



# Lojistik İncelemede Ayrımsama Performansının Değerlendirilmesi

Mehmet GÜRCAN<sup>1\*</sup>, Mehmet Onur KAYA<sup>2</sup>, Nurhan HALİSDEMİR<sup>3</sup>

<sup>1</sup> Fırat Üniversitesi, Fen Fakültesi, İstatistik Bölümü, Elazığ, Türkiye (ORCID: 0000-0002-3641-8113)

<sup>2</sup> Fırat Üniversitesi, Tıp Fakültesi, Biyoistatistik ve Tıp Bilişimi AD, Elazığ, Türkiye (ORCID: 0000-0001-8052-0484)

<sup>3</sup> Fırat Üniversitesi, Fen Fakültesi, İstatistik Bölümü, Elazığ, Türkiye (ORCID: 0000-0003-2151-7917)

(İlk Geliş Tarihi 18 Kasım 2019 ve Kabul Tarihi 19 Aralık 2019)

(DOI: 10.31590/ejosat.654731)

**ATIF/REFERENCE:** GürCAN, M., Kaya, M. O., & Halisdemir, N. (2019). Lojistik İncelemede Ayrımsama Performansının Değerlendirilmesi. *Avrupa Bilim ve Teknoloji Dergisi*, (17), 1008-1013.

## Öz

Lojistik inceleme uygulamalı istatistiğin önemli materyallerinden birisidir. Veri yapısına bağlı olarak ayrımsama performansı önemli ölçüde etkilenebilir. Ayrımsama performansının istenen düzeyde olması kullanılan bağlantı fonksiyonunun esnekliğine doğrudan bağlıdır. Lojistik modeller de genellikle bağlantı fonksiyonu olarak büyüme eğrileri kullanılmaktadır. Kullanılan büyüme eğrisi veriyi adapte edilebilirse modelin ayrımsama performansı önemli ölçüde artabilir. Bazı durumlarda lojistik modellerde parametre tahminleri en çok olabilirlik yöntemiyle yapılamamaktadır. Bu durum araştırmacılar için önemli bir güçlük oluşturmaktadır. Yapılan çalışmada lojistik ayrımsamanın performansını artırabilmek için bağlantı fonksiyonu olarak Richard eğrisi kullanılmıştır. Buna bağlı olarak lojistik ayrımsamanın nasıl yapılacağı incelenmiştir. Büyüme eğrileri genel olarak birinci mertebeden diferansiyel denklemin çözümü olarak elde edilir. Dolayısıyla eğrinin parametreleri integral sabitleridir. İntegral sabitinin doğrusal olasılık modeline bağlı olarak modele yerleştirilmesi teorik olarak önemli bir bulgudur. Bunun yanı sıra diferansiyel denklemin oluşturduğu fonksiyonların seçiminde bulunulması da mümkündür. İlk olarak çalışmada lojistik ayrımsamada kullanılan Richard bağlantı fonksiyonunda bu şekilde bulunan parametrenin hesaplanması amaçlanmıştır. Eğrinin büküm parametresi hatalı sınıflanan gözlemler incelenerek bağımsız değişken değerleri ile ilişkilendirilerek, lojistik ayrımsama dinamik bir yapıya kavuşturulmuştur. Lojistik inceleme için tasarlanan bu yapılar anksiyete bozukluğu olgularına ait 20, 40 ve 60 birimlik veri setlerinde değerlendirilerek yöntemin uygulanabilirliği ve ayrımsama performansı gösterilmiştir. İki grup lojistik modelde, yani bağımlı değişkenin iki değer aldığı lojistik modelde doğru sınıflama oranları 20, 40 ve 60 birimlik olgular için sırasıyla, %72,5, %66,3 ve %67,5 olarak bulunmuştur. Buna karşılık Richard bağlantı fonksiyonu kullanıldığında, hatalı sınıflanan gözlemler için eğrinin büküm parametresinin uygun seçilebilmesiyle bu oranlar sırasıyla, %82,5, 78,75 ve % 80'e kadar artırılabilmiştir. Yapılan uygulamadan da görüldüğü kadarıyla bağlantı fonksiyonunun büküm noktasının kontrol edilebilmesi lojistik incelemeyi dinamik bir yapıya sahip kılmaktadır. Buna ilaveten büküm parametresinin değiştirilebilmesi doğrusal olasılık modeline dokunulmaksızın yapılabilen ve bu sayede hatalı bulunan gözlemler içerisinde doğru sınıflama yapılabilir. Bu bağlamda büküm parametresinin bağımsız değişken değerlerine bağlı olarak hesaplanması bağımlı değişkenin olasılık değerlerini etkileyecek artımlarda bağlantı fonksiyonunun bu değişimi doğru bir şekilde yansıtmasını sağlar. Dolayısıyla bu durumların tümü dikkate alındığında lojistik ayrımsama modeli dinamik bir yapıya sahip olmuş olur.

