

T.C
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

RICHARD LİNK FONKSİYONLARI
DOKTORA TEZİ

Arzu EKİNCİ DEMİRELLİ

Anabilim Dalı: İstatistik

Programı: İstatistik Teorisi

Danışman: Doç. Dr. Mehmet GÜRCAN

MAYIS-2015

T.C
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

RİCHARD LİNK FONKSİYONLARI

DOKTORA TEZİ

Arzu EKİNCİ DEMİRELLİ
(092133203)

Anabilim Dalı: İstatistik
Programı : İstatistik Teorisi

Danışmanı: Doç. Dr. Mehmet GÜRCAN

Tezin Enstitüye Verildiği Tarih: 08 Nisan 2015

MAYIS-2015

T.C
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

RİCHARD LİNK FONKSİYONLARI

DOKTORA TEZİ

Arzu EKİNCİ DEMİRELLİ
(Enstitü No: 092133203)

Tezin Enstitüye Verildiği Tarih: 08.04.2015

Tezin Savunulduğu Tarih: 11.05.2015

Tez Danışmanı:

Doç. Dr. Mehmet GÜRCAN (F.Ü.)

Diğer Jüri Üyeleri:

Doç.Dr.Yüksel Öner (O.M.Ü.)

Doç.Dr. Cemil ÇOLAK (İ.Ü.)

Doç.Dr. Sinan ÇALIK (F.Ü.)

Yrd.Doç.Dr. Nurhan HALİSDEMİR (F.Ü.)

MAYIS-2015

ÖNSÖZ

Lisans eğitimim sonrasında gerek yüksek lisans, gerek doktora çalışmalarında engin bilgi ve tecrübeleri ile bana araştırma olanağı sağlayan ve bu konuda çalışmaya yönlendiren, çalışmalarımın her aşamasında benden bilgi ve yardımlarını esirgemeyen ve verdiği önerilerle beni yönlendiren değerli hocam sayın Doç. Dr. Mehmet Gürcan'a,

Aynı şekilde tez çalışmalarım süresince her türlü zorluğu paylaştığım ve elinden gelen yardımı benden esirgemeyen değerli arkadaşım Dr. Esra Pamukçu'ya,

Yine çalışmalarım esnasında desteklerini sürekli hissettiğim değerli hocalarım sayın Yrd. Doç. Dr. Nurhan Halisdemir, sayın Doç. Dr. Mahmut Işık ve değerli bölüm başkanım sayın Doç. Dr. Sinan Çalık'a,

Hayatta tek gayeleri, tek çocukları beni her şeyden önce mutlu sonra başarılı görmek olan, hiçbir sıkıntıyı dert ettirmeyen hayattaki en büyük gücüm annem ve babama,

Üniversite yıllarından beri birlikte yol alıp, hayata birlikte tutunarak bugünlere geldiğim sevgili eşim Op. Dr. Salih Demirelli'ye, aramıza katıldığı günden beri bana mutlulukların en güzeli annelik duygusunu yaşatan oğlum İsmail Aras'a sonsuz teşekkür ederim.

Arzu EKİNCİ DEMİRELLİ
MAYIS-2015

İÇİNDEKİLER

	<u>Sayfa No</u>
ÖNSÖZ.....	I
İÇİNDEKİLER.....	II
ÖZET	III
SUMMARY.....	IV
TABLolar LİSTESİ.....	V
1. GİRİŞ.....	1
2. MATERYAL VE METOT.....	4
2.1. Gompertz Büyüme Eğrileri.....	4
2.2. Bertalanffy Büyüme Eğrileri	6
2.3. Richard Büyüme Eğrileri.....	6
2.4. Richard Eğri Ailesinin Sürekli Deformasyonu	9
2.5. Çoklu Grup Lojistik Ayrımsama.....	11
3. BULGULAR.....	13
3.1. Richard Yoğunluk Fonksiyonu	13
3.2. Alt Aralıklarda Tanımlanan Richard Dağılımı	14
3.3. Pozitif Yarı Eksende Tanımlı Richard Dağılımından Üretilen Sayılar	17
3.4. Büyüme Modellerinde Richard Link Fonksiyonunun Kullanılması.....	24
3.5. Büyüme Eğrilerinde Duraklama Anlarının Belirlenmesi	29
3.6. Büyüme Eğrisinin Duraklama Anlarına Göre Parçalanması	32
3.7. Sarmal Richard Dağılımı	33
3.8. Uygulama: Karınca Verileri ve Dağılım Fonksiyonunun Sürekli Deformasyonu Kullanılarak Dairesel Richard Dağılımına Ait Uygulama	36
4. SONUÇ VE TARTIŞMA	41
KAYNAKLAR	43
ÖZGEÇMİŞ.....	47

ÖZET

Richard link fonksiyonları temelde büyüme eğrileri ya da büyüme modelleri içerisinde yer alan uygulamalı istatistiğin önemli bir materyalidir. Çalışmamızda Richard link fonksiyonlarının genel yapısı üzerinde durularak bu yapı olasılık teorisi için önemli olan dağılım fonksiyonu özelliklerini ihtiva edecek şekilde genişletilmiştir. Bunun yanı sıra büyüme eğrilerinin temel özelliklerine Richard link fonksiyonunun sağladığı bir takım yeni bulgularda eklenmiştir.

Çalışmada elde edilen teorik sonuçlar büyüme modelleri açısından oldukça önemlidir. Özellikle literatürde günümüze kadar bir büyüme modelinde büyümenin yavaşladığı anlar model üzerinde belirlenmemektedir. Bunun nedeni bilinen modellerin sürekli bir artış göstererek büyümenin tamamlanmasından sonra stabil hale gelmesinden kaynaklanmaktadır. Çalışmada bu probleme uygun bir şekilde çözüm getirilebilmiştir.

Yine çalışmanın literatüre kazandırdığı yenilik pozitif yarı ekseninde tanımlı bir Richard dağılımının sarmal hale getirilerek dairesel verilerde kullanılmasıdır. Bu kazanım çalışmanın son bölümünde karınca verileri ile yapılan bir uygulama ile desteklenmiştir. Ayrıca buna ilaveten uygulama içerisinde Richard ailesinin sürekli deformasyonu da gerçekleştirilmiştir.

Anahtar kelimeler: Richard link fonksiyonu, Dağılım fonksiyonu, Büyüme modelleri, Sarmal dağılımlar,

SUMMARY

Richard Link Functions

Richard link functions are an important component of statistical applications in growth models or in the basis of growth curves. In this, study, by focussing on the general structure of Richard link functions, the structure is expanded in such a way as to include the features of distribution function important for possibility theory. In addition, a set of new findings provided by the Richard link function have been added to the basic features of the growth curves.

The theoretical results obtained in this study are extremely important in terms of growth models. To date in literature, the point at which growth has slowed in a growth model has not been defined in the model. The reason for this is that by showing a continuous increase, the known models are seen to become stable after growth has been completed. The current study has been able to resolve this problem in an appropriate way.

Another new aspect of this study is that by making the identified Richard distribution spiral in the positive half-axis, cyclical data was used. This benefit was supported by an application made with Ant data in the final section of the study. However, continuous deformation of the Richard family occurred within this additional application.

Key Words: Richard link function, Distribution function, Growth models, Wrapped distribution

TABLolar LİSTESİ

	<u>Sayfa No</u>
Tablo 1. $k = 1$ ve $m = 0.5$ için $Richard(1,0.5, x)$ dağılımından üretilen sayısal değerler.	18
Tablo 2. $k = 1$ ve $m = 1.5$ için $Richard(1,1.5, x)$ dağılımından üretilen sayısal değerler.	19
Tablo 3. $k = 1$ ve $m = 2.5$ için $Richard(1,2.5, x)$ dağılımından üretilen sayısal değerler.	19
Tablo 4. $k = 1$ ve $m = 3.5$ için $Richard(1,3.5, x)$ dağılımından üretilen sayısal değerler.	20
Tablo 5. $k = 1$ ve $m = 6$ için $Richard(1,6, x)$ dağılımından üretilen sayısal değerler. ...	20
Tablo 6. $k = 1$ ve $m = 10$ için $Richard(1,10, x)$ dağılımından üretilen sayısal değerler.	21
Tablo 7. $k = 1$ ve $m = 20$ için $Richard(1,20, x)$ dağılımından üretilen sayısal değerler.	21
Tablo 8. $k = 1$ ve $m = 30$ için $Richard(1,30, x)$ dağılımından üretilen sayısal değerler.	22
Tablo 9. $k = 0.2$ ve $m = 2.5$ için $Richard(0.2,2.5, x)$ dağılımından üretilen sayısal değerler.....	22
Tablo 10. $k = 0.02$ ve $m = 2.5$ için $Richard(0.02,2.5, x)$ dağılımından üretilen değerler.	23
Tablo 11. $k = 0.5$ ve $m = 2.5$ için $Richard(0.5,2.5, x)$ dağılımından üretilen değerler. .	23
Tablo 12. $k = 0.05$ ve $m = 2.5$ için $Richard(0.05,2.5, x)$ dağılımından üretilen değerler.	24
Tablo 13. Karınca verileri.....	37
Tablo 14. Sarmal dağılımın şekil parametreleri.....	38
Tablo 15. Dağılımdan tahmin edilen olasılık değerleri.....	39
Tablo 16. m değerlerinin tahmininde kullanılan R program kodları.....	40

1. GİRİŞ

İstatistik literatüründe link fonksiyonları ilk olarak lojistik regresyon analizinde kullanılmaya başlanmıştır. Regresyon analizinde bağımlı ve bağımsız değişkenler arasında var olan ilişkinin modelinin kurulması amaçlanmaktadır. Ancak bağımlı ve bağımsız değişkenlerin sayısal değişim aralıkları birbirleri cinsinden olmadığında bu iki değişkeni aynı değişim aralığında tanımlayabilmek için link fonksiyonlarından yararlanılmaktadır. Bu amaç için kullanılan bilinen link fonksiyonlarından en çok kullanılanı lojistik link fonksiyonudur. Lojistik link fonksiyonu reel düzlemin $\{(x, y): x \in R, 0 \leq y \leq 1\}$ bandında yerleştiğinden tüm reel ekseninde değer alabilen bağımsız değişken değerlerini $[0,1]$ aralığına daraltmaktadır. Lojistik link fonksiyonunun yanı sıra aynı amaçla S (sigmoidal) eğrilerin birçoğu bu amaçla kullanılabilir. Bu durumda link fonksiyonu, büyüme eğrisi ve dağılım fonksiyonu aynı türlü eğriler olmaktadır. Bu amaçla kullanılan S-eğrilerinin analizin yapısına bağlı olarak kullanışlı olabilmesi için oldukça esnek bir yapıya sahip olması gerekmektedir. Bu çalışmadaki amacımız, S-eğrileri içerisinde oldukça yüksek bir esnekliğe sahip olan Richard eğrilerinin link fonksiyonu olarak kullanılabilmesini sağlayarak Richard link fonksiyonlarının yapısal özelliklerini istatistiksel açıdan değerlendirmektir.

Richard büyüme modeli ilk olarak F. J. Richard tarafından 1959 yılında kaleme alınan “A Flexible Growth Function for Empirical Use” başlıklı çalışmada yer almış ve bu konuyla ilgili olan araştırmacıların incelemesine sunulmuştur. Daha sonraları bu konuyla ilgili literatür genişletilmiş ve başta A. Gregorczyk (1998) olmak üzere bir çok araştırmacı önemli çalışmalar yapmıştır.

Çalışmanın amacına yönelik olarak çalışmanın ilk kısımlarında lojistik regresyon analizi ve büyüme modelleri hakkında kısa tanımlayıcı bilgiler sunulacaktır.

Genel olarak büyüme eğrileri büküm noktasının olup olmamasına göre iki sınıfa ayrılabilir. Şayet büyüme eğrisi büküm noktasına sahipse S-eğrisi olup belli bir noktaya kadar artan eğime sahip olacak büküm noktasından sonra azalan bir eğim gösterecektir. Büyüme eğrisi büküm noktasına sahip değil ise başlangıçtan itibaren azalan bir eğime sahip olacaktır. Bu tip büyüme eğrilerinde başlangıçta eğrinin eğimi maksimum olup ilerleyen değerler için gittikçe sifira yaklaşır. Her iki tip büyüme eğrisi de artan değerlere karşılık üst sınırdaki bir asimptota sahip olacaktır. Ancak büküm noktasına sahip olan bir büyüme eğrisi alt sınırdaki da bir asimptota sahip olur. Büyüme eğrilerinde eğrinin belli

zaman dilimindeki değerleriyle ilgilenildiğinden analizlerde bu asimptot değerleri dikkate alınmamaktadır. Buna karşılık asimptot değerleri eğrinin büküm noktasını ve azalan eğime sahipse eğimin azalma hızını etkilediğinden oldukça önemlidir.

Lojistik regresyon modellerinin kullanılması 1845’li yıllara kadar dayanmaktadır. Lojistik analiz ilk önceleri sürekli, pozitif değer alan, artan ve sınırlı bir değişkenin zamana bağlı olarak incelenmesi amacıyla geliştirilmiştir. On dokuzuncu yüzyıl istatistikçilerinin araştırmalarında yer alan bu özelliklere sahip en önemli sosyoekonomik değişken nüfus miktarıdır. Bu bakımdan $y = y(t)$, $t > 0$ anındaki nüfus miktarı olmak üzere, nüfus artışının nüfus büyüklüğüne bağlı olacağı düşünülerek aşağıdaki diferansiyel model düzenlenmiştir,

$$\frac{dy(t)}{dt} = c y(t) \quad (1.1)$$

Burada c reel sabittir. Diferansiyel denklemin çözümü a integral sabiti olmak üzere aşağıdaki şekilde elde edilir,

$$y(t) = a \exp(c t) \quad (1.2)$$

Bulunan çözüm integral sabitinin seçimine göre bir eğri ailesi vermektedir. Bu eğri ailesi içerisinde hangi eğrinin model olarak kullanılacağına karar verilmesi aşamasında istatistiksel analizler devreye girmektedir. Bu kısım ile ilgili literatürdeki detaylı çalışma Cramer (1991)’de bulunmaktadır.

