

İntiharı Etkileyen Sosyal ve Ekonomik Faktörlerin Beta Regresyon Analizi ile Belirlenmesi

Elif BULUT
Ondokuz Mayıs Üniversitesi
elif@omu.edu.tr
ORCID ID: 0000-0001-8278-1821

Veysel Gökhan AYDIN
Ondokuz Mayıs Üniversitesi
veyselgokhanaydin@gmail.com
ORCID ID: 0000-0002-8121-745X

Araştırma Makalesi	DOI: 10.31592/aeusbed.657624
Geliş Tarihi: 23.12.2019	Revize Tarihi: 19.06.2020
	Kabul Tarihi: 13.07.2020

Atf Bilgisi

Bulut, E. ve Aydın, V. G. (2020). İntiharı etkileyen sosyal ve ekonomik faktörlerin beta regresyon analizi ile belirlenmesi. *Ahi Evran Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 6(2), 422-436.

ÖZ

İntihar, süre gelen bir kavram olup, psikolojik, sosyolojik ve ekonomik etkenlerdeki değişimlerden kolayca etkilenebilmektedir. İntihar oranları ülkeden ülkeye değiştiği gibi, aynı toplum içerisinde de farklı kesimlerde farklı şekillerde ve oranlarda görülebilmektedir. Bu çalışma da intihara etki ettiği düşünülen bazı ekonomik ve sosyal yapıdaki değişkenler ele alınmış, değişkenlerin anlamlılığı beta regresyon analizi ve çoklu doğrusal regresyon analizi ile saptanmıştır. Çalışmada, 2015 yılı ve 2016 yılına ait intihar oranları, nüfus, işsizlik oranı ve alkol tüketimi gibi değişkenler analizde kullanılmıştır. Çalışmanın yapıldığı tarih itibarıyla eksiksiz olarak ve en güncel hali ile yalnızca 2016 yılına kadar tespit edilmiş olduğundan 2015 ve 2016 yılı verileri üzerinde çalışılmıştır. Bağımlı değişkenin intihar oranları olarak (0,1) arasında oransal olduğu durumlarda beta regresyon analizinin çoklu doğrusal regresyon analizi ile karşılaştırıldığında daha iyi tahminler yaptığı gösterilmiştir. Ayrıca analiz türlerine göre değişkenlerin anlamlılığı test edilmiş ve kişi başı alkol tüketimi, işsizlik oranı ve işgücü katılım oranının intihar üzerinde etkisi olduğu saptanmıştır.

Anahtar Kelimeler: İntihar, beta regresyon, çoklu doğrusal regresyon, hata kareler ortalamasının karekökü.

Determination of Social and Economic Factors Affecting Suicide Via Beta Regression Analysis

ABSTRACT

Suicide is an ongoing concept that is easily affected by psychological, sociological and economic factors. Suicide rates changing from country to country as in the society can be seen in different ways and proportions. In this study, some variables that are thought to affect suicide in economical and social structure are discussed. Significance of variables are obtained with beta regression analysis and multiple linear regression analysis. In this study, suicide rates, population, unemployment rate and alcohol consumption for 2015 and 2016 are used as variables. Data used in this study has been obtained completely and currently since 2016. Since suicide rates are used as dependent variable and the dependent variable is continuous on the interval (0,1), beta regression analysis makes better estimates than multiple linear regression analysis. The study has also found that per capita alcohol consumption, unemployment rate and labor force participation rate had a significant effect on suicide.

Keywords: Suicide, beta regression, multiple linear regression, root mean square error.

Giriş

Günümüz şartlarında intihar önemli toplumsal problemler arasında yer almaktadır. Bu sorun sadece günümüzde yaşanan bir sorun olmaktan öte süregelen bir olgu olarak karşımıza çıkmaktadır. Bazı çalışmalarda intihar tamamıyla toplumsal bir olgu olarak tanımlanmıştır. Toplum yapılarındaki pozitif ya da negatif değişiklikler intihar oranlarının artmasına ya da azalmasına sebep olmaktadır. Toplumsal bir olgu olarak intihara etki eden değişkenler bireysel olabileceği gibi ekonomik ve ülke bazlı değişkenlerde olabilmektedir.

İntihar oranları ülkeden ülkeye değiştiği gibi aynı toplum içerisinde de farklı kesimlerde farklı şekillerde ve oranlarda görülebilmektedir. Bu çalışmada, toplumların ekonomik ve sosyal

yapılarındaki bazı değişkenler ele alınarak bu değişkenlerin intihar oranları üzerinde anlamlı etkileri olup olmadıkları incelenmiştir. Çalışma kapsamında 30 farklı ülkenin intihar oranları ve bazı ekonomik-sosyal göstergeleri kullanılarak bu etki tespit edilmeye çalışılmıştır. Türkiye için yapılan benzer çalışmalar olmakta olup çoğunlukla işsizlik etkeni üzerinde çokça durulmuştur. Bunun sonucu olarak işsizlik oranları ile intihar oranları arasında anlamlı bir ilişki olduğu gösterilmiştir. Sözü geçen çalışmalarda genellikle VAR analizi, ki-kare testi, t-testi gibi yaygın kullanılan yöntemler tercih edilmiştir. Yapılan bu çalışmada regresyon analiz yöntemlerinden beta regresyonu kullanılmış, değişkenler arasındaki ilişki modellenmeye çalışılmış ve etkisi önemli olan değişkenler elde edilmiştir. Çalışmada, 2015 yılına ait veriler analiz edilmiş olup oluşturulan modele göre 2016 yılı intihar oranları tahmin edilmiş ve 2016 yılına ait gerçek intihar oranları ile karşılaştırılmıştır. Çalışmaya ek olarak çoklu doğrusal regresyon analizi eklenmiş olup, veri yapısına göre beta regresyon analizinin daha uygun bir yöntem olduğu gösterilmiştir.

