

## Bayesci Yaklaşım ile Regresyon Modellerinde Parametre Tahmini

Samet KAYA<sup>1</sup>, Esin KÖKSAL BABACAN<sup>2\*</sup>

### Öz

Regresyon analizi, aralarında sebep- sonuç ilişkisi bulunan iki veya daha fazla değişken arasındaki ilişkiyi modellemek ve incelemek için kullanılan istatistiksel bir yöntemdir. Klasik istatistikte, bilinmeyen parametreler sabit birer değer olarak alınırken Bayesci istatistikte birer rasgele değişken olarak göz önüne alınır ve bunların da kendilerine ait dağılımları olduğu varsayımı kullanılır. Önsel dağılım olarak adlandırılan bu dağılım bilgisi ve örneklem bilgisi kullanılarak parametrelere ilişkin sonsal dağılım elde edilir. Parametre ile ilgili her türlü sonuç çıkarımı bu sonsal dağılım kullanılarak yapılır. Bu çalışmada ilk olarak alınan bir model için simülasyonla veriler üretilmiş ve üretilen bu veriler kullanılarak model parametreleri klasik regresyon ve Bayesci regresyon kullanılarak tahmin edilmiş ve her iki yöntemin sonuçları karşılaştırılmıştır. Daha sonra literatürde verilen teslim süresi verileri için benzer karşılaştırma işlemi yapılmıştır. Yapılan karşılaştırma sonucunda her iki yöntemle elde edilen tahminlerin benzer olduğu sonucuna ulaşılmıştır. Simülasyonlar için Matlab programı ve Bayesci regresyon sonuçları için Winbugs programı kullanılmıştır.

**Anahtar Kelimeler:** MCMC (Markov Chain Monte Carlo), Bayesci çıkarım, Regresyon analizi, Winbugs

## Bayesian Approach to Parameter Estimation in Regression Models

### Abstract

Regression analysis is a statistical method used to model and analyze the relationship between variables. The main objective of regression analysis is to estimate unknown parameters in the regression model. In classical statistics, the least-squares estimation method is used to estimate these unknown parameters. On the other hand, in Bayesian statistics, parameters are considered as random variables and they have their distribution. By using this distribution information and sample information, the posterior distribution of the parameters is obtained. Any results related to the parameter are made by using this posterior distribution. This study aims to compare the estimation results which are obtained by classical regression and Bayesian regression. For this purpose, first, the data are obtained for the considering model by the simulation and the results of both methods are compared. Then, a similar comparison is made for the real dataset given in the literature. The Matlab program is used for the simulations and the Winbugs program is used for Bayesian regression results.

**Keywords:** MCMC (Markov Chain Monte Carlo), Bayesian Approach, Regression Analysis, Winbugs

<sup>1</sup>Ankara Üniversitesi, İstatistik Bölümü, Ankara, Türkiye, sametky\_1301@hotmail.com

<sup>2</sup>Ankara Üniversitesi, İstatistik Bölümü, Ankara, Türkiye, ekoksal@science.ankara.edu.tr

<sup>1</sup><https://orcid.org/0000-0002-6937-8138>

<sup>2</sup><https://orcid.org/0000-0002-9649-5276>

## 1. Giriş

Bayes Teoremi ilk olarak 1793 yılında Bayes tarafından ortaya atılmıştır. Klasik istatistikten farklı olarak Bayesci istatistikte parametreler birer rasgele değişken olarak göz önüne alınır. Bu nedenle, parametrelerin de kendilerine ait dağılımları vardır. Bayes yaklaşımında, ilk aşamada modeldeki parametreler için alınan bu dağılımlar önsel dağılım olarak bilinir (Gelman and et al., 2004). Daha sonra örneklem bilgisi ve önsel dağılım bilgisi kullanılarak parametrelere ilişkin bilgiler güncellenir ve sonsal dağılımlar elde edilir. Parametrelere ilişkin tüm sonuç çıkarımları bu sonsal dağılımlar kullanılarak yapılır.

Regresyon analizine Bayesci yaklaşım önsel bilgi kullanımına izin verir. Bayesci regresyon ile örneklem sayısının az olduğu durumlarda bile iyi sonuçlar elde edilir (Wundervald, 2019). Bayesci sonuç çıkarımında, parametreler için bilgi vermeyen önseller kullanıldığında elde edilen sonuçlar klasik yöntemlerle elde edilen sonuçlara benzer olmaktadır (Gamerman and Lopes, 2006). Tek biçimli (uniform) önsel dağılımlar, düz (flat) önsel dağılımlar ve Jeffreys'in önsel dağılımları bu tür önsel dağılımlara örnek olarak verilebilir (Gamerman and Lopes, 2006). Bilgi veren önseller kullanıldığında ise birçok durumda sonsal dağılımın açık formunun elde edilmesinde zorluklarla karşılaşmaktadır. Bu nedenle bilgi veren önsellerin kullanıldığı durumda Bayesci regresyonda sonuç çıkarımı yapabilmek için geliştirilen simülasyon yöntemleri kullanılır. Bu simülasyon yöntemleri ile analitik olarak elde edilemeyen sonsal dağılımlardan örneklem çekilir ve bu örneklem kullanılarak sonuç çıkarımına gidilir. Bu simülasyon yöntemleri Markov Chain Monte Carlo (MCMC) yöntemleri olarak bilinir (Gamerman and Lopes, 2006).

Bu çalışmanın amacı istatistiğin önemli bir konusu olan regresyon analizinde Bayesci sonuç çıkarımının nasıl yapıldığını araştırmak ve uygulama çalışması ile parametre tahmin sonuçlarını klasik yöntemlerle elde edilen sonuçlarla karşılaştırmaktır. Bu amaçla ikinci bölümde doğrusal regresyon analizi için basit ve çoklu doğrusal regresyon modelleri hatırlatılmıştır. Üçüncü bölümde Bayesci sonuç çıkarımı ana hatları ile verilmiştir. Dördüncü bölümde regresyon modellerinde Bayesci yaklaşım ile sonuç çıkarımının nasıl yapılacağı anlatılmıştır. Bilgilendirici ve bilgilendirici olmayan önsellerin kullanımı ile sonsal dağılımların nasıl elde edileceği verilmiştir. Beşinci bölümde Bayesci sonuç çıkarımında yaygın kullanıma sahip olan sonsal dağılımlardan örneklem çekme yöntemlerinden Gibbs algoritması verilmiştir. Altıncı bölüm uygulama çalışmalarına ayrılmıştır. İlk olarak alınan bir regresyon modeline ilişkin farklı parametre değerleri için simülasyonla veriler üretilip parametre tahminleri Bayesci yöntemle ve klasik en küçük kareler yöntemiyle yapılmış ve elde edilen sonuçlar karşılaştırılmıştır. Daha sonra literatürde bulunan gerçek veriler kullanılarak sonuç çıkarımları yapılmıştır. Son bölümde çalışmadan elde edilen sonuçlar tartışılmıştır.

## 2. Doğrusal Regresyon Analizi

### 2.1. Basit doğrusal regresyon modeli

$x$  bağımsız değişken ve  $y$  bağımlı değişken olmak üzere basit doğrusal regresyon modeli,

$$y = \beta_0 + \beta_1 x + \varepsilon \quad (1)$$

biçiminde ifade edilir. Bu denklemde,  $\beta_0$  kesim noktası,  $\beta_1$  eğim ve  $\varepsilon$  rasgele hata terimidir. Regresyon analizinin önemli bir amacı modelde bulunan bilinmeyen parametreleri tahmin etmektir. Buna göre uydurulan model,

$$\hat{y} = b_0 + b_1 x \quad (2)$$

olur. Denklemdeki,  $\hat{y}$ ,  $x$  değişkenine karşılık gelen  $y$  değişkeninin uydurulmuş ya da kestirilmiş değeridir. Regresyon modellerinde parametre tahmini yapabilmek için modelde yer alan hata teriminin sağlaması gereken bir takım varsayımlar vardır. Bu varsayımlar,

i.  $E(\varepsilon) = 0$

ii.  $Var(\varepsilon) = \sigma^2$

iii.  $Cov(\varepsilon_i, \varepsilon_j) = 0, i \neq j$  için

dir. Klasik yaklaşımda,  $\beta_0, \beta_1$  bilinmeyen parametreleri en küçük kareler yöntemi kullanılarak kestirilir. Bu yöntemde,  $\beta_0$  ve  $\beta_1$  bilinmeyen regresyon parametreleri,  $y_i$  gözlemleri ile uydurulacak regresyon doğrusu arasındaki farkların karelerinin toplamı minimum olacak şekilde kestirilir. Buna göre,

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \quad (3)$$

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{\sum_{i=1}^n y_i(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (4)$$

olarak bulunur (Erar, 2013).

