

BAYESCİ GRAFİK MODELLERİ

BAYESIAN GRAPHICAL MODELS

ÖZGÜR TUĞRUL

105626

Hacettepe Üniversitesi
Fen Bilimleri Enstitüsü Yönetmeliğinin
İSTATİSTİK Anabilim Dalı İçin Öngördüğü
BİLİM UZMANLIĞI TEZİ
olarak hazırlanmıştır.

2001

105626
T.C. YÜKSEKÖĞRETİM KURULU
DOKÜMANTASYON MERKEZİ

BAYESCI GRAFİK MODELLERİ

Özgür Tuğrul

Hacettepe Üniversitesi, Fen Fakültesi İstatistik Bölümü

ÖZ

Karmaşık modellerin analizinde elde edilen sonsal dağılımlar çok boyutlu olduğundan, analitik yollar marjinal dağılımlara ulaşmada çoğu kez uygun olmamaktadır. Analitik çözümlerin uygun olmadığı durumlarda iteratif yöntemler kullanılabilir. Bu çalışmanın amacı, karmaşık modellerin analizinde son yıllarda ele alınan Bayesci pratik çözümlerin tanıtılması; Bayesci Grafiksel Modeller ile Gibbs Örnekleme'ne dayalı Markov Zinciri Monte Carlo (MCMC) yöntemlerinin incelenmesi ve BUGS (Bayesian Inference Using Gibbs Sampling) yazılımı ile karmaşık modellerin iteratif yöntemlerle nasıl çözümlenebildiğinin gösterilmesidir.

İncelenen modellerdeki parametreler arasındaki bağımlılık yapısının anlaşılması, karmaşık modellerin çözümlenmesinde hayati bir önem taşımaktadır. Grafik modeller ile deşifre edilen bu yapı, koşullu bağımsızlık ilkesine dayalı bir bakış açısıyla, bileşik dağılımların çarpımsal olarak ifade edilmesine olanak sağlamakta ve böylece hesaplamalarda büyük kolaylıklar elde edilebilmektedir. Markov Zinciri simülasyon teknikleri ile marjinal sonsal dağılımlardan örneklemeler iteratif olarak elde edilebilmekte ve böylece marjinal dağılımlara ilişkin çıkarsamalar yapılabilmektedir.

Çalışmada BUGS programı kullanılarak ele alınan üç farklı model için grafiksel sunumlar elde edilmiş ve modeller Gibbs Örnekleme kullanılarak işletilmiştir. Model parametreleri için marjinal sonsal dağılımlar elde edilmiştir. Klasik tahminlerle Gibbs örneklemesinden elde edilen tahminler, çoklu doğrusal bir regresyon modeli üzerinde incelenmiştir. İterasyon sayıları, başlangıç değerleri ve parametrelerin önsel dağılımları üzerindeki değişimlerin model sonuçları üzerindeki etkileri incelenmiştir. İterasyon sayısı arttıkça modelden elde edilen sonuçların asimtotik olarak gerçek sonuçlara yaklaştığı ve yine başlangıç değerlerinin sonuçlar üzerinde çok etkili olmadığı görülmüştür.

Anahtar Kelimeler: Bayesci grafik modeller, Bayesci yaklaşım, Gibbs Örnekleme, Markov Zinciri Monte Carlo yöntemi, marjinal sonsal dağılım.

Danışman: Doç. Dr. Gül ERGÜN , Hacettepe Üniversitesi, Fen Fakültesi, İstatistik Bölümü, Olasılık Teorisi ve Olasılık Süreçleri Anabilim Dalı

BAYESIAN GRAPHICAL MODELS

Özgür Tuğrul

Hacettepe University, Faculty of Science, Department of Statistics

ABSTRACT

Posterior distributions obtained by Bayesian approach usually have high dimensions when the models are rather complicated. Therefore, to reach the marginal distributions from the models is analytically intractable. The aim of the study is to introduce the newest Bayesian practical techniques for the complicated models and to investigate Bayesian graphical model and Gibbs Sampling which is a Markov Chain Monte Carlo method together. BUGS package is used to show how the complicated models are solved iteratively.

To understand the conditional structure of the models is very crucial for analysing the complicated models. Graphical models can be used to have a visual information of the structure of these models. Conditional independence allows us to factorize the joint distributions. Samples can be drawn iteratively from the marginal distributions of the model parameters by Markov Chain Monte Carlo techniques. Then statistical inference can be made easier for the marginal distributions.

Three different models are investigated in the study. BUGS package is used to have the visual representations of the three models. Gibbs samplings is applied for these models to obtain the marginal distributions of the model parameters. The result obtained from Gibbs samplings are compared with the classical results in the multiple regression model. The changes in the number of iterations, initial values and the prior distributions of the models parameters are investigated. When the number of the iteration increases, the results are very close to true values. It is also seen in the study that the initial values are not so important if the number of iterations is high.

Keywords: Bayesian graphical models, Bayesian approach, Gibbs Sampling, Markov Chain Monte Carlo techniques, Marginal Posterior Distribution.

Advisor: Doç. Dr. Gül ERGÜN, Hacettepe University, Faculty of Science, Department of Statistics, Probability Theory and Stochastic Processes Section

İÇİNDEKİLER DİZİNİ

Sayfa

1. GİRİŞ	1
2. BAYESÇİ ÇIKARSAMA.....	3
2.1 Giriş.....	3
2.2 Bayes Teoremi.....	3
2.2.1 Önsel, Sonsal ve Öngörü Dağılımları	5
2.2.2 Bilgiyi Özetleme	10
2.3 Eşlenik Dağılımlar	11
2.3.1 Üstel Aile İçin Eşlenik Dağılımlar.....	12
2.3.2 Eşleniklik ve Regresyon Modelleri.....	17
2.3.3 Koşullu Eşleniklik.....	20
2.4 Hiyerarşik Modeller.....	22
2.5 Dinamik Modeller	25
2.5.1 Ardışık Çıkarıma.....	27
3. BAYESÇİ GRAFİK MODELLEME	29
3.1 Giriş.....	29
3.2 Bayesci Grafik Modellerin Tanıtımı.....	29
3.2.1 Klasik Grafikselle Modeller.....	35
3.2.2 Bayesci Grafik Modelleri.....	36
3.2.2.1 Dinamik Doğrusal Modelin Grafikselle Sunumu	36
3.3 Bayesci Grafik Modellerinde Çıkarımlar	39
3.3.1 Model Kurma	39
3.4 Markov Zinciri Simülasyonu	42
3.5 Gibbs Örneklemesi.....	43
3.5.1 Tanımı ve Özellikleri.....	45
3.5.2 Örneklem Oluşturma	51
3.5.3 Tarama Stratejileri	52
3.5.4 Örneklem Kullanılması	52
3.6 BUGS Paket Programı	54
3.6.1 BUGS Dosyası ve Özellikleri.....	54
3.6.2 Nesnelere ve Özellikleri.....	55
3.6.3 BUGS'da Çözümleme	56
4. BAYESÇİ GRAFİK MODEL UYGULAMALARI.....	60
4.1 Çoklu Doğrusal Regresyon Modeli için Bayesci Grafik Modelleme	60
4.1.1 Modelin İşletimi	62
4.2 Hiyerarşik Çoklu Doğrusal Regresyon Modeli	69
4.3 Poisson Gamma Hiyerarşik Modeli	72
5. SONUÇLAR	76
KAYNAKLAR.....	77

ŞEKİLLER DİZİNİ

Sayfa

Şekil 3.1: Farklı Grafik Modeller	31
Şekil 3.2: Yönlendirilmiş Döngüsel Olmayan Grafik	34
Şekil 3.3: Dinamik Doğrusal Modelin Grafikselsunumu	38
Şekil 3.4: Örnek Bir Bayesci Grafikselsel Model	39
Şekil 3.5: Dört Bağımsız Koldan Gibbs Örneklemeesi	48
Şekil 3.6: İki Boyutlu Parametre Uzayında Gibbs Örnekleminin Tipik Yörüngesi ..	48
Şekil 3.7: Grafikselsel Gösterim : Node	55
Şekil 3.8: Grafikselsel Gösterim : Plate.....	55
Şekil 4.1: Çoklu Doğrusal Regresyon Modelinin Grafikselsel Yapısı	61
Şekil 4.2: Önsel Dağılımı Farklı Grafik Model.....	66
Şekil 4.3: Hiyerarşik Model.....	71
Şekil 4.4: Poisson – Gamma Hiyerarşik Modeli.....	73

TABLULAR DİZİNİ

Tablo 4.1: Önsel Dağılımlar.....	62
Tablo 4.2: 10.000 İterasyon Sonunda Model Tahminleri.....	63
Tablo 4.3: 100.000 İterasyon Sonunda Model Tahminleri.....	64
Tablo 4.4: 1.000.000 İterasyon Sonunda Model Tahminleri.....	64
Tablo 4.5: 10.000 iterasyon Sonunda Model Tahminleri.....	65
Tablo 4.6: 100.000 iterasyon Sonunda Model Tahminleri.....	65
Tablo 4.7: 1.000.000 iterasyon Sonunda Model Tahminleri.....	65
Tablo 4.8: 10.000 İterasyon Sonunda Model Tahminleri.....	67
Tablo 4.9: 100.000 İterasyon Sonunda Model Tahminleri.....	67
Tablo 4.10: 10.000 İterasyon Sonunda Model Tahminleri.....	68
Tablo 4.11: 100.000 İterasyon Sonunda Model Tahminleri.....	68
Tablo 4.12: 1.000.000 İterasyon Sonunda Model Tahminleri.....	71
Tablo 4.13: 10.000 İterasyon Sonunda Model Tahminleri.....	74
Tablo 4.14: 100.000 İterasyon Sonunda Model Tahminleri.....	74

1. GİRİŞ

İstatistikte amaç, bilinmeyen parametreler hakkında çıkarsamalar yapmaktır. Bayesci yaklaşımda tüm bilinmeyen miktarlar rastlantı değişkenidir ve burada amaç, veriler elde edildiğinde bu miktarların (parametreler) dağılımlarının elde edilmesidir. Sonsal dağılımların elde edilmesi burada önemli bir adımdır ancak son nokta değildir. Bu dağılımdan anlamlı bir bilginin çıkarılması ve bu çalışmayla ilgili olarak sonuçların yorumlanması gerekir. Bu da genellikle ortalama, ortanca ve tepe değeri gibi nokta tahminlerinin ya da aralık tahminlerinin değerlendirilmesiyle ilgilidir.

Bayesci yaklaşımda yukarıda belirtilen amaç çerçevesinde iki tür çözümlene söz konusudur. Birincisi, analitik çözümlenmeler. İkincisi ise sayısal çözümlenmelerdir. Burada analitik çözümlene öncelikle tercih edilmesi gereken bir yöntem olmalıdır. Ancak çoğu kez elde edilen bileşik dağılımın karmaşık olması, istenen sonuçlara analitik olarak ulaşılmasını zorlaştırmaktadır. Analitik çözümlenmelerde, incelenen modeldeki bilinmeyen parametre sayısının fazla olması ya da parametrelerin önsel dağılımlarının karmaşıklığı önemli integral sorunları yaratabilir. Bu sorunun çözümlenmesinde boyut indirgeme gibi bazı çözüm yolları önerilebilir. Boyut indirgeme yöntemi modeldeki önemsiz görülen bazı parametreleri ilgi dışı bırakarak, incelenecek parametre sayısını azaltmaya yönelik bir yoldur. Ancak bu durum da çoğu kez karşılaşılan problemi çözmeye yeterli olmamaktadır. Analitik yaklaşımların tıkanıdığı bu gibi durumlarda bazı yaklaştırma yöntemlerine başvurulabilir. Bu tez çalışmasında stokastik simülasyon temelli Bayesci grafiksel modelleme ele alınacaktır.

Modeldeki miktarlar arasındaki bağımlılık yapısının anlaşılması, karmaşık modellerin çözümlenmesinde hayati bir önem taşımaktadır. Grafiksel modeller bir istatistiksel modele temel teşkil eden koşullu bağımsızlık varsayımlarını şekilsel olarak ifade ederler, bu da zor modellerin deşifre edilmesini kolaylaştırır. Günümüzde ele alınan problemlerin birçoğu çok boyutlu modellerle ilgilidir. Buna örnek olarak aşağıdaki hiyerarşik model gösterilebilir:

$$y | \beta_1, \phi \sim N(X_1 \beta_1, \phi^{-1} I_n)$$

$$\beta_1 | \beta_2 \sim N(X_2 \beta_2, C)$$

$$\beta_2 \sim N(b, B)$$

$$\phi \sim G(n_0 / 2, n_0 \sigma_0^2 / 2)$$

Markov Zinciri Monte Carlo (MCMC) yöntemi, yukarıda belirtilen hiyerarşik model gibi birçok karmaşık modelde ortaya çıkan yüksek boyutlu dağılımların simülasyon problemlerine bir çözüm sağlar ve son yıllarda oldukça yoğun olarak kullanılmaktadır. Karmaşık modellerde, elde edilen bileşik sonsal dağılımdan örneklemin doğrudan elde edilmesi çok zordur. Bu nedenle bu dağılımdan bilgiye ulaşılması çoğu kez mümkün olmamaktadır. Düşük boyutlu problemlerde bazı örnekleme teknikleri kullanılabilir ya da bileşik veya marjinal dağılımlara ilişkin beklenen değerler bazı yakınsama teknikleri kullanılarak tahmin edilebilir. Ancak bu tür yakınsamaya dayalı simülasyonlar çıkarsama amacı için çok uygun olmamaktadır. Markov Zinciri Simülasyonu çok karmaşık bir teknik olmasına rağmen birçok problemin (hiyerarşik modellerin sonsal dağılımları da dahil olmak üzere) çözümünde kolay ve güvenilir sonuçlar vermektedir. Monte Carlo Markov Zinciri yöntemleri birçok farklı algoritma içermektedir. Burada ise sadece Gibbs örnekleme ele alınmıştır. Gibbs örnekleme tekniğinin seçilmesinde ana neden, tez çalışmasının uygulama bölümünde kullanılan BUGS yazılımının Gibbs örneklemesine dayalı olmasıdır.

Çalışmada Bayesci yaklaşımın tarihçesi ve ilkeleri ikinci bölümde verilmiştir. Grafik modellerin tanıtımı, grafik modellerde çıkarsama yöntemleri ve tez çalışmasında kullanılan BUGS paketi tezin üçüncü bölümünde yer almıştır. BUGS paket programına dayalı olarak kurulan Bayesci grafik modeller ve bu modellerin işletimine ilişkin sonuçlar ise, çalışmanın uygulama kısmı olan dördüncü bölümde verilmiştir. Tez çalışmasında elde edilen sonuçlar ise beşinci bölümde verilmiştir.

2. BAYESCI ÇIKARSAMA

2.1 Giriş

Bu bölümde İstatistiksel çıkarsama için Bayesci yaklaşım kısaca tanıtılmış, çıkarsama probleminde modelin oluşturulması için gerekli parçaların nasıl formüle edileceği ve uygun bilgilerin süreçten nasıl elde edileceği gösterilmiştir. Model sınıfları için daha detaylı teknikler izleyen alt bölümlerde incelenmiştir. Burada özellikle doğrusal regresyon modelleri ve onun genelleşmiş biçimleri olan genelleştirilmiş doğrusal modeller, hiyerarşik modeller, dinamik doğrusal modeller ve dinamik genelleştirilmiş doğrusal modeller ele alınmıştır.

Standart Bayesci bakış açısı ve tarihsel gelişimi Kesim 2.2'de sunulmuştur. Bayesci modellemede önemli bir rol oynayan, özel bir dağılım ailesi olan eşlenik dağılımlar ailesi Kesim 2.3'te verilmiştir. Bu dağılımların kullanımları, normal regresyon modelleri koşulları altında verilmiştir. Sonraki bölümlerde hiyerarşik ve dinamik modeller incelenmiş; bu modeller, normal gözlemler ve genelleştirilmiş normal gözlemler durumları için ele alınmıştır.

2.2 Bayes Teoremi

İstatistikte koşullu olasılık tanımı ilk olarak Bayes ve Laplace'ın çalışmalarında yer almıştır. Toplam olasılık formülü ve Bayes teoreminin 18'inci yüzyılın son yarısında ortaya atılmasına karşın, hem 19'uncu yüzyıl hem de 20'inci yüzyılın ilk yarısı Bayesci fikirler açısından oldukça karanlık bir döneme sahip olmuştur. Bayesci düşüncenin yeniden önem kazanması DeFinette, Savage, Jeffreys ve öteki bazı istatistikçiler tarafından sağlanmıştır. Günümüzde bilimsel öğrenme ve karar vermede, Bayesci yaklaşım önemli bir ilgi odağı olmaktadır ve özellikle son 30 yılda, Bayesci görüşlere dayalı bilimsel çalışmaların sayısında büyük bir artış gözlenmektedir. İstatistiksel çözümlenmelerde kullanılan Bayesci yaklaşımın kuramsal ayrıntıları için başvurulabilecek kaynak kitaplardan bazıları, Savage (1954), Jeffreys (1962), Lindley (1965), DeGroot (1970), Berger (1985), Lee (1989), Press (1989), Gamerman ve Migon (1993), O'Hagan (1994), Robert (1994) ve Smith ve Bernardo (1994)'dur (Ergün, 1995).

İstatistikte Bayesci yaklaşım, model parametreleri için önsel bilgiler ile örneklem bilgilerinin bir araya getirildiği ve parametrelere ilişkin çıkarımların yapıldığı doğal bir yoldur. Burada parametrelere ilişkin önsel bilgiler, uygun dağılımlarla modelleme sürecine dahil edilirler. Bu nedenle Bayesci yaklaşımda önsel bilgilerden önsel dağılımlara, sonsal dağılımlardan da sonsal bilgilere bir geçiş söz konusudur (O'Hagan, 1986). Uygulamalarda önsel dağılımların kolayca elde edilememesi sıkça karşılaşılan bir sorundur. Bu sorunların giderilmesinde skorlama kuralı, belirsiz (vague) ya da referans önsellerin kullanılması gibi çeşitli yollar önerilebilir.

Bayesci yaklaşım, klasik ya da sıklık yaklaşımında olduğu gibi, olasılık yoğunluk veya olasılık fonksiyonu $f(x|\theta)$ ile verilen x gözlemlerinin varlığı üzerine dayanmaktadır. Örneğin, fiziksel bir nicelik olan θ 'nin ölçümlerinden oluşan bir seri ve ölçüm hatalarının oluşturduğu e_i 'nin $N(0, \sigma^2)$ ile tanımlandığını ve σ^2 'nin bilindiği bir durumu düşünelim (Örnek 1). Bu durumda tanımlanacak model,

$$X_i = \theta + e_i, \quad i=1, \dots, n \quad (2.1)$$

ve gözlemlerin olabirlik fonksiyonu aşağıdaki şekilde olacaktır:

$$f(x|\theta) = \prod_{i=1}^n f_N(x_i; \theta, \sigma^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2} \frac{(x_i - \theta)^2}{\sigma^2}\right\}$$

Burada θ niceliğinin basit bir indeksten daha önemli olduğunu belirtmek gerekir. Ayrıca bilinmeyen parametre sayısı birden fazla olabilir.

İzleyen kesimde Bayes teoreminin uygulanması ve önsel dağılım tanımlamalarındaki bakış açıları verilmektedir.

2.2.1 Önsel, Sonsal ve Öngörü Dağılımları

Bayesci yaklaşımda bilinmeyen θ parametresi bir raslantı değişkeni olarak muamele görür. Bu nedenle, θ hakkında yapılacak çıkarsamada ilgili önsel bilgiler dağılım formunda modele dahil edilir. Diğer bir ifade ile, bu önsel bilgiler, analizler sırasında biçimsel olarak kullanılır. Bu noktada Bayesci istatistikçiler ile sıklıkçılar/klasik istatistikçiler ayrışır. Gözlemlerin, parametresi θ olan bir olasılık yoğunluk (ya da olasılık) fonksiyonundan çekilmiş olduğu bir istatistiksel çıkarsama probleminde, klasik istatistikçilerin amacı, gerçek değeri Ω parametre uzayında olduğu varsayımı altında $f(x|\theta)$ fonksiyonundan gelen gözlemlere dayalı olarak, bilinmeyen θ 'nın tahminini yapmaktır. Bayesci istatistikçiler ise, $f(x|\theta)$ 'den gözlemleri elde etmeden önce, θ parametresi hakkında bilgi ve deneyimleri kullanarak, Ω parametre uzayında θ için bir dağılım tanımı yapabilmektedirler. Klasik istatistikçiler bu bilgiyi, gözlenmediği ve bu yüzden subjektif olduğu için kabul etmezler. Bayesci yaklaşım ise ulaşılan bu bilgiyi değersiz bile olsa, bir $p(\theta)$ olasılık yoğunluk fonksiyonu ile analize dahil eder.

Önsel bilgilerin uygun bir biçimde çözümlemelere dahil edilmesine yönelik birçok değişik fikir vardır; konuyla ilgili daha fazla bilgi Efron (1986), Lindley (1978) ve Smith (1984)'ten edinilebilir. Bununla birlikte, önerilen bu fikirler arasındaki fark iyi bir şekilde anlaşılır ve uygun durumda gerekli olan kullanılırsa, bu tür bir tartışmanın anlamsız olacağı açıktır. Karar vermede en önemli nokta, Bayes teoreminin uygulanabilirliğidir. Bayesci yaklaşımın analitik olarak uygulanmasında θ 'nın boyutu önemli bir noktadır. Ancak son yıllarda gelişen teknoloji bu kısıtları da ortadan kaldırmıştır. Bayesci yaklaşımın birçok karışık problemin çözümünde kullanılmasını sağlayan teknikler geliştirilmiştir. Bu yüzden bu bölüm sadece Bayesci çıkarsamanın temel elemanlarının tanıtılmasına ayrılmıştır.

Yukarıda da anlatıldığı gibi, Bayesci yaklaşımda iki ana parça vardır; buna göre ilk parça gözlemlenen örneklem dağılımı $f(x|\theta)$ ve ikinci parça da, θ 'nın önsel dağılımı $p(\theta)$ 'dir. Bu ikinci dağılım bazı sabitlerin yardımıyla X 'in bir dağılımı şeklinde de yazılabilir. Bazı durumlarda bu dağılımı θ parametresinden ayırmak yararlı olabilir. Bu durumda bu sabitler dağılımın parametrelerinin parametreleri

oldukları için **hiperparametreler** olarak adlandırılır. Başlangıçta hiperparametrelerin bilindiği varsayılır.

Bayesci çıkarsamanın ilk bileşeni θ 'nın bir fonksiyonu olan $l(\theta) = f(x|\theta)$ olabilirlik fonksiyonudur. Bu fonksiyon, θ 'nın gözlemlenmiş x değerinin kullanılmasını sağlar. İkinci bileşen ise, θ 'nın x gözlenmeden önceki önsel yoğunluk fonksiyonu olarak adlandırılır. Bayesci yaklaşımda gözlemlerin olabilirlik fonksiyonu ile θ parametresinin önsel dağılımı elde edildikten sonra, koşullu olasılık tanımından (toplam olasılık özelliği ve Bayes teoremi) θ parametresinin sonsal dağılımı ve öngörü dağılımı elde edilir. Model parametrelerine ilişkin sonsal dağılımın elde edildiği Bayes formülü aşağıda verilmiştir:

$$p(\theta | x) = \frac{f(x|\theta)p(\theta)}{f(x)} \quad (2.2)$$

Eşitlik (2.2)'de yer alan $f(x)$, X 'in marjinal olasılık yoğunluk fonksiyonu olup; aşağıdaki şekilde elde edilir:

$$f(x) = \int f(x|\theta)p(\theta)d\theta \quad (2.3)$$

Burada $f(x)$, elde edilen sonsal dağılımın integralinin bir eşiğ olmasını sağlayan bir sabit terim olduğundan, Eşitlik (2.2) orantısal form ile aşağıdaki şekilde ifade edilebilir:

$$\pi(\theta) \propto f(x|\theta) p(\theta),$$

Sonsal Dağılım \propto Olabilirlik * Önsel Dağılım .

Bayes teoreminin basit bir uygulaması Kesim 2.2'de verilen Örnek1 üzerinde verilebilir. Burada, bilinmeyen θ 'nın önsel dağılımı $\theta \sim N(\mu, \tau^2)$ ile tanımlanabilir. Bu dağılımda, μ ve τ^2 , ele alınan modelin hiperparametreleri olup, bilinen

sabitlerdir. Olabilirlik fonksiyonu aşağıdaki şekilde ifade edildiğinde, θ 'nın sonsal dağılımı eşitlik (2.4)'deki gibi elde edilir:

$$l(\theta) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2} \frac{(x_i - \theta)^2}{\sigma^2}\right\} \propto \exp\left\{-\frac{n}{2\sigma^2} (\bar{x} - \theta)^2\right\}$$

burada \bar{x} , x 'lerin aritmetik ortalamasıdır. Böylece sonsal dağılım

$$\begin{aligned} \pi(\theta) &\propto \exp\left\{-\frac{1}{2} \frac{(\bar{x} - \theta)^2}{\sigma^2 / n}\right\} \exp\left\{-\frac{1}{2} \frac{(\theta - \mu)^2}{\tau^2}\right\} \\ &\propto \exp\left\{-\frac{1}{2} \left[\frac{(\theta - \mu_1)^2}{\tau_1^2} + \frac{(\bar{x} - \mu)^2}{n^{-1}\sigma^2 + \tau^2} \right]\right\} \end{aligned} \quad (2.4)$$

Burada $\tau_1^{-2} = n\sigma^{-2} + \tau^{-2}$ ve $\mu_1 = \tau_1^2(n\sigma^2\bar{x} + \tau^2\mu)$ 'dür.