Literatürde Beta regresyon analizinin diğer analizlerde karşılaştırılmasının yapıldığı yerli ve yabancı çalışmalar yer almaktadır. Ferrari ve Cribari-Neto (2004), çalışmalarında oranları modellemede beta regresyon analizinden yararlanmışlardır. Konuyu teorik olarak incelemiş ve 1956 yılına ait bir veri seti ile çalışmalarını uygulama ile desteklemişlerdir. Korkmaz, Germir, Dündar, Şen, Karta ve Şahbudak (2015), çalışmalarında dünya ülkelerini ekonomik benzerliklerine göre sınıflandırarak enflasyonu etkileyen faktörleri saptamak için kümeleme analizi ve beta regresyon analizini kullanmışlardır. Net vergiler ve nüfusun enflasyon oranı üzerinde anlamlı etkisi olduğunu göstermişlerdir. Mandal, Srivastav ve Simonovic (2016), Campbell River Havzası, British Columbia, Kanada'da yağışların istatistiksel olarak küçültülmesi için beta regresyonunun kullanımına dair bir çalışma yapmışlardır. Önerdikleri ölçek küçültme modelinin yerel yağış değişkenliğinde zamansal ve mekânsal değişkenliğinin etkisini farklı lokasyonlarda iyi tanımladığını göstermişlerdir. Türkan ve Özel (2017), çalışmalarında Türkiye'de eğitim vermekte olan 43 üniversite üzerinden çalışmalarını yürütmüş olup etkinliğe etki eden faktörlerin belirlenmesinde beta regresyon modelini kullanmışlardır. Çalışma sonunda h-indexi, lisansüstü mezun sayısı ve tıp fakültesi varlığının üniversitelerin etkinliği üzerinde etkisi olduğunu tespit etmişlerdir. Ünlü ve Aktaş (2017) çalışmalarında gelir, sağlık, eğitim gibi göstergelerle Türkiye 2015 refah endeksi verilerini modellemede Beta regresyon analizinden yararlanmışlardır. Çalışmalarında 40 adet bağımsız değişkenden 22 tane bağımsız değişkenin mutluluk üzerinde anlamlı etkisinin olduğunu göstermişlerdir. Chen, Chiu ve Chang (2017), konteyner nakliye endüstrisinde hizmet özellikleri ve müşteriye elde tutma olasılığı arasındaki ilişkiyi saptamada beta regresyon analizini kullanmışlardır. Fiyat ve iskonto özellikleri, kişisel satış ve müşteri ilişkilerinin müşteriye elde tutma olasılığı üzerinde anlamlı etkisi olduğunu göstermişlerdir. Atasoy (2018), tez çalışmasında göç hızına etki edebilecek 26 değişken belirlemiş olup bu değişkenleri 8 il için elde ettiği veriler üzerinde beta regresyon modeli ile incelemiş ve göç hızına etki eden değişkenleri tespit etmiştir. Koç (2019), çalışmasında 2017 yılına ait veriler üzerinde beta regresyon modelini kullanarak bazı parametrelerin olumlu, bazı parametrelerin olumsuz etkisini tespit etmiştir. Zaman, Dündar ve Aydın (2019), çalışmalarında gini katsayısı üzerinde etkiye sahip olan makro ekonomik değişkenleri beta regresyon modeli kullanarak tespit etmeye çalışmışlardır. Çalışma sonucunda gini katsayısı üzerinde anlamlı etkiye sahip olan değişkenler tespit edilip yorumlanmıştır. Gayawan, Fasusi ve Bandyopadhyay (2020), Nijerya'da yaptıkları bir çalışmada çocuk ölüm oranına dahil olmayan bir kadının olasılığını hesaplamada beta regresyon analizinden yararlanmışlardır.

İntihar konusu literatürde gerek sosyal tabanlı gerekse sayısal boyutlu olarak çalışılmıştır. Ekici, Savaş ve Çıtak (2001), çalışmalarında mevcut intihar olgularında sosyal güvencenin olup olmaması, göç etmeni ve diğer psikososyal etkilerin intihar olgusuna bir etkisi olup olmadığını araştırmışlardır. Araştırma sonucunda bu etmenlerin intihar riskini arttıran etmenler arasında yer aldıklarını tespit etmişlerdir. Topbaş (2007), çalışmasında 1974-2005 yılları arasında Türkiye'deki işsizlik ve intihar verilerini incelemiş olup analizleri sonucunda işsizliğin intihar üzerinde anlamlı bir etkiye sahip olduğunu tespit etmiştir. Tunalı ve Özkaya (2016), yaptıkları çalışmada intihar ve işsizlik arasındaki ilişkiyi 1980-2014 dönemi verileri kullanarak analiz etmişlerdir. Çalışma sonucunda bu iki olgu arasında çift yönlü bir nedensellik ilişkisi saptamışlardır. Amini, Ahmadinia, Poorolajal ve Moqaddasi Amiri (2016), intihar için yüksek riskli grupları saptamada destek vektör makineleri, karar ağacı ve yapay sinir ağları analizlerinden yararlanmışlardır. Destek vektör makinelerinin en iyi

sınıflama modeline sahip olduğunu göstermişlerdir. Fazel, Wolf, Larsson, Mallett ve Fanshawe (2019), ağır akıl hastalığında intihar tahminini çalışmışlardır. Çalışmalarında çok değişkenli lojistik regresyon analizini kullanmış ve en yüksek risk faktörünün önceden kendine zarar verme ve erkek olma gibi faktörler olduğunu göstermişlerdir. Miché, Studerus, Meyer, Gloster, Beesdo-Baum, Wittchen ve Lieb (2020), toplumdaki ergenlerde ve genç yetişkenlerde intihar girişimlerinin tahmini için regresyon yöntemlerini ve makine öğrenme metodunu kullanmışlardır. Çalışmada kullandıkları iki algoritmanında intihar girişimini iyi şekilde açıkladığını göstermişlerdir.

Çalışmada beta regresyon analizinin istatistiksel bir yöntem olarak veri yapısına daha uygun bir analiz olduğu gösterilmeye çalışılmış ve intihar verileri ile çalışma desteklenmiştir. Çalışma yöntem, bulgular ve sonuç bölümlerinden oluşmaktadır.

Dünya genelindeki intihar oranlarının sebeplerine bakıldığında bunların bir kısmının finansal, bir kısmının sosyal ve bir kısmının ise diğer sebeplerden kaynaklandığı ifade etmek mümkündür. Bu bağlamda çalışma kapsamında araştırılan, sosyal ve finansal değişkenlerin intihar oranlarının üzerinde etkiye sahip olup olmadığını ortaya çıkarmaktır.

Yöntem

Çalışmaya konu olan etki, değişkenler arasında matematiksel model kurabilmek ve katkının büyüklüğünü araştırmak adına regresyon analizi yöntemlerinden yararlanılarak araştırılmıştır. Çalışma kapsamında sosyal ve ekonomik değişkenlerin intihar oranları üzerindeki etkisinin tespit edilmesinin yanı sıra literatürde yaygın olarak kullanılan çoklu doğrusal regresyon analizi ile veri yapısına uygun olan ve sosyal bilimlerde de kullanılan beta regresyon analizleri tahminleme açısından karşılaştırılmaya çalışılmıştır. Hangi analiz yönteminin tahmin gücünün yüksek olduğunun gösterilmesi de çalışmanın diğer bir araştırma problemi olmaktadır.

