BAYES AÇIORTAY REGRESYON TEKNİĞİ VE BİR UYGULAMA

YÜKSEK LİSANS TEZİ

Ece ÖZGÖREN

Danışman
Doç. Dr. Sinan SARAÇLI

İSTATİSTİK ANABİLİM DALI

Ağustos 2019

Bu tez çalışması 18.FEN.BİL.29 numaralı proje ile
Afyon Kocatepe Üniversitesi BAPK tarafından desteklenmiştir.

AFYON KOCATEPE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

BAYES AÇIORTAY REGRESYON TEKNİĞİ VE BİR UYGULAMA

Ece ÖZGÖREN

Danışman
Doç. Dr. Sinan SARAÇLI

İSTATİSTİK ANABİLİM DALI

Ağustos 2019

TEZ ONAY SAYFASI

Ece ÖZGÖREN tarafından hazırlanan "Bayes Açıortay Regresyon Tekniği ve Bir Uygulama" adlı tez çalışması lisansüstü eğitim ve öğretim yönetmeliğinin ilgili maddeleri uyarınca 19/08/2019 tarihinde aşağıdaki jüri tarafından **oy birliği** ile Afyon Kocatepe Üniversitesi Fen Bilimleri Enstitüsü **İstatistik Anabilim Dalı'nda YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Danışman	: Doç. Dr. Sinan SARAÇLI	**İmza**
Başkan	: Prof. Dr. Birdal ŞENOĞLU Ankara Üniversitesi, Fen Fakültesi	
Üye	:Doç. Dr. Sinan SARAÇLI Afyon Kocatepe Üniversitesi, Fen-Edebiyat Fakültesi	
Üye	:Prof. Dr. İsmet DOĞAN Afyonkarahisar Sağlık Bilimleri Üniv., Tıp Fakültesi	

Afyon Kocatepe Üniversitesi

Fen Bilimleri Enstitüsü Yönetim Kurulu'nun

........./......../........ tarih ve

…………………sayılı kararıyla onaylanmıştır.

....

Prof. Dr. İbrahim EROL

Enstitü Müdürü

BİLİMSEL ETİK BİLDİRİM SAYFASI
Afyon Kocatepe Üniversitesi

Fen Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmasında;

- Tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- Görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- Başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- Atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- Kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- Ve bu tezin herhangi bir bölümünü bu üniversite veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

19/08/2019

Ece ÖZGÖREN

ÖZET

Yüksek Lisans Tezi

BAYES AÇIORTAY REGRESYON TEKNİĞİ VE BİR UYGULAMA

Ece ÖZGÖREN

Afyon Kocatepe Üniversitesi

Fen Bilimleri Enstitüsü

İstatistik Anabilim Dalı

Danışman: Doç. Dr. Sinan SARAÇLI

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 açıortay regresyon denklemleri hesaplanmıştır. Daha önceki çalışmalarda Tip II regresyon teknikleri arasında en iyi performansı sergileyen tekniğin açıortay 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 açıortayı alınarak Bayes açıortay denklemi hesaplanmıştır. Önsel ve mevcut bilgi verisine dayalı sonsal dağılımları elde etmek amacıyla farklı örneklem hacimlerinden yararlanılmış ve Bayes regresyon denklemlerinin performansları HKO ve GE kriterlerine göre karşılaştırılmıştır.

Araştırma bulgularına göre n=100, n=75 ve n=50 birimlik örneklemlerde Bayes açıortay tekniğinin performansının en düşük HKO değerine sahip olduğu, dolayısıyla mevcut veri setine ait bu örneklem hacimleri için en iyi performansı sergilediği belirlenmiştir.

2019, viii + 87 sayfa

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

ABSTRACT

M.Sc. Thesis

BAYESIAN BİSECTOR REGRESSION AND AN APPLICATION

Ece Özgören

Afyon Kocatepe University

Graduate School of Natural and Applied Sciences

Department of Statistics

Supervisor: Assoc. Prof. Sinan SARAÇLI

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 and RE criteria.

The results of the study indicates that performance of Bayesian bisector technique has the minimum MSE for the sample sizes n=100, n=75 and n=50 which means that the performance of Bayesian bisector technique is the best for these sample sizes for the related data set.

2019, viii + 87 pages

Keywords: Bayesian Approaches, Type II Regression, Bayesian Regression, Measurement Error Models

TEŞEKKÜR

Bu tezin hazırlanması sürecinde, bana danışmanlık yaparak, yol gösteren, bilgisini her zaman içtenlik ve özenle aktaran, ayrıca zorlandığım zamanlarda manevi desteğini her zaman hissettiğim danışmanım Sayın Doç. Dr. Sinan Saraçlı'ya en içten teşekkürlerimi sunarım.

Yapmış olduğum çalışmada yardımları ile bana destek olan, bilgilerini esirgemeyen ve sonuca ulaşmamda büyük payı olan saygıdeğer hocam Prof. Dr. Birdal Şenoğlu'na teşekkürlerimi sunarım.

Beni bu günlere getiren, maddi-manevi desteğini hiçbir zaman esirgemeyen, zor ve güzel zamanlarımı her zaman paylaştığım, en büyük desteği aldığım değerli aileme, anneme, babama ve kardeşime sonsuz teşekkürlerimi sunarım.

Yüksek lisansa başladığım günden itibaren yanımda olan, bana her türlü konuda içtenlikle yardımcı olan, takıldığım ve zorlandığım zamanlarda vermiş olduğu manevi destek ile beni güçlendiren kıymetli arkadaşım Berkalp Tunca'ya teşekkürü borç bilirim.

Lisans ve yüksek lisans eğitimim boyunca yanlarında yaşadığım, bana her konuda anne ve baba içtenliği ile destek olan, maddi ve manevi desteklerini esirgemeyen dedem, babaannem ve anneanneme sonsuz teşekkürlerimi sunarım.

Ayrıca bu tezin oluşmasına katkı sağlayan 18.FEN.BİL.29 proje numaralı BAPK'ne teşekkürlerimi sunarım.

Ece ÖZGÖREN

AFYONKARAHİSAR, 2019

İÇİNDEKİLER DİZİNİ

Sayfa

SİMGELER ve KISALTMALAR DİZİNİ

Simgeler

μ	Kitle Ortalaması
σ	Kitle Standart Sapması
β_0	Regresyon Denkleminin Sabit Katsayısı
β_1	Regresyon Denkleminin Eğim Katsayısı
$\hat{\beta}_0$	Regresyon Denkleminin Sabit Katsayısı için EKK Tahmin Edicisi
$\hat{\beta}_1$	Regresyon Denkleminin Eğim Katsayısı için EKK Tahmincisi
$\bar{\bar{\beta}}_0$	Regresyon Denkleminin Sabit Katsayısı için Bayes Tahmincisi
$\bar{\bar{\beta}}_1$	Regresyon Denkleminin Eğim Katsayısı için Bayes Tahmin Edicisi
$\bar{\beta}_0$	Bayes Regresyon Denkleminin Önsel Dağılım için Sabit Katsayı Tahmincisi
$\bar{\beta}_1$	Bayes Regresyon Denkleminin Önsel Dağılım için Eğim Katsayı Tahmincisi
EKK(X|Y)	X Bağımlı Y Bağımsızken Regresyon Modeli
EKK(Y|X)	Y Bağımlı X Bağımsızken Regresyon Modeli
exp	Üstel
n	Örneklem Hacmi
N	Kitle Hacmi
$\bar{\bar{v}}\bar{\bar{s}}^2$	Bayes Regresyon modeline ait Hata Kareler Toplamı
s^2	Y Gözlem Değerinin Varyansı
x_i	X Değişkeninin i. Gözlem Değeri
y_i	Y Değişkeninin i. Gözlem Değeri
$\hat{Y}_i$	Y Değişkeninin i. Tahmin değeri
$\bar{Y}$	Y Değişkeninin Ortalaması
$\bar{X}$	X Değişkeninin Ortalaması
ε	Hata Terimi
$\propto$	Proportional-Oransal

Kısaltmalar

EKK	En Küçük Kareler
EÇO	En Çok Olabilirlik Yöntemi
HKO	Hata Kareler Ortalaması
HKT	Hata Kareler Toplamı
MCMC	Markov Zinciri Monte Carlo Algoritması
MA	Major Axis-Majör Eksen
DEM	Deming
RMA	Reduced Major Axis-İndirgenmiş Majör Eksen
MSE	Mean Square Error-Hata Kareler Ortalaması
GE	Göreli Etkinlik

ŞEKİLLER DİZİNİ

Sayfa

ÇİZELGELER DİZİNİ

Sayfa

1. GİRİŞ

Bayesci yaklaşımın özü, ele alınan bir çalışmada, konu ile ilgili gerçekleşen tüm bilginin analize dahil edilmesine dayanmaktadır. Gözlem verileri ile birlikte, araştırmacının konuya ilişkin önceden edindiği bilgiler ve sübjektif yorumlarının da analize dahil edilmesi Bayesci yaklaşımı klasik yaklaşımdan ayıran en önemli özelliktir. Bayesci yaklaşımda esas amaç bilinmeyen parametrelere ilişkin tahminler yapmaktır. Parametre tahminlerinin elde edilmesi sürecinde araştırmacının önceden edindiği bilgileri analize dahil etmemesi bilgi kaybına neden olabilmektedir. Ancak Bayesci yaklaşım ile kurulan matematiksel modellerde önsel bilgi ile mevcut veri setine ait bilgiler birleştirilerek sonsal dağılım hesaplanabilmektedir. Böylece ilgilenilen çalışma ile ilgili bilgi kaybı olmadan yeniden modeller oluşturulabilmektedir. Buna ek olarak, Bayesci yaklaşım tahmin edilmek istenen parametrelerin de bir dağılımı olduğunu varsayarak birer rasgele değişken olarak kabul eder. Böylece yapılan analizlerde parametreler için Bayesci yaklaşım ile doğrudan olasılık ifadeleri belirlenmiş olur. Bayesci yaklaşımın diğer bir önemli unsuru ise küçük örneklem hacimlerinde yapılan parametre tahminlerinde klasik yaklaşımlara göre daha iyi sonuçlar vermesidir. 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. Bu nedenle klasik regresyon ile oluşturulan modellere kıyasla her iki ölçüm hatasını da içeren Tip II Regresyon tekniğinin kullanılması daha uygun olacaktır. Literatürde yer alan Tip II regresyon teknikleri: EKK-Açıortay Regresyon Tekniği, Majör Eksen Regresyon tekniği, İndirgenmiş Majör Eksen Regresyon tekniği, Deming Regresyon tekniği, Passing-Bablok Regresyon tekniği ve York Regresyon tekniğidir. Bu tekniklerin hepsinin ortak noktası bağımsız değişken olan X'lerin de ölçüm hatası değerlerini içeren modeller kurulabilmesidir.

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 ve GE kriterlerine göre değerlendirilecektir.

2. LİTERATÜR BİLGİLERİ

2.1 Bayes Yaklaşımı

Bu bölümde Bayes yaklaşımının tarihsel gelişim süreci, Bayes teoremi, Bayes Tahmin edicisi, Bayesci yaklaşımın temelleri ve istatistiksel kavramlara ilişkin yorumları, önsel dağılım, sonsal dağılım ve Bayesci yaklaşım ile klasik yaklaşım arasındaki kavramsal farklılıklar ile ilgili konulara yer verilerek açıklanacaktır.

2.1.1 Tarihsel Gelişim Süreci

Bayesci olasılık kuramı, matematiksel istatistik kuramının bir dalıdır. Bu kuram belirsizlik taşıyan herhangi bir durumun modelini oluşturmak, bu durumla ilgili evrensel doğruları ve gözlemleri kullanarak sonuçlar üretmek amacıyla kullanılır. Bilimsel karar yöntemlerinden biri olan Bayesci yaklaşım, olasılıklı bir bilginin incelemesinde objektif bir bakış açısını esas alır. Bu yaklaşım bilimsel gerçekten çok, bilginin aşamalarına odaklanır. Bu kuramın dayanağı, koşullu olasılıklarla ilgili olan Bayes Kuralı'dır (Çevik 2009).

İstatistik gelişirken temel olarak iki farklı felsefi yaklaşımın belirginleştiği görülmektedir. Klasik (veya Frekansçı, Berkeley istatistiği) yaklaşım ve Bayesci yaklaşım. Bu disiplinin başlangıç aksiyomlarının yorumlanmasında, pek çok konu ve kavramın ele alınışında bu yaklaşımlardan biri diğerine alternatif olmuştur. Ancak zaman içinde, bu konuda çalışanların tutumuna bağlı olarak ve belki de algılanması daha kolay olması nedeniyle, istatistik alanında klasik yaklaşım daha fazla hâkim olmuştur (Ekici 2005).

Bayesyen yaklaşımın ilk temeli, İngiltere Tunbridge Wells'de yaşayan, bir rahip ve matematikçi olan Thomas Bayes (1702(?)-1761) 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 (Stigler 1983).

İstatistik yazını incelendiğinde, 18. yüzyılın sonlarından 20. yüzyılın başlarına kadar istatistiksel çıkarsamanın Bayesci yaklaşımın etkisinde olduğu görülmektedir (Ekici 2005). Yine aynı zaman aralığında Thomas Bayes ’in çalışmaları Laplace tarafından oldukça geliştirilerek olasılık teorisine büyük katkı sağlamıştır. Daha sonra, Bayes ve Laplace tarafından kullanılan çıkarım yöntemleri ve kavramlar R. A Fisher, J. Neyman, E. S Pearson, A. Wald ve birçok bilim adamı tarafından kullanılarak teorik istatistikte daha detaylı ve başarılı çalışmalar ortaya çıkmasında etkili olmuştur (Anscombe 1961).

Unwin (2004)’ e göre 1920 ve 50’ li yıllar arasında F. Ramsey, B. de Finetti ve L. Savage gibi ünlü isimlerin temsil ettiği bir matematik okulu, frekansçıların bakış açılarının çok dar olduğu görüşü üzerinden yeni bir olasılık kuramı geliştirmiştir. Olasılık kuramının matematiksel formalizmine bağlı kalan bu kişiler, matematiksel olasılıklarını düşünme biçimimizi yeniden tanımlamaya çalışmışlardır. Bu bilim adamlarının araştırmaları bugün olasılıkların Bayesyen yorumu diye adlandırılan kuramla sonuçlanmıştır (Çelebioğlu ve Yaman 2005).

Bayesci yaklaşımın popülaritesini yitirmeye başladığı yıllarda F. P Ramsey’ in “Truth and Probability” (“Gerçeklik ve Olasılık”) adlı denemesi ile bu olasılık teorisi yeniden önem kazanmıştır. 1950’li yıllarda B. de Finetti ve H. Jeffreys oldukça farklı bir bakış açısı ile yaptıkları kapsamlı çalışmalarla bu yaklaşıma katkı sağlamışlardır. Ayrıca Bayesci yaklaşımda hipotez testi ilk defa H. Jeffreys (1961) tarafından kullanılmıştır. Matematiksel istatistikçiler için böyle bir teorinin en kapsamlı çalışması H. J Savage tarafından yapılmıştır. R. Schlaifer iş ve endüstri dünyasında oluşan problemlere karşı Bayesci yaklaşım ile yaptığı çalışmalar ile yeni çözümler getirmiştir. Böylelikle Bayes teoremi bilim adamlarınca geliştirilmiş veya yeniden yorumlanmış ve istatistiksel metodolojide bir devrim niteliğinde gelişmeler yaşanmıştır (Anscombe 1961).

2.1.2 Bayes Teoremi

Bayes yaklaşımı, özü 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 (Altındağ 2015).

Bayes teoremi koşullu olasılıkların hesaplanmasında kullanılan matematiksel bir formüle dayanmaktadır. A ve B olayları U örnek uzayında herhangi iki olay ve $P(A) \neq 0$ ve $P(B) \neq 0$ olmak üzere, B olayı biliniyorken A olayının koşullu olasılığı $P(A|B)=P(A \cap B)/P(B)$'dir. Buna göre $P(A \cap B)= P(B)\ P(A|B)$ olur. Benzer şekilde A olayı biliniyorken B olayının gerçekleşmesi olasılığı $P(B|A)=P(A \cap B)/P(A)$'dır. $P(B|A)$'da $P(A \cap B)$ yerine $P(B)\ P(A|B)$ yazılabilir. Bu durumda $P(B|A)=[P(B)\ P(A|B)]P(A)$ yazılabilir. $P(A \cap B)$ ise A ve B olaylarının birlikte gerçekleşmesi olasılığını ifade eder. A ve B olaylarının olasılıkları arasındaki bu ilişki "Bayes Teoremi" olarak bilinmektedir.

Bayes teoremi A olayı biliniyorken B'nin olasılığını hesaplamaya yaradığı gibi, B olayı biliniyorken A'nın olasılığını hesaplamaya da yarar. B olayı $B_1, B_2, \ldots, B_n$ gibi n tane ayrık olayın bir kümesi ise $P(B_i)>0$ ve $P(A_i)>0$ VE A, B'de bir olay olmak üzere $A = (B_1 \cap A) \cup (B_2 \cap A) \cup \ldots \cup (B_n \cap A)$ yazılabilir. $P(B_i \cap A)$ olayları ayrık olaylar olduklarından $P(A) = \sum_{i=1}^{n} P(B_i \cap A)$'dır. Yazılan koşullu olasılık ifadesine göre $P(B_i \cap A)= P(B_i)\ P(A|B_i)$'dir. Buna göre $P(A) = \sum_{i=1}^{n} P(B_i) P(A \mid B_i)$ şeklinde yazılabilir. Bu durumda A olayı biliniyorken B_i olayının gerçekleşmesi Bayes teoremine göre Eşitlik 2.1'de verildiği gibi gösterilir.

$$P(B_i \mid A) = \frac{P(B_i \cap A)}{P(A)} = \frac{P(B_i)P(A \mid B_i)}{P(A)} = \frac{P(B_i)P(A \mid B_i)}{\sum_{i=1}^{n} P(B_i)P(A \mid B_i)} \qquad (2.1)$$

Burada, $P(B_i)$ olasılıkları önsel olasılık, $P(A|B_i)$ olasılığı ise olabilirlik fonksiyonunu temsil etmektedir Bu iki bilginin birleştirilmesi ile oluşan $P(B_i|A)$ koşullu olasılığı ise sonsal olasılığı ifade eder. Bayesyen çıkarsamada $P(B|A)$ ifadesindeki B, analize konu olan parametrenin belirli bir değerini temsil etmektedir. Bayes teoremi genel olarak Eşitlik 2.2'de verildiği gibi ifade edilebilir (Karagöz 2004, Kendall and Buckland 2009, Miller and Miller 2014, Ntzoufras 2009).

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)} \propto P(B \mid A)P(A) \tag{2.2}$$

Bayes teoreminin uygulanabilmesi için $P(B_i)$ olarak verilen önsel bilgi olasılıklarının bilinmesi gerekmektedir. Bu teoremden hareketle parametre tahmini yapılabilir (Gündoğdu 2016). Bayesci yaklaşımda olasılık yoğunluk fonksiyonu $p(y \mid \theta)$ ile tanımlı istatistiksel bir model kullanılarak $y = \{y_1, \dots, y_n\}$ verisine ait θ parametresinin tahmini elde edilmek istenir. Bayesci yaklaşımda bu çıkarımın temelleri aşağıda verilmiştir (Cengiz vd. 2012).

1. θ için, $\pi(\theta)$ ile gösterilen bir olasılık dağılımı formüle edilir. Bu dağılım, önsel dağılım ya da sadece önsel olarak adlandırılır. Önsel dağılım, veri bilinmeden önceki parametre hakkındaki bilgileri ifade eder.
2. Gözlemlenmiş y veri seti için, θ verildiğinde y'nin dağılımını tanımlayan bir $p(y \mid \theta)$ olabilirlik fonksiyonu belirlenir.
3. Önsel ve olabilirlik güncellenerek $p(\theta \mid y)$ sonsal dağılımı hesaplanır ve θ hakkındaki bütün istatistiksel çıkarımlar sonsal dağılımdan elde edilir.

Yukarıda bahsedilen 3 adım Bayes Teoremi kullanılarak Eşitlik 2.3'de görüldüğü gibi formüle edilebilir.

$$p(\theta \mid y) = \frac{p(\theta, y)}{p(y)} = \frac{p(y \mid \theta)\pi(\theta)}{p(y)} = \frac{p(y \mid \theta)\pi(\theta)}{\int p(y \mid \theta)\pi(\theta)\, d\theta} \tag{2.3}$$

Burada $p(y)$, sonsal dağılımı (sürekli veriler için integral, kesikli veriler için toplam alınarak) 1'e eşitleyen "normalleştirme" katsayısı olarak adlandırılır. θ parametresinin olabilirlik fonksiyonu $p(y \mid \theta)$ eşdeğer olarak $L(\theta \mid y)$ orantılı fonksiyonu şeklinde de yazılabilmektedir. Sonuç olarak $p(y \mid \theta) \propto L(\theta \mid y)$ şeklinde ifade edilebilir (Congdon 2003). Ayrıca $p(y)$ marjinal dağılımı bir integraldir ve integral sonlu olduğu sürece, bu integralin değeri sonsal dağılım hakkında herhangi bir ilave bilgi vermez. Bu nedenle $p(y \mid \theta)$ denklemde verilen orantılı formda Eşitlik 2.4'de belirtildiği üzere rasgele sabit olarak yazılabilir (Cengiz vd. 2012).

$$p(\theta \mid y) \propto L(\theta)\pi(\theta) \qquad (2.4)$$

Bayes teoremi en genel haliyle:

$$sonsal\ olasılık \propto önsel\ olasılık\ .\ olabilirlik\ fonksiyonu(likelihood) \qquad (2.5)$$

şeklinde ifade edilmektedir (Robert *et al.* 2009).

Olabilirlik fonksiyonu Bayes teoreminde çok önemli bir rol oynar. Çünkü bu fonksiyon y verisi ile θ parametresine ait önsel bilginin elde edilebilmesini sağlamaktadır. Bu nedenle θ parametresine ait tüm örneklem bilgisini temsil ettiği kabul edilebilir (Box and Tiao 1973).

2.1.3 Bayes Tahmin Edicisi

Olasılık dağılımları $F = \{f(.;\theta\colon \theta \epsilon \Theta\}$ ailesinde olasılık yoğunluk fonksiyonlarını indisleyen θ parametresi bilinmeyen bir sabit olarak ele alındığı gibi bir olasılık dağılımına sahip rasgele değişken olarak da ele alınmaktadır. İlki klasik, ikincisi Bayes yaklaşımı olarak isimlendirilmektedir. Klasik yaklaşımda θ parametresi, parametre kümesinde bilinmeyen bir değer olmak üzere $X_1,X_2,\ldots,X_n$ örneklemindeki X_i' ler bağımsız ve aynı $f(.;\theta)$ olasılık yoğunluk fonksiyonuna sahip rasgele değişkenler olarak ele alınır. Bayes yaklaşımında bir rasgele değişken olan θ 'nın dağılımı önsel (prior) dağılım olarak adlandırılır. Genellikle bu öznel (subjective) bir dağılım olmak üzere $X_1,X_2,\ldots,X_n$ örneklemi, θ rasgele değişkeni ile koşullandırılmış dağılımdan alınmış gibi düşünülmektedir, yani $i = 1,2,\ldots,n$ için X_i' ler bağımsız ve her biri $f(.;\theta)$ olasılık (yoğunluk) fonksiyonuna sahip rasgele değişkenler olarak ele alınmaktadır. Bu durumda, örneklemin olasılık (yoğunluk) fonksiyonu Eşitlik 2.6'da verildiği gibi gösterilir:

$$f(x_1, x_2, \ldots, x_n) = f(x_1 \mid \theta)f(x_2 \mid \theta) \ldots f(x_n \mid \theta) = \prod_{i=1}^{n} f(x_i \mid \theta) \qquad (2.6)$$

ve θ 'nın önsel dağılımının olasılık (yoğunluk) fonksiyonu π olmak üzere:

$$f_{X_1,X_{2,\dots,}X_n,\theta}(x_1,x_2,\dots,x_n,\theta) = f(x_1 \mid \theta)f(x_2 \mid \theta)\dots f(x_n \mid \theta)\pi(\theta) \qquad (2.7)$$
$$= \prod_{i=1}^{n} f(x_i \mid \theta)\pi(\theta)$$

Eşitlik 2.7 elde edilir. $X = (X_1, X_2, \dots, X_n)'$ ve $x = (x_1, x_2, \dots, x_n)'$ vektör gösterimleri ile:

$$f(x \mid \theta) = f(x_1 \mid \theta)f(x_2 \mid \theta)\dots f(x_n \mid \theta) = \prod_{i=1}^{n} f(x_i \mid \theta)\pi \qquad (2.8)$$

$$f_{X,\theta}(x \mid \theta) = f(x,\theta)\pi(\theta) \qquad (2.9)$$

Eşitlik 2.8 ve 2.9 yazılabilir. Buradan X' in marjinal dağılımının olasılık (yoğunluk) fonksiyonu Eşitlik 2.10 olmak üzere:

$$f_X(x,\theta) = \int_{\Theta} f(x \mid \theta)\pi(\theta)\, d\theta \qquad (2.10)$$

örneklem değerleri verildiğinde θ'nın koşullu dağılımının, yani θ'nın sonsal (posterior) dağılımının olasılık (yoğunluk) fonksiyonu Eşitlik 2.11'de verilmiştir.

$$\pi(\theta \mid x) = \frac{f_{X,\theta}(x \mid \theta)}{f_X(x)} = \frac{f(x,\theta)\pi(\theta)}{f_X(x)} \qquad (2.11)$$

Sonsal dağılım, bir anlamda, önsel dağılımın gözlenen değerler ile güncellenmesi olarak düşünülebilir. Sonsal dağılımın beklenen değeri, bu gözlemleri üreten dağılımdaki θ için bir tahmin edici olarak kullanılabilir.

$$\hat{\theta}_B = \int_{\Theta} \theta\, \pi(\theta \mid x)\, d\theta \qquad (2.12)$$

Eşitlik 2.12'de belirtilen $\hat{\theta}_B$ istatistiğine θ' nın Bayes tahmin edicisi denir.