θ 'ya bağlı olmayan çarpım terimi (2.4) nolu denklemden çıkartıldığında aşağıdaki nihai yapıya ulaşılır:

$$\pi(\theta) \propto \exp\left\{-\frac{1}{2} \frac{(\theta - \mu_1)^2}{\tau_1^2}\right\} \quad (2.5)$$

Buna göre, θ 'nın sonsal dağılımı $\pi(\theta) \sim N(\mu_1, \tau_1^2)$ olur.

Önsel dağılımın varyansı olan τ^2 'nin değeri arttıkça, önsel dağılımdan elde edilen bilginin analizler üstündeki etkisinin azalacağı eşitlik (2.4)'ten görülmektedir. Limit durumunda ($\tau^2 \rightarrow \infty$), zayıf önsel dağılım $p(\theta) \propto k$ olur ve bu durumda da sonsal dağılım klasik sonuca yakınsayarak, $\pi(\theta) = N(\bar{x}, \sigma^2 / n)$ formuna ulaşılır.

Bilgi içermeyen önsel dağılımı belirlemek konusunda pek çok karışık fikir vardır. Bu özellik çoğu kez uygun olmayan dağılımlara yol açabilir. Bu tür dağılımlar, olasılık teorisi aksiyomlarından biri olan toplamın 1 olması koşulunu sağlamazlar. Örneğin yukarıdaki örnekte ($\tau^2 \rightarrow \infty$) için, $\int p(\theta)d\theta \neq 1$ olur. Bununla birlikte, özellikle

çok deęişkenli durumlarda, bilgi içermeyen önsel dağılımlar hakkında pekçok farklı tanımlamalar vardır. En çok kabul edilen tanımlamalardan biri Jeffreys'in $p(\theta) \propto |I(\theta)|^{1/2}$ şeklinde verilen önsel dağılımıdır. Burada $I(\theta)$, θ için beklenen Fisher bilgi matrisidir ve aşağıdaki gibi tanımlanır.

$$I(\theta) = E \left[-\frac{\partial^2 \log f(x|\theta)}{\partial \theta \partial \theta'} \middle| \theta \right]$$

Jeffreys (1961), önsel bilgiyi kullanan bir Bayesci çıkarsama teorisi formüle etmiştir ve bunun parametrik dönüşümlerinde deęişmezliğini göstermiştir (Gamerman, 1997). Bu teorem çerçevesinde, konum parametresi θ için önsel dağılım, $p(\theta) \propto k$; ölçek parametresi σ için ise, $p(\sigma) \propto \sigma^{-1}$ şeklinde ifade edilir. σ ve θ parametrelerinin her ikisinin de bilinmedięi durumda ise, $p(\theta, \alpha) \propto \sigma^{-1}$ şeklinde bileşik bir önsel yoğunluk yine Jeffreys tarafından önerilmiştir.

Belirsiz önsel dağılım için bir başka seçkin belirleme de Bernardo (1979) tarafından yapılmıştır ve Berger ve Bernardo (1992), Berger ve Mendoza (1983) çalışmalarında bu yolu daha ayrıntılı olarak incelenmiştir (Gamerman, 1997). Parametre sayısının fazla olduęu, çok boyutlu durumlarda bu yaklaşım parametre vektörünü gruplara ayırır ve dięer yaklaşımlarda gözlenen bazı güçlükleri önler. Fakat boyut sayısı arttıęı için önsel dağılımın belirlenmesi güçleşir.

Genel olarak, önsel dağılımların belirsizlięi sorun yaratır. Fakat Bayes teoreminin uygulanması sonucunda, bu önseller bilinen bir sonsal dağılıma ulaşırlar ve böylece çıkarsama daha kolay olur. Sonsal dağılımın belirli olmadığı bazı istisnai durumlar vardır. Bu birçok model için, önemsiz olmaktan uzak, oldukça ciddi bir problemdir. Bu modeller için kesin çıkarsama yapılamaz ve yaklaşımlar bizi birbiriyle uyuşmayan sonuçlara götürebilir. Bu durumda önerilen yol, belirsiz önsellerin kullanımından mümkün olduğunca uzak durmak ya da onları çok dikkatli kullanmaktır.

Bayesci çıkarsamanın bir başka önemli unsuru da $f(x)$ yoğunluk fonksiyonu ile verilen X 'in marjinal fonksiyonudur (Eşitlik 2.3). Bu dağılım gözlenen x için beklenen dağılımı sağlar. Buna göre, $f(x) = [E(f(x|\theta))]$ 'dir ve beklenen değer θ 'nın önsel dağılımından yararlanarak elde edilir. Benzer bir çıkarsama x gözlendikten sonra ileride gözlenecek y 'nin öngörüsü için de kullanılabilir. Burada öngörüler, $y|X$ fonksiyonunun dağılımına dayandırılabilir. Eğer y ve X koşullu olarak θ 'dan bağımsız ise, aşağıdaki eşitlik yazılabilir (Lee, 1989):

$$f(y|x) = \int f(y, \theta | x) d\theta = \int f(y | \theta) d\theta \pi(\theta) d\theta \quad (2.6)$$

Burada, öngörü dağılımı, örneklem dağılımından elde edilir. Eşitlik (2.6)'da θ 'nın sonsal dağılımına dikkat edilmelidir. X ve y arasında koşullu bağımsızlık, $X' = (x_1, \dots, x_n)$ ve $y = (x_{n+1}, \dots, x_{n+m})$ olduğu durumda sağlanır. Bu durumda öngörü fonksiyonu kitlenin gelecekteki değerlerini öngörmek için kullanılabilir.

Öngörü dağılımı, çıkarsamaya öngörüsel yaklaşımın temelini oluşturur. Bu yaklaşım Aitchinsen ve Dunsmore (1975) ve Geisser (1993) tarafından detaylı olarak anlatılmıştır (Gamerman, 1997). Parametreler hakkında çıkarsama yapılırken, parametreler gözlenemediğinden bu karşılaştırmayı yapmak mümkün değildir. Yine de önemli parametreler için, aşağıdaki eşitlikle elde edilecek anlamlı olasılıksal ifadeler bulmak mümkündür.

$$\pi(\theta_i) = \int \pi(\theta_1 \dots \theta_d) d\theta_{-i} \quad (2.7)$$

Burada, $\theta_{-i} = (\theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_d)$ vektörü, θ_i elemanı çıkarılmış vektördür. Her θ_i için mümkün koşullu dağılımlar aşağıdaki gibi ifade edilebilir:

$$\pi(\theta_i | \theta_j, j \in C) = \pi(\theta_i, \theta_j, j \in C) | \pi(\theta_j, j \in C) \text{ 'dir.}$$

Burada $C = \{1, \dots, i-1, i+1, \dots, p\}$ 'dir. Burada en önemli nokta, $\theta_i | \theta_{i-1}$ 'in θ_i 'nin tam koşullu fonksiyonu olarak adlandırılması ve $\pi_i(\theta_i)$ ile gösterilmesidir. Yukarıdaki tanımlamalar her θ_i 'nin vektör olduğu durumlar için genişletilebilir.

2.2.2 Bilgiyi Özetleme

Önsel dağılım elde edildikten sonra, içerdği bilgi değişik biçimlerde özetlenebilir. Önsel dağılımdaki değişimi göstermek için konum ve dağılım ölçüleri hesaplanabilir. Temel konum ölçüleri ortalama, ortanca ve tepe değeri; dağılım ölçüleri ise, varyans, standart sapma, ortalama mutlak sapma, çeyrek değerler vs. dir. Önsel dağılımın ortalaması, θ 'nın beklenen değerini; tepe değeri, fonksiyonu maksimum yapan değeri ve ortanca da parametre uzayını iki eşit parçaya bölen değeri tanımlamaktadır.

Çok değişkenli durumda ise varyans bir matrisle tanımlanır ve standart sapma, köşegen öğelerinin kareköklerinden oluşan bir vektördür. Ortanca dışında bütün bu ölçüler, bileşik ve marjinal dağılımlar için hesaplanabilir. Ortanca ise, sadece tek değişkenli durumlar için hesaplanabilir. Burada adı geçen ölçülerin ve bunların karar kuramıyla ilişkileri Gamerman ve Migon'da (1993) ayrıntılı bir şekilde verilmiştir (Gamerman, 1997).

Çok değişkenli uzayda, marjinal yoğunluklar parametre uzayının bir elemanı üzerinde çıkarsama yapabilmek için yararlıdır. Marjinal yoğunluklar Eşitlik 2.7'de verilen integral ile elde edilebilirler. Elde edilen marjinal olasılık dağılımları, parametrelerin olasılık aralıklarıyla bilgiyi özetlemek için kullanışlıdır. Eğer $\int_C \pi(\theta) d\theta = 1 - \alpha$ ise C, bu θ parametresi için $(1 - \alpha)$ güven aralığıdır. Verilen bir α değeri için, aralığın içinde olmayan noktalardan daha yüksek bir önsel dağılım içeren C aralığı, en küçük aralıktır. Bu aralıklar daha yüksek önsel yoğunluk aralıkları (highest density region, HDR) olarak adlandırılırlar.

Elde edilen sonsal dağılımın yapısına bağlı olarak birçok özetleme yöntemi kullanılır. Kesin çıkarsamalarda analitik yollar uygulanabilirken; bazı durumlarda yaklaşımlara başvurulabilir. Kesim 2.2'de verilen örnek1'deki θ parametresinin önsel dağılımının bu kez (normal dağılım yerine) μ konum ve τ^2 dağılım ölçüleri ile Cauchy dağılımı olduğunu düşünelim. Bu durumda sonsal dağılım,

$$\pi(\theta) \propto \exp\left\{-\frac{1}{2} \frac{(\bar{x} - \theta)^2}{\sigma^2/n}\right\} \frac{1}{\tau^2 + (\theta - \mu)^2} \quad (2.8)$$

olur ve elde edilen bu dağılım, daha basit bir biçime indirgenemez. Bir önceki sonucun tersine elde edilen dağılım bilinen bir dağılım değildir ve özetleyici bilgiler analitik olarak hesaplanamaz. Bununla birlikte sonsal dağılımın basit sayısal hesaplamalarla özetlenmesi istendiğinde daha dikkatli olunması gerekmektedir. (2.8) nolu ifade, özetleyici bilgilerin elde edilmesinin zor olduğunun bir göstergesidir. Bunun nedeni de Cauchy dağılımının uzun kuyruklu bir dağılım olması ve uzun kuyruklu dağılımların genelde yeterli bilgi olmadığı durumlarda kullanılmasıdır.

2.3 Eşlenik Dağılımlar

Kesim 2.2'de verilen Örnek1'de görüldüğü üzere, önsel ve sonsal dağılımların aynı dağılım ailesinden olmaları, çıkarsama problemlerinde büyük işlem kolaylıkları sağlar. Önselden sonsala geçişte tek değişim sadece hiperparametrelerde olur ve θ 'nın dağılımı, rutin olarak yeni gelen gözleme bağlı olarak, kolaylıkla güncelleştirilir.

Örnek 1 ile verilen modelde $x_1 = (x_{11}, \dots, x_{1n})'$, $N(\theta, \sigma^2)$ alınmış bir örneklem olsun. Buna göre θ için sonsal dağılım, $\tau_1^{-2} = n_1\sigma^{-2} + \tau^{-2}$, $\mu_1 = \tau_1^2(n_1\sigma^{-2}\bar{x}_1 + \tau^{-2}\mu)$ ve \bar{x}_1 ortalamaya sahip olan $N(\mu_1, \tau_1^2)$ dağılımıdır. Eğer yeni bir örneklem $x_2 = (x_{21}, \dots, x_{2n})'$ aynı kitleden çekilirse x_1 bilindiğinde θ 'nın sonsal dağılımı, x_1 ve x_2 'nin koşullu fonksiyonu olur. Yukarıda görüldüğü gibi, önsel dağılım normal ise sonsal dağılım da normal olur. Bu durumda (μ_1, τ_1^2) ile yeni önsel dağılım (μ_2, τ_2^2) hiperparametreleri arasındaki ilişki aşağıdaki şekilde olur:

$$\tau_2^{-2} = n_2\sigma^{-2} + \tau_1^{-2}$$

$$\mu_2 = \tau_2^2(n_2\sigma^{-2}\bar{x}_2 + \tau_1^{-2}\mu_1)$$

Burada ele alınan önsel dağılım, genel anlamda kolay anlaşılır, iyi bilinen bir dağılımdır. Şayet sonsal dağılım aynı ailedense, önsel dağılımın özellikleri sonsal dağılımda da aynen korunur ve veriyi özetlemek oldukça kolaylaşır.

Güncelleştirmeden sonra dağılımın korunmasına eşleniklik denir. Eğer her önsel dağılım $p \in P$ ve gözlenen herhangi bir dağılım için $f \in F$ ise, P gözlenen model F 'e eşlenik denir ve sonsal dağılım da $\pi \in P$ olur. Bu kural her P ailesi, bütün süreklili dağılım aileleri ve nadiren tek dağılımı içeren kısıtlı aileler için de geçerlidir.

Eşlenik dağılımların birçok avantajlarına rağmen, eşlenik sonsal dağılımlarla çalışılırken dikkatli olunması gerekir. Burada kolay hesaplamaların bir bedeli olarak, önsel dağılımın şekli üzerinde bazı kısıtlamalar getirilmiştir. Çoğu kez mevcut durum için eşlenik önsel dağılım uygun bir dağılım değildir. Bu durumda hesaplamalardaki kolaylık önemli bir avantaj olmaktan çıkar ve sonsal dağılıma ulaşmak zorlaşır. Uygulamalı istatistikçiler için gerçeğe uygunlukla, yöntemin kolaylığı arasında daima bir çelişki vardır. Bu sorunun kesin bir cevabı olmamakla birlikte, burada birkaç örnekle hangi durumda hangi yaklaşımın daha uygun olacağına dair temel bilgiler verilecektir.

2.3.1 Üstel Aile İçin Eşlenik Dağılımlar

X raslantı değişkeni, tek parametrelili üstel dağılım ailesine mensup bir raslantı değişkeni ise, $f(x|\theta)$ olasılık fonksiyonu aşağıda verilen yapıya sahiptir:

$$f(x|\theta) = a(x)\exp\{\phi(\theta)t(x) + b(\theta)\} \quad (2.9)$$

Eşitlik (2.9) ile verilen üstel dağılım ailesinin önemi, modelleme kullanılan pek çok dağılımın bu sınıfa ait olmasıdır. Üstel aileye mensup birçok dağılım Eşitlik (2.9)'da verilen yapıda aşağıda verilmiştir:

(a) Normal Dağılım (σ^2 biliniyor; $X \sim N(\theta, \sigma^2)$)

$$\phi(\theta) = \frac{\theta}{\sigma^2}, t(x) = x, b(\theta) = -\frac{\theta^2}{2\sigma^2}, a(x) = \exp\left\{-\frac{x^2}{\sigma^2}\right\}$$

(b) Beta Dağılımı ($X \sim B(n, \theta)$)

$$\phi(\theta) = \log\left(\frac{\theta}{1-\theta}\right), t(x) = x, b(\theta) = n \log(1-\theta), a(x) = \binom{n}{x}$$

(c) Üstel dağılım ($X \sim \exp(\theta)$)

$$\phi(\theta) = -\theta, t(x) = x, b(\theta) = \log(\theta), a(x) = 1$$

(d) Poisson dağılımı ($X \sim \text{Poisson}(\theta)$)

$$\phi(\theta) = \log(\theta), t(x) = x, b(\theta) = -\log(\theta), a(x) = 1$$

Üstel aileye dahil olmayan önemli dağılımlar, uniform dağılım, t dağılımı ve bazı kesikli dağılımlardır. Eşitlik (2.9) ile verilen açıklama, birçok benzer duruma uyarlanabilir. $\phi(\cdot)$ ve $t(\cdot)$ için kesin bir tanımlama olmamakla birlikte, herhangi bir c sabitinin çarpılması ya da bölünmesine göre, $\phi(\theta)t(x)$ çarpımı sabit kalacak biçimde değişebilir. Üstel dağılım için, alternatif bir tanımlama olarak, $\phi(\theta) = \theta$ ve $t(x) = -x$ olarak alınabilir. Birçok durumda $t(x) = x$ 'dir. $x(t)$ ve $\theta(\phi)$ ters dönüşümlerinin olduğu varsayılırsa, t'nin yoğunluk fonksiyonu aşağıdaki biçime sahiptir:

$$f^\bullet(t | \theta) = a^\bullet(t) \exp\{\phi t + b^\bullet(\theta)\}$$

Burada $a^\bullet(t) = [x(t)] \left| \frac{dx(t)}{dt} \right|$ ve $b^\bullet(\phi) = b[\theta(\phi)]$ 'dir.

Üstel ailenin her iki tanımı da bu dönüşümler tanımlandıktan sonra kullanılabilir. Bu çalışmada kolaylık açısından ikinci gösterim tercih edilmiştir. Buna göre üstel aile yoğunluk fonksiyonu şu biçimde yazılabilir:

$$f(x|\theta) = a(x)\exp\{\theta x + b(\theta)\} \quad (2.10)$$

Eşitlik (2.10)'dan beklenen değer ve varyans aşağıdaki şekilde elde edilir:

$$E(x|\theta) = \mu = -b'(\theta) \text{ ve } \text{var}(x|\theta) = -b''(\theta) \text{ olur.}$$

Burada b' ve b'' , birinci ve ikinci dereceden türevleri göstermektedir ve b 'nin ikinci dereceden türevlenebilir bir fonksiyon olduğu varsayılmıştır. b'' fonksiyonu bazen varyans fonksiyonu yerine tercih edilir (McCullagh ve Nelder, 1988).

θ parametresi için tanımlanan önsel $p(\theta) = k(\alpha, \beta)\exp\{x\theta + \beta b(\theta)\}$ biçiminde olsun. Burada normalleştirme parametresi k , α ve β hiperparametrelerine bağlıdır. (2.10) ifadesinde tanımlanan bir gözlem elde edildiğinde, Bayes teoremi ile aşağıda verilen sonsal dağılıma ulaşılır:

$$\begin{aligned} \pi(\theta) &\propto f(x|\theta)p(\theta) \\ &\propto \exp\{\theta x + b(\theta)\}\exp\{\alpha\theta + \beta b(\theta)\} \\ &= \exp\{(\alpha + x)\theta + (\beta + 1)b(\theta)\} \\ &= \exp\{\alpha_1\theta + \beta_1 b(\theta)\} \end{aligned} \quad (2.11)$$

Burada $\alpha_1 = \alpha + x$ ve $\beta_1 = \beta + 1$ 'dir. Eşitlik (2.11)'de elde edilen dağılım, eşitlik (2.10) ile aynı dağılımdır. Genel anlamda aynı dağılım ailesine üyelerler. Eşlenik önsel dağılım $\theta \sim \text{CP}(\alpha, \beta)$ ile gösterilirse, (2.11) eşitliğinde verilen sonsal dağılım da $\theta/x \sim \text{CP}(\alpha_1, \beta_1)$ şeklinde yazılabilir.

Yukarıdaki çıkarsamalar, olabilirlik fonksiyonu ile aynı biçimdeki önsel dağılımları aramanın bizi eşlenik dağılıma kolayca götürdüğünü göstermektedir.

Gözlenen modelin λ parametresi ile Poisson dağılımlı olduğu ve λ için önsel dağılımın da $\text{Gamma}(\alpha, \beta)$ olduğu bir durum için, Bayes teoremi uygulandığında aşağıdaki sonuç elde edilir:

$$\begin{aligned}\pi(\lambda) &\propto e^{-\lambda} \lambda^x \lambda^{\alpha-1} e^{-\beta\lambda} \\ &\propto \lambda^{\alpha+x-1} e^{-(\beta+1)\lambda}\end{aligned}$$

Ulaşılan sonuç, λ için sonsal dağılımın $\text{Gamma}(\alpha+1, \beta+1)$ olduğunu göstermektedir. Buradan Gamma ailesinin Poisson gözlemsel modeli için eşlenik aile olduğunu görebiliriz.

Üstel dağılım ailesinde daha pekçok eşlenik durum söz konusudur. Bunlar Gamerman ve Migon (1993) çalışmasında ayrıntılı olarak verilmiştir. Bayes teoreminin ardışık yapısından dolayı (bir önceki adımın sonsalının, bir sonraki adımın önseli olduğu), tekrarlar bir örneklemin incelenmesi için başarılı bir biçimde uygulanabilir. Eğer böyle bir güncelleştirme işlemi sonucunda hala aynı dağılım ailesi korunuyorsa, bu durum n adım sonra da korunacaktır. Bu yüzden tek gözlem için eşleniklik kavramı, n gözlem için eşleniklik kavramına denktir. Eşitlik (2.11)'in her adımında, Bayes teoremine göre x_i 'nin her gözlemi ile güncelleştirme yapmak modelin hiperparametrelerini azaltır. Buna göre güncellenen parametreler, $\alpha_i = \alpha_{i-1} + x_i$ ve $\beta_i = \beta_{i-1} + 1$, $i=1, \dots, n$ ve $\alpha_0 = \alpha$, $\beta_0 = \beta$ 'dir ve n büyüklüğünde bir örneklem gözlendikten sonra önsel ve sonsal dağılımlar aynı dağılım sınıfına ait olur. Burada sonsal dağılımın parametreleri, $\alpha_n = \alpha + \sum x_i$ ve $\beta_n = \beta + n$ olur. Bundan dolayı model eşleniktir. Bu durum $\theta | x_1 \dots x_n \sim CP(\alpha_n, \beta_n)$ olarak yazılabilir.

Üstel aile tanımı, çok parametreli durum için genelleştirilebilir. $f(x|\theta)$, aşağıdaki biçimde ayrıştırılabiliyorsa, aşağıda belirtilen k parametrelili bir üstel aileye dahildir.

$$f(x|\theta) = a(x) \exp \left\{ \sum_{i=1}^k \theta_i X_i + b(\theta) \right\} \quad (2.12)$$

Burada, $X = (x_1, \dots, x_k)'$ ve $\theta = (\theta_1, \dots, \theta_k)'$ dir.

Parametre sayısının birden fazla olduğu durum için şu örnek verilebilir: $X = (x_1, \dots, x_k)'$ raslantı değişkeni vektörü n indeksi ve p_1, \dots, p_k olasılıkları ile

multinomial dağılıma sahip olsun ve $\sum x_i = n$ ve $\sum p_i = 1$ koşulu sağlansın. Buna göre bileşik olasılık fonksiyonu aşağıdaki biçimde ifade edilir:

$$f(x | p_1 \dots p_k) = \frac{n!}{\prod_{i=1}^k x_i!} \prod_{i=1}^k p_i^{x_i}, \quad 0 \leq x_i \leq n, \sum_i x_i = n$$

Olasılık fonksiyonundaki $\theta = (\theta_1, \dots, \theta_{k-1})'$ ların kanonik parametreleri $\theta_i = \log(p_i / p_k), i = 1, \dots, k-1$ ile eşitlik (2.12)'deki biçimde yazılabilir.

Aynı şekilde bir eşlenik dağılımlar ailesi oluşturulduğunda, önsel dağılım $p(\theta)$ aşağıdaki gibi tanımlanır:

$$p(\theta) = k(\alpha, \beta) \exp \left\{ \sum_{i=1}^k \alpha_i \theta_i + \beta b(\theta) \right\} \quad (2.13)$$

Burada normalleştirme parametresi k yine hiperparametreler $\alpha = (\alpha_1, \dots, \alpha_k)'$ ve β 'ya bağlıdır. Eşitlik (2.12) ve eşitlik (2.13) Bayes formülünde kullanıldığında, aşağıdaki sonuca ulaşılır:

$$\begin{aligned} \pi(\theta) &\propto f(x | \theta) p(\theta) \\ &\propto \exp \left\{ \sum_{i=1}^k \theta_i x_i + b(\theta) \right\} \exp \left\{ \sum_{i=1}^k \theta_i \alpha_i + \beta b(\theta) \right\} \\ &= \exp \left\{ \sum_{i=1}^k (\alpha_i + x_i) \theta_i + (\beta + 1) b(\theta) \right\} \\ &= \exp \left\{ \sum_{i=1}^k (\alpha_{1i} \theta_i + \beta_1 b(\theta)) \right\} \end{aligned}$$

Burada $\alpha_{1i} = \alpha_i + x_i, i=1, \dots, k$ ve $\beta_1 = \beta + 1$ 'dir. Görüldüğü üzere, π ve p yoğunlukları (2.12) modeli ile aynı biçime sahiptirler ve aynı sınıfa ait oldukları için model eşleniktir.

$\theta = (\theta_1, \dots, \theta_k)'$ rastlantı vektörü $\sum \theta_i = 1$ kısıtı altında $\alpha = (\alpha_1, \dots, \alpha_k)$ parametresi ile Dirichlet dağılımına sahip olsun ve $D(\alpha)$ ile gösterilsin. Bu durumda yoğunluk fonksiyonu,

$$f_D(\theta) = \frac{\Gamma(\sum_{i=1}^k \alpha_i)}{\prod_{i=1}^k \Gamma(\alpha_i)} \prod_{i=1}^k \theta_i^{\alpha_i - 1}, \quad 0 < \theta_i < 1, (\alpha_i > 0), i = 1, \dots, k$$

dır. $p=(p_1. . . p_k)$ 'lar için önsel dağılım olan Dirichlet dağılımı ile multinomial dağılım Bayes teoremi ile birleştirildiğinde, sonsal dağılım $\pi(p)=D(\alpha_1)$ olur ve burada $\alpha_1 = \alpha + x'$ 'tir. Böylece Dirichlet ailesinin multinomial model için eşlenik aile olduğu gösterilmiş olur (Gamerman, 1997).