Evren ve Örneklem

Çalışma evrenini Dünya Bankası Açık Veri (World Bank Open Data) veri tabanında yer alan ülkeler oluşturmaktadır. Ülke sayısının çok olmasından ötürü ve veri tabanlarında çalışmada yer alan değişkenlerin tüm ülkeler için temin edilememesinden dolayı az sayıda ülke ile çalışma yürütülmek durumunda kalmıştır. Değişkenlerin tam olarak temin edilebildiği ülkelerden rastgele seçim yapılmıştır. Araştırmanın örneklemini olarak 30 farklı ülke çalışma kapsamına alınmıştır. Çalışmada yer alan ülkeler tanınırlık ve gelişmişlik düzeyine göre (gelişmiş, gelişmekte olan ve az gelişmiş ülkeler) kaynaklardan rastgele ve objektif bir yaklaşımla seçilmiştir. Araştırma kapsamında değerlendirmeye alınan 2015 ve 2016 yılı alkol tüketim oranları Dünya Sağlık Örgütü'nün [World Health Organization (WHO)] kendi web sitesinde yayınlamış olduğu "Alkol ve Sağlık Hakkında Küresel Durum Raporu 2018" (Global Status Report on Alcohol and Health 2018) başlıklı rapordan elde edilmiştir (Poznyak ve Rekve, 2018). İntihar oranları ve diğer finansal, sosyal değişkenlere ait veriler ise Dünya Bankası'na ait veri bankasındaki Dünya Kalkınma Göstergeleri (World Development Indicators) açık veri tabanı kullanılarak elde edilmiştir.

Verilerin Elde Edilmesi

Çalışmanın bu bölümünde kullanılan verilerin tamamı Dünya Bankası (The World Bank Databank) ve Dünya Sağlık Örgütü [World Health Organization (WHO)] açık erişim veri tabanından temin edilmiştir. Bu verilerde 30 farklı ülke çalışma kapsamına alınmıştır. Araştırmada 1 tane bağımlı ve 12 tane bağımsız değişken yer almaktadır. Değişkenlerin ve ülkelerin az sayıda olmasının başlıca nedeni değerlendirmeye alınan bağımlı ve bağımsız değişkenlerin ilgili kaynaklarda en güncel hali ile yalnızca 2016 yılına kadar tespit edilmiş olmasıdır. Bu nedenle 2015 ve 2016 yılı verileri üzerinde çalışılmıştır.

Verilerin Analizi

Beta Regresyon Analizi

Uygulayıcılar sıklıkla regresyon modellerini diğer değişkenlerle ilişkili olduğu algılanan verileri analiz etmede kullanmaktadır. Özellikle doğrusal regresyon modeli sıkça kullanılmaktadır. Ancak, bağımlı değişkenin (0,1) aralığında kısıtlı olduğu durumlarda doğrusal regresyon modeli uygun olmamaktadır. Çünkü ilgilenilen değişken için, değişkenin alt ve üst sınır değerlerini geçen tahminler verebilir. Mümkün çözüm, bağımlı değişkeni dönüştürmek olup gerçek çizgide değerleri varsayılabilir kılmak ve sonrasında dönüştürülmüş değişkenin ortalamasının dış değişken setine dayalı olarak bir doğrusal tahminleyici şeklinde modellenebilmesidir. Bu yaklaşım bazı sakıncalara sahip olabilmektedir. Bunlardan biri model parametrelerinin orijinal değişkenlere göre kolaylıkla yorumlanamaması olmaktadır. Diğer bir eksiklik ise oranların ölçümü asimetrik özellik gösterebilir ve bundan dolayı normal varsayımına dayanan çıkarımlara yanıltıcı olabilir. Bu durumda istenen bağımlı değişkenin standart birim aralığında $0 < y < 1$ ölçümlendiği durumlar için kullanılan regresyon modeli ile çalışmaktadır. Önerilen model bağımlı değişkenin beta dağıldığı varsayımına dayanmaktadır. Beta dağılımı, dağılımı indeksleyen iki parametrenin değerlerinde yoğunluğu oldukça farklı şekiller aldığından oranları modellemek için çok esneklerdir (Ferrari ve Cribari-Neto, 2004).

Beta yoğunluğu;

$$f(y; p, q) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} y^{p-1}(1-y)^{q-1} \quad 0 < y < 1$$

şeklinde tanımlanmaktadır. Burada $p > 0$, $q > 0$ ve dağılımın şeklini kontrol eden bilinmeyen parametrelerdir. y bağımlı değişken ve $\Gamma(\cdot)$ gamma fonksiyonudur (Ferrari ve Cribari-Neto, 2004).

y 'nin ortalaması ve varyansı sırasıyla;

$$E(y) = \frac{p}{p+q} \quad ve$$
$$var(y) = \frac{pq}{(p+q)^2(p+q+1)} \quad 'dir.$$

Dağılımın modu hem p hem de q 1'den büyük olduğunda mevcuttur (Ferrari ve Cribari-Neto, 2004).

$$mod(y) = \frac{p-1}{p+q-2} \quad 'dir.$$

Beta dağılımı, belirsizlik veya olasılığın tesadüfi değişiminde, kesir gibi durumlarda uygulama alanına sahiptir. Bu yüzden bu dağılım tesadüfi değişkenin (0,1) aralığında sürekli olduğu finansal bilimler veya eğitim bilimleri gibi sosyal bilimlerde yaygın uygulama alanına sahiptir (Cepeda-Cuervo, 2015).

Beta dağılım fonksiyonu Beta (p,q) ortalama ve varyansın bir fonksiyonu olarak tekrar parametrelendirilebilir. Beta dağılımının bu karakterizasyonu beta regresyon modelinin tanımında uygun olabilmektedir (Yellareddygarı, Pasche, Taylor, Hua ve Gudmestad, 2016).

İlk tekrar parametrelendirmede;

$$p = \mu\phi \quad ve \quad q = \phi(1-\mu)$$
$$\sigma^2 = \frac{\mu(1-\mu)}{\phi+1}$$

Bu durumda θ , μ 'nün sabit değerleri için hassas parametreleri olarak, θ 'nın büyük değerleri ise y 'nin varyansının küçülen değerlerine karşılık gelecek şekilde yorumlanır. Böylece;

$$f(y; \mu, \theta) = \frac{\Gamma(\theta)}{\Gamma(\mu\theta)\Gamma((1-\mu)\theta)} y^{\mu\theta-1} (1-y)^{(1-\mu)\theta-1} \quad 0 < y < 1$$

$$0 < \mu < 1 \text{ ve } \theta > 0 \text{ olur.}$$

Yüzdelik bağımlı değişkenin beta dağıldığı varsayımında, bir beta regresyon modeli dizayn edilebilir.

$y_1, y_2, \dots, y_n \sim B(\mu, \theta)$ [$y \sim B(\mu, \theta)$] beta yoğunluğunda tesadüfi örneklem olsun. O zaman beta regresyonu;

$$g(\mu_i) = \beta_0 + x_{i1}\beta_1 + \dots + x_{ik}\beta_k = \eta_i \quad i = 1, 2, \dots, n$$

x_{i1}, \dots, x_{ik} 'ler açıklayıcı değişkenler, $\beta_0, \beta_1, \dots, \beta_k$ ise her bir açıklayıcı değişkene karşılık gelen tahmin edilen kesişim ve katsayılarıdır. n örneklem genişliği ve η_i 'de gözlemin doğrusal kestiricisidir.

Burada $g(\cdot)$ doğrusal kestirici ve bağımlı değişkeni bağlayan bağıntı fonksiyonudur (Yellareddygarı vd., 2016).

Çoklu Doğrusal Regresyon

Veri setinde iki veya daha fazla bağımsız değişkenin yer aldığı durumlarda, bağımsız değişkenler ile bağımlı değişken arasındaki matematiksel model çoklu doğrusal regresyon analizi ile kurulmaktadır.

