

T.C.

HASAN KALYONCU ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

EĞİTİMDE ÖLÇME VE DEĞERLENDİRME ANA BİLİM DALI

EĞİTİMDE ÖLÇME VE DEĞERLENDİRME TEZLİ YÜKSEK LİSANS PROGRAMI

**AYNI ÖZELLİĞİ ÖLÇEN İKİ TESTİN FARKLI YÖNTEMLERLE ELDE EDİLEN
SINIFLAMA DOĞRULUKLARININ KARŞILAŞTIRILMASI**

YÜKSEK LİSANS TEZİ

HAZIRLAYAN

NESİR DENİZ TEMEKOĞLU

GAZİANTEP - 2021

T.C.
HASAN KALYONCU ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
EĞİTİMDE ÖLÇME VE DEĞERLENDİRME ANA BİLİM DALI
EĞİTİMDE ÖLÇME VE DEĞERLENDİRME TEZLİ YÜKSEK LİSANS PROGRAMI

**AYNI ÖZELLİĞİ ÖLÇEN İKİ TESTİN FARKLI YÖNTEMLERLE ELDE EDİLEN
SINIFLAMA DOĞRULUKLARININ KARŞILAŞTIRILMASI**

YÜKSEK LİSANS TEZİ

HAZIRLAYAN
NESİR DENİZ TEMEKOĞLU

TEZ DANIŞMANI
DOÇ. DR. UFUK AKBAŞ

GAZİANTEP - 2021

ÖNSÖZ

Ders dönemindeki katkılarından dolayı Prof. Dr. Şener BÜYÜKÖZTÜRK, Prof. Dr. Selahattin GELBAL, Doç. Dr. Erkan Hasan ATALMIŞ, Dr. Öğr. Üyesi Burcu GÜRKAN, Arş. Gör. Merve YILDIRIM SEHERYELİ ve aynı zamanda tez danışmanım olan Doç. Dr. Ufuk AKBAŞ hocalarıma çok teşekkür ederim.

Gerek ders döneminde kattıklarıyla, gerek tez döneminde gösterdiği ilgi ve alakasıyla, yönlendirmeleriyle her zaman yanımda olduğunu hissettiğim, mütevazı kişiliğini örnek aldığım, motivasyon sağlayıcı ve esprili konuşmalarından destek bulduğum değerli hocam Dr. Çiğdem REYHANLIOĞLU KEÇEOĞLU'na çok teşekkür ederim.

Veri toplama sürecinde desteğini esirgemeyen, tez yazım sürecinde manevi desteklerini hep yanımda hissettiğim değerli meslektaşlarıma çok teşekkür ederim.

Tatlı sohbetleri, misafirperver kişilikleri ve gülyüzleriyle Gaziantep'teki ailem bildiğim Canan ÇELİK ve Harun ÇELİK'e çok teşekkür ederim.

Tanıştığımız günden beri birbirimizi hep olumlu yönlendirdiğimiz, fikirlerine, pratikliğine, desteğine her daim ihtiyaç duyduğum, üniversitenin bana en güzel kazanımlarından biri olan, 'ortanca kardeşim', can yoldaşım, imkânsız görünen pek çok şeyi birlikte inanıp birlikte başardığımız Şerife CEYLAN'a çok teşekkür ederim.

Ve elbette attığım her adımda koşulsuz yanımda olduklarına emin olduğum, bugünlere gelmeme sebep, benim için hep en iyiyi istediklerini bildiğim, maddi veya manevi desteklerini daima hissettiğim başta canım annem olmak üzere kıymetli aileme çok teşekkür ederim.

ÖZET

Temekođlu, N. D. (2021). *Aynı Özelliđi Ölçen İki Testin Farklı Yöntemlerle Elde Edilen Sınıflama Doğruluklarının Karşılaştırılması*, Yüksek Lisans Tezi, Gaziantep.

Bu arařtırmada geleneksel standart belirleme yöntemlerine alternatif olarak sunulan üç farklı istatistiksel yöntemin karşılaştırılması amaçlanmıştır. Bu doğrultuda 7. sınıf seviyesine uygun olarak arařtırmacı tarafından geliştirilen matematik başarı testinin iki farklı formu veri toplama aracı olarak kullanılmıştır. Bu formlar aynı özelliđi ölçen çoktan seçmeli ve kısa cevaplı maddelerden oluşmaktadır. Veriler 2020-2021 eğitim öğretim yılında Gaziantep'teki iki farklı devlet okulundaki 159'u kız ve 167'si erkek olmak üzere toplam 326 öğrenciden elde edilmiştir. Milli Eğitim Bakanlığı'nın geçme puanı olarak belirlediđi 45 puanı ölçüt kabul edilmiş, uygulanan testlerden geçme/kalma durumları lojistik regresyon, kümeleme ve ROC analizi yöntemleriyle incelenerek her bir yonteme ve her bir forma ilişkin sınıflama doğrulukları elde edilmiştir. Arařtırmanın bulgularında, yöntemlere göre bakıldığında ÇSF için de (%95.9), KCF için de (%93.2) en doğru sınıflama yüzdesi ROC analiziyle elde edilmiştir. Formlara göre bakıldığında ise her üç yöntem için de ayrı ayrı elde edilen sınıflama doğrulukları istatistiksel olarak manidar bir farklılık göstermemiştir.

Anahtar Sözcükler: Standart Belirleme, Sınıflama Doğruluđu, Lojistik Regresyon Analizi, Kümeleme Analizi, ROC Analizi

ABSTRACT

Temekođlu, N. D. (2021). *Comparison of the Accuracy of Classification Obtained by Different Methods of Two Tests Measuring the Same Feature*, Master Thesis, Gaziantep.

In this study, it was aimed to compare three different statistical methods presented as an alternative to traditional standard setting methods. Accordingly, two different forms of the mathematics achievement test developed by the researcher in accordance with the 7th grade level were used as data collection tools. These forms consist of multiple choice and short answer items that measure the same features. The data were obtained from a total of 326 students, 159 of which were girls and 167 boys, in two different state schools in Gaziantep in the 2020-2021 academic year. 45 points determined as the passing score by the Ministry of National Education were accepted as the criterion, and the classification accuracies for each method and each form were obtained by examining the pass / fail situations from applied tests with logistic regression, cluster and ROC analysis methods. In the findings of the study, when considering the methods, the most accurate classification percentage for MCF (95.9%) and for SAF (93.2%) was obtained by ROC analysis. Considering the forms, the classification accuracies obtained separately for each of the three methods did not show a statistically significant difference.

Keywords: Standart Setting, Classification Accuracy, Logistic Regression Analysis, Cluster Analysis, ROC Analysis

İÇİNDEKİLER

ÖNSÖZ	i
ÖZET	ii
ABSTRACT.....	iii
İÇİNDEKİLER.....	iv
TABLO LİSTESİ	vi
ŞEKİL LİSTESİ	vii
KISALTMALAR.....	viii
BİRİNCİ BÖLÜM	
GİRİŞ	1
1.1. Problem Durumu	1
1.2. Araştırmanın Amacı Ve Önemi	7
1.3. Problem Cümlesi	8
1.4. Alt Problemler.....	8
1.5. Sayıtlar.....	9
1.6. Sınırlılıklar	9
1.7. Tanımlar.....	9
İKİNCİ BÖLÜM	
KAVRAMSAL ÇERÇEVE VE İLGİLİ ARAŞTIRMALAR	10
2.1. Standart Belirleme Yöntemleri	10
2.1.1. Geleneksel Yöntemler.....	10
2.1.2. Alternatif Yöntemler.....	12
2.2. İlgili Araştırmalar.....	21
ÜÇÜNCÜ BÖLÜM	
YÖNTEM.....	27
3.1. Araştırma Modeli	27
3.2. Çalışma Grubu	27
3.3. Verilerin Toplanması.....	28
3.4. Veri Toplama Araçları.....	28

3.5. Verilerin Analizi Ve Yorumu	28
3.5.1. Birinci Ve İkinci Alt Problemler İçin Verilerin Analizi Ve Yorumu	29
3.5.2. Üçüncü Alt Problem İçin Verilerin Analizi Ve Yorumu	32
3.5.3. Dördüncü Alt Problem İçin Verilerin Analizi Ve Yorumu.....	32
3.5.4. Beşinci Alt Problem İçin Verilerin Analizi Ve Yorumu	32
DÖRDÜNCÜ BÖLÜM	
BULGULAR VE YORUM.....	34
4.1. Çoktan Seçmeli Formun Lojistik Regresyon Analizi, Kümeleme Analizi ve ROC Analizi ile Belirlenen Sınıflama Doğruluklarına İlişkin Bulgular.....	34
4.2. Kısa Cevaplı Formun Lojistik Regresyon Analizi, Kümeleme Analizi ve ROC Analizi ile Belirlenen Sınıflama Doğruluklarına İlişkin Bulgular	39
4.3. Çoktan Seçmeli ve Kısa Cevaplı Formun Lojistik Regresyon Analizi ile Belirlenen Sınıflama Doğruluklarına İlişkin Bulgular.....	45
4.4. Çoktan Seçmeli ve Kısa Cevaplı Formun Kümeleme Analizi ile Belirlenen Sınıflama Doğruluklarına İlişkin Bulgular.....	45
4.5. Çoktan Seçmeli ve Kısa Cevaplı Formun ROC Analizi ile Belirlenen Sınıflama Doğruluklarına İlişkin Bulgular.....	45
BEŞİNCİ BÖLÜM	
SONUÇ VE ÖNERİLER.....	47
5.1. Sonuçlar.....	47
5.2. Öneriler.....	49
KAYNAKÇA	51
EKLER	57
Ek 1. Uzman Görüş Formu.....	57
Ek 2. Çoktan Seçmeli Matematik Başarı Testi.....	58
Ek 3. Kısa Cevaplı Matematik Başarı Testi	62
Ek 4. Araştırma Uygulama İzni	65

TABLO LİSTESİ

Tablo 1. AUC Değerlerinin Yorumlanması	20
Tablo 2. Çalışma Grubuna İlişkin Frekans ve Yüzdeler	27
Tablo 3. ÇSF'ye İlişkin LR Analizi Başlangıç Modeli İterasyon Öyküsü	34
Tablo 4. ÇSF'ye İlişkin LR Analizi Başlangıç Modeline Göre Elde Edilen İlk Sınıflandırma Durumu	34
Tablo 5. ÇSF'ye İlişkin LR Analizi Bağımsız Değişkenin Bulunduğu Model İçin İterasyon Öyküsü	35
Tablo 6. ÇSF'ye İlişkin LR Analizi Model Katsayılarına Dair Omnibus Testi	35
Tablo 7. ÇSF'ye İlişkin LR Analizi Sonuç Modelinin Özeti	36
Tablo 8. ÇSF'ye İlişkin LR Analizi Sonucu Elde Edilen Sınıflandırma Tablosu	36
Tablo 9. ÇSF'ye İlişkin Kümelerde Yer Alan Öğrencilerin Frekans ve Yüzdeleri	37
Tablo 10. ÇSF'ye İlişkin Kümeleme Analizi Sınıflandırma Tablosu	37
Tablo 11. ÇSF'ye İlişkin ROC Analizinden Elde Edilen Değerler	38
Tablo 12. Birinci Alt Probleme İlişkin Bulgular	39
Tablo 13. KCF'ye İlişkin LR Analizi Başlangıç Modeli İterasyon Öyküsü	39
Tablo 14. KCF'ye İlişkin LR Analizi Başlangıç Modeline Göre Elde Edilen İlk Sınıflandırma Durumu	40
Tablo 15. KCF'ye İlişkin LR Analizi Bağımsız Değişkenin Bulunduğu Model İçin İterasyon Öyküsü	40
Tablo 16. KCF'ye İlişkin LR Analizi Model Katsayılarına Dair Omnibus Testi	41
Tablo 17. KCF'ye İlişkin LR Analizi Sonuç Modelinin Özeti	41
Tablo 18. KCF'ye İlişkin LR Analizi Sonucu Elde Edilen Sınıflandırma Tablosu	42
Tablo 19. KCF'ye İlişkin Kümelerde Yer Alan Öğrencilerin Frekans ve Yüzdeleri	42
Tablo 20. KCF'ye İlişkin Kümeleme Analizi Sınıflandırma Tablosu	43
Tablo 21. KCF'ye İlişkin ROC Analizinden Elde Edilen Değerler	43
Tablo 22. İkinci Alt Probleme İlişkin Bulgular	44
Tablo 23. ÇSF ve KCF'nin LR Yöntemiyle Elde Edilen Sınıflama Doğrulukları	45
Tablo 24. ÇSF ve KCF'nin Kümeleme Analizi Yöntemiyle Elde Edilen Sınıflama Doğrulukları	45
Tablo 25. ÇSF ve KCF'nin ROC Analizi Yöntemiyle Elde Edilen Sınıflama Doğrulukları ..	46

ŞEKİL LİSTESİ

Şekil 1. ROC Eğrisi Örneği	20
Şekil 2. ÇSF'ye İlişkin ROC Eğrisi	38
Şekil 3. KCF'ye İlişkin ROC Eğrisi.....	44



KISALTMALAR

ÇSF: Çoktan Seçmeli Form

KCF: Kısa Cevaplı Form

LR: Lojistik Regresyon

ROC: Receiver Operating Characteristics

AUC: Area Under Curve



BİRİNCİ BÖLÜM

GİRİŞ

Araştırmanın bu bölümünde problem durumuna, araştırmanın amacı ve önemine, araştırmanın problemine ve alt problemlerine, sayıltularına, sınırlılıklarına ve tanımlara yer verilmiştir.

1.1. Problem Durumu

Değişen, gelişen ve günden güne zorlaşan dünya koşulları içerisinde ayakta kalabilmek için insanların kendilerini yenilemeleri gerekmektedir. Hâlihazırda sahip oldukları olumlu özelliklerini, ilgi ve yeteneklerini geliştirmeleri, olumsuz özelliklerini olumlu hâle çevirmeleri veya sahip olmaları beklenen özellikleri edinmeleri, bu koşullara uyum sağlayabilmeleri açısından fazlaca önem taşımaktadır. Çünkü her geçen gün nitelikli bireylere duyulan ihtiyaç artmaktadır. Yaşamın her alanında, her sektörde belirlenen standartlara, istenen koşullara uygun bireylere ihtiyaç duyulmaktadır. Gerek özel kuruluşlar gerekse kamu kurum ve kuruluşları, içerisinde buldukları sektörün veya meslek alanının gerekliliklerine uygun, nitelikli bireyleri istihdam etmek istemektedirler. Bireyler ‘uygun aday’ olma yolunda gayret gösterirken, kurum ve kuruluşlar da ‘doğru aday’ seçmek durumundadırlar. Adayların uygun olup olmadığına ise kurum veya kuruluşların belirledikleri standartlar yön vermektedir. Burada standartla anlatılmak istenen bireylerdeki başarı veya yetenek düzeylerini belirleyebilmek adına tanımlanan özelliklerin tamamıdır. Standartların doğru tanımlanmasının önemi, adayların doğru aday olup olmadığı konusunda kesin bir yargıya varma sürecinde etkin bir rol oynayacak olmasından kaynaklanmaktadır. Alınan kararların yerinde ve isabetli olması, bu kararların alınmasında dayanak oluşturan bilgilerin yeterli ve sağlam olmasıyla yakından ilgilidir (Baykul, 2015). Eğitim kurumlarının öğrenci seçerken, işveren kurumlar çalışan ararken hatta ev sahipleri evlerini kiralarken kendilerine, kurumlarına veya meslek alanına uygun olarak bir dizi kural belirlemekte ve bu doğrultuda karşılaştıkları adayları elemektedir. Bu yüzden yanlış standartların belirlenmesi demek yanlış adayların seçilmesi anlamına gelmektedir. Başka bir deyişle, bir seçme gerçekleştirilecekse, öncelikle ölçme işlemi, ardından da belirlenen standartlar doğrultusunda bir karara varmak amacıyla değerlendirme işlemi yapılmaktadır. Bunun sonucunda ölçme işlemi uygulanan kişiler standartlara uygun ise ‘doğru aday’ olmakta, aksi takdirde elenmektedir.

Bir karar alma, yargıya varma süreci olarak değerlendirme, öncesinde mutlaka bir ölçme işlemi ile temellendirilmiştir. Alanyazına bakıldığında bu kavramlar farklı şekillerde tanımlanmaktadır. Ölçme Baykal'a (2007) göre "Olayların, öznelerin ve nesnelerin önceden belirlenmiş kurallara göre simgelerle özetlenmesi" olarak tanımlanırken; Büyüköztürk, Çakmak, Akgün, Karadeniz ve Demirel'e (2019) göre "Birey ya da nesnelerin niteliklerinin uygun araçlar kullanılarak gözlenmesi ve gözlem sonuçlarının sembollerle ifade edilmesi" olarak tanımlanmaktadır. Değerlendirme ise kararların, ölçmeden elde edilen sonuçlar ile bir ölçütün karşılaştırılması sonucunda verilmesi işlemidir (Turgut ve Baykul, 2019). Yani bir değerlendirme sürecinde ölçme işlemi, ölçüt ve ölçme işleminden elde edilen sonucun ölçüt ile karşılaştırılmasıyla varılan karar yer almaktadır. Açıkça ifade edilmese bile, her değer yargısının temelinde bir ölçme sonucu ile bir ölçütün karşılaştırılması yer alır (Tekin, 2019). Karar aşamasında ölçme sonucunun karşılaştırıldığı ölçüt yani standart ne kadar iyi belirlenirse, ilgili alanın amaçlarıyla ne kadar uyum gösterirse, varılan karar da o kadar isabetli olacaktır. Çünkü başarılı/başarısız veya geçti/kaldı sınıflamaları da bu doğrultuda belirlenmektedir. Söz konusu ölçütler bazen mutlak, bazen ise bağıl olabilmektedir. Mutlak ölçüt esas alınarak yapılan değerlendirmeler "mutlak değerlendirme", bağıl ölçüt esas alınarak yapılan değerlendirmeler de "bağıl değerlendirme" olarak tanımlanmaktadır (Gipps, 1994; Nitko, 1996; Atkin ve diğerleri, 2001). Bu kavramlara bakıldığında, mutlak değerlendirmede bireyin grup içerisindeki başarısından ziyade bireysel olarak başarısının ön planda olduğunu söylemek mümkündür. Çünkü bu değerlendirme türünde öğrencilerin her birinin başarısının, gruptaki diğer kişilerin başarısından bağımsız bir şekilde değerlendirilmesi söz konusudur (Nartgün, 2007). Alanyazına bakıldığında mutlak değerlendirme, bir alanda yeterli görülebilmek için farklı yöntemlerle ortaya konmuş olan minimum bir standart (kesme puanı) doğrultusunda yapılan değerlendirme şeklinde ifade edilmektedir (Erkuş, 2006). Bağıl değerlendirmede ise bir bireyin grup içerisindeki diğer bireylerle karşılaştırıldığındaki başarısı ön plandadır. Yani bağıl değerlendirmede kullanılan ölçütleri öğretmenden ziyade öğrenciler belirlemektedir (Tekin, 2019). Örneğin, yapılan herhangi bir sınavda öğrencilerin başarı ortalamasına bakılarak ve standart sapmayı birim olarak çeşitli notlara karşılık gelen puan aralıklarını belirlemek bağıl değerlendirme iken; toplamda 100 puan alınabilecek bir testten 70 puan ve üzeri alanları başarılı kabul etmek mutlak değerlendirme değildir.

Elbette bu süreçte ölçütü belirlenen ölçme aracının da birtakım niteliklere sahip olması beklenmektedir. Ölçme aracında bulunması istenen bu nitelikler, ölçme işleminin sonuçlarının daha güçlü dayanaklarla sunulabilmesini, ölçme aracının da verilen kararların da daha

uygulanabilir olmasını ve ölçümlerin daha hatasız olmasını sağlayacaktır. Yapılacak herhangi bir ölçme işleminde kullanılan ölçme aracında olması gereken bu nitelikler geçerlik, güvenirlik ve kullanılabilirlik olarak üç başlık altında toplanabilir (Tekin, 2019). En genel tanımıyla geçerlik, bir ölçme aracının, ölçülmek istenen özelliği başka değişkenlerle karıştırmadan ölçebilmesi (Thorndike ve Hagen, 1961; Turgut ve Baykul, 2019), yani başka bir deyişle bu ölçme aracının, kullanıldığı amaca hizmet etme derecesidir. Eğitimde kullanılan ölçme araçları için söz konusu olabilecek geçerlik türleri kapsam geçerliği, yapı geçerliği, yordama geçerliği ve görünüş geçerliği olarak dört temel başlık altında toplanabilir. Bu geçerlik türlerini özetlemek gerekirse;

Kapsam Geçerliği: Ölçme aracındaki her bir maddenin ölçülmek istenen konu ve davranışların bir örneklemini olduğunun gösterilmesidir (Cronbach ve Meehl, 1955). Çünkü bir ölçme aracı ile öğrenciye kazandırılmak istenen tüm davranışların yoklanması mümkün değildir. Ancak ölçme aracındaki maddeler, ölçülmek istenen davranışların kapsam olarak bir temsilini oluşturmalıdır. Bu temsili derecesi de kapsam geçerliliği olarak tanımlanmaktadır.

Yapı Geçerliği: Bir ölçme aracının psikolojik bir yapıyı ya da özelliği, ne derecede ortaya koyabildiğidir (Anastasi, 1988). Yani yapı geçerliği, testteki maddelerin puzzle parçaları gibi kendi içlerinde bir uyum sağlamaları ve bir aradayken bir bütün oluşturabilmeleri olarak düşünülebilir.

