BULANIK LOJİSTİK REGRESYONDA PARAMETRE TAHMİNLERİ İÇİN YENİ BİR YAKLAŞIM VE BİR UYGULAMA

Gültekin ATALIK

Yüksek Lisans Tezi

İstatistik Anabilim Dalı

Ağustos - 2014

Bu tez çalışması Anadolu Üniversitesi Bilimsel Araştırma Projeleri Komisyonu Başkanlığı tarafından desteklenmiştir. Proje No: 1307F285

ANADOLU ÜNİVERSİTESİ

JÜRİ VE ENSTİTÜ ONAYI

Gültekin ATALIK'ın "Bulanık Lojistik Regresyonda Parametre Tahminleri için Yeni Bir Yaklaşım ve Bir Uygulama" başlıklı **İstatistik** Anabilim Dalındaki, Yüksek Lisans Tezi 20.08.2014 tarihinde, aşağıdaki jüri tarafından Anadolu Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin ilgili maddeleri uyarınca değerlendirilerek kabul edilmiştir.

		Adı Soyadı	**İmza**
Üye (Tez Danışmanı)	**:**	**Doç. Dr. Sevil ŞENTÜRK**	
Üye	**:**	**Prof. Dr. Veysel YILMAZ**	
Üye	**:**	**Doç. Dr. Zerrin AŞAN GREENACRE**	

Anadolu Üniversitesi Fen Bilimleri Enstitüsü Yönetim Kurulu'nun …………….. tarih ve …………….. sayılı kararıyla onaylanmıştır.

Enstitü Müdürü

ÖZET

Yüksek Lisans Tezi

BULANIK LOJİSTİK REGRESYONDA PARAMETRE TAHMİNLERİ İÇİN YENİ BİR YAKLAŞIM VE BİR UYGULAMA

Gültekin ATALIK

Anadolu Üniversitesi
Fen Bilimleri Enstitüsü
İstatistik Anabilim Dalı

Danışman: Doç. Dr Sevil ŞENTÜRK
2014, 94 sayfa

Bağımlı değişken ile bağımsız değişkenler arasındaki sebep sonuç ilişkisini tanımlayan modellere, regresyon modelleri denir. Ele alınan bağımlı değişken kategorik ise bu durumda kullanılan regresyon modeline ise lojistik regresyon adı verilmektedir. Lojistik regresyonun uygulanabilmesi için belirli varsayımların sağlanması gerekmektedir. Bulanık mantık ve istatistik teorilerinin birleşmesinden doğan "Bulanık Lojistik Regresyon" ise, bu varsayımların sağlanamadığı ya da yapısı gereği bulanıklık içeren veriler söz konusu olduğunda kullanılabilen bir yöntemdir. Bu tez çalışmasında bulanık lojistik regresyon modelinin teorik yapısından bahsedilmiş ve bu modelin bilinmeyen parametrelerinin tahmini için yeni bir yaklaşım önerilmiştir. Önerilen yaklaşım doğum ağırlığı verilerine uygulanmış ve modelin geçerliliği uyum iyiliği kriteri olan Ortalama üyelik derecesine göre yorumlanmıştır.

Anahtar Kelimeler: Bulanık Lojistik Regresyon, MDM Kriteri, Tanaka Regresyon Modeli, Revize Tanaka Regresyon Modeli

ABSTRACT

Master of Science Thesis

A NEW APPROACH FOR PARAMETER ESTIMATION IN FUZZY LOGISTIC REGRESSION AND AN APPLICATION

Gültekin ATALIK

Anadolu University
Graduate School of Sciences
Department of Statistics

Supervisor: Assoc. Prof. Dr. Sevil ŞENTÜRK
2014, 94 pages

Models define the relationship between dependent and independent variables called regression models. However, if the considered dependent variable is categorical, it's named logistic regression. To be able to apply logistic regression, certain hypothesis' should be provided. Composed of the fuzzy set and statistical theories, fuzzy logistic regression is a method can be used when these hypothesis' can't be provided or data include structural vagueness. In this thesis fuzzy logistic regression model's theoretical structure mentioned and a new approach is suggested for this model's unknown parameter's estimation. Suggested approach is applied to birth weight and it's validity is interpreted by the help of a goodness of fit test called Mean Degree Membership.

Keywords: Fuzzy Logistic Regression, MDM Criteria, Tanaka Regression Model, Revised Tanaka Regression Model

TEŞEKKÜR

Yüksek lisans çalışmam boyunca sürekli yanımda olan, bana her zaman her türlü konuda bilgi ve deneyimlerini esirgemeyen çok değerli hocam ve tez danışmanım **Doç. Dr. Sevil ŞENTÜRK**'e en içten teşekkürlerimi sunarım.

Ayrıca çalışmamın gerçekleşmesinde katkılarıyla bana yön veren hocalarım **Prof. Dr. Veysel YILMAZ**'a ve **Doç. Dr. Zerrin AŞAN GREENACRE**'ye en içten teşekkürlerimi sunarım.

Hayatımın her aşamasında sürekli yanımda olan, sevgi ve sıcaklılarını hiç eksik etmeyen ve her türlü konuda desteklerini benden esirgemeyen, sevgili **aileme** en içten teşekkürlerimi sunarım.

Yüksek lisans eğitimim boyunca özveri ve desteğini bir an olsun eksik etmeyen, daima yanımda olan gülüşü ile içimi ısıtan hayat arkadaşım değerli eşim **Özlem UÇAR ATALIK**'a teşekkür ederim.

İÇİNDEKİLER

Sayfa

ŞEKİLLER DİZİNİ

ÇİZELGELER DİZİNİ

1. GİRİŞ

Bağımlı değişken ile bağımsız değişkenler arasındaki ilişkiyi matematiksel olarak tanımlayan yönteme regresyon denilmektedir. Regresyon yönteminde istatistiksel analizlerde sıkça kullanılmaktadır. Klasik regresyon analizinde aralarında ilişki kurulacak değişkenler sayısal olmak zorundadır. Doğrusal regresyon analizi çeşitli varsayımlara dayalıdır. Bu varsayımlar sağlanmadığında, Bulanık Doğrusal Regresyon analizi değişkenler arasındaki ilişkiyi modelleme ve gelecekteki durumlar için tahmin yapmada daha başarılı olabilmektedir.

Lojistik regresyon analizi ise bağımlı değişkenin kategorik olduğu durumlarda kullanılan bir regresyon yöntemidir. Ele alınan bağımlı değişken "olumlu - olumsuz", "hasta - hasta değil" gibi kategorilerden oluştuğunda lojistik regresyon analizi tercih edilmektedir. Lojistik regresyon analizi uygulamasının ve yorumlanmasının kolay olmasından dolayı sıklıkla tercih edilen bir yöntemdir.

Aristo ile başlayan klasik (iki değerli) mantık, mantık sistemleri ve bilimin geçen yüzyıllar boyunca ilerlemesiyle önce üç değerli mantığa ardından çok değerli mantığa ve en nihayetinde bulanık mantığa dönüşmüştür. Bulanık mantık, tüm mantık sistemlerini kapsayan bir mantık türüdür.

İnsan düşünce ve deneyimlerinin etkin olduğu sistemlerin modellenmesi zordur. Çünkü bu sistemler belirsiz bir yapıya sahiptirler. Belirsizliği anlama ve tanımlama da çeşitli teoriler kullanılmaktadır. Belirsizliği modellemede olasılık teorisi ve istatistiksel yöntemler sıkça kullanılmaktadır. Ancak günlük yaşamda karşılaşılan belirsizliklerin hepsi rassallıkla açıklanamayabilir. Rassal olmayan belirsizlikleri tanımlamak için sayısal belirsizlikleri dikkate alan istatistik ve olasılık teorilerini kullanmak doğru olmayabilir. Bu tür belirsizlikleri modelleme de Zadeh tarafından 1960'lı yıllarda önerilmiş bulanık küme teorisi daha başarılı olmaktadır. Bulanık küme teorisi, nitel değişkenleri sayısal olarak değerlendirmemize olanak vermektedir.

Bulanık lojistik regresyon analizi, klasik lojistik regresyon analizinin varsayımlarının sağlanamadığı ya da verilerin yapısı gereği bulanık olduğu durumlarda kullanılan bulanık küme teorisine dayalı bir regresyon yöntemidir.

Bu çalışmanın amacı, bulanık lojistik regresyon analizini ayrıntılı bir biçimde incelemek ve model parametrelerini tahmin etmek için yeni bir yaklaşım sunmaktır. Pourahmad ve ark. (2011a) tarafından önerilen Bulanık Lojistik Regresyon yönteminde model parametrelerinin tahmini Tanaka'nın Bulanık Regresyon modeline dayalı olarak elde edilmiştir. He ve ark. (2007) ise yaptıkları çalışmalarında Tanaka'nın Bulanık Regresyon modelini gözden geçirmiş ve o modelden daha etkili bir model önermişlerdir. Bu çalışmada bulanık lojistik regresyon modelinin parametrelerinin tahmininde He ve ark. (2007) tarafından önerilen Tanaka'nın Bulanık Regresyon Modelinin Gözden Geçirilmiş Metodu ile "Tanaka'nın Revize Bulanık Regresyon Modeline Dayalı Bulanık Lojistik Regresyon" analizinin teorik yapısının oluşturulması ve uygulamasının gerçekleştirilmesi literatürde ilk defa gerçekleştirilmiştir.

Çalışma 6 bölümden oluşmaktadır ve birinci bölümde literatür taramasına yer verilmiştir.

Çalışmanın ikinci bölümünde Bulanık Küme Teorisi yer almaktadır. Mantık ve bulanık mantıktan bahsedilmiş, kısaca olabilirlik teorisine değinilmiş ve daha sonra olasılık ve olabilirlik teorisinin karşılaştırılması yapılmıştır. Ayrıca bulanık kümelerden, bulanık sayılar ve bu sayılar üzerinde tanımlı cebirsel işlemlerden bahsedilmiştir. Bu bölümde son olarak durulaştırmanın tanımı verilmiş ve kısaca durulaştırma yöntemlerine değinilmiştir.

Üçüncü bölümde ise bulanık regresyonun tanımı yapılmış daha sonra posibilistik ve en küçük karelere dayalı bulanık regresyon yöntemlerinden bahsedilmiştir.

Çalışmanın özgün kısmını oluşturan dördüncü bölümde öncelikle bulanık lojistik regresyon yönteminin genel tanımı verilmiştir. Daha sonra Nagar ve Srivastava (2008) ve Pourahmad ve arkadaşları (2011a) tarafından önerilmiş olan Bulanık Lojistik Regresyon modellerinden bahsedilmiştir. Pourahmad ve arkadaşları tarafından önerilmiş olan bulanık lojistik regresyon yönteminin katsayılarının elde edilmesi için yeni bir yaklaşım önerilmiştir.

Beşinci bölümde ise uygulama yer almaktadır. Uygulamada Bulanık lojistik regresyon yöntemi bebeklerin doğum ağırlığı verilerine uygulanmıştır. Bulanık lojistik regresyon modelinin katsayıları hem Pourahmad ve ark. (2011a) tarafından önerilen yöntemle hem de ilk defa bu çalışmada önerilen Tanaka'nın Revize Bulanık Regresyon Modeline Dayalı Bulanık Lojistik Regresyon yaklaşımıyla elde edilmiştir. Farklı *h* seviyelerinde iki yaklaşım için bulanık lojistik regresyon modelleri oluşturulmuş ve bir uyum iyiliği kriteri olan "Ortalama Üyelik Derecesi" bu modeller için hesaplanmıştır.

Çalışmanın son bölümünde ise beşinci bölümde uygulaması yapılan yaklaşımlar karşılaştırılmış ve bu yaklaşımlara ait yorumlara yer verilmiştir.

1.1. Literatürde Yapılan Çalışmalar

Zadeh tarafından bulanık küme teorisinin tanıtılmasından günümüze kadar, bulanık bilgiyi dikkate alan bulanık regresyon modelleri pek çok bilim dalında başarıyla uygulanmıştır. Bulanık regresyon modeli ile yapılan çalışmalar aşağıdaki gibi özetlenebilir.

Tanaka ve ark. (1982), ilk bulanık regresyon çalışmasını önermişlerdir. Bu çalışmada bağımlı ve bağımsız değişkenlerin bulanık olmadığı, fakat sistem bilgisinin bulanık olduğu varsayılmaktadır. Model parametrelerinin tahmininde kullanılan amaç fonksiyonu bağımlı değişkenin tahmin değerinin yayılımının minimizasyonuna dayanmaktadır. Analiz doğrusal programlama tekniği kullanılarak çözümlenmektedir.

Diamond (1988), klasik en küçük kareler regresyon yöntemine benzer bulanık en küçük kareler regresyon yöntemini bulanık üçgen sayılar üzerinde metrik tanımlaması yaparak ortaya koymuştur.

Moskowitz ve Kim (1993), bulanık doğrusal regresyonda üyelik fonksiyonları şekilleri, h değeri ve bulanık parametrelerin yayılmaları arasındaki ilişkiyi belirlemişlerdir. Üyelik fonksiyonunun aldığı şekil ve h değerine göre yayılmanın duyarlılığını araştırmışlardır.

Redden ve Woddall (1996), Sakawa ve Yano'nun modelini ortogonal en küçük kareler yardımıyla geliştirmişler ve yeni bir doğrusal programlama problemi önermişlerdir.

Chang ve Lee (1996), aykırı değer olması durumu için üyelik dereceleriyle ağırlıklandırma yapan ve karar verici ile etkileşime dayanan genelleştirilmiş bulanık ağırlıklandırılmış en küçük kareler yöntemi ileri sürmüşlerdir.

Yang ve Ko (1996), ağırlıklandırılmış bulanık en küçük kareler çözümlemesinin iteratif algoritmasını ileri sürmüşlerdir. İki aşamadan oluşan bu algoritmada ilk olarak gözlemlerin sınıf üyeliklerini veren bulanık sınıflama yöntemi seçilir, daha sonra üyeliklerin bu değerleri ağırlıklar olarak kullanılır.

Chang ve ark. (1997), Tayvan'da pis su arıtma tesisinin maliyet tahmininin yapılması için bulanık regresyon yöntemini geliştirerek, bulanık amaçlı programlama tekniği yöntemini önermişlerdir.

Wang ve Tsaur (2000a), bulanık regresyon modelinde bulanık aralıklar üzerinde durarak aralık regresyon analizinden bahsetmişlerdir.

Wang ve Tsaur (2000b), Tanaka tarafından tanımlanmış kesin bağımsız değişken ve bulanık bağımlı değişkenli problemin çözümü için bulanık en küçük kareler yöntemini önermiştir.

Özelkan ve Duckstein (2000) Arizona'da bulunan Walnut Gulch boşaltma havzasının Lucky Hills alt havzasıyla yağış akış ilişkisini modellemişlerdir. Modellemede Peters (1994) tarafından önerilen bulanık aralıklı bulanık doğrusal regresyon yöntemini kullanmışlar ve bu yöntemi geliştirmişlerdir. Klasik regresyon, bulanık aralıklı bulanık doğrusal regresyon ve önerilen bulanık regresyon yöntemleri örnek havza gözlem verilerine uygulanmış ve uygulama sonunda önerilen bulanık aralıklı bulanık doğrusal regresyon modelinin en iyi performansı gösterdiği gözlemlenmiştir.

Ishibuchi ve Nii (2001), bulanık regresyon yönteminin bazı kısıtlarından söz etmişler ve simetrik üçgen bulanık sayı tipinde olan katsayıların asimetrik üçgen ve yamuk sayı tipine uzanmasını yapmışlardır.

Chang ve Ayyub (2001), çalışmalarında kapsamlı bir literatür taraması yapılarak bulanık regresyon ve klasik regresyon arasındaki farkları tanımlamışlar ve bulanık regresyonun üç yaklaşımını özetlemişlerdir. Bu yaklaşımlardan ilki en uygun ölçüt ile bulanıklığın minimizasyonu esasına dayanır. İkinci yaklaşımda uygun kriter olarak en küçük kareler kullanılmaktadır ve makalede iki yöntem özetlenmiştir. Üçüncü yaklaşım aralıklı regresyon analizi olarak tanımlanmaktadır.

Lee ve Chen (2001), genelleştirilmiş bulanık doğrusal regresyon modelini sunmuşlar ve modelin bulanık parametrelerini belirlemek için doğrusal olmayan programlama modeli önermişlerdir.

Özelkan ve Duckstein (2001), Arizona'da Walnut Gulch havzasında yağış akış modellemesi için klasik regresyon ve bulanık regresyon yöntemlerini uygulamışlardır. Uygulama sonucunda ortaya çıkan sapmaların bulanık doğrusal regresyon yöntemleriyle minimuma indirgenmesi sağlanmıştır.

Tseng ve Tzeng (2002), bulanık regresyon ve mevsimsel zaman serisinin avantajlı yönlerini birleştiren bulanık mevsimsel ARIMA (SARIMA) yöntemi önermişlerdir.

Tran ve Duckstein (2002), klasik ve bulanık regresyon modellerinin sırasıyla merkezi eğilim ve olabilirlik özelliklerini birleştiren çok amaçlı bulanık regresyon modeli sunmuştur.

Wu ve Tseng (2002), en küçük kareler yaklaşımı ile bulanık parametre tahminli bulanık regresyon modeli sunmuşlardır.

Yang ve Lin (2002a), bulanık girdi ve bulanık çıktı değişkenleri için bulanık en küçük kareler yaklaşımı altında iki tahminleme yöntemi önermişlerdir. Heterojen veri kümesi ve aykırı değerleri belirlemek için kümeleme analizinden yararlanmışlardır.

Yang ve Liu (2002b), etkileşimli bulanık doğrusal regresyon modelleri için bulanık en küçük kareler algoritmasını önermişlerdir. Bu algoritmalar basit regresyon için aykırı değerlere karşı dayanıklıdır. Bu algoritmalarda ortogonal koşullar optimizasyon problemine kısıt olarak eklenmiştir.

Wu (2003), bulanık küme teorisindeki çözüm benzerliği yardımıyla regresyon parametrelerinin bulanık tahminlerini elde eden bir yöntem önermiştir.

Nasrabadi ve Nasrabadi (2004), bulanık ve bulanık olmayan bağımlı değişken ile bulanık ve bulanık olmayan bağımsız değişkenli bulanık doğrusal regresyon modelini ele almış, matematiksel programlama tekniğine dayanarak bir tahminleme yöntemi önermişlerdir.

Hojati ve ark. (2005), sadece bağımsız değişkenlerin bulanık, hem bağımlı hem bağımsız değişkenlerin bulanık olduğu iki durum için bulanık düşünüş altında hesaplanan yeni bir yöntem önermişlerdir.

Modarres ve ark. (2005), kesin girdi ve bulanık çıktıya sahip bulanık regresyon modellerindeki parametreleri tahmin etmek için bir matematiksel programlama yöntemi önermişlerdir. Yöntem, hataların karelerinin minimize edilmesine dayanmaktadır.

Yücel (2005), tez çalışmasında bulanık regresyon yöntemini kullanarak, Türkiye'de 1980-2004 yıllarında kayıt dışı ekonominin tahminine yönelik bir uygulama sunmuştur.

Stahl, (2006), çalışmasında doğrusal bulanık regresyon modelini bulanık bağımsız değişken ve bulanık parametreler kullanarak kurmuştur. Model parametrelerinin tahminini en küçük kareler metodu kullanarak yapmıştır.

He ve ark. (2007), üretkenlik, müşteri memnuniyeti ve karlılık arasındaki ilişkiyi klasik regresyon, Tanaka regresyon ve yeni bir bulanık regresyon yaklaşımı kullanarak belirlemeye çalışmışlardır. Klasik en küçük kareler yöntemi kullanılarak model tahmin edilmiştir, fakat bu yöntemle üretkenlik, karlılık ve müşteri memnuniyeti gibi kesin olmayan veriler için regresyon katsayılarının tahmin edilmesi bazı kısıtlar doğurmuştur. Bu nedenle mevcut Tanaka'nın bulanık doğrusal regresyon yönteminde bu kısıtlar dikkate alınarak "Revize Tanaka bulanık doğrusal regresyon" yöntemi oluşturulmuştur.

Düzyurt (2008), tez çalışmasında klasik regresyon ve Bulanık regresyon hakkında temel bilgiler verdikten sonra Tanaka regresyon, Revize Tanaka regresyon ve en küçük kareler doğrusal regresyon çözümlemesi ile uygulamalar yapmıştır.

Lu ve Wang (2009), çalışmalarında tahmin edilen bağımlı değişkenin yayılımlarının gözlenen bağımlı değişkenin yayılımlarına uyduğu bir geliştirilmiş bulanık doğrusal regresyon modeli önermişlerdir ve dört adet sayısal örnek kullanarak modelin etkinliğini göstermişlerdir.

Berry-Stölzle ve ark. (2010), çalışmalarında Tanaka regresyon ve He ve ark. (2007) tarafından önerilmiş olan Revize Tanaka regresyona dayalı regresyon katsayılarının bulanıklığını ölçen bir test geliştirmişlerdir. Regresyon katsayılarının yayılımlarını, ilgili bağımsız değişkenler ve bağımlı değişken arasındaki ilişkinin bulanıklığının bir ölçüsü olarak gören bir istatistik olarak yorumlamaktadırlar. Geliştirdikleri testi Alman sigorta şirketleri verilerine uygulamışlardır.

Chou ve ark. (2013), çalışmalarında bulanık regresyon tahmin modelini uluslararası uçak kargo piyasasının talep tahminini elde etmek için kullanmışlardır.

Literatürde Bulanık Lojistik Regresyon Analizi ile çok sınırlı sayıda çalışma bulunmaktadır. Bu alanda yapılan çalışmalar aşağıda gösterildiği gibidir.

Nagar ve Srivastava (2008) çalışmalarında istatistiksel lojistik regresyon yöntemi ile bulanık mantık konseptini birleştirerek ikili bağımlı değişkenin tahmini için bir adaptif yöntem önermişlerdir. Önerdikleri modeli kanser verileri üzerinde denemişlerdir.

Dom ve ark. (2008), çalışmalarında klasik regresyon yöntemini bulanık mantık konseptiyle birleştirerek ikili bağımlı değişkenin tahmini için bir öğrenme algoritması önermişlerdir. Önerilen algoritma doğrusal dönüşüm ve bulanık regresyon içermektedir.

Pourahmad ve ark. (2011a), özellikle tıp alanında sıklıkla kullanılan lojistik regresyon yönteminin bazı hastalık teşhislerinde yetersiz kalmasından dolayı yeni terim olan "posibilistik odds"'ve bu terime dayalı bulanık lojistik regresyonu tanıtmışlardır. Bulanık lojistik regresyonun katsayı tahminlerinden ve model uyum iyiliği testlerinden bahsetmişler ve önerdikleri modeli şeker hastalığı verileri üzerinde denemişlerdir.

Pourahmad ve ark. (2011b), bulanık lojistik regresyon yönteminin katsayılarının tahmini için en küçük karelere dayalı bir yaklaşım önermiş ve modelin uygulamasını Lupus Eritematozus hastalığı verileri üzerinde denemişlerdir. Ayrıca modelin uyum iyiliğinin ölçülmesi için "capability index" adı verilen bir kriterden bahsetmişlerdir.

Namdari ve ark. (2014), çocukların iştahlarında folik asidin etkisini araştırmak için bulanık lojistik regresyon ve sıralı lojistik regresyonu kullanmışlar ve sonuçları yorumlamışlardır.

2. BULANIK MANTIK VE BULANIK KÜME TEORİSİ

Bulanık küme teorisi, tanımlanması güç veya anlaşılması zor olan kavramlara bir meydana gelme derecesi atayarak, onları belirli hale getirmeyi amaçlamaktadır. (Yücel, 2005). Bulanık lojistik regresyon yönteminin daha iyi anlaşılabilmesi için bu bölümde öncelikle bulanık mantık kavramı tanımlanacaktır. Sonrasında olasılık ve olabilirlik teorilerinden kısaca bahsedilerek, iki teori arasındaki farklar açıklanacaktır. Son olarak bulanık kümeler, bulanık sayılar, bulanık kümeler üzerindeki işlemler tanımlanacaktır.

2.1. Mantık ve Bulanık Mantık

Doğru öncüllerden doğru sonuçlar çıkartma biçimlerini inceleyen bilim dalına mantık denilmektedir. Mantığın asıl konusu akıl yürütmeler ve çıkarımlardır. Geleneksel mantık sistemleri, sadece belirli koşullarda oluşan, kesin doğruluk değerleri doğru ya da yanlıştan birine sahip önermelerle ilgilenmektedirler. Belirsizliklerle ilgilenmemektedirler. Geleneksel mantık sistemlerinde üçüncü bir durumun gerçekleşmesinin imkânsız olduğu varsayılmakta ve bu tür durumlara ise "paradoks" adı verilmektedir. Bir başka ifadeyle mantıkta, doğruluk önermeleri {"Yanlış", "Doğru"} ya da sayısal olarak {0,1} kümesinin elemanlarıyla ilişkilendirilen bir küme olarak görülebilir.

Aristoteles ile başlayan klasik mantıkta 3 temel ilke vardır. Bu ilkelerden ilki "Özdeşlik" ilkesidir. Bu ilkeye göre bir şey ne ise odur. İkincisi "Çelişmezlik" ilkesidir. Diğer bir deyişle bu ilkeye göre, bir şey hem kendi, hem de başka bir şey olamaz. İlkelerden sonuncusu ise "Üçüncünün Olmazlığı" ilkesidir. Bu ilke bir şeyin ya A ya da A'nın olmayanı olduğunu diğer bir ifadeyle üçüncü bir durumun var olamayacağını söyler. Aristoteles'ten sonra insanlık klasik mantık ilkelerini kabullenmiş bu mantığı kullanarak matematik, klasik fizik gibi bilim dalları üzerinde çalışmışlardır (Baykal ve Beyan 2004).

İki değerli klasik mantık, 19. yüzyılın başlarından itibaren filozof ve teorik matematikçilerin ürettikleri paradoksları açıklamakta yetersiz kalmıştır. Çünkü Aristo mantığı gerçek dünyayı bütünüyle tasvir etmekten uzaktır. Doğadaki her oluşumun bir meydana gelme derecesi vardır. Bu anlamda herhangi bir önermenin yalnızca doğruya da yalnızca yanlış olması gerekliliği, ikili mantığın gelişerek çok değerli mantığa dönüşmesine sebep olmuştur. Çok değerli mantığın en ilkel hali olan üç değerli mantık, önermelerin {0,1} değerlerinin yanında, {0.5} değerini de almasını sağlamıştır ve böylece değer kümesi {0, 0.5, 1} olarak geliştirilmiştir. Değer kümesindeki {0} ögesi önermenin kesinlikle yanlış olduğunu, {0.5} ögesi belirsiz olduğunu ve {1} ögesi de kesinlikle doğru olduğunu ifade etmektedir.

Polonyalı mantık bilimcisi Lukasiewicz, üç değerli mantıktan yola çıkarak çok değerli mantığı bütünüyle ele almıştır ve sonsuz değerli mantığı geliştirmiştir (Yücel 2005).