Aynı kaynakta bulunan diğer bir önemli bilgi ise, modele bir üst sınır çizgisinin eklenmesidir. Büyümenin reel bir α sabitini aşmaması durumunda diferansiyel denklem aşağıdaki şekilde düzenlenecek,

$$\frac{dy(t)}{dt} = c y(t) [\alpha - y(t)] \quad (1.3)$$

ve çözüm, $\alpha = 1$ için aşağıdaki şekilde olacaktır,

$$y(t) = \frac{\exp(c_1 + c_2 t)}{1 + \exp(c_1 + c_2 t)} \quad (1.4)$$

Burada c_1 ve c_2 reel sabitlerdir. Yukarıdaki eşitliğin tersi düşünüldüğünde aşağıdaki eşitlik elde edilir,

$$\ln \frac{y}{1-y} = c_1 + c_2 t = A(t) \quad (1.5)$$

Burada $A(t)$ doğrusal olasılık modeli olarak adlandırılır. Doğrusal olasılık modelinin ikinci türevinin sıfıra eşit olduğu noktanın bulunması önemlidir. Bu nokta doğrusal olasılık modelinin büküm noktası olup lojistik ayrımsamanın yapılabilmesi için gereklidir. Birçok doğrusal olasılık modelinde büküm noktası $y = \alpha/2$ olarak elde edilmektedir.

Yukarıda anlatılan modelle ilgili literatür on dokuzuncu yüzyıl kaynaklarında yer almaktadır. Bu yüzyılda önemli bir araştırmacı olan Quetelet kendi öğrencisi olan Verhulst ile birlikte yürüttükleri araştırmalarında diferansiyel denklemi oluşturarak lojistik inceleme anlamında ilk modeli kurmayı başarmışlardır. Bu aşamadan sonra, yarım yüzyıl kadar, uzun bir süre araştırmacıların dikkatini çekmeyen lojistik model yirminci yüzyılın başlarında Miner ve Pearl tarafından tekrar ele alınarak popülerliğini kazanmıştır. Yine aynı yüzyılda Reed, Berkson, Forrel, Aitchison ve Brown gibi araştırmacılar tarafından incelenen konu yirminci yüzyılın ortalarında neredeyse günümüzdeki konumuna kavuşturulmuştur [38]. Yirminci yüzyıl literatürünün bu konuyla ilgili son halkalarını Theil ve J. A. Anderson almaktadır. Theil lojistik model yapısını çoklu grup lojistik ayrımsamaya taşıyarak genişletmiş, Anderson ise model parametre tahminlerinin hangi şartlar altında var olacağını ispatlamıştır.

Yirmi birinci yüzyıla gelindiğinde M. Gürcan ve Y. Öner tarafından Richard link fonksiyonu artık çoklu gurup lojistik ayrımsamada da kullanılmaya başlanmıştır. 2001 yılında yapılan bir çalışmada Richard link fonksiyonu kullanılarak olabilirlik fonksiyonu oluşturulmuş ve olabilirlik fonksiyonunun hangi şartlar altında çözümünün var olduğu ispatlanmıştır [18].

Günümüz biliminde bu konu hala güncelliğini korumakta olup uygulama alanları genişledikçe bilim dünyasının ve bu konu ile ilgili araştırmacıların daha fazla ilgisini çekmektedir.

2. MATERYAL VE METOT

Diferansiyel denklem formunda verilen büyüme modelleri genel olarak aşağıdaki denkleme genelleştirilebilir,

$$\frac{dy(x)}{dx} = g(y)[h(\alpha) - h(y)] \quad (2.1)$$

Burada h ve g fonksiyonları

$$h(0) = g(0) = 0 \quad (2.2)$$

şartını sağlayacak şekilde artan fonksiyonlardır. Ayrıca $x \rightarrow \infty$ ve $x \rightarrow -\infty$ şartları altında devamdaki limit özelliği sağlanmaktadır [38],

$$y'(x) \rightarrow 0 \quad (2.3)$$

Özel olarak $g(y) = cy$ ve $h(y) = y$ olarak seçelim. Bu durumda h ve g fonksiyonları lineer fonksiyonlar olup (2.2) özelliğini sağlarlar. Bu özel durumda (2.1) denklemi doğal olarak (1.3) denklemine dönüşecektir. Bu sayede denklemin çözümü (1.4) eşitliği ile verilen lojistik model olmaktadır. Şimdi devamda h ve g fonksiyonlarının çeşitli seçimleriyle elde edilen büyüme eğrilerini ele alalım.

2.1. Gompertz Büyüme Eğrileri

Gompertz büyüme modeli ilk olarak 1825 yılında ünlü bilim adamı Benjamin Gompertz tarafından bilim literatürüne kazandırılmıştır. Benjamin Gompertz 5 mart 1779 yılında Londra'da doğmuş ve 86 yaşında 14 temmuz 1865 yılında yine Londra'da hayatını kaybetmiştir. Yaptığı en önemli çalışmalarından biri olan ölüm oranı kanunu daha sonraları Gompertz büyüme eğrileri veya büyüme modeli olarak ün kazanmıştır. Benjamin Gompertz'in ölüm oranı eşitlik (1.1) ile verilen büyüme oranı $c \in R$ olmak üzere aşağıdaki şekilde tanımlanmıştır,

$$K^c = \frac{1}{y(t)} \quad (2.4)$$

Burada K reel sabittir. Bu model daha sonraki yıllarda hatta günümüzde bile önemli olmasının en önemli faktörlerinden birisi büyümeyi gösteren esnekliğinin kullanışlı olması yani, modelin büyüme verisine uygunluğunun lojistik büyüme eğrisinden daha güçlü olmasıdır. Bu özelliğinden dolayı Gompertz büyüme modeli 1964 yılında A. K. Laird tarafından kanserli hücrelerin incelenmesinde kullanılmıştır [41]. A. K. Laird'in tümör büyüme modeli aşağıdaki şekilde oluşturulmuştur,

$$y(t) = K \exp\{\log(y(0)/K) \exp(-ct)\} \quad (2.5)$$

Dikkat edilecek olursa $t \rightarrow \infty$ için $y(t) \rightarrow K$ maksimum tümör büyüklüğü ve $t = 0$ için başlangıçtaki tümör büyüklüğü $y(0)$ olarak alınmıştır.

Şimdi büyüme modellerinin üretici olan (2.1) denkleminde h ve g fonksiyonlarını aşağıdaki şekilde seçelim,

$$h(y) = \ln y, \quad g(y) = cy \quad (2.6)$$

Bu durumda (2.1) eşitliği ile verilen diferansiyel denklem,

$$\frac{dy(t)}{y(t)} = c y(t) \ln\left(\frac{\alpha}{y(t)}\right), \quad c > 0, \alpha > 0 \quad (2.7)$$

şekline dönüşür. Denklemin çözümü aşağıdaki şekilde elde edilecektir,

$$y(t) = \alpha \exp[-C \exp(-ct)] \quad (2.8)$$

Burada $C \in R$ diferansiyel denklemin çözümünde bulunan integral sabitidir.

2.2. Bertalanffy Büyüme Eğrileri

Karl Ludwing von Bertalanffy (1901-1972) Viyana yakınlarında bulunan Atzgerdorf bölgesinde doğmuştur. Bugün bile sistem teorisinin duayenlerinden biri olarak görülen ünlü bilim adamı Bertalanffy açık sistemlerin kurucusu olarak bilinmektedir. Bertalanffy 1934 yılında yayınlanan “Büyüme Yasası Konusunda Çalışmalar” başlıklı yayınında günümüzde Bertalanffy büyüme modeli olarak adlandırılan modeli literatüre kazandırmıştır [6]. Bertalanffy büyüme modeli (2.1) eşitliğinde,

$$h(y) = y^{1-m}, \quad g(y) = (1 - m)^{-1}y^m \quad (2.9)$$

şeklinde alınacak olursa,

$$\frac{dy(t)}{y(t)} = c(1 - m)^{-1}y^m(\alpha^{1-m} - y^{1-m}) \quad (2.10)$$

diferansiyel denkleminin çözümünden aşağıdaki formda elde edilir,

$$y(t) = [\alpha^{1-m} - C \exp(-ct)]^{1/(1-m)} \quad (2.11)$$

Burada $\alpha, C, c, m \in R$ reel sabitler olmak üzere dört adet parametre bulunmaktadır. Bu sabitler aşağıdaki şekilde seçilecek olursa model lojistik modele indirgenecektir,

$$m = 2, \quad C = \frac{1}{\alpha} \quad (2.12)$$

2.3. Richard Büyüme Eğrileri

Richard eğrileri lojistik modeller içerisinde özel bir yere sahip olup bu modelde parametre sayısı fazla olduğundan lojistik ayrımsamada daha iyi ayrımsama yapabilmektedir. Richard eğrisi (2.1) ile verilen genel diferansiyel denklem yapısında,

$$h(y) = y^{m-1} \quad (2.13)$$

$$g(y) = k(m - 1)^{-1}\alpha^{1-m}y \quad (2.14)$$

şeklinde seçilecek olursa diferansiyel denklem aşağıdaki formda elde edilecektir,

$$\frac{dy(t)}{y(t)} = k(m-1)^{-1} \alpha^{1-m} y (\alpha^{m-1} - y^{m-1}) \quad (2.15)$$

Buradan denklemin çözümü olan Richard büyüme modeli,

$$y(t) = \alpha \{1 + (m-1) \exp(-kt)\}^{1/(1-m)} \quad (2.16)$$

elde edilir. Bu modelde α büyüme eğrisi için bir üst sınır ve $k > 0$ reel sabittir. Modelin değişkeni (1.5) eşitliği ile verilen doğrusal olasılık modeli ile değiştirilebilir. Bağımsız değişkenlerin lineer bir kombinasyonu olan $A(x)$ modelde kullanılacak olursa (2.14) ile verilen $g(y)$ fonksiyonu aşağıdaki şekilde seçilmelidir,

$$g(y) = \frac{A'(x)y}{(m-1)\alpha^{m-1}} \quad (2.17)$$

Bu durumda (2.15) ile verilen diferansiyel denklem devamdaki şekli alacaktır,

$$\frac{dy}{dx} = \frac{A'(x)y}{(m-1)\alpha^{m-1}} \{\alpha^{m-1} - y^{m-1}\} \quad (2.18)$$

Bu denklemin çözümü (2.16) çözümüne benzer olarak aşağıdaki şekilde elde edilebilir,

$$y(x) = \alpha \{1 + (m-1) \exp(-A(x))\}^{\frac{1}{1-m}} \quad (2.19)$$

Bu model $m = 2$ alındığında standart lojistik modele rahatlıkla dönüşecektir. Her büyüme eğrisi gibi Richard eğrisi de lojistik ayrımsama da kullanılabilir. Bunun için büyüme eğrilerinin büküm noktaları dikkate alınır. Büküm noktasının altında kalan noktalar sıfır gurubuna üstünde kalan noktalar ise bir gurubuna seçilmektedir. Richard eğrisinin büküm noktasını bulabilmemiz için ikinci türevini incelemeliyiz. Richard eğrisinin ikinci türevi aşağıdaki formdadır,

$$\frac{d^2y}{dx^2} = \frac{A'(x)}{(m-1)\alpha^{m-1}} \{\alpha^{m-1} - my^{m-1}\} \frac{dy}{dx} \quad (2.20)$$

İkinci türev sıfıra eşitlenecek olursa aşağıdaki değer bulunacaktır,

$$y_0 = \alpha m^{1/(1-m)} \quad (2.21)$$

Bulunan bu değer dikkate alındığında eğrinin büküm noktası aşağıdaki gibi bulunur,

$$x_0 = \frac{1}{A'(x)} \ln \frac{A'(x)}{m-1} \quad (2.22)$$

Dikkat edilecek olursa (2.16) modeli $t > 0$ reel değişkenine göre düzenlendiğinden zamana bağlı bir büyümeyi incelemek için tasarlanmıştır. Buna karşılık (2.19) modeli zaman parametresinden ayrılarak yapılan gözlemlerin belirlediği değişkenlere bağlı tasarlanmıştır. Bu durumda (2.19) modeli bir lojistik ayrımsama modeli olarak kullanılabilir. Bu çalışmanın devamında da Richard eğrileri ayrımsama için tasarlanmıştır. Diferansiyel denklemin çözümünden elde edilen lojistik bir eğrinin parametresi integral sabitidir. İntegral sabitinin doğrusal olasılık modeline bağlı olarak modele yerleştirilmesi yanı sıra diferansiyel denklemini oluşturan fonksiyonların seçiminde bulunması da mümkündür. Lojistik ayrımsamada kullanılan Richard link fonksiyonunda bu şekilde bulunan parametrenin hesaplanması da mümkündür. Büküm parametresi hatalı sınıflanmış gözlemler incelenerek bağımsız değişken değerleri ile ilişkilendirilerek lojistik ayrımsama dinamik bir yapıya kavuşturulabilir.

Richard link fonksiyonları kullanılarak oluşturulan ayrımsama modellerinde parametre tahmininde bulunmak bilinen standart yöntemlerle oldukça güçtür. Bu hesaplamaları yapabilmek için bilgisayar programlarına ve ileri seviyede yazılan algoritmalara ihtiyaç duyulmaktadır. Bu nedenle on dokuzuncu yüzyılın tamamında ve yirminci yüzyılın ilk yarısına kadar bu incelemeler yeteri kadar gelişmemiş ancak yirminci yüzyılın ikinci yarısında ve bundan sonra hızlı bir gelişme göstermiştir. Günümüzde geliştirilmiş olan algoritmalar sayesinde bu hesaplamalar çok rahat bir şekilde yapılabilmektedir.

Model parametreleri tahmin edilirken aralık veya nokta tahminleri kullanılabilir. Aralık tahmini kullanıldığında, parametre tahmini için belirlenen aralıkta genetik

algoritmalar kullanılarak küçük bir hata payı ile ardışık hesaplama yapan algoritmalar kullanılabilir. Bunun için gerekli olan en önemli unsur dar bir aralık tahminidir.