## 2.2.Çoklu doğrusal regresyon modeli

Çoklu doğrusal regresyon modelinde,  $y$  bağımlı değişken ve onunla doğrusal ilişkiye sahip  $k$  tane  $x$  bağımsız değişkeni göz önüne alınır. Buna göre model,

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon = \beta_0 + \sum_{j=1}^k \beta_j x_j + \varepsilon \quad (5)$$

biçiminde ifade edilir. Burada  $\varepsilon$ , rasgele hata bileşeninin yukarıda verilen varsayımları sağlandığı kabul edilir. Çoklu doğrusal regresyon modelleri matris formunda yazıldığında sonuçlar daha kısa bir şekilde ifade edilir. Buna göre,  $\beta = (\beta_0, \beta_1, \dots, \beta_k)'$  ile parametre vektörü gösterilmek üzere model,

$$y = X \beta + \varepsilon$$

biçiminde olup, bu matrisler

$$\begin{bmatrix} n & \sum_{i=1}^n x_{i1} & \sum_{i=1}^n x_{i2} & \dots & \sum_{i=1}^n x_{ik} \\ \sum_{i=1}^n x_{i1} & \sum_{i=1}^n x_{i1}^2 & \sum_{i=1}^n x_{i1}x_{i2} & \dots & \sum_{i=1}^n x_{i1}x_{ik} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n x_{ik} & \sum_{i=1}^n x_{ik}x_{i1} & \sum_{i=1}^n x_{ik}x_{i2} & \dots & \sum_{i=1}^n x_{ik}^2 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_{i1}y_i \\ \vdots \\ \sum_{i=1}^n x_{ik}y_i \end{bmatrix}$$

gibidir. Buna göre  $\beta$  için en küçük kareler tahmin edicisi,

$$\hat{\beta} = (X'X)^{-1}X'y \quad (6)$$

biçiminde elde edilir (Erar, 2013).

## 3. Bayesci Sonuç Çıkarımı

Bayesci yaklaşım, 1763'de Thomas Bayes tarafından ortaya konulan ve matematiksel istatistiğin önemli bir teoremi olan Bayes teoremine dayanır (Gelman and et al., 2004). Bayesci yaklaşımda,  $\theta$  parametresi kendine ait dağılımı olan bir rasgele değişken olarak düşünülür. Bu amaçla, ilk olarak  $\theta$  parametresi için  $\Pi(\theta)$  ile gösterilen önsel (prior) dağılım belirlenir.

Daha sonra, göz önüne alınan model ile ilgili gözlenen  $y$  verileri kullanılarak  $L(\theta|y)$  ile ifade edilen olabilirlik fonksiyonu oluşturulur. Alınan önsel dağılım bilgisi ve veri bilgisi güncellenerek  $\theta$

parametresine ilişkin sonsal dağılım elde edilir. Buna göre artık,  $\theta$  hakkında tüm istatistiksel çıkarımlar elde edilen bu sonsal dağılım kullanılarak yapılır. Bayesci sonuç çıkarımı için bahsedilen adımlar Bayes teoremi kullanılarak

$$p(\theta|y) = \frac{p(\theta,y)}{p(y)} = \frac{L(\theta|y)\Pi(\theta)}{p(y)} = \frac{L(\theta|y)\Pi(\theta)}{\int L(\theta|y)\Pi(\theta)d\theta} \quad (7)$$

biçiminde ifade edilir. Burada,  $\int L(\theta|y)\Pi(\theta)d\theta$  ifadesi sonsal dağılım için normalleştirme sabitidir ve integral sonlu olduğu sürece değeri sonsal dağılım hakkında herhangi bir ilave bilgi içermez. Bu nedenle sonsal dağılım önsel dağılım ve olabilirliğin çarpımı

$$p(\theta|y) \propto L(\theta|y)\Pi(\theta) \quad (8)$$

biçiminde orantısal olarak yazılabilir (Gamerman and Lopes, 2006). Yani, Bayesci yaklaşım kısaca var olan bilginin yeni bilgi ile güncellenmesi olarak ifade edilebilir.

Bayesci yaklaşımın klasik yöntemlere göre pek çok avantajı vardır. Daha önceki yıllarda Bayesci yaklaşımla sonuç çıkarımı yapabilmek için kullanılan uygulaması zor olan nümerik integrasyon yöntemleri bilgisayar teknolojileri geliştikçe yerini simülasyon tabanlı örnekleme yöntemlerine bırakmış ve bu sayede, Bayesci yaklaşım biyoistatistik, ekonometri, genetik haritalama gibi pek çok alanda uygulanmaya başlanmıştır (Albert ve Chib, 1993; Heath and et al., 1997). Bayesci yöntemler, model kestiriminde en kapsamlı ve sağlam yaklaşımı sağlarlar (Congdon, 2004). Bayesci yaklaşım ile yapılan tahminlerde en çok olabilirlikteki gibi asimptotik Normallik varsayımına ihtiyaç duyulmaz. Çünkü, Bayesci yaklaşımda örnekleme tabanlı yöntemler ile parametreler için tam bir dağılıma ulaşılır ve Normallik varsayımına ihtiyaç duyulmaz. Tüm çıkarımlar elde edilen sonsal dağılım kullanılarak yapılır. Ayrıca sonsal dağılım hesaplanırken önsel bilgiler de kullandığı için Bayesci yaklaşım klasik yöntemlere göre daha doğru tahminler verir.

#### 4. Bayesci Regresyon

Bayesci sonuç çıkarımında parametreler birer rasgele değişken olarak kabul edildiğinden, regresyona ait  $(\beta_0, \beta_1, \dots, \beta_k)$  ve  $\sigma^2$  parametreleri de birer rasgele değişkendir ve kendilerine ait olasılık dağılımları vardır.

#### 4.1. Basit doğrusal regresyon modeline Bayesci yaklaşım

Bayesci yaklaşımın basit doğrusal regresyon modellerine uygulanması,

*Sonsal dağılım*  $\propto$  *önsel dağılım*  $\times$  *olabilirlik fonksiyonu*

$$P(\beta_0, \beta_1, \sigma^2 | y, X) = P(\beta_0, \beta_1, \sigma^2) \times P(y, X | \beta_0, \beta_1, \sigma^2) \quad (9)$$

biçimindedir (Ekici, 2005). Basit doğrusal regresyon modeli,

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (10)$$

olmak üzere, Bayesci sonuç çıkarımı için kullanılacak varsayımlar klasik doğrusal regresyon modeli için kullanılan varsayımlar ile aynıdır.  $\varepsilon_i \sim N(0, \sigma^2)$  dağılımına sahip rasgele hatalardır. Regresyon analizinde,  $x$  sabit ya da rasgele değişken olabilir. Eğer,  $X$  rasgele değişken ise hata terimi  $\varepsilon_i$ 'lerden bağımsız olduğu varsayımı yapılmalıdır. Yani bu gibi durumlarda, Bayesci sonuç çıkarımı için bir varsayıma daha ihtiyaç vardır. Bu varsayım, önsel dağılım parametrelerinin  $X$ 'lerden ve birbirlerinden bağımsız olmasıdır. Klasik regresyonda,  $X$ 'in veri olduğu durumda  $y$ 'nin koşullu olasılığı hakkında  $X$ 'in dağılımının bilgi sağlamadığı varsayılır. Bayesci sonuç çıkarımında ise model  $X$ 'in önsel dağılımı olan  $g(X|\varphi)$ 'de içermelidir. Bu şekildeki bir model için ortak önsel dağılım,

$$P(y, X | \beta_0, \beta_1, \sigma^2, \varphi) = P(y | X, \beta_0, \beta_1, \sigma^2) \times g(X | \varphi) \times P(\beta_0, \beta_1, \sigma^2, \varphi) \quad (11)$$

$$P(\beta_0, \beta_1, \sigma^2, \varphi) = P(\varphi) \times P(\beta_0, \beta_1, \sigma^2) \quad (12)$$

olarak ifade edilebilir.  $P(y | X, \beta_0, \beta_1, \sigma^2)$ 'yi belirleyen  $\beta_0, \beta_1, \sigma^2$  parametreleri ve  $g(X | \varphi)$ 'yi belirleyen  $\varphi$  parametreleri birbirinden bağımsızdır. Bu durumda,

$$P(\beta_0, \beta_1, \sigma^2, \varphi | y, X) = P(\beta_0, \beta_1, \sigma^2 | y, X) \times P(\varphi | X)$$

$$P(\beta_0, \beta_1, \sigma^2, \varphi | y, X) = [P(\beta_0, \beta_1, \sigma^2) \times P(y | X, \beta_0, \beta_1, \sigma^2)] \times [P(X | \varphi) \times P(\varphi)]$$

biçiminde yazılabilir. Önsel dağılımların bağımsız olması varsayımından faydalanarak, ortak sonsal dağılım çarpanlarına ayrılarak;

$$P(\beta_0, \beta_1, \sigma^2, \varphi|y, X) \propto P(\beta_0, \beta_1, \sigma^2) \times P(y|X, \beta_0, \beta_1, \sigma^2) \quad (13)$$

biçiminde yazılabilir. Sonsal dağılım, denklem (13)'deki gibi oransallık kullanılarak ifade edildiğinde, bilgi kaybı olmadan klasik regresyonda olduğu gibi  $X$ 'in dağılımına olan bağımlılık ortadan kaldırılmış olur (Ekici, 2005).  $X$ ,  $\beta_0$ ,  $\beta_1$  ve  $\sigma^2$  veri iken,  $y$ 'nin ortalaması,

$$E(y_i|x_i, \beta_0, \beta_1, \sigma^2) = \beta_0 + \beta_1 x_i, i = 1, 2, \dots, n$$

ve varyansı,

$$Var(y_i|x_i, \beta_0, \beta_1, \sigma^2) = \sigma^2, i = 1, 2, \dots, n$$

olup bağımsız Normal dağılır. Bu varsayımlara göre olabilirlik fonksiyonu,

$$P(y|X, \beta_0, \beta_1, \sigma^2) \propto \frac{1}{\sigma^n} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2\right]$$

olarak ifade edilir.

