

# MARMARA ÜNİVERSİTESİ ÖĞRENCİLERİNİN KREDİ KARTI SAHİBİ OLMALARINI ETKİLEYEN FAKTÖRLERİN BAYESÇİ LOJİSTİK REGRESYON YARDIMIYLA İNCELENMESİ

Esin AVCI\*

## ÖZET

*Bayesci yaklaşım, verilerden elde edilen yeni bilgi ile önceden bilinen bilginin derlenmesi ile oluşan bir yöntemdir. Bu çalışmada klasik yaklaşıma alternatif olan Bayesci yaklaşım, ikili sonuç değişkeni ile etki eden değişken(ler) arasındaki sebep-sonuç ilişkisini ortaya çıkaran Lojistik regresyona uygulanmıştır. Bu amaçla Marmara üniversite öğrencilerinin kredi kartına sahip olma durumlarına etki eden sosyoekonomik ve demografik faktörler incelenmiştir.*

**Anahtar Kelimeler:** Kredi kartı sahipliği, Bayesci lojistik regresyon, WinBUGS.

## 1. GİRİŞ

Kredi kartı; mülkiyeti kendilerine ait olmak üzere banka ya da finansal kuruluşların müşterilerine önceden belirlenen limitlerde, anlaşmalı işyerlerinden yurtiçi ve yurtdışında mal ve hizmet satın alma ile nakit ödeme birimleri veya otomatik ödeme makinelerinden nakit çekimlerde kullanılmak amacıyla verilen karttır. Artık modern dünyada çağdaş bir ödeme sistemi olan ve “plastik para” olarak adlandırılan kredi kartı; kredi kartını veren banka veya kuruluşun açtığı krediye istinaden kart sahibinin gereksinim duyduğu mal veya hizmeti o anda bir ödeme yapmadan satın alınmasına ve bedelini daha sonra herhangi ek bir mali külfet yüklenmeksizin ödeme yapmasına imkan veren bir ödeme aracıdır (Çavuş, 2006).

Kredi kartlarının, nakit dolaşım ihtiyacını azaltması, ekonominin kayıt altına tutulmasını kolaylaştırarak kayıt dışı hareketlerin önlenmesi, genel ekonominin kartlar sayesinde kağıt yükünden kurtulması ve tasarruf-yatırım akışının hızlanması sonucu ticari faaliyetlerin canlanması gibi ekonomik faydaları da söz konusudur (Turgay ve Başgül, 2007).

Yeni gelişmeler paralelinde, kredi kartı sektörü de Türkiye’de hızlı bir büyüme trendi içerisine girmiştir. Türkiye’de hızla büyüyen kredi kartı sektörü, üniversite öğrencilerini de müşterileri arasına dahil edebilmek için yarışmaktadır. Üniversite öğrencilerinin kredi kartı kullanım tercihlerinin, sorunlarının, bu konudaki tutum ve davranışlarının tespit edilmesi bankalar açısından büyük önem taşımaktadır. Çünkü teknolojik gelişmelerle birlikte öğrencilerin kredi kartı sahipliği üzerinde etkili olan faktörler de değişiklik göstermektedir (Keskin ve Koparan, 2010).

Sosyal bilimlerde özellikle sosyo-ekonomik araştırmalarda, incelenen değişkenlerin bazıları hassas ölçülemlerle ölçülmekle beraber, bazıları da olumlu-olumsuz, başarılı-başarısız, evet-hayır gibi iki şıklı verilerden oluşmaktadır. İki şıklı veriler, kategorik verilerin en yaygın olarak kullanılan şeklidir. Bağımlı değişkenin iki şıklı kategorik veriler olması durumunda bağımsız değişkenle (veya değişkenlerle) bağımlı değişken

\*Yrd. Doç. Dr., Giresun Üniversitesi, Fen Edebiyat Fakültesi, İstatistik Bölümü, Giresun, e-posta: [esin.avci@giresun.edu.tr](mailto:esin.avci@giresun.edu.tr)

arasındaki sebep-sonuç ilişkisini incelerken Lojistik regresyon analizi kullanılır (Oktay vd., 2009). Lojistik regresyon modeli ilk olarak 1944 yılında Berkson tarafından kullanılmıştır. Cox (1970) bu modeli gözden geçirerek çeşitli uygulamalarını yapmış, özet gelişmeler ise ilk Anderson (1979, 1983) tarafından verilmiştir. Pregibon (1981) iki grup Lojistik modelde etkin (influential) aykırı (outlier) gözlemleri ve belirleme ölçütlerini, Lesaffre (1986), Lesaffre ve Albert (1989) ise çoklu grup Lojistik modellerde etkin ve aykırı gözlemlerle belirleme ölçülerini incelemiştir.

Lojistik regresyon modellerinin yaygın bir biçimde kullanılır hale gelmesi, katsayı tahmin yöntemlerinin geliştirilmesi ve Lojistik regresyon modellerinin daha ayrıntılı incelenmesine sebep olmuştur. Cornfield (1962), Lojistik regresyondaki katsayı tahmin işlemlerinde diskriminant fonksiyonu yaklaşımını ilk kez kullanarak popüler hale getirmiştir. Robert vd. (1987) Lojistik regresyonda standart Ki-kare, olabilirlik oran ( $G^2$ ), “pseudo” en çok olabilirlik tahminleri, uyum mükemmelliği ve hipotez testleri üzerine araştırmalar yapmışlardır. Duffy (1990) Lojistik regresyonda hata terimlerinin dağılışı ve parametre değerlerinin gerçek değerlere yaklaşmasını incelemiştir. Başarır (1990) klinik verilerde çok değişkenli Lojistik regresyon analizi ve ayırimsama sorunu üzerine çalışmıştır.

