

T.C.
MARMARA ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
EKONOMETRİ ANABİLİM DALI
EKONOMETRİ BİLİM DALI

**OECD ÜLKELERİNDE DAYANIKLI ÖĞRENCİLERİN AKADEMİK BAŞARISINI
ETKİLEYEN FAKTÖRLERİN NONPARAMETRİK BAYESYEN
REGRESYON TAHMİNİ**

Doktora Tezi

DERYA TOPDAĞ

İstanbul, 2023

T.C.
MARMARA ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
EKONOMETRİ ANABİLİM DALI
EKONOMETRİ BİLİM DALI

**OECD ÜLKELERİNDE DAYANIKLI ÖĞRENCİLERİN AKADEMİK BAŞARISINI
ETKİLEYEN FAKTÖRLERİN NONPARAMETRİK BAYESYEN
REGRESYON TAHMİNİ**

Doktora Tezi

DERYA TOPDAĞ

Danışman: PROF. DR. EBRU ÇAĞLAYAN AKAY

İstanbul, 2023

ÖZET

OECD ÜLKELERİNDE DAYANIKLI ÖĞRENCİLERİN AKADEMİK BAŞARISINI ETKİLEYEN FAKTÖRLERİN NONPARAMETRİK BAYESYEN REGRESYON TAHMİNİ

Ekonometrik analizlerde klasik yaklaşıma bir alternatif olarak Bayesyen yaklaşımın sıklıkla kullanıldığı görülmektedir. Bayesyen yaklaşım, klasik yaklaşımdan farklı olarak rastgele parametrelerle birlikte küçük örneklem için daha doğru tahminler sağlamaktadır. Nonparametrik yaklaşımda hata dağılımı ve ortalama fonksiyonunun fonksiyonel formu hakkında kısıtlayıcı parametrik varsayımlar yapılmamaktadır. Nonparametrik Bayesyen modeller kısıtlayıcı parametrik ilk dağılım varsayımını esneterek verileri modellemek için daha gerçekçi bileşenler sağlamaktadır. Nonparametrik Bayesyen yaklaşım sayesinde tüm rastgele etkiler desteklenerek model tabanlı bir kümeleme analizi sağlanmaktadır. Böylece klasik yöntemlerin hesaplama sorunuyla karşılaştıkları çok sayıda rastgele etki içeren karmaşık modellerin herhangi bir hesaplama sorunuyla karşılaşmadan analiz edilmesine olanak sağlanmaktadır.

Bu tez çalışmasında 16 OECD ülkesindeki dayanıklı öğrencilerin akademik başarısını etkileyen faktörler nonparametrik Bayesyen regresyon modelleri yardımıyla analiz edilmektedir. Tahmin sonuçları, akademik beklenti ve özgüveninin dayanıklı öğrencilerin matematik başarısını pozitif etkileyen en önemli ve tutarlı faktörler olduğunu göstermektedir. Cinsiyet parametresine ilişkin bulgular, dayanıklı öğrenciler için cinsiyet etkisinin varlığını göstermektedir. Zorbalığın tüm ülkelerdeki öğrencilerin matematik performansı üzerinde negatif etkiye sahip olduğu tespit edilmiştir. Ülkeler arası elde edilen bulgular akademik dayanıklılığı artırmaya yönelik politikaların ülkeler bağlamında özel olarak ele alması gerektiğini göstermektedir.

Anahtar Kelimeler: Bayesyen Çıkarım, Nonparametrik Bayesyen Regresyon, Lineer Bağımlı Dirichlet Süreç Modeller (LDDP), Dayanıklı Öğrenciler, TIMSS

ABSTRACT

NONPARAMETRIC BAYESIAN REGRESSION ESTIMATION OF FACTORS AFFECTING THE ACADEMIC ACHIEVEMENT OF RESILIENT STUDENTS IN OECD COUNTRIES

Econometric analysis shows that the Bayesian approach is often used as an alternative to the classical approach. Unlike the classical approach, the Bayesian approach provides more accurate estimates for small samples along with random parameters. The nonparametric approach does not make restrictive parametric assumptions about the functional form of the mean function and the distribution of errors. Nonparametric Bayesian models provide more realistic components for modeling data by stretching the strict parametric prior distribution assumption. By means of the nonparametric Bayesian approach, a model-based clustering analysis is provided by supporting all random effects. Thus, it is possible to analyse complex models with a large number of random effects, where classical methods encounter computational issues, without any computational issues.

In this dissertation, nonparametric Bayesian regression models are used to analyse the factors affecting the academic success of resilient students in 16 OECD countries. The estimation results indicate that academic expectation and self-confidence are the most significant and consistent factors affecting the mathematics achievement of resilient students. Findings concerning the gender parameter indicate a gender effect for resilient students. Bullying has been found to have a negative impact on the mathematical achievement of students in all countries. Cross-country findings indicate that policies to increase academic resilience should be addressed specifically in the context of countries.

Keywords: Bayesian Inference, Nonparametric Bayesian Regression, Linear Dependent Dirichlet Process Models (LDDP), Resilient Students, TIMSS

İÇİNDEKİLER

ÖZET.....	i
ABSTRACT	ii
TABLolar LİSTESİ	vi
ŞEKİLLER LİSTESİ.....	vii
KISALTMALAR	viii
1.GİRİŞ	1
2.BAYESYEN YAKLAŞIM.....	5
2.1. Bayes Teoremi.....	7
2.1.1. Olasılıklar Oranı ve Bayes Faktörü	9
2.1.2. Bayesyen İlk Dağılımlar.....	11
2.1.2.1. Eşlenik İlk Dağılımlar	12
2.1.2.2. Bilgi Vermeyen İlk Dağılımlar.....	14
2.1.2.2.1. Uniform İlk Dağılım.....	14
2.1.2.2.2. Jeffrey İlk Dağılımı	15
2.1.2.2.3. Referans İlk Dağılımlar	16
2.1.2.3. Bilgi Veren İlk Dağılımlar.....	17
2.1.2.4. Hibrit İlk Dağılımlar.....	18
2.1.2.4.1. Maksimum Entropi İlk Dağılımlar	18
2.1.2.4.2. Histogram İlk Dağılımlar	19
2.2. Simülasyon Yöntemleri	20
2.2.1. Monte Carlo Entegrasyonu.....	21
2.2.2. Son Dağılım Simülasyonları için Algoritmalar	22
2.2.2.1. Red Örnekleme.....	23

2.2.2.2. Önem Örneklemesi	23
2.2.3. Markov Zinciri Monte Carlo Yöntemleri	25
2.2.3.1. Markov Zincirlerinin Genel Özellikleri.....	27
2.2.3.1.1. İndirgenemezlik	28
2.2.3.1.2. Yenileme	28
2.2.3.1.3. Homojenlik.....	29
2.2.3.1.4. Durağanlık	30
2.2.3.1.5. Ergodiklik.....	31
2.2.3.2. Metropolis-Hastings Algoritması	32
2.2.3.3. Gibbs Örneklemesi	33
2.2.3.4. Hibrit Monte Carlo	34
2.2.4. Bayesyen Hiyerarşik Modeller	36
3.NONPARAMETRİK BAYESYEN YAKLAŞIM VE DIRICHLET SÜREÇLER	38
3.1. Dirichlet Süreçleri	42
3.2. Dirichlet Süreç Karışım Modelleri (DPM).....	46
3.3.Nonparametrik Bayesyen Regresyon Modelleri.....	51
3.3.1. Bağımlı Nonparametrik Süreçler.....	54
3.3.1.1. Bağımlı Dirichlet Süreç Modeli (DDP).....	56
3.3.1.1.1. ANOVA Bağımlı Dirichlet Süreç Modeli (ANOVA-DDP).....	60
3.3.1.1.2. Lineer Bağımlı Dirichlet Süreç Modeli (LDDP)	62
3.4. Nonparametrik Bayesyen Regresyon Modelleri için Son Dağılım Algoritmaları.....	66
3.4.1. Escobar ve West (1995) Gibbs Algoritması	66
3.4.2. MacEachern ve Müller (1998) Algoritması.....	68
3.4.3. Neal (2000) Metropolis-Hastings Algoritması	69
3.4.4. Neal (2000) Yardımcı Parametrelere Sahip Gibbs Algoritması	71
3.4.5. Kalli vd.(2011) Kesit Örnekleyici	73

4.UYGULAMA: OECD ÜLKELERİNDE DAYANIKLI ÖĞRENCİLERİN MATEMATİK BAŞARISINI ETKİLEYEN FAKTÖRLERİN BELİRLENMESİ.....	75
4.1. Akademik Dayanıklılık	76
4.2. Veri Seti ve Değişkenler.....	83
4.3. Bulgular	92
4.3.1. Öğrenci Özellikleri	94
4.3.2.Öğrenci Tutumu ve Beklentisi.....	98
4.3.3.Okul İklimi	102
4.3.4. Optimal Kümeleme Algoritması	109
5.SONUÇ	112
KAYNAKÇA	116
6.EKLER	132
Ek 1: Kolmogorov-Smirnov testi ve Q-Q grafiği.....	132
Ek 2: Dirichlet Süreç Karışım Model Parametrelerinin Marjinal Son Dağılım Özeti.....	132
Ek 3: Model Sabit Terimi Son Dağılım Özeti	133
Ek 4: Regresyon Hata Varyans Parametresi İçin Son Dağılım Özeti.....	134
Ek 5: Dirichlet Süreç Temel Dağılım Parametrelerinin Son Dağılım Özeti	135

TABLÖLAR LİSTESİ

Tablo 1: Farklı Olabilirlik Fonksiyonları için Eşlenik İlk Dağılımlar.....	13
Tablo 2: Son Dağılım Özellikleri	22
Tablo 3: Ülkelere Göre Düşük SES’e Sahip Öğrencilerin ve Akademik Dayanıklı Öğrencilerin Yüzdesi	86
Tablo 4: Değişken Tanımları.....	88
Tablo 5: Tanımlayıcı İstatistikler	89
Tablo 6: Ülkelere Göre Tanımlayıcı İstatistikler.....	91
Tablo 7: Model Uyum İstatistikleri	94
Tablo 8: Cinsiyet Parametresinin Marjinal Son Dağılım Tahmini.....	95
Tablo 9: Yaş Parametresinin Marjinal Son Dağılım Tahmini.....	97
Tablo 10: Akademik Beklenti Parametresinin Marjinal Son Dağılım Tahmini	99
Tablo 11: Özgüven Parametresinin Marjinal Son Dağılım Tahmini.....	101
Tablo 12: Okula Ait Hissetme Parametresinin Marjinal Son Dağılım Tahmini	103
Tablo 13: Pozitif Öğretmen Beklentisi Parametresinin Marjinal Son Dağılım Tahmini	105
Tablo 14: Zorbalık Parametresinin Marjinal Son Dağılım Tahmini	107
Tablo 15: Benzer Parametre Tahmin Sonuçları	110
Tablo 16: Optimal Kümeleme Sonuçları.....	110

ŞEKİLLER LİSTESİ

Şekil 1: Jeffrey İlk Dağılımının Değişmezliği.....	16
Şekil 2: Evde Eğitim Kaynaklarına Göre Ortalama Matematik Başarısı	85
Şekil 3: Dayanıkl Öğrencilerin Matematik Başarısının Kernel Yoğunluk Tahmini	90
Şekil 4: Ülkelere Göre Dayanıkl Öğrencilerin Matematik Puan Dağılımı.....	91
Şekil 5: Model Parametrelerine Ait İz Grafiği	93

KISALTMALAR

ABD	: Amerika Bileşik Devletleri
APA	: Amerika Psikoloji Derneği
ANOVA	: Varyans Analizi
BNP	: Noparametrik Bayesyen Modeller
DP	: Dirichlet Süreçler
DDP	: Bağımlı Dirichlet Süreçleri
DPM	: Dirichlet Süreç Karışımı
DI	: Demokrasi Endeksi
GII	: Cinsiyet Eşitsizliği Endeksi
GNI	: Kişi Başı Gayri Safi Milli Gelir
HDI	: İnsani Gelişmişlik Endeksi
HER	: Evde Eğitim Kaynakları Endeksi
HMC	: Hibrit Monte Carlo
IEA	: Uluslararası Eğitim Başarısı Değerlendirme Derneği
LDDP	: Lineer Bağımlı Dirichlet Süreç Modeller
MC	: Monte Carlo
MCMC	: Markov Zinciri Monte Carlo
MCSE	: Monte Carlo Standart Hatalar
MYS	: Ortalama Eğitim Yılı
NAEP	: Ulusal Eğitim İlerleme Değerlendirme Programı
NCES	: Ulusal Eğitim İstatistikleri Merkezi
OECD	: Ekonomik Kalkınma ve İşbirliği Örgütü
PISA	: Uluslararası Öğrenci Değerlendirme Programı
SDG	: Sürdürülebilir Kalkınma Hedefleri
SES	: Sosyoekonomik Statü
SSE	: Hata Kareler Toplamı
STEM	: Bilim, Teknoloji, Mühendislik ve Matematik Alanları
TIMSS	: Uluslararası Matematik ve Fen Eğilimleri Araştırması
UNDP	: Birleşmiş Milletler Kalkınma Programı
UNESCO	: Birleşmiş Milletler Eğitim, Bilim ve Kültür Örgütü
UNICEF	: Birleşmiş Milletler Çocuklara Yardım Fonu
WDI	: Dünya Kalkınma Göstergeleri

1.GİRİŞ

Bayesyen yaklaşım ilk olarak inançları güncellemek için Bayes'in kendi adını taşıyan temel teoremin kullanılmasıyla birlikte ortaya çıkmıştır. 20. yüzyılda klasik istatistiğin hızlı gelişimiyle birlikte Bayesyen yöntemler daha yavaş bir gelişme gösterse de 1990'lardan bu yana simülasyon tabanlı tekniklerin gelişmesiyle birlikte Bayesyen yaklaşımın literatüre sağlam bir şekilde yerleştiği görülmektedir. Günümüzde istatistiğin uygulandığı birçok uygulamalı ve temel araştırma alanlarında Bayesyen yaklaşımın kullanıldığı yöntemler önemli bir yere sahiptir.

Bayesyen yaklaşım ile klasik yaklaşım arasındaki temel fark iki yaklaşım için olasılığın farklı şekilde tanımlanmasından kaynaklanmaktadır. Bu fark olasılığın belirli bir sonucun uzun vadeli sıklığı olarak mı yoksa bilginizin belirsizliği olarak mı ele alındığı ile ilgilidir. Bunun sonucu olarak da klasik yaklaşım olasılığı uzun vadeli frekansla ilişkilendirirken, Bayesyen yaklaşım olasılığı subjektif belirsizlik deneyimi olarak yorumlamaktadır. Ayrıca klasik yaklaşımda tümdengelim mantık ile bir sebepten başlanarak sonuçlar elde edilmek istenmektedir ve bu akıl yürütme genelden özele doğru çalışmaktadır. Bayesyen yaklaşımda ise tümevarımsal çıkarım, özelden genele doğru çalışmaktadır. Tümevarımsal mantık sayesinde Bayesyen yaklaşım, gerçek hayattaki belirsizliklerin belirli bir inanç seviyesi ile modellenmesini sağlamaktadır.

Bayesyen yaklaşımın klasik yaklaşımdan en önemli farklılıklarından biri de parametrelerin rastgele nicelikler olduğu ve parametre hakkında ön bilgiler içerebileceği varsayımına dayanmasıdır. Bayesyen yaklaşım, Bayes teoremini kullanarak gözlemlenen verilerden gelen bilgiler ile ön bilgileri birleştirerek parametrelerin güncellenmiş son dağılımının elde edilmesini sağlar. Böylece, tamamen veri odaklı klasik yaklaşımın aksine Bayesyen analizler yalnızca eldeki verileri değil, aynı zamanda parametreler için mevcut bilgileri de kullanarak daha sağlam bir yaklaşım sunar. Bu yaklaşım etkilerin büyüklüğü ve belirsizliği hakkındaki bilgilerin aşamalı olarak gelişmesine olanak tanır ve böylece bilimsel ilerlemeyi kümülatif hale getirir.

Sosyal bilimlerde özellikle özel bir grup ile çalışıldığında, araştırma sorusuna ve seçtiğimiz alt örnekleme bağlı olarak küçük örneklemeler ortaya çıkmaktadır. Klasik yaklaşımda (en çok olabilirlik tahmini gibi) küçük örneklem genellikle düşük istatistik güç ve sapmalı tahminlere neden olmaktadır. Bayesyen yaklaşımın kullanılmasının önemli motivasyonlarından biri de küçük örneklemeler söz konusu olduğunda, ön bilgileri dağılımlar yoluyla modele dahil ederek parametrelerin daha doğru tahmin edilmesine olanak sağlamasıdır. Yani Bayesyen yaklaşımda tahmin kesinliği örneklem büyüklüğü ile

sınırlanmamaktadır. Böylece Bayesyen yöntemler küçük örneklem boyutlarına sahip verileri modellemek için daha donanımlı olmaktadır.

Ayrıca Bayesyen analiz, istatistiksel modelleme ve sonuçların yorumlanması açısından geleneksel çıkarımlara göre daha kapsamlı ve esnek bir yaklaşım sunar. Klasik yaklaşımda bilinmeyen sabit parametrelerin gerçek değerlerine yaklaşmak için tahminciler kullanır ve genellikle modelin gerektirdiği varsayımların karşılanıp karşılanmadığına bağlıdır. Bayesyen yaklaşım parametrelerin tam bir son dağılımını sunar ve çıkarım için en çok olabilirlik gibi birçok klasik yaklaşımda olduğu gibi asimptotik Normallik varsayımına dayanmamaktadır. Bayesyen çıkarım, belirli istatistiksel problemler ve modeller için tasarlanmış çeşitli tahmin yöntemlerine dayanan klasik yaklaşımın aksine, tüm modellere uygulanabilen tek bir olasılık kuralına, Bayes teoremine dayanmaktadır. Bu durum Bayesyen yaklaşımı evrensel kılar ve uygulanıp yorumlanmasını büyük ölçüde kolaylaştırır.

Diğer taraftan standart regresyon yaklaşımları, konum ve ölçek gibi bazı sonlu sayıda yanıt dağılımı özelliklerinin bağımsız değişkenin bir fonksiyonu olarak değiştiğini varsaymaktadır. Ancak, hata dağılımının tahmin uzayı üzerinde değişmeyen bir şekle sahip olduğunu varsaymak her zaman uygun değildir. Yanıtların dağılımının, bağımsız değişkenlerle, sonlu boyutlu bir fonksiyonel form tarafından makul bir şekilde temsil edilemeyeceği uygulamalı araştırmalarda görülmektedir. Bu nedenle fonksiyonel form ve hataların dağılımı hakkında kısıtlayıcı varsayımlar yapmayan daha genel yaklaşımlara ihtiyaç duyulmaktadır. Bu durum, nonparametrik regresyon modellerinin gelişmesine yol açmıştır. Nonparametrik regresyon modellerinin birçok konuda yararlı olduğu söylenebilir. Bu yaklaşımın en büyük katkısı fonksiyonel form varsayımlarına daha az bağımlı olan tahminci ve çıkarım yöntemleri sunmasıdır.

Bayesyen çerçevede, parametrik bir model varsayımı, son derece güçlü bir ilk dağılım fikri olarak görülebilir. Bu nedenle parametrik Bayesyen modelleme, olasılığı tüm yoğunlukların zayıf bir alt kümesine atayan bir ilk dağılımda ısrar etmeye eşdeğerdir. Bu durumda önceki inançları özetleyen ve tüm belirsizlikleri barındıracak kadar büyük olan bir ilk dağılım gerekmektedir. Yani geleneksel parametrik dağılım ailesinden çok daha geniş bir dağılım ailesine ihtiyaç vardır. Bu durum sonsuz sayıda parametreye izin vererek daha esnek modeller ve üretici yaklaşımlar sunan nonparametrik Bayesyen (Nonparametric Bayesian Regression, BNP) modellerin gelişmesinin temel motivasyonunu oluşturmuştur.

Nonparametrik Bayesyen yaklaşım, model seçimi sorununa farklı bir çözüm sunmaktadır. BNP yaklaşımı farklı karmaşıklığa sahip modelleri karşılaştırmak yerine, karmaşıklığı verilere uyarlayarak tek bir modele sığdırmaktadır. Yani, verilerin modelin karmaşıklığını belirlemesine izin vererek bu sorunu ortadan kaldırmaktadır. Nonparametrik Bayesyen yaklaşım sınırsız karmaşıklığa sahip bir model

kullanarak eksik uyum sorunu hafifletilirken, Bayesci yaklaşım hesaplaması veya parametreler üzerinden tüm son dağılıma yakınsama yaklaşımı aşırı uyum sorununu hafifleterek nonparametrik model seçimine de bir alternatif sunmaktadır. Ayrıca bu modeller, yerel davranışın yaygın bir şekli olan aykırı değerleri ele almada özellikle iyidir. Bunu, aykırı değerleri modelin farklı bir bileşeniyle modellemek yerine, verilerdeki neredeyse her örüntüyü yakalayabilen esnek ilk dağılımı kullanarak yapmaktadır.

Diğer taraftan bağımlı nonparametrik süreçler, ortak değişkenle gelişen yanıt dağılımlarına izin verdiğinden, bireysel durumlar hakkında çıkarımı kolaylaştırır ve ortak değişkenin farklı seviyeleri için farklı şekillerde tahmin dağılımların oluşmasına izin vermektedir. Bunu normal hata dağılımını bağımlı Dirichlet süreçlerinin bir karışımı kullanılarak gelişen bir hata dağılımı ile değiştirerek yapar. Tüm bu genellemeler hataların değişen varyansının modellenmesini sağlar. Böylece genelleştirilmiş doğrusal modelin de ilerisine geçilmesine izin verir. Sonuçta bağımlı Dirichlet süreçler, ortak değişken verildiğinde yanıtın hem ortalama yapısının hem de dağılım şeklinin uygun biçimde modellenmesine izin vererek verileri modellemek için daha gerçekçi bileşenler sağlar. Böylece çarpıklık, aykırı değerler, kalın kuyruklar (tümsekler) ve iki modluluk gibi normallikten sapmaların modellenmesi sağlanır.

Bu tez çalışmasında, nonparametrik Bayesyen regresyon modeli olan lineer bağımlı Dirichlet süreç (Linear Dependent Dirichlet Process, LDDP) modeli kullanılmıştır. Dirichlet süreç karışım modelleri, özellikle gruplanma problemlerinde kullanılan olasılıksal bir modeldir. Bu modeller farklı gruplar üzerinde modeller oluşturulmasına, parametrelerin gruplar arasında farklılaşmasına ve grup düzeyindeki etkileri tahmin ederken belirsizliğin hesaba katılmasına izin vermektedir. Ayrıca Bayesyen yaklaşım model parametreleriyle birlikte rastgele etkileri de tahmin edilmesine olanak sağlar. Nonparametrik Bayesyen yaklaşım sayesinde tüm rastgele etkilerin desteklenir ve model tabanlı bir kümeleme analizi sağlanır. Böylece klasik yöntemlerin hesaplama sorunuyla karşılaştıkları çok sayıda rastgele etki veya hiyerarşik yapı içeren karmaşık modellerin herhangi bir hesaplama sorunuyla karşılaşmadan analiz edilmesine olanak sağlanır.

Bu tez çalışmasında amaç, Bayesyen yaklaşımdaki kısıtlayıcı parametrik ilk dağılım varsayımını esneterek verileri modellemek için daha gerçekçi ve esnek bileşenler sağlayan nonparametrik Bayesyen modellerin bir uygulamasını sunmaktır. Uygulamada genellikle küçük örneklem boyutlarıyla ilişkilendirilen düşük sosyoekonomik statüye sahip olmalarına rağmen zorlukları yenerek akademik olarak başarılı olan özel bir öğrenci grubuna odaklanılmaktadır. Bu bağlamda, 16 OECD ülkesindeki dayanıklı öğrencilerin akademik başarısını etkileyen faktörler karmaşık modellerin esnek bir tahmini sunan nonparametrik Bayesyen regresyon modelleri yardımıyla incelenmektedir. Böylece daha fazla düşük sosyoekonomik statüye sahip öğrencinin zorlukları yenmesi için neler yapılabileceği konusunda çıkarımlar yapılmaktadır.

Tez planına göre bu çalışma giriş bölümü ile birlikte beş bölümden oluşmaktadır. İkinci bölüm olan Bayesyen yaklaşım bölümü, Bayes teoremi ve simülasyon yöntemleri temel başlıkları çerçevesinde oluşturulmuştur. Bu bölümde ilk olarak Bayesyen yaklaşım ve klasik yaklaşım arasındaki farklılıklar ve üstünlükler vurgulanmıştır. Daha sonra Bayesyen yaklaşımın önemli bir özelliği olan Bayesyen ilk dağılımlar ele alınmış ve güncel Hibrit ilk dağılımlar açıklanmıştır. Yine bu bölümde Bayesyen yaklaşımın ilerlemesinde önemli bir yere sahip simülasyon yöntemleri; Metropolis-Hastings Algoritması, Gibbs Örnekleme ve Hibrit Monte Carlo algoritmaları anlatılmıştır. Üçüncü bölüm nonparametrik Bayesyen yaklaşım ve Dirichlet süreçler çerçevesinde oluşturulmuştur. Burada ilk olarak nonparametrik Bayesyen yaklaşımın farklılığı ve üstünlüğü vurgulanmıştır. Ardından Dirichlet süreçler açıklanmış ve Dirichlet karışım modelleri anlatılmıştır. Yine bu bölümde nonparametrik Bayesyen regresyon modelleri, bağımlı nonparametrik süreçler ve bağımlı Dirichlet süreç modeli açıklanmıştır. Daha sonra nonparametrik Bayesyen regresyon modelleri çerçevesinde ANOVA bağımlı Dirichlet süreç modeli ve lineer bağımlı Dirichlet süreç modeli anlatılmıştır. Bu bölümde son olarak nonparametrik Bayesyen regresyon modelleri için son dağılım algoritmaları sunulmuştur. Uygulama bölümü olan dördüncü bölümde ilk olarak akademik dayanıklılık kavramı ve ilgili literatüre yer verilmiştir. Ardından analizde ele alınan veri seti ve değişkenler tanıtılmıştır. Son olarak nonparametrik Bayesyen regresyon analizi gerçekleştirilmiş ve analize ilişkin bulgular sunulmuştur. Çalışma, elde edilen bulguların değerlendirmesi ile sonlandırılmıştır.

2.BAYESYEN YAKLAŞIM

Olasılığın Bayes felsefesinde temel bir inanç derecesi olarak yorumlanması, 1713'te J. Bernoulli de dâhil olmak üzere Ars Conjectandi'nin çalışmasıyla uzun bir tarihe sahiptir. Çıkarım problemlerine Bayesyen yaklaşım için ilk adım, Thomas Bayes'in ölümünden sonra 1763'te "*Şans Doktrinde Problem Çözme Yöntemi Üzerine Bir Deneme*" başlıklı makalesini yayınlayan Richard Price tarafından atılmıştır (Türkman vd. 2019: 5).

1990'a kadar Bayes yöntemlerinin istatistiksel analize uygulanması çok sınırlıdır, çünkü gerekli matematiksel hesaplamalar analitik olarak sadece çok basit modeller için uygulanabilmektedir. Yirminci yüzyılın ilk yarısında, güçlü bilgisayarların artan erişilebilirliği ile Bayes modelleri için yeni hesaplama algoritmalarının geliştirilmesi, karmaşık gerçek problemlerde Bayes yöntemlerinin kullanımına olanak sağlamıştır. Günümüzde BUGS, JAGS, WinBUGS ve R gibi Bayes analizine yardımcı olmak için büyüyen bir yazılım yelpazesinin olması Bayesyen yöntemlerin gelişmesinde önemli rol oynamaktadır.

Bayesyen çıkarımının istatistiksel modelleme ve veri analizinde birtakım avantajları vardır. Bayesyen yaklaşım, sağduyulu yorumlarla daha uyumlu olan hipotezlerde parametreler ve olasılık değerleri üzerinde güvenilir aralıklar sağlamaktadır. Son zamanlarda elde edilen bilgi sentezi kavramlarına uygun olarak inançları güncellemek için verilerden öğrenme sürecini resmileştirmenin bir yolunu sunmaktadır (Congdo 2006: 2).

Olasılık belirsizliğin bir ölçüsüdür. Belirsizlik tamamen ortadan kaldırılamadığı için modellenmesi gerekmektedir. Belirsizlik durumunda sadece objektif olasılıklar verilir, ancak subjektif olasılıklar belirlenemez. Bu nedenle, verilerin analizini geliştirmek için kişisel inançları kullanabilme fikrine dolaylı olarak dayanan subjektif Bayes yaklaşımı geliştirilmiştir (Goldstein, Wooff 2007: 4). Klasik yaklaşım yorumu bir tekrarlanabilir deney dizisinin mümkün olduğu durumlarla sınırlıdır. Belirsizliği ölçmek istediğimiz birçok olayın olasılığı için frekans yorumlanması mümkün değildir (Cowles 2013: 18).

Klasik ve Bayesyen yaklaşımda olasılık kavramları arasındaki farkın bu kadar önemli olmasının nedeni, parametrelerin farklı şekilde ele alınmasıdır. Parametreler bilinmeyen şeyleri temsil eder ve genellikle verilerin ortaya çıktığı anakütlenin özellikleri olarak düşünülebilir. Klasik yaklaşımda parametreler sabit olarak ele alınırken, Bayesci perspektiften parametreler rastgele bir değişken olarak kabul edilir ve bu nedenle bunlar hakkında olasılık ifadeleri kullanılır (O'Hagan 2004: 41). Diğer bir açıdan, klasik yaklaşım tahminciyi rastgele bir değişken olarak yorumlarken, Bayesyen yaklaşım parametreyi rastgele

değişken olarak yorumlamaktadır (Koop 2003: 20). Klasik yaklaşımda koşullandırma parametreler üzerineyken, çoğaltma (birleştirme) veri üzerindedir. Bayesyen yaklaşımda ise koşullandırma veri üzerine ve çoğaltma (birleştirme) parametreler üzerinedir (Carlin, Louis 2009: 5).

Klasik ve Bayesyen yöntemler arasındaki temel fark Bayesyen yaklaşımının parametreler hakkında bir ön bilgi kullanmasıdır. Bu ön bilgilerin sübjektif olması sebebiyle Bayesyen yaklaşım eleştirilmektedir. Yani, Bayesyen yöntemler olasılığın sübjektif yorumuna dayanmaktadır. İlke olarak aslında sübjektiflik Bayesyen yöntemin gücüdür, çünkü farklı gözlemciler (ön bilgiler) sayesinde elde edilebilecek son dağılımları incelememize olanak sağlamaktadır. Ayrıca Bayesyen yöntemler bu ön bilgileri içerdikleri için mevcut tüm bilgileri kullanmaktadır (O'Hagan 2004: 36).

Klasik yöntemler genellikle asimptotik sonuçlara dayanır ve bu durumlar büyük örnekler için geçerlidir. Bayesyen yöntemler özellikle küçük örneklerde klasik yöntemler kullanıldığında engel teşkil edebilen asimptotik özelliklere dayanmaz. Böylelikle Bayesyen yöntemler küçük örneklem boyutlarına sahip verileri modellemek için daha donanımlı olmaktadır. Küçük örneklem konusu oldukça önemlidir ve Bayesyen yöntemler sonlu örnek düzeltmeleri ve serbestlik derecesi ayarlamalarına duyulan gereksinimi ortadan kaldırmaktadır. Ayrıca parametrelerin örnekleme dağılımı ile ilgili varsayımlar yapılmasına da gerek yoktur, çünkü Markov Zinciri Monte Carlo (MCMC, Markov Chain Monte Carlo) yaklaşımı her parametrenin yeniden örnekleme algoritmaları yoluyla dağılımını sağlamaktadır (McNeish 2016: 750-754).