**Anahtar Kelimeler:** Lojistik Model, Richard Modeli, Ayrımsama Sorunu.

## Assessment of Discrimination Performance in Logistics

### Abstract

Logistics analysis is one of the important materials of applied statistics. Discrimination performance can be significantly affected depending on the data structure. The desired level of discrimination performance is directly dependent on the flexibility of the link

\* Sorumlu Yazar: Fırat Üniversitesi, Fen Fakültesi, İstatistik Bölümü, Elazığ, Türkiye ORCID: 0000-0002-3641-8113, [mgurcan@firat.edu.tr](mailto:mgurcan@firat.edu.tr)

function used. In logistics models, growth curves are generally used as link functions. If the growth curve used can be adapted to the data, the discrimination performance of the model can be significantly improved. In some cases, parameter estimation in logistics models cannot be obtained by likelihood method. This issue a significant problem for researchers. In this study, we investigated how to do logistic discrimination by using Richard link functions. A parameter of logistic curve obtained from the solution of differential equation is an integral constant. The finding of the integral constant is possible that by means of linear probability model or the selection of functions that consist of the differential equation. Richard link function used logistic discrimination have intended to use for the calculation of the parameter. The incorrectly classified observations are examined by twist parameter of the curve. So, this twist parameter associated with the independent variable values and logistic discrimination adapt to a dynamic structure. This adaptation of applicability in method have emphasized a data set which contain people that have anxiety disorders 60 unit individuals with have not disorders 60 unit individuals. It is found to be correct classification rate of 67.5% when the dependent variable has binary logistic model. While using Richard link function, it is found correct classified rate up to %100 by means of the suitable twist parameter of the curve. As shown in the analysis, Richard link function has dynamic structure. The exchangeable of the inflection parameters can be made protecting Linear probability model, and thus in incorrectly of the observations can be correct classification of observations. The calculation of twist parameter as the independent variable values will affect the value of the probability, but as using the link function will be tolerate this change. Thus, the logistic discrimination model have converged a dynamic structure.

**Keywords:** Logistic model, Richard model, a problem of discrimination.

## 1. Giriş

Lojistik regresyon modellerinin kullanılması 1845’li yıllara kadar dayanmaktadır. Bu alandaki ilk çalışmalar  $y = y(x)$  eğrisinin diferansiyelinin  $c$  reel sabit olmak üzere aşağıdaki şekilde düzenlenmesiyle başlamıştır,

$$\frac{dy(x)}{dx} = cy(x)$$

Burada amaç fonksiyonun artım hızının yine fonksiyonun o noktadaki değerine bağlayabilmektir. Örnek olarak bir bölgedeki nüfus artış hızı doğal olarak bölgenin mevcut nüfusuna bağlı olacaktır. Ancak fonksiyonun artımı belli bir noktada doyuma ulaşabilmeli ve belli bir andan sonra sabitleşebilmelidir.