2.3.2 Eşleniklik ve Regresyon Modelleri

Kesim 2.3.1'de verilen sonuçlar normal regresyon modeli için genişletilebilir. Burada olasılıksal tanımı bilinen ve diğer değişkenlerden etkilenen bağımlı bir değişken gözlenir. Ele alınan bağımsız değişkenlerin bağımlı (cevap) değişken üzerinde bir çok değişik etki biçimleri vardır. Burada öncelikle doğrusal bir yapı irdelenecektir.

Bağımlı değişkenin normal dağıldığı durumda,

$$y_i \sim N(x_{i1}\beta_1 + \dots + x_{id}\beta_d, \sigma^2), \quad i = 1, \dots, n$$

ile tanımlanan $y = (y_1, \dots, y_n)'$ gözlemleri ile bir doğrusal regresyon modeli elde edilir. Burada $x_{i1}. . . x_{id}$, d sayıda değişkenin i . gözlemleridir ve $\beta_1. . . \beta_d$ ise, bu değişkenlere ait regresyon katsayılarıdır. Bağımsız gözlemler için model,

$$y = N(x\beta; \sigma^2 I_n) \tag{2.14}$$

matris biçiminde yazılır. Burada, $x = \begin{bmatrix} x_{11} & \cdot & \cdot & x_{1d} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ x_{n1} & \cdot & \cdot & x_{nd} \end{bmatrix}$ ve $\beta = \begin{bmatrix} \beta_1 \\ \cdot \\ \cdot \\ \beta_D \end{bmatrix}$ 'dir.

Bağımsız değişkenlere ilişkin metinde matrisin her zaman tam ranklı ($d \leq n$) olduğu varsayılır. Bu varsayımı sağlamayan durumlarda basit matematiksel işlemlerle bu bozulmuş ortadan kaldırılabilir. Bazı durumlarda homojen varyanslılık hipotezi daha genel bir biçimle değiştirilir. Burada $\text{var}(y) = \text{köş}(\sigma_1^2, \dots, \sigma_n^2)$ 'dir. Hatta daha genel olarak $\text{var}(y) = \Sigma$ kullanılabilir. Σ , $n \times n$ boyutlu simetrik pozitif tanımlı matristir.

Bayesci bakış açısında model (2.14) bir sonsal dağılıma ulaşarak tamamlanır. Bu durumda β parametresine ve model varyansına ilişkin önsel dağılımlar,

$$\beta | \sigma^2 \sim N(b_0, \sigma^2 \beta_0) \text{ ve } \sigma^2 \sim IG\left(\frac{n_0}{2}, \frac{n_0 S_0}{2}\right)$$

dir. Bu arada $n_0 S_0 / S^2 \sim \chi^2_{n_0}$ olduğunu hatırlamak iyi olur. Bu eşlenik durumu kullanmadan önce bir başka parametreleştirme uygulanabilir. Buna göre, σ^2 , $\phi = \sigma^{-2}$ olarak değiştirilebilir. ϕ 'ye göre önsel dağılım, $n_0 S_0 \phi \sim \chi^2_{n_0}$ olduğundan, bileşik önsel dağılım,

$$\begin{aligned} p(\beta, \phi) &\propto p(\beta | \phi) p(\phi) \\ &\propto \phi^{d/2} \exp\left\{-\frac{\phi}{2}(\beta - b_0)' B_0^{-1}(\beta - b_0)\right\} \phi^{(n_0/2)-1} \exp\left\{-\frac{\phi}{2} n_0 S_0\right\} \\ &\propto \phi^{[(n_0+d)/2]-1} \exp\left\{-\frac{\phi}{2}[n_0 S_0 + (\beta - b_0)' B_0^{-1}(\beta - b_0)]\right\} \end{aligned} \quad (2.15)$$

dır. Eşitlik (2.14) olabilirlik β ve ϕ cinsinden aşağıdaki gibi yazılabilir:

$$\begin{aligned}
 l(\beta, \phi) &= f(y|\beta, \phi) \\
 &\propto \phi^{-n/2} \exp\left\{-\frac{\phi}{2}(y - X\beta)'(y - X\beta)\right\} \\
 &\phi^{-n/2} \exp\left\{-\frac{\phi}{2}[Q(\beta) + S_e]\right\}. \tag{2.16}
 \end{aligned}$$

Burada $Q(\beta) = (\beta - \hat{\beta})'X'X(\beta - \hat{\beta})$, $\hat{\beta} = (X'X)^{-1}X'Y$ ve

$S_e = (y - X\hat{\beta})'(y - X\hat{\beta}) = \sum_{i=1}^n e_i^2$ $e_i = y_i - \hat{\mu}_i$ ve $\hat{\mu}_i = x_{i1}\hat{\beta}_1 + \dots + x_{id}\hat{\beta}_d$ dir. Eşitlik

(2.16)'dan $\hat{\beta}$, β 'nin bir fonksiyonudur ve olabilirlik fonksiyonunu maksimum yapar.

Eşitlik (2.15) ve (2.16) birleştirildiğinde sonsal dağılıma ulaşılır:

$$\begin{aligned}
 \pi(\beta, \phi) &\propto l(\beta, \phi)p(\beta, \phi) \\
 &\propto \phi^{[(n_0+n+d)/2]-1} \exp\left\{-\frac{\phi}{2}[n_0S_0 + S_e + (\beta - b_0)'B_0^{-1}(\beta - b_0) + Q(\beta)]\right\} \\
 &\propto \phi^{d/2} \exp\left\{-\frac{\phi}{2}(\beta - b_1)'B_1^{-1}(\beta - b_1)\right\} \phi^{\frac{n_1}{2}-1} \exp\left\{-\frac{\phi}{2}n_1S_1\right\} \tag{2.17}
 \end{aligned}$$

Burada, $n_1 = n_0 + n$, $B_1^{-1} = B_0^{-1} + X'X$,

$n_1S_1 = n_0S_0 + (y - Xb_1)'y + (b_0 - b_1)'B_0^{-1}b_0$ ve $b_1 = B_1(B_0^{-1}b_0 + X'y)$ 'tir. Eşitlik (2.15) ile (2.17) karşılaştırıldığında önsel ve sonsal dağılımların aynı biçime sahip oldukları ve aynı eşlenik aileye ait oldukları açıkça görülebilir (Gamerman, 1997).

Genel doğrusal modellerin birçok uygulaması vardır, özel durumlar ve maksimum olabilirlik tahmini McCullagh ve Nelder (1988) çalışmasında ayrıntılı bir biçimde verilmiştir. Bayesci çıkarsama için parametre vektörü β için bir önsel dağılım tanımlanmalıdır. Normal regresyonun bir uzantısı önseli $N(b_0, B_0)$ olarak almaktır. Ne yazık ki bu dağılım eşlenik değildir ve bu nedenle kesin Bayesci çıkarsama yapmak mümkün olmaz. Aynı yorum nadir görülen ve kullanışlı olmayan ve istisnaları olan diğer önsel dağılımlar için de geçerlidir. Bu çıkarsama işlemlerinin uygulanamadığı durumlardan birini gösteren güzel bir örnektir.

2.3.3 Koşullu Eşleniklik

Eşlenik dağılım elde etmekteki güçlük modelin karmaşıklığı ve parametre uzayının büyümesi ile daha da artar. Kesim 2.3.2'de belirtildiği gibi genel doğrusal modellerin önemli bir sınıfı eşlenikliğe uymaz. Eşitlik (2.14) ile verilen normal regresyon modelini düşünelim. Burada $\beta \sim N(b_0, B_0)$ ve $\phi \sim G(n_0/2, n_0 S_0/2)$ önsel dağılımları bağımsız dağılımlardır. Bir önceki duruma göre farklılık β ve ϕ arasındaki koşullu varyansta $\beta|\phi$ 'deki bağımlılığın kalkmasıdır. Buna göre sonsal dağılım aşağıdadır:

$$\begin{aligned}
 \pi(\beta, \phi) &\propto l(\beta, \phi) p(\beta) p(\phi) \\
 &\propto \phi^{n/2} \exp\left\{-\frac{\phi}{2}[(\beta - \hat{\beta})' X' X (\beta - \hat{\beta}) + S_e]\right\} \\
 &\propto \exp\left\{-\frac{1}{2}[(\beta - b_0)' B_0^{-1} (\beta - b_0)]\right\} \phi^{(n_0/2)-1} \exp\left\{-\frac{n_0 S_0 \phi}{2}\right\} \\
 &\propto \phi^{(n+n_0/2)-1} \exp\left\{-\frac{\phi}{2}[(\beta - \hat{\beta})' X' X (\beta - \hat{\beta}) + S_e + n_0 S_0]\right\} \\
 &\propto \exp\left\{-\frac{1}{2}[(\beta - b_0)' B_0^{-1} (\beta - b_0)]\right\} \tag{2.18}
 \end{aligned}$$

Burada önsel dağılımdan herhangi bir tanımlanabilir biçim yakalamak mümkün değildir. Bu nedenle eşleniklik yoktur. Aslında hala bir eşleniklik yakalamak mümkündür. Çünkü önsel dağılım, sonsal dağılımın özel bir durumudur. Bu tür düşüncenin pratik değeri vardır. Eşleniklik kavramı modeli sadece analitik olarak incelenebilecek bir duruma götürdüğü zaman yararlıdır. Ancak burada böyle bir durum söz konusu değildir.

Eşleniklik olmamasına rağmen, $\beta|\phi$ ve $\phi|\beta$ koşullu sonsal dağılımlarını aşağıdaki gibi elde etmek mümkündür.

$$\begin{aligned}\pi(\beta, \phi) &\propto \exp\left\{-\frac{1}{2}\left[(\beta - \hat{\beta})'X'X(\beta - \hat{\beta}) + (\beta - b_0)'B_0^{-1}(\beta - b_0)\right]\right\} \\ &\propto \exp\left\{-\frac{1}{2}\left[(\beta - b_\phi)'B_\phi^{-1}(\beta - b_\phi)\right]\right\}\end{aligned}$$

Burada $b_\phi = B_\phi(B_0^{-1}b_0 + \phi X'y)$ ve $B_\phi^{-1} = B_0^{-1} + \phi X'X$ 'tir. Buradan $\beta|\phi \sim N(b_\phi, \beta_\phi)$ biçiminde olur. β ve ϕ arasındaki bağımsızlık ile, $\beta|\phi \sim N(b_0, \beta_0)$ olur. Bu durumda β koşullu eşleniktir.

Aynı biçimde $\phi|\beta$ 'nin sonsal dağılımı,

$$\pi(\phi, \beta) \propto \phi^{(n+n_0)/2-1} \exp\left\{-\frac{\phi}{2}\left[(\beta - \hat{\beta})'X'X(\beta - \hat{\beta}) + S_e + n_0 S_0\right]\right\}$$

olur ve bu yüzden $\phi|\beta \sim \text{Gamma}(n_1/2, n_1 S_\beta/2)$ olur. Burada, $n_1 = n + n_0$ ve $S_\beta = (\beta - \hat{\beta})'X'X(\beta - \hat{\beta}) + S_e + n_0 S_0$ 'dir. Yine burada koşullu önsel dağılım ile koşullu sonsal dağılım aynı biçime sahiptir. Bu yüzden her iki koşullu dağılım için de eşleniklik vardır.

Koşullu önsel dağılımlar marjinal önsel dağılımları kolaylaştırır. Önsel dağılımlar bağımsız olmasa bile koşullu olarak eşlenik olması mümkündür.

Parametre vektörü $\theta = (\theta_1, \dots, \theta_d)'$ için eğer tam koşullu önsel dağılımlar $(p_i(\theta_i) = p(\theta_i | \theta_{-i}))$ ve tam koşullu sonsal dağılımlar $(\pi_i(\theta_i) = \pi(\theta_i | \theta_{-i}))$ aynı dağılımlar ailesine aitlerse koşullu eşleniklik söz konusu olur. Diğer durumlarda sadece birkaç öge koşullu eşlenik olur.

Koşullu eşleniklik, olasılık özelliklerinin sadece belli durumlarda sağlandığı pek çok karmaşık modelde bile mevcuttur. Boyut sayısı arttığında parametre uzayında bileşik önsel dağılım oluşturmak oldukça zordur. Bu nedenle parametre vektörünün elemanları arasında koşullu bağımsızlıkların araştırılması gereklidir. Bu da küçük boyutlu parametre uzayının elemanları hakkında olasılıksal önermeler biçimindedir. Bunları tanımak ve sınıflamak daha kolaydır. Bu yaklaşımla model oluşturma çalışmanın ileri bölümlerinde kullanılacaktır. Analitik yolların tıkandığı durumlarda çıkarsama yaparken, tam koşullu dağılımların ve mümkün olduğunda koşullu eşlenikliğin önemi artar.

2.4 Hiyerarşik Modeller

Kesim 2.3.2' de verilen Bayesci normal regresyon modelinde bağımlı değişken y ile açıklayıcı değişkenler $x_1 \dots x_p$ arasındaki ilişki, regresyon katsayıları β vektörü ile verilmişti. Pek çok durumda β hakkındaki nicel olasılık ifadeleri modele oturtulamaz.

Örneğin, $y_{ij} \sim N(\beta_i, \sigma^2)$, $j = 1 \dots n, i = 1 \dots d$ farklı β ortalamalı fakat aynı varyanslı d tane gruptan alınmış gözlemler olsun. Bu model, $y = (y_{11} \dots y_{1n_1} \dots y_{d1} \dots y_{dn_d})'$ ve veri matrisi $X = köş(1_{n_1} \dots 1_{n_d})$ ile regresyon modelinin özel durumudur. Burada 1_m m boyutlu, 1 'lerden oluşan vektördür. Modelde (β, σ^2) için önsel dağılımlar tamamlanır. β_i 'ler, $i = 1 \dots d$ arasında önsel bağımsızlık olduğu düşünülebilir. Eğer d grubun benzer özellikleri varsa, makul bir alternatif ortalamanın, ortalamalardan oluşan bir kitleden çekildiği düşünülebilir. Bu kitle hipotetiktir ve homojen olduğu varsayılır. Burada ele alınan kitlenin normal bir kitle olduğu varsayıldığında, $\beta_1 \dots \beta_d$ $N(\mu, \tau^2)$ 'den alınmış bir örneklem olur. Modelde (μ, τ^2) için önsel dağılımla

tanımlanır. Modelde ele alınan önsellerin tümü hiyerarşik bir yapıda aşağıda verilmiştir:

$$\begin{aligned} \text{Birinci aşama} & : & \beta | \mu, \tau^2 & \sim N(1_d \mu, \tau^2 I_d) \\ \text{İkinci aşama} & : & \mu & \sim N(b_0, \beta_0) \\ \sigma^2 & \sim F_\sigma & \text{ve } \tau^2 & \sim F_\tau \end{aligned}$$

F_σ ve F_τ bağımsız dağılımlardır. Model parametreleri $(\beta, \mu, \sigma^2, \tau^2)$ için bileşik önsel yoğunluk fonksiyonu şu şekilde yazılabilir:

$$P(\beta, \mu, \sigma^2, \tau^2) = p(\beta | \mu, \tau^2) p(\mu) p(\sigma^2) p(\tau^2) \quad (2.19)$$

Eşitlik (2.19)'da verilen önsel dağılım iki aşamada verilmiştir. Bu model pekçok farklı durum için genelleştirilebilir. Normal regresyon modeli için genelleştirme Lindley ve Smith (1972) tarafından aşağıdaki gibi verilmiştir (Gamerman, 1997):

$$Y | \beta_1, \phi \sim N(X_1 \beta_1, \phi^{-1} I_n)$$

$$\beta_1 | \beta_2 \sim N(X_2 \beta_2, C)$$

$$\beta_2 \sim N(b, B)$$

$$\phi \sim \text{Gamma}(n_0/2, n_0 \sigma_0^2/2) \quad (2.20)$$

Yukarıdaki modelde önsel iki aşamada verilmiştir. Probleme bağlı olarak modelin uygun açıklamasını yapabilmek için daha çok aşama gerekebilir. Yeni eklenen denklemlerle biçim değişmeden kalır.

$$\beta_j | \beta_{j+1} \sim N(X_{j+1} \beta_{j+1}, C_j)$$

Genelde aşama sayısı arttıkça, dağılımları belirlemek de zorlaşır. Nadiren, modellerin üçten fazla aşaması vardır ve zaten yaygın bir düşünce de yüksek aşamadaki önsellerin fazla bilgi içermediğidir.

Eşitlik (2.20)'de verilen C ve B matrislerinin bilinen matrisler olduğu varsayılmıştır. Bu varsayım genelde mantıklı bir varsayım değildir. C ve B varyansları olabirliğin saçılımına göre önsel dağılımın saçılımını verir. Dolayısıyla analizler hala C ve B matrislerinin değeri üzerinde koşullu kalırlar. Bu iki matrisin bilinmediği durumda yapılacak işlemler Gamerman ve Migon tarafından (1993)'da verilmiştir.

Aşağıdaki işlemler kolay anlaşılması için iki aşamalı model üzerinde yoğunlaşmıştır. Fakat k-aşamalı modeller için de aynı şekilde uygulanabilir. Belirtilmesi gereken ilk nokta problemdeki bütün değişkenlerin aşağıda verilen bileşik dağılım yapısına etkisidir.

$$p(y, \beta_1, \beta_2, \phi) = p(y | \beta_1, \phi) p(\beta_1 | \beta_2) p(\beta_2) p(\phi)$$

Modelin hiyerarşik karakteri koşullu yapıları belirlemede oldukça başarılıdır. Ancak, hesaplamalar hala kolay değildir ve β_1 ve ϕ 'nin marjinal sonsal dağılımlarının elde edilmesi çoğu kez mümkün değildir. Kass ve Steffey (1989), George, Markov ve Smith (1993) çalışmalarında hiyerarşik modellerdeki bu sorunları incelemişlerdir. Albert (1996) birinci aşama önsel dağılım için diğer eşlenik olmayan durumların kullanılmasının mümkün olup olmadığını tartışmıştır. Hiyerarşik modellerin daha geniş bir uzantısı ise, Albert (1998) tarafından incelenmiştir (Gamerman, 1997). Eşitlik (2.20)'de tanımlanan iki aşamalı regresyon modeli ve genel doğrusal modeller, genel doğrusal hiyerarşik modellere ulaşmak için kullanılabilir.

$$y_i | \mu_i \sim EF(\mu_i), \quad i = 1 \dots n$$

$$\eta = X_1 \beta_1$$

$$\beta_1 | \beta_2 \sim N(X_2 \beta_2, C)$$

$$\beta_2 \sim N(b, B)$$

(2.21)

Burada $\eta = (\eta_1 \dots \eta_n)$ ve $\eta_i = g(\mu_i)$, $i = 1 \dots n$ olup, EF ile ifade edilen dağılımın normal olduğu düşünülürse, model 2.20 elde edilir.

2.5 Dinamik Modeller

Zamana göre değişen parametrelili, zaman serisi modellemesi ve regresyona uygun büyük bir modeller sınıfı ilk kez Harrison ve Stevens (1976) tarafından verilmiştir ve geliştirilen bu modeller West ve Harrison (1997) çalışmalarında ayrıntılı olarak anlatılmıştır. Dinamik modeller gözlem denklemi ve sistem denklemi olmak üzere iki denklem ile aşağıdaki şekilde tanımlanır:

$$y_t = F_t' \beta_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_t^2) \quad (2.22)$$

$$\beta_t = G_t \beta_{t-1} + w_t, \quad w_t \sim N(0, W_t) \quad (2.23)$$

Burada β_t , $n \times 1$ boyutlu süreç parameter vektörü;

F_t , t zamanında bilinen $n \times 1$ boyutlu bağımsız değişkenler vektörü;

G_t , $n \times n$ boyutlu sabit geçiş (sistem) matrisi;

ε_t ve w_t , sırasıyla sıfır ortalamalı ve değerleri t noktasında bilinen σ_t^2 ve W_t varyanslı normal dağılıma sahip rasgele değişkenler;

y_t , t noktasında sürecin gözlemidir.

Ayrıntıları West ve Harrison (1997)' de verilen dinamik doğrusal modeller aşağıdaki özellikler ile tanımlanır:

- Parametrik bir model yapısına sahiptirler,
- Her t noktasında parametreler ve gözlemlere ilişkin olasılıksal bilgiler (önsel ve sonsal dağılımlar) vardır,
- Parametrelerin zaman boyunca gelişimini açıklayan ardışık bir model yapısına sahiptirler,
- Öngörüler olasılık dağılımları biçiminde elde edilirler (Ergün, 1995).

Dinamik doğrusal model, nitel ve nicel bilgilerin birleştirilmesiyle çok boyutlu parametrelili önsel dağılımın belirlenmesi için güzel bir örnektir. Sistem denklemleri parametreler arasındaki ilişki hakkında nitel bilgi sağlar. Nicel bilgi ise, β_1 'in önsel dağılımından ve W_t değişim hatalarından elde edilir. Önsel dağılımın tam tanımı bu iki bilginin birleşimiyle ortaya çıkar.

En basit zaman serisi modeli birinci dereceden modeldir. Bu model literatürde durgun (steady) model olarak adlandırılır ve aşağıdaki şekilde tanımlanır:

$$\begin{aligned} Y_t &= \beta_t + \varepsilon_t, & \varepsilon_t &\sim N(0, \sigma_t^2) \\ \beta_t &= \beta_{t-1} + w_t, & w_t &\sim N(0, W_t) \end{aligned}$$

Burada sistem denklemleri basit bir rasgele yürüyüş modelidir. Durgun model, birinci derece serinin zaman trendini gösteren düz fonksiyona Taylor serisi yaklaşımı olarak da düşünülebilir. Bu model stok kontrolü, üretim planı ve finansal veri analizi için uygundur. Gözlem ve sistem varyansı zaman içinde değişebilir. Bu da sistemin değişkenliğini açıklamada büyük bir çalışma alanı sağlar.

Doğrusal büyüme modeli, zamana göre değişen serinin büyüme düzeyini gösteren fazladan bir β_2 parametresi olduğu için daha uğraştırıcıdır. Doğrusal büyüme modeli aşağıdaki denklemlerle tanımlanır:

$$\begin{aligned} Y_t &= \beta_{1,t} + \varepsilon_t, & \varepsilon_t &\sim N(0, \sigma_t^2) \\ \beta_{1,t} &= \beta_{1,t-1} + \beta_{2,t} w_{1,t} \\ \beta_{2,t} &= \beta_{2,t-1} + w_{2,t}, & w_t = (w_{1,t}, w_{2,t})' &\sim N(0, W_t) \end{aligned}$$

Bu modelde $F_t = (1, 0)$ ve

$$G_t = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \text{ biçiminde yazılabilir.}$$

F_t ve G_t 'nin seçimi istenen modele ve tanımlanmak istenen serinin doğasına göre yapılır. Modelin tam olarak belirlenmiş olması için, Eşitlik (2.22) ve (2.23)'deki σ_t^2 ve W_t varyanslarının tam olarak tanımlanması gerekir. Genelde zamana göre değişmez oldukları farz edilir ve σ_t^2 uygulamaya giren W_t 'den daha büyük kabul edilir. Kesin çıkarsamalarda ise, W_t çoğu kez bilinmez ve tahmin edilmesi gerekir. Bu durum işlemleri imkansızlaştırır. Bu durumdan korunmak için, öznel iskonto faktörünün kullanılması önerilir (West ve Harrison, 1997).

2.5.1 Ardışık Çıkarsama

Dinamik modellerin temel görüşlerinden biri, herhangi bir t zamanında çıkarsamaların $\beta_t | y^t$ 'nin dağılımına dayandırılmasıdır. Ardışık çıkarsamada bilgiler zaman içinde yenilenir. Burada üç temel işlem yer alır; gelişme; tahmin ve güncelleme. Tanımlanan sıraya göre aşağıdaki işlemler verilmiştir:

$t-1$ zamanında β 'nin sonsal dağılımının $\beta_{t-1} | y^{t-1} \sim N(m_{t-1}, c_{t-1})$ olduğunu düşünelim. Sistem denklemini $\beta_t | \beta_{t-1} \sim N(G_t \beta_{t-1}, w_t)$ olarak yazılabilir. Sistem denkleminde, t noktasında β 'nin önsel dağılımı aşağıdaki biçimde elde edilir:

$$\beta_t | y^{t-1} \sim N(a_t, R_t) \quad (2.24)$$

Burada, $a_t = G_t m_{t-1}$ ve $R_t = G_t C_{t-1} G_t' + w_t$ dir. Modelden bir adım ileri öngörü dağılımı aşağıdadır:

$$y_t | y^{t-1} \sim N(f_t, Q_t) \quad (2.25)$$

Burada $f_t = F_t' a_t$ ve $Q_t = F_t' R_t F_t + \sigma_t^2$ dir.

t noktasında y_t değeri elde edildiğinde, Bayes teoremi ile model parametresi β hakkındaki bilgiler aşağıdaki şekilde güncelleştirilir.

$$p(\beta_t | y^t) = p(\beta_t | y_t, y^{t-1}) \alpha p(y_t | \beta_t) p(\beta_t | y^{t-1})$$

Güncellenen sonsal dağılım aşağıda verilmiştir:

$$\beta_t | y^t \sim N(m_t, C_t) \quad (2.26)$$

Burada $m_t = a_t + A_t e_t$ ve $C_t = R_t - A_t A_t' Q_t$ ve $A_t = R_t F_t' / Q_t$ ve $e_t = y_t + f_t$ 'dir.

Bu bölümde Bayesci yaklaşımla ilgili bazı temel bilgilere yer verilmiş, modelleme çalışmalarında sıkça kullanılan bazı modeller ve bu modellerde karşılaşılan sorunlar incelenmiştir. İzleyen bölümde sayısal çözümlere esaslı modelleme çalışmaları tanıtılacaktır.



3. BAYESCİ GRAFİK MODELLEME

3.1 Giriş

Bu bölümde, analitik yolların yetersiz olduğu durumlarda kullanılan Bayesci yöntemler ele alınacak ve Bölüm 2'de de zaman zaman irdelenen parametre boyutunun artması ile ortaya çıkan sorunların, son yıllarda geliştirilen bu yöntemlerle nasıl çözümleneceği açıklanacaktır.