Son yıllarda klasik yöntemlerdeki kısıtlamalar Bayesci yaklaşıma olan ilgiyi arttırmıştır. Bayesci yaklaşım subjektif düşüncenin temel taşı olarak kabul edilen Bayes teoremine dayanarak geliştirilmiştir ve yaklaşıma göre parametreler klasik yaklaşımdaki gibi sabit olarak değil, olasılığa bağlı olarak tanımlanır. Dolayısı ile her bir parametreye ilişkin bir dağılım söz konusudur. Bu olasılıklar “kanaat derecesi (degrees of belief)” olarak tanımlanmaktadır. Bir başka deyişle parametreler rasgele değişken olarak ele alınmaktadır. “Bayesci” kelimesi de, parametre tahminleri için yapılan çıkarsamalarda Thomas Bayes’in teoreminden faydalanılmasından doğmuştur (İbrahim vd., 2001).

Bayesci yaklaşım, karmaşık veriyi modellemede önsel (prior) bilgiye başvurma esnekliği nedeniyle klasik yöntemlere göre oldukça avantajlıdır (İbrahim vd., 2001; Wong vd., 2005). Önsel bilginin elde edilmesi, Bayesci çıkarsamada önemli rol oynar. Önsel dağılımın sonuçları ne kadar değiştireceği model seçim kriterleri ile tespit edilmelidir. Önsel dağılımlar, açıklayıcı olan (informative) ve açıklayıcı olmayan (noninformative) olmak üzere iki temel gruba ayrılır. Açıklayıcı önsel bilgiler, daha önce yapılmış çalışmalardan elde edilen bilgiler, geçmiş deneyimler olarak belirlenirken, açıklayıcı olmayan önsel bilgiler ise parametrenin tanımlı olduğu aralık bilgisi dışında herhangi bir bilginin olmaması olarak tanımlanmaktadır. Bayesci yaklaşım, önsel bilgiler ışığında, gözlenen verinin subjektif yorumuna dayanır ve buradan hareketle elde edilen yeni bilginin bileşimi olan sonsal dağılımla açıklanır (Congdon, 2006; Wong vd., 2005). Hsu ve Leonard (1995) Lojistik regresyon fonksiyonlarında Bayes tahminlerinin elde edilmesi işlemleri üzerine çalışmışlar ve Bayesci Lojistik regresyonda Monte Carlo dönüşümünün kullanılabilirliğini göstermişlerdir.

Çalışmanın amacı, Bayes tanımından yola çıkarak var olan bilginin yeni bilgi ile güncellenmesini göstermektir. Bu amaçla 2011 yılında Gaziosmanpaşa ve İnönü üniversite öğrencilerinin kredi kartı sahipliğini etkileyen faktörleri belirleme çalışmasında uygulanan 24 soruluk anket, Marmara üniversitesinde okuyan 200 öğrenciye uygulanmıştır. Marmara üniversitesi öğrencilerinden derlenen veriler, önsel bilgi olarak kullanılan Gaziosmanpaşa ve İnönü üniversite öğrencilerinin verileri ile

birleştirilerek Bayesci Lojistik regresyon uygulanmış ve kredi kartı sahipliğine etki eden faktörler belirlenmiştir.

## 2. YÖNTEM

### 2.1 Lojistik regresyon

İstatistiksel modellerin kullanıldığı birçok bilimsel araştırmada sonuçların analiz edilmesinde en çok lineer olmayan modeller kullanılmaktadır. Lojistik regresyon modeli lineer-olmayan modellerden en önemlilerindedir. Lojistik regresyon modeli bağımlı değişkenin kesikli; bağımsız değişkenlerin ise kesikli veya sürekli olması durumunda bağımlı değişkenle bağımsız değişkenler arasındaki sebep-sonuç ilişkisinin ortaya konulmasında kullanılmaktadır (Eskandari ve Meshkani, 2006).

Lojistik regresyon analizi, diskriminant analizi ve çoklu regresyon analizinden farklı olarak bağımsız değişkenlerin dağılımına ilişkin araştırmacılarca karşılanması gereken varsayımlar gerektirmez (Tabachnick ve Fidell, 1996). Bir başka deyişle bağımsız değişkenlerin normal dağılımı, doğrusallık ve varyans-kovaryans matrislerinin eşitliği gibi varsayımların karşılanması gerekmez. Dolayısıyla da Lojistik regresyonun diğer iki teknikten çok daha esnek olduğu ifade edilebilir. Lojistik regresyonun yansız ve sapsız istatistikler ortaya koyması için büyük örneklem gerektirdiği bildirilmektedir. Özellikle bağımlı değişkenin ikiden fazla kategorisinin olduğu durumlarda, geçerli bir hipotez testi için, her bağımsız değişkende en az 50 kişilik bir grup büyüklüğüne ihtiyaç vardır. Bazı kaynaklarda bu sayının her bağımsız değişken için minimum 20, toplamda minimum 60 olması gerektiği vurgulanmaktadır. Diğer yandan örneklem büyüklüklerinin aynı olması durumunda, bağımlı değişkenin her bir kategorisinde bağımsız değişkenlerin çok değişkenli normalliğe sahip olması, her bir kategori için varyans ve kovaryansların eşitliği varsayımlarının karşılanması durumunda, daha önce de değinildiği gibi diskriminant analizi, Lojistik regresyon analizine tercih edilmelidir. Bununla birlikte, Lojistik regresyon analizi ile yapılan çözümlerden elde edilen matematiksel modelin yorumlanmasının daha kolay olduğunu belirtmekte yarar vardır (Akkuş ve Çelik, 2004; Grimm ve Yarnold, 1995; Kalaycı, 2005; Leech vd., 2005; Poulsen ve French, 2008; Tabachnick ve Fidell, 1996; Tatlıdil, 1996).

