

Bayes Açığortay Regresyon Tekniği ve Bir Uygulama

Ece ÖZGÖREN¹, Sinan SARAÇLI*²

¹Eskişehir Teknik Üniversitesi, Fen Fakültesi, İstatistik Bölümü, 26470, Eskişehir, Türkiye

²Afyon Kocatepe Üniversitesi, Fen-Edebiyat Fakültesi, İstatistik Bölümü, 03200, Afyonkarahisar, Türkiye

(Alınış / Received: 10.03.2021, Kabul / Accepted: 11.10.2021, Online Yayınlanma / Published Online: 25.12.2021)

Anahtar Kelimeler
Bayesci Yaklaşım,
Tip II Regresyon,
Bayes Regresyon,
Ölçüm Hatalı Modeller
İstatistiksel Modelleme

Özet: Bu çalışmanın amacı Bayes Tip II regresyon tekniğinin performansını incelemektir. Bu amaçla gerçek bir veri seti üzerinde Bayes yaklaşımı yardımı ile basit doğrusal regresyon ve açığortay regresyon denklemleri hesaplanmıştır. Daha önceki çalışmalarda Tip II regresyon teknikleri arasında en iyi performansı sergileyen tekniğin açığortay tekniği olarak belirtilmesinden dolayı mevcut veri seti için sırasıyla X ve Y değişkenleri bağımlı değişken olarak ele alınarak regresyon denklemleri elde edilmiş, daha sonra elde edilen bu iki regresyon denkleminin açığortayı alınarak Bayes açığortay denklemi hesaplanmıştır. Önsel ve mevcut bilgi verisine dayalı sonsal dağılımları elde etmek amacıyla farklı örneklem hacimlerinden yararlanılmış ve Bayes regresyon denklemlerinin performansları HKO kriterine göre karşılaştırılmıştır. Araştırma bulgularına göre n=100 ve n=30 birimlik örneklemelerde Bayes açığortay tekniğinin performansının daha düşük HKO değerine sahip olduğu, dolayısıyla mevcut veri setine ait bu örneklem hacimleri için en iyi performansı sergilediği belirlenmiştir.

Bayesian Bisector Regression Technique and an Application

Keywords
Bayesian Approaches,
Type II Regression,
Bayesian Regression,
Measurement Error Models
Statistical Modeling

Abstract: The purpose of this study is to examine the performance of Bayesian Type II regression Analysis. With this purpose, simple linear regression and bisector regression equations are calculated by the help of Bayesian approach on a real data set. Because in the earlier studies its mentioned that the best technique among Type II regression techniques is the Bisector regression technique, regression equations are obtained by considering the X and Y variables as the dependent variable respectively and then the Bayesian Bisector equation is calculated by bisecting these two regression lines. Different sample sizes are considered to obtain the posterior distribution based on prior and likelihood information and then the performances of Bayesian regression equations are compared according to MSE criteria. The results of the study indicates that performance of Bayesian bisector technique has the minimum MSE for the sample sizes n=100 and n=30 which means that the performance of Bayesian bisector technique is the best for these sample sizes for the related data set.

1. Giriş

Bayesci yaklaşımın ilk temeli, İngiltere Tunbridge Wells'de yaşayan, bir rahip ve matematikçi olan Thomas Bayes tarafından düşünülmüş ve ölümünden 2 yıl sonra (1963) arkadaşı Richard Price' in Bayes' in çalışma kâğıtlarını bulması ve ardından yayınladığı bir makale olan "An Essay Towards Solving a Problem in the Doctrine of Chances" ile ortaya konmuştur [1].

Bayesci yaklaşımın özü, ele alınan bir çalışmada, konu ile ilgili gerçekleşen tüm bilginin analize dahil edilmesine dayanmaktadır. Bu yaklaşımda parametre,

olasılık dağılımına sahip bir rasgele değişken olarak kabul edilmekte ve parametre ile ilgili çıkarsama, önsel bilgi ile mevcut bilgi birleştirilerek oluşturulan sonsal dağılım aracılığı ile yapılmaktadır. Bununla birlikte klasik yaklaşımda ise parametre bilinmeyen bir sabit olarak kabul edilerek, parametre tahmini sadece mevcut veriler ile yapılmaktadır [2].

Ayrıca bu yaklaşımın diğer bir önemli unsuru ise küçük örneklem hacimlerinde yapılan parametre tahminlerinde klasik yaklaşımlara göre daha iyi sonuçlar vermesidir. Klasik yaklaşımda zaman ve maliyet tasarrufu sağlamak amacıyla daha küçük

*İlgili yazar: ssaracali@aku.edu.tr

örneklem tercih edilirken parametre tahminleri güvenilirliğini yitirebilir ancak Bayesci yaklaşımda küçük örneklem ile güvenilir sonuçlar elde etmek mümkündür [3].

Kısaca Bayes yaklaşımı, aslı Bayes Teoremine dayandırılarak yapılandırılmış bir yaklaşım sistemidir. Bayes teoremi, olasılık kuramı içinde incelenen önemli bir konudur. Bu teorem bir rasgele değişken için olasılık dağılımı içinde koşullu olasılıklar ile marjinal olasılıklar arasındaki ilişkiyi gösterir [4]. İfade edilen tüm bu özellikler klasik yaklaşım yerine Bayesci yaklaşımın kullanılması için önemli bir unsur oluşturmaktadır.