Genellikle çevreyle ilgili modellerde parametrelerin tahmin aralıkları oldukça güçleşmektedir. Bu ise yapılan nokta tahminlerinin modelin uyumunu güçleştirmesini sağlamaktadır. Bu gibi durumlarda kullanılabilir yöntemlerden biri Bayes çıkarımı ve Bootstrap tekniğidir. (Holmes, 2001). (Gelman et al., 1995)

2.4. Richard Eğri Ailesinin Sürekli Deformasyonu

Veri yapısına bağlı olarak link fonksiyonları ayırsamayı bazen istenilen düzeyde yapmayabilmektedir. Genellikle bu durum ilgilenilen bağımsız değişkenlerin çokluğundan ve veri yapısının dağılımından kaynaklanmaktadır. Seçilen lojistik ayırsama modeli verinin bir kısmına uymakta ancak azımsanmayacak geri kalan kısmına uymamaktadır. Bu gibi durumlarda daha esnek yapıya sahip olan link fonksiyonları tercih edilebilir. Ancak link fonksiyonunun esnekliğinin yeterli olmadığı veri gruplarında birden fazla model kullanılabilir. Bu aynı link fonksiyonu ailesi içerisinde model parametrelerinin veri gruplarına adapte edilmesi ile mümkün olabilmektedir. Bu bölümdeki amacımız Richard link fonksiyonunun şekil parametresini değiştirerek modelin veriye uyumunun artırılmasını sağlamaktır. Şekil parametresi değiştirildikçe Richard eğrisi düzlemin daraltılmış bir bandında soldan sağa doğru ötelemeyi gerçekleştirecektir.

Şimdi (2.16) eşitliğinde verilmiş olan Richard link fonksiyonu reel düzlemin aşağıda verilmiş olan kapalı bir alt aralığında yer almaktadır,

$$S = \{(x; y): x \in I, 0 \leq Y \leq \alpha\} \quad (2.23)$$

İndis kümesi olarak $I = [0,1]$ aralığını alalım. $m \neq 1$ ve $m \in R$ parametresi yardımıyla aşağıdaki değişkeni tanımlayalım,

$$\tau = \frac{1}{1+e^{-m}} \quad (2.24)$$

Tanımlanan bu değişken $\tau \in [0,1]$ indis aralığında yer alacaktır. Bu durumda Richard link fonksiyonlarının deformasyonunu aşağıda tanımlanan $T(t, \tau): R \times I \rightarrow S$ dönüşümü ile gerçekleştirebiliriz,

$$T(t, \tau) = f_\tau(t) \quad (2.25)$$

Burada $f_\tau(t)$ fonksiyonu (2.16) eşitliği ile verilen Richard link fonksiyonunun τ parametresine göre düzenlenmiş halidir.

Bu durumda $\tau = 0$ ve $\tau = 1$ için sırasıyla aşağıdaki sınır eğrileri elde edilecektir,

$$f_0(t) = y(t, \tau = 0) \quad (2.26)$$

$$f_1(t) = y(t, \tau = 1) \quad (2.27)$$

Öteleme, başlangıç eğrisi $f_0(x)$ 'den başlayarak $f_1(x)$ eğrisine kadar sürdürülebilir. Bu işlem Richard link fonksiyonları için önemli bir özelliktir. Model için aranan optimal link fonksiyonu aşağıda tanımlanan eğri ailesi içerisinde bulunabilir,

$$R = \{f_\tau(x)\}_\tau \quad (2.28)$$

Bu sonuç model arayışını doğal olarak R kümesi üzerinde bir optimizasyon problemine dönüştürmüştür. Elde edilen sonuç lojistik ayırimsama problemi açısından literatürde oldukça önemli bir bulgudur. Bu çalışma Y. Öner, M. Gürcan ve N. Halisdemir tarafından 2005 yılında gerçekleştirilmiş ve çalışmada aşağıdaki sonuçlara erişilmiştir [32],

- a. R kümesi içerisinde tüm Richard link fonksiyonları birbirlerine denktir,
- b. $0 < \tau_i < 1$ ve $\tau_1 > \tau_2 > \dots > \tau_k > \dots$ olmak üzere,

$$E_i^k = \{x : f_{\tau_i}(x) < k, k \in R\} \quad (2.29)$$

şeklinde tanımlanan kümeler ölçülebilendir ve aralarında $E_1^k \subset E_2^k \subset \dots$ artan bir ilişki mevcuttur. Reel sayıların keyfi bir aralığı A olmak üzere, A kümesi $u_i = (E_i - E_{i-1})A$ kümelerinin birleşimi olarak yazılabilir.

- c. R konveks kümedir.

2.5. Çoklu Grup Lojistik Ayrımsama

Düzlemde S-eğrisi şeklinde tanımlanan link fonksiyonlarının tamamı iki grup ayrımsama yapabilmektedir. Ayrımsamada grup sayısı ikiden fazla olduğu durumda ayrımsama çoklu grup lojistik ayrımsama problemi adını almaktadır. Çoklu grup lojistik ayrımsama problemi iki grup ayrımsamaya göre daha karmaşık bir yapıya sahiptir. Bunun nedeni link fonksiyonları içerisinde hiçbir eğrinin birden fazla büküm noktasına sahip olmamasıdır. Literatürde bu problem grupların ikişerli olarak birbirleriyle karşılaştırılması yardımıyla çözülebilmektedir. Bunun için ilk önce sabit bir referans grubu alınmalı ve seçilen referans grubuyla diğer grupların arasında ikili ayrımsama yapılarak referans grubuna dâhil olacak veriler belirlenir. Daha sonra aynı işlemler diğer gruplar arasında sürdürülerek verilerin gruplara dâhil olma olasılıkları bulunur. Verinin hangi gruba dâhil olma olasılığı daha büyükse veri o gruba ait olmalıdır.

Varsayalım ki ayrımsama için lojistik veri B_1, \dots, B_g gibi g -tane grupta yapılmak istensin. Keyfi bir y_i gözleminin B_j grubuna ait olma olasılığını aşağıdaki şekilde gösterelim,

$$a_{ij} = Pr\{y_i \in B_j\} \quad (2.30)$$

Bu durumda bağımsız değişkenlerin lineer bir kombinasyonu olan $A(x)$ her bir grup için farklı katsayılarla $A_1(x), \dots, A_g(x)$ g -tane alınarak aşağıdaki şekilde bir link fonksiyonuna eşitlenmektedir,

$$A_j(x_i) = \log(a_{ij}/a_{i1}) \quad (2.31)$$

Burada $j = 1, 2, \dots, g$ olmak üzere a_{ij} olasılıklarını maksimum yapan x_i gözlemi B_j grubuna atanmaktadır. x_i gözleminin B_j grubuna ait olabilmesi için $A_j(x_i) - A_r(x_i)$ farkı tüm $r = 1, 2, \dots, g$ ve $r \neq j$ değerleri için pozitif olmalıdır.

Bağımsız değişkenlerin $A_1(x), \dots, A_g(x)$, g -tane lineer olasılık modeli göz önüne alındığında a_{ij} olasılığı aşağıdaki şekilde yazılabilir,

$$a_{ij} = y_j(x_i) \left\{ \sum_{r=1}^g y_r(x_i) \right\}^{-1} \quad (2.32)$$

Burada $y_j(x_i)$ ile (2.19) eşitliğinde $A(x)$ yerine $A_j(x_i)$ kullanılan Richard link fonksiyonu gösterilmiştir. Buna göre logaritmik olabilirlik fonksiyonu aşağıdaki şekilde kurulur,

$$\text{Log}L = \sum_{j=1}^g \sum_{i \in B_j} \log a_{ij} \quad (2.33)$$

Bu durumda olabilirlik fonksiyonunu maksimum yapan $A_1(x), \dots, A_g(x)$ lineer olasılık modelleri çoklu grup lojistik ayırmsamada kullanılabilir. (2.33) ile verilen logaritmik olabilirlik fonksiyonunun hangi durumlarda çözümünün var olacağı M. Gürcan ve Y. Öner tarafından incelenmiştir. Bu araştırmada $A_j(x_i) - A_r(x_i)$ farklarına göre değerlendirme yapılarak veri pozitif olması, sıfıra eşit olması ve negatif olmasına göre üç sınıfa ayrılmıştır. Şayet veri kümesi içindeki tüm gözlemler için bu fark pozitif oluyorsa veri grubu tam ayırmsamaya sahiptir. Bu durumda olabilirlik fonksiyonunun maksimum değeri mevcut değildir. Bazı gözlemler için bu fark sıfıra eşit oluyorsa yine olabilirlik fonksiyonunun maksimum değeri mevcut değildir. Ancak olabilirlik fonksiyonunun maksimum değerinin olabilmesi için veri kümesi içerisindeki en az bir gözlemden bu fark negatif olmalıdır [19].

3. BULGULAR

Literatürde kullanılan link fonksiyonlarının büyük bir çoğunluğu düzlemde S-eğrisi formunda yer almaktadır. Bunun nedeni lineer bir regresyon modelinin olasılık terimi haline getirilerek bağımlı değişkenin olasılık dağılımına uygun şekle getirilebilmesidir. Bu özellikten yararlanarak ilk olarak devamdaki kısımda Richard link fonksiyonunu dağılım fonksiyonu haline getirmeye çalışacağız.

3.1. Richard Yoğunluk Fonksiyonu

(2.19)'da verilen Richard link fonksiyonu $\alpha = 1$ parametresi ile aşağıdaki formda alınabilir,

$$F(x) = C\{1 + (m - 1)\exp(-kx)\}^{1/(1-m)} \quad (3.1)$$

Bu durumda link fonksiyonunun birinci türevi şu şekilde olacaktır,

$$\frac{dF(x)}{dx} = Ck\{1 + (m - 1)\exp(-kx)\}^{m/(1-m)} \exp(-kx) \quad (3.2)$$

Şimdi bu fonksiyonun $x \geq 0$ pozitif yarı ekseninde yoğunluk fonksiyonu olabilmesi için aşağıdaki şartı sağlayacak şekilde sabitini bulalım,

$$\int_0^{\infty} C k\{1 + (m - 1)\exp(-kx)\}^{m/(1-m)} \exp(-kx) dx = 1 \quad (3.3)$$

$$C = k\{1 - m^{1/(1-m)}\}^{-1}$$

Bu durumda Richard yoğunluk fonksiyonu aşağıdaki şekilde yazılabilir,

$$f(x) = k\{1 - m^{1/(1-m)}\}^{-1}\{1 + (m - 1)\exp(-kx)\}^{m/(1-m)} \exp(-kx) \quad (3.4)$$

Buradan Richard dağılım fonksiyonu yoğunluk fonksiyonu yardımıyla aşağıdaki formda olacaktır,

$$F(x) = (1 - m^{1/(1-m)})^{-1} \{ (1 + (m-1)\exp(-kx))^{1/(1-m)} - m^{1/(1-m)} \} \quad (3.5)$$

Bu dağılım fonksiyonunun özel durumda $m = 0$ için k parametrelili üstel dağılım fonksiyonuna dönüşecektir.

3.2. Alt Aralıklarda Tanımlanan Richard Dağılımı

Pozitif yarı ekseninde Richard dağılımının yoğunluk ve olasılık fonksiyonunu bir önceki bölümde elde etmiştik. $X > 0$, $X \sim Richard(k, m, x)$ şartı altında bu tesadüfi değişkenin beklenen değeri elde edilememektedir. Ancak bu dağılımın ortalamasının bulunması istatistiksel analizdeki uygulamalar için oldukça önem arz etmektedir. Dağılım fonksiyonu yardımıyla

$$EX = \int_0^{\infty} (1 - F(x)) dx \quad (3.6)$$

integrali beta integraline indirgenmekte ve bu durumda beta dağılımının parametresi sıfıra eşit çıkmaktadır. Parametresi sıfır olan beta dağılımı tanımlı olmadığından bu beklenen değer hesaplanamamaktadır. Bu sorundan dolayı Richard dağılımının alt aralıklarda tanımlanması bu problemi ortadan kaldırabilmektedir. Şimdi $\alpha_1 T \leq X \leq \alpha_2 T$ olmak üzere $X \sim Richard(\alpha_1 T, \alpha_2 T, k, m, x)$ dağılımının yoğunluk ve olasılık fonksiyonunu elde edelim. Bunun için aşağıdaki integralin değerini hesaplayalım,

$$\int_{\alpha_1 T}^{\alpha_2 T} k \exp(-kx) \{1 + (m-1)\exp(-kx)\}^{m/(1-m)} dx = c_2^{1/(1-m)} - c_1^{1/(1-m)} \quad (3.7)$$

Burada c_1 ve c_2 sabitleri devamdaki şekilde olmaktadır,

$$c_1 = 1 + (m-1)\exp(-k\alpha_1 T)$$

$$c_2 = 1 + (m-1)\exp(-k\alpha_2 T)$$

Bu durumda yoğunluk fonksiyonu aşağıdaki şekilde yazılır,

$$f(x) = \frac{k \exp(-kx)}{c_2^{1/(1-m)} - c_1^{1/(1-m)}} \{1 + (m-1) \exp(-kx)\}^{m/(1-m)} \quad (3.8)$$

Buradan dağılım fonksiyonu devamdaki formda olacaktır,

$$F(x) = \frac{1}{c_2^{1/(1-m)} - c_1^{1/(1-m)}} \left\{ (1 + (m-1)e^{-kx})^{1/(1-m)} - (1 + (m-1)e^{-k\alpha_1 T})^{1/(1-m)} \right\} \quad (3.9)$$

Dikkat edilecek olursa $\alpha_1 = 0$ ve $\alpha_2 \rightarrow \infty$ için yoğunluk ve olasılık fonksiyonlarının $X > 0$, şartı altında $X \sim Richard(k, m, x)$ dağılımının yoğunluk ve olasılık fonksiyonları ile aynı olacağı görülmektedir.