Çok geniş bir yelpaze içeren bu bölümde, Grafik modeller ve bu modellerle ilgili bazı uygulamalar Kesim 3.2'de verilecektir. Analitik yaklaşımlara bir alternatif olan Monte Carlo Markov Zinciri simülasyonu ise Kesim 3.3'te anlatılacaktır. Gibbs örnekleme Kesim 3.5 ve tez çalışmasının uygulama bölümünde kullanılan BUGS istatistik paketi ise Kesim 3.6'da kısaca tanıtılacaktır.

3.2 Bayesci Grafik Modellerin Tanıtımı

Modern uygulamalı istatistik son yıllarda birçok alanda ileri düzeyli kavramsal ve hesapsal araçlar kullanmaktadır. Bu alanlar;

- (a) Genelleştirilmiş doğrusal modellerin bir uzantısı olan ve daha karmaşık bağımlılık yapılarını içeren modellerin ele alınması (örneğin hiyerarşik modeller, dinamik doğrusal modeller veya yapay sinir ağları),
- (b) Gözlenen ve gözlenmeyen miktarlar için olasılık modelleri inşa eden Bayesci yöntemlerin kullanılması ve ardından verilerin ışığı altında bu miktarlara ilişkin koşullu olasılık dağılımları hakkında çıkarımlar yapılması,
- (c) Bayesci ve olabilirlik temelli yaklaşımların her ikisini de MCMC (simülasyon temelli) yöntemleriyle birlikte kullanarak sayısal integrallerin uygulanması ve
- (d) Çok yoğun kullanılmamakla beraber, bir modelin koşullu bağımsızlık yapısının bir gösterimsel sunumu olan grafiklerin kullanılmasıdır.

Tez çalışmasının amacı, yukarıda belirtilen dört çalışma alanının Bayesci grafik modelleme başlığı altında nasıl bir araya getirildiğini araştırmak; parametre ve gözlemler için koşullu bağımsızlık varsayımlarının varlığını incelemek ve genel

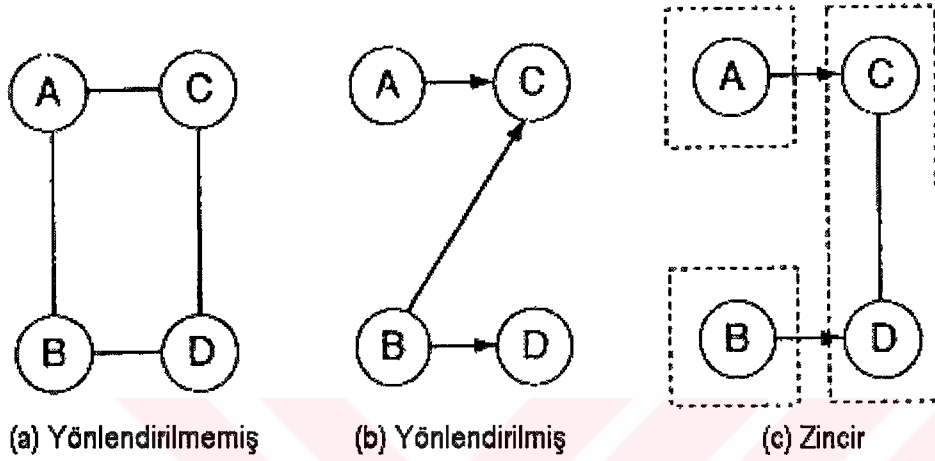
olarak MCMC metotlarına dayalı bir hesaplama stratejisinin nasıl işlediğini araştırmaktır.

Bu dört alanın bir araya getirilmesi yeni bir düşünce değildir. Grafikselleştirme sunumları kullanılarak olasılıkların verimli olarak hesaplanabilmesi için yöntemler son yıllarda geliştirilmiştir. Lauritzen ve Spiegelhalter (1988), Lauritzen ve Wermuth (1989), Pearl (1988), Lange (1992) ve Gilks (1993) çalışmaları bu bakış açısına yönelik ön adımlardır. Gilks'in (1994) BUGS projesi de bu yakınsamaya dayalıdır. Burada ele alınan her bir alan kendi içinde çok geniş bir literatüre sahip olduğundan, bu çalışmada sadece temel ve önemli olan kaynaklar verilecektir. Her bir makalede ele alınan problem kısaca aşağıdaki gibi tanımlanabilir. Çalışmalarda bağımlılık yapısı grafikselleştirilerek çok boyutlu dağılımlar sözkonusudur ve çalışmaların odak noktası, sağlanabilen tüm bilgiye dayalı olarak rastlantı değişkenlerinin bileşik dağılımının ya da marjinal dağılımlarının elde edilmesidir. Burada temel fikir, olasılıkları verimli olarak koşturmak için grafiğin topolojisini kullanmaktır. Tipik olarak bu grafikselleştirme sunuma "ağ" da denilebilir. Buna göre "ağ", rastlantı değişkenlerini gösteren düğümlerin ve bu düğümleri birbirine bağlayan ayrıtların bütünüdür. Grafikselleştirme modeller bir istatistiksel modele temel teşkil eden koşullu bağımsızlık varsayımlarını şekilsel olarak temsil ederler ve üç önemli fonksiyonu yerine getirirler. Burada birinci fonksiyon, modelin nitel, cebirsel olmayan esaslarının ulaşılabilir bir tanımının elde edilebilmesi; ikincisi, daha ileri koşullu bağımsızlık ifadelerini kanıtlamak için biçimsel bir temelin elde edilmesi ve üçüncüsü, çıkarsama problemlerine hesaplanabilir çözümlerle doğrudan bir bağlantının sağlanabilmesidir. O halde bir grafik model, basit olarak koşullu bağımsızlık varsayımlarını temsil eden bir araçtır. Öyle ki, David (1979)'a göre, modele ilişkin özellikler koşullu bağımsızlık aksiyomları yerine grafik modellerle tanımlanabilir (Pearl, 1988). Bununla birlikte, nitel varsayımların etkileyici bir görsel sunumunu sağlamanın bazı avantajları da vardır. Bu sayede, istatistiksel modelin yapısı ve anlamı matematiksel ifadelerle gerek kalmadan tartışılabilir ve anlatılabilir.

Bir grafik modelde, her bir rastgele miktar bir düğümlerle gösterilir ve bu düğümler birbirine yönlü veya yönsüz bağlarla bağlanır. Bu gibi bağların varlığı/yokluğu koşullu bağımsızlık varsayımlarını gösterir ve grafik tümüyle yönlü veya karışık

(zincir) olabilir. Bu kesimde, grafik modellerin kısa ve teorik olmayan bir incelemesi yapılacaktır. Daha detaylı bilgi için Lauritzen (1989), Whittaker (1990) ve Frydenberg (1990) çalışmalarından yararlanılabilir.

Grafik modellere ilişkin bazı örnekler aşağıdaki şekillerle verilmiştir:



Şekil 3.1: Farklı Grafik Modeller

Burada V , bir grafiğin düğümleriyle temsil edilen bir rastgele miktarlar kümesi olsun. Şekil 3.1'de gösterilen 3 farklı grafiğin her biri V 'nin öğelerini ilgilendiren koşullu bağımsızlık varsayımlarının bir kümesini temsil ediyor olarak yorumlanabilir. Burada ilgililenen konu, V 'de yer alan rastlantı değişkenlerinin çarpanlara ayrılmış biçiminin ve dolayısıyla bileşik dağılımın ayrıştırılarak tanımlanmasıdır. Şekil 3.1'de (a) ile tanımlanan grafik, yönsüz grafik olarak adlandırılır. Literatürde yönsüz grafikler, Gaussian dağılımlarda veya logaritmik-doğrusal modellerin kullanıldığı çok değişkenli analizlerde ilk kez ortaya çıkmıştır (Whittaker, 1990). Burada elde edilen yönsüz bağlantılar, grafiğin en yakın komşuları arasındaki birleşimleri ve korelasyonları temsil etmektedir. Bu grafiği, V değişken kümesi ile ilgili farklı koşullu bağımsızlık varsayımlarını göstermek için kullanabiliriz. Örneğin : "yerel Markov özelliğine göre, grafikte en yakın komşuları verilen bir düğüm, diğer tüm düğümlerden bağımsızdır (Bu özellik, Gibbs örnekleme uygulanırken Markov özelliklerinin kısmen uygun olduğunu göstermektedir). Bundan dolayı Şekil 3.1(a)'da A, B ve C'ye koşullu olarak D'den

bağımsız kabul edilmektedir ve Dawid (1990) notasyonu ile aşağıdaki şekilde ifade edilmektedir (Spielgelhalter ve arkadaşları, 1996):

$$A \perp D \setminus B, C$$

Burada \perp , bağımsızlık (diklik) ifadesi için kullanılmıştır. Bileşik olasılık yoğunluk fonksiyonu $p(V)$ 'nin her yerde pozitif olması sağlandığında, yerel Markov özelliği, hem koşullu bağımsızlığın diğer olası grafik yorumlarına, hem de $p(V)$ 'in aşağıdaki çarpanlara ayrılmış haline denktir. Eğer C , tam bağlantılı alt grafikler kümesi olarak tanımlanırsa, bu durumda;

$$p(V) = \prod_{C \in c} \phi_C(v_c)$$

elde edilir. Bu nedenle grafik, bileşik dağılımın tümü birleştirilmiş olan düğüm grupları üzerinde tanımlı, tek olmayan $\phi_C(v_c)$ terimlerine bölünebildiği bir "Markov Sahası" tanımlanmıştır. Buna göre, Şekil 3.1(a)'dan aşağıdaki faktörizasyon elde edilmiştir:

$$p(V) = \phi(A, B)\phi(B, D)\phi(A, C)\phi(C, D)$$

Yönlü grafikler, Şekil 3.1(b)'de verilmiştir. Bu grafikler, ağ analizinde yıllardır kullanılmakta olup, yakın geçmişte "olasılıksal uzman sistemler" başlığında geniş olarak yer almıştır (Pearl, 1988; Lauritzen ve Spiegelhalter, 1988).

Zincir Grafikleri ise, Şekil 3.1(c)'de yer almıştır. Bileşik dağılımın tanımı için kullanılan bu tür bir grafik, döngüsel olmayan bir grafik adını da alır. Bir diğer ifade ile grafikte yönlü daireler (döngüler) bulunmamaktadır. Bu; "aile", "torun", "ata" gibi soyağacıyla birleşime tanıklık yapan belirli terimlerin yorumlanmasını sağlar. Döngüsel olmayan grafik ile ilgili bir diğer örnek, Şekil 3.2'de verilmiştir.

Grafiklerde gösterilen koşullu bağımsızlık varsayımları, etki ailesi bilinen herhangi bir 'v' düğümünün, ataları olmayan diğer tüm düğümlerden bağımsız olduğunu

T.C. YATIRIM MENKUL DEĞERLER A.Ş.
BÜYÜKSAKALYOLU
MİLLİ MECLİS BİNASI
KAT: 5
SİGORTACI YOLU
E-10660
06530

gösteren “yerel bir Markov özelliği” ile tam olarak açıklanmaktadır. Örneğin, Şekil 3.1(b)'deki bu durum aşağıdaki ifade ile belirtilmiştir:

$$D \perp\!\!\!\perp A \setminus C, B$$

Buna göre, bileşik olasılık yoğunluk fonksiyonu aşağıdaki biçimde parçalanarak ifade edilebilir:

$$p(V) = \prod_{v \in V} p(v \setminus v' \text{nin ailesi}) \quad (3.1)$$

Eşitlik (3.1)'e ilişkin daha genel bir ifade aşağıdaki gibi elde edilir:

$$p(V) = \prod_t p(\Delta_t \mid \Delta_t' \text{nin ailesi}) \quad (3.2)$$

Bu grafiklerde yer alan doğrudan veya doğrudan olmayan bağlar, nihai aşamada birleşirler ve bu nedenle bu grafikler, sosyal bilimlerde yer alan önemli hipotez modellemelerinde potansiyel bir kullanıma sahiptirler.

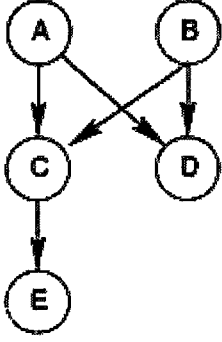
Burada Δ_t ailesine mensup olan herhangi bir u 'dan, Δ_t 'nin bir ögesine yönlü bir bağlantı varsa, modeldeki her bir blok, yönlü döngüsel olmayan (acyclic) bir grafikte tek bir düğüm gibi görülebilir. Bu durumda Eşitlik (3.2)'deki her blok aşağıdaki şekilde ifade edilebilir:

$$p(\Delta_t \mid \Delta_t' \text{nin ailesi}) = \prod_{C \in C_t} \phi_C(v_C) \quad (3.3)$$

Sonuçta, bileşik dağılım için aşağıdaki çarpanlara ayrılmış ifade elde edilir:

$$p(v) = p(A) p(B) p(C, D/A, B)$$

Burada, $P(C, D/A, B) = \phi(A, B) \phi(B, D) \phi(A, C) \phi(C, D)$ olarak yazılabilir. Yönlendirilmiş döngüsel olmayan grafik modele ilişkin ikinci bir örnek ise Şekil 3.2'de verilmiştir:



Şekil 3.2: Yönlendirilmiş Döngüsel Olmayan Grafik

Bütün bağlantıların yönlendirildiği ve hiçbir döngüye izin verilmeyen yönlendirilmiş döngüsel olmayan grafikte yer alan ve düğümlerle temsil edilen A, B, C, D ve E rastgele miktarları Şekil 3.2'de görülmektedir. Burada, A ve B büyükanne ve büyükbaba; C ve D bunların çocukları ve E' de, C'nin çocuğunun genleridir. Genetik benzeşim doğal olarak çocuklar, atalar ve torunlar gibi terimlerin belirli kullanımını sağlar ve yönlendirilmiş bağlantılar Mendel'in kendi standart kalıtım yasalarının bir kişinin genetik özelliklerinin sadece ailesininkilere dayandırıldığı nitel bakış açılarını göstermektedir. Bundan dolayı eğer bir kişinin ailesinin genleri tamamıyla biliniyorsa, o kişi bir torun olmadıkça başka hiçbir şey o kişinin genleri hakkında herhangi bir bilgi vermeyecektir. Buna göre, bilinmeyen miktar 'v' ile ilgili olasılıksal bir yapı incelendiğinde, bu değişkeni temsil eden bir düğümün ailesi şayet biliniyorsa bu durumda bu düğümün diğer tüm torun olmayanlardan koşullu olarak bağımsız olacağını yorumlamak en doğal bakış açısı olacaktır. Bu basit yorumun büyük bir gücü olduğu ortaya çıkmaktadır. İlk olarak David tarafından 1979 yılında bu düşünce bir grup koşullu bağımsızlık aksiyomlarına uygulanmıştır (Spielgelhalter, 1998). Buna göre, Şekil 3.2'den $E \perp (A, B, D) / C$ olduğunu görebiliriz. Burada \perp koşullu bağımsızlık anlamındadır ve bu ifade E düğümünün C'nin varlığı altında A,B ve D'den bağımsız olduğunu ifade etmektedir. Burada kullanılan koşullu bağımsızlıkları okuma kuralları Pearl (1988) ve Lauritzen (1990) tarafından türetilmiştir. Bu durumda Şekil 3.2'de verilen modele ilişkin bileşik dağılım grafikteki yerel bağımlılıkları tanımlayan bir dizi terimin çarpımı olarak aşağıdaki şekilde ifade edilebilir:

$$P(A,B,C,D,E) = P(A).P(B).P(C|A,B).P(D|A,B).P(E|C)$$

Bileşik dağılımın bu şekilde ifade edilebilmesi MCMC metotlarının kullanımında son derece yararlı olacaktır.

Bu çalışmada ele alınan grafik modeller Bayesci idolojiyi benimseyen ve kullanan grafik modeller olmasına rağmen, klasik grafik modeller de mevcuttur. İzleyen alt kesimde klasik grafik modellerine ilişkin kısa bir literatür bilgisine yer verilecektir. Burada yer alan kaynaklar Spiegelhalter (1998) çalışmasında bahsedilen çalışmalardır.

3.2.1 Klasik Grafikselle Modeller

Klasik yönlendirilmiş grafiklerin, gözlemlenebilir rastgele miktarların bağımlılık yapılarını resmetmede kullanımı ilk olarak Path analizi diyagramlarına dayanır. Bu grafiklerin yapısal modellerin yorumlanmasında uzun bir geçmişi vardır. En son gelişmeler ise, değişkenler arasındaki yönlendirilmemiş birleşimlerin incelendiği yönlendirilmemiş grafiklerde (Whittaker, 1990) ve yönlendirilmiş ve yönlendirilmemiş bağlantıların karışımından oluşan zincir grafiklerde (Frydenberg, 1990; Lauritzen ve Wermuth, 1990) yoğunlaşmıştır. Bu modellerde de burada incelenen koşullu bağımsızlık yorumlamaları vardır. Burada geleneksel bir yaklaşım olan en çok olabilirlik yöntemi ile modeller tahmin edilmekte ve seçenek yapılar olabilirlik oranı testleri ile karşılaştırılmaktadır. Klasik yaklaşıma ilişkin en son gelişmeler için Lauritzen (1986)'da verilmiştir. Olasılıksal uzman sistemleri de, Bayesci ağlar olarak bilinen yönlendirilmiş grafik gösterimini değişkenler için kullanmışlardır (Pearl, 1988; Lauritzen ve Spiegelhalter, 1988). Burada amaç mevcut gözlemlerden yola çıkarak gözlemlenmemiş miktarların koşullu olasılık dağılımını hesaplamaktır. Pearl (1988) çalışması birçok hayati gelişmeye ve anlayışa yol açmıştır. Bu gelişmeler, grafik üzerinde yerel hesaplamalara dayanarak koşullu olasılıkların bulunması için geliştirilen teknikleri, bir model belirlemek ve bu model üzerinde çıkarsamalar yapmak amacıyla kod geliştirmek için basit bir grafik arabirimi kullanmayı ve son olarak koşullu olasılık dağılımlarının parametrelerinin de modeldeki düğümler olabildiğinin anlaşılması çalışmalarını kapsamaktadır.

3.2.2 Bayesci Grafik Modelleri

İkinci bölümde belirtildiği gibi, Bayesci yaklaşım, bütün bilinmeyen miktarların rastlantı değişkenleri olarak kabul edilmesi ilkesine dayanır. Bu nedenle, Bayesci grafik modellerinde potansiyel değişkenler ve gözlemlenebilir miktarlar gibi parametrelerin de grafikte düğümler olarak yer alması doğal bir adımdır. Birçok araştırmacı, bu gösterimi karmaşık problemlerin çözümü için benimsemiştir. Hiyerarşik modeller için Lange (1992), log-lineer modeller için Madigan ve York (1994), Kalman filtresi için Normand ve Tritchler (1992), model tanımlama ve çıkarsama problemleri için Spiegelhalter (1998), ölçüm hatası içeren problemleri için Richardson ve Gilks (1993) çok değişkenli zaman dizileri için Queen ve Smith (1993), uzaysal haritalama modelleri için Bernardinelli ve arkadaşları (1997), durum-kontrol çalışmaları için Breslow (1996), parametrik olmayan riziko regresyonu için Arjas ve Liu (1996) çok tepeli dağılımlarda örnekleme çalışmaları için Abraham ve Haines (1999) bu duruma örnek çalışmalar olarak verilebilir. Gilks ve arkadaşları, 1996'daki çalışmasında problemin yapısını açıklamak için Bayesci grafikleri kullanmıştır. Buntine (1994) çalışmasında, yapay sinir ağlarını da kapsayan birçok standart istatistiksel model için grafiksel yorumlar açıklamıştır.

Başlangıçta, standart bir Bayesci analizin grafik kuramına ne eklediği açık olmayabilir. Herhangi bir model grafik ile gösterilebilir. Ancak Bayesci yaklaşım ile, koşullu bağımsızlık varsayımlarının yapısı daha kolay bir şekilde çözümlenebilmekte ve dolayısıyla büyük bir denklem kümesine başvurmadan sonuca daha kolay ulaşılabilmektedir. Bu bağlamda, yerel hesaplama metoduna doğrudan bağlantı kurulabilir. Dinamik doğrusal modeller için grafik modelin nasıl kurulacağı bir örnek olarak izleyen kesimde kısaca verilecektir.

3.2.2.1 Dinamik Doğrusal Modelin Grafiksel Sunumu

Bir Bayes ağı, yönlü bağlantıları nedensel etkiyi gösteren bir dairesel grafikdir. Bayes ağının amaçları, bir rastlantı değişkenleri kümesine bağımlılıkların kavramsal bir portresini sunmak ve grafikteki koşullu olasılıkların güncellenmesi için belirli bir metot sağlamaktır. Burada, Bayesci dinamik doğrusal model grafik

yardımıyla kavramsal olarak açıklanacaktır. Kesim 2.5'te ayrıntılı bir şekilde tanıtilan model, aşağıda verilmiştir:

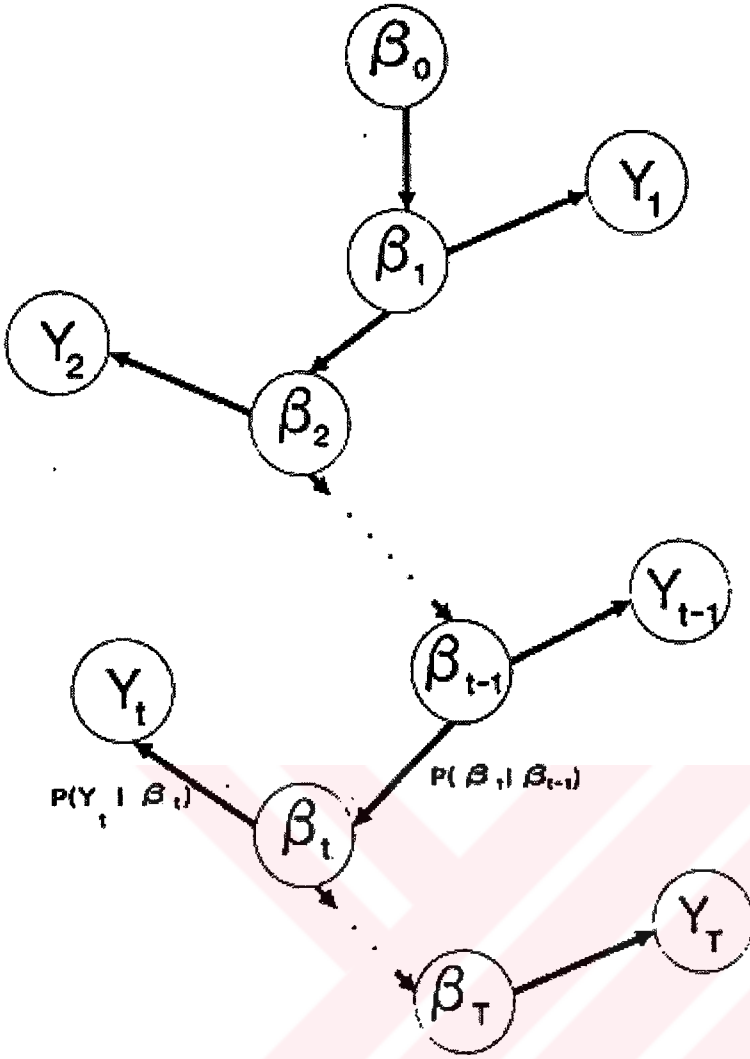
$$\text{Gözlem Denklemi} : y_t = F_t' \underline{\beta}_t + \varepsilon_t$$

$$\text{Sistem Denklemi} : \underline{\beta}_t = G_t \underline{\beta}_{t-1} + w_t$$

Burada, $t=1,2,\dots,T$ için $w_t \sim N(0, W_t)$ $\varepsilon_t \sim N(0, \sigma^2_t)$ ve w_t , ε_t 'den bağımsızdır. Bütün $t>0$ 'lar için, σ^2_t , F_t , Y_t ve G_t değerlerinin bilindiği varsayılmıştır. β_t parametreleri çıkarsamaların ilgi merkezleridir. Yukarıda tanımlanan sistemin, aşağıda belirtilen C_T rastlantı değişkenlerinin bir kümesi olduğu düşünülebilir:

$$C_T = \{ \{\beta_0 \rightarrow \beta_1\}, \{\beta_1 \rightarrow Y_1\}, \\ \{\beta_0 \rightarrow \beta_1\}, \dots, \{\beta_{T-1} \rightarrow \beta_T\}, \{\beta_T \rightarrow Y_T\} \}.$$

Burada elde edilen bu tanımlama dinamik modelin koşullu bağımsızlık yapısının anlaşılmasında önemli bir noktadır. Buna göre, dinamik doğrusal modelin grafiksel gösterimi Şekil 3.3'deki gibi elde edilir.