Klasik ve Bayesyen yöntemler, iki farklı mantıksal çıkarımdan oluşmaktadır; tümdengelim ve tümevarım yaklaşımı. James Bernoulli (1713), şans oyunları için geçerli olan tümdengelim mantığı ile günlük yaşam için gerekli olan tümevarımsal mantık arasındaki farkı algılayarak sorunu ilk dile getiren kişidir (Sivia, Skilling, 2006: 8). Klasik yaklaşımda tümdengelim mantık ile bir sebepten başlanarak sonuçlar elde edilmek istenmektedir ve bu akıl yürütme genelden özele doğru çalışmaktadır (Koch, 2007: 3). Ancak çoğu zaman, belirli etkiler gözlemlendikten sonra, altta yatan nedenler bulunmak istenmektedir. Bayesyen yöntemlerde tümevarımsal çıkarım, özelden genele doğru çalışmaktadır. Öncelikle birçok gözlem yapılır, bir örüntü fark edilir ve bir genelleme yapılır; yani genel ilkeleri belirli örneklerden çıkarmayı hedeflemektedir (Donovan, Mickey 2019: 39). Bayesyen yaklaşım, tümevarımsal mantık sayesinde gerçek hayattaki belirsizliklerin belirli bir inanç seviyesi ile modellenmesini sağlamaktadır.

2.1. Bayes Teoremi

İngiliz matematikçi Thomas Bayes'in adını taşıyan Bayes teoremi, koşullu olasılığı belirlemek için kullanılan matematiksel bir kuraldır. Bayes teoremi, ilk olasılıkları (ilk dağılımları), ilgili modeller için son olasılıkları (son dağılımları) hesaplama olasılığı ile birleştirmektedir (Cowles 2013: 21). Son dağılım Bayes teoremi kullanılarak ilk dağılım güncellenerek hesaplanmaktadır.

Bayes teoremi hem koşullu olasılıkların hesaplanmasını basitleştirdiği hem de olasılığın sübjektif yorumunun özelliklerini açıkladığı için önemlidir. Bayes teoreminin inşası için öncelikle marjinal olasılık, koşullu olasılık ve çarpma kuralı gibi kavramların açıklanması gerekmektedir.

A ve B, bir deneyinin tüm olası sonuçlarını içeren aynı örnek uzay U'nun iki alt kümesi ($A \subseteq U$ ve $B \subseteq U$) olarak temsil edilebilecek iki rastgele olay olsun. B olayının tümleyeni ($\bar{B}$), B'de olmayan sonuçların kümesidir. Ortak olay düzleminde bir A olayının olasılığına marjinal olasılık denir. Olasılık aksiyomları kullanılarak marjinal olasılık şu şekilde yazılabilir;

$$P(A) = P(A \cap B) + P(A \cap \bar{B}) \quad (2.1)$$

A olayı meydana geldiğinde B olayının koşullu olasılığı, A'nın zaten bilindiği durumda B'nin gerçekleşme şansını ölçer. $P(A) > 0$ olmak üzere, A ve B olayları için koşullu olasılık

$$P(B | A) = \frac{P(A \cap B)}{P(A)} \quad (2.2)$$

şeklinde tanımlanmaktadır. Burada iki A ve B olayının kesişimi ($A \cap B$), hem A hem de B'deki sonuçların kümesidir (Shemyakin, Kniazev 2017: 4). A olayı meydana geldiği için koşullu olasılıklar bölünmüş evrendeki ($U_r = A$), olasılıklar olup toplamı 1'e ölçeklendirilmiştir. Yani artık yalnızca A olayı içindeki sonuçları dikkate alınmaktadır. A ve B olayları bağımsız olduğunda, $P(A \cap B) = P(A).P(B)$ olup, A olayı hakkındaki bilgi B olayının meydana gelme olasılığını etkilemediği için koşullu olasılık $P(B|A) = P(B)$ şeklinde yazılmaktadır (Lynch 2007: 11).

A ve B olaylarının rolleri tersine çevrildiğinde, gözlemlenemeyen bir B olayı ile verilen gözlemlenebilir bir A olayının koşullu olasılık ilişkisi şu şekilde yazılmaktadır:

$$P(A | B) = \frac{P(A \cap B)}{P(B)} \quad (2.3)$$

Bu şekildeki koşullu olasılık ilişkisi bu formada kullanılmadığı için, ortak olasılık $P(A \cap B)$ 'yi bulmak için çarpma kuralı kullanılmaktadır. Çarpma kuralı, kesirden temizlenen koşullu olasılığı

tanımlanmaktadır ve bileşik olaylara olasılık atamak için kullanılır. Böylece çarpma kuralı kullanılarak ortak olasılık $P(A \cap B)$ şu şekilde ifade edilmektedir:

$$P(A \cap B) = P(B)P(A|B) \quad (2.4)$$

Bir A olayı meydana geldiği bilindiğinde, ($U_r = A$), A olayının marjinal olasılığı (2.1), koşullu olasılık tanımını (2.2)'de yerine yazıldığında;

$$P(B|A) = \frac{P(A \cap B)}{P(A \cap B) + P(A \cap \bar{B})} \quad (2.5)$$

eşitliği elde edilir. Burada ortak olasılıkların her birini bulmak için çarpma kuralı uygulanırsa, $P(A \cap B) = P(A|B)P(B)$ ve $P(A \cap \bar{B})P(\bar{B})$ olmak üzere;

$$P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\bar{B})P(\bar{B})} \quad (2.6)$$

Eşitlik (2.6) Bayes kuralını vermektedir (Cowles 2013: 19). Genellikle evreni bölen ikiden fazla bir dizi olay vardır. $B_1 \cup B_2 \dots \cup B_n = U$ olmak üzere $B_1, B_2, \dots B_n$ gibi n tane gözlemlenemeyen olay olduğunu varsayalım. Örnek alan (evreni) bölen B olayları ile A olayı olduğunda, toplam olasılık yasası gereği A olayının olasılığı $P(A) = \sum_{j=1}^n P(A \cap B_j)$ şeklinde $(A \cap B_1), \dots (A \cap B_n)$ ayrı parçaların olasılıkları toplamı olarak yazılmaktadır. Her bir ortak olasılık $P(A \cap B_j)$ için çarpma kuralı kullanılarak $P(A) = \sum_{j=1}^n P(B_j) \cdot P(A|B_j)$ elde edilir. Her bir $P(B_j|A)$ koşullu olasılığı $P(A)$ olasılığına bölünerek bulunur. Böylece genelleştirilmiş Bayes teoremi şu şekilde ifade edilir;

$$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{\sum_{j=1}^n P(A|B_j)P(B_j)} \quad (2.7)$$

Bu ifade koşullu olasılık tanımından gelmektedir. $P(B_i)$ 'ye B_i olayının ilk olasılığı denilmektedir. $P(B_i|A)$, B_i olayının son olasılığı olarak adlandırılmaktadır. Bayes evreninde, gözlemlenemeyen olaylar $B_1, B_2, \dots B_n$ evreni yatay bölerken, gözlemlenen A olayı dikey bölmektedir. $P(A)$ olasılığı $P(A \cap B_i)$ sütundan aşağı doğru toplanarak bulunur. Her bir $P(A \cap B_i)$, $P(B_i)$ ilk dağılım ile $P(A|B_i)$ olabilirliğin çarpımı olarak; $P(A \cap B_i) = P(B_i) \cdot P(A|B_i)$ bulunur. Dolayısıyla Bayes teoremi, yani $P(B_i|A)$, B_i olayının son olasılığı; ilk dağılım ile olabilirlik çarpımının ilk dağılım ve olabilirlik çarpımının toplamına bölünmesiyle bulunmaktadır (Bolstad, Curran 2017: 79). Bu nedenle son dağılım, ilk dağılım ve olabilirliğin çarpımı ile orantılıdır şeklinde ifade edilmektedir.

Bayes teoremi ile bir A olayı meydana geldiğinde, bir rastgele değişken gözlenecek ve gözlemlendiğimiz değerler göz önüne alındığında, gözlemlenemeyen B olayı için rastgele değişkenin koşullu olasılık dağılımı belirlenmiş olacaktır.

Bayes modelleri tipik olarak, bilinmeyen d boyutlu $\theta = (\theta_1, \theta_2, \dots, \theta_d)$ parametre vektöründeki çıkarımlarla ilgilidir. θ parametresi için Bayesyen çıkarım olasılık ifadeleri açısından ifade edilmektedir. Örneklem büyüklüğü n olan D (veri) gözlemlendiğinde, bilinmeyen θ hakkında olasılık ifadelerini tanımlamak için öncelikle θ ve D için ortak olasılık dağılımını sağlayan bir model ile başlanır. Burada ortak olasılık yoğunluk fonksiyonu sırasıyla, ilk olasılık dağılımı (prior), $P(\theta)$ ve örnekleme dağılımı (olabilirlik) $P(D|\theta)$ gibi iki yoğunluğun bir ürünü olarak şu şekilde yazılmaktadır:

$$P(\theta, D) = P(\theta)P(D|\theta) \quad (2.8)$$

Bayes teoremi ve koşullu olasılığın temel özellikleri kullanılarak D değeri gözlemlendiğinde θ için son dağılım (posterior) aşağıdaki gibi ifade edilebilir:

$$P(\theta|D) = \frac{P(\theta, D)}{P(D)} = \frac{P(\theta)P(D|\theta)}{P(D)} \quad (2.9)$$

burada θ sürekli olduğu durumda, $P(D) = \int P(\theta).P(D|\theta)d\theta$ olur. (2.9) eşitliğinden $p(D)$ marjinal olasılık faktörü, D sabit olduğundan ve θ parametresine bağımlı olmadığından çıkarılabilir. Böylece θ parametre vektörü için normalize edilmemiş son dağılım (2.10)'nun sağındaki ifadeden ilk dağılım ve olabilirliğin çapımı ile orantılı olacak şekilde bulunur (Gelman vd. 2004: 8).

$$P(\theta|D) \propto P(D|\theta)P(\theta) \quad (2.10)$$

Bayesci çıkarımda yer alan öğrenme süreci, verileri gözlemlmeden önce parametreler hakkındaki ilk dağılım ifadeleri, yani ön bilgiler ile elde edilen verilerden elde edilen bilgileri birleştiren güncellenmiş ya da son bilgiye göre değişmektedir (Congdon 2006: 3).

2.1.1. Olasılıklar Oranı ve Bayes Faktörü

Matematiksel olasılıklar, veriler tarafından iki hipoteze verilen karşılaştırmalı bir olasılık oranı biçiminde birbirleriyle karşılaştırıldığında anlam taşımaktadır. Olasılık oranı, Bayes faktörünün en basit şeklidir (Goodman 1999b: 1007).

Bir A olayının olasılık oranı (odds), olayın olasılığının tümleyeninin olasılığına bölünmesi ile bulunur:

$$odds(A) = \frac{P(A)}{P(\bar{A})} \quad (2.11)$$

Burada A olayının tümleyeninin olasılığı, yani meydana gelmeme olasılığı $P(\bar{A}) = (1 - P(A))$ olup, bu ifade denklem (2.11) de yerine yazılıp düzenlendiğinde bir olayın olasılığı ile olasılık oranı arasındaki bire bir ilişki şu şekilde ifade edilmektedir: (Bolstad, Curran 2017: 75)

$$P(A) = \frac{odds(A)}{(1 + odds(A))} \quad (2.12)$$

Bayes teoremi, biri verileri özetleyen diğeri inancı temsil eden iki bileşenden meydana gelmektedir. Bayes faktörü, verilerin konuştuğu indekstir ve denklemin sübjektif kısmından ayrıdır. Bayes faktörü, iki hipotezin verileri ne kadar iyi tahmin ettiğinin bir karşılaştırmasıdır (Goodman 1999b: 1005).

Bayesyen model karşılaştırması, Bayes faktörlerine dayalı bir model seçim yöntemidir (Goodman 1999a: 996). D (veri) gözlemlendiğinde, M_i, M_j alternatif modellerinin varsayımlarına karşılık gelen H_i ve H_j hipotezleri verildiğinde, H_i lehine (ve H_j 'ye karşı) Bayes faktörü son olasılığın ilk olasılığa oranı olarak bulunmaktadır: (Bernardo, Smith 2000: 389)

$$B_{ij}(D) = \frac{P(D|M_i)}{P(D|M_j)} = \left\{ \frac{P(M_i|D)}{P(M_j|D)} \right\} / \left\{ \frac{P(M_i)}{P(M_j)} \right\} \quad (2.13)$$

Kelimeler ile yukarıdaki eşitlik, Son olasılık oranı = Bayes faktörü x İlk olasılık oranı şeklinde ifade edilmektedir. Yani Bayes faktörü ilk olasılık oranını (prior odds) son olasılık oranına (posterior odds) dönüştürmektedir (Held Ott 2018: 3).

Bayes faktörü, D verilerinin H_i olasılığını nasıl artırdığına veya azalttığına dair doğrudan nicel bir ölçü sağlamaktadır. Böylece $B_{ij}(D) > 1$ olması D verisi verildiğinde H_i hipotezinin (M_i modelinin) nispeten daha makul olduğunu göstermektedir. Bu nedenle, Bayes faktörü (veya logaritması) genellikle kanıtın gücü veya kanıt ağırlığı olarak adlandırılmaktadır (Bernardo, Smith 2000: 390).

Aynı veriye ait birkaç farklı model düşünüldüğünde Bayes faktörleri (2.14)'de görüldüğü gibi bu modellerin son olasılığının hesaplanmasını sağlamaktadır. D verisi gözlemlendiğinde, (K+1) alternatif modelin varsayımlarına karşılık gelen $H_0, H_1, \dots, H_k$ hipotez olduğu kabul edilsin. Her bir $H_1, \dots, H_k$ sırasıyla H_0 ile karşılaştırılarak Bayes faktörleri $B_{10}, \dots, B_{K0}$ elde edilir. $B_{ij}(D) = P(D|H_i)/P(D|H_j)$ olmak üzere, H_k son olasılığı;

$$P(H_k | D) = \frac{\alpha_k B_{k0}}{\sum_{r=0}^K \alpha_r B_{r0}} \quad (2.14)$$

elde edilir. Burada $B_{00} = 1 = \alpha_0$ ve $\alpha_k = P(H_k)/P(H_0)$ olmak üzere H_0 ' a karşı $H_k (k = 0, \dots, K)$ için ilk olasılık oranıdır. Tüm ilk olasılık oranları $\alpha_k = 1$ ' e eşit alınabilir ancak farklı α_k değerleri rakip modellerin uygun olup olmadığı hakkında ilk bilgileri yansıtmak için kullanılabilir (Kass, Raftery 1995: 784).

En fazla kanıta sahip olan hipotez, maksimum matematiksel olasılığa sahip olan hipotezdir, bu da gözlemlenen verileri en iyi şekilde tahmin ettiği anlamına gelmektedir (Goodman 1999b: 1007). Aynı veri kümesine ait farklı modeller arasından seçim yapılırken en yüksek son olasılığa sahip model uygun model olarak seçilmektedir (Bernardo, Smith 2000: 389).

Bayes faktörü ve olasılık oranı arasındaki temel fark; olasılık oranı genellikle basit hipotezlerin karşılaştırmalarıyla sınırlıdır, ancak Bayes faktörü, bileşik hipotezler için kanıtların ölçülmesine izin veren Bayes teoreminin mekanizmasını kullanmaktadır (Goodman 1999b: 1009).

2.1.2. Bayesyen İlk Dağılımlar

Bayesyen yaklaşım, model parametreleri hakkındaki belirsizliği ölçmek için mevcut tüm bilgileri kullanmak için mantıksal olarak tutarlı bir çerçeve sağlamaktadır. $\theta = (\theta_1, \dots, \theta_p)$ parametre vektörü ve $y = (y_1, \dots, y_n)$ gözlemlenen değişken matrisi olmak üzere, Bayes teoremi şu şekilde ifade edilmektedir:

$$P(\theta | y) \propto L(\theta | y)\pi(\theta) \quad (2.15)$$

Burada,

$\pi(\theta)$: θ parametresini ilk dağılımı,

$L(\theta|y)$: θ 'nın benzerlik fonksiyonu,

$m(y) = \int L(\theta|y)\pi(\theta)d\theta$: y 'nin marjinal yoğunluğu

$P(\theta|y) = \frac{L(\theta|y)\pi(\theta)}{m(y)}$: y veri gözlemlendiğinde θ 'nın son dağılımını göstermektedir.

Son dağılım ilk dağılımlar ile orantılı olup, ön bilgilerin ve gözlemlenen verilerdeki yeni bilgilerin hesaplanmasından sonra parametreler hakkındaki belirsizliği nicelleştirmektedir (Reich, Ghosh 2019: 22).

Bayesyen çıkarım, ilk dağılımlar sayesinde ilk inançların olasılık olarak tanımlanmasına izin vermektedir. $\pi(\theta)$ ilk dağılımlar verilen veriye bağlı değildir, yani veriler görülmeden önce θ hakkındaki bildiklerimizi özetlemektedir (Koop 2003: 3-6).

İlk dağılımlar ile araştırmacı verileri görmeden önce sahip olduğu tüm bilgileri modele dâhil etmeyi amaçlamaktadır. Dolayısıyla, yorumlanması veya hesaplanması kolay belirli ilk dağılım seçmek yaygın olsa bile ilk dağılımlar herhangi bir şekilde olabilir (Koop 2003: 18).

İlk olasılık dağılımının belirlenmesi Bayesyen yaklaşımın hem en tartışmalı hem de en önemli yönlerinden biridir. Doğru veya en iyi ilk dağılım kavramı yoktur, bunun yerine ilk dağılım veri biriktikçe güncellenen Bayesyen istatistiksel çıkarım sürecinin başlangıcı olarak görülmektedir. Bayes modellerini belirtmek, bilinmeyen parametreler için zorunlu olarak bir ilk dağılımın seçilmesini gerektirmektedir ve bu seçim prensip olarak yinelemeli yöntemleri başlatmaya benzemektedir. Son dağılımının, ilk olasılık dağılımdaki genel koşullar altında gerçek değere yakınsaması beklenmektedir (Reich, Ghosh 2019: 41).

Bu tez çalışmasında, çeşitli modeller için eşlenik, bilgi vermeyen, bilgi veren ve hibrit ilk dağılımlar olmak üzere genel yaklaşımlar incelenmektedir.

2.1.2.1. Eşlenik İlk Dağılımlar

Bir olasılık yoğunluk fonksiyonu, benzerlik fonksiyonu ile çarpıldıktan sonra ilk olasılık fonksiyonu ile aynı dağılım ailesine ait olan bir son olasılık yoğunluk fonksiyonu elde ediliyorsa eşlenik ilk dağılım olarak adlandırılmaktadır (Koch 2007: 59). Bu ilk dağılımlar, benzerlik fonksiyonu ile aynı fonksiyonel forma sahip olduğu için ön bilgiler olabilirlik fonksiyonu bilgisi ile benzer şekilde yorumlanmaktadır (Koop 2003: 18).

Eşlenik ilk dağılımlar şu şekilde tanımlanmaktadır; $L(\theta) = P(y|\theta)$, y gözlemine dayalı bir benzerlik fonksiyonunu göstermek üzere; ilk dağılım $P(\theta)$ G 'de olduğunda, son dağılımı $P(\theta|y)$ tüm y 'ler için G 'nin içindeyse bu G sınıf dağılıma $L(\theta)$ 'ye göre eşleniktir denilmektedir. Uygulamada $L_y(\theta)$ benzerliğine ait daha küçük G dağılım kümeleri bulunmaya çalışılmaktadır (Held, Bové 2014: 180). Farklı olabilirlik fonksiyonlarına karşılık gelen eşlenik ilk dağılımların özeti Tablo 1 de verilmektedir.

Tablo 1: Farklı Olabilirlik Fonksiyonları için Eşlenik İlk Dağılımlar

Benzerlik Fonksiyonu dağılımı	Eşlenik Dağılımlar	İlk Son Olasılık Dağılımı
$y \theta \sim \text{Bin}(n, \theta)$	$\theta \sim \text{Be}(\alpha, \beta)$	$\theta y \sim \text{Be}(\alpha + y, \beta + n - y)$
$y \theta \sim \text{Geom}(\theta)$	$\theta \sim \text{Be}(\alpha, \beta)$	$\theta y \sim \text{Be}(\alpha + 1, \beta + y - 1)$
$y \lambda \sim \text{Pareto}(\lambda \text{ için})$	$\lambda \sim G(\alpha, \beta)$	$\lambda y \sim G(\alpha + y, \beta + e)$
$y \theta \sim \text{Üstel}(\theta)$	$\theta \sim G(\alpha, \beta)$	$\theta y \sim G(\alpha + 1, \beta + y)$
$y \mu \sim N(\mu, \sigma^2 \text{ bilinmiyor})$	$\mu \sim N(v, \tau^2)$	$\mu y \sim N(\frac{\kappa y}{\kappa + \delta}, (\kappa + \delta)^{-1})$
$y \sigma^2 \sim N(\mu \text{ bilinmiyor}, \sigma^2)$	$\sigma^2 \sim \text{TersG}(\alpha, \beta)$	$\sigma^2 y \sim \text{TersG}(\alpha + \frac{1}{2}, \beta + \frac{1}{2}(y - \mu)^2)$

Kaynak: Held, Bové, 2014, *Applied statistical inference: Likelihood and Bayes*, s.181

İlk dağılımlar eşlenik olduğunda ilk gözlem gözlendikten sonra ortaya çıkan son dağılım aynı dağılım ailesine ait olma tanımı gereği, bir sonraki gözlem için yeni ilk olasılık görevi görmektedir. Daha sonra ikinci gözlemi içeren yeni son olasılık dağılımı yine aynı eşlenik aileye ait olup bu işlemler son gözleme kadar devam etmektedir. Verilerin bu şekilde sıralı işlenmesinde değişen tek şey dağılım parametreleridir (Held, Bové 2014:181).

Eşlenik ilk dağılımlar, inançları ifade etmek için, iki dağılım arasındaki matematiksel bir ilişkiye atıfta bulunduğu için bu olasılıklar en uygun seçimler olarak görülmektedir. Bu ilk dağılımlar ön bilgiler mevcut olduğunda yararlıdır ve ön bilgi olmadığı zamanda kullanılabilir. Diğer taraftan eşlenik ilk dağılımlar benzersiz, tek değildir. Ayrıca, tüm benzerlik fonksiyonları bilinen bir eşlenik içermediği için eşlenik ilk dağılımlar bu açıdan sınırlıdır ve birçok eşlenik çifti daha az parametreye sahip küçük örnekler için geçerlidir. Bu sınırlamalar, zengin istatistiksel modeller oluşturmak için bir yaklaşım sağlayan hiyerarşik modelleme ve Gibbs örnekleme yoluyla azaltılmıştır (Reich, Ghosh 2019: 42).

Bayesyen analizlerde, MCMC simülasyon teknikleri ortaya çıkmadan önce, önerilen modellerin tahmin edilmesi eşlenik ilk dağılımlar olmadan çok zorlayıcı olmaktaydı. Yeni gelişen yöntemler sayesinde durum artık böyle değildir, ancak eşlenik ilk dağılımlar uygulamada halen yararlıdır ve mükemmel bir açıklama aracı olarak kullanılmaktadır (Gill 2015: 104).

2.1.2.2. Bilgi Vermeyen İlk Dağılımlar

Bayesyen yaklaşımın temel faydası, ön bilgileri analize dâhil etme yeteneğidir. Herhangi bir ön bilgi olmadığında, klasik yaklaşım yerine Bayesyen yaklaşım uygulamanın olası motivasyonu; parametrelerin rastgele olması varsayımına izin vermesidir. Bu motivasyon, özellikle çok sayıda parametreye sahip karmaşık modellerde, bazı parametreler için ön bilgi kullanılırken diğer parametreler için ön bilgi kullanılmaması durumunda özellikle önemli olmaktadır. Bilgi vermeyen ilk dağılımların belirli matematiksel ilkelere dayalı olarak yapılandırılmasına yönelik birçok yaklaşım vardır. Bu yaklaşımlar tarafından türetilen bilgi vermeyen ilk dağılımlara sahip Bayes yöntemleri, genellikle klasik istatistik yöntemleri ile tutarlı sonuçlar vermektedir (Shemyakin, Kniazev 2017: 61).

Bilinmeyen parametre vektörü hakkında herhangi bir bilgi mevcut değilse, bilgi vermeyen ilk olasılık yoğunluk fonksiyonları kullanılmaktadır. Yani, ilgili olay hakkında önceden hiçbir öznel bilginin kesinlikle bilinmediği durumdur. Bilinmeyen θ parametresi hakkında önceden hiçbir şey bilinmiyorsa, $\theta - \infty$ ile $+\infty$ arasında değerler alabilir. C bilgisinin koşulu altında bilgi vermeyen ilk dağılım $\pi(\theta|C)$ 'nin şu şekilde olduğu varsayılır;

$$\pi(\theta|C) \propto \text{sabit} \quad -\infty < \theta < \infty \quad (2.16)$$

Burada $\int_{-\infty}^{+\infty} \pi(\theta|C)d\theta = 1$ şartı yerine getirilmediği için, yoğunluk fonksiyonunu normalize edilemez ve bu nedenle bu ilk dağılımlar uygunsuz (improper) bir yoğunluk fonksiyonudur. Burada normalize edilecek bir benzerlik fonksiyonu seçilirse son olasılık dağılımının normalleşmesi sağlanabilir. Buradan, Jeffreys (1961) dönüşüm değişmezliğine dayanarak özel durumlar olarak (2.16) eşitliğindeki yoğunluk fonksiyonlarını içeren bilgi vermeyen ilk dağılımlar elde etmek için genel bir formül türetilmiştir (Koch 2007: 56-57).

2.1.2.2.1. Uniform İlk Dağılım

Bayesyen analizlerde ön bilgilerin kullanımındaki en önemli nokta, ilk dağılımın kişisel ön yargı veya bilimde sübjektiflik olarak eleştirilen faktörleri değil gerçek bilgileri yansıtmayı sağlamaktır. Bu sübjektiflik eleştirisine yanıt olarak, ilk bilginin olmadığı bir durumu temsil ettiği varsayılan geleneksel çözümler önerilmiştir. θ parametresini 0 ile 1 aralığında olması gereken bir oran olduğu varsayıldığında, θ hakkındaki bilgisizliği ifade etmenin bir yolu; tüm olası θ değerlerine eşit bir ağırlık veren uniform (tek düze) ilk yoğunluk seçmektir (O'Hagan 2004: 37-38).

Önceden θ oranının ne olduğu hakkında bir fikir yoksa herhangi bir değeri diğerinden daha fazla tercih etmeyen uniform ilk dağılım şu şekilde ifade edilir;

$$\pi(\theta) = 1 \quad 0 \leq \theta \leq 1 \quad (2.17)$$

Bu durumda, ilk dağılımlar θ 'nın tüm değerleri için sabit olduğundan son dağılımın şekli etkilenmez ve son dağılım benzerlik fonksiyonu ile orantılı olur (Bolstad, Curran 2017: 151). Parametre alanı sınırlı değilse, uniform ilk dağılımın seçilmesi uygun değildir (Held, Bové 2014: 185).

2.1.2.2.2. Jeffrey İlk Dağılımı

Bazı durumlarda, ön bilgi eksikliğinden dolayı parametre hakkında fazla bilgi vermeyen bir ilk dağılım seçmek faydalı olabilir. Jeffrey ilk dağılımı, yeniden parametreleştirmeye göre değişmeyen bir ilk dağılım oluşturan yöntemdir. Bu, bir yöntemin objektif olması için gerekli bir koşuldur. Farklı parametrelendirmeye göre farklı sonuçlar alınır, yani parametreleştirme seçimine bağlı ise yöntem sübjektiftir. θ için tek değişkenli Jeffrey ilk dağılımı;

$$\pi(\theta) \propto \sqrt{I(\theta)} \quad (2.18)$$

olarak tanımlanır ve burada $I(\theta)$, θ 'nın beklenen Fisher bilgisidir. Beklenen Fisher bilgisi,

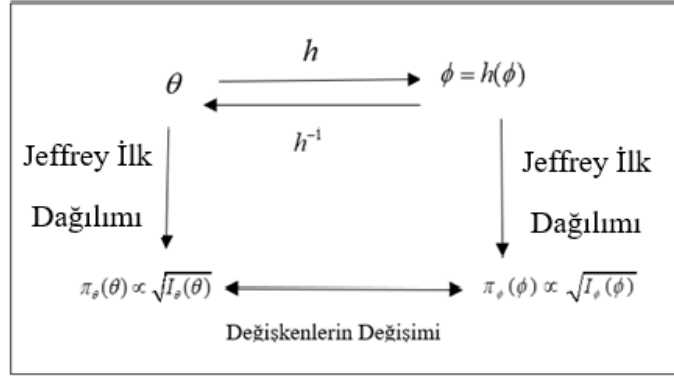
$$I(\theta) = -E\left(\frac{d^2 \log L(\theta | y)}{d\theta^2}\right) \quad (2.19)$$

şeklinde ifade edilir ve beklenti $y|\theta$ ile ilgilidir. Bu yapının temel özelliği parametre dönüşümüne göre değişmemesi ve $I(\theta)$ 'daki beklenen değerlerin veriye olan bağımlılığı ortadan kaldırması sonucunda ilk dağılımların veriye (y) bağlı olmamasıdır. Olabilirlik fonksiyonu $L(\theta|y)$ belirlendikten sonra Jeffrey ilk dağılımı belirlenir, bu anlamda aynı benzerlik fonksiyonu ile çalışıldığında aynı ilk dağılımlar ve aynı son dağılıma ulaşacağı için Jeffrey ilk dağılımı objektiftir (Reich, Ghosh 2019: 59-60).

Jeffrey ilk dağılımı, θ değerlerinin birebir dönüşümü altında değişmezdir. Eğer, $\pi_\theta(\theta) \propto \sqrt{I_\theta(\theta)}$ ise, ve θ bire bir dönüşümle yeniden parametrelendirilirse, (h 'ın tersi alınabilir olduğu varsayımı altında) $\phi = h(\theta)$ 'nın yoğunluk fonksiyonu $\pi_\phi(\phi) \propto \sqrt{I_\phi(\phi)}$ olur. Burada, beklenen Fisher bilgisi $I_\theta(\theta)$, bilinmeyen θ parametresi ile ilgili olarak verideki (ortalama) bilginin bir ölçüsüdür ve $I_\phi(\phi)$, ϕ 'nın beklenen Fisher bilgisini göstermektedir. Şekil 1'de θ 'dan $I_\phi(\phi)$, 'ye iki yol kullanılarak geçilmektedir,

ancak her iki yol da aynı sonucu ($\pi_\phi(\phi) \propto \sqrt{I_\phi(\phi)}$) vermektedir. Burada, Jeffrey ilk dağılımının değişmezlik özelliği açık bir şekilde görülmektedir (Held, Bové 2014: 86).

Şekil 1: Jeffrey İlk Dağılımının Değişmezliği



Kaynak: Held, Bové, 2014, *Applied statistical inference: Likelihood and Bayes*, s.187

Jeffrey ilk dağılımı tek parametrelili problemlere önemli ölçüde makul çözümler sağlarken, çok parametrelili karmaşık problemler için uygun olmadığı belirtilmektedir (Koop vd. 2007: 81).