#### 4.1.1. Bilgi vermeyen önsel dağılım ile sonuç çıkarımı

Varyansın bilinmediği durumda  $\beta_0$ ,  $\beta_1$  ve  $\sigma$  için en yaygın kullanılan bilgi vermeyen önsel dağılım;

$$P(\beta_0, \beta_1, \sigma) = P(\beta_0, \beta_1) \times P(\sigma) \propto \frac{1}{\sigma}, -\infty < \beta_0, \beta_1 < \infty, 0 < \sigma < \infty \quad (14)$$

biçimindedir. Buradan sonsal dağılım,

$$P(\beta_0, \beta_1, \sigma|y, X) \propto \frac{1}{\sigma^{n+1}} \exp\left\{-\frac{1}{2\sigma^2} [v s^2 + n \times (\beta_0 - \hat{\beta}_0)^2 + (\beta_1 + \hat{\beta}_1)^2 \sum_{i=1}^n x_i^2 + 2(\beta_0 - \hat{\beta}_0) \times (\beta_1 + \hat{\beta}_1) \sum_{i=1}^n x_i]\right\} \quad (15)$$

$$v = n - 2, \quad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}, \quad \hat{\beta}_1 = \frac{\sum_{i=1}^n y_i (x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad s^2 = \frac{\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2}{v}$$

biçiminde yazılır (Zellner, 1971). Denklem (15) ile sonsal dağılımın,  $\hat{\beta}_0$  ve  $\hat{\beta}_1$  ortalamalı  $\sigma^2 (X'X)^{-1}$  varyans- kovaryans matrisli çok değişkenli normal dağılım olduğu görülür. Buradan  $\beta_0, \beta_1$  için marjinal sonsal dağılım ortak dağılımın  $\sigma$ 'ya göre integrali alınarak,

$$P(\beta_0, \beta_1 | y, X) = \int_0^\infty P(\beta_0, \beta_1, \sigma | y, X) d\sigma \propto [vs^2 + n \times (\beta_0 - \hat{\beta}_0)^2 + (\beta_1 + \hat{\beta}_1)^2 \sum_{i=1}^n x_i^2 + 2(\beta_0 - \hat{\beta}_0) \times (\beta_1 + \hat{\beta}_1) \sum_{i=1}^n x_i]^{-\frac{n}{2}} \quad (16)$$

biçiminde elde edilir. Bu ifade, iki değişkenli *Student-t* dağılımı ile aynı yapıdadır. Başka bir deyişle  $\beta_0$  ve  $\beta_1$  parametrelerinin marjinal sonsal dağılımı, *Student-t* dağılımına uyum sağlamaktadır.

$\beta_0$  ve  $\beta_1$  parametrelerine ilişkin sonuç çıkarımı yapılırken, parametrelerin her birinin marjinal sonsal dağılımı *Student-t* olasılık yoğunluk fonksiyonunun özelliklerine dayanarak,

$$P(\beta_0 | y, X) \propto \left[ v + \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{s^2} (\beta_0 - \hat{\beta}_0)^2 \right]^{-\frac{(v+1)}{2}}, \quad -\infty < \beta_0 < \infty \quad (17)$$

$$P(\beta_1 | y, X) \propto \left[ v + \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{s^2} (\beta_1 - \hat{\beta}_1)^2 \right]^{-\frac{(v+1)}{2}}, \quad -\infty < \beta_1 < \infty \quad (18)$$

gibi elde edilir (Zellner, 1971).

$\sigma$ 'da bilinmeyen bir değişken olarak ele alındığından marjinal sonsal dağılımını elde etmek için, denklem (15) ile elde edilen birleşik sonsal dağılımın  $\beta_0$  ve  $\beta_1$ 'e göre integrali alınır. Bunun sonucunda;

$$P(\sigma | y, X) \propto \frac{1}{\sigma^{v+1}} \exp\left(-\frac{vs^2}{2\sigma^2}\right), \quad 0 < \sigma < \infty \quad (19)$$

ifadesine ulaşılır. Burada  $\sigma$ 'nın Ters-Gamma dağılımına sahip olduğu görülür. Varyans bilindiği durumda, elde edilen sonsal dağılım modelinde parametre olarak katsayılar kalır. Model denklem (16) yapısında olur. Hataların Normal dağıldığı varsayımı altında lineer regresyon modellerinde parametrelerin sonsal dağılımının çok değişkenli *t* dağılımı olduğunu ve varyansın dağılımının Ters-Gamma dağılımı olduğunu 1973 yılında Box ve Tiao yaptıkları çalışmada ispatlamışlardır.

Genel olarak, bilgi vermeyen önsel dağılımla çalışıldığında, nokta tahmini ve aralık tahmini gibi sonuç çıkarımları, en küçük kareler yöntemi ile aynı olacaktır (Gamerman and Lopes, 2006).

#### 4.1.2. Bilgi veren önsel dağılım ile sonuç çıkarımı

Birçok durumda, araştırmacı önceden yapılan araştırmalardan gelen önsel bilgiye sahiptir. Daha önceki denemelerden bilgi edinilmesi yoluyla oluşturulan önsel dağılım veriye dayalı önsel dağılım



iken, kişisel gözlemler veya teorik dayanaklar ile oluşturulan önsel dağılım ise veriye dayalı olmayan önsel dağılım olarak adlandırılır.

Önsel bilgiyi ifade eden önsel dağılım, olabilirlik fonksiyonu ile birleşebilecek uygunlukta bir matematiksel yapıya sahip olursa sonuç çıkarımı daha iyi yapılır. Çünkü, bu şekilde elde edilen sonsal dağılım da sonuç çıkarımı yapabilmek için uygun bir fonksiyonel yapıya sahip olur. Bu özellikte olan önsel dağılımlar “Doğal Eşlenik Önsel Dağılım” veya “Eşlenik Önsel Dağılım” olarak adlandırılır. Olabilirlik fonksiyonu ile birleştiğinde ortaya çıkan sonsal dağılım, önsel dağılım ile aynı dağılım sınıfına ait ise böyle dağılımlara “Eşlenik Önsel Dağılım” denir. “Doğal Eşlenik Önsel Dağılım” ise olabilirlik fonksiyonu ile birleştiğinde ortaya çıkan sonsal dağılım ile aynı dağılımdır ve buna ek olarak olabilirlik fonksiyonu ile de aynı yapıdadır. Doğal Eşlenik aile ile önsel dağılım oluşturulurken ilk olarak olabilirlik fonksiyonu yeterli istatistikler cinsinden yazılmalıdır (Raiffa and Schlaifer, 1968).

Varyans bilinmiyorken regresyon modeli için olabilirlik fonksiyonu,

$$P(y|X, \beta_0, \beta_1, \sigma) \propto P_1(y, \sigma|X, \beta_0, \beta_1) \times P_2(y|\sigma)$$

$$P_1(y, \sigma|X, \beta_0, \beta_1) = \exp \left\{ -\frac{1}{2\sigma^2} [n(\beta_0 - \hat{\beta}_0)^2 + (\beta_1 - \hat{\beta}_1)^2 \sum_{i=1}^n x_i^2 + 2(\beta_0 - \hat{\beta}_0)(\beta_1 - \hat{\beta}_1) \sum_{i=1}^n x_i] \right\} \quad (20)$$

$$P_2(y|\sigma) = \frac{1}{\sigma^n} \exp \left\{ -\frac{y^2}{2\sigma^2} \right\} \quad (21)$$

biçimindedir. (20) ile verilen dağılım Normal dağılıma uyar,  $\beta_0$  ve  $\beta_1$  yeterli istatistiktir. (21) ile verilen dağılım Ters- Gamma dağılımına uymaktadır ve burada,  $\sigma$  yeterli bir istatistiktir.

