



HACETTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Eğitim Bilimleri Ana Bilim Dalı
Eğitimde Ölçme ve Değerlendirme Programı

FARKLI KOŞULLARDAKİ KAYIP VERİ ORANININ İÇ TUTARLIĞA ETKİSİ



Ayşe BAYHAN

Yüksek Lisans Tezi

Ankara, 2018



Liderlik, arařtırma, inovasyon, kaliteli eęitim ve deęiřim ile

Daha ileriye... En İyiyeye...



HACETTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Eğitim Bilimleri Ana Bilim Dalı
Eğitimde Ölçme ve Değerlendirme Programı

FARKLI KOŞULLARDAKİ KAYIP VERİ ORANININ İÇ TUTARLIĞA ETKİSİ

THE EFFECT OF MISSING DATA RATE ON INTERNAL CONSISTENCY WITHIN
DIFFERENT CONDITIONS

Ayşe BAYHAN

Yüksek Lisans Tezi

Ankara, 2018

Kabul ve Onay

Eđitim Bilimleri Enstitüsü M¼d¼rl¼đ¼ne,
AyŒe BAYHAN'nın hazırladıđı "Farklı KoŒullardaki Kayıp Veri Oranının İç Tutarlılıđa Etkisi" baŒlıklı bu çalıŒma j¼rimiz tarafından **Eđitim Bilimleri Ana Bilim Dalı, Eđitimde Ölçme ve Deđerlendirme Bilim Dalında Yüksek Lisans Tezi** olarak kabul edilmiŒtir.

J¼ri BaŒkanı Dr. Ogr. Üyesi H. Deniz
G¼LLEROĐLU

İmza

J¼ri Üyesi (DanıŒman) Prof. Dr. Nuri DOĐAN

İmza

J¼ri Üyesi Dr. Ogr. Üyesi K¼bra ATALAY
KABASAKAL

İmza

Enstit¼ Y¼netim Kurulunun
.../.../... Tarihtli ve
sayılı kararı.

Bu tez Hacettepe Üniversitesi Lisans¼st¼ Eđitim, Öğretim ve Sınav Y¼netmeliđi'nin ilgili maddeleri uyarınca yukarıdaki j¼ri üyeleri tarafından **28 / 06 / 2018** tarihinde uygun gör¼lm¼Œ ve Enstit¼ Y¼netim Kurulunca / / tarihinde kabul edilmiŒtir.

Prof. Dr. Ali Ekber ŒAHİN
Eđitim Bilimleri Enstit¼s¼ M¼d¼r¼

Öz

Kayıp veri sorunu, yapılan çalışmalarda araştırmacıların karşılaştığı önemli bir sorundur. Yaygın kullanılan istatistiksel analiz programları tam veri setleri ile çalıştığından araştırmacılar kayıp verinin varlığında bu sorunu gidermeye yönelik çözümler aramaktadırlar. Her ne kadar alanyazında yer alan çeşitli yöntemler kullanılarak veri atama yoluyla tam veri setleri oluşturulsa da kullanılan ölçeğin psikometrik özelliklerinin etkileneceği bilinen bir gerçektir. Bu çalışmada farklı oranlarda kayıp veri içeren veri setleri için hesaplanan Cronbach α değerindeki değişim örneklem büyüklüğü, test uzunluğu, dağılım biçimi ve puanlama türü değişkenleri açısından incelenmiştir. Çalışmada kullanılan tam veri setleri Wingen 3 programında belirlenen koşullar altında 20'li tekrarlar hâlinde üretilmiş ve bu 20'li tekrarlardan random olarak seçilen veri setlerine faktör analizi uygulanarak verilerin tek boyutlu olduğu belirlenmiştir. Oluşturulan tam veri setlerinin her biri R programında yazılan kodlarla rassal olarak kayıp veri mekanizmasına uygun olarak %5, %10 ve %20 oranında değer silinerek Cronbach α güvenilirlik değerinin hesaplanacağı eksik veri setleri elde edilmiştir. Çalışma, kullanılan veriler benzetim yolu ile üretildiğinden simülatif; belirli koşullar altında güvenilirliğin nasıl değiştiğini kestirme amacı taşıdığından ilişkisel bir çalışmadır. Bulgulara göre farklı puanlama türleri için(kategorik ve likert) örneklem büyüklüğü arttıkça Cronbach α değerindeki değişimin azaldığı görülmüştür. Benzer şekilde madde sayısı arttıkça Cronbach α değerindeki değişim azalmaktadır. Farklı oranlarda kayıp veri içeren veri setleri için hesaplanan Cronbach α değerindeki değişimin dağılım biçimine (normal dağılım, sağa çarpık dağılım, sola çarpık dağılım) bağlı olmadığı ortaya konmuştur. Örneklem büyüklüğü, testin uzunluğu ve dağılım biçimi farklı olsa dahi kayıp veri oranı arttıkça Cronbach α değerindeki değişim de artmaktadır.

Anahtar sözcükler: kayıp veri analizi, güvenilirlik, kayıp veri oranı

Abstract

Missing data is problematical for researchers. As common statistical analysis programs depend on full data sets, researchers need solutions to eliminate missing data problem. Although full data sets are created via data imputation by using several methods in literature, it is known that psychometric features of the scale will be affected. In this study, the change in Cronbach α value for data sets having missing data in different rates, has been analyzed in terms of sample size, test length, distribution manner and scoring method. Full data sets in the study have been created as 20-repetitions under conditions designated in Wingen 3 program, and by applying factor analysis to data sets chosen randomly from 20-repetitions, it has been determined that data are one-dimensional. Values have been deleted randomly in the rates of 5%, 10% and 20% in accordance with missing data mechanism with codes typed in R program, where each full data set is created, and missing data sets, where Cronbach α reliability value will be calculated, have been obtained. This study is simulative as the data are created via simulation method and also relational and aims to determine reliability changes. For different scoring methods, as sample size increases, change in Cronbach α value decreases. As the number of items increases, change in Cronbach α decreases. Even if sample size, test length and distribution manner are different, as the rate of missing data increases, change in Cronbach α value increases.

Key words: missing data analysis, reliability, missing data rate

Teşekkür

Tez yazım sürecimde her türlü sorumu hoşgörüyüyle cevaplayan, ortaya çıkan sorunların çözümünde bana yol gösteren, bilgi ve desteğini esirgemeyen, eleştiri ve önerileriyle çalışmama katkıda bulunan ve sonsuz sabrıyla tezimi bitirmemi sağlayan değerli hocam Sayın Prof. Dr. Nuri DOĞAN'a;

Hayatımın her anında maddi ve manevi desteğini esirgemeyen, anlayış ve sabır gösteren, bugünlere gelmem için emek harcıyıp çaba gösteren annem Şener DURMAZ, babam Zeki DURMAZ ve kardeşim Burak DURMAZ'a;

Bu süreçte bana sabrı ve sevgisiyle destek olan eşim Kemal BAYHAN ve akşamları erken uyuyarak ders çalışmama izin veren, hayatıma anlam katan oğlum Tunç BAYHAN'a;

Tez yazım sürecinde sorunlarımın çözülmesinde yardımcı olan Arş. Gör. Nermin KIBRISLIOĞLU UYSAL ve Arş. Gör. Sümeyra SOYSAL'a;

Bu süreçte desteğini hep yanımda hissettiğim beni yalnız bırakmayan arkadaşlarım Ahmet Korkut ÇOLAK, Aylin KOCAAĞA, Deniz ÜNLÜ ve Nihal SOYYİĞİT'e;

Eğitim hayatım boyunca benimle bilgilerini paylaşan, hayatıma katkı sağlayan bütün öğretmenlerime çok teşekkür ederim.

Ayşe BAYHAN

İçindekiler

Öz.....	ii
Abstract.....	iii
Teşekkür.....	iv
Tablolar Dizini.....	vii
Şekiller Dizini.....	viii
Simgeler ve Kısaltmalar Dizini.....	x
Bölüm 1 Giriş.....	1
Problem Durumu.....	3
Araştırmanın Amacı ve Önemi.....	4
Araştırma Problemi.....	4
Sınırlılıklar.....	5
Bölüm 2 Araştırmanın Kuramsal Temeli ve İlgili Araştırmalar.....	6
Kayıp Veri Çalışmalarının Tarihsel Gelişimi.....	6
Kayıp Veri Mekanizmaları.....	9
Kayıp Veri Sürecinde Rassallığın Sorgulanması.....	10
Kayıp Veri İle Başa Çıkma Yöntemleri.....	10
İlgili Araştırmalar.....	16
Bölüm 3 Yöntem.....	19
Benzetim(Simülasyon) Koşulları.....	19
Verilerin Analizi.....	20
Bölüm 4 Bulgular ve Yorumlar.....	21
Araştırmanın 1. Alt Problemine Yönelik Bulgular.....	21
Araştırmanın 2. Alt Problemine Yönelik Bulgular.....	29
Bölüm 5 Sonuç, Tartışma ve Öneriler.....	38
Likert Tipi Verilere Ait Sonuçlar.....	38
İkili Verilere Ait Sonuçlar.....	38

Kaynaklar	43
EK-A: Likert Tipi Veriler İin Trok Testi Sonuları.....	49
EK-B: İkili Veriler İin Trok Testi Sonuları	52
EK-C: Tretilmiř Veriler iin rnek Faktr Analizi Sonuları	55
EK-: Etik Beyanı.....	56
EK-D: Yksek Lisans Tez alıřması Orijinallik Raporu.....	57
EK-E: Thesis Originality Report.....	58
EK-F: Yayımlama ve Fikrİ Mlkiyet Hakları Beyanı	59



Tablolar Dizini

Tablo 1 <i>Kayıp Veri İle Başa Çıkma Yöntemleri</i>	11
---	----



Şekiller Dizini

Şekil 1. Veri matrisi gösterimi.	2
Şekil 2. Normal dağılımlarda örneklem büyüklüğü iç tutarlık katsayısı ilişkisi(likert)	21
Şekil 3. Sağa çarpık dağılımlarda örneklem büyüklüğü iç tutarlık katsayısı ilişkisi(likert)	22
Şekil 4. Sola çarpık dağılımlarda örneklem büyüklüğü iç tutarlık katsayısı ilişkisi(likert)	22
Şekil 5. 5 maddelik veri setleri için dağılım biçimi iç tutarlık katsayısı ilişkisi(likert)	24
Şekil 6. 10 maddelik veri setleri için dağılım biçimi iç tutarlık katsayısı ilişkisi(likert)	24
Şekil 7. 25 maddelik veri setleri için dağılım biçimi iç tutarlık katsayısı ilişkisi(likert)	25
Şekil 8. 50 maddelik veri setleri için dağılım biçimi iç tutarlık katsayısı ilişkisi(likert)	25
Şekil 9. Normal dağılımlarda test uzunluğu iç tutarlık katsayısı ilişkisi(likert).....	27
Şekil 10. Sağa çarpık dağılımlarda test uzunluğu iç tutarlık katsayısı ilişkisi(likert)	27
Şekil 11. Sola çarpık dağılımlarda test uzunluğu iç tutarlık katsayısı ilişkisi(likert)	28
Şekil 12. Normal dağılımlarda örneklem büyüklüğü iç tutarlık katsayısı ilişkisi(ikili)	30
Şekil 13. Sağa çarpık dağılımlarda örneklem büyüklüğü iç tutarlık katsayısı ilişkisi(ikili).....	30
Şekil 14. Sola çarpık dağılımlarda örneklem büyüklüğü iç tutarlık katsayısı ilişkisi(ikili).....	31
Şekil 15. 5 maddelik veri setleri için dağılım biçimi iç tutarlık katsayısı ilişkisi(ikili)	32
Şekil 16. 10 maddelik veri setleri için dağılım biçimi iç tutarlık katsayısı ilişkisi(ikili)	32
Şekil 17. 25 maddelik veri setleri için dağılım biçimi iç tutarlık katsayısı ilişkisi(ikili)	33
Şekil 18. 50 maddelik veri setleri için dağılım biçimi iç tutarlık katsayısı ilişkisi(ikili)	33

Şekil 19. Normal dağılımlarda test uzunluğu iç tutarlık katsayısı ilişkisi(ikili)	35
Şekil 20. Sağa çarpık dağılımlarda test uzunluğu iç tutarlık katsayısı ilişkisi(ikili).	35
Şekil 21. Sola çarpık dağılımlarda test uzunluğu iç tutarlık katsayısı ilişkisi(ikili)..	36



Simgeler ve Kısaltmalar Dizini

ROK: Rassal Olarak Kayıp

TROK: Tamamıyla Rassal Olarak Kayıp



Bölüm 1

Giriş

Geçmişten günümüze yapılan birçok araştırma için toplanan verilerde eksiklikler bulunur. Veri setlerinde yer alan bu eksiklikler kayıp veriler olarak adlandırılır. Longford (2005) kayıp veriyi, en yalın hâliyle toplanması planlanan veriler ile toplanabilen veriler arasındaki fark olarak tanımlamıştır (s.13). Araştırmacıların öncelikli amacı eksiksiz veri seti elde etmek olsa dahi bu her zaman mümkün olmayabilir. Bir araştırmanın kayıp veri içermesinin birçok sebebi olabilir. Bu sebepler yanıtlayıcıların soruya ilişkin fikre sahip olmaması, cevap verdiği hâlde hatalı kodlama yapması, verilen süreyi etkin değerlendirememesi gibi katılımcılardan kaynaklanırken veri toplama aracının sorunlu olması, araştırmacının veri girişinde hata yapması gibi katılımcılardan bağımsız da ortaya çıkabilmektedir. Özetle, kayıp veri sorunu günümüze kadar gelmiş ve engellenmesi mümkün olmayan bir sorundur.

Her ne sebeple ortaya çıkarsa çıksın kayıp veri içeren veri setinin yapısı bozulacak ve bu veri setinde hesaplanan istatistiksel analiz sonuçları genelleme yapılmasına uygun olmayacaktır; çünkü bu istatistiksel yöntemler eksiksiz veri setleri için tanımlanmıştır. Bununla beraber hâlihazırda kullanılan veri analizi programları da tam veri setleri ile çalışmaktadır.

Standart istatistiksel yöntemler dikdörtgensel veri setlerini analiz etmek için geliştirilmiştir. Genellikle veri matrisinin satırları, gözlemleri; sütunları ise değişkenleri yansıtır. Veri matrisindeki girdiler yaş, gelir gibi sürekli değişkenleri veya eğitim derecesi, ırk, cinsiyet gibi kesikli (kategorik) değişkenleri temsil eden değerlerdir. Ancak bu girdiler her zaman gözlemlenebilir değerlerden oluşmaz. Bazı araştırmalardaki katılımcılar gelirlerini bildirmeyi reddeder, bazı araştırmalarda ise deneysel adımlara bağlı olmayan mekanik aksaklıklar yüzünden değerler gözlenemez. Gözlenemeyen değerler veri setine kayıp veri olarak kodlanır ve böylece kayıp veri içeren veri setleri meydana gelir (Little & Rubin, 1987).

İstatistiksel araştırmalarda veri toplama araçlarıyla elde edilen verileri göstermek için veri matrisleri kullanılır. Daha önce de bahsedildiği gibi bu matriste satırlar, gözlemleri; sütunlar ise değişkenleri gösterir. Veri matrisi n gözlemin bulunduğu satırlar ve m değişkenin bulunduğu sütunlardan oluşan dikdörtgensel bir

yapıdır. Bu yapıda değişkenlere ait gözlemler her bir hücreyi oluşturmaktadır. Bir hücredeki veriler satırlar için i sütunlar için j indisi kullanılarak a_{ij} biçiminde gösterilir. Veri matrisleri en genel haliyle Şekil 1’de gösterilmiştir.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2j} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3j} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & a_{i3} & \dots & a_{ij} & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mj} & \dots & a_{mn} \end{bmatrix}$$

Şekil 1. Veri matrisi gösterimi.

Bu veri matrisinde a_{ij} değerleri eksiksiz yer alıyorsa veya eksik olduğu hâlde belirli yöntemlerle tamamlandıysa A matrisi tam veri seti olarak, a_{ij} değerlerinden bazıları yer almıyorsa diğer bir deyişle bu matris eksik değerlerden oluşuyorsa A matrisi kayıp değer içeren veri seti olarak adlandırılabilir.

Problem Durumu

Herhangi bir sebeple eksik verilerden oluşan kayıp veri seti için hesaplanacak istatistiksel analizler sonuçlar için genelleme yapılmasında etkili olmayacaktır. Bu durumun nedeni kayıp verilerin oluşma sebebinin bilinmemesi olabileceği gibi, araştırmacının kayıp veri içeren durumları araştırmasından çıkararak (silme yöntemi gibi) var olan örnekleme küçültmesi de olabilir. Bir veri seti kayıp veri içerdiğinde güvenilirliğin hesaplanması zorlaşacak hatta güvenilirlik azalacaktır. Veri setindeki kayıp veri oranı, testin uzunluğu, dağılım biçimi, kayıp veri türü, puanlama yöntemi gibi durumlar hesaplanan güvenilirlik değerini etkileyecektir. Son yıllarda araştırmacılar kayıp veri sorunu ile baş edebilmek için kayıp veri atama yöntemlerini kullanmaktadırlar. Bir veri setinin herhangi bir sebeple kayıp veri barındırması ölçme araçlarının psikometrik özelliklerini olumsuz etkileyecektir (Alpar, 2011). Büyük örneklem ile yapılan çalışmaların kayıp veri içermesi kaçınılmazdır. Araştırmacılar kayıp veri içeren veri setlerini araştırmadan çıkarırlarsa örneklem büyüklüğü olumsuz etkilenecek bu durum da yapılan analizlerin yordama gücünü olumsuz etkileyecektir. Bu çalışmada en temel olarak kayıp veri oranının güvenilirlikle ilişkisi incelenmiştir. Bugüne kadar yapılan çalışmalarda kayıp verinin önemli bir sorun olduğu ortaya konulmuş (Demir ve Parlak, 2012), kayıp veri atama yöntemlerinin etkililik düzeylerinin karşılaştırılmış (Akbaş ve Tavşancıl, 2015; Aslan, 2010; Baygöl, 2007; Demir, 2013; Sarı, 2012; Sezgin ve Çelik, 2013; Şahin Kürşad, 2014; Öztumur, 2014) ancak araştırma sonuçlarının en önemli özelliklerinden biri olan güvenilirliğe olan etkisine pek fazla değinilmemiştir. Kayıp veri ve güvenilirlik ile ilgili yapılan çalışmalar kayıp veri atama yöntemlerinin güvenilirliğe etkisinin incelendiği çalışmalarla sınırlı kalmıştır (Çokluk ve Kayrı, 2011; Demir, 2013; Şahin Kürşad, 2014).

Ölçme sonuçlarının psikometrik özelliklerinin kayıp veri oranından ne düzeyde etkilendiği ile ilgili çalışma yapılmadığı görüldüğünden bu çalışmada ölçme sonuçlarının önemli özelliklerinden biri olan güvenilirlik ile kayıp veri ilişkisi incelenecektir. Araştırma sonucunda kayıp değer içeren veri setleri için hesaplanan güvenilirlik değerlerinin (Cronbach α) dağılım biçimi, testin uzunluğu, puanlama türü ve örneklem büyüklüğü değişkenlerinden nasıl etkilendiği ortaya çıkarılmıştır.