Tip II Regresyon tekniklerinin kullanılmasındaki amaç, bağımlı ve bağımsız değişkenlerin ölçümüne ilişkin oluşabilecek tüm ölçüm hatalarını analize dahil ederek bir model oluşturabilmektir. Günümüzde kullanılan klasik regresyon tekniklerinde bağımsız değişkenlerden kaynaklanabilecek hatalar göz ardı edilerek sadece bağımlı değişken(ler)den kaynaklanan hatalar ile modeller oluşturulmaktadır. Ancak günlük hayatta bağımsız değişkenlerin de ölçüm hataları içerdiği bilinmektedir [5,6]. Bu nedenle klasik regresyon ile oluşturulan modellere kıyasla her iki ölçüm hatasını da içeren Tip II Regresyon tekniğinin kullanılması daha uygun olacaktır.

Bu çalışmada ise, Tip II regresyon tekniklerinden biri olan EKK-Açıortay tekniğine farklı bir bakış açısı getirilerek, parametre tahminleri EKK tekniği yerine Bayes yaklaşımı ile yapılarak Bayes-Açıortay tekniği ile regresyon denklemleri oluşturulacaktır. Ayrıca ilk defa kullanılan bu tekniğin performansı HKO kriterlerine göre değerlendirilecektir.

2. Materyal ve Metot

2.1. Regresyon Analizi

Regresyon analizi, değişkenler arasındaki neden sonuç ilişkisini bağımlı ve bağımsız değişkenler aracılığı ile açıklamaktadır. Regresyon analizi uygulamalarında bir bağımlı ve bir bağımsız değişken bulunması durumunda basit regresyon modeli, bir bağımlı değişken ve birçok bağımsız değişken bulunması durumunda ise çoklu regresyon modeli kullanılmaktadır [7].

Kurulan bir matematiksel regresyon modelinin doğrusal olup olmamasına göre farklı varsayımlar bulunmaktadır. Doğrusal regresyon modelleri bağımlı ve bağımsız değişken fonksiyonlarının aralarındaki ilişkinin doğrusal olması durumunda kullanılmaktadır [7]. Bu varsayımlar altında kurulan modelin parametre tahminlerine dair çeşitli yöntemler geliştirilmiştir. Bu yöntemler, en çok olabilirlik tahmin

edici (EÇO), en küçük kareler tahmin edicisi (EKK), momentler tahmin edicisi ve Bayesci tahmindir [8].

Buna ek olarak, Regresyon teknikleri, sadece bağımlı değişkenin değerinin kestirilmesinde ortaya çıkabilecek hatanın dikkate alınması veya ele alınacak olan tüm değişkenlerin elde edilmesinde oluşabilecek hataların dikkate alınması bakımından incelendiğinde Tip I ve Tip II regresyon tekniği olarak iki bölüme ayrılmaktadır [5].

Tip II Regresyon tekniğinde hata, Saraçlı (2008) tarafından yapılan çalışmada şu şekilde açıklanmıştır; "Tip II regresyon tekniklerinde genel olarak minimize edilmek istenen hata, gerçek değerlerden ε ve δ büyüklüklerinde, çeşitli ölçüm hatası sebepleri ile yanlış ölçülmüş olan x_i ve y_i gözlem değerlerinin, tahmin edilmek istenen regresyon doğrusu üzerinde yer alan tahmini değerlerine olan dik ya da belirli bir açı ile olan uzaklıklarının karesidir."

2.1.1 En Küçük Kareler (EKK) Açığortay Tekniği

Bu teknik, sırası ile Y ve X değişkenlerini bağımlı ve bağımsız değişken olarak dikkate alarak iki ayrı regresyon doğrusunun açıortay doğrusunu elde ederek çözümlemeyi gerçekleştirir. Elde edilen bu açıortay doğrusu hem X hem de Y'deki hataları dikkate aldığından klasik EKK tekniğinden ayrılarak Tip II Regresyon tekniği olarak adlandırılır [9]. Saraçlı (2008) çalışmasında, Tip II Regresyon teknikleri arasında EKK-Açıortay tekniğinin diğer tekniklere göre daha iyi bir sonuç verdiğini ortaya koymuştur.

EKK(X|Y) regresyon doğrusu için elde edilen eğim katsayısı $\hat{\beta}_1$ ve EKK(Y|X) regresyon doğrusu için elde edilen eğim katsayısı $\hat{\beta}_2$ olmak üzere, EKK-Açıortay doğrusuna ait eğim katsayısı, sabit katsayı ve bu katsayılarla ilişkin varyans ve kovaryanslar Eşitlik (1)-(4)'de verildiği gibi hesaplanabilir [9].

$$\hat{\beta}_{1AO} = (\hat{\beta}_1 + \hat{\beta}_2)^{-1} \left[\hat{\beta}_1 \hat{\beta}_2 - 1 + \sqrt{(1 + \hat{\beta}_1^2)(1 + \hat{\beta}_2^2)} \right] \quad (1)$$

$$\hat{\beta}_{0AO} = \bar{y} - \hat{\beta}_{0A1} \bar{x} \quad (2)$$

$$\begin{aligned} & Var(\hat{\beta}_{AO}) \\ &= \frac{\hat{\beta}_{AO}^2}{(\hat{\beta}_1 + \hat{\beta}_2)^2 (1 + \hat{\beta}_1^2) (1 + \hat{\beta}_2^2)} \left[(1 + \hat{\beta}_2^2)^2 Var(\hat{\beta}_1) \right. \\ &+ 2(1 + \hat{\beta}_1^2)(1 + \hat{\beta}_2^2) Cov(\hat{\beta}_1, \hat{\beta}_2) + (1 + \hat{\beta}_1^2)^2 Var(\hat{\beta}_2) \left. \right] \quad (3) \end{aligned}$$

$$Cov(\hat{\beta}_1, \hat{\beta}_2) = (\hat{\beta}_1 S_{xx}^2)^{-1} \left\{ \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) [y_i - \bar{y} - \hat{\beta}_1(x_i - \bar{x})] [y_i - \bar{y} - \hat{\beta}_2(x_i - \bar{x})] \right\} \quad (4)$$