Yordama Geçerliği: Bilinen bir durumdan hareketle, bilinmeyen durumlara ilişkin tahminde bulunma işlemi şeklinde tanımlanabilir ve ölçütün bu özelliği taşımaması, yapılan yordamaların da hatalı olmasına sebep olacaktır (Turgut, 1992; Tekin, 2019; Baykul, 2015; Aiken, 2000).

Görünüş Geçerliği: Bir testin görünüşünün, ölçmek istediği şeyle uyumlu olması olarak tanımlanabilir (Tekin, 2019). Yani görünüş geçerliği merkezine testin ne ölçtüğünden ziyade, ne ölçüyor görüldüğünü almaktadır.

Ölçme araçlarında bulunması istenen özellikler denilince akla gelen önemli bir diğer kavram da hiç şüphesiz güvenirliktir. Güvenirlik her şeyden önce geçerlik için gerekli koşul olması bakımından önem taşımakta, hatta aslında bu özelliğiyle geçerliğin bir parçası olmaktadır. Bir ölçme aracının geçerli olabilmesi için ön koşul, o ölçme aracının güvenilir olmasıdır. Başka bir deyişle, bir ölçme aracının güvenilir olması, geçerli olması için gereklidir ama yeterli değildir. Yani bir ölçme aracı güvenilir olduğu hâlde geçerli olmayabilir ancak

güvenilir olmayan bir ölçme aracı geçerli de olamaz. Güvenirlik kavramının farklı anlamlarda tanımları da mevcuttur. Kararlılık anlamında bakıldığında güvenirlik, aynı ölçme aracından uygun zaman aralıkları belirlenerek elde edilen ölçme sonuçlarının benzer veya aynı olması olarak tanımlanabilir. Tutarlılık anlamında bakıldığında güvenirlik, ölçme aracının tekrar tekrar kullanıldığında benzer veya aynı sonuçları vermesi olarak tanımlanabilir. Duyarlılık anlamında bakıldığında ise güvenirlik, ölçme aracıyla yapılan ölçümlerin daha hassas olabilmesi için birimlerin olabildiğince küçük alınması olarak tanımlanabilir. İçerdiği hatalar anlamında bakıldığında güvenirliğin, bir ölçme aracının, kaynağı bilinmediği için giderilemeyen tesadüfi hatalardan arınlık derecesi olarak tanımlanması da mümkündür (Turgut ve Baykul, 2019). Ölçme sonuçlarına karışan hatalar sabit hata, sistematik hata ve tesadüfi hata olmak üzere üç başlık altında incelenebilir (Baykul, 2015; Özçelik, 1992; Turgut, 1992).

Sabit Hata: Her bir ölçme işleminde aynı miktarda yapılan hatadır (Büyüköztürk, Çakmak, Akgün, Karadeniz ve Demirel, 2019).

Sistematik Hata: Tüm ölçmeler için sabit olmayan ancak her bir ölçme için yordanabilir nitelikte olan, puanlayıcı yanlılıklarını yansıtan hatalardır (Tekin, 2019).

Tesadüfi Hata: Kaynağı, yönü (pozitif mi negatif mi) ve bir ölçmedeki miktarı sabit ve sistematik hataların aksine net olarak belirlenemeyen hatalardır (Baykul, 2015).

Güvenirlikte temel alınan hata tesadüfi hatalardır. Bunun temelinde, bu hataların ölçme sonuçlarına pozitif veya negatif olarak yansıyabilmesi, kontrol altına alınamaması, hatanın nedeninin ölçmeyi yapan kişinin dikkatsizliği, ölçme ortamı, şans başarısı veya ölçme yapılan bireylerin o anki psikolojileri gibi pek çok etkenden kaynaklanabilmesi yatmaktadır (Büyüköztürk, Çakmak, Akgün, Karadeniz ve Demirel, 2019).

Bir ölçme aracında olması beklenen bir diğer özellik ise kullanılabilirliktir. Kavram olarak kullanılabilirlik, bir ölçme sürecinde baştan sona kadar ölçme aracının sahip olması gereken birtakım nitelikleri kendi içinde barındırmaktadır. Tekin (2019) bu kavramı, bir ölçme aracının geliştirilmesi, çoğaltılması, uygulanması ve puanlanması bakımından kolay ve ekonomik olması olarak tanımlamaktadır. Ölçme araçlarının sahip olması beklenen bu özellikleri taşıması, bu ölçme araçlarının kullanıldığı amaca hizmet etme sürecinde atılan adımların, verilen kararların yerinde olması açısından da büyük önem taşımaktadır.

Eđitim s¼recinde verilecek kararların, ge¼erliđi ve g¼venirliđi y¼ksek olan ¼l¼mlere dayanması ¼nem tařıtmaktadır. Nihayetinde eđitim sisteminin ama¼larından biri de ¼đrencilerin sahip oldukları ¼zelliklere ve sergiledikleri performanslara g¼re bir ¼st eđitim kurumuna yerleřtirilmesidir. Bu ama¼ dođrultusunda eđitim kurumları kendi standartlarını belirlemekte ve belirlenen standartlara uygun ¼đrencileri se¼mektedir. S¼z konusu standartların dođru belirlenememesi, uygun ¼đrencilerin se¼ilememesine veya ¼đrenciler hakkında yanlıř kararlar alınmasına, akabinde de s¼re¼te birtakım aksaklıkların yařanmasına sebep olacaktır. Kesme puanına dayalı olarak yapılan deđerlendirme iřlemlerinde, testi alan bireylerin sınıflanması s¼z konusu olduđundan, bu sınıflamanın dođruluđu, ger¼eđi yansıtması daha da ¼nemli bir h¼l almaktadır. Bu t¼r durumlarda olumsuzlukların yařanmaması adına, her bir basamakta yapılan ¼l¼me ve deđerlendirme iřlemlerinin titizlikle y¼r¼t¼lmesi gerekmektedir. ¼¼nk¼ ¼l¼me ve deđerlendirme ¼alıřmaları, ge¼erli ve g¼venilir bilgilerin ¼retilmesi ve bu bilgilerden yola ¼ıkılarak objektif ve yerinde kararların alınması bakımından b¼y¼k ¼nem tařıtmaktadır (Atılđan, 2014). Bu ¼alıřmalar birtakım ¼l¼me ara¼ları, y¼ntem ve teknikler kullanılarak yapılmaktadır. Eđitimde alınan kararların dođru ve isabetli olması da uygun ¼l¼me ara¼larının kullanılmasıyla bađlantılıdır (Turgut, 1992).

Eđitim alanında ¼đrencilerle ilgili kararların verilmesinde kullanılan ¼l¼me ara¼ları genellikle bařarı testleridir. Bařarı testlerini, i¼erdikleri maddelerin t¼r¼ne g¼re yazılı yoklamalar, s¼zl¼ yoklamalar, kısa cevaplı testler, eřleřtirme maddeleri, dođru-yanlıř testleri, ¼oktan se¼meli testler olmak ¼zere altı grupta toplamak m¼mk¼nd¼r (Tekin, 2019). Bařarı testlerinde en ¼ok kullanılan madde t¼r¼ ¼oktan se¼meli maddelerdir. ¼oktan se¼meli maddeler, hazırlanışı a¼ısından kolay olmayan maddelerdir. Bu maddeleri hazırlayan ¼đretmenlerin soru hazırlama becerisine, ¼l¼len davranıřlarla ilgili ¼đrencilerin sahip olabilecekleri kavram yanılıđları bilgisine hakim olmaları, hatta se¼enekleri de bu dođrultuda ¼zenle hazırlamaları gerekmektedir. Bu y¼zden bu maddelerin hazırlık s¼re¼leri zaman alıcıdır. Ancak bu maddeler ¼đrencilerin yazı yazmasını gerektirmeyip yalnızca okuma ve iřaretleme yapmalarını gerektirdiđinden, ¼oktan se¼meli maddelerden oluřan ¼l¼me ara¼ları hem yazı g¼zelliđi gibi tesad¼fi hataların ¼l¼me sonu¼larına karıřmamasını sađladıđından hem de kısa s¼rede daha ¼ok soru yanıtlanabilmesine olanak tanıdıđından kapsam ge¼erliđi y¼ksek olan ¼l¼me ara¼larıdır. Aynı zamanda puanlama ařamasında kabul edilen cevaplar ¼znellik i¼ermediđinden objektif puanlamalar yapılabilmekte, hatta cevaplar optik formlara iřaretlendiđi takdirde puanlama optik okuyucular tarafından hızla yapılabilmektedir. Bu da fazla sayıda bireyin d¼hil edildiđi ¼l¼melerde olduk¼a avantajlı olmaktadır. ¼zetle ¼oktan

seçmeli maddeler, diğer maddelere göre daha kolay uygulanabilmesi, objektif olması, geçerli ve güvenilir gözlemler yapmaya imkân tanınması nedeniyle yaygın olarak kullanılmaktadır (Cronbach, 1990; Conway, 2006; Pawar, 2004). Fakat özellikle sonunda bireyler açısından ciddi kararların alındığı sınavlarda, hakkında olumlu karar verilen bireylerin ölçülen ve sahip olduğu görülen özelliklere gerçekten sahip olması önem taşımaktadır. Bu da çoktan seçmeli maddelerin en çok eleştiri alan yönlerinden şans başarısını akla getirmektedir. Cevaplayıcının doğru cevabı şansla bulması ve ilgili maddeden aldığı puanı şansla kazanması şans başarısı olarak tanımlanmaktadır (Turgut ve Baykul, 2019). Bir bireyin tahmin yoluyla veya tamamen şans eseri bir maddeyi doğru cevaplama, o maddenin ölçtüğü özelliğe sahip olduğu anlamına gelmemesine rağmen, bu maddelerden puan alması, o özelliğe sahip olduğu anlamını taşımaktadır. Bu açıdan bakıldığında, ilgili maddeyi gerçekten ölçülen özelliğe sahip olduğu için doğru yapan bir bireyle, bu bireyin bir farkı kalmamaktadır. Çoktan seçmeli maddelerin özellikle şans başarısı kaynaklı olarak aldığı birtakım eleştiriler, insanları farklı madde türlerini araştırmaya yöneltmiştir. Kısa cevaplı maddeler de bu madde türlerinden biridir. Cevabının bir sözcük, bir sayı veya kısa bir cümle ile sınırlandırıldığı bu maddeler, açık uçlu maddelerdeki gibi öğrencinin cevaba ilişkin pek çok şey yazmasını gerektirmeyecek şekilde yapılanmıştır. Bu da verilen cevapların değerlendirildiği aşamada nesneliği sağladığından, kısa cevaplı maddelerin objektif maddeler olduğu düşünülebilir. Hatta kısa cevaplı maddeler, cevabı öğrenci tarafından hatırlanıp yazmaya dayanan tek objektif madde türüdür demek de mümkündür (Tekin, 2019). Ayrıca istenen cevapların kısa olması, verilen süre içerisinde daha çok maddenin yanıtlanmasına olanak tanıyacağından, kapsam geçerliği yüksek sonuçlar elde edilmesini sağlayacaktır. Benzer şekilde, yanıtlayıcıların doğru cevabı görerek hatırlaması sonucu bulması veya tahminle cevaplayarak şans eseri bulması gibi bir ihtimal bulunmamaktadır. Bu durum da, elde edilen puanların geçerlik ve güvenilirliğine olumlu katkı sağlamaktadır. Bunun yanında, çoktan seçmeli maddelere kıyasla, doğru cevabı hatırlayarak yazmanın daha üst düzey beceri gerektirdiği göz önüne alındığında, kısa cevaplı maddelerin ezber ve bilgiye dayalı davranışlardan daha fazlasını ölçebileceği söylenebilir. Örneğin, aşağıda verilen maddeyle en az uygulama düzeyindeki bir davranış ölçülmektedir (Tekin, 2019):

Madde

Aşağıdaki cümlenin öğelerinin (altı çizilmiş kısımların) adlarını, onlar için ayrılan boşluklara yazınız.

“İhtiyar delikanlı, polisten kaçan çocuğu yakalayiverdi.”

(A)

(B)

(C)

Maddenin Cevabı

(A) Özne

(B) Nesne

(C) Yüklem

Öte yandan, kısa cevaplı maddelerin hazırlanma süreci de diğer madde türlerine göre daha kolaydır ve gerekli teknikler kısa sürede edinilebilir (Turgut ve Baykul, 2019). Kısa cevaplı maddeler, madde içerisinde verilen bir boşluğu doldurmayı gerektirebileceği gibi, öğrencinin kısa bir tanım yazmasını da gerektirebilir. Bu aşamada eksik cümle olarak yapılandırılabilir olan kısa cevaplı maddeler, soru cümlesi olarak da yapılandırılabilir. Yani kısa cevaplı maddeler, gerekli esaslar dikkate alındığında hem hazırlama, hem uygulama hem de puanlama açısından avantajlı maddelerdir.

Özetle, eğitim alanında kritik kararların alınmasında en çok kullanılan ölçme araçları olan başarı testlerinde daha çok tercih edilen çoktan seçmeli maddeler ile bu madde türüne iyi bir alternatif olabilecek niteliklere sahip olan kısa cevaplı maddeler ele alınmış ve iki madde türünün detayları açıklanmıştır. Öte yandan, öğrencilerin akademik olarak başarı düzeyinin belirlenmesinde veya öğrenciler hakkında karar verilmesinde çeşitliliğin sağlanması önemlidir. Böyle olduğu durumlarda da hangi ölçüm ile nasıl karar verileceği noktasında farklılıklar ortaya çıkabilir. Belli bir özelliğe yönelik değerlendirmenin farklı şekillerde yapılması hâlinde elde edilen sonuçların isabetliliğinin / tutarlılığının incelenmesi gerekmektedir. Bu sebeple sınıflama tutarlılığına ilişkin üç farklı yöntemin karşılaştırılması ele alınmıştır.

1.2. Araştırmanın Amacı ve Önemi

Bu çalışmada aynı özelliği ölçen bir testin farklı iki formunun sınıflama doğruluklarının üç farklı istatistiksel yöntemle karşılaştırılması amaçlanmıştır. Bu doğrultuda her bir formda farklı yöntemlerle gerçekleştirilen sınıflamalar karşılaştırılmıştır. Ayrıca öğrencilerin başarılı sayılmasında formun çoktan seçmeli veya kısa cevaplı olmasının etkisinin belirlenmesi de çalışmanın amaçları arasındadır. Bu sebeple farklı formlardan aynı yöntemle gerçekleştirilen sınıflamalar da karşılaştırılmıştır. Geleneksel yöntemlere alternatif olabilecek üç farklı istatistiksel yöntemin seçilmesi ile elde edilen sonuçların daha geçerli ve güvenilir olması

hedeflenmektedir. Araştırmaya yön verecek olan problem cümlesi ve alt problemler de bu amaçlar göz önüne alınarak oluşturulmuştur.

Çağımızda hızla ilerleyen teknoloji, günden güne yaşamımızın her alanında karşımıza çıkmaktadır. Bazen eski alışkanlıklarımızı terk etmek istemesek de, uzun veya uğraştırıcı da olsa en kısa yol olarak görüp yine bildiklerimizi kullanma eğilimimiz olsa da teknolojik gelişmeler hem çok fazla alanda kendini göstermekte hem de bu gelişmeleri takip etmek zorunda bırakacak kadar hızlı gerçekleşmektedir. Geçmişten günümüze bakıldığında bunun en bariz örneklerinden biri şüphesiz ki bilgisayar kullanımınıdır. Bilgisayarın önemli alanlarda kullanılmaya başlamasıyla birlikte, sağladığı kolaylıklar da daha yakından görülmüş ve zamanla günlük yaşamın veya iş yaşamının bir parçası hâline gelmiştir. Bu çalışmada geleneksel standart belirleme yöntemlerine alternatif olabilecek istatistiksel yöntemlerin seçilmesinin ilham kaynağı da bununla açıklanabilir. Elbette geleneksel yöntemlerin standart belirlemedeki yeri ve önemi yadsınamaz. Ancak bu yöntemlerin merkezinde uzman görüşlerinin yer alması, elde edilen sonuçların objektiflikleriyle ilgili bir soru işaretini de beraberinde getirmektedir. Özellikle standart belirleme yöntemlerinde zaman, maliyet ve uygulama aşamasının tamamında kolaylık sağlayan bir yöntemin olmaması, seçilecek olan standart belirleme yönteminin önemini bir kat daha artırmaktadır. Bu çalışmanın da üç farklı istatistiksel yöntemi karşılaştırması ve bunu aynı özelliği ölçen iki farklı form üzerinden gerçekleştirmesi açısından literatüre katkı sağlayacağı düşünülmektedir. Ayrıca her ne kadar merkezi sınavlarda çok mümkün olmasa da, sınıf içi değerlendirmeler veya daha küçük gruplara yapılan ölçmelerde kısa cevaplı maddeler de kullanılabileceği için, bu tür uygulamalara da ışık tutan bir çalışma olacağı düşünülmektedir.

1.3. Problem Cümlesi

Çoktan seçmeli ve kısa cevaplı olarak iki farklı formu geliştirilen bir matematik testinin lojistik regresyon analizi, kümeleme analizi ve ROC analizi ile belirlenen sınıflama doğrulukları birbirinden farklı mıdır?

1.4. Alt Problemler

- 1) Çoktan seçmeli formun lojistik regresyon analizi, kümeleme analizi ve ROC analizi ile belirlenen sınıflama doğrulukları farklılaşmakta mıdır?
- 2) Kısa cevaplı formun lojistik regresyon analizi, kümeleme analizi ve ROC analizi ile belirlenen sınıflama doğrulukları farklılaşmakta mıdır?

- 3) Çoktan seçmeli ve kısa cevaplı formun lojistik regresyon analizi ile belirlenen sınıflama doğrulukları farklılaşmakta mıdır?
- 4) Çoktan seçmeli ve kısa cevaplı formun kümeleme analizi ile belirlenen sınıflama doğrulukları farklılaşmakta mıdır?
- 5) Çoktan seçmeli ve kısa cevaplı formun ROC analizi ile belirlenen sınıflama doğrulukları farklılaşmakta mıdır?

1.5. Sayıtlar

Öğrenciler tarafından test maddelerine verilen cevapların bilgi düzeylerini yansıttıkları varsayılmıştır. Ayrıca pandemi döneminden dolayı çevrimiçi ortamda toplanan verilerde öğrencilerin vermek istedikleri cevapları doğru bir şekilde sisteme girebildikleri, ilgili maddeleri cevaplayabilecek kadar cihaz (bilgisayar/tablet/telefon) kullanma becerisine sahip oldukları varsayılmıştır. Yine öğrencilerin cevaplama anları kontrol altına alınmadığı için başka kaynaklardan yararlanmadıkları ve test maddelerinin başkası tarafından değil, öğrencilerin kendileri tarafından cevaplandığı varsayılmıştır.

1.6. Sınırlılıklar

Pandemi sürecinin uzaması nedeniyle tez yazma sürecinde özellikle veri toplama aşamasında ciddi aksaklıklar yaşanmıştır. Bu yüzden sürecin başında hem kâğıt-kalem yoluyla elden, hem de çevrimiçi olarak toplanması planlanan veriler, yalnızca çevrimiçi ortamda toplanmıştır. Bazı öğrencilerin cevaplarını sisteme nasıl gireceklerine dair cihaz kullanma becerisi temelli sorun yaşamasından ötürü veri kayıpları yaşanmıştır.

1.7. Tanımlar

Kesme Puanı: İstenilen performans düzeyine ulaşanlarla, bu düzeye henüz ulaşamayanların ayırt edildiği performans noktasıdır. Minimum geçme puanı (MGP) da denilmektedir.

İKİNCİ BÖLÜM

KAVRAMSAL ÇERÇEVE VE İLGİLİ ARAŞTIRMALAR

Başarı testlerinde belirlenen ölçütlere yani standartlara göre bireylerin istendik davranışlara sahip olup olmadıkları veya hangi düzeyde sahip oldukları ortaya konmakta, sonunda da bireyler hakkında başarılı-başarısız, geçti-kaldı, yeterli-yetersiz gibi yargılara varılmaktadır. Başka bir deyişle, bireyler aldıkları puanlara göre başarı düzeyleri açısından farklılaştıkları için bu doğrultuda sınıflandırılmaktadır (Cizek, 1993). Yapılan bu sınıflamanın doğruluğu ölçüsünde elde edilen ölçme sonuçları amaca daha iyi hizmet etmektedir. İstenilen bir performans düzeyine ulaşmış olanlarla henüz ulaşamamış olanların ayrıldığı performans noktası ise kesme puanı olarak adlandırılmaktadır (Gündeğer, 2012). Standart belirleme, bireylerin farklılaştığı yeri tespit etmeye dayandığından ve bu doğrultuda önemli kararlar alındığından, bu süreçte kullanılacak olan standart belirleme yönteminin seçilmesi de özenle yapılmalıdır (Tülübaş, 2009).

2.1. Standart Belirleme Yöntemleri

Standart belirleme yöntemleri en temelde geleneksel ve alternatif yöntemler olmak üzere iki başlık altında incelenebilir. Alternatif yöntemlerde, bu çalışmada kullanılan lojistik regresyon, kümeleme ve ROC analizleri detaylı olarak ele alınmıştır.

2.1.1. Geleneksel Yöntemler

Geleneksel standart belirleme yöntemleri Jaeger (1991) tarafından öğrenci merkezli ve test merkezli yöntemler olmak üzere iki alt başlıkta detaylandırılmıştır. Bu başlıklar, yargıcıların kararlarındaki odak noktalarını vurgulamaktadır. Öğrenci merkezli yöntemlerde öğrencilerin yeterlik düzeylerine yönelik kararlar alınırken, test merkezli yöntemlerde testin içeriğine yönelik kararlar alınmaktadır.

