

**SAKLI MARKOV MODELİNİN FARKLI DAĞILIMLAR İÇİN
İNCELENMESİ**

**INVESTIGATION OF HIDDEN MARKOV MODEL FOR
VARIOUS DISTRIBUTIONS**

CEREN EDA CAN

PROF. DR. GÜL ERGÜN

Tez Danışmanı

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin

İstatistik Anabilim Dalı için Öngördüğü

DOKTORA TEZİ olarak hazırlanmıştır.

2016

CEREN EDA CAN' in hazırladığı "**Saklı Markov Modelinin Farklı Dağılımlar için İncelenmesi**" adlı bu çalışma aşağıdaki jüri tarafından **İSTATİSTİK ANABİLİM DALI'** nda **DOKTORA TEZİ** olarak kabul edilmiştir.

Prof. Dr. Serkan ERYILMAZ

Başkan

.....

Prof. Dr. Gül ERGÜN

Danışman

.....

Doç. Dr. Yeliz YOLCU OKUR

Üye

.....

Doç. Dr. Serpil AKTAŞ ALTUNAY

Üye

.....

Doç. Dr. Ayten YİĞİTER

Üye

.....

Bu tez Hacettepe Üniversitesi Fen Bilimleri Enstitüsü tarafından **DOKTORA TEZİ** olarak onaylanmıştır.

Prof. Dr. Salih Bülent ALTEN
Fen Bilimleri Enstitüsü Müdürü

ETİK

Hacettepe Üniversitesi Fen Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmada,

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

beyan ederim.

16 / 05 / 2016

CEREN EDA CAN

ÖZET

SAKLI MARKOV MODELİNİN FARKLI DAĞILIMLAR İÇİN İNCELENMESİ

Ceren Eda CAN

Doktora, İstatistik Bölümü

Tez Danışmanı: Prof. Dr. Gül ERGÜN

Mayıs 2016, 155 sayfa

Markov modellerinde doğrudan gözlemlenebilen sistemlere ilişkin problemler ele alınmaktadır. Ancak, gerçek uygulamalarda ilgilenilen sistemin doğrudan gözlemlenebilir olması mümkün olmayabilir. Doğrudan gözlemlenemeyen ve sadece başka bir stokastik süreç aracılığıyla dolaylı olarak gözlemlenebilen bu sistemler, Markov modellerinin uygulama alanlarını daraltan önemli bir sorundur. Markov modellerinin bir uzantısı olarak bilinen saklı Markov modeli, kısmen gözlemlenebilen sistemlerde istatistiksel çıkarsamalar yapılmasında güçlü ve esnek bir matematiksel yapıya sahiptir. Bu çalışmada, saklı Markov modelinin matematiksel yapısı ve çalışma ilkesi, Poisson ve beta dağılım aileleri varsayımı altında detaylı olarak incelenmiştir. Tez çalışması, Poisson saklı Markov modelinin Türkiye deprem verisine ilk kez uygulanmasıdır. 1900 ve 2012 yılları arasında Bilecik' in merkez olduğu 100 km yarıçapındaki bir alanda meydana gelmiş, büyüklüğü 4 ve daha fazla olan depremlerin yıllık frekansları modellenerek, gelecek 35 yıllık bir periyod için deprem riski tahmin edilmiştir. Elde edilen sonuçlar, belli bir zaman içerisinde meydana gelen depremlerin frekanslarının stokastik modellenmesinde standart model olarak kullanılan Poisson süreci ile karşılaştırılmıştır. Poisson saklı Markov modeline kıyasla, Poisson sürecinin aşırı yayılım gösteren ve serisel bağımlılık içeren gözlemlerin açıklanmasında çok yetersiz kaldığı ve aşırı uyum sorunu yarat-

tığı belirlenmiştir. Tez çalışmasında, ayrıca beta dağılım ailesi varsayımı altında saklı Markov modeline ilk kez Bayesci yaklaşım adapte edilmiştir. Böylece, klasik yaklaşım altında analitik olarak elde edilemeyen beta dağılım ailesi şekil parametrelerinin Bayesci tahminlerine ulaşılmıştır. Modelin Bayesci çözümlenmesi sayesinde, klasik yaklaşımda alternatif olarak kullanılabilen sayısal yöntemlerde sıklıkla karşılaşılan yakınsaklık sorunları da önlenmiştir. Saklı Markov modeli parametrelerinin tam koşullu dağılımlarının elde edilebilmesi özelliğinden dolayı, modelin Bayesci çözümlenmesinde Gibbs örnekleme kullanılmıştır. Eşlenik önsel kullanımıyla da, tam koşullu sonsal dağılımlar kapalı formda kolaylıkla elde edilebilmiştir. Burada, özellikle beta dağılımı şekil parametreleri için Bayyari' nin 1985 yılında önerdiği bileşik eşlenik önsel dağılımdan yararlanılmıştır. Gibbs örnekleme içinde, şekil parametrelerine ilişkin adımda Metropolis-Hastings algoritması kullanılmıştır. Bayesci beta saklı Markov modelinin işletiminde ise, Amerika Birleşik Devletleri' nin 1990-01 : 2015-11 dönemlerindeki işsizlik oranları kullanılmıştır. 2015 yılı Aralık ve 2016 yılı Ocak ayları için işsizlik oranlarının öngörü yoğunlukları tahmin edilerek öngörü dağılımları elde edilmiştir. Öngörü dağılımı, gerçek değerlerle oldukça tutarlı bulunmuştur. Tez çalışmasında yapılan tüm analizler için, R programında özgün kodlar yazılmıştır.

Anahtar Kelimeler: Saklı Markov Modeli, Markov Zincirleri, Poisson Süreci, Poisson Dağılımı, Beta Dağılımı, Monte Carlo Yöntemi, Markov Zinciri Monte Carlo Yöntemleri.

ABSTRACT

INVESTIGATION OF HIDDEN MARKOV MODEL FOR VARIOUS DISTRIBUTIONS

Ceren Eda CAN

Doctor of Philosophy, Department of Statistics

Supervisor: Prof. Dr. Gül ERGÜN

May 2016, 155 pages

Markov models consider the problems of the systems which are directly observable. However, in the real-world applications, the underlying system can be unobservable; but, can only be observed through another stochastic process. The evaluation of these partially observed systems is the major problem restricting the applications areas of Markov models. As an extension of Markov models, hidden Markov model has a powerful and flexible mathematical structure to make statistical inferences on partially observable systems. In this study, mathematical structure of hidden Markov model and its working principle are investigated in detail under Poisson and beta distribution families. The study is the first application of Poisson hidden Markov model for Turkish earthquake data. The annual frequencies of earthquakes occurred within a radius of 100 km area centered on Bilecik, from 1990 to 2012 with magnitudes ≥ 4 are modeled by using Poisson hidden Markov model. The earthquake hazard in the area is determined by forecasting the annual frequencies of earthquakes for the next 35 years. Also, Poisson hidden Markov model is compared to Poisson process which is a traditional way for stochastic modeling the occurrences within a certain time interval. Contrary to the Poisson hidden Markov model, Poisson process becomes seriously insufficient to explain over-dispersed and serially dependent data and causes over-fitting. The study is also the first time that

Bayesian approach is adapted to hidden Markov model under the family of beta distributions. In classical approach, the shape parameters of beta distribution can not be obtained analytically. In order to overcome the problem of analytical insolubility, the estimates of the shape parameters are obtained under Bayesian approach. Bayesian approach also prevents the convergence problems which result from numerical methods. Hidden Markov model allows the full conditional distributions to be obtained for all model parameters. Due to this property, Gibbs sampling is used for generating samples from the joint posterior distribution of parameters. Employing the conjugate prior distributions for model parameters, the full conditional posterior distributions are obtained easily in closed form. In 1985, Bayyari proposed a joint conjugate prior distribution for the shape parameters of beta distribution. In this study, Bayyari' s joint conjugate prior distribution is used for obtaining the full conditional posterior distribution of the shape parameters in closed form. Metropolis-Hastings algorithm is carried out within Gibbs sampling for the shape parameters as well. For the application of beta hidden Markov model, the unemployment rates of the United States of America for the period of 1990-01 and 2015-11 are modeled under Bayesian approach. The forecast distributions are obtained for the unemployment rates in December, 2015 and January, 2016. The actual values for terms are found to be quite consistent with the forecast distributions. The original codes are written in R for all analysis.

Keywords: Hidden Markov Model, Markov Chains, Poisson Process, Poisson Distribution, Beta Distribution, Monte Carlo Method, Markov Chain Monte Carlo Methods.

TEŞEKKÜR

Doktora çalışmalarımın her aşamasında destek ve yardımlarını benden esirgemeyen; ilham veren fikirleri ve değerli önerileri ile beni yönlendiren; yaşam felsefesi ile çok yönlü bir bilim insanı olmayı öğreten danışman hocam Sayın Prof. Dr. Gül ERGÜN' e,

Yaşadığım tüm zorluklarda beni cesaretlendiren; büyük bir özveri, sabır ve içtenlikle destek olan; sevgisini, inancını ve güvenini hep hissettiğim; hayatımdaki varlığıyla bana güç veren sevgili hocam Doç. Dr. Serpil AKTAŞ ALTUNAY' a,

Bayesci yaklaşımı öğrenmem ve saklı Markov modeline adapte etmem de engin bilgi ve deneyimlerini benimle paylaşan; öncü fikir ve önerileriyle istatistiksel modellemeye olan bakış açımı değiştiren ve ufkumu genişleten değerli hocam Prof. Dr. Refik SOYER' e,

Tez çalışmam sırasında sevgileriyle ve dostluklarıyla hep yanımda olan; hoşgörü ve sabırla desteklerini hiç eksik etmeyen, en zor zamanlarımı benimle paylaşarak çalışmalarımın moral desteğini sağlayan değerli arkadaşlarım Didem EGEMEN' e, Gökçen ALTUN' a ve Emrah ALTUN' a,

Tez konum hakkında araştırma yapmak amacıyla Amerika Birleşik Devletleri' nde kaldığım süre zarfında verdiği maddi destekten dolayı TÜBİTAK' a,

Son olarak, hayatımın her aşamasında bana destek veren; sevgi ve hoşgörüleriyle beni her zaman motive eden; inanç ve güvenleriyle beni güçlendiren annem Meyva CAN, babam Zülküf CAN ve canım kardeşlerim İlkın Seda CAN ve Nilay Nida CAN' a sonsuz teşekkürlerimi sunarım.

İÇİNDEKİLER

	<u>Sayfa</u>
ÖZET	i
ABSTRACT	iii
TEŞEKKÜR	v
İÇİNDEKİLER	vi
ÇİZELGELER	viii
ŞEKİLLER	x
SİMGELER VE KISALTMALAR	xii
1. GİRİŞ	1
2. BAYESÇİ ÇIKARSAMA	5
2.1. Önsel Dağılımlar	8
2.2. Sonsal Dağılımın Özetlenmesi	12
3. MARKOV ZİNCİRLERİ	17
4. STOKASTİK BENZETİM YÖNTEMLERİ	24
4.1. Monte Carlo Yöntemi	24
4.2. Markov Zinciri Monte Carlo Yöntemleri	26
4.2.1. Metropolis-Hastings Algoritması	32
4.2.2. Gibbs Örnekleme	39
5. SAKLI MARKOV MODELİ	45
5.1. Modelin Tanımı ve Bileşenleri	46
5.2. Model Varsayımları ve Süreçlerin Bazı Özellikleri	49
5.3. Modelin Üç Temel Problemi ve Çözüm Yöntemleri	52

5.3.1. Deęerlendirme Problemi ve özüm Yöntemi	52
5.3.2. Kodlama Problemi ve özüm Yöntemi	58
5.3.3. Öğrenme Problemi ve özüm Yöntemi	61
6. POISSON SAKLI MARKOV MODELİ	69
7. BETA SAKLI MARKOV MODELİ	75
7.1. Bayesci Beta Saklı Markov Modeli	78
7.1.1. Beta Saklı Markov Modelinde Bayesci Parametre Tahminlerinin Elde Edilmesi	79
7.1.2. Gözlem Deęerlerinin Öngörülmesi	86
7.1.3. Modelin Boyut Sayısının Tahmini	87
8. UYGULAMA	98
8.1. Poisson Saklı Markov Modelinin Uygulanması	98
8.2. Beta Saklı Markov Modelinin Uygulanması	109
9. SONUÇ VE TARTIŞMA	127
KAYNAKLAR	131
ÖZGEÇMİŞ	138

ÇİZELGELER

	<u>Sayfa</u>
Çizelge 4.1. Yakınsaklık Testleri.....	30
Çizelge 7.1. Jeffreys' in Kuralı ile BF_{12} Bayes Faktörünün Model Seçimindeki Yorumu.....	88
Çizelge 8.1. Model Karşılaştırması-1	100
Çizelge 8.2. Model Karşılaştırması-2	100
Çizelge 8.3. Ortalama ve Varyans Değerlerinin Karşılaştırılması	101
Çizelge 8.4. 3-Durumlu PSMM ile Poisson Sürecinin Model Uyumu Karşılaştırması	102
Çizelge 8.5. 3-Durumlu PSMM' nin En Çok Olabilirlik Parametre Tahminleri	103
Çizelge 8.6. Gelecek 35 Yıl için Saklı Durumlara İlişkin Olasılıklar	104
Çizelge 8.7. 2013-2017 Yıllarında Deprem Riski	105
Çizelge 8.8. 2013-2017 Yıllarında Beklenen Deprem Sayısı ve Güven Aralıkları .	106
Çizelge 8.9. 1-3 Yıl İçindeki Deprem Riski Açısından 3-Durumlu PSMM ve Poisson Sürecinin Karşılaştırılması.....	106
Çizelge 8.10. 4-6 Yıl İçindeki Deprem Riski Açısından 3-Durumlu PSMM ve Poisson Sürecinin Karşılaştırılması.....	108
Çizelge 8.11. Rakip Modeller için Bilgiler	112
Çizelge 8.12. $m = 2$ için Yakınsaklık Testleri (790000 İterasyon Üzerinden)	112
Çizelge 8.13. 2-Durumlu Beta-SMM Parametrelerinin Bayesci Tahminleri	113
Çizelge 8.14. 2-durumlu Beta-SMM için Log-Marjinal Olabilirlik Değerinin Hesaplanması	114
Çizelge 8.15. $m = 3$ için Yakınsaklık Testleri (1190000 İterasyon Üzerinden)	115

Çizelge 8.16.	$m = 3$ için Yakınsaklık Testleri (1071000 İterasyon Üzerinden)	116
Çizelge 8.17.	3-Durumlu Beta-SMM Parametrelerinin Bayesci Tahminleri	117
Çizelge 8.18.	3-durumlu Beta-SMM için Log-Marjinal Olabilirlik Değerinin Hesaplanması	118
Çizelge 8.19.	$m = 4$ için Yakınsaklık Testleri (1990000 İterasyon Üzerinden)	119
Çizelge 8.20.	$m = 4$ için Yakınsaklık Testleri (1194000 İterasyon Üzerinden)	121
Çizelge 8.21.	4-Durumlu Beta-SMM Parametrelerinin Bayesci Tahminleri	123
Çizelge 8.22.	4-durumlu Beta-SMM için Log-Marjinal Olabilirlik Değerinin Hesaplanması	124
Çizelge 8.23.	Rakip Modellerin Karşılaştırılması	125
Çizelge 8.24.	Öngörü Dağılımlarının Elde Edilmesi için Geçen Süreler	125

ŞEKİLLER

	<u>Sayfa</u>
Şekil 2.1. Bayesci Tahmin Süreci	7
Şekil 2.2. İki Tepeli Bir Sonsal Dağılımda Bayesci Aralık Tahminleri.....	15
Şekil 4.1. $\Gamma(5, 4)$ Öneri Dağılımı Kullanıldığında Oluşan Zayıf Zincir	37
Şekil 4.2. $\Gamma(3, 0.5)$ Öneri Dağılımı Kullanıldığında Oluşan Karma Zincir	38
Şekil 4.3. χ^2_6 Öneri Dağılımı Kullanıldığında Oluşan Karma Zincir	38
Şekil 4.4. X_1 Raslantı Değişkeni için Oluşturulan Karma Zincir	42
Şekil 4.5. X_2 Raslantı Değişkeni için Oluşturulan Karma Zincir	43
Şekil 4.6. Doğrudan Örnekleme ile Gibbs Örneklemesinin Karşılaştırılması	43
Şekil 5.1. Saklı Markov Modelinde Süreçler Arasındaki Bağımlılık Yapısı.....	50
Şekil 5.2. İleri-Yön Değişkeninin Oluşumu	56
Şekil 5.3. Geri-Yön Değişkeninin Oluşumu	57
Şekil 6.1. m -Durumlu Poisson-SMM' nin Çalışma İlkesi	70
Şekil 7.1. m -Durumlu Beta-SMM' nin Çalışma İlkesi.....	75
Şekil 8.1. 1990-01 : 2015:11 Dönemleri ABD İşsizlik Oranları	109
Şekil 8.2. 1990-01 : 2015:11 Dönemleri ABD İşsizlik Oranlarının Otokorelasyon Fonksiyonu	110
Şekil 8.3. 1990-01 : 2015:11 Dönemlerinde X Saklı Sürecinin Bulunduğu Durumlar	114
Şekil 8.4. 1990-01 : 2015:11 Dönemlerinde X Saklı Sürecinin Bulunduğu Durumlar	118

Şekil 8.5. 1990-01 : 2015:11 Dönemlerinde X Saklı Sürecinin Bulunduğu Durumlar	124
Şekil 8.6. Y_{T+1} ve Y_{T+2} için Öngörü Dağılımları	126



SİMGELER VE KISALTMALAR

Simgeler

m	Saklı Durum Sayısı
S	Kesikli Sonlu Durum Kümesi
π	İlk Olasılık Vektörü
δ	Denge Dağılımı
$\lambda^{(m)}$	Poisson Dağılım Ailesi Oran Parametreleri Vektörü
$\alpha^{(m)}, \beta^{(m)}$	Beta Dağılım Ailesi Şekil Parametreleri Vektörleri
$B(\cdot, \cdot)$	Beta Fonksiyonu
$\psi(\cdot)$	Digamma Fonksiyonu

Kısaltmalar

SMM	Saklı Markov Modeli
MCMC	Markov Zinciri Monte Carlo
d.	Dağılımda Yakınsama
h.h.k.	Hemen Hemen Kesinlikle Yakınsama
EM	Beklenti En Büyükleme
E-Adımı	Beklenti Adımı
M-Adımı	En Büyükleme Adımı
Poisson-SMM	Poisson Saklı Markov Modeli
AIC	Akaike Bilgi Kriteri
BIC	Bayes Bilgi Kriteri
Beta-SMM	Beta Saklı Markov Modeli
EBeta	Eşlenik Beta Dağılımı
BF	Bayes Faktörü
SBIC	Schwarz Bayesci Bilgi Kriteri
PSMM	Poisson Saklı Markov Modeli
ABD	Amerika Birleşik Devletleri
OECD	Ekonomik Kalkınma ve İşbirliği Örgütü

1. GİRİŞ

Gerçek hayatta birçok olay zaman boyunca rasgele değişen ve gelişen sistemlerin varlığında oluşur. Dinamik yapıda olan bu sistemlerin davranışlarının incelenmesinde stokastik süreçler kuramı kullanılır. Günümüzde birçok bilim dalında bu tür olaylara ilişkin problemlerin çözümlenmesinde ve bu olaylara açıklık getirilmesinde stokastik modellemeden yararlanılır.

Dinamik sistemlerin zaman boyunca hareketleri çoğu zaman stokastik bir model olan Markov modelleri ile açıklanmaktadır. Markov modellerinde incelenen sistemin durumları ve bu durumlara ilişkin olasılık fonksiyonları bilinmektedir. Sistemdeki her durum gözlemlenebilir bir olaya karşılık gelmektedir. Böylece, sistemin herhangi bir t anında içinde bulunduğu durum bir gözlem olarak tanımlanabilmekte ve bunun sonucu olarak, sistem doğrudan gözlemlenebilmektedir. Ancak, birçok uygulamada sistemin gözlemlenebilir olması çoğu zaman mümkün değildir. Bu durum, Markov modellerinin uygulama alanını önemli ölçüde daraltan bir sorundur.

Saklı Markov modeli (SMM), sistemin doğrudan gözlemlenemediği stokastik modelleme problemleri için uygun bir alternatiftir. Markov modellerinin bir uzantısı olarak kabul edilmektedirler. Standart SMM, iki katmanlı bir stokastik süreçtir. Alt katmanda durumları doğrudan gözlemlenemeyen bir Markov zinciri; üst katmanda ise, bu Markov zincirinin etkisi altında olan ve gözlemleri üreten bir stokastik süreç vardır. Burada sistemin izlediği süreç alt katmandaki Markov zinciridir ve saklı süreç olarak adlandırılır. Üst katmandaki sürece ise, gözlemlenebilir süreç adı verilir. Sistem sadece gözlemlenebilen süreç tarafından izlenebilir. Tüm istatistiksel çıkarımlar gözlemlenebilen bu sürecin çıktıları üzerinden yapılmaktadır.

SMM, ilk olarak 1966 yılında Baum ve Petrie [1] tarafından “*sonlu durumlu Markov zincirlerinin olasılıksal fonksiyonları*” olarak tanıtılmıştır. Markov zincirlerinin olasılıksal fonksiyonu olması nedeniyle, SMM durağan olmayan sistemlerdeki rasgele değişimleri saptamada ve sistemin zaman içerisindeki gelişiminde beklenmeyen yapıların ortaya çıkartılmasında Markov modellerine kıyasla daha güçlü ve esnek bir yapıya sahiptir. Özellikle, aşırı yayılım gösteren ve serisel bağımlılık içeren verilerin modellenmesinde oldukça başarılıdır. SMM, zengin bir matematiksel yapıya sahiptir. Modelin dayandığı mantık ve teorik temel, birçok alanda ele alınan problemlere kolayca uyarlanabilmekte-

dir. Ağır ve karmaşık bir teorik yapıya sahip olmasına rağmen, günümüzde bilgisayar teknolojisinin ve bilişim sistemlerinin hızla artan gelişimi, SMM' nin kolaylıkla uygulanmasına olanak sağlamaktadır. Modelin teorisi problemlere uygun bir şekilde entegre edildiğinde, çok iyi sonuçlar alınmaktadır.

SMM' nin matematiksel yapısı ve çalışma ilkesi literatürde birçok modelle ilişkilidir. Örneğin, SMM' de iki süreç arasındaki bağımlılık yapısı dinamik bir Bayes ağı oluşturmaktadır. Bu bağlamda, SMM, dinamik Bayes ağının bir alt sınıfıdır. Bir başka örnek verilecek olursa, SMM' deki saklı sürecin durumları ile karışım modellerindeki (mixture models) karışım bileşenleri birbirine benzemektedir. Bu iki modelde en önemli fark, karışım bileşenleri birbirinden bağımsız iken, saklı sürecin durumları bir Markov zinciri oluşturmaktadır. Bu nedenle, karışım modelleri verideki aşırı yayılımı açıklayabilirken, verinin içerdiği serisel bağımlılığı açıklayamamaktadır. SMM, karışım modellerinin bu açığını kapatmaktadır. Literatürde SMM' ye benzer diğer modellerden bazıları ise, durum-uzay modelleri (state-space models), Markov-değişimli modeller (Markov-switching models) ve Markov-modüleli Poisson süreçleri (Markov-modulated Poisson processes)' dir.

SMM' nin teorisine ilişkin ilk uygulamalar, Baker [2], Jelinek, Bahl ve Mercer [3], Bahl ve Jelinek [4] ve Jelinek [5] tarafından "*konuşma tanıma*" üzerine yapılmıştır. Günümüzde SMM çok çeşitli alanlarda uygulanmakla birlikte, modelin uygulandığı alanlar her geçen gün hızla artmaktadır. Modelin uygulanmasında öne çıkan bazı çalışmalar, konuşma tanıma [6]-[8]; finans ve ekonomide [9]-[12]; yazılım güvenilirliğinde [13], [14]; trafik mühendisliğinde [15]; biyolojide [16]; dil modelleme ve metin analizinde [17], [18]; meteorolojide [19]-[22]; biyoinformatikte [23]-[26]; biyofizikte ve biyokimyada [27]-[29]; gen motifi bulmada [30]; protein katlanmasında [31]; metamorfik virüs bulmada [32] vb. olarak söylenebilir.

Tez çalışmasında, farklı dağılım aileleri varsayımı altında SMM' nin incelenmesi, bu dağılımların model ve elde edilecek tahminler üzerindeki etkilerinin tartışılması, mevcut SMM' nin işletimine kolaylık sağlayacak bazı uyarlamaların ve kullanılan algoritmalarda bazı iyileştirmelerin önerilmesi amaçlanmıştır. Burada, özellikle gözlemlerin Poisson ve beta dağılımlı olduğu süreçler için SMM ile nasıl bir modelleme yapılacağı gerçek veriler üzerinden incelenmiştir. Modelin uygulanmasında R programından yararlanılması ve ilgili kodların geliştirilmesi hedeflenmiştir.

Tez çalışmasının ilerleyen bölümlerinde ilk olarak Bayesci çıkarsamaya ilişkin bilgiler Bölüm 2' de yer verilmiştir. Beta dağılım ailesi parametreleri klasik yaklaşım altında analitik olarak elde edilememektedir. Bu problemin giderilmesinde beta dağılım ailesi varsayımı altında SMM' nin Bayesci çözümlemesi yapılmıştır. Bu nedenle, Bölüm 2' de verilen bilgiler SMM' nin Bayesci çözümlemesinde kullanılan terminolojiye temel oluşturmaktadır.

Bölüm 3' te Markov zincirlerine ilişkin genel bir bilgi verilmiştir. Bu bilgiler, SMM' nin alt katmanında yer alan Markov zincirinin davranışının daha iyi anlaşılması amacıyla verilmiştir. Ayrıca, SMM' nin Bayesci çözümlemesinde kullanılan stokastik benzetim yöntemleride Markov zincirlerine dayanmaktadır. Bu bağlamda Bölüm 3, SMM' nin hem klasik yaklaşım hem de Bayesci yaklaşım altında çözümlenmesinde kullanılan süreçlere ve teorik terimlere açıklık getirmektedir.

Bölüm 4' te SMM' nin Bayesci çözümlemesinde kullanılan stokastik benzetim yöntemleri detaylı olarak incelenmiştir.

Bölüm 5' te ilk olarak Standart SMM' nin tanımı yapılmış ve model bileşenleri açıklanmıştır. Ardından, model varsayımları verilmiştir. Son olarak, SMM' nin uygulanmasında ortaya çıkan üç temel problem ayrıntılı olarak ele alınmış ve bu problemlerin çözümleri anlatılmıştır.

Bölüm 6' da Poisson dağılım ailesi altında SMM incelenmiştir. Klasik yaklaşım altında Poisson dağılım ailesi parametrelerinin nasıl elde edileceği açıklanmıştır.

Bölüm 7' de, beta dağılım ailesi varsayımı altında SMM incelenmiştir. Bu bölümde, öncelikle klasik yaklaşım altında beta dağılım ailesi parametrelerinin analitik olarak elde edilemediği gösterilmiştir. Ardından, beta dağılım ailesi varsayımı altında SMM' ye Bayesci yaklaşımın nasıl adapte edileceği detaylı olarak açıklanmıştır. Son olarak ise, SMM' de bir model seçim problemi olan optimal saklı durum sayısının belirlenmesi ele alınmış ve Bayesci yaklaşım altında modelin optimal boyut sayısının nasıl belirleneceği anlatılmıştır.

Bölüm 8' de, Poisson ve beta dağılım aileleri varsayımı altında SMM, iki ayrı gerçek veri seti üzerinde uygulanmıştır. Poisson dağılımı için, klasik yaklaşım altında; beta dağılımı için ise, Bayesci yaklaşım altında çözümlenme yapılarak ilgili modelin parametreleri tah-

min edilmiştir ve gözlemlere ilişkin öngörülerde bulunulmuştur.

Bölüm 9' da ise, tez çalışmasında elde edilen sonuçlar özetlenmiş ve gelecek çalışmalarda hedeflenenler belirtilmiştir.



2. BAYESÇİ ÇIKARSAMA

Çoğu araştırmada gelecek verilerin öngörülmesi ilgi çekse de, araştırmalarda daha önemli bir hedef verilerin yaratıldığı süreç hakkında bilgi sahibi olmaktır. Bu hedef doğrultusunda, istatistiksel modellemeye başvurulur ve oluşturulan modelde yer alan parametrelerin tahminine ihtiyaç duyulur. Parametre tahmini iki farklı yaklaşım altında yapılabilir. Bunlar, *klasik yaklaşım* ve *Bayesci yaklaşım*dır. Bu iki yaklaşım arasındaki en önemli farklılık parametre tanımıdır.

Klasik yaklaşımda bilinmeyen model parametreleri sabit değerler olarak kabul edilirler. Klasik yaklaşım, örneklem bilgisine dayanır ve tamamen objektiftir. Gözlenen veriler, tekrarlanabilir rasgele örneklemdir ve model parametreleri tekrarlanabilir olan bu süreçte değişmezlerdir. Modellemede sadece, olabilirlik fonksiyonu olarak adlandırılan " $P(\text{Veriler} \mid \text{Parametreler})$ " objektif bilgisi kullanılır. Parametrelere ilişkin subjektif bilgi kabul edilmez ve çıkarsamalarda olabilirlik fonksiyonu tek kaynak olarak doğrudan kullanılır. Klasik yaklaşımda nokta tahmin edicilerinin sağlaması gereken birçok özellik mevcuttur. Bir parametrenin nokta tahmin edicisi, yansız, en küçük varyanslı, tutarlı ve yeterli olmalıdır. Bunlara ek olarak tahmin edicilerin sağlam olması diğer bir değişle, aykırı değerlerden de etkilenmemesi beklenir. Belirtilen bu özelliklerin sağlanabilmesi için, büyük örneklem genişliğine ihtiyaç duyulur. Ancak, örneklem genişliğinin artması çoğu kez zaman ve maliyetin artmasına yol açar.

Bayesci yaklaşımda ise, bilinmeyen model parametreleri birer raslantı değişkeni olarak kabul edilirler. Bu nedenle, klasik yaklaşımın aksine parametrelerin nokta tahmin edicilerinin, yansızlık, tutarlılık ve minimum varyanslılık gibi özellikleri taşımaları gerekmez. Burada, parametrelerin yerine gözlenen veriler sabit olarak kabul edilirler. Bu yaklaşımda hedef, veriler kullanılarak parametreler hakkında bilgi sahibi olunmasıdır. Bu süreçte, klasik yaklaşımdan farklı olarak objektif bilgi ile yetinilmeyip, objektif bilginin subjektif bilgi ile bütünleşmesinden yararlanır. Parametreye ilişkin tüm bilgi kullanıldığı için, çıkarsama sürecinde herhangi bir bilgi kaybı olmamaktadır.

Subjektif bilgi kaynaklarından biri, geçmişteki benzer çalışmaların sonuçlarının kullanılmasıdır. Ayrıca, bir teoriden veya gerçekliği kabul görmüş bir savdan gelen bilgi ya da araştırmacının inaç ve tecrübelerine dayanarak oluşturduğu öznel düşünce de önemli diğer kaynak olabilir. Model parametreleri hakkındaki subjektif bilgi, önsel bilgi olarak

adlandırılır ve bu bilgi ışığında “ $P(\text{Parametreler})$ ” önsel dağılımı oluşturulur. Örnekleme den gelen objektif bilgi ise, Bayesci tahmin sürecine olabilirlik fonksiyonu olarak katılır. Bayesci yaklaşımda örneklem büyüklüğü için bir kısıt yoktur. Az sayıda veri ile çalışılabilir. Öyle ki, hiç veri gözlemlenmeden sadece önsel dağılım üzerinden parametreler hakkında bilgi sahibi olunabilir. Bu nedenle, örneklem büyüklüğü için bir kısıt olmaması, zaman ve maliyet tasarrufu sağlar.

Bayesci yaklaşımın temel taşı olan Bayes teoremi kullanılarak, olabilirlik fonksiyonu önsel bilgi ile birleştirilir ve parametreler için “ $P(\text{Parametreler} \mid \text{Veriler})$ ” sonsal dağılım elde edilir. Koşullu olasılık ve toplam olasılık formülüne dayanan Bayes teoremi, 18. yüzyılın sonlarında Thomas Bayes tarafından ortaya konulmuştur. S örneklem uzayı, ayrık olan k tane olaydan oluşsun. Bu durumda, $S = \{A_1, A_2, \dots, A_k\}$ ’dir. Örneklem uzayında rasgele bir B olayı tanımlansın. Herhangi bir A_i ($i = 1, \dots, k$) olayı için, B ve A_i olaylarının aynı anda ortaya çıkması olasılığı koşullu olasılık tanımından

$$P(B \cap A_i) = P(B \mid A_i)P(A_i) \quad (2.1)$$

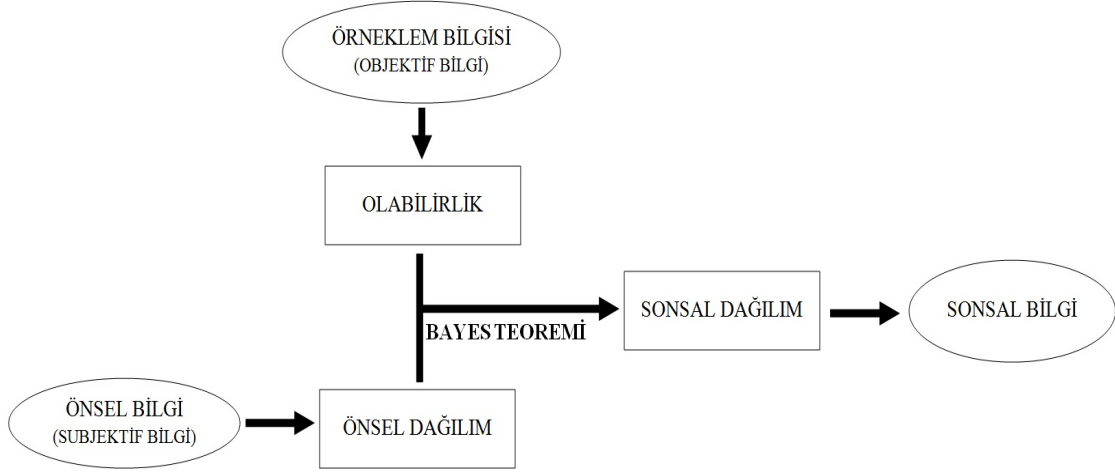
olarak hesaplanır. B olayının ortaya çıkma olasılığı ise, aşağıdaki gibi elde edilir.

$$\begin{aligned} P(B) &= P(B \cap A_1) + P(B \cap A_2) + \dots + P(B \cap A_k) \\ &= P(B \mid A_1)P(A_1) + P(B \mid A_2)P(A_2) + \dots + P(B \mid A_k)P(A_k) \end{aligned} \quad (2.2)$$

Eşitlik 2.2, toplam olasılık formülüdür. Eşitlik 2.1 ve Eşitlik 2.2’de verilen koşullu olasılık ve toplam olasılık formüllerinden yararlanılarak Bayes teoremi matematiksel olarak ifade edilir. B olayının ortaya çıktığı bilindiğinde, herhangi bir A_i olayının ortaya çıkması olasılığı aşağıdaki gibi elde edilir.

$$\begin{aligned} P(A_i \mid B) &= \frac{P(A_i \cap B)}{P(B)} \\ &= \frac{P(B \mid A_i)P(A_i)}{P(B \mid A_1)P(A_1) + P(B \mid A_2)P(A_2) + \dots + P(B \mid A_k)P(A_k)} \end{aligned} \quad (2.3)$$

Eşitlik 2.3, Bayes formülü olarak adlandırılır. Bayes teoremi kullanılarak elde edilen sonsal dağılım üzerinden parametreler tahmin edilir ve bu parametreler hakkında istenilen istatistiksel çıkarımlar yapılır. Bayesci tahmin süreci aşağıda yer alan şema ile ifade edilebilir:



Şekil 2.1. Bayesci Tahmin Süreci

Bayesci yaklaşım altında modelleme yapılırken tüm belirsizlikler olasılık dağılımları şeklinde ifade edilir. \mathbf{y} gözlemlerini açıklamak amacıyla kurulan istatistiksel modelde, ilgi bir ya da birden fazla parametre üzerine yoğunlaşabilir. θ , bilinmeyen model parametreleri olsun. Model parametreleri kesikli ya da sürekli yapıda olabilirler. Yapılacak tüm istatistiksel çıkarımlar model parametrelerinin sonsal dağılımına bağlıdır. $f(\mathbf{y}) \neq 0$ olmak üzere, θ 'nın sonsal dağılımı Bayes teoremi kullanılarak aşağıdaki gibi elde edilir [33], [34].

$$f(\theta | \mathbf{y}) = \frac{f(\mathbf{y}, \theta)}{f(\mathbf{y})} = \begin{cases} \frac{f(\mathbf{y} | \theta) f(\theta)}{\int_{R_\theta} f(\mathbf{y} | \theta) f(\theta) d\theta}, & \theta \text{ Sürekli durumda ise} \\ \frac{f(\mathbf{y} | \theta) f(\theta)}{\sum_{R_\theta} f(\mathbf{y} | \theta) f(\theta)}, & \theta \text{ Kesikli Durumda ise} \end{cases} \quad (2.4)$$

Eşitlik 2.4'te $f(\theta)$, θ 'nın önsel dağılımı; $f(\mathbf{y} | \theta)$, sabit \mathbf{y} gözlemleri için θ 'nın olabilirlik fonksiyonudur. R_θ , θ 'nın tanım kümesidir. $f(\theta)$ önsel dağılımı, \mathbf{y} gözlemlenmeden önce, θ 'ya dair ön bilgiyi temsil eder. Olabilirlik fonksiyonu ise, gözlemlerden gelen bilgiyi içermektedir. Paydadaki $f(\mathbf{y})$, \mathbf{y} 'nin marjinal dağılımıdır ve θ 'yı içermemektedir. Bu nedenle, $f(\mathbf{y})$, sonsal dağılımın integralini veya toplamını bire eşitleyen sabit bir terimdir ve normalleştirme katsayısı olarak kabul edilir. Buna göre, θ 'nın sonsal dağılımı orantısal olarak aşağıdaki gibi ifade edilebilir:

$$\underbrace{f(\theta | \mathbf{y})}_{\text{Sonsal Dağılım}} \propto \underbrace{f(\mathbf{y} | \theta)}_{\text{Olabilirlik Fonksiyonu}} \underbrace{f(\theta)}_{\text{Önsel Dağılım}} \quad (2.5)$$

Eşitlik 2.5' te θ ' ya ilişkin önsel bilginin örneklemden gelen bilgiyle birleştirilmesi sonucu

θ ' nın sonsal dağılımına ulaşıldığı görülür. Bu nedenle, elde edilen sonsal dağılım, θ hakkında tüm bilgiyi içerir.

2.1. Önsel Dağılımlar

θ parametresi hakkında herhangi bir gözlem yapmadan önceki önsel bilginin tanımlanması ve bu bilginin bir önsel dağılıma aktarılması Bayesci yaklaşımda en zor konulardan biridir. Önsel dağılım, veriler gözlemlenmeden önce, parametreler hakkındaki belirsizliğin bir ölçüsüdür. Bu dağılımın tanımlanmasında, uzman görüşleri, yapılan geçmiş çalışmalardan elde edilen bilgiler ve parametrelerin ait olduğu kitlenin özellikleri kullanılmalıdır.

Bayesci tahmin sürecinde önsel dağılım aracılığıyla kullanılan önsel bilgi, klasik yaklaşımda bir dezavantaj olarak kabul edilir. Ancak, uygun bir önsel dağılımın seçilmesiyle, dezavantaj olarak kabul edilen bu bilgi avantaja çevrilebilir. Bunun sonucunda da elde edilen parametre tahminlerinin varyansları daha küçük ve aralık tahminleri de daha dar elde edilir. Sonsal dağılım, parametreye ilişkin önsel bilgi ve veriden gelen bilgiyi içermesi nedeniyle önsel dağılıma kıyasla daha az değişkendir. Dağılımdaki değişkenliğin bir ölçüsü olan varyans değerleri karşılaştırıldığında, önsel dağılımın varyansı her zaman sonsal dağılımın varyansından büyüktür [35]. Parametrelere ilişkin önsel varyans, sonsal varyans cinsinden aşağıdaki gibi yazılabilir [36]:

$$\begin{aligned}
 V[\theta] &= E[V[\theta | \mathbf{y}]] + V[E[\theta | \mathbf{y}]] & (2.6) \\
 &= E[E[\theta^2 | \mathbf{y}] - (E[\theta | \mathbf{y}])^2] + E[(E[\theta | \mathbf{y}])^2] - (E[E[\theta | \mathbf{y}]])^2 \\
 &= E[\theta^2] - E[(E[\theta | \mathbf{y}])^2] + E[(E[\theta | \mathbf{y}])^2] - (E[\theta])^2 \\
 &= E[\theta^2] - (E[\theta])^2
 \end{aligned}$$

Eşitlik 2.6' da verilerden gelen bilgi ile birlikte parametrelere ilişkin belirsizliğin azaldığı görülmektedir. Bayesci yaklaşımda uygun bir önsel dağılım seçildiğinde, klasik yaklaşıma kıyasla, daha güvenilir tahminler elde edilmektedir. Önsel dağılımın ortalaması ise, sonsal dağılımın ortalaması cinsinden aşağıdaki gibi bulunur [36]:

$$E[\theta] = \int_{R_\theta} \theta f(\theta) d\theta = \int_{R_\theta} \theta \left(\int_{R_y} f(\theta, \mathbf{y}) d\mathbf{y} \right) d\theta$$

$$\begin{aligned}
&= \int_{R_y} \left(\int_{R_\theta} \theta f(\theta | \mathbf{y}) d\theta \right) f(\mathbf{y}) d\mathbf{y} \\
&= \int_{R_y} E[\theta | \mathbf{y}] f(\mathbf{y}) d\mathbf{y} \\
&= E[E[\theta | \mathbf{y}]]
\end{aligned} \tag{2.7}$$

Eşitlik 2.7' de önsel ortalama, mümkün verilerin dağılımı üzerinden tüm olası sonsal ortalamaların ortalaması olarak elde edilmiştir. Burada, model parametrelerinin sürekli yapıda olduğu kabul edilmiştir. Aynı durum, kesikli yapıdaki model parametreleri için de geçerlidir.

Önsel dağılımın seçiminin yanı sıra, örneklem büyüklüğü de sonsal dağılım üzerinde etkilidir. Örneklem büyüklüğü önsel dağılımı etkilememesine rağmen; önsel dağılımın sonsal dağılıma katkısını etkilemektedir. Örneklem büyüklüğü arttıkça önsel dağılımın sonsal dağılım üzerindeki etkisi azalmaktadır. Çünkü, gözlenen verilerin bir fonksiyonu olan olabilirlik fonksiyonu örneklem büyüklüğü arttıkça daha sivri olmakta ve önsel dağılım olabilirlik fonksiyonuna göre daha basık kalmaktadır. Bu durumda, önsel dağılımın sonsal dağılıma katkısı çok fazla olmaz. Basık olan önsel dağılımın kuyrukları uzun olacağından sonsal dağılımı çok fazla etkilemez. Örneklem büyüklüğü azaldıkça, olabilirlik fonksiyonu önsel dağılıma göre daha basık olur ve sonsal dağılım önsel dağılım tarafından çok fazla etkilenir [37]. Bu durum parametre tahminleri üzerinde etkili olacağı için, önsel dağılım dikkatli bir şekilde seçilmelidir. Önsel dağılımlar, parametre hakkında bilgi içermelerine, belirsizlik durumlarına ve eşlenik olmalarına göre sınıflandırılabilirler.

Parametreye ilişkin bilgi içermeye durumlarına göre, önsel dağılımlar iki gruba ayrılır. Bunlar, bilgi içeren (informative) önseller ve bilgi içermeyen (noninformative) önsellerdir.

Bilgi içeren önsel dağılımlar, parametrelerin geldikleri kitlenin özelliklerine, önceden gözlemlenmiş veri kümesinden gelen bilgiye, geçmiş deneyimlere ve uzman görüşlerine dayanılarak belirlenebilir. Bu önsel dağılımlar, sonsal dağılım üzerinde oldukça etkilidir. Ayrıca, olabilirlik fonksiyonu bu önsel dağılımlara baskın değildir. Bu nedenlerle, bilgi içeren önsel dağılımlar oldukça dikkatli bir şekilde belirlenmelidir. Uygun bir önsel dağılım belirlenmemesi durumunda, dağılımın matematiksel yapısı sonsal dağılımı daha karmaşık bir yapıya dönüştürebilir ve/veya yanlış parametre tahminlerinin elde edilmesine

neden olabilir.

Bilgi içermeyen önsel dağılımlar, model parametreleri hakkında herhangi bir bilgi olmadığı veya bilgi var ama elde edilmesi zor ya da öyle bir bilginin dağılımını matematiksel olarak formüle etmenin olanaklı olmadığı durumlarda kullanılırlar. Bu önsel dağılımlar, genel olarak referans önseller olarak adlandırılır. Referans önseller, sonsal dağılım üzerinde en az etkiye sahiptir. Bilgi içermeyen bir önsel dağılım kullanılması durumunda, Bayesci yaklaşım klasik yaklaşıma yakınsar. Bunun nedeni, parametrelere ilişkin kullanılan bilginin sadece olabilirlik fonksiyonu aracılığıyla verilerden sağlanan bilgi olmasıdır. Referans önsel dağılımların yoğunluk fonksiyonları, belirsiz (vague), dağınık/yaygın (diffuse) ve düz (flat) olarak tanımlanmaktadır [36]. Bilgi içermeyen önsel dağılımın belirlenmesinde, Jeffreys [38]' in 1961 yılında ortaya attığı kuraldan yararlanılabilir. Jeffreys'in kuralına göre, parametrelerin önsel dağılımı Fisher bilgi matrisinin determinantının karekökü ile orantılıdır. Buna göre, Jeffreys'nin önseli aşağıdaki gibi ifade edilebilir [36], [39].

$$f(\boldsymbol{\theta}) \propto |J(\boldsymbol{\theta})|^{1/2} \quad (2.8)$$

Eşitlik 2.8' de $J(\boldsymbol{\theta})$, model parametreleri için Fisher bilgi matrisidir. Fisher bilgi matrisi, \mathbf{y} gözlenen verilerin dağılımından elde edilen model parametrelerine ilişkin bilgi miktarının ölçüsüdür ve aşağıdaki gibi tanımlanmıştır [39].

$$J(\boldsymbol{\theta}) = -E \left[\frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \ln f(\mathbf{y} | \boldsymbol{\theta}) \middle| \boldsymbol{\theta} \right] = - \int_{R_{\mathbf{y}}} \left(\frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \ln f(\mathbf{y} | \boldsymbol{\theta}) \right) f(\mathbf{y} | \boldsymbol{\theta}) d\mathbf{y}.$$

Model parametrelerinin tanım bölgesinde önsel dağılımın yoğunluk fonksiyonunun (olasılık fonksiyonunun) integrali (toplamı) alındığında sonucun bire eşit olup olmasına göre önsel dağılımlar iki gruba ayrılır. Bunlar, tam (proper) önseller ve tam olmayan (improper) önsellerdir. Parametrenin tanım bölgesinde önsel dağılımın integrali (toplamı) bire eşitse, bu önsel dağılım tam önsel dağılımdır [37]:

$$\begin{aligned} \boldsymbol{\theta} \text{ sürekli ise, } & \int_{R_{\boldsymbol{\theta}}} f(\boldsymbol{\theta}) d\boldsymbol{\theta} = 1. \\ \boldsymbol{\theta} \text{ kesikli ise, } & \sum_{R_{\boldsymbol{\theta}}} f(\boldsymbol{\theta}) = 1. \end{aligned}$$

Parametrenin tanım bölgesinde belirlenen önselin integrali (toplamı) sonsuza eşit veya mevcut değil ise, bu önsel dağılım tam olmayan önsel olarak adlandırılır [37]. Herhangi

bir θ_i parametresi için tanımlanan aşağıdaki önsel yoğunluklar tam olmayan önsellere örnektirler [37]:

$$f(\theta_i) = a , -\infty < \theta_i < +\infty , a > 0$$

$$f(\theta_i) = a \theta_i^{-1} , 0 < \theta_i < +\infty , a > 0$$

Bayesci yaklaşımda sonsal dağılımın mutlaka tam olması gerekmektedir. Çünkü, tam olmayan sonsal dağılım üzerinden hiçbir çıkarsama yapılamaz. Tam olmayan önsel dağılımlar, tam olmayan sonsal dağılımlara neden olabilir. Bu nedenle, tam olmayan önseller çok dikkatli bir şekilde belirlenmelidir ya da mümkün olduğunca kullanımından uzak durulmalıdır. Tam olmayan bir önsel dağılım kullanıldığında, sonsal dağılımın tamlığının sağlanması gerekir. Bu durumda, tüm \mathbf{y} değerleri için normalleştirme katsayısı sonlu ise, sonsal dağılımda tamdır.

Önsel dağılımlar, sonsal dağılımla aynı dağılım ailesinden gelme durumlarına göre, iki gruba ayrılır. Bunlar, eşlenik (conjugate) önseller ve eşlenik olmayan önsellerdir. Önsel dağılım ve sonsal dağılım aynı dağılım ailesinin bir üyesi iseler, bunlara eşlenik dağılımlar denir ve önsel dağılım, ilgili olabilirlik fonksiyonu için eşlenik önsel olarak adlandırılır. Eşlenik önsel dağılım kullanıldığında, sonsal dağılım analitik olarak elde edilebilir. Sonsal dağılım önsel dağılımla aynı dağılım ailesine mensup olmakla birlikte sonsal dağılımın parametreleri örneklemden gelen bilgiyle güncellenmiş bir şekilde elde edilir. Böylece, Bayesci çıkarsamada büyük hesaplama kolaylığı sağlarlar. Olabilirlik fonksiyonu üstel ailenin bir üyesi ise, ilgili dağılımın parametreleri için eşlenik önsel dağılım mevcuttur [36], [39]. Olabilirlik fonksiyonu aşağıda verilen biçimde yazılabiliyorsa üstel ailenin bir üyesidir [36].

$$f(\mathbf{y} | \boldsymbol{\theta}) = \left(\prod_{i=1}^n h(y_i) \right) (g(\boldsymbol{\theta}))^n \exp \left(\boldsymbol{\phi}(\boldsymbol{\theta})^\top \sum_{i=1}^n u(y_i) \right)$$

Burada, $h(\cdot)$ ve $u(\cdot)$ sadece \mathbf{y} gözlem değerlerine bağlı; $g(\cdot)$ ve $\boldsymbol{\phi}(\cdot)$ ise sadece $\boldsymbol{\theta}$ model parametrelerine bağlı sürekli fonksiyonlardır. Eşlenik olmayan önsel dağılımların kullanılması oldukça karmaşık sonsal dağılımlara neden olabilir. Özellikle, yüksek boyutlu ve karmaşık modellerde sonsal dağılımların analitik olarak elde edilmesini ve bu dağılımlardan çıkarsama yapılmasını oldukça güçleştirir.

2.2. Sonsal Dağılımın Özetlenmesi

Bayesci yaklaşımda, sonsal dağılımın elde edilmesi birinci amaçtır. İkinci amaç ise, elde edilen sonsal dağılım üzerinden ihtiyaç duyulan istatistiksel çıkarımlar yapılarak parametrelere veya gelecek gözlem değerlerine ilişkin bilgiye ulaşmaktır. Sonsal dağılımın içerdiği bilgi birçok farklı şekilde özetlenebilir. Parametreler hakkında sonsal bilgiye ulaşmak amacıyla, sonsal dağılımın karakteristik özelliklerini ortaya koyan özetleyici istatistikler hesaplanabilir. Her bir parametre için, belirlenen bir α güven düzeyinde aralık tahminleri elde edilebilir. Ayrıca, sonsal dağılım grafiksel olarak özetlenebilir. Parametre uzayının yüksek boyutlu olması durumunda, sonsal dağılımın grafiksel gösterimi genellikle uygun bir seçim olmaz. Bu durumda, grafiksel sunum yerine özetleyici istatistikler daha çok tercih edilirler.

Önsel dağılıma göre, sonsal dağılımdaki değişimi göstermek amacıyla konum ve yayılım ölçüleri hesaplanabilir. Konum ölçüleri arasından en çok kullanılanlar, sonsal ortalama, sonsal ortanca, sonsal tepe değeri ya da değerleridir. Yayılım ölçüleri için ise, sonsal varyans, sonsal standart sapma, sonsal ortalama mutlak sapma, sonsal çeyrek değerler genişliği (interquartile range) ve diğer sonsal yüzdeler hesaplanabilir. Ayrıca, parametreler için sonsal kovaryans değerleride bulunabilir.

Kurulan bir modelde k tane parametre olduğu varsayalım ve bu modelde parametreler $\theta = \{\theta_1, \theta_2, \dots, \theta_k\}$ parametre vektörü ile gösterilsin. Parametrelere ilişkin sonsal değerlerin hesaplanabilmesi için, marjinal sonsal dağılımlara ulaşılması gerekmektedir. İlgi dışı tüm parametreler üzerinden bileşik sonsal dağılım integrallenerek marjinal sonsal dağılımlara geçilir. Herhangi bir θ_i parametresi için marjinal sonsal dağılım aşağıdaki gibi elde edilir:

$$f(\theta_i | \mathbf{y}) = \int_{R_{\theta^{-i}}} f(\theta | \mathbf{y}) d\theta^{-i} \quad (2.9)$$

Eşitlik 2.9' da $\theta^{-i} = \{\theta_j : j = 1, 2, \dots, k \text{ ve } j \neq i\}$ olmak üzere, $R_{\theta^{-i}}$, θ^{-i} ' nin tanım aralığıdır. Parametre sayısına bağlı olarak, $f(\theta | \mathbf{y})$ bileşik sonsal dağılımı yüksek boyutlu ve karmaşık yapıda olabilir. Bu durumda, $f(\theta_i | \mathbf{y})$ marjinal sonsal dağılımın elde edilmesinde Eşitlik 2.9' da verilen integralin analitik çözümüne ulaşamayabilir. Bayesci çıkarımda analitik çözümsüzlüğün giderilmesinde sayısal integrasyon yöntemlerine başvurulur. Böylece, analitik olarak elde edilemeyen bilgilerin yaklaşık değer-

lerine ulaşılır. Örneklemeye temelli olan stokastik benzetim yöntemleri kullanılarak bu integralin ve daha birçok Bayesci hesaplamaların yaklaşık çözümüne ulaşılabilir. Stokastik benzetim yöntemlerinden standart Monte Carlo yöntemi ve Markov zinciri Monte Carlo yöntemleri en çok tercih edilenlerdir. Bu yöntemler, tez çalışmasının ilerleyen bölümlerinde anlatılmıştır.

Marjinal sonsal dağılımın ortalaması, ilgili parametrenin bir nokta tahmini olmaktadır. Sonsal dağılımın sağa veya sola çarpık olması; tek veya çok tepeli olmasına göre, parametrenin nokta tahmini sonsal tepe değeri (veya değerleri) ya da sonsal ortanca değeri olabilir. Çok tepeli bir sonsal dağılım için birden fazla sonsal tepe değeri bulunur. Herhangi bir θ_i parametresi için sonsal ortalama, sonsal tepe değeri, sonsal ortanca ve sonsal varyans aşağıda verildiği gibi hesaplanmaktadır:

$$\text{Sonsal ortalama} : E[\theta_i | \mathbf{y}] = \int_{R_{\theta_i}} \theta_i f(\theta_i | \mathbf{y}) d\theta_i$$

$$\text{Sonsal tepe değeri} : \hat{\theta}_i = \arg \max_{\theta_i} f(\theta_i | \mathbf{y})$$

$$\text{Sonsal ortanca} : \int_{-\infty}^{\tilde{\theta}_i} f(\theta_i | \mathbf{y}) d\theta_i = 0.5 \text{ integralini sağlayan } \tilde{\theta}_i \text{ değeri' dir.}$$

$$\text{Sonsal varyans} : V[\theta_i | \mathbf{y}] = \int_{R_{\theta_i}} (\theta_i - E(\theta_i | \mathbf{y}))^2 f(\theta_i | \mathbf{y}) d\theta_i$$

Herhangi θ_i ve θ_j parametreleri için sonsal kovaryans değeri aşağıda verilmiştir:

$$Kov(\theta_i, \theta_j) = \int_{R_{\theta_i}} \int_{R_{\theta_j}} (\theta_i - E[\theta_i | \mathbf{y}]) (\theta_j - E[\theta_j | \mathbf{y}]) f(\theta_i, \theta_j | \mathbf{y}) d\theta_j d\theta_i \quad (2.10)$$

Eşitlik 2.10' da θ_i ve θ_j parametrelerinin bileşik sonsal dağılımı yer almaktadır. Bu parametrelerin bileşik sonsal dağılımı aşağıdaki gibi elde edilir:

$$f(\theta_i, \theta_j | \mathbf{y}) = \int_{R_{\theta^{-i,j}}} f(\boldsymbol{\theta} | \mathbf{y}) d\boldsymbol{\theta}^{-i,j} \quad (2.11)$$

Eşitlik 2.11' de $\boldsymbol{\theta}^{-i,j} = \{\theta_h : h = 1, 2, \dots, k, h \neq i \text{ ve } h \neq j\}$ olmak üzere, $R_{\boldsymbol{\theta}^{-i,j}}$, $\boldsymbol{\theta}^{-i,j}$ ' nin tanım aralığıdır.

