



T.C.
AKDENİZ ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
EĞİTİM BİLİMLERİ ANA BİLİM DALI

YÜKSEK
LİSANS TEZİ

EĞİTİM ARAŞTIRMALARINDA
KAYIP VERİ DURUMUNDA FARKLI
KAPPA KATSAYILARININ
İNCELENMESİ

HALİME TOPBAŞ

EĞİTİMDE ÖLÇME VE DEĞERLENDİRME
TEZLİ YÜKSEK LİSANS PROGRAMI

Antalya, 2022

AKDENİZ ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
EĞİTİM BİLİMLERİ ANA BİLİM DALI
ÖLÇME VE DEĞERLENDİRME TEZLİ YÜKSEK
LİSANS PROGRAMI

EĞİTİM ARAŞTIRMALARINDA KAYIP VERİ
DURUMUNDA FARKLI KAPPA KATSAYILARININ
İNCELENMESİ

YÜKSEK LİSANS TEZİ

Hazırlayan
Halime TOPBAŞ

Tez Danışmanı
Doç. Dr. BİLAL BARIŞ ALKAN

Antalya, 2022

DOĞRULUK BEYANI

Yüksek lisans tezi olarak sunduğum bu çalışmayı, bilimsel ahlak ve geleneklere aykırı düşecek bir yol ve yardıma başvurmaksızın yazdığımı, yararlandığım eserlerin kaynakçalardan gösterilenlerden oluştuğunu ve bu eserleri her kullanışında alıntı yaparak yararlandığımı belirtir; bunu onurumla doğrularım. Enstitü tarafından belli bir zamana bağlı olmaksızın, tezimle ilgili yaptığım bu beyana aykırı bir durumun saptanması durumunda, ortaya çıkacak tüm ahlaki ve hukuki sonuçlara katlanacağımı bildiririm.

10/06/2022

Halime TOPBAŞ

TEŞEKKÜR

Tüm ders ve tez dönemlerim boyunca akademik bilgi ve birikiminden faydalanmamı sağlayan, sosyal ve eğitsel manada gereken tüm fırsat ve imkanları sunan değerli tez danışmanım Doç. Dr. Bilal Barış ALKAN' a, sonrasında ise öğrenim hayatımın tüm kademelerinde gerekli özveri ve desteklerini hiç esirgemedi sunan annem ve babama, bu süreçte yardımları ve ilgisiyle destekçim olan kız kardeşim Zübeyde TOPBAŞ' a sevgi, saygı ve şükranlarımı iletmek isterim.



ÖZET

EĞİTİM ARAŞTIRMALARINDA KAYIP VERİ DURUMU İÇİN FARKLI KAPPA KATSAYILARININ İNCELENMESİ

TOPBAŞ, Halime

Yüksek Lisans, Eğitim Bilimleri Ana Bilim Dalı,

Tez Danışmanı: Doç. Dr. Bilal Barış Alkan

Haziran 2022, 50 sayfa

Bu çalışmada puanlayıcılar arası uyumu incelemek için araştırmalarda yaygın olarak kullanılan Cohen Kappa, Gwet Kappa, Silme Yöntemli Kappa, Fleiss Kappa, Light Kappa ve çalışma kapsamında önerilen Gwet Kappa'ya dayalı Light Kappa yöntemlerinin kayıp veri durumunda nasıl değişim göstereceğinin incelenmesi amaçlanmaktadır. Çalışmada, farklı kayıp oranlarında (%5, %10, %15, %20, %25, %30) MCAR yöntemi ile kayıp değer ataması yapılarak elde edilen farklı senaryolar için iki ve daha fazla puanlayıcı durumlarında uyum tabloları oluşturulmuş ve kappa katsayıları hesaplanmıştır. Elde edilen bulgulara göre karşılaştırmalar yapılmış ve kayıp değer durumunda hangi kappa türünün daha iyi sonuç verdiği incelenmiştir. İki puanlayıcı olduğunda, katılımcı sayısı 10000, kayıplık oranı %30 Gwet Kappa katsayısının daha iyi sonuç verdiği gözlemlenmiştir. Kayıp veri durumunda Liste bazında silmeye dayanan kappa sık tercih edilen bir yöntemdir ancak veri kaybına yol açtığından Gwet Kappa kullanılması önerilir. Gwet Kappa kayıp değerleri ayrı bir kategoride hesaplamaya dahil ettiğinden sonuçlar daha anlamlı çıkmıştır. İkiden fazla puanlayıcı olması durumunda, katılımcı sayısı 100, kayıplık oranları %15 ve %20 olduğunda; katılımcı sayısı 1000, kayıplık oranı %5 olduğunda; katılımcı sayısı 10000, kayıplık oranları %5 ve %10 olduğunda çalışma kapsamında önerilen Gwet Kappa'ya Dayalı Light Kappa yaklaşımının daha iyi performans gösterdiği görülmüştür. Gerçek veri kümelerinden elde edilen bulgulara göre 3 puanlayıcı için Fleiss Kappa, Light Kappa ve Gwet Kappa'ya Dayalı Light Kappa hesaplamalarında çalışma kapsamında önerilen Gwet Kappa'ya Dayalı Light Kappa yaklaşımının tüm kayıplık oranlarında da oldukça iyi olduğu söylenebilir.

Anahtar Sözcükler: Kayıp değer, Cohen Kappa, Gwet Kappa, Fleiss Kappa, Light Kappa, Gwet Kappa'ya dayalı Light Kappa

ABSTRACT

INVESTIGATION OF DIFFERENT KAPPA COEFFICIENTS FOR MISSING DATA STUATION IN EDUCATIONAL RESEARCH

TOPBAŞ, Halime

Master's Degree, Department of Educational Sciences

Supervisor: Assoc. Prof. Dr. Bilal Barış Alkan

June 2022, 50 pages

In this study, it is aimed to examine how the Cohen Kappa, Gwet Kappa, Listwise Deletion Kappa, Fleiss Kappa, Light Kappa and Light Kappa methods based on Gwet Kappa, which are widely used in research to examine the agreement between raters in the case of missing data. In the study, for different scenarios obtained by assigning missing values with the MCAR method at different missing rates (5%, 10%, 15%, 20%, 25%, 30%), agreement tables were created for two or more rater cases, and the kappa coefficients were calculated. Comparisons were made according to the findings and it was examined which kappa type gave better results in case of missing value. It has been observed that when there are two raters, the number of participants is 10000 and the missing rate 30% Gwet Kappa coefficient gives better results. In case of missing data, Kappa based on Listwise deletion is a frequently preferred method, but it is recommended to use Gwet Kappa as it causes data loss. Since Gwet Kappa included the missing values in the calculation in a separate category, the results were more significant. In case of more than two raters, when the number of participants is 100, and the loss rates are 15% and 20%; when the number of participants is 1000 and the missing rate is 5%; When the number of participants was 10000, the loss rates were 5% and 10%, it was seen that the Light Kappa approach based on the Gwet Kappa suggested in the study performed better. According to the findings obtained from the real datasets, it can be said that the Light Kappa Based on Fleiss Kappa, Light Kappa and Gwet Kappa calculations for 3 raters, proposed within the scope of the study, are quite good in all missing rates as well.

Keywords: *Missing data, Cohen Kappa, Gwet Kappa, Fleiss Kappa, Light Kappa, Light Kappa based on Gwet Kappa*

İÇİNDEKİLER

DOĞRULUK BEYANI	iii
TEŞEKKÜR.....	i
ÖZET	ii
ABSTRACT	iii
TABLolar LİSTESİ	vii
ŞEKİLLER LİSTESİ	viii
SİMGELER VE KISALTMALAR LİSTESİ	ix

BÖLÜM I

GİRİŞ

1.1. Problem Durumu.....	1
1.2. Araştırmanın Amacı.....	2
1.2.1. Problem Cümlesi	2
1.3. Araştırmanın Önemi.....	2
1.4. Tanımlar	3

BÖLÜM II

KAVRAMSAL ÇERÇEVE İLE İLGİLİ ARAŞTIRMALAR

2.1. Kayıp Veri.....	4
2.2. Kayıp Veri Mekanizmaları.....	5
2.3. Cohen Kappa.....	8
2.3.1. Gwet Kappa	10
2.3.2. Liste Bazında Silmeye Dayanan Kappa	11
2.4. Fleiss Kappa Katsayısı	11

2.5. Light Kappa.....	14
2.6. İlgili Çalışmalar.....	14
2.6. R Yazılımı	17

BÖLÜM III

YÖNTEM

3.1. Araştırmanın Modeli	19
3.1.1. Araştırmanın Materyali	19
3.2. Önerilen Yaklaşım: Kayıp Veri Durumunda Gwet Kappa’ya Dayalı Light Kappa (GKDLK).....	21
3.3. Veri Analizi Aracı.....	22
3.4. Veri Analizi Süreci.....	22
3.4.1. “irr” Paketi İçerisinde Bulunan Veri Kümelerinin Analize Hazırlanması ..	22
3.4.2. Gwet Kappa’ya Dayalı Light Kappa Hesabı için Veri Kümelerinin Analize Hazırlanması.....	22

BÖLÜM IV

BULGULAR

4.1. Tam Veri Durumu için Üretilen Veri Kümesinden Elde Edilen Bulgular.....	23
4.2. Kayıp Veri Durumu için Üretilen Veri Kümesinden Elde Edilen Bulgular	24
4.3. ‘diagnoses’ ve ‘anxiety’ Veri Kümelerinde Kappa Katsayısı Hesabından Elde Edilen Bulgular	26

BÖLÜM V

SONUÇ, TARTIŞMA VE ÖNERİLER

5.1. Sonuç ve Tartışma.....	28
5.2. Öneriler	30
KAYNAKÇA.....	31
EKLER	34
Ek-1: Veri Üretme ve Karşılaştırmaya İlişkin R Kodları	34
ÖZGEÇMİŞ	36
BİLDİRİM.....	37
İNTİHAL RAPORU	38

TABLÖLAR LİSTESİ

Tablo 2.1. Üç Kategoriye Ayrılmış İkili Sınıflandırma	8
Tablo 2.2. Kappa Değerleri Yorumu (Landis and Koch,1977)	9
Tablo 2.3. Kayıp Değerlendirme için Bir Kategori ve Üç Genel Kategoriye Ayrılmış İkili Sınıflandırma	10
Tablo 2.4. Fleiss Kappa Katsayısını Hesaplamak İçin Değerlendirme Sonuçları	13
Tablo 3. 1. Üç Puanlayıcı Tarafından Sınıflandırılmış ‘diagnoses’ Veri Kümesi	20
Tablo 3. 2. Üç Puanlayıcı Tarafından Sınıflandırılan ‘anxiety’ Veri Kümesi	21
Tablo 4.1. Farklı Veri Setlerinden Elde Edilen Uyum Oranları	23
Tablo 4.2. Farklı Senaryolara Göre Oluşturulan Veri Kümelerine Kayıp Veri Ataması Yapıldıktan Sonra Elde Edilen Sonuçlar (2 Puanlayıcı Durumu)	24
Tablo 4. 3. Farklı Senaryolara Göre Oluşturulan Veri Kümelerine Kayıp Veri Ataması Yapıldıktan Sonra Elde Edilen Sonuçlar (3 Puanlayıcı Durumu)	25
Tablo 4.4. R Yazılımında Bulunan ‘diagnoses’ Veri Kümesine Kayıp Veri Ataması Yapıldıktan Sonra Elde Edilen Sonuçlar	26
Tablo 4.5. R Yazılımında Bulunan ‘anxiety’ Veri Kümesine Kayıp Veri Ataması Yapıldıktan Sonra Elde Edilen Sonuçlar	27