2.1.2.2.3. Referans İlk Dağılımlar

Bernardo (1979), tarafından önerilen referans ilk dağılımlar, belirli bir probleme uygunlukları düşünülmeden otomatik seçimler olarak kullanılmak için tasarlanmıştır. Referans ilk dağılımlar, örnek boyut sonsuza yaklaştıkça, Kullback-Leibler mesafesi ile hesaplanan ilk dağılım ve son dağılım arasındaki farkı, eksik bilgiyi, maksimize eden ilk dağılımlardır (Koop vd. 2007: 82).

Bu tanım, ilk dağılım $\pi(\theta)$ ve son dağılım $P(\theta|y)$ gibi iki yoğunluk fonksiyonu arasındaki sapmanın hesaplanmasını gerektirmektedir. Bu hesaplama için Kullback-Leibler (KL) sapması kullanılmaktadır:

$$KL(p, \pi | y) = E \{ \log [f(\theta | y)] - \log [\pi(\theta)] \} \quad (2.20)$$

Kullback-Leibler sapması verilere bağlı olduğundan ilk dağılım türetmek için doğrudan kullanılmamaktadır. Referans ilk dağılımları, beklenen bilgi kazancını maksimize eden $\int KL(p, \pi | y) m(y) dy$, π ilk yoğunluk fonksiyonudur, burada $m(y)$, verilerin marjinal dağılımıdır (Reich, Ghosh 2019: 62).

Referans ilk dağılımlar, karmaşık bir parametre vektörünü ilgili parametre ve bozucu (nuisance) parametre olarak ikiye bölünebilmesi fikrine dayanır ve böylelikle bozucu parametre problemi ile başa çıkarak daha yüksek dereceli problemlerin sistematik olarak çözülmesini sağlamaktadır (Gill 2015: 112).

2.1.2.3. Bilgi Veren İlk Dağılımlar

Bilinmeyen parametre vektörü θ için bilgi vermeyen ilk dağılımların aksine, burada θ parametresi ile ilgili ön bilgilerin mevcut olduğu varsayılmaktadır (Koch 2007: 104). Bilgi veren ilk dağılımlar, araştırmacıların ellerindeki ön bilgileri analize dâhil etmesine izin veren ilk olasılıklardır. Ön bilgilerin istatistiksel çıkarımda rol oynaması gerekliliği bu yaklaşımın gerçekçi ve makul olduğunu göstermektedir (Gill 2015: 115).

Bilinmeyen θ parametresinin muhtemel yoğunluğu ile ilgili kanıt sağlayan önceki çalışmalar bulunuyorsa, bunlar mevcut çalışma için ilk dağılımların oluşturulmasında bir ön analiz şeklinde kullanılabilir (Congdon 2020: 27).

İbrahim ve Chen (2000) daha önce yapılmış çalışmalardaki verileri kullanan bilgi veren ilk dağılımları önermektedir. Mevcut çalışma verileri, n örneklem boyutu ve y gözlemlenen veri olmak üzere $D = (n, y)$ ile ifade edilsin, ayrıca θ parametre vektörü olmak üzere, mevcut çalışmanın benzerlik fonksiyonu $L(\theta|D)$ ile gösterilsin. Diğer taraftan $D_0 = (n_0, y_0)$ ile gösterilen önceki çalışmadan elde edilen eski veri olduğu varsayılın. θ için ilk önceki dağılım olarak adlandırılan $\pi_0(\theta|.)$, D_0 eski veri gözlemlenmeden önce eski ilk dağılımları gösterebilir. α_0 verildiğinde, mevcut çalışma için güçlü ilk dağılımlar şu şekilde tanımlanmaktadır:

$$\pi(\theta | D_0, \alpha_0) \propto L(\theta | D_0)^{\alpha_0} \pi_0(\theta | c_0) \quad (2.21)$$

Burada c_0 , ilk önceki dağılım için bir hiperparametreyi gösterir ve bu parametre $\pi_0(\theta|c_0)$ 'ın tüm ilk dağılım üzerindeki etkisini kontrol etmektedir. Genellikle $0 \leq \alpha_0 \leq 1$ arasında sınırlanan α_0 mevcut çalışmanın benzerlik fonksiyonuna göre eski verileri ağırlıklandıran skaler bir parametreyi göstermektedir. α_0 parametresi $\pi(\theta|D_0, \alpha_0)$ benzerliği üzerindeki geçmiş verilerin etkisini kontrol ederek, araştırmacının geçmiş verilerin mevcut çalışma üzerindeki etkisini kontrol etmesine izin vermektedir. Böyle bir kontrol, önceki ve mevcut çalışma arasında heterojenlik olduğunda veya iki çalışmanın örneklem büyüklükleri arasındaki fark çok büyük olduğunda önem arz etmektedir. α_0 'ın

temel görevlerinden biri de θ parametresi için ilk dağılımının kuyruklarının ağırlığını kontrol etmektir ve α_0 küçüldükçe (2.21) ilk dağılımının kuyrukları ağırlaşmaktadır (İbrahim, Chen 2000: 47-48).

2.1.2.4. Hibrit İlk Dağılımlar

Bazı ilk dağılım tanımları, bilgi veren ve bilgi vermeyen ilk dağılımları karıştırarak bilgi geçişlerinin yansıtılabilen ilk dağılımlar şeklinde oluşturulmuştur. Bu ilk dağılımlar, bilgi veren ilk dağılım ve bilgi vermeyen ilk dağılımlar karıştırılarak inşa edildiği için hibrit olarak adlandırılmaktadır. Burada önemli nokta, ilk dağılımlarının çok esnek özelliklere sahip olması ve böylece bu ilk dağılımlar farklı düzeylerde nicel ya da nitel bilgileri yansıtacak şekilde uyarlanabilir olmasıdır (Gill 2015:129). Bu bölümde, ilk dağılımlar için önceki genel yaklaşımlar yerine belirli model problemlerine çözüm sağlayan bazı yeni ilk dağılım yapıları açıklanmaktadır.

2.1.2.4.1. Maksimum Entropi İlk Dağılımlar

Entropi, bilgi teorisiyle doğrudan bir ilişkiye sahiptir ve olasılık dağılımının doğasında bulunan belirsizlik miktarını ölçmektedir (Berger 1980: 75). Maksimum entropi yaklaşımı, bilinmeyen parametreler hakkında ilk olasılık yoğunluk fonksiyonları türetmek için kullanılmaktadır. Bilinmeyen bir olay hakkında bilgi elde etmek için deney yapılarak, deney gerçekleşmeden önce var olan belirsizlik ortadan kaldırılmaktadır. Deney ile ortadan kaldırılan belirsizlik aslında deneyden elde edilen ön bilgilere karşılık gelmektedir (Koch 2007: 58).

İlk olasılık dağılımının momentler ya da benzer nicelikler gibi bazı özelliklerinin bilinmesi durumunda, ilk dağılımların beklenen değer şeklinde yazılabildiği varsayımı altında;

$$E^{\pi} [g_k(\theta)] = \omega_k \quad (2.22)$$

($k = 1, \dots, K$) olmak üzere, bu kısıtlamaları karşılayan bir ilk dağılım π 'yi seçmenin bir yolu Jaynes (1983) tarafından geliştirilen maksimum entropi yöntemidir. Bilgi teorisindeki belirsizliğin bir ölçüsü olarak entropi şu şekilde tanımlanmaktadır:

$$\varepsilon(\pi) = -\sum_i \pi(\theta_i) \log(\pi(\theta_i)) \quad (2.23)$$

Burada entropiyi maksimize eden π ilk dağılımları, π aracılığı ile θ hakkında elde edilen ön bilgileri en aza indirmektedir. (2.22)'deki moment kısıtlaması altında maksimum entropi dağılımı, yoğunluk ile ilişkilidir:

$$\pi^*(\theta_i) = \frac{\exp\left\{\sum_1^K \lambda_k g_k(\theta_i)\right\}}{\sum_j \exp\left\{\sum_1^K \lambda_k g_k(\theta_j)\right\}} \quad (2.24)$$

λ_k sayıları Lagrange çarpanları olarak (2.22)'den türetilmektedir. π ilk dağılımlar üzerinde bir sınırlama yoksa maksimum entropi dağılımı sonlu parametre uzayı $\Theta = (\theta_1, \dots, \theta_k)$ üzerindeki uniform dağılım olmaktadır (Robert 2007: 110).

Maksimum entropi ilk dağılımlarının avantajı, kısıtlamalar altında sınırlı bilgilerin elde edilebilmesi ve ilk dağılımların belirli parametrik şekli için, mümkün olan maksimum entropi ile ilk dağılımlar üretebilmesidir. Burada amaç, bildiğimiz şeyi yansıtmaktır, ancak bunu yaparken ilk dağılımların üretilmesine mümkün olduğunca az müdahale edilmelidir. Kısıtlamalar göz önüne alındığında, mümkün olduğunca dağınık bir ilk olasılık dağılımı üretilmektedir, yani belirlenen $\pi(\theta)$, θ hakkında minimum bilgi veren ilk dağılım olmaktadır (Gill 2015: 132).

2.1.2.4.2. Histogram İlk Dağılımlar

Johnson ve Albert (1999), bir ilk dağılım değerlendirilmesi için parametre hakkındaki ön bilgileri daha iyi temsil eden bir olasılık dağılımının kaba bir şeklini oluşturma fikrine dayanan histogram ilk dağılımını önermektedir. Histogram ilk dağılım yaklaşımı parametrik varsayımlar gerektirmediği ve ön bilgileri belirlemede esneklik sağladığı için çekici bir yöntem olarak görülmektedir (Johnson, Albert 1999: 13).

Parametre uzayı Θ olmak üzere, θ için histogram ilk dağılımlar, genellikle Θ 'nin kümeler halinde aralıklara bölünmesi ve her alt kümenin olasılığının belirlenmesi ile elde edilmektedir (Robert 2007: 109). Histogram ilk dağılımı, $\pi(\theta)$ belirlendikten sonra düzgünleştirilmiş bir yoğunluğa dönüştürülebilmektedir (Congdon 2020: 27).

2.2. Simülasyon Yöntemleri

İstatistiksel çıkarımda özellikle yüksek boyutlu karmaşık modeller söz konusu olduğunda, sayısal hesaplama problemlerinden dolayı, bazı durumlarda belirli bir paradigma ile ilişkili tahminleri doğrudan analitik olarak hesaplamak mümkün olmamaktadır.

Matematiksel hesaplamaların mümkün olmadığı; integral alınacak fonksiyon için önemli bölgenin tespit edilememesi, karmaşık modellerde çok boyutlu integrallerin hesaplanmaması gibi durumlarda alternatif bir yaklaşım olarak stokastik simülasyon yöntemleri kullanılmaktadır (Robert, Casella 2008: 62). Bayesyen yaklaşımda stokastik simülasyon genellikle dağılımın momentlerini ve olasılıklarını tahmin etme sorunlarına yardımcı olan bir araç olarak kullanılmaktadır (Shemyakin, Kniazev 2017: 86).

Bayesyen hesaplama, örnekleme dağılımını tahmin eden son olasılık dağılımını tahmin etmek için tasarlanmıştır (Reich, Ghosh 2019: 69). Belirli bir model için eşlenik ilk dağılımlar kabul edildiğinde, son dağılım da ilk dağılım ile aynı aileden olduğu için bilinen yöntemler ile örneklenebilmektedir. Ancak günümüzde eşlenik ilk dağılımlar çok tercih edilmediği için, son dağılımın örneklenmesi daha karmaşık ve zorlayıcı hale gelmektedir (Press 2003: 119).

Hesaplama problemleri önemlidir, çünkü model karşılaştırması ve tahmini ile ilgili denklemler doğrudan veya dolaylı olarak integraller içermektedir (Koop 2003: 7). Bayes teoremine bakıldığında, son dağılımının belirlenmesi genellikle yüksek boyutlu integrallerin hesaplanmasına yani (2.25) eşitliğinin paydasına bağlı olduğu görülmektedir. Ayrıca son dağılımın özelliklerinin (momentler, nicelikler gibi) özetlenmesi daha fazla integral hesaplaması gerektirmektedir.

$$P(\theta | y) = \frac{L(\theta | y)\pi(\theta)}{\int L(\theta | y)\pi(\theta)d\theta} \quad (2.25)$$

Böyle yüksek boyutlu integralleri değerlendirmek için çok yüksek boyutlu modellere (on parametreden fazla) bile bilgi sağlayan ve programlanması nispeten daha kolay olan Monte Carlo entegrasyon yöntemleri geliştirilmiştir (Carlin, Louis 2000: 133).

Bölümün devamında, öncelikle doğrudan simülasyonun mümkün olduğu durumlar, ardından son dağılımından doğrudan simülasyonunun mümkün olmadığı durumları ve MCMC yöntemleri incelenmektedir. Bu bölümde ele alınan simülasyon yöntemlerinin çoğu son dağılımının özelliklerini özetleyerek, bilinmeyen parametreler hakkındaki belirsizliğini tam olarak ölçülmesini sağlamaktadır.

2.2.1. Monte Carlo Entegrasyonu

Simülasyon çalışması, istatistiksel çıkarım problemlerinin çoğunda, model özelliklerinin analitik olarak elle hesaplanması imkânsız olan bir son dağılıma yol açması durumunda, aynı dağılım özelliklerine sahip bir dizi simüle edilmiş değer üretilmesini sağlamaktadır. Böylece, zor bir hesaplama problemi yerine bu simüle edilen değerler kullanılarak son dağılım tanımlanmaktadır (Gill 2015: 275).

Monte Carlo yöntemleri, ilgili çıkarımı, olasılık dağılımlarından simüle edilmiş rasgele örnekleri içeren hesaplamalara dayanarak değerlendirmektedir. Monte Carlo yöntemi şu şekilde çalışır:

$$\int g(\theta)p(\theta|y)d\theta = E[g(\theta)|y] \quad (2.26)$$

θ parametre vektörü ve y gözlemlenen veri olmak üzere $p(\theta|y)$ ' e göre beklenen değer var olduğu varsayımı altında, (2.26) formundaki bir integrale yaklaşma problemi düşünüldüğünde, birçok son dağılımın özeti integrali alınabilen $g(\theta)$ fonksiyonları için (2.26) şeklinde ifade edilebilmektedir. Monte Carlo yönteminin (2.26)' daki integrale yaklaştırma ilkesi; son dağılım $p(\theta|y)$ 'den rasgele bir örnek $\theta^{(1)}, \dots, \theta^{(n)}$ oluşturmak ve bir yaklaşım olarak $g(\theta)$ fonksiyonunun örnek ortalamasını önermektir:

$$\hat{E}[g(\theta)|y] = \frac{1}{n} \sum_{i=1}^n g(\theta^{(i)}) \quad (2.27)$$

Burada, ortalama $\hat{E}[g(\theta)|y]$, büyük sayılar yasası ($n \rightarrow \infty$) ile neredeyse kesin olarak (yani üretilen her dizi için) $E[g(\theta)|y]$ 'e yakınsamaktadır. Tahmin edicinin kesinliği, Monte Carlo ortalamasının standart hatasıyla değerlendirilebilmektedir:

$$\frac{1}{\sqrt{n(n-1)}} \left\{ \sum_{i=1}^n \left[g(\theta_i) - \frac{1}{n} \sum_{i=1}^n g(\theta_i) \right]^2 \right\}^{1/2} \quad (2.28)$$

Burada örnek büyüklüğü araştırmacı tarafından belirlenmektedir ve yaklaşımın kalitesi örnek sayısı artıkça artmaktadır (Türkman vd. 2019: 43-44).

Monte Carlo yöntemi, son dağılımdan çekilen örneklere $\theta^{(i)} \sim p(\theta|y)$ dayalı olarak bilinmeyen parametre θ hakkında ortalama ve momentler gibi bazı nicelikleri tahmin etmemizi sağlamaktadır. Bu nicelikler şu şekilde ifade edilmektedir (Ando 2010: 64):

Tablo 2: Son Dağılım Özellikleri

Son Dağılım Özellikleri	Monte Carlo Yaklaşımı	θ Tahmini
Son dağılım ortalaması	$\frac{1}{n} \sum_{i=1}^n \theta^{(i)}$	$\bar{\theta}_n = \int \theta p(\theta y) d\theta$
Son dağılım mod	$\arg \max_i p(\theta^{(i)} y)$	$\hat{\theta}_n = \arg \max_{\theta} p(\theta y)^1$
Belirli bir Q bölgesinde son dağılım olasılığı	$\frac{1}{n} \sum_{i=1}^n I(\theta^{(i)} \in Q)^2$	$p(\theta \in Q y)$
Marjinal son olasılık dağılımı	$\frac{1}{n} \sum_{i=1}^n p(\theta_1, \theta_2^{(i)} y)$	$p(\theta_1 y) = \int p(\theta_1, \theta_2 y) d\theta_2$

Monte Carlo yaklaşımı yalnızca son dağılımdan analitik olarak bilgi elde etmek mümkün olmadığında kullanılmaktadır. Monte Carlo yönteminin kesinliği, verideki gözlem sayısına bakılmaksızın, üretilen örneklerin boyutu tarafından kontrol edilmektedir ve yaklaşım hataları olasılıksal olarak ölçülebilmektedir (Gamerman, Lopes 2006: 98).

2.2.2. Son Dağılım Simülasyonları için Algoritmalar

Son dağılım simülasyon algoritmaları için iki temel yaklaşım vardır. Birinci yaklaşım, son dağılıma benzer ve bağımsız (i.i.d.) çizimlerle örnek oluşturabilen algoritmalar. Rasgele örnek olarak $\theta^{(m)}$ vektörünü doğrudan $p(\theta|I)$ yoğunluğundan (i.i.d.) örnek almak için bir yöntem türetilmezken, $p(\theta|I)$ ya benzer bir $p(\theta|S)$ yoğunluğundan (i.i.d.) örnekler simüle edilebilmektedir. Böylece ilgili ω parametresi ile ilgili birçok bilgi elde edilmektedir. Birinci yaklaşım için iki yöntem vardır: Red örnekleme ve Önem örnekleme. Burada sadece son dağılım doğrudan örneklemediği durumlarda kullanılan algoritmalar dâhil edilmektedir. İkinci yaklaşım, MCMC yöntemleri olarak adlandırılan, son dağılımdan bağımsız olmayan ve aynı şekilde dağılmamış ancak son dağılıma yakınsayan Markov zinciri oluşturulmasına dayalı algoritmalar içermektedir (Geweke 2005: 110).

¹ argmax, fonksiyon değerlerinin maksimize eden noktalarıdır.

² Burada I gösterge fonksiyonudur.

2.2.2.1. Red Örneklemesi

Son dağılım simülasyonunun ilk geliştirilen algoritmalarından biri olan kabul-red örneklemesi genellikle red örneklemesi olarak bilinmektedir ve matematikçi John Von Neuman (1951) tarafından geliştirilmiştir. Bu algoritma, rastgele örneklerin (i.i.d.) çekilebildiği bir kaynak yoğunluk (aday yoğunluk) $h(\theta)$ verildiğinde, bu yoğunluktan rastgele çekimler elde edilmesi ve son dağılıma ait olmayanların atılması fikrine dayanmaktadır.

K, 1 den büyük bir sabit olmak üzere $h(\theta)$ fonksiyonunun mevcut olduğu varsayımı altında;

$$p(\theta | y) \leq K h(\theta) \quad (2.29)$$

olsun. Burada K sabitini görevi θ 'nın tüm değerleri için eşitsizliğin sağlanmasını sağlamaktır. θ 'nın son dağılımı $p(\theta|y)$ den örnek üretmek için red örneklem algoritması prosedürü aşağıdaki adımlardan oluşmaktadır:

- 1- $h(\theta)$ 'dan θ örnekleme çekilsin ve bu çekim θ^* ile gösterilsin.
- 2- Kabul olasılığı için $\alpha = \frac{p(\theta^*|y)}{Kh(\theta^*)}$ oranı hesaplanır.
- 3- α , olasılığıyla, $p(\theta|y)$ son dağılımdan bir çekilen bir örnek olarak θ^* kabul edilir.

θ^* reddedilirse, (1) adımına geri dönülür ve red algoritmasının adımları birçok kez tekrarlanarak, son dağılımdan bir örnek üretilir. Reddedip etmemeye karar vermek için, $u \sim U[0,1]$ tek düze dağılımdan bir u gözlemi çekilir, eğer ($u \leq \alpha$) ise θ^* örneği kabul edilir ($u > \alpha$) ise θ^* reddedilir (Rachev vd. 2008: 64).

2.2.2.2. Önem Örneklemesi

Önem örneklemesi red örneklemesinde olduğu gibi kaynak yoğunluğundan çekimlerin sadece bir kısmını kabul etmek yerine, bu çekimlerin hepsini kullanarak çekimleri uygun şekilde ağırlıklandırarak yaklaşık $E[h(\omega)|I]$ elde edilmesini sağlamaktadır. $p(\omega|y)$ son dağılımı ile yakından ilişkili bir dizi $\omega^{(m)}$ simülasyonu için önem örneklem algoritması kullanılarak ω 'nin son dağılımı ile ilgili bilgiler elde edilmektedir.

Önem örneklemede, herhangi bir $c_s > 0$ için kaynak yoğunluğu $p(\theta|S)$ ve kaynak yoğunluğunun isteğe bağlı kerneli $k(\theta|S) = c_s p(\theta|S)$ ile gösterilsin. Herhangi bir $c_I > 0$ için ilgili yoğunluğun isteğe bağlı kerneli $k(\theta|I) = c_I p(\theta|I)$ ile ifade edilsin. Önem örnekleme tarafından momentlerin yaklaşıklarının hesaplanması şu şekilde gerçekleşir: $\theta^{(m)}, \omega^{(m)}$ çekimlerinin sırasıyla $\theta^{(m)} \sim p(\theta|S)$ ve $\omega^{(m)} \sim p(\omega|\theta^{(m)}, I)$ ile bağımsız ve aynı şekilde dağıldığı varsayılın. $h: \Omega \rightarrow \mathbb{R}^1$ olmak üzere, ağırlıklandırma fonksiyonu $\omega(\theta) = k(\theta|I)/k(\theta|S)$ şeklinde tanımlansın. Aşağıdaki birkaç ek koşul altında;

- 1- $E[h(\omega)|I] = \bar{h}$ mevcut,
- 2- $var[h(\omega)|I] = \sigma^2$ mevcut,
- 3- $p(\theta|S)$ yoğunluğu Θ içerir,
- 4- $\omega(\theta)$, Θ ile üstten sınırlıdır,

yaklaşık momentler bulunabilir. Buradan, çekilen örneklem büyüklüğü M 'in yeterince büyük olduğu varsayılırsa verilen (1) ve (3) koşulları altında

$$\bar{h}^{(M)} = \frac{\sum_{m=1}^M \omega(\theta^{(m)}) h(\omega^{(m)})}{\sum_{m=1}^M \omega(\theta^{(m)})} \xrightarrow{\text{yaklaşır}} \bar{h}, \quad (2.30)$$

olarak hesaplanır. Ayrıca verilen (1) ve (4) koşulları altında;

$$M^{1/2} (\bar{h}^{(M)} - \bar{h}) \rightarrow N(0, \tau^2),$$

$$\hat{\tau}^{2(M)} = \frac{M \sum_{m=1}^M [h(\omega^{(m)}) - \bar{h}^{(M)}]^2 \omega(\theta^{(m)})^2}{\left[\sum_{m=1}^M \omega(\theta^{(m)}) \right]^2} \xrightarrow{\text{yaklaşır}} \tau^2, \quad (2.31)$$

olarak hesaplanmaktadır. Red örneklemesine göre, önem örneklemesinin avantajı yaklaşık momentlere uygulanmasının daha kolay olmasıdır (Geweke 2005: 116).

Son dağılım simülasyonları için önem örnekleme algoritması günümüzde çok tercih edilmemesine rağmen, daha karmaşık modeller için geliştirilen algoritmalara temeldeki fikirler açısından birçok uygulamada yararlı olmaya devam etmektedir (Türkman vd. 2019: 50).

2.2.3. Markov Zinciri Monte Carlo Yöntemleri

Klasik Monte Carlo (MC) yöntemleri ile son dağılımın karmaşıklığına bağlı olarak, hedef dağılımdan (i.i.d.) örnekler oluşturularak ya da inşası hedef dağılımı içeren bazı özel örnekleme dağılımları kullanılarak $E[g(\theta)|y]$ gibi son dağılımın özetlerinin değerlendirilmesi gerçekleştirilmektedir. Özellikle 1990'lardan beri, daha karmaşık ve yüksek boyutlu son dağılımlar içeren analizler için, hedef dağılıma eşit bir ergodik dağılıma sahip olacak şekilde inşa edilmiş homojen bir Markov zincirinin simülasyonuna dayanan MCMC yöntemi kullanmak daha yaygın hale gelmiştir (Türkman vd. 2019: 90).

Bu yöntemdeki matematiksel araç, adını icat eden Rus matematikçi Andrey Markov'dan (1856–1922) alan, uzayda rasgele hareket eden bir parçacık, belirli bir tür stokastik süreç olan Markov zinciridir. Bayesyen istatistik açısından, rasgele hareket eden parçacığın yörüngesi bir Monte Carlo algoritmasının çıktısıdır ve buradaki uzay bir son dağılım $p(\theta|y)$ desteğini sağlayan bir parametre uzayını ifade etmektedir. Burada önemli olan şey parametre uzayları üzerine Markov zincirleri oluşturmanın mümkün olmasıdır (Jackman 2009: 172).

Stokastik süreç, $\{\theta^{(t)}: t \in T\}$ ile gösterilen durum uzayında (Θ) tanımlanan ardışık rastgele niceliklerin kümesi olarak tanımlanmaktadır. Burada durum uzayı Θ , ilgilenilen bir parametre vektörünün desteğini temsil eden parametre uzayı olarak ifade edilebilir. T genellikle ardışık, çift zaman aralıklarını gösteren pozitif tamsayılar kümesidir. Böylece Markov zinciri, t anında, yeni bir duruma geçiş olasılığı sadece sürecin mevcut $\theta^{(t)}$ durumuna bağlı olan ve bu nedenle koşullu olarak $\theta^{(0)}, \theta^{(1)}, \dots, \theta^{(t-1)}$ durumlarından bağımsız olan bir stokastik süreç olarak tanımlanmaktadır. Mevcut durum verildiğinde, geçmiş ve gelecek durumlardan bağımsız olan Markov zinciri şu şekilde ifade edilebilir;

$$p(\theta^{(t)} \in A | \theta^{(0)}, \theta^{(1)}, \dots, \theta^{(t-2)}, \theta^{(t-1)}) = p(\theta^{(t)} \in A | \theta^{(t-1)}) \quad (2.32)$$

burada A , bir olay için tam durum uzayında tanımlanmış herhangi bir kümedir ve (2.32) deki olasılıklar t ve A 'ya bağlıdır. Bu durumda, Markov zinciri, sadece son dönemde nerde bulunduğunu hatırlayarak durum uzayında dolaşmaktadır. Böylece, zincir en sonunda en yüksek yoğunluğa sahip durum uzayı bölgesini bularak bu dağılımdan ilgilenilen son dağılımı özetlemek için kullanılan bağımsız olmayan bir örnek üretecektir. Burada zincirin mevcut durumu göz önüne alındığında, durum uzayındaki diğer noktalara hareket etme olasılıkları için geçiş sürecinin tanımlanması gerekmektedir (Gill 2015: 334).

Bu geiş süreci mevcut zincir durumuna baėlı olarak belli bařka bir duruma geme olasılıėının elde edilmesini saėlayan genel bir mekanizma olarak geiş kerneli $K(\theta, A)$ tarafından tanımlanmaktadır (Robert, Casella 2004: 208).

Bu durumda bir geiş kerneli $K(\theta, A)$ řu řekilde tanımlanabilir;

- 1- tüm $\theta \in \Theta$ için, $K(\theta, .)$ Θ parametre uzayı üzerinde bir olasılık daėılımıdır,
- 2- tüm $A \subset \Theta$ için, $K(\theta, A)$ fonksiyonu, tüm θ 'lar için deėerlendirilebilir olmalıdır.

$\Theta = \{\theta_1, \theta_2, \dots, \theta_r\}$ kesikli durum uzayı için $P(\theta_i, \theta_j)$ ile verilen (i, j) . elemanlı bir geiş matrisi P tanımlanmaktadır. Θ , r elemanlı sonlu ise, geiş matrisi P řu řekilde verilmektedir:

$$P = \begin{bmatrix} P(\theta_1, \theta_1) & \dots & P(\theta_1, \theta_r) \\ \vdots & & \vdots \\ P(\theta_r, \theta_1) & \dots & P(\theta_r, \theta_r) \end{bmatrix}_{r \times r} \quad (2.33)$$

Bu tür matrislere stokastik matris denilmektedir ve burada satır zincirin bu dönemde nerede olduėunu, sütün ise zincirin bir sonraki dönemde nereye gittiėini göstermektedir. P geiş matrisinin elemanları herhangi bir olasılık daėılımında olması gerektiėi gibi; her $i, j \in A$ için $P(\theta_i, \theta_j) \geq 0$, ve her i . satır için: $\sum_{j=1}^r p(\theta_i, \theta_j) = 1$, iki temel özelliėe sahip olmalıdır (Gamerman, Lopes 2006: 116). Durum uzayı sürekli olduėunda K geiş kerneli kořullu olasılık yoėunluk fonksiyonu $p(\theta|\theta_i)$ řeklinde ifade edilir ve mevcut durum θ_i verildiėinde tüm $\theta \in A$ için uygun řekilde tanımlanmıř bir yoėunluėu göstermektedir (Gill 2015: 335).

$P^m(x, y)$ ile gösterilen m adımda x durumundan y durumuna geiş olasılıkları, tam olarak m adımla x durumundan y durumuna hareket eden bir zincir olasılıėı ile verilmektedir. Geiş çekirdeėi sayesinde, kesikli bir durum uzayı için 0 zamanında m adımla x durumundan y durumuna geiş olasılıėı bu iki durum arasındaki geiş olasılıklarınnın arpımı řeklinde hesaplanabilmektedir. $m > 2$ için řu řekilde elde edilebilir:

$$\begin{aligned}
P^m(x, y) &= P(\theta^{(m)} = y \mid \theta^{(0)} = x) \\
&= \sum_{x_1} \dots \sum_{x_{m-1}} P(\theta^{(m)} = y, \theta^{(m-1)} = x_{m-1}, \dots, \theta^{(1)} = x_1 \mid \theta^{(0)} = x) \\
&= \sum_{x_1} \dots \sum_{x_{m-1}} P(\theta^{(m)} = y \mid \theta^{(m-1)} = x_{m-1}), \dots, P(\theta^{(1)} = x_1 \mid \theta^{(0)} = x) \\
&= \underbrace{\sum_{x_1} \sum_{x_2} \dots \sum_{x_{m-1}}}_{\text{tüm olası durumlar}} \underbrace{P(x, x_1)P(x_1, x_2)P(x_{m-1}, y)}_{\text{geçiş olasılıklarının çarpımı}}
\end{aligned} \tag{2.34}$$

Son eşitlik, $P^m(x, y)$ elemanlarını içeren matrisin de bir stokastik matris olduğunu ve P geçiş matrisinin elemanlarının m durumlarının matris çarpımı ile elde edilen P^m ile verildiğini göstermektedir (Gamerman, Lopes 2006: 116).

