMF'ye göre daha iyi performans göstermiştir.

Sıddıkoğlu, Erdoğan, Gökmen, Kutlay (2019), kayıp veri yöntemlerinin WHODAS-2.0 ölçeğinin psikometrik özellikleri üzerine etkisini incelediği çalışmada çok kategorili puanlanan veri setleri farklı oranlarda eksiltilerek kayıp veriler oluşturulmuştur. Yapılan analizler doğrultusunda, maddelerin çalışma dışında tutulmasının olumsuz sonuçlara yol açtığı, yanıt fonksiyonu kullanılarak değer atama yapıldığında elde edilen güvenilirlik ölçütlerinin tam veri setinden elde edilen güvenilirlik ölçütlerine oldukça benzer çıktığı bulunmuştur.

İlgili çalışmalara genel olarak bakıldığında, genel olarak sürekli veya çok kategorili puanlanan veri setlerinde kayıp veri yöntemleri ile ilgili çalışmaların gerçekleştirildiği görülmektedir. Bu çalışmalarda psikometrik özellikler içinde güvenilirlik kestirimlerine yer verilse de genellikle KTK baz alınarak çalışmalar yürütülmüş, MTK'ya pek yer verilmemiştir. Bu çalışmada iki kategorili veri setleri üzerinde kayıp veri tamamlama yöntemlerinin Cronbach alfa ve MTK temelli ampirik güvenilirlik katsayısı üzerine etkisi incelenmiştir.

Bölüm 3

Yöntem

Bu araştırma farklı şartlar altında kayıp veriyle başa çıkma yöntemlerinin güvenilirlik üzerine etkisinin incelendiği simülasyon bir araştırmadır. Bu bağlamda çalışma, temel araştırma biçimindedir.

Simülasyon Verisi

Bu araştırmada, test uzunluğu 20 ve 40 olan, $n=200$, 500 ve 1000 birimlik yapay veri setleri olmak üzere toplamda 6 koşul altında veri setleri üretilmiştir. Her bir koşul için bu işlem 25 replikasyon yapılmıştır. Verilen üretilmesi ve eksiltilmesi R 3.6.3 programında kod yazılarak yapılmıştır. Verilerin analizinde R 3.6.3 ve SPSS 22.0 programları kullanılmıştır.

Simülasyon Koşulları

İki kategorili puanlanan maddelerden oluşan testlerde madde sayısının 20'nin altında olduğu durumlarda, ölçülmek istenen özelliğin süreklilik göstermesi zorlaşacağından (Aiken, 2009) alt sınır olan 20 ve madde sayısının artması durumunda güvenilirlik kestirimlerini incelemek ve karşılaştırma yapmak amacıyla 40 maddelik veri setleri oluşturulmuştur.

Kayıp veri ile ilgili yapılan çalışmalar incelendiğinde (Tavşancıl ve Akbaş, 2014; Kürşad ve Nartgün, 2015; Bayhan, 2018; Tamcı, 2018) örneklem büyüklüklerinin 100 ve 1000 arasında değiştiği görülmektedir. Bu nedenle bu çalışmada küçük, orta ve büyük örneklemi temsil etmeleri için 200, 500 ve 1000 kişilik veri setleri oluşturulmuştur.

Kayıp veri ile ilgili yapılan çalışmalar incelendiğinde (Enders, 2003; Demir, 2013; Acock, 2005; Işıkoğlu, 2017; Tamcı, 2018) kayıp veri oranlarının ise genellikle %5 ve %50 arasında olduğu görülmektedir. Bu çalışmada %5, %10 ve %20 oranlarında eksiltme yapılmıştır.

Kayıp veriler oluşturulurken 2-parametrelili logistik model kullanılmış, madde ayırteciçiliği(a) $\text{meanlog}=0.0$ ve $\text{sdlog}=0.3$ alınarak log normal dağılıma ve madde güçlüğü(b) katsayıları ortalama 0, standart sapma 1 alınarak normal dağılıma göre

Rstudio programında yapay olarak üretilmiştir. Birey parametrelerinin dağılımı için veri setleri üretilirken ortalama 0, standart sapma 1 olacak şekilde kod yazılmıştır.

Verilerin üretilmesi ve analizi aşamasında 25 replikasyon yapılmıştır. 25 replikasyondan sonra sonuçların birbirine oldukça benzer olması, kayıp veri yöntemlerinden çoklu atama yöntemi ile kayıp verileri tamamlamanın örneklem büyüklüğünün ve kayıp veri oranının fazla olduğu durumlarda oldukça uzun sürmesi nedenleriyle replikasyon sayısı minimumda tutulmuştur.

Verilerin Analizi

Verilerin analizi sürecinde 6 koşul altında üretilen her bir replikasyon için veri setlerinin Cronbach-alfa güvenilirlik katsayısı ve MTK temelli ampirik güvenilirlik katsayısı hesaplanmıştır. Literatüre bakıldığında ikili puanlanan maddeler içeren veri setlerinin parametrelerini hesaplamada MTK'ya dayalı modellerin sıklıkla kullanıldığı görülmektedir (Little ve Rubin, 1987). Tam veri setlerinden elde edilen güvenilirlik kestirimleri referans değerler olarak, tamamlanmış veri setlerinden elde edilen güvenilirlik kestirimleri ile karşılaştırmada kullanılmıştır.

Daha sonra, her bir replikasyon için, 20 ve 40 soruluk 200, 500 ve 1000 kişilik örnekleme sahip tam veri setlerinin her birinden %5, %10 ve %20 oranlarında kayıp veri içeren 18 koşul altında veri setleri oluşturulmuştur. Eksiltmeler tamamen rassallık göz önüne alınarak oluşturulmuş olsa da, oluşturulan veri setlerinin TROK mekanizmasına sahip olup olmadığını kontrol etmek için tam veri setlerinden elde edilen güvenilirlik katsayıları ile eksik veri setlerinden elde edilen güvenilirlik katsayıları karşılaştırılmıştır. Bu karşılaştırmada, iki ilişkisiz ölçmelerden elde edilen puanların birbirlerinden anlamlı bir farklılık gösterip göstermediğini test eden (Büyüköztürk, 2002) Mann Whitney U-Testi kullanılmıştır. Burada da tam veri setleri ve eksik veri setleri olmak üzere iki grup yer aldığından, bu testin kullanılması uygun görülmüştür. Aşağıda Tablo 1 ve Tablo 2'de Mann Whitney U-Testi sonuçlarına yer verilmiştir.

Tablo 1

Veri Setlerinin (Alfa) Güvenirlik Değerlerinin U-Testi Sonucu

Grup	n	Sıra Ortalaması	Sıra Toplamı	U	P
tamveri	6	12.17	73	52	.923
eksikveri	18	12.61	227		

Tablo 1'e göre sıra ortalamaları arasındaki farkın 0.05 alfa düzeyinde anlamlı olmaması kayıp verilerin tamamen rassal olarak dağıldığını göstermektedir.

Tablo 2

Veri Setlerinin (MTK Temelli Ampirik) Güvenirlik Değerlerinin U-Testi Sonucu

Grup	n	Sıra Ortalaması	Sıra Toplamı	U	P
tamveri	6	13.67	82	47	.673
eksikveri	18	12.11	218		

Tablo 2'ye göre sıra ortalamaları arasındaki farkın 0.05 alfa düzeyinde anlamlı olmaması kayıp verilerin tamamen rassal olarak dağıldığını göstermektedir.

Eksiltelen veri setleri DS, sıfır atama, BM, RA ve ÇA yöntemleri kullanılarak tamamlanmıştır. DS yönteminde içinde kayıp veri bulunan gözlemler analize dahil edilmemiştir. Bunun sonucunda örneklem sayılarında ciddi azalmalar meydana gelmiştir. Dizin silme yönteminin uygulanması sonucu kalan örneklem sayılarına ait sonuçlar, replikasyon sonuçları da dikkate alınarak, Tablo 3'te verilmiştir.