Alt aralıkta tanımlanan Richard dağılımının beklenen değerini yoğunluk fonksiyonunu kullanarak devamdaki şekilde elde edelim,

$$EX = \frac{k}{c_2^{1/(1-m)} - c_1^{1/(1-m)}} \int_{\alpha_1 T}^{\alpha_2 T} x e^{-kx} \{1 + (m-1) \exp(-kx)\}^{m/(1-m)} dx \quad (3.10)$$

Kısmi integrasyon kullanılırsa integral devamdaki formda olur,

$$EX = \frac{1}{C} \left\{ \alpha_2 T c_2^{1/(1-m)} - \alpha_1 T c_1^{1/(1-m)} - \int_{\alpha_1 T}^{\alpha_2 T} \{1 + (m-1) e^{-kx}\}^{1/(1-m)} dx \right\} \quad (3.11)$$

Burada C sabiti aşağıdaki şekilde alınmaktadır,

$$C = c_2^{1/(1-m)} - c_1^{1/(1-m)}$$

Beklenen değer içindeki integral sayısal olarak hesaplanacak olursa aşağıdaki şekilde elde edilebilir,

$$I = \frac{1}{n} \sum_{j=1}^n \left\{ 1 + (m-1) \exp\left(-k(\alpha_1 T + \frac{jT}{n}(\alpha_2 - \alpha_1))\right) \right\}^{1/(1-m)}, n \rightarrow \infty \quad (3.12)$$

Şimdi pozitif yarı eksende tanımlı $X \sim Richard(k, m, x)$ tesadüfi değişkenin yoğunluk ve olasılık fonksiyonu arasında var olan ilişkiyi kullanarak aşağıdaki teoremi ifade edelim,

Teorem: Pozitif yarı eksende tanımlı Richard dağılımının yoğunluk ve olasılık fonksiyonları sırasıyla $f(x)$ ve $F(x)$ olsun. Bu durumda aşağıdaki eşitlik sağlanır,

$$f(x) = \frac{k}{(m-1)(1-m^{1/(1-m)})} \{u(x) - u^m(x)\} \quad (3.13)$$

Burada $u(x)$ fonksiyonu aşağıdaki şekildedir,

$$u(x) = (1 - m^{1/(1-m)})F(x) + m^{1/(1-m)} \quad (3.14)$$

Buna ilaveten pozitif yarı eksende tanımlı Richard dağılımının beklenen değeri aşağıdaki eşitlikle ifade edilebilir,

$$EX = \frac{(1-m)}{k} + \frac{1}{1-m^{1/(1-m)}} \int_0^{\infty} (1 - u^m(x)) dx \quad (3.15)$$

İspat: $f(x)$ ve $F(x)$ fonksiyonları arasındaki eşitlik tanımlarından kolaylıkla elde edilebilmektedir. Bu eşitlik yoğunluk fonksiyonunun dağılım fonksiyonu cinsinden ifade edilmesini sağlamaktadır. Bu eşitliğin her iki tarafında $x > 0$ için integral alınacak olursa yoğunluk fonksiyonunun özelliğinden devamdaki integral yazılabilir,

$$\int_0^{\infty} (u(x) - u^m(x)) dx = \frac{(m-1)(1-m^{1/(1-m)})}{k} \quad (3.16)$$

Burada $u(x)$ fonksiyonunda dağılım fonksiyonunun ifadesi yazılacak olursa,

$$F(x) = \frac{u(x)-m^{1/(1-m)}}{1-m^{1/(1-m)}} \quad (3.17)$$

eşitliği elde edilir. Buna göre $1 - F(x)$ ifadesi aşağıdaki şekilde yazılabilir,

$$1 - F(x) = \frac{1-u(x)}{1-m^{1/(1-m)}} \quad (3.18)$$

Bu durumda dağılımın beklenen değeri aşağıdaki formda yazılır,

$$EX = \frac{1}{1-m^{1/(1-m)}} \int_0^{\infty} (1 - u(x)) dx \quad (3.19)$$

$u(x)$ ve $u^m(x)$ fonksiyonlarının farklarının integrali yardımıyla aşağıdaki eşitliği yazabiliriz,

$$(1 - m^{1/(1-m)})EX - \int_0^{\infty} (1 - u^m(x)) dx = \frac{(1 - m)(1 - m^{1/(1-m)})}{k} \quad (3.20)$$

Buna göre istenen sonuç aşağıdaki eşitlikten elde edilecektir,

$$EX = \frac{1}{1 - m^{1/(1-m)}} \left\{ \frac{(1 - m)(1 - m^{1/(1-m)})}{k} + \int_0^{\infty} (1 - u^m(x)) dx \right\} \quad (3.21)$$

Elde edilen eşitlikten $m = 0$ için $Exp(k)$ parametrelü üstel dağılımın beklenen değeri olan $1/k$ rahatlıkla elde edilebilmektedir. Aynı işlemler alt aralıklarda tanımlanan Richard dağılımları için de yapılabilmektedir.

3.3. Pozitif Yarı Eksende Tanımlı Richard Dağılımından Üretilen Sayılar

Richard dağılımı m ve k olmak üzere temelde iki parametreye sahiptir. Bu kısımda bu parametrelerin sabit değerleri seçilerek pozitif yarı eksende tanımlanan dağılımın sayısal değerleri incelenmeye çalışılacaktır. Üretilen sayı tabloları % 1'lik olasılık

artımlarıyla düzenlenecektir. Dağılımdan üretilen sayılar aşağıdaki eşitlik yardımıyla elde edilebilir,

$$x = \frac{-1}{k} \text{Ln} \frac{(u(1 - m^{1/(1-m)}) + m^{1/(1-m)})^{1-m} - 1}{m - 1} \quad (3.22)$$

Burada $0 \leq u \leq 1$ olasılık değeridir.

Tablo 1. $k = 1$ ve $m = 0.5$ için *Richard*(1,0.5, x) dağılımından üretilen sayısal değerler.

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0	0.015	0.030	0.045	0.060	0.075	0.090	0.105	0.120	0.135
0.1	0.151	0.166	0.181	0.197	0.212	0.228	0.244	0.259	0.275	0.291
0.2	0.307	0.323	0.340	0.356	0.373	0.389	0.406	0.423	0.440	0.458
0.3	0.475	0.493	0.510	0.528	0.546	0.565	0.583	0.602	0.621	0.640
0.4	0.660	0.679	0.699	0.720	0.740	0.761	0.728	0.803	0.825	0.847
0.5	0.870	0.893	0.916	0.939	0.964	0.988	1.013	1.039	1.065	1.091
0.6	1.118	1.146	1.174	1.203	1.233	1.264	1.295	1.327	1.360	1.394
0.7	1.429	1.466	1.503	1.542	1.581	1.623	1.666	1.711	1.757	1.806
0.8	1.857	1.910	1.966	2.026	2.088	2.155	2.226	2.302	2.386	2.473
0.9	2.570	2.678	2.798	2.933	3.089	3.273	3.498	3.788	4.195	4.890
									$\bar{X} =$	1.131

Tablo 2. $k = 1$ ve $m = 1.5$ için $Richard(1,1.5,x)$ dağılımından üretilen sayısal değerler.

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0	0.018	0.037	0.056	0.075	0.093	0.112	0.131	0.150	0.169
0.1	0.188	0.207	0.226	0.245	0.264	0.283	0.302	0.322	0.341	0.361
0.2	0.380	0.400	0.420	0.440	0.460	0.480	0.500	0.520	0.541	0.561
0.3	0.528	0.603	0.624	0.645	0.667	0.688	0.710	0.732	0.754	0.777
0.4	0.799	0.822	0.845	0.868	0.892	0.916	0.940	0.964	0.989	1.014
0.5	1.040	1.065	1.092	1.118	1.145	1.173	1.200	1.229	1.258	1.287
0.6	1.317	1.348	1.379	1.411	1.443	1.476	1.510	1.545	1.581	1.618
0.7	1.656	1.694	1.734	1.776	1.818	1.862	1.908	1.955	2.004	2.055
0.8	2.109	2.165	2.223	2.285	2.350	2.419	2.493	2.571	2.656	2.747
0.9	2.847	2.957	3.079	3.217	3.375	3.562	3.789	4.081	4.491	5.188
									$\bar{X} =$	1.293

Tablo 3. $k = 1$ ve $m = 2.5$ için $Richard(1,2.5,x)$ dağılımından üretilen sayısal değerler.

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0	0.021	0.042	0.063	0.084	0.105	0.126	0.147	0.168	0.189
0.1	0.211	0.232	0.253	0.275	0.296	0.317	0.339	0.361	0.382	0.404
0.2	0.426	0.448	0.470	0.492	0.514	0.536	0.558	0.581	0.604	0.626
0.3	0.649	0.672	0.695	0.719	0.742	0.766	0.790	0.814	0.838	0.862
0.4	0.887	0.912	0.937	0.962	0.988	1.014	1.040	1.066	1.093	1.120
0.5	1.148	1.175	1.203	1.232	1.261	1.290	1.320	1.350	1.381	1.412
0.6	1.144	1.476	1.510	1.543	1.578	1.613	1.648	1.685	1.723	1.761
0.7	1.801	1.841	1.883	1.926	1.970	2.016	2.063	2.112	2.163	2.216
0.8	2.271	2.329	2.389	2.453	2.520	2.590	2.666	2.746	2.832	2.925
0.9	3.026	3.138	3.261	3.401	3.561	3.749	3.978	4.272	4.683	5.382
									$\bar{X} =$	1.394

Tablo 4. $k = 1$ ve $m = 3.5$ için $Richard(1,3.5, x)$ dağılımından üretilen sayısal değerler.

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0	0.022	0.045	0.068	0.091	0.113	0.136	0.159	0.182	0.205
0.1	0.228	0.251	0.274	0.297	0.320	0.343	0.366	0.389	0.413	0.436
0.2	0.46	0.483	0.507	0.531	0.554	0.578	0.602	0.627	0.651	0.675
0.3	0.700	0.724	0.749	0.774	0.799	0.824	0.850	0.875	0.901	0.927
0.4	0.953	0.980	1.006	1.033	1.060	1.088	1.115	1.143	1.172	1.200
0.5	1.229	1.258	1.288	1.318	1.348	1.379	1.410	1.442	1.474	1.507
0.6	1.540	1.574	1.608	1.644	1.679	1.716	1.753	1.791	1.830	1.870
0.7	1.911	1.953	1.996	2.040	2.086	2.133	2.182	2.232	2.284	2.338
0.8	2.395	2.454	2.515	2.580	2.648	2.720	2.797	2.878	2.966	3.060
0.9	3.162	3.275	3.400	3.541	3.702	3.891	4.122	4.416	4.829	5.529
									$\bar{X} =$	1.475

Tablo 5. $k = 1$ ve $m = 6$ için $Richard(1,6, x)$ dağılımından üretilen sayısal değerler.

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0	0.025	0.051	0.077	0.103	0.129	0.155	0.181	0.207	0.233
0.1	0.259	0.285	0.311	0.337	0.363	0.389	0.415	0.442	0.468	0.494
0.2	0.521	0.547	0.574	0.601	0.628	0.655	0.682	0.709	0.736	0.763
0.3	0.791	0.818	0.846	0.874	0.902	0.930	0.958	0.987	1.015	1.044
0.4	1.073	1.102	1.132	1.161	1.191	1.222	1.252	1.283	1.314	1.345
0.5	1.377	1.409	1.441	1.474	1.507	1.540	1.574	1.608	1.643	1.678
0.6	1.714	1.751	1.788	1.825	1.864	1.903	1.943	1.983	2.025	2.067
0.7	2.111	2.155	2.200	2.247	2.295	2.345	2.396	2.449	2.503	2.560
0.8	2.619	2.680	2.744	2.811	2.882	2.956	3.035	3.118	3.208	3.305
0.9	3.409	3.524	3.652	3.794	3.958	4.150	4.382	4.679	5.093	5.796
									$\bar{X} =$	1.597

Tablo 6. $k = 1$ ve $m = 10$ için $Richard(1,10,x)$ dağılımından üretilen sayısal değerler.

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0	0.029	0.058	0.087	0.116	0.145	0.175	0.204	0.233	0.262
0.1	0.291	0.321	0.350	0.379	0.409	0.438	0.468	0.497	0.527	0.557
0.2	0.586	0.616	0.646	0.676	0.706	0.736	0.766	0.796	0.827	0.857
0.3	0.888	0.919	0.949	0.980	1.012	1.043	1.074	1.106	1.137	1.500
0.4	1.201	1.234	1.266	1.299	1.332	1.365	1.398	1.432	1.466	1.863
0.5	1.535	1.570	1.605	1.641	1.677	1.713	1.750	1.787	1.824	2.279
0.6	1.901	1.941	1.980	2.021	2.062	2.104	2.146	2.190	2.234	2.279
0.7	2.325	2.372	2.420	2.470	2.520	2.573	2.626	2.682	2.739	2.798
0.8	2.859	2.923	2.990	3.059	3.132	3.209	3.290	3.376	3.468	3.567
0.9	3.675	3.792	3.921	4.067	4.232	4.426	4.661	4.960	5.377	6.082
									$\bar{X} =$	1.7805

Tablo 7. $k = 1$ ve $m = 20$ için $Richard(1,20,x)$ dağılımından üretilen sayısal değerler.

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0	0.034	0.068	0.102	0.136	0.170	0.205	0.239	0.273	0.307
0.1	0.341	0.376	0.410	0.444	0.479	0.513	0.548	0.582	0.617	0.651
0.2	0.686	0.720	0.755	0.790	0.825	0.860	0.895	0.930	0.965	1.001
0.3	1.036	1.071	1.107	1.143	1.179	1.215	1.251	1.287	1.323	1.36
0.4	1.397	1.434	1.471	1.508	1.546	1.583	1.621	1.660	1.698	1.737
0.5	1.776	1.815	1.855	1.895	1.935	1.976	2.017	2.059	2.101	2.143
0.6	2.186	2.230	2.274	2.318	2.364	2.409	2.456	2.504	2.552	2.601
0.7	2.651	2.702	2.755	2.808	2.863	2.919	2.976	3.036	3.097	3.160
0.8	3.225	3.292	3.363	3.436	3.513	3.593	3.678	3.768	3.864	3.966
0.9	4.077	4.198	4.331	4.480	4.649	4.846	5.084	5.387	5.807	6.515
									$\bar{X} =$	2.0005

Tablo 8. $k = 1$ ve $m = 30$ için $Richard(1,30,x)$ dağılımından üretilen sayısal değerler.

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0	0.037	0.074	0.112	0.149	0.186	0.224	0.261	0.298	0.336
0.1	0.373	0.411	0.448	0.486	0.523	0.561	0.598	0.636	0.673	0.711
0.2	0.749	0.787	0.825	0.863	0.900	0.939	0.977	1.015	1.053	1.091
0.3	1.130	1.169	1.207	1.246	1.285	1.324	1.363	1.402	1.442	1.481
0.4	1.521	1.561	1.601	1.641	1.682	1.722	1.763	1.804	1.846	1.887
0.5	1.929	1.971	2.014	2.057	2.100	2.143	2.187	2.232	2.276	2.321
0.6	2.367	2.413	2.460	2.507	2.555	2.604	2.653	2.703	2.754	2.806
0.7	2.859	2.912	2.967	3.023	3.080	3.139	3.199	3.261	3.324	3.389
0.8	3.457	3.527	3.600	3.675	3.754	3.837	3.924	4.016	4.114	4.219
0.9	4.332	4.455	4.590	4.741	4.913	5.112	5.352	5.657	6.079	6.789
$\bar{X} =$										2.14

Tablo 9. $k = 0.2$ ve $m = 2.5$ için $Richard(0.2,2.5,x)$ dağılımından üretilen sayısal değerler.