Şekil 3.3: Dinamik Doğrusal Modelin Grafiksel Sunumu

Benzer olarak β_{t-1} 'i β_t 'ye bağlayan bağlantı da sistem denkleminde elde edilir. Sistem denklemini aşağıdaki şekilde formüle edilebilir:

$$P(\beta_t | \beta_{t-1}) = \left(\frac{1}{\sqrt{2\pi}} \right)^p |W_t|^{-1/2} \exp \left\{ -\frac{1}{2} (\beta_t - G_t \beta_{t-1})' W_t^{-1} (\beta_t - G_t \beta_{t-1}) \right\}$$

Kök düğümüne ilişkin dağılım ise aşağıdaki gibi verilmiştir:

$$\pi(\beta_0) = \left(\frac{1}{\sqrt{2\pi}} \right)^p |S_0|^{-1/2} \exp \left[-\frac{1}{2} (\beta_0 - \mu_0)' S_0^{-1} (\beta_0 - \mu_0) \right]$$

Burada grafiksel olarak incelenen model için tahmin ve öngörüler yapılabilir. Normand ve Tritchler (1992) çalışmalarında dinamik model parametrelerine ilişkin

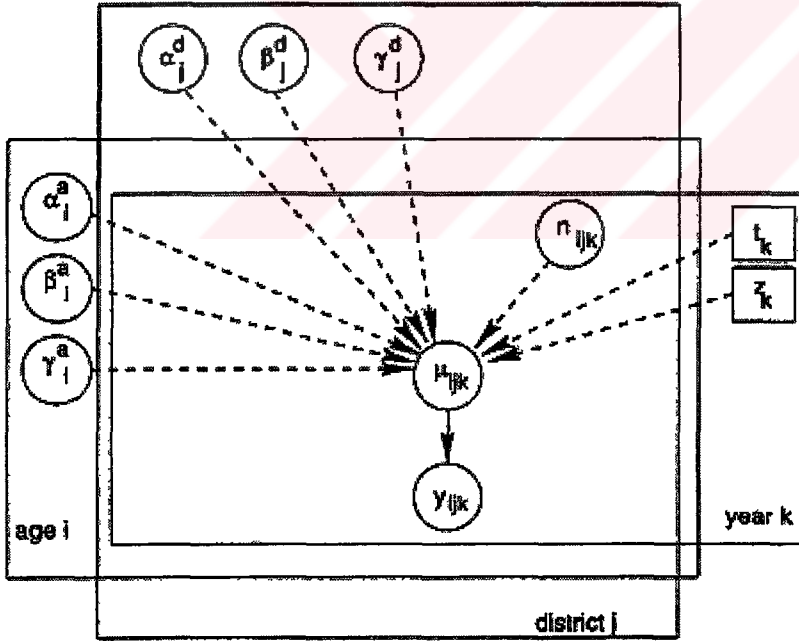
bu sonsal tahminlerin ve gözlemlere ilişkin öngörülerin nasıl elde edildiğini incelemiştir. Ancak yazarların kullandıkları yaklaşım tez çalışmasının hedeflediği noktadan uzak olduğu için burada verilmemiştir.

3.3 Bayesci Grafik Modellerinde Çıkarımlar

Bayesci grafik modellerde çıkarımda öncelikle modelin tasarlanması gerekir. Ardından hesaplama aşamasına geçilir. Hesaplama aşamasında analitik ya da benzetim tabanlı yöntemlere başvurulur.

3.3.1 Model Kurma

Burada genel yaklaşım, bütün miktarlar arasındaki nitel ilişkileri bir grafik ile göstermek ve sonra da ailesi verilen her bir düğümün koşullu dağılımlarını elde etmektir. Bu işlem, bileşik dağılımı elde edilmesi için gerekli terimleri sağlar.



Şekil 3.4: Örnek Bir Bayesci Grafiksel Model

Birçok istatistiksel örnekte, ele alınan modelin incelenmesi ve modelin düzeltilmesi hayati bir bileşen olmasına rağmen, genellikle, grafiğin yapısının sağlam bir bilgiye dayandığı şeklinde bir ön kabul vardır. Model yapısı hakkında kararsızlık olduğu

durumlarda arařtırmacıların başvurabileceđi birçok model seçimi tekniđi bulunmaktadır ve bunlara ilişkin adımlar Buntine (1994) alıřmasında verilmiřtir. Bazı durumlarda belirgin bir modelin seçilmesine yönelik olarak belirgin bir ihtiyaç yoktur. Bu gibi durumlarda, Bayes faktörlerinin seçenek grafik modellerin ortalamasının alınmasında nasıl kullanılabilceđi Madigan ve Raftery (1994)'da gösterilmiřtir (Spielgelhalter, 1998)..

Bayesci grafik modelde tüm miktarlara ilişkin bileřik dađılımın belirlenmesinden sonra, V kümesi, X ve θ olarak ikiye bölünebilir. Burada X , veri olarak gözlenen miktarlar olup; $\theta = \theta_1, \theta_2, \dots, \theta_p$ 'da kalan kısmıdır. ıkarsama ařamasında problem, verilen bir X için, Q 'nun tek tek bileřenlerinin ve belli fonksiyonlarının kořullu olasılık dađılımının nasıl hesaplanacađının belirlenmesidir. θ çođu zaman grafiđin üstünde ve altında yer aldıđından, bileřik dađılım arpanları, θ 'nın önsel dađılımı $p(\theta)$ ve $p(X|\theta)$ olabilirliđi olarak ikiye ayrılır. Bu nedenle ařađıda verilen sonsal dađılımı elde etmek için Kesim 2.2'de verilen Bayes Teoremi'nin kullanması gerekir.

$$p(\theta|x) \propto p(x|\theta) p(\theta)$$

Burada $p(\theta|x)$, bileřik sonsal dađılım olup; marjinal dađılımlar ařađıdaki integral yardımıyla elde edilir:

$$p(\theta_k \setminus x) = \int p(\theta \setminus x) d\theta_1 d\theta_2 \dots d\theta_{k-1} d\theta_{k+1} \dots d\theta_p.$$

Burada arařtırmacıların karřısına ıkan en önemli sorun, θ 'nın çok yüksek boyutlu olmasıdır. ıkarsama için herbir bileřenin yukarıdaki řekilde tek tek marjinal dađılımlarına ulařmak ve elde edilen bu dađılımlardan bilgiye inmek çođu zaman sıkıntılı bir süreç gerektirmektedir. Bu tür zor integral problemleri için birçok teknik kullanılmaktadır. Genel olarak Bayesci yaklařımda bileřik sonsal ve marjinal sonsal dađılımlara ulařmada birçok hesaplama yöntemi mevcuttur. Bunlar,

- Tam ıkarsamalar,
- Analitik Yaklařımlar ve
- Simülasyon Tabanlı özömlerlerdir.

Tam çıkarsamalarda, basit olabilirlikler ve standart eşlenik analizlerle, kapalı formda çözümlere ulaşılmanın mümkün olacağı söylenebilir. Örneğin, olasılıksal uzman sistemler bağlamı içinde, gözlemlenebilen X 'ler genellikle kesiklidir ve Q 'nın elemanları da multinomial dağılımlı parametrelerdir. Bu durumda tüm bilgilerin gözlemlenmesi ve bağımsız Dirichlet dağılımlarının öncelik olarak kabul edilmesi şartında basit çözümlerin var olabileceği Lauritzen ve Spiegelhalter (1988) çalışmasında gösterilmiştir. Benzer bir tam çıkarsamaya dayalı hesaplama yöntemi, eşlenik bağlamları normal olabilirlik fonksiyonları için Lindley ve Smith (1972) çalışmasında; Kalman filtresi modelleri için ise, Normand ve Tritchler (1992)'da görülebilir (Spiegelhalter, 1998). Ancak tam çıkarsamaya dayalı hesaplamaların gerçek uygulamalarda çok kullanışlı olmadığı söylenebilir. Bunun bir nedeni, uygulamalarda daha karmaşık model yapılarının karşıya çıkması ve Kesim 2.3'te de belirtildiği gibi, eşlenik analizlerin her zaman uygulanamaz olmasıdır.

Bu nedenle, Bayesci çıkarsama için esas olarak iki tür çözümlenmenin söz konusu olduğu söylenebilir. Bunlar, analitik çözümlenmeler ve benzetim tabanlı sayısal çözümlenmelerdir. Analitik yaklaşımla ilgili, özellikle hiyerarşik model sınıflarını ele almaya yarayan birçok çalışma mevcuttur (Gelman ve arkadaşları, 1992). Analitik çözümlenme **her zaman öncelikle** tercih edilmesi gereken bir yöntem olmalıdır. Ancak çoğu kez elde edilen bileşik dağılımların karmaşık olması, istenen sonuçlara analitik olarak ulaşılmasını zorlaştırmaktadır. İntegral güçlükleri nedeniyle, marjinal dağılımlara ulaşamamaktadır. Analitik yaklaşımların önünün tıkandığı bu durumlarda stokastik simülasyonlara başvurulabilir.

Bayesci grafik modellerde bilinmeyen parametre sayısı fazla iken, analitik yaklaşımlarla ulaşılabilen sonuçlara ulaşmanın bir diğer yolu simülasyon tabanlı hesaplamalara başvurmaktır. Tez çalışmasında Markov zinciri simülasyonu kullanılmıştır. Bu nedenle bu yöntem Kesim 3.4'te tanıtılacaktır.

3.4 Markov Zinciri Simülasyonu

Karmaşık modellerde, sonsal dağılımlardan örneklemin doğrudan elde edilmesi çok zor olduğundan, model parametrelerine ilişkin çıkarsama güçlükleri çoğu kez gündeme gelir. Düşük boyutlu problemlerde, bazı örnekleme tekniklerinin kullanılması ya da marjinal tepe değerlerine ilişkin yaklaşık tahminlerin elde edilmesi mümkündür. Burada elde edilecek yaklaşık tahminler için her bir düğümün çok değişkenli normal ya da t dağılımına sahip olduğu varsayılmalıdır (Gamerman, 1997). Ancak bu tür yakınsamaya dayalı simülasyonlar çıkarsama amacı için çok uygun olmamaktadır. Buna karşın, yakınsayan bu dağılımdan çekilecek bir örneklem, hedef dağılımı sağlayacak bir Markov Zincirini işletmek için uygun bir başlangıç noktası olarak kullanılabilir.

Karmaşık integrallerin çözümlenmesinde simülasyon bazlı metotlarının kullanımında son yıllarda büyük bir artış gözlenmektedir. Özellikle Markov Zinciri Monte Carlo (MCMC) yöntemleri en fazla kullanılan simülasyon yöntemleri olmuştur (örneğin Smith ve Roberts, 1993; Gilks ve arkadaşları, 1996a; Gamerman, 1997; Richardson ve Robert, 1999; Han ve Carlin, 2000; Brooks, Giudici ve Philippe, 2001). MCMC metotlarına bu ismin verilmesinin nedeni, bilinmeyen miktarlar için benzetilmiş değerlerin er geç durağan bir dağılımı (bu istenilen sonsal dağılım) bir Markov Zinciri ile takip etmesidir.

En basit anlamda MCMC metodunda her bir bilinmeyen parametre için başlangıç değerleri söz konusudur ve koşullu olasılık dağılımından her bir değişken sırayla benzetilerek grafikte taranır. Benzetilmiş her bir değer, eski değeri ile yer değiştirir. Bu işlemde yeterli sayıda iterasyonundan sonra, Markov Zinciri'nin durağan dağılımına erişildiği varsayılır ve kar düğümleri için gelecekte benzetilecek değerler izlenir.

Markov Zinciri Simülasyonu çok karmaşık bir teknik olmasına rağmen, birçok problemin (hiyerarşik modellerin sonsal dağılımları da dahil olma üzere) çözümünde kolay ve güvenilir sonuçları vermektedir. Burada temel ilke tekniğin doğru bir şekilde uygulanmasıdır.

Markov zinciri simülasyonundaki ana fikir, durağan bir dağılıma yakınsayan (θ uzayında) bir rastgele yürüyüş benzetimi oluşturmaktır. Buradaki durağan dağılım $p(\theta/y)$ bileşik sonsal dağılımıdır. Çalışmada $p(\theta/y)$ dağılımı ayrıca hedef dağılımı da olarak adlandırılacaktır.

Simülasyonun yeterince uzun işletimi dağılımdan çekilen örneklemin durağan dağılıma daha etkin bir şekilde yakınsamasını sağlayacaktır. Bu nedenle iterasyon sayısının fazla alınması, dağılımdan elde edilecek örneklemin hedef dağılıma yakınsamasını sağlayacaktır. Sonuç olarak Markov Zinciri simülasyon teknikleri Bayesci sonsal dağılımlardan örneklemelerin çekilmesi için kullanışlı tekniklerdir.

Markov Zinciri Simülasyon metotları arasında en çok kullanılanları aşağıdadır. Bunlar:

- Metropolis algoritması ve
- Gibbs örneklemesidir.

Tez çalışmasında sadece Gibbs örneklemesi ele alınmıştır. Burada bu tekniğin seçilmesinin ana nedeni grafik modellerin kurulmasında kullanılan BUGS istatistik paketinin Gibbs örneklemesine dayalı olmasıdır. Kesim 3.5'te Gibbs örneklemesi tanıtılacaktır.

3.5 Gibbs Örneklemesi

Gibbs örneklemesi kavramı uzaktan algılama süreçlerinden alınmış bir kavramdır. Bu kavramda Gibbs sözcüğü örnekleme için ilgili sonsal bir Gibbs dağılımıdır. Uzaktan algılama terminolojisine göre Gibbs dağılımı aşağıdaki şekilde tanımlanmıştır:

$$f(x_1, \dots, x_n) \propto \exp\left[-\frac{1}{kT} E(x_1, \dots, x_n) \right]$$

Burada k : pozitif bir değişmez,

T : İlgili sistemin sıcaklığı,

E : Sistemin enerjisi ve

x_i : Sistemdeki i. bileşen için ilgili olunan bir karakterdir ($i=1, 2, \dots, n$).

Buna göre x_i , i. parçacığın pozisyonu, yoğunluğu v.s. gibi tanımlamalar olabilir. E enerji fonksiyonu, potansiyel fonksiyon V'nin bir toplamı olarak belirtilmiştir. Bu toplamlar bölgesel özelliklere dayalı bir olasılık tanımlanmasına yol açmaktadır. Bu da bileşenler arasında uzaysal etkilenmelerin modellenmesinde çok kullanışlıdır.

Geman ve Geman (1984), bu modelleme problemini özellikle örnekleme şeması ve Markov rastgele alanları ile karşılaştırmak amacıyla ele almıştır. Geman ve Geman'ın örnekleme planı, kısmi tanımlanmayı kapsayan koşullu yapıları keşfetmiştir. Geman ve Geman'ın çalışması çok değerli bir çalışma olmasına rağmen, iyi bir istatistik dergisinde yayınlanamamıştır. Bu nedenle Bayesci problemlerin çözümlemesinde bu metodun tanıtılmasında bir gecikme olmuştur. Gelfand ve Smith (1990), Geman ve Geman'ın keşfettiği bu örnekleme planını istatistik camiasına ilk tanıtan yazarlar olmuştur. Bazı yazarlar, örneğin Robert (1994), Gibbs örneklemesinin kullanılmasının yanlışlığı (kavram açısından) neden olabileceğini savunmuş ve adının Bayesci örnekleme ile değiştirilmesini önermiştir. Gelfand ve Smith (1990) çalışmasında ayrıca Gibbs örnekleme planı ile veri çoğaltma algoritmalarını ve Sampling-Importance Resampling'i (Monte Carlo) karşılaştırmıştır (Gelman ve arkadaşları, 1992). Grafik modellerle Gibbs örnekleme arasındaki önemli bağlantı koşullu bağımsız ilkesidir. Grafik modeller ile Gibbs örnekleme arasındaki bağlantı ilk kez Pearl'in (1988) yapay zeka ile ilgili çalışmasında değinilmiştir.

Bir grafik modelde yer alan herhangi bir v düğümü için koşullu dağılım $p(v/v'$ 'nin ailesi) şeklinde ifade edilebilir. Böylece v'nin örnekleme yapılırken, sadece v'nin ebeveynleri, çocukları ve akrabaları olan düğümleri düşünmek durumundayız. Örneğin Şekil 3.2'de verilen grafik modelde, C miktarına ilişkin bir örnekleme istendiğinde aşağıda verilen ilişkinin göz önüne alınması gereklidir. Burada, C'yi içeren $p(C/A,B,D,E)$ şeklinde bir koşullu yapı, $p(C/A,B,E)$ ile ifade edilebilir. Bu nedenle D'nin değerinin konuyla herhangi bir ilgisi bulunmamaktadır.

Gibbs örnekleme algoritması ile bazı temel özellikleri izleyen alt kesimde tanıtılacaktır.

3.5.1 Tanımı ve Özellikleri

Gibbs örnekleme, geçiş (transition) olasılıkları tam koşullu dağılımlar ile biçimlendirilmiş bir Markov Zinciri Monte Carlo (MCMC) planıdır. İlgili olunan dağılım $\pi(\theta)$ olsun. Burada $\theta = \theta_1, \theta_2, \dots, \theta_k$ ' dir ve her $\pi(\theta_i)$ 'nin koşullu dağılımı, $\pi(\theta_i) = \pi(\theta_i / \theta_{-i})$, $i = 1, \dots, k$ şeklinde bilinen dağılımlar olarak ifade edilir.

Burada çözülmesi istenen problem, doğrudan veri üretme planlarının pahalı, karışık ya da müsait olmadığı zamanlarda π 'den örnekleme çekme (veri üretme) işlemidir. Gibbs örnekleme, tüm koşullu dağılımlar birbirini izleyecek (ardıl) şekilde bir veri üretme planına dayalı olan bir alternatif çözüm sağlar. Gibbs örnekleme aşağıda verilen adımlarla tanımlanır :

1-) $j = 1$ de iterasyonu başlatmak ve başlangıç değerlerini tanımlamak.

$$\theta^{(0)} = (\theta_1^{(0)}, \dots, \theta_k^{(0)})'$$

2-) $\theta^{(j)} = (\theta_1^{(j)}, \dots, \theta_k^{(j)})'$ 'nin yeni değerlerini ardışık olarak aşağıdaki şekilde elde etmek.

$$\theta_1^{(j)} \sim \pi(\theta_1 / \theta_2^{(j-1)}, \dots, \theta_k^{(j-1)})$$

$$\theta_2^{(j)} \sim \pi(\theta_2 / \theta_1^{(j)}, \theta_3^{(j-1)}, \dots, \theta_k^{(j-1)})$$

.

.

.

$$\theta_k^{(j)} \sim \pi(\theta_k / \theta_1^{(j)}, \dots, \theta_{k-1}^{(j)})$$

3-) Sayaç numarasını j'den j+1'e deęiřtirip, 2. adıma dnmek ve aynı iřlemlere yakınsama saęlanana dek devam etmek.

Yakınsama saęlandığı zaman, $\theta^{(j)}$ 'nin bu sonuç deęeri $\bar{\theta}$ 'den bir alıntı olacaktır. İterasyon sayısı arttıkça, bu zincir denge kořuluna yaklařır (limit durumunda denge daęılımına ulařıldıđı varsayılır). Dolayısıyla elde edilen yakınsamanın yaklařık olarak saęlandığı varsayılır.

zetlenirse, Gibbs rneklemesinde parametre vektr θ , k adet alt bileřene ya da alt vektrlere ayrılır. Gibbs rneklemesinin her bir iterasyonu θ 'nın alt vektr civarında olur, her bir alt set iin tm dięer deęerlere kořullu olarak rnekleme ekilir. Bu nedenle her t iterasyonunda k adım vardır ve her bir θ_j^t ařađıda, verilen kořullu daęılımdan ekilir.

$$P(\theta_j / \theta_j^{t-1}, y)$$

Burada,

$$\theta_j^{t-1} = (\theta_1^t, \dots, \theta_{j-1}^{t-1}, \theta_{j+1}^{t-1}, \dots, \theta_k^{t-1}) \text{ 'dir.}$$

Bylece her θ_j alt vektr, θ 'nın t. iterasyonda daha nceden gncelleřtirilmiř dięer bileřenleri ve t-1'deki deęerlerine kořullu olarak gncelleřtirilmektedir. Genel olarak, iteratif simlasyonlarda ıkarsama, dięer Bayesci simlasyonlardakine benzerdir. Buna gre bu tr hesaplamalarda da ama,

- $P(\underline{\theta} / y)$ 'den gelen tm ekimleri (simle edilmiř) kullanarak, sonsal yoęunluęu zetleme, eyrek deęerleri, momentleri ve dięer zet istatistikleri hesaplamak ve
- Gzlemlere iliřkin ngrler yapmaktır.

Gibbs rneklemesi ile ilgili basit bir rnek ařađıda verilmiřtir. Daha karmařık uygulama da bu rnekten sonra verilecektir.

Ortalaması (θ_1, θ_2) ve kovaryans matrisi $\begin{bmatrix} 1 & p \\ p & 1 \end{bmatrix}$ olan iki deęişkenli bir normal

dağılımdan gelen (y_1, y_2) gözlemlerini ele alalım. Burada model parametresi θ vektörü için uniform önseli düşünöldüğünde, θ 'ya ilişkin sonsal dağılım aşağıdaki şekilde ifade edilir:

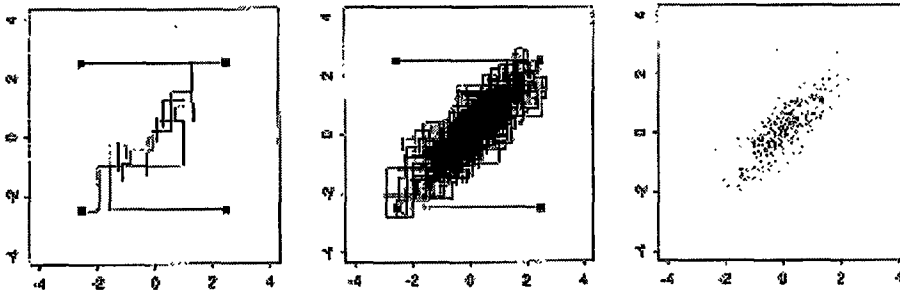
$$\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} / y \sim N \left(\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \begin{pmatrix} 1 & p \\ p & 1 \end{pmatrix} \right)$$

Burada elde edilen dağılımdan örnek elde edilmesi çok kolay olmasına rağmen Gibbs örnekleme metodunun işleyişinin görölməsi için Gibbs örnekleme örneği olarak uygulanabilir. (θ_1, θ_2) 'nin Gibbs örnekleminin oluşturulması için, koşullu sonsal dağılımlara ihtiyaç vardır. Bu dağılımlar aşağıdaki biçimlerde ifade edilmiştir:

$$\theta_1 / \theta_2, y \sim N(y_1 + P(\theta_2 - y_2), 1 - p^2)$$

$$\theta_2 / \theta_1, y \sim N(y_2 + P(\theta_1 - y_1), 1 - p^2)$$

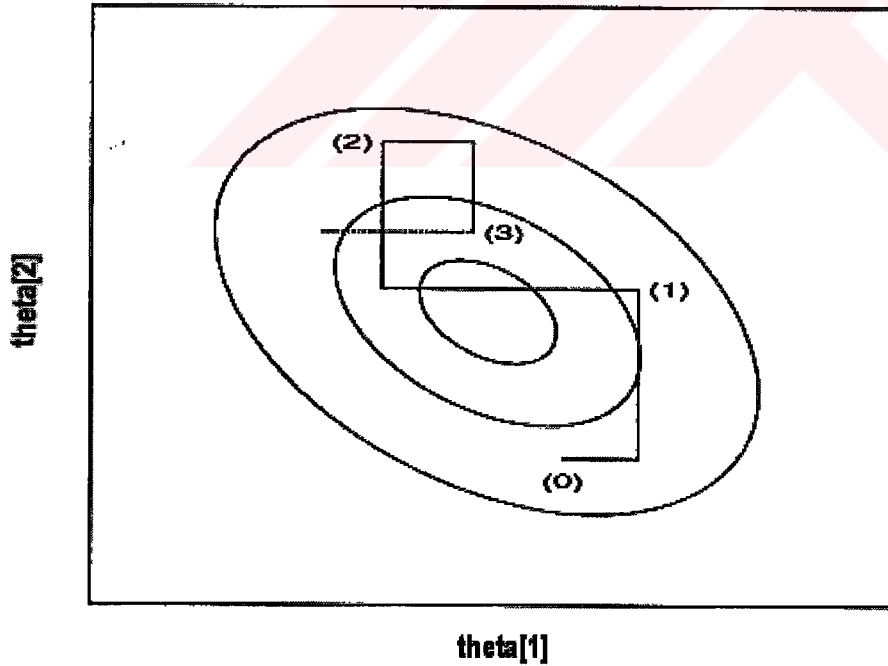
Buna göre Gibbs örnekleme yukarıda verilen normal dağılımlardan örnekleme oluşturur. $p=0.8$ ve $(y_1, y_2) = (0,0)$ olduğunda; 4 farklı bağımsız başlama noktasından (Bunlar: $(-2.5, -2.5)$; $(-2.5, 2.5)$; $(2.5, -2.5)$; $(2.5, 2.5)$) elde edilen sonuçlar aşağıdaki grafikte verilmiştir (Gelman ve arkadaşları, 1992):



Şekil 3.5: Dört Bağımsız Koldan Gibbs Örneklemesi

π ile ifade edilen dağılımdan n büyüklüğünde bir örneklemin elde edilmesinin en bariz yolu, benzeşim elde edinceye kadar n 'yi tekrarlamaktır. Alternatif olarak, benzeşim elde edildikten sonra zincirden yapılan tüm çizimler sabit dağılımdan çıkar. Böylelikle bu zincirden elde edilen n ardıl değer π 'den bir örneklem oluşturur. Örneklemin nasıl oluşturulacağı dair daha detaylı bilgiler Kesim 3.5.4'te ele alınacaktır.

Gibbs örnekleme zincirinin tipik yörüngesi aşağıdaki Şekil 3.5'de verilmiştir. Tüm iterasyonlar d 'nin θ bileşenlerinin ordinatları boyunca hareket etmesiyle tamamlanır. Benzetim tanıları d çok büyük olabileceği için karmaşıktır. Benzetimin tüm parametre bileşelerinin ortak dağılımının tüm θ değerleri için ortak art alana benzemesi gerektiği anlamına gelen bir dağılımda gerçekleşmesi gerekir. Bu usandırıcı gerçekleştirme denemekten çok uzaktır (Gelman ve arkadaşları, 1992).