Araştırmanın Amacı ve Önemi

Başarı testleri, tutum ölçekleri, anketler gibi yaygın olarak kullanılan ölçme araçlarında yer alan sorular herhangi bir sebeple cevaplanmadığında meydana gelen kayıp değerlerin ölçme sonuçlarından biri olan güvenilirliği olumsuz etkileyeceği düşünülmektedir. Bu araştırma sonucunda kayıp veri oranı ile güvenilirliğin ilişkisi ortaya çıkarılacak; kayıp veri ile dağılım biçimi, puanlama türü, örneklem büyüklüğü ve test uzunluğu gibi değişkenleri etkileşimleri de ortaya konarak hangi durumlarda ne tür sonuçlar elde edilebileceğine ilişkin araştırmacılara deneysel bir bulgu sağlanacaktır.

Ölçek geliştirenler, ölçme sonucunu kullanan uygulayıcılar elde ettikleri bir güvenilirlik değerinin kayıp verilerden ve kayıp verilerle diğer değişkenlerin etkileşiminden ne düzeyde etkilendiğini tahmin edebilecekler ve buna göre gerekli önlemleri alma yoluna gideceklerdir. Elde edilen güvenilirlik değerini etkileyen faktörleri bilmek, bu etkiyi ortadan kaldırmak için alınması gereken önlemleri belirlemeyi kolaylaştıracaktır.

Araştırma Problemi

Kayıp veri oranı ve kayıp verilerin puanlama tipi, örneklem büyüklüğü, test uzunluğu ve dağılım biçimi değişkenleri ile etkileşiminin ölçme sonuçlarının güvenilirlik değerine etkisi nasıldır?

Alt problemler.

1. Farklı oranlarda kayıp veri içeren Likert tipi veri seti için hesaplanan iç tutarlık değerleri nasıl değişmektedir?
 - 1.1. Farklı oranlarda kayıp veri içeren veri setlerinde örneklem büyüklüğüne göre iç tutarlık katsayısı nasıl değişmektedir?
 - 1.2. Farklı oranlarda kayıp veri içeren veri setlerinde dağılım biçimine göre iç tutarlık katsayısı nasıl değişmektedir?
 - 1.3. Farklı oranlarda kayıp veri içeren veri setlerinde testin uzunluğuna göre iç tutarlık katsayısı nasıl değişmektedir?

2. Farklı oranlarda kayıp veri içeren ikili puanlanan veri seti için hesaplanan iç tutarlık değerleri nasıl değişmektedir?
 - 2.1. Farklı oranlarda kayıp veri içeren veri setlerinde örneklem büyüklüğüne göre iç tutarlık katsayısı nasıl değişmektedir?
 - 2.2. Farklı oranlarda kayıp veri içeren veri setlerinde dağılım biçimine göre iç tutarlık katsayısı nasıl değişmektedir?
 - 2.3. Farklı oranlarda kayıp veri içeren veri setlerinde testin uzunluğuna göre iç tutarlık katsayısı nasıl değişmektedir?

Sınırlılıklar

- Bu araştırma, araştırmada kullanılan benzetim koşulları ile sınırlıdır.
- İkili puanlama ve beşli Likert tipi puanlama ile sınırlıdır.
- İç tutarlık katsayısı olarak Cronbach α hesaplanması ile sınırlıdır.

Bölüm 2

Araştırmanın Kuramsal Temeli ve İlgili Araştırmalar

Kayıp Veri Çalışmalarının Tarihsel Gelişimi

Üzerinde araştırma yapılması gereken konular, bu araştırmayı yapabilecek bireyler ve kullanılan ölçme araçları var oldukça kayıp veriler de var olmaya devam edecektir. Kayıp veri tam veri setlerine yönelik geliştirilen istatistik programların kullanımı için önemli bir sorun olmasına rağmen ortaya çıkması 1970'leri bulmuştur. Bu yıllarda Afifi ve Elashof (1966) "Çok Değişkenli İstatistiklerde Eksik Gözlemler" adlı çalışmasında sadece eksik veri içermeyen gözlemleri kullanmış, Dempster, Laird ve Rubin (1977) kayıp veri sorunu için maksimum benzerlik ilkesine dayanan beklenti maksimizasyonu yöntemini ileri sürmüşlerdir. Rubin (1976) kayıp verilerin analizinde kayıp veri mekanizmalarının rolünü açıklamış ve Little ve Rubin (1987) kayıp veri mekanizmalarını tamamıyla rassal olarak kayıp, rassal olarak kayıp ve ihmal edilemez kayıp olmak üzere üç kategoride sınıflandırmıştır. Bu sınıflama ile kayıp verilerin yapısının belirlenip bu yapıya uygun kayıp veri mekanizmalarının kullanılacağı öngörülmüştür. Kayıp veri içeren veri setleriyle analiz yapabilmek için silme ve atama yöntemleri kullanılmaya başlandıkça bu yöntemlerin etkililiği hakkında yapılan çalışmalar da hız kazanmıştır. Cool (2000) liste bazında silme ve çift yönlü silme yöntemlerinin, örneklem büyüklüğünü azaltarak istatistiksel analizlerin yordama gücünü olumsuz etkilediğini ve çalışmanın geçerliği hakkında bir tehdit unsuru olduğunu belirtmiştir. Bunun yanı sıra atama yöntemleri kullanıldığında örneklem büyüklüğü değişmediğinden istatistiksel analizlerini yordama gücünün de değişmeyeceğini vurgulayarak atama yöntemlerinin silme yöntemlerine göre daha elverişli olduğunu göstermiştir. Enders ve Bandalos (2001) liste bazında silme ve çift yönlü silme yöntemleri ile benzer tepki örüntüsüne göre atama yöntemlerini doğrulayıcı faktör analizi ile incelemişlerdir.

Ehman, Liou, Harwell ve Peng (2002) eğitim araştırmalarında kayıp verilerin nasıl ele alındığını anlamak için 1998-2002 yılları arasında yayımlanan 11 dergide 1087 nicel araştırmayı incelemişlerdir. Bu araştırmalardan 305(%28)'i araştırmada kayıp verinin mevcut olmadığını belirtmiş, 587(%54)'si araştırmaların kayıp veri içerdiğini göstermiş kalan 195(%18)'i ise kayıp verilerin mevcut olup olmadığıyla ilgili bilgi vermemiştir. Kayıp veri içerdiğini belirten çalışmaların kayıp veri ile baş etmek

için %90'ında liste bazında silme, %7'sinde çiftler bazında silme kullanıldığı görülmüştür. Bu da o dönemde revaçta olan SPSS® veya SAS® gibi istatistiksel paket programları kullananların kayıp verilerle başa çıkmada silme yöntemlerine daha çok güvendiğini açıkça göstermektedir. Kalan %3'lük kısımda ise kayıp verilerle baş etmede diğer yöntemlerin kullanıldığı görülmüştür. Silme yöntemlerine olan güvenin kırılması ise 1999'da yayınlanan APA Task Report ile ortaya çıkan 4 araştırmada daha yeni ve daha temel olan beklenti maksimizasyonu yöntemini kullandığını göstermesiyle olmuştur. Tüm bu sonuçlar o dönemde eğitim araştırmacılarının deneysel çalışmalarda hâlen ilkel yöntemlerle çalıştığını, yeni yöntemleri kullanmadıklarını ve hakemli dergilerin de onları ilkel silme yöntemlerinden uzaklaşmaya teşvik etmediklerini göstermektedir.

Araştırmaların kayıp veri içermesinin bir sorun olduğunun ve bu sorunu çözmek için kullanılan yöntemlerin ortaya çıkmasıyla kayıp veri atama tekniklerinin karşılaştırıldığı çalışmalar hız kazanmıştır. Allison (2001) kayıp veriyle baş etme yöntemlerinin karşılaştırdığı çalışmasında maximum benzerlik ve çoklu atama yöntemlerinin geleneksel yöntemlere göre daha iyi sonuçlar verdiğini açıklamıştır.

Bu alanda çalışmalar yapan Aslan (2010), Aydilek (2013), Demir (2013), Graham (2009), Graham (2012), Nakai ve Ke (2011), Pigott (2001), Sarı (2012), Schlomer, Bauman ve Card (2010), Zhu (2014) farklı istatistiksel yaklaşımları temele aldıkları çalışmalarında silme yöntemleri yerine atama yöntemlerinin kullanılması gerektiğini ve bu yöntemlerin içerisinde de genellikle çoklu atama yönteminin daha kullanışlı olduğunu ortaya koymuşlardır. Benzer şekilde Baraldi ve Enders (2010) modern kayıp veri analizi yöntemlerini inceledikleri çalışmalarında liste bazında silme ve ortalama atama gibi geleneksel yöntemlerin yerine en çok olabilirlik ve çoklu atama gibi modern yöntemlerin daha iyi sonuç verdiklerini ortaya koymuşlardır. Modern kayıp veri analiz yöntemlerini inceledikleri bu çalışma yöntemlerin uygulama adımlarını ve sonuçlarını ortaya koyarak araştırmacılara yol gösterici olmuştur.

Horton ve Kleinman (2007) kayıp veri ile baş etme yöntemlerinin uygulanabileceği çeşitli istatistiksel yazılımların performanslarını karşılaştırırken birtakım eksikleri olsa da kayıp veri atama yöntemlerinin kullanılması gerektiğini vurgulamışlardır. Misztal (2012) ise R paket programında kayıp veri ile baş etme yöntemlerini ele almış ve çoklu atama yönteminin genellikle daha iyi performans

sergilediğini belirtmiştir. Soley Bori (2013) yaptığı çalışmasında kayıp veri mekanizmalarının SAS ve STATA başta olmak üzere diğer yazılımlara nasıl uygulanacağını açıklamıştır.

Görüldüğü üzere kayıp veriler ile ilgili çalışmalar kayıp verilerin tanımlanması, kayıp veri mekanizmalarının karşılaştırılması, bu mekanizmaların istatistik programlarına uygulanması şeklinde süregelmiştir. Ülkemizde ise Oğuzlar (2001) yapılan araştırmalarda kayıp veri sorununa değinerek bu soruna çözüm önerileri geliştirmek amacıyla çalışma yapmıştır. Bal (2003) türetilmiş veri setleriyle yürüttüğü çalışmasında kayıp veri mekanizmalarını karşılaştırmış ve bunun yanında verilerdeki mod, medyan ve standart sapmadaki değişimleri incelemiştir. Pembegül (2009) kayıp veri sorununu ihmal edilebilir veya ihmal edilemez olarak iki farklı durum içerisinde ihmal edilemeyen kayıp veriler için 2x2 boyutunda olumsuzluk tablo çözümlemesi yapmıştır. Kaspar (2011) parametrik olmayan veri atama yöntemi olan eğilim puanı(propensity skor) veri atama yöntemi ile diğer veri atama yöntemlerini karşılaştırdığı çalışmasında eğilim puanı veri atama yönteminin az sayıda kayıp veri olduğu durumlarda iyi sonuç verdiğini belirtmiştir. Alkan (2012) kayıp veri atama yöntemlerini cox regresyon analizi kullanarak karşılaştırdığı çalışmasında örneklem büyüklüğü arttıkça kayıp veri atama yöntemlerinin kullanıldığı veriler ile tam veri seti kullanılarak yapılan analizler arasındaki farkın azaldığını göstermiştir.

Demir ve Parlak (2012) yapılan çalışmalarda kayıp veri sorununun ne düzeyde dikkate alındığını ve sıklıkla hangi yöntemlerin kullanıldığını belirlemek için Türkiye’de, araştırma kapsamında belirledikleri ölçütleri karşılayan dört eğitim dergisinde, 2009-2011 yılları arasında yayımlanan toplam 708 makaleden istatistiksel analiz süreci içerdiğini tespit ettikleri 405 makaleyi incelemiş ve 31 tanesinde kayıp veri sorununun var olduğunu ve bunlardan 7 tanesinde kayıp veri atama yöntemi kullanıldığını tespit etmişlerdir. Bu da Türkiye’de henüz kayıp veri sorununa yeterince önem verilmediğini göstermektedir. Sezgin ve Çelik (2013) kayıp veri sorunu için kullanılan yöntemlerin birbirine karşı üstünlük sağlayamayacağını her yöntemin araştırmada kullanılacak verinin yapısına göre farklı oranlarda başarı göstereceğini ileri sürerek araştırmacının kendine en uygun yöntemi seçmesi gerektiğini vurgulamışlardır. Benzer şekilde Koçak ve Çokluk Bökeoğlu (2017) kayıp veriler baş etme yöntemlerinin model veri ve madde model uyumuna etkisini araştırdıkları çalışmalarında tüm kayıp veri mekanizmaları için

kayıp veri oranı arttıkça kayıp veri ile baş etme yöntemlerinin düşük performans gösterdiğini ortaya koymuşlardır.

Kayıp Veri Mekanizmaları

Araştırmacılar araştırma sürecinde kayıp veri sorunu ile karşı karşıya kaldıklarında öncelikle kayıp veri durumunun Little ve Rubin (2002) tarafından tanımlanan kayıp veri mekanizmalarından hangisine ait olduğunu belirlemelidir. Little ve Rubin kayıp veri mekanizmalarını “Tamamıyla Rassal Olarak Kayıp (TROK), Rassal Olarak Kayıp (ROK) ve İhmal Edilemez Kayıp” olmak üzere üç başlık altında toplamıştır.

Tamamıyla rassal olarak kayıp (TROK) (Missing completely at random).

Tamamıyla rassal olarak kayıp durumunda veri setindeki bir değişkenin kayıp değer içermesi, veri setinde bulunan diğer değişkenlerle veya değişkenin kendisi ile ilişkili değildir. Bu mekanizmada kayıp verinin oluşması şans faktörüyle yakından ilgilidir (Sinharay, Stern ve Russel, 2001). X(yaratıcı düşünme becerisi) ve Y(ders başarısı) olmak üzere iki değişkenden oluşan bir veri setinde Y değişkeninin kayıp veri barındırması ne X değişkenine ne de Y değişkenine bağlıdır. Başka bir deyişle ders başarısı değişkenin kayıp veri içermesi yaratıcı düşünme becerisi ile alakalı olmadığı gibi ders başarısıyla da alakalı değildir. Allison (2003), bu mekanizmayı, X daima gözlenen ve Y kayıp veri içeren değişken olmak şartıyla Eşitlik 1’de gösterildiği gibi formülize etmiştir:

$$P(Y \text{ kayıp} / X, Y) = P(Y \text{ kayıp}) \quad (1)$$

Rassal olarak kayıp (ROK) (Missing at random). Rassal olarak kayıp durumunda durumunda veri setindeki bir değişkenin kayıp değer içermesi, veri setindeki kayıp veri içeren değişkenin dışındaki değişkenlerle ilişkilidir. X(cinsiyet) ve Y(kilo) olmak üzere iki değişkenden oluşan bir veri setinde Y değişkeninin kayıp veri içermesi Y değişkeninin kendisine bağlı değilken X değişkenine bağlıdır. Başka bir deyişle kadınlar erkeklere göre kilolarını söylemekte çekimser olduklarından kilo değişkenin kayıp veri içermesi cinsiyetle alakalıdır (Buhi ve diğerleri 2008). Allison (2003), bu mekanizmayı, X daima gözlenen ve Y kayıp veri içeren değişken olmak şartıyla Eşitlik 2’de gösterildiği gibi formülize etmiştir:

$$P(Y \text{ kayıp} / X, Y) = P(Y \text{ kayıp} / X) \quad (2)$$

İhmal edilemez kayıp (Nonignorable missing). İhmal edilemez kayıp durumunda veri setindeki bir değişkenin kayıp değer içermesi, veri setinde kayıp veri içeren değişkenin kendisi ile ilişkilidir. Bir testte yer alan sorunun yanlış sorulmasından dolayı doğru cevaba ulaşamadığı durumlarda oluşan kayıp veri değişkenin kendisi ile alakalıdır (Sezgin ve Çelik 2013). Allison (2003), bu mekanizmayı, X daima gözlenen ve Y kayıp veri içeren değişken olmak şartıyla Eşitlik 3'te gösterildiği gibi formülize etmiştir:

$$P(Y \text{ kayıp} / Y, X) = P(Y \text{ kayıp} / Y) \quad (3)$$

Kayıp Veri Sürecinde Rassallığın Sorgulanması

Tanımlanan kayıp veri mekanizmalarına uygun olarak kayıp veri analiz yöntemini seçmeden önce kayıp verilerin rastgele dağılıp dağılmadığının sorgulanması gerekir. Bu sorgulamayı yapmak için kullanılacak 3 yöntem vardır. Bunlardan ilki, veri setindeki değişkene ait gözlemlerin kayıp veri içerenler ve kayıp veri içermeyenler olarak iki gruba ayrılarak ilgilenilen değişken açısından bu iki grup arasında anlamlı bir fark olup olmadığının t testi kullanılarak belirlenmesidir. Çıkan farkın anlamlı olması kayıp veri sürecinde rassal dağılımın olmadığını gösterir. İkincisi, veri setindeki değişkenlerin kayıp değer içerenler ve kayıp değer içermeyenler olmak üzere iki gruba ayrılıp tam verilerin 1, kayıp verilerin 0 olarak kodlanması ve bu değişkenler arasındaki Pearson korelasyon katsayısı hesaplanmasıdır. Hesaplanan korelasyon katsayıları her bir değişken çifti için kayıp veriler arasındaki ilişki miktarının derecesini belirtir. Küçük korelasyon katsayısı rassallığı işaret eder. Hesaplanan korelasyon değeri değişken çiftlerinin tümünde düşükse veri yapısı TROK, değişken çiftlerinin bazılarında düşükse veri yapısı ROK olarak tanımlanır. Üçüncüsü ise Little'in TROK testi olarak adlandırılan bir ki-kare testidir. Bu test sonucunda anlamlı fark çıkmazsa veri yapısı TROK olarak kabul edilir (Baygül, 2007; Kaspar, 2011; Şahin Kürşad, 2014).

Kayıp Veri İle Başa Çıkma Yöntemleri

Kayıp veri içeren veri setlerini düzenlemek için kullanılan yöntemler aşağıdaki tabloda listelenmiştir.

Tablo 1

Kayıp Veri İle Başa Çıkma Yöntemleri

KAYIP VERİ YÖNTEMLERİ	SİLME YÖNTEMLERİ	Liste Bazında Silme
	ATAMA YÖNTEMLERİ	Çiftler Bazında Silme
		Yaklaşık Değer Atama
		Ortalama Atama
		Hot/Cold Deck Atama
		Regresyon Ataması
		En Çok Olabilirlik
		Beklenti Maksimizasyonu
		Çoklu Atama

Daha önce de bahsedildiği gibi tam veri setleriyle çalışan ve yaygın kullanılan istatistik paket programlarını kullanabilmek için kayıp verileri düzenlemek gereklidir. Bu gerekliliği ortaya çıkaran araştırmacılar kayıp veri ile baş etme yöntemlerini tanımlamışlardır. Bu tanımlamalar aşağıda özetlenmiştir.

Silme yöntemleri. Bu yöntemler kayıp veri içeren verilerin (değişken veya bireylerin) analize dâhil edilmeyerek araştırmadan çıkarılmasına dayanır. Kullanıcılar için pratik bir yöntem olarak görünse de çok fazla kayıp veri içeren veri setlerinin olduğu araştırmalarda kullanılması örneklem büyüklüğünü küçülteceğinden araştırma sonuçları genellemeyi olumsuz etkileyecektir. Bu yöntem kullanıldığında veri setlerinin sayısı azalacağından (örneklem küçüleceğinden) yöntemi kullanan araştırmacılar ellerinde kalan verilerin yapılacak analizler için yeterli ve uygun olmasını göz önünde bulundurmalıdır (Baygül, 2007). Silme yöntemleri Liste Bazında Silme ve Çiftler Bazında Silme olarak ikiye ayrılmıştır.