Bayesci yaklaşımda, parametrelerin sonsal belirsizliğinin özetlenmesinde aralık tahminleri oldukça önemlidir. Klasik yaklaşımda, parametreler için belirlenen bir $\%100(1 - \alpha)$ güven düzeyinde güven aralığı (confidence interval) bulunur ve aralık tahmininin yorumu, " tahmin edilen aralığın parametreyi içermesi olasılığı $\%100(1 - \alpha)$ ' dir. " şeklinde yapılır. Burada, parametre bir sabittir ve tahmin edilen aralığın sınırları birer raslantı değişkenidir. Bayesci yaklaşımda ise, aralık tahmininin yorumu farklıdır. Bu yorum, " parametrenin tahmin edilen aralığa düşmesi olasılığı $\%100(1 - \alpha)$ ' dir. " şeklinde yapılır. Parametre bir raslantı değişkenidir ve tahmin edilen aralığın sınırları birer sabittir. Bayesci çıkarsamada, iki çeşit aralık tahmini vardır. Bunlar,

- i- $\%100(1 - \alpha)$ Sonsal Bayes güven aralığı (posterior credible interval) ve
- ii- $\%100(1 - \alpha)$ En yüksek sonsal yoğunluk aralığı (highest posterior density interval)' dir.

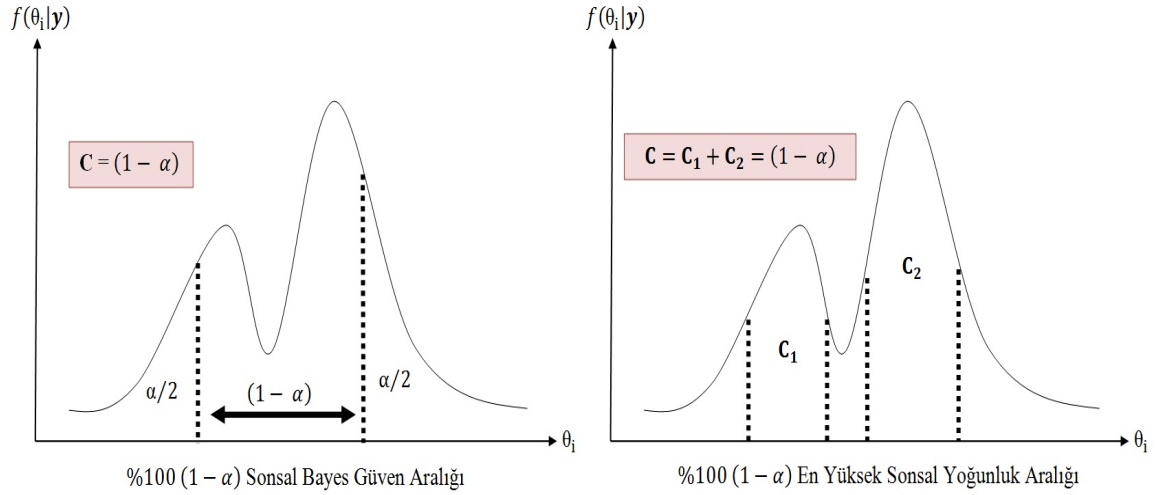
$\mathcal{C} \subset R_{\theta}$ olmak üzere, \mathcal{C} kümesi aralık tahminini gösterebilir. $\%100(1 - \alpha)$ sonsal Bayes güven aralığı, aşağıdaki eşitliği sağlayan \mathcal{C} kümesidir [40].

$$\int_{\mathcal{C}} f(\theta | \mathbf{y}) d\theta = 1 - \alpha \quad (2.12)$$

Eşitlik 2.12' de görülen \mathcal{C} kümesinin, sonsal olasılığın tam olarak $\%100(\alpha/2)$ ' lik kısmının üstünde ve altında olan değerlerin oluşturduğu aralık olduğu görülmektedir. Aşağıda verilen iki koşul sağlandığında, \mathcal{C} kümesi, $\%100(1 - \alpha)$ en yüksek sonsal yoğunluk aralığı olarak tanımlanır [37]:

- i. \mathcal{C} kümesi, sonsal olasılığın $\%100(1 - \alpha)$ ' lik kısmını içerir: $P(\theta \in \mathcal{C} | \mathbf{y}) = 1 - \alpha$.
- ii. \mathcal{C} kümesindeki değerlerin sahip olduğu yoğunluk aralık dışında bulunanlardan daha büyüktür: $\theta_1 \in \mathcal{C}$ ve $\theta_2 \notin \mathcal{C}$ için, $f(\theta_1 | \mathbf{y}) \geq f(\theta_2 | \mathbf{y})$.

Sonsal dağılımın tek tepeli ve simetrik olması durumunda, bu iki aralık aynı bölgeyi temsil etmektedir. Tek tepeli fakat çok çarpık bir sonsal dağılımda, bu iki aralık birbirinden oldukça farklıdır [36]. Çok tepeli dağılımlarda ise, en yüksek sonsal yoğunluk aralığının bulunması zor olsada, sonsal Bayes güven aralığına kıyasla parametre hakkında daha çok bilgi taşımaktadır ve parametrenin aralık tahmini için daha güvenilir sonuçlar vermektedir. Burada, θ_i parametresinin iki tepeli bir marjinal sonsal dağılıma sahip olduğu varsayalım. Bu dağılımda Bayesci aralık tahminlerinin tanımları aşağıda gösterilmiştir.



Şekil 2.2. İki Tepeli Bir Sonsal Dağılımda Bayesci Aralık Tahminleri

Şekil 2.2.' de en yüksek sonsal yoğunluk aralığının, C_1 ve C_2 olmak üzere, iki aralıktan oluştuğuna dikkat edilmelidir. Çok tepeli dağılımlarda, en yüksek sonsal yoğunluk aralığı birden fazla aralıktan oluşmaktadır. Bayesci yaklaşımda, aralık tahminleri genellikle sonsal dağılım üzerinden simülasyon yapılarak hesaplanır. Sonsal dağılımın analitik olarak kapalı formu elde edilebildiği durumda ise, dağılım fonksiyonu yardımıyla sonsal Bayes güven aralığı elde edilebilir.

Sonsal dağılım üzerinden parametreler için ihtiyaç duyulan tüm istatistiksel çıkarımlar yapılırken; ayrıca değeri bilinmeyen ve henüz gözlemlenmemiş gözlem değişkeni için, sonsal öngörü dağılımı da elde edilebilir. Diğer bir deyişle, \mathbf{y} verisi gözlemlendikten sonra, bilinmeyen $\tilde{\mathbf{y}}$ gelecek gözlemi tahmin edilebilir. Bunun için, \mathbf{y}' ye koşullu olarak $\tilde{\mathbf{y}}$ ' nin dağılımına ihtiyaç duyulur. Bu dağılıma $\tilde{\mathbf{y}}$ ' nin öngörü dağılımı denir ve aşağıdaki gibi elde edilir [36]:

$$\begin{aligned}
 f(\tilde{\mathbf{y}} | \mathbf{y}) &= \int_{R_{\theta}} f(\tilde{\mathbf{y}}, \theta | \mathbf{y}) d\theta \\
 &= \int_{R_{\theta}} f(\tilde{\mathbf{y}} | \mathbf{y}, \theta) f(\theta | \mathbf{y}) d\theta \\
 &= \int_{R_{\theta}} f(\tilde{\mathbf{y}} | \theta) f(\theta | \mathbf{y}) d\theta
 \end{aligned} \tag{2.13}$$

θ bilindiğinde, \mathbf{y} ve $\tilde{\mathbf{y}}$ birbirinden koşullu olarak bağımsız olacaktır. Bunun bir sonucu olarak, Eşitlik 2.13' e ulaşılır. $f(\tilde{\mathbf{y}} | \mathbf{y})$ sonsal öngörü yoğunluğu, parametrelerin sonsal dağılımı üzerinden bir beklenen değere dönüşmüştür. Parametrelerin sonsal

dağılımından gelen bilgiden yararlanılarak, sonsal öngörü dağılımının beklenen değeri ve varyansı elde edilebilir. Sonsal öngörü ortalaması,

$$E[\tilde{\mathbf{y}} | \mathbf{y}] = E[E[\tilde{\mathbf{y}} | \boldsymbol{\theta}, \mathbf{y}] | \mathbf{y}] = E[\boldsymbol{\theta} | \mathbf{y}] \quad (2.14)$$

olarak bulunur. Eşitlik 2.14' te bulunan sonuca göre, gelecek herhangi bir gözlemin beklenen değeri parametrelerin sonsal ortalamasına eşittir [36]. Sonsal öngörü varyansı ise,

$$\begin{aligned} V[\tilde{\mathbf{y}} | \mathbf{y}] &= E[V[\tilde{\mathbf{y}} | \boldsymbol{\theta}, \mathbf{y}] | \mathbf{y}] + V[E[\tilde{\mathbf{y}} | \boldsymbol{\theta}, \mathbf{y}] | \mathbf{y}] \\ &= E[V[\tilde{\mathbf{y}} | \boldsymbol{\theta}] | \mathbf{y}] + V[\boldsymbol{\theta} | \mathbf{y}] \end{aligned} \quad (2.15)$$

olarak elde edilir. Eşitlik 2.15' e göre gelecek bir gözleme ilişkin varyans, kurulan modelden gelen varyans ile parametrelerin sonsal varyansının toplamıdır [36].

Tez çalışmasında, gözlemlerin beta dağılım ailesinden geldiği bir saklı Markov modelinin tahmin edilmesinde yakınsama sorunundan dolayı yetersiz kalan klasik tahmin yöntemleri yerine Bayesci yaklaşım kullanılmıştır. Bu nedenle, bu bölümde Bayesci yaklaşımın temelleri verilmiştir.

Saklı Markov modelinin teorisi Markov zincirlerine dayanmaktadır. Tez çalışmasında modelin Bayesci çözümlenmesinde kullanılacak olan stokastik benzetim yöntemlerinin temelinde de Markov zincirleri teorisi yatmaktadır. Bu nedenlerle, klasik ve Bayesci yaklaşım altında modelin teorik alt yapısının ve işleyişinin daha iyi anlaşılabilmesi amacıyla, izleyen bölümde Markov zincirleri teorisine bir giriş yapılacaktır.

3. MARKOV ZİNCİRLERİ

Markov süreçleri, Rus matematikçi Andrei Andreivich Markov [41]' un 1906 yılında stokastik süreçler üzerine tamamen teorik olarak yaptığı çalışma sonucunda ortaya çıkmıştır ve Markov'un ismine atfen adlandırılmışlardır. Markov [42], teorik bulgularını ilk kez Alexander Pushkin' in Eugene Onegin adlı şiir-romanındaki ilk 20.000 sesli ve sessiz harflerin durumunu inceleyerek uygulamıştır ve sonuçlarını 1913 yılında yaptığı çalışmayla özetlemiştir. Markov'un geliştirdiği teknik olasılık teorisini yeni bir yöne kaydırmıştır ve stokastik süreçlerin birçok farklı alanda uygulanmasını sağlamıştır.

Markov süreci, Markov özelliği taşıyan bir stokastik süreçtir. Stokastik süreç, rasgele değişkenlerin oluşturdukları $X = \{X_t, t \in T\}$ şeklinde bir topluluk olarak düşünülebilir. Burada T , dizin kümesi olarak adlandırılır. X_t raslantı değişkeni, t zamanında sürecin içinde bulunduğu durumu gösterir. Zaman, sürecin parametresidir. Stokastik süreç, kesikli ya da sürekli parametrelili olabilir. Dizin kümesi, $T = \{0, 1, 2, \dots\}$ gibi sayılabilen sonsuzlukta elemandan oluşuyorsa sürece kesikli parametrelili süreç denir. Dizin kümesi şayet $T = \{t : -\infty < t < +\infty\}$ gibi sayılamayacak sonsuzlukta elemandan oluşuyorsa, bu sürece de sürekli parametrelili süreç denir. Tez çalışmasında ele alınan modelin kesikli zaman periyodunda incelenmesi nedeniyle, bu bölümde sadece kesikli parametrelili Markov süreci üzerinde durulmuştur.

X_t raslantı değişkeninin alacağı tüm olası durumların oluşturduğu kümeye durum uzayı denir ve S ile gösterilir. Durum uzayı, sonlu ya da sayılabilir sonsuzlukta durum içeriyorsa kesikli yapıdadır ve sürece, kesikli durumlu süreç denir. Durum uzayı, sayılamayacak sonsuzlukta durum içeriyorsa sürekli yapıdadır ve sürece, sürekli durumlu süreç denir.

T dizin kümesinde herhangi bir n sayıda zaman noktası, $t_1 < t_2 < \dots < t_n$ kümesini oluştursun. $X_{t_1}, X_{t_2}, \dots, X_{t_{n-1}}$ ' lerin verilen değerlerine göre, X_{t_n} ' in koşullu dağılımı sadece $X_{t_{n-1}}$ ' in değerine bağlı ise, X süreci Markov özelliğini sağlar ve Markov süreci olarak adlandırılır. X sürecinin Markov özelliği aşağıda verildiği gibi tanımlanır [43], [44]:

$$P(X_{t_n} = x_{t_n} \mid X_{t_1} = x_{t_1}, X_{t_2} = x_{t_2}, \dots, X_{t_{n-1}} = x_{t_{n-1}}) = P(X_{t_n} = x_{t_n} \mid X_{t_{n-1}} = x_{t_{n-1}}) \quad (3.1)$$

Eşitlik 3.1' de x_{t_i} ($i = 1, \dots, n$) değeri, X_{t_i} ' nin t_i zamanında bulunduğu durumdur ve durum uzayının bir elemanıdır. Durum uzayı kesikli yapıda olduğunda, Markov sürecine Markov zinciri denir. Durum uzayı sonlu ise, sonlu Markov zinciri; sonsuz ise, sonsuz

Markov zinciri olarak adlandırılır.

Markov zincirinin durumlar arasındaki hareketleri, koşullu olasılık olarak tanımlanan geçiş olasılıklarıyla belirlenir. $T = 0, 1, 2, \dots$ olmak üzere, kesikli parametrelili bir $\{X_t, t \in T\}$ Markov zinciri olsun. Herhangi bir $t \in T$ anında S_i durumundan tek bir adımda S_j durumuna geçiş olasılığı aşağıdaki gibi tanımlanır [43], [44]:

$$p_{ij} = P(X_{t+1} = S_j | X_t = S_i) , S_i, S_j \in \mathcal{S} \quad (3.2)$$

Eşitlik 3.2' de verilen geçiş olasılıklarının dağılımı P geçiş olasılıkları matrisi ile gösterilir. P matrisi aşağıda verilen özellikleri sağlar [43], [44]:

i. $0 \leq p_{ij} \leq 1, \forall S_i, S_j \in \mathcal{S}$

ii. $\sum_{\forall S_j \in \mathcal{S}} p_{ij} = 1, \forall S_i \in \mathcal{S}$

Herhangi bir $t \in T$ anında S_i durumunda olan sürecin n ($n = 1, 2, \dots$) adım sonra S_j durumunda olması olasılığına n -adım geçiş olasılığı denir ve aşağıdaki gibi gösterilir [43], [44]:

$$p_{ij}^{(n)} = P(X_{t+n} = S_j | X_t = S_i) , S_i, S_j \in \mathcal{S} \quad (3.3)$$

n -adım geçiş olasılıkları dağılımı ise, $P^{(n)}$ n -adım geçiş olasılıkları matrisi ile gösterilir. $P^{(n)}$ n -adım geçiş olasılıkları matrisinin tüm elemanları pozitif olacak şekilde bir n değeri mevcut ise, bu zincire düzenli (regular) Markov zinciri denir [44]. Eşitlik 3.3' te verilen geçiş olasılıkları, sadece n zaman aralığına bağlı, t zaman noktasına bağlı değilse, bu zincire homojen Markov zinciri denir. Homojen olmayan Markov zincirinde ise, durumlar arası geçiş olasılıkları zincirin bulunduğu zaman noktasına bağlı olarak değişir. Eşitlik 3.3' te verilen n -adım geçiş olasılığı aşağıdaki biçimde yazılabilir [43], [44]:

$$p_{ij}^{(n)} = \sum_{\forall S_k \in \mathcal{S}} p_{ik}^{(r)} p_{kj}^{(n-r)} , r \in \{0, 1, \dots, n\} \quad (3.4)$$

Eşitlik 3.4' te verilen ifadeye Chapman-Kolmogorov eşitliği denir. Chapman-Kolmogorov eşitliğinin matris formu, $P^{(n)} = P^{(r)} P^{(n-r)}$ olarak yazılır. $n = 1, 2, 3, 4, 5$ için aşağıdaki sonuçlara ulaşılır [44]:

$$P^{(1)} = P^1 = P$$

$$P^{(2)} = PP = P^2$$

$$P^{(3)} = PP^{(2)} = P^3$$

$$P^{(4)} = PP^{(3)} = P^4$$

$$P^{(5)} = PP^{(4)} = P^5$$

Bu durumda, Chapman-Kolmogorov eşitliğinin genel bir sonucu olarak n -adım geçiş olasılıkları matrisi $P^{(n)} = P^n$ olarak hesaplanır. Zincirin herhangi bir $t \in T$ anında S_i durumunda bulunması olasılığı aşağıdaki gibi tanımlanır:

$$\pi_i(t) = P(X_t = S_i) , S_i \in S \quad (3.5)$$

Eşitlik 3.5' te verilen olasılıkların dağılımı, $\pi(t)$ vektörü ile gösterilir. Bu vektörün elemanları toplamı bire eşittir:

$$\sum_{\forall S_i \in S} \pi_i(t) = 1 , \forall t \in T$$

Markov süreci, $t = 0$ anında başlamaktadır. Sürecin başladığı andaki olasılıklara, ilk olasılıklar denir ve aşağıdaki biçimde tanımlanır:

$$\pi_i(0) = P(X_0 = S_i) , S_i \in S \quad (3.6)$$

Eşitlik 3.6' da verilen başlangıç durumlarının olasılık dağılımı olan $\pi(0)$ vektörü, ilk olasılık vektörü olarak adlandırılır. Geçiş olasılıkları matrisi ve ilk olasılık vektörü bilindiğinde, Markov zinciri tamamiyle tanımlanmış olur. Bu durumda, zincirin olası gerçekleştirmelerinin olasılıksal hesaplamaları yapılabilir. Zincirin t zamanında S_i durumunda bulunması olasılığı Chapman-Kolmogorov eşitliğinden yararlanılarak aşağıdaki gibi hesaplanabilir:

$$\begin{aligned} \pi_i(t+1) &= \sum_{\forall S_j \in S} P(X_t = S_j) P(X_{t+1} = S_i | X_t = S_j) \\ &= \sum_{\forall S_j \in S} \pi_j(t) P(X_{t+1} = S_i | X_t = S_j) \end{aligned} \quad (3.7)$$

Eşitlik 3.7' de verilen ifade matris formunda, $\pi(t+1) = \pi(t)P$ olarak yazılır. $\pi(t)$ vektörü, $\pi(0)$ ilk olasılık vektörü ve P geçiş olasılıkları matrisi kullanılarak hesaplanabilir. Bu durum, Chapman-Kolmogorov eşitliğinin genel bir sonucu olarak aşağıdaki gibi elde edilir:

$$\begin{aligned}
\pi(t) &= \pi(t-1) P \\
&= (\pi(t-2) P) P = \pi(t-2) P^2 \\
&= (\pi(t-3) P) P^2 = \pi(t-3) P^3 \\
&= (\pi(t-4) P) P^3 = \pi(t-4) P^4 \\
&\vdots \\
\pi(t) &= \pi(0) P^t
\end{aligned}$$

Markov zincirinde durum uzayının elemanları farklı şekillerde sınıflandırılabilir. Herhangi bir $n \geq 0$ sonlu tamsayısı için $p_{ij}^{(n)} > 0$ ise, S_i durumundan S_j durumuna ulaşılabilir. Eğer sonlu sayıda geçişle S_i durumundan S_j durumuna ulaşılabilir ve S_j durumundan da S_i durumuna ulaşılabilir ise, bu iki durum birbiriyle bağlantılıdır [43], [44], [45]. Durum uzayındaki bir ya da daha çok durumdan oluşan bir \mathcal{C} kümesi tanımlansın. \mathcal{C} kümesinin dışındaki bir durumdan bu kümedeki herhangi bir duruma ulaşamıyorsa, \mathcal{C} kümesine kapalı küme denir [43], [45]. Bu durumda, kapalı kümedeki her durum sadece o kümedeki öteki durumlarla bağlantılıdır. Birbiriyle bağlantılı durumlar, durum uzayında aynı kümede yer alır. Eğer kapalı küme tek durum içeriyorsa, bu duruma yutucu (absorbing/trapping) durum denir [43], [45]. S_i , yutucu bir durum ise, $p_{ii} = 1$ ve $p_{ij} = 0$, $i \neq j$ ' dir. Diğer bir ifadeyle, zincir yutucu duruma geldiğinde bu durumdan diğer durumlara asla geçemez ve hep yutucu durumda kalır [45]. Durum uzayındaki tüm durumlar birbiriyle bağlantılı ise, durum uzayı alt kümelere bölünemez ve durumların tümü bir kapalı küme oluşturur. Durum uzayı bölünmeyen zincirlere indirgenemez (irreducible) Markov zinciri denir [43], [44], [45]. Diğer bir ifadeyle, $\forall S_i, S_j \in \mathcal{S}$ için $p_{ij}^{(n)} > 0$ mevcut ise, Markov zinciri indirgenemezdir.

Herhangi bir $t \in \mathcal{T}$ zamanında S_i durumunda olan Markov zincirinin n adımda S_j durumuna ilk geçiş (first passage) olasılığı aşağıdaki şekilde tanımlanır [43], [45]:

$$f_{ij}^{(n)} = P(X_{t+n} = S_j, X_r \neq S_j, 0 < r < (t+n) \mid X_t = S_i)$$

Burada, $i \neq j$ için $f_{ij}^{(0)} = 0$ dır. Ayrıca, $f_{ij}^{(1)} = p_{ij}$ olur. Zincirin S_i durumunda olduğu verildiğinde, S_j durumuna ilk geçiş olasılığı, f_{ij} ; ilk geçiş zamanının beklenen değeri ise, μ_{ij} ile gösterilir. Bunlar,

$$f_{ij} = \sum_{n=1}^{\infty} f_{ii}^{(n)}$$

$$\mu_{ij} = \sum_{n=1}^{\infty} n f_{ii}^{(n)}$$

olarak bulunur [43], [45]. $f_{ij}^{(n)}$ ilk geçiş olasılığına benzer şekilde, herhangi bir $t \in T$ zamanında S_i durumunda olan Markov zincirinin n adımda S_i durumuna ilk dönüş olasılığı aşağıdaki gibi tanımlanır [44], [45]:

$$f_{ii}^{(n)} = P(X_{t+n} = S_i, X_r \neq S_i, 0 < r < (t+n) \mid X_t = S_i) \quad (3.8)$$

Burada, $f_{ii}^{(0)} = 0$ dır ve $f_{ii}^{(1)} = p_{ii}$ olur. Zincirin S_i durumunda olduğu verildiğinde, S_i durumuna geri dönüş olasılığı, f_{ii} ; geri dönüş zamanının beklenen değeri ise, μ_{ii} ile gösterilir. Bunlar,

$$f_{ii} = \sum_{n=1}^{\infty} f_{ii}^{(n)}$$

$$\mu_{ii} = \sum_{n=1}^{\infty} n f_{ii}^{(n)}$$

olarak hesaplanır [44], [45].

Herhangi bir S_i durumunda olan Markov zincirinin S_i durumuna geri dönmesi kesin ($f_{ii} = 1$) ise, bu duruma geri dönüş (recurrent/persistent) durum denir. Geri dönüş durum için geri dönüş zamanının beklenen değeri sonlu ($\mu_{ii} < \infty$) ise, bu duruma artı-geri dönüş (nonnull/positive-recurrent) durum; sonsuz ($\mu_{ii} = \infty$) ise, etkisiz-geri dönüş (null-recurrent) durum denir. Bir duruma geri dönüş olasılığı birden küçük ($f_{ii} < 1$) ise, bu duruma geçiş (transient) durumu denir. S_i durumunda olan zincir S_i durumuna yalnızca $d, 2d, 3d, \dots$ zamanlarında ($d > 1$ ve tamsayı) geri dönüyorsa, S_i durumuna d periyodu/döngüsü ile periyodik/döngüsel (periodic) durum denir. Bir durumun periyodu bir ($d = 1$) ise, bu durum periyodik değildir. Markov zincirinin periyodik durumlar arası geçişi sistematik bir seyir gösterir. Bu nedenle, periyodik durumlar Markov zincirinin hareketini deterministik bir yapıya dönüştürür ve zincir durağan dağılımına yakınsayamaz. Zincirin tüm durumları periyodik değilse, zincir periyodik değildir. Periyodik olmayan ve artı-geri dönüş bir durum, ergodik/döngel (ergodic) durum olarak adlandırılır. Markov zincirinin tüm durumları ergodik ise, bu zincire ergodik Markov zinciri denir [43], [44], [45].

İstatistiksel modellemede stokastik süreçler için durağanlık oldukça önemli bir özelliktir. Özellikle, Bayesci istatistikte Markov zinciri Monte Carlo yöntemleri kullanılarak belli bir dağılımdan örnek değerler çekilmek istendiğinde öncelikle oluşturulan zincirin durağan dağılımına yakınsaması kontrol edilir. Durağan dağılıma yakınsayan bir süreç başlangıç değerlerinden bağımsızdır. Durağan dağılımdan bir örnek elde edildiğinde sonraki tüm örnekler de bu dağılımdan gelir. Daha açık bir ifadeyle, sürecin sahip olduğu dağılım korunur. Bu nedenlerle, yakınsama sağlandıktan sonraki zincir değerleri üzerinden çıkarımlar rahatlıkla yapılabilir. Aksi durumda, yapılacak çıkarımlara güvenilmemelidir.

Aşağıda verilen özellikleri sağlayan bir δ satır vektörü, P geçiş olasılıkları matrisine sahip olan Markov zincirinin durağan (stationary) dağılımıdır [44]:

- i. $\forall S_i \in \mathcal{S}$ için, $\delta_i \geq 0$ 'dır.
- ii. $\sum_{\forall S_i \in \mathcal{S}} \delta_i = 1$ 'dir.
- iii. $\delta_i = \sum_{\forall S_k \in \mathcal{S}} \delta_k p_{ki}$ 'dir. Bu eşitlik, matris gösterimiyle $\delta = \delta P$ biçiminde de ifade edilebilir.

Markov zincirinin birden fazla durağan dağılımı olabilir. Eğer durum uzayı birden fazla kapalı kümeden oluşuyorsa, her küme için farklı bir durağan dağılım mevcuttur. Tek bir durağan dağılım için gerek ve yeter koşul, Markov zincirinin indirgenemez ve ergodik olmasıdır [43], [44]. Çünkü, indirgenemez ve ergodik Markov zincirinde durum uzayı tek bir kapalı kümeden oluşmaktadır.

Markov zincirinin uzun dönem davranışının belirlenmesinde $n \rightarrow \infty$ da zincirin durumları arasındaki hareketleri olasılıksal olarak incelenir. $n \rightarrow \infty$ durumunda zincirin n -inci iterasyondaki olasılık dağılımına, zincirin limit dağılımı denir. Bir Markov zincirinin limit dağılımı, her zaman mevcut değildir. Sonlu durum uzayına sahip bir Markov zinciri düzenli ise, limit dağılımı mevcuttur. Markov zincirinin limit dağılımı mevcut ise, bu dağılım aynı zamanda zincirin durağan dağılımıdır. Fakat, bunun tersi doğru değildir. Periyodik olmayan ve indirgenemez sonlu Markov zinciri düzenlidir. Bu durum gereği, indirgenemez ve ergodik olan sonlu Markov zincirinin limit dağılımı, aynı zamanda zincirin tek olan durağan dağılımına yakınsar. Sonlu Markov zincirinin durum uzayı, $\mathcal{S} = \{S_1, \dots, S_m\}$ kümesi olarak belirlensin. Markov zincirinin limit dağılımı aşağıdaki

biçimde tanımlanır [43], [44], [45]:

$$\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \delta_j > 0, \quad j = 1, 2, \dots, m \quad (3.9)$$

ya da

$$|p_{ij}^{(n)} - \delta_j| \xrightarrow{n \rightarrow \infty} 0, \quad j = 1, 2, \dots, m \quad (3.10)$$

Eşitlik 3.9 ve Eşitlik 3.10' da $n \rightarrow \infty$ için, $p_{ij}^{(n)}$ n -adım geçiş olasılığının zincirin S_i durumunda bulunması başlangıç koşulundan bağımsız olduğu görülür. n adım sayısı yeterince büyük olduğunda, $P^{(n)}$ ' den çekilen örnek değerleri δ durağan dağılımından çekilen örnek değerlerdir ve zincirin başlangıçta bulunduğu durumun bir önemi yoktur. Limit dağılımının durağan dağılıma yakınsaması matris formunda aşağıdaki gibi ifade edilir:

$$\lim_{n \rightarrow \infty} P^{(n)} = \begin{matrix} 1 \\ 2 \\ 3 \\ \vdots \\ m \end{matrix} \begin{bmatrix} \delta \\ \delta \\ \delta \\ \vdots \\ \delta \end{bmatrix}$$

İndirgenemez ve ergodik sonlu Markov zincirinin $\delta = [\delta_i]_{1 \times m}$ durağan dağılımı aşağıdaki şekilde elde edilebilir [44], [45]:

$$\delta_i = \frac{1}{\mu_{ii}}, \quad j = 1, 2, \dots, m \quad (3.11)$$

Eşitlik 3.11' de μ_{ii} , zincirin S_i durumuna geri dönüş zamanının beklenen değeridir.

Tez çalışmasında, beta dağılım ailesi varsayımı altında incelenen SMM' nin parametre tahminlerinin analitik olarak elde edilmesi, modelin yapısından dolayı mümkün değildir. Bu nedenle, klasik yaklaşıma dayalı olarak SMM' nin çözümlenmesi, beta dağılım ailesinden gelen gözlemler için uygun değildir. Burada karşılaşılan analitik çözümsüzlüğün aşılmasında Bayesci yaklaşım tercih edilmiştir. Ancak, Beta-SMM' nin Bayesci çözümlenmesinde ağır integrasyon sorunlarıyla karşılaşmıştır. Tezin izleyen bölümünde SMM' de karşılaşılan integrasyon problemlerinin aşılmasında kullanılan stokastik benzetim yöntemleri tanıtılacaktır.

4. STOKASTİK BENZETİM YÖNTEMLERİ

Bayesci istatistikte, normalleştirme katsayısının hesaplanması, marjinal sonsal dağılımların elde edilmesi ve bu dağılımlardan istatistiksel çıkarımlar yapılması genellikle karmaşık ve çok boyutlu fonksiyonların integrallerinin hesaplanmasını gerektirir. Bu integrallerin analitik çözümü her zaman mümkün değildir ve bu durum integrasyon problemi oluşturur. Özellikle, parametre sayısının artması ve eşlenik olmayan önsel dağılımların kullanılmasıyla sonsal dağılım daha karmaşık bir hale gelir. Bu nedenle, Bayesci çözümlemede sıklıkla integrasyon problemiyle karşılaşılır.

Bu problemin aşılmasında, sayısal integrasyon yöntemleri kullanılır. Bu yöntemler, deterministik yöntemler ve benzetim yöntemleri olarak ikiye ayrılabilir. Deterministik yöntemlerde, integralin boyut sayısı arttıkça kullanılan yöntemle ilişkin hata değeri de artar ve elde edilen sayısal sonuçların doğruluğu önemli oranda düşer [33], [46]. Bu durumda, örnekleme temelli yöntemler kullanılarak ilgili integralin yaklaşık değerine ulaşılır. Benzetim yöntemlerinde, integralin boyut sayısına bakılmaksızın, çekilen örneklem büyüklüğü artırılarak yöntemle ilişkin hata oranı düşürülür.

Tez çalışmasında, Beta-SMM'nin Bayesci çözümlemesinde, rasgele örneklemeyle dayalı olan standart Monte Carlo yöntemi ve Markov zincirlerine dayalı olan Markov Zinciri Monte Carlo yöntemleri kullanılmıştır. Bu nedenle, izleyen bölümlerde bu yöntemler anlatılmıştır.

4.1. Monte Carlo Yöntemi

Monte Carlo yöntemi, karmaşık ve çok boyutlu fonksiyonların integrallerinin hesaplanması amacıyla, Metropolis ve Ulam [47] ve Von Neumann [48] tarafından ortaya konulmuştur [46] [49]. Yöntemde temel amaç,

$$E_f[h(\mathbf{X})] = \int_{\mathcal{R}_X} h(\mathbf{x})f(\mathbf{x})d\mathbf{x} \quad (4.1)$$

şeklinde tanımlanan integralin çözümünün yaklaşık değerinin hesaplanmasıdır [39], [50]. $f(\mathbf{x})$, \mathcal{R}_X aralığında tanımlı \mathbf{X} raslantı değişkeninin olasılık yoğunluk fonksiyonu ve $h(\mathbf{x})$, \mathbf{X} raslantı değişkeninin herhangi bir fonksiyonudur. Monte Carlo yöntemi, $f(\mathbf{x})$ olasılık yoğunluk fonksiyonundan çekilen $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ rasgele örnekleme dayanmak-

tadır. Eşitlik 4.1' de verilen integralin çözümünün yaklaşık değeri

$$\bar{h}_n = \frac{1}{n} \sum_{i=1}^n h(\mathbf{x}_i) \quad (4.2)$$

şeklinde tanımlanan örneklem ortalaması olarak elde edilir [33], [39], [50], [51]. Çekilen örnekler bağımsız olduğu için, $n \rightarrow \infty$ durumunda güçlü büyük sayılar yasası gereği elde edilen \bar{h}_n örneklem ortalaması $E_f[h(\mathbf{X})]$ kitle ortalamasına aşağıdaki gibi hemen hemen kesinlikle yakınsar.

$$\bar{h}_n \xrightarrow{n \rightarrow \infty} E_f[h(\mathbf{X})]$$

$E_f[h^2(\mathbf{X})] < \infty$ olduğunda, \bar{h}_n Monte Carlo tahmin edicisinin standart hatası hesaplanabilir. Bu durumda, Monte Carlo yönteminin hatası ölçülerek, \bar{h}_n ' in yakınsama hızı değerlendirilir. \bar{h}_n tahminine ilişkin standart hata değeri aşağıdaki gibi elde edilir [33], [39]:

$$V(\bar{h}_n) = \frac{V_f[h(\mathbf{X})]}{n} = \frac{\int_{\mathcal{R}_x} (h(\mathbf{x}) - E_f[h(\mathbf{X})])^2 f(\mathbf{x}) d\mathbf{x}}{n} \quad (4.3)$$

$\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ bağımsız örnekleme kullanılarak, Eşitlik 4.3' te verilen standart hata değeri,

$$\hat{V}(\bar{h}_n) = \frac{1}{n} \left(\frac{1}{n-1} \sum_{i=1}^n (h(\mathbf{x}_i) - \bar{h}_n)^2 \right)$$

olarak tahmin edilir [34], [51], [52]. Merkezi limit teoremi gereği, $n \rightarrow \infty$ olduğunda aşağıdaki durum elde edilir.

$$\frac{\bar{h}_n - E_f[h(\mathbf{X})]}{\sqrt{\hat{V}(\bar{h}_n)}} \xrightarrow{n \rightarrow \infty} N(0, 1) \quad (4.4)$$

Eşitlik 4.4' te verilen merkezi limit teoreminin bir sonucu olarak, $E_f[h(\mathbf{X})]$ için %100(1- α) güven aralığı aşağıda verilmiştir [34], [50]:

$$\left(\bar{h}_n - Z_{\alpha/2} \sqrt{\hat{V}(\bar{h}_n)}, \bar{h}_n + Z_{\alpha/2} \sqrt{\hat{V}(\bar{h}_n)} \right)$$

Burada, $Z_{\alpha/2}$, $\frac{\alpha}{2}$ önem düzeyindeki standart normal dağılım tablo değeridir. Monte Carlo integrasyonu, $\mathcal{O}(1/\sqrt{n})$ yakınsaklığı göstermektedir [46], [51]. Bu nedenle, hatayı yarıya indirmek için örneklem büyüklüğünü dört katına çıkartmak gerekir.

Eşitlik 4.1' de verilen integralde $f(x)$ olasılık yoğunluk fonksiyonundan rasgele örneklem çekmek her zaman kolay olmayabilir. Bu durumda, Monte Carlo yöntemi yerine önem örnekleme tercih edilir. Bunun için, tanım aralıkları aynı olacak şekilde $f(x)$ ' e kabaca benzeyen ve örneklem çekmenin daha kolay olduğu bir olasılık yoğunluk fonksiyonu gerekmektedir. Bu olasılık yoğunluk fonksiyonuna, önem fonksiyonu denir. İlgili integralin çözümünün tahmin edilmesinde önem örnekleme başarısı belirlenen önem fonksiyonuna bağlıdır. $q(x)$, önem fonksiyonu olarak kabul edilsin. Eşitlik 4.1' de verilen integralin tahminine aşağıdaki gibi ulaşılır.

$$\int_{\mathcal{R}_X} h(\mathbf{x})f(\mathbf{x})d\mathbf{x} = \int_{\mathcal{R}_X} h(\mathbf{x}) \left\{ \frac{f(\mathbf{x})}{q(\mathbf{x})} \right\} q(\mathbf{x})d\mathbf{x} = E_q \left[\frac{h(\mathbf{X})f(\mathbf{X})}{q(\mathbf{X})} \right] \cong \frac{1}{n} \sum_{i=1}^n \frac{h(\mathbf{x}_i)f(\mathbf{x}_i)}{q(\mathbf{x}_i)} \quad (4.5)$$

Eşitlik 4.5' te kullanılan $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ bağımsız örnekleme $q(x)$ önem fonksiyonundan çekilmiştir.

4.2. Markov Zinciri Monte Carlo Yöntemleri

Monte Carlo yönteminde, rasgele örnekleme yapılacak olan $f(x)$ dağılımından doğrudan örneklem çekmek her zaman mümkün olmayabilir. Bu durumda, alternatif bir yöntem olan önem örnekleme kullanılarak Eşitlik 4.1' de tanımlanan integralin çözümü tahmin edilebilir. Fakat, önem örnekleme etkinliği önem fonksiyonunun seçimine bağlıdır. Boyut sayısı arttıkça, önem fonksiyonunun seçimi giderek zorlaşır; hatta $f(x)$ ' i temsil edecek uygun bir dağılım bulunamayabilir. Bu nedenle, $f(x)$ ' ten rasgele örneklem yerine bağımlı örneklem üreten benzetim yöntemlerine başvurulur. Bu yöntemler, Markov zinciri benzetimine dayalı olan Markov Zinciri Monte Carlo (MCMC) yöntemleridir. MCMC yöntemleri ile karmaşık ve çok boyutlu sonsal dağılımlardan kolaylıkla örneklem çekilebilmektedir. Çekilen bağımlı örnek değerlerden yararlanılarak sonsal dağılımlar üzerinden istenilen istatistiksel çıkarımlar yapılır.

Eşitlik 4.1' de yer alan $f(x)$ dağılımından doğrudan rasgele örneklem çekilemediğinden dolayı, MCMC yöntemlerinden yararlanılarak durağan dağılımı $f(x)$ olan indirgenemez ve ergodik bir $\{\mathbf{X}^{(t)}\}_{t \geq 0}$ Markov zinciri oluşturulur. Sürecin parametresi olan t , burada iterasyon sayısını göstermektedir. Bu nedenle, üs olarak yazılmıştır. $t \rightarrow \infty$ olduğunda

oluşturulan Markov zinciri, dağılımda \mathbf{X} raslantı değişkenine yakınsar [53].

$$\mathbf{X}^{(t)} \xrightarrow[d]{t \rightarrow \infty} \mathbf{X} \sim f(\mathbf{x}) \quad (4.6)$$

Burada, d. harfi, dağılımda yakınsamayı ifade etmektedir.

İndirgenemez ve ergodik Markov zincirinden gelen $\{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(T)}\}$ bağımlı örneklem değerleri aşağıda verilen Ergodik teoremi sağlar [33], [46], [50]:

$$E_f[|h(\mathbf{X})|] < \infty \implies \bar{h}_T = \frac{1}{T} \sum_{t=1}^T h(\mathbf{x}^{(t)}) \xrightarrow[h.h.k.]{T \rightarrow \infty} E_f[h(\mathbf{X})] \quad (4.7)$$

Eşitlik 4.7' de T , toplam iterasyon sayısıdır ve \bar{h}_T , ergodik ortalama olarak adlandırılır. Ergodik ortalama, $T \rightarrow \infty$ durumunda Eşitlik 4.1' de verilen integralin çözümüne hemen hemen kesinlikle yakınsar. Bu yakınsaklık, h.h.k. ile kısaca ifade edilmiştir.

MCMC yöntemleriyle birden fazla Markov zinciri oluşturularak bağımsız örneklem elde edilebilir. Böylece, bağımsız örneklem üzerinden tahminler yapılabilir. Her biri T_0 uzunluğunda n tane bağımsız Markov zinciri oluşturulsun. Her bir zincirden T_0 -ıncı değerler alınsın. Bu değerlerle oluşturulan $\mathbf{x}^{(T_0)} = (\mathbf{x}_{T_0}^{(1)}, \mathbf{x}_{T_0}^{(2)}, \dots, \mathbf{x}_{T_0}^{(n)})$, bağımsız örnektir. $\mathbf{x}^{(T_0)}$ bağımsız örneği kullanarak Eşitlik 4.1' de verilen integralin çözümünün yaklaşık değeri aşağıdaki gibi hesaplanır [33], [50]:

$$\bar{h}_{n, T_0} = \frac{1}{n} \sum_{i=1}^n h(\mathbf{x}_{T_0}^{(i)}) \quad (4.8)$$

Burada, nT_0 değerden $n(T_0 - 1)$ tanesi kullanılmadığından dolayı önemli bir zaman kaybı oluşur. Zaman maliyeti düşünüldüğünde, ergodik teoreme dayanılarak tek bir Markov zinciri gerçekleştirilerek ilgili integralin çözümünün iyi bir tahmin değeri elde edilir. Çünkü, ergodik teorem, bağımsız örneklem varsayımı altında geçerli olan güçlü büyük sayılar yasasının bağımlı olan örneklem için karşılığıdır.

MCMC yöntemlerinin uygulanmasında, uygun başlangıç değerlerinin belirlenmesi, başlangıç değer etkilerinin yok edildiği (burn-in) iterasyon sayısına karar verilmesi, simülasyonun ne kadar uzunlukta çalıştırılması gerektiğinin saptanması ve hedef dağılıma yakınsamanın tespit edilmesi önemli konulardır.

MCMC yöntemlerinde oluşturulan zincir indigenemez olması nedeniyle, başlangıç değerlerinin durağan dağılımı üzerinde etkisi yoktur ve herhangi bir değer olabilir. Başlangıç değerleri, sadece zincirin performansını ve durağan dağılıma yakınsama hızını etkiler. Zincirin performansı, hızlı veya yavaş karışmasına göre değerlendirilir. MCMC yöntemleri uygulandığına, iki çeşit zincir oluşur. Bunlar, zayıf ve karma zincirlerdir. Zayıf zincirde, iterasyonlarda üretilen değerler hedef dağılımın tanım kümesinin belli değerlerinde sabitlenmektedir ve/veya tanım kümesinin sadece belli bölgelerinde hareketlilik göstermektedir. Bu nedenle, üretilen zincir değerleri hedef dağılımı iyi bir şekilde temsil edemez ve zayıf zincirin performansı düşük olur. Karma zincirde ise, iterasyonlarda üretilen değerler hedef dağılımın tanım kümesini kapsayacak şekilde rasgele hareket etmektedir. Bu nedenle, karma zincirdeki değerler hedef dağılımı daha iyi temsil eder ve zincirin performansı yüksek olur. Hızlı karışan (fast-mixing) zincirlerde, başlangıç değerlerinin etkileri önemli değildir. Çünkü, bu zincirler başlangıç değerlerinin etkilerinden hızlı bir şekilde kurtulmaktadır. Yavaş karışan (poor-mixing) zincirlerde ise, başlangıç değerleri dikkatli seçilmelidir. Çünkü, başlangıç değerlerinin etkilerinin yok edilmesi için bu zincirlerde çok fazla iterasyona ihtiyaç duyulabilir ve zincirlerin durağan dağılımlarına yakınsamaları oldukça yavaşlayabilir.

Zincirin nasıl karıştığına görülmesi ve yakınsama durumunun görsel olarak değerlendirilmesi amacıyla iz grafiği (traceplot), hareketli ortalama grafiği (running mean plot) ve otokorelasyon fonksiyonu grafiğinden yararlanılır. İz grafiği, üretilen zincir değerlerinin ilgili oldukları iterasyon sayısına karşı grafiğidir. İz grafiği sayesinde, üretilen değerlerin karma veya zayıf bir zincir oluşturduğu rahatlıkla görülebilir. Hareketli ortalama grafiği, her bir iterasyon değerine karşılık ilgili iterasyona kadar üretilen zincir değerlerinin ortalamalarının grafiğidir. Bu grafikte ortalamaların kararlı hale geldiği iterasyon sayısına bakılır. Belirlenen bu iterasyon sayısında yakınsamanın gerçekleştiği ve başlangıç değer etkilerinin yok olduğu kabaca söylenebilir. Otokorelasyon grafiği ise, üretilen değerler arasındaki bağımlılık yapısını göstermektedir. MCMC yöntemlerinde bağımlı örneklem üretilmesine rağmen; gecikmelere ilişkin otokorelasyon değerlerinin büyük olması zincirin iyi karışmadığını ve zayıf bir zincir olduğunu göstermektedir. Düşük otokorelasyon değerleri ise, zincirin performansının yüksek olduğunu gösterir. Yakınsama problemi bulunmayan zincirlerde otokorelasyon değerlerinin düşük olması beklenir. Zincirin yüksek bağımlılık içerdiği durumda durağan dağılıma yakınsamanın sağlanması için, uzun bir zincir oluşturulmalıdır. Zaman maliyeti ve depolama

alanı kısıtı düşünülduğünde, yüksek otokorelasyon değerleri istenmeyen bir durumdur. Özellikle, uzun dönem bağımlılığı gösteren yüksek gecikme değerlerinde otokorelasyon değerlerinin düşük düzeyde olması beklenir. Otokorelasyonu düşürmek amacıyla zincirde seyreltme (thinning) yapılabilir. Seyreltme işleminde, üretilen zincir değerlerinin sadece k -ıncı ($k \geq 1$ ve pozitif tamsayı) örnek değerleri korunup diğer örnek değerler atılmaktadır. Burada k değerine otokorelasyon fonksiyonuna bakılarak karar verilir. Zincirin seyreltilmesi depolama alanı sorununa çözüm getirmesine rağmen, zaman maliyeti açısından bir dezavantaj sağlayabilir. Çünkü, üretilen zincirin $\% \left[\frac{k-1}{k} \right]$ sı atılıp göz ardı edilmektedir. Bu durum, yeterli sayıda örneklem büyüklüğüne ulaşmaya dek uzun bir zincir oluşturulmasına sebebiyet verir.

Zincirin durağan dağılıma yakınsamasının tespitinde görsel denetimin yanısıra yakınsaklık testleride (convergence diagnostics) kullanılır. Görsel denetim yaparak başlangıç değer etkileri arıdıktan sonraki iterasyonlar üzerinden zincirin durağan dağılıma ulaşıp ulaşmadığı kontrol edilir. Yakınsamanın tespitinde, MCMC yöntemlerinde üretilen zincir sayısı önemlidir. Bazı testler tek zincir varlığında; bazıları ise, birden çok zincir varlığında gerçekleştirilmektedir. Tek ya da birden fazla paralel zincir üretmenin avantaj ve dezavantajları vardır. Örneğin, üretilen tek bir uzun zincirin sonunda durağan dağılıma yakınsama, aynı anda üretilen kısa zincirlerinkine kıyasla, daha iyi olacaktır. Bu durumda, birçok kısa zincir üretimi zaman kaybına neden olabilir. Buna karşın, tek bir zincir er ya da geç tüm tanım kümesinde hareketlilik gösterecek olmasına rağmen; belirlenen iterasyon sayısına bağlı olarak zincirin tanım kümesinin önemli bir kısmında hareketlilik gösterememe ihtimali mevcuttur. Bu durumda, farklı başlangıç değerleri ile aynı anda üretilen birçok zincir sayesinde bu ihtimal ortadan kaldırılabilir. Ayrıca, zincirlerin tanım kümesindeki hareketlilikleri rahatlıkla incelenerek karışma durumlarında kolayca belirlenebilir.

Aşağıdaki tabloda yakınsaklık testlerine ilişkin özet bilgilere yer verilmiştir [54]. Tabloda testlerin kullanılmasında gerekli zincir sayısı, raslantı değişkeninin boyut sayısı (tek/çok değişkenli) ve hangi MCMC yöntemlerinde uygulanabilir oldukları belirtilmiştir.

Çizelge 4.1. Yakınsaklık Testleri

Yakınsaklık Testi	Zincir Sayısı	Boyut Sayısı	Uygulanabilirlik
Gelman-Rubin (1992)	Birden fazla	Tek değişkenli	Tüm MCMC
Raftery-Lewis (1992)	Tek	Tek değişkenli	Tüm MCMC
Geweke (1992)	Tek	Tek değişkenli	Tüm MCMC
Roberts (1992,1994)	Birden fazla	Çok değişkenli	Bazı MCMC
Ritter-Tanner (1992)	İkisinde	Çok değişkenli	Sadece Gibbs
Zellner-Min (1995)	Tek	Çok değişkenli	Bazı MCMC
Liu-Liu-Rubin (1992)	Birden fazla	Çok değişkenli	Sadece Gibbs
Garren ve Smith (1993)	Birden fazla	Tek değişkenli	Sadece Gibbs
Johnson (1994)	Birden fazla	Çok değişkenli	Bazı MCMC
Heidelberger-Welch (1983)	Tek	Tek değişkenli	Tüm MCMC
Mykland-Tierney-Yu (1995)	Tek	Çok değişkenli	Bazı MCMC
Yu (1994)	Tek	Çok değişkenli	Bazı MCMC
Yu-Mykland (1994)	Tek	Tek değişkenli	Tüm MCMC

Tez çalışmasında incelenen Beta-SMM' nin işletiminde her parametre için tek zincir üretilmiştir. Ardından, üretilen her bir zincirin durağan dağılımına yakınsaklığının kontrolünde ise, Geweke [55] ve Heidelberger-Welch [56] yakınsaklık testleri kullanılmıştır. Bu nedenle, bu testler aşağıda kısaca anlatılmıştır. Geweke ve Heidelberger-Welch yakınsaklık testlerinin uygulanmasında da R programında bulunan **CODA** paketinden yararlanılmıştır.

Geweke'nin yakınsaklık testinde herhangi bir MCMC yönteminden üretilen zincirin baştan n_A ve sondan n_B genişliklerinde örneklem değerleri alınarak bu iki örneklemin ortalamaları karşılaştırılır. Ortalamalar arasında istatistiksel olarak anlamlı bir farklılık yoksa, yakınsama probleminin olmadığı ve bu iki örneklem değerlerinin zincirin durağan dağılımından çekildiği sonucuna varılır. Testin gerçekleştirilmesinde kullanılan test istatistiğine Geweke Z-skoru adı verilmiştir. Geweke Z-skorunun elde edilmesi için zincirin başından ve sonundan alınan örneklemelerin ortalamaları sırayla aşağıdaki gibi hesaplanır [57]:

$$\bar{X}_A = \frac{1}{n_A} \sum_{t=1}^{n_A} x_t \quad \text{ve} \quad \bar{X}_B = \frac{1}{n_B} \sum_{t=N-n_B+1}^N x_t$$

Burada, N zincirin toplam uzunluğudur ve $(n_A + n_B) < N$ dir. (n_A/N) ve (n_B/N) oranları sabit ve $\{X_t, t = 1, 2, \dots, N\}$ zinciri durağan ise, Z_G Geweke Z-skoru asimtotik olarak standart normal dağılıma sahiptir [57]:

$$Z_G = \frac{(\bar{X}_A - \bar{X}_B)}{\sqrt{\frac{1}{n_A} \hat{S}_X^A(0) + \frac{1}{n_B} \hat{S}_X^B(0)}} \xrightarrow[N \rightarrow \infty]{d.} N(0, 1)$$

$\hat{S}_X^A(0)$ ve $\hat{S}_X^B(0)$, sırasıyla $\{x_1, \dots, x_{n_A}\}$ ve $\{x_{N-n_B+1}, \dots, x_N\}$ örneklemelerinin spektral yoğunluk kestirimleridir. $|Z_G| > Z_{\alpha/2}$ ise, $\%100(1 - \alpha)$ güvenirlikle ortalamalar arasında istatistiksel olarak anlamlı bir fark söz konusu olur. Bu durumda, durağan dağılıma yakınsamanın $t = n_A$ zamanına kadar henüz sağlanmadığına karar verilir ve zincirin $\%100(n_A/N)$ ' lık kısmı atılarak geriye kalan kısım üzerinden yeniden Geweke yakınsama testi uygulanır. Üretilen zincirin en fazla ilk $\%40$ ' lık kısmı atılabilir. Aksi durumda, daha fazla iterasyon gerçekleştirilerek zincirin uzatılması gerekmektedir. Geweke [55], testin uygulanmasında üretilen zincirin ilk $\%10$ ' luk ($n_A = 0.1N$) ve son $\%50$ ' lik ($n_B = 0.5N$) kısımlarının alınmasını önermiştir.

Heidelberger-Welch yakınsaklık testi ise, iki testten oluşmaktadır. İlki durağanlık testi; ikincisi ise, yarı genişlik testidir. Durağanlık testi uygulanmadan yarı genişlik testi uygulanamaz. Durağanlık testinde, “*örneklem değerleri Markov zincirinin durağan dağılımından gelmektedir.*” yokluk hipotezi test edilir. Bunun için, Cramer-von-Mises test istatistiğinden yararlanılır. Durağanlık testi önce zincirin tamamına uygulanır. Şayet yokluk hipotezi reddedilir ise, yokluk hipotezi kabul edilene kadar zincirin ilk $\%10$, $\%20$, $\%30$, $\%40$ ya da en çok ilk $\%50$ ' lik kısmı atılana kadar test tekrar edilir. Zincirin ilk $\%50$ ' lik kısmı atılmasına rağmen, yokluk hipotezi kabul edilemez ise, daha fazla iterasyon gerçekleştirilerek zincirin uzatılması gerekmektedir. Durağanlık testi aşaması geçildikten sonra, zincirin bu testi geçen kısmına yarı genişlik testi uygulanır. Yarı genişlik testinde, ortalamanın belirlenen bir doğrulukta tahmin edilebilmesi için zincirin kabul edilebilir uzunlukta olup olmadığı incelenir. İlk olarak bir ϵ (genellikle 0.1) değeri belirlenir. Ardından ortalamanın $\%100(1 - \alpha)$ güven aralığının yarı genişliği hesaplanır. En son olarak, yarı genişlik değeri ortalamanın tahminine bölünerek bir oran elde edilir. Bu oran, ϵ değeri ile karşılaştırılır. Elde edilen oran daha küçük ise, yarı genişlik testi geçilir. Bu durumda, ortalamanın ϵ kesinliğinde tahmininde zincir kabul edilebilir uzunluktadır. Yarı genişlik testi geçilemediğinde, ϵ değeri arttırılabilir ya da daha fazla iterasyon gerçekleştirilerek zincir uzatılır ve Heidelberger-Welch yakınsaklık testi yeniden yapılabilir.