Daha sonraları bu diferansiyel denklemin çözümünü oluşturan eğrilerin sınırlı olabilmesi için reel bir  $\alpha$  üst sınırı denkleme eklenerek aşağıdaki şekilde düzenlenmiştir,

$$\frac{dy(x)}{dx} = cy(x)[\alpha - y(x)]$$

Elde edilen bu son denklemin çözümünden  $\alpha = 1$  için aşağıdaki lojistik fonksiyon elde edilir [1, 2],

$$y(x) = \frac{\exp(c_1 + c_2x)}{1 + \exp(c_1 + c_2x)}$$

Burada  $c_1$  ve  $c_2$  reel sabitlerdir. Yukarıdaki eşitliğin tersi düşünüldüğünde doğrusal olasılık modeli aşağıdaki şekle dönüşür,

$$\begin{aligned} \ln \frac{y}{1-y} &= c_1 + c_2x \\ &= A(x) \end{aligned}$$

Doğrusal olasılık modelinin ikinci türevinin sıfıra eşit olduğu noktanın bulunması önemlidir. Bu nokta doğrusal olasılık modelinin büküm noktası olup lojistik ayırmsamanın yapılabilmesi için gereklidir. Birçok doğrusal olasılık modelinde büküm noktası  $y = \alpha/2$  olarak elde edilmektedir.

Diferansiyel denklem formunda verilen lojistik modeller genel olarak aşağıdaki şekilde genelleştirilebilir [3],

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

Burada  $h$  ve  $g$  fonksiyonları  $h(0) = g(0) = 0$  şartını sağlayacak şekilde artan fonksiyonlardır. Ayrıca  $x \rightarrow \infty$  ve  $x \rightarrow -\infty$  için  $y'(x) \rightarrow 0$  özelliği sağlanmaktadır. Özel olarak  $g(y) = cy$  ve  $h(y) = y$  olarak seçilirse denklemin çözümünden doğrusal olasılık modeli elde edilebilecektir. Yine bu fonksiyonların özel seçimleriyle Gompertz, Von Bertalanffy ve Richard lojistik modelleri gibi özel modeller elde edilebilmektedir. Richard eğrileri lojistik modeller içerisinde özel bir yere sahip olup bu modelde parametre sayısı fazla olduğundan lojistik ayırmsamada daha iyi ayırmsama yapabilmektedir. Richard eğrisi genel diferansiyel denklem yapısında

$$\begin{aligned} h(y) &= y^{m-1} \\ g(y) &= \frac{A'(x)y}{(m-1)\alpha^{m-1}} \end{aligned}$$

fonksiyonlarının özel seçimleriyle aşağıdaki diferansiyel denklemin

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

çözümünden aşağıdaki şekilde elde edilir,

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

Bu modelde  $m = 2$  alındığında standart lojistik model rahatlıkla elde edilebilir. Richard eğrisinin ikinci türevi aşağıdaki formdadır,

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

Bu eşitlik kullanılarak  $y''(x) = 0$  için  $y_0 = \alpha m^{1/(1-m)}$  değeri dikkate alınacak olursa

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

noktası bulunmuş olur. Bu nokta Richard eğrisinin büküm noktasıdır.

## 2. Materyal ve Metot

### 2.1. Richard Modellerinde Parametrelerin Tahmini

Lineer olmayan modellerde model parametrelerinin tahmini için sıklıkla kullanılan yöntem en küçük kareler çözümüne lineer yaklaşımdır. Bu yaklaşımı Richard modeli için kullanacak olursak Richard eğrisini sadece lineer olacak şekilde Taylor serisine açmamız yeterli olacaktır. Bu durumda Richard eğrisini yaklaşık olarak temsil eden birinci dereceli Taylor polinomu aşağıdaki şekilde elde edilir [3,6],

$$y(x) \approx y(A_0(x, \theta)) + y'(A_0(x, \theta))(\theta - \theta_0)$$

Bu yaklaşım en küçük kareler çözümünde kullanılacak olursa hata kareler toplamı aşağıdaki şekilde olur,

$$HKT = \sum_{j=1}^n \{y_j - [y(A_0(x, \theta)) + y'(A_0(x, \theta))(\theta - \theta_0)]\}^2$$

Hata kareler toplamının parametrelere göre türevlerinden elde edilen denklem sistemi çözülerek modelin katsayıları bulunabilir. Yöntem ardışık olarak tekrarlanarak her defasında çözüme daha yakın değerler elde edilecektir.