Liste bazında silme. Kayıp veri mekanizmasının TROK olduğu durumlarda sıklıkla kullanılan bir yöntemdir (Allison, 2009). Bu yöntem herhangi bir değişken için kayıp veri içeren gözlemin veya bireyin çalışmadan çıkarılmasına dayanır (Howell, 2007). Bu yöntemin örneklem kaybının az olduğu durumlarda veya büyük örneklemelerin olduğu durumlarda kullanılması önerilmektedir. Bu durumlarda sonuçlar dikkate değer şekilde etkilenmeyecektir (Acuna ve Rodriguez, 2004). Bu

yöntemin dezavantajı küçük örneklerde yapılan parametre tahminlerinde yanlılığa yol açması (Demir, 2013) ; avantajı ise yapısal eşitlik modelinden log-lineer analizlere kadar birçok analiz yönteminde özel bir hesaplama gerektirmeden kullanılabilmesidir (Allison, 2002).

Çiftler bazında silme. Kayıp veri mekanizmasının TROK olduğu durumlarda kullanılan bir yöntemdir (Allison, 2009). Bu yöntem eldeki tam veriler kullanılarak ortalama, standart sapma gibi dağılımı tanımlayan ve korelasyon, kovaryans gibi dağılımla ilişkili değerlerin hesaplanmasına dayanır (Allison, 2002). Bu yöntemin liste bazında silme yönteminden farkı kayıp veri içeren bireyin veya gözlemin analizden çıkarılması yerine sadece kayıp veri içeren durumun analizden çıkarılmasıdır (Howell, 2007). Yöntemin dezavantajı örneklem büyüklüğüne çok duyarlı olmasıdır. Yetersiz örneklerde hesaplanan kovaryans veya korelasyon değerinin pozitif tanımlı olmaktan çıkarak regresyon hesaplanmasını engellemesi (Baygöl, 2007) ; avantajı ise kayıp veri sürecinde daha fazla verinin kullanılmasını sağlayıp daha az örneklem değişikliğine sebep olarak yapılan tahminlerin daha küçük standart hataya sahip olmasını sağlamasıdır (Şahin Kürşad, 2014).

Yapılan çalışmalar göstermektedir ki liste bazında silme yöntemi değişkenler arasındaki korelasyon katsayısının yüksek olduğu durumlarda etkili iken çiftler bazında silme yöntemi korelasyon katsayısının düşük olduğu durumlarda etkilidir (Baygöl, 2007; Demir, 2013; Öztumur, 2014).

Atama Yöntemleri. Bu yöntemler kayıp veri içeren veri setlerinde kayıp verinin bulunduğu hücreye çeşitli durumlar göz önüne alınarak bir değer atanmasına dayanır. Bu yöntemler örneklem büyüklüğünü etkilememesi, istatistiksel hesaplamalara uygun olması ve parametre tahminlerinin yanlı olmasına engel olması açısından tercih edilmektedir (Kim ve Curry, 1977).

Yaklaşık değer atama. Bu yöntem her bir kayıp veriye belirlenen bir değer atanması ve kayıp veri oluşmamış gibi analizler yapılması ile ilgilidir. Kayıp veri yerine atanacak değerler belirlenirken tüm serinin ortalaması, kayıp verinin üstünde ve altında bulunan tam verilerin ortalaması, kayıp verinin altında ve üstünde bulunan tam verilerin medyanı kullanılabilir (Kalaycı, 2011). Bu yöntemin avantajı örneklem büyüklüğünü küçültmemesi sayılabilir ancak bu yöntemin varyans-kovaryans kestiriminde yanlılık üretmesi büyük bir dezavantajdır (Demir, 2013).

Ortalama atama. Kayıp veri mekanizmasının TROK olduğu durumlarda kullanılır. Bu yöntemde her bir kayıp veriye eldeki verilerin ortalamasının atanarak analizler yapılmasına dayanır. Kayıp veri yerine verilerin ortalamasının atanması, atama yöntemleri içerisinde en pratik olanıdır. Bu yöntemin avantajı veriler rastgele dağılıma sahipken en iyi değeri vermesi iken (Osborne, 2013) dezavantajı ise kayıp veriler yerine ortalamayı atayarak veriler arasındaki değişkenliği azalttığından standart hataların azalmasına, varyansın olduğundan küçük hesaplanmasına, kovaryansın negatif eğilimde çıkmasına sebep olmasıdır (Şahin Kürşad, 2014).

Hot/Cold deck atama. Bu yöntemde kayıp verinin yeri daha önceki gözlemler veya dış kaynaklar kullanılarak belirlenen bir değer ile doldurulur. Bu değerler belirlenirken kayıt altına alınmış güncel veriler kullanılır (Schoier, 2004). Bu yöntemi isimlendirirken kayıp veriler yerine kullanılacak değer belirlendiği kaynak dikkate alınır. Kayıp veri yerine kullanılacak değer verinin ait olduğu gruptan türetiliyorsa Hot deck atama; farklı bir gruptan türetiliyorsa Cold deck atama olarak isimlendirilir (Baygül, 2007). Hot deck atama yönteminin avantajı değişkenlerin ölçüm düzeyini değiştirmemesi (sürekli ve kategorik) ve kavramsal basitlik sağlaması iken dezavantajı ise benzerlik kavramının belirlemede güçlük yaşanmasıdır. Benzerlik kavramını belirleyebilecek bir yazılım oluşturulursa bu güçlük durumu ortadan kalkacaktır (Oğuzlar, 2001). Cold deck ataması ise yerine ortalamayı koyma yöntemine oldukça benzemektedir. Bu iki yöntem arasındaki fark atanacak değer belirlenmesi için seçilen kaynaktır (Öztemur, 2014). Araştırmacı cold deck atama yöntemini uygularken kayıp veriler yerine kullanacağı değer verilerin ortalamasına ait değerden daha geçerli olacağına emin olmalıdır. Bu yöntemde de tıpkı ortalama atama yöntemindeki gibi kayıp veri yerine atanan değer verilerin değişkenliğini azaltacaktır (Alpar, 2011).

Regresyon ataması. Bu yöntem verilerin TROK ve ROK olduğu durumlarda kullanılır. Bu yöntem kayıp verilerin olduğu veri setini bağımlı değişken diğer veri setini bağımsız değişken kabul ederek kayıp veriler için tahmin yürütülmesine dayanır. Bu tahmin yürütülürken kayıp veri içermeyen(tam) veri seti için bir regresyon denklemi yazılarak bu denklemden elde edilen sonuçlara göre kayıp veriler yerine atama yapılır (Baraldi ve Enders, 2010; Dural, 2010; Tabachnick ve Fidell, 2015). Bu yöntem kullanılırken araştırmacıların dikkat etmesi gereken önemli nokta değişkenler arasındaki ilişkinin yeterli seviyede olması ve bağımlı değişkenin

bağımsız değişken tarafından açıklanabiliyor olmasıdır. Eğer bu iki durum sağlanmıyorsa araştırmacılar regresyon ataması yerine ortalama atamayı kullanmalıdır (Bal, 2003). Bu yöntemin avantajı atama yapılacak değişkende yer alan her kayıp veri için birbirinden farklı bağımsız değişken grubu kullanmasıdır. Ayrıca bir değişkende yer alan kayıp veri diğer değişkenlere bağımlı olduğundan her seferinde farklı bir tahmin değeri ortaya çıkacak ve böylece kayıp veri içeren değişkenler için varyans ve kovaryans değeri korunmuş olacaktır (Oğuzlar, 2001). Bir diğer avantajı ise her ne kadar çok fazla hesaplama işlemi gerektirse de SPSS paket programı eklentisiyle araştırmacılara uygulama kolaylığı sağlamasıdır. Bu yöntemin dezavantajı ise kayıp veri yerine kullanılacak değer tahmin edilirken kayıp veri dışındaki veriler kullanıldığından veri setinde hâlihazırda var olan ilişkinin daha da kuvvetlenmesi ve değişkenliğin azalmasıdır (Baygöl, 2007; Brown ve Kros, 2003; Öztemur, 2014). Ek olarak bu değişkenliğin azalması regresyon katsayılarının standart hatalarının daha az hesaplanmasına sebep olacaktır. Bu olumsuzluğu gidermek ve standart hatalardaki negatif yanlılığı azaltmak için SPSS programında her bir kayıp veri için regresyon ataması yaparken bir de hata terimi eklenmelidir (Howell, 2007). Bu hata terimi eklenmezse varyasyon olması gereken değerden düşük çıkacaktır (Baygöl, 2007). Bunların yanı sıra yöntemin diğer bir dezavantajı ise veri tahmini yapılırken herhangi bir sınırlama olmadığından elde edilen tahmini değerlerin belirlenen aralığın dışında kalabilmesidir (Alpar, 2011; Brown ve Kros, 2003).

En çok olabilirlik. Bu yöntem kayıp veri tahmininde kullanılacak veriler seçilirken gözlemlerin olasılığını maksimum yapacak verilerin seçilmesine dayanmaktadır. Bu bir kayıp veri atama yöntemi olmaktan çok parametre kestiriminde olasılık temelli bir yaklaşım olarak nitelendirilmektedir (Demir ve Parlak, 2012). Bu yaklaşım verilerin yapısının ROK olduğu durumlarda, normallik varsayımının sağlandığı durumlarda, yapısal eşitlik modellemesinde kullanılmakta ve iyi bir tahmin değeri oluşturmayı sağlamaktadır (Allison, 2003). Benzer şekilde verilerin yapısı TROK olduğu durumlarda da yanlı olmayan parametre tahminleri oluşturmaktadır (Enders ve Bandalos, 2001).

Beklenti maksimizasyonu. Bu yöntem Demspiter, Laird ve Rubin tarafından 1977 yılında ortaya atılmıştır. BM yöntemi aşamalı ve tekrarlamalı bir yöntem olup B aşamasında kayıp veri için en iyi olasılık kestirimleri belirlenirken M aşamasında

belirlenen bu deęerler atandıęında oluřan standart sapma, korelasyon, ortalama gibi deęerleri iliřkin kestirimler belirlenir. Bu tekrarlamalı bir sũreç olup deęerlerdeki deęiřimin ˆnemsenecek kadar kũçũlmesiyle son bulur (Alpar, 2011). Allison (2009), beklenti maksimizasyonunun uygulama adımlarını řu řekilde aēıklamıřtır:

1. Liste bazında silme ve çiftler bazında silme yöntemleri kullanılarak ortalama, varyans ve kovaryans deęerleri hesaplanır.
2. Her bir kayıp veri mekanizması için ayrı ayrı olacak řekilde regresyon denklemi yazılarak kayıp deęerler için tahmin yapılır. Bu ařamada, bir ˆnceki ařamada hesaplanan ortalama, varyans ve kovaryans deęerleri de kullanılır.
3. Regresyon denklemleri kullanılarak kayıp veri yerine atanacak deęerlerin tahmini yapılır.
4. Gerçek veriler ve atama yapılan veriler kullanılarak ortalama, varyans ve kovaryans deęerleri yeniden hesaplanır.

İkinci ařamaya geri dˆnũlerek parametre kestirimleri arasındaki fark ˆnemsenecek dũzeye gelene kadar tekrarlar devam ettirilir. Bu yöntemin en ˆnemli avantajı kullanımının pratik olması ve birden fazla istatistiksel paket program için kullanılabilmesidir (řahin Kũrřad, 2014). Bir dięer avantajı ise veriler TROK ve ROK yapıda olduęunda hesaplanacak parametre tahminlerini yansız řekilde yapabilmesidir (Enders, 2010; Schafer ve Graham, 2002). Yöntemin dezavantajı ise ok fazla kayıp veri ieren veri setlerinde parametre kestirimleri arasındaki fark ˆnemsenecek kadar azalana dek iřlem yapmanın ok uzun bir sũreç olmasıdır (Bennett, 2001). Bu yöntemin bir dięer dezavantajı da beklenti maksimizasyonu algoritmasının standart hata kestirimi ˆretmemesidir (Demir, 2013).

oklu atama. Bu yöntem Rubin tarafından 1987’de ortaya atılmıřtır. Kayıp veriler için tek bir atama yöntemi ile belirlenen deęerin kullanılması yerine, iki veya daha ok yöntemin birlikte kullanılması sonucu elde edilen deęerin kullanılmasına dayanır. Bˆylelikle oklu atama yöntemi karma bir kestirim deęer elde etmeyi amalar. Bu deęer iki veya daha fazla yöntemle elde edilmiř tahmini deęerlerin ortalamasıdır. (Kaspar, 2011). Bu yöntem $m > 1$ sayıda tam veri seti oluřacak řekilde kayıp veri atamasının yapılması, oluřan m farklı veri setinin analiz edilmesi ve ortaya

çıkan sonuçların bir araya getirilmesi olmak üzere üç adımda gerçekleşir (Schafer ve Graham, 2002). Yöntem kullanıldığında veri dağılımının özelliklerinin yansıtacak şekilde kayıp veri atanması, parametre kestirimlerinin etkililiğini artıracaktır (Demir, 2013). Bu yöntemin bir diğer avantajı da analizde yer alan değişkenlerin normalliği ihlal etmesi durumunda dahi güçlü ve üstün tahminler sunmasıdır (Özdemir, 2014). Her ne kadar bu yöntem geleneksel veri analiz programlarında dahi kullanılabilir olsa da, veri atamasının genellikle seçkisiz biçimde yapılması farklı yazılımlar için farklı sonuçlar ortaya çıkmasına sebep olması bir dezavantajdır (Allison, 2009). Demir (2013) kayıp veriler ile ilgili bilgilerin birbirinde kopuk ve dağınık olması durumunda çoklu atama yöntemini kullanmanın istenilen düzeyde anlamlı sonuç vermeyeceğini ve bu durumda basit atama yöntemlerinin kullanılmasının daha yararlı olacağını belirtmiştir.

İlgili Araştırmalar

Çokluk ve Kayrı (2011) kayıp değerlere yaklaşık değer atama yöntemlerinin güvenilirlik ve geçerliğe etkisini inceledikleri çalışmada tam veri setinde görülen tek faktörlü yapının atama yöntemlerinin kullanıldığı kayıp veri setinde de görüldüğünü belirlemiş ancak atama yöntemlerinin kullanıldığı veri setinde açıklanan varyans oranı, öz değerler ve Cronbach α iç tutarlık katsayısının düştüğünü gözlemlemişlerdir.

Demir (2013) kayıp veri varlığında iki kategorili maddelerden oluşan testlerin psikometrik özelliklerini incelediği çalışmada silmeye dayalı yöntemlerin testin güvenilirliğini artırma, atama yöntemlerinin ise düşürme eğiliminde olduğu, çoklu atama ve en çok olabilirlik yöntemlerinin ise daha ölçülü olduğu sonucuna ulaşmıştır.

Weaver ve Maxwell (2014) SPSS kullanıcıları için kayıp veri içeren veri setlerinde açılımlı faktör analizi ve güvenilirlik hesaplanmalarında kullanabilecekleri yöntem tanımlamışlardır.

Şahin Kürşad (2014) sıklıkla kullanılan kayıp veri atama yöntemlerini betimsel istatistik, geçerlik ve güvenilirlik açısından karşılaştırdığı çalışmada tek faktörlü ve normal dağılıma sahip gerçek verileri kullanmıştır. 9 farklı yöntemin ele alındığı bu çalışmada kayıp veri içeren veri setleri ile tam veri setleri için hesaplanan güvenilirlik katsayılarının az da olsa farklılık gösterdiği sonucuna ulaşılmıştır. Farklı

örneklem büyüklükleri için hesaplanan güvenilirlik değerleri kayıp veri oranı arttıkça tam veri setinden elde edilen değerlerden farklı çıkmıştır. Kayıp veri oranının yüksek olduğu durumlarda atama yöntemlerinin kullanıldığı veri setlerinde hesaplanan güvenilirlik katsayıları tam veri setlerinde hesaplanan güvenilirlik katsayılarından düşük çıkarken kayıp verinin düşük olduğu durumlarda liste bazında silme yöntemi kullanılan veri setlerinde hesaplanan güvenilirlik değerinin tam veri setinde hesaplanan güvenilirlik değerinden yüksek çıktığı görülmüştür.

Akbaş (2014) farklı örneklem büyüklüklerinde ve farklı kayıp veri örüntülerinde ölçeklerin psikometrik özelliklerini kayıp veri ile başa çıkma teknikleri altında incelediği çalışmasında, Cronbach α değerlerinin örneklem büyüklüğü ve madde sayısı açısından birbirine benzer kestirim oluşturduğunu ancak farklı kayıp veri oranları açısından seçkisiz olmayan kayıp şartında negatif yönlü kestirim oluşturduğu sonucuna varmıştır.

Nartgün (2015) kayıp veri ile baş etmede kullanılan beş farklı yöntemin betimsel istatistik, geçerlik ve güvenilirlik açısından karşılaştırdığı çalışmasında, farklı kayıp veri oranları için çoklu atama ve regresyon yöntemlerinin kullanıldığı durumlarda hesaplanan güvenilirlik değerinin, tam veri setinde hesaplanan değerle uyumlu olduğu ancak liste bazında silme yönteminin kullanıldığı çalışmalarda hesaplanan güvenilirlik değerinin tam veri setinde hesaplanan değerden büyük ölçüde farklılık gösterdiğini tespit etmiştir.

Soğuksu ve Alıcı (2016) eşdeğer yarılar güvenirliliğinin farklı homojenlik düzeylerindeki örneklem büyüklüklerinde, test uzunluğuna, yarıya bölme yöntemlerine ve güvenilirlik kestirme tekniklerine göre inceledikleri çalışmalarında test uzunluğu ve örneklem büyüklüğü arttıkça eşdeğer yarılar güvenirliliğinin yüksek, standart hataların ise düşük çıktığı sonucuna ulaşmışlardır.

Bu çalışmada kayıp veri oranı ile güvenilirlik arasındaki ilişki incelenmiştir. Alanyazın incelendiğinde kayıp veri güvenilirlik ilişkisine değinen çalışmaların sayıca az olduğu görülmektedir. Güvenirlik ve kayıp veri ikisi ile ilgili yapılan çalışmalar tam veri setleri ile eksik veri setleri için kayıp veri atama yöntemlerinin karşılaştırılması çerçevesinde ilerlemiştir. Bu çalışmada farklı olarak herhangi bir kayıp veri atama yöntemi kullanılmayacak ve veri setlerindeki kayıp veri oranının örneklem büyüklüğü, testin uzunluğu, dağılım biçimi ve puanlama durumu değişkenleri

açısından güvenilirliđi nasıl etkilediđi tartiřılacaktır. Yapılan çalıřmalarda bu deđiřkenlerin bir arada kullanılmadıđı görölmüřtür.



Bölüm 3

Yöntem

Bu çalışma belirli koşullar altında güvenilirliğin nasıl etkilendiğini belirleme amacı taşıdığından ilişkisel araştırma grubunda, kullanılan veriler benzetim yolu ile üretildiğinden simülatif araştırma grubundadır.

Benzetim(Simülasyon) Koşulları

Madde sayısının belirlenmesi. Aiken (1985), iki kategorili puanlama tekniği kullanılan testlerde madde sayısının 20'den az olması durumunda test ile ölçülmek istenen özelliğin süreklilik gösterme ihtimalinin azalacağını belirtmiştir. Bu bilgi dikkat alınarak madde sayısı hem 20 maddenin altında hem de üstünde olacak şekilde 5,10, 25, 50 olarak belirlenmiştir.