ŞEKİLLER LİSTESİ

Şekil 2.1.Okul Öncesi Öğrencisine Öğretmeni Tarafından Verilen Resim.....	6
Şekil 2.2. Birinci Durum	6
Şekil 2.3. İkinci Durum	7
Şekil 2.4. Üçüncü Durum	7



SİMGELER VE KISALTMALAR LİSTESİ

κ : Kappa Katsayısı

MCAR: Tamamen Rastgele Kayıplık

MAR: Rastgele Kayıplık

MNAR: Rastgele Olmayan Kayıplık

MSE: Ortalama Karesel Hata

P_0 : Cohen'in Kappası Gözlemlenen Uyumun Yüzdesi

P_e : Cohen'in Kappası Beklenen Uyumun Yüzdesi

P_{gu} : Gwet'in Kappası Gözlemlenen Uyumun Yüzdesi

P_{bu} : Gwet'in Kappası Beklenen Uyumun Yüzdesi

κ_g : Gwet'in Kappa Katsayısı

BÖLÜM I

GİRİŞ

Araştırmanın bu bölümünde problem durumu, araştırmanın amacı ve problemlerine yer verilmiştir.

Eğitim araştırmalarında ve uygulamalarında, birimlerin (kişiler, nesneler) sınıflanmasına ve değerlendirilmesine sık sık ihtiyaç duyulur. Örneğin; Öğrencilerin öğretmenleri tarafından yeterliklerini değerlendirmek için ödevlerin sınıflandırılması, Ruh sağlığı sorunları olan kişilerin bir psikolog tarafından zihinsel bozukluk sınıflarına tahsis edilmesi, Okul yöneticilerinin eğitim durumu, mesleki kıdem, yaş değişkenleri gibi faktörlere göre sınıflandırılması gibi. Yapılan bu sınıflandırmalar ve değerlendirmelerde doğru ve güvenilir sonuçlar elde etmek önemlidir. Üst düzey becerilerin ölçülmesinde elde edilen puanların tutarlılığı hakkında sağlıklı yorumlar yapmak her zaman mümkün değildir. Bu anlamda puanlayıcılar arası güvenilirliği ölçmek için kullanılan istatistikler vardır. Cohen Kappa, Fleiss Kappa, Light Kappa bunlara örnektir. Hesaplamaya dahil edilen kappa katsayıları tam veri durumunda iyi sonuç vermektedir ancak puanlayıcıların bazı puanlamaları eksik bıraktığı durumlarda kayıp değerler oluşmaktadır. Puanlayıcıların değerlendirmelerinden biri veya her ikisi eksikse kayıp kabul edilir. Kayıp değerler puanlayıcılar arası uyumu belirlemekte sorunlara yol açmaktadır. Kayıp değerlerin yer aldığı veri kümesinde, puanlayıcılar arası uyumun belirlenmesinde kullanılacak yöntemin seçilmesi üzerine çalışmalar son yıllarda artış göstermektedir.

1.1. Problem Durumu

Kategorilere göre yapılan sınıflandırmalar araştırma amaçları için de yararlı olabilir. Kategorilerin tanımı belirsizse veya talimatlar net değilse puanlamaların güvenilirliği düşük olabilir. Değerlendirmelerin doğru ve yüksek güvenilirlikte olması için araştırmacılar genellikle puanlayıcıların aynı birim grubunu değerlendirmelerini istemektedir. Puanlayıcılar arasındaki

uyum, güvenilirliğin bir göstergesi olarak kullanılır. (Blackman ve Koval, 2000; McHugh, 2012; Shiloach ve diğerleri, 2010; Wing ve diğerleri, 2002).

İki eksik değerlendirmeye sahip veya tek bir değerlendirme eksik olan birimler “kayıp” kategorisinde sınıflandırılır. Ölçme ve değerlendirme çalışmalarında puanlayıcılar arası uyum incelenirken kayıp değer sorunu ile karşılaşilmektedir. Kayıp değerler araştırmalarda oldukça yaygındır ve verilerden çıkarılabilecek sonuçlar üzerinde dikkate değer bir etkisi olabilir. Kayıp değerleri ihmal etmek bir alternatiftir ancak çoğu zaman hatalı çıkarımlara yol açabilir. Bu çalışmada kayıp değer durumunda, farklı kappa katsayılarının nasıl ilişkili olduğunu, kayıplığın derecesi ve güvenilirlik üzerinde etkisi araştırılmıştır.

1.2. Araştırmanın Amacı

Eğitim araştırmalarında sık sık kayıp verilerle karşılaşilmektedir. Değerlendirmelerin doğru ve yüksek güvenilirlikte olması araştırmacılar için önem teşkil etmektedir. Kayıp veriyi ihmal etmek çoğu zaman bilgi kaybına yol açtığından tercih edilmez. Bu çalışmada eğitim araştırmalarında puanlayıcılar arası uyumu veren kappa katsayılarının incelenmesi amaçlanmaktadır. Kayıp veri durumunda kappa katsayılarının nasıl hesaplanacağı ve kayıp veri sorununa nasıl çözüm getirileceği ile ilgili çalışmalar yapılacaktır. Kayıp verilerle başa çıkabilmek için farklı stratejiler ve çözüm yolları aranacaktır.

1.2.1. Problem Cümlesi

Eğitim araştırmalarında kayıp veri durumunda puanlayıcılar arası uyumun belirlenmesinde kullanılan farklı kappa katsayılarının performansları arasında farklılık var mıdır?

1.3. Araştırmanın Önemi

Eğitim araştırmalarında kayıp değer sorunu ile oldukça sık karşılaşilmektedir. Alan yazındaki çalışmalar incelendiğinde ikiden fazla puanlayıcı olması durumunda, kayıp değer varlığında nasıl bir yöntem kullanılacağı üzerine fikir birliği yoktur ve çalışmalar halen devam etmektedir. Bu bağlamda, çalışmanın ikiden fazla puanlayıcı olması durumunda, kayıp veri varlığında kullanılacak yeni bir yaklaşım önerdiği düşünüldüğünde alan yazındaki ilgili literatüre katkı sağlayacağından araştırmanın önemi büyüktür.

1.4. Tanımlar

MCAR (Missingness Completely At Random): Herhangi bir veri üyesinin eksik olmasına neden olan olaylar hem gözlemlenebilir değişkenler hem de gözlemlenemeyen ilgili parametrelerden bağımsızsa ve rastgele gerçekleşirse, bu veri kümesindeki değerler tamamen rastgele eksiktir.

MAR (Missing At Random): Eksiklik rastgele olmadığında ancak eksikliğin değişkenler tarafından tamamen açıklanabildiği durumlarda oluşur.

MNAR (Missingness Not A Random): Ne MAR ne de MCAR durumunda olmayan verilerdir.



BÖLÜM II

KAVRAMSAL ÇERÇEVE VE İLGİLİ ARAŞTIRMALAR

Bu bölümde araştırmanın konusu ile ilgili yurtiçi ve yurtdışındaki akademik çalışmalar yer verilmiştir.

2.1. Kayıp Veri

İstatistiksel çalışmalarda kayıp veri ile çok sık karşılaşılmaktadır. Kayıp veri sorunu, istatistiksel analizlerde önemli bir tartışma alanıdır. Allison (2002) ve Graham'a (2009) göre araştırmacılar tarafından kullanılan analitik süreçlerin, tam veri kümeleri ile yapılandırılmış olması bu tartışmalara sebep olmuştur. Yapılan bu çalışmalar kayıp veri durumu için bir çözüm mekanizması üretmemiştir.

Matris şeklinde hazırlanan veri setlerinde satırlar gözlemleri, sütunlar ise değişkenleri temsil etmektedir. Kayıp veri, veri kümesindeki herhangi bir hücrede değer olmamasıdır. Bir gözlemi temsil eden hücre eksik kalırsa kayıp veri oluşur. Kayıp veri çeşitli nedenlerle ortaya çıkabilir. Kayıp veri durumu çalışmayı yürüten kişiden, veri toplanan kişinin bilerek cevap vermemesinden, eksik bilgilendirmeden, verilerin gözlenememesinden, soruların anlaşılamaması gibi sebeplerden ortaya çıkabilir. Bu çalışmada, her iki puanlayıcı tarafından sınıflandırılmamış veya yalnızca bir puanlayıcı tarafından sınıflandırılmışsa verilerin kayıp olduğu düşünülmektedir.

Araştırmacılar için kayıp verilerle çalışmak sorunlara yol açabilmektedir. Bu durum araştırmacıları tam veri setleri üzerinde çalışmaya yönlendirmektedir. Bazı araştırmalarda eksiksiz veri kümeleri elde edilebilmesi mümkün olmadığından araştırmacılar genellikle, kayıp verilerin 'ihmal edilebilir (ignorable)' olduğunu düşünmektedir ve verileri analiz dışında bırakmaktadırlar (Rubin, 1976; Allison, 2002).

Eğitim araştırmaları içerisinde, nitel yaklaşımla hazırlanan araştırmalar bulunduğu gibi çoğunluk olarak istatistiksel analiz süreçleri içeren araştırmalar olduğu görülmektedir. Bu

durum eğitim arařtırmalarında daha çok istatistiksel analizler gerektiren veri kümeleri ile çalışıldığını göstermek için yeterli değildir. İstatistiksel analiz süreçleri içeren arařtırmalarda genellikle tam veri kümesi üzerinde çalışılmaktadır. Eksiksiz bir veri kümesi elde etmek, maliyetli ve ekstra çaba gerektiren zor bir iřtir. Yanlılık ve hatalı kestirim risklerinin en aza indirilmesi için de tam veri kümesi ile çalışmak avantajlıdır. Arařtırmalarda kayıp veri sorunu bulunduđu kesinleşen az sayıda arařtırma tespit edilebilmiştir. Bu arařtırmaların büyük çoğunluğunda kayıp veri mekanizmasının dikkate alınmadığı ve herhangi bir kayıp veri yönteminin kullanılmadığı görülmüřtür