2.1.2 Bilgi Veren Önsel Dağılım Kullanılarak Uygulanan Bayes Regresyon Tekniği

Bilgi veren önsel dağılım ile analiz yapılması halinde araştırmacı çoğu zaman tahmin edilmek istenen parametreye dair bilgiye sahip değildir. Ancak araştırmacının önsel bilgisi daha önceki yapılan araştırmalardan ve teorik bilgisine dayalı olarak oluşturulabilmektedir. Bu aşamada önsel bilgiyi en iyi şekilde yansıtan ve benzerlik fonksiyonu ile matematiksel olarak daha kolay bir şekilde birleşebilen dağılımlar kullanmak mümkündür. Birçok durumda, “Doğal Eşlenik Önsel Dağılımlar” bu özellikleri sağlayacak yapıya sahiptir. Doğal Eşlenik Önsel Dağılım ile olabilirlik fonksiyonunun birleşmesi ile ortaya çıkan sonsal dağılım, Doğal Eşlenik Önsel ile aynı dağılıma sahip olacaktır [10].

Bu bölümde doğal eşlenik önsel dağılım kullanılarak, bilgi veren önsel dağılım yardımı ile sonsal dağılım elde edilecektir. Bayes teoremi doğrusal regresyon modeline uyarlanmadan önce kısaca aşağıdaki gibi yazılı olarak ifade edilebilir.

$$\text{sonsal dağılım} \propto \text{olabilirlik fonksiyonu} \times \text{önsel dağılım} \quad (5)$$

Yukarıda ifade edilen teoreme dayanarak regresyon modelinin oluşturulabilmesi için öncelikle önsel dağılım oluşturulmalı ve daha sonra olabilirlik fonksiyonu ile birleştirilerek sonsal dağılım elde edilmelidir. Bu aşamada elde edilen Bayes ilkeleri Lineer Doğrusal Regresyon modeline uyarlanarak Lineer Bayes Regresyon Modeli elde edilebilmektedir.

$$y = X\beta + \varepsilon \quad (6)$$

Lineer Regresyon modeli olan Eşitlik (6)'ya göre, y , $(T \times 1)$ büyüklüğünde bağımlı değişkenlerin gözlem vektörü, X , K açıklayıcı değişkenleri üzerinde $(T \times K)$ boyutlu bir gözlem matrisi, ε , $\varepsilon \sim N(0, \sigma^2)$ dağılıma sahip $(T \times 1)$ boyutunda bir hata vektörü ve β ve σ hakkında bilgi edinmek istediğimiz parametreleri ifade etmektedir. Burada $\theta = (\beta', \sigma)'$ olmak üzere Eşitlik (5) yardımı ile Eşitlik (7) elde edilebilir [10].

$$g(\beta, \sigma | y) \propto l(\beta, \sigma | y) g(\beta, \sigma) \quad (7)$$

Bayesci yaklaşım için olabilirlik fonksiyonu Eşitlik (8)'de belirtildiği gibi yazılabilir [11].

$$l(\beta, \sigma | y) \propto \frac{1}{(2\pi)^{T/2}} \frac{1}{\sigma^T} \exp \left\{ -\frac{1}{2\sigma^2} [(y - X\beta)'] (y - X\beta) \right\} \quad (8)$$

Burada $(b', \hat{\sigma}^2)$ yeterli istatistik olmak üzere Eşitlik (9) yazılabilir.

$$b = (X'X)^{-1} X'y, \hat{\sigma}^2 = (y - Xb)'(y - Xb)/v, \quad (9)$$

$$v = T - K$$

Eşitlik (9)'da belirtilen ifadeler EKK yöntemi ile elde edilen parametre tahmin edicileridir. Yukarıdaki denklemler kullanılarak elde edilecek olan $\bar{\beta}$, \bar{v} ve \bar{s}^2 ifadeleri önsel dağılıma ilişkin tahmin edicileri $\bar{\beta}$, \bar{v} ve \bar{s}^2 ifadeleri ise sonsal dağılıma ilişkin tahmin edicileri göstermektedir [12; 13].

Bayesci yaklaşım kullanılarak önsel dağılıma ilişkin β ve σ parametresinin ortalama ve varyans değerleri aşağıdaki eşitlikler ile elde edilir.

$$E[\beta | \sigma] = E[\beta] = \bar{\beta}, \text{cov}[\beta | \sigma] = A^{-1}\sigma^2, \quad (10)$$

$$A = (X'X)^{-1} X'$$

$$E[\sigma] = \frac{\Gamma[(\bar{v} - 1)/2]}{\Gamma(\bar{v}/2)} \left(\frac{\bar{v}}{2}\right)^{1/2} \bar{s}, \quad (11)$$

$$E[\sigma^2] = \frac{\bar{v}\bar{s}^2}{\bar{v} - 2}$$

A simetrik bir matris ve $A = A^{1/2}A^{1/2}$ olmak üzere Eşitlik (12) yazılırsa;

$$w = \begin{pmatrix} A^{1/2}\bar{\beta} \\ y \end{pmatrix}, W = \begin{pmatrix} A^{1/2} \\ X \end{pmatrix} \quad (12)$$