2.1.1.1. Öğrenci Merkezli Yöntemler

Öğrenci merkezli standart belirleme yöntemlerinde yargıcıların öğrencilerin yeterlik düzeyleri hakkında bilgi sahibi olmaları gerekmektedir. Çünkü bu yöntemlerde öğrencilerin testi almadan önce yargıcılar tarafından yetenek düzeylerine göre gruplandırılması söz konusudur. Zıt gruplar yöntemi ve sınır grup yöntemi, öğrenci merkezli olan standart belirleme yöntemlerindedir.

Zıt Gruplar Yöntemi

Testi alan bireyleri iyi tanıyan yargıcıların, bu grupta yeterli ve yetersiz olarak düşündükleri kişilerden oluşan iki zıt gruba aynı test uygulanmaktadır. Brandon'a (2002) göre, her bir gruptan elde edilen toplam puanların dağılımlarından hareketle kesme puanı belirlenmektedir. Testi alan bireylerin yeterli-yetersiz sınıflandırılma olasılıklarındaki eşitlik durumunda ise lojistik regresyonla kesme puanı belirlenebilmektedir.

Sınır Grup Yöntemi

Yargıcıların testi alacak olan bireyleri yeterli, sınırdaki ve yetersiz olmak üzere üç gruba ayırmasına ve sınır grupta yer alan öğrencilerin testten elde ettikleri puanların aritmetik ortalamasının veya ortancasının kesme puanı olarak belirlenmesine dayanan bir yöntemdir (Livingston ve Zieky, 1982).

2.1.1.2. Test Merkezli Yöntemler

Test merkezli standart belirleme yöntemlerinde testin bütünü ile test maddeleri odak noktası olmaktadır. Bu yöntemlerden olan Angoff, Nedelsky ve Ebel yöntemleri, kısaca açıklanmıştır.

Angoff Yöntemi

Yargıcılardan testteki her bir maddenin sınırdaki bir öğrenci tarafından doğru yanıtlanma olasılığının yüzde kaç olduğunu tahmin etmeleri istenir (Cizek ve Bunch, 2007). Tüm maddelere yönelik elde edilen bu kestirimlerin toplanıp madde sayısına bölünerek o yargıcıya ait kesme puanı belirlenir. Testin kesme puanı ise, tüm yargıcıların kesme puanlarının ortalaması alınarak belirlenmektedir. Evet/Hayır (Yes/No) yöntemi, Genişletilmiş Angoff yöntemi ve Değiştirilmiş Angoff yöntemi gibi varyasyonları bulunmaktadır.

Nedelsky Yöntemi

Yargıcılarından sınırdaki öğrencilerin testteki her bir maddenin seçeneklerinden kaç tanesini eleyebileceklerini tahmin etmeleri istenir (Livingston ve Zieky, 1982). Elenmeyen seçenek sayılarına bakılarak bu öğrencilerin ilgili maddeyi doğru cevaplama olasılıkları hesaplanmaktadır. Testteki tüm maddeler için hesaplanan olasılıkların toplamı, o yargıcının en düşük geçme puanını; yargıcıların en düşük geçme puanlarının ortalaması da testin geçme puanını oluşturmaktadır. Seçenekler üzerinden kesme puanı belirleme söz konusu

olduğundan, bu yöntem yalnızca çoktan seçmeli maddelerden oluşan testler için uygulanabilir.

Ebel Yöntemi

Yargıcıların testteki her bir maddeyi uygunluk ve zorluk bakımından değerlendirdikleri bir yöntemdir. Uygunluk açısından 'gerekli, önemli, kabul edilebilir ve tartışılabilir' olmak üzere dört düzey yer alırken, zorluk açısından 'kolay, orta ve zor' olmak üzere üç düzey yer almaktadır (Poggio, Glasnapp ve Eros, 1982; Ebel, 1972). Yargıcılar testin yapısını dikkate alarak her bir maddeyi 3x4'lük tabloya yerleştirdikten sonra, sınırdaki öğrenciler dikkate alınarak her bir maddenin doğru cevaplandırma yüzdeleri belirlenir. Daha sonra tablonun her bir hücresinde yer alan madde sayısı ile yüzde çarpılarak hücreye ilişkin bir puan elde edilir ve bu puanların toplamı yargıcılara ait minimum geçme puanını verir (Livingston ve Zieky, 1982). Tüm yargıcılar için elde edilen puanların ortalaması ise testin geçme puanını verir.

2.1.2. Alternatif Yöntemler

Geleneksel yöntemlerde yargıcıların (uzmanların) kanılarına dayanan bir süreç işlediğinden, bu yöntemlerin daha öznel olduğunu söylemek mümkündür. Alternatif yöntemler ise diğerlerine kıyasla daha nesnel çünkü istatistiksel yöntemlerle yürütülen bir süreç söz konusudur. Çok değişkenli istatistiksel analizlerin amaçlarından biri olan sınıflama, kesme puanı belirlemeden önce ne kadar yerinde kararlar alınacağıyla ilgili, belirlenecek olan kesme puanının ne kadar isabetli olacağıyla ilgili önemli ipuçları elde etmemizi sağlamaktadır. Bu çalışmada da sınıflama yapma amacıyla çok değişkenli istatistiksel tekniklerden Lojistik Regresyon Analizi, Kümeleme Analizi ve ROC Analizinden yararlanılmıştır. Bu yöntemlerin detaylı açıklamalarına aşağıda yer verilmiştir.

2.1.2.1. Lojistik Regresyon Analizi

Bilimsel araştırmaların amaçlarından bir tanesi de, olaylar arasındaki neden-sonuç ilişkilerini ortaya koymaktır. Elde edilen sonuç yordanan (bağımlı) değişken, bu sonuç üzerinde etkili olan nedenler ise yordayıcı (bağımsız) değişkenler olarak ele alınmaktadır (Çokluk, Şekercioğlu ve Büyüköztürk, 2018). Bu değişkenlerin yapısı da analizde kullanılacak olan modeli belirlerken farklılıklara yol açmaktadır. Örneğin, doğrusal regresyon modellerinde bağımlı ve bağımsız değişkenlerin sürekli olması gerekmektedir. Ancak özellikle sosyal bilimlerde yapılan araştırmalara ilişkin elde edilen verilerdeki bağımlı

değişkenler kategoriktir. Örneğin, bir kişinin bir gruba üye olması veya olmaması, çalışır durumda veya çalışmıyor olması, bir öğrencinin bir programı tamamlaması veya tamamlamaması, başarılı ya da başarısız olması, bir bireyde öğrenme güçlüğüne sahip olması ya da olmaması şeklinde olabilir (Kılıç, 2000; Peng, Lee ve Ingersoll, 2002). Böyle durumlarda kullanılabilecek modellerden biri lojistik regresyon analizidir. Lojistik regresyon analizi, hem bağımlı değişkenin sürekli veya nicel olmak yerine kategorik veya sınıflamalı olduğu durumlarda uygulanabilir (Mertler ve Vannatta, 2005), hem de basit ya da çoklu doğrusal regresyon analizlerinin aksine, bağımlı ve bağımsız değişkenlerle hata varyanslarının normal dağılımı gibi varsayımların karşılanmasını gerektirmez. Ayrıca lojistik regresyon analizinin kategorik veri analizinde önemli bir yere sahip olmasının ardında, regresyon tipi mantığıyla ve diğer analizlere göre daha kullanışlı olmasının avantaj sağlaması yatmaktadır (Kılıç, 2000).

Lojistik regresyon analizi, adını bağımlı değişkene uygulanan logit dönüşümden almaktadır. Bu durum, hem tahmin hem de yorumlama sürecinde birtakım farklılıklara neden olmaktadır (Hair, Black, Babin, Anderson ve Tatham, 2006). Bu analizde bağımlı değişken iki kategorili bir nitel değişken ise, analiz ‘İkili Lojistik Regresyon (Binary Logistic Regression)’ adını alır (Can, 2019). Örneğin, öğrencilerin bir dersi geçme durumu ile ilgili olarak başarılı ve başarısız şeklinde sınıflandırıldığı durumlarda ikili lojistik regresyon kullanılmaktadır. Eğer bağımlı değişken ikiden fazla kategorili bir nitel değişken ise yapılacak olan analiz ‘Çok Kategorili Regresyon Analizi (Multinomial Logistic Regression)’ adını alır (Can, 2019). Örneğin, kişilerin sahip oldukları arabanın benzinli, dizel, LPG’li olarak sınıflandırıldığı durumlarda çok kategorili regresyon analizi kullanılmaktadır. Eğer bağımlı değişken ‘düşük, orta, yüksek sosyoekonomik düzey’ gibi sıralama ölçeğinde bir değişken ise sıralama ölçeğinde ise bu defa da analiz ‘Sıralı Regresyon Analizi (Ordinal Logistic Regression)’ adını alır (Can, 2019).

Lojistik regresyon analizinde, her bir birey için bağımlı değişkende yer alan kategorilerin (başarılı/başarısız, geçti/kaldı, yeterli/yetersiz gibi) gerçekleşme olasılığının hesaplanması söz konusudur. Başka bir deyişle, bu analiz bireylerin hangi gruba dâhil olduklarına ilişkin bir üyelik tahmini yapmayı amaçlar. Bu sayede bireylerin ait oldukları grupları doğru bir şekilde belirleyerek doğru bir sınıflama yapmak amaçlanmaktadır. Bu doğrultuda analizin sınıflandırma amacı olduğu gibi, bağımlı ve bağımsız değişkenlerin arasındaki ilişkileri belirlemek olduğu da ifade edilebilir (Mertler ve Vannatta, 2005).

Lojistik regresyon analizi, özellikle son yıllarda daha yaygın olarak kullanılmaktadır. Bu durumun sebepleri arasında, analize dâhil olan bağımlı değişkenin kategorik olmasına olanak tanınması, bağımsız değişkenlerin dağılımına ilişkin araştırmacılarca karşılanması gereken sayıltılar gerektirmemesi (Tabachnick ve Fidell, 1996) bulunmaktadır. Çünkü bu özellikler, lojistik regresyon analizini, benzer diğer analizlere göre daha esnek hâle getirmekte ve daha çok alanda uygulanabilmesine olanak tanımaktadır. Ayrıca analiz sonuçlarındaki parametrelerin kolaylıkla yorumlanabilmesi ve lojistik regresyon analizi yapmaya imkân tanıyan pek çok paket programın olması da bu analizin giderek yaygınlaşmasının temel sebepleri arasında gösterilebilir.

Bununla birlikte Lojistik Regresyon analizi ile elde edilen istatistiklerin yansız ve sapmasız olabilmesi için, örneklemlerin de büyük olması gerektiği belirtilmektedir. Özellikle ikiden fazla kategoriye sahip olan bir bağımlı değişkenin söz konusu olduğu durumlarda, hipotez testinin geçerli olabilmesi adına, her bir bağımsız değişken için en az 50 kişilik grup büyüklüğüne gereksinim duyulmaktadır. Farklı kaynaklarda bu sayının her bir bağımsız değişken için minimum 20, toplamda minimum 60 olması gerektiği belirtilmektedir. Bununla birlikte, Lojistik Regresyon analizi ile yapılan çözümlemede elde edilen matematiksel modelin yorumlanmasının daha kolay olduğunu belirtmekte fayda vardır (Akkuş ve Çelik, 2004; Grimm ve Yarnold, 1995; Kalaycı, 2005; Leech, Baret ve Morgan, 2005; Poulsen ve French, 2008; Tabachnick ve Fidel, 1996; Tatlıdil, 1996; Akt: Çokluk ve diğerleri, 2018).

2.1.2.2. Kümeleme Analizi

Kümeleme analizi, bir araştırmada ele alınan birimler aralarındaki benzerlikler doğrultusunda bu birimleri belirli gruplarda toplayarak sınıflandırmaya, ilgili birimlerde bulunan ortak özellikleri tespit etmeye ve bu sınıflara ilişkin genel tanımlar yapmaya olanak tanıyan bir yöntemdir (Kaufman ve Rousseuw, 1990). Değişkenleri, birbiriyle benzerlik gösteren alt gruplara ayırarak sınıflama yapmaya olanak tanınması sebebiyle pek çok alanda kümeleme analizinden yararlanılabilmektedir. Örneğin, eğitim alanında kişilere yönelik alt kişilik kalıpları belirleyerek bu kalıplarla uyum gösterecek şekilde eğitim programları geliştirme amacıyla kümeleme analizinden faydalanılabilir (Özdamar, 2004). Kümeleme analizinde amaç, herhangi bir gruplama yapılmamış olan veri matrisindeki gözlemleri benzerlikleri ölçüsünde gruplamaktır. Yapılan analiz sonuçlarında ise kendi içlerinde homojen, kendi aralarında heterojen grupların elde edilmesi beklenmektedir. Kümeleme analizinde verilerin normal dağılıma sahip olması gerektiği varsayımı olmakla beraber

normallik varsayımı prensipte kalmakta, uzaklık değerlerinin normalliği yeterli bulunmaktadır. Ayrıca kovaryans matrisine dair bir varsayım belirtilmemektedir (Tatlídil, 2002).

Kümeleme analizinde veri setinde yer alan gözlemlerin benzerliğine karar vermek için uzaklık matrisleri oluşturulmakta ve bu matrisler oluşturulurken birtakım uzaklık ölçüleri kullanılmaktadır. Bu uzaklık ölçülerinden bazılarına ilişkin açıklamalara aşağıda yer verilmiştir.

Euclid Uzaklığı

Euclid (Öklit) uzaklığı, en yaygın olarak kullanılan uzaklık ölçüsü olmakla birlikte çok boyutlu uzaydaki geometrik uzaklık olarak tanımlanabilir (Tatlídil, 2002).

Manhattan (City-Blok) Uzaklığı

Verilerin matristeki koordinatları arasındaki farkın mutlak değerleri toplanarak hesaplanan bir uzaklık ölçüsüdür.

Mahalanobis Uzaklığı

Mahalanobis uzaklık ölçüsü, doğrudan birleştirme yapmaya olanak sağlayan standart bir ölçüdür. Aralarında ilişki bulunan iki değişken için, hem standart sapmaları hem de bu iki değişken arasındaki kovaryans veya korelasyonu dikkate alan Mahalanobis uzaklığının kullanılması gerekir.

Kümeleme yöntemlerini en temelde “hiyerarşik kümeleme yöntemleri” ve “hiyerarşik olmayan kümeleme yöntemleri” olmak üzere iki başlık altında toplamak mümkündür.

2.1.2.2.1. Hiyerarşik Kümeleme Yöntemleri

Hiyerarşik kümeleme yöntemleri tipik olarak küçük örneklemeler ($n < 250$) için uygundur (Çokluk, Şekercioğlu ve Büyüköztürk, 2018). Bu yöntemler, incelenen veri setinde kaç grup olduğunun başlangıçta bilinmediği durumlarda uygulanabilen yöntemler olduğu, ayrıca veri setinde daha öncesinde gözlenmemiş olan ilişkileri de ortaya koyabilme imkânı verdiği belirtilmektedir (Anderberg, 1973). İki alt başlıkta toplanabilen bu yöntemlerden ilki ayırıcı (bölünmeli / ayrımlı / divisive) yöntemler, ikincisi ise birleştirici (toplamalı / agglomerative) yöntemlerdir.

2.1.2.2.1.1. Ayırıcı (Bölünmeli / Ayrımlı / Divisive) Yöntemler

Bu yöntemlerde veri setindeki tüm gözlemler başlangıçta tek bir küme olarak düşünülür. Daha sonra küme sayısı bir indirgenerek benzerlik matrisinin tekrar oluşturulması sonucunda benzer birimler bir araya getirilerek aşamalı bir şekilde, her bir gözlem ayrı ayrı birer küme olana kadar kümeleme işlemi gerçekleştirilir (Everitt, 1971).

2.1.2.2.1.2. Birleştirici (Toplamalı / Agglomerative) Yöntemler

Bu yöntemlerde ise veri setindeki gözlemlerin her biri başlangıçta ayrı ayrı birer küme olarak düşünülür. Daha sonra küme sayısı bir indirgenerek benzerlik matrisinin tekrar oluşturulması sonucunda benzer birimler bir araya getirilerek aşamalı bir şekilde, tüm gözlemler tek bir küme olana kadar kümeleme işlemi gerçekleştirilir (Everitt, 1971). Birleştirici yöntemlerde veri setindeki gözlemlerin birbiri ile birleştirilmesi sırasında farklı yaklaşımlar kullanılmaktadır ve bu yaklaşımların en yaygın olanları şöyledir:

Tek Bağlantı Kümeleme Yöntemi (En Yakın Komşuluk Yöntemi / Single Linkage / Nearest Neighbour Method)

Hiyerarşik kümeleme analizleri arasında en basit olan bu yöntemde uzaklıklar matrisi kullanılarak birbirine en yakın birey veya kümeler birleştirilmekte ve bu işlem art arda tekrarlanarak kümeleme işlemi yapılmaktadır. Bu yöntemin birbirinden yeterince ayrı olan kümeleri belirlemede iyi, birbirleri arasındaki farkı az olan kümeleri belirlemede ise yetersiz kaldığı bilinmektedir.

Tam Bağlantılı Kümeleme Yöntemi (En Uzak Komşuluk Yöntemi / Complete Linkage / Farthest Neighbour Method)

Tek bağlantı tekniğinden farklı olarak her bir kümedeki eleman çiftleri arasındaki uzaklıktan en fazla olanı ele alınmaktadır, geri kalan yönlerinde ise tek bağlantı tekniğiyle yüksek oranda benzerlik göstermektedir (Gren, 1989; Akt: Çelik, 2013).

Ortalama Bağlantı Kümeleme Yöntemi (Average Linkage Method)

Kümeleme işlemi tek bağlantı yönteminde birbirine olan uzaklıkları en az, tam bağlantı yönteminde ise birbirine olan uzaklıkları en fazla olan noktalardan başlanarak gerçekleştirildiğinden, bu yöntemler uç değerlere oldukça duyarlıdır. Ortalama bağlantı yöntemi de bu dezavantaj göz önüne alınarak tercih edilen bir yöntemdir. Çünkü bu

yöntemde, kümeleme ölçütü olarak oluşan kümeler içerisindeki birimlerin birbirine olan ortalama uzaklıkları kullanılmaktadır.

Ward's Bağlantı Kümeleme Yöntemi (Ward Linkage Method)

Diğer adı en küçük varyans yöntemi olan bu yöntemde amaç, adından da anlaşılacağı gibi oluşan kümeler içindeki varyansın minimum olmasını sağlamaktır. Bu amaca yönelik olarak da, diğer yöntemlerde olduğu gibi kümeler arası uzaklığın değil de, kümeler içi hata kareler toplamının hesaplanmasına dayanmaktadır. Hesaplanan bu değer minimum olmasıyla, kümeler içinde maksimum homojenlik sağlanacaktır. Birbirine en yakın homojenlikte olan gruplar birleştirilerek yeni kümeler oluşturulabilmekte ve böylelikle, birbiri arasında heterojenliği maksimum olan kümeler elde edilebilmektedir (Hair, Black, Babin, Anderson ve Tatham, 2006).

Merkezi Bağlantı Kümeleme Yöntemi (Kitle Merkezi Yöntemi / Centroid Linkage Method)

Merkezi bağlantı yönteminde, adından da anlaşılacağı gibi iki küme arasındaki benzerliğin ölçütü, bu kümelerin merkezleri arasındaki uzaklıktır. Çünkü küme merkezleri, kümedeki gözlemlerin ortalama değeri olarak nitelendirilebilir. Burada sözü edilen küme merkezleri sabit değildir, kümelerde yeni birleşmeler oldukça merkezler de yer değiştirmektedir (Çokluk, Şekercioğlu ve Büyüköztürk, 2018). Bu yöntem de, ortalama bağlantı yönteminde olduğu gibi diğer hiyerarşik yöntemlere kıyasla, uç değerlerden daha az etkilenmektedir (Hair, Black, Babin, Anderson ve Tatham, 2006).

2.1.2.2.2. Hiyerarşik Olmayan Kümeleme Yöntemleri

Hiyerarşik tekniklere kıyasla daha büyük veri kümelerine uygulanabilen hiyerarşik olmayan kümeleme yöntemlerinin, (Johnson ve Wichern, 1988) hiyerarşik olan kümeleme yöntemlerine göre belirgin farkları bulunmaktadır. Bu farkların başında da küme sayısının önceden belirlenmiş olması gelmektedir. Başlangıçta küme sayısı ile ilgili ön bilginin olduğu durumlarda ya da araştırmacının analiz öncesinde anlamlı olacak küme sayısını belirlediği durumlarda, hiyerarşik yöntemler yerine hiyerarşik olmayan yöntemler kullanılır (Anderberg, 1973). Diğer bir farklılık ise, dendrogram tablosu oluşturulmaması ve bir uzaklık veya benzerlik ölçüsü kullanılmamasıdır. Tüm bu farklılıklar, hiyerarşik olmayan yöntemlerin, hiyerarşik olan yöntemlere göre çeşitli açılardan avantajlı olmasını sağlamaktadır. Bu avantajlar kısaca şöyle sıralanabilir (Hair ve diğerleri, 2006):

1. Sonular duyarlılıđı verilerdeki aykırı deęerlerden, kullanılan uzaklık lusunden, alakasız veya uygunsuz deęiřkenlerin dâhil edilmesinden daha az etkilenmektedir.
2. Hiyerarřik olmayan yntemler son derece byk veri kmelerini algoritma yapıları geređi daha hızlı ve rahat bir biimde analiz edebilir, nk tm gzlemler arasında benzerlik matrislerinin hesaplanmasını gerektirmez. Bunun yerine, her gzlemin sadece kme ađırlık merkezlerine olan benzerliđini dikkate alır. Kmeler arasında gzlemlerin yeniden atanmasına izin veren optimizasyon algoritmaları bile her boyuttaki veri kmesine kolayca uygulanabilir.