Tablo 3

DS Yöntemi Sonrası Veri Gruplarının Örneklem Sayıları

Test Uzunluğu	Örneklem Büyüklüğü	5%	10%	20%
20	200	61-83	16-33	0-4
	500	164-200	50-76	1-12
	1000	333-376	97-142	7-16
40	200	17-36	0-6	0
	500	56-76	4-13	0-1
	1000	111-150	8-18	0-2

Tablo 3'te görüldüğü üzere 20 soruluk veri setlerinde, 200 kişilik örneklem büyüklüğünde, dizin silme yöntemiyle %5 oranında kayıp bulunduğu, örneklem büyüklüğü 61 ile 83 arasında, %10 oranında kayıp bulunduğu 16 ile 33 arasında, %20 oranında kayıp bulduğundaysa 0 ile 4 arasında değişmektedir. 500 kişilik örneklem büyüklüğünde, dizin silme yöntemiyle %5 oranında kayıp bulunduğu örneklem büyüklüğü 164 ile 200 arasında, %10 oranında kayıp bulunduğu 50 ile 76 arasında, %20 oranında kayıp bulduğundaysa 1 ile 12 arasında değişmektedir. 1000 kişilik örneklem büyüklüğünde, dizin silme yöntemiyle %5 oranında kayıp bulunduğu örneklem büyüklüğü 333 ile 376 arasında, %10 oranında kayıp bulunduğu 97 ile 142 arasında, %20 oranında kayıp bulunduğu 7 ile 16 arasında değişmektedir.

40 soruluk veri setlerinde 200 kişilik örneklem büyüklüğünde dizin silme yöntemiyle %5 oranında kayıp bulunduğu, örneklem büyüklüğü 17 ile 36 arasında, %10 oranında kayıp bulunduğu 0 ile 6 arasında değişmekte ve %20 oranında kayıp bulunduğu elimizde hiç veri kalmamaktadır. 500 kişilik örneklem büyüklüğünde, dizin silme yöntemiyle %5 oranında kayıp bulunduğu örneklem büyüklüğü 56 ile 76 arasında, %10 oranında kayıp bulunduğu 4 ile 13 arasında değişmekte, %20 oranında kayıp bulunduğu 0 veya 1 veri kalmaktadır. 1000 kişilik örneklem büyüklüğünde, dizin silme yöntemiyle %5 oranında kayıp bulunduğu büyüklüğü 111 ile 150 arasında, %10 oranında kayıp bulunduğu 8 ile 18 arasında, %20 oranında kayıp bulunduğu 0 ile 2 arasında değişmektedir.

Dizin silme yöntemi sonucu elde edilen örneklem büyüklüklerinde güvenilirlik kestirimleri yapmak yanlış ve hatalı olacağından, bu yöntem analizlere dahil edilmemiştir.

Bir diğer kayıp veri yöntemi olan BM'de test uzunluğu 40 ve kayıp veri oranı %20 olduğunda, iki kategorili veri setlerine atama yapılamamaktadır. Bu nedenle BM yöntemi ile, tüm replikasyonlarda, 18 koşul altında eksiltilmiş veri setlerinin 15'i tamamlanabilmiştir. Bu da yöntemin sınırlılığını yansıtmaktadır.

Sıfır atama, RA ve ÇA yöntemleriyle tüm veri setleri tamamlanmış ve analizler yapılmıştır. ÇA yönteminde her bir eksik veri grubu için 10 farklı veri seti atanmış ve bu veri setlerinin güvenilirliklerinin ortalaması alınarak kullanılmıştır.

Kayıp verilere atamalar yapıldıktan sonra her bir veri setinin Cronbach alfa ve MTK temelli ampirik güvenilirlik katsayıları hesaplanmış ve aralarında anlamlı bir farklılık bulunup bulunmadığını kontrol etmek için iki veya daha fazla ilişkisiz ölçmelerden elde edilen puanlar arasında anlamlı bir farklılık olup olmadığını test eden Kruskal-Wallis veya iki ilişkisiz ölçmelerden elde edilen puanların birbirlerinden anlamlı bir farklılık gösterip göstermediğini test eden Mann Whitney U-Testi kullanılmıştır.

Kayıp veri oranı (%5, %10, %20; grup sayısı=3), örneklem büyüklüğü (200, 500, 1000; grup sayısı=3), kayıp veri yöntemleri (tam veri, sıfır atama, regresyon, çoklu atama, beklenti maksimizasyonu; grupsayısı=5) ele alındığında, karşılaştırmada Kruskal-Wallis Testi kullanılmıştır. Test uzunluğu (20, 40; grup sayısı=2) ele alındığında, Mann Whitney U-Testi kullanılmıştır.

Veriler Kruskal-Wallis ve Mann Whitney U analizleri için hazırlanırken, her bir alt problem için farklı gruplara göre düzenlenmiştir. Cronbach alfa ve MTK temelli ampirik olarak ayrılmak şartıyla güvenilirlik kestirimlerinin betimsel istatistiklerine Tablo 4, Tablo 5 ve Tablo 6'da yer verilmiştir.

Tablo 4

Veri Setlerinin Betimsel İstatistik Sonuçları (Test Uzunluğu)

	Test Uzunluğu	Ortalama	Standart Sapma	Çarpıklık	Basıklık
ALFA	20soru	.750	.071	-.781	-.358
	40soru	.832	.058	-.818	-.711
MTK	20soru	.797	.034	-.444	-.886
	40soru	.865	.035	-.706	-.907

Tablo 4'e bakıldığında Cronbach alfa güvenilirlik kestirimlerine ilişkin ortalamaların 20 ve 40 soruluk veri setlerinde sırasıyla 0.750 ve 0.83, standart sapmanın 0.071 ve 0.058, çarpıklığın -0.781 ve -0.818, basıklığın ise -0.358 ve -0.711 olduğu görülmektedir. MTK temelli ampirik güvenilirlik kestirimlerine ilişkin ortalamaların 20 ve 40 soruluk veri setlerinde sırasıyla 0.797 ve 0.865, standart sapmanın 0.034 ve 0.035, çarpıklığın -0.444 ve -0.706, basıklığın ise -0.886 ve -0.907 olduğu görülmektedir.

Tablo 5

Veri Setlerinin Betimsel İstatistik Sonuçları (Örneklem Büyüklüğü)

	Örneklem Büyüklüğü	Ortalama	Standart Sapma	Çarpıklık	Basıklık
ALFA	200	.834	.045	-.331	-.677
	500	.700	.063	-.435	-.371
	1000	.812	.057	-.261	-.961
MTK	200	.859	.036	-.227	-1.008
	500	.787	.035	-.110	-1.239
	1000	.842	.042	-.125	-1.422

Tablo 5'e bakıldığında Cronbach alfa güvenilirlik kestirimlerine ilişkin ortalamanın 200,500 ve 1000 örneklem büyüklüğüne sahip veri setlerinde sırasıyla 0.834, 0.700 ve 0.812, standart sapmanın 0.045, 0.063, 0.057, çarpıklığın -0.331, -0.435, -0.261, basıklığın -0.677, -0.371, -0.961 olduğu görülmektedir. MTK temelli ampirik güvenilirlik kestirimlerine ilişkin ortalamanın 200,500 ve 1000 örneklem büyüklüğüne sahip veri setlerinde sırasıyla 0,859, 0,787, 0,842, standart sapmanın0.036, 0,035, 0.042, çarpıklığın-0.227, -0.110, -0.125, basıklığın -1.008, -1.239, -1.422 olduğu görülmektedir.

Tablo 6

Veri Setlerinin Betimsel İstatistik Sonuçları (Kayıp Oranı)

	Kayıp Oranı(%)	Ortalama	Standart Sapma	Çarpıklık	Basıklık
ALFA	5%	.794	.075	-.457	-.685
	10%	.789	.079	-.644	-.239
	20%	.769	.087	-.619	-.135
MTK	5%	.837	.047	-.140	-1.014
	10%	.833	.047	-.157	-.925
	20%	.821	.051	-.079	-.672

Tablo 6'ya bakıldığında Cronbach alfa güvenilirlik kestirimlerine ilişkin ortalamanın %5, %10, %20 oranlarında kayıp veriye sahip veri setlerinde sırasıyla 0.794, 0.789, 0.769, standart sapmanın 0.075, 0.079, 0.087, çarpıklığın -0.457, -0.644, -0.619, basıklığın -0.685, -0.239, -0.135 olduğu görülmektedir. MTK temelli ampirik güvenilirlik kestirimlerine ilişkin ortalamanın %5, %10, %20 oranlarında kayıp veriye sahip veri setlerinde sırasıyla 0.837, 0.833, 0.821, standart sapmanın 0.047,

0.047, 0.051, çarpıklığın -0.140, -0.157, -0.079, basıklığın -1.014, -0.925, -0.672 olduğu görülmektedir.