Olmak üzere;

Parametre tahmincileri;

$$\bar{\beta} = (W'W)^{-1} W'w = (A + X'X)^{-1} (A\bar{\beta} + X'y) \quad (13)$$

$$\bar{v}\bar{s}^2 = \bar{v}\bar{s}^2 + (w - W\bar{\beta})'(w - W\bar{\beta}) = \bar{v}\bar{s}^2 + y'y + \bar{\beta}'A\bar{\beta} - \bar{\beta}'(X'X)\bar{\beta} \quad (14)$$

$$\bar{v} = T + \bar{v} \quad (15)$$

ve Bayes regresyon modeli;

$$g(\beta, \sigma | y) \propto \sigma^{-\bar{v}-K-1} \exp \left\{ -\frac{1}{2\sigma^2} \left[\bar{v}\bar{s}^2 + (w - W\bar{\beta})'(w - W\bar{\beta}) + (\beta - \bar{\beta})'W'W(\beta - \bar{\beta}) \right] \right\} \quad (16)$$

$$g(\beta, \sigma | y) \propto \sigma^{-K} \exp \left[-\frac{1}{2\sigma^2} (\beta - \bar{\beta})'(A + X'X)(\beta - \bar{\beta}) \right] \cdot \sigma^{-(\bar{v}+1)} \exp \left(-\frac{\bar{v}\bar{s}^2}{2\sigma^2} \right) \quad (17)$$

$$g(\beta, \sigma | y) \propto g(\beta | \sigma, y) \cdot g(\sigma | y) \quad (18)$$

elde edilir. Burada β parametresi için koşullu yoğunluk fonksiyonu $\bar{\beta}$ ortalamalı, $\sigma^2(A + X'X)^{-1}$ kovaryans matrisi ile “çok değişkenli normal” dağılıma sahiptir ve σ parametresi için de marjinal sonsal yoğunluk fonksiyonu \bar{v} ve \bar{s}^2 parametreleri ile “Ters-Gamma” dağılımına uymaktadır [10].

Bu çalışmada, gerçek veriler üzerinde Bayes-Açırtay tekniği ile farklı örneklem hacimlerinde regresyon modeli oluşturularak, HKO kriterine göre performansın değerlendirilmesi amaçlanmıştır. Bu amaçla, Saraçlı ve Çelik'in (2012) çalışmalarında [14] kullandığı, Van Yüzüncüyıl Araştırma Hastanesi Acil Servisine gelen 100 hastadan derlenen koltuk altı (aksiller) ve kulaktan ateş ölçer aletleri yardımı ile elde edilen gerçek veriler kullanılmıştır. Regresyon çözümlemesi yapılırken X değişkenine ait terimler koltuk altı ve Y değişkenine ait terimler ise kulaktan elde edilen veriler olarak kullanılmıştır. Verilerin analizi gerçekleştirilmeden önce ise normal dağılıma uygunluğu Kolmogorov-Smirnov testi ile incelenmiştir. Test sonucunda verilerin normal dağılıma uygunluk gösterdiği belirlenmiştir.

Bayes tekniği ile elde edilen Açığortay regresyon denklemlerinin HKO değerlerinin hesaplanması amacı ile sırasıyla örneklem hacimleri 100 ve 30 olarak seçilmiştir. 100 birimlik veri seti, 30 önsel bilgi verisi ve 70 birimlik mevcut veri, ilk 30 gözlem değeri ise 5-25, 10-20 ve 15-15 şeklinde önsel bilgi verisi ve mevcut veri olarak ayrılarak regresyon modelleri oluşturulmuştur. Çalışmada önsel dağılımın belirlenmesi aşamasında ise, mevcut veri setinin dağılımının normal olması ve matematiksel açıdan regresyon denkleminin oluşturulmasında uygunluk sağlaması sebebi ile önsel dağılım olarak “Doğal Eşlenik Önsel Dağılım” seçilmiştir. Önsel dağılıma ilişkin β parametresinin ortalama ve kovaryans değerleri Eşitlik (19) ile, σ parametresine ait ortalama ve varyans değerleri ise Eşitlik (20) ile hesaplanmıştır.

$$E[\beta | \sigma] = E[\beta] = \bar{\beta}, \text{cov}[\beta | \sigma] = A^{-1}\sigma^2, \quad (19)$$

$$A = (X'X)^{-1}X'$$

$$E[\sigma] = \frac{\Gamma[(\bar{v} - 1)/2]}{\Gamma(\bar{v}/2)} \left(\frac{\bar{v}}{2}\right)^{1/2} \bar{s}, \quad (20)$$

$$E[\sigma^2] = \frac{\bar{v}\bar{s}^2}{\bar{v} - 2}$$

Burada $\bar{\beta}$ ve \bar{s} değerleri, önsel verilere ait EKK tekniği ile elde edilen tahmin değerleri ve σ^2 , önsel verilerle oluşturulan regresyon denkleminin ait hata değerini ifade etmektedir.

Çalışmada önsel dağılım için parametre değerlerine ait ortalama ve varyans değerlerinin belirlenmesinin ardından sonsal dağılımın oluşturulabilmesi amacı ile

STATA paket programı kullanılmıştır. Kullanılan programda, önsel verilere ait bilgiler ve mevcut veri bilgisi birleştirilerek Bayes(Y/X) ve Bayes(X/Y) doğruları hesaplanmıştır.