Örneklem büyüklüğünün belirlenmesi. Nunnally ve Bernstein (1994), ölçek çalışmalarında 300 kişilik bir örneklemin yeterli olacağını belirtirken; Comrey ve Lee (1992), 100 kişilik örnekleme zayıf, 200 kişilik örnekleme orta, 300 kişilik örnekleme iyi, 500 kişilik örnekleme çok iyi ve 1000 kişilik örnekleme mükemmel olarak nitelendirmiştir. Bu bilgiler ışığında örneklem büyüklüğü 100, 250, 500, 1000 olarak belirlenmiştir.

Kayıp veri oranının belirlenmesi. Literatür incelendiğinde kayıp veriler ile ilgili yapılan çalışmalarda kayıp veri oranının sıklıkla %5, %10, %15 olarak seçildiği görülmektedir. Bu çalışmada ise kayıp veri oranları %5, %10 ve %20 olarak belirlenmiştir.

Araştırmanın veri setleri Wingen 3 programında belirtilen koşullar dikkate alınarak üretilmiştir. Veri setlerinin üretilmesi sırasında yukarıda belirtilen simülasyon koşullarına dayanarak madde sayısı 5, 10, 25, 50; kişi sayısı 100, 250, 500, 1000; puanlama tekniği 1-0, Likert; dağılım biçimi normal, sağa çarpık, sola çarpık ve kayıp veri oranı %5, %10, %20 olarak belirlenmiştir.

Sonuç olarak tek boyutlu iki kategorili ve Likert, 20 tekrarlı 1920 adet tam veri seti oluşturulmuştur. Oluşturulan veri setlerinin tek boyutluluğu SPSS paket programı kullanılarak yapılan açımlayıcı faktör analizi ile test edilmiştir. Faktör analizi sonucu özdeğer ve açıklanan varyans değerleri incelendiğinde veri setlerinin tek boyutlu olduğu gözlenmiştir. Oluşturulan veri setlerinin her biri için R

programında yazılan kodlarla, tamamen rassal kayıp veri özelliğinde olacak şekilde %5, %10, %20 oranlarında kayıp veri setleri oluşturularak toplamda 5760 veri seti ile çalışılmıştır. Tamamen rastlantısal kayıp veri yapısı için R programında, satır=birey (m) ve sütun=madde (n) çarpımına dayalı ve 0-1 arasında değerler alan bir olasılık matrisi oluşturulmuştur. Bu matriste kayıp veri oranına karşılık gelen olasılık değerlerinin mxn yapısındaki ana veri setinde karşılık gelen değerleri kayıp veri olarak tanımlanmış ve veri kümesinden silinmiştir.

Verilerin Analizi

Bu araştırmada kayıp değerlere sıfır ataması yapılan 5760 veri seti için Excel 2010 paket programı kullanılarak Likert tipi ve ikili veri setleri için Cronbach α değerleri hesaplanmıştır. İkili veri setleri için hesaplanan Cronbach α değerleri KR-20 değerlerine eşit olduğundan yapılan yorumlar KR-20 için de geçerlidir. Her 20'li tekrar için hesaplanan Cronbach α değerlerinin standart sapması hesaplanarak ortalama hataların karekökü elde edilmiştir. Farklı değişkenler için belirlenen problem ve alt problemlere uygun olarak oluşturulan grafiklerde Cronbach α değerlerinin standart sapmaları kullanılmıştır.

Rubin (1987), standart hatanın verilen formülle hesaplanabileceğini Eşitlik 4'te ortaya koymuştur;

$$\sqrt{\frac{1}{M} \sum_k s_k^2 + \left(1 + \frac{1}{M}\right) \left(\frac{1}{M-1}\right) \sum_k (b_k - \bar{b})^2}$$

(4)

M = Örneklem Sayısı

b_k = k. Veri setinden elde edilen parametre tahmini

s_k = k. Veri setinden elde edilen standart hata

Bölüm 4

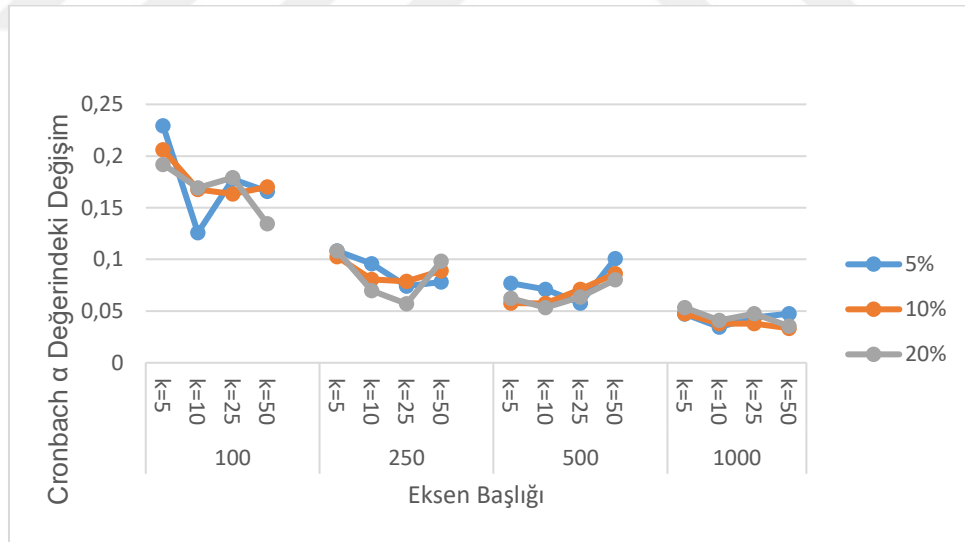
Bulgular ve Yorumlar

Bu bölümde araştırma sonucu elde edilen bulgulara yer verilmiştir. Farklı oranlarda oluşturulmuş kayıp veri setleri için hesaplanan güvenilirlik değerlerinin örneklem büyüklüğü, soru sayısı ve dağılım biçimine göre değişimi raporlanacaktır.

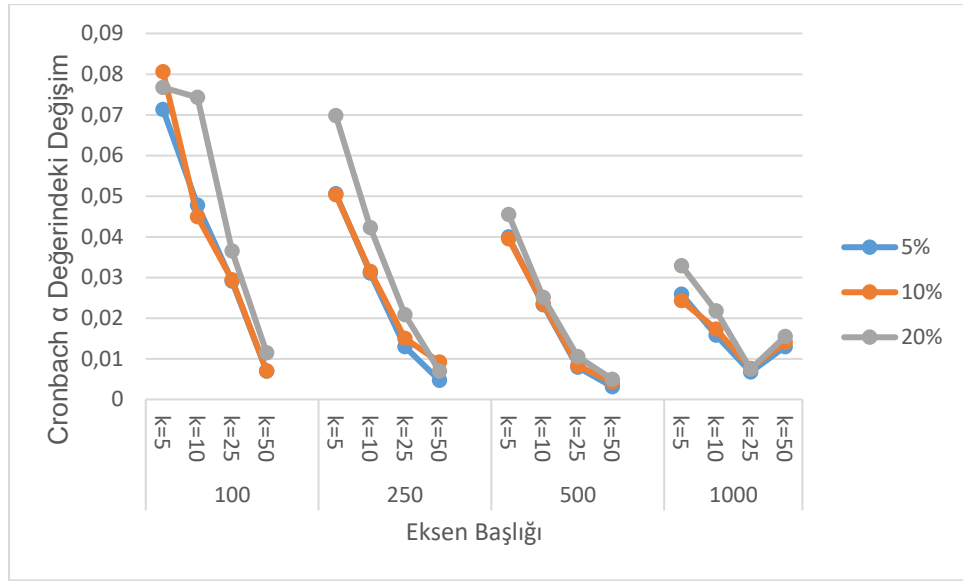
Araştırmanın 1. Alt Problemine Yönelik Bulgular

Bu bölümde problem cümlesine çözüm bulmak amacıyla oluşturulan ilk alt problem olan “Farklı oranlarda kayıp veri içeren Likert tipi veri seti için hesaplanan iç tutarlık değerleri nasıl değişmektedir?” sorusuna cevap aramak için elde edilen bulgular örneklem büyüklüğü, soru sayısı ve dağılım biçimi alt başlıklarında tartışılmıştır.

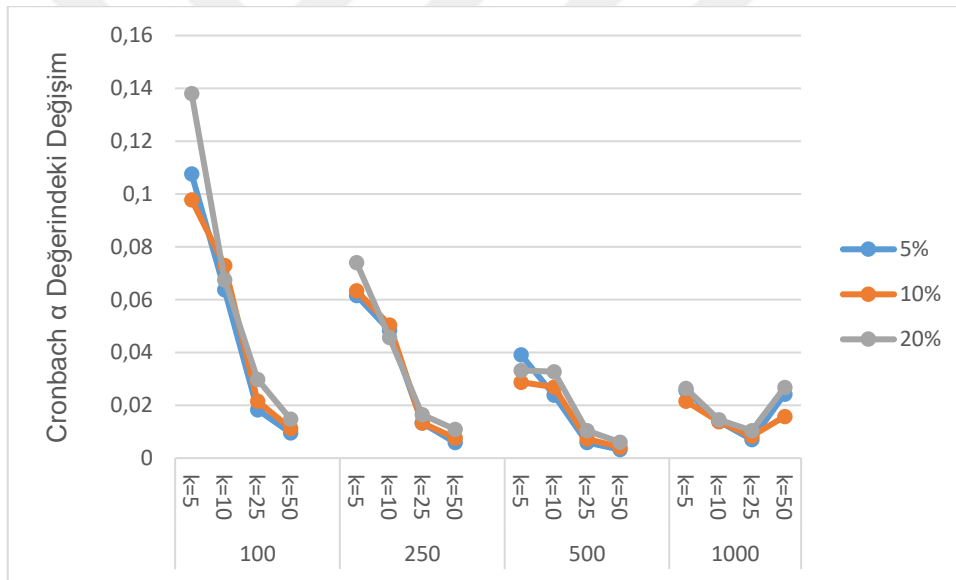
“Farklı oranlarda kayıp veri içeren Likert tipi veri setlerinde örneklem büyüklüğüne göre iç tutarlık katsayısı nasıl değişmektedir?” sorusunu ait bulgular.



Şekil 2. Normal dağılımlarda örneklem büyüklüğü iç tutarlık katsayısı ilişkisi(likert)



Şekil 3. Sağa çarpık dağılımlarda örneklem büyüklüğü iç tutarlık katsayısı ilişkisi(likert)



Şekil 4. Sola çarpık dağılımlarda örneklem büyüklüğü iç tutarlık katsayısı ilişkisi(likert)

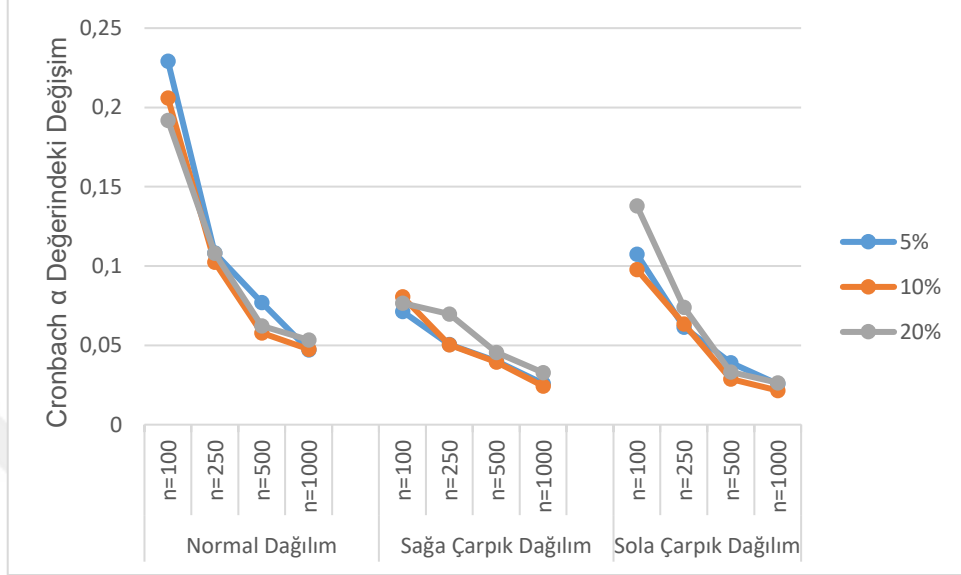
Şekil 2 incelendiğinde normal dağılımlarda örneklem büyüklüğü arttıkça Cronbach α değerinin değişiminin azaldığı görülmektedir. Benzer şekilde madde sayısı arttıkça Cronbach α değerindeki değişimin azalması beklenmektedir. Ancak bu değişim 100 kişilik örnekleme 5 maddelik testten 10 maddelik teste geçişte azalırken 10 maddelik testten 25 maddelik teste geçerken artmakta, 25 maddelik testten 50 maddelik teste geçerken ise azalmaktadır. Madde sayısı dikkate alındığında normal dağılımlı verilerde artış veya azalışın madde sayısına bağlı bir

sistematik izlemediği görülmektedir. 250 kişilik, 500 kişilik ve 1000 kişilik örneklemelerde de bu artış ve azalışlar belirli bir düzene sahip değildir. Ancak büyük örneklemelere gidildikçe Cronbach α değerindeki değişimin tüm kayıp veri oranlarında birbirine yaklaştığı söylenebilir. Diğer yandan kayıp veri oranı arttıkça Cronbach α değerindeki değişim de artmaktadır.

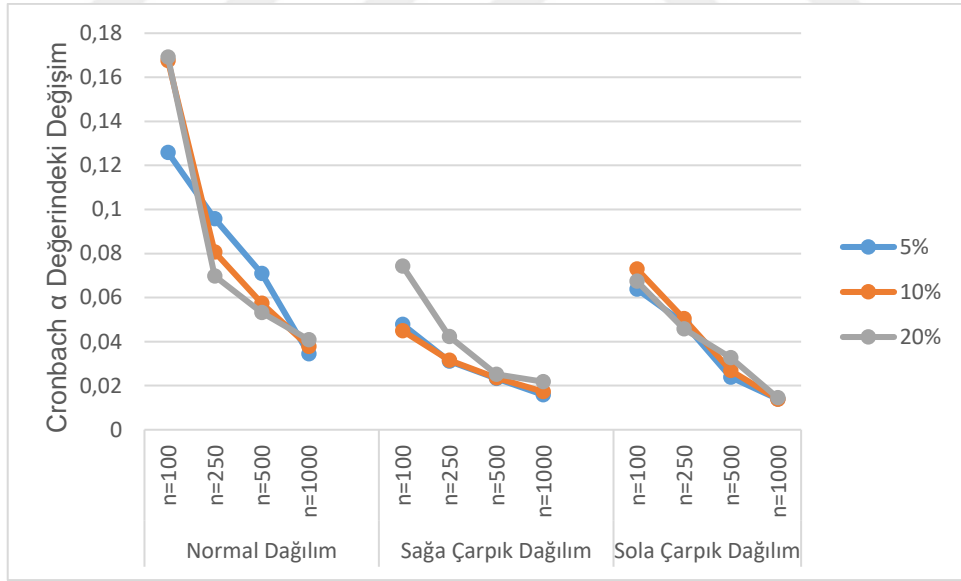
Şekil 3 incelendiğinde sağa çarpık dağılımlarda örneklem büyüklüğü arttıkça Cronbach α değeri için değişimin azaldığı görülmektedir. Benzer şekilde madde sayısı arttıkça Cronbach α değerindeki değişimin azaldığı görülmektedir. Bu azalış sadece 1000 kişilik örneklemde 25 maddelik veri setinden 50 maddelik veri setine geçerken artış şeklinde kendini göstermiştir. Ancak sağa çarpık dağılım için hesaplanan sapma zaten oldukça küçüktür. Dolayısıyla söz konusu değişimler fark yaratan değişimler değildir ya da tesadüfilikle açıklanabilir. Öte yandan kayıp veri oranı arttıkça Cronbach α değerlerindeki değişim de artmaktadır.

Şekil 4 incelendiğinde sola çarpık dağılımlarda örneklem büyüklüğü arttıkça Cronbach α değeri için değişim azalmaktadır. Benzer şekilde madde sayısı arttıkça Cronbach α değerindeki değişimin azalması beklenmektedir. Ancak bu değişim 1000 kişilik örneklemde 25 maddelik testten 50 maddelik teste geçişte bütün kayıp veri oranları için artış olarak kendini göstermektedir. Sola çarpık dağılım için hesaplanan sapma zaten oldukça küçüktür. Dolayısıyla söz konusu değişimler fark yaratan değişimler değildir ya da tesadüfilikle açıklanabilir. Diğer yandan kayıp veri oranı arttıkça standart sapma artmakta, Cronbach α değerindeki değişim de artmaktadır. Mutlak kayıp veri oranına göre sapmalardaki artış hiç de fazla değildir.

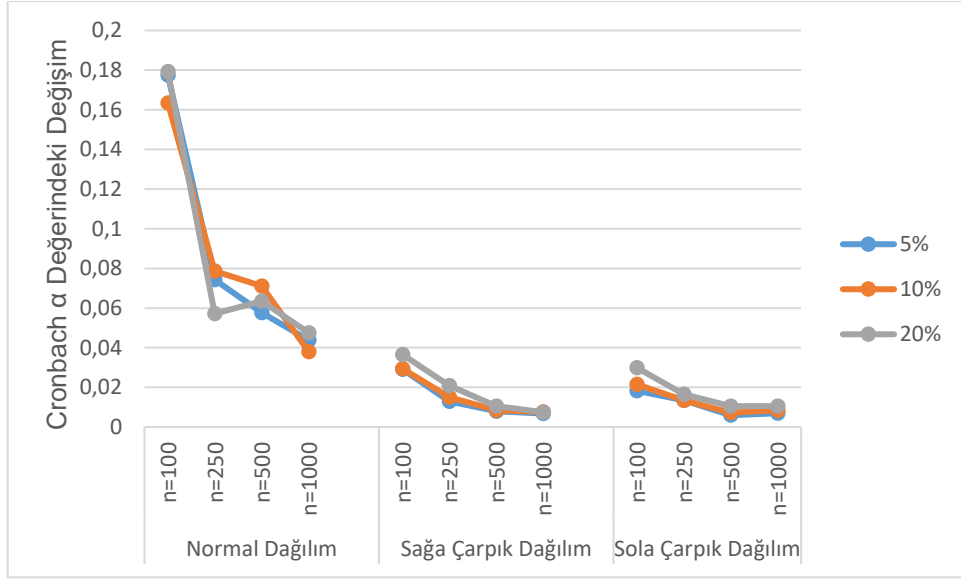
“Farklı oranlarda kayıp veri içeren Likert tipi veri setlerinde dağılım biçimine göre iç tutarlık katsayısı nasıl değişmektedir?” sorusuna ait bulgular.