Çoklu doğrusal regresyon analizinde matrisel notasyon kullanılmakta ve değişkenler arasında $Y = X\beta + \epsilon$ şeklinde eşitlik kurulmaktadır. Daha fazla ayrıntı için Sınıksaran (2011), "Teori ve İstatistiksel Yöntemler" kitabına bakılabilir. Özçiftçi (2020) çalışmasında, sosyal medya özelliklerinin algılanan değere etkisini tespit etmek amacıyla çoklu regresyon analizinden faydalanmıştır. Yaygın bir kullanım alanına sahip olan ve her bilim dalında sıkça kullanılan ve bilinen analizlerden biri olan çoklu doğrusal regresyon analizi çalışmada kullanıldığı için başlık açılmış olup, konu anlatımında ayrıntıya girilmemiştir.

Bulgular

Çalışmada yer alan ekonomik ve sosyal bağımsız değişkenler ile intihar oranları bir arada değerlendirilerek 2015 yılı verileri üzerinden regresyon modeli kurularak 2016 yılı intihar oranları tahmin edilmesi amaçlanmaktadır. Analizler için RStudio programı kullanılmıştır.

Aşağıdaki tabloda çalışmada yer alan bağımlı ve bağımsız değişkenler ile kısaltmaları verilmektedir.

Tablo 1
Çalışmada Yer Alan Değişkenlere Ait Kısaltmalar

Değişken	Kısaltma
İntihar Oranı	Y
Kişi Başı Millî Gelir	X ₁
Gayri Safi Millî Hasıla	X ₂

İhracat Miktarı	X_3
Gayri Safi Yurtiçi Hasıla	X_4
Enflasyon Oranı	X_5
Nüfus Miktarı	X_6
Kişi Başı Alkol Tüketimi	X_7
Vergi Oranı	X_8
İşgücü Katılım Oranı	X_9
İşsizlik Oranı	X_{10}
Kişi Başı Sigara Tüketimi	X_{11}
Beslenme Olanaklarına Ulaşım Zorluğu	X_{12}

Belirlenen amaç doğrultusunda belirtilen kaynaklardan 2015 ve 2016 yılı verileri toplanmıştır. Verilerin birimleri arasındaki farklılıkları ortadan kaldırmak ve daha net bir değerlendirme yapabilmek adına verilerde standartlaştırma işlemi yapılmıştır.

Tablo 2’de 2015 verileri için veri setinde yer alan ekonomik ve sosyal göstergelerin korelasyon matrisi ve anlamlılık dereceleri görülmektedir.

Tablo 2
Bağımsız Değişkenlere Ait Korelasyon Matrisi

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
X_1	1	0,992	0,177	0,998	-0,478	-0,312	0,373	-0,158	0,204	-0,071	-0,154	-0,493
p		0,000	0,349	0,000	0,008	0,093	0,042	0,404	0,279	0,710	0,416	0,006
X_2	0,992	1	0,127	0,994	-0,456	-0,319	0,364	-0,127	0,188	-0,062	-0,158	-0,490
p	0,000		0,502	0,000	0,011	0,085	0,048	0,505	0,321	0,745	0,406	0,006
X_3	0,177	0,127	1	0,165	-0,205	0,528	0,162	0,418	0,122	-0,157	0,142	-0,033
p	0,349	0,502		0,383	0,276	0,003	0,394	0,022	0,521	0,407	0,455	0,864
X_4	0,998	0,994	0,165	1	-0,482	-0,314	0,385	-0,137	0,191	-0,063	-0,164	-0,497
p	0,000	0,000	0,383		0,007	0,092	0,035	0,470	0,313	0,740	0,385	0,005
X_5	-0,478	-0,456	-0,205	-0,482	1	0,128	-0,313	-0,089	-0,044	-0,044	0,281	0,142
p	0,008	0,011	0,276	0,007		0,499	0,092	0,640	0,816	0,818	0,133	0,454
X_6	-0,312	-0,319	0,528	-0,314	0,128	1	-0,154	0,45	0,021	-0,222	-0,243	0,535
p	0,093	0,085	0,003	0,092	0,499		0,416	0,012	0,911	0,239	0,197	0,002
X_7	0,373	0,364	0,162	0,385	-0,313	-0,154	1	0,154	0,015	0,134	0,279	-0,483
p	0,042	0,048	0,394	0,035	0,092	0,416		0,418	0,938	0,481	0,136	0,007
X_8	-0,158	-0,127	0,418	-0,137	-0,089	0,450	0,154	1	-0,488	0,314	0,136	0,033
p	0,404	0,505	0,022	0,470	0,640	0,012	0,418		0,006	0,091	0,473	0,864
X_9	0,204	0,188	0,122	0,191	-0,044	0,021	0,015	-0,488	1	-0,692	-0,088	0,011
p	0,279	0,321	0,521	0,313	0,816	0,911	0,938	0,006		0,000	0,645	0,956
X_{10}	-0,071	-0,062	-0,157	-0,063	-0,044	-0,222	0,134	0,314	-0,692	1	0,256	-0,297
p	0,710	0,745	0,407	0,740	0,818	0,239	0,481	0,091	0,000		0,172	0,111
X_{11}	-0,154	-0,158	0,142	-0,164	0,281	-0,243	0,279	0,136	-0,088	0,256	1	-0,331
p	0,416	0,406	0,455	0,385	0,133	0,197	0,136	0,473	0,645	0,172		0,074
X_{12}	-0,493	-0,49	-0,033	-0,497	0,142	0,535	-0,483	0,033	0,011	-0,297	-0,331	1
p	0,006	0,006	0,864	0,005	0,454	0,002	0,007	0,864	0,956	0,111	0,074	

* Koyu renkli korelasyon değerleri %5 önem seviyesinde anlamlı olarak tespit edilmiştir.

Tablo 2 incelendiğinde bağımsız değişkenlerin kendi aralarındaki korelasyon değerlerinden yalnızca X_1 (kişi başı millî gelir), X_2 (gayri safi millî hasıla) ve X_4 (gayri safi yurtiçi hasıla) arasında yüksek çıktığı bunların dışındaki bağımsız değişkenlerin kendi aralarındaki korelasyonun çok daha düşük olduğu tespit edilmiştir.

2015 yılı verileri ile beta regresyon modeli kurulduğunda farklı link fonksiyonları için bilgi kriterleri (AIC ve BIC) aşağıdaki Tablo 3'deki şekilde tespit edilmiştir.

Tablo 3
Beta Regresyon Modelinde Farklı Link Fonksiyonları İçin Bilgi Kriterleri

Link Fonksiyonu	AIC	BIC
logit	-92,31602	-72,69925
probit	-93,21182	-73,59505
cloglog	-91,85109	-72,23433
cauchit	-86,13141	-66,51464
log	-91,35715	-71,74039
loglog	-94,15459	-74,53782

Akaike Bilgi ölçütü (AIC) ve Bayesçi benzeri, Bayesçi bilgi ölçütü (BIC), sıradan en küçük kareler yönteminden daha karışık olan modellerde model seçim işlemleri için kullanılmaktadır (Montgomery, 2013). Kurulan olası modellerden en küçük AIC ve BIC değerlerini içeren modelin veri setini en iyi açıklayan model olduğu kabul edilmektedir (Kaya ve Yeşilova, 2012).

