

YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

DOĞRUSAL OLMAYAN BAYESÇİ REGRESYON VE
YÜKSEK FREKANSLI SES SİSTEMLERİNDE BİR UYGULAMA

İstatistikçi Meral ÇEVİK

FBE İstatistik Anabilim Dalında

Hazırlanan

YÜKSEK LİSANS TEZİ

Tez Danışmanı :Yrd. Doç. Atıf EVREN (YTÜ)

İSTANBUL,2009

İÇİNDEKİLER

Sayfa

SİMGE LİSTESİ	iv
KISALTMA LİSTESİ	vi
ŞEKİL LİSTESİ	vii
ÇİZELGE LİSTESİ	viii
ÖNSÖZ	ix
ÖZET	x
ABSTRACT	xi
1. GİRİŞ	1
2. BAYESÇİ İSTATİSTİK	3
2.1 Bayes Kuralı	3
2.2 Bayesçi Yöntemin Temelleri ve İstatistiksel Kavramlara İlişkin Yaklaşımı	6
2.3 Bayesçi Tahmin Problemi	11
2.3.1 Kayıp Fonksiyonu	11
2.3.2 Önsel ve Sonsal Dağılımlar	13
2.3.2.1 Önsel Dağılımın Belirlenmesi	13
2.3.2.2 Sonsal Dağılımın Oluşturulması	14
2.3.2.3 Eşlenik Dağılımlar	14
2.3.2.3.a Bernoulli Dağılımının Parametresinin Tahmini:	14
2.3.2.3.b Bir Poisson Dağılımından Örneklemeye:	16
2.3.2.3.c Bir Normal Dağılımın Parametrelerinin Tahmini	16
2.3.2.3.d Üstel Dağılımdan Örneklemeye	18
3. REGRESYON	20
3.1 Doğrusal Regresyon	21
4. BASİT DOĞRUSAL REGRESYONA BAYESÇİ YAKLAŞIM	24
4.1 Bayesçi Basit Doğrusal Regresyon Modeli	25
4.1.1 Belirsiz Ön Olasılık Yoğunluk Fonksiyonu ile Son Olasılık Yoğunluk Fonksiyonu Çıkarımı	27
4.1.2 Bilgi Verici Ön Olasılık Yoğunluk Fonksiyonu ile Son Olasılık Yoğunluk Fonksiyonu Çıkarımı	27
4.2 Bayesçi Çoklu Doğrusal Regresyon	29
4.2.1 Belirsiz Ön Bilgi Durumunda Çoklu Doğrusal Bayesçi Regresyon	29
4.2.2 Bilgi Verici Önsel Dağılım Fonksiyonu ile Çoklu Regresyon Modeli	30

5.	DOĞRUSAL OLMAYAN BAYESÇİ YAKLAŞIM.....	32
5.1	Benzer Fonksiyonlar	32
5.2	En İyi Model	32
5.2.1	Bayesçi Metodlar	32
5.3	Bayesçi Metodoloji.....	33
5.4	Allais Paradoksu	35
5.5	Doğrusal Olmayan Modeller İçin Ek Bağlantılar	36
5.6	Sınıflandırma İçin Doğrusal Olmayan Modeller	38
5.7	Doğrusal Olmayan Bayesçi Regresyon	43
5.7.1	Olabilirlik Fonksiyonu	45
5.7.2	Önsel Dağılım.....	45
5.7.3	Sonsal Dağılım.....	45
5.7.4	Bayesçi Hesaplama: Metropolis-Hastings Algoritması.....	46
5.7.4.1	Bağımsız Zincir Metropolis-Hastings Algoritması	48
5.7.4.2	Rastlantısal Yürüyüş Zincir Metropolis- Hastings Algoritması	51
5.7.4.3	Metropolis –within- Gibbs.....	53
5.7.5	Model Uygunluğu İçin Bir Ölçüm: Sonsal Tahmin P-Değeri	54
5.7.6	Model Karşılaştırması: Gelfand-Dey Metodu	58
5.7.7	Tahmin	61
5.7.8	Metropolis Hastings Algoritması İçin Deneysel Çalışma	61
5.8	Büyüme Modellerindeki Parametrelerin Analizinde Bayesçi Yöntem	67
5.8.1	Büyüme Modelleri, Yöntemler ve Malzemeler	67
5.9	Doğrusal Olmayan Bayesçi Regresyon İçin Genel Çıkarsama	68
6.	UYGULAMA	70
6.1	Veriler ve Ön Bilgilendirme	70
6.2	Doğrusal Olmayan Bayesçi Regresyon Yaklaşımı	71
6.3	Model ve Varsayımlar	72
6.4	Bazı Kısıtlamalar	73
6.5	Bayesçi Formüller.....	73
6.5.1	Önsel Dağılım	73
6.5.2	Olabilirlik Fonksiyonu.....	74
6.5.3	Birleşik Sonsal Yoğunluk Fonksiyonu	74
6.6	Gerçekleştirilen Analizler	75
6.6.1	Normal Dağılıma Uygunluk Kontrolü.....	75
6.6.2	Otokorelasyon Kontrolü	76
6.6.3	Önsel Dağılıma Ait Değerler	76
6.6.4	EKK Yöntemi ile Elde Edilen Parametre Tahminleri (Deneysel Veri).....	77
6.6.5	Sonsal Dağılıma Ait Değerler	77
6.7	Modelin Yorumlanması	79
7.	SONUÇ VE ÖNERİLER.....	80
	KAYNAKLAR.....	82
	EKLER	84

Ek 1 Yüksek Frekanslı Seslere ait veriler.....	84
Ek 2 Tahminler ve Kalıntılar	86
Ek 3 Metropolis Hastings fonksiyonu	88
Ek 4 Gibbs Sampler.....	90
ÖZGEÇMİŞ.....	94

SİMGE LİSTESİ

a, b	Düzgün dağılımının parametresi
f_A	Görelî frekans
$f_x(t)$	Olasılık yoğunluk fonksiyonu
$f_{xy}(.,.)$	Birikimli olasılık yoğunluk fonksiyonu
$f_x(x)$	Marjinal yoğunluk fonksiyonu
$F_x(x)$	($=F(x)$) Dağılım fonksiyonu
$F_{xc}(.)$	Sürekli dağılım fonksiyonu
$F_{xd}(.)$	Kesikli dağılım fonksiyonu
h	Gerçek doğrusunun ters fonksiyonu
J	Jacobiyen matris
$L_{göz}(v)$	Bayesçi formül
$m_k(t)$	k. dereceden moment türeten fonksiyon
n, p	Binom dağılımının parametresi
p	Bernoulli ve Geometrik dağılımının parametresi
$p(X \lambda)$	Olasılık yoğunluk fonksiyonu
$p(y \theta)$	Olabilirlik fonksiyonu
$p(\theta)$	Önsel fonksiyon
$p(\theta y)$	Sonsal fonksiyon
$P(A B)$	A' nın B' ye göre koşullu olasılığı
$P(X)$	Olasılık ölçüsü
$X(.)$	Rastlantı değişkeni
$X^{-1}(.)$	t değişkenine göre X' in tersi
v	değişim katsayısı, aynı zamanda ki-kare dağılımının parametresi
$\widehat{var}(\hat{\theta}_{ML})$	Maksimumum olabilirlik varyansının tahmini
$V(X)$	Varyans
Z	Standart rastlantı değişkeni
α	Üstel dağılımının parametresi
α, β	Gama dağılımının parametresi

α_3	Çarpıklık ölçüsü
α_4	Basıklık ölçüsü
ϕ	Birikimli dağılım fonksiyonu
$\varphi(t)$	Karakteristik fonksiyonu
μ_k	Merkezi moment(ortalamaya göre k. dereceden moment)
μ_x	(E(X))Beklenen değer
μ, σ	Normal dağılımın parametresi
θ	Ortalama
$\hat{\theta}_{ML}$	Maksimum olabilirlik ortalaması
σ	Standart sapma
σ_x	Standart sapma
$\nu(\nu)$	Gauss Dağılımının varyansı
Σ	Kovaryans matrisi
Σ_E	Hata terimlerinin kovaryans matrisi

KISALTMA LİSTESİ

CES	Üretim fonksiyonu
DORM	Doğrusal olmayan regresyon modeli
EKK(OLS)	En küçük kareler yöntemi
GLM	Genelleştirilmiş doğrusal modeller (Generalized Linear Models)
MCMC	Markov Zinciri Monte Carlo Algoritması
MMF	Morgan-Mercer-Flodin
MSE	Kalıntı kareleri ortalaması
MSR	Ortalama açıklanan kareler toplamı
o.y.f.	Olasılık yoğunluk fonksiyonu
SSE	Kalıntı kareleri toplamı (Residual sum of squares)
SSR	Açıklanan kareler toplamı (Regression sum of squares)

ŞEKİL LİSTESİ

Şekil 3.1 Gerçek ve tahmin edilen regresyon doğruları	22
Şekil 5.1 Çarpıklık için Sonsal Tahmin Dağılımı	64
Şekil 5.2 Basıklık için Sonsal Tahmin Dağılımı	65
Şekil 6.1 Y için Doğrusal Olmayan Regresyon.....	78
Şekil 6.2 Y değerinin tahmincisi	78
Şekil 6.3 Modelin kalıntıları.....	79

ÇİZELGE LİSTESİ

Çizelge 2.1 Bayeſçi yöntemler ile frekansçı yöntemlerin karşılaſtırılması.....	6
Çizelge 1.2 Bayeſçi yaklaşım ve klasik yaklaſıma göre işleyen süreçler	7
Çizelge 2.3 Örnekleme dağılımlarının ön ve son dağılımları.....	18
Çizelge 5.1 Bağımsızlık zinciri ve rassal yürüyüş yöntemlerinden elde edilen ortalama ve standart sapma değerleri	62
Çizelge 6.1 Kolmogorov-Smirnov test istatistikleri.....	75
Çizelge 6.2 Önsel dağılıma ait test istatistikleri olur.....	75
Çizelge 6.3 Kolmogorov-Smirnov test istatistikleri.....	76
Çizelge 6.4 Önsel dağılıma ait test istatistikleri olur.....	76
Çizelge 6.5 Önsel dağılıma ait genel değerler	76
Çizelge 6.6 Y değişkeni için doğrusal olmayan regresyon değerleri	76
Çizelge 6.7 Önsel Ortalama Değerleri.....	76
Çizelge 6.8 Önsel Standart Sapma Değerleri	77
Çizelge 6.9 EKK regresyon değerleri.....	77
Çizelge 6.10 Ekk yöntemi ile elde edilen parametre tahminleri.....	77
Çizelge 6.11 Sonsal ortalama değeri	77
Çizelge 6.12 Sonsal parametre değerleri	77
Çizelge 6.13 Sonsal parametre değerleri ile oluşturulan modele ait bazı istatistiksel değerler	77

ÖNSÖZ

Bu tez çalışmasında, tahmin yöntemleri ile ilgili genel bilgi verilerek, Doğrusal Olmayan Bayesçi Regresyon üzerinde durulmuştur. Bazı kavramların, Türkçe’ de tam karşılıkları olmaması ya da birden fazla karşılıkları olması nedeniyle, kavram karmaşası olmaması açısından, bu sözcüklerin yanına parantez içerisinde İngilizce karşılıkları da yazılmıştır.

Bu tezin hazırlanmasında bana danışmanlık yaparak bilgi ve desteğini esirgemeyen değerli hocam Sayın Yrd. Doç. A. Atıf Evren’ e teşekkürü bir borç bilirim.

Ayrıca tezi hazırladığım dönem içerisinde manevi desteğini esirgemeyen ailem ve dostlarıma da çok teşekkür ederim.

Meral ÇEVİK

ÖZET

Bayesçi Yaklaşım uygulamaları son yıllarda risk ve klinik çalışmalarında yoğun olarak kullanılmaya ve yeterli bilgi bulunmadığı durumlar için tercih edilebilir bir yöntem halini almıştır. Bayesçi yaklaşımda da klasik yöntemlerde olduğu gibi, görsel açıdan doğrusal ve doğrusal olmayan yapı olmak üzere, tahmin geliştirme için iki genel regresyon türü mevcuttur. Son yıllarda Türkiye’ de yoğun olarak Doğrusal Bayesçi Regresyon uygulamaları gerçekleştirilmektedir.

Bu tezin amacı ise, henüz uygulaması nadir olan doğrusal olmayan Bayesçi regresyonu bir uygulama ile açıklayarak, bu konudaki çalışmalara katkı sağlamaktır. Doğrusal Olmayan Bayesçi regresyona giriş yapılmadan önce Bayesçi İstatistik ve regresyon analizi kısaca anlatılmış, doğrusal ve doğrusal olmayan Bayesçi regresyon bu kapsamda değerlendirilmiştir. Son olarak, konunun iyi anlaşılması için yüksek frekanslı seslere dair verilerle doğrusal olmayan Bayesçi regresyon analizi uygulaması yapılmıştır.

Anahtar kelimeler: Bayesçi istatistik, regresyon, doğrusal olmayan regresyon, Bayesçi doğrusal regresyon, doğrusal olmayan Bayesçi regresyon, Metropolis-Hastings Algoritması

ABSTRACT

BAYESIAN ANALYSIS OF NONLINEAR REGRESSION AND ITS APPLICATIONS (ULTRASONIC CALIBRATION OF NONLINEAR REGRESSION AND ITS APPLICATIONS)

Bayesian Approach applications have begun to be used for risk and clinical studies in recent years. It has become a preferable method in case insufficient information exists. Similar to classical methods, there exist two basic regression types in Bayesian approach which are visually linear and nonlinear structures. In Turkey, linear Bayesian regressions applications are mostly practiced.

The aim of this research is to contribute to the studies about this subject by explaining nonlinear Bayesian regression with an application. Prior to introducing Nonlinear Bayesian regression, estimation methods are emphasized and basic information about regression is given. Afterwards, Bayesian Statistics is explained and linear and nonlinear Bayesian regression are evaluated in this content. Finally, a nonlinear Bayesian regression application regarding to ultrasonic calibration is given in order to make the subject more understandable.

Key words: Bayesian statistics, regression, nonlinear regression, bayesian linear regression, nonlinear bayesian regression, Metropolis-Hastings Algoritms

1. GİRİŞ

İstatistiğin temel konularından biri olan tahminin gerçekleştirilebilmesi için bir regresyon modelinin kurulabilmesi ve bu modelin, gerçekten tahmini gerçekleştirilecek olayı açıklayıp açıklayamadığı sorusu, bu konu üzerine birçok araştırmanın yapılmasına neden olmuştur.

Regresyon modelinin fonksiyonel bir yapı ile ifade edilerek, bu fonksiyonun şeklinin değerlendirilmesi ile regresyonun doğrusal olup olmamasına ilişkin varsayımlar ortaya çıkar. Regresyon modelinin doğrusal olup olmamasına göre çeşitli teknikler aracılığı ile parametre tahmini değerlendirmeleri yapılabilmektedir. Bu yöntemlerin en sık kullanılanları; En Küçük Kareler Yöntemi, Maksimum Olabilirlik Yöntemi, Momentler Yöntemi ve Bayesçi Yöntemlerdir.

Bu yüksek lisans tezinin amacı, son dönemlerde kullanımı artmaya başlayan, doğrusal olmayan regresyon modellerinin Bayesçi yöntemler ile değerlendirilmesine bir yenisini ekleyerek, bu yöntem ile elde edilen sonuçların, en iyi sonuç olarak gösterilip gösterilemeyeceğini incelemektir. Doğrusal olmayan regresyon analizi karmaşık bir yapıya sahiptir. Modele uygun parametre tahminleri, klasik regresyon analizindeki gibi çözümü rahat denklemler ile elde edilemezler. Çeşitli simülatlörlere ihtiyaç duyulabilir. Simülatlörlerde parametre tahminleri için iterasyonlar kullanılır. Bu yöntemler için en önemli adım, başlangıç değerinin oluşturulmasıdır. Bu değerler araştırma sırasında elde edilebileceği gibi, önceki araştırmalarda elde edilen değerler de kullanılabilir. Yapılan bütün bu araştırmalar büyük örneklem için geçerlidir. Sonuç olarak, doğrusal olmayan regresyon modeli, doğrusal yaklaşıma en iyi ve en uygun olduğu durumda geçerli olacaktır.

Bu tez çalışmasının 2. Bölüm’ünde istatistiğin bayesçi yaklaşım ile değerlendirilmesi yapılmıştır. Bayesçi istatistiğin, çıkış kaynağı olan Bayes kuralı ve bu çerçevede önsel ve sonsal dağılım ile olabilirlik fonksiyonundan bahsedilmiştir.

3. bölümde, regresyon kavramı genel bir çerçeve içerisinde değerlendirilerek 4. bölümde bayesçi doğrusal regresyonun daha iyi anlaşılabilmesi için doğrusal regresyon kısaca anlatılmıştır.

Bölüm 4’te Bayesçi doğrusal regresyon basit ve çoklu regresyon sınıflandırmasına göre anlatılmıştır.

5. Bölüm’de ise esas konu olan, doğrusal olmayan bayesçi regresyon Metropolis-Hastings Algoritması, Gelfand- Dey Metodu, büyüme modellerine değinilerek anlatılmış ve 6. Bölüm’

de bu konu ile ilgili Yüksek Frekanslı Ses Sistemlerine dair verilerle bir uygulama yapılmıştır. Bu uygulamada, deneysel veriler, Excel, Mathcad, MATLAB ve XLStat programları kullanılarak analiz edilmiştir. Biri bağımlı, biri bağımsız olmak üzere, 2 değişkenli, bir regresyon modeli oluşturulmuş, yapılmış olunan çalışma sonucunda elde edilen değerlerin, istatistiksel olarak analizi gerçekleştirilmiştir.

2. BAYESÇİ İSTATİSTİK

Bayesçi olasılık kuramı, matematiksel istatistik kuramının bir dalıdır. Bu kuram belirsizlik taşıyan herhangi bir durumun modelini oluşturmak, bu durumla ilgili evrensel doğruları ve gözlemleri kullanarak sonuçlar üretmek amacıyla kullanılır.

Bilimsel karar yöntemlerinden biri olan Bayesçi yaklaşım, olasılıklı (kesin olmayan) bir bilginin incelemesine objektif bir bakış açısını esas alır. Bu yaklaşım bilimsel gerçekten çok, bilginin aşamalarına odaklanır. Bu yaklaşımın tam olarak ifadesinin bulunması 1930' ları kadar gecikmiştir. Bu teorem genetik danışmanlar tarafından, pedigrî bilgileri ile klinik verilerin birleştirilmesinde ve taşıyıcılık risklerinin hesaplanmasında yaygın olarak kullanılmıştır. Bununla birlikte, bilgisayar destekli tanısal yaklaşımda kullanımı giderek artmaktadır. Bayes hesapları çalışma düzeninden bağımsız olduğundan, klinik analizde diğer bazı kenar analizlerine göre daha esnektir ve geleceğe ait olasılıkları hesaplayabilme olanağına sahiptir.

Bu kuramın dayanağı, koşullu olasılıklarla ilgili olan Bayes Kuralı' dır.

2.1 Bayes Kuralı

Bayesçi yaklaşımın dayanak noktası, X ve Y olaylarının koşullu olasılıklarına ilişkin bir teoremdir. Bu teoreme göre;

$$P(X, Y) := P(X, Y) \cdot P(Y)$$

$$P(X, Y) := P(X, Y) \cdot P(X) \quad (2.1)$$

Yukarıdaki denklem yeniden düzenlendiğinde Bayes Kuralı ortaya çıkar.

$$P(X, Y) = P(Y | X) * P(X) / P(Y) \quad (2.2)$$

$P(X|Y)$: X' in Y kanıtından sonraki olasılığı (Sonsal)

$P(X)$: X' in Y kanıtından önceki olasılığı (Önsel)

$P(Y|X)$: Y kanıtının X olayının gerçekleşmesi için oluşma olasılığı (Olabilirlik)

Bu kurala göre X' in olasılık dağılımı, Y hakkında elde edilen bilgiler ışığında güncellenebilir.

X ve Y' nin koşullu olasılıkları, çeşitli araştırmalarda verilerin değerlendirme ve birleştirme biçimi ile Bayesçi uygulama alanlarını çoğaltmaktadır. Özellikle banka ve genetik gibi

sektörlerde, yakın zamana kadar yeterli derecede veri seti bulunmamaktadır. Mevcut verilerle yapılan analizlerin, geçmişteki verilerle desteklenerek bir takım sonuçlara varılması elde edilen sonuçların güvenilirliği açısından daha uygun olmaktadır. Bu konudaki veri eksikliği sorununa, Bayesçi yöntemler, nitel ve nicel bilgileri, istatistiksel temellere dayandırarak birleştirebildiği için, uygun bir çözüm sunabilmektedir.

Klasik istatistikte, bir modelin parametrelerinin sabit olduğu fakat bilinmediği varsayılır. Ancak Bayesçi yaklaşımda, modelin parametrelerinin de rassal değişken olduğu varsayılır. X ' in gözlemlenmiş verileri, θ ' nın ise modelin parametrelerini ifade ettiği varsayılınsın. Bayes teoremi, X verilerinin gözlemlenmesinden sonra, θ ' nın X ' e bağlı koşullu olasılık dağılımının oluşturulmasında kullanılır. Bayes teoremi kullanılarak söz konusu dağılıma şu şekilde ulaşılır.

$$P(\theta | X) = [P(\theta)P(X | \theta)] / P(X) \quad (2.3)$$

$$P(\theta | X) = [P(\theta)P(X | \theta)] / \int [P(\theta)P(X | \theta)d\theta] \quad (2.4)$$

(2.4)' deki olasılık standartlaştırılmış olabilirlik olarak adlandırılır. (2.3) ve (2.4) denklemlerindeki $P(X)$, bir ölçeklendirme sabiti olarak görev yaptığından Bayes teoremi yaygın olarak şu şekilde ifade edilmektedir.

$$P(\theta | X) \propto P(\theta) * P(X | \theta) \quad (2.5)$$

Diğer bir ifadeyle,

Son dağılım $\propto$ ön dağılım* olabilirlik fonksiyonu

Risk yönetimi açısından bakıldığında, Bayesçi tahmin objektif veriler ile subjektif verileri bir araya getiren bir parametre tahmin yöntemidir. Bu terimler;

Objektif veriler: İncelemesi yapılan birimin iç verileri olarak tanımlanır.

Subjektif veriler: Bir dış veri tabanından elde edilen (tam olarak diğer verilere uygun olmayan) verilerdir. Uzman görüşleri ile oluşturulan risk skorları veya yapılan senaryo analizleri olabilir.

Ön olasılık dağılımı (önsel=prior): Verinin gözlemlenmesinden önce, parametre değerlerine (θ) ilişkin var olan kanılar ile oluşturulan olasılık dağılımıdır. Diğer bir ifadeyle, θ hakkındaki bilgi birikimini (ya da tahmini) yansıtır. Subjektif veriler, ön olasılık dağılımının oluşturulmasında kullanılır.

Olabilirlik fonksiyonu: Parametrelerin veri olarak değerlendirilmesiyle örneklemin olabilirliğini ifade eder. Objektif veriler kullanılarak olabilirlik fonksiyonu oluşturulur.

Son dağılım (sonsal=posterior): Ön olasılık dağılımı ve olabilirlik fonksiyonunun çarpımı ile ifade edilir. Yani ön kanıları ve örneklemden elde edilen bilgiyi, bir araya getiren bir dağılımdır. Bayes teoremi ön olasılıkların yenilenmesini sağlar.

Sonsal dağılımın ortalaması, kitle büyüklüğü için yapılacak en iyi tahmindir. Bununla birlikte sonsal dağılımın tepe değeri, örnekleme dağılımının en yüksek olabilirlik tahmincisidir

Koşullu olasılık: Bir olayın, diğer bir olayın varlığına bağlı olasılığı olarak tanımlanabilir. Örneğin bir hastalığın varlığına ya da yokluğuna ait eldeki mevcut bilgi.

Birleşik olasılık: İki olayın birlikte gözlenme olasılığı olarak tanımlanır. Ön olasılıkla koşullu olasılığın çarpımıdır.

Ön Olasılık Dağılımı → Ek (Yeni) Bilgi → Bayes Teoremi → Son Olasılık Dağılımı

Eğer gerek duyulursa, son dağılım kullanılarak parametrenin değerinin tahmin edilmesi mümkün olur. Buna “Bayesçi Tahmin” denir.

Örneklem büyüklüğü tahminleri, ardışık Bayes yaklaşımı ile daha iyi tahmin edilmektedir. Bu sonuçların tutarlılığı N' in alt ve üst değerlerinin doğru belirlenmesi ya da başka bir ifade ile N' in önsel düzgün dağılımının iyi tanımlanmasına bağlıdır.

Bunlarla birlikte Bayesçi yaklaşımda, önceki çalışmaların bilgileri (regresyon katsayıları) kolaylıkla modele ilave edilebilir. Önceki kısıtlamalar, bilinmeyenlerin olası değerlerine eklenebilir veya önsel dağılım önceki çalışmaların istatistiklerine bağlı olarak oluşturulur. Bayesçi istatistikte gözlem sayısı arttıkça, sonsal dağılımın önsel dağılımdan etkilenmesi giderek azalır ve bahis oranı (odds ratio) daha çok gözlemlenen değerlerden etkilenmeye başlar.

Bayesçi yöntemlerin klasik yöntemlerden farkı ise Çizelge 2.1’ deki gibi özetlenebilir.

Çizelge 2.1 Bayesçi yöntemler ile frekansçı yöntemlerin karşılaştırılması

<i>Konu</i>	<i>Klasik yöntemler</i>	<i>Bayesçi yöntemler</i>
Araştırmadan önce var olan bilgi	Araştırma düzeninde informal olarak kullanılır.	Önsel dağılım oluşturularak formal biçimde kullanılır.
İlgilenilen parametrenin sunumu	Sabit bir değer olarak	Bir olasılık dağılımına sahip bilinmeyen bir kantite olarak
Temel soru	Veriler belli bir parametre değerine ne ölçüde uygunluk gösterir?	Belli bir parametre değeri verilere ne ölçüde uygunluk gösterir?

2.2 Bayesçi Yöntemin Temelleri ve İstatistiksel Kavramlara İlişkin Yaklaşımı

İstatistikte, mevcut olan ilişkinin determinist mi yoksa olasılıklı mı olduğu kabul edilen yaklaşıma göre değişmektedir.

Bu düşünsel çeşitlilikler zaman içerisinde belirginleşerek iki yaklaşımda kutuplaşmıştır. Klasik yaklaşım tümdengelim yöntemi ile paralellikler gösterirken, Bayesçi yaklaşım, tümevarım yöntemiyle paralellik gösterir. Ayrıca her zaman bu ayrım pek belirgin değildir. Klasik yaklaşım nedensellik ilkesinin deterministik düşüncesine yaklaşırken, Bayesçi yaklaşım olasılıklı yorumuna yakındır. Önemli bazı istatistiksel konu ve kavramların Bayesçi yaklaşımla ifadesi aşağıdaki gibidir:

Belirsizliğin değerlendirilmesi : Bayesçi yaklaşımı benimseyenler, belirsizlikle ifade edilen bir teoriyi tamamen kabul etmenin veya körü körüne reddetmenin bilgi oluşturma sürecine pek bir katkı sağlamayacağı, aksine önemli sayılabilecek yanılsamalara yol açacağı şeklinde eleştiriler getirebilirler. Ancak olasılığın hesaplanması ile bir derece olsun bu yanılgılardan sıyrılabilineceği düşüncesine sahiptirler.

Olasılık : Olasılığın Mantıksal Teori, Klasik Teori, Frekansçı Teori, Sübjektif Teori, gibi teorilerle ortaya konan birbirinden farklı tanımları mevcuttur. Bayesçi yaklaşım bunlardan Sübjektif tanımı kabul etmektedir. Bu yaklaşımı esas alarak gelişen Bayesçi istatistikte bir olayın olasılığı, o olaya ilişkin inanç derecesi (ön bilgi) ile denemeden elde edilen sonuçların (verinin) birleştirilmiş halidir. Bir araya getirme işlemi, Bayes teoremine, dolayısıyla koşullu olasılığa dayanmaktadır. Bayesçi istatistikte olasılık “tümevarım olasılığı”dır. Amaç, en

kesine en yakın sonuca ulaşmaktır.

Çizelge 1.2 Bayesçi yaklaşım ve klasik yaklaşıma göre işleyen süreçler

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

“t” bir teori ve “g” onunla ilgili gözlemler ise farklı olasılık yaklaşımlarına göre süreç Çizelge 1.2’ deki gibi özetlenir;

Parametre : Parametre Bayesçi yaklaşımda olasılık dağılımı olan bir rastlantı değişkeni ile eş değerdedir. Dolayısıyla parametrenin tahmincisi için de bir ön olasılık dağılımı belirlenir. Mevcut veri ile birleştirilerek, parametre tahmincisinin son olasılık dağılımı elde edilir. Özetle, Bayesçi yaklaşımda parametre ile ilgili tüm çıkarsama işlemleri, son dağılıma dayanarak yapılır. Klasik yaklaşımda ise parametre, bilinmeyen bir sabit olarak görülür. Parametre tahmini sadece mevcut veriye dayandırılarak hesaplanır.

Nokta Tahmini : İstatistiğin tahmin problemlerinden biri de nokta tahminidir. Bayesçi yaklaşımda ilgilenilen parametrenin nokta tahmini genellikle sonsal dağılımın ortalamasıdır (posterior mean). Karar teorisi açısından bakıldığında, zarar fonksiyonunun beklenen değeri, optimum tahmini verir: y gözlemler, θ parametre ve $\hat{\theta}$ bu parametrenin gözlemlerden tahmini ($\hat{\theta} = \hat{\theta}(y)$) iken, son olasılık dağılımı $p(\theta | y)$, kayıp fonksiyonu $L = L(\hat{\theta}, \theta)$, dır. Optimum tahmin ise $EL(\hat{\theta}, \theta) = \int_{-\infty}^{+\infty} L(\hat{\theta}, \theta) p(\theta | y) d\theta$ ’ dır.

Nokta tahmini başlığı altında istatistikte tahminin değerlendirilmesinde kullanılan, Bayes ve Klasik yaklaşım için birbirinden farklı önem derecelerine sahip bazı kıstaslara da değinilmiştir;

Yeterlilik : İlgilenilen parametreye ilişkin, sahip olunabilecek tüm bilgiyi içerisinde bulunduran istatistik, yeterli istatistiktir. Bayesçi ve Klasik yaklaşım bu tanımda hemfikirdir.

Fakat yeterlilik, Klasik nokta-tahmin teorisinde sadece bir varsayımken, Bayesçi hesaplamada kanıtı mevcuttur. Aslında Klasik yaklaşım görüşüyle pek bağdaşmaz. Çünkü yeterliliğin özünde, parametrenin koşullu olasılığı vardır. Bir parametrenin olasılığının olması da Klasik görüşe taban tabana zıttır.

Parametre θ , örnek değerleri $X = X_1, X_2, \dots, X_n$ iken t de, $t = t(X)$, X 'in fonksiyonu bir yeterli istatistiği gösteriyorken;

Yeterliliğin anlamını açıklayan matematiksel ilişkiler;

1-) $P(X | t \& \theta) = P(X | t)$ diğer bir ifadeyle,

$$\left(f(X_1, X_2, \dots, X_n | t) = \frac{f(X_1, X_2, \dots, X_n, t)}{g(t)} = \frac{f(X_1, X_2, \dots, X_n)}{g(t)} \right) \text{ ve}$$

2-) $P(\theta | X) = P(\theta | t)$ 'dir.

1-) $P(X | t \& \theta) = P(X | t)$; olabilirlik (likelihood) ilkesi açısından yazılan ifadeye göre, parametre (θ) ve yeterli istatistik (t) veri iken, X 'leri gözlemenin olasılığı; sadece yeterli istatistik veri iken X 'leri gözlemenin olasılığına eşittir. Dolayısıyla X 'in koşullu olasılık dağılımı, parametreden bağımsızdır. Parametre olmadan da “yeterli” istatistik aynı ilişkiyi sağlar.