Şekil 3.6: İki Boyutlu Parametre Uzayında Gibbs Örnekleminin Tipik Yörüngesi

Merkezi eğriler art alan yoğunluğunun kontür çizgilerini temsil etmektedir. Parantezdeki noktalar zincirin olası yörüngesini göstermektedir.

Burada öncelikle bazı temel gerçekler ortaya konulmalıdır. Birincisi, daha öncede belirtildiği gibi, Gibbs örnekleme bir Markov zincirini tanımlar. Buna göre, j tekrarlı görüntüsünde olası değişiklik sadece j -1 basamağındaki zincir değerlerine bağlıdır. Ayrıca, geçişler sadece zincir değerlerindeki tekrardan etkilendiği için, zincir homojendir. Buna göre geçiş olasılık fonksiyonu aşağıdaki şekilde de ifade edilir:

$$p(\theta, \phi) = \prod_{i=1}^d \pi(\phi_i | \phi_1, \dots, \phi_{i-1}, \theta_{i+1}, \dots, \theta_d) \quad (3.4)$$

Bu fonksiyon sadece θ ve ϕ zincir değerlerindeki tekrarlara bağlıdır. Yapılan her bir değişiklik geri çevrilebilirse de, bileşenlerin tümünü tarayan bu zincirin kendisi geri döndürülemez.

Peter Green, tersine çevrilmiş dizideki bileşenlerin yeniden taranmasını takiben bileşenlerin taranmasını içeren bir zincirde geri dönüşüm olabileceğine dikkat çekmektedir.(Gamerman,1997) Diğer önemli sonuç, Eşitlik (3.4) ile tanımlanan geçiş fonksiyonunda bir zincirin denge dağılımının π olduğudur. Sürekli durumda ise aynı tartışma geçerli olamaz ancak algoritma mekaniği aynıdır. Geçiş çekirdeği $p(\theta, \phi)$ olan bir Markov zinciri π^α ile sağlanmalıdır ve zincir indirgenemez özellikte olmalıdır. İndirgenemezlik pozitif artalan ihtimali taşıyan her A dizisi için $P(x,A)>0$ olduğu kontrol edilerek her bir uygulamaya uyarlanabilir.

Durağanlığı kontrol etmek için, parametre iki boyutlu kabul edilsin ve $\theta=(\theta_1, \theta_2)$ marjinal kısıtlayıcı yoğunlukları olan $\pi^\alpha(\theta_1)$ ve $\pi^\alpha(\theta_2)$ bir durum olsun. θ_1 'in tam durumsal ve kısıtlayıcı dağılımı, $\pi^\alpha(\theta_1)$ dir. Geçiş çekirdeği Eşitlik (3.4)'den $\phi=(\phi_1, \phi_2)$ 'in θ 'inkilerle aynı boyutları olan bileşenlerinin olduğundan aşağıdaki şekilde basitleştirilebilir (Gamerman, 1997).

$$p(\theta, \phi) = \pi(\phi_1 | \theta_2) \pi(\phi_2 | \phi_1)$$

Burada, $\int \int \pi(\phi_2|\phi_1)\pi^\alpha(\theta_1|\theta_2)d\theta_1d\phi_2 = \int \pi(\phi_2|\phi_1)d\phi_2 \times \int \pi^\alpha(\theta_1|\theta_2)d\theta_1 = 1$ olduğundan,

$$\begin{aligned}\pi^\alpha(\theta_2) &= \int \int \pi(\phi_2|\phi_1)\pi^\alpha(\theta_1|\theta_2)\pi^\alpha(\theta_2)d\theta_1d\phi_2 \\ &= \int \int \pi(\phi_2|\phi_1)\pi^\alpha(\theta)d\theta_1d\phi_2\end{aligned}\quad (3.5)$$

elde edilir. ϕ_2 göz önüne alınarak ϕ_1 'in marjinal kısıtlayıcı yoğunluğu,

$$\begin{aligned}\pi^\alpha(\phi_1) &= \int \int p(\theta, \phi)\pi^\alpha(\theta)d\theta d\phi_2 \\ &= \int \int \int \pi(\phi_1|\theta_2)\pi(\phi_2|\phi_1)\pi^\alpha(\theta)d\theta_1d\theta_2d\phi_2 \\ &= \int \pi(\phi_1|\theta_2)\pi^\alpha(\theta_2)d\theta_2 \quad \text{'dir}\end{aligned}\quad (3.6)$$

Burada son eşitlik (3.5)'ten çıkar. (3.6)'yı karşılayan tek dağılımın $\pi(\theta_1|\theta_2) = \pi^\alpha(\phi_1/\phi_2)$ eşitliği olması gerekir. Aynı tartışma $\pi(\phi_2/\phi_1) = \pi^\alpha(\phi_2/\phi_1)$ 'i elde etmek için de kullanılabilir. Aynı tartışma, d bileşen bloklarına bölünmüş θ için de izlenebilir, çünkü bunlar daima θ_1 ve θ_2 gibi iki blokla yeniden düzenlenebilir. Bu, tüm kısıtlayıcı koşulların art alandaki koşullarca ortaya konulduğu anlamına gelir. Genellikle $\pi^\alpha = \pi$ olmasını garantilemez. Bununla birlikte, Besag (1974) çalışmasında, çok katı olmayan koşullarda bütün dağılımlar dizisinin bir ortak dağılımı belirlediğini göstermiştir. Dolayısıyla, geçiş çekirdeği (3.4) olan Markov zinciri, π ilgi dağılımına benzer ve tekrarlı örneklem tasarısı bu dağılımdan bir değer oluşturur. Gibbs örnekleme için benzeşim koşulları Smith (1984) ve Tierney (1994) tarafından ortaya konulmuştur (Gamerman, 1997).

Gibbs örnekleminde benzeşimi sağlayan kuramsal sonuçlara rağmen, uygulama ele alınan modellerin karmaşıklığı yüzünden zor olabilir. Örneklemin benzeşimini karakterize etmek zorlaşır. Sıralı ve tekrarlı bir metod düşünüldüğünde, metodun etkinliğini artıracak uygulama stratejileri hesaplama maliyeti üzerinde oldukça etkili olabilir. Etkinlik büyük ölçüde tekrar sayısını ve her bir tekrarda gereken aritmetik

işlem sayısının azaltılmasına bağlıdır. Verilen teknikler bir önceki bölümde tanımlandığı ve Gibbs örnekleme ile gösterildiği üzere temel MCMC yöntemleriyle ilişkilidir.

3.5.2 Örneklem Oluşturma

π dağılımından n büyüklüğünde bir örneklem elde etmenin iki yolu bir önceki kesimde sunulmuştu. Burada açıkça görülen nokta, n zincirinde benzeşim elde edene dek işlem yapmak ve her bir n zincirinden örneklem elemanları almanın gerekliliğidir. Eğer zincirler bağımsız olarak başlamışsa, örneklemeler, π 'den bağımsız değerlerden oluşur. Başlangıç değerlerinin hepsi farklıysa bağımsızlığı elde etmek daha kolaydır.

Başka bir yöntem de tek bir zinciri alıp, sonuçlar elde etmektir. Benzeşimi takiben, π denge dağılımı ile bütün zincir değerlerinin marjinal dağılımları elde edilir. Böylece, n büyüklüğünde bir örneklem bu zincirdeki n ardıl değerlerinden elde edilebilir. Buradaki güçlük örneklem elemanlarının zincirin bağımlılığı yüzünden artık bağımsız olmayışlarıdır. Zincirin oto korelasyonu çok yüksek ve örneklem bunu red edecek kadar geniş değilse bazı sorunlar ortaya çıkabilir. Bu durumlarda, zincirlerin tüm parametre aralığını uygun biçimde ve yeterli şekilde öretecek hale gelmesi çok uzun zaman alabilir.

Bağımsızlığı veren alternatif yaklaşım da, k . tekrarlarının her birindeki örneklem zincir değerini almaktır. **Markovcu** işlemlerde sadece ilk sıra bağımlılığı vardır. Tekrarlar arası açıldıkça zincir değerleri giderek korelasyonunu yitirir ve yeterince büyük bir k aralık değeri için, zincirler bağımsız olur.

Bağımsız örneklem yaklaşımı Gelfand ve Smith (1990) tarafından önerilmiştir ve bu yol bazı araştırmacılarca kullanılmıştır. Tek zincir yaklaşımı ise Geyer (1992) tarafından önerilmiştir. Her k . aralığın örneklenmesi ise, Raftery ve Lewis (1992) tarafından irdelenmiştir. Gelman ve Rubin (1992a), tek zincirlerin farklı limitler için bilimsel benzetim ortalaması veren bir örnek çalışmalarıyla, Gelman ve Rubin (1992b), desteklenen bağımsız zincirlerin az sayıda kullanımını önermişlerdir

(Gelman ve arkadaşları, 1998). Uygulamalarda n paralel zincirleri kullanmanın hesap açısından yetersiz ve gereksiz olduğu konusunda genel bir uyuşma sağlanmışken, bu alanda genel bir uzlaşma yoktur. Esas tartışma bir kaç paralel zincire gerek olup olmadığıdır. Eğer zincirin benzeşim özellikleri iyi anlaşılırsa, o halde tek bir zincir yeterlidir. Bu özellikleri elde etmek zor olduğundan, öngörümüz bir kaç paralel pilot zincirin kullanılmasını söylemektedir. Ortak değerleri çabucak belirlemeleri halinde, tek bir zincir oldukça güvenli biçimde kestirim için geniş bir örneklem oluşturmak üzere de kullanılabilir.

3.5.3 Tarama Stratejileri

Uygulamalarda tüm iterasyonlar genellikle $1 \rightarrow 2 \rightarrow \dots \rightarrow d$ gibi aynı belirleyici düzendeki bileşenlerin güncellenmesini içerir. θ bileşenlerinin taranması veya güncellenmesi için bir çok başka olası stratejiler vardır.

Geman ve Geman (1984), zincir süresiz olarak kullanıldığında tüm bileşenlerin incelendiğini garantileyen tarama planları için, ortak dağılıma olan benzerliği göstermiştir (Gelman ve arkadaşları, 1992). Belli bir sırada her bir bileşen görüldüğü ve ters sırada yeniden tarandığı için geri döndürülebilir. Gibbs örnekleme, bu özelliğe uyar. Bu durumda, ikinci güncellemelerden ve stratejiler arasındaki kıyaslamalardan oluşan her iterasyon, bunu göz önünde bulundurmalıdır.

Roberts ve Sahu (1997) her bir iterasyonda $\{1, \dots, d\}$ değişiminin seçildiği ve bileşenlerin bu sırada tarandığı bir tesadüfi değişim taramasını ele alır. Zeger ve Karim (1991) ise, bazı bileşenlerin sadece k . tekrarında tarandığı bir Gibbs örneklem tasarısını anlatır.

3.5.4 Örneklemin Kullanılması

Örnekleme oluşturmak için seçilen tasarım ne olursa olsun, $\theta_1, \dots, \theta_n$ vektörlerinin örnekleminin kullanılmasından sonra, π dağılımı elde edilir. Tek bir Markov Zincirinden elde edilen ardışık değerlerin olduğu daha genel durumu düşünelim.

θ 'nın i . bileşeni örnekleme $\theta_{1i}, \dots, \theta_{ni}$ 'den elde edilir. Herhangi bir reel fonksiyon $\psi = t(\theta)$ 'nin marjinal noktası ya da aralık toplamı, örnekleme dayalı ve onları karşılayan kestirimlerle tahmin edilir. Bu kestirimin kalitesi, MCMC kestirimiyle ilgili güven aralıklarının oluşturduğu merkezi limit teoremince değerlendirilebilir (Gamerman, 1997).

Böylelikle, ψ ortalaması, $\psi_j = t(\theta_j), j = 1, \dots, n$ olmak üzere

$E(\psi) = \psi = (1/n) \sum_{j=1}^n \psi_j$ ile kestirilir. Benzer biçimde ψ 'nin varyansı

$\sigma_{\psi}^2 = \text{Var}(\psi) = E(\psi^2) - [E(\psi)]^2$ 'dir. Burada σ_{ψ}^2 ,

$\hat{\sigma}_{\psi}^2 = \hat{E}(\psi^2) - [E(\psi)]^2 = \frac{1}{n} \sum_{j=1}^n (\psi_j - \hat{\psi})^2$ örneklem varyansı ile kestirilir. $n-1$, n

paydasının yerini alabilir, ancak n bu değişikliği önemsiz hale getirecek kadar genişdir. n ardışık değerleri olan bir örnekleme, her bir k . tekrarı gözardı ederek elde edilmiş bir $m=n/k$ değerleri örnekleme arasında tercih yapma problemi yeniden ele alalım. Yarı bağımsız alt örneklerin oluşturulduğunu da varsayalım.

$\hat{\psi}_1, \dots, \hat{\psi}_k$ ile gösterilen alt örneklem ortalamaları, $\hat{\psi} = (1/k) \sum_{j=1}^k \hat{\psi}_j$ ile gösterilen de

bütün örneklemin ortalaması olsun. $E(\psi)$ 'nin kestiricileri $k+1$ adettir. ve bunlar tutarlı kestiricilerdir. Bu her j için $\text{Var}(\bar{\psi}) \leq \text{Var}(\bar{\psi}_j)$ olarak gösterilebilir. Bu, bağımsız örneklemin indirgeme etkinliğine karşılık elde edildiği anlamına gelir.

Marjinal yoğunluklar $\pi(\theta_i)$, θ_i örneklem değer histogramı ile kestirilebilir. Daha iyi kestirimler durumsal dağılımlarla elde edilebilir. $\pi(\theta_i) = \int \pi(\theta_i | \theta_{-i}) \pi(\theta_{-i}) d\theta_{-i}$ olduğu anımsanarak, bir Monte Carlo kestirimi aşağıdaki eşitlikle elde edilir:

$$\hat{\pi}(\theta_i) = \frac{1}{n} \sum_{j=1}^n \pi(\theta_i | \theta_{j,-i}) \quad (3.7)$$

Burada $\theta_{j,-i}, j = 1, \dots, n$ marjinal $\pi(\theta_{-i})$ 'den elde edilen örneklem olup, $\hat{\pi}$ 'nin tutarlı bir kestirici olduğu ve her bir θ_i değeri için, Merkezi Limit Teoremi'ne uyduğu

söylenbilir (Gamerman, 1997). Bu kestiriciler sürekli parametreler için her zaman süreklidir. Daha da önemlisi, Ψ 'nin şekline dair bilgiye dayalıdır. Bu nedenle, Gelfand ve Smith (1990) buna Rao-Blackwell yoğunluk kestirimi adını verir. Bu, kestiricilerin yeterli istatistiklerle daima geliştirildiğini (örneklem varyansını düşürme bağlamında) öne süren Rao-Blackwell teoremine bir göndermedir. Bu sonucu bağımsız örneklem bağlamında yoğunluk kestirimi için doğrulamışlardır. Markov zinciri örnekleme için sonucun genel kanıtı Liu, Wong ve Kong (1994) tarafından yapılmıştır. Aynı fikir, buradaki kazançlar çok fazla olmasa da $E[t(\theta_i)] = \frac{1}{n} \sum_{j=1}^n E[t(\theta_i) | \theta_{j,-i}]$ ile $t(\theta_i)$ momentleri için daha iyi kestirimler yapmak üzere kullanılabilir (Gamerman, 1997).

3.6 BUGS Paket Programı

BUGS (Bayesian Inference Using Gibbs Sampling) programı, Bayesci grafiksel modellerdeki karmaşık koşullu dağılımların (minimum kullanıcı girdisi ile) otomatik olarak türetilmesi için Gibbs örneklemesinin bir uygulanması olarak yazılmıştır. Modeller özellikle grafikleri tanımlamak için betimleyici bir dille belirtilmiş. Bu program aşağıdaki adımları izler:

Bu bölümde BUGS programının teknik özellikleri ve BUGS programında, bir modelin tanımlanmasından, modele ait çıkarsama sonuçlarının elde edilmesine kadar olan adımlar ve her adımda yapılması gereken işlemler anlatılacaktır.

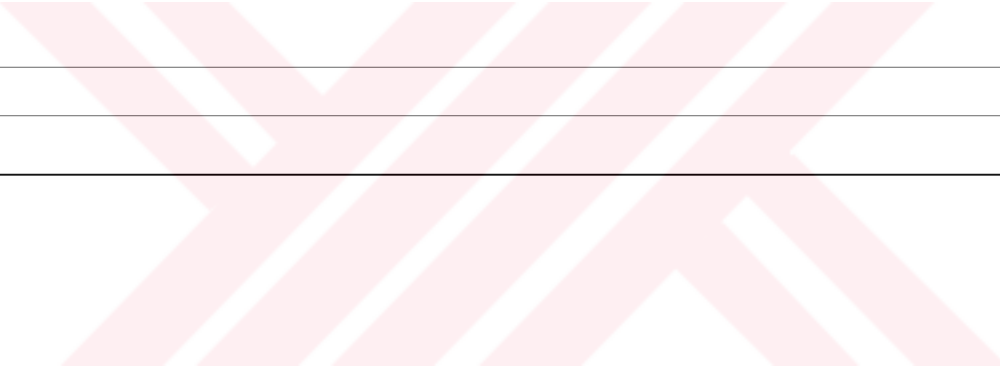
3.6.1 BUGS Dosyası ve Özellikleri

BUGS programında kullanılan dosya, bileşik (compound) bir dosyadır. Burada bileşik dosya ile anlatılmak istenen, (text) yazı ya da grafik türü bileşenlerin hepsini içinde bulundurabilen bir dosya olma özelliğidir. BUGS'ın bu tip bir dosyalama sisteminin olması kullanıcıya, kurduğu modelin kodunu, grafiksel gösterimini, çözümlene için kullanacağı verileri, başlangıç değerlerini ve elde ettiği sonuçları (grafikler ya da tablolar) tek bir dosya içinde saklaması kolaylığını sağlamaktadır.

3.6.2 Nesnelere ve Özellikleri

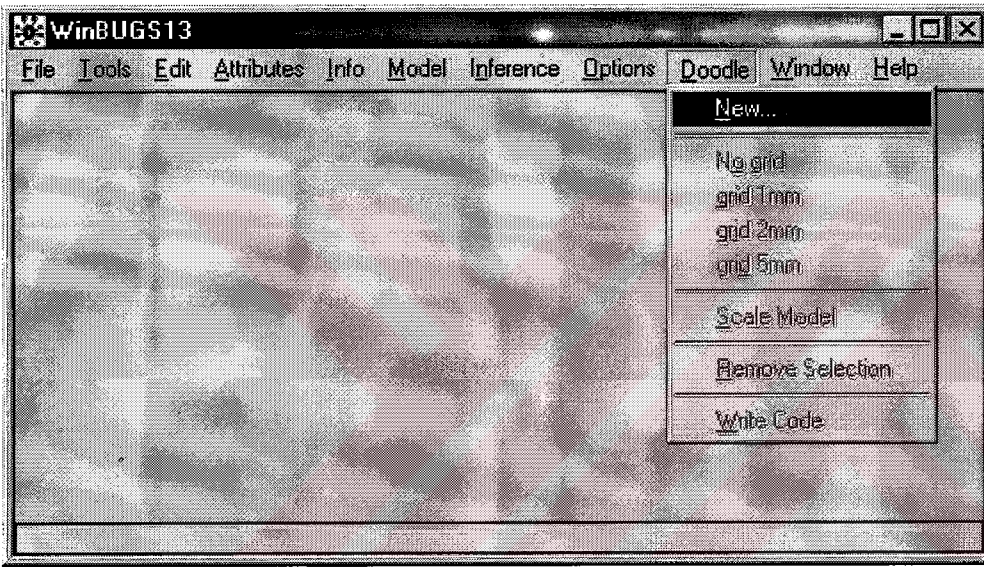
Modelin grafiksel gösterimi, node, plate gibi nesnelere de (kavramların da) BUGS'da yer almasına neden olmuştur. Node olarak adlandırdığımız nesnelere, modeldeki parametreler veya sabitlerdir. Grafiksel gösterimleri türlerine göre değişim gösterir. Sabitler dikdörtgen, parametreler elips şeklindedir. Node'larla ilgili ayrıntılı bilgiler Model Tanımlama bölümünde verilecektir.

name:		type:	stochastic	density:	dnorm
mean	0.0	precision	1.0E-6	lower bound	upper bound

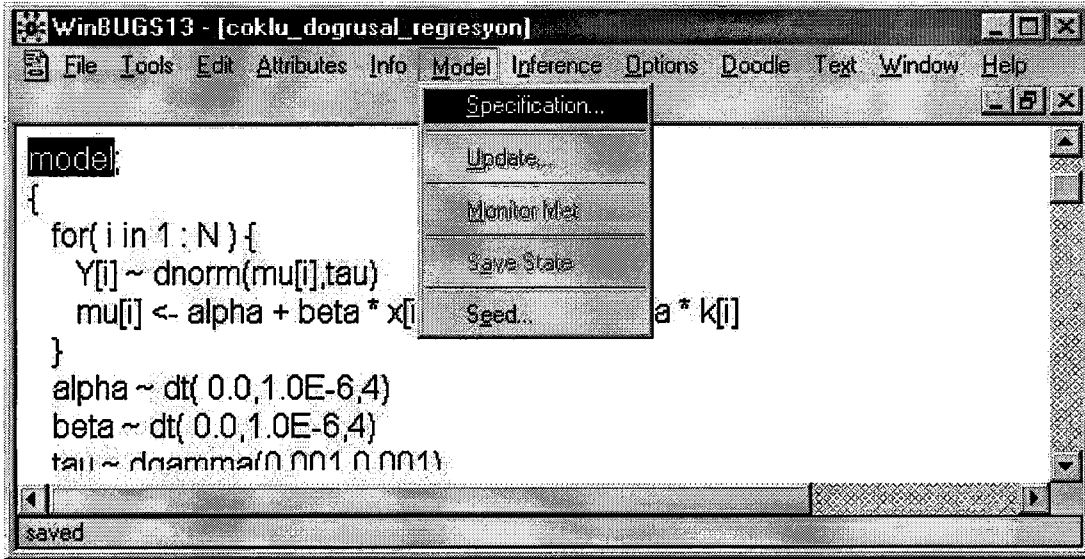


3.6.3 BUGS'da Çözümleme

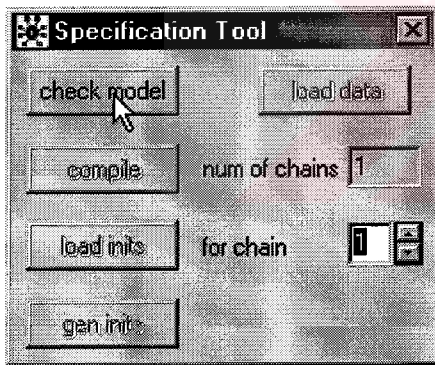
BUGS ile çözümleme yapmak için öncelikle modelin BUGS kodu ile ifade edilebilmesi gerekmektedir. Model tanımlamayla ilgili iki yöntem mevcuttur. Modeli doğrudan kod olarak yazabileceğimiz gibi, grafiksel gösterimini oluşturduğumuz modelin BUGS koduna dönüştürülmesini de sağlayabiliriz. "**Doodle**" menüsünden "**New**" seçeneğine girerek yeni bir doodle yaratıp, yaratılan bu doodle üzerine node ve plate'ler oluşturarak kurmak istediğimiz modelin grafiksel gösterimini oluşturabiliriz. Oluşturduğumuz grafiksel modelin BUGS kodunu ise aynı menüdeki "**Write Code**" seçeneği ile elde edebiliriz.



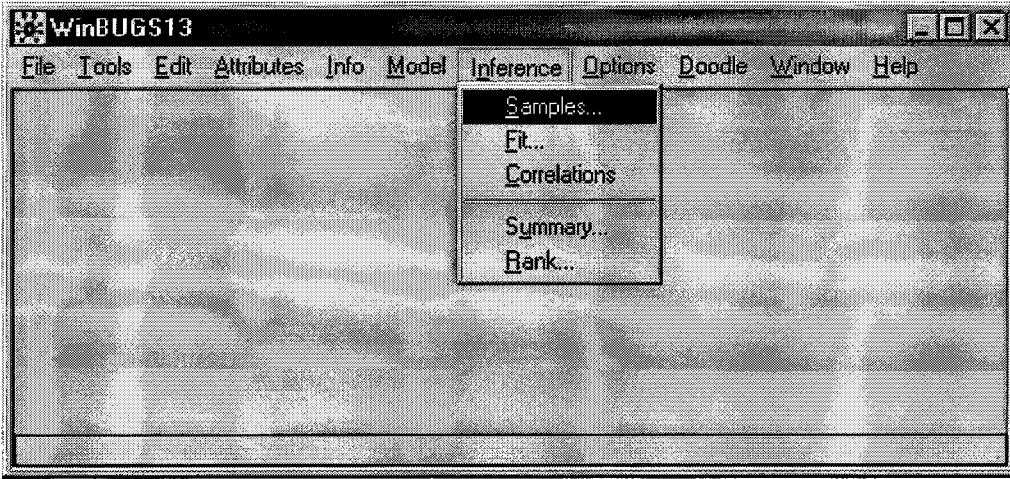
Bir grafiksel modelin BUGS'a tanıtılması, için modelin kodunun yazılması ve daha sonra yazılan bu kodda model kelimesi seçilerek, "**Model**" menüsünden "**Specification**" bölümü seçilmelidir.