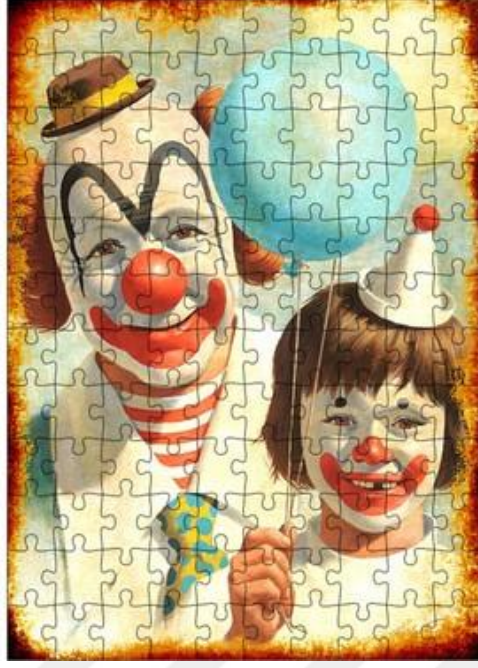
2.2. Kayıp Veri Mekanizmaları

Kayıp veri mekanizması, analiz yöntemlerinden uygun olanının belirlenmesinde ve sonuçlarının yorumlanmasında anahtar rol oynamaktadır. Kayıp veri mekanizmaları genellikle

Tamamen Rastgele Eksik Olan Veriler (Missing Completely at Random - MCAR): Herhangi bir veri kümesinin eksik olmasına neden olan olaylar hem gözlemlenebilir hem deęişkenlerden hem de gözlemlenemeyen ilgili parametrelerden bağımsızsa ve tamamen rastgele gerçekleşirse, deęerler tamamen rastgele (MCAR) eksiktir. Veriler MCAR olduđuunda, veriler üzerinde yapılan analiz tarafsızdır; ancak, veriler nadiren MCAR'dır. Rastgele Eksik Olan Veriler (Missing at Random - MAR), eksiklik rastgele olmadığında, ancak eksikliğin tam bilginin bulunduđu deęişkenler tarafından tamamen açıklanabildiği durumlarda oluşur. MAR istatistiksel olarak doęrulanması imkânsız bir varsayımdır. Örneğin, öğrencilerin bir sınav kaygısı anketini doldurma olasılıklarının daha düşük olması, ancak bunun kaygı düzeyleriyle ilgisi yoktur. Rastgele Eksik Olmayan Veriler (Missing Not at Random - MNAR): Tamamen rastgele eksiklik ve rastgele eksiklik olmayan verilerdir. Örnekle açıklamak gerekirse, öğrenciler sınav kaygısı nedeniyle anketi dolduramazlarsa bu gerçekleşen bir durumdur (İnan, S., 2001).

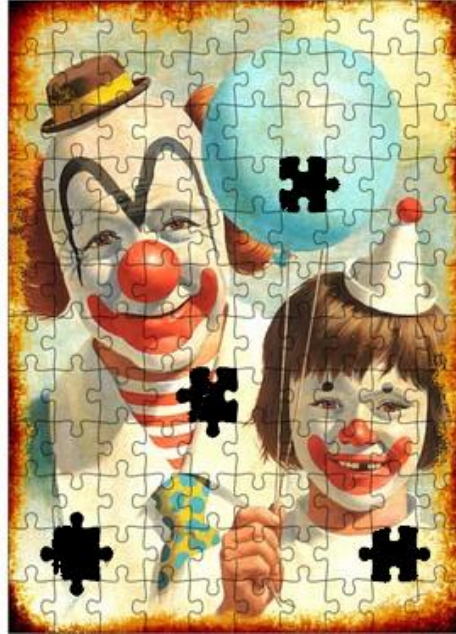
MCAR ile her deęerlendirmenin kayıp olarak etiketlenme şansı eşittir. MNAR ile ise yalnızca ilk kategoriyle ilişkili deęerlendirmelerin kayıp olarak adlandırılmasına izin verilmektedir ve bunların her birinin kayıp olarak yeniden etiketlenme şansı vardır.

Kayıp veri mekanizmalarını görselleřtirmek için bir örnek kurgulanmıştır. 5 yařındaki okul öncesi öğrencisine öğretmeni zihinsel ve kas gelişiminin pekiřmesi amacıyla bir yapboz verdiđi düşünölsün.



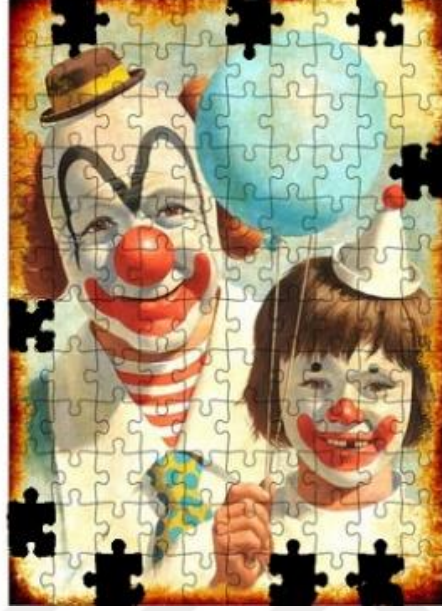
Şekil 2.1.Okul Öncesi Öğrencisine Öğretmeni Tarafından Verilen Resim

Şekil 2.1’de yapboz resminde her parça renklerle tanımlanabilir. Öğrenci birinci durumda yapboz parçalarını bazılarını alır ve eksik parçalar oluşur. Bu durum rasgele parçaları alarak kendiliğinden oluşur.



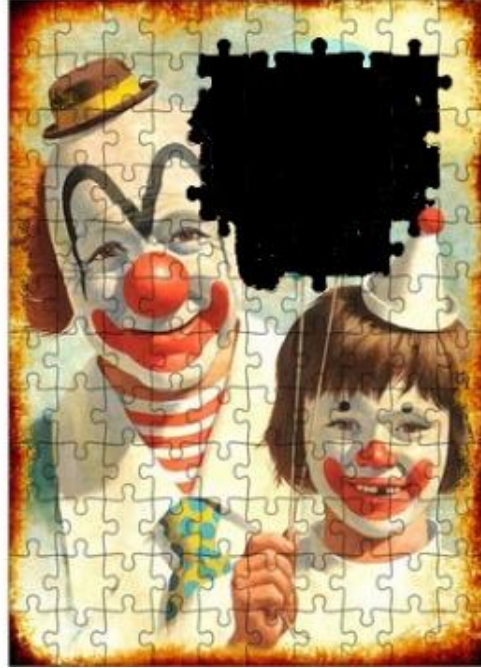
Şekil 2.2. Birinci Durum

İkinci durumda ise yapbozun kenarlarında bulunan parçaları çıkarmak daha kolay olduğundan bu parçaları çıkarır.



Şekil 2.3. İkinci Durum

Üçüncü durumda ise balonla uçabileceğini düşünür. Balon resminin olduğu yapboz parçalarını çıkararak bunu deneyimlemek ister.



Şekil 2.4. Üçüncü Durum

İlk durumda kayıp yapboz parçaları tamamen rastgele olduğundan MCAR durumuna örnektir. İkinci durum MAR ve üçüncü durum rastgele olmayan bir kayıplık olarak MNAR'dır.

2.3. Cohen Kappa

Cohen'in kappa katsayısı iki Puanlayıcı arasındaki karşılaştırmalı uyumun güvenilirliğini ölçen bir istatistik yöntemidir. Kappa istatistiği 1960 yılında Cohen tarafından tanıtıldı. Kappa en yaygın kullanılan anlaşma ölçüsüdür. Puanlayıcılar arasında uyum kavramını belirlemek için, sütunların farklı puanlayıcıları temsil ettiği ve satırların puanlayıcıların veri topladığı değişkenleri temsil ettiği bir matris oluşturuldu (Tablo 1). Aynı N grubunu, önceden tanımlanan k kategorilerinden birine bağımsız olarak sınıflandıran A ve B olmak üzere iki puanlayıcı olduğunu varsayalım.

Tablo 2.1

Üç Kategoriye Ayrılmış İkili Sınıflandırma

Puanlayıcı A	Puanlayıcı B			Toplam
	Kategori 1	Kategori 2	Kategori 3	
Kategori 1	p_{11}	p_{12}	p_{13}	p_{1+}
Kategori 2	p_{21}	p_{22}	p_{23}	p_{2+}
Kategori 3	p_{31}	p_{32}	p_{33}	p_{3+}
Toplam	p_{+1}	p_{+2}	p_{+3}	

Tablo 2.1'deki p_{11} , p_{22} ve p_{33} , puanlayıcılar arasındaki uyumu gösterir. Diğer hücreler puanlayıcılar arasındaki anlaşmazlığı yansıtır. Kappa katsayısı (κ) iki niceliğin bir fonksiyonudur:

- (i) Gözlemlenen uyumun yüzdesi: Her iki puanlayıcının üzerinde anlaştığı öğelerin oranıdır.

$$P_0 = \sum_{i=1}^k p_{ii} \quad (1)$$

- (ii) Beklenen uyumun yüzdesi: Sınıflandırmaların istatistiksel bağımsızlığı altında gözlenen uyumun değeridir.

$$P_e = \sum_{i=1}^k p_{i+} p_{+i} \quad (2)$$

Kappa katsayısı;

$$\kappa = \frac{P_o - P_e}{1 - P_e} \quad (3)$$

Katsayının maksimum değerinin 1 olmasını sağlamak için, $P_o - P_e$, $1 - P_e$ 'nin maksimum değerine bölünür. Kappa'nın değeri genellikle 0 ile 1 arasındadır. Landis ve Koch (1977), kappa değerinin yorumlanması için Tablo 2.2'deki yönergeleri önerdiler;

Tablo 2.2

Kappa Değerleri Yorumu (Landis ve Koch,1977)

Kappa Değerleri	Uyum Düzeyi
.0 ile .2	Çok az uyum
.2 ile .4	Az uyum
.4 ile .6	Orta uyum
.6 ile .8	İyi uyum
.8 ile 1	Çok iyi uyum

Bu yönergelerin ve diğer herhangi bir yönerge kümesinin genellikle keyfi kabul edildiğine dikkat edilmelidir. Alanyazında farklı kappa katsayısı yorumlarına rastlanmaktadır. Kappa değerlerinin yorumlanması uygulama alanlarında farklı anlamlara sahip olabilir.

Gözlemlenen uyumun yüzdesinin yüksek olmasına rağmen, satır x sütun tablolarındaki büyük dengesizlikler nedeniyle küçük bir kappa katsayısının elde edilmesi, Satır-Sütun toplamalarında asimetrik dengesizlik olduğunda kappa katsayısının simetrik dengesizlik olması durumuna göre daha yüksek bulunabilmesi kappa istatistiğinin sınırlılıklarındandır.

Cohen'in kappasının kayıp veri durumunda hesaplamalar yapılmasını sağlayan farklı versiyonları vardır. Bu çalışmada, Gwet Kappa, Liste Bazında Silmeyi temel alan kappa ele alınmıştır.

2.3.1. Gwet Kappa

Gwet Kappa, Cohen kappanın kayıp verileri işleyebilen bir türünü sunmaktadır. Kayıp verileri ayrı bir kategoride değerlendirerek hesaplamaya dahil etmektedir. Tablo 2.3'de üç kategoriye ayrılmış ve 1 eksik değerlendirmeye sahip ikili sınıflandırma gösterilmektedir. Puanlayıcı A ve puanlayıcı B tarafından değerlendirilen tabloda p_{11} , p_{22} , p_{33} puanlayıcılar arası uyumu göstermektedir. Eksik kategorisinde belirtilen p_{m1} , p_{m2} , p_{m3} , p_{m+} , puanlayıcı B tarafından değerlendirilen ancak puanlayıcı A tarafında eksik değerlendirmeye sahip olduğunu, p_{1m} , p_{2m} , p_{3m} , p_{+m} değerleri ise puanlayıcı A tarafından değerlendirilen ancak puanlayıcı B tarafından eksik değerlendirmeye sahip olduğunu göstermektedir.