1956 yılında Amerika Birleşik Devletlerinde düzenlenen bir konferansta ilk kez “Bulanık Küme” yaklaşımından bahsedilmiştir. Ancak bu konudaki ilk ciddi adım 1965 yılında Lotfi A. Zadeh tarafından atılmıştır. Zadeh tarafından yayınlanan makalede bulanık mantık ve bulanık küme kuramı ortaya atılmıştır. Zadeh bu çalışmasında insan düşüncesinin kesin olmadığını belirtmiştir. Zadeh’e göre insan düşüncesinin büyük çoğunluğunu bulanıktır. İnsan mantığı açık, kapalı, sıcak soğuk, 0 ve 1 gibi kesin ifadelerin yanı sıra, az açık, az kapalı, serin, ılık gibi ara değerleri de göz önüne almaktadır. Bulanık mantık klasik mantığın aksine, iki seviyeli değil çok seviyeli işlemleri kullanmaktadır.

Bulanık mantığın genel özellikleri Zadeh tarafından aşağıdaki gibi ifade edilmiştir:

- Bulanık mantıkta, kesin değerlere dayanan düşünme yerine, yaklaşık düşünme kullanılır.
- Bulanık mantıkta her şey [0,1] aralığında belirli bir derece ile gösterilir.
- Bulanık mantıkta, bilgi büyük, küçük, çok, az gibi dilsel ifadeler şeklindedir.
- Bulanık çıkarım işlemi dilsel ifadeler arasında tanımlanan kurallar ile yapılır.
- Her mantıksal sistem bulanık olarak ifade edilebilir.

- Bulanık mantık matematiksel modeli çok zor elde edilen sistemler için çok uygundur (Elmas 2003; Şanlı 2005).

2.1.1. Bulanık mantık uygulamaları

İnsan düşünce yapısı belirsiz bilgiyle yaklaşık karar verebilen en üstün mantığa sahiptir. Makineler ise sadece ikili sayı (bit) değerleriyle girilen bilgileri işleyerek çalışırlar, dolayısıyla bu özellikten yoksundurlar. Günümüzde bulanık mantık uygulamalarıyla yazılan programlar sayesinde, akıllı ve insana ihtiyaç duymadan çalışabilen makineler üretilebilmektedir. Söz konusu yazılımlar uzmanlarca oluşturulan kural tabanlarının eğer - ise şeklinde makinelere işlenmesi ile oluşmaktadırlar. Böylece makinelerin zeka seviyesi yükseltilmiş olunur. Bu sayede kirliliğine göre çamaşır yıkayan çamaşır makineleri, süpürdüğü yüzeyin tozluluğuna ve zeminin sertliğine göre emme gücünü otomatik ayarlayan süpürgeler ve daha pek çok akıllı elektronikler üretilebilmektedir. Bulanık mantık yaklaşımı makinelere insanların insan düşünce yapısını işleyebilme ve insanların deneyimlerinden ve önsezilerinden yararlanarak çalışabilme yetisi verir. Bu yeteneği kazandırırken sayısal ifadeler yerine sembolik ifadeler kullanılmaktadır. İşte bu sembolik ifadelerin makinelere aktarılması matematiksel bir temele dayanmaktadır. Bu matematiksel temel bulanık mantık kümeler kuramı ve buna dayanan bulanık mantıktır (Şen 2001; Ergün 2011).

Bulanık mantık aşağıdaki bilim dalları ve alanlarında uygulanmıştır:

- Yapay zeka
- Uzman sistemler
- Kontrol teorisi
- Kalite kontrol
- Çok amaçlı karar verme
- Ürün planlaması, seçimi
- Optimum sistem planlanması
- Taşıma, ulaşım
- Network

- Oyunlar kuramı
- Çevre yönetimi
- Bankacılık finansı
- Ziraat

2.2. Kesinlik, Bulanıklık ve Belirsizlik

Kesinlik önermelerin iki değerli olduğu durumu kapsar. Bir önerme ya doğrudur ya da yanlıştır. Bu iki sonuç arasında başka bir değer alamaz. Klasik küme teorisi keskindir. Herhangi bir eleman bir kümeye ya aittir, ya da değildir.

Bir amacı aktaran, bir kavramı anlatan veya bir sistemi tanıtan ifadelerdeki belirsizliğe veya kesin olmama haline ise “bulanıklık” denilmektedir. İnsanların zihinsel düzeydeki algılama farklılıkları, onların sübjektif davranışları, ifade ve amaçlarındaki belirsizlikler, bulanıklık kavramı ile açıklanabilir (Özkan 2003; Gülcan 2012). Bulanıklığı diğer bir ifadeyle incelenen bir konunun araştıran kişi tarafından tam, kesinlikle bilinmemesi durumunda sahip olunan eksik ve belirsiz bilgilerin tümü olarak tanımlayabiliriz.

Belirsizliği ise rassallık ve bulanıklık olmak üzere iki kategoride ele alabiliriz. Rassal karakterde olan belirsizlikler istatistiki yöntemlerle ele alınabilmektedirler. Rassallık bir olayın meydana gelmesindeki belirsizliğin sayısal ölçütü olarak tanımlanabilir. Rassallığın en önemli özelliği, sonuçların ortaya çıkmasında tamamen şans faktörlerinin etkin olmasıdır (Baykal ve Beyan 2004).

Ancak var olan tüm belirsizliklerin hepsi rasgele karakterde değildir. Rassal karakterde olmayan olaylar için, özellikle sözel belirsizlikler söz konusu olduğunda bulanık mantık teorisinden yararlanılmaktadır. (Ross ve ark. 2002; Gülcan 2012).

İstatistik ve Bulanık Mantık teorileri arasındaki fark iki teorinin belirsizlik kavramına bakış açılarından kaynaklanmaktadır. İstatistik ve Olasılık teorisi belirsizlikleri şansa bağlı belirsizlikler olarak kabul etmektedir. Bulanık Mantık ise belirsizliği belirsiz bir bilginin gerekliliğine inanma derecesi olarak kabul etmektedir.

Belirsizlik içeren bir sürecin rassal olaylardan ve rassal olmayan olaylardan oluştuğunu düşünürsek, istatistik rassal belirsizlikleri konu alırken, bulanık mantık ise rassal olmayan belirsizlikleri konu alarak olayların sadece oluşum durumları ile ilgilenmektedir. Sonuç olarak; rassallık kavramı olayların oluşundaki kesin olmamayı ifade etmektedir. Bulanık mantık ise, olayların olup olmadığını değil hangi dereceye kadar olduğunu ölçmektedir (Baykal ve Beyan 2004). Bulanık mantık ve istatistik bilimi arasındaki ilişkiyi daha iyi anlayabilmek için olabilirlik teorisi hakkında bilgi sahibi olmak gerekmektedir. Bu yüzden ilerleyen bölümlerde olabilirlik teorisine kısaca değinilecektir.

2.3. Olabilirlik Teorisi

Bu bölümde olabilirlik teorisinden kısaca bahsedilecek ve olabilirlik ölçümü, olabilirlik dağılımı, ortak ve marjinal olabilirlik dağılımlarıyla, koşullu olabilirlik dağılımlarının tanımları verilecektir. Olabilirlik teorisini daha iyi anlayabilmek için öncelikle bulanık ölçümlerin tanımlanması gerekmektedir.

Verilmiş bir X evrensel kümesi ve X evrensel kümesinin alt kümelerinin ailesi $\mathbb{C}$ verilmiş olsun. $\langle X, \mathbb{C} \rangle$ ‘de tanımlı bir fonksiyon olan bulanık ölçüm (2.1) ‘de gösterildiği gibidir.

$$g : \mathbb{C} \rightarrow [0,1] \tag{2.1}$$

Bulanık ölçüm aşağıda gösterilen gereklilikleri sağlamaktadır. (Klir ve Yuan 1995).

$$g(\varnothing) = 0, g(X) = 1 \tag{2.2}$$

$$\forall A, B \in \mathbb{C} \text{ için eğer } A \subseteq B \text{ ise } g(A) \leq g(B) \tag{2.3}$$

$\mathbb{C}$ ‘deki herhangi bir artan dizi $A_1 \subset A_2 \subset \ldots$ için eğer $\bigcup_{i=1}^{\infty} A_i \in \mathbb{C}$ ise

$$\lim_{i \to \infty} g(A_i) = g\left(\bigcup_{i=1}^{\infty} A_i\right) \tag{2.4}$$

$\mathbb{C}$’deki herhangi bir azalan dizi $A_1 \supset A_2 \supset \ldots$ için eğer $\bigcap_{i=1}^{\infty} A_i \in \mathbb{C}$ ise

$$\lim_{i\to\infty} g(A_i) = g\left(\bigcap_{i=1}^{\infty} A_i\right) \tag{2.5}$$

Bulanık ölçümü tanımladıktan sonra şimdi kanıt teorisinin dayalı olduğu iki eş ölçüm olan inanç ve plausibility ölçümlerinin tanımlanması gerekmektedir.

Verilen sonlu bir X evrensel kümesinde "inanç ölçümü"

$$Bel : P(X) \to [0,1] \tag{2.6}$$

şeklinde tanımlanan bir fonksiyondur. Burada $Bel(\varnothing)=0$ ve $Bel(X)=1$'dir. İnanç ölçümleri (2.7)'den dolayı "süper toplamsal" olarak adlandırılırlar.

$$\begin{aligned} Bel(A_1 \cup A_2 \cup \ldots \cup A_n) \geq & \sum_j Bel(A_j) - \sum_{j<k} Bel(A_j \cap A_k) \\ & + \ldots + (-1)^{n+1} Bel(A_1 \cap A_2 \cap \ldots \cap A_n) \end{aligned} \tag{2.7}$$

Verilen sonlu bir X evrensel kümesinde "plausibility ölçümü"

$$Pl : P(X) \to [0,1] \tag{2.8}$$

şeklinde tanımlanan bir fonksiyondur. Burada $Pl(\varnothing)=0$ ve $Pl(X)=1$'dir.

Gereklilik ölçümleri ise (2.9)'deki eşitsizlikten dolayı "alt toplamsal" olarak adlandırılırlar.

$$\begin{aligned} Pl(A_1 \cap A_2 \cap \ldots \cap A_n) \leq & \sum_j Pl(A_j) - \sum_{j<k} Pl(A_j \cup A_k) \\ & + \ldots + (-1)^{n+1} Pl(A_1 \cup A_2 \cup \ldots \cup A_n) \end{aligned} \tag{2.9}$$

İnanç ölçümü ve plausibility ölçümleri birbirlerine basit bir eşitlik yardımıyla dönüştürülebilir.

Her bir inanç ölçümü, bir plausibility ölçümü ile aşağıdaki gibi gösterilebilir.

$$Pl(A) = 1 - Bel(\bar{A}) \quad \forall A \in P(X) \tag{2.10}$$

Benzer şekilde, plausibility ölçümü bir inanç ölçümü ile (2.11)'deki gibi ifade edilir.

$$Bel(A) = 1 - Pl(\bar{A}) \quad \forall A \in P(X) \tag{2.11}$$

İnanç ve plausibility ölçümleri "basit olasılık ataması" adı verilen bir fonksiyonla kolayca karakterize edilebilirler. Basit olasılık ataması (2.12)'da gösterildiği gibidir.

$$m : \mathrm{P}(X) \to [0,1] \tag{2.12}$$

Basit olasılık ataması için aşağıdaki özellikler geçerlidir.

$$\sum_{A \in P(X)} m(A) = 1 \,, \; m(\varnothing) = 0 \tag{2.13}$$

İnanç ve plausibility ölçümleri, basit olasılık ataması yardımıyla aşağıdaki eşitliklerdeki gibi belirlenebilirler.

$$Bel(A) = \sum_{B \mid B \subseteq A} m(B) \tag{2.14}$$

$$Pl(A) = \sum_{B \mid A \cap B \neq \varnothing} m(B) \tag{2.15}$$

$A \in P(X)$ için $m(A) > 0$ her küme m'nin "merkez elemanı" (focal element) olarak adlandırılır. F merkez elemanların kümesi ve m ilgili basit atamayı gösterdiğinde $\langle F, m \rangle$ çiftine "kanıt gövdesi" adı verilir (Klir ve Yuan 1995).

Olabilirlik teorisi, bulanık küme teorisinin genişletilmesiyle geliştirilmiş, belirsizliklerin modellenmesinde kullanılan matematiksel bir araçtır ve olabilirlik dağılımlarına dayanır. Olasılık teorisine bir alternatif veya olasılık teorisinin tamamlayıcısı olarak düşünülebilir (Uçal Sarı 2012). Olabilirlik teorisi inanç ve plausibility ölçümlerinin eşi olan iki ölçüme dayalıdır. Bu ölçümler "gereklilik (necessitiy)" ölçümü ve "olabilirlik (possibility)" ölçümüdür. Gereklilik ölçümü Nec ve olabilirlik ölçümü Pos ile gösterilir. Gereklilik ve olabilirlik ölçümleri aşağıdaki gibi tanımlanmıştır.

Nec $\langle X, \mathbb{C} \rangle$'de bir bulanık ölçümü göstersin. Nec sadece ve sadece $\mathbb{C}$'deki herhangi bir $\{A_k | k \in K\}$ öyle ki $\bigcap_{k \in K} A_k \in \mathbb{C}$ ailesi için (2.16) sağlanıyorsa gereklilik ölçütü olarak adlandırılır. Burada K keyfi bir gösterge kümesidir.

$$Nec\left(\bigcap_{k \in K} A_k\right) = \inf_{k \in K} Nec\left(A_k\right) \tag{2.16}$$

Benzer şekilde Pos $\langle X, \mathbb{C} \rangle$'de bir bulanık ölçümü göstersin. K keyfi bir gösterge kümesi olmak üzere Pos sadece ve sadece $\mathbb{C}$'deki herhangi bir $\{A_k | k \in K\}$ öyle ki $\bigcup_{k \in K} A_k \in \mathbb{C}$ ailesi için sağlanıyorsa gereklilik ölçütü olarak adlandırılır.

$$Pos\left(\bigcup_{k \in K} A_k\right) = \sup_{k \in K} Pos\left(A_k\right) \tag{2.17}$$

Gereklilik ölçümü özel inanç ölçümü ve olabilirlik ölçümü özel olabilirlik ölçümleri olduklarından, aşağıdaki özellikleri sağlarlar.

$$Nec(A) + Nec(\bar{A}) \leq 1 \tag{2.18}$$

$$Pos(A) + Pos(\bar{A}) \geq 1 \tag{2.19}$$

$$Nec(A) = 1 - Pos(\bar{A}) \tag{2.20}$$

Olabilirlik teorisinde kullanılan ve önemli olan denklemlerin elde edilmesi için aşağıdaki teoreme ihtiyaç vardır.

Teorem 2.1. Verilen kanıt gövdesi $\langle F, m \rangle$ iç içe geçmiş olsun. O halde ilgili inanç ve plausibility ölçümleri $\forall A, B \in P(X)$ için aşağıdaki özellikleri sağlar.

$$Bel(A \cap B) = \min[Bel(A), Bel(B)] \tag{2.21}$$

$$Pl(A \cup B) = \max[Bel(A), Bel(B)] \tag{2.22}$$

Teorem 2.1 yardımıyla olabilirlik teorisinin basit denklemleri elde edilmiş olur. Bu denklemler aşağıdaki gibidir.

$$Nec(A \cap B) = \min[Nec(A), Nec(B)] \tag{2.23}$$

$$Pos(A \cup B) = \max[Pos(A), Pos(B)] \tag{2.24}$$

Ayrıca (2.23) ve (2.24)'den aşağıdaki özellikler elde edilmiş olur.

$$\min[Nec(A), Nec(\bar{A})] = 0 \tag{2.25}$$

$$\max[Pos(A), Pos(\bar{A})] = 1 \tag{2.26}$$

Olabilirlik ve gereklilik ölçütleri aşağıdaki teoremde ifade edildiği gibi birbirlerini kısıtlamaktadırlar.

Teorem 2.2. Bütün $A \in P(X)$ için $P(X)$'deki herhangi bir gereklilik ölçütü Nec ve ilgili olabilirlik ölçümü Pos aşağıdaki özellikleri sağlamaktadır.

$$Nec(A) > 0 \Rightarrow Pos(A) = 1 \tag{2.27}$$

$$Pos(A) < 1 \Rightarrow Nec(A) = 0 \tag{2.28}$$

$P(X)$'de verilen bir olabilirlik ölçümü Pos için $r: X \to [0,1]$ şeklinde verilen ve aşağıdaki özelliği sağlayan fonksiyona, ilgili Pos ölçümü için "olabilirlik dağılımı fonksiyonu" adı verilir.

$$r(x) = Pos(\{x\}), \quad \forall x \in X \tag{2.29}$$

Olabilirlik teorisinin bir önemli özelliği, her olabilirlik ölçümünün sadece ilgili olabilirlik dağılımı fonksiyonuyla gösterilmesidir. Sonlu bir evrensel küme için bu özellik aşağıdaki teorem yardımıyla gösterilmiştir.

Teorem 2.3. Sonlu kuvvet kümesi $P(X)$'de tanımlı her olabilirlik ölçümü Pos, (2.30) ile gösterilen formül yardımıyla ilgili olabilirlik dağılımı fonksiyonuyla gösterilir.

$$\begin{gathered} r: X \to [0,1] \\ Pos(A) = \max_{x \in A} r(x), \quad \forall A \in P(X) \end{gathered} \tag{2.30}$$

Olabilirlik dağılımları tanımladıktan sonra ortak olabilirlik dağılımının ve marjinal olabilirlik dağılımlarının tanımlamalarının yapılması gerekmektedir.

Kartezyen çarpımı $X \times Y$'de tanımlı r "ortak olabilirlik dağılımını" göstersin. r'nin izdüşümleri olan r_X ve r_Y 'ye "marjinal olabilirlik dağılımı" adı verilir ve eşitliklerdeki gibi tanımlanır.

$$r_X(x) = \max_{y \in Y} r(x,y), \ \forall x \in X \tag{2.31}$$

$$r_Y(y) = \max_{x \in X} r(x,y), \ \forall y \in Y \tag{2.32}$$

X ve Y'de tanımlı ve r_X, r_Y olabilirlik dağılımı fonksiyonlarıyla gösterilen iç içe kanıt gövdeleri, bütün $x \in X$ ve $y \in Y$ için sadece ve sadece (2.33)'deki eşitliği sağlıyorsa "ilişkisiz" olarak adlandırılırlar.

$$r(x,y) = \min[r_X(x), r_Y(y)] \tag{2.33}$$

X ve Y'de tanımlı marjinal olasılık dağılımları r_X, r_Y olsun. $X \times Y$'de tanımlı ortak olabilirlik dağılımı fonksiyonu olan r (2.33)'de gösterildiği gibi ilişkisiz r_X, r_Y ile tanımlanmış olsun. Pos_X, Pos_Y ve Pos sırasıyla ilgili r_X, r_Y ve r için olabilirlik ölçümlerini göstermek üzere Pos, Pos_X ve Pos_Y yardımıyla aşağıdaki gibi hesaplanabilir.

$$Pos(A \times B) = \min\left[Pos_X(A), Pos_Y(B)\right], \ \forall A \in P(X), \forall B \in P(Y) \quad (2.34)$$

Buradaki $Pos(A \times B), Pos_X(A)$ ve $Pos_Y(B)$ aşağıda gösterildiği gibidir.

$$\begin{aligned} Pos_X(A) &= \max_{x \in A} r_X(x), \\ Pos_Y(B) &= \max_{y \in B} r_Y(y), \\ Pos(A \times B) &= \max_{x \in A, y \in B} r(x, y) \end{aligned} \quad (2.35)$$

Bütün $x \in X$ ve $y \in Y$ için $X \times Y$'deki koşullu olabilirlikler $r_{X|Y}(x|y)$, $r_{Y|X}(y|x)$ ile gösterilsin. İki marjinal posibilistik kanıt gövdesi, eğer sadece ve sadece aşağıdaki eşitlikler sağlanıyorsa "bağımsız" olarak adlandırılır (Klir ve Yuan 1995).

$$r_{X|Y}(x|y) = r_X(x) \quad (2.36)$$

$$r_{Y|X}(y|x) = r_Y(y) \quad (2.37)$$

Yukarıda olabilirlik teorisi hakkında gerekli bilgiler verilmiştir. İlerleyen bölümde olasılık teorisi ile olabilirlik teorisi arasındaki farklar kısaca anlatılacaktır.

2.4. Olasılık Teorisi ile Olabilirlik Teorisinin Karşılaştırılması

Olabilirlik teorisi, inanç ve plausibility ölçümlerinin özel versiyonları olan iki eş ölçüme dayalı ilken, olasılık teorisi kanıt teorisinin inanç ve plausibility ölçümlerinin eşit olduğu bir alt bölgesiyle kesişmektedir. Bu durum ilgili kanıt gövdesinin yapısındaki temel bir farktan kaynaklanmaktadır. Olasılıksal kanıt gövdeleri tek kümelerden (singleton) oluşurken, olabilirliksel kanıt gövdeleri iç içe kümeler ailesidir.

İki teorinin matematiksel özelliklerindeki bu farklılıklar, her bir teoriyi belirsizliğin belirli bir tipini modellemede daha uygun yapmaktadır. Bilindiği gibi, olasılık teorisi, sınıf frekanslarının bilindiği ya da kanıtın bağımsız rassal deneylerin uzun denemelerinden elde edilen çıktıya dayandığı durumlarda belirsizliği modelleme de ideal bir araçtır. Diğer yandan, olabilirlik teorisi bulanık önermelerle ifade edilmiş eksik bilgiyi şekillendirmede uygundur.

İki teori arasındaki bir temel fark ise, bu teorilerin toplam bilgisizliği (total ignorance) ifade ediş şekilleridir.

Olabilirlik teorisinde toplam bilgisizlik kanıt teorisindeki gibi ifade edilir. Yani;

$$m(X)=1 \text{ ve } m(A)=0, \forall A \neq X \qquad (2.38)$$

Benzer şekilde;

$$r(X)=1, \forall x \in X \qquad (2.39)$$

biçiminde gösterilir.

Olasılık teorisinde ise, evrensel kümedeki uniform olasılık dağılımıyla ifade edilir. $m(\{x\})=1/|X|$. Olasılık ve olabilirlik teorileri arasındaki matematiksel farklar Çizelge 2.1. ile gösterilmiştir (Klir ve Yuan 1995).

Olasılık ve olabilirlik arasındaki farka başka bir penceren bakılırsa, olasılık ve olabilirlik arasındaki en önemli farkın bir olayın olasılığının karşıt olayın olasılığını kesin olarak belirtebilmesi olduğu görülmektedir. Belirli bir olayın olabilirliği ve o olayın karşıt olayının olabilirliğine zayıf olarak bağlıdır. Benzer şekilde bir olayın gerekliliği ile karşıt olayının gerekliliğine zayıf olarak bağlıdır. Diğer bir ifadeyle bir olayının belirsizliği tanımlanırken hem gereklilik hem de olabilirlik ölçütlerinin ikisinin birden değerlerine ihtiyaç duyulur. Olasılık ölçütleri kesin fakat farklılaştırılmış bilgi bileşenlerine uygulanırken, olabilirlik ölçütleri kesin olmayan fakat tutarlı bilgi bileşenlerini de yansıtmaktadır. Olabilirlik fonksiyonları sözü edilen özelliğinden dolayı öznel belirsizlikleri ifade etmede daha yararlıdır. (Dubois ve Prade 1988; Uçal Sarı 2012)

Çizelge 2.1. Olasılık ve olabilirlik teorilerinin matematiksel olarak karşılaştırılması (Klir ve Yuan 1995)

Olasılık Teorisi	**Olabilirlik Teorisi**
Tek tip ölçüme dayanır: Olasılık ölçüsü	İki tip ölçüme dayanır Gereklilik ve Olabilirlik ölçüsü
Kanıt gövdesi tek kümeden oluşmaktadır.	Kanıt gövdesi iç içe geçmiş alt kümeler ailesinden oluşmaktadır.
Olasılık dağılımı fonksiyonu ile tek bir gösterimi vardır. $p: X \to [0,1]$ iken $P(A) = \sum_{x \in A} p(x)$	Olabilirlik dağılım fonksiyonu ile bir tek gösterimi vardır. $\mathrm{r}: X \to [0,1]$ iken $Pos(A) = \max_{x \in A} r(x)$
Normalleştirme: $\sum_{x \in X} p(x) = 1$	Normalleştirme: $\max_{x \in A} r(x) = 1$
Toplanabilirlik: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$	Min/max kuralları: $Nec(A \cap B) = \min[Nec(A), Nec(B)]$ $Pos(A \cup B) = \max[Pos(A), Pos(B)]$
Uygulanmamaktadır.	$Nec(A) = 1 - Pos(\bar{A})$ $Pos(A) < 1 \Rightarrow Nec(A) = 0$ $Nec(A) > 0 \Rightarrow Pos(A) = 1$
$P(A) + P(\bar{A}) = 1$	$Pos(A) + Pos(\bar{A}) \geq 1$, $Nec(A) + Nec(\bar{A}) \leq 1$ $\max\left[Pos(A), Pos(\bar{A})\right] = 1$ $\min\left[Nec(A), Nec(\bar{A})\right] = 0$
Toplam bilgisizlik: $p(x) = 1/\lvert X \rvert, \quad \forall x \in X$	Toplam bilgisizlik: $r(X) = 1, \forall x \in X$
Koşullu olasılıklar: $p_{X\|Y}(x\|y) = p(x,y) / p_Y(y)$ $p_{Y\|X}(y\|x) = p(x,y) / p_X(x)$	Koşullu olasılıklar: $r_{X\|Y}(x\|y) = \begin{cases} r_X(x), & r_X(x) < r_Y(y) \\ [r_Y(y), 1], & r_X(x) \geq r_Y(y) \end{cases}$ $r_{Y\|X}(y\|x) = \begin{cases} r_Y(y), & r_Y(y) < r_X(x) \\ [r_X(x), 1], & r_Y(y) \geq r_X(x) \end{cases}$

Çizelge 2.2. (Devam) Olasılık ve olabilirlik teorilerinin matematiksel olarak karşılaştırılması (Klir ve Yuan 1995)

Olasılıksal ilişkisizlik: $p(x,y)=p_X(x).p_Y(y) \qquad (a)$ Olasılıksal bağımsızlık: $p_{X\mid Y}(x\mid y)=p_X(x)$ $p_{Y\mid X}(y\mid x)=p_Y(y) \qquad (b)$ $(a)\Leftrightarrow(b)$	Olabilirliksel ilişkisizlik: $r(x,y)=\min\left[r_X(x),r_Y(y)\right] \qquad (c)$ Olabilirliksel bağımsızlık: $r_{X\mid Y}(x\mid y)=r_X(x)$ $r_{Y\mid X}(y\mid x)=r_Y(y) \qquad (d)$ $(d)\Rightarrow(c)$ ama $(c)\not\Rightarrow(d)$

Çizelgede olasılık teorisindeki olasılık fonksiyonunun ve normalleştirmenin gösterimi kesikli rassal değişken için verilmiştir.

2.5. Bulanık Kümeler

Küme kavramı klasik kümeler ve bulanık kümeler olarak karşımıza çıkmaktadır. Klasik kümeler üye olma ve üye olmama ilişkisi çerçevesinde geliştirilmişlerdir. Bu yaklaşıma göre istediğimiz özelliğe sahip olan bir birey, bir kümeye ya aittir ya da değildir. Bu tür kümeleri ifade etmekte ise karakteristik fonksiyonlardan yararlanılmaktadır. Karakteristik fonksiyon her bir elemana 1 ve 0 değerlerinden birini üyelik durumuna göre atayarak evrensel küme üzerinde tanımlanan ve bizim ilgilendiğimiz özelliğe sahip elemanların oluşturduğu kümeyi belirlemektedir.