2-) $P(\theta | X) = P(\theta | t)$; bu ifade ise başka bir açıdan (normal koşullu olasılık ile) bakarak aynı kavramı açıklar. Gözlemler (X) veri iken parametrenin (θ) olasılığı; yeterli istatistik (t) veri iken elde edilen parametre olasılık değerine eşittir. Yani gözlemler yerine “yeterli” istatistik de parametrenin olasılığını açıklarsa, değer aynı olacaktır.

Yukarıdaki eşitliklerden hareketle, Bayes Teoremine göre yeterliliğin ispatı θ , t ve X 'in her değeri için,

$$P(X | \theta \& t) = \frac{P(\theta | X \& t)P(X | t)}{P(\theta | t)}$$

olur. Burada $P(\theta | X \& t)$ olasılığı, $P(\theta | X)$ 'e eşittir. Çünkü t , X 'den hesaplanmıştır. Teoreme yerine konursa,

$$P(X | \theta \& t) = \frac{P(\theta | X)P(X | t)}{P(\theta | t)} \quad (2.6)$$

olur. Böylelikle yeterliliğin ispatı yapılmış olur. Yeterlilik, Bayesçi yaklaşım için oldukça önemli bir ölçüttür. Ön dağılımın elde edilmesinde bu özellikten yararlanılır.

Yansızlık : Tahmincinin beklenen değerinin, parametrenin gerçek değerine eşit olması demektir: θ parametre ve $\hat{\theta}$ bu parametrenin tahmincisi iken, tahminci yansız ise $E(\hat{\theta}) = \theta$ 'dır.

Klasik yaklaşımda iyi bir tahminci yansızdır. Öte yandan Bayesçi tahminciye (son dağılımın ortalamasına) bakıldığında yansız olmadığı görülüyor. Bu ispat şöyle gösterilebilir;

Y gözlemlerden elde edilen tahmin, $f(y | \theta)$ bu gözlemlerin olasılık yoğunluk fonksiyonu, θ ise bir parametredir. $E_{\theta}(Y) = \theta$ Klasik yaklaşımın yansızlık ifadesidir. Bayesçi nokta tahmini (son dağılımın ortalaması) aşağıdaki gibidir;

$$E(\theta | y) = \int \theta P(\theta | y) d\theta \quad (2.7)$$

O halde yukarıda belirtilen son dağılımın ortalamasının yansız olması için, şu eşitliği sağlaması gerekmektedir;

$$E_{\theta}[E(\theta | Y)] = \int \left[\int \theta P(\theta | y) d\theta \right] f(y | \theta) dy = \theta. \text{ (Burada tahminci olarak } Y \text{ yerine, son dağılımın ortalaması konulmuştur.)}$$

Y tahmincisinin varyansının tanımından yola çıkılarak, Bayesçi tahmincinin, beklenen değerinin parametre değerine eşit olup olmadığına bakılmaktadır.

Aşağıdaki ifade, θ ' ya göre koşullu hale getirilirse;

$$E[(Y - \theta)^2] = E[Y^2 - 2Y\theta + \theta^2]$$

$$= E[E((Y^2 - 2Y\theta + \theta^2) | \theta)]$$

Beklenen değeri alınır.

$$= E[E(Y^2 | \theta) - 2\theta^2 + \theta^2]$$

Burada $(E(Y | \theta) = E_{\theta}Y = \theta)$ varsayılır.

$$= E[E(Y^2 | \theta) - \theta^2] = E(Y^2) - E(\theta^2).$$

Benzer biçimde Y' ye göre koşullu olursa;

$$E[(Y - \theta)^2] = E[E((Y^2 - 2Y\theta + \theta^2) | Y)]$$

$$= E(Y^2 - 2Y^2 + E(\theta^2 | Y))$$

Burada $E(\theta | Y) = Y$ varsayılır.

$$= E(\theta^2) - E(Y^2) \quad (2.8)$$

Her iki hesaplamının yansızlık değerlendirmesi yapılırsa;

Eğer $E(\theta^2) = E(Y^2)$ eşitliği sağlanıyorsa, başka bir ifade ile $[E(Y - \theta)^2] = 0$ ise, tahminci Y, parametre θ ' ya eşit demektir.

Etkinlik : Klasik yaklaşım içinde anlamlı sayılabilecek değerlendirme ölçütlerinden biri olan etkinlik ile, tahmincinin mümkün olan en küçük varyansa sahip olması şartı aranmaktadır. Klasik yaklaşımda yansız tahmincilerden biri tercih edilecekse, varyansı küçük olan, başka bir ifadeyle olasılık dağılımı daha dar aralıkta yayılan, tahminci seçilir. Bilindiği gibi, ortalama hata karesi de etkinliği ölçmektedir.

Kesinlik : Kesinlik (precision) varyans ile ters orantılıdır. Bir değerlendirme kısıtı olarak kesinlik, istatistikte tahmin kriterleri olarak kabul edilen diğer kavramlarla karşılaştırılarak daha iyi anlatılabilir. Kesinlik, hatanın olasılıklı kısmını içerir ve sistematik hatanın yapılmadığı varsayılırsa, gerçek değerden sapmayı ifade eder. Tümüyle Bayesçi çatıda yapılan bir analizde varyans yerine kesinlik alınarak, gerçek değer olasılığı hesaplanır.

Güven Aralığı : İstatistikteki tahmin probleminin temel konulardan biri de, aralık tahminidir. Güven aralıklarına ilişkin bilgi, başka bir ifadeyle parametrenin belirli değerler arasında yer aldığı (veya Bayesçi yaklaşıma göre yer alma olasılığının) bilgisi anakütleyi tanımlama ve tanımlamada yardımcıdır. Bu noktada belirtmek gerekir ki Bayesçi yaklaşım, Klasik yaklaşımın “güven aralığı” tanımlaması yerine, “güvenilir aralık” (credible interval), “Bayesçi aralık” (Bayesian interval) veya “En Yüksek Son Yoğunluk Bölgesi veya Aralığı” (Highest Posterior Density Region or Interval) tanımlamalarını kullanmaktadır.

Aslında tanımlamada farklı olduğu gibi yorumlamada da farklıdır; Bayesçi yaklaşımda, son dağılımın ortalaması için, örneğin 0,95 güven düzeyinde bir güven aralığından söz ediliyorsa, bu aralığın, son dağılımın ortalamasını içermesi olasılığı - son dağılımın ortalamasını içereceğine olan inanç derecesi- %95'tir.

Hipotez Testi : İstatistikte anakütleyi anlamaya ve çözmeye çalışırken, parametrelere ilişkin iddiaların yer aldığı bazı önermelerden yararlanır. Hipotez olarak adlandırılan bu önermelerde ortaya atılan ifadeler, 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 anakütlenin belirsizliği kısmi

olarak aydınlatılmış olunur.

Bayes, hipotezlere olasılık tayin eder. Bunun nedeni, Bayesçi yaklaşımın olasılık teorisine daha geniş bir mantıksal açıdan bakmasıdır. Klasik Aristo mantığına göre bu değer 0 veya 1 olması değil, 0-1 aralığında değerler alması söz konusudur. Yani bir hipotez kabul ya da reddedilmez, onun sahip olduğuna inanılan olasılığı belirlenir. Ancak bu noktada belirtmek gerekir ki, Bayesçi 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.

2.3 Bayesçi Tahmin Problemi

Bilinmeyen bir θ ve rasgele seçilen $X_1, \dots, X_n$ parametrelerinin oluşturduğu $f(x|\theta)$ dağılımı olsun. θ değeri, Ω aralığındaki gerçek değerlerin tahminidir. Ω aralığı sınırlı ya da sınırsız olabilir. Sonuç olarak, θ değeri örnekten seçilen değerler sonucunda oluşur.

$X_1, \dots, X_n$ gibi rastlantı değişkenlerinin oluşturduğu olasılık kümesinin tahmini değerleri, $\delta(X_1, \dots, X_n)$ gerçek değerli fonksiyonuna bağlı olarak θ 'nın tahminci olması sağlanır. Diğer bir ifadeyle, θ değeri, $\delta(X_1, \dots, X_n)$ kümesinin oluşturduğu Ω aralığında yer almalıdır.

Tahminci ve tahmin arasındaki ilişki özetlenecek olunursa, tahminci $(\delta(X_1, \dots, X_n))$, $(X_1, \dots, X_n)$ 'nin birikimli olasılıklarının ve rastlantı değişkenlerinin fonksiyonlarına bağlıdır. Tahmin ise özel $x_1, \dots, x_n$ değerleriyle oluşturulmuş $\delta(x_1, \dots, x_n)$ dağılımına bağlı olarak değişir. $X = X_1, \dots, X_n$ ve $x = x_1, \dots, x_n$ vektörleri sıklıkla kullanılacaktır. $\delta(X)$ notasyonu yerine ise δ ifadesi yer alacaktır.

2.3.1 Kayıp Fonksiyonu

Gerçek θ değeri, δ 'nin iyi bir tahminci olmasını sağlar. İyi bir tahminci “ $\delta - (\theta)$ ” değerinin sıfır olduğu durumlar için söz konusu olur. Her olasılık değeri için “ $\theta \in \Omega$ ” ve her tahmin değeri içinde “ $a \in \Omega$ ” olmalıdır. θ değeri ve onun tahmincisi olan a değeriyle arasındaki uzaklık “ $L(\theta, a)$ ” fonksiyonuyla ifade edilir. Bu fonksiyonun beklenen değeri ise,

$$E[L(\theta, a)] = \int L(\theta, a) \xi(\theta) d\theta \quad (2.9)$$

ifadesiyle tanımlanır. Bu ifade de $\xi(\theta)$ önsel dağılım olarak tanımlanır. Bu işlemler yapılarak minimum değer hesaplanmaya çalışılır. Bir tahmin değeri oluşturmaya çalışan L

fonksiyonu minimize edilmiş bir kayıp fonksiyonu olarak adlandırılır.

Bayes Tahmincisinin Tanımı

$x \in X$ (X bir vektör) iken, θ 'nın Ω aralığında son dağılımı $\xi(\theta, x)$ ile ifade edilir. Bu durumda istatistiksel “a” tahmincisi kullanılarak aşağıdaki gibi bir beklenen kayıp fonksiyonu oluşturulur:

$$E[L(\theta, a) | x] = \int L(\theta, a) \xi(\theta | x) d\theta \quad (2.10)$$

Bu fonksiyon nedeniyle istatistikçiler minimum olan bir tahminciyi tercih etmelidirler.

X vektörünün elemanı olan x olasılık değeri için, beklenen minimum kayıp değeri olacak a tahmincisi için $\delta^*(x)$ fonksiyonu tanımlansın. Aynı zaman da bu fonksiyon Θ 'nın özel bir tahmincisi olsun. Bu tahminci θ 'ın *Bayes tahmincisi* olarak adlandırılır. Bayes tahmincisi için aşağıdaki beklenen değer hesaplanır:

$$E[L(\theta, \delta^*(x)) | x] = \min E[L(\theta, a) | x] \quad (a \in \Omega \text{ için}) \quad (2.11)$$

Sonuç olarak, $X = (X_1, \dots, X_n)$ ‘den çekilen rasgele örneklerin oluşturduğu tahmin problemi, Ω aralığındaki bilinmeyen değerlerin θ parametresinin oluşturduğu dağılım ile çözümlenebilir. Verilen her kayıp fonksiyonu $L(\theta, a)$ ve her önsel dağılım $\xi(\theta)$ için θ 'nın Bayes tahmincisi, X 'deki her x olasılık değeri için $\delta^*(x)$ tahmincisiyle hesaplanır. Buradan da anlaşılacağı gibi Bayes tahmincisi, kayıp fonksiyonu ve θ 'nın oluşturduğu önsel dağılıma bağlı olarak oluşturulur.

Hata Karelerinin Kayıp Fonksiyonu: Hata kareleri kayıp fonksiyonu, tahmin problemindeki kayıp fonksiyonu aracılığıyla bulunur. Bu fonksiyon aşağıdaki gibi gösterilebilir:

$$L(\theta, a) = (\theta - a)^2 \quad (2.12)$$

Hata karelerinin kayıp fonksiyonu kullanıldığı zaman, x değeri için Bayes tahmini $\delta^*(x)$, minimum a değerli $E[(\theta - a)^2 | x]$ beklenen değeriyle bulunur.

$(\theta - a)^2$ beklenen değeri θ 'nın son dağılımına uygun hesaplandığı zaman, bu beklenen değeri θ 'nın sonsal dağılımının ortalamasına $[\delta^*(X) = E(\theta | x)]$ uygun seçilecek olan a olasılığının, minimum olmasını sağlayacaktır. Buradan da anlaşıldığı üzere hata karelerinin kayıp fonksiyonu kullanılarak Bayes tahmincisi $\delta^*(X) = E(\theta | X)$ 'ne eşit olacaktır.

2.3.2 Önsel ve Sonsal Dağılımlar

2.3.2.1 Önsel Dağılımın Belirlenmesi

Bayesçi bir tahmin gerçekleştirebilmek için öncelikle, Bayesçi tahminin gereklerinin nasıl belirleneceğine karar verilmelidir. Buradaki asıl sorun, kısaca olasılık fonksiyonu $f(x|\theta)$ olarak verilen bir anakütleden çekilen gözlemleri temel alarak, θ 'nin asıl değerinin parametreler uzayı Ω 'nin neresinde var olabileceğini saptamaya çalışmaktır.

Birçok problemde, $f(x|\theta)$ ile ilgili herhangi bir gözlem yapılmadan önce, deney yapan bir kişi, önceki bilgilerini gözden geçirerek, Ω uzayında bir olasılık dağılımı oluşturmak suretiyle, θ 'in değerinin nerede bulunabileceğini tahmin edebilir. Diğer bir deyişle, herhangi bir deneysel veri elde edilmeden ya da gözlemlenmeden önce, deneyi yapan kişinin geçmiş tecrübeleri ve bilgi birikimi onu, θ 'nın Ω 'nin diğer bölgelerine oranla, belli başlı bazı bölgelerinde bulunduğu inanmaya sevk edecektir. θ 'nin Ω uzayında belirli bölgelerde bulunma olasılıklarından hareketle bir ön dağılım düşünülebilir. Bu dağılıma θ 'nin önsel dağılımı denir.

İstatistikte, önsel dağılım kavramı çok fazla tartışma yaratmaktadır. Bir araştırmacının θ 'nin gerçek değerinin nerede olabileceğine dair öznel inançları bulunması bağlamında, bu dağılımın bir öznel olasılık dağılımı olduğu savı bazı istatistikçiler tarafından ileri sürülür. Öte yandan, bir önsel dağılımın istatistik alanında kullanılan diğer herhangi bir olasılık dağılımından farksız olduğuna ve olasılık teorisinin tüm kurallarının bir önsel dağılım için de geçerli olacağına inanılır.

Diğer bazı istatistikçiler birçok problemde θ 'nin olasılık dağılımından söz etmenin doğru olmadığı yorumunda bulunur. Çünkü θ 'nin gerçek değeri daha çok deney yapan kişi tarafından belirlenen sabit bir sayı olarak düşünülür. Bu istatistikçiler; ancak θ 'nin geçmişteki olası değerleri ile ilgili yoğun bir ön bilgi sahibi olduğunda, θ parametresinde bir önsel dağılımın kullanılabileceği düşünmektedirler. Ancak o zaman, iki farklı istatistik değerinin kullanılacak uygun bir önsel dağılım üzerinde uzlaşması mümkün olacaktır.

Ayrıca Kesikli ve Sürekli Önsel Dağılımlara ait bazı problemlerde, θ parametresi sadece sonlu sayıda değişik değer alabilir. Diğer problemlerde, θ parametresi reel sayılar kümesi üzerinde ya da reel sayılar kümesinin bir aralığında sonsuz sayıda değer alabilir. Bu durumda bu dağılımın olasılık yoğunluk fonksiyonuna, θ 'nin önsel olasılık yoğunluk fonksiyonu denir.

2.3.2.2 Sonsal Dağılımın Oluşturulması

$X_1, \dots, X_n$ 'nin; olasılık fonksiyonunun $f(x|\theta)$ olduğu bir anakütleden; rassal bir örnek oluşturduğu varsayalım. Ayrıca parametre θ 'nin değerinin bilinmediği ve θ 'nin önsel olasılık fonksiyonunun da $\xi(\theta)$ olduğu düşünölsün. Bu şartlara bağılı $X_1 \dots X_n$ 'in olasılık fonksiyonu aşğıdaki gibi yazılabilir:

$$g_n(x) = \int_{\Omega} f_n(x|\theta) \xi(\theta) d(\theta) \quad (2.13)$$

Benzer bir şekilde $X_1=x_1$, $X_n=x_n$ verildiğinde, θ 'nin koşullu olasılık yoğunluk fonksiyonu, $X_1 \dots X_n$ 'nin bileşik olasılık yoğunluk fonksiyonuna ve θ 'nin $x_1 \dots x_n$ marjinal bileşik olasılık yoğunluk fonksiyonuna bölünerek elde edilir.

$$\xi(\theta|x) = \frac{f_n(x|\theta) \xi(\theta)}{g_n(x)} \quad (2.14)$$

Buradaki koşullu olasılık yoğunluk fonksiyonuna sonsal dağılım denir. Çünkü bu, $X_1 \dots X_n$ 'in değerleri gözlemlendikten sonra elde edilen θ 'nin dağılımıdır. Sonsal olasılık yoğunluk fonksiyonu $\xi(\theta|x)$; $X_1=x_1$, $X_n=x_n$ gözlemlendikten sonra θ 'nin bu görelî olabîrlîğini temsil eder.

2.3.2.3 Eşlenik Dağılımlar

Bazı durumlarda θ parametresinin önsel dağılımının bilinmesi, sonsal dağılımının tahmin edilmesini sağlamaktadır. Örneğîn, θ 'nin önsel dağılımı beta ise sonsal dağılımı da bir diğêr beta dağılımına uymaktadır. Bu durumda önsel ve sonsal dağılımlar eşlenik (conjugate) dağılımlar olarak nitelendirilebilir. Bunun için aşğıdaki dağılımlar incelenmiştir.

2.3.2.3.a Bernoulli Dağılımının Parametresinin Tahmini:

θ parametresinin değerinin bilinmediğı bir Bernoulli dağılımından rassal bir örnek alınsın. Eğêr θ 'nin önsel dağılımı bir beta dağılımıysa, o zaman herhangi bir rassal örnek yardımı ile, θ 'nin sonsal dağılımı da yine bir beta dağılımı olacaktır.

$X_1 \dots X_n$ 'in, θ parametresinin değerinin bilinmediğı bir Bernoulli dağılımından rassal bir örnek olsun ($0 < \theta < 1$). Ayrıca θ 'un önsel dağılımının, verilen α ve β ($\alpha > 0$ ve $\beta > 0$)

parametreleriyle, bir beta dağılımı olduğu varsayalım. $y = \sum_{i=1}^n x_i$ olsun.

Bu durumda önsel olasılık yoğunluk fonksiyonu için aşğıdaki eşitlik söz konusu olabilir.

$$\xi(\theta) \propto \theta^{\alpha-1} (1-\theta)^{\beta-1}, 0 < \theta < 1$$

Sonsal olasılık yoğunluk fonksiyonu $f_n(x|\theta) \propto \xi(\theta)\xi(\theta|x)$ 'ya oranlanabildiğinden, şu sonuç çıkmaktadır:

$$\xi(\theta|x) \propto \theta^{\alpha+y-1} (1-\theta)^{\beta+n-y-1} \quad 0 < \theta < 1$$

Sabit bir faktör dışında, bu bağıntının sağ tarafı, $\alpha+y$ ve $\beta+n-y$ parametrelerine sahip bir beta dağılımının olasılık yoğunluk fonksiyonuna eşittir. Ve yine bu nedenle bu dağılımın Bayes tahmincisi:

$$\delta^1(X) := \frac{\alpha + \sum_{i=1}^n X_i}{\alpha + \beta + n}$$

ifadesiyle bulunur.

2.3.2.3.b Bir Poisson Dağılımından Örnekleme:

$X_1 \dots X_n$ Poisson dağılımında θ 'nin ortalama değerinin bilinmediği rassal bir örnek oluştursun ($\theta > 0$). θ 'nin önsel dağılımı α ve β 'nin ($\alpha > 0$ ve $\beta > 0$) belli parametrelerle gama dağılımı olduğu varsayılın. Yine θ 'nin sonsal dağılımı, $X_i = x_i$ ($i=1, \dots, n$) verildiğinde, parametreler $\alpha + \sum_{i=1}^n x_i$ ve $\beta + n$ ile gama dağılımıdır. Aynı şekilde, $y = \sum_{i=1}^n x_i$ olsun. Bu

durumda, $f_n(x|\theta) \propto e^{-n\theta} \theta^y$ bağıntısı oluşur. Bu bağıntıda x 'i içeren ancak θ 'ya bağlı olmayan bir faktör sağ taraftan atılır. Ayrıca θ 'nin önsel olasılık yoğunluk fonksiyonu

$$\xi(\theta) \propto \theta^{\alpha-1} e^{-\beta\theta} \quad \theta > 0$$

$$\xi(\theta|x) \propto \theta^{\alpha+y-1} e^{-(\beta+n)\theta}$$

biçiminde ifade edilebilir. Olasılık dağılım fonksiyonu ($\theta|x$) sonsal, $f_n(x|\theta) \propto \xi(\theta)$ 'ya oransal olduğu için $\xi(\theta|x) \propto \theta^{\alpha+y-1} e^{-(\beta+n)\theta}$ uyar. Bu bağıntının sağ tarafı, $\alpha + y$ ve $\beta + n$ parametreleriyle sabit bir faktör dışında, bir gama dağılımının olasılık yoğunluk fonksiyonu olarak varsayılabilir.

1.3.2.3.c Bir Normal Dağılımın Parametrelerinin Tahmini

$X_1 \dots X_n$, ortalama θ 'nin değerinin bilinmediği ($-\infty < \theta < \infty$) ve σ^2 varyansının değerinin bilindiği ($\sigma^2 > 0$) bir normal dağılımdan rasgele bir örnek oluşturduğu varsayılın. θ 'nin önsel dağılımının, ortalama μ 'nin ve σ^2 varyansının belli değerleriyle bir normal dağılım olduğu düşünölsün. θ 'nin sonsal dağılımı,

$$\mu_1 = \frac{\sigma^2 \mu + n \bar{x}_n}{\sigma^2 + n}$$

parametrelili normal bir dağılıma uymaktadır. Olasılık fonksiyonu $f_n(x|\theta)$ ise

$$f_n(x|\theta) \propto \exp \left[-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \theta)^2 \right]$$

formölasyonu ile ifade edilir. Burada sabit faktör sağ taraftan atıldıktan sonra ifadeyi dönöştürmek için şu eşitlik kullanılır:

$$\sum_{i=1}^n (x_i - \theta)^2 = n(\theta - \bar{x}_n)^2 + \sum_{i=1}^n (x_i - \bar{x}_n)^2$$

ve böylece $x_1 \dots x_n$ içeren ancak $f_n(x|\theta)$ ' ya bağlı olmayan bir faktör olarak ortaya çıkar.

Sonuç olarak $f_n(x|\theta)$ aşağıdaki şekilde yeniden ifade edilebilir:

$$f_n(x|\theta) \propto \exp\left[-\frac{1}{2\sigma^2}(\theta - \bar{x}_n)^2\right]$$

Önsel olasılık yoğunluk fonksiyonu $\xi(\theta)$ şekline sahip olduğu için, onu takiben sonsal olasılık yoğunluk fonksiyonu aşağıdaki ifadeyle gösterilirken:

$$\xi(\theta) \propto \exp\left[-\frac{1}{2\nu^2}(\theta - \mu)^2\right]$$

ve daha önceden belirtildiği gibi

$$\xi(\theta|x) \propto \exp\left\{-\frac{1}{2}\left[\frac{n}{\sigma^2}(\theta - \bar{x}_n)^2 + \frac{1}{\nu^2}(\theta - \mu)^2\right]\right\}$$

ifadesi elde edilir. Bu denklemin sağ tarafındaki son terim θ içermediği için, oransallık faktörü içersinde düşünülebilir ve

$$\frac{n}{\sigma^2}(\theta - \bar{x}_n)^2 + \frac{1}{\nu^2}(\theta - \mu)^2 = \frac{1}{\nu_1^2}(\theta - \mu_1)^2 + \frac{n}{\sigma^2 + n\nu^2}(\bar{x}_n - \mu)^2$$

bağıntısı elde edilir. Bu bağıntının sağ tarafı, sabit faktör dışında, ortalaması μ ve varyansı σ^2 olan normal bir olasılık yoğunluk fonksiyonu olarak değerlendirilebilir. Bu nedenle, θ ' nın sonsal dağılımı teoremden belirtildiği gibidir. θ ' nın sonsal dağılımının ortalaması, aşağıdaki gibi yeniden yazılabilir:

$$\xi(\theta|x) \propto \exp\left[-\frac{1}{2\nu^2}(\theta - \mu_1)^2\right]$$

ν ve σ^2 ' nin değişmez değerleri için, daha geniş örnek büyüklüğü n , daha büyük x 'e verilen değişim ağırlığı olacaktır.

ν ve n ' nin değişmez değerleri için, örnekteki her gözlem σ^2 ' nin varyansından ne kadar geniş olursa, x 'e verilen değişim ağırlığı o kadar küçük olacaktır.

σ^2 ve n 'nin değişmez değerleri için, önsel dağılımın σ^2 varyansı ne kadar geniş olursa, x 'e verilen değişim ağırlığı o kadar geniş olacaktır.

Sonuç olarak μ_1 için,

$$\mu_1 = \frac{\sigma^2}{\sigma^2 + n\sigma^2} \mu + \frac{n\sigma^2}{\sigma^2 + n\sigma^2} \bar{x}_n$$

eşitliği oluşacaktır.

2.3.2.3.d Üstel Dağılımdan Örneklem

$X_1, \dots, X_n$, θ parametresinin ($\theta > 0$), bir üstel dağılıma uyduğu varsayalım. (θ 'nın önsel dağılımının α ve β 'nin ($\alpha > 0$ ve $\beta > 0$) belli olduğu parametrelerle, bir gama dağılımı olduğu varsayalım.) θ 'nın sonsal dağılımı, $X_i = x_i$ ($i=1, \dots, n$) verildiğinde, $\alpha+n$ ve $\beta+y$ parametreleriyle bir gama dağılımıdır.

$$y = \sum_{i=1}^n x_i$$

olsun. Önsel olasılık yoğunluk fonksiyonu $f_n(x|\theta) \propto \theta^n e^{-\theta y}$, $\xi(\theta) \propto \theta^{\alpha-1} e^{-\beta\theta}$, $\theta > 0$ şeklindedir. Bu nedenle

$$\xi(\theta|x) \propto \theta^{\alpha+n-1} e^{-(\beta+y)\theta} \quad \theta > 0$$

ifadesindeki orantısallık ilişkisi elde edilir. Bu bağıntının sağ tarafı, sabit faktör dışında, $\alpha+n$ ve $\beta+y$ parametreleriyle bir gama dağılımının olasılık yoğunluk fonksiyonuna eşittir.

Yukarıda örnekleri verilen dağılımların önsel ve sonsal dağılımları bir tablo ile gösterilecek olursa;

Çizelge 2.3 Örneklem dağılımlarının ön ve son dağılımları

Örneklem dağılımı	Ön dağılım	Son dağılım
Bernoulli	Beta	Beta
Poisson	Gama	Gama
Negatif Binomial	Beta	Beta
Üstel	Gama	Gama
Normal	Normal	Normal

Özetle ifade edilecek olunursa, bir ön dağılımın olabilirlik fonksiyonu ile işleme girdikten sonra bir son dağılımın oluşturulmasıyla bayesçi tahmin oluşturulur. Bayesçi tahmin, Bayes

teoreminden yola çıkarak t mdengelim y ntemiyle eksik verilerin olumsuzluęunu ortadan kaldırarak bir yargı oluřmasını saęlar.

3. REGRESYON

Bir konu hakkında mevcut olan iki ve ya daha fazla değişken arasında ilişki olup olmadığını, varsa bu ilişkilerin ne yönde olduğunu veya bu ilişkilerin derecesinin ölçülebilmesi ihtimali, bu konu üzerinde araştırma yapma ihtiyacı doğurmuştur.

İstatistiksel anlamda iki veya daha fazla değişken arasındaki ilişki, bunların mevcut değerleri arasındaki bağıllık olarak anlaşılır. Şöyle ki X değişkeninin sahip olduğu değerlere göre, Y değişkeninin değerleri artı veya eksi yönde değişiyorsa bu iki değişken arasında bir ilişkinin varlığından söz edilebilir. Bu iki değişken arasında mevcut olabilecek bir ilişkinin varlığının nasıl gösterilebileceği akla gelen soruların başında yer alır. Bunu matematiksel bir ifade ile formülize ederek, ilişkinin gösterilmesi için ilk başlangıç yapılır. Eğer bu değişkenlerden birine ait değerler biliniyorsa diğer değişkene ait değerler bunlardan yola çıkılarak tahmin edilebilir.

Değişkenler arasındaki ilişki, bir neden-sonuç ilişkisidir. Ancak bu neden sonuç ilişkisi her zaman basit denklemlerle ifade edilemeyebilir. Öyleki bazı durumlarda hangi değişkenin neden, hangi değişkenin sonuç durumunu karşıladığı tam olarak kestirilemez.

Değişkenler arasındaki ilişkinin incelenmesi, bilim açısından hep bir öneme sahip olmuştur. Öyle ki, karşılaşılan tüm sorunlarda, iki veya daha fazla değişkenin arasındaki ilişkiye bağlıdır. Örneğin bir öğrencinin zekâ düzeyi ile başarı düzeyi, bir partiye oy vermekle kişinin eğitim düzeyi gibi.

Yukarıdaki açıklamalardan anlaşılacağı üzere, değişkenler arasındaki ilişkinin matematiksel ifadesi diğer bir deyişle fonksiyonel ifadesi, birbirlerini nasıl etkiledikleri ve bu etkinin gücü bilinmek istenmektedir. Değişkenler arasındaki ilişkinin nasıl bir matematiksel denklem ile ifade edileceği ve bu denklem sonucunda ilişkinin doğrusal ya da eğrisel olacağı sorusunun cevabını, ilişkinin fonksiyonel şekli verir. İlişkinin yönü ise aradaki bağlantının aynı yönde mi yoksa zıt yönde mi değişime neden olduğunu gösterir. İlişkinin derecesi ise bu değişkenler arasındaki bağın kuvvetini açıklar. İlişkinin fonksiyonel şekli regresyon çözümlemesini, ilişkinin derecesi ise korelasyon çözümlemesini ortaya çıkarır.

Tarihsel gelişimi incelendiğinde 19. yy' da regresyon kavramı ortaya çıkmıştır. Bu kavramın ortaya çıkmasına, kalıtım üzerinde yapılan araştırmalar vesile olmuştur. Regresyon kavramı, daha sonra tahminleme işleminde kullanılan bir teknik olarak gündeme gelmiştir. Özetle regresyon, bilinenlerden yararlanılarak bilinmeyen durumların tahmin edilmesinde kullanılan

bir tekniktir.

Regresyon çözümlemesinde “sonuç” durumunu oluşturan değişkene “bağlı değişken”, “neden” durumunu oluşturan değişkene ise “bağımsız değişken” denir. Bağımsız değişkenlerin sayısı bir veya birden fazla olabilir.

İlişkinin fonksiyonel şekli açısından regresyon, doğrusal ve doğrusal olmayan regresyon olarak ikiye ayrılır.

3.1 Doğrusal Regresyon

Regresyon çözümlemesinde, bağımsız değişken sayısı iki veya daha fazla olabilir. Modelde yer alan bağımsız değişken sayısı bir ise bu regresyona basit regresyon, iki ve daha fazla ise çoklu regresyon modeli olarak ifade edilir. Diğer taraftan değişkenler arasındaki ilişkinin şekli (doğrusal olup olmadığı) de önem taşımaktadır. Burdan da anlaşılacağı üzere değişkenler arasındaki ilişki, doğrusal ise bu regresyon çözümlemesi doğrusal regresyon modeli olarak ifade edilir.