$A(x, \theta) = a + bx$  şeklinde alınsın. Bu durumda denklem sistemi aşağıdaki şekilde olacaktır [4, 5, 7],

$$\sum_{j=1}^n \left\{ y_j - [y(a_0, b_0) + \frac{dy}{da}(a - a_0) + \frac{dy}{db}(b - b_0)] \right\} \frac{dy(a_0, b_0)}{da} = 0$$
$$\sum_{j=1}^n \left\{ y_j - [y(a_0, b_0) + \frac{dy}{da}(a - a_0) + \frac{dy}{db}(b - b_0)] \right\} \frac{dy(a_0, b_0)}{db} = 0$$

Richard eğrilerinde model katsayılarını barındıran doğrusal yapı  $A(x)$  kısmıdır. Bunun yanı sıra modelde doğrusal yapının dışında birde  $m \neq 1$  katsayısı bulunmaktadır. Eğri bu parametreye göre üstel olarak düzenlendiğinden bu parametre lineer yaklaşım yardımıyla tahmin edilememektedir. Bu parametrenin seçimi modelin kullanımını zorlaştırmakla birlikte modelin daha iyi tahmin yapabilme gücünü de artırmaktadır. Hatalı sınıflanan bazı gözlem değerlerini doğru sınıflayabilen bir  $m$  değeri bulmak mümkündür. Bu bakımdan sınıflama güclüğü bulunan sınır noktaları göz önüne alındığında bazı noktalar için  $m$  değeri bulunabilir.  $m$  değerinin belli bir aralıkta sürekli olarak deformasyonu analize bu kolaylığı sağlayabilecektir. Bir  $\tau$  parametresini,

$$\tau = \frac{1}{1 + e^{-m}}$$

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.

### 2.2. Veri Yapısı

Çalışmanın verileri, Eylül 2010 ile Kasım 2012 tarihleri arasında, Uludağ Üniversitesi Tıp Fakültesi Psikiyatri Bölümünden toplanmıştır. Çalışmaya katılmış olan 120 olgunun, 60'ı kontrol ve 60'ı ise anksiyete bozukluğu olan hasta grubunu oluşturmaktadır. Çalışmada bağımlı değişken ( $y$ ), hastalık olup olmaması iken, bağımsız değişkenler ise yaş ( $x_1$ ), cinsiyet ( $x_2$ ), medeni durum ( $x_3$ ), meslek ( $x_4$ ), eğitim durumu ( $x_5$ ), kronik hastalık ( $x_6$ ), ailede psikiyatrik bozukluk ( $x_7$ ), ailede kronik hastalık ( $x_8$ ), STAI-I ( $x_9$ ) ve STAI-II ( $x_{10}$ ) ölçekleridir.

## 3. Araştırma Sonuçları ve Tartışma

Richard eğrilerinde mevcut olan parametreler klasik bağlantı fonksiyonlarına göre daha fazla olduğundan iki grup ayrımsama yapılan lojistik verilerde daha optimal sonuçlar vermektedir. Bundan dolayı yapılacak olan analizde lojistik bağlantı fonksiyonu olarak