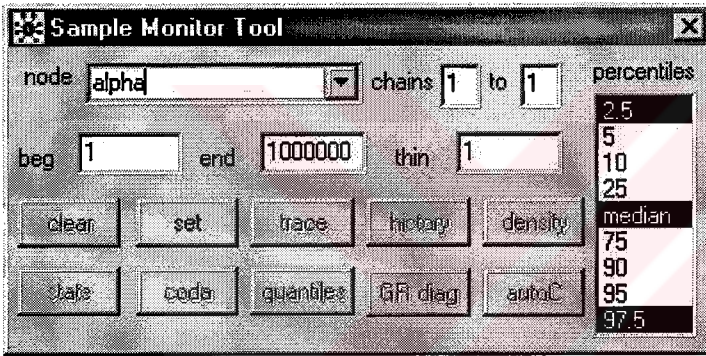
Gelen form üzerinde "**check model**" düğmesi tıklanmalı, sonra BUGS'a girilen verilerin "**list**" kısmı seçilip "**load data**" düğmesine basılarak veriler yüklenmeli ve model "**compile**" düğmesine basılarak compile edilmelidir. Bu aşamadan sonra yine list formatında girilmiş ilk değerler de "**load inits**" düğmesine basılarak aynı şekilde yüklenip model BUGS tarafından çözümlenmeye hazır hale getirilir.



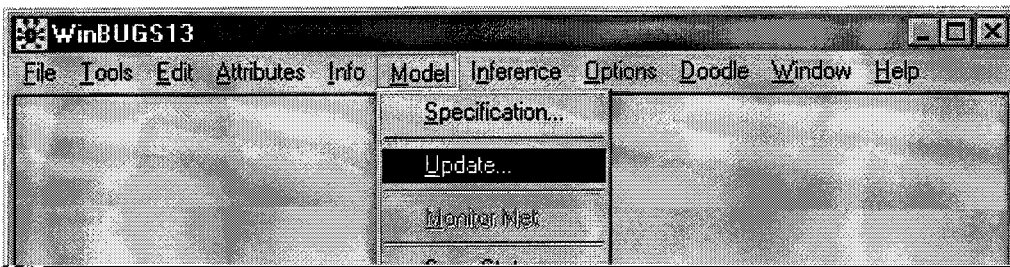
Bu işlemlerden sonra "**Inference**" menüsünden "**Samples...**" seçeneğine girilir.



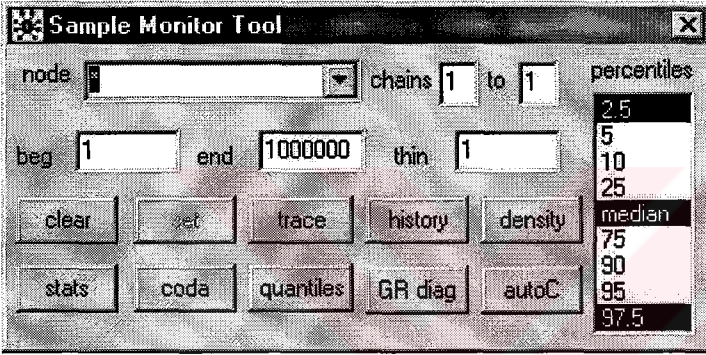
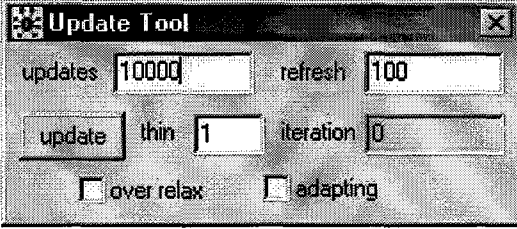
Gelen formda modeldeki parametreler (alpha, beta) girilip "set" edilir.



Bu aşamadan sonra "Model" menüsünden "Update" seçeneğine girilir.



Bu formda iterasyon sayısı updates bölümüne girilerek update düğmesine basılır. Update işlemi bittiğinde, tekrar "Inference" menüsündeki "Samples..." seçeneğine dönülüp "node" bölümüne * yazılarak set edilen tüm parametreler için "stats, trace, history, density" vs sonuçları alınır.



Çalışmanın izleyen bölümünde Bayesci grafik modelleme tekniği üç farklı model ele alınarak incelenmiştir. Burada modellerin koşullu bağımsızlık yapıları grafiklerle ifade edilmiş ve modeller BUGS programı ile işletilmiştir. Model parametrelerine ilişkin çıkarsamalar her bir model çalışmasının son kısmında verilmiştir.

4. BAYESCİ GRAFİK MODEL UYGULAMALARI

Tez çalışmasının uygulama bölümünde üç model ele alınmıştır. Bu modeller,

- Çoklu doğrusal regresyon modeli,
- Hiyerarşik çoklu doğrusal regresyon modeli,
- Poisson-Gamma hiyerarşik modelidir.

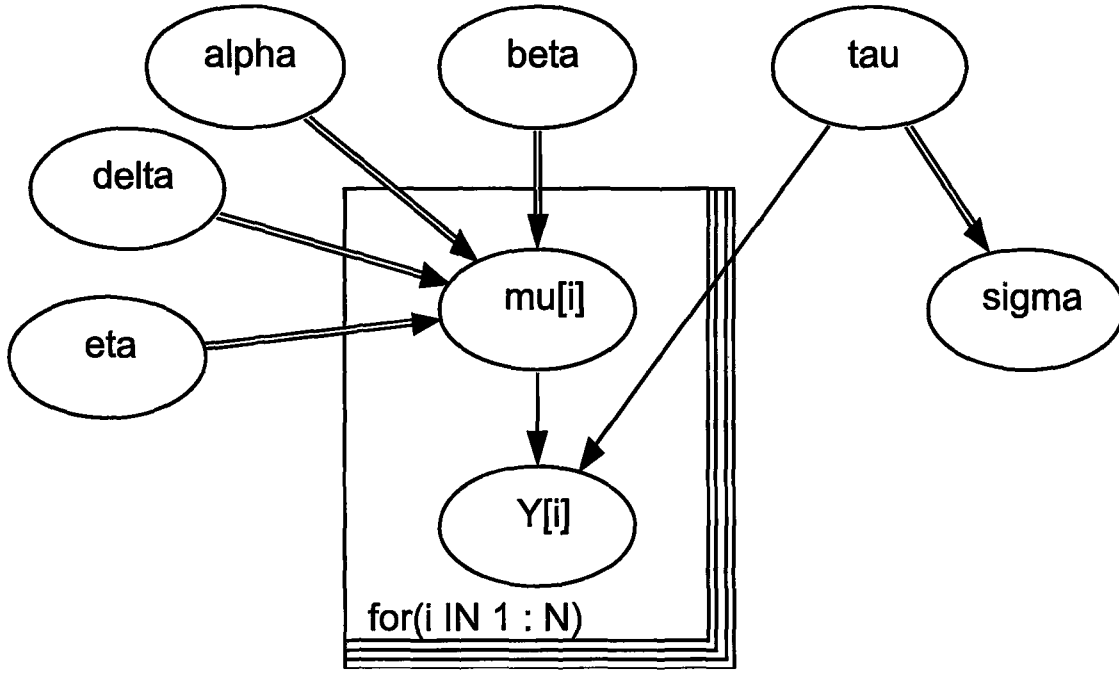
Burada herbir model için grafiksel yapı elde edilmiş ve dolayısıyla Bayesci grafik model tanımları oluşturulmuştur. Modellerde yer alan düğümler için önsel dağılımların tanımlamaları yapılmış ve modeller BUGS kodları ile ifade edilmiştir. Çalışmada modeller Gibbs örnekleme ile işletilmiştir. Burada ilk iki modelin işletiminden elde edilen sonuçlar, analitik sonuçlarla karşılaştırılmıştır.

4.1 Çoklu Doğrusal Regresyon Modeli için Bayesci Grafik Modelleme

Çalışmada ele alınan gözlem değişkeni $y_i \sim N(\mu_i, \sigma^2)$ ve $\mu_i = \alpha + \beta(x_i) + \delta(Z_i) + \eta(K_i)$ olup; çoklu regresyon modeli aşağıdaki şekilde tanımlanmıştır:

$$y_i = \alpha + \beta(x_i) + \delta(Z_i) + \eta(K_i) + \sigma^2 \quad (4.1)$$

Çoklu regresyon modelinde yer alan α , β , δ , η ve τ ($1/\sigma^2$) parametreleri Bayesci düşünce çerçevesinde birer rastlantı değişkenleri olup; herbiri oluşturulacak olan grafik modelde birer düğüm olarak yer almışlardır. Tez çalışmasında sadece üç bağımsız değişken içeren bir regresyon modeline ilişkin sonuçlara yer verilmiştir. Ancak, bağımsız değişken sayısının farklı olduğu birçok model BUGS programında ele alınmış; Gibbs örnekleme tahminleri ile en küçük kareler tahminleri karşılaştırılmıştır. Tüm bu denemelerde de benzer sonuçlar elde edilmiştir. Çoklu doğrusal regresyon modelinin grafiksel olarak sunumu aşağıdadır:



Şekil 4.1: Çoklu Doğrusal Regresyon Modelinin Grafiksel Yapısı

Şekil (4.1)'den model parametreleri $(\alpha, \beta, \eta, \delta, \tau)$ için bileşik önsel yoğunluk fonksiyonu şu şekilde yazılabilir:

$$P(\alpha, \beta, \eta, \delta, \tau) = p(\alpha)p(\beta)p(\eta)p(\delta)p(\tau) \quad (4.2)$$

Modelde yer alan parametrelerin birbirinden bağımsız olmaları, bileşik önsel dağılımın Eşitlik (4.2)'de verildiği gibi kolayca ayrıştırılabilmesine olanak sağlamıştır. Eşitlik (4.1) ile verilen modelde çoklu doğrusal bağlantı (açıklayıcı değişkenler arasındaki ilişki durumu) gibi sorunlar olsaydı, yukarıda tanımlanan önsel dağılımın daha karmaşık bir yapı içermesi kaçınılmaz olurdu. Modelde yer alan parametreler için kullanılan önsel dağılım formları aşağıdaki tabloda verilmiştir:

Tablo 4.1: Önsel Dağılımlar

Parametre	Uygun Dağılım	f(.)
α	Normal (μ, σ^2)	$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
β	Normal (μ, σ^2)	$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
η	Normal (μ, σ^2)	$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
δ	Normal (μ, σ^2)	$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
τ	Gamma(r, φ)	$\varphi^r \tau^{r-1} e^{-\varphi\tau} / \Gamma(r); \tau > 0$

Çoklu regresyon modelinin BUGS programında tanımlanabilmesi için geliştirilen BUGS kodu aşağıda verilmiştir:

model

```
{
  for( i in 1 : N ) {
    Y[i] ~ dnorm(mu[i],tau)
    mu[i] <- alpha + beta * (x[i]) +delta * (z[i]) +eta * (k[i])
  }
  tau ~ dgamma(0.001,0.001) sigma <- 1 / sqrt(tau)
  alpha ~ dnorm(0.0,1.0E-6)
  beta ~ dnorm(0.0,1.0E-6)
  delta ~ dnorm(0.0,1.0E-6)
  eta ~ dnorm(0.0,1.0E-6)
}
```

4.1.1 Modelin İşletimi

Burada, grafik modeli ve BUGS program kodu oluşturulan regresyon modeli BUGS programında işletilmiş, modelde yer alan parametrelere ilişkin sonsal tahminler farklı başlangıç değerleri kullanılarak elde edilmiştir. Model tahminlerine 10.000; 100.000 ve 1.000.000 iterasyon sonucunda ulaşılmıştır. BUGS programına veriler aşağıdaki komut ile girilmiştir:

- list(x=c(22,26,45,37,28,50,56,34,60,40),z=c(2,2,3,4,4,3,6,3,5,3),
k=c(4,8,7,0,2,10,8,8,2,6), Y= c(16,17,26,24,22,21,32,18,30,20), N = 10)

Çalışmada öncelikle Gibbs örneklemesinden elde edilen model tahminleri üzerinde başlangıç değerlerinin ne anlamda önemli olduğu incelenmiştir. Bu amaçla ilk olarak aşağıdaki başlangıç değerleri ele alınmıştır. Model tahminleri farklı iterasyon sayıları için tablolar halinde aşağıda sunulmuştur:

BAŞLANGIÇ DEĞERLERİ
(alpha = -1, beta = 1, delta=0 , eta= 0, tau = 1)

EKK ile elde edilen tahmin

$$\hat{Y} = 7,596 + 0,194 X + 2,338 Z - 0,163 K$$

(2,595) (0,088) (0,908) (0,244)

Kesim 3.5.1'te belirtildiği gibi, model parametrelerine başlangıç değerlerinin atanması Gibbs Örneklemesinin uygulanabilmesinde ilk adımdır. Bu çalışmada tek zincirli MCMC ele alındığından, tarama ve denge dağılımına ulaşma sürecinde her bir parametre için sadece bir başlangıç değeri ele alınmıştır. 10.000 iterasyon sonunda elde edilen sonuçlar aşağıdaki tabloda sunulmuştur.

Tablo 4.2: 10.000 İterasyon Sonunda Model Tahminleri

Parametre	Ortalama	Standart Sapma	MC Hata	Ortanca
alpha	7.46	2.909	0.01179	7.416
beta	0.1999	0.09415	0.005792	0.1989
delta	2.303	0.9538	0.0562	2.291
eta	-0.1615	0.2753	0.01113	-0.1604
tau	0.2533	0.1414	0.002961	0.2185

Tabloda yer alan ilk sütun, parametrelerin sonsal dağılımlarının beklenen değerlerini; ikinci sütun varyanslarını; üçüncü sütun simülasyon sonucuna ilişkin hataları ve son sütun ise, ortanca değerlerini vermektedir. Burada alpha, beta, delta ve eta parametrelerinin sonsal dağılımı normal olup; tau'nun sonsal dağılımı da gamma dağılımıdır (eşlenik yapı nedeniyle). Parametre tahminlerinin yukarıda kutu içine alınmış EKK tahminlerine oldukça yakın oldukları söylenebilir. MC hataların da düşük değerler aldıkları görülmektedir. Bu da simülasyon sonucuna duyulacak güveni bir ölçüde desteklemektedir. Ortanca değerlerinin ilk dört parametre için ortalama değerlere yakın çıkması elde edilen dağılımın simetrikliği

hakkında bilgiyi pekiştirmektedir. Tau parametresinin dağılımı ise, pozitif çarpık bir dağılımdır. İterasyon sayısının artırılmasının sonuçlar üzerinde etkisinin gözlenmesi için, yukarıdaki işlemler 100.000 ve ardından 1.000.000 iterasyonlar için de yapılmıştır. Sonuçlar aşağıdaki tablolarda verilmiştir:

Tablo 4.3: 100.000 İterasyon Sonunda Model Tahminleri

Parametre	Ortalama	Standart Sapma	Mc Hata	Ortanca
alpha	7.504	3.205	0.04975	7.515
beta	0.1925	0.1097	0.002446	0.1931
delta	2.368	1.129	0.02589	2.361
eta	-0.1559	0.301	0.004915	-0.1571
tau	0.241	0.1396	0.001279	0.2147

Tablo 4.4: 1.000.000 İterasyon Sonunda Model Tahminleri

Parametre	Ortalama	Standart Sapma	Mc Hata	Ortanca
alpha	7.617	3.19	0.01514	7.613
beta	0.1951	0.1081	7.658E-4	0.195
delta	2.333	1.12	0.008039	2.327
eta	-0.1656	0.3002	0.001506	-0.1649
tau	0.2413	0.1393	4.134E-4	0.2131

İterasyon sayısının artması model sonuçlarına duyulan güveni daha da pekiştirmiştir. Burada, MChatalarının Tablo 4.2'ye göre daha küçük olduğu görülmektedir. Ayrıca, bir önceki tabloya göre, EKK tahminlerine daha da yakın sonuçlar elde edilmiştir. Ancak, iterasyon sayısının artırılmasının zaman açısından bilgisayarlara ek bir yük getirdiği de açıktır. Örneğin Pentium III işlemcili 64 MB RAM'e sahip bir bilgisayarda, 10.000 iterasyon birkaç saniye zaman alırken; 1.000.000 iterasyon aynı bilgisayarda 15 dakikalık bir zaman almaktadır.

Farklı bir başlangıç değerleri seti için elde edilen sonuçlar aşağıda sunulmuştur.

BAŞLANGIÇ DEĞERLERİ
(alpha = 0, beta = 0, delta=1 , eta= 1, tau = 1.2)

EKK ile elde edilen tahmin

$$\hat{Y} = 7,596 + 0,194 X + 2,338 Z - 0,163 K$$

(2,595) (0,088) (0,908) (0,244)

10.000 iterasyon sonunda elde edilen sonuçlar aşağıdaki tabloda sunulmuştur.

Tablo 4.5: 10.000 iterasyon Sonunda Model Tahminleri

Parametre	Ortalama	Standart Sapma	Mc Hata	Ortanca
alpha	7.519	2.842	0.1188	7.441
beta	0.1996	0.0926	0.005721	0.1989
delta	2.295	0.9626	0.05765	2.289
eta	-0.1641	0.2763	0.01156	-0.161
tau	0.2532	0.1413	0.002995	0.2282

100.000 iterasyon sonunda elde edilen sonuçlar aşağıdaki tabloda sunulmuştur.

Tablo 4.6: 100.000 iterasyon Sonunda Model Tahminleri

Parametre	Ortalama	Standart Sapma	Mc Hata	Ortanca
alpha	7.532	3.199	0.04961	7.518
beta	0.1925	0.1095	0.002443	0.1931
delta	2.367	1.13	0.02595	2.36
eta	-0.1562	0.3011	0.004931	-0.1573
tau	0.241	0.1396	0.001279	0.2147

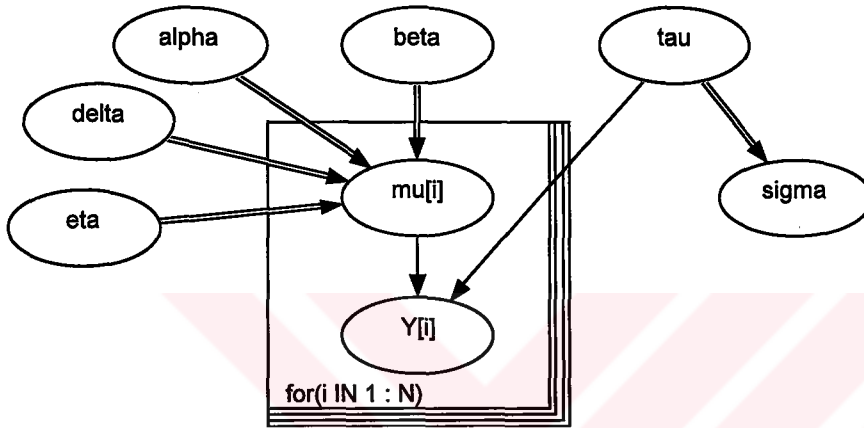
1.000.000 iterasyon sonunda elde edilen sonuçlar aşağıdaki tabloda sunulmuştur.

Tablo 4.7: 1.000.000 iterasyon Sonunda Model Tahminleri

Parametre	Ortalama	Standart Sapma	Mc Hata	Ortanca
alpha	7.617	3.19	0.01513	7.613
beta	0.1951	0.1081	7.657E-4	0.195
delta	2.322	1.12	0.008041	2.327
eta	-0.1656	0.3002	0.001506	-0.1649
tau	0.2413	0.1393	4.133E-4	0.2151

Burada da iterasyon sayısının artmasının model sonuçlarına duyulan güveni pekiştirdiği görülmüştür. Bir önceki denemede olduğu gibi, burada da EKK tahminlerine yakın sonuçlar elde edilmiştir. Sonuç olarak başlangıç değerlerinin farklılığının model tahminleri üzerinde çok önemli değişiklikler yapmadığı söylenebilir. İterasyon sayısının büyük alınması işlemlerdeki geçmiş kayıtlara verilen önemi azaltmakta ve dolayısıyla başlangıç değerleri çok etkili olamamaktadır.

Çalışmada ayrıca $(\alpha, \beta, \eta, \delta, \tau)$ parametrelerinin önsel dağılımların farklılığının model sonuçları üzerine olası etkisi incelenmiştir. Beş parametre içeren bu modelde, α , β , η ve δ parametreleri simetrik bir önsel dağılıma sahip rastlantı değişkenleridir ve normal dağılımla modelde temsil edilmişlerdir. Bu kesimde her bir parametre için t dağılımı tanımlanmış olup; varyans bileşeni için ise yine gamma dağılımı kullanılmıştır. Grafik model ve modelin BUGS kodu aşağıda verilmiştir:



Şekil 4.2: Önsel Dağılımı Farklı Grafik Model

```

model;
{
  for( i in 1 : N ) {
    Y[i] ~ dnorm(mu[i],tau)
    mu[i] <- alpha + beta * x[i] + delta * z[i] + eta * k[i]
  }
  alpha ~ dt( 0.0,1.0E-6,4)
  beta ~ dt( 0.0,1.0E-6,4)
  tau ~ dgamma(0.001,0.001)
  sigma <- 1 / sqrt(tau)
  delta ~ dt( 0.0,1.0E-6,4)
  eta ~ dt( 0.0,1.0E-6,4)
}

```

Yukarıda tanımlanan model, $\alpha = -1$, $\beta = 1$, $\delta = 0$, $\eta = 0$, $\tau = 1$ başlangıç değerleri ile işletildiğinde, 10.000 iterasyon sonunda elde edilen sonuçlar aşağıdaki tabloda sunulmuştur.

Tablo 4.8: 10.000 İterasyon Sonunda Model Tahminleri

Parametre	Ortalama	Standart Sapma	MC Hata	Ortanca
alpha	7.6670	3.296	0.1605	7.805
beta	0.1974	0.115	0.0079	0.190
delta	2.2960	1.068	0.0679	2.289
eta	-0.1744	0.298	0.0116	-0.178
tau	0.2395	0.137	0.0037	0.2183

Tabloda yine ilk sütun, parametrelerin sonsal dağılımlarının beklenen değerlerini; ikinci sütun varyanslarını; üçüncü sütun simülasyon sonucuna ilişkin hataları ve son sütun ise, ortanca değerlerini vermektedir. Burada alpha, beta, delta ve eta parametrelerinin sonsal dağılımı t dağılımı olup; tau'nun sonsal dağılımı da gamma dağılımıdır. Burada alpha, beta, delta ve eta parametreleri için eşlenik olmayan bir dağılım seçilmiştir, bu da ele alınan modelin analitik yolla çözümünü hayli zorlaştırır. Ancak simülasyon ile çözüme gidildiğinden burada herhangi bir sıkıntı yaşanmamıştır. Parametre tahminlerinin yine EKK tahminlerine oldukça yakın oldukları söylenebilir. Burada ele alınan simetrik dağılım, normal dağılıma göre daha ağır kuyruklara sahiptir; bu nedenle sonsal dağılımın s.sapmaları nispeten daha büyük elde edilmiştir. MC hataların da bir önceki sonuçlara benzer olarak, düşük değerler aldıkları söylenebilir. Ortanca değerlerinin ilk dört parametre için ortalama değerlere yakın çıkması elde edilen dağılımın simetrikliği hakkında bilgiyi pekiştirmektedir. Tau parametresinin dağılımı ise, pozitif çarpık bir dağılımdır. Yukarıdaki işlemler 100.000 iterasyonlar için de yapılmıştır. Sonuçlar aşağıdaki tabloda verilmiştir:

Tablo 4.9: 100.000 İterasyon Sonunda Model Tahminleri

Parametre	Ortalama	Standart Sapma	MC Hata	Ortanca
alpha	7.675	3.069	0.04507	7.676
beta	0.197	0.1063	0.00218	0.195
delta	2.296	1.088	0.02214	2.308
eta	-0.172	0.2917	0.00422	-0.169
tau	0.242	0.1390	0.00125	0.2165

Bir önceki denemede olduğu gibi, burada da EKK tahminlerine yakın sonuçlar elde edilmiştir. Sonsal varyanslarda ve MC hatalarda nispi bir küçülme gözlenmiştir. Çalışmada ayrıca, tau parametresinin de dağılımı değiştirilmiş; gamma dağılımı

yerine benzer karakteristikte olan ki-kare dağılımı kullanılmıştır. Modelin işletimi için yazılan BUGS kodu aşağıda sunulmuştur:

```

model;
{
  for( i in 1 : N ) {
    Y[i] ~ dnorm(mu[i],tau)
    mu[i] <- alpha + beta * x[i] + delta * z[i] + eta * k[i]
  }
  alpha ~ dt( 0.0,1.0E-6,4)
  beta ~ dt( 0.0,1.0E-6,4)
  tau ~ dchisqr(4)
  sigma <- 1 / sqrt(tau)
  delta ~ dt( 0.0,1.0E-6,4)
  eta ~ dt( 0.0,1.0E-6,4)
}

```

Yukarıda tanımlanan model, $\alpha = -1$, $\beta = 1$, $\delta = 0$, $\eta = 0$, $\tau = 1$ başlangıç değerleri ile işletildiğinde, 10.000 iterasyon ve 100.000 iterasyon sonunda elde edilen sonuçlar aşağıdaki tablolarda sunulmuştur.

Tablo 4.10: 10.000 İterasyon Sonunda Model Tahminleri

Parametre	Ortalama	Standart Sapma	MC Hata	Ortanca
alpha	7.700	2.354	0.1202	7.770
beta	0.1941	0.078	0.0045	0.1918
delta	2.321	0.741	0.0434	2.322
eta	-0.172	0.208	0.0088	-0.171
tau	0.387	0.171	0.0038	0.362

Tablo 4.11: 100.000 İterasyon Sonunda Model Tahminleri

Parametre	Ortalama	Standart Sapma	MC Hata	Ortanca
alpha	7.607	2.251	0.033	7.612
beta	0.1939	0.077	0.0016	0.1934
delta	2.336	0.795	0.0171	2.341
eta	-0.163	0.218	0.0031	-0.162
tau	0.388	0.172	0.0013	0.7921

Elde edilen tahminler ile EKK tahminleri diğer tablolarda olduğu gibi büyük benzerlikler içermektedir. Burada elde edilen tek farklı sonuç, tau parametresinin dağılımına yöneliktir. Tablo 4.11'de tau parametresinin dağılımın ortancası ortalamasından oldukça büyük bir değer çıkmıştır. Bu da sola çarpık bir sonsal

dağılım oluşumunu vurgulamaktadır. Ancak burada elde edilecek sonsal dağılımın sağa çarpık bir dağılım olan ki-kare dağılımı olması yönündedir.