Bu çalışmada bir veri grubunun analizlerine yer verilmiştir. Yapılan replikasyonlar çıkan sonuçları destekler niteliktedir.

Bölüm 4

Bulgular ve Yorumlar

Bu bölümde alt problemlerin sırasına göre araştırma bulgularına yer verilmiştir.

Eksiksiz veri setlerinden kayıp veriler oluşturulduktan sonra kayıp veri içeren veri setlerinin Cronbach alfa ve MTK temelli ampirik güvenilirlik katsayıları hesaplanmış Tablo 7 ve Tablo 8'de verilmiştir.

Tablo 7

Veri Setlerinin Kayıp Veri Oranlarına Göre Cronbach Alfa Güvenirlik Katsayıları

Test Uzunluğu	Örneklem Büyüklüğü	%0	%5	%10	%20
20	200	.809	.807	.803	.551
	500	.671	.676	.520	.854
	1000	.781	.778	.791	.766
40	200	.886	.891	.880	.874
	500	.769	.676	.856	.854
	1000	.876	.892	.883	.874

Tablo 7 incelendiğinde, 20 soruluk veri setlerinde 200 kişilik örneklem büyüklüğünde farklı oranlarda kayıp veri bulunduğunda Cronbach alfa güvenilirlik katsayısı 0.551 ve 0.809 arasında, 500 kişilik örneklem büyüklüğünde 0.520 ve 0.854 arasında, 1000 kişilik örneklem büyüklüğünde 0.766 ve 0.791 arasında değişmektedir. 40 soruluk veri setlerinde 200 kişilik örneklem büyüklüğünde farklı oranlarda kayıp veri bulunduğunda Cronbach alfa güvenilirlik katsayısı 0.874 ve 0.891 arasında, 500 kişilik örneklem büyüklüğünde 0.676 ve 0.856 arasında, 1000 kişilik örneklem büyüklüğünde 0.874 ve 0.892 arasında değişmektedir.

Tablo 8

Veri Setlerinin Kayıp Veri Oranlarına Göre MTK Temelli Ampirik Güvenirlik Katsayıları

Test Uzunluğu	Örneklem Büyüklüğü	0%	5%	10%	20%
20	200	.839	.833	.803	.906
	500	.764	.806	.520	.785
	1000	.814	.808	.825	.781
40	200	.901	.897	.893	.887
	500	.826	.821	.814	.788
	1000	.888	.883	.878	.895

Tablo 8'e bakıldığında, 20 soruluk veri setlerinde 200 kişilik örneklem büyüklüğünde farklı oranlarda kayıp veri bulunduğunda MTK temelli ampirik güvenirlilik katsayısı 0.803 ve 0.906 arasında, 500 kişilik örneklem büyüklüğünde 0.520 ve 0.806 arasında, 1000 kişilik örneklem büyüklüğünde 0.781 ve 0.825 arasında değişmektedir. 40 soruluk veri setlerinde 200 kişilik örneklem büyüklüğünde farklı oranlarda kayıp veri bulunduğunda MTK temelli ampirik güvenirlilik katsayısı 0.887 ve 0.901 arasında, 500 kişilik örneklem büyüklüğünde 0.788ve 0.826 arasında, 1000 kişilik örneklem büyüklüğünde 0.878 ve 0.895 arasında değişmektedir.

Birinci Alt Probleme İlişkin Bulgular

Bu bölümde *“Farklı örneklem büyüklüğünde ve farklı test uzunluğundaki eksiksiz veri setleri ile bu veri setlerinin farklı oranlarda eksiltilmesiyle oluşan veri setlerine uygulanan dizin silme, sıfır atama, regresyon atama, beklenti maksimizasyonu ve çoklu atama yöntemleri ile tamamlanarak elde edilen yeni veri setleri arasında Cronbach alfa güvenirlilik kestirimleri açısından bir farklılık var mıdır?*

a. Test uzunluğu dikkate alınarak karşılaştırıldığında Cronbach alfa güvenirlilik kestirimleri arasında bir farklılık var mıdır?

b. Örneklem büyüklüğü dikkate alınarak karşılaştırıldığında Cronbach alfa güvenirlilik kestirimleri arasında bir farklılık var mıdır?

c. Kayıp veri oranı dikkate alınarak karşılaştırıldığında Cronbach alfa güvenirlilik kestirimleri arasında bir farklılık var mıdır?” sorusuna cevap aranmaktadır.

Eksiksiz veri setlerinin Cronbach alfa güvenilirlik katsayıları hesaplanmış ve referans değerler olarak karşılaştırmada kullanılmıştır. Daha sonra tam veri setlerinden %5, %10 ve %20 oranlarında kayıp veriler içeren veri setleri oluşturulmuş ve DS, sıfır atama, BM, RA ve ÇA yöntemleri ile tamamlamalar yapılmıştır. Tablo 1’de gösterildiği üzere DS yöntemi sonrasında veri setlerinde ciddi oranda azalmalar meydana geldiğinden, DS yöntemi analizlere dahil edilmemiştir.

Çalışmada ele alınan simülasyon koşullarında kestirilen Cronbach alfa katsayıları Tablo 9’da verilmiştir.

Tablo 9

Tamamlanmış Veri Setlerinin Cronbach Alfa Güvenirlik Katsayıları

Yöntem		200				500				1000			
		0%	5%	10%	20%	0%	5%	10%	20%	0%	5%	10%	20%
Eksiksiz Veri	20	.809				.671				.781			
	40	.886				.769				.876			
0 atama	20		.797	.779	.728		.644	.600	.545		.756	.736	.682
	40		.872	.856	.833		.747	.727	.686		.862	.844	.809
Regresyon	20		.809	.805	.805		.667	.662	.664		.780	.779	.766
	40		.886	.881	.880		.765	.772	.766		.877	.874	.875
Çoklu Atama	20		.807	.814	.800		.669	.663	.663		.781	.778	.776
	40		.885	.881	.877		.768	.773	.766		.875	.875	.873
BM	20		.805	.811	.786		.671	.662	.644		.775	.772	.754
	40		.885	.881	*		.764	.770	*		.873	.871	*

*. Veri setleri tamamlanamadığından güvenilirlik kestirimleri yapılamamıştır.

Tablo 9’da görüldüğü üzere 20 soru ve 200 kişilik örneklemin yer aldığı eksiksiz veri setinin Cronbach alfa güvenilirlik katsayısı 0.809’dur. Farklı oranlarda eksiltmeler ve farklı yöntemlerle veri setlerinin tamamlanmaları sonucunda güvenilirlik katsayıları 0.728 ve 0.814 arasında değişmektedir. En yakın kestirimler regresyon yönteminde görülmektedir. 40 soru ve 200 kişilik örnekleme sahip eksiksiz veri setinin Cronbach alfa güvenilirlik katsayısı 0.886’dır. Farklı oranlarda eksiltmeler ve farklı yöntemlerle veri setlerinin tamamlanmaları sonucunda güvenilirlik katsayıları 0.833 ve 0.886 arasında kestirilmiştir. En yakın kestirimler

regresyon ve BM yöntemlerinde görülmektedir. 20 soru ve 500 kişilik örneklemin yer aldığı eksiksiz veri setinin Cronbach alfa güvenilirlik katsayısı 0.671'dir. Farklı oranlarda eksiltmeler ve farklı yöntemlerle veri setlerinin tamamlanmaları sonucunda güvenilirlik katsayıları 0.6 ve 0.669 arasında değişmektedir. En yakın kestirimler çoklu atama yönteminde görülmektedir. 40 soru ve 500 kişilik örneklemin yer aldığı eksiksiz veri setinin Cronbach alfa güvenilirlik katsayısı 0.769'dur. Farklı oranlarda eksiltmeler ve farklı yöntemlerle veri setlerinin tamamlanmaları sonucunda güvenilirlik katsayıları 0.727 ve 0.773 arasında değişmektedir. En yakın kestirimler çoklu atama yönteminde görülmektedir. 20 soru ve 1000 kişilik örneklemin yer aldığı eksiksiz veri setinin Cronbach alfa güvenilirlik katsayısı 0.781'dir. Farklı oranlarda eksiltmeler ve farklı yöntemlerle veri setlerinin tamamlanmaları sonucunda güvenilirlik katsayıları 0.736 ve 0.781 arasındadır. En yakın kestirimler çoklu atama yönteminde görülmektedir. 40 soru ve 1000 kişilik örnekleme sahip eksiksiz veri setinin Cronbach alfa güvenilirlik katsayısı 0.876'dır. Farklı oranlarda eksiltmeler ve farklı yöntemlerle veri setlerinin tamamlanmaları sonucunda güvenilirlik katsayıları 0.809 ve 0.877 arasında değişmektedir. En yakın kestirimler regresyon yönteminde görülmektedir.