Lojistik regresyon analizi adını, bağımlı değişkene uygulanan logit dönüşümünden (logit transformation) almaktadır. Bu durum aynı zamanda hem kestirim, hem de yorumlama sürecinde bazı farklılıklara neden olur (Hair vd., 2006). Lojistik regresyon analizi, bağımlı değişkenin ölçüldüğü ölçek türüne ve bağımlı değişkenin seçenek sayısına göre üçe ayrılmaktadır. Eğer bağımlı değişken iki seçeneğe sahip bir kategorik değişken ise “İkili Lojistik Regresyon Analizi (Binary Logistic Regression Analysis)” adını alır. Örneğin bir akademik programı bitirme durumuna göre öğrencilerin başarılı ve başarısız olarak nitelendirilmesi durumunda ikili Lojistik regresyon uygulanır. Eğer bağımlı değişken ikiden çok kategorili (düzeyli) sınıflamalı bir değişken ise “Çok Kategorili/Düzeyli İsimli Lojistik Regresyon Analizi (Multinomial Logistic Regression Analysis)” adını alır. Örneğin üç farklı akademik programda öğrenim görmekte olan öğrencilerden oluşan bir bağımlı değişkenin olması durumunda, çok düzeyli isimli Lojistik regresyon uygulanır. Eğer bağımlı değişken sıralama ölçeğiyle elde edilmiş ise, bu durumda da “Sıralı Lojistik Regresyon Analizi (Ordinal Logistic Regression Analysis)” kullanılır. Örneğin öğrencilerin öğrenim gördükleri akademik



programdaki başarılarının “düşük”, “orta” ve “yüksek” olarak gruplandığı durumda sıralı Lojistik regresyon uygulanır. Lojistik regresyon, “tek değişkenli Lojistik regresyon (bağımsız değişkenin tek olduğu durum)” ve “çok değişkenli Lojistik regresyon (bağımsız değişkenin iki veya daha fazla olduğu durum)” olarak sınıflandırma yapılmaktadır (Stephenson, 2008).

Doğrusal regresyon analizinde bağımlı değişkenin değeri kestirilirken, Lojistik regresyon analizinde bağımlı değişkenin alacağı değerlerden birinin gerçekleşme olasılığı kestirilir. Bu olasılık değeri aşağıdaki model kullanılarak elde edilir.

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)}{1 + \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)} \quad (1)$$

Burada  $x = (x_1, x_2, \dots, x_p)$  bağımsız değişkenler vektörü ve  $\beta = (\beta_1, \beta_2, \dots, \beta_p)$  bağımsız değişkenlere ait parametre vektörünü göstermektedir.

İkili değer alan bağımlı değişken  $y(0,1)$  için Eşitlik 1’de verilen ifade, verilen  $x$  değeri için  $y$ ’nin 1’e eşit olma koşullu olasılığını vermektedir. Bu olasılık,

$$\pi(x) = P(y = 1|x) \quad (2)$$

eşitliği ile sağlanır. Benzer biçimde,

$$1 - \pi(x) = P(y = 0|x) \quad (3)$$

eşitliği  $y$ ’nin 0’ı alma koşullu olasılığını göstermektedir.

Lojistik regresyon modelinde parametre tahmini yapılabilmesi için olabilirlik fonksiyonu öncelikle oluşturulmalıdır.  $y(0,1)$  bağımlı değişkeni  $\pi(x)$  başarı olasılığı ile Bernoulli dağılmaktadır. Yukarıdaki eşitlik 1 ve 2’den  $i$ . birim için  $y_i = 1$  olduğunda olabilirlik fonksiyonuna katkısı  $\pi(x_i)$  ve  $y_i = 0$  olduğunda olabilirlik fonksiyonuna katkısı  $1 - \pi(x_i)$  kadardır. Buna göre  $i$ . birimin olabilirlik fonksiyonuna katkısı,

$$L(\beta_i) = \pi(x_i)^{y_i} (1 - \pi(x_i))^{(1-y_i)} \quad i=1,2,\dots,n \quad (4)$$

eşitliği ile ifade edilir. Gözlemlerin birbirinden bağımsız olduğu varsayıldığında olabilirlik fonksiyonu eşitlik 4’teki her bir birimin çarpılmasıyla elde edilir:

$$L(y|\beta) = \prod_{i=1}^n \pi(x_i)^{y_i} (1 - \pi(x_i))^{(1-y_i)} \quad i=1,2,\dots,n \quad (5)$$

Burada  $n$  birim sayısını göstermektedir.

Lojistik regresyon modelinin parametre tahmininde klasik yöntemler olarak; En Çok Olabilirlik, Yeniden Ağırlıklandırılmış Tekrarlı En Küçük Kareler ve tekrarlı veri

durumunda Minimum lojit Ki-Kare yöntemleri kullanılmaktadır (Murat ve Işığışık, 2007). Klasik yöntem alternatif olarak Bayesci yöntemde kullanılmaktadır.

## 2.2 Lojistik Regresyon için Bayesci Yaklaşım

Bayesci yaklaşım, verilerden elde edilen yeni bilgi ile önceden bilinen bilginin derlenmesi ile oluşan bir yöntemdir (Wong vd., 2005). Klasik yaklaşımda tahmin yöntemi sadece veriden derlenen en çok olabilirlik fonksiyonuna dayanırken, Bayesci yaklaşımda klasik yöntemde elde edilen olabilirlik, parametre hakkında bilinen önsel bilgiyi ( $p(\beta)$ ) değiştirmek için kullanılmaktadır. Buna göre Bayesci yaklaşıma göre parametre tahmini;

$$p(\beta / y) \propto L(y|\beta) \times p(\beta) \quad (6)$$

eşitliği ile verilen sonsal dağılım ile elde edilmektedir. Burada sonsal dağılım, parametre hakkındaki önsel bilgi ile veriden elde edilen olabilirlik fonksiyonunun birleşmesinden oluşan güncel bilgidir.