Ayrıca, genel fikri (2.34) eşitliğine dayanan Chapman-Kolmogorov denklemleri ardışık olayların olasılıkla birbirine nasıl bağlandığını belirlemektedir. $P^1(x, y) = P(x, y)$ ve $P^0(x, y) = I(x, y)$ olmak üzere, kesikli durum uzayları için Chapman-Kolmogorov denklemleri için şu şekildedir:

$$\begin{aligned}
P^{m+n}(x, y) &= \sum_z P(\theta^{(m+n)} = y \mid \theta^{(n)} = z, \theta^{(0)} = x) P(\theta^{(n)} = z \mid \theta^{(0)} = x) \\
&= \sum_z P^n(x, z) P^m(z, y)
\end{aligned} \tag{2.35}$$

(2.35) eşitliği bir dizi geçiş matrisi çarpımı olarak yazılabilir. Bu yüksek geçiş olasılıklar ile daha yüksek geçiş matrisleri oluşturulabilir ve bu matrisler $P^{m+n} = P^m \cdot P^n$ ve $P^{n+1} = P^n \cdot P$ gibi özellikleri sağlayarak, yenilemeli olasılıkların ara dağılımlara bağlı olarak istenilen şekilde parçalara ayrıştırılmasını sağlamaktadır (Gill 2015: 338; Gamerman, Lopes 2006: 116).

2.2.3.1. Markov Zincirlerinin Genel Özellikleri

Markov Zinciri Monte Carlo yaklaşımı göz önüne alındığında, Markov zincirleri için temel sonuçlar hakkında bilgi sahibi olmadan, bu yöntemi ve Bayes uygulamalarının ayrıntılarını tam olarak anlamak mümkün olmamaktadır. Algoritma açısından düşünüldüğünde, bir Markov zincirinin, simüle edilen modelin kabul edilebilir bir yaklaşımını bulmak için çok iyi kararlılık özelliklerine sahip olması

gerekmektedir. Bu nedenle aşağıdaki bölümde Markov zincirlerinin genel özellikleri: indirgenemezlik, yenilenme, homojenlik, durağanlık ve ergodiklik başlıkları altında özetlenmektedir.

2.2.3.1.1. İndirgenemezlik

İndirgenemezlik özelliği, Markov zincirinin x_0 veya μ gibi başlangıç koşullarına duyarlılığının ilk ölçüsü olarak bilinmektedir. Markov Zinciri Monte Carlo algoritmalarının kurulumunda çok önemlidir, çünkü kesin yakınsama özelliğine sahiptir ve bu sayede kabul edilebilir başlangıç koşullarını belirlemek için gerekli olan geçiş operatörünün ayrıntılı çalışmasına ihtiyaç duyulmamaktadır (Robert, Casella 2004: 213).

Geçiş kerneli P 'ye sahip bir Markov zinciri olduğu varsayılın, bu zinciri kullanan bir MCMC denemesinin başarılı olması için örnek yol ortalaması;

$$\bar{f}_n = \frac{1}{n+1} \sum_{i=0}^n f(X_i), \quad (2.36)$$

şeklinde olmak zorundadır. Bu beklenen değer mevcut olduğunda herhangi bir dağılım için beklenen değer f 'ye yakınsamaktadır. Bu özelliğin sağlanması için temel gereklilik, zincirin durum uzayının tüm kısımlarına ulaşabilmesi yani zincirin indirgenemez olmasıdır. Genel durum uzayı Markov zincirleri için indirgenemezlik bir dağılıma göre tanımlanmaktadır. Bir Markov zinciri, $A \subset \Theta$ kümesi için $\varphi(A) > 0$ olması durumunda Θ parametre uzayı üzerindeki bir φ olasılık dağılımı için φ -indirgenemez olduğu tüm $x \in \Theta$ için şu şekilde gösterilmektedir:

$$P_x \{ \tau_A < \infty \} > 0 \quad (2.37)$$

Bir Markov zinciri φ -indirgenemezse, zincir için φ indirgenemezlik dağılımı olarak adlandırılmaktadır. Bir zincir indirgenemezse, birçok farklı indirgenemezlik dağılımına sahiptir (Gilks vd. 1997: 62).

2.2.3.1.2. Yenileme

Bir zincir için indirgenemezlik, durum uzayındaki tüm kümelere ulaşılabilceği anlamına gelmektedir. Yenileme, bu tür kümelerin hepsine hemen hemen tüm başlangıç noktalarından sonsuz sıklıkta erişilebilmesi özelliği olarak ifade edilmektedir. Maksimum indirgenemezlik dağılımı ψ 'ye sahip

indirgenemez bir Markov zinciri, eğer $\psi(A) > 0$ olan herhangi bir $A \subset \Theta$ kümesi için aşağıdaki iki koşul sağlanıyorsa yenilenme özelliğine sahip olduğu söylenebilir:

- 1- Tüm x 'ler için $P_x\{X_n \in A \text{ sonsuz sıklıkla}\} > 0$
- 2- Yaklaşık olarak tüm x 'ler için $P_x\{X_n \in A \text{ sonsuz sıklıkla}\} = 1$ olmalıdır.

İndirgenemez tekrarlayan bir zincir, eğer değişmez bir olasılık dağılımına sahipse pozitif olarak tekrarlanır, aksi takdirde zincir geçersiz (null) bir yenileme olmaktadır. Burada, (1) koşulu her yerde olduğu sürece (2) koşulunun boş bir küme üzerinde başarısız olması durumu ortaya çıkmaktadır. Harris tekrarı kavramı, (2) koşulunun her yerde geçerli olmasını gerektirerek bu sorunu ortadan kaldırmak için, tüm $x \in \Theta$ için $\psi(A) > 0$ ve $A \subset \Theta$ olmak üzere;

- 3- $P_x\{X_n \in A \text{ sonsuz sıklıkla}\} = 1$ olacak şekilde tasarlanmıştır (Gilks vd. 1997: 63-66).

Kesikli ve sonlu durum uzaylarında indirgenemez bir Markov zinciri, zincirin sonsuz sıklıkta sınırsız süre boyunca A 'ya uğrama olasılığı sıfır değilse, tek bir nokta ya da tanımlanmış bir nokta kümesi olan belirli bir duruma göre yenilenen olarak adlandırılmaktadır. Bu durum, A 'ya beklenen dönüş sayısının sonsuz olduğunu söyler ve (1) koşulu ile ifade edilmektedir. Ancak sınırsız ve sürekli durum uzaylarında A 'yı sonsuz sıklıkta ziyaret etme olasılığını kesinleştiren daha kararlı bir yenileme olan Harris yenileme tanımı yapılmıştır. Bu durumda, (3) eşitliğinde ifade edildiği gibi, zincirin A 'ya sonsuz sıklıkta uğrama olasılığının bire eşit olduğunu söyler ve A kümesinin Harris tekrarlayan olduğu söylenir. Bir zincir kapalı, indirgenemez yenilenen Harris durumları kümesine girdiğinde, orada kalır ve her bir durumu bir olasılıkla ziyaret ederek her noktadaki yakınsamayı neredeyse kesin olan yakınsama ile değiştirir (Gill 2015: 341).

2.2.3.1.3. Homojenlik

Belirli bir mevcut duruma koşullu olunduğunda, gelecekteki durumların olasılığı geçmişe bağlı olmaz ve bu koşullu bağımsızlık özelliğine sahip bir süreç Markov zinciri olarak tanımlanır. $\theta \in \Theta$ ve $A \subset \Theta$ bir küme olmak üzere bir Markov zinciri, tüm A olayları ve $t \geq 0$ için

$$p(\theta^{(t+1)} \in A | \theta^{(0)}, \theta^{(1)}, \dots, \theta^{(t)}) = p(\theta^{(t+1)} \in A | \theta^{(t)}) \equiv K_t(\theta, A), \quad (2.38)$$

şeklinde ifade edilir ve burada olasılık $K_t(\theta, A)$, t anında tek adımlı geçiş fonksiyonunu göstermektedir. Benzer şekilde $A = (-\infty, \omega]$ olmak üzere tüm $\theta, \omega \in \Theta$ için bir Markov zinciri koşullu dağılım fonksiyonları ile şu şekilde tanımlanabilmektedir:

$$F_{\theta_{(t+1)}}(\omega | \theta^{(0)}, \dots, \theta^{(t)}) = F_{\theta_{(t+1)}}(\omega | \theta^{(t)}) \equiv F_t(\theta, \omega). \quad (2.39)$$

Burada geçiş fonksiyonu t ' ye göre değişmez olduğunda, yani Markov zinciri t değerine bağlı olmayan bir geçiş olasılığına sahip olduğunda, $K(\theta, A)$ veya $F(\theta, \omega)$ yazılabilir. Bu durumda Markov zinciri homojenlik özelliğine sahip olur (Türkman vd. 2019: 91).

2.2.3.1.4. Durağanlık

Durağanlık Markov zincirinin sahip olması gereken önemli bir özelliktir. Çünkü, bir Markov zinciri durağan ise, durum uzayında (Θ) nerede olursak olalım, durum uzayının tüm A bölgelerine olasılık atamaları yapmak için aynı (durağan) dağılım π 'yi kullandığımız anlamına gelmektedir (Jackman 2009: 178).

Simülasyon bağlamında Markov zincirleri için bir problem, zincirin asimptotik davranışının, adımların veya yenilemelerin sayısı $t \rightarrow \infty$ olarak incelenmesidir. Bu durumda çözüm, Markov zinciri için durağan bir π dağılımıdır. Kesikli ve sürekli durumlarda π dağılımının, θ_i durumundan θ_j durumuna geçiş olasılıkları $P(\theta_i, \theta_j)$ olan bir zincirin durağan dağılımı olması her $\theta_j \in \Theta$ ve $A \subset \Theta$ için, şu şekilde ifade edilmektedir:

$$\sum_{\theta_i \in \Theta} \pi(\theta_i) P(\theta_i, \theta_j) = \pi(\theta_j) \quad (2.40)$$

$$\int_{\Theta} \pi(\theta_i) P(\theta_i, \theta_j) d(\theta_i) = \pi(\theta_j).$$

Buradan, geçiş olasılıkları ile çarpma ve mevcut nokta için değerlendirme (toplama ve integral alma işlemi) aynı marjinal dağılımı üretmektedir. Matris gösteriminde $\pi = \pi P$ olarak yazılabilir ve herhangi bir t adımdaki marjinal dağılım π ise, sonraki adımdaki dağılım yine $\pi = \pi P$ olmaktadır. Yani, zincir π 'nin zincir dağılımı olduğu aşamaya ulaştığında, zincir bu dağılımı sonraki tüm aşamalar için korur ve bu da zincir durağan uzaklığa ulaştığında marjinal dağılımın sabit kaldığını gösterir. Bu dağılım, değişmez veya denge dağılımı olarak da bilinmektedir (Gamerman, Lopes 2006: 121).

Markov zincirleri için durağanlık özelliğine sahip olmasının önemli sonuçları şu şekilde özetlenebilir:

- 1- Son dağılım $P(\theta|y)$ bilinmeden Markov zinciri başlatılabilir,
- 2- Uygun bir geçiş kerneli $K(\theta, .)$ ile bir Markov zinciri oluşturulabilir,
- 3- Yeterince büyük sayıda zincir yenilemesinden sonra, $K(\theta, .)$ parametre uzayı üzerinde, sadece ilgili son dağılıma $P(\theta|y)$ karşılık gelenlere değil aynı zamanda zincirin tekrarlanan yenilemelerine de olasılık atamaları üretir.

Böylece, yeterince büyük T ve $r > 0$ için Markov zinciri $\{\theta^{(T+r)}\}$ 'nin yörüngesi, durağan yoğunluktan yani bu durumda ilgili $P(\theta|y)$ son dağılımdan çekilen örneklerden oluşur (Jackman 2009:178).

2.2.3.1.5. Ergodiklik

Markov zincirleri için, durağan dağılımların olduğu ancak sınırlayıcı dağılımların olmadığı durumlar vardır. Sınırlayıcı sonuçlar oluşturmak için, bir Markov zincirinin periyodunun tanımlanması gerekmektedir. Periyod, aynı zincir döngüsünün tekrarını oluşturmak için geçen süre olarak tanımlanmaktadır. Periyodiklik ise bir Markov zincirinin durağan bir yoğunluğa yaklaşamaması durumudur. $d\theta_i$ ile gösterilen θ_i durumunun periyodu, kümenin en büyük ortak böleni olarak şu şekilde ifade edilmektedir; $\{t \geq 1: P^t(\theta_i, \theta_i) > 0\}$.

Buradan $P(\theta_i, \theta_i) > 0$ ve $d\theta_i = 1$ ise indirgenemez bir zincirin durumları aynı periyoda aittir. Periyodik olmama, geçiş olasılıklarının yakınsamasını sağlamak için gerekli bir koşuldur. θ_i durumu periyodik değilse, $d\theta_i = 1$ ve buna ek olarak bu durum pozitif Harris yenilenen ise durumun ergodik olduğu söylenmektedir. Yani tek başına yenileme özelliği, zincirin, marjinal dağılımının sabit kaldığı ve ilgili son dağılım ile aynı kaldığı bir duruma ulaşmasını sağlamak için yeterli olmamaktadır. Bir zincirin tüm durumları ergodik ise zincir ergodiktir. Böylece, tüm θ_i ve θ_j durumları için Markov zincirleri,

$$\lim_{t \rightarrow \infty} P^t(\theta_i, \theta_j) = \pi(\theta_j) \quad (2.41)$$

özelliğini taşımaktadır. Zincirin ergodikliği oluşturulduktan sonra, bu Markov zincirinin durağan bir dağılımı için, zincir rastgele bir başlangıç dağılımından hangi koşullar altında bu durağan dağılıma yakınsadığı ergodik teoremi ile ifade edilebilir. Gerçek değerli bir $g(\theta)$ fonksiyonunun ergodik ortalaması $\overline{g}_t = \frac{1}{t} \sum_{i=1}^t g(\theta_i)$ olmak üzere, belirli dağılım π için, zincir ergodik ve $E_\pi[g(\theta)] < \infty$ ise, $t \rightarrow \infty$ giderken neredeyse kesin olarak,

$$\bar{g}_t \rightarrow E_\pi [g(\theta)], \quad (2.42)$$

yakınsamaktadır. Sonuç olarak, ergodik özelliğine sahip bir Markov zincirinden çekilen örneklerle, ilgili son dağılım parametrelerinin tutarlı tahminleri elde edilebilmektedir (Gamerman, Lopes 2006: 125; Jackman 2009: 187).

2.2.3.2. Metropolis-Hastings Algoritması

Metropolis – Hastings algoritması, Metropolis vd. (1953) tarafından geliştirilen ilk Markov Zinciri Monte Carlo yöntemidir. Burada örnek almak için özel bir dağılıma ihtiyaç duyulmaz ve son dağılım $P(\theta|y)$ için uygulanabilir (Koch 2007: 216).

Genel olarak, Markov zinciri Monte Carlo yöntemlerinin çalışma prensibi şu şekilde ifade edilebilir. İlgili yoğunluk $P(\theta|y)$ verildiğinde, durağan dağılımı $P(\theta|y)$ olan bir Markov kerneli $K(\theta,.)$ oluşturulmaktadır. Daha sonra bu kernel fonksiyonu kullanılarak bir Markov zinciri $(\theta^{(t)})$ oluşturulur ve böylece ergodik teoremi ile son dağılımın özellikleri (integraller) yaklaşık olarak hesaplanmaktadır. Buradaki en önemli nokta, keyfi bir $P(\theta|y)$ yoğunluğu ile ilişkili bir $K(\theta,.)$ kerneli oluşturmaktır. Metropolis – Hastings algoritması, bu tür kernel fonksiyonların üretilmesini sağlayan yöntemlerden biridir. İlgili yoğunluk $P(\theta|y)$ ve koşullu yoğunluk $q(.|\theta)$ ile Metropolis – Hastings algoritması, bir geçiş kerneli aracılığıyla bir Markov zinciri $(\theta^{(t)})$ üretmektedir (Robert, Casella 2008: 171).

Metropolis – Hastings algoritması, belirtilen Bayes son dağılıma yakınsamak için bir kabul/red kuralı kullanan rastgele bir yürüyüşün uyarlamasıdır. Algoritma şu şekilde ilerlemektedir:

- 1- $P(\theta^0|y) > 0$ olan bir başlangıç değeri $\theta^{(0)}$ seçilir,
- 2- $t = 1, 2, \dots, T$ için;
- a) t anında bir öneri (proposal) yoğunluğu $q_t(\theta^*|\theta^{(t-1)})$ 'dan bir aday çekim θ^* alınır,

Metropolis algoritması için, öneri dağılımı simetrik olmalıdır ve tüm θ_a, θ_b ve t için $q_t(\theta_a|\theta_b) = q_t(\theta_b|\theta_a)$ koşulunu sağlamalıdır.

- b) Yoğunluklar oranı r hesaplanır;

$$r = \frac{P(\theta^* | y)}{P(\theta^{(t-1)} | y)}. \quad (2.43)$$

c) Bir kabul olasılığı $\alpha = \min \{r, 1\}$ hesaplanır ve α olasılığı ile $\theta^{(t)} = \theta^*$ ve $(1 - \alpha)$ olasılığı ile $\theta^{(t)} = \theta^{(1-t)}$ örneği alınır.

3- Bir ve ikinci adımlar T kez tekrarlanarak $\theta^{(T)}$ simüle dizisi elde edilir.

Buradan, $g(\theta)$ ortalama, varyans gibi son dağılımın özelliklerini gösteren ilgili fonksiyon olmak üzere, $g(\theta^{(1)}), \dots, g(\theta^{(T)})$ çekimlerin ortalaması alınarak; $\widehat{g}_T = \frac{1}{T} \sum_{t=1}^T g(\theta^{(t)})$ ilgili fonksiyon için $E[g(\theta)|y]$ tahminleri elde edilmektedir (Gelman vd. 2004: 296; Koop 2003: 93).

Burada, başlangıç değerleri, maksimum olabilirlik tahmini veya başka bir yöntemle elde edilmektedir. MCMC teorisi, algoritma tarafından üretilen Markov zincirinin yakınsadığı durağan dağılımının seçilen başlangıç değerinden bağımsız olarak ilgili son dağılım olacağını söylemektedir. İkinci adımda çekilen değer aday olarak kabul edilir çünkü otomatik olarak ilgili dağılımdan bir çekiliş olarak kabul edilmez ve red örneklemesinde olduğu gibi değerlendirilmesi gerekmektedir. Red örnekleme ile tek fark, burada öneri yoğunluğunun son dağılımı kapsamına gerek yoktur ve öneri yoğunluğu örnekleme kolay olan (normal, uniform, vb.) herhangi bir biçimde olabilir. Metropolis – Hastings algoritması, tüm ortak yoğunluk fonksiyonunu kullanarak, olasılık dağılımından örnekler üreten bir algoritmadır. Reddetme örnekleme gibi diğer örnekleme yöntemlerine göre Metropolis – Hastings algoritmasının önemli bir avantajı, çok değişkenli dağılımlarla çalışabilmektedir (Lynch 2007: 108).

2.2.3.3. Gibbs Örnekleme

Çoğu istatistiksel modelde olduğu gibi, θ yüksek boyutlu olduğunda son dağılım $P(\theta|y)$ 'den doğrudan örnekleme daha önce tanımlanan örnekleme algoritmaları veya Metropolis-Hastings algoritması için mümkün olmamaktadır. Bu durumda, yüksek boyutlu son dağılım $P(\theta|y)$ 'den örnekleme problemini, birbiriyle ilişkili daha düşük boyutlu örnekleme problemlerine ayıran algoritmalara ihtiyaç duyulmaktadır. Daha düşük boyutlu örnekleme yaklaşımı ile yüksek boyutlu yoğunluktan örnek çekmek yerine, birlikte ortak yoğunluğu tanımlayan daha düşük koşullu yoğunluklardan örnekleme yapılmaktadır. Bu yaklaşım, en yaygın kullanılan MCMC algoritmalarından biri olan Gibbs örnekleycisidir (Jackman 2009: 214).

Gibbs örnekleme, geçiş kernelinin tam koşullu dağılımlarla oluşturulduğu bir MCMC yaklaşımıdır. İlgili dağılımın $P(\theta|y)$ ve parametre vektörü $\theta = (\theta_1, \dots, \theta_d)'$ d bloklara ya da alt vektörlere bölündüğü varsayılın. Burada θ_i bileşenlerinin her biri bir skaler, vektör veya bir matris olabilir. Ayrıca, $i = 1, \dots, d$ için tam koşullu dağılımların $P_i(\theta_i|y) = P(\theta_i|y, \theta_{-i})$ tamamen bilindikleri ve örneklenebildikleri göz önünde bulundurulmalıdır. İlgili dağılım P 'den doğrudan rastgele çekimler

karmaşık, maliyetli ya da kısaca mevcut olmadığında, ancak tam koşullu dağılımlar P_i 'den çekimler mümkün olduğunda Gibbs örnekleme kullanılmaktadır. Gibbs örnekleme, tam koşullu dağılımlardan ardışık çekimlere dayanan alternatif bir üretim şeması sağlamaktadır. Gibbs örneklemesinin adımları şu şekilde tanımlanabilir:

1.adım: Başlangıç değeri $\theta^{(0)} = (\theta_1^{(0)}, \dots, \theta_d^{(0)})'$ seçilir ve $j = 1$ için zincirinin yenileme sayacı başlatılır,

2. adım: Art arda değer üretim yoluyla $\theta^{(j-1)}$ 'den yeni bir $\theta^{(j)} = (\theta_1^{(j)}, \dots, \theta_d^{(j)})'$ değeri elde edilir,

$$\theta_1^{(j)} \sim P(\theta_1 | y, \theta_2^{(j-1)}, \dots, \theta_d^{(j-1)}),$$

$$\theta_2^{(j)} \sim P(\theta_2 | y, \theta_1^{(j)}, \theta_3^{(j-1)}, \dots, \theta_d^{(j-1)}), \quad (2.44)$$

$\vdots$

$$\theta_d^{(j)} \sim P(\theta_d | y, \theta_1^{(j)}, \theta_2^{(j)}, \dots, \theta_{d-1}^{(j)});$$

3.adım: j sayacı $j + 1$ olarak değiştirilir ve yakınsama sağlanana kadar ikinci adıma dönülür.

Yakınsama sağlandığında, elde edilen değer $\theta^{(j)} = (\theta_1^{(j)}, \dots, \theta_d^{(j)})'$, P 'den bir çekimdir. Yenileme sayısı arttıkça zincir denge durumuna yaklaşır ve yakınsamanın yaklaşık olarak geçerli olduğu varsayılır. Yakınsama dağılımında olduğu için, tüm parametre bileşenlerinin ortak dağılımının tüm θ değerleri için ortak son dağılıma yakınsaması gerektiği anlamına gelmektedir (Gamerman, Lopes 2006: 142).

2.2.3.4. Hibrit Monte Carlo

Gelişmiş Markov zincir simülasyon algoritmaları, genellikle yardımcı değişkenler yardımı ile Metropolis-Hastings algoritması çerçevesi üzerine inşa edilmektedir. Gibbs örnekleme ve Metropolis algoritması gibi MCMC algoritmalarındaki verimsizlik rastgele yürüme davranışı ve gecikmiş yakınsama, simülasyonlar ilgili dağılıma yaklaşırken çizimlerin uzun zaman almasıdır. Özellikle yüksek boyutlu ilgili dağılımlarda, Hibrit Monte Carlo (HMC) algoritması, Metropolis algoritmasındaki yerel rastgele yürüme davranışını baskılayan yardımcı değişkenler ekleyerek, ilgili dağılıma çok daha hızlı yaklaşılması sağlanmaktadır. İlgili alandaki her q_j bileşeni (q parametre vektörü) için, Hibrit Monte Carlo q ile aynı boyutlu bir yardımcı momentum vektörü p ekler. Burada momentum, q 'daki sıçramanın beklenen mesafesini ve yönünü vermektedir. Böylece birbirini izleyen sıçramalar aynı yönde olma eğilimindedir ve simülasyon mümkün olduğunda q 'nın alanı içinde hızlı hareket ederler. Hibrit Monte

Carlo fizikteki Hamilton dinamikleri modeliyle ilgili olduğu için aynı zamanda Hamilton Monte Carlo olarak da adlandırılmaktadır (Gelman vd. 2004: 344).

Bir MCMC yöntemi inşa etmek için ilk adım olarak, örneklemek istenilen olasılık dağılımı için bir Hamilton fonksiyonunun tanımlanması gerekmektedir. İlgili değişkenlere (q) karşılık gelen konum değişkenlerine ek olarak bağımsız normal dağılıma sahip yardımcı momentum değişkenleri (p) tanımlanmaktadır. Hibrit Monte Carlo yaklaşımı, bu momentum değişkenleri için basit güncellemeleri, sıçrama (leapfrog) yöntemiyle uygulanan Hamilton dinamiklerine göre bir yörünge hesaplayarak yeni bir durumun önerildiği Metropolis güncellemeleri ile değiştirmektedir.

Hamilton dinamiği, her ikisi de d boyutlu ilgili parametre vektörü q ve momentum vektörü p üzerinde çalışmaktadır. Hibrit Monte Carlo için genellikle şu şekilde yazılabilen Hamilton fonksiyonu kullanılmaktadır;

$$H(q, p) = U(q) + K(p). \quad (2.45)$$

Buradan, ilgili q parametreleri için son dağılım, $U(q)$ potansiyel enerji fonksiyonu kullanılarak kanonik bir dağılım olarak şu şekilde tanımlanmaktadır;

$$U(q) = -\log[\pi(q)L(q|y)], \quad (2.46)$$

burada $\pi(q)$ ilk dağılım ve $L(q|y)$, y verileri için olabilirlik fonksiyonudur. $K(p)$ kinetik enerji olarak adlandırılır ve şu şekilde tanımlanmaktadır;

$$K(p) = p^T M^{-1} p / 2. \quad (2.47)$$

Burada $M = \sigma^2 I$ olmak üzere, simetrik ve pozitif tanımlı bir matristir. Bu form, sıfır ortalamalı çok değişkenli normal dağılımının M kovaryans matrisi ile olasılık yoğunluğunun logaritma değerlerine karşılık gelmektedir. Hibrit Monte Carlo, sadece yoğunluk fonksiyonunun değerlendirildiği $\mathbb{R}^d$ üzerindeki sürekli dağılımlardan örnekleme için kullanılmaktadır. Hibrit Monte Carlo algoritmasının her yenilemesinin iki adımı vardır:

1. adım: Mevcut q değerlerinden bağımsız olarak, momentum değişkenleri için yeni p^* değerleri, $K(p)$ ile ifade edilen çok değişkenli normal dağılımdan rastgele olarak çekilir,
2. adım: Yeni bir durum önermek için Hamilton dinamikleri kullanılarak bir Metropolis güncellemesi gerçekleştirilir. Aday (q^*, p^*) oluşturmak için, şu an ki durum (q, p) ile başlatılarak Hamilton dinamiği sıçrama (leapfrog) yöntemi kullanılarak L kez ε adımları için simüle edilir.

Bu önerilen yeni durum ya kabul edilir ya da mevcut durum H fonksiyonu kullanılarak standart bir Metropolis kararı ile Markov zincirinin bir sonraki durumu olarak kabul edilir:

$$\min[1, e^{(-H(q^*, p^*) + H(q, p))}] = \min[1, e^{(-U(q^*) + U(q) - K(p^*) + K(p))}] \quad (2.48)$$

Burada L ve ε , en iyi performans elde etmek için ayarlanması gereken kritik öneme sahip algoritma parametreleridir (Neal 2011: 114-124).

Yalnızca mantıklı durum önerileri sunulduğu için Hibrit Monte Carlo her öneriyi her zaman kabul etmektedir. Hibrit Monte Carlo, Metropolis veya Gibbs örneklemesine göre hesaplama açısından daha maliyetlidir. Ancak öneri yoğunlukları çok daha verimli olduğu için Hibrit Monte Carlo son dağılımı açıklamak için çok sayıda örneğe ihtiyaç duymamaktadır. Hibrit Monte Carlo algoritmasının en önemli özelliği diğer algoritmaların çalışmadığı binlerce parametreye sahip çok düzeyli modeller için çok verimli bir şekilde örnekleme yapabilme yeteneğine sahip olmasıdır (McElreath 2019: 283).

2.2.4. Bayesyen Hiyerarşik Modeller

Önceki bölümde belirtildiği gibi Bayesyen analizde, birçok istatistiksel uygulama karmaşık modellerin tanımlanmasını gerektirmektedir. Genellikle gerçek hayattaki modeller, uygulamaya özel kısıtlamalarla belirlenen temel bir yapıya göre bağlanması gereken çok sayıda parametrenin kullanılmasını gerektirmektedir. Bayesyen yaklaşım, hiyerarşik modeller ile bu tür modellerin analizine imkân sağlamaktadır (Hamada vd. 2008: 69).

Analiz birimleri bir anakütle içindeki belirli veri kümelerinden alındığında, aynı kümeden gelen bireyler diğer kümelerden gelen bireylere göre birbirlerine daha çok benzedikleri için artık bağımsız olarak kabul edilemezler. Bu durumda yani gözlemler bağımsız değilse benzerlik fonksiyonu hesaplanamaz. Bu nedenle hiyerarşik modeller oluşturmanın bir nedeni bağımsızlık varsayımı sağlanmadığında, çoğunlukla standart hatalarda ortaya çıkan sapmaları gidermektir (Lynch 2007: 232).

Sosyal bilimlerde Bayesyen yaklaşımın kullanımındaki artışının temel nedenlerinden biri hiyerarşik modellerin yaygın olarak kullanılmaya başlanmasıdır. Hiyerarşik veya çok düzeyli bir model, parametrelere bir olasılık modeli verilen bir regresyon olarak tanımlanmaktadır. Bu ikinci seviye model verilerden tahmin edilen hiperparametre olarak adlandırılan kendi parametrelerine sahiptir. Hiyerarşik modellerin iki temel ögesi; değişen katsayılar ve bu değişen katsayılar için verilen bir olasılık modelidir (Gelman, Hill 2006: 3).

