ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

Nimet TÜRKER

LİNEER REGRESYON MODELİNDE BAYES TAHMİN EDİCİLER

İSTATİSTİK ANABİLİM DALI

ADANA, 2013

ÖZ

YÜKSEK LİSANS TEZİ

LİNEER REGRESYON MODELİNDE BAYES TAHMİN EDİCİLER

Nimet TÜRKER

ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
İSTATİSTİK ANABİLİM DALI

Danışman : Prof. Dr. Selahattin KAÇIRANLAR

Yıl: 2013, Sayfa: 76

Jüri : Prof. Dr. Selahattin KAÇIRANLAR

: Yrd. Doç. Dr. Gülesen ÜSTÜNDAĞ ŞİRAY

: Prof. Dr. H. Altan ÇABUK

Çoklu iç ilişki problemini gidermek için en küçük kareler (EKK) tahmin ediciye alternatif olarak birçok yanlı tahmin edici önerilmiştir. Bunlardan biri Bayes tahmin edicidir. Bu çalışmada Bayesyen analizi gerçekleştirmede temel yapıların neler olduğuna ve nasıl elde edildiğine yer verilmiştir. Ayrıca, lineer regresyon modelinde Bayesyen tahmin yöntemleri incelenmiştir. Çalışmanın ilerleyen kısımlarında Bayesyen yaklaşım gecikmesi dağıtılmış modellere taşınarak, bu modellerde Liu ve ridge tipi tahmin ediciler ele alınmıştır. Ridge-tipi Bayesyen tahmin ediciler sınıfına ait olan Shiller ve Lindley tahmin ediciler, Gruber (2012) tarafından tanımlanan Liu-tipi Bayesyen tahmin ediciyle birlikte kullanılarak iki tane yeni Liu-tipi Bayesyen tahmin edici elde edilmiştir.

Anahtar Kelimeler: En Küçük Kareler Tahmin Edici, Bayesyen Tahmin Edici, Ridge-tipi Tahmin Edici, Liu-tipi Tahmin Edici, Bayesyen Liu-tipi Tahmin Edici

ABSTRACT

PhD THESIS

BAYES ESTIMATORS IN LINEAR REGRESSION MODEL

Nimet TÜRKER

ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES
DEPARTMENT OF STATİSTİCS

Supervisor : Prof. Dr. Selahattin KAÇIRANLAR

Year: 2013, Pages: 76

Jury : Prof. Dr. Selahattin KAÇIRANLAR

: Asst. Prof. Dr. Gülesen Üstündağ ŞİRAY

: Prof. Dr. H. Altan ÇABUK

To remedy the problem of multicollinearity some biased estimators are proposed is an alternative to the ordinary least squares (OLS) estimator. One of these estimators is Bayes estimator. In this study, the main structures and methods to carry out the Bayesian analysis are indicated. Also, Bayesian estimation methods in linear regression model are investigated. Subsequently, Liu and ridge type estimators are examined by carrying the Bayesian approach to the distributed lag models. Two new Liu-type Bayesian estimators are obtained by using Shiller and Lindley estimator, which are belong to Ridge-type Bayesian estimators, with Liu-type Bayesian introduced by Gruber (2012).

Key Words: Least Squares Estimator, Bayesian Estimator, Ridge-type Estimator, Liu-type Estimator, Bayesian Liu-type Estimator

TEŞEKKÜR

Tez çalışmam süresince hem bir bilim insanı olarak, hem de hayat yaklaşımıyla kendime örnek edindiğim, bilgisini ve deneyimlerini her zaman çok cömertçe paylaşan saygıdeğer hocam Prof. Dr. Selahattin KAÇIRANLAR'a en içten teşekkürlerimi sunarım.

Beni bugünlere getiren, hiçbir zaman sevgisini ve desteğini esirgemeyen, hayatın zorluklarına karşı cesaretli ve kararlı bir duruş sergilemeyi öğreten, her başarımı benden daha çok hak eden sevgili annem Sevgi TÜRKER'e, sevgili babam Cemal Ata TÜRKER'e ve sevgili abim Mesut TÜRKER'e teşekkür ederim.

Birlikte yürüdüğümüz yolda bana her zaman destek olan ve hayatımın sonuna kadar da destek olacağını bildiğim, çok sevdiğim eşim Alper ÖZBAY'a ve kıymetli ailesine teşekkür ederim.

Bütün çalışma arkadaşlarıma ve hayatımı güzelleştiren sevgili dostlarıma teşekkür ederim.

İÇİNDEKİLER **SAYFA**

ÖZ I
ABSTRACT II
TEŞEKKÜR III
İÇİNDEKİLER ..... … IV
SİMGELER VE KISALTMALAR VI
1. GİRİŞ 1
2. BAYESYEN ANALİZ 7
2.1. Bayes Teoremi 7
2.2. Bayesyen Tahmin 8
2.3. Ön Dağılımın Belirlenmesi 12
2.3.1. Bilgi Vermeyen Ön Dağılımlar 12
2.3.1.1. Dikdörtgen Ön Dağılım 13
2.3.1.2. Jeffreys 'in Ön Dağılımı 13
2.3.1.3. Referans Ön Dağılım 14
2.3.1.4. Belirsiz Ön Dağılım 15
2.3.2. Eşlenik Ön Dağılım 15
2.3.3. Maksimum Entropi Ön Dağılımı 17
2.3.4. Hiyerarşik Ön Dağılım 18
3. REGRESYONUN BAYESYEN ANALİZİ 19
3.1. Stokastik Kısıtlı Mixed Regresyon Modeli 19
3.2.Eşlenik Ön Dağılım Varlığında Bayesyen Analiz 22
3.3. Belirsiz Ön Dağılım Varlığında Bayesyen Analiz 27
3.4. Bayesyen Regresyon Tahmin Edici 28
3.5. Bayesyen Tahmin Edicilerin Özellikleri 29
3.6. Lindley 'in Bayesyen Regresyon Modeli 33
4. LİNEER BAYES TAHMİN EDİCİ 38
4.1. Temel Bilgiler 38

4.2. EKK Tahmin Edici ve Özellikleri 41
4.3. Bayes Tahmin Edici 44
4.4. Bayes Tahmin Edicinin Alternatif Formları 51
4.4.1. X Matrisi Tam Ranklı İse 51
4.4.2. X Matrisi Tam Ranklı Değilse 54
5. GECİKMESİ DAĞITILMIŞ MODELLERDE BAYESYEN LİU VE RİDGE TİPİ TAHMİN EDİCİLER 58
5.1. Giriş 58
5.2. Stokastik Olmayan Ön Bilgi Altında Tahmin 60
5.3. Stokastik Ön Bilgi Altında Tahmin 63
5.3.1. Stokastik Ön Bilginin Bayesyen Yorumu 64
5.4. Bayesyen Liu Tipi Tahmin Ediciler 67
6. SONUÇLAR VE ÖNERİLER 72
KAYNAKLAR 73
ÖZGEÇMİŞ 76

SİMGELER VE KISALTMALAR

BLUE : En İyi Lineer Yansız Tahmin Edici (Best Linear Unbiased Estimator)

EKK : En Küçük Kareler (Ordinary Least Squares)

MLE : En Çok Olabilirlik Tahmin Edici (Maximum Likelihood Estimator)

MCMC : Markov Zinciri Monte Carlo Simülasyonu (Monte Carlo Markov Chain)

GLS : Genelleştirilmiş En Küçük Kareler (Generalized Least Squares)

MSE : Hata Kareler Ortalaması (Mean Squared Error)

RLS : Kısıtlı En Küçük Kareler (Restricted Least Squares)

1. GİRİŞ

İstatistikte temel olarak iki farklı felsefi yaklaşımın ortaya çıktığı görülmektedir. Bunlar klasik yaklaşım ve Bayesyen yaklaşımdır. İstatistikte pek çok konu ve kavramın ele alınışında bu yaklaşımlar birbirine alternatif olmuştur. Zaman içinde, bu konuda çalışanların fikirleri nedeniyle ve belki de anlaşılması daha kolay olduğundan, istatistik alanında klasik yaklaşım daha fazla kullanılmaya başlanmıştır.

İstatistiğe Bayesyen yaklaşımında, konu ile elde edilebilen tüm bilgilerin analizlerde kullanılması gerekmektedir. Kişisel deneyimler, daha önceki benzer araştırmalardan elde edilmiş sonuçlar ve örneklem bilgisi Bayesyen yaklaşımla yapılan analizlerde kullanılabilir. Bu bölümde, Bayesyen yaklaşımın gelişimiyle ilgili tarihsel sürece yer verilmiştir. Ayrıca, bazı istatistiksel kavramlar Bayesyen yaklaşım düşüncesiyle açıklanmıştır.

Bayesyen yaklaşım, özellikle son yıllarda, istatistiğin hemen hemen tüm alanlarında ve pek çok ekonometrik çalışmada yoğun bir biçimde kullanılmaktadır. Thomas Bayes'in koşullu olasılıkların hesaplanması için önerdiği teorem, Bayes yaklaşımının temelini oluşturmaktadır. Bu teoreme göre A olayının gerçekleşmesi olasılığı bilindiğinde, B olayının da gerçekleşmesi olasılığını hesaplamak mümkündür. Uzun süre yukarıda bahsedilen amaç için kullanılan Bayes teoremi, daha sonra karar verme tekniklerinde uygulanmaya başlamıştır. Karar verme tekniklerinde temel amaç yanlış karar verme riskini olabildiğince küçük yapmak olduğundan, kişisel deneyim ve bilgilerin de karar verme sürecine katılması mantıklı bir yaklaşım olmaktadır. Böylece, yapılan istatistiksel çalışmalarda araştırmacı tarafından kişisel deneyimlere sık sık başvurulmaktadır.

Regresyon tekniklerinde daha etkili parametre tahminleri yapmak ve en önemlisi rezidü kareler ortalamasını küçültmek gibi amaçlarla Bayesyen yaklaşım kullanılarak çalışmalar yapılmıştır. Yapılan bu çalışmalarında, kişisel bilgilerin ne olacağı ve bu bilgilerin olasılık dağılımı biçiminde nasıl ifade edileceği belirtilmemiş, genelde kolaylıkla oluşturulan dağılımlar kullanılmıştır. Kişisel bilgileri dağılımlar yardımıyla ifade etmenin yolları da yapılan araştırmanın yapısına ve araştırmacıya göre değişebilmektedir.

Bayesyen yaklaşım ilk defa, İngiltere 'de yaşayan bir rahip ve matematikçi olan Thomas Bayes (?-1761) tarafından yazılan ve ölümünden birkaç yıl sonra arkadaşı Richard Price 'ın bulup yayınladığı bir denemeyle("An Essay Towards Solving a Problem in the Doctrine of Changes") ortaya konmuştur. Bu denemeyle bugünkü Bayes dünyasının çekirdeği olan Bayes teoremi bulunmuştur.

18. yüzyılın sonlarından 20. yüzyılın başlarına kadar istatistiksel çalışmaların Bayesyen yaklaşımın etkisinde olduğu görülmektedir. 1764 'te Bayes teoremi bu çalışmalara katkıda bulunurken, yakın dönemde bağımsız bir şekilde başlayıp devam eden daha detaylı ve özgün analizleriyle Laplace etkili olmuştur. Ayrıca 1900 'lerin ortalarında H. Jeffreys objektif yaklaşımlı Bayesyen analizi geliştirmiş ve Bayesyen regresyon konusunda ilk çalışmayı yapmıştır. 1950 'lerden sonra Bayesyen model seçimi ve hipotez testi geliştirilmiştir. Böylelikle teorik altyapı daha ayrıntılı yapılandırılmıştır. Fakat bazı matematiksel yapıların çözüme kavuşturulamaması önemli bir engel oluşturmuştur ve model seçiminde elde edilen bazı son (posterior) yapıların integralleri çözülememiştir. Ancak W.K. Hastings, Markov Zinciri Monte Carlo tekniğinin istatistik alanındaki uygulamalarını geliştirmiştir ve bu sorun aşılmıştır. Bu nedenle artık modern Bayesyen istatistik, simülasyon tekniklerine bağlı olan uygulamalarla kendini daha iyi açıklamaktadır. Bilgisayar teknolojisinin gelişmesiyle Bayesyen istatistiğin matematikle ilgili alanlarında karşılaşılan güçlükler de ortadan kalkmıştır.

Örneklem bilgisi ve ön bilginin kullanılmasıyla lineer regresyon modellerinde Bayesyen yaklaşım ile alternatif tahmin yöntemleri ele alınmıştır. Lindley ve Smith (1971) lineer modeller için Bayes tahminleri uygulamalarıyla incelemişlerdir. Zellner (1971) ekonometride Bayesyen sonuç çıkarımı ele almıştır. Farklı modeller altında Bayesyen tahminleri incelemiştir. Vinod ve Ullah (1981) regresyonun Bayesyen analizini ele alarak ridge tahmin edicinin Bayesyen ve Bayesyen olmayan yorumlarını vermişlerdir. Shu-Lu (1993) Stein tipinin Bayes tahmin edicilerini ve minimax tahmin edici ailesinin Bayes analizini ele almıştır. Kubokawa ve Srivastava lineer regresyon modelinde çoklu iç ilişki olması durumunda, temel bileşenler regresyon ve ridge regresyonun minimax tahmin edicisinin Bayes analizini incelemiştir. Yine, Kubokawa ve Srivastava çok değişkenli lineer regresyon

modellerinde, çoklu iç ilişki olması durumunda minimax tahmin edicinin Bayesyen analizini ele almıştır. Burada normallik varsayımı altında Bayes ridge regresyon tahmin ediciler 3 farklı shrinkage tipi fonksiyonla incelenmiştir. Ekici (2005) Bayesyen regresyon ve WinBUGS ile bir uygulama adlı yüksek lisans tezi ile Bayesyen regresyonu incelemiş ve bir uygulama ile desteklemiştir. Giles (2010) eşlenik (bileşik) ön dağılımın kullanılmasıyla lineer regresyonda Bayes tahminleri inceleyerek, Bayes ve kısıtlı Bayes tahmin edicilerin karşılaştırılmalarını ele almıştır. Elde edilen sonuçlar ridge regresyona uygulanmıştır. Gruber (2012) lineer Bayes, mixed ve minimax tahmin edicileri inceleyerek ridge tipi tahmin edicilerin ve Özkale ve Kaçıranlar (2007) tarafından tanımlanan iki parametreli tahmin ediciler sınıfının bazı üyelerinin birer Bayes, mixed ve minimax tahmin ediciler olduğunu göstermiştir.

Bayesyen istatistiğe göre bir teori veya önerme ile ilgili belirsizlik olasılıklarla ifade edilmelidir. Klasik yaklaşım belirsizliklere deterministik davranır. Varsayımlar doğrultusunda söz konusu belirsizliği, orada iddia edilen ilişkiyi, sıklıklarına göre değerlendirerek kabul edilmesi ya da edilmemesi yönünde karar verir. Bayesyen yaklaşımı benimseyenler, belirsizlikle ifade edilen bir teoriyi kabul etmenin veya körü körüne reddetmenin bilgi oluşturma sürecine pek bir katkı sağlamayacağı, aksine önemli sayılabilecek yanılsamalara yol açacağı konusunda eleştiriler getirebilirler.

Bayesyen istatistikte bir olayın olasılığı, o olaya ilişkin ön bilgi (prior) ile denemeden elde edilen sonuçların (verinin) birleştirilmiş halidir. Bir araya getirme işlemi, Bayes teoremine dolayısıyla da koşullu olasılığa dayanmaktadır.

Bayesyen yaklaşımda parametre olasılık dağılımı olan bir rasgele değişken gibi düşünülmektedir. Bu doğrultuda parametre için bir ön olasılık dağılımı belirlenir ve mevcut veri ile birleştirilerek parametrenin son olasılık dağılımı (posterior) elde edilir. Bayesyen yaklaşımda parametre ile ilgili tüm çıkarımlar son dağılıma dayanarak yapılır.

İstatistiğin temel konularından biri nokta tahminidir. Bayesyen yaklaşımda ilgilenilen parametrenin nokta tahmini genellikle son dağılımın ortalamasıdır (posterior mean).

İlgilenilen parametreye ilişkin mümkün olan tüm bilgiyi içeren istatistik, yeterli istatistiktir. Yeterlilik klasik yaklaşımda olduğu gibi Bayesyen yaklaşımda da oldukça önemlidir. Çünkü ön dağılımın elde edilmesinde bu özellikten yararlanılır. Aslında yeterlilik klasik yaklaşım biçimiyle pek uyumlu değildir. Çünkü yeterliliğin özünde parametrenin koşullu olasılığı vardır. Bir parametrenin olasılığının olması da klasik görüşe zıttır.

Yansızlık, tahmin edicinin beklenen değerinin, parametrenin gerçek değerine eşit olması demektir. Klasik yaklaşımda iyi bir tahmin edici yansızdır. Öte yandan Bayesyen tahmin edici (son dağılımın ortalaması) yanlıdır. Bayesyen analizde bu sonucun değeri yoktur. Çünkü Bayesyen yaklaşım en başta parametrenin tanımını farklı yapmaktadır. Tahmini Bayesyen yaklaşıma göre elde edip, klasik yaklaşımın kriterlerine göre değerlendirmek pek doğru değildir.

Genellikle yanlı tahmin edicilerin yansız tahmin edicilere göre hata kareler ortalaması daha küçüktür. Örneğin, ridge tahmin edici de yanlı bir tahmin edicidir ama bazı k değerleri için yansız tahmin edicininkinden daha küçük bir hata kareler ortalaması vardır. Şunu da söylemek gerekir ki tahminde karşılaşılan bazı sorunlar yanlı bir tahmin edici kullanıldığında ortadan kalkabilmektedir. Kısaca yansızlık, Bayesyen yaklaşım açısından bakıldığında hiç de önemli olmayan bir değerlendirme ölçütüdür.

Tutarlılık kriteri tahmin edicinin limitteki özelliğidir. Örneklem büyüklüğü arttıkça tahmin edicinin değeri, gerçek parametre değerine yaklaşır. Klasik yaklaşımda tutarlılık oldukça önemli bir özelliktir. Bayesyen yaklaşımda ise tutarlılığın ilgilenilmesi gereken bir kriter olup olmadığı tartışılmaktadır.

Kesinlik, Bayesyen yaklaşım için önemli bir kriterdir. Kesinlik (precision), varyans ile ters orantılıdır. Tam bir Bayesyen analizde varyans yerine kesinlik alınarak, gerçek değerin olasılığı hesaplanır.

İstatistikte ilgilenilen temel konulardan biri de aralık tahminidir. Klasik yaklaşımda aralık tahminleri yalnızca örneklem bilgisine dayanılarak yapılmaktadır. Ancak Bayesyen yaklaşımda, klasik aralık tahmini yorumlanırken de parametre bir rasgele değişken olarak kabul edilir. Örneğin, son dağılımın ortalaması için 0,95 güven düzeyinde bir güven aralığından söz ediliyorsa, "bu aralığın son dağılımın

ortalamasını içermesi olasılığı %95′tir" şeklinde yorum yapılır. Bayesyen yaklaşımda bu aralığın içinde bulunan her noktanın sahip olduğu olasılık yoğunluğunun, aralık dışında bulunan noktaların sahip olduğu olasılık yoğunluğundan daha büyük olduğu belirtilmiştir (Box and Tiao, 1973). Bayesyen yaklaşım, klasik yaklaşımın "güven aralığı" (confidence interval) tanımlaması yerine, "güvenilir aralık" (credible interval), "Bayesyen aralık" (Bayesian interval) veya "en yüksek son yoğunluk bölgesi veya aralığı" (highest posterior density region or interval) tanımlamalarını kullanmaktadır.