Tablo 2.3

Kayıp Değerlendirme için Bir Kategori ve Üç Genel Kategoriye Ayrılmış İkili Sınıflandırma

Puanlayıcı A	Puanlayıcı B				Toplam
	Kategori 1	Kategori 2	Kategori 3	Eksik	
Kategori 1	p_{11}	p_{12}	p_{13}	p_{1m}	p_{1+}
Kategori 2	p_{21}	p_{22}	p_{23}	p_{2m}	p_{2+}
Kategori 3	p_{31}	p_{32}	p_{33}	p_{3m}	p_{3+}
Eksik	p_{m1}	p_{m2}	p_{m3}	p_m	p_{m+}
Toplam	p_{+1}	p_{+2}	p_{+3}	p_{+m}	1

Gwet'in formülasyonunda yalnızca iki derecelendirmeye sahip birimler dahil edilmiştir.

(i) Gözlemlenen uyumun yüzdesi;

$$P_{gu} = \frac{\sum_{i=1}^k p_{ii}}{\sum_{i=1}^k \sum_{j=1}^k p_{ij}} \quad (4)$$

(ii) Beklenen uyumun yüzdesi;

$$P_{bu} = \frac{\sum_{i=1}^k p_{i+i}}{(1-P_{m+})(1-P+m)} \quad (5)$$

Gwet'in kappa katsayısı;

$$\kappa_g = \frac{P_{gu}-P_{bu}}{1-P_{bu}} \quad (6)$$

2.3.2. Liste Bazında Silmeye Dayanan Kappa

Kayıp verilerle baş etmenin üçüncü bir yolu, her iki puanlayıcı tarafından sınıflandırılmayan tüm birimleri silmek (veya yok saymak) ve klasik Cohen'in kappasını uygulamaktır (Strijbos Stahl, 2007). Bu nedenle, bu yaklaşıma dayanan kappa, liste bazında silmeye dayanan kappa olarak adlandırılmaktadır.

2.4. Fleiss Kappa Katsayısı

İki puanlayıcı arasındaki uyumu ele alırken Cohen'in kappa katsayısı kullanılmaktadır. Uyumun ölçüldüğü puanlayıcı sayısı ikiden fazla ise Fleiss'in kappa katsayısı kullanılmalıdır. İkiden fazla puanlayıcı olması veya değişkenlerin ikiden fazla seviyesi olma durumlarında kullanılır. Bu nedenle Cohen'in kappasından daha geniş kullanım alanı vardır. Fleiss kappa katsayısı da nominal değişkenlerde kullanılabilir. Fleiss kappa katsayısının hesaplanma adımları:

Adım 1. N değerlendirilecek durumların sayısı; n toplam puanlayıcı sayısı ve k ise değerlendirmede kullanılacak kategori sayısı olsun. Önce j-inci kategoriye ait değerlendirme oranı P_j olarak hesaplanır.

$$p_j = \frac{1}{N_n} \sum_{i=1}^N n_{ij} \quad 1 = \frac{1}{n} \sum_{j=1}^k n_{ij} \quad (10)$$

Adım 2. Uyum gösteren p_i değerleri hesaplanır.

$$p_i = \frac{1}{n(n-1)} \sum_{j=1}^k n_{ij}(n_{ij} - 1) \quad (11)$$

$$= \frac{1}{n(n-1)} (\sum_{j=1}^k (n_{ij}^2 - n_{ij})) \quad (12)$$

$$= \frac{1}{n(n-1)} (\sum_{j=1}^k n_{ij}^2) \quad (13)$$

Adım 3. Elde edilen P_i uyum değerlerinin ortalaması, $\bar{P}$ hesaplanır.

$$\bar{P} = \frac{1}{N} \sum_{i=1}^N p_i \quad (14)$$

$$= \frac{1}{Nn(n-1)} (\sum_{i=1}^N \sum_{j=1}^k n_{ij}^2 - Nn) \quad (15)$$

Adım 4. Son olarak, toplam ortalama olan $\bar{P}_e$ ve Fleiss 'in kappa katsayısı hesaplanır.

$$\bar{P}_e = \sum_{j=1}^k p_j^2 \quad (16)$$

$$\kappa_F = \frac{\bar{P} - \bar{P}_e}{1 - \bar{P}_e} \quad (17)$$

Fleiss kappa katsayısını hesaplanma aşamalarının daha net anlaşılması için bir örnek üzerinden değerlendirmeler yapılmıştır.

Örnek: Bir okulda 14 rehberlik ve psikolojik danışmanın 10 öğrenciyi incelediği düşünölsün. Her rehberlik ve psikolojik danışman her öğrenciyi 5 kategoriden birinde değerlendirmiştir. Değerlendirme sonuçları Tablo 2.4'de verilmiştir.

Tablo 2.4*Fleiss Kappa Katsayısını Hesaplamak İçin Değerlendirme Sonuçları*

Öğrenci	Kategori					P_i
	1	2	3	4	5	
1	0	0	0	0	14	1.000
2	0	2	6	4	2	0.253
3	0	0	3	5	6	0.308
4	0	3	9	2	0	0.440
5	2	2	8	1	1	0.330
6	7	7	0	0	0	0.462
7	3	2	6	3	0	0.242
8	2	5	3	2	2	0.176
9	6	5	2	1	0	0.286
10	0	2	2	3	7	0.286
Toplam	20	28	39	21	32	
P_j	0.143	0.200	0.279	0.150	0.229	

N= 10, n=14, k=5 olmak üzere tüm gözlemler için toplam 140 ve P_i için toplam 3.780 hesaplanmıştır.

$$p_1 = \frac{(0 + 0 + 0 + 0 + 2 + 7 + 3 + 2 + 6 + 0)}{140} = 0.143$$

İkinci satır için;

$$= \frac{1}{14(14 - 1)} (0^2 + 2^2 + 6^2 + 4^2 + 2^2 - 14) = 0.253$$

$\bar{P}$ değerini hesaplamak için, P_i toplamı bulunur.

$$= 1.000 + 0.253 + + 0.286 = 3.780$$

Formül adımlarına göre Fleiss'in kappa katsayısına ulaşılır.

$$\bar{P} = \frac{1}{((10)((14)(14 - 1)))} ((3.780)(14)(14 - 1)) = 0.378$$

$$\bar{P}_e = 0.143^2 + 0.200^2 + 0.279^2 + 0.150^2 + 0.229^2 = 0.213$$

$$\kappa_F = \frac{0.378 - 0.213}{1 - 0.213} = 0.210$$

Elde edilen sonucun Landis ve Koch (1977) yorumuna göre adil bir uyum olduğu çıkarımı yapılmıştır.

2.5. Light Kappa

Puanlayıcılar arasındaki iki değişkenli kappa katsayılarının tüm mümkün kombinasyonlarının ortalamasına eşittir.

h puanlayıcı sayısı, R düzey sayısı, n gözlem sayısı ve κ_{ij} i . ve j . Puanlayıcılar için hesaplanan Cohen κ katsayısı olmak üzere ($i= 1,2,\dots,h$ ve $j= 1,2,\dots,h$), Cohen κ katsayısının genelleştirilmiş hali olan Light κ katsayısı Eşitlik (18)'de verilmiştir.

$$\kappa_L = \frac{2}{h(h-1)} \sum_{i=1}^{h-1} \sum_{j=i+1}^h \kappa_{ij} \quad (18)$$

Light kappa katsayısı ikili gruplar halinde hesaplanır $h(h - 1) / 2$ tane Cohen'in kappa katsayılarının ortalamasıdır.

2.6. İlgili Çalışmalar

Dalkıran (2006), çalışmasında keman eğitimi yarıyıl sonu sınavlarında yapılan ölçme ve değerlendirme sürecinde mevcut durumu saptamak ve eğitim-öğretimi olumlu yönde etkileyecek geçerliği, güvenirliği ve kullanışlılığı yüksek olan ölçme aracı geliştirmeyi amaçlamıştır. Araştırmayı 15 üniversitenin Eğitim fakültelerine bağlı Güzel sanatlar Eğitimi bölümlerinin Müzik Eğitimi Anabilim Dalında (MEABD) görev yapan 54 keman eğitimcisi ile bu dersi alan 737 keman öğrencisi oluşturmuştur. Ölçme aracında, puanlayıcılar arası tutarlığın belirlenmesi amacı ile Kappa hesabından yararlanılmıştır.

Vanbelle ve Albert (2009), “İki Bağımsız Puanlayıcı Grubu Arasındaki Anlaşma” adlı çalışmasında nominal ölçekle sınıflandırılmış iki bağımsız puanlayıcı grubu arasındaki uyum derecesini değerlendirmek için bir uyum katsayısı önermişlerdir. Bu katsayı, iki puanlayıcı arasındaki uyumu ölçmek için klasik Cohen’in kappa katsayısını genişletmişlerdir. Örneklem varyansı Jackknife yöntemi ile belirlenmiştir. Sonuç olarak, bu çalışmada önerilen grup içi heterojenliği hesaba katarak iki bağımsız puanlayıcı arasındaki uyumu ölçmektedir. Kullanılan yöntem Cohen’in kappa katsayısının doğal bir uzantısıdır ve benzerlikler gösterir.

Bağ ve diğerleri (2010), “2x2 Tablolarda Gözlemciler/Gözlemler Arası Uyumun Değerlendirilmesi” adlı çalışmasında ise diş hekimliğinde yapılan bazı gözlemler ya da gözlemciler arasındaki uyumun belirlenmesi amacına dayanmaktadır. Bu amaçla geliştirilmiş basit uyum katsayısı (p_0), Cohen’in kappa katsayısı (κ) ve PABAK katsayısı kullanılmıştır. Örneklerle bu katsayılar incelenmiştir. Çalışmada incelemeler sonucunda 2 sorun ortaya çıkmıştır. Birinci sorun ‘var’ ve ‘yok’ kategorilerinin göreceli olasılıklarından κ etkilenmiştir ve bu olasılıklar arasındaki fark Prevelans indeksi (PI) olarak adlandırılmıştır. Prevelans indeksi yüksek değerde olduğu zaman, Prevelans indeksinin küçük veya sıfır olması durumuna göre daha düşük çıkmıştır. İkinci sorun Gözlemci A ve B, farklı değerlendirmeler yaptığında yanlışlık oluşmaktadır. Yanlışlık büyük olduğu zaman kappa katsayısı, yanlışlık küçük olduğu veya hiç olmadığı duruma göre daha yüksek çıkmıştır. Sonuç olarak 2x2 tablolarda uyumu göstermek için basit uyum katsayısı yerine, sıklıkla şans düzeltmesi yapan κ katsayısı kullanılmakta, ancak iki sorun nedeniyle yorumlanmasında güçlükler oluşmaktadır.