Basit doğrusal regresyon modeli aşağıdaki gibi ifade edilebilir:

$$Y_i = \alpha + \beta X_i + u_i \quad (3.1)$$

Yukarıdaki ifade tek bir bağımsız değişken içeren stokastik bir modeldir. Bu modelde yer alan α ve β parametrelerinin değerlerini bulabilmek için X bağımsız değişkeni ve Y bağımlı değişkenine ait gözlemlere ihtiyaç vardır.

X ve Y değişkenlerinin içerisinde bulunduğu anakütleye ait bütün veriler elde edilemeyeceği için, bu değişkenleri karşılayabilecek bir örneklem oluşturulmalıdır. Böylece α ve β parametrelerinin kesin değerleri bulunamasa bile, bu parametrelerin değerlerinin birer tahmini olan “a ve b” değerleri bu örneklemde yararlanılarak bulunabilir.

Söz konusu olabilecek anakütle içerisinde birer “ α ve β ” değeri bulunurken, bu anakütleden çekilen her bir örneklem için ayrı birer “a ve b” değeri bulunmaktadır. İşte bu “a ve b” normal bölünmeye sahip olup, beklenen değeri sırasıyla “ α ve β ” dır. Eğer bu değerler bulunmak isteniyorsa, anakütle içerisinde tek bir örneklem alınarak, bu örneklem yardımıyla anakütle parametreleri tahmin edilebilmektedir.

Regresyon modeline dahil edilemeyen diğer değişkenler, modelde “u” ile ifade edilir ve hata terimi olarak adlandırılır. “u” hata terimini gözlemlemek hiçbir zaman mümkün olamaz.

Bunun için hata terimleri için belirli bazı varsayımları göz önünde bulundurmak gerekir. Y ile X değişkenleri arasındaki gerçek ilişki,

$$Y_i = \alpha + \beta X_i + u_i \quad (3.2)$$

iken gerçek regresyon doğrusu

$$E(Y_i) = \alpha + \beta X_i \quad (3.3)$$

eşitlikleri ile ifade edilir. Bununla birlikte, tahmin edilen ilişki,

$$Y_i = a + bX_i + e_i \quad (3.4)$$

denklemleri ile ifade edilir. Tahmin edilen regresyon doğrusu ise :

$$\hat{Y}_i = a + bX_i \quad (3.5)$$

Yukarıdaki eşitliklerde;

Y_i : Y değişkeninin gözlemlenen değerini,

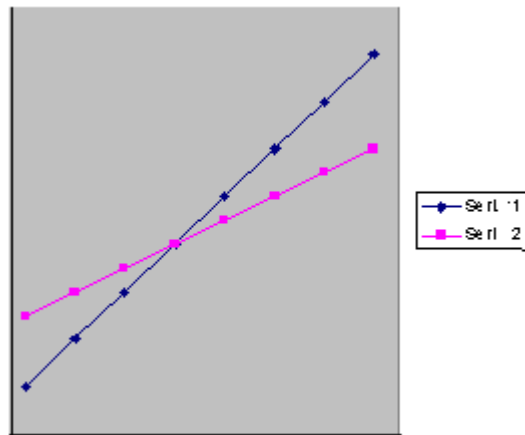
$\hat{Y}_i$: X değişkeninin belli bir değeri veri iken, Y değişkeninin tahmin edilen değeri,

a : α gerçek kesim noktasının tahminini,

b : β gerçek değerinin tahminini

e : u hata teriminin gerçek değerinin tahminini

simgeler. Gerçek ve tahmin edilen regresyon doğruları aşağıdaki doğrularla ifade edilebilir.



Şekil 3.1 Gerçek ve tahmin edilen regresyon doğruları

Seri 1 : $\hat{Y}_i = a + bX_i$

Seri 2 : $E(Y_i) = \alpha + \beta X_i$

Doğrusal regresyon modeli bazı varsayımlara dayanır. Söz konusu varsayımlar kısaca;

- Hata teriminin bölünmesi,
- Hata terimi ile serbest değişkenler arasındaki ilişki,
- Serbest değişkenler arasındaki ilişki,

Hata terimleri ile ilgili stokastik varsayımlar aşağıdaki gibi özetlenebilir:

1. Hata terimi rassal bir değişkendir.
2. Hata teriminin ortalaması sıfırdır.
3. Hata teriminin varyansı X değerlerine göre değişmez, yani sabittir.
4. Hata terimi normal bölünmeye sahiptir. Diğer bir ifadeyle hata teriminin değerleri kendi ortalamaları etrafında simetrik bir yapı izlerler.
5. Hata terimlerinin ardışık değerleri birbirlerinden bağımsızdır. (Yani birbirlerini izleyen hata terimleri arasında otokorelasyon yoktur.)
6. Serbest değişken değerleri sabit sayılar olup, hata terimi serbest değişkenden bağımsızdır.
7. Serbest değişkende ölçme hatası yoktur.
8. Serbest değişkenler arasında tam veya kuvvetli doğrusal bağlılık yoktur.
9. İlişkinin belirlenişi doğrudur.
10. Modelin matematiksel kalıbı tektir.

4. BASİT DOĞRUSAL REGRESYONA BAYESÇİ YAKLAŞIM

Bazı durumlarda x ve y gibi iki değişken arasındaki ilişki, bir model ile ifade edilmek veya bu ilişkiyi tanımlayacak bir eşitlik bulunmak istenebilir. Çoğu zaman ilişki kurulacak değişkenlerden biri olabilecek, y değişkeninin tahminine yardımcı olabilecek bir x değişkeni kullanılması uygun görülür.

$i=1,2,\dots,n$ için (x_i, y_i) nokta çiftlerinin sıralandığı n tane gözlemi içeren bir veri seti mevcut olsun. X değişkeni, tahmini oluşturulacak, hata teriminin dışında kalan bağımsız bir değişken olsun. Y değişkeni ise X değişkenine bağlı, hata terimini içeren bağımlı bir değişkeni ifade etsin. Bu ilişkinin yönü apsis ve ordinat değerlerine sırasıyla x ve y değerleri verilerek oluşturulabilir.

İki boyutlu bir grafikte bu iki değişken arasındaki bağlantı ifade edilebilir. Ancak bu ifadenin daha açık ifade edilebilmesi ve bu bağlantı için kullanılacak değişkenlerin arasındaki ilişkiyi regresyon modeli ile ifade edilebilmesi için, mevcut olan veri setine uygun eşitliklerin oluşturulması gerekir. Bu doğrusal ilişki, iki değişkenden oluşan basit bir eşitlik ile, ve bağımsız X değişkeni ile bağımlı Y değişkeni arasındaki ilişki bir doğru ile ifade edilebilir. Bu doğrunun parametreleri, β katsayısı ile α_0 sabit değeri olarak adlandırılabilir. Oluşturulabilecek bütün doğrular, bir ihtimal içerir.

Bazı kriterlerle, bilinmeyen parametrelerin tahminleri tanımlanabilir. En sık kullanılan kriter sistemi en küçük kareler yönteminin kriterleridir. Bu yöntemde, kalıntıların karelerinin toplamı minimize edilerek parametre değerleri bulunur. Bu durum, basit doğrusal regresyonda kullanılmaktadır. Bazı durumlarda regresyonun logaritmik olarak kurulması sonucunda bağımlı değişkenin logaritma aracılığıyla tahmini gerçekleştirilebilir.

Bazı veriler aracılığıyla olasılık modeli kurulamadığı için, değişkenler arasındaki etkileşim hakkında bir yorum yapılamayabilir. Bu verilerin nasıl bir regresyon modelinde, hangi bağımsız değişkenin, bağımlı değişkeni nasıl etkilediği değerlendirilir ve buna uygun bir regresyon modeli oluşturulur. Oluşturulan bu model ile parametrelerin etkileri incelenir. Mevcut verilere eklenen yeni gözlemler, parametrelerin oluşturduğu dağılımın türünü değiştirebilir. Bu durum önemlidir. Bu ilişkinin istenilen bilgileri verebilmesi için, Bayesçi yaklaşımın bazı kriterlerinden faydalanılması gerekebilir. Oluşturulan modeldeki parametreler üzerinde Bayesçi yaklaşımın sağladığı olanaklar incelenmelidir.

Bayesçi yaklaşım aşağıdaki gibi özetlenebilir,

sonsal $\propto$ *önsel* * *olabilirlik*

Bu model için oluşturulmuş önsel bir dağılıma ve bunun olabilirlik fonksiyonuna ihtiyaç vardır.

β ve $\alpha_{\bar{x}}$ için birikimli olabilirlik

Ekoloji gibi bazı bilimlerde, istatistiksel ilişkileri, bir taslak oluşturarak uygun nedenlere bağlayan yöntemler, Bayesçi yöntemler olarak tanımlanabilir. Bayesçi regresyon yöntemleri asıl konuyu teşkil etmektedir.

Bayesçi regresyon yöntemi, normal dağılımda yer alan rastlantı değişkenleri için Bayes teoreminden yararlanılarak oluşturulmuştur. (Gelman et al,1995) Bayes yöntemi daha önce, hiç olasılık dağılımları için kullanılmamıştı. Bu yöntem, ilk denemelerle birlikte olasılıkların alt kümelerinin kısıtlarını belirlemiş oldu.

Çeşitli matris yöntemleriyle bu yöntem, çok daha iyi açıklanabilir. Öncelikli olarak Bayesçi yaklaşımın temel bileşenlerinden söz etmek gerekir.

4.1 Bayesçi Basit Doğrusal Regresyon Modeli

$$Y_i = \beta_1 + \beta_2 X_i + U_i, \quad i = 1, 2, \dots, n \quad (4.1)$$

Y_i = bağımlı değişkenin i'inci gözlemi

X_i = bağımsız değişkenin i'inci gözlemi

U_i = hata payının i'inci gözlenmemiş değeri

doğrusal klasik regresyonun varsayımlarına uygun doğrusal bir modelle çalışılmaktadır.

$Y = (y_1, y_2, \dots, y_n)$ bağımlı değişkene ait gözlem değerlerine, $X = (X_1, X_2, \dots, X_n)$ bağımsız değişkene ait gözlem değerleri olduğu göz önünde bulundurulursa, birleşik olasılık yoğunluk fonksiyonu (3.2) ifadesi ile yazılabilir.

$$P(Y, X | \beta_1, \beta_2, \sigma^2, \theta) = P(Y | X, \beta_1, \beta_2, \sigma^2) g(X | \theta) \quad (4.2)$$

Burada θ , X için marjinal olasılık yoğunluk fonksiyonunun parametrelerini göstermektedir.

X_i , nin değerlerinin dağılımı β_1 , β_2 ve σ ile aynı olmadığından β_1 , β_2 ve σ için benzerlik fonksiyonu, $P(Y, X | \beta_1, \beta_2, \sigma^2)$ ' den bulunabilir. Y_i , verilen X_i , β_1 , β_2 ve σ^2 değerlerinden bağımsız bir dağılıma sahip olacaktır.

$$E(Y_i | X_i, \beta_1, \beta_2, \sigma^2) = \beta_1 + \beta_2 X_i$$

$$\text{Var}(Y_i | X_i, \beta_1, \beta_2, \sigma^2) = \sigma^2 \quad i = 1, 2, \dots, n \quad (4.3)$$

Olabilirlik fonksiyonu,

$$P(Y / X, \beta_1, \beta_2, \sigma) \propto \frac{1}{\sigma^n} \exp \left[-\frac{1}{2\sigma^2} \sum (Y_i - \beta_1 - \beta_2 X_i)^2 \right] \quad (4.4)$$

gösterilir.

Son olasılık yoğunluk fonksiyonun meydana gelmesi için 2 tane varsayıma ihtiyaç duyulmaktadır. Bunlardan ilki yetersiz ön bilgiye sahip olunacak, ikincisi ise belirsiz bir ön dağılıma sahip olması gerekmektedir.

$$P(\beta_1, \beta_2, \sigma) \propto 1/\sigma \quad -\infty < \beta_1, \beta_2 < +\infty \quad 0 < \sigma < +\infty \quad (4.5)$$

β_1 , β_2 ve σ 'ya ait olan ön olasılık yoğunluk fonksiyonu (β_1 , β_2 ve $\log \sigma$ 'nın tekdüze ve bağımsız dağıldığı varsayımı altında) son olasılık dağılım fonksiyonu ile birleştiğinde β_1 , β_2 ve σ için birleşik son olasılık dağılımı gösterimi şu şekildedir.

$$P(\beta_1, \beta_2, \sigma | Y, X) \propto \frac{1}{\sigma^{n+1}} \exp \left[-\frac{1}{2\sigma^2} \sum (Y_i - \beta_1 - \beta_2 X_i)^2 \right] \quad (4.6)$$

Bu ifade açılırsa, son olasılık yoğunluk fonksiyonunu ifadesi,

$$\begin{aligned} \sum (Y_i - \beta_1 - \beta_2 X_i)^2 &= v s^2 + n(\beta_1 - \hat{\beta}_1)^2 + (\beta_2 - \hat{\beta}_2)^2 \sum X_i^2 + 2(\beta_1 - \hat{\beta}_1)(\beta_2 - \hat{\beta}_2) \sum X_i \\ \hat{\beta}_1 &= \bar{Y} - \hat{\beta}_2 \bar{X} \quad \hat{\beta}_2 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2} \quad s^2 = v^{-1} \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2 \end{aligned} \quad (4.7)$$

ile tanımlanmış olur. (3.4) nolu ifadeyi aynı zamanda şu biçimde de yazmak mümkündür.

$$\sum (Y_i - \beta_1 - \beta_2 X_i)^2 = \sum \left\{ (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i) - [(\beta_1 - \hat{\beta}_1) + (\beta_2 - \hat{\beta}_2) X_i] \right\}^2$$

Eğer (3.4)'deki ifade (3.3)'da yerine yazılırsa,

$$\begin{aligned} P(\beta_1, \beta_2, \sigma | Y, X) &\propto \frac{1}{\sigma^{n+1}} \\ &\exp \left\{ -\frac{1}{2\sigma^2} \left[v s^2 + n(\beta_1 - \hat{\beta}_1)^2 + (\beta_2 - \hat{\beta}_2)^2 \sum X_i^2 + 2(\beta_1 - \hat{\beta}_1)(\beta_2 - \hat{\beta}_2) \sum X_i \right] \right\} \end{aligned} \quad (4.8)$$

elde edilir. β_1 ve β_2 ve için bilinen σ ile koşullu son olasılık yoğunluk fonksiyonu, iki değişkenli normal dağılıma uymaktadır ve,

$$\sigma^2 \begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix}^{-1} = \sigma^2 \begin{bmatrix} \frac{\sum X_i^2}{n \sum (X_i - \bar{X})^2} & \frac{-\bar{X}}{\sum (X_i - \bar{X})^2} \\ \frac{-\bar{X}}{\sum (X_i - \bar{X})^2} & \frac{1}{\sum (X_i - \bar{X})^2} \end{bmatrix} \quad (4.9)$$

biçimindeki varyans-kovaryans matrisine sahiptir.

4.1.1 Belirsiz Ön Olasılık Yoğunluk Fonksiyonu ile Son Olasılık Yoğunluk Fonksiyonu Çıkarımı

Ön bilginin dağılık veya belirsiz olduğu durumlarda, β ve $\log \sigma$ elemanları, $P(\beta, \sigma) \propto 1/\sigma$, $-\infty < \beta_i < \infty$, $0 < \sigma < \infty$ $i = 1, 2, \dots, k$ olarak ifade edilir. β ve σ parametreleri için bileşik sonsal olasılık yoğunluk fonksiyonu,

$$P(\beta, Y, X) \propto \frac{1}{\sigma^{n+1}} \exp \left\{ -\frac{1}{2\sigma^2} \left[v s^2 + (\beta - \hat{\beta})' X' X (\beta - \hat{\beta}) \right] \right\} \quad (4.10)$$

olarak elde edilir. Bu dağılım, çok değişkenli normal olasılık yoğunluk fonksiyonu ile ifade edilir, σ^2 nadiren bilindiği için $(X'X)^{-1}\sigma^2$ genellikle değerlendirilemez. β elemanları için marjinal sonsal olasılık yoğunluk fonksiyonu dağılımını oluşturmak için σ 'ya göre integral alındığında,

$$P(\beta, \sigma / Y, X) = \int_0^\infty P(\beta, \sigma / Y, X) d\sigma \propto \left[v s^2 + (\beta - \hat{\beta})' X' X (\beta - \hat{\beta}) \right]^{-n/2} \quad (4.11)$$

sonucuna ulaşılır. Bu ifade, p hakkında değerlendirmeler yapmak için kullanılır. σ için marjinal sonsal olasılık yoğunluk fonksiyonunun β 'ya göre (4.11) ifadesinden integral yoluyla elde edilişi,

$$P(\sigma / Y, X) = \int_{-\infty}^\infty \dots \int_{-\infty}^\infty P(\beta, \sigma / Y, X) d\beta \propto \frac{1}{\sigma^{v+1}} \exp \left(-\frac{v s^2}{2\sigma^2} \right) \quad (4.12)$$

şeklindedir. Bu dağılım dönüşük (inverted) gama olasılık yoğunluk fonksiyonu biçimindedir. (Kahyaoğlu, 1996)

4.1.2 Bilgi Verici Ön Olasılık Yoğunluk Fonksiyonu ile Son Olasılık Yoğunluk Fonksiyonu Çıkarımı

Dağılık ön olasılık yoğunluk fonksiyonu altında β ve σ parametreleri için elde edilen bileşik sonsal olasılık yoğunluk fonksiyonu olan

$$P(\beta, \sigma / Y, X) \propto \frac{1}{\sigma^{n+1}} \exp \left\{ -\frac{1}{2\sigma^2} \left[v s^2 + (\beta - \hat{\beta})' X' X (\beta - \hat{\beta}) \right] \right\} \quad (4.13)$$

ifadesi önsel olasılık yoğunluk fonksiyonu olarak kullanıldığında ve örneklem n_1 ve n_2 biçiminde iki bölümde tanımlanırsa, elde edilecek yeni ifade şu biçimdedir.

$$P(\beta, \sigma / Y_1, X_1) \propto \frac{1}{\sigma^{n+1}} \exp \left[-\frac{1}{2\sigma^2} (Y_1 - X_1 \beta)' (Y_1 - X_1 \beta) \right] \propto \frac{1}{\sigma^{n+1}} \exp \left[\frac{1}{2\sigma^2} \left[v_1 s_1^2 + (\beta - \hat{\beta}_1)' X_1' X_1 (\beta - \hat{\beta}_1) \right] \right] \quad (4.14)$$

Burada, $v_1 = n_1 - k$, $\hat{\beta}_1 (X_1' X_1)^{-1} X_1' Y$ ve $v_1 s_1^2 = (Y_1 - X_1 \hat{\beta}_1)' (Y_1 - X_1 \hat{\beta}_1)$ olarak ele alınmıştır.

Bu ifade, ön olasılık yoğunluk fonksiyonu kabul edildiğinde birinci örnek için β normal dağılıyor ve bunun yanı sıra σ biliniyorsa, ortalaması $\hat{\beta}_1$, kovaryans matrisi $(X_1' X_1)^{-1} \sigma$ olacaktır. σ için marjinal bir olasılık yoğunluk fonksiyonunu ise, v_1 ve s_1^2 parametreleriyle dönüşük gama formundadır. (4.15)'ün ikinci satırından hareketle,

$$P(\beta / \sigma, \hat{\beta}_1, s_1^2) \propto \frac{1}{\sigma^k} \exp \left[-\frac{1}{2\sigma^2} (\beta - \hat{\beta}_1)' X_1' X_1 (\beta - \hat{\beta}_1) \right] \quad (4.15)$$

yazılabilir.

İkinci örnek için olabilirlik fonksiyonu ise,

$$\ell(\beta / \sigma, Y_2, X_2) \propto \frac{1}{\sigma^{n_2}} \exp \left[-\frac{1}{2\sigma^2} (Y_2 - X_2 \beta)' (Y_2 - X_2 \beta) \right] \quad (4.16)$$

olarak elde edilir.

Bu örnek için β ve σ 'nın, birinci örnekteki değerlerle aynı parametre olduğu varsayımı altında (4.15)'deki önsel olasılık yoğunluk fonksiyonu (4.16)'daki olabilirlik fonksiyonu ile birleştirilirse, sonsal olasılık yoğunluk fonksiyonu olan,

$$P(\beta, \sigma / Y_1, X_1, X_2) \propto \frac{1}{\sigma^{n_1+n_2+1}} \exp \left[-\frac{1}{2\sigma^2} (Y_1 - X_1 \beta)' (Y_1 - X_1 \beta) + (Y_2 - X_2 \beta)' (Y_2 - X_2 \beta) \right] \quad (4.17)$$

elde edilir.

Böylece ilk örneğin önsel olasılık yoğunluk fonksiyonu ve ikinci örneğin olabilirlik fonksiyonu yardımıyla sonsal olasılık yoğunluk fonksiyonuna ulaşılmıştır. Bu (4.17)'de verilen ifade ile aynı dağılıma sahiptir. (Zellner, 1985)

4.2 Bayeşçi Çoklu Doğrusal Regresyon

$Y = (y_1, y_2, \dots, y_m)$ gözlemleri, m tane değişken üzerine $n \times m$ gözlem matrisi boyutuyla, $Y = Xp + U$ biçiminde modele alındığında, burada X , k bağımsız değişken üzerine verilen gözlemlerin k 'ninci sırasındaki bir $n \times k$ matrisini gösterir.

$\beta = (\beta_1, \beta_2, \dots, \beta_m)$ ise, regresyon parametreleri üzerine $k \times m$ boyutlu matristir. $U = (u_1, u_2, \dots, u_m)$ hata terimlerinin $n \times m$ boyutlu matrisidir. U ' ların birbirinden bağımsız dağıldığını, her birinin sıfır vektör ortalamalı ve pozitif tanımlı $m \times m$ boyutlu kovaryans matrisi Σ ile normal dağılıma sahip olduğu varsayılmaktadır.

Bu varsayımlar ışığında verilen X , β , Σ için Y 'nin ön olasılık yoğunluk fonksiyonu aşağıdaki gibidir:

$$\begin{aligned}
 P(Y / X, \beta, \Sigma) &\propto |\Sigma|^{-n/2} \exp \left[-\frac{1}{2} \text{tr}(Y - X\beta)'(Y - X\beta) \Sigma^{-1} \right] \\
 (Y - X\beta)'(Y - X\beta) &= (Y - X\hat{\beta})'(Y - X\hat{\beta}) + (\beta - \hat{\beta})' X' X (\beta - \hat{\beta}) \\
 (Y - X\beta)'(Y - X\beta) &= S + (\beta - \hat{\beta})' X' X (\beta - \hat{\beta}) \\
 \hat{\beta} &= (X' X)^{-1} X' Y \\
 S &= (Y - X\hat{\beta})'(Y - X\hat{\beta})
 \end{aligned} \tag{4.18}$$

Buradan, β ve Σ için olabilirlik fonksiyonu çıkarımı,

$$\ell(\beta, \Sigma / Y, X) \propto |\Sigma|^{-n/2} \exp \left[-\frac{1}{2} \text{tr} S \Sigma^{-1} - \frac{1}{2} \text{tr} (\beta - \hat{\beta})' X' X (\beta - \hat{\beta}) \Sigma^{-1} \right] \tag{4.19}$$

olacaktır. (Zellner, 1985)

4.2.1 Belirsiz Ön Bilgi Durumunda Çoklu Doğrusal Bayeşçi Regresyon

Σ ' nın, $m(m+1)/2$ ayırıcı elemanları ve p 'nin parametre değerleri hakkında ön olarak az bir bilgi veya belirsizlik söz konusu olduğunda, β ve Σ ' nın elemanlarının bağımsız dağıldığı varsayımı altında önsel olasılık yoğunluk fonksiyonu çıkarımı,

$$P(\beta, \Sigma) = P(\beta) P(\Sigma) \tag{4.20}$$

şeklindedir. Bayes kuralına göre sonsal dağılım ise aşağıdaki gibidir:

$$p(\beta_0, \dots, \beta_k, \sigma^2 | Y, X) \propto p(\beta_0, \dots, \beta_k, \sigma^2) \times \prod_i p(y_i | \mu_i, \sigma^2) \tag{4.21}$$

Standart bilgi vermeyen önsel $(\beta, \log \sigma^2)$ üzerinde sabittir. Bu da $p(\beta, \log \sigma^2) \propto \sigma^{-2}$, a eşittir.

Parametrelerin koşullu sonsal olasılık yoğunluk fonksiyonları aşağıdaki gibi gösterilebilir.

$$P(\beta_1 / \Sigma, Y, X) \propto \frac{1}{\sigma_{11}^{k/2}} \exp \left[-\frac{1}{2\sigma_{11}} (\beta_1 - \hat{\beta}_1)' X' X (\beta_1 - \hat{\beta}_1) \right] \quad (4.22)$$

$\hat{\beta}_1$ ortalama ve $(X'X)^{-1}\sigma_{11}$ varyans kovaryans matrisli çok değişkenli normal dağılımı, Σ için marjinal sonsal olasılık yoğunluk fonksiyonu,

$$P(\Sigma / Y, X) \propto |\Sigma|^{-v/2} \exp \left(-\frac{1}{2} \text{tr} \Sigma^{-1} S \right) \quad (4.23)$$

Burada $v = n-k+m+1$ 'dir. v açıklayıcı değişken sayısıdır.

4.2.2 Bilgi Verici Önsel Dağılım Fonksiyonu ile Çoklu Regresyon Modeli

Burada en önemli sorun, ön bilgiyi göstermede kullanılacak olasılık yoğunluk fonksiyonunun bütün durumlar için uygun olup olmadığıdır.

Eğer çok değişkenli regresyon modeli için basit bir bileşik önsel dağılım kullanılırsa, varyans-kovaryans matrisi bu dağılımdan etkilenecektir. Dolayısıyla parametre varyanslarına yansiyacaktır. Bu problemten kaçınmanın çözümü olarak, modelin tüm katsayıları için genel çok değişkenli normal olasılık yoğunluk fonksiyonu kullanımı önerilmektedir. Fakat buradaki sorun ise, böyle genel çok değişkenli normal olasılık yoğunluk fonksiyonunun, basit bileşik olasılık yoğunluk fonksiyonu kullanımında olduğu gibi bir işlemsel uygunluk sağlayamayışıdır. Bu matematiksel zorluk, bir çok durumda yine de göz ardı edilebilir. Çünkü basit bir olasılık yoğunluk fonksiyonu kullanılmasını gerektiren durumlarla da karşılaşilmektedir. Bu bilgiler ışığında ön olasılık yoğunluk fonksiyonu (4.24)' teki gibi yazılabilir.

$$P(\beta / \Sigma^{-1}) \propto |\Sigma^{-1}|^{-(m+1)/2} \exp \left[-\frac{1}{2} (\beta - \bar{\beta})' C^{-1} X (\beta - \bar{\beta}) \right] \quad (4.24)$$

$\hat{\beta}$, $k \times 1$ boyutunda bir vektör ve önsel olasılık yoğunluk fonksiyonunun ortalamasıdır. Araştırmacı tarafından belirlenir. C ise $k \times k$ boyutunda önsel varyans-kovaryans matrisidir. E elemanları hakkında önsel bir bilgi olduğu varsayılmaktadır ve daha önce kullanılan belirsiz önsel olasılık yoğunluk fonksiyonu ele alınırsa, (4.19)' daki olabilirlik fonksiyonuyla birleştirilerek, bileşik sonsal olasılık yoğunluk fonksiyonu,

$$P(\beta, \Sigma^{-1} / Y) \propto |\Sigma^{-1}|^{(n-m-1)/2} \exp \left\{ -\frac{1}{2} \text{tr} \Sigma^{-1} [S + (\beta - \hat{\beta})' X' X (\beta - \hat{\beta})] \right\}$$

$$P(\beta, \Sigma^{-1} / Y) \propto |\Sigma^{-1}|^{(n-m-1)/2} \exp \left[-\frac{1}{2} (\beta - \hat{\beta})' X' X (\beta - \hat{\beta}) \right] \quad (4.25)$$

olacaktır. Burada, $\beta = (\beta_1, \beta_2, \dots, \beta_m)$ olarak $\beta' = (\beta'_1, \beta'_2, \dots, \beta'_m)$ ve $\hat{\beta} = (X'X)^{-1}X'Y$ olarak, belirlenmiştir. (Kahyaoğlu, F., 1996)

Bu ifade Σ^{-1} için birleştirildiğinde,

$$P(\beta, / Y) \propto \left| S + (\beta - \hat{\beta})' X' X (\beta - \hat{\beta}) \right|^{-n/2} \exp \left[-\frac{1}{2} (\beta - \bar{\beta})' C^{-1} (\beta - \bar{\beta}) \right] \quad (4.26)$$

B için sonsal olasılık yoğunluk fonksiyonu elde edilir ve bu ifade, genelleştirilmiş çoklu Student t dağılımına uyar (Zellner, 1985, s.239).

5. DOĞRUSAL OLMAYAN BAYESÇİ YAKLAŞIM

Tanımlanmaya çalışılan, bilinmeyen girdi ve çıktı değişkenleri arasında oluşması beklenen, fonksiyonel formun doğru olması için bir veri analizinin yapılması gerekir. Veri analizi hakkında bir takım bilgiler mevcuttur. Ancak bu bilgilerin nasıl kullanılabileceği, çeşitli analizlerin karşılaştırılmasıyla mümkün olabilir.

Verilerin etkilerinin ne yönde olduğu yani doğrusal olup olmadığı ve hangi yöntemlerle analiz edileceğine karar vermeden önce, bir takım sorular sorulabilir. Bu sorular:

1. Kullanılabilecek veya benzer fonksiyon modellemesi ne olmalıdır?
2. Gerçek sonuçlara en yakın değerin bulunduğu nasıl anlaşılabilir?
3. Değerlendirme yapılacak konu üzerinde mevcut olan nitel ve nicel veriler arasında nasıl bir ilişki kurularak konuya dahil edilebilir?

5.1 Benzer Fonksiyonlar

Çok sayıda veriye sahip bir araştırmadan elde edilecek tahminlerden, hangisinin gerçeğe daha yakın olduğu gösterilebilir. Çoğunlukla lineer ve genelleştirilmiş lineer modeller, yapay sinir ağları, dalgacıklar, karar ağaçları ve çekirdek fiziğinde benzer fonksiyonlar kullanılır. Tahminci ve cevap arasındaki ilişki için, belirgin modellerin oluşmasını sağlarlar. X ve Y gibi biri bağımlı biri bağımsız iki değişken olsun. Bu iki değişken arasında zayıf ve güçlü değişkenin tanımlanması, olasılıksal ifadesi birkaç modellemeye örneklendirilebilir.

5.2 En İyi Model

Oluşturulan bir model, gerçeğin benzeri olmalıdır. Sadece bir modelin bulunmadığı, gerçek veri analizinin de gerçekleştirilemediği durumlarda, girdi ve çıktıları bir bütün olarak ele almak gerekir. Bu durum, birden fazla modelin yer aldığı tahmin sistemlerinde, en iyi modelin hangisi olduğunun bulunması için bir alternatif oluşturur.