Richard eğrisinin kullanılması lojistik ayırimsama yüzdesini daha yüksek hale getirecektir. Veri lojistik bağlantı fonksiyonu yardımıyla ayırimsamaya tabi tutulduktan sonra doğrusal olasılık modelleri aşağıdaki şekilde elde edilmiştir,

$$A_{20}(x) = 2,0335 - 0,0189x_1 + 0,4792x_2 - 0,0343x_3 - 0,9792x_4 - 0,9645x_5 + 1,1460x_6 + 0,2098x_7 - 0,0095x_8 - 0,2480x_9 + 0,2447x_{10}$$

$$A_{40}(x) = 0,7663 - 0,0024x_1 + 0,1888x_2 - 0,0503x_3 - 0,5845x_4 - 0,0107x_5 + 0,1240x_6 + 0,0225x_7 - 0,7878x_8 - 0,0002x_9 + 0,0257x_{10}$$

$$A_{60}(x) = -1,6260 - 0,0124x_1 + 0,7535x_2 - 0,0107x_3 - 0,8051x_4 - 0,1976x_5 + 0,0312x_6 + 0,8865x_7 - 0,7417x_8 + 0,0535x_9 + 0,0295x_{10}$$

Bu durumda lojistik model aşağıdaki şekilde tahmin edilmiştir,

$$\hat{y}_{20} = \{1 + 0,2383 \exp(-A_{20}(x))\}^{-4,1963}$$

$$\hat{y}_{40} = \{1 + 0,6426 \exp(-A_{40}(x))\}^{-1,5561}$$

$$\hat{y}_{60} = \{1 + 0,2424 \exp(-A_{60}(x))\}^{-4,1249}$$

Tahmin modeli yardımıyla elde edilen doğrusal olasılık modelinin değerleri ve y değerlerinin tahmini olasılıkları aşağıdaki tabloda olduğu gibi hesaplanmaktadır,

Tablo 1. 60 birimlik veri için hatalı sınıflanan gözlemlerin tahmin değerleri

Sıra No:	Gözlem No:	A(x)	$\hat{y}(x)$	y	m	$\hat{y}_m$
1	1	1,4492	0,7969	0	-1,2901	0,0508
2	2	2,4982	0,9222	0	-2,3698	0,0728
3	4	3,6440	0,9744	0	-6,2645	0,1904
4	10	0,8056	0,6558	0	0,1426	0,1248
5	13	1,4492	0,7969	0	-1,6163	0,1929
6	16	2,4982	0,9222	0	-2,3948	0,1410
7	25	2,0780	0,8846	0	-3,8431	0,1707
8	29	1,8835	0,8621	0	-1,1769	0,0617
9	32	3,6440	0,9744	0	-6,2645	0,1544
10	35	2,0780	0,8846	0	-3,8409	0,1208
11	50	0,8316	0,6625	0	-0,6679	0,0878
12	57	1,8835	0,8621	0	-0,6879	0,1968
13	63	-0,6827	0,1995	1	76,3826	0,9363
14	65	-0,1438	0,3627	1	94,2889	0,9437
15	71	-1,0627	0,1126	1	47,5533	0,9057
16	78	-0,1425	0,3635	1	54,3558	0,9161
17	84	-0,7899	0,1726	1	94,8564	0,9474
18	87	-0,3482	0,2972	1	82,7804	0,9465
19	89	0,3132	0,4920	1	98,8593	0,9545
20	91	-0,3753	0,2890	1	88,2001	0,9449
21	96	-0,0519	0,3932	1	64,1917	0,9283
22	100	-0,4119	0,2777	1	81,4546	0,9398
23	108	-0,4119	0,2777	1	95,3915	0,9469
24	111	-0,4308	0,2712	1	96,5696	0,9530

Hazırlanan tabloya dikkat edildiğinde 24 adet gözlem analiz sonucu hatalı olarak sınıflanmıştır. Hatalı sınıflanan gözlemlerden 12 tanesi sıfır grubunda olması gerekirken bir grubuna atanmış 12 tanesi ise bir grubuna atanması gerekirken sıfır grubuna atanmıştır. Geri kalan 96 adet gözlem ise doğru olarak sınıflandırılmıştır. Bu durumda kullanılan lojistik modelin doğru sınıflama oranı %80 olmaktadır.