Demir ve Parlak (2012), çalışmasında Türkiye’de eğitim araştırmalarında kayıp veri sorununa yönelik olarak hangi yöntemlerin daha yaygın bir şekilde kullanıldığının belirlenmesini amaçlamıştır. Nitel bir araştırma olup döküman tarama yöntemi kullanılmıştır. Araştırma kapsamında Türkiye’de dört eğitim dergisinde, 2009-2011 yılları arasında yayımlanan toplam 708 makale, üç araştırmacı tarafından eş zamanlı olarak incelenmiştir. İncelen araştırmaların 405’i, istatistiksel analiz süreçleri içermekle birlikte bu araştırmaların ancak 31’inde kayıp veri sorunu bulunduğu saptanmıştır. 31 araştırmadan 7 tanesinde kayıp veri yöntemi kullanıldığı, fakat bu kullanımın çok da bilinçli olarak yapılmadığı görülmüştür. Araştırmanın sonucunda Türkiye’de eğitim araştırmalarında ve araştırmaların raporlanmasında önemli eksiklikler bulunmuştur. Metodoloji ve kayıp veri yöntemleri konusunda yeterli düzeyinin beklenenin altında olduğu görülmüştür.

Yıldıztekin (2014), çalışmasında Klasik test kuramı ve Genellenebilirlik kuramına göre puanlayıcılar arası tutarlılığı karşılaştırmıştır. Ankara ilindeki bir ortaokulun 7.sınıfında öğrenim gören 84 öğrenciye uygulanan ve problem çözme becerisini ölçen açık uçlu sorular, analitik ve bütünsel dereceleri iki ayrı puanlama anahtarı kullanılarak 5 farklı matematik öğretmeni tarafından puanlanmıştır. Klasik test kuramından, Pearson momentler çarpımı korelasyon katsayısı (PMÇKK), Spearman sıra farkları korelasyon katsayısı (SSFKK), Cronbach Alpha, Kappa ve Krippendorff Alpha katsayıları ile Genellenebilirlik kuramından $b \times m \times p$ çapraz deseninde değişkenlik kaynakları ve yüzdeleri belirlenerek güvenilirlik analizleri yapılmıştır. Puanlayıcılar arası tutarlık düzeyinin yüksek olduğu ve analitik dereceli puanlama anahtarı ile elde edilen puanların tutarlığının bütünsel dereceli puanlama anahtarı ile edilen puanların tutarlığından göreceli olarak daha yüksek olduğu belirlenmiştir.

Kılıç (2015), “Kappa Testi” adlı çalışmasında, gözlemciler arasındaki uyumun şans eseri olabileceğini dikkate alarak kappa hesaplaması yapar. Bu nedenle gözlemciler arası uyumun yüzde hesabından daha güçlü sonuç vereceği belirtilmiştir. Bir klinikte görev yapan iki psikiyatristin 50 hastayı ayrı ayrı değerlendirdiği bir örnek üzerinden kappa değeri hesaplanmıştır. Bulunan değerler Landis ve Koch (1977) yorumuna göre incelenmiştir. Elde edilen sonuca göre iki psikiyatrist arasında tanı uyumu zayıf düzeyde kabul edilmiştir. Çalışmadaki en önemli eleştiri ise iki gözlemci arasındaki toplam uyumlu sonuç sayısı aynı olduğu durumlarda farklı sonuçlar verebilmesidir. Kategori sayısı ne kadar küçük olursa hesaplanan kappa değerinin o kadar büyük olduğuna değinilmiştir.

Vanbelle (2017), “Çok Düzeyli Verilerde Elde Edilen Bağımlı Kappa Katsayılarının Karşılaştırılması” başlıklı çalışmasında çok düzeyli veriler üzerinde elde edilen çeşitli bağımlı kappa katsayılarını karşılaştırmış, çok sınırlı sayıda denek olması durumunda her zaman yeterli olmayan daha gelişmiş modelleme tekniklerine iki basit alternatif sunmuştur. Giriş bölümünde çok düzeyli kappa katsayılarının tanımı yapılmış ve birkaç bağımlı kappa katsayıları karşılaştırılmıştır. Daha sonra Tip I hata oranının davranışını incelemek için, sabit marjinal olasılık dağılımı ve değişken çiftleri arasında kappa katsayısı ile çok düzeyli bağımlı kategorik değişken simüle edilmiştir. Her simülasyon şeması için, ortalama karesel hata, ortalama standart hata kayıt edilmiş ve şekil oluşturulmuştur. Simülasyonlarda gösterilenlere göre sonuç olarak, az sayıda küme ve küçük küme boyutları için bile verilerin hiyerarşik yapısının hesaba katılması şiddetle tavsiye edilmiştir.

De Raadt ve diğerkleri (2019), alıřmalarında Cohen'in kappa'nın kayıp verileri işleyen üç türünü sunmaktadırlar. Gwet'in kappası, Normal Kategori Kappa ve Silme Yöntemli Kappadır (Listwise Deletion Kappa). Veri, değerklendirmelerden biri veya ikisi birden bir birim eksikse kayıp kabul edilir. 2x2 ve 3x3 uyum tabloları oluşturularak, 10000 tekrar ile simülasyon alıřması yapılmıştır. Ortalama karesel hata (MSE) ve yanlılık değerkleri hesaplanmıştır. İyi performans gösterdiğinden ve hesaplanması kolay olduğundan, kayıplığın tamamen rastgele veya rastgele olmadığı varsayılabilirse, kayıp değerklendirmelerin listeye göre silinmesine dayanan kappa katsayısının kullanılması önerilmiştir.

Alruhaymi & Kim (2021) araştırmasında kayıp veri sorunun nasıl ortaya çıktığını ve kayıp veri sorununda verilerin nasıl analiz edilmesi gerektiğini açıklamıştır. Kayıp verileri rastgele kayıp(MAR), rastgele tamamen eksik(MCAR), rastgele değilk eksik(MNAR) olarak gruplandırmıştır. Eksik verileri tahmin etmek için geleneksel ve modern tekniklerden bahsetmiştir. Bu teknikler, liste bazında silme, ikili silme, modern atama yöntemleri şeklindedir. Sonuç olarak herhangi bir test veya deneysel projede kayıp verilerin her zaman sınırlayıcı olduğunu, yanlıř ıkarımlar yapılmasına neden olabileceğini vurgulanmıştır.

2.6. R Yazılımı

Bilimsel arařtırmalarda veri analizi sürecinde yapılan özümlemeler ve elde edilen bulgular önem taşımaktadır. Nicel verilerin özümlemesinde ise çoğunlukla istatistiksel paket programlarından faydalanılmaktadır. Bu paket programlardan para ile temin edilenlerin yanı sıra, ücretsiz ve pek çok özümleme tekniğı kullanımına olanak sağlayan R yazılımının kullanımı yaygınlaşmıştır.

R yazılımının temelini Becker ve Chambers tarafından geliştirilmiş olan “S” dili oluşturmaktadır. R yazılımı daha önce piyasaya sunulan ve ücretli bir yazılım olan S-PLUS'ın gelişmiş bir versiyonudur (Er ve Sönmez,2005; İlk,2011).

İstatistik programlarından para ile temin edilenlerin yanı sıra, yazılımların dayandığı kodları geri planda saklayarak arařtırmacılara bazı kısa yolların bulunduğu bir ara yüz sağlar. Kullanıcı burada ilgili sekmeleri işaretleyerek ve tıklayarak gerekli analizleri gerçekleştirir ancak kullanıcı yazılımın geri planda işleyen ve analizin yapılmasını sağlayan kodları göremez. Açık kaynak kodlu olan R yazılımı ise bu kodları kullanıcı ile paylaşır ve dünyanın herhangi bir yerinde herhangi birisinin yazılımın gelişmesine katkı sağlamasına olanak tanır. Bu özellik

R yazılımının çok dinamik ve sürekli gelişen bir yapıya sahip olmasını sağlar. (Doğan,C.D.,& Uluman,M. 2016)

R yazılımının diğer programlara göre avantajları olduğu bilinmektedir. Bunlardan en önemlisi, veri bilimindeki analiz süreçlerini gerçekleştirme ve veri görselleştirmede hızlı ve güçlü olmasıdır. Açık kaynaklı ücretsiz bir yazılım olması tercih edilmesini sağlayan diğer bir etkidir. Birçok özelleştirilmiş pakete erişim imkânı tanır.

R yazılımının avantajları bulunmaktadır. Bunlardan birincisi nicel analizlerin hesaplanmasına olanak tanıyan güçlü bir progamlama diline sahiptir. İkincisi ise kullanıcılara herkes tarafından ulaşılabilir istatistiksel paketleri sunar ve imkan tanır. Üçüncüsü ise daha önce belirtildiği gibi R'ın açık kaynak kodlu bir yazılım olmasıdır. (Doğan,C.D ve Uluman,M., 2016)



BÖLÜM III

YÖNTEM

3.1. Araştırmanın Modeli

Çalışmanın modeli bilimsel araştırma türlerinden temel araştırmadır. Temel araştırmalar; kuramlara dayanarak, varsayımlar geliştirerek, bunları test ederek, sonuçlarını bilimsel olarak yorumlayarak ve ortaya çıkararak yapılan araştırmalardır. Bu çalışmada, kayıp değer durumunda, iki veya daha fazla puanlayıcılar arası uyumu incelemek için araştırmalarda yaygın olarak kullanılan Cohen Kappa, Gwet Kappa, Silme Yöntemli Kappa, Fleiss Kappa, Light Kappa ve çalışma kapsamında önerilen Gwet Kappa'ya Dayalı Light Kappa yöntemlerinin performansları karşılaştırılmıştır.

3.1.1. Araştırmanın Materyali

Araştırmanın materyalini farklı kayıp oranları ve farklı senaryolara göre üretilen veri kümeleri oluşturmaktadır. Aynı zamanda R yazılımında bulunan “*irr*” paketi içerisinde bulunan iki gerçek veri kümesinden yararlanılmıştır.

3.1.1.1. Farklı Kayıp Oranlarında ve Farklı Durumlarda Senaryolara Göre Veri Oluşturma

R yazılımında “*sample*” fonksiyonu kullanılarak rastgele veriler üretildi. Kişi sayısı $N=100$, 1000 , 10000 olmak üzere ve iki puanlayıcı için uyumu gösteren $0,1$ 'li veri setleri oluşturuldu. Bu aşamada üretilen senaryolarda tam veri setlerinde kayıp değerler oluşturuldu. Bu değerler oluşturulurken ‘*add_missing*’ fonksiyonu kullanılarak %5, %10, %15, %20, %25, %30 oranlarında MCAR ile kayıp değerler atandı.

Fonksiyon çalıştırıldığında her seferinde farklı sonuçlar elde edildi. Karşılaştırma yapabilmek için farklı olasılıklarda veri setleri oluşturuldu.