- a. Bu aşamada veriler 20 ve 40'ar soru olmak üzere gruplanarak analize devam edilmiştir. Yapılan analizler sonucunda 0.05 alfa düzeyinde hesaplanan $z=-4.72$ değeri kritik değer olan ± 1.96 aralığının dışında olduğundan farklı test uzunluklarına sahip veri setlerinin Cronbach alfa güvenilirlik kestirimlerinin sıra ortalamaları arasında manidar bir farklılık bulunmuştur ($U=257$, $p<0.05$). Etki büyüklüğünün 0.361 olması aradaki farkın güçlü olduğunu göstermektedir. Ayrıca 20 soruluk veri setinin sıra ortalamasının 26.59, 40 soruluk veri setinin sıra ortalamasının 50.36 olduğu ve bu değerler göz önüne alındığında 40 soruluk veri setinin daha güvenilir sonuçlar verdiği görülmüştür.

Veri setlerine uygulanan yöntemler karşılaştırıldığında $\chi^2=4.436$ değeri kritik değer olan 9.488'den düşük olduğundan veri setlerinin Cronbach alfa güvenilirlik kestirimlerinin sıra ortalamaları arasında 0.05 alfa düzeyinde manidar bir farklılık bulunmamıştır ($\chi^2=4.436$, $p>0.05$). Yöntemlerin sıra ortalamalarına bakıldığında tam verinin sıra ortalamasının 44.75 olduğu ve 42.00 ile 41.36 sıra ortalaması ile bu

değere en yakın sonuçların çoklu atama ve regresyon yöntemlerinde olduğu görülmektedir. Ayrıca 29.44 ile en farklı değer sıfır atama yönteminde ortaya çıkmıştır.

- b. Bu aşamada veriler örneklem büyüklüğü bakımında 200, 500 ve 1000 olmak üzere 3 gruba ayrılmış ve analizlere devam edilmiştir. Yapılan analizler sonucunda $\chi^2=43.600$ değeri kritik değer olan 5.991'den büyük olduğundan farklı örneklem büyüklüklerine sahip veri setlerinin Cronbach alfa güvenilirlik kestirimlerinin sıra ortalamaları arasında 0.05 alfa düzeyinde manidar bir farklılık bulunmuştur ($\chi^2=43.600$, $p<0.05$). Hangi gruplar arasında farklılık olduğunu bulmak amacıyla yapılan ikili karşılaştırmalar sonucunda 200 ve 500, 500 ve 1000 kişilik örnekleme sahip veri setleri arasında 0.05 alfa düzeyinde manidar farklılık olduğu görülmüştür. Etki büyüklüğünün 0.538 olması aradaki farkın güçlü olduğunu göstermektedir. Sıra ortalamalarına bakıldığında 200, 500 ve 1000 kişilik veri setlerinin sıra ortalamalarının sırasıyla 54.66, 15.32 ve 44.02 olduğu görülmektedir. Buradan 200 kişilik veri setlerinin, diğer veri setlerine göre daha güvenilir sonuçlar verdiği ortaya çıkmaktadır.

Veri setlerine uygulanan yöntemler karşılaştırıldığında $\chi^2=4.436$ değeri kritik değer olan 9.488'den düşük olduğundan veri setlerinin Cronbach alfa güvenilirlik kestirimlerinin sıra ortalamaları arasında 0.05 alfa düzeyinde manidar bir farklılık bulunmamıştır ($\chi^2=4.436$, $p>0.05$). Yöntemlerin sıra ortalamalarına bakıldığında tam verinin sıra ortalamasının 44.75 olduğu ve 42.00 ile 41.36 sıra ortalaması ile bu değere en yakın sonuçların çoklu atama ve regresyon yöntemlerinde olduğu görülmektedir. Ayrıca 29.44 ile en farklı değer sıfır atama yönteminde ortaya çıkmıştır.

- c. Bu aşamada veriler kayıp veri oranları bakımında %5, %10 ve %20 olmak üzere 3 gruba ayrılmış ve analizlere devam edilmiştir. Yapılan analizler sonucunda $\chi^2=1.092$ değeri kritik değer olan 5.991'den düşük olduğundan veri setlerinin farklı örneklem büyüklüklerine sahip veri setlerinin Cronbach alfa güvenilirlik kestirimlerinin sıra ortalamaları arasında 0.05 alfa düzeyinde manidar bir farklılık bulunmamıştır ($\chi^2=1.092$, $p>0.05$). Sıra

ortalamalarına bakıldığında %5, %10 ve %20 oranlarında kayıp içeren veri setlerinin sıra ortalamalarının sırasıyla 46.40, 45.35 ve 39.83 olduğu görülmektedir. Buradan %5 oranında kayıp veri içeren veri setlerinin, diğer veri setlerine göre daha güvenilir sonuçlar verdiği ortaya çıkmaktadır.

Veri setlerine uygulanan yöntemler karşılaştırıldığında $\chi^2=5.092$ değeri kritik değer olan 9.488'den düşük olduğundan veri setlerinin Cronbach alfa güvenilirlik kestirimlerinin sıra ortalamaları arasında 0.05 alfa düzeyinde manidar bir farklılık bulunmamıştır ($\chi^2=5.092$, $p>0.05$). Yöntemlerin sıra ortalamalarına bakıldığında eksiksiz veri setinin sıra ortalamasının 50.75 olduğu ve bunu 47.28 ve 46.58 ile çoklu atama ve regresyon yöntemlerinin en yakın sonuçlar verdiği görülmektedir. 33.39 sıra ortalaması ile sıfır atama yöntemi en uzak sonuçları veren yöntemdir.

İkinci Alt Probleme İlişkin Bulgular

Bu bölümde “*Farklı örneklem büyüklüğünde ve farklı test uzunluğundaki eksiksiz veri setleri ile bu veri setlerinin farklı oranlarda eksiltilmesiyle oluşan veri setlerine uygulanan dizin silme, sıfır atama, regresyon atama, beklenti maksimizasyonu ve çoklu atama yöntemleri ile tamamlanarak elde edilen yeni veri setleri arasında MTK temelli ampirik ampirik güvenilirlik kestirimleri açısından bir farklılık var mıdır?*”

a. *Test uzunluğu dikkate alınarak karşılaştırıldığında arasında MTK temelli ampirik ampirik güvenilirlik kestirimleri arasında bir farklılık var mıdır?*

b. *Örneklem büyüklüğü dikkate alınarak karşılaştırıldığında arasında MTK temelli ampirik ampirik güvenilirlik kestirimleri arasında bir farklılık var mıdır?*

c. *Kayıp veri oranı dikkate alınarak karşılaştırıldığında arasında MTK temelli ampirik ampirik güvenilirlik kestirimleri arasında bir farklılık var mıdır?”* sorusuna cevap aranmaktadır.

Eksiksiz veri setlerinin MTK temelli ampirik güvenilirlik katsayıları hesaplanmış ve referans değerler olarak karşılaştırmada kullanılmıştır. Daha sonra tam veri setlerinden %5, %10 ve %20 oranlarında kayıp veriler içeren veri setleri oluşturulmuş ve DS, sıfır atama, BM, RA ve ÇA yöntemleri ile tamamlamalar

yapılmıştır. Tablo 1’de gösterildiği üzere DS yöntemi sonrasında veri setlerinde ciddi oranda azalmalar meydana geldiğinden DS yöntemi analizlere dahil edilmemiştir.