Bayes tahmin edicileri, $T(X)$ bir istatistik olmak üzere, seçilmiş bir $k(\theta, T)$ kayıp fonksiyonuna dayalı Eşitlik 2.13'de belirtilmiş olan risk fonksiyonu için,

$$R(\theta, T) = \int k(\theta, T(X)) f(x \mid \theta)\, dx \tag{2.13}$$

Eşitlik 2.14 ile belirtilen integralin beklenen değerini en küçük yapan tahmin edicidir.

$$R(T) = \int_{\Theta} R(\theta, T)\, \pi(\theta)\, d\theta \tag{2.14}$$

Kısaca özetlenecek olursa $T = \{T: R(T) < \infty\}$ olmak üzere, bir $T^* \in \mathrm{T}$ için Eşitlik 2.15'e göre:

$$R(T^*) \leq R(T) \quad , \quad T \in \mathrm{T} \tag{2.15}$$

T^* tahmin edicisine Bayes tahmin edicisi denir. Bu anlamda T^* tahmin edicisi,

$$k(\theta, T) = (T - \theta)^2 \tag{2.16}$$

Eşitlik 2.16 ile belirtilen kayıp fonksiyonuna dayalı olan Bayes tahmin edicisidir (Öztürk vd. 2006).

2.1.4 Bayesci Yaklaşımın Temelleri Ve İstatistiksel Kavramlara İlişkin Yorumları

İstatistikte bir olasılığın objektifliği değerlendirilen süreç hakkında önceden oluşan hiçbir bilginin göz ardı edilmemesiyle sağlanır. Süreç hakkında önceden edinilen bu bilgileri kullanmamak bilgi kaybına sebep olmaktadır. Bayesci yaklaşım süreç hakkında önceden edinilmiş bilgi ile mevcut veri setine ait bilginin birleştirilerek sonsal dağılımın olasılığının hesaplanmasına olanak sağlar. Bu özellik klasik yaklaşım yerine Bayes yaklaşımının kullanılması için önemli bir sebeptir (Bolstad 2007).

Belirsizlik durumunda karar alma halinde birbirine alternatif oluşturacak hareketlerin sonuçları tahmini olarak bilinir. Gerçek durum, hakkında tam bir ihmalin olduğunu varsayan belirsizlik halinde karar verme modellerinden farklı olarak kısmi ihmal durumunda karar alma modeli, karar vericinin durumların gerçekleşme olasılıkları hakkında kısmen bilgi sahibi olduklarını veya kısmi bir ihmalin olduğunu varsayar. Karar vericinin konu ile ilgili bilgisi ve geçmiş tecrübelerinin olduğunu kabul eden bu varsayım ile kurulan model oldukça gerçekçidir, bu sayede karar alıcı problemin formüle edilmesinde ve durumların gerçekleşme olasılıklarının saptanmasında aktif bir rol oynar. Karar vericinin konu ile ilgili bilgisi, geçmiş tecrübeleri ve sezgisinin sayılarla ve matematiksel bir model ile ifade edilmesi olan bu olasılıklara sübjektif veya kişisel olasılıklar denir. Sübjektif olasılıklarla çalışan olasılık teorisine objektif veya klasik istatistik teorisinden farklı olarak sübjektif olasılık teorisi veya Bayesci istatistik denmektedir (Kurtuluş 1998).

2.1.4.1 Belirsizliğin Değerlendirilmesi

Bayesci istatistiğe göre bir teori olsun, bir önerme veya bir nedensellik ilişkisi olsun, her kapsamdaki belirsizlik olasılıklarla ifade edilmelidir. Yaklaşımın asıl fikri budur. Klasik yaklaşım belirsizliklerde deterministik davranır. Varsayımlar doğrultusunda söz konusu belirsizliği, orada iddia edilen ilişkiyi, sıklıklarına göre değerlendirerek, kabul edilmesi ya da edilmemesi yönünde karar verir. Bu bağlamda Bayesci yaklaşımı benimseyenler, belirsizlikle ifade edilen bir teoriyi tamamen kabul etmenin veya körü körüne reddetmenin bilgi oluşturma sürecine pek bir katkı sağlamayacağı, aksine önemli sayılabilecek yanılsamalara yol açacağı seklinde eleştiriler getirebilirler. Öte yandan bu eleştirileri getirirken kendilerini de "Bayesci yaklaşımda belirsiz olan ilişki, olasılığının hesaplanması suretiyle bir derece aydınlatılmış olur. Böylelikle bilgi ve karar sürecine daha net, yanıltıcı olmayan ilaveler yapar." seklinde savunabilirler (Ekici 2009).

2.1.4.2 Olasılık

Bayesci yaklaşımda olasılık, Sübjektif olarak yorumlanmaktadır. Sübjektif yaklaşımı esas alarak gelişen Bayesçi istatistikte bir olayın olasılığı, araştırmacının konuya ilişkin önceden edindiği bilgiler (önsel bilgi) ile denemelerden elde edilen sonuçların (verinin) birleştirilmiş halidir. Bir araya getirme işlemi, Bayes teoremine, dolayısıyla koşullu olasılığa dayanmaktadır. Bayesçi istatistikte olasılık "tümevarım olasılığıdır. Amaç, en kesine en yakın sonuç ile ulaşmaktır (Çevik 2009).

Bayesci istatistik tümevarım yöntemini kullanmaktadır, bundan dolayı Bayesci istatistikte amaç en kesin olan bilgiye, yani "1" olasılığına ulaşmaktır. Klasik yaklaşımda ise tümdengelim yöntemi kullanıldığından dolayı esas olan amaç yanlışlama yöntemi ile "0" olasılığını elde etmektir (Ekici 2009).

Çizelge 2.1 Bayesci Yaklaşım ve Klasik Yaklaşıma ilişkin İşleyen Olasılık Süreçleri (Çevik 2009, Ekici 2009)

Bayesci Yaklaşım	Klasik yaklaşım
Varsayımsız deneme	Varsayımlarla deneme
Doğrulama	Yanlışlama
$\frac{1}{2} \ll p(t,g) < 1$	$p(t,g) = 0$

Çizelge 2.1'e göre "t" bir teori ve "g" onunla ilgili elde edilen gözlemler ise farklı olasılık yaklaşımlarına göre süreç çizelgede ifade edildiği gibidir (Çevik 2009, Ekici 2009).

Bayesci yaklaşım hipotez testleri, güven aralıkları, nokta tahminleri gibi istatistiksel yöntemlere yer vermektedir. Ancak klasik yaklaşımda parametreler için kullanılan tek olasılık tanımı uzun dönem göreceli frekanstır (Bolstad 2007). Bu olasılık tanımı bir olayın gerçekleşme olasılığı, o olayın çok sayıda tekrarlanan denemelerinin elde edilen sıklık değerini ifade etmektedir (Ekici 2009).

2.1.4.3 Parametre

Bayesci ve klasik yaklaşımdaki en önemli fark parametrenin tanımıdır. Klasik yaklaşımda parametre kitlenin sabit bir özelliği olarak tanımlanır. Klasik analizlerde amaç parametrenin gerçek değerini tahmin etmek ve tahminler etrafındaki güven aralığını oluşturmaktır. Standart hata bu sürecin ayrılmaz bir parçasıdır ve tekrarlanan örneklere ait tahminlerin değişkenliğini temsil eder. Farkın daha iyi anlaşılması için %95 güven aralığı yorumu incelenecek olursa klasik yapıda %95 olasılıkla parametre A ile B arasına düşer yorumu yanlıştır. Çünkü herhangi bir örneklemden hesaplanan güven aralığı parametreyi ya içerir ya da içermez. Bu yorum yerine güven aralığı, tekrarlanan örnekler üzerinde aralığın beklenen performansı olarak tanımlanır. Örneğin kitleden 100 örnek çekilerek ve her bir örnekten parametre tahmini etrafında %95 güven aralığı oluşturulsun. Bu durumda aralıkların 95 tanesinin kitle parametresini içermesi beklenir. Klasik yapıda, olasılık ifadesi veriler için geçerli olmasına karşın parametre için geçerli değildir (Alkan 2012).

Bayesci yaklaşımda parametre, olasılık dağılımına sahip bir rasgele değişken olarak kabul edilmektedir. Bu yaklaşımda 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 (Bolstad 2007).

2.1.4.4 Nokta tahmini

Ana kitle hakkında bilgi veren karakteristik değere parametre denir. Ancak ana kitleyi gözlemlemek her zaman mümkün olmadığından, parametre değerlerine de ulaşmak mümkün değildir. Bu nedenle ana kitleden elde edilen örnekler yardımıyla ana kitle parametreleri hakkında tahminde bulunmak mümkün olabilmektedir. Bir parametre tahmini ya nokta tahminler ya da aralık tahminler olarak elde edilebilir. Ancak her iki durumun temelinde nokta tahmin bulunmaktadır. Bu nokta tahminlerin parametre değerini en iyi şekilde temsil etmesi için tahmin edicilerin yansızlık, yeterlilik, tutarlılık ve etkinlik özelliklerine sahip olması gerekmektedir (Gasım 2013).

Yansızlık: Tahmin edicilerin yansızlığı, tahmin teorisinde en çok kullanılan özelliklerden biridir. Bir parametre için birden çok yansız tahmin edici bulunabildiği gibi bazen yansız tahmin ediciler olmayabilir. Örneğin, $X_1, X_2 \dots, X_n$, parametresi θ olan kitleden bir örneklem, T de θ'nın herhangi bir tahmin edicisi olmak üzere, her $\theta \epsilon \Theta$ için

$E_\theta(T) = \theta$ oluyorsa, T tahmin edicisine θ parametresi için yansız bir tahmin edicidir denir. Burada yan, $E_\theta(T) = -\theta)$ şeklinde ifade edilmektedir (Akdi 2014).

Yeterlilik: Bilinmeyen θ parametresi için yeterli bir tahmin edici, örneklem içinde parametre hakkındaki bilgiyi özetleyen tahmin edicidir. Yani, örneklem içinde parametre hakkında ne kadar bilgi varsa, hiçbir bilgi kaybı olmadan özetleyen bir tahmin edici, o parametre için yeterlidir. Örneğin, $X_1, X_2 \ldots, X_n$, olasılık veya olasılık yoğunluk fonksiyonu $f(x,\theta)$ olan bir kitleden bir örneklem olsun. $T(X_1, X_2 \ldots, X_n) = T(\underline{X})$ tahmin edicisi θ için yeterli ise, θ hakkında $X_1, X_2 \ldots, X_n$ örneklemine bağlı olarak elde edilen bir sonuç sadece $T(\underline{X})$ tahmin edicisinin değerine bağlıdır. Burada $T(\underline{X})$ bilindiğinde $\underline{X}$'in koşullu dağılımı θ parametresine bağlı değil ise, $T(\underline{X})$'e θ için yeterli bir tahmin edici denir (Akdi 2014).

Tutarlılık: Tahmin edici örneklemin bir fonksiyonu olup, aynı zamanda örneklem hacmine de bağlıdır. Örneğin, $X_1, X_2 \ldots, X_n$ olasılık veya olasılık yoğunluk fonksiyonu $f(x,\theta)$ olan kitleden bir örneklem, $T_n = T_n(\underline{X})$ de θ'nın herhangi bir tahmin edicisi olmak üzere, $n \to \infty$ iken $T_n \xrightarrow{p} \theta$ ise T_n tahmin edicilerinin dizisi θ için tutarlıdır denir (Akdi 2014).

Etkinlik: Tahmin edicilerde aranan özelliklerden biri de etkinliktir. Bu özellik tahmin edicileri karşılaştırmak için de kullanılır. Bir parametre için birçok tahmin edici önerilebilir. Burada tahmin edicilerin ifade edilen yansızlık, yeterlilik, tutarlılık özellikleri ile beraber etkin olması da beklenir. Eğer mevcut ise, en etkin tahmin edici kullanılmak istenir. Örneğin T_n ve $T_n{}^*$ herhangi bir θ parametresi için iki tahmin edici olmak üzere, her θ için $Var_\theta(T_n) \leq Var_\theta(T_n{}^*)$ oluyorsa T_n tahmin edicisi $T_n{}^*$ tahmin edicisine göre daha etkindir denir (Akdi 2014).

Nokta tahmini üzerinde çalışılan kitlenin bir veya birden fazla özelliği hakkında bilgi sahibi olmak için bir yığının örnekten tahmin edilmesi sorununun çözümlenmesi için kullanılan yöntemlerden biridir. Nokta tahmininde iki yöntem vardır. Bilinmeyen parametreler için tahmin ediciler elde etmek; ikincisi; olası tahmin ediciler arasından en iyi tahmin ediciler bulmaktır. Nokta tahmin edicilerini bulmak için çeşitli yöntemler geliştirilmiştir (Gencer 2016). Bu yöntemler En Çok Olabilirlik yöntemi (EÇO), Momentler yöntemi, En Küçük Kareler yöntemi (EKK) ve Bayes yöntemidir.

Nokta tahminini parametrenin bir tahmincisi ve gerçek örnek veriler için aldığı değerin bir değeri olarak adlandırılır. Bayesci istatistikte söz konusu parametrenin nokta tahmini çoğunlukla sonsal dağılımın medyanı veya ortalamasıdır. Bayesci yaklaşımda sonsal (posterior) dağılımın ortalaması Bayesci tahmini olarak kullanılır çünkü sonsal (posterior) dağılım hata kareler ortalamasını en aza indirir. Bu durum sonsal dağılımın en uygun tahmin edici olacağı anlamına gelir (Bolstad 2007).

Bayesci yaklaşımda sonsal dağılımın ortalaması, parametreye ait nokta tahminini ifade etmektedir. Bu durumda sonsal dağılımın varyansının karekökü θ parametresine ait sonsal standart sapma değerinin vermektedir. Bayesci istatistikte bilinmeyen parametre değerine ilişkin hesaplanan nokta tahmini, klasik yaklaşımda kullanılan MLE tahmin yöntemine benzemektedir (Gosh *et al.* 2006).

İstatistiksel nokta tahmini bilinmeyen bir parametreye ait en iyi sonucu veren tahmin yöntemidir. Nokta tahmini A=Θ ($\theta \in \Theta$ olmak üzere) için bir karar verme problemi olarak değerlendirilebilir. Karar fonksiyonları verileri nokta tahminlerine eşleştirir ve ayrıca tahmin ediciler olarak da adlandırılabilirler. Nokta tahmini, çeşitli karar ilkelerinin etkilerini incelemek için etkili ve basitleştirilmiş bir yöntemdir (Parmigiani *et al*. 2009).

Kayıp fonksiyonu, tahmin edilen α değerinin gerçek değerin θ olması durumunda oluşacak hatayı ölçer. A=Θ ve kayıp fonksiyonunun kuadratik (ikinci dereceden) formu $L(\theta,\alpha) = (\theta - \alpha)^2$ olarak yazılabilir ve temeli Gauss'a kadar dayanan bu fonksiyon kolayca analitik olarak yorumlanabilir. Ayrıca kuadratik kayıp değeri ile risk fonksiyonu Eşitlik 2.17'de verildiği üzere iki parçaya ayrılabilmektedir (Parmigiani *et al*. 2009):

$$R(\theta,\delta) = \int_x L(\theta,\delta) f(\,x \mid \theta\,)\,dx = \int_x (\delta(x)-\theta)^2 f(\,x \mid \theta\,)\,dx$$
$$= \int_x [(\delta(x) - E[(\delta(x) \mid \theta] - \theta]^2 f(\,x \mid \theta\,)\,dx \qquad (2.17)$$
$$= Var[(\delta(x) \mid \theta] + \; [E[(\delta(x) \mid \; \theta] - \theta]^2$$

Ayrıştırmadaki ilk terim karar kuralının bir değişkenidir, ikinci terim ise yansızlığını ifade etmektedir.

Basitçe ifade edilecek olursa Bayes kuralı yardımı ile $\delta^*(x) = E(\theta \mid x)$, sonsal ortalama olarak ifade edilebilir. Bunun nedeni sonsal kaybın Eşitlik 2.18'de ki gibi olmasıdır.

$$L_x = \int_{\Theta} (\theta - \; \alpha)^2 \pi_x \,d\theta = \int_{\Theta} [(\theta - E[\theta \mid x]) + (\alpha - E[\theta \mid x])]^2\, \pi_x \,d\theta$$
$$= \int_{\Theta} (\theta - E[\theta \mid x])^2\, \pi_x \,d\theta + (\alpha - E[\theta \mid x])^2 \qquad (2.18)$$
$$= Var[\theta \mid x] + (\alpha - E[\theta \mid x])^2$$

Sonuç olarak kayıp fonksiyonu değeri $\alpha^* = E[\theta \mid x]$ değerini alarak en aza indirgenir ve α^* ile ilişkili sonsal beklenen kayıp θ 'nın sonsal varyansıdır. Kayıp fonksiyonu, $\pi(\theta \mid x)$ önsel dağılımının medyanının $\delta^*(x)$ şeklinde ifade edilmesi halinde; $L(\alpha,\theta) = \; |\theta - \alpha|$ şeklinde yazılabilir (Parmigiani *et al*. 2009).

Yaygın olarak kullanılan bir tahmin paradigması maksimum olabilirlik ilkesine dayanmaktadır. Bunu göstermek amacı ile θ parametrelerinden oluşan $\Theta=(\theta_1, \theta_2,\ldots)$ örnek uzayı ele alınacak olursa ilgili kayıp fonksiyonu Eşitlik 2.19'da verildiği gibi yazılabilir.

$$L(\alpha,\theta) = I_{[\alpha \neq \theta]} \qquad (2.19)$$

Burada α^* yine θ parametresi için nokta tahminini temsil eder ve sonsal beklenen kayıp $\alpha^* = mod(\,\theta \mid x\,) \equiv \theta^0$ denklemi ile minimize edilir (Parmigiani *et al*. 2009).

Kısacası Bayes teoremi için nokta tahmini,y gözlemler ve $\theta \in \Theta$ olmak üzere bir $\theta \in \Omega$ parametresi için önerilen bir $\widehat{\theta}$ tahmin edicisi iken, sonsal dağılımın olasılık yoğunluk fonksiyonu $p(\theta \mid y)$, kayıp fonksiyonu ise $L(\widehat{\theta}\ , \theta)$'dır. En iyi tahmin ise Eşitlik 2.20'de verilmiştir (Ekici 2005).

$$E\left(L(\widehat{\theta}, \theta)\right) = \int_{-\infty}^{+\infty} L(\widehat{\theta}\ , \theta)\ \ p(\theta \mid y)\, d\theta \tag{2.20}$$

2.1.4.5 Güven Aralığı

Anakitle parametresinin tahmini tek bir değer yerine bir değerler aralığı içinde verilmesi istenebilir. Belli bir güvenle bu aralığın tahmin edilmek istenen parametreyi içerdiği söylenebilir. $1-\alpha$ olarak ifade edilen güven düzeyi parametrenin gerçek değerinin $1-\alpha$ olasılıkla belirlenen aralık içinde olduğunu ifade eder. Bununla birlikte aralığın α olasılıkla parametreyi içermemesi de muhtemeldir. Bir anakitle katsayısının aralık tahmin edicisi, o anakitle katsayısının içine düşebileceği bir aralığı örneklem bilgisine dayanarak belirlemenin kuralıdır. Buna karşılık gelen tahmine de aralık tahmini denir (Güner 2014).

θ parametresi için Bayesci aralık tahminleri, klasik çıkarsamanın güven aralığı tahminine benzemektedir. Burada incelenen konu ölçmek istenilen bir aralık için sonsal dağılımın ortalama değerini içerip içermediğidir. Güven aralıkları frekansçı istatistiklerin bir aralığı bulmaya çalıştığı değerlerdir. Bayesci yaklaşım için güven aralığı ifadesini aşağıdaki şekilde ifade etmek mümkündür (Gosh *et al.* 2006):

$0 < \alpha < 1$ olmak üzere, θ parametresi için $(1-\alpha) \times \%100$ güven aralığı ve $C \in \Theta$ için Eşitlik 2.21'e göre hesaplanır.

$$P\{C|X = x\} = 1 - \alpha \tag{2.21}$$

Burada C bir aralık ifadesi olmak üzere, Θ parametre uzayının alt kümesi, $(1-\alpha) \times \%100$ ifadesi ise θ parametresi için hesaplanan rasgele aralıkların $\%100$ 'nün gerçek parametre değerlerini içerdiği anlamlarına gelmektedir. θ parametresi sürekli bir rasgele değişken olmak üzere, $\theta^{(1)}$ ve $\theta^{(2)}$ parametreleri için, $\alpha_1 \times \%100$ ve $(1 - \alpha_2) \times \%100$ olarak yazıldığında $\alpha_1 + \alpha_2 = \alpha$ ve $C = [\theta^{(1)}, \theta^{(2)}]$ olarak ifade edilebilir. Genel olarak eşit kuyruklu aralıklarda, $P(C \mid X = x) = 1 - \alpha$ olasılığında, $\alpha_1 + \alpha_2 = \alpha/2$ olarak alınabilir. θ parametresinin ayrık olması halinde sonsal olasılığın $(1 - \alpha)$ değeri için bir güven aralığı bulmak zordur. Bu durumda Eşitlik 2.22 kullanılmaktadır.

$$P(C \mid X = x) \geq 1 - \alpha \tag{2.22}$$

Bayesci yaklaşımda sonsal dağılım için istenilen aralığın değerlerine ulaşabilmek mümkündür. Burada bu aralık "En Yüksek Sonsal Yoğunluk" olarak adlandırılmaktadır. θ parametresi için sonsal yoğunluğun tekdüze olduğu varsayılarak, θ için "En Yüksek Sonsal Yoğunluk Aralığı" Eşitlik 2.23 yardımı ile hesaplanabilir.

$$C = \{\theta : \pi(\theta \mid X = x) \geq k\} \tag{2.23}$$

Genellikle, kullanılan tahmin edicinin örnekleme dağılımı ortalama olarak gerçek değere eşit, yaklaşık olarak normaldir. Örneğin önsel dağılımın normal ve σ^2 varyansının bilindiği bir normal dağılıma sahip örneklem için "En Yüksek Sonsal Yoğunluk Aralığı" Eşitlik 2.24'de verildiği gibidir.

$$\begin{gathered} C = sonsal\ dağılımın\ ortalaması \\ \pm z_{\alpha/2}.sonsal\ dağılımın\ standart\ sapması \end{gathered} \tag{2.24}$$

Burada $z_{\alpha/2}$ değeri normal dağılım için kritik değeri ifade etmektedir.