Hiyerarřik olmayan kmeleme yntemleri bařlıđı altında fazlaca yntemden bahsedilmekle birlikte, en yaygın kullanılan iki tanesi K-Ortalamlar yntemi ile en ok olabilirlik yntemidir.

2.1.2.2.2.1. K-Ortalamlar Yntemi

İlk olarak MacQueen tarafından 1967 yılında ortaya atılan en eski kmeleme algoritmalarından biridir. Hiyerarřik olmayan kmeleme yntemleri arasındaki en sık kullanılan yntem K-Ortalamlar (K-Means) kmeleme yntemi olduđundan, literatrde hiyerarřik olmayan yntem, K-Ortalamlar yntemi olarak bile anılmaktadır. Bu algoritma, her verinin yalnızca bir kmeye dâhil olabilesine izin vermektedir (Han ve Kamber, 2006). Ayrıca bu yntemde kme sayısı yani K deęeri, en az iki, en fazla ise veri setindeki gzlem sayısına eřit olacak řekilde belirlenmektedir (Alpar, 2011). K-Ortalamlar ynteminde takip edilen algoritmalar sırasıyla řu řekildedir (Pektař, 2013):

1. Kme merkezi olarak bařlangıta rastgele K adet birim seilir.
2. Kme merkezinin dıřındaki birimler, hangi uzaklık lleri belirlendiyse bu dođrultuda bařlangıtaki kme merkezlerinin ait oldukları kmelere atanır.
3. Bařlangıta oluřturulan K adet kmedeki deęiřkenlerin ortalamaları alınarak yeni kme merkezleri belirlenir.
4. Birimlerin yeni oluřturulan kme merkezlerine olan uzaklıkları hesaplanır ve her birim en yakın olduđu kmeye atanır.
5. Kme merkezlerinden bir ncekine olan uzaklıklar ile yeni oluřturulana olan uzaklıklar karřılařtırılır.
6. Uzaklıklar kabul edilebilir dzeyde ise drdnc adıma dnlr.
7. ok byk bir deęiřikliđin olmadığı durumlarda ise iterasyona son verilir.

2.1.2.2.2. En Çok Olabilirlik Yöntemi

Veri setindeki her bir gözlemin en büyük olabilirlik değeri verecek şekilde daha önceden belirlenen kümelere atanması temeline dayanmaktadır.

2.1.2.3. ROC (Receiver Operating Characteristics) Analizi

ROC analizi ilk olarak II. Dünya Savaşı sırasında Britanya'da radarda tespit edilen sinyallerin tanımlanmasında kullanılmak üzere geliştirilmiştir. Daha sonra 1967 yılında ROC analizinin tıpta karar vermede kullanımı önerilmiş ve 1969 yılında medikal görüntüleme cihazlarında kullanılmıştır. Sonraki yıllarda ise tıpta tanı testlerinin performansını değerlendirme amacıyla kullanımı giderek yaygınlaşmıştır. Bu analizin en önemli çıktılarından biri olan ROC eğrisi, dikey ekseninde duyarlılık (sensitivity) ve yatay ekseninde 1-özgüllük (1-specificity) yer alacak şekilde çizilen eğridir. Bu eğri, bir tanı testinin yanlış pozitif ve gerçek pozitif oranları arasındaki ilişkinin grafiksel bir gösterimidir. Yani, duyarlılık ve 1-özgüllük ile oluşturulmuş bir fonksiyonun grafiksel şeklidir. Burada duyarlılık, gerçekten hastalıklı bir bireyin pozitif bir test sonucuna sahip olma olasılığını; özgüllük (seçicilik) ise, gerçekten hastalıklı bir bireyin, negatif bir test sonucuna sahip olma olasılığını ifade etmektedir (Krzanowski ve David, 2009).

Günümüzde ise bu analiz, eğitim alanındaki uygulamalarda da karşımıza çıkmaya başlamıştır. ROC analizi ile ortaya çıkan gelişmeler, istatistiksel sonuçların değerlendirilmesi ve kıyaslanmasına duyulan ihtiyaçtan doğmaktadır.

ROC analizi, bireylerin testlerdeki performanslarını belirlemek amacıyla kullanılabilirliği gibi, farklı istatistiksel modellerin veya testlerin ve tahmin modellerinin doğruluğunu değerlendirmek amacıyla da kullanılabilir (Obuchowski, Lieber ve Wians, 2004). Öte yandan ROC eğrisinin kesme puanı belirlemede de kullanılabilirliği üzerinde durulmuştur (Kaftandjieva, 2010). Bu eğrinin önemli göstergelerinden birisi de AUC değeridir. AUC (area under curve) yani eğri altında kalan alan;

$$AUC = \int_0^1 ROC(c)dc$$

şeklinde tanımlanır ve değer aralığı,

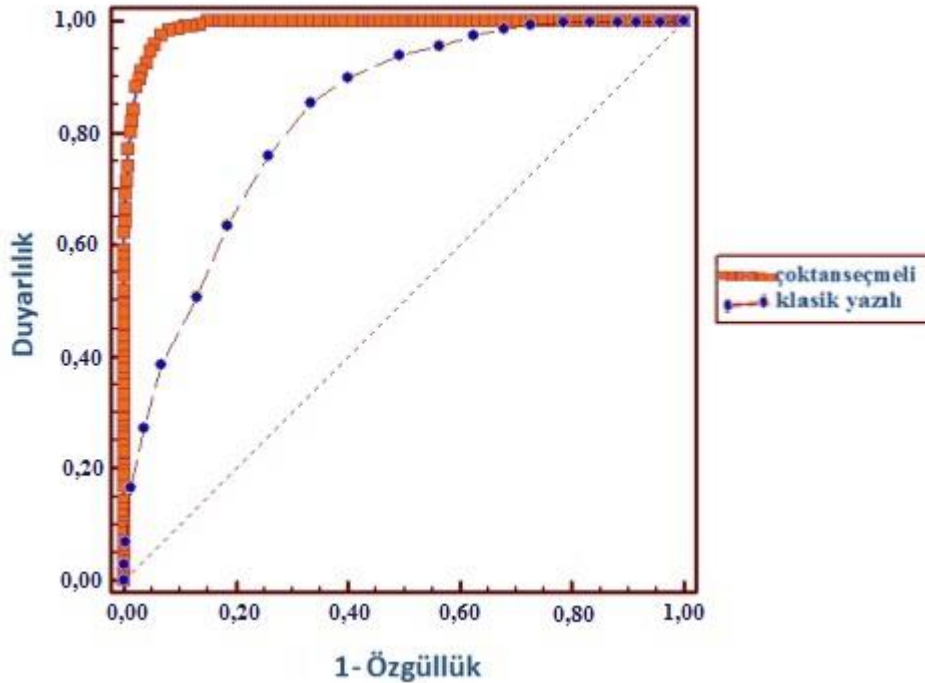
$0,50 \leq AUC \leq 1,00$ 'dır.

Bu değerin 1,00'a eşit olması testin bireyleri kesin olarak ayırt ettiğini gösterirken, 0,50'ye yaklaşması testin ayırt ediciliğinin giderek azaldığını, 0,50'ye eşit olması ise testin ayırt edici olmadığını ifade etmektedir (Alpar, 2010). Tüm AUC değerlerinin yorumlanması Tablo 1'de gösterilmiştir (Obuchowski, 1994).

Tablo 1. AUC Değerlerinin Yorumlanması

AUC = 0,50	Test ayırım gücüne sahip değildir.
0,50 < AUC < 0,70	Test “zayıf” bir ayırım gücüne sahiptir.
0,70 ≤ AUC ≤ 0,80	Test “kabul edilebilir” bir ayırım gücüne sahiptir.
0,80 ≤ AUC ≤ 0,90	Test “mükemmel” bir ayırım gücüne sahiptir.
0,90 < AUC	Test “olağanüstü” bir ayırım gücüne sahiptir.

Tablo 1'de de gösterildiği gibi AUC değeri, kendi tanım aralığı içerisinde beş kategoriye ayrılarak yorumlanabilmektedir. Şekil 1'de yapılan bir araştırmada elde edilen iki farklı ROC eğrisi gösterilmiştir (Keçeoğlu, 2016).



Şekil 1. ROC Eğrisi Örneği

Testin ‘şans çizgisi’ olarak adlandırılan köşegen, Şekil 1'deki (0,0) noktasından (1,1) noktasına kadar 45°lik açıyla çizilmektedir. Şans çizgisiyle ROC eğrisinin çakışık olması, yanlış pozitif oran ile doğru pozitif oran değerlerinin birbirine eşit olduğunu göstermektedir.

ROC eğrisinin ideal olanı ise (0,0), (0,1) ve (1,1) noktalarının birleştirilmesiyle oluşmaktadır. Dolayısıyla eğri (0,1) noktasına yaklaştığı ölçüde testin bireyleri doğru sınıflama gücü (*duyarlılığı / sensitivity*) artmaktadır. Bir başka deyişle, ROC eğrisinin altında kalan alan yani AUC değeri 1,00'a yaklaştıkça yapılan sınıflamada bireylerin ayrımı da mükemmelliğe doğru gitmekte; 0,50'ye yaklaştıkça ise yapılan sınıflama kullanışsız hâle gelmektedir. Bu bilgiler ışığında, Şekil 1'deki çoktan seçmeli teste ilişkin ROC eğrisinin altında kalan alanın yani AUC değerinin, klasik yazılı testine ilişkin ROC eğrisinin AUC değerinden büyük olduğunu söylemek mümkündür.

2.2. İlgili Araştırmalar

Dattalo (1995), Lojistik Regresyon analizinin Diskriminant analizine güçlü bir alternatif olarak ortaya çıktığını vurgulamıştır. Yöntemlerin her ikisinin de iyi birer sınıflama yöntemi olmasına rağmen, Lojistik Regresyon analizinin, Diskriminant analizine göre daha sıkı (*parsimonious*) ve yorumlamasının daha kolay olduğunu belirtmiştir. Ancak uygulamalarda kullanılacak yönteme karar verilirken, analizin amacının (sınıflama ya da betimleme gibi), örneklemin özelliklerinin, varsayımların karşılanması ve hangi yöntemin verdiği analiz sonuçlarının daha anlamlı, daha kullanışlı olduğunu dikkate alınması gerektiği üzerinde durmuştur.

Hess, Subhiyah ve Giordano (2007), kümeleme analizinin testi alanları, maddelere verdikleri yanıtlara göre homojen gruplara ayırarak testin minimum geçme puanını belirlemede kullanışlı bir istatistiksel yöntem olabileceği üzerinde durduğu bir çalışma gerçekleştirmişlerdir. Başka bir deyişle, uzman yargısına dayanan temellerle geçme puanının belirlendiği Ebel ve Nedelsky gibi metotlardansa, kümeleme analizinin daha kullanışlı olabileceğini vurgulamışlardır. Sınıflama amacıyla da kullanılan kümeleme analizi ile Angoff (1-0), Ebel ve Nedelsky tekniklerini kullanarak kişilerle ilgili alınan geçti/kaldı kararlarının doğruluklarını belirlemişlerdir. Bu amaçla 652 kişiden elde ettikleri verilerin analizi sonucunda, Ebel ve Nedelsky yöntemleri için yargıcı kararları ile belirlenen geçti/kaldı sınıflamaları benzer bulunmuştur. Angoff (1-0) kesme puanının doğruluğu için kümeleme analizi kullanılmıştır. Test merkezli sınav sonucunda Angoff ve kümeleme analizi teknikleri ile geçme notu için elde edilen kesme puanlarının birbirleri ile tutarlılıklarının yüksek olduğu ortaya konmuştur.

Vivo ve Franco (2008) akademik başarıya ilişkin bağımsız değişkenlerin doğruluğunun değerlendirilmesi ve ROC analizinin üniversiteye girişte etkili olan faktörlerin

uygunluğunun belirlenmesine ilişkin bir araştırma gerçekleştirmişlerdir. Araştırmanın temelde iki amacı bulunmaktadır. Bu amaçlardan biri her bir kademeden lisans öğrencilerinin akademik performans sınıflandırmasında üniversiteye giriş faktörlerinin geçerliliğinin değerlendirilmesidir. Bir diğeri ise öğrenci başarısını tanımlamak amacıyla kesme puanı belirlemektir. Araştırmanın örneklemini, iki farklı İspanyol Devlet üniversitesindeki Ekonomi ve İşletme bölümlerine, kayıtlarını üniversite sınavını geçmelerinin hemen ardından yaptıran öğrenciler oluşturmaktadır. Üniversiteye giriş sırasında akademik performansı sınıflandırmak amacıyla dört farklı faktör belirlenmiştir. Bu faktörler öğrencilerin ortaöğretim notlarına ilişkin ortalamaları, üniversite girişindeki genel sınava ilişkin notları, üniversite girişindeki özel kısım sınavına ilişkin notları, ortaöğretim ve üniversite girişindeki sınava ilişkin ağırlıklandırılmış puanı şeklindedir. Ekonomi ve işletme alanlarındaki öğrencilerin başarılı veya başarısız şeklinde sınıflandırılmasında, ortaöğretim not ortalaması, ortaöğretim ve üniversite giriş sınavı ağırlıklandırılmış puan faktörlerinden elde edilen verilerin daha doğru olduğu, ekonomi alanında yapılan sınıflandırmanın da işletmeye kıyasla daha doğru olduğu ortaya konmuştur. Bunlarla birlikte, alanların her ikisi için de bu faktörlerin tahmin edildiği ölçüde iyi yordayıcılar olmadığı ifade edilmiştir.

2007 yılında Uşak'taki kamu hastanelerinde hasta memnuniyetini ölçmeyi amaçlayan bir anketten elde edilen ve 364 hastayı kapsayan veri seti üzerinden araştırmasını gerçekleştiren Tayyar (2010), hasta memnuniyetini tahmin etmede üç farklı yöntem karşılaştırması yapmıştır. Bu araştırmada Yapay Sinir Ağları, Lojistik Regresyon ve Ayırma Analizi yöntemlerini kullanmıştır. Ortalama tahmin oranları sırasıyla %78, %75,3 ve %75 olarak bulunmuştur. Bu sonuçlar ışığında LR ve ayırma analizine ilişkin oranların birbirine oldukça yakın olduğu; yapay sinir ağlarının ise en iyi ortalama tahmin oranına sahip olduğu ifade edilmiştir. Sonuç olarak, yapay sinir ağlarının diğer iki yönteme göre hasta memnuniyetini tahmin etmede daha iyi bir orana sahip olduğu ortaya konmuştur.

İşlem karakteristik (ROC) eğrisi analizi ve eğri altında kalan alanların (AUC) karşılaştırılmasına ilişkin bir araştırma yapan Tokmak ve Bek (2010), hipertansiyon saptanmasında kullanılan iki tanı testine ait elde ettikleri verileri binormal ve nonparametrik yöntemlere göre karşılaştırmışlardır. Binormal yöntemle karşılaştırma sonucunda iki testin tanı koyabilme açısından farklarının istatistiksel olarak manidar olduğunu ortaya koymuşlardır. Diğer yandan, nonparametrik yöntemle karşılaştırma sonucunda iki tanı testinin istatistiksel olarak manidar bir farkının olmadığı görülmüştür. AUC değerinin binormal yöntemde değişebildiği ancak nonparametrik yöntemde eğri altında kalan alanın gerçek

değerinin hesaplanabildiği saptanmış ve bu sebeple nonparametrik yöntemin binormal yönteme tercih edilebileceği vurgulanmıştır.

Köksal (2011) 18 yaş ve üzerindeki 1345 kişiyi çalışma grubu olarak belirlediği araştırmasında, bağımlı değişken olarak mutluluk algısı ile öznel ve nesnel nitelikli sosyo-demografik göstergeleri almıştır. Bu değişkenler için kestirilen farklı lojistik regresyon modelleri içinden en yüksek kestirim gücüne sahip modelin tespit edilmesinde, mevcut seçim ölçütlerine alternatif olarak ROC eğrisi kestiriminden de faydalanılabileceğini göstermeyi amaçlamıştır. Araştırmada ROC eğrisi kestirimine göre model seçiminin, Cox&Snell R², Nagelkerke R², Olabilirlik Oran Testi ve Hosmer-Lemeshow Uyum İyiliği Testi'ne alternatif olarak yapılabileceği hem kuramsal olarak, hem de yapılan uygulamalarla ortaya konmuştur. Verilerle iyi uyuma sahip olan, en yüksek ikinci R² değerine sahip olan ve içerdiği değişkenlere ve parametre manidarlıklarına ilişkin diğerlerine göre öznel modelin daha iyi olduğu sonucu; en yüksek kestirim gücüne sahip olan modelin, ROC eğrisine göre %95 güvenle öznel model olduğu bulgusuyla da desteklenmiştir. ROC eğrisi kestiriminin, kestirilen modeller içinden en iyi ayırım gücüne sahip olan modelin tespit edilmesinde, en uygun seçim ölçütü olduğu saptanmıştır. En etkin modelin tespitinde model seçim ölçütü olarak AUC kestirimlerinin göz önüne alınmasının daha uygun olacağı da vurgulanmıştır.

Reyhanlıoğlu Keçeoğlu (2012), Türkiye'de bulunan bir yükseköğretim kurumundaki 1708 öğrenciye uygulanan İngilizce hazırlık sınıfı muafiyet sınavının sonuçlarından elde ettiği veriler üzerinde Lojistik Regresyon analizi, Ayırma (Diskriminant) analizi ve ROC analizi yöntemleriyle belirlediği kesme puanlarının, yöntemle göre değişip değişmediğini incelediği bir çalışma yapmıştır. Kurumun kendi yaptığı sınavda geçme puanı olarak belirlenen 65 iç ölçütü, ulusal düzeyde yapılan sınavlarda (ÜDS veya KPDS) geçme puanı olarak belirlenen 75 ise dış ölçütü temsil etmektedir. Araştırmanın bulguları doğrultusunda iç ölçüte göre, lojistik regresyon ve ayırma analizi yöntemlerinde kişileri başarılı/başarısız olarak sınıflama açısından manidar bir farklılık bulunmamıştır. Aynı yöntemlerle dış ölçüte ilişkin belirlenen kesme puanları arasındaki fark lojistik regresyon analizinin lehine bulunmuştur. İç ölçüte ilişkin, lojistik regresyon ve ROC analizi yöntemleriyle belirlenen kesme puanları arasında ROC analizi lehine bir farklılık bulunurken; aynı yöntemlerle dış ölçüte ilişkin belirlenen kesme puanları arasındaki fark ise lojistik regresyon analizinin lehine bulunmuştur. Ayırma ve ROC analizi yöntemleriyle belirlenen kesme puanları arasında da gerek iç gerekse dış ölçüte ilişkin bulunan farkın ROC analizi lehine olduğu ortaya konmuştur.

Taşdemir (2013) sınıf öğretmenliği öğrencilerinin test başarılarının farklı standart belirleme yöntemleri ile incelendiği, kesme puanlarının belirlendiği ve bu yöntemlerle belirlenen sınıflandırma doğruluklarını konu alan bir araştırma gerçekleştirmiştir. Bu amaçla standart belirleme yöntemlerinden Angoff (1-0), Nedelsky ile sınır değer saptama yöntemleri İşlem Karakteristik Eğrisi (ROC) ve Aralık Tahmini ile belirlenen sınıflama doğruluklarını incelemiştir. Araştırmanın çalışma grubunu, Türkçe ve matematik dersi almış olma ölçütünü karşılayan, sınıf öğretmenliği programının üçüncü sınıfındaki 211 öğrenci oluşturmaktadır. 2006-2012 yılları arasında uygulanan Kamu Personeli Seçme Sınavı'ndaki (KPSS) Türkçe ve matematik alt testlerinden tesadüfi olarak seçilen toplam 30 soruluk bir test ölçme aracı olarak kullanılmıştır. Araştırmada kullanılan dört farklı yöntemle ilişkin kesme puanları ayrı ayrı belirlenmiştir. ROC analizinin, matematik testinden başarılı/başarısız olan öğrencileri sınıflandırmada doğru sonuçlar verdiği ifade edilmiştir. Buna ek olarak, hem kesme puanı belirleme amaçlı yapılan çalışmalarda, hem de öğrencilerin sınıflandırılmasında daha doğru kararlar alınabilmesi için genellikle kullanılan Angoff (1-0) ve Nedelsky teknikleri gibi bu analizin de tercih edilebileceği belirtilmiştir.

Tavakol ve Dennick (2012) 320 öğrenciden elde ettikleri veri seti ile standart belirleme üzerine ROC analizi yöntemini kullanarak yaptıkları çalışmada veri toplama aracı olarak 23 maddeden oluşan çoktan seçmeli bir fizyoloji testi kullanmışlardır. Testi alanların geçti/kaldı sınıflandırması yapılırken, kesme puanı olarak 11 puanı kullanılmıştır. Bu puan, standart belirleyiciler tarafından tahmin edilmiştir. ROC analiziyle ortaya konulan 10,5 kesme puanı değerinin, kesme puanı belirleyicileri tarafından ortaya konan 11,0 kesme puan değerinden çok uzak olmadığı ve ROC analizi sonuçlarının öğrencilerin geçti/kaldı kararlarının verilmesinde kritik olan kesme değerine ilişkin en yakın bilgiyi sağladığı ifade edilmiştir. Bunlara ek olarak, ROC analizinde test merkezli değer yöntemlerin aksine duyarlılık ve özgüllük (seçicilik) değerleri dikkate alındığından, sonuçların daha objektif olduğu bilgisi üzerinde durulmuştur.