SMM'nin Bayesci çözümlemesinde model parametrelerinin sonsal dağılımı karmaşık ve çok boyutlu integraller içermektedir. Tez çalışmasında, model parametrelerinin tahmin edilmesi amacıyla sonsal dağılımdan örnekler çekilirken MCMC yöntemlerinden yararlanılmıştır. MCMC yöntemleri arasında en çok kullanılanlar, Metropolis-Hastings algoritması ve Gibbs örneklemesidir. SMM' de elde edilen sonsal dağılımın yapısı Gibbs örneklemesinin gerçekleştirilebilmesi için uygundur. Gerçekleştirilen Gibbs örneklemesinin içinde Metropolis-Hastings algoritmasına ihtiyaç duyulmuştur. Bu nedenle, bu MCMC yöntemleri çalışmanın alt bölümlerinde anlatılmıştır.

4.2.1. Metropolis-Hastings Algoritması

Metropolis-Hastings algoritması, Metropolis vd. [58] tarafından geliştirilen Metropolis algoritmasının geliştirilmiş hali olarak, ilk olarak Hastings [59] tarafından tanıtılmıştır. Bu algorithmada temel amaç, doğrudan rasgele örneklem çekmenin mümkün olmadığı, karmaşık ve çok boyutlu $f(\mathbf{x})$ dağılımından örneklem çekmektir. Bunun için, durağan dağılımı $f(\mathbf{x})$ olan ergodik bir $\{X^{(t)}\}_{t \geq 0}$ Markov zinciri yaratılır.

Burada, $f(\mathbf{x})$, hedef dağılım (target distribution) olarak adlandırılır. Hedef dağılımdan doğrudan örneklem çekilemediği için, bu dağılıma benzeyen ve örneklem çekmenin daha kolay olduğu bir öneri dağılımı (proposal distribution) belirlenir. Oluşturulacak zincirinin aday noktaları, öneri dağılımı kullanılarak üretilir. Metropolis-Hastings algoritmasının performansı, öneri dağılımının seçimine bağlıdır. Algoritmanın performansı iyi ise, karma bir zincir; kötü ise, zayıf bir zincir elde edilir. Öneri dağılımın yayılımı, oluşturulan zincirin davranışını etkiler [60].

Metropolis-Hastings algoritmasında, oluşturulacak Markov zincirinin durağan dağılımı bilinmemektedir; ancak bu durağan dağılımını oluşturan P geçiş olasılıkları dağılımı bilinmemektedir. Bu nedenle, P' nin tanımlanması gerekir. Herhangi bir t anında Markov zinciri $X^{(t)} = \mathbf{x}$ değerini alsın. $X^{(t+1)}$ zincir değeri için aday nokta $q(\cdot | \mathbf{x})$ öneri dağılımından çekilir. Çekilen \mathbf{x}^* aday nokta, $\alpha(\mathbf{x}, \mathbf{x}^*)$ olasılığı ile $X^{(t+1)} = \mathbf{x}^*$ olarak kabul edilir. Aksi durumda, zincir aynı değerde kalır ve $X^{(t+1)} = \mathbf{x}$ olur. Bu durumda, \mathbf{x} zincir değerinden \mathbf{x}^* aday değerine geçiş olasılığı aşağıdaki şekilde tanımlanır [33], [46], [53], [60], [61]:

$$P(\mathbf{x}^* | \mathbf{x}) = q(\mathbf{x}^* | \mathbf{x}) \alpha(\mathbf{x}, \mathbf{x}^*) + \mathbb{1}(\mathbf{x} = \mathbf{x}^*) (1 - r(\mathbf{x})) \quad (4.9)$$

Burada, $1(\cdot)$ gösterge fonksiyonudur ve aşağıdaki gibi tanımlıdır.

$$1(\mathbf{x} = \mathbf{x}^*) = \begin{cases} 1, & \mathbf{x} = \mathbf{x}^* \text{ ise,} \\ 0, & \mathbf{x} \neq \mathbf{x}^* \text{ ise.} \end{cases}$$

$r(\mathbf{x}) = \int_{\xi} q(\xi | \mathbf{x}) \alpha(\mathbf{x}, \xi) d\xi$ dir. Eşitlik 4.9' da verilen P geçiş olasılıkları dağılımı, $f(\mathbf{x})$ hedef dağılımı ile birlikte

$$f(\mathbf{x})P(\mathbf{x}^* | \mathbf{x}) = f(\mathbf{x}^*)P(\mathbf{x} | \mathbf{x}^*) \quad (4.10)$$

tersine çevrilebilirlik (reversibility) koşulunu sağlar. P 'nin tersine çevrilebilirlik özelliği kullanılarak,

$$\int P(\mathbf{x}^* | \mathbf{x}) f(\mathbf{x}) d\mathbf{x} = \int P(\mathbf{x} | \mathbf{x}^*) f(\mathbf{x}^*) d\mathbf{x} = f(\mathbf{x}^*) \int P(\mathbf{x} | \mathbf{x}^*) d\mathbf{x} = f(\mathbf{x}^*) \quad (4.11)$$

eşitliğine ulaşılır. Dolayısıyla, $f(\mathbf{x})$, P için durağan dağılımdır. Eşitlik 4.10 sağlandığında, Markov zinciri durağanlığa ulaşır ve üretilen zincir değerleri $f(\mathbf{x})$ hedef dağılımından örneklerle karşılık gelir. Burada, zincirin durağanlığa ulaşması için gereken adım sayısının belirlenmesi çok önemlidir. Yakınsama testleri ile zincirin durağanlığa ulaşıp ulaşmadığı kontrol edilir. k , başlangıç değerlerinin etkilerinin yitirilmesinde gerekli adım sayısı ise, $(\mathbf{x}^{(k+1)}, \mathbf{x}^{(k+2)}, \dots, \mathbf{x}^{(k+n)})$ bağımlı örnekleme hedef dağılımdan çekilmiş n birimlik örnekleme olur.

Eşitlik 4.9' da verilen geçiş olasılıkları kullanılarak, $\mathbf{x} \neq \mathbf{x}^*$ için tersine çevrilebilirlik koşulu aşağıdaki gibi yazılabilir:

$$f(\mathbf{x})q(\mathbf{x}^* | \mathbf{x})\alpha(\mathbf{x}, \mathbf{x}^*) = f(\mathbf{x}^*)q(\mathbf{x} | \mathbf{x}^*)\alpha(\mathbf{x}^*, \mathbf{x}) \quad (4.12)$$

Öneri dağılımından aday değerlerin üretilme aşamasında aşağıda yer alan iki durumla karşılaşılır:

$$\text{Durum 1 : } f(\mathbf{x})q(\mathbf{x}^* | \mathbf{x}) > f(\mathbf{x}^*)q(\mathbf{x} | \mathbf{x}^*),$$

$$\text{Durum 2 : } f(\mathbf{x})q(\mathbf{x}^* | \mathbf{x}) < f(\mathbf{x}^*)q(\mathbf{x} | \mathbf{x}^*).$$

Eşitlik 4.12 sağlanacak şekilde, yukarıda verilen durumlar için $\alpha(\mathbf{x}, \mathbf{x}^*)$ ve $\alpha(\mathbf{x}^*, \mathbf{x})$ kabul olasılıkları tanımlanır. Birinci durumda, zincirde \mathbf{x}' ten \mathbf{x}^* ' a geçişler, \mathbf{x}^* ' dan \mathbf{x}' e

geçişlerden daha fazladır. Eşitlik 4.12' ün sağlanması için, $\alpha(\mathbf{x}^*, \mathbf{x})$ kabul olasılığı olabildiğince büyük tanımlanmalıdır. Bu nedenle, $\alpha(\mathbf{x}^*, \mathbf{x}) = 1$ olarak alınır ve $\alpha(\mathbf{x}, \mathbf{x}^*) = [\pi(\mathbf{x}^*)q(\mathbf{x} | \mathbf{x}^*)/\pi(\mathbf{x})q(\mathbf{x}^* | \mathbf{x})]$ olarak elde edilir. İkinci durumda ise, zincirde \mathbf{x}' ten $\mathbf{x}^{*'} a$ geçişler, $\mathbf{x}^{*'} dan$ $\mathbf{x}' e$ geçişlerden daha azdır. Bu kez, $\alpha(\mathbf{x}, \mathbf{x}^*) = 1$ olarak alınır ve $\alpha(\mathbf{x}^*, \mathbf{x}) = [\pi(\mathbf{x})q(\mathbf{x}^* | \mathbf{x})/\pi(\mathbf{x}^*)q(\mathbf{x} | \mathbf{x}^*)]$ olarak elde edilir.

Yukarıda verilen iki durum dikkate alındığında, \mathbf{x}^* aday noktasının kabul edilmesi olasılığı aşağıdaki gibi hesaplanır [34], [50], [60], [62]:

$$\alpha(\mathbf{x}, \mathbf{x}^*) = \begin{cases} \min \left\{ 1, \frac{f(\mathbf{x}^*)q(\mathbf{x} | \mathbf{x}^*)}{f(\mathbf{x})q(\mathbf{x}^* | \mathbf{x})} \right\} & , f(\mathbf{x})q(\mathbf{x}^* | \mathbf{x}) > 0, \\ 1 & , f(\mathbf{x})q(\mathbf{x}^* | \mathbf{x}) = 0. \end{cases} \quad (4.13)$$

Sonuç olarak, Metropolis-Hastings algoritmasının işleyişi aşağıda verilen adımlarla ifade edilebilir [34], [35], [36], [60], [63]:

Adım 1: $f(\mathbf{x}^{(0)}) > 0$ olacak şekilde herhangi bir $\mathbf{x}^{(0)}$ başlangıç değeri belirlenmesi,

Adım 2: İterasyon sayacı $t = 0$ olarak atanması,

Adım 3: $q(\cdot | \mathbf{x}^{(t)})$ öneri dağılımından bir \mathbf{x}^* aday değeri seçilmesi,

Adım 4: $[0, 1]$ aralığında tanımlı tekbiçimli dağılımdan rasgele bir u değeri seçilmesi,

Adım 5: \mathbf{x}^* aday değeri için, $\alpha(\mathbf{x}^{(t)}, \mathbf{x}^*)$ kabul edilme olasılığının hesaplanması,

Adım 6: $u \leq \alpha(\mathbf{x}^{(t)}, \mathbf{x}^*)$ ise, $\mathbf{X}^{(t+1)} = \mathbf{x}^*$; aksi durumda, $\mathbf{X}^{(t+1)} = \mathbf{x}^{(t)}$ değerinin atanması ve

Adım 7: $t = 1, 2, \dots, N$ için 3. adıma dönülmesidir.

Yukarıda verilen algoritmanın çalıştırılmasıyla $(\mathbf{x}^{(0)}, \mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(N)})$ bağımlı örnekleme elde edilir. Burada, k -ncı iterasyon itibariyle başlangıç değeri etkilerinin yok edildiği varsayılır. Bu durumda, ergodik teoremi kullanılarak $(\mathbf{x}^{(k)}, \mathbf{x}^{(k+1)}, \mathbf{x}^{(k+2)}, \dots, \mathbf{x}^{(N)})$ zinciri üzerinden ilgili tahminler yapılır.

Metropolis-Hastings algoritmasında kullanılan öneri dağılımının uygunluğu iki şekilde değerlendirilebilir. İlk olarak, her iterasyonda elde edilen kabul olasılığı değerlerine bakılabilir. Kabul olasılığı tek değişkenli dağılımda, $[0.45-0.50]$; çok değişkenli dağılımda

ise, $[0.25 - 0.30]$ olduğunda öneri dağılımının uygun olduğu söylenebilir [60]. İterasyonlarda çok küçük kabul olasılıklarının elde edilmesi, zayıf bir zincir oluştuğunu gösterir. İkinci olarak, iz grafiğine bakılabilir. Böylece, zincirin hedef dağılımın tanım kümesindeki hareketliliği incelenerek, zayıf veya karma bir zincir olup olmadığına karar verilebilir.

Tez çalışmasında, Metropolis-Hastings algoritmasının temel işleyişinin görülmesi amacıyla, tek değişkenli bir dağılım örnek alınmış ve bunun için, ters- $\chi^2(\nu, s^2)$ dağılımı kullanılmıştır. Burada, ν , serbestlik derecesini ve s^2 , ölçek parametresini göstermektedir. Dağılımın olasılık yoğunluk fonksiyonu aşağıda vermiştir [36]:

$$f(x) = \frac{\left(\frac{\nu}{2}\right)^{\left(\frac{\nu}{2}\right)}}{\Gamma\left(\frac{\nu}{2}\right)} s^\nu x^{-\left(\frac{\nu}{2}+1\right)} e^{-\left(\frac{\nu s^2}{2x}\right)}, \quad x \geq 0$$

Burada amaç, ters- $\chi^2(7, 4)$ dağılımından Metropolis-Hasting algoritması ile örneklem çekmek olsun. Metropolis-Hastings algoritmasının her bir iterasyonunda kabul olasılığı hesaplanırken, $f(x)$ dağılımında x değerinden bağımsız olan katsayılar sadeleşecektir. Bu nedenle, hesaplamalarda kolaylık sağlaması açısından bu katsayılar ihmal edilebilir. Bu durumda, $C = \frac{\left(\frac{\nu}{2}\right)^{\left(\frac{\nu}{2}\right)}}{\Gamma\left(\frac{\nu}{2}\right)} s^\nu$ olmak üzere, ters- $\chi^2(7, 4)$ hedef dağılımı aşağıdaki şekilde yazılabilir.

$$f(x) = C x^{-\left(\frac{7}{2}+1\right)} e^{-\left(\frac{7*4}{2x}\right)}, \quad x \geq 0$$

Algoritmanın uygulanmasında kullanılacak öneri dağılımı, $q(\cdot | x^{(t)}) = q_1(\cdot)$ olarak seçilmiştir. $q_1(x)$ dağılımı, zincirin bulunduğu $x^{(t)}$ noktasından bağımsız olarak x^* aday değerini üretmektedir. Bu forma uygun olan bir öneri dağılımı kullanılarak gerçekleştirilen Metropolis-Hastings algoritması sonucunda bağımsız bir örneklem elde edilir. Diğer bir ifadeyle, bağımsız bir zincir oluşur [62]. Bu durumda, Metropolis-Hastings algoritması çok iyi ya da çok kötü işleyebilir. Algoritmanın çok iyi işlemesi için, $q_1(x)$ öneri dağılımı $f(x)$ hedef dağılımının iyi bir yaklaşımı olmalıdır. Burada en güvenli yol, kuyukları hedef dağılımının kuyuklarından daha kalın olan bir $q_1(x)$ öneri dağılımı seçmektir. Çünkü, hedef dağılıma göre daha ince kuyuklu bir öneri dağılımı seçildiğinde, aday noktalarının çoğu kuyuk bölgesinden gelmeyecektir ve bu durumda düşük kabul olasılıkları elde edilerek zayıf bir zincir oluşacaktır. Hedef dağılıma göre daha kalın kuyuklu bir öneri dağılımı seçildiğinde ise, red edilen toplam aday nokta sayısı azalacaktır ve karma bir zincir oluşacaktır.

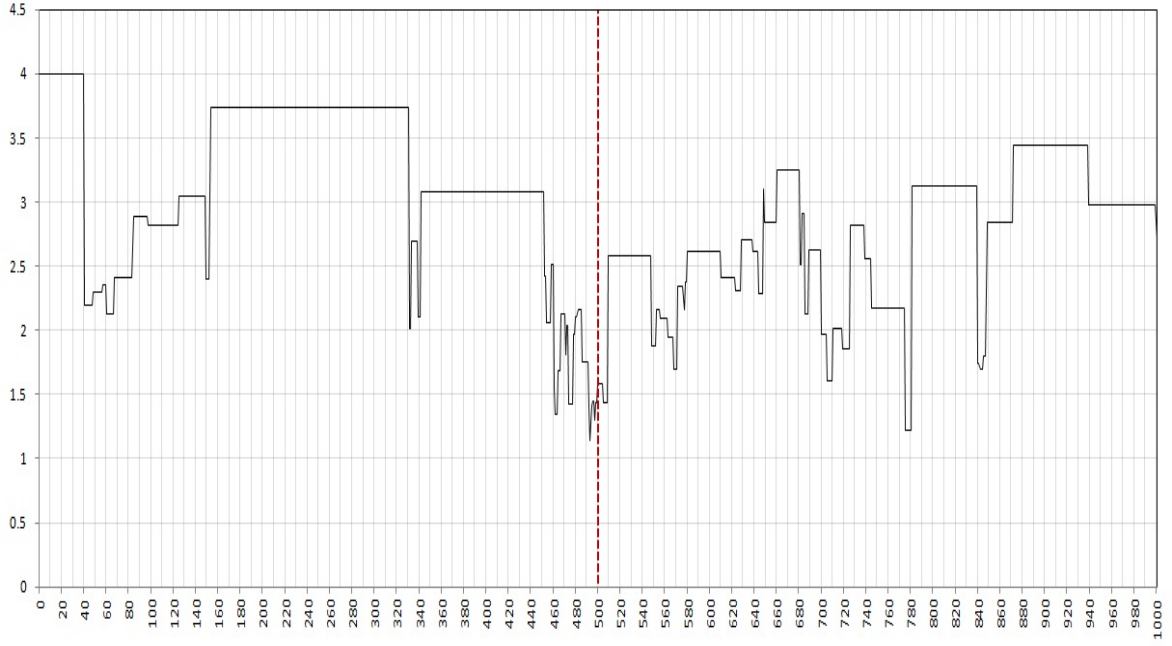
Bu algoritmanın işleyişinin açıklanması için, $q_1(x)$ öneri dağılımı, $Gamma(5, 4)$ olarak seçilmiştir. Burada, 5, şekil parametresini (n) ve 4, oran parametresini (λ) göstermektedir. Kabul olasılığının hesaplanmasında, $Gamma(n, \lambda)$ dağılımında x' ten bağımsız katsayılar ihmal edilmiş ve olasılık yoğunluk fonksiyonu aşağıdaki gibi kullanılmıştır.

$$q_1(x) \propto x^{(n-1)} e^{-\lambda x}, \quad x > 0$$

Burada, $x^{(0)}$ başlangıç değeri, 4 alınmıştır. Metropolis-Hastings algoritmasının ilk iterasyonu için $Gamma(5, 4)$ dağılımdan türetilen aday nokta, $x^* = 0.69380$ ' dir. Buna göre, bu aday noktanın kabul edilme olasılığı aşağıdaki gibi elde edilebilir:

$$\begin{aligned} \alpha(x^{(0)}, x^*) &= \min \left(\frac{f(x^*) q_1(x^{(0)})}{f(x^{(0)}) q_1(x^*)}, 1 \right) \\ &= \min \left(\left(\frac{0.69380^{-\left(\frac{7}{2}+1\right)} e^{-\left(\frac{7*4}{2*0.69380}\right)}}{4^{-\left(\frac{7}{2}+1\right)} e^{-\left(\frac{7*4}{2*4}\right)}} \right) \left(\frac{4^{(5-1)} e^{-4*4}}{0.69380^{(5-1)} e^{-4*0.69380}} \right), 1 \right) \\ &= \min (0.0000003021, 1) \\ &= 0.0000003021 \end{aligned} \tag{4.14}$$

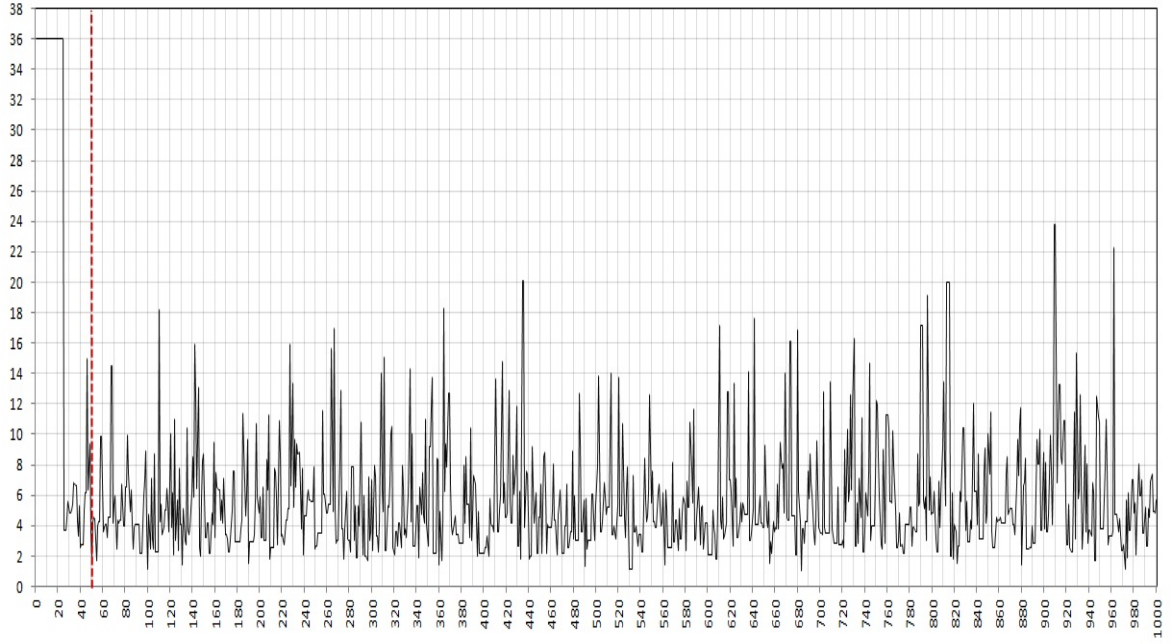
Eşitlik 4.14' te görüldüğü gibi, 0.69380 aday noktasının kabul olasılığı çok küçük bulunmuştur. Kabul olasılığının çok küçük elde edilmesi, zincirin aynı değerde kalacağını bir göstergesidir. $x^{(1)}$ zincir değerinin belirlenmesi amacıyla, $\mathcal{U}(0, 1)$ tekbiçimli dağılımdan rasgele bir $u = 0.96687$ değeri seçilmiştir. $u > \alpha(x^{(0)}, x^*)$ olmasından dolayı, $x^{(1)} = x^{(0)}$ olur. Burada, zincir aynı değerde kalmıştır. İkinci iterasyonda, $Gamma(5, 4)$ dağılımdan elde edilen aday nokta, $x^* = 0.81314$ ' tür. Bu aday noktanın kabul edilme olasılığı ise, $\alpha(x^{(1)}, x^*) = 0.00000244$ olarak hesaplanmıştır. Kabul olasılığı yine çok küçük bulunmuştur. $x^{(2)}$ zincir değerinin belirlenmesi amacıyla, $\mathcal{U}(0, 1)$ tekbiçimli dağılımdan rasgele bir $u = 0.59099$ değeri seçilmiştir. $u > \alpha(x^{(1)}, x^*)$ olmasından dolayı, $x^{(2)} = x^{(1)}$ olur. Zincir, yine aynı değerde kalmıştır. Toplam 1000 iterasyon gerçekleştirilmiştir. Uygulamada elde edilen zincir aşağıda vermiştir:



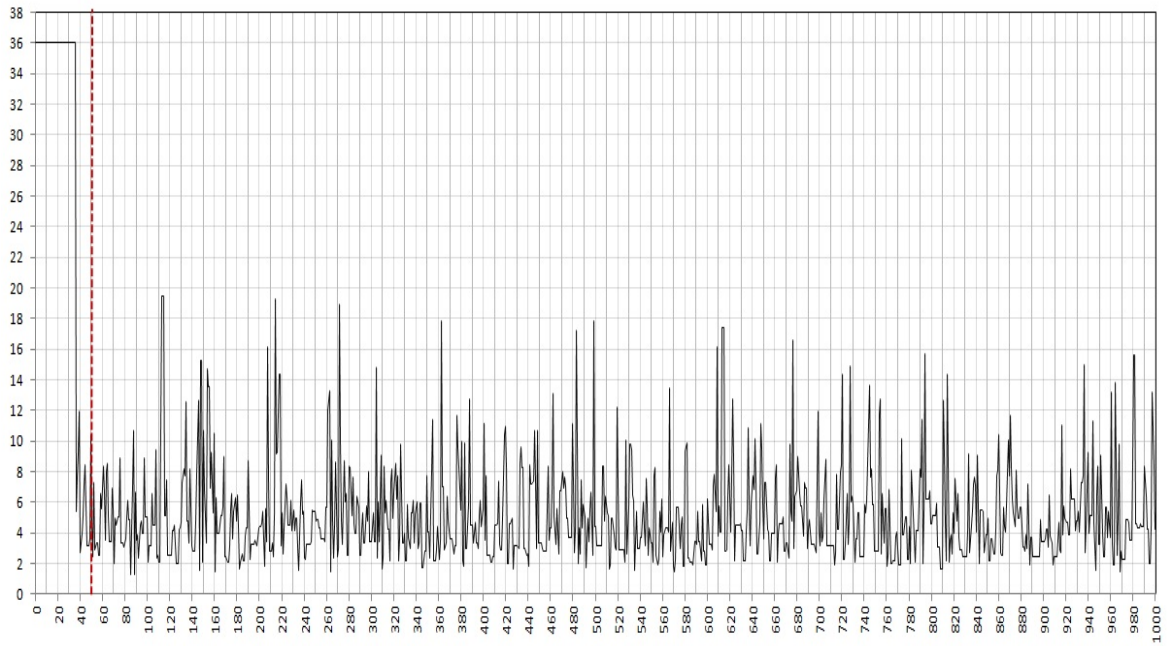
Şekil 4.1. $Gamma(5,4)$ Öneri Dağılımı Kullanıldığında Oluşan Zayıf Zincir

Başlangıç değerlerinin etkileri olduğu düşünülerek, ilk 500 iterasyon Şekil 4.1.' de kırmızı çizgiyle bölünmüştür. Şekil 4.1.' de düz-yatay çizgiler oldukça sık gözlenmiştir. Bunun nedeni, birçok iterasyonda zincirin aynı değerde kalmasıdır. Hedef dağılımın tanım kümesi zincir tarafından yeteri ölçüde taranamamıştır ve zayıf bir zincir oluşmuştur. Sonuç olarak, $Gamma(5,4)$ dağılımı, $ters-\chi^2(7,4)$ dağılımı için uygun bir öneri dağılımı değildir. Bu iki dağılımın olasılık yoğunluk fonksiyonlarının grafiklerine bakıldığında, $Gamma(5,4)$ 'ün yayılımının $ters-\chi^2(7,4)$ 'ün yayılımını kapsamadığı görülebilir.

$Ters-\chi^2(7,4)$ dağılımının olasılık yoğunluk grafiğine bakıldığında, $Gamma(3,0.5)$ ' in veya χ_6^2 ' nin daha uygun öneri dağılımları olduğu söylenebilir. Bu dağılımların kuyrukları, hedef dağılıminkilerden daha kalındırlar. Bu nedenle, Metropolis-Hastings algoritmasının uygulanmasıyla karma zincirlerin oluşması beklenir. Çalışmada, $Gamma(3,0.5)$ ve χ_6^2 dağılımlarının her biri için toplam 1000 iterasyon gerçekleştirilmiştir. Burada, $x^{(0)}$ başlangıç değeri, 36 olarak alınmıştır. Bu değer, $ters-\chi^2(7,4)$ dağılımının kuyruk bölgesinde yer alan bir uç değerdir. Elde edilen zincirler aşağıda verilmiştir:



Şekil 4.2. $Gamma(3, 0.5)$ Öneri Dağılımı Kullanıldığında Oluşan Karma Zincir



Şekil 4.3. χ_6^2 Öneri Dağılımı Kullanıldığında Oluşan Karma Zincir

Şekil 4.2. ve Şekil 4.3. incelendiğinde, başlangıç değer etkilerinin ilk 50 iterasyonda yok olduğu görülmektedir. İlk 50 iterasyondan sonrasına bakıldığında, oluşturulan zincirler hedef dağılımın tanım kümesini kapsayacak şekilde hareket etmiştir. Zincir, belli değerlerde kalmamıştır ve tanım kümesinin sadece belli bölgelerinde hareketlilik göstermemiştir. Bu nedenle, Metropolis-Hastings algoritması sonucu her iki öneri dağılımında karma bir zincir oluşmuştur.

4.2.2. Gibbs Örneklemesi

Gibbs örneklemesi, uzaktan algılama problemlerinde görüntü işleme modelleri üzerine çalışan Geman ve Geman [64] tarafından ilk kez tanıtılmıştır. Bu çalışmanın temelini, Metropolis vd. [58], Hastings [59] ve Peskun [65] oluşturmuştur. Geman ve Geman [64] çalışması, mühendislik alanında etkili bir makale olmasına rağmen, bir istatistik dergisinde yayınlanmamış olmasından dolayı, istatistiksel problemlerin çözümünde kullanımı geç olmuştur. Geman ve Geman'ın örnekleme yönteminden etkilenen Gelfand ve Smith [66], Bayesci çıkarsamada Gibbs örneklemesinin çeşitli istatistik problemlerine uygulanabilirliğini 1990 yılındaki çalışmalarında göstermişlerdir. Gelfand ve Smith [66], Gibbs örneklemesini istatistik camiasına tanıtan ilk yazarlardır. Bu çalışma ile birlikte, Bayesci yöntemlerde, istatistiksel hesaplamalarda ve stokastik süreçlerde Gibbs örnekleme ve Metropolis-Hastings algoritması gibi benzetim yöntemlerinin kullanımına olan ilgi artmıştır [33], [36], [39], [61], [67].

Gibbs örneklemesi, yaygın olarak kullanılan bir Markov zinciri Monte Carlo yöntemidir. Metropolis-Hastings algoritmasının özel bir durumudur. Gibbs örneklemesinde, üretilen aday nokta için hesaplanan kabul olasılığı her zaman birdir. Dolayısıyla, üretilen her değer kabul edilir [33], [68].

Bu yöntem, çok boyutlu ve karmaşık yapıda olan bileşik dağılımından, ilgili dağılıma gerek duyulmadan, örneklem çekebilme olanağı sağlar [67]. Hedef dağılımı yerine, bileşenlere ilişkin tam koşullu dağılımlar kullanılır. Böylece, tek değişkenli dağılımlar kullanılarak, durağan dağılımı hedef dağılıma yakınsayan ergodik bir Markov zinciri üretilir. Çok boyutlu karmaşık bir problem, basit ve daha az boyutlu problemlere dönüştürülür [35]. Gibbs örneklemesinde, tam koşullu dağılımların tamamen bilindiği ve bu dağılımlardan örneklem çekilebildiği varsayılır [35], [39], [61]. Tam koşullu dağılımın formu, bilinen bir dağılım (Normal, Gamma, vd.) ise, doğrudan örnekler üretilir. Diğer durumlarda, red örnekleme (rejection sampling), uyarlamalı red örnekleme (adaptive rejection sampling) ve Metropolis-Hastings algoritması gibi benzetim yöntemleri kullanılarak dolaylı bir şekilde örnekler çekilebilir [35].

Gibbs örneklemesinin işleyişini anlamak için, $\mathbf{X} = (X_1, X_2, \dots, X_p)$ ($p > 1$) raslantı değişkeni vektörü tanımlansın. Çok boyutlu ve karmaşık yapıda olan $f(\mathbf{x})$ bileşik dağılımından örneklem çekilmek istensin. Bunun için, $f(x_i | \mathbf{x}_{-i})$ ($i = 1, 2, \dots, p$) tam

koşullu dağılımların bilinmesi ve bu dağılımlardan doğrudan ya da dolaylı olarak örnekler üretilebilmesi gerekmektedir. Burada, $\mathbf{X}_{-i} = (X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_p)$ ' dir. İterasyon sayısı, $g = 0, 1, 2, \dots, N$ olarak gösterilsin. Gibbs algoritması aşağıda verilen adımlarla tanımlanır [34], [35], [39], [51], [61]:

Adım 1: $\mathbf{x}^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_p^{(0)})$ başlangıç değerinin belirlenmesi,

Adım 2: İterasyon sayacı $g = 1$ olarak atanması,

Adım 3: Aşağıda verilen tam koşullu dağılımlardan $\mathbf{X}^{(g)} = (X_1^{(g)}, X_2^{(g)}, \dots, X_p^{(g)})$ vektörünün oluşturulması,

$$X_1^{(g)} \sim f(X_1 | X_2^{(g-1)}, X_3^{(g-1)}, X_4^{(g-1)}, \dots, X_p^{(g-1)})$$

$$X_2^{(g)} \sim f(X_2 | X_1^{(g)}, X_3^{(g-1)}, X_4^{(g-1)}, \dots, X_p^{(g-1)})$$

$$X_3^{(g)} \sim f(X_3 | X_1^{(g)}, X_2^{(g)}, X_4^{(g-1)}, \dots, X_p^{(g-1)})$$

⋮

$$X_{p-1}^{(g)} \sim f(X_{p-1} | X_1^{(g)}, X_2^{(g)}, X_3^{(g)}, X_4^{(g)}, \dots, X_{p-2}^{(g)}, X_p^{(g-1)})$$

$$X_p^{(g)} \sim f(X_p | X_1^{(g)}, X_2^{(g)}, X_3^{(g)}, X_4^{(g)}, \dots, X_{p-1}^{(g)})$$

Adım 4: $g < N$ olduğu durumda, g değeri bir arttırılarak 3. adıma dönülmesidir.

Yukarıdaki algoritmadan üretilen $(\mathbf{x}^{(0)}, \mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(N)})$ dizisi, geçiş olasılıkları aşağıda verilen bir Markov zinciri oluşturur.

$$P(\mathbf{X}^{(g+1)} | \mathbf{X}^{(g)}) = \prod_{i=1}^p f(X_i^{(g+1)} | X_j^{(g+1)}, j < i \text{ ve } X_j^{(g)}, j > i) \quad , \quad g = 0, 1, 2, \dots, N \quad (4.15)$$

Eşitlik 4.15' te verilen geçiş olasılıkları dağılımı ile üretilen Markov zinciri, başlangıç değerlerine bağlıdır. Gibbs örnekleme yeterli uzunlukta çalıştırıldığında ($N \rightarrow \infty$), başlangıç değerlerinin etkileri tamamiyle yok edilir. k -ıncı iterasyonda başlangıç değerlerinin etkilerinin yok olduğu ve zincirin durağan dağılıma yakınsadığı varsayalım. Bu durumda, $(\mathbf{x}^{(k)}, \mathbf{x}^{(k+1)}, \mathbf{x}^{(k+2)}, \dots, \mathbf{x}^{(N)})$ örnekleme değerleri, $f(\mathbf{x})$ hedef dağılımından çekilen bağımlı örnek değerler olur [35], [39], [67].

Tez çalışmasında bu algoritmanın işleyişinin açık bir şekilde ortaya konulması için örnek olarak iki değişkenli bir normal dağılım, $(X_1, X_2) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ ele alınmıştır. Burada, $\boldsymbol{\mu}$,

beklenen değer vektörünü ve Σ , varyans-kovaryans matrisini göstermektedir:

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

Buna göre, $-\infty \leq x_1, x_2 < +\infty$ olmak üzere, dağılımın bileşik olasılık yoğunluk fonksiyonu aşağıda vermiştir [69]:

$$f(x_1, x_2) = \frac{e^{\left[-\frac{1}{2(1-\rho^2)} \left\{ \frac{(x_1-\mu_1)^2}{\sigma_1^2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2} - \frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} \right\} \right]}}{2\pi\sigma_1\sigma_2\sqrt{(1-\rho^2)}} \quad (4.16)$$

Gibbs örneklemesinin uygulanabilmesi için, X_1 ve X_2 raslantı değişkenlerinin tam koşullu sonsal dağılımlarına ihtiyaç duyulur. Eşitlik 4.16' da verilen bileşik olasılık yoğunluk fonksiyonundan yararlanılarak, tam koşullu sonsal dağılımlar kapalı formda aşağıdaki gibi elde edilir [69]:

$$(X_1 | X_2) \sim \mathcal{N} \left(\mu_1 + \rho \frac{\sigma_1}{\sigma_2} (x_2 - \mu_2), (1 - \rho^2) \sigma_1^2 \right) \quad (4.17)$$

$$(X_2 | X_1) \sim \mathcal{N} \left(\mu_2 + \rho \frac{\sigma_2}{\sigma_1} (x_1 - \mu_1), (1 - \rho^2) \sigma_2^2 \right) \quad (4.18)$$

Çalışmada ele alınan örnek dağılım aşağıdaki şekilde tanımlanmıştır:

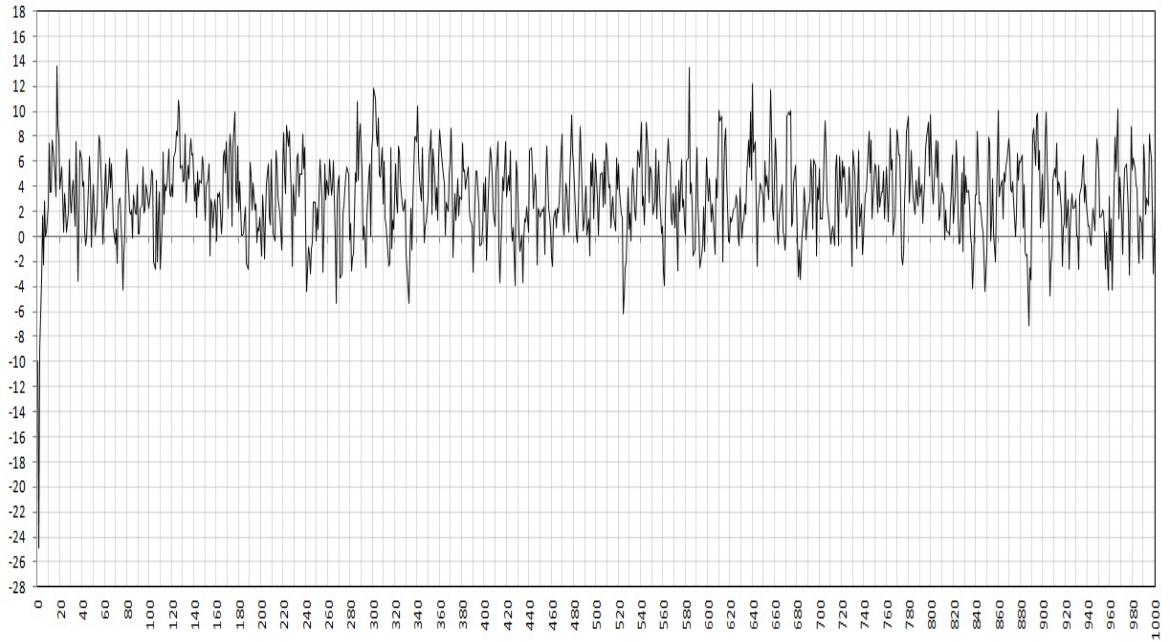
$$(X_1, X_2) \sim \mathcal{N} \left(\begin{bmatrix} 3 \\ 5 \end{bmatrix}, \begin{bmatrix} 10 & 3 \\ 3 & 2 \end{bmatrix} \right) \quad (4.19)$$

Gibbs örneklemesinin uygulanmasında, başlangıç değerleri $X_1^{(0)} = -10$ ve $X_2^{(0)} = -15$ alınmıştır. $(-10, -15)$ noktası, hedef dağılımın kuyruk bölgesinde yer alan bir uç değerdir. Verilen örnek için Gibbs örneklemesinin işleyişi, aşağıdaki adımlarla tanımlanmıştır:

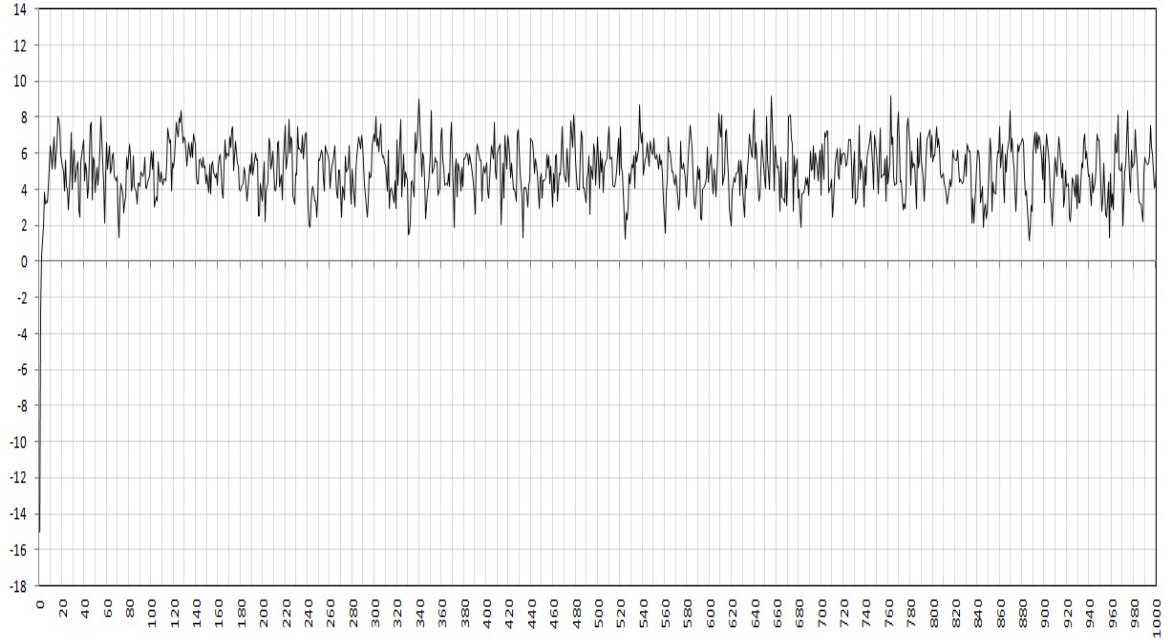
- 1- Eşitlik 4.17'e göre, $f(x_1^{(1)} | x_2^{(0)} = -15)$ ' ten $x_1^{(1)} = -24.86088$ değeri üretilmiş,
- 2- Eşitlik 4.18'e göre, $f(x_2^{(1)} | x_1^{(1)} = -24.86088)$ ' den $x_2^{(1)} = -2.49843$ değeri üretilmiş,
- 3- Böylece, ilk iterasyon için $(x_1^{(1)}, x_2^{(1)}) = (-24.86088, -2.49843)$ örnek değeri üretilmiş,
- 4- Eşitlik 4.17'e göre, $f(x_1^{(2)} | x_2^{(1)} = -2.49843)$ ' ten $x_1^{(2)} = -9.00922$ değeri üretilmiş,

- 5- Eşitlik 4.18'e göre, $f(x_2^{(2)} | x_1^{(2)} = -9.00922)$ ' dan $x_2^{(2)} = 0.32668$ değeri üretilmiş,
- 6- Böylece, ikinci iterasyon için $(x_1^{(2)}, x_2^{(2)}) = (-9.00922, 0.32668)$ örnek değeri üretilmiş ve
- 7- Algoritmada bu şekilde devam edilerek, $((x_1^{(1)}, x_2^{(1)}), \dots, (x_N^{(1)}, x_N^{(1)}))$ örneklemini oluşturulmuştur

Uygulama için $N = 1000$ iterasyon gerçekleştirilmiştir. X_1 ve X_2 için oluşan zincirler aşağıda sunulmuştur:

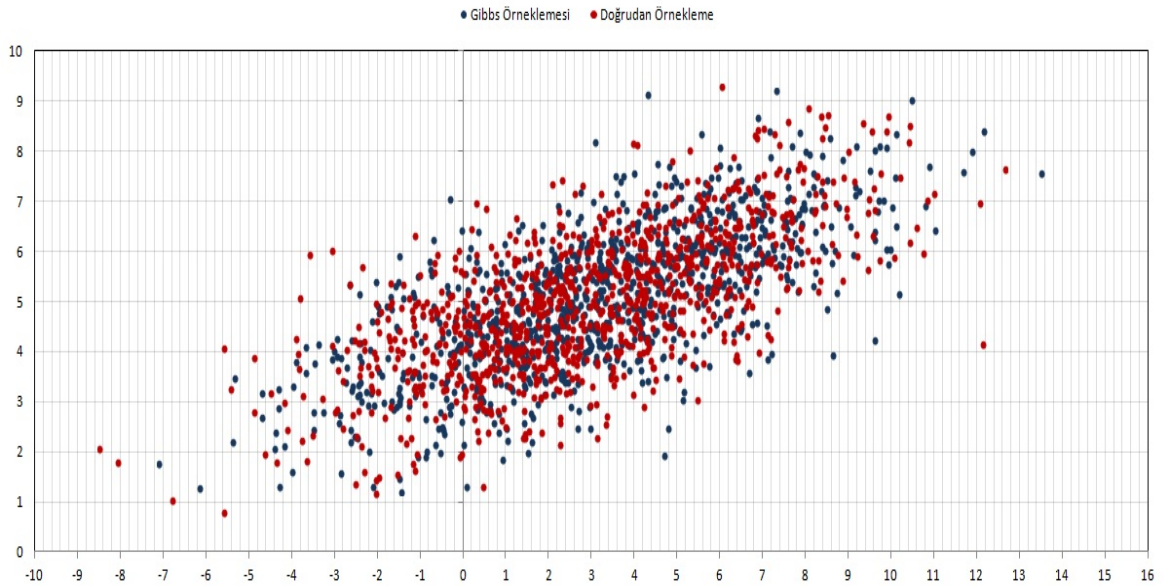


Şekil 4.4. X_1 Raslantı Değişkeni için Oluşturulan Karma Zincir



Şekil 4.5. X_2 Raslantı Değişkeni için Oluşturulan Karma Zincir

Şekil 4.4. ve Şekil 4.5. incelendiğinde, $(-10, -15)$ başlangıç değerinin etkilerinin kısa sürede yok olduğu görülmektedir. Her iki raslantı değişkeni için de karma zincir olduğu görülmektedir. Gibbs örneklemesinin performansının daha somut bir şekilde görülebilmesi amacıyla, başlangıç değer etkilerinin giderilmesi için ilk 20 iterasyondan elde edilen örnek değerler zincirden atılmıştır ve Eşitlik 4.19' da verilen iki değişkenli normal dağılımdan doğrudan 980 birimlik örneklem çekilmiştir. Doğrudan örnekleme ve Gibbs örneklemesinden elde edilen toplam 1960 $(980 \cdot 2)$ örnek değerlerin saçılım grafikleri aşağıda verilmiştir.



Şekil 4.6. Doğrudan Örnekleme ile Gibbs Örneklemesinin Karşılaştırılması

Şekil 4.6.' da görüldüğü gibi, doğrudan örnekleme ve Gibbs örnekleme ile üretilen değerler birbirleriyle uyumludur. Gibbs örnekleme, başarılı bir şekilde işlemektedir.

Bu bölümde tanıtılan stokastik benzetim yöntemleri, son yıllarda birçok karmaşık istatistiksel modellerin Bayesci çözümlemesinde oldukça etkili sonuçlar vermiştir. Burada yer alan bilgiler, Bayesci SMM için kullanılan bilgilerdir. Tez çalışmasında, teorisi ve uygulaması hayli zor olan SMM için bu yöntemler tahmin sürecinde kullanılacaktır.

Tez çalışmasının izleyen bölümünde, öncelikle standart saklı Markov modelinin yapısı ve bileşenleri tanıtılacak; modelin üç temel problemi çözümleriyle açıklanacaktır. Burada tanıtılan modelin parametrelerinin tahmininde öncelikle klasik bir yöntem olan olabilirlik fonksiyonunun en büyüklenmesi işleminin nasıl uygulandığı verilecektir.



5. SAKLI MARKOV MODELİ

Saklı Markov modeli, stokastik bir model olup, Markov modellerinin bir uzantısıdır. Markov modellerinde ele alınan sistemin durumları ve bu durumlara ilişkin olasılık fonksiyonları bilinir. Sistemin herhangi bir t zamanındaki çıktısı gözlem olarak adlandırılır. Bu nedenle, Markov modellerinde sistemin sahip olduğu süreç gözlemler aracılığıyla rahatlıkla izlenebilir. Ancak, gerçek uygulamalarda ilgilenilen sistemin izlenebilir olması çoğu zaman mümkün değildir. Bu durum, Markov modellerinin uygulama alanlarını daraltan önemli bir sorundur. Saklı Markov modeli, sistemin durumlarının gözlemlenemediği stokastik modelleme problemleri için uygun bir alternatif sunar. Burada, “saklı” ifadesi sistemin izlediği sürecin çıktılarının gözlemlenememesinden dolayı kullanılmaktadır. Diğer bir ifadeyle, incelenen sistemin durumları saklı bir süreç tarafından açıklanmaktadır. “Markov” ifadesi ise, saklı sürecin Markov özelliği taşımasından kaynaklanmaktadır. Modelde gözlemler, sistemin izlediği sürecin etkisi altında olan başka bir sürecin çıktısı olarak tanımlanmaktadır. Burada, gözlemlenebilen bu süreç aracılığıyla, sistemin saklı durumuna ilişkin problemler incelenmekte ve tüm istatistiksel çıkarımlar gözlemlenebilen bu süreç üzerinden yapılmaktadır.

Saklı Markov modelinin teorisi, ilk olarak 1966 yılında Baum ve Petrie [1] tarafından ortaya konulmuş ve “*sonlu durumlu Markov zincirlerinin olasılıksal fonksiyonları*” olarak tanımlanmıştır. 1960’ların sonları ve 1970’lerin başı itibarıyla, saklı Markov modelleri Baum ve arkadaşları tarafından bir dizi klasik çalışma (Baum ve Egon [70], Baum ve Sell [71], Baum vd. [72] ve Baum [73]) ile geliştirilmiştir.

Saklı Markov modellerinin teorisine ilişkin ilk uygulamalar Baker [2], Jelinek, Bahl ve Mercer [3], Bahl ve Jelinek [4] ve Jelinek [5] tarafından 1970’li yılların ortalarında “*konuşma tanıma*” üzerine yapılmıştır. Modelin matematiksel yapısının zengin olması ve uygun bir şekilde kullanıldığında iyi sonuçlar vermesine rağmen, teorisinin farklı alanlara uygulanması zaman almıştır. Bu durumun en önemli nedenleri, modelin teorisi üzerine yapılan çalışmaların yoğunluklu olarak matematiksel dergilerde yayınlanması ve modeli kendi çalışma alanlarına entegre etmek isteyenlerin modelin teorisini anlamakta ve uygulamakta materyal sıkıntısı yaşaması olarak söylenebilir. Bilgisayar teknolojisinin gelişmesi ile birlikte saklı Markov modelinin uygulama alanlarında gelişme kaydedilmiştir. Günümüzde, saklı Markov modeli bir çok farklı alanda uygulanmaktadır. Modelin uygulanma alanları her geçen gün daha da artmaktadır.

5.1. Modelin Tanımı ve Bileşenleri

Markov modellerinde, belirlenen bir zaman aralığında sistemin içinde bulunduğu durumlar gözlemleri oluşturmaktadır. Dolayısıyla, Markov modellerinde sistem doğrudan gözlemlenebilmektedir. Sistemin durumları gözlemlenemediğinde, Markov modelleri kullanılarak sisteme ilişkin istatistiksel çıkarımlar yapılamaz. Bu durumda, Markov modellerinden daha esnek bir yapıya sahip olan saklı Markov modelleri kullanılır.

Standart Saklı Markov modeli, *kesikli zamanlı, homojen, sonlu-durumlu Markov zinciri* üzerine kurulu stokastik bir modeldir. SMM' de sistemin herhangi bir anda içinde bulunduğu durum doğrudan gözlemlenemez. Sistemin durumları, *saklı bir süreç* tarafından belirlenmektedir. Sistem belli bir durumda iken, bu durumun tetiklediği gözlem ortaya çıkar. Dolayısıyla, saklı sürecin etkisi altında olan başka bir süreç tarafından sistem gözlemlenebilmektedir.

Kesikli dizin kümesi $T = \{1, 2, 3, \dots, T\}$ olmak üzere, **saklı Markov modeli**, kesikli zamanlı ikili stokastik süreç $(\mathbf{X}, \mathbf{Y}) = (X_t, Y_t)_{t \in T}$ olarak tanımlanır [74], [75], [76]. Bu süreçler,

- $\mathbf{X} = (X_t)_{t \in T}$ süreci, kesikli zamanlı, homojen, sonlu-durumlu Markov zinciridir ve saklı süreç olarak tanımlanır. X_t , t anında sistemin içinde bulunduğu durumdur.
- $\mathbf{Y} = (Y_t)_{t \in T}$ süreci, \mathbf{X} saklı sürecinin etkisi altında ortaya çıkan gözlemlenebilen süreç olarak tanımlanmaktadır. Y_t , t anında meydana gelen gözlemdir.

SMM, aşağıdaki bileşenlerden oluşur [6]:

- **m Durum Sayısı**

\mathbf{X} saklı sürecinin Markov zinciri olması nedeniyle, sistemin durum uzayı kesiklidir. Standart SMM' de durum uzayı sonludur. m , durum uzayındaki toplam durum sayısıdır. m tane saklı durum içeren bir sistem için, m -durumlu SMM incelenir. Burada, $\mathcal{S} = \{S_1, S_2, \dots, S_m\}$, kesikli sonlu durum kümesi olsun. Standart SMM' de durumlar, ergodik yapıdadır ve birbirinden bağımsızdırlar. Durumların sayısı, SMM'nin boyutunu gösterir. Saklı Markov modeli incelenirken, "Gözlemlerin en iyi şekilde açıklanmasında, gerçekte kaç tane duruma ihtiyaç vardır?" sorusunun cevabı çok önemlidir. Çünkü, sistemin durum sayısına bağlı olarak, yapılacak istatistiksel çıkarımlar değişir.

- **Durum Geçiş Olasılıkları Dağılımı**

S durum kümesi üzerinde tanımlanan $P = [p_{ij}]_{m \times m}$ matrisi ile gösterilmektedir. p_{ij} elemanı, sistemin S_i durumundan S_j durumuna geçmesi olasılığıdır. $\forall S_i, S_j \in S$ için geçiş olasılıkları,

$$p_{ij} = P(X_{t+1} = S_j | X_t = S_i)$$

şeklinde ifade edilir. Standart SMM' de X saklı süreci homojen bir Markov zinciri olması nedeniyle, durum geçiş olasılıkları zaman içerisinde değişmez ve bu olasılıklar gözlemlerden bağımsızdırlar. P matrisi aşağıdaki özellikleri sağlar:

i. X saklı sürecinin ergodik bir Markov zinciri olması nedeniyle, durumlar kendi içlerinde herhangi bir durumdan diğer tüm durumlara ulaşılabilir şekilde bağlantılıdır. Diğer bir ifadeyle, sistemin her bir durumundan tüm durumlara geçiş mümkündür. Bu nedenle, $\forall 1 \leq i, j \leq m$ için, $p_{ij} > 0$ olur. Ergodik olmayan X saklı süreci için, bir ya da daha fazla $p_{ij} = 0$ olur.

ii. $\forall 1 \leq i \leq m$ için, $\sum_{j=1}^m p_{ij} = 1$ ' dir.

P matrisinin özellikleri dikkate alındığında, sistemin durumlarının *katlı-terimli dağılıma* sahip olduğu görülür.

- **Gözlemler Kümesi**

Y süreci, kesikli ya da sürekli değerler alabilir. Y ' nin X ' e koşullu dağılımı, tek bir (kesikli ya da sürekli) parametrik aileye aittir. Burada, *Poisson, Binom, Normal, Gamma, Beta vb.* dağılım aileleri düşünülebilir. İlgilenilen dağılım ailesinin parametreleri, X saklı süreci tarafından belirlenir. Bu durumda, Y ' nin marjinal dağılımı, belirlenen parametrik ailenin ilişkili karma dağılımı olur.

- **Gözlem Olasılıkları**

Sistemin bulunduğu duruma bağlı olarak gözlemler oluşur. Sistemin durumu bilindiğinde, ortaya çıkan gözlem diğer gözlemlerden bağımsızdır. Gözlem olasılıkları, $B = [b_i(y)]_{m \times M}$ matrisi ile gösterilir. Y sürecinin tanım kümesine göre, M değeri tanımlanır. $b_i(y)$ elemanı, sistem S_i durumunda iken, y gözleminin ortaya çıkması olasılığıdır. Sürekli bir dağılım ailesi ele alındığında, $b_i(y)$ bir olasılık yoğunluk değeridir. Herhangi bir $t \in T$ zamanında $\forall S_i \in S$ için gözlem olasılıkları

$$b_i(y) = P(Y_t = y | X_t = S_i)$$

şeklinde ifade edilir. B matrisi aşağıdaki özellikleri sağlar.

i. $\forall S_i \in \mathcal{S}$ ve $\forall y \in \mathcal{R}_Y$ için, $b_i(y) \geq 0$ ' dir.

ii. Kesikli gözlemler için, $\sum_{\mathcal{R}_Y} b_i(y) = 1$ ' dir. Sürekli gözlemler için, $\int_{\mathcal{R}_Y} b_i(y) dy = 1$ ' dir.