Klasik küme kavramında, A kümesi X evrensel kümesinde klasik bir kümeyi temsil etmek üzere A kümesinin karakteristik fonksiyonu χ_A ile ifade edilmektedir. Klasik bir A kümesini karakteristik fonksiyon yardımıyla aşağıdaki şekilde ifade etmek mümkündür (Klir ve Yuan 1995).

$$\chi_A : \mathrm{X} \rightarrow \{0,1\}$$
$$\forall \mathrm{x}\in\mathrm{X}, \quad \chi_A(\mathrm{x})=\begin{cases}1\ , & \mathrm{x}\in\mathrm{A}\\ 0\ , & \mathrm{x}\notin\mathrm{A}\end{cases} \qquad (2.40)$$

Söz konusu fonksiyonda görüldüğü gibi A kümesine ait elemanlar 1 değerini alırken, ait olmayan elemanlar ise 0 değerini almaktadır.

Görüldüğü gibi klasik küme teorisinde kümelerin sınırları ve küme elemanlarının sıfatları kesin olmaktadır. Fakat gerçek hayatta her kümenin sınırları ve bu kümelere ait her elemanın sıfatı o kadar kesin olmamaktadır. Böylece klasik küme anlayışının gerçek hayatta karşılaşılan bazı durumlarda yetersiz olduğu görülebilmektedir.

Klasik küme teorisine karşın, bulanık küme teorisi ise bize gerçek hayatta belirsizliklerin ölçülmesinde güçlü ve anlamlı araçlar sunmakta ve doğal dildeki belirsiz kavramların anlamlı bir şekilde ifade edilmesini de sağlamaktadır. Bulanık küme teorisinde, bulanık kümeleri içeren bir evrensel küme içerisindeki elemanların üyelik geçişi dereceli olmaktadır. Eğer bir eleman herhangi bir kümeye ait olacaksa, o elemanın o kümeye ait olma derecesi de söz konusu olmaktadır. Bu derecelendirme bulanık kümelerin sınırlarına belirsizlik özelliğini katmaktadır (Klir ve Yuan 1995).

Zadeh (1965) bulanık kümeyi, sürekli dizi halindeki üyelik derecelerine sahip nesnelerden oluşan bir sınıf olarak tanımlamıştır. Bu tip bir küme, her bir nesneye 0 ile 1 arasında bir üyelik derecesi atayan bir üyelik fonksiyonu ile tanımlanır. Burada 0 sayısı, ilgili nesnenin bulanık kümenin üyesi olmadığını; 1 sayısı, ilgili nesnenin bulanık kümenin tam üyesi olduğunu ve bu iki değer arasındaki herhangi bir sayı ilgili nesnenin bulanık kümeye üyelik derecesini veya kısmi üyeliğini göstermektedir (Özkan 2003; Uçal Sarı 2012).

X evrensel kümesinde A bulanık kümesinin üyelik fonksiyonu $\mu(A)$ olarak ifade edilmek üzere, bulanık bir A kümesi

$$\forall x \in X \text{ için } ; \ \mu_{\tilde{A}}(x) : X \rightarrow [0,1] \tag{2.41}$$

olarak ifade edilebilmektedir. Burada $\mu_{\tilde{A}}(x)$ x elemanının A bulanık kümesine üyelik değerini göstermektedir. $\mu_{\tilde{A}}(x)$; A'nın elemanlarının istenilen özelliği hangi ölçüde sağladığının ifadesi olmaktadır.

Bulanık bir küme, bir nesneyi ve bu nesnenin ilgili kümeye üyelik derecesini gösteren sıralı çiftler halinde ifade edilir:

$$\tilde{A} = \left(x, \mu_{\tilde{A}}(x)\right), \ \forall x \in X \tag{2.42}$$

Bulanık bir A kümesi, evrensel kümenin sonlu olması halinde (2.43)'teki biçimde gösterilir:

$$\tilde{A} = \sum_{i}^{n} \frac{\mu_{\tilde{A}}(x_i)}{x_i} = \frac{\mu_{\tilde{A}}(x_1)}{x_1} + \frac{\mu_{\tilde{A}}(x_2)}{x_2} + \ldots + \frac{\mu_{\tilde{A}}(x_n)}{x_n} \tag{2.43}$$

Evrensel kümenin sonlu olması halinde ise bulanık A kümesi (2.44) 'te gösterildiği gibi ifade edilir:

$$\tilde{A} = \int \frac{\mu_{\tilde{A}}(x_i)}{x_i} \tag{2.44}$$

Yukarıdaki gösterimlerde kullanılan $\sum, \int$ ve + işaretleri alışılagelmiş cebirsel anlamlarını ifade etmemektedirler. Toplam ve integral işaretleri, bulanık çiftlerin sırasıyla kesikli ve sürekli evrenlerde bir araya getirilmesini ifade eder. + simgesi ise bulanık sayı çiftlerinin birleşimini gösteren bir simgedir (Özkan 2003; Uçal Sarı 2012).

2.5.1. Bulanık küme kısımları

Bulanık sayılar ve bu sayılar üzerinde tanımlı işlemlerin daha iyi anlaşılabilmesi için öncelikle bulanık kümenin kısımlarından bahsedilmesi gerekmektedir. Bulanık kümenin kısımları destek (dayanak), α (alfa) kesim, yükseklik, normallik ve dış bükey kavramlarından oluşmaktadır. Şimdi bu kısımlardan sırasıyla kısaca bahsedilecektir.

A kümesi bir bulanık küme olsun. A'nın tüm elemanlarının üyelik dereceleri 0'dan büyük küme A'nın **destek (support) kümesi** olarak ifade edilir.

$$Destek(A) = \left\{ x \in X \,\middle|\, \mu_{\tilde{A}}(x) > 0 \right\} \tag{2.45}$$

Üyelik derecelerinin maksimum değeri bulanık kümenin **yüksekliği** olarak ifade edilir.

$$Yük(A) = Max\, \mu_{\tilde{A}}(x) \qquad x \in X \tag{2.46}$$

A bulanık kümesinin yüksekliği 1 ise A bulanık kümesine **normaldir** denir. Matematiksel olarak,

$$Max\, \mu_{\tilde{A}}(x) = 1 \qquad x \in X \tag{2.47}$$

ise A bulanık kümesi normaldir.

α **kesim kümesi** A_α, üyelik dereceleri α'dan az olmayan üyelerden oluşan kümedir. Burada α, $[0,1]$ aralığında keyfi bir değerdir.

$$A_\alpha = \left\{ x \in X \,\middle|\, \mu_{\tilde{A}}(x) \geq \alpha \right\} \tag{2.48}$$

Seviye kümesi ise α'lar aracılığıyla elde edilir. Dolayısıyla seviye kümesi (2.49)'te gösterildiği gibi ifade edilir.

$$\Lambda_A = \left\{ \alpha \,\middle|\, \mu_{\tilde{A}}(x) = \alpha, \alpha \geq 0, x \in X \right\} \tag{2.49}$$

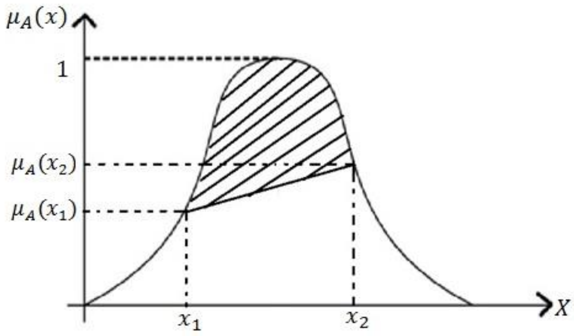
X evrensel kümesinin n boyutlu Öklid vektör uzayında $(\Re^n)$ tanımlandığı varsayılsın. Eğer bütün α kesim kümeleri konveks ise, bu α kesim kümeleriyle **bulanık küme konvekstir**. Başka bir deyişle;

$$t = \lambda r + (1-\lambda)s \qquad r, s \in \Re^n, \qquad \lambda \in [0,1] \tag{2.50}$$

olduğunda bir

$$\mu_{\tilde{A}}(t) \geq \min\left[\mu_{\tilde{A}}(r), \mu_{\tilde{A}}(s)\right] \tag{2.51}$$

ilişkisi varsa, A bulanık kümesi konvekstir (Lee 2005; Özdemir 2013). Yukarıda matematiksel olarak ifade edilen konveks bulanık sayının grafiksel gösterimi Şekil 2.1.'de verilmiştir.



Şekil 2.1. Konveks ve normal bulanık küme (Özdemir 2013)

2.5.2. Genişleme prensibi

Genişleme prensibi, (extension principle) Zadeh (1978) tarafından ortaya konulmuştur. Bu prensip, bulanık küme teorisinde kullanılan en önemli araçların başında gelmektedir. Genişleme prensibi kesin matematiksel kavramların bulanık küme çerçevesinde genelleştirilmesini sağlamaktadır (Dubois ve Prade 1980).

Verilen bir f fonksiyonu $X_1 \times X_2 \times \ldots \times X_n$ 'deki elemanları Y uzayına götürsün. Diğer bir deyişle; $y \in Y$ ve $\forall i$ için $x_i \in X$ iken $y = f(x_1, x_2, \ldots, x_n)$ olsun. Y içerisindeki bulanık A kümesi (2.52)'de tanımlanmıştır.

$$\tilde{A} = \left\{ \left(y, \mu_A(y)\right) \middle| y = f(x_1, x_2, \ldots, x_n),\ (x_1, x_2, \ldots, x_n) \in X \right\} \tag{2.52}$$

Yukarıda tanımlanan bulanık kümenin üyelik fonksiyonu (2.53) ile tanımlanmıştır (Uçal Sarı 2012).

$$\mu_{\tilde{A}}(y)=\begin{cases} y=\sup\limits_{f(x_1,x_2,\dots,x_n)} \min\left(\mu_{\tilde{A}_1}(x_1),\dots,\mu_{\tilde{A}_n}(x_n)\right) & ,\ \text{eğer } f^{-1}\neq\varnothing \\ 0 & ,\text{diğer durumlarda} \end{cases} \tag{2.53}$$

2.6. Bulanık Sayılar ve İşlemler

Bulanık sayılar, bulanık kümelerin özel bir alt kümesidir. Bulanık kümelerde geçerli olan birleşim, kesişim, α-kesim, genişleme prensibi gibi teorik küme işlemleri bulanık sayılara da kolayca uygulanabilir. Bulanık sayıların kullanım alanları arasında bulanık regresyon, bulanık programlama ve bulanık karar verme ön plana çıkmaktadır (Özkan 2003; Uçal Sarı 2012).

Normal ve dışbükey bir bulanık kümenin alfa kesimi kapalı bir küme ise **bulanık sayı** olarak adlandırılmaktadır. Matematiksel olarak ifade edilecek olursa; A bulanık bir küme ve $x \in A$ olmak üzere aşağıda gösterilen koşullar sağlanıyorsa x bir bulanık sayıdır.

$$Max\,\mu_{\tilde{A}}(x)=1 \tag{2.54}$$

$$A_\alpha \in [0,1] \tag{2.55}$$

$$\mu_{\tilde{A}}\left[\lambda x_1+(1-\lambda)x_2\right]\geq \min\left[\mu_{\tilde{A}}(x_1),\mu_{\tilde{A}}(x_2)\right] \tag{2.56}$$

Eşitlik (2.54) bulanık sayının normal olmasını, (2.55) bulanık kümenin alfa kesiminin kapalı olmasını ve (2.56) ise bulanık kümenin dış bükey olmasını ifade etmektedir.

Çeşitli bulanık sayı tipleri olmasına rağmen en sık kullanılan bulanık sayılar; aralık tipi bulanık sayılar, LR tipi bulanık sayılar, üçgen bulanık sayılar, yamuk bulanık sayılar, Gauss bulanık sayılar ve üstel bulanık sayılardır. Bu tezde sırasıyla önce aralık tipi bulanık sayılar, daha sonra LR tipi, üçgen ve yamuk bulanık sayılar tanımlanacak ve bu sayılar üzerinde bulanık aritmetik işlemler gösterilecektir.

2.6.1. Aralık tipi bulanık sayılar

Üyelik değerleri aralık şeklinde gösterilen bulanık sayılara “aralık bulanık sayı” denir. Bu tip sayılar için temel cebirsel işlemler aşağıda gösterildiği gibidir.

Toplama:

$$[a;b](+)[c;d]=[a+c;b+d] \tag{2.57}$$

Çıkarma:

$$[a;b](-)[c;d]=[a-d;b-c] \tag{2.58}$$

Çarpma:

$$[a;b](\bullet)[c;d]=\begin{bmatrix}\min(a*c;a*d;\mathrm{b}*c;b*d); \\ \max(a*c;a*d;\mathrm{b}*c;b*d)\end{bmatrix} \tag{2.59}$$

Bölme:

$$[a;b](/)[c;d]=\begin{bmatrix}\min(a\div c;a\div d;\mathrm{b}\div c;b\div d); \\ \max(a\div c;a\div d;\mathrm{b}\div c;b\div d)\end{bmatrix},\quad c\neq 0, d\neq 0 \tag{2.60}$$

Ters alma:

$$[a;b]^{-1}=[\min(1\div a;1\div b);\max(1\div a;1\div b)],\quad a\neq 0, b\neq 0 \tag{2.61}$$

Bir skaler λ ile çarpma:

$$\lambda(\bullet)[a;b]=[\lambda;\lambda](\bullet)[a;b]=[\lambda*a;\lambda*b],\ \lambda>0 \tag{2.62}$$

2.6.2. L-R tipi bulanık sayılar

L-R tip bulanık sayılar Dubois ve Prade (1978) tarafından ortaya atılmıştır. Bir $\tilde{M}$ bulanık sayısının L-R tip bulanık sayı olarak adlandırılabilmesi için aşağıdaki eşitliğin sağlanması gerekmektedir.

$$\mu_{\tilde{M}}(x) = \begin{cases} L\left((m-x)/\alpha\right), \text{eğer } x \leq m,\ \alpha > 0, \\ R\left((x-m)/\beta\right), \text{eğer } x \geq m,\ \beta > 0. \end{cases} \tag{2.63}$$

Eşitlik (2.63)'deki m, $\tilde{M}$ bulanık sayısının orta değerini, L; sol taraf referansını, R; sağ taraf referansını göstermektedir.

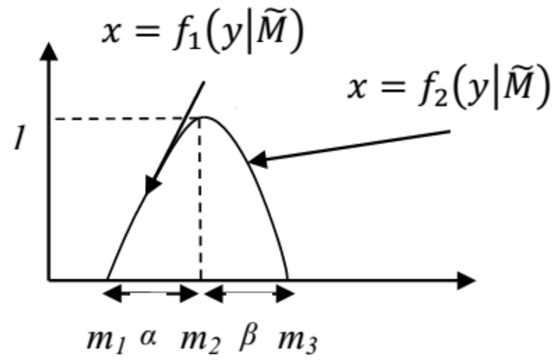
Referans fonksiyonları olarak anılan $L(x)$ ve $R(x)$ fonksiyonları, $\Re^{+} \rightarrow [0,1]$ olmak üzere aşağıdaki koşulları sağlar.

i. $L(x) = L(-x), R(x) = R(-x)$

ii. $L(0) = 1, R(0) = 1$

iii. L ve R $[0,\infty]$ aralığında artmayandır.

α, β ise sırasıyla sol ve sağ yayılım olarak adlandırılmaktadır. Eğer sol ve sağ yayılımlar sıfıra eşit olursa $\tilde{M}$ bulanık sayısı kesin bir sayı haline dönüşür. M bulanık sayısı sembolik olarak $\tilde{M} = (m;\alpha;\beta)_{LR}$ biçiminde gösterilmektedir. L - R tipi bulanık $\tilde{M}$ sayısının grafiksel gösterimi Şekil 2.2'deki gibidir (Dubois ve Prade 1978.).



Şekil 2.2. L-R tipi bulanık sayının üyelik fonksiyonu (Uçal Sarı 2012)

L-R tipi bulanık sayılara ait temek cebirsel işlemler (2.64)-(2.74)‘de gösterildiği gibidir (Dubois ve Prade 1980).

Toplama:

$$\tilde{M}\left(m;\alpha;\beta\right)_{LR}(+)\,\tilde{N}\left(\mathrm{n};\gamma;\delta\right)_{LR} = \left(m+n;\alpha+\gamma;\beta+\delta\right)_{LR} \qquad (2.64)$$

Çıkarma:

$$\tilde{M}\left(m;\alpha;\beta\right)_{LR}(-)\,\tilde{N}\left(\mathrm{n};\gamma;\delta\right)_{LR} = \left(m-n;\alpha+\delta;\beta+\gamma\right)_{LR} \qquad (2.65)$$

Çarpma:

$\tilde{M}>0,\ \tilde{N}>0$ iken

$$\tilde{M}\left(m;\alpha;\beta\right)_{LR}(\bullet)\,\tilde{N}\left(\mathrm{n};\gamma;\delta\right)_{LR} = \left(mn;m\gamma+n\alpha;m\delta+n\beta\right)_{LR} \qquad (2.66)$$

$\tilde{M}<0,\ \tilde{N}>0$ iken

$$\tilde{M}\left(m;\alpha;\beta\right)_{LR}(\bullet)\,\tilde{N}\left(n;\gamma;\delta\right)_{LR} = \left(mn;n\alpha-m\delta;n\beta+m\gamma\right)_{LR} \qquad (2.67)$$

$\tilde{M} < 0,\ \tilde{N} < 0$ iken

$$\tilde{M}\left(m;\alpha;\beta\right)_{LR}(\bullet)\tilde{N}\left(n;\gamma;\delta\right)_{LR} = \left(mn;-n\beta-m\delta;-n\alpha-m\gamma\right)_{LR} \tag{2.68}$$

Bölme:

$n \neq 0$ olmak üzere; $\tilde{M} > 0,\ \tilde{N} > 0$ iken

$$\tilde{M}\left(m;\alpha;\beta\right)_{LR}(/)\tilde{N}\left(n;\gamma;\delta\right)_{LR} = \left(\frac{m}{n};\frac{\delta m+\alpha n}{n^2};\frac{\gamma m+\beta n}{n^2}\right)_{LR} \tag{2.69}$$

$\tilde{M} < 0,\ \tilde{N} > 0$ iken

$$\tilde{M}\left(m;\alpha;\beta\right)_{LR}(/)\tilde{N}\left(n;\gamma;\delta\right)_{LR} = \left(\frac{m}{n};\frac{\alpha n-\gamma m}{n^2};\frac{\beta n-\delta m}{n^2}\right)_{LR} \tag{2.70}$$

$\tilde{M} < 0,\ \tilde{N} < 0$ iken

$$\tilde{M}\left(m;\alpha;\beta\right)_{LR}(/)\tilde{N}\left(n;\gamma;\delta\right)_{LR} = \left(\frac{m}{n};\frac{-\beta n-\delta m}{n^2};\frac{-\alpha n-\gamma m}{n^2}\right)_{LR} \tag{2.71}$$

Ters alma:

$$\left(m;\alpha;\beta\right)^{-1}{}_{LR} = \left(m^{-1};m\beta^{-2};m\alpha^{-2}\right) \tag{2.72}$$

Bir skaler λ ile çarpma:

$$\lambda > 0,\ \lambda \in \mathbb{R} \qquad \lambda(\bullet)\tilde{M}\left(m;\alpha;\beta\right)_{LR} = \left(\lambda m;\lambda\alpha;\lambda\beta\right)_{LR} \tag{2.73}$$

$$\lambda < 0,\ \lambda \in \mathbb{R} \qquad \lambda(\bullet)\tilde{M}\left(m;\alpha;\beta\right)_{LR} = \left(\lambda m;-\lambda\beta;-\lambda\alpha\right)_{LR} \tag{2.74}$$

2.6.3. Üçgen bulanık sayılar

En çok tercih edilen bulanık sayılardan biri olan üçgen bulanık sayı $\tilde{M} = (m_1; \mathrm{m}_2; \mathrm{m}_3)$ şeklinde ifade edilmektedir. Buradaki m_1 bulanık sayının alt sınırı, m_2 bulanık sayının merkezi ve m_3 bulanık sayının üst sınırını göstermek üzere üçgen bulanık sayı eşitlikteki gibi tanımlanır.

$$\mu_{\tilde{M}}(x) = \begin{cases} 0, & x < m_1 \\ \dfrac{x - m_1}{m_2 - m_1}, & m_1 \leq x \leq m_2 \\ \dfrac{m_3 - x}{m_3 - m_2,} & m_2 < x \leq m_3 \\ 0, & x > m_3 \end{cases} \tag{2.75}$$

$\tilde{M}(m_1; m_2; m_3)$ ve $\tilde{N}(n_1; n_2; n_3)$ olmak üzere üçgen bulanık sayılara ait cebirsel işlemler ilerleyen bölümde gösterilmiştir (Chen ve ark. 1992).

Toplama:

$$\tilde{M}(+)\tilde{N} = (m_1 + n_1; m_2 + n_2; m_3 + n_3) \tag{2.76}$$

Çıkarma:

$$\tilde{M}(-)\tilde{N} = (m_1 - n_3; m_2 - n_2; m_3 - n_1) \tag{2.77}$$

Çarpma:

$\tilde{M} > 0,\ \tilde{N} > 0$ iken

$$\tilde{M}(\bullet)\tilde{N} = (m_1 n_1; m_2 n_2; m_3 n_3) \tag{2.78}$$

$\tilde{M} < 0,\ \tilde{N} > 0$ iken

$$\tilde{M}(\bullet)\tilde{N}=(m_1n_3;m_2n_2;m_3n_1) \tag{2.79}$$

$\tilde{M}<0,\ \tilde{N}<0$ iken

$$\tilde{M}(\bullet)\tilde{N}=(m_3n_3;m_2n_2;m_1n_1) \tag{2.80}$$

Bölme:

$\tilde{M}>0,\ \tilde{N}>0$ iken

$$\tilde{M}(/)\tilde{N}=\left(\frac{m_1}{n_3};\frac{m_2}{n_2};\frac{m_3}{n_1}\right) \tag{2.81}$$

$\tilde{M}<0,\ \tilde{N}>0$ iken

$$\tilde{M}(/)\tilde{N}=\left(\frac{m_3}{n_3};\frac{m_2}{n_2};\frac{m_1}{n_1}\right) \tag{2.82}$$

$\tilde{M}<0,\ \tilde{N}>0$ iken

$$\tilde{M}(/)\tilde{N}=\left(\frac{m_3}{n_1};\frac{m_2}{n_2};\frac{m_1}{n_3}\right) \tag{2.83}$$

Ters alma:

$$\tilde{M}^{-1}=\left(\frac{1}{m_3};\frac{1}{m_2};\frac{1}{m_1}\right) \tag{2.84}$$

Bir skaler λ ile çarpma:

$$\lambda>0,\ \lambda\in\mathbb{R} \qquad \lambda(\bullet)\tilde{M}(m_1;m_2;m_3)=(\lambda m_1;\lambda m_2;\lambda m_3) \tag{2.85}$$

$$\lambda<0,\ \lambda\in\mathbb{R} \qquad \lambda(\bullet)\tilde{M}(m_1;m_2;m_3)=(\lambda m_3;\lambda m_2;\lambda m_1) \tag{2.86}$$

2.6.4. Yamuk bulanık sayılar

Yamuk bulanık sayı $\tilde{M} = (m_1; m_2; m_3; m_4)$ olmak üzere, yapısında üyelik derecesi 1'e eşit olan pek çok nokta barındırır. Yamuk bulanık sayı aşağıdaki gibi bir üyelik fonksiyonu yardımı ile ifade edilmektedir.

$$\mu_{\tilde{M}}(x) = \begin{cases} 0, & x < m_1 \\ \dfrac{x - m_1}{m_2 - m_1}, & m_1 \le x \le m_2 \\ 1, & m_2 < x \le m_3 \\ \dfrac{m_4 - x}{m_4 - m_3}, & m_3 < x \le m_4 \\ 0, & x > a_4 \end{cases} \qquad (2.87)$$

$\tilde{M}(m_1; m_2; m_3; m_4)$ ve $\tilde{N}(n_1; n_2; n_3; n_4)$ olmak üzere üçgen bulanık sayılara ait cebirsel işlemler aşağıdaki eşitlikler yardımıyla gösterilmiştir (Chen ve ark. 1992).

Toplama:

$$\tilde{M}(+)\tilde{N} = (m_1 + n_1; m_2 + n_2; m_3 + n_3; m_4 + n_4) \qquad (2.88)$$

Çıkarma:

$$\tilde{M}(-)\tilde{N} = (m_1 - n_4; m_2 - n_3; m_3 - n_2; m_4 - n_1) \qquad (2.89)$$

Çarpma:

$\tilde{M} > 0,\ \tilde{N} > 0$ iken

$$\tilde{M}(\bullet)\tilde{N} = (m_1 n_1; m_2 n_2; m_3 n_3; m_4 n_4) \qquad (2.90)$$

$\tilde{M} < 0,\ \tilde{N} > 0$ iken

$$\tilde{M}(\bullet)\tilde{N} = (n_1 m_4; n_2 m_3; n_3 m_2; n_4 m_1) \tag{2.91}$$

$\tilde{M} < 0,\ \tilde{N} < 0$ iken

$$\tilde{M}(\bullet)\tilde{N} = (m_4 n_4; m_3 n_3; m_2 n_2; m_1 n_1) \tag{2.92}$$

Bölme:

$n \neq 0$ iken $\tilde{M} > 0,\ \tilde{N} > 0$ iken

$$\tilde{M}(/)\tilde{N} = \left(\frac{m_1}{n_4}; \frac{m_2}{n_3}; \frac{m_3}{n_2}; \frac{m_4}{n_1}\right) \tag{2.93}$$

$\tilde{M} < 0,\ \tilde{N} > 0$ iken

$$\tilde{M}(/)\tilde{N} = \left(\frac{m_4}{n_4}; \frac{m_3}{n_3}; \frac{m_2}{n_2}; \frac{m_1}{n_1}\right) \tag{2.94}$$

$\tilde{M} < 0,\ \tilde{N} > 0$ iken

$$\tilde{M}(/)\tilde{N} = \left(\frac{m_4}{n_1}; \frac{m_3}{n_2}; \frac{m_2}{n_3}; \frac{m_1}{n_4}\right) \tag{2.95}$$

Ters alma:

$$\tilde{M}^{-1} = \left(\frac{1}{m_4}; \frac{1}{m_3}; \frac{1}{m_2}; \frac{1}{m_1}\right) \tag{2.96}$$

Bir skaler λ ile çarpma:

$$\lambda > 0,\ \lambda \in \mathbb{R} \qquad \lambda(\bullet)\tilde{M} = (\lambda m_1; \lambda m_2; \lambda m_3; \lambda m_4) \tag{2.97}$$

$$\lambda < 0,\ \lambda \in \mathbb{R} \qquad \lambda(\bullet)\tilde{M} = (\lambda m_4; \lambda m_3; \lambda m_2; \lambda m_1) \tag{2.98}$$

2.7. Durulaştırma ve Durulaştırma Yöntemleri

Bulanık sayılar ile yapılan işlemler sonucunda elde edilen yeni bulanık kümeden bir çıkarım yapılması gereklidir. Başka bir ifadeyle, bulanık sayılarla yapılan işlemlerden sonra elde edilen bulanık çıktı kümesinin kesin bir değere dönüştürülmesi gerekmektedir. Durulaştırma, bulanık bir kümeyi tek değerli kesin bir miktara indirgeme işlemidir (Roychowdhury ve Pedrycz 2001; Özdemir 2013). Literatürde çok sayıda durulaştırma yöntemi vardır. Bu tezde durulaştırma yöntemlerinden en sık kullanılan ağırlık merkezi (COG - center of gravitiy) ve ağırlıklı ortalama (WA – weighted average) durulaştırma yöntemlerinden bahsedilecektir.