Bayesci yaklaşım, teorik çatı altında verinin önsel bilgi ile birleşiminin temel ve doğal bir yol sağlaması, asimptotik yaklaşım olmaksızın, veriden şartlı ve kesin çıkarımlar vermesi, küçük örnek büyüklüklerinde kullanılabilmesi, parametre tahminleri ve hipotez testlerinde doğrudan çıkarımlar yapması, kayıp veri ve hiyerarşik modeller için kolaylıkla uygulanabilmesi gibi birçok avantajının yanı sıra; önsel bilginin seçiminde kesin bir yöntem olmaması ve özellikle parametre sayısının fazla olduğu durumlarda hesaplama zorluğu dezavantajları arasında sayılabilmektedir (Cengiz vd., 2012).

Bir parametrenin önsel bilgisi, veri elde edilmeden önce parametre hakkındaki bilgileri ifade eder. Bu bilgiler bir dağılım ile ifade edilir. Bu nedenle Bayesci yaklaşıma göre parametreler klasik yaklaşımdaki gibi sabit olarak değil, olasılığa bağlı olarak tanımlanır. Önsel bilgi olmaksızın Bayesci çıkarsama yapılamaz. Önsel bilgiler; bilgi verici (informative) ve bilgi verici olmayan (non-informative) olmak üzere sınıflandırılırlar. Bilgi verici olmayan önseller, sonsal dağılım üzerinde minimum etkiye sahiptir. Bu önseller daha objektiftir olduklarından birçok istatistikçi tarafından kullanılırlar. Ancak, parametre hakkındaki toplam belirsizliğin objektif önselle verilmesi her zaman uygun değildir. Bazı durumlarda, objektif önseller kullanıcıyı yanlış sonsallara yönlendirebilir. Bilgi verici önseller, olabilirlik fonksiyonu tarafından etkisi azaltılmayan önsel dağılımdır. Bu tip önseller deneyimlerden, benzer geçmiş çalışmalardan ve uzman görüşlerinden elde edilen bilgi çerçevesinde belirlenirler. Sonsal dağılım üzerinde oldukça etkili oldukları için önsel dağılım belirlenirken çok dikkatli olunmalıdır. Genel olarak parametre hakkındaki belirsizliği en iyi açıklayacak dağılım;

$$\beta_j \sim N(\mu_j, \sigma_j^2) \quad j=1,2,\dots,p \quad (7)$$

Bilgi olamaması halinde en çok, sıfır ortalamalı ve olabildiğince büyük varyans seçilmelidir. Varyans için 100 ile 10000 arasında bir aralık tercih edilmektedir (Rashwan vd., 2012).

Lojistik regresyon için Bayesci yaklaşım (6) Eşitliğine göre ifade edilirse,  $n$  gözlem için olabilirlik fonksiyonu,

$$L(y|\beta) = \left[ \left( \frac{\exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)}{1 + \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)} \right)^{y_i} \left( 1 - \frac{\exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)}{1 + \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)} \right)^{(1-y_i)} \right] \quad (8)$$

Bilinmeyen  $\beta$  parametrelerine ait önsel bilginin  $\beta_j \sim N(\mu_j, \sigma_j^2)$  dağılımlı olduğu varsayıldığında sonsal dağılım,

$$p(\beta|y) = \prod_{i=1}^n \left[ \left( \frac{\exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)}{1 + \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)} \right)^{y_i} \left( 1 - \frac{\exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)}{1 + \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)} \right)^{(1-y_i)} \right] \quad (9)$$

$$\times \prod_{j=1}^p \frac{1}{\sqrt{2\pi\sigma_j}} \exp \left\{ -\frac{1}{2} \left( \frac{\beta_j - \mu_j}{\sigma_j} \right)^2 \right\}$$

elde edilir (Acquah, 2013). Eşitlik (9)'un analitik çözümü bulunmamaktadır. İstatistiksel çıkarım yapabilmek için her bir parametrenin marjinal dağılımlarının elde edilmesi gerekmektedir. Bunun için çoklu integrallerle işlem yapılması gerekmektedir. Uygulamada sonsal dağılımlardan istatistiksel çıkarımlar yapmak için simülasyon yöntemlerinden Markov Zinciri Monte Carlo (MCMC) yönteminin kullanımını yaygınlaştırmıştır. MCMC yöntemi, ilgilenilen sonsal dağılımdan başarılı bir şekilde bir önekinine bağlı örneklemeler seçmektedir. MCMC yöntemi ücretsiz olarak indirilebilen WinBUGS paket programı yardımıyla kolaylıkla uygulanmaktadır. BUGS (Spiegelhalter vd., 1996), özel olarak MCMC yönteminin tam olasılık modellerine uygulanması için kullanılan ve kod yazımına dayanan bir programdır. Bu program Bayesci analizi sağladığından tüm parametreler rasgele değişken olarak ele alınmaktadır.

### 3. BULGULAR

Bu çalışmada, Marmara Üniversitesi'nde 2013 yılında eğitim görmekte olan öğrencilerin kredi kartı sahipliği belirlemede etkili olan sosyo-ekonomik ve demografik faktörler Bayesci Lojistik regresyon yardımıyla belirlenmeye çalışılmıştır. Bu amaçla, 2011 yılında Gaziosmanpaşa ve İnönü üniversite öğrencilerinin kredi kartı sahipliğini etkileyen faktörleri belirleme çalışmasında uygulanan 24 soruluk anket, Marmara Üniversitesi Haydarpaşa ve Göztepe kampüsü'nde okuyan 200 öğrenciye uygulanmış ve veriler derlenmiştir. Anketin genel güvenilirlik katsayısı Cronbach Alfa katsayısı 0.733 olarak elde edilmiştir. Bu değer, kullanılan anketin oldukça güvenilir olduğunu göstermektedir.