3.1.1.2. “irr” Paketi İçerisinde Bulunan Gerçek Veri Kümeleri

Çalışmada R yazılımında “irr” paketi içerisinde bulunan ‘*diagnoses*’, ‘*anxiety*’ veri kümeleri kullanılmıştır. Tablo 3.1’de 30 hastanın her bir puanlayıcı tarafından psikiyatrik tanıları 1.Depresyon, 2. Kişilik Bozukluğu, 3.Şizofreni, 4. Nevroz (Sinir Hastalıkları), 5.Diğer olmak üzere kaydedilmiş ve sınıflandırılmıştır. Tablo 3.2’de üç puanlayıcı tarafından derecelendirilen anksiyete verileri verilmiştir. Burada kullanılan derecelendirme ölçeği; 1 hiç endişeli değil, 6 son derece endişeli şeklinde ifade edilmektedir.

Tablo 3. 1

Üç Puanlayıcı Tarafından Sınıflandırılmış ‘diagnoses’ Veri Kümesi

	Puanlayıcı 1	Puanlayıcı2	Puanlayıcı3
1	4	4	4
2	2	2	2
3	2	3	3
4	5	5	5
5	2	2	2
6	1	1	3
7	3	3	3
8	1	1	3
9	1	1	4
10	5	5	5
11	1	4	4
12	1	2	4
13	2	2	2
14	1	4	4
15	2	2	4
16	3	3	3
17	1	1	1
18	1	1	1
19	2	2	4
20	1	3	3
21	5	5	5
22	2	4	4
23	2	2	4
24	1	1	4
25	1	4	4
26	2	2	2
27	1	1	1
28	2	2	4
29	1	3	3
30	5	5	5

Tablo 3. 2*Üç Puanlayıcı Tarafından Sınıflandırılan 'anxiety' Veri Kümesi*

	Puanlayıcı1	Puanlayıcı2	Puanlayıcı3
1	3	3	2
2	3	6	1
3	3	4	4
4	4	6	4
5	5	2	3
6	5	4	2
7	2	2	1
8	3	4	6
9	5	3	1
10	2	3	1
11	2	2	1
12	6	3	2
13	1	3	3
14	5	3	3
15	2	2	1
16	2	2	1
17	1	1	3
18	2	3	3
19	4	3	2
20	3	4	2

3.2. Önerilen Yaklaşım: Kayıp Veri Durumunda Gwet Kappa'ya Dayalı Light Kappa (GKDLK)

Puanlayıcı sayısı iki olduğunda Cohen kappa'nın kayıp verileri işleyen bir türü olan Gwet Kappa, kayıp değerleri ayrı bir kategoride değerlendirerek hesaplamaya dahil etmektedir. Puanlayıcı sayısı ikiden fazla olması ve kayıp değer varlığında, puanlayıcılar arası uyumun belirlenmesi için tam veri durumunda ikiden fazla puanlayıcı varlığında kullanılan Light Kappa istatistiğinin ve kayıp değer durumunda iki puanlayıcı için kullanılan Gwet Kappa istatistiği ile birlikte değerlendirilmesi sonucunda Eşitlik (19) elde edilmiş ve bu eşitlik Gwet Kappa'ya Dayalı Light kappa (GKDLK) olarak adlandırılmıştır.

$$\kappa_{GKDLK} = \frac{2}{h(h-1)} \sum_{i=1}^{h-1} \sum_{j=i+1}^h \kappa_g(i, j) \quad (19)$$

Burada, h puanlayıcı sayısını ve $\kappa_g(i, j)$ ise $i = 1, 2, \dots, h$ ve $j = i+1, \dots, h$ için Gwet kappa istatistiklerini gösterir.

3.3. Veri Analizi Aracı

Araştırmada açık kaynak kodlu R (R-4.1.3.) yazılımı kullanılmıştır. R diğer birçok özelliğin yanı sıra veri manipülasyonu, programlama ve grafiksel gösterimi bir arada sunan entegre bir yazılımdır (Venables ve Smith, 2014).

3.4. Veri Analizi Süreci

3.4.1. “irr” Paketi İçerisinde Bulunan Veri Kümelerinin Analize Hazırlanması

R yazılımında “irr” paketi içerisinde bulunan ‘*diagnoses*’, ‘*anxiety*’ veri kümeleri analize hazırlanmıştır. Tablo 3.1’de 30 hastanın her bir puanlayıcı tarafından psikiyatrik tanıları 1. Depresyon, 2. Kişilik Bozukluğu, 3. Şizofreni, 4. Nevroz (Sinir Hastalıkları), 5. Diğer olmak üzere kaydedilmiş ve sınıflandırılmıştır. Tablo 3.2’de üç puanlayıcı tarafından derecelendirilen anksiyete verileri verilmiştir. Burada kullanılan derecelendirme ölçeği; 1 hiç endişeli değil, 6 son derece endişeli şeklinde ifade edilmektedir. Gerçek veri kümelerine farklı kayıp oranlarına göre (%5, %10, %15, %20, %25, %30) MCAR ile kayıp değerler eklenmiştir. Fleiss kappa, Light Kappa ve Gwet Kappa’ya Dayalı Kappa değerleri hesaplanmıştır.

3.4.2. Gwet Kappa’ya Dayalı Light Kappa Hesabı için Veri Kümelerinin Analize Hazırlanması

R yazılımında “*sample*” fonksiyonu kullanılarak üretilen rastgele veriler ve “irr” paketi içerisinde bulunan gerçek veri kümeleri Gwet Kappa’ya Dayalı Light Kappa hesabı için hazırlandı. 3 puanlayıcı durumunda değerlendirilen ve farklı kayıp oranlarına göre kayıp değerler eklenen veri kümelerine önerilen yaklaşım olan Light Kappa istatistiği ve Gwet Kappa istatistiğinin birlikte değerlendirilmesi sonucu oluşan eşitlik uygulanmıştır. Puanlayıcı 1 ve 2, Puanlayıcı 1 ve 3 ve Puanlayıcı 2 ve 3 olmak üzere ikili gruplar halinde hesaplamalar yapılmıştır.

BÖLÜM IV

BULGULAR

Bu bölümde, alt problemlere ilişkin araştırma bulguları yer almaktadır. Araştırmadan elde edilen bulgulara dayalı olarak ulaşılan sonuçlar değerlendirilmiştir.

4.1. Tam Veri Durumu için Üretilen Veri Kümesinden Elde Edilen Bulgular

Çalışmada iki puanlayıcı tarafından puanlanan farklı veri kümeleri oluşturularak her bir veri kümesi için Cohen kappa katsayısı hesaplanmıştır. Tam veri durumu için hesaplamalar yapılmıştır. Elde edilen sonuçlar Tablo 4.1’de gösterilmiştir.

Tablo 4.1

Farklı Veri Setlerinden Elde Edilen Uyum Oranları

	2 Puanlayıcı İçin Uyum Oranları			
	1	2	3	4
P_{11}	.35	.40	.20	.20
P_{12}	.20	.20	.25	.35
P_{21}	.15	.15	.30	.20
P_{22}	.30	.25	.25	.25
κ	.30	.28	.50	.15

Tablo 4.1’de tam veri durumunda iki puanlayıcı için 100 katılımcı sayısı olmak üzere ‘sample’ fonksiyonu kullanılarak 4 farklı veri kümesi oluşturulmuştur ve Cohen’in kappa katsayıları hesaplanmıştır. Bu hesaplama yapılırken P_{11} , P_{22} satırlarındaki her iki puanlayıcının anlaştığı gözlemlenen uyumun yüzdesi oluşturulmuştur. İki puanlayıcının farklı değerlendirmelerde bulunduğu beklenen uyumun yüzdesini oluşturmak için diğer satırlardan yararlanılmıştır. Kappa değerleri sırasıyla .30, .28, .50, .15 bulunmuştur.

4.2. Kayıp Veri Durumu için Üretilen Veri Kümesinden Elde Edilen Bulgular

Tablo 4.2’de ise iki Puanlayıcı için farklı katılımcı sayısı ve farklı kayıp oranları ile rasgele üretilen veri kümeleri kayıp veri durumunda karşılaştırılmıştır. %5, %10, %15, %20, %25, %30 oranlarında MCAR durumuna göre kayıp veriler oluşturulmuştur. Kayıp veri durumunda kappa çeşitleri kullanılarak hesaplamalar yapılmıştır. %M kayıp veri oranını belirtmektedir. %5, %10, %15, %20, %25 ve %30 kayıplık durumda Cohen Kappa ve Gwet Kappa katsayıları hesaplanmıştır. R yazılımında Cohen Kappa katsayısı hesaplanırken kayıp değerler silinerek sonuç alınır ancak Gwet kappa hesabında kayıp değerler ayrı bir kategoride hesaplamaya dahil edilir.

Tablo 4.2

Farklı Senaryolara Göre Oluşturulan Veri Kümelerine Kayıp Veri Ataması Yapıldıktan Sonra Elde Edilen Sonuçlar (2 Puanlayıcı Durumu)

N (kişi sayısı)	%M	Cohen Kappa	Z	p- değeri	Gwet Kappa	z	p- değeri
100	5	.144	1.250	.210	.157	1.250	.218
100	10	.112	.987	.324	.119	.901	.267
100	15	.005	.050	.960	.018	0	1
100	20	.049	.433	.665	.154	.407	.614
100	25	-.048	-.427	.669	-.002	-.495	.629
100	30	.076	.712	.476	.261	.619	.536
1000	5	.005	.148	.883	.005	.141	.008
1000	10	-.001	-.048	.961	-.002	-.049	.261
1000	15	-.009	-.266	.791	-.011	-.254	.79
1000	20	.026	.749	.454	.199	.741	.455
1000	25	.033	.949	.343	.286	.94	.245
1000	30	.014	.419	.675	.317	.421	.675
10000	5	.004	.370	.711	.005	.362	.113
10000	10	.007	.687	.492	.082	.854	.492
10000	15	-.002	-.222	.825	-.001	-.201	.925
10000	20	-.000	-.023	.982	-.000	-.021	.982
10000	25	.000	.020	.984	.000	0	.984
10000	30	.001	.146	.884	.000	.010	.808

Tablo 4.3’de üç Puanlayıcı için farklı katılımcı sayısı ve farklı kayıp oranları ile rasgele üretilen veri kümeleri kayıp veri durumunda karşılaştırılmıştır. %5, %10, %15, %20, %25, %30 oranlarında MCAR durumuna göre kayıp veriler oluşturulmuştur. Kayıp veri durumunda kappa çeşitleri kullanılarak hesaplamalar yapılmıştır. %M kayıp veri oranını belirtmektedir. %5, %10, %15, %20, %25 ve %30 kayıplık durumda Fleiss Kappa, Light Kappa ve Gwet Kappa’ya Dayalı Light kappa katsayıları hesaplanmıştır. P değeri istatistiksel anlamlılığı ifade etmektedir.