Önsel verilere ait bilgiler için ilk olarak, bilgi içeren önsel dağılım ile lineer Bayes Regresyon modelinin oluşturulması kısmında belirtildiği üzere β parametresinin dağılımı normal dağılım, σ parametresinin dağılımı ise ters-gamma dağılımı seçilerek bu iki parametre değeri için de hesaplanan ortalama ve varyans değerleri dikkate alınmış, ikinci olarak da X ve Y değişkenleri için önsel verilerin ortalama ve varyans değerleri hesaplanarak çözümleme gerçekleştirilmiştir.

Son olarak elde edilen önsel bilgi kullanılarak mevcut verilere bağlı olarak Bayes Regresyon denklemi elde edilmiştir.

Bayes tekniği ile elde edilen Açığortay denklemini oluşturmak amacı ile Bayes(Y/X) (Y bağımlı, X bağımsız değişken) ve Bayes(X/Y) (X bağımlı, Y bağımsız değişken) doğruları için sabit katsayı ve eğim katsayısı değerleri Eşitlik (21)' de belirtildiği gibi hesaplanmıştır.

$$\hat{\beta}_1 = \frac{n \sum x_i y_i - \sum(x_i) \sum(y_i)}{n \sum x_i^2 - \sum(x_i)^2}, \quad (21)$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

Oluşturulan Bayes(Y/X) ve Bayes(X/Y) regresyon doğrularının eğim katsayıları sırası ile $\hat{\beta}_1$ ve $\hat{\beta}_2$ olarak alınarak Bayes-Açırtay regresyon doğrusunun eğim katsayısı ve sabit katsayı Eşitlik (1), (2)'de belirtildiği gibi hesaplanmış olup bu katsayılar bağlı Bayes-Açırtay regresyon denklemi Eşitlik (22)'de belirtildiği gibi hesaplanmıştır.

$$Y_{iBAYESAO} = \hat{\beta}_0 + \hat{\beta}_{AO} \bar{x}_i \quad (22)$$

Elde edilen Bayes(Y/X), Bayes(X/Y) ve Bayes-Açırtay regresyon denklemlerinin mevcut veri seti için sergiledikleri performansları ise Eşitlik (25)'de verilen HKO kriteri yardımıyla değerlendirilmiştir.

$$HKO = \frac{\sum(Y_i - \hat{Y}_i)^2}{n - k} \quad (25)$$

3. Bulgular

Farklı örneklem hacimlerinde, önsel bilgi verisi ve mevcut veri hacimlerinin farklı olarak seçilmesi ile oluşturulan regresyon doğruları için hesaplanan önsel dağılımlara ait parametre değerlerinin ortalama ve varyans değerleri Tablo 1. ve 3'de verilmiştir. Ayrıca Bayes(Y/X), Bayes(X/Y) ve Bayes Açığortay regresyon tekniklerine ilişkin sabit katsayı, eğim katsayısı ve HKO değerleri Tablo 1. Ve 3'de belirtilmiştir.

Tablo 1. n=100 iken Bayes(Y/X) ve Bayes(X/Y) regresyon doğruları için hesaplanan önsel dağılıma ait parametrelerin ortalama ve varyans değerleri.

		β		σ	
		Ortalama	Varyans	Ortalama	Varyans
30*	Bayes(Y/X)	0,968	0,0035	0,1172	0,014
	Bayes(X/Y)	0,938	0,0031	0,1126	0,013
40*	Bayes(Y/X)	1,004	0,0019	0,107	0,0116
	Bayes(X/Y)	0,929	0,0016	0,102	0,01005
50*	Bayes(Y/X)	0,981	0,0014	0,1016	0,01043
	Bayes(X/Y)	0,953	0,0014	0,1016	0,01043

* Önsel bilgi verisi için örneklem hacmi

Tablo 2. n=100 iken Bayes-Açıortay, Bayes(Y/X) ve Bayes(X/Y) regresyon teknikleri için hesaplanan sabit katsayı, eğim katsayısı ve HKO değerleri.

	30*-70**			40*-60**			50*-50**		
	β_0	β_1	HKO	β_0	β_1	HKO	β_0	β_1	HKO
Bayes _(Y X)	0,944	0,976	0,009651	0,993	0,969	0,008491	0,972	0,961	0,009083
Bayes _(X Y)	0,911	0,979	0,009503	0,917	0,979	0,009126	0,943	0,978	0,009829
Bayes _{AO}	0,761	0,975	0,008408	0,798	0,974	0,008404	0,798	0,974	0,008404

* Önsel bilgi verisi için örneklem hacmi

** Mevcut veri seti için örneklem hacmi

Tablo 1. ve Tablo 2. incelendiğinde, 100 birimlik veri seti 30 birimlik önsel bilgi ve 70 birimlik mevcut bilgi verisi olarak ayrıldığında, Bayes(Y/X) tekniği için önsel verilere ait β 'nin dağılımı Normal(0,968, 0,0035), σ 'nın dağılımı Ters-Gamma(0,1172, 0,014) ve X bağımsız değişkeninin dağılımı Normal(37,06,0,1251) iken sonsal dağılımın HKO değeri 0,009651, Bayes(X/Y) tekniği için önsel verilere ait β 'nin dağılımı Normal(0,938, 0,0031), σ 'nın dağılımı Ters-Gamma(0,1126, 0,013) ve Y bağımsız değişkeninin dağılımı Normal(36,84, 0,1291) iken sonsal dağılımın HKO değeri 0,009503 ve Bayes(Y/X) ve Bayes(X/Y) teknikleri ile oluşturulan regresyon doğrularının açıortayı alınarak elde edilen Bayes-Açıortay tekniğinin HKO değeri 0,008408 olarak hesaplanmıştır.