En yüksek sonsal yoğunluk aralığının, içerisindeki her noktanın olasılık yoğunluğu, dışındaki noktaların her birinden daha büyük olmasını sağlayacak şekilde ve belirli bir olasılık içeriği için mümkün olan en küçük hacme sahip olması gerekmektedir. Kısacası En yüksek sonsal yoğunluk aralığı, $P(\theta \mid x)$ sonsal dağılımı için θ ilgilenilen parametrenin k-boyutlu bir vektörü, y ise n boyutlu bir gözlem vektörü olmak üzere, θ parametresinin parametre uzayının belirli bir $\bar{R}$ alt bölgesinde bulunması olasılığı Eşitlik 2.25'e göre hesaplanır (Box and Tiao 1973).

$$P\{\theta \epsilon\, \bar{R} \mid y\} = \int_{\bar{R}} P(\theta \mid y)\, d\theta = 1 - \alpha \qquad (2.25)$$

Parametreler için güven aralıklarının hesaplanması konusunda klasik yaklaşım ile Bayesci yaklaşımı için oldukça farklı iki kavram mevcuttur. Klasik yaklaşım için güven aralığı, çok sayıda tekrarlanan denemeler sonucunda (1-α) anlamlılık düzeyinde bilinmeyen parametrenin gerçek değerinin bu aralık içerisinde olması olasılığını ifade eder. Bayesci yaklaşımda ise "güven aralığı" terimini yerine "En yüksek sonsal yoğunluk aralığı" terimi kullanılmaktadır ve bilinmeyen parametrenin gerçek değerinin direk olarak (1-α) anlamlılık düzeyinde bu aralıkta bulunması olasılığını ifade etmektedir. Bunun nedeni ise Bayesci yaklaşımda parametre değerinin bir rasgele değişken olarak görülmesidir (Kendall and Buckland 2009).

2.1.4.6 Hipotez Testi ve Bayes Faktörü

Anlamlılık testleri olarak da adlandırılan hipotez testlerinde araştırmacının ortaya attığı iddiaya göre ilgilenilen parametrenin belirlenen değere eşit olup olmadığı veya istenilen aralıkta olup olmadığı test edilir. H_0 Hipotezi, yokluk hipotezi olarak adlandırılmaktadır. Buna karşın kurulan H_1 hipotezi ise alternatif hipotez olarak adlandırılmaktadır (Gosh *et al*. 2006).

Tek yönlü hipotez testinde yokluk hipotezi, alternatif hipotezin sonsal olasılığına karşı test edilir. Bayesci yaklaşımda öncelikle tek yönlü etkiyi test etmek için kullanılan tek taraflı hipotez testini kullanarak Bayesci yöntem ile α-anlamlılık düzeyinde elde edilecek gösterim Eşitlik 2.26'da belirtilmiştir. (Gosh *et al*. 2006).

$$H_0\colon \theta \in \Theta_0 \text{ yokluk hipotezine karşı,} H_1\colon \theta \in \Theta_1 \text{ alternatif hipotezi.} \qquad (2.26)$$

Eğer Θ_0 ve Θ_1 parametre uzayları yokluk ve alternatif hipotezleri ile aynı boyutta ise, önsel olasılık yoğunluk fonksiyonunu seçmek ve belirlemek kolaydır. Önsel olasılıklar belirlendikten sonra, sonsal olasılıkların yanı sıra sonsal olasılıkların oranı da Eşitlik 2.27 ile belirlenebilir (Gosh *et al*. 2006).

$$P\{\Theta_0 \mid x\} \big/ P\{\Theta_1 \mid x\} \tag{2.27}$$

Daha sonra burada H_0 hipotezine karşı neyin kanıt oluşturduğuna karar vermek için bir olasılık veya aralık sınırı koyulur. Bayesci kurala göre belirlenecek bu değerlerin "0 – 1" arasında değişen değerler alması beklenir. Bayesci yaklaşımda elde edilen sonsal dağılımların sahip oldukları olasılıklar ile karşılaştırma yapılmaktadır, karar verilirken son olasılık değeri fazla olan hipotez, doğru kabul edilecek olan hipotezdir. Ayrıca elde edile bu sonsal olasılık değeri ∝-anlamlılık seviyesinden düşük ise bu yokluk hipotezi red edilir (Gosh *et al*. 2006).

Hipotez testinde en önemli kısım önsel olasılıkları belirleyebilmektir. Eğer önsel olasılıklar yanlış belirlenmiş ise sonsal olasılıkların tayini de yanlış olacaktır. Tarafsız ve nesnel bir seçim yapabilmek için Θ_0 ve Θ_1 'e eşit olasılıklar verilmelidir. Önsel dağılımı belirtmek için aşağıdaki alternatif yöntem kullanılırsa daha iyi bir sonuç alınabilir (Gosh *et al*. 2006).

π_0 ve $1 - \pi_0$, Θ_0 ve Θ_1 için önsel olasılık değerleri olsun. θ parametresi için Θ_i altında $g_i(\theta)$ önsel olasılık ifadesi olmak üzere Eşitlik 2.28 ve Eşitlik 2.29 yazılabilmektedir (Gosh *et al*. 2006).

$$\int_{\Theta_i} g_i(\theta)\, d\theta = 1 \tag{2.28}$$

$$\pi(\theta) = \pi_0 g_0(\theta) \big/ \{\theta \in \Theta_0\} + (1 - \pi_0) g_1(\theta) \big/ \{\theta \in \Theta_1\} \tag{2.29}$$

Böylelikle Θ_0 ve Θ_1 'in aynı boyutta olmasına gerek duyulmaz. Daha sonra verilen bir X gözleminin π önsel dağılımı altında olasılık yoğunluk fonksiyonu Eşitlik 2.30 ile hesaplanabilir (Gosh *et al*. 2006):

$$\begin{aligned} m_\pi(x) = \int_{\Theta} f(x \mid \theta)\pi(\theta)\, d\theta = \pi_0 \int_{\Theta_0} f(x \mid \theta) g_0(\theta)\, d\theta + 1 \\ - \pi_0 \int_{\Theta_1} f(x \mid \theta) g_1(\theta)\, d\theta \end{aligned} \tag{2.30}$$

ve böylece $X = x$ olmak üzere θ parametresi için sonsal olasılık yoğunluk fonksiyonu:

$$\pi(\theta \mid x) = \frac{f(x \mid \theta)\pi(\theta)}{m_\pi(x)} \tag{2.31}$$

Eşitlik 2.31 ile elde edilir. Eşitlik 2.31 yardımı ile H_0 ve H_1hipotezleri için Eşitlik 2.32 ve Eşitlik 2.33 elde edilebilir.

$$\begin{aligned} P^\pi(H_0 \mid x) = P^\pi(\Theta_0 \mid x) &= \frac{\pi_0}{m_\pi(x)} \int_{\Theta_0} f(x \mid \theta) g_0(\theta)\, d\theta \\ &= \frac{\pi_0 \int_{\Theta_0} f(x \mid \theta) g_0(\theta)\, d\theta}{\pi_0 \int_{\Theta_0} f(x \mid \theta) g_0(\theta)\, d\theta + 1 - \pi_0 \int_{\Theta_1} f(x \mid \theta) g_1(\theta)\, d\theta} \end{aligned} \tag{2.32}$$

ve

$$\begin{aligned} P^\pi(H_1 \mid x) = P^\pi(\Theta_1 \mid x) &= \frac{\pi_1}{m_\pi(x)} \int_{\Theta_1} f(x \mid \theta) g_1(\theta)\, d\theta \\ &= \frac{\pi_1 \int_{\Theta_1} f(x \mid \theta) g_1(\theta)\, d\theta}{\pi_1 \int_{\Theta_1} f(x \mid \theta) g_1(\theta)\, d\theta + 1 - \pi_0 \int_{\Theta_1} f(x \mid \theta) g_1(\theta)\, d\theta} \end{aligned} \tag{2.33}$$

Eşitlik 2.32 ve Eşitlik 2.33 ışığında H_1 'e göre H_0 'ın Bayes faktörü Eşitlik 2.34'de verildiği gibi yazılabilir.

$$BF_{01} = \frac{\int_{\Theta_0} f(x \mid \theta) g_0(\theta)\, d\theta}{\int_{\Theta_1} f(x \mid \theta) g_1(\theta)\, d\theta} \tag{2.34}$$

Açıkça görüleceği üzere, $BF_{10} = \frac{1}{BF_{01}}$ 'dir. Ayrıca H_0 'ın H_1 'e göre olasılık (odds) oranı ise Eşitlik 2.35'de verilmiştir.

$$\left(\frac{\pi_0}{1 - \pi_0}\right) BF_{01} \tag{2.35}$$

Böylece $\pi_0 = \pi_1 = \frac{1}{2}$ ise bilgi içermeyen önsel dağılım için bu oran Bayes Faktörüne eşittir ve BF_{01} değeri ne kadar küçükse H_0 hipotezini reddetme olasılığı o kadar artar. Sonuç olarak elde edilen bu oranda olasılık değeri yükseldikçe hipotezi red etme olasılığı azalmaktadır (Gosh *et al*. 2006).

2.1.5 Önsel Dağılım

Bayesci istatistikte örneklemden elde edilen bilginin yanı sıra önsel bilgi dağılımının da kullanılması klasik yöntemlere göre bazı farklılıklara ve problemlere sebep olmuştur. Bunlardan en önemlisi parametre fonksiyonunun önsel bilgi dağılımının seçimini etkilediğidir. Bayesci istatistikte kullanılan önsel bilgi, araştırmayı yapan kişinin sahip olduğu bilgi miktarına göre üç farklı gruba ayrılmıştır. Bunlar, eşlenik (conjugated) önsel dağılım, bilgi vermeyen (noninformative) önsel dağılım ve sübjektif (bilgi içeren) önsel dağılımlardır. Bilgi vermeyen önsel dağılımı kullanılması durumunda yapılan analizler klasik yaklaşım ile yapılan analizlere göre benzer sonuçlar vermektedir. Fakat sübjektif ve eşlenik önsel dağılımın kullanılması halinde ortaya çıkan sonuçlar klasik yaklaşımdan farklılık göstermektedir (Akar ve Gündoğdu 2014).

Bayesci yaklaşımda önsel dağılım yardımı ile sonsal dağılım fonksiyonu belirlendiğinden, önsel dağılımın belirlenmesi ve kullanılması önemli konudur. Önsel dağılımın olmadığı bir durumda Bayesci çıkarım ve istatistiksel modelleme yapılamaz. Önsel dağılımın sınıflandırılmasında diğer bir kısıt ise, kesikli veya sürekli dağılımlarda integral veya toplam olasılık değerinin “1” değerine eşit olup olmamasıdır. İntegral veya toplam olasılık değerinin “1” değerine ulaşması halinde bu tür önsel olasılıklar tam (proper) önsel olasılık, ulaşılmaması halinde tam olmayan (improper) önsel olasılık olarak adlandırılırlar (Karadağ 2011).

Bayesci çıkarsamanın en önemli özelliklerinden birisi de örneklem hacmi arttıkça önsel dağılımın sonsal dağılım üzerindeki etkisinin azalmasıdır. Kısaca özetlenecek olursa, örneklem hacminin değişmesi olabilirlik fonksiyonunun basıklık veya sivrilik özelliğini değiştirebilmektedir. Örneğin; olabilirlik fonksiyonu sivri ve önsel dağılım fonksiyonu basık ise sonsal dağılım fonksiyonu da sivri olmaktadır. Sonuç olarak önsel dağılımın sonsal dağılıma etkisi olmamakla birlikte Klasik yaklaşıma benzer sonuçlar vermektedir (Box and Tiao 1973).

2.1.5.1 Bilgi Vermeyen Önsel Dağılım

Bilgi vermeyen önsel dağılım parametre hakkında çok az bilgiyi açıklamaktadır (Saçaklı 2011). Bilgi içermeyen önsel dağılım, yapılan analiz çalışmalarında en çok kullanılan yöntemlerden biridir. Bu yöntem, kullanılan parametrenin tanımlı olduğu aralık dışında başka hiçbir bilginin olmaması ve bu parametre için elde edilen önsel bilginin sonsal dağılımı üzerinde en az etkiye sahip olması durumunda kullanılmaktadır (Avcı 2012, Zerey 2018). Bilgi içermeyen önsel dağılımlara örnek olarak:

- Düzgün (üniform) önsel dağılımlar
- Belirsiz (diffuse/vague) önsel dağılımlar
- Jeffrey’ in önsel dağılımı

verilebilir (Murat 2012).

Bilgi içermeyen önsel dağılımların hemen hemen hepsi tanım bölgesindeki integralin “1” e eşit olması koşulunu sağlamaz. Yapılan uygulamalarda önsel dağılım tam olmasa bile sonsal dağılımlar her zaman tam olarak elde edilir. Genel olarak bilgi içermeyen önsel dağılım Eşitlik 2.36’da belirtildiği üzere orantısal formül ile tanımlanabilir.

$$f(\theta) \propto sabit \tag{2.36}$$

Parametreye ait önsel bilginin tanımlanmasında kullanılan Eşitlik 2.36’nın, elde edilecek olan sonsal bilginin olabilirlik (likelihood) fonksiyonuna orantılı olacağı anlamına gelir. Daha önce belirtilen $f(\theta \mid y) \propto f(y \mid \theta)\pi(\theta)$ eşitliği Bayes formülünden yola çıkılarak Eşitlik 2.37’da belirtildiği şekilde yeniden yazılabilir.

$$f(\theta \mid y) \propto f(\theta) \times f(y \mid \theta) \tag{2.37}$$

Burada,

$$f(\theta) = a \qquad a \epsilon A \tag{2.38}$$

şeklinde tanımlandığında, bilgi içermeyen önsel dağılımlar parametrenin tanımlı olduğu A aralığı bilgisi dışında hiçbir bilgiye sahip olunmadığı durumda kullanılan önsel dağılımlardır. Bu tür bir önsel dağılım kullanıldığında sonsal dağılım ile orantısal olarak olabilirlik fonksiyonunun a katına eşit olmaktadır ve Eşitlik 2.39'da belirtilmiştir:

$$f(\theta \mid y) \propto a \times f(y \mid \theta) \tag{2.39}$$

Bu durumda θ parametresinin tanım aralığı belirtilen A aralığının dışına çıkmamaktadır. Böylece bilgi içermeyen önsel dağılımların kullanılması halinde elde edilen sonuçlar klasik yaklaşıma göre benzer sonuçlar vermektedir. Ayrıca bu tür önsel dağılımlarda eğer A aralığının sınırlarından en az bir tanesi sonsuza gidiyorsa normalleştirme katsayısı Eşitlik 2.40'da verilen integrale eşit olur ve hesaplanması mümkün olmaz.

$$f(y) = \int f(\theta)\, d\theta = \infty \tag{2.40}$$

Bu durumda elde edilen bilgi içermeyen önsel dağılımda tam olmayan (improper) dağılım olmaktadır.

Bilgi içermeyen önsel dağılımlarda A aralığının sınırlarının herhangi birinin sonsuza gitmemesi durumunda ise oluşacak sonsal dağılım Eşitlik 2.41'de belirtilmiştir (Karadağ 2011, Sunar 2009, Box and Tiao 1973).

$$f(\theta \mid y) \sim N(\theta, \vartheta^2 (X'X)^{-1}) \tag{2.41}$$

Düzgün (üniform) önsel dağılımlar: Yapılan bir araştırmada θ parametresine ait az veya hiçbir önsel bilgi bulunmuyorsa Laplace-Bayes önermesi kullanılır. Bu önerme Eşitlik 2.42'de verildiği gibidir (Murat 2012):

$P(\theta \mid Y, M) \propto L$ ve $\theta \epsilon R$ olmak üzere:

$$P(\theta \mid Y, M) \propto P(Y \mid \theta, M) P(\theta) \tag{2.42}$$

Düzgün önsel dağılımlarda belirlenen aralıkta θ parametresine aynı olasılık değerleri atanır (Gündoğdu 2016).

Jeffrey ‘in Önsel Dağılımı: Bilgi vermeyen önsel dağılım oluşturulabilmesi amacı ile H. Jeffrey tarafından 1961 yılında elde edilen bu teorem, Fisher’in bilgi matrisinin determinantının karekökü alınarak elde edilmektedir. Bu teoreme göre Eşitlik 2.43 yazılabilir.

$$f(\theta) \propto \left|J(\theta)\right|^{1/2} \tag{2.43}$$

Bir tahmin edicinin, en çok olabilirlik tahmin edicisinin duyarlılığını ölçen Fisher bilgi matrisi ise Eşitlik 2.44’de verildiği gibi tanımlanabilir.

$$J(\theta) = -\int \left[\frac{d^2 lof(y \mid \theta)}{d\theta^2}\right] f(y \mid \theta)\, dy \tag{2.44}$$

Elde edilen Eşitlik 2.44 yardımı ile herhangi bir bilgi vermeyen önsel dağılım elde edilebilir (Gill 2008; Karadağ 2011). Farklı dağılımlar için elde edilen bazı Jeffrey önsel dağılımları Çizelge 2.2’de verilmiştir.

Çizelge 2.2 Farklı Dağılımlar İçin Jeffrey Önsel Dağılımları **(**Gill 2008; Karadağ 2011).

Olabilirlik Fonksiyonu	**Parametre(ler)**	**Önsel Dağılım**
Bernoulli	$\theta = p$	$(p(1-p)^{-1/2}$
Binom	$\theta = p$	$(p(1-p)^{-1/2}$
Poisson	$\theta = \lambda$	$\lambda^{-1/2}$
Normal (σ^2 biliniyor)	$\theta = \mu$	1
Normal (μ biliniyor)	$\theta = \sigma^2$	σ^{-1}
Normal	$\theta = \pi, \sigma^2$	σ^{-1}

Belirsiz önsel dağılım: Parametrelerin sahip olduğu dağılım çeşidine göre bu parametrenin ölçüsünün şeklini veren parametre değişmektedir. Kısacası parametre şekil parametresine sahipse Gamma dağılımı, konum parametresine sahipse Normal dağılım alınabilir. Büyük varyansın belirsiz durumla eş olması fikrine dayanarak şekil parametrelerine uygun değerler atanır ve düzgün önsel dağılıma uyan ve geniş bir aralıkta yer alan bilgi içermeyen bir önsel dağılım oluşturulur.

- $p(\theta) \sim N(\mu, \sigma^2)$ olması durumunda σ^2 (varyans) değerine oldukça büyük değerler atanır.
- $p(\theta) \sim G(\alpha, \beta)$ olması durumunda β parametresine oldukça küçük değerler atanır.

Bu şekilde oluşturulan belirsiz önsel dağılımın en önemli özelliği uygun (proper) dağılım olmasıdır (Öztürk 2014).

2.1.5.2 Eşlenik Önsel Dağılımlar

Bayesci çıkarsamada zorluk yaşanan konulardan biri de bilinmeyen parametrenin sonsal olasılık dağılımının analitik olarak çözülebilmesidir. Çıkarsama yapılırken matematiksel işlem zorlukları nedeni ile istatistikçiler “Eşlenik Önsel Dağılım” kavramını geliştirmişlerdir. Önsel olasılık dağılımı ile olabilirlik fonksiyonunun eşlenik olması halinde sonsal olasılık dağılımı da kolaylıkla elde edilebilir. Bu dağılım önsel olasılık dağılımı ile ilk olasılık dağılımının aynı dağılım ailesine sahip olmasını sağlamakta ve araştırmacılara uygulama kolaylığı sağlamaktadır (Gill 2008; Saçaklı 2011). Karadağ (2011) tarafından oluşturulan aynı aileden olan eşlenik dağılımlara ait bilgiler Çizelge 2.3’de verilmiştir.

Çizelge 2.3 Eşlenik Önsel ve Sonsal Dağılımlar

Olabilirlik Fonksiyonu	Parametre(ler)	Önsel Dağılım	Sonsal Dağılım
Bernoulli	p	Beta	Beta
Binom	p	Beta	Beta

Çizelge 2.3 (Devam) Eşlenik Önsel ve Sonsal Dağılımlar

Poisson	λ	Gamma	Gamma
Üstel	λ	Gamma	Gamma
Normal (σ^2 biliniyor)	μ	Normal	Normal
Normal (μ biliniyor)	σ^2	Ters Gamma	Ters Gamma
Normal	μ, σ^2	Normal-Gamma	Normal-Gamma
Tek Biçimli	a,b	Pareto	Pareto
Gamma	α,β	Gamma	Gamma

2.1.5.3 Sübjektif Önsel Dağılım

Sübjektif (bilgi veren) önsel dağılım olabilirlik fonksiyonuna etkisi olmayan fakat sonsal dağılımı etkileyen bir dağılımdır. Araştırmacının önceki çalışmalardan edinmiş olduğu bilgi, önceki deneyimler Bayesci çıkarsamanın kanıtlanabilirliğini ve gücünü artırır (Öztürk 2014). Bilgi veren önsel dağılımın oluşturulmasında kullanılan bilgiler, önceden elde edilmiş verilerden, araştırmacının araştırmaya dair önceki bilgilerinden ve uzman bir görüşün yapacağı değerlendirmelerle elde edilen bilgiler ile oluşturup elde bulunan mevcut veriler ile birleştirilerek güncellenir (Zerey 2018).

2.1.6 Sonsal Dağılım

Bayesci istatistiksel çıkarsamada önsel dağılımın belirlemesi ne kadar önem arz ediyorsa, sonsal dağılımın belirlenmesi de o kadar önemlidir. Yapılan araştırmalarda gözlemlenen en önemli problemlerden biri sonsal dağılımın belirlenmesinde gerekli olan integrallerin çözümünde analitik çözüm yoluna ulaşılamaması veya seçilen önsel dağılımların eşlenik aileye mensup olmamasıdır. Analizlerin daha kolay bir hale gelmesi ve çözüme ulaştırılabilmesi için Markov Zinciri Monte Carlo (MCMC) metodu geliştirilerek, kullanımı Bayesci istatistikçiler tarafından yaygın hale gelmiştir (Cengiz vd. 2012).

Markov Zinciri Monte Carlo (MCMC) Yöntemi: MCMC yöntemi, belirlenen hedef dağılım olarak adlandırılan belirli bir olasılık dağılımından örnekleme yapmak için uygulanan bir simülasyon yöntemidir. Markov zincirleri değişmeyen dağılımın hedef dağılım olduğu özelliği üzerine inşa edilmiş bir yöntemdir. MCMC yöntemleri Metropolis-Hastings algoritması ve Gibbs örnekleme algoritmasıdır. Bu iki algoritmanın karışımı ile Markov zincirleri elde edilebilir ve bu zincir çok sayıda simüle edilerek hedef dağılımın bir korelasyonlu örneği elde edilebilir. Böylece MCMC simülasyonu yüksek boyutlu örnekler elde etmede oldukça kolaylık sağlar (Chib 2001).

Bayesci çıkarsamada önsel dağılın belirlenmesinde eşlenik önsel dağılımı kullanılması halinde sonsal dağılımın standart dağılım olmaması halinde hesaplanmak istenen parametre değerleri bulunamayacaktır. Bu durumda Bayesci çıkarımda parametre tahminleri için simülasyon tekniklerinin kullanılması gerekmektedir. Bayesci çıkarsamada Markov zincirinin kullanılması halinde Monte Carlo integrasyonunun da kullanılması gerekmektedir (Güner 2014).

Markov zinciri $t = (1,2,\ldots,n)$ boyutlu $\theta^{\,t-1}$ önceki rasgele değişkenine bağlı $\theta^{\,t}$ rasgele değişkenlerinin oluşturduğu bir dizidir. Belirlenen bir $f(\theta)$ karmaşık hedef fonksiyonunun Monte Carlo integrasyonu $E[f(\theta)]$ beklenen değerine bağlıdır ve Eşitlik 2.45'de belirtilmiştir.

$$E[f(\theta)] = \int_S g(\theta)p(\theta)\,d\theta \tag{2.45}$$

Fakat buradaki problem belirlenen bu karmaşık $f(\theta)$ karmaşık hedef fonksiyonunun integralinin hesaplanmasındaki zorluktur. Sonuç olarak sonsal dağılımın elde edilebilmesi için $f(\theta)$ olasılık yoğunluk fonksiyonundan, MCMC yöntemi temeline dayanarak, örnekler çekilirse Monte Carlo integrasyonu tamamlanır ve Eşitlik 2.46'daki gibi yaklaşık değer:

$$E[f(\theta)] = \int_S g(\theta)p(\theta)\,d\theta \cong \frac{1}{n}\sum_{t=1}^{n} g(\theta)^t\, E[f(\theta)] = \int_S g(\theta)p(\theta)\,d\theta \tag{2.46}$$

elde edilir. Burada $g(\theta)$, θ rasgele değişkeninin ilgilenilen bir fonksiyonu, $p(\theta)$ θ' nın bir olasılık yoğunluk fonksiyonu ve θ^{t} ise belirlenen S aralığında $p(\theta)$ olasılık yoğunluk fonksiyonundan çekilen örneklerdir. Sonuç olarak MCMC yöntemi ile ana kitleden çekilen örneklem değerlerinin ortalamasının hesaplanması ile karmaşık $f(\theta)$ hedef fonksiyonunun beklenen değeri hesaplanmış olur. Bu yöntem ile Bayesci çıkarsamada sonsal dağılımın değeri yaklaşık olarak bulunabilir (Gilks *et al.* 1996, Gasım 2013).

a) Gibbs Örnekleme Algoritması

Gibbs örnekleme algoritması adını Amerikan fizikçi W.Gibbs' den sonra çalışmalarında yer veren Geman ve Geman (1984) tarafından geliştirilmiştir. Bu algoritma Metropolis-Hastings algoritmasının özel bir durumudur (Gilks *et al.* 1996). Gibbs örnekleme algoritması her örnekleme modelindeki her bir parametre değeri için koşullu olasılık dağılımları ile ortak sonsal dağılımın ayrıştırılmasını gerektirir. Gibbs örneklemesinde, parametrelerin birbirlerine çok bağımlı olmadığı durumlarda koşullu dağılımların tamamını örneklemek daha kolay olacaktır. Ayrıca Gibbs örneklemesi, Metropolis-Hastings algoritmasında olduğu gibi yardımcı dağılımlar kullanmadığından araştırmacılar tarafından daha çok tercih edilir durumdadır (SAS 2009).