Tablo 4. 3

Farklı Senaryolara Göre Oluşturulan Veri Kümelerine Kayıp Veri Ataması Yapıldıktan Sonra Elde Edilen Sonuçlar (3 Puanlayıcı Durumu)

N (kişi sayısı)	% M	Fleiss kappa	z	p- değeri	Light kappa	z	p- değeri	Gwet Kappa’ya Dayalı Light kappa	z	p- değeri
100	5	.143	1.250	.212	.144	1.230	.218	.144	1.210	.222
100	10	.104	.901	.367	.112	.961	.337	.118	.961	.317
100	15	0	0	1	.005	.049	.961	.02	.030	.989
100	20	.046	.407	.684	.049	.429	.668	.247	.048	.635
100	25	-.057	.499	.618	-.048	-.411	.681	-.035	-.429	.681
100	30	.068	.619	.536	.076	.693	.488	.079	.714	.238
1000	5	.005	.145	.885	.005	.148	.883	.032	.148	.872
1000	10	-.001	-.049	.961	-.001	.048	.961	-.011	-.042	.961
1000	15	-.009	-.266	.79	-.009	.265	.791	-.017	.264	.799
1000	20	.026	.748	.455	.026	.747	.455	.026	.745	.455
1000	25	.033	.94	.347	.033	.947	.344	.063	.99	.344
1000	30	.014	.419	.675	.014	.418	.676	.028	.416	.677
10000	5	.004	.368	.713	.004	.370	.711	.044	.421	.708
10000	10	.007	.687	.492	.007	.687	.492	.009	.687	.492
10000	15	-.002	-.222	.825	-.002	-.222	.825	-.002	-.222	.825
10000	20	-.000	-.023	.982	-.000	-.023	.982	-.000	-.023	.982
10000	25	.000	.020	.984	.000	.020	.984	.000	.020	.989
10000	30	.001	.145	.885	.001	.146	.884	.001	.145	.881

4.3. ‘diagnoses’ ve ‘anxiety’ Veri Kümelerinde Kappa Katsayısı Hesabından Elde Edilen Bulgular

30 hastanın 3 puanlayıcı tarafından 5 nominal kategoriye kaydedilip sınıflandırıldığı veri kümesinde, puanlayıcılar arası uyumu ölçmek için kappa katsayıları hesaplanmıştır. Kappa değeri 0.651 olarak bulunmuştur. Landis ve Koch (1977), yorumuna göre puanlayıcılar arasında iyi uyum olduğu çıkarımı yapılmıştır. ‘anxiety’ veri kümesinde ise 3 puanlayıcı tarafından 1 hiç endişeli değil, 6 son derece endişeli olmak üzere sınıflandırılmıştır. Kappa hesaplamaları sonucunda 0.189 olarak bulunmuştur. Elde edilen bu sonuç puanlayıcılar arasında zayıf uyum olduğunu göstermektedir.

Veri kümelerinin kayıp veri durumunda nasıl sonuçlanacağını karşılaştırmak amacıyla %5, %10, %15, %20, %25, %30 oranlarında kayıp değerler atanmıştır. Fleiss Kappa, Light Kappa, Gwet Kappa’ya Dayalı Light Kappa değerleri R yazılımında hesaplanmıştır. Elde edilen bulgular Tablo 4.4 ve Tablo 4.5’de gösterilmektedir. Tablo 4.4 ve Tablo 4.5 incelendiğinde kayıplık oranları %5, %10, %15, %20, %25, %30 durumu için Fleiss kappa, Light Kappa ve Gwet Kappa’ya Dayalı Kappa önemlidir ($p < 0,05$).

Tablo 4.4

R Yazılımında Bulunan ‘diagnoses’ Veri Kümesine Kayıp Veri Ataması Yapıldıktan Sonra Elde Edilen Sonuçlar

M(%)	Fleiss Kappa	z	p-değeri	Light Kappa	z	p-değeri	Gwet Kappa’ya Dayalı Light Kappa	z	p-değeri
5	.456	7.71	0	.487	.427	.668	.513	.475	0
10	.428	16.3	0	.456	-.211	.883	.586	.124	0
15	.510	7.68	0	.533	.482	.712	.548	.482	0
20	.473	6.52	0	.500	.145	.883	.509	.145	0
25	.507	6.27	0	.531	.416	.483	.547	.501	0
30	.479	5.24	0	.511	.408	.983	.674	.483	0

Tablo 4.5

R Yazılımında Bulunan ‘anxiety’ Veri Kümesine Kayıp Veri Ataması Yapıldıktan Sonra Elde Edilen Sonuçlar

M(%)	Fleiss Kappa	z	p- değeri	Light Kappa	z	p- değeri	Gwet Kappa’ya Dayalı Light Kappa	z	p- değeri
5	-.015	-.211	.833	.011	.240	.214	.003	.302	0
10	-.020	-.264	.791	.001	.146	.884	-.014	.149	0
15	-.039	-.465	.642	-.009	.255	.891	.005	.267	0
20	-.067	-.647	.518	.006	.647	.482	.000	.708	0
25	-.058	-.448	.654	.001	.274	.866	.000	.001	0
30	-.090	-.589	.556	0	0	1	.011	.002	0

BÖLÜM V

SONUÇ, TARTIŞMA VE ÖNERİLER

Bu bölümde araştırmanın amaçları doğrultusunda analiz sonuçlarına ve önerilere yer verilmiştir.

5.1. Sonuç ve Tartışma

R yazılımında kayıp veri mekanizmaları ile oluşturulan veri setlerinde puanlayıcılar arası uyum incelenmiştir. Puanlayıcılar arası uyum incelenirken Cohen Kappa, Gwet Kappa, Fleiss Kappa, Light Kappa ve çalışma kapsamında önerilen Gwet Kappa'ya Dayalı Light Kappa istatistikleri kullanılmıştır. İki puanlayıcı için kayıp veri durumunda Cohen Kappa ve Gwet Kappa, üç puanlayıcının olduğu durumda Fleiss Kappa, Light Kappa ve Gwet Kappa'ya Dayalı Light Kappa katsayıları hesaplanmıştır.

Elde edilen kappa katsayıları sonuçlarına göre değerler 1' e yakın olduğu durumda puanlayıcıların iyi uyum içerisinde olduğu söylenebilir. Tam veri durumunda katılımcı sayısı 100, puanlayıcı sayısı iki olduğunda, uyum oranlarını belirlemek için hesaplanan kappa katsayıları Tablo 4.1'de verilmiştir. Kappa değerleri sırasıyla .30, .28, .50, .15 bulunmuştur. Puanlayıcılar arası uyum düzeyleri Landis ve Koch (1977) yorumuna göre sırasıyla “az uyum”, “az uyum”, “orta uyum”, “çok az uyum” şeklindedir. Katılımcı sayısı arttıkça kappa katsayısının daha iyi sonuç verdiği gözlemlenmiştir. İki puanlayıcı durumunda Tablo 4.2'de belirtilen kayıp değere sahip veri kümeleri için kappa versiyonlarında, katılımcı sayısının 10000, kayıplık yüzdesinin %30 olduğu durumda sonuçlar neredeyse aynıdır. İki den fazla puanlayıcı olması durumunda, katılımcı sayısı 100, kayıplık oranları %15 ve %20 olduğunda; katılımcı sayısı 1000, kayıplık oranı %5 olduğunda; katılımcı sayısı 10000, kayıplık oranları %5 ve %10 olduğunda çalışma kapsamında önerilen Gwet Kappa'ya Dayalı Light Kappa yaklaşımının daha iyi performans gösterdiği görülmüştür. Tablo 4.4 ve Tablo 4.5'de gerçek veri kümelerine kayıp

değer ataması yapıldıktan sonra elde edilen kappa katsayıları sonuçları verilmiştir. Gerçek veri kümelerinden elde edilen bulgulara göre 3 puanlayıcı için Fleiss Kappa, Light Kappa ve Gwet Kappa'ya Dayalı Light Kappa hesaplamalarında yeni yaklaşım olan Gwet Kappa'ya Dayalı Light Kappa tüm kayıplık oranlarında daha anlamlıdır.

Eğitim araştırmalarında kayıp verilerle başa çıkmak için çeşitli yollar bulunmaktadır. Kayıp verilerle uğraşmak için Cohen'in kappa'sının birkaç versiyonu bulunmaktadır. Cohen'in kappa katsayısının versiyonları; Liste Boyunca Silmeye Dayanan Kappa, Gwet Kappa. Bunlardan birincisi kayıp verinin bulunduğu satırların hepsinin silindiği liste bazında silmeye dayanan kappa yöntemidir. Ancak bu durum bazı katılımcıların göz ardı edilmesine sebep olmuştur. R yazılımında veri kümelerine bu yöntem uygulandığında neredeyse veri kümesinin büyük bir çoğunluğu silinmiştir. Bu yöntem doğru değerlendirme ve karşılaştırma yapmayı zorlaştıran bir stratejidir. İkinci bir yöntem olan Gwet'in kappasında kayıp değerleri ayrı bir kategori ile sınıflandırılarak hesaplamaya dahil etmektedir. Gwet'in kappa formülasyonunda, gözlemlenen uyumun yüzdesinin hesaplanmasında yalnızca iki değerlendirmeye sahip birimler dahil edilmiştir. Kayıp veriler daha kesin tahminler yapmak için kullanılır. Eksik değerlendirmeye sahip birimler, beklenen uyum yüzdesinin hesaplanmasında kullanılmıştır. İki eksik değerlendirme tamamen hesaplama dışında bırakılmıştır. Bu nedenle Gwet'in kappası hesaplamada daha iyi performans göstermiştir. Liste bazında Silmeye Dayanan Kappa katsayısında ise puanlayıcılar tarafından bir puanlama veya her iki puanlama eksik olduğu durumda tamamen silerek hesaplama dışında bırakılır.

Puanlayıcı sayısının ikiden fazla olduğu durumlarda Fleiss kappa hesabından yararlanılabilir. R'de bulunan `kappam.fleiss()` [irr paketi] fonksiyonu, Fleiss kappa'yı kategorik veriler üzerinde puanlayıcılar arası uyumun bir indeksi olarak hesaplamak için kullanılır. Rastgele oluşturulan veri kümelerinde Fleiss kappa performansı, diğer versiyonlardan yüksektir ancak R yazılımında bulunan '*diagnoses*' ve '*anxiety*' veri kümelerinde yüksek performans gösterememiştir. Yeni bir yaklaşım olarak çalışma kapsamında oluşturulan Gwet Kappa'ya dayanan Light Kappa'da ise sonuçların daha iyi olduğu söylenebilir.

Literatürde kayıp veri durumunda genellikle iki veya üç kappa katsayısı üzerinden tartışmalar yapılmıştır (Gwet (2012); Gwet (2014); Raadt vd. (2019)). Veri kümelerinde kayıp veri problemi ve çözümüne ilişkin çalışmalar (Baraldi ve Enders, 2010; Enders, 2010; Peugh ve Enders, 2004), genellikle nicel veri kümelerinde yer alan kayıplık üzerine yapılmıştır. Literatür incelendiğinde kayıp veri durumunda puanlayıcılar arası uyumun araştırıldığı çalışmaların son

derece kısıtlılığı olduğu görülmüştür. Bu bağlamda bu çalışmada ikiden fazla puanlayıcı olması durumunda kayıplık problemine çözüm üzerine odaklanılmıştır.