Koyuncu (2015) psikolojik ölçeklerde ROC analiziyle standart belirmeye yönelik olarak yaptığı yüksek lisans çalışmasında ölçek olarak öğretmen adayları için KPSS Sürekli Kaygı Ölçeği'ni (ÖAKSKÖ) kullanmıştır. Çalışmasında gerek geliştirilmiş gerekse uyarlanmış ölçeklerde kesme puanının ROC analiziyle nasıl belirleneceğini göstermeyi amaçlamış ve çalışmasını 2013-2014 Bahar Dönemi Gazi Üniversitesi Gazi Eğitim Fakültesi'nde öğrenim görmekte olan 554 kız ve 170 erkek öğrenci olmak üzere toplam 724 öğrenciden elde ettiği veri setiyle gerçekleştirmiştir. Çalışmasında hem ilgili ölçeğe ilişkin en

uygun kesme puanının ROC analizi sonucu kaç olduğunu incelemiş, hem de bu kesme puanının cinsiyete ve örneklem büyüklüğüne göre farklılaşma durumunu ele almıştır. Çalışmanın bulgularına bakıldığında, genel olarak ÖAKSKÖ'nin yorumlanmasında kullanılacak olan kesme puanı 28,5 olarak hesaplanmıştır. Cinsiyete göre seçilen kesim noktası ile gerçek durumun uyum yüzdesi kızlarda %71; erkeklerde %74 olarak bulunmuştur. ÖAKSKÖ için kesme puanları ile eğri altında kalan alan (AUC) değerlerinin, örneklem büyüklükleri rastgele seçildiğinde farklılaştığı ve elde edilen kesme puanının uygulanan gruba göre farklılık gösterebileceği ifade edilmiştir. Belirlenen kesme puanları aynı olsa bile, duyarlılık, özgüllük ve AUC değerlerinin farklı olabileceği gösterilmiştir.

Sarkın (2017), üniversite öğrencilerine yönelik kaygı düzeyi belirleyen bir ölçeğin kesme puanının Angoff, ROC ve Sınır Grup yöntemleriyle belirlenmesine ilişkin bir araştırma gerçekleştirmiştir. Araştırma verilerini bir devlet üniversitesinin eğitim fakültesinde öğrenim gören ve Kamu Personeli Seçme Sınavı'na girecek olan 290 öğrenciden elde etmiştir. Araştırmanın sonucunda ise kesme puanı belirlenecek olan ölçme aracı, çoklu puanlanan bir psikolojik ölçek olduğunda ROC analizinin kullanılmasının, diğer yöntemlerde olmayan duyarlılık, özgüllük ve pozitif/negatif yordayıcı değerlerinin bulunması sebebiyle daha ayrıntılı ve nesnel bilgiler sağladığı ortaya konmuştur.

Şata ve Karip (2017) 'akıllı telefon bağımlılık ölçeği-kısa versiyonu'nu kullandıkları çalışmalarında bu ölçeğin ergen bireyler için Türk kültürüne uyarlanması ile psikometrik özelliklerinin incelenmesini amaçlamışlardır. Bu amaçlar doğrultusunda, 138'i kız, 106'sı erkek olan toplam 244 katılımcıdan elde ettikleri verileri kullanmışlardır. Sonrasında ilgili ölçeğin güvenilirlik ve geçerliğine ilişkin kanıtların sağlandığını ortaya koymuşlardır. Bunun yanında çalışma bir ölçek uyarlama çalışması olduğundan, gerekli olan standart belirleme adımlarını da izlemişlerdir. Bu doğrultuda, çalışmada iki aşamalı kümeleme analizi kullanılmıştır. Bu analiz sonucunda bağımlı ve bağımlı değil olmak üzere iki tane küme oluşmuş, dolayısıyla bir tane kesme puanı belirlenmiştir. Bu puan belirlenirken, ölçekten alınan puanlara göre, düşük ortalamaya sahip olan kümedeki en yüksek puan ile yüksek ortalamaya sahip olan kümedeki en düşük puanının ortalaması alınmıştır. Farklı cinsiyetteki öğrenciler için benzer işlemler ayrı ayrı da yapılmıştır. Her iki cinsiyete ilişkin kesme puanları, iki aşamalı kümeleme analizi kullanılarak hesaplanmış ve 29,5 olarak belirlenmiştir.

Çakıroğlu ve Soylu (2019) yaptıkları çalışmada dört farklı okulda öğrenim görmekte olan 10-18 yaş arası 1161 öğrenciden elde ettikleri verileri kullanmışlardır. 'İnternet Oyun

Oynama Bozukluęu Ölçeęi' için Kümeleme analizi ve ROC analizi kullanılarak, ölçeęin kesme puanı 69 olarak elde edilmiştir.

Şata, Çorbacı ve Koyuncu (2019), 'mobil öğrenmeye hazırbulunuşluk' ölçeęinin Türk kültürüne uyarlamasını yapmayı ve psikometrik özelliklerini incelemeyi amaçladıkları çalışmalarında; 165 kız ve 36 erkek olmak üzere toplam 201 katılımcıdan elde ettikleri verileri kullanmışlardır. Öncelikle ilgili ölçeęin güvenilirlik ve geçerliğine ilişkin kanıtlar ortaya konmuştur. Sonrasında, ölçek uyarlama çalışmalarında gerekli olan standart belirleme adımları izlenmiştir. Bu doğrultuda, çalışmada iki aşamalı kümeleme analizi kullanılmıştır. İki aşamalı kümeleme analizi sonucunda iki küme ortaya çıkmış, dolayısıyla bir tane kesme puanı belirlenmiştir. Kesme puanı, ölçekten alınan puanlar doğrultusunda, düşük ortalamaya sahip kümenin en yüksek puanı ile yüksek ortalamaya sahip kümenin en düşük puanının ortalaması olarak bulunmuştur. İki aşamalı kümeleme analizi ile hesaplanan ilgili ölçeęin kesme puanı 89,0 olarak belirlenmiştir. Bir başka deyişle, 89 puan ve üstü alan katılımcıların yüksek düzeyde mobil öğrenmeye istekli olduęu, 89 puanın altında alan katılımcıların ise düşük düzeyde mobil öğrenmeye istekli olduęu ortaya konmuştur.

Sonuç olarak LR analizinin farklı yöntemlerle karşılaştırmalı olarak yapılan analizler sonucunda standart belirlemede iyi bir alternatif olduęu, yorumlanması kolay sonuçlar verdięi ortaya konmuştur. Kümeleme analizinin kesme puanı belirlemede kullanıldığı çok fazla çalışmaya rastlanmasa da, kesme puanı belirlemede kümeleme analizinin de kullanılabileceęinin ortaya konduęu çalışmalar mevcuttur. ROC analizi de önceki çalışmalarda farklı yöntemlerle karşılaştırmalı olarak kullanılmış ve dięer yöntemlerde bulunmayan duyarlılık, özgüllük, pozitif/negatif yordayıcı deęerlerinin daha ayrıntılı ve nesnel sonuçlar elde etmede avantaj sağladığı ortaya konmuştur.

ÜÇÜNCÜ BÖLÜM

YÖNTEM

Bu bölümde araştırmanın türü, çalışma grubu, verilerin toplanması, veri toplama araçları, verilerin analizi ve çözümlenmesi yer almaktadır.

3.1. Araştırma Modeli

Bu çalışmada, standart belirlemeye yönelik olan üç farklı istatistiksel yöntem, bireyleri sınıflandırma doğrulukları açısından karşılaştırılmıştır. Bu karşılaştırmalar yapılırken, bireylerin ÇSF ve KCF den geçme-kalma durumlarıyla genel durumda geçme-kalma durumları incelenmiştir. İlişkisel araştırmalar, iki ya da daha fazla değişken arasındaki ilişkinin, bu değişkenlere bir müdahalede bulunulmadan ele alındığı araştırmalardır (Ary, Jacobs, Sorensen, ve Razavieh, 2010). İlişkisel türdeki araştırmalarda değişkenlerin ilişki düzeyleri, istatistiksel testler kullanılarak ölçülmeye çalışılır. Üç farklı istatistiksel yöntemin etkilerinin incelendiği bu çalışmada da bireylerin aynı özelliği ölçen iki testteki geçme-kalma durumlarıyla genel sonuçları arasındaki bağlantının ilgili üç yöntemle yapılan analizlerle ortaya konulması ve elde edilen sınıflama doğruluklarının yöntemsel açıdan karşılaştırılması amaçlandığından, araştırma model olarak ilişkiseldir.

3.2. Çalışma Grubu

Bu araştırmanın evreni, çalışmada kullanılan başarı testleri 7. sınıf seviyesine uygun olduğundan, Türkiye'deki 7. sınıf öğrencilerinden oluşmaktadır. Araştırmanın çalışma grubunu ise, 2020-2021 eğitim-öğretim yılında Gaziantep ilinin Şahinbey ilçesindeki iki devlet ortaokulunda öğrenim görmekte olan 7. sınıf öğrencileri oluşturmaktadır. 326 kişiden oluşan çalışma grubuna ilişkin frekans ve yüzdeler Tablo 2'de gösterilmiştir.

Tablo 2. Çalışma Grubuna İlişkin Frekans ve Yüzdeler

	Cinsiyet					
	Kız		Erkek		Toplam	
	f	%	f	%	f	%
Okul-1	28	54,9	23	45,1	51	100
Okul-2	131	47,6	144	52,4	275	100
Toplam	159	48,8	167	51,2	326	100

Tablo 2’de görüldüğü gibi Okul-2’den daha fazla katılımcıya ulaşılmış olmakla beraber cinsiyete göre dengeli bir dağılımın söz konusu olduğu görülmektedir.

3.3. Verilerin Toplanması

Araştırma için Gaziantep ilinin Şahinbey ilçesine bağlı iki devlet okulunda öğrenim gören 7. sınıf öğrencilerinden elde edilen veriler kullanılmıştır. Bu veriler salgın döneminde olunması sebebiyle çevrimiçi ortamda toplanmıştır. Araştırma kapsamında kullanılan başarı testleri Google forms kullanılarak öğrencilere ulaştırılmıştır. Her iki başarı testinin de hem deneme uygulaması hem de nihai uygulaması bu platform üzerinden gerçekleştirilmiştir. Analiz aşamasında aynı özelliği ölçen iki farklı testi alan bireylerin eşleştirilebilmesi için demografik özellik olarak okul adı ve okul numarasına yer verilmiştir. Araştırmaya dâhil olan grubu betimleyebilmek için testi alan bireylerin cinsiyetleri de yine demografik özellik olarak istenmiştir.

3.4. Veri Toplama Araçları

Araştırmada 7. sınıf seviyesine uygun olarak aynı özelliği ölçen iki farklı formda test geliştirilmiştir. Bu testlerden bir tanesi Ek-2’de verilen ve çoktan seçmeli 14 maddeden oluşan çoktan seçmeli form (ÇSF), diğeri de Ek-3’te verilen ve kısa cevaplı 24 maddeden oluşan kısa cevaplı formdur (KCF). Yapılan deneme uygulamaları sonunda başlangıçta 16 madde olarak geliştirilen ÇSF’de ayırt edicilik indeksleri Microsoft Excel 2010 paket programı yardımıyla hesaplanmış ve 0.20’nin altında olan iki madde testten çıkarılmıştır. Yine başlangıçta 25 madde olarak geliştirilen KCF’de ise hazırlanan uzman görüş formu (Ek-1) ile alınan uzman görüşleri sonucunda uygun bulunmaması sebebiyle bir madde testten çıkarılmıştır. Böylelikle nihai forma alınmak üzere ÇSF ve KCF sırasıyla 14 ve 24 madde olarak düzenlenmiştir. Uygulamaya ilişkin gerekli izinler alınmıştır (Ek-4).

3.5. Verilerin Analizi ve Yorumu

Araştırmanın bulgularını elde etmek için geliştirilen 7. sınıf başarı testinin aynı özelliği ölçen iki farklı formundan (ÇSF ve KCF) alınan puanlar araştırmanın bağımsız değişkenlerini oluşturmaktadır. Öğrenciler her iki formdan da 100 üzerinden puan almış ve aldıkları bu puanların ortalaması, genel ortalamalarını oluşturmuştur. Başarı testi Milli Eğitim Bakanlığı’na bağlı devlet okullarında uygulandığından, mevcut sistemdeki geçme puanı olan 45 ölçüt olarak alınmıştır. Bu ölçüte göre genel durumu 45 ve üzeri olan öğrenciler ‘geçti’, 45’in altında olan öğrenciler ise ‘kaldı’ olarak adlandırılmıştır. Veri dosyasında öğrencilerin

başarı testinin her iki formunda yer alan her bir maddeye verdikleri doğru cevap ‘1’ ile, yanlış cevap ise ‘0’ ile kodlanmıştır.

3.5.1. Birinci ve İkinci Alt Problemler İçin Verilerin Analizi ve Yorumu

‘Çoktan seçmeli formun lojistik regresyon analizi, kümeleme analizi ve ROC analizi ile belirlenen sınıflama doğrulukları farklılaşmakta mıdır?’

‘Kısa cevaplı formun lojistik regresyon analizi, kümeleme analizi ve ROC analizi ile belirlenen sınıflama doğrulukları farklılaşmakta mıdır?’

Birinci ve ikinci alt problemler için sırasıyla ÇSF ve KCF verilerinin analizi üç farklı istatistiksel yöntemle yapılmış ve her bir yöntemden bir sınıflama doğruluğu yüzdesi elde edilmiştir. Sonrasında ise, elde edilen bu yüzdeler oran testi kullanılarak karşılaştırılmış ve yüzdeler arasındaki farkın istatistiksel olarak manidar olup olmadığı belirlenmiştir.

İlk olarak LR analizi için SPSS 20.0 paket programı ile varsayımların kontrolü yapılmıştır. Bu analizin varsayımları aşağıda detaylı olarak açıklanmıştır (Can, 2019).

Bağımlı değişken iki kategorili olmalıdır: Araştırmanın bağımlı değişkeni olan ‘genel durum’ yani başarı durumu, geçti/kaldı şeklinde iki kategoriden oluşmaktadır.

Bağımsız değişken(ler) sürekli (aralık ve oran ölçeği) yapıda ve(ya) kategorik yapıda olabilir: Araştırmanın bağımsız değişkenleri olan “ÇSF” ve “KCF” değişkenleri başarı testinden elde edilen puanlardan oluşup sürekli yapıya sahiptir.

Gözlemlerin bağımsızlığı sağlanmış olmalıdır: Veri setinde bulunan gözlemler 326 farklı kişiden elde edilmiştir.

Bağımsız değişkenler arasında çoklu doğrusal bağlantı olmamalıdır: Bağımsız değişkenler arasındaki korelasyon yüksek düzeyde olmamalıdır. Araştırmadaki bağımsız değişkenler arasındaki korelasyon da 0,66 olarak bulunmuştur. Çoklu doğrusal bağlantı ise değişkenler arasında oldukça yüksek korelasyon olduğunda ($r=0,90$ ve daha yüksek) ortaya çıkmaktadır (Pallant, 2017).

Aşırı uç değerler olmamalıdır: Araştırmanın verileri incelenerek uç değer analizi yapılmıştır. Herhangi bir uç değere rastlanmamıştır.

Elde edilen bulgular doğrultusunda 326 öğrenciden toplanan verilere SPSS 20.0 paket programıyla LR analizi uygulanmıştır. LR analizi çıktılarında, başlangıç modeline ait

iterasyon öyküsü ve sınıflama tablosu ile bağımsız değişkenin de dâhil edildiği modele ilişkin iterasyon öyküsü, Omnibus testi sonuçları, Pseudo-R² ölçüleri ve sınıflama tablosu raporlanmıştır. Omnibus testinden elde edilen X^2 değerinin anlamlı olması bağımsız değişkenlerin başlangıç modeline eklenmesi ile bağımlı değişkeni yordama gücünün artacağını göstergesidir. Öte yandan çoklu doğrusal regresyon analizinde bağımsız değişkenlerin bağımlı değişkenin ne kadarını açıkladığını ifade etmek için R² determinant katsayısından yararlanılmaktadır. Ancak LRA doğrusal bir regresyon analizi olmadığı için kurulan LR modelinde bağımlı değişkendeki değişimin bir başka ifade ile varyansın bağımsız değişkenler tarafından açıklanma derecesini ortaya koyma amacıyla Pseudo-R² olarak adlandırılan farklı R² ölçülerinden yararlanır. Bunlar, “Pseudo-R²” ölçülerinden “Cox&Snell’s R²” ve “Nagelkerke R²” ölçüleridir. Bu ölçülere göre lojistik model tarafından açıklanan varyansın 1,00 olması mükemmel uyumun göstergesidir. Her iki Pseudo-R² ölçüsünün de alabileceği en küçük değer 0’dır.

İkinci olarak kümeleme analizi için SPSS 20.0 paket programı ile varsayımların kontrolü yapılmıştır. Bu analizde araştırmayı yapan kişilerin düşünmesi gereken iki önemli husus, ‘örneklem temsili’ ve ‘değişkenler arası çoklu doğrusal bağlantı’ problemidir (Hair ve diğerleri, 2006). Araştırmacılar çok nadir olarak evrenin tam sayısına sahip olduklarından, bu araştırmada da olduğu gibi örneklem üzerinde çalışılır (Çokluk, Şekercioğlu ve Büyüköztürk, 2018). Çoklu bağlantı problemi ile ilgili varsayım testi sonuçları yukarıda LR analizi varsayımlarının testinde belirtilmişti. Ayrıca kümeleme analizinin varsayımlarına ilişkin olarak normallik varsayımı olmakla beraber, bu varsayımın teoride kaldığı ve uzaklık değerlerinin normalliğinin yeterli görüldüğü belirtilmektedir (Tatlıdil, 2002). Kümeleme analizindeki hiyerarşik olmayan yöntemler, örneklem büyüklüğünün 250’den büyük olduğu ve küme sayısının önceden belli olduğu veri setleri için daha uygun görülmektedir (Çokluk, Şekercioğlu ve Büyüköztürk, 2018). Bu yöntemler arasında K-Ortalamalar kümeleme analizi, yaygın olarak kullanıldığından bu çalışmada da kümeleme yöntemi olarak tercih edilmiştir. Pektaş’a (2013) göre, K-Ortalamalar yönteminin uygulanabilmesi için iki koşul sağlanmalıdır:

1. Veri setindeki değişkenler en az aralık ölçeğinde olmalıdır.
2. Oluşturulacak olan küme sayısı başlangıçta biliniyor olmalıdır.

Elde edilen bulgular doğrultusunda 326 öğrenciden toplanan verilere SPSS 20.0 paket programıyla hiyerarşik olmayan yöntemlerden biri olan K-Means kümeleme analizi

uygulanmıştır. Kümeleme analizi çıktılarında, önceden iki (geçti/kaldı) olarak belirlenen kümelerdeki kişi sayılarına ilişkin frekans ve yüzdeler ile sınıflama tablosu raporlanmıştır.

Üçüncü olarak ROC analizi için Jamovi paket programı kullanılmıştır. Bu analiz, çıktılarında verdiği ROC eğrisinin altında kalan alanın büyüklüğü ölçüsünde yapılan sınıflamanın doğruluğunu göstermektedir. Ölçme aracının istenilen ayrımı hiç yapamadığı durumlarda ROC eğrisi altında kalan alanı gösteren bu indeksin beklenen değeri .50 iken, istenilen ayırım mükemmel olarak yapıldığında ise bu değer 1.00 olmaktadır. ROC eğrisi altında kalan bu alanın büyüklüğünü gösteren bu değer .50 ile 1.00 arasında olması beklenmektedir (Kanık ve Erden, 2003). Ölçme aracının doğru sınıflama tahminlerinin, ROC eğrisinin altında kalan alanın büyüklüğü ölçüsünde iyi olduğu söylenebilir (Hanley ve McNeil, 1982; Grove, 2006).

Elde edilen bulgular doğrultusunda 326 öğrenciden toplanan verilere Jamovi paket programıyla ROC analizi uygulanmıştır. Analiz çıktılarında, sonuçların doğruluk oranını gösteren doğruluk (accuracy) değeri, özgüllük değeri, duyarlılık değeri, AUC değeri ile ROC eğrisi raporlanmıştır. Buradaki doğru sınıflama yüzdesini AUC değeri göstermektedir.

Son olarak elde edilen sınıflama doğrulukları arasındaki farkı ve bu farkın istatistiksel manidarlığını sınamak için oran testi kullanılmıştır. Kullanılan üç istatistiksel yöntemin sınıflama sonuçları ikili olarak karşılaştırılmış ve bulgular bu doğrultuda raporlanmıştır. Oran testi, n değeri 30'dan küçük ise t istatistiği; 30'dan büyük ise z istatistiği hesaplanarak gerçekleştirilmektedir. Bu araştırmada n değeri 326'ya eşit olduğundan, oran testinde z istatistiği hesaplanmıştır. Yapılan oran testinde kullanılan formül Eşitlik 1'de verilmiştir.

$$Z = \frac{P_1 - P_2}{\hat{\sigma}_{\Delta p}}$$

$$\hat{\sigma}_{\Delta p} = \sqrt{\frac{P_1 \times q_1}{n_1} + \frac{P_2 \times q_2}{n_2}}$$

Eşitlik 1. Oran Testi Formülü

p₁: Birinci oran

q₁: 1-p₁

n_1 : Birinci örneklem sayısı

p_2 : İkinci oran

q_2 : $1-p_2$

n_2 : İkinci örneklem sayısı

Eşitlik 1'deki formülde p değerleri, dolayısıyla q değerleri 0 ile 1 arasında olacak şekilde alınmaktadır. Bu formülle hesaplanan z değeri, 650 (n_1+n_2-2) serbestlik derecesi ve .05 manidarlık düzeyindeki z değeri ile ($Z_{tablo}=1.96$) karşılaştırılmıştır. Eğer formül ile hesaplanan z değeri, Z_{tablo} değerinden küçük ise farkın istatistiksel olarak manidar olmadığı, büyük ise farkın istatistiksel olarak manidar olduğu sonucuna ulaşılmıştır.