5.2.1 Bayesçi Metodlar

Önsel nitel ve nicel verilerle belirli modeller geliştirilebilir. Örneğin, bir regresyonun düzgünlüğünün derecesi bilgisi, 2. türevinde sağlanan düzgün modellerin sınıflandırılmasında kullanılır. Bu gibi yaklaşımlar, Bayesçi yaklaşımların başlangıcını teşkil eder. Nitel ve nicel bilgiler ışığında, önsel dağılımlar oluşturulur ve bu dağılımların bilgisi dahilinde Bayes teoreminin yardımıyla, sonsal bir dağılım meydana getirilir. Bu durum, aşağıdaki gibi ifade edilebilir:

*Sonsal $\propto$ Önsel * Olabilirlik*

5.3 Bayesçi Metodoloji

Doğrusal olmayan regresyon modeli için gözlenen radyasyon değerleri verilerek, sezilerden yola çıkılarak Bayesçi formülü, $L_{göz}(v)$ ile ifade edilirse,

$$L_{göz}(v) = F(v; \beta, \eta) + E(v) \quad (5.1)$$

Gözlenen sistem, gürültü veya hata ile tanımlanan durum parametreleri ve gözlenen radyasyon değerleri arasındaki ilişki baz alınarak, $F(v; \beta, \eta)$ ve Gauss dağılımının varyansı olarak bilinen $v(v)$ ifadesiyle oluşturulan tahmin, $E(v)$ tanımlandığı durumlar için yukarıdaki formülasyon oluşturulabilir. Hata değerleri bağımsız tahmin edilir, bu nedenle hata terimlerinin kovaryans matrisleri, Σ_E , n değerinin kesikli dalga sayılarının kaydedildiği numaralar olarak belirttiği durumlar için, köşegendir ($\Sigma_E = \text{Diag}(v(v_1), \dots, v(v_n))$). Gauss oldukça akla yatkın bir tahmindir.

(5.1)'deki model, parametrelerin doğrusal olmayan bir fonksiyonu olan F ile akıl yürütmeyi sağlayan doğrusal olmayan bir modeldir. Bu parametreler β ve η vektörleri ile ifade edilmektedir.

Öncelikle yukarıdaki radyasyon örneği ile doğrusal olmayan bayesçi regresyonun nasıl ifade edilebileceğine dair bir fikir verilmek istenmiştir. Doğrusal olmayan regresyonunun, tahmindeki uygulamalarla tercihinin yapılabilme durumunun oluşması, sırasıyla açıklanmaya çalışılacaktır.

Risk analizi gibi çeşitli analizlerde, algıya dayalı tahminler yapılır. Ancak bunun bazı dezavantajları vardır. Oluşturulan dağılımın parametreleri için, her zaman güvenilir sonuçlar elde edilemeyebilir.

Algısal yanlışlıkların, objektif dağılım parametrelerine etkisinin olup olmadığı açık değildir. Bir dağılıma ait parametreler bilinmiyorsa, bu dağılımın tahmin edilebilmesi için bayesçi istatistik teknikler mevcuttur. Tekniğin genel özelliği, belirgin olamayan bir olaya önsel bir olasılık atfederken, bu olayın olma hali ile ilgili pek çok bilginin kullanılmasına gerek görülmemesidir. Bu bağlamda, değerlendirmeye tabi tutulan önsel olasılık (prior probability), objektif olasılık dağılımının parametreleri hakkında yürütülen bir tahmindir. Bu subjektif değerlendirmede iki varsayım ortaya çıkar. İlk varsayım, olayların olma halleri ile ilgili

simetrik bilgi eşit olasılıklarla sonuçlanır hükmüdür. Bu varsayım, simetrik iki olaya aynı olasılıkların atfedilmesi anlamındadır. İkinci olarak, eğer olayların olma uzayı k sayıda simetrik partisyonlara bölünürse (partitioned), her olayın olma olasılığı $1/k$ olur. Schmeidler, Bayesçi olasılık atfetmede benimsenen bu kurallardan ilkinin kabul etmekte birlikte, ikincisini kabul etmez. Schmeidler' a göre iki olayın olmasına ilişkin bilgi simetrik olduğunda, bu olaylara subjektif bir önsel olasılık atfederken eşit olasılıklar verilir. Dikkat edilecek olunursa, burada olasılıkların eşit olması, olayların birbirinden bağımsız veya bağımlı olması ile ilgili bir kavram olmayıp, doğrudan simetrik bilginin benzer olaylara eşit olasılık verilmesine neden olmasıdır.

Schmeidler atfedilen subjektif olasılığın veya toplanmaz olasılığın (nonadditive probability), objektif olasılık veya toplanabilir olasılık (additive probability) dağılımında içerilmeyen daha başka bilgileri içerdiğini bu nedenle, atfedilen subjektif olasılığın objektif olasılıktan daha çok bilgi taşıyacağını ileri sürmektedir. Bir önceki örnekte, eğer simetrik ve bütünler iki belirsiz olaya $3/7$ değeri verilecek olursa, $1/7$ karar vericinin bu olasılığa olan güvenini gösterecektir.

Schmeidler, birbirini bütünler olaylardan birine verilecek olasılığın objektif bir olasılık olmak zorunda olmadığını ileri sürer. Bu olasılığı atfederken kişi, gerçek dağılımın parametreleri hakkında bütün bilgiye sahip değildir. O halde, kendi yargısını kullanarak bu olay için bir olasılık değeri verebilir. İşte bu olasılık $3/7$ ise ve olayların bütünlerliğinden önce, olayın simetrik olduğu bilgisi kullanılacak olunursa, diğer olayın subjektif olasılığı $3/7$ olacaktır. O halde, artık $1/7$ olasılık, kişinin verdiği bu subjektif değere olan kendi güvenini gösterecektir.

Bu istatistik teorisinde, önemli bir felsefi tartışma dile getirilmektedir. Eğer öğrenmenin, varolan objektif dağılımı değiştirdiği kabul edilirse, verilen subjektif olasılıkların objektif dağılım parametreleri üzerine etkisi olmayacağını ileri sürmek doğru bir önerme olmaz. Subjektif ve objektif dağılımlar arasındaki bazı varyasyonların hesaplanabilir olduğu kabul edilir. Ancak, olasılıkları belirleyen olaylar ve bu olayların meydana geldiği örnek uzayın tanımı, bu varyasyonların birbirleriyle hiç ilgisinin olmadığını gösterir. Örneğin, insanın koyun yetiştirmeyi öğrenmesi, koyun nüfusunun dağılım parametrelerini etkilemesi doğaldır. Ancak, koyun sayısının dağılım parametrelerindeki bu değişim, Jüpiter' in yörüngesini belirleyen parametrelere etkisi olmayacaktır. O halde, sonlu uzaylarda dağılımı oluşturan parametreler, verilen subjektif olasılıklardan etkilenebilir. Çünkü her subjektif olasılık bir eylemi gerektirir ve bu eylemin olayları oluşturan sonlu uzayda bir değişime neden olması beklenir. Ancak, sonsuz uzayda oluşan olaylar için benzer hipotezi sürdürmek olanaksızdır.

Bu nedenle, simetrik olaylara atfedilen, subjektif olasılıklar toplanabilir (additive) olamazlar ve eğer, olaylar birbirini bütünler ise, simetrik olarak verilen olasılık toplanamaz (nonadditive). Simetrik olarak değerlendirilen bu olasılık, sadece bu olasılığı veren kişinin bu olasılığa karşı kendi güvenini gösterecektir. Ancak, kendine olan inanca bağlı bu güven düzeyi irrasyonel değildir. Good bu güveni rasyonel inanç derecesi (rational degree of belief) olarak adlandırır. Bu inanç rasyoneldir, çünkü kişisel yargılar kullanılarak en kötü tutarsızlıklar çıkarıldıktan sonra ortada kalan derecesi, subjektif olasılığı belirlerken kullanılır.

Subjektif olasılık, olaylar üzerine oluşan kişisel tercihlerden hesaplanmaktadır. Ancak bu hesaplama için kişisel tercihlerin sıralanmasında bazı düzenlilik varsayımları yapılmalıdır. Bu varsayımlardan üçü von Neumann-Morgenstern varsayımlarıdır. Bunlar; geçişlilik, süreklilik ve bağımsızlık varsayımlarıdır. Buna ek olarak, tercihlerin olayları meydana getiren hallerden bağımsız olduğu (state- independence of preferences) ve dağılımların bozulmazlığı (nondegeneracy) varsayılmalıdır. Dağılımın bozulmazlığı, parametrik durağanlık için bir önkoşuldur. Tercihlerin sıralanmasında, ön koşul asimetri ile bozulmazlık arasında bir ilişki vardır. Asimetri tercihlerin sıralanmasında düzenliliği sağlar ve olasılık içeren hallerde dağılımın bozulmazlığı ön koşulunu gerektirir. Bu derece kısıtlanmış tercihlere dayalı subjektif olasılık belirlemesi, gerçek bir davranışı açıklamaya yönelik olmaktan çok, teorik önermenin gerektirdiği sonuç için gerekli tutarlılığı, sağlamaya çalışan bir zorlama gibi gözükmemektedir.

Stokastik geçişlilik, diğer etkenler aynı olmak koşulu ile kesinlik ve risk arasındaki tercihin değişmeyeceğini varsayar. Bu varsayım tercihlerin değişmezliği hükmünün (invariance of preferences), zayıf bir uzantısıdır. Çünkü tercihlerin değişmezliği hükmü, yapılan seçimlerin kararın olduğu çerçeveden bağımsız olduğunu kabul etmektedir. Tversky bu anlamda geçişliliğin toplanabilir farklar kuralına (additive difference rule) uymasının ancak ve ancak matematiksel beklentiye tanımlayan olasılıkların doğrusal olmasıyla mümkün olduğunu söylemektedir. Oysa literatürde, stokastik transitivitenin dayandığı kesinlik ve risk arasındaki tercihlerin değişmezliği varsayımı, gözlemsel olarak doğrulanmamaktadır. Bu ikilem, Allais Çelişkisi (Allais Paradox) olarak anılır.

5.4 Allais Paradoksu

İnsanların belirsizlik ortamındaki davranışları, araştırmacılar açısından önemlidir. İnsanın olasılık içeren hallerdeki davranışları, genellikle oyun teorisi çerçevesi içerisinde incelenmiştir. Bu konuda yerleşik düşünce, insanların belirsizlik içeren bir oyuna katılması

için, oyunun beklenen değerinin pozitif olması veya eğer bir giriş maliyeti var ise, beklenen değer bu maliyetten yüksek olması gerektiği şeklindedir. Ancak, uygulamada insanların beklenen değer pozitif olması halinde bile, bu gibi olasılık içeren oyunları oynamada istekli olmadıkları görülmüştür. Bu çelişki St.Petersburg paradoksu olarak adlandırılır. İlk kez 18. yy' da Bernoulli tarafından bu çelişki araştırılmıştır.

5.5 Doğrusal Olmayan Modeller İçin Ek Bağlantılar

Gerçekleştirilen çoğu deneyin amacı, her deney için seçilmiş olan $\xi_1, \xi_2, \dots, \xi_p$ değişkenleri ve çıktı olarak oluşan y değerinin hatası arasındaki ilişkiyi anlayabilmektir. Gerçekleştirilmek istenen deney için oluşturulabilecek model aşağıdaki gibi yazılabilir;

$$y = \eta(\xi | \theta) + \varepsilon \quad (5.2)$$

Yukarıdaki denklem $\xi' = (\xi_1, \xi_2, \dots, \xi_p)$ olduğu durumlar için geçerlidir. ε sıfır ortalama değerli hatayı ve $\theta' = (\theta_1, \dots, \theta_k)$ tahmin edilmiş bilinmeyen parametreler için k boyutlu vektörleri ifade eder. Eğer θ parametresi doğrusal ise, θ için oluşturulabilecek doğrusal model aşağıdaki gibi olur;

$$\eta(\xi | \theta) = x_1\theta_1 + \dots + x_k\theta_k \quad (5.3)$$

$\xi_1, \xi_2, \dots, \xi_p$, $x_1, x_2, \dots, x_k$ değerlerinin oluşturduğu fonksiyonları tanımladığı durumlar için. Diğer taraftan, θ parametresi için bu model doğrusal olmayan model olarak ifade edilebilir.

y_i , n denemenin i . gözleminde oluşan bir değişken olsun. $\xi_i' = (\xi_{i1}, \xi_{i2}, \dots, \xi_{ip})$, p tane değişkenden oluşsun ve ε_i de uygun hata terimi olsun. Mevcut olan bu değerlerle aşağıdaki denklem oluşur;

$$y_i = \eta(\xi_i | \theta) + \varepsilon_i, i=1, \dots, n \quad (5.4)$$

$E(\varepsilon_i) = 0$ veya $E(y_i) = \eta(\xi_i | \theta)$ eşitliklerinin söz konusu olduğu durumlar için. Aşağıdaki gibi bir üssel dağılım için $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$ hata terimlerinin bağımsız olduğu farz edilsin.

$$p(\varepsilon | \sigma, \beta) = w(\beta)\sigma^{-1} \exp \left[-c(\beta) \left| \frac{\varepsilon}{\sigma} \right|^{2/(1+\beta)} \right], -\infty < \varepsilon < \infty \text{ için} \quad (5.5)$$

Bu nedenle, (5.4) modeli hesaba katılmazsa, θ parametresinin doğrusal ya da doğrusal olmayan, (θ, σ, β) değerleri için olabilirlik fonksiyonu aşağıdaki gibi ifade edilebilir.

$$I(\theta, \sigma, \beta | y) \propto [w(\beta)]^n \sigma^{-n} \exp \left[-c(\beta) \sum_{i=1}^n \left| \frac{\varepsilon_i}{\sigma} \right|^{2(1+\beta)} \right] \quad (5.6)$$

$\varepsilon_i = y_i - E(y_i) = y_i - \eta(\xi_i | \theta)$ eşitliği, $w(\beta)$ ve $c(\beta)$ fonksiyonları ile,

$$p(y | \theta, \sigma, \beta) = w(\beta) \sigma^{-1} \exp \left[-c(\beta) \left| \frac{y - \theta}{\sigma} \right|^{2/(1+\beta)} \right], \quad -\infty < y < \infty \quad (5.7)$$

olması için aşağıdaki eşitliklerin gerçekleşmesi gerekir.

$$c(\beta) = \left\{ \frac{\Gamma \left[\frac{3}{2}(1+\beta) \right]}{\Gamma \left[\frac{1}{2}(1+\beta) \right]} \right\}^{1/(1+\beta)} \quad (5.8)$$

ve

$$w(\beta) = \frac{\left\{ \Gamma \left[\frac{3}{2}(1+\beta) \right] \right\}^{1/2}}{(1+\beta) \left\{ \Gamma \left[\frac{1}{2}(1+\beta) \right] \right\}^{3/2}} \quad (\sigma > 0), (-\infty < \theta < \infty), (-1 < \beta \leq 1) \text{..için.} \quad (5.9)$$

θ ortalamalı ve σ standart sapmalı bir örneklemin parametreleridir. Aynı zamanda β , normal olmayan bir dağılımın basıklık ölçüsünü ifade etmektedir.

Eğer θ ve $\log \sigma$ 'nin bağımsız ve düzgün önsel dağılıma sahip olduğu varsayılırsa verilen bir β değeri için θ ve σ 'nın birikimli sonsal dağılımı aşağıdaki gibi ifade edilebilir.

$$p(\theta, \sigma | \beta, y) \propto \sigma^{-(n+1)} \exp \left[-c(\beta) \sum_{i=1}^n \left| \frac{\varepsilon_i}{\sigma} \right|^{2(1+\beta)} \right], \quad (\sigma > 0), (-\infty < \theta_j < \infty), j = 1, \dots, k \text{ için} \quad (5.10)$$

σ ve θ 'nın birikimli dağılımının birleştirilmesiyle aşağıdaki basit forma ulaşılır.

$$p(\theta | \beta, y) = J(\beta)^{-1} [M(\theta)]^{\frac{1}{2}n(1+\beta)}, \quad (-\infty < \theta_j < \infty), j = 1, \dots, k \text{ için}$$

$$M(\theta) = \sum_{i=1}^n |y_i - \eta(\xi_i | \theta)|^{2/(1+\beta)} \quad \text{ve} \quad J(\beta), \quad \text{nin} \quad \text{normalleştirilmiş} \quad \text{integrali}$$

$$J(\beta) = \int_R [M(\theta)]^{\frac{1}{2}n(1+\beta)} d\theta \quad (R : (-\infty < \theta_j < \infty) \dots \text{ve} \dots j = 1, \dots, k) \text{ ifadesiyle bulunur.}$$

Eğer β , (θ, σ) 'nin bağımsız rastlantı değişkenlerinin önsel bir ifadesiyse, verilen önsel $p = (\beta)$ olasılığı için β 'nin sonsal dağılımı aşağıdaki ifadeyle bulunur.

$$p(\beta | y) \propto p_u(\beta | y)p(\beta), -1 < \beta \leq 1, \quad (5.11)$$

$$\text{Ve } p_u(\beta | y) \propto \frac{\Gamma\left[1 + \frac{1}{2}n(1 + \beta)\right]}{\left\{\Gamma\left[1 + \frac{1}{2}(1 + \beta)\right]\right\}^n} J(\beta) \text{ olduğu durumlar için } (-1, 1) \text{ aralığında düzgün bir}$$

önsel dağılımın esaslarına dayandırılarak β 'nin sonsal dağılımı oluşturulmuş olur. Son olarak, β , θ 'nin bilinmeyen marjinal sonsal dağılımını ifade ettiği zaman;

$$p(\theta | y) = \int_{-1}^1 p(\theta | \beta, y) p(\beta | y) d\beta \propto \int_{-1}^1 p(\theta | \beta, y) p_u(\beta | y) p(\beta) d\beta \quad (-\infty < \theta_j < \infty), j = 1, \dots, k$$

için eşitliği oluşur.

Aşağıdaki iki varsayım için de yukarıdaki eşitlikler kullanılabilir:

- a) θ 'ya ilişkin kurulan model ya lineerdir ya da değildir.
- b) Hatanın dağılımı normaldir.

5.6 Sınıflandırma İçin Doğrusal Olmayan Modeller

Doğru sınıflandırmayı yapabilmek, istatistikî açıdan büyük önem taşır. Bunun için istatistiksel deneyimlerin iyileştirilmesi gerekir. Sınıflandırmanın kapsadığı belirli durumlar vardır. Örneğin C olarak adlandırılmış olan sınıfın, altında yer alan gözlemler veya C kategorisinin altında yer alan gözlemlerin sebep değişkenlerinin etiketi olduğu yerler, sınıflandırmanın kapsam alanı içerisine girer. Genel anlamda, 1, 2, ..., C olarak adlandırılmış bu sınıfların basit bir sıralaması yoktur. Daha dikkatli kontrol edebilmek için, sadece iki değişkenli, C=2 olarak, bir sınıflandırma yapılsın. Bu özel durum için genellikle kullanılan 0 ve 1 sınıf etiketi olarak kullanılsın. Daha sonra ikiden fazla değişken için C>2 durumu değerlendirmeye alınmalıdır.

Veriler için asıl merak edilmesi gereken durum, $y \in \{0, 1\}$, $x = \{x_1, x_2, \dots, x_p\}$ tahmini değerlerine benzetilmesinden etkilenmesidir. Fakat genellikle tahmini değerler ve modelden alınan cevaplar arasındaki ilişkinin yapısı çok açık değildir. Bu nedenle, verilerle tanımlanmış olmalıdır. Olasıdır ki sınıflandırma ile ilgili şu ana kadar kabul edilmiş yaklaşımlar, genellikle (y,x) olarak adlandırılan veri noktaları, örneklem dağılımında cevap bulunmasını istenen

noktalar. Aşağıda verilen denklemler bu noktaların nasıl kullanılabildiğine dair bir fikir edinilmesini sağlayacaktır;

$$p(y | \beta) = Br\{\mu(\beta)\} = \{\mu(\beta)\}^y \{1 - \mu(\beta)\}^{1-y}, y = 0,1, \quad (5.12)$$

$$\mu(\beta) = h(\beta_0 + \sum_{j=1}^p \beta_j x_j) \text{ olduğu durumlar için.} \quad (5.13)$$

Burada h , $[0,1]$ aralığında yer alan grafiğin gerçek doğrusunun ters fonksiyonunu ifade eder. Bu, mevcut aralıkta yer alan $\mu(\beta)$, Bernoulli dağılımının parametresi olduğu durumunun kesinleştirmek için gereken bir durum olmalıdır. Bunun gerçekleşebilmesi için aşağıdaki ifadenin gerçekleştirilebilmesi gerekir;

$$h(\eta) = \frac{\exp(\eta)}{1 + \exp(\eta)}, \quad (5.14)$$

$$h(\eta) = \phi(\eta), \quad (5.15)$$

$$h(\eta) = \exp[-\exp(-\eta)], \quad (5.16)$$

$$h(\eta) = 1 - \exp[-\exp(\eta)], \quad (5.17)$$

ϕ , verilen standart normal rastlantı değişkenlerinin birikimli dağılımının fonksiyonu olarak bulunduğu durumlar için

$$\phi(\eta) = \int_{-\infty}^{\eta} \frac{1}{\sqrt{2\pi}} \exp(-x^2 / 2) dx. \quad (5.18)$$

Yukarıda $h(\eta)$ ile verilmiş olan modeller, lojistik, probit, log-log ve bütünler(tamamlayıcı) log-log gibi doğrusal olmayan modeller için kullanılmaktadır.

Bu ifadelerle h , η , doğrusal tahmininin bir fonksiyonu olduğu kabul edilmektedir. η aşağıdaki gibi hesaplanmaktadır;

$$\eta = \beta_0 + \sum_{j=1}^p \beta_j x_j. \quad (5.19)$$

h ' nin seçimi (5.14) ve (5.15) denklemlerinde en genel biçimde ifade edilmesine rağmen, kişisel kontrol yapılırken bir problem haline gelebilmektedir. Sınıflandırma için (5.12) ve (5.13) denklemlerinden (5.20)' deki gibi bir eşitlik oluşturulabilir;

$$p(\beta | D) \propto p(D | \beta) p(\beta) \\ = \left[\prod_{i=1}^n \{h(\eta_i)\}^{\sum y_i} \{1 - h(\eta_i)\}^{n - \sum y_i} \right] p(\beta), \quad (5.20)$$

η_i değerinin, doğrusal tahminin bir değeri olarak ifade edildiği durumlar için. Ne yazık ki asıl sorun, alınan h fonksiyonunun hangi yapıda olduğudur. $p(\beta)$ gibi bir ön bilgiyi belirleyebilmek için bir yol yoktur, öyleki sonsal bilgi de aynı yapıdadır. Örneğin, yukarıdaki model için önsel bir eşlenik mevcut değildir. Özellikle bilinmeyen ilgili tahmincilerin sayılarına dayanarak modelin boyutlarının bulunması beklendiğinde, hesaplama problemlerinin başında bu durum yer almaktadır. Öncelikle bilinen analitik yapıda olan $p(\beta | D)$, sonsal dağılımın etkisindeki benzer yapıdaki önsellerin kullanılması, yapılacak değerlendirmelerde önemli bir yere sahiptir. Bu, örnekleme algoritmalarının etkisi için sırasıyla yapılan, marjinal olabilirlik için kapalı bir şekilde anlatıldığı kabul edilmektedir.

Tahminci uzayının (anakütlesinin) sabit olduğu farz edildiğinde, önsel eşleniğin eksikliği belirgin olarak hesaplamayı etkilemeyecektir. Bu durumda, örneklem yaklaşımlarından faydalanılabilir. Ancak bilinmeyen sayısının artabilir durumda olması gerçekleştirilecek uygulamalarda sorunlar oluşturabilecektir. Bazı durumlarda tahminci çıkartmak ya da eklemek sorunu çözebilmektedir. Ancak bu her durum için söz konusu olmamaktadır. Bütün regresyon modellerinde de öncelikli olduğu gibi, şimdi de tahminci kümesinde meydana gelebilecek bir değişim durumunda, β katsayısının mümkün olan en iyi şekilde güncellenmesinin nasıl bir yöntemle tanımlanacağı belirlenmelidir. Bu sadece katsayıları, örneklemenin bir yöntemi olmasına rağmen, bazı ayrıntılarla daha kullanışlı sonuçlar elde edilebilmesini sağlar.

z , rastlantı değişkenleri olarak verilmiş olsun,

$$z = \eta + \epsilon$$

η doğrusal bir tahminci ve ϵ ise 0 ortalama, sabit varyanslı standart normal bir rastlantı değişkeni olsun. Eğer

$$p(y = 1 | z, \beta) = \begin{cases} 1, \dots z > 0 \\ 0, \dots \text{diğer durumlarda} \end{cases}$$

ve $p(y = 0 | z, \beta) = 1 - p(y = 1 | z, \beta)$ olduğu farzedilirse, marjinal dağılım aşağıdaki gibi bulunur.

$$\begin{aligned}
p(y=1|\beta) \\
&= p(y=1|z>0, \beta)p(z>0|\beta) + p(y=1|z<0, \beta)p(z<0|\beta) \\
&= p(z<0|\beta) = p(\epsilon > \eta|\beta) = \Phi(\eta),
\end{aligned}$$

ve $p(y=0|\beta) = 1 - \Phi(\eta)$ olur. Bu nedenle, katsayılardaki koşulun sonucunun dağılımı, $\Phi(\eta)$ 'nin başarı olasılığıyla Bernoulli olur. Böylece, katsayılarla verilen cevabın aynı dağılımının bulunduğu görülür. Bu nedenle, yardımcı z değişkeninin içeriği, bu parametre üzerinde yok sayıldığında, orijinal modelden bir farkı yoktur.

Eğer bu yöntem uygulanırsa, gözlemlerin her biri, oluşturulan modelde n tane değişkenin eklenmesinin tanımlanmasıyla, model için ileriye doğru bir adım atılmış olmasını sağlar. Verilen z için yardımcı değişken önemlidir, bu durum β katsayısının güncelleştirilmesini kolaylaştırır. Bilinen hata varyansı ile x bir doğrusal regresyon oluşturur. Bir bakıma yardımcı değişken ile daha kolay çözülebilecek bir regresyonda sınıflandırma problemine dönüştürülür. Sınıflandırma için ek adım, z 'nin kendisine rassal bir genel modelleme sürecinin, bir parçası olarak güncellenmesi şeklinde olmalıdır.

$D = \{y_i, x_i\}_1^n$, yani bütün veri kümesi ayrıntılı olarak değerlendirilerek, β üzerinde $N(m, V)$ önseli ve $z = (z_1, z_2, \dots, z_n)$ yardımcı değişkenler kümesi olduğu varsayalım. Bu şartlar altında z 'nin doğru üzerinde, nasıl gözükebileceği tanımlansın. z 'nin katsayılarının koşulu aşağıdaki gibi ifade edilebilir;

$$p(\beta|z) \propto p(z|\beta)p(\beta).$$

Bununla birlikte z_i ile verilen cevaplar ve regresyon varyansının bilindiği durumlarda, normal doğrusal modelin katsayılarının sonsal dağılımı ile aynı olduğu gözlemlenir;

$$z_i = \eta_i + \epsilon_i = \beta_0 + \sum_{j=1}^p \beta_j x_{ij} + \epsilon_i$$

Bu nedenle, $p(\beta|z) = N(m^*, V^*)$,

$$m^* = (B'B + V^{-1})^{-1}(V^{-1}m + B'z),$$

$$V^* = (B'B + V^{-1})^{-1}.$$

olduğu durumlarda doğrulanır. Aynı zamanda, bilinen hata varyansları için verilerin marjinal olabilirliği aşağıdaki gibi ifade edilir,

$$p(D | z) \propto \frac{|V^*|^{1/2} \exp(-b^*)}{|V|^{1/2} (2\pi)^{n/2}},$$

$b^* = (z'z + m'V^{-1}m - (m^*)'(V^*)^{-1}m^*)/2$ olduğu durumlar için. Bilindiği üzere Gama dağılımı için eşlenik önsel dağılım aracılığıyla, v^{-1} , in koşullu sonsal dağılımı, diğer rastlantısal bulgular ile birlikte

$$\begin{aligned} p(v^{-1} | D, \beta, \sigma^2, \theta) &= p(v^{-1} | \beta, \sigma^2) \propto p(\beta | v^{-1}, \sigma^2) p(v^{-1}) \\ &= Ga\left(\delta_1 + \frac{1}{2}k, \delta_2 + \frac{1}{2\sigma^2} \sum_1^k \beta_i^2\right) = Ga(\delta_1^*, \delta_2^*) \end{aligned}$$

denklemleri ile ifade edilir. Bu denklemlerdeki $V = vI$ ve Gama dağılımındaki δ_1 ve δ_2 parametreleri olarak alınan, v^{-1} için hiyerarşik önseli olarak farzedilebilir. Bu nedenle, gerçekçi bir şekilde, sadece $p(\beta | z, v)$, $N(m^*, V^*)$ ile ifade edildiğinde aşağıdaki denklem oluşur,

$$p(v^{-1} | \beta) = Ga\left(\delta_1 + \frac{1}{2}k, \delta_2 + \frac{1}{2} \sum_1^k \beta_i^2\right)$$

Böylece bir sınıflandırma modelinde, verilen β için v gibi, verilen z değerleri içinde β örneğinin nasıl bulunabileceği anlaşılmış olur. Bununla birlikte, $z | \beta$ örnekleme ile z merkezden uzaklaşmayı sağlar. z' nin dağılımı Bayes teoreminden faydalanılarak,

$$p(z_i | y_i = 1, \beta) \propto p(y_i = 1 | z_i, \beta) p(z_i | \beta), \quad (5.21)$$

elde edilir, fakat bilindiği üzere $y_i = 1$, $z_i > 0$ ve $p(z_i | \beta) = N(\eta_i, 1)$ için, aşağıdaki koşullar oluşturulur,

$$p(z_i | y_i = 1, \beta) \propto \begin{cases} N(\eta_i, 1), \dots z_i > 0 \\ 0, \dots \dots \dots \text{diğer durumları için} \end{cases}$$

Bu nedenle, z_i kesikli normal dağılımdır. Benzer şekilde,

$$p(z_i | y_i = 0, \beta) \propto \begin{cases} N(\eta_i, 1), \dots z_i < 0 \\ 0, \dots \dots \dots \text{diğer durumları için} \end{cases} \text{ olur.}$$

Bu nedenle, η_i ortalamasıyla kesikli(truncated) normal dağılımdan, benzerini içeren verilerle verilmiş olan z_i genelleştirilebilir. Esas Normal dağılım ile gerçekleştirilebilmesine rağmen,

gerçek doğrunun yanlış yarısındaki benzer değerler belirgin olarak etkili değildir. Bu nedenle esas normal dağılımdan daha etkilidir.

Şimdiye kadar η' in doğrusal olduğu farzedilerek çeşitli ifadeler oluşturuldu, doğrusal olmadığı durumlarda ise aşağıdaki gibi ifade edilebilir.

$$\eta(x) = \beta_0 + \sum_{i=1}^k \beta_i B_i(x) = B\beta \quad (5.22)$$

Temel fonksiyonun verilerden ve çeşitli boyutların olasılıklarının, bazı θ vektörünün parametrelerle izah edilerek tanımlanmış olduğu durumlar için.

5.7 Doğrusal Olmayan Bayesçi Regresyon

Şu ana kadar doğrusal olmayan Bayesçi değerlendirmeler için sınıflandırma gibi, parametre tahminlerinde kullanılacak olan verilerin örnekleme açısından önemli çeşitli özelliklerinden bahsedilmiştir.

Genelde kullanılan doğrusal regresyon modelleri;

$$y_i = \beta_1 + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i$$

İfadesinde olur, $i = 1, \dots, N$ ' e kadar olduğu durumlarda. Bu model sadece doğrusal bağımlı ve açıklayıcı değişkenlerin arasındaki ilişkide kullanılır durumda değildir, dolayısıyla doğrusallık için dönüşüm yapılabilir. Cobb-Douglass üretimi olan çıktı ile ilişkili fonksiyon, y ; girdiler ise $x_2, \dots, x_k$ iken

$$y = \alpha_1 x_2^{\beta_2}, \dots, x_k^{\beta_k} \quad (5.23)$$

Eğer eşitliğin iki tarafının logaritması alınır ve hata terimi eklenirse, aşağıdaki gibi bir regresyon modeli oluşur;

$$\ln(y_i) = \beta_1 + \beta_2 \ln(x_{i2}) + \dots + \beta_k \ln(x_{ik}) + \varepsilon_i$$

$\beta_1 = \ln(\alpha_1)$ olduğu durumlar için. Bu tanımlama, açıklayıcı ve bağımlı değişkenlerin logaritması ve küçük farklarını ifade etmektedir.