Yeterli istatistiklerin oluşturduğu bu yapıdan doğal eşlenik önsel dağılıma geçildiğinde, notasyonda da gerekli değişiklikler yapılarak doğal eşlenik önsel dağılım tahmin edicilerini ifade etmek gerekir. Buna göre yeterli istatistiklerin yerini alacak önsel dağılım tahmin edicileri  $\bar{\beta}_0, \bar{\beta}_1, \bar{v}, \bar{s}^2$  ile, önsel dağılıma ait gözlem sayısı  $t$  ile ve gözlemler  $\sum_{j=1}^t x_j$  ( $j = 1, 2, \dots, t$ ) ile gösterilsin. Veriye-dayalı ön bilgi olduğu ve önsel dağılım ile örneklem dağılımının varyanslarının birbirine eşit olduğu varsayımı ile  $\beta_0, \beta_1$  parametreleri için önsel dağılım,

$$P(\beta_0, \beta_1|y, X) = \frac{1}{\sigma^t} \exp \left\{ -\frac{1}{2\sigma^2} [t(\beta_0 - \bar{\beta}_0)^2 + (\beta_1 - \bar{\beta}_1)^2 \sum_{j=1}^t x_j^2 + 2(\beta_0 - \bar{\beta}_0) \times (\beta_1 - \bar{\beta}_1) \sum_{j=1}^t x_j] \right\} \quad (22)$$

standart sapma parametresi için önsel dağılım,

$$P(\sigma|y) = \frac{1}{\sigma^{\bar{v}+1}} \exp \left\{ -\frac{\bar{v}\bar{s}^2}{2\sigma^2} \right\} \quad (23)$$

şeklinde elde edilir. Denklem (22) ve Denklem (23)'deki ifadelerin birleşmesiyle ve normalleştirme sabitlerinin atılmasıyla ortak önsel olasılık yoğunluk fonksiyonu elde edilir. Bu elde edilen dağılım, Normal-Gama dağılımı olur ve

$$P(\beta_0, \beta_1, \sigma|y, X) = \sigma^{-t-\bar{v}-1} \exp \left\{ -\frac{1}{2\sigma^2} \left[ \bar{v}\bar{s}^2 + t(\beta_0 - \bar{\beta}_0)^2 + (\beta_1 - \bar{\beta}_1)^2 \sum_{j=1}^t x_j^2 + 2(\beta_0 - \bar{\beta}_0) \times (\beta_1 - \bar{\beta}_1) \sum_{j=1}^t x_j \right] \right\} \quad (24)$$

biçiminde ifade edilir. Bu ifadenin  $\sigma$ 'ya göre integrali alınarak,  $\beta$ 'lar için marjinal önsel olasılık yoğunluk fonksiyonu elde edilir. Benzer şekilde,  $\beta$ 'ya göre integral alınarak  $\sigma$  için önsel olasılık yoğunluk fonksiyonu elde edilir. Yapılan işlemler sonucunda  $\beta$  parametreleri için elde edilen marjinal önsel yoğunluğun çok değişkenli *Student-t* dağılımı olduğu görülür (Zellner, 1971).

Sonsal dağılımın olasılık yoğunluk fonksiyonunu elde etmek için Bayes teoremine göre doğal eşlenik önsel dağılım ile olabilirlik fonksiyonu çarpılır. Buna göre,

$$P(\beta_0, \beta_1, \sigma|y, X) \propto \sigma^{-n} \exp \left\{ -\frac{1}{2\sigma^2} \left[ v\bar{s}^2 + n(\beta_0 - \hat{\beta}_0)^2 + (\beta_1 - \hat{\beta}_1)^2 \sum_{i=1}^n x_i^2 + 2(\beta_0 - \hat{\beta}_0) \times (\beta_1 - \hat{\beta}_1) \sum_{i=1}^n x_i \right] \right\} \times \sigma^{-t-\bar{v}-1} \exp \left\{ -\frac{1}{2\sigma^2} \left[ \bar{v}\bar{s}^2 + t(\beta_0 - \bar{\beta}_0)^2 + (\beta_1 - \bar{\beta}_1)^2 \sum_{j=1}^t x_j^2 + 2(\beta_0 - \bar{\beta}_0) \times (\beta_1 - \bar{\beta}_1) \sum_{j=1}^t x_j \right] \right\} \quad (25)$$

olup, parametrelere bağlı olmayan ifadelerin atılması ve gerekli işlemlerin yapılması sonucunda ortak sonsal dağılımın son hali,

$$P(\beta_0, \beta_1, \sigma|y, X) \propto \sigma^{-n-t-\bar{v}-1} \exp \left\{ -\frac{1}{2\sigma^2} \left[ (\bar{v}\bar{s}^2 + v\bar{s}^2) + T \left( \beta_0 - \frac{n\hat{\beta}_0 - t\bar{\beta}_0}{T} \right)^2 + \left( \sum_{i=1}^n x_i^2 + \sum_{j=1}^t x_j^2 \right) \times \left( \beta_1 - \frac{(\hat{\beta}_1 \sum_{i=1}^n x_i^2) + (\bar{\beta}_1 \sum_{j=1}^t x_j^2)}{\sum_{i=1}^n x_i^2 + \sum_{j=1}^t x_j^2} \right)^2 \right] \right\} \quad (26)$$

( $T=n+t$ ) olarak elde edilir.

Ortak sonsal dağılım, Normal- Gama dağılımı biçimindedir. Burada,  $\beta_0$  ve  $\beta_1$  için marjinal sonsal dağılım,  $\bar{\beta}_0$  ve  $\bar{\beta}_1$  ortalamalı ve  $\sigma^2 \left( \frac{1}{(\sum_{i=1}^n x_i^2 + \sum_{j=1}^t x_j^2)} \right)$  varyanslı Normal dağılım olur.  $\sigma$  için marjinal sonsal dağılım ise,  $\bar{v} \bar{s}^2$  parametrelili Ters- Gama dağılımı olur (Zellner, 1971).

Varyans biliniyorsa, varyans artık bir parametre olmayacağından modelde sadece katsayı parametreleri için hesaplama yapmak gerekecektir. Bu nedenle bu kısımda sadece temel farklardan bahsedilecektir.

Normal dağılım için  $\sigma$  bilindiğinde olabilirlik fonksiyonunu denklem (20)'de,

$$P(y, \sigma | X, \beta_0, \beta_1) \propto \exp \left\{ -\frac{1}{2\sigma^2} [n(\beta_0 - \hat{\beta}_0)^2 + (\beta_1 - \hat{\beta}_1)^2 \sum_{i=1}^n x_i^2 + 2(\beta_0 - \hat{\beta}_0) \times (\beta_1 - \hat{\beta}_1) \sum_{i=1}^n x_i] \right\} \quad (27)$$

gibi elde edilmişti. Bu yapıyla uygun düşecek olan önsel dağılım da Normal dağılımdır;

$$P(\beta_0, \beta_1 | y, X) = \exp \left\{ -\frac{1}{2\sigma^2} [t(\beta_0 - \bar{\beta}_0)^2 + (\beta_1 - \bar{\beta}_1)^2 \sum_{j=1}^t x_j^2 + 2(\beta_0 - \bar{\beta}_0) \times (\beta_1 - \bar{\beta}_1) \sum_{j=1}^t x_j] \right\} \quad (28)$$

Bu iki dağılımın çarpımından elde edilecek sonsal dağılımın olasılık yoğunluk fonksiyonu ise;

$$P(\beta_0, \beta_1, \sigma | y, X) \propto \exp \left\{ -\frac{1}{2\sigma^2} \left[ T \left( \beta_0 - \frac{n\hat{\beta}_0 - t\bar{\beta}_0}{T} \right)^2 + (\sum_{i=1}^n x_i^2 + \sum_{j=1}^t x_j^2) \times \left( \beta_1 - \frac{(\hat{\beta}_1 \sum_{i=1}^n x_i^2) + (\bar{\beta}_1 \sum_{j=1}^t x_j^2)}{\sum_{i=1}^n x_i^2 + \sum_{j=1}^t x_j^2} \right)^2 \right] \right\} \quad (29)$$

olup görüldüğü gibi bu dağılım yine Normal dağılımdır ve elde edilen sonsal dağılımın ortalaması Bayesci tahmini verecektir.

#### 4.2. Çoklu doğrusal regresyon modeline bayesci yaklaşım

$\beta$ ,  $k \times 1$  boyutlu regresyon katsayıları vektörü olmak üzere,

$$y = X\beta + \varepsilon \quad (30)$$

modeli göz önüne alınsın.  $\beta$  ve  $\sigma$  parametre ve  $X$  biliniyor iken  $y$  için olabilirlik fonksiyonu;

$$P(y|X, \beta, \sigma) \propto \frac{1}{\sigma^n} \exp \left[ -\frac{1}{2\sigma^2} (y - X\beta)'(y - X\beta) \right] \quad (31)$$

biçiminde ifade edilir.

$$s^2 = \frac{(y - X\hat{\beta})'(y - X\hat{\beta})}{v}, v = n - k, \hat{\beta} = (X'X)^{-1}X'y$$

olmak üzere olabilirlik fonksiyonu,

$$P(y|X, \beta, \sigma) \propto \frac{1}{\sigma^n} \exp \left\{ -\frac{1}{2\sigma^2} [vs^2 + (\beta - \hat{\beta})'X'X(\beta - \hat{\beta})] \right\} \quad (32)$$

biçiminde yazılır (Ekici, 2005).