• Başlangıç Durum Dağılımı

$\pi = [\pi_i]_{1 \times m}$ ilk olasılık vektörü ile gösterilir. π_i elemanı, başlangıç anında sistemin S_i durumunda bulunması olasılığıdır. $\forall S_i \in \mathcal{S}$ için başlangıç olasılıkları

$$\pi_i = P(X_1 = S_i)$$

şeklinde ifade edilir. π vektörü, $\sum_{i=1}^m \pi_i = 1$ koşulunu sağlar. Standart SMM' de X saklı süreci ergodik bir Markov zinciridir. Bu nedenle, başlangıç durum dağılımı bu sürecin durağan dağılımı olarak alınabilir. X saklı süreci indirgenemez ve ergodik bir Markov zinciri olması nedeniyle, tek durağan dağılıma sahiptir. X ' in durağan dağılımı, $\delta = [\delta_i]_{1 \times m}$ vektörüyle gösterilsin. δ , aşağıda verilen iki şekilde elde edilebilir [76]:

i. $\delta = \delta P$ eşitliği sağlanacak biçimde durağan dağılıma ulaşılabilir.

ii. $\mathbf{1} = [1]_{1 \times m}$, birlerden oluşan bir satır vektörü ve $U = [1]_{m \times m}$, "1" değerlerinden oluşan bir matris olsun. I_m birim matrisi olmak üzere, $\delta (I_m - P + U) = \mathbf{1}$ eşitliğinin çözümü olarak durağan dağılıma ulaşılabilir.

Yukarıda verilen model bileşenleri dikkate alındığında, SMM' yi tanımlayan parametreler P , B ve π ' dir. Bunlar, $\Theta = \{P, B, \pi\}$ parametre kümesi ile gösterilir. SMM' de bilinmeyen değişkenler sadece model parametreleri değildir. X sürecinin doğrudan gözlemlenememesi nedeniyle, modelde $\mathbf{X}^{(T)} = (X_1, \dots, X_T)$ durum dizisi de bilinmemektedir. Θ model parametreleri ve $\mathbf{X}^{(T)}$ durum dizisi bilinen SMM' de, $\mathbf{Y}^{(T)} = (Y_1, Y_2, \dots, Y_T)$ gözlem dizisi aşağıdaki adımlar izlenerek oluşur [77], [6].

i. $t = 1$ olarak alınır.

ii. π başlangıç durum dağılımına göre, X Markov zincirinin başlangıç durumu $X_1 = S_i$ olarak seçilir.

iii. t anında sistem S_i durumundayken, bu durumda belirtilen parametrelere sahip ilgili dağılımdan Y_t gözlem değeri oluşturulur.

- iv. P geçiş olasılıkları dağılımına göre, S_i durumundan S_j durumuna geçilir ve $X_{t+1} = S_j$ olur.
- v. $t < T$ olduğunda, t değeri bir artırılarak iii-üncü adıma gidilir. Aksi halde, işlem sonlandırılır.

5.2. Model Varsayımları ve Süreçlerin Bazı Özellikleri

Standart SMM' nin teorisi, üç ana varsayım üzerine kurulmuştur. Bu varsayımlar aşağıda belirtilmiştir [43]:

1. Markov Varsayımı

X ' in birinci derece Markov zinciri olduğu varsayılır. Birinci derece Markov özelliği gereği, herhangi bir $0 < t_1 < t_2 < \dots < t_{n-1} < t_n < T$ kümesi için, $X_{t_1}, X_{t_2}, \dots, X_{t_{n-1}}$ verildiğinde X_{t_n} ' in koşullu olasılık dağılımı yalnızca $X_{t_{n-1}}$ değerine bağlıdır. X ' in Markov özelliği aşağıdaki eşitlikte gösterilmiştir:

$$P(X_{t_n} | X_{t_1}, X_{t_2}, \dots, X_{t_{n-1}}) = P(X_{t_n} | X_{t_{n-1}}) \quad (5.1)$$

Eşitlik 5.1' de görüldüğü gibi, sistemin bir sonraki adımda bulunacağı durum geçmişten bağımsız olup, yalnızca mevcut duruma bağlıdır.

2. Durağanlık Varsayımı

X saklı süreci, durağan geçiş olasılıklı Markov zinciridir. Bu durum, $\forall S_i, S_j \in \mathcal{S}$ ve $\forall t_1, t_2 \in \mathcal{T}$ için

$$p_{ij} = P(X_{t_1+1} = S_j | X_{t_1} = S_i) = P(X_{t_2+1} = S_j | X_{t_2} = S_i)$$

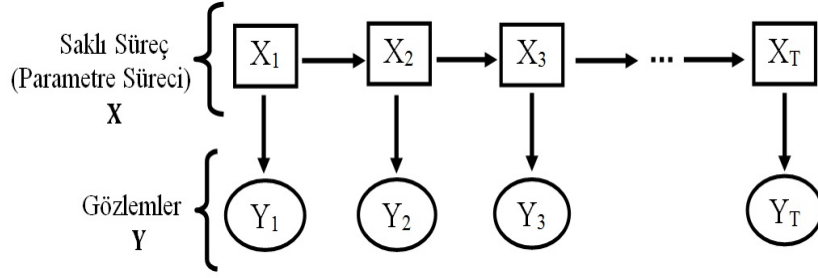
olarak tanımlanır. Bu nedenle, sistemin durumları arasındaki geçiş olasılıkları zamandan bağımsız olup, sadece zaman aralığına bağlıdır.

3. Gözlemlerin Koşullu Bağımsızlığı Varsayımı

$\forall t \in \mathcal{T}$ için, X_t değeri bilindiğinde, Y_t gözlem değeri diğer gözlemlerden ve saklı süreçten istatistiksel olarak bağımsızdır. Bu durumda, $\mathbf{X}^{(T)}$ durum dizisi verildiğinde, $\mathbf{Y}^{(T)}$ gözlem dizisinin olasılığı,

$$P(\mathbf{Y}^{(T)} | \mathbf{X}^{(T)}) = \prod_{t=1}^T P(Y_t | X_t)$$

olarak elde edilebilir. X ve Y süreçleri arasındaki bağımlılık yapısı, aşağıda grafiksel olarak belirtilmiştir [75], [78]:



Şekil 5.1. Saklı Markov Modelinde Süreçler Arasındaki Bağımlılık Yapısı

Burada, gözlemlerin koşullu olarak birbirinden bağımsız olduğu görülmektedir.

SMM' de X saklı sürecindeki serisel bağımlılığın açıklanması amacıyla kovaryans ve otokorelasyon fonksiyonlarına ihtiyaç duyulur. Durağan ve indirgenemez X saklı süreci, $\{1, 2, \dots, m\}$ üzerinde değerler alsın. $\mathbf{v} = [1 \ 2 \ \dots \ m]$ satır vektörü ve $\mathbf{V} = \mathbf{Köş}(1, 2, \dots, m)$ köşegen matrisi tanımlansın. k negatif olmayan herhangi bir tam sayı olmak üzere, kovaryans fonksiyonu

$$\begin{aligned} \text{Kov}(X_t, X_{t+k}) &= E(X_t X_{t+k}) - E(X_t) E(X_{t+k}) \\ &= \delta \mathbf{V} \mathbf{P}^k \mathbf{v}^\top - (\delta \mathbf{v}^\top)^2 \end{aligned} \quad (5.2)$$

olarak hesaplanır [76]. Eşitlik 5.2' de verilen kovaryans değeri, zamandan bağımsızdır ve sadece k gecikme sayısına bağlıdır. Buradan, X ' in durağan bir süreç olduğu görülmektedir. Burada, \mathbf{P} geçiş olasılıkları matrisinin köşegenleştirilebilir bir matris olduğu varsayılınsın. \mathbf{P} ' nin birden farklı özdeğerleri, $(\lambda_2, \dots, \lambda_m)$ olarak gösterilsin. $\mathbf{\Lambda} = \mathbf{Köş}(1, \lambda_2, \dots, \lambda_m)$ özdeğerler matrisi ve \mathbf{H} bu özdeğerlere ilişkin özvektörler matrisi olmak üzere, \mathbf{P} matrisine özdeğer ayrıştırması yapılsın. Bu durumda, $\mathbf{P} = \mathbf{H} \mathbf{\Lambda} \mathbf{H}^{-1}$ olarak elde edilir. $1 \times m$ boyutlu $\mathbf{a} = \delta \mathbf{V} \mathbf{H}$ satır vektörü ve $m \times 1$ boyutlu $\mathbf{b} = \mathbf{H}^{-1} \mathbf{v}^\top$ sütun vektörü tanımlansın. Özdeğer ayrıştırması sonucunda, Eşitlik 5.2' de verilen kovaryans fonksiyonu aşağıdaki gibi hesaplanabilir [76]:

$$\begin{aligned} \text{Kov}(X_t, X_{t+k}) &= \delta \mathbf{V} \mathbf{H} \mathbf{\Lambda}^k \mathbf{H}^{-1} \mathbf{v}^\top - (\delta \mathbf{v}^\top)^2 \\ &= \mathbf{a} \mathbf{\Lambda}^k \mathbf{b} - a_1 b_1 \\ &= \sum_{i=2}^m a_i b_i \lambda_i^k \end{aligned} \quad (5.3)$$

X sürecinin varyansı, $V(X_t) = \sum_{i=2}^m a_i b_i$ olarak bulunur. Bu durumda, otokorelasyon fonksiyonu aşağıdaki gibi elde edilir [76]:

$$\rho(k) = \frac{\text{Kov}(X_t, X_{t+k})}{\sqrt{V(X_t) V(X_{t+k})}} = \frac{\sum_{i=2}^m a_i b_i \lambda_i^k}{\sum_{i=2}^m a_i b_i} \quad (5.4)$$

Burada, X ' in durağan bir süreç olması nedeniyle, $V(X_t) = V(X_{t+k})$ olacaktır. Eşitlik 5.4' te verilen $\rho(k)$ değeri, P geçiş olasılıkları matrisinin birden farklı özdeğerlerinin k -nıncı kuvvetlerinin ağırlıklı ortalamasıdır.

SMM' de gözlemlere ait özelliklerin belirlenmesi için, Y sürecine ilişkin beklenen değer, varyans, kovaryans ve otokorelasyon fonksiyonlarının hesaplanması gerekmektedir. Herhangi bir $t \in T$ zamanında, sistemin $S_i \in S$ durumu için ilgili parametrik dağılımın beklenen değeri μ_i ve varyansı σ_i^2 olsun. Bu durumda, Y sürecinin beklenen değeri aşağıdaki gibi elde edilir [76]:

$$E(Y_t) = \sum_{i=1}^m E(Y_t | X_t = S_i) P(X_t = S_i) = \sum_{i=1}^m \mu_i \delta_i = \delta \mu^T$$

Burada, $\mu = [\mu_i]_{1 \times m}$ beklenen değer vektörüdür. Herhangi bir $g(Y_t)$ fonksiyonunun beklenen değeri,

$$E(g(Y_t)) = \sum_{i=1}^m E(g(Y_t) | X_t = S_i) P(X_t = S_i)$$

olarak bulunur [76]. Buna göre, Y sürecinin varyansı aşağıdaki gibi hesaplanır [76]:

$$V(Y_t) = E(Y_t^2) - [E(Y_t)]^2 = \sum_{i=1}^m \delta_i (\sigma_i^2 + \mu_i^2) - (\delta \mu^T)^2$$

Herhangi bir $g(Y_t, Y_{t+k})$ fonksiyonunun beklenen değeri ise, aşağıda verilmiştir [76]:

$$E(g(Y_t, Y_{t+k})) = \sum_{i=1}^m \sum_{j=1}^m E(g(Y_t, Y_{t+k}) | X_t = S_i, X_{t+k} = S_j) P(X_{t+k} = S_j | X_t = S_i) P(X_t = S_i)$$

X Markov zincirinin k -adım geçiş olasılıkları matrisi, $\mathbf{P}^{(k)} = [p_{ij}^{(k)}]_{m \times m}$ olarak tanımlansın. $(Y_t Y_{t+k})$ çarpımı için beklenen değer aşağıdaki gibi hesaplanır [76]:

$$\begin{aligned} E(Y_t Y_{t+k}) &= \sum_{i=1}^m \sum_{j=1}^m E(Y_t Y_{t+k} | X_t = S_i, X_{t+k} = S_j) P(X_{t+k} = S_j | X_t = S_i) P(X_t = S_i) \\ &= \sum_{i=1}^m \sum_{j=1}^m E(Y_t | X_t = S_i) E(Y_{t+k} | X_{t+k} = S_j) P(X_{t+k} = S_j | X_t = S_i) P(X_t = S_i) \\ &= \sum_{i=1}^m \sum_{j=1}^m \mu_i \mu_j p_{ij}^{(k)} \delta_i \\ &= \boldsymbol{\delta} \mathbf{K} \mathbf{P}^k \boldsymbol{\mu}^\top \end{aligned}$$

Burada, $\mathbf{K} = \mathbf{Köş}(\mu_1, \mu_2, \dots, \mu_m)$ olarak tanımlanmıştır. Y sürecinin kovaryans fonksiyonu,

$$\text{Kov}(Y_t Y_{t+k}) = E(Y_t Y_{t+k}) - E(Y_t) E(Y_{t+k}) = \boldsymbol{\delta} \mathbf{K} \mathbf{P}^k \boldsymbol{\mu}^\top - (\boldsymbol{\delta} \boldsymbol{\mu}^\top)^2$$

ve otokorelasyon fonksiyonu,

$$\rho^{(k)} = \frac{\text{Kov}(Y_t, Y_{t+k})}{\sqrt{V(Y_t) V(Y_{t+k})}} = \frac{\boldsymbol{\delta} \mathbf{K} \mathbf{P}^k \boldsymbol{\mu}^\top - (\boldsymbol{\delta} \boldsymbol{\mu}^\top)^2}{\sum_{i=1}^m \delta_i (\sigma_i^2 + \mu_i^2) - (\boldsymbol{\delta} \boldsymbol{\mu}^\top)^2}$$

olarak elde edilir [76].

5.3. Modelin Üç Temel Problemi ve Çözüm Yöntemleri

Standart SMM' nin gözlenmiş verilere uygulanması aslında üç temel problemin çözümüdür. Bunlar, değerlendirme problemi (evaluation problem), kodlama problemi (decoding problem) ve öğrenme problemi (learning problem)' dir. İzleyen alt bölümlerde, modelin üç temel problemi ve çözüm yöntemleri verilmiştir. Bunun için, Ibe [43]; Zucchini ve MacDonald [76]; Rabiner ve Juang [77]; Rabiner [6] ve Alpaydin [79] kaynaklarından yararlanılmıştır.

5.3.1. Değerlendirme Problemi ve Çözüm Yöntemi

Değerlendirme problemi, $P(\mathbf{Y}^{(T)} | \boldsymbol{\Theta})$ gözlem dizisi olasılığının hesaplanmasıdır. Problemin çözümü, m -durumlu SMM' nin $\mathcal{L}(\boldsymbol{\Theta}; \mathbf{Y}^{(T)})$ olabilirlik fonksiyon değerini vermektedir. Hesaplanan olabilirlik değeri, modelle gözlemlerin ne kadar uyduğu bir göster-

gesidir. Olabilirlik değeri arttıkça, gözlemlerin modele olan uyumu da artmaktadır. Bu bağlamda, modelin saklı durum sayısına karar verilmesinde, problemin çözümü gözlemlerle en iyi uyuşan modelin seçilmesini sağlayacaktır.

$\mathcal{L}(\Theta; \mathbf{Y}^{(T)})$ olabilirlik fonksiyonu doğrudan hesaplanabilir. $P(\mathbf{Y}^{(T)}, \mathbf{X}^{(T)} | \Theta)$ bileşik olasılığı, olası tüm durum dizileri üzerinden toplandığında olabilirlik değeri aşağıdaki gibi elde edilir:

$$\begin{aligned}\mathcal{L}(\Theta; \mathbf{Y}^{(T)}) &= P(\mathbf{Y}^{(T)} | \Theta) \\ &= \sum_{\text{Tüm } \mathbf{X}^{(T)}} P(\mathbf{Y}^{(T)}, \mathbf{X}^{(T)} | \Theta) \\ &= \sum_{\text{Tüm } \mathbf{X}^{(T)}} P(\mathbf{Y}^{(T)} | \mathbf{X}^{(T)}, \Theta) P(\mathbf{X}^{(T)} | \Theta)\end{aligned}\quad (5.5)$$

Eşitlik 5.5' te verilen herhangi bir $\mathbf{X}^{(T)}$ durum dizisinin olasılığı aşağıdaki gibi hesaplanır:

$$P(\mathbf{X}^{(T)} | \Theta) = \pi_{x_1} p_{x_1 x_2} p_{x_2 x_3} p_{x_3 x_4} \cdots p_{x_{T-2} x_{T-1}} p_{x_{T-1} x_T}\quad (5.6)$$

Gözlemlerin koşullu bağımsızlığı varsayımına göre, $\mathbf{X}^{(T)}$ durum dizisi bilindiğinde $\mathbf{Y}^{(T)}$ gözlem dizisi olasılığı,

$$P(\mathbf{Y}^{(T)} | \mathbf{X}^{(T)}, \Theta) = \prod_{t=1}^T P(Y_t | X_t, \Theta) = b_{x_1}(Y_1) b_{x_2}(Y_2) b_{x_3}(Y_3) \cdots b_{x_{T-1}}(Y_{T-1}) b_{x_T}(Y_T)\quad (5.7)$$

şeklinde hesap edilir. Eşitlik 5.6 ve Eşitlik 5.7' den yararlanılarak olabilirlik fonksiyonu,

$$\mathcal{L}(\Theta; \mathbf{Y}^{(T)}) = \sum_{\text{Tüm } \mathbf{X}^{(T)}} \pi_{x_1} b_{x_1}(Y_1) p_{x_1 x_2} b_{x_2}(Y_2) p_{x_2 x_3} b_{x_3}(Y_3) \cdots b_{x_{T-1}}(Y_{T-1}) p_{x_{T-1} x_T} b_{x_T}(Y_T)\quad (5.8)$$

olarak elde edilir. Herhangi bir $\mathbf{X}^{(T)}$ durum dizisi için, Eşitlik 5.8 sözel olarak aşağıdaki gibi açıklanır:

- Sistem başlangıç anında, π_{x_1} olasılığı ile X_1 durumunda bulunur. X_1 durumundayken, $b_{x_1}(Y_1)$ olasılığı ile Y_1 gözlemi üretilir.
- $(t = 1)$ ' den $(t = 2)$ zamanına geçerken, sistem $p_{x_1 x_2}$ geçiş olasılığı ile X_1 durumundan X_2 durumuna geçer. X_2 durumunda, $b_{x_2}(Y_2)$ olasılığı ile Y_2 gözlemi üretilir.
- $(t = 2)$ ' den $(t = 3)$ zamanına geçerken, sistem $p_{x_2 x_3}$ geçiş olasılığı ile X_2 du-

rumundan X_3 durumuna geçer. X_3 durumunda, $b_{x_3}(Y_3)$ olasılığı ile Y_3 gözlemi üretilir.

- Bu işleyiş, sistem $a_{x_{T-1}x_T}$ geçiş olasılığıyla X_{T-1} durumundan X_T durumuna geçerek $b_{x_T}(Y_T)$ olasılığı ile Y_T gözlemini üretene kadar devam eder.

Eşitlik 5.8' de verilen olabilirlik değerini **doğrudan hesaplamak için $2Tm^T$ dereceden bir hesaplama yapılması gerekmektedir.** Daha açık bir ifadeyle, $\mathbf{Y}^{(T)}$ gözlem dizisindeki her bir gözlem değerine m tane saklı durumdan herhangi biri aracılığıyla geçiş yapılabilir. T tane gözlem değeri olmasından dolayı, m^T adet saklı durum dizisinden $\mathbf{Y}^{(T)}$ gözlem dizisi üretilebilir. Her bir olası saklı durum dizisi için, $(2T - 1)$ adet çarpma işlemi yapılmaktadır. Tüm olası saklı durum dizilerinin olasılıkları toplamı için, $(m^T - 1)$ sayıda toplama işlemi yapılır. Sonuç olarak, olabilirlik değerinin **doğrudan hesaplanabilmesi için $(2T - 1)m^T$ sayıda çarpma ve $(m^T - 1)$ sayıda toplama işlemi yapmak gereklidir.** Çarpma ve toplama işlemlerinin çok fazla olması, olabilirlik değerinin hesaplanışını karmaşıktırarak zorlaştırmaktadır. m ve T ' nin küçük değerleri için bile, doğrudan hesaplama yöntemi elverişli değildir. Örneğin, sistemdeki saklı durumların sayısı $m = 3$ ve gözlemlerin sayısı $T = 50$ olduğunda, olabilirlik değerinin doğrudan hesaplanması için gerekli işlem sayısı yaklaşık olarak $2 \cdot 50 \cdot 3^{50} \approx 7 \cdot 10^{25}$ dir.

İşlem yükünün hafifletilerek olabilirlik değerinin daha etkili bir şekilde hesaplanması amacıyla, ileri-yön ve geri-yön algoritmaları geliştirilmiştir. Bu algoritmalar, doğrudan hesaplama göre, daha az karmaşıktır. Olabilirlik değerinin hesaplanması için, bu algoritmalarından herhangi biri kullanılabilir. Algoritmalar aşağıda açıklanmıştır.

İleri-Yön Algoritması

İleri-yön algoritması, $t \in \mathbf{T}$ ve $S_i \in \mathbf{S}$ olmak üzere tanımlanan $\alpha_t(i)$ ileri değişkeni üzerine kurulmuştur. $\alpha_t(i)$ değişkeni, sistemin t anında S_i durumunda olması ve $\mathbf{Y}^{(t)} = (Y_1 Y_2 \dots Y_{t-1} Y_t)$ kısmi gözlem dizisinin oluşması olasılığı olarak tanımlanır.

$$\alpha_t(i) = P(Y_1, Y_2, \dots, Y_{t-1}, Y_t, X_t = S_i | \Theta) = P(\mathbf{Y}^{(t)}, X_t = S_i | \Theta)$$

Herhangi bir t anındaki ileri-yön değişkenleri, $1 \times m$ boyutlu

$$\boldsymbol{\alpha}_t = \left[\alpha_t(1) \ \alpha_t(2) \ \dots \ \alpha_t(m-1) \ \alpha_t(m) \right]$$

vektörü ile gösterilir. İleri-yön algoritmasında, ileri değişkeni kullanılarak tümevarımsal bir yöntemle değerlendirme probleminin çözümüne aşağıdaki adımlarla ulaşılır.

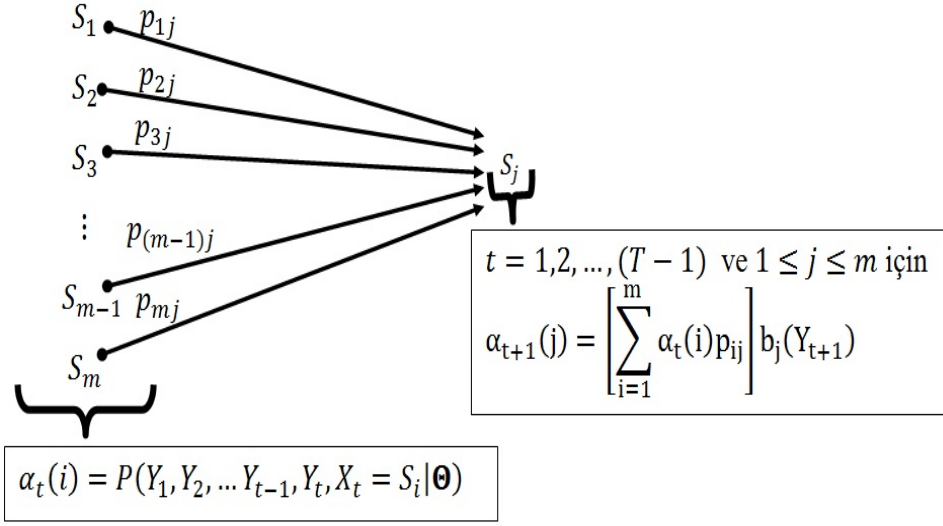
1. Başlangıç: $\forall 1 \leq i \leq m$ için, $\alpha_1(i)$ ileri değişkeni aşağıdaki gibi hesaplanır:

$$\alpha_1(i) = P(Y_1, X_1 = S_i | \Theta) = P(X_1 = S_i | \Theta) P(Y_1 | X_1 = S_i, \Theta) = \pi_i b_i(Y_1)$$

2. Yineleme: $t = 1, 2, \dots, (T-1)$ ve $\forall 1 \leq j \leq m$ için, $\alpha_{t+1}(j)$ ileri değişkeni aşağıdaki gibi hesaplanır:

$$\begin{aligned} \alpha_{t+1}(j) &= P(\mathbf{Y}^{(t+1)}, X_{t+1} = S_j | \Theta) \\ &= P(\mathbf{Y}^{(t+1)} | X_{t+1} = S_j, \Theta) P(X_{t+1} = S_j | \Theta) \\ &= P(Y_{t+1} | X_{t+1} = S_j, \Theta) P(\mathbf{Y}^{(t)} | X_{t+1} = S_j, \Theta) P(X_{t+1} = S_j | \Theta) \\ &= P(Y_{t+1} | X_{t+1} = S_j, \Theta) P(\mathbf{Y}^{(t)}, X_{t+1} = S_j | \Theta) \\ &= \left[\sum_{i=1}^m P(\mathbf{Y}^{(t)}, X_t = S_i, X_{t+1} = S_j | \Theta) \right] P(Y_{t+1} | X_{t+1} = S_j, \Theta) \\ &= \left[\sum_{i=1}^m P(\mathbf{Y}^{(t)}, X_{t+1} = S_j | X_t = S_i, \Theta) P(X_t = S_i | \Theta) \right] P(Y_{t+1} | X_{t+1} = S_j, \Theta) \\ &= \left[\sum_{i=1}^m P(\mathbf{Y}^{(t)} | X_t = S_i, \Theta) P(X_{t+1} = S_j | X_t = S_i, \Theta) P(X_t = S_i | \Theta) \right] \\ &\quad P(Y_{t+1} | X_{t+1} = S_j, \Theta) \\ &= \left[\sum_{i=1}^m P(\mathbf{Y}^{(t)}, X_t = S_i | \Theta) P(X_{t+1} = S_j | X_t = S_i, \Theta) \right] P(Y_{t+1} | X_{t+1} = S_j, \Theta) \\ &= \left[\sum_{i=1}^m \alpha_t(i) p_{ij} \right] b_j(Y_{t+1}) \end{aligned}$$

$\alpha_{t+1}(j)$ ileri değişkeninin nasıl oluştuğu aşağıdaki şekilde grafiksel olarak açıklanmıştır.



Şekil 5.2. İleri-Yön Değişkeninin Oluşumu

Şekil 5.2.'de t zamanında sistemin m sayıda saklı durumundan S_j durumuna nasıl geçildiği görülür.

3. **Sonlandırma:** T zamanı için bulunan ileri değişkenler kullanılarak, olabilirlik değeri aşağıdaki gibi hesaplanır:

$$P(\mathbf{Y}^{(T)} | \boldsymbol{\theta}) = \sum_{i=1}^m P(\mathbf{Y}^{(T)}, X_T = S_i | \boldsymbol{\theta}) = \sum_{i=1}^m \alpha_T(i)$$

İleri-yön algoritması kullanılarak olabilirlik değerinin hesaplanmasındaki işlem yükü, $m^2 T^2$ ye indirilmiştir. İleri-yön algoritmasında, $m(m+1)(T-1) + m$ sayıda çarpma ve $m(m-1)(T-1)$ sayıda toplama işlemi yapılmaktadır. Bu durumda, işlem yükünde $2m^{(T-2)}$ kat tasarruf sağlanır.

Geri-Yön Algoritması

Geri-yön algoritması, $t \in T$ ve $S_i \in S$ olmak üzere tanımlanan $\beta_t(i)$ geri değişkeni üzerine kurulmuştur. $\beta_t(i)$ değişkeni, sistemin t anında S_i durumunda olduğu bilindiğinde, $\mathbf{Y}_{t+1}^{(T)} = (Y_{t+1} Y_{t+2} \dots Y_{T-1} Y_T)$ kısmi gözlem dizisinin oluşması olasılığı olarak tanımlanır.

$$\beta_t(i) = P(Y_{t+1}, Y_{t+2}, \dots, Y_{T-1}, Y_T | X_t = S_i, \boldsymbol{\theta}) = P(\mathbf{Y}_{t+1}^{(T)} | X_t = S_i, \boldsymbol{\theta}) \quad (5.9)$$

Herhangi bir t anındaki geri-yön değişkenleri, $1 \times m$ boyutlu

$$\boldsymbol{\beta}_t = [\beta_t(1) \beta_t(2) \dots \beta_t(m-1) \beta_t(m)]$$

vektörü ile gösterilir. Geri-yön algoritmasında, geri değişkeni kullanılarak tümevarımsal bir yöntemle değerlendirme probleminin çözümüne aşağıdaki adımlarla ulaşılır:

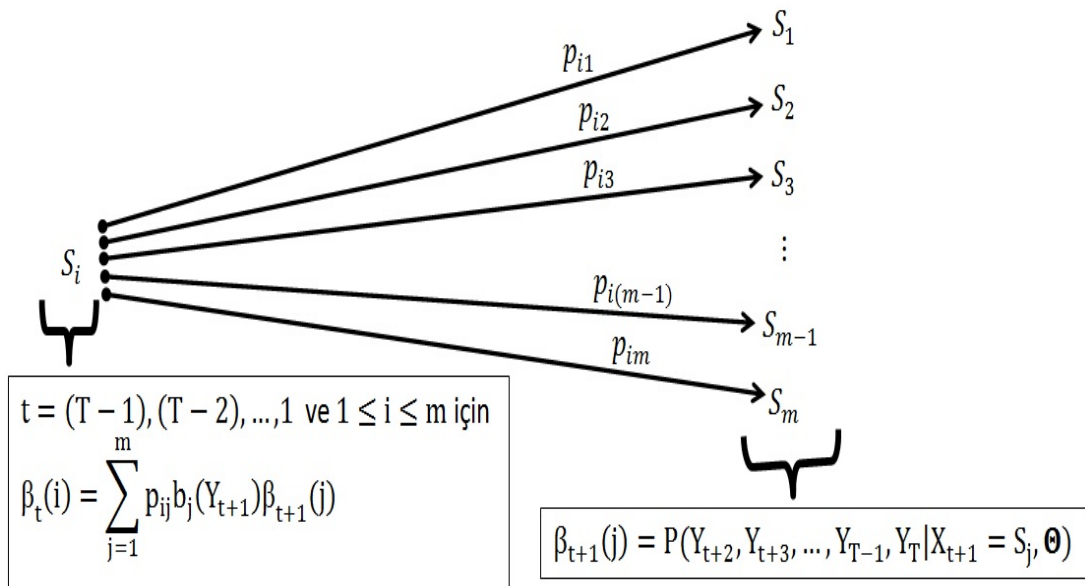
1. **Başlangıç:** T zamanında, $\forall 1 \leq i \leq m$ için, geri değişkenine keyfi olarak bir değeri verilir.

$$\beta_T(i) = P(Y_{T+1} | X_T = S_i, \Theta) = 1 \quad (5.10)$$

2. **Yineleme:** $t = (T - 1), (T - 2), \dots, 1$ ve $\forall 1 \leq i \leq m$ için, $\beta_t(i)$ geri değişkeni aşağıdaki gibi hesaplanır.

$$\begin{aligned} \beta_t(i) &= P(\mathbf{Y}_{t+1}^{(T)} | X_t = S_i, \Theta) \\ &= \sum_{j=1}^m P(\mathbf{Y}_{t+1}^{(T)}, X_{t+1} = S_j | X_t = S_i, \Theta) \\ &= \sum_{j=1}^m P(X_{t+1} = S_j | X_t = S_i, \Theta) P(\mathbf{Y}_{t+1}^{(T)} | X_t = S_i, X_{t+1} = S_j, \Theta) \\ &= \sum_{j=1}^m P(X_{t+1} = S_j | X_t = S_i, \Theta) P(Y_{t+1} | X_t = S_i, X_{t+1} = S_j, \Theta) \\ &\quad P(\mathbf{Y}_{t+2}^{(T)} | X_t = S_i, X_{t+1} = S_j, \Theta) \\ &= \sum_{j=1}^m P(X_{t+1} = S_j | X_t = S_i, \Theta) P(Y_{t+1} | X_{t+1} = S_j, \Theta) P(\mathbf{Y}_{t+2}^{(T)} | X_{t+1} = S_j, \Theta) \\ &= \sum_{j=1}^m p_{ij} b_j(Y_{t+1}) \beta_{t+1}(j) \end{aligned}$$

$\beta_t(i)$ geri değişkeninin nasıl oluştuğu aşağıdaki şekilde grafiksel olarak açıklanmıştır.



Şekil 5.3. Geri-Yön Değişkeninin Oluşumu

Şekil 5.3.' te t zamanında sistemin S_i durumundan m sayıda saklı duruma nasıl geçtiği görülür.

3. **Sonlandırma:** ($t = 1$) başlangıç zamanı için bulunan geri değişkenleri kullanılarak olabilirlik değeri aşağıdaki gibi hesaplanır.

$$P(\mathbf{Y}^{(T)} | \Theta) = \sum_{i=1}^m \pi_i b_i(Y_1) \beta_1(i) \quad (5.11)$$

Geri-yön algoritması kullanılarak olabilirlik değerinin hesaplanmasındaki işlem yükü, m^2T ' ye indirilmiştir. Dolayısıyla, ileri-yön ve geri-yön algoritmaları arasında işlem yükü bakımından fark yoktur. Bu iki algoritmadan herhangi biri kullanılarak olabilirlik değeri hesaplanabilir; hatta İleri-yön ve geri-yön değişkenleri birlikte kullanılarak olabilirlik değeri elde edilebilir. Herhangi bir t anında, ileri-yön ve geri-yön değişkenlerinin çarpımları

$$\alpha_t(i)\beta_t(i) = P(\mathbf{Y}^{(t)}, X_t = S_i | \Theta) P(\mathbf{Y}_{t+1}^{(T)} | X_t = S_i, \Theta) = P(\mathbf{Y}^{(T)}, X_t = S_i | \Theta)$$

olarak bulunur. Bu durumda, olabilirlik fonksiyonu aşağıdaki gibi hesaplanabilir:

$$\mathcal{L}(\Theta; \mathbf{Y}^{(T)}) = \sum_{i=1}^m \alpha_t(i)\beta_t(i) = \alpha_t \beta_t^T$$

5.3.2. Kodlama Problemi ve Çözüm Yöntemi

Kodlama problemi, $\mathbf{Y}^{(T)}$ gözlem dizisi ve Θ model parametreleri verildiğinde, belirtilen gözlemleri en yüksek olasılıkla üreten optimal $\mathbf{X}^{(T)}$ saklı durum dizisinin belirlenmesidir. Bu problemin çözümü, modelin saklı olan kısmını açığa çıkarmaktadır.

Belirli bir gözlem dizisine ait optimal saklı durum dizisi farklı yaklaşımlar altında elde edilebilir. Kodlama probleminde, optimal saklı durum dizisinin tanımı çok önemlidir. Çünkü, birden fazla optimallik kriteri tanımlanabilir ve tanımlanan kriterlere göre elde edilecek çözüm farklılaşır. Burada, verilmiş bir gözlem dizisi için her bir $t \in T$ zamanında tek başına en yüksek olasılıkla gerçekleşen saklı durum seçilebilir. Bu yaklaşım, yerel kodlama (local decoding) olarak adlandırılır. Yerel kodlamada problemin çözümüne ulaşmak için $\gamma_t(i)$ değişkeni tanımlanır. $\gamma_t(i)$ değişkeni, Θ model parametreleri ve $\mathbf{Y}^{(T)}$ gözlem

dizisi bilindiğinde, t anında sistemin S_i durumunda bulunması olasılığıdır:

$$\gamma_t(i) = P(X_t = S_i | \mathbf{Y}^{(T)}, \Theta)$$

$\alpha_t(i)$ değişkeni, $\mathbf{Y}^{(t)}$ kısmi gözlem dizisini ve t anında sistemin S_i durumunda bulunmasını açıklarken; $\beta_t(i)$ değişkeni, sistemin t anında S_i durumunda bulunduğu bilindiğinde, geriye kalan $\mathbf{Y}_{t+1}^{(T)}$ kısmi gözlem dizisini açıklamaktadır. Bu nedenle, $\gamma_t(i)$ değişkeni, ileri ve geri değişkenleri cinsinden aşağıdaki gibi ifade edilebilir:

$$\gamma_t(i) = P(X_t = S_i | \mathbf{Y}^{(T)}, \Theta) = \frac{P(\mathbf{Y}^{(T)}, X_t = S_i | \Theta)}{P(\mathbf{Y}^{(T)} | \Theta)} = \frac{\alpha_t(i) \beta_t(i)}{\left[\sum_{i=1}^m \alpha_t(i) \beta_t(i) \right]}$$

Burada, $P(\mathbf{Y}^{(T)} | \Theta)$ gözlem dizisi olasılığı bir normalleştirme katsayısıdır ve $\forall t \in \mathbf{T}$ zamanında $\gamma_t(\cdot)$ değişkenini bir olasılık ölçüsü yapar. Dolayısıyla, her bir zaman noktasında aşağıdaki eşitlik sağlanır:

$$\sum_{i=1}^m \gamma_t(i) = 1$$

Optimal durum dizisi, her bir $t \in \mathbf{T}$ için aşağıdaki en büyükleme argümanı gerçekleştirilerek elde edilir:

$$X_t^* = \arg \max_{1 \leq i \leq m} [\gamma_t(i)]$$

Bu işlem sonucunda, her bir t zamanında $P(X_t = S_i | \mathbf{Y}^{(T)}, \Theta)$ koşullu olasılığını tek başına en büyükleyen X_t^* durumu seçilerek $\mathbf{X}^{*(T)} = (X_1^* X_2^* \dots X_T^*)$ optimal durum dizisi oluşturulur. Yerel kodlamada elde edilen optimal durum dizisi ile ilgili problemler oluşabilir. Örneğin, X saklı süreci ergodik olmayan bir Markov zinciri olduğunda bazı durumlar arası geçiş olasılıkları sıfır olabilir. Bu durumda, yerel kodlamada durum dizilerinin oluşumu dikkate alınmadan sadece her bir t anında en olası durum belirlendiği için elde edilen optimal durum dizisi gerçekte geçerli olmayan bir durum dizisi olabilir. Bu problemin giderilmesi için optimallik kriteri modifiye edilir. Birden fazla zaman noktası için en çok olası saklı durum dizileri bulunabilir. Örneğin, durum çiftlilerinin (üçlülerinin vb.) olasılığı en büyüklenecek şekilde kodlama probleminin çözümüne ulaşılabilir.

Kodlama probleminin çözümünün elde edilmesinde en çok kullanılan yaklaşım, en iyi tekil durum dizisinin bulunmasıdır. Verilmiş bir gözlem dizisi için tüm zaman noktaları

üzerinden durum dizileri olasılığı en büyüklendiğinde en iyi tekil durum dizisi elde edilir. Bu yaklaşıma, genel kodlama (global decoding) denilmektedir. En iyi tekil durum dizisine, $P(\mathbf{X}^{(T)} | \mathbf{Y}^{(T)}, \Theta)$ koşullu olasılığı en büyüklenerek ulaşılır. Tüm durum dizileri üzerinden bu koşullu olasılığın en büyüklenmesi işlemi m^T tane fonksiyonun değerlendirilmesini gerektirmektedir. Çok küçük T değerleri dışında, bu işlemlerin gerçekleştirilmesi mümkün değildir. Koşullu olasılık tanımından, aşağıdaki doğru orantı geçerlidir:

$$\max (P(\mathbf{X}^{(T)} | \mathbf{Y}^{(T)}, \Theta)) \propto \max (P(\mathbf{X}^{(T)}, \mathbf{Y}^{(T)} | \Theta)) \quad (5.12)$$

Bu durumda, $P(\mathbf{X}^{(T)}, \mathbf{Y}^{(T)} | \Theta)$ bileşik olasılığı en büyüklenerek en iyi tekil durum dizisine ulaşılır. Bunun için, bir dinamik programlama yöntemi olan Viterbi algoritması kullanılmaktadır. Viterbi algoritması, 1967 yılında Andrew Viterbi [80] tarafından önerilmiştir. Bu algoritma, kesikli dizin kümeli sonlu durumlu Markov süreçlerinin durum dizilerinin tahmin edilmesinde kullanılmaktadır. Hem durağan hem de durağan olmayan Markov süreçleri için uygulanabilir. Bu algoritma aşağıda açıklanmıştır.

Viterbi Algoritması

Viterbi algoritmasının gerçekleştirilebilmesi için $\delta_t(i)$ değişkeni tanımlanır. Bu değişken, $\mathbf{Y}^{(t)}$ kısmi gözlem dizisini açıklayan ve t anında S_i durumunda sonlanan tek bir $\mathbf{X}^{(t)}$ kısmi durum dizisi için en yüksek olasılık değeri olarak ifade edilir:

$$\begin{aligned} \delta_t(i) &= \max_{X_1, \dots, X_{t-1}} P(X_1, X_2, \dots, X_{t-1}, X_t = S_i, Y_1, Y_2, \dots, Y_t | \Theta) \\ &= \max_{\mathbf{X}^{(t-1)}} P(\mathbf{X}^{(t-1)}, X_t = S_i, \mathbf{Y}^{(t)} | \Theta) \end{aligned}$$

Bir sonraki zaman noktası için, en yüksek olasılık değeri aşağıdaki gibi hesaplanabilir:

$$\delta_{t+1}(j) = \left[\max_{1 \leq i \leq m} \delta_t(i) p_{ij} \right] b_j(Y_{t+1})$$

Optimal durum dizisine ulaşmak için, $\delta_{t+1}(j)$ değerini her t ve j için en büyükleyen saklı durumun sürekli olarak izlenmesi gerekir. Bu işlem, $\psi_t(j)$ sıralı dizini kullanılarak yapılır. Viterbi algoritmasının adımları aşağıda verilmiştir:

1. **Başlangıç:** $\forall 1 \leq i \leq m$ için, $t = 1$ zamanında $\delta_1(i)$ değişkeni aşağıdaki gibi

hesaplanır ve sıralı dizinin değeri sıfır olarak atanır:

$$\delta_1(i) = \pi_i b_i(Y_1)$$

$$\psi_1(i) = 0$$

2. **Yineleme:** $\forall 2 \leq t \leq T$ ve $1 \leq j \leq m$ için, aşağıdaki değişkenler hesap edilir:

$$\delta_t(j) = \max_{1 \leq i \leq m} [\delta_{t-1}(i) p_{ij}] b_j(Y_t)$$

$$\psi_t(j) = \arg \max_{1 \leq i \leq m} [\delta_{t-1}(i) p_{ij}]$$

3. **Sonlandırma:** Y_T son gözlem değeri için saklı durum aşağıdaki gibi bulunur:

$$P^* = \max_{1 \leq i \leq m} [\delta_T(i)]$$

$$X_T^* = \arg \max_{1 \leq i \leq m} [\delta_T(i)]$$

4. **Durum Dizisi Geri İzleme:** T zaman periyodunda geri yönde ilerlenerek, her bir t zamanı için en iyi tekil durum aşağıdaki gibi elde edilir:

$$X_t^* = \psi_{t+1}(X_{t+1}^*)$$

Böylece, $\mathbf{X}^{*(T)} = (X_1^* X_2^* \dots X_T^*)$ en iyi tekil durum dizisine ulaşılır.

5.3.3. Öğrenme Problemi ve Çözüm Yöntemi

Öğrenme problemi, $\mathcal{L}(\Theta; \mathbf{Y}^{(T)})$ olabilirlik fonksiyonu en büyüklenecek şekilde model parametrelerinin tahmin edilmesidir. $\mathcal{L}(\Theta; \mathbf{Y}^{(T)})$, çok boyutlu ve Θ parametrelerinin karmaşık bir fonksiyonudur. Olabilirlik fonksiyonunun bilinmeyen parametrelere göre ayrı ayrı türevi alındığında ortak çözüm için saklı durumlara ilişkin bilgiye ihtiyaç duyulur. X saklı sürecine ilişkin bilginin bilinmemesi nedeniyle, parametrelerin en çok olabilirlik tahmin edicileri olabilirlik fonksiyonundan doğrudan hesaplanamaz. Bu nedenle, öğrenme probleminin analitik bir çözümü yoktur. Olabilirlik fonksiyonunu yerel olarak en büyükleme iteratif yöntemler kullanılarak model parametreleri tahmin edilir.

SMM parametrelerinin klasik yaklaşımla tahmin edilmesinde, Baum vd. [72] tarafından iteratif bir yöntem olan Baum-Welch algoritması geliştirilmiştir. Bu algoritma, sistemin

izlediği X saklı sürecinin homojen bir Markov zinciri olduğu varsayımı altında, Θ parametrelerinin en çok olabilirlik tahmin edicilerinin elde edilmesi için tasarlanmıştır. X saklı sürecinin durağan olması şart değildir. Baum-Welch algoritması, Dempster vd. [81] tarafından tanımlanan beklenti en büyükleme (Expectation Maximization - EM) yönteminin özel bir durumudur. EM algoritması, kayıp veri durumunda model parametrelerinin en çok olabilirlik tahminlerini hesaplamak amacıyla kullanılan iteratif bir yöntemdir. SMM' de gözlemlenemeyen $X^{(T)}$ saklı durum dizisi kayıp veri olarak düşünülebilir. Model parametrelerinin tahmininde alternatif olarak, kesitsel K -ortalamalar yöntemi (Segmental K-Means Method) veya gradyan yöntemler (Gradient Methods) gibi iteratif yöntemler de kullanılabilir. Ancak, iteratif yöntemler, parametre tahminleri için mutlak en büyük (global maximum) değeri vermeyi garanti etmez. Parametrelerin başlangıç değerlerine bağlı olarak, farklı yerel en büyük (local maximum) değerlere de ulaşılabilir. Bu nedenle, parametrelerin en uygun başlangıç değerleri seçilmelidir.

Baum-Welch algoritması, parametrelerin en çok olabilirlik tahminlerini bulmak için EM algoritmasını kullanmaktadır. Bu nedenle, model parametrelerinin elde edilmesi EM algoritması üzerinden ele alınmıştır.

EM algoritmasında, kayıp ve gözlenen verilerin birleşiminden oluşan tam veri kullanılmaktadır. SMM' de $X^{(T)}$ saklı durum dizisi ve $Y^{(T)}$ gözlemleri tam veriyi oluşturmaktadır. Tam veri olabilirlik fonksiyonu aşağıdaki gibi yazılır:

$$\begin{aligned} L(\Theta | X^{(T)}, Y^{(T)}) &= P(X^{(T)}, Y^{(T)} | \Theta) \\ &= P(X^{(T)} | Y^{(T)}, \Theta) P(Y^{(T)} | \Theta) \end{aligned} \quad (5.13)$$

Eşitlik 5.13' ün logaritması alındığında, tam veri log-olabilirlik fonksiyonuna aşağıdaki gibi elde edilir:

$$\begin{aligned} \ln L(\Theta | X^{(T)}, Y^{(T)}) &= \ln P(X^{(T)} | Y^{(T)}, \Theta) + \ln P(Y^{(T)} | \Theta) \\ &= \ln P(X^{(T)} | Y^{(T)}, \Theta) + \ln L(\Theta | Y^{(T)}) \end{aligned} \quad (5.14)$$

Model parametrelerinin en çok olabilirlik tahmin edicileri, $\ln L(\Theta | Y^{(T)})$ olabilirlik fonksiyonunu en büyükleyen değerlerdir. Bu durumda, Eşitlik 5.14' ten olabilirlik fonksiyonu

çekilerek aşağıdaki gibi yazılır:

$$\ln L(\Theta | Y^{(T)}) = \ln L(\Theta | X^{(T)}, Y^{(T)}) - \ln P(X^{(T)} | Y^{(T)}, \Theta) \quad (5.15)$$

$g \in \{1, 2, \dots, G\}$, iterasyon sayısı olsun. Parametrelerin başlangıç değeri $\Theta^{(0)}$, g -inci iterasyondaki tahmin değeri $\Theta^{(g)}$ ve en çok olabilirlik tahmin edicisi $\hat{\Theta}$ olarak gösterilsin. $Y^{(T)}$ gözlemleri ve $\Theta^{(g)}$ tahmin değeri bilindiğinde, X saklı sürecin dağılımı üzerinden Eşitlik 5.15' in beklenen değeri alındığında aşağıdaki eşitlik elde edilir:

$$E_{X|Y^{(T)}, \Theta^{(g)}} [\ln L(\Theta | Y^{(T)})] = E_{X|Y^{(T)}, \Theta^{(g)}} [\ln L(\Theta | X^{(T)}, Y^{(T)})] - E_{X|Y^{(T)}, \Theta^{(g)}} [\ln P(X^{(T)} | Y^{(T)}, \Theta)] \quad (5.16)$$

Eşitlik 5.16' nın sol tarafı X saklı sürecinden bağımsızdır. Bu nedenle, sabit olarak işlem görür ve beklenen değer fonksiyonundan etkilenmez. Bu durumda, Eşitlik 5.16 aşağıdaki gibi yeniden düzenlenebilir:

$$\begin{aligned} \ln L(\Theta | Y^{(T)}) &= \int_{X|Y^{(T)}, \Theta^{(g)}} [\ln L(\Theta | X^{(T)}, Y^{(T)})] P(X | Y^{(T)}, \Theta^{(g)}) dX - \\ &\quad - \int_{X|Y^{(T)}, \Theta^{(g)}} [\ln P(X^{(T)} | Y^{(T)}, \Theta)] P(X | Y^{(T)}, \Theta^{(g)}) dX \\ &= Q(\Theta | \Theta^{(g)}) - H(\Theta | \Theta^{(g)}) \end{aligned} \quad (5.17)$$

Eşitlik 5.17' de $[-H(\Theta | \Theta^{(g)})]$ fonksiyonu, $(X | Y^{(T)}, \Theta^{(g)})$ ' nin entropi değeridir. Burada, $Q(\Theta | \Theta^{(g)})$ koşullu beklenen değerini en büyükleyen model parametreleri aynı zamanda olabilirlik fonksiyonunda en büyükleyecektir:

$$\arg \max_{\Theta} Q(\Theta | \Theta^{(g)}) = \arg \max_{\Theta} \ln L(\Theta | Y^{(T)})$$

Bu durumda, model parametrelerinin en çok olabilirlik tahmin edicileri, $Q(\Theta | \Theta^{(g)})$ koşullu beklenen değeri en büyüklenecek bulunur. EM algoritmasının her yinelemesi iki adımdan oluşur. Bunlar, beklenti adımı (E-adımı) ve en büyükleme adımı (M-adımı)' dir. Bu adımlar aşağıda açıklanmıştır:

- **E-Adımı:** $\Theta^{(g)}$ bir önceki adımda hesaplanan parametrelerin tahminleri ve $Y^{(T)}$

gözlemleri verildiğinde, aşağıdaki koşullu beklenen değer hesaplanır:

$$Q(\Theta | \Theta^{(g)}) = E_{\mathbf{X}|\mathbf{Y}^{(T)}, \Theta^{(g)}} [\ln L(\Theta | \mathbf{X}^{(T)}, \mathbf{Y}^{(T)})]$$

- **M-Adımı:** E-adımında hesaplanan koşullu beklenen değeri en büyük yapan parametreler tahmin edilir:

$$\Theta^{(g+1)} = \arg \max_{\Theta} Q(\Theta | \Theta^{(g)})$$

EM algoritmasının adımları aşağıda verilmiştir:

Adım 1: $\Theta^{(0)}$ parametrelerinin başlangıç değerlerinin ve $\mathbf{Y}^{(T)}$ gözlemlerinin verilmesi,

Adım 2: E-adımında, $Q(\Theta | \Theta^{(0)})$ koşullu beklenen değerinin hesaplanması,

Adım 3: M-Adımında, E-adımında hesaplanan koşullu beklenen değeri en büyük yapan $\Theta^{(1)}$ tahmin değerinin elde edilmesi ve

Adım 4: İkinci ve üçüncü adımların belirli bir yakınsaklık kriteri sağlanıncaya kadar tekrar edilmesidir. Burada $\varepsilon > 0$ olmak üzere, yakınsaklık kriteri

$$L(\Theta^{(g+1)} | \mathbf{X}, \mathbf{Y}) - L(\Theta^{(g)} | \mathbf{X}, \mathbf{Y}) > \varepsilon$$

olarak tanımlanır. Yukarıdaki eşitsizlik sağlandığı sürece 2. adıma geri dönülür. Aksi durumda, en son elde edilen tahmin değerleri en çok olabirlik tahmin değerleri olarak kabul edilir.

Gerçekleşen her iterasyon sonucunda $L(\Theta | \mathbf{X}^{(T)}, \mathbf{Y}^{(T)})$ tam veri olabirlik fonksiyonunun değeri azalmaz. Bunun sonucu olarak, her iterasyonda $P(\mathbf{Y}^{(T)} | \Theta)$ gözlem olasılığı artmaktadır. Algoritmadan elde edilen $(\Theta^{(1)}, \Theta^{(2)}, \Theta^{(3)}, \dots)$ tahmin dizisi ise, $\hat{\Theta}$ en çok olabirlik tahmin edicisine yakınsamaktadır.

EM algoritmasının uygulanabilmesi için, X saklı sürecine ilişkin aşağıda verilen u ve ν gösterge değişkenleri tanımlanır:

$$u_i(t) = \begin{cases} 1, & X_t = S_i \text{ ise,} \\ 0, & X_t \neq S_i \text{ ise.} \end{cases}$$

$$\nu_{ij}(t) = \begin{cases} 1, & X_t = S_i \text{ ve } X_{t+1} = S_j \text{ ise,} \\ 0, & X_t \neq S_i \text{ ve } X_{t+1} \neq S_j \text{ ise.} \end{cases}$$

Bu deęişkenlerden yararlanılarak saklı Markov modelinin tam veri log-olabilirlik fonksiyonu ařaęıdaki gibi elde edilir:

$$\begin{aligned} \ln L(\Theta | \mathbf{X}^{(T)}, \mathbf{Y}^{(T)}) &= \ln P(\mathbf{X}^{(T)}, \mathbf{Y}^{(T)} | \Theta) \\ &= \ln \left[\pi_{x_1} b_{x_1}(Y_1) p_{x_1 x_2} b_{x_2}(Y_2) \dots b_{x_{T-1}}(Y_{T-1}) p_{x_{T-1} x_T} b_{x_T}(Y_T) \right] \\ &= \ln \left[\pi_{x_1} \prod_{t=1}^{T-1} p_{x_t x_{t+1}} \prod_{t=1}^T b_{x_t}(Y_t) \right] \\ &= \ln \pi_{x_1} + \sum_{t=1}^{T-1} \ln p_{x_t x_{t+1}} + \sum_{t=1}^T \ln b_{x_t}(Y_t) \\ &= \sum_{i=1}^m u_i(1) \ln \pi_i + \left[\sum_{i=1}^m \sum_{j=1}^m \left(\sum_{t=1}^{T-1} \nu_{ij}(t) \right) \ln p_{ij} \right] + \\ &\quad + \sum_{i=1}^m \sum_{t=1}^T u_i(t) \ln b_i(Y_t) \end{aligned} \quad (5.18)$$

SMM için gerekleřtirilen EM algoritmasının E-adımı ve M-adımı ařaęıda kısaca aıklanmıřtır:

E-Adımı

Eřitlik 5.18' den elde edilen tam veri log-olabilirlik fonksiyonundan yararlanılarak, $Q(\Theta | \Theta^{(g)})$ kořullu beklenen deęeri ařaęıdaki gibi hesaplanır:

$$\begin{aligned} Q(\Theta | \Theta^{(g)}) &= E_{\mathbf{X} | \mathbf{Y}^{(T)}, \Theta^{(g)}} \left[\ln L(\Theta | \mathbf{X}^{(T)}, \mathbf{Y}^{(T)}) \right] \\ &= \sum_{i=1}^m E[u_i(t)] \ln \pi_i + \sum_{i=1}^m \sum_{j=1}^m \left(\sum_{t=1}^{T-1} E[\nu_{ij}(t)] \right) \ln p_{ij} + \sum_{i=1}^m \sum_{t=1}^T E[u_i(t)] \ln b_i(Y_t) \end{aligned}$$

Burada, u ve ν gosterge deęişkenlerinin kořullu beklenen deęerleri ařaęıdaki gibi elde edilirler:

$$\begin{aligned} E_{\mathbf{X} | \mathbf{Y}^{(T)}, \Theta^{(g)}} [u_i(t)] &= P(X_t = S_i | \mathbf{Y}^{(T)}, \Theta^{(g)}) \\ E_{\mathbf{X} | \mathbf{Y}^{(T)}, \Theta^{(g)}} [\nu_{ij}(t)] &= P(X_t = S_i, X_{t+1} = S_j | \mathbf{Y}^{(T)}, \Theta^{(g)}) \end{aligned}$$

u' nun kořullu beklenen deęeri, kodlama probleminde tanımlanan $\gamma_t(i)$ deęişkenine eřittir. ν' nün kořullu beklenen deęeri ise, $\xi_t(i, j) = P(X_t = S_i, X_{t+1} = S_j | \mathbf{Y}^{(T)}, \Theta)$

değişkeni olarak tanımlansın. Bu durumda, Q fonksiyonu aşağıdaki gibi elde edilir:

$$Q(\Theta | \Theta^{(g)}) = \sum_{i=1}^m \gamma_1(i) \ln \pi_i + \sum_{i=1}^m \sum_{j=1}^m \left(\sum_{t=1}^{T-1} \xi_t(i, j) \right) \ln p_{ij} + \sum_{i=1}^m \sum_{t=1}^T \gamma_t(i) \ln b_i(Y_t) \quad (5.19)$$

M-Adımı

Eşitlik 5.19' da tanımlanan Q fonksiyonunu en büyükleyen Θ model parametreleri elde edilir. Eşitliğin sağ tarafındaki ilk terimin en büyüklenmesi, π başlangıç durum vektörünü; ikinci terimin en büyüklenmesi, P geçiş olasılıkları matrisini; üçüncü terimin en büyüklenmesi ise, $S_i \in \mathcal{S}$ dağılım ailesi parametrelerini verir.