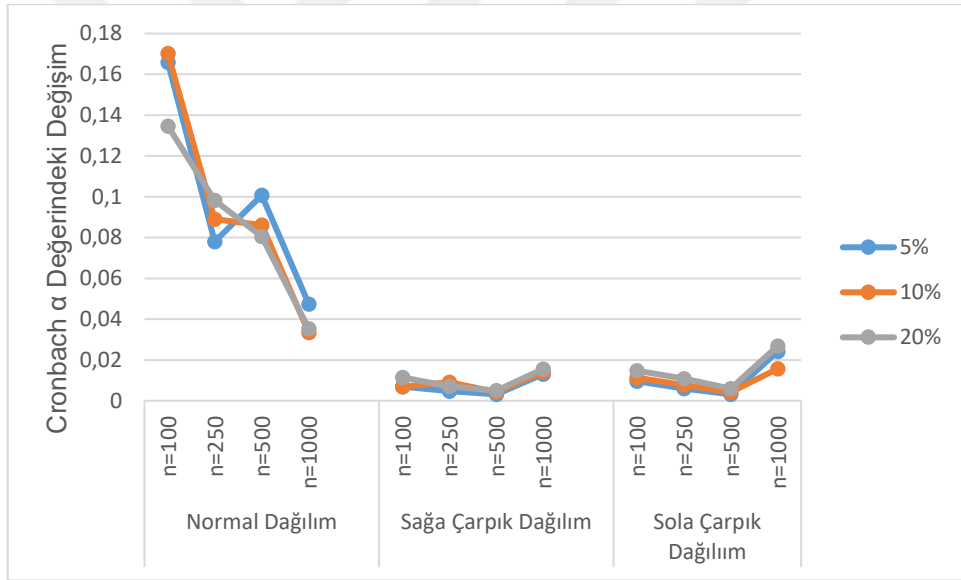
Şekil 5. 5 maddelik veri setleri için dağılım biçimi iç tutarlık katsayısı ilişkisi(likert)



Şekil 6. 10 maddelik veri setleri için dağılım biçimi iç tutarlık katsayısı ilişkisi(likert)



Şekil 7. 25 maddelik veri setleri için dağılım biçimi iç tutarlık katsayısı ilişkisi(likert)



Şekil 8. 50 maddelik veri setleri için dağılım biçimi iç tutarlık katsayısı ilişkisi(likert)

Şekil 5 incelendiğinde hem normal hem de çarpık dağılım biçimlerinde büyük örneklemelere gidildikçe Cronbach α değeri için değişim azalmaktadır. Cronbach Alfa değeri değişimindeki azalış normal dağılım içeren veri setlerinde daha hızlı gerçekleşirken sola çarpık ve sağa çarpık dağılımlarda daha yavaş gerçekleşmektedir. Cronbach α değerinde örneklem büyüklüğüne göre değişimin çarpık dağılımlarda daha az olması dağılımların homojen olmasına (test varyansının düşük olmasına) dayanabilir. Diğer yandan çoğu koşulda kayıp veri oranı arttıkça

standart sapma artmakta, dolayısıyla Cronbach α deęerindeki deęişim de aynı şekilde artmaktadır.

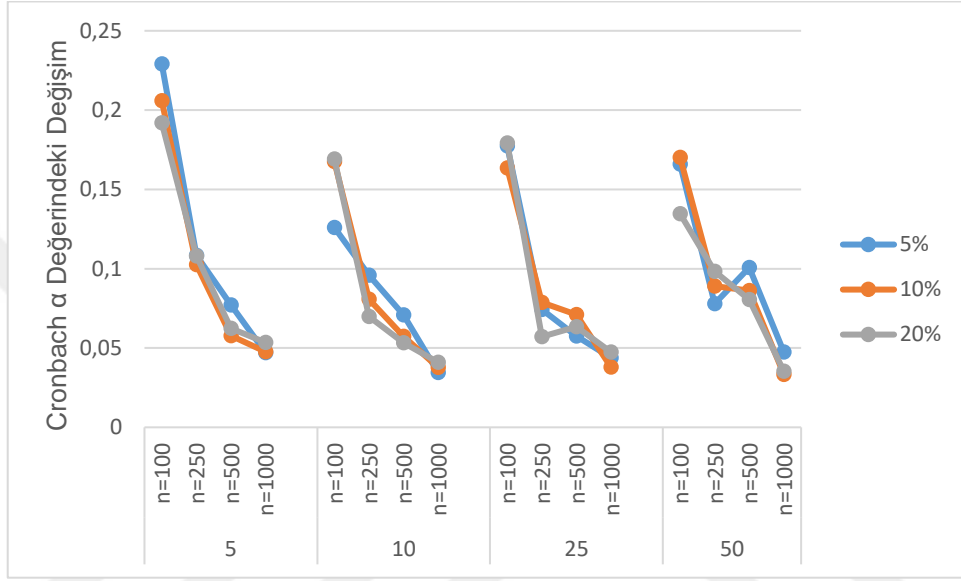
Şekil 6'ya bakıldığında hem normal hem de çarpık dağılımlarda Cronbach α deęerindeki deęişim azalmaktadır. Cronbach α deęeri deęişimindeki azalış 5 maddelik veri setlerinde olduęu gibi normal dağılım içeren veri setlerinde daha hızlı gerçekleşirken sola çarpık ve sağa çarpık dağılımlarda daha yavaş gerçekleşmektedir. Cronbach α deęerinde örneklem büyüklüğüne göre deęişimin çarpık dağılımlarda daha az olması dağılımların homojen olmasına (test varyansının düşük olmasına) dayanabilir. Diğer yandan çoęu koşulda kayıp veri oranı arttıkça standart sapma artmakta, dolayısıyla Cronbach α deęerindeki deęişim de aynı şekilde artmaktadır.

Şekil 7'ye bakıldığında kayıp veri oranı arttıkça hem normal dağılım hem de çarpık dağılımlarda Cronbach α deęerindeki deęişim azalmaktadır. Cronbach Alfa deęeri deęişimindeki azalış 5 ve 10 maddelik veri setlerinde olduęu gibi normal dağılım içeren veri setlerinde daha hızlı gerçekleşirken sola çarpık ve sağa çarpık dağılımlarda daha yavaş gerçekleşmektedir. Diğer yandan sağa çarpık dağılım ve sola çarpık dağılım için Cronbach α deęerindeki deęişim için hesaplanan deęerler birbirine çok yakındır. Benzer şekilde örneklem büyüklüğü arttıkça Cronbach α deęerindeki deęişim yine azalmaktadır. Ayrıca sağa çarpık ve sola çarpık dağılımlarda örneklem büyüklüğü deęişse de Cronbach α sapma deęerlerinin birbirine oldukça yakın olduęu görülmektedir. Diğer yandan 500 örneklem büyüklüğüne ulaşıldıktan sonra kayıp veri oranının etkisi azalmakta Cronbach α daha kararlı hâle gelmektedir.

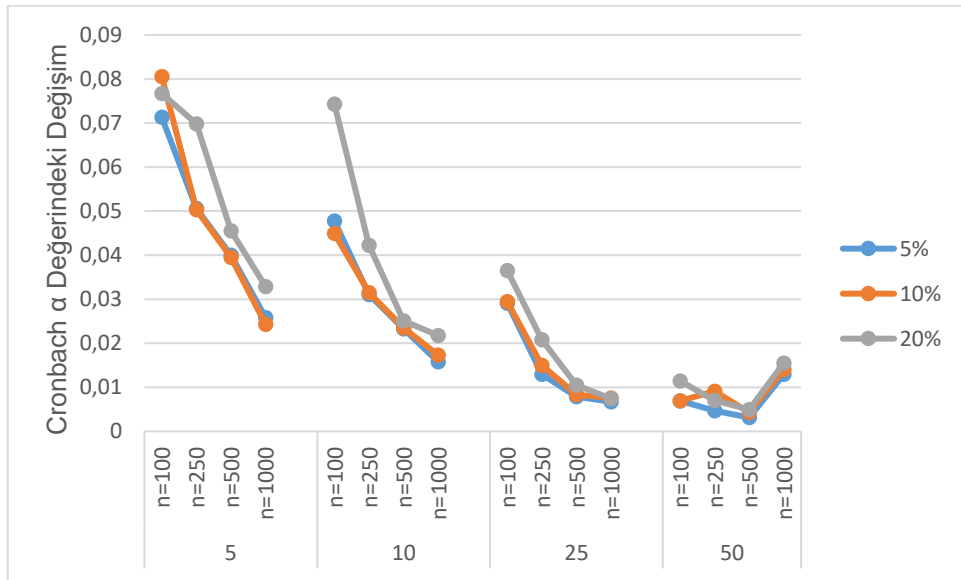
Şekil 8 incelendiğinde normal dağılım biçiminde Cronbach α deęeri için deęişim azalmaktadır. Sağa çarpık ve sola çarpık dağılım biçimlerinde 100 maddelik veri setinden 500 maddelik veri setine kadar Cronbach α deęerlerindeki deęişim azalırken ani bir şekilde 500 maddelik veri setinden 1000 maddelik veri setine geçildiğinde bu deęişim artmaktadır. Bu deęişim tıpkı 5, 10, 25 maddelik veri setlerinde olduęu gibi normal dağılım içeren veri setlerinde daha hızlı ilerlerken sola çarpık ve sağa çarpık dağılımlarda daha yavaş ilerlemektedir. Aynı şekilde örneklem büyüklüğü arttıkça Cronbach α deęerindeki deęişim azalmaktadır. Bu genellemeyi bozan tek durum çarpık dağılımlarda 500 maddelik veri setleri ile 1000 maddelik veri

setleri arasında artış olmaktadır. Diğer yandan kayıp veri oranı arttıkça standart sapma artmakta, Cronbach α değerindeki değişim aynı şekilde artmaktadır.

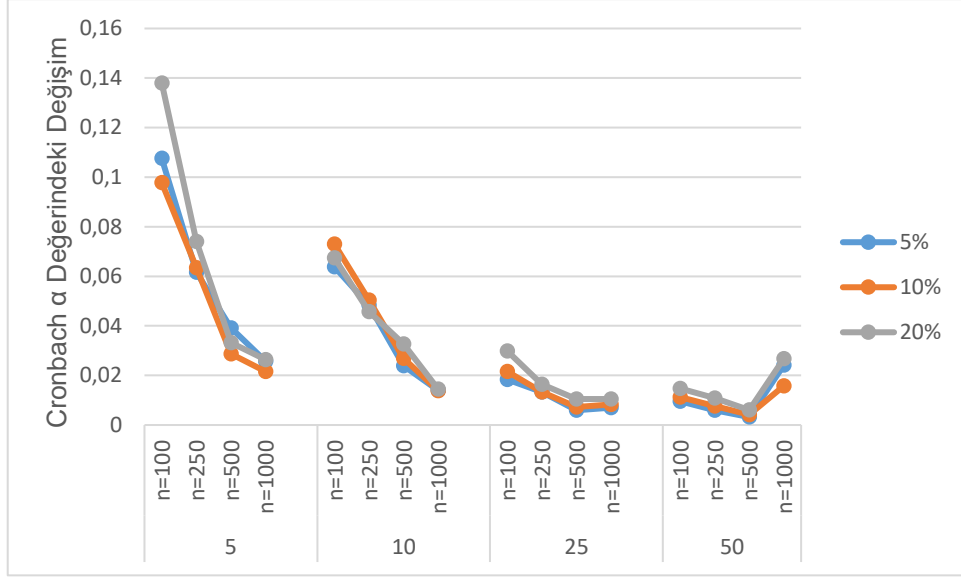
“Farklı oranlarda kayıp veri içeren Likert tipi veri setlerinde testin uzunluğuna göre iç tutarlık katsayısı nasıl değişmektedir?” sorusuna ait bulgular.



Şekil 9. Normal dağılımlarda test uzunluğu iç tutarlık katsayısı ilişkisi(likert)



Şekil 10. Sağa çarpık dağılımlarda test uzunluğu iç tutarlık katsayısı ilişkisi(likert)



Şekil 11. Sola çarpık dağılımlarda test uzunluğu iç tutarlık katsayısı ilişkisi(likert)

Şekil 9 incelendiğinde normal dağılımlarda testin uzunluğu arttıkça Cronbach α değeri için değişimin azaldığı görülmektedir. Madde sayısının farklı olduğu durumların çoğunda Cronbach α değerindeki değişim benzer bir sistematiğe azalmaktadır. Örneklem büyüklüğünün artmasının Cronbach α değerinin değişimini azalttığı normal dağılım temelli analizlerle de doğrulanmaktadır. Normal dağılımlı veriler için yapılan analiz sonuçlarının çoğunda kayıp veri oranının artışı ile Cronbach α değerinin değişiminin arttığı görülmektedir. Normal dağılımlarda Cronbach α değeri üzerinde kayıp veri oranı, örneklem büyüklüğü ve dağılım biçiminin önemli olduğu söylenebilir.

Şekil 10 incelendiğinde sağa çarpık dağılımlarda testin uzunluğu arttıkça Cronbach α değeri için değişimin azaldığı görülmektedir. Bu azalışlar %5 ve %10 oranında kayıp verilerde daha yavaş ve birbirine yakın değerlerde olurken %20 oranında kayıp verilerde daha hızlı gelişmektedir. Örneklem büyüklüğünün artmasının Cronbach α değerinin değişimini azalttığı sağa çarpık dağılım temelli analizlerle de doğrulanmaktadır. Sağa çarpık dağılımlı veriler için yapılan analiz sonuçlarının tamamına yakınında kayıp veri oranının artışı Cronbach α değerinin değişiminin arttığı görülmektedir. Madde sayısının azaltıldığı durumların çoğunda Cronbach α değerindeki değişim benzer bir sistematiğe azalmaktadır Burada 50 soruluk veri setlerinde azalış olması beklenirken çok küçük bir artış gözlenmektedir. Ancak bu artışın çok düşük olması durumun tesadüfle açıklanabileceğini

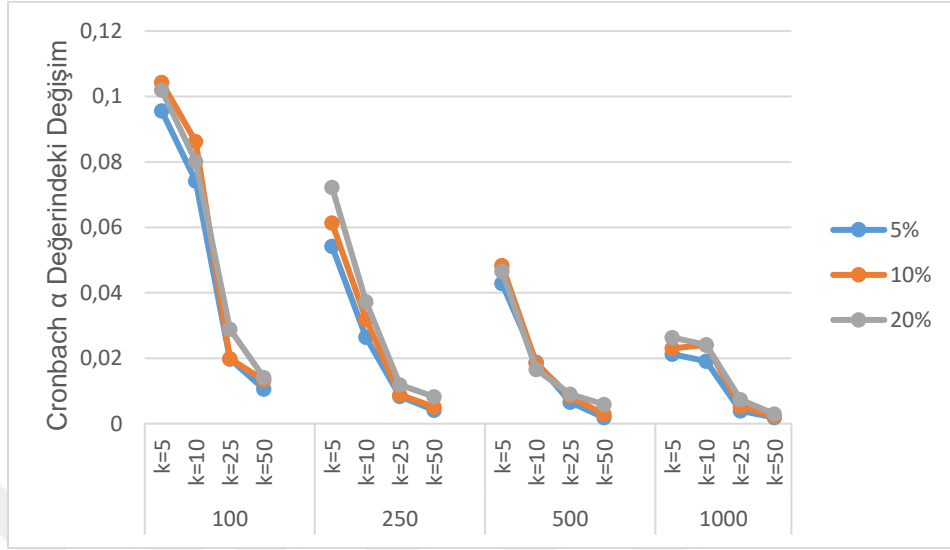
düşündürmektedir. Normal dağılımlarda Cronbach α değeri üzerinde kayıp veri oranı, örneklem büyüklüğü ve dağılım biçiminin önemli olduğu söylenebilir.

Şekil 11 incelendiğinde sola çarpık dağılımlarda testin uzunluğu arttıkça Cronbach α değeri için değişimin azaldığı görülmektedir. Bu azalışlar %5 ve %10 oranında kayıp verilerde daha yavaş ve birbirine yakın değerlerde olurken %20 oranında kayıp verilerde daha hızlı ilerlemektedir. Örneklem büyüklüğünün artmasının Cronbach α değerinin değişimini azalttığı sağa çarpık dağılım temelli analizlerle de doğrulanmaktadır. Sağa çarpık dağılımlı veriler için yapılan analiz sonuçların tamamına yakınında kayıp veri oranının artışı Cronbach α değerinin değişiminin arttığı görülmektedir. Madde sayısının azaltıldığı durumların çoğunda Cronbach α değerindeki değişim benzer bir sistematiğe azalmaktadır. Burada 50 soruluk veri setlerinde azalış olması beklenirken çok küçük bir artış gözlenmektedir. Ancak bu artışın çok düşük olması durumun tesadüfle açıklanabileceğini düşündürmektedir. Normal dağılımlarda Cronbach α değeri üzerinde kayıp veri oranı, örneklem büyüklüğü ve dağılım biçiminin önemli olduğu söylenebilir.

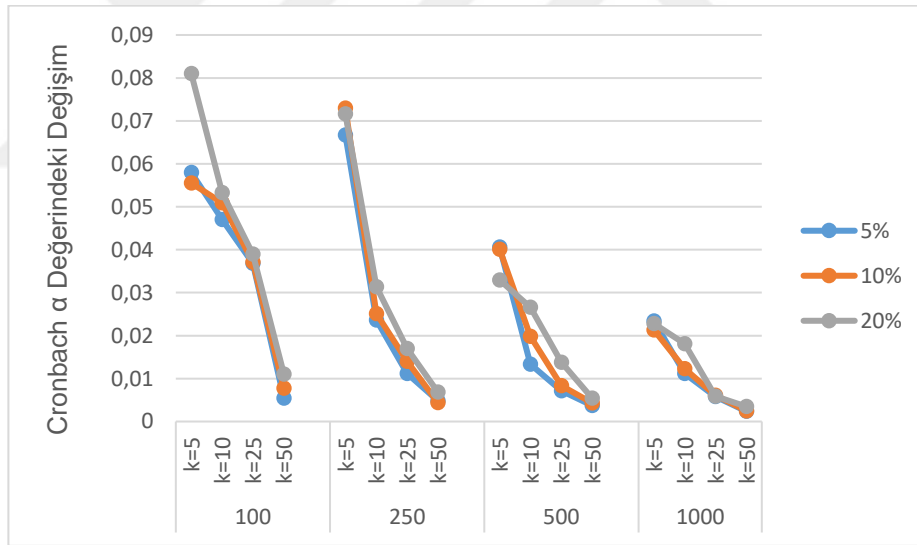
Araştırmanın 2. Alt Problemine Yönelik Bulgular

Bu bölümde problem cümlesine çözüm bulmak amacıyla oluşturulan ikinci alt problem olan “Farklı oranlarda kayıp veri içeren ikili veri seti için hesaplanan iç tutarlık değerleri nasıl değişmektedir?” sorusuna cevap aramak için elde edilen bulgular örneklem büyüklüğü, soru sayısı ve dağılım biçimi alt başlıklarında tartışılmıştır.

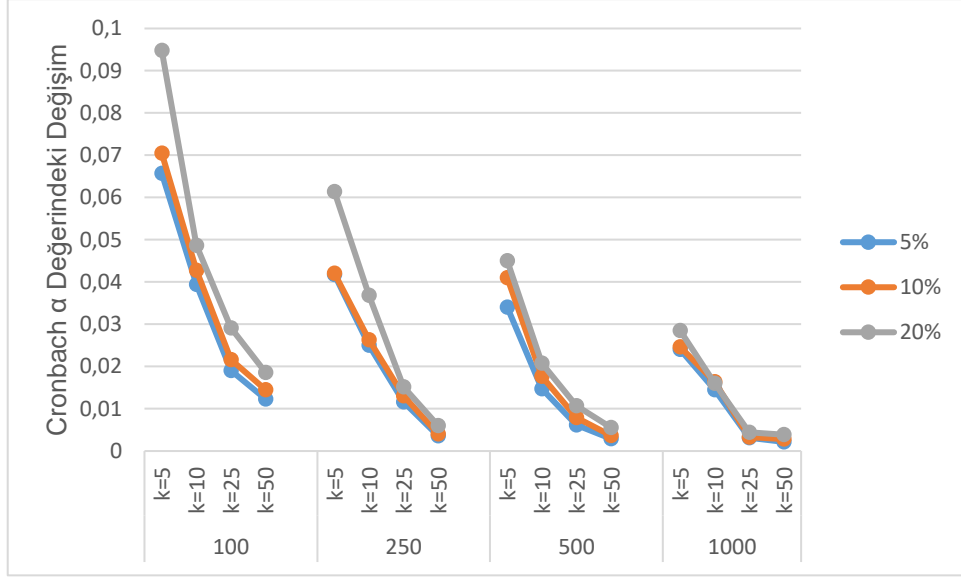
“Farklı oranlarda kayıp veri içeren ikili veri setlerinde örneklem büyüklüğüne göre iç tutarlık katsayısı nasıl değişmektedir?” sorusunu ait bulgular.