#### 4.2.1. Bilgi vermeyen önsel dağılım ile sonuç çıkarımı

$\beta$  ve  $\sigma$  için belirsiz önsel dağılım,

$$P(\beta, \sigma) \propto \frac{1}{\sigma} , \quad -\infty < \beta_i < \infty, \quad 0 < \sigma < \infty, \quad i = 1, 2, \dots, k \quad (33)$$

biçiminde göz önüne alınsın. Bu durumda, sonsal dağılım olabilirlik fonksiyonu ve önsel dağılım göz önüne alınarak orantısal olarak,

$$P(\beta, \sigma|y, X) \propto \frac{1}{\sigma^{n+1}} \exp \left\{ -\frac{1}{2\sigma^2} [vs^2 + (\beta - \hat{\beta})'X'X(\beta - \hat{\beta})] \right\} \quad (34)$$

biçiminde yazılabilir. Denklem (34)'den,  $\sigma$  ve  $\beta$  için koşullu sonsal olasılık yoğunluk fonksiyonunun,  $\beta$  ortalamalı ve  $\sigma^2(X'X)^{-1}$  varyanslı ( $k$ -boyutlu) çok değişkenli Normal dağılım olduğu görülür. Buradan  $\sigma$ 'ya göre integral alınarak  $\beta$  için marjinal sonsal olasılık yoğunluk fonksiyonu,

$$P(\beta|y) = \int_0^\infty P(\beta, \sigma|y) d\sigma = \int_0^\infty P(\beta|\sigma, y) P(\sigma|y) d\sigma$$

$$P(\beta|y) \propto \left[ 1 + \frac{1}{v} (\beta - \hat{\beta})' \frac{X'X}{s^2} (\beta - \hat{\beta}) \right]^{-\frac{(k+v)}{2}} \quad (35)$$

biçiminde elde edilir. Elde edilen bu dağılımın çok değişkenli *Student-t* dağılımı şeklinde olduğu görülür. Tek tek parametrelere göre çıkarım yapılmak isteniyor ise istenilen  $\beta_i$  parametresi dışındaki  $\beta$ 'lara göre integral alınmalıdır.  $\sigma$  için marjinal sonsal olasılık yoğunluk fonksiyonu ise,

$$P(\sigma|y) = \int_0^{\infty} P(\beta, \sigma|y)d\beta$$

$$P(\sigma|y) = \frac{1}{\sigma^{v+1}} \exp\left\{-\frac{vs^2}{2\sigma^2}\right\} \quad (36)$$

biçiminde elde edilir (Ekici, 2005). Görüldüğü gibi, önsel bilgi dahil edilmediğinde yani bilgi vermeyen önsel dağılımlar kullanıldığında sonsal dağılımın ortalaması olan Bayes tahmin edicisi en küçük kareler tahmin edicisi ile aynı olacaktır.

#### 4.2.2. Bilgi veren önsel dağılım ile sonuç çıkarımı

Regresyonun belirtilen varsayımlarından ve Bayesci sonuç çıkarımının doğasından dolayı doğal eşlenik önsel dağılım, *Normal-Gamma* formunda bir dağılımdır. Matris notasyonu ile doğal eşlenik önsel dağılımın yapısı,

$$P(\beta|\sigma) \propto \sigma^{-k} \exp\left\{-\frac{1}{2\sigma^2} (\beta - \bar{\beta})' \Sigma (\beta - \bar{\beta})\right\} \quad (37)$$

$$P(\sigma|y) \propto \sigma^{-\bar{v}-1} \exp\left\{-\frac{\bar{v}s^2}{2\sigma^2}\right\}, \bar{v} = t - k \quad (38)$$

biçimindedir. Denklem (37) ve denklem (38) birleştirilerek ortak önsel dağılım,

$$P(\beta|\sigma) \propto \sigma^{-k-\bar{v}-1} \exp\left\{-\frac{1}{2\sigma^2} \left(\bar{v}s^2 + (\beta - \bar{\beta})' \Sigma (\beta - \bar{\beta})\right)\right\} \quad (39)$$

olarak yazılır.  $\bar{\beta}$ ,  $\bar{v}$ ,  $\bar{s}^2$  ve  $\Sigma$  önsel dağılımın parametrelerinin tahmin edicileridir. Burada,  $\Sigma$  pozitif tanımlı simetrik bir matrisidir.

$\sigma$  biliniyor iken  $\beta$  vektörü,  $k$  değişkenli ortalaması  $\bar{\beta}$  ve varyans-kovaryans matrisi  $\sigma^2 \Sigma^{-1}$  olan *Normal* dağılımdır.  $\beta$  biliniyor iken  $\sigma$ 'nın dağılımı *Ters-Gamma* dağılımıdır. Buna göre olabilirlik fonksiyonu,

$$P(y|\beta, \sigma) \propto \frac{1}{\sigma^n} \exp\left\{-\frac{1}{2\sigma^2} \left(vs^2 + (\beta - \hat{\beta})' X'X (\beta - \hat{\beta})\right)\right\}, v = n - k \quad (40)$$

ve sonsal dağılım,

$$P(\beta, \sigma|y) \propto \sigma^{-n-k-\bar{v}-1} \exp \left\{ -\frac{1}{2\sigma^2} ((vs^2 + \bar{v}\bar{s}^2) + (\beta - \bar{\beta})' \Sigma (\beta - \bar{\beta}) + (\beta - \hat{\beta})' X'X (\beta - \hat{\beta})) \right\} \quad (41)$$

dır. Sonsal dağılımı tüm terimleriyle istenilen biçime getirmek için birkaç işlem yapılması gerekir. Denklem (41)'da önsel dağılımın tahmin edicileri ve en küçük kareler tahmin edicileri ile oluşturulan karesel ifadeler

$$W = (\beta - \bar{\beta})' \Sigma (\beta - \bar{\beta}) + (\beta - \hat{\beta})' X'X (\beta - \hat{\beta})$$

ile gösterilir ise bu terimler ayrı ayrı,

$$\begin{aligned} (\beta - \bar{\beta})' \Sigma (\beta - \bar{\beta}) &= \beta' \Sigma \beta - 2\bar{\beta}' \Sigma \beta + \bar{\beta}' \Sigma \bar{\beta} \\ (\beta - \hat{\beta})' X'X (\beta - \hat{\beta}) &= \beta' X'X \beta - 2\hat{\beta}' X'X \beta + \hat{\beta}' X'X \hat{\beta} \end{aligned} \quad (42)$$

biçiminde yazılır. Her iki satırda  $\beta$ 'yı içermeyen ifadeler göz ardı edilirse denklem (42) ile verilen iki ifadenin toplamı;

$$W = \beta' \Sigma \beta - 2\bar{\beta}' \Sigma \beta + \beta' X'X \beta - 2\hat{\beta}' X'X \beta = \beta' (\Sigma + X'X) \beta - 2(\bar{\beta}' \Sigma + \hat{\beta}' X'X) \beta \quad (43)$$

ile verilir. Burada,

$$M = \Sigma + X'X \text{ ve } m = \bar{\beta}' \Sigma + \hat{\beta}' X'X \quad (44)$$

alınmak üzere Denklem (44)'deki eşitliklerden faydalanarak, denklem (43) kısaca

$$W = \beta' M \beta - 2m' \beta \quad (45)$$

biçiminde ifade edilir. Denklem (45)'teki ifadenin  $(b^2 - 2mb)$ 'ye benzer bir yapıda olduğu görülür. "Kareye tamamlama" işlemi ile  $(b^2 - 2mb) = (b - m)^2 - m^2$  eşitliğine ulaşılır. Aynı mantık ile denklem (45)'teki ifade,

$$W = \beta' M \beta - 2m' \beta = (\beta - M^{-1}m)' M (\beta - M^{-1}m) - m' M^{-1}m \quad (46)$$

biçiminde olur. Burada,

$$M^{-1}m = \bar{\beta} \text{ ve } M^{-1} = (\Sigma + X'X)^{-1} \quad (47)$$

denilirse sonsal dağılımın ortalaması,  $\bar{\beta} = (\Sigma + X'X)^{-1} (\bar{\beta}\Sigma + \hat{\beta}X'X)$  eşitliğini sağlar. Denklem (47) ile verilen ifadeye göre denklem (46)'daki ifade,

$$W = \beta'M\beta - 2m'\beta = (\beta - \bar{\beta})'(\Sigma + X'X)(\beta - \bar{\beta}) - (\bar{\beta}\Sigma + \hat{\beta}X'X)'\bar{\beta} \quad (48)$$

biçimini alır. Denklem (48) ile elde edilen sonuç, denklem (41)'de tekrar yerine koyulur ise,

$$P(\beta, \sigma|y) \propto \sigma^{-n-k-\bar{v}-1} \exp \left\{ -\frac{1}{2\sigma^2} \left( (vs^2 + \bar{v}\bar{s}^2) + (\beta - \bar{\beta})'(\Sigma + X'X)(\beta - \bar{\beta}) - (\bar{\beta}\Sigma + \hat{\beta}X'X)'\bar{\beta} \right) \right\} \quad (49)$$

elde edilir. Parametreyi içermediği için denklem (42)'de göz ardı edilen ifadeler, denklem (49)'a ilave edilir ve

$$P(\beta, \sigma|y) \propto \sigma^{-n-k-\bar{v}-1} \exp \left\{ -\frac{1}{2\sigma^2} \left( (vs^2 + \bar{v}\bar{s}^2) + (\beta - \bar{\beta})'(\Sigma + X'X)(\beta - \bar{\beta}) + \bar{\beta}\Sigma\bar{\beta} + \hat{\beta}X'X\hat{\beta} - (\bar{\beta}\Sigma + \hat{\beta}X'X)'\bar{\beta} \right) \right\} \quad (50)$$