2.7.1. Ağırlık merkezi (COG – center of gravity) yöntemi

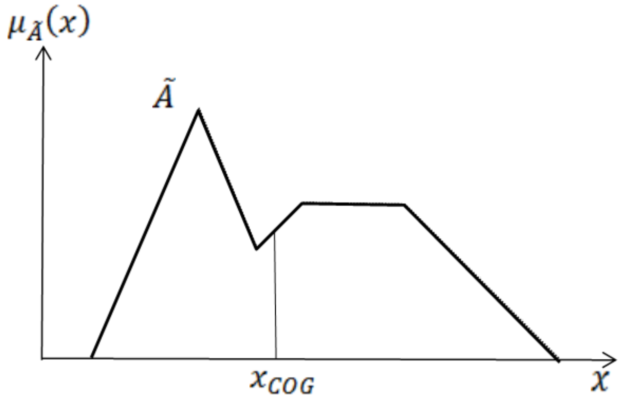
Literatürde en çok kullanılan durulaştırma yöntemlerinden biri olan bu yönteme göre, üyelik fonksiyonu altındaki alanın ağırlık merkezi aşağıda gösterildiği gibi hesaplanır.

$$X_{COG} = \frac{\int_X x\mu_{\tilde{A}}(x)\,dx}{\int_X \mu_{\tilde{A}}(x)\,dx} \tag{2.99}$$

Ele alınan üyelik fonksiyonu kesikli ise ağırlık merkezi yöntemine göre durulaştırma (2.100) ‘te gösterildiği gibi yapılır.

$$X_{COG} = \frac{\sum_{i=1}^{n} x_i\mu_{\tilde{A}}(x_i)\,dx}{\sum_{i=1}^{n} \mu_{\tilde{A}}(x_i)\,dx} \tag{2.100}$$

Bu yöntemin grafiksel gösterimi Şekil 2.3‘te gösterildiği gibidir (Uçal Sarı 2012).



Şekil 2.3. Ağırlık merkezi durulaştırma yöntemi (Uçal Sarı 2012)

2.7.2. Ağırlık ortalama (WA – weighted avarage) yöntemi

Sadece simetrik üyelik fonksiyonlarına sahip bulanık kümelerin durulaştırılmasında kullanılabilen bu yöntemde, sürecin çıktısını oluşturan her bir üyelik fonksiyonunun en büyük değerini aldığı değerlerin ağırlıklı ortalaması alınır. Yöntemin matematiksel gösterimi (2.101)'teki gibidir (Uçal Sarı 2012).

$$X_{WA} = \frac{\sum \bar{x}\mu_{\tilde{A}}(\bar{x})}{\sum \mu_{\tilde{A}}(\bar{x})} \tag{2.101}$$

3. BULANIK REGRESYON

İstatistikte, ilgilenilen bağımlı (açıklanan) ve bağımsız (açıklayıcı) değişkenler arasında bir matematiksel model kurmak hayati bir önem taşımaktadır. Bu amaçla çeşitli modelleme yöntemleri geliştirilmiştir. Bu yöntemlerden biri ise regresyon analizidir. Regresyon analizi, ele alınan bağımlı değişken ve bağımsız değişkenler arasında matematiksel bir model oluşturmayı amaçlamaktadır. Bağımlı y değişkeni ile tek bir bağımsız x değişkeni ile oluşturulan doğrusal regresyon modeli şu şekildedir.

$$y = \beta_0 + \beta_1 x + \varepsilon \tag{3.1}$$

Burada β_0 kesim noktası ve β_1 eğim noktası, ε ise hata terimidir.

Aralarında ilişki aranacak değişkenler çoğunlukla karmaşık ilişki ağlarına sahip olmakta ve bağımlı değişkeni etkileyen birden çok bağımsız değişkenler olabilmektedir. Bu durumlarda çoklu doğrusal regresyon modeli kullanılmaktadır. k tane bağımsız değişkenden oluşan çoklu doğrusal regresyon modeli ise şu şekildedir.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon \tag{3.2}$$

Dünya karmaşık yapılarla örülüdür. Bu karmaşıklığın kaynakları genel olarak belirsizlik, kesin düşünceden yoksunluk ve karar verilemeyiştir. İnsan düşüncesinin tam olarak olgunlaşmamış olmasından kaynaklı belirsizlikler birçok teknik, sosyal ve ekonomik konularda her zaman bulunur. Değişik şekillerde ortaya çıkabilen belirsizlik ve karmaşıklık gibi tam ve kesin olmayan bilgi kaynaklarına bulanık kaynaklar adı verilmektedir (Şen 2001;Düzyurt 2008). Zadeh (1965) tarafından geliştirilmiş olan bulanık küme teorisi bulanık kaynaklarla başa çıkmada etkin bir yöntem olarak kullanılmaktadır.

Klasik regresyon analizinin pek çok uygulama alanı olmasına rağmen;

- Gözlem sayısının yetersiz olması (küçük örneklem)

- Varsayımların doğruluğunu sınarken karşılaşılan güçlükler
- Bağımlı ve bağımsız değişkenler arasındaki ilişkideki bulanıklık
- Olayların ortaya çıkmasındaki belirsizlik
- Doğrusallık varsayımının bozulması

gibi durumlarda problemler meydana gelmektedir. Bu tür problemler bulanık regresyon analizi ile çözümlenebilmektedir (Shapiro 2004). Bulanık mantık teorisine dayanılarak geliştirilmiş yöntemlerden biri olan bulanık regresyon, klasik regresyon analizinin varsayımlarının sağlanamadığı durumlarda gerçek yaşam problemlerine uygulanabilirliği açısından daha çok tercih edilmektedir (Düzyurt 2008).

Bulanık regresyon modelinin teorik yapısı ilk olarak 1982 yılında Tanaka ve ark. tarafından ortaya atılmıştır (Tanaka ve ark. 1982). Bulanık regresyon analizi zaman içerisinde bir çok araştırmacı tarafından geliştirilmiştir. Genel anlamda bulanık regresyon, iki kategoride sınıflandırılabilir:

i. Değişkenler arasındaki ilişkilerden kaynaklı bulanıklık içeren bulanık regresyon.
ii. Değişkenler arasındaki ilişkiden değil değişkenkerin kendilerinin bulanık olduğu durumlarda bulanık regresyon.

Diğer bir bakış açısıyla bulanık regresyon şu şekilde sınıflandırılabilir:

i. Olabilirlik konseptine dayalı posibilistik regresyon analizi olarak bulanık regresyon
ii. En küçük karaler yaklaşımına dayalı bulanık regresyon (Taheri 2003).

İlerleyen bölümlerde çeşitli bulanık regresyon modelleri ve bu modellerin teorik yapıları verilecektir.

3.1. Tanaka'nın Bulanık Regresyon Modeli

Tanaka ve ark. (1982) tarafından geliştirilmiş olan bulanık regresyon modelinde, regresyon katsayıları bulanık sayılardır. Bu bulanıklık, bağımsız değişkenler arasındaki ilişkiden kaynaklanmaktadır. Regresyon katsayıları bulanık sayılar olduğundan, tahmin edilen bağımlı değişken Y'de bulanık sayıdır (Tanaka ve ark. 1982). Gözlemlenen ve tahmin edilen değerler arasındaki sapmaların ölçüm hatasından kaynaklandığı varsayılan klasik regresyon analizinin tersine, bulanık regresyonda bu sapmanın bulanık parametrelerle ifade edilen sistem yapısından kaynaklandığı varsayılmaktadır (Kahraman ve ark. 2006). Bulanık regresyondaki amaç gözlemlenen tüm bulanık veriler için uygun bir regresyon modeli oluşturmaktır. Basit bir bulanık regresyon denklemi aşağıda gösterilmiştir.

$$\tilde{Y} = \tilde{A}_0 + \tilde{A}_1 X \tag{3.3}$$

Burada $\tilde{Y}$ bağımlı değişken, $\tilde{A}_j$ ise bulanık katsayılardır. Bulanık katsayı $\tilde{A}_j = \left(c_j, s_j\right)$, merkezi değeri c_j ve yayılım değeri s_j olan simetrik üçgensel bulanık sayı olarak ifade edilir.

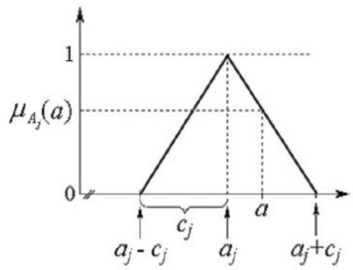
n tane bağımsız değişkenli temel bir bulanık regresyon denklemi aşağıdaki gibi gösterilmektedir.

$$\tilde{Y} = \tilde{A}_0 + \tilde{A}_1 X_1 + ... + \tilde{A}_n X_n = \tilde{A}\mathrm{X} \tag{3.4}$$

Burada $\mathrm{X} = \left[\mathrm{X}_0, \mathrm{X}_1, ..., \mathrm{X}_n\right]^{\mathrm{T}}$ bağımsız değişkenler vektörü, $\tilde{A} = \left[\tilde{A}_{0,}\tilde{A}_1, ..., \tilde{A}_n\right]$ ise $\tilde{A}_j = \left(c_j, s_j\right)$ şeklinde simetrik üçgensel bulanık katsayı vektörleridir. Bu tezde bağımsız değişken sayısı n ile ve gözlem sayısı m ile gösterilmek üzere bulanık $\tilde{A}_j = \left(c_j, s_j\right)$ sayısının üyelik fonksiyonu aşağıdaki gibi tanımlanmaktadır.

$$\mu_{\tilde{A}_j}(a_j)=\begin{cases}1-\dfrac{|c_j-s_j|}{s_j}, & c_j-s_j\le a_j\le c_j+s_j \;,\; j=1,2,...,n\\ 0 & ,\quad \text{dd}\end{cases} \tag{3.5}$$

Bulanık regresyon modelinin simetrik parametreleri Şekil 3.1.'de gösterildiği gibidir.



Şekil 3.1. Simetrik bulanık parametreler (Shapiro 2004)

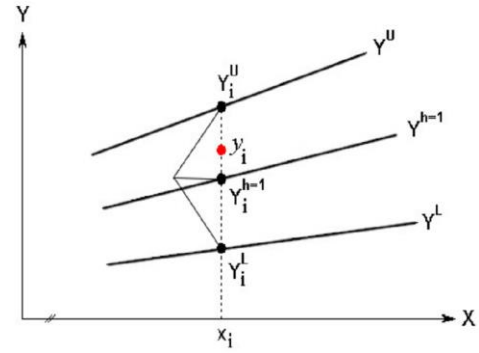
Bulanık regresyon modeli üçgen bulanık sayılar yardımıyla şu şekilde yazılabilir.

$$\tilde{Y}_i=(c_0,s_0)+(c_1,s_1)X_{1i}+...+(c_n,s_n)X_{ni} \tag{3.6}$$

Bulanık sayı $\tilde{Y}_i$ 'nın üyelik fonksiyonu genişleme prensibi yardımıyla aşağıdaki gibi hesaplanabilir.

$$\mu(\tilde{Y}_i)=\begin{cases}1-\dfrac{|Y_i-X^tc|}{s^t|X|} & ,\; X\neq 0\\ 1 & ,\; X=0,\, Y\neq 0,\, i=1,2,...,m\\ 0 & ,\; X=0,\, Y=0\end{cases} \tag{3.7}$$

Burada $s^t=(s_0,s_1,...,s_n)$ ve $c=(c_0,c_1,...,c_n)$ 'dir. Bağımlı değişkenin değeri $\tilde{Y}_i=\left(Y_i^U,Y_i^{h=1},Y_i^L\right)$ şeklindedir ve Şekil 3.2'deki gibi gösterilmektedir. Bulanık sayının alt sınır değeri $Y_i^L = \sum_{j=0}^{n}\left(c_j - s_j\right)\mathrm{X}_{ij}$, merkez değeri $Y_i^{h=1} = \sum_{j=0}^{n} c_j \mathrm{X}_{ij}$ ve üst sınır değeri $Y_i^U = \sum_{j=0}^{n}\left(c_j + s_j\right)\mathrm{X}_{ij}$ şeklinde hesaplanmaktadır. Alt ve üst regresyon doğrularının gösterimi Şekil 3.2.'de olduğu gibidir.



Şekil 3.2. Bulanık Regresyon Aralıkları (Shapiro 2004)

Bulanıklığı minimize edecek şekilde bulanık regresyon analizinin uygulanabilmesi için, $\tilde{Y}_i$ bulanık sayısının (3.8)'deki amaç fonksiyonu ve (3.9) 'daki kısıta göre minimize edilmesi gerekmektedir. Eşitsizlik (3.9)'a göre kısıtlar her Y_i gözlem değerinin $\tilde{Y}_i$ 'ya en az h derecesi ile bağlı olmasını gerektirir. h derecesi tahmin edilen regresyon denkleminin veri setine uyumunun bir göstergesi olarak yorumlanabilir (Wang ve Tsaur 2000a).

$$Min\, c^t \left|X\right| = Min \sum_{j=0}^{n}\left(c_j \sum_{i=1}^{m}\left|x_{ij}\right|\right) \qquad (3.8)$$

$$1 - \frac{\left|Y_i - \mathrm{X}^t c\right|}{s^t\left|\mathrm{X}\right|} \geq h$$

$$i = 1, 2, \ldots, m \quad j = 1, 2, \ldots, n \qquad (3.9)$$

Tanaka ve ark. (1982) tarafından önerilmiş olan bulanık regresyon modelinin bilinmeyen parametrelerinin tahmin edilebilmesi için (3.10) ile gösterilen amaç fonksiyonunun aşağıda verilmiş olan üç kısıt altında çözülmesi gerekmektedir.

$$\min \mathrm{S} = \mathrm{ns}_0 + s_1 \sum_{j=1}^{n} \left|\mathrm{X}_j\right| \qquad (3.10)$$

Kısıtları altında; $s_0 \geq 0,\ s_1 \geq 0$ (3.11)

$$\sum_{j=0}^{l} c_i \mathrm{X}_{ij} + (1-h) \sum_{j=0}^{l} s_j \left|\mathrm{X}_{ij}\right| \geq Y_i + (1-h) e_i \quad i = 1, 2, \ldots, n \qquad (3.12)$$

$$\sum_{j=0}^{l} c_i \mathrm{X}_{ij} - (1-h) \sum_{j=0}^{l} s_j \left|\mathrm{X}_{ij}\right| \leq Y_i - (1-h) e_i \quad i = 1, 2, \ldots, n \qquad (3.13)$$

Eşitlik (3.10)'daki amaç fonksiyonu toplam bulanıklığı minimize etmektedir. Eşitlik (3.12) ve (3.13) gözlemlenen bulanık veri değeri $\tilde{Y}_i = (y_i, \mathrm{e}_i)$ ile ilgilidir ve y_i bulanık merkezi, e_i ise bulanık yayılım ölçüsünü göstermektedir. Gözlemlenen veri kesin ise (bulanık değilse), verinin bulanık yayılım ölçüsünün değeri sıfırdır. Böylece belirli bir kesin sayı bulanık sayının özel bir durumu haline gelmiş olmaktadır (Chang ve Ayyub 2001). Bulanık regresyon modelinin katsayılarının tahmininde kullanılan kısıtlarda h termininin varlığı göze çarpmaktadır. İlerleyen bölümde bu terimden kısaca bahsedilecektir.

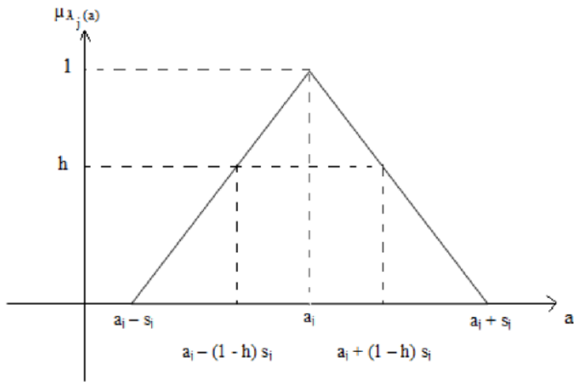
3.1.1. h-inanç derecesi

Klasik regresyon analizinde model kurulurken yapılması muhtemel hatalar tek bir hata teriminde toplanırken, bulanık regresyonda ise hata, modelin tüm katsayılarına dağıtılmaktadır. Diğer bir ifadeyle, bulanık regresyonda her bir parametre belirli bir bulanıklık seviyesinde tahmin edilmiş olmaktadır. Söz konusu bulanıklık seviyesine "*h* terimi" adı verilir. *h* inanç derecesi $[0,1]$ aralığında değerler almaktadır.

h değerinin belirlenmesi araştırmacıya bırakılmıştır. Araştırmacı, çalışmanın başında veri setinin eksik, yarım veya tam olması durumuna göre *h* değerini belirlemektedir (Moskowitz ve Kim, 1993).

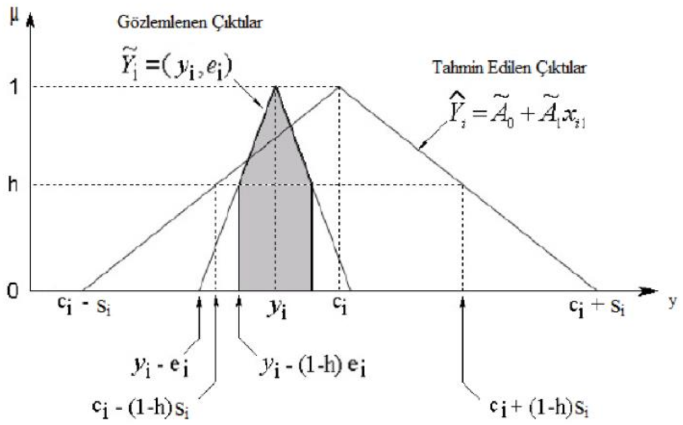
Bu bilgilere ek olarak, h'nin seçiminde genellikle tavsiye edilen değer ise 0,5 değeridir. Fakat bu değer araştırmaya bağlı olarak değişiklik gösterebilmektedir. Çalışmalarda farklı h değerlerinde analizler yapıp karar vermek gerekmektedir (Düzyurt 2008). Çünkü belirli bir veri seti için *h* inanç derecesinin artmasıyla bulanık regresyon katsayılarının yayılımları da artmaktadır (Shapiro 2004).

h inanç derecesinin grafiksel olarak gösterimi aşağıdaki gibidir.



Şekil 3.3. Simetrik üyelik fonksiyonunda inanç derecesi (Kaya 2010)

İnanç derecesinin değeri bire yaklaştıkça $\tilde{Y}_i$ parametresinin bulanık aralıkları daralır, *h*'nin değeri sıfıra yaklaştıkça ise bulanık aralıklar genişler. Aşağıdaki şekilde simetrik üçgen bir üyelik fonksiyonu için *h* değerinin sıfıra ve bire çok yakın değerler almaması gerektiği gösterilmiştir.



Şekil 3.4. Simetrik Üyelik Fonksiyonu: Gözlemlenen ve Tahmin Karşılaştırılması (Shapiro 2004)

Bulanık regresyon için tahmin edilen parametreler yardımıyla h değeri elde edilebilir. Bu değer “h^* ” ile gösterilmek üzere h’yi tahmin sonuçlarından göstermek için aşağıdaki formül uygulanmaktadır (Kaya 2010).

$$h_i^* = 1 - \frac{(c_i - y_i)}{s_i} \tag{3.14}$$

3.2. Tanaka’nın Revize Bulanık Regresyon Modeli

He ve ark. (2007) tarafından yapılan bir çalışma da Tanaka’nın bulanık regresyon modelinde h_i değerinin sadece kestirilen $\tilde{y}_i$ ‘lerin yayılımlarına $\left(s_0 + \sum_{j=1}^{m} s_j \left|x_{ji}\right|\right)$ bağlı olmadığı, ayrıca $\tilde{y}_i$ ’nin merkezi ile gözlemlenen y_i arasındaki uzaklığa da $\left(c_0 + \sum_{j=1}^{m} c_j x_{ji}\right)$ bağlı olduğunu farketmişlerdir. Bulanık regresyonda amaç $\tilde{y}_i$ ’lerin dağılımlarının toplamının minimize edilmesidir. He ve ark. (2007) bu amaca ulaşmak için ve daha iyi bir model uyumu elde edebilmek için yeni bir amaç fonksiyonu geliştirilmiştir.

$$\text{Min}: \ \text{SF}+\sum_{i=1}^{m}\text{d}_i = \sum_{i=1}^{m}\left(s_0+\sum_{j=1}^{n}s_j\left|x_{ji}\right|\right)+\sum_{i}^{m}d_i \tag{3.15}$$

Kısıtları altında;

$$0 \le h_i = {1-d_i}\bigg/\left(s_0+\sum_{j=1}^{m}s_j\left|x_{ji}\right|\right)$$

$$(3.16)$$

$$d_i = \left|y_i - \left(c_0+\sum_{j=1}^{m}c_j x_{ji}\right)\right| \quad i=1,2,\ldots,\text{m}$$

$$s_j \ge 0, \quad j=0,1,2,\ldots,\text{n}$$

Eşitlik (3.15) bulanık parametrelerin tahmininde kullanılan amaç fonksiyonunu, (3.16) ise bu amaç fonksiyonuna bağlı kısıtları göstermektedir. Tanaka'nın revize bulanık regresyon modelinde sistem bulanıklığının Tanaka'nın bulanık regresyon modeline göre daha düşük olduğu görülmüştür (He ve ark. 2007).

Bulanık regresyon analizinde amaç veriye en uygun modeli belirlemek olduğundan uyum iyiliğinin ölçüsü olan $\bar{h}$ değerinin hesaplanması gerekmektedir. $\bar{h}$ değeri tahmin edilen h_i^* değerlerinin ortalamasıdır ve (3.17) ile gösterildiği gibi hesaplanabilir (Tanaka ve ark. 1982).

$$\bar{h} = \sum_{i=}^{m}h_i^* \Big/ m \tag{3.17}$$

$\bar{h}$ değeri klasik regresyon analizindeki R^2 değerine benzeyen bir uyum ölçüdür ve oluşturulan modelin veri setine uyumunu göstermektedir. $\bar{h}$ değeri sıfır ile bir arasında değerler almaktadır. $\bar{h}$ 'nın değeri arttıkça oluşturulan bulanık regresyon modelinin veri setine uyumu da artmış olmaktadır.

He ve ark. (2007) tarafından önerilmiş olan modelin Tanaka'nın bulanık regresyon modeline göre daha iyi sonuçlar verdiği ve bu modelden elde edilen sistem ortalama uyumluluk derecesi $\bar{h}$ 'nın Tanaka modeline göre artış sağladığı gözlemlenmiştir (He ve ark. 2007).

3.3. Bulanık En Küçük Kareler Regresyonu

Bulanık en küçük kareler regresyonu, 1988 yılında Diamond tarafından bulanık regresyon parametrelerini tahmin etmek için ortaya atılmıştır (Wang ve Tsaur 2000b). Bulanık en küçük kareler regresyonu, klasik regresyonda olduğu gibi tahmin değeri ile gözlem değeri arasındaki farkı minimize etmeyi amaçlamaktadır. Bulanık en küçük kareler regresyonu, tahmin edilen $\tilde{y}_i$ bulanık bağımlı değişken değeri ile gözlemlenen y_i değerleri arasındaki mesafeyi minimize etmektedir.

Celmins (1987), Diamond (1988), Savic ve Pedrycz (1991) ve Chang ve Ayyub (1993) tarafından bulanık en küçük kareler regresyonuna farklı bakış açıları getirilmiştir. Celmins (1987), bulanık veri ve bir model arasında bir ölçüm tanımlamış ve bu ölçümü model uygunluk kriteri olarak kullanmıştır. Diamond (1988), en küçük kareler yöntemi geliştirmiştir. Savic ve Pedrycz (1991), minimum bulanıklık kriterini klasik regresyonla birleştirerek bulanık en küçük kareler regresyonu için bütünleşmiş bir yaklaşım geliştirmiştir. Chang ve Ayyub (1993) bulanık en küçük karelerin standart sapma, korelasyon katsayısı gibi güvenilirlik konularına dikkat çekmiştir (Chang ve Ayyub 2001).

$\tilde{Y}_i = \tilde{A}_0 + \tilde{A}_1 X_{1i}$ şeklindeki tek değişkenli bir bulanık regresyon modelinde $\tilde{y}_i$ bulanık bağımlı değişken değeri ile gözlemlenen y_i değerleri arasındaki fark (mesafe) (3.18) ile tanımlanmıştır (Wang ve Tsaur 2000b).

$$\mathrm{r}\left(\tilde{A}_0 + \tilde{A}_1 X_{1i}, Y_i\right) = \sum d\left(\tilde{A}_0 + \tilde{A}_1 X_{1i}, Y_i\right)^2 = \left(c_0 + s_0 + c_1 X_{i1} + s_1 X_{i1} - y_i - e_i\right)^2 + \left(c_0 + c_1 X_{i1} - y_i - e_i\right)^2 + \left(c_0 - s_0 + c_1 X_{i1} - s_1 X_{i1} - y_i - e_i\right)^2 \tag{3.18}$$

Eşitlik (3.18)'in c_0, c_1 ve s_0, s_1 değerlerine göre türevleri alınıp sıfıra eşitlenerek, bulanık en küçük kareler regresyonunun parametreleri bulunmuş olur. Eşitlik (3.18)'de türev alındığında, parametrelerin sıfırdan büyük olup olmamalarına göre dört durum ortaya çıkmaktadır.