Çalışmada ele alınan simülasyon koşullarında kestirilen MTK temelli ampirik güvenilirlik katsayıları Tablo 10’da verilmiştir.

Tablo 10

Tamamlanmış Veri Setlerinin MTK temelli ampirik Güvenirlik Katsayıları

Yöntem	200				500				1000				
	0%	5%	10%	20%	0%	5%	10%	20%	0%	5%	10%	20%	
Eksiksiz Veri	20	.839			.764				.814				
	40	.901			.826				.888				
sıfır atama	20		.831	.811	.780		.759	.742	.712		.799	.789	.757
	40		.889	.877	.860		.815	.805	.780		.881	.870	.848
Regresyon	20		.841	.834	.827		.761	.755	.768		.815	.814	.809
	40		.901	.897	.896		.826	.829	.822		.889	.887	.889
Çoklu Atama	20		.839	.838	.831		.764	.764	.762		.815	.814	.814
	40		.901	.897	.895		.828	.829	.821		.888	.888	.887
BM	20		.839	.833	.830		.761	.768	.754		.812	.810	.801
	40		.900	.897	*		.825	.827	*		.887	.886	*

*. Veri setleri tamamlanamadığından güvenilirlik kestirimleri yapılamamıştır.

Tablo 10’da görüldüğü üzere 20 soru ve 200 kişilik örneklemin yer aldığı eksiksiz veri setinin MTK temelli ampirik güvenilirlik katsayısı 0.839’dır. Farklı oranlarda eksiltmeler ve farklı yöntemlerle veri setlerinin tamamlanmaları sonucunda güvenilirlik katsayıları 0.780 ve 0.841 arasında değişmektedir. En yakın kestirimler çoklu atama yönteminde görülmektedir. 40 soru ve 200 kişilik örnekleme sahip eksiksiz veri setinin MTK temelli ampirik güvenilirlik katsayısı 0.901’dir. Farklı oranlarda eksiltmeler ve farklı yöntemlerle veri setlerinin tamamlanmaları sonucunda güvenilirlik katsayıları 0.860 ve 0.901 arasında değişmektedir. En yakın kestirimler regresyon yönteminde görülmektedir. 20 soru ve 500 kişilik örneklemin yer aldığı eksiksiz veri setinin MTK temelli ampirik güvenilirlik katsayısı 0.764’dir. Farklı oranlarda eksiltmeler ve farklı yöntemlerle veri setlerinin tamamlanmaları sonucunda güvenilirlik katsayıları 0.712 ve 0.768 arasında değişmektedir. En yakın kestirimler çoklu atama yöntemine görülmektedir. 40 soru ve 500 kişilik örnekleme

sahip eksiksiz veri setinin MTK temelli ampirik güvenilirlik katsayısı 0.826'dir. Farklı oranlarda eksiltmeler ve farklı yöntemlerle veri setlerinin tamamlanmaları sonucunda güvenilirlik katsayıları 0.780 ve 0.829 arasında değişmektedir. En yakın kestirimler regresyon yönteminde 20 soru ve 1000 kişilik örneklemin yer aldığı eksiksiz veri setinin MTK temelli ampirik güvenilirlik katsayısı 0.814'dir. Farklı oranlarda eksiltmeler ve farklı yöntemlerle veri setlerinin tamamlanmaları sonucunda güvenilirlik katsayıları 0.757 ve 0.815 arasındadır. En yakın kestirimler çoklu atama yönteminde görülmektedir. 40 soru ve 1000 kişilik örnekleme sahip eksiksiz veri setinin MTK temelli ampirik güvenilirlik katsayısı 0.888'dir. Farklı oranlarda eksiltmeler ve farklı yöntemlerle veri setlerinin tamamlanmaları sonucunda güvenilirlik katsayıları 0.848 ve 0.889 arasında değişmektedir. En yakın kestirimler çoklu atama yönteminde görülmektedir.

- a. Bu aşamada veriler 20 ve 40'ar soru olmak üzere gruplanarak analize devam edilmiştir. Yapılan analizler sonucunda 0.05 alfa düzeyinde hesaplanan $z=-5.838$ değeri kritik değer olan ± 1.96 aralığının dışında olduğundan farklı test uzunluklarına sahip veri setlerinin Cronbach alfa güvenilirlik kestirimlerinin sıra ortalamaları arasında manidar bir farklılık bulunmuştur ($U=151.5$, $p<0.05$). Etki büyüklüğünün 0.496 olması aradaki farkın güçlü olduğunu göstermektedir. Ayrıca 20 soruluk veri setinin sıra ortalamasının 23.88, 40 soruluk veri setinin sıra ortalamasının 53.29 olduğu ve bu değerler göz önüne alındığında 40 soruluk veri setinin daha güvenilir sonuçlar verdiği görülmüştür.

Veri setlerine uygulanan yöntemler karşılaştırıldığında $\chi^2=4.042$ değeri kritik değer olan 9.488'den düşük olduğundan veri setlerinin MTK temelli ampirik güvenilirlik kestirimlerinin sıra ortalamaları arasında 0.05 alfa düzeyinde manidar bir farklılık bulunmamıştır ($\chi^2=4.042$, $p>0.05$). Yöntemlerin sıra ortalamalarına bakıldığında tam veri grubunun sıra ortalamasının 43.17 olduğu ve ona en yakın değerlerin 42.08 ve 41.31 ile çoklu atama ve regresyon yöntemlerinde olduğu belirlenmiştir. 29.67 sıra ortalaması ile sıfır atama yönteminin tam veri grubuna en uzak değeri verdiği görülmektedir.

- b. Bu aşamada veriler örneklem büyüklüğü bakımında 200, 500 ve 1000 olmak üzere 3 gruba ayrılmış ve analizlere devam edilmiştir. Yapılan

analizler sonucunda $\chi^2=31.412$ değeri kritik değer olan 5.991'den büyük olduğundan farklı örneklem büyüklüklerine sahip veri setlerinin MTK temelli ampirik güvenilirlik kestirimlerinin sıra ortalamaları arasında 0.05 alfa düzeyinde manidar bir farklılık bulunmuştur ($\chi^2=31.412$, $p<0.05$). Hangi gruplar arasında farklılık olduğunu bulmak amacıyla yapılan ikili karşılaştırmalar sonucunda 200 ve 500, 500 ve 1000 kişilik örnekleme sahip veri setleri arasında 0.05 alfa düzeyinde manidar farklılık olduğu görülmüştür. Etki büyüklüğünün 0.415 olması aradaki farkın güçlü olduğunu göstermektedir. Sıra ortalamalarına bakıldığında 200, 500 ve 1000 kişilik veri setlerinin sıra ortalamalarının sırasıyla 54.14, 19.78 ve 40.08 olduğu görülmektedir. Buradan 200 kişilik veri setlerinin, diğer veri setlerine göre daha güvenilir sonuçlar verdiği ortaya çıkmaktadır.

Veri setlerine uygulanan yöntemler karşılaştırıldığında $\chi^2=4.042$ değeri kritik değer olan 9.488'den düşük olduğundan veri setlerinin MTK temelli ampirik güvenilirlik kestirimlerinin sıra ortalamaları arasında 0.05 alfa düzeyinde manidar bir farklılık bulunmamıştır ($\chi^2=4.042$, $p>0.05$). Yöntemlerin sıra ortalamalarına bakıldığında tam veri grubunun sıra ortalamasının 43.17 olduğu ve ona en yakın değerlerin 42.08 ve 41.31 ile çoklu atama ve regresyon yöntemlerinde olduğu belirlenmiştir. 29.67 sıra ortalaması ile sıfır atama yönteminin tam veri grubuna en uzak değeri verdiği görülmektedir.