İstatistikte kitleyi anlamaya ve çözmeye çalışırken parametrelerle ilgili iddiaların yer aldığı bazı önermelerden yararlanılır. Hipotez olarak adlandırılan bu önermelerde ortaya atılanlar, araştırmacının karar vermesini sağlayan bir test sürecinden geçer. Hipotez testi sürecinde parametrenin iddia edilen değere eşit olup olmadığı veya sözü edilen aralıkta bulunup bulunmadığı test edilir ve böylelikle kitle ile ilgili belirsizlik bir parça aydınlatılmış olur.

Klasik yaklaşımda hipotez testi için birinci ve ikinci tip hatalar kullanılır. Birinci tip hata, H_0 doğru iken H_1′in kabul edilmesi; ikinci tip hata, H_0 yanlış iken H_0′ın kabul edilmesidir.

Bayesyen yaklaşım, hipotezlere olasılık tayin eder. Bu, Bayesyen yaklaşımın olasılık teorisine daha geniş bir mantıksal açıdan bakması nedeniyledir. Ancak bu noktada belirtmek gerekir ki, Bayesyen yaklaşım çift taraflı hipotez testi söz konusu ise zayıftır. Çünkü sürekli bir dağılımda parametrenin "0" a eşit olma olasılığı "0" dır. Buna önerilen çözüm, örneğin regresyon katsayısı için "0" a yakın bir değer almaktır.

Bayesyen istatistikte hipotez testi yapmak için izlenen iki yol vardır. İlki, hipotezler ile önsel olasılıkların birleştirilerek sonsal olasılıkların bulunmasıdır. Hipotezler,

$$H_0: \ \theta = \theta_0 \qquad H_1: \ \varphi = \varphi_1$$

biçiminde verilmiş ise sonsal olasılıkların oranı,

$$K_{01} = \frac{P(H_0/y)}{P(H_1/y)} = \frac{P(H_0)P(y/\theta=\theta_0)}{P(H_1)P(y/\varphi=\varphi_1)} \quad (1.1)$$

ile bulunur. Zellner (1971) 'e göre (1.1) oranının birden büyük olması durumunda H_0 hipotezi reddedilemez. İkinci olarak Zellner (1971), hipotezin reddedilip reddedilmeyeceğini kayıp fonksiyonu açısından incelemiştir. Sonsal bilgiler yardımı ile bulunan H_0 ve H_1 hipotezlerinin reddedilmemesi için beklenen kayıplar $E(L/H_0)$ ve $E(L/H_1)$ olmak üzere,

$$E(L/H_0) \leq E(L/H_1)$$

oluyorsa H_0 hipotezinin reddedilemeyeceği belirtilmiştir.

2. BAYESYEN ANALİZ

Bu bölümde Bayesyen yaklaşımın temelini oluşturan, Bayes Teoremi açıklanmıştır. Daha sonra Bayesyen analizi gerçekleştirmede temel yapıların neler olduğuna ve nasıl elde edildiğine yer verilmiştir. Son olarak Bayesyen yaklaşım için oldukça önemli olan ön dağılımın belirlenmesine ilişkin tekniklere değinilmiştir.

2.1. Bayes Teoremi

Bayes Teoremi koşullu olasılık tanımından yola çıkarak açıklanır. Bir örneklem uzayında A ve B gibi iki olay için, B verilmişken A 'nın gerçekleşmesi koşullu olasılığı,

$$P(A/B) = P(A \cap B)/P(B) \quad , \quad P(B) > 0 \tag{2.1}$$

şeklindedir. Benzer şekilde,

$$P(B/A) = P(A \cap B)/P(A) \quad , \quad P(A) > 0 \tag{2.2}$$

şeklindedir. (2.1) ve (2.2) ifadeleri birlikte,

$$P(A/B)P(B) = P(A \cap B) = P(B/A)P(A) \tag{2.3}$$

şeklinde yazılabilir.

(2.3) 'de sağlanan sonuç genelleştirilerek ifade edilirse Bayes Teoremi elde edilir. A, aynı anda gerçekleşmesi mümkün olmayan $A_1, A_2, \dots, A_k$ ayrık olaylarından oluşan bir örneklem uzayı olsun. Bu örneklem uzayında bir de B olayı olsun. Buna göre B verilmişken, herhangi bir A_i olayının gerçekleşmesi olasılığı,

$$P(A_i/B) = P(A_i)P(B/A_i)/P(B) \tag{2.4}$$

ile ifade edilir. (2.4) 'de paydadaki marjinal dağılım $P(B)$ aşağıdaki gibi yazılabilir:

$$P(B) = P(A_1)P(B/A_1)+\ldots+P(A_k)P(B/A_k) = \sum_{j=1}^{k} P(A_j)P(B/A_j)$$

Eşitliklerin düzenlenmesiyle,

$$P(A_i/B) = P(A_i)P(B/A_i)/\sum_{j=1}^{k} P(A_j)P(B/A_j) \quad , \quad (i = 1,2,\ldots,k) \quad (2.5)$$

elde edilir. Böylelikle (2.5) ifadesi Bayes Teoremini gösterir.

Teorem, ön olasılıkların veriden elde edilen bilgilerle güncellendiğini ifade eder. Bu teoremi uygulayabilmek için ön olasılık değerleri olan $P(A_i)$ değerlerini bilmek gereklidir. Değişik bir bakış açısıyla yaklaşılırsa teoremin, ilgilenilen olaya göre, olasılıkların hesabında tersine bir gidişe izin verdiği de söylenebilir.

2.2. Bayesyen Tahmin

Bayes yaklaşımında, (2.1) ve (2.2) 'ye benzer olarak,

$$p(y/\theta) = p(y,\theta)/p(\theta) \quad (2.6)$$

$$p(\theta/y) = p(y,\theta)/p(y) \quad (2.7)$$

temel eşitlikleri kullanılır. Burada θ, hakkında analiz yapılan parametreyi göstermektedir. Bayesyen yaklaşımda, θ parametresi rasgele değişken olarak alındığından (2.6) ve (2.7) ifadeleri (2.1) ve (2.2) ifadeleri ile tutarlı olmaktadır. (2.6) eşitliğinden $p(y,\theta)$ ifadesi (2.7) eşitliğinde yerine yazılarak $p(y) \neq 0$ için,

$$p(\theta/y) = p(\theta)p(y/\theta)/p(y) \quad (2.8)$$

bulunur. Burada $p(y)$ terimine normalleştirme sabiti denir. Bu sabiti ortadan kaldırarak (2.8) eşitliği,

$$p(\theta/y) \propto p(\theta)p(y/\theta) \tag{2.9}$$

biçiminde verilir. Buradaki $\propto$ simgesi, (2.9) ifadesinin sol tarafı ile sağ tarafının orantılı olduğunu belirtmektedir. $p(y/\theta)$, örneklem bilgisini temsil eden olabilirlik fonksiyonunu gösterir. $p(\theta)$, θ parametresi için önsel dağılım olup önsel bilgiye karşılık gelmektedir. Kişinin araştırmaya başlamadan önce sahip olduğu kişisel bilgiler ya da daha önceki benzer çalışmalardan elde edeceği mantıklı tüm bilgiler önsel bilgi (prior information) olarak adlandırılır. Önsel bilgi ile örneklem bilgisinin birleşimi (2.9) eşitliği yardımı ile sağlanarak, elde edilen bilgi de sonsal bilgi (posterior information) olarak adlandırılmaktadır. $p(\theta/y)$, y verildiğinde θ parametresi için sonsal dağılımdır. Bayesyen analizde sonsal bilgiler yapılan çıkarımları elde etmede doğrudan kullanılırlar. Böylece Bayes Teoremi

$$sonsal\ oyf \propto \text{önsel}\ oyf \times olabilirlik\ fonksiyonu \tag{2.10}$$

biçiminde tanımlanabilir.

Parametreyle ilgili çıkarım süreci yaygın olarak kullanılan normal dağılım varsayımı altında şöyle gerçekleşmektedir:

y, normal bir kitleden çekilen n adet bağımsız gözlemi temsil etsin. μ parametresi kitle ortalaması, σ^2 ise kitle varyansı olsun. Burada varyansın bilindiği ve ön dağılımın da normal dağılım olduğu varsayılsın. $\bar{\mu}$ ön dağılımın ortalamasını ve $\bar{\sigma}^2$ ön dağılımın varyansını göstersin. μ 'nün son dağılımı $p(\mu/y,\sigma^2)$,

$$p(\mu/y,\sigma^2) \propto p(\mu)p(y/\mu,\sigma^2)$$

ifadesi yardımıyla bulunur. μ 'nün ön dağılımının $p(\mu)\sim N(\bar{\mu},\bar{\sigma}^2)$ olduğu verilmişti. Öyleyse,

$$p(\mu) = \frac{1}{\sqrt{2\pi\bar{\sigma}}} exp\left[-\frac{1}{2\bar{\sigma}^2}(\mu - \bar{\mu})^2)\right] \quad (2.11)$$

olur. Parametreye ilişkin başlangıç tahminlerinin, $\bar{\mu}$ ve $\bar{\sigma}^2$ ile analize dahil edildiği (2.11) eşitliğinden görülmektedir.

Benzerlik fonksiyonu,

$$p(y/\mu,\sigma^2) = \prod_{i=1}^{n} p(y_i/\mu,\sigma^2) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2\right]$$

şeklindedir. Burada eşitliğe örneklem ortalaması $\hat{\mu}$ 'yı dahil etmek için,

$$p(y/\mu,\sigma^2) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left((y_i-\hat{\mu})-(\mu-\hat{\mu})\right)^2\right]$$

$$= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}[(y_i-\hat{\mu})^2 - 2(y_i-\hat{\mu})(\mu-\hat{\mu}) + (\mu-\hat{\mu})^2]\right] \quad (2.12)$$

düzenlemesi yapılır. Gözlemlerle örneklem ortalamasının farkları toplamı,

$$\sum_{i=1}^{n}(y_i-\hat{\mu})(\mu-\hat{\mu}) = 0$$

bulunur. Bu nedenle benzerlik fonksiyonu,

$$p(y/\mu,\sigma^2) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}[(y_i-\hat{\mu})^2 - n(\mu-\hat{\mu})^2]\right] \quad (2.13)$$

olur. Serbestlik derecesi $v = n - 1$ iken, örneklem varyansı

$$s^2 = \frac{1}{v}\sum_{i=1}^{n}(y_i-\hat{\mu})^2$$

şeklindedir. Bu eşitliklere göre (2.13) eşitliğinde gerekli düzenlemeler yapılırsa benzerlik fonksiyonunun son hali,

$$p(y/\mu,\sigma^2) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n exp\left[-\frac{1}{2\sigma^2}(vs^2 - n(\mu-\hat{\mu})^2)\right] \quad (2.14)$$

olur.

Ön bilgiyi yansıtan ön dağılım (2.11) ve örneklem bilgisini yansıtan benzerlik fonksiyonu (2.14) çarpılırsa son dağılım,

$$p(\mu/y,\sigma^2) = \frac{1}{\sqrt{2\pi}\bar{\sigma}} exp\left[-\frac{1}{2\bar{\sigma}^2}(\mu-\bar{\mu})^2)\right]\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n exp\left[-\frac{1}{2\sigma^2}(vs^2 - n(\mu-\hat{\mu})^2)\right]$$

olur. $\frac{1}{\sqrt{2\pi}\bar{\sigma}}$ ve $\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n$ sabitleri ile $exp\left(-\frac{1}{2\sigma^2}vs^2\right)$ sabiti, oransallık ifadesi kullanıldığında kaybolur ve

$$p(\mu/y,\sigma^2) \propto exp\left\{-\frac{1}{2}\left[\frac{(\mu-\bar{\mu})^2}{\bar{\sigma}^2} - \frac{n}{\sigma^2}(\mu-\hat{\mu})^2\right]\right\} \quad (2.15)$$

elde edilir. (2.15) 'de varyans terimleri bir tarafta, ortalama terimleri diğer tarafta olacak şekilde düzenleme yapılırsa,

$$p(\mu/y,\sigma^2) \propto exp\left[-\left(\frac{\bar{\sigma}^2+\sigma^2/n}{2\bar{\sigma}^2\sigma^2/n}\right)\left(\mu - \frac{\hat{\mu}\bar{\sigma}^2+\bar{\mu}\sigma^2/n}{\bar{\sigma}^2+\sigma^2/n}\right)^2\right] \quad (2.16)$$

elde edilir. (2.16) 'dan görüldüğü gibi son dağılım da normal dağılımdır. Son dağılımın ortalamasının $\bar{\bar{\mu}} = \frac{\hat{\mu}\bar{\sigma}^2+\bar{\mu}\sigma^2/n}{\bar{\sigma}^2+\sigma^2/n}$ olduğu, (2.16) 'dan görülmektedir.

$$\bar{\bar{\mu}} = \frac{\hat{\mu}\bar{\sigma}^2+\bar{\mu}\sigma^2/n}{\bar{\sigma}^2+\sigma^2/n} = \frac{(\hat{\mu}\bar{\sigma}^2)/\bar{\sigma}^2}{(\bar{\sigma}^2+\sigma^2/n)/\bar{\sigma}^2} + \frac{(\bar{\mu}\sigma^2/n)/(\sigma^2/n)}{(\bar{\sigma}^2+\sigma^2/n)/(\sigma^2/n)}$$

$$= \frac{\hat{\mu}}{1+\frac{\sigma^2/n}{\bar{\sigma}^2}} + \frac{\bar{\mu}}{1+\frac{\bar{\sigma}^2}{\sigma^2/n}} \quad (2.17)$$

şeklinde yeniden yazılabilir. Bu sonuca göre, son dağılımın ortalaması ($\bar{\bar{\mu}}$), ön dağılımın ortalaması $\bar{\mu}$ ile örneklem ortalaması $\hat{\mu}$ 'nin ağırlıklı ortalamasıdır.

Parametreyi tanımlamaya yardımcı olan ortalama ve varyans ölçüleri, kesinlikler (precision) anlamında da ifade edilebilir. Kesinlik, Bayesyen yaklaşım için önemli bir kavramdır. Parametrenin tahmininin kesinliğini ifade eder. Kesinlik, varyans ile ters orantılıdır. Bu nedenle varyans azaldıkça kesinlik "1" e yaklaşacaktır. $\bar{h} = \frac{1}{\bar{\sigma}^2}$ ön dağılımın kesinliğini, $h = \frac{1}{\sigma^2/n}$ ise örneklemin kesinliğini gösterir. Son dağılımın ortalaması (2.17) bu kesinliklere göre yazılırsa,

$$\bar{\bar{\mu}} = \hat{\mu}\frac{h}{\bar{h}+h} + \bar{\mu}\frac{\bar{h}}{\bar{h}+h}$$

elde edilir. Bu nedenle son dağılımın ortalaması, kesinliklere göre yorumlanabilir. Ön bilginin kesinliği arttıkça son dağılımın ortalamasında ön dağılımın ortalaması baskın olurken, örneklemin kesinliği arttıkça örneklemden elde edilen ortalama değeri son dağılımın ortalamasında daha baskın olur. Son dağılımın varyansı $\bar{\bar{\sigma}}^2$,

$$Var(\mu) = \left(\frac{\bar{\sigma}^2\sigma^2/n}{\bar{\sigma}^2+\sigma^2/n}\right)$$

olur (Zellner, 1971).

Kesinlikler ile ifade edilirse, $Var(\mu) = \frac{1}{\bar{h}+h}$ şeklindedir. Buna göre son dağılımın kesinliği $\bar{\bar{h}}$, ön dağılımın kesinliği ile örneklemin kesinliğinin toplamıdır. Bu durum, $\bar{\bar{h}} = h + \bar{h}$ şeklinde ifade edilebilir.

Özet olarak, son dağılımın ortalamasının elde edilişinde, ön dağılım ile örneklem bilgisi kesinlikleri ölçüsünde uzlaşır (Ekici, 2005).

2.3. Ön Dağılımın Belirlenmesi

2.3.1 Bilgi Vermeyen Ön Dağılımlar

Bilgi vermeyen ön dağılımların parametreyi açıklama gücü çok azdır. Diğer yandan bu tür ön dağılımlar, Bayesyen yaklaşımla ilgili yapılan sübjektiflik

eleştirilerini yok eder. Genel olarak belirtilmelidir ki bilgi vermeyen ön dağılımlardan hiçbiri, tam bilgisizlik (ignorance) durumunu yansıtmaz.
Bu gruptaki ön dağılımlar şunlardır:

2.3.1.1 Dikdörtgen Ön Dağılım

Dikdörtgen ön dağılımda, belirlenen aralıkta parametreye aynı olasılık değerleri atanır. θ, dikdörtgen ön dağılıma sahip bir parametre ise θ 'nın olasılık yoğunluk fonksiyonu,

$$0 \leq \theta \leq m \text{ iken } p(\theta) = c = \frac{1}{m}$$

şeklindedir. Bu dağılım genellikle parametrenin belirli bir aralıkta yer aldığı, sınırlandırılabildiği ve oran şeklinde ifade edildiği durumlarda kullanılabilir. Dikdörtgen ön dağılım, tam bilgisizlik durumunu yansıtmaz. Örneğin, parametre $[0, m]$ aralığında yer alıyorsa, m sonsuza gittikçe ön dağılım daha az bilgi verir duruma gelir.

Reel eksende $[-\infty, \infty]$ aralığında, θ 'nın tüm değerleri için $p(\theta) = c$ iken, dikdörtgen ön dağılım "uygunsuz" dur (improper). Çünkü, bu aralıkta olasılık yoğunluk fonksiyonunun integrali sonsuz çıkar ve bu da olasılıklar toplamının 1 'e eşit olma aksiyomunu bozar. "Uygun" (proper) dağılımlarda bu toplam 1 'e eşit olur. Uygunsuz ön dağılımlar hesaplamada güçlük çıkarır. Öte yandan, böyle bir ön dağılımdan elde edilen son dağılım uygunsuz olmayabilir.

Dikdörtgen ön dağılımın bir diğer zayıf yanı da "değişmezlik" (invariant) özelliğini taşımamasıdır. Parametre dönüşüme uğradığında elde edilen yeni ön dağılım, dikdörtgen ön dağılım olmayabilir.

2.3.1.2. Jeffreys 'in Ön Dağılımı

Jeffreys, parametre belirli bir aralıkta sınırlı olsun ya da olmasın ($[-\infty, \infty]$ veya $[0, \infty]$), ön dağılımı bir sabite eşitler. Bu durumda da belirlenen dikdörtgen ön

dağılım, uygunsuz olur. Jeffreys, çalışmasının ilerleyen aşamasında $\theta \epsilon R \; ve \; \sigma > 0$ olmak üzere

$$f(y/\theta) = \sigma^{-1} f\left(\frac{(y-\theta)}{\sigma}\right) \tag{2.18}$$

benzerlik fonksiyonu ile

$$I(\theta) = -E_\theta \left[\frac{\partial^2 log f(y/\theta)}{\partial \theta^2}\right] \tag{2.19}$$

Fisher 'in bilgi matrisini kullanır. Bu benzerlik fonksiyonu ile bilgi matrisi,

$$I(\theta) = -E_\theta \begin{bmatrix} \frac{\partial^2}{\partial \theta^2}\left(-log\sigma - \frac{(y-\theta)^2}{2\sigma^2}\right) & \frac{\partial^2}{\partial \theta \partial \sigma}\left(-log\sigma - \frac{(y-\theta)^2}{2\sigma^2}\right) \\ \frac{\partial^2}{\partial \theta \partial \sigma}\left(-log\sigma - \frac{(y-\theta)^2}{2\sigma^2}\right) & \frac{\partial^2}{\partial \sigma^2}\left(-log\sigma - \frac{(y-\theta)^2}{2\sigma^2}\right) \end{bmatrix}$$

$$= -E_\theta \begin{bmatrix} -\frac{1}{\sigma^2} & \frac{2(\theta-y)}{\sigma^3} \\ \frac{2(\theta-y)}{\sigma^3} & \frac{1}{\sigma^2} - \frac{3(y-\theta)^2}{\sigma^4} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{2}{\sigma^2} \end{bmatrix} \tag{2.20}$$

şeklinde elde edilir (James O. Berger, s. 88). Ön dağılım, elde edilen bilgi matrisinin determinantının karekköküdür.