Şekil 12. Normal dağılımlarda örneklem büyüklüğü iç tutarlık katsayısı ilişkisi(ikili)



Şekil 13. Sağa çarpık dağılımlarda örneklem büyüklüğü iç tutarlık katsayısı ilişkisi(ikili)



Şekil 14. Sola çarpık dağılımlarda örneklem büyüklüğü iç tutarlık katsayısı ilişkisi(ikili)

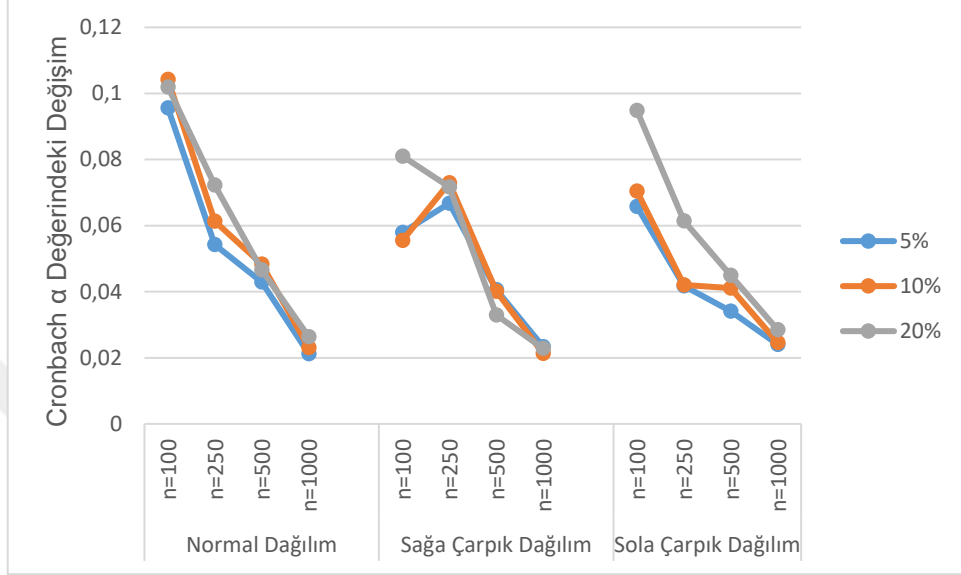
Şekil 12 incelendiğinde normal dağılımlarda örneklem büyüklüğü arttıkça Cronbach α değerindeki değişimin azaldığı görülmektedir. Madde sayısı yani testin uzunluğu arttıkça Cronbach α değerindeki değişimin azalması beklenmektedir. Şekil 12'deki sonuçlar incelendiğinde bütün örneklem büyüklüklerinde madde sayısı arttıkça Cronbach α değerindeki değişimin beklentilere uygun olarak azaldığı görülmektedir. Öte yandan koşulların çoğunda kayıp veri oranı arttıkça Cronbach α değerindeki değişimin de arttığı görülmektedir.

Şekil 13 incelendiğinde sağa çarpık dağılımlarda örneklem büyüklüğü arttıkça Cronbach α değeri için değişim azaldığı görülmektedir. Benzer şekilde madde sayısı arttıkça Cronbach α değerindeki değişimin azaldığı görülmektedir. Diğer yandan kayıp veri oranı arttıkça Cronbach α değerlerindeki değişimin arttığı görülmektedir. Bu artış küçük örneklerde daha hızlı gelişirken örneklem büyüklüğü arttıkça yavaşlamaktadır.

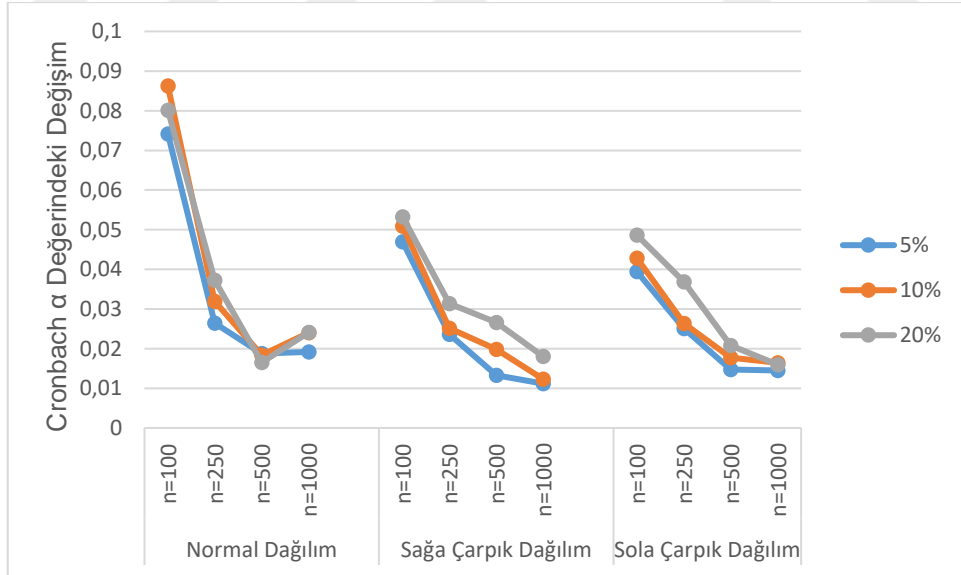
Şekil 14 incelendiğinde sola çarpık dağılımlarda örneklem büyüklüğü arttıkça Cronbach α değerindeki değişimin azaldığı görülmektedir. Testin uzunluğu arttıkça Cronbach α değerindeki değişimin azalması beklenen bir durumdur. Yeterli madde sayısına ulaşıldığında Cronbach α değerindeki değişim için hesaplanan değerlerin birbirine yakın olduğu görülmektedir. Öte yandan kayıp veri oranı arttıkça standart sapma artmakta, dolayısıyla Cronbach α değerindeki değişim artmaktadır. Bu artış küçük örneklerde daha belirgin gözlenebilirken örneklem büyüklüğü arttıkça

birbirine yakın değerlerde seyretmiştir. Sonuç olarak örneklem büyüklüğü arttıkça Cronbach α kararlı hale gelmektedir.

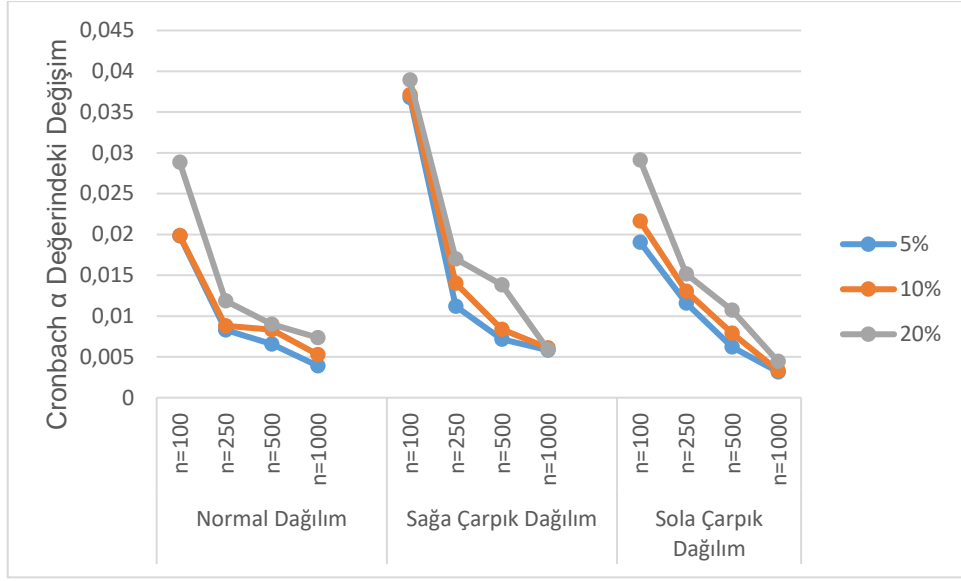
“Farklı oranlarda kayıp veri içeren ikili veri setlerinde dağılım biçimine göre iç tutarlık katsayısı nasıl değişmektedir?” sorusuna ait bulgular.



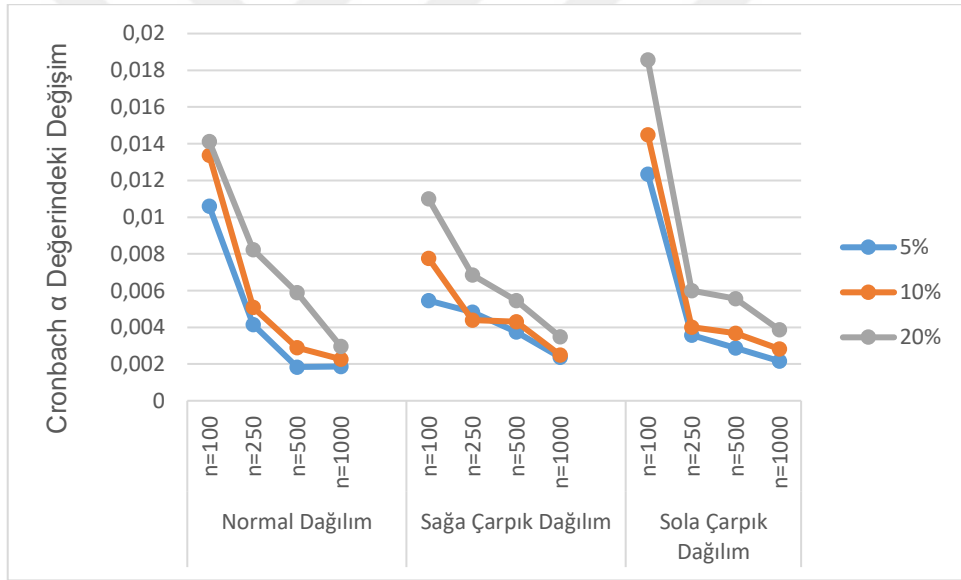
Şekil 15. 5 maddelik veri setleri için dağılım biçimi iç tutarlık katsayısı ilişkisi(ikili)



Şekil 16. 10 maddelik veri setleri için dağılım biçimi iç tutarlık katsayısı ilişkisi(ikili)



Şekil 17. 25 maddelik veri setleri için dağılım biçimi iç tutarlık katsayısı ilişkisi(ikili)



Şekil 18. 50 maddelik veri setleri için dağılım biçimi iç tutarlık katsayısı ilişkisi(ikili)

Şekil 15 incelendiğinde hem normal hem de çarpık dağılım biçimlerinde Cronbach α değerindeki değişim örneklem büyüklüğü arttıkça azalmaktadır. Cronbach α değeri değişimdeki azalış normal dağılım içeren veri setlerinde daha hızlı gerçekleşirken çarpık dağılımlarda daha yavaş gerçekleşmektedir. Diğer yandan kayıp veri oranı arttıkça standart sapma artmakta, dolayısıyla Cronbach α değerindeki değişim de artmaktadır.

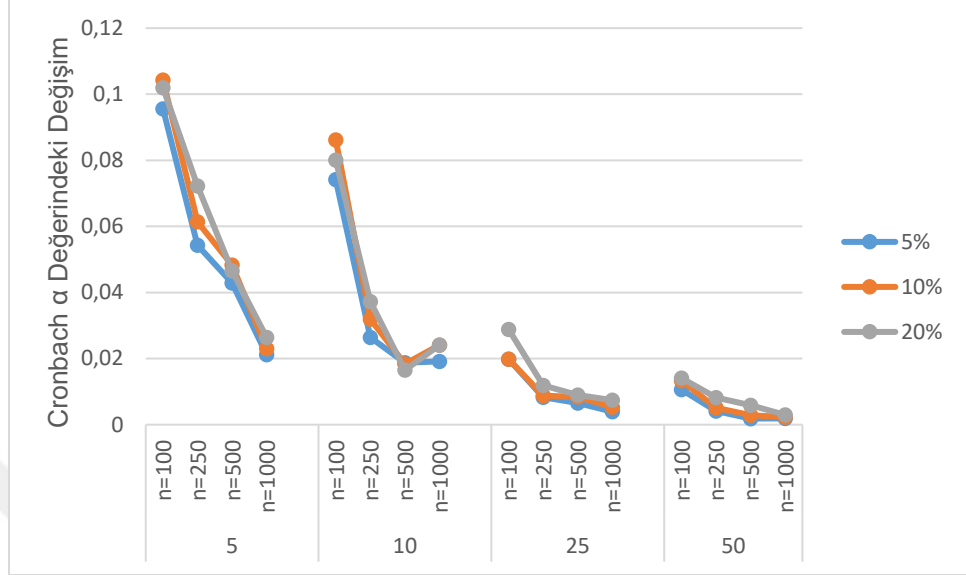
Şekil 16 incelendiğinde hem normal dağılım hem de çarpık dağılımlarda Cronbach α değeri için değişim azalmaktadır. Aynı şekilde örneklem büyüklüğü

arttıkça Cronbach α deęerindeki deęişimin azaldığı görölmektedir. Dięer yandan Cronbach α deęerindeki deęişim normal daęılımlarda 500 kişilik örneklemden 1000 kişilik örnekleme geçişte artış şeklinde kendini göstermektedir. Ancak bu deęişim fark yaratan bir deęişim deęildir ve tesadüfilikle açıklanabilir. Öte yandan kayıp veri oranı arttıkça standart sapma artmakta, dolayısıyla Cronbach α deęerindeki deęişim artmaktadır.

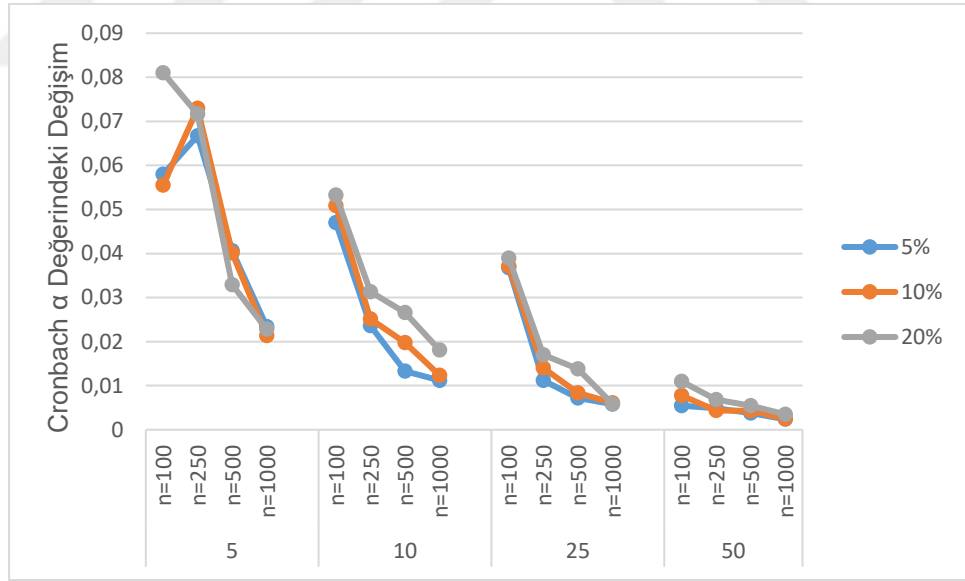
Şekil 17'ye bakıldığında hem normal daęılım hem de çarpık daęılımlarda Cronbach α deęerindeki deęişim azalmaktadır. Cronbach α deęeri deęişimindeki azalış 5 maddelik veri setlerinde olduđu gibi normal daęılım içeren veri setlerinde daha hızlı gerçekleşirken sola çarpık ve sağa çarpık daęılımlarda daha yavaş gerçekleşmektedir. Benzer şekilde örneklem büyüklüğü arttıkça Cronbach α deęerindeki deęişimin azaldığı görölmektedir. Dięer yandan kayıp veri oranı arttıkça Cronbach α deęerindeki deęişimin arttığı görölmektedir. Ayrıca hem normal hem de sağa çarpık ve sola çarpık daęılımlarda örneklem büyüklüğü arttıkça Cronbach α deęerlerinin birbirine oldukça yakın olduđu görölmektedir. Öte yandan 1000 örneklem büyüklüğüne ulaşıldıktan sonra kayıp veri oranının etkisi azalmakta Cronbach α deęeri daha kararlı hale gelmektedir.

Şekil 18 incelendiğinde hem normal daęılım hem de sağa çarpık ve sola çarpık daęılımlarda Cronbach α deęerindeki deęişim azalmaktadır. Benzer şekilde örneklem büyüklüğü arttıkça Cronbach α deęerindeki deęişimin azaldığı görölmektedir. Dięer yandan kayıp veri oranı arttıkça Cronbach α deęerindeki deęişimin arttığı görölmektedir. Dięer yandan kayıp veri oranı arttıkça Cronbach α deęerindeki deęişim deęerlerinin birbirine oldukça yakın olduđu görölmektedir.

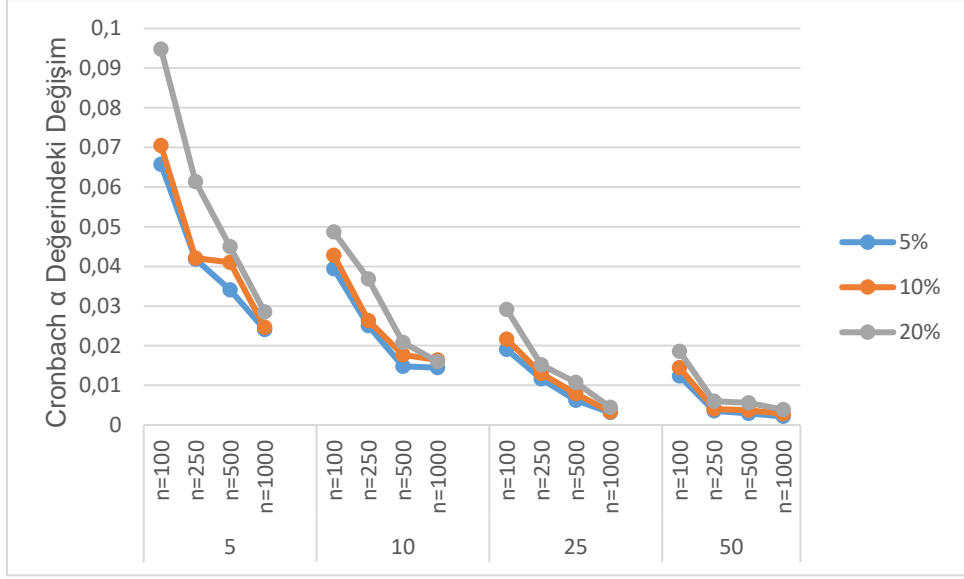
“Testin uzunluğuna göre farklı oranlarda kayıp veri içeren ikili veri setlerinde iç tutarlık katsayısı nasıl değişmektedir?” sorusuna ait bulgular.