elde edilir. Burada, Denklem (50) ile yapılan birkaç işlem sonucunda,

$$P(\beta, \sigma|y) \propto \sigma^{-n-k-\bar{v}-1} \exp \left\{ -\frac{1}{2\sigma^2} \left( (vs^2 + \bar{v}\bar{s}^2) + (\beta - \bar{\beta})'(\Sigma + X'X)(\beta - \bar{\beta}) + \frac{\Sigma X'X(\bar{\beta} - \hat{\beta})'(\bar{\beta} - \hat{\beta})}{(\Sigma + X'X)} \right) \right\} \quad (51)$$

elde edilir. Denklem (51)'deki sonsal dağılımın varyansının parametrelerine ilişkin,

$$\bar{v}\bar{s}^2 = (vs^2 + \bar{v}\bar{s}^2) + \frac{\Sigma X'X(\bar{\beta} - \hat{\beta})'(\bar{\beta} - \hat{\beta})}{(\Sigma + X'X)}, \bar{v} = n + \bar{v} \quad (52)$$

eşitliği elde edilir. (51)'deki dağılım yeterli istatistiklerine göre yazılırsa,

$$\propto \sigma^{-k} \exp \left[ -\frac{1}{2\sigma^2} \left( (\beta - \bar{\beta})' (\Sigma + X'X) (\beta - \bar{\beta}) \right) \right] \cdot \sigma^{-\bar{v}-1} \exp \left( -\frac{\bar{v}\bar{s}^2}{2\sigma^2} \right) \quad (53)$$

elde edilir. Denklem (53)'e göre ortak sonsal dağılım,  $\bar{\beta}$  ortalamalı ve  $\sigma^2(\Sigma + X'X)^{-1}$  varyanslı çok değişkenli *Normal* dağılımdır.  $\sigma$  için marjinal sonsal dağılımın ise  $\bar{v}\bar{s}^2$  parametrelili *Ters-Gamma* dağılımı olduğu görülür.  $\beta$  için marjinal sonsal olasılık yoğunluk fonksiyonu çok değişkenli *Student-t* dağılımıdır ve aşağıdaki gibi

$$P(\beta|y) \propto \left[ 1 + \frac{1}{\bar{v}} \left( (\beta - \bar{\beta})' \frac{(\Sigma + X'X)}{\bar{s}^2} (\beta - \bar{\beta}) \right) \right]^{-\frac{\bar{v}-k}{2}} \quad (54)$$

edilir (Ekici, 2005).

Bu bölümde, doğrusal regresyon modellerinde Bayesci yaklaşım ile sonuç çıkarımı yapmak için kullanılacak teorik sonuçlar anlatılmıştır. Fakat bir önceki bölümde de ifade edildiği gibi günümüzde bilgisayar teknolojilerinin de gelişmesi ile birlikte Bayesci sonuç çıkarımı yapılırken MCMC olarak adlandırılan stokastik simülasyon yöntemlerinden faydalanılmaktadır. Bu yöntemler kullanılarak sonsal dağılımın kapalı hali elde edilemese bile dağılımdan çok sayıda örneklem çekilebilmekte ve bu örneklem kullanılarak sonuç çıkarımı yapılabilmektedir. Bu çalışmada sonuç çıkarımı için MCMC yöntemlerinden Gibbs algoritması anlatılmıştır.

## 5. Gibbs örnekleme algoritması

İstenilen bir olasılık dağılımından birbirinden bağımsız rasgele sayılar üretmek için Monte Carlo simülasyon yöntemi kullanılır. MCMC yöntemleri ile ise, üretilen her yeni değer bir önceki değere bağlı olarak zincir değeri biçiminde üretilir. Eğer bu zincir yeterince uzun çalışırsa ilgilenilen sonsal dağılım istenilen halini bulur. Elde edilen örneklem kullanılarak sonsal dağılıma ilişkin ortalama, medyan gibi istenilen istatistikler hesaplanabilir, Bayesci güvenilir aralıklar bulunabilir.

En çok kullanılan MCMC yöntemleri Metropolis, Metropolis-Hasting ve Gibbs algoritmalarıdır. Metropolis ve Metropolis-Hasting algoritmasının özel bir hali olan Gibbs algoritması, ilk olarak Amerikalı fizikçi Josiah W. Gibbs, sonrasında ise Geman (1984) tarafından isimlendirilmiştir. Bayesci çıkarım için Gibbs örneklemesini ilk kez Gelfand ve arkadaşları kullanmıştır. Gibbs algoritması kullanılarak örnekleme yapılırken, değişkenler arasındaki ilişki ile



ilgili bilgiye ihtiyaç duyulur, çünkü Gibbs örnekleme koşullu olasılıklara dayanır. Metropolis algoritmasında olduğu gibi yardımcı bir dağılıma gerek duymaz.

$\theta = (\theta_1, \dots, \theta_k)'$  parametre vektörü,  $p(y|\theta)$  olabirlik fonksiyonu,  $\pi(\theta)$  önsel dağılım olsun. Buna göre,  $\pi(\theta_i|\theta_j, i \neq j, y)$  sonsal dağılımı,

$$\pi(\theta_i|\theta_j, i \neq j, y) \propto p(y|\theta) \pi(\theta)$$

gibi yazılabilir. Gibbs örnekleme için algoritma adımları aşağıdaki gibidir:

1.  $t = 0$  için keyfi bir  $\theta^{(0)} = \{\theta_1^{(0)}, \dots, \theta_k^{(0)}\}$  başlangıç değeri seçilir.

2.  $\theta'$  nın her bir bileşeni

$$\begin{aligned} &\pi(\theta_1|\theta_2^{(t)}, \dots, \theta_k^{(t)}, y)' \text{ den } \theta_1^{(t+1)}, \\ &\pi(\theta_2|\theta_1^{(t+1)}, \theta_3^{(t)}, \dots, \theta_k^{(t)}, y)' \text{ den } \theta_2^{(t+1)}, \\ &\dots\dots\dots \\ &\pi(\theta_k|\theta_1^{(t+1)}, \dots, \theta_{k-1}^{(t+1)}, y)' \text{ den } \theta_k^{(t+1)} \end{aligned}$$

şeklinde elde edilir.

3.  $t = t + 1$  alınır ve  $T$  istenilen örneklem büyüklüğü olmak üzere  $t < T$  ise 2.nci adıma gidilir değilse işlem bitirilir (Steyvers, 2015).

MCMC simülasyon yöntemlerinde, başlangıç değerlerinin etkileri zincirin ilk adımlarında görülür. Daha ileri adımlarda zincir yavaş yavaş başlangıç durumunu unuttur ve durağanlaşır. Analizlerde başlangıç değerlerinin etkisinden kurtulabilmek için zincirden elde edilen ilk gözlemler ihmal edilir. Bunlara, burn in (yakma) değerleri denir (Link and Barker, 2010). Literatürde, başlangıç değerlerinin etkisinden kurtulmak için zincir uzunluğunu belirlemeye yarayan birçok yöntem vardır. Bu yöntemlerden en çok kullanılanı, çekilen örneklemin grafiğini inceleyerek zincir uzunluğuna karar vermektir (Gilks and et al., 1996; Walsh, 2004).

## 6. Uygulama Çalışması

Uygulama çalışmasında,

$$y = \beta_0 + \beta_1 x_1 + \varepsilon$$

biçiminde basit doğrusal regresyon modeli göz önüne alınmıştır. Sırasıyla,  $\beta_0 = 1, 1.5, 3$  ve  $\beta_1 = 2, 3, 2$  ve  $x$  değerleri için  $\varepsilon \sim N(0,1)$  olmak üzere Matlab programı yardımıyla örneklem sayıları  $n = 20, 30$  olarak alınmış ve  $y$  değerleri simülasyon ile türetilmiştir. İlk olarak, Matlab programı yardımıyla  $\beta_0$  ve  $\beta_1$  parametreleri klasik regresyon yöntemi ile tahmin edilmiştir. Daha sonra aynı veriler kullanılarak WinBugs programında bilgi veren önsel olarak  $\beta_0$  ve  $\beta_1$  lar için  $Normal(0, 0.001)$  ve varyans için  $Gamma(0.001, 0.001)$  dağılımları alınmış ve Bayesci yöntem ile parametre tahminleri yapılmış ve sonuçlar Tablo 1 ile verilmiştir. Bayesci yöntemle parametre tahmininde kullanılan Markov zinciri için 11000 örneklem çekilmiş ve örneklemin baştan ilk 1000 verisi burn in değerleri olarak alınıp işlemlere dahil edilmemiştir. Tahmin sonuçlarını karşılaştırabilmek için klasik yöntemde güven aralıkları, Bayesci yöntemde ise güvenilir aralıklar hesaplanmıştır.