100 birimlik veri seti 40 birimlik önsel bilgi verisi ve 60 birimlik mevcut veri olarak ayrıldığında, Bayes(Y/X) tekniği ile hesaplanan önsel verilere ait β 'nin dağılımı Normal(1,004, 0,0019), σ 'nın dağılımı Ters-Gamma(0,107, 0,0116) ve X bağımsız değişkeninin dağılımı Normal(37,15, 0,1415) iken sonsal dağılıma ait HKO değeri 0,008491, Bayes(X/Y) tekniği için önsel verilere ait β 'nin dağılımı Normal(0,929, 0,0016), σ 'nın dağılımı Ters-Gamma(0,102, 0,01005) ve Y bağımsız değişkeninin dağılımı Normal(36,94, 0,1530) iken sonsal dağılımın HKO değeri 0,009126 olarak hesaplanmıştır. Ayrıca Bayes(Y/X) ve Bayes(X/Y) tekniği ile oluşturulan regresyon doğrularının açıortayı alınarak oluşturulan Bayes-Açıortay tekniğinin HKO değeri 0,008404 olarak hesaplanmıştır.

100 birimlik veri seti 50 birimlik önsel bilgi verisi ve 50 birimlik mevcut veri olarak ayrıldığında ise Bayes(Y/X) tekniği için önsel verilere ait β 'nin dağılımı Normal(0,981, 0,0014), σ 'nın dağılımı Ters-Gamma(0,1016, 0,01043) ve X bağımsız değişkeninin

dağılımı Normal(37,11, 0,1457) iken sonsal dağılımın HKO değeri 0,009083, Bayes(X/Y) tekniği için önsel verilere ait β 'nin dağılımı Normal(0,953, 0,0014), σ 'nın dağılımı Ters-Gamma(0,1016, 0,01043) ve Y bağımsız değişkeninin dağılımı Normal(36,92, 0,1457) iken sonsal dağılımın HKO değeri 0,009829 olarak hesaplanmıştır. Buna ek olarak Bayes(Y/X) ve Bayes(X/Y) tekniği ile elde edilen regresyon doğrularının açıortayı alınarak oluşturulan Bayes-Açıortay tekniğinin HKO değeri 0,008404 olarak hesaplanmıştır.

Tablo 2.'ye göre 100 birimlik veri seti 40 birimlik önsel bilgi verisi ve 60 birimlik mevcut veri olarak ayrıldığında, elde edilen Bayes(Y/X) ve Bayes(X/Y) tekniklerinin daha düşük HKO değerlerine sahip olduğu gözlemlenmiştir. Ayrıca 100 birimlik veri seti 40*-60** ve 50*-50** olarak ayrıldığında Bayes-Açıortay tekniğinin HKO değerlerinin eşit ve 30*-70** alınarak oluşturulan Bayes-Açıortay tekniğine göre daha az hataya sahip oldukları görülmektedir. Genel olarak HKO değerleri incelendiğinde ise Bayes-Açıortay tekniğinin HKO değerlerinin Bayes(Y/X) ve Bayes(X/Y) teknikleri için hesaplanan HKO değerlerine göre daha düşük olduğu tespit edilmiştir.

Tablo 3. ve Tablo 4.'e göre, 30 birimlik veri seti 5 birimlik önsel bilgi verisi ve 25 birimlik mevcut veri olarak ayrıldığında önsel bilgi verisine ait β parametresinin dağılımı Normal(0,696, 0,0087), σ parametresinin dağılımı Ters-Gamma(0,3069, 0,015) ve X bağımsız değişkeninin dağılımı Normal(36,94, 0,1144) iken Bayes(Y/X) tekniğinin HKO değeri 0,013328 olarak hesaplanmıştır. Önsel bilgi verisine ait β parametresinin dağılımı Normal(1,363, 0,0342), σ parametresinin dağılımı Ters-Gamma(0,4341, 0,03) ve Y bağımsız değişkeninin dağılımı Normal(36,66, 0,0584) iken ise Bayes(X/Y) tekniğinin HKO değeri 0,012726 olarak hesaplanmıştır. Ayrıca Bayes(Y/X) ve

Tablo 3. n=30 iken Bayes(Y/X) ve Bayes(X/Y) regresyon doğruları için hesaplanan önsel dağılıma ait parametrelerin ortalama ve varyans değerleri.

		β		σ	
		Ortalama	Varyans	Ortalama	Varyans
5*	Bayes(Y/X)	0,696	0,0087	0,3069	0,015
	Bayes(X/Y)	1,363	0,0342	0,4341	0,03
10*	Bayes(Y/X)	0,865	0,0126	0,1162	0,0147
	Bayes(X/Y)	1,019	0,0175	0,1263	0,0173
15*	Bayes(Y/X)	0,885	0,0031	0,2985	0,00945
	Bayes(X/Y)	1,072	0,0046	0,3338	0,0118

* Önsel bilgi verisi için örneklem hacmi

Tablo 4. n=30 iken Bayes-Açıortay, Bayes(Y/X) ve Bayes(X/Y) regresyon teknikleri için hesaplanan sabit katsayı, eğim katsayısı ve HKO değerleri.

	5*-25**			10*-20**			15*-15**		
	β_0	β_1	HKO	β_0	β_1	HKO	β_0	β_1	HKO
Bayes _(Y X)	0,621	0,978	0,013328	0,725	0,975	0,013033	0,864	0,973	0,019648
Bayes _(X Y)	0,813	0,984	0,012726	0,796	0,984	0,012786	1,042	0,976	0,0163
Bayes _{AO}	0,488	0,981	0,012802	0,543	0,979	0,012797	0,728	0,975	0,012785

* Önsel bilgi verisi için örneklem hacmi

** Mevcut veri seti için örneklem hacmi

Bayes(X/Y) teknikleri ile elde edilen regresyon doğrularının açıortayı alınarak oluşturulan Bayes-Açıortay tekniğinin HKO değeri 0,012802 olarak hesaplanmıştır.