Yukarıdaki Tablo 3'de de görüldüğü gibi loglog fonksiyonu en küçük bilgi kriterini veren link fonksiyonu olarak tespit edilmiş ve sonraki aşamalarda loglog fonksiyonu kullanılmıştır.

Tablo 4
Beta Regresyon Analizi Sonuçları

Değişkenler	Tahmin	Standart Hata	Z Değeri	P Değeri
Kesişim	-0,717250	0,023745	-30,206	< 2e-16***
X ₁	-0,838179	0,502858	-1,667	0,095548
X ₂	0,134889	0,261648	0,516	0,606179
X ₃	-0,017354	0,047177	-0,368	0,712978
X ₄	0,689834	0,521671	1,322	0,186050
X ₅	0,053534	0,033330	1,606	0,108232
X ₆	0,042434	0,049209	0,862	0,388510
X ₇	0,192237	0,033065	5,814	6.1e-09***
X ₈	0,001499	0,044900	0,033	0,973362
X ₉	-0,087407	0,039921	-2,190	0,028560*
X ₁₀	-0,140708	0,038094	-3,694	0,000221***
X ₁₁	0,014653	0,037451	0,391	0,695613
X ₁₂	-0,075090	0,041521	-1,808	0,070529

*** 0,001; *0,05; Log-olabilirlik: 61,08; Pseudo R²: 0,7798

Beta regresyon analizinde regresyon tahmini en çok olabilirliğe dayandığından çıktılarda, uygulanan modele ait log-olabilirlik ve pseudo R² değerleri verilmektedir. Pseudo R² değeri, ortalama için doğrusal tahminleyici ve link dönüşümlü bağımlı değişken arasındaki ilişkinin karesi olarak ifade edilmektedir ve $0 \leq \text{pseudo R}^2 \leq 1$ değerini almaktadır. Değerin 1'e yakın olması doğrusal tahminleyici ve link dönüşümlü bağımlı değişken arasında mükemmel uyum olduğunu göstermektedir.

Tablo 4 incelendiğinde X₇ (kişi başı alkol tüketimi), X₉ (işgücü katılım oranı) ve X₁₀ (işsizlik oranı) bağımsız değişkenlerinin %5 önem seviyesinde intihar bağımlı değişkeni üzerinde anlamlı etkiye (p değeri<0,05) sahip oldukları tespit edilmiştir. Özellikle kişi başı alkol tüketiminin ve işsizlik oranının $\alpha=0,001$ önem seviyesinde geçerli çıkması aradaki ilişkinin anlamlılık derecesinin ne kadar yüksek olduğunu ortaya koymaktadır.

2015 yılı bağımlı ve bağımsız değişkenleri kullanılarak beta regresyon analizi yapılmış olup anlamlı çıkan değişkenlerle model kurulmuştur. Kurulan modelde 2016 yılı bağımsız değişkenleri yer

arak 2016 yılı intihar oranları tahmin edilmiştir. Tahmin edilen 2016 yılına ait intihar oranları ve gerçek 2016 yılı intihar oranları Tablo 5’de verilmiştir.

Tablo 5
Beta Regresyon Analizine Ait 2016 Yılı Tahmini Değerler

	Gerçek İntihar Oranları	Tahmini İntihar Oranları	Fark
Avustralya	0,132	0,158	-0,026
Belçika	0,207	0,203	0,004
Çin	0,097	0,104	-0,007
Danimarka	0,128	0,156	-0,028
Mısır	0,040	0,038	0,002
Finlandiya	0,159	0,156	0,003
Fransa	0,177	0,187	-0,010
Almanya	0,136	0,234	-0,098
Macaristan	0,191	0,206	-0,015
Hindistan	0,163	0,141	0,022
İtalya	0,082	0,113	-0,031
Japonya	0,185	0,149	0,036
Kazakistan	0,225	0,104	0,121
Kore Cumhuriyeti	0,269	0,174	0,095
Kırgızistan	0,083	0,094	-0,011
Malezya	0,055	0,052	0,003
Yeni Zelanda	0,121	0,151	-0,030
Norveç	0,122	0,120	0,002
Pakistan	0,029	0,066	-0,037
Polonya	0,162	0,197	-0,035
Portekiz	0,140	0,164	-0,024
Rusya	0,310	0,184	0,126
İspanya	0,087	0,078	0,009
İsveç	0,148	0,130	0,018
İsviçre	0,172	0,167	0,005
Türkiye	0,073	0,052	0,021
Birleşik Arap Emir.	0,028	0,052	-0,024
İngiltere	0,089	0,188	-0,099
Amerika	0,153	0,160	-0,007
Vietnam	0,073	0,115	-0,042

Tablo 5’de yer alan tahmini değerler ve gerçek değerler arasındaki farkların T testi ile anlamlı olup olmadığı yani kurulan regresyon modeli ile yapılan tahmini değerler ile gerçek intihar oranları arasında anlamlı bir farklılığın olup olmadığı araştırılmıştır. Elde edilen sonuçlar Tablo 6’da belirtilmiştir.

Tablo 6
2016 Yılı Beta Regresyon Tahmini Değerleri İle Gerçek Değerler Arası T-Testi Sonuçları

	T Değeri	P Değeri
Farklar	-0,214	0,832
Test Değeri = 0		

H₀: Hata ortalaması sıfırdır

H₁: Hata ortalaması sıfırdan farklıdır

Hipotezleri sınıandığında $p > 0,05$ olduğundan yokluk hipotezi kabul edilmiş, hata ortalamasının sıfır olduğu gösterilmiştir. Beta regresyon analizi ile kurulan model ile yapılan tahminleme sonuçları

ile gerçek sonuçlar arasında anlamlı bir farklılığın olmadığı tespit edilmiştir. Buna göre kurulan modelin geçerli bir tahminleme gücüne sahip olduğu söylenebilmektedir.

İkinci aşamada verilere çoklu doğrusal regresyon analizi uygulanmıştır. Beta regresyon analizindeki adımların izlendiği analizde elde edilen sonuçlar aşağıda yer almaktadır.

Tablo 7
Çoklu Doğrusal Regresyon Anova Tablosu

Model	Kareler Toplamı	df	Kareler Ortalaması	F	P
Regresyon	0,109	12	0,009	3,976	0,005
Artık	0,039	17	0,002		
Toplam	0,147	29			

Modele ait R^2 değeri 0,737 olarak bulunmuştur.