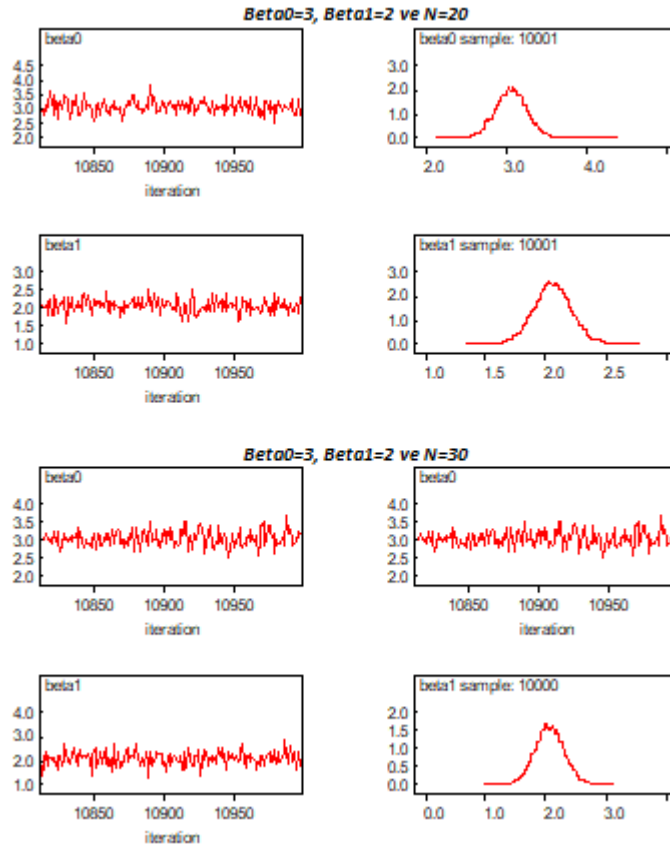
**Tablo 1.** Simülasyon sonuçları

Parametre Değerleri	n	Klasik Regresyon Yöntemi		Bayesci Regresyon Yöntemi	
		$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$
$\beta_0=1$ $\beta_1=2$	20	1.0516	2.0247	1.049	2.021
	GA	[0.6660, 1.4372]	[1.5183, 2.5311]	[0.6661, 1.430]	[1.513, 2.521]
	SH	0.18354	0.24104	0.002057	0.002773
	30	1.0685	2.0564	1.065	2.055
	GA	[0.6074, 1.5297]	[1.5524, 2.5604]	[0.6089, 1.529]	[1.552, 2.558]
	SH	0.12513	0.14604	0.002687	0.002463
$\beta_0=1.5$ $\beta_1=3$	20	1.5053	3.0666	1.503	3.063
	GA	[1.0848, 1.9257]	[2.5364, 3.5968]	[1.085, 1.918]	[2.535, 3.586]
	SH	0.20012	0.25237	0.002168	0.002903
	30	1.5265	3.0933	1.524	3.092
	GA	[1.1945, 1.8585]	[2.7628, 3.4239]	[1.195, 1.850]	[2.763, 3.421]
	SH	0.16207	0.16136	0.001965	0.001615
$\beta_0=3$ $\beta_1=2$	20	3.0673	2.056	3.064	2.054
	GA	[2.6586, 3.4760]	[1.7277, 2.3842]	[2.659, 3.470]	[1.731, 2.378]
	SH	0.19453	0.15624	0.002152	0.00178
	30	3.0499	2.0579	3.047	2.056
	GA	[2.6643, 3.4356]	[1.5569, 2.5590]	[2.667, 3.430]	[1.556, 2.550]
	SH	0.18827	0.14461	0.002275	0.002448

\*SH: klasik yöntem için Standart Hata, Bayesci yöntem için MC Standart Hatası; GA: klasik yöntem için %95 güven düzeyinde güven aralığı, Bayesci yöntem için %95 lik güvenilir aralık

Tablo 1'e bakıldığında klasik yöntemle elde edilen basit doğrusal regresyon parametre tahminleri ile Bayesci yöntemle elde edilen parametre tahminleri benzer sonuçlar vermiştir ve elde edilen bu sonuçlar simülasyonla verilen gerçek değerlere oldukça yakındır. Bu parametre tahminlerine ilişkin güven aralıklarına bakıldığında Bayesci yöntemle elde edilen güvenilir aralıkların klasik yöntemle elde edilen güven aralıklarına göre daha dar bir aralık oluşturduğu görülmektedir. Ayrıca, klasik ve Bayesci yöntemle elde edilen güven aralıkları 0 değerini içermediklerinden parametrelerin modele katkıları önemlidir sonucu çıkarılabilir.

Bayesci yöntemle elde edilen parametre tahminlerine örnek olması açısından,  $\beta_0 = 3$ ,  $\beta_1=2$  olduğu durum göz önüne alınmış ve  $n=20$ , 30 için elde edilen parametre değerlerine ilişkin grafik (iz grafiği) ve sonsal yoğunluk grafikleri aşağıdaki gibi elde edilmiştir.



Şekil 1.  $\beta_0 = 3$  ve  $\beta_1=2$  ve  $n=20$  ve 30 değerleri için iz ve sonsal yoğunluk grafikleri

Sonuçlara göre parametre değerleri belli bir ortalama etrafında (simülasyonla verilen gerçek değer etrafında) salınım göstermekte ve sonsal olasılık yoğunluk grafikleri de Normal dağılıma benzerlik göstermektedir.

Gerçek veri ile uygulama çalışması için ise (Erar, 2013)'te verilen teslim süresi verileri kullanılmıştır. Problem aşağıdaki gibidir.

“Bir meşrubat şişeleme fabrikası, dağıtım sistemindeki otomatik satış makinelerinin servis güzergâhlarını incelemektedir. Güzergâhtaki bir operatörün, satış mağazalarından birindeki otomatik

satış makinesine bakım yapılabilmesi için yaklaşık olarak gereken süre kestirilmek istenmektedir. Bakım hizmetleri, içecek ürünleriyle makineleri stoklama, küçük çaplı bakım ya da temizliği içermektedir. Bu çalışma için endüstri mühendisi iki önemli değişkenin teslim süresini ( $y$ ) etkilediğini öne sürmüştür: Stoklanmış ürün sayısı ( $x_1$ , teslim hacmi) ve güzergâhtaki operatörün yürüyerek kattettiği mesafe ( $x_2$ ). Mühendis teslim süresi için 25 gözlem almıştır.” Bu gözlemler Tablo 2 ile verilmiştir. Teslim süresi verileri için,

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

çoklu doğrusal regresyon modelinin uygun olduğu düşünülmektedir.

**Tablo 2.** Teslim süresi verileri

Gözlem verileri	Teslim süresi (y)	Teslim hacmi ( $x_1$ )	Mesafe ( $x_2$ )	Gözlem verileri	Teslim süresi (y)	Teslim hacmi ( $x_1$ )	Mesafe ( $x_2$ )
1	16.68	7	560	14	19.75	6	462
2	11.50	3	220	15	24.00	9	448
3	12.03	3	340	16	29.00	10	776
4	14.88	4	80	17	15.35	6	200
5	13.75	6	150	18	19.00	7	132
6	18.11	7	330	19	9.50	3	36
7	8.00	2	110	20	35.10	17	770
8	17.83	7	210	21	17.90	10	140
9	79.24	30	1460	22	52.32	26	810
10	21.50	5	605	23	18.75	9	450
11	40.33	16	688	24	19.83	8	635
12	21.00	10	215	25	10.75	4	150
13	13.50	4	255				

Burada,  $x_1$  ile teslim hacmi,  $x_2$  ile mesafe alınmak üzere, Bayesci sonuç çıkarımı için, WinBUGS programında, önsel olarak, modeldeki tüm parametreler için, ortalaması 0 ve kesinliği 0.001 (kesinlik= 1/varyans) olan Normal dağılım alınmıştır. Modelin kesinliği (tau) için, alfa ve teta parametresi çok küçük olan Gamma dağılımı alınmıştır. Bayesci yöntemle parametre tahmininde kullanılan Markov zinciri için 11000 örneklem çekilmiştir ve alınan örneklemin baştan ilk 1000 verisi burn in değerleri olarak atılmış, analize dahil edilmemiştir. Buna göre elde edilen sonuçlar aşağıdaki gibidir.

node	mean	sd	MC error	2.5%	median	97.5%
beta0	2.356	1.158	0.01253	0.09126	2.348	4.658
beta1	1.613	0.1787	0.001848	1.261	1.612	1.965
beta2	0.01441	0.003745	4.07E-5	0.007083	0.01439	0.02196
tau	0.09396	0.02862	3.218E-4	0.04629	0.09128	0.1576

Şekil 2. Özet istatistikler

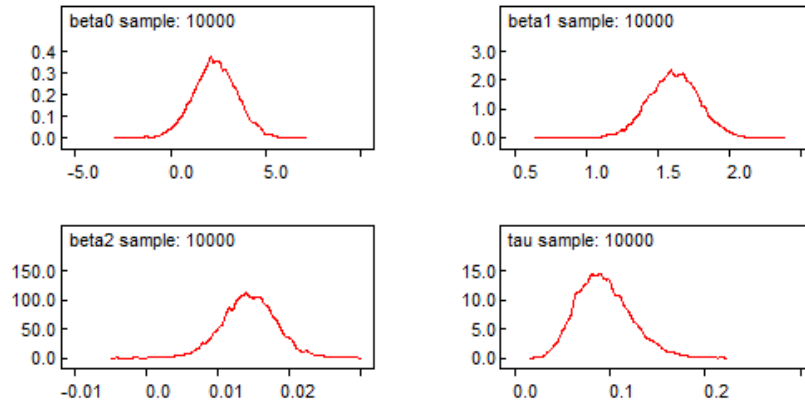
Modeldeki MC hataları oldukça küçük elde edilmiştir. Burada model aşağıdaki yapıda oluşturulmuş;

$$\hat{y} = \text{beta0} + \text{beta1} \times x1 + \text{beta2} \times x2$$

ve elde edilen sonuçlara göre parametre tahminleri;