Gibbs örneklemesindeki temel amaç, koşullu ilişkiler yardımı ile tahmin sürecini periyodik olarak çalıştırarak, parametrelerin birleşik dağılımlarına ulaşabilmektir (Demirci 2016). Algoritmanın işleyiş süreci ise aşağıdaki gibi ifade edilmektedir.

$\theta = (\theta_1, \theta_2, \dots, \theta_k)'$ parametre vektörü, $p(y \mid \theta)$ olabilirlik fonksiyonu ve $\pi(\theta)$ önsel dağılım olmak üzere, $\pi(\theta_i \mid \theta_j, i \neq j, y)$' nin sonsal koşullu dağılımının tamamı sonsal olasılık yoğunluk fonksiyonu ile orantılı olacaktır. Buradan Eşitlik 2.47 elde edilebilir.

$$\pi(\theta_i \mid \theta_j, i \neq j, y) \propto p(y \mid \theta)\pi(\theta) \qquad (2.47)$$

Gibbs örneklemesi yapılırken iterasyonlar kullanılır ve son elde edilen parametrelerden koşullu örneklem çekimi yapılarak bir döngü oluşturulur (Saçaklı 2011). Bu döngü aşağıdaki gibi tanımlanabilir:

1. $t = 0$ olmak üzere başlangıç değeri için keyfi bir θ^0 $=(\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_k^{(0)})$ değeri alınır.
2. Her bir θ bileşeni aşağıdaki gibi oluşturulur;

   $\pi(\theta_1 \mid \theta_2^{(t)}, \dots, \theta_k^{(t)}, y)$ ‘den $\theta_1^{(t+1)}$

   $\pi(\theta_2 \mid \theta_1^{(t)}, \theta_3^{(t)} \dots, \theta_k^{(t)}, y)$ ‘den $\theta_2^{(t+1)}$

   $\vdots$

   $\pi(\theta_k \mid \theta_1^{(t)}, \dots, \theta_{k-1}^{(t+1)}, y)$ ‘den $\theta_k^{(t+1)}$

   elde edilir.
3. $t = t + 1$ alınır ve $t < T$ (istenilen örneklem örneklem büyüklüğü) ise 2. adıma gidilir. Aksi takdirde döngü tamamlanır (SAS 2009, Öztürk 2014).

Gibbs örnekleme algoritmasının tam olarak çalışması için gerekli olan önemli noktalar vardır. Bunlar burn-in (yanma) periyodu, iterasyon sayısı, Markov Zinciri örnek sayısı ve başlangıç değeridir.

i. Burn-in periyodu: Bir Markov zinciri örnekleminin başlangıç bölümünün çıkarılmasıdır. Böylece başlangıç değerinin sonsal dağılım değeri üzerindeki etkisi en aza indirilmiş olur. Örneğin, hedef dağılımın $N(0,1)$ olduğu ve Monte Carlo başlangıç değerinin 10^6 değeri ile başlatıldığı varsayılsın. Zincir, birkaç yinelemede “0” değerine yaklaşacaktır. Bununla birlikte sonsal ortalama hesaplamasında 10^6 değeri etrafındaki örneklerin dahil edilmesi ortalama tahmininde önemli bir yanlılığa neden olacaktır. Markov zinciri sonsuz bir süre boyunca çalıştırılırsa başlangıç değerinin etkisi sıfıra düşer. Uygulamalarda sonsuz sayıda örneklem çekme mümkün olmadığından t tekrar sayısından sonra zincir hedef dağılıma ulaşır ve başlangıç bölümü atılarak sonsal çıkarım için en uygun örnekler kullanılır. Burada t değeri en uygun yanma sayısıdır (SAS 2009).

ii. Çoğu araştırmada sonsal dağılımın en uyun şekilde belirlenebilmesi için 1000 iterasyon gereklidir (Alkan 2012).

iii. Araştırmalarda genellikle Marcov zinciri başlangıç değeri için en çok olabilirlik tahmin değeri kullanılır (Alkan 2012).

b) Metropolis-Hastings Algoritması

Bu algoritma, N. C. Metropolis tarafından geliştirilmiştir.. Bu algoritma araştırmacıya oldukça basit ve pratik bir çözüm sağlamaktadır. Metropolis-Hastings algoritması normalleştirme sabiti değeri olarak bilinen herhangi bir boyuttaki keyfi kompleks bir dağılımdan rasgele örnekler üretmek için kullanılmaktadır.

$f(\theta \mid y)$ tek değişkenli bir olasılık yoğunluk fonkiyonu ve f' den elde edilmek istenen t.inci örnek θ^t olmak üzere Metrapolis-Hastings algoritmasını kullanabilmek için başlangıç değeri olarak θ^0 alınır. $q(\theta_{yeni} \mid \theta^t)$ simetrik bir dağılım kabul edilsin. θ^t 'den θ_{yeni} elde edebilmenin olabilirliği ile θ_{yeni}'den geriye doğru θ^t elde etmenin olabilirliği aynı olmalıdır. Kısacası elde edilmek istenen eşitlik $q(\theta_{yeni} \mid \theta^t) = q(\theta^t \mid \theta_{yeni})$'dir. Bu algoritma aşağıdaki gibi özetlenebilir:

1. $f(\theta^0) > 0$ koşulunu sağlayan $t = 0$ başlangıç noktası ve θ^0 alınır.
2. Önerilen dağılım $q(. \mid \theta^t)$ kullanılarak θ_{yeni} elde edilir.
3. Bir $r = min\left\{\frac{f(\theta_{yeni}|y)}{\theta^t|y}\right\}$ noktası seçilir.
4. $U(0,1)$ dağılımından bir u üretilir.
5. $u < r$ olması halinde $\theta^{t+1} = \theta_{yeni}$ olur, aksi takdirde $\theta^{t+1} = \theta^t$ olarak alınır.
6. $t = t + 1$ alınması halinde $t < T$ (istenilen örnek sayısı) 2. adıma tekrar dönülür. Aksi takdirde işlem tamamlanmış olur (Zerey 2018).

2.1.7 Bayesci Yaklaşım İle Klasik Yaklaşım Arasındaki Kavramsal Farklılıklar

Bayesci yaklaşım, klasik yönteme göre kavramsal olarak çıkarım yapmak için kullanılan daha basit bir yöntemdir ve Bayesci bakış açısı klasik bakış açısına göre bir dizi avantaj sunar (Bolstad 2007). Bu avantajlardan bazıları:

- Klasik yaklaşım tümdengelim yöntemi ile paralellikler gösterirken, Bayesci yaklaşım tümevarım yöntemiyle paralellik gösterir. Ayrıca her zaman bu ayrım pek net olmamakla birlikte klasik yaklaşım nedensellik ilkesinin deterministik yorumuna yakın görünürken, Bayesci yaklaşım olasılıklı yorumuna yakındır (Ekici 2009).

- Bayesci istatistik süreç hakkında önceden edinilen bilgilerin (prior) de kullanılmasını sağlar. Bu bilginin kullanımı da süreç hakkında oluşacak bilgi kaybını önler (Bolstad 2007).

- Bayesci istatistik "sübjektiflik" ilkesine dayanırken, klasik istatistik "objektiflik" ilkesine dayanır. Bayes kuramı süreç hakkında edinilen hiçbir bilgiyi göz ardı etmediği için sübjektif olarak değerlendirilir (Bolstad 2007).

- Klasik istatistikte bilinmeyen parametreler tahmin edilirken genellikle en çok olabilirlik ya da en küçük kareler yöntemleri kullanılır. Bayesci istatistik, hipotez testlerine, nokta ve güven aralığı tahminlerine bir seçenek sunar. Bayesci istatistiğin temelinde konu ile ilgili tüm bilgilerin çözümlemeye katılması varsayımı yer almaktadır ve Bayesci uygulamalarda ilgilenilen parametre için çıkarsama yapmak temel amaçtır (Güner 2014).

- Klasik yaklaşımda aralık kestirimlerinin yorumu aralığın parametreyi içermesi olasılığı üzerine yapılırken, Bayesci yaklaşımda bu yorum parametrenin aralığa düşmesi olasılığı ile ilgilidir. Bayesci yaklaşımda aralığın içinde bulunan her noktanın sahip olduğu olasılık yoğunluğu, aralık dışında bulunanlardan daha büyük olduğunda, bu aralığa en yüksek sonsal yoğunluk aralığı adı verilir (Güner 2014).

- Bayes teoremi klasik yaklaşımlarla kestirimi yapılamayan modellerde kullanılabilme imkânı sunar. Ayıca diğer bir özelliği ise güven aralığından olasılık hesaplayabilmesidir (Yurtçu 2018).

- Klasik yaklaşımda parametre için bir aralık tahmin yapılırken, Bayesci yaklaşımda sonsal dağılım yardımıyla parametre tahmininin olasılığı elde edilir. Bayesci yaklaşımda tahminin bir aralık olarak değil doğrudan hesaplanabilmesi uygulamada avantaj sağlar (Avcı 2012).

- Klasik yaklaşımda karmaşık modellerin parametrelerini tahmin etmek oldukça zordur. Bayesci yaklaşım ise parametre tahmininde marjinal sonsal dağılımı kullandığından, parametre tahminlerini elde etmek daha kolaydır (Avcı 2012).

- Klasik yaklaşımda maliyet ve zaman tasarrufu sağlamak için küçük örneklemler tercih edilir ve bu nedenle parametre tahminlerinin güvenilirliği yitirilebilir. Ancak Bayesci yaklaşımda küçük örneklemler ile güvenilir sonuçlar elde etmek mümkündür(Avcı 2012).

- Bayesci yaklaşımda farklı zaman dilimlerinde toplanmış veriler, parametrenin sonsal dağılımının oluşturulmasında kullanılabilir. Bu durum yaklaşımın esnekliği konusunda araştırmacıya yardımcı olur (Avcı 2012).

Bu iki yaklaşım arasındaki temel farklılıklar, Carlin ve Louis (2000), Ekici (2005),Gill (2002), Turanlı ve Güriş (2005) ve Wille (2003) tarafından yapılan çalışmalar Uludağ (2013) tarafından derlenerek Çizelge 2.4'de kısaca özetlenmiştir.

Çizelge 2.4 Bayes yaklaşımı ile klasik yaklaşımın farklılaştığı temel konular

BAYES YAKLAŞIMI	**KLASİK YAKLAŞIM**
Temel Yaklaşım Farklılıkları	
Tümevarım yönteminin genel mantığını ele alır.	Tümdengelim yönteminin mantığını ele alır.
Nedensellik ilkesinin olasılıklı yorumuna uygundur.	Nedensellik ilkesinin belirleyici yorumuna uygundur.
Olasılığın öznel tanımını benimser.	Olasılığın nesnel tanımını benimser.
Doğrulama yolu ile "1" olasılığına ulaşmayı esas alır.	Yanlışlama yolu ile "0" olasılığına ulaşmayı esas alır.

Çizelge 2.4 (Devam) Bayes yaklaşımı ile klasik yaklaşımın farklılaştığı temel konular

Karar Sürecine Bakış Açısı Farklılıkları	
Karar sürecinde belirsizlik kavramı, olasılık kavramına dayanarak açıklanmaya çalışılmaktadır.	Karar sürecinde belirsizlik kavramı, belirleyici bakış açısı çerçevesinde açıklanmaya çalışılmaktadır.
Yöntem ve Analiz Farklılıkları	
Parametre, olasılık dağılımına sahip bir değişkeni ifade eder.	Parametre, bilinmeyen bir sabiti ifade eder.
Son dağılıma ait aritmetik ortalama, parametrenin nokta tahminini ifade eder.	Parametrenin en iyi tahmini, denemeler sonucunda elde edilir.
Belirli bir aralığın, son dağılımın ortalamasını içermesi olasılığı ile ilgilenir. Bu güvenilir aralık gibi ifadelerle tanımlanır.	Belirli bir aralığın, parametrenin gerçek değerini içine alması ile ilgilenir. Bu güven aralığı ile ifade edilir.
Hipotezlere ilişkin olasılıkların incelenmesiyle, hipotezlerin karşılaştırılması söz konusudur.	Parametrelerin ilgili değer ya da aralıkta olup olmadıklarını incelenmesiyle, hipotezlerin test edilmesi söz konusudur.
Analizde uzman görüşü, benzer ya da geçmişteki veri vb. önsellerden yararlanılmaktadır.	Analizde uzman görüşü, benzer ya da geçmişteki veri vb. önsellerden yararlanılmamaktadır.

2.2 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 (Çiftci 2009).

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 (Çiftci 2009). Bu varsayımlar altında kurulan modelin parametre tahminlerine dair çeşitli yöntemler geliştirilmiştir. Bu yöntemler, en çok olabilirlik tahmin edicisi (EÇO), en küçük kareler tahmin edicisi (EKK), momentler tahin edicisi ve Bayesci tahmindir (Çevik 2009).

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 (Saraçlı 2008). Bu bölümde regresyon modelinin oluşturulmasında Tip I ve Tip II regresyon tekniği açıklanacak ve parametrelerin tahmin yöntemlerinde ise En Küçük Kareler tahmin edicisi ve Bayesci tahmin yöntemleri uygulanarak regresyon modellerinin oluşturulması gösterilecektir.

2.2.1 Tip I Regresyon Analizi

Basit doğrusal regresyon modeli bir bağımsız değişken (X) ve bir bağımlı değişken (Y) içerir ve bağımlı değişkenin gerçek ortalamasında, bağımsız değişkenin değeri arttıkça ve azaldıkça ne gibi değişimler olacağını ifade eder. Basit doğrusal regresyon (Tip I) modeli en basit hali ile Eşitlik 2.48'de ki gibi ifade edilebilir (Rawlings *et al*. 2001).

$$Y_i = \beta_0 + \beta_1 X_{\dot{\mathrm{I}}} + \varepsilon_{\dot{\mathrm{I}}} \qquad \dot{\mathrm{I}} = 1,2,\ldots,n \tag{2.48}$$

Eşitlik 2.48'e göre:

β_0: $Doğrunun\ y\ eksenini\ kestiği\ nokta, sabit\ katsayı, regresyon\ katsayısı$
β_1: $Doğrunun\ eğimi, eğim\ katsayısı$
ε_i: $Rassal\ hata\ terimidir.$

Burada β_0 ve $\beta_1$tahmini yapılacak olan parametrelerdir (Öztürk 2003). Ancak elde edilen regresyon denkleminin kestirim amaçlı kullanılabilmesi için ε_i ($\varepsilon_i = Y_i - \hat{Y}_i$) hata terimlerinin rassal olup, dağılımının $\varepsilon_i \sim N(0, \sigma^2)$,birbirinden bağımsız olması ($cov(\varepsilon_i . \varepsilon_j = 0$), bağımlı değişkenler ile hata terimlerinin arasında korelasyon olmaması gibi bazı varsayımları sağlaması gerekmektedir (Fox 1997; Alma ve Vupa 2008).

Basit Doğrusal Regresyon Parametrelerinin En Küçük Kareler Tahmin Edicisi: EKK Tahmin edicisi β_0 ve β_1 parametrelerinin tahmininde en sık kullanılan yöntemlerden biridir. Regresyon denkleminde yer alan β_0 ve β_1 parametrelerinin tahmin edicileri $\hat{\beta}_0$ ve $\hat{\beta}_1$ olmak üzere tahmin edilen lineer regresyon doğrusunun denklemi Eşitlik 2.49'da ki gibi olacaktır.

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i \tag{2.49}$$

EKK yönteminin temeli, gözlem değerlerinin regresyon doğrusuna olan uzaklıklarının karelerinin toplamını en küçük yapmaktır. $\hat{\beta}_0$ ve $\hat{\beta}_1$ parametre tahmin değerleri için $L = \sum \varepsilon_i{}^2$ değerini minimum yapan değer, Eşitlik 2.50-2.59 izlenerek elde edilebilir.

$$\varepsilon_i = Y_i - (\hat{\beta}_0 + \hat{\beta}_1 X_i) \tag{2.50}$$

ve

$$\varepsilon_i = Y_i - \hat{Y}_i \qquad\qquad i = 1,2, \dots, n \tag{2.51}$$

olmak üzere her iki tarafın karesi alınır:

$$L = \sum_{i=1}^{n} \varepsilon_i{}^2 = \sum_{i=1}^{n} \left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_\text{İ}\right)^2 \tag{2.52}$$

$\hat{\beta}_0$ ve $\hat{\beta}_1$ parametre değerleri için Eşitlik 2.53 ve 2.54'de verildiği üzere L ' nin öncelikle $\hat{\beta}_0$ daha sonra $\hat{\beta}_1$ parametresine göre türevi alınır ve "0" değeri için eşitlenir.

$$\frac{dL}{d\hat{\beta}_0} = 2\sum_{i=1}^{n} \left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_\text{İ}\right)(-1) = 0 \tag{2.53}$$

$$\frac{dL}{d\hat{\beta}_1} = 2\sum_{i=1}^{n} \left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_\text{İ}\right)(-X_\text{İ}) = 0 \tag{2.54}$$

bu denklemler yardımı ile Eşitlik 2.55'de verilen değerler elde edilebilir.

$$\sum_{i=1}^{n} \left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_\text{İ}\right) = 0 \qquad ve \qquad \sum_{i=1}^{n} (X_\text{İ} Y_i - \hat{\beta}_0 X_\text{İ} - \hat{\beta}_1 X_\text{İ}{}^2 = 0 \tag{2.55}$$

Buradan,

$$n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} X_\text{İ} = \sum_{i=1}^{n} Y_i \qquad ve \quad \hat{\beta}_0 \sum_{i=1}^{n} X_\text{İ} + \hat{\beta}_1 \sum_{i=1}^{n} X_\text{İ}{}^2 = \sum_{i=1}^{n} X_\text{İ} Y_i \tag{2.56}$$

Eşitlik 2.56 yazılabilir ve bu iki denklemin çözümü ile tahmin edilen $\hat{\beta}_0$ ve $\hat{\beta}_1$ parametre değerleri Eşitlik 2.57 ile ifade edildiği gibi:

$$\hat{\beta}_1 = \frac{\sum (X_\text{İ} - \bar{X})(Y_i - \bar{Y})}{\sum (X_\text{İ} - \bar{X})^2} = \frac{\sum x_i y_i}{\sum x_i{}^2} \qquad \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} \tag{2.57}$$

ve $x_i = X_\text{İ} - \bar{X}$, $y_i = (Y_i - \bar{Y})$ alınırsa,

$$\sum x_i{}^2 = \sum X_\text{İ}{}^2 - \frac{\sum (X_\text{İ})^2}{n} \qquad \sum x_i y_i = \frac{\sum (X_\text{İ}) \sum (Y_i)}{n} \tag{2.58}$$

Eşitlik 2.58 için,

$$\hat{\beta}_1 = \frac{\sum X_{\mathrm{i}} Y_i - \frac{\sum(X_{\mathrm{i}}) \sum(Y_i)}{n}}{\sum X_{\mathrm{i}}{}^2 - \frac{\sum(X_{\mathrm{i}})^2}{n}} \qquad \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} \tag{2.59}$$

Eşitlik 2.59 elde edilebilir (Rawlings *et al*. 2001).

2.2.2 Tip II Regresyon

Tip I Regresyon tekniği ile kurulan regresyon modellerinde X bağımsız değişkeninin ölçümünün hatasız bir şekilde yapıldığı varsayılır. Bu tür regresyon modellerinde elde edilen hatanın Y bağımlı değişkenine ait ölçüm değerlerinden kaynaklandığı varsayılır. Burada elde edilmek istenen hata, Y gözlem değerleri ile tahmini yapılmak istenen regresyon doğrusu üzerinde yer alan Y tahmin değerleri arasındaki dikey uzaklığın karesi ile minimize edilerek hesaplanır (Saraçlı 2008).

Tip II Regresyon modelinin oluşturulmasında ise hem X bağımsız değişkeninin hem de Y bağımlı değişkeninin gözlem değerlerinin ölçülmesinde oluşabilecek hata dikkate alınır. Bu işlem kısaca şöyle açıklanabilir:

$$x_i = X_{\mathrm{i}} + \varepsilon_{\mathrm{i}} \tag{2.60}$$

$$y_i = Y_i + \delta_i \tag{2.61}$$

Eşitlik 2.60 ve 2.61'den de görüldüğü üzere gözlemlenen x_i ve y_i değerleri, X_{i} ve Y_i gerçek değerleri ile ε_{i}, δ_i hata değerlerini içererek hesaplanmışlardır (Linnet 1993).

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."

Tip II Regresyon teknikleri hakkında literatür bilgileri bu bölümde açıklanmıştır.

2.2.2.1 En Küçük Kareler (EKK) Açıortay 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 (Isobe *et al*. 1990). Saraçlı (2008) çalışmasında, Tip II Regresyon teknikleri arasında EKK-Açıortay tekniğinin diğer tekniklere göre daha iyi bir sonuç verdiğini ortaya koyulmuştur.

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 katsayılara ilişkin varyans ve kovaryanslar Eşitlik 2.62-2.65'te verildiği gibi hesaplanabilir (Isobe *et al*. 1990).

$$\hat{\beta}_{1AO} = (\hat{\beta}_1 + \hat{\beta}_2)^{-1}\left[\hat{\beta}_1\hat{\beta}_2 - 1 + \sqrt{\left(1+\hat{\beta}_1{}^2\right)\left(1+\hat{\beta}_2{}^2\right)}\right] \qquad (2.62)$$

$$\hat{\beta}_{0AO} = \bar{y} - \hat{\beta}_{0A1}\,\bar{x} \qquad (2.63)$$

$$Var\left(\hat{\beta}_{AO}\right) = \frac{\hat{\beta}_{AO}{}^2}{(\hat{\beta}_1+\hat{\beta}_2)^2\left(1+\hat{\beta}_1{}^2\right)\left(1+\hat{\beta}_2{}^2\right)}$$
$$\left[\left(1+\hat{\beta}_2{}^2\right)^2 Var\left(\hat{\beta}_1\right) + 2\left(1+\hat{\beta}_1{}^2\right)\left(1+\hat{\beta}_2{}^2\right)Cov\left(\hat{\beta}_1,\hat{\beta}_2\right) + \left(1+\hat{\beta}_1{}^2\right)^2 Var\left(\hat{\beta}_2\right)\right] \qquad (2.64)$$

$$Cov\left(\hat{\beta}_1,\hat{\beta}_2\right) = \left(\hat{\beta}_1 S_{xx}^2\right)^{-1}$$
$$\left\{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})\left[y_i-\bar{y}-\hat{\beta}_1(x_i-\bar{x})\right]\left[y_i-\bar{y}-\hat{\beta}_2(x_i-\bar{x})\right]\right\} \qquad (2.65)$$

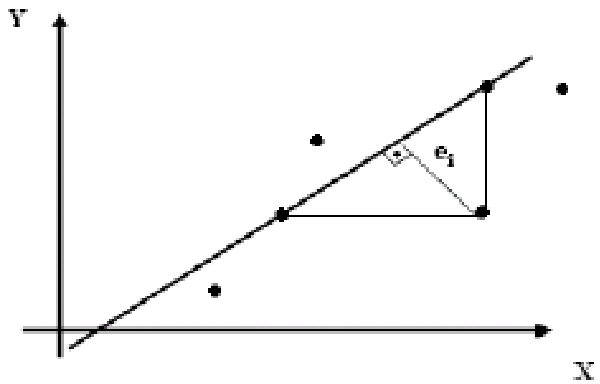
2.2.2.2 Majör Eksen (MA) Ortogonal Regresyon Tekniği

Ortogonal regresyon tekniği ilk kez Adcock (1878) tarafından ortaya konulmuştur (Keleş ve Altun 2016). Major Eksen olarak da adlandırılan Ortogonal Regresyon tekniğinin kullanımı daha sonraki dönemlerde ise Pearson tarafından 1901 yılında önerilmiştir (Isobe *et al.* 1990).