2.3.1.3. Referans Ön Dağılım

Berger-Bernardo "referans ön dağılım" terimini genel anlamında kullanmıştır. Onların geliştirdiği yönteme göre ilgilenilen parametre için referans ön dağılım, son dağılım ile ön dağılım arasındaki Kullback-Leibler uzaklığını maksimize etmekle bulunur.

$Y_1, Y_2, \dots, Y_n$ birbirinden bağımsız ve özdeş dağılan n tane rastgele değişken olmak üzere, ${Y_1}^n = (Y_1, Y_2, \dots, Y_n)$ olsun. $\pi(\theta/{y_1}^n)$ son dağılım, $\pi(\theta)$ ise ön dağılım olmak üzere Kullback-Leibler uzaklığı,

$$K_n\left(\pi(\theta/{y_1}^n),\pi(\theta)\right) = \int log\left[\frac{\pi(\theta/{y_1}^n)}{\pi(\theta)}\right]\pi(\theta/{y_1}^n)d\theta \quad (2.21)$$

şeklindedir (Robert E. Kass, Larry Wasserman, s.1350). $K_n^{\pi} = E\left(K_n\left(\pi(\theta/{y_1}^n),\pi(\theta)\right)\right)$ ifadesi, Kullback-Leibler uzaklığının beklenen değeridir. Amaç bu ifadeyi maksimum yapmak yani $K_{\infty}^{\pi} = \lim_{n\to\infty} K_n^{\pi}$ değerini hesaplamaktır. Ancak genellikle K_{∞}^{π} sonsuzdur. Bu sorunu ortadan kaldırmak için öncelikle sonlu uzaklık olan K_n^{π} 'yi maksimum yapan ön dağılımlar (π_n) bulunur. Daha sonra bu ön dağılımlara karşılık gelen son dağılımların limiti bulunur. Sonuç olarak bu şekilde elde edilen ön dağılım, limitte Bayes teoreminden son dağılım üreten bir dağılım olur.

2.3.1.4. Belirsiz Ön Dağılım

Parametrenin dağılımının türüne göre, bu parametrenin dağılımının konum ve şekil parametresi değişmektedir. Büyük varyansın belirsizliğe eşit olması fikrine dayanarak, bu dağılımların şekil parametrelerine öyle değerler atanır ki oldukça geniş bir aralıkta yer alan, neredeyse dikdörtgen ön dağılım kadar düz, bilgi vermeyen bir ön dağılım oluşturulur (H. Raiffa ve R. Schlaifer, s.63). Örneğin, konum parametresi için normal dağılım ve şekil parametresi için gama dağılımı alınsın.

$p(\theta)\sim N(\bar{\mu},\bar{\sigma}^2)$ olup, $\bar{\sigma}^2$ 'ne büyük değer atanır.

$p(\sigma)\sim G(\alpha,\beta)$ olup, α ve β 'ya (0,001 gibi) çok küçük değerler atanır.

Bu şekilde oluşturulan belirsiz ön dağılımın en güçlü yanı, uygun dağılım olmasıdır. Dikdörtgen ön dağılım yerine belirsiz ön dağılım alınarak, hesaplamalarda kolaylık sağlanır.

2.3.2. Eşlenik Ön Dağılım

Önsel dağılım herhangi bir aileden tanımlanmış olsun. Herhangi n genişlikteki örneklem ve bu örneklemdeki gözlemlerin değerleri ile bulunacak sonsal

dağılım da önsel dağılımın tanımlandığı ailede yer alıyorsa, bu aileye eşlenik (conjugate) aile denilmektedir (De Groot, 1970).

Eşlenik ön dağılım, ön dağılım ile benzerlik fonksiyonu birleştirildiğinde elde edilen son dağılımın da ön dağılım ile aynı fonksiyon formunda olduğu bir dağılımdır. Bu dağılım, benzerlik fonksiyonunun dağılım yapısından faydalanarak belirlenebilir. Benzerlik fonksiyonu yeterli istatistiklerine göre çarpanlarına ayrılır. Hesaplanan yeterli istatistiklerin dağılımları, oluşturulacak eşlenik ön dağılıma temel teşkil eder.

Eşlenik ön dağılım seçmekle, her zaman aynı evrenden yeni örneklem bilgisinin alınması mümkün olmaktadır. Böylece ilgilenilen parametreye ilişkin ön bilgiler yeniden gözden geçirilebilmekte ve daha tutarlı sonuçlara ulaşılmaktadır. Eşlenik ön dağılımla benzerlik fonksiyonundan son dağılımın elde edilmesi, Bayesyen yaklaşım için matematiksel kolaylık sağlamaktadır.

Bazı dağılımlardan alınan örneklemler ve bunlara karşılık gelen sonsal dağılımların yapısı aşağıda kısaca açıklanmıştır.

1) $x_1, x_2, \dots, x_n$ bilinmeyen θ parametreli Bernoulli dağılımından alınan bir rastgele örneklem olsun. θ 'nın önsel dağılımı ise a ve b parametreleri ile Beta dağılımı olsun. $X_i = x_i$ $(i = 1,2, \dots, n)$ için $y = \sum x_i$ olmak üzere θ 'nın sonsal dağılımı da $a + y$ ve $b + n - y$ parametreleri ile Beta dağılımıdır.

2) $x_1, x_2, \dots, x_n$ bilinmeyen ortalamalı Poisson dağılımından alınan bir rastgele örneklem olsun. . θ 'nın önsel dağılımı ise a ve b parametreleri ile Gamma dağılımı olsun. $X_i = x_i$ $(i = 1,2, \dots, n)$ için θ 'nın sonsal dağılımı da $a + \sum x_i$ ve $b + n$ parametreleri ile Gamma dağılımıdır.

3) $x_1, x_2, \dots, x_n$ r ve θ parametreli negatif binom dağılımından alınan bir rastgele örneklem olsun. r, sıfırdan büyük olup özel bir değere sahiptir. θ 'nın ise değeri bilinmemektedir. θ 'nın önsel dağılımı ise a ve b parametreleri ile Beta dağılımı olsun. $X_i = x_i$ $(i = 1,2, \dots, n)$ için θ 'nın sonsal dağılımı $a + rn$ ve $b + \sum x_i$ parametreleri ile Beta dağılımıdır.

4) $x_1, x_2, \dots, x_n$ bilinmeyen θ parametreli üstel dağılımdan alınan bir rastgele örneklem olsun. θ 'nın önsel dağılımı ise a ve b parametreleri ile Gamma dağılımı

olsun. $X_i = x_i$ $(i = 1,2, \dots, n)$ için θ 'nın sonsal dağılımı da $a + n$ ve $b + \sum x_i$ parametreleri ile Gamma dağılımıdır.

5) $x_1, x_2, \dots, x_n$ σ^2 bilinen varyansı ve μ bilinmeyen ortalaması ile normal dağılımdan alınan bir rastgele örneklem olsun. İlgilenilen θ parametresinin $N(\mu_0, \sigma_0^2)$ dağılımına sahip olduğu varsayılırsa, θ parametresine ait sonsal dağılımın da $\mu_1 = \sigma_1^2 \left(\frac{\mu_0}{\sigma_0^2} + \frac{nx}{\sigma^2}\right)$ ortalaması ve $\sigma_1^2 = \sigma_0^{-2} + n\sigma^{-2}$ varyansına sahip normal dağılım olduğu gösterilmiştir (Lindley,1965).

2.3.3. Maksimum Entropi Ön Dağılımı

Entropi, olasılık dağılımının doğası gereği taşıdığı belirsizliği ölçer (Berger, 1985). Başka bir deyişle, gözlemlerin belirsizliğini sayısallaştırır.

Entropi ön dağılımı, ön dağılım parametrelerinin dağılımına ilişkin belirsizliğin nispi seviyesini tanımlar. Birbirinden farklı ön dağılımlar tarafından sağlanan belirsizlik veya mutlak kesinlik aynı şekilde modellenir. Bu nedenle esnek olduğu söylenebilir. Ancak parametre dönüşümü sonucu değişmezlik özelliğine sahip değildir. Dolayısıyla sınırlı uygulama alanı vardır (Berger, 1985).

Kesikli parametre $\theta = (\theta_1, \theta_2, \dots, \theta_n)$ için entropi $H(\theta)$ ise,

$$H(\theta) = -\sum P(\theta_i) log P(\theta_i) \tag{2.22}$$

şeklindedir. Entropinin (2.22) 'deki matematiksel ifadesiyle iki uç durum tespiti yapılabilir. Biri entropinin minimumunu (bilgi olması hali), diğeri maksimumunu (bilgisizliği) ifade eder. Parametrenin belirlenen bir değeri alma olasılığı 1, yani $P(\theta_k) = 1$ ise; $i \neq k$ olmak üzere $P(\theta_i) = 0$ alınır, yani parametrenin belirlenen başka bir değeri alma olasılığına "0" atanır. Burada parametreyle ilgili tam bilgi verilmiş olur ve dolayısıyla belirsizlik "0" olur. Bu durum,

$$H(\theta) = -\sum P(\theta_i) log P(\theta_i) = 0$$

şeklinde ifade edilebilir.

Ön dağılım dikdörtgen ön dağılım olarak alınırsa, örneğin $P(\theta_i) = \frac{1}{n}$ tayin edilirse, bu ön dağılımın entropisi maksimuma ulaşmış olur. Bu da,

$$H(\theta) = -\sum \frac{1}{n} \log\left(\frac{1}{n}\right) = logn \tag{2.23}$$

şeklinde ifade edilebilir. (2.23) 'de elde edilen entropi, bilgi vermeyen ön dağılım ile aynıdır. Parametrenin gözlenen değerlerinin sayısı arttıkça (yani $n \rightarrow$), $P(\theta_i) = \frac{1}{n}$ değeri azalır.

Sözü edilen iki yapı arasında düşünülünce, entropi ön dağılım için $H(\theta) \leq logn$ şeklinde sınır oluşturulur (Berger, 1985).

2.3.4. Hiyerarşik Ön Dağılım

Ön dağılım parametre değerlerinin belirsizliğiyle başa çıkmada bir yol, bu parametrelere ilave bir ön dağılım atamaktır. Bu yeni ön dağılımın parametreleri hiperparametre olarak bilinmektedir. Bu nedenle hiyerarşik ön dağılım, bazı kaynaklarda hiper ön dağılım olarak da adlandırılmaktadır. Hiyerarşik ön dağılımı belirlemede kullanılacak aşama sayısıyla ilgili bir sınırlama yoktur. Ancak uygulamada iki aşamadan fazlası nadiren kullanışlıdır.

Hiyerarşik ön dağılımla oluşturulmuş modelleme, MCMC simülasyon tekniğinin geliştirilmesiyle etkin olarak kullanılmaya başlanmıştır.

3. REGRESYONUN BAYESYEN ANALİZİ

3.1. Stokastik Kısıtlı Mixed Regresyon Modeli

Mixed regresyon modeli Theil ve Goldberger tarafından $R\beta = r$ lineer kısıtları kullanılarak 1961 yılında geliştirilmiştir. Bu lineer kısıtlar

$$r = R\beta + v \tag{3.1}$$

şeklinde ifade edilebilir. Burada r, $m \times 1$ tipinde bir vektördür ve $m \leq p$ olarak verilmiştir. Kısıtları belirleyen R matrisi $m \times p$ tipindedir ve v; $m \times 1$ tipinde 0 ortalamalı, $\sigma_v^2\Omega$ varyans kovaryans matrisli normal dağılıma sahip hatalar vektörüdür. v ve u vektörlerinin birbirlerinden bağımsız olduğu varsayılmıştır. $u \sim N(0, \sigma^2 I_T)$ iken $y = X\beta + u$ modeli (3.1) eşitliği ile birleştirildiğinde

$$\begin{bmatrix} y \\ r \end{bmatrix} = \begin{bmatrix} X \\ R \end{bmatrix} \beta + \begin{bmatrix} u \\ v \end{bmatrix} \tag{3.2}$$

$$y_M = X_M\beta + u_M \tag{3.3}$$

modelleri elde edilir. (3.2) modeli, (3.1) modeli ile ifade edilen ön bilgi ve örneklem bilgisini ($y = X\beta + u$) birleştirmektedir. (3.3) modelinde $u_M \sim N(0, \Omega_M)$ 'dır. Ayrıca

$$\Omega_M = \begin{bmatrix} \sigma^2 I_T & 0 \\ 0 & \sigma_v^2\Omega \end{bmatrix} \text{ ve } \Omega_M^{-1} = \begin{bmatrix} \frac{1}{\sigma^2} I_T & 0 \\ 0 & \frac{1}{\sigma_v^2}\Omega^{-1} \end{bmatrix} \tag{3.4}$$

matrisleri $P'\Omega_M P = I$ ve $\Omega_M^{-1} = PP'$ olacak şekildeki pozitif tanımlı matrislerdir. Burada P, $(m+T) \times (m+T)$ tipinde dönüşüm matrisidir. Yani, $P'u_M \sim N(0, I)$ şeklinde ifade edilebilir.

$$P'y_M = P'X_M\beta + P'u_M \tag{3.5}$$

dönüştürülmüş modeline EKK yöntemi uygulanabilir.

$$(P'y_M - P'X_M\beta)'(P'y_M - P'X_M\beta)$$

$$= (y_M - X_M\beta)'\Omega_M^{-1}(y_M - X_M\beta)$$

$$= \frac{(y-X\beta)'(y-X\beta)}{\sigma^2} + \frac{1}{\sigma_v^2}(r - R\beta)'\Omega^{-1}(r - R\beta)$$

ifadesi β 'ya göre minimum yapılırsa EKK tahmin edici

$$b_M = [(P'X_M)'(P'X_M)]^{-1}(P'X_M)'P'y_M \tag{3.6}$$

$$= (X_M'\Omega_M^{-1}X_M)^{-1}X_M'\Omega_M^{-1}y_M$$

olarak elde edilir. (3.6) 'daki Aitken genelleştirilmiş EKK tahmin edici, bilinen Ω_M matrisi ile minimum varyanslı yansız tahmin edici olma özelliğini sağlar (BLUE). b_M, mixed tahmin edici olarak adlandırılır. $\lambda = \frac{\sigma^2}{\sigma_v^2}$ varyansların oranı olmak üzere

$$X_M'\Omega_M^{-1}X_M = [X' \quad R']\begin{bmatrix} \frac{1}{\sigma^2}I_T & 0 \\ 0 & \frac{1}{\sigma_v^2}\Omega^{-1} \end{bmatrix}\begin{bmatrix} X \\ R \end{bmatrix}$$

$$= \frac{1}{\sigma^2}X'X + \frac{1}{\sigma_v^2}R'\Omega^{-1}R$$

olduğundan mixed tahmin edici b_M 'nin alternatif bir formu

$$b_M = \left(\frac{1}{\sigma^2}X'X + \frac{1}{\sigma_v^2}R'\Omega^{-1}R\right)^{-1}\left(\frac{X'y}{\sigma^2} + \frac{R'\Omega^{-1}r}{\sigma_v^2}\right) \tag{3.7}$$

$$= (X'X + \lambda R'\Omega^{-1}R)^{-1}(X'y + \lambda R'\Omega^{-1}r)$$

$$= (X'X + \lambda R'\Omega^{-1}R)^{-1}\left(X'Xb + \lambda R'\Omega^{-1}Rb_p\right)$$

şeklindedir. Burada $b = (X'X)^{-1}X'y$, $y = X\beta + u$ modelindeki β 'nın EKK tahmin edicisi ve $b_p = (R'\Omega^{-1}R)^{-1}R'\Omega^{-1}r$ ise $r = R\beta + v$ kısıtındaki β 'nın önbilgi tahmin edicisidir. (3.7) 'deki yapı b_p önbilgi tahmin edicisi ile b örneklem tahmin edicisinin ağırlıklandırılmış matris kombinasyonu şeklinde mixed tahmin ediciyi ifade etmektedir. Yani,

$$W_1 = (X'X + \lambda R'\Omega^{-1}R)^{-1}X'X$$

ve

$$W_2 = \lambda(X'X + R'\Omega^{-1}R)^{-1}R'\Omega^{-1}R$$

olmak üzere

$$b_M = W_1 b + W_2 b_p \tag{3.8}$$

formundadır.

(3.7) 'deki üçüncü eşitlik kısmı

$$b_M = (X'X + \lambda R'\Omega^{-1}R)^{-1}\left[X'X\left(b - b_p\right) + (\lambda R'\Omega^{-1}R + X'X)b_p\right] \tag{3.9}$$

$$= b_p + (X'X + \lambda R'\Omega^{-1}R)^{-1}X'X\left(b - b_p\right)$$

şeklinde düzenlenerek b_M 'nin alternatif bir formu elde edililebilir.

(3.7) ya da (3.9) 'dan görüldüğü gibi $\sigma_v^2 \to \infty$ veya $\lambda \to \infty$ ise yani bilgi vermeyen ön bilgi kullanılırsa $b_M = b$ olur. Ayrıca $\lambda \to \infty$ ise yani $\sigma_v^2 \to 0$ ise $b_M = b_p$ olur.

Özel bir durum olarak, değerleri bilinen bir β_0 vektörü ile $r = R\beta + v$ kısıtı $r = R\beta_0$ formunda yazılabilir. Bu durumda,

$$b_p = (R'\Omega^{-1}R)^{-1}R'^{\Omega^{-1}}R\beta_0 = \beta_0$$

olur ve (3.9) eşitliğinden,

$$b_M = \beta_0 + (X'X + \lambda R'\Omega^{-1}R)^{-1}X'X(b - \beta_0) \tag{3.10}$$

elde edilir. Örneğin, $r = 0$ ve $R = I$ iken β_0 (3.10) 'da yerine konulduğunda,

$$b_M = \left(X'X + \lambda I_p\right)^{-1}X'y \tag{3.11}$$

olur. Genellikle tek sabitli ve p açıklayıcı değişkenli bir modelde σ^2 'nin yansız tahmini için $S^2 = (y - X\beta)^{-1}(y - X\beta)(T - p)^{-1}$ kullanılır. $\sigma^2\Omega$ 'nın bilindiği varsayımıyla,

$$\hat{\beta}_M = \left[\frac{1}{S^2}X'X + \frac{1}{\sigma_v^2}R'\Omega^{-1}R\right]^{-1}\left[\frac{1}{S^2}X'y + \frac{1}{\sigma_v^2}R'\Omega^{-1}R\right] \tag{3.12}$$

b_M 'nin tahmini elde edilir. Ancak bu tahmin edici yansız olmasına rağmen varyans kovaryans matrisi b_M 'ninkinden farklıdır. Çünkü (3.12) 'de S^2 stokastiktir.