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0	0.105	0.210	0.315	0.421	0.526	0.632	0.737	0.843	0.949
0.1	1.055	1.161	1.268	1.375	1.482	1.589	1.697	1.805	1.913	2.021
0.2	2.130	2.240	2.350	2.460	2.571	2.682	2.794	2.907	3.020	3.134
0.3	3.248	3.363	3.479	3.596	3.713	3.832	3.951	4.071	4.192	4.314
0.4	4.437	4.562	4.687	4.814	4.942	5.071	5.202	5.334	5.468	5.603
0.5	5.740	5.879	6.019	6.162	6.306	6.453	6.602	6.753	6.907	7.063
0.6	7.222	7.384	7.550	7.718	7.890	8.065	8.244	8.428	8.616	8.808
0.7	9.006	9.209	9.418	9.632	9.854	10.083	10.319	10.564	10.819	11.083
0.8	11.359	11.647	11.949	12.266	12.600	12.954	13.330	13.731	14.161	14.627
0.9	15.134	15.690	16.309	17.006	17.807	18.748	19.893	21.360	23.416	26.911
$\bar{X} =$										6.989

Tablo 10. $k = 0.02$ ve $m = 2.5$ için $Richard(0.02, 2.5, x)$ dağılımından üretilen değerler.

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0	1.052	2.105	3.158	4.212	5.266	6.321	7.378	8.436	9.495
0.1	10.556	11.619	12.685	13.752	14.822	15.895	16.971	18.050	19.133	20.219
0.2	21.309	22.404	23.503	24.606	25.715	26.829	27.948	29.073	30.204	31.342
0.3	32.486	33.638	34.796	35.963	37.138	38.321	39.513	40.714	41.925	43.147
0.4	44.379	45.622	46.877	48.144	49.423	50.716	52.023	53.345	54.682	56.035
0.5	57.404	58.792	60.197	61.622	63.068	64.534	66.023	67.536	69.073	70.637
0.6	72.228	73.849	75.500	77.183	78.901	80.656	82.449	84.284	86.163	88.088
0.7	90.064	92.093	94.180	96.329	98.544	100.832	103.198	105.649	108.193	110.839
0.8	113.596	116.477	119.495	122.665	126.007	129.544	133.301	137.313	141.619	146.273
0.9	151.340	156.907	163.095	170.068	178.071	187.481	198.930	213.605	234.167	269.112
									$\bar{X} =$	65.89

Tablo 11. $k = 0.5$ ve $m = 2.5$ için $Richard(0.5, 2.5, x)$ dağılımından üretilen değerler.

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0	0.042	0.084	0.126	0.168	0.210	0.252	0.295	0.337	0.379
0.1	0.422	0.464	0.507	0.550	0.592	0.635	0.678	0.722	0.765	0.808
0.2	0.852	0.896	0.940	0.984	1.028	1.073	1.117	1.162	1.208	1.253
0.3	1.299	1.345	1.391	1.438	1.485	1.532	1.580	1.628	1.677	1.725
0.4	1.775	1.824	1.875	1.925	1.976	2.028	2.080	2.133	2.187	2.241
0.5	2.296	2.351	2.407	2.464	2.522	2.581	2.640	2.701	2.762	2.825
0.6	2.889	2.953	3.020	3.087	3.156	3.226	3.297	3.371	3.446	3.523
0.7	3.602	3.683	3.767	3.853	3.941	4.033	4.127	4.225	4.327	4.433
0.8	4.543	4.659	4.779	4.906	5.040	5.181	5.332	5.492	5.664	5.850
0.9	6.053	6.276	6.523	6.802	7.122	7.499	7.957	8.544	9.366	10.764
									$\bar{X} =$	2.79

Tablo 12. $k = 0.05$ ve $m = 2.5$ için $Richard(0.05, 2.5, x)$ dağılımından üretilen değerler.

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0	0.421	0.842	1.263	1.684	2.106	2.528	2.951	3.374	3.798
0.1	4.222	4.647	5.074	5.501	5.929	6.358	6.788	7.220	7.653	8.087
0.2	8.523	8.961	9.401	9.842	10.286	10.731	11.179	11.629	12.081	12.536
0.3	12.994	13.455	13.918	14.385	14.855	15.328	15.805	16.285	16.770	17.258
0.4	17.751	18.248	18.750	19.257	19.769	20.286	20.809	21.338	21.872	22.414
0.5	22.961	23.516	24.079	24.649	25.227	25.813	26.409	27.014	27.629	28.255
0.6	28.891	29.539	30.200	30.873	31.560	32.262	32.979	33.713	34.465	35.235
0.7	36.025	36.837	37.672	38.531	39.417	40.333	41.279	42.259	43.277	44.335
0.8	45.438	46.590	47.798	49.066	50.403	51.817	53.320	54.925	56.647	58.509
0.9	60.536	62.763	65.238	68.027	71.228	74.992	79.572	85.442	93.667	107.645
									$\bar{X} =$	27.96

3.4. Büyüme Modellerinde Richard Link Fonksiyonunun Kullanılması

Büyüme eğrisi modellerinde kullanılan eğriler düzlemde S-eğrisi olarak adlandırılmaktadır. Reel uygulamalarda bir büyüme modeli temelde iki şekilde incelenmektedir. Birincisi zamana bağlı olarak gözlemlenen bir büyüme verisinin modellenmesi ikincisi ise ilgilenilen bir takım bağımsız değişkenlerin etkileri sonucunda gözlemlenen bir büyüme verisinin modellenmesidir. Zamana bağlı olarak gözlemlenen büyüme verilerinde gözlemlenen canlının veya bir organizmanın ilgilenilen boyut büyüklüğü dikkate alınmaktadır. Sağlıklı çocuk gelişiminde boy uzunluğu, kilo miktarı gibi değişkenler bu tip verilere örnek olarak gösterilebilir. İlgili bir takım bağımsız değişkenlerin değişimi sonucu meydana gelen artımların incelenmesine ise daha çok farmakolojik araştırmalarda rastlanmaktadır. Örnek olarak bir ilacın etkin olan maddesinin doza bağlı olarak kandaki emilimi veya verilen ilaca bağlı olarak hastada gözlemlenen değişiklikler verilebilir.

Buna ilaveten sıklıkla lojistik regresyon uygulamalarında da büyüme eğrileri kullanılmaktadır. İki grup lojistik ayrımında reel değişkenlerin sıfır ve bir gibi iki değer alan bağımlı değişkenle eşlenmesinde link fonksiyonları oldukça önemlidir. S-eğrisinin verilere uygun olarak yapılandırılmasında eğrinin esnekliğinin önemi oldukça fazladır. Bundan dolayı gerek büyüme eğrilerinde gerekse lojistik regresyon analizinde kullanılan link fonksiyonlarının esnekliğini artırıcı birden fazla parametreyle kullanılması önemlidir.

S-eğrileri sürekli artan bir eğilim göstermesine karşılık belli bir seviyeden sonra sabitleşmelidir. Doğal olarak büyümenin sonsuza kadar devam etmesi beklenemez. Ancak yapıya bağlı olarak S-eğrisindeki büyümenin ilk anlarda fazla daha sonra ise yavaşlayarak sabitleştiği gözlenmelidir. Bu durumda eğrinin bir bükülme noktası mevcut olmalıdır. Genelde lojistik ayrımsama problemlerinde bu bükülme noktası esas alınarak ayrımsama yapılmaktadır.

Bazı büyüme modellerinde S-eğrisi artan bir grafik gösterirken belli anlarda duraklamaya sahip olmaktadır. Bu anlarında belirlenmesi büyümenin incelenmesi açısından oldukça önemlidir. Veriden gözlemlenemeyen bu anlar büyüme modelinde modelin yapısına bağlı olarak bazı modellerde gözlemlenebilmektedir. Bir sonraki kısımda bu özelliğe değinilecektir.

Sonuç olarak gerek büyümenin gerekse lojistik ayrımsamanın daha sağlıklı bir şekilde yapılabilmesi için seçilen link fonksiyonunun esnekliğinin yeterince fazla olması yani bu esnekliği gösterebilecek parametre çokluğunun modelde bulunması oldukça önemlidir. Bilinen link fonksiyonlarının içerisinde bu esnekliğe en fazla sahip olan link fonksiyonu Richard link fonksiyonudur. Nadir olarak veri yapısının özelliğinden kaynaklanan bazı durumlarda diğer büyüme eğrilerinin de uyumu yeterli olabilmektedir.

Bu alt bölümde Richard link fonksiyonunun parametrelerinin veriye uygun olarak nasıl belirlenebileceğini göstermeye çalışacağız.

Model parametrelerinin tahmin edilmesinde kullanılan en önemli yöntem hiç şüphe yoktur ki en küçük kareler yöntemidir. Gauss Markov teoremine göre model parametrelerinin en küçük kareler tahmini minimum hatalı tahmindir. Bu teorem Johann Carl Friedrich Gauss ve Andrey Markov tarafından ispatlanmıştır. Johann Carl Friedrich Gauss 30 Nisan 1777 tarihinde dünyaya gelmiş alman matematikçidir. Ölümü 23 Şubat 1855 tarihindedir. 78 yıllık hayatı boyunca matematik bilimine birçok hizmette bulunmuş ender bilim adamlarından birisidir. Andrey Markov ise 14 Mayıs 1856 tarihinde doğmuş ve 66 yıllık ömrü boyunca sayısız birçok alanda bilime hizmet ederek 20 Haziran 1922 tarihinde hayatını kaybetmiştir. Lineer modeller için en iyi yansız tahmin edici olarak adlandırılan en küçük kareler yöntemi bu gün dahi sıklıkla kullanılan en etkili yöntemdir.

Yöntemin içeriğinde lineer denklem sistemi değişken sayısından fazla sayıda denklem içeriyorsa kesin çözüm olmadığından çözüm denklemleri minimum hatayla sağlayacak değerlerin bulunmasına dayanmaktadır. Lineer denklem sisteminin katsayılar matrisi bu durumda karesel olmayıp boyuna uzunlamasına bir matris olmaktadır. Ters

alınmadığı için amaç denklem sistemini soldan katsayılar matrisinin transpozu ile çarparak katsayılar matrisini karesel hale getirmektir. Gerçekten bu sayede lineer denklem sistemi karesel hale gelmekte ve katsayılar matrisinin tersi alınarak çözüm bulunabilmektedir.

Bazı uygulamalarda değişken sayısı fazla buna karşılık denklem sistemini oluşturan denklem sayısı değişken sayısından az olabilmektedir. Bu durumda denklem sistemi soldan katsayılar matrisinin transpozu ile çarpıldığında oluşan karesel matris tam ranklı olmayıp determinantı sıfır çıkmaktadır. Bu durumda tersi alınmamakta ve denklem sistemi yine çözülememektedir. Bu gibi durumlarda ise katsayılar matrisinin asal köşegen elemanlarına minimum pozitif bir sayı eklenerek bu sorun ortadan kaldırılabilir. Bu şekilde elde edilen regresyon modellerine Rich regresyon modelleri denilmektedir.

Sonuç olarak en küçük kareler yöntemi gerek lineer modellerde ve gerekse lineerleştirilebilen modellerde halen kullanılan ve bilinen en etkili yöntemdir. Lineerleştirme yöntemi sayısal hesaplamalarda çözülemeyen denklemlerin ardışık hesaplamalarla yaklaşık çözümünün bulunmasını sağlayan etkili ve önemli bir yöntemdir. Bu yöntemde denklemin lineer olmayan kısmı genellikle Taylor serisine açılarak serinin lineer olan ilk iki teriminin kullanılması sayesinde lineer hale getirilmektedir. Bu sayede çözülen denklemden bulunan sayısal değer bir sonraki adımda Taylor serisinde tekrar kullanılarak denklem bir kere daha çözülmektedir. İşlemler aynı adımlarla tekrarlanarak çözüm durağanlaşmaya kadar devam edilmektedir. Yöntem temelde lineer olmayan bir fonksiyonun reel bir değerinin zaten yaklaşık olarak değerlendirilmesinden kaynaklanmaktadır. Örnek olarak üstel veya trigonometrik bir ifadenin kesin değeri zaten bulunamamaktadır.

Richard link fonksiyonunda k parametresinin tahmin edilebilmesi için hata kareler toplamı oluşturularak sözü edilen yöntem kullanılabilir. Bu durumda Richard link fonksiyonunun i -inci gözlem değeri için değeri aşağıdaki şekilde olmak üzere,

$$\hat{y}_i = (1 + (m - 1) \exp(-k_0 x_i))^{1/(1-m)} \quad (3.23)$$

Taylor polinomu kullanılarak lineerleştirilmiş hali,

$$\hat{y}_i \approx \hat{y}_i(k_0) + \frac{d\hat{y}_i(k_0)}{dk} (k - k_0) \quad (3.24)$$

şeklinde kullanılabilir. Buradan hata kareler toplamı fonksiyonu ve türevi aşağıdaki şekilde elde edilecektir,

$$HKT = \sum_j (y_j - \hat{y}_j)^2 \quad (3.25)$$

$$\frac{dHKT}{dk} = \sum_j (y_j - \hat{y}_j) \frac{d\hat{y}_i(k_0)}{dk} = 0 \quad (3.26)$$

Bu denklemin çözümünden k parametresi bulunarak bulunan bu değer Taylor açılımında tekrar kullanılarak aynı işlemler tekrarlanacaktır. Tüm bu işlemler yapılırken m parametre değeri belirlenen sabit olarak kabul edilir. Bunun aksine k parametresinin sabitlenen değeri için m parametre değeri bulunmak isteniyorsa hata kareler toplamından m parametresine göre türev alınarak denklem çözülmelidir. Bunun için Taylor açılımı aşağıdaki formda olmalıdır,

$$\hat{y}_i \approx \hat{y}_i(m_0) + \frac{d\hat{y}_i(m_0)}{dm} (m - m_0) \quad (3.26)$$

Burada m parametresine göre türev,

$$\frac{d\hat{y}_i}{dm} = (1 + (m - 1) \exp(-k_0 x_i))^{1/(1-m)} \ln(1 + (m - 1) \exp(-k_0 x_i)) e^{-kx} \quad (3.27)$$

şeklinde olacaktır. Buradan hata kareler toplamının m parametresine göre türevinden elde edilen denklem devamdaki şekilde olur,

$$\frac{dHKT}{dm} = \sum_j (y_j - \hat{y}_j) \frac{d\hat{y}_i(m_0)}{dm} = 0 \quad (3.28)$$

Hata kareler toplamı fonksiyonunda \hat{y}_j değerinin lineerleştirilmiş şeklinin kullanıldığına dikkat edilmelidir.