3.5.2. Üçüncü Alt Problem İçin Verilerin Analizi ve Yorumu

“Çoktan seçmeli ve kısa cevaplı formun lojistik regresyon analizi ile belirlenen sınıflama doğrulukları farklılaşmakta mıdır?”

ÇSF ve KCF için LR analizi ile hesaplanan sınıflama doğrulukları arasındaki farkın istatistiksel olarak manidarlığı, oran testi kullanılarak belirlenmiştir. Yapılan LR analizi sürecindeki adımlar ve oran testinin detayları, birinci ve ikinci alt problemler için verilerin analizi ve yorumu kısmında açıklanmıştır.

3.5.3. Dördüncü Alt Problem İçin Verilerin Analizi ve Yorumu

“Çoktan seçmeli ve kısa cevaplı formun kümeleme analizi ile belirlenen sınıflama doğrulukları farklılaşmakta mıdır?”

ÇSF ve KCF için kümeleme analizi ile hesaplanan sınıflama doğrulukları arasındaki farkın istatistiksel olarak manidarlığı, oran testi kullanılarak belirlenmiştir. Yapılan kümeleme analizi sürecindeki adımlar ve oran testinin detayları, birinci ve ikinci alt problemler için verilerin analizi ve yorumu kısmında açıklanmıştır.

3.5.4. Beşinci Alt Problem İçin Verilerin Analizi ve Yorumu

“Çoktan seçmeli ve kısa cevaplı formun ROC analizi ile belirlenen sınıflama doğrulukları farklılaşmakta mıdır?”

ÇSF ve KCF için ROC analizi ile hesaplanan sınıflama doğrulukları arasındaki farkın istatistiksel olarak manidarlığı, oran testi kullanılarak belirlenmiştir. Yapılan ROC analizi sürecindeki adımlar ve oran testinin detayları, birinci ve ikinci alt problemler için verilerin analizi ve yorumu kısmında açıklanmıştır.



DÖRDÜNCÜ BÖLÜM

BULGULAR VE YORUM

Bu bölümde araştırmanın her bir alt problemine ilişkin elde edilen bulgulara ve yorumlar yer almaktadır. Bulguların sunulmasında, araştırmadaki alt problemlerin sırası izlenmiştir.

4.1. Çoktan Seçmeli Formun Lojistik Regresyon Analizi, Kümeleme Analizi ve ROC Analizi ile Belirlenen Sınıflama Doğruluklarına İlişkin Bulgular

Bu alt problemin çözümü için ilk olarak ÇSF'nin LR analizi ile elde edilen sınıflama sonuçlarına ait çıktıların tablolaştırılmasına ve yorumlanmasına yer verilmiştir. Analiz dosyasında öğrencilerin ÇSF'den ve genel durumda geçme durumları '1' ile, kalma durumları ise '0' ile kodlanmıştır. Tablo 3'te başlangıç modeli (Block 0: Beginning Block) için iterasyon öyküsü verilmiştir.

Tablo 3. ÇSF'ye İlişkin LR Analizi Başlangıç Modeli İterasyon Öyküsü

İterasyon		-2Log Olabilirlik	Katsayılar
		(-2LL)	Sabit
Adım 0	1	386.358	.883
	2	386.076	.948
	3	386.075	.949

Tablo 3 incelendiğinde, -2LL değerinin sıfıra yaklaştıkça mükemmel uyuma işaret ettiği düşünülürse, başlangıç modelinden elde edilen 386.075 değerinin oldukça yüksek olduğu söylenebilir. Tablo 4'te LR analizi sonucunda başlangıç modeline göre elde edilen ilk sınıflandırma durumu gösterilmiştir.

Tablo 4. ÇSF'ye İlişkin LR Analizi Başlangıç Modeline Göre Elde Edilen İlk Sınıflandırma Durumu

Gerçek/Gözlenen Durum	Kestirilen Durum		
	Geçti	Kaldı	Doğru Sınıflandırma Yüzdesi
Geçti	0	91	0

Kaldı	0	235	100.0
Toplam Doğru Sınıflandırma Yüzdesi			72.1

Tablo 4 incelendiğinde, başlangıç modeline göre sınıflandırma sonuçları doğrultusunda, testi alan tüm öğrencilerin durumu ‘kaldı’ olarak sınıflandırıldığında, doğru sınıflandırma yüzdesinin %72.1 olduğu sonucu çıkarılmaktadır. Tablo 5’te bağımsız değişkenin bulunduğu model için iterasyon öyküsü verilmiştir.

Tablo 5. ÇSF’ye İlişkin LR Analizi Bağımsız Değişkenin Bulunduğu Model İçin İterasyon Öyküsü

İterasyon		Katsayılar	
Adım	-2LL	Sabit	ÇSF
1	207.961	3.485	-.067
2	158.084	5.922	-.117
3	141.626	8.212	-.165
4	138.322	9.784	-.197
5	138.121	10.295	-.208
6	138.120	10.335	-.208
7	138.120	10.336	-.208

Başlangıçta Tablo 3’te verilen, yalnızca sabit terimini içeren başlangıç modeli için olan iterasyon öyküsünden farklı olarak Tablo 5’te daha güçlü kestirimler yapabilmek adına bağımsız değişkenin dâhil edildiği modele ait iterasyon öyküsü verilmiştir. Tablo 3’te 386.075 olan -2LL değerinin, Tablo 5’te 138.120’ye kadar düştüğü görülmektedir. Yani sadece sabit terimin bulunduğu başlangıç modeline, daha güçlü kestirimler yapabilmek için bağımsız değişkenler dâhil edildiğinde -2LL farkının 247.955 (386.075-138.120) olduğu görülmektedir. Bu durumda -2LL değerinin sıfıra yaklaştıkça model uyumunun mükemmele yaklaştığı dikkate alındığında, bağımsız değişkenin modele eklenmesinin anlamlı bir değişmeye neden olacağını göstermektedir. Tablo 6’da model katsayılarına ilişkin Omnibus Testi sonuçları verilmiştir.

Tablo 6. ÇSF’ye İlişkin LR Analizi Model Katsayılarına Dair Omnibus Testi

Adım	Ki-kare	sd	P
1	Adım 247.956	1	.00

Blok	247.956	1	.00
Model	247.956	1	.00

Tablo 6'daki model için ki-kare değerine ilişkin p değerinin manidar çıkması ($p < .05$), bağımsız değişken ile bağımlı değişkenin arasındaki ilişkinin varlığına işaret etmektedir. Aynı zamanda model ki-kare istatistiğinin manidarlığı, bağımsız değişkenin dâhil edildiği modelin olasılıkları kestirme gücünün, yalnızca sabit terimle oluşturulan başlangıç modeline göre anlamlı şekilde ve olumlu yönde farklılaştığı şeklinde yorumlanabilir. Ayrıca buradaki p değerinin .05'ten küçük olması sonucu reddedilen H_0 hipotezinin 'başlangıç modeliyle yordayıcı değişkenin dâhil edilmesiyle oluşan model arasında fark yoktur' şeklinde olması da, bağımsız değişkenle oluşturulan modelin başlangıç modeline göre daha güçlü olasılık kestirimleri yapılabileceğini ortaya koymaktadır. Tablo 7'de sonuç modelinin özeti verilmiştir.

Tablo 7. ÇSF'ye İlişkin LR Analizi Sonuç Modelinin Özeti

Adım	-2LL	Cox&Snell R ²	Nagelkerke R ²
1	138.120	.533	.767

Tablo 7'deki Cox&Snell R² ve Nagelkerke R² değişkenleri incelendiğinde, bağımsız değişkenin analize girmesi ile genel durum bağımlı değişkenindeki varyansın sırasıyla %53.3'ünün ve %76.7'sinin açıklandığı söylenebilir. Nagelkerke R² değerinin Cox&Snell R² değerinden daha yüksek olduğu da açıkça görülmektedir. Tablo 8'de lojistik regresyon modeli sonucu elde edilen sınıflandırma tablosu verilmiştir.

Tablo 8. ÇSF'ye İlişkin LR Analizi Sonucu Elde Edilen Sınıflandırma Tablosu

Gerçek/Gözlenen Durum	Kestirilen Durum		
	Geçti	Kaldı	Doğru Sınıflandırma Yüzdesi
Geçti	81	10	89.0
Kaldı	19	216	91.9
Toplam Doğru Sınıflandırma Yüzdesi			91.1

Tablo 8 incelendiğinde başlangıçtaki sınıflandırma tablosunda (Tablo 4) model, öğrencilerin %72.1'ini doğru sınıflandırırken, bağımsız değişken dâhil edildikten sonra elde edilen modelin doğru sınıflandırma yüzdesi %91.1'e yükselmiştir.

Bu alt problemin çözümü için ikinci olarak ÇSF'nin kümeleme analizi ile elde edilen sınıflama sonuçlarına ait çıktıların tablolaştırılmasına ve yorumlanmasına yer verilmiştir. Analiz dosyasında öğrencilerin ÇSF'den geçme durumları '1' ile, kalma durumları ise '0' ile kodlanmıştır. Elde edilen kümeleme analizinde geçen öğrenciler kümesi '1' ile, kalan öğrenciler kümesi ise '2' ile kodlanmıştır. Tablo 9'da her bir kümede yer alan öğrencilere ilişkin frekans ve yüzdeler gösterilmiştir.

Tablo 9. ÇSF'ye İlişkin Kümelerde Yer Alan Öğrencilerin Frekans ve Yüzdeleri

Küme	f	%
1	100	30.7
2	226	69.3
Toplam	326	100

Tablo 9 incelendiğinde, 100 öğrencinin geçen öğrenciler kümesinde, 226 öğrencinin ise kalanlar kümesinde yer aldığı görülmektedir. Bu öğrenci sayılarının yüzde karşılıkları ise sırasıyla %30,7 ve %69,3 şeklindedir. ANOVA sonuçları dikkate alındığında ise $p=.00$ bulunduğundan, ÇSF değişkeninin kümeleme analizinde anlamlı bir etkiye sahip olduğu ($p<.05$) sonucu çıkarılmaktadır. Başka bir ifadeyle, bu değişkendeki geçti/kaldı durumu kümelerine göre farklılık göstermektedir. Tablo 10'da kümeleme analizinin sonuçları ile genel durum (45 puanın altı/üstü) arasında çapraz tablo yapılmış ve elde edilen sınıflandırma tablosu verilmiştir.

Tablo 10. ÇSF'ye İlişkin Kümeleme Analizi Sınıflandırma Tablosu

Gerçek/Gözlenen Durum	Kestirilen Durum		
	Geçti	Kaldı	Doğru Sınıflandırma Yüzdesi
Geçti	81	19	81.0
Kaldı	10	216	95.6
Toplam Doğru Sınıflandırma Yüzdesi			91.1

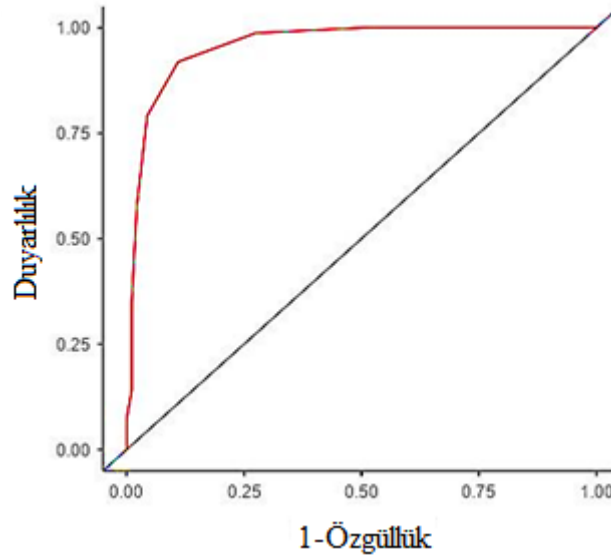
Tablo 10 incelendiğinde, ÇSF değişkeni ile elde edilen sınıflandırma tablosunda öğrencilerin %91.1'inin doğru sınıflandırıldığı görülmektedir.

Bu alt problemin çözümü için son olarak ÇSF'nin ROC analizi ile elde edilen sınıflama sonuçlarına ait çıktıların tablolaştırılmasına ve yorumlanmasına yer verilmiştir. Analiz dosyasında öğrencilerin ÇSF'den geçme durumları '1' ile, kalma durumları ise '0' ile kodlanmıştır. Tablo 11'de ROC analizinden elde edilen doğruluk (accuracy), özgüllük (specificity), duyarlılık (sensitivity) ve eğri altında kalan alanın büyüklüğünü gösteren AUC değerleri verilmiştir.

Tablo 11. ÇSF'ye İlişkin ROC Analizinden Elde Edilen Değerler

Doğruluk	Özgüllük	Duyarlılık	AUC
0.911	0.890	0.919	0.959

Tablo 11 incelendiğinde, ROC analizi yöntemiyle elde edilen sınıflama sonuçlarının doğruluk oranının yaklaşık olarak %91.1 olduğu görülmektedir. Bunun yanı sıra, özgüllük için %89 ve duyarlılık için %91.9 değerleri elde edilmiştir. Şekil 2'de ROC analizi çıktılarında yer alan eğri verilmiştir.



Şekil 2. ÇSF'ye İlişkin ROC Eğrisi

ÇSF için ROC eğrisi altında kalan alan istatistiksel olarak anlamlıdır ($p < .001$) ve Tablo 10'daki AUC değerinden de anlaşılacağı üzere %95.9 olarak bulunmuştur. Bu değer, yapılan sınıflamanın doğruluğunun göstergesi olduğu düşünüldüğünde, doğruluk ve AUC değerlerinin birbirine yakın olduğu sonucuna ulaşılabilir.

Bu alt probleme ilişkin tüm bulgular Tablo 12’de özetlenmiştir.

Tablo 12. Birinci Alt Probleme İlişkin Bulgular

Form	Yöntem	Sınıflama Doğruluğu
ÇSF	LR	%91.1
	Kümeleme	%91.1
	ROC Analizi	%95.9

Tablo 12’ye bakıldığında ÇSF’ye ilişkin LR, kümeleme ve ROC analizi yöntemleriyle yapılan sınıflama doğruluklarının sırasıyla %91.1, %91.1 ve %95.9 olduğu görülmektedir. LR ve kümeleme analizi sonuçları arasında fark bulunmazken, LR ve ROC analizi sonuçları arasında ROC analizi lehine %4.8’lik bir fark bulunmuştur. Bu fark, iki örnek arasındaki oranın farkının ve bu farkın öneminin test edildiği oran testi sonucunda istatistiksel olarak manidar bulunmuştur ($z=2.498 > z_{tablo}=1.96$). Benzer şekilde kümeleme ve ROC analizi sonuçları arasında ROC analizi lehine %4.8’lik bir fark bulunmuştur. Bu fark, iki örnek arasındaki oranın farkının ve bu farkın öneminin test edildiği oran testi sonucunda istatistiksel olarak manidar bulunmuştur ($z=2.498 > z_{tablo}=1.96$).

4.2. Kısa Cevaplı Formun Lojistik Regresyon Analizi, Kümeleme Analizi ve ROC Analizi ile Belirlenen Sınıflama Doğruluklarına İlişkin Bulgular

Bu alt problemin çözümü için ilk olarak KCF’nin LR analizi ile elde edilen sınıflama sonuçlarına ait çıktıların tablolaştırılmasına ve yorumlanmasına yer verilmiştir. Analiz dosyasında öğrencilerin KCF’den ve genel durumda geçme durumları ‘1’ ile, kalma durumları ise ‘0’ ile kodlanmıştır. Tablo 13’te başlangıç modeli (Block 0: Beginning Block) için iterasyon öyküsü verilmiştir.

Tablo 13. KCF’ye İlişkin LR Analizi Başlangıç Modeli İterasyon Öyküsü

İterasyon		-2Log Olabilirlik	Katsayılar
		(-2LL)	Sabit
Adım 0	1	386.358	.883
	2	386.076	.948
	3	386.075	.949

Tablo 13 incelendiğinde, -2LL değerinin sıfıra yaklaştıkça mükemmel uyuma işaret ettiği düşünülürse, başlangıç modelinden elde edilen 386.075 değerinin oldukça yüksek olduğu söylenebilir. Tablo 14’te LR analizi sonucunda başlangıç modeline göre elde edilen ilk sınıflandırma durumu gösterilmiştir.

Tablo 14. KCF’ye İlişkin LR Analizi Başlangıç Modeline Göre Elde Edilen İlk Sınıflandırma Durumu

Gerçek/Gözlenen Durum	Kestirilen Durum		
	Geçti	Kaldı	Doğru Sınıflandırma Yüzdesi
Geçti	0	91	0
Kaldı	0	235	100.0
Toplam Doğru Sınıflandırma Yüzdesi			72.1

Tablo 14 incelendiğinde, başlangıç modeline göre sınıflandırma sonuçları doğrultusunda, testi alan tüm öğrencilerin durumu ‘kaldı’ olarak sınıflandırıldığında, doğru sınıflandırma yüzdesinin %72.1 olduğu sonucu çıkarılmaktadır. Tablo 15’te bağımsız değişkenin bulunduğu model için iterasyon öyküsü verilmiştir.

Tablo 15. KCF’ye İlişkin LR Analizi Bağımsız Değişkenin Bulunduğu Model İçin İterasyon Öyküsü

İterasyon Adım	Katsayılar		
	-2LL	Sabit	KCF
1	232.938	3.446	-.064
2	194.714	5.642	-.108
3	185.344	7.351	-.142
4	184.285	8.168	-.158
5	184.264	8.303	-.161
6	184.264	8.306	-.161
7	184.264	8.306	-.161

Başlangıçta Tablo 13’te verilen, yalnızca sabit terimini içeren başlangıç modeli için olan iterasyon öyküsünden farklı olarak Tablo 15’te daha güçlü kestirimler yapabilmek adına bağımsız değişkenin dâhil edildiği modele ait iterasyon öyküsü verilmiştir. Tablo 13’te

386.075 olan -2LL değerinin, Tablo 15’te 184.264’e kadar düştüğü görülmektedir. Yani sadece sabit terimin bulunduğu başlangıç modeline, daha güçlü kestirimler yapabilmek için bağımsız değişkenler dâhil edildiğinde -2LL farkının 201.811 (386.075-184.264) olduğu görülmektedir. Bu durumda -2LL değerinin sıfıra yaklaştıkça model uyumunun mükemmele yaklaştığı dikkate alındığında, bağımsız değişkenin modele eklenmesinin anlamlı bir değişmeye neden olacağını göstermektedir. Tablo 16’da model katsayılarına ilişkin Omnibus Testi sonuçları verilmiştir.

Tablo 16. KCF’ye İlişkin LR Analizi Model Katsayılarına Dair Omnibus Testi

Adım	Ki-kare	sd	p	
	Adım	201.811	1	.00
1	Blok	201.811	1	.00
	Model	201.811	1	.00

Tablo 16’daki model için ki-kare değerine ilişkin p değerinin manidar çıkması ($p < .05$), bağımsız değişken ile bağımlı değişkenin arasındaki ilişkinin varlığına işaret etmektedir. Aynı zamanda model ki-kare istatistiğinin manidarlığı, bağımsız değişkenin dâhil edildiği modelin olasılıkları kestirme gücünün, yalnızca sabit terimle oluşturulan başlangıç modeline göre anlamlı şekilde ve olumlu yönde farklılaştığı şeklinde yorumlanabilir. Ayrıca buradaki p değerinin .05’ten küçük olması sonucu reddedilen H_0 hipotezinin ‘başlangıç modeliyle yordayıcı değişkenin dâhil edilmesiyle oluşan model arasında fark yoktur’ şeklinde olması da, bağımsız değişkenle oluşturulan modelin başlangıç modeline göre daha güçlü olasılık kestirimleri yapılabileceğini ortaya koymaktadır. Tablo 17’de sonuç modelinin özeti verilmiştir.

Tablo 17. KCF’ye İlişkin LR Analizi Sonuç Modelinin Özeti

Adım	-2LL	Cox&Snell R^2	Nagelkerke R^2
1	184.264	.462	.665

Tablo 17’deki Cox&Snell R^2 ve Nagelkerke R^2 değişkenleri incelendiğinde, bağımsız değişkenin analize girmesi ile genel durum bağımlı değişkenindeki varyansın sırasıyla %46.2’sinin ve %66.5’inin açıklandığı söylenebilir. Nagelkerke R^2 değerinin Cox&Snell R^2

değerinden daha yüksek olduğu da açıkça görülmektedir. Tablo 18’de lojistik regresyon modeli sonucu elde edilen sınıflandırma tablosu verilmiştir.

Tablo 18. KCF’ye İlişkin LR Analizi Sonucu Elde Edilen Sınıflandırma Tablosu

Gerçek/Gözlenen Durum	Kestirilen Durum		
	Geçti	Kaldı	Doğru Sınıflandırma Yüzdesi
Geçti	61	30	67.0
Kaldı	10	225	95.7
Toplam Doğru Sınıflandırma Yüzdesi			87.7

Tablo 18 incelendiğinde başlangıçtaki sınıflandırma tablosunda (Tablo 14) model, öğrencilerin %72.1’ini doğru sınıflandırırken, bağımsız değişken dâhil edildikten sonra elde edilen modelin doğru sınıflandırma yüzdesi %87.7’ye yükselmiştir.