Toplam 120 adet gözlemden 24 tanesi modelde istenilen tahmini olasılık değerini vermemektedir. Ancak üstel kısmı  $A(x)$  şeklinde hesaplanan Richard eğrisi tanımlı olduğu bölge üzerinde  $m$  parametresi değiştirilerek hareket ettirilebilmektedir. Bu hareket ettirme esnasında hatalı olarak sınıflanan 24 adet gözlemin tamamı  $m$  değerinin belirli sabitlerinde istenilen olasılık değerlerine sahip olabilmektedirler. Bu durumda hatalı gözlemleri yeniden analiz ederek doğru sınıflama şansına sahip olmaktadır.  $\hat{y}$  değeri Richard eğrisi yardımıyla hesapladığımız tahmini olasılık değeri olmak üzere aşağıdaki denklemi Richard eğrisinden yazabiliriz,

$$\hat{y}^{1-m} + (1 - m) \exp(-A(x)) - 1 = 0$$

Bu denklemin  $m$  parametresine göre çözümü incelenen gözlem için doğru sınıflamayı yapabilen modeli oluşturacaktır. Denklemin ifadesine dikkat edilecek olursa  $m = 1$  denklemin aşık çözümüdür. Ancak  $m \neq 1$  olması gerektiğinden denklemin 1 değerinden farklı çözümlerinin bulunması gerekmektedir. Bazı olasılık değerleri için bu çözüm bulunabilmekte bazı değerler için ise bulunamamaktadır. Tabloda denklemin tüm hatalı gözlemleri için  $m$  değeri hesaplanabilmektedir. Bu sayede hatalı sınıflanan gözlem sayısı 24 gözlemden 0 gözleme indirgenmiş ve doğru sınıflama yüzdesi ise %80'den %100'e yükseltilebilmiştir. Bu gözlemlerin  $m$  değerleri ve olasılık tahminleri tablonun son iki sütununda yer almaktadır.

Tabloda yer alan  $m$  değerleri ilgili gözlemlerin değişkenleri dikkate alındığında aşağıdaki gibi lineer şekilde ifade edilebilmektedir,  $m = 62,004 + 0,605X_1 - 2,359X_2 - 4,748X_3 + 14,144X_4 + 23,463X_5 - 3,581X_6 - 24,440X_7 + 6,576X_8 - 3,230X_9 + 0,817X_{10}$

Var olan bu ilişkinin Richard bağlantı fonksiyonunda kullanılması  $m$  parametresinin gözlem değerine göre değişmesini ve eğrinin daha optimal ayrımsama yapabilmesini sağlamaktadır.

Diğer taraftan çalışmadan elde edilen doğru sınıflandırma oranları ROC analizi ile değerlendirildiğinde elde edilen bulgular aşağıdaki tabloda yer almaktadır.

Tablo 2. Bağlantı Fonksiyonlarına Ait ROC Analizi Sonuçları

Değişkenler	Eğri Altında Kalan Alan (AUC)	%95 Güven Sınırları		p-değeri
		Alt Sınır	Üst Sınır	
Lojistik Bağlantı Fonksiyonu	0,780	0,698	0,862	<0,001
Richard Bağlantı Fonksiyonu	0,791	0,707	0,875	<0,001

Tablo 2 incelenecek olursa, her iki bağlantı fonksiyonunun da eğri altında kalan alan değerlerinin istatistiksel olarak anlamlı olduğu görülecektir. Bununla beraber Richard fonksiyonuna ait alanın daha büyük olduğu da görülmektedir.

## 4. Sonuç

Tıbbi bir veri incelenirken amaç hastalığın teşhis edilmesi olduğunda veya bazı değişkenlerin hastalığın teşhisinde önemli olup olmadığını belirlenmesi olduğunda yapılan en önemli analizlerden birisi lojistik ayrımsama. Bu çalışmada da lojistik ayrımsama tekniği Richard bağlantı fonksiyonları yardımıyla yapılmaya çalışılmıştır.