5.2. Öneriler

İki puanlayıcı tarafından puanlanan kayıp değerlerin bulunduğu veri kümelerinde kappa çeşitlerinden Gwet'in kappa katsayısı kayıp değerleri hesaba katarak işlem yaptığından daha iyi sonuç verdiği görülmüştür. Üç puanlayıcı tarafından puanlanan veri kümelerinde ise kayıp veri durumunda önerilen yaklaşım olan Gwet Kappa'ya Dayalı Light Kappa daha iyi sonuç vermiştir. Bu yaklaşım, kayıp veri durumunda iyi bir tercih olacaktır. Silme yöntemlerinde olduğu gibi veriyi yok saymak hatalı yorumlamalara yol açabilir. Çünkü büyük ölçüde veri kaybına yol açmaktadır. Araştırmalarda kayıp veri durumunda, liste bazında silmeye dayanan kappa sık tercih edilmektedir ancak daha doğru sonuçlar elde etmek için Gwet Kappa'ya Dayalı Light kappa önerilebilir.

Gwet kappa'ya Dayalı Light kappa katsayısı, Light kappa ve Fleiss kappa katsayısına göre daha iyi yorumlamalar yapılmasına yardımcı olmuştur. Bu nedenle kappa katsayıları arasından Gwet Kappa'ya Dayalı Light Kappa katsayısı tercih edilebilir. Son olarak, eğitim araştırmalarında kayıp veri sorunu ve kappa katsayısı ile ilgili çalışmalar halen devam etmektedir. Bu nedenler bu tür çalışmalara yer verilmesi tavsiye edilmektedir. Çalışmanın bu anlamda eğitim araştırmalarına yön vereceği öngörülmektedir.

KAYNAKÇA

- Allison, P.D. (2002). Missing Data. California: Sage Publication, Inc
- Alruhaymi, A. Z. ve Kim, C. J. (2021). Study on the Missing Data Mechanisms and Imputation Methods. *Open Journal of Statistics*, 11(4), 477-492.
- Bilgen, Ö. B. ve Doğan, N. (2017). Puanlayıcılar arası güvenirlik belirleme tekniklerinin karşılaştırılması. *Journal of Measurement and Evaluation in Education and Psychology*, 8(1), 63-78.
- Bağ,H.,Karabulut,E.,Alpar,R. (2010). ‘2x2 Tablolarda Gözlemciler/Gözlemler Arası Uyumun Değerlendirilmesi’, *Hacettepe Diş Hekimliği Fakültesi Dergisi*, 34 (2), 46-52
- Çiftçi,S., “Rastgele olmayan kayıp verili değişkenler içeren genelleştirilmiş doğrusal modellerin analizi için Bayesiyen model havuzu”,Orta Doğu Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Ana Bilim Dalı
- Dalkıran, E. (2006). Keman eğitiminde performansın ölçülmesi. *Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 5(2), 116-136.
- Demir, E. ve Parlak, B. (2012). Türkiye’de eğitim araştırmalarında kayıp veri sorunu. *Journal of Measurement and Evaluation in Education and Psychology*, 3(1), 230-241.
- De Raadt, A., Warrens, M. J., Bosker, R. J., ve Kiers, H. A. (2019). Kappa coefficients for missing data. *Educational and psychological measurement*, 79(3), 558-576.

- Doğan, C. D. ve Uluman, M. (2016). İstatistiksel veri analizinde R yazılımı ve kullanımı. *Ilkogretim Online*, 15(2).
- Elishakoff, I (2003), “Notes on philosophy of the monte carlo method”, *International Applied Mechanics*, Vol.39, No.7, pp.753-762
- Foster, D. ve G.Arthur (1982), “Average neutronic properties of prompt fission products”, Los Alamos National Laboratory Report LA-9168-MS.
- Fleiss’in kappa katsayısı.(2020,Nisan 8). İçinde Wikipedia.
- Graham, J.W. (2009). Missing data analysis: making it work in the real world. *Annual Review of Psychology*, s.60, 4, sy.549-576
- İnan, S. (2001). Normal, multinomial, üssel (exponential) ve gamma dağılım gösteren veri yapıları ve eksik veri tiplerinde (MCAR, MAR, MNAR) tamamlama algoritmalarının parametre tahminleri üzerine etkileri (Master's thesis, Fen Bilimleri Enstitüsü).
- Kılıç,S.(2015). ‘Kappa Testi’,Gülhane Askeri Tıp Fakültesi
- Landis, J. R. ve Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174.
- McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, 22,276-282.
- Mielke, P., Bloch, D., ve Kraemer, H. (1989). Cohen's Kappa. *Biometrics*, 45(4), 1329-1330.
- Retrieved July 15, 2021, from <http://www.jstor.org/stable/2531787>
- <http://www.datanovia.com/en/lessons/inter-rater-reliability-analyses-quick-r-codes/>
- <https://www.datanovia.com/en/lessons/cohens-kappa-in-r-for-two-categorical-variables/>

- Vanbelle, S., ve Albert, A. (2009). Agreement between two independent groups of raters. *Psychometrika*, 74, 477-491.
- Vanbelle, S. (2017). Comparing dependent kappa coefficients obtained on multilevel data. *Biometrical Journal*, 59(5), 1016-1034.
- Yıldıztekin, B. (2014). Klasik test kuramı ve genellenebilirlik kuramından puanlayıcılar arası tutarlılığın farklı yöntemlere göre karşılaştırılması.
- Warrens, M. J. (2015). Five ways to look at Cohen's kappa. *Journal of Psychology & Psychotherapy*, 5(4), 1.

EKLER

Ek-1: Veri Üretme ve Karşılaştırmaya İlişkin R Kodları

```
xtab<-as.table(rbind(c(35,20),c(15,30)))  
diagonal.counts<-diag(xtab)  
N<-sum(xtab)  
row.marginal.props<-rowSums(xtab)/N  
col.marginal.props<-colSums(xtab)/N  
 $P_0$ <-sum(diagonal.counts)/N  
 $P_e$ <-sum(row.marginal.props*col.marginal.props)  
 $k <- (P_0 - P_e)/(1 - P_e)$   
install.packages("SimDesign")  
library(SimDesign)  
head(ymiss<-add_missing(diagnoses),10)  
kappam.fleiss (v)  
kappam.fleiss (v, exact=TRUE)  
kappam.fleiss (v, detail=TRUE)  
kappam.light(df2)  
kappam.light(veri)  
kappa2(anxiety[,c("rater1","rater2")],weight="equal")  
kappa2(diagnoses[,2:3])  
veri=matrix (sample(c(0,1),200,replace=T),ncol=2)  
veri=matrix (sample(c(0,1),2000,replace=T),ncol=2)
```

```
veri=matrix (sample(c(0,1),20000,replace=T),ncol=2)

head(y<-add_missing(veri),5)

head(y<-add_missing(veri),10)

head(y<-add_missing(veri),15)

head(y<-add_missing(veri),20)

head(y<-add_missing(veri),25)

head(y<-add_missing(veri),30)

df1<-diagnoses

df1[c(1,5),2]<-NA

df1[c(2,6),1]<-NA

df2<-na.omit(df1)

kappa2(df2[,c("degerlendirici1","degerlendirici2")],weight="unweighted")

df3<-anxiety

kappa2(df4[,1:2],"squared")
```

ÖZGEÇMİŞ

Kişisel Bilgiler

Adı Soyadı : Halime TOPBAŞ

Eğitim Durumu

Lisans Öğrenimi : Akdeniz Üniversitesi, Eğitim Fakültesi,
İlköğretim Matematik Öğretmenliği

Yüksek Lisans Öğrenimi : Akdeniz Üniversitesi, Eğitim Bilimleri
Enstitüsü, Eğitimde Ölçme ve
Değerlendirme Tezli Yüksek Lisans
Programı

Bildiği Yabancı Diller İngilizce

Tarih 10/06/2022

BİLDİRİM

Hazırladığım tezin/raporun tamamen kendi çalışmam olduğunu ve her alıntıya kaynak gösterdiğimi taahhüt eder, tezimin/raporumun kağıt ve elektronik kopyalarının Akdeniz Üniversitesi Eğitim Bilimleri Enstitüsü arşivlerinde aşağıda belirttiğim koşullarda saklanmasına izin verdiğimi onaylıyorum:

- ☐ Tezimin/Raporumun tamamı her yerden erişime açılabilir.

Tezim/Raporum sadece Akdeniz Üniversitesi yerleşkelerinden erişime açılabilir.

- ☒ Tezimin/Raporumun 3 yıl süreyle erişime açılmasını istemiyorum. Bu sürenin sonunda uzatma için başvuruda bulunmadığım takdirde, tezimin/raporumun tamamı her yerden erişime açılabilir.

10.06.2022

Halime TOPBAŞ

İNTİHAL RAPORU

Turnitin Orijinallik Raporu

İşleme kondu: 12-May-2022 22:49 +03
NUMARA: 1834874493
Kelime Sayısı: 6545
Gönderildi: 1

EĞİTİM ARAŞTIRMALARINDA KAYIP VERİ DURUMU İÇİ... Halime
Topbaş tarafından

Benzerlik Endeksi	Kaynağa göre Benzerlik
%16	Internet Sources: %16 Yayınlar: %2 Öğrenci Ödevleri: %5

[alıntılar çıkar](#) [bibliyografayı dahil et](#) [küçük eşleşmeleri çıkar](#) mod: [raporu hızlı görüntüle \(klasik\)](#) [Change mode](#) [yazdır](#) [yenile](#) [İndir](#)

3% match (09-Ağu-2015 tarihli internet)
<http://epod-online.org>

2% match (08-Tem-2019 tarihli internet)
<http://www.openaccess.hacettepe.edu.tr:8080>

2% match (30-Nis-2016 tarihli internet)
<http://dergipark.ulakbim.gov.tr>

1% match (17-Eyl-2011 tarihli internet)
<http://dishekerji.hacettepe.edu.tr>

1% match ()
de Raadt, Alexandra, Warrens, Matthijs J., Bosker, Roel J., Kiers, Henk A. L.. "Kappa Coefficients for Missing Data", 'SAGE Publications', 2019

1% match (21-Eyl-2018 tarihli internet)
<http://www.narsanat.com>

1% match (02-May-2019 tarihli internet)
<https://www.ejmanager.com/mnstems/8/8-1441823131.pdf>

<1% match (13-Mar-2022 tarihli internet)
https://acikbilim.yok.gov.tr/bitstream/handle/20.500.12812/617286/yokAcikBilim_10083564.pdf?isAllowed=y&sequence=-1

<1% match (12-Oca-2022 tarihli internet)
https://acikbilim.yok.gov.tr/bitstream/handle/20.500.12812/40045/yokAcikBilim_10336030.pdf?isAllowed=y&sequence=-1

<1% match (13-Oca-2022 tarihli internet)
https://acikbilim.yok.gov.tr/bitstream/handle/20.500.12812/703122/yokAcikBilim_10185487.pdf?isAllowed=y&sequence=-1

<1% match (12-Oca-2022 tarihli internet)
https://acikbilim.yok.gov.tr/bitstream/handle/20.500.12812/40202/yokAcikBilim_10229825.pdf?isAllowed=y&sequence=-1

<1% match (15-Nis-2022 tarihli internet)
https://acikbilim.yok.gov.tr/bitstream/handle/20.500.12812/222242/yokAcikBilim_10300140.pdf?isAllowed=y&sequence=-1