Bayesyen yaklaşımda ilgilendiğimiz temel şey, ilk dağılımlar ile benzerlik fonksiyonunun çarpımı ile orantılı bir son dağılımı elde etmektir: $P(\theta|y) \propto L(y|\theta)\pi(\theta)$ Hiyerarşik modeller için, θ

parametresinin $\pi(\psi)$ ilk dağılımı ile verilen başka bir bilinmeyen ψ parametresine bağlı olduğu varsayalım. Bu durumda son dağılım hesaplaması şu şekildedir:

$$P(\theta, \psi | y) \propto L(y | \theta) \pi(\theta | \psi) \pi(\psi). \quad (2.49)$$

Burada θ 'nın ilk dağılımı $\pi(\theta | \psi)$ vardır, ancak bu ilk dağılım hiperparametre olarak adlandırılan kendi ilk olasılığı $\pi(\psi)$ olan başka bir parametreye koşulludur. Her iki parametre için çıkarım ilgili marjinal yoğunluklara bakılarak aşağıdaki gibi elde edilebilir;

$$P(\theta | y) = \int_{\psi} P(\theta, \psi | y) d\psi, \quad (2.50)$$

$$P(\psi | y) = \int_{\theta} P(\theta, \psi | y) d\theta.$$

Eşitlik (2.49)'dan ψ 'yi başka bir parametreye $\pi(\psi | \zeta)$ koşullandırarak yeni hiper ilk dağılım $\pi(\zeta)$ hesaplamaya eklenebilir. Bu durumda son olasılıklar, bu ilk dağılımlar eşitliğinin sağ tarafına eklenerek,

$$P(\theta, \psi, \zeta | y) \propto L(y | \theta) \pi(\theta | \psi) \pi(\psi | \zeta) \pi(\zeta), \quad (2.51)$$

şeklinde hesaplanır ve burada ζ en yüksek seviyeli parametredir ve bu nedenle koşulsuz olan tek hiper ilk dağılımdır. Hiyerarşik modelleri, hiper ilk dağılım düzeyinden veri seviyesine kadar şu şekilde gösterebiliriz: (Gill 2015: 425).

$$\begin{array}{c} \pi(A, B) \searrow \\ \underbrace{\pi_{\lambda}(\alpha, \beta)}_{\text{İlk dağılım}} \rightarrow \underbrace{\lambda}_{\text{Parametre}} \rightarrow \underbrace{y}_{\text{Gözlemlenen veri}} \\ \underbrace{\pi(C, D)}_{\text{Hiper ilk dağılım}} \nearrow \end{array} \quad (2.52)$$

Genel olarak hiyerarşik modeller çok daha karmaşık yapılar mevcut olduğunda diğer (hiyerarşik olmayan) modellerden daha esnek yaklaşım sunmaktadır. Hiyerarşik modellerin önemli bir özelliği belirli bir gruba atıfta bulunan her parametrenin benzer özelliklere sahip diğer gruba karşılık gelen parametrelerden güç alabilmesidir. Bu, özellikle bazı gruplarda az sayıda gözlem olduğunda oldukça faydalı olabilmektedir. Ayrıca hiyerarşik yaklaşımda son dağılım ilgili parametrelerin farklı ilk dağılım seçimlerinin ortalaması alınarak hesaplandığı için sübjektiflik azaltılır ve böylece daha sağlam bir analiz sağlanır. Son olarak hiyerarşik yapı karmaşık veri kümeleri için son dağılım sayesinde, modelin yorumlanmasını ve hesaplanmasını basitleştirir (Ntzoufras 2009).

3.NONPARAMETRİK BAYESYEN YAKLAŞIM VE DIRICHLET SÜREÇLER

Nonparametrik Bayesyen (BNP) modeller temel olarak sonsuz sayıda parametreye izin veren esnek modellemeye ve üretici yaklaşımlara odaklanmaktadır. Nonparametrik Bayesyen yöntemler, arkada gözlemlenen verilerin koşullu olarak bağımsız olduğu gizli değişkenlere sahip gizli değişken modelleri döndürmektedir. Buradaki her model, gizli değişkenler içeren verilerin üretken bir sürecini ifade etmektedir. Bu süreç gizli ve gözlenen rastgele değişkenlerin ortak olasılık dağılımını belirler. Nonparametrik Bayesyen modellerin diğer Bayesyen modellerden farkı, bu gizli yapının verilerle birlikte büyüdüğünü varsaymasıdır. Model karmaşıklığı, karışım bileşenlerin sayısı son dağılımın bir parçasıdır ve önceden belirtilmesi gerekmez, veri analizinin bir parçası olarak belirlenmektedir (Gershman Blei 2012: 2).

Nonparametrik modeller, model boyutunun veri boyutuyla büyümesine izin veren model seçimine bir yaklaşım oluşturmaktadır. Bu durumda, model seçimi, yani uygun bir karmaşıklık düzeyinde bir modelin nasıl seçileceği ile ilgili sorunu ortaya çıkarmaktadır. Model seçim metrikleri genellikle iki terim içerir; eksik uyum ve aşırı uyum. Modelin gereğinden az sayıda değişken içermesi eksik uyum olarak adlandırılır ve regresyon tahmincilerinin sapmalı olmasına neden olabilir. Diğer taraftan modelin gereğinden fazla değişken içermesi durumu aşırı uyum olarak adlandırılır ve hata terimlerinin varyansının artmasına neden olabilir. Buradaki strateji sapma kare ve varyans toplamını en aza indirmeye çalışmaktır. Model karmaşıklığına bağlı olarak sapmayı azaltmak genellikle varyansı artırdığından ve varyansı azaltmak sapmayı artıracığından toplamalarının en aza indirilmesi bir denge gerektirecektir (Çağlayan Akay, Kangallı Uyar, 2017).

Ceza fonksiyonu kullanılması gibi varyans sapma dengesini sağlamanın farklı yolları vardır ancak model karmaşıklığını ayarlamak en yaygın kullanılan yöntemdir. Bu yöntem mevcut eğitim verilerinin miktarına göre değişen bir model seçimine yol açar, yani daha fazla veri için daha karmaşık model kullanılır. Bu şekilde, eğitim setinin boyutu sonsuza giderken elde edilen performansın optimuma yaklaşacağı garanti edilebilir. Böylece artan model karmaşıklığı ile sapma azalacağından, artan veri miktarları nedeniyle varyans da azalacaktır (model karmaşıklığındaki artışın yeterince yavaş olması şartıyla). Pratik kurallar bazen belirli bir boyuttaki eğitim seti ile bir modelin ne kadar karmaşık olması gerektiğine karar vermek için kullanılır. Bayesyen yaklaşım açısından bakıldığında, model karmaşıklığını eğitim verilerinin miktarına göre ayarlamanın bir önemi yoktur. Bayesçi bir model tanımlar, bir ilk dağılım seçer, veri toplar, son dağılımı hesaplar ve sonra tahminlerde bulunur. Bayesyen yaklaşım açısından, ne kadar veri toplandığına bağlı olarak modeli veya ilk dağılımı değiştirmek için herhangi bir varsayıma ihtiyaç duyulmamaktadır. Model ve ilk dağılım bin gözlem için doğruysa, on

gözlem için de doğru olduğu varsayılır (sadece yanlış bir ilk dağılım kullanılması daha az gözlemle daha önemli etkiye neden olabilir) (Neal 1995: 10).

Diğer taraftan nonparametrik Bayesyen modeller model seçimi sorununa farklı bir yaklaşım sunmaktadır. Nonparametrik Bayesyen yaklaşım farklı karmaşıklığa sahip modelleri karşılaştırmak yerine, karmaşıklığı verilere uyarlayarak tek bir modele sığdırmaktadır. Yani, verilerin modelin karmaşıklığını belirlemesine izin vererek bu sorunu ortadan kaldırmaktadır³. Ayrıca nonparametrik Bayesyen modeller ek veri gözlemlendikçe karmaşıklığın artmasına izin vermektedir. Örneğin, kümelemeye yönelik klasik modelleme yaklaşımı veri analizi öncesinde küme sayısının belirtilmesini gerektirmektedir. Nonparametrik Bayesyen yaklaşım, gözlemlenen verileri modellemek için kaç kümeye ihtiyaç olduğunu tahmin eder ve gelecekteki verilerin daha önce görülmemiş kümelemelerinin oluşmasına izin vermektedir (Gershman, Blei 2012: 3). Ayrıca nonparametrik Bayesyen yaklaşım nonparametrik modelleme ve model seçimine bir alternatif sunmaktadır. Sınırsız karmaşıklığa sahip bir model kullanarak eksik uyum sorunu hafifletilirken, Bayesci yaklaşım hesaplaması veya parametreler üzerinden tüm son dağılıma yakınsama yaklaşımı aşırı uyum sorununu hafifletmektedir (Teh 2011: 2).

Sonuç olarak, doğru karmaşıklığa sahip model seçimi istatistiksel modellemede genellikle önemli bir konudur. George Box istatistiksel modelleme hakkında kısaca şöyle söylemiştir: *“Bütün modeller yanlıştır, ancak bazıları faydalıdır”* (Keele 2008: 1). Yani sonsuz sayıda bileşenli karışım modelleri küçük veri kümelerine de uygulanabilmektedir (Neal 1995: 10).

Nonparametrik Bayesyen modeller rastgele ölçülerin sonsuz boyutlu uzayına bir ilk dağılım yerleştirdikleri için “Nonparametrik Bayes” olarak adlandırılmaktadır. Buradaki sonsuz boyutlu parametre genellikle bir fonksiyon veya ölçüdür. Bilinmeyen miktarın veya parametrelerin sonsuz boyutlu olması bu durumu ifade etmek için kullanılmaktadır (Ghosal, van der Vaart 2017: 2). Nonparametrik Bayesyen modellerde parametre uzayı olası tüm çözümlerin kümesi olarak seçilmektedir (Orbanz, Teh 2010: 2).

Parametrik modeller, veri üretme mekanizması hakkında kısıtlayıcı varsayımlar yapar ve bu da çıkarımda ciddi sapmaya neden olabilir. Bayesçi çerçevede, parametrik bir model varsayımı, son derece güçlü bir ilk dağılım fikri olarak görülebilir. Bu nedenle parametrik Bayesyen modelleme, olasılığı tüm yoğunlukların zayıf bir alt kümesine atayan bir ilk dağılımda ısrar etmeye eşdeğerdir (Ghosal, van der Vaart 2017: 3). Bu durumda dağılım uzayında tam desteğe sahip bir ilk dağılım (önceki inançları özetleyen ve tüm belirsizlikleri barındıracak kadar büyük olması) gerekmektedir (Hjort vd. 2010: 24). Yani geleneksel parametrik dağılım ailesi çok daha geniş bir aile ile değiştirilmektedir. İlk dağılımın

³ Karmaşıklık sorunu, Dirichlet süreç (Dirichlet Process, DP) ilk dağılımından seçilen rastgele dağılım fonksiyonlarının kesikli kütlelerin sayılabilir sonsuzluğu ile ilgilidir (Walker 2007:45).

desteği y_i 'nin dağılımının tutarlı bir şekilde tahmin edilmesine izin vermeli ve yeterince büyük olmalı ki y_i dağılımı göz önüne alındığında $\theta|y_i$ koşullu dağılımı hakkındaki belirsizliği doğru şekilde yansıtsın. Bu da sonsuz boyutlu parametrik olmayan bir ilk dağılım kullanılmasını gerektirmektedir. Bu nedenle ilk dağılım için tam destek nonparametrik Bayes yaklaşımda önemli motivasyonlardan biridir (MacEachern 2016: 446).

Bu modellerin tipik kullanımları, yerel özellikleri yakalayan nonparametrik bir yapı ile verilerin global özelliklerini yakalayan nispeten düşük boyutlu çıkarım terimlerini içermesidir. Modeller, yerel davranışın yaygın bir şekli olan aykırı değerleri ele almada özellikle iyidir ve son dağılımın incelenmesi genellikle modelin bir bölümünün aykırı değerleri takip ettiğini gösterir. Bu durum, aykırı değerlerin genel analiz üzerindeki etkisini azaltmak için robust çıkarıma olan ihtiyacı göstermektedir. Nonparametrik Bayesyen modeller aykırı değerlere karşı robust olan ilgili parametreleri tanımlamak için çıkarım fonksiyonlarının kullanılmaktadır. Bu durumda, çıkarım fonksiyonu ilgilenilen parametreyi tanımlamaktadır ve kullanılan yöntemler bu parametrenin dağılımının incelenmesine izin vermektedir. Bu modeller verilerin iyi ve kötü bileşenlerini (aykırı değerleri) modellemeye veya belirli bir çıkarımı hedefleyen nispeten esnek olmayan bir model kullanmaya çalışmak yerine, problemi yoğunluk tahmini olarak yeniden biçimlendirilir. Temel fikir, aykırı değerleri modelin farklı bir bileşeniyle modellemek yerine, verilerdeki nerdeyse her örüntüyü yakalayabilen esnek bir ilk dağılım kullanmaktır. Biçimsel olarak, bu yöntemler, dağılım fonksiyonunun formu hakkında tam bir belirsizliği ifade etmek için dağılım fonksiyonları uzayına bir ilk dağılım yerleştirir. Yoğunluk tahmini için kullanılan modeller yüksek veya sonsuz boyutludur ve veriler için geniş örüntüye uyabilir⁴. Sonuç olarak, bir kuyruk alanında aykırı veriler gözlemlendiğinde (bazı durumlar veri yığınının dışındadır), son dağılım onun kütlelerinin çoğunu, uzaktaki verilerin yakınında bir tümsek (veya tümsekler) olan dağılımlara atar. Son dağılımdan elde edilen yoğunluk tahmin sonuçları hem iyi verilerin genel modelini hem de aykırı değerlerin bu modelden sapmasını gösterir. Nonparametrik Bayes modellerinin önemli bir özelliği, tamamen esnek yüksek boyutlu bir modelin kullanılması ile aykırı değerler için bir bileşen içeren bir modelin yazılması arasında bir hibrit oluşturmastır (Lee, MacEachern 2014: 477- 486).

Tüm Bayes yöntemlerinin altında yatan temel varsayım, modeli belirleyen parametrelerin rastgele bir değişken olmasıdır. Rastgele dağıtılmış bir parametrenin var olduğuna dair Bayes varsayımı, bir varsayım değil, veri özelliklerinin matematiksel bir sonucudur. Bu bağlamda sonsuz boyutlu

⁴ Bu modeller, nonparametrik Bayes modeller başlığı altına girer. Yoğunluk tahmini için popüler modeller, Dirichlet sürecine dayanan Dirichlet karışım (Dirichlet Process Mixture, DPM) modelleridir. Bu yöntemler, herhangi bir dağılımın tutarlı bir şekilde tahmin edilmesine izin verecek şekilde tasarlanmıştır.

parametreye sahip nonparametrik Bayesyen modellerin inşasının temelini oluşturan değiştirilebilirlik kavramı açıklanmaktadır.

Aynı olasılık uzayında (Y, Ω) bir değişken dizisi $Y_1, Y_2, \dots, Y_n$ değişkenlerin sırası değiştiğinde bu değişkenlerin ortak dağılımları değişmez ise değiştirilebilir olarak tanımlanmaktadır. Yani, P ortak dağılım ve σ herhangi bir $\{1, n\}$ permütasyonu ise, o zaman:

$$P(Y_1 = y_1, Y_2 = y_2, \dots, Y_n = y_n) = P(Y_1 = y_{\sigma(1)}, Y_2 = y_{\sigma(2)}, \dots, Y_n = y_{\sigma(n)}) \quad (3.1)$$

$Y_1, Y_2, \dots, Y_n$ her $n \geq 1$ için değiştirilebilirse, sonsuz bir dizi $Y_1, Y_2, \dots$ sonsuz değiştirilebilirdir. Değiştirilebilirlik, değişkenlerin kendi aralarında bağımlı olsalar da endekslerine bağlı olmadıkları varsayımını ifade etmektedir⁵. Eğer θ , temeldeki dağılımı parametreleştiriyorsa ve θ üzerinde bir ilk dağılım varsayılırsa, o zaman $Y_1, Y_2, \dots$ üzerinde θ ile marjinalleştirilmiş olarak ortaya çıkan marjinal dağılım yine değiştirilebilir olacaktır. Ünlü De Finetti teoremine atıfta bulunulan temel bir sonuç; eğer $Y_1, Y_2, \dots$ sonsuz değiştirilebilir ise o zaman her $n \geq 1$ için rastgele bir θ parametresi vardır:

$$P(Y_1, Y_2, \dots, Y_n) = \int P(\theta) \prod_{i=1}^n P(Y_i | \theta) d\theta \quad (3.2)$$

Değiştirilebilirlik varsayımı temel olarak θ rastgele gizli parametresinin ve hiyerarşik bir Bayes modelinin varlığını göstermektedir. De Finetti'nin Teoreminde, θ 'nın sonsuz boyutlu olabileceğini vurgulamak önemlidir (bu rastgele bir ölçüdür), bu nedenle hiyerarşik Bayes modeli (3.2) tipik olarak nonparametrik bir model olur (Orbanz, Teh 2010: 6). De Finetti Teoremi değiştirilebilirlik varsayımı ile birlikte nonparametrik Bayesyen yaklaşımdaki sonsuz boyutlu rastgele parametrenin varlığını ispatlamaktadır.

Sonlu boyutlarda, bir olasılık modeli genellikle bir yoğunluk fonksiyonu veya olasılık kütle fonksiyonu ile tanımlanmaktadır. Nonparametrik Bayesyen modellerde, parametre uzayı sonsuz boyuta sahip bir Bayes modelini ifade etmektedir. Yani nonparametrik Bayes modelini tanımlamak için öncelikle sonsuz boyutlu bir uzayda bir olasılık dağılım (ilk dağılım) tanımlamak gerekmektedir. Sonsuz boyutlu bir parametre uzayı üzerindeki bir dağılım, yolları parametre uzayında olan stokastik bir süreç olduğundan nonparametrik Bayesyen modelleri tanımlamak için stokastik süreçlere ihtiyaç duyulmaktadır. Nonparametrik Bayesyen modeller Dirichlet süreci gibi stokastik süreçler içermektedir ve genellikle içerdikleri stokastik süreçlere göre adlandırılmaktadır (Orbanz 2014: 3). Bu bağlamda ilk olarak nonparametrik Bayesyen modellerin inşasında önemli bir yere sahip olan Dirichlet süreçleri ve bu süreçlere ilişkin modeller açıklanmaktadır.

⁵ Değiştirilebilirlik, i.i.d.'den çok daha zayıf bir varsayımdır; i.i.d. olan değişkenler otomatik olarak değiştirilebilir.

3.1. Dirichlet Süreçleri

Nonparametrik Bayesyen modeller için iki ana yaklaşım vardır. Birinci yaklaşım, regresyon fonksiyonu gibi bir fonksiyona odaklanır ve geleneksel lineer regresyonu çok daha esnek bir regresyonla değiştirir. İkinci yaklaşım ise olasılık dağılımlarına odaklanmaktadır. Geleneksel parametrik dağılım ailesini çok daha geniş bir aile ile değiştirir. Parametrik bir aileye bir ilk dağılımı yerleştirmek yerine, yani dağılımı yöneten parametreler için bir olasılık modeli yazmak yerine, dağılımların kendileri için bir olasılık modeli yazar. Burada ikinci yaklaşıma ait nonparametrik Bayesyen modeller takip edilmektedir (MacEachern 2016: 448).

Bayesyen yaklaşım istatistiksel problemlerde birçok yönden verimli olmasına rağmen, ilk zamanlarda nonparametrik problemlerin çözümünde başarısız olmuştur. Bunun temel nedeni, nonparametrik problemlerde belirli bir örnek uzay verildiğinde bir dizi olasılık dağılımı olarak ele alınan parametre uzayı üzerinde uygulanabilir ilk dağılımların bulunamamasıdır. Nonparametrik Bayesyen yaklaşımdaki büyük ilerleme Ferguson'un (1973) Dirichlet sürecini (Dirichlet Process, DP) tanımlayan makalesi tarafından sağlanmıştır.

Ferguson (1973), nonparametrik Bayesyen yöntemler üzerine ilk makalesinde dağılım fonksiyonları uzayına bir dağılım yerleştirmeyi sağlayan Dirichlet sürecini tanımlamaktadır. Sınırlandırılmış bir parametrik aile yerine dağılımlar için geniş uzayda çalışmanın faydaları açıktır; çarpıklık, aykırı değerler, kalın kuyruklar (tümsekler) ve iki modluluk gibi normallikten sapmaları modellemeyi sağlamaktadır (MacEachern 1999, 2000; Richardson, Hartman 2018: 2).

Ferguson (1973)'e göre nonparametrik problemler için bir ilk dağılımın sahip olması gereken iki özelliği vardır:

- 1- İlk dağılımın desteği, örnek uzaydaki olasılık dağılımları uzayındaki bazı uygun topolojiye göre büyük olmalıdır.
- 2- Gerçek olasılık dağılımından bir gözlem örneği verildiğinde son dağılımlar analitik olarak izlenilebilir olmalıdır (Ferguson 1973: 209).

Ferguson (1973)'de Dirichlet sürecini, analitik olarak izlenebilen son dağılım ve geniş bir desteğe sahip dağılımlar üzerinde bir ilk dağılım olarak resmîleştirmiştir. Dirichlet süreci, nonparametrik Bayesyen modellerde merkezi bir rol oynamaktadır.

Rastgele fonksiyonlar, olasılık ölçüleri ve daha genel olarak sonsuz boyutlu rastgele nesneler üzerindeki olasılık dağılımları, stokastik süreçler olarak adlandırılmaktadır (Orbanz, Teh 2010: 2). Dirichlet süreci, nonparametrik Bayesyen modellerinde, özellikle Dirichlet süreç karışım modellerinde (sonsuz karışım modelleri) kullanılan stokastik bir süreçtir. Dağılımlar üzerine bir dağılımı ifade eder, yani bir Dirichlet sürecinden her çekimin kendisi bir dağılımdır. Dirichlet sürecinden çekilen dağılımlar kesiklidir ve sınırlı sayıda parametre kullanılarak tanımlanamaz, bu nedenle sınıflandırma nonparametrik bir model olarak sınıflandırılır (Teh 2011: 1).

Dirichlet sürecini tanımlamak için öncelikle Dirichlet dağılımının özelliklerin açıklanması gerekmektedir. Burada genel bir Dirichlet dağılımı tanımı verilmektedir. Bayesyen yaklaşımda Dirichlet dağılımı, çok değişkenli dağılımın parametreleri için bir eşlenik ilk dağılım olarak bilinmektedir. Daha açık ifade ile Dirichlet dağılımı, Beta dağılımının çok değişkenli bir genellemesidir. En iyi bilinen eşlenik dağılım örneği genellikle yazı tura dizilerini modellemek için kullanılan Beta-Binom dağılımıdır (beta-binom model):

$$\begin{aligned} p &\sim \text{Beta}(\alpha, \beta) \\ Y_i | p &\sim \text{Bernoulli}(p) \end{aligned} \quad (3.3)$$

Binom dağılımı, bir dizi Bernoulli denemesinde başarı olasılığını gösterirken, Beta dağılımı, her deneme için başarı olasılığının ilk olasılık dağılımını temsil etmektedir ve böylece bağlantı oluşturulmaktadır. Çok değişkenli Dirichlet model ise, beta-binom modelini gözlemlenebilir veriler için daha fazla kategoriye genişletmektedir:

$$\begin{aligned} p &= (p_1, \dots, p_k) \sim \text{Dirichlet}(\alpha_1, \dots, \alpha_k) \\ Y_i | p &\sim \text{multinomial}(p) \end{aligned} \quad (3.4)$$

Dirichlet- Multinomial dağılımının daha açık bir ifadesi; Beta dağılımı tek bir olasılık (p) olasılığını modellerken, Dirichlet, kesinlik parametresi olarak adlandırılan ve her seçim için ağırlıkları temsil eden α parametresiyle, birbirini dışlayan çoklu seçimlerin olasılıklarını modellemektedir (MacEachern 2016: 448-450).

Ferguson (1973), Dirichlet dağılımını kullanarak olasılık ölçüleri uzayı üzerinde bir ilk dağılım olarak Dirichlet süreç modeli tanımlamıştır. Aslında burada sonlu boyutlu parametre uzaylarının kullanımına bir alternatif olarak, sonsuz boyutlu bir G uzayına bir olasılık dağılımı yerleştirilmektedir. $G, \mathcal{X}$ 'in alt kümelerinin bir σ -alanı olan $\mathcal{B}$ indeks kümesine sahip bir stokastik süreç olarak tanımlansın. Yani G bir Borel uzayında $(\mathcal{X}, \mathcal{B})$ rastgele bir olasılık ölçüsü olsun. G için bir olasılık dağılımı yazmak için tüm n ve $B_1, \dots, B_n \in \mathcal{B}$ için $(G(B_1), \dots, G(B_n))$ ortak dağılımının belirlenmesi gerekmektedir. α , $(\mathcal{X}, \mathcal{B})$ üzerinde sıfırdan farklı negatif olmayan sonlu bir ölçü olsun. Her $n > 0$ tamsayı ve $\mathcal{X}$ 'in

ölçülebilir $B_1, \dots, B_n$ için, $(G(B_1), \dots, G(B_n))$ 'nin ortak dağılımı α parametresi ile $Dir(\alpha_1, \dots, \alpha_n)$ Dirichlet dağılımına sahiptir denilir ve şu şekilde ifade edilmektedir (Schervish 1995: 52):

$$(G(B_1), \dots, G(B_n)) \sim Dir(\alpha_1, \dots, \alpha_n) \quad (3.5)$$

$\mathcal{X}$ 'in ölçülebilir herhangi bir parçası için $(G(B_1), \dots, G(B_n))$ rastgele vektörü $(\alpha G_0(B_1), \dots, \alpha G_0(B_n))$ parametreleri ile Dirichlet dağılımına sahipse, $\mathcal{B}$ 'nin B elemanları tarafından indekslenen bir stokastik süreç G için, $(\mathcal{X}, \mathcal{B})$ üzerinde $\alpha > 0$ ve G_0 parametrelerine sahip bir Dirichlet süreci olduğu söylenir. Yani (α, G_0) parametrelerine sahip Dirichlet süreci, $(\mathcal{X}, \mathcal{B})$ 'nin her (ölçülebilir) sonlu bölümü $(B_1, \dots, B_n)$ için her (ölçülebilir) B kümesine $G(B)$ olasılığını atayan $(\mathcal{X}, \mathcal{B})$ ölçülebilir uzay üzerinde tanımlanan rastgele olasılık ölçüsü (dağılım) G olarak tanımlanır (Müller vd. 2015: 8).

Dirichlet süreci, rastgele olasılık dağılımları için bir olasılık modelidir; G_0 temel dağılımı ve kesinlik parametresi α olmak üzere iki parametreyle indekslenir. G_0 temel dağılımı ve kesinlik parametresi α ile dağılan bir Dirichlet süreci $DP(\alpha, G_0)$, şu şekilde gösterilmektedir:

$$(G(B_1), \dots, G(B_n)) \sim Dir(\alpha G_0(B_1), \dots, \alpha G_0(B_k)) \quad (3.6)$$

Temel dağılım G_0 , Dirichlet sürecinin ortalamasıdır: herhangi bir ölçülebilir küme $B \subset (\mathcal{X}, \mathcal{B})$ için

$$E[G(B)] = G_0(B) \quad (3.7)$$

olur. G_0 , bir dağılım fonksiyonu ve G deki tahmini bir ilk dağılım olarak düşünülmektedir. Kesinlik parametresi α , ters varyans olarak düşünülebilir:

$$Var[G(B)] = \frac{G_0(B)[1 - G_0(B)]}{(1 + \alpha)} \quad (3.8)$$

α kesinlik parametresi ilk dağılımın G_0 civarında ne kadar yoğunlaştığını kontrol ederek G 'in belirsizliğini belirlemektedir. α ne kadar büyükse, varyans o kadar küçüktür ve G , G_0 civarında yoğunlaşır; kütesinin çoğunu ortalama etrafında yoğunlaştıracaktır. α , Dirichlet sürecinin ortalaması etrafındaki kütle konsantrasyonunu olarak tanımlandığı için $\alpha \rightarrow \infty$ yaklaştıkça süreç herhangi bir ölçülebilir B kümesi için $G(B) \rightarrow G_0(B)$ olacaktır. Kesinlik parametresi, bir nonparametrik Bayesyen modelde Dirichlet sürecini nonparametrik bir ilk dağılım olarak kullanırken ilk dağılımın gücüne ve bu güç örnek boyutunun birimleri ile ölçülebildiği için kütle parametresine atıfta bulunarak güç parametresi olarak da adlandırılmaktadır⁶(Teh 2011: 3).

⁶ Dirichlet sürecini parametreleştiren sonlu pozitif ölçü α , genellikle toplam kütle M ve G_0 dağılım fonksiyonuna bağlıdır. Böylece kesinlik parametresi α ; $M > 0$ o.ü $\alpha((-\infty, x]) = MG_0$ şeklinde tanımlanır ve toplam kütle parametresi olarak da adlandırılabilir (Maceachern 2007:486).

Dirichlet süreçleri dağılımlar üzerine bir dağılımdır, yani bir Dirichlet sürecinden her bir çekimin kendisi (G) yine bir dağılımdır. Dirichlet süreçlerinin önemli bir özelliği; Dirichlet sürecinden çekilen bir dağılım, G_0 temel dağılım düzgünleştirilmiş olsa bile her zaman kesikli dağılım olacaktır. Dirichlet sürecinden çekilen G kesikli dağılımı $G \sim DP(\alpha, G_0)$ şeklinde ifade edilmektedir. Böylece nonparametrik Bayesyen modelde sonsuz boyutlu bilinmeyen G parametresi üzerine bir ilk dağılım modelinin tanımlayan $G \sim DP(\alpha, G_0)$ Dirichlet süreç ilk dağılımı tanımlanmıştır (Müller vd. 2015: 8).

Dirichlet sürecinden rastgele bir G dağılımı çektiğimizi ve bağımsız olarak G'den N rastgele değişken η_n çizdiğimizi varsayalım:

$$G | G_0, \alpha \sim DP(G_0, \alpha) \quad (3.9)$$

$$\eta_n \sim G, \quad n \in \{1, \dots, N\}.$$

Farklı η_n 'lerin aynı değerleri aldığı konumlara pozitif olasılık atanır ve temeldeki rastgele G dağılımı, bir olasılıkla kesiklidir. Bu durum G'nin sonsuz atomik ölçüler toplamı olarak inşa edildiği Dirichlet sürecinin çubuk-kırma temsiliinde görülmektedir (Blei, Jordan 2006: 122).

Ferguson (1973) temel teoremi: G , G_0 temel dağılım ve α parametresi ile bir Dirichlet süreci ise ve $y_1, \dots, y_n$, G'den çekilen bir örnek ise, $y_1, \dots, y_n$, verildiğinde G'nin son dağılımı da $\alpha G_0 + \sum_{i=1}^n \delta_{y_i}$ ile bir Dirichlet sürecidir. Burada δ_y , y noktasına verilen ağırlık ölçüsünü ifade etmektedir (Ferguson 1973: 210):

$$G | y_1, \dots, y_n \sim DP(\alpha G_0 + \sum_{i=1}^n \delta_{y_i}) \quad (3.10)$$

DP model (3.9) altındaki son dağılım DP açık olarak yeniden yazılırsa;

$$G | y_1, \dots, y_n \sim DP(\alpha + n, \frac{\alpha}{\alpha + n} G_0 + \frac{n}{\alpha + n} \frac{\sum_{i=1}^n \delta_{y_i}}{n}) \quad (3.11)$$

Son dağılımın, temel dağılım G_0 ve ampirik dağılım $\sum_{i=1}^n \delta_{y_i}/n$ arasında ağırlıklı bir ortalama olduğu görülmektedir. Temel dağılım ile ilişkili ağırlık α ile orantılı iken ampirik dağılım, n gözlem sayısı ile orantılı ağırlığa sahiptir. Bu nedenle α parametresi ilk dağılımla ilgili kütle olarak yorumlanmaktadır (Teh 2011: 5).

Ferguson (1973) temel teoremi Dirichlet sürecinin, (i.i.d.) örnekleme göre eşlenik olduğunu; yani, G üzerinde bir Dirichlet süreç ile örnekleme modeli (3.9) altında, G için son dağılım yine bir Dirichlet

süreci olmaktadır. Bu ilerleme, Bayes yöntemiyle kümülatif bir dağılımın tahmin edilmesi sorununu çözmektedir (Müller vd. 2015: 10).