Şekil 19. Normal dağılımlarda test uzunluğu iç tutarlık katsayısı ilişkisi(ikili)



Şekil 20. Sağa çarpık dağılımlarda test uzunluğu iç tutarlık katsayısı ilişkisi(ikili)



Şekil 21. Sola çarpık dağılımlarda test uzunluğu iç tutarlık katsayısı ilişkisi(ikili)

Şekil 19 incelendiğinde normal dağılımlarda testin uzunluğu arttıkça Cronbach α değerindeki değişimin azaldığı görülmektedir. Benzer olarak örneklem büyüklüğü arttıkça Cronbach α değerindeki değişimin azaldığı görülmektedir. 10 maddelik veri setleri incelendiğinde 500 kişilik örneklemden 1000 kişilik örnekleme geçişte Cronbach α değerlerindeki değişimde azalış olması beklenirken çok küçük bir artış gözlenmektedir. Ancak bu artışın çok düşük olması durumun tesadüfle açıklanabileceğini düşündürmektedir. Farklı test maddeleri için çoğu koşulda kayıp veri oranı arttıkça Cronbach α değerlerindeki değişimin arttığı görülmektedir.

Şekil 20 incelendiğinde sağa çarpık dağılımlarda testin uzunluğu arttıkça Cronbach α değeri için değişimin azaldığı görülmektedir. Bu azalış 5 maddelik veri setleri için 100 kişilik örneklemden 250 kişilik örnekleme geçişte %20 oranında kayıp veri için normalken, %5 ve %10 oranındaki kayıp veri için bozulmakta azalması beklenirken artış şeklinde gerçekleşmektedir. Normal dağılımlarda olduğu gibi kayıp veri oranı arttıkça farklı sayıda test maddeleri için hesaplanan Cronbach α değerlerindeki değişimin arttığı görülmektedir. Diğer yandan farklı test uzunluklarında, örneklem büyüklüğü arttıkça Cronbach α değerlerindeki değişimin azaldığı görülmektedir.

Şekil 21'e bakıldığında sola çarpık dağılımlarda testin uzunluğu arttıkça Cronbach α değeri için değişim azaldığı görülmektedir. 50 maddeden oluşan veri setlerinde 100 kişilik örneklem büyüklüğünden sonra kayıp veri oranı arttıkça hesaplanan Cronbach α değerlerindeki değişimin birbirine oldukça yakın olduğu

görülmektedir. Büyük örneklerde madde sayısı arttıkça kayıp veri oranının etkisinin azaldığı söylenebilir.



Bölüm 5

Sonuç, Tartışma ve Öneriler

Likert Tipi Verilere Ait Sonuçlar

Farklı oranlarda kayıp veri içeren Likert tipi veri setlerinde örneklem büyüklüğü arttıkça Cronbach α değerleri için hesaplanan değişim azalmaktadır.

Farklı oranlarda kayıp veri içeren Likert tipi veri setlerinde 5 maddelik testler ele alınmadan düşünüldüğünde madde sayısı arttıkça (testin uzunluğu arttıkça) normal dağılımlar için Cronbach α değerlerindeki değişim beklenmedik şekilde artmaktadır. Sağa çarpık dağılımlar için madde sayısı arttıkça Cronbach α değerindeki değişim azalmaktadır. Sola çarpık dağılımlar için madde sayısı arttıkça madde sayısı arttıkça Cronbach α değerindeki değişim azalmaktadır.

Farklı oranlarda kayıp veri içeren Likert tipi veri setlerinde her üç dağılım biçimi için Cronbach α değerindeki değişim azalmaktadır.

Farklı oranlarda kayıp veri içeren Likert tipi veri setlerinde kayıp veri oranı arttıkça Cronbach α değerindeki değişim artmaktadır.

İkili Verilere Ait Sonuçlar

Farklı oranlarda kayıp veri içeren ikili veri setlerinde örneklem büyüklüğü arttıkça Cronbach α değerleri için hesaplanan değişim azalmaktadır.

Farklı oranlarda kayıp veri içeren ikili veri setlerinde madde sayısı arttıkça (testin uzunluğu arttıkça) arttıkça Cronbach α değerindeki değişim normal dağılım ve sağa çarpık dağılımlarda azalmakta ancak sola çarpık dağılımlarda düzenli şekilde ilerlememektedir.

Farklı oranlarda kayıp veri içeren ikili veri setlerinde her üç dağılım biçimi için Cronbach α değerindeki değişim azalmaktadır.

Farklı oranlarda kayıp veri içeren ikili veri setlerinde kayıp veri oranı arttıkça Cronbach α değerindeki değişim artmaktadır.

Likert tipi veri setlerinde örneklem büyüklüğü arttıkça Cronbach α değerlerindeki değişimin azalması beklenen bir durumdur. Bu sonuç Soğuksu ve Alıcı (2016) tarafından yapılan çalışma ile benzerlik göstermektedir. Benzer şekilde madde sayısı arttıkça Cronbach α değerlerindeki değişimin azalması beklenen bir durumdur. Ancak normal dağılımlarda 100 kişilik örneklem için 5 maddelik teste hesaplanan Cronbach α değişimindeki değer, 10 maddelik testlerden daha yüksek iken 10 maddelik testler için hesaplanan değerler 25 maddelik testlerden daha yüksek, 25 maddelik testler için hesaplanan değerler ise 50 maddelik testlerde hesaplanan değerlerden daha yüksektir. 1000 kişilik örneklem grubunda 25 maddelik veri setinde hesaplanan Cronbach α değişimindeki değer beklenen şekilde 50 maddelik veri setinde hesaplanan değerden yüksek değil aksine düşüktür.

Sola çarpık dağılımlarda aynı sağa çarpık dağılımda olduğu gibi 1000 kişilik örnekleme 25 maddelik veri setinde hesaplanan değer 50 maddelik veri setinde hesaplanan değerden düşük çıktığı görülmüştür. Sonuç olarak çarpık dağılımlarda büyük örneklemlerde madde sayısı arttıkça hesaplanan Cronbach α değerlerindeki değişim beklenenin dışında seyretmektedir. Araştırmacılar bu konuda dikkatli olmalı kayıp veri içeren çarpık dağılımlar için genelleme yapılamadığını göz önünde bulundurmalarıdır. Hem normal dağılım hem de çarpık dağılımlarda kayıp veri oranı arttıkça beklendiği şekilde Cronbach α değerindeki değişim de arttığı görülmüştür. Dolayısıyla Cronbach α değerindeki değişim dağılım biçiminden etkilenmediği söylenebilir. Bu durum yapılan diğer çalışmalarla benzerlik göstermektedir (Chen, Wang ve Chen 2012; Çoklu ve Kayrı 2011; Şahin Kürşad 2014).

Likert tipi veri setlerinde Cronbach α değerlerindeki değişimin azalması normal dağılıma sahip veri setlerinde daha keskin biçimde olurken gerek sağa çarpık gerek sola çarpık dağılımlarda çok net biçimde gözlenememektedir. Başka bir deyişle sağa çarpık dağılım ve sola çarpık dağılım için Cronbach α değerindeki değişim için hesaplanan değerler birbirine çok yakındır ve oldukça yavaş bir şekilde azalmaktadır. Likert tipi veri setlerinde testin uzunluğu arttıkça Cronbach α değerlerindeki değişim azalmakta olduğu görülmüştür. Ancak bu azalış %5 ve %10 oranında kayıp veri içeren veri setlerinde birbirine yakın değerler alırken %20 oranında kayıp veri içeren veri setlerinde daha keskin bir şekilde gözlenebilmektedir.

İkili veri setlerinde hem normal dağılımlarda hem de sağa çarpık ve sola çarpık dağılımlarda örneklem büyüklüğü arttıkça Cronbach α değerlerindeki değişim

azalmaktadır. Bu zaten beklenen bir durumdur. Bu deęişim normal daęılım ieren veri setlerinde daha hızlı gerekleşmekte ancak arpık daęılımlarda daha yavaş gerekleşmektedir. Öte yandan saęa arpık daęılımda 100 kişilik örneklemden 250 kişilik örnekleme geişte bu deęişimin azalması beklenirken artması genellemeyi bozmuştur. Benzer şekilde arpık daęılımlarda Cronbach α deęeri iin deęişim azalırken normal daęılımda bu azalış 500 kişilik örneklemden 1000 kişilik örnekleme geişte artış şeklinde gerekleşmiştir. Bu durumu arpık daęılımlar iin genelleme yapılamayacağını göstermektedir.

Madde sayısı arttıka testin uzunluęu artacağından Cronbach α deęerlerindeki deęişimin azalması beklenmektedir. Normal daęılıma ve saęa arpık daęılıma sahip veri setlerinde bu beklenti karşılanırken sola arpık veri setlerinde hesaplanan deęerler, 250, 500 ve 1000 kişilik örneklemlerde sorunsuz bir şekilde azalırken; 100 kişilik örnekleme 5 maddelik veri seti ile 25 maddelik veri seti arasında olması gerektięi gibi azalmış ancak 25 maddelik veri setinden 50 maddelik veri setine geişte beklenmedik şekilde artış göstermektedir. Bu durum kayıp veri ieren sola arpık daęılımlarla ilgili analiz yapılırken göz önünde bulundurulmalıdır. Kayıp veri oranı arttıka hesaplanan Cronbach α deęerindeki deęişim aynı şekilde artmaktadır. Bu artış küçük örneklemlerde daha keskin gözlenirken örneklem büyüklüęü arttıka deęerler birbirine yaklaşmaktadır. İkili veri setlerinde hem normal daęılımlarda hem de saęa arpık ve sola arpık daęılımlarda testin uzunluęu arttıka Cronbach α deęerindeki deęişimin azaldığı genellemesi yapılabilir. Benzer şekilde birbirinden farklı sayıda madde ieren veri setleri iin kayıp veri oranı arttıka Cronbach α deęerlerindeki deęişim de artmaktadır.

Likert Tipi Veri Setleri İle İkili Veri Setlerinin Karşılaştırılması

Hem normal daęılıma sahip Likert tipi veri setlerinde hem de normal daęılıma sahip ikili veri setlerinde örneklem büyüklüęü arttıka Cronbach α deęerinin deęişiminin azaldığı görülmektedir. Benzer olarak hem normal daęılıma sahip Likert tipi veri setlerinde hem de normal daęılıma sahip ikili veri setlerinde kayıp veri oranı arttıka Cronbach α deęerindeki deęişimin arttığı görülmektedir. Öte yandan normal daęılıma sahip Likert tipi veri setlerinde Cronbach α deęerinin deęişiminin madde sayısına baęlı bir sistematik izlemedięi görülürken, normal daęılıma sahip ikili veri

setlerinde madde sayısı arttıkça Cronbach α deęerindeki deęişimin azaldığı görölmektedir.

Hem saęa çarpık dağılıma sahip Likert tipi veri setlerinde hem de saęa çarpık dağılıma sahip ikili veri setlerinde örneklem büyüklüęü arttıkça Cronbach α deęerinin deęişiminin azaldığı görölmektedir. Benzer şekilde hem saęa çarpık dağılıma sahip Likert tipi veri setlerinde hem de saęa çarpık dağılıma sahip ikili veri setlerinde madde sayısı arttıkça Cronbach α deęerindeki deęişimin azaldığı görölmektedir. Dięer yandan Hem saęa çarpık dağılıma sahip Likert tipi veri setlerinde hem de saęa çarpık dağılıma sahip ikili veri setlerinde kayıp veri oranı arttıkça Cronbach α deęerlerindeki deęişim de arttığı görölmektedir.

Hem sola çarpık dağılıma sahip Likert tipi veri setlerinde hem de sola çarpık dağılıma sahip ikili veri setlerinde örneklem büyüklüęü arttıkça Cronbach α deęerinin deęişiminin azaldığı görölmektedir. Benzer şekilde hem sola çarpık dağılıma sahip Likert tipi veri setlerinde hem de sola çarpık dağılıma sahip ikili veri setlerinde madde sayısı arttıkça Cronbach α deęerindeki deęişimin azaldığı görölmektedir. Öte yandan hem sola çarpık dağılıma sahip Likert tipi veri setlerinde hem de sola çarpık dağılıma sahip ikili veri setlerinde kayıp veri oranı arttıkça Cronbach α deęerlerindeki deęişim de arttığı görölmektedir.

Öneriler

Araştırmacılara öneriler. Bu çalışma ele alınan benzetim koşullarıyla sınırlandırılmıştır. Araştırmacılar bu benzetim koşullarını deęiştirerek özellikle dağılım biçimi ve madde uzunluęu deęişkenlerindeki farklılaşmayla güvenilirlik ilişkisini inceleyebilirler.

Bu çalışmada sadece kayıp veri içeren veri setleri ile çalışılmıştır. Başka bir çalışmada tam veri setlerinde hesaplanan güvenilirlik deęerleri ile kayıp veri içeren eksik veri setlerinde hesaplanan güvenilirlik deęerleri karşılaştırılabilir.

Farklı bir çalışmada okullardaki şubelerin öğrenci sayısı dikkate alınarak en düşük örneklem büyüklüęü 25 kabul edilip artan şekilde örneklerle çalışıldığında hesaplanan deęerlerin bu çalışma ile benzer şekilde olup olmayacağı incelenebilir.

Araştırmacılar yapacakları çalışmalarda, çarpık dağılımlardaki kayıp veri durumunu farklı uzunluktaki testler için inceleyebilirler.

Kullanıcılara öneriler. Kayıp deęer ieren veri seti ile alıřan kullanıcılar, kayıp veri oranının analiz sonuçlarını etkilediđini gz nne alarak mmknse kayıp verinin az olduđu durumlarla alıřmalarına devam etmelidirler.

Kayıp deęer ieren veri setlerinde rneklem byklđ ve madde sayısı arttıķa Cronbach α deęerindeki deęiřim azalacađı gz nne alınarak alıřmalarını bu ynde dzenleyebilirler.



Kaynaklar

- Acuna, E., & Rodriguez, C. (2004). The treatment of missing values and its effect on classifier accuracy. *Studies in classification, data analysis and knowledge organisation* (s. 639-647). Springer, Berlin, Heidelberg.
- Afifi, A. A., & Elashoff, R. M. (1966). Missing observations in multivariate statistics I. review of the literature. *Journal of the American Statistical Association*(61), 595-604.
- Aiken, L. R. (1985). *Psychological testing and assessment*. Allyn & Bacon.
- Allison, P. D. (2002). *Missing Data*. Thousands Oaks: Sage Publication.
- Allison, P. D. (2003). Missing data techniques for structural equation modeling. *Journal of Abnormal Psychology*(112), 545-557.
- Allison, P. D. (2009). *Missing data* (Sage University Paper Series on Quantitative Applications in the Social Sciences, 72-89). London: Sage Publication.
- Alkan N. (2012). *Kayıp verili cox regresyon yöntemine Bayesci bir yaklaşım*. (Yayımlanmamış Doktora Tezi). Ondokuz Mayıs Üniversitesi, Samsun.
- Alpar, R. (2011). *Çok değişkenli istatistiksel yöntemler*. Ankara: Detay Yayıncılık.
- Akbaş, U. (2014). Farklı örneklem büyüklüğü ve kayıp veri örüntülerinde ölçeklerin psikometrik özelliklerinin kayıp veri bas etme teknikleri ile incelenmesi. Yayımlanmamış Doktora Tezi, Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Akbaş, U., & Tavşancıl, E. (2015). Farklı örneklem büyüklüklerinde ve kayıp veri örüntülerinde ölçeklerin psikometrik özelliklerinin kayıp veri baş etme teknikleri ile incelenmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 6(1), 38-57.
- Aslan, S. (2010). Comparison of missing value imputation methods for meteorological time series data. (Yayımlanmamış Yüksek Lisans Tezi). Middle East Technical University, Ankara.
- Aydilek, İ. B. (2013). Veri kümelerindeki eksik değerlerin yeni yaklaşımlar kullanılarak hesaplanması. (Yayımlanmamış Doktora Tezi) Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Konya.

- 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.
- Baraldi, A. D., & Enders, C. K. (2010). An introduction to modern missing data analyses. *Journal of School Psychology*(48), 5-37.
- Baygül, A. (2007). Kayıp veri analizinde sıklıkla kullanılan etkin yöntemlerin değerlendirilmesi. (Yayımlanmamış Yüksek Lisans Tezi). İstanbul Üniversitesi, Sağlık Bilimleri Enstitüsü, İstanbul.
- Bennett, D. A. (2001). How can i deal with missing data in my study? *Australian and New Zealand Journal of Public Health*(25), 464-469.
- Buhi, E. R., Goodson, P., & Neilands, T. B. (2008). Out of sight, noot out of mind: Strategies for handling missing data. *American Journal Health Behaviour*(32), 83-92.
- Chen, F. S., Wang, S., & Chen, Y. C. (2012). A simulation study using EFA and CFA programs based the impact of missing data on test dimensionality. *Expert Systems with Applications: An International Journal*, 39(4), 4026-4031.
- Comrey, A. L., & Lee, H. B. (2013). *A first course in factor analysis*. Psychology Press.
- Cool, A. L. (2000). A Rewiev of Methods for Dealing with Missing Data.
- Çokluk Bökeoğlu, U., & Koçak, D. (2017). Kayıp Veriyle Baş Etme Yöntemlerinin Model Veri Uyumu Ve Madde Model Uyumuna Etkisi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 200-223.
- Çokluk, Ö., & Kayrı, M. (2011). Kaayıp değerlere yaklaşık değer atama yöntemlerinin ölçme araçlarının geçerlik ve güvenilirliği üzerindeki etkisi. *Kuram ve Uygulamada Eğitim Bilimleri*(11), 289-309.
- Demir, E. (2013). Kayıp verilerin varlığında çoktan seçmeli testlerde madde ve test parametrelerinin kestirilmesi. *Eğitim Bilimleri Araştırma Dergisi*(3), 47-68.
- Demir, E. (2013). Kayıp verilerin varlığında iki kategorili puanlanan maddelerden oluşan testlerin psikometrik özelliklerinin incelenmesi. (Yayımlanmamış Doktora Tezi). Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.

- Demir, E., & Parlak, B. (2012). Türkiye eğitim arařtırmalarında kayıp veri sorunu. *Eđitimde ve Psikolojide Ölçme ve Deđerlendirme Dergisi*(3), 230-241.
- Dempster, A. P. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society. Series B (methodological)*, 1-38.
- Dural, S. (2010). Farklı kayıp veri tekniklerinin çok göstergeli örtük gelisme modelleri üzerindeki etkisi. (Yayımlanmamıs Doktora Tezi). Ege Üniversitesi, Sosyal Bilimler Enstitüsü, İzmir.
- Enders, C. K. (2010). *Applied missing data analysis*. New York: Guilford Press.
- Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling: A Multidisciplinary*, 430-457.
- Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*(60), 549-576.
- Graham, J. W. (2012). *Missing Data : Analysis and Design*. New York: Springer.
- Horton, N. J., & Kleinman, K. P. (2007). Much Ado About Nothing: A Comparison of Missing Data Methods. *American Statistical Association*(61), 79-90.
- Howell, D. C. (2007). The Treatment of Missing Data. W. Outhwaite, & S. P. Turner içinde, *The SAGE handbook of social science methodology* (s. 208-224). Los Angeles: Sage Publications.
- Kalaycı, Ş. (2018). *SPSS Uygulamalı Çok Deđişkenli İstatistik Teknikleri*. Ankara: Dinamik Akademi Yayın Dađıtım.
- Kaspar, E. Ç. (2011). Kayıp veriler ve kayıp veriler için bir çoklu veri atama yöntemi: Prospensity skor. (Yayımlanmamıs Doktora Tezi). Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.
- Kim, J., & Curry, J. (1977). The Treatment of Missing Data in Multivariate Analysis. *Sociological Methods & Research*(6), 215-240.
- Kros, J. F., & Brown, M. L. (2003). Data mining and the impact of missing data. *Industrial Management & Data System*, 103(8), 611-621.