Marmara Üniversitesi Haydarpaşa ve Göztepe kampüsü'nde okuyan 200 öğrencinin kredi kartı sahipliğini belirlemede etkili olabilecek değişkenler; sınıf düzeyi, cinsiyet, yaş, kardeş sayısı, öğretim türü, kredi-burs durumu, anne ve babanın eğitim durumları, annenin çalışma durumu, ailenin aylık geliri, öğrencinin aylık geliri ve harcama tutarı gibi sosyo-ekonomik ve demografik göstergeler ele alınmıştır. Araştırmada yer verilen kredi kart sahipliği (Bağımlı değişken) ve etki eden sosyo-ekonomik ve demografik

göstergelere (bağımsız değişkenlere) ait belirleyici istatistikler Tablo 1'de verilmektedir.

**Tablo 1. Belirleyici istatistikler**

<i>Değişkenler</i>	<i>Sayı</i>	<i>Yüzde</i>
<b>KKS</b> (Kredi Kartı Sahipliliği)		
1=Evet	92	46
2=Hayır	108	54
<b>SNF</b> (Sınıf Düzeyi)		
1=1. Sınıf	35	17.5
2=2. Sınıf	63	31.5
3=3. Sınıf	31	15.5
4=4. Sınıf	71	35.5
<b>CNS</b> (Cinsiyet)		
0=Bayan	117	58.5
1=Erkek	83	41.5
<b>YAS</b> (Yaş)		
1=17-20	66	33
2=21-23	111	55.5
3=24 ve üzeri	23	11.5
<b>KRDS</b> (Kardeş Sayısı)		
1=1-2	134	67
2=3-4	66	33
3=5 ve üzeri kardeş	0	0
<b>OGT</b> (Öğretim Türü)		
1=1. Öğretim	178	89
2=2. Öğretim	22	11
<b>KRDB</b> (Kredi veya Burs)		
1=Hayır	80	40
2=Evet	120	60
<b>BEGT</b> (Baba Eğitim Durumu)		
1=İlköğretim	73	36.5
2=Lise	65	32.5
3=Önlisans	13	6.5
4=Lisans	45	22.5
5=Yüksek Lisans	4	2
<b>AEGT</b> (Anne Eğitim Durumu)		
1=İlköğretim	114	57
2=Lise	58	29
3=Önlisans	4	2
4=Lisans	22	11
5=Yüksek Lisans	2	1
<b>ACLS</b> (Anne Çalışma Durumu)		
1=Hayır	149	74.5
2=Evet	51	25.5
<b>AGLR</b> (Ailenin Aylık Geliri)		
1=500 TL'den az	10	5
2=501-1000 TL	39	19.5
3=1001-1500 TL	43	21.5
4=1501-2000 TL	38	19

5=2001-2500 TL	26	13
6=2501 TL'den fazla	44	22
<b>OGLR (Öğrencinin Aylık Geliri) TL</b>		
1=300 TL'den az	74	37
2=301-600 TL	50	25
3=601 TL'den fazla	76	38
<b>OHRC (Öğrencinin Aylık Harcama Tutarı) TL</b>		
1=300 TL'den az	58	63
2=301-600 TL	23	25
3=601 TL'den fazla	11	12

Bayesci yaklaşımın uygulanmasında, ele alınan sosyo-ekonomik ve demografik göstergelere ait parametrelerin dağılımlarının (önsel bilginin) belirlenmesi amacıyla, 2011 yılında Yayar ve arkadaşları tarafından Gaziosmanpaşa ve İnönü Üniversitelerinde okuyan öğrencilerin kredi kart sahipliğine etki eden faktörlerin saptanması çalışmasından yararlanılmıştır. Bu çalışmada, her iki üniversite verilerini birleştiren (Model-3) klasik Lojistik regresyon analizi sonucunda elde edilen parametre tahmin ve bu tahminlere ait standart hatalar kullanılmış ( $\beta \sim N(\hat{\beta}_{2011}, \sigma_{2011}^2)$ ) ve Tablo 2'de verilmiştir.

**Tablo 2. Açıklayıcı değişkenlerin parametreleri için önsel bilgiler**

Değişkenler	$\hat{\beta}_{2011}$	$\sigma_{\hat{\beta}_{2011}}$
Sabit	-4.359	0.687
SNF	0.022	0.089
CNS	0.051	0.163
YAS	0.635	0.160
KRDS	-0.173	0.132
OGT	0.277	0.169
KRDB	0.291	0.185
BEGT	-0.006	0.093
AEGT	0.050	0.150
ACLS	0.670	0.279
AGLR	0.122	0.074
OGLR	0.366	0.211
OHRC	0.367	0.214

Tablo 2'de verilen önsel bilgiler kullanılarak, Bayesci Lojistik regresyon yöntemi WinBUGS programında macro yazılarak uygulanmıştır. Açıklayıcı değişkenlerin parametrelerine ait sonsal dağılımın sonuçları Tablo 3'te verilmiştir.



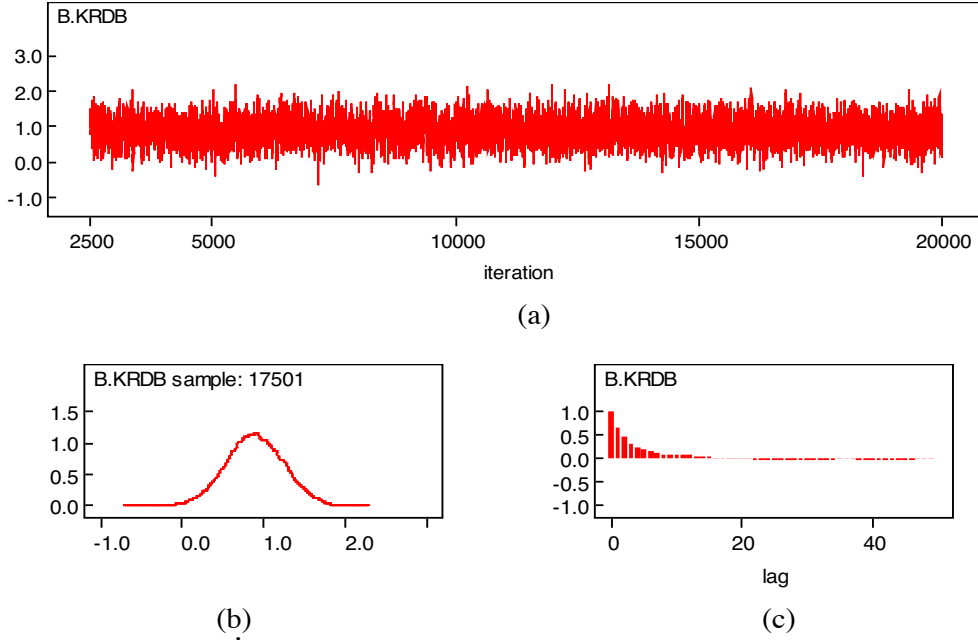
**Tablo 3. Bayesci Lojistik regresyon analiz sonuçları**