<u>1.Durum:</u> $c_0 > 0$, $c_1 > 0$ ve s_0, s_1 kısıtsız ise

Bulanık farkın c_0, c_1 ve s_0, s_1'e göre 1. dereceden türev alınıp sıfıra eşitlenirse c_0, c_1 ve s_0, s_1 değerleri elde edilmiş olur.

$$\partial r(\tilde{A}_0 + \tilde{A}_1 X_{i1}, Y_i) / \partial c_0 = 2\sum_i \{c_0 - s_0 + (c_1 - s_1)X_{i1} - (y_i - e_i) + c_0 + s_0 + (c_1 + s_1)X_{i1} - (y_i + e_i) + c_0 + c_1 X_{i1} - y_i\} \tag{3.19}$$

$$\partial r(\tilde{A}_0 + \tilde{A}_1 X_{i1}, Y_i) / \partial s_0 = 2\sum_i \{c_0 - s_0 + (c_1 - s_1)X_{i1} - (y_i - e_i)(-1) + c_0 + s_0 + (c_1 + s_1)X_{i1} - (y_i + e_i) + c_0 + c_1 X_{i1} - y_i\} \tag{3.20}$$

$$\partial r(\tilde{A}_0 + \tilde{A}_1 X_{i1}, Y_i) / \partial c_1 = 2\sum_i X_{i1}\{c_0 - s_0 + (c_1 - s_1)X_{i1} - (y_i - e_i) + c_0 + s_0 + (c_1 + s_1)X_{i1} - (y_i + e_i) + c_0 + c_1 X_{i1} - y_i\} \tag{3.21}$$

$$\partial r(\tilde{A}_0 + \tilde{A}_1 X_{i1}, Y_i) / \partial s_1 = 2\sum_i X_{i1}\{-[c_0 - s_0 + (c_1 - s_1)X_{i1} - (y_i - e_i)] + [c_0 + s_0 + (c_1 + s_1)X_{i1} - (y_i + e_i)]\} \tag{3.22}$$

Denklemler sıfıra eşitlendiğinde c_0, c_1, s_0 ve s_1 değerleri bulunmuş olur.

$$c_1 = (N\sum_i X_{i1}y_i - \sum_i X_{i1}\sum_i y_i) / (N\sum_i X_{i1}^2 - (\sum_i X_{i1})^2) \tag{3.23}$$

$$c_0 = \bar{y} - c_1\bar{X} \tag{3.24}$$

$$s_1 = (N\sum_i X_{i1}e_i - \sum_i X_{i1}\sum_i e_i) / (N\sum_i X_{i1}^2 - (\sum_i X_{i1})^2) \tag{3.25}$$

$$s_1 = \bar{e} - c_1\bar{X} \tag{3.26}$$

<u>2.Durum:</u> $c_0 < 0$, $c_1 > 0$ ve $s_0 < 0$, s_1 kısıtsız ise

$$c_1 = (N\sum_i X_{i1}y_i - \sum_i X_{i1}\sum_i y_i)/(N\sum_i X_{i1}{}^2 - (\sum_i X_{i1})^2) \tag{3.27}$$

$$c_0 = \bar{y} - c_1\bar{X} \tag{3.28}$$

$$s_1 = (N\sum_i X_{i1}e_i - \sum_i X_{i1}\sum_i e_i)/(N\sum_i X_{i1}{}^2 - (\sum_i X_{i1})^2) \tag{3.29}$$

$$s_1 = -(\bar{e} - c_1\bar{X}) \tag{3.30}$$

<u>3.Durum:</u> $c_0 > 0$, $c_1 < 0$ ve s_0 kısıtsız, $s_1 < 0$ ise

$$c_1 = (N\sum_i X_{i1}y_i - \sum_i X_{i1}\sum_i y_i)/(N\sum_i X_{i1}{}^2 - (\sum_i X_{i1})^2) \tag{3.31}$$

$$c_0 = \bar{y} - c_1\bar{X} \tag{3.32}$$

$$s_1 = (N\sum_i X_{i1}e_i - \sum_i X_{i1}\sum_i e_i)/(N\sum_i X_{i1}{}^2 - (\sum_i X_{i1})^2) \tag{3.33}$$

$$s_1 = \bar{e} - c_1\bar{X} \tag{3.34}$$

<u>4.Durum:</u> $c_0 < 0$, $c_1 < 0$ ve $s_0 < 0$, $s_1 < 0$ ise

$$c_1 = (N\sum_i X_{i1}y_i - \sum_i X_{i1}\sum_i y_i)/(N\sum_i X_{i1}{}^2 - (\sum_i X_{i1})^2) \tag{3.35}$$

$$c_0 = \bar{y} - c_1\bar{X} \tag{3.36}$$

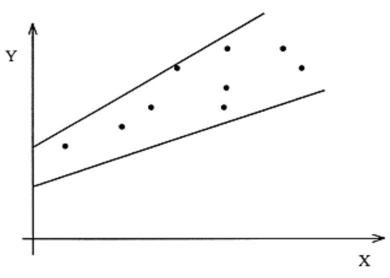
$$s_1 = (N\sum_i X_{i1}e_i - \sum_i X_{i1}\sum_i e_i)/(N\sum_i X_{i1}{}^2 - (\sum_i X_{i1})^2) \tag{3.37}$$

$$s_1 = -(\bar{e} - c_1\bar{X}) \tag{3.38}$$

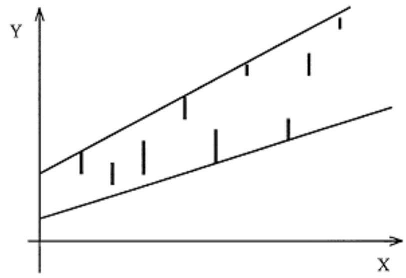
Bulanık en küçük kareler regresyonunu uygulamak, değişken sayısı arttığında zorlaşmaktadır. s_0, s_1 ve e_i değerleri sıfır olduğunda, bulanık en küçük kareler yönteminden elde edilen parametreler, klasik en küçük kareler yönteminden elde edilen parametrelerle aynı olmaktadır. Bu yöntemde, minimize edilen regresyon aralıkları Tanaka'nın bulanık regresyon modeline göre daha dardır. Bulanık en küçük kareler regresyonu tahmin yönünden avantajlı olmasına rağmen, oldukça uzun hesaplamalar gerektirmesi bu yöntemin dezavantajıdır (Wang ve Tsaur 2000b).

3.4. Aralık Regresyonu

Bir diğer bulanık regresyon analizi metodu olan aralık regresyonunda, bulanık veri ve bulanık regresyon katsayıları aralık sayı gibi davranmaktadır. Bu metotta, aralık işlemleri bulanık regresyona eklenir ve regresyon katsayıları (sistem parametreleri) aralıklar halinde tanımlanır. Bulanık regresyon katsayıları, tüm bulanık çıktılar bulanık regresyon modelindeymiş gibi hesaplanmaktadır. Şekil 3.5.'de kesin X ve kesin Y verileri için regresyon aralığı, Şekil 3.6.'da ise kesin X ve bulanık Y için regresyon aralığı gösterilmiştir.



Şekil 3.5. Belirli X ve belirli Y değerleri için aralık regresyon modeli (Chang ve Ayyub 2001)



Şekil 3.6. Belirli X ve bulanık Y değerleri için aralık regresyon modeli (Chang ve Ayyub 2001)

$\tilde{Y} = \tilde{A}_0 + \tilde{A}_1 X$ şeklindeki tek değişkenli bir bulanık regresyon için regresyon aralıkları, aşağıdaki doğrusal programlama probleminin çözülmesiyle elde edilmektedir.

$$\min\ ns_0 + s_1 \sum_{i=1}^{n} X_i \tag{3.39}$$

Kısıtları altında; $s_0 > 0,\ s_1 > 0$

$$\left(c_0 - s_0\right) + \left(c_1 - s_1\right) X_i \le Y_{i,L} \tag{3.40}$$

$$\left(c_0 + s_0\right) + \left(c_1 + s_1\right) X_i \ge Y_{i,\mathrm{U}} \quad i = 1, 2, \ldots, \mathrm{m} \tag{3.41}$$

$Y_{i,L}$ ve $Y_{i,\mathrm{U}}$ sırasıyla her bir bulanık veri için alt ve üst sınır değerleridir. Eşitlik (3.39)'deki amaç fonksiyonu, toplam bulanık büyüklükleri minimize edecek şekilde sonuçlanmaktadır. Eşitlik (3.40) ve (3.41)'taki kısıtlar ise, bulanık verileri bulanık regresyon modeli ile sınırlamak için kullanılmaktadır (Chang ve Ayyub 2001).

4. BULANIK LOJİSTİK REGRESYON

Regresyon analizi, bağımlı değişken ve bağımsız değişkenler arasındaki ilişkiyi araştıran ve bu ilişkiyi matematiksel olarak modelleyen istatistiksel bir yöntemdir. Bu analizde, aralarında ilişki kurulacak değişkenler sayısal olmak zorundadır. Regresyon analizi hataların sıfır ortalama ve sabit varyans ile normal dağıldığı ve hataların ilişkisiz (otokorelasyonsuz) olduğu varsayımlarının sağlandığı durumlarda kullanılabilir (Montgomery ve ark. 2001). Aralarında matematiksel ilişki kurulacak değişkenler her zaman sayısal olmayabilir. Bu durumda, lojistik regresyon klasik regresyona alternatif olarak kullanılabilir.

Lojistik regresyon analizini doğrusal regresyon analizinden ayıran en önemli özellik, lojistik regresyonda bağımlı değişkenin kategorik olmasıdır. Lojistik regresyon ve doğrusal regresyon arasındaki bu fark hem parametrik model seçimine hem de varsayımlara yansımaktadır. Lojistik regresyon analizinde de, doğrusal regresyon analizindeki gibi bazı değişken değerlerine dayanarak kestirim yapılmaya çalışılmaktadır. Doğrusal regresyon analizi ve lojistik regresyon analizi arasındaki farkları ifade etmek gerekirse;

i. Doğrusal regresyon analizinde bağımlı değişken sürekli değerler alırken, lojistik regresyonda bağımlı değişken kesikli bir değer olmaktadır.
ii. Doğrusal regresyon analizinde bağımlı değişkenin değeri, lojistik regresyonda ise bağımlı değişkenin alabileceği değerlerden birinin gerçekleşme olasılığı tahmin edilmektedir.
iii. Doğrusal regresyon analizinde bağımlı değişkenlerin çoklu normal dağılım göstermesi şartı aranırken, lojistik regresyonun uygulanabilmesi için bağımsız değişkenlerin dağılımına ilişkin bir zorunluluk bulunmamaktadır. (Kleinbaum ve Klein, 2002)

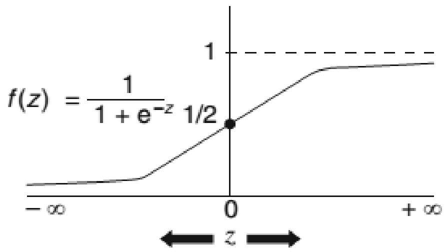
İstatistikteki çok değişkenli yöntemlerden biri olan lojistik regresyon, bu yöntemlerin sınıflama özelliğine de hizmet etmektedir (Özdamar 2004). İstatistikte sınıflama yapmak amacıyla kullanılabilecek diğer modeller kümeleme ve diskriminant analizidir. Kümeleme analizinde bireylerin (gözlemlerin) atanacağı grup sayısı bilinmezken, diskriminant ve lojistik regresyon analizinde ise bilinmektedir.

İlerleyen bölümlerde önce lojistik regresyon modelinin nasıl oluşturulduğundan, model parametrelerinin tahmin yöntemlerinden, daha sonra, yukarıda sözü geçen lojistik regresyon modellerinin matematiksel yapısından bahsedilecektir.

Lojistik modelin temelini oluşturan ve $f(z)$ ile gösterilen lojistik fonksiyon aşağıdaki gibi tanımlanır.

$$f(z) = \frac{1}{1+e^{-z}} \tag{4.1}$$

Şekil 4.1'den z'nin tanım aralığının $(-\infty, +\infty)$ arasında olduğu görülmektedir. $f(z)$ lojistik fonksiyonunun 0 ile 1 arasında bir değişim aralığına sahip olması ve olasılık modelinde uygulama kolaylığı sağlaması bu fonksiyonun tercih edilmesindeki önemli nedenlerden biridir (Kleinbaum ve Klein 2002; Le 2003).



Şekil 4.1. Lojistik fonksiyonun şekli (Kleinbaum ve Klein 2002)

Lojistik fonksiyondan lojistik modeli elde edebilmek için z'nin α , β ve X 'lerin doğrusal toplamı olarak $z = \alpha + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p$ biçiminde yazılması gerekmektedir. Gerekli dönüşüm yapıldıktan sonra lojistik model (4.2)'deki gibi elde edilmektedir (Kleinbaum ve Klein 2002).

$$\mathrm{f}(\mathrm{z})=\frac{1}{1+\mathrm{e}^{-\mathrm{z}}}=\frac{1}{1+\mathrm{e}^{-\left(\alpha+\sum_{k=1}^{p}\beta_k X_k\right)}} \tag{4.2}$$

Bir regresyon probleminde önemli nokta, bağımsız değişkenin herhangi bir değeri verildiğinde bağımlı değişkenin bu değere karşılık gelen ortalama değerinin elde edilmesidir. Bu niceliğe "koşullu beklenen değer (ortalama)" denir ve $\mathrm{E}(Y \mid X)$ şeklinde gösterilir. Buradaki "Y" bağımlı değişkeni, "X" ise bağımsız değişkeni göstermektedir. Lojistik dağılım kullanıldığında koşullu beklenen değeri göstermek için $\pi(x)=\mathrm{E}(Y \mid X)$ ifadesi kullanılmaktadır. Böylece lojistik regresyon modeli (4.3) ile ifade edilebilir.

$$\pi(x)=\frac{e^{\beta_0+\beta_1 x}}{1+e^{\beta_0+\beta_1 x}} \tag{4.3}$$

Koşullu beklenen değerin 0 ile 1 arasında bir olasılık değeri alabilmesi için bağımlı değişken $\pi(x)$'i $(-\infty,+\infty)$ aralığında tanımlı hale getirecek bir dönüşümün uygulanması gerekmektedir. Eşitlik (4.3)'te verilen lojistik regresyon modelinde $\pi(x)/\left[1-\pi(x)\right]$ dönüşümü yapıldığında bağımlı değişkenin sınırları $(0,+\infty)$ olmaktadır. Bağımlı değişkenin sınırlarını $(-\infty,+\infty)$ aralığına dönüştürmek için $\pi(x)/\left[1-\pi(x)\right]$ oranının doğal logaritmasının alınması gerekmektedir. Bu dönüşüm "lojit dönüşümü" olarak adlandırılmaktadır. Lojit dönüşümü sonucu elde edilen yeni bağımlı değişken, bağımsız değişkenlerin doğrusal bir fonksiyonu olmaktadır. Bu durum (4.4)'den görülebilmektedir.

$$g(x)=\ln\left[\frac{\pi(x)}{1-\pi(x)}\right]=\beta_0+\beta_1 x \tag{4.4}$$

$\pi(x)/\left[1-\pi(x)\right]$ oranına "odds oranı" denir. Odds oranı bir olayın gerçekleşmesi olasılığının, gerçekleşmeme olasılığına oranı olarak tanımlanabilir.

$g(x)$, doğrusal regresyon modelinde istenen özelliklerin çoğuna sahiptir. Lojit $g(x)$ parametreleri, (β_0, β_1) bakımından doğrusaldır ve x'in aldığı değerlere bağlı olarak $(-\infty, +\infty)$ aralığındadır.

Doğrusal regresyon modelinde, bağımlı değişken için yapılan gözlem $y = E(Y|x) + \varepsilon$ şeklinde gösterilmektedir. Burada ε 'a "hata terimi" adı verilir ve gözlemin koşullu ortalamadan sapma miktarını ifade etmektedir. Doğrusal regresyon analizinde hata terimi ε'un 0 ortalama ve bağımsız değişkenin her bir düzeyi için sabit bir varyansla normal dağıldığı varsayılmaktadır $\left(\varepsilon \sim N(0, \sigma^2)\right)$. Verilen x için bağımlı değişken y'nin koşullu dağılımı, $E(Y|x)$ ortalamasına ve sabit bir varyansa sahip bir normal dağılımdır. Ancak bağımlı değişken iki kategoriden oluştuğunda ise durum farklı olmaktadır. Bu durumda verilen x için bağımlı değişken $y = \pi(x) + \varepsilon$ şeklinde ifade edilmektedir. ε'un bir veya iki olası değeri söz konusudur;

- Eğer y=1 ise $\pi(x)$ olasılıkla $\varepsilon = 1 - \pi(x)$,
- Eğer y=0 ise $1 - \pi(x)$ olasılıkla $\varepsilon = -\pi(x)$ değerini almaktadır.

Böylece ε, 0 ortalamalı ve $\pi(x)\left[1 - \pi(x)\right]$ varyanslı bir dağılıma sahip olmaktadır. Sonuç değişkeni Y'nin koşullu dağılımı, $\pi(x) = \mathrm{E}(Y \mid X)$ koşullu dağılımına göre bir binom dağılımına sahiptir (Hosmer ve Lemeshow 2000).

x_i i. birimin bağımsız değişken için aldığı değer olmak üzere ve y_i iki değer alan bağımlı değişkeninin değerini göstermek üzere, $(x_i, y_i), i = 1, 2, ..., n$ çifti ile gösterilen n bağımsız gözlemin olduğu varsayılsın. Ayrıca bağımlı değişken belirli bir olayın varlığını gösteren 1 ve yokluğunu gösteren 0 ile kodlanmış olsun. Eşitlik (4.3)'teki lojistik regresyon modelinin veri setine uyması için denklemdeki β_0 ve β_1 parametrelerinin tahmin edilmesi gerekmektedir (Hosmer ve Lemeshow 2000).

Lojistik regresyon analizinde bilinmeyen parametrelerin tahminleri En Çok Olabilirlik Yöntemi (Maximum Likelihood Method), Yeniden Ağırlıklandırılmış

İteratif En Küçük Kareler Yöntemi ve Minimum Lojit Ki-Kare yöntemlerinden biriyle yapılabilir. Bu tezde lojistik regresyon analizinde parametre tahmin yöntemlerinden en sık kullanılanı olan En Çok Olabilirlik Yöntemi'nden bahsedilecektir.

En çok olabilirlik yöntemi, bilinmeyen parametrelerin gözlenmiş veri kümesinden elde edilmesi olasılığını maksimum kılan değerler verir. Bu yöntemin uygulanabilmesi için öncelikle en çok olabilirlik fonksiyonu adı verilen fonksiyonun bulunması gerekir. Bu fonksiyonu maksimum yapan değerler, bu parametrelerin en çok olabilirlik kestirimleri olarak seçilir. Böylece, gözlemlenmiş değerlere en yakın kestirimler belirlenmiş olur (Hosmer ve Lemeshow 2000). Lojistik regresyon için bu fonksiyon ilerleyen bölümlerde gösterildiği gibi bulunmaktadır.

Eğer bağımlı değişken y'nin 0 ya da 1 değerlerini aldığı varsayılırsa, (4.3) 'te verilen $\pi(x)$ ifadesi herhangi bağımsız değişkenin değeri verildiğinde y'nin 1'e eşit olması koşullu olasılığını ve $1-\pi(x)$ niceliği, bağımsız değişkenin değeri verildiğinde y'nin 0'a eşit olma koşullu olasılığını verir. (x_i, y_i) çiftlerinin $y_i = 1$ olduğunda olabilirlik fonksiyonuna katkısı $\pi(x_i)$ kadar, $y_i = 0$ olduğunda olabilirlik fonksiyonuna katkısı $1-\pi(x_i)$ kadardır. (x_i, y_i) çiftlerinin olasılık fonksiyonuna katkısı şu şekildedir.

$$\pi(x_i)^{y_i}\left[1-\pi(x_i)\right]^{1-y_i} \tag{4.5}$$

Gözlem değerleri bağımsız olduğunda, olabilirlik fonksiyonu (4.5)'de verilen ifadelerin çarpımı olarak yazılabilir ve şu şekilde gösterilir:

$$l(\beta) = \prod_{i=1}^{n}\left[\pi(x_i)^{y_i}\left[1-\pi(x_i)\right]\right] \tag{4.6}$$

En çok olabilirlik metodunda (4.6)' daki ifadeyi maksimum yapan β değeri kestirim değerini vermektedir. Ayrıca, matematiksel olarak (4.6)'nın doğal logaritmasıyla çalışmak daha kolaydır. Bu ifade “log-olabilirlik fonksiyonu” olarak isimlendirilir ve şu şekilde gösterilir:

$$L(\beta) = \ln\left[l(\beta)\right] = \sum_{i=1}^{n}\left\{y_i \ln\left[\pi(x_i)\right] + (1-y_i)\ln\left[1-\pi(x_i)\right]\right\} \qquad (4.7)$$

Eşitlik (4.7) ile verilen log-olabilirlik fonksiyonunu maksimum yapan β değerlerini bulmak için β_0 ve β_1’e göre türev almak ve bu ifadeleri sıfıra eşitlemek gerekmektedir. Bu denklemler “olabilirlik eşitlikleri” olarak bilinir ve şu şekildedir:

$$\sum\left[y_i - \pi(x_i)\right] = 0 \qquad (4.8)$$

$$\sum x_i\left[y_i - \pi(x_i)\right] = 0 \qquad (4.9)$$

Doğrusal regresyon analizi söz konusu olduğunda olabilirlik eşitlikleri doğrusal olmalarından dolayı kolayca çözülmektedirler. Ancak lojistik regresyon analizinde bu eşitlikler β_0 ve β_1’e göre doğrusal olmadıklarından, denklemlerin çözümü için özel metotlar gerekmektedir. Bu metotlar doğası gereği iteratiftirler. Eşitlik (4.8) ve (4.9)’un iteratif yöntemlerle çözümüyle elde edilen β değerleri en çok olabilirlik tahmini olarak adlandırılır ve $\hat{\beta}$ ile gösterilir (Hosmer ve Lemeshow 2000).

Regresyon denklemleri kurulduktan sonra yapılması gereken ilk iş modeldeki bağımsız değişkenlerin istatistiksel olarak anlamlılığının kontrol edilmesidir. Bu kontrol bağımsız değişkenlere ait ilgili katsayılar için istatistiksel hipotezlerin kurulmasını ve test edilmesini kapsamaktadır. Lojistik regresyonda katsayıların anlamlılığının testi için kullanılan yöntemler;

- Olabilirlik oran testi
- Wald testi
- Skor testidir.

Bu tezde en çok kullanılan iki yöntem olan Olabilirlik oran testi ve Wald testi kısaca açıklanacaktır.

Wald testi maksimum olabilirlik kestirimi olan eğim katsayısı $\hat{\beta}_1$'in kendi standart hatasının kestirimiyle karşılaştırılmasından elde edilir. Lojistik regresyon modeli için bu test istatistiği aşağıdaki biçimde tanımlanabilir.

$$W = \frac{\hat{\beta}_1}{S\hat{E}\left(\hat{\beta}_1\right)} \tag{4.10}$$

Eğim parametresini gösteren $\mathrm{H}_0 : \beta_1 = 0$ hipotezi için W istatistiği standart normal dağılım göstermektedir (Hosmer ve Lemeshow 2000). Normal dağılım gösteren Wald istatistiğinin karesi alındığında da test uygulanabilmektedir. Bilindiği gibi Z istatistiğinin karesi, 1 serbestlik dereceli ki-kare dağılımına uymaktadır. Bu durumda Wald istatistiği (4.11)'de görüldüğü şeklini almaktadır.

$$W^2 = \left(\frac{\hat{\beta}_1}{S\hat{E}\left(\hat{\beta}_1\right)} \right)^2 \tag{4.11}$$

Olabilirlik oran testinin uygulanabilmesi için, lojistik regresyonda bağımsız değişkenin dahil olduğu ve olmadığı modellerdeki gözlenen değerler, tahmin edilen değerlerle karşılaştırılır. Bu karşılaştırmanın temeli log - olabilirlik fonksiyonuna dayanmaktadır. Bu fonksiyon aşağıdaki gibi tanımlanır:

$$D = -2\ln\left[\frac{\text{Üzerinde durulan modelin olabilirliği}}{\text{Doymuş modelin olabilirliği}}\right] \tag{4.12}$$

Değişken sayısı kadar parametre içeren modele "doymuş model" adı verilir. Eşitlik (4.12)'deki parantez içindeki ifade "olabilirlik oranı (likelihood ratio)" olarak adlandırılır. Eşitlik (4.12) dikkatli incelendiğinde "-2ln" teriminin varlığı göze çarpmaktadır. Bu ifade matematikseldir ve hipotez testinde kullanılmak üzere dağılımı bilinen bir nicelik elde edilmesi için gereklidir. D (deviance) istatistiğine “olabilirlik oran testi” denilmektedir. Bu istatistik doğrusal regresyondaki hata kareler toplamına benzemektedir. Olabilirlik oran testi için kullanılan D istatistiği aşağıda görüldüğü gibi hesaplanır:

$$D = -2\sum_{i=1}^{n}\left[y_i \ln\left(\frac{\hat{\pi}_i}{y_i}\right) + (1-y_i)\ln\left(\frac{1-\hat{\pi}_i}{1-y_i}\right)\right] \tag{4.13}$$

Bağımsız değişkenin anlamlı olup olmadığını anlamak için, ilgili bağımsız değişkenin dahil olduğu ve olmadığı eşitliklerin D değerleri birbiriyle karşılaştırılır. Bağımsız değişkenin modelde olmasından kaynaklanan, D istatistiğindeki değişim (4.14)’teki gibi elde edilir.

$$G = D(\text{değişkensiz model}) - D(\text{değişkenli model}) \tag{4.14}$$

G istatistiğini hesaplamak için farkı alınacak olan D değerlerinin yukarıda belirtilmiş her iki durumu için de doymuş modelin olabilirlikleri aynı olduğundan, G istatistiği (4.15)’te gösterildiği gibi olur.

$$G = -2\ln\left[\frac{(\text{değişkensiz modelin olabilirliği})}{(\text{değişkenli modelin olabilirlik})}\right] \tag{4.15}$$

G istatistiği tüm bağımsız değişkenlerin anlamlılığının sınanmasında kullanılabilmektedir. $\mathrm{H}_0 : \beta_i = 0$ hipotezi test edilirken 1 serbestlik dereceli ki-kare dağılımıyla karşılaştırılarak karar verilir (Kleinbaum ve Klein 2002).

Lojistik regresyon modelleri bağımlı değişkenin yapısına göre üçe ayrılmaktadır. Bu modellerden ilki olan İkili Lojistik Regresyon Modeli, kategorik bağımlı değişkenin iki durumlu (örn: var-yok) olduğu durumda kullanılmaktadır. Çoklu Lojistik Regresyon modeli ise kategorik bağımlı değişkenin çok kategorili ve sınıflayıcı ölçekle ölçülmüş (örn: medeni durum: evli – bekar - boşanmış) olduğu durumlarda kullanılmaktadır. Bağımlı değişken çok kategorili ve kategorileri arasında sıralı bir yapı söz konusu ise (örn: likert tipi ölçekler, az-orta-çok) Sıralı Lojistik Regresyon modelleri kullanılmaktadır (Kaşko 2007).

Yukarıda sözü geçen modellerden biri olan ikili lojistik regresyon bağımlı değişkenin düzeylerinin iki veya iki değere indirgenmiş olduğu durumlarda kullanılmaktadır. Bağımsız değişken ya da değişkenlerle iki düzeyli bağımlı değişkenin kategorileri arasındaki ilişkiyi göstermektedir. Bağımlı değişken, binom değişkeni gibi var - yok, olumlu - olumsuz, başarılı - başarısız gibi değerler almaktadır. Başarılı (olumlu, var) değeri 1 ile başarısız (olumsuz, yok) değeri 0 ile kodlanmaktadır. Bu yöntemde, bağımsız değişkenler iki grupta incelenmektedir. İkili lojistik regresyon modelinde ele alınan açıklayıcı değişkenler kategorik değişken ise faktör değişken, sürekli değişken ise de ortak (covariate) değişken olarak ifade edilmektedirler (Ayhan 2006).

İkili lojistik regresyon modeli;

$$P\left(Y_j=1\right)=\frac{e^{\beta_0+\beta_1 x_{j1}+......+\beta_k x_{jk}}}{1+e^{\beta_0+\beta_1 x_{j1}+......+\beta_k x_{jk}}} \tag{4.16}$$

olarak ifade edilmektedir. Burada; $P\left(Y_j=1\right)$ j.birimin 1. kategoride olması veya 1. kategoriyi seçme olasılığıdır.