3.2. Eşlenik Ön Dağılım Varlığında Bayesyen Analiz

$u \sim N(0, \sigma^2 I)$ olmak üzere y 'nin $X\beta$ ortalama vektörlü, $\sigma^2 I$ varyans kovaryans matrisli çok değişkenli normal dağılıma sahip olduğunu düşünelim. Yani $y \sim N(X\beta, \sigma^2 I)$ olsun. y 'nin olasılık fonksiyonu (verinin olasılığı),

$$f(y/\beta, \sigma, X) = \frac{1}{(2\pi)^{\frac{T}{2}}\sigma^T}\exp\left[-\frac{1}{2\sigma^2}(y - X\beta)\,(y - X\beta)\right]$$

şeklinde yazılabilir. Bu da y ve X üzerine verilmiş bir veri kümesi için β, σ ya bağlı benzerlik fonksiyonu olarak ifade edilebilir. Yani,

$$L(\beta, \sigma) = L(\beta, \sigma/y) = f(y/\beta, \sigma, X)$$

olur. Benzerlik fonksiyonundaki karesel formu,

$$(y - X\beta)\,(y - X\beta) = (y - Xb)\,(y - Xb) + (\beta - b)\,X\,X(\beta - b)$$

$$= SSE(b) + (\beta - b)\,X\,X(\beta - b)$$

şeklinde yazabiliriz. Böylece benzerlik fonksiyonu,

$$L(\beta, \sigma) = \frac{1}{(2\pi)^{p/2}\sigma^T} \exp\left[-\frac{1}{2\sigma^2}\{SSE(b) + (\beta - b)\,X\,X(\beta - b)\}\right]$$

$$\propto \sigma^{-T} \exp\left[-\frac{1}{2\sigma^2}\{SSE(b) + (\beta - b)\,X\,X(\beta - b)\}\right] \qquad (3.13)$$

biçiminde yeniden yazılabilir. Burada $\propto$ oransallığı göstermektedir. Sabitler atılarak ifade basitleştirilmiştir. İki tür durumla karşılaşılabilir:

-σ biliniyorsa:

Bu durumda (3.13) eşitliği

$$L(\beta) \propto \sigma^{-T} exp\left[-\frac{1}{2\sigma^2}\{(\beta - b)\,X\,X(\beta - b)\}\right]$$

olarak yazılabilir. Çünkü $SSE/2\sigma^2$ ifadesi β yı içermemektedir. Benzerlik fonksiyonunun bu yapısı nedeniyle β nın ön dağılımı için normal dağılım

kullanılabileceği önerilmiştir. β nın ön dağılımı için, β_0 önsel ortalama vektörü ve $\sigma^2\Omega$ önsel varyans kovaryans matrisli çok değişkenli normal dağılım,

$$p(\beta) = \frac{1}{(2\pi)^{p/2}|\sigma^2\ |^{1/2}} \exp\left[-\frac{1}{2\sigma^2}(\beta-\beta_0)\,\Omega^{-1}(\beta-\beta_0)\right]$$

$$\propto \exp\left[-\frac{1}{2\sigma^2}(\beta-\beta_0)\,\Omega^{-1}(\beta-\beta_0)\right] \qquad (3.14)$$

şeklindedir. Daha sonra Bayes teoremi kullanılarak sonsal dağılım, ön dağılım ile benzerlik fonksiyonunun çarpımı şeklinde yazılır ve oransallık ifadesiyle yeniden düzenlenir. Yani,

$$p(\beta/y) \propto p(\beta)L(\beta)$$

$$\propto exp\left[-\frac{1}{2\sigma^2}\{(\beta-b)'X'X(\beta-b)+(\beta-\beta_0)'\Omega^{-1}(\beta-\beta_0)\}\right] \qquad (3.15)$$

olarak ifade edilir.

$$(\beta-b)'X'X(\beta-b)+(\beta-\beta_0)'\Omega^{-1}(\beta-\beta_0)$$

$$=\beta'(X'X+\Omega^{-1})\beta-2\beta'(X'Xb+\Omega^{-1}\beta_0)+c$$

$$=(\beta-b^*)'\Omega^{*-1}(\beta-b^*)+c^* \qquad (3.16)$$

olarak yazılabilir. Burada $c = b'X'Xb+\beta_0'\Omega^{-1}\beta_0$ şeklindedir ve β 'yı içermez. Ayrıca,

$$b^* = (X'X+\Omega^{-1})^{-1}(X'Xb+\Omega^{-1}\beta_0)$$

$$\Omega^* = (X'X+\Omega^{-1})^{-1}$$

$$c^* = c - b^{*\prime}\Omega^{*-1}b^* \qquad (3.17)$$

şeklindedir ve bu ifadeler de β'yı içermemektedir. (3.16) kullanılarak sonsal yoğunluk,

$$p(\beta/y) \propto exp\left[-\frac{1}{2\sigma^2}(\beta - b^*)'\Omega^{*-1}(\beta - b^*)\right] \quad (3.18)$$

şeklinde yazılabilir. Burada $(2\pi)^{-T/2}|\sigma^2\Omega^*|^{-1/2}$ normallik sabitidir. Sonsal yoğunluk da b^* ortalama vektörlü, Ω^* varyans kovaryans matrisli çok değişkenli normal dağılımdır ve bu genellikle $\beta \sim N(b^*, \sigma^2\Omega^*)$ şeklinde ifade edilir.

Önsel ve sonsal dağılımların normal dağılım olması nedeniyle, yukarıda verilen β'nın ön dağılımı eşlenik ön dağılım olarak adlandırılır. Genel olarak,

$$b_1^* = \left(\frac{X'X}{\sigma^2} + \frac{\Omega^{-1}}{\sigma_\beta^2}\right)^{-1}\left(\frac{X'X}{\sigma^2} + \frac{\Omega^{-1}\beta_0}{\sigma_\beta^2}\right)$$

$$\Omega_1^* = \left(\frac{X'X}{\sigma^2} + \frac{\Omega^{-1}}{\sigma_\beta^2}\right)^{-1} \quad (3.19)$$

olmak üzere β'nın ön dağılımı $N(\beta_0, \sigma_\beta^2\Omega)$ ise β'nın son dağılımı $N(b_1^*, \sigma^2\Omega_1^*)$ olur. Ayrıca $\sigma_\beta^2 = \sigma^2$ için, $b_1^* = b^*$ ve $\Omega_1^* = \Omega^*$ olur.

-σ bilinmiyorsa:

Bu durumda (3.13) 'deki benzerlik fonksiyonuna yakın bir ön dağılım söz konusudur. a>0, l>0 olmak üzere eşlenik ön dağılım,

$$p(\beta, \sigma) = p(\beta/\sigma)p(\sigma)$$

$$\propto \sigma^{-(l+p)}|\Omega|^{-1/2}exp\left[-\frac{1}{2\sigma^2}[(l-1)a^2 + (\beta - \beta_0)'\Omega^{-1}(\beta - \beta_0)]\right] \quad (3.20)$$

şeklinde ortak bir dağılımdır. Verilen bir σ ile β ′nın çok değişkenli koşullu normal dağılımı,

$$p(\beta/\sigma) \propto |\Omega|^{-1/2}\sigma^{-p}exp\left[-\frac{1}{2\sigma^2}[(\beta-\beta_0)'\Omega^{-1}(\beta-\beta_0)]\right]$$

şeklindedir. l>0 ve a>0 olmak üzere,

$$p(\sigma) \propto \sigma^{-1}exp\left[-\frac{(l-1)a^2}{2\sigma^2}\right]$$

ters gama dağılımıdır. (3.20) 'deki ön dağılım Normal-Gama formunda bir ön dağılımdır. $p(\beta,\sigma)$ ′nın σ>0 için integrali alındığında β 'nın marjinal dağılımı çok değişkenli student t dağılımı olacaktır. (3.13) 'deki benzerlik fonksiyonunu (3.20) 'deki ön dağılım ile çarparak ve (3.16) 'yı kullanarak β ve σ ′nın son dağılımı,

$$p(\beta,\sigma/y) \propto \sigma^{-n}exp\left[-\frac{1}{2\sigma^2}\{c_0+(\beta-b^*)'\Omega^{*-1}(\beta-b^*)\}\right] \qquad (3.21)$$

elde edilir. Burada $n=(T+l+p)$ ve $c_0=SSE(b)+(l-1)a^2+c^*$ 'dır. (3.21) nın σ ′ya göre integrali alındığında β ′nın marjinal son dağılımı,

$$p(\beta/\sigma) \propto [c_0+(\beta-b^*)'\Omega^{*-1}(\beta-b^*)]^{-(n-1)/2} \qquad (3.22)$$

olur. Bu da b^* ortalama vektörlü, $\frac{c_0}{n-p-1}\Omega^*$ varyans kovaryans matrisli çok değişkenli t dağılımının yoğunluk fonksiyonudur. Yani, β 'nın son yoğunluk fonksiyonu~çok değişkenli $t\left(b^*,\frac{c_0}{n-p-1}\Omega^*\right)$ olur. Önsel ve sonsal dağılımların çok değişkenli t dağılımı olması nedeniyle, yukarıda verilen β ′nın ön dağılımı eşlenik ön dağılım olarak adlandırılır.

3.3. Belirsiz Ön Dağılım Varlığında Bayesyen Analiz

β ve σ parametreleri hakkında bir şey bilmediğimiz zaman β ve σ 'nın ortak ön dağılımı,

$$p(\beta,\sigma) = p(\beta/\sigma)p(\sigma)$$

$$\propto \frac{1}{\sigma} \tag{3.23}$$

şeklindedir. Burada $p(\beta/\sigma)$ bir sabittir ve $p(\sigma)$, $1/\sigma$ ile orantılıdır. Bu da, (3.20) deki β ve σ'nın ön dağılımlarının varyanslarının $\Omega^{-1} \to 0$ ve $a \to 0$ olmasıyla elde edilebilir. (3.23) 'deki ön dağılım "belirsiz" ön dağılım ya da "uygunsuz" ön dağılım olarak bilinir. Uygunsuz denmesinin sebebi şudur: Olasılık yoğunluk fonksiyonunun integrali alındığında sonsuz çıkar ve bu da olasılıklar toplamının 1'e eşit olma aksiyomunu bozar. Belirsiz ön dağılım kullanılarak son dağılım,

$$p(\beta,\sigma/y) \propto \sigma^{-(T+1)} exp\left[-\frac{1}{2\sigma^2}\{SSE(b) + (\beta-b)'X'X(\beta-b)\}\right] \tag{3.24}$$

şeklinde elde edilir. Böylece verilen bir σ ile β 'nın son dağılımı$\sim N(b,\sigma^2\, X'X)$ olur. Ancak, σ ′ya göre integral alındığında β 'nın son dağılımı,

$$p(\beta/y) \propto [\{SSE(b) + (\beta-b)'X'X(\beta-b)\}]^{-T/2} \tag{3.25}$$

olur. Bu da b ortalama vektörlü, $\frac{SSE(b)}{T-p}(X'X)^{-1}$ varyans kovaryans matrisli çok değişkenli t dağılımıdır. σ hakkında yorum yapabilmek için ortak son dağılımın β 'ya göre integrali alınmalıdır.

3.4. Bayesyen Regresyon Tahmin Edici

β ve σ üzerine eşlenik ve belirsiz ön dağılımlar kullanıldığında β'nın marjinal son dağılımı çok değişkenli t dağılımı olarak elde edilmişti. Böylece β parametresi, çok değişkenli t dağılımının olasılık yoğunluk fonksiyonunun bilinen özellikleri kullanılarak analiz edilebilir. Bu simetrik son yoğunluk fonksiyonunun ortalaması moduna eşittir ve parametrenin Bayes tahmin edicisini sağlamaktadır. Bayes tahmin edici karesel kayıp fonksiyonuna göre beklenen kaybın minimum olduğu bir tahmin edicidir (MELO). (3.25) 'den β ve σ hakkında belirsiz ön bilgi olduğu varsayılırsa β'nın Bayes tahmin edicisi EKK ya da MLE tahmin edici ile aynıdır, yani $b = (X'X)^{-1}X'y$ olur. Aslında bu beklenen bir durumdur çünkü belirsiz ön bilgi kullanıldığında son yoğunluk fonksiyonunda olabilirlik fonksiyonu, ön dağılım fonksiyonuna üstün gelir.

β ve σ üzerine eşlenik ön dağılım kullanılması durumunda ise β'nın Bayes tahmin edicisi (3.16) 'dan,

$$b^* = (X'X + \Omega^{-1})^{-1}(X'Xb + \Omega^{-1}\beta_0) \tag{3.26}$$

$$= (X'X + \Omega^{-1})^{-1} X'Xb + (X'X + \Omega^{-1})^{-1}\Omega^{-1}\beta_0 \tag{3.27}$$

şeklindedir. b^* tahmin edicisi, veriye dayanan ML tahmin edicisi b ve önsel ortalama β_0 'ın ağırlıklandırılmış matris kombinasyonudur. (3.27) 'daki ağırlık matrislerinin toplamı birim matrisi verir. Ayrıca, b ve β_0 'ın ağırlık matrisleri sırasıyla b ve β 'nın varyans kovaryans matrislerinin normalleştirilmiş tersleridir. $V(b) = \sigma^2(X'X)^{-1}$ ve $V(\beta) = \sigma^2$ olduğundan $V(b)^{-1} + V(\beta)^{-1} = \frac{1}{\sigma^2}(X'X + \Omega^{-1})$ olur. Bu nedenle b için normalleştirilmiş ağırlık onun kesinlik matrisi olan $[V(b)^{-1} + V(\beta)^{-1}]^{-1}[V(b)]^{-1} = (X'X + \Omega^{-1})^{-1} X'X$ ile verilir. Benzer biçimde β_0 için ağırlık da onun kesinlik matrisidir. β'nın son dağılımının kesinlik matrisi de $V(b)^{-1} + V(\beta)^{-1} = (\sigma^2\Omega^*)^{-1}$ şeklindedir. Ayrıca (3.26) eşitliğindeki ikinci paranteze X'Xβ_0 ekleyip çıkartılarak b^* Bayesyen tahmin edici,

$$b^* = (X'X + \Omega^{-1})^{-1}[X'X(b - \beta_0) + (X'X + \Omega^{-1})\beta_0] \quad (3.28)$$

$$= \beta_0 + (X'X + \Omega^{-1})^{-1}X'X(b - \beta_0) \quad (3.29)$$

şeklinde yazılabilir. Ya da $X'X = X'X + \Omega^{-1} - \Omega^{-1}$ eşitliği kullanılarak,

$$b^* = \beta_0 + (I - (X'X + \Omega^{-1})^{-1}\Omega^{-1})(b - \beta_0) \quad (3.30)$$

elde edilebilir. (3.26) ve (3.29) 'daki Bayesyen tahmin edici b^* 'ın her iki formu da $R = I$,$\lambda = 1$ ve $r = \beta_0$ için (3.7) ve (3.9) eşitliklerindeki mixed tahmin edici (b_M) formlarına karşılık gelmektedir. Bunun anlamı ise $p(\beta/\sigma)$ önselinin β_0 ortalamalı, $\sigma^2\Omega$ varyans kovaryans matrisli çok değişkenli normal dağılıma sahip olduğudur. Böylece $E_v = 0$ ve $E_{vv'} = \sigma^2\Omega$ olmak üzere $\beta = \beta_0 + v$ olarak yazılabilir. Bu da $r = \beta_0$ ve $R = I$ olmak üzere $r = R\beta + v$ stokastik kısıtları altında $y = X\beta + u$ modelinden bahsettiğimizi gösterir. Aslında β_0 üzerine koşul konularak,

$$\frac{(y-X\beta)'(y-X\beta)}{\sigma^2} + \frac{1}{\sigma^2}(\beta - \beta_0)'\Omega^{-1}(\beta - \beta_0)$$

ifadesinin β 'ya göre minimum yapılmasıyla mixed tahmin edici olarak adlandırılan GLS tahmin edici elde edilir. Burada mixed tahmin edici b^* Bayesyen tahmin ediciye eşittir. Bu minimizasyon (3.5) 'daki minimizasyona karşılık gelmektedir.

Bayesyen regresyon sonuçları ile mixed regresyon sonuçları aynı olsa bile yorumları farklıdır. Bayesyen regresyon modelinde β bir rasgele değişken iken mixed regresyon modelinde değildir.

3.5. Bayesyen Tahmin Edicilerin Özellikleri

Normalde tam bir Bayesyen analiz için Bayesyen tahmin edicilerin klasik özelliklerini bilmeye gerek yoktur. Bununla birlikte, birçok tahmin edici Bayesyen

tahmin edici ile kıyaslandığından bayesyen tahmin edicinin özelliklerini incelemek faydalı olacaktır. Burada b ve b^* tahmincilerinin özellikleri incelenecektir.

$T \to \infty$ iken, $(X'X + \Omega^{-1})^{-1}X'X \equiv (\frac{X'X}{T} + \frac{\Omega^{-1}}{T})^{-1}\frac{X'X}{T}$ ifadesinin birim matrise, $(X'X + \Omega^{-1})^{-1} \equiv (\frac{X'X}{T} + \frac{\Omega^{-1}}{T})^{-1}\frac{1}{T}$ ifadesinin ise sıfıra yakınsadığı kolaylıkla gösterilebilir. Çünkü $T \to \infty$ iken, $\frac{X'X}{T}$ matrisi pozitif tanımlı bir matristir ve $\frac{\Omega^{-1}}{T}$ matrisi 0 matrisidir. Böylece, büyük örneklemler için b^*'ın b ye eşit olduğu (3.26.) eşitliğinden görülebilir. Bu nedenle b^* tutarlı, asimptotik olarak yansız ve asimptotik olarak etkindir.

Kullanılan ön bilgi doğru ise b^* MSE kriteri bakımından yeterli bir tahmin edicidir. Ayrıca eşlenik ön dağılım varlığında b^*, karesel kayıp fonksiyonuna göre beklenen kaybın minimum olduğu bir tahmin edicidir (MELO). Öte yandan, uygunsuz ön dağılım kullanıldığında b^* MELO tahmin edici olmasına rağmen üç ya da daha fazla açıklayıcı değişken varlığında MSE kriteri bakımından yeterli değildir. b yansız, dahası bir BLUE tahmin edicidir. Öte yandan b^* lineer homojen olmayan bir tahmin edicidir ve yanlıdır. (3.26) eşitliğinin her iki tarafının beklenen değeri alınırsa yanlılık,

$$E(b^*) - \beta = -(X'X + \Omega^{-1})^{-1}\Omega^{-1}(\beta - \beta_0)$$

$$= -\Omega^*\Omega^{-1}\beta^* \quad (14)$$

olarak elde edilir. Burada $\Omega^* = (X'X + \Omega^{-1})^{-1}$ ve $\beta^* = (\beta - \beta_0)$ şeklindedir.