Klasik EKK tekniğinde, değişkenlerin bağımlı veya bağımsız olması önem taşırken, Ortogonal Regresyon tekniğinde hangi değişkenin bağımlı, hangi değişkenin bağımsız olacağı ile ilgilenilmez. Bu teknik her zaman EKK($Y|X$) doğrusu ile EKK($X|Y$) regresyon doğruları arasında yer alır (Amman and Ness 1988; Saraçlı 2008).

MA Regresyon tekniğinde minimize edilmek istenen hata, aşağıdaki şekilde görüleceği üzere gözlem değerlerinin tahmin edilen regresyon doğrusuna olan dik uzaklıkları alınarak hesaplanmaktadır. MA tekniği birçok akademik çalışmada Ortogonal teknik olarak da isimlendirilmektedir (Saraçlı 2011).

Sokal and Rohlf (1995)'e göre, EKK tekniği ile oluşturulan regresyon modelinde gözlemlenen değerler ile oluşturulan regresyon doğrusu arasındaki artıkların karelerinin toplamı minimize edilir, Major eksen regresyon tekniğinde ise tüm noktalar ile regresyon doğrusu arasındaki dikey uzaklıkların kareleri toplamı minimize edilir (Mesple *et al.* 1996). Bu durum Şekil 2.1'de gösterilmiştir (Saraçlı 2011).



Şekil 2.1 Ortogonal Regresyon Tekniğinde minimize edilmek istenen hata

Ortogonal Regresyon tüm gözlem noktalarından regresyon doğrusuna kadar kareli dik mesafelerin toplamını en aza indirir ve eşit hata varyanslarını dikkate alır. Bu teknik için hata kareler toplamı aşağıdaki Eşitlik 2.66'da verildiği gibi hesaplanmaktadır (Wu *et al.* 2018);

$$HKT = \sum_{i=1}^{N} [(x_i - X_i)^2 + (y_i - Y_i)^2] \tag{2.66}$$

Ortogonal Regresyon tekniği ile tahmini yapılmak istenen regresyon doğrusu için eğim katsayısı ve varyansı sırasıyla Eşitlik 2.67 ve Eşitlik 2.68 ile elde edilir ((Isobe *et al.* 1990).

$$\hat{\beta}_{MA} = \frac{1}{2}\left[\left(\hat{\beta}_2 - \hat{\beta}_1^{\,-1}\right) + Sign(S_{xy})\sqrt{4 + \left(\hat{\beta}_2 - \hat{\beta}_1^{\,-1}\right)^2}\right] \tag{2.67}$$

$$Var(\hat{\beta}_{MA}) = \frac{\hat{\beta}_{MA}^{\,2}}{4\hat{\beta}_1^{\,2} + \left(\hat{\beta}_1\hat{\beta}_2 - 1\right)^2}$$
$$\left[\hat{\beta}_1^{\,-2} Var(\hat{\beta}_1) + 2Cov(\hat{\beta}_1, \hat{\beta}_2) + \hat{\beta}_1^{\,2} Var(\hat{\beta}_2)\right] \tag{2.68}$$

Eşitlik 2.68'de tahmin edilen regresyon doğrusunun eğim katsayısı Fuller (1987) tarafından düzenlenerek Eşitlik 2.69'da görüleceği üzere daha sade bir şekilde ifade edilmiştir (Caroll and Ruppert 1994).

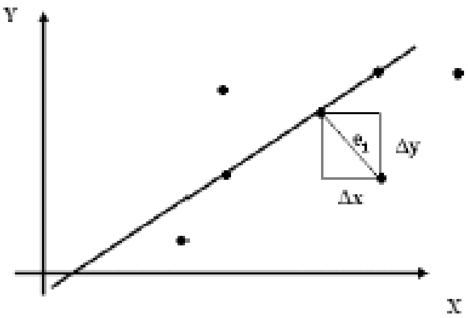
$$\hat{\beta}_{MA} = \frac{\sigma_y^2 - \sigma_x^2 + \left[\left(\sigma_y^2 - \sigma_x^2\right)^2 + 4\sigma_{xy}^2\right]^{1/2}}{2\sigma_{xy}^2} \tag{2.69}$$

2.2.2.3 İndirgenmiş Majör Eksen (RMA) Regresyon Tekniği

İndirgenmiş Majör Eksen (Reduced Major Axis - RMA) regresyon tekniği 20. yüzyılda birçok istatistikçi tarafından incelenmiş ve Kermack and Handle (1950) tarafından öne sürülmüş, literatürde "Uyarlanmış veya Azaltılmış Majör Eksen" olarak da adlandırılan bir tekniktir (Li 2012). Bu teknik ölçek bağımlılığını azaltması amacı ile ortaya koyulmasına rağmen birçok istenmeyen özelliğe de sahiptir (Isobe *et al*. 1990).

RMA Regresyon tekniğinde Y değişkeninin bağımlı, X değişkeninin bağımsız olması halinde oluşacak EKK$(Y|X)$ doğrusu ile X değişkeninin bağımlı, Y değişkeninin bağımsız olması halinde oluşacak EKK$(X|Y)$ regresyon doğrularının eğim katsayılarının geometrik ortalaması hesaplanmaktadır (Li 2012).

RMA Regresyon tekniğinde minimize edilmek istenen hata, aşağıdaki şekilde görüleceği üzere dikey ve yatay uzaklıkların geometrik ortalaması şeklindedir (Saraçlı 2011).



Şekil 2.2 RMA Regresyon Tekniğinde minimize edilmek istenen hata (Saraçlı 2011).

Şekil 2.2'ye göre oluşan iki üçgen birbirine göre simetriktir, ayrıca X ve Y değişkenlerinin yer değiştirmesi iki üçgenin de alanını değiştirmez (Li 2012). Burada minimize edilmek istenen hata kareler toplamı Eşitlik 2.70 ile hesaplanır (Wu *et al*. 2018);

$$HKT = \sum_{i=1}^{N} [(x_i - X_i) + (y_i - Y_i)] \tag{2.70}$$

RMA Regresyon Tekniği ile elde edilecek regresyon doğrusunun eğim katsayısı ve varyansı Eşitlik 2.71 ve Eşitlik 2.72 ile hesaplanmaktadır (Isobe *et al*. 1990, Li 2012).

$$\hat{\beta}_{RMA} = Sign(S_{XY})\sqrt{\hat{\beta}_1, \hat{\beta}_2} = Sign(S_{XY})\sqrt{\hat{\beta}_{EKK(Y|X)}\hat{\beta}_{EKK(X|Y)}}$$

$$= Sign(S_{XY})\sqrt{(S_{XY}/S_{XX})(S_{YY}/S_{XY})} = Sign(S_{XY})\sqrt{S_{YY}/S_{XX}} \qquad (2.71)$$

$$= Sign(S_{XY})\sqrt{S_{YY}/S_{XX}}$$

$$Var(\hat{\beta}_{RMA}) = \frac{1}{4}\left[\frac{\hat{\beta}_2}{\hat{\beta}_1}Var(\hat{\beta}_1) + 2Cov(\hat{\beta}_1, \hat{\beta}_2) + \frac{\hat{\beta}_1}{\hat{\beta}_2}Var(\hat{\beta}_2)\right] \qquad (2.72)$$

2.2.2.4 Deming Regresyon Tekniği

Deming Regresyon Tekniği, ünlü Alman kaliteci W. Edwards Deming' in 1943 yılında yazdığı "Statistical Adjustment of Data" adlı kitabında En Küçük Kareler metodundaki problemler ve eksiklikleri ele alması ile ortaya konmuştur (Saraçlı 2008).

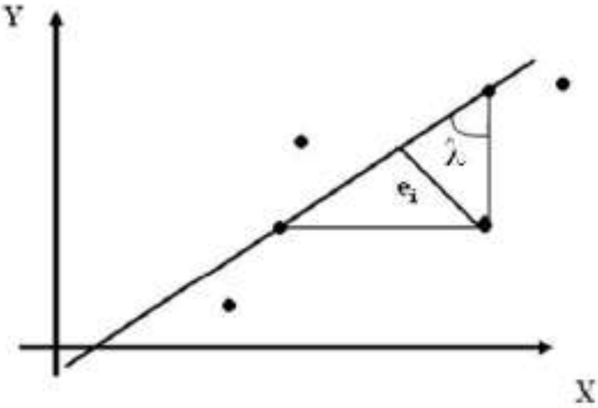
Deming Regresyon Tekniği hem X hem de Y değişkenlerinin ölçüm hatalarını içeren iki boyutlu verilerin düz bir çizgi ile belirtilmesini sağlayan bir tekniktir. Bu teknik, yalnızca Y değişkenlerinin hatalarını içeren Basit Doğrusal Regresyon Tekniğinden farklıdır. Deming Regresyonu, iki ölçüm metodu arasındaki farklılıkları araştırmak için özellikle klinik kimyadaki metot karşılaştırma çalışmaları için kullanılmaktadır (İnt.Kay.1).

Deming Regresyon, ölçüm hata oranının $\lambda = S_{ex}^2/S_{ey}^2$ (veya $\lambda = V(\varepsilon_i)/V(\delta_i)$) gibi bir sabit değere eşit olduğunu varsayar. Uygulanan bu işlem, çalışmayı yapan kişinin bilinen bir hata oranını işleme dahil etmesini veya verilerden hata oranının hesaplanabilmesi için her gözlem değeri için birden fazla ölçüm yapılmasını gerektirir (İnt.Kay.1). Ayrıca $\lambda = 1$ olması halinde gözlem değerlerinden regresyon doğrusuna çizilen dikey ve yatay uzaklıklar ikizkenar dik üçgen oluşturmakta ve Deming Regresyon, Ortogonal Regresyon ile aynı sonuçları vermektedir (İnt.Kay,1, Saraçlı 2008).

Deming Regresyon tekniğinde minimize edilmek istenen Hata Kareler Toplamı (HKT) Eşitlik 2.73 ile hesaplanmaktadır (İnt.Kay.1).

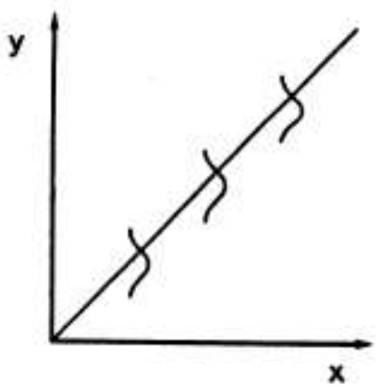
$$HKT = \sum_{i=1}^{N} [(x_i - X_i)^2 + \lambda(y_i - Y_i)^2]\, y_i = Y_i + \delta_i \quad (2.73)$$

Deming Regresyon analizinde gözlem değerlerinin, hesaplanan regresyon doğrusuna λ açısı ile gelen uzaklığının karesi minimize edilmeye çalışılır (Linnet 1998; Saraçlı 2008). Bu işlem Şekil 2.3'de görsel olarak ifade edilmiştir.

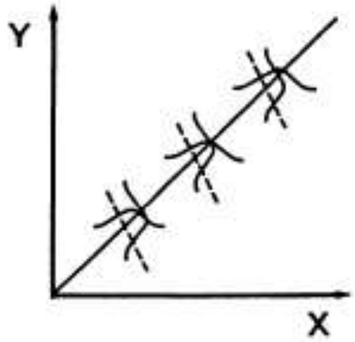


Şekil 2.3 Deming Regresyon Tekniği İle Minimize Edilmek İstenen Hata (Saraçlı 2011).

EKK ve Deming Regresyon tekniğine ait X ve Y değişkenine ait hataların dağılımı Şekil 2.4 ve 2.5'de gösterilmiştir.



Şekil 2.4 EKK Regresyon Tekniği ile oluşturulan regresyon doğrusu için, X değişkeni hata içermiyorken Y değişkeni için sabit Gauss hata dağılımına ilişkin grafik (Linnet 1993).



Şekil 2.5 Deming Regresyon Tekniği ile elde edilen hataların dağılımına ilişkin grafik (Linnet 1993).

Deming regresyon yönteminde, X ve Y metotlarının hataların Gaussian (normal) dağıldığı varsayılır (Akkoca 2012).

Deming Regresyon Tekniği ile tahmin edilmek istenen regresyon denklemine ait eğim katsayısı $\hat{\beta}_1$ ve regresyon sabit katsayısı $\hat{\beta}_0$ Eşitlik 2.74 ve Eşitlik 2.78 ile hesaplanmaktadır (Saraçlı vd. 2009).

$$\hat{\beta}_1 = \frac{(\lambda q - u) + \sqrt{(u - \lambda u)^2 + 4\lambda p^2}}{2\lambda p} \tag{2.74}$$

$$u = \sum_{i=1}^{N} (x_i - \bar{x})^2 \tag{2.75}$$

$$q = \sum_{i=1}^{N} (y_i - \bar{y})^2 \tag{2.76}$$

$$p = \sum_{i=1}^{N} (x_i - \bar{x})\,(y_i - \bar{y}) \tag{2.77}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \tag{2.78}$$

Deming Regresyon Tekniğindeki tüm tahminler için standart hatalar “Jackknife Metodu” kullanılarak hesaplanır. Jackknife Metodu parametrik olmayan bir tekniktir ve

Deming Regresyon Tekniği için yeterli performans gösterdiği Linnet (1990) tarafından ifade edilmiştir (İnt.Kay.1). Bu metoda ile hesaplanan Deming Regresyon katsayılarına ilişkin hataların hesabı Eşitlik 2.79-2.82 ile açıklanmıştır (Linnet 1990, Saraçlı vd. 2009).

$$b_i = kb - (k-1)b^i \qquad i = 1, \dots, k \tag{2.79}$$

$$\tilde{b} = \sum_{i=1}^{k} \frac{b_i}{k} \tag{2.80}$$

$$V_J(\hat{\beta}_{DEM}) = \sum_{i=1}^{k} \left(b_i - \tilde{b}\right)^2 / (k-1) \tag{2.81}$$

$$SE(\hat{\beta}_{DEM}) = \sqrt{V_J(\hat{\beta}_{DEM})/k} \tag{2.82}$$

Parvin (1984) yaptığı bir çalışmada Deming tekniği ile hesaplanan eğim katsayısı, “Deming’in Eğim Katsayısı” veya “Doğrusal Fonksiyonel Eğim Katsayısı” (b_{LF}) ve EKK tekniği ile hesaplanan eğim katsayısı (b_{LS}) arasındaki ilişki Eşitlik 2.83 ile ifade edilmiştir.

$$b_{LF} = \frac{1}{2}\left\{\frac{b_{LS}}{r^2} - \frac{\lambda}{b_{LS}} + \left[\left(\frac{b_{LS}}{r^2} - \frac{\lambda}{b_{LS}}\right)^2 + 4\lambda\right]^2\right\} \tag{2.83}$$

Eşitlik 2.83, bilinen bir r ve λ değeri için, b_{LF} ve b_{LS} arasındaki kesin ilişkiyi tanımlar ve matematiksel olarak yorumlanması aşağıdaki gibidir.

- b_{LF} değeri her zaman b_{LS} değerinden daha büyüktür.
- λ değeri büyüdükçe, Deming tekniği ile hesaplanan eğim katsayısı (b_{LF}), EKK tekniği ile hesaplanan eğim katsayısı (b_{LS})’ ye yaklaşır.
- λ değeri küçüldükçe, (b_{LF}) değeri b_{LS}/r^2 değerine yaklaşacaktır. Bu olgu eksenleri tersine çevrilmiş EKK eğim katsayısı tahminine benzemektedir.

- r değeri küçüldükçe, b_{LF} ve b_{LS} değerleri arasındaki fark artarak değerler birbirinden uzaklaşır.
- λ değeri yaklaşık olarak "1" değerine ulaştığında b_{LF} ve b_{LS} değerleri arasındaki fark λ değerine karşı duyarlı hale gelir (Parvin 1984).

2.2.2.5 Passing-Bablok Regresyon Tekniği

Passing ve Bablok 1983 yılında klasik doğrusal regresyon varsayımlarına ihtiyaç duymayan bir yöntem geliştirerek iki ölçüm yönteminin karşılaştırılmasını sağlayan bir regresyon yöntemi geliştirmişlerdir (İnt.Kay.4).
Passing- Bablok Regresyon yönteminin farklılıklarını incelemek adına öncelikle Klasik EKK varsayımları yazılacak olursa (Acıtaş 2014, İnt.Kay.3, Alma ve Vupa 2008; Fox 1997, Türkay 2004):

- Hata terimleri Normal dağılıma sahip ve homojendir.
- Hata terimlerinin beklenen değeri "0" , varyansları sabit ve σ^2 değerine eşittir.
- Hata terimleri birbirinden bağımsız ve aynı dağılıma sahiptir.
- Hata terimleri arasında ardışık bağımlılık (otokorelasyon) yoktur.
- Hata terimleri ile açıklayıcı (bağımsız) değişkenler arasında bir ilişki yoktur.
- Açıklayıcı değişkenler arasında çoklu doğrusal bağlantı yoktur.
- Ölçüm hatasının varyansı sabittir.
- Regresyon denkleminde Y bağımlı değişkenini açıklayan X bağımsız değişkenleri hata içermez.
- Regresyon modeli aykırı değer içermez.

şeklindedir. Passing-Bablok Regresyon tekniğinde verilerin dağılımına ilişkin özel bir varsayım gerekliliği olmayan bir regresyon modeli öne sürülmüştür. Bu parametrik olmayan regresyon tekniği gözlem değerlerinin sıralanmasına dayanmaktadır. Fakat burada işlemlerin hesaplanması oldukça uzun zaman almaktadır. Bu regresyon tekniğinde seçilen X (bağımsız değişken) referans yöntemi ile, Y (bağımlı değişken) test yöntemi verilerinin atanması birbirinden bağımsız gerçekleşmektedir (İnt.Kay.5).

Passing ve Bablok (1983) önerdiği regresyon modeli Eşitlik 2.84-2.87'de verilmiştir.

$$y_i = a + \beta x_i + \eta_i \qquad (2.84)$$

$$x_i = A + By_i + \xi_i \qquad (2.85)$$

ve

$$x_i = x_i^* + \xi_i \qquad (2.86)$$

$$y_i = y_i^* + \eta_i \qquad (2.87)$$

Burada η_i ve ξ_i hata terimleri aynı dağılıma sahip olmaktadır. Ayrıca x_i^* ve y_i^* gözlem değerlerine ilişkin tahmin edilen değerlerdir. Passing-Bablok yöntemi a ve β ($A = -a/\beta, B = 1/\beta$ olarak hesaplanan) katsayılarının hesaplanmasını ve bu katsayılara ilişkin güven aralığı hesaplamalarının yapılmasını sağlar. Burada bu iki değerin incelenmesi ile iki yöntemin karşılaştırılması yapılabilir (İnt.Kay.4).

Passing-Bablok Regresyon tekniğinde eğim katsayısı β_1 $0/0$ veya -1 ile sonuçlanan çifter hariç, tüm olası veri noktası çiftlerinden oluşturulabilecek tüm eğimlerin medyanı olarak hesaplanır. Tahmin edilen sabit β_0 katsayısı ise ($Y_i - \beta_1 X_i$) değerinin medaynı olarak hesaplanır (İnt.Kay.2).

Passing-Bablok Regresyonun özellikleri ise aşağıdaki gibidir (İnt.Kay.2):

- X ve Y değişkenleri yüksek ve pozitif korelasyona sahiptir.
- X ve Y değişkenleri arasındaki ilişki doğrusaldır.
- X ve Y'nin dağılımları hakkında özel bir varsayım yapılamaz (varyansları hariç).
- Hata terimlerinin dağılımlarının Gaussian (Normal) olması şartı yoktur (Saraçlı 2008).
- β_1 ve β_0 parametreleri parametrik olmayan esaslara göre tahmin edilir.

Passing-Bablok Regresyon tekniğine göre tahmin edilen regresyon denklemine ilişkin eğim katsayısı ve sabit katsayı hesabı Eşitlik 2.88-2.93'de açıklanmıştır (Passing and Bablok 1983).

$$S_{ij} = \frac{y_i - y_j}{x_i - x_j} \qquad 1 \leq i \leq j \leq n \; için, \tag{2.88}$$

$$S_{ij} = \frac{{y_i}^* - {y_j}^* + \eta_i - \eta_i}{{x_i}^* - {x_j}^* + \xi_\mathrm{i} - \xi_\mathrm{i}} \tag{2.89}$$

$${y_i}^* = \alpha + \beta {x_i}^* \quad ve \quad d_{ij} = {x_i}^* - {x_j}^* \quad olmak\, üzere, \tag{2.90}$$

$$S_{ij} = \frac{\beta d_{ij} + (\eta_i - \eta_i)}{d_{ij} + (\xi_\mathrm{i} - \xi_\mathrm{i})} = \beta \frac{d_{ij} + \frac{\eta_i - \eta_i}{\beta}}{d_{ij} + (\xi_\mathrm{i} - \xi_\mathrm{i})} = \beta \frac{d_{ij} + z_{ij}}{d_{ij} + {z_{ij}}'} \tag{2.91}$$

Burada z_{ij} ve ${z_{ij}}'$ bağımsız ve aynı dağılıma sahiptir. Burada S_{ij} değerleri bağımsız olmadığı için β'nın yanlı bir tahmin edicisi olacaktır. Bu yüzden işleme devam edilir:

$$\beta = \begin{cases} S_{\frac{N+1}{2}+K} & , N çift\; ise \\ \frac{1}{2}\left(S_{\left(\frac{N}{2}+K\right)} + S_{\left(\frac{N}{2}+1+K\right)}\right) & , N\; tek\; ise \end{cases} \tag{2.92}$$

$$\alpha = med\{y_i - \beta x_i\} \tag{2.93}$$

Burada K, $S_{ij} \leq -1$ olan S_{ij} değer değerlerinin sayısıdır.

2.2.2.6 York Regresyon Tekniği

Tip II Regresyon tekniklerinde bağımlı ve bağımsız değişkenlerin hataları göz önünde bulundurularak bir regresyon modeli oluşturulur. York Regresyon tekniğinde, York (1969)'un belirttiği üzere X ve Y değişkenlerinin hataları analize dahil edilirken aynı zamanda bu değişkenler arasındaki korelasyon ve ağırlık değerlerinin hesabına ilişkin farklı metodlar uyarlanmıştır (Saylor *et al*. 2006).

York (1969)' un çalışmasında yer alan ve minimize edilmek istenen Hata Kareler Toplamı (HKT) aşağıdaki Eşitlik 2.94'de verilmiştir (Saylor *et al*. 2006).

$$HKT = \Big[\omega(X_i)(x_i - X_i)^2 - 2r_i\sqrt{\omega(X_i)\omega(Y_i)} \times (x_i - X_i)(y_i - Y_i) + \omega(Y_i)(y_i - Y_i)^2\Big] \times \frac{1}{(1 - {r_i}^2)} \qquad (2.94)$$

Burada r_i, x_i ve y_i değişkenleri arasındaki korelasyon katsayısını ifade etmektedir. Bazı iterasyonlara ihtiyaç duyan York Regresyon tekniğinin eğim katsayısı ise aşağıdaki Eşitlik 2.95'de verildiği gibidir (Saylor *et al*. 2006).

$$b = \frac{\sum_{i=1}^{N} W_i\beta_i(y_i - \bar{y})}{\sum_{i=1}^{N} W_i\beta_i(x_i - \bar{x})} \qquad (2.95)$$

Buradan W_i ve β_i değerleri sırası ile Eşitlik 2.96 ve Eşitlik 2.97 ile hesaplanmaktadır (Saylor et al. 2006).

$$W_i = \frac{\omega(x_i)\omega(y_i)}{\omega(x_i) + b^2\omega(y_i) - 2br_i\sqrt{\omega(x_i)\omega(y_i)}} \qquad (2.96)$$

$$\beta_i = W_i\left[\frac{(x_i - \bar{x})}{\omega(x_i)} + \frac{b(y_i - \bar{y})}{\omega(x_i)} - b(x_i - \bar{x}) + (y_i - \bar{y})\frac{r_i}{\sqrt{\omega(x_i)\omega(y_i)}}\right] \qquad (2.97)$$

Burada,

$$\bar{X} = \frac{\sum_{i=1}^{N} W_i x_i}{\sum_{i=1}^{N} W_i} \qquad (2.98)$$

$$\bar{Y} = \frac{\sum_{i=1}^{N} W_i y_i}{\sum_{i=1}^{N} W_i} \qquad (2.99)$$

Eşitlik 2.98 ve Eşitlik 2.99 sırası ile $\bar{X}$ ve $\bar{Y}$ değerlerini ifade etmektedir. Saraçlı (2008) ağırlık değerlerine ilişkin ($\omega(x_i)$, $\omega(y_i)$), "York Regresyonda standart olarak 1, gözlem değerlerinin hataları arasındaki korelasyon katsayısı olan r_i değeri ise 0 olarak alınır (Simülasyon çalışmalarında geçerlidir).” yorumunda bulunmuştur.