Parametreler	$\hat{\beta}$	$\sigma_{\hat{\beta}}$	MC HATA	MC Hata/ $\sigma_{\hat{\beta}}$	Alt sınır %2.5	Üst sınır %97.5	Başlama	Örnek	$Exp(\hat{\beta})$
B.Sabit	-3.591	0.8186	0.0311	0.0380	<b>-5.206</b>	<b>-1.964</b>	2500	17501	0.0276*
B.SNF	0.499	0.1965	0.0063	0.0319	<b>0.1207</b>	<b>0.8976</b>	2500	17501	1.6471*
B.CNS	0.687	0.3487	0.0060	0.0173	<b>0.0053</b>	<b>1.373</b>	2500	17501	1.9877*
B.YAS	-0.344	0.346	0.0117	0.0338	-1.04	0.313	2500	17501	0.7089
B.KRDS	0.0091	0.3434	0.0106	0.03078	-0.6644	0.6762	2500	17501	1.0091
B.OGT	0.6514	0.5493	0.0060	0.0109	-0.4143	1.748	2500	17501	1.9182
B.KRDB	0.8901	0.3501	0.0062	0.0177	<b>0.2048</b>	<b>1.577</b>	2500	17501	2.4354*
B.BEGT	0.1316	0.1679	0.0041	0.0243	-0.1956	0.4618	2500	17501	1.1407
B.AEGT	-0.258	0.2279	0.0055	0.0243	-0.713	0.1885	2500	17501	0.7726
B.ACLS	0.5046	0.4372	0.0076	0.0173	-0.3485	1.368	2500	17501	1.6563
B.AGLR	0.1061	0.1308	0.0036	0.0271	-0.1419	0.3696	2500	17501	1.1119
B.OGLR	0.6807	0.2	0.0047	0.0235	<b>0.2911</b>	<b>1.085</b>	2500	17501	1.9753*
B.OHRC	0.2989	0.3430	0.0086	0.0247	-0.3734	0.9712	2500	17501	1.3484

Sonsal dağılımın yakınsama sağlaması ve başlangıç değerlerinin etkisinin en aza indirilmesi için, Markov zincirinde elde edilen örneklemin başlangıç kısmı çıkarılır. Uygulamada 20000 iterasyonun ilk 2500 iterasyonu çıkarılmış ve elde edilen Bayesci Lojistik regresyon sonuçları Tablo 3'te verilmiştir. Tablo 3 incelendiğinde MC hatanın, sonsal standart hataya oranı %5'ten küçük olduğu için, Thumb kuralına göre yeterli iterasyon sayısına ulaşıldığı söylenir. %2.5 ve %97.5 güven aralıkları incelendiğinde, "0" değerini içeren aralığa sahip olan parametrelerin kredi kart sahipliği üzerine bir etkisi olmadığı söylenir. Buradan hareketle, kredi kart sahipliğine yaş, kardeş sayısı, öğretim türü, baba eğitim durumu, anne eğitim durumu, anne çalışma durumu, ailenin gelir durumunun önemli etkilerinin olmadığı saptanmıştır.

Sınıf düzeyi, cinsiyet, kredi veya burs alma durumu, öğrencinin gelirinin kredi kart sahipliğine %5 önem seviyesinde önemli etkisi olduğu saptanmıştır. Sınıf düzeyindeki bir birimlik artış 1.65 kat veya %65 oranında, erkek öğrencilerin bayan öğrencilere göre yaklaşık 2 kat veya % 99 oranında, kredi ve burs alanların almayanlara göre yaklaşık 2.5 kat veya %104 oranında ve öğrencinin gelirindeki bir birimlik artışın yaklaşık 2 kat veya %98 oranında kredi kart sahipliğini arttırdığı görülmektedir.

Bayesci yaklaşımda kullanılan MCMC yönteminin doğru sonuçlar verildiğine karar verebilmek için sonsal dağılıma yakınsamanın gerçekleştiğinin saptanması gerekir bu amaçla; iz, sonsal yoğunluk ve otokorelasyon grafiklerinden yararlanılır. Kredi kart sahipliğine etki eden değişkenlerden kredi veya burs alma durumuna ait ilgili sırasıyla iz (a), Kernel yoğunluk (b) ve otokorelasyon (c) grafikleri Şekil 1'de verilmiştir.



Şekil 1. İz, Kernel yoğunluk ve otokorelasyon grafikleri

İz grafiğinde salınım fazlalığı ve sıklığı sonsal dağılıma yakınsamanın hızlı bir şekilde gerçekleştiğini, Kernel yoğunluk grafiğinde ise, çan şeklinde oluşan görünümün sonsal dağılıma ulaşıldığını, Otokorelasyon grafiği ise Markov zinciri örnekleri arasındaki bağımlılığı ölçmekte, bu nedenle düşük korelasyon yakınsamanın sağlandığını belirtmektedir.

Klasik ve Bayesci yaklaşımın karşılaştırılması amacıyla Akaike Bilgi Kriteri (AIC) ve Bayesci Bilgi Kriteri (BIC) değerleri bütün veriler hesaplanmış ve Tablo 4'te verilmiştir.