Dirichlet sürecinden yapılan çekimleri, kesikli bir rastgele olasılık ölçüsü olarak, G , her zaman nokta kütlelerinin ağırlıklı toplamı olarak yazılmaktadır. Sethuraman (1994), Dirichlet sürecinin bu şekilde, çubuk-kırma (Stick-Breaking) temsili olarak adlandırılan yapıcı bir tanımını sağlayarak bunu kesinleştirmiştir. $DP(\alpha, G_0)$ 'dan sonsuz boyutlu rastgele bir G dağılımı çekmek için çubuk-kırma temsili şu şekilde ifade edilir:

$$G(\bullet) = \sum_{k=1}^{\infty} w_k \delta_{\theta_k}(\bullet) \quad (3.12)$$

Burada θ_k konumları temel dağılım G_0 'dan (i.i.d.) çekimlerdir ($\theta_k \sim G_0$). Ağırlıklar $w_k = \sum_{k=1}^{\infty} w_k$ ve $v_1, v_2, \dots \sim \text{Beta}(1, \alpha)$ olmak üzere $w_k = v_k \prod_{l=1}^{k-1} (1 - v_l)$ şeklinde oluşturulur. Çubuk kırma yapısı için birim uzunlukta bir çubuk olduğu varsayılır. Daha sonra çubuğun rastgele bir v_1 noktasında kırılmasıyla $w_1 = v_1$ ağırlığı elde edilir ve çubuk artık $(1 - v_1)$ uzunluğuna sahiptir. İkinci ağırlık kalan çubuktan bir v_2 oranının kırılmasıyla $w_2 = v_2(1 - v_1)$ olarak elde edilir. İşlem bu şekilde devam eder ve bütün ağırlıklar dizisi $w_1, w_2, \dots$ oluşturulur. Çubuk-kırma süreci Dirichlet sürecinden $w_1, w_2, \dots$ ağırlıklarının nasıl inşa edildiğini göstermektedir. w üzerindeki çubuk-kırma inşası bazen $w \sim GEM(\alpha)$ şeklinde gösterilmektedir (Teh 2011: 8). Sethuraman (1994) inşası θ_k konumlarının G_0 dağılımından (i.i.d.) örnekler olduğunu ve w_k ağırlıkların ise yeniden ölçeklendirilmiş bir Beta dağılımı, $v_k \sim \text{Beta}(1, \alpha)$ olduğunu gösterir. Dirichlet sürecinin en yaygın uygulaması, verilerin karışım modelleri kullanılarak kümelenmesidir. Dirichlet sürecinin nonparametrik yapısı sayesinde, bu modeller sonsuz sayıda bileşen içeren karışım modellerine dönüşmektedir.

3.2. Dirichlet Süreç Karışım Modelleri (DPM)

Dirichlet süreci, bir olasılıkla kesikli dağılımlar üretmektedir ve bu durum sürekli yoğunluk tahmini sınırlamaktadır. Bu sınırlama, karıştırma dağılımı olarak Dirichlet sürecini kullanan karışım modeli inşa edilerek kaldırılmaktadır (Müller vd. 2015: 11). Buradaki temel düşünce Dirichlet sürecinden çekilen karışım parametrelerini kernel (çekirdek) yardımı ile düzgünleştirmektir (Teh 2011: 9). Nonparametrik Bayesyen yaklaşım ile sürekli yoğunluk tahmini Dirichlet süreçlerinin kesikli yapısı otomatik olarak gizli değişkenler içeren karışım modellerini ortaya çıkarmaktadır.

Karışım modelleri en temel gizli değişken modellerinden biridir. Bir karışım modeli, verilerin kümelenildiğini ve her veri noktasının atandığı kümeyle ilişkili bir dağılımdan çekildiğini varsayar.

Modelin gizli deęişkenleri, küme atamaları ve her bir küme dağılımına ilişkin parametrelerdir. Veriler gözlemlendiğinde karışım modelinin son dağılımı, kümelemeler ve parametreler üzerinde koşullu bir dağılımdır. Bu koşul, verilerin olası bir gruplamasını ve her grubun özelliklerini tanımlar. Burada, karışım modellerinin çerçevesini göstermek için Gauss karışım model bileşenleri ele alınmaktadır. Daha genel olarak, karışım modelleri diğer dağılımların doğrusal kombinasyonlarını içerebilir. Gauss dağılımı gibi daha temel dağılımların lineer kombinasyonları alınarak oluşturulan bu tür karışımlar, karışım dağılımları olarak bilinen olasılık modelleri olarak formüle edilebilir. Temel olarak K sayıda Gauss yoğunluklarının bir karışımı olarak karışım modeli şu şekilde ifade edilmektedir:

$$p(y) = \sum_{k=1}^K w_k N(y | \mu_k, \Sigma_k) \quad (3.13)$$

Bu model Gauss karışım model olarak adlandırılmaktadır. Her bir Gauss yoğunluğu $N(y|\mu_k, \Sigma_k)$ karışım bileşeni olarak adlandırılır ve kendi μ_k ortalama ile Σ_k kovaryansına sahiptir. w_k parametreleri karıştırma katsayısı (karışım oranı) olarak adlandırılır. Bu katsayılar $\sum_{k=1}^K w_k = 1$, tüm k için $w_k \geq 0$ ve $0 \leq w_k \leq 1$ özelliklerini sağlar. Böylece, karıştırma katsayılarının olasılık olma gereksinimlerini karşıladığını görülmektedir (Bishop 2006: 11). Bu modeldeki gizli deęişkenler her bir veri noktası için karışım bileşenleri ve karışım oranıdır. Veriler gözlemlendiğinde, Gauss karışım modeli ile karışım bileşenlerinin, karışım oranlarını ve verilerin kümelenme şeklinin son dağılımı tahmin edilerek analiz edilir. Bu son dağılım, gözlemlenen verideki gizli yapıyı ortaya çıkarır; veri noktalarını K gruplara ayırarak kümeler ve her grubun konumunu (ortalamasını) tanımlar (Blei 2014: 207).

Karışım modelleri, veri setinin grup yapısını anlamak ve anakütle dağılımını esnek bir şekilde tahmin etmek için kullanılır. Karışım modelleri veri üreten dağılımı deęiştirerek birçok veri türünü barındırabilir. Bayes karışım modelleri ise ayrıca karıştırma dağılımı ve kümeleme parametreleri üzerine bir ilk dağılım varlığını gerektirmektedir. Veriler sonlu karışım model ile analiz edildiğinde gizli kümelerin sayısının önceden belirlenmesi gerekmektedir. Ancak birçok veri analizi için bu sayı bilinmemektedir. Nonparametrik Bayes karışım modelleri sonsuz sayıda gizli küme olduğunu ve bunların sonlu sayıda gözlemlenen verileri oluşturmak için kullanıldığını varsayarak bu sorunu çözmektedir. Bu varsayımlar altında, son dağılım, küme sayısı, verilerin kümelere atanması ve her kümeyle ilişkili parametreler üzerinde bir dağılım sağlamaktadır (Gershman, Blei 2012: 3).

Böylece nonparametrik Bayesyen karışım modelleri model tabanlı bir kümeleme analizi sağlar. Model tabanlı kümeleme analizi verilerin farklı alt popülasyonlardan oluşan bir popülasyondan geldiğini varsayar. Model tabanlı kümeleme analizinde alt popülasyonların her biri ayrı modellenir ve sonlu karışım modelleri kullanılarak bu alt popülasyonların bir karışımı olarak modellenir (Mun vd. 2008; Raftery, Dean 2006). Bu karışım modelleri konuların farklı homojen gruplar halinde model tabanlı kümelenmesini sağladıkları ve verilerin esnek modellemesini sağlamaları açısından önemlidir.

Sonlu karışım modeli, verilerin sonlu sayıda dağılımın karışımından geldiğini varsayar. Bu karışım modelleri $f(y) = \sum_{k=1}^K w_k f(y|\theta_k)$ formundaki y verisi üzerinde bir yoğunluk fonksiyonu tanımlamaktadır; burada w_k karışım oranı ve θ_k , k bileşen ile ilişkili parametrelerdir. Sonlu boyutlu parametre uzayı Θ üzerinde tanımlı bir olasılık dağılımı G verildiğinde, G 'ye göre bu yoğunluk nonparametrik şekilde integral olarak şu şekilde ifade edilebilir:

$$f(y) = \int f_{\theta}(y)G(\theta)d\theta \quad (3.14)$$

Burada $G = \sum_{k=1}^K w_k \delta_{\theta_k}$, karışım modelinin tüm parametrelerini içeren kesikli bir karıştırma dağılımıdır. Nonparametrik Bayesyen karışım modeller bunun yerine sayılabilir sonsuz sayıda atomdan oluşan karışım dağılımlarını kullanmaktadır; $G = \sum_{k=1}^{\infty} w_k \delta_{\theta_k}$. Bu da sonsuz sayıda bileşene sahip karışım modellerin ortaya çıkmasına yol açmaktadır⁷. Bayesyen yaklaşım, karıştırma dağılımı üzerinde bir ilk dağılıma ihtiyaç duymaktadır. En yaygın kullanılan ilk dağılım Dirichlet süreçleri olduğu için elde edilen karışım modeli Dirichlet karışım modeli (Dirichlet Process Mixture, DPM) olarak adlandırılmaktadır (Orbanz, Teh 2010: 3).

Kernel yoğunluk fonksiyonu f_{θ} ile dönüşüm, sürekli bir dağılım tanımlamak için G 'nin kesikli nokta kütlelerinin düzgünleştirilmesini sağlar. Böylece Dirichlet süreç karışım modelleri sürekli yoğunluk tahminine yönelik bir çözüm sunmuştur (Müller vd. 2015: 13).

Dirichlet ilk dağılımı ile DPM modeller, niteliksel olarak kernel yöntemine benzer, ancak farklı σ_i^2 varyanslarının kullanılması yardımıyla örnek uzay boyunca farklı derecelerde düzgünleştirme sağlayan tahmin dağılımları üretmektedir. Böylece DPM modeller global düzgünleştirmeye (σ^2) karşı yerel düzgünleştirme (σ_i^2) ile belirsizliği değerlendirir⁸. Bu yapı son dağılımda, yakın olan y_i ve y_j veri noktaları için $\pi_i = (\mu_i, \sigma_i^2)$ ve π_j bireysel parametrelerinin ortak değerlerini güçlü bir şekilde destekleyecek şekildedir, böylece yerel yapıyı tahmin etmek için bilgiyi örnek uzayda yerel olarak birleştirir (Escobar, West 1995: 577).

Burada amaç belirli bir gözlem kümesinin çekildiği sürekli yoğunluğu modellemektir. Yani bir dizi gözlem $\{y_1, \dots, y_n\}$ bir dizi gizli parametre $\{\theta_1, \dots, \theta_n\}$ kullanarak modellenir. Her bir θ_i , G den bağımsız ve aynı şekilde çekilir ve her y_i , θ_i ile parametrelendirilmiş f_{θ_i} düzgünleştirilmiş yoğunluğa (normal

⁷ k sayıları sonsuzdur ve bu nedenle model, sonsuz sayıda kümeye bölünmüş bir anakütleyi temsil eder, ancak, n sonlu bir örnek için bu kümelerin en fazla n tanesi gözlemlenebilir.

⁸ Yerel düzgünleştirme derecesini kontrol eden varyans ve kesinlik parametreleri tanımlanır.

kernel, i.n.d) sahiptir. Karışım modeli (3.14), karıştırma dağılımı G üzerinde bir DP ilk dağılımı ile hiyerarşik bir model olarak eşdeğer şekilde yazılabilir:

$$\begin{aligned} y_i | \theta_i &\overset{ind}{\sim} f_{\theta_i} \\ \theta_i | G &\overset{iid}{\sim} G \\ G | \alpha, G_0 &\sim DP(\alpha, G_0) \end{aligned} \quad (3.15)$$

Hiyerarşik model, her birime özgü yeni gizli değişken θ_i 'ler ortaya çıkarır. Bu hiyerarşik model ile eşitlik (3.14)'deki modele eşdeğer olduğu görülmektedir. Dirichlet süreç karışım modelinde, veriler değil karışım parametreleri bir Dirichlet sürecinden örneklenmektedir. G kesikli olduğundan birden fazla θ_i aynı anda aynı değeri alabilir ve yukarıdaki model aynı θ_i değerine sahip y_i 'lerin aynı kümeye ait olduğu bir karışım modeli olarak görülmektedir (Teh 2011: 9).

Dirichlet süreç karışım modelleri, çubuk-kırma yapısı kullanılarak karışım modellerinin olağan temsili ile daha uyumlu hale getirilebilir. Burada önemli nokta karışım bileşenlerin bir Dirichlet sürecinden (G) çekimler olmasıdır. Böylece G 'ni yapısı gereği kesikli olması otomatik olarak θ_i 'lerin kümelemesini oluşturmaktadır. Dirichlet süreci burada karıştırma dağılımı görevi görmektedir. Dirichlet süreç karışım modelleri için çubuk-kırma yapısı aşağıdaki süreçler takip edilerek inşa edilmektedir. Öncelikle $G = \sum_{k=1}^{\infty} w_k \delta_{\theta_k^*}$ kesikli rastgele karıştırma dağılımı inşası için Dirichlet süreç çubuk-kırma yapısına benzer aşağıdaki adımlar takip edilir:

- 1- G_0 'dan θ_1^* çekilir
- 2- $Be(1, \alpha)$ dağılımından v_1 çekilir,
- 3- $w_1 = v_1$ yazılır,
- 4- G_0 'dan θ_2^* çekilir,
- 5- $Be(1, \alpha)$ dağılımından v_2 çekilir,
- 6- $w_2 = v_2(1 - v_1)$ yazılır ...

Süreç atom konumları ve ağırlıklar oluşturularak bu şekilde devam eder. k sayıları sonsuzdur ancak n sonlu bir örnek için bu kümelerin en fazla n tanesi gözlemlenebilir. Yani öncelikle temel dağılım G_0 belirlenir, sonra G_0 'dan atom konumları (θ_k^*) örneklenir ve son olarak atom ağırlıkları (w_k) Dirichlet sürecinden çekilir. z_i, w_k olasılığı ile k değerini alan bir küme atama değişkeni olsun. $G = \sum_{k=1}^{\infty} w_k \delta_{\theta_k^*}$ ve $\theta_i = \theta_{z_i}^*$ olmak üzere model (3.15) yeniden şu şekilde ifade edilebilir:

$$w | \alpha \sim GEM(\alpha)$$

$$\theta_k^* | G_0 \sim G_0$$

(3.16)

$$z_i | w \sim Mult(w)$$

$$y_i | z_i, \{\theta_k^*\} \sim f(\theta_z^*)$$

Karışım modelinde, w karıştırma oranı, θ_k^* kümeleme parametreleridir, $f(\theta_k^*)$, k kümesindeki veriler üzerindeki dağılımdır ve G_0 kümeleme parametreleri üzerindeki ilk dağılımdır. Modeldeki son satır $y_i | z_i, \theta_k^* \sim f(\theta_z^*)$, verilerin artık kümelenmiş parametrelere göre modellendiğini göstermektedir (Teh 2011: 10).

Birçok veri analizi uygulaması için kesikli yapı uygun değildir. DPM modelleri, sürekli bir kernele sararak örnekleme dağılımındaki bu kesikli yapıyı düzeltir. Son olarak tipik bir DPM modeli şu şekildedir;

$$y_i \stackrel{ind}{\sim} f$$

$$f(y) = \int f(y | \mu) dG(\mu) \quad (3.17)$$

$$G \sim DP(\alpha, G_0)$$

yani, rastgele karıştırma dağılımı G üzerinde bir Dirichlet süreç ilk dağılımına sahip bir karışım. Birçok uygulama, ortak bir kovaryans matrisi S ile normal bir kernel $f(y|\mu) = N(\mu, S)$ kullanır bu da $f(y) = \sum_{k=1}^{\infty} w_k N(\mu_k, S)$ normallerinin kesikli bir karışımına neden olur. Dirichlet karışım modellerinin en önemli özelliği hesaplama basitliği ve son dağılım simülasyon algoritmalarının boyuttan bağımsız olmasıdır (De Iorio vd. 2004: 205).

Antoniak (1974) çalışmasına göre hiyerarşik model (3.15) altında, G üzerindeki son dağılım, $p(G|y_1, \dots, y_n)$, Dirichlet süreçlerin bir karışımıdır, yani gizili parametre θ_i 'ye göre karışan Dirichlet süreç modellerinin bir karışımıdır. $y_i | \theta_i \sim f_{\theta_i}$ $i = 1, \dots, n$, $\theta_i \sim G$ ve $G \sim DP(\alpha, G_0)$ ise son dağılım şu şekilde ifade edilmektedir:

$$G | y \sim \int DP(\alpha G_0 + \sum_{i=1}^n \delta_{\theta_i}) dp(\theta | y) \quad (3.18)$$

Uygun bir kernel fonksiyonu seçimi, temeldeki örnek uzaya bağlıdır. Altta yatan yoğunluk fonksiyonu gerçek çizginin ($\mathbb{R}$) tamamında tanımlanırsa, konum ve ölçeğe (μ, σ) göre bir kernel uygun olmaktadır. Birim aralıkta, beta dağılımları esnek iki parametrelili bir aile oluşturmaktadır. Pozitif yarımlı çizgide gama, Weibull veya log-normal dağılımların karışımları kullanılmaktadır (Müller vd. 2015: 12).

Nonparametrik Bayes karışım modeli olan Dirichlet süreç karışım modeller sonsuz bir karışım modelidir; sayılabilir şekilde sonsuz sayıda kümeye sahip bir karışım modelidir. Dirichlet süreç karışım modeli verilerden küme sayısını çıkarır ve yeni veri noktaları gözlemlendikçe küme sayısının artmasına izin verir (Gershman, Blei 2012: 2). Bu, verileri modellemek için sabit sayıda küme kullanan sonlu bir karışım modelinden farklıdır. Dirichlet süreç karışım modellerinde temel fikir, bir karışım bileşenini G 'deki her bir atomla ilişkilendirmektir. Veri noktalarını karışım bileşenleriyle ilişkilendiren gösterge değişkeni ise son dağılımın bir ürünü olarak ortaya çıkar ve veri kümeleri üzerinde bir olasılık dağılımı verir (Teh vd. 2006: 1567).

3.3.Nonparametrik Bayesyen Regresyon Modelleri

Standart regresyon yaklaşımları, konum ve ölçek gibi bazı sonlu sayıda yanıt dağılımı özelliklerinin bağımsız değişkenin (parametrik veya nonparametrik) bir fonksiyonu olarak değiştiğini varsayar. Ancak, hata dağılımının tahmin uzayı üzerinde değişmeyen bir şekle sahip olduğu bir konum/ölçek gösterimini varsaymak her zaman uygun değildir. Yanıtların dağılımının, bağımsız değişkenlerle, sonlu boyutlu bir fonksiyonel form tarafından makul bir şekilde temsil edilemeyeceği uygulamalı araştırmalarda görülmektedir. Bu nedenle daha genel yaklaşımlara ihtiyaç duyulmaktadır. Bu durum, nonparametrik regresyon modellerinin gelişmesine yol açmıştır. Nonparametrik regresyon, geleneksel regresyon modellerinin bir uzantısı olarak görülebilir, burada, $\mathcal{X}$ 'teki bazı öğelerden ve $\mathcal{Y}$ 'deki karşılık gelen bir dizi yanıtta başlanarak, amaç x verildiğinde y 'nin dağılımını modellemektir (Quintana vd. 2020:2).

Bir regresyon eğrisi bağımsız değişken x ile yanıt değişkeni y arasındaki genel bir ilişkiyi tanımlamaktadır. Regresyon ilişkisi şu şekilde modellenir:

$$y_i = f(x_i) + \varepsilon_i \quad (3.19)$$

Burada f bilinmeyen regresyon fonksiyonu ve ε_i hata terimidir. Hata terimi ε_i , $E(\varepsilon_i | x) = 0$ ve $E(\varepsilon_i | x) = \sigma_x^2$ özelliklerine sahiptir. Regresyon analizinin amacı, bilinmeyen f fonksiyonu tahmin etmektir. Bu tahmin iki yolla yapılabilir. İlki ortalama fonksiyon f için önceden belirlenmiş bir

fonksiyonel form varsayan ve fonksiyonel formun sonlu bir parametre kümesiyle tamamen tanımlandığını varsayan parametrik yaklaşımdır. İkincisi, belirli bir fonksiyonel form varsayımı yapmadan f fonksiyonunu nonparametrik bir şekilde tahmin etmeye çalışır. Fonksiyonel formu önceden belirlenmiş parametrik model düşük boyutlu olurken, nonparametrik yaklaşım bilinmeyen regresyon ilişkilerini analiz etmede esnek bir araç sunar. Bu bağlamda hem ortalama fonksiyon hem de hata dağılımı için fonksiyonel form önceden belirlenmemiştir. Böylece nonparametrik yaklaşım bilinmeyen regresyon fonksiyonu için daha esnek bir fonksiyonel form varsayımı yapmaktadır (Hardle 1994: 4-9).

Nonparametrik regresyon modellerinin birçok konuda yararlı olduğu söylenebilir. Bu yaklaşımların en büyük katkısı fonksiyonel form varsayımlarına daha az bağımlı olan tahminci ve çıkarım yöntemleri sunmasıdır. Yani regresyon eğrisinin fonksiyonel formu için kısıtlayıcı varsayım yapmamasıdır. Ayrıca eğer regresyon eğrisi hakkında hiçbir önsel model bilgisi mevcut değilse, nonparametrik analiz, regresyon ilişkisinin basit parametrik formlarını önermede faydalı olmaktadır (Yatchew 1998: 669).

Standart lineer regresyon modelleri, x 'in doğrusal bir kombinasyonu olarak modellenen bir ortalama ile $y|x$ için bir normal dağılım varsayımı altında ilerler. Bu fikrin üstel ailelere daha da genişletilmesi, ortalama tepkinin bir dönüşümünün x 'in doğrusal bir kombinasyonu olarak modellendiği genelleştirilmiş doğrusal modeller (GLM) sınıfına yol açmaktadır. GLM, doğrusal modelin bir bağlantı fonksiyonu⁹ aracılığıyla yanıt değişkeni ile ilişkilendirilmesine izin vererek doğrusal regresyonu genelleştirir. Klasik doğrusal modeller Y 'nin bileşenleri için μ ortalama ve sabit varyans σ^2 ile normal dağılım ve bağlantı fonksiyonu için birim fonksiyon varsayımı yapmaktadır. Genelleştirilmiş doğrusal modeller iki uzantıya izin verir; ilk olarak, bileşen dağılım, normal dağılım dışında bir üstel aileden gelebilir ve ikinci olarak, bağlantı fonksiyonu herhangi bir türevlenebilir monoton fonksiyon olabilir. Genelleştirilmiş doğrusal modeller, normallik varsayımını üstel ailelere genişletse de temelde parametrik bir yapıya bağlı kalır (McCullagh, Nelder 1983: 27).

Genelleştirilmiş doğrusal modellerde, bağlantı fonksiyonu doğrusal tahmin edicinin doğrusal olmayan bir dönüşümü olsa da modelin sistematik bileşeninin doğrusal kaldığı görülmektedir. Doğrusal tahmin, modelin sistematik bileşenini oluşturur. Modelin hem bağlantı fonksiyonu hem de stokastik bileşeni farklıdır ancak doğrusal $\eta = X'\beta$ tahmincisi kullanılır. Bu nedenle, doğrusal olmayan modeller olduğu varsayılan modellerin çoğu, doğrusallık varsayımını korur. Yani modelin değişkenlerde doğrusal olmama olasılığını engelleyen hiçbir şey yoktur. Gerçek veri oluşturma süreci doğrusal değilse, hiçbir

⁹ Bağlantı fonksiyonu doğrusal tahminci ve dağılım fonksiyonu ortalaması arasındaki ilişkiyi sağlamaktadır.

bağlantı fonksiyonu bu tanımlama hatasını düzeltmez¹⁰ (Keele 2008: 3-5). Bu yanlış tanımlama problemi yanlış fonksiyonel form varsayımı sonucunda ortaya çıkmaktadır. Çünkü yanlış fonksiyonel form özünde bir tanımlama hatasıdır¹¹. Nonparametrik regresyon, değişkenler arasındaki ilişki için önceden belirlenmiş herhangi bir parametrik form varsaymak yerine verilerin en uygun fonksiyonel formu belirlemesine izin vererek bu problemi çözmektedir.

Nonparametrik Bayesyen regresyon genelleştirilmiş doğrusal modelin nonparametrik versiyonuna odaklanmaktadır. Yani tahmin edicilerin etkisinin ortalama, kantil veya genelleştirilmiş bir doğrusal modeldeki parametreler gibi yanıt dağılımının bazı belirli fonksiyonlarını değiştirmek için sınırlı olduğu fikrinin ötesine geçmeyi içermektedir (Quintana vd. 2020: 2).

Ortalama fonksiyonu, hata dağılımı veya her ikisi üzerindeki kısıtlayıcı parametrik varsayımları gevşetilmek istendiğinde, regresyon problemleri doğal olarak nonparametrik Bayes yöntemler gerektirir. Rastgele ortalama fonksiyonlar için ilk dağılımlar, hata dağılımı için nonparametrik yoğunluk tahmininin kullanımı ve son olarak hem ortalama fonksiyon hem de hata dağılımı nonparametrik bir şekilde modellendiğinde tamamen nonparametrik regresyon için nonparametrik Bayes yöntemler mevcuttur. Bir $x_i \in \mathcal{X}$ ortak değişkeninin fonksiyonu olarak bir y_i yanıtını açıklamanın genel regresyon problemi şu şekilde ifade edilir:

$$y_i = f(x_i) + \varepsilon_i \quad (3.20)$$

Burada f bilinmeyen bir ortalama fonksiyonudur ve ε_i , genellikle bağımsız olduğu varsayılan hata terimidir. f fonksiyonu ve hata dağılımı sonlu boyutlu bir parametre vektörü θ tarafından indekslenirse, problem geleneksel parametrik regresyona, örneğin normal lineer regresyona indirgenir. Ancak sonlu boyutlu θ parametre vektörü kısıtlaması yoksa nonparametrik regresyon yöntemlere uzanır. Parametrik modelin gevşemesi olarak nonparametrik regresyonun bu tanımı, üç tip nonparametrik Bayesyen regresyon modelinin tanımlamasını gerektirmektedir:

- 1- Hata dağılımını modellemek için nonparametrik bir ilk dağılım kullanıldığında yani; $\varepsilon_i|G \sim G$ olduğunda ve G için bir nonparametrik Bayes ilk dağılımı belirlenir,
- 2- Ortalama f fonksiyonu üzerinde nonparametrik bir ilk dağılım varsayılır,

¹⁰ Genelleştirilmiş doğrusal modeller çerçevesinin ima ettiği gibi, lojistik veya Poisson regresyonu gibi modeller, yanlış bir fonksiyonel formdan kaynaklanan yanlış tanımlamaya eşit derecede eğilimlidir.

¹¹ Doğrusal regresyonda, yanlış tanımlama yalnızca, ihmal edilen değişken diğer değişkenlerle ilişkiliyse katsayı tahminlerinin sapmalı olmasına neden olur. Ancak doğrusal olmayan bağlantı fonksiyonuna sahip çoğu model için, model yanlış belirlenirse tüm parametreler sapmalı olur.

3- Koşullu dağılımlar kümesi üzerinde bir nonparametrik ilk dağılım kullanıldığında, yani; $\mathcal{G} = \{G_x, x \in \mathcal{X}\}$ ailesi üzerinde bir nonparametrik Bayes ilk dağılımına sahip $y_i | G, x_i \sim G_{x_i}$ modeli tanımlanır.

Bu üç yaklaşıma sırasıyla nonparametrik hata dağılımlar, nonparametrik ortalama fonksiyonlar ve tamamen nonparametrik regresyon olarak adlandırılmaktadır. Sonuncusu, yoğunluk regresyonu olarak da bilinmektedir ve tahmin edicilerin bir fonksiyonu olarak yanıt dağılımının tüm şeklinin değişmesine izin verilen yaklaşımları içermektedir (Müller vd. 2015:52).

Ele alınan nonparametrik Bayesyen regresyon problemi, $y_i | G_{x_i} \sim G_{x_i}$ olduğu varsayıldığında ortaya çıkmaktadır. İlgilenilen parametre, tahmin ediciye bağlı rastgele olasılık ölçülerinin tüm kümesidir, $\mathcal{G} = \{G_x, x \in \mathcal{X}\}$. Burada G_x , elemanları x tahmincilerinin değerleriyle esnek bir şekilde değişebilen, yanıt örnek uzayı Y , üzerinde tanımlanan bir olasılık ölçüsüdür, yani dağılımın tüm şekli x ile değişmektedir. Bayesci bir bakış açısından, nonparametrik regresyon modeli, G için bir ilk dağılım tanımlayarak tamamlanır (Quintana vd. 2020: 2).

Bu kurulum altında Bayesyen nonparametrik regresyon, ilgili rasgele olasılık ölçüleri $\mathcal{G}$, ailesi üzerinde çıkarıma indirgenir. $\mathcal{G}$ üzerinde anlamlı bir ilk dağılım, x boyunca bir süreklilik kavramı dahil olmak üzere G_x 'lerin bağımlılığını içermektedir. Rastgele olasılık ölçümleri ailesi için en temel ilk dağılım modeli bağımlı Dirichlet süreçlerdir (Müller vd. 2015: 65).

3.3.1. Bağımlı Nonparametrik Süreçler

Mevcut nonparametrik Bayesyen modeller ya nonparametrik dağılımın tek bir gerçekleştirmesi ya da nonparametrik dağılımların bağımsız gerçekleştirmeleri olan dağılımlar koleksiyonuyla sınırlıdır. İlk durumda, dağılımların az sayıda parametreyle, konumlar ve ölçeklerle farklılık göstermesine izin verilebilir, ancak bu durumda dağılımlar birçok yönden aynıdır. İkinci durumda, dağılımlar, nonparametrik ilk dağılımları kontrol eden hiperparametreler aracılığıyla birbirine bağlanabilir. Ancak, bu hiperparametrelere bağlı olarak, gerçekleşen dağılımlar bağımsızdır. İhtiyaç duyulan şey, birbirine benzer fakat aynı olmayan nonparametrik dağılımların bir koleksiyonunu yakalamaya izin veren genel rastgele etkiler dağılımı için bir model yazmaktır. Buradaki modelleme stratejisi rastgele dağılımların gerçekleştirmelerinin bağımlı olmasını gerektirmektedir (MacEachern 1999: 2).

Bağımlı nonparametrik süreçler, nonparametrik Bayes modellerinin geliştirilmesinde Dirichlet süreçlerinden sonra önemli bir adımdır. Bağımlı Dirichlet sürecinin geliştirilmesiyle birlikte, ortak değişkenin rastgele dağılımı indekslediği sürekli bir ortak değişken uzayı ile çalışılmaktadır. Böylece

tek bir bilinmeyen dağılımın ötesine geçilmesine, bağımlı nonparametrik dağılımlar uzayına girilmesine izin verilmektedir. Burada, normal hata dağılımı, bağımlı Dirichlet süreçlerinin bir karışımı kullanılarak gelişen bir hata dağılımı ile değiştirilir. Hata dağılımları arasındaki bağımlılığın oluşturulduğu indeks, ortak değişkenin kendisidir. Tüm bu genellemeler, hata dağılımına bir ölçek parametresinin dahil edilmesine izin verir. Böyle bir parametrenin dahil edilmesi, hataların değişen varyansının modellenmesini sağlar. Böylece genelleştirilmiş doğrusal modelin de ilerisine geçmeye izin verir (MacEachern 2000: 27).

Bağımlı nonparametrik süreçler, ortak değişkenin her sabit seviyesinde nonparametrik ve aynı zamanda ortak değişkenin bir fonksiyonu olarak yanıt dağılımının sürekliliğini sağlayan bir yanıt dağılımı kavramını yakalamayı sağlamaktadır. Bu modeller, ortak değişkenle "gelişen" yanıt dağılımlarına izin verdiğinden, bireysel durumlar hakkında çıkarımı kolaylaştırır ve ortak değişkenin farklı seviyeleri için farklı şekillerde tahmin dağılımların oluşmasına izin vermektedir. Bağımlı nonparametrik süreçler, verileri modellemek için daha gerçekçi bileşenler sağlama arzusuyla motive edilir; ortak değişken verildiğinde yanıtın hem ortalama yapısının hem de dağılım şeklinin uygun biçimde modellenmesine izin veren bileşenler elde edilmek istenmektedir (MacEachern 2001: 2).