Araştırmada yöntemleri karşılaştırmak adına 2015 yılı verilerine çoklu doğrusal regresyon analizi uygulanmış ve modele dahil edilen değişkenlerle birlikte R^2 değeri verilmiştir. Kurulan modelin anlamlılığı sınamada;

H_0 : Doğrusal model anlamsızdır

H_1 : Doğrusal model anlamlıdır

şeklinde hipotez kurulmuş olup $p < 0,05$ olduğundan dolayı H_0 red edilmiş ve kurulan modelin anlamlı yani geçerli olduğu kabul edilmiştir.

Tablo 8
2015 Yılı Çoklu Doğrusal Regresyon Analizi Sonuçları

Değişkenler	Tahmin	Standart Hata	T Değeri	P Değeri
(Constant)	0,136	0,009	15,592	0,000***
X_1	-0,206	0,186	-1,107	0,284
X_2	0,014	0,099	0,137	0,892
X_3	-0,005	0,018	-0,272	0,789
X_4	0,188	0,196	0,962	0,350
X_5	0,020	0,012	1,574	0,134
X_6	0,007	0,018	0,388	0,703
X_7	0,050	0,012	4,117	0,001*
X_8	0,000	0,017	0,010	0,992
X_9	-0,022	0,015	-1,485	0,156
X_{10}	-0,038	0,014	-2,761	0,013*
X_{11}	0,007	0,014	0,504	0,621
X_{12}	-0,015	0,014	-1,070	0,300

*** 0,001; *0,05

Tablo 8 incelendiğinde X_7 (kişi başı alkol tüketimi) ve X_{10} (işsizlik oranı) bağımsız değişkenlerinin %5 önem seviyesinde bağımlı değişken olan intihar üzerinde anlamlı etkiye (p değeri $< 0,05$) sahip oldukları tespit edilmiştir. Yapılacak tahmin için oluşturulacak modelde bu değişkenler yer alacaktır. Tablo 9’da kurulan modellerle elde edilen tahmini değerler ve gerçek değerler yer almaktadır. Ayrıca hata olarak isimlendirilen, gerçek değer ile tahmin değeri arasındaki farkta tabloda yer almaktadır.

Tablo 9
Çoklu Doğrusal Regresyona Ait 2016 Yılı Tahmini Değerler

	Gerçek İntihar Oranları	Tahmini İntihar Oranları	Fark
Avustralya	0,132	0,172	-0,040
Belçika	0,207	0,170	0,037
Çin	0,097	0,138	-0,041
Danimarka	0,128	0,164	-0,036
Mısır	0,040	-0,034	0,074
Finlandiya	0,159	0,141	0,018
Fransa	0,177	0,154	0,023
Almanya	0,136	0,225	-0,089
Macaristan	0,191	0,188	0,003
Hindistan	0,163	0,136	0,027
İtalya	0,082	0,069	0,013
Japonya	0,185	0,163	0,022
Kazakistan	0,225	0,140	0,085
Kore Cumhuriyeti	0,269	0,187	0,082
Kırgızistan	0,083	0,097	-0,014
Malezya	0,055	0,064	-0,009
Yeni Zelanda	0,121	0,179	-0,058
Norveç	0,122	0,140	-0,018
Pakistan	0,029	0,056	-0,027
Polonya	0,162	0,180	-0,018
Portekiz	0,140	0,140	0,000
Rusya	0,310	0,188	0,122
İspanya	0,087	0,022	0,065
İsveç	0,148	0,140	0,008
İsviçre	0,172	0,192	-0,020
Türkiye	0,073	0,004	0,069
Birleşik Arap Emir.	0,028	0,121	-0,093
İngiltere	0,089	0,193	-0,104
Amerika	0,153	0,169	-0,016
Vietnam	0,073	0,181	-0,108

Tablo 9’da yer alan tahmini değerler ve gerçek değerler arasındaki farkların T testi ile anlamlı olup olmadığı yani kurulan regresyon modeli ile yapılan tahmini değerler ile gerçek intihar oranları arasında anlamlı bir farklılığın olup olmadığı araştırılmıştır. Elde edilen sonuçlar Tablo 10’da belirtilmiştir.

Tablo 10
2016 Yılı Çoklu Doğrusal Regresyon Tahmini Değerleri İle Gerçek Değerler Arası T-Testi Sonuçları

	T Değeri	P Değeri
Farklar	-0,136	0,893
Test Değeri = 0		

H₀: Hata ortalaması sıfırdır

H₁: Hata ortalaması sıfırdan farklıdır

Hipotezleri sınanıldığında $p > 0,05$ olduğundan yokluk hipotezi kabul edilmiş, hata ortalamasının sıfır olduğu gösterilmiştir. Çoklu doğrusal regresyon analizi ile kurulan model ile yapılan tahminleme sonuçları ile gerçek sonuçlar arasında anlamlı bir farklılığın olmadığı tespit edilmiştir. Buna göre kurulan modelin geçerli bir tahminleme gücüne sahip olmasını sağlayan varsayımın sağlanmış olduğu söylenebilmektedir.

İki farklı model kullanılan araştırmalarda modelin tahminleme de hata payının daha düşük olması, hangi modelin daha geçerli olduğunu belirlemektedir. Beta regresyon analizi ve çoklu doğrusal regresyon analizi sonucunda kurulan modellerden geçerli olanı saptamak adına “Hata Kareler Ortalamasının Karekökü” (Root Mean Square Error – RMSE) istatistiğinden yararlanılmıştır. RMSE’ye ait formülasyon aşağıda verilmektedir. Elde edilen sonuçlar Tablo 11’de yer almaktadır.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - y'_i)^2}{n}}$$

Tablo 11
RMSE Sonuçları

Yöntem	RMSE Değeri
Beta Regresyon Analizi	0,049
Çoklu Doğrusal Regresyon Analizi	0,057

Tablo 11’deki sonuçlara göre beta regresyon analizi daha düşük RMSE değerine sahiptir. Bu da kurulan modelin daha düşük hata payı ile tahminleme yaptığını ifade etmektedir. Bu nedenle verinin (0,1) arasında oransal olduğu durumlarda çoklu doğrusal regresyon analizi yerine beta regresyon analizini tercih etmek daha doğru sonuçlar elde edilmesini sağlamaktadır.

Sonuç, Tartışma ve Öneriler

Günlük hayat veri setleri çoğunlukla normal dağılım şartını sağlayamamaktadır. Şartın sağlanmadığı durumda beta dağılımı çarpık dağılıma sahip veri setlerinde esnek kullanıma sahiptir. Literürde beta regresyon analizinin yüzdeler verilerde normal regresyon analizinden daha iyi sonuçlar verdiği görülmektedir. Yellareddygarı ve diğerleri (2016), çalışmalarında depolama sırasında patates yumrular içinde pembe çürüme gelişimini tahminlemede istatistiksel olarak beta regresyon analizinin doğrusal regresyondan daha iyi sonuçlar verdiğini göstermişlerdir. Ünlü ve Aktaş (2017), refah endeksi verilerini modellemede beta regresyon analizinden yararlanmışlardır. Beta regresyon analizinin doğrusal regresyon analizine göre daha doğru ve etkin sonuçlar verdiğini ortaya koymuşlardır.