Richard link fonksiyonunun yapı itibari ile kullanılan iki parametresinden k parametresinin eğrinin şekli ile ilgili, m parametresinin ise eğrinin konumuyla ilgili

olduğunu biliyoruz. Bir önceki alt kısımda yer alan tablolara dikkat edildiğinde Richard link fonksiyonundan üretilen sayıların m parametresi değiştikçe buna bağlı olarak ortalamasının pek fazla etkilenmediğini ancak ortalamasının k parametresine göre değiştiğini görebilmekteyiz.

Doğal olarak Richard link fonksiyonundan elde edilen dağılımın ortalamasının bulunamaması bize bu yorumu analitik olarak yapma kolaylığını sağlayamamaktadır. Ancak görsel olarak üretilen değerlerin ortalamasının etkin bir şekilde k parametresine bağlı olduğunu buna karşılık m parametresinden çok az miktarda etkilendiğini görebilmekteyiz. Bu link fonksiyonunun analitik yapısından kaynaklanan önemli bir özelliktir.

Model tahmini yapılırken m parametresi sabit tutulduğunda k parametresine göre uyum aranabilmektedir. Bu sayede model için uygun bir k değeri bulunduğunda verilerin modele daha iyi uyumunu sağlayabilmek açısından modeli m parametresine göre deforme edebiliriz.

Doğal olarak her bir veri değeri için farklı bir m parametre değeri bulmak model yapısını karmaşıktır. Ancak incelemenin özü buna dayanmaktadır. Modelin optimal uyumu için bu yapılmalıdır. Buna karşılık modelin belli bir hata payı göz önüne alınabiliyorsa modele uymayan aykırı değerler için bu işlemi yapmak daha uygun olabilmektedir.

Öteleme parametresini aşağıdaki şekilde almamız bunun için yeterli olacaktır. Bu durumda τ parametresini,

$$\tau = \frac{1}{1 + e^{-m}} \quad (3.29)$$

olarak alacak olursak $m \in (-\infty, \infty)$ için $\tau \in (0,1)$ aralığında değişir. Bu ise bize hatalı sınıflanan bazı gözlemler için m değerinin doğru sınıflama yapabilecek şekilde değiştirilebileceğini garanti etmektedir.

3.5. Büyüme Eğrilerinde Duraklama Anlarının Belirlenmesi

Büyüme eğrisi modellerinde kullanılan S-eğrileri genellikle üstel yapıda sürekli fonksiyonlar olup artan olduğu aralıklarda büküm noktalarına sahip değildirler. Ancak bazı büyüme verilerinde büyümenin gerçekleşmediği duraklama anları bulunabilmektedir. Böyle bir büyüme verisini herhangi bir büyüme modeliyle açıklamaya çalıştığımızda veride var olan bu duraklama anları kaybolmaktadır. Bu ise analizde oldukça önemli olan bir bilgi kaybına neden olmaktadır. Özellikle farmakolojik verilerde herhangi bir ilacın meydana getirdiği etkiyi gözlemlerken meydana gelen etkinin artmaması, etkinin istenen seviyeye gelememesini sağlamaktadır. Bu durumda araştırmacı bu duraklama anını tespit ettiğinde kullanılan ilacın dozunu artırabilir veya bunun nedenini ilaca bağlı olarak araştırma imkânına sahip olabilir.

Bir büyüme eğrisi verisinde var olan duraklama anlarının tespit edilmesi ancak incelenen verilere ait modelin ikinci türevinin sifıra eşit olabileceği noktaların bulunmasıyla mümkün olabilmektedir. Bu noktalar bir büyüme eğrisinin ikinci türevi artan olduğu aralıkta hiçbir zaman sifıra eşit olmayacağından büyüme eğrisi yardımıyla elde edilemez. Bu noktaların tespiti için verilere farklı bir modelin yaklaşımı gerekmektedir. Bu ise modelin yaklaşımının değerlendirilebildiği lineer pozitif operatörler yardımıyla yapılabilir. Genelde modelin yaklaşımı modelin uyum iyiliği olan regresyon kareler toplamı veya hata kareler toplamı yardımıyla kıyaslanmaktadır. Ancak lineer pozitif operatörlerde modelin yaklaşımı model değeri ile gerçek fonksiyon değerinin farkının mutlak değerinin maksimum sınırının bulunmasıyla yapılmaktadır. Bu durumda fonksiyon ile operatörün sabit değerinin farkı değerlendirilirken fonksiyonun türevlerinin de kullanılmasını mümkün kılabilir. Bu konu ile bilinen en önemli ilk kaynak Voronovskaja'nın asimptotik formülüdür. 1932 yılında yapılan bu çalışmada aşağıdaki asimptotik yaklaşım elde edilmiştir,

$$B_n(f, x) = f(x) + \frac{x(1-x)}{2n} f''(x) + O(1/n), \quad n \rightarrow \infty \quad (3.30)$$

Burada $B_n(f, x)$, n -inci dereceden Bernstein polinomu, $f(x)$, $0 \leq x \leq 1$ sürekli ve ikinci mertebeden türevlenen bir fonksiyon ve $O(1/n)$ ise $n \rightarrow \infty$ için sifıra yaklaşan hata terimidir. E. Voronovskaja bu çalışmasını Moskova bilimler akademisinin önemli bir

yayını olan Doklud Akademia Nauk SSSR dergisinde yayınlamıştır. Daha sonraki yıllarda bu konu ile ilgili bazı çalışmalara literatürde rastlanabilmektedir. Özellikle 2008 yılında yapılan bir çalışmada Bulgar araştırmacı Gancho T. Tachev, Voronovskaja'nın asimptotik formülünü farklı şekilde yorumlayarak bazı sonuçlara ulaşabilmiştir.

Konunun en geniş ele alınış şekli M. Gürcan tarafından 2011 yılında yapılan bir çalışmayla ortaya konulmuştur. Bu çalışmada fonksiyon ile herhangi keyfi mertebeden türevlerinin lineer kombinasyonu ile yaklaşım operatörünün değerinin farkının mutlak değeri bir üst sınır ile sınırlandırılmıştır. A. A. Shahbazov'un da içinde bulunduğu bu çalışmada önemli birçok sonuç elde edilebilmiş ve literatüre kazandırılmıştır. A. A. Shahbazov'un 2007 yılında ölümünden sonra bu çalışma Shahbazov'un anısına ithafen yayınlanmıştır. Şimdi devamda kısaca bu ifadelerden bahsedelim.

Kesikli parametrelili $\{X_n: n \geq 1\}$ stokastik sürecinde her bir $n \geq 1$ için X_n tesadüfi değişkeni (n, x) parametrelili dağılıma sahip olsun. Yaklaşımında kullanılacak olan lineer pozitif operatör,

$$L_n(f, x) = Ef(X_n) < \infty \quad (3.31)$$

olarak tanımlansın. Burada f fonksiyonu $[a, b]$ aralığında tanımlı reel değerli sürekli bir fonksiyon olup m -inci türevi $f^{(m)} \in C[a, b]$ olarak alınsın. X_n tesadüfi değişkenininin k -inci momentini aşağıdaki şekilde gösterelim,

$$b_k = b_k(n, x), \quad k = 0, 1, \dots, m \quad (3.32)$$

Bu durumda tek pozitif m tamsayıları için aşağıdaki yaklaşım yazılabilmektedir,

$$\left| L_n(f, x) - \sum_{k=0}^m \frac{f^{(k)}(x)}{k!} b_k \right| \leq \frac{1}{m!} (\sqrt{b_{2m}} + \delta b_{m+1}) \omega_m(1/\delta) \quad (3.33)$$

Burada $\omega_m(1/\delta)$ fonksiyonu, $f^{(m)}$ fonksiyonunun süreklilik modülüdür,

$$\omega_m(1/\delta) = \sup_{|x-y|<\delta} |f(x) - f(y)| \quad (3.34)$$

Benzer şekilde çift pozitif m tamsayıları için aşağıdaki yaklaşım yazılabilmektedir,

$$\left| L_n(f, x) - \sum_{k=0}^m \frac{f^{(k)}(x)}{k!} b_k \right| \leq \frac{1}{m!} (b_m + \delta \sqrt{b_2 b_{2m}}) \omega_m(1/\delta) \quad (3.35)$$

Dikkat edilecek olursa toplamın terimleri fonksiyonun kendisi ile türevlerinden oluşmaktadır. Bu durum Voronovskaja'nın ifadesinin keyfi mertebeli türevlere genişletilmiş şeklidir. Bu kısımda bizim kullanacağımız ifade sadece ikinci türevle ilgili olduğundan toplamı $m = 2$ için yazmamız yeterli olacaktır. Tesadüfi değişkene ait b_k momentleri de merkezi momentler olarak kullanılırsa istenen ifade rahatlıkla elde edilebilmektedir,

$$\left| L_n(f, x) - f(x) - \frac{b_2}{2} f''(x) \right| \leq \frac{1}{2} (b_2 + \delta \sqrt{b_2 b_4}) \omega_2(1/\delta) \quad (3.36)$$

Burada merkezi birinci moment sıfır olduğundan fonksiyonun birinci mertebeden türevi toplamda bulunmamaktadır. Toplam ifadesi sadece fonksiyonun kendisi ve ikinci türevinden ibaret olmaktadır. Bu ifade $S_n \sim \text{Binom}(n, x)$ olmak üzere $X_n = S_n/n$ olarak alındığında,

$$b_2 = \frac{x(1-x)}{n} \quad (3.37)$$

şeklinde elde edilir. b_4 merkezi momenti ise aşağıdaki şekilde sınırlandırılabilir,

$$b_4 = E\left(\frac{S_n}{n} - x\right)^4 = \frac{E(S_n - nx)^4}{n^4} \leq \frac{3}{4n^2} \quad (3.38)$$

Elde edilen bu ifadeler ilgilenilen eşitsizlikte yerine yazılacak olursa aşağıdaki sonuç elde edilecektir,

$$\left| B_n(f, x) - f(x) - \frac{x(1-x)}{2n} f''(x) \right| \leq \frac{1 + \sqrt{3}}{8n} \omega_2(1/\sqrt{n}) \quad (3.39)$$

Burada $B_n(f, x)$ Bernstein polinomu olup $\delta = \sqrt{n}$ ve $x(1-x) \leq 1/4$ olarak alınmıştır. Elde edilen bu son eşitsizlikten fonksiyonun ikinci türevi için alt ve üst sınır değerleri aşağıdaki şekilde elde edilecektir,

$$\begin{aligned} -\frac{1+\sqrt{3}}{8n} \omega_2 + B_n(f, x) - f(x) &\leq \frac{x(1-x)}{2n} f''(x) \\ &\leq \frac{1+\sqrt{3}}{8n} \omega_2 + B_n(f, x) - f(x) \end{aligned} \quad (3.40)$$

Elde edilen bu daraltılmış aralıkta fonksiyonun ikinci türevi sıfır değerini alabiliyorsa bu aralıkta bir büküm noktasına sahip olacaktır. Büküm noktaları büyüme modellerinde duraklama anlarına karşılık gelmektedir.

3.6. Büyüme Eğrisinin Duraklama Anlarına Göre Parçalanması

Richard link fonksiyonu aynı zamanda bir büyüme eğrisi modeli olduğu için hiçbir noktada büküme sahip olamamaktadır. Ancak pozitif yarı ekseninde tanımlanan Richard dağılımı başlangıç noktasında büküme sahip olabilmektedir. Bu özelliğinden dolayı alt aralıklarda tanımlanan Richard dağılımları peş peşe eklenerek tam bir pozitif ekseni örtebilir. Doğal olarak alt aralıklarda tanımlanan dağılımların peş peşe eklenerek yeni bir büyüme modelinin oluşturulması dağılım fonksiyonu yapısını bozacak ancak büyüme eğrisi modeline duraklama anlarının yerleştirilebilmesine olanak sağlayacaktır.

Büyüme eğrisinin pozitif yarı ekseninde tespit edilen duraklama anları t_1, t_2, \dots olsun. Richard dağılımının tanımlı olduğu alt aralıkları $A_j = [\alpha_j T, \alpha_{j+1} T]$, $j = 0, 1, \dots$ olarak alalım. Bu durumda ilk duraklama anı ve ardışık duraklama anları aşağıdaki şekilde alt aralıkların uç noktalarına eşlensin,

$$\alpha_0 T = 0, \alpha_j T = t_j, j = 1, 2, \dots$$

A_j alt aralığında tanımlı Richard dağılımının olasılık fonksiyonu F_j , $j = 0, 1, \dots$ olsun. $I(A_j)$ alt aralığın karakteristik fonksiyonu olmak üzere büyüme eğrisi modeli devamdaki biçimde yazılabilmektedir,

$$F(x) = \sum_j [j + F_j(x)] I(A_j) \quad (3.41)$$

Benzer şekilde elde edilen bu büyüme eğrisi modeli zamana bağlı periyodik verilerin incelenebilmesi için de kullanılabilir. Ancak bu durumda serinin tespit edilen kırılma noktalarında oluşturulan çekirdek model azalan trend aralıklarında $1 - F$ şeklinde kullanılmalıdır. Artan ve azalan trendlerin birlikte ele alındığı durumda modelin çekirdek kısmı,

$$1 - \alpha + (2\alpha - 1)F \quad (3.42)$$

şeklinde alınabilir. Burada α parametresi azalan trend için $\alpha = 0$, artan trend için ise $\alpha = 1$ olarak alınacaktır. $\alpha = 1/2$ olması durumu ise serinin durağanlığını gösterecektir. α parametresinin dağılımı belirlenebildiğinde serinin durağanlık testi yapılabilir.