4.2 Hiyerarşik Çoklu Doğrusal Regresyon Modeli

Burada Kesim 4.1'de ele alınan model, hiperparametrelerin bilinmediği bir durum için ele alınacaktır. Eşitlik (4.1) ile sunulan modelde bilinmeyen parametre sayısı beş olarak kabul edilmiş ve herbir parametrenin hiperparametrelerinin bilindiği varsayılmıştır. Ancak, çoğu zaman parametrelerin dağılımlarını karakterize eden bu miktarlar bilinmez ve birer rastlantı değişkeni olarak modelde yer alırlar.

Eşitlik 4.1'de ele alınan çoklu regresyon modeli, önsellerin tümü hiyerarşik bir yapıda olacak şekilde aşağıda verilmiştir:

$$\begin{aligned} \text{Birinci aşama} & : \theta_i | v_i, \vartheta_i^2 \sim N(v_i, \vartheta_i^2) \\ \text{İkinci aşama} & : v_i \sim N(k_{i0}, V_{i0}) \\ \tau & \sim \text{Gamma}(a, b) \text{ ve } \vartheta_i^2 \sim \text{Gamma}(c_i, d_i) \end{aligned}$$

Burada $\theta=(\alpha, \beta, \delta, \eta)$ parametrelerini temsil etmektedir. Yukarıda verilen hiyerarşik modelde varyans bileşenlerine ait gamma dağılımları bağımsız dağılımlardır. Buna göre, model parametreleri $(\underline{\theta}, v, \vartheta^2, \tau)$ için bileşik önsel yoğunluk fonksiyonu şu şekilde yazılabilir:

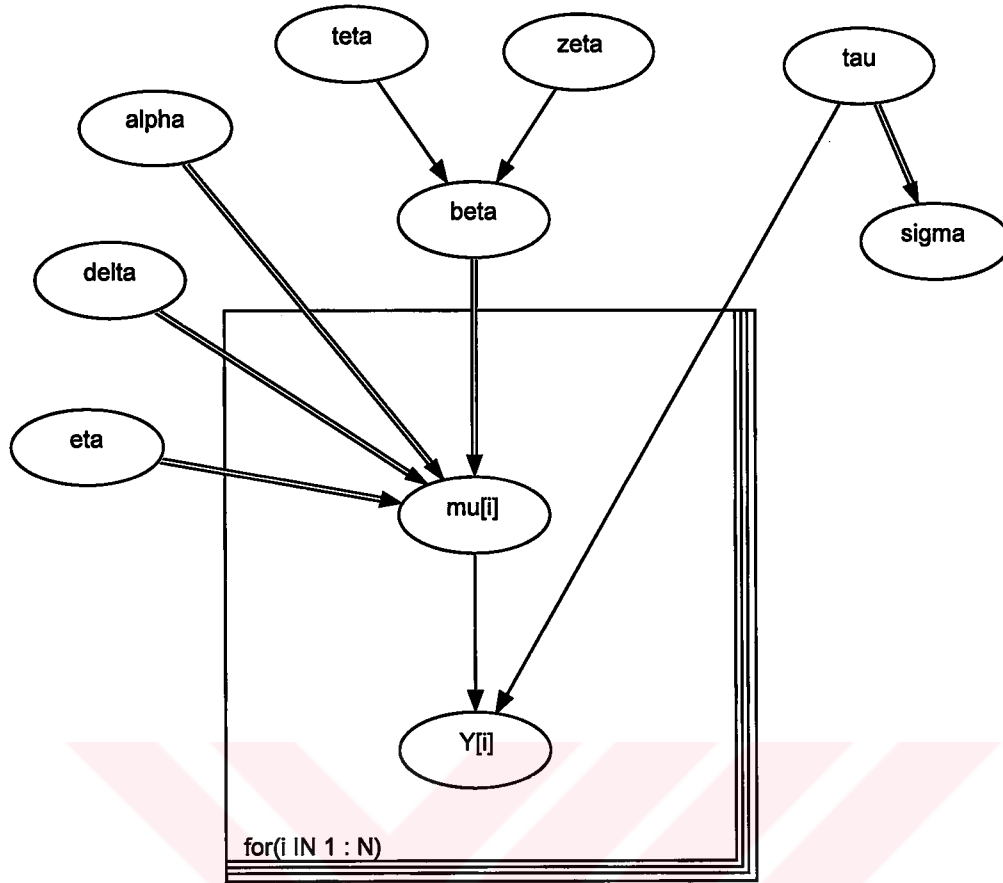
$$P(\underline{\theta}, v, \vartheta^2, \tau) = p(\underline{\theta} | v, \vartheta^2) p(v) p(\vartheta^2) p(\tau) \quad (4.2)$$

Daha açık bir şekilde, en genel anlamıyla hiyerarşik model aşağıdaki biçimde yazılabilir:

$$\begin{aligned} P(\alpha, \beta, \lambda, \eta, v_1, \vartheta_1^2, v_2, \vartheta_2^2, v_3, \vartheta_3^2, v_4, \vartheta_4^2, \tau) = & p(\alpha | v_1, \vartheta_1^2) p(\beta | v_2, \vartheta_2^2) p(\lambda | v_3, \vartheta_3^2) p(\eta | v_4, \vartheta_4^2) \\ & p(v_1) p(\vartheta_1^2) p(v_2) p(\vartheta_2^2) p(v_3) p(\vartheta_3^2) p(v_4) p(\vartheta_4^2) p(\tau) \end{aligned}$$

Modelde bilinmeyen hiperparametre sayısı ne kadar fazla olursa, modelin o denli karmaşık olacağı görülmektedir. Ayrıca Eşitlik (4.2) de yer alan hiperparametrelerin de ortalama ve varyanslarının da bilinmediği durumlar söz konusu olabilir. Bu gibi durumlarda kurulacak hiyerarşik modelde aşama sayısı artacaktır. Kesim 2.4' te de belirtildiği gibi aşama sayısı arttıkça dağılımların belirlenmesi güçleşmektedir. Ayrıca yaygın bir düşünce de yüksek aşamadaki önsellerin fazla bilgi içermediğidir. Tez çalışmasında sadece bir parametre için hiperparametrelerin bilinmediği durum ele alınmıştır. Bu durum diğer parametreler için de uygulanabilir. Gibbs örnekleme tahminlerinin bu gibi bir durumda EKK sonuçlarından ne ölçüde farklılık içereceğini görmek bu çalışmanın amaçlarından biridir. Çalışmada tanımlanan hiyerarşik model grafiksel olarak Şekil 4.3 ile verilmiştir. Burada model β parametresinin hiperparametrelerinin bilinmediği durum ele alınmıştır. Bilinmeyen hiperparametrelerden teta, beta parametresinin ortalaması olduğundan modele normal önsel ile; zeta da varyans olduğundan gamma önseliyle tanımlanmıştır. Modelin işletimi için geliştirilen BUGS kodu aşağıdadır:

```
model
{
  for( i in 1 : N ) {
    Y[i] ~ dnorm(mu[i],tau)
    mu[i] <- alpha + beta * (x[i]) +delta * (z[i]) +eta * (k[i])
    tau ~ dgamma(0.001,0.001) sigma <- 1 / sqrt(tau)
    alpha ~ dnorm(0.0,1.0E-6)
    beta ~ dnorm(0.0,1.0E-6)
    delta ~ dnorm(0.0,1.0E-6)
    eta ~ dnorm(0.0,1.0E-6)
    teta ~ dnorm(0.0,1.0E-6)
    zeta ~ dgamma(0.001,0.001) }
}
```



Şekil 4.3: Hiyerarşik Model

Tez çalışmasında ele alınan model , $\alpha = -1$, $\beta = 1$, $\delta = 0$, $\eta = 0$, $\tau = 1$, $\theta = 0$, $\zeta = 1$ başlangıç değerleri girilerek işletilmiştir. 1.000.000 iterasyon sonucunda elde edilen tahminler aşağıdaki tabloda özetlenmiştir:

Tablo 4.12: 1.000.000 İterasyon Sonunda Model Tahminleri

Parametre	Ortalama	Standart Sapma	Mc Hata	Ortanca
alpha	7.592	3.150	0.01456	7.588
beta	0.1931	0.1071	7.37E-4	0.193
delta	2.345	1.103	0.007605	2.347
eta	-0.1605	0.2961	0.001415	-0.1604
tau	0.2421	0.1394	4.068E-4	0.216
teta	-0.3298	301.5	0.291	0.1922
zeta	47.15	212.1	0.4121	0.01272

Gibbs örnekleme sonuçları ile EKK sonuçları karşılaştırıldığında, her iki tahminin de çok yakın oldukları görülebilir. Tablo 4.12'de hiperparametrelere ait dağılımların çok büyük değişikliklere sahip oldukları söylenebilir. Ayrıca bu iki parametreye ilişkin MC hatalar da diğer sonuçlara kıyasla daha büyük olarak elde edilmiştir. Bu sonuç süpriz bir sonuç değildir. Kesim 2.4'te de belirtildiği gibi, hiyerarşik modellemede bu gibi durumlarla karşılaşılması olasıdır. Hiperparametrelere ilişkin bilgi zayıflığı bu tip bir sonucun önemli bir nedenidir. Ancak burada önemli bir sonuç, hiperparametrelerin dağılımlarının bu denli zayıf olmasına rağmen, beta parametresine ilişkin tahminin doğru olmasıdır.

4.3 Poisson Gamma Hiyerarşik Modeli

Çalışmanın bu kesiminde Poisson Gamma Hiyerarşik modeli ele alınacaktır. Model kısaca anlatılacak ve grafiksel yapısı tanıtılıp; BUGS kodu geliştirilecektir. Model yapay bir veri seti için işletilecektir.

Klasik anlamda bir Poisson modeli, belirli bir zaman aralığında değişmez bir ortalama geliş oranına göre gerçekleşen olay sayılarını ifade etmektedir. Bu modelde ortalama oran, dağılımın parametresi olarak yorumlanmaktadır. Gözlem değişkeninin Poisson dağılımlı olduğu bir Bayesci model aşağıdaki şekilde tanımlanır:

$$y_i / \theta_i \sim \text{Poisson}(\theta_i)$$

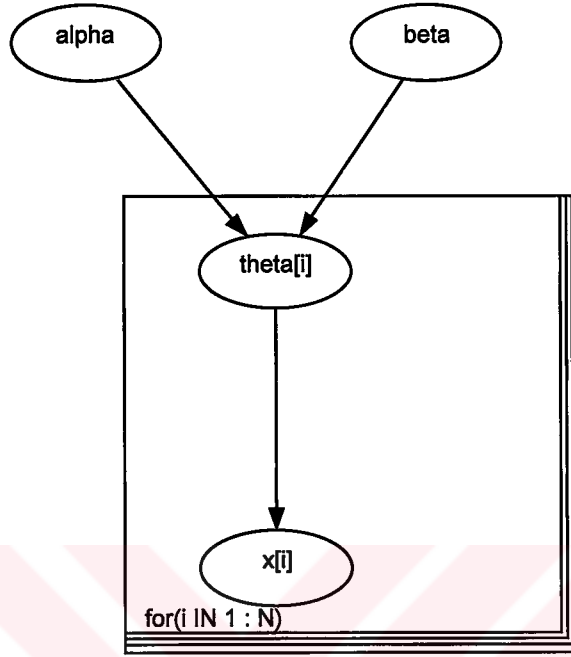
Eşlenik yapıdan ve parametrenin fiziksel özelliğinden dolayı, gamma önsel dağılımı θ parametresi için kullanılır ve aşağıdaki biçimde ifade edilir:

$$\theta_i \sim \text{Gamma}(\alpha, \beta)$$

Burada hiyerarşik bir yapı düşünüldüğünden α ve β hiperparametreleri için de önsel dağılım tanımlaması yapılabilir. Çalışmada α ve β parametreleri için,

$$\begin{aligned} \alpha &\sim \text{Üstel}(a) \\ \beta &\sim \text{Gamma}(b, c) \end{aligned}$$

tanımlamaları yapılmıştır. Burada a , b ve c pozitif değişmezler olup; modelin işletiminde $a=1$, $b=0.1$ ve $c=1$ olarak alınmıştır. Poisson-Gamma hiyerarşik modelinin grafiksel olarak sunumu aşağıdadır:



Şekil 4.4: Poisson – Gamma Hiyerarşik Modeli

Yukarıda tanımlanan modelin işletimi için geliştirilen BUGS kodu aşağıdadır.

```
model
{
  for (i in 1 : N) {
    theta[i] ~ dgamma(alpha, beta)
    x[i] ~ dpois(theta[i])
  }
  alpha ~ dexp(1)
  beta ~ dgamma(0.1, 1.0)
}
```

Poisson-Gamma modelinin işletimi için 10 genişliğinde aşağıda verilen yapay bir veri seti girilmiştir.

```
list( x = c( 5, 1, 5, 14, 3, 19, 1, 1, 4, 22), N = 10))
```

Model parametrelerinden $\alpha=1$ ve $\beta=1$ başlangıç değerleri alındığında, 10.000 iterasyon sonucu model tahminleri aşağıda verilmiştir:

Tablo 4.13: 10.000 İterasyon Sonunda Model Tahminleri

Parametre	Ortalama	Standart Sapma	MC Hata	Ortanca
alpha	1.077	0.5368	0.01507	0.9776
beta	0.1461	0.09161	0.002598	0.1288
teta[1]	5.298	2.191	0.02278	4.983
teta[2]	1.792	1.281	0.01583	1.511
teta[3]	5.285	2.153	0.02173	5.016
teta[4]	13.21	3.525	0.03797	12.86
teta[5]	3.55	1.763	0.01808	3.278
teta[6]	17.66	4.097	0.04807	17.31
teta[7]	1.78	1.281	0.01735	1.496
teta[8]	1.775	1.283	0.01625	1.481
teta[9]	4.383	1.964	0.02104	4.076
teta[10]	20.23	4.370	0.05329	19.91

Gibbs Örnekleme yöntemi uygulanarak elde edilen sonsal tahminler, α için başlangıç değerine yakın çıkmışken; β için aynı şeyi söylemek kolay değildir. θ parametresine ilişkin tahminler ise, daha çok örneklem bilgisinden etkilenmiş görülmektedir. Bunun önemli nedenlerinden biri, bir önceki uygulamada da karşılaşılan problemin ortaya çıkmasıdır. Tanımlanan önsel dağılımların zayıf olmaları, θ 'nın güncelleştirilmesinde örneklem bilgisinin ön plana çıkmasına neden olmuştur. Aynı başlangıç değerleri ile işletilen modelin 100.000 iterasyon sonucundaki tahminleri aşağıdadır:

Tablo 4.14: 100.000 İterasyon Sonunda Model Tahminleri

Parametre	Ortalama	Standart Sapma	MC Hata	Ortanca
alpha	1.083	0.4916	0.003927	0.9933
beta	0.1457	0.08143	6.508E-4	0.1305
teta[1]	5.313	2.17	0.007342	5.016
teta[2]	1.806	1.293	0.004921	1.518
teta[3]	5.31	2.17	0.006631	5.013
teta[4]	13.21	3.46	0.01176	12.91
teta[5]	3.558	1.776	0.005704	3.275
teta[6]	17.59	4.042	0.0154	17.26
teta[7]	1.804	1.293	0.004913	1.514
teta[8]	1.796	1.287	0.004789	1.51
teta[9]	4.439	1.986	0.006505	4.141
teta[10]	20.23	4.349	0.01651	19.9

Burada da Tablo 4.13'e benzer sonuçlar elde edilmiş olup, iterasyon sayısının artırılması sadece MC Hata'larının küçülmesini sağlamıştır. Poisson-Gamma

hiyerarşik modelinin buradaki gibi bir yapay veri seti yerine gerçek bir veri seti ile işletilmesi, model parametrelerinin daha anlamlı yorumlanmasını sağlayacaktır. Bu amaçla kurulan bu modelin ileride gerçek bir uygulamada kullanılması önerilebilir.

Sonuç olarak; kurulan her üç model için grafiksel sunumlar elde edilmiş, modeller için BUGS kodları geliştirilmiştir. Ardından modeller farklı iterasyon sayıları, farklı başlangıç değerleri ve farklı önsel dağılımlar ile işletilmiştir. Elde edilen tahminlerin bilinen tahminlerle karşılaştırılmasında ilk iki model sonuçları kullanılmış, tahminlerin analitik sonuçlara olan yakınlığı gözlenmiştir. İterasyon sayısının artırılmasının bilgisayarlara ek bir zaman yükü getirmesine karşın daha hassas tahminler elde edilmesini sağladığı görülmüştür. Çalışmaya ilişkin genel sonuç izleyen bölümde verilecektir.



5. SONUÇLAR

Bu tez çalışmasının amacı, standart istatistiksel tekniklerin doğal bir uzantısı olan Bayesci grafiksel modellerin tanıtımıdır. Burada özellikle modellerin yapısal gösterimi ve çıkarsama yapabilmek için kısmi hesaplama yöntemleri üzerinde durulmuştur.

Ancak, burada ele alınan tekniklerin bazı sıkıntılarının da gözardı edilmemesi gerekmektedir. Birincisi, Bayesci yaklaşımın doğasından gelen bir problem olan önsel dağılımın tanımlanma sorunudur. Çoğu zaman önsel dağılımla ilgili kararsızlık sonuçları olumsuz ve yanlış yönde etkileyebilmektedir. İkincisi ise, MCMC metotların doğruluğunun test edilmesinin zor olan bir simülasyona dayalı olmasıdır (Bu teknik bir yaklaşım tekniğidir). Elde edilen çözümler analitik olarak elde edilmediğinden hata kaçınılmazdır. Üçüncüsü, MCMC tekniklerinin bir çok modelin tahmininde ve sonuçlar üzerindeki etkilerinin görülmesinde uygun bir teknik olmamasıdır. Dördüncüsü, bir model için parametrelerin belirlenmesi simülasyon sonuçları üzerinde çok önemli etkilere sahiptir. Bu nedenle uygun parametrelendirme, sonuçlar üzerinde önemli etkiye sahiptir.

Bütün bu sıkıntılara rağmen, grafiksel yaklaşım modellerin çok iyi bir şekilde anlaşılmasını sağlamakta ve bu yaklaşımın istatistik bilimine çok büyük faydalar getireceğine inanılmaktadır.

KAYNAKLAR

1. Abraham K. J. ve Haines, L. M., 1999, "A New Technique for Sampling Multi-Modal Distributions", Dept. of Physics and Astronomy, Iowa State University, Ames I . Email: abraham@iastate.edu, haines@stat.unp.ac.za. Available as file with non-standard extension. Abstract also available as file with non-standard extension.
2. Berger, J.O. ve Bernardo, J.M., 1992, On the Development of Reference Priors, Bayesian Stat. 4, Oxford University Press.
3. Berger, J.O. ve Mendoza, A., 1983, The Robust Bayesian Viewpoint, in Robustness in Bayesian Statistics, ed. J.Kadane, Amsterdam.
4. Bernardo, J.M., 1979, Reference Posterior Distribution for Bayesian Inference, J. Royal Stats. Soc. B, 41, 113-147.
5. Besag, J., 1974, Spatial Interaction and the Statistical Analysis of Lattice Systems, J. Royal Stats. Soc. B, 36, 192-236.
6. Breslow, N., 1996, Biostatistics and Bayes, Statistical Science, 5, 269-298.
7. Brooks, S.P., Giudici, P. ve Philippe, A., 2001, Nonparametric Convergence Assessment for MCMC Model Selection, University of Cambridge. Email: steve@statslab.cam.ac.uk. Available as Postscript. Abstract also available as Hypertext.
8. Dawid, A.P., 1979, Conditional Independence in Statistical Theory, J. Royal Stat. Soc. B, 41, 1-31.
9. DeGroot , M. H., 1970, Optimal Statistical Decisions, New York: Mc Graw-Hill.
10. Efron, B., 1986, Why isn't Everyone a Bayesian?, American Statistician, 40,1-11.
11. Ergün, G., 1995, Devingen Doğrusal Modeller ve Bayesci Öngörüler Üzerine Bir Çalışma, Yayınlanmamış Doktora Tezi, H.Ü. İstatistik Bölümü.
12. Frydenberg, D., 1990, The chain Graph Markov Property, J. Scand. Stats. 17, 333-353.
13. Gamerman, D., 1997, Markov Chain Monte Carlo, Chapman & Hall.
14. Gelfand, A. E. ve Smith, A. F. M., 1990, Bayesian Computation, New York: Wiley

15. Gelman, A. ve Rubin, D.R., 1992b, Inference from Iterative Simulation Using Multiple-Sequences, *Statistical Science*, 7, 457-511.
16. Gelman, A. ve Rubin, D.R., 1992a, A Simple Series From the Gibbs Sampler Provides a False Sense of Security, in *Bayesian Statistics 4*, Oxford: Oxford University Press, 625-631.
17. Geman, S. ve Geman, D., 1984, Stochastic relaxation, Gibbs Distributions and the Bayesian restoration of images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, 721-741.
18. Geyer, C.J., 1992, Practical Markov Chain Monte Carlo, *Statistical Science*, 7, 473-511.
19. Gilks, W.R., 1993, Modelling Complexity: Application of Gibbs Sampling in Medicine, *J. Royal Stats. Soc. B*, 55, 39-102.
20. Gilks, W.R., 1994, Adaptive Rejection Metropolis Sampling within Gibbs Sampling, *Applied Statistics*, 44, 455-472.
21. Gilks, W.R., Richardson, S. ve Spiegelhalter, D., 1996a, *Practical Markov Chain Monte Carlo*, New York: Chapman & Hall.
22. Han, C. ve Carlin, B.P., 2000, MCMC methods for computing Bayes factors: a comparative review, Division of Biostatistics, University of Minnesota. Email: brad@biostat.umn.edu.
23. Lange, K., 1992, Normal Independent Distributions and Their Applications in Robust regression, *Journal of Computational and Graphical Stat.*, 2, 175-198.
24. Lauritzen, S. ve Wermuth, 1989, The EM Algorithm for Graphical Association Models with Missing Data, *Comp. Stat. And data Anal.*, 19, 191-201.
25. Lauritzen, S., 1986, *Graphical Models*, Oxford: Clarendon.
26. Lauritzen, S.L. ve Spiegelhalter, D.J., 1988, Local Computations with Probabilities on Graphical Structures and Their Application to Expert Systems, *J. Royal Stats. Soc. B*, 50, 157-224.
27. Lindley, D.V. ve Smith, A.F.M., 1972, Bayes Estimates for Linear Models, *J. Royal Stats. Soc. B*, 34, 1-41.
28. Lindley, D.V., 1978, *Bayesian Statistics, A Review*, Society for Industrial and Applied Mathematics.

29. Liu, J., Wong, W. ve Kong, A., 1994, Correlation Structure and Converge Rate of the Gibbs Sampler: Application to the Comparison of Estimators and Augmentation Schemes, *Biometrika*, 81, 27-40.
30. Madigan, D. ve York, A.E., 1994, Model selection and Accounting for Model Uncertainty in Graphical Models using Occam's Windows, *JASA*, 89, 1535-1546.
31. McCullagh, P. ve Nelder, J.A., 1988, *Generalized Linear Models*, New York: Chapman & Hall.
32. Normand, S. ve Tritchler, D., 1992, Parameter Updating in a Bayes Network, *JASA*, 87, 1109-1117.
33. O'Hagan, A., 1986, *Probability: Methods and Measurement*, New York: Chapman & Hall.
34. Pearl, J., 1988, *Probabilistic Inference in Intelligent Systems*, San Mateo: Morgan Kaufmann.
35. Queen, C. M. ve Smith, J. Q., 1993, Multiregression Dynamic Models, *J. R. Statist. Soc. B*, 55, 849-870.
36. Raftery, A.E. ve Lewis, J., 1992, Inference for the Binomial N Parameter: A Hierarchical Bayes Approach, *Biometrika*, 75, 223-228.
37. Richardson, S. ve Gilks, W.R., 1993, A Bayesian Approach to Measurement Error Problems in Epidemiology Using Conditional Independence Models, *American Journal of Epidemiology*, 138, 430-442.
38. Smith, A.F.M. ve Robert, G.O., 1994, Bayesian Computation via the Gibbs Sampler and Related Markov Chain Monte Carlo, *J. Royal Stats. Soc. B*, 55, 3-102.
39. Smith, A.F.M. ve Roberts, G.O., 1993, Bayesian Computation via the Gibbs Sampler and Related Markov Chain Monte Carlo Methods, *J. Royal Stats. Soc. B*, 55, 3-23.
40. Smith, A.F.M., 1984, Bayesian Approaches to Outliers and Robustness, in *Specifying Statistical Models*, ed. JP Fdoren, Mmouchart.
41. Spiegelhalter, D.J., 1998, Bayesian Graphical Modelling: A Case-study in Monitoring Health Outcomes, *Applied Stat.*, 47, 115-133.
42. Spiegelhalter, D.J., Thomas, A. ve Best, N.G., 1996, Computation on Bayesian Graphical Models, *Bayesian Statistics* 5, 407-425.

43. West, M. ve Harrison, P.J., 1997, Bayesian Forecasting and Dynamic Models, 2nd. Edition, Springer Verlag.
44. Whittaker, J., 1990, Graphical Models in Applied Multivariate Analysis, Chichester: Wiley.
45. Zeger, S.L. ve Karim, M.R., 1991, Generalized Linear Models with Random Effects: A Gibbs Sampling Approach, JASA, 86, 79-86.