Bu alt problemin çözümü için ikinci olarak KCF’nin kümeleme analizi ile elde edilen sınıflama sonuçlarına ait çıktıların tablolaştırılmasına ve yorumlanmasına yer verilmiştir. Analiz dosyasında öğrencilerin KCF’den geçme durumları ‘1’ ile, kalma durumları ise ‘0’ ile kodlanmıştır. Elde edilen kümeleme analizinde geçen öğrenciler kümesi ‘1’ ile, kalan öğrenciler kümesi ise ‘2’ ile kodlanmıştır. Tablo 19’da her bir kümede yer alan öğrencilere ilişkin frekans ve yüzdeler gösterilmiştir.

Tablo 19. KCF’ye İlişkin Kümelerde Yer Alan Öğrencilerin Frekans ve Yüzdeleri

Küme	F	%
1	102	31.3
2	224	68.7
Toplam	326	100

Tablo 19 incelendiğinde, 102 öğrencinin geçen öğrenciler kümesinde, 224 öğrencinin ise kalanlar kümesinde yer aldığı görülmektedir. Bu öğrenci sayılarının yüzde karşılıkları ise sırasıyla %31.3 ve %68.7 şeklindedir. ANOVA sonuçları dikkate alındığında ise $p=.00$ bulunduğundan, KCF değişkeninin kümeleme analizinde anlamlı bir etkiye sahip olduğu ($p<.05$) sonucu çıkarılmaktadır. Başka bir ifadeyle, bu değişkendeki geçti/kaldı durumu

kümelere göre farklılık göstermektedir. Tablo 20’de kümeleme analizinin sonuçları ile genel durum arasında çapraz tablo yapılmış ve elde edilen sınıflandırma tablosu verilmiştir.

Tablo 20. KCF’ye İlişkin Kümeleme Analizi Sınıflandırma Tablosu

Gerçek/Gözlenen Durum	Kestirilen Durum		
	Geçti	Kaldı	Doğru Sınıflandırma Yüzdesi
Geçti	75	27	73.5
Kaldı	16	208	92.9
Toplam Doğru Sınıflandırma Yüzdesi			86.8

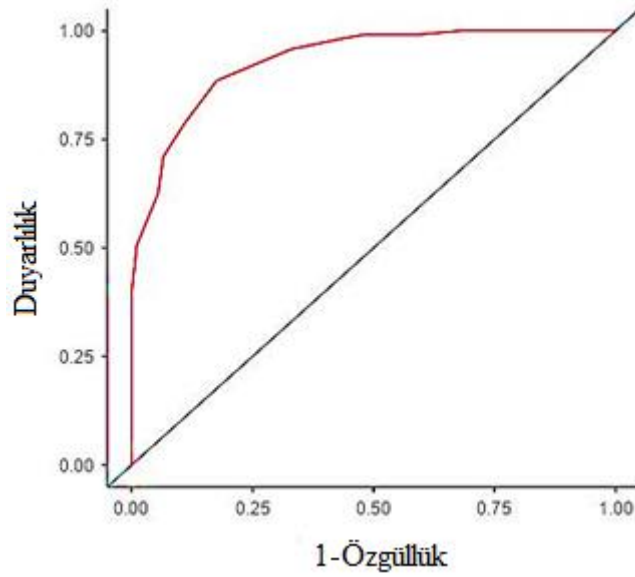
Tablo 20 incelendiğinde, KCF değişkeni ile elde edilen sınıflandırma tablosunda öğrencilerin %86.8’inin doğru sınıflandırıldığı görülmektedir.

Bu alt problemin çözümü için son olarak KCF’nin ROC analizi ile elde edilen sınıflama sonuçlarına ait çıktıların tablolaştırılmasına ve yorumlanmasına yer verilmiştir. Analiz dosyasında öğrencilerin KCF’den geçme durumları ‘1’ ile, kalma durumları ise ‘0’ ile kodlanmıştır. Tablo 21’de ROC analizinden elde edilen doğruluk (accuracy), özgüllük (specificity), duyarlılık (sensitivity) ve eğri altında kalan alanın büyüklüğünü gösteren AUC değerleri verilmiştir.

Tablo 21. KCF’ye İlişkin ROC Analizinden Elde Edilen Değerler

Doğruluk	Özgüllük	Duyarlılık	AUC
0.877	0.670	0.957	0.932

Tablo 21 incelendiğinde, ROC analizi yöntemiyle elde edilen sınıflama sonuçlarının doğruluk oranının yaklaşık olarak %87.7 olduğu görülmektedir. Bunun yanı sıra, özgüllük için %67 ve duyarlılık için %95.7 değerleri elde edilmiştir. Şekil 3’te ROC analizi çıktılarında yer alan eğri verilmiştir.



Şekil 3. KCF'ye İlişkin ROC Eğrisi

KCF için ROC eğrisi altında kalan alan istatistiksel olarak anlamlıdır ($p < .001$) ve Tablo 10'daki AUC değerinden de anlaşılacağı üzere %93.2 olarak bulunmuştur. Bu değer, yapılan sınıflamanın doğruluğunun göstergesi olduğu düşünüldüğünde, doğruluk ve AUC değerlerinin birbirine yakın olduğu sonucuna ulaşılabilir.

Bu alt probleme ilişkin tüm bulgular Tablo 22'de özetlenmiştir.

Tablo 22. İkinci Alt Probleme İlişkin Bulgular

Form	Yöntem	Sınıflama Doğruluğu
	LR	%87.7
KCF	Kümeleme	%86.8
	ROC Analizi	%93.2

Tablo 22'ye bakıldığında KCF'ye ilişkin LR, kümeleme ve ROC analizi yöntemleriyle yapılan sınıflama doğruluklarının sırasıyla %87.7, %86.8 ve %93.2 olduğu görülmektedir. LR ve kümeleme analizi sonuçları arasında LR lehine %0.9'luk bir fark bulunmuştur. Bu farkın manidar olmadığı tespit edilmiştir ($z = .345 < z_{tablo} = 1.96$). Benzer şekilde LR ve ROC analizi sonuçları arasında ROC analizi lehine %5.5'lik bir fark bulunmuştur. Bu fark, oran testi sonucunda istatistiksel olarak manidardır ($z = 2.4 > z_{tablo} = 1.96$). Kümeleme ve ROC analizi

sonuçları arasında ise yine ROC analizi lehine %6.4'lük bir fark söz konusudur. Bu fark, oran testi sonucunda istatistiksel olarak manidardır ($z=2.74 > z_{\text{tablo}}=1.96$).

4.3. Çoktan Seçmeli ve Kısa Cevaplı Formun Lojistik Regresyon Analizi ile Belirlenen Sınıflama Doğruluklarına İlişkin Bulgular

ÇSF ve KCF'nin LR yöntemiyle bulunan sınıflama doğrulukları sırasıyla birinci ve ikinci alt problemlerde verilmişti. Tablo 23'te elde edilen sonuçlar özetlenmiştir.

Tablo 23. ÇSF ve KCF'nin LR Yöntemiyle Elde Edilen Sınıflama Doğrulukları

Form	Sınıflama Doğrulukları
ÇSF	%91.1
KCF	%87.7

Tablo 23'e bakıldığında LR yöntemiyle elde edilen sınıflama doğrulukları arasında ÇSF lehine %3.4'lük bir fark bulunmuştur. Bu fark, oran testi sonucunda istatistiksel olarak manidar bulunmamıştır ($z=1.41 < z_{\text{tablo}}=1.96$).

4.4. Çoktan Seçmeli ve Kısa Cevaplı Formun Kümeleme Analizi ile Belirlenen Sınıflama Doğruluklarına İlişkin Bulgular

ÇSF ve KCF'nin kümeleme analizi yöntemiyle bulunan sınıflama doğrulukları sırasıyla birinci ve ikinci alt problemlerde verilmişti. Tablo 24'te elde edilen sonuçlar özetlenmiştir.

Tablo 24. ÇSF ve KCF'nin Kümeleme Analizi Yöntemiyle Elde Edilen Sınıflama Doğrulukları

Form	Sınıflama Doğrulukları
ÇSF	%91.1
KCF	%86.8

Tablo 24'e bakıldığında kümeleme analizi yöntemiyle elde edilen sınıflama doğrulukları arasında ÇSF lehine %4.3'lük bir fark söz konusudur. Bu farkın, oran testi sonucunda istatistiksel olarak manidar olmadığı tespit edilmiştir ($z=1.76 < z_{\text{tablo}}=1.96$).

4.5. Çoktan Seçmeli ve Kısa Cevaplı Formun ROC Analizi ile Belirlenen Sınıflama Doğruluklarına İlişkin Bulgular

ÇSF ve KCF'nin ROC analizi yöntemiyle bulunan sınıflama doğrulukları sırasıyla birinci ve ikinci alt problemlerde verilmiştir. Tablo 25'te elde edilen sonuçlar özetlenmiştir.

Tablo 25. ÇSF ve KCF'nin ROC Analizi Yöntemiyle Elde Edilen Sınıflama Doğrulukları

Form	Sınıflama Doğrulukları
ÇSF	%95.9
KCF	%93.2

Tablo 25'e bakıldığında ROC analizi yöntemiyle elde edilen sınıflama doğrulukları arasında ÇSF lehine %2.7'lik bir fark söz konusudur. Bu farkın, oran testi sonucunda istatistiksel olarak manidar olmadığı tespit edilmiştir ($z=1.52 < z_{\text{tablo}}=1.96$).



BEŞİNCİ BÖLÜM

SONUÇ VE ÖNERİLER

Bu bölümde araştırmanın bulgularına dayalı olarak elde edilen sonuçlar özetlenmiş ve ileride yapılacak olan araştırmalara yönelik önerilere yer verilmiştir.

5.1. Sonuçlar

Bu araştırmada geleneksel olmayan standart belirleme yöntemlerinden lojistik regresyon analizi, kümeleme analizi ve ROC analizi olmak üzere üç farklı istatistiksel yöntemin karşılaştırılması yapılmıştır. Araştırma kapsamında, araştırmacı tarafından 7. sınıf seviyesine uygun aynı özelliği ölçen çoktan seçmeli ve kısa cevaplı olmak üzere iki farklı test geliştirilmiştir. Geliştirilen bu testlerin her biri için öğrencilere ait sınıflama doğrulukları ilgili üç analizle belirlenmiştir. Yöntemlerden elde edilen sınıflama doğruluklarının arasındaki farkın istatistiksel olarak anlamlı olup olmadığı, oran testi ile karşılaştırma yapılarak test edilmiştir. Literatüre bakıldığında benzer araştırmaların genellikle İngilizce dersine ilişkin testler üzerinden gerçekleştirildiği görülmüştür. Ayrıca bu testlerdeki madde türleri de çoktan seçmeli ve açık uçlu maddelerden oluşmaktadır. Ayrıca bu madde türlerinin bulunduğu testler, öğrencileri farklı açılardan ölçmeyi hedeflemektedir. Ancak bu çalışmada geliştirilen her iki test de aynı davranışları ölçmeye yönelik olarak geliştirilmiş ve farklı türde maddelerden oluşturulmuştur.

Nihai formu toplam 14 maddeden oluşan ÇSF'nin LR analizi ile yapılan sınıflama doğruluğu %91,1, K-Ortalamalar kümeleme analizi ile yapılan sınıflama doğruluğu %91,1 ve ROC analizi ile yapılan sınıflama doğruluğu %95,9 olarak bulunmuştur. Elde edilen bu yüzelere bakıldığında LR ve kümeleme yöntemleri arasında bir fark bulunmazken; LR ve ROC analizlerinin sonuçları arasındaki fark istatistiksel olarak manidar ve ROC analizi lehinedir. Benzer şekilde kümeleme ve ROC analizlerinin sonuçları arasındaki fark da istatistiksel olarak manidar ve ROC analizi lehinedir. Yani tüm bu sonuçlara bakıldığında, ÇSF verilerine ilişkin öğrencilerin geçti/kaldı şeklindeki sınıflandırılmasına dair en doğru sonuç, ROC analizi ile elde edilmiştir. Benzer çalışmalarda, bir sınıflama söz konusu olduğunda, ROC analizinin kullanılmasının diğer analizlere göre daha iyi sonuçlar vereceği düşünülmektedir.

Nihai formu toplam 24 maddeden oluşan KCF'nin LR analizi ile yapılan sınıflama doğruluğu %87,7, K-Ortalamalar kümeleme analizi ile yapılan sınıflama doğruluğu %86,8 ve

ROC analizi ile yapılan sınıflama doğruluğu %93,2 olarak bulunmuştur. Elde edilen bu yüzelere bakıldığında LR ve kümeleme yöntemleri arasındaki fark istatistiksel olarak manidar bulunmamıştır. Buna karşılık LR ve ROC analizlerinin sonuçları arasındaki fark istatistiksel olarak manidar ve ROC analizi lehinedir. Benzer şekilde kümeleme ve ROC analizlerinin sonuçları arasındaki fark da istatistiksel olarak manidar ve ROC analizi lehinedir. Yani tüm bu sonuçlara bakıldığında, KCF verilerine ilişkin öğrencilerin geçti/kaldı şeklindeki sınıflandırılmasına dair en doğru sonuç, yine ROC analizi ile elde edilmiştir. Böylelikle bir başarı testinden elde edilen sınıflama doğrulukları düşünüldüğünde, madde türü çoktan seçmeli veya kısa cevaplı olduğunda en iyi sonucu ROC analizinin verdiği görülmüştür. Başka bir deyişle, madde türü değiştiğinde, sınıflama doğruluğuna ilişkin en iyi sonucu veren analiz yönteminin değişmediği ifade edilebilir.

ÇSF'nin LR analizi ile yapılan sınıflama doğruluğu %91,1 iken, KCF'nin aynı analizle yapılan sınıflama doğruluğu %87,7 olarak hesaplanmıştır. Bu iki yüzde arasındaki fark istatistiksel olarak manidar bulunmamıştır. Yani öğrencilerin geçti/kaldı sonuçlarına yönelik olarak yapılan sınıflamada ÇSF'nin veya KCF'nin kullanılması, analiz yöntemi LR olduğunda istatistiksel açıdan manidar bir fark ortaya koymamaktadır.

ÇSF'nin K-Ortalamlar kümeleme analizi ile yapılan sınıflama doğruluğu %91,1 iken, KCF'nin aynı analizle yapılan sınıflama doğruluğu %86,8 olarak hesaplanmıştır. Bu iki yüzde arasındaki fark istatistiksel olarak manidar bulunmamıştır. Yani öğrencilerin geçti/kaldı sonuçlarına yönelik olarak yapılan sınıflamada ÇSF'nin veya KCF'nin kullanılması, analiz yöntemi K-Ortalamlar kümeleme analizi olduğunda istatistiksel açıdan manidar bir fark ortaya koymamaktadır.

ÇSF'nin ROC analizi ile yapılan sınıflama doğruluğu %95,9 iken, KCF'nin aynı analizle yapılan sınıflama doğruluğu %93,2 olarak hesaplanmıştır. Bu iki yüzde arasındaki fark istatistiksel olarak manidar bulunmamıştır. Yani öğrencilerin geçti/kaldı sonuçlarına yönelik olarak yapılan sınıflamada ÇSF'nin veya KCF'nin kullanılması, analiz yöntemi ROC analizi olduğunda istatistiksel açıdan manidar bir fark ortaya koymamaktadır.

Özetle, sınıflama doğruluklarına ilişkin aynı testin farklı istatistiksel yöntemlerle yapılan analizleri doğrultusunda en iyi sonucun ROC analizi ile elde edildiği ifade edilebilir. Ancak aynı özelliği ölçen ve farklı madde türlerinden oluşan testlerin aynı istatistiksel yöntemle analizlerinin yapılmasıyla elde edilen sonuçlar arasında manidar bir farkın söz konusu olmadığı görülmüştür. Köksal (2011) yaptığı çalışmada ayırım gücü en iyi olan

modelin belirlenmesine ilişkin en iyi kestirimi ROC analizi ile yaptığını ortaya koymuştur. Bu bağlamda iki çalışmanın bulgularının birbiri ile paralellik gösterdiği söylenebilir. Benzer şekilde kesme puanı belirlenmesinde ROC analizinin kullanılabilceği sonucuna ulaşan Taşdemir (2013) gibi, bu çalışmada da sınıflama doğruluğuna ilişkin analizlerde ROC analizinin kullanılabilceği, hatta en iyi sonuçların yine bu analizle elde edilebileceği sonucuna ulaşılmıştır. Bu anlamda iki çalışmanın sonuçlarının benzer olduğunu söylemek mümkündür.

5.2. Öneriler

Pandemi döneminde veri toplama sürecinde yaşanan zorluklar nedeniyle ÇSF ve KCF'nin uygulanma esnasında öncelik sıralarının değiştirilmesinin, süreci daha da karmaşıklaştıracağı düşünüldüğünden, böyle bir uygulama gerçekleştirilememiştir. Verilerin yüz yüze toplanabilceği veya veri toplama sürecinin örnekleme bağı olarak daha iyi kontrol altında tutulabilceği başka bir çalışmada, testlerin öncelik sırası da örneklemleri farklılaştıran bir değişken olarak çalışmaya dâhil edilebilir.

Bu çalışmada aynı özelliği ölçen ve iki farklı formu geliştirilen başarı testleri kullanılmıştır. Aynı analiz yöntemleriyle başarı testleri yerine psikolojik ölçek(ler) kullanılarak benzer çalışmalar yapılabilir.

2020-2021 eğitim ve öğretim yılı ilk döneminde pandemi sebebiyle yüz yüze eğitim alamayan 7. sınıf öğrencilerinin karne notlarının belirlenmesinde canlı ders performansları, EBA üzerinden verilen ödevlerin yapılması gibi kriterler göz önüne alınmıştır. Ancak bu kriterlerin, maddi imkânsızlıklar, teknolojik cihaz kullanmadaki yetersizlikler, internet erişimine ilişkin yaşanan sorunlar nedeniyle tüm öğrenciler açısından standart olmayan kriterler olduğu düşünülmektedir. Bu sebeple araştırmaya, öğrencilerin ilgili dönemdeki matematik not ortalamaları değişken olarak dâhil edilmemiştir. Fakat tüm öğrenciler açısından fırsat eşitliğinin sağlandığının düşünüldüğü bir dönemde, matematik not ortalamasının da bağımsız değişken olarak araştırmaya dâhil edildiği benzer bir çalışma yapılabilir.

Bu çalışmada geleneksel standart belirleme yöntemlerine alternatif olan istatistiksel yöntemlerden LR, kümeleme ve ROC analizi seçilmiştir. Benzer bir araştırma, başka istatistiksel yöntemlerin karşılaştırılması ile de yapılabilir. Bunun ötesinde, geleneksel olan yöntemlerle istatistiksel yöntemlerin karşılaştırıldığı bir çalışma da yapılabilir.

Geliştirilen başarı testlerinde madde türü olarak çoktan seçmeli ve kısa cevaplı maddelerin kullanıldığı bu çalışmanın benzeri, farklı madde türleri ile oluşturulmuş testler geliştirilerek yapılabilir.

7. sınıf seviyesindeki öğrencilerden elde edilen veriler üzerinde çalışıldığı bu araştırmaya benzer şekilde farklı sınıf seviyesindeki öğrencilerden elde edilen veriler üzerinde çalışılabilir, hatta elde edilen sonuçlar, bu araştırmanın sonuçlarıyla karşılaştırılabilir.

Veri toplama aracı olarak sayısal bir ders olan matematik ile ilgili testler geliştirilmiş ve çalışma bu doğrultuda yürütülmüştür. Benzer bir çalışmada farklı disiplinlerde geliştirilen testler kullanılabilir, hatta sözel ve sayısal alanda geliştirilen farklı testler kullanılırsa, elde edilen bulgular bu doğrultuda karşılaştırmalı olarak yorumlanabilir.

Öğrenci başarısının ve onların akademik performans düzeylerine ilişkin değişimlerin doğru bir biçimde tespit edilebilmesi veya onlar hakkında kritik kararlar alma aşamasında standart belirlemenin önemli olması nedeniyle buna benzer çalışmaların yapılması önerilmektedir.

Bu çalışmada aynı özelliği ölçen iki başarı testi için madde türü farklı olsa da, en iyi sınıflama doğruluklarının ROC analizi ile elde edildiği görüldüğünden, benzer çalışmalarda da benzer amaçlar söz konusu olduğunda ROC analizinin kullanılması önerilmektedir.

KAYNAKÇA

- Aiken, R. L. (2000). *Psychological testing and assessment* (10th Ed.). Boston: Allyn and Bacon Inc.
- Alpar, R. (2010). *Uygulamalı istatistik ve geçerlik-güvenirlilik*. Ankara: Detay Yayıncılık.
- Alpar, R. (2011). *Uygulamalı çok değişkenli istatistiksel yöntemler*. Ankara: Detay Yayıncılık.
- Anastasi, A. (1988). *Psychological testing*. New York: Macmillan Publishing Company.
- Ary, D., Jacobs, L. C., Sorensen, C., ve Razavieh, A. (2010). *Introduction to research in education*. (8. Baskı) Wadsworth: Cengage Learning.
- Atılğan, H. (Ed.). (2014). *Eğitimde ölçme ve değerlendirme*. Ankara: Anı Yayıncılık.
- Atkin, J. M., Black, P. ve Coffey, J. (2001). *Classroom assessment and the national science education standarts*. National Academy Pres. Washington.
- Baykul, Y. (2015). *Eğitimde ve psikolojide ölçme: klâsik test teorisi ve uygulanması* (3. Baskı). Ankara: Pegem Yayınları.
- Brandon, P. R. (2002). Two versions of the contrasting-groups standard-setting method: A review. *Measurement and Evaluation in Counseling and Development*, 35(3), 167-181.
- Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö. E., Karadeniz, Ş. ve Demirel, F. (2019). *Bilimsel araştırma yöntemleri* (26. Baskı). Ankara: Pegem Yayınları.
- Can, A. (2019). *SPSS ile bilimsel araştırma sürecinde veri analizi* (7. Baskı). Ankara: Pegem Yayınları.
- Cizek, G. J. (1993). Reconsidering standards and criteria. *Journal of Educational Measurement*, 30.
- Cizek, G. J., ve Bunch, M. B. (2007). *Standart setting: a guide to establishing and evaluating performance standards on tests*. Sage Publications.
- Conway, M. J. (2006). *How to collect data: measurement and evaluation*. New York: American Society for Training and Development. Workplace Learning and Performance Press.