Bu çalışmada bağımlı değişkenin (0,1) arasında oransal olduğu ve/veya beta dağılım gösterdiği durumda beta regresyon analizi ve çoklu doğrusal regresyon analizi yapılmıştır. İstatistiksel olarak iki yöntemin tahmin gücünün karşılaştırıldığı bu çalışmada elde edilen sonuçlar veri yapısına göre beta regresyon analizinin daha iyi çalıştığını göstermiştir. Yöntemler karşılaştırılırken intihar verileri kullanılmıştır. İntihar verileri farklı analiz yöntemleri ile çalışılmış beta regresyon analiziyle yapılmış bir çalışmaya rastlanmamıştır. Analizler sonucunda kişi başı alkol tüketiminin, işsizlik oranının ve işgücü katılım oranının intihar üzerinde etkisi olduğu görülmüştür. Çalışmada elde edilen bulgular literatürde yer alan çalışmalarla uyumaktadır. Topbaş (2007), çalışmasında işsizliğin intihar üzerinde anlamlı bir etkiye sahip olduğunu tespit etmiştir. Darvishi, Farhadı, Haghtalab ve Poorolajal (2015), yaptıkları çalışmada, dengesiz alkol tüketiminin intihar düşünce riskini anlamlı ölçüde etkilediğini göstermişler ve de intiharı arttırdığına dair kanıtlar ortaya koymuşlardır. Tunalı ve Özkaya (2016), çalışmalarında intihar ve işsizlik arasında çift yönlü bir nedensellik ilişkisi saptamışlardır. Durmaz (2016), çalışmasında ülkelerin gelişmişlik düzeylerine bağlı olarak özellikle kadınların istihdam sorununun dünya genelinde ortak bir sorun olmaya devam ettiğini ve ekonomik önlemlerin bu durumun önüne geçebileceğini vurgulamıştır.

Çalışmada elde edilen bulgulara göre yüzdeler verilerde beta regresyon analizinin daha iyi çalıştığı gösterilmiştir. Elde edilen bulgulara dayanarak analizler sonucunda ortak değişken olarak karşımıza çıkan alkol tüketimi ve işsizlik oranının intihar üzerinde istatistiksel olarak anlamlı bir

etkiye sahip olduğu görülmüştür. Literatüre yer alan çalışmalara dayanarak, ülke politikasına bağlı olarak değişkenlik gösteren işsizlik oranındaki düşüşün ve bireysel bir etken olan alkol tüketimindeki azalmanın intihar oranını düşük seviyelere çekmekte etkili olacağı tahmin edilmektedir. Farklı ekonomik ve sosyal göstergeler kullanılarak farklı ülkeler için ya da farklı bir çalışma grubu için ilerleyen çalışmalarda araştırmayı genişletmek mümkün olabilecektir.

Açıklamalar: Bu çalışma, Dr. Öğr. Üyesi Elif Bulut danışmanlığında Veysel Gökhan Aydın'a ait savunmak üzere tamamlanmış yüksek lisans tezinden farklı veri kullanılarak türetilmiş bir çalışmadır.

Kaynaklar

- Amini, P., Ahmadiania, H., Poorolajal, J. and Moqaddasi Amiri, M. (2016). Evaluating the high-risk groups for suicide: a comparison of logistic regression, support vector machine, decision tree and artificial neural network. *Iranian Journal of Public Health*, 45(9), 1179-1187.
- Atasoy, A. (2018). *Türkiye’de net göç hızına etki eden faktörlerin beta regresyonu ile incelenmesi*. Yüksek Lisans Tezi, Kafkas Üniversitesi, Kars.
- Chen, K.K., Chiu, R-H and Chang, C-T. (2017). Using beta regression to explore the relationship between service attributes and likelihood of customer retention for the container shipping industry. *Transportation Research Part E*, 104, 1-16.
- Cuervo-Cepeda, E. (2015). Beta regression models: joint mean and variance modeling. *Journal of Statistical Theory and Practice*, 9(1), 134-145.
- Darvishi, N., Farhadi, M., Haghtalab, T. and Poorolajal, J. (2015). Alcohol-related risk of suicidal ideation, suicide attempt, and completed suicide: A Meta-Analysis. *PLOS ONE*, 10(5), 1-14.
- Durmaz, Ş. (2016). İşgücü piyasasında kadınlar ve karşılaştıkları engeller. *Ahi Evran Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 2(3), 37-60.
- Ekici, G., Savaş, H. ve Çıtak, S. (2001). İntihar riskini arttıran psikososyal etmenler. *Anadolu Psikiyatri Dergisi*, 2(4), 204-212.
- Fazel, S., Wolf, A., Larsson, H., Mallett, S. and Fanshawe, T.R. (2019). The prediction of suicide in severe mental illness: development and validation of a clinical prediction rule (OxMIS). *Translational Psychiatry*, 9(1), 98.
- Ferrari, S. and Cribari, F. (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, 31(7), 799-815.
- Gayawan, E., Fasusi, O.D and Bandyopadhyay, D. (2020). Structured additive distributional zero augmented beta regression modeling of mortality in Nigeria. *Spatial Statistics*, 35, 1-14.
- Kaya, Y. ve Yeşilova, A. (2012). Karışıklı ikili lojistik regresyon modeline ilişkin bir uygulama. *Dokuz Eylül Üniversitesi Mühendislik Fakültesi Mühendislik Bilimleri Dergisi*, 14(1), 39-47.
- Koç, T. (2019). Türkiye’de boşanma oranlarını etkileyen faktörlerin beta regresyon modeli ile belirlenmesi. *Avrasya Uluslararası Araştırmalar Dergisi*, 7(16), 1111-1117.
- Korkmaz, M., Germir, H.N., Dündar, E., Şen, E., Karta, N. ve Şahbudak, E. (2015). Dünya ülkelerindeki enflasyonu etkileyen faktörlerin ve genel iktisadi yapının iki aşamalı kümeleme ve beta regresyon analizi ile belirlenmesi. *Uluslararası Hakemli Ekonomi Yönetimi Araştırmaları Dergisi*, 6, 16-26.