- c. Bu aşamada veriler kayıp veri oranları bakımında %5, %10 ve %20 olmak üzere 3 gruba ayrılmış ve analizlere devam edilmiştir. Yapılan analizler sonucunda $\chi^2=1.446$ değeri kritik değer olan 5.991'den düşük olduğundan farklı örneklem büyüklüklerine sahip veri setlerinin MTK temelli ampirik güvenilirlik kestirimlerinin sıra ortalamaları arasında 0.05 alfa düzeyinde manidar bir farklılık bulunmamıştır ($\chi^2=1.446$, $p>0.05$). Sıra ortalamalarına bakıldığında %5, %10 ve %20 oranlarında kayıp içeren veri setlerinin sıra ortalamalarının sırasıyla 47.43, 44.65 ve 39.46 olduğu görülmektedir. Buradan %5 oranında kayıp veri içeren veri setlerinin, diğer veri setlerine göre daha güvenilir sonuçlar verdiği ortaya çıkmaktadır.

Veri setlerine uygulanan yöntemler karşılaştırıldığında $\chi^2=4.529$ değeri kritik değer olan 9.488'den düşük olduğundan veri setlerinin MTK temelli ampirik güvenilirlik kestirimlerinin sıra ortalamaları arasında 0.05 alfa düzeyinde manidar bir farklılık bulunmamıştır ($\chi^2=4.529$, $p>0.05$). Yöntemlerin sıra ortalamalarına bakıldığında eksiksiz veri setinin sıra ortalamasının 49.17 olduğu ve buna en yakın değerlerin 47.81 ve 47.03 ile çoklu atama ve regresyon yöntemlerinde olduğu görülmektedir. 33.67 ile en farklı sonucun sıfır atama yönteminde olduğu da çıkarımlar arasındadır.

Bölüm 5

Sonuç, Tartışma ve Öneriler

Bu bölümde araştırmanın bulgularına dayanarak elde edilen sonuçlara, sonuçlara yönelik tartışma ve önerilere yer verilmiştir.

Sonuç

Bu çalışmada farklı koşullar altında kayıp veri yöntemlerinin güvenilirlik üzerine etkisi incelenmiştir. Bu amaca yönelik yapılan analizlerin sonuçları aşağıdaki gibidir.

1. Farklı test uzunlukları dikkate alınarak yapılan karşılaştırmada 40 soruluk veri setlerinin Cronbach alfa ve MTK temelli ampirik güvenilirlik katsayılarının 20 soruluk veri setine göre daha iyi sonuçlar verdiği gözlenmiştir.
2. Farklı test uzunluklarına sahip veri setlerine uygulanan kayıp veri tamamlama yöntemleri karşılaştırıldığında yöntemler arasında anlamlı bir farklılık bulunmadığı ancak eksiksiz veri setlerine en yakın sonuçların çoklu atama ve regresyon yöntemlerinden elde edildiği sonucuna varılmıştır.
3. Örneklem büyüklükleri dikkate alındığında 500 kişilik veri seti ile 200 ve 1000 kişilik veri setleri arasında anlamlı bir farklılık bulunduğu, 200 ve 1000 kişilik veri setlerinin sıra ortalamalarına bakılarak 500 kişilik veri setine oranla daha güvenilir sonuçlar verdiği ortaya çıkmıştır.
4. Örneklem büyüklükleri farklı olan veri setlerine uygulanan yöntemler karşılaştırıldıklarında aralarında anlamlı bir farklılık bulunmamakla beraber regresyon ve çoklu atama yöntemlerinin eksiksiz veri setlerine en yakın sonuçları verdiği gözlenmiştir.
5. Farklı oranlarda kayıp veri içeren veri setleri arasında %5 kayıp veri oranına sahip veri setlerinin %10 ve %20 oranında kayıp veri içeren veri setlerine kıyasla daha güvenilir sonuçlar verdiği ancak aralarında anlamlı bir farklılık bulunmadığı gözlenmiştir.

6. Farklı oranlarda kayıp veri içeren veri setlerine uygulanan kayıp veri yöntemlerinden eksiksiz veri setine en yakın sonuçları regresyon ve çoklu atama yönteminin verdiği, ancak yöntemler arasında anlamlı bir farklılık bulunmadığı sonucuna varılmıştır.
7. Bu sonuçlar göz önüne alınarak tüm koşullar altında en iyi sonuçları ÇA ve RA yöntemlerinin verdiği, onları BM yöntemini takip ettiği, en uzak sonucu sıfır atama yönteminin verdiği gözlenmiştir.
8. DS yöntemi veri setlerini hatalı kestirimlere neden olacak ölçüde daralttığı için uygun bir yöntem olarak görülmemektedir.

Tartışma

Bu çalışmada farklı test uzunluğundaki ve örneklem büyüklüğündeki veri setlerinden elde edilen farklı oranlarda kayıp veri içeren veri setlerine uygulanan birtakım kayıp veri yöntemleri arasında Cronbach alfa ve MTK temelli ampirik güvenilirlikler açısından anlamlı bir farklılık olup olmadığı incelenerek tartışılmıştır.

Eksiksiz veri setleri ile %5, %10 ve %20 oranlarında eksiltilecek kayıp veri içeren yeni veri setleri oluşturulmuş ve bu veri setleri DS, sıfır atama, BM, ÇA ve RA yöntemleri kullanılarak tamamlanmıştır. DS yöntemi sonrasında özellikle %20 oranında kayıp veri bulunması durumunda örneklem büyüklüklerinde ciddi eksilmeler ortaya çıktığından bu yöntemin kullanılmasının her zaman mümkün olmayacağı görülmüştür.

BM yönteminde verinin iki kategorili olması ve test uzunluğuyla da ilişkili olarak her koşulda atama yapılamadığı görülmüş ve bu da yöntemin sınırlılığı olarak rapor edilmiştir.

Kayıp veri içeren veri setleri farklı yöntemlerle tamamlandıktan sonra Cronbach alfa ve MTK temelli ampirik güvenilirlik katsayıları hesaplanmış ve çeşitli koşullar altında güvenilirlik kestirimleri arasında anlamlı farklılıklar olup olmadığı incelenmiştir.

Test uzunluğu dikkate alındığında 20 ve 40 soruluk veri setlerinin güvenilirlikleri arasında anlamlı bir farklılık bulunduğu ve sıra ortalamaları incelendiğinde 40 soruluk veri setlerinin daha güvenilir sonuçlar verdiği görülmüştür. Ayrıca farklı test uzunluğundaki veri setlerine uygulanan yöntemler arasında anlamlı

bir farklılık olmadığı ve sıra ortalamaları incelendiğinde çoklu atama ve regresyon yöntemlerinin eksiksiz veri setlerine en yakın sonuçlar verdiği görülmüştür. Sıfır atama yöntemi eksiksiz veri setine en uzak sonuçları vermiştir.

Örneklem büyüklükleri dikkate alındığında 500 kişilik örneklem büyüklüğüne sahip veri setleri ile 200 ve 1000 kişilik veri setleri arasında anlamlı bir farklılık bulunduğu ortaya çıkmıştır. Ancak bu veri setlerine uygulanan yöntemler arasında anlamlı bir farklılık bulunmaması da varılan sonuçlar arasındadır. Sıra ortalamalarına bakıldığında eksiksiz veri setlerine en yakın sonuçları regresyon ve çoklu atama yöntemlerinden elde edildiği de görülmektedir. Sıfır atama yöntemi eksiksiz veri setine en uzak sonuçları vermiştir.

Kayıp veri yüzdeleri dikkate alınarak bir karşılaştırma yapıldığında yüzde oranları arasında anlamlı bir farklılığın bulunmadığı ancak en güvenilir sonuçların %5 oranında kayıp veri içeren veri setlerinde olduğu gözlenmiştir. Veri setlerine uygulanan kayıp veri yöntemleri arasında anlamlı bir farklılık olmadığı, eksiksiz veri setine en yakın sonuçların çoklu atama ve regresyon yöntemlerinden, en uzak sonucun sıfır atama yönteminden elde edildiği de varılan sonuçlar arasındadır.

Bütün bu bulgular dikkate alındığında kayıp veri yöntemlerinden RA ve ÇA yöntemlerinin eksiksiz veri setlerine en yakın sonuçlar verdiği, bu yöntemleri BM yönteminin takip ettiği görülmüştür. En uzak sonucu sıfır atama yönteminin verdiği de görülmektedir. KTK'da sıkça kullanılan Cronbach alfa katsayısı ve MTK temelli ampirik güvenilirlik kestirimlerinin de birbirlerini destekler yönde sonuçlar verdiği görülmüştür.