- Little, R. J., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: Wiley .
- Little, R. J., & Rubin, D. B. (2002). *Statistical Analysis With Missing Data*. Canada: John Wiley and Sons.
- Longford, N. (2005). *Missing data and small-area estimation: Modern analytical equipment for the survey statistician*. Springer Science & Business Media.
- Misztal, M. (2012). Imputation of Missing Data Using R Package. *Acta Universitatis Lodzianis Folia Oeconomica*(269), 132-144.
- Nakai , M., & Ke, W. (2011). Review of the Methods for Handling Missing Data in Longitudinal Data Analysis. *Journal of Math. Analysis*, 1-13.
- Nartgün, Z. (2015). Comparison of Various Methods Used in Solving Missing Data Problems in terms of Psychometric Features of Scales and Measurement Results under Different Missing Data Conditions. *International Online Journal of Educational Sciences*(7), 252-265.
- Nunnally, J., & Bernstein, I. (1994). *Psychometric Theory* (3 b.). New York: McGraw-Hill.
- Oğuzlar, A. (2001). Alan arařtırmalarında kayıp deęer problemi ve çözüm önerileri. *Ulusal Ekonometri ve İstatistik Sempozyumu Çukurova Üniversitesi İİBF Ekonometri Bölümü*, (s. 19-22). Adana.
- Osborne, J. W. (2013). *Best practices in data cleaning*. Thousand Oaks: SAGE Publications.
- Öztemur, B. (2014). Kayıp veri yöntemlerinin farklı deęişkenler altında varyans analizi (t-testi,anova) parametreleri üzerine etkisinin incelenmesi. (Yayımlanmamış Yüksek Lisans Tezi). Abant İzzet Baysal Üniversitesi, Eğitim Bilimleri Enstitüsü, Bolu.
- Pembegül A. (2009). İhmal edilemeyen kayıp veri varlığında olumsuzluk tablo çözümlenmeleri. (Yayımlanmamış Yüksek Lisans Tezi). Hacettepe Üniversitesi, İstatistik Bölümü, Ankara.

- Peng, J. C.-Y., Harwell, M., Liou, S. M., & Ehman, L. H. (2006). Advances in Missing Data Methods and Implications for Educational Research. S. S. Sawilowsky içinde, *Real Data Analysis* (s. 31-78). New York.
- Pigott, T. D. (2001). A Review of Methods for Missing Data. *Educational Research and Evaluation*(4), 353-383.
- Rubin, D. R. (1976). Inference and missing data. *Biometrika*, 581-592.
- Sarı, İ. K. (2012). Karma ayrıştırma analizinde kayıp gözlem tahmin yöntemlerinin değerlendirilmesi. (Yayımlanmamış Yüksek Lisans Tezi). Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Konya.
- Schafer, J. L., & Graham, J. W. (2002). Missing data: our view of the state of the art. *Missing data: our view of the state of the art.*, 7(2), 144-177.
- Schlomer, G. L., Bauman, S., & Card, N. A. (2010). Best Practices for Missing Data Management in Counseling Psychology. *American Psychological Association*, 1-10.
- Schoier, G. (2004). On partial nonresponse situations:the hot deck imputation. *52nd Session of ISI University of Helsinki*.
- Sezgin, E., & Çelik, Y. (2013). Veri Madenciliğinde Kayıp Veriler İçin Kullanılan Yöntemlerin Karşılaştırılması. *Akademik Bilişim 2013 – XV. Akademik Bilişim Konferansı Bildirileri*, (s. 612-614). Akdeniz Üniversitesi, Antalya.
- Sinharay, S., Stern, H. S., & Russell, D. (2001). The use of multiple imputation for the analysis of missing data. *Psychological Methods*(6), 317-329.
- Soğuksu, Y. B., & Alıcı, D. (2016). Eşdeğer Yarılar Güvenirliğinin Farklı Homojenlik Düzeylerindeki Örneklem Büyüklüklerinde, Test Uzunluğuna, Yarıya Bölme Yöntemlerine ve Güvenirlik Kestirme Tekniklerine Göre İncelenmesi. *16*(1).
- Soley Bori, M. (2013). Dealing with missing data: Key assumptions and. *Technical Report*(4).
- Şahin Kürşad, M. (2014). Sıklıkla kullanılan kayıp veri yöntemlerinin betimsel istatistik, güvenilirlik ve geçerlik açısından karşılaştırılması. (Yayımlanmamış Yüksek Lisans Tezi). Abant İzzet Baysal Üniversitesi, Bolu.

- Tabachnick, B. G., & Fidell, L. S. (2015). Çok Değişkenli İstatistiklerin Kullanımı. (M. Baloğlu, Çev.) Nobel Yayın Dağıtım.
- Weaver, B., & Maxwell, H. (2014). Exploratory factor analysis and reliability analysis with missing data: A simple method for SPSS users. *The Quantitative Methods for Psychology*, 10(2), 143-152.
- Zhu, X. (2014). Comparison of Four Methods for Handling Missing Data in Longitudinal Data Analysis through a Simulation Study. *Open Journal of Statistics*, 933-944.



EK-A: Likert Tipi Veriler İçin Trok Testi Sonuçları

	N	K	Kayıp Veri Oranı	p
Likert Tipi Normal Dağılıma Sahip Verilerde Trok Testi Sonuçları	100	5	5%	0,50
			10%	0,50
			20%	0,58
		10	5%	0,51
			10%	0,54
			20%	0,60
		25	5%	0,41
			10%	0,63
			20%	0,68
		50	5%	0,58
			10%	0,61
			20%	0,92
	250	5	5%	0,70
			10%	0,49
			20%	0,04
		10	5%	0,67
			10%	0,59
			20%	0,72
		25	5%	0,62
			10%	0,40
			20%	0,60
		50	5%	0,68
			10%	0,53
			20%	0,64
	500	5	5%	0,45
			10%	0,77
			20%	0,84
		10	5%	0,59
			10%	0,44
			20%	0,65
		25	5%	0,43
			10%	0,38
			20%	0,54
		50	5%	0,67
			10%	0,56
			20%	0,60
1000	5	5%	0,68	
		10%	0,53	
		20%	0,69	
	10	5%	0,45	
		10%	0,42	
		20%	0,88	
	25	5%	0,26	
		10%	0,66	
		20%	0,44	
	50	5%	0,55	
		10%	0,49	
		20%	0,57	

	N	K	Kayıp Veri Oranı	p
Likert Tipi Sağa Çarpık Dağılıma Sahip Verilerde Trok Testi Sonuçları	100	5	5%	0,30
			10%	0,33
			20%	0,51
		10	5%	0,34
			10%	0,20
			20%	0,44
		25	5%	0,76
			10%	0,38
			20%	0,66
		50	5%	0,59
			10%	0,82
			20%	0,93
	250	5	5%	0,05
			10%	0,41
			20%	0,80
		10	5%	0,35
			10%	0,32
			20%	0,31
		25	5%	0,16
			10%	0,49
			20%	0,61
		50	5%	0,66
			10%	0,53
			20%	0,64
	500	5	5%	0,53
			10%	0,62
			20%	0,21
		10	5%	0,23
			10%	0,83
			20%	0,14
		25	5%	0,13
			10%	0,60
			20%	0,45
		50	5%	0,45
			10%	0,62
			20%	0,60
	1000	5	5%	0,60
			10%	0,68
			20%	0,79
		10	5%	0,45
			10%	0,75
			20%	0,57
		25	5%	0,84
			10%	0,37
			20%	0,45
		50	5%	0,59
			10%	0,55
			20%	0,57

	N	K	Kayıp Veri Oranı	p
Likert Tipi Sola Çarpık Dağılımlarda Trok Testi Sonuçları	100	5	5%	0,61
			10%	0,86
			20%	0,73
		10	5%	0,61
			10%	0,39
			20%	0,16
		25	5%	0,58
			10%	0,36
			20%	0,68
		50	5%	0,50
			10%	0,82
			20%	0,94
	250	5	5%	0,50
			10%	0,76
			20%	0,58
		10	5%	0,20
			10%	0,47
			20%	0,74
		25	5%	0,70
			10%	0,59
			20%	0,60
		50	5%	0,22
			10%	0,49
			20%	0,65
	500	5	5%	0,18
			10%	0,18
			20%	0,39
		10	5%	0,44
			10%	0,32
			20%	0,29
		25	5%	0,18
			10%	0,70
			20%	0,55
		50	5%	0,49
			10%	0,60
			20%	0,59
1000	5	5%	0,96	
		10%	0,85	
		20%	0,74	
	10	5%	0,72	
		10%	0,54	
		20%	0,23	
	25	5%	0,12	
		10%	0,55	
		20%	0,43	
	50	5%	0,54	
		10%	0,56	
		20%	0,57	

EK-B: İkili Veriler İçin Trok Testi Sonuçları

	N	K	Kayıp Veri Oranı	p
İkili Normal Dağılımlarda Trok Testi Sonuçları	100	5	5%	0,37
			10%	0,25
			20%	0,42
		10	5%	0,36
			10%	0,73
			20%	0,64
		25	5%	0,54
			10%	0,43
			20%	0,66
		50	5%	0,48
			10%	0,64
			20%	0,92
	250	5	5%	0,84
			10%	0,38
			20%	0,46
		10	5%	0,77
			10%	0,42
			20%	0,08
		25	5%	0,79
			10%	0,37
			20%	0,51
		50	5%	0,42
			10%	0,45
			20%	0,64
	500	5	5%	0,40
			10%	0,33
			20%	0,53
		10	5%	0,69
			10%	0,25
			20%	0,44
		25	5%	0,52
			10%	0,50
			20%	0,50
		50	5%	0,45
			10%	0,57
			20%	0,60
	1000	5	5%	0,21
			10%	0,64
			20%	0,05
		10	5%	0,82
			10%	0,43
			20%	0,56
25		5%	0,09	
		10%	0,54	
		20%	0,48	
50		5%	0,57	
		10%	0,50	
		20%	0,57	

	N	K	Kayıp Veri Oranı	p
İkili Sağa Çarpık Dağılımlarda Trok Testi Sonuçları	100	5	5%	0,20
			10%	0,94
			20%	0,34
		10	5%	0,17
			10%	0,35
			20%	0,88
		25	5%	0,37
			10%	0,23
			20%	0,67
		50	5%	0,67
			10%	0,65
			20%	0,96
	250	5	5%	0,70
			10%	0,67
			20%	0,38
		10	5%	0,39
			10%	0,18
			20%	0,74
		25	5%	0,22
			10%	0,39
			20%	0,58
		50	5%	0,48
			10%	0,49
			20%	0,64
	500	5	5%	0,33
			10%	0,61
			20%	0,12
		10	5%	0,57
			10%	0,23
			20%	0,66
		25	5%	0,72
			10%	0,49
			20%	0,44
		50	5%	0,22
			10%	0,47
			20%	0,60
	1000	5	5%	0,27
			10%	0,49
			20%	0,63
		10	5%	0,35
			10%	0,93
			20%	0,57
25		5%	0,28	
		10%	0,48	
		20%	0,59	
50		5%	0,51	
		10%	0,47	
		20%	0,57	

	N	K	Kayıp Veri Oranı	p
İkili Sola Çarpık Dağılımlarda Trok Testi Sonuçları	100	5	5%	0,09
			10%	0,75
			20%	0,15
		10	5%	0,51
			10%	0,84
			20%	0,45
		25	5%	0,41
			10%	0,53
			20%	0,70
		50	5%	0,56
			10%	0,82
			20%	0,96
	250	5	5%	0,70
			10%	0,42
			20%	0,89
		10	5%	0,69
			10%	0,51
			20%	0,69
		25	5%	0,32
			10%	0,57
			20%	0,45
		50	5%	0,89
			10%	0,46
			20%	0,64
	500	5	5%	0,87
			10%	0,76
			20%	0,32
		10	5%	0,36
			10%	0,90
			20%	0,49
		25	5%	0,71
			10%	0,38
			20%	0,50
		50	5%	0,25
			10%	0,52
			20%	0,60
	1000	5	5%	0,87
			10%	0,65
			20%	0,60
		10	5%	0,85
			10%	0,92
			20%	0,58
		25	5%	0,59
			10%	0,32
			20%	0,36
		50	5%	0,24
			10%	0,48
			20%	0,57

EK-C: Türetilmiş Veriler için Örnek Faktör Analizi Sonuçları

	Veri Grubu	Özdeğer	Açıklanan VaryansYüzdesi
1-0 Tipi Veriler	normal_5_500	1,602	32,043
	normal_10_1000	2,335	23,549
	normal_25_250	4,981	19,924
	normal_50_100	8,592	17,184
	sag_5_100	1,844	36,888
	sag_10_500	2,706	27,061
	sag_25_1000	5,009	20,036
	sag_50_500	8,640	17,280
	sol_5_250	1,553	31,056
	sol_10_500	2,378	23,775
	sol_25_100	4,915	19,662
	sol_50_1000	9,042	18,084
	Likert Tipi Veriler	normal_5_250	1,321
normal_10_500		1,625	16,208
normal_25_100		4,965	19,860
normal_50_1000		7,344	14,688
sag_5_500		1,504	30,073
sag_10_1000		2,173	21,728
sag_25_250		4,258	17,031
sag_50_100		9,352	18,704
sol_5_100		1,647	32,945
sol_10_500		2,372	23,717
sol_25_1000		4,692	18,769
sol_50_500		9,514	19,027

EK-Ç: Etik Beyanı

Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmada,

- tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- görsel, işitsel ve yazılı bütün bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- atıfta bulunduğum eserlerin bütününe kaynak olarak gösterdiğimi,
- kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- bu tezin herhangi bir bölümünü bu üniversitede veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

24/09/2018

A. Bayhan

Ayşe BAYHAN

EK-D: Yüksek Lisans Tez Çalışması Orijinallik Raporu

24/09/2018

HACETTEPE ÜNİVERSİTESİ

Eğitim Bilimleri Enstitüsü

Eğitimde Ölçme ve Değerlendirme Ana Bilim Dalı Başkanlığına,

Tez Başlığı : Farklı Koşullardaki Kayıp Veri Oranının İç Tutarlığa Etkisi
Yukarıda başlığı verilen tez çalışmamın tamamı (kapak sayfası, özetler, ana bölümler, kaynakça) aşağıdaki filtreler kullanılarak **Turnitin** adlı intihal programı aracılığı ile kontrol edilmiştir. Kontrol sonucunda aşağıdaki veriler elde edilmiştir:

Rapor Tarihi	Sayfa Sayısı	Karakter Sayısı	Savunma Tarihi	Benzerlik Oranı	Gönderim Numarası
15/05/2018	70	42661	28/6/2018	%10	964131979

Uygulanan filtreler:

1. Kaynaklar hariç
2. Alıntılar dâhil
3. 5 kelimedenden daha az örtüşme içeren metin kısımları hariç

Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü Tez Çalışması Orijinallik Raporu Alınması ve Kullanılması Uygulama Esasları'nı inceledim ve çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan eder, gereğini saygılarımla arz ederim.

Ad Soyadı: Ayşe BAYHAN

Öğrenci No.: N11228665

Ana Bilim Dalı: Eğitimde Ölçme ve Değerlendirme

Programı: Tezli Yüksek Lisans

Statüsü: Y.Lisans Doktora Bütünleşik Dr.

A. Bayhan

DANIŞMAN ONAYI

UYGUNDUR

(Prof. Dr. Nuri DOĞAN, İmza)

EK-E: Thesis Originality Report

24/09/2018

HACETTEPE UNIVERSITY
Graduate School Of Educational Sciences
To The Department Of Measurement and Evaluation in Education

Thesis Title : The Effect Of Missing Data Rate On Internal Consistency Within Different Conditions

The whole thesis that includes the *title page, introduction, main chapters, conclusions and bibliography section* is checked by using **Turnitin** plagiarism detection software take into the consideration requested filtering options. According to the originality report obtained data are as below.

Time Submitted	Page Count	Character Count	Date of Thesis Defense	Similarity Index	Submission ID
15/05/2018	70	42661	28/6/2018	%10	964131979

Filtering options applied:

1. Bibliography excluded
2. Quotes included
3. Match size up to 5 words excluded

I declare that I have carefully read Hacettepe University Graduate School of Educational Sciences Guidelines for Obtaining and Using Thesis Originality Reports; that according to the maximum similarity index values specified in the Guidelines, my thesis does not include any form of plagiarism; that in any future detection of possible infringement of the regulations I accept all legal responsibility; and that all the information I have provided is correct to the best of my knowledge.

I respectfully submit this for approval.

Name Lastname: Ayşe BAYHAN

Student No.: N11228665

Department: Measurement and Evaluation in Education

Program: Masters with thesis

Status: Masters Ph.D. Integrated Ph.D.

A. Bayhan

ADVISOR APPROVAL


APPROVED
(Prof. Dr. Nuri DOĞAN, Signature)

EK-F: Yayınlama ve Fikrî Mülkiyet Hakları Beyanı

Enstitü tarafından onaylanan lisansüstü tezimin/raporumun tamamını veya herhangi bir kısmını, basılı (kâğıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe Üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanılması zorunlu metinlerin yazılı izin alınarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Yükseköğretim Kurulu tarafından yayınlanan "**Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına ilişkin Yönerge**" kapsamında tezim aşağıda belirtilen koşullar haricince YÖK Ulusal Tez Merkezi / H.Ü. Kütüphaneleri Açık Erişim Sisteminde erişime açılır.

- o Enstitü/Fakülte yönetim kurulu kararı ile tezimin erişime açılması mezuniyet tarihinden itibaren 2 yıl ertelenmiştir. ⁽¹⁾
- o Enstitü/Fakülte yönetim kurulunun gerekçeli kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren ... ay ertelenmiştir. ⁽²⁾
- o Tezimle ilgili gizlilik kararı verilmiştir. ⁽³⁾

24/09 /2018


Ayşe BAYHAN

"Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge"

- (1) Madde 6. 1. Lisansüstü teze ilgili patent başvurusu yapılması veya patent alma sürecinin devam etmesi durumunda, tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü Üzerine enstitü veya fakülte yönetim kurulu iki yıl süre ile tezin erişime açılmasının ertelenmesine karar verebilir.
- (2) Madde 6.2. Yeniteknik, materyal ve metotların kullanıldığı, henüz makaleye dönüşmemiş veya patent gibi yöntemlerle korunmamış ve internette paylaşılması durumunda 3. şahıslara veya kurumlara haksız kazanç; imkânı oluşturabilecek bilgi ve bulguları içeren tezler hakkında tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakültenin uygun görüşü Üzerine üniversite yönetim kurulu tarafından verilen karar ile altı ayı aşmamak üzere tezin erişime açılması engellenebilir.
- (3) Madde 7. 1. Ulusal çıkarları veya güvenliği ilgilendiren, emniyet, istihbarat, savunma ve güvenlik, sağlık vb. konulara ilişkin lisansüstü tezlerle ilgili gizlilik kararı, tezin yapıldığı kurum tarafından verilir*. Kurum ve kuruluşlarla yapılan işbirliği protokolü çerçevesinde hazırlanan lisansüstü tezlere ilişkin gizlilik kararı ise, ilgili kurum ve kuruluşun önerisi ile enstitü veya fakültenin uygun görüşü Üzerine üniversite yönetim kurulu tarafından verilir. Gizlilik kararı verilen tezler Yükseköğretim Kuruluna bildirilir.
Madde 7.2. Gizlilik kararı verilen tezler gizlilik süresince enstitü veya fakülte tarafından gizlilik kuralları çerçevesinde muhafaza edilir, gizlilik kararının kaldırılması halinde Tez Otomasyon Sistemine yüklenir

* Tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu tarafından karar verilir.