Çoklu lojistik regresyon yöntemi; bağımlı değişken sınıflayıcı olduğunda ve 3 düzeyden (kategoriden) oluştuğunda kullanılan bir tekniktir. Sınıflayıcı bağımlı değişkenin kategorileri doğal bir sıraya sahip değildir. Bu yüzden kodlama yapılırken kategoriler için doğal bir sıra izlemek şart değildir.

Sınıflayıcı lojistik regresyon modeli, ikili lojistik regresyon modelinin genelleştirilmiş halidir. Bu modelde ikiden fazla kategori olduğu için ilk veya son kategori referans kategorisi olarak belirlenir. Belirlenen referans kategorisinin diğer kategorilerle ilişkisini incelemek için olasılık kurallarından yararlanılır.

Sınıflayıcı lojistik regresyon modeli;

$$\mathrm{P}\left(Y_j=h\right)=\frac{e^{\beta_{h0}+\beta_{h1}\mathrm{X}_{j1}+...+\beta_{hk}\mathrm{X}_{jk}}}{1+\sum_{h=1}^{M-1} e^{\beta_0+\beta_1\mathrm{X}_{j1}+...+\beta_k\mathrm{X}_{jk}}} \tag{4.17}$$

olarak ifade edilmektedir. Burada $\mathrm{P}\left(Y_j = h\right)$ j. birimin h. kategoride olma olasılığıdır.

$$\mathrm{P}\left(Y_j = h_0\right) = \frac{1}{1 + \sum_{h=1}^{M-1} e^{\beta_0 + \beta_1 \mathrm{X}_{j1} + \ldots + \beta_k \mathrm{X}_{jk}}} \tag{4.18}$$

Burada $\mathrm{P}\left(Y_j = h_0\right)$ j. birimin referans kategorisinde olma olasılığıdır. Ayrıca k bağımsız değişken sayısını ve M ise bağımlı değişkenin sahip olduğu düzey sayısını ifade etmektedir. (Ayhan 2006).

Bağımlı değişkenin düzeyleri sıralı ölçekle ölçüldüğünde kullanılan tekniğe ise sıralı lojistik regresyon adı verilmektedir. Kategoriler arasında doğal bir sıra vardır (küçük<orta<büyük) ve bu sıra göze alınarak kodlama yapılmalıdır. Gelir düzeyleri (düşük gelir, orta gelir, yüksek gelir), başarı düzeyleri (düşük, orta, yüksek), hastalığın şiddeti (az şiddetli, orta şiddetli, çok şiddetli) sıralı ölçekle ölçülmüş değişkenlere örnek olarak verilebilir. Sıralı lojistik regresyon modeli, gözlemlenebilir bir kategorik değişkenin altında gözlemlenemeyen bir gizli değişkenin olduğu varsayımına dayandırılmaktadır. Sıralı lojistik regresyon modeli genel olarak (4.19)'daki gibi gösterilir.

$$link(\gamma_j) = \tau_j - \sum \beta_k' x_k \tag{4.19}$$

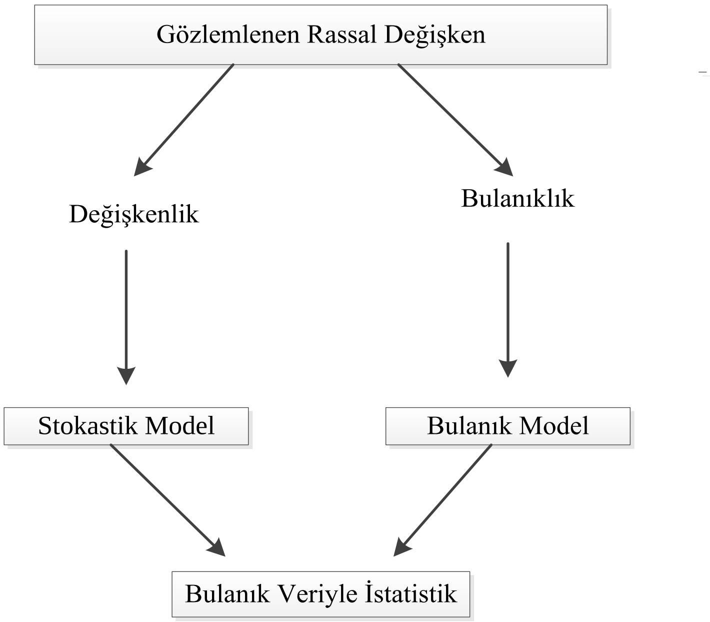
Ancak açıklayıcı değişkenlerin farklı değerleri (açıklayıcı değişken kategorik ise) bağımlı değişkenin farklı kategorilerinde daha yüksek oranda yer alıyorsa (4.20) ile gösterilen genelleştirilmiş sıralı lojistik regresyon modeli kullanılır:

$$link\left(\gamma_j\right) = \frac{\tau_j - \left[\beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k\right]}{\exp(\theta_0 + \theta_1 z_1 + \ldots + \theta_l z_l +)} \tag{4.20}$$

Burada, γ_j , j. kategori için birikimli olasılık değeri, τ_j j. kategorinin eşik değeri, $\beta_1, \ldots, \beta_k$ regresyon katsayıları, yer parametreleri için $x_1, \ldots, x_k$ açıklayıcı değişkenler ve k açıklayıcı değişken sayısıdır. β ve θ bilinmeyen yer ve ölçek parametreleri vektörüdür. Ayrıca τ_j bilinmeyen kesme noktaları vektörü ve ölçek parametreleri z_l için açıklayıcı değişkenlerdir (Ayhan 2006).

Kategorik verilerin analizi için önemli bir model olan lojistik regresyon modelinin çok geniş uygulama alanı vardır ve bu alan giderek artmaktadır. İlk başlarda sadece biyomedikal çalışmalarda kullanılsa da son 20 yıldır sosyal bilimler araştırmaları ve pazarlama alanlarında çoğunlukla kullanıldığı görülmektedir. Son dönemlerde lojistik regresyon işletme uygulamalarında popüler bir araç haline gelmiştir. Bazı kredi değerlendirme kuruluşları kendilerine gelen başvuruların kredilendirmeye değer olup olmadığını lojistik regresyon modelini kullanarak belirlemektedir (Agresti 2002).

Lojistik regresyonun kullanım alanlarından olan sosyal bilimler ve pazarlama alanlarındaki anket çalışmaları insan duygu ve düşüncesini içermektedir. İnsan duygu ve düşüncesi bulanık bir yapıya sahiptir. Benzer şekilde tıp çalışmalarında da hastalığın hastaya verdiği ağrı (acı) ölçülürken de “az, orta, çok” gibi dilsel terimler kullanılmaktadır (Pourahmad ve ark. 2011a). Bu gibi dilsel terimler kişiden kişiye değişiklik gösterdiği için bulanık küme olarak görülebilirler. Ayrıca sayısal ölçekle ölçülmelerine rağmen (1:az, 2:orta gibi) bu bulanık kümeler arasındaki sınır kesin değildir. Örneğin; bir kişinin diabet hastası olup olmadığının belirlenmesinde kullanılan oral glukoz tolerans testinde, yüksek kan şekerini belirlemek için kullanılan kesim noktası iki saatlik plazma (kan) glukozu için 140 mg/dl’dir. Bu değer kesin bir sınır değildir. Hastalığa bağlı olarak sınırın komşuluğundaki vakalar bulanık olabilir. Bu tür gözlemler arasındaki ilişkiyi modellemek için kesin gözlemlere ve varsayımlara dayalı istatistiksel yöntemler yetersiz kalmaktadır. Kategorik bağımlı değişkenle, bağımsız değişkenler arasındaki ilişkiyi modelleyen lojistik regresyon yönteminin pratikte kullanılamadığı pek çok durum vardır. Bu durumlara bir örnek verilmesi gerekirse, lojistik regresyonun en çok kullanıldığı alanlardan olan tıp alanındaki kimi araştırmalarda bazı hastalıklar için iyi tanımlanmış ve tamamiyle kabul edilmiş bir kriterin olmayışı araştırmacıların bulanık verilerle karşılaşmalarına neden olmaktadır (Pourahmad ve ark. 2011a). Gözlemlenen rassal değişken belirsizlik (uncertainty) kaynaklarından ya değişkenliği (variability) ya da bulanıklığı (fuzzy) içerebilir. Söz konusu rassal değişken sadece belirsizlik içeriyorsa bulanık modellerle, sadece değişkenlik içeriyorsa stokastik modellerle tanımlanabilir. Ancak bu rassal değişken belirsizlik kaynaklarından her ikisini de içeriyorsa, bu durumda bulanık istatistik modelleriyle tanımlanabilir. Bu durum Şekil 4.2 ile ifade edilmiştir.



Şekil 4.2. Değişkenlik ve Bulanıklık (Viertl 2011)

Bulanık mantık ve istatistik teorileri birbirini tamamlayan teorilerdir. Klasik istatistiksel yöntemlerin varsayımlarının sağlanamadığı ya da istatistiksel yöntemlerin yetersiz kaldığı durumlarda, bu iki teorinin birleşmesinden doğan yöntemler verileri çözmede ve modellemede başarılı olmaktadırlar. Bahsedilen nedenlerden dolayı, bulanık regresyon ve lojistik regresyonun analizinin birleşmesiyle "Bulanık Lojistik Regresyon Analizi" ortaya çıkmıştır. Bulanık lojistik regresyon analizi, klasik lojistik regresyon analizinin varsayımlarının sağlanamadığı ve bulanık verinin ortaya çıktığı durumlarda kullanılan ve kategorik bağımlı değişkeni modelleyen bir bulanık yaklaşımdır.

Bulanık lojistik regresyon yaklaşımı ilk olarak Nagar ve Srivastava (2008) tarafından "Adaptive Fuzzy Regression Model for the Prediction of Dichotomous Response Variables Using Cancer Data: A Case Study" adlı çalışmayla ortaya atılmıştır. Pourahmad ve ark. (2011a) ise bulanık lojistik regresyona yeni bir yaklaşım getirmişlerdir. Bu tezde önce Nagar ve Srivastava'nın Bulanık Lojistik Regresyonundan daha sonra Pourahmad ve arkadaşları tarafından geliştirilmiş olan Bulanık Lojistik Regresyon yaklaşımları anlatılacaktır.

4.1. Nagar ve Srivastava'nın Bulanık Lojistik Regresyonu

Nagar ve Srivastava (2008) çalışmalarında var olan tahminleme tekniklerinin istatistiksel teknikler ve Yapay Sinir Ağları (ANN), Destek Vektör Makinesi (SVM), Bulanık Mantık, k – en yakın komşular (k- NN), Bulanık Sinir Ağları (FNN), Bulanık Regresyon gibi yapay zeka tekniklerini içerdiğinden bahsetmişlerdir.

Ayrıca bu tekniklerin bazı kısıtlama ve dezavantajları olduğundan ve ikili değişkene sahip verilerden anlamlı sonuçlar çıkarabilmek için yeni bir model olması gerektiğinden bahsetmişlerdir ve aşağıdaki sorulara cevap aramışlardır.

- Yapay sinir ağları tekniklerini kullanarak, tahmin tutarlılığını nasıl arttırabiliriz?
- Bağımlı bağımsız değişkenler arasındaki belirsiz ilişkileri modellemek için hangi teknik kullanılabilir?
- İkili sonuç değişkeninin tahmini için nasıl bir model önerilebilir?
- Çok değişkenli bir ortamda bağımlı değişken ve bağımsız değişkenler arasındaki doğrusal olmayan ilişkiyi nasıl analiz edebiliriz?

Bu soruların bir sonucu olarak, bulanık mantık konsepti ile istatistiksel lojistik regresyonun birleşiminden bulanık lojistik regresyon ortaya çıkmıştır.

Bulanık lojistik regresyon modeli Tanaka'nın bulanık regresyon modeline dayanmaktadır. Bilindiği gibi Tanaka'nın regresyon modeli, doğrusal fonksiyonlara uygulanabilmektedir. İkili sonuç değişkeninin doğrusallık varsayımını bozmasından dolayı bir dönüşüm yapılması gerekmektedir. Bulanık lojistik regresyona ulaşabilmek için önerilen algoritma aşağıda gösterilmiştir (Nagar ve Srivastava 2008).

Adım 1: Regresyon fonksiyonunun doğrusallığının kontrolü

Adım 2: Doğrusal dönüşüm

Adım 3: Bağımsız verilerin işlenmesi (hazırlığı)

Adım 4: Doğrusal programlama yaklaşımı ile analizin yapılması

Adım 5: Yapılan dönüşüm ile sonuçları yorumlamak

4.2. Pourahmad ve Arkadaşlarının Tanaka'nın Bulanık Regresyon Modeline Dayalı Bulanık Lojistik Regresyon

Lojistik regresyon analizi, diğer istatistiksel yöntemler gibi bazı varsayımlara dayalıdır ve bu varsayımlar bazen uygulamada bazı problemlerin ortaya çıkmasına neden olmaktadır. Örneğin, belirli risk faktörlerine (bağımsız değişkenlere) dayalı klinik çalışmalarında uygun araç gereç eksikliğinden ya da iyi tanımlanmış bir kriter olmamasından dolayı doktorlar teşhislerinden şüphe edebilmektedirler. Böylece hastayı iki gruptan birine atayamamaktadırlar. Bu durumdaki ikili yanıt değişkeninin gözlemleri belirsizdir ve değişkenler arasındaki ilişki klasik lojistik regresyonun uygulanabilmesi için yeterince kesin değildir. Bağımlı değişkendeki belirsizlikten dolayı bireyin 1 kategorisine ait olması olasılığı $\left(p = P\left(Y = 1\right)\right)$ ve odds oranı $\left(p/\left(1-p\right)\right)$ hesaplanamaz. Genellikle, böylesi belirsiz gözlemler, lojistik regresyon analizinde modele dahil edilmez. Eğer bütün gözlemler belirsiz ise, başarı olasılığı ile bağımsız değişkenler arasındaki ilişki klasik istatistiksel yaklaşımlarla modellenemez. Bu tür belirsizlik, rassallık ve olasılıktan kaynaklanmamaktadır. Böylesi gözlemleri modelleyebilmek için olabilirlik (possibility) gibi belirsizliğin başka bir boyutlarını göz önüne almak gerekmektedir. Bir uzmana danışarak, her bir gözlem (vaka, durum) 1. kategorinin önceki elemanlarından kabul edilmiş kriterle karşılaştırılırsa ve bu kategoriye ait olması tutarlılık derecesi not edilirse, yeni bir istatistiksel terim olan "posibilistik odds" tanımlanmış ve modellenmiş olur.

Eğer her belirsiz (bulanık) gözlem için bilinen karakteristiğe ait tutarlılık derecesi μ_i ile ve tümleyeni $1-\mu_i$ $\left(0 \le 1-\mu_i \le 1\right)$ ile gösterilirse, ilgili bulanık gözlemin belirli bir özelliğe sahip olması olabilirliğinin olmamasına oranını gösteren $\mu_i / (1-\mu_i)$ "posibilistik odds" olarak adlandırılır (Pourahmad ve ark. 2011a). Bulanık lojistik regresyon analizinin uygulanabilmesi için gerekli olan posibilistik oddsun tanımı verildiğine göre artık bulanık lojistik regresyon modeli elde edilebilir. Bulanık lojistik regresyon modelinin nasıl elde edileceği ilerleyen bölümlerde gösterilmiştir.

4.2.1. Bulanık lojistik regresyon modelinin elde edilmesi

$X_i = (x_{i1}, x_{i2}, ..., x_{in}),\ i = 1,2,...,\mathrm{m}$ şeklinde bir veri seti göz önüne alınsın. X_i i. birim için yaş, cinsiyet, ağırlık gibi bağımsız değişkenlerden oluşan gözlemlerin vektörü olsun. μ_i, i. birimin ilgili bağımlı değişken için istenilen özelliğe sahip olması olabilirliğini göstersin. Yani, $\mu_i = Pos(Y_i = 1)$'dir. Böylece, bulanık lojistik regresyon şu şekilde tanımlanmış olur.

$$\tilde{W}_i = \tilde{b}_0 + \tilde{b}_1 x_{i1} + ... + \tilde{b}_n x_{in} \quad i = 1,2,...,\mathrm{m} \tag{4.21}$$

Buradaki $\tilde{b}_0, \tilde{b}_1, ..., \tilde{b}_n$ üçgen bulanık sayı olan model parametreleri ve $\tilde{W}_i = \ln\left(\mu_i / (1-\mu_i)\right)$ ise posibilistik oddsların logaritmik dönüşümlerinin tahmincisidir. Hesaplamayı kolaylaştırmak için $\tilde{b}_j = \left(b_j^c, s_j^L, s_j^R\right)_T$, $j = 1,2,...,n$ üçgen bulanık sayı olarak ele alınacaktır. Model parametreleri üçgen bulanık sayı olduğundan, bulanık çıktı olan $\tilde{W}_i$'da üçgen bulanık sayıdır ve $\tilde{W}_i = \left(f_i^c(x), f_{is}^L(x), f_{is}^R(x)\right)$ biçiminde gösterilmektedir. Buradaki;

$$\begin{aligned} f_i^c(x) &= a_0^c + a_1^c x_{i1} + ... + a_n^c x_{in}, \\ f_{is}^L(x) &= s_0^L + s_1^L x_{i1} + ... + s_n^L x_{in}, \\ f_{is}^R(x) &= s_0^R + s_1^R x_{i1} + ... + s_n^R x_{in}, \end{aligned} \tag{4.22}$$

şeklindedir. Kestirilmiş bulanık çıktının üyelik fonksiyonu (4.23)'teki gibi gösterilebilir.

$$
\tilde{W}_i\left(w_i\right)=\begin{cases} 1-\dfrac{f_i^c(x)-w_i}{f_{is}^L(x)}, & f_i^c(x)-f_{is}^L(x)\le w_i \le f_i^c(x) \\ 1-\dfrac{w_i-f_i^c(x)}{f_{is}^R(x)}, & f_i^c(x)< w_i \le f_i^c(x)+f_{is}^R(x) \end{cases} \tag{4.23}
$$

Eğer $s_i^L = s_i^R = s_i$ olursa, bu üçgen bulanık sayıya simetrik üçgen bulanık sayı denir. Bu durumda $\tilde{W}_i$ için $f_{is}^L(x) = f_{is}^R(x) = f_i^s(x)$ elde edilmiş olur.

Bilindiği gibi, $\tilde{W}_i$ i. birimin ilgili özelliğe sahip olmasının posibilistik oddsunun doğal logaritmasıdır. Model parametreleri tahmin edildikten sonra, genişleme prensibinden yararlanarak $\exp\left(\tilde{W}_i(x)\right)$, $x>0$ posibilistik odsunun üyelik fonksiyonunu (4.24)'te gösterildiği gibi elde edebiliriz.

$$
\exp\left(\tilde{W}_i(x)\right)=\tilde{W}_i\left(\ln\left(x\right)\right)=\begin{cases} 1-\dfrac{f_i^c(x)-\ln\left(x\right)}{f_{is}^L(x)}, & f_i^c(x)-f_{is}^L(x)\le \ln\left(x\right)\le f_i^c(x) \\ 1-\dfrac{\ln\left(x\right)-f_i^c(x)}{f_{is}^R(x)}, & f_i^c(x)<\ln\left(x\right)\le f_i^c(x)+f_{is}^R(x) \end{cases}
$$

(4.24)

Eşitlik (4.24)'den yararlanarak, yeni gelen bulanık bir gözlemin posibilistik oddsu bulunabilir (Pourahmad ve ark. 2011a).

4.2.2. Bulanık lojistik regresyon katsayılarının tahmini

$\tilde{b}_j$, $j=0,1,2,...,n$ bulanık katsayılarının tahmini için Tanaka'nın bulanık regresyon yönteminden yararlanılacaktır. Bulanık katsayıların tahmininde uygulanan temel düşünce elde edilen modelin bulanıklığının minimize edilmesidir. Bulanık katsayıların elde edilmesinde aşağıdaki varsayımlar kullanılmaktadır:

- Her gözlem w_i, $\tilde{W}_i$'de en az h kadar üyelik derecesine sahip olmalıdır. Yani, $\tilde{W}_i(w_i) \geq h$ olmalıdır. Buradaki w_i, $w_i = \ln\left(\frac{\mu_i}{1-\mu_i}\right)$ şeklindedir.
- Bulanık katsayılar model bulanıklığını minimize edecek şekilde olmalıdır. Bulanık sayıların bulanıklığı yayılımlarının artmasıyla artmaktadır. Bu yüzden, bulanık çıktıların yayılımlarının minimize edilmesiyle, modelin bulanıklığının minimize edilmiş olur.

Bulanık katsayıların belirlenebilmesi için amaç fonksiyonu (4.25) ile gösterilen doğrusal programlama probleminin çözülmesi gerekmektedir.

$$Z = m\left(s_0^L + s_0^R\right) + \sum_{j=1}^{n}\left[\left(s_j^L + s_j^R\right)\sum_{i=1}^{m} x_{ij}\right] \tag{4.25}$$

Yukarıdaki amaç fonksiyonunun kısıtları ise;

$$\begin{aligned} &1 - \frac{f_i^c(x) - w_i}{f_{is}^L(x)} \geq h \Rightarrow (1-h)s_0^L + (1-h)\sum_{j=1}^{n} s_j^L x_{ij} - a_0^c - \sum_{j=1}^{n} a_j^c x_{ij} \geq -w_i \\ &1 - \frac{w_i - f_i^c(x)}{f_{is}^R(x)} \geq h \Rightarrow (1-h)s_0^R + (1-h)\sum_{j=1}^{n} s_j^R x_{ij} + a_0^c + \sum_{j=1}^{n} a_j^c x_{ij} \geq w_i \end{aligned} \tag{4.26}$$

şeklindedir. Bulanık lojistik regresyon katsayılarının elde edilebilmesi için (4.25) ile gösterilen amaç fonksiyonu (4.26) kısıtları altında minimize edilmelidir (Pourahmad ve ark. 2011a).

4.2.3. Bulanık lojistik regresyon için uyum iyiliği kriterleri

Diğer istatistiksel yöntemlerde olduğu gibi, bulanık kurallara dayalı modellerinde bazı kriterlerle değerlendirilmesi gerekmektedir. Diğer bir ifadeyle oluşturulan modelin veri setine uyup uymadığının kontrol edilmesi gerekmektedir.

Literatürde pek çok uyum iyiliği kriteri bulunmasına rağmen bu çalışmada bulanık lojistik regresyon modeli için önerilmiş olan uyum iyiliği testlerinden

"Ortalama Üyelik Derecesi (Mean Degree of Memberships) ve "Hata Kareler Ortalaması (Mean of Squares Errors)"ından bahsedilecektir.

w_i i. birey için bağımlı değişkenin gözlemlenen değerini ve $\tilde{W}_i$ ise kestirilen değerini gösterdiğinde "Ortalama Üyelik Derecesi" (4.27) 'deki formül yardımıyla hesaplanabilir.

$$MDM = \frac{1}{m}\sum_{i=1}^{m}\tilde{W}_i\left(w_i\right) = \frac{1}{m}\exp\left(\tilde{W}_i\left(\frac{\mu_i}{1-\mu_i}\right)\right) \tag{4.27}$$

MDM kriterinin maksimum değeri 1 ve minimum değeri ise 0'dır. Ortalama üyelik derecesi 0 ile 1 arasında değerler almakta ve ortalama üyelik derecesinin 1'e yakın değerleri oluşturulan modelin veri setine uyumunun iyi olduğu göstermektedir.

Modelin uyum iyiliğini değerlendirmede kullanılan bir diğer ölçüt ise "Ortalama Hata Kareler" değeridir ve şu şekilde hesaplanabilir.

$$MSE = \frac{1}{m}\sum_{i=1}^{m}\left[def\left(\exp\left(\tilde{W}_i\right)\right) - \left(\frac{\mu_i}{1-\mu_i}\right)\right] \tag{4.28}$$

Eşitlik (4.28)'de görülen $def\left(\exp\left(\tilde{W}_i\right)\right.$ ifadesi, $\exp\left(\tilde{W}_i\right)$'nın durulaştırılmasını ifade etmektedir (Pourahmad ve ark. 2011a).

4.3. Tanaka'nın Revize Bulanık Regresyon Modeline Dayalı Bulanık Lojistik Regresyon

Pourahmad ve ark. (2011a) tarafından önerilmiş olan Bulanık Lojistik Regresyon analizinde, modelin parametrelerinin tahmininde Tanaka ve ark. (1982) tarafından önerilmiş olan bulanık regresyon modeli kullanılmaktadır.

Daha önceden belirtildiği gibi He ve ark. (2007) tarafından yapılan çalışmada, Tanaka'nın bulanık regresyon modelindeki h_i değerinin sadece kestirilen $\tilde{y}_i$ 'lerin yayılımlarına bağlı olmadığı, ayrıca $\tilde{y}_i$'nin merkezi ile gözlemlenen y_i arasındaki uzaklığa da bağlı olduğun farkedilmiştir. Bu yüzden Tanaka regresyon modelindeki amaç fonksiyonunu geliştirerek yeni bir model önerilmiştir. Önerilen yaklaşımda, Tanaka modeline göre sistem bulanıklığının düşüş sağladığı ve sistemin ortalama uyumluluk derecesinin ($\bar{h}$'nın) ise arttığı görülmüştür. Bu sebepten dolayı, Pourahmad ve ark. (2011a) tarafından önerilmiş olan bulanık lojistik regresyon modelinin katsayılarının tahmini için Tanaka yaklaşımına göre daha iyi sonuçlar veren bu yaklaşımın kullanılması literatürde ilk defa bu çalışmada önerilmiştir.

Böylece, aşağıda gösterilen bulanık lojistik regresyon modelinin bilinmeyen parametrelerinin tahmini yeni bir yaklaşım kullanılarak gerçekleştirilecektir.

$$\tilde{W}_i = \tilde{b}_0 + \tilde{b}_1 x_{i1} + \ldots + \tilde{b}_n x_{in} \quad i = 1,2,\ldots,\mathrm{m} \tag{4.29}$$

Eşitlik (4.29) ile gösterilen bulanık lojistik regresyon modelinin parametreleri (4.30) ile gösterilen amaç fonksiyonu ve (4.31) ile gösterilen kısıtlar yardımıyla elde edilebilir.

$$\mathrm{Min:}\ \ \mathrm{SF}+\sum_{i=1}^{m}\mathrm{d}_i = \sum_{i=1}^{m}\left(s_0 + \sum_{j=1}^{n} s_j \left|x_{ji}\right|\right) + \sum_{i}^{m} d_i \tag{4.30}$$

$$\begin{aligned} &0 \le h_i = 1 - d_i \Big/ \left(s_0 + \sum_{j=1}^{m} s_j \left|x_{ji}\right|\right) \\ &d_i = \left|y_i - \left(c_0 + \sum_{j=1}^{m} c_j x_{ji}\right)\right| \quad i = 1,2,\ldots,\mathrm{m} \\ &s_j \ge 0, \quad j = 0,1,2,\ldots,\mathrm{n} \end{aligned} \tag{4.31}$$

Söz konusu model uyum iyiliği kriteri olarak $\bar{h}$'yı kullanmaktadır. Pourahmad ve ark. (2011a) tarafından önerilen uyum iyiliği kriteri olan MDM aslında $\bar{h}$ ile aynı şeyi ifade etmektedir. Bu sebeple bu çalışmada uyum iyiliği kriteri olarak, MDM kriteri kullanılacaktır.