Birçok veri setinde ortaya çıkan iki tema, bağımlı nonparametrik süreçlerin gelişimini motive etmektedir:

İlk tema, ortak değişken verildiğinde yanıtın koşullu dağılımının doğası gereği genellikle parametrik olmamasıdır. Bu bağlamda, genellikle parametrik formların dışına çıkma ihtiyacı doğmaktadır. Yanıt dağılımı sabit bir ortak değişken seviyesinde bir anlık görüntü olarak düşünülürse, bu anlık görüntü nonparametrik dağılıma olan ihtiyacı gösterebilir. Ancak birçok veri seti, yanıt dağılımını oluşturmak için ortak değişkenlerin her bir düzeyi için ampirik dağılıma bakabilecek kadar çok tekrara sahip değildir. Bunun yerine, ortak değişken seviyeleri arasında bilgi toplamayı sağlayacak bir modele ihtiyaç vardır.

İkinci tema, ortak değişken düzeyi değiştikçe koşullu yanıt dağılımlarının alışılmadık şekillerde değişmesidir. Dağılımdaki değişikliklerin basit bir göstergesi, genellikle bir çarpıklık ölçüsü izlenerek sağlanmaktadır. Ortak değişken değiştikçe yanıt dağılımının çarpıklığı değiştiğinde, değişen yanıt dağılımlarının buna izin vermesi gerekmektedir. Gelişen bir yanıt dağılımı, yanıt dağılımının nonparametrik bir biçim gerektirmesi gelişen parametrik olmayan dağılımların bir ailesini verir (MacEachern 2001: 3).

Bağımlı nonparametrik süreçler, gelişen nonparametrik dağılımlara yönelik temel Bayesyen yaklaşım ortaya koyar ve bir yan ürün olarak, nonparametrik rastgele etki dağılımlarının koleksiyonları için bir model sınıfı sağlar. Bu gelişen dağılımlar, dağılım fonksiyonları uzayına bir ilk dağılım yerleştiren

Bayes yöntemlerinin bir genellemesini temsil etmektedir. Böylece yalnızca nonparametrik bir formda tanımlanabilen, nonparametrik bir konum ölçek ailesi olarak tanımlanamayan ve ortak değişken değiştikçe düzgün bir şekilde değişmeyen koşullu dağılımlar için mantıklı modeller yazılmasını sağlamaktadır (MacEachern 1999: 2).

Bayesyen modelleme yaklaşımı büyük problemleri daha yönetilebilir bileşenlere bölmeye, problemin her bir bileşeni için bir ilk dağılım tasarlanmasına ve bileşenlerin hiyerarşik bir şekilde bağlanmasına izin vermektedir. Bağımlı nonparametrik süreçler de nispeten basit bir şekilde daha karmaşık Bayes modellerinin parçaları olarak kullanılabilir¹². Modern Bayes modellerini yönlendiren hiyerarşik modelleme tekniği, büyük bir modeli daha yönetilebilir bileşenlere böler ve ardından bu bileşenleri basit biçimlerde birbirine bağlar. Bağımlı nonparametrik süreçler, hiyerarşik modelin bileşenleri için ikameler sağlayarak, bileşenler arasındaki basit bağlantıları korurken bileşenlerin daha karmaşık bir şekilde etkileşime girmesine izin verir (MacEachern 2000: 27).

3.3.1.1. Bağımlı Dirichlet Süreç Modeli (DDP)

MacEachern (1999, 2000), nonparametrik Bayesyen regresyon problemini ele almak için bağımlı Dirichlet süreç modelini geliştirdi. Bağımlı Dirichlet süreç (Dependent Dirichlet Process, DDP) modelini, ortak değişkene bağlı rastgele olasılık dağılımlarının esnek bir sınıfı olarak tanımladı. Bağımlı Dirichlet süreci, Dirichlet sürecini tek bir dağılım fonksiyonuna bir ilk dağılım yerleştirme aracından, bir dağılım fonksiyonları koleksiyonuna bir ilk dağılım yerleştirme aracına genişletir. Böylece dağılım fonksiyonlarının koleksiyonu ortak değişkenler kümesi tarafından indekslenir (MacEachern 2001: 5). Burada ihtiyaç duyulan yapı, ortak değişken değiştikçe sürekli bir şekilde gelişen nonparametrik dağılımların bir koleksiyonuna bir olasılık dağılımı yerleştirmek için bir mekanizmadır. Bu özelliklere sahip bir model oluşturma anahtarı, önceki bölümde tanımlanan Sethuraman (1994) Dirichlet süreç inşasıdır. Bağımlı Dirichlet sürecinin basit formu için θ_k vektörünün çekimlerini indeks kümesi ortak değişken uzayı olan stokastik sürecin gerçekleştirilmesiyle değiştirilir. Bu gerçekleştirmelerin sonsuz bir dizisi, Dirichlet sürecindeki sonsuz konumlar kümesinin yerini alır (MacEachern 2016: 461).

MacEachern (1999) gerçekleştirmeleri bağımlı olan rastgele dağılımların toplamı için bir olasılık modeli tanımlayarak Dirichlet süreçlerini bağımlı Dirichlet süreçlerine genelleştirir. Burada G , G_x ile

¹²Bağımlı Dirichlet süreçler, modellemenin yüksek boyutlu bir uzayda ve çıkarımın daha düşük boyutlu bir uzayda gerçekleştiği modelleme yöntemlerini geliştirmiştir (MacEachern 2016: 461).

değiştirilerek ortak değişkenlere bağımlılığı içerdiği için, artık ilk dağılım belirleme problemi, ortak değişkene bağlı karıştırma dağılımı $\{G_x, x \in \mathcal{X}\}$ koleksiyonunun modellenmesi ile ilgili hale gelir. G_x, x tarafından indekslenmiş rastgele dağılımların bir kümesi olsun. Rastgele dağılımların toplamı şu şekilde tanımlanır:

$$G_x(\bullet) = \sum_{k=1}^{\infty} \underbrace{\left\{ V_{xk} \prod_{l < k} [1 - V_{xl}] \right\}}_{w_{xk}} \delta_{\theta_{xk}}(\bullet) \quad (3.21)$$

Burada V_{xk} , her $k \in \mathbb{N}$ için $\mathcal{X}$ indeks kümesi ve $Beta(1, \alpha_x)$ marjinal dağılımlarına sahip $[0,1]$ değerli bağımsız stokastik süreçlerdir. θ_{xk} , her $k \in \mathbb{N}$ için $\mathcal{X}$ indeks kümesi ve G_{0x} marjinal dağılımları ile bağımsız stokastik süreçlerdir.

Bağımlı Dirichlet süreçlerde G_x 'nin marjinal dağılımı, temel dağılım G_{0x} ve α_x kesinlik parametresi ile her $x \in X$ için bir Dirichlet sürecini takip eder; $G_x \sim DP(\alpha_x, G_{0x})$. G_x boyunca bağımlılık θ_{xk} atom konumlarının farklı x seviyelerine göre bağlı olduğu varsayılarak tanımlanır. Bağımlı Dirichlet süreç modellerinde θ_{xk} atom konumları temel G_0 dağılımına göre değil stokastik bir sürece göre ortak değişkenlerle değiştirilerek G 'nin ortak değişkenlere istenilen bağımlılığı sağlanır (Quintana vd., 2020: 8). Yani bağımlı Dirichlet süreçlerin standart Dirichlet süreçlerinden temel farkı atom konumlarının bazı temel G_0 dağılımlarının aksine temel stokastik süreç G_{0x} 'den elde edilen çekimler olmasıdır (Richardson, Hartman 2018: 5).

Tek boyutlu θ bağımlılığını yönlendiren fikir, bir ortak değişkenin varlığında, Dirichlet sürecinin konumu olan θ 'nın, θ_x ile gösterilen stokastik bir sürecin örnek yolu ile değiştirilebileceğidir. Bu örnek yol, ortak değişkenin her bir değerindeki konumu sağlar. Benzer şekilde, Sethuraman çubuk-kırma inşası ile θ değerine atanan ağırlığı üreten beta rastgele değişkeni, V, V_x ile gösterilen bir süreçle değiştirilebilir. Bu süreç, ortak değişkenin her bir değeri için θ_x değerine atanan ağırlığı üretir. Daha yüksek boyutlu modellerde, aynı fikir, $\theta_x = (\theta_{x1}, \dots, \theta_{xp})$ veren stokastik süreçlerin koleksiyonları yoluyla çok değişkenli konumlara genişletilir.

MacEachern (2000), her $k \geq 1$ için $Z_x^{v_k} = \{Z_k^v(x) : x \in \mathcal{X}\}$ ve $Z_x^{\theta_k} = \{Z_k^\theta(x) : x \in \mathcal{X}\}$ bağımsız iki stokastik süreç kümesinin dönüşümlerini kullanarak (birincisi $\{v_{xk}\}$ ağırlıkları ikincisi ise, $\{\theta_{xk}\}$ konumlarını tanımlamak için kullanılır) bağımlı Dirichlet süreçlerin standart yapısını tanımladı. Bir stokastik süreç $Z_x = (Z_x, x \in \mathcal{X})$ indeks kümesi¹³ $\mathcal{X}$, durum uzayı S ve Z_x arasındaki bağımlılık ilişkileri ile tanımlanır. Stokastik bir süreç, matematiksel kümeler tarafından indekslenen rastgele

¹³ Rastgele değişkenleri indekslemek için kullanılan küme, indeks kümesi olarak adlandırılır. İndeks kümesi, sonlu sayılabilir, doğal sayılar gibi reel sayıların ($\mathbb{R}$) alt kümeleri veya daha genel bir uzay olabilir.

değişkenlerin bir koleksiyonu¹⁴ olarak tanımlanır. Stokastik sürecin her rastgele değişkeni bu indeks kümedeki bir öge ile benzersiz bir şekilde ilişkilendirilir.

Z_x için durum uzayı $\mathbb{R}^p$ ve indeks kümesi $\mathcal{X} = \mathbb{R}$ olsun. Stokastik süreçler Z_x ile karşılık gelen bir stokastik süreç θ_x arasında bir bağlantı ve stokastik süreçler U_x ve V_x arasında benzer bir bağlantı kurar. U_x süreci gerçek değerlidir. Aşağıdaki tanımda G_x, p -boyutlu rastgele dağılım fonksiyonlarının koleksiyonunu temsil etmektedir. Kesinlik parametresi $\alpha_x, x \in \mathcal{X}$ tüm değerleri için sonlu ve pozitifdir, ancak x ile değişebilir. Temel kümülatif dağılım fonksiyonları kümesi G_{0x} 'dir. Bu kümülatif dağılım fonksiyonlarının her biri bir p -boyutlu dağılım fonksiyonudur. Stokastik süreç Z_x ve U_x sırasıyla konumlara ve olasılıklara dönüştürülen çekimleri sağlar. $T_{Z,\theta,x}$ dönüşümleri her $x \in \mathcal{X}$ için Z_x 'in θ_x 'e eşleşmesini gösterir. $T_{U,V,x}$ dönüşümleri her $x \in \mathcal{X}$ için U_x 'in V_x 'e eşleşmesini gösterir. Dönüşümler, Z_x, F_x gibi bir dağılıma sahipse, $T_{Z,\theta,x}$ 'in her $x \in \mathcal{X}$ için G_{0x} dağılımına sahip olacağı şekilde olmalıdır. Benzer şekilde, eğer U_x 'nin F_x gibi bir dağılımı varsa, o zaman $T_{U,V,x}$ 'in $Beta(1, \alpha_x)$ dağılımına sahip olması gerekir. Tüm bu dönüşümlerin ölçülebilir olduğu varsayılmaktadır. Bu açıklamalar ışığında bağımlı Dirichlet süreçleri şu şekilde tanımlanmaktadır:

$$G_x \sim DDP(\alpha_x, G_{0x}, Z_x, U_x, T_{Z,\theta,x}, T_{U,V,x}) \quad (3.22)$$

Burada Z_k ve $U_k, k = 1, 2, \dots$ için Z_x ve U_x stokastik süreçlerin bağımsız gerçekleştirmeleridir. θ_k ve V_k süreçleri sırasıyla Z_k ve U_k dönüşümleri tarafından belirlenir. G_x dağılımı her $x \in \mathcal{X}$ için $G_x(\theta) = \sum_{k=1}^{\infty} w_{xk} \theta_{xk}$ şeklinde kesikli dağılım olarak tanımlanır.

MacEachern (2000), $\{V_x\}$ ve $\{\theta_x\}$ için istenilen marjinal dağılımları sağlamak için ters dönüşüm yöntemine başvurmuştur. Z_x 'ten θ_x 'e standart dönüşüm kümülatif dağılım fonksiyonuna dayanan her bir noktada gerçekleşen bireysel koordinatlı dönüşümle tanımlanır:

$$\theta_{ix} = G_{0x}^{-1}(F_{ix}(Z_{ix})) \quad (3.23)$$

Burada $i = 1, \dots, p$ ve her $x \in \mathcal{X}$ için F_{ix}, Z_x 'in i . koordinatının marjinal kümülatif dağılım fonksiyonudur. F_i 'nin sürekli bir dağılım olduğu varsayılır ve G_{0x}, θ_x için temel kümülatif dağılım fonksiyonunun i . koordinatıdır. Burada ters kümülatif dağılım fonksiyonu tanımlanır. $F_{ix}(Z_{ix})$ uniform dağılımı izleyeceği için, θ_x marjinal G_{0x} dağılımını takip edecektir. Benzer şekilde U_x 'den V_x 'e standart dönüşüm şu şekilde yapılır:

¹⁴ Koleksiyondaki her rastgele değişken, durum uzayı olarak bilinen aynı matematiksel uzayda değerler alır. Bu durum uzayı örneğin tamsayılar, reel sayılar veya n boyutlu Öklid uzayı olabilir.

$$V_{ix} = G_{Bx}^{-1}(F_x(U_x)) \quad (3.24)$$

Burada sürekli bir dağılım olduğu varsayılan F_x , U_x 'in marjinal kümülatif dağılım fonksiyonudur¹⁵. Böylece her x için marjinal dağılımı G_{0x} elde edilmiş olur. Sonuçta bağımlı Dirichlet süreçler için G_x marjinal dağılımı; $G_x \sim DP(\alpha_x, G_{0x})$ Dirichlet sürecini takip eder (MacEachern 1999, 2000).

Sonuç, stokastik süreçlerin sayılabilir bir karışımıdır. Her bir ortak değişken x değerinde, rastgele dağılım G_x , sayılabilir bir karışımıdır. G_x 'in marjinal dağılımı bir Dirichlet sürecidir. Bütünüyle bakıldığında sonuç, olasılık dağılım değerli bir stokastik süreçtir. Bu kesikli dağılımlar daha sonra Dirichlet sürecinin karışım modelini oluşturmak için düzgünleştirildiği şekilde düzgünleştirilebilir (MacEachern, 2016: 461).

En yaygın kullanılan bağımlı Dirichlet süreç modeller, ortak değişken bağımlılığının ya konumlarda ya da ağırlıklarda ortaya çıktığını varsayar, diğerini ortak değişken bağımlılığı sergilemeyen bir rastgele değişkenler topluluğu olarak bırakır. Burada ele aldığımız model, hesaplama avantajları sunan tek-ağırlıklı bağımlı Dirichlet süreç modellerdir¹⁶. MacEachern tek- ağırlıklı DDP modeli olarak adlandırılan x değerleri arasındaki ortak ağırlıkları şu şekilde tanımladı:

$$G_x(\bullet) = \sum_{k=1}^{\infty} \underbrace{\left\{ V_k \prod_{l < k} [1 - V_l] \right\}}_{w_k} \delta_{\theta_{kx}}(\bullet) = \sum_{k=1}^{\infty} w_k \delta_{\theta_{kx}}(\bullet) \quad (3.24)$$

Burada V_k, x 'in tüm seviyelerinde ortak olan Beta $(1, \alpha)$ rastgele değişkenleridir ve sabit ağırlık $\sum_{k=1}^{\infty} w_k = 1$ çubuk -kıırma inşasındaki ile aynı şekilde oluşturulur. θ_{kx} , $\mathcal{X}$ indeks kümesi ve G_{0x} marjinal dağılımına sahip bağımsız stokastik süreçlerdir. Literatürde, tek-ağırlıklı DDP bağımlı Dirichlet süreçlerinin en popüler biçimidir, bunun sebebi son dağılım simülasyonun Dirichlet süreçler için mevcut olan aynı tür örnekleme algoritmaları kullanılarak uygulanabilmesidir (Quintana vd. 2020: 8)

θ_{kx} 'nin örnek yolu ortak değişkenin her değerindeki konumunu sağlar ve böylece rastgele dağılımlar arasındaki bağımlılık derecesi, $\{G_x, x \in \mathcal{X}\}$, ortak değişken düzeyi tarafından yönetilir. G_x boyunca

¹⁵ Z_x indeks kümesi $\mathcal{X} = \mathbb{R}$ üzerinde sıfır ortalama, σ^2 sabit varyans ile Gauss süreci olsun. F_i ve G_{Bx} sırasıyla $N(0,1)$ ve $Beta(1, \alpha_x)$ dağılımlarının kümülatif dağılım fonksiyonlarını gösterir.

¹⁶ Hem konumlarda hem de ağırlıklarda bağımlı olan DDP'ler, çoğunlukla uygulamanın gerektirdiği hesaplama karmaşıklığı nedeniyle daha az yaygın olarak bulunur. Bu durum lineer bağımlı Dirichlet süreç modeller (LDDP) için de geçerlidir (De Iorio vd. 2004).

bağımlılık k 'dan bağımsız olarak, θ_{xk} konumlarının farklı x seviyelerine bağlı olduğu varsayılararak tanımlanır. Her bir x değişkeni için $\{\theta_{xk}, x \in \mathcal{X}\}$ bağımlılığı açık olarak şu şekilde tanımlanır:

$$\begin{aligned} x=1: G_1 &= w_1 \delta_{\theta_{11}} + w_2 \delta_{\theta_{12}} + \dots \\ x=2: G_2 &= w_1 \delta_{\theta_{21}} + w_2 \delta_{\theta_{22}} + \dots \\ x=3: G_3 &= w_1 \delta_{\theta_{31}} + w_2 \delta_{\theta_{32}} + \dots \\ &\dots \end{aligned} \quad (3.25)$$

Sonuçta her $x \in \mathcal{X}$ için bağımlı Dirichlet süreçler için G_x marjinal dağılımı; $G_x \sim DP(\alpha, G_{0x})$ Dirichlet sürecini takip eder ve $\theta_{xk} \sim G_{0x}$ (i.i.d.) çekimlerdir. İlk bağımlı Dirichlet süreç modeli ANOVA-DDP modellerdir.

3.3.1.1.1. ANOVA Bağımlı Dirichlet Süreç Modeli (ANOVA-DDP)

De Iorio vd. (2004) bağımlı Dirichlet sürecinin birden çok kategorik ortak değişkeni kapsayacak şekilde genişleterek, karmaşık hiyerarşik Bayes modellerini modellemede bir bileşen olarak kullanmıştır. $\mathcal{G} = \{G_x, x \in \mathcal{X}\}$ 'in bir kategorik ortak değişken x tarafından indekslenen bir rastgele dağılımlar dizisi olsun. Burada rastgele dağılımlar arasındaki bağımlılık, θ_{xk} konumları için bir ANOVA tipi olasılık modeli uygulanarak tanımlanmaktadır. Açıklama kolaylığı için, $x = (v, w)$, $v \in \{1, 2, \dots, V\}$ ve $w \in \{1, 2, \dots, w\}$ olmak üzere iki değişkenli bir ortak değişken olsun. ANOVA-DDP modeli, $\mathcal{G}$ üzerinde ANOVA tipi bir bağımlılık yapısı tanımlamasına izin verir. İstenen bağımlılık yapısı, θ_{xk} konumlarına aşağıdaki olasılık modeli uygulanarak elde edilir:

$$\theta_{xk} = M_k + A_{vk} + B_{wk} \quad (3.26)$$

Yani konumlar üzerine $\{\theta_{xk}, x = (v, w)\}$ bir ANOVA model varsayarak $x = (v, w)$ için bağımlılık üretir. Burada $M_k \sim p_M^0(M_k)$, yani genel ortalama, $A_{vk} \sim p_{A_v}^0(A_{vk})$ ve $B_{wk} \sim p_{B_w}^0(B_{wk})$ ise v ve w ortak değişken düzeyleri için ana etkilerdir¹⁷. $\mathcal{G} = \{G_x, x \in \mathcal{X}\}$ üzerindeki birleşik olasılık modeli, ANOVA DDP(α, p^0) olarak adlandırılır; burada α , kesinlik parametresidir ve p^0 , eşitlikteki ANOVA etkilerinin temel ölçüsüdür. Marjinal olarak her $x = (v, w)$ için G_x , kesinlik parametresi α ve $(p_M^0, p_{A_v}^0, p_{B_w}^0)$ dağılımlarının sarılmasıyla (konvolüsyon)¹⁸ tanımlanan temel ölçü G_{0x} ile bir Dirichlet sürecini takip

¹⁷Burada $M_k \sim N(0, \tau_m^2)$, $A_v \sim N(0, \tau_v^2)$ ve $B_w \sim N(0, \tau_w^2)$, $A_{0k} \equiv B_{0k} \equiv 0$ olur.

¹⁸ Konvolüsyon fonksiyonel analizde bir fonksiyonun şeklinin diğer fonksiyon tarafından nasıl değiştiğini gösteren üçüncü bir fonksiyon $(f * g)$ üreten $(f$ ve $g)$ üzerinde bir integral işlemidir.

eder. Rastgele dağılımlar arasındaki bağımlılık, x boyunca nokta kütlelerinin θ_{xk} kovaryans yapısı ile tanımlanır. ANOVA DDP modelinde M_k , bir "genel ortalama" olarak yorumlanırken, A_v ve B_w , v ve w ortak değişken düzeyleri için "ana etkiler" olarak yorumlanır. Modelin önemli bir özelliği, model belirlemesinin ve hesaplamasının θ_{xk} 'nin boyutundan bağımsız olmasıdır¹⁹. ANOVA yapısı, vektör değerli θ_{xk} için tanımlanır ve çıkarım, θ_{xk} boyutundaki vektörlerin koşullu son dağılımlarının analitik ulaşılabilirliğini kullanan bir son dağılım simülasyon şemasıyla uygulanır.

Bu yapıdaki $\mathcal{X}$ ortak değişken uzayının kesikli olduğu görülmektedir. De Iorio vd. (2004), modele sürekli bir yanıtı yerleştirmek için Dirichlet süreç karışım modellerinde yapıldığı gibi yapıyı sürekli bir çekirdeğe, tipik olarak normal bir kernele sararak genişletti²⁰. Aynı nedenlerle, ANOVA DDP modeline ek bir normal karışım eklenmiştir. Bu model ortak normal varyans ϕ ve ANOVA DDP parametreleri için uygun hiper ilk dağılımlar ile şu şekilde ifade edilmektedir:

$$(y_i | x_i = x) \sim f_x(y_i),$$

$$f_x(y) = \int N(y | \mu, \phi) dG_x(\mu), \quad (3.27)$$

$$G_x \sim \text{ANOVA DDP}(\alpha, G_{0x}).$$

Yani bu model y_i verisinin bilinmeyen karıştırma dağılımı üzerinde bir Dirichlet süreç ilk dağılımına sahip ANOVA modellerin bir karışımından örneklendiğini söylemektedir (De Iorio vd. 2004: 207).

Daha genel bir ifade ile q -boyutlu rastgele dağılımlar için (3.27) eşitliği verilsin ve $\beta_k = [M_k, A_{2k}, \dots, A_{V_k}, B_{2k}, \dots, B_{W_k}]$, k . nokta konumuna karşılık gelen ANOVA parametrelerinin $(qx1)$ matrisini gösterebilir. d_i , bağımsız değişken x için tasarım vektörünü gösterebilir, yani $x_i = x$ için $\theta_{xk} = \beta_k d_i$ olur. Özel olarak bu yaklaşım $k \geq 1$ için $\theta_{xk} = \beta_k d_x$ kullanır, burada $\{\beta_k: k \geq 1\}$ G_0 dağılımına sahip (i.i.d.) rastgele vektörlerin bir dizisidir ve d_x gözlemlenen faktör değişken x ' in genel bir kombinasyonuna karşılık gelen tasarım vektörüdür. ANOVA-DDP modellerinin fikri, $\{\theta_{xk}, x \in \mathcal{X}\}$ konumları için basit doğrusal regresyonlar şeklinde ortak değişken bağımlılığını kodlamaktır. Bu durumda bağımlılık G_x ile şu şekilde tanımlanır;

¹⁹ Model kolaylıkla p -boyutlu bağımsız değişken vektörü $x = (x_1, \dots, x_n)$ içi genelleştirilebilir.

²⁰ Marjinal olarak $\mathcal{G}$ ailesinin her bir elemanı neredeyse kesin olarak kesikli olduğundan (çünkü bir DP'den çekimler), kesikli sonuçlar için modeller genellikle DP'leri sürekli bir kernele sarmak üzerine kuruludur. Böylece kendisi sürekli bir dağılım olan sürekli dağılımların bir karışımı elde edilir. ANOVA-DDP modelinde normal kernel tam olarak bu rolü oynar (Quintana vd. 2020: 10).

$$G_x(\bullet) = \sum_{k=1}^{\infty} w_k \delta_{\beta_k d_x}(\bullet) \quad (3.28)$$

yani, $\beta_k d_x$ doğrusal modellerinin bir Dirichlet süreç karışımı elde edilir. $\mathcal{G} = \{G_x, x \in \mathcal{X}\}$ koleksiyonunun her bir elemanı, konumları $\{\beta_k: k \geq 1\}$ ile verilen bir Dirichlet ilk dağılımına sahiptir. Bu notasyon kullanarak temel dağılım $(p_M^0, p_{A_v}^0, p_{B_w}^0) = G_{0x}$ ve $G_x = \sum_{k=1}^{\infty} w_k \delta_{\beta_k d_x}$ rastgele karıştırma dağılımı ile model (3.28) yeniden şu şekilde yazılabilir:

$$y_i | G_{x_i} \stackrel{ind.}{\sim} \int N(y_i | \beta d_{x_i}, S) dG_x(\beta)$$

$$y_i | \beta_i, \phi \sim N(\beta_i d_{x_i}, \phi)$$
(3.29)

$$\beta_1, \dots, \beta_n | G \stackrel{iid}{\sim} G$$

$$G \sim DP(\alpha, G_{0x})$$

Bu model, ϕ parametresi için uygun bir ilk dağılım ile tamamlanır. Eğer ϕ bir skaler ise $\phi \sim Ga(a, b)$ veya bir matris ise $\phi \sim Wishart(v, S)$ dağılımına sahiptir (Quintana vd. 2020:9). Tanımlanan yapıda ANOVA DDP modeli, kesikli ortak değişkenler gerektirir. Birçok veri analizinde bu durum, sürekli değişkenleri dahil etmek istediğimizde modelin uygulanabilirliğini sınırlar. De Iorio vd. (2009), ANOVA-DDP yapısının bir başka uzantısı olan lineer DDP (LDDP) modelleri inşa ettiler. Bu modellerde veriler kesikli ve sürekli ortak değişkenler içerdiği için bu sınırlamayı ortadan kaldırmaktadır.

3.3.1.1.2. Lineer Bağımlı Dirichlet Süreç Modeli (LDDP)

De Iorio vd. (2009), ANOVA DDP modelini bir ortak değişken kümesinin doğrusal kombinasyonu ile lineer bağımlı Dirichlet süreç modellerini (Linear Dependent Dirichlet Process, LDDP) genişletilir. Burada v kategorik ve z sürekli değişken olmak üzere $x = (v, z)$ iki değişkenli bir ortak değişken olsun. Rastgele dağılımlar arasındaki bağımlılık, nokta konumlarına doğrusal bir model uygulanarak elde edilir:

$$\theta_{xk} = M_k + A_{vk} + \lambda_k z \quad (3.30)$$

Burada $M_k \sim p_M^0(M_k)$, $A_{vk} \sim p_{A_v}^0(A_{vk})$ ve $\lambda_k \sim p_{\lambda}^0(\lambda_k)$. Standart bir doğrusal modelde olduğu gibi, λ_k bir eğim katsayısı olarak yorumlanabilir. Model α parametresi, eğim katsayısı λ ve ANOVA etkiler

üzerindeki p^0 temel dağılım ile parametrelendirilir. Yani G_x modeli parametreleştirir. Marjinal olarak, her $x = (v, z)$ için, rastgele dağılım G_x bir DP'yi takip eder. ANOVA DDP durumunda olduğu gibi, model (3.30), ortak değişkenler için θ_{xk} nokta konumlarının kovaryans yapısını belirleyerek x 'e bağımlılığı tanımlamaktadır. $\mathcal{G}$ üzerindeki ortak olasılık modeli $\{G_x, x \in \mathcal{X}\} \sim \text{Lineer DDP}(\alpha, p^0)$ şeklinde ifade edilir. Model, birden fazla sürekli ortak değişkene genelleştirilebilir.

Genel olarak her i için x_i , kategorik ve sürekli ortak değişkenlerin vektörü olsun. $y = y_1, \dots, y_n, [0, \infty)$ üzerinde tanımlanan sürekli negatif olmayan bir rastgele değişken olsun. $f(\cdot | \mu, \sigma^2)$ parametrik yoğunlukların keyfi bir konum ölçek ailesini gösterebilir. Her x_i , için verilerin karışım şu şekilde tanımlanır:

$$f_{x_i=x}(y_i | x, G_x) = \int f(y_i | \mu, \sigma^2) dG_x(\mu, \sigma^2) \quad (3.31)$$

$$G_x \sim \text{Lineer DDP}(\alpha, G_{0x})$$

G_x , modeli parametreleştirir ve w_k , tüm x 'ler için sabit olan ağırlıklar²¹ olmak üzere $G_x = \sum_{i=1}^{\infty} w_k \delta_{\theta_{xk}}$ ile DDP çerçevesi kullanılarak ortak değişken değerleri arasındaki bağımlılık tanımlanır. v kategorik ve z sürekli olmak üzere $x = (v, z)$ ortak değişkenler olsun. $\beta_k = [M_k, A_{2k}, \dots, A_{V_k}, \lambda_k]$ k . nokta konumu için denklem (3.30) de görülen parametrelerin satır vektörünü gösterebilir. d_i , ANOVA etkisini ve x_i 'ye karşılık gelen sürekli ortak değişkenin değerini temsil eden bir tasarım vektörünü gösterebilir. Böylece denklem (3.31) $x = x_i$ için $\theta_{xk} = \beta_k d_i$ olarak yazılabilir ve ikiden fazla ortak değişkene genelleme yapılabilir. Model (3.31) $G_x(\cdot) = \sum_{k=1}^{\infty} w_k N(\cdot | \beta_k d_i, \sigma_k^2)$ ile doğrusal modellerin bir Dirichlet süreç karışımı olarak şu şekilde yazılabilir²²:

$$f_{x_i=x}(y_i | G) \equiv f_x(y | \zeta) \equiv \int f(y | \beta d_i, \sigma^2) dG(\beta, \sigma^2) \quad (3.32)$$

$$G \sim \text{DP}(\alpha, G_0)$$

Burada $\zeta = G$ tüm model parametrelerini ve $G_0 = (p_M^0, p_{A_v}^0, p_{\lambda}^0, p_{\sigma^2}^0)$ temel dağılımı göstermektedir. Bu modelin, (3.31) modeli ile eş değer olması; $(\beta, \sigma^2) \sim G$ ve $\theta_{xk} = \beta_k d_i$, için temel ölçünün $(\theta, \sigma^2) \sim G_{0x}$

²¹ Lineer DDP kullanılırken, tahmin yoğunluğundaki ağırlıklar Dirichlet süreci tarafından ortaya çıkan kümelerin boyutuna göre belirlenir ve her küme farklı bir regresyon yapısını temsil eder (Cruz -Marcelo vd. 2010: 23).