Tablo 4. Klasik ve Bayesci Lojistik regresyon için AIC ve BIC değerleri

Kriter	Klasik	Bayesci
AIC	301.9775	267.44
BIC	344.8556	310.3181

Tablo 4'te yer alan sonuçlar incelendiğinde AIC ve BIC değerleri içinde en küçük değer Bayesci Lojistik regresyon yaklaşımına ait olduğu görülür.

#### 4. SONUÇ

Bu çalışmada, var olan bilginin güncellenmesini sağlayan Bayes yaklaşımının kullanılması amaçlanmıştır. Bu amaçla, 2011 yılında Gaziosmanpaşa ve İnönü üniversite öğrencilerine kredi kartı sahipliğine etki eden faktörlerin belirlenmesi için uygulanan anket formu Marmara Üniversitesi öğrencilerine uygulanmıştır. Kredi kartı sahipliğine etki eden faktörler, Bayesci Lojistik regresyon yaklaşımı ile analiz edilmiştir. Önsel bilgiler, 2011 yılında Gaziosmanpaşa ve İnönü Üniversiteleri için uygulanan klasik Lojistik regresyon analizi sonuçlarından elde edilmiştir. Bayesci Lojistik regresyon yaklaşımının yeterli yakınsama sağlaması ile ilgili Thumb kuralı, iz

grafığı, Kernel yoğunluk grafığı ve otokorelasyon grafikleri verilmiştir. Klasik yöntem ile karşılaştırılmasında AIC ve BIC kriterleri kullanılmıştır.

2011 yılında Yayar ve arkadaşları tarafından Gaziosmanpaşa ve İnönü Üniversitelerinde okuyan öğrencilerin kredi kart sahipliğine etki eden faktörlerin saptanması çalışmasında her iki üniversite verilerini birleştiren (Model-3) klasik Lojistik regresyon analizi sonucunda elde edilen parametre tahmin ve bu tahminlere ait standart hatalar kullanılarak belirlenen önsel bilgiler yardımıyla uygulanan Bayesci Lojistik regresyon sonucunda yeterli yakınsamanın Thumb kuralı, iz grafığı, Kernel yoğunluk grafığı ve otokorelasyon grafikleri ile sağlandığı görülmüştür.

Kredi kartı sahipliğine etki eden faktörlerin belirlenmesinde sınıf düzeyi, cinsiyet, yaş, kardeş sayısı, öğretim türü, kredi-burs durumu, anne ve babanın eğitim durumları, annenin çalışma durumu, ailenin aylık geliri, öğrencinin aylık geliri ve harcama tutarı gibi sosyo-ekonomik ve demografik göstergeler ele alınmıştır.

%5 önem düzeyine göre belirlenen güven aralığında sıfır'ı içermeyen (anlamlı bulunan) faktörlerin; sınıf düzeyi, cinsiyet, kredi veya burs alma durumu, öğrencinin gelirinin kredi kart sahipliğine önemli etkisi olduğu saptanmıştır. Sınıf düzeyindeki artış 1.65 kat, erkek öğrencilerin bayan öğrencilere göre yaklaşık 2 kat, kredi ve burs alanların almayanlara göre yaklaşık 2.5 kat ve öğrencinin gelirindeki bir birimlik artışın yaklaşık 2 kat veya kredi kart sahipliğini arttırdığı görülmektedir. Öğrencinin kredi kartı sahip olmalarında kendi gelirinin etkili ancak aile gelir durumunun etkili olmamasının nedeni, öğrencilerin yarı zamanlı çalışmasına bağlanabilir.

Klasik ve Bayesci yaklaşımın karşılaştırılması amacıyla AIC ve BIC değerleri hesaplanmış ve en küçük AIC ve BIC değerlerinin Bayesci yaklaşıma ait olduğu saptanmıştır. Dolayısıyla bu çalışma için Bayesci yaklaşımın tercih edilebileceği göstermiştir.

Önsel bilgiye başvurma esnekliği nedeniyle Bayesci yaklaşım, klasik yöntemlere göre oldukça avantajlı olduğu yorumu yapılabilmektedir.

## 5. KAYNAKLAR

Acquah, H. D., 2013. Bayesian Logistic Regression Modelling Via Markov Chain Monte Carlo Algorithm, Journal of Social and Development Sciences, 4(4), 193-197.

Akkuş, Z., Çelik, M. Y., 2004. Lojistik Regresyon ve Diskriminant Analizi Yöntemlerinde Önemli Ölçütler. VII. Ulusal Biyoistatistik Kongresinde sunulan bildiri, Mersin Üniversitesi, Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Mersin.

Anderson, H. A., 1983. Robust Inference Using Logistic Models, Bulletin of International Statistical Institute, 48, 25-53.

Anderson, J. A., 1979. Multivariate Logistic Compounds. Biometrika, 66, 17-191.

Başarı, G., 1990. Çok Değişkenli Verilerde Ayrımsama Sorunu ve Lojistik Regresyon Analizi, (Uygulamalı İstatistik Doktora Tezi). Hacettepe Üniversitesi., 1-36, Ankara.

Cengiz, M. A., Terzi, E., Şenel, T., Murat, N., 2012. Lojistik Regresyonda Parametre Tahmininde Bayesci bir Yaklaşım, Afyon Kocatepe Üniversitesi Fen Bilimleri Dergisi, 12, 15-22.

Congdon, P., 2006. Bayesian Statistical Modelling, John Wiley & Sons, England.

Cornfield, J., 1962. Joint Dependence of the Risk of Coronary Heart Disease on Serum Cholesterol and Systolic Blood Pressure: A Diskrimant Function Analysis, Federation Proceedings, 21, 58-61.

Cox, D. R., 1970. Analysis of Binary Data. London: Chapman and Hall (2nd ed. 1989 with E.J. Snell).