4.4. En Küçük Karelere Dayalı Bulanık Lojistik Regresyon

Klasik en küçük kareler yaklaşımının bulanık sayılar uzayında tanımlanmış bir uzaklık ölçümüne dayalı olarak genişletilmesiyle "Bulanık En Küçük Kareler" yaklaşımı elde edilmiş olur. Bulanık en küçük kareler yaklaşımında kullanılabilecek pek çok uzaklık ölçümü vardır. Bu ölçümlerden bir tanesi Xu ve Li (2001) tarafından önerilmiştir ve aşağıdaki eşitlikte gösterilmiştir.

E bulanık sayılar uzayını göstermek üzere keyfi $u,v \in E$ için $f(\alpha)$ fonksiyonuna bağlı u ve v arasındaki uzaklık;

$$d(u,v) = \left[\int_0^1 f(\alpha) d^2\left((u)_\alpha,(v)_\alpha\right) d\alpha\right]^{\frac{1}{2}} \tag{4.32}$$

ile gösterilmektedir. Eşitlik (4.32)'taki $d^2\left((u)_\alpha,(v)_\alpha\right)$ ise (4.33)'da gösterildiği gibidir.

$$d^2\left((u)_\alpha,(v)_\alpha\right) = \left[a_1(\alpha)-b_1(\alpha)\right]^2 + \left[a_2(\alpha)-b_2(\alpha)\right]^2 \tag{4.33}$$

Buradaki $(u)_\alpha,(v)_\alpha$ ise $(u)_\alpha = \left[a_1(\alpha),a_2(\alpha)\right]$ ve $(v)_\alpha = \left[b_1(\alpha),\mathrm{b}_2(\alpha)\right]$ şeklinde gösterilen u ve v'nin α kesimleridir.

Bulanık lojistik regresyonun katsayılarını en küçük karelerle tahmin edebilmek için $\tilde{w}_i$ ve $\tilde{W}_i$ arasındaki hata kareler ortalamasının (SSE) minimize edilmesi gerekmektedir. Eşitlik (4.32) ile gösterilen d uzaklığı kullanılarak

$$SSE = \sum_{i=1}^{m}\left(d\left(\tilde{w}_i,\tilde{W}_i\right)\right)^2 \tag{4.34}$$

$$d\left(\tilde{w}_i,\tilde{W}_i\right) = \left[\int_0^1 f(\alpha) d^2\left((\tilde{w}_i)_\alpha,(\tilde{W}_i)_\alpha\right) d\alpha\right]^{\frac{1}{2}} \tag{4.35}$$

elde edilmiş olur.

$$\tilde{W}_i = \ln\frac{\mu_i}{1-\mu_i} = \tilde{A}_0 + \tilde{A}_1 x_{i1} + \ldots + \tilde{A}_n x_{in} \quad i = 1,2,\ldots,\mathrm{m} \tag{4.36}$$

Eşitlik (4.36) ile gösterilen modelde ki bulanık parametre $A_j = \left(a_j, s_j\right)_T$ şeklinde olduğu varsayılırsa, kestirilmiş bulanık çıktı $\tilde{W}_i$, $\tilde{W}_i = \left(f(a), f(s)\right)$ şeklinde gösterilen simetrik üçgen bulanık sayı olur ve $f_i(a) = a_0 + a_1 x_{i1} + \ldots + a_n x_{in}$, $f_i(s) = s_0 + s_1 x_{i1} + \ldots + s_n x_{in}$ şeklinde gösterilir. Bulanık en küçük karelere dayalı parametreleri elde edebilmek için bulanık çıktı $\tilde{W}_i$ ve $\tilde{w}_i$ 'nin α kesimlerinin hesaplanması gerekmektedir. İlgili işlemler yapılarak aşağıdaki eşitlikler elde edilir.

$$\left(\tilde{W}_i\right)_\alpha = \left[(\alpha-1) f_i(s) + f_i(a), (1-\alpha) f_i(s) + f_i(a)\right] \tag{4.37}$$

$(\mu_i)_\alpha = [b_1, b_2]$ ‘ya bağlı olarak $(\tilde{w}_i)_\alpha$ ’yı hesaplarsak

$$(\tilde{w}_i)_\alpha = \left[\ln\frac{b_1}{1-b_1}, \ln\frac{b_2}{1-b_2}\right] \tag{4.38}$$

elde edilmiş olur. Böylece;

$$d^2\left((\tilde{w}_i)_\alpha, (\tilde{W}_i)_\alpha\right) = \left[\ln\frac{b_1}{1-b_1} - (\alpha-1) f_i(s) - f_i(a)\right]^2 + \left[\ln\frac{b_2}{1-b_2} - (1-\alpha) f_i(s) - f_i(a)\right]^2 \tag{4.39}$$

Eşitlik (4.39)’yı, (4.34)’de yerine koyarsak;

$$SSE=\sum_{i=1}^{m}\int_0^1 f(\alpha)\begin{bmatrix}\left[\ln\frac{b_1}{1-b_1}-(\alpha-1)f_i(s)-f_i(a)\right]^2+\\ \left[\ln\frac{b_2}{1-b_2}-(1-\alpha)f_i(s)-f_i(a)\right]^2\end{bmatrix}d\alpha \tag{4.40}$$

Parametreleri elde edebilmek için (4.40)'ın a_j ve s_j 'ye göre kısmi türevlerinin alınıp sıfıra eşitlenmesi gerekmektedir. Yani $\frac{\partial SSE}{\partial a_j}=0$ ve $\frac{\partial SSE}{\partial s_j}=0$ olmalıdır. Gerekli işlemler yapıldıktan sonra;

$$\begin{aligned}&\sum_{i=1}^{m}\left(\int_0^1 2\alpha x_{ij}\cdot\left[2f_i(a)-\ln\frac{b_1}{1-b_1}-\ln\frac{b_2}{1-b_2}\right]d\alpha\right)=0, && j=0,1,\dots,n\\ &\sum_{i=1}^{m}\left(\int_0^1 2\alpha(1-\alpha)x_{ij}\cdot\left[2(1-\alpha)f_i(s)+\ln\frac{b_1}{1-b_1}-\ln\frac{b_2}{1-b_2}\right]d\alpha\right)=0, && j=0,1,\dots,n\end{aligned} \tag{4.41}$$

Eşitlik (4.41)'teki nicelikler, bir uzman tarafından belirlenmiş olan başarı olabilirliğine (μ_i) bağlıdır. Gerekli dönüşümler yapılır ve integraller hesaplanırsa aşağıdaki eşitlikler elde edilmiş olur.

$$\begin{aligned}a_0\sum_{i=1}^{m}x_{i0}x_{ij}+a_1\sum_{i=1}^{m}x_{i1}x_{ij}+\dots+a_n\sum_{i=1}^{m}x_{in}x_{ij}&=\sum_{i=1}^{m}z_ix_{ij},\; j=0,1,\dots,n\\ s_0\sum_{i=1}^{m}x_{i0}x_{ij}+s_1\sum_{i=1}^{m}x_{i1}x_{ij}+\dots+s_n\sum_{i=1}^{m}x_{in}x_{ij}&=\sum_{i=1}^{m}k_ix_{ij},\; j=0,1,\dots,n\end{aligned} \tag{4.42}$$

Burada ki $x_{i0}=1$ ve, z_i, k_i her bir birim için integral hesaplarının sonucunu göstermektedir. Eşitlik (4.42) matris formatında gösterilecek olursa, $Aa=Z$ ve $As=K$ olur. Buradaki;

$$A = X'X, \qquad X = \begin{pmatrix} 1 & x_{11} & \dots & x_{1n} \\ 1 & x_{21} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ 1 & x_{m1} & \dots & x_{mn} \end{pmatrix}_{m \times (n+1)} \tag{4.43}$$

$$a = (a_0, a_1, \dots, a_n)^T, \qquad Z = \left(\sum_{i=1}^{m} z_i x_{i0}, \sum_{i=1}^{m} z_i x_{i1}, \dots, \sum_{i=1}^{m} z_i x_{in}, \right)^T \tag{4.44}$$

$$s = (s_0, s_1, \dots, s_n)^T, \qquad K = \left(\sum_{i=1}^{m} k_i x_{i0}, \sum_{i=1}^{m} k_i x_{i1}, \dots, \sum_{i=1}^{m} k_i x_{in}, \right)^T \tag{4.45}$$

şeklindedir. Eğer $Rank(X) = n+1$ ise o halde A matrisine pozitif tanımlı matris denir ve A^{-1} hesaplanabilir. Yani $A^{-1}K \geq 0$ ise, minimizasyon probleminin tek çözümü vardır ve (4.46) ‘teki gibi hesaplanabilir (Pourahmad ve ark. 2011b)

$$a = A^{-1}Z, \;\; s = A^{-1}K \tag{4.46}$$

5. UYGULAMA

Bu tez çalışmasının uygulama bölümünde, teorik yapısı daha önce verilen bulanık lojistik regresyon analizinin doğum ağırlıkları veri setine uygulanmasına yer verilmiştir. İlgili veriler için bulanık lojistik regresyon katsayılarının tahmini öncelikle Pourahmad ve ark. (2011a) tarafından önerilen yöntemle gerçekleştirilmiştir. Sonrasında ise bu tez çalışmasında ilk defa önerilen Tanaka'nın Revize Modeline Dayalı Bulanık Lojistik Regresyon ile bilinmeyen parametrelerinin tahmini yapılmıştır. İki modelden elde edilen parametreler için uyum iyiliği kriterleri hesaplanmış ve sonuçlar yorumlanmaya çalışılmıştır. Son aşamada ise önerilen yaklaşımın geçerliliğini göstermek adına Pourahmad ve ark. (2011a)'nın çalışmasında kullanılan veri seti, Tanaka'nın revize modeline dayalı bulanık lojistik regresyon ile modellenmiştir.

Çalışmanın veri setinde ele alınacak olan bebeklerin doğum ağırlığını etkileyen pek çok faktör bulunmaktadır. Kırımi ve Pençe (1999), yaptıkları çalışmada hamilelik süresince sigara kullanımının doğum ağırlığını etkisini araştırmışlar ve sigara içen annelerin bebeklerinin doğum ağırlıklarının daha düşük olduğunu belirtmişlerdir. Hirve ve Ganatra (1994), ise düşük doğum ağırlığını etkileyen faktörleri araştırmış ve annenin sosyo ekonomik durumunun, annelik yaşının, hamilelik süresince kanamanın, annenin boyunun doğum ağırlığına etkisi olduğunu belirtmişlerdir. Bircan (2004), doğum ağırlığını etkileyen faktörleri lojistik regresyon analizi ile incelemiş ve bebeğin cinsi, doğum haftası, annenin beslenme şekli ve annenin boyunun doğum ağırlığını etkileyen faktörler olduğu belirlenmiştir.

Bu bilgiler ışığında bu çalışmada ele alınan bağımsız değişkenler bebeklerin doğum ağırlığı, annenin yaşı, annenin beslenme alışkanlığı ve sigara kullanımı olarak belirlenmiştir. Bağımlı değişken ise ele alınan bağımsız değişkenlere göre bebeğin doğum ağırlığının normal doğum ağırlığına sahip olup - olmadığıdır.

Bebeklerin doğum ağırlığının belirlenmesinde klasik lojistik regresyon analizinin uygulanabilir olduğu düşünülebilir. Ancak, ele alınan verilerin doğası gereği bağımsız değişkenler arasındaki ilişkiden kaynaklı bir bulanıklık söz konusudur. Bu durumda ele alınan veri seti modellemek için bulanık lojistik regresyon yönteminin kullanılması gerekmektedir.

Bu çalışmada dördüncü bölümde belirtildiği gibi bulanık lojistik regresyon analizinin uygulanabilmesi için posibilistik oddsların hesaplanması gerekmektedir. Bu amaçla ele alınan bağımsız değişkenleri değerlendirebilmek için kadın hastalıkları ve doğum uzmanı ile görüşülmüş ve bağımlı değişkenin olabilirliğinin belirlenebilmesi adına uzman görüşü alınmıştır.

Analizde kullanılan 25 birimlik örneklem için 2013 yılına ilişkin değerler Kayseri ilinde bulunan bir hastanenin veri bankasından elde edilmiştir. Ele alınan değişkenleri tanımlarsak;

X_1 : Annenin yaşını göstermektedir ve sürekli bir değişkendir.

X_2 : Bebeğin doğum ağırlığını göstermekte olup, sürekli bir değişkendir.

X_3 : Annenin hamilelik süresince sigara kullanımını gösteren kategorik bir değişkendir. Sigara içmeyen anneler 0, sigara içen anneler 1 ile kodlanmıştır.

X_4 : Hamilelik süresince annenin beslenmesine dikkat edip etmediğini gösteren kategorik bir değişkendir. Beslenmesine dikkat eden anneler 1 ile etmeyenler ise 0 ile kodlanmıştır.

İlgili değişkenler ve bu değişkenlere ait posibilistik oddslar Çizelge 5.1.'de gösterilmiştir.

Çizelge 5.1. Doğum ağırlıklarına ilişkin veriler

No:	X_1 (yıl)	X_1 (gr)	X_2	X_3	μ_i	$\mu_i / 1-\mu_i$	$W_i = ln\left(\frac{\mu_i}{1-\mu_i}\right)$
1	21	2400	1	0	0,03	0,0309	-3,4761
2	22	3200	1	0	0,05	0,0526	-2,9444
3	23	3400	0	0	0,08	0,0870	-2,4424
4	27	3300	1	1	0,10	0,1111	-2,1972
5	27	3500	1	1	0,12	0,1364	-1,9924
6	22	3500	1	0	0,14	0,1628	-1,8153
7	28	3600	0	0	0,16	0,1905	-1,6582
8	28	3600	1	1	0,19	0,2346	-1,4500
9	32	3500	1	1	0,20	0,2500	-1,3863
10	35	3500	1	0	0,22	0,2821	-1,2657
11	25	3500	1	0	0,25	0,3333	-1,0986
12	27	3800	0	0	0,28	0,3889	-0,9445
13	18	3800	1	0	0,30	0,4286	-0,8473
14	22	3750	1	1	0,33	0,4925	-0,7082
15	31	4000	0	0	0,38	0,6129	-0,4896
16	27	3500	1	0	0,42	0,7241	-0,3228
17	38	4000	1	0	0,45	0,8182	-0,2007
18	21	4000	0	1	0,48	0,9231	-0,0804
19	28	4000	1	0	0,52	1,0833	0,0800
20	29	4250	1	1	0,59	1,4390	0,3640
21	20	4400	1	0	0,75	3,0000	1,0986
22	28	4500	1	0	0,79	3,7619	1,3249
23	27	4500	0	0	0,81	4,2632	1,4500
24	30	5000	1	0	0,91	10,1111	2,3136
25	25	5500	1	0	0,95	19,0000	2,9444

Uygulamada kullanılan bulanık lojistik regresyon modeli, girdisi kesin çıktısı bulanık olan bir modeldir ve teorik olarak aşağıda gösterildiği gibi ifade edilmektedir.

$$\tilde{W}_i = \tilde{b}_0 + \tilde{b}_1 x_{i1} + \tilde{b}_2 x_{i2} + \tilde{b}_3 x_{i3} + \tilde{b}_4 x_{i4} \quad i = 1,2,\ldots,25 \tag{5.1}$$

$$\tilde{b}_0 = \left(b_0^c, \mathrm{s}_0\right), \tilde{b}_1 = \left(b_1^c, \mathrm{s}_1\right), \ldots, \tilde{b}_4 = \left(b_4^c, \mathrm{s}_4\right) \tag{5.2}$$

Eşitlik (5.1)'de gösterilen bulanık regresyon modelinin parametrelerinin Pourahmad ve ark. (2011a) tarafından önerilmiş olan yaklaşıma göre tahmini için (5.3)'deki amaç fonksiyonunun (5.4)'deki kısıtlara göre çözülmesi gerekmektedir.

$$Z = m\left(s_0^L + s_0^R\right) + \sum_{j=1}^{n}\left[\left(s_j^L + s_j^R\right)\sum_{i=1}^{m} x_{ij}\right] \tag{5.3}$$

$$1-\frac{f_i^c(x)-w_i}{f_{is}^L(x)} \ge h \Rightarrow (1-h)s_0^L + (1-h)\sum_{j=1}^{n} s_j^L x_{ij} - a_0^c - \sum_{j=1}^{n} a_j^c x_{ij} \ge -w_i$$
$$1-\frac{w_i - f_i^c(x)}{f_{is}^R(x)} \ge h \Rightarrow (1-h)s_0^R + (1-h)\sum_{j=1}^{n} s_j^R x_{ij} + a_0^c + \sum_{j=1}^{n} a_j^c x_{ij} \ge w_i$$
(5.4)

Çizelge 5.1’de gösterilmiş olan veriler için (5.3)‘deki amaç fonksiyonu aşağıdaki gibi hesaplanabilir.

$$\begin{aligned} Z &= 2\left(25s_0 + s_1\sum_{i=1}^{25} X_{1i} + s_2\sum_{i=1}^{25} X_{2i} + s_3\sum_{i=1}^{25} X_{3i} + s_4\sum_{i=1}^{25} X_{4i}\right) \\ &= 2\left(25s_0 + 661s_1 + 96000s_2 + 18s_3 + 7s_4\right) \end{aligned} \quad (5.5)$$

Yukarıda gösterilen amaç fonksiyonunun 50 kısıt altında (25 gözlem x 2) çözülmesi gerekmektedir. Eşitlik (5.4)’teki kısıtlar dikkatli incelenirse “h” teriminin varlığı göze çarpmaktadır. h teriminin analiz yapılmadan önce karar verici tarafından belirlendiğinden daha önce bahsedilmişti. Çizelge 5.1’deki ilk ve son veri için $h = 0,5$ alındığında kısıtlar aşağıda gösterildiği gibi elde edilebilir.

İlk veri için;

$$\begin{aligned} &c_0 + 21\times c_1 + 2400\times c_2 + 1\times c_3 + 0\times c_4 - 0.5\times s_0 - 0.5\times 21\times s_1 - 0.5\times 2400\times s_2 \\ &-0.5\times 1\times s_3 + 0.5\times 0\times s_4 \le -3,4761 \end{aligned} \quad (5.6)$$

$$\begin{aligned} &c_0 + 21\times c_1 + 2400\times c_2 + 1\times c_3 + 0\times c_4 + 0.5\times s_0 + 0.5\times 21\times s_1 + 0.5\times 2400\times s_2 \\ &+0.5\times 1\times s_3 + 0.5\times 0\times s_4 \ge -3,4761 \end{aligned} \quad (5.7)$$

Son veri için;

$$\begin{aligned} &c_0 + 25\times c_1 + 5500\times c_2 + 1\times c_3 + 0\times c_4 - 0.5\times s_0 - 0.5\times 25\times s_1 - 0.5\times 5500\times s_2 \\ &-0.5\times 1\times s_3 - 0.5\times 0\times s_4 \le 2,9444 \end{aligned} \quad (5.8)$$

$$c_0 + 25\times c_1 + 5500\times c_2 + 1\times c_3 + 0\times c_4 + 0.5\times s_0 + 0.5\times 25\times s_1 + 0.5\times 5500\times s_2$$
$$+0.5\times 1\times s_3 + 0.5\times 0\times s_4 \geq 2,9444 \tag{5.9}$$

Bulanık lojistik regresyon modelinin parametrelerinin tahmini "Lingo 14.0" yazılımı ile yapılmış ve c bulanık sayıların merkezini ve s yayılımını göstermek üzere izleyen sonuçlar elde edilmiştir: $c_0 = 0,0000$, $s_0 = 2,1674$, $c_1 = -0,1652$, $s_1 = 0,0556$, $c_2 = 0,0009$, $s_2 = 0,00$, $c_3 = -0,1000$, $s_3 = 0,8275$, $c_4 = 0,3395$, $s_4 = 0,00$

$h = 0,5$ alındığında sistem bulanıklığı $Z = 211,63$ olarak bulunmuş ve model aşağıdaki gibi elde edilmiştir.

$$\begin{aligned}\tilde{W}_i &= \left(0.0000, 2.1674\right) + \left(-0.1652, 0.0556\right) X_1 + 0,0009 X_2 \\ &+ \left(-0.1000, 0.8275\right) X_3 + 0,3395 X_4\end{aligned} \tag{5.10}$$

Eşitlik (5.10) yardımıyla Çizelge 5.1.'deki 8. gözlemin posibilistik odds oranını hesaplayabiliriz.

$$\begin{aligned}\tilde{W}_7 &= \left(0.0000, 2.1674\right) + \left(-0.1652, 0.0556\right)\times 28 + 0,0009\times 3600 \\ &+ \left(-0.1000, 0.8275\right)\times 1 + 0,3395\times 1 = \left(0.0000, 2.1674\right) + \left(-4.6256, 1.5568\right) \\ &+ \left(3.2400, 0.0000\right) + \left(-0.1000, 0.8275\right) + \left(0.3395, 0.0000\right) = \left(-1.1461, 4,5517\right)\end{aligned} \tag{5.11}$$

Yukarıdaki denklemin sonucu elde edildikten sonra genişleme prensibi yardımıyla 8. gözlemin üyelik fonksiyonu aşağıdaki gibi hesaplanabilir.

$$\exp\left(\tilde{W}_7(x)\right)=\begin{cases}1-\dfrac{-1{,}15-\ln(x)}{4{,}55}, & -5{,}7\le \ln(x)\le -1{,}15\ \left(0{,}0033\le x\le 0{,}32\right)\\ 1-\dfrac{\ln(x)+1{,}15}{4{,}55}, & -1.15<\ln(x)\le 3{,}41\ \left(0{,}32\le x\le 30{,}13\right)\end{cases}$$

(5.12)

8. gözlem için tahmin edilen posibilistik odds oranının değeri 0,32 olarak bulunmuştur. Bu değer, bebeğin doğum ağırlığının normal olması olabilirliğinin olmaması olabilirliğine oranını göstermektedir. İlgili ters dönüşümler yapıldığında, bu bebeğin doğum ağırlığının normal doğum ağırlığına sahip olduğu 0,24 olabilirlikle söylenebilmektedir.

Bulanık lojistik regresyonun bir katkısı da yeni gözlemler için posibilistik odds oranlarının hesaplanabilmesidir. Örneğin $\left(x_1=24,\ x_2=4500, x_3=1, x_4=0\right)$ şeklinde yeni bir birey geldiğini varsayalım. Eşitlik (5.11) ile gösterilen modele göre bu yeni bireyin posibilistik oddsu şu şekilde hesaplanabilir.

$$\tilde{W}_{yeni}=\left(0.0000,2.1674\right)+\left(-0.1652,0.0556\right)\times 24+0{,}0009\times 4500$$
$$+\left(-0.1000,0.8275\right)\times 1+0{,}3395\times 0=\left(0.0000,2.1674\right)+\left(-3.9648,1.3344\right)$$
$$+\left(4.0500,0.0000\right)+\left(-0.1000,0.8275\right)+\left(0.0000,0.0000\right)=\left(-0.0148,4{,}3293\right)$$

(5.13)

Eşitlik (5.13) ile elde edilen $\tilde{W}_{yeni}$ değeri işlemleri kolaylaştırmak için yaklaşık olarak $\tilde{W}_{yeni}\cong(-0.015,4.33)$ alınırsa yeni gözlemin posibilistik oddsu (5.14) ile gösterildiği gibi elde edilebilir.

$$\exp\left(\tilde{W}_{yeni}(x)\right)=\begin{cases}1-\dfrac{-0{,}015-\ln(x)}{4{,}33}, & -4.345\le \ln(x)\le -0{,}015\ \left(0{,}013\le x\le 0{,}98\right)\\ 1-\dfrac{\ln(x)+0{,}015}{4{,}33}, & -0.015<\ln(x)\le 4{,}315\ \left(0{,}013\le x\le 74{,}81\right)\end{cases}$$

(5.14)

Böylece yeni gelen bireyin posibilistik odds değerininn 0,98 olduğu görülmektedir. Posibilistik odds hesaplandıktan sonra ters dönüşüm yapılırsa, ilgili değişkenler için bebeğin doğum ağırlığının normal doğum ağırlığına sahip olduğu 0,50 olabilirlikle söylenebilmektedir.

Bir önceki bölümde anlatılan bulanık lojistik regresyon uyum iyiliği kriterlerinden biri olan "Ortalama Üyelik Derecesi (MDM)" klasik regresyondaki belirlilik katsayısına $\left(R^2\right)$ benzer bir kriter olup, oluşturulan modelin veri setine uyumunun bir göstergesidir. Yukarıda oluşturulan model için "Ortalama Üyelik Derecesi" aşağıdaki gibi hesaplanmıştır.

$$MDM = \frac{1}{m}\sum_{i=1}^{m}\tilde{W}_i\left(w_i\right) = \frac{1}{25}\times 17,96 \Rightarrow MDM = 0,72 \qquad (5.15)$$

Ortalama üyelik derecesi 0,72 olarak bulunmuştur. Bu değer, oluşturulan bulanık lojistik regresyon modelinin ele alınan veri setini modellemede iyi olduğu şeklinde yorumlanabilir. Diğer bir ifadeyle, MDM kriteri bağımlı değişkenin bulanık ortamda bağımsız değişkenler tarafından 0,72 oranında açıklandığı göstermektedir.

Pourahmad ve ark. (2011a) tarafından önerilen yaklaşıma göre bulanık lojistik regresyon katsayıları hesaplandıktan sonra, bu tez çalışmasının asıl uygulama konusu olan ve bu tezde önerilen yaklaşım olan Tanaka'nın Revize modeliyle bulanık lojistik regresyon denkleminin katsayı tahminleri, ilgili veri seti kullanılarak gerçekleştirilecektir. Eşitlik (5.16) ile gösterilen bulanık lojistik regresyon modelinin katsayılarının tahmini için (5.17)'teki amaç fonksiyonunun (5.18)'teki kısıtlara göre çözülmesi gerekmektedir.

$$\tilde{W}_i = \tilde{b}_0 + \tilde{b}_1 x_{i1} + \tilde{b}_2 x_{i2} + \tilde{b}_3 x_{i3} + \tilde{b}_4 x_{i4} \quad i = 1,2,\ldots,25 \qquad (5.16)$$

$$Z = \mathrm{SF} + \mathrm{d}_i = \sum_{i=1}^{n}\left(s_0 + \sum_{j=1}^{m} s_j\left|x_{ji}\right|\right) + \sum_{i}^{n} d_i \qquad (5.17)$$

$$0 \le h_i = {}^{1-d_i}\Big/\left(s_0 + \sum_{j=1}^{m} s_j \left|x_{ji}\right|\right)$$

$$d_i = \left| y_i - \left(c_0 + \sum_{j=1}^{m} c_j x_{ji}\right)\right| \quad i = 1,2,\ldots,n \tag{5.18}$$

Eşitlik (5.17)'teki amaç fonksiyonu aşağıdaki gibi elde edilebilir.

$$\begin{aligned} Z &= 2\left(25s_0 + s_1\sum_{i=1}^{25} X_{1i} + s_2\sum_{i=1}^{25} X_{2i} + s_3\sum_{i=1}^{25} X_{3i} + s_4\sum_{i=1}^{25} X_{7i}\right) + \sum_{i=1}^{25} d_i \\ &= 2\left(25s_0 + 661s_1 + 96000s_2 + 18s_3 + 7s_4\right) + d_1 + d_2 + \ldots + d_{25} \end{aligned} \tag{5.19}$$

Yukarıdaki amaç fonksiyonunun veri setindeki ilk ve son veri için kısıtları (5.20) ve (5.21) ile gösterildiği gibidir.