Ambley, Omar ve Royston (2007) yaptıkları çalışmada örneklem büyüklükleri ve kayıp veri oranları değişiklik göstermesine rağmen farklı kayıp veri yöntemlerinin arasından çoklu atama yönteminin en iyi sonuçlar verdiğini belirtmiştir. Misztal (2012) yaptığı çalışmada liste bazında silme yönteminin en uzak, çoklu atama yöntemininse en iyi sonuçlar verdiği bilgisine ulaşmıştır. Demir (2013) iki kategorili maddelerden testlerin psikometrik özelliklerini karşılaştırdığı çalışmasında dizin silme ve sıfır atama yöntemlerinin bu veri türüne uygun olmadığı, en iyi sonuçların en çok olabilirlik ve çoklu atama yöntemlerinden elde edildiği sonucuna ulaşmıştır. Zhu (2014) çalışmasında farklı koşullar altında yapılan karşılaştırmalarda TROK varsayımı altında en iyi sonucun liste bazında silme ve çoklu atama, ROK

varsayımında çoklu atama yöntemlerinden elde edildiği bilgisine ulaşmıştır. Tavşancıl ve Akbaş (2014), yürüttüğü çalışmanın sonucunda DS yönteminin önemli sorunlara yol açabileceği, BM ve ÇA yöntemlerinin daha yüksek performans göstereceği sonucuna varmıştır. Kürşad ve Nartgün (2015) yaptığı çalışmasının sonucunda en iyi sonuçların BM, ÇA ve RA yöntemlerinin verdiğini, en az benzerlik gösteren sonuçların DS yöntemine ait olduğunu söylemiştir. Işıkoğlu (2017) yaptığı çalışmada örneklem büyüklüğü ve kayıp veri oranına göre değişiklik göstermesine rağmen kayıp veri yöntemlerinden BM, ÇA ve RA yöntemlerinin iyi sonuçlar verdiğini söylemiştir. Tamcı (2018), çalışmasının sonucunda ÇA yönteminin daha iyi çalıştığını belirtmiştir. Görüldüğü üzere kayıp veri ile ilgili birtakım çalışmalarda benzer sonuçlar elde edilmiştir.

Yurdugül (2006) yaptığı çalışmada KTK'da yer alan parametrelerin MTK kullanılarak kestirilebileceğini belirtmiş ve analizler sonucunda güvenilirlik kestirimlerinin yaklaşık eşit değerler verdiğini belirtmiştir. Akyıldız ve Şahin (2017) yapılan çalışmada KTK ve MTK'ya dayalı kestirimlerin benzer olduğunu söylemiştir. Yukarıdaki çalışmalar, bu çalışma sonucunda ulaşılan Cronbach alfa katsayısı ve MTK temelli ampirik güvenilirlik kestirimlerinin de birbirlerine benzer sonuçlar vermesini desteklemektedir.

Öneriler

Bu bölüm araştırmanın sonuçlarına yönelik ve araştırmacılara yönelik öneriler olmak üzere iki başlıkta sunulacaktır.

Araştırmanın sonuçlarına yönelik öneriler. Bu araştırmadan elde edilen bulgulara dayanarak,

1. Test maddelerinin sayısının fazla olması daha güvenilir sonuçlara yol açmaktadır. Bu nedenle test uzunluğunun fazla olduğu çalışmalar tercih edilebilir.
2. 200 kişilik veri setlerinde 500 ve 1000 kişilik veri setlerine kıyasla daha güvenilir sonuçlar elde edilmiştir. Araştırmacıların bu örneklem büyüklüklerini dikkate alarak çalışmalarını önerilebilir.

3. Farklı test uzunluklarında, farklı örneklem büyüklüklerinde ve farklı oranlarda kayıp veri içeren veri setlerinde ÇA ve RA yöntemleri ile tamamlamaların yapılması tercih edilebilir.
4. DS ve sıfır atama yöntemlerinin iki kategorili puanlanan maddelerin yer aldığı veri setlerinde kayıp verileri tamamlamak için uygun yöntemler olmadığı söylenebilir.
5. Test uzunluğunun ve kayıp veri oranının fazla olduğu durumlarda BM yönteminin kullanılmaması önerilebilir.

Araştırmacılara yönelik öneriler.

1. Bu çalışmada simülasyon veri kullanılmıştır. Bundan sonraki çalışmalar gerçek veriler üzerinden yürütülebilir.
2. Bu çalışma iki kategorili veriler ile yürütülmüştür. Araştırmacılar tarafından çoklu kategorili verilerle çalışılabilir.
3. Bu çalışmada 20 ve 40 soruluk veri setleri kullanılmış ve bu veri setlerine kayıp veri atama yöntemleri uygulanmıştır. Bundan sonraki araştırmalarda farklı test uzunlukları üzerinde çalışılabilir.
4. Bu çalışma %5, %10 ve %20 oranlarında kayıp veri içeren veri setleri üzerinde yürütülmüştür. Araştırmacılar tarafından farklı oranlarda yöntemlerin ne gibi sonuçlar vereceği incelenebilir.
5. Bu çalışmada TROK mekanizması ele alınmıştır. Bundan sonraki çalışmalarda diğer kayıp veri mekanizmaları üzerinde çalışılabilir.
6. Bu çalışmada kayıp veri atama yöntemlerinden DS, ÇA, sıfır atama, BM ve RA yöntemleri incelenmiştir. Araştırmacılar tarafından farklı yöntemler araştırmalara dahil edilebilir.

Kaynaklar

- Acock, A. C. (2005). Working with missing values. *Journal of Marriage and family*, 67(4), 1012-1028.
- Aiken, L. R. (2009). *Psychological testing and assessment*. Pearson Education India.
- Akyıldız, M., & Şahin, M. D. (2017). Açıköğretimde kullanılan sınavlardan Klasik Test Kuramına ve Madde Tepki Kuramına göre elde edilen yetenek ölçülerinin karşılaştırılması. *Açıköğretim Uygulamaları ve Araştırmaları Dergisi*, 3(4), 141-159.
- Allison, P. D. (2001). *Missing data* (Vol. 136). Sage publications.
- Allison, P. D. (2003). Missing data techniques for structural equation modeling. *Journal of abnormal psychology*, 112(4), 545.
- Allison, P. D. (2009). Missing data . (*Sage University Paper Series on Quantitative Applications in the Social Sciences*, 72-89). London: Sage Publication.
- 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.
- Akin Arıkan, C., & Soysal, S. (2018). Investigation of Reliability Coefficients According to Missing Data Imputation Methods. *Hacettepe üniversitesi eğitim fakültesi dergisi-Hacettepe university journal of education*, 33(2), 316-336.
- Bal, C. (2003). *Çok gruplu veri setlerinde eksik gözlem sorununun çözümlenmesi ve sağlık alanında bir uygulama* (Yayımlanmamış Doktora Tezi). Osmangazi Üniversitesi, Sağlık Bilimleri Enstitüsü, Eskişehir.
- Baygül, A. (2007). *Kayıp veri analizinde sıklıkla kullanılan etkin yöntemlerin değerlendirilmesi* (Yayımlanmamış Doktora Tezi). Osmangazi Üniversitesi, Sağlık Bilimleri Enstitüsü, Eskişehir.
- Bayhan, A. (2018). *Farklı Koşullardaki Kayıp Veri Oranının İç Tutarlığa Etkisi* (Yüksek Lisans Tezi). Erişim adresi: <https://tez.yok.gov.tr/UlusalTezMerkezi/giris.jsp>