Çavuş, M. F., 2006. Bireysel Finansmanın Temininde Kredi Kartları: Türkiye’de Kredi Kartı Kullanımı Üzerine bir Araştırma, Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 15, 173-187.

Duffy, D. E., 1990. On Continuity-corrected Residuals in Logistic Regression, Biometrika, 77, 287-293.

Eskandari, F., Meshkani M. R., 2006. Bayesian Logistic Regression Model Choice Via Laplace-Metropolis Algorithm, Journal of the Iranian Statistical Society (JIRSS), 5(1-2), 9-24.

Grimm, L. G., Yarnold, P. R., 1995. Reading and Understanding Multivariate Statistics, Washington D.C.: American Psychological Association.

Hair, J. F., Black, W. C., Babin, B., Anderson, R. E., Tatham, R. L., 2006. Multivariate Data Analysis (6th ed), Upper Saddle River, NJ: Prentice-Hall.

Hsu, J. S., Leonard, T., 1995. Hierarchical Bayesian Semiparametric Procedures for Logistic Regression, Biometrika, 84, 85-93.

Ibrahim, J. G., Chen, M. H., Sinha, D., 2001. Bayesian Survival Analysis, Springer-Verlag, New York.

Kalaycı, Ş., 2005. SPSS Uygulamalı Çok Değişkenli İstatistik Teknikleri, Ankara: Asil Yayın Dağıtım.

Keskin, D., Koparan, E., 2010. Üniversite Öğrencilerinin Kredi Kartı Sahipliliğini Belirleyen Faktörler, Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 5(1), 111-129.

Leech, N.L., Barrett, K.C., Morgan, G.A., 2005. SPSS for Intermediate Statistics: Use and Interpretation (2nd ed), Mahwah, NJ: Lawrence Erlbaum Associates.

Lesaffre, E., 1986. Logistic Discriminant Analysis with Applications in Electrocardiography, Phd Thesis, Katholieke Universiteit Leuven, Belgium, 354. Unpublished.



Lesaffre, E., Albert, A., 1989. A Multiple Group Logistic Regression Diagnostics, *Applied Statistics*, 38 (3), 425-440.

Murat, D., Işığçok, E., 2007. 2007 Seçim Döneminde Ekonomik ve Siyasi Duruma İlişkin Beklentiler: Bursa uygulaması, 8. Türkiye Ekonometri ve İstatistik Kongresi 24-25 Mayıs 2007. İnönü Üniversitesi, Malatya.

Oktay, E., Özen, Ü., Alkan, Ö., 2009. Kredi Kart Sahipliğinde Etkili Olan Faktörlerin Araştırılması: Erzurum örneği, *Dokuz Eylül Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 24(2), 1-22.

Poulsen, J., French, A., 2008. Discriminant Function Analysis. <http://userwww.sfsu.edu/~efc/classes/biol710/discrim/discrim.pdf> adresinden 22 Kasım 2008 tarihinde edinilmiştir.

Pregibon, D., 1981. Logistic Regression Diagnostics, *The Annals of Statistics*, 9, 705-724.

Rashwan, N. I., El Dereny, M., 2012. The Comparison Between Results of Application Bayesian and Maximum Likelihood Approaches on Logistic Regression Model for Prostate Cancer Data, *Applied Mathematical Sciences*, 6 (23), 1143-1158.

Robert, G., Rao, N. K., Kumar, S., 1987. Logistic Regression Analysis of Sample Data, *Biometrika*, 35, 58-79.

Spiegelhalter, D., Thomas, A., N. Best, N. Gilks, W., 1996. BUGS 0.5: Examples Volume 1, MRC Biostatistics Unit, Institute of Public Health, Cambridge.

Stephenson, B., 2008. Binary Response and Logistic Regression Analysis, [www.public.iastate.edu/~stat415/stephenson/stat415\\_chapter3.pdf](http://www.public.iastate.edu/~stat415/stephenson/stat415_chapter3.pdf) adresinden 22 Kasım 2008 tarihinde edinilmiştir. Şekercioğlu, G. (2008). Fatalizm.

Tabachnick, B. G., Fidell, L. S., 1996. *Using Multivariate Statistics* (3rd ed.), New York, USA: Harper Collins College Publishers.

Tathdil, H., 1996. *Uygulamalı Çok Değişkenli İstatistiksel Analiz*, Ankara: Engin Yayınları.

Turgay, O., Başgöl, N., 2007. Önemli Bir Finansman Kaynağı Olarak Kredi Kartları: Kredi Kartlarının Kart Sahiplerinin Harcamaları Üzerindeki Etkilerini Belirlemeye Yönelik Burdur İlinde Bir Araştırma, *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 12, 215-226.

Wong, M. C. M., Lam, K. F., Lo, E. C. M., 2005. Bayesian Analysis of Clustered Interval-Censored Data, *J Dent Res*, 84(9), 817-821.

Yayar, R., Karaca, S. S., Turkut, A., 2011. Üniversite Öğrencilerinin Kredi Kart Sahibi Olmaları Üzerinde Etkili Olan Faktörler: Gaziosmanpaşa ve İnönü Üniversitelerinden Ampirik Bulgular, *Akademik Yaklaşımlar Dergisi*, 2(1), 152-169.



## INVESTIGATION OF FACTORS EFFECTIVE ON CREDIT CARD OWNERSHIP OF MARMARA UNIVERSITY STUDENTS BY BAYESIAN LOGISTIC REGRESSION

### ABSTRACT

*Bayesian approach is a method which combines new information obtained from data and previous knowledge. In this study, Bayesian approach, which is an alternative to classical approach, is applied to logistic regression, which explores the effects of covariate(s) on binary dependent variable. To this aim, socioeconomic and demographic factors that are effective on credit card ownership of Marmara University students are examined.*

**Keywords:** Credit card ownership, Bayesian logistic regression, WinBUGS