30 birimlik veri setinin 10 birimlik önsel bilgi verisi ve 20 birimlik mevcut veri olarak ayrıldığı durumda, önsel bilgi verisine ait β parametresinin dağılımı Normal(0,865, 0,0126), σ parametresinin dağılımı Ters-Gamma(0,1162, 0,0147) ve X bağımsız değişkeninin dağılımı Normal(37,02, 0,0876) iken Bayes(Y/X) tekniğine ait HKO değeri 0,013033 olarak hesaplanmıştır. Önsel bilgi verisine ait β parametresinin dağılımı Normal(1,019, 0,0175), σ parametresinin dağılımı Ters-Gamma(0,1263, 0,0173) ve Y bağımsız değişkeninin dağılımı Normal(36,76, 0,0744) iken Bayes(X/Y) tekniğinin HKO değeri 0,012786 olarak hesaplanmıştır. Buna ek olarak Bayes(Y/X) ve Bayes(X/Y) tekniği ile oluşturulan regresyon doğrularının açıortayı alınarak oluşturulan Bayes-Açıortay tekniğinin HKO değerinin 0,012797 olduğu görülmektedir.

30 birimlik veri seti, 15 birimlik önsel bilgi verisi ve 15 birimlik mevcut veri olarak ayrıldığında ise, önsel bilgi verisine ait β parametresinin dağılımı Normal(0,885, 0,0031), σ parametresinin dağılımı Ters-Gamma(0,2985, 0,00945) ve X bağımsız değişkeninin dağılımı Normal(37,03, 0,1742) iken Bayes(Y/X) tekniğinin HKO değeri 0,019648 olarak hesaplanmıştır. Önsel bilgi verisine ait β parametresinin dağılımı Normal(1,072, 0,0046), σ parametresinin dağılımı Ters-Gamma(0,3338, 0,0118) ve Y bağımsız değişkeninin dağılımı Normal(36,75, 0,1438) iken ise Bayes(X/Y) tekniğinin HKO değeri 0,0163 olarak hesaplanmıştır. Ayrıca Bayes-Açıortay regresyon tekniğinin HKO değerinin 0,012785 olduğu görülmektedir.

Tablo 4.'e göre, Bayes(Y/X) tekniği için hesaplanan HKO değerinin, 30 birimlik veri seti 10*-20** şeklinde

ayrıldığında daha düşük olduğu görülmektedir. Bayes(X/Y) tekniğinin HKO değeri ise 75 birimlik veri seti 5*-25** şeklinde ayrıldığında daha düşük olduğu belirlenmiştir. Bayes-Açıortay tekniğinin HKO değerinin önsel bilgi verisinin örneklem hacmi arttıkça azaldığı ve 30 birimlik veri seti 5*-25** şeklinde ayrıldığında en düşük HKO değerine sahip olduğu görülmektedir. Ayrıca 30 birimlik veri setinin 5*-15** ve 10*-20* şeklinde ayrıldığı durumlarda Bayes(X/Y) tekniği için hesaplanan HKO değerinin Bayes(Y/X) ve Bayes-Açıortay tekniklerinin HKO değerine göre daha düşük olduğu görülmektedir. Buna ek olarak, 30 birimlik veri setinin 15*-15** şeklinde ayrılması halinde Bayes-Açıortay tekniği için hesaplanan HKO değerinin Bayes(Y/X) ve Bayes(X/Y) tekniklerinin HKO değerlerine göre daha düşük olduğu belirlenmiştir.

4. Tartışma ve Sonuç

Bu çalışmada, Bayesci yaklaşım açıklanarak, bu yaklaşım aracılığı ile sadece bağımlı değişkenler değil, aynı zamanda bağımsız değişkenlerinin hatalarını da analize dahil eden Tip II regresyon tekniklerinden biri olan Açıortay tekniği kullanılarak parametre tahminleri yapılmış ve Bayes-Açıortay tekniği ortaya konmuştur. Yapılan analizlerde farklı örneklem hacimleri, bu örneklem hacimlerinde farklı önsel bilgi veri hacmi ve farklı mevcut veri hacimleri alınarak Bayes-Açıortay tekniğinin performansına ilişkin değerlendirmeler yapılmıştır. Çalışmada elde edilen bulgular gerçek veri setinden elde edilmiş olup bir simülasyon yapılmamıştır.

Bayes-Açıortay tekniğinin performansının değerlendirilmesi amacı ile yapılan analizler sonucunda regresyon doğrularının HKO değerleri hesaplanmıştır. Elde edilen sonuçlar doğrultusunda Bayes-Açıortay tekniği için elde edilen HKO değerlerinin, 30 birimlik veri seti için 5*-25** ve 10*-20** olarak ayrılan örneklerle kurulan Bayes(X/Y)

modeli hariç, Bayes(Y/X) ve Bayes(X/Y) regresyon doğruları için hesaplanan HKO değerlerine göre daha düşük olduğu ve performansının daha iyi olduğu gözlemlenmiştir. Daha önce yapılan araştırmalarda Bayes tekniğinin küçük örneklem hacimlerinde daha iyi sonuçlar verdiği tespit edilmiş olup yapılan bu çalışmada büyük örneklem hacminde (n=100) tüm veri setleri için Bayes-Açığortay tekniğinin daha düşük HKO değerine sahip olduğu gözlemlenmiştir.