Başlangıç durum vektörünün en çok olabilirlik tahmin edicisi aşağıdaki gibi hesaplanır:

$$\frac{\partial}{\partial \pi} \left(\sum_{i=1}^m \gamma_1(i) \ln \pi_i \right) = 0 \quad (5.20)$$

Burada, $\pi_m = 1 - (\pi_1 + \pi_2 + \dots + \pi_{m-1})$ olarak yazıldığında ve Eşitlik 5.20' de verilen türev herhangi bir π_j ($j = 1, \dots, m-1$) için sıfıra eşitlendiğinde,

$$\begin{aligned} \frac{\partial}{\partial \pi_j} \left(\sum_{i=1}^m \gamma_1(i) \ln \pi_i \right) &= 0 \\ \gamma_1(j) \left(\frac{1}{\pi_j} \right) + \gamma_1(m) \left(\frac{-1}{1 - \pi_1 - \dots - \pi_{m-1}} \right) &= 0 \\ \frac{\gamma_1(j)}{\pi_j} - \frac{\gamma_1(m)}{\pi_m} &= 0 \\ \frac{\pi_j}{\gamma_1(j)} &= \frac{\pi_m}{\gamma_1(m)} \end{aligned}$$

sonucuna ulaşılır. Elde edilen sonuçlar ortak çözüm olarak yazıldığında, aşağıdaki orantı oluşur:

$$\frac{\pi_1}{\gamma_1(1)} = \frac{\pi_2}{\gamma_1(2)} = \dots = \frac{\pi_{m-1}}{\gamma_1(m-1)} = \frac{\pi_m}{\gamma_1(m)} = \frac{\sum_{i=1}^m \pi_i}{\sum_{i=1}^m \gamma_1(i)} = 1$$

Bu orantıya göre, i -inci ($i = 1, \dots, m$) başlangıç durum olasılığının en çok olabilirlik tahmin edicisi,

$$\hat{\pi}_i = \gamma_1(i) = P(X_1 = S_i | \mathbf{Y}^{(T)}, \boldsymbol{\Theta}^{(g)}) \quad (5.21)$$

olarak bulunur.

Geçiş olasılıkları matrisinin en çok olabilirlik tahmin edicisi aşağıdaki gibi hesaplanır:

$$\frac{\partial}{\partial \mathbf{P}} \left(\sum_{i=1}^m \sum_{j=1}^m \left(\sum_{t=1}^{T-1} \xi_t(i, j) \right) \ln p_{ij} \right) = 0 \quad (5.22)$$

$k = 1, \dots, m$ olmak üzere, $p_{km} = 1 - (a_{k1} + a_{k2} + \dots + a_{k(m-1)})$ olarak yazıldığında ve Eşitlik 5.22' de verilen türev, herhangi bir p_{kl} ($l = 1, \dots, m - 1$) geçiş olasılığı için sıfıra eşitlendiğinde,

$$\frac{\partial}{\partial p_{kl}} \left(\sum_{i=1}^m \sum_{j=1}^m \left(\sum_{t=1}^{T-1} \xi_t(i, j) \right) \ln p_{ij} \right) = 0$$

$$\sum_{t=1}^{T-1} \xi_t(k, l) \left(\frac{1}{p_{kl}} \right) + \xi_t(k, m) \left(\frac{-1}{p_{km}} \right) = 0$$

$$\frac{\sum_{t=1}^{T-1} \xi_t(k, l)}{p_{kl}} - \frac{\sum_{t=1}^{T-1} \xi_t(k, m)}{p_{km}} = 0$$

$$\frac{p_{kl}}{\sum_{t=1}^{T-1} \xi_t(k, l)} = \frac{p_{km}}{\sum_{t=1}^{T-1} \xi_t(k, m)}$$

sonucuna ulaşılır. Elde edilen sonuçlar ortak çözüm olarak yazıldığında aşağıdaki orantı oluşur:

$$\frac{p_{k1}}{\sum_{t=1}^{T-1} \xi_t(k, 1)} = \dots = \frac{p_{k(m-1)}}{\sum_{t=1}^{T-1} \xi_t(k, (m-1))} = \frac{p_{km}}{\sum_{t=1}^{T-1} \xi_t(k, m)} = \frac{\sum_{l=1}^m p_{kl}}{\sum_{t=1}^{T-1} \sum_{l=1}^m \xi_t(k, l)} = \frac{1}{\sum_{t=1}^{T-1} \gamma_t(k)}$$

Bu orantıya göre, p_{ij} ($i, j = 1, \dots, m$) durum geçiş olasılığının en çok olabilirlik tahmin edicisi,

$$\hat{p}_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i, j)}{\sum_{t=1}^{T-1} \gamma_t(i)} = \frac{S_i \text{ durumundan } S_j \text{ durumuna yapılan geçişlerin beklenen sayısı}}{S_i \text{ durumundan yapılan geçişlerin beklenen sayısı}}$$

olarak bulunur.

Eşitlik 5.19' un sağ tarafındaki üçüncü terimde yer alan $b_i(Y_t)$ gözlem olasılığı aracılığıyla, dağılım ailesi parametrelerinin en çok olabilirlik tahmin edicileri elde edilir. **Üçüncü terimin ilgili dağılım parametrelerine göre en büyüklenmesi işlemi, $b_i(\cdot)$ ' ye bağlı olarak kolay ya da zor olabilir. Örneğin, Normal ve Poisson dağılım ailelerinde dağılım parametreleri analitik olarak elde edilebilmektedir. Buna karşın, beta, gamma ve negatif binom gibi dağılım ailelerinde, üçüncü terimin dağılım parametrelerine göre en büyüklenmesi işlemi analitik olarak mümkün değildir. Bu durumda, M-adımının bu kısmında sayısal yöntemler kullanılarak dağılım parametrelerine ulaşılması gerekmektedir.**

Bu bölümde saklı Markov modelleri ve kullanılan standart tahmin yöntemleri genel hatlarıyla tanıtılmıştır. Burada verilen metodoloji, ele alınan dağılım ailesi Normal ya da Poisson olduğunda kapalı formda sonuçların elde edilmesine olanak sağlamaktadır. Klasik tahmin yöntemleri kullanılarak, Poisson saklı Markov modeliyle ülkemizin bir bölgesinde deprem sayılarının tahmini Can vd. [82] çalışmasıyla ortaya konulmuştur. Bu çalışma, jeolojik açıdan önemli fay hatlarıyla kuşatılmış ülkemizde, SMM' nin bu alana uygulandığı ilk çalışmadır. Bunun yanı sıra Normal dağılım ailesi içinde saklı Markov modeli uygulaması Can ve Ergün [83] ile yapılmıştır. Her iki çalışmada da model özellikleri için özgün kodlar yazılmıştır.

Tez çalışmasının izleyen bölümlerinde, bir kesikli bir de sürekli dağılım için SMM ele alınmıştır. Özellikle, Poisson ve beta dağılım aileleri varsayımı altında model yapılarının nasıl olacağı, parametre tahminleri ve gelecek gözlem değerlerinin öngörülmesi için nasıl bir metodolojinin izlenebileceği verilmiştir.

6. POISSON SAKLI MARKOV MODELİ

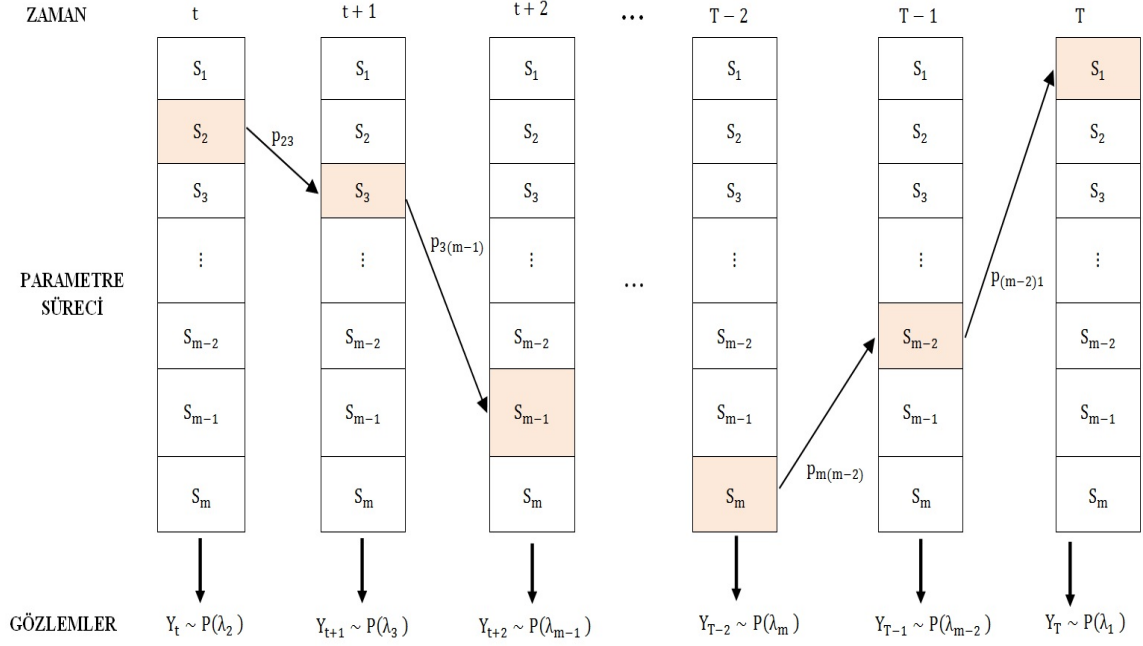
Poisson dağılımı, sayımla elde edilen verilerin açıklanmasında kullanılan geleneksel bir yaklaşımdır. Bu nedenle, belli bir zaman içerisinde meydana gelen olayların (deprem, salgın hastalıklar, sel, tsunami gibi) frekanslarının stokastik modellenmesinde Poisson süreçleri tercih edilmektedir. Ancak, belleksizlik özelliği taşıması nedeniyle, serisel bağımlılık içeren verilerin zaman içindeki değişimlerinin açıklanmasında Poisson süreçleri yetersiz kalmaktadır [76]. Ayrıca, Poisson dağılımının ortalama ve varyans değerlerinin birbirine eşit olma özelliğinden dolayı, aşırı yayılım gösteren verilerin modellenmesinde uygun değildir [76]. Ardışık zaman aralıklarında toplanan verilerdeki bağımlılık yapısının gösterilmesinde en yaygın olarak Markov zincirleri kullanılmaktadır. Poisson Saklı Markov modeli (Poisson-SMM), serisel bağımlılık içeren ve aşırı yayılım gösteren frekans verilerinin açıklanmasında ve bu özellikteki verilerin zaman içerisindeki değişimlerinin öngörülmesinde etkili bir modeldir.

Poisson-SMM (PSMM)'de, gözlemler Poisson dağılım ailesinden türetilmektedir. Poisson dağılım ailesinin oran parametresi, homojen bir Markov zinciridir ve X saklı sürecini oluşturur. Herhangi bir t anında, saklı olan oran parametresinin tanımlayacağı Poisson dağılımından Y_t değeri gözlenir.

m -durumlu Poisson-SMM'de Poisson dağılım ailesine ilişkin oran parametreleri $\lambda^{(m)} = (\lambda_1, \lambda_2, \dots, \lambda_m)$ vektörü olarak tanımlansın. Bu durumda, X Markov zincirinin kesikli durum kümesi $S = \{S_i : S_i = \lambda_i, 1 \leq i \leq m\}$ olur. Poisson-SMM'de herhangi bir t anında sistemin içinde bulunduğu saklı durum bilindiğinde, gözlemlerin olasılık fonksiyonu aşağıdaki şekilde tanımlanır:

$$\begin{aligned} b_i(y) &= e^{-\lambda_i} \frac{\lambda_i^y}{y!}, \quad y = 0, 1, 2, \dots \\ &= 0, \quad \text{diğer } y \text{ değerleri için} \end{aligned}$$

Burada, $b_i(y) = P(Y_t = y \mid X_t = S_i)$ koşullu olasılıktır. Poisson-SMM'nin çalışma ilkesi Şekil 6.1.'de verilmiştir:



Şekil 6.1. m -Durumlu Poisson-SMM'nin Çalışma İlkesi

Şekil 6.1.' de herhangi bir $t \in T$ anında sistem S_i ($i = 1, 2, \dots, m$) durumunda iken, bu durumun tetiklediği gözlemlerin S_i parametrelili Poisson dağılımına uygun olarak meydana geldiği görülür. Diğer bir ifadeyle, $(Y_t | X_t = S_i) \sim P(\lambda_i)$ ' dir. Şekildeki renkli kutular sistemin $t \in T$ zamanında içinde bulunduğu durumu göstermektedir. Örneğin, $(t + 2)$ zamanında sistem S_{m-1} durumundadır. Bu durumda, Y_{t+2} gözlemi λ_{m-1} oran parametresine sahip olan Poisson dağılımına uygun olarak meydana gelecektir.

Poisson dağılım ailesi varsayımı altında SMM üzerinden istatistiksel çıkarımlar yapılabilmesi için model parametrelerinin tahmin edilmesi gerekmektedir. Bölüm 5.3.3.' te yer alan SMM'nin öğrenme probleminde Poisson dağılım parametreleri hariç diğer tüm model parametrelerinin EM algoritması kullanılarak nasıl tahmin edildiği gösterilmiştir. EM algoritmasında, $\lambda^{(m)}$ oran parametrelerinin en çok olabilirlik tahminleri, Eşitlik 5.19' da üçüncü terimin en büyüklenmesiyle elde edilebilir. Poisson dağılımı için üçüncü terim aşağıdaki gibidir:

$$\begin{aligned}
 \sum_{i=1}^m \sum_{t=1}^T \gamma_t(i) \ln b_i(Y_t) &= \sum_{i=1}^m \sum_{t=1}^T \gamma_t(i) \ln \left(e^{-\lambda_i} \frac{\lambda_i^{y_t}}{y_t!} \right) \\
 &= \sum_{i=1}^m \sum_{t=1}^T \gamma_t(i) (-\lambda_i + y_t \ln \lambda_i - \ln(y_t!)) \quad (6.1)
 \end{aligned}$$

Eşitlik 6.1' i en büyükleyen oran parametreleri ise aşağıdaki eşitliğin çözümü olarak

bulunur:

$$\frac{\partial}{\partial \lambda^{(m)}} \left(\sum_{i=1}^m \sum_{t=1}^T \gamma_t(i) (-\lambda_i + y_t \ln \lambda_i - \ln(y_t!)) \right) = 0 \quad (6.2)$$

Eşitlik 6.2' de herhangi bir λ_i ($i = 1, 2, \dots, m$) için türev alındığında bu parametrenin en çok olabilirlik tahmin edicisi aşağıdaki gibi elde edilir.

$$\begin{aligned} \frac{\partial}{\partial \lambda_i} \left(\sum_{i=1}^m \sum_{t=1}^T \gamma_t(i) (-\lambda_i + y_t \ln \lambda_i - \ln(y_t!)) \right) &= 0 \\ \sum_{t=1}^T \gamma_t(i) \left(-1 + \frac{y_t}{\lambda_i} \right) &= 0 \\ \frac{1}{\lambda_i} \sum_{t=1}^T \gamma_t(i) y_t - \sum_{t=1}^T \gamma_t(i) &= 0 \\ \hat{\lambda}_i &= \frac{\sum_{t=1}^T \gamma_t(i) y_t}{\sum_{t=1}^T \gamma_t(i)} \end{aligned} \quad (6.3)$$

Eşitlik 6.3' te elde edilen sonuca göre, Poisson-SMM' nin oran parametreleri $\gamma_t(\cdot)$, $t \in \mathbf{T}$ değişkenleri kullanılarak tahmin edilmektedir. Böylece, Poisson-SMM' de tüm model parametreleri analitik olarak elde edilebilmektedir.

Modelde π başlangıç durum vektörü, P geçiş olasılıkları matrisinin limit hali olarak düşünülebilir. Bu durumda, başlangıç durum vektörü durağan dağılımı temsil eder ve $\pi = \pi P$ eşitliğinden yararlanılarak hesaplanır. Bu bakış açısı altında, m -durumlu Poisson SMM' de tahmin edilecek toplam parametre sayısı m^2 olmaktadır.

SMM' de verileri en iyi şekilde açıklayacak saklı durum sayısına karar verilmesi gerekmektedir. Çünkü, sistemin durum sayısına bağlı olarak yapılacak istatistiksel çıkarsamalar değişecektir. $\mathcal{L}(\Theta; \mathbf{Y}^{(T)})$ modelin olabilirlik değeri dikkate alındığında, artan m değeri ile birlikte modelin veriye olan uyumuda artmaktadır. Artan m değeri sonucunda, bu uyum artışıyla birlikte model parametrelerinin sayısında da karesel bir artış oluşmaktadır. Modelin veriye olan uyum derecesi ile model parametrelerinin sayısı dengelenmelidir. Bu nedenle, model seçimi için bir kritere ihtiyaç vardır.

Klasik yaklaşım altında en çok kullanılan model seçim kriterleri, Akaike Bilgi Kriteri (AIC) ve Bayes Bilgi Kriteri (BIC)' dir. Bu bilgi kriterlerinden herhangi birinden yararlanılarak,

en iyi m değerine karar verilir. En küçük bilgi kriteri değerine sahip olan model en iyi modeldir. Bilgi kriteri değerleri aşağıda verilmiştir [76]:

$$AIC = -2\log \mathcal{L}(\Theta; \mathbf{Y}^{(T)}) + 2p$$

$$BIC = -2\log \mathcal{L}(\Theta; \mathbf{Y}^{(T)}) + p \log T$$

Burada, p , model parametrelerinin ve T , verilerin sayısıdır. AIC' de ilk terim modelin uyum ölçüsünü göstermektedir ve artan m değeriyle birlikte azalmaktadır; ikinci terim ise, ceza terimidir ve artan m değeriyle artmaktadır. $T > e^2$ için, BIC' de ceza teriminin kriter üzerindeki ağırlığı AIC' dekine göre daha fazladır.

Modeldeki saklı durum sayısına karar verilmesi amacıyla olabilirlik oran testi de gerçekleştirilebilir [84]. SMM' de olabilirlik oran testi, $j = 1, 2, \dots, (m - 1)$ durum sayıları için yapılır. Teste ilişkin yokluk ve seçenek hipotezleri sırasıyla aşağıda verilmiştir:

H_0 : Saklı durumların sayısı j ' dir.

H_1 : Saklı durumların sayısı $(j + 1)$ ' dir.

Test istatistiğinin hesaplanması amacıyla hipotezlerin temsil ettiği modellerin olabilirlik değerleri bulunur. H_0 doğru olduğunda log-olabilirlik değeri, $\log \mathcal{L}_0(\Theta; \mathbf{Y}^{(T)})$ ' dir. H_1 doğru olduğunda log-olabilirlik değeri ise, $\log \mathcal{L}_1(\Theta; \mathbf{Y}^{(T)})$ ' dir. Olabilirlik oranı testine ilişkin test istatistiği aşağıda verilmiştir:

$$\chi_{\text{Test}}^2 = -2 \left[\log \mathcal{L}_0(\Theta; \mathbf{Y}^{(T)}) - \log \mathcal{L}_1(\Theta; \mathbf{Y}^{(T)}) \right] \sim \chi_{sd=sd_1-sd_0}^2$$

Burada, sd , serbestlik derecesidir. sd_0 değeri, j -durumlu SMM' deki parametre sayısıdır. sd_1 değeri ise, $(j + 1)$ -durumlu SMM' deki parametre sayısıdır. $\chi_{\text{Test}}^2 \leq \chi_{\frac{\alpha}{2}, sd}^2$ ya da $\chi_{\text{Test}}^2 \geq \chi_{(1-\frac{\alpha}{2}), sd}^2$ ise, H_0 hipotezi reddedilir ve %100(1 - α) güvenirlikle $(j + 1)$ -durumlu SMM, j -durumlu SMM' ye kıyasla, veriye daha uygundur.

Poisson-SMM' de gelecek bir zaman aralığında oluşacak olan gözlemlere ve sistemin içinde bulunacağı saklı durumlara ilişkin olasılık dağılımları elde edilebilir. Bunun için, Eşitlik 5.8' de verilen $\mathcal{L}(\Theta; \mathbf{Y}^{(T)})$ olabilirlik fonksiyonu matris formunda aşağıdaki gibi yeniden yazılır:

$$L(\Theta; \mathbf{Y}^{(T)}) = \pi PA(Y_1)PA(Y_2)PA(Y_3) \dots PA(Y_T)\mathbf{1}^\top$$

Burada, $\mathbf{1} = [1]_{1 \times m}$, "1" değerlerinden oluşan bir satır vektörüdür. $\mathbf{A}(y)$ köşegen matristir ve aşağıdaki gibi tanımlanır:

$$\mathbf{A}(y) = \begin{bmatrix} A_1(y) & & 0 \\ & \ddots & \\ 0 & & A_m(y) \end{bmatrix}$$

$\forall 1 \leq i \leq m$ için, $A_i(y) = P(Y_t = y | X_t = S_i)$ olasılık değeridir. Olabilirlik fonksiyonu, değerlendirme probleminin çözüm yöntemlerinden biri olan ileri-yön algoritmasında tanımlanan ileri-yön değişkenleri kullanılarak $L(\Theta; \mathbf{Y}^{(T)}) = \alpha_T \mathbf{1}^\top$ şeklinde ifade edilebilir. $\alpha_T, \alpha_t(i)$ ($1 \leq i \leq m$) ileri-yön değişkenlerine ilişkin vektördür.

$\mathbf{Y}^{(T)}$ gözlem dizisi bilindiğinde, k (pozitif tamsayı) dönem sonra meydana gelecek gözlemlerin tahmin dağılımı olabilirliklerin oranı cinsinden aşağıdaki gibi hesaplanabilir [76]:

$$\begin{aligned} P(Y_{T+k} = y | \mathbf{Y}^{(T)}, \Theta) &= \frac{P(\mathbf{Y}^{(T)}, Y_{T+k} = y | \Theta)}{P(\mathbf{Y}^{(T)} | \Theta)} \\ &= \frac{\pi \mathbf{P}\mathbf{A}(Y_1) \mathbf{P}\mathbf{A}(Y_2) \mathbf{P}\mathbf{A}(Y_3) \dots \mathbf{P}\mathbf{A}(Y_T) \mathbf{P}^{(k)} \mathbf{A}(y) \mathbf{1}^\top}{\pi \mathbf{P}\mathbf{A}(Y_1) \mathbf{P}\mathbf{A}(Y_2) \mathbf{P}\mathbf{A}(Y_3) \dots \mathbf{P}\mathbf{A}(Y_T) \mathbf{1}^\top} \\ &= \frac{\alpha_T \mathbf{P}^{(k)} \mathbf{A}(y) \mathbf{1}^\top}{\alpha_T \mathbf{1}^\top} \\ &= \phi_T \mathbf{P}^{(k)} \mathbf{A}(y) \mathbf{1}^\top \end{aligned}$$

k -ıncı dönemde meydana gelecek gözlemlerin tahmin dağılımı, $\mathbf{A}(y)$ matrisinin köşegen elemanlarının karma dağılımı şeklinde ifade edilebilir:

$$P(Y_{T+k} = y | \mathbf{Y}^{(T)}, \Theta) = \sum_{i=1}^m \varsigma_i(k) A_i(y)$$

Burada, $\varsigma_i(k), \phi_T \mathbf{P}^{(k)}$ vektörünün i -nci elemanıdır. \mathbf{X} saklı sürecinin denge dağılımı δ olmak üzere, k değerinin çok büyük olması durumunda k -ıncı dönemde meydana gelecek olan gözlemlerin olasılık dağılımı,

$$\lim_{k \rightarrow +\infty} P(Y_{T+k} = y | \mathbf{Y}^{(T)}, \Theta) = \lim_{k \rightarrow +\infty} \phi_T \mathbf{P}^{(k)} \mathbf{A}(y) \mathbf{1}^\top = \delta \mathbf{A}(y) \mathbf{1}^\top$$

olarak elde edilir [76].

$\mathbf{Y}^{(T)}$ gözlem dizisi bilindiğinde, k -ıncı dönemde sistemin içinde bulunacağı saklı durumların ($S_i, 1 \leq i \leq m$) tahmin dağılımı ise, aşağıdaki gibi elde edilir [76]:

$$\begin{aligned} P(X_{T+k} = S_i | \mathbf{Y}^{(T)}, \Theta) &= \frac{P(\mathbf{Y}^{(T)}, X_{T+k} = S_i | \Theta)}{P(\mathbf{Y}^{(T)} | \Theta)} \\ &= \frac{\alpha_T \mathbf{P}^{(k)}[, i]}{\alpha_T \mathbf{1}^\top} \\ &= \phi_T \mathbf{P}^{(k)}[, i] \end{aligned}$$

Burada, $\mathbf{P}^{(k)}[, i]$ vektörü, $\mathbf{P}^{(k)}$ matrisinin i -nci kolunudur. k değerinin büyük olması durumunda, \mathbf{X} saklı süreci dengeye ulaşmaktadır ve sistem, δ denge dağılımına göre davranmaktadır.



7. BETA SAKLI MARKOV MODELİ

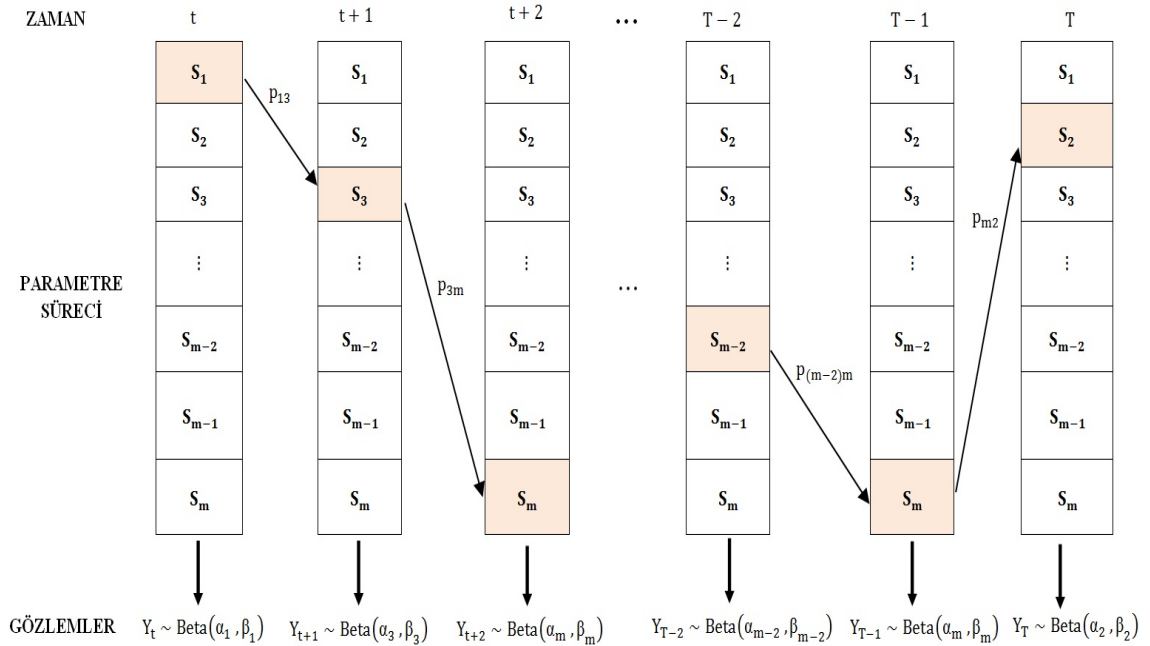
Beta saklı Markov modelinde (Beta-SMM), gözlemler Beta dağılım ailesinden türetilmektedir. Beta dağılım ailesinin şekil parametreleri, homojen bir Markov zinciridir ve X saklı sürecini oluşturur. Buna göre, herhangi bir t anında, saklı olan şekil parametrelerinin tanımlayacağı Beta dağılımından Y_t değeri gözlenir.

Beta dağılımı, iki tane şekil parametresine sahiptir. Bu durumda, m -durumlu Beta-SMM' de beta dağılım ailesine ilişkin şekil parametreleri $\alpha^{(m)} = (\alpha_1, \alpha_2, \dots, \alpha_m)$ ve $\beta^{(m)} = (\beta_1, \beta_2, \dots, \beta_m)$ vektörleri ile tanımlansın. X Markov zincirinin kesikli yapıdaki durum kümesi ise, $\mathcal{S} = \{S_i : S_i = (\alpha_i, \beta_i), 1 \leq i \leq m\}$ olur. Beta-SMM' de herhangi bir t anında sistemin içinde bulunduğu saklı durum bilindiğinde, gözlemlerin olasılık yoğunluk fonksiyonu aşağıdaki şekilde tanımlanır:

$$b_i(y) = \frac{y^{\alpha_i-1}(1-y)^{\beta_i-1}}{B(\alpha_i, \beta_i)}, \quad 0 < y < 1$$

$$= 0, \quad \text{diğer } y \text{ değerleri için}$$

Burada, $b_i(y) = f(y | X_t = S_i)$ olasılık yoğunluk fonksiyonudur. $B(\cdot, \cdot)$ ise, beta fonksiyonudur ve $B(\alpha_i, \beta_i) = \frac{\Gamma(\alpha_i)\Gamma(\beta_i)}{\Gamma(\alpha_i + \beta_i)}$ dir. Beta-SMM' nin çalışma ilkesi şematik olarak aşağıda verilmiştir:



Şekil 7.1. m -Durumlu Beta-SMM' nin Çalışma İlkesi

Şekil 7.1.' de herhangi bir $t \in \mathbf{T}$ anında sistem S_i ($i = 1, 2, \dots, m$) durumunda iken, bu durumun tetiklediği gözlemlerin S_i parametrelili Beta dağılımına uygun olarak meydana geldiği görülür. Diğer bir ifadeyle, $(Y_t | X_t = S_i) \sim \text{Beta}(\alpha_i, \beta_i)$ ' dir. Şekildeki renkli kutular sistemin $t \in \mathbf{T}$ zamanında içinde bulunduğu durumu göstermektedir.

Beta-SMM üzerinden istatistiksel çıkarsamalar yapılabilmesi için, model parametrelerinin tahmin edilmesi gerekmektedir. Dağılım parametreleri hariç diğer tüm model parametrelerinin EM algoritması kullanılarak nasıl tahmin edildiği Bölüm 5.3.3.' te yer alan öğrenme probleminde gösterilmiştir. EM algoritmasında, $\alpha^{(m)}$ ve $\beta^{(m)}$ şekil parametrelerinin en çok olabilirlik tahminleri, Eşitlik 5.19' da üçüncü terimin en büyüklenmesiyle elde edilir. Beta dağılımı için üçüncü terim aşağıdaki gibidir:

$$\begin{aligned} \sum_{i=1}^m \sum_{t=1}^T \gamma_t(i) \ln b_i(y_t) &= \sum_{i=1}^m \sum_{t=1}^T \gamma_t(i) \ln \left(\frac{y_t^{\alpha_i-1} (1-y_t)^{\beta_i-1}}{B(\alpha_i, \beta_i)} \right) \\ &= \sum_{i=1}^m \sum_{t=1}^T \gamma_t(i) ((\alpha_i - 1) \ln y_t + (\beta_i - 1) \ln(1 - y_t) - \ln B(\alpha_i, \beta_i)) \end{aligned} \quad (7.1)$$

Eşitlik 7.1' i en büyükleyen şekil parametreleri ise aşağıdaki eşitliğin çözümü olarak bulunur:

$$\frac{\partial^2}{\partial \alpha^{(m)} \partial \beta^{(m)}} \left(\sum_{i=1}^m \sum_{t=1}^T \gamma_t(i) ((\alpha_i - 1) \ln y_t + (\beta_i - 1) \ln(1 - y_t) - \ln B(\alpha_i, \beta_i)) \right) = 0 \quad (7.2)$$

Herhangi bir (α_i, β_i) ikilisi için Eşitlik 7.2 aşağıdaki gibi yazılır:

$$\frac{\partial^2}{\partial \alpha_i \partial \beta_i} \left(\sum_{t=1}^T \gamma_t(i) ((\alpha_i - 1) \ln y_t + (\beta_i - 1) \ln(1 - y_t) - \ln B(\alpha_i, \beta_i)) \right) = 0 \quad (7.3)$$

Eşitlik 7.3' te α_i ve β_i parametrelerine göre sırasıyla türev alındığında izleyen en çok olabilirlik denklemleri elde edilir:

$$\begin{aligned} \left[\sum_{t=1}^T \gamma_t(i) \ln y_t \right] - T \frac{\partial}{\partial \alpha_i} \ln B(\alpha_i, \beta_i) &= 0 \\ \left[\sum_{t=1}^T \gamma_t(i) \ln(1 - y_t) \right] - T \frac{\partial}{\partial \beta_i} \ln B(\alpha_i, \beta_i) &= 0 \end{aligned} \quad (7.4)$$

Beta fonksiyonunun α_i ve β_i 'ye göre türevleri ise, aşağıda verilmiştir:

$$\begin{aligned}\frac{\partial}{\partial \alpha_i} \ln B(\alpha_i, \beta_i) &= -\frac{\partial}{\partial \alpha_i} \ln \Gamma(\alpha_i + \beta_i) + \frac{\partial}{\partial \alpha_i} \ln \Gamma(\alpha_i) + \frac{\partial}{\partial \alpha_i} \ln \Gamma(\beta_i) \\ &= -\psi(\alpha_i + \beta_i) + \psi(\alpha_i) + 0 \\ &= \psi(\alpha_i) - \psi(\alpha_i + \beta_i)\end{aligned}$$

$$\begin{aligned}\frac{\partial}{\partial \beta_i} \ln B(\alpha_i, \beta_i) &= -\frac{\partial}{\partial \beta_i} \ln \Gamma(\alpha_i + \beta_i) + \frac{\partial}{\partial \beta_i} \ln \Gamma(\alpha_i) + \frac{\partial}{\partial \beta_i} \ln \Gamma(\beta_i) \\ &= -\psi(\alpha_i + \beta_i) + 0 + \psi(\beta_i) \\ &= \psi(\beta_i) - \psi(\alpha_i + \beta_i)\end{aligned}$$

Burada, $\psi(\cdot)$, digamma fonksiyonudur ve gamma fonksiyonunun logaritmik türevi olarak tanımlanır. Diğer bir ifadeyle, $z > 0$ için $\psi(z) = \frac{d}{dz} \ln \Gamma(z) = \frac{\Gamma'(z)}{\Gamma(z)}$, dir. Beta fonksiyonunun türevleri Eşitlik 7.4' te yerine konulduğunda, Beta-SMM için şekil parametrelerinin en çok olabilirlik denklemleri aşağıdaki gibi elde edilmektedir [85], [86], [87].

$$\begin{aligned}\psi(\hat{\alpha}_i) - \psi(\hat{\alpha}_i + \hat{\beta}_i) &= T^{-1} \sum_{t=1}^T \gamma_t(i) \ln y_t \\ \psi(\hat{\beta}_i) - \psi(\hat{\alpha}_i + \hat{\beta}_i) &= T^{-1} \sum_{t=1}^T \gamma_t(i) \ln(1 - y_t)\end{aligned}\tag{7.5}$$

$\hat{\alpha}_i$ ve $\hat{\beta}_i$ en çok olabilirlik tahminleri, Eşitlik 7.5' te verilen denklem sisteminin çözümüdür. Bu denklem sisteminin digamma fonksiyonlarını içermesi nedeniyle, şekil parametrelerinin tahminleri kapalı formda elde edilememektedir. EM algoritmasının M-adımında, uygun sayısal yöntemlerden yararlanılarak Beta-SMM için dağılım parametreleri tahmin edilmelidir. Newton-Raphson yöntemi kullanılabilir. Fakat, Newton-Raphson yöntemi başlangıç değerlerine karşı oldukça duyarlıdır ve yakınsamanın sağlanmasını garanti etmez [86].

Beta-SMM' de, dağılım parametrelerinin tahmin sürecinde standart yaklaşımın uygulanmasının zor olduğu açıkça görülmektedir. Kapalı formda elde edilemeyen parametrelerin en çok olabilirlik tahminlerinin, sayısal yöntemlerle elde edilmesi, alternatif bir yol olarak önerilebilir. Ancak bu yöntemler, saklı Markov modellerinde önemli bir yakınsama sorununa ya da işleyişin daha da karmaşık olmasına yol açabilir. Burada belirtilen gerekçeler nedeniyle Beta-SMM ile ilgili uygulamaların az olması şaşırtıcı bir durum değildir. Sayısal yöntemler kullanıldığında, başlangıç değerlerinin sonuçlar üzerine

olumsuz etkileri de uygulamalarda karşılaşılabilecek bir diğer sorundur. Bu problemler, parametre tahminlerinin kesinliğini ve doğruluğunu azaltıcı unsurlardır. Model üzerinden yapılacak istatistiksel çıkarımların güvenilirliği de olumsuz yönde etkileneceği açıktır. Sonuç olarak, Beta dağılım ailesi için kurulan saklı Markov modellerinde, en çok olabilirlik tahminlerinin elde edilmesindeki çözümsüzlüğün sayısal yöntemlerle aşılması uygulanabilir; ancak tavsiye edilen bir yol değildir. SMM'nin çok boyutlu ve karmaşık yapıda olmasından dolayı, hatalı tahmin değerlerine ulaşılması mümkündür.

Tez çalışmasında yukarıda belirtilen sorunların aşılmasında SMM' nin Bayesci çözümlenmesi ele alınmıştır. Son yıllarda özellikle bilgisayar teknolojisindeki gelişime paralel olarak birçok alanda karmaşık, çok boyutlu modellerin çözümlenmesinde Bayesci yaklaşım başarıyla kullanılmaktadır. Stokastik benzetim yöntemlerinin etkin olarak uygulanmasıyla; başlangıç değer etkilerinden ve yakınsaklık problemlerinden uzak; daha etkin sonuçlara ulaşılması da mümkün görünmektedir. Bu nedenle, klasik istatistiksel çıkarımlar altında Beta-SMM için yaşanan bu sorunların yok edilmesi için önerilen Bayesci Beta-SMM izleyen bölümde tanıtılacaktır.

7.1. Bayesci Beta Saklı Markov Modeli

Bu bölümde, Beta-SMM' nin Bayesci çözümlenmesi ele alınmıştır. İlk olarak, Bayesci yaklaşım altında model parametrelerinin nasıl elde edileceği gösterilmiştir. Ardından, gelecek dönem gözlem değerlerinin tahmini anlatılmıştır. Son olarak, gözlemleri en iyi açıklayan modele karar verilmesi amacıyla, Bayesci yaklaşım altında Beta-SMM için marjinal olabilirliğin nasıl hesaplanacağı açıklanmıştır.

Bu bölümde yer alan Beta-SMM' nin Bayesci çözümlenmesi, literatürde beta dağılım ailesi için yapılmış olan ilk çalışmadır. Çalışmada kullanılan Bayesci analizin teknik detayları için, Pievatolo, Ruggeri ve Soyer [13]' in 2012 yılında yaptıkları çalışma temel alınmıştır. Pievatolo vd. [13], üstel dağılım varsayımı altında SMM' nin Bayesci çözümlenmesini yapmışlardır. Ayrıca, model seçimi için Bayes faktörü hesaplanmasında, Chib [88]' in 1995 yılında önerdiği yol takip edilmiştir. Bu çalışmada önerilen strateji, SMM' de optimal saklı durum sayısının belirlenmesinde kullanılabilir. Tez çalışmasında Beta-SMM' deki X Markov zincirinin durum sayısının belirlenmesinde Chib [88]' in çalışması baz alınmıştır. Bunun yanısıra, kurulan modelde beta dağılım ailesi parametrelerinin Bayesci tahmin sürecinde Bayyari [89]' nin 1985 yılında önerdiği yol benimsenmiştir.

Yazar çalışmasında, Beta dağılımının şekil parametreleri için bir bileşik eşlenik önsel dağılım önermiştir. İşlemlerin daha yalın olabilmesine olanak sağlayacağı düşünülen bu bileşik eşlenik önsel dağılım, Beta-SMM' de kullanılmıştır.

Tez çalışmasının izleyen kesimlerinde Beta-SMM' de tahmin, öngörü ve saklı durum sayılarının belirlenmesine ilişkin oluşturulan teorik yapı verilecektir.

7.1.1. Beta Saklı Markov Modelinde Bayesci Parametre Tahminlerinin Elde Edilmesi

SMM' de temel amaç, $Y^{(T)}$, T dönemlik gözlem değerleri verildiğinde, Y_{T+k} , $k = 1, 2, \dots$ gelecek dönem gözlem değerinin kestirilmesidir. Örneğin, $(T + 1)$ zamanında saklı süreç i durumunda olduğu bilindiğinde, Y_{T+1} gözlem değeri $Beta(\alpha_i, \beta_i)$ dağılımından türetilir. Burada, $X^{(T)}$ saklı süreç bilgisi bilinmemektedir. Bu nedenle, parametre tahmin süreci $X^{(T)}$ saklı durum dizisinin tahminini de gerektirmektedir. Klasik yaklaşımdan farklı olarak, Beta-SMM' de bilinmeyen tüm değişkenler $\Theta = (P, \alpha^{(m)}, \beta^{(m)}, X^{(T)})$ parametre vektörü olarak tanımlanmıştır.

Bayesci yaklaşım altında m -durumlu Beta-SMM' nin parametrelerinin tahmini için öncelikle, Θ ' nin bileşik sonsal dağılımı elde edilmelidir. Ardından, elde edilen bileşik sonsal dağılımdan örnekler çekilerek model parametrelerinin tahminlerine ulaşılr. Bileşik sonsal dağılıma ulaşmak için, durum geçiş olasılıkları dağılımı ve beta dağılımının şekil parametreleri için önsel dağılımlar belirlenir. Burada, eşlenik önsel bir dağılımın seçilmesi sayısal integrasyon işlemlerine gerek duyulmadan sonsal dağılımın kolayca elde edilmesini sağlar. Ortaya çıkan sonsal dağılım, önsel dağılımla aynı parametrik dağılım ailesinin bir üyesi olur.

SMM' de X Markov zincirinin durum geçiş olasılıkları matrisi, birbirinden bağımsız olan $P_i = [p_{ij}]_{1 \times m}$, $(i = 1, 2, \dots, m)$ satır vektörlerinden oluştuğunda, P geçiş olasılıkları matrisi aşağıdaki gibi ifade edilir:

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1m} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2m} \\ \vdots & \vdots & \vdots & & \vdots \\ p_{m1} & p_{m2} & p_{m3} & \cdots & p_{mm} \end{bmatrix} = \begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_m \end{bmatrix}$$

P 'nin satır toplamları bire eşit olması nedeniyle, SMM' de P_i 'ler katlı-terimli dağılıma sahiptir. Bayeci istatistikte katlı-terimli dağılımın eşlenik önseli, Dirichlet dağılımıdır. Bu nedenle, satır vektörlerinin önsel dağılımı Dirichlet dağılımı olarak belirlenir. Buna göre, $P_i \sim \mathcal{D}(\nu_{i1}, \dots, \nu_{im})$ 'dir ve $\forall i \in \{1, 2, \dots, m\}$ için $\nu_i = (\nu_{i1}, \dots, \nu_{im})$ hiperparametreleri önceden belirlenmiş değerlerdir. P_i 'nin önsel dağılımı aşağıda verilmiştir:

$$f(P_i) = \frac{\Gamma\left(\sum_{j=1}^m \nu_{ij}\right)}{\prod_{j=1}^m \Gamma(\nu_{ij})} \prod_{j=1}^m p_{ij}^{\nu_{ij}-1} \quad (7.6)$$

$$f(P_i) \propto \prod_{j=1}^m p_{ij}^{\nu_{ij}-1} \quad (7.7)$$

Eşitlik 7.6' da p_{ij} 'den bağımsız olan değerler normalleştirme katsayısı olarak kabul edilir ve ihmal edilebilirler. Bu durumda, Eşitlik 7.7 yaklaşık olarak elde edilir. P_i 'lerin birbirinden bağımsız olması sonucu, P 'nin bileşik önsel yoğunluğu aşağıdaki gibi elde edilir:

$$f(P) = \prod_{i=1}^m f(P_i)$$

Geçiş olasılıkları matrisi elemanları, Y sürecinden bağımsızdır ve sadece X saklı sürecine bağlıdır.

Bayarri [89], genel bir bilgi olarak, beta dağılımının şekil parametreleri için bileşik eşlenik önsel bir dağılım önermiştir. Herhangi bir $S_i = (\alpha_i, \beta_i)$, ($i = 1, 2, \dots, m$) durumu için önerilen eşlenik önsel dağılım aşağıda verilmiştir:

$$f(\alpha_i, \beta_i) = k(d_i, a_i, b_i) \frac{e^{-a_i \alpha_i} e^{-b_i \beta_i}}{[B(\alpha_i, \beta_i)]^{d_i}}, \quad \alpha_i > 0 \text{ ve } \beta_i > 0 \quad (7.8)$$

$$\propto \frac{e^{-a_i \alpha_i} e^{-b_i \beta_i}}{[B(\alpha_i, \beta_i)]^{d_i}} \quad (7.9)$$

Eşitlik 7.8' de $k(d_i, a_i, b_i)$ fonksiyonu normalleştirme katsayısıdır ve ihmal edilebilir. Bu durumda, Eşitlik 7.9 üzerinden işlemler yapılabilir. $d_i > 0$, $a_i > 0$ ve $b_i > 0$ hiperparametreleri önceden belirlenmiş değerlerdir. d_i , a_i ve b_i parametreleriyle Eşitlik 7.8' de verilen bileşik olasılık yoğunluk fonksiyonuna sahip S_i saklı durumu, tez çalışmasında $S_i \sim EBeta(d_i, a_i, b_i)$ olarak gösterilmiştir. $EBeta$, eşlenik beta dağılımını ifade etmek-

tedir. $i \neq j$ ve $i, j \in \{1, 2, \dots, m\}$ olmak üzere, $\mathcal{S}_i \perp \mathcal{S}_j$ ' dir. Şekil parametrelerinin birbirinden bağımsız olması sonucu, $(\boldsymbol{\alpha}^{(m)}, \boldsymbol{\beta}^{(m)})$ ' in bileşik önsel yoğunluğu aşağıdaki gibi elde edilir:

$$f(\boldsymbol{\alpha}^{(m)}, \boldsymbol{\beta}^{(m)}) = \prod_{i=1}^m f(\alpha_i, \beta_i)$$

Tüm şekil parametreleri, geçiş olasılıkları matrisi elemanlarından bağımsızdırlar.

m -durumlu Beta-SMM' nin $\mathcal{L}(\boldsymbol{\Theta}; \mathbf{Y}^{(T)})$ olabilirlik fonksiyonu aşağıdaki gibi elde edilir:

$$\begin{aligned} \mathcal{L}(\boldsymbol{\Theta}; \mathbf{Y}^{(T)}) &= P(\mathbf{Y}^{(T)} | \boldsymbol{\Theta}) \\ &= \pi_{x_1} f(y_1 | \alpha_{x_1}, \beta_{x_1}) \left[\prod_{t=2}^T P(X_t | X_{t-1}) f(y_t | \alpha_{x_t}, \beta_{x_t}) \right] \\ &= \pi_{x_1} \left(\frac{Y_1^{\alpha_{x_1}-1} (1-Y_1)^{\beta_{x_1}-1}}{B(\alpha_{x_1}, \beta_{x_1})} \right) \left[\prod_{t=2}^T P(X_t | X_{t-1}) \left(\frac{Y_t^{\alpha_{x_t}-1} (1-Y_t)^{\beta_{x_t}-1}}{B(\alpha_{x_t}, \beta_{x_t})} \right) \right] \end{aligned}$$

$\mathbf{Y}^{(T)}$ gözlemleri verildiğinde, $\boldsymbol{\Theta}$ ' nın bileşik sonsal dağılımı Bayes teoreminin uygulanmasıyla aşağıdaki gibi elde edilir:

$$\begin{aligned} f(\boldsymbol{\Theta} | \mathbf{Y}^{(T)}) &\propto \mathcal{L}(\boldsymbol{\Theta}; \mathbf{Y}^{(T)}) f(\boldsymbol{\Theta}) \\ &\propto \pi_{x_1} \left(\frac{Y_1^{\alpha_{x_1}-1} (1-Y_1)^{\beta_{x_1}-1}}{B(\alpha_{x_1}, \beta_{x_1})} \right) \left[\prod_{t=2}^T P(X_t | X_{t-1}) \left(\frac{Y_t^{\alpha_{x_t}-1} (1-Y_t)^{\beta_{x_t}-1}}{B(\alpha_{x_t}, \beta_{x_t})} \right) \right] \\ &\quad \left[\prod_{i=1}^m f(\mathbf{P}_i) f(\alpha_i, \beta_i) \right] \\ &\propto \pi_{x_1} \left(\frac{Y_1^{\alpha_{x_1}-1} (1-Y_1)^{\beta_{x_1}-1}}{B(\alpha_{x_1}, \beta_{x_1})} \right) \left[\prod_{t=2}^T P(X_t | X_{t-1}) \left(\frac{Y_t^{\alpha_{x_t}-1} (1-Y_t)^{\beta_{x_t}-1}}{B(\alpha_{x_t}, \beta_{x_t})} \right) \right] \\ &\quad \left[\prod_{i=1}^m \left(\prod_{j=1}^m p_{ij}^{\nu_{ij}-1} \right) \left(\frac{e^{-a_i \alpha_i} e^{-b_i \beta_i}}{[B(\alpha_i, \beta_i)]^{d_i}} \right) \right] \end{aligned} \quad (7.10)$$

Eşitlik 7.10' da görüldüğü gibi $f(\boldsymbol{\Theta} | \mathbf{Y}^{(T)})$, çok boyutlu ve karmaşık bir fonksiyondur. Bu nedenle, $\boldsymbol{\Theta}$ ' nın bileşik sonsal dağılımının kapalı formunu elde etmek mümkün değildir ve $f(\boldsymbol{\Theta} | \mathbf{Y}^{(T)})$ dağılımından doğrudan örneklem çekilmesi imkansızdır. Bileşik sonsal dağılım, daha düşük boyutlu ve basit yapıda olan tam koşullu sonsal dağılımlara dönüştürülebilir. Eşlenik önsel dağılımların kullanılmasının bir sonucu olarak, Beta-SMM' de her bir parametrenin tam koşullu sonsal dağılımı kapalı formda elde edilebilmektedir. Bu durumda, tam koşullu sonsal dağılımlardan kolaylıkla örneklem çekilebilir. Tam koşullu dağılımlardan örneklem çekmeye dayalı olan Gibbs örnekleme

kullanılarak, $f(\Theta | Y^{(T)})$ bileşik sonsal dağılımından örnek değerler üretilebilir.

Modelde $P \perp\!\!\!\perp Y^{(T)}$ ve $P \perp\!\!\!\perp (\alpha^{(m)}, \beta^{(m)})$ bağımsızlıklarının mevcut olmasından dolayı, Θ 'nin bileşik sonsal dağılımı, model parametrelerin tam koşullu sonsal dağılımlarının çarpımı olarak aşağıdaki şekilde yazılabilir.

$$f(\Theta | Y^{(T)}) = f(X^{(T)} | \alpha^{(m)}, \beta^{(m)}, P, Y^{(T)}) \times f(\alpha^{(m)}, \beta^{(m)} | X^{(T)}, Y^{(T)}) \times f(P | X^{(T)}) \quad (7.11)$$

Eşitlik 7.11' de yer alan tam koşullu sonsal dağılımlar aşağıda elde edilmiştir.

P_i 'lerin birbirinden bağımsız olması nedeniyle, P 'nin tam koşullu sonsal yoğunluğu aşağıdaki şekilde hesaplanır:

$$f(P | X^{(T)}) = \prod_{i=1}^m f(P_i | X^{(T)})$$

Burada, P_i 'nin tam koşullu sonsal dağılımı aşağıdaki gibi elde edilir:

$$\begin{aligned} f(P_i | X^{(T)}) &\propto \underbrace{p(X_1 = x_1, X_2 = x_2, \dots, X_T = x_T | P_i)}_{\text{Olabilirlik Fonksiyonu}} \underbrace{f(P_i)}_{\text{Önsel Dağılım}} \\ &\propto \left[\begin{array}{ccc} \sum_{t=1}^{T-1} \mathbb{1}(X_t = i, X_{t+1} = 1) & \sum_{t=1}^{T-1} \mathbb{1}(X_t = i, X_{t+1} = m) & \\ p_{i1} & \dots p_{im} & \end{array} \right] \left[\prod_{j=1}^m p_{ij}^{\nu_{ij}-1} \right] \\ &\propto \left[\prod_{j=1}^m \sum_{t=1}^{T-1} \mathbb{1}(X_t = i, X_{t+1} = j) \right] \left[\prod_{j=1}^m p_{ij}^{\nu_{ij}-1} \right] \\ &\propto \prod_{j=1}^m p_{ij}^{\left[\left(\nu_{ij} + \sum_{t=1}^{T-1} \mathbb{1}(X_t = i, X_{t+1} = j) \right) - 1 \right]} \end{aligned} \quad (7.12)$$

Burada, $\mathbb{1}(\cdot, \cdot)$, gösterge fonksiyonudur ve aşağıdaki gibi tanımlanır:

$$\mathbb{1}(X_t = i, X_{t+1} = j) = \begin{cases} 1, & X_t = i \text{ ve } X_{t+1} = j \text{ ise,} \\ 0, & \text{diğer durumlarda.} \end{cases}$$

Eşitlik 7.12' de elde edilen $f(P_i | X^{(T)})$ fonksiyonunun biçimi Dirichlet dağılımına aittir.

Bu nedenle, P_i ' nin tam koşullu sonsal dağılımı,

$$(P_i | \mathbf{X}^{(T)}) \sim \mathcal{D} \left(\nu_{ij} + \sum_{t=1}^{T-1} \mathbf{1}(X_t = i, X_{t+1} = j) ; j = 1, 2, \dots, m \right) \quad (7.13)$$

olarak bulunur. Bu durum, Dirichlet dağılımının eşlenik önsel bir dağılım olmasının güzel bir sonucudur.