Eşitlik 2.97’de görüldüğü üzere W_i ve β_i, b ifadesinin birer fonksiyonudur ve iterasyon yapılarak belirlenmelidir. Her bir veri çifti için verilen $\omega(x_i)$ ve $\omega(y_i)$ ağırlıkları ve korelasyon katsayısı ile Ortogonal Regresyon tekniği veya EKK tekniği ile hesaplanmış olan bir başlangıç b değeri seçilir (Saraçlı 2008). b’nın ardışık değerleri önceden belirlenmiş bir tolerans dahilinde olana kadar aşağıdaki adımlar aşağıdaki gibidir (Saylor *et al.* 2006):

1. b, $\omega(x_i)$ ve $\omega(y_i)$ değerleri kullanılarak her bir gözlem değeri için Eşitlik 2.96’da gösterildiği üzere W_i değerleri belirlenir.
2. Tüm (x_i, y_i) gözlem değerleri için W_i değerleri yardımıyla $\bar{x}$ ve $\bar{y}$ ortalama değerleri hesaplanır.
3. Eşitlik 2.97’de gösterildiği üzere, tüm (x_i, y_i) gözlem değerleri için β_i değerleri hesaplanır.
4. Eşitlik 2.95’de gösterildiği üzere yeni bir b tahmin değeri hesaplanır ve 1. Adıma geri dönülür.
5. York regresyon tekniği için, sabit katsayı değeri ise $\alpha = \bar{y} - b\bar{x}$ şeklinde hesaplanır.

York Regresyon tekniğinde tahmin edilen regresyon denkleminin katsayılarına ilişkin standart hatalar Eşitlik 2.100 ve Eşitlik 2.101 ile hesaplanmaktadır (York *et al.* 2004; Saraçlı 2008):

$$\sigma_\alpha^2 = \frac{1}{\sum W_i} + \bar{x}^2\sigma_b^2 \tag{2.100}$$

$$\sigma_b^2 = \frac{1}{\sum W_i\,(x_i - \bar{x})^2} \tag{2.101}$$

2.3 Bayes Regresyon

Bayesci regresyonda Bayes kestirimi, önsel parametre kestirimi ile en küçük kareler (EKK) kestiriminin ağırlıklı ortalaması olarak ifade edilmektedir. En küçük kareler yönteminden daha duyarlı olduğu bilinen Bayes tahmin edicisi önsel dağılımın kullanılmasından dolayı daha yanlı bir kestirici olmaktadır (Yardımcı 2000).

Bayesci yaklaşımda regresyona ait parametre daha önce belirtildiği gibi bir sabit değil, bir rassal değişken olarak kabul edilir. Kısacası klasik regresyonda regresyon parametresi ($\hat{\beta} \sim N(\beta, \sigma^2 (X'X)^{-1}$) olarak ifade edilirken, Bayesci regresyonda $P(\beta \mid \sigma) \sim N(\hat{\beta}, \sigma^2 (X'X)^{-1})$ olarak ifade edilir. Açıkça görüldüğü üzere $(\beta_1, \beta_2, \dots, \beta_k)$ ve σ^2 parametreleri olasılık dağılımları olan birer rasgele değişkendir. Doğrusal regresyon modeli için Bayesci çıkarsama uygulandığında daha önce de belirtildiği gibi $sonsal\ dağılım \propto olabilirlik\ fonksiyonu \times önsel\ dağılım$ ilkesi kullanılarak istenilen matematiksel modeller elde edilebilecektir. Bu ifade Eşitlik 2.102 ile matematiksel olarak förmülüze edilmiştir (Judge *et al*. 1985).

$$P(\beta_1, \beta_2, \sigma^2 \mid y, X) = P(y, X \mid \beta_1, \beta_2, \sigma^2) \times P(\beta_1, \beta_2, \sigma^2) \qquad (2.102)$$

Basit Doğrusal Regresyon Modeli İçin Bayesci Yaklaşım: Bayesci çıkarsama için Eşitlik 2.105'de verilen lineer regresyon modeli ele alınacaktır.

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \qquad i = 1,2,\dots,n \qquad (2.103)$$

Eşitlik 2.103'de belirtilen regresyon modelinde y_i bağımlı değişkeni, x_i bağımsız değişkeni, ε_i ise hata terimini ifade etmektedir. Bayesci doğrusal regresyon modelinin varsayımları (Gasım 2013):

1. Hata terimlerinin dağılımı $\varepsilon_i \sim N(0, \sigma^2)$ 'dir.
2. Rasgele değişkenler bağımsızdır.
3. Katsayılara göre doğrusaldır.
4. Ölçme hatası yoktur.

Regresyonda X bir sabit veya rasgele değişken olarak adlandırılmaktadır. Bir regresyon problemindeki sayısal veri hem X i hem de y' yi içerir. Böylece tam bir Bayesci modeli bir parametre vektörü ψ ile indekslenen X, $P(X \mid \psi)$ için bir dağılım içerir ve bu nedenle önsel dağılım olan $P(\psi, \beta_0, \beta_1, \sigma^2)$ ile $P(y, X \mid \beta_0, \beta_1, \sigma^2, \psi)$ ortak dağılım olasılığının değerlerini içerir. Kısaca ifade edilmek istenirse X'in dağılımını içeren dağılım fonksiyonu $P(X \mid \psi)$'dir. Standart regresyon modelinde X biliniyorken y'nin koşullu dağılımı hakkında hiçbir bilgi olmadığı varsayılarak X'in dağılımı için; β_0, β_1 için belirlenen $P(y \mid X, \beta_0, \beta_1, \sigma^2)$ ve ψ için belirlenen $P(y \mid X, \psi)$ fonksiyonlarının bağımsız olduğu varsayılır.

Bu nedenle Bayes bakış açısına göre, bir regresyon modelinin tanımlayıcı özelliği X'in $(\beta_0, \beta_1, \sigma^2, \psi)$ hakkında verilen bilgiyi gözardı etmesidir. Bu durum şöyle açıklanabilir, β_0, β_1 ve ψ' nın önsel dağılımlarının bağımsız olması halinde $(P(\beta_0, \beta_1, \psi) = P(\beta_0, \beta_1, \sigma^2). P(\psi))$ sonsal dağılımları da Eşitlik 2.104 ve Eşitlik 2.105'de belirtildiği gibi olacaktır.

$$P(\psi, \beta_0, \beta_1, \sigma^2 \mid X, y) = P(\psi \mid X) P(\beta_0, \beta_1, \sigma^2 \mid X, y) \qquad (2.104)$$

$$P(\beta_0, \beta_1, \sigma^2 \mid X, y) \propto P(\beta_0, \beta_1, \sigma^2) P(y \mid X, \beta_0, \beta_1, \sigma^2) \qquad (2.105)$$

Yukarıda görüldüğü üzere bu oransallık ifadesi ile Klasik regresyon modelinde olduğu gibi X'in dağılımına olan bağımlılık giderilmiş olur (Gelman *et al*. 1995).

Klasik lineer Bayes regresyon modeli $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ ifadesinde X, β_0, β_1 $(k = 1,2)$ ve σ^2 veri iken, y_i $(i = 1,2, \ldots n)$ gözlem değerleri ile normal dağılır. Kısaca ifade etmek gerekirse Eşitlik 2.106 elde edilebilir (Gelman *et al*. 1995).

$$y_i \mid \beta_0, \beta_1, \sigma, X \sim N(X\beta_0\beta_1, \sigma^2) \qquad (2.106)$$

Bayesci lineer regresyon denklemi bu şartlar altında ifade edilirse y_i bağımlı değişkeni için beklenen değeri (ortalaması) ve varyansı Eşitlik 2.107 ve Eşitlik 2.108'de belirtilen:

$$E(y_i \mid x_i, \beta_0, \beta_1, \sigma^2) = \beta_0 + \beta_1 x_i \qquad (2.107)$$

$$Var(y_i \mid x_i, \beta_0, \beta_1, \sigma^2) = \sigma^2 \quad (2.108)$$

için benzerlik fonksiyonu Eşitlik 2.109'da verildiği gibidir (Gasım 2013).

$$P(\beta_0, \beta_1, \sigma \mid y, X) \propto \frac{1}{\sigma^n} exp\left[-\frac{1}{2\sigma^2}\sum(y_i - \beta_0 - \beta_1 x_i)^2\right] \quad (2.109)$$

Bilgi veren önsel dağılım ile Doğrusal Bayes Regresyon Modelinin oluşturulması: 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. Her iki durumda da bu tür bir ön bilgiyi analize dâhil etmek için Bayesci analize ihtiyaç duyulmaktadır. Bu bölümde doğal eşlenik önsel dağılım kullanılarak, bilgi veren önsel dağılım yardımı ile sonsal dağılım elde edilecektir (Judge *et al*. 1985).

Genel olarak β ve σ parametrelerinin tahminleri için ne tür bir olasılık yoğunluk fonksiyonu kullanıldığı önemli değildir. Seçilebilecek bir dizi olasılık yoğunluk fonksiyonu ile önsel dağılım kullanılarak sonsal dağılım oluşturulabilir. Ancak pratikte 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. Ayrıca burada bu önsel dağılım, benzerlik fonksiyonuyla da aynı yapıya sahiptir. Parametreler için Doğal Eşlenik Önsel Dağılımı elde etmek için öncelikle olabilirlik fonksiyonu, yeterli istatistik olarak ifade edilmeli ve daha sonra bilinmeyen parametrelerin bir fonksiyonu olarak fonksiyonel biçimi elde edilmelidir (Judge *et al*. 1985).

Bayes Regresyon modelini oluşturabilmek adına öncelikli olarak Bayes teoreminden yola çıkılırsa; θ ilgilendiğimiz parametrenin bir vektörü, y ise $f(y \mid \theta)$ yoğunluk fonksiyonundan alınan bir örneklem gözlem vektörü olmak üzere $f(y \mid \theta)$, θ için olabilirlik fonksiyonu ile cebirsel olarak aynıdır ve θ ile ilgili tüm örneklem bilgisini içerir. Eğer θ rasgele bir vektör olarak kabul edilirse Bayes teoremine dayanarak Eşitlik 2.110 yazılabilir.

$$h(\theta, y) = f(y \mid \theta)g(\theta) = g(\theta \mid y)f(y) \tag{2.110}$$

Yukarıda $h(\theta, y)$ ortak yoğunluk fonksiyonunu, $g(\theta)$ önsel yoğunluk fonksiyonunu, $g(\theta \mid y)$ sonsal yoğunluk fonksiyonunu, $f(y)$, y vektörünün marjinal fonksiyonunu ifade etmektedir. Burada Eşitlik 2.110 yeniden düzenlenerek Eşitlik 2.111 yazılabilir.

$$g(\theta \mid y) = \frac{f(y \mid \theta)g(\theta)}{f(y)} \tag{2.111}$$

Elde edilen Eşitlik 2.111, Bayes teoremi olarak bilinmektedir. Burada θ parametresi için sonsal yoğunluk fonksiyonu $g(\theta \mid y)$'dir. Çünkü bu fonksiyon için θ, y gözlem değerlerine ilişkin tüm örneklem bilgisini özetlemektedir. Eğer θ için $f(y)$ fonksiyonu bir sabit olarak kabul edilip $f(y \mid \theta)$, $l(\theta \mid y)$ olabilirlik fonksiyonu biçimine dönüştürülürse Eşitlik 2.111, Eşitlik 2.112 şeklinde ifade edilebilir.

$$g(\theta \mid y) = l(\theta \mid y)g(\theta) \tag{2.112}$$

Eşitlik 2.112, en basit haliyle Eşitlik 2.113'ü ifade etmektedir.

$$sonsal\ dağılım \propto olabilirlik\ fonksiyonu \times önsel\ dağılım \tag{2.113}$$

Burada olabilirlik fonksiyonu "örneklem bilgisi" olarak da ifade edilebilmektedir. Bu aşamada elde edilen Bayes ilkeleri Lineer Doğrusal Regresyon modeline uyarlanarak Lineer Bayes Regresyon Modeli elde edilebilmektedir.

$$y = X\beta + \varepsilon \tag{2.114}$$

Lineer Regresyon modeli olan Eşitlik 2.114'e göre, y, (T×1) büyüklüğünde bağımlı değişkenlerin gözlem vektörü, X, K açıklayıcı değişkenleri üzerinde (T×K) boyutlu bir gözlem matrisi, ε, $\varepsilon \sim N(0, \sigma^2)$ dağılıma sahip (T×1) boyutunda bir hata vektörü ve β ve

σ hakkında bilgi edinmek istediğimiz parametreleri ifade etmektedir. Burada $\theta = (\beta', \sigma)'$ olmak üzere Eşitlik 2.112 yardımı ile Eşitlik 2.115 elde edilebilir.

$$g(\beta, \sigma \mid y) \propto l(\beta, \sigma \mid y) g(\beta, \sigma) \qquad (2.115)$$

Burada $g(\beta, \sigma \mid y)$, y gözlem değerleri biliniyorken β ve σ bilinmeyen parametrelerine ait sonsal dağılımı, $l(\beta, \sigma \mid y)$, olabilirlik fonksiyonunu, $g(\beta, \sigma)$ önsel bilgiyi ifade etmektedir. Örneklem bilgisini özetleyen olabilirlik fonksiyonu Bayesci çıkarsamada önemli bir rol oynar (Judge *et al.* 1985). Klasik yaklaşımın varsayımlarından biri hata terimi olan ε'nin normal dağıldığıdır. Aynı varsayım altında Bayesci yaklaşım için olabilirlik fonksiyonu Eşitlik 2.116'da belirtildiği gibi yazılabilir (Genç vd. 2010).

$$l(\beta, \sigma \mid y) \propto \frac{1}{(2\pi)^{T/2}} \frac{1}{\sigma^T} \, exp\left\{-\frac{1}{2\sigma^2}[(y - X\beta)'(y - X\beta)]\right\} \qquad (2.116)$$

Burada $(b', \hat{\sigma}^2)$ yeterli istatistik olmak üzere Eşitlik 2.117 yazılabilir.

$$b = (X'X)'X'y \qquad \hat{\sigma}^2 = (y - Xb)'(y - Xb)/v \qquad v = T - K \qquad (2.117)$$

Eşitlik 2.119'da belirtilen ifadeler EKK yöntemi ile elde edilen parametre tahmin edicileridir. Bu bilgiler aşığında, parametreden bağımsız ifadelerin denklemden çıkarılması ve yukarıdaki ifadelerin Eşitlik 2.116'da yerine yazılması ile oluşacak olabilirlik fonksiyonu Eşitlik 2.118'de verilmiştir.

$$l(\beta, \sigma \mid y) \propto \frac{1}{\sigma^T} \, exp\left\{-\frac{1}{2\sigma^2}[v\hat{\sigma}^2 + (\beta - b)'X'X(\beta - b)]\right\} \qquad (2.118)$$

Bayesci yaklaşımda (β', σ) parametrelerine ilişkin önsel bilgi ile mevcut verilerle oluşturulan olabilirlik fonksiyonu birleştirilerek sonsal yoğunluk fonksiyonu oluşturulabilir. Eşitlik 2.118 incelendiğinde, $l(\beta, \sigma \mid y)$ "Normal-Gamma" fonksiyonu biçimindedir. Çünkü:

$$l(\beta, \sigma \mid y) \propto h_1(\beta \mid \sigma, y).h_2(\sigma \mid y) \qquad (2.119)$$

Eşitlik 2.119' da verildiği üzere iki parçaya ayrılacak olursa:

$$h_1(\beta \mid \sigma, y) = \; exp\left\{-\frac{1}{2\sigma^2}[(\beta - b)'X'X(\beta - b)]\right\} \tag{2.120}$$

$$h_2(\sigma \mid y) \; = \frac{1}{\sigma^T} exp\left(\frac{v\hat{\sigma}^2}{2\sigma^2}\right) \tag{2.121}$$

Eşitlik 2.120 ve Eşitlik 2.121 elde edilecektir. Görüldüğü üzere (β', σ) için Doğal eşlenik önsel dağılım, σ biliniyorken β'nın koşullu dağılımı "çok değişkenli normal" ve σ'nın marjinal dağılımı "Ters-Gamma"dır (Judge *et al*. 1985).

Yukarıda elde edilen Eşitlik 2.120 ve 2.121 yardımı ile önsel dağılıma ilişkin β ve σ parametrelerinin tahmin edicilerini belirlemek gerekmektedir. Aşağıdaki eşitliklerde kullanılacak olan $\bar{\beta}$, $\bar{v}$ ve $\bar{s}^2$ ifadeleri önsel dağılıma ilişkin tahmin edicileri $\bar{\bar{\beta}}$, $\bar{\bar{v}}$ ve $\bar{\bar{s}}^2$ ifadeleri ise sonsal dağılıma ilişkin tahmin edicileri göstermektedir (Tiao and Zellner 1964; Gasım 2013).

β ve σ parametrelerinin önsel yoğunluk fonksiyonları aşağıdaki eşitlikler yardımı ile elde edilmektedir (Judge *et al*. 1985). Eşitlik 2.116 göz önünde bulundurularak Eşitlik 2.122 ve Eşitlik 2.123 yazılabilir.

$$g(\beta \mid \sigma) = \frac{1}{(2\pi)^{K/2}} \frac{1}{\sigma^K} \, |A|^{1/2} exp\left[-\frac{1}{2\sigma^2}(\beta - \bar{\beta})'A(\beta - \bar{\beta})\right] \tag{2.122}$$

$$g(\sigma) = \frac{2}{\Gamma(\bar{v}/2)} \left(\frac{\bar{v}\bar{s}^2}{2}\right)^{\bar{v}/2} \frac{1}{\sigma^{\bar{v}+1}} exp\left(-\frac{\bar{v}\bar{s}^2}{2\sigma^2}\right) \tag{2.123}$$

Eşitlik 2.123'ün açılarak gerekli oransallık işlemlerinin yapılması ile:

$$g(\sigma) = \frac{1}{\sigma^{\bar{v}+1}} \; exp\left(-\frac{\bar{v}\bar{s}^2}{2\sigma^2}\right) \tag{2.124}$$

Eşitlik 2.124 yazılabilir. Burada Eşitlik 2.125 ile β parametresinin, Eşitlik 2.126 ile de σ parametresinin ortalama ve varyans değerleri elde edilebilir.

$$E[\beta \mid \sigma] = E[\beta] = \bar{\beta} \qquad cov[\beta \mid \sigma] = A^{-1}\sigma^2 \qquad A = (X'X)^{-1}X' \tag{2.125}$$

$$E[\sigma] = \frac{\Gamma[(\overline{v}-1)/2]}{\Gamma(\overline{v}/2)}\left(\frac{\overline{v}}{2}\right)^{1/2}\bar{s} \qquad E[\sigma^2] = \frac{\overline{v}\bar{s}^2}{\overline{v}-2} \tag{2.126}$$

Ayrıca yukarıda görüldüğü üzere Eşitlik 2.122 ve 2.123'de görülmekte olan önsel parametreler, Eşitlik 2.120 ve 2.121'de görülen yeterli istatistiklerin yerini almıştır. Burada elde edilen Eşitlik 2.122 ve Eşitlik 2.123'in birleştirilmesi ile "Normal-Gamma" dağılımına sahip ortak önsel ortak yoğunluk fonksiyonu elde edilebilmektedir. Bu işlem aşağıdaki şekilde gerçekleştirilmektedir, Bayes teoreminden yola çıkarak:

$$g(\beta,\sigma) = g(\beta \mid \sigma).g(\sigma) \tag{2.127}$$

ve Eşitlik 2.122 ve Eşitlik 2.123'de belirtilen ifadeler Eşitlik 2.127'de yerine yazılacak olursa, gerekli işlemlerin yapılmasının ardından Eşitlik 2.128 elde edilebilir.

$$g(\beta,\sigma) \propto \sigma^{-K-\overline{v}-1} exp\left\{-\frac{1}{2\sigma^2}\left[\overline{v}\bar{s}^2 + \left(\beta-\bar{\beta}\right)'A\left(\beta-\bar{\beta}\right)\right]\right\} \tag{2.128}$$

Eşitlik 2.128'in σ parametresine göre integralinin alınması ile β parametresi için marjinal önsel yoğunluk fonksiyonu elde edilecektir. Bu fonksiyonun dağılımı "çok değişkenli t dağılımına" uymaktadır ve Eşitlik 2.129'da belirtildiği gibidir (Judge *et al.* 1985).

$$g(\beta) \propto \left[1 + \frac{1}{\overline{v}}\left(\beta-\bar{\beta}\right)'\frac{A}{\bar{s}^2}\left(\beta-\bar{\beta}\right)\right]^{-(K+\overline{v})/2} \tag{2.129}$$

Bu denkleme bağlı olarak β parametre vektörünün bireysel ögelerinin her biri $v = T - K$ serbestlik dereceli "tek değişkenli t" dağılımına uymaktadır (Genç vd. 2010).

Hakkında bilgi edinilmek istenilen parametrelerin önsel dağılımının oluşturulmasından sonraki en önemli aşama sonsal dağılımın oluşturulmasıdır. Bayes teoremi yardımı ile sonsal dağılım, yukarıda elde edilen ortak önsel yoğunluk fonksiyonu ($g(\beta,\sigma)$) ile olabilirlik fonksiyonunun ($l(\beta,\sigma \mid y)$) birbiri ile çarpılarak birleştirilmesi ile oluşturulur. Elde edilen bu fonksiyon "Birleşik sonsal olasılık yoğunluk fonksiyonu" olarak adlandırılmaktadır (Judge *et al.* 1985).