Şimdi b ve b^*'ın varyans kovaryans matrisleri ve MMSE karşılaştırmaları incelensin.

$$b^* - E(b^*) = -\beta^* + (I - \Omega^*\Omega^{-1})(b - \beta_0) + \Omega^*\Omega^{-1}\beta^* \quad (3.31)$$

$$= (I - \Omega^*\Omega^{-1})(b - \beta) \quad (3.32)$$

ve

$$(b^* - Eb^*)(b^* - Eb^*)' = (I - \Omega^*\Omega^{-1})(b-\beta)(b-\beta)'(I - \Omega^{-1}\Omega^*) \qquad (3.33.)$$

şeklindedir. (3.33) eşitliğinin iki tarafının da beklenen değeri alınırsa,

$$V(b^*) = (I - \Omega^*\Omega^{-1})V(b)\ (I - \Omega^{-1}\Omega^{*\prime})$$

$$= \sigma^2(I - \Omega^*\Omega^{-1})(X'X)^{-1}(I - \Omega^{-1}\Omega^{*\prime})$$

$$= \sigma^2\Omega^*X'X\Omega^* \qquad (3.34)$$

elde edilir. Ayrıca

$$\Omega^{*-1} = (X'X + \Omega^{-1})$$

olduğundan

$$\Omega^{-1} = (\Omega^{*-1} - X'X) \text{ ve } \Omega^*\Omega^{-1} = (I - \Omega^*X'X)$$

elde edilir.

$$\Delta = V(b) - V(b^*) = \sigma^2[(X'X)^{-1} - \Omega^*X'X\Omega^*]$$

olsun. $b^{*\prime}$ın b 'den daha küçük varyansa sahip olması için Δ pozitif tanımlı bir matris olmalıdır. Bu ise, $V(\eta'b) \geq V(\eta'b^*)$ ya da her $\eta \neq 0$ için $\eta'\,\Delta\eta \geq 0$ ifadesinin gerçeklenmesiyle sağlanır. Bu nedenle,

$$\eta'\,\Delta\eta = \sigma^2[\eta'(X'X)^{-1}\eta - \eta'\Omega^*X'X\Omega^*\eta] \qquad (3.35)$$

yazılır. Eşdeğer olarak,

$$\xi' \Omega^{*-1} \Delta \Omega^{*-1} \xi = \sigma^2 \xi' [\Omega^{*-1} (X'X)^{-1} \Omega^{*-1} - X'X] \xi$$

yazılabilir. Burada $\xi = \Omega^* \eta'$dır.

$$\Omega^{*-1} (X'X)^{-1} \Omega^{*-1} = \Omega^{-1} (X'X)^{-1} \Omega^{-1} + X'X + 2\Omega^{-1}$$

olduğundan (3.35) 'in sağ tarafı, $\sigma^2 \xi' (\Omega^{-1} (X'X)^{-1} \Omega^{-1} + 2\Omega^{-1} \xi)$ şeklinde yazılabilir. Ω^{-1} ve $\Omega^{-1} (X'X)^{-1} \Omega^{-1}$ matrisleri pozitif tanımlı matrislerdir ve bu nedenle

$$\sigma^2 \xi' (X'X)^{-1} \Omega^{-1} + X'X + 2\Omega^{-1} > 0$$

olur. Buradan her $\eta \neq 0$ için $\eta' \Delta \eta \geq 0$ elde edilir. Böylece $V(\eta' b) > V(\eta' b^*)$ olur. Özel olarak $V(b_i) > V(b_i^*),\ i = 1,2, \dots, p$ şeklindedir. O halde b^* yanlı bir tahmin edici olmasına rağmen b'den daha küçük varyansa sahiptir.

Şimdi b^* ve b'nin MMSE kriterine göre karşılaştırılmasını inceleyelim. $\Delta^* = MtxMSE(b) - MtxMSE(b^*)$ ifadesi pozitif tanımlı bir matris olduğunda ya da her $\eta \neq 0$ için $MSE(\eta' b) \geq MSE(\eta' b^*)$ eşitsizliği sağlandığında b^*, b'ye tercih edilir.

$$MtxMSE(b) = V(b) + [Bias(b)][Bias(b)]' = \sigma^2 (X'X)^{-1}$$

ve

$$MtxMSE(b^*) = V(b^*) + [Bias(b^*)][Bias(b^*)]'$$

$$= \sigma^2 \Omega^* X'X \Omega^* + \Omega^* \Omega^{-1} \beta^* \beta^{*\prime} \Omega^{-1} \Omega^{*\prime} \qquad (3.36)$$

şeklindedir. Her $\eta \neq 0$ için $\eta' \Delta^* \eta \geq 0$ ise Δ^* pozitif tanımlı bir matristir. Alternatif olarak $b^{*\prime}$ın EKK tahmin edicisinden üstün olduğunu gösterebilmek için,

$$\xi' \left[\Omega^{*-1}(X'X)^{-1}\Omega^{*-1} - X'X - \frac{\Omega^{-1}\beta^*\beta^{*\prime}\Omega^{-1}}{\sigma^2} \right] \xi \geq 0 \qquad (3.37)$$

eşitsizliğinin sağlanması gerekir.

$$\theta_0 = \frac{(\xi\prime\Omega^{-1}\beta^*\beta^{*'}\Omega^{-1}\xi)}{2\sigma^2\xi\prime[\Omega^{-1}(X'X)^{-1}\Omega^{-1}+2\Omega^{-1}]\,\xi} \leq \frac{1}{2} \qquad (3.38)$$

olduğunda yukarıdaki ifade gerçeklenir. Ayrıca (3.37), sıfırdan farklı her ξ için

$$\theta = \mathrm{Sup}_{\xi}\,\theta_0$$

$$= \frac{\beta^{*\prime}\left[(X'X)^{-1}+2\Omega\right]^{-1}\beta^*}{2\sigma^2} \leq \frac{1}{2} \qquad (3.39)$$

eşitsizliği sağlandığında geçerlidir. Burada $\beta^* = \beta - \beta_0\,'$dır. Yani, $\theta \leq \frac{1}{2}$ olduğunda b^*, b 'ye tercih edilir.

3.6. Lindley 'in Bayesyen Regresyon Modeli

Ön bilgi seçiminden etkilenecek bir yapıya sahip olan β parametre vektörü olduğunda, $y = X\beta + u$ modelinin Bayesyen analizini göz önüne alalım. $\beta\,'$nın yapısı üzerine en basit haliyle,

$$E(\beta) = A\beta_0 \qquad (3.40)$$

formu örnek verilebilir. Burada A, $p \times p_0$ tipinde bilinen bir matris ve β_0, $p_0 \times 1$ tipinde hiperparametredir. β, bilinen varyans kovaryans matrisi Ω_0 ile çok değişkenli normal dağılıma sahip olsun. Böyle ise yukarıdaki yapı $\beta\,'$nın ön dağılımının$\sim N(A\beta_0, \sigma^2\Omega_0)$ olduğunu vurgulamaktadır. Örneğin,bu tip bir yapı gecikmesi dağıtılmış polinom şeklindeki modellerde ve çoklu regresyon modellerinde görülmektedir.

Birçok uygulamadaki süreçte β_0 hiperparametresinin yapısının,

$$E(\beta_0) = A_1\mu \tag{3.41}$$

formunda olduğu varsayılmaktadır. A_1 ve μ sırasıyla $p_0 \times p_1$ ve $p_1 \times 1$ tipinde olup bunlar bilinen matris ve vektördür. Genellikle p_1, p 'den daha küçüktür ve birçok uygulamada da 1'dir. β_0 hiperparametresinin bilinen Ω_1 varyans kovaryans matrisli normal dağılıma sahip olduğu varsayılmıştır, yani $\beta_0 \sim N(A_1\mu, \sigma^2\Omega_1)$ 'dir. Problemimiz yeniden, $y = X\beta + u$ genel Bayesyen regresyon modeli için,

$$y \sim N(X\beta, \sigma^2 I)$$

$$\beta \sim N(A\beta_0, \sigma^2\Omega_0)$$

$$\beta_0 \sim N(A_1\mu, \sigma^2\Omega_1) \tag{3.42}$$

varsayımları ile analiz yapmaktır. (3.42) 'deki genel Bayesyen modeli, Lindley (1971) ve Lindley ve Smith (1972) 'e göre ön dağılım yapıları hiyerarşik (üç aşamalı) lineer regresyon modelinden oluşmaktadır. İki aşamalı hiyerarşik ön dağılımların birinci aşaması lineer regresyondaki parametreler arasındaki ilişkiyi tanımlar, ikincisi ise bu ilişkinin yapısı hakkındaki bilgiyi tanımlar. $A = I$ alınarak sadece birinci aşama ön dağılım kullanılıyorsa, önceki bölümlerde tanımlanan Bayesyen regresyon modeli elde edilir.

(3.40) ve (3.41) 'deki yapıların lineer ve dağılımlarının normal olması nedeniyle yukarıdaki model,

$$y \sim N(X\beta, \sigma^2 I)$$

$$\beta \sim N(A_2\mu, \sigma^2\Omega_2);\ A_2 = AA_1;\ \Omega_2 = A\Omega_1A' + \Omega_0$$

modeline indirgenir. Çünkü burada $v \sim N(0, \Omega_0)$, $w \sim N(0, \Omega_1)$ ve $E_{vw'} = 0$ olmak üzere, $\beta = A\beta_0 + v$ ve $\beta_0 = A_1\mu + w$ olarak ifade edilebilir. Böylece $\beta = AA_1 + v + A\,w$ olup $AA_1\mu$ ortalamalı, $\sigma^2(A\Omega_1 A' + \Omega_0)$ varyans kovaryans matrislidir. (3.18) 'den y verisi ve σ^2, Ω_0, Ω_1 koşulları ile β 'nın sonsal dağılımı,

$$p(\beta/y, \sigma^2, \Omega_0, \Omega_1) \propto exp\left[-\frac{1}{2\sigma^2}(\beta - b_L^*)'\Omega_L^{*-1}(\beta - b_L^*)\right] \tag{3.43}$$

şeklinde yazılabilir. Burada b_L^*, Lindley ve b EKK tahmin ediciyi göstermek üzere,

$$b_L^* = \left(X'X + {\Omega_2}^{-1}\right)^{-1}\left(X'Xb + {\Omega_2}^{-1}A_2\mu\right)$$

$$\Omega_L^* = \left(X'X + {\Omega_2}^{-1}\right)^{-1} \tag{3.44}$$

elde edilir. b_L^* 'ın farklı bir formu da,

$$b_L^* = A_2\mu + \left[I - \left({\Omega_2}^{-1} + X'X\right)^{-1}{\Omega_2}^{-1}\right](b - A_2\mu) \tag{3.45}$$

şeklindedir. Bu sonsal yoğunluk b_L^* ortalama vektörlü ve Ω_L^* varyans kovaryans matrisli çok değişkenli normal dağılımdır. Böylece (3.44) 'den β 'nın Lindley Bayesyen tahmin edicisi,

$$b_L^* = [X'X + (A\Omega_1 A' + \Omega_0)^{-1}]^{-1}[X'Xb + (A\Omega_1 A' + \Omega_0)^{-1}AA_1\mu] \tag{3.46}$$

olur. Burada (3.45) 'deki ${\Omega_2}^{-1}$ formu için,

$${\Omega_2}^{-1} = (A\Omega_1 A' + \Omega_0)^{-1}$$

$$= {\Omega_0}^{-1} - {\Omega_0}^{-1}A(A'{\Omega_0}^{-1}A + {\Omega_1}^{-1})^{-1}A'{\Omega_0}^{-1} \tag{3.47}$$

şeklinde matris tersinin eşitliği kullanılmıştır.

$\Omega_1 \to \infty$ olduğu durumda yani β_0 üzerinde belirsiz ön dağılım söz konusuysa (3.47) eşitliği,

$$\Omega_2{}^{-1} = \Omega_0{}^{-1} - \Omega_0{}^{-1}A(A'\Omega_0{}^{-1}A)^{-1}A'\Omega_0{}^{-1} \tag{3.48}$$

formundadır ve Ω_1 terimini içermez. Burada,

$$\Omega_2{}^{-1}A = \Omega_0{}^{-1}A - \Omega_0{}^{-1}A\left(A'^{\Omega_0{}^{-1}}A\right)^{-1}\left(A'\Omega_0{}^{-1}A\right) = 0$$

olması nedeniyle (3.44) 'deki b_L^*,

$$b_L^* = \left(X'X + \Omega_2{}^{-1}\right)^{-1}X'y$$

$$= [I + (\Omega_2 X'X)^{-1}]^{-1}b \tag{3.49}$$

şeklinde basitleşir. Buradaki $(\Omega_2 X'X)^{-1}$ terimi, EKK tahmin edicisi b için bir düzeltme terimi olarak düşünülebilir. Bu formun benzer bir şekli Stein tipi ve Ridge tahmin edicilerde de görülmektedir. Bu tahmin ediciler Bayesyen olmayan yöntemlerle James ve Stein (1961) ve Hoerl ve Kennard (1970) tarafından elde edilmiştir. Böylece Lindley 'in hiperparametre modeli Bayesyen olmayan bu tahmin edicilere Bayesyen bir bakış açısı sağlamaktadır.

Eğer (3.42) 'de β yerine β_0 ile ilgilenilirse ve $\Omega_1 \to \infty$ ise verilen σ^2,Ω_0, Ω_1 ve y için β_0 ′ın sonsal yoğunluğu elde edilebilir. Buradan β_0 ′ın sonsal dağılımı,

$$b_0^* = \{A'((X'X)^{-1} + \Omega_0)^{-1}\}^{-1}A'((X'X)^{-1} + \Omega_0)^{-1}b \tag{3.50}$$

ortalamalı çok değişkenli normal dağılım olacaktır. Ayrıca (3.45) 'den $y \sim N(XA\beta_0, \sigma^2 I + \sigma^2 X\Omega_0 X')$ olmasıyla da b_0^* elde edilebilir. Böylece, β_0 belirsiz ön dağılıma sahipse β_0 ′ın sonsal ortalaması,

$$b_0^* = [A'X'(I + X\Omega_0 X')^{-1}XA]^{-1}A'X'(I + X\Omega_0 X')^{-1}y$$

şeklinde genelleştirilmiş EKK tahmin edicisidir. Burada $(I + X\Omega_0 X')^{-1}$ ifadesinin (3.47) 'deki formu kullanılmış ve ifade (3.50) ile benzer hale gelmiştir. (3.48) 'deki matris tersi kullanılarak,

$$X'X + \Omega_2{}^{-1} = \left(X'X + \Omega_0{}^{-1}\right) - \left(\Omega_0{}^{-1}A\right)\left(A'\Omega_0{}^{-1}A\right)^{-1}(A'\Omega_0{}^{-1})$$

elde edilir ve böylece Lindley 'in Bayesyen tahmin edicisi $b_L^{*\prime}$ın farklı bir formu,

$$b_L^* = \left(X'X + \Omega_0{}^{-1}\right)^{-1}(X'Xb + \Omega_0{}^{-1}Ab_0^*)$$

$$= Ab_0^* + [I + (\Omega_0 X'X)^{-1}]^{-1}(b - Ab_0^*) \qquad (3.51)$$

olarak elde edilir.

$b_L^{*\prime}$ın bu formunun b ve Ab_0^* tahmin edicilerinin ağırlıklandırılmış matris kombinasyonu olduğu görülmektedir. Ayrıca (3.51) 'in ikinci satırı, $b_L^{*\prime}$ın EKK tahmin edicisi b 'yi $Ab_0^{*\prime}$a daralttığını gösteriyor. Bunun sonucu olarak, b_L^* ve ridge tahmin ediciler arasında bir benzerlik olduğu (3.49) 'dan görülse de $(\Omega_2 = 1)$, daralmanın doğrultusu bakımından farklı oldukları da (3.51) 'den görülmektedir. Ridge regresyonda büzülme 0 'a doğru olurken, b_L^* Ab_0^* yönünde büzülmektedir. $b_L^{*\prime}ın$ (3.51) 'deki formu, $y = X\beta + u$ lineer regresyon modelinde β'nın ön dağılımı$\sim N(Ab_0^*, \Omega_0)$ olarak alındığında elde edilen sonsal ortalamaya benzerdir. Bu nedenle (3.51) 'deki b_L^*, (3.26) 'daki $b^{*\prime}$ın formuna çok yaklaşmaktadır. Fakat bunlar farklı ön dağılım yapılarına sahip olduğundan yorumlamaları da farklı olacaktır.

4. LİNEER BAYES TAHMİN EDİCİ

Bu bölümde Gruber (2012) tarafından, en genel haliyle lineer Bayesyen tahmin edici ve alternatif formları incelenmiştir. Ayrıca Bayesyen Liu tipi tahmin ediciler sınıfı tanımlanmıştır.

4.1. Temel Bilgiler

X, $s \leq m$ ranklı $n \times m$ boyutlu ve $U{:}\, m \times s$, $V{:}m \times (m-s)$, $S'{:}n \times s$, $T'{:}n \times (n-s)$ matrisler olmak üzere X 'in tekil değer ayrışımı,

$$X = [S' \quad T'] \begin{bmatrix} \Lambda^{1/2} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} U' \\ V' \end{bmatrix} = S'\Lambda^{1/2}U' \tag{4.1}$$

şeklindedir. Burada $[S' \quad T']$ ve $[U \quad V]$ ortogonal matrislerdir ve $\Lambda{:}s \times s$ boyutlu, $X'X$ matrisinin özdeğerlerinden oluşan köşegen matristir. λ_i özdeğerleri,

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_s$$

olacak biçimde sıralanmıştır.

$$X'X = U\Lambda^{1/2}SS'\Lambda^{1/2}U' = U\Lambda U' \tag{4.2}$$

$$XUU' = S'\Lambda^{1/2}U'UU' = S'\Lambda^{1/2}U' = X \tag{4.3}$$

ve

$$XV = S'\Lambda^{1/2}U'V = 0 \tag{4.4}$$

olduğu açıktır.

x ve y m boyutlu vektörler, A ise $s \leq m$ ranklı $m \times n$ tipinde bir matris olsun. Regresyon problemlerinin çözümünde sıklıkla

$$Ax = y \tag{4.5}$$

denklem sisteminin çözümüne ihtiyaç duyulmaktadır. $n = m$ ve $s = m$ iken $x = A^{-1}y$ olduğunu biliyoruz. (4.5) denklem sisteminin bir çözümü olsun. G, $m \times n$ tipinde bir matris olsun. Bu denklem sisteminin çözümü,

$$x = Gy \tag{4.6}$$

biçiminde olacaktır. Buradaki G, genelleştirilmiş invers (ters) olarak adlandırılır. (4.6) çözümü tek çözüm olmak zorunda değildir.

Tanım 4.1. $AGA = A$ eşitliği sağlanıyorsa, G matrisi A matrisinin bir genelleştirilmiş inversidir.

Tanım 4.2. Tanım 4.1. 'e ek olarak aşağıdaki koşulları sağlayan bir genelleştirilmiş inverse, Moore-Penrose invers olarak adlandırılır.

1. $GAG = G$
2. AG matrisi simetriktir.
3. GA matrisi simetriktir.

Bir A matrisinin genelleştirilmiş inversi A^- ile Moore-Penrose inversi A^+ ile gösterilir. Bir A matrisinin birden fazla genelleştirilmiş inversi olabilir fakat bir tane Moore-Penrose inversi vardır.

Teorem 4.1: X, s ranklı, $n \times m$ tipinde bir matris ve X 'in tekil değer ayrışımı $X = S'\Lambda^{1/2}\mathrm{U}'$ olsun.

1. $X^+ = U\Lambda^{-1/2}\mathrm{S}$ 'dir.
2. X 'in her genelleştirilmiş inversi X^- için $U'X^-S' = \Lambda^{-1/2}$ 'dir.
3. $A_1 = U'X^-T'$, $A_2 = V'X^-S'$ ve $A_3 = V'X^-T'$ olmak üzere

$$X^- = [U \quad V]\begin{bmatrix}\Lambda^{-1/2} & A_1\\ A_2 & A_3\end{bmatrix}\begin{bmatrix}S\\ T\end{bmatrix}$$

$$= X^+ + VA_2S + UA_1T + VA_3T$$

şeklinde ifade edilebilir.

4. C_1, C_2 bilinen matrisler ve $C_3 = C_2 \Lambda C_1$ olsun. Tanım 4.2 'deki birinci koşulu sağlayan X matrisinin genelleştirilmiş inversine yansıyan genelleştirilmiş invers denir.

$$M = [U \quad V] \begin{bmatrix} \Lambda^{-1/2} & C_1 \\ C_2 & C_3 \end{bmatrix} \begin{bmatrix} S \\ T \end{bmatrix}$$

X 'in bir yansıyan genelleştirilmiş inversidir.

5. Bir J matrisinin Moore Penrose inversini J^+ ile gösterelim. N, $n \times n$ tipinde tekil olmayan bir matris ise

$$(UNU')^+ = UN^{-1}U' \tag{4.3}$$

olup, (4.1) ve (4.3) eşitliklerinden

$$(X'X)^+ = U\Lambda^{-1}U'$$

elde edilir.

Teorem 4.1 'den aşağıdaki sonuçlara ulaşılabilir:

Sonuç 4.1.