Şekil parametrelerinin birbirinden bağımsız olması nedeniyle, $(\alpha^{(m)}, \beta^{(m)})$ ' nin tam koşullu bileşik sonsal yoğunluğu aşağıdaki şekilde hesaplanır.

$$f(\alpha^{(m)}, \beta^{(m)} | \mathbf{X}^{(T)}, \mathbf{Y}^{(T)}) = \prod_{i=1}^m f(\alpha_i, \beta_i | \mathbf{X}^{(T)}, \mathbf{Y}^{(T)})$$

Burada, S_i ' nin tam koşullu sonsal dağılımı aşağıdaki gibi bulunur.

$$\begin{aligned} f(\alpha_i, \beta_i | \mathbf{X}^{(T)}, \mathbf{Y}^{(T)}) &\propto \underbrace{\left[\prod_{t=1}^T \mathbb{H}(X_t = i) f(y_t | X_t = x_t, \alpha_i, \beta_i) \right]}_{\text{Olabilirlik Fonksiyonu}} \underbrace{f(\alpha_i, \beta_i)}_{\text{Önsel Dağılım}} \\ &\propto \left[\prod_{t=1}^T \mathbb{H}(X_t = i) \left(\frac{y_t^{\alpha_i - 1} (1 - y_t)^{\beta_i - 1}}{\text{B}(\alpha_i, \beta_i)} \right) \right] \left[\frac{e^{-a_i \alpha_i} e^{-b_i \beta_i}}{[\text{B}(\alpha_i, \beta_i)]^{d_i}} \right] \\ &\propto \frac{e^{\left[\alpha_i \ln \left(\prod_{t=1}^T \mathbf{1}(X_t = i) y_t \right) + \beta_i \ln \left(\prod_{t=1}^T \mathbf{1}(X_t = i) (1 - y_t) \right) \right]} e^{-a_i \alpha_i} e^{-b_i \beta_i}}{\left[\text{B}(\alpha_i, \beta_i) \right]^{\left(a_i + \sum_{t=1}^T \mathbf{1}(X_t = i) \right)}} \\ &\propto \frac{e^{-\left(a_i - \sum_{t=1}^T \mathbf{1}(X_t = i) \ln(y_t) \right) \alpha_i} e^{-\left(b_i - \sum_{t=1}^T \mathbf{1}(X_t = i) \ln(1 - y_t) \right) \beta_i}}{\left[\text{B}(\alpha_i, \beta_i) \right]^{\left(a_i + \sum_{t=1}^T \mathbf{1}(X_t = i) \right)}} \end{aligned} \quad (7.14)$$

Burada, $\mathbb{H}(\cdot, \cdot)$ fonksiyonu ve $\mathbf{1}(\cdot)$ gösterge fonksiyonu aşağıdaki gibi tanımlanırlar:

$$\mathbb{H}(X_t = i) = \begin{cases} f(y_t | X_t = S_i), & X_t = i \text{ ise,} \\ 1, & X_t \neq i \text{ ise,} \end{cases}$$

$$\mathbf{1}(X_t = i) = \begin{cases} 1, & X_t = i \text{ ise,} \\ 0, & X_t \neq i \text{ ise,} \end{cases}$$

Eşitlik 7.14' te elde edilen $f(\alpha_i, \beta_i | \mathbf{X}^{(T)}, \mathbf{Y}^{(T)})$ fonksiyonunun biçimi Eşitlik 7.8' de verilen eşlenik beta dağılım ailesine aittir. Bu durumda, S_i ' nin tam koşullu sonsal dağılımı

$$(\alpha_i, \beta_i | \mathbf{X}^{(T)}, \mathbf{Y}^{(T)}) \sim EBeta(d_i^*, a_i^*, b_i^*) \quad (7.15)$$

olmaktadır. Eşitlik 7.15' teki sonsal dağılım parametreleri aşağıdaki gibidir:

$$\begin{aligned} d_i^* &= d_i + \sum_{t=1}^T \mathbb{1}(X_t = i) \\ a_i^* &= a_i - \sum_{t=1}^T \mathbb{1}(X_t = i) \ln(y_t) \\ b_i^* &= b_i - \sum_{t=1}^T \mathbb{1}(X_t = i) \ln(1 - y_t) \end{aligned}$$

$\mathbf{X}^{(-t)} = \{X_s : s \in \mathbf{T} \text{ ve } s \neq t\}$ olarak tanımlansın. Saklı durumlar dizisinin bir Markov zinciri olması nedeniyle, $\mathbf{X}^{(T)}$ ' nin tam koşullu sonsal olasılığı aşağıdaki gibi hesaplanır:

$$p(\mathbf{X}^{(T)} | \boldsymbol{\alpha}^{(m)}, \boldsymbol{\beta}^{(m)}, \mathbf{Y}^{(T)}, \mathbf{P}) = \prod_{t=1}^T p(X_t | X^{(-t)}, \alpha_{X_t}, \beta_{X_t}, \mathbf{Y}^{(T)}, \mathbf{P})$$

Burada, $\mathbf{X}^{(T)}$ saklı durumlar dizisinin tam koşullu sonsal olasılığının hesaplanması amacıyla, $\forall X_t, t \in \mathbf{T}$ için tam koşullu sonsal olasılık bulunmalıdır. X_1, X_T ve $X_t, t = 2, \dots, (T-1)$ saklı durumları için tam koşullu sonsal olasılıklar sırasıyla aşağıda verildiği gibi hesaplanır [13]:

$$\begin{aligned} p(X_1 | X^{(-1)}, \alpha_{X_1}, \beta_{X_1}, \mathbf{Y}^{(T)}, \mathbf{P}) &\propto f(y_1 | X_1, \alpha_{X_1}, \beta_{X_1}) P(X_2 | X_1) \\ &\propto \left(\frac{y_1^{\alpha_{X_1}-1} (1-y_1)^{\beta_{X_1}-1}}{B(\alpha_{X_1}, \beta_{X_1})} \right) P(X_2 | X_1) \end{aligned} \quad (7.16)$$

$$\begin{aligned} p(X_T | X^{(-T)}, \alpha_{X_T}, \beta_{X_T}, \mathbf{Y}^{(T)}, \mathbf{P}) &\propto P(X_T | X_{T-1}) f(y_T | X_T, \alpha_{X_T}, \beta_{X_T}) \\ &\propto P(X_T | X_{T-1}) \left(\frac{y_T^{\alpha_{X_T}-1} (1-y_T)^{\beta_{X_T}-1}}{B(\alpha_{X_T}, \beta_{X_T})} \right) \end{aligned} \quad (7.17)$$

$$\begin{aligned} p(X_t | X^{(-t)}, \alpha_{X_t}, \beta_{X_t}, \mathbf{Y}^{(T)}, \mathbf{P}) &\propto P(X_t | X_{t-1}) f(y_t | X_t, \alpha_{X_t}, \beta_{X_t}) P(X_{t+1} | X_t) \\ &\propto P(X_t | X_{t-1}) \left(\frac{y_t^{\alpha_{X_t}-1} (1-y_t)^{\beta_{X_t}-1}}{B(\alpha_{X_t}, \beta_{X_t})} \right) P(X_{t+1} | X_t) \end{aligned} \quad (7.18)$$

Burada, $(\alpha_{X_h}, \beta_{X_h})$ değerleri, X Markov zincirinin $h \in \mathbf{T}$ zamanında içinde bulunduğu saklı duruma ilişkin şekil parametreleridir.

Tez çalışmasında, Beta-SMM' de tüm parametreler için tam koşullu sonsal dağılımlar kapalı formda elde edilmiştir. Elde edilen tam koşullu sonsal dağılımlara bakıldığında, sadece P ve $X^{(T)}$ parametreleri için doğrudan örneklem çekilebilmektedir. Bu durumda, $(\alpha^{(m)}, \beta^{(m)})$ parametrelerinin tam koşullu bileşik sonsal dağılımından dolayı olarak örneklem çekilmelidir. Bunun için, Metropolis-Hastings algoritması veya kabul-red örnekleme kullanılabılır. Bu çalışmada, Metropolis-Hasting algoritması tercih edilmiştir. Gibbs örnekleme içinde Metropolis-Hastings algoritması kullanılarak, $f(\alpha^{(m)}, \beta^{(m)} | X^{(T)}, Y^{(T)})$ tam koşullu bileşik sonsal yoğunluğundan dolayı olarak örneklem çekilmiştir. Beta-SMM için, Gibbs örneklemesinin adımları aşağıdaki şekilde uyarlanmıştır:

Adım 1: $X_0^{(T)} = (X_1^0, X_2^0, \dots, X_T^0)$ gibi bir başlangıç saklı durum vektörünün belirlenmesi,

Adım 2: $X^{(T)}$ verildiğinde, P_i 'nin tam koşullu sonsal dağılımından (Eşitlik 7.13) durum geçiş olasılıkları matrisinin oluşturulması,

Adım 3: $X^{(T)}$ ve $Y^{(T)}$ verildiğinde, $(\alpha^{(m)}, \beta^{(m)})$ ' in tam koşullu bileşik sonsal dağılımından (Eşitlik 7.15) Metropolis-Hastings algoritması kullanılarak beta dağılımı şekil parametrelerinin türetilmesi,

Adım 4: Üretilen tüm değerler verildiğinde, saklı durumlar dizisi için yeni örnekler Eşitlik 7.16 - Eşitlik 7.18 kullanılarak oluşturulması,

Adım 5: Durağan bileşik sonsal örneklem elde etmek için 2 - 4. adımların tekrarlanması.

Yukarıda Beta-SMM için uyarladığımız Gibbs algoritması sonucu, Θ 'nın bileşik sonsal dağılımından bağımlı örnekler elde edilir. Gibbs örnekleme ile oluşturulan zincirin ergodik olması nedeniyle, başlangıç değerleri durağan dağılım üzerinde etkili değildir. Bu nedenle, $X_0^{(T)}$ başlangıç değerlerinin seçimi önemli değildir. Bu değerler, sadece zincirin performansını ve durağan dağılımına yakınsama hızını etkiler. İterasyon sayısı arttıkça zincir denge durumuna yaklaşır. Gibbs örnekleme yeterli uzunlukta çalıştırıldığında başlangıç değerlerinin etkileri yok olur. Başlangıç değerlerinin etkileri yok olduğunda, üretilen zincir değerleri durağan hale gelir. Diğer bir ifadeyle, zincir denge durumuna ulaşır. Zincir denge durumuna ulaştığında, Gibbs örneklemesinden elde edilen örneklem değerleri $f(\Theta | Y^{(T)})$ hedef dağılımından çekilen değerlere karşılık gelir. Bu ne-

denle, durağan dağılıma yakınsama gerçekleşene kadar geçmesi gereken iterasyon sayısı belirlenmelidir. Belirlenen bu iterasyondan sonra Gibbs örneklemesinden elde edilen örneklem değerleri kullanılarak, Beta-SMM' nin parametre tahminleri hesaplanır. Herhangi bir parametre için tam koşullu sonsal dağılımının tek tepeli olduğu durumda, sonsal ortalama veya sonsal tepe değeri; çok tepeli olduğu durumda ise, sonsal tepe değeri bir yüksek yoğunluk noktasıdır. Bu yüksek yoğunluk noktalarından uygun olan herhangi biri hesaplanarak ilgili parametrenin Bayesci tahminine ulaşılır. Ayrıca, durağan zincir değerleri kullanılarak standart sapma, standart hata, yüzde değerler ve Bayes güven aralıkları gibi parametrelerin özetleyici istatistikleri de hesaplanabilir.

7.1.2. Gözlem Değerlerinin Öngörülmesi

m -durumlu Beta-SMM' de gelecek dönemler için oluşacak gözlem değerleri kestirilebilir. k , pozitif bir tamsayı olsun. $Y^{(T)}$ gözlemlendikten sonra meydana gelecek herhangi bir Y_{T+k} gözlemi için sonsal öngörü yoğunluğu aşağıdaki gibi hesaplanır:

$$f(Y_{T+k} | Y^{(T)}) = \sum_{i=1}^m \int_{\Theta} f(Y_{T+k} | \alpha_i, \beta_i) P^{(k)}(X_{T+k} = S_i | X_T) f(\Theta | Y^{(T)}) d\Theta \quad (7.19)$$

Burada, $f(\Theta | Y^{(T)})$, model parametrelerinin bileşik sonsal yoğunluğudur. $P^{(k)}(X_{T+k} | X_T)$ ise, k -adım geçiş olasılık değeridir. Eşitlik 7.19' da verilen integral çok boyutlu ve karmaşık bir yapıdadır. Bu nedenle, $f(Y_{T+k} | Y^{(T)})$ sonsal öngörü yoğunluğu doğrudan hesaplanamaz. Öngörü yoğunluğu, model parametrelerinin bileşik sonsal dağılımına yakınsama sağlandıktan sonraki Gibbs örneklemesinden elde edilen örnek değerler kullanılarak Monte Carlo yöntemiyle tahmin edilebilir. G değeri, başlangıç değerlerinin etkisi yok edildikten sonraki toplam iterasyon sayısı olsun. Bu durumda, sonsal öngörü yoğunluğu,

$$f(Y_{T+k} | Y^{(T)}) \approx \frac{1}{G} \sum_{g=1}^G \sum_{i=1}^m f(Y_{T+k} | (\alpha_i, \beta_i)^{(g)}) P^{(k)}(X_{T+k} = S_i | X_T^{(g)}) \quad (7.20)$$

olarak tahmin edilir. $X_T^{(g)}$ saklı durumu ve $(\alpha_i, \beta_i)^{(g)}$ şekil parametreleri, Gibbs örneklemesinin g -inci iterasyonundaki örnek değerleridir. Eşitlik 7.20' de verilen yaklaşımdan yararlanılarak, seçilen $Y_{T+k} > 0$ değerleri için sonsal öngörü yoğunlukları tahmin edilir. Elde edilen sonsal öngörü yoğunluk tahminleri kullanılarak, sonsal öngörü dağılımına ulaşılır ve k -ıncı gelecek dönem gözlem değerine ilişkin ilgili çıkarsamalar yapılır.

7.1.3. Modelin Boyut Sayısının Tahmini

Saklı Markov modelinin boyutu, X Markov zincirinin durum sayısı tarafından belirlenir. Model üzerinden yapılacak istatistiksel çıkarsamalar, saklı durumların sayısına bağlı olarak değişir. Bu nedenle, öncelikle $Y^{(T)}$ gözlemlerini en iyi açıklayacak saklı durum sayısı belirlenmelidir. SMM' nin boyut sayısının belirlenmesi, bir model seçim problemi olarak görülür. SMM' nin Bayesci analizinde, modelin boyut sayısının belirlenmesi önemli bir sorundur. Bayesci yaklaşımda model seçimi, Bayes faktörü (BF) kullanılarak gerçekleştirilir.

Genel olarak, birbirine rakip \mathcal{M}_1 ve \mathcal{M}_2 modelleri tanımlandığında ve D gözlemleri kullanılarak bu modellerin karşılaştırılması istendiğinde; Bayes teoreminden yararlanılarak, \mathcal{M}_i ($i = 1, 2$) modelinin sonsal olasılık değeri aşağıdaki gibi hesaplanır [90], [91].

$$P(\mathcal{M}_i | D) = \frac{f(D | \mathcal{M}_i)P(\mathcal{M}_i)}{f(D | \mathcal{M}_1)P(\mathcal{M}_1) + f(D | \mathcal{M}_2)P(\mathcal{M}_2)} \quad (7.21)$$

Modellerin sonsal olasılıkları toplamı, $P(\mathcal{M}_1 | D) + P(\mathcal{M}_2 | D) = 1$ dir. Eşitlik 7.21' de $P(\mathcal{M}_i)$, \mathcal{M}_i modelinin önsel olasılığıdır. Modellerin önsel olasılıkları toplamı da $P(\mathcal{M}_1) + P(\mathcal{M}_2) = 1$ dir. $P(D | \mathcal{M}_i)$ ise, \mathcal{M}_i modeli için gözlemlerin marjinal olasılık değeridir. Modellerin sonsal olasılıkları oranlanığında, sonsal odds değeri elde edilir. \mathcal{M}_2 ' ye karşı \mathcal{M}_1 ' in sonsal odds değeri aşağıdaki gibi hesaplanır:

$$\frac{P(\mathcal{M}_1 | D)}{P(\mathcal{M}_2 | D)} = \left[\frac{f(D | \mathcal{M}_1)}{f(D | \mathcal{M}_2)} \right] \left[\frac{P(\mathcal{M}_1)}{P(\mathcal{M}_2)} \right] \quad (7.22)$$

Eşitlik 7.22' nin sağ tarafındaki birinci terim, modellerin marjinal olasılıklarının oranıdır. İkinci terim ise, önsel odds değeridir. Bayes faktörü, rakip iki modelin marjinal olasılıklarının oranı olarak tanımlanır. \mathcal{M}_1 modelinin \mathcal{M}_2 modeline karşı Bayes faktörü, BF_{12} olarak gösterilir. BF_{12} değeri aşağıda verilmiştir:

$$BF_{12} = \frac{f(D | \mathcal{M}_1)}{f(D | \mathcal{M}_2)} \quad (7.23)$$

Eşitlik 7.23' te verilen BF_{12} değeri, \mathcal{M}_2 ' ye kıyasla \mathcal{M}_1 ' in lehine gözlemler tarafından sağlanan kanıtın ölçüsüdür. Eşitlik 7.22, sözel olarak

$$\text{Sonsal Odds} = \text{Bayes Faktörü} \times \text{Önsel Odds}$$

şeklinde ifade edilebilir. Bu durumda, Bayes faktörü, sonsal odds' un önsel odds' a oranı olarak hesaplanabilir. Modellerin tercih edilebilirliği açısından herhangi bir ön bilgi olmaması durumunda, önsel olasılıklar eşit olarak alınır. Yani, $P(\mathcal{M}_1) = P(\mathcal{M}_2) = 0.5$ ' tir. Dolayısıyla, Bayes faktörü sonsal odds değerine eşit olur.

Bilimsel teorilerin karşılaştırılmasında Bayes faktörünün kullanımı, ilk olarak Jeffreys [92] tarafından önerilmiştir [91]. Jeffreys [38], [92], hipotez testlerinde Bayesci yaklaşımı geliştirmiştir ve Bayes faktörünü hipotez testlerinde ilk olarak kullanmıştır [90]. Model seçiminde Bayes faktörünün kullanımı ise ilk olarak Schwarz [93] ve Raftery [94] tarafından ortaya konulmuştur [90]. Bayes faktörünün yorumlanmasında, Jeffreys [38] tarafından bir kural geliştirilmiştir. Jeffreys' in kuralı, günümüzde Bayes faktörü yorumunda yaygın olarak kullanılmaktadır. Alternatif olarak, Kass ve Raftery [90] tarafından önerilen kuralda kullanılabilir. BF_{12} değeri için Jeffreys' in kuralı aşağıdaki tabloda verilmiştir [38].

Çizelge 7.1. Jeffreys' in Kuralı ile BF_{12} Bayes Faktörünün Model Seçimindeki Yorumu

Bayes Faktörü Değeri	Yorumu
$BF_{12} < 0.01$	\mathcal{M}_2 lehine kesin kanıt mevcuttur.
$0.01 < BF_{12} < 0.03$	\mathcal{M}_2 lehine çok kuvvetli kanıt mevcuttur.
$0.03 < BF_{12} < 0.1$	\mathcal{M}_2 lehine kuvvetli kanıt mevcuttur.
$0.1 < BF_{12} < 0.32$	\mathcal{M}_2 lehine makul kanıt mevcuttur.
$0.32 < BF_{12} < 1$	\mathcal{M}_2 lehine zayıf kanıt mevcuttur.
$1 < BF_{12} < 3.2$	\mathcal{M}_1 lehine zayıf kanıt mevcuttur.
$3.2 < BF_{12} < 10$	\mathcal{M}_1 lehine makul kanıt mevcuttur.
$10 < BF_{12} < 32$	\mathcal{M}_1 lehine kuvvetli kanıt mevcuttur.
$32 < BF_{12} < 100$	\mathcal{M}_1 lehine çok kuvvetli kanıt mevcuttur.
$BF_{12} > 100$	\mathcal{M}_1 lehine kesin kanıt mevcuttur.

$BF_{12} > 1$ durumunda, \mathcal{M}_2 ' ye kıyasla \mathcal{M}_1 tercih edilir ve BF_{12} değeri büyüdükçe artan derecede \mathcal{M}_1 tercih edilir. $BF_{12} < 1$ durumunda ise, \mathcal{M}_1 ' e kıyasla \mathcal{M}_2 tercih edilir ve BF_{12} değeri azaldıkça artan derecede \mathcal{M}_2 tercih edilir [91]. Bayes faktörünün değerleri ve detaylı yorumları, Çizelge 7.1.' de görülmektedir.

Bayes Faktörünün hesaplanabilmesi için, rakip modellerin marjinal olabilirlik değerleri bilinmelidir. Marjinal olabilirlik, model parametrelerinin sonsal dağılımının nor-

malleştirme katsayısıdır. Herhangi bir \mathcal{M} modelinin $f(\mathbf{D} \mid \mathcal{M})$ marjinal olabilirliği, olabilirlik fonksiyonunun önsel yoğunluğa göre parametre kümesi üzerinden integrali alınarak aşağıdaki gibi hesaplanır [88], [90], [91]:

$$f(\mathbf{D} \mid \mathcal{M}) = \int_{\Theta_{\mathcal{M}}} f(\mathbf{D} \mid \Theta_{\mathcal{M}}, \mathcal{M}) f(\Theta_{\mathcal{M}} \mid \mathcal{M}) d\Theta_{\mathcal{M}} \quad (7.24)$$

$\Theta_{\mathcal{M}}$, \mathcal{M} modeline ait parametre kümesidir. $f(\mathbf{D} \mid \Theta_{\mathcal{M}}, \mathcal{M})$, modelin olabilirlik fonksiyonu; $f(\Theta_{\mathcal{M}} \mid \mathcal{M})$, model parametrelerinin önsel yoğunluğudur. Bayes faktörünün analitik olarak hesaplanabilmesi, Eşitlik 7.24' te verilen integralin hesaplanmasını gerektirir. Modelin parametre sayısı arttıkça, bu integralin hesaplanması oldukça güçleşir. Modelin yapısına göre, hem çok boyutlu hem de karmaşık bir integral işlemi gerçekleştirilmesi gerekebilir. Bu nedenle, Eşitlik 7.24' te verilen integralin hesaplanması kolay olmamakla birlikte analitik çözümü her zaman mümkün değildir.

Marjinal olabilirlik değerinin analitik olarak hesaplanmasının mümkün olmadığı durumda, Laplace yöntemi, Monte Carlo yöntemi, önem örnekleme, Gauss kareleştirme ve sonsal dağılım üzerinden benzetim yöntemi en çok kullanılan yöntemlerdir [90], [95], [96]. Ayrıca, gözlem sayısının fazla olduğu durumda, Schwarz [93] tarafından geliştirilen Schwarz Bayesci bilgi kriteri (SBIC) kullanılarak Bayes faktörünün kaba bir tahmini de elde edilebilir [90], [95].

Gözlem ve model parametrelerinin sayısı orta düzeyde olduğunda, Laplace yöntemiyle elde edilen marjinal olabilirlik tahminleri Bayes faktörü için genellikle yeterli yaklaşımı sağlamaktadır. Elde edilen Bayes faktörü tahmininden yapılacak yorumlar yeterince doğru olmaktadır. Kass, Tierney ve Kadane [97]' nin ayrıntılı olarak belirttiği koşullar altında, gözlem sayısı arttıkça marjinal olabilirlik değerinin tahmini için görelî hata $\mathcal{O}(n^{-1})$ olarak elde edilmektedir. Bayes faktörünün hem pay hem de paydası için Laplace yöntemi uygulandığında ortaya çıkan Bayes faktörü tahmininin görelî hata değeri de $\mathcal{O}(n^{-1})$ olmaktadır. Dolayısıyla, gözlem sayısı arttıkça yöntemin başarısı artmaktadır. Kass ve Raftery [90], gözlem ve parametre sayısı için kabaca bir tahminde bulunmuşlardır. Gözlemlerin, parametre sayısının beş katından az olduğu durumda yöntemin iyi sonuç vermediğini; fakat yirmi katından fazla olduğu durumda yöntemin iyi çalıştığını belirtmişlerdir [90], [95], [96]. Bu durumda, gözlem sayısı az ve parametre sayısı fazla olduğunda Laplace yöntemi başarılı bir şekilde çalışmayacaktır.

Monte Carlo yöntemi kullanılarak marjinal olabilirlik değeri ise, aşağıdaki gibi tahmin edilir [90], [95]:

$$\hat{f}(\mathbf{D} | \mathcal{M}) = \frac{1}{G} \sum_{g=1}^G f(\mathbf{D} | \boldsymbol{\Theta}_{\mathcal{M}}^{(g)}, \mathcal{M}) \quad (7.25)$$

Eşitlik 7.25' te $\{\boldsymbol{\Theta}_{\mathcal{M}}^{(g)} : g = 1, 2, \dots, G\}$ değerleri, $f(\boldsymbol{\Theta}_{\mathcal{M}} | \mathcal{M})$ önsel dağılımdan çekilen örnek değerlerdir. Önsel dağılımdan doğrudan örneklem çekilemediğinde, benzetim yöntemlerinden biri kullanılarak dolaylı olarak örneklem çekilebilir. Sonsal dağılım, önsel dağılıma göre daha konsantre olduğunda $\boldsymbol{\Theta}_{\mathcal{M}}^{(g)}$ parametre değerlerinin çoğu için olabilirlik değeri küçük bulunur. Bu durum, benzetim sürecini oldukça verimsizleştirir. Çünkü, $\hat{f}(\mathbf{D} | \mathcal{M})$ tahmin değeri, büyük olan az sayıdaki olabilirlik değerlerinin etkisinde kalır. Bunun bir sonucu olarak, tahmin değerinin varyansı büyük olur ve normal dağılıma yakınsaması yavaşlar [90], [98].

Monte Carlo yönteminin kesinliğinin geliştirilmesinde önem örnekleme kullanılabılır. Uygun bir $f^*(\boldsymbol{\Theta}_{\mathcal{M}} | \mathcal{M})$ önem fonksiyonu belirlenerek marjinal olabilirlik değeri aşağıdaki gibi tahmin edilebilir [90], [99]:

$$\hat{f}(\mathbf{D} | \mathcal{M}) = \frac{\sum_{g=1}^G w_g f(\mathbf{D} | \boldsymbol{\Theta}_{\mathcal{M}}^{(g)}, \mathcal{M})}{\sum_{g=1}^G w_g} \quad (7.26)$$

Burada $\{\boldsymbol{\Theta}_{\mathcal{M}}^{(g)} : g = 1, 2, \dots, G\}$ değerleri, $f^*(\boldsymbol{\Theta}_{\mathcal{M}} | \mathcal{M})$ önem dağılımından çekilen değerlerdir. g -inci önem ağırlığı,

$$w_g = \frac{f(\boldsymbol{\Theta}_{\mathcal{M}}^{(g)} | \mathcal{M})}{f^*(\boldsymbol{\Theta}_{\mathcal{M}}^{(g)} | \mathcal{M})}$$

olarak tanımlanmıştır. Eşitlik 7.26' da verilen marjinal olabilirliğin tahmin değeri, benzetim tutarlıdır. $G \rightarrow +\infty$ durumunda, $\hat{f}(\mathbf{D} | \mathcal{M})$ tahmin değeri $f(\mathbf{D} | \mathcal{M})$ gerçek değerine hemen hemen kesinlikle yakınsar. Yani, G değeri arttıkça elde edilen tahmin değeri o kadar iyi olur.

Gauss kareleştirme yöntemi ise, model parametre sayısının orta düzeyde olduğu durumlarda marjinal olabilirliğin tahmininde etkili bir şekilde çalışmaktadır. Kass ve Raftery [90], parametre sayısının kabaca dokuzdan az olması gerektiğini belirtmişlerdir.

Eşitlik 7.24' te verildiği gibi marjinal olabilirlik değeri, olabilirlik fonksiyonunun model parametrelerinin önsel dağılımına göre integrali alınarak bulunmaktadır. Yukarıda anlatılan yöntemler parametrelerin önsel dağılımından çekilen örnek değerleri kullanarak marjinal olabilirlik değerini tahmin etmektedir. Fakat, model parametreleri sonsal dağılımdan çekilen örnek değerler üzerinden tahmin edilmiştir ve model parametrelerinin sonsal örnekleri mevcuttur. $f(\Theta_{\mathcal{M}} | \mathbf{D}, \mathcal{M})$ sonsal dağılımından çekilen örneklem üzerinden marjinal olabilirlik değerinin tahmin edilmesi zordur [88].

Newton ve Raftery [100], sonsal dağılımdan elde edilen örnek değerleri kullanarak marjinal olabilirlik değerini olabilirlik değerlerinin harmonik ortalaması olarak tahmin etmişlerdir [88], [90], [95].

$$\hat{f}_{NR}(\mathbf{D} | \mathcal{M}) = \left\{ \frac{1}{G} \sum_{g=1}^G \left(\frac{1}{f(\mathbf{D} | \Theta_{\mathcal{M}}^{(g)}, \mathcal{M})} \right) \right\}^{-1} \quad (7.27)$$

Eşitlik 7.27' de $\{\Theta_{\mathcal{M}}^{(g)} : g = 1, 2, \dots, G\}$ değerleri, $f(\Theta_{\mathcal{M}} | \mathbf{D}, \mathcal{M})$ sonsal dağılımından çekilen değerlerdir. $\Theta_{\mathcal{M}}^{(g)}$ örnek değerleri, Markov zinciri Monte Carlo yöntemlerinden biri kullanılarak elde edilebilir. $\hat{f}_{NR}(\mathbf{D} | \mathcal{M})$, marjinal olabilirlik değerinin benzetim tutarlı bir tahmin edicisidir. $G \rightarrow +\infty$ durumunda, $\hat{f}_{NR}(\mathbf{D} | \mathcal{M})$ tahmin değeri, $f(\mathbf{D} | \mathcal{M})$ gerçek değerine hemen hemen kesinlikle yakınsar. Fakat, genellikle Merkezi limit teoremini sağlamaz ve kararsız bir tahmin edicidir. Çünkü, bazı $\Theta_{\mathcal{M}}^{(g)}$ parametre değerleri için olabilirlik değerleri küçük bulunur ve bu değerler $\hat{f}_{NR}(\mathbf{D} | \mathcal{M})$ tahmin değeri üzerinde büyük etkiye sahiptir. Olabilirliğin tersi sonlu varyansa sahip değildir [88], [90], [95]. Newton ve Raftery [100]' nin geliştirdiği tahmin edicinin kararsızlığının giderilmesi ve merkezi limit teoremini sağlaması için tahmin edicinin üzerinde birkaç değişiklik önerilmiştir.

Gelfand ve Dey [101], $\hat{f}_{NR}(\mathbf{D} | \mathcal{M})$ ' nin üzerinde küçük bir değişiklik yaparak marjinal olabilirlik değeri için aşağıdaki tahmin ediciyi önermişlerdir [88], [90], [95]:

$$\hat{f}_{GD}(\mathbf{D} | \mathcal{M}) = \left\{ \frac{1}{G} \sum_{g=1}^G \left(\frac{h(\Theta_{\mathcal{M}}^{(g)})}{f(\mathbf{D} | \Theta_{\mathcal{M}}^{(g)}, \mathcal{M}) f(\Theta_{\mathcal{M}}^{(g)} | \mathcal{M})} \right) \right\}^{-1} \quad (7.28)$$

Eşitlik 7.28' de $\{\Theta_{\mathcal{M}}^{(g)} : g = 1, 2, \dots, G\}$ değerleri, $f(\Theta_{\mathcal{M}} | \mathbf{D}, \mathcal{M})$ sonsal dağılımdan çekilen değerlerdir. $h(\cdot)$, olabilirlik ve önsel yoğunluğun çarpımından daha ince kuyruklara sahip olan bir dağılımdır. $\hat{f}_{GD}(\mathbf{D} | \mathcal{M})$, marjinal olabilirliğin kararlı bir tahmin edicisidir. Yansız ve tutarlıdır. Merkezi limit teoremini sağlar. Düşük boyutlu bir parametre uza-

yında $h(\cdot)$ dağılımının iyi bir seçimiyle çok doğru bir tahmin değeri elde edilebilir. Fakat, yüksek boyutlu parametre uzayında uygun bir $h(\cdot)$ dağılımı belirlemek zordur ve marjinal olabilirliğin kötü bir tahmin değeri elde edilebilir [88], [90], [95].

Bu durumda, sonsal dağılımdan elde edilen örnek değerler kullanılarak marjinal olabilirliğin tahmininde daha iyi sonuçlar veren bir yönteme ihtiyaç duyulmuştur. Chib [88], marjinal olabilirliğin hesaplanmasında bir yöntem geliştirmiştir. Chib' in yöntemiyle hesaplanan tahmin değerleri için yukarıda belirtilen problemlerden hiç biri görülmemektedir. Bu yöntem, Gibbs örneklemesinin kullanıldığı ve tüm parametrelerin tam koşullu sonsal dağılımlarının bilindiği durumlarda kullanılabilir. Saklı Markov modelinin Bayesci çözümlemesinde Chib' in yönteminin nasıl uygulanacağı ilk olarak 2006 yılında Hock ve Soyer [102] ve 2012 yılında Pievatolo, Ruggeri ve Soyer [13] tarafından gösterilmiştir.

Aşağıda belli bir modelin marjinal olabilirliğinin hesaplanması üzerinden Chib' in yöntemi anlatılacaktır. Bu nedenle, gösterimin sadeleştirilmesi açısından \mathcal{M} model simgesi kaldırılmıştır. Marjinal olabilirlik değeri, olabilirlik ile önsel yoğunluğun çarpımının sonsal yoğunluğa bölünmesiyle bulunabilir. Chib, marjinal olabilirliğin bu tanımından faydalanmıştır. Tanıma göre, marjinal olabilirlik değeri aşağıdaki gibi yazılır [88]:

$$f(\mathbf{D}) = \frac{f(\mathbf{D} | \Theta)f(\Theta)}{f(\Theta | \mathbf{D})} \quad (7.29)$$

Burada, Θ ' nın seçimi önemli değildir. Herhangi bir Θ değeri için, Eşitlik 7.29 sağlanmaktadır. Chib, yöntemin etkinliği açısından sonsal dağılımın yüksek yoğunlukta olduğu bölgeden bir Θ değeri seçilmesini önermiştir. Çünkü, yüksek yoğunluklu bölgeden seçilen bir değer düşük yoğunluklu kuyruk bölgesinden seçilen bir değere göre, sonsal dağılımı çok daha doğru temsil edecektir. Chib' in yönteminde, Θ değeri için bir yüksek yoğunluk noktasının seçimi $\hat{f}(\mathbf{D})$ tahmin değerinin doğruluğunu iyileştirmektedir. Gibbs örneklemesiyle her bir parametre için elde edilen sonsal örneklem üzerinden hesaplanan sonsal tepe değeri ve en çok olabilirlik tahmin değeri sonsal dağılımın birer yüksek yoğunluk noktalarıdır. Düşük yoğunluk noktası olmaması koşuluyla, sonsal ortalama değeride seçilebilir [88]. Sonsal ortalamanın bir düşük yoğunluk noktası olması, ilgili parametrenin çok tepeli bir sonsal dağılıma sahip olmasından kaynaklanır.

Marjinal olabilirlik değerinin hesaplanması amacıyla, bir Θ^* yüksek yoğunluk noktası seçilsin. Eşitlik 7.29' da $f(\mathbf{D} | \Theta^*)$ olabilirlik değeri ve $f(\Theta^*)$ önsel yoğunluğu doğru-

dan hesaplanabilirken; $f(\Theta^* | D)$ sonsal yoğunluğu doğrudan hesaplanamaz. Sonsal yoğunluk, yeniden Gibbs örnekleme gerçekleştirilerek tahmin edilir. Bunun sonucunda, marginal olabilirlik değerinin tahmini aşağıdaki gibi elde edilir [13], [88]:

$$\hat{f}(D) = \frac{f(D | \Theta^*)f(\Theta^*)}{\hat{f}(\Theta^* | D)} \quad (7.30)$$

Eşitlik 7.30' da verilen $\hat{f}(D)$, kararlı bir tahmin edicidir. Bu tahmin değerine ilişkin tüm hata, sonsal yoğunluğun Gibbs örneklemeyle tahmin edilmesi sürecinde meydana gelmektedir.

Tez çalışmasında, Beta-SMM' de boyut sayısına karar verme aşamasında yukarıda tanımlanan Bayes faktöründen yararlanılmıştır. Chib' in yöntemi kullanılarak m -durumlu Beta-SMM' nin marjinal olabilirlik değerinin nasıl tahmin edileceği araştırılmıştır. Beta-SMM için, Eşitlik 7.30' da verilen marjinal olabilirlik tahmin değeri aşağıdaki gibi yazılabilir:

$$\hat{f}(\mathbf{y}^{(T)}) = \frac{\overbrace{f(\mathbf{y}^{(T)} | \boldsymbol{\alpha}^{*(m)}, \boldsymbol{\beta}^{*(m)}, \mathbf{X}^{*(T)})}^{\text{Olabilirlik}} \overbrace{f(\boldsymbol{\alpha}^{*(m)}, \boldsymbol{\beta}^{*(m)}) p(\mathbf{X}^{*(T)} | \mathbf{P}^*) f(\mathbf{P}^*)}^{\text{Önsel Yoğunluklar}}}{\underbrace{\hat{f}(\boldsymbol{\alpha}^{*(m)}, \boldsymbol{\beta}^{*(m)}, \mathbf{P}^*, \mathbf{X}^{*(T)} | \mathbf{y}^{(T)})}_{\text{Sonsal Yoğunluk}}} \quad (7.31)$$

Eşitlik 7.31' de olabilirlik fonksiyonu, önsel ve sonsal yoğunluklar ayrı ayrı belirtilmiştir. Burada, $\boldsymbol{\alpha}^{*(m)}$, $\boldsymbol{\beta}^{*(m)}$, \mathbf{P}^* ve $\mathbf{X}^{*(T)}$ değerleri, Gibbs örnekleme sonucu elde edilen parametrelerin sonsal tepe değerleridir. Eşitlik 7.31' in payında bulunan tüm terimler, sonsal tepe değerleri kullanılarak hesaplanır.

SMM' de $\mathbf{X}^{(T)}$ durum dizisi verildiğinde, $\mathbf{Y}^{(T)}$ gözlemleri koşullu olarak birbirlerinden bağımsız olduğu görülür. Bu özellikten yararlanıldığında, olabilirlik değeri beta yoğunlukları çarpımına eşit olarak aşağıdaki gibi bulunur:

$$\begin{aligned} f(\mathbf{y}^{(T)} | \boldsymbol{\alpha}^{*(m)}, \boldsymbol{\beta}^{*(m)}, \mathbf{X}^{*(T)}) &= \prod_{t=1}^T f(y_t | \boldsymbol{\alpha}^{*(m)}, \boldsymbol{\beta}^{*(m)}, \mathbf{X}^{*(T)}) \\ &= \prod_{t=1}^T \frac{y_t^{\alpha_{X_t^*} - 1} (1 - y_t)^{\beta_{X_t^*} - 1}}{B(\alpha_{X_t^*}, \beta_{X_t^*})} \end{aligned} \quad (7.32)$$

Eşitlik 7.31' de önsel yoğunluklar, parametreler için önerilen ilgili önsel dağılımlar kullanılarak hesaplanır. SMM' de S_i durumları birbirinden bağımsızdır. Dolayısıyla, beta

dağılımı şekil parametrelerinin bileşik önsel yoğunluğu aşağıdaki gibi elde edilir:

$$\begin{aligned} f(\boldsymbol{\alpha}^{*(m)}, \boldsymbol{\beta}^{*(m)}) &= \prod_{i=1}^m f(\alpha_i^*, \beta_i^*) \\ &= \prod_{i=1}^m \frac{e^{-a_i \alpha_i^*} e^{-b_i \beta_i^*}}{[B(\alpha_i^*, \beta_i^*)]^{d_i}} \end{aligned} \quad (7.33)$$

$X^{*(T)}$ sonsal tepe değeri, bir Markov zinciridir. Bu nedenle, $p(X^{*(T)} | \mathbf{P}^*)$ önsel olasılığı aşağıdaki gibi hesaplanır:

$$\begin{aligned} p(X^{*(T)} | \mathbf{P}^*) &= P(X_1^*, X_2^*, \dots, X_T^* | \mathbf{P}^*) \\ &= p(X_1^* | \mathbf{P}^*) p(X_2^* | X_1^*, \mathbf{P}^*) p(X_3^* | X_2^*, \mathbf{P}^*) \dots p(X_T^* | X_{T-1}^*, \mathbf{P}^*) \\ &= \underbrace{\pi_{X_1^*}^*}_{\text{Başlangıç Olasılığı}} \underbrace{P(X_2^* | X_1^*) P(X_3^* | X_2^*) \dots P(X_T^* | X_{T-1}^*)}_{\text{Geçiş Olasılıkları}} \end{aligned} \quad (7.34)$$

Eşitlik 7.34' te $\pi_{X_1^*}^*$ değeri, saklı sürecin başlangıç anında X_1^* durumunda olması olasılığıdır. Bu değer, $\boldsymbol{\pi}^* = (\pi_1^*, \pi_2^*, \dots, \pi_m^*)$ başlangıç olasılık vektörünün bir elemanıdır. $\boldsymbol{\pi}^*$ vektörü, \mathbf{P}^* geçiş olasılıkları matrisinin durağan dağılımı olarak hesaplanır. Eşitlik 7.34' te belirtilen geçiş olasılıkları ise, \mathbf{P}^* geçiş olasılıkları matrisinin elemanlarıdır.

SMM' de \mathbf{P}_i satır vektörlerinin birbirinden bağımsız olmasından dolayı, $f(\mathbf{P}^*)$ önsel yoğunluğu Dirichlet yoğunlukları çarpımı olarak aşağıdaki şekilde bulunur:

$$\begin{aligned} f(\mathbf{P}^*) &= \prod_{i=1}^m f(\mathbf{P}_i^*) \\ &= \prod_{i=1}^m \left[\frac{\Gamma\left(\sum_{j=1}^m \nu_{ij}\right)}{\prod_{j=1}^m \Gamma(\nu_{ij})} \prod_{j=1}^m p_{ij}^{*(\nu_{ij}-1)} \right] \end{aligned} \quad (7.35)$$

Eşitlik 7.35' te p_{ij}^* sonsal tepe değeri, herhangi bir $t \in \mathbf{T}$ zamanında \mathcal{S}_i durumundan \mathcal{S}_j durumuna geçiş olasılığıdır.

Eşitlik 7.31' in paydasındaki sonsal yoğunluk değeri doğrudan hesaplanamadığı için, bu yoğunluk yeniden Gibbs örnekleme kullanılarak tahmin edilir. Bu sonsal yoğunluk aşağıdaki gibi parçalanabilir:

$$f(\boldsymbol{\alpha}^{*(m)}, \boldsymbol{\beta}^{*(m)}, \mathbf{P}^*, \mathbf{X}^{*(T)} | \mathbf{Y}^{(T)}) = f(\boldsymbol{\alpha}^{*(m)}, \boldsymbol{\beta}^{*(m)} | \mathbf{X}^{*(T)}, \mathbf{Y}^{(T)}) f(\mathbf{P}^* | \mathbf{X}^{*(T)}) p(\mathbf{X}^{*(T)} | \mathbf{Y}^{(T)}) \quad (7.36)$$

Eşitlik 7.36' da $f(\boldsymbol{\alpha}^{*(m)}, \boldsymbol{\beta}^{*(m)} | \mathbf{X}^{*(T)}, \mathbf{Y}^{(T)})$ ve $f(\mathbf{P}^* | \mathbf{X}^{*(T)})$ sonsal yoğunlukları doğrudan hesaplanabilirken, $p(\mathbf{X}^{*(T)} | \mathbf{Y}^{(T)})$ sonsal olasılığı doğrudan hesaplanamaz. Bu sonsal olasılık değerinin tahmin edilebilmesi için, yeniden Gibbs örnekleme gerçekleştirilir. Gibbs örneklemesinden elde edilen örneklem üzerinden Monte Carlo yöntemi uygulanarak $\hat{p}(\mathbf{X}^{*(T)} | \mathbf{Y}^{(T)})$ tahmini elde edilir. Tüm bu işlemlerin sonucunda, Eşitlik 7.31' in paydasının $\hat{f}(\boldsymbol{\alpha}^{*(m)}, \boldsymbol{\beta}^{*(m)}, \mathbf{P}^*, \mathbf{X}^{*(T)} | \mathbf{y}^{(T)})$ tahmin değerine ulaşılır.

S_i ' lerin birbirinden bağımsız olmasından dolayı, beta dağılımı şekil parametrelerinin tam koşullu bileşik sonsal yoğunluğu aşağıdaki gibi hesaplanır:

$$f(\boldsymbol{\alpha}^{*(m)}, \boldsymbol{\beta}^{*(m)} | \mathbf{X}^{*(T)}, \mathbf{Y}^{(T)}) = \prod_{i=1}^m f(\alpha_i^*, \beta_i^* | \mathbf{X}^{*(T)}, \mathbf{Y}^{(T)}) \quad (7.37)$$

Burada, $f(\alpha_i^*, \beta_i^* | \mathbf{X}^{*(T)}, \mathbf{Y}^{(T)})$ bileşik sonsal yoğunluğu Eşitlik 7.14' te verilmiştir.

P_i satır vektörlerinin birbirinden bağımsız olması sonucu, geçiş olasılıkları matrisinin tam koşullu sonsal yoğunluğu aşağıdaki gibi hesaplanır:

$$f(\mathbf{P}^* | \mathbf{X}^{*(T)}) = \prod_{i=1}^m f(P_i^* | \mathbf{X}^{*(T)}) \quad (7.38)$$

Burada, $f(P_i^* | \mathbf{X}^{*(T)})$ sonsal yoğunlukları Eşitlik 7.12' de verilmiştir.

Eşitlik 7.36' da doğrudan hesaplanamayan $p(\mathbf{X}^{*(T)} | \mathbf{Y}^{(T)})$ sonsal olasılığı çarpım kuralı kullanılarak aşağıdaki gibi yazılabilir:

$$p(\mathbf{X}^{*(T)} | \mathbf{Y}^{(T)}) = p(X_1^* | \mathbf{Y}^{(T)}) p(X_2^* | X_1^*, \mathbf{Y}^{(T)}) p(X_3^* | \mathbf{X}^{*(2)}, \mathbf{Y}^{(T)}) \dots p(X_t^* | \mathbf{X}^{*(t-1)}, \mathbf{Y}^{(T)}) \dots \\ \dots p(X_{T-1}^* | \mathbf{X}^{*(T-2)}, \mathbf{Y}^{(T)}) p(X_T^* | \mathbf{X}^{*(T-1)}, \mathbf{Y}^{(T)}) \quad (7.39)$$

$\mathbf{X}^{(t-1)} = (X_1, X_2, \dots, X_{t-1})$ olarak tanımlanmıştır. Eşitlik 7.39' da ilk terimden sonraki terimlere indirgenmiş koşullular denir. İlk terim aşağıdaki gibi hesaplanır:

$$p(X_1^* | \mathbf{Y}^{(T)}) = \int p(X_1^* | \mathbf{Y}^{(T)}, \boldsymbol{\alpha}^{(m)}, \boldsymbol{\beta}^{(m)}, \mathbf{X}^{(s>1)}, \mathbf{P}) \\ p(\boldsymbol{\alpha}^{(m)}, \boldsymbol{\beta}^{(m)}, \mathbf{X}^{(s>1)}, \mathbf{P} | \mathbf{Y}^{(T)}) d\boldsymbol{\alpha}^{(m)} d\boldsymbol{\beta}^{(m)} d\mathbf{X}^{(s>1)} d\mathbf{P} \quad (7.40)$$

$\mathbf{X}^{(s>1)} = (X_2, X_3, \dots, X_T)$ olarak tanımlanmıştır. Eşitlik 7.40' ta verilen integral çok

boyutlu ve karmaşık yapıdadır. Bu nedenle, Monte Carlo yönteminden yararlanılarak $p(X_1^* | \mathbf{Y}^{(T)})$ sonsal olasılığı aşağıdaki gibi tahmin edilir:

$$\begin{aligned} p(X_1^* | \mathbf{Y}^{(T)}) &\approx \frac{1}{G} \sum_{g=1}^G p\left(X_1^* | (\boldsymbol{\alpha}^{(m)})^{(g)}, (\boldsymbol{\beta}^{(m)})^{(g)}, (\mathbf{X}^{(s>1)})^{(g)}, \mathbf{P}^{(g)}, \mathbf{Y}^{(T)}\right) \\ &\approx \frac{1}{G} \sum_{g=1}^G \pi_{X_1^* | (\boldsymbol{\alpha}^{(m)})^{(g)}, (\boldsymbol{\beta}^{(m)})^{(g)}, (\mathbf{X}^{(s>1)})^{(g)}, \mathbf{P}^{(g)}, \mathbf{Y}^{(T)}}^{(g)}. \end{aligned} \quad (7.41)$$

Eşitlik 7.40' ta $\mathbf{Y}^{(T)}$ gözlemleri verildiğinde model parametrelerinin dağılımı mevcuttur. Bu nedenle, $(\boldsymbol{\alpha}^{(m)})^{(g)}$, $(\boldsymbol{\beta}^{(m)})^{(g)}$, $(\mathbf{Y}^{(s>1)})^{(g)}$ ve $\mathbf{P}^{(g)}$ değerleri, model parametrelerinin tahmin edilmesinde gerçekleştirilen Gibbs örneklemesinin g -inci iterasyonundaki parametre değerleridir. $\pi_{X_1^*}^{(g)}$ olasılığı, g -inci iterasyonda saklı sürecin X_1^* durumunda başlaması olasılığıdır. $\pi^{(g)}$ bağılangıç durum dağılımı, g -inci iterasyondaki $\mathbf{P}^{(g)}$ geçiş olasılıkları matrisinin durağan dağılımı olarak hesaplanır. Eşitlik 7.39' da $t \geq 2$ için indirgenmiş koşullular aşağıdaki gibi hesaplanır:

$$\begin{aligned} p\left(X_t^* | \mathbf{X}^{*(t-1)}, \mathbf{Y}^{(T)}\right) &= \int p\left(X_t^* | \boldsymbol{\alpha}^{(m)}, \boldsymbol{\beta}^{(m)}, \mathbf{P}, \mathbf{X}^{(s>t)}, \mathbf{X}^{*(t-1)}, \mathbf{Y}^{(T)}\right) \\ &\quad p\left(\boldsymbol{\alpha}^{(m)}, \boldsymbol{\beta}^{(m)}, \mathbf{P}, \mathbf{X}^{(s>t)} | \mathbf{X}^{*(t-1)}, \mathbf{Y}^{(T)}\right) d\boldsymbol{\alpha}^{(m)} d\boldsymbol{\beta}^{(m)} d\mathbf{X}^{(s>t)} d\mathbf{P} \end{aligned} \quad (7.42)$$

Burada, $\mathbf{X}^{(s>t)} = (X_{t+1}, X_{t+2}, \dots, X_T)$ olarak tanımlanmıştır. Eşitlik 7.42' deki integral çok boyutlu ve karmaşık yapıdadır. Bu nedenle, Monte Carlo yöntemi kullanılarak bu integralin çözümü tahmin edilir. Buradaki Monte Carlo yönteminde $\{(\boldsymbol{\alpha}^{(m)})^{(g)}, (\boldsymbol{\beta}^{(m)})^{(g)}, (\mathbf{P})^{(g)}, (\mathbf{X}^{(s>t)})^{(g)} : g = 1, 2, \dots, G\}$ Gibbs örnekleme parametre değerleri kullanılamaz. Çünkü, bu parametre değerleri, sadece $\mathbf{Y}^{(T)}$ gözlemleri bilindiğinde elde edilen tam koşullu sonsal dağılımlardan türetilmişlerdir. Eşitlik 7.42' de ise, hem $\mathbf{X}^{*(t-1)}$ durum dizisi hem de $\mathbf{Y}^{(T)}$ gözlemleri bilindiğinde elde edilen tam koşullu sonsal dağılımlardan türetilen $\boldsymbol{\alpha}^{(m)}$, $\boldsymbol{\beta}^{(m)}$, \mathbf{P} ve $\mathbf{X}^{(s>t)}$ model parametrelerinin değerleri kullanılarak Monte Carlo yöntemi uygulanmalıdır. Bunun için yeniden Gibbs örnekleme gerçekleştirilmelidir. Gibbs örneklemesinde kullanılacak olan tam koşullu sonsal dağılımlar aşağıda verilmiştir:

$$f(\boldsymbol{\alpha}^{(m)}, \boldsymbol{\beta}^{(m)} | X_t, \mathbf{X}^{(s>t)}, \mathbf{X}^{*(t-1)}, \mathbf{Y}^{(T)}) \quad (7.43)$$

$$f(\mathbf{P} | X_t, \mathbf{X}^{(s>t)}, \mathbf{X}^{*(t-1)}) \quad (7.44)$$

$$p(X_t | \boldsymbol{\alpha}^{(m)}, \boldsymbol{\beta}^{(m)}, \mathbf{P}, \mathbf{X}^{(s>t)}, \mathbf{X}^{*(t-1)}, \mathbf{Y}^{(T)}) \quad (7.45)$$

$$p(X_h | \boldsymbol{\alpha}^{(m)}, \boldsymbol{\beta}^{(m)}, \mathbf{P}, X_t, \mathbf{X}^{(-h, s>t)}, \mathbf{X}^{*(t-1)}, \mathbf{Y}^{(T)}), \quad h = (t+1), (t+2), \dots, T \quad (7.46)$$

$\mathbf{X}^{(-h,s>t)} = \{X_s : s > t \text{ ve } s \neq h\}$ olarak tanımlanmıştır. Eşitlik 7.43 - Eşitlik 7.46' da verilen tam koşullu sonsal dağılımlar üzerinden Gibbs örneklemesi gerçekleştirilir. Burada, $\mathbf{X}^{*(t-1)}$ durum dizisi, model parametrelerinin tahmininde gerçekleştirilen Gibbs örneklemesi sonuçları kullanılarak elde edilen sonsal tepe değerleridir. Eşitlik 7.43 - Eşitlik 7.46' da ki tam koşullu sonsal dağılımlardan çekilen örnek değerler $p(\boldsymbol{\alpha}^{(m)}, \boldsymbol{\beta}^{(m)}, \mathbf{P}, \mathbf{X}^{(s>t)} \mid \mathbf{X}^{*(t-1)}, \mathbf{Y}^{(T)})$ dağılımından gelir. Bu durumda, Monte Carlo yöntemi kullanılarak $p(X_t^* \mid \mathbf{X}^{*(t-1)}, \mathbf{Y}^{(T)})$ ' nin tahmini aşağıdaki gibi hesaplanır:

$$p(X_t^* \mid \mathbf{X}^{*(t-1)}, \mathbf{Y}^{(T)}) \approx \frac{1}{\hat{G}} \sum_{k=1}^{\hat{G}} p(X_t^* \mid (\boldsymbol{\alpha}^{(m)})^{(k)}, (\boldsymbol{\beta}^{(m)})^{(k)}, \mathbf{P}^{(k)}, (\mathbf{X}^{(s>t)})^{(k)}, \mathbf{X}^{*(t-1)}, \mathbf{Y}^{(T)}) \quad (7.47)$$

Burada, \hat{G} iterasyon sayısıdır. $(\boldsymbol{\alpha}^{(m)})^{(k)}$, $(\boldsymbol{\beta}^{(m)})^{(k)}$, $\mathbf{P}^{(k)}$ ve $(\mathbf{X}^{(s>t)})^{(k)}$ değerleri, Gibbs örneklemesinde k -nıncı iterasyonunda parametrelerin aldıkları değerlerdir. $\forall t \in \{2, 3, \dots, T\}$ için indirgenmiş koşullu olasılıkların Eşitlik 7.47' de verilen tahmin değeri hesaplanır. Böylece, Eşitlik 7.36' da verilen tüm terimlerin değerleri elde edilir. Sonuç olarak, $f(\boldsymbol{\alpha}^{*(m)}, \boldsymbol{\beta}^{*(m)}, \mathbf{P}^*, \mathbf{X}^{*(T)} \mid \mathbf{Y}^{(T)})$ sonsal yoğunluğunun tahmin değerine ulaşılır. Sonrasında, m -durumlu Beta-SMM' nin $f(\mathbf{Y}^{(T)})$ marjinal olabilirlik değerinin tahmini hesaplanır.