- Baykul, Y. (2015). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması*. Pegem Akademi.
- Brown, M. L., & Kros, J. F. (2003). Data mining and the impact of missing data. *Industrial Management & Data Systems*.
- Büyükkıdık, S., & Atar, H. Y. (2018). Çok Kategorili Madde Tepki Kuramı Modellerinin Karşılaştırılması. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, 38(2), 663-692.
- Büyüköztürk, Ş. (2002). *Sosyal bilimler için veri ve analizi el kitabı: istatistik, araştırma deseni, SPSS uygulamaları ve yorum*. Pegem A Yayıncılık.
- Cheema, J. (2012). *Handling missing data in educational research using SPSS* (Doctoral dissertation).
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Holt, Rinehart and Winston, 6277 Sea Harbor Drive, Orlando, FL 32887.
- Culbertson, M. J. (2011). Is it wrong? Handling missing responses in IRT.
- Çakmur, H. (2012). Araştırmalarda Ölçme-Güvenilirlik-Geçerlilik. *TAF Preventive Medicine Bulletin*, 11(3).
- Çokluk, Ö., & 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.
- Demir, E., & Parlak, B. (2012). Türkiye’de eğitim araştırmalarında kayıp veri sorunu. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 3(1), 230-241.
- Demir, E. (2013). Kayıp verilerin varlığında çoktan seçmeli testlerde madde ve test parametrelerinin kestirilmesi: SBS örneği. *Eğitim Bilimleri Araştırmaları Dergisi*, 3(2), 47-68.
- Enders, C., & Bandalos, D. (2001). The relative performance of full information maksimum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling*, 8(3), 430-457.
- Enders, C. K. (2003). Using the expectation maximization algorithm to estimate coefficient alpha for scales with item-level missing data. *Psychological methods*, 8(3), 322.

- 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.
- Enders, C. (2010). *Applied missing data analysis*. USA: The Guilford Press.
- Ercan, İ., & Kan, İ. (2004). Ölçeklerde güvenirlik ve geçerlik.
- Erkuş, A. (2012). *Psikolojide ölçme ve ölçek geliştirme*. Ankara: Pegem Akademi Yayınları.
- Finch, W. H. (2010). Imputation methods for missing categorical questionnaire data: A comparison of approaches. *Journal of Data Science*, 8(3), 361-378.
- Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual review of psychology*, 60, 549-576.
- Gürsakal, N. (2001). Sosyal bilimlerde araştırma yöntemleri. *Bursa: Uludağ Üniversitesi Yayınevi*.
- Hohensinn, C. & Kubinger K. D. (2011). On the impact of missing values on item fit and the model validness of the Rasch model. . (tarih yok). *Psychological Test and Assessment Modeling*, , 53, 380-393.
- Howell, D. (2007). The treatment of missing data. Outhwaite, W., Turner, S.P. (eds.), *The SAGE handbook of social science methodology* . Sage Publications.
- Işıkoğlu, M. A. (2017). *Kayıp veri ile baş etme yöntemlerinin ölçme değişmezliğine etkisi açısından karşılaştırılması* (Yayımlanmamış Yüksek Lisans Tezi). Hacettepe Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Köse, İ., & Öztemur, B. (2014). Kayıp veri ele alma yöntemlerinin t-testi ve ANOVA parametreleri üzerine etkisinin incelenmesi. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*.
- Kürşad, M., & Nartgün, Z. (2015). Kayıp veri sorununun çözümünde kullanılan farklı yöntemlerin ölçeklerin geçerlik ve güvenirliği bağlamında karşılaştırılması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 6(2), 254-267.
- Little Roderick, J. A., & Rubin Donald, B. (1987). Statistical analysis with missing data. *Hoboken, NJ: Wiley*, 65.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Routledge.

- McKnight, P., McKnight, K., Sidani, S. S., & Figueredo, A. (2007). *Missing data: A gentle introduction*. Guilford Press.
- Milanzi, E., Molenberghs, G., Alonso, A., Verbeke, G., & De Boeck, P. (2015). Reliability measures in item response theory: Manifest versus latent correlation functions. *British Journal of Mathematical and Statistical Psychology*, 68(1), 43-64.
- Misztal, M. (2012). Imputation of missing data using R package.
- Molenberghs, G., & Verbeke, G. (2006). *Models for discrete longitudinal data*. Springer Science & Business Media.
- Olinsky, A., Chen, S., & Harlow, L. (2003). The comparative efficacy of imputation methods for missing data in structural equation modeling. *European Journal of Operational Research*, 151(1), 53-79.
- Peng, C., Harwell, m., Liou, S., & Ehman, L. (2006). *Advances in missing data methods and implications for educational research*. *Real data analysis*, 3178.
- Rubin, D. (1987). *Multiple imputation for nonresponse in surveys*. New York: John Wiley & Sons.
- Sıddıkođlu, D., Erdođan, B. D., Gökmen, D., & Kutlay, Ő. (2019). A Study of the Approaches to Treating Missing Data in the Data Collected with the WHODAS 2.0 Scale. *Ankara Universites Tip Fakultesi Mecmuasi= Journal of Ankara University Faculty of Medicine*, 72(1), 1.
- Tabachnick, B., & Fidell, L. (2013). *Using multivariate statistics*. (6th edition). Boston: Pearson.
- Tamcı, P. (2018). *Kayıp Veriyle Başa Çıkma Yöntemlerinin Deđişen Madde Fonksiyonu Üzerindeki Etkisinin İncelenmesi* (Yüksek Lisans Tezi). Erişim adresi: <https://tez.yok.gov.tr/UlusalTezMerkezi/giris.jsp>
- Tavşancıl, E. T. D., & Akbaş, U. Y. *Farklı örneklem büyüklüklerinde ve kayıp veri örüntülerinde ölçeklerin psikometrik özelliklerinin kayıp veri baş etme teknikleri ile incelenmesi* (Doktora Tezi). Erişim adresi: <https://tez.yok.gov.tr/UlusalTezMerkezi/giris.jsp>
- Turgut, F. (1977). Tutumların ölçülmesi. Hacettepe Üniversitesi Teksir Materyal, 7.
- Turgut, M. F. (1993). Eğitimde ölçme ve değerlendirme metotları. M. Fuat Turgut.

- Van Ginkel, J., Sijtsma, K., Van der Ark, L., & Vermunt, J. (2010). Incidence Of Missing Item Scores In Personality Measurement, and Simple Item – Score Imputation. *Methodology: European Journal of Research Methods for the Social Sciences*, 6 (1).
- Yurdugül, H. (2006). Obtaining Classical Reliability Terms from Item Response Theory in Multiple Choice Tests. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 39(2), 27-44.
- Zhu, X. (2014). Comparison of four methods for handing missing data in longitudinal data analysis through a simulation study. *Open Journal of Statistics*, 4(11), 933.

EK-A: Etik Komisyonu Onay Bildirimi



T.C.
HACETTEPE ÜNİVERSİTESİ
Rektörlük



Sayı : 35853172-300
Konu :

EĞİTİM BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜNE

İlgi : 12.03.2019 tarihli ve 51944218-300/00000502214 sayılı yazı.

Enstitünüz Eğitim Bilimleri Ana Bilim Dalı Eğitimde Ölçme ve Değerlendirme Bilim Dalı yüksek lisans öğrencilerinden İzel ÇELİK'in Doç. Dr. Burcu ATAR danışmanlığında yürüttüğü "Kayıp Veri İle Baş Etme Yöntemlerinin Güvenirlilik Üzerine Etkisi Açısından Karşılaştırılması" başlıklı tez çalışması Üniversitemiz Senatosu Etik Komisyonunun 19 Mart 2019 tarihinde yapmış olduğu toplantıda incelenmiş olup, etik açıdan uygun bulunmuştur.

Bilgilerinizi ve gereğini saygılarımla rica ederim.

e-imzalıdır
Prof. Dr. Rahime Meral NOHUTCU
Rektör Yardımcısı

Evrensel elektronik imzalı suretine <https://belgedogrulama.hacettepe.edu.tr> adresinden 0c1d761b-af10-4638-014e-b3780f716fba kodu ile erişebilirsiniz.
Bu belge 5070 sayılı Elektronik İmza Kanunu'na uygun olarak Güvenli Elektronik İmza ile imzalanmıştır.

Hacettepe Üniversitesi Rektörlük 06100 Sıhhiye-Ankara
Telefon:0 (312) 305 3001-3002 Faks:0 (312) 311 9992 E-posta: yazim@hacettepe.edu.tr İnternet
Adresi: www.hacettepe.edu.tr

Değerli Dilek İLPAZ