Cronbach, L. J. (1990). *Essentials of psychological testing*. (5th Ed.). New York: Harper Collins Publishers Inc.

Cronbach, L. J. ve Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281–302.

Çakıroğlu, S. & Soylu, N. (2019). İnternet oyun oynama bozukluğu ölçeğinin türkçe uyarlama, geçerlik ve güvenilirlik çalışması. *Türk Psikiyatri Dergisi*, 30(2), 130-136.

Çelik, Ş. (2013). Kümeleme analizi ile sağlık göstergelerine göre türkiye'deki illerin sınıflandırılması. *Doğuş Üniversitesi Dergisi*, 14(2), 175-194.

Çokluk, Ö., Şekercioglu, G. ve Büyüköztürk, Ş. (2018). *Sosyal bilimler için çok değişkenli istatistik SPSS ve LISREL uygulamaları* (5. Baskı). Ankara: Pegem Yayınları.

Dattalo, P. (1995). A comparison of discriminant analysis and logistic regression. *Journal of Social Service Research*, 19(3), 121-144.

Ebel, R. L. (1972). *Essentials of educational measurement*. New Jersey: Prentice Hall.

Erkuş, A. (2006). *Sınıf öğretmenleri için ölçme ve değerlendirme: kavramlar ve uygulamalar*. Ekinoks Yayınları. Ankara.

Gipps, C. (1994). *Beyond testing: towards a theory of educational assessment*. Routledge Falmer Press. Philadelphia. USA.

Grove, W. M. (2006). Mathematical aspects of diagnosis. *United States of America*, 50-75.

Gündeğer, C. (2012). *Angoff, yes/no ve ebel standart belirleme yöntemlerinin karşılaştırılması*. Yüksek lisans tezi, Hacettepe Üniversitesi, Ankara.

Hair, J. F., Black, W. C., Babin, B., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis* (6th Edition). Upper Saddle River, NJ: Prentice-Hall.

Han, J. & Kamber, M. (2006). *Data Mining concepts and techniques*. Morgan Kaufmann Publishers Inc.

Hanley, J. A. & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143, 29-36.

Hess, B., Subhiyah, R. G., & Giordano, C. (2007). Convergence between cluster analysis and the angoff method for setting minimum passing scores on credentialing examinations. *Evaluation & The Health Professions*, 30(4), 362-375.

Johnson, R. A. & Wichern, D. W. (1988). *Applied multivariate statistical analysis* (2nd Edition). Prentice Hall, Englewood Cliffs, New Jersey.

Kaftandjieva, F. (2010). *Methods for setting cut scores in criterion-referenced achievement tests a comparative analysis of six recent methods with an application to tests of reading in EF*, EALTA, Cito, Arnhem.

Kanık, E. A. & Erden, S. (2003). Tanı testlerinin değerlendirilmesinde ROC eğrisinin kullanımı. *Mersin Üniversitesi Tıp Fakültesi Dergisi*, 3, 260-264.

Kaufman, L., Rousseeuw, P.J. (1990). *Finding groups in data: an introduction to cluster analysis*, New York: John Wiley and Sons.

Kılıç, S. (2000). *Lojistik regresyon analizi ve pazarlama araştırmalarında bir uygulama*. Yayınlanmamış yüksek lisans tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.

Koyuncu, M. S. (2015). *Psikolojik ölçeklerde ROC analizi yöntemiyle standart belirleme*. Yüksek lisans tezi, Gazi Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.

Köksal, B. (2011). *Regresyon analizinde ROC eğrisi kestirimi ile model seçimi*. Yüksek lisans tezi, Marmara Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul.

Krzanowski, J.W. & David, J.H. (2009). *ROC curves for continuous data*. Chapman and Hall/CRC.

Livingston, S. A., & Zieky, M. J. (1982). *Passing scores: a manual for setting standards of performance on educational and occupational tests*. New Jersey: Educational Testing Service.

Mertler, C. A. & Vannatta, R. A. (2005). *Advanced and multivariate statistical methods: practical application and interpretation* (3rd Edition). Glendale, CA: Pyrczak Publishing.

Nartgün, Z. (2007). Aynı puanlar üzerinden yapılan mutlak ve bağıl değerlendirme uygulamalarının notlarda farklılık oluşturup oluşturmadığına ilişkin bir inceleme. *Ege Eğitim Dergisi* (8)1: 19-40

- Nitko, A.J. (1996). *Educational assessment of students*. (2nd Edition). Prentice Hall. New Jersey.
- Obuchowski, N.A. (1994). Computing sample size for receiver operating characteristic studies. *Investigative Radiology*, 29: 238-43.
- Obuchowski, N. A., Lieber, M. L. & Wians, F. H. (2004). ROC curves in clinical chemistry: uses, misuses and possible solutions. *Clinical Chemistry*, 50(7), 1118-25.
- Özçelik, D. A. (1992). *Ölçme ve değerlendirme* (Genişletilmiş 2. Baskı). Ankara: ÖSYM Yayınları.
- Özdamar, K. (2004). *Paket programlar ile istatistiksel veri analizi-2* (5. Baskı). Eskişehir: Kaan Kitabevi.
- Pallant, J. (2017). *SPSS kullanma kılavuzu: SPSS ile adım adım veri analizi* (2. Baskı), çeviren: Sibel Balcı-Berat Ahi. Ankara: Anı Yayıncılık.
- Pawar, M. S. (2004). *Data collecting methods and experiences: a guide for social researchers*. India: Sterling Publishers.
- Pektaş, A. O. (2013). *SPSS ile veri madenciliği*. İstanbul: Dikeyksen Yayın Dağıtım.
- Peng, C. Y. J., Lee, K. L., & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *The Journal of Educational Research*, 96(1), 3-14.
- Poggio, J. P., Glasnapp D. R. & Eros P. I. (1982). An Evaluation of Contrasting Groups Methods for Setting Standards. *Paper presented at The Annual Meeting of the American Educational Research Association*. New York: USA.
- Reyhanlıoğlu Keçeoğlu, Ç., Gelbal, S. & Doğan N. (2016). ROC eğrisi yöntemi ile kesme puanının belirlenmesi. *The Journal of Academic Social Science Studies*. 50 (2), 553-562.
- Reyhanlıoğlu Keçeoğlu, Ç. (2012). *Kesme puanının belirlenmesinde lojistik regresyon analizi, diskriminant analizi ve ROC eğrisi yöntemlerinin karşılaştırılması*. Yüksek lisans tezi, Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.

Şahin Sarkın, D. B. (2017). *Çoklu puanlanan ölçeklerde angoff, ROC ve sınır grup yöntemleri ile kesme puanının belirlenmesi*. Doktora tezi, Ankara Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.

Şata, M. & Karip, F. (2017). Akıllı telefon bağımlılığı ölçeği-kısa versiyonunun ergenler için türk kültürüne uyarlanması. *Cumhuriyet Uluslararası Eğitim Dergisi*, 6(4), 426-440.

Şata, M., Çorbacı, E. C., & Koyuncu, M. S. (2019). Mobil öğrenmeye hazırbulunmuşluk ölçeği'nin uyarlanması: geçerlik ve güvenirlik çalışması. *Gazi Üniversitesi Eğitim Fakültesi Dergisi*, 39(3), 1513-1533.

Taşdemir, F. (2013). *Angoff (1-0), nedelsky ve sınır değerleri saptama yöntemleri ile bir testin sınıflama doğruluklarının incelenmesi*. Doktora tezi, Ankara Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.

Tatlıdil, H. (2002). *Uygulamalı çok değişkenli istatistiksel analiz*. Ankara: Ziraat Matbaacılık A.Ş.

Tavakol, M. & Dennick, R. (2012). Standard setting: the application of the receiver operating characteristic method. *International Journal of Medical Education*, 3, 198-200.

Tayyar, N. (2010). Müşteri memnuniyeti tahmininde yapay sinir ağları, lojistik regresyon ve ayırma analizinin performanslarının karşılaştırılması. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 15(1), 339-355.

Tekin, H. (2019). *Eğitimde ölçme ve değerlendirme* (27. Baskı). Ankara: Yargı Yayınları.

Thorndike R. L.ve Hagen, E. (1961). *Measurement and evaluation in psychology and education*. Newyork: John Wiley and sons.

Tokmak, L. ve Bek, Y. (2010). İşlem karakteristik eğrisi analizi ve eğri altında kalan alanların karşılaştırılması. *OMÜ Deneysel ve Klinik Tıp Dergisi*, 27, 58-65.

Turgut, M. F. (1992). *Eğitimde ölçme ve değerlendirme metotları* (9. Baskı). Ankara: Saydam Matbaacılık

Turgut, M. F. ve Baykul, Y. (2019). *Eğitimde ölçme ve değerlendirme* (8. Baskı). Ankara: Pegem Akademi Yayıncılık.

Tülübaş, G. (2009). *Psikolojik testlerde angoff ve sınır grup yöntemleri ile kesme puanlarının belirlenmesi*, Yüksek lisans tezi, Hacettepe Üniversitesi, Ankara.

Vivo, J.M. ve Franco, M. (2008). How does one assess the accuracy of academic success predictors? ROC analysis applied to university entrance factors. *International Journal of Mathematical Education in Science and Technology*, 39(3), 325–340.



EKLER

EK 1. Uzman Görüş Formu

Sayın hocam,

Ortaokul 7. sınıf düzeyinde yürütülmekte olan Matematik dersi kapsamında aşağıdaki kazanımlar işlenmiştir.

- Tam sayılarla toplama ve çıkarma işlemlerini yapar, ilgili problemleri çözer.
- Toplama işleminin özelliklerini akıcı işlem yapmak için birer strateji olarak kullanır.
- Tam sayılarla çarpma ve bölme işlemlerini yapar.
- Tam sayıların kendileri ile tekrarlı çarpımını üslü nicelik olarak ifade eder.
- Rasyonel sayıları tanır ve sayı doğrusunda gösterir.
- Rasyonel sayıları ondalık gösterimle ifade eder.
- Devirli olan ve olmayan ondalık gösterimleri rasyonel sayı olarak ifade eder.
- Rasyonel sayıları sıralar ve karşılaştırır.

Sizden, bu kapsamı ölçmek amacıyla hazırlanmış olan maddeleri hem dersin kapsamı hem de madde yazma ilkelerine uygunluk açısından incelemeniz ve eleştirilerinizi/önerilerinizi sunmanız beklenmektedir. Lütfen, her bir madde için madde kökünün (ve tüm seçeneklerin) uygunluğunu değerlendiriniz ve görüşünüzü form üzerindeki ilgili kısımları işaretleyerek belirtiniz. Doğru cevaplar soruların altındaki kısımlarda koyu yazılarak belirtilmiştir. Öneri ve eleştirilerinizi “Açıklama” kutucuklarına yazabilirsiniz.

Katkılarınız için teşekkür ederiz.

	MADDE KÖKÜ VE (VARSA) SEÇENEKLERİ		Açıklama
1		Uygun	Uygun değil
	MADDENİN DOĞRU CEVABI	<input type="checkbox"/>	<input type="checkbox"/>
			Açıklama

EK 2. Çoktan Seçmeli Matematik Başarı Testi

1. Aşağıda verilen eşitliklerden hangisi doğrudur?

- A) $-6^2 = (-6)^2$
B) $(-4)^3 = (+8)^2$
C) $(-5)^0 = 0$
D) $-7^3 = (-7)^3$

2. Bir alışveriş merkezinin 3. katında bulunan İrem, arabasını almak için -2. katta bulunan otoparka inmiştir.

Buna göre, İrem arabasını almak için kaç kat aşağı inmiştir?

- A) -6 B) -4 C) 1 D) 5

3. Yanda Akın, Feride ve Hilal'in ellerinde tuttıkları kartonların üzerinde yazan rasyonel sayılar birbirine eşittir.

$\frac{-3}{5}$ <i>Akın'ın kartı</i>	$-\frac{3}{5}$ <i>Feride'nin kartı</i>	$\frac{?}{?}$ <i>Hilal'in kartı</i>
--	---	--

Buna göre, Hilal'in elindeki kartonda yazan rasyonel sayı aşağıdakilerden hangisidir?

- A) $\frac{-5}{3}$
B) $\frac{3}{-5}$
C) $\frac{3}{5}$
D) $\frac{5}{3}$

4. Tolga Öğretmen, sınıfı dörder kişilik gruplara ayırarak her bir gruba ödev olarak bir işlem veriyor. Gruplardan birine verdiği işlem ve bu gruptaki öğrencilerin sırasıyla yaptığı işlemler aşağıda verilmiştir.

İşlem: $8 + 9 + (-8) = ?$

Emine: $8 + (-8) + 9$

Serdar: $[8 + (-8)] + 9$

Kerem: $0 + 9$

Sare: 9

Buna göre, öğrencilerin kullandıkları toplama işlemi özellikleriyle ilgili aşağıdakilerden hangisi doğrudur?

- A) Emine ters eleman özelliğini kullanmıştır.
- B) Kerem birleşme özelliğini kullanmıştır.
- C) Sare etkisiz eleman özelliğini kullanmıştır.
- D) Serdar birim eleman özelliğini kullanmıştır.

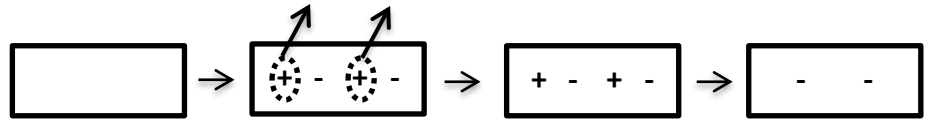
5. Aşağıda verilen ifadelerden hangisi rasyonel sayı değildir?

- A) 3,2 B) $\frac{5}{6}$ C) $\frac{8}{0}$ D) 0

6. Aşağıda verilen işlemlerden hangisinin sonucu diğerlerinden farklıdır?

- A) $25 : (-1)$
- B) $(-25) \cdot 1$
- C) $-25 : (-1)$
- D) $(+25) \cdot (-1)$

7. Yanda sayma pullarıyla modellenen işlem



aşağıdakilerden hangisidir?

- A) $(+1) \cdot (-2)$
- B) $(-2) \cdot (+1)$
- C) $(+2) : (-1)$
- D) $(-2) : (+1)$

8. Aşağıdaki rasyonel sayılardan hangisi sayı doğrusunda en solda yer alır?

- A) $\frac{-35}{6}$ B) -5 C) $\frac{24}{7}$ D) 4

9. Aşağıdaki sayılardan hangisinin toplama işlemine göre tersi kendisine eşittir?

- A) -2 B) -1 C) 0 D) 1

10.

<p>SORU $-5 - (+2) = ?$</p>

Yukarıda verilen işlemin doğru sonucu aşağıdakilerden hangisidir?

- A) -7 B) -3 C) 3 D) 7

11. Aşağıdaki rasyonel sayılardan hangisi sayı doğrusunda 0 ile -1 arasında bulunur?

- A) $\frac{-3}{4}$ B) $\frac{-9}{8}$ C) $-1\frac{2}{3}$ D) $\frac{-1}{0}$

12.

$\frac{-65}{140}$ $\frac{150}{300}$ $\frac{34}{45}$

Yukarıdaki rasyonel sayıların büyükten küçüğe doğru sıralanışı aşağıdakilerden hangisidir?

- A) $\frac{-65}{140} > \frac{150}{300} > \frac{34}{45}$ B) $\frac{34}{45} > \frac{150}{300} > \frac{-65}{140}$
C) $\frac{150}{300} > \frac{-65}{140} > \frac{34}{45}$ D) $\frac{150}{300} > \frac{34}{45} > \frac{-65}{140}$

13.

$\frac{-32}{125}$

Yukarıda verilen rasyonel sayının ondalık gösterimi aşağıdakilerden hangisidir?

- A) 0,256
B) 0,246
C) -0,246

D) -0,256

14.

$$\frac{-14}{25}$$

Yukarıda verilen rasyonel sayının ondalık gösterimi aşağıdakilerden hangisidir?

- A) -0,56
- B) -0,46
- C) 0,46
- D) 0,56

EK 3. Kısa Cevaplı Matematik Başarı Testi

1. Bir alışveriş merkezinin 3. katında bulunan İrem, arabasını almak için -2. katta bulunan otoparka inmiştir. **Buna göre, İrem'in arabasını almak için kaç kat aşağı indiğini yazınız.**

Cevabımız:

2. $-\frac{3}{5}$ rasyonel sayısına denk bir rasyonel sayı yazınız.

Cevabımız:

3. Tolga Öğretmen, sınıfı dörder kişilik gruplara ayırarak her bir gruba ödev olarak bir işlem veriyor. Gruplardan birine verdiği işlem ve bu gruptaki öğrencilerin sırasıyla yaptığı işlemler aşağıda verilmiştir.

İşlem: $8 + 9 + (-8) = ?$

Emine: $8 + (-8) + 9$

Serdar: $[8 + (-8)] + 9$

Kerem: $0 + 9$

Sare: 9

Buna göre, toplama işleminin etkisiz eleman özelliğini kullanan kişinin ismini yazınız.

Cevabımız:

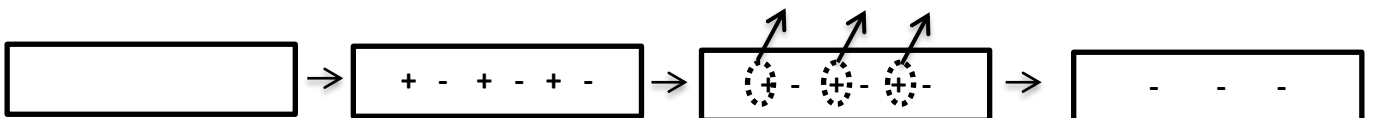
4. 'Paydası sıfır olan sayılar rasyonel sayı değildir.' ifadesinden yola çıkarak rasyonel olmayan bir ifade yazınız.

Cevabımız:

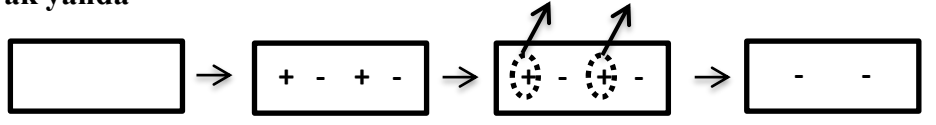
5. $(-25) : (-1)$ işleminin sonucunu yazınız.

Cevabımız:

6. **Örnek:** $(-3) \cdot (+1)$ işleminin sayma pullarıyla modellenmesi aşağıda verilmiştir.



Yukarıdaki örneği dikkate alarak yanda sayma pullarıyla modellenen çarpma işlemini yazınız.



Cevabınız:

7. Sayı doğrusunda 2'nin sağında yer alan bir rasyonel sayı yazınız.

Cevabınız:

8. Toplama işlemine göre tersi kendisine eşit olan sayıyı yazınız.

Cevabınız:

9. $(-5)^3$ üslü ifadesini tekrarlı çarpım şeklinde yazınız.

Cevabınız:

10. $-5 - (+2)$ işleminin sonucu kaçtır?

Cevabınız:

11. Sayı doğrusunda 0 ile -1 arasında bulunan bir rasyonel sayı yazınız.

Cevabınız:

12. $\frac{-65}{140}$, $\frac{150}{300}$, $\frac{34}{45}$ rasyonel sayılarını büyükten küçüğe doğru sıralayınız.

Cevabınız:

13. $\frac{-32}{125}$ rasyonel sayısının ondalık gösterimini yazınız.

Cevabınız:

14. $-1, \blacksquare = \frac{-14}{9}$ eşitliğinde devirli ondalık gösterimdeki \blacksquare yerine gelmesi gereken sayıyı yazınız.

Cevabınız:

15. $\frac{-14}{25}$ rasyonel sayısının ondalık gösterimini yazınız.

Cevabımız:

16. $(-2020)^0$ üslü ifadesinin değerini yazınız.

Cevabımız:

17. $(-7) + (-7)$ işleminin sonucunu yazınız.

Cevabımız:

18. $9 + \triangle$ işleminin sonucu çarpma işleminin yutan elemanına eşittir. **Buna göre, \triangle ile ifade edilen sayıyı bulunuz.**

Cevabımız:

19. $-0,\overline{8}$ devirli ondalık gösterimini rasyonel sayı olarak yazınız.

Cevabımız:

20. Bir şehirde belli bir zaman diliminde ölçülen hava sıcaklığı 2°C 'dir. **Bu şehirdeki hava sıcaklığının toplama işleminin etkisiz elemanına eşit olması için kaç $^{\circ}\text{C}$ düşmesi gerekir?**

Cevabımız:

21. $105 + (-105)$ işleminin sonucunu yazınız.

Cevabımız:

22. Bir binanın -3. katındaki otoparkında bulunan Leyla'nın 4. kattaki bir dairede bulunan Gül'ün yanına gitmesi için kaç kat çıkması gerektiğini yazınız.

Cevabımız:

23. $(218) : (-1)$ işleminin sonucunu yazınız.

Cevabımız:

24. $0,\overline{23}$ devirli ondalık gösterimini rasyonel sayı olarak yazınız.

Cevabımız:

EK 4. Arařtırma Uygulama İzni