Beta-SMM' nin boyut sayısına karar vermek amacıyla, farklı m değerleri için Chib' in yöntemi kullanılarak marjinal olabilirlik değerleri yukarıda anlatıldığı şekilde tahmin edilir. Burada, SMM' de boyut sayısı, $m \in \{2, 3, \dots\}$ olabilir. $i, j \in \{2, 3, \dots\}$ ve $i \neq j$ olmak üzere, i ve j indisleri Beta-SMM' nin boyut sayısını gösterdiğinde, $\forall i < j$ durumu için, i -durumlu Beta-SMM' nin j -durumlu Beta-SMM' ye karşı BF_{ij} Bayes faktörü tahmin edilebilir. Ardından, Çizelge 7.1.' de verilen Jeffreys' in kuralı kullanılarak Bayes faktörleri yorumlanır ve gözlemleri en iyi açıklayan boyut sayısına karar verilir.

İzleyen bölümde, Poisson-SMM ve Beta-SMM gerçek veriler kullanılarak işletilecek ve her iki model için elde edilen sonuçlar verilecektir.

8. UYGULAMA

Tez çalışmasında, Poisson ve beta dağılım aileleri varsayımı altında saklı Markov modeli incelenmiş; Beta-SMM için teorik yapı geliştirilmiştir. SMM, X Markov sürecinin durum kümesini oluşturan dağılım ailesi parametrelerinin analitik olarak elde edilebilme durumlarına göre, klasik ve Bayesci yaklaşım altında ele alınmıştır. Poisson dağılım ailesi parametrelerinin tahminlerine kapalı formda ulaşılabilmesi nedeniyle, Poisson-SMM klasik yaklaşım altında modellenmiştir. Beta dağılım ailesi parametrelerinin kapalı formda tahminlerine ulaşılamaması nedeniyle, Beta-SMM Bayesci yaklaşım altında modellenmiştir. Bu bölümde, belirlenen gerçek veri setleri üzerinden Poisson-SMM ve Beta-SMM işletilerek elde edilen sonuçlar yorumlanmıştır. Her iki modelin işletiminde özgün kodlar geliştirilmiştir. İlk olarak, klasik yaklaşım altında Poisson-SMM; daha sonra Bayesci yaklaşım altında Beta-SMM uygulanmıştır.

Poisson-SMM' nin uygulanmasında belirli bir yerde ve zaman aralığında gerçekleşen deprem frekansları modellenerek deprem riski tahmin edilmiştir. Poisson-SMM' nin Türkiye deprem verisine ilk kez uygulanması Can vd. [82] tarafından yapılmıştır. Bu çalışma, aynı zamanda modelin Türkiye verisine ilk kez uygulanmasıdır. Bayesci Beta-SMM' nin işletiminde ise, beta dağılımının tanım kümesi gereği, $[0 - 1]$ aralığında değişen oran değerleri kullanılmıştır. Bu tez çalışması, beta dağılım ailesi varsayımı altında saklı Markov modelinin ilk Bayesci çözümlemesidir.

8.1. Poisson Saklı Markov Modelinin Uygulanması

İnsan yaşamı ve yapılar üzerinde ciddi oranda hasar verici etkisi olması nedeniyle, deprem önemli bir doğal tehlikedir. Günümüzde, deprem doğasının ve davranışının anlaşılması, büyük depremlerin oluşturacağı tehlikenin belirlenmesinde ve gerekli önlemlerin alınmasında oldukça önemlidir. Belli bir yerde ve zamanda gerçekleşen deprem sayıları, genellikle aşırı yayılım ve serisel bağımlılık içermektedir. Bu durum, Poisson-SMM' nin, özellikle Poisson süreçlerine kıyasla, deprem frekanslarını esnek ve etkili bir şekilde modellemesine ve gelecek deprem frekanslarını başarılı bir şekilde tahmin etmesine olanak sağlar. Poisson-SMM, belirtilen ardışık zaman aralıklarında gerçekleşen depremlerin frekanslarının Poisson dağılım ailesinden türetildiği ve bu ailenin oran parametresinin homojen bir Markov zinciri olduğu bir modeldir.

Burada, Poisson-SMM kullanılarak 1900 ve 2012 yılları arasında Bilecik' in merkez olduğu 100 km yarıçapındaki bir alanda meydana gelmiş, büyüklüğü 4 ve daha fazla olan depremlerin yıllık frekansları modellenmiştir. Gelecek 35 yıllık bir periyod (2013-2047) için, her yıl 4 ve daha fazla büyüklükte oluşacak olan depremlerin frekansları tahmin edilerek çalışma sahası için deprem riski belirlenmiştir. Ayrıca Poisson-SMM' den elde edilen tahmin sonuçları, Poisson süreçleri ile yapılan modellemelerden elde edilen sonuçlarla karşılaştırılmaktadır. Tüm analizler R programında ilgili kodlar yazılarak gerçekleştirilmiştir.

Poisson-SMM' nin işletiminde çalışma sahası olarak seçilen Bilecik, Türkiye' nin en büyük iki kenti olan İstanbul ve Ankara arasında yer almaktadır. Dolayısıyla, önemli karayolu ve demiryolunun geçtiği ve buna bağlı olarak köprü, viyadük, tünel gibi birçok ulaşım yapısının inşa edileceği bir yer olma özelliğine sahiptir. Bunlara ek olarak, dünyanın sismik açıdan en aktif bölgelerinden biri olan Kuzey Anadolu Fay Zonu yakınında bulunmaktadır. Tüm bu özellikler dikkate alınarak, Bilecik ve çevresi çalışma sahası olarak belirlenmiştir. Poisson-SMM için kullanılan deprem verileri Boğaziçi Üniversitesi Kandilli Rasathanesi ve Deprem Araştırma Enstitüsü kataloglarından [103] sağlanmıştır.

Poisson-SMM ile çalışılan sahadaki deprem riskinin belirlenebilmesi için ilk olarak modelin saklı durumlarının sayısına karar verilmesi gerekmektedir. Saklı durumların sayısının $m \in \{1, 2, 3, 4, 5, 6\}$ olduğu kabul edilsin. $m = 1$ durumu, bir tane oran parametresi olduğunu göstermekle birlikte Poisson sürecini ifade etmektedir. Çalışmada, en çok olabilirlik tahmin yönteminden yararlanılarak Poisson sürecine ilişkin oran parametresi tahmin edilmiştir ve oran parametresinin tahmini kullanılarak olabilirlik değeri hesaplanmıştır. $m \geq 2$ durumunda ise, EM algoritmasından yararlanılarak m -durumlu Poisson-SMM' nin parametreleri tahmin edilmiştir ve parametre tahminlerinden olabilirlik değeri hesaplanmıştır. Aşağıdaki tabloda modellere ilişkin parametre sayıları (p), negatif log-olabilirlik değerleri, AIC ve BIC değerleri verilmiştir:

Çizelge 8.1. Model Karşılaştırması-1

Model	p	$-\log \mathcal{L}(\Theta; Y^{(T)})$	AIC	BIC
Poisson Süreci	1	557.4164	1116.833	1119.56
2-durumlu PSMM	4	186.1575	382.3151	395.952
3-durumlu PSMM	9	172.2728	366.5457	396.547
4-durumlu PSMM	16	167.6581	373.3161	425.1365
5-durumlu PSMM	25	166.3246	390.6492	469.7434
6-durumlu PSMM	36	162.4184	406.8368	518.6597

Çizelge 8.1.' de Poisson sürecine ilişkin bilgi kriteri değerleri, PSMM' dekilere kıyasla, büyüktür. Bu durum, ilgili deprem frekanslarının modellenmesinde Poisson-SMM' nin daha etkili olduğunu göstermektedir. En iyi modele karar vermede bilgi kriterleri farklı modelleri önermiştir. AIC' ye göre, 3-durumlu PSMM; BIC' ye göre, 2-durumlu PSMM en iyi model olarak bulunmuştur. Burada, olabilirlik oran testi yapılarak SMM' deki saklı durum sayısına karar verilebilir. Bunun için, %95 güven düzeyinde olabilirlik oran testleri yapılmıştır ve testlere ilişkin sonuçlar aşağıdaki tabloda verilmiştir:

Çizelge 8.2. Model Karşılaştırması-2

	H_0	H_1	sd	χ^2_{Test}	P-Değeri	Karar
1	$m = 1$	$m = 2$	3	742.5178	< 0.01	H_0 Red
2	$m = 2$	$m = 3$	5	27.7694	< 0.01	H_0 Red
3	$m = 3$	$m = 4$	7	9.2294	0.2366	H_0 Kabul
4	$m = 4$	$m = 5$	9	2.667	0.9760	H_0 Kabul
5	$m = 5$	$m = 6$	11	7.8124	0.7300	H_0 Kabul

Çizelge 8.2.' de P-değerlerine bakıldığında, 2-durumlu PSMM, Poisson sürecine göre; 3-durumlu PSMM, 2- ve 4-durumlu PSMM' ye göre; 4-durumlu PSMM, 5-durumlu PSMM' ye göre; 5-durumlu PSMM ise, 6-durumlu PSMM' ye göre daha iyi bir modeldir. Beş testin sonuçları dikkate alındığında, 3-durumlu PSMM' nin en iyi model olduğu sonucuna varılır. Bu durumda, tüm istatistiksel çıkarımlar 3-durumlu PSMM üzerinden yapılır. 3-durumlu PSMM' de gözlemler üç farklı Poisson dağılımından türetilir. Herhangi bir t anında sistemin içinde bulunduğu duruma ilişkin Poisson dağılımından t anındaki gözlem değeri türetilir.

Aşağıda ele alınan modellerde yıllık deprem frekanslarının ortalama ve varyans değerleri, gözlemlerin ortalama ve varyans değerleriyle karşılaştırılmıştır:

Çizelge 8.3. Ortalama ve Varyans Değerlerinin Karşılaştırılması

	Ortalama	Varyans
Gözlemler	2.2301	71.9823
Poisson Süreci	2.2301	2.2301
2-durumlu PSMM	2.2419	66.7821
3-durumlu PSMM	2.2460	67.6174
4-durumlu PSMM	2.1210	62.0436
5-durumlu PSMM	2.1577	63.8925
6-durumlu PSMM	2.1577	69.0683

Çizelge 8.3.' te gözlemlerin ortalama ve varyans değerlerine bakıldığında, aşırı yayılımın (varyans > ortalama) mevcut olduğu görülmektedir. Poisson sürecinde, yıllık deprem frekanslarının ortalama ve varyans değerleri eşittir. Bu nedenle, Poisson süreci gözlemlerin açıklanmasında ve gelecek gözlem değerlerinin tahmininde çok yetersiz kalacaktır. Gözlemlerin ortalama ve varyans değerleriyle olan uyumu ve yıllık deprem frekanslarındaki aşırı yayılımı karşılama derecesi açısından, 3-durumlu PSMM en iyidir. Bu durum, Çizelge 8.2.' de verilen kararı da desteklemektedir.

Gözlemlerin açıklanmasında model performanslarının daha net görülebilmesi amacıyla, Poisson süreci ve en iyi model seçilen 3-durumlu PSMM için yıllık deprem frekanslarının beklenen değerleri hesaplanmıştır. Bu değerler aşağıda verilmiştir:

Çizelge 8.4. 3-Durumlu PSMM ile Poisson Sürecinin Model Uyumu Karşılaştırması

Deprem Frekansları	Gözlenen Toplam Yıl Sayısı	Beklenen Toplam Yıl Sayısı	
		3-Durumlu PSMM	Poisson Süreci
0	56	56.5577	12.1496
1	32	30.4497	27.0948
2	11	11.1066	30.2119
3	5	5.3278	22.4584
4	4	3.2296	12.521
5	0	1.8323	5.5846
6	0	0.8929	2.0757
7	1	0.3749	0.6613
8	1	0.1379	0.1843
9	0	0.0451	0.0457
10	0	0.0133	0.0102
≥ 11	3	3.0323	0.0025
TOPLAM	113	113	113

Çizelge 8.4.' te yıllık deprem frekanslarının gözlenen ve beklenen değerlerine göre, 3-durumlu PSMM' nin verilere uyumunun Poisson sürecine göre son derece başarılı olduğu görülmektedir. Örneğin, 113 yıl içerisinde toplam 56 yılda 4 ve daha fazla büyüklükte hiç deprem meydana gelmemiştir. Poisson süreciyle modelleme yapıldığında bu değer 12 yıl olarak tahmin edilmektedir. Dolayısıyla, Poisson süreci ilgili değeri gerçek değerinin çok altında tahmin (under-estimate) etmektedir. 3-durumlu PSMM ise, bu değeri 56 yıl olarak tahmin etmiştir. Başka bir örnek verilecek olursa, 113 yıl içerisinde toplam 5 yılda 4 ve daha fazla büyüklükte olan 3 deprem meydana gelmiştir. Poisson sürecine göre, bu değer 22.4584 olarak tahmin edilirken; 3-durumlu PSMM' ye göre, 5.3278 olarak tahmin edilmiştir. Bu örnekte ise, Poisson süreci ilgili değeri gerçek değerinin çok üstünde tahmin (over-estimate) etmektedir. Diğer gözlenen değerlere de bakıldığında, Poisson sürecinin gözlemlerin modellenmesinde çok yetersiz kaldığı; 3-durumlu PSMM' nin ise, ölçülen değerlere uyumunun çok iyi olduğu görülmektedir.

3-durumlu PSMM parametrelerinin en çok olasılık tahminleri EM algoritması kullanılarak elde edilmiştir. Bu tahminler aşağıda verilmiştir:

Çizelge 8.5. 3-Durumlu PSMM' nin En Çok Olabilirlik Parametre Tahminleri

Parametreler	En Çok Olabilirlik Tahminleri
Oran Parametreleri	$\lambda^{(3)} = (\lambda_1, \lambda_2, \lambda_3) = (0.4956, 2.9425, 50.6667)$
Geçiş Olasılıkları Matrisi	$P = \begin{bmatrix} 0.8591 & 0.1219 & 0.0190 \\ 0.5743 & 0.3568 & 0.0689 \\ 0.6858 & 0.3142 & 0.000 \end{bmatrix}$
Denge Dağılımı	$\delta = (\delta_1, \delta_2, \delta_3) = (0.8072, 0.1660, 0.0268)$

Çizelge 8.5.' te görüldüğü gibi Poisson dağılım ailesinin oran parametreleri 0.4956, 2.9425 ve 50.6667 olarak tahmin edilmiştir. Bu oran parametreleri, X Markov zincirinin durum kümesini oluşturmaktadır. Markov zincirinin durumlarına ilişkin geçiş olasılıkları ise, Çizelge 8.5.' te P matrisi ile verilmiştir. Markov zincirinin denge dağılımı, δ vektörüdür. Denge dağılımı, $\delta = \delta P$ eşitliğinden hesaplanmıştır. Çizelge 8.5.' te verilen δ ' ya göre, X süreci en yüksek olasılıkla λ_1 durumunda; en düşük olasılıkla λ_3 durumunda olacaktır. Tez çalışmasında yapılacak tüm istatistiksel çıkarsamalarda Markov zincirinin başlangıç durum dağılımı olarak P matrisinin denge dağılımı alınmıştır.

3-durumlu PSMM ile X saklı sürecinin gelecek 35 yılda (2013-2047) bulunacağı durumların olasılıkları tahmin edilmiştir. Yıllara göre, λ_1 , λ_2 ve λ_3 saklı durumlara ilişkin olasılıklar aşağıda verilmiştir:

Çizelge 8.6. Gelecek 35 Yıl için Saklı Durumlara İlişkin Olasılıklar

YIL	2013	2014	2015	2016	2017	2018	2019
λ_1	0.855572	0.820176	0.810677	0.808129	0.807445	0.807261	0.807212
λ_2	0.124787	0.154952	0.163045	0.165217	0.165799	0.165956	0.165998
λ_3	0.019640	0.024872	0.026278	0.026655	0.026756	0.026783	0.026790
TOPLAM	1	1	1	1	1	1	1
YIL	2020	2021	2022	2023	2024	2025	2026
λ_1	0.807199	0.807195	0.807194	0.807194	0.807194	0.807194	0.807194
λ_2	0.166009	0.166012	0.166013	0.166013	0.166013	0.166013	0.166013
λ_3	0.026792	0.026793	0.026793	0.026793	0.026793	0.026793	0.026793
TOPLAM	1	1	1	1	1	1	1
YIL	2027	2028	2029	2030	2031	2032	2033
λ_1	0.807194	0.807194	0.807194	0.807194	0.807194	0.807194	0.807194
λ_2	0.166013	0.166013	0.166013	0.166013	0.166013	0.166013	0.166013
λ_3	0.026793	0.026793	0.026793	0.026793	0.026793	0.026793	0.026793
TOPLAM	1	1	1	1	1	1	1
YIL	2034	2035	2036	2037	2038	2039	2040
λ_1	0.807194	0.807194	0.807194	0.807194	0.807194	0.807194	0.807194
λ_2	0.166013	0.166013	0.166013	0.166013	0.166013	0.166013	0.166013
λ_3	0.026793	0.026793	0.026793	0.026793	0.026793	0.026793	0.026793
TOPLAM	1	1	1	1	1	1	1
YIL	2041	2042	2043	2044	2045	2046	2047
λ_1	0.807194	0.807194	0.807194	0.807194	0.807194	0.807194	0.807194
λ_2	0.166013	0.166013	0.166013	0.166013	0.166013	0.166013	0.166013
λ_3	0.026793	0.026793	0.026793	0.026793	0.026793	0.026793	0.026793
TOPLAM	1	1	1	1	1	1	1

Çizelge 8.6.' da X saklı sürecinin kısa sürede dengeye ulaştığı görülmektedir. Özellikle, 2016 yılı ve sonrasında saklı durumlara ilişkin olasılıklar, Çizelge 8.5.' te verilen δ denge dağılımındakilere çok yakındır. Çizelge 8.6.' da 35 yıllık zaman periyodunun tamamında λ_1 saklı durumu en yüksek olasılığa sahiptir. Bu nedenle, X saklı sürecinin 35 yıl boyunca λ_1 durumunda kalacağı söylenebilir. Daha açık bir ifadeyle, gelecek 35 yıl boyunca Bilecik' in merkez olduğu 100 km yarıçapındaki bir alanda 4 ve daha

fazla büyüklükte meydana gelecek depremler $\lambda_1 = 0.4956$ oran parametrelili Poisson dağılımına göre oluşacaktır.

3-durumlu PSMM ile gelecek 35 yılda 4 ve daha fazla büyüklükte meydana gelecek depremlere ilişkin risk değerleri hesaplanmıştır. 2018 yılı ve sonrasındaki yıllar için deprem frekanslarına ait olasılıklar 2017 yılındaki değerlere çok yakındır. Bu nedenle, aşağıda sadece 2013-2017 yılları için olasılıklar verilmiştir. Ayrıca, deprem sayısının 5' ten büyük olduğu durumlarda çok küçük olasılıklar elde edilmesi nedeniyle, frekans sayısı 0-5 olarak alınmış ve olasılıklar aşağıdaki gibi özetlenmiştir:

Çizelge 8.7. 2013-2017 Yıllarında Deprem Riski

Deprem Frekansı	Yıllar				
	2013	2014	2015	2016	2017
0	0.527809	0.507836	0.502476	0.501038	0.500652
1	0.277675	0.27167	0.270057	0.269625	0.269509
2	0.092496	0.096734	0.097871	0.098176	0.098258
3	0.038516	0.044833	0.046527	0.046982	0.047104
4	0.021864	0.026779	0.028097	0.028451	0.028546
5	0.012226	0.015145	0.015928	0.016138	0.016194
TOPLAM	0.970586	0.962996	0.960957	0.96041	0.960263

Çizelge 8.7.'deki değerlere göre, Bilecik' in merkez olduğu 100 km yarıçapındaki bir alanda 2013-2047 yıllarında %96 ile %97 arasında bir olasılıkla 0-5 tane 4 ve daha fazla büyüklükte deprem meydana gelmesi beklenmektedir. Özellikle, hiç deprem olmaması ve bir deprem olması durumları diğer deprem sayılarına kıyasla daha yüksek bir olasılığa sahiptir. 5' ten fazla deprem meydana gelmesi durumu ise, %3 – %4 gibi düşük bir olasılıkla beklenmektedir.

2013-2047 yılları için hesaplanan deprem riski değerleri dikkate alınarak, bu yıllara ilişkin beklenen deprem sayıları ve güven aralıkları hesaplanmıştır. 2017 yılı ve sonrasındaki yıllar için çok yakın değerler elde edilmesi nedeniyle, burada sadece 2013-2017 yıllarına ilişkin sonuçlar verilmiştir.

Çizelge 8.8. 2013-2017 Yıllarında Beklenen Deprem Sayısı ve Güven Aralıkları

Yıl	Beklenen Deprem Sayısı	%95 Güven Aralığı
2013	1.785	[0.435, 3.135]
2014	2.121	[0.610, 3.631]
2015	2.211	[0.660, 3.762]
2016	2.235	[0.674, 3.796]
2017	2.242	[0.678, 3.806]

Çizelge 8.8.'deki değerlere göre, her yıl ortalama iki tane 4 ve daha fazla büyüklükte depremin gerçekleşmesi beklenmektedir.

Belli bir zaman diliminde meydana gelen olayların frekanslarının stokastik modellenmesinde geleneksel bir yaklaşım olarak Poisson süreçleri kullanılmaktadır. Bu bağlamda, geleneksel bir yaklaşım olan Poisson süreci ile yıllık deprem frekanslarının modellenmesinde en iyi model seçilen 3-durumlu PSMM deprem riski tahminleri açısından karşılaştırılmıştır. Böylece, bu iki modelin gözlem değerlerine ilişkin yapılacak çıkarsamalarda yaratacakları farklılıklar ortaya konulmuştur. Sonuçlar aşağıdaki tablolarda verilmiştir:

Çizelge 8.9. 1-3 Yıl İçindeki Deprem Riski Açısından 3-Durumlu PSMM ve Poisson Sürecinin Karşılaştırılması

Deprem Frekansı	1 Yıl İçinde		2 Yıl İçinde		3 Yıl İçinde	
	3-Durumlu PSMM	Poisson Süreci	3-Durumlu PSMM	Poisson Süreci	3-Durumlu PSMM	Poisson Süreci
0	0.527809	0.107519	0.268040	0.011560	0.134684	0.001243
1	0.277675	0.239777	0.284403	0.051561	0.215292	0.008316
2	0.092496	0.267362	0.173466	0.114986	0.190201	0.027817
3	0.038516	0.198747	0.095212	0.170952	0.134993	0.062035
4	0.021864	0.110806	0.057098	0.190619	0.092144	0.103757
5	0.012226	0.049421	0.035451	0.170039	0.062882	0.138832
6	0.005943	0.018369	0.020757	0.126401	0.041506	0.154804
7	0.002494	0.005852	0.011159	0.080539	0.025858	0.147954
TOPLAM	0.979023	0.997852	0.945585	0.916657	0.897560	0.644758

Çizelge 8.9.' a göre,

- 1 yıl içinde 3-durumlu PSMM' de %97.902; Poisson sürecinde %99.785,
- 2 yıl içinde 3-durumlu PSMM' de %94.559; Poisson sürecinde %91.666,
- 3 yıl içinde 3-durumlu PSMM' de %89.756; Poisson sürecinde %64.476,

olasılıkla 0-7 tane 4 ve daha fazla büyüklükte deprem meydana gelmesi beklenmektedir. Küçük frekans değerleri için elde edilen olasılıkların azalması durumunda deprem riski artmaktadır. Çünkü, daha fazla sayıda depremin meydana gelmesi beklenecektir. Çizelge 8.9.' daki koyu hücrelere dikkat edildiğinde, yıllar geçtikçe Poisson sürecinin, 3-durumlu PSMM' ye kıyasla, daha büyük frekans değerlerine yüksek olasılık verdiği görülmektedir. Genel olarak ilk üç yıl içerisinde 3-durumlu PSMM, 0-3 frekans değerlerine daha çok ağırlık vermiştir. Bu durumda, Poisson süreci ile modelleme yapıldığında deprem riskinin daha fazla olması beklenmektedir. Bunun sonucu olarak, aşırı uyum sorunu oluşmaktadır.

Çizelge 8.10. 4-6 Yıl İçindeki Deprem Riski Açısından 3-Durumlu PSMM ve Poisson Sürecinin Karşılaştırılması

Deprem Frekansı	4 Yıl İçinde		5 Yıl İçinde		6 Yıl İçinde	
	3-Durumlu PSMM	Poisson Süreci	3-Durumlu PSMM	Poisson Süreci	3-Durumlu PSMM	Poisson Süreci
0	0.067482	0.000134	0.033785	0.000014	0.016911	0.000002
1	0.144184	0.001192	0.090373	0.000160	0.054340	0.000021
2	0.166569	0.005317	0.128882	0.000893	0.092186	0.000138
3	0.146384	0.015810	0.135525	0.003320	0.113042	0.000617
4	0.115185	0.035258	0.122204	0.009255	0.115582	0.002063
5	0.086839	0.062903	0.101958	0.020640	0.106490	0.005522
6	0.063084	0.093519	0.080822	0.038357	0.091743	0.012315
7	0.043697	0.119175	0.061075	0.061100	0.074919	0.023539
8	0.028793	0.132885	0.044014	0.085161	0.058286	0.039371
9	0.018156	0.131709	0.030369	0.105509	0.043390	0.058534
10	0.011040	0.117489	0.020177	0.117647	0.031062	0.078321
11	0.006507	0.095277	0.012973	0.119256	0.021485	0.095271
12	0.003727	0.070825	0.008098	0.110813	0.014410	0.106231
13	0.002076	0.048599	0.004917	0.095048	0.009395	0.109341
14	0.001125	0.030966	0.002908	0.075702	0.005966	0.104503
15	0.000594	0.018415	0.001677	0.056274	0.003696	0.093220
TOPLAM	0.905443	0.979471	0.879759	0.899150	0.852902	0.729008

Çizelge 8.10.' a göre,

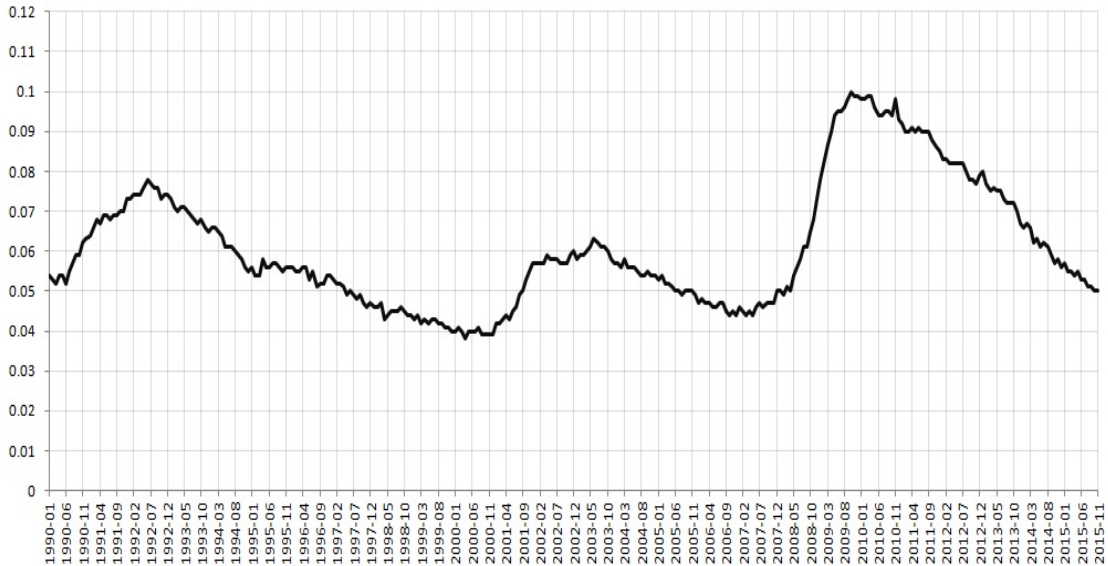
- 4 yıl içinde 3-durumlu PSMM' de %90.544; Poisson sürecinde %97.947,
- 5 yıl içinde 3-durumlu PSMM' de %87.976; Poisson sürecinde %89.915,
- 6 yıl içinde 3-durumlu PSMM' de %85.290; Poisson sürecinde %72.901

olasılıkla 0-15 tane 4 ve daha fazla büyüklükte depremin meydana gelmesi beklenmektedir. Çizelge 8.10.' da koyu hücelere dikkat edildiğinde, 3-durumlu PSMM' nin küçük frekans değerlerine yüksek olasılık verdiği; Poisson sürecinin ise, daha büyük frekans değerlerine yüksek olasılık verdiği görülmektedir. Buna göre, 3-durumlu PSMM' ye kıyasla Poisson süreci deprem riskini daha fazla göstermektedir. Sonuç olarak, aşırı yayılım ve serisel bağımlılık içeren frekans değerlerinin modellenmesinde Poisson sürecinin, Poisson-SMM' ye kıyasla, çok yetersiz kalmakla birlikte aşırı uyum sorunu yaratmaktadır.

8.2. Beta Saklı Markov Modelinin Uygulanması

Beta-SMM' de, gözlemlerin Beta dağılım ailesinden geldiği varsayılması nedeniyle, modelin işletiminde kullanılacak veriler $[0, 1]$ tanım aralığında değer almalıdır. Beta dağılımı, genellikle yüzde değerlerin modellenmesinde kullanılmaktadır. Bir yüzde değer olması nedeniyle, Beta-SMM' nin işletiminde işsizlik oranının kullanılması tercih edilmiştir. Bu bölümde, ilgili veri seti üzerinden Beta-SMM' nin Bayesci çözümlemesi Gibbs örnekleme kullanılarak gerçekleştirilmiş ve gözlemler için öngörü değerleri elde edilmiştir. Beta-SMM' nin Bayesci çözümlemesinde yapılan tüm analizler, R programında ilgili kodlar yazılarak gerçekleştirilmiştir. Sadece, Gibbs örnekleme çıktılarının, $f(\Theta | Y^{(T)})$ bileşik sonsal dağılıma yakınsama durumunu kontrol etmek amacıyla, R programında bulunan **Coda** paketindeki Heidelberg - Welch ve Geweke yakınsama testlerinden yararlanılmıştır.

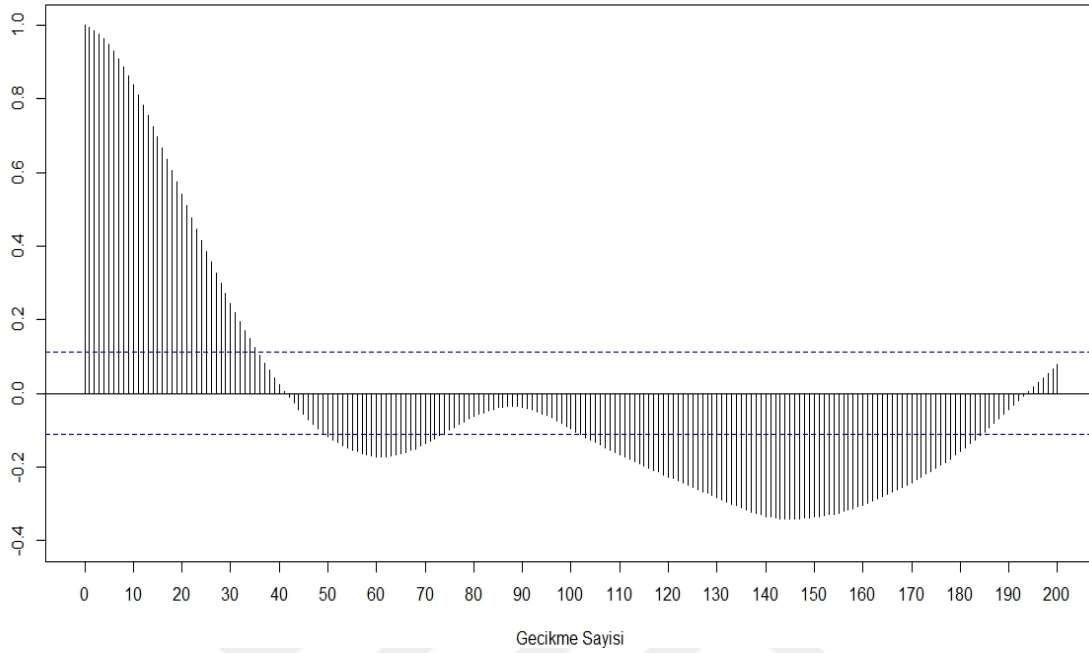
Tez çalışmasında, Amerika Birleşik Devletleri' nin (ABD) 1990-01 : 2015-11 dönemlerindeki işsizlik oranları üzerinden Beta-SMM' nin Bayesci çözümlemesi yapılmıştır. Veriler, ekonomik kalkınma ve işbirliği örgütünün (OECD) [104] veri tabanından alınmıştır. 311 döneme ait aylık işsizlik oranları, %3.8 ile %10 arasında değişmektedir. Bu değerler aşağıda gösterilmiştir:



Şekil 8.1. 1990-01 : 2015:11 Dönemleri ABD İşsizlik Oranları

Verilen dönemlere ilişkin işsizlik oranlarının ortalaması, %6.098' dir ve varyansı, %0.024' tür. Varyans değerinin çok küçük olması, verilerin tanım aralığından kaynaklan-

maktadır. $[0,1]$ tanım aralığındaki bir veri seti için, %0.024' lük gibi bir değişimin modellemeye etkisi analizler sonucunda görülecektir. İşsizlik oranlarına ilişkin otokorelasyon fonksiyonu aşağıda verilmiştir:



Şekil 8.2. 1990-01 : 2015:11 Dönemleri ABD İşsizlik Oranlarının Otokorelasyon Fonksiyonu

Şekil 8.2.' de ilk 35 gecikme ve diğer birçok gecikmeye ait otokorelasyon değerleri güven sınırını aşmıştır. Gözlemler arasında aşırı derecede serisel bağımlılık mevcuttur.

SMM ile modellemede ilk olarak, gözlemleri en iyi açıklayan saklı durum sayısına karar verilerek modelin boyut sayısı belirlenmelidir. Bunun için, X sürecinin farklı durum sayılarına sahip olduğu kabul edilir. Modelin Bayesci çözümlemesinin yapılması nedeniyle, farklı m değerlerine sahip rakip Beta-SMM' ler için Bayes faktörleri hesaplanır. Bayes faktöründe yer alan marjinal olabilirlik değerlerinin hesaplanmasında ise, Chib [88]' in yöntemi kullanılmıştır. Böylece, her bir modelin parametrelerinin sonsal dağılımından çekilen örnek değerler üzerinden ilgili modelin marjinal olabilirlik değeri tahmin edilmiştir. Karmaşık bir yapıya sahip olan SMM için, Chib' in yöntemi hem zamandan hem de bilgisayarın depolama alanı bakımından tasarruf sağlamaktadır.

Tez çalışmasında, X saklı sürecinin durum sayısına $m \in \{2, 3, 4\}$ olarak karar verilmiştir. Bilgisayardaki hesaplama zamanının çok fazla olması ve bilgisayarın depolama alanı kısıtı nedeniyle, $m \geq 5$ değerleri için Beta-SMM'nin Bayesci çözümlemesi yapıla-

mamıştır. Belirlenen her bir m değeri için ayrı Gibbs örnekleme gerçekleştirilmiştir. Gibbs örnekleme sonucu oluşturulan zincirlerde başlangıç değerlerinin etkisini artırmak amacıyla, zincirlerde belli sayıda iterasyon göz ardı edilmiştir. Göz ardı edilecek iterasyon sayısına, Gibbs örnekleme çıktıları görsel olarak incelenerek karar verilmiştir. Zincirlerin görsel denetiminde, iz grafiği, hareketli ortalama grafiği ve otokorelasyon fonksiyonu grafiğinden yararlanılmıştır. Başlangıç değerlerinin etkileri artıldıktan sonraki iterasyonlar üzerinden zincirlerin durağan dağılıma ulaşip ulaşmadığı, yakınsaklık testlerinden yararlanılarak kontrol edilmiştir. Zincirlerin yakınsaklık durumlarının kontrolünde, Heidelberger - Welch ve Geweke yakınsama testleri kullanılmıştır. $f(\Theta | Y^{(T)})$ durağan dağılıma yakınsama sağlandıktan sonraki zincir değerleri kullanılarak, model parametrelerinin Bayesci tahminleri elde edilmiştir. Parametrelerin ilgili tam koşullu sonsal dağılımlarının bir yüksek yoğunluk noktası olan sonsal tepe değerleri, parametrelerin Bayesci tahminleri olarak kabul edilmiştir. Sonrasında, parametrelerin Bayesci tahminleri üzerinden ilgili modeller için, $f(Y^{(T)})$ marjinal olabilirlik değerleri tahmin edilerek Bayes faktörleri hesaplanmıştır. Bayes faktörü tanımı gereği, $\ln f(Y^{(T)})$ logaritmik marjinal olabilirlik değeri en büyük olan model gözlemleri en iyi açıklayan modeldir. m^* en iyi saklı durum sayısına karar verildikten sonra, m^* -durumlu Beta-SMM üzerinden gelecek dönemlere ilişkin işsizlik oranları tahmin edilmiştir.

Modelde saklı durum sayısı arttıkça parametre sayısı da artmaktadır. m -durumlu Beta-SMM' de, toplam $m(m + 2)$ tane parametre mevcuttur. Tüm m değerleri düşünüldüğünde, toplam parametre sayısı (47 tane) çok fazla olmaktadır. Bu nedenle, parametrelere ilişkin Gibbs çıktılarının iz grafikleri, hareketli ortalama grafikleri ve otokorelasyon grafikleri burada verilmemiştir. Bu grafikler görsel olarak incelenerek, başlangıç değer etkilerinin artıldığı iterasyon sayısı, ilk 10000 iterasyon olarak belirlenmiştir. Ayrıca, parametre sayısı arttıkça, zincirin durağan dağılıma yakınsaması için daha çok iterasyona ihtiyaç duyulmuştur. Bu nedenle, Gibbs örneklemesindeki toplam iterasyon sayısı, m değerine göre farklılık göstermiştir. Her bir m değerine ilişkin model parametrelerinin sayısı, gerçekleştirilen toplam iterasyon sayısı ve süreleri aşağıdaki gibidir:

Çizelge 8.11. Rakip Modeller için Bilgiler

Saklı Durum Sayısı	Parametre Sayısı	Toplam İterasyon Sayısı	Süre (Saat)
$m = 2$	8	800000	5.01
$m = 3$	15	1200000	10.43
$m = 4$	24	2000000	22.30
Toplam	47		

Çizelge 8.11.' de görüldüğü gibi, m değeri arttıkça ihtiyaç duyulan iterasyon sayısı ve buna bağlı olarak hesaplama zamanı büyük oranda artmaktadır. Aşağıda, her bir $m \in \{2, 3, 4\}$ değeri için sırayla Gibbs örneklemesi gerçekleştirilmiştir.

$m = 2$ durumunda, ilk 10000 iterasyon atıldıktan sonra geriye kalan 790000 iterasyon üzerinden yakınsaklık testleri yapılmış ve bu testlere ilişkin sonuçlar aşağıda verilmiştir:

Çizelge 8.12. $m = 2$ için Yakınsaklık Testleri (790000 İterasyon Üzerinden)

		Heidelberg-Welch Testi $\alpha = 0.05$					Geweke Testi	
		Durağanlık Testi		Yarı Genişlik Testi ($\epsilon = 0.01$)			İlk %10 ve Son %50	
		Başlangıç İterasyonu	P-Değeri	Ortalama	Yarı Genişlik	Sonuç	Z-Skoru	P-Değeri
α	α_1	1	0.879	2.04	0.00491	$< \epsilon$	0.4276	0.6689
	α_2	1	0.492	2.05	0.00494	$< \epsilon$	-0.6841	0.4939
β	β_1	1	0.259	2.67	0.00749	$< \epsilon$	-0.1487	0.8818
	β_2	1	0.666	2.67	0.00749	$< \epsilon$	-0.7079	0.4790
P	p_{11}	1	0.874	0.794	0.00131	$< \epsilon$	-0.2983	0.7655
	p_{12}	1	0.874	0.206	0.00131	$< \epsilon$	0.2983	0.7655
	p_{21}	1	0.757	0.206	0.00131	$< \epsilon$	-1.0370	0.2997
	p_{22}	1	0.757	0.794	0.00131	$< \epsilon$	1.0370	0.2997

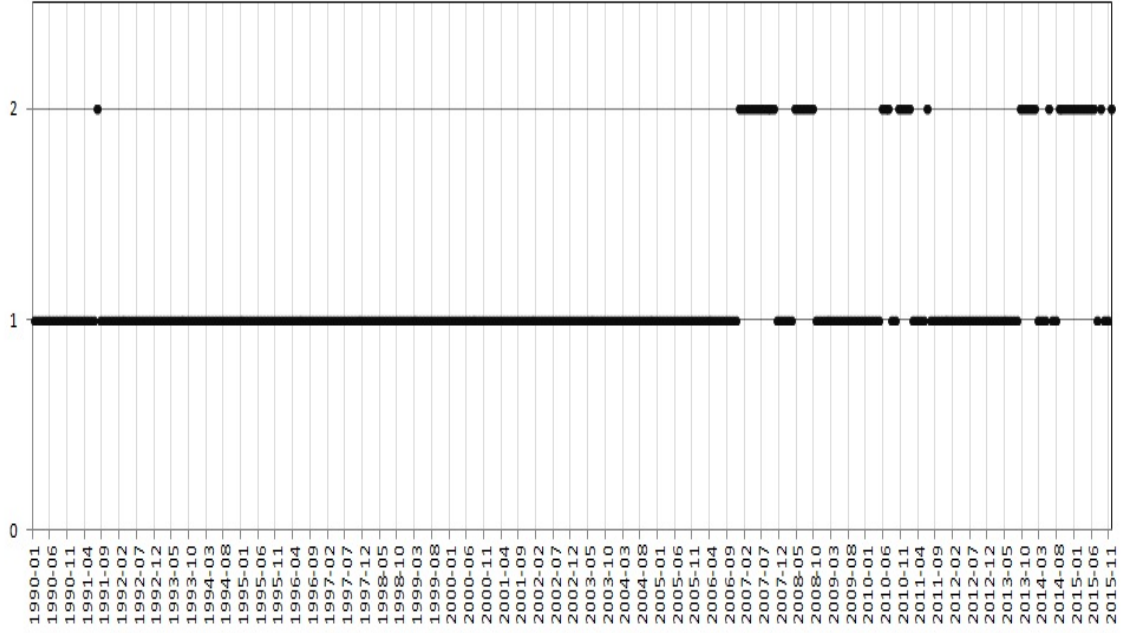
Çizelge 8.12.' de durağanlık testi P-değerlerine bakıldığında, $\forall P > 0.05$ ' tir. Dolayısıyla, %95 güvenirlikle 10001-800000-inci iterasyondaki Gibbs çıktıları, $f(\Theta | Y^{(T)})$ bileşik sonsal dağılımdan gelmektedir. Yarı genişlik testine göre, tüm parametrelerde yarı genişliğin sonsal ortalamaya oranı, ϵ değerinden küçük bulunmuştur. Bu durumda, parametrelerin sonsal ortalamalarının ϵ kesinliğinde tahmin edilmesinde, 790000 iteryondan oluşan zincir kabul edilebilir uzunluktadır. Geweke testine göre ise, tüm

parametrelerde $P > 0.05$ ' tir. %95 güvenirlikle Gibbs çıktılarının ilk %10' luk ve son %50' lik kısımları arasındaki fark istatistiksel olarak anlamsız bulunmuştur. Uygulanan iki yakınsaklık testine göre, ilk 10000 iterasyon sonucunda durağan dağılıma yakınsama sağlandığı %95 güvenirlikle söylenebilir. 2-durumlu Beta-SMM parametrelerinin Bayesci tahminleri, 10001-800000-inci iterasyonlardaki örnek değerlerin sonsal tepe değerleri olarak elde edilir. Model parametrelerinin Bayesci tahminleri ve %95 en yüksek sonsal yoğunluk aralıklarının alt-üst sınırları aşağıda verilmiştir:

Çizelge 8.13. 2-Durumlu Beta-SMM Parametrelerinin Bayesci Tahminleri

Parametre	Sonsal Tepe Değeri	Alt Sınır	Üst Sınır
α	α_1	0.75271	0.00651 4.68219
	α_2	0.73924	0.01022 4.69737
β	β_1	1.32095	$1.69274e - 05$ 6.17938
	β_2	1.34572	$4.57336e - 06$ 6.17405
P	p_{11}	0.99325	0.14888 1
	p_{12}	0.00675	$1.85272e - 08$ 0.85112
	p_{21}	0.00672	$2.16365e - 08$ 0.85203
	p_{22}	0.99328	0.14797 1
π	π_1	0.49887	— —
	π_2	0.50113	— —

Çizelge 8.13.' te beta dağılım ailesine ilişkin parametrelerin Bayesci tahminlerinin birbirlerine çok yakın olduğu görülmektedir. 2-durumlu Beta-SMM' de, X saklı sürecinin bulunduğu durumlar ise aşağıdaki gibi tahmin edilmiştir:



Şekil 8.3. 1990-01 : 2015:11 Dönemlerinde X Saklı Sürecinin Bulunduğu Durumlar

Şekil 8.3.' te X sürecinin genellikle birinci durumda bulunduğu görülmektedir. Bu durumda, gözlemlerin genellikle Beta (α_1, β_1) dağılımından geldiği sonucuna varılır. Parametrelerin hesaplanan sonsal tepe değerleri kullanılarak, Eşitlik 7.31' de verilen marjinal olabilirlik değeri tahmin edilir. Burada, $f(\mathbf{y}^{(T)})$ ' nin paydasındaki $p(\mathbf{X}^{*(T)} | \mathbf{Y}^{(T)})$ olasılığı 500 iterasyondan oluşan ikinci bir Gibbs örnekleme yapılarak tahmin edilmiştir. Bu işlem, 32.67 dakika sürmüştür. Marjinal olabilirliğin pay ve paydasında yer alan yoğunlukların logaritmaları alınarak aşağıda verilmiştir:

Çizelge 8.14. 2-durumlu Beta-SMM için Log-Marjinal Olabilirlik Değerinin Hesaplanması

PAY	PAYDA
$\ln f(\mathbf{y}^{(T)} \boldsymbol{\alpha}^{*(m)}, \boldsymbol{\beta}^{*(m)}, \mathbf{X}^{*(T)}) = -23.85002$	$\ln f(\boldsymbol{\alpha}^{*(m)}, \boldsymbol{\beta}^{*(m)} \mathbf{X}^{*(T)}, \mathbf{Y}^{(T)}) = -720.2098$
$\ln f(\boldsymbol{\alpha}^{*(m)}, \boldsymbol{\beta}^{*(m)}) = -16.68935$	$\ln f(\mathbf{P}^* \mathbf{X}^{*(T)}) = -32.24034$
$\ln p(\mathbf{X}^{*(T)} \mathbf{P}^*) = -107.6683$	$\ln \hat{p}(\mathbf{X}^{*(T)} \mathbf{Y}^{(T)}) = -87.70020$
$\ln f(\mathbf{P}^*) = 0$	
TOPLAM _{PAY} = -148.2077	TOPLAM _{PAYDA} = -840.1503

Çizelge 8.14.' teki değerlere göre, 2-durumlu Beta-SMM için, $\ln f(\mathbf{y}^{(T)})$ log-marjinal olabilirlik değeri 691.9426 olarak elde edilir.

$m = 3$ durumunda, ilk 10000 iterasyon atıldıktan sonra geriye kalan 1190000 iterasyon üzerinden yakınsaklık testleri yapılmış ve bu testlere ilişkin sonuçlar aşağıda verilmiştir:

Çizelge 8.15. $m = 3$ için Yakınsaklık Testleri (1190000 İterasyon Üzerinden)

		Heidelberg-Welch Testi $\alpha = 0.05$					Geweke Testi	
		Durağanlık Testi		Yarı Genişlik Testi ($\epsilon = 0.01$)			İlk %10 ve Son %50	
		Başlangıç İterasyonu	P-Değeri	Ortalama	Yarı Genişlik	Sonuç	Z-Skoru	P-Değeri
α	α_1	119001	0.0951	1.87	0.00468	$< \epsilon$	2.8809	0.0040
	α_2	1	0.9666	1.87	0.00450	$< \epsilon$	-0.4308	0.6666
	α_3	1	0.6671	1.87	0.00451	$< \epsilon$	0.4941	0.6212
β	β_1	1	0.1410	2.46	0.00701	$< \epsilon$	-1.2112	0.2258
	β_2	1	0.7710	2.46	0.00700	$< \epsilon$	0.5526	0.5805
	β_3	1	0.6080	2.45	0.00698	$< \epsilon$	0.3956	0.6924
P	p_{11}	1	0.1550	0.577	0.00170	$< \epsilon$	-1.689	0.0912
	p_{12}	1	0.1100	0.211	0.00098	$< \epsilon$	1.9690	0.0490
	p_{13}	1	0.2140	0.212	0.00099	$< \epsilon$	0.8195	0.4125
	p_{21}	1	0.1570	0.212	0.00098	$< \epsilon$	-1.186	0.2356
	p_{22}	1	0.3540	0.577	0.00160	$< \epsilon$	0.7244	0.4688
	p_{23}	1	0.3290	0.212	0.00098	$< \epsilon$	0.0015	0.9988
	p_{31}	1	0.0892	0.212	0.00098	$< \epsilon$	-1.6480	0.0994
	p_{32}	1	0.9830	0.212	0.00098	$< \epsilon$	0.3931	0.6942
	p_{33}	1	0.3610	0.576	0.00161	$< \epsilon$	0.7484	0.4542

Çizelge 8.15.' teki durağanlık testinde, $\forall P > 0.05$ olmasına karşın, α_1 parametresi için başlangıç iterasyonu 119001 olarak bulunmuştur. Bu durumda, α_1 parametresine ilişkin zincirin ilk %10' luk kısmı olan 119000 adım analiz dışı bırakılmalıdır. α_1 için kalan 1071000 iterasyondaki Gibbs çıktıları, %95 güvenirlikle $f(\Theta | Y^{(T)})$ bileşik sonsal dağılımdan gelmektedir. Yarı genişlik testine göre, tüm parametrelerde yarı genişliğin sonsal ortalamaya oranı ϵ değerinden küçük bulunmuştur. Bu durumda, α_1 dışındaki parametrelerin sonsal ortalamalarının ϵ kesinliğinde tahmin edilmesinde, 1190000 iterasyondan oluşan zincir kabul edilebilir uzunluktadır. α_1 için ise, 1071000 iterasyondan oluşan zincir kabul edilebilir uzunluktadır. Geweke testine göre, α_1 ve p_{12} parametrelerine ilişkin zincirlerin ilk %10' luk ve son %50' lik kısımları arasındaki fark %95 güvenir-

likle istatistiksel olarak anlamlı bulunmuştur. Bu nedenle, zincirin ilk %10'luk kısmı olan 119000 iterasyon atılarak, kalan 1071000 iterasyon üzerinden yakınsaklık testleri yeniden yapılmıştır. Sonuçlar aşağıda verilmiştir:

Çizelge 8.16. $m = 3$ için Yakınsaklık Testleri (1071000 İterasyon Üzerinden)

				Heidelberg-Welch Testi $\alpha = 0.05$			Geweke Testi	
				Durağanlık Testi		Yarı Genişlik Testi ($\epsilon = 0.01$)	İlk %10 ve Son %50	
		Başlangıç İterasyonu	P-Değeri	Ortalama	Yarı Genişlik	Sonuç	Z-Skoru	P-Değeri
α	α_1	1	0.096	1.87	0.00468	$< \epsilon$	0.9192	0.3580
	α_2	1	0.926	1.87	0.00474	$< \epsilon$	0.6404	0.5219
	α_3	1	0.667	1.87	0.00477	$< \epsilon$	-0.3852	0.7001
β	β_1	1	0.159	2.46	0.00739	$< \epsilon$	0.1132	0.9099
	β_2	1	0.591	2.46	0.00739	$< \epsilon$	-0.8688	0.3850
	β_3	1	0.641	2.45	0.00735	$< \epsilon$	0.3118	0.7552
P	p_{11}	1	0.155	0.577	0.00170	$< \epsilon$	-0.6194	0.5357
	p_{12}	1	0.293	0.211	0.00103	$< \epsilon$	0.1136	0.9096
	p_{13}	1	0.227	0.212	0.00104	$< \epsilon$	0.9090	0.3634
	p_{21}	1	0.185	0.212	0.00103	$< \epsilon$	-0.3313	0.7404
	p_{22}	1	0.283	0.577	0.00170	$< \epsilon$	-0.4018	0.6878
	p_{23}	1	0.285	0.212	0.00104	$< \epsilon$	0.9858	0.3242
	p_{31}	1	0.213	0.212	0.00103	$< \epsilon$	-0.2966	0.7668
	p_{32}	1	0.917	0.212	0.00103	$< \epsilon$	-0.9615	0.3363
	p_{33}	1	0.386	0.576	0.00169	$< \epsilon$	0.7705	0.4410

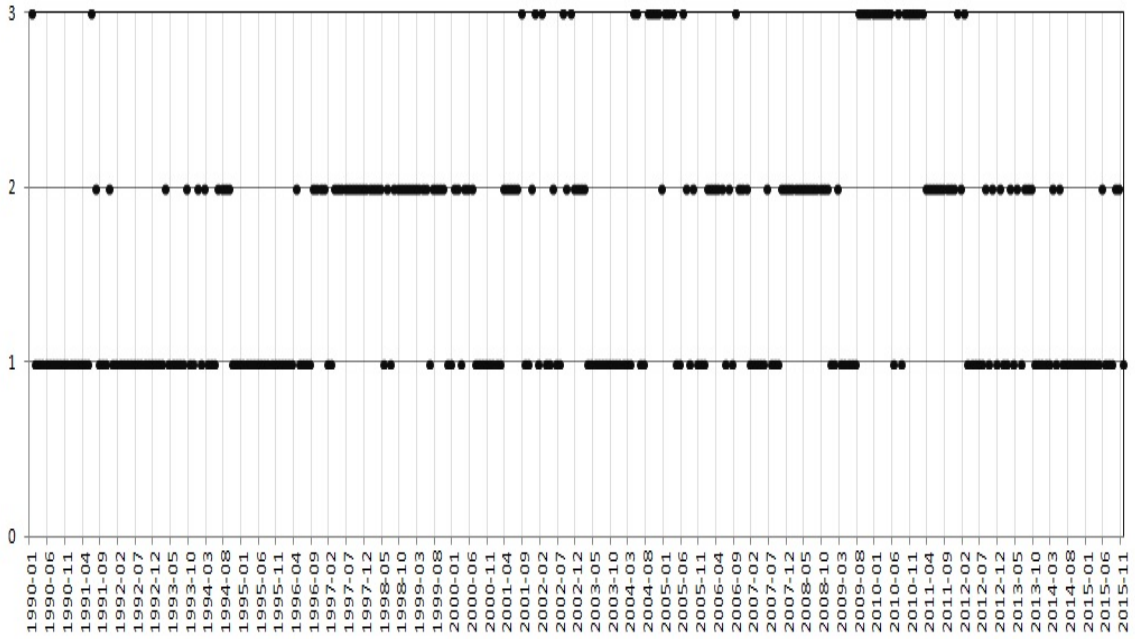
Çizelge 8.16.' da durağanlık testi sonuçlarına göre, 1071000 iterasyondan elde edilen örneklem, %95 güvenirlikle $f(\Theta | Y^{(T)})$ bileşik sonsal dağılımdan gelmektedir. Yarı genişlik testine göre, tüm parametrelerin sonsal ortalamalarının ϵ kesinliğinde tahmin edilmesinde, 1071000 iterasyondan oluşan zincir %95 güvenirlikle kabul edilebilir uzunluktadır. Geweke testine göre ise, %95 güvenirlikle tüm parametrelere ilişkin zincirlerin ilk %10' luk ve son %50' lik kısımları arasındaki fark istatistiksel olarak anlamlı değildir. Uygulanan iki yakınsaklık testine göre, ilk 129000 iterasyon sonucunda durağan dağılıma yakınsamanın sağlandığı %95 güvenirlikle söylenebilir. 3-durumlu Beta-SMM parametrelerinin Bayesci tahminleri, 129001-1200000-üncü iterasyonlardaki

örnek değerlerin sonsal tepe değerleri olarak elde edilir. Model parametrelerinin Bayesci tahminleri ve %95 en yüksek sonsal yoğunluk aralıklarının alt-üst sınırları aşağıda verilmiştir:

Çizelge 8.17. 3-Durumlu Beta-SMM Parametrelerinin Bayesci Tahminleri

Parametre	Sonsal Tepe Değeri	Alt Sınır	Üst Sınır
α	α_1	0.81281	0.02609 4.44103
	α_2	0.73627	0.03685 4.43987
	α_3	0.74269	0.05068 4.46714
β	β_1	1.21550	0.00069 5.84414
	β_2	1.12725	0.01630 5.87149
	β_3	1.06929	0.00942 5.85869
P	p_{11}	0.98342	0.04071 0.99974
	p_{12}	0.00922	$8.85557e - 09$ 0.75250
	p_{13}	0.00736	$2.52401e - 08$ 0.75418
	p_{21}	0.00910	$5.97459e - 08$ 0.75399
	p_{22}	0.98342	0.04101 0.99987
	p_{23}	0.00748	$3.23148e - 08$ 0.75398
	p_{31}	0.00907	$5.14285e - 08$ 0.75222
	p_{32}	0.00911	$8.31564e - 08$ 0.75414
	p_{33}	0.98182	0.04102 0.99985
	π	π_1	0.35399
π_2		0.35612	— —
π_3		0.28989	— —

Çizelge 8.17.' de beta dağılım ailesine ilişkin parametrelerin Bayesci tahminlerinin birbirlerine çok yakın olduğu görülmektedir. 3-durumlu Beta-SMM' de, X sürecinin bulunduğu durumlar ise aşağıdaki gibi tahmin edilmiştir:



Şekil 8.4. 1990-01 : 2015:11 Dönemlerinde X Saklı Sürecinin Bulunduğu Durumlar

Şekil 8.5.' te X sürecinin, birinci ve ikinci durumda daha çok bulunduğu görülmektedir. Hesaplanan sonsal tepe değerleri kullanılarak, modelin marjinal olabilirlik değeri tahmin edilmiştir. Burada, marjinal olabilirliğin paydasındaki $p(\mathbf{X}^{*(T)} | \mathbf{Y}^{(T)})$ olasılığı 500 iteryasyondan oluşan ikinci bir Gibbs örnekleme yapılarak tahmin edilmiştir. Bu işlem, 45.23 dakika sürmüştür. Marjinal olabilirliğin pay ve paydasında yer alan yoğunlukların logaritmaları alınarak aşağıda verilmiştir:

Çizelge 8.18. 3-durumlu Beta-SMM için Log-Marjinal Olabilirlik Değerinin Hesaplanması

PAY	PAYDA
$\ln f(\mathbf{y}^{(T)} \boldsymbol{\alpha}^{*(m)}, \boldsymbol{\beta}^{*(m)}, \mathbf{X}^{*(T)}) = 1.883767$	$\ln f(\boldsymbol{\alpha}^{*(m)}, \boldsymbol{\beta}^{*(m)} \mathbf{X}^{*(T)}, \mathbf{Y}^{(T)}) = -767.5184$
$\ln f(\boldsymbol{\alpha}^{*(m)}, \boldsymbol{\beta}^{*(m)}) = -23.11298$	$\ln f(\mathbf{P}^* \mathbf{X}^{*(T)}) = -222.0267$
$\ln p(\mathbf{X}^{*(T)} \mathbf{P}^*) = -482.2948$	$\ln \hat{p}(\mathbf{X}^{*(T)} \mathbf{Y}^{(T)}) = -255.5969$
$\ln f(\mathbf{P}^*) = 2.079442$	
TOPLAM_{PAY} = -501.4446	TOPLAM_{PAYDA} = -1245.142

Çizelge 8.18.' deki değerlere göre, 3-durumlu Beta-SMM için, $\ln f(\mathbf{y}^{(T)})$ log-marjinal olabilirlik değeri 743.6974 olarak tahmin edilmiştir.