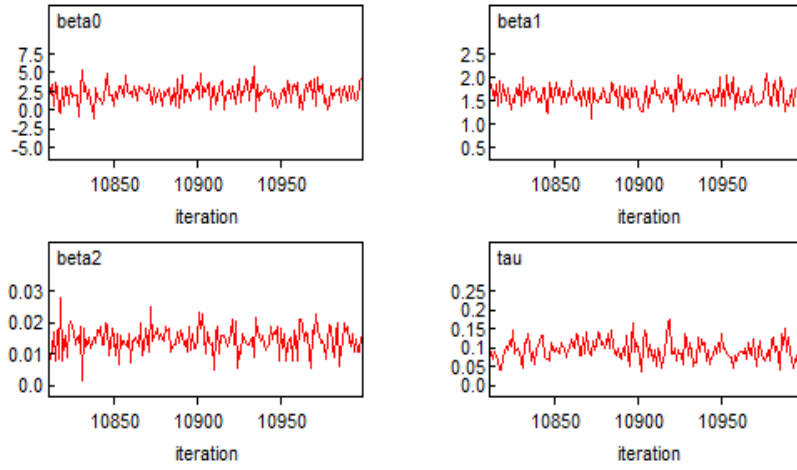
$$\beta = \begin{bmatrix} \text{beta0} \\ \text{beta1} \\ \text{beta2} \end{bmatrix} = \begin{bmatrix} 2.356 \\ 1.61 \\ 0.01459 \end{bmatrix}$$

olarak bulunmuştur. Sonsal yoğunluk fonksiyonlarının grafikleri aşağıda verilmiştir.



Şekil 2. Tahmin edicilerin Yoğunluk Fonksiyonu

Şekillere göre parametrelerin sonsal dağılımlarının Normal dağılıma yakın olduğu görülmektedir. Kesinlik için ise çarpık bir dağılım olduğu görülmektedir.



Şekil 3. Tahmin edicilerin İz Grafikleri

Şekil 3’te verilen iz grafiklerine bakıldığında örneklemenin belli bir ortalama etrafında gezindiği görülmektedir. Verilere ilişkin elde edilen klasik tahmin sonuçları ve Bayesci tahmin sonuçları aşağıda Tablo 3 ile özetlenmiştir.

Tablo 3. Teslim verilerine ilişkin sonuçlar

	Klasik Regresyon	%95 Güven Düzeyinde Güven aralıkları	p-değeri	Standart Hata	Bayesci Regresyon	%95 lik Güvenilir aralıklar	Monte Carlo Standart hatası
$\hat{\beta}_0$	2.341	[0.0658, 4.6161]	0.044	1.097	2.356	[0.1313, 4.658]	0.01253
$\hat{\beta}_1$	1.6159	[1.2618, 1.9700]	0.0001	0.1707	1.610	[1.266, 1.963]	0.001848
$\hat{\beta}_2$	0.01438	[0.00689, 0.0218]	0.001	0.00361	0.01459	[0.0077, 0.0224]	4.07E-5
$R^2$	0.9559				0.9609		

Bayesci yönteme göre elde edilen regresyon modeli,

$$\hat{y} = 2.356 + 1.610 \times x_1 + 0.01459 \times x_2$$

gibidir. Teslim süresinde kullanılan mesafe ve teslim hacmi değişkenlerinin önemli olduğu görülmüştür. Modele göre, diğer değişkenler sabit tutulduğunda operatörün yürüyerek kattığı mesafedeki bir birimlik artış teslim süresinde 1.610’luk bir süre artışına sebep olur. Modele göre, diğer değişkenler sabit tutulduğunda stoklanmış ürün sayısındaki bir birimlik artış teslim süresinde 0.01459’luk bir süre artışına sebep olur.

## 7. Sonuç

Bu çalışmada, ilk olarak basit ve çoklu doğrusal regresyon modelleri verilmiştir. Bu regresyon modellerine ait varsayımlar hatırlatılmış ve en küçük kareler yöntemi ile parametrelerin tahmin edicileri elde edilmiştir. Daha sonra Bayesci yaklaşım anlatılmıştır. Bayesci sonuç çıkarımı için gerekli adımlar kullanılarak regresyon modellerinde parametrelere ait sonsal dağılımların elde edilişi açıklanmıştır. Daha sonra Bayesci sonuç çıkarımında yaygın olarak kullanılan MCMC yöntemlerinden Gibbs örnekleme algoritması anlatılmıştır.

Simülasyon çalışması için Matlab programı kullanılarak göz önüne alınan basit doğrusal regresyon modeli için  $(x, y)$  ikilileri türetilmiştir. Bu veriler kullanılarak klasik ve Bayesci yöntemler ile regresyon parametreleri tahmin edilmiş ve güven aralıkları hesaplanmıştır. İki yöntem kullanılarak elde edilen sonuçlar Tablo 1 ile verilmiştir. Tablo 1 incelendiğinde her iki yöntemle elde edilen tahmin sonuçlarının birbirine oldukça yakın olduğu görülmüştür. Bayesci yaklaşım ile bulunan güvenilir aralıkların klasik yöntemle elde edilenlere göre daha dar bir aralık oluşturduğu gözlenmiştir.

Simülasyon çalışmasından sonra literatürde verilmiş olan gerçek veriler için hem klasik hem de Bayesci regresyon ile sonuçlar elde edilmiştir. Bu sonuçlara göre Bayesci yaklaşımla bulunan güvenilir aralıkların klasik regresyon modeline göre elde edilenden daha dar olduğu görülmüştür.

Simülasyon ve gerçek veri sonuçlarına göre, Bayesci regresyon modeli ile elde edilen tahmin sonuçları klasik yöntemle elde edilen tahmin sonuçlarına oldukça yakındır. Fakat Bayesci yöntemle elde edilen tahmin aralıkları klasik yöntemle göre daha dardır. Bu sonuçlara göre, Bayesci regresyon modeli yaklaşımının klasik regresyon modeline alternatif olarak kullanılabilceği söylenebilir.

### **Yazarların Katkısı**

Tüm yazarlar çalışmaya eşit katkıda bulunmuştur.

### **Çıkar Çatışması Beyanı**

Yazarlar arasında herhangi bir çıkar çatışması bulunmamaktadır.

### **Araştırma ve Yayın Etiği Beyanı**

Yapılan çalışmada araştırma ve yayın etiğine uyulmuştur.

**Kaynaklar**

- Albert, J., H., Chib, S. (1993). Bayesian Analysis of Binary and Polychotomous Response Data, *Journal of the American Statistical Association*, Vol. 88, No. 422, pp. 669- 679.
- Box, G., E., P., Tiao, G.,C. (1973). *Bayesian Inference in Statistical Analysis*, John Wiley And Sons, Inc.
- Congdon, P. (2004). *Bayesian Statistical Modelling*. John Wiley& Sons Inc., West Sussex.
- Ekici, O. (2005). *Bayesyen Regresyon ve WinBUGS ile Bir Uygulama*, İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, Yüksek lisans tezi, 115s, İstanbul.
- Erar, A., M. (2013). *Doğrusal Regresyon Analizine Giriş*. Nobel Yayıncılık, Ankara
- Gamerman, D., Lopes, H. F. (2006). *Markov Chain Monte Carlo Stochastic Simulation for Bayesian Inference* (Second Edition), London.
- Geman, S., Geman, D. (1984). Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions On Pattern Analysis And Machine Intelligence*, 721-741
- Gelman A., Carlin, J., B., Stern, H., S., Rubin, D., B. (2004). *Bayesian Data Analysis*. Chapman-Hall, Florida.
- Gilks, W., R., Richardson, S., Spiegelhalter, D. J. (1996). *Markov Chain Monte Carlo in Practice*, London.
- Heath, S., Snow, G., Thompson, E., Tseng, C., Wijsman, E. (1997). MCMC segregation and linkage analysis. *Genet. Epidemiol.*, 14: 1011-1016. [https://doi.org/10.1002/\(SICI\)1098-2272\(1997\)14:6<1011:AID-GEPI75>3.0.CO;2-L](https://doi.org/10.1002/(SICI)1098-2272(1997)14:6<1011:AID-GEPI75>3.0.CO;2-L)
- Link ,W. A., Barker, R. J. (2010). *Bayesian Inference With Ecological Applications*, Academic Press, Boston, United States.
- Raiffa, H., Schlaifer, R. (1968). *Applied Statistical Decision Theory*, America, MIT Press.
- Steyvers, M. (2015). *Advanced Matlab: Exploratory Data Analysis and Computational Statistics*.
- Walsh, B. (2004). *Markov Chain Monte Carlo and Gibbs Sampling*. <http://nitro.biosci.arizona.edu/courses/EEB519A-2007/pdfs/Gibbs.pdf>
- Wundervald, B. (2019). *Bayesian Linear Regression*, Technical Report, June 2019, <https://Brunaw.Com/Phd/Bayes-Regression/Report.pdf>, doi: 10.13140/RG.2.2.28385.97121
- Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics*. John Wiley & Sons, New York.