3.7. Sarmal Richard Dağılımı

Tesadüfi değişkenin periyodik olduğu durumlarda veya gözlem değerlerinin periyodik olduğu durumlarda kullanılacak uygun olan dağılımlar dairesel dağılımlardır. Dairesel dağılımlar son yıllarda oldukça üzerinde durulan önemli konulardan birisidir. Bu konuyla ilgili ilk tanımlama sarmal (wrapped) yoğunluk fonksiyonudur. $X > 0$ ve $[0, 2\pi]$ periyod uzunluğuna sahip olan ve periyodik değer alan sürekli bir değişken olsun. Yoğunluk fonksiyonu $f(x)$ olmak üzere yoğunluk fonksiyonunun pozitif yarı eksen üzerinden integrali bir olmalıdır. Buna göre pozitif yarı ekseni 2π uzunluğundaki parçalara ayırdığımızı düşünürsek aşağıdaki toplamı rahatlıkla yazabiliriz,

$$[0, \infty) = [0, 2\pi) + [2\pi, 4\pi) + \dots = \sum_{k=0}^{\infty} [2k\pi, (2k+2)\pi) \quad (3.43)$$

Bu durumda pozitif yarı eksende yoğunluk fonksiyonunun integrali alt aralıkların integrallerinin toplamı şeklinde yazılabilecektir,

$$\int_0^{\infty} f(x) dx = \sum_{k=0}^{\infty} \int_{k\pi}^{(2k+2)\pi} f(x) dx$$

$$= \sum_{k=0}^{\infty} \int_0^{2\pi} f(\varphi + 2k\pi) d\varphi = 1 \quad (3.44)$$

Elde edilen bu sonuçtan sarmal yoğunluk fonksiyonu aşağıdaki şekilde tanımlanabilmektedir,

$$f_w(\varphi) = \sum_{k=0}^{\infty} f(\varphi + 2k\pi) \quad (3.45)$$

Bu yoğunluk fonksiyonu tüm reel ekseninde tanımlanmak istendiğinde devamdaki şekilde rahatlıkla yazılabilmektedir,

$$f_w(\varphi) = \sum_{k=-\infty}^{\infty} f(\varphi + 2k\pi) \quad (3.46)$$

Yoğunluk fonksiyonunun periyodik veriye göre uyarlanması ile yoğunluk fonksiyonunun $[-\pi, \pi]$ aralığında daraltılması farklı işlemler olup birincisinde yoğunluk fonksiyonu periyod uzunluğuna göre parçalanırken ikincisinde ise bir link fonksiyonu kullanılarak fonksiyonun değişim aralığı ilgilenilen aralığa daraltılmaktadır. Bu şekilde kullanılan en önemli link fonksiyonu,

$$X = \frac{\cos\varphi}{1 - \sin\varphi} \quad (3.47)$$

topolojik eş yapılı (homeomorfizm) dönüşümdür. Bu dönüşüm pozitif reel eksenini $[-\pi/2, \pi/2]$ aralığına daraltmaktadır.

Uygulanan yöntem her türlü yoğunluk fonksiyonu için yapılabilir şekilde gözükse de tesadüfi değişkenin gözlem değerlerinin uygun olan sarmal dağılıma uyum göstermesi gerekmektedir. Bu nedenle uygulamada sıklıkla kullanılan düzgün, normal, Cauchy gibi dağılımların sarmal şekilde düzenlenmiş biçimleri bulunmaktadır. Aşağıda sırasıyla sarmal düzgün, sarmal normal ve sarmal Cauchy dağılımlarının yoğunluk fonksiyonları verilmiştir,

$$WU(\varphi) = \frac{1}{2\pi} \quad (3.48)$$

$$WN(\varphi, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \exp\left[-\frac{(\varphi - \mu - 2k\pi)^2}{2\sigma^2}\right] \quad (3.49)$$

$$\begin{aligned} WC(\varphi, \varphi_0, \gamma) &= \sum_{k=-\infty}^{\infty} \frac{\gamma}{\pi(\gamma^2 + (\varphi + 2k\pi - \varphi_0)^2)} \\ &= \frac{1}{2\pi} \frac{\sinh\gamma}{\cosh\gamma - \cos(\varphi - \varphi_0)} \end{aligned} \quad (3.50)$$

Bunlara ilaveten sarmal dağılımların en yaygın kullanılanlarından birisi de von Misses dağılımıdır. Bu dağılım Besses fonksiyonunun özelliğinden yararlanılarak elde edilmiştir. Dağılımın yoğunluk fonksiyonu aşağıdaki şekildedir,

$$Wf(x) = \frac{\exp(K\cos(x - \mu))}{2\pi I_0(K)} \quad (3.51)$$

Burada $I_0(K)$ sıfırıncı dereceli modifiye edilmiş Bessel fonksiyonudur.

Şimdi bu kısımda pozitif yarı eksende tanımlı olan Richard dağılımı için aynı uygulamayı yapmaya çalışalım. $X > 0$ ve $[0, 2\pi]$ periyod uzunluğuna sahip olan bir tesadüfi değişken olsun. Richard yoğunluk fonksiyonu $f(x)$ olmak üzere sarmal yoğunluk fonksiyonu aşağıdaki şekilde yazılacaktır,

$$Wf(\theta) = \sum_{n=0}^{\infty} \frac{ke^{-k(\theta+2n\pi)}}{1 - m^{1/(1-m)}} (1 + (m-1)e^{-k(\theta+2n\pi)})^{m/(1-m)} \quad (3.52)$$

Sarmal Richard yoğunluk fonksiyonunun $[0, 2\pi]$ aralığındaki integrali,

$$\begin{aligned} &\int_0^{2\pi} Wf(\theta) d\theta \\ &= \sum_{n=0}^{\infty} \frac{\{(1 + (m-1)e^{-(n+1)2k\pi}\}^{1/(1-m)} - (1 + (m-1)e^{-2nk\pi}\}^{1/(1-m)})\}}{1 - m^{1/(1-m)}} \end{aligned}$$

$$= \frac{1}{1 - m^{1/(1-m)}} \lim_{n \rightarrow \infty} [(1 + (m - 1)e^{-(n+1)2k\pi})^{1/(1-m)} - m^{1/(1-m)}] = 1 \quad (3.53)$$

olacak şekilde elde edilebilir. Sarmal Richard dağılımı ile alt aralıklarda tanımlanan Richard dağılımları yapısal olarak benzerlik göstermektedirler. Her bir alt aralık aynı periyot uzunluğuna sahip olarak alındığında ve dağılımlar, aralıklar birbirini takip edecek şekilde devam ettirildiğinde pozitif yarı ekseninde tanımlı sarmal Richard dağılımı benzer şekilde elde edilmiş olacaktır. Ancak bu işlemde katsayılar uygun şekilde değiştirilmelidir.

3.8 Uygulama: Karınca Verileri ve Dağılım Fonksiyonunun Sürekli Deformasyonu Kullanılarak Dairesel Richard Dağılımına Ait Uygulama

İstatistiksel analizin veri türleri içerisinde sürekli veriler ve bunların incelenmesi önemli bir bölümü oluşturmaktadır. Sürekli veriler içerisinde bir çember üzerinde yer alan açısal veriler sürekli veri türleri cinsinden sayılmakla birlikte sürekli verilerin analiz metotları ile incelenememektedir. Yirminci yüzyılın ikinci yarısından itibaren kompleks analizin yöntemleri kullanılarak istatistiksel analiz metotları içerisinde açısal verilere ait yeni yöntemler geliştirilebilmiştir. Bu yöntemlerin en önemlilerinin başında sarmal hale getirilmiş dağılım fonksiyonları gelmektedir. Bu bölümde de karınca verileri kullanılarak bu verilere uyan sarmal Richard dağılımı oluşturulacaktır. Karınca verileri karıncaların gidiş yönlerini gösteren açısal değerlerdir. Uygulamada 100 adet karıncadan gözlenen açı değerleri kullanılacaktır. Veriler aşağıdaki tabloda gösterilmiştir,

Tablo 13. Karınca verileri

Sıra no:	Açı değeri	Frekansı	Sıra no:	Açı değeri	Frekansı	Sıra no:	Açı değeri	Frekansı
1	10	1	11	130	1	21	230	2
2	30	2	12	140	3	22	250	1
3	40	1	13	150	5	23	260	1
4	50	1	14	160	7	24	270	1
5	60	2	15	170	6	25	280	2
6	70	1	16	180	17	26	290	1
7	80	1	17	190	8	27	300	2
8	90	1	18	200	13	28	330	1
9	110	2	19	210	8	29	350	1
10	120	2	20	220	5	30	360	1

Yukarıdaki tabloda gözlenen verilerin ortalaması 181.1 olarak hesaplanır. Sarmal Richard dağılımı sonsuz toplam olarak tanımlandığından hesaplamalarda toplam ifadesi belli terim sayısında bırakılmalıdır. Bu verilerde toplamın ilk beş terimi hesaplamaya dahil edilmiştir. Diğer terimler 10^{-5} den küçük olduğundan toplamda göz ardı edilmiştir. Bu durum dikkate alınarak sarmal Richard dağılımının büyüme oranı $k = 1/5(181.1) = 0.001$ olarak alınmıştır. Verinin deneysel dağılım fonksiyonu oluşturularak birikimli olasılıklar ve dağılımın şekil parametreleri aşağıdaki tabloda hesaplanmıştır,

Tablo 14. Sarmal dağılımın şekil parametreleri

Sıra no:	Açı değeri	Olasılığı	Birikimli olasılığı	m değeri
1	10	0.01	0.01	0.001
2	30	0.02	0.03	0.001
3	40	0.01	0.04	0.001
4	50	0.01	0.05	0.001
5	60	0.02	0.07	0.015
6	70	0.01	0.08	0.016
7	80	0.01	0.09	0.017
8	90	0.01	0.1	0.019
9	110	0.02	0.12	0.022
10	120	0.02	0.14	0.037
11	130	0.01	0.15	0.04
12	140	0.03	0.18	0.072
13	150	0.05	0.23	0.137
14	160	0.07	0.3	0.25
15	170	0.06	0.36	0.337
16	180	0.17	0.53	0.66
17	190	0.08	0.61	0.817
18	200	0.13	0.74	1.1
19	210	0.08	0.82	1.275
20	220	0.05	0.87	1.385
21	230	0.02	0.89	1.422
22	250	0.01	0.9	1.42
23	260	0.01	0.91	1.431
24	270	0.01	0.92	1.443
25	280	0.02	0.94	1.48
26	290	0.01	0.95	1.492
27	300	0.02	0.97	1.53
28	330	0.01	0.98	1.517
29	350	0.01	0.99	1.515
30	360	0.01	1	1.527

Sarmal Richard dağılımından m değerine göre hesaplanan tahmini olasılıklar ve hesaplanan mutlak hatalar aşağıdaki tabloda verilmiştir,

Tablo 15. Dağılımdan tahmin edilen olasılık değerleri

Sıra no:	Birikimli olasılığı	Tahmini olasılık	Hata
1	0.01	0.0109	0.0009
2	0.03	0.0304	0.0004
3	0.04	0.04	0
4	0.05	0.049	0.001
5	0.07	0.0705	0.0005
6	0.08	0.0804	0.0004
7	0.09	0.0902	0.0002
8	0.1	0.1007	0.0007
9	0.12	0.1206	0.0006
10	0.14	0.1400	0
11	0.15	0.1506	0.0006
12	0.18	0.1807	0.0007
13	0.23	0.2303	0.0003
14	0.3	0.3049	0.0049
15	0.36	0.3604	0.0004
16	0.53	0.5307	0.0007
17	0.61	0.6101	0.0001
18	0.74	0.7408	0.0008
19	0.82	0.8200	0
20	0.87	0.8704	0.0004
21	0.89	0.8907	0.0007
22	0.9	0.900	0
23	0.91	0.9102	0.0002
24	0.92	0.9202	0.0002
25	0.94	0.9401	0.0001
26	0.95	0.9501	0.0001
27	0.97	0.9702	0.0002
28	0.98	0.9805	0.0005
29	0.99	0.99007	0.00007
30	1	0.9999	0.00001

Yukarıdaki tabloda gösterilen mutlak hatalardan hata kareler toplamı 0.000030565 olarak hesaplanmıştır. Aşağıda m değerlerinin tahmin edilmesi için kullanılan program kodları bulunmaktadır,

Tablo 16. *m* değerlerinin tahmininde kullanılan R program kodları

```
fy<-c(0.001, 0.001, 0.001, 0.001, 0.001, 0.001, 0.008, 0.015, 0.016, 0.017, 0.019, 0.0205, 0.022, 0.0295, 0.037, 0.04,
0.0506, 0.0612, 0.072, 0.085, 0.098, 0.111, 0.124, 0.137, 0.1531, 0.1692, 0.1853, 0.2014, 0.2175, 0.2336, 0.25, 0.2645,
0.279, 0.2935, 0.308, 0.3225, 0.337, 0.356, 0.375, 0.394, 0.413, 0.432, 0.451, 0.47, 0.489, 0.508, 0.527, 0.546, 0.565,
0.584, 0.603, 0.622, 0.641, 0.66, 0.6796, 0.6992, 0.7188, 0.7384, 0.758, 0.7776, 0.7972, 0.817, 0.8387, 0.8604, 0.8821,
0.9038, 0.9255, 0.9472, 0.9689, 0.9906, 1.0123, 1.034, 1.0557, 1.0774, 1.1, 1.1218, 1.1436, 1.1654, 1.1872, 1.209,
1.2308, 1.2526, 1.275, 1.297, 1.319, 1.341, 1.363, 1.385, 1.4035, 1.422, 1.42, 1.431, 1.443, 1.4615, 1.48, 1.492, 1.511,
1.53, 1.517, 1.515, 1.527)
n=length(y)
z=rep(0,n)
b<-rep(0,n)
for (j in 1:n){
x[j]=(j-1)/(n-1)
for (i in 1:n){
a[i]=choose(n-1, (i-1))*(x[j]^(i-1))*((1-x[j])^(n-i))*y[i]
z=a
b[j]<-sum(z)
}
}
b
fb<-c(0.001, 0.00100424854855888, 0.00113802068173531, 0.00180052470531673, 0.00331245927552061,
0.00561234466720902, 0.00838380601428741, 0.0113284685182807, 0.0143278009497619, 0.0174547303680771,
0.0209027370847066, 0.0249023722901871, 0.0296610362110081, 0.0353324205344642, 0.0420082109392453,
0.0497225982302075, 0.0584628663102387, 0.0681821513670121, 0.0788119593236446, 0.090272659524541,
0.102480767036972, 0.115352725455213, 0.128805909458406, 0.142758282392712, 0.157128299759201,
0.171836241843541, 0.186807400389937, 0.20197672874415, 0.217293945409542, 0.232727806031232,
0.248268361977316, 0.263926443276231, 0.279730213222078, 0.295719283042687, 0.311937391834409,
0.328424930901183, 0.345212569244935, 0.362316942121271, 0.379738888957667, 0.397464200508521,
0.415466386671651, 0.433710697700441, 0.452158557035177, 0.470771670693021, 0.489515303072647,
0.508360474910409, 0.527285078790988, 0.546274079123699, 0.56531905359152, 0.584417351016014,
0.603571107228603, 0.622786298057135, 0.642071934198787, 0.661439427865225, 0.680902093577291,
0.700474693315699, 0.720172908941884, 0.740012631745047, 0.760009004778741, 0.780175233187815,
0.800521273572347, 0.821052598207721, 0.841769273810421, 0.862665575765474, 0.883730273180382,
0.90494758551022, 0.92629866375343, 0.947763331625242, 0.969321770759828, 0.990955864989829,
1.01265002206074, 1.0343914332208, 1.05616986537496, 1.0779771616252, 1.09980662371865, 1.12165235771704,
1.14350850162406, 1.16536806205628, 1.18722092128796, 1.20905049727267, 1.23082861276662,
1.25250843003824, 1.27401592439839, 1.29524139143039, 1.31603392292173, 1.33620341959782,
1.35553583247618, 1.37382658264772, 1.39093267738779, 1.40683458178288, 1.42168547683249,
1.43581376218089, 1.4496449807207, 1.46353207294668, 1.47752718544692, 1.49117339258061,
1.50341619659133, 1.51276809120929, 1.51806161943386, 1.52053050253067, 1.527)
```