Eşitlik 2.116 ile Eşitlik 2.128'in çarpılması ile oluşacak birleşik sonsal olasılık yoğunluk fonksiyonu Eşitlik 2.130a-2.130c'de adımsal olarak ifade edilmiştir.

$$g(\beta,\sigma \mid y) \propto \sigma^{-T-K-\overline{v}-1}$$
$$exp\left\{-\frac{1}{2\sigma^2}\left[\overline{v}\bar{s}^2 + \left(\beta-\bar{\beta}\right)' A\left(\beta-\bar{\beta}\right) + (y-X\beta)'(y-X\beta)\right]\right\} \quad (2.130a)$$

$$g(\beta,\sigma \mid y) = \sigma^{-T-K-\overline{v}-1}$$
$$exp\left\{-\frac{1}{2\sigma^2}\left[\overline{v}\bar{s}^2 + \begin{pmatrix} A^{1/2}\bar{\beta} - A^{1/2}\beta \\ y - X\beta \end{pmatrix}' \begin{pmatrix} A^{1/2}\bar{\beta} - A^{1/2}\beta \\ y - X\beta \end{pmatrix}\right]\right\} \quad (2.130b)$$

$$g(\beta,\sigma \mid y) = \sigma^{-T-K-\overline{v}-1} exp\left\{-\frac{1}{2\sigma^2}\left[\overline{v}\bar{s}^2 + (w-W\beta)'(w-W\beta)\right]\right\} \quad (2.130c)$$

A simetrik bir matris ve $A = A^{1/2}A^{1/2}$ olmak üzere Eşitlik 2.131 yazılabilir.

$$w = \begin{pmatrix} A^{1/2}\bar{\beta} \\ y \end{pmatrix} \qquad W = \begin{pmatrix} A^{1/2} \\ X \end{pmatrix} \quad (2.131)$$

Eşitlik 2.130c'ye göre Bayes regresyon modeli için sonsal dağılıma ilişkin parametre tahmin edicileri Eşitlik 2.132-2.134'de verildiği gibidir.

$$\bar{\bar{\beta}} = (W'W)^{-1}W'w = (A + X'X)^{-1}(A\bar{\beta} + X'Xb) \quad (2.132)$$

$$\begin{aligned} \bar{\bar{v}}\bar{\bar{s}}^2 &= \bar{v}\bar{s}^2 + \left(w - W\bar{\bar{\beta}}\right)'\left(w - W\bar{\bar{\beta}}\right) \\ &= \bar{v}\bar{s}^2 + y'y + \bar{\beta}'A\bar{\beta} - \bar{\bar{\beta}}'(X'X)\bar{\bar{\beta}} \end{aligned} \tag{2.133}$$

$$\bar{\bar{v}} = T + \bar{v} \tag{2.134}$$

Aynı zamanda Eşitlik 2.116’dan, Eşitlik 2.118’in elde edilişi gibi, Eşitlik 2.130c’de önsel bilgiye ait değerler yerine yazılarak β parametresine ilişkin sonsal dağılım tahmin edicisi, ortalama ve varyans değeri bulunabilmektedir. Bu ifadeye göre Eşitlik 2.135-2.137 yazılabilir (Judge *et al*. 1985).

$$\begin{gathered} g(\beta, \sigma \mid y) \propto \sigma^{-\bar{\bar{v}}-K-1} \\ exp\left\{-\frac{1}{2\sigma^2}\left[\bar{v}\bar{s}^2 + \left(w - W\bar{\bar{\beta}}\right)'\left(w - W\bar{\bar{\beta}}\right) + (\beta - \bar{\bar{\beta}})'W'W(\beta - \bar{\bar{\beta}})\right]\right\} \end{gathered} \tag{2.135}$$

$$\begin{gathered} g(\beta, \sigma \mid y) \propto \sigma^{-K} \\ exp\left[-\frac{1}{2\sigma^2}\left(\beta - \bar{\bar{\beta}}\right)'(A + X'X)\left(\beta - \bar{\bar{\beta}}\right)\right].\sigma^{-(\bar{\bar{v}}+1)}exp\left(-\frac{\bar{\bar{v}}\bar{\bar{s}}^2}{2\sigma^2}\right) \end{gathered} \tag{2.136}$$

$$g(\beta, \sigma \mid y) \propto g(\beta \mid \sigma, y).g(\sigma \mid y) \tag{2.137}$$

Elde edilen Birleşik son olasılık yoğunluk fonksiyonunun dağılımı da “Normal-Gamma”dır. Çünkü σ biliniyorken, β parametresi için koşullu yoğunluk fonksiyonu $\bar{\bar{\beta}}$ ortalamalı, $\sigma^2(A + X'X)^1$ kovaryans matrisi ile “çok değişkenli normal” dağılıma sahiptir ve σ parametresi için de marjinal sonsal yoğunluk fonksiyonu $\bar{\bar{v}}$ ve $\bar{\bar{s}}^2$ parametreleri ile “Ters-Gamma” dağılımına uymaktadır (Judge *et al*. 1985).

β parametresine ait marjinal sonsal yoğunluk fonksiyonu olarak ifade edilen Eşitlik 2.138 “Çok değişkenli t” dağılımına uymaktadır ve Eşitlik 2.136’nın σ parametresine göre integral alınarak elde edilir (Judge *et al*. 1985).

$$g(\beta \mid y) = \propto \left[1 + \frac{1}{\bar{\bar{v}}}(\beta - \bar{\bar{\beta}})' \frac{(A + X'X)}{\bar{\bar{s}}^2}(\beta - \bar{\bar{\beta}})\right]^{-(K+\bar{\bar{v}})/2} \quad (2.138)$$

β parametresine ait tek bir eleman için, örneğin β_1, "tek değişkenli t" dağılımına sahip olacaktır.

$$g(\beta_1 \mid y) = \propto \left[1 + \frac{1}{\bar{\bar{v}}} \frac{\beta_1 - \bar{\bar{\beta}}_1}{\bar{\bar{s}}\sqrt{c^{11}}}\right]^{-(1+\bar{\bar{v}})/2} \quad (2.139)$$

Eşitlik 2.139'da $\bar{\bar{\beta}}_1$, $\bar{\bar{\beta}}$ parametre vektöründeki ilk elemandır ve c^{11} ise $(A + X'X)^{-1}$ matrisindeki ilk çapraz (diagonal) elemandır.

Elde edilen Birleşik sonsal olasılık yoğunluk fonksiyonu ile parametreler hakkındaki mevcut bilgi durumunu (önsel bilgi ve örneklem bilgisi) temsil eder ve istenirse nokta tahminleri, aralık tahminleri veya hipotez testleri uygulanabilir. Ayrıca sonsal dağılımın ortalaması olan $\bar{\bar{\beta}}$,önsel dağılımın ortalama değeri $\bar{\beta}$ ve örneklem tahmin edicisi β 'nın ağırlıklı ortalaması olarak kabul edilmektedir (Judge *et al*. 1985).

3. MATERYAL ve METOT

Bu çalışmada, Bayesci yaklaşım, Bayes Regresyon ve Tip II Regresyon konularında literatür bilgilerine yer verilerek, gerçek veriler üzerinde Bayes-Açıortay tekniği ile farklı örneklem hacimlerinde regresyon modeli oluşturularak, HKO ve Göreli Etkinlik (GE) kriterlerine göre performansının değerlendirilmesi amaçlanmıştır. Bu amaçla, Saraçlı ve Çelik'in (2012) çalışmalarında 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 Açıortay regresyon denklemlerinin HKO ve GE değerlerinin hesaplanması amacı ile sırasıyla örneklem hacimleri 100, 75, 50 ve 30 olarak seçilmiştir. 100 birimlik veri seti, 30 önsel bilgi verisi ve 70 birimlik mevcut veri, 40 birimlik önsel bilgi verisi ve 60 birimlik mevcut veri ve en son 50 birimlik önsel bilgi verisi ve 50 birimlik mevcut veri olmak üzere 3 şekilde ayrılmıştır. 100 birimlik veri setinin ilk 75 gözlem değeri aynı şekilde 15-60, 25-50 ve 35-40, ilk 50 gözlem değeri 10-40, 20-30 ve 25-25, 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 literatür bilgilerinde de verildiği üzere, 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 3.1 ile, σ parametresine ait ortalama ve varyans değerleri ise Eşitlik 3.2 ile hesaplanmıştır.

$$E[\beta \mid \sigma] = E[\beta] = \bar{\beta} \qquad cov[\beta \mid \sigma] = A^{-1}\sigma^2 \qquad A = (X'X)^{-1}X' \qquad (3.1)$$

$$E[\sigma] = \frac{\Gamma[(\overline{v}-1)/2]}{\Gamma(\overline{v}/2)}\left(\frac{\overline{v}}{2}\right)^{1/2} \bar{s} \qquad\qquad E[\sigma^2] = \frac{\overline{v}\bar{s}^2}{\overline{v}-2} \qquad (3.2)$$

Burada $\bar{\beta}$ ve $\bar{s}$ değerleri, önsel verilere ait EKK tekniği ile elde edilen tahmin değerleri ve σ^2 , önsel verilerle oluşturulan regresyon denklemine 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 literatür bölümünde yer alan, 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 Açıortay 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 3.3' de belirtildiği gibi hesaplanmıştır.

$$\hat{\beta}_1 = \frac{n\sum x_i y_i - \sum(x_i)\sum(y_i)}{n\sum x_i{}^2 - \sum(x_i)^2} \qquad\qquad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1\,\bar{x} \qquad (3.3)$$

Oluşturulan Bayes(Y/X) ve Bayes(X/Y) regresyon doğrularının eğim katsayıları sırası ile a ve b olarak alınarak Bayes-Açıortay regresyon doğrusunun eğim katsayısı ve sabit katsayı ve de bu katsayılara bağlı Bayes-Açıortay regresyon denklemi Eşitlik 3.4, 3.5 ve 3.6'da belirtildiği gibi hesaplanmıştır.

$$\hat{\beta}_{AO} = (a+b)^{-1}\left[ab - 1 + \sqrt{(1+a^2)(1+b^2)}\right] \tag{3.4}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_{AO}\,\bar{x} \tag{3.5}$$

$$y_{iEKKAO} = \hat{\beta}_0 + \hat{\beta}_{AO}\,\bar{x}_i \tag{3.6}$$

Elde edilen Bayes(Y/X), Bayes(X/Y) ve Bayes-Açıortay regresyon denklemlerinin mevcut veri seti için sergiledikleri performansları ise Eşitlik 3.7'de verilen HKO kriteri ve Eşitlik 3.8'de verilen GE yardımıyla değerlendirilmiştir.

$$HKO = \frac{\sum\left(Y_i - \hat{Y}_i\right)^2}{n-k} \tag{3.7}$$

$$GE = \frac{Regresyon\ tekniğinin\ HKO\ değeri}{Açıortay\ regresyon\ tekniğinin\ HKO\ değeri} \times 100 \tag{3.8}$$

Göreli etkinlik değeri yorumlanırken, ele alınan regresyon tekniğinin etkinliği açıortay tekniğine göre kıyaslanmış olup, GE değeri 100'den büyük olan tekniklerin performansının açıortay tekniğinden daha yüksek olduğu, 100'den küçük olan tekniklerin performansının ise açıortay tekniğinden düşük olduğu anlamına gelmektedir.

4. BULGULAR

Farklı örneklem hacimlerinde, önsel bilgi verisi ve mevcut veri hacimlerinin farklı olarak seçilmesi ile hesaplanan Bayes(Y/X), Bayes(X/Y) ve Bayes Açıortay regresyon tekniklerine ilişkin sabit katsayı, eğim katsayısı ve HKO değerleri Çizelge 4.1, 4.3, 4.5 ve 4.7'de verilmiştir. Ayrıca Bayes yaklaşımı ile oluşturulan regresyon doğruları için hesaplanan önsel dağılımlara ait parametre değerlerinin ortalama ve varyans değerleri Çizelge 4.2, 4.4, 4.6 ve 4.8'de verilmiştir.

Çizelge 4.1 n=100 iken Bayes(Y/X) ve Bayes(X/Y) regresyon doğruları için hesaplanan önsel dağılıma ait parametrelerin ortalama ve varyans değerleri

		$\boldsymbol{\beta}$		$\boldsymbol{\sigma}$	
		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

Çizelge 4.2 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****		
	$\hat{\boldsymbol{\beta}}_0$	$\hat{\boldsymbol{\beta}}_1$	HKO	$\hat{\boldsymbol{\beta}}_0$	$\hat{\boldsymbol{\beta}}_1$	HKO	$\hat{\boldsymbol{\beta}}_0$	$\hat{\boldsymbol{\beta}}_1$	HKO
Bayes$_{(Y\|X)}$	0,944	0,976	0,009651	0,993	0,969	0,008491	0,972	0,961	0,009083

Çizelge 4.2 (Devam) n=100 iken Bayes-Açıortay, Bayes(Y/X) ve Bayes(X/Y) regresyon teknikleri için hesaplanan sabit katsayı, eğim katsayısı ve HKO değerleri

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

Çizelge 4.1 ve Çizelge 4.2 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'nın 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'nın 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'nın 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'nın 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'nın 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'nın 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.

Çizelge 4.2'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.

Çizelge 4.3 n=75 iken Bayes(Y/X) ve Bayes(X/Y) regresyon doğruları için hesaplanan önsel dağılıma ait parametrelerin ortalama ve varyans değerleri

		$\boldsymbol{\beta}$		$\boldsymbol{\sigma}$	
		Ortalama	**Varyans**	**Ortalama**	**Varyans**
15*	**Bayes(Y/X)**	0,885	0,0031	0,2985	0,009455
	Bayes(X/Y)	1,072	0,0046	0,3338	0,01182
25*	**Bayes(Y/X)**	0,966	0,0032	0,3558	0,0131
	Bayes(X/Y)	0,96	0,0032	0,3558	0,0131
35*	**Bayes(Y/X)**	1,012	0,0021	0,3521	0,01314
	Bayes(X/Y)	0,925	0,0018	0,3771	0,01205

* Önsel bilgi verisi için örneklem hacmi

Çizelge 4.4 n=75 iken Bayes-Açıortay, Bayes(Y/X) ve Bayes(X/Y) regresyon teknikleri için hesaplanan sabit katsayı, eğim katsayısı ve HKO değerleri

	15*-60**			25*-50**			35*-40**		
	$\hat{\beta}_0$	$\hat{\beta}_1$	HKO	$\hat{\beta}_0$	$\hat{\beta}_1$	HKO	$\hat{\beta}_0$	$\hat{\beta}_1$	HKO
Bayes$_{(Y\|X)}$	0,866	0,973	0,011285	0,949	0,971	0,01209	1,002	0,969	0,010411
Bayes$_{(X\|Y)}$	1,040	0,976	0,010252	0,937	0,978	0,012319	0,913	0,979	0,01119
Bayes$_{AO}$	0,770	0,975	0,009633	0,770	0,974	0,009633	0,789	0,974	0,009630

* Önsel bilgi verisi için örneklem hacmi

** Mevcut veri seti için örneklem hacmi

Çizelge 4.3 ve Çizelge 4.4 incelendiğinde, 75 birimlik veri seti 15 birimlik önsel bilgi verisi ve 60 mevcut veri olarak ayrıldığı durumda önsel bilgi verisine ait β parametresinin dağılımı Normal(0,885, 0,0031), σ parametresinin dağılımı Ters-Gamma(0,2985, 0,009455) ve X bağımsız değişkeninin dağılımı Normal(37,03, 0,1742) iken Bayes(Y/X) tekniğinin HKO değeri 0,011285 olarak hesaplanmıştır. Ayrıca önsel bilgi verisine ait β parametresinin dağılımı Normal(1,072, 0,0046), σ parametresinin dağılımı Ters-Gamma(0,3338, 0,01182) ve Y bağımsız değişkeninin dağılımı Normal(36,75, 14,38) iken Bayes(X/Y) tekniğinin HKO değeri 0,010252, Bayes(Y/X) ve Bayes(X/Y) teknikleri ile oluşturulan regresyon doğrularının açıortayı alınarak oluşturulan Bayes-Açıortay tekniğinin HKO değeri ise 0,009633 olarak hesaplanmıştır.

75 birimlik veri seti 25 birimlik önsel bilgi verisi ve 50 birimlik mevcut veri olarak ayrıldığı durumda önsel bilgi verisine ait β parametresinin dağılımı Normal(0,966, 0,0032), σ parametresinin dağılımı Ters-Gamma(0,3558, 0,0131) ve X bağımsız değişkeninin dağılımı Normal(37,03, 0,1510) iken Bayes(Y/X) tekniğine ait HKO değeri 0,01209 olarak hesaplanmıştır. Buna ek olarak önsel bilgi verisine ait β parametresinin dağılımı Normal(0,966, 0,0032), σ parametresinin dağılımı Ters-

Gamma(0,3558, 0,0131) ve Y bağımsız değişkeninin dağılımı Normal(36,79, 0,1519) iken Bayes(X/Y) tekniğinin HKO değeri 0,012319, Bayes-Açıortay tekniğinin HKO değeri ise 0,009633 olarak hesaplanmıştır.

75 birimlik veri seti, 35 birimlik önsel bilgi verisi ve 40 birimlik mevcut veri olarak ayrıldığında ise önsel bilgiye verisine ait β parametresinin dağılımı Normal(1,012, 0,0021), σ parametresinin dağılımı Ters-Gamma(0,3521, 0,01314) ve X bağımsız değişkeninin dağılımı Normal(37,12, 0,1616) iken Bayes(Y/X) tekniğinin HKO değeri 0,010411 olarak hesaplanmıştır. Buna ek olarak önsel bilgi verisine ait β parametresinin dağılımı Normal(0,925, 0,0018), σ parametresinin dağılımı Ters-Gamma(0,3771, 0,01205) ve Y bağımsız değişkeninin dağılımı Normal(36,91, 0,1768) iken Bayes(X/Y) tekniğinin HKO 0,01119, Bayes-Açıortay tekniğinin HKO değeri ise 0,009630 olarak hesaplanmıştır.

Çizelge 4.4'e göre Bayes(Y/X) tekniği için hesaplanan HKO değerinin, 75 birimlik veri seti 35*-40** şeklinde ayrıldığında daha düşük olduğu görülmektedir. Bayes(X/Y) tekniğinin HKO değerinin ise 75 birimlik veri seti 15*-60** şeklinde ayrıldığında daha düşük olduğu belirlenmiştir. Bayes-Açıortay tekniğinin HKO değerinin ise 15*-60** ve 25*-50** şeklinde ayrılarak hesaplanan örneklem büyüklüklerinde aynı olduğu, 35*-40** şeklinde ayrılarak hesaplanan örneklem büyüklüklerinde ise HKO değerinin az bir farkla dahi olsa daha düşük olduğu görülmektedir. Genel olarak inceleme yapıldığında ise 15*-60**, 25*-50** ve 35*-40** için hesaplanan Bayes-Açıortay tekniğinin HKO değerlerinin, Bayes(Y/X) ve Bayes(X/Y) tekniği için hesaplanan HKO değerlerine göre daha düşük olduğu görülmektedir.

Çizelge 4.5 n=50 iken Bayes(Y/X) ve Bayes(X/Y) regresyon doğruları için hesaplanan önsel dağılıma ait parametrelerin ortalama ve varyans değerleri

		$\boldsymbol{\beta}$		$\boldsymbol{\sigma}$	
		Ortalama	**Varyans**	**Ortalama**	**Varyans**
10*	**Bayes(Y/X)**	0,865	0,0126	0,1162	0,0147
	Bayes(X/Y)	1,019	0,0175	0,1263	0,0173

Çizelge 4.5 (Devam) n=50 iken Bayes(Y/X) ve Bayes(X/Y) regresyon doğruları için hesaplanan önsel dağılıma ait parametrelerin ortalama ve varyans değerleri

20*	**Bayes(Y/X)**	0,943	0,0043	0,119	0,0146
	Bayes(X/Y)	0,975	0,0047	0,1235	0,0158
25*	**Bayes(Y/X)**	0,959	0,0036	0,1181	0,0143
	Bayes(X/Y)	0,959	0,0036	0,1181	0,0143

* Önsel bilgi verisi için örneklem hacmi

Çizelge 4.6 n=50 iken Bayes-Açıortay, Bayes(Y/X) ve Bayes(X/Y) regresyon teknikleri için hesaplanan sabit katsayı, eğim katsayısı ve HKO değerleri

	10*-40**			**20*-30****			**25*-25****		
	$\hat{\beta}_0$	$\hat{\beta}_1$	HKO	$\hat{\beta}_0$	$\hat{\beta}_1$	HKO	$\hat{\beta}_0$	$\hat{\beta}_1$	HKO
Bayes$_{(Y\|X)}$	0,720	0,976	0,010362	0,915	0,971	0,010889	0,935	0,971	0,012592
Bayes$_{(X\|Y)}$	0,799	0,983	0,010175	0,941	0,979	0,010417	0,936	0,979	0,010716
Bayes$_{AO}$	0,569	0,979	0,009882	0,736	0,975	0,009887	0,736	0,975	0,009887

* Önsel bilgi verisi için örneklem hacmi

** Mevcut veri seti için örneklem hacmi

Çizelge 4.5 ve Çizelge 4.6 incelendiğinde, 50 birimlik veri seti 10 birimlik ön bilgi verisi ve 40 birimlik mevcut veri olarak ayrıldığında, 10 birimlik 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) tekniğinin HKO değeri 0,010362olarak 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 0,010175 olarak hesaplanmıştır.

Ayrıca Bayes-Açıortay regresyon doğrusunun HKO değerinin 0,009882 olduğu görülmektedir.

50 birimlik veri seti, 20 birimlik önsel bilgi verisi ve 30 birimlik mevcut veri olarak ayrıldığında, önsel bilgi verisine ait β parametresinin dağılımı Normal(0,943, 0,0043), σ parametresinin dağılımı Ters-Gamma(0,119, 0,0146) ve X bağımsız değişkeninin dağılımı Normal(37,07, 0,1529) iken Bayes(Y/X) tekniğinin HKO değeri 0,010889 olarak hesaplanmıştır. Önsel bilgi verisine ait β parametresinin dağılımı ise Normal(0,975, 0,0047), σ parametresinin dağılımı Ters-Gamma(0,1235, 0,0158) ve Y bağımsız değişkeninin dağılımı Normal(36,82, 0,1479) iken ise Bayes(X/Y) tekniğinin HKO değeri 0,010417 olarak hesaplanmıştır. Ayrıca Bayes-Açıortay tekniğinin HKO değerinin 0,009887 olduğu görülmektedir.

50 birimlik veri seti 25 birimlik önsel bilgi verisi ve 25 birimlik mevcut veri olarak ayrıldığında ise önsel bilgi verisine ait β parametresinin dağılımı Normal(0,959, 0,0036), σ parametresinin dağılımı Ters-Gamma(0,1181, 0,0143) ve X bağımsız değişkeninin dağılımı Normal(37,05, 0,1433) iken Bayes(Y/X) tekniğinin HKO değeri 0,012592 olarak hesaplanmıştır. Önsel bilgi verisine ait β parametresinin dağılımı Normal(0,959, 0,0036), σ parametresinin dağılımı Ters-Gamma(0,1181, 0,0143) ve Y bağımsız değişkeninin dağılımı Normal(36,82, 0,1432) iken Bayes(X/Y) tekniğinin HKO değeri 0,010716 olarak hesaplanmıştır. Ayrıca Bayes-Açıortay tekniğinin HKO değerinin ise 0,009887 olarak hesaplandığı görülmektedir.

Çizelge 4.6 genel olarak incelendiğinde, 50 birimlik veri seti 10*-40** şeklinde ayrıldığında Bayes(Y/X), Bayes(X/Y) ve Bayes-Açıortay tekniklerinin HKO değerlerinin, 20*-30** ve 25*-25** şeklinde ayrılarak elde edilen Bayes(Y/X), Bayes(X/Y) ve Bayes-Açıortay regresyon tekniklerinin HKO değerlerine göre daha düşük olduğu görülmektedir. Bayes-Açıortay tekniğinin HKO değerinin ise 20*-30** ve 25*-25** şeklinde ayrılarak hesaplanan örneklem büyüklüklerinde aynı olduğu, 10*-40** şeklinde ayrılarak hesaplanan örneklem büyüklüklerinde ise HKO değerinin az bir farkla dahi olsa daha düşük olduğu görülmektedir. Ayrıca 10*-40**, 20*-30** ve 25*-25** için hesaplanan Bayes-Açıortay tekniğinin HKO değerlerinin, Bayes(Y/X) ve

Bayes(X/Y) teknikleri için hesaplanan HKO değerlerine göre daha düşük olduğu görülmektedir.

Çizelge 4.7 n=30 iken Bayes(Y/X) ve Bayes(X/Y) regresyon doğruları için hesaplanan önsel dağılıma ait parametrelerin ortalama ve varyans değerleri

		$\boldsymbol{\beta}$		$\boldsymbol{\sigma}$	
		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

Çizelge 4.8 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****		
	$\hat{\boldsymbol{\beta}}_0$	$\hat{\boldsymbol{\beta}}_1$	HKO	$\hat{\boldsymbol{\beta}}_0$	$\hat{\boldsymbol{\beta}}_1$	HKO	$\hat{\boldsymbol{\beta}}_0$	$\hat{\boldsymbol{\beta}}_1$	HKO
$\mathbf{Bayes_{(Y\|X)}}$	0,621	0,978	0,013328	0,725	0,975	0,013033	0,864	0,973	0,019648
$\mathbf{Bayes_{(X\|Y)}}$	0,813	0,984	0,012726	0,796	0,984	0,012786	1,042	0,976	0,0163
$\mathbf{Bayes_{AO}}$	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

Çizelge 4.7 ve Çizelge 4.8'e göre, 30 birimlik veri seti 5 birimlik önsel bilgi verisi ve 25 birimlik mevcut veri olarak ayrıldığına 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 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.

Çizelge 4.8’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.

Çizelge 4.9 Bayes Yaklaşımı ile oluşturulan regresyon denklemlerinin GE değerleri

		Bayes$_{(Y\|X)}$	**Bayes$_{(X\|Y)}$**
n=100	**30*-70***	114,78	113,02
	40*-60*	101,03	108,58
	50-50*	108,07	116,96
n=75	**15*-60***	117,14	106,42
	25*-50*	125,51	127,88
	35*-40*	108,106	116,20
n=50	**10*-40***	104,87	102,97
	20*-30*	110,14	105,36
	25*-25*	127,36	108,37
n=30	**5*-25***	104,11	99,40
	10*-20*	101,84	99,92
	15*-15*	153,68	127,49

* Önsel bilgi verisi için örneklem hacmi

** Mevcut veri seti için örneklem hacmi

Çizelge 4.9'a göre, Bayes(Y/X) ve Bayes(X/Y) tekniklerinin performansının, farklı örneklem ve kullanılan farklı önsel bilgi ve mevcut verilere ait örneklem hacimlerinde Bayes Açıortay tekniğine göre daha yüksek olduğu görülmektedir. Ayrıca Bayes(X/Y) tekniğinin performansının, örneklem hacminin 30 olduğu durumda 5*-25** ve 10*-20** olarak ayrılmış veri setlerinde Bayes açıortay tekniğine göre daha düşük olduğu tespit edilmiştir.