Doğrusal hale getirilemeyen bazı fonksiyonlar mevcuttur. Gerçek doğrusal olmayan fonksiyonel yapının bir örneği, üretim fonksiyonunun değişimiyle verilsin.

$$y_i = \left(\sum_{j=1}^k \gamma_j x_{ij}^{\gamma_{k+1}} \right)^{\frac{1}{\gamma_{k+1}}} \quad (5.24)$$

ε ve y sırasıyla bağımlı değişkenlerinin N vektörlü kümenin hataları ve gözlemleri, ve X ise $N \times k$ boyutlu matris kümesinde, k açıklayıcı değişkenlerinin gözlemleri olsun. Bu durumda, aşağıdaki standart varsayımlar sıralanır;

1) ε , $N(0_N, h^{-1}I_N)$ kümesi içerisinde.

2) X ' in bütün elemanları, herbiri belli (örneğin rastlantı değişkeni değilse) ya da eğer rastlantı değişkenleri ise X rastlantı değişkenleri $p(X | \lambda)$ olasılık yoğunluk fonksiyonuyla ε ' nun bütün elemanlarının hiç içermediği parametrelerin vektörünü temsil ettiği durumlarda.

Temel düşünce, genel doğrusal olmayan regresyon modeli nasıl ifade edilebileceğidir.

$$y_i = f(X_i, \gamma) + \varepsilon_i$$

X_i , X ' in i . satırı, $f(\cdot)$ X_i ve γ parametresinin bir vektörüne bağlı, bir fonksiyon olsun. Bu açıklamalarla birlikte, yukarıdaki notasyon matris formu ile ifade edilecek olunursa;

$$y = f(X, \gamma) + \varepsilon \quad (5.25)$$

$f(X, \gamma)$, $f(X_i, \gamma)$ ile verilmiş olan i . elemanlarının fonksiyonunun N vektörü olduğu durumlar için. Sonsal bir algoritmanın tam ifadesi $f(\cdot)$ yapısına bağlı olacak ve (5.25) modeli (5.24) modelinden önce değerlendirilecektir.

Doğrusal olmayan regresyon modeli, doğru sonucun bulunabilmesi için önemli bir adımdır. Fakat herhangi bir modelde kabul edilebilecek tekniklerin bir sayısının tanımlanması için bir olasılık verilmesi gereklidir. Doğrusal regresyon modeli, analitik sonsal sonuçları içeren bazı durumlardaki olasılığın çok özel bir durumu olabilir. Önsel dağılımlar, analitik sonuçların kabul edilebilirliğine engel oluşturmalarına rağmen, bazı normal doğrusal regresyon modelleri için kabul edilebilir özel teknikler mevcuttur. (Örneğin Gibbs örnekleme) Fakat birçok model kullanılan bu özel metodları kabul etmez ve bu durum, herhangi bir modelde kullanılabilecek kapsamlı metodların geliştirilmesi için önemlidir. Doğrusal olmayan regresyon modeli, benzer doğrusal regresyon modellerine küçük eklemeler bağlamındaki bazı kapsamlı metodların tanımlanmasını kabul eder. Sonsal simülasyon ile ilgili olarak Metropolis-Hastings algoritmaları ile sonsal simülatörler sınıflandırılması gibi bir örnek mevcuttur. Bu algoritmalarından sırası geldikçe bahsedilecektir. Kapsamlı metodların öncelikli incelemesi

1994 yılında Gelfand ve Dey tarafından geliştirilmiş olan marjinal olabilirlik ile değerlendirmeye tabi tutulmuş ve sonsal tahmindeki p- değerinin bir modelde hangi eşitlikle nasıl ölçülebileceği üzerine yorumlarda bulunulmuştur.

5.7.1 Olabilirlik Fonksiyonu

Çoklu normal dağılımın tanımında kullanılan doğrusal olmayan modeldeki olabilirlik fonksiyonu aşağıdaki gibi ifade edilebilir:

$$p(y|\gamma, h) = \frac{h^{\frac{N}{2}}}{(2\pi)^{\frac{N}{2}}} \left\{ \exp \left[-\frac{h}{2} \{y - f(X, \gamma)\}' \{y - f(X, \gamma)\} \right] \right\} \quad (5.26)$$

Doğrusal regresyon modeli ile OLS (En küçük kareler) miktarları açısından da doğal eşlenik önsel bir form yukarıda önerilen bu ifade ile yazmak mümkündür. Burada, böyle bir basitleştirme olmazsa, $f(\cdot)$ çok karmaşık formlardan oluşur.

5.7.2 Önsel Dağılım

Önsel olasılık, $f(\cdot)$ fonksiyonunun ne olduğuna ve γ' 'nın nasıl yorumlanacağına bağlıdır. Araştırmacılar, bu parametreler için uygun değerlerin ne olabileceğine ilişkin önsel bilgiye sahip olabilecektir. Bu nedenle önsel elastikiyet, deneysel içeriklere bağlıdır. Çeşitli tartışmalar sonucunda oluşturulan genel olarak nitelendirilmiş bazı elde edilen ürünler, $p(\gamma, h)$ olarak tanımlanan basit önsel bilgi ile birlikte bilgi içermeyen önsel dağılım, doğrusal regresyon modeli ile aşağıdaki gibi ifade edilebilir:

$$p(\gamma, h) \propto \frac{1}{h} \quad (5.27)$$

Bu önsel γ ve $\ln(h)$ için Düzgün dağılıma sahiptir. Birçok durumda, doğrusal olmayan regresyon modelinin parametreleri için de bu mantıklı, bilgi içermeyen önsel bilgi kullanılabilir.

5.7.3 Sonsal Dağılım

Sonsal dağılım, önsel kere olabilirlik ile orantılıdır ve aşağıdaki gibi yazılabilir:

$$p(\gamma, h|y) \propto p(\gamma, h) \frac{h^{\frac{N}{2}}}{(2\pi)^{\frac{N}{2}}} \left\{ \exp \left[-\frac{h}{2} \{y - f(X, \gamma)\}' \{y - f(X, \gamma)\} \right] \right\} \quad (5.28)$$

Genel olarak, bu ifadenin, $p(\gamma, y)$, $f(\cdot)$ için basit bir şekilde çözümü yoktur ve herhangi bir

bilinen dağılım yapısı içerisinde de değerlendirilemez. (5.27)'deki bilgi içermeyen önsel kullanıldığında, hata payı, h , temel aşamasında analitiksel olarak bütünleştirilerek, γ için aşağıdaki ifade oluşturulur:

$$p(\gamma | y) \propto \left[\{y - f(X, \gamma)\}' \{y - f(X, \gamma)\} \right]^{-\frac{N}{2}} \quad (5.29)$$

5.7.4 Bayesçi Hesaplama: Metropolis-Hastings Algoritması

Sonsal dağılımla ilişkili analitik sonuçlar, sonsal bir simülatör ile elde edilir. $f(\cdot)$ 'in bazı durumları için Gibbs örnekleri türetilir. Bazı durumlarda, $p(\gamma | y)$ için uygun yaklaşım kendisi olarak önerilebilir ve öyleyse, örneklem etkisi kullanılabilir. Burada, üçüncü ihtimal olan Metropolis-Hastings algoritmasına değinilecektir. Bu algoritmalarda, geniş çaptaki çeşitliliğe sahip modeller için sonsal simülatör oluşturmanın nasıl sınıflandırılacağı üzerinde durulacaktır. Bu algoritmada da önceki algoritmalarda kullanılmış olan parametrelerin bir vektörü olan θ ve bunun fonksiyonları olan $p(y | \theta)$, $p(\theta)$ ve $p(\theta | y)$ diğer bir ifadeyle sırasıyla olabilirlik, önsel ve sonsal olasılık fonksiyonları kullanılacaktır.

Metropolis- Hastings algoritması, etkili örneklem ile bazı benzerliklere sahiptir. Bunlardan faydalı olabileceklerden biri yaklaşık olasılıkların mevcut olması fakat diğeri rassal grafikler nedeniyle sonsal dağılımın oluşturulmasında belirli zorluklara sahip olmasıdır. Etki örnekleminde (importance sampling) önem fonksiyonu olarak adlandırılmış olan fonksiyon, Metropolis- Hastings algoritmasında örnek üretici yoğunluk fonksiyonu (candidate generating density) olarak adlandırılmaktadır. θ^* , $q(\theta^{(s-1)}; \theta)$ ile belirtilmiş olan yoğunluktan elde edilen görüntüyü ifade etsin. Bu notasyon, $\theta^{(s-1)}$ ifadesine bağlı olan yoğunluk, θ rastlantı değişkeninin alınan aday görüntüsü θ^* 'dır şeklinde de ifade edilebilir. Diğer bir ifadeyle, Gibbs örnekleyicisi gibi bugünkü görüntünün önceki görüntü ile bağlantılı olduğu sonucuna varılır. Bu nedenle Metropolis- Hastings algoritması, Gibbs örnekleyicisi gibi Markov Zinciri Monte Carlo Algoritması (MCMC) ve zincirdeki gibi görüntü değerleri alır ($s=1, \dots, S$ için $\theta^{(s)}$ değerlerinin oluşması gibi)

Etki örneklemiyle, farklı görüntülerin ağırlıklandırılması ile sonsal dağılımdan farklılaştırma gerçekleştirilebilir. Metropolis- Hastings ile eşit ağırlıktaki görüntüler elde edilebilirken, uygun görüntülerin hepsi elde edilememektedir. Diğer bir ifade ile $g(\cdot)$ ilgili fonksiyon ise, $E[g(\theta) | y]$ ifadesi, $\hat{g}_s$ sembolü ile bu fonksiyonun tahminini kapsayabilir. Bu tahmin aşağıdaki ifade ile hesaplanabilir:

$$\hat{g}_s = \frac{1}{S} \sum_{r=1}^S g(\theta^{(s)}) \quad (5.30)$$

Metropolis-Hastings Algoritması genellikle aşağıdaki yapıları içerisinde barındırır:

0. Adım: $\theta^{(0)}$ gibi bir başlangıç değeri seçilir.
1. Adım: $q(\theta^{(s-1)}; \theta)$ yoğunluk fonksiyonundan elde edilen uygun görüntü θ^* alınır.
2. Adım: $\alpha(\theta^{(s-1)}, \theta^*)$ kabul edilebilir bir olasılık hesaplanır.
3. Adım: $\alpha(\theta^{(s-1)}, \theta^*)$ olasılığıyla $\theta^{(s)} = \theta^*$ ve $1 - \alpha(\theta^{(s-1)}, \theta^*)$ olasılığıyla da $\theta^{(s)} = \theta^{(s-1)}$ eşitlikleri kurulur.
4. Adım: Sırasıyla 1, 2 ve 3. adımlar tekrarlanır.
5. Adım: S' nin görüntülerinin $g(\theta^{(1)}), \dots, g(\theta^{(S)})$ ortalaması alınır.

Bu adımlar ile herhangi bir ilgili fonksiyon için $E(g(\theta) | y)$ beklenen değerinin bir tahmininin yapılması sağlanır.

Gibbs örnekleyicisi gibi, Metropolis-Hastings algoritması da genellikle $\theta^{(0)}$ başlangıç değerini seçerek değerlendirmeyi başlatır. Bu başlangıç değerinin etkisinden emin olmak için 0' a eşitlenir, S_0 başlangıç değerini çıkarmak için genellikle bu yöntem kullanılır. Gibbs örnekleyicisi için sunulan MCMC tanısı, ele alınan değerlerin yeterli sayıda olduğundan ve çıkarılan başlangıç değerinin yeterli olduğundan emin olabilmek için Metropolis-Hastings algoritmasıyla da değerlendirme yapılabilir.

Kabul edilebilir olasılık için belirli bir formül olarak, kısaca $\alpha(\theta^{(s-1)}, \theta^*)$ verilir. Ancak ilk olarak, bu olasılığın sahip olduğu olumlu özelliklerden söz etmek gerekir. Örnek üretici yoğunluk, sonsal için tanımlanmamıştır ve bu nedenle, parametre uzayındaki her bir alanın doğru görüntülerinin sayısı kolay bir şekilde elde edilemez. Metropolis-Hastings algoritması, çekilen her aday içerisinde kabul edilmeyenler için, doğruluğu neyin belirleyeceğinin bulunmasını sağlar. Bu algoritmaya göre, sonsal olasılığın yüksek olduğu durumlarda benimsenen olasılık yüksek ve sonsal olasılığın düşük olduğu durumlarda ise benimsenen olasılık düşük olarak türetilir. Eğer $\theta^{(s-1)}$, düşük sonsal olasılığın bir ifadesi ise algoritma, $\theta^{(s-1)}$ ' dan hızla uzağa hareket etme eğilimi gösterecektir. Fakat eğer $\theta^{(s-1)}$, yüksek sonsal olasılığa sahip ise, algoritma aynı şekilde kalacaktır. Yüksek sonsal olasılıklı bir noktada

bulunan algoritma, dolaylı olarak etkili örnekleyicinin ağırlıklandırılmış görüntüsüne benzer davranışlarda bir ağırlık verir. Benzer sonuçlar, θ^* içinde gerçekleştirilir. Verilen $\theta^{(s-1)}$ için, eğer $\theta^{(s-1)}$ 'dan daha yüksek bir sonsal olasılık alanına sahip ise θ^* ifadesi yüksek olasılıkla kabul edilmiş olan, uygun değer olarak alınmak istenir. θ^* 'nın düşük olasılıklı örneklemde seçilen değerlerinin, yüksek olasılıkla reddedilmiş olması istenecektir.

Düşük sonsal olasılığın artışa doğru geçtiği alanlarda bağlantı kurularak taşınması için oluşan eğilimde θ^* ve $\theta^{(s-1)}$ değerlerinin bağlı olduğu kabul edilen olasılık için bir önsezi oluşturması gerekmektedir. Yüksek sonsal olasılık alanlarında bir bağlantı kurulması yerine, düşük olasılıklı alanlar üzerinde inceleme yapılması daha çok tercih edilen bir durumdur. Kabul edilen olasılığın bir yolu, bağlantılardan anlamlar çıkarmaktır. Ancak, düşük sonsal olasılıklardan yükseğe taşınamaz. Kabul edilen olasılığın bulunması gereken yapı aşağıdaki gibi ifade edilebilir:

$$\alpha(\theta^{(s-1)}, \theta^*) = \min \left[\frac{p(\theta = \theta^* | y) q(\theta^*; \theta = \theta^{(s-1)})}{p(\theta = \theta^{(s-1)} | y) q(\theta^{(s-1)}; \theta = \theta^*)}, 1 \right] \quad (5.31)$$

(4.31) nolu eşitlikte, minimum alınmıştır. Çünkü olasılık değeri 1' den küçük bir değer almalıdır.

Etki örnekleyicisi gibi, Metropolis-Hastings algoritması, ilk bakışta herhangi bir sonsal simülasyon problemi için mükemmel bir çözüm olarak görünür. Sonrasında, $E(g(\theta) | y)$ 'nin tahmininin oluşumunda kullanılan, $s=1, \dots, S$ için $\theta^{(s)}$ değişkeninin sıklığını içeren (5.31) denkleminde kullanılmış aday görüntü kabul edilir veya reddedilir ve $q(\theta^{(s-1)}; \theta)$ gibi herhangi bir uygun yoğunluktan rastlantısal bir örnek olarak ifade edildiği görülür. Ne yazık ki bunu bir örnekle ifade etmek kolay değildir. Eğer aday üretici yoğunluk, iyi seçilememişse, oluşan aday görüntülerinin tümü reddedilir ve zincir, zamanının büyük bölümü için belirli noktalarda saplanmış bir şekilde kalır. Bu nedenle, seçilen aday üretici yoğunluk ve MCMC ifadesi, algoritmanın yöneliminin doğrulanması ile daha itinalı bir şekilde oluşturulmalıdır. Aday üretici yoğunluğun seçimi için çok sayıda strateji oluşturulması mümkündür. Bu stratejilerden iki tanesi üzerinde durulacaktır.

5.7.4.1 Bağımsız Zincir Metropolis-Hastings Algoritması

İsminde ifade edildiği gibi, Bağımsız Zincir Metropolis-Hastings algoritması, doğru bağımsız değişkeni seçebilmek için aday geliştirme yoğunluk fonksiyonunda kullanılır. Bu algoritma,

$q(\theta^{(s-1)}; \theta) = q^*(\theta)$ ’dır ve aday geliştirme yoğunluk fonksiyonu $\theta^{(s-1)}$ ifadesine bağlı değildir. Böyle bir uygulama, sonsal dağılım için varolan uygun tahminlerin yapılması gereken durumlarda faydalıdır. Bu uygun tahmin, aday üretici yoğunluk fonksiyonu gibi kullanılabilir. Eğer böyle bir şey yapılmışsa, kabul edilen olasılık basit bir ifade ile:

$$\alpha(\theta^{(s-1)}, \theta^*) = \min \left[\frac{p(\theta = \theta^* | y) q^*(\theta = \theta^{(s-1)})}{p(\theta = \theta^{(s-1)} | y) q^*(\theta = \theta^*)}, 1 \right] \quad (5.32)$$

eşitliği ile gösterilir. Bağımsız Zincir Metropolis-Hastings algoritması, etkili örnekleme ile yakından ilişkilidir. Bu tespit ile görülebilirki, eğer etkili örnekleme ağırlıkları için benzer ağırlıklandırmalar yapılmış ise:

$$w(\theta^A) = \frac{p(\theta = \theta^A | y)}{q^*(\theta = \theta^A)}$$

kabul edilen olasılık aşağıdaki gibi yazılabilir:

$$\alpha(\theta^{(s-1)}, \theta^*) = \min \left[\frac{w(\theta^*)}{w(\theta^{(s-1)})}, 1 \right]$$

Bir deyişle, kabul edilen olasılık, sadece eski ve aday değişkenlerde değerlendirilen etkili örnekleme ağırlıklarının bir oranıdır.

Doğrusal olmayan regresyon modeli açısından ise, bu algoritmanın faydası, uygun tahmin yoğunluk fonksiyonunun ifadesi olan $f(\cdot)$ ’nin bulunmasına bağlıdır. Tahmin yoğunluk fonksiyonunun seçimi için eksiksiz genel bir uygulama yoktur. Ancak bir seçim yapılacaksa, MCMC için bir sonuç algoritması kullanılabilir.

Ortak bir uygulama olarak, $q^*(\theta)$ bulmak için bir temelde, frekansçı maksimum olabilirlik sonuçlarının kullanımını içerir. Frekansçı yaklaşım, maksimum olabilirlik tahmincisini, asimptotik normallikle ve asimptotik kovaryans ile aşağıdaki gibi ifade eder:

$$\text{var}(\hat{\theta}_{ML}) = I(\theta)^{-1}$$

Olabilirlik fonksiyonunun logaritmasının 2. türevinin beklenen değerinin tersi olarak tanımlanan bilgi matrisi, $I(\theta)$ olduğu yerlerde (y için bir beklenti mevcut iken):

$$I(\theta) = -E \left[\frac{\partial^2 \ln(p(y | \theta))}{\partial \theta \partial \theta'} \right]$$

ifadesi ile tanımlanır. Diğer bir ifade ile, eğer örnek çapı büyük ise, bilgi matrisinin tersi, $p(y|\theta)$ 'nın biçiminin, iyi bir fikir olduğu ortaya çıkar. Eğer bilgi matrisi direkt hesaplanamıyorsa, farklı nümeriksel uygulamalar ile $\frac{\partial^2 \ln(p(y|\theta))}{\partial \theta \partial \theta'}$ ifadesi hesaplanır ve bunun için $\widehat{\text{var}}(\hat{\theta}_{ML})$ ile ifade edilmiş olan, $\text{var}(\hat{\theta}_{ML})$ için yapılan tahmininin içeriği kullanılır. $\frac{\partial^2 \ln(p(y|\theta))}{\partial \theta \partial \theta'}$ ifadesini için bazı ek terimler Hessian olarak ifade edilir ve Bayesçi için, eğer örnek büyüklüğü oldukça büyük ve nispeten bilgi içermeyen bir önsel dağılıma sahipse, sonsal dağılım yaklaşık olarak $\hat{\theta}_{ML}$ ortalamalı ve yine yaklaşık olarak $\widehat{\text{var}}(\hat{\theta}_{ML})$ kovaryans matrisli bir Normal dağılıma uyum sağlar. Bazı modeller için, bilgisayar yazılımlarında maksimum olabilirliğin bazı değerlerinin direkt hesaplanması mümkündür. Alternatif olarak, MATLAB gibi bazı bilgisayar paketleri sayesinde özel fonksiyonların optimum hesaplanması da mümkün olabilir. Bu nedenle, maksimum olabilirliğe ait nicelikler bu programlar aracılığı ile hesaplanabilir. Eğer, olabilirlik fonksiyonunun maksimize etmeye ve $\widehat{\text{var}}(\hat{\theta}_{ML})$ ifadesinin bulunması için sıradan bir yazılıma ihtiyaç varsa, sonsal dağılımın maksimizasyonu alınır ve $\widehat{\text{var}}(\hat{\theta}_{ML})$ değeri yerine yaklaşık değerini bulmak için sonsal dağılımın 2. dereceden türevi alınabilir. Eğer önsel dağılıma ait bilgi kullanılıyorsa, sonsal dağılıma dair daha uygun sonuçlar elde edilebilir. Asimptotik sonuçlar, $f_N(\theta | \hat{\theta}_{\max}, \widehat{\text{var}}(\hat{\theta}_{\max}))$ fonksiyonun yaklaşık değeriyle oluşturulan sonsal dağılımı akla getirir. Aşağıdaki malzeme olarak kullanılan, en fazla olasılık sonuçları ile ilgili tahminler dayanak noktası olur, ancak eğer mevcutlarsa $\hat{\theta}_{\max}$ ve $\widehat{\text{var}}(\hat{\theta}_{\max})$ yer değiştirebilirler.

$q^*(\theta) = f_N(\theta | \hat{\theta}_{ML}, \widehat{\text{var}}(\hat{\theta}_{ML}))$ ile ilgili ayarlar, bazı durumlarda daha iyi çalışabilirler. Fakat aday geliştirici yoğunluk fonksiyonu gibi t dağılımının kullanımı da yaygındır ve bunun için $q^*(\theta) = f_t(\theta | \hat{\theta}_{ML}, \widehat{\text{var}}(\hat{\theta}_{ML}), v)$ oluşturulur. Bunun nedeni ise, bu uygulamada, araştırmacılar tarafından sonsal dağılımın en yoğunlaştığı tarafın, aday geliştirici yoğunluk fonksiyonu için çok önemli olduğu bulunmuştur. Geweke (1989) etkili örnekleminin bu kadar önemli olmasını başlıca neden olarak belirlemiştir, ancak aynı mantık Bağımsız Zincir Metropolis-Hastings algoritmasını da barındırmaktadır.

Normal dağılım, çok zayıf kuyruklara sahiptir. t- dağılımı da, v ' nün küçük değerleri için çok yoğun kuyruklara sahiptir. t- dağılımının faydalı özellikleri, $v \rightarrow \infty$ giderken Normal

dağılıma yaklaşır ve kuyruğun yoğun olduğu durumlarda ν küçük değerler alır. Bazı durumlarda, sonsal dağılımın baskın aday geliştirici yoğunluk fonksiyonunun kuyruklarında yer alan, ν değerinin bulunması sonsal dağılımın kontrolü ile mümkün olabilir. Ancak genelde araştırmacılar tarafından, ν için küçük bir değer seçilir ve algoritmanın yönelimini sağlamak için MCMC tanısı kullanılır.

Bu durum bir anlamda, uygun olmayan geliştirilmiş örnek değerlerine, t- dağılımının kullanıldığı durumlarda baskı oluşturmaları açısından önemlidir. Örneğin, sonsal dağılım çoklu modele sahip ise, tek modelli t- dağılımı çok iyi çalışmaz. Ayrıca eğer sonsal dağılım, sınırlı bir aralıkta tanımlanmışsa(örneğin Gama dağılımı sadece pozitif reel sayılar için tanımlanır.), tüm reel sayılarda tanımlanmış olan t dağılımı çok iyi çalışmayabilir.

$\widehat{\text{var}}(\hat{\theta}_{ML})'$ yi elde etmek, (5.26) ifadesindeki her bir fonksiyonun 2. dereceden diferansiyelleştirerek veya bazı bilgisayar programlarının alt yordamlarında yer alan, nümeriksel türevleme kullanılarak mümkün olabilir. Bu adımlar, $f(\cdot)$ fonksiyonunun yapısının tahminlerine bağlı olabilir, bu nedenle böyle bir durumda doğrusal olmayan regresyon modeli hakkında hiçbir yorum yapılamayabilir.

Bağımsız Zincir Metropolis- Hastings algoritması veya etkili örnekleme kullanılarak yoğunluk fonksiyonu oluşturulur. Bununla birlikte, birçok model asimptotik, örneklem büyüklüğü sonsuza giden ve sonsal dağılımı Normal dağılıma yaklaşan bir yapıya sahip olabilir. Bu durumlarda bulunan modeller için, eğer örneklem büyüklüğü makul bir büyüklüğe sahipse, $f_t(\theta | \hat{\theta}_{ML}, \widehat{\text{var}}(\hat{\theta}_{ML}), \nu)$ sonsal dağılımı da mantıklı sonuçlar dahilinde tahmin etmek daha uygun olacaktır. Aşağıdaki rastlantısal algoritmada bu strateji kullanılacaktır.

5.7.4.2 Rastlantısal Yürüyüş Zincir Metropolis- Hastings Algoritması

Sonsal dağılım için uygun bir tahmin oluşturulamadığı durumlarda, Rastlantısal Yürüyüş Zincir Metropolis- Hastings Algoritması kullanılabilir. Sezgisel olarak, Bağımsız Zincir Metropolis- Hastings Algoritmasıyla, sonsal dağılıma benzer bir yoğunluk ifadesi elde edilebilir ve kabul edilebilir olasılık, sonsal ve tahmin dağılımları arasındaki farklılıkların doğrulanması için kullanılır. Bağımsız Zincir Metropolis- Hastings Algoritmasıyla, hiçbir girişim sonsal dağılımın çeşitli alanlarında orantılı görüntülerle almak, aday geliştirici yoğunluk fonksiyonu geniş bir alandan seçilmesi, sonsal dağılımın tahmini için değildir.

Biçimsel olarak, Rastlantısal Yürüyüş Zincir Metropolis-Hastings Algoritmasının aday geliştirici tahmin değişkenine göre;

$$\theta^* = \theta^{(s-1)} + z \quad (5.33)$$

z' nin artık rastlantısal değişken olarak ifade edildiği durumlarda. (5.33) ifadesindeki aday olarak ifade edilen değişken, rastlantısal yürüyüşe göre üretilir. Burada, adaylar geçerli noktalardan rastgele yönlerde gösterilir. Kabul edilebilir olasılık, uygun yönlü bağlantılarla kesinleştirilir. θ^* ve $\theta^{(s-1)}$, (5.33) ifadesinde simetrik olarak girilsin ve bu nedenle, ifadede $q(\theta^*; \theta = \theta^{(s-1)}) = q(\theta^{(s-1)}; \theta = \theta^*)$ eşitliği ortaya çıkacaktır. Bu kabul edilebilir olasılığın basit ifadesi, aşağıdaki gibi tanımlanır;

$$\alpha(\theta^{(s-1)}, \theta^*) = \min \left[\frac{p(\theta = \theta^* | y)}{p(\theta = \theta^{(s-1)} | y)}, 1 \right] \quad (5.34)$$

ve net bir rasgele yürüyüş zinciri, daha sonra olasılık bölgelerine doğru hareket eğilimi içerisinde görülebilir.

z için yoğunluk seçimini, aday üretici yoğunluk fonksiyonunun belirli yapısı tanımlar. Çoklu Normal dağılım ortak ve uygun seçimdir. Bu durumda, ($\theta^{(s-1)}$, ortalamadır.) (5.33) ifadesi Normal dağılımın ortalamasını tanımlar ve araştırmacılar Σ olarak ifade edilen kovaryans matrisini seçmelidirler. Normal dağılım için mevcut notasyon açısından:

$$q(\theta^{(s-1)}; \theta) = f_N(\theta | \theta^{(s-1)}, \Sigma) \quad (5.35)$$

Bütün araştırmacılar, Σ 'nin seçimine ihtiyaç duyarlar. Bu, kabul edilebilme ihtimali için, ne çok düşük ne de çok yüksek eğilimli değerler seçilmelidir. Eğer kabul edilebilir olasılık, çok düşükse, aday görüntüleri hemen hemen her zaman reddedilir ve nadiren, sadece hareket eğilimi ile bağlantılıdır. Eğer bu zincir sonsal dağılım üzerinde değerlendirilecekse, bu S 'nin büyük olmasını ima etmesi gibi iyi bir durum değildir. Küçük kabul edilebilir olasılıklar, Σ 'un çok büyük olduğunu ve en iyi aday görüntüsünün sonsal dağılımın çok olası olduğu alanlarda sonsal dağılımın kuyruklarının dışında üretildiğini belirtir. Diğer uç noktalarda, Σ 'nin çok küçük bir değeri ifade ettiği durumlardaki gibi, yakınlarında bulunan kabul edilebilir olasılık iyi değildir. Eğer, Σ çok küçükse θ^* ve $\theta^{(s-1)}$ kabul edilebilir olasılığın yakınlarındaki değerlerden birinin bir başkasına eğilimi olacaktır. Bu durumda da, S , zincir ile bütünüyle sonsal dağılımın araştırılarak kesinleştirilen, uygulama olanağı sağlamayan bir büyüklüğe sahip olabilir.

Optimum kabul oranı verebilecek, genel bir kural yoktur. Sonsal ve aday geliştirici yoğunluk fonksiyonlarının Normal dağıldığı özel durumlarda, optimum kabul oranı, yüksek boyutlu

problemlerin düşük değerleriyle tek boyutlu problemlerde, 0.45 olarak hesaplanmıştır. Boyutlar sonsuza yaklaşırken, optimum kabul olasılığı 0.23' e yaklaşır. Kabul edilen başka bir kural, bu olasılığı 0.5 olarak alır. Genelde, bu bölgede kabul olasılığına yakınlaşmayı sağlamak için Σ seçilmiş ise, yanlış olasılık seçilebilir. Ancak genelde bu algoritmanın doğrulanması için, MCMC tanısı kullanılır.

Σ' nın, 0.2 ve 0.5 aralığında ortalama kabul olasılığını sağlamak için, θ' nın skaler olduğu yerlerde (dolayısıyla Σ' da skaler olur), seçim yapabilmek için yeterli bilginin elde edilebilmesini sağladığı söylenmektedir. Bu durumda, nedensel kabul olasılığı bulunana kadar Σ için çeşitli değerlerle Rastlantısal Yürüyüş Metropolis-Hastings algoritmasıyla, basit denemeler gerçekleştirilebilir. Σ , $\frac{p(p+1)}{2}$ değerinde olana kadar, böyle durumlar için p-boyutlu böyle yaklaşımlar zor olabilir. Bu gibi durumlarda, Ω , θ' nın sonsal kovaryans matrisinin tahmini ve c nümerik değerler aldığı durumlar için $\Sigma = c\Omega$ kümesi için daha iyi çalışır. Uygun bir kabul edilebilir olasılık değeri bulunana kadar, c için farklı değerler denenebilir. Bu yaklaşım $\text{var}(\theta | y)$ nın tahmini, Ω ifadesinin bulunmasını gerektirir. Bu iki yolla gerçekleştirilebilir. Üzerinde durulacak ilk nokta, en basit yol olan, $\Sigma = cI_p$ ile başlayarak düzenlemek ve (eğer yeterince değişken mevcut değilse, tamamen faydalı olabilecek 0.000001 ile 0.99999 arasında yer alacak bir olasılık değeri) kabul edilebilir olasılık için, faydalı olabilecek bir c değeri bulmaktır. c' nin bu değeri Ω' nın kalıntı tahmininde kullanılan değer olabilir. $\Sigma = c\Omega$ kümesi oluşturularak nedensel kabul edilebilir olasılığın oluşturulmasında c için yeni bir değer bulunabilir. Σ' nun en iyi değeri için bulunan Ω' nın, iyi bir değeri kullanılabilir. Bu süreç, Σ için iyi ve uygun bir değer bulunana kadar tekrar edilir. Rassal Yürüyüş Metropolis- Hastings Algoritması' nın kullanılmasıyla bu tarz araştırmalarda ek program kullanmaya gerek kalmaz.