4. SONUÇ VE TARTIŞMA

İstatistik bilim dalında önemli bir yere sahip olan regresyon analizi, link fonksiyonları, büyüme eğrileri ve lojistik ayırimsama problemleri günümüz biliminde halen araştırılmaya devam edilen ve birçok önemli araştırmacının üzerinde durduğu konuların başında gelmektedir. Uygulama sahasının geniş olması konuya ayrı bir önem kazandırmaktadır. Bu konuyla ilgili geliştirilen yeni yöntemler başta ekonomi, sağlık bilimleri, endüstri, sosyal bilimler olmak üzere diğer bilim dallarının araştırmacıları tarafından dikkatle takip edilmektedir.

Ekonomi ile ilgili alanlarda büyüme eğrisi denildiğinde, ekonomik büyüme, ticari firmaların sermaye artışı, üretim artışı gibi analizler ilk akla gelen inceleme sahalarıdır.

Sağlık bilimlerinde büyüme eğrisi veya lojistik ayırimsama denildiğinde, sağlıklı bireyin gelişim çağındaki büyümesinin incelenmesi, kan tahlillerinde kandaki incelenen maddenin artışı, teşhis için gelen bireyin ölçümlerine dayanarak hasta olup olmamasının kararı gibi analizler ilk akla gelen inceleme alanlarıdır. Bu gibi örnekler diğer sahalarda da sıklıkla verilebilir.

Yapılan çalışmada Richard link fonksiyonları incelenmiştir. Richard link fonksiyonlarının gerek büyüme eğrileri olsun gerekse lojistik ayırimsama olsun her iki alanda da başarılı sonuç vermesinin önemli bir nedeni eğrinin esnek bir yapıya sahip olmasıdır. Bu özelliği modelin veriye uyumunu artırarak istatistiksel incelemenin daha sağlıklı sonuç vermesini sağlamaktadır.

Çalışmada sırasıyla aşağıdaki analizlere yer verilmiştir,

1. $[0, \infty)$ aralığında tanımlanan Richard yoğunluk ve dağılım fonksiyonu,
2. Reel eksenin alt aralıklarında tanımlanan Richard yoğunluk ve dağılım fonksiyonu,
3. Büyüme modellerinde Richard link fonksiyonunun kullanılması,
4. Büyüme eğrilerinde duraklama anlarının belirlenmesi,
5. Büyüme eğrilerinin duraklama anlarına göre parçalanması,
6. Sarmal hale getirilmiş Richard link fonksiyonunun incelenmesi,
7. Richard link fonksiyonunun reel düzlemin bir bandında ötelenmesi,
8. Richard link fonksiyonunun şekil parametresinin ötelemeye göre değiştirilmesi.

Yukarıda sekiz başlık altında toplanan analizler konunun daha ilerilere taşınabilmesi için sunulan basamaklardır. Özellikle sarmal şekle getirilen Richard dağılımı istatistik literatürü içerisinde yeni yeni esaslandırılmaya başlayan dairesel veriler için oldukça kullanışlı bir sarmal dağılımdır. Diğer sarmal dağılımlardan ayrılan önemli özelliği dağılım fonksiyonunun katsayısının açık şekilde yazılabilmesidir. Bunun aksine önemli sarmal dağılımlardan olan Von Misses dağılımının normalleyen katsayısı Bessel denkleminde alınan Bessel fonksiyonudur. Bessel fonksiyonunun analitik yapısının zorluğu bu dağılım fonksiyonunun kullanılabilirliğini güçleştirmektedir.

Bunun yanı sıra alt aralıklarda tanımlanan Richard dağılımlarının sürekli bir şekilde birbirlerine eklenebilmesi de bu dağılım için oldukça önemli bir özelliktir. Bu sayede bir zaman serisi verisinin kırılma noktalarına göre incelenebilmesi analize oldukça büyük bir değer katmaktadır.

Ayrıca Richard link fonksiyonlarının ötelenmesinde şekil parametresinin değişken olarak alınması verinin dağılımının yanı sıra şekil parametresinin dağılımının da önemini ortaya koymuştur.

KAYNAKLAR

- [1] Abramowitz, M. and Stegun, I.A. (Eds.) (1964) "Handbook of Mathematical Functions, Applied Mathematics Series", Vol. 55. National Bureau of Standards, US Department of Commerce, Washington, DC.
- [2] Albert, A. & Anderson, J. A. (1984) "On the Existence of Maximum Likelihood Estimates in Logistic Regression Models" *Biometrika*, 71, 1, 1-10.
- [3] Aber, M. S., & McArdle, J. J. (1991) "Latent growth curve approaches to modeling the development of competence". In M. Chandler & M. Chapman (Eds.), *Criteria for competence: Controversies in the conceptualization and assessment of children's abilities* Hillsdale, NJ: Lawrence-Erlbaum (231-258).
- [4] A. Gregorczyk (1998) "Richards Plant Growth Model" *J. Agronomy and Crop Science*, 181, 243-247.
- [5] Beltran J.J., Butts W.T. Jr., Olson T.A. and Koger M. (1992), "Growth patterns of two lines of angus Calve selected using predicted growth parameters". *J. Anim. Sci.* 70, 734-741.
- [6] Bertalanffy, L. von, (1934). "Untersuchungen über die Gesetzlichkeit des Wachstums", I. Allgemeine Grundlagen der Theorie mathematische und physiologische Gesetzlichkeiten des Wachstums bei Wassertieren. *Arch. Entwicklungsmech*, 131, 613-652.
- [7] Cramer, J.S. (2001) "An Introduction to Logit Model for Economists" ikinci baskı, yayıncı: Edward Arnold , ISBN 0953339467, 9780953339464, 113 sayfa.
- [8] Cramer, H. (1946) "Mathematical Methods of Statistics" Princeton University Press: New Jersey.
- [9] Delgado J. and Pena J.M., (2003) "A Linear Complexity Algorithm for the Bernstein Basis", In *IEEE Int. Conf. on Geometric Modeling and Graphics (GMAG)*, 3, 162-167.
- [10] G.Tachev, (2008), "Voronovskaja's Theorem Revisited", in *Journal of Mathematical Analysis and Applications*, 343, 399-404.
- [11] G.Tachev, (2013) "The Distance between Bezier Curve and its Control Polygon", 7380 p. Jubilee Collection of Papers Dedicated to the 60th Anniversary of Prof. Mihail M. Konstantinov.

- [12] G.Tachev, (2010) "On the second moment of rational Bernstein functions", *Journal of Computational Analysis and Applications*, JoCAAA, 12, no:2, 471-479.
- [13] G.Tachev, (2012) "Voronovskaja Theorem for Schoenberg operator", *Mathematical Inequalities and Applications*, MIA, 15, no:1, 49-59.
- [14] G.Tachev, (2012) "New Estimates in Voronovskaja's theorem", *Numerical Algorithms*, 59, 119-129.
- [15] G.Tachev, (2012) "The Complete Asymptotic Expansion for Bernstein Operators", *Journal of Math. Analysis and Applications*, JMAA, 385, 1179-1183.
- [16] G.Tachev, (2012) "Pointwise Approximation by Bernstein Polynomials", *Bulletin of the Australian Mathematical Society*, BAMS, 85, no: 2, 353-358.
- [17] G.Tachev, (2012) "The Rate of Approximation by Rational Bernstein Functions in Terms of Second Order Moduli of Continuity", *Numerical Functional Analysis and Applications*, 33(2), 206-215.
- [18] G. Tachev, (2014) "Approximation of Bounded Continuous Functions by Linear Combinations of Phillips Operators", *Demonstratio Mathematica*, vol.4.
- [19] Gurcan M. and Oner Y. (2001) "On The Existence Problem in Logistic Regression Models By Alternative Form". *Adv.& Appl. in Stat*; 165-174.
- [20] Gurcan M. and Colak C. (2009) "The Polynomial Approach in Growth Curves" *Turkiye Klinikleri Bioistatistik Dergisi*, 1, 2, 54-62.
- [21] Gurcan M., Colak C. and Orman M.N. (2010) "Bernstein Polynomial Approach Against to Some Frequently Used Growth Curve Models on Animal Data", *Pak. J. Statist.* 26, 3, 509-516.
- [22] Gurcan M. and Colak C. (2011) "Generalization of Korovkin Type Approximation by Appropriate Random Variables & Moments and an Application in Medicine", *Pak. J. Statist.* 27, 3, 283-297.
- [23] Gompertz, Benjamin (1825). "On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies". *Philosophical Transactions of the Royal Society of London* 115, 513–585.
- [24] H.Gonska and G.Tachev, (2009), "A Quantitative Variant of Voronovskaja's Theorem", *Results in Mathematics*, 53, 287-294.
- [25] H.Gonska and G.Tachev, (2011) "Gruss-type inequality for positive linear operators with second order Ditzian-Totik moduli", *Mat. Vesnik*, 63, no: 4, 247-252.

- [26] Hox, J. (2002) "Multilevel Analysis. Techniques and Applications", Copyright 2010 by Routledge of the Taylor & Francis 380 sayfa.
- [27] I. I. Gigan and A. V. Skorohod (1965). "Introduction to the Theory of Random Processes" Fizmatgiz.
- [28] Joel L. Horowitz and Enno Mammen (2004), "Nonparametric Estimation of an Additive Model with a Link Function", Volume 32, Number 6, 2412-2443.
- [29] McArdle, J. J. & Epstein, D. (1987). "Latent Growth Curves With in Developmental Psychology, Child Development", JSTOR, 58, 110-133.
- [30] McArdle, J. & Nesselroade, J. R. (2003). "Growth Curve Analysis in Contemporary Psychological Research". J. A. Schinka & W. F. Velicer (Eds.), Handbook of Psychology Vol. 2, 447-480, Hoboken, NJ: Wiley & Sons.
- [31] Oztuna D., Ates C., Gultekin S.S. and Genc Y. (2011) "Estimation of Area Under Receiver Operating Characteristic (ROC) Curve for Clustered Data", Turkiye Klinikleri Bioistatistik Dergisi, 3, 2, 57-62.
- [32] Oner Y., Gurcan M. and Halisdemir N. (2005) "On Continuous Deformation of Richards Family", International Journal of Pure and Applied Mathematics, 375-377.
- [33] Paul C. Hoel, Sidney C. Port, Charles J. Stone (1972). "Introduction to Stochastic Processes" Houghton Miffling Company, Boston.
- [34] Richard F.J. (1959) "A Flexible Growth Function for Emprical Use" J. Exp. Bot. 10, 290-300.
- [35] Roush W.B., Dozier W.A. 3rd, and Branton S.L. (2006) "Comparison of Gompertz and Neural Network Models of Broiler Growth", Poult Sci. 85, 4, 794-801.
- [36] Shahbazov, A. (2005) "Olasılık Teorisine Giriş", Birsen Yayınevi, İstanbul.
- [37] Sayer, A. G. & Willett, J. B. (1998). "A Cross-Domain Model for Growth in Adolescent Alcohol Expectancies", Multivariate Behavioral Research, 33, 509-543.
- [38] Seber G.A.F. and Wild C.J. (1989) "Non Linear Regression", Jhon Wiley and Sons, Inc. New York.
- [39] Sorin Gal and Gancho Tachev, (2012) "On the Constant in The Lower Estimate for the Bernstein Operator", Mathematica Balkanica, 26.
- [40] Scott Menard (2002) "Applied Logistic Analysis", Sage Publications Inc. California, 106 p.

- [41] Laird A. K. (1964) "Dynamics of Tumor Growth". Br J of Cancer 18, 3, 490–502.
- [42] Y.C. Lei, S.Y. Zhang (2004), "Features and Partial Derivatives of Bertalanffy-Richards Growth Model in Forestry", Nonlinear Analysis: Modelling and Control, Vol. 9, No: 1, 65–73.

ÖZGEÇMİŞ

1983 yılında Elazığ'da doğan Arzu Ekinci Demirelli, ilk orta ve lise öğrenimini Elazığ'da tamamladıktan sonra 2001 yılında Anadolu Üniversitesi İstatistik bölümünü kazanmıştır. 2006 yılında mezun olan yazar, 2007 yılında Şekerbak T.A.Ş Elazığ Şubesi'nde Ticari Krediler Müşteri Temsilcisi olarak göreve başlamıştır. Bankadaki görevinin yanı sıra 2008 yılında Fırat Üniversitesi İstatistik anabilim dalında yüksek lisansa başlamıştır. 2010 yılı Ocak ayında yüksek lisansı bitirdikten sonra aynı yılın Şubat ayında doktora başlamıştır. 2011 yılında Ankara Ümitköy Şubesinde Yetkili Müşteri Temsilcisi olarak görev aldıktan sonra 2012 Eylül ayında bankadaki görevinden istifa ederek, doktora çalışmalarına ağırlık veren yazar, evli ve bir çocuk annesidir.

Arzu Ekinci Demirelli