Lojistik regresyon analizinin diğer ayrımsama tekniklerine göre birçok üstün yönleri bulunmasına karşılık lojistik bağlantı fonksiyonu ve doğrusal olasılık modeli birbirlerine eşitlendiğinde bağlantı fonksiyonu dönüşümünün yapısından dolayı bağımsız değişkenin küçük farklılıklar göstermesi durumunda olasılık değerleri bu farklılıkları tolere edememektedir. Bu yüzden lojistik ayrımsamada aşırı hatalara neden olunabilir. Bu problemin çözülmesi oldukça güçtür. Bu aşamada araştırmacının en kolay alternatifi bağımsız değişken sayısının azaltılması yani aşırı farklılıklara neden olan ve tolere edilemeyen değişkenlerin modelden çıkartılması olmaktadır. Tıbbi bir veri yapısında herhangi bir hastalığın teşhisinde bağımsız değişkenlerin doğru olarak seçilmesi oldukça önemlidir. Lineer model yapısında bağımsız değişken sayısının artırılması modelin uyumunu artırırken ve aynı zamanda yanlış seçilmiş bir bağımsız değişkenin modelde etkisinin olmaması sağlanabilirken lojistik ayrımsamada bu mümkün olmamaktadır. Bu yüzden lojistik ayrımsamalarda ilgili bağımsız değişkenlerin seçimi oldukça önemlidir.

İncelemesi yapılan 60 anksiyete bozukluğu ve 60 da kontrol olmak üzere 120 olgunun verilerinin lojistik bağlantı fonksiyonu ile ayrımsaması yapıldığında modelin doğru ayrımsama yapma oranı düşük (%67,5) çıkmaktadır. Bu aşamada Richard bağlantı fonksiyonu kullanıldığında doğru sınıflama oranı %80'e ulaşmaktadır. Yapılan uygulamada hatalı sınıflanan gözlemler tekrar ele alınarak bağlantı fonksiyonunun büküm parametresinin uygun seçimiyle ayrımsama performansını büyük ölçüde artırılmıştır. Yapılan analizdeki gibi Richard bağlantı fonksiyonu, lojistik modeli dinamik bir yapıya çevirir. Büküm parametresinin değiştirilebilmesi önemli bir teorik bulgudur. Bu sayede doğrusal olasılık modeline dokunulmaksızın hatalı bulunan gözlemlerin içerisinde doğru sınıflama gerçekleştirilebilir. Büküm parametresinin bağımsız değişken değerlerine bağlı olarak hesaplanması olasılık değerini etkileyecek

artımlarda bağlantı fonksiyonunun bu değişimi doğru bir şekilde yansıtabilmesini sağlar. Bu sayede lojistik ayrımsamanın performansı artırılabilecektir.

## **Kaynakça**

- [1] Öner, Y., Gürcan, M. & Halisdemir, N. (2005). On Continuous Deformation of Richards Family. *International Journal of Pure and Applied Mathematics*, 375-377.
- [2] Gürcan, M. & Öner Y. (2001) On The Existence Problem in Logistic Regression Models By Alternative Form", *Adv.& Appl. in Stat.*, 165-174.
- [3] Seber, G.A.F. & Wild, C.J. (1989). *Non Linear Regression*. Jhon Wiley and Sons, Inc. New York.
- [4] Beltran, J.J., Butts, W.T. Jr., Olson, T.A. & Koger, M. (1992). Growth patterns of two lines of angus Calve selected using predicted growth parameters. *J. Anim. Sci.*, 70, 734-741.
- [5] Roush, W.B., Dozier, W.A. 3rd, & Branton, S.L. (2006). Comparison of Gompertz and neural network models of broiler growth. *Poult Sci.*, 85(4), 794-7.
- [6] Delgado J. & Pena J.M. (2003). A linear complexity algorithm for the bernstein basis. In *IEEE Int. Conf. on Geometric Modeling and Graphics (GMAG'03)*, 162-167.
- [7] Kaya, MO. & Kilic, MB. (2008) *Lojistik Modelde Bernstein Polinomlarının Kullanılması*. VI. İstatistik Günleri Sempozyumu, Samsun.