Çalışmada önsel bilgi ve mevcut verilerin örneklem hacimlerinin artırılıp azaltılması sonucunda her bir veri seti için hesaplanan regresyon doğrularının HKO değerlerine ilişkin genel olarak Bayes-Açığortay tekniğinin klasik tekniğe göre yakın veya daha iyi sonuçlar verdiği söylenebilmektedir. Bazı veri setlerinde önsel bilgi veri hacminin küçük olduğu ve mevcut veri hacminin daha büyük olduğu durumda HKO değerinin düşük olduğu, bazılarında ise önsel bilgi veri hacminin büyük, mevcut veri hacminin küçük olduğu durumlarda HKO değerinin düşük olduğu gözlemlenmiştir.

Genel olarak herhangi bir regresyon modeli oluşturmada gerekli varsayımların sağlanması halinde Tip I Regresyon tekniği kullanılmakta ve oluşturulan regresyon modeline ilişkin parametre tahmin etmede klasik yöntemler kullanılmaktadır. Fakat gerçekte elde edilen bağımsız değişkenlerden oluşabilecek hataların da analize dahil edilebilmesini sağlayan Tip II Regresyon tekniklerinin kullanılması yapılan analizlerde avantaj sağlamaktadır. Tip II regresyon tekniklerinden biri olan Açığortay tekniğinde parametre tahmin yöntemi olarak EKK tekniği kullanılmaktadır. Bu çalışmada ise, EKK tekniği yerine, Bayes yaklaşımı aracılığıyla parametre tahmini yapılarak ilk defa Bayes-Açığortay tekniği ortaya konmuştur. Ayrıca ortaya konulan bu tekniğin literatüre önemli bir katkı sağlayacağı öngörülmektedir.

Etik Beyanı

Bu çalışma, Ece Özgören'in Sinan Saraçlı danışmanlığında Afyon Kocatepe Üniversitesi Fen Bilimleri Enstitüsünde tamamlanan yüksek lisans tezinden üretilmiş olmakla birlikte çalışmada "Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Etiği Yönergesi" kapsamında uyulması gerekli tüm kurallara uyulduğunu, bahsi geçen yönergenin "Bilimsel Araştırma ve Yayın Etiğine Aykırı Eylemler" başlığı altında belirtilen eylemlerden hiçbirinin gerçekleştirilmediğini taahhüt ederiz.

Kaynakça

- [1] Stigler, S. M. 1983. Who Discovered Bayes's Theorem. *The American Statistician*, 37(4), 290-296.
- [2] Bolstad, W. M. 2007. *Introduction to Bayesian Statistics*. 2nd edition. Wiley – Interscience Hamilton, New Zealand, 463s.
- [3] Avcı, E. 2012. *Bayesci Sağkalım Analizi ve Meme Kanseri Verileri Üzerine Bir Uygulama*. Mimar Sinan Güzel Sanatlar Üniversitesi, Fen Bilimleri Enstitüsü, Doktora 82s, Tezi, İstanbul.
- [4] Altındağ, İ. 2015. *Bayesci Doğrusal Olmayan Yapısal Eşitlik Modeli*. Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi, 150s, Konya.
- [5] Saraçlı, S. 2008. *Ölçüm Hatalı Modellerde Doğrusal Regresyon Tekniklerinin Karşılaştırılması-Monte Carlo Simülasyon Çalışması*. Eskişehir Osmangazi Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi, 90s, Eskişehir.
- [6] Özgören, E. 2019. *Bayes Açığortay Regresyon Tekniği ve Bir Uygulama*. Afyon Kocatepe Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, 87s, Afyonkarahisar.
- [7] Çiftçi, F. 2009. *Regresyonda Alternatif Metotlar*. Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi, 103s, Konya.
- [8] Çevik, M. 2009. *Doğrusal Olmayan Bayesci Regresyon ve Yüksek Frekanslı Ses Sistemlerinde Bir Uygulama*. Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, 94s, İstanbul.
- [9] Isobe, T., Feigelson, E.D., Akritas, M.G., Babu, G.J. 1990. *Linear Regression in Astronomy I. The Astrophysical Journal*, 364, 104-113.
- [10] Judge, G.G, Griffiths, W.E., Hill, R.C., Lütkepohl, H., Lee, T. 1985. . 2nd edition. *The Theory and Practice Of Econometrics*. John Wiley&Sons. Canada, 999s.
- [11] Genç, A., Karadavut, U., Palta, Ç. 2010. *Lineer Olmayan Bayesci Regresyon ve Tarım Alanında Bir Uygulama*. TÜBAV Bilim Dergisi, 3(3), 250-258.
- [12] Tiao B., Zellner, A. 1964. *Bayes's theorem and the use of prior knowledge in regression analysis*. *Biometrika*, 51(1/2), 219-230.
- [13] Gasım, N. 2013. *Bayesci Model İle Doğrusal Regresyon Modellerinin Karşılaştırılması Üzerine Bir Uygulama*, Dokuz Eylül Üniversitesi, Sosyal Bilimler Enstitüsü, Yüksek Lisans Tezi, 91s, İzmir.
- [14] Saraçlı, S., Çelik, E. 2012. *Metot Karşılaştırma Çalışmalarında Bland-Altman ve Tip II Regresyon Analizinin Karşılaştırılması*. Düzce Üniversitesi Sağlık Bilimleri Enstitüsü Dergisi, 2(1), 11-14.