²² DDP modeller DP modellerin bir karışımı olarak temsil edilebilir ve bu durum son dağılım simülasyon yöntemlerini uygulamada yardımcı olur. Ancak önerilen DDP modeller standart DP karışım modellerinin ötesine geçer. DPM modeller sabit, sonlu bir ortak değişken kümesinde tanımlanırken DDP modeli çok daha geniş bir uzayda tanımlanabilir (De Iorio vd. 2004: 208).

olması gerektiğini göstermektedir. Yani bu durum temel ölçünüm tanımı ile doğrulanır. Gizli değişkenler β_i ve σ_i^2 için model (3.32) hiyerarşik bir model olarak yeniden şu şekilde yazılır:

$$y_i | x_i, \beta_i, \sigma_i^2 \sim f(\cdot | \beta_i d_i, \sigma_i^2)$$

$$\beta_i, \sigma_i^2 | G \sim G \quad (3.33)$$

$$G \sim DP(\alpha, G_0).$$

Burada parametrik aile $f(\cdot | \mu, \sigma^2)$ normal ise, y_i gözlemleri bilinmeyen karıştırma dağılımı üzerinde bir Dirichlet süreç ilk dağılımına sahip doğrusal modellerin bir karışımından örneklenir²³. Temel dağılım G_0 ve kernelin koşullu eşleniği nedeniyle, normal veya log- normal kernellerin kullanımı, hesaplamayı büyük ölçüde kolaylaştırmaktadır²⁴. Lineer DDP model (3.33), koşullu eşlenik bir temel dağılım, $G_0 = (p_M^0, p_{A_v}^0, p_\lambda^0, p_{\sigma^2}^0)$, belirlenerek tamamlanır:

$$G_0(\beta, \sigma^2) = N(\beta | \mu, B) IG(\sigma^2 | s/2, sS/2) \quad (3.34)$$

Bu temel dağılım için eşlenik hiper ilk dağılımlar şu şekildedir: $S \sim Ga(0.5a_0, 0.5a_0/R)$, $\mu \sim N(m, A)$ ve $B^{-1} \sim Wish(c_0, (c_0 C_0)^{-1})$. Gama ve Wishart dağılımı, $E(B^{-1}) = C^{-1}$ ve $E(S) = R$ olacak şekilde parametrelendirilir. Son olarak, kesinlik parametresi $\alpha \sim G(a_\alpha, b_\alpha)$ için bir Gama ilk dağılımı olduğu varsayılır²⁵. Burada $N(\cdot | \mu, s)$, (μ, s) parametreleri ile normal olasılık yoğunluk fonksiyonunu göstermektedir. $Ga(\cdot | a, b)$, a şekil parametresi ve a/b beklenen değer ile gama olasılık yoğunluk fonksiyonunu ve $Wish(f, A)$, f serbestlik derecesi ve fA beklenen değer ile Wishart dağılımını göstermektedir (De Iorio vd. 2009: 763-765).

$D_n\{(y_i, x_i)\}_{i=1}^n$ verisinin $j = 1, \dots, N_j$ tarafından N_j gruba ayrıldığı varsayılınsın. Her j grubu için $y_{i(j)}, j$ grubunun i . yanıt gözlemini göstermektedir. $y_j = (y_{i(j)})_{i(j)}^{n_j}$, $x_{i(j)}^T$ ortak değişken vektörünün n_j satırları gözlenen tasarım matrisi $X_j = (x_{1(j)}^T, \dots, x_{n(j)}^T)$ ' ye karşılık gelen n_j yanıt gözleminin sütun

²³ Örneğin, parametrik aile log normal ise, log lineer modellerin bir karışımı elde edilir.

²⁴ Log-normal kernel seçimi, log- dönüştürülmüş verilere dayalı normal bir kernel kullanılmasına eşdeğerdir. Weibull gibi alternatif kerneller kullanmak da mümkündür, ancak eşlenikliğin olmaması hesaplamaları daha maliyetli hale getirecek ve daha az verimli MCMC şemalarının kullanılmasını gerektirecektir.

²⁵ Burada kesinlik parametresi α 'nın, gözlemlerin düştüğü küme sayısı üzerinde bir dağılıma neden olduğuna dikkat edilmelidir.

vektörü olsun. Bu durumda gruplara ayrılmış veriler için Lineer DDP modeli genel olarak şu şekilde tanımlanmaktadır²⁶:

$$[y_{i(j)}]_{i(j)}^{n_j} | X_j \sim f(y_j | X_j; \varsigma) \quad j = 1, \dots, N_j$$

$$f(y_j | X_j; \varsigma) = \sum_{k=1}^{\infty} \left\{ \prod_{i(j)=1}^{n_j} n(y_{i(j)} | x_{i(j)}^T \beta_k, \sigma_k^2) \right\} w_k$$

$$w_k = v_k \prod_{l=1}^{k-1} (1 - v_l)$$

$$v_k | \alpha \sim Be(1, \alpha)$$

$$\beta_k | \mu, T \sim N(\mu, T)$$

$$\sigma_k^2 \sim IG(a_0/2, a_0/2)$$

$$\mu, T \sim N(\mu | 0, r_0 I_{p+1}) IW(T | p + 3, s_0 I_{p+1})$$

$$\alpha \sim Ga(a_\alpha, b_\alpha) \tag{3.35}$$

Bu denklem sisteminde, j bir grubu, n_j , j grubundaki gözlem sayısını ve N_j toplam grup sayısını göstermektedir. Bu model lineer regresyon modellerinin bir Dirichlet süreç karışımıdır. Lineer DDP model, tüm model parametreleri, $\zeta \in \Omega_\zeta$ için ortak olasılık yoğunluk fonksiyonu olarak tanımlanan $\pi(\zeta)$ ilk dağılım belirlenerek tamamlanır:

$$\begin{aligned} \pi(\varsigma) = & \prod_{k=1}^{\infty} be(v_k | 1, \alpha) n(\beta_k | \mu, T) ig(\sigma_k^2 | a_0/2, a_0/2) \\ & \times n(\mu | 0, r_0 I_{p+1}) iw(T | p + 3, s_0 I_{p+1}) ga(\alpha | a_\alpha, b_\alpha) \end{aligned} \tag{3.36}$$

Bu model, kesinlik parametresi α ve temel dağılım $G_0(\beta, \sigma^2) = N(\beta | \mu, T) IG(\sigma^2 | a_0/2, a_0/2)$ sahip bir DP (α, G_0) ilk dağılımı atanan katsayılar ve hata varyans parametreleri için $G(\beta, \sigma^2)$ karıştırma dağılımına dayanmaktadır (De Iorio vd. 2009; Karabatsos 2017).

²⁶ Burada $N(\mu, T)$ ve $N(\mu | 0, r_0 I_{p+1})$ çok değişkenli normal dağılım ve IW ters-Wishart dağılımını göstermektedir.

3.4. Nonparametrik Bayesyen Regresyon Modelleri için Son Dağılım Algoritmaları

Nonparametrik Bayes yöntemlerin parametrik Bayes yöntemlere göre önemli bir avantajı belirsizliği dağılım fonksiyonları uzayları düzeyinde dahil etme yeteneğine sahip olmasıdır. Ancak, bu esnekliğin bir hesaplama maliyeti vardır. Nonparametrik Bayesyen modellerinin son yıllardaki hızlı gelişiminin çoğu, simülasyon tabanlı hesaplama yöntemlerindeki, özellikle Markov Zinciri Monte Carlo yöntemlerindeki ilerlemelerin doğrudan bir sonucu olmuştur. MCMC yöntemleri, genel amaçlı algoritmaların oluşturulmasına izin verir ve birçok farklı uygulama ve modelleme bağlamında DP ilk dağılımı altında son dağılım örnekleme için başarıyla kullanılmıştır. Bir önceki bölümde görüldüğü gibi DDP modeller, Dirichlet süreçlerin karışımı olarak temsili edildiği için burada DPM modeller için önerilen son dağılım algoritmaları DDP modeller için ortak olarak kullanılmaktadır (Müller vd. 2015: 70). Bu bölümün devamında nonparametrik Bayesyen modeller için kullanılabilecek algoritmalara yer verilmektedir.

3.4.1. Escobar ve West (1995) Gibbs Algoritması

Dirichlet süreç karışım modellerinin kullanımı, bileşen dağılımlarının parametrelerinin son dağılımından veya karışım bileşenlerinin gözlemlerle ilişkilerinin son dağılımından örnekleme için Markov zincir yöntemlerinin geliştirilmesiyle hesaplama açısından uygulanabilir hale gelmiştir. Eşlenik olma özelliği, hesaplama kaynaklarının yetersiz olduğu zamanlar için son dağılımların üretilmesine olanak sağladığı için nonparametrik Bayesyen yaklaşımda da önemli bir rol oynamaktadır (Neal 2000: 1)

Öncelikle algoritmalar için kullanılacak notasyon tanımlanmaktadır. G_0 sürekli temel ilk dağılım olmak üzere; c_i , hangi gizli kümenin gözlem y_i ile ilişkili olduğunu gösterir. Her $c = \{c_1, \dots, c_n\}$ kümesi için $\phi_c = \{\phi_1, \dots, \phi_k\}$ bu kümedeki gözlemlerin dağılımını belirler. Yani ϕ , farklı θ_i 'lerin kümesini gösterir ve burada $k \leq n$, $\theta = (\theta_1, \dots, \theta_n)$ vektöründeki farklı öğelerin sayısını göstermektedir. $i = 1, \dots, n$ için $\theta_i = \phi_c$ ise $c = \{c_1, \dots, c_n\}$, $c_i = c$ ile tanımlanan grupta göstergesi vektörünü gösterecek ve $c_i = c$ 'lerin sayısı n_c olsun. Aynı grupta göstergesi c_i 'lere karşılık gelen θ_i 'lere atıfta bulunmak için küme terimi kullanılmaktadır. Daha önce tanımlanan n_c , c . kümenin boyutudur: $n_c = \{i: c_i = c\}$. Kümeler için karıştırma oranı p_c 'ye simetrik bir Dirichlet ilk dağılımı atanır.

Karıştırma oranları p_c üzerinden integral alarak, aşağıdaki formun koşullu olasılıklarının ürünü olarak c_i için ilk dağılım yazılabilir:

$$P(c_i = c \mid c_1, \dots, c_{i-1}) \rightarrow \frac{n_{i,c}}{i-1+\alpha} \quad (3.37)$$

$$P(c_i \neq c_j \text{ tüm } j < i \mid c_1, \dots, c_{i-1}) \rightarrow \frac{\alpha}{i-1+\alpha}$$

Burada $n_{i,c} = \{c_j = c : j < i\}$, $j < i$ için c 'ye eşit olan c_j sayısını ifade etmektedir. Denge dağılımı olarak bu son dağılıma sahip bir Markov zincirini simüle ederek $(\theta_1, \dots, \theta_n)$ 'nin son dağılımından örnekler çekilir. Bu tür yöntemler, eşlenik ilk dağılımlar kullanıldığında Gibbs örneklemesine dayanmaktadır.

Diriclet karışım modelleri için örnekleme ilk direkt yaklaşım, $j \neq i$ için $(\theta_{-i}$ olarak yazılan) hem veri hem de θ_j verildiğinde koşullu dağılımdan her bir θ_i için değerler tekrar tekrar çekilerek elde edilerek inşa edilmektedir²⁷. Bu koşullu dağılım θ_{-i} üzerine koşullu ilk dağılım ve $F(y_i, \theta_i)$ olarak yazılan benzerlik fonksiyonunun birleştirilmesiyle elde edilir. θ_{-i} üzerine koşullu ilk dağılım şu şekilde ifade edilir²⁸:

$$\theta_i \mid \theta_{-i} \sim \frac{1}{n-1+\alpha} \sum_{j \neq i} \delta(\theta_j) + \frac{\alpha}{n-1+\alpha} G_0 \quad (3.38)$$

Benzerlik ile birleştirildiğinde, Gibbs örneklemesinde kullanım için aşağıdaki koşullu dağılımı verir:

$$\theta_i \mid \theta_{-i}, y_i \sim \sum_{j \neq i} q_{i,j} \delta(\theta_j) + r_i H_i \quad (3.39)$$

Burada H_i , tek y_i gözlemi için $F(y_i, \theta)$ benzerlik fonksiyonu ve G_0 ilk dağılımına dayanan θ için son dağılımı ifade etmektedir. $q_{i,j}$ ve r_i değerleri şu şekilde tanımlanır:

$$q_{i,j} = bF(y_i, \theta_j) \quad (3.40)$$

$$r_i = b\alpha \int F(y_i, \theta) dG_0(\theta)$$

²⁷ Gösterimlerde, θ_{-i} , $(-i)$ alt indisi, vektörün i . elemanı içermediği anlamına gelmektedir.

²⁸ Bu ifade Dirichlet sürecinin Polya urn şeması temsilinden gelmektedir. Bu koşullu ilk dağılım gözlemler değiştirilebilir olduğundan, i 'nin n gözlemin sonuncusu olduğu düşünülerek, $\theta_i \mid \theta_1, \dots, \theta_{i-1} \sim \frac{1}{i-1+\alpha} \sum_{j=1}^{i-1} \delta(\theta_j) + \frac{\alpha}{i-1+\alpha} G_0$ eşitliğinden türetilir. Burada $\delta(\theta)$ tek θ noktasında yoğunlaşan dağılımdır.

Burada b, $\sum_{j \neq i} q_{i,j} + r_i = 1$ olacak şekilde seçilmelidir. Bu Gibbs örnekleme yönteminin uygulanabilir olması için, r_i 'yi tanımlayan integralin hesaplanması ve H_i 'den örnekleme uygulanan işlemler olması gerekir. Verilen F benzerliği için G_0 eşlenik ilk dağılım olduğunda bu hesaplamalar mümkün olmaktadır. Bu Gibbs örnekleme algoritması şu şekilde özetlenebilir:

Algoritma 1: Markov zincir durumu $\theta_1, \dots, \theta_n$ 'dan oluşsun. Aşağıdaki adımlar tekrarlanarak örnekleme yapılır:

- 1- Adım: θ için bir başlangıç değerleri belirlenir, makul başlangıç değeri (3.39)'deki bireysel koşullu son dağılımdan (H_i) alınan örneklerdir.
- 2- Adım: $(\theta_1|\theta_{-1}, y)$ ardından $(\theta_2|\theta_{-2}, y)$ ve $(\theta_n|\theta_{-n}, y)$ 'ye kadar olan dağılımdan çekimler yaparak, her bir aşamada koşullandırma vektörlerine eklenen en son örnekleme θ_{-i} değerlerinin ilgili elemanları ile θ 'nın öğeleri sırasıyla örneklenir.
- 3- Adım: Adım 2'ye dönülür ve yakınsama oluncaya kadar yinelemeli olarak devam edilir (Escobar, West 1995: 579).

Bu algoritma ergodik bir Markov zinciri üretir, ancak son dağılıma yakınsama oldukça yavaş olabilir ve bu yüzden örnekleme verimsiz olabilir. Burada problem, genellikle aynı θ ile ilişkili yüksek olasılığa sahip gözlem gruplarının var olmasıdır. Algoritma aynı anda birden fazla gözlem için θ 'yı değiştiremeyeceğinden, böyle bir gruptaki gözlemler için θ değerlerinde bir değişiklik nadiren gerçekleşebilir. Eşitlik (3.40)'deki integral genellikle analitik olarak izlenebilir olmadığı için, bu algoritma G_0 ilk dağılımı eşlenik olmadığında kolayca uygulanmaz. Ayrıca H_i son dağılımdan örnekleme ilk dağılım eşlenik olmadığında zor olabilir (Neal 2000: 5).

3.4.2. MacEachern ve Müller (1998) Algoritması

Gibbs örneklemesine dayalı yöntemler, eşlenik ilk dağılımlara dayanan modeller için kolayca uygulanabilir, ancak birçok bağlamda uygun olduğu gibi, eşlenik olmayan ilk dağılım kullanıldığında, basit Gibbs örnekleme genellikle zor bir sayısal entegrasyonun gerçekleştirilmesini gerektirmektedir.

MacEachern ve Müller (1998), G_0 'dan çekilen ϕ için yardımcı değerlerin geçerli bir Markov zincir örnekleme için tanımlamak için kullanılmasına izin veren bir çerçeve sunmuştur. İlk yaklaşımlarda c_i 'nin sayısal değerleri, yalnızca hangi gözlemlerin aynı bileşenle ilişkili olduğunu gösterdikleri süreçte önemli olarak kabul edilmiştir. MacEachern ve Müller, c_i 'ye dağılımları atamak için, ϕ_c için değerler koleksiyonundan haritalamaya hizmet eden daha özel şemalar inşa etmişlerdir. Böyle bir şema, θ_i için

aynı dağılımı üretir, ancak farklı örnekleme algoritmalarına neden olmaktadır. MacEachern ve Müller'in “no gaps” algoritması, $i = 1, \dots, n$ için c_i 'nin 1'den k 'ya kadar olan tamsayılar kümesini kapsaması gerektiğinde ortaya çıkar. Burada k , farklı c_i sayısıdır, ancak başka türlü sınırlanmaz. Bu yaklaşımda Gibbs örnekleme basitleştirilmiştir; r_i integralinin değerlendirilmesi, basit benzerlik değerlendirmeleri ile değiştirilmiştir. Gibbs örnekleme göz önünde bulundurularak “no gaps” algoritması şu şekilde özetlenebilir:

Algoritma 2: Markov zincir durumu $c_1, \dots, c_n$ ve $\phi = (\phi_c : c \in \{c_1, \dots, c_n\})$ oluşsun. Aşağıdaki adımlar tekrarlanarak örnekleme yapılır:

- 1- Adım: $i = 1, \dots, n$ için; k^- , tüm $j \neq i$ için farklı c_j sayısı olsun ve bu c_j 'ler $\{1, \dots, k^-\}$ kümesinden değerler alsın. Eğer tüm $j \neq i$ için $c_j \neq c_i$ ise, $k^-(k^- + 1)$ olasılıkla c_i 'yi değişmeden bırakın. Aksi takdirde $j \neq i$ için $c_j \neq c_i$ ise, c_i 'yi $k^- + 1$ olarak etiketlenir veya bazı $j \neq i$ için $c_j = c_i$ ise G_0 'dan ϕ_{k^-+1} için bir değer çekilir. Ardından aşağıdaki olasılıklar kullanılarak $\{1, \dots, k^- + 1\}$ 'den c_i için yeni bir değer çekilir:

$$P(c_i = c \mid c_{-i}, y_i, \phi_1, \dots, \phi_{k^-+1}) = \begin{cases} bn_{-i,c} F(y_i, \phi_c) & \text{eğer } 1 \leq c \leq k^- \\ b[\alpha / (k^- + 1)] F(y_i, \phi_c) & \text{eğer } c = k^- + 1 \end{cases} \quad (3.41)$$

Burada b uygun normalleştirme sabitidir. Yapı artık yalnızca bir gözlemle ilişkili olan ϕ_c 'leri içerecek şekilde değiştirilir.

- 2- Adım: Tüm $c = \{c_1, \dots, c_n\}$ için; $\phi_c \mid y_i$ ögesinden $c_i = c$ olacak şekilde yeni bir değer çekilir veya bu dağılımı değişmez bırakan başka ϕ_c güncellemesi yapılır.

Bu algoritma F için G_0 'nın eşlenik olup olmadığına bakmaksızın, G_0 'dan örnekleme yapıldığı ve $F(y_i, \theta)$ hesaplandığı sürece uygulanabilir (Maceachern, Müller 1998: 228). Diğer taraftan, algoritmanın c_i 'yi diğer tüm c_j 'lerden farklı bir değere ayarlama mekanizmasında (yani yeni oluşturulan bir karışım bileşenine bir gözlem atamak için) verimsizlik ortaya çıkmaktadır (Neal 2000: 8).

3.4.3. Neal (2000) Metropolis-Hastings Algoritması

Eşlenik olmayan ilk dağılımlar için diğer temel yaklaşım, teklif dağılımı olarak koşullu ilk dağılımı kullanarak c_i 'yi güncellemek için Metropolis-Hastings algoritmasını kullanmaktır. Bu yaklaşım bir

Dirichlet süreç karışım modeli için bir algoritma üretmek için kullanılır. Bu durumda c_i için koşullu ilk dağılım şu şekilde tanımlanır:

$$\begin{aligned} \text{Eğer bazı } j \neq i \text{ için } c = c_j \text{ ise: } P(c_i = c | c_{-i}) &= \frac{n_{-i,c}}{n-1+\alpha} \\ \text{tüm } j \neq i \text{ için } c_i \neq c_j \text{ ise: } P(c_i \neq c_j | c_{-i}) &= \frac{\alpha}{n-1+\alpha} \end{aligned} \quad (3.42)$$

Bu koşullu ilk dağılımlar c_i 'yi güncellemek için teklif dağılımı olarak kullanılırsa c_i^* adayı $\{c_1, \dots, c_n\}$ içinde değilse ϕ için G_0 'dan ilişkili bir değer çekmek gerekecektir. Eğer mevcut c_i diğer herhangi bir c_j 'ye eşit değilse, c_i^* 'yi c_i ile aynı seçme olasılığının sıfır olduğuna dikkat edilmelidir. Yani c_i^* diğer herhangi bir c_j 'den farklı seçildiğinde, mevcut c_i değil, diğer c_j 'den bile farklı olsa her zaman yeni bir bileşen olacaktır. Her bir c_i için birden fazla Metropolis-Hastings güncellemesi yapılır. Metropolis-Hastings algoritması şu şekilde özetlenebilir:

Algoritma 3: Markov zincir durumu $c_1, \dots, c_n$ ve $\phi = (\phi_c : c \in \{c_1, \dots, c_n\})$ oluşsun. Aşağıdaki adımlar tekrarlanarak örnekleme yapılır:

1- Adım: $i = 1, \dots, n$ için c_i 'nin aşağıdaki güncellemesi R kez tekrarlanır. Eşitlik (0.94) ile verilen c_i koşullu ilk dağılımından c_i^* , adayı çekilir. Eğer bu c_i^* , $\{c_1, \dots, c_n\}$ 'de değilse, G_0 'dan $\phi_{c_i^*}$ için bir değer seçilir. c_i için teklif dağılımı olarak bu koşullu ilk dağılım kullanılırsa, $\alpha(c_i^*, c_i)$ kabul olasılığı şu şekilde hesaplanır;

$$\alpha(c_i^*, c_i) = \min \left[1, \frac{F(y_i, \phi_{c_i^*})}{F(y_i, \phi_{c_i})} \right] \quad (3.43)$$

ve bu olasılıkla c_i 'nin yeni değeri c_i^* olarak ayarlanır. Aksi takdirde, c_i 'nin yeni değeri eski değerle aynı olsun.

2- Adım: Tüm $c \in \{c_1, \dots, c_n\}$ için; $c_i = c$ olacak şekilde $\phi_c | y_i$ 'den yeni bir değer çekilir veya bu dağılımı değişmez bırakan başka ϕ_c güncellemesi yapılır.

R birden büyükse, daha önce hesaplanan F değerlerini yeniden kullanarak hesaplama süresinden tasarruf etmek mümkündür. c_i^* ile c_i aynı olduğu durumda F 'nin hesaplanması da atlanabilir. Bu nedenle bir c_i 'yi güncellemek için gereken F hesaplamalarının sayısı $(R + 1)$ 'den fazla değildir²⁹. Metropolis-Hastings yöntem (Algoritma 3) davranışı, karşılık gelen Gibbs örnekleme yöntemlerinden (Algoritma

²⁹ Karşılaştırma açısından, Gibbs örnekleme (Algoritma 1) ve “no gaps” algoritması (Algoritma 2) için bir gereken F hesaplamalarının sayısı yaklaşık olarak $j \neq i$ o.ü farklı c_j sayısı artı bir olarak hesaplanabilir.

1) ve “no gaps yönteminden (Algoritma 2) önemli ölçüde farklıdır. Bu diğer yöntemler, c_i için yeni bir değere karar verirken tüm karışım bileşenlerini dikkate alırken, Metropolis-Hastings yönteminin, c_i 'yi birkaç gözlemle ilişkili bir bileşenden ziyade birçok gözlemle ilişkili bir bileşene değiştirmeyi düşünmesi daha olasıdır (Neal 2000: 10).

3.4.4. Neal (2000) Yardımcı Parametrelere Sahip Gibbs Algoritması

Neal (2000) çalışmasında yardımcı parametrelerin eklenmesiyle genişletilmiş bir duruma Gibbs örnekleme uygulayarak eşlenik olmayan ilk dağılımlara sahip modeller için yardımcı parametrelere sahip Gibbs örnekleme yaklaşımını tanıtmıştır. Bu yaklaşım, MacEachern ve Müller (1998) yaklaşımına benzer, ancak yardımcı parametrelerin yalnızca geçici olarak var olduğunun kabul edilmesi açısından farklıdır; bu durum algoritmaların oluşturulmasında daha fazla esneklik sağlamaktadır.

Yardımcı değişken yöntemlerinin temel fikri x 'in marjinal dağılımının π_x olduğu (x, y) için bazı π_{xy} dağılımdan örnekleme yaparak x için bir π_x dağılımından örnekleme yapılabilmesidir. Bu fikir Markov zincir simülasyonu sırasında oluşturulan ve atılan yardımcı değişkenleri barındıracak şekilde genişletilmiştir. Markov zincirinin kalıcı durumu x olacaktır, ancak aşağıdaki formun bir güncellemesi sırasında geçici olarak bir y değişkeni tanıtılacaktır³⁰:

- 1- π_{xy} tarafından tanımlandığı gibi, x verildiğinde koşullu dağılımından y için bir değer çizin.
- 2- π_{xy} değişmez bırakan bazı (x, y) güncellemelerini gerçekleştirin.
- 3- Yalnızca x değerini bırakarak y 'yi atın.

Bu yaklaşım G_0 'ya göre integral almak zorunda kalmadan Dirichlet süreç karışım modelinde c_i 'yi güncellemek için kullanılmaktadır³¹. Markov zincirinin kalıcı durumu, c_i ve ϕ_c 'dan oluşacaktır, ancak c_i güncellendiğinde, diğer gözlemlerle ilişkili olmayan bileşenlerin parametreleri için olası değerleri temsil eden geçici yardımcı değişkenler tanımlanır. Daha sonra c_i bu yardımcı parametreleri içeren

³⁰ π_x, π_{xy} altında x 'in marjinal dağılımı olduğu sürece, x için bu güncellenmenin π_x ' değişmez bırakmaktadır. π_x 'e yakınsayacak bir Markov zinciri oluşturmak için π_x 'i değişmez bırakan diğer güncellemelerle birlikte farklı yardımcı değişkenler içerebilen bu tür birkaç güncelleme birleştirilebilir.

³¹ Önceki algoritmalar, eşlenik olmayan durumda değerlendirilmesi zor olan G_0 üzerinden integral almayı gerektirmektedir. Burada G_0 üzerinden integral almak yerine, her biri G_0 'dan bağımsız olarak çekilen yeni bir bileşene karşılık gelen m yardımcı parametre eklenir.

dağılıma göre Gibss örnekleme ile güncellenir. c_i , diğer gözlemlerle ilişkili bileşenlerden biri veya tanımlan yardımcı bileşenlerden biri olması gerektiğinden, bu olasılıkların göreceli olasılıklarını değerlendirerek Gibbs örnekleme yapılır. c_i için yeni bir değer seçildikten sonra, bir gözlemle ilişkili olmayan tüm ϕ değerleri atılır. Bu yardımcı parametrelerle Gibss örnekleme algoritması şu şekilde özetlenebilir:

Algoritma 4: Markov zincir durumu $c_1, \dots, c_n$ ve $\phi = (\phi_c: c \in \{c_1, \dots, c_n\})$ oluşsun. Aşağıdaki adımlar tekrarlanarak örnekleme yapılır:

1- Adım: $i = 1, \dots, n$ için: $k^-, i \neq j$ için farklı c_j sayısı ve $h = k^- + m$ olsun. Bu c_j 'ler $\{1, \dots, k^-\}$ içindeki değerlerle etiketlenir. Bazı $i \neq j$ için $c_i = c_j$ ise, $k^- < c \leq h$ olan bu ϕ_c için G_0 'dan bağımsız değerler çekilmesine izin verilir. Tüm $i \neq j$ için $c_i \neq c_j$ ise, c_i , $k^- + 1$ etiketine sahip olur ve $k^- + 1 < c \leq h$ olan bu ϕ_c için G_0 'dan bağımsız değerler çekilmesine izin verilir. Aşağıdaki olasılıklar kullanarak $\{1, \dots, h\}$ 'den c_i için yeni bir değer çekilir:

$$P(c_i = c | c_{-i}, y_i, \phi_1, \dots, \phi_h) = \begin{cases} b \frac{n_{-i,c}}{n-1+\alpha} F(y_i, \phi_c) & 1 \leq c \leq k^- \\ b \frac{\alpha/m}{n-1+\alpha} F(y_i, \phi_c) & k^- < c \leq h \end{cases} \quad (3.44)$$

Burada $n_{-i,c}, j \neq i$ için c 'ye eşit c_j sayısıdır ve b uygun normalleştirme sabitidir. Durumu, yalnızca şu anda bir veya daha fazla gözlemle ilişkilendirilmiş olan ϕ_c 'yi içerecek şekilde değiştirin.

2- Adım: Tüm $c \in \{c_1, \dots, c_n\}$ için: $c_i = c$ olacak şekilde $\phi_c | y_i$ 'den yeni bir değer çekilir veya bu dağılımı değişmez bırakan başka ϕ_c güncellemesi yapılır.

Burada $m = 1$ olduğunda, Algoritma 4, Algoritma 2'ye çok benzer. Aradaki fark, c_i 'yi diğer gözlemlerle paylaşılan bir bileşenden yeni bir tek bileşene değiştirme olasılığının Algoritma 4 ile yaklaşık $k^- + 1$ kat daha fazla olmasıdır ve aynısı ters değişim için de geçerlidir. α , küçük olduğunda diğer değişikliklerin olasılıkları daha az etkilendiğinden, bu açık bir fayda gibi görülmektedir. Bununla birlikte, Algoritma 4 tarafından tanımlanan Markov zincirinin denge dağılımı, Algoritma 2'yi uygulamak için bir Monte Carlo yaklaşımının kullanıldığı durumun aksine, herhangi bir m değeri için kesin olarak doğrudur (Neal 2000: 14).

3.4.5. Kalli vd.(2011) Kesit Örnekleyici

Walker (2007) ve Kalli vd.(2011) Dirichlet süreç karışım modelleri için bir kesit örnekleyici (slice sampler) geliştirdiler. Bu yaklaşım daha önce açıklanan algoritmaların aksine, rastgele olasılık ölçüsü G 'yi marjinalleştirmez (Müller vd. 2015: 23).

Diğer açıklanan algoritmalar, sonsuz boyut