1.

$$(X'X)^+ = U\Lambda^{-1}U'$$

2. $X'X$ matrisinin her genelleştirilmiş inversi için,

$$U'(X'X)^- U = \Lambda^{-1}$$

eşitliği doğrudur.

3.

$$(X'X)^{-} = [U \quad V]\begin{bmatrix}\Lambda^{-1} & B_1\\ B_2 & B_3\end{bmatrix}\begin{bmatrix}U'\\ V'\end{bmatrix} \tag{4.4}$$

şeklinde yazılabilir.

4.

$$H = [U \quad V]\begin{bmatrix}\Lambda^{-1} & B_1\\ B_2 & B_3\end{bmatrix}\begin{bmatrix}S\\ T\end{bmatrix}$$

matrisi $X'X$ 'in bir genelleştirilmiş inversidir. H matrisi, $B_1 = 0$ alındığında Tanım 4.2 'nin ikinci ve üçüncü koşullarını sağlar. $B_2 = B_1\,\Lambda {B_1}'$ iken birinci koşul da sağlanır.

Tanım 4.1 ve Tanım 4.2 'deki koşullar kontrol edilerek aşağıdaki sonuçlara ulaşılabilir:

N tekil olmayan bir matris ve P ortogonal bir matris olsun.

1.

$$(UNU')^{+} = UN^{-1}U'$$

2.

$$(PAP')^{+} = PA^{+}P'$$

3.

$$\begin{bmatrix}A^{+} & 0\\ 0 & B^{+}\end{bmatrix} = \begin{bmatrix}A & 0\\ 0 & B\end{bmatrix}^{+}$$

4.2. Ekk Tahmin Edici Ve Özellikleri

X, $s \leq m$ ranklı bir matris olmak üzere

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

lineer modelini ele alalım. X tam ranklı ise EKK tahmin edici

$$b = (X'X)^{-1}X'y$$

şeklindedir. X tam ranklı değilse EKK tahmin edici

$$b = G\,X'y$$

şeklindedir. Buradaki G matrisi, $X'X$ matrisinin genelleştirilmiş inversidir. Bir matrisin birden fazla genelleştirilmiş inversi olabilir. Bu nedenle regresyon parametreleri birden fazla EKK tahmin ediciye sahip olabilir.

G_1 ve G_2 $X'X$ matrisinin genelleştirilmiş inversleri olsunlar. $p'\beta$ parametrik fonksiyonunun EKK tahmin edicileri

$$p'\hat{\beta}_1 = p'G_1X'y$$

$$p'\hat{\beta}_2 = p'G_2X'y$$

eşitlikleriyle ifade edilebilir.

$p'\beta$ parametrik fonksiyonunun bir tek tahmin ediciye sahip olduğunu düşünelim. O zaman, $p'\beta$ parametrik fonksiyonuna tahmin edilebilir parametrik fonksiyon denir.

$VX = 0$ ve $XUU' = X$ olduğu dikkate alınarak (4.4) 'den

$$b = (X'X)^{+}X'y + UU'GVV'X'y + VV'GUU'X' + VV'GVV'X'y$$

$$= (X'X)^{+}X'y + VV'GUU'X'y$$

yazılabilir.

$$p'\beta = p'(X'X)^{+}X'y$$

olması için gerek ve yeter koşul

$$p'V = 0 \qquad (4.5)$$

denkleminin sağlanmasıdır. Yani, tahmin edilebilir parametrik fonksiyonlar sınıfı (4.5) denklemini sağlamalıdır.

Farebrother (1976,1978) 'a göre; d, s boyutlu bir vektör olmak üzere $p'\beta$ 'nın tahmin edilebilir olması için gerek ve yeter koşul

$$p = Ud \tag{4.6}$$

denkleminin sağlanmasıdır.

Diğer tanımlara göre (Searle (1971)); tahmin edilebilir bir parametrik fonksiyon, gözlemlerin bir lineer fonksiyonunun beklenen değeridir. Yani,

$$p'\beta = E(c'y) = c'X\beta \tag{4.7}$$

ya da

$$p' = c'X \tag{4.8}$$

eşitlikleri sağlandığında $p'\beta$ tahmin edilebilir olacaktır. Şimdi, (4.5) 'i kullanarak (4.6) eşitliğini ispatlayalım. $[U \quad V]$ ortogonal bir matris olduğundan

$$p' = p'(UU' + VV') = p'UU' + p'VV' = p'UU'$$

elde edilir. $d = U'p$ alınırsa (4.6) gösterilmiş olur. (4.6) kullanılarak (4.8) eşitliği ispatlansın. $X = U\Lambda^{1/2}S$ şeklindedir. Buradan $U = X'S'\Lambda^{-1/2}$ elde edilir. $p = Ud$ ise, $p = X'S'\Lambda^{-1/2}d$ olur. $p' = c'X$ olması için $p = X'c$ sağlanmalıdır. $c = S'\Lambda^{-1/2}d$ alınırsa (4.8) gösterilmiş olur. Son olarak, (4.8) 'i kullanılarak (4.5) ispatlansın. $p' = c'X$ ve $XV = 0$ olduğundan $p'V = c'XV = 0$ elde edilir.

X matrisi tam ranklı olduğunda (4.7) denkleminin bir tek çözümü vardır. O halde; X 'in tam ranklı olması durumunda her parametrik fonksiyon tahmin edilebilir olacaktır.

Teorem 4.2:

$$E(\varepsilon) = 0 \; ve \; D(\varepsilon) = \sigma^2 I$$

olmak üzere

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

lineer regresyon modelini ele alalım. $L'y$ 'nin $p'\beta$ parametrik fonksiyonunun yansız tahmin edicisi olması için gerek ve yeter koşul $p'\beta$ 'nın tahmin edilebilir olmasıdır.

İspat:

$$E(L'y) = L'X\beta = p'\beta$$

eşitliğinin sağlanması için gerek ve yeter koşul $p' = L'X$ olmasıdır. Bu durumda $p'\beta$ parametrik fonksiyonu tahmin edilebilir olacaktır.

4.3. Bayes Tahmin Edici

y, gözlemlerden oluşan n boyutlu bir vektör; X, $s \leq m$ ranklı $n \times m$ tipinde bir matris; β, m boyutlu parametrelerin bir vektörü ve ε, n boyutlu hatalar vektörü olmak üzere,

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

modelini düşünelim. β, sırasıyla ortalaması ve varyans kovaryans matrisi

$$E(\beta) = \theta \; ve \; D(\beta) = H$$

olan bir rasgele değişken olsun. Burada H, negatif tanımlı olmayan bir matristir.

$$E(\varepsilon/\beta) = 0 \; ve \; D(\varepsilon/\beta) = \sigma^2 I$$

olduğu varsayılsın. Lineer Bayes tahmin ediciyi bulmak için

$$v = Var(p'\beta - a - L'y)$$

varyansını minimum yapan ve

$$E(p'\beta - a - L'y) = 0 \tag{4.9}$$

koşulunu sağlayan $a + L'y$ lineer fonksiyonu hesaplanmalıdır. (4.9) 'dan

$$a = (p' - L'X)\theta \tag{4.10}$$

bulunur.

$$v = Var(p'\beta - a - L'y)$$

$$= Var(p'\beta) + Var(L'y) - 2cov(p'\beta, L'y) \tag{4.11}$$

olup, (4.11) 'i minimum yapmak için $Var(L'y)$ ve $cov(p'\beta, L'y)$ değerleri hesaplanırsa

$$Var(L'y) = L'Var(y)L$$

$$= L'Var(X\beta + \varepsilon)L$$

$$= L'[Var(X\beta) + Var(\varepsilon)]L$$

$$= L'[XVar(\beta)X' + \sigma^2 I]L$$

$$= L'[XHX' + \sigma^2 I]L$$

ve

$$cov(p'\beta, L'y) = E\big[\big(p'\beta - E(p'\beta)\big)\big(L'y - E(L'y)\big)'\big]$$

$$= E\big[p'\big(\beta - E(\beta)\big)\big(y'L - E(y'L)\big)\big]$$

$$= p'E\big[\big(\beta - E(\beta)\big)\big(y' - E(y')\big)\big]L$$

$$= p'cov(\beta, y)L$$

elde edilir.

$$cov(\beta, y) = E\big[\big(\beta - E(\beta)\big)\big(y - E(y)\big)'\big]$$

$$= E[(\beta - \theta)(y' - \theta'X')]$$

$$= E[\beta(\beta'X' + \varepsilon')] - \theta\theta'X' - \theta\theta'X' + \theta\theta'X'$$

$$= E(\beta\beta'X') + E(\beta\varepsilon') - \theta\theta'X'$$

$$= [E(\beta\beta') - \theta\theta']X'$$

$$= E[(\beta - \theta)(\beta - \theta)']X'$$

$$= HX'$$

olup,

$$Var(p'\beta - a - L'y) = p'Hp + L'(XHX' + \sigma^2 I)L - 2p'HX'L \qquad (4.12)$$

elde edilir.

(4.12) 'nin L 'ye göre türevi alınıp bu türev sıfıra eşitlenirse, işlem sonucunda,

$$(XHX' + \sigma^2 I)L = XHP$$

ve

$$L' = p'HX'(XHX' + \sigma^2 I)^{-1} \tag{4.13}$$

bulunur. (4.10) ve (4.13) 'den lineer Bayes tahmin edici

$$p'\hat{\beta}_c = p'\theta + p'HX'\,(XHX' + \sigma^2 I)^{-1}(y - X\theta) \tag{4.14}$$

şeklindedir.

(4.12) 'deki optimizasyon problemini çözmenin alternatif bir yolu aşağıda verilmiştir. c,

$$E(p'\beta - a - c'b) = 0 \tag{4.15}$$

eşitliğini sağlayan bir vektör olsun. b, EKK tahmin edici ve a da bir sabit olmak üzere

$$v = Var(p'\beta - a - c'b)$$

varyansı minimum yapılmalıdır.

$$v = Var(p'\beta) + Var(c'b) - 2cov(p'\beta, c'b) \tag{4.16}$$

$$= p'Hp + c'[UU'HUU' + \sigma^2(X'X)^{+}]c - 2p'HUU'c$$

elde edilir. (4.15) 'den

$$a = p'\theta - cUU'\theta$$

olduğu görülür. (4.16) ifadesinin c 'ye göre türevi alınıp bu türev sıfıra eşitlenirse

$$[UU'HUU' + \sigma^2(X'X)^+]c = p'HUU'$$

elde edilir. Burada $UU'FUU' + \sigma^2(X'X)^+$ matrisi negatif tanımlı değildir. (4.15) ve (4.16) 'dan

$$p'\beta = p'\theta + p'HUU'[UU'HUU' + \sigma^2(X'X)^+]^+(b - \theta)$$

elde edilir.

Örnek:

$$y_{1j} = \mu + \varepsilon_{1j} \qquad 1 \le n \le n_1$$
$$y_{2j} = \mu + \varepsilon_{2j} \qquad 1 \le n \le n_2$$

lineer modeli ele alınsın. $i = 1{,}2$ için

$$E(\varepsilon_i) = 0 \; ve \; D(\varepsilon_i) = \sigma^2$$

olsun. $E(\mu_i) = 0 \; ve \; Var(\mu_i) = 1$ olduğu verilsin. Burada μ_1 ve μ_2 bğımsızdır. $(\mu_1 - \mu_2)$ 'nin $c_1\bar{Y}_{1.} + c_2\bar{Y}_{2.}$ formundaki Bayes tahmin edicisi aşağıdaki şekilde bulunabilir.

$$E\left(c_1\bar{Y}_{1.} + c_2\bar{Y}_{2.} - (\mu_1 - \mu_2)\right) = -a_0 = 0$$

olmak üzere

$$g(c_1, c_2) = Var\big((c_1\bar{Y}_{1.} + c_2\bar{Y}_{2.}) - (\mu_1 - \mu_2)\big)$$

$$= c_1^2 Var(\bar{Y}_{1.}) + c_2^2 Var(\bar{Y}_{2.}) - 2c_1 cov(\bar{Y}_{1.}, \mu_1) + 2c_2 cov(\bar{Y}_{2.}, \mu_2)$$

varyansı minimum yapılmalıdır. Bunun için,

$$\frac{\partial g}{\partial c_1} = 2c_1 Var(\bar{Y}_{1.}) - 2\, cov(\bar{Y}_{1.}, \mu_1) = 0$$

$$\frac{\partial g}{\partial c_2} = 2c_2 Var(\bar{Y}_{2.}) + 2\, cov(\bar{Y}_{2.}, \mu_2) = 0$$

türevleri hesaplanır ve buradan

$$c_1 = \frac{cov(\bar{Y}_{1.}, \mu_1)}{Var(\bar{Y}_{1.})}$$

ve

$$c_2 = -\frac{cov(\bar{Y}_{2.}, \mu_2)}{Var(\bar{Y}_{2.})}$$

bulunur. Ayrıca,

$$\bar{Y}_{1.} = \frac{\sum_{j=1}^{n_2}(\mu_1 + \varepsilon_{1j})}{n_2} = \frac{n_2\mu_1 + \sum_{j=1}^{n_2}\varepsilon_{1j}}{n_2} = \mu_1 + \bar{\varepsilon}_{1.}$$

$$\bar{Y}_{2.} = \frac{\sum_{j=1}^{n_2}(\mu_2 + \varepsilon_{2j})}{n_2} = \frac{n_2\mu_2 + \sum_{j=1}^{n_2}\varepsilon_{2j}}{n_2} = \mu_2 + \bar{\varepsilon}_{2.}$$

olduğundan

$$\text{cov}(\bar{Y}_{1.}, \mu_1) = cov(\mu_1 + \bar{\varepsilon}_{1.}, \mu_1)$$

$$= E[(\mu_1 + \bar{\varepsilon}_{1.})\mu_1] - E(\mu_1 + \bar{\varepsilon}_{1.})E(\mu_1)$$

$$= E(\mu_1^2) + E(\bar{\varepsilon}_{1.}\mu_1) - [E(\mu_1) + E(\bar{\varepsilon}_{1.})]E(\mu_1)$$

$$= E(\mu_1^2) - [E(\mu_1)]^2$$

$$= Var(\mu_1)$$

$$= 1$$

bulunur.

Benzer şekilde $\text{cov}(\bar{Y}_{2.}, \mu_2) = 1$ olur.

$$\text{Var}(\bar{Y}_{1.}) = Var(\mu_2 + \bar{\varepsilon}_{2.}) = Var(\mu_2) + Var(\bar{\varepsilon}_{2.}) = 1 + \frac{\sigma^2}{n_1}$$

ve

$$\text{Var}(\bar{Y}_{2.}) = 1 + \frac{\sigma^2}{n_2}$$

olacağından

$$c_1 = \frac{cov(\bar{Y}_{1.},\mu_1)}{Var(\bar{Y}_{1.})} = \frac{1}{\sigma^2/_{n_1}+1}$$

ve

$$c_2 = \frac{cov(\bar{Y}_{2.},\mu_2)}{Var(\bar{Y}_{2.})} = -\frac{1}{\sigma^2/_{n_2}+1}$$

elde edilir. Buradan, $(\mu_1 - \mu_2)$ için Bayes tahmin edici

$$(\mu_1 - \mu_2)_b = \frac{1}{\sigma^2/_{n_1}+1}\bar{Y}_{1.} - \frac{1}{\sigma^2/_{n_2}+1}\bar{Y}_{2.}$$

bulunur.

4.4. Bayes Tahmin Edicinin Alternatif Formları

4.4.1. X matrisi tam ranklı ise

Teorem 4.3: X matrisi tam ranklı ve H pozitif tanımlı bir matris olsun. Bayes tahmin edicinin alternatif formları

$$p'\hat{\beta} = p'\theta + p'[X'X + \sigma^2 H^{-1}]^{-1}X'(y - X\theta) \tag{4.17}$$

$$= p'[X'X + \sigma^2 H^{-1}]^{-1}[X'y + H^{-1}\theta] \tag{4.18}$$

$$= p'\theta + p'F[\sigma^2(X'X)^{-1} + H]^{-1}(b - \theta) \tag{4.19}$$

$$= p'(X'X)^{-1}[H + \sigma^2(X'X)^{-1}]^{-1}\theta\sigma^2 + p'H[H + \sigma^2(X'X)^{-1}]^{-1}b \tag{4.20}$$

,

$$= p'b - p'\,(X'X)^{-1}[H + \sigma^2(X'X)^{-1}]^{-1}(b - \theta)\sigma^2 \tag{4.21}$$

şeklindedir. Burada $p'b = p'(X'X)^{+}X'y$, EKK tahmin edicidir.

İspat: Teoremin ispatı için

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1} \tag{4.22}$$

matris eşitliği kullanılacaktır.

1. (3.22) 'de $A = \sigma^2 I$, $B = X$, $C = F$ ve $D = X'$ alınırsa

$$(\sigma^2 I + XHX')^{-1} = \frac{1}{\sigma^2}I - \frac{X}{\sigma^2}\left(H^{-1} + \frac{1}{\sigma^2}X'X\right)^{-1}\frac{X'}{\sigma^2} \tag{4.23}$$

elde edilir. (3.23), (3.14) 'de yerine konulursa,

$$p'\hat{\beta} = p'\theta + p'HX'\left[\frac{1}{\sigma^2}I - \frac{X}{\sigma^2}\left(H^{-1} + \frac{1}{\sigma^2}X'X\right)^{-1}\frac{X'}{\sigma^2}\right](y - X\theta) \tag{4.24}$$

$$= p'\theta + p'HX'\left[I - X(\sigma^2 H^{-1} + X'X)^{-1}\frac{X'}{\sigma^2}\right](y - X\theta)$$

$$= p'\theta + p'HX'(y - X\theta) - p'HX'X(\sigma^2 H^{-1} + X'X)^{-1}\frac{X'}{\sigma^2}(y - X\theta)$$

$$= p'\theta + p'HX'(y - X\theta) - p'\frac{HX'X}{\sigma^2}\left(H^{-1} + \frac{1}{\sigma^2}X'X\right)^{-1}\frac{X'}{\sigma^2}(y - X\theta)$$

elde edilir. Şimdi,

$$I - \sigma^2 H^{-1}(\sigma^2 H^{-1} + X'X)^{-1} = X'X(X'X + \sigma^2 H^{-1})^{-1} \tag{4.25}$$

eşitliğinin sağlandığı gösterilmelidir. (4.22) 'de $A = I$, $B = \sigma^2 H^{-1}$, $C^{-1} = X'X$ ve $D = I$ alındığında,

$$I - \sigma^2 H^{-1}(\sigma^2 H^{-1} + X'X)^{-1} = [I + \sigma^2 H^{-1}(X'X)^{-1}]^{-1}$$

$$= [(X'X + \sigma^2 H^{-1})(X'X)^{-1}]^{-1}$$

$$= (X'X)(X'X + \sigma^2 H^{-1})^{-1}$$

olduğu görülür. (4.25) eşitliği (4.24) 'de yerine koyulursa,

$$p'\hat{\beta} = p'\theta + p'HX'(y - X\theta) - p'\frac{HX'X}{\sigma^2}\left(H^{-1} + \frac{1}{\sigma^2}X'X\right)^{-1}\frac{X'}{\sigma^2}(y - X\theta)$$

$$= p'\theta + p'HX'(y - X\theta) - p'\frac{HX'X}{\sigma^2}\left(\frac{\sigma^2 H^{-1} + X'X}{\sigma^2}\right)^{-1}\frac{X'}{\sigma^2}(y - X\theta)$$

$$= p'\theta + p'HX'(y - X\theta) - p'HX'X(\sigma^2 H^{-1} + X'X)^{-1}\frac{X'}{\sigma^2}(y - X\theta)$$

$$= p'\theta + p'H[I - X'X(\sigma^2 H^{-1} + X'X)^{-1}]\frac{X'}{\sigma^2}(y - X\theta)$$

$$= p'\theta + p'H[I - I + \sigma^2 H^{-1}(\sigma^2 H^{-1} + X'X)^{-1}]\frac{X'}{\sigma^2}(y - X\theta)$$

$$= p'\theta + p'H[\sigma^2 H^{-1}(\sigma^2 H^{-1} + X'X)^{-1}]\frac{X'}{\sigma^2}(y - X\theta)$$

$$= p'\theta + p'H\sigma^2 H^{-1}(\sigma^2 H^{-1} + X'X)^{-1}\frac{X'}{\sigma^2}(y - X\theta)$$

$$= p'\theta + p'(\sigma^2 H^{-1} + X'X)^{-1}X'(y - X\theta)$$

ifadesi elde edilir. Bu ifade (4.17) 'deki Bayesyen tahmin edicidir.