Araştırma için diğer bir alternatif, Bağımsız Zincir Metropolis - Hastings Algoritması'nın tartışmaları altında tanımlanan, maksimum olabilirlik varyansının tahmini olan, $\text{var}(\hat{\theta}_{ML})$ için Ω kümesidir.

5.7.4.3 Metropolis –within- Gibbs

Metropolis-Hastings Algoritması, $p(\theta | y)$ için sonsal dağılımın simülatörünün oluşturulmasını sağlar. Sonsal dağılımın simülatörü için Gibbs örnekleyicisi kullanılabilir. Bağımsız Normal Gama önsel dağılımıyla Normal Doğrusal Regresyon modeli için, bir Gibbs

örnekleyicisi, Gama dağılımı için $p(h|y, \beta)$ ve Normal dağılım olan $p(\beta|y, h)$ için uygulanmasını kolaylaştırır. Doğrusal olmayan regresyon modelinde, h için bilgi içermeyen veya bağımsız Gama önsel dağılımı, $p(h|y, \gamma)$ ile bir Gama dağılımını ima etmektedir. Ancak, $p(\gamma|y, h)$, (5.28)' deki değere orantılı olacaktır ve bu nedenle uygun bir görüntüye sahip yoğunluk fonksiyonunun modeli alınmayacaktır. İlk bakışta $p(\gamma|y, h)$ ve $p(h|y, \gamma)$ ifadelerini içeren Gibbs örnekleyicisi ile doğrusal olmayan regresyon modeli için, tanımlama yapılamayacağını ima etmektedir. Ancak, simülasyon ile oluşturulan değerlerin sonuçlarının $p(\gamma|y, h)$ için Metropolis-Hastings Algoritması ile kullanılırsa, $s = 1, 2, \dots, S$ için $\gamma^{(s)}$ ve $h^{(s)}$ değerleri, geçerli sonsal simülasyon görüntüleri olacaktır. Gibbs Örnekleyicisinde kullanılan sonsal koşullar, ya (veya her ikisi) $p(\theta_{(1)} | y, \theta_{(2)})$ ve $p(\theta_{(2)} | y, \theta_{(1)})$ için Metropolis –Hastings Algoritmasında kullanılmak için kabul edilir. Eğer Gibbs örnekleyicisi bu iki koşulu içerirse, bu durum doğrulanmış olur. Böyle Metropolis-within –Gibbs Algoritmaları, koşullamalarının çoğunun basit görüntülerinden oluşan sonsal dağılımlara sahip, fakat bir veya ikisinde uygun yapılara sahip olmayan modellerinde çok yaygındır. İkinci koşullamalar için, Metropolis-Hastings Algoritmaları kullanılabilir.

5.7.5 Model Uygunluğu İçin Bir Ölçüm: Sonsal Tahmin P-Değeri

Model karşılaştırması için tipik Bayesçi metodu, tamamen belirlenmiş iki modelin göreceli olasılığı olan sonsal bahis oranıdır. Ancak, bir araştırmacının modelin performansını belirli alternatif bir model ile göreceli olarak araştırmak yerine, mutlak olarak araştırmak isteyeceği bazı durumlar vardır. Ayrıca, araştırmacının hazırlıksız, bilgi içermeyen önsel kullanmak isteyeceği birçok durum vardır ve eğer böyle bir önsel, tüm modellerde ortak olmayan parametreler üzerinde kullanılırsa, sonsal bahis oranı (odds) anlamsız olabilir. Böyle durumların tümünde, sonsal tahmin p-değeri yaklaşımı sonsal bahis oranına mantıklı bir alternatiftir.

Sonsal tahmin p-değeri yaklaşımını uygulamak için, gözlemlenen veri olan y ile çalışılmakta olan modelden çıkartılabilecek gözlemlenebilir veri $y^\dagger$ 'yi ayırt etmek önemlidir. $g(\cdot)$ bir faiz fonksiyonu olsun. Veriyi gördükten sonra $p(g(y^\dagger)|y)$, modelin $g(y^\dagger)$ hakkında söylediği herşeyi özetler. Başka bir ifadeyle, bu modelin üretebileceği veri kümelerinin tipini anlatır. Gözlemlenen veri için $g(y)$, doğrudan hesaplanabilir. Eğer $g(y)$, $p(g(y^\dagger)|y)$ ' nin uç noktalarındaysa, o zaman model $g(y)$ ' yi açıklamakta başarısız olur. Aslında, sonsal tahmin p-değeri, açık olarak gözlemlenmiş olandan ziyade daha uç özelliklere sahip bir veri kümesi

sağlayan bir modelin olasılığıdır

$p(g(y^\dagger)|y)$, tahmini çıkarım için kullanılan metoda çok benzer olan simülasyon metotları kullanılarak hesaplanabilir. Bu, (5.36) ifadesi ile yazılabilir;

$$p(g(y^\dagger)|y) = \int p(g(y^\dagger)|\theta, y)p(\theta|y)d\theta = \int p(g(y^\dagger)|\theta)p(\theta|y)d\theta \quad (5.36)$$

Burada son eşitlik, θ' ya bağlı olan şu gerçekten çıkar; asıl data $y^\dagger$ hakkında ek bir bilgi sağlamamaktadır. Sonsal simülatör $p(\theta|y)$ ' den çıkarımlar sağlar ve yapay veriyi sadece, tahmin için kullanılan veriye eş olacak şekilde, verilen parametre değeri için olan modele benzeterek $p(g(y^\dagger)|\theta)$ 'ye, benzetilebilir.

Sonsal tahmin p-değeri iki farklı yolla kullanılabilir. İlk olarak, mutlak anlamda veri üretmesinin ne kadar mümkün olduğunu gösteren uygunluk ölçümü olarak kullanılabilirler. İkinci olarak, farklı modelleri karşılaştırmak için kullanılabilirler. Yani, eğer bir model diğerinden düşük olan sonsal tahmin p-değeri sağlarsa, bu ilk modele karşı bir kanıttır. Ancak, Bayesçi'lerin çoğu sonsal bahis oranlarını tercih eder çünkü ikincisi bilgi içermeyen önseller kullanılmadığı takdirde sonsal bahis oranını anlamsız kılar veya açıklanmalarını zorlaştırır.

Sonsal tahmin p-değeri yaklaşımı bir faiz fonksiyonu, $g(\cdot)$, seçimi gerektirir. Doğru bir $g(\cdot)$ seçeneği, deneysel uygulamaya bağlı olarak değişecektir. Pratik bir örnek olarak, doğrusal olmayan regresyon modeline dönelim. Bu model için mevcut olan,

$$y_i^\dagger = f(X_i, \gamma) + \varepsilon_i$$

$i=1, \dots, N$. Alternatif olarak, hatalar hakkında yapılmış olan varsayımlar verilmiş olsun,

$$p(y^\dagger | \gamma) = f_N(y^\dagger | f(X, \gamma), h^{-1}I_N) \quad (5.37)$$

Burada $f(X, y)$, (5.25)' de tanımlanan N-vektörüdür. Modelin parametrelerinin verilmiş değerleri için, $y^\dagger$ 'nin benzer değerlerinin oldukça basit olduğuna dikkat edilmelidir, yalnızca 'Çok Değişkenli Normal Dağılım' dan gelen çıkarımları içermektedir. Bu kolaylık birçok model için ortaktır ve sonsal tahmin p-değerlerinin birçok değişik durumda kolaylıkla hesaplanmasını sağlar.

(5.27) ifadesinde verilen bilgi içermeyen önselli doğrusal olmayan regresyon modeli için, h dışardan entegre edilebildiği için (5.37) ifadesi daha da kolaylaştırılabilir. Özellikle, (5.28) ifadesinden (5.29) ifadesine geçmek için gerekli olanla hemen hemen eş (identical) bir türetme kullanılarak (5.38)' deki eşitlik gösterilebilir;

$$p(y^\dagger | \gamma) = f_t(y^\dagger | f(X, \gamma), s^{-2} I_N, N) \quad (5.38)$$

burada

$$s^{-2} = \frac{[y - f(X, \gamma)]' [y - f(X, \gamma)]}{N} \quad (5.39)$$

Dolayısıyla, y' ye bağımlı olarak, $y^\dagger$ 'nın çıkarımları çok değişkenli t dağılımı kullanılarak alınabilir. Bu çıkarımlar, bu modelin üretebileceği veri kümesinin tipini yansıttığı gibi açıklanabilir. Sonsal tahmin p-değeri yaklaşımının kullandığı fikir şudur; eğer model mantıklı bir model ise, asıl gözlemlenen veri kümesi, model tarafından çoğunlukla üretilen tipte olmalıdır.

Frekansçılar, hataların OLS tahminlerini, ϵ_i , hesaplar ve bunlar kalıntılar olarak adlandırılır. Bu kalıntıların özellikleri, modelin altındaki varsayımların mantıklı olup olmadığını açığa çıkarmak için araştırılabilir. Bayesçi bağlamında, $i=1, \dots, N$ için hatalar şöyle tanımlanır;

$$\epsilon_i = y_i - f(X_i, \gamma)$$

Bu hataların çeşitli özellikleri olduğu varsayalım. Özellikle, i.i.d $N(0, h^{-1})$ olduğu varsayılmıştır. Bu varsayımlar belirli bir veri kümesinde mantıksız olabilir ve dolayısıyla araştırmacı bunları test etmek isteyebilir. "Hatalar i.i.d Normal dağılır." şeklindeki kısa ifade birçok varsayım içerir (örneğin, varsayımlarının birbirinden bağımsız oldukları, ortak bir değişkene (varyans gibi) sahip oldukları, vb.) ve araştırmacı araştırmak için bunlardan herhangi birini seçebilir. Normal' in iki özelliği, simetrik olması ve kuyruklarının belirli şekilde olmasıdır. İstatistikte, Normal dağılımı çarpıklık göstermez ve kuyruklarının belirli bir basıklığı vardır. Çarpıklık ve basıklık dağılımın üçüncü ve dördüncü momentleri üzerinden ölçülür ve standart Normal için (yani $N(0,1)$) üçüncü moment sıfırdır ve dördüncü moment üçtür. Bu yüzden Normallik varsayımı, yaygın olarak kullanılan çarpıklık ve basıklık ölçülerinin ikisinin de sıfır olmasını gerektirir:

$$\text{Çarpıklık} = \frac{\sqrt{N} \sum_{i=1}^N \epsilon_i^3}{\left[\sum_{i=1}^N \epsilon_i^2 \right]^{\frac{3}{2}}} \quad (5.40)$$

ve

$$\text{Basıklık} = \frac{N \sum_{i=1}^N \varepsilon_i^4}{\left[\sum_{i=1}^N \varepsilon_i^2 \right]^2} - 3 \quad (5.41)$$

Çarpıklığın ve basıklığın bu ölçüleri, ε_i gözlemlenemez olduğu için, doğrudan hesaplanamaz. Frekansçı, ε_i ' yi önceki formüldeki uygun OLS kalıntıları ile yer değiştirir ve sonucu çarpıklık veya basıklık için bir test yürütmek için kullanılır. Kalıntılardaki çarpıklık veya basıklık bulgusu Normallik varsayımının uygun olmadığını belirtir.

Formal olarak,

$$E[\text{Çarpıklık} | y] = E \left\{ \frac{\sqrt{N} \sum_{i=1}^N [y_i - f(X_i, \gamma)]^3}{\left[\sum_{i=1}^N [y_i - f(X_i, \gamma)]^2 \right]^{\frac{3}{2}}} | y \right\}$$

bir kez sonsal simülator elde edildiğinde açık biçimde hesaplanabilen birşeydir. Yani, *Çarpıklık* basit olarak model parametrelerinin (ve verinin) bir fonksiyonudur ve dolayısıyla, sonsal ortalama i , herhangi bir faiz fonksiyonun sonsal ortalama i ile aynı yoldan hesaplanabilir. $E[\text{Basıklık}|y]$ aynı biçimde hesaplanabilir. Normallik varsayımı mantıklı ise, $E[\text{Basıklık}|y]$ ve $E[\text{Çarpıklık}|y]$ nin ikisi de kabaca sıfır olmalıdır.

$E[\text{Çarpıklık}|y]$ ve $E[\text{Basıklık}|y]$ gözlemlenen verinin birer fonksiyonudur ve sonsal simülator kullanılarak hesaplanabilir. Her gözlemlenebilir veri için, $y^\dagger$, $E[\text{Çarpıklık} | y^\dagger]$ ve $E[\text{Basıklık} | y^\dagger]$ aynı biçimde hesaplanabilir. Eğer bu sondaki fonksiyonları, geniş çeşitlilikteki gözlemlenebilir veri kümesi için hesaplanırsa, çarpıklık ve basıklık için ayrı ayrı bu modelin üretebileceği değerler dağılımlarını elde edilebilir. Eğer, $E[\text{Çarpıklık} | y]$ ve $E[\text{Basıklık} | y]$ ' den herhangi biri $E[\text{Çarpıklık} | y^\dagger]$ ve $E[\text{Basıklık} | y^\dagger]$ ' nin dağılımının kuyruklarından uzakta ise bu Normallik varsayımına karşı bir kanıttır. Şu da üzerinde durulmaya değerdir ki, $E[\text{Çarpıklık} | y]$ ve $E[\text{Basıklık} | y]$ ' nin ikisi de basitçe birer sayıdır, oysa $E[\text{Çarpıklık} | y^\dagger]$ ve $E[\text{Basıklık} | y^\dagger]$ ' nin ikisi de olasılık dağılımları (5.36) ifadesi ile hesaplanan rassal değişkenlerdir. Daha önceki notasyona göre, $g(y) = E[\text{Çarpıklık} | y]$ veya $E[\text{Basıklık} | y]$ ve $g(y^\dagger) = E[\text{Çarpıklık} | y^\dagger]$ veya $E[\text{Basıklık} | y^\dagger]$ ' dir.

Pratikte, bilgi içermeyen önsel kullanan doğrusal olmayan regresyon modelde çarpıklık için,

sonsal tahmin p-değerlerini hesaplayan bir program şu adımları içerecektir. Basıklık durumu aynı yolla yapılabilir. Bu adımlar bir sonsal simülatörün türetildiğini varsayar (yani sonsaldan çıkarımlar üreten bir Metropolis-Hastings algoritması).

Adım 1: Sonsal simülatörü kullanarak bir çıkarım, $y^{(s)}$, alınır.

Adım 2: (5.38) ifadesi kullanılarak $p(y^\dagger | \gamma^{(s)})$ bir veri kümesi üretilir, $y^{\dagger(s)}$.

Adım 3: $i = 1, \dots, N$ için $\varepsilon_i^{(s)} = y_i - f(X_i, \gamma^{(s)})$ kurun ve Çarpıklık^(s) yi elde etmek için (5.40) ifadesi bu noktada hesaplanır.

Adım 4: $i = 1, \dots, N$ için $\varepsilon_i^{\dagger(s)} = y_i^{\dagger(s)} - f(X_i, \gamma^{(s)})$ kurun ve Çarpıklık^{†(s)} yi elde etmek için (5.40) ifadesi bu noktada hesaplanır.

Adım 5: S kere 1., 2., 3. ve 4. adımları tekrar edilir.

Adım 6: $E[\text{Çarpıklık}|y]$ 'yi tahmin etmek için Çarpıklık⁽¹⁾, ..., Çarpıklık^(s) S değerlerinin ortalaması alınır.

Adım 7: 6. adımdaki $E[\text{Çarpıklık}|y]$ tahmininizden daha küçük olan Çarpıklık^{†(1)}, ..., Çarpıklık^{†(s)} S değerlerinin oranı hesaplanır. Eğer bu sayı 0.5' den küçükse, bu sizin sonsal tahmin p-değeri tahmininizdir. Aksi takdirde, sonsal tahmin p-değeri; bir eksi bu sayıdır.

Unutulmamalıdır ki, çarpıklık için sonsal tahmin p-değeri 0.05' e eşitse denilebilirki, "Bu model çarpıklık ölçülerini zamanın sadece yüzde beşinde gözlemlenenden daha büyük olarak üretir. Dolayısıyla, bu modelin gözlemlenen veriyi üretmesi mümkün değildir."

5.7.6 Model Karşılaştırması: Gelfand-Dey Metodu

Bilgi içeren önselleri ortaya çıkarmak mantıklı ise ve araştırmak istediğiniz iki veya daha fazla model varsa, sonsal olasılık oranları model karşılaştırması tercih edilen metod olarak kalır. Doğrusal olmayan regresyon modeli durumunda, $f(\cdot)$ için farklı seçenekleri karşılaştırmak istenebilir. Bu tipik olarak, iç içe yerleştirilemeyen (non-nested) model karşılaştırmasını içerecektir. Alternatif olarak, doğrusal olmayan ile doğrusal regresyon modellerini karşılaştırmak istenebilir. Doğrusal olmayan bir modelle sadece ekstradan bir karmaşıklık gerçekten buna değer ise çalışmak istenir. (5.24)' ifadesinde verilen CES üretim fonksiyonu durumunda, eğer $\gamma_{k+1} = 1$ ve doğrusal model doğrusal olmayan modele yerleştirilecekse, model doğrusallığa indirgenir. Yani, $M_1: \gamma_{k+1} = 1$ ve $M_2: \gamma$ kısıtlamasızca mevcut olmuş olacaktır. Doğrusal olmayan regresyon modeli için, eğer γ için olan önsel basit

bir formda ise, Savage-Dickey yoğunluk oranı genellikle kolayca hesaplanabilir. Örneğin, eğer önsel şu formda ise:

$$p(\gamma, h) \propto \frac{p(\gamma)}{h}$$

o zaman γ' nin sonsalı (5.29) ifadesinde verilen form, $p(\gamma)$ olacaktır. Ama eğer önsel daha karmaşıksa o zaman Savage-Dickey yoğunluk oranını kolayca ölçmek imkânsız olabilir.

Sonuçta, iç içe yerleştirilemeyen (non-nested) model karşılaştırmasında veya Savage-Dickey yoğunluk oranının kolayca hesaplanamadığı iç içe geçmiş (nested) model karşılaştırmasında, sonsal olasılıkların bahis oranını hesaplamak için başka, daha genel bir metoda ihtiyaç vardır. Gelfand ve Dey (1994) metodu bu durumlarda kullanılabilir.

Gelfand-Dey metodu, parametre vektörü θ' ya bağlı olan bir M_i modeli için marjinal olabilirliğinin tersinin, belirli bir $g(\cdot)$ seçeneği için $E[g(\theta) | y, M_i]$ olarak yazılabilmesi gerçeğine dayanır. Bununla birlikte, Metropolis-Hastings algoritması gibi, sonsal simülatörleri de bu miktarları kesin olarak tahmin etmek için oluşturulmuştur. Aşağıdaki teorem $g(\cdot)$ ' nin gerekli seçimini sağlar.

Teorem 5.1: Marjinal Olabilirlik Hesaplaması için Gelfond-Dey Metodu

Θ alanı üzerinde tanımlı M_i modeli için $p(\theta | M_i)$, $p(y | \theta, M_i)$ ve $p(\theta | y, M_i)$ sırasıyla önsel, olabilirlik ve sonsalı ifade etsin. Eğer $f(\theta)$, Θ içinde desteklenen herhangi bir o.y.f ise, o zaman,

$$E\left[\frac{f(\theta)}{p(\theta | M_i)p(y | \theta, M_i)} | y, M_i\right] = \frac{1}{p(y | M_i)} \quad (5.42)$$

İspat

$$\begin{aligned} E\left[\frac{f(\theta)}{p(\theta | M_i)p(y | \theta, M_i)} | y, M_i\right] &= \int \frac{f(\theta)}{p(\theta | M_i)p(y | \theta, M_i)} p(y | \theta, M_i) \\ &= \int \frac{f(\theta)}{p(\theta | M_i)p(y | \theta, M_i)} \frac{p(\theta | M_i)p(y | \theta, M_i)}{p(y | M_i)} d\theta \\ &= \frac{1}{p(y | M_i)} \int f(\theta) d\theta = \frac{1}{p(y | M_i)} \end{aligned}$$

Bu teorem dikkate değer şekilde güçlüdür, şöyle ki, herhangi bir o.y.f. $f(\theta)$ için, kolayca (5.43) ifadesi ile kurulabilir:

$$g(\theta) = \frac{f(\theta)}{p(\theta | M_i)p(y | \theta, M_i)} \quad (5.43)$$

ve sonsal simülör çıktısını, herhangi bir sonsal faiz özelliği için olduğu gibi aynı yolla, $E[g(\theta) | y, M_i]$ ' yi tahmin etmek için kullanabilir. Ama Gelfand-Dey metodunun iyi çalışması için, $f(\theta)$ çok dikkatli seçilmelidir. Mümkün olan her θ değeri için, sonlu olmalıdır. Tıpkı etki örnekleminin, önem (importance) fonksiyonunun dikkatli şekilde seçilmesini gerektirdiği gibi, ve Metropolis-Hastings algoritmasının dikkatli bir üreten yoğunluk adayı seçimi gerektirdiği gibi, Gelfand-Dey metodu da $f(\theta)$ ' nin dikkatli seçilmesini gerektirir.

Geweke (1999), $f(\theta)$ seçimi için pratikte iyi çalışan şu stratejiyi önerir. Bu stratejide $f(\cdot)$, kuyrukları kesik Normal yoğunluk fonksiyonu olmalıdır. Genellikle Normal yoğunluk fonksiyonunun kuyruklarında, $\frac{f(\theta)}{p(\theta | M_i)p(y | \theta, M_i)}$ ' nin sonlu olduğunu doğrulamanın zor olması, bu kesmenin gerekçesidir. Kuyruklar kesilerek, problem oluşturma ihtimali olan bölgelerde, $f(\theta)$, sıfıra eşitlenir. Formal olarak, $\hat{\theta}$ ve $\hat{\Sigma}$, $E(\theta | y, M_i)$ ' nin tahminleri olsun ve $\text{var}(\theta | y, M_i)$ sonsal simülörden elde edilmiş olsun. Buna ek olarak, bazı olasılıklar için, $p \in (0,1)$ için Θ , $f(\theta)$ ' in desteği olsun ve şöyle tanımlansın:

$$\hat{\Theta} = \{\theta : (\hat{\theta} - \theta)' \hat{\Sigma}^{-1} (\hat{\theta} - \theta) \leq \chi_{1-p}^2(k)\} \quad (5.44)$$

burada $\chi_{1-p}^2(k)$, k bağımsızlık dereceli Ki-kare dağılımının $(1-p)$ ' inci yüzdesidir ve k , θ ' daki eleman sayısıdır. Sonrasında Geweke (1999), $f(\theta)$ ' nin Θ bölgesinde kesilmiş çok değişkenli Normal yoğunluk olmasını önerir,

$$f(\theta) = \frac{1}{p(2\pi)^{\frac{k}{2}}} |\hat{\Sigma}|^{\frac{1}{2}} \exp\left[-\frac{1}{2}(\hat{\theta} - \theta)' \hat{\Sigma}^{-1} (\hat{\theta} - \theta)\right] 1(\theta \in \hat{\Theta}) \quad (5.45)$$

burada (1), indiktor fonksiyonudur. En iyi çalışma için p ' nin küçük değerlerde olması umulur(örn. $p=0.01$), böylece marjinal olabilirliği tahmin ederken daha çok çıkarım dahil edilebilir. Ama, Geweke (1999)' in dikkat çektiği gibi, p için birçok değişik değer denemenin

ekstra bedeli çok düşüktür. Herhangi bir faizin sonsal özelliğinin, bir sonsal simülator kullanılarak tahmin edildiği gibi, nümerik standart bir hata da standart yoldan hesaplanabilir ve Gelfand-Dey marjinal olabilirlik tahmininin kesinliğini ölçmek için kullanılabilir.

Teorem 5.1' de gösterilen marjinal olabilirliği hesaplamak için, genel Gelfand-Dey metodunun hemen hemen her model için çalıştığı varsayılır. Pratikte, gerekli tek koşul sonsal simülatorün mevcut olmasıdır ve $p(\theta|M_i)$ ile $p(y|\theta,M_i)$ ' nin bilinmesidir. Sonda belirtilenler çözülmesi zor koşullardır. Çünkü bazı durumlarda önsel ve/veya olabilirliğin tüm o.y.f' lerini değil yalnızca kernelleri bilinebilir. Böyle durumlarda, Gelfand-Dey metodu kullanılamaz. Geweke (1999)'in Gelfand-Dey metodu uygulaması sadece sonsalın desteği (5.21) de tanımlanmış bölgeyi (regionı) içeriyorsa çalışır (yani $\hat{\Theta} \in \Theta$). Eğer durum böyle değilse, Geweke (1999) küçük yollarla bu uygulamanın nasıl değiştirilebileceği hakkında bazı tavsiyeler sunar.

5.7.7 Tahmin

Bu modeldeki tahmin edilebilir sonuç, $p(\gamma|y)$ çıkarımları Metropolis-Hastings algoritması tarafından sağlanır ve buna bağlı olarak, herhangi bir tahmini faiz fonksiyonu tahmini elde etmek için, $p(y^*|y, \gamma^{(s)})$ çıkarımları (5.32) ifadesindeki gibi alınabilir ve ortalaması alınır. Örneğin, (4.25) ifadesinde verilen bilgi içermeyen önsel için şu gerçekleştirilebilir:

$$p(y^*|y, \gamma) = f_t(y^* | f(X^*, \gamma), s^{-2}I_T, N)$$

buradaki s^{-2} (5.39) ifadesi ile tanımlanmıştır. Dolayısıyla, γ' ye bağımlı olan, y^* ' ın çıkarımları kolaylıkla elde edilebilir.

5.7.8 Metropolis Hastings Algoritması İçin DeneySEL Çalışma

Doğrusal olmayan regresyon modelinde Bayesçi çıkarım için, mikroekonomik bir uygulama kullanılmıştır. Çıktı y , girdilerin sınıfı x_1 ve belli başlı x_2 ' ye eşlik eden $N=123$ olan bir veri seti kullanılır. Yani, herbir değişken, standart sapmaya bölünür. Artık değişkenlerin standart hatası olduğundan, böyle bir standardizasyon bazen yapılır. Örneğin doğrusal bir regresyon modelinde, bir katsayı, β_j , şu şekilde açıklanabilir: "Eğer açıklayıcı katsayı, j , bir standart hata kadar artarsa, bağımlı değişken β_j , standart sapma kadar artma eğilimi gösterecektir.

(5.24)' de verilen CES formunun bir üretim fonksiyonu varsayalım:

$$y_i = \gamma_1 + (\gamma_2 x_{i1}^{\gamma_4} + \gamma_3 x_{i2}^{\gamma_4})^{\frac{1}{\gamma_4}} + \varepsilon_i$$

Toplanabilir kesim noktası(Additive intercept), basit bir $\gamma_4 = 1$ sınırlamasının regresyon modelini sağlaması göz önüne alınarak seçilir.

(5.27)' de verilen bilgi içermeyen önselleri kullanan, Bağımsız Zincir ve Rasgele Yürüyüş Zinciri Metropolis-Hastings Algoritmasını kullanarak, bu modeldeki sonsal çıkarımlar karşılaştırılarak başlanmıştır. (5.31)' de verilen $\gamma = (\gamma_1, \dots, \gamma_4)'$ için marjinal sonsal üzerinde yoğunlaşılacaktır. Bu iki algoritma için, yoğunluk üreten adayı kurmak ve ilk önce $\hat{\gamma}_{\max}$ ' i bulmak için bir optimizasyon programı kullanılır ($p(\gamma|y)$ ' nin modu). Optimizasyon programı, $\hat{\gamma}_{\max}$ ' in oluşturulmasında, kullanılabilecek olan Hessian için bir tahmin sağlar.

Çizelge 5.1, herbir algoritmanın basit bir uygulaması üzerine kurulmuş sonuçları gösterir. Bu iki algoritmanın, sadece yoğunluk üreten adayları ve dolayısıyla kabul olasılıkları farklıdır. Bağımsız Zincir algoritması $f_t(\gamma | \hat{\gamma}_{\max}, \hat{\text{var}}(\hat{\gamma}_{\max}), 10)$ yoğunluğundan ve Rasgele Yürüyüş Algoritması $f_N(\gamma | \gamma^{(s-1)}, \hat{\text{var}}(\hat{\gamma}_{\max}))$ yoğunluğundan aday çıkarımları alır. (candidate generating density)

Çizelge 5.1 Bağımsızlık zinciri ve rassal yürüyüş yöntemlerinden elde edilen ortalama ve standart sapma değerleri

	Bağımsızlık Zinciri		Rassal Yürüyüş	
	<i>Ortalama</i>	<i>Standart Sapma</i>	<i>Ortalama</i>	<i>Standart Sapma</i>
γ_1	1.04	0.05	1.02	0.06
γ_2	0.73	0.08	0.72	0.12
γ_3	0.97	0.12	1.00	0.16
γ_4	1.33	0.23	1.36	0.29

Örneğin, $\hat{\text{var}}(\hat{\gamma}_{\max})$ ' yi Metropolis-Hastings algoritmasının ilk akışı kullanılarak elde edilen $\text{var}(\gamma|y)$ yaklaşımı ile yer değiştirebilir. Veya her bir aday üretim yoğunluğu yerine

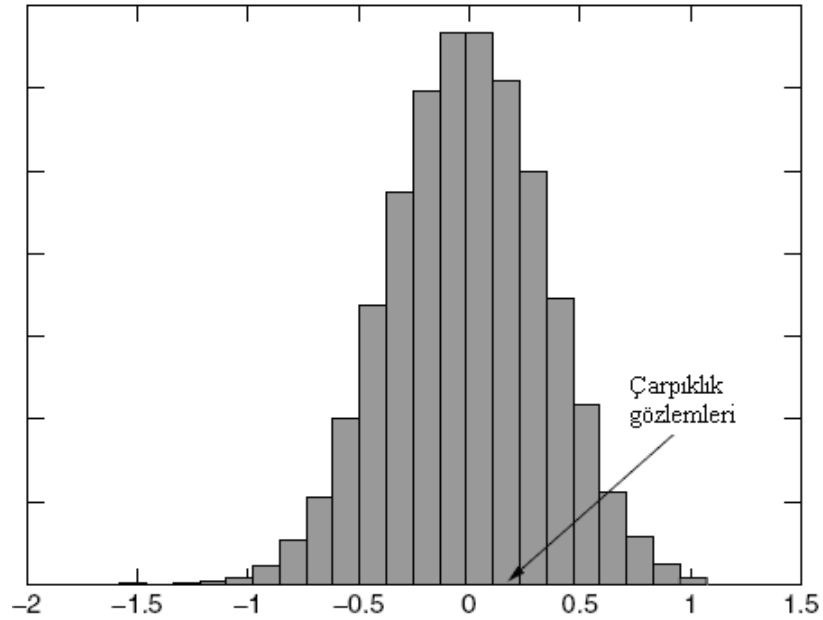
$c \text{ var}(\hat{\gamma}_{\max})$ ' ı kullanılarak ve sonra algoritmanın performansını arttırmak için skaler c' nin çeşitli değerleri ile deney yapılabilir. Bağımsız Zincir Metropolis-Hastings algoritması için, aday üretim yoğunluğunda bağımsız parametre derecesinin değişik değerleri ile deney yapılabilir. Her durumda da $S=25\ 000$ kullanılmıştır.

Bağımsız Zincir ve Rasgele Yürüyüş Algoritmaları için, %7.4 ve %20.6 sırasıyla aday çıkarsamaları kabul edilir. Çizelge 5.1, bu iki farklı algoritmanın basit uygulamalarının kabaca aynı sonucu verdiğini belirtmektedir. Sonsal standart sapmalarla, tahmin edilmiş sonsal ortalamalar birbirlerine çok yakındır. Tahmin edilmiş olan sonsal standart sapmalar, algoritmalar üzerinde farklılık göstermesine karşılık, yine de yakın değerlere sahiptirler. Daha kesin tahminler için, aday üretim yoğunlukları iyi ayarlamalı veya tekrarlamaların sayısı artırılmalıdır.