5. TARTIŞMA 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.

Bayes-Açıortay tekniğinin performansının değerlendirilmesi amacı ile yapılan analizler sonucunda regresyon doğrularının HKO ve GE 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 gözlemlenmiştir. Ayrıca örneklem hacmi küçüldükçe hesaplanan tüm regresyon doğrularının HKO değerlerinin düştüğü tespit edilmiştir.

Çalışmada önsel bilgi ve mevcut verilerin örneklem hacimlerinin artırılıp azaltılması sonucunda her bir veri seti için hesaplanan regresyon doğrularının HKO değerlerine ilişkin genel bir yargıdan söz edilememektedir. 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.

Bayes Regresyon için doğrusal regresyon modeli kullanılmıştır. Regresyon modelinin oluşması aşamasında ele alınan gerçek veri setinin normal dağılıma uygunluk göstermesi ve matematiksel açıdan daha kolay uygulanabilirlik sağlaması nedeni ile önsel dağılım olarak “ Doğal Eşlenik Önsel Dağılım” seçilmiştir. Doğal Eşlenik önsel dağılım için ise literatürde belirtildiği üzere σ parametresinin dağılımı Ters-Gamma, β parametresinin dağılımı Normal dağılım ve olabilirlik fonksiyonunun dağılımı ise

Normal olarak seçilmiştir. Bayes-Açıortay tekniği ile regresyon doğrusunun hesaplanması seçilen bu dağılımlara ilişkin ortalama, standart hata ve varyans değerlerinin hesaplanarak analize dahil edilmesi ile gerçekleştirilmiştir.

Literatürde yer alan diğer Bayes yaklaşımı çalışmaları incelenecek olursa:

Genç vd. (2010)'da Lineer olmayan Bayesci regresyon ile tarım alanında yaptıkları bir çalışma ile doğrusal olmayan regresyon modeli ile kurulan modele ilişkin parametre tahminlerini Bayesci yaklaşım ve En Küçük kareler yöntemi ile elde ederek, iki yöntemi karşılaştırmışlardır. Elde edilen sonuçlara göre iki yöntemin birbirine yakın sonuçlar verdiğini ve Bayesci yaklaşımın tarımsal verilerde kullanılabileceğini belirtmişlerdir.

Ekici (2009) yaptığı bir çalışmada Bayesci yaklaşım ve Klasik yaklaşıma ilişkin istatistiki açıdan kavramsal farklılıklarını ortaya koymuş ve ayrıca Bayesci yaklaşımın objektif ve sübjektif yorumlarına değinmiştir.

Cengiz vd. (2012)'de yaptıkları bir çalışmada gerçek veriler için Lojistik regresyon için parametre tahmin yöntemi olarak Bayesci yaklaşımı kullanmışlardır. Ayrıca Bayesci yaklaşım için sonsal dağılım belirleme yöntemleri olan Gibbs örnekleme algoritması ve Metropolis-Hastings algoritmalarını kullanarak Bayesci yaklaşım ile Klasik yaklaşımın karşılaştırmalarını yapmışlardır. Sonuç olarak AIC ve BIC (Bayesci Bilgi Kriteri) kriterlerine göre Bayesci yaklaşımın çok daha iyi sonuçlar verdiğini gözlemlemişelerdir.

Cengiz vd. (2013) uyguladıkları bir çalışmada doğrusal regresyon modeli için Markov Zinciri Monte Carlo yakınsama kriterlerini Bayesci yaklaşım ve Klasik yaklaşıma göre değerlendirmişlerdir. Sonuç olarak AIC ve BIC bilgi kriterlerine göre Klasik ve Bayes çoklu regresyon modelleri için, Bayes yaklaşımın çok daha üstün olduğunu belirtmişlerdir.

Akar ve Gündoğdu (2014)'de yaptıkları bir çalışmada Bayesci yaklaşımın su ürünleri alanındaki kullanım olanaklarını incelemişlerdir. Balıkçılık alanı ile ilgili olan boy ve ağırlık verileri için doğrusal regresyon modeli oluşturarak parametre ve güven aralığı tahminlerini Bayesci yaklaşım ve EKK yöntemi ile belirlemişlerdir. Yapılan çalışmanın analiz sonuçlarında Bayesci yaklaşım ile elde edilen sonuçların daha uygun ve güvenilirliğinin daha yüksek olduğunu elde etmişlerdir.

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 sağlamaktır. Tip II regresyon tekniklerinden biri olan Açıortay tekniğinde parametre tahmin yöntemi olarak EKK tekniği kullanılmaktadır. Bu çalışmada ise, EKK tekniği yerine, Bayes yaklaşımı aracılığıyla parametre tahmini yapılarak ilk defa Bayes-Açıortay tekniği ortaya konmuştur. Ayrıca ortaya konulan bu tekniğin literatüre önemli bir katkı sağlayacağı öngörülmektedir.

6. KAYNAKLAR

Acıtaş, Ş. (2014). İki Fazlı Doğrusal Regresyon Modelinde Değişim Noktasının Dayanıklı Tahmini. Doktora Tezi, Anadolu Üniversitesi, Fen Bilimleri Enstitüsü, Eskişehir.

Adcock, R.J. (1878). A Problem in Least Squares. *Annals of Mathematics,* **5(2):** 53-54.

Akar, M. ve Gündoğdu, S. (2014). Bayes Teorisinin Su Ürünlerinde Kullanım Olanakları. *Journal of Fisheries Science,* **8(1):** 8-16.

Akdi, Y. (2014). Matematiksel İstatistiğe Giriş. Gazi Kitabevi, Ankara.

Akkoca, G. (2012). Klinik Araştırmalarda Sürekli Sonuçlu Ölçüm Tekniklerinin Uyumunun İncelenmesinde Kullanılan İstatistiksel Yöntemler. Yüksek Lisans Tezi, Ankara Üniversitesi, Sağlık Bilimleri Enstitüsü, Ankara.

Alkan, N. (2012). Kayıp Verili Cox Regresyon Yöntemine Bayesci Bir Yaklaşım. Doktora Tezi, Ondokuz Mayıs Üniversitesi, Fen Bilimleri Enstitüsü, Samsun.

Alma, Ö. ve Vupa, Ö. (2008). Regresyon Analizinde Kullanılan En Küçük Kareler ve En Küçük Medyan Kareler Yöntemlerinin Karşılaştırılması. *SDÜ Fen Edebiyat Fakültesi Fen Dergisi,* **3(2):** 219-229.

Altındağ, İ. (2015). Bayesci Doğrusal Olmayan Yapısal Eşitlik Modeli. Doktora Tezi, Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Konya.

Amman, L. and Ness, J.V. (1988). A Routine for Converting Regression Algorithms in to Corresponding Orthogonal Regression Algorithms. *ACM Transactions on Mathematical Software,* **14(1):** 76-87.

Anscombe, F. J. (1961). Bayesian Statistics. *The American Statistician*, **15(1):** 21-24.

Avcı, E. (2012). Bayesci Sağkalım Analizi ve Meme Kanseri Verileri Üzerine Bir Uygulama. Doktora Tezi, Mimar Sinan Güzel Sanatlar Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.

Bolstad, W. M. (2007). Introduction to Bayesian Statistics. Wiley – Interscience, 2. edition, Hamilton, New Zealand.

Box, G.E. and Tiao, G.C. 1973. Bayesian Inference In Statistical Anlysis. Addison Wesley Publishing Company, USA.

Carlin, B.P. and Louis, T.A. (2000). Bayes Empirical Bayes Methods for Data Analysis. Chapman & Hall/CRC, Florida.

Carroll, R.J. and Ruppert, D. (1994). The Use and Misuse of Orthogonal Regression in Linear Errors in Variables Models. *The American Statistician,* **50(1):** 1-6.

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

Cengiz, M.A., Şenel, T., Terzi, E. ve Terzi, Y. (2013). Doğrusal Regresyonda Markov Zincirleri Monte Carlo Yakınsama Kriterlerinin Karşılaştırılması. *International Anatolia Academic Online Journal.* **1(1):** 27-33.

Chib, S. (2001). Markov Chain Monte Carlo Methods: Computation and Inference. In: Heckman,J.J, and Leamer E., (Eds), Handbook Econometrics. Elseiver Science B.V, Volume 5, Chapter 57, St. Louis, USA.

Congdon, P. (2003). Applied Bayesian Modelling. John Willey&Sons, LTD, England.

Çelebioğlu, N. ve Yaman, A.C. (2005). Tanrı'nın Olasılığı. Yedinci kapı yayınları, 1. Baskı, İstanbul.

Çevik, M. (2009). Doğrusal Olmayan Bayesci Regresyon ve Yüksek Frekanslı Ses Sistemlerinde Bir Uygulama. Yüksek Lisans Tezi, Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.

Çiftci, F. (2009). Regresyonda Alternatif Metotlar. Doktora Tezi, Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Konya.

Demirci, M. (2016). Bayes Teoremi ve Teoremin İşletme Bölümünde Uygulamaları. *The Journal Of Academic Social Science Studies***, 43:** 439-462.

Ekici, O. (2005). Bayesyen Regresyon ve Win BUGS ile Bir Uygulama. Yüksek Lisans Tezi, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.

Ekici, O. (2009). İstatistikte Bayesyen ve Klasik Yaklaşımın farklılıkları. *Balıkesir Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, **12(21):** 89-101.

Fox J. (1997). Applied Regression Analysis: Linear Models and Related Methods. Sage Publication, USA.

Fuller, W. A., (1987), Measurement Error Models. John Wiley&Sons, New York.

Gasım, N. (2013). Bayesyen Model İle Doğrusal Regresyon Modellerinin Karşılaştırılması Üzerine Bir Uygulama. Yüksek Lisans Tezi, Dokuz Eylül Üniversitesi, Sosyal Bilimler Enstitüsü, İzmir.

Gelman, A., Carlin, J.B., Stern, H.S, Dunson, D.B., Vehtari, A. and Rubin, D.B., (1995). Bayesian Data Analysis. A Chapman, Hall Book, 3. edition, New York, US.

Geman, S. and Geman, D. (1984). Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of İmages. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6:** 721-741.

Gencer, G. (2016). Bazı Dağılımlar İçin En Çok Olabilirlik ve Kayıp Fonksiyonları Altında Bayes Tahmin Edicilerinin Performanslarının Karşılaştırılması. Yüksek Lisans Tezi, Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Konya.

Genç, A., Karadavut, U. and Palta, Ç (2010). Lineer Olmayan Bayesci Regresyon ve Tarım Alanında Bir Uygulama. *TÜBAV Bilim Dergisi*. **3(3):** 250-258.

Gilks, W.R., Richardson, S. and Spiegelhalter, D.J. (1996). Introducing Markov Chain Monte Carlo. Champan and Hall, London.

Gill, J. (2002*)*. Bayesian methods a social and behavioral sciences approach. CRC Press, 1. edition, Florida.

Gill, J. (2008). Bayesian Method: A Social and Behavioral Sciences Approach. Chapman&Hall/CRC, 2. edition, USA.

Gosh, J.K., Delampady, M. and Samanta, T. (2006). An Introduction to Bayesian Analysis Theory and Methods. Springer Science+Business Media, LLC, New York, USA.

Gündoğdu, S. (2016). Balıklarda Büyüme Parametrelerinin Bayesyen İstatistiksel Yöntemle Tahmini. Doktora Tezi, Çukurova Üniversitesi, Fen Bilimleri Enstitüsü, Adana.

Güner, A. (2014). Bayesci Yaklaşımda Eşlenik Aileleri Önseli ile Jeffreys Önselinin Karşılaştırılması. Yüksek Lisans Tezi, Osmangazi Üniversitesi, Fen Bilimleri Enstitüsü, Eskişehir.

Isobe, T., Feigelson, E.D., Akrıtas, M.G. and Babu, G.J. (1990). Linear Regression in Astronomy I. *The Astrophysical Journal*. **364:** 104-113.

Jeffreys, H. (1961). The Theory of Probability. Oxford University, 3. edition, Oxford, U.K.

Judge, G., Grıffıths, W.E., Hıll, R.C., Lütkepohl, H. and Lee., T., (1985). The Theory and Practice Of Econometrics. John Willey&Sons, 2. edition, Canada.

Karadağ, Ö. (2011). Bayesci Hiyerarşik Modeller. Yüksek Lisans Tezi, Hacettepe Üniversitesi, Ankara.

Karagöz, K. (2004). Öngörü ve Zaman Serileri Analizinde Bayesyen Yaklaşım. Doktora Tezi, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.

Keleş, T. ve Altun, M. (2016). Comparison of Classical Linear Regression and Orthogonal Regression with Respect to the Sum of Squared Perpendicular Distances. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, **7(2):** 296-308.

Kendall, M. and Buckland, W. (2009). Introduction to Bayesian Anlysis, Appendix 2, Draft version.

Kurtuluş, K. (1998). Pazarlama Araştırmaları, Avcıol Basım, 6. baskı, İstanbul.

Li, J. (2012). Multivariate Generalization of Reduced Major Axis Regression. PhD Thesis, Arizona State University, Arizona.

Linnet, K. (1990). Estimation of the Linear Relationship between the Measurements of Two Methods with Proportional Errors. *Statistics in Medicine*, **9(12):** 1463-1473.

Linnet, K. (1993). Evaluation of Regression Procedures for Methods Comparison Studies. *Clinical Chemistiry*, **39(3):** 424-432.

Linnet, K., (1998). Performance of Deming regression analysis in case of misspecified analytical error ratio in method comparison studies. *Clinical Chemistry*, **44(5):** 1024-1031.

Mesple, F., Trousellier, M., Casellas, C. and Legendre, P. (1996). Evaluation of Simple Criteria to Qualify A Simulation. *Ecological Modeling*, **88:** 9-18.

Miller, I. and Miller, M. (2014). John E. Freund's Mathematical Statistics with Applications. Pearson Education Limited, 5. edition, USA.

Murat, N. (2012). Yapısal Eşitlik Modellerde Parametre Tahminlerinde Klasik ve Bayesci Bir Yaklaşım. Doktora Tezi, Ondokuz Mayıs Üniversitesi ,Fen Bilimleri Enstitüsü, Samsun.

Ntzoufras, L. (2009). Bayesian Modeling Using Winbugs. A John Willey & Sons, Inc., Publication, Hoboken, New Jersey, Canada.

Öztürk, F., Akdi, Y., Aydoğdu H. ve Karabulut, İ. (2006). Parametre Tahmini ve Hipotez Testi. Bıçaklar Kitabevi, Ankara.

Öztürk, L. (2003). Doğrusal Regresyonda Sağlam Kestirim Yöntemleri ve Karşılaştırılmaları. Doktora Tezi, Mimar Sinan Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.

Öztürk, Z. (2014). Genelleştirilmiş Lineer Karma Modellerde Bayesci Yaklaşımın Kullanımı. Doktora Tezi, Ondokuz Mayıs Üniversitesi, Samsun.

Parmigiani, G., Inoue, L.Y.T. and Lopes, H.F. (2009). Decision Theory Principles and Approaches, A John Willey and Sons, Ltd, Publication, USA.

Parvin, C.A. (1984). A Direct Comparison of Two Slope-Estimation Techniques Used in Method Comparison Studies. *Clincal Chemistry.*, **30(5):** 751-754.

Passing, H. and Bablok, W. (1983). A New Biometrical Procedure for Testing the Equality of Measurements from Two Different Analytical Methods. Application of linear regression procedures for method comparison studies in Clinical Chemistry, Part I. *Journal of Clinical Chemistry & Clinical Biochemistry,* **21:** 709-720.

Pearson, K. (1901). On Lines and Planes of Closest Fit to Systems of Points in Space. *Philosophical Magazine,* **2:** 559-572.

Rawlings, J.O., Pantula,S.G. and Dickey, D.A. (2001). Applied Regression Anlysis: A Research Tool. Springer Science&Business Media, 2. edition,USA..

Robert, C.P., Chopin, N. and Rousseau, J. (2009). Harold Jeffrey's Theory of Probability Revisited. *Statistical Science,* **24(2):**141-172.

Saçaklı, S.İ. (2011). Gelişmiş ve Gelişmekte Olan Piyasalarda Hisse Senedi Getiri Volatilitelerinin Bayes Yaklaşımı İle Analizi. Doktora Tezi, Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.

Saraçlı, S. (2008). Ölçüm Hatalı Modellerde Doğrusal Regresyon Tekniklerinin Karşılaştırılması-Monte Carlo Simülasyon Çalışması. Doktora Tezi, Eskişehir Osmangazi üniversitesi, Fen Bilimleri Enstitüsü, Eskişehir.

Saraçlı, S. (2011). Tip II Regresyon Tekniklerinin Monte-Carlo Simülasyonu ile Karşılaştırılması. e-Journal of New World Sciences Academy**. 6(2):** 26-35.

Saraçlı, S. ve Çelik, E. (2012). Metot Karşılaştırma Çalışmalarında Bland-Altman ve Tip II Regresyon Analizinin Karşılaştırılması. Düzce Üniversitesi Sağlık Bilimleri Enstitüsü Dergisi, **2(1):** 11-14.

Saraçlı, S., Doğan, İ. ve Doğan, N. (2009). Medikal Metod Karşılaştırma Çalışmalarında Deming Regresyon Tekniği. *Türkiye Klinikleri Biyoistatistik Dergisi,* **1(1):** 9-15.

SAS. (2009). SAS/STAT 9.2 User's Guide. Cary, NC, 7745, 2. edition, USA.

Saylor, R.D., Edgerton, E.S. and Hartsell, B.E. (2006). Linear regression techniques for use in the EC tracer method of secondary organic aerosol estimation, *Atmospheric Environment,* **40:** 7546-7556.

Sokal, R.R. and Rohlf, F.J. (1995). Biometry. W.H Freeman Company, 3. edition, New York.

Stigler, S. M. (1983). Who Discovered Bayes's Theorem. *The American Statistician,* **37(4):** 290-296.

Sunar, Ö. (2009). Ortalamalardaki Değişimi İzleyen Bayesci Süreç Kontrolü. Doktora Tezi, Hacettepe Üniversitesi, Ankara.

Tiao B. and Zellner, A. (1964). Bayes's theorem and the use of prior knowledge in regression analysis. *Biometrika,* **51(1/2):** 219-230.

Turanlı, M. ve Güriş, S. (2005). Temel istatistik. Der Yayınları, 4. Baskı, İstanbul.

Türkay, H. (2004). Doğrusal Regresyon Modellerinin Robust (Dayanıklı) Yöntemlerle Tahmini ve Karşılaştırılmalı Uygulamaları. Doktora Tezi, İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul.

Uludağ, S. (2013). Denetçi Yargısının Nesnelleştirmesinde Bayes Yaklaşımı Ve Bir Uygulama. Doktora tezi, Anadolu Üniversitesi, Sosyal Bilimler Enstitüsü, Eskişehir.

Wille, F.J. (2003). Auditing Using Bayesian Decision Analysis. PhD Thesis, Vrije University, Amsterdam.

Wu, C. and Yu, J.Z. (2018). Evaluation of Linear Regression Techniques for Atmospheric Applications: The Importance of Appropriate Weighting. *Atmospheric Measurement Techniques,* **11:** 1233-1250.

Yardımcı, A. (2000). Doğrusal Regresyonda Değişken Seçimine Bayesci Yaklaşımların Karşılaştırılması. Doktora Tezi, Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, Ankara.

York, D. (1969). Least squares fitting of a straight line with correlated errors. *Earth and Planetary Letters*, **5:** 320-324.

York, D., Evensen, N.M., Martinez, M.L. and Delgado, D.B. (2004). Unified equations for the slope, intercept, and Standard errors of the best straight line. *American Journal of Physics*, **72(3):** 367-375.

Yurtçu, M. (2018). Parametrik Olmayan Bayes Yöntemiyle Ortak Değişkenlere Göre Yapılan Test Eşitlemelerinin Karşılaştırılması. Doktora tezi, Hacettepe Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.

Zerey, G. (2018). Klasik Bayesci Portföy Seçim Modellerinin Bist 30 Üzerinde Uygulanması. Doktora Tezi, Ondokuzmayıs Üniversitesi, Fen Bilimleri Enstitüsü, Samsun.

İnternet Kaynakları

1) http://ncss.wpengine.netdna-cdn.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Deming_Regression.pdf, 20.04.2019
2) http://ncss.wpengine.netdna-cdn.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Passing-Bablok_Regression_for_Method_Comparison.pdf, 28.04.2019
3) https://help.xlstat.com/customer/en/portal/articles/2062362-run-passing-bablok-regression-to-compare-methods?b_id=9283, 15.04.2019
4) https://www.xlstat.com/en/solutions/features/passing-and-bablok-regression, 15.04.2019.
5) http://alfresco.ubm-us.net/alfresco_images/pharma/2014/08/22/fc6cdb11-0339-480a-bb9d-48c5782bce9d/article-7276.pdf, 17.04.2019

ÖZGEÇMİŞ

Adı Soyadı : Ece Özgören

Doğum Yeri ve Tarihi : Ankara-09.09.1993

Yabancı Dili : İngilizce

İletişim (Telefon/E-posta) : ecem_ozgoren@hotmail.com

Eğitim Durumu (Kurum ve Yıl)

Lise : Afyon Kocatepe Anadolu Lisesi (2007-2011)

Lisans : Ankara Üniversitesi (2011-2016)

Yüksek Lisans : Afyon Kocatepe Üniversitesi (2017-)

Çalıştığı Kurum/Kurumlar ve Yıl:

Yayınları (SCI ve diğer):

Özgören E., Saraçlı S., *Comparison of Bayes and Ordinary Least Squares Based Type II Regression*, 11th International Statistics Days Conference, 3 - 7 October 2018 Muğla, Turkey.

Özgören E., Saracli S., Tunca B**.**, Performance of Type II Regression in Time Series Analysis, *International Conference on Data Science, Machine Learning and Statistics*, 26-29 June 2019, Van/Turkey.

Diğer Konular:

EKLER

Ek.1. Çalışmada kullanılan ve 100 hastadan derlenen koltuk altı (aksiller) ve kulaktan ateş ölçer aletleri yardımı ile elde edilen gerçek veriler aşağıdaki tabloda verilmiştir.

X(Koltuk Altı)	Y(Kulak İçi)	X(Koltuk Altı)	Y(Kulak İçi)	X(Koltuk Altı)	Y(Kulak İçi)
37	36,7	37,2	37	36,6	36,5
36,5	36,3	37,2	37	36,5	36,4
36,7	36,5	37	36,9	37	37
37	36,8	37,6	37,4	37,2	37,1
37,5	37	37	36,8	36,9	36,8
37,4	37,2	37,3	37,2	36,8	36,7
37,2	37	36,9	36,7	37,3	37
37,2	37	36,5	36,5	38,5	38,3
36,8	36,6	36,7	36,6	37	36,9
36,9	36,5	37,5	37,4	36,6	36,5
36,7	36,5	36,8	36,6	36,5	36,4
38	37,6	36,7	36,5	36,8	36,5
37,5	37,1	37	36,8	37,1	37
36,5	36,2	37,6	37,4	37	37
36,6	36,3	36,8	36,7	37,5	37,4
36,8	36,5	37	36,8	37,2	37
37,2	37	37,2	37	36,7	36,5
37,3	37,3	36,3	36,1	37,3	37,1
37,6	37,4	36,5	36,5	37	36,8
37,1	37	37,1	37	36,7	36,5
36,7	36,5	37	36,8	37,4	37,3
37,5	37,4	36,8	36,7	37	37
36,8	36,6	36,4	36,4	36,8	36,7
36,7	36,5	36,4	36,4	36,7	36,5
37,1	37	36,5	36,4	37	36,9
36,9	36,7	36,8	36,5	37,4	37,2
37	36,9	37,3	37,1	37,6	37,5
37,2	37,2	37,5	37,3	37,2	37
37,4	37,2	38	37,9	36,9	36,7
37	36,8	37,9	37,7	37,4	37,3
37,5	37,4	37,4	37,3	37	36,8
37,5	37,3	37	36,8	36,6	36,5
37,8	37,7	36,7	36,5	37,2	37
38	37,8				