2. (1) 'deki ispata benzer şekilde

$$I - (X'X + \sigma^2 H^{-1})^{-1}X'X = (X'X + \sigma^2 H^{-1})^{-1}\sigma^2 H^{-1}$$

olduğu kullanılarak (3.17) 'den (3.18) eşitliği elde edilebilir.

3.

$$(X'X + \sigma^2 H^{-1})^{-1} = H[\sigma^2 (X'X)^{-1} + H]^{-1}(X'X)^{-1} \tag{4.26}$$

ve

$$X'(y - X\theta) = (X'X)(X'X)^{-1}X'(y - X\theta) = (X'X)(b - \theta) \tag{4.27}$$

şeklindedir.

(4.19) alternatif formu da benzer şekilde, (4.26) ve (4.27) ifadelerini (4.17) 'de yerlerine koyarak elde edilir.

4. Dördüncü alternatif form (4.20),

$$I - H[\sigma^2(X'X)^{-1} + H]^{-1} = \sigma^2(X'X)^{-1}[\sigma^2(X'X)^{-1} + H]^{-1} \tag{4.28}$$

ifadesini (3.19) 'da yerine koyarak elde edilir.
5**.** (4.21), (4.28) 'i (4.20) 'de yerine koyarak elde edilebilir.

4.4.2. X matrisi tam ranklı değilse

Teorem 4.4: X matrisi tam ranklı olmasın. $U'HU$ pozitif tanımlı ve $p'\beta$ parametrik fonksiyonu tahmin edilebilir olsun. Bayes tahmin edicinin alternatif formları

$$p'\hat{\beta} = p'\theta + p'[X'X + \sigma^2(UU'HUU')^+]^+X'(y - X\theta) \tag{4.29}$$

$$= P'[X'X + \sigma^2(UU'HUU')^+]^+[X'y + (UU'HUU')^+\theta] \tag{4.30}$$

$$= p'\theta + p'F[\sigma^2(X'X)^+ + UU'FUU']^+(b - \theta)$$

$$= p'(X'X)^+[UU'FUU' + \sigma^2(X'X)^+]^+\theta\sigma^2$$

$$+p'UU'FUU'[UU'FUU' + \sigma^2(X'X)^+]^+$$

$$= p'b - p'(X'X)^+[UU'FUU' + \sigma^2(X'X)^+]^+(b - \theta)\sigma^2 \tag{4.31}$$

eşitlikleriyle ifade edilir.
İspat: Teoremin ispatı için $y = X\beta + \varepsilon$ lineer modeli tam ranklı olacak şekilde yeniden parametrize edilsin. $XUU' = X$ olduğu göz önüne alınırsa

$$y = X\beta + \varepsilon \tag{4.32}$$

$$= XUU'\beta + \varepsilon$$

$$= XU\gamma + \varepsilon$$

olur. Burada $\gamma = U'\beta$ olup, önsel varsayımlar

$$E(\gamma) = U'\theta \; ve \; D(\gamma) = U'HU$$

şeklindedir. $p'\beta$ parametrik fonksiyonu tahmin edilebilir olduğundan $p = Ud$ olacak şekilde bir d vektörü vardır. $U'X'XU = 1$ olduğu göz önüne alınsın. Teorem 4.3 'den, yeniden parametrize edilmiş model için Bayes tahmin edicinin alternatif formları

$$d'\hat{\gamma} = d'U'\theta + d'[\Lambda + \sigma^2(U'HU)^{-1}]^{-1}X'(y - XU'\theta) \quad (4.33)$$

$$= d'[\Lambda + \sigma^2(U'HU)^{-1}]^{-1}[U'Xy + (U'HU)^{-1}U'\theta\sigma^2]$$

$$= d'U'\theta + d'U'HU[\sigma^2\Lambda^{-1} + U'HU]^{-1}(g - U'\theta)$$

$$= d'\Lambda^{-1}[U'HU + \sigma^2\Lambda^{-1}]^{-1}U'\theta + d'U'HU[U'HU + \sigma^2\Lambda^{-1}]^{-1}g$$

$$= d'g - d'\Lambda^{-1}[U'HU + \sigma^2\Lambda^{-1}]^{-1}(g - U'\theta)\sigma^2 \quad (4.34)$$

şeklindedir. Burada $g = \Lambda^{-1}U'X'y$, (4.32) modelindeki γ için EKK tahmin edicidir.

$X'X = U\,\Lambda U'$ ve $(X'X)^+ = U\Lambda^{-1}U'$ ve pozitif tanımlı bir A matrisi için $UA^{-1}U' = (UAU')^+$ olduğu göz önüne alınırsa,

$$U'[\Lambda + \sigma^2(U'HU)^{-1}]^{-1}U = [X'X + \sigma^2(UU'HUU')^+]^+ \quad (4.35)$$

ve

$$U'[U'HU + \sigma^2\Lambda^{-1}]^{-1}U = [UU'FUU' + \sigma^2(X'X)^+]^+ \quad (4.36)$$

elde edilir. $UU' = I$ ve $g = \Lambda^{-1}U'X'y = U'U\Lambda^{-1}U'X'y = U'b$ olup, (4.35) ve (4.36), (4.33) 'den (4.34) 'e kadar olan eşitliklerde yerlerine konulursa, (4.29) 'dan (4.31) 'e kadar olan sonuçlar elde edilir. Buradaki işlemler Teorem 4.3 'ün ispatındaki gibi olduğundan tekrar yapılmayacaktır.

(4.30) formu, ridge tahmin edicinin genelleştirilmiş halidir. (4.31) formu, empirical Bayes tahmin edicileri formülize ederken ve Liu tahmin edicilerin bazı özelliklerini ispatlarken faydalı olmaktadır. Sonuç olarak (4.30) ve (4.31) formundaki tahmin ediciler ridge-type tahmin ediciler olarak adlandırılır.

F, negatif tanımlı olmayan bir matris olmak üzere

$$H = \frac{1}{1-d}[UU'FUU' + d(X'X)^+\sigma^2]$$

ve $M = (UU'HUU')^+$ ise,

$$M = (1-d)[UU'FUU' + d(X'X)^+\sigma^2]^+$$

olur. Bu ifadeler (3.31) 'de yerine konulursa Liu-type tahmin edici,

$$p'\hat{\beta}_{LT} = p'b - p'(X'X)^+\left[\sigma^2(X'X)^+ + \frac{1}{1-d}(UU'FUU' + d\sigma^2(X'X)^+)\right]^+(b-\theta)\sigma^2$$

$$= p'b - (1-d)p'(X'X)^+[\sigma^2(X'X)^+ + UU'FUU']^+(b-\theta)\sigma^2$$

$$= dp'b + (1-d)p'\hat{\beta}_c \quad (4.37)$$

elde edilir. Bu tahmin edici,

$$E(\beta) = \theta \; ve \; D(\beta) = F$$

varsayımları sağlandığında EKK tahmin edici ile lineer Bayes tahmin edicilerin bir konveks kombinasyonudur. Lineer Bayes tahmin edicinin alternatif formları kullanılarak (4.37)

$$p'\hat{\beta}_{LT} = dp'b + (1-d)p'[X'X + \sigma^2(UU'FUU')^+]^+[X'y + \sigma^2(UU'FUU')^+\theta]$$

$$= p'[X'X + \sigma^2(UU'FUU')^+]^+[(X'y + \sigma^2(UU'FUU')^+\theta)$$

$$+d\sigma^2(UU'FUU')^+(b-\theta)] \qquad (4.38)$$

ifadeleriyle yeniden yazılabilir.

X matrisinin tam ranklı olması durumunda $\theta = 0$ ve $F = \frac{\sigma^2}{k}I$ iken iki parametreli Liu tahmin edici elde edilir. Bu nedenle (4.38), EKK tahmin edici ile ridge tipi tahmin edicilerin bir konveks kombinasyonudur. Sonuçta, (4.37) ve (4.38) Liu tipi tahmin ediciler olarak adlandırılır.

5. GECİKMESİ DAĞITILMIŞ MODELLERDE BAYESYEN LİU VE RİDGE TİPİ TAHMİN EDİCİLER

Bu bölümde Bayesyen yaklaşım, gecikmesi dağıtılmış modellere taşınacaktır. Ayrıca, bu modellerde Liu ve ridge tipi tahmin ediciler ele alınacaktır. Shiller ve Lindley tahmin ediciler kullanılarak iki yeni Bayesyen Liu tipi tahmin edici tanımlanacaktır.

5.1. Giriş

Gecikmesi sonlu dağıtılmış model,

$$y_t = \beta_0 x_t + \beta_1 x_{t-1} + ... + \beta_p x_{t-p} + u_t \qquad t = p+1,.....,T$$

$$= \sum_{i=0}^{p} \beta_i x_{t-i} + u_t \qquad , \qquad u_t \sim IN\left(0, \sigma_u^2\right) \tag{5.1}$$

şeklindedir. β_i katsayıları gecikme ağırlıkları olarak adlandırılır. (5.1) modeli matris formunda

$$y = X\beta + u \tag{5.2}$$

şeklinde yazılabilir. Burada

$$y = \begin{bmatrix} y_{p+1} \\ y_{p+2} \\ \vdots \\ y_T \end{bmatrix}, \ \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}, \ X = \begin{bmatrix} x_{p+1} & x_p & \dots & x_1 \\ x_{p+2} & x_{p+1} & \dots & x_2 \\ \vdots & \vdots & \ddots & \vdots \\ x_T & x_{T-1} & \dots & x_{T-p} \end{bmatrix}, \ u = \begin{bmatrix} u_{p+1} \\ u_{p+2} \\ \vdots \\ u_T \end{bmatrix}$$

formundadır.

Bağımsız değişkenler arasında çoklu iç ilişki problemi varsa, (5.1) modelinin direkt olarak bilinen EKK yöntemiyle tahmin edilmesi durumunda bazı problemlerle

karşılaşılır. Gecikmesi dağıtılmış modellerdeki bu tür problemleri çözmek amacıyla çeşitli alternatif yöntemler önerilmiştir. Bu modellerin hemen hemen hepsi (5.1) 'deki β 'ların davranışı hakkında bazı önbilgilerin tanımlanmasını gerektirmektedir. Genel olarak bu önbilgi stokastik ve stokastik olmayan düzeltilmiş önbilgi şeklinde sınıflandırılmaktadır (Vinod ve Ullah, 1981; Gujarati,1999).

Irving Fisher (1937) ilk olarak stokastik olmayan düzeltilmiş ön bilgiyi

$$\beta_i = (p+1-i)\alpha \qquad 0 \le i \le p \tag{5.3}$$

$$= 0 \qquad i > p$$

formunda vermiştir. Burada α herhangi bilinmeyen parametredir. (5.1) 'de (5.3) 'ün yerleştirilmesiyle,

$$y_t = \left[\sum_{i=0}^{p}(p+1-i)x_{t-i}\right]\alpha + u_t \tag{5.4}$$

$$= z_t\alpha + u_t$$

elde edilir. Böylece α 'nın EKK tahmini (5.4) numaralı modelden bulunabilir ve (5.3) 'ün kullanılmasıyla β_i 'lerin tahmini bulunabilir.

β_i üzerinde lineer stokastik olmayan önbilginin genelleştirilmesi,

$$\beta_i = \alpha_0 + \alpha_1 i + \alpha_2 i^2 + \ldots + \alpha_r i^r \qquad p \ge r \ge 0 \tag{5.5}$$

şeklinde r-inci dereceden bir polinom olarak yazılabilir. β_i gecikme ağırlıklarının bu yapısı Almon (1965) tarafından verilmiştir. Bu yüzden Almon gecikme polinomu olarak bilinmektedir. Tekrar (5.5) 'in (5.1) 'de yerleştirilmesiyle α 'ların tahminlerini ve (5.5) 'in kullanılmasıyla β_i 'lerin tahminlerini bulabiliriz. (5.5) matris formunda,

$$\beta = A\alpha \tag{5.6}$$

şeklinde yazılabilir. Burada $A:(p+1)\times(r+1)$ tipinde bir matris ve $\alpha:(r+1)\times 1$ tipinde bir vektör olmak üzere;

$$A=\begin{bmatrix} 1 & 0 & 0 & \dots 0 \\ 1 & 1 & 1 & \dots 1 \\ \vdots & \vdots & \vdots & \ddots \vdots \\ 1 & p & p^2 & \cdots p^r \end{bmatrix}, \qquad \alpha=\begin{bmatrix} \alpha_0 \\ \vdots \\ \alpha_r \end{bmatrix}$$

formundadır. X ve A matrislerinin ranklarının $(p+1)<(T-p)$ ve $(r+1)<(p+1)$ olduğu varsayılmaktadır. Eğer $r<p$ ise A ' nın rankı $(r+1)$ olur.

5.2. Stokastik Olmayan Ön Bilgi Altında Tahmin

(5.6) ile verilen β üzerinde stokastik olmayan önbilgi altında, (5.2) 'deki β yı Almon metodu ile tahmin edebiliriz. Bunu yapmanın iki yolu vardır. (5.6), (5.2)'de yerine yazılırsa,

$$y = XA\alpha + u$$

$$= Z\alpha + u \tag{5.7}$$

elde edilir. α ' nın EKK tahmin edicisi,

$$\hat{\alpha} = (Z'Z)^{-1}Z'y \tag{5.8}$$

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

dir. β 'nın Almon tahmin edicisi bu durumda ,

$$\hat{\beta}_A = A\hat{\alpha} \tag{5.9}$$

olur. $\hat{\beta}_A$ en iyi lineer yansız tahmin edicidir (BLUE).

$\hat{\beta}_A$ 'yı elde etmenin diğer bir yolu ise Kısıtlı En Küçük Kareler (RLS) yöntemi uygulamaktadır. A 'nın kolonları lineer bağımsız olduğundan ,

$$M = I - A\left(A'A\right)^{-1}A' \tag{5.10}$$

şeklinde idempotent bir matris tanımlanabilir. Bu durumda $\beta = A\alpha$ olması

$$M\beta = 0 \tag{5.11}$$

olmasını gerektirir. (4.2) ile (4.11) 'in birlikte çözülmesiyle RLS tahmin edicisi

$$\hat{\beta}_R = b - S^{-1}M'\left[MS^{-1}M'\right]^{+}M\,b \tag{5.12}$$

şeklinde elde edilir. Burada b , (5.2) modelinden elde edilen $b = S^{-1}X'y$ şeklinde EKK tahmin edicisidir ve $+$ genelleştirilmiş tersi gösterir. M matrisi, Terasvirta (1976) nın kullanılmasıyla,

$$M = I - A\left(A'A\right)^{-1}A' = R'\left(RR'\right)^{-1}R = M^2 \tag{5.13}$$

şeklinde yazılabilir. Burada $R:(p-r)\times(p+1)$ tipinde aşağıdaki gibi bilinen bir matristir.

$$R=\begin{bmatrix} (-1)^0\binom{r+1}{0} & (-1)^1\binom{r+1}{1} & \cdots & (-1)^{r+1}\binom{r+1}{r+1} & 0 & \cdots & 0 \\ 0 & (-1)^0\binom{r+1}{0} & \cdots & (-1)^{r}\binom{r+1}{r} & (-1)^{r+1}\binom{r+1}{r+1} & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & \cdots & (-1)^{r+1}\binom{r+1}{r+1} & \end{bmatrix}$$

$M\beta = 0$ olması aynı zamanda,

$$R\beta = 0 = RA\alpha \tag{5.14}$$

olmasını gerektirir. (5.2) ile (5.14) birlikte çözülürse RLS tahmin edicisi

$$\hat{\beta}_R = b - S^{-1}R'\left[RS^{-1}R'\right]^{-1}Rb \tag{5.15}$$

şeklinde tahmin edilir. Bu ifade, (5.9) ve (5.12) ile özdeştir. $R\beta = 0$ doğru değilse $\hat{\beta}_R$ 'de yanlı olacaktır.

Almon tahmin edicinin bazı dezavantajlarından dolayı karşılaşılan problemleri gidermek için Hoerl ve Kennard (1970) 'ın ridge tahmin edici yaklaşımı alternatif olarak ele alınmıştır (Maddala ,1974; Vinod ve Ullah ,1981; Chanda ve Maddala ,1984).

(5.7) modelinden α 'nın Ridge tahmin edicisi

$$\hat{\alpha}_k = \left(Z'Z + kI\right)^{-1}Z'y$$

$$= \left(A'SA + kI\right)^{-1}A'X'y$$

ve

$$\hat{\beta}_k = A\hat{\alpha}_k \tag{5.16}$$

şeklinde elde edilir.

5.3. Stokastik Ön Bilgi Altında Tahmin

Almon tahmin edici p ve r'nin gerçek değerleri bilindiğinde ve $\beta = A\alpha$ formu doğru ise BLUE özelliğine sahiptir. p ve r'ye bağlı herhangi bir yanlış bilgi Almon tahmin edicisinin etkinliğini kaybetmesine neden olabilir. Ayrıca p'nin değeri doğru olarak belirlense bile gecikme ağırlıkları beklenen r-inci dereceden olmayabilir. Bu gibi nedenlerden dolayı önbilginin kesin olmadığını varsayarak,

$$\beta = A\alpha + v \tag{5.17}$$

formu daha esnek bir yaklaşım olabilir. $v \sim N\left(0, \sigma_v^2 I\right)$ dağılımına sahiptir. $RA\alpha = 0$ olmasından dolayı β üzerindeki stokastik kısıt,

$$R\beta = Rv \tag{5.18}$$

$Rv \sim N\left(0, \sigma_v^2 RR'\right)$ dağılımına sahiptir. Bundan sonra $R\beta = Rv$ stokastik önbilgisi altında (5.2) ile verilen genel lineer model tahmin edilebilir. Bunun için,

$$S = \frac{(y - X\beta)'(y - X\beta)}{\sigma_u^2} + \frac{(R\beta)'(RR')^{-1}(R\beta)}{\sigma_v^2}$$

eşitliği β'ya göre minimize edilmelidir. Bunun sonucunda elde edilen tahmin edici,

$$\hat{\beta}_M = \left(\frac{X'X}{\sigma_u^2} + \frac{R'(RR')R}{\sigma_v^2} \right)^{-1} \left(\frac{X'y}{\sigma_u^2} \right) = (X'X + \lambda M)^{-1} X'y \tag{5.19}$$

şeklinde Theil-Goldberger' in mixed tahmin edicisidir. Burada $\lambda = \sigma_u^2 / \sigma_v^2$ ve $M = I - A(A'A)^{-1}A$ şeklindedir.

(5.18) 'deki kovaryans matrisi $Rv \sim N(0, \sigma_v^2 I)$ varsayılırsa, $R\beta = Rv$ kısıtı altında β'nın tahmin edicisi Shiller (1973) tarafından,

$$\hat{\beta}_{SH} = (X'X + \lambda R'R)^{-1} X'y \tag{5.20}$$

olarak tanımlanmıştır.