İlk veri için;

$$\begin{aligned} &d_1(s_0 + 21s_1 + 2400s_2 + s_3) \le 0.5 \\ &d_1 = \left|-3.4761 - c_0 - 21c_1 - 2400c_2 - c_3\right| \end{aligned} \tag{5.20}$$

Son veri için;

$$\begin{aligned} &d_1(s_0 + 25s_1 + 5500s_2 + s_3) \le 0.5 \\ &d_1 = \left|2.9444 - c_0 - 25c_1 - 5500c_2 - c_3\right| \end{aligned} \tag{5.21}$$

Lingo 14 yazılımı yardımıyla elde edilen katsayılar şu şekildedir: $c_0 = 0,0000$ $s_0 = 2,3599$, $c_1 = -0,0713$, $s_1 = 0,0000$, $c_2 = 0,0004$, $s_2 = 0,0006$, $c_3 = -0,8202$, $s_3 = 0,4898$, $c_4 = 0,0672$, $s_4 = 0,0000$

Eşitlik (5.22) ile gösterilen bulanık lojistik regresyon modelinin $h = 0,5$ alındığında sistem bulanıklığı $Z = 288,053$ olarak elde edilmiştir.

$$\begin{aligned} \tilde{W}_i &= \left(0.0000, 2,3599\right) - 0.0713X_1 + \left(0.0004, 0,0006\right) X_2 \\ &+ \left(-0.8202, 0.4898\right) X_3 + 0,0672X_4 \end{aligned} \tag{5.22}$$

Bulanık lojistik regresyon parametrelerinin tahmini yapıldıktan sonra oluşturulan modelin veri setine uyumunun ölçüsü olan uyum iyiliği kriterinin hesabına geçilebilir. Tanaka'nın revize modeline dayalı bulanık lojistik regresyon modeli için Ortalama üyelik derecesi (MDM) kriteri aşağıda gösterildiği gibi hesaplanmıştır.

$$MDM = \frac{1}{m}\sum_{i=1}^{m}\tilde{W}_i\left(w_i\right) = \frac{1}{25}\times 19,47 \Rightarrow MDM = 0,78 \tag{5.23}$$

Bulanık lojistik regresyon analizinde parametre tahmini için önerilen yaklaşımla elde edilen MDM kriterinin değeri 0,78 olarak bulunmuştur. MDM kriteri bağımlı değişkenin bulanık ortamda bağımsız değişkenler tarafından 0,78 oranında açıklandığı göstermektedir. Diğer bir deyişle, oluşturulan bulanık lojistik regresyon modelinin ele alınan doğum ağırlığı veri setini modellemede iyi olduğu şeklinde yorumlanabilir.

Farklı h değerleri için hem Pourahmad ve ark. (2011a) önerdiği, hem de bu tezde önerilen yaklaşımla bulanık lojistik regresyon katsayıları tahmin edilmiştir.

Model I: Pourahmad ve arkadaşlarının önerdiği modeli, Model II: bu tezde önerilen yaklaşımı göstermek üzere $h = 0,5$ ve $h = 0,7$ değerleri için iki modelle elde edilen katsayılar, sistem bulanıklığı ve bu modellere ait MDM kriterinin değerleri sırasıyla Çizelge 5.2, Çizelge 5.3. ve Çizelge 5.4. ile gösterilmiştir.

Çizelge 5.2. $h = 0,5$ değeri için model katsayıları ve sistem bulanıklığı

	(c_0, s_0)	(c_1, s_1)	(c_2, s_2)	(c_3, s_3)	(c_4, s_4)	**Z**
Model I	(0.00,2.17)	(-0.16,0.05)	(-0.009,0.00)	(-0.10,0.83)	(0.339,0.00)	211,6
Model II	(0.00,2.359)	(-0.07,0.00)	(0.0004,0.0006)	(-0.82,0.49)	(0.672,0.00)	288,7

Çizelge 5.3. $h=0,7$ değeri için model katsayıları ve sistem bulanıklığı

	(c_0, s_0)	(c_1, s_1)	(c_2, s_2)	(c_3, s_3)	(c_4, s_4)	**Z**
Model I	(0.00,1.548)	(-0.16, 0.039)	(0.0009,0.00)	(-0.1,0.59)	(0.339,0.00)	151,1
Model II	(0.00,1.38)	(-0.073, 0.00)	(0.0003, 0,0004)	(-0.63,0.69)	(0.056,0.00)	214,6

Çizelge 5.4. $h=0,5$ ve $h=0,7$ için MDM kriterinin değerleri

	$h = 0,5$	$h = 0,7$
Model I	0,72	0,61
Model II	0,78	0,72

Çizelge 5.4 incelendiğinde bu tezde önerilen yaklaşım için hesaplanmış olan MDM kriteri değerinin her iki h seviyesinde de daha yüksek olduğu görülmektedir. Bulanık ortamda bağımlı değişkenin bağımsız değişkenler tarafından hangi ölçüde açıklandığını gösteren bir uyum iyiliği kriteri olan MDM'ye göre Tanaka'nın Revize Modeline Dayalı Bulanık Lojistik regresyon yaklaşımının daha iyi sonuçlar verdiği görülmüştür.

Tanaka'nın revize modeline dayalı bulanık lojistik regresyon yaklaşımın geçerliliğini göstermek adına, bu model Pourahmad ve ark. (2011a) tarafından yapılan çalışmada kullanılan şeker hastalığı verilerine uygulanmıştır. Çalışmada kullanılan değişkenler aşağıda gösterildiği gibidir.

X_1 : Cinsiyeti göstermektedir ve kadınlar 1 ile erkekler 0 ile kodlanmıştır.

X_2 : 2 saatlik tokluk kan şekerini (mg/dl) göstermektedir.

X_3 : Bireylerin yaşını göstermektedir.

X_4 : Vücut kitle indeksini $\left(\text{kg/m}^2\right)$ göstermektedir.

X_5 : Aile geçmişini göstermektedir. Anne, baba, kardeşlerinden birinde şeker hastalığı olan bireyler 1 ile olmayanlar 0 ile kodlanmıştır.

Çalışmada bulanık lojistik regresyon modelinin bilinmeyen parametreleri Tanaka'nın bulanık regresyon modelin dayanarak hesaplanmıştır. h=0,6 için elde edilen model (5.24)'de gösterildiği gibidir.

$$\tilde{W}_i = -15.8833 + (0.4884, 0.5841) X_1 + 0.0927 X_2 + 0.0727 X_3$$
$$(-0.1142, 0.0194) X_4 + (0.4940, 1.1305) X_5 \quad (5.24)$$

Yukarıda gösterilen modelin uyum iyiliğinin göstergesi olarak MDM kriteri hesaplanmış ve bu değer 0,70 olarak bulunmuştur.

Pourahmad ve ark. (2011a) tarafından yapılan çalışmada kullanılan veriler bu tezde önerilmiş olan katsayı tahmin yaklaşımıyla h=0,6 için yeniden hesaplanmış ve elde edilen model aşağıda gösterilmiştir.

$$\tilde{W}_i = -15.985 + (0.4087, 0.6438) X_1 + 0.0919 X_2 + 0.0689 X_3$$
$$(-0.1000, 0.0203) X_4 + (0.5556, 1.037) X_5 \quad (5.25)$$

Yukarıda oluşturulmuş bulanık lojistik regresyon için uyum iyiliği kriteri olan MDM hesaplanmış ve 0,73 olarak bulunmuştur. Elde edilen modelin bu veri seti için de daha iyi sonuç verdiği görülmektedir.

6. SONUÇ VE ÖNERİLER

Bilimsel araştırmalarda incelenen olaylar matematiksel modellerle tanımlanabilmektedir. İstatistikte modelleme amacıyla kullanılabilecek yöntemlerden birisi de regresyon analizidir. Regresyon analizi, ele alınan bağımlı değişken ve bağımsız değişkenler arasındaki ilişkiyi matematiksel olarak modellemektedir. Eğer ele alınan bağımlı değişken kategorik ise, bağımlı ve bağımsız değişkenler arasındaki ilişkiyi matematiksel olarak modellemenin yolu lojistik regresyon analizidir. Lojistik regresyon analizinin etkin olarak kullanıldığı alanlar ise tıp ve sosyal bilimlerdir.

Sosyal bilimlerde yapılan çalışmalar kişilerin görüş ve değerler yargılarını içermektedir. İnsan düşünce yapısı ve algısı belirsizlik içermektedir. Benzer şekilde tıp alanında ilgilenilen veriler doğası gereği belirsizlik içerebilirler. Belirsizlik, rassallık ve bulanıklık olarak iki türlü ortaya çıkmaktadır. Bunlardan ilki olan rassallık istatistiksel yöntemlerle modellenebilir. Bulanıklık ise bulanık mantık yaklaşımı kullanılarak modellenmektedir. Bulanık mantık ve istatistik teorileri birbirlerini tamamlayan teorilerdir. İki teorinin birleşmesinden doğan yöntemler istatistik teorisindeki varsayımların sağlanamadığı ya da insan düşünce yapısından kaynaklı belirsizlikleri içeren durumlarda ele alınan verileri modellemede etkin ve başarılı olmaktadırlar.

Yukarıda bahsedilen iki teorinin birleşmesinden doğan yöntemlerden biri olan bulanık lojistik regresyon, klasik lojistik regresyon analizinin varsayımlarının sağlanamadığı ya da yapısı gereği belirsizlik içeren veriler söz konusu olduğunda kullanılan bir yöntemdir.

Bu tez çalışmasında amaçlanan Pourahmad ve ark. (2011a) tarafından bulanık lojistik regresyon modelinin bilinmeyen parametrelerinin tahmininde kullanılması için önerilmiş Tanaka'nın regresyon modeline alternatif olan, başka bir yaklaşımı literatürde ilk defa önermektir. Bu kapsam da öncelikle tezin ikinci bölümünde bulanık mantıktan, bulanık küme işlemlerinden ve olasılık teorisi ile olabilirlik teorisi arasındaki farklardan bahsedilmiştir. Bulanık küme teorisiyle ilgili bilgi verildikten sonra, bulanık küme teorisine dayalı olan bulanık regresyon yöntemleri tanıtılmıştır. Çalışmanın dördüncü bölümünde ise bu tez çalışmasının konusu olan bulanık lojistik regresyon analizinin teorik yapısına ve bu tez çalışmasında ilk defa önerilmiş olan Tanaka'nın revize modeline dayalı bulanık lojistik regresyon analizine yer verilmiştir.

Çalışmanın uygulama kısmında hem Pourahmad ve ark. (2011a) tarafından hem de bu tez çalışmasında önerilmiş olan yaklaşım doğum ağırlığı verilerine farklı h seviyelerinde uygulanmış ve sonuçlar bir uyum iyiliği kriteri olan MDM'ye göre yorumlanmaya çalışılmıştır. $h = 0,5$ için oluşturulan lojistik regresyon modelinde Pourahmad ve ark.(2011a) tarafından önerilen yöntem için MDM kriterinin değeri 0,72 ve Tanaka'nın revize modeline dayalı bulanık lojistik regresyon modeli için ise bu değer 0,78 olarak hesaplanmıştır. Benzer şekilde $h = 0,7$ değeri için bu tezde önerilen yaklaşım için MDM kriteri 0,72, diğer yöntem için ise 0,61 olarak bulunmuştur. MDM kriteri oluşturulan modelin ele alınan veri setine uyumunun ölçüsünü gösteren bir kriter olduğuna göre, bu tez çalışmasında önerilmiş olan Tanaka'nın revize modeline dayalı bulanık lojistik regresyon yaklaşımının ele alınan veri setini modellemede daha iyi olduğu söylenebilir.

Önerilen yaklaşımın geçerliliğin gösterilmesi adına Pourahmad ve ark. (2011a) tarafından yapılmış olan çalışmada kullanılmış olan veri setine Tanaka'nın revize bulanık lojistik regresyon analizi uygulanmış ve hesaplanan MDM kriterinin değerinin daha yüksek olduğu görülmüştür.

Çalışmanın ilerleyen dönemlerinde amaçlananlardan ilki ise bulanık lojistik regresyon modelinde kullanılan bulanık sayı tiplerini değiştirerek bulanık sayı tiplerinin parametrelere olan etkisini incelemektir.

Bulanık lojistik regresyon analizinin uygulanabilmesi için uzman görüşü alınması gerekmektedir. Bu zorunluluğu en aza indirmek için bir veri seti için uzman görüşü alındıktan sonra, bu uzman görüşünü sinir ağları ile eğiterek yeni gelecek veriler için uzman görüşünü tahmin etmek amaçlananlar arasındadır. Ayrıca ilerleyen çalışmalarda, bulanık lojistik regresyon analizindeki bilinmeyen parametrelerin tahmini için farklı bulanık regresyon yaklaşımlarının önerilmesi de amaçlanmaktadır.

Sonuç olarak hem doğası gereği bulanıklık içeren veri setleri hem de klasik istatistiksel yöntemlerin varsayımlarının sağlanamadığı durumlarda kullanılan yöntemlerden biri olan bulanık lojistik regresyon analizinin katsayılarının tahmini için yeni bir yaklaşım önerilmiş ve bu yaklaşımın geçerliliği karşılaştırılmıştır. Literatürde yeni olan bulanık lojistik regresyon yönteminin bilinmeyen parametrelerinin tahmini için yeni yaklaşımlar geliştirilip, bu yaklaşımların geçerliliği karşılaştırılabilir.

KAYNAKLAR

Agresti, A. (2002), *Categorical Data Analysis*, John Wiley ve Sons, New York, A.B.D.

Ayhan, S. (2006), *Sıralı Lojistik Regresyon Analiziyle Türkiye'deki Hemşirelerin İş Bırakma Niyetini Etkileyen Faktörlerin Belirlenmesi*, Yüksek Lisans Tezi, Osmangazi Üniversitesi, Fen Bilimleri Enstitüsü, Eskişehir.

Başarır, G. (1990), *Çok Değişkenli Verilerde Ayrımsama Sorunu ve Lojistik Regresyon Analizi*, Yüksek Lisans Tezi, Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, Ankara.

Baykal, N. ve Beyan, T. (2004), *Bulanık Mantık İlke ve Temelleri.* Bıçaklar Kitabevi, Ankara.

Berry-Stölzle, T. R., Koissi, M. C. ve Shapiro, A. F. (2010), "Detecting fuzzy relationships in regression models: The case of insurer solvency surveillance in Germany", *Insurance: Mathematics and Economics, 40*(3), 554-567.

Bircan, H. (2004), "Lojistik Regresyon Analizi: Tıp Verileri Üzerine Bir Uygulama", *Kocaeli Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 185-208.

Chang, N.-B., Chen, Y. L. ve Wang, S. F. (1997), "A fuzzy interval multiobjective mixed integer programming approach for the optimal planning of solid waste management systems", *Fuzzy Sets and Systems, 89*(1), 35-60.

Chang, P. T. ve Lee, E. S. (1996), "A generalized fuzzy weighted least-squares regression", *Fuzzy Sets and Systems, 82*(3), 289-298.

Chang, Y.-H. O. ve Ayyub, B. M. (2001), "Fuzzy Regression Methods- a Comparative Assessment", *Fuzzy Sets and Systems*(119), 187-203.

Chen, S. J., Hwang, C. L. ve Hwang, F. P. (1992), *Fuzzy multiple attribute decision making: methods and applications.* Springer-Verlag, Berlin, Almanya

Chou, T. Y., Liang, G. S. ve Han, T. C. (2013), "Application of fuzzy regression on air cargo volume forecast", *Quality & Quantity, 47*(2), 897-908.

Diamond, P. (1988), "Fuzzy Least Squares", *Information Sciences, 46*(3), 141-157.

Dom, R. M., Kareem, S. A., Razak, A. ve Abidin, B. (2008), "A Learning System Prediction Method Using Fuzzy Regression", *Proceedings of the International MultiConference of Engineers and Computer Scientists*, Hong Kong, 69-72.

Dubois, D.ve Prade, H. (1978), "Operation on fuzzy numbers", *International Journal of System Science, 9*(6), 613-626.

Dubois, D. ve Prade, H. (1980), *Fuzzy Sets and System: Theory and Applications.* Academic Press Inc, New York, A.B.D.

Dubois, D. ve Prade, H. M. (1988), *Possibility Theory: An Approach to Computerized Processing of Uncertainty,* Plenum Press, New York, A.B.D.

Düzyurt, S. (2008), *Bulanık Regresyon ile Tahmin ve Bir Uygulama*, Yüksek Lisans Tezi, Gazi Üniversitesi, Fen Bilimleri Enstitüsü, Ankara.

Elmas, Ç. (2003), *Bulanık mantık denetleyiciler,* Seçkin Yayıncılık, Ankara.

Ergün, E. (2011), *Bulanık Regresyon ve Yarıparametrik Toplamsal Regresyon Üzerine Bİr Örnek Uygulama Çalışması*, Yüksek Lisans Tezi, Muğla Üniversitesi, Fen Bilimleri Enstitüsü, Muğla.

Girginer, N. ve Cankuş, B. (2008), "Tramvay Yolcu Memnuniyetinin Lojistik Regresyon Analiziyle Ölçülmesi: Estram Örneği", *Yönetim ve Ekonomi, 15*(1), 181-193.

Gülcan, B. (2012), *Bulanık Doğrusal Programlama ve Bir Bisküvi İşletmesinde Optimum Ürün Fomrülü Oluşturma*, Yüksek Lisans Tezi, Karamanoğlu Mehmetbey Üniversitesi, Sosyal Bilimler Enstitüsü, Karaman.

He, Y. Q., Chan, L. K. ve Wu, M. L. (2007), "Balancing productivity and consumer satisfaction for profitability: Statistical and fuzzy regression analysis", *European Journal of Operational Resarch, 176*(1), 252-263.

Hirve, S. S. ve Ganatra, B. R. (1994), "Determinants Of Low Birth Weight: A Commun,ty Based Prospective Cohort Study", *Indian Pediatrics, 31*(10), 1221-1225.

Hojati, M., Bector, C. R. ve Smimou, K. (2005), "A simple method for computation of fuzzy linear regression", *European Journal of Operational Research, 166*(1), 172-184.

Hosmer, D. W. ve Lemeshow, S. (2000), *Applied Logistic Regression,* John Wiley ve Sons, New York, A.B.D.

Ishibuchi, H. ve Nii, M. (2001), "Fuzzy regression using asymmetric fuzzy coefficients and fuzzified neural networks", *Fuzzy Sets and Systems, 119*(2), 273–290.

Kahraman, C., Beşkese, A. ve Bozbura, F. T. (2006), "Fuzzy Applications in Industrial Engineering", *Fuzzy Regression Aprroaches and Applications*, (Ed:Kahraman C), Springer - Verlag, New York, A.B.D., 589-615

Kaşko, Y. (2007), *Çoklu Bağlantı Durumunda İkili (Binary) Lojistik Regresyon Modelinde 1. Tip Hata ve Testin Gücü*, Yüksek Lisans Tezi, Ankara Üniversitesi, Fen Bilimleri Enstitüsü, Ankara.

Kaya, H. S. (2010), *Bulanık Regresyon ve Bir Uygulama*, Yüksek Lisans Tezi, Süleyman Demirel Üniversitesi, Sosyal Bilimler Enstitüsü, Isparta.

Kırımi, E. ve Pençe, S. (1999), "Gebelikte Sigara Kullanımının Fetus ve Plasentanın Gelişimine Etkisi", *Van Tıp Dergisi, 6*(1), 28-30.

Kleinbaum, D. G. ve Klein, M. (2002), *Logistic Regression, A Self-Learning Text*, Springer - Verlag, New York, A.B.D.

Klir, G. J. ve Yuan, B. (1995), *Fuzzy Sets and Fuzzy Logic Theory and Applications.* Prentice-Hall Inc., New Jersey, A.B.D.

Le, C. T. (2003), *Introductory Biostatistics,* John Wiley ve Sons, New Jersey, A.B.D.

Lee, H. T. ve Chen, S. H. (2001), "Fuzzy regression model with fuzzy input and output data for manpower forecasting", *Fuzzy Sets and Systems, 119*(2), 205-213.

Lee, K. H. (2005), *First Course on Fuzzy Theory and Applications,* Springer - Verlag, New York, A.B.D.

Lu, J. ve Wang, R. (2009), "An enhanced fuzzy linear regression model with more flexible spreads", *Fuzzy Sets and Systems*, *160*(17), 2505 – 2523.

Modarres, M., Nasrabadi, E. ve Nasrabadi, M. M. (2005), "Fuzzy linear regression models with least square errors", *Applied Mathematics and Computation, 163*(2), 977–989.

Montgomery D.C., Peck E. A., ve Vining G.G. (2001), *Introduction to Linear Regression Analysis*, John Wiley ve Sons, New Jersey, A.B.D.

Moskowitz, H. ve Kim, K. J. (1993), "On Assesing The h Value in Fuzzy Linear Regression", *Fuzzy Sets and Systems*(58), 304.

Nagar, P. ve Srivastava, S. (2008), "Adaptive Fuzzy Regression Model for the Prediction of Dichotomous Response Variables Using Cancer Data: A Case Study", *Journal of Applied Mathematics, Statistics and Informatics(JAMSI), 4*(2), 183-191.

Namdari, M., Abadi, A., Taheri, S. M., Rezaei, M., Kalantari, N., ve Omidvar, N. (2014), "Effect of folic acid on appetite in children: Ordinal logistic and fuzzy logistic regressions", *Nutrition, 30*(3), 274–278.

Nasrabadi, M. M. ve Nasrabadi, E. (2004), "A mathematical-programming approach to fuzzy linear regression analysis", *Applied Mathematics and Computation, 155*(3), 873-881.

Özdamar, K. (2004), *Paket Programlar İle İstatistiksel Veri Analizi,* Kaan Kitabevi, Eskişehir.

Özdemir, Ö. (2013), *Dağılım ve Sinir Ağı Tabanlı Bulanık Zaman Serisi Modelleri* Doktora Tezi, Anadolu Üniversitesi, Fen Bilimleri Enstitüsü, Eskişehir.

Özelkan, E. C. ve Duckstein, L. (2000), "Multi-objective fuzzy regression: a general framework", *Computers & Operations Research, 27*(7-8), 635-652.

Özelkan, E. C. ve Duckstein, L. (2001), "Fuzzy conceptual rainfall-runoff models", *Journal of Hydrology, 253*(1-4), 41-68.

Özkan, M. M. (2003), *Bulanık Hedef Programlama,* Ekin Kitabevi, Bursa.

Peters, G. (1994), "Fuzzy linear regression with fuzzy intervals", *Fuzzy Sets and Systems, 63*(1), 43-55.

Pourahmad, S., Ayatollahi, S. M. ve Taheri, S. M. (2011a), "Fuzzy Logistic Regression: A New Possibilistic Model And Its Application In Clinical Vague Status", *Iranian Journal of Fuzzy Systems, 8*(1), 1-17.

Pourahmad, S., Ayatollahi, S. M. ve Taheri, S. M. (2011b), "Fuzzy Logistic Regression Based on The Least Squares Approach with Application in Clinical Studies", *Computers and Mathematics with Applications*(62), 3353-3365.

Redden, D. T. ve Woodall, W. H. (1996), "Further examination of fuzzy linear regression", *Fuzzy Sets and Systems, 79*(2), 203-211.

Ross, T., Booker, J. M. ve Parkinson, W. J. (2002), *Fuzzy Logic and Probability Applications: Bridging the Gap,* ASA-SIAM, Philadelphia, A.B.D.

Roychowdhury, S. ve Pedrycz, W. (2001), "A Survey of Defuzzification Strategies", *International Journey of Intelligent Systems, 16*(6), 679-695.

Shapiro, A. F. (2004), "Fuzzy Regression and the Term Structure of Interest Rates Revisited", *Proceedings of the 14th Annual International AFIR Colloquium.* Boston, 29-45.

Stahl, C. (2006), "A strong consistent least-squares estimator in a linear fuzzy regression model with fuzzy parameters and fuzzy dependent variables", *Fuzzy Sets and Systems, 157*(19), 2593 – 2607.

Şanlı, K. (2005), *Bulanık Robust Regresyon Çözümlemesi*, Doktora Tezi, Ankara Üniversitesi Fen Bilimleri Enstitüsü, Ankara.

Şen, Z. (2001), *Bulanık Mantık ve Modelleme İlkeleri.* Bilge Kültür Yayınevi, İstanbul.

Taheri, S. M. (2003), "Trends in Fuzzy Statistics", *Austrian Journal of Statistics, 32*(3), 239-257.

Tanaka, H., Uejima, S. ve Asai, K. (1982), "Lineer Regression Analysis with Fuzzy Model", *Ieee Transactions On Systems, Man, and Cybernetics, 12*(6), 903-907.

Tran, L. ve Duckstein, L. (2002), "Multiobjective fuzzy regression with central tendency and possibilistic properties", *Fuzzy Sets and Systems, 130*(1), 21-31.

Tseng, F.-M. ve Tzeng, G. H. (2002), "A fuzzy seasonal ARIMAmodel for forecasting", *Fuzzy Sets and Systems, 126*(3), 367-376.

Uçal Sarı, İ. (2012), *Yatırım Analizinde Bulanık Model Önerileri*, Doktora Tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.

Viertl, R. (2011). *Statistical Methods for Fuzzy Data.* John Wiley ve Sons, New York, A.B.D.

Wang, H. F. ve Tsaur , R. C. (2000a), "Insight of a Fuzzy Regression Model", *Fuzzy Sets and Systems*(112), 355-369.

Wang, H. F. ve Tsaur, R. C. (2000b), "Theory and Methodology Resolution of fuzzy regression model", *Resolution of fuzzy regression model*(126), 637-650.

Wu, B. ve Tseng, N.-F. (2002), "A new approach to fuzzy regression models with application to business cycle analysis", *Fuzzy Sets and Systems, 130*(1), 33-42.

Wu, H. C. (2003), "Linear Regression Analysis for Fuzzy Input and Output Data Using the Extension Principle", *Computers and Mathematics with Applications, 45*(12), 1849-1859.

Xu, R. ve Li, C. (2001), "Multidimensional least-squares fitting with a fuzzy model", *Fuzzy Sets and Systems*(119), 215–223.

Yang, M. S. ve Ko, C. H. (1996), "On a class of fuzzy c-numbers clustering procedures for fuzzy data", *Fuzzy Sets and Systems, 84*(1), 49-60.

Yang, M.-S. ve Lin, T. S. (2002a), "Fuzzy least-squares linear regression analysis for fuzzy input–output data", *Fuzzy Sets and Systems, 126*(3), 389-399.

Yang, M.-S. ve Liu, H. H. (2002b), "Fuzzy least-squares algorithms for interactive fuzzy linear regression models", *Fuzzy Sets and Systems, 135*(2), 305-316.

Yücel, L. İ. (2005), *Bulanık Regresyon: Türkiye'de 1980-2004 Döneminde Kayıt Dışı Ekonominin Bulanık Yöntemlerle Tahminine İlişkin Bir Uygulama*, Yüksek Lisans Tezi, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.

Zadeh, L. A. (1965), "Fuzzy Sets", *Information and Control, 8*, 338-353.

Zadeh, L. A. (1978), "Fuzzy Sets As a Basis for a Theory of Possibility", *Fuzzy Sets and Systems, 1*(1), 3-28.