- Mandal. S., Srivastav, R.K. and Simonovic, S.P. (2016). Use of beta regression for statistical downscaling of precipitation in the Campbell River basin, British Columbia, Canada. *Journal of Hydrology*, 538, 49-62.
- Miché, M., Studerus, E., Meyer, A.H. and Gloster, T. (2020). Prospective prediction of suicide attempts in community adolescents and young adults, using regression methods and machine learning. *Journal of Affective Disorders*, 265, 570-578.
- Montgomery, D.C., Peck, E.A. and Vining, G.G. (2013). *Doğrusal regresyon analizine giriş* (Çev. Ö. V. Çilengiroğlu, Ö. G. Alma, E. Bulut, A. K. Şehirlioğlu). Ankara: Nobel Yayınevi.
- Özçiftçi, V. (2020). Sosyal medyanın tüketicilerin satın alma niyeti üzerindeki etkisi. *Ahi Evran Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 6(1), 206-223.
- Simiksaran, E. (2011). *Teori ve uygulamalarıyla istatistiksel yöntemler*. İstanbul: Türkmen Kitabevi.
- Swearingen C., Tilley B.C., Adams R.J. and Rumbolt Z. (2011). Application of beta regression to analyze ischemic stroke volume in ninds rt-pa clinical trials. *Neuroepidemiology*, 37(2), 73-82.
- Topbaş G. (2007). İşsizlik ve intihar ilişkisi: 1975-2005 var analizi. *Karamanoğlu Mehmetbey Üniversitesi Sosyal ve Ekonomik Araştırmalar Dergisi*, 2, 161-172.
- Tunalı H. ve Özkaya S. (2016). Türkiye’de işsizlik – intihar ilişkisinin analizi. *Kırklareli Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 5(2), 56-70.
- Türkan S. ve Özel G. (2017). 2014-2015 öğretim yılında Türkiye’de devlet üniversitelerinin etkinlikleri ve etkinliğe etki eden faktörlerin belirlenmesi. *Eğitim ve Bilim*, 42(191), 307-322.
- Ünlü H. ve Aktaş S. (2017). Beta regression for the indicator values of well-being index for provinces in Turkey. *Journal of Engineering Technology and Applied Sciences*, 2(2), 101-111.
- Poznyak V. and Rekve D. (2018). *Global status report on alcohol and health 2018* (Technical Report). World Health Organization web sayfası: <https://apps.who.int/iris/bitstream/handle/10665/274603/9789241565639-eng.pdf?ua=1>. (Erişim Tarihi: 02.03.2020)
- The World Bank. Data Bank World Development Indicators. The World Bank web sayfası: <https://databank.worldbank.org/reports.aspx?source=2>. (Erişim Tarihi: 02.03.2020)
- Yellareddygarı S. KR., Pasche J.S., Taylor R.J., Hua S. and Gudmestad N.C. (2016). Beta regression model for predicting the development of pink rot in potato tubers during storage. *The American Phytopathological Society*, 100(6), 1118-1124.
- Zaman T., Dunder E. ve Aydın S. (2019). Gini katsayısını etkileyen faktörlerin beta regresyon yöntemi yardımı ile belirlenmesi. *Erzincan Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 12(1), 235-240.

Extended Abstract

Introduction

Suicide is settled among the major social problems. This problem is not only a problem recently, it is still an ongoing phenomenon. As suicide has changed country to country, it can also be seen in different ways and proportions within a community. In this study, some variables in the economic and social structures of the societies were examined whether they have significant or non-significant effect on the suicide levels of the societies. Within the scope of the study, this effect was tried to be determined by using suicide rates and some economic-social indicators of 30 different countries.

Method

Linear regression model is not suitable where the dependent variable is limited in the range of (0,1). Because it can give estimates that exceed its lower and upper limit values for the variable of interest. In this case, it is suitable to work with the regression model where the dependent variable is measured in the standard unit range, $0 < y < 1$. The proposed model is based on the assumption that the dependent variable is beta distributed. Beta distribution is very flexible to model the proportions since the density of the two parameters indexing the distribution takes quite different shapes. Beta distribution has the application field such as fraction and random change of uncertainty or probability. Therefore, this distribution has widespread application in educational, social and financial sciences where the random variable is continuous in the range of (0,1). Beta (p, q) can be reparameterized again as a function of mean and variance. This characterization of beta distribution may be appropriate in the definition of the beta regression model.

In the first reparameterization;

$$p = \mu\varnothing \text{ and } q = \varnothing(1 - \mu)$$

$$\sigma^2 = \frac{\mu(1 - \mu)}{\varnothing + 1}$$

In this case, \varnothing is interpreted as the precision parameters for the constant values of μ , and the large values of \varnothing are interpreted as corresponding to the shrinking values of the variance of y . Then;

$$f(y; \mu, \varnothing) = \frac{\Gamma(\varnothing)}{\Gamma(\mu\varnothing)\Gamma((1 - \mu)\varnothing)} y^{\mu\varnothing-1}(1 - y)^{(1-\mu)\varnothing-1} \quad 0 < y < 1$$

$$0 < \mu < 1 \text{ and } \varnothing > 0 \text{ olur.}$$

Assuming that the percentage of dependent variable is beta distributed, a beta regression model can be designed. $y_1, y_2, \dots, y_n \sim B(\mu, \varnothing)$ [$y \sim B(\mu, \varnothing)$] is a random sample in beta density. Then the beta regression is defined as follows.

$$g(\mu_i) = \beta_0 + x_{i1}\beta_1 + \dots + x_{ik}\beta_k = \eta_i \quad i = 1, 2, \dots, n$$

x_{i1}, \dots, x_{ik} 's are explanatory variables, $\beta_0, \beta_1, \dots, \beta_k$ are estimated intersections and coefficients corresponding to each independent variable. n is the sample size and η_i is the linear predictor of the observations.

Multiple linear regression analysis was used as an another method in this study. Where two or more independent variables are included in the data set, the mathematical model between independent

variables and dependent variable is established by multiple linear regression analysis. Since multiple linear regression analysis, which is one of the well-known analyzes, is widely used in every field of science and is used in the study, the title has been opened and no details have been entered in the narrative.

Findings

All the data used in this study were obtained from World Bank Databank and World Health Organization (WHO). 30 different countries were examined in this study with 1 dependent and 12 independent variables. The main reason for the less number of variables and countries is that the dependent and independent variables included in the evaluation were identified completely for the date 2016. For this reason, the data of 2015 and 2016 were studied. It is aimed to estimate suicide rates for 2016 by establishing a regression model based on the data for 2015 by evaluating economic and social independent variables and suicide rates.

When multiple linear regression analysis results were examined, it was found that the independent variables such as alcohol consumption per person, labor force participation rate and unemployment rate had a significant effect on the suicide-dependent variable at 5% significance level. In particular, the fact that alcohol consumption per capita and unemployment rate was valid at $\alpha = 0.001$ significance level reveals the high level of significance of the relationship.

Beta regression analysis results showed that alcohol consumption per person and unemployment rate had a significant effect on the dependent variable at the 5% significance level. These variables were included in the model developed for the estimation. The estimated values were obtained for each regression model and compared with the actual values. RMSE statistic was used to determine the validity of the models. This statistics shows the estimation power of the model.

According to the results, beta regression analysis has lower RMSE value. This means that the established model makes estimation with a lower error margin. Therefore, in cases where the data is proportional between (0,1), preferring beta regression analysis instead of multiple linear regression analysis provides more accurate results.

Results, Discussion and Conclusion

Daily life data sets almost do not show normal distribution requirement in analysis. If the requirement is not met, the beta distribution has flexible use in data sets with skewed distribution. It is seen that beta regression analysis gives better results in percentage data than normal regression analysis. In the study, it was found that per capita alcohol consumption, unemployment rate and labor force participation rate had an effect on suicide. Based on the findings, it is estimated that the decrease in the unemployment rate and per capita alcohol consumption will be effective in reducing the suicide rate.