$\hat{\beta}_M$ ve $\hat{\beta}_{SH}$ arasındaki fark Shiller'in $R\beta = 0$ olmasını sağlayan $\beta = A\alpha$ ile başlaması ve $R\beta = v$ gibi stokastik kısıtlar belirlemek için v gibi stokastik bir vektör kullanmasından kaynaklanmaktadır. Böylece $\beta = A\alpha$' yı ya da $R\beta = 0$'ı stokastik yaparak fark yaratılır.

5.3.1. Stokastik Ön Bilginin Bayesyen Yorumu

Stokastik bilgiyi ele almanın en etkili ve anlamlı yolu Bayesyen yaklaşımdır. Gecikme ağırlığı vektörü β normal önbilgiye sahip olduğu zaman,

$$\beta \sim N\left(A\alpha, \sigma_v^2 I \right) \ , \ \sigma_v^2 = \sigma_\beta^2$$

ve örneklem bilgisi,

$$y \sim N(X\beta, \sigma_u^2 I)$$

olur. Böylece sonsal ortalama, σ_v^2 ve α biliniyorken

$$\hat{\beta}_B = \left[\frac{X'X}{\sigma_u^2} + \frac{I}{\sigma_v^2}\right]^{-1}\left[\frac{X'y}{\sigma_u^2} + \frac{A\alpha}{\sigma_v^2}\right] \tag{5.21}$$

şeklindedir. Burada $\hat{\beta}_B$, karesel kayıp ve normal önbilgi altında, β'nın Bayes tahmin edicisini göstermektedir. Ancak β üzerinde belirsiz önbilgi varsa $\hat{\beta}_B = (X'X)^{-1}X'y = b$, yani EKK tahmin edici olur.

α parametresinin nadir olarak bilinmesinden dolayı, α'nın normal dağıldığını varsayarak bir adım ileri gidilebilir. Böylece Lindley ve Smith' in Bayesyen modeli,

$$y \sim N(X\beta, \sigma_u^2 I)$$

$$\beta \sim N(A\alpha, \sigma_v^2 I)$$

$$\alpha \sim N(\overline{\alpha}, \sigma_\alpha^2 I)$$

olur. Bu durumda β'nın sonsal yoğunluğu,

$$D_0^{-1} = \left[\frac{X'X}{\sigma_u^2} + \frac{I}{\sigma_v^2} - \frac{A}{\sigma_v^2}\left[\frac{A'A}{\sigma_v^2} + \frac{I}{\sigma_\alpha^2}\right]^{-1}\frac{A'}{\sigma_v^2}\right]$$

$$d_0 = \left[\frac{X'y}{\sigma_u^2} + \left[\frac{I}{\sigma_v^2} - \frac{A}{\sigma_v^2}\left[\frac{A'A}{\sigma_v^2} + \frac{I}{\sigma_\alpha^2}\right]\frac{A'}{\sigma_v^2}\right]A\overline{\alpha}\right]$$

ve $\beta^* = D_0 d_0$ olmak üzere $N(\beta^*, D_0)$ 'dır. Ancak α için belirsiz önbilgi kullanılırsa, yani $\sigma_\alpha^2 \to \infty$ ise β^*,

$$\begin{aligned}\hat{\beta}_L = \beta^* &= (X'X + \lambda M)^{-1} X'y \\ &= (I + \lambda (X'X)^{-1} M)^{-1} b, \qquad \lambda = \frac{\sigma_u^2}{\sigma_v^2}\end{aligned} \tag{5.22}$$

formunda basitleşir. Burada $b = (X'X)^{-1} X'y$ olan EKK tahmin edici, $M = I - A(A'A)^{-1} A'$ ve L alt indisi de Lindley 'i göstermektedir. (5.13)'ün kullanılmasıyla β^*, (5.19) 'daki mixed tahmin ediciyle aynı olur. Ancak yorumları farklıdır. (5.22) eşitliği,

$$\begin{aligned}\beta^* &= (X'X + \lambda I)^{-1} (X'Xb + \lambda A\alpha^*) \\ &= A\alpha^* + [I + \lambda (X'X)^{-1}]^{-1} (b - A\alpha^*)\end{aligned} \tag{5.23}$$

formunda da yazılabilir. Burada,

$$\alpha^* = \left[A'\left[(X'X)^{-1} + \lambda^{-1} I\right]^{-1} A\right]^{-1} A'\left[(X'X)^{-1} + \lambda^{-1} I\right]^{-1} b \tag{5.24}$$

şeklinde α'nın sonsal ortalamasıdır. (5.23) ve (5.24) eşitliklerinden, $\beta^* = A\alpha + v$ modelinin EKK tahmin edicisinin,

$$\alpha^* = (A'A)^{-1} A'\beta^* \tag{5.25}$$

olduğu görülebilir.

β^*'ın sonsal ortalaması, b ve $A\alpha^*$ tahmin edicilerinin ağırlıklı matris kombinasyonu gibidir. Ayrıca β^*, EKK tahmin edici b'yi $A\alpha$'nın sonsal ortalaması olan $A\alpha^*$ yönünde büzer. (5.22) 'deki λ ile ilgili olarak Lindley (1971) ve Lindley Smith (1972), β'nın EKK tahminiyle başlayan $(\lambda = 0)$ ve σ_u^2 ve σ_v^2 için $(y - Xb)'(y - Xb)/n_1$, $(b - Aa)'(b - Aa)/n_2$ şeklinde tahminlerini kullanan iteratif bir yöntem önermişlerdir. Burada $\lambda = 0$ için (bkz. (5.25)) $a = (A'A)^{-1}A'b$ ve n_1, n_2 sabitlerdir. λ, σ_u^2 ve σ_v^2'nin bu tahminlerinin kullanılmasıyla,

$$\hat{\lambda} = \frac{gs}{b'Mb}, \qquad s = (y - Xb)'(y - Xb),$$

şeklinde yazılabilir. Burada g keyfi bir sabit ve $M = I - A(A'A)^{-1}A'$'dır.

Ancak, ridge tahmin edicisi ve Lindley ve Smith tarafından verilen genişletilmiş şekli gecikmesi dağıtılmış modeller için çok umut verici değildir (Maddala ,1974). Bu nedenle alternatif tahmin yöntemlerine ihtiyaç duyulmaktadır.

5.4. Bayesyen Liu Tipi Tahmin Ediciler

$$y = X\beta + \varepsilon \quad \varepsilon \sim (0, \sigma^2 I) \tag{5.26}$$

lineer regresyon modelini göz önüne alalım. Burada, y, gözlemlerden oluşan n boyutlu bir vektör; X, $s \leq m$ ranklı $n \times m$ tipinde bir matris; β, m boyutlu parametrelerin bir vektörü ve ε, n boyutlu hatalar vektörüdür. Çoklu iç ilişki olması durumunda kullanılan ridge tipi tahmin ediciler bir çok yazar tarafından çalışılmıştır. Liu (1993) de ridge tahmin ediciye benzer formda fakat yanlılık parametresi farklı olan

$$\hat{\beta}_L = (X'X + I)^{-1}(X'y + db), \qquad 0 < d < 1 \tag{5.27}$$

tahmin ediciyi tanımlamıştır. Özkale ve Kaçıranlar (2007) bu tahmin edicinin iki parametreli versiyonunu,

$$\hat{\beta}_{OK} = (X'X + kI)^{-1}(X'y + kdb), \quad 0 < d < 1 \tag{5.28}$$

şeklinde tanımlamıştır. (5.26) modelindeki β 'nın, sırasıyla ortalaması ve varyans kovaryans matrisi

$$E(\beta) = \theta \; ve \; D(\beta) = H \tag{5.29}$$

olan bir rasgele değişken olduğu varsayılsın. Burada H, negatif tanımlı olmayan bir matristir. Ayrıca,

$$E(\varepsilon/\beta) = 0 \; ve \; D(\varepsilon/\beta) = \sigma^2 I \tag{5.30}$$

olduğu varsayılsın. Rao (1973) ve Gruber (2012) 'de lineer Bayes tahmin edici

$$\hat{\beta}_c = \theta + HX' \, (XHX' + \sigma^2 I)^{-1}(y - X\theta) \tag{5.31}$$

şeklinde verilmiştir.

X, $s \leq m$ ranklı $n \times m$ boyutlu ve $U{:}\, m \times s$, $V{:}m \times (m-s)$, $S'{:}n \times s$, $T'{:}n \times (n-s)$ matrisler olmak üzere X 'in tekil değer ayrışımı,

$$X = [S' \quad T'] \begin{bmatrix} \Lambda^{1/2} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} U' \\ V' \end{bmatrix} = S' \Lambda^{1/2} U'$$

şeklindedir. Burada $[S' \quad T']$ ve $[U \quad V]$ ortogonal matrislerdir ve $\Lambda{:}s \times s$, $X'X$ matrisinin özdeğerlerinden oluşan köşegen matristir. λ_i özdeğerleri, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_s$ olacak biçimde sıralanmıştır.

$$X'X = U\Lambda^{1/2}SS'\Lambda^{1/2}U' = U\Lambda U' \tag{5.32}$$

$$XUU' = S'\Lambda^{1/2}U'UU' = S'\Lambda^{1/2}U' = X$$

ve

$$XV = S'\Lambda^{1/2}U'V = 0$$

olduğu açıktır. Bir J matrisinin Moore Penrose inversini J^{+} ile gösterelim. N, $n \times n$ tipinde singüler olmayan bir matris ise

$$(UNU')^{+} = UN^{-1}U' \tag{5.33}$$

olup, (5.32) ve (5.33) eşitliklerinden

$$(X'X)^{+} = U\Lambda^{-1}U'$$

elde edilir. $U'HU$ pozitif tanımlı olduğundan, $\hat{\beta}_c$ Bayes tahmin edicinin alternatif formları Gruber (1998, 2010)'da

$$\hat{\beta}_c = [X'X + \sigma^2(UU'HUU')^{+}]^{+}[X'y + (UU'HUU')^{+}\theta\sigma^2] \tag{5.34}$$

ve

$$\hat{\beta}_c = b - (X'X)^{+}[UU'HUU' + \sigma^2(X'X)^{+}]^{+}(b - \theta)\sigma^2 \tag{5.35}$$

şeklinde verilmiştir. (5.34) formu ridge tahmin edicinin genel formudur. (5.35) formu Bayes tahmin edicilerin genel formudur. Sonuç olarak (5.34) ve (5.35) formundaki tahmin ediciler ridge-tipi Bayesyen tahmin ediciler olarak adlandırılır. $\hat{\beta}_{SH}$ ve $\hat{\beta}_L$ nin (5.34) formunda tahmin ediciler olduğu kolayca görülür.

F, negatif tanımlı olmayan bir matris olmak üzere

$$H = \frac{1}{1-d}[UU'FUU' + d(X'X)^{+}\sigma^2] \tag{5.36}$$

ve

$M = (UU'HUU')^{+}$ ise,

$$M = (1-d)[UU'FUU' + d(X'X)^{+}\sigma^2]^{+} \tag{5.37.}$$

olur. Bu ifadeler (5.35) 'de yerine konulursa Liu-tipi tahmin edici,

$$\hat{\beta}_{LT} = b - (X'X)^{+}\left[\sigma^2(X'X)^{+} + \frac{1}{1-d}(UU'FUU' + d\sigma^2(X'X)^{+})\right]^{+}(b-\theta)\sigma^2 \tag{5.38}$$

$$= b - (1-d)(X'X)^{+}[\sigma^2(X'X)^{+} + UU'FUU']^{+}(b-\theta)\sigma^2$$

$$= db + (1-d)\hat{\beta}_c$$

elde edilir. Bu tahmin edici,

$$E(\beta) = \theta \; ve \; D(\beta) = F$$

varsayımları sağlandığında EKK tahmin edici ile lineer Bayes tahmin edicilerin bir konveks kombinasyonudur.

Şimdi ridge-tipi Bayesyen tahmin ediciler sınıfına ait olan $\hat{\beta}_{SH}$ ve $\hat{\beta}_{L}$'nin (5.38) 'de kullanılmasıyla, sırasıyla Shiller tahmin ediciyle EKK tahmin edicinin ve Lindley tahmin ediciyle EKK tahmin edicinin konveks kombinasyonları yardımıyla aşağıdaki tahmin edicileri gecikmesi dağıtılmış modeller için tanımlayabiliriz:

$$\hat{\beta}_{LT}(SH) = db + (1-d)\,\hat{\beta}_{SH} \tag{5.39}$$

$$\hat{\beta}_{LT}(L) = db + (1-d)\,\hat{\beta}_{L} \tag{5.40}$$

Sadece ridge, Lindley ve Shiller tahmin edicilerin gecikmesi dağıtılmış modeller için çok umut verici olmadığını (Maddala ,1974) 'den biliyoruz. Bu nedenle (5.39) ve (5.40) 'da tanımlanan alternatif tahmin yöntemlerinin gecikmesi dağıtılmış modellerde karşılaştırması gelecekteki çalışmalarda ele alınacaktır.

6. SONUÇLAR VE ÖNERİLER

Bu çalışmada, regresyona Bayesyen yaklaşımın temelleri açıklanmıştır. Ön dağılım türlerine yer verilmiştir ve ön dağılımın belirlenmesinde nasıl bir yol izleneceği anlatılmıştır. Regresyona Bayesyen yaklaşımın teorisi üzerinde çalışılmıştır. Ayrıca, lineer regresyon modelinde Bayesyen tahmin yöntemleri incelenmiştir. En genel haliyle lineer Bayesyen tahmin edici ve alternatif formları incelenmiştir. Gruber (2012) tarafından tanımlanan Bayesyen Liu tipi tahmin ediciler sınıfı ele alınmıştır. Bayesyen yaklaşım, gecikmesi dağıtılmış modellere taşınmıştır ve bu modellerde Liu ve ridge tipi tahmin ediciler üzerinde çalışılmıştır. Shiller ve Lindley tahmin ediciler kullanılarak iki yeni Bayesyen Liu tipi tahmin edici tanımlanmıştır. Tanımlanan alternatif tahmin yöntemlerinin gecikmesi dağıtılmış modellerde karşılaştırması gelecekteki çalışmalarda ele alınacaktır.

KAYNAKLAR

ALMON, S., 1965. The Distributed Lag between Capital Appropriations and Expenditures. Econometrica, Vol 30, 96-178.

AKDENİZ, F. and KAÇIRANLAR, S., 1995. On the almost unbiased generalized Liu estimator and unbiased estimation of the bias and mse. Communications in Statistics-Theory and Methods, 24 (7), 1789–1797.

BAYES, T., 1958. Essay Towards Solving a Problem in The Doctrine of Chances. Biometrica, Vol. 45, 293-315. (Reproduction of 1763 paper)

BERGER, J. O., 1985. Statistical Decision Theory and Bayesian Analysis. 2. b., New York, Springer-Verlag Inc., 617 s.

BOX, G. E. P. and TİAO, G. C., 1992. Bayesian Inference In Statistical Analysis. Wiley Classics Library Edition, New York, John Wiley&Sons, 588 s.

CHANDA, A.K. and MADDALA, G.S., 1984. Ridge estimators for distributed lag models. Communications in Statistics - Theory and Methods, Vol 13, Issue2, 217-225.

EKİCİ, O., 2005. Bayesyen Regresyon ve Winbugs ile Bir Uygulama. İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, Ekonometri Anabilim Dalı Yüksek Lisans Tezi, 106 s.

FAREBROTHER, R. W., 1976. Further results on the mean square error of ridge regression. Journal of the Royal Statistical Society, Series B (Methodological) 38(3):248–250.

FİSHER, I., 1937. Income in theory and income taxation practice , Econometrica, 5, 1-55.

GİLES, D., E., 2010. Bayesian Estimation of a Possibly Mis-Specified Linear Regression Model, Econometrics Working Paper EWP1004, ISSN 1485-6441.

GROß, J., 2003. Restricted ridge estimation. Stat. Prob. Lett. 65 , 57-64.

GRUBER, M. H. J., 2012. Liu and Ridge Estimators-A Comparison. Communications in Statistics-Theory and Methods, 41:3739-3749.

GUJARATİ, D. N., 1999. Temel Ekonometri. (Çev.Ü. Şenesen, G. G. Şenesen). Literatür Yayıncılık.

HOERL, A.E. and KENNARD, R.W., 1970. Ridge regression: biased estimation for non - orthogonal problems. Technometrics 12, 69-82.

JAMES, W. and STEİN, C., 1961. Estimation with Quadratic Loss. Proceedings Fourth Berkeley Symposium in Mathematical Statistics and Probability, 1, 361-379

KUBOKAWA,T., SRİVASTAVA, M., S., 2002. Minimax Empirical Bayes Ridge-Principal Component Regression Estimators,CIRJE-F,170 (Discussion paper).

LINDLEY, D. V. SMİTH, F. M., 1971. Bayes Estimates for the Linear Model. Journal of the Royal Statistical Society, Ser.B, 34, 1-41.

LIU, K., 1993. A New Class of Biased Estimate in Linear Regression. Commun. in Statistics - Theory and Methods , 22(2), 393-402.

MADDALA, G.S., 1974. Ridge Estimator for Distributed Lag Models. NBER Working Paper Series, no:69.

MORRİS, H. D., 1970. Optimal Statistical Decisions. Wiley Classics Library Edition, New York, John Wiley&Sons, 489 s.

ÖZKALE, M. R. and KAÇIRANLAR, S., 2007. The restricted and unrestricted two parameter estimators. Communications in Statistics Theory and Methods, Vol 36(15), 2707-2725.

RAİFFA, H. and SCHLAİFER, R., 1968. Applied Statistical Decision Theory, America. MIT Press, 356 s.

SAKALLIOĞLU, S. ; KAÇIRANLAR, S. and Akdeniz, F., 1996. A Note on Combining Ridge and Least Squares Estimator. Journal of Institute of Mathematics and Computer Science (Math. Series), Vol. 9, No. 2, 193-198.

SEARLE, S. R., 1971. Linear Models. John Wiley and Sons, New York, 529 s.

SHILLER, R. J., 1973, A Distributed Lag Estimator Derived From Smoothness Priors. Econometrica, 429-449

SHU-LU, W., 1993. Estimations in the normal regression empirical Bayes model. Communication in Stat.- Theory and Methods, 22(6), 1773-1794.

SRİVASTAVA, M., S., KUBOKAWA, T., 2002. Minimax Multivariate Empirical Bayes Estimators Under Multicollinearity. Journal of Multivariate Analysis, 93, 394-416.

TERAVİSTA, T., 1976. A Note on Bias in Almon Distributed Lag Estimators. Econometrica. 1317-1322.

VINOD, H., D. ,ULLAH, A., 1981. Recent Advances in Regression Methods, Marcel Dekker, New York

YEO, S.J. and TRİVERDİ, P.K., 1989. On Using Ridge - Type Estimators for a Distributed Lag Model. Oxford Bulletin of Economics and Statistics, Vol 51, Issue 1, 85-90.

YARDIMCI, A., 1992. Çoklu Bağlantılı Çoklu Doğrusal Regresyonda Bayes Yaklaşımı. Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, 72 s.

ZELLNER, A., 1971. An Introduction to Bayesian Inference in Econometrics. John Wiley and Sons, New York.

ÖZGEÇMİŞ

01/03/1989 yılında Adana'da doğdu. İlk, orta ve lise öğrenimimi Adana'da tamamladı. 2007 yılında başladığı Çukurova Üniversitesi, Fen-Edebiyat Fakültesi, Matematik Bölümü'nden ve Çift anadal yaptığı İstatistik Bölümü'nden 2011 yılında mezun oldu. Aynı yıl İstatistik Bölümü'nde yüksek lisansa başladı. Yüksek lisans eğitimine halen devam etmektedir. 2012 yılından itibaren İstatistik Bölümü'nde Araştırma Görevlisi olarak çalışmaktadır.