Çizelge 5.1' deki sonuçlar hatalı, bilgi içermeyen önseller kullanılarak elde edilmiştir. Bu nedenle, kurulmuş modeli bir diğeriyle karşılaştırmak için sonsal olasılık oranı kullanılamaz, bunun yerine, modelin veriye ne kadar uygun olup olmadığını göstermesi için sonsal tahmin p -değerinin kullanılması daha uygun olmuştur. Kısaca, her bir yenilemede, yapay bir veri seti elde edilir. Çarpıklık ve basıklık ölçüleri için (5.40) ve (5.41) ifadeleri kullanılarak hesaplanır. Böylece, bu ölçüler için sonsal tahmin yoğunluğu elde edilebilir. Formüllerle ifade edilmesi gerekirse, $y^\dagger$ rassal olduğundan, rassal değişkenler olan $E[\text{Çarpıklık} \mid y^\dagger]$ ve $E[\text{Basıklık} \mid y^\dagger]$ hesaplanır. Gözlemlenen veri için ayrıca, y rassal olmadığından, rassal değerler olmayan çarpıklık ve basıklık ölçülerini, $E[\text{Çarpıklık} \mid y]$ ve $E[\text{Basıklık} \mid y]$, hesaplanabilir. $E[\text{Çarpıklık} \mid y^\dagger]$ ve $E[\text{Basıklık} \mid y^\dagger]$ ' nin sonsal tahmin yoğunluklarında, $E[\text{Çarpıklık} \mid y]$ ve $E[\text{Basıklık} \mid y]$ ' nin elde edilmesi, modelin veriye ne kadar uygun olduğunu anlamak için bilgi sağlar.

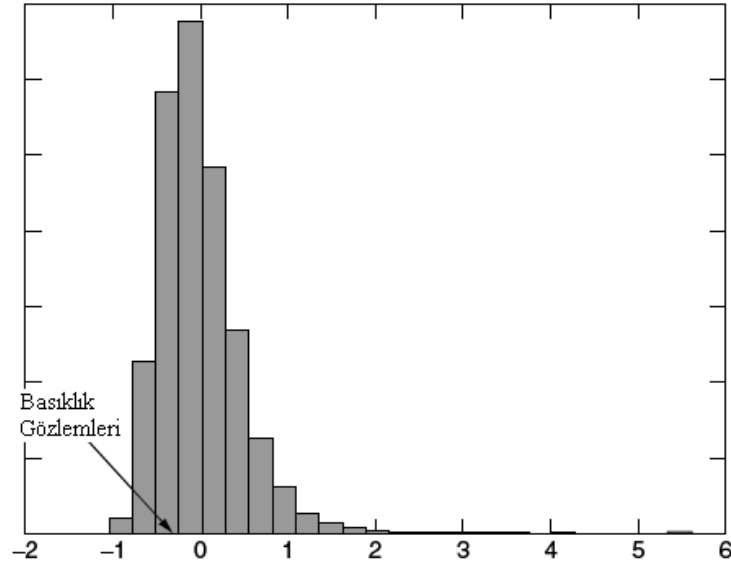
5.1 ve 5.2 şekilleri, bu sonsal tahmin yoğunlukları için yaklaşımları grafikte ifade etmiştir. Ayrıca bu grafiklerde $E[\text{Çarpıklık} \mid y]$ ve $E[\text{Basıklık} \mid y]$ ' yi sırasıyla "gözlemlenmiş çarpıklık" ve "gözlemlenmiş basıklık"olarak belirtilmiştir. Daha net bir ifade ile açıklanacak olunursa, çarpıklık (veya basıklık) için sonsal tahmin yoğunluğu, çalışmakta olunan modelin çarpıklık (veya basıklık) için hangi değeri üretmeye eğilimli olduğunu anlatır. Örneğin, Şekil 5.1 bu deneysel çalışmada kullanılan, doğrusal olmayan regresyon modelinin hemen hemen her zaman çarpıklık ölçüsünün, mutlak değer içinde olandan daha küçük veri setleri ürettiğini ifade etmektedir. Eğer mevcut veri, mutlak değer içinde olandan daha büyük bir çarpıklık değeri ürettiyse, bu durum, modelin kullanılacak bu veri seti ile uyumlu olmadığını gösteren bir kanıttır. Aslında, Şekil 5.1'de çizilen yoğunluğun merkezinin yanında bulunan $E[\text{Çarpıklık}$

$|y|=0.17$ değerine karşılık gelen p-değeri 0.37' dir, yani bu model tarafından üretilmiş olan yapay veri kümesinin %37' si mevcut verinin, ortaya koyduğundan daha büyük çarpıklık derecesi ortaya koyacaktır. Böylelikle, mevcut veri kümesi doğrusal olmayan regresyen modeli ile tutarlı bir çarpıklık derecesi ortaya koyacaktır. Benzer düşünceler, basıklık için sonsal tahmin yoğunluğu için de geçerlidir. Gözlemlenmiş veri bu modelin üretebileceği basıklık ölçüsünün tipleriyle tutarlı olan $E[\text{Basıklık} | y] = -0.37$ basıklık ölçüsünü sağlar. Aslında, sonsal tahmin p-değeri 0.38' dir ve bu, bu modelden üretilmiş yapay verinin %38' inin aşırı basıklık dereceleri olduğunu belirtir. Buna göre, şu sonuca varılır ki, bu deneysel çalışmada kullanılan doğrusal olmayan regresyon modeli veriye iyi uyar (en azından çarpıklık ve basıklık açısından).



Şekil 5.1 Çarpıklık için Sonsal Tahmin Dağılımı

Şimdiye kadar, deneysel çalışma y üzerine yoğunlaştırılmış ve diğer parametre h için birşey söylenmemiştir. Eğer h ile ilgili sonsal bir sonuç elde edilmek isteniyorsa, $h^{(s)}$ çıkarımlarını sağlayan bir sonsal simülatöre ihtiyaç vardır. Dahası, marjinal olabilirlik hesaplamasının Gelfand-Dey metodu (Teorem 5.1'de belirtildiği gibi) $\theta = (\gamma', h)'$ parametre vektörünün tamamının çıkarımlarını gerektirir. Bu sebeplerden, kısaca sırasıyla $p(h | y, \gamma)$ ve $p(\gamma | y, h)$ den elde edilen bir Metropolis-within-Gibbs algoritması türetilmiştir.



Şekil 5.2 Basıklık için Sonsal Tahmin Dağılımı

Doğrusal bir regresyon modelinin uygun olup olmadığı araştırılarak, sonsal bahis oranlarını kullanan model karşılaştırması yorumlanmıştır. Dolayısıyla, koşullu model $M_1 : \gamma_4 = 1$, koşulsuz model ise $M_2 : \gamma_4 \neq 1$ olur. Bu iki modeli karşılaştırılarak sonsal bahis oranını hesaplamak için, ön bir bilgi içeren önsel gereklidir. Herbir model için de bağımsız Normal-Gama önseli kullanılır. M_2 için, γ için önsel; h için olan önselden bağımsızdır ve birincisi

$$\gamma \sim N(\underline{\gamma}, \underline{V})$$

iken, ikincisi şöyledir

$$h \sim G(\underline{s}^2, \underline{v})$$

$\underline{\gamma} = (1, 1, 1, 1)'$, $\underline{V} = 0.25I_4$, $\underline{v} = 12$ ve $\underline{s}^{-2} = 10.0$ seçilir. Verilmiş olan emek ve kapital marjinal ürünlerinin, muhtemel miktarları ve veriyi standart hale getirme yoluna göre bu seçimler mantıklıdır, ama göreceli olarak bilgi içermez.

M_1 için aynı önseli kullanılır, artık sadece üç elemanı olan y hariçtir ve bu nedenle $\underline{\gamma}$ ve $\underline{V}$ son sıraya sahiptirler. Sıra ve kolon sırasıyla dahil edilmez.

Bu önsel ile doğrusal olmayan regresyon modeli için, bir Metropolis-within-Gibbs algoritması üretilir. Böyle bir algoritmayı kurmak için, ilk önce $p(h | y, \gamma)$ ve $p(\gamma | y, h)$ türetilmelidir:

$$h | y, \gamma \sim G(\bar{s}^{-2}, \bar{v}) \quad (5.46)$$

Burada

$$\underline{v} = N + \underline{v}$$

ve

$$\underline{s}^{-2} = \frac{[y - f(X, \gamma)]' [y - f(X, \gamma)] + \underline{v} s^2}{\underline{v}}$$

(5.28)' i kullanarak ve $p(\gamma | y, h) \propto p(\gamma, h | y)$ ' a dikkat ederek, görülebilir ki

$$p(\gamma | y, h) \propto \exp \left[-\frac{h}{2} \{y - f(X, \gamma)\}' \{y - f(X, \gamma)\} \right] \exp \left[-\frac{1}{2} (\gamma - \underline{\gamma})' \underline{V}^{-1} (\gamma - \underline{\gamma}) \right] \quad (5.47)$$

Bu koşullu sonsal yoğunluk, basit bir ifade ile ouşturulacak yapıya sahip değildir. Ama $p(\gamma | y, h)$ için bir Metropolis-Hastings algoritması kullanılabilir, bu Gama dağılımının çıkarsamaları ile birleşerek Metropolis-within-Gibbs algoritmasını sağlar. Bu deneysel uygulamada, (5.47) ifadesinden elde etmek için bir Rassal Yol Zincir Metropolis-Hastings algoritması kullanılır. Bu algoritma Çizelge 5.1' in son iki kolonundaki sayıları yaratmak için kullanılan algoritma ile özdeştir, yalnız kabul edilebilir olasılığı (5.47) ifadesi kullanılarak hesaplanmıştır.

M_1 için geliştirilenler gibi, sonsal simülatörlerden gelen çıktılar ve M_2 verilmiş olsun, Gelfand ve Dey kullanarak marjinal olabilirliği hesaplamak daha basittir. Sonsal simülatorun çıkarımları bir kez elde edildiğinde, bu çıkarımlarda olabilirliği ve önseli hesaplamak gerekir. Herbir model için de önseller bağımsız Normal-Gama olduğu için önsel ölçüm, bileşeni oluşturması kolaydır. Önsel, olabilirlik ve $f(\cdot)$, herbir sonsal çıkarımda bir kez hesaplandığında, $g(\cdot)$ hesaplanabilir ve marjinal olabilirliğin tahminini sağlaması için sonucun ortalaması alınabilir.

İki sonsal simülatör (yani doğrusal model için Gibbs örneği ve doğrusal olmayan model için Metropolis-within-Gibbs), tekrar için başlangıç değeri $S_0=2500$ ve yinelenen cevap $S_1=25000$ kullanılarak çalıştırılmıştır. Marjinal olabilirlikler $p=0.01, 0.05$ ve 0.10 kullanılarak hesaplanmıştır. Bu üç kesim değeri için Bayes faktörlerinin tahmin sonuçları sırasıyla 1.067, 1.075 ve 1.065'dir. Bu üç tahminin birbirine yakın olmaları, sonsal simülatörlerin güvenilir sonuçlar verdiğine dair formal olmayan kanıt sağlamaktadır.

Bayes faktörleri birbirine çok yakındır, bu doğrusal olan ve olmayan regresyon modellerinin eşit destek aldığını belirtir. γ_4 'ün sonsal ortalaması bir standart sapmadan 1.0 kadar büyük olduğunda, bu sonuç oldukça mantıklıdır.

5.8 Büyüme Modellerindeki Parametrelerin Analizinde Bayesçi Yöntem

Örneğin orman büyüme modellerinde, büyük ölçüde orman ekosisteminin dinamik modellenmesi ve özel ağaçların gelişiminin simülasyonu kullanılır. Model tahmininin güvenilirliğini tanımlamak için, modelin birkaç defa koşturulması gerekir. Birçok değerlendirme için, modelin birkaç defa tekrar edilerek denenmesi, modeldeki parametrelerin korelasyonlarının ve dağılımlarının tahmini ile gerçek modelin (gerçeğe yakın) bulunmasını sağlayacağına dair bir sonuç elde edilmiştir.

Orman büyüme modelleri için, kullanılan büyüme modellerinin çoğu doğrusal değildir. Doğrusal olmayan modellerdeki parametrelerin tahmininde kullanılan tahmin yöntemleri için, asimptotik çoklu normal dağılımların parametreleri ile oluşturulduğu farzedilir. Fakat birçok doğrusal olmayan modelle, parametrelerin birikimli dağılımları, çoklu normal dağılımdan farklı olabilir. Çoklu normal dağılımlar, modellerin gerçek dağılımlarını yansıtmayabilirler.

Yakın zamanda, Bayesçi yöntemlerle hesaplama, doğrusal ve doğrusal olmayan modeller için geliştirilmiştir. Bayesçi Red yöntemi, Bayesçi yöntemlerin hesaplamalarından biridir. Doğrusal olmayan modellerin parametrelerinin, birikimli dağılımlarının tahminlerinin sağlanması mümkündür. Bayesçi Red (rejection) Yönteminin, doğrusal olmayan modellerin tahmininde kullanılan geleneksel yöntemlerden daha güçlü olmasının, en önemli avantajlarından birisi budur.

5.8.1 Büyüme Modelleri, Yöntemler ve Malzemeler

Morgan- Mercer- Flodin ailesinin fonksiyonu ve Richard fonksiyonu doğrusal olmayan sigmoid büyüme fonksiyonudur. Morgan-Mercer- Flodin ailesinin fonksiyonu (MMF) genel formu;

$$y = \frac{\beta\gamma + \alpha x^\delta}{\gamma + x^\delta} \quad (5.48)$$

Bu sigmoid büyüme modeli, bitkilerin büyüme modellerinde geniş ölçüde kullanılır. Modelde x değişkeni, bağımsız değişken ve y değişkeni, bağımlı değişkeni ifade eder. Modelde 4 parametre bulunmaktadır: α , β , γ ve δ

Richards büyüme modelinin formu,

$$y = A(1 - b \bullet e^{-Kt})^{1/(1-M)} \quad (5.49)$$

Bağımlı ve bağımsız değişkenler sırasıyla, y ve t dir. Richards fonksiyonunda da dört parametre vardır: A , b , K ve M . Çoğunlukla b parametresi 1' in analitik çözümüdür. Bu araştırma için, b parametresi 1' in bir sabiti olarak farzedilir.

2 tahmin yöntemi de her bir model için, farklı parametre tahminleri varyans ve korelasyon matrislerinin oluşturulmasını şart koşar.

Modellerin parametrelerinin tahmininde 2 farklı yöntem uygulanmasına rağmen, modellerin tahmin sonuçları benzer değerler elde edilebileceği, şimdiye kadar yapılmış analizlerde gözlemlenmiştir. Morgan-Mercer-Flodin fonksiyonu için, iki yöntemde de kullanılmış olan β ve γ parametrelerinin, varyansı ve parametre tahminleri, nispeten büyük ölçüde farklı olabilir. Doğrusal olmayan regresyon kullanılırken β ve γ parametrelerinin değişim dereceleri (CV) değerlendirilir. Daha sonra bu parametrelerin ortalaması ve varyansı, değerlendirmeye tabi tutulur. MMF' nin tüm parametrelerinin varyansı için, doğrusal olmayan regresyon kullanılarak büyük tahminler oluşturulabilir. Doğrusal olmayan regresyonda, tahmin için kullanılan MMF' nin korelasyon matrisi ile parametreler arasındaki bağlantı kontrol edilir. Bu şekilde Bayesçi Red yöntemi ile değerlendirme yapılmış olmaktadır.

5.9 Doğrusal Olmayan Bayesçi Regresyon İçin Genel Çıkarsama

Doğrusal olmayan regresyon modelinde değerlendirme yapmak için Bayesçi yaklaşımlar anlatılmıştır. Doğrusal olmayan yaklaşım, basit bilgi içermeyen önseller kullanılsa dahi, analitik sonuçların mevcut olmadığı anlamına gelir. Ayrıca genellikle sonsal simülasyon için, bir Gibbs örneği geliştirmek mümkün değildir. Bu özellikler, sonsal simülasyon için çok genel algoritmalar sınıfına bir giriş yapmak amacıyla kullanılır: Metropolis-Hastings sınıfı algoritmalar. En çok kullanılan değişkenlerden Bağımsız Zincir ve Rassal Yol Zincir Metropolis-Hastings algoritmaları detaylarıyla anlatılmıştır. Doğrusal olmayan regresyon modelini, model veya modelleri ölçmek için yeni bir metot olan sonsal tahmin p-değerini sunmak için bir taslak olarak kullanıldı. Bu bir modelin uygunluğunu ölçmek için genel bir araçtır. Marjinal olabilirlik hesaplaması için genel bir araç olan Gelfand-Dey metodu da anlatılmıştır.

Sonsal simülatörler açısından, Monte Carlo integrasyonunu, önem örnekleme, Gibbs

örnekleme ve Metropolis-Hastings algoritması anlatılmıştır. Model karşılaştırması için, marjinal olabilirlik hesaplamasında Savage-Dickey yoğunluk oranı gibi kolay, ama daha kapsamlı olan metodların yanı sıra Gelfand-Dey metodu anlatılmıştır. Sonsal tahmin p-değeri yaklaşımı da tanıtılmıştır. Bu yaklaşım en iyi model uygunluğunu ölçmek için bir yol olarak açıklanabilir, ama ayrıca modelleri karşılaştırmak için de kullanılabilir. Artık geniş çeşitlilikte araçlara sahip ve bazı modeller için birden fazla araç kullanılabilir. Örneğin doğrusal olmayan regresyon modelinde önem örneklemini veya bir Metropolis-Hastings algoritmasını kullanmak mümkündür. Model karşılaştırması için sonsal bahis oranı veya sonsal tahmin p-değeri kullanılabilir. Yapılan seçimler, verilen model için tek (veya hatta en iyi) çözüm olmak zorunda değildir.

Son olarak, Büyüme modellerinin parametrelerinin, marjinal ve birikimli dağılımlarının tahmininde kullanılan, doğrusal olmayan regresyon ve sonuçlar özetlenmiştir.

6. UYGULAMA

Uygulamaya öncelikle deneysel verilerin analizi ile başlanacak. Deneysel veriler, Excel, Mathcad, MATLAB ve XLStat programları kullanılarak analiz edilmiştir. Biri bağımlı, biri bağımsız olmak üzere, 2 değişkenli, bir regresyon modeli oluşturulmuş, yapılmış olunan çalışma sonucunda elde edilen değerlerin, istatistiksel olarak analizi gerçekleştirilmiştir. Bu sonuçlar, teorik olarak ifade edilmek için diğer subjektif analiz ile birleştirilerek, elde edilen değerler yorumlanır. Önceki benzer araştırmalar ve kişisel deneyimlerde çıkan sonuçlar, analizin yapılabildiği alanların, kendilerine olan kısıtlamaları ve örneklem bilgisi, Bayes yaklaşımı ile analiz edilerek parametreler dahil edildiği ve bu sayede bütün bilgilerin kullanılabileceği sonuçlara ulaşılması sağlanmıştır. Böylece, oluşturulması hedeflenen son olasılık yoğunluk fonksiyonuna ulaşılır. Bu sayede sonsal olasılık yoğunluk fonksiyonu, parametrelerine ait bilgilerin oluşmasını sağlayarak, hem önbilgiyi (önseli) hem de örnek bilgisini birleştirir. Bu yaklaşım sayesinde kişilerden elde edilmiş olan önbilginin, daha net ve doğru sonuçların oluşmasındaki bilimsel etkisi vurgulanmıştır. Normal dağılıma sahip örneklem bilgisi, ön bilgi ile birleştirildiğinde sonsal olasılık yoğunluk fonksiyonu da normal dağılıma uyacaktır. Çalışmada normal dağılıma uyan örneklem bilgisi kullanılarak yüksek frekanslı sesleri etkileyen parametrelerin analizi yapılmıştır.

6.1 Veriler ve Ön Bilgilendirme

NIST kurumunun yapmış olduğu Yüksek frekanslı seslerin verilerini içeren, bir çalışmadan elde edilmiştir. <http://www.itl.nist.gov/div898/strd/nls/data/chwirut2.shtml> adresinden bu verilere ulaşılabilir. Bağımlı değişken yüksek frekans ölçümleri ve bağımsız değişken (tahminci) metal uzaklıklarıdır.

y = yüksek frekans cevabı, bağımlı değişken; x = metal uzaklığı, bağımsız değişken olmak üzere 2 değişkenden oluşan bir model oluşturulacaktır. Deneylerden elde edilmiş 54 gözlem bulunmaktadır. Doğrusal olmayan bu veriler üzerinde en küçük kareler yaklaşımı ile Doğrusal Olmayan Bayesçi regresyon yaklaşımı olmak üzere iki yöntemde değerlendirme yapılacaktır. Yüksek frekanslı seslere ait uygulamada kullanılacak olan veriler EKLER bölümünde Çizelge Ek 1' de yer almaktadır.

Yüksek frekanslı seslerin etkileri, hem deneysel veri aynı zamanda subjektif verilere dayalı olarak da değerlendirilebilir. Tezde ilgilenilmiş olunan Bayes teorisi, alt yapı bakımıyla birbiri ile örtüşmektedir. Yüksek frekanslı sesleri, metal uzaklığının mı etkilediğini, klasik

istatistiksel metotlarla da analiz ederek, kişisel verilere dayandırılarak yapılmış olunan analizleri de değerlendirmeye tabi tutup, Bayesçi yaklaşımla analiz edilecektir.

6.2 Doğrusal Olmayan Bayesçi Regresyon Yaklaşımı

Öncelikli olarak uygulamada kullanılması gereken örnekleme ve önsel dağılımların matris formları ifade edilmelidir. Aşağıdaki gibi özetlenebilir:

1. Adım: Örnekleme Dağılımı aşağıdaki gibi ifade edilebilir;

$$Y_{ij} \sim N(\alpha_i + \beta_i x_{ij}, \sigma^2), \quad i=1, \dots, k=30 \text{ ve } j=1, \dots, n=5$$

2. Adım: Önsel Dağılım'ın matris formu aşağıdaki gibidir;

$$\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} \sim N \left(\begin{bmatrix} \alpha_0 \\ \beta_0 \end{bmatrix} \Sigma \right), \quad i=1, \dots, k$$

3. Adım: Hiper önsel' in matris formu aşağıdaki gibidir;

$$\sigma^2 \sim IG(a, b)$$

$$\begin{pmatrix} \alpha_0 \\ \beta_0 \end{pmatrix} \sim N \left(\begin{bmatrix} \eta_0 \\ \eta_0 \end{bmatrix} C \right)$$

$$\Sigma^{-1} \sim W((\rho R)^{-1}, \rho), \quad E(\Sigma^{-1}) = R^{-1} \text{ ve } \text{var}(\Sigma) \propto \rho^{-1}$$

3 adımda örnekleme, önsel ve hiper önsel yapıların matris formları ifade edilmiştir. Doğrusal olmayan regresyon modellerinde, Bayesçi yöntemin uygulanabilirliğini gözlemlemek için, Metropolis-Hastings Algoritması'ndan faydalanılmalıdır. Bu algoritmanın uygulanmasındaki amaç, $p(\mu, \sigma, m_1 | \text{veriler})$ yaklaşık olasılığını elde etmektir.

Bu algoritma için önsel dağılımlar;

$$m_1 \sim \text{Gama}(a_0, b_0)$$

$$\mu \sim N(c_0, d_0)$$

$$\sigma^2 \sim IG(e_0, f_0)$$

Önsel için olabilirlik fonksiyonları;

$$p(\mu, \sigma, m_1 | \text{veriler}) \propto f(\text{veriler} | \mu, \sigma, m_1) \pi(\mu, \sigma, m_1)$$

$$\propto \left\{ \prod_{i=1}^k [h(w_i)]^{y_i} [1-h(w_i)]^{n_i-y_i} \right\} \frac{m_1^{a_0-1}}{\sigma^{2(e_0+1)}} \\ \times \exp \left[-\frac{1}{2} \left(\frac{\mu - c_0}{d_0} \right)^2 - \frac{m_1}{b_0} - \frac{1}{f_0 \sigma^2} \right]$$

Sonsal dağılımda parametreler değişken olarak kullanılır. Bu durumda;

$$p(\theta | veri) \propto \left\{ \prod_{i=1}^k [h(w_i)]^{y_i} [1-h(w_i)]^{n_i-y_i} \right\} \exp(a_0 \theta_3 - 2e_0 \theta_2) \\ \times \exp \left[-\frac{1}{2} \left(\frac{\theta_0 - c_0}{d_0} \right)^2 - \frac{\exp(\theta_3)}{b_0} - \frac{\exp(-2\theta_2)}{f_0} \right]$$

olarak ifade edilir.

Uygulama için önsel dağılımın, Gama dağılımına sahip olduğu varsayılmış, ancak daha sonra normal dağılım ile çalışmanın daha uygun olduğu görülmüştür.

6.3 Model ve Varsayımlar

Bayesçi regresyon ile doğru örnekleme oluşturmak için 200 iterasyon gerçekleştirilerek bir uygulama oluşturulmuştur. Bu iterasyonlar sonucunda yakınsaklık değeri 0,00001 elde edilmiştir. Gerçekleştirilen analizler sonucunda, en uygun model olarak, N boyutlu gözlem vektörü y için yüksek frekanslı ses modeli şöyledir:

$$Y = \alpha_1 + \alpha_2 \exp(\alpha_3 X) + \varepsilon \quad (6.1)$$

Burada α_1 , g boyutlu sabit etkiler vektörü, X, kovaryetelerin merkezi değerini içeren $N \times \alpha_2$ boyutlu bir matris α_2 ve α_3 , kovaryetelere ait α boyutlu regresyon katsayıları vektörü ve ε , N boyutlu şansa bağlı hatalar vektörüdür. Yukarıdaki (6.1) ifadesi ile oluşturulmuş modelde N tane frekans ölçümü olduğu ve bunların toplam 54 gözlemden elde edildiği varsayılmaktadır. Yüksek frekanslı sesleri aşırı ölçüde etkileyen harici unsur olmadığı ve metal uzunluklarından etkilendiği varsayılmıştır. Diğer varsayımlar ise şöyledir:

$$E(y | \alpha_1, \alpha_2, \alpha_3, \sigma_x^2, \sigma_\varepsilon^2) = \alpha_1 + \alpha_2 \exp(\alpha_3 X),$$

$$E(X) = 0, E(\varepsilon) = 0,$$

$$Var(X | \sigma_x^2) = G = I_x \sigma_x^2,$$

$$\text{Var}(\varepsilon \mid \sigma_\varepsilon^2) = R = I_N \sigma_\varepsilon^2$$

$$\text{Var}(y \mid \alpha_1, \alpha_2, \alpha_3, \sigma_x^2 \sigma_\varepsilon^2) = I_N \sigma_\varepsilon^2$$

$$\text{Cov}(X, \varepsilon \mid \sigma_x^2, \sigma_\varepsilon^2) = 0$$

$$E = \begin{bmatrix} \mathbf{Y} \\ \mathbf{X} \\ \varepsilon \end{bmatrix} = \begin{bmatrix} \alpha_1 + \alpha_2 \exp(\alpha_3 X) \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}$$

6.4 Bazı Kısıtlamalar

Bazı uygulamalarda, varyans unsurları yerine bunların oranları veya fonksiyonları hakkında yorumlamaların yapılması gerekebilir. γ , metal uzunlukları ve hata varyans unsurlarının oranı,

$\frac{\sigma_x^2}{\sigma_\varepsilon^2}$, olarak ifade edilsin. Bu durumda, metal uzunluklarının derecesi, k_2 , γ 'nin artan bir

fonksiyonu olup $k_2 = 4/(1 + \gamma - 1)$ şeklinde verilir. γ varyansların bir oranı olduğundan,

pozitifdir. Ayrıca, (1)'deki yüksek frekanslı ses modelinin kullanımı γ üzerine bir üst sınır

konulmasını zorunlu kılar, bu w ile gösterilsin. Böylece $0 \leq \gamma \leq w$ veya $0 \leq \sigma_x^2 \leq w \sigma_\varepsilon^2$ sınırları

elde edilir. Oran olması nedeniyle, metal uzunluklarının dereceleri 0 ile 1 arasında olmalıdır.

Bu kısıtlamayı dikkate almayan bir uygulama, metal uzunlukları derecesi tahminlerinin kendi parametre sınırları dışında değerler almasına yol açabilir.

6.5 Bayesçi Formüller

6.5.1 Önsel Dağılım

Modeli eksiksiz olarak tanımlamak için, ek varsayımlarda bulunmak gerekir. Böylece varyans değerlerinin tahmini, daha doğru bir şekilde ifade edilecektir. Öncelikle, modeldeki $\alpha_1, \alpha_2, \alpha_3$

σ_x^2 ve σ_ε^2 parametreleri için, önsel bir dağılım kabul edilmesi zorunludur. Bu modeldeki α_1

α_2 ve α_3 parametreleri için düzgün önsel dağılım olduğu kabul edilecektir.

$$f(\alpha_1, \alpha_2, \alpha_3) = \text{sabit} \quad (6.2)$$

Bu ifadenin anlamı, başlangıç değerlerinin dağılımı hakkında çok fazla bilgiye sahip olunmadığıdır. Bir sonraki adım olarak, X değerlerinin, normal dağılıma uyduğu varsayılır.

$$X | \sigma_X^2 \sim N_X(0, I_X \sigma_X^2) \quad (6.3)$$

Varyans değerleri, σ_X^2 ve σ_ε^2 'nin önsel dağılımlarının birleşik yoğunluk fonksiyonunun aşağıdaki gibi olduğu varsayılmaktadır:

$$f(\sigma_X^2, \sigma_\varepsilon^2 | v_X, s_X^2, v_\varepsilon, s_\varepsilon^2) \propto (\sigma_X^2)^{-\frac{1}{2}(v_X+2)} (\sigma_\varepsilon^2)^{-\frac{1}{2}(v_\varepsilon+2)} \exp\left[-\frac{1}{2}\left(\frac{v_X s_X^2}{\sigma_X^2} + \frac{v_\varepsilon s_\varepsilon^2}{\sigma_\varepsilon^2}\right)\right] \quad (6.4)$$

Bu fonksiyondan σ_X^2 ve σ_ε^2 'nin birbirinden bağımsız ve ters χ^2 dağılımları, $\chi^2(v_X, v_X s_X^2)$ ve $\chi^2(v_\varepsilon, v_\varepsilon s_\varepsilon^2)$, izledikleri anlaşılmaktadır. Bu nedenle s_X^{-2} ve s_ε^{-2} , sırasıyla σ_X^{-2} ve σ_ε^{-2} 'nin önsel beklenen değerleri olarak değerlendirilebilirler, bunun yanısıra v_X ve v_ε serbestlik derecesine eşdeğer doğruluk (precision) parametreleridir.

Diğer dağılımların kullanılması halinde, Gibbs örnekleme ile varyans unsurlarının çözümlenmesi karmaşık hale gelecektir. Şimdiye kadar gerçekleştirilmiş Gibbs örnekleme çalışmalarında genellikle düzgün dağılım kullanılmıştır (Wang ve ark., 1994; Jensen ve ark., 1994). Ancak bu durumda, Hobert (1994)' in belirtmiş olduğu gibi, varyans unsurları için düzgün dağılımlar kullanıldığında birleşik dağılım her zaman oluşturulamaz ve Gibbs örneklemesinden elde edilen sonuçların güvenilirliği azalmaktadır. Oluşturulan yüksek frekanslı sesler modeli doğrusal olmayan bir yapı izlediğinden Gibbs örnekleme kullanılamayacaktır.