$m = 4$ durumunda, ilk 10000 iterasyon atıldıktan sonra geriye kalan 1990000 iterasyon üzerinden yakınsaklık testleri yapılmış ve bu testlere ilişkin sonuçlar aşağıda verilmiştir:

Çizelge 8.19. $m = 4$ için Yakınsaklık Testleri (1990000 İterasyon Üzerinden)

				Heidelberg-Welch Testi $\alpha = 0.05$			Geweke Testi			
				Durağanlık Testi		Yarı Genişlik Testi ($\epsilon = 0.01$)	İlk %10 ve Son %50			
				Başlangıç İterasyonu	P-Değeri	Ortalama	Yarı Genişlik	Sonuç	Z-Skoru	P-Değeri
α	α_1	1	0.169	1.770	0.00381	$< \epsilon$	-1.1988	0.2306		
	α_2	1	0.823	1.770	0.00383	$< \epsilon$	1.0269	0.3045		
	α_3	1	0.860	1.770	0.00384	$< \epsilon$	0.7239	0.4691		
	α_4	1	0.197	1.770	0.00381	$< \epsilon$	-0.1552	0.8767		
β	β_1	1	0.643	2.300	0.00591	$< \epsilon$	0.9720	0.3311		
	β_2	1	0.428	2.300	0.00590	$< \epsilon$	-0.0099	0.9921		
	β_3	1	0.167	2.300	0.00592	$< \epsilon$	-1.4094	0.1587		
	β_4	1	0.524	2.300	0.00590	$< \epsilon$	-0.9114	0.3621		
P	p_{11}	1	0.648	0.443	0.00130	$< \epsilon$	0.2998	0.7643		
	p_{12}	1	0.843	0.186	0.00066	$< \epsilon$	-0.5708	0.5681		
	p_{13}	1	0.336	0.186	0.00065	$< \epsilon$	-0.0509	0.9594		
	p_{14}	1	0.671	0.186	0.00065	$< \epsilon$	0.0344	0.9726		
	p_{21}	1	0.196	0.186	0.00066	$< \epsilon$	0.9095	0.3631		
	p_{22}	1	0.540	0.441	0.00131	$< \epsilon$	-0.8815	0.3780		
	p_{23}	1	0.797	0.186	0.00066	$< \epsilon$	0.4707	0.6379		
	p_{24}	1	0.984	0.186	0.00065	$< \epsilon$	0.3714	0.7103		
	p_{31}	1	0.083	0.186	0.00065	$< \epsilon$	3.0660	0.0022		
	p_{32}	1	0.103	0.186	0.00065	$< \epsilon$	1.8250	0.0680		
	p_{33}	1	0.130	0.441	0.00130	$< \epsilon$	-2.9990	0.0027		
	p_{34}	1	0.793	0.186	0.00066	$< \epsilon$	0.9864	0.3239		
	p_{41}	597001	0.177	0.185	0.00078	$< \epsilon$	2.0820	0.0373		
	p_{42}	1	0.664	0.185	0.00065	$< \epsilon$	0.6573	0.5110		
	p_{43}	1	0.259	0.186	0.00066	$< \epsilon$	-0.1233	0.9019		
	p_{44}	597001	0.267	0.444	0.00157	$< \epsilon$	-1.3610	0.1735		

Çizelge 8.19.'daki durağanlık testinde, $\forall P > 0.05$ olmasına karşın, p_{41} ve p_{44} para-

metreleri için başlangıç iterasyonu 597001 olarak bulunmuştur. Bu durumda, bu parametrelere ilişkin zincirlerin ilk %30' luk kısmı olan 597000 adım analiz dışı bırakılmalıdır. p_{41} ve p_{44} parametreleri için kalan 1393000 iterasyondaki Gibbs çıktıları, %95 güvenirlikle $f(\Theta | Y^{(T)})$ bileşik sonsal dağılımdan gelmektedir. Yarı genişlik testine göre, tüm parametrelerde yarı genişliğin sonsal ortalamaya oranı ϵ değerinden küçük bulunmuştur. Bu durumda, p_{41} ve p_{44} dışındaki parametrelerin sonsal ortalamalarının ϵ kesinliğinde tahmin edilmesinde, 1990000 iterasyondan oluşan zincir kabul edilebilir uzunluktadır. p_{41} ve p_{44} için ise, 1393000 iterasyondan oluşan zincir kabul edilebilir uzunluktadır. Geweke testine göre, %95 güvenirlikle p_{31} , p_{33} ve p_{41} parametrelerine ilişkin zincirlerin ilk %10' luk ve son %50' lik kısımları arasındaki fark istatistiksel olarak anlamlı bulunmuştur. Bu durumda, hem durağanlık hem de Geweke testi sonuçlarına göre, parametrelere ilişkin zincirlerin ilk %30'luk kısmı atılarak, kalan 1393000 iterasyon üzerinden yakınsaklık testleri yeniden yapılmıştır. Ancak, ilk %30' luk kısmın atılmasına rağmen, Geweke testinde anlamlı farklılıklar bulunmuştur. Bu durumda, ilk %40' luk kısım olan 796000 iterasyon atılmıştır ve kalan 1194000 iterasyon üzerinden yakınsaklık testleri yeniden yapılmıştır. Sonuçlar aşağıda verilmiştir:

Çizelge 8.20. $m = 4$ için Yakınsaklık Testleri (1194000 İterasyon Üzerinden)

				Heidelberg-Welch Testi $\alpha = 0.05$			Geweke Testi		
				Durağanlık Testi		Yarı Genişlik Testi ($\epsilon = 0.01$)	İlk %10 ve Son %50		
				Başlangıç P-Değeri	Ortalama	Yarı Genişlik	Sonuç	Z-Skoru	P-Değeri
				İterasyonu					
α	α_1	1	0.222	1.770	0.00492	$< \epsilon$	-0.3629	0.7167	
	α_2	1	0.882	1.770	0.00492	$< \epsilon$	0.4871	0.6262	
	α_3	1	0.664	1.770	0.00494	$< \epsilon$	-0.1331	0.8941	
	α_4	1	0.759	1.770	0.00493	$< \epsilon$	1.1691	0.2424	
β	β_1	1	0.606	2.300	0.00762	$< \epsilon$	0.4046	0.6958	
	β_2	1	0.089	2.290	0.00764	$< \epsilon$	-0.5718	0.5675	
	β_3	1	0.340	2.300	0.00769	$< \epsilon$	-1.6354	0.1020	
	β_4	1	0.607	2.300	0.00764	$< \epsilon$	-1.5069	0.1318	
P	p_{11}	1	0.495	0.442	0.00169	$< \epsilon$	0.3315	0.7403	
	p_{12}	1	0.775	0.185	0.00084	$< \epsilon$	0.8100	0.4179	
	p_{13}	1	0.678	0.187	0.00084	$< \epsilon$	-1.1690	0.2424	
	p_{14}	1	0.584	0.186	0.00084	$< \epsilon$	-0.3006	0.7637	
	p_{21}	1	0.105	0.186	0.00085	$< \epsilon$	1.7180	0.0858	
	p_{22}	1	0.363	0.441	0.00168	$< \epsilon$	-0.9477	0.3433	
	p_{23}	1	0.626	0.186	0.00084	$< \epsilon$	-0.3383	0.7351	
	p_{24}	1	0.931	0.186	0.00084	$< \epsilon$	0.4673	0.6403	
	p_{31}	1	0.418	0.186	0.00084	$< \epsilon$	0.5312	0.5953	
	p_{32}	1	0.159	0.186	0.00084	$< \epsilon$	-0.5295	0.5965	
	p_{33}	1	0.540	0.442	0.00168	$< \epsilon$	-0.1088	0.9134	
	p_{34}	1	0.606	0.187	0.00084	$< \epsilon$	0.2105	0.8333	
	p_{41}	1	0.359	0.185	0.00084	$< \epsilon$	0.8314	0.4057	
	p_{42}	1	0.405	0.185	0.00084	$< \epsilon$	1.7700	0.0767	
	p_{43}	1	0.757	0.186	0.00085	$< \epsilon$	0.2583	0.7962	
	p_{44}	1	0.250	0.444	0.00169	$< \epsilon$	-1.4360	0.1510	

Çizelge 8.20.' de durağanlık testi sonuçlarına göre, 1194000 iterasyondan elde edilen örneklem, %95 güvenirlikle $f(\Theta | Y^{(T)})$ bileşik sonsal dağılımdan gelmektedir. Yarı genişlik testine göre, tüm parametrelerin sonsal ortalamalarının ϵ kesinliğinde tah-

min edilmesinde, 1194000 iterasyondan oluşan zincir %95 güvenirlikle kabul edilebilir uzunluktadır. Geweke testine göre ise, %95 güvenirlikle tüm parametrelere ilişkin zincirlerin ilk %10' luk ve son %50' lik kısımları arasındaki fark istatistiksel olarak anlamlı değildir. Uygulanan iki yakınsaklık testine göre, ilk 806000 iterasyon sonucunda durağan dağılıma yakınsamanın sağlandığı %95 güvenirlikle söylenebilir. 4-durumlu Beta-SMM parametrelerinin Bayesci tahminleri, 806001-2000000-uncu iterasyonlardaki örnek değerlerin sonsal tepe değerleri olarak elde edilir. Model parametrelerinin Bayesci tahminleri ve %95 en yüksek sonsal yoğunluk aralıklarının alt-üst sınırları aşağıda verilmiştir:

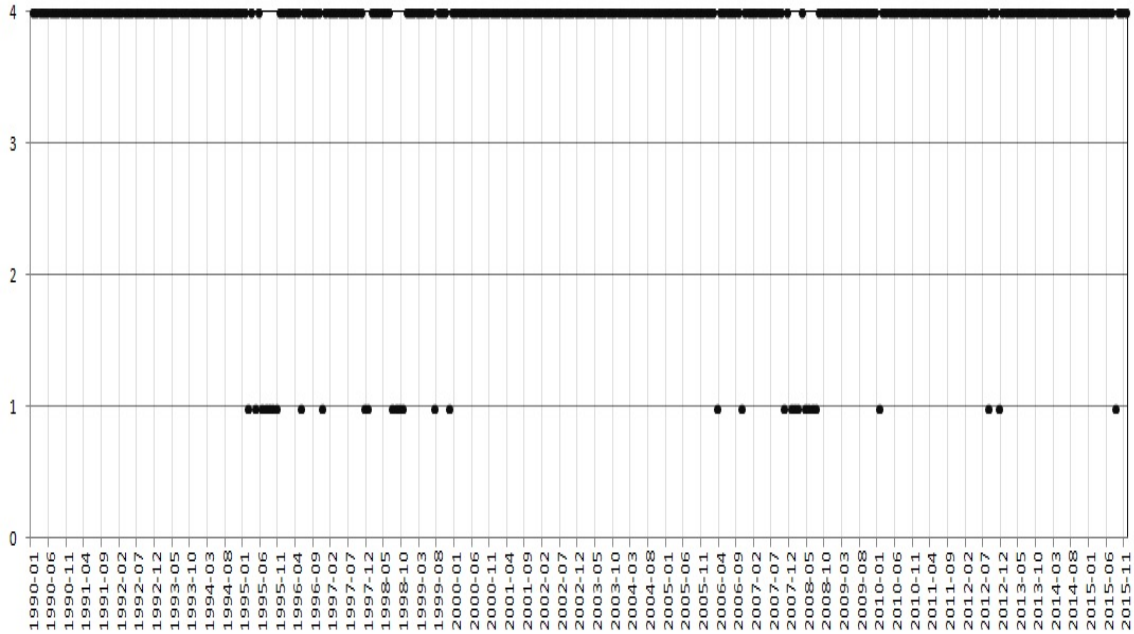


Çizelge 8.21. 4-Durumlu Beta-SMM Parametrelerinin Bayesci Tahminleri

Parametre	Sonsal Tepe Değeri	Alt Sınır	Üst Sınır
α	α_1	0.75761	0.05864 4.29202
	α_2	0.78710	0.07092 4.29262
	α_3	0.78489	0.07017 4.29914
	α_4	0.74524	0.06142 4.27428
β	β_1	0.99071	0.05160 5.64335
	β_2	0.96939	0.05117 5.63386
	β_3	1.00865	0.04465 5.64699
	β_4	1.16466	0.05080 5.64614
P	p_{11}	0.97896	0.02179 0.99647
	p_{12}	0.00806	$1.06289e - 08$ 0.64818
	p_{13}	0.00792	$2.03641e - 09$ 0.65104
	p_{14}	0.00505	$7.47210e - 09$ 0.64380
	p_{21}	0.00792	$1.27329e - 08$ 0.64111
	p_{22}	0.97871	0.02229 0.99682
	p_{23}	0.00800	$1.83609e - 09$ 0.64964
	p_{24}	0.00537	$3.06586e - 09$ 0.64940
	p_{31}	0.00805	$1.20240e - 08$ 0.65018
	p_{32}	0.00806	$3.17418e - 08$ 0.64815
	p_{33}	0.97882	0.02196 0.99660
	p_{34}	0.00507	$2.02457e - 08$ 0.65107
	p_{41}	0.00824	$2.40409e - 08$ 0.64835
	p_{42}	0.00805	$1.96071e - 10$ 0.65055
	p_{43}	0.00810	$1.45156e - 10$ 0.65113
	p_{44}	0.97561	0.02183 0.99649
π	π_1	0.27663	— —
	π_2	0.27455	— —
	π_3	0.27403	— —
	π_4	0.17480	— —

Çizelge 8.21.'de beta dağılım ailesine ilişkin parametrelerin Bayesci tahminlerinin birbirlerine çok yakın olduğu görülmektedir. Bu durum, diğer saklı durum sayılarına ilişkin

Beta-SMM' lere benzer bir sonuçtur. 4-durumlu Beta-SMM' de, X sürecinin bulunduğu durumlar ise aşağıdaki gibi tahmin edilmiştir:



Şekil 8.5. 1990-01 : 2015:11 Dönemlerinde X Saklı Sürecinin Bulunduğu Durumlar

Şekil 8.5.' te X saklı süreci, ikinci ve üçüncü durumda hiç bulunmamıştır. Genellikle dördüncü durumda bulunmuştur. Parametrelerin Bayesci tahminleri olan sonsal tepe değerleri kullanılarak, modelin marjinal olabilirlik değeri hesaplanmıştır. Marjinal olabilirliğin paydasındaki $p(\mathbf{X}^{*(T)} | \mathbf{Y}^{(T)})$ olasılığı ise, 500 iterasyondan oluşan ikinci bir Gibbs örnekleme yapılarak tahmin edilmiştir. Bu işlem, 58.67 dakika sürmüştür. Marjinal olabilirliğin pay ve paydasında yer alan yoğunlukların logaritmaları alınarak aşağıda verilmiştir:

Çizelge 8.22. 4-durumlu Beta-SMM için Log-Marjinal Olabilirlik Değerinin Hesaplanması

PAY	PAYDA
$\ln f(\mathbf{y}^{(T)} \boldsymbol{\alpha}^{*(m)}, \boldsymbol{\beta}^{*(m)}, \mathbf{X}^{*(T)}) = 9.026224$	$\ln f(\boldsymbol{\alpha}^{*(m)}, \boldsymbol{\beta}^{*(m)} \mathbf{X}^{*(T)}, \mathbf{Y}^{(T)}) = -764.9988$
$\ln f(\boldsymbol{\alpha}^{*(m)}, \boldsymbol{\beta}^{*(m)}) = -29.68383$	$\ln f(\mathbf{P}^* \mathbf{X}^{*(T)}) = -73.7568$
$\ln p(\mathbf{X}^{*(T)} \mathbf{P}^*) = -190.0389$	$\ln \hat{p}(\mathbf{X}^{*(T)} \mathbf{Y}^{(T)}) = -122.5617$
$\ln f(\mathbf{P}^*) = 7.167038$	
TOPLAM_{PAY} = -203.5294	TOPLAM_{PAYDA} = -961.3174

Çizelge 8.22.' deki değerlere göre, 4-durumlu Beta-SMM için, $\ln f(\mathbf{y}^{(T)})$ log-marjinal olabilirlik değeri 757.7879 olarak tahmin edilmiştir.

Her bir $m \in \{2, 3, 4\}$ saklı durum sayısına ilişkin Beta-SMM' ler için tahmin edilen log-marjinal olabilirlik değerleri kullanılarak, rakip modeller için Bayes faktörleri aşağıdaki gibi elde edilmiştir:

Çizelge 8.23. Rakip Modellerin Karşılaştırılması

	Bayes Faktörü Değeri
BF_{32}	$6.699e + 22$
BF_{43}	90671.64834

Bayes faktörleri çok büyük bulunmuştur. Çizelge 8.23.' teki BF_{32} değerine göre, 3-durumlu model 2-durumlu modele kıyasla veriler tarafından daha çok desteklenmektedir ($BF_{32} > 100$). 4-durumlu model ise, 3-durumlu modele kıyasla çok daha iyidir ($BF_{43} > 100$). Bu durumda, verilerin en iyi şekilde açıklanmasında 4-durumlu Beta-SMM tercih edilir. Bu model üzerinden işsizlik oranları için öngöründe bulunulur.

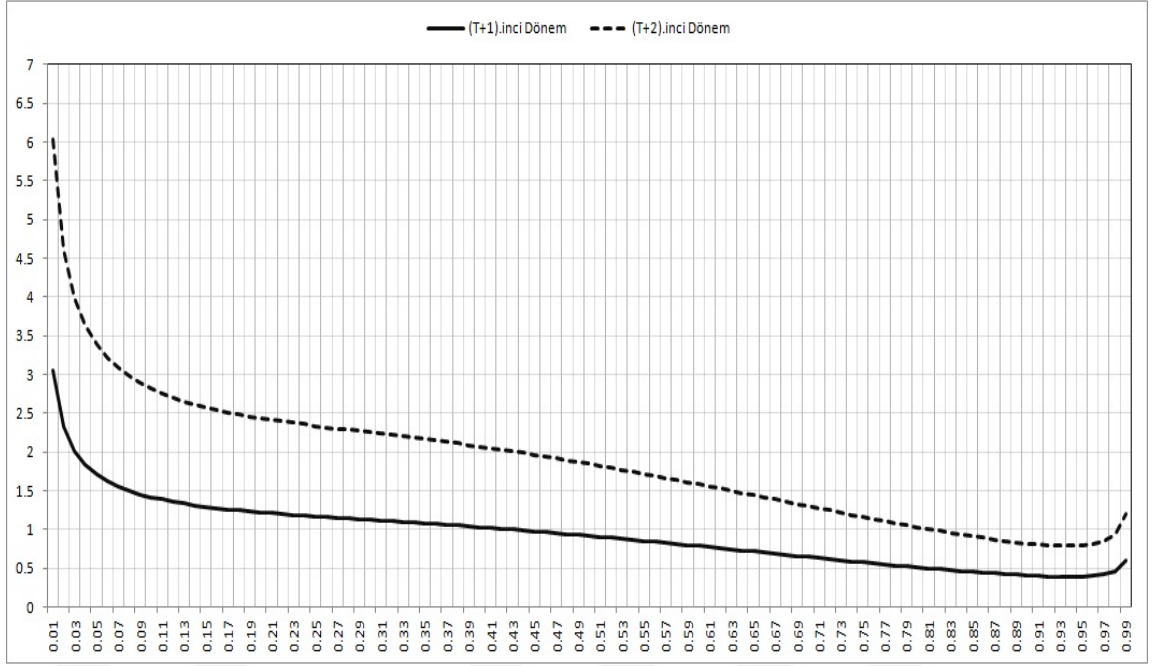
Burada önemli bir nokta, saklı durum sayısı arttıkça log-marjinal olabilirlik değeri artmaktadır. Bu durumda, 5-durumlu Beta-SMM' nin Bayesci çözümlemesi yapılarak bu model için de $\ln f(Y^{(T)})$ değeri tahmin edilmelidir. Ancak, bu modelin Bayesci çözümlemesi mevcut bilgisayar imkanlarını zorlamaktadır ve depolama alanı sıkıntısı nedeniyle sonuç alınamamaktadır. Bu nedenle, bu model için Bayesci çözümleme yapılamamıştır.

4-durumlu Beta-SMM kullanılarak, gelecek iki dönem için işsizlik oranlarının olasılık yoğunluk değerleri 1194000 iterasyondan elde edilen Gibbs çıktıları üzerinden Monte Carlo yöntemi kullanılarak tahmin edilmiştir. Öngörü dağılımlarının elde edilmesinde geçen hesaplanma süreleri aşağıda verilmiştir:

Çizelge 8.24. Öngörü Dağılımlarının Elde Edilmesi için Geçen Süreler

	Hesaplanma Süresi (Saat)
Y_{T+1}	5.76
Y_{T+2}	7.00

Çizelge 8.24.'te hesaplanma sürelerinin fazla olduğu görülmektedir. Bu nedenle, ikiden fazla dönem için öngörü dağılımları elde edilmemiştir. 2015:12 ve 2016:01 dönemleri için işsizlik oranı öngörü dağılımları aşağıdaki gibi elde edilmiştir.



Şekil 8.6. Y_{T+1} ve Y_{T+2} için Öngörü Dağılımları

Şekil 8.6.' da verilen öngörü dağılımlarına göre, yüksek yoğunluklu bölgeler işsizlik oranlarının düşük olduğu yerlerdir. Özellikle, %5 ve daha düşük olan değerler en yüksek yoğunluğa sahiptir. Bu durumda, 2015 yılı Aralık ve 2016 yılı Ocak aylarında ABD' de işsizlik oranının %5' in altında olması olasılığı daha yüksektir denilebilir. **Şekil 8.6.' da elde edilen öngörü yoğunluğu grafiğine uygun olarak, ABD işsizlik oranı 2015 yılı Aralık ayında %5 ve 2016 yılı Ocak ayında %4.9 olmuştur [104]. Dolayısıyla, Beta-SMM ile yapılan öngörü iyi ve tutarlı bir sonuç vermiştir.**

9. SONUÇ VE TARTIŞMA

Tez çalışmasında, rasgele değişen ve gelişen, doğrudan gözlemlenemeyen, sadece başka bir sürecin çıktıları aracılığıyla dolaylı olarak gözlemlenebilen ve durağan olmayan sistemlerin stokastik modellenmesinde kullanılan saklı Markov modelinin teorik yapısı ve işleyişi farklı dağılım aileleri varsayımı altında detaylı olarak incelenmiştir. Çalışmada, iki farklı dağılım ailesi belirlenmiştir. Bunlar, Poisson ve beta dağılım aileleridir. Poisson dağılımı, kesikli yapıda olan verilerin stokastik modellenmesinde temel bir dağılım olması nedeniyle seçilmiştir. Beta dağılımı ise, SMM' de beta dağılım ailesi şekil parametrelerinin tahmininde yaşanan soruna çözüm getirilmesi amacıyla ele alınmıştır. Bu doğrultuda, Poisson ve beta dağılımı varsayımlarının SMM üzerindeki etkileri tartışılmıştır. Model üzerinden yapılan istatistiksel tahmin ve çıkarsamalar detaylı olarak açıklanmıştır. Poisson ve beta dağılım aileleri varsayımı altında SMM' nin gerçek veri setleri üzerinden işletilmesi amacıyla, R programında ilgili kodlar yazılmıştır.

Poisson dağılım ailesi varsayımı altında SMM' nin tüm parametreleri klasik yaklaşım altında analitik olarak elde edilebilmektedir. Uygulamada veri sıkıntısı da yaşanmaması nedeniyle, klasik yaklaşım altında Poisson-SMM ele alınmıştır. Belirli bir zaman aralığında gerçekleşen olaylara ilişkin frekansların stokastik modellenmesinde geleneksel bir yaklaşım olarak Poisson süreçleri tercih edilmektedir. Bu nedenle, tez çalışmasında Poisson-SMM' nin gerçek bir veri seti üzerinden işletiminde rakip model olarak Poisson süreci seçilmiştir. Aşırı yayılım gösteren ve serisel bağımlılık içeren frekans verilerinin stokastik modellenmesinde bu iki modelin başarısı karşılaştırılmıştır. Burada, Poisson-SMM' nin işletiminde Türkiye deprem verisi kullanılmıştır. Poisson-SMM' nin Türkiye deprem verisine ilk kez uygulanması Can vd. [82] tarafından gerçekleştirilmiştir. Bu çalışma, aynı zamanda modelin Türkiye verisine ilk kez uygulanmasıdır. Uygulamada, 1900 ve 2012 yılları arasında Bilecik' in merkez olduğu 100 km yarıçapındaki bir alanda meydana gelmiş ve büyüklüğü 4 ve daha fazla olan depremlerin yıllık frekansları modellenmiştir. Gelecek 35 yıllık bir periyod için, her yıl 4 ve daha fazla büyüklükte oluşacak olan depremlerin frekansları tahmin edilerek çalışma sahası için deprem riski belirlenmiştir. Bilgi kriteri değerleri ve olabilirlik oran testlerinden elde edilen sonuçlara göre, en iyi model olarak 3-durumlu Poisson-SMM belirlenmiştir. 3-durumlu Poisson-SMM ve Poisson sürecinin gözlemleri açıklamadaki başarıları karşılaştırıldığında, Poisson sürecinin çok yetersiz kaldığı ve Poisson-SMM' nin oldukça başarılı olduğu sonucuna

varılmıştır. Bu durumun iki nedeni vardır. Bunlardan biri, Poisson sürecinin belleksizlik özelliği taşımasıdır. Belleksizlik özelliğine sahip bir süreç, serisel bağımlılık içeren verilerin zaman içindeki değişimlerinin açıklanmasında yetersiz kalmaktadır. Diğer ise, Poisson dağılımının ortalama ve varyans değerlerinin birbirine eşit olmasıdır. Poisson dağılımının bu özelliği, aşırı yayılım gösteren verilerin modellenmesinde uygun değildir. 3-durumlu Poisson-SMM' ye göre, çalışma sahasında her yıl ortalama iki tane depremin gerçekleşeceği tahmin edilmiştir. Bu tahmine göre, çalışma sahasında deprem riskinin düşük olduğu sonucuna varılır. Daha detaylı bir bilgi verilirse, gelecek 35 yıl için her yıl %96 ile %97 arasında olasılıkla 0-5 tane deprem olması beklenmektedir. Ayrıca, iki modelde çalışma sahası için deprem riskleri yıllara göre birikimli olarak (1-6 yıl içinde) tahmin edilmiştir. Tahminlere göre, 3-durumlu Poisson-SMM' ye kıyasla, Poisson süreci büyük frekans değerlerine daha yüksek; küçük frekans değerlerine ise, daha düşük olasılık vermiştir. Bunun sonucu olarak, Poisson süreci deprem riskinin daha fazla olduğunu öngörmekte ve aşırı uyum sorunu oluşturmaktadır.

Poisson dağılım ailesinin aksine, beta dağılım ailesi varsayımı altında SMM' nin parametreleri klasik yaklaşım altında tahmin edildiğinde önemli bir sorunla karşılaşmaktadır. EM algoritmasının en büyükleme adımında beta dağılım ailesinin şekil parametrelerine ilişkin en çok olabilirlik denklemleri digamma fonksiyonlarını içermektedir. Bu nedenle, şekil parametrelerinin en çok olabilirlik tahminleri analitik olarak elde edilemez. Sayısal yöntemler kullanılarak, şekil parametrelerinin yaklaşık değerlerine ulaşılabilir. Fakat, bu yöntemlerde başlangıç değerlerine oldukça duyarlıdır. Başlangıç değerlerinin etkileri, elde edilecek tahminlerin ve sonrasında yapılacak çıkarsamaların doğruluğunu ve kesinliğini ciddi oranda olumsuz etkileyebilir. Belirtilen sorunların aşılması amacıyla, tez çalışmasında Beta-SMM' nin Bayesci çözümlenmesi önerilmiştir. Bu bağlamda, Beta-SMM' ye Bayesci yaklaşımın nasıl adapte edileceği detaylı olarak açıklanmıştır. Bu tez çalışması, beta dağılım ailesi varsayımı altında SMM' nin ilk Bayesci çözümlemesidir. Modeldeki tüm parametreler için eşlenik önsel dağılımlar kullanılmıştır. Bu sayede, sonsal dağılımlar kapalı formda kolaylıkla elde edilebilmiştir. Özellikle, beta dağılım ailesinin şekil parametrelerinin sonsal dağılımı için Bayyari [89]' nin önerdiği bileşik eşlenik önsel dağılımdan yararlanılmıştır. Beta-SMM' de model parametrelerinin bileşik sonsal dağılımı çok boyutlu ve karmaşık bir yapıdadır. Bu nedenle, bileşik sonsal dağılımdan doğrudan örnekleme yapılamamaktadır. Fakat, SMM' nin teorik yapısı Gibbs örneklemesinin kullanımına uygundur. Modelin bu özelliğinden

yola çıkılarak, tez çalışmasında Gibbs örnekleme kullanılarak model parametrelerinin tahmin edilmesi uygun görülmüştür. Ayrıca, beta dağılımının şekil parametreleri için Gibbs örnekleme içerisinde Metropolis-Hastings algoritmasına ihtiyaç duyulmuştur. Modeldeki her bir parametre için, Gibbs örneklemesinde oluşturulan zincir durağan dağılıma yakınsadıktan sonraki zincir değerleri kullanılarak ilgili parametre tahminlerine ulaşılır. Bu durağan örneklem üzerinden istenilen tüm istatistiksel çıkarımlara gidilir. Bayesci yaklaşım altında Beta-SMM' nin optimal boyut sayısının belirlenmesinde ise, Bayes faktörü kullanılmıştır. Çünkü, SMM' de optimal boyut sayısının belirlenmesi bir model seçim problemidir ve Bayesci yaklaşım altında model seçiminde Bayes faktörü kullanılır. Bayes faktöründe yer alan marjinal olabilirlik değerleri ise, Chib [88]' in önerdiği yöntem kullanılarak tahmin edilmiştir. Tez çalışmasında, Bayesci yaklaşım altında Beta-SMM' nin işletiminde ise, ABD' nin 1990-01 : 2015-11 dönemlerindeki işsizlik oranları kullanılmıştır. Beta-SMM' de model parametrelerinin Bayesci tahminlerinin elde edilmesi amacıyla gerçekleştirilen Gibbs örneklemesinde toplam iterasyon sayısı saklı durum sayısına bağlı olarak değişmiştir. Saklı durum sayısına bağlı olarak artan toplam parametre sayısı, Gibbs örneklemesinde oluşturulan zincirlerin durağan dağılımına yakınsamasını yavaşlatmıştır. Dolayısıyla, saklı durum sayısı arttıkça Gibbs örneklemesinde gerçekleştirilen toplam iterasyon sayısı da ciddi oranda artmıştır. Bilgisayardaki hesaplama zamanının çok fazla olması ve bilgisayarın depolama alanı kısıtı nedeniyle, saklı durum sayısı 5 ve daha fazla olduğu Beta-SMM' ler için Bayesci çözümleme yapılamamıştır. Saklı durum sayısı 2, 3 ve 4 olduğu Beta-SMM' ler için model parametrelerinin Bayesci tahminleri birbirine çok yakın bulunmuştur. Bu durumun iki nedeni olduğu düşünülmektedir. Bunlardan biri, işsizlik oranının varyans değerinin çok küçük olmasıdır. Diğeri ise, Metropolis-Hastings algoritmasında üretilen şekil parametrelerinin başlangıç değerlerinden çok fazla etkilenmesi ve bunun sonucu olarak zayıf bir zincir oluşmasıdır. Chib [88]' in yöntemine göre tahmin edilen log-marjinal olabilirlik değerleri saklı durum sayısı arttıkça büyümüştür. Bu nedenle, Bayes faktörleri tahmin edildiğinde 3-durumlu modelin 2-durumlu modele kıyasla; 4-durumlu modelin ise, 3-durumlu modele kıyasla verilerin açıklanmasında çok daha iyi olduğu sonucuna varılmıştır. En iyi model olarak 4-durumlu Beta-SMM belirlenmiştir. 4-durumlu Beta-SMM üzerinden 2015 yılı Aralık ayı ve 2016 yılı Ocak ayı için öngörü dağılımları elde edilmiştir. Öngörü dağılımlarına göre, gelecek iki dönem için %5' in altında bir işsizlik oranı yüksek olasılıkla beklenmektedir. İşsizlik oranının gerçek değerleri ise, 2015 yılı Aralık ayında %5 ve 2016 yılı Ocak ayında %4.9 olmuştur. Dolayısıyla, Beta-SMM' nin Bayesci çözümlenmesinden

elde edilen öngörü dağılımı gerçek değerlerle tutarlı olduğu söylenebilir.



KAYNAKLAR

- [1] Baum, L.E., Petrie, T., Statistical Inference for Probabilistic Functions of Finite State Markov Chains, *The Annals of Mathematical Statistics*, 37, 1554-1563, **1966**.
- [2] Baker, J.K., The Dragon System-An Overview, *IEEE Transactions Acoustics, Speech and Signal Processing*, Vol. ASSP-23, 24-29, **1975**.
- [3] Jelinek, F., Bahl, L.R., Mercer, R.L., Design of a Linguistic Statistical Decoder for Recognition of Continuous Speech, *IEEE Transactions on Information Theory*, 21(3), 250-256, **1975**.
- [4] Bahl, L.R., Jelinek, F., Decoding for Channels with Insertions, Deletions and Substitutions with Applications to Speech Recognition, *IEEE Transactions on Information Theory*, 21(4), 404-411, **1975**.
- [5] Jelinek, F., Continuous Speech Recognition by Statistical Methods, *Proceedings of the IEEE*, 64(4), 532-556, **1976**.
- [6] Rabiner, L.R., A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, *Proceedings of the IEEE*, 77(2), 257-285, **1989**.
- [7] Rabiner, L.R., Juang B.-H., *Fundamentals of Speech Recognition*, Prentice-Hall, **1993**.
- [8] Jelinek, F., *Statistical Methods for Speech Recognition*, MIT Pres, **1997**.
- [9] MacDonald, I.L., Zucchini, W., *Hidden Markov Models and Other Models for Discrete-Valued Time Series*, Chapman & Hall/CRC Monographs on Statistics & Applied Probability, **1997**.
- [10] Hamilton, J.D., *Time Series Analysis*, Princeton University Press, **1994**.
- [11] Rydén, T., Teräsvirta, Asbrink, S., Stylized facts of daily returns series and the Hidden Markov Model, *Journal of Applied Econometrics*, 13, 217-244, **1998**.
- [12] Kim, C.-J., Nelson, C.R., *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*, MIT Press, **1999**.
- [13] Pievatolo, A., Ruggeri, F., Soyer, R., A Bayesian Hidden Markov Model for Imperfect Debugging, *Reability Engineering and System Safety*, 103, 11-21, **2012**.
- [14] Landon, J., Özekici, S., Soyer, R., A Markov Modulated Poisson Model for Software Reability, *European Journal of Operational Research*, 229(2), 404-410, **2013**.
- [15] Dainotti, A., Pescapé, A., Rossi, P. S., Palmieri, F., Ventre, G. , Internet Traffic Modelling by Means of Hidden Markov Models, *Computer Networks*, 52(14), 2645-2662, **2008**.

- [16] Leroux, B.G., Puterman, M.L., Maximum-Penalized-Likelihood Estimation for Independent and Markov-Dependent Mixture Models, *Biometrics*, 48(2), 545-558, **1992**.
- [17] Cave, R.L., Neuwirth, L.P., Hidden Markov Models for English, *Proceedings of the Symposium on the Application of Hidden Markov Models to Text and Speech*, (Editors: Ferguson, J.D.), Princeton, 16-56, **1980**.
- [18] Charniak, E., *Statistical Language Learning*, MIT Press, **1993**.
- [19] Zucchini, W., Guttorp, P., A Hidden Markov Model for Space-Time Precipitation, *Water Resources Research*, 27(8), 1917-1923, **1991**.
- [20] Sansom, J., A Hidden Markov Model for Rainfall Using Breakpoint Data, *Journal of Climate*, 11(1), 42-53, **1998**.
- [21] Bellone, E., Hughes, J. P., Guttorp P., A Hidden Markov Model for Downscaling Synoptic Atmospheric Patterns to Precipitation Amounts, *Climate Research*, 15, 1-12, **2000**.
- [22] Robertson, A.W., Kirshner, S., Smyth, P., Downscaling of Daily Rainfall Occurrence over Northeast Brazil Using a Hidden Markov Model, *Journal of Climate*, 17, 4407-4424, **2004**.
- [23] Baldi, P., Hunkapiller, T., Chauvin, Y., McClure, M.A., Hidden Markov Models of Biological Primary Sequence Information, *Proceedings of the National Academy of Sciences of the United States of America*, 91(3), 1059-1063, **1994**.
- [24] Krogh, A., Brown, M., Mian, I. S., Sjölander, K., Haussler, D., Hidden Markov Models in Computational Biology: Applications to Protein Modeling, *Journal of Molecular Biology*, 235, 1501-1531, **1994**.
- [25] Durbin, R., Eddy, S., Krogh, A., Mitchison, G., *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids*, Cambridge University Press, **1998**.
- [26] Koski, T., *Hidden Markov Models for Bioinformatics*, Kluwer Academic Publishers, **2001**.
- [27] Fredkin, D. R., Rice, J.A., Bayesian Restoration of Single-channel Patch Clamp Recordings, *Biometrics*, 48(2), 427-448, **1992**.
- [28] Fredkin, D. R., Rice, J.A., Maximum Likelihood Estimation and Identification Directly from Single-Channel Recordings, *Proceedings: Biological Science*, The Royal Society, 249(1325), 125-132, **1992**.
- [29] Ball, F.G., Cai, Y., Kadane, J.B. and O'Hagan, A., Bayesian Inference for Ion-Channel Gating Mechanisms Directly from Single-Channel Recordings, Using Markov Chain Monte Carlo, *Proceedings: Mathematical, Physical and Engineering Sciences*, The Royal Society, 455(1988), 2879-2932, **1999**.

- [30] Wong, K.-C., Chan, T.-M., Peng, C., Li, Y., Zhang, Z., DNA Motif Elucidation Using Belief Propagation, *Nucleic Acids Research*, 41(16), e153, **2013**.
- [31] Stigler, J., Ziegler, F., Gieseke, A., Gebhardt, J.C.M., Rief, M., The Complex Folding Network of Single Calmodulin Molecules, *Science*, 334(6055), 512-516, **2011**.
- [32] Wong, W., Stamp, M., Hunting for Metamorphic Engines, *Journal of Computer Virology and Hacking Techniques*, 2(3), 211-229, **2006**.
- [33] Robert, C., Caselle, G., *Monte Carlo Statistical Methods*, Springer, Second Edition, **2004**.
- [34] Gill, J., *Bayesian Methods: A Social and Behavioral Sciences Approach*, Chapman & Hall/CRC Statistics in the Social Behavioral Sciences Series, Third Edition, **2014**.
- [35] Carlin, B.P., Louis, T. A., *Bayesian Methods for Data Analysis*, Chapman & Hall/CRC Texts in Statistical Science, Third Edition, **2008**.
- [36] Gelman, A., Carlin J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B., *Bayesian Data Analysis*, Chapman & Hall/CRC Texts in Statistical Science, Third Edition, **2014**.
- [37] Box, G.E.P., Tiao, G.C., *Bayesian Inference in Statistical Analysis*, Wiley, First Edition, **1992**.
- [38] Jeffreys, H., *The Theory of Probability*, Oxford, Third Edition, **1961**.
- [39] Robert, C.P., Casella G., *Introducing Monte Carlo Methods with R (Use R!)*, Springer-Verlag, **2010**.
- [40] Berger, J. O., *Statistical Decision Theory and Bayesian Analysis*, Springer Verlag, Second Edition, **1993**.
- [41] Markov, A.A., Rasprostranenia Zakona Bol'shih Chisel na Velichiny, Zavisyaschie Drug ot Druga (Extension of Law of Large Numbers to Dependent Quantities), *Izvestiya Fiziko-Matematicheskogo Obschestva Pri Kazanskom Universitete*, 15(2), 135-156, **1906**.
- [42] Markov, A.A., An Example of Statistical Investigation of the Text Eugene Onegin Concerning the Connection of Samples in Chains (in Russian), *Bulletin of the Imperial Academy of Sciences of St. Petersburg*, 7(3), 153-162, **1913**.
- [43] Ibe, O.C., *Markov Process for Stochastic Modeling*, Elsevier Academic Press, **2009**.
- [44] Pinsky, M.A., Karlin, S., *An Introduction to Stochastic Modeling*, Academic Press, Fourth Edition, **2011**.
- [45] Papoulis, A., Pillai, S.U., *Probability, Random Variables, Stochastic Processes*, McGraw-Hill Europe, Fourth Edition, **2002**.

- [46] Robert, C.P., *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*, Springer Verlag, Second Edition, **2007**.
- [47] Metropolis, N., Ulam, S., The Monte Carlo Method, *Journal of the American Statistical Association*, 44(247), 335-341, **1949**.
- [48] Von Neumann, J., Various Techniques Used in Connection with Random digits, *Journal Resources of The National Bureau of Standards - Applied Mathematics Series*, 12, 36-38, **1951**.
- [49] Sobol, I.M., *A Primer for the Monte Carlo Method*, CRC Press, **1994**.
- [50] Gosh, J.K., Delampady, M., Samanta, T., *An Introduction to Bayesian Analysis: Theory and Methods*, Springer Texts in Statistics, **2009**.
- [51] Givens, G.H., Hoeting, J.A., *Computational Statistics*, Wiley Series in Computational Statistics, Second Edition, **2013**.
- [52] Albert, J., *Bayesian Computation with R (Use R!)*, Springer, Second Edition, **2009**.
- [53] Bernardo, J.M., Smith, A.F.M., *Bayesian Theory*, Wiley, **2000**.
- [54] Cowles, M.K. and Carlin, B.P., Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review, *Journal of the American Statistical Association*, 91(434), 883-934, **1996**.
- [55] Geweke, J., Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments, *Bayesian Statistics 4*, (Editors: Bernardo, J.M., Berger, J.O., Dawid, A.P., Smith, A.F.M.), Oxford University Press, 169-193, **1992**.
- [56] Heidelberger, P., Welch, P.D., Simulation Run Length Control in the Presence of an Initial Transient, *Operations Research*, 31, 1109-1144, **1983**.
- [57] Brooks, S.P., Roberts, G.O., Convergence Assessment Techniques for Markov Chain Monte Carlo, *Statistics and Computing*, 8, 319-335, **1998**.
- [58] Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A., Teller, E., Equations of State Calculations by Fast Computing Machines, *Journal of Chemical Physics*, 21(6), 1087-1091, **1953**.
- [59] Hastings, W.K., Monte Carlo Sampling Methods Using Markov Chains and Their Applications, *Biometrika*, 57, 97-109, **1970**.
- [60] Chib, S., Greenberg, E., Understanding the Metropolis-Hastings Algorithm, *The American Statistician*, 49(4), 327-335, **1995**.
- [61] Gamerman, D., Lopes, H.F., *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, Chapman & Hall/CRC Texts in Statistical Science, Second Edition, **2006**.
- [62] Tierney, L., Markov Chains for Exploring Posterior Distributions, *The Annals of Statistics*, 22(4), 1701-1762, **1994**.

- [63] Liu, J.S., *Monte Carlo Strategies in Scientific Computing*, Springer Series in Statistics, **2008**.
- [64] Geman, S., Geman, D., Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, 721-741, **1984**.
- [65] Peskun, P., Optimum Monte Carlo Sampling Using Markov Chains, *Biometrika*, 60, 607-612, **1973**.
- [66] Gelfand, A.E., Smith, A.F.M., Sampling Based Approaches to Calculating Marginal Densities, *Journal of the American Statistical Association*, 85, 398-409, **1990**.
- [67] Casella, G., George, E.I., Explaining the Gibbs Sampling, *Journal of the American Statistical Association*, 46(3), 167-174, **1992**.
- [68] Brooks, S.P., Markov Chain Monte Carlo Method and its Applications, *The Statistician*, 47(1), 69-100, **1998**.
- [69] Zelterman, D., *Applied Multivariate Statistics with R*, Springer Series in Statistics for Biology and Health, **2015**.
- [70] Baum, L.E., Egon, J.A., An Inequality with Applications to Statistical Estimation for Probabilistic Functions of a Markov Process and to a Model for Ecology, *Bulletin of the American Mathematical Society*, 73, 360-363, **1967**.
- [71] Baum, L.E., Sell, G.R., Growth Functions for Transformations on Manifolds, *Pacific Journal of Mathematics*, 27(2), 211-227, **1968**.
- [72] Baum, L.E., Petrie, T., Soules, G., Weiss, N., A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains, *The Annals of Mathematical Statistics*, 41(1), 164-171, **1970**.
- [73] Baum, L.E., An Inequality and Associated Maximization Technique in Statistical Estimation for Probabilistic Functions of Markov Processes, *Inequalities*, 3, 1-8, **1972**.
- [74] Ephraim, Y., Merhav, N., Hidden Markov Processes, *IEEE Transactions on Information Theory*, 48(6), 1518-1569, **2002**.
- [75] Cappe, O., Moulines, E., Ryden, T., *Inference in Hidden Markov Models*, Springer Science+Business Media, **2005**.
- [76] Zucchini, W., MacDonald, I.L., *Hidden Markov Models for Time Series: An Introduction Using R*, Chapman&Hall/CRC Monographs on Statistics & Applied Probability, **2009**.
- [77] Rabiner, L.R., Juang, B.H., An Introduction to Hidden Markov Models, *IEEE ASSP Magazine*, 3(1), 4-16, **1986**.

- [78] Ghahramani, Z., An Introduction to Hidden Markov Models and Bayesian Networks, *International Journal of Pattern Recognition and Artificial Intelligence*, 15(1), 9-42, **2001**.
- [79] Alpaydin, E., *Introduction to Machine Learning*, The MIT Press, Second Edition, **2010**.
- [80] Viterbi, A.J., Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm, *IEEE Transactions on Information Theory*, 13(2), 260-269, **1967**.
- [81] Dempster, A.P., Laird, N.M., Rubin, D.B., Maximum Likelihood for Incomplete Data via the EM Algorithm, *Journal of the Royal Statistical Society, Series B (Methodological)*, 39(1), 1-38, **1977**.
- [82] Can, C.E., Ergün, G., Gökçeoğlu, C., Prediction of Earthquake Hazard by Hidden Markov Model (around Bilecik, NW Turkey), *Central European Journal of Geosciences*, 6(3), 403-414, **2014**.
- [83] Can, C.E. and Ergün, G., Risk Assessment of the Stock Exchange National 100 Index by Hidden Markov Model, *9th International Statistics Day Symposium (ISDS 2014) Programme and Submitted Abstracts Book*, (Editor: Aşır Genç), 10 May - 14 May, Antalya, Turkey, 30, **2014**.
- [84] Orfanogiannaki, K., Karlis, D., Papadopoulos, G.A., Identifying Seismicity Levels via Poisson Hidden Markov Models, *Pure and Applied Geophysics*, 167(8), 919-931, **2010**.
- [85] Johnson, N.L., Kotz, S., *Continuous Univariate Distributions Volume 2*, Houghton Mifflin Company, **1970**.
- [86] Johnson, N.L., Kotz, S., Balakrishnan, N., *Continuous Univariate Distributions Volume 2*, Wiley Series in Probability and Statistics, Second Edition, **1995**.
- [87] Nguyen, T.T., Maximum Likelihood Estimators of the Parameters in a Beta Distribution, *Handbook of Beta Distribution and Its Applications*, (Editors: Gupta, A.K., Nadarajah, S.), CRC Press, 229-236, **2004**.
- [88] Chib, S., Marginal Likelihood From the Gibbs Output, *Journal of the American Statistical Association*, 90(432), 1313-1321, **1995**.
- [89] Bayyari, M.J., Bayesian Inference on the Parameters of a Beta Distribution, *Statistics & Decisions*, Supplement Issue No. 2, 17-21, **1985**.
- [90] Kass, R.E., Raftery, A.E., Bayes Factors, *Journal of the American Statistical Association*, 90(430), 773-795, **1995**.
- [91] Raftery, A.E., Bayesian Model Selection in Social Research, *Sociological Methodology*, 25, 111-163, **1995**.
- [92] Jeffreys, H., Some Tests of Significance, Treated by the Theory of Probability, *Proceedings of the Cambridge Philosophy Society*, 31, 203-222, **1935**.

- [93] Schwarz, G., Estimating the Dimension of a Model, *The Annals of Statistics*, 6(2), 461-464, **1978**.
- [94] Raftery, A.E., Choosing Models for Cross-Classifications, *American Sociological Review*, 51(1), 145-146, **1968**.
- [95] Kass, R.E., Raftery, A.E., *Bayes Factors and Model Uncertainty*, Technical Report No: 254, Department of Statistics, GN-22, University of Washington, Seattle, Washington, 98195, USA, **1993**.
- [96] Rozenkranz, S.L., Raftery, A.E., *Covariate Selection in Hierarchical Models of Hospital Admission Counts: A Bayes Factor Approach*, Technical Report No: 268, Department of Statistics, GN-22, University of Washington, Seattle, Washington, 98195, USA, **1994**.
- [97] Kass, R.E., Tierney L. and Kadane, J.B., The Validity of Posterior Expansions Based on Laplace's Method, *Bayesian and Likelihood Methods in Statistics and Econometrics: Essays in Honor of George Barnard*, (Editors: Geisser, S., Hodges, J.S., Press, S.J., Zellner, A.), Elsevier Science Publishers B.V. (North-Holland), 473-488, **1990**.
- [98] Gelfand, A.E., Model Determination Using Sampling-Based Methods, *Markov Chain Monte Carlo in Practice*, (Editors: Gilks, W.R., Richardson, S., Spiegelhalter, D.J.), Chapman & Hall, 146-161, **1996**.
- [99] Geweke, J., Bayesian Inference in Econometric Models Using Monte Carlo Integration, *Econometrics*, 57(6), 1317-1339, **1989**.
- [100] Newton, M.A., Raftery, A.E., Approximate Bayesian Inference with the Weighted Likelihood Bootstrap, *Journal of the Royal Statistical Society, Series B*, 56(1), 3-48, **1994**.
- [101] Gelfand, A.E., Dey, D.K., Bayesian Model Choice: Asymptotics and Exact Calculations, *Journal of the Royal Statistical Society, Series B*, 56(3), 501-514, **1994**.
- [102] Hock, M., Soyer, R., A Bayesian Approach to Signal Analysis of Pulse Trains, *Bayesian Process Monitoring, Control and Optimization*, (Editors: Colosimo, B.M., Del Castillo, E.), Chapman & Hall, 215-243, **2006**.
- [103] UIDM (Boğaziçi Üniversitesi Kandilli Rasathanesi, Deprem Araştırma Enstitüsü Ulusal Deprem İzleme Merkezi), <http://www.koeri.boun.edu.tr/> (Mart, 2013).
- [104] Organization for Economic Co-operation and development (OECD) Data Base, <https://data.oecd.org/unemp/unemployment-rate.htm#indicator-chart> (Aralık, 2015)

ÖZGEÇMİŞ

Kimlik Bilgileri

Adı Soyadı : Ceren Eda CAN

Doğum Yeri : Ankara

Medeni Hali : Bekar

E-Posta : cerencan@hacettepe.edu.tr

Adresi : Hacettepe Üniversitesi, İstatistik Bölümü, Beytepe Kampüsü,
06800, Çankaya-Ankara

Eğitim

Lise : 1998-2001 Halide Edip Lisesi

Lisans : 2002-2007 Hacettepe Üniversitesi İstatistik Bölümü

Yüksek Lisans : 2007-2010 Hacettepe Üniversitesi İstatistik Bölümü

Doktora : 2010-2016 Hacettepe Üniversitesi İstatistik Bölümü

Yabancı Dil ve Düzeyi

İngilizce : (02 Mayıs 2010) KPDS: 82 Puan

İş Deneyimi

Araştırma Görevlisi (2007-...) Hacettepe Üniversitesi, İstatistik Bölümü

Deneyim Alanları

Saklı Markov Modelleri, Markov Modelleri, Poisson Süreçleri, Bayesci İstatistik, Olasılık Teorisi, Stokastik Modelleme, Finansal Matematik

Tezden Üretilmiş Projeler ve Bütçesi

—

Tezden Üretilmiş Yayınlar

Can, C.E., Ergün, G., Gökçeoğlu, C., Prediction of Earthquake Hazard by Hidden Markov Model (around Bilecik, NW Turkey), *Central European Journal of Geosciences*, 6(3), 403-414, **2014**.

Tezden Üretilmiş Tebliğ ve/veya Poster Sunumu ile Katıldığı Toplantılar

Can, C.E., Ergün, G., An Application of Hidden Markov Model for Turkish Earthquake Data, *EURO-INFORMS Joint International Meeting: 26-th European Conference on the Operational Research (EURO2013)*, 1 July - 4 July, Rome, Italy, **2013**

Can, C.E., Ergün, G., Gökçeoğlu, C., Bilecik Çevresinde Deprem Tehlikesinin Saklı Markov Modeli ile Tahmini, *2-nd Turkish Conference on Earthquake Engineering and Seismology (TDMSK 2013)*, 25 Eylül- 27 Eylül, Antakya/Hatay, Turkey, **2013**.