6.5.2 Olabilirlik Fonksiyonu

Hata değerlerinin normal dağıldığı varsayımı ile aşağıdaki olabilirlik fonksiyonu oluşturulur.

$$f(Y | \alpha_1, \alpha_2, \alpha_3, X, \sigma_\varepsilon^2) \propto (\sigma_\varepsilon^2)^{-\frac{1}{2}N} \exp\left\{-\frac{1}{2\sigma_\varepsilon^2} \left[(Y - \alpha_1 - \alpha_2 \exp(\alpha_3 X))' (Y - \alpha_1 - \alpha_2 \exp(\alpha_3 X))\right]\right\} \quad (6.5)$$

Yukarıda (6.2)-(6.5) eşitliklerinde verilen varsayımlar, olabilirliğe dayalı bir çok yöntemde kullanılanlarla aynıdır.

6.5.3 Birleşik Sonsal Yoğunluk Fonksiyonu

Birleşik sonsal olasılık yoğunluk fonksiyonu, yukarıda (6.2)-(6.5)' de belirtilen önsel ve şartlı dağılımların çarpımı olarak yazılabilir. Gözlemler ve önsel bilgi verildiğinde $\alpha_1, \alpha_2, \sigma_X^2$ ve

σ_ε^2 parametrelerinin birleşik yoğunluk fonksiyonu aşağıdaki gibi verilir:

$$f(\alpha_1, \alpha_2, \alpha_3, X, \sigma_X^2, \sigma_\varepsilon^2 | Y, v_X, s_X^2, v_\varepsilon, s_\varepsilon^2) = F$$

$$F \propto f(Y | \alpha_1, \alpha_2, \alpha_3, X, \sigma_\varepsilon^2) f(\alpha_1, \alpha_2, \alpha_3) f(X | \sigma_X^2) f(\sigma_X^2, \sigma_\varepsilon^2 | v_X, s_X^2, v_\varepsilon, s_\varepsilon^2)$$

$$\propto (\sigma_\varepsilon^2)^{-\frac{1}{2}(N+v_\varepsilon+2)} (\sigma_X^2)^{-\frac{1}{2}(X+v_X+2)} \exp\left[-\frac{1}{2\sigma_X^2}(X^2 + v_X s_X^2)\right]$$

$$\times \exp\left\{-\frac{1}{2\sigma_\varepsilon^2}\left[(Y - \alpha_1 - \alpha_2 \exp(\alpha_3 X))^2 + v_\varepsilon s_\varepsilon^2\right]\right\} \quad (6.6)$$

Önsel dağılım, olabilirlik ve sonsal dağılım oluşturulduktan sonra verilerin bu fonksiyon ve dağılımlara göre analizi gerçekleştirilecektir.

6.6 Gerçekleştirilen Analizler

6.6.1 Normal Dağılıma Uygunluk Kontrolü

Dağılımın normal dağılıma uyup uymadığını kontrol etmek için Kolmogorov-Smirnov uygunluk testi yapılmıştır:

Çizelge 6.1 Kolmogorov-Smirnov test istatistikleri

D	0,169
p-değeri	0.082
Alfa	0,05

Normal dağıldığı varsayımı altında, önsel dağılımın belirli istatistiksel değerleri;

Çizelge 6.2 Önsel dağılıma ait test istatistikleri olur.

İstatistik	Veri	Parametreler
Ortalama	2365,921	2365,921
Varyans	3680056,477	3680056,477
Çarpıklık	0,274	0,000
Basıklık	-1,002	0,000

Dağılımın üssel dağılıma uyup uymadığını kontrol etmek için Kolmogorov-Smirnov uygunluk testi yapılmıştır:

Çizelge 6.3 Kolmogorov-Smirnov test istatistikleri

D	0,151
p-değeri	0.065
Alfa	0,0567

Üssel dağıldığı varsayımı altında, önsel dağılımın belirli istatistiksel değerleri;

Çizelge 6.4 Önsel dağılıma ait test istatistikleri olur.

İstatistik	Veri	Parametreler
Ortalama	2254,942	2254,942
Varyans	3680056,477	3680056,477
Çarpıklık	0,256	0,000
Basıklık	-1,017	0,000

6.6.2 Otokorelasyon Kontrolü

Durbin-Watson istatistiği = 2.269133

6.6.3 Önsel Dağılıma Ait Değerler

Çizelge 6.5 Önsel dağılıma ait genel değerler

Değişken	Gözlemler	Eksik Veri	Mevcut Veri	Minimum	Maksimu	Ortalama
Y	54	0	54	37500,000	929000,00	315640,741
X	54	0	54	0,500	6000,000	2365,921

Çizelge 6.6 Y değişkeni için doğrusal olmayan regresyon değerleri

Gözlemler	54
Serbestlik Derecesi	51
R ²	0.937
SSE	232301866702,577
MSE	4554938562,796
RMSE	67490,285

Çizelge 6.7 Önsel Ortalama Değerleri

315640.7	2340.921
----------	----------

Çizelge 6.8 Önsel Standart Sapma Değerleri

263130,261	1918,347
------------	----------

6.6.4 EKK Yöntemi ile Elde Edilen Parametre Tahminleri (DeneySEL Veri)

Çizelge 6.9 EKK regresyon değerleri

Hata Kareler Toplamı	5.1304802941E+02
Hataların Standart Sapması	3.1717133040E+00
Serbestlik Derecesi	51
Gözlem Sayısı	54

Çizelge 6.10 Ekk yöntemi ile elde edilen parametre tahminleri

	Oluşturulmuş Değerler	
	Parametre	Standart Sapma
β_1	304067,925	1908,1916
β_2	217500501200150	140015,275
β_3	-164,564	7,35

6.6.5 Sonsal Dağılıma Ait Değerler

Çizelge 6.11 Sonsal ortalama değeri

304067,925	2410,552
------------	----------

Çizelge 6.12 Sonsal parametre değerleri

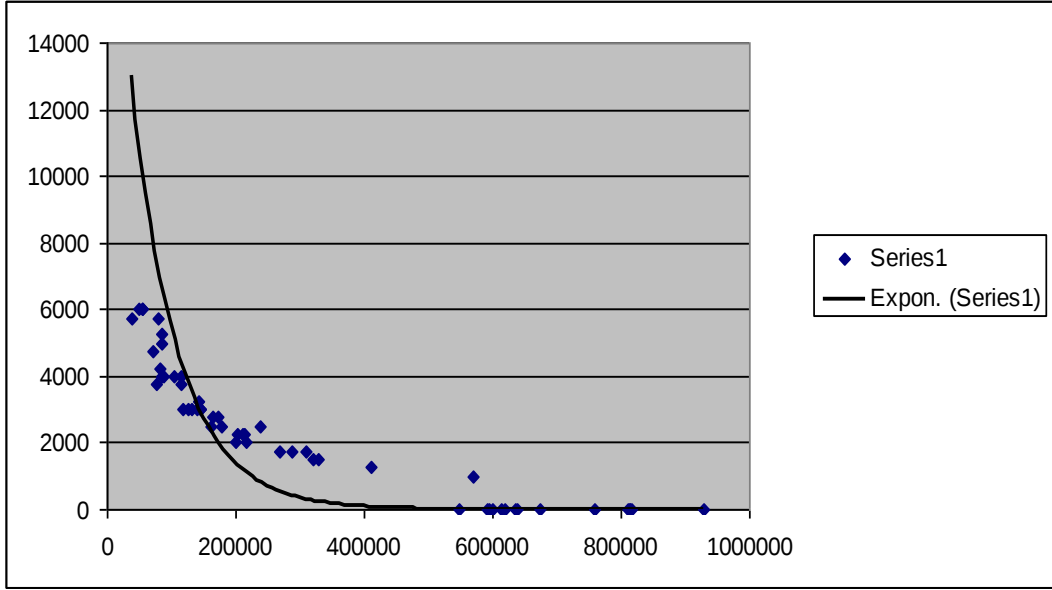
Parametre	Değer	Standart Hata
α_1	252035.092	1679,198
α_2	177125424991826	131200,972
α_3	-39.176	6,78

Elde edilen sonsal parametre değerleri ile model aşağıdaki gibi olur:

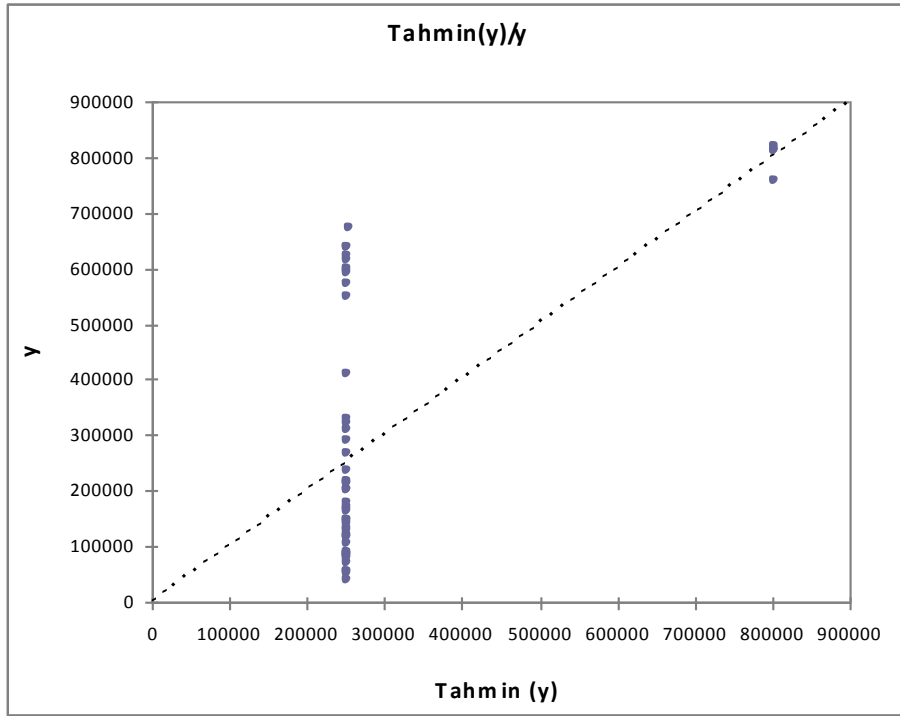
$$Y = 252035.092 + 177125424991826 \cdot \text{Exp}(-39.176 \cdot X)$$

Çizelge 6.13 Sonsal parametre değerleri ile oluşturulan modele ait bazı istatistiksel değerler

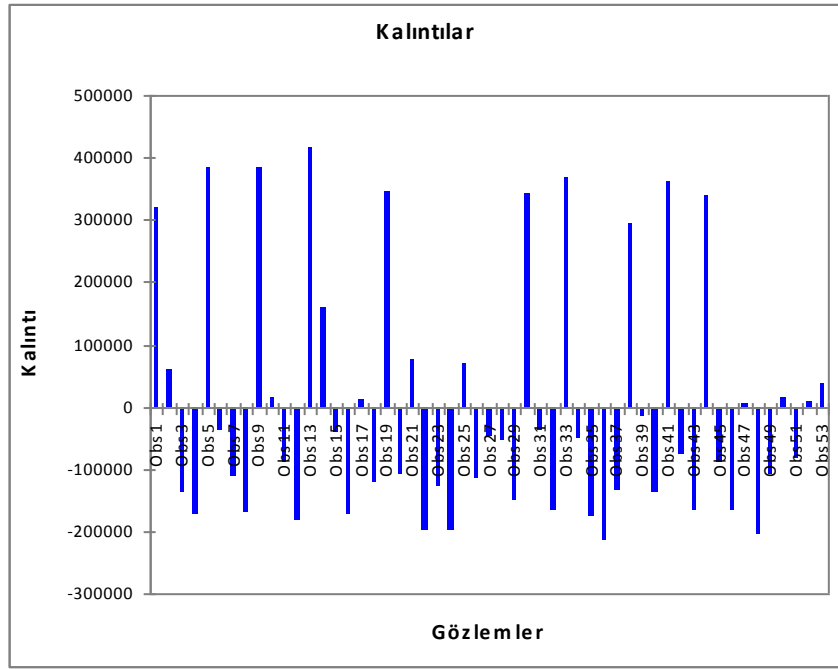
Çarpıklık Katsayısı	0,251
Basıklık Katsayısı	-0,992
p – değeri	0,001



Şekil 6.1 Y için Doğrusal Olmayan Regresyon



Şekil 6.2 Y değerinin tahmincisi



Şekil 6.3 Modelin kalıntıları

6.7 Modelin Yorumlanması

Kullanılacak modelin hangi dağılıma daha uygun olduğuna dair bir fikre sahip olmak için Kolmogorov-Smirnov ve Ki-kare testleri gibi uygunluk testleri yapılmıştır. Bu testler sonucunda, oluşan p değeri α değerinden büyük olduğu için, H_0 =Normal Dağılıma Uyar, hipotezi kabul edilir. Bunun yanısıra Jarque -Bera test istatistiği de değerlendirmeye tabii tutulmuş, elde edilen değer 2.643872 bulunmuştur. 0,05 anlamlılık düzeyinde bu değer kritik değeri aşmadığı için Normal dağılıma uyduğu kabul edilir. Bu iki test sonucunda, gerçekleştirilen işlemlerin hepsi, değişkenlerin ve parametrelerin Normal Dağılıma uyduğu varsayımı altında gerçekleştirilmiştir. Bu veri setini tecrübe sahibi olan insanların düşüncelerine ve daha önceki deneyimlerine dayanılarak yapılmıştır.

Bunun yanısıra otokorelasyon testinde Durbin Watson test istatistiği 2.269133 olarak hesaplanmıştır. Bu değer, N=54 için çizelge değerleri ile karşılaştırıldığında otokorelasyonun olmadığına karar verilmiştir.

Şimdi de değişen varyans olup olmadığına karar verilir. Bunun için White testine karar verilmiştir. $0.11 \cdot 54 = 5.94 < 24,882$ çizelge değeri hesaplanan değer çizelge değerinden küçük olduğundan sabit varyans varsayımı geçerlidir.

7. SONUÇ VE ÖNERİLER

Doğrusal olmayan yaklaşım, basit bilgi içermeyen önseller kullanılsa da, analitik sonuçların mevcut olmadığı anlamına gelmektedir. Doğrusal olmayan regresyon modelini ölçmek için yeni bir metot olan sonsal tahmin p-değeri kullanılmıştır. p değeri ile modelin uygunluğu kontrol edilebilir. Bu yaklaşım ayrıca, modelleri karşılaştırmak için de kullanılabilir. Ancak yapılan seçimler, verilen model için tek çözüm olmak zorunda değildir.

Doğrusal olmayan Bayesçi yaklaşım değerlendirilirken, öncelikle dağılım hakkında bir fikir sahibi olunmadığı göz önünde bulundurulmalıdır. Bu nedenle, doğrusal değerlendirmelerin denenmiş olması, doğrusal olmayan yöntemlerde daha uygun olan çalışmalarını gerçekleştirmek açısından kolaylık sağlar.

Yüksek Frekanslı Seslere ait veriler üzerinde yapılan bu çalışma, Doğrusal Olmayan Bayesçi Yaklaşım ile değerlendirilmiştir. Bu değerlendirme sonucunda elde edilen parametreler, anlamlı bulunmuştur. Doğrusal Olmayan Bayesçi Yaklaşım ile elde edilen sonuçlar En Küçük Kareler Yaklaşımıyla elde edilen sonuçlarla karşılaştırılmış ve Doğrusal Olmayan Bayesçi Yaklaşım ile elde edilen parametrelerin, En Küçük Kareler Yaklaşımı ile elde edilen parametrelere göre daha düşük standart sapmaya sahip oldukları görülmüştür. (Örneğin β_1 parametresi için EKK ile standart sapma değeri 1908,192 iken, Doğrusal Olmayan Bayesçi regresyon ile hesaplanan standart sapma değeri 1679,185' dir.) Parametrelerin düşük standart sapmaya sahip olması, kurulmuş modelin anakütle değerlendirmesi için daha uygun olduğu anlamını içerir. Ayrıca bir başka bakış açısıyla, bu parametrelerin modelde kullanılmış olmasının, Yüksek Frekanslı Seslerin değerlendirilmesi için gerekli olduğu varsayımı kuvvet kazanır. Sonuç olarak, Bayesçi Yaklaşım, En Küçük Kareler yöntemine göre daha tutarlı ve daha güvenilir sonuçlar vermektedir.

Sonsal model uygunluğu için hesaplanan p değeri 0,001 gibi bir değer olduğu için parametrelerin modele uygun olduğu görülür. Bu nedenle, bu modelin p değeri daha yüksek diğer modellere göre daha avantajlı olduğu kabul edilir.

Yüksek frekanslı sesler üzerinde, etkili olan parametreleri, En Küçük Kareler Yaklaşımına dayandırılarak analiz edip, çıkan sonuçlara göre değerlendirme yapmak, daha geniş bir frekans aralığı bulunmasına neden olur. Ancak, istenilen daha dar bir aralığın elde edilerek, kesin sonuca yakın değerlerin elde edilmesidir. Yeni yüksek frekanslı ses sistemleri oluşturulurken, bu sistemleri oluşturan şahısların yargıları göz önünde bulundurulur. Ona göre

uygun mekân ve malzemeler seçilir. Ayrıca yüksek frekanslı sesler üzerinde etkisi çok olan metal uzunlukları, uzmanların seçimine göre kullanılır. Bu nedenle, subjektif değerlendirmelerin her zaman için daha başarılı sonuçların elde edilmesine ön ayak olacağı sonucuna varılır.

Buradan da anlaşılabacağı üzere, doğrusal olmayan Bayesçi regresyon modeli, yeterli ön bilgi bulunmadığı durumlar için, öncelikli tercih edilecek yaklaşımlardan biri haline gelmiştir. EKK gibi yöntemler daha çok kullanılmış olmasına rağmen, daha iyi sonuçların elde edilmesini sağlamıştır.(Örneğin, küçük standart sapma değerlerine sahip olunması...) Subjektif değerlendirmeler ile edinilmiş deneyimlerin yapılan araştırmalarda eksik bilginin yerini doldurmasını sağlayacaktır.

KAYNAKLAR

- Birkes, D. Ve Dodge, Y., (1993), "Alternative Methods of Regression", Wiley&Sons,LTD, England
- Box, G.E.P and G.C. Tiao, (1992), "Bayesian Inference in Statistical Analysis", Addison Wesley, England
- Bradley P.Carlin and Thomas A.Louis, (2000), "Bayes and Empirical Bayes Methods for Data Analysis"; Chapman –Hall/Blue Diamonds Academy Yayınları
- Congdon, Peter, (2003), "Applied Bayesian Modelling", John Wiley&Sons,LTD, England
- De Groot,M.H, (1970), "Optimal Statistical Decisions";McGraw-Hill Inc.,Newyork
- De Groot,M.H, (1985), "Probability and Statistics"; Addison Wesley, England
- Denison,David G.T, Holmes,Christoper C., Mallick,Bani K., Smith,Adrian F.M., (2002), Bayesian Methods for Nonlinear Classification and Regression, John Wiley&Sons,LTD, England
- Evren.A. Atıf, (2006), "Olasılık Ve Çok Değişkeni Olasılık Dağılımlarına Giriş", YTÜ Yayınları, İstanbul
- Kahyaoğlu,F., (1999), "Regresyonda BayesYaklaşımı ve Cobb-Douglass Üretim Fonksiyonu Üzerine Bir Uygulama," Doktora Tezi
- Lee,P.M., (1989), "Bayesian Statistics;An Introduction"; Oxford.
- Lemos, J. M., (2009), "Bayesian Parameter Estimation in Nonlinear Dynamic Regression Models", INESC-ID Technical Report
- Leonard,Thomas, Hsu,John S.J., (2001), "Bayesian Methods", Cambridge University Press, U.K.
- Mood, A.M.; Graybill; F.A.; Boes, D. C. 1974; "Introduction to the Theory of Statistics" McGraw-Hill International Editions, Statistics Series, Third Edition,
- Neter, J., Kutner, M.H., Nachtsheim, C.J., Wasserman, W., (1993) "Applied Linear Statistical Methods", Wiley, England
- Press, S.J. et al; (2003) "Subjective and Objective Bayesian Statistics : Principles, Models, and Applications" ; Wiley-Interscience, England
- Zellner,A, (1971), "An Introduction to Bayesian Inference in Econometrics"; John Wiley , England

İNTERNET KAYNAKLARI

- [1]http://limnology.wisc.edu/regime/appendix_14jul03.pdf
- [2]http://209.85.135.132/search?q=cache:0PuBsV-REl8J:dondurma.ksu.edu.tr/semprkonf/tbt4s_Cd/tbt4s_ozet/print/firat.doc+do%C4%9Frusal+olmayan+bayes&cd=4&hl=tr&ct=clnk
- [3]http://en.wikipedia.org/wiki/Bayesian_inference
- [4]<http://www.abelard.org/briefings/bayes.htm>
- [5]<http://www.fon.hum.uva.nl/.../CDROM/Literature/LOTwinterschool2006/bidug.pnl.gov/presentations/PEP/>
- [6]<http://www.bayesian.org/>
- [7]<http://amath.colorado.edu/seminars/2003fall/PrGöztatSlides/linreg.pdf>
- [8]http://209.85.135.104/search?q=cache:FfxMo6M_XpEJ:home.uchicago.edu
- [9]<http://www.aris.com>
- [10]<http://hal.archives-ouvertes.fr/docs/00/32/70/81/PDF/Supelec425.pdf>
- [11]<http://jnm.snmjournals.org/cgi/reprint/43/6/762.pdf>
- [12]<http://www.mdpi.org/sensors/papers/s7060905.pdf>
- [13]<http://www.u.arizona.edu/~hirano/696E/ln10.pdf>

EKLER**Ek 1 Yüksek Frekanslı Seslere ait veriler**

Y	X
929000,00	0,50
571000,00	1000,00
310500,00	1750,00
115875,00	3750,00
80250,00	5750,00
636000,00	0,88
214000,00	2250,00
142500,00	3250,00
84750,00	5250,00
638000,00	0,75
268000,00	1750,00
164625,00	2750,00
71250,00	4750,00
673000,00	0,63
410000,00	1250,00
211500,00	2250,00
81750,00	4250,00
815000,00	0,50
131200,00	3000,00
599000,00	0,75
146200,00	3000,00
329000,00	1500,00
54400,00	6000,00
125600,00	3000,00
54400,00	6000,00
320000,00	1500,00
139500,00	3000,00
758000,00	0,50
200000,00	2000,00
104200,00	4000,00
595000,00	0,75
216700,00	2000,00
85500,00	5000,00
620000,00	0,75
202000,00	2250,00
77600,00	3750,00

37500,00	5750,00
118100,00	3000,00
547000,00	0,75
237000,00	2500,00
115500,00	4000,00
613000,00	0,75
177000,00	2500,00
87400,00	4000,00
592000,00	0,75
163000,00	2500,00
86200,00	4000,00
810000,00	0,50
48700,00	6000,00
146200,00	3000,00
817000,00	0,50
171700,00	2750,00
813000,00	0,50
289000,00	1750,00

Ek 2 Tahminler ve Kalıntılar

Uygulamada yapılan analizler sonucunda oluşan tahmin ve kalıntı değerleri aşağıdaki gibidir

Gözlemler	0.5=X	929000=Y	Tahminci(Y)	Kalıntılar
Göz1	1000,000	571000,000	252035,092	318964,908
Göz2	1750,000	310500,000	252035,092	58464,908
Göz3	3750,000	115875,000	252035,092	-136160,092
Göz4	5750,000	80250,000	252035,092	-171785,092
Göz5	0,880	636000,000	252035,281	383964,719
Göz6	2250,000	214000,000	252035,092	-38035,092
Göz7	3250,000	142500,000	252035,092	-109535,092
Göz8	5250,000	84750,000	252035,092	-167285,092
Göz9	0,750	638000,000	252065,845	385934,155
Göz10	1750,000	268000,000	252035,092	15964,908
Göz11	2750,000	164625,000	252035,092	-87410,092
Göz12	4750,000	71250,000	252035,092	-180785,092
Göz13	0,630	673000,000	255420,010	417579,990
Göz14	1250,000	410000,000	252035,092	157964,908
Göz15	2250,000	211500,000	252035,092	-40535,092
Göz16	4250,000	81750,000	252035,092	-170285,092
Göz17	0,500	815000,000	803287,030	11712,970
Göz18	3000,000	131200,000	252035,092	-120835,092
Göz19	0,750	599000,000	252065,845	346934,155
Göz20	3000,000	146200,000	252035,092	-105835,092
Göz21	1500,000	329000,000	252035,092	76964,908
Göz22	6000,000	54400,000	252035,092	-197635,092
Göz23	3000,000	125600,000	252035,092	-126435,092
Göz24	6000,000	54400,000	252035,092	-197635,092
Göz25	1500,000	320000,000	252035,092	67964,908
Göz26	3000,000	139500,000	252035,092	-112535,092
Göz27	0,500	758000,000	803287,030	-45287,030
Göz28	2000,000	200000,000	252035,092	-52035,092
Göz29	4000,000	104200,000	252035,092	-147835,092
Göz30	0,750	595000,000	252065,845	342934,155
Göz31	2000,000	216700,000	252035,092	-35335,092
Göz32	5000,000	85500,000	252035,092	-166535,092
Göz33	0,750	620000,000	252065,845	367934,155
Göz34	2250,000	202000,000	252035,092	-50035,092
Göz35	3750,000	77600,000	252035,092	-174435,092
Göz36	5750,000	37500,000	252035,092	-214535,092
Göz37	3000,000	118100,000	252035,092	-133935,092
Göz38	0,750	547000,000	252065,845	294934,155
Göz39	2500,000	237000,000	252035,092	-15035,092
Göz40	4000,000	115500,000	252035,092	-136535,092
Göz41	0,750	613000,000	252065,845	360934,155
Göz42	2500,000	177000,000	252035,092	-75035,092
Göz43	4000,000	87400,000	252035,092	-164635,092
Göz44	0,750	592000,000	252065,845	339934,155
Göz45	2500,000	163000,000	252035,092	-89035,092
Göz46	4000,000	86200,000	252035,092	-165835,092

Göz47	0,500	810000,000	803287,030	6712,970
Göz48	6000,000	48700,000	252035,092	-203335,092
Göz49	3000,000	146200,000	252035,092	-105835,092
Göz50	0,500	817000,000	803287,030	13712,970
Göz51	2750,000	171700,000	252035,092	-80335,092
Göz52	0,500	813000,000	803287,030	9712,970
Göz53	1750,000	289000,000	252035,092	36964,908

Ek 3 Metropolis Hastings fonksiyonu

#metropolis example: bivariate normal density with bivariate normal jumping kernel

```
metropolis_function(theta=c(1,1),mu=c(0,0),Sigma=diag(1,2)){
  thetastar_simulate.multnorm(theta,Sigma)
  Prec_solve(Sigma)
  ratio_ exp(-1/2*(t(thetastar-mu )%*%Prec%*%(thetastar-mu) -
t(theta-mu)%*%Prec%*%(theta-mu)))
  u_ runif(1)
  if(u <= ratio) {theta_thetastar}
  test_ (u <= ratio)
  return(theta)
}

iterate_function(theta0=c(0,0),NN){
  theta_matrix(NA,2,NN)
  test_NULL
  theta[,1]_theta0
  for(n in 2:(NN-1)){
    theta[,n]_metropolis(theta=theta[,n-1],mu=c(0,0),Sigma=diag(1,2))}
  return(theta)
}

plotmcmc_function(w=0,NN){
  if (w == 1) postscript("/home/biostats/fdominic/course/mcmc.ps")
  par(oma=c(0,0,2,0))
  I_iterate(theta0=c(0,0),NN)
  I1_iterate(theta0=c(3,3),NN)
  I2_iterate(theta0=c(3,-3),NN)
```

```

I3_iterate(theta0=c(-3,3),NN)
I4_iterate(theta0=c(-3,-3),NN)
plot(I[1,],I[2,],type="l",xlim=c(-4,4), ylim=c(-4,4),xlab="theta1",ylab="theta2")
points(0,0,pch=0,cex=4)
lines(I1[1,],I1[2,],lty=2)
points(3,3,pch=0,cex=4)
lines(I2[1,],I2[2,],lty=2)
points(3,-3,pch=0,cex=4)
lines(I3[1,],I3[2,],lty=2)
points(-3,3,pch=0,cex=4)
lines(I4[1,],I4[2,],lty=2)
points(-3,-3,pch=0,cex=4)
par(oma=c(0,0,0,0))
if (w == 1) dev.off()
}

```

Ek 4 Gibbs Sampler

#Approximating the joint posterior distribution $p(\theta, \mu, \sigma, \tau | \text{data})$ for the hierarchical normal model described in section 9.8, pag 285, Gelman book

#list of functions

##LIST OF THE FUNCTION-----

D_data()

I_initial.values(D)

I_one.iteration(I,D)

#gibbs.chain_markov(2000,100,2,I)

#-----

##DATA AND STARTING POINTS-----

data_function(){

J_4

#data-set

yy_rbind(c(62,60,63,59,NA,NA,NA,NA),

c(63,67,71,64,65,66,NA,NA), #use scan command for bigger data-sets

c(68,66,71,67,68,68,NA,NA),

c(56,62,60,61,63,64,63,59))

#prior hyperparameters

AA1_ 3#tausqr and sigmasqr have a IG dist

AA2 _ 3 #with mean close to rough estimates a

BB1_ 26 # a big variance

BB2_ 10

output_list(J=J,yy=yy,AA1=AA1,BB1=BB1,AA2=AA2,BB2=BB2)

return(output)

}

```

initial.values_function(D){
  yy_D$yy
  J_D$J
  AA1_D$AA1
  AA2_D$AA2
  BB1_D$AA1
  BB2_D$BB1
  theta_apply(yy,1,mean,na.rm=T)
  sigmasqr_mean(c(var(yy[1,!is.na(yy[1,])]),
  var(yy[2,!is.na(yy[2,])]),
  var(yy[3,!is.na(yy[3,])]),
  var(yy[4,!is.na(yy[4,])]))))
  mu_mean(theta)
  tausqr_var(theta)
  output_list(theta=theta,mu=mu,sigmasqr=sigmasqr,tausqr=tausqr)
  return(output)
}

```

#####ONE ITERATION OF THE GIBBS-SAMPLER-----

```

one.iteration_function(I,D){
  J_D$J
  yy_D$yy
  AA1_D$AA1
  AA2_D$AA2
  BB1_D$BB1
  BB2_D$BB2
  theta_I$theta

```

```

mu_I$mu

sigmasqr_I$sigmasqr

tausqr_I$tausqr

n_NULL

#Draw thetaji's
for(j in 1:J){
  n[j]_length(yy[j,!is.na(yy[j,])])
  vvv _ (1/tausqr+n[j]/sigmasqr)^(-1)
  mmm _ (mu/tausqr+sum(yy[j,!is.na(yy[j,])])/sigmasqr)*vvv
  theta[j]_rnorm(1,mmm,sqrt(vvv))
}

#Draw sigmasqr
num_NULL
for(j in 1:J){
  yysqr_sum(yy[j,]^2,na.rm=T)
  num[j]_yysqr+n[j]*theta[j]^2-2*sum(yy[j,],na.rm=T)*theta[j]
}

num_sum(num)

AA1_AA1+sum(n)/2

BB1_BB1+num/2

sigmasqr _ 1/(rGama(1,AA1)/BB1)

#Draw mu
mu_rnorm(1,mean(theta),sqrt(tausqr/J))

#Draw tausqr
AA2_AA2 + J/2

BB2_ BB2 + .5*sum(theta-mu)^2

```

```
tausqr_1/(rGama(1,AA2)/BB2)  
return(theta,sigmasqr,mu,tausqr)  
}
```

ÖZGEÇMİŞ

Doğum tarihi 15.01.1983

Doğum yeri Tunceli

Lise 1997-2001 Ümraniye Lisesi (YDA)

Lisans 2001-2006 Yıldız Teknik Üniversitesi Fen – Edebiyat Fak.
İstatistik Bölümü

Yüksek Lisans 2006- ... Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü
İstatistik Anabilim Dalı

Çalıştığı Kurum(lar)

2006-2007 Eagle Global Lojistik (Deniz İhracat Operasyon)
2007-Devam ediyor C/S Enformasyon Tekn. Ltd. Şti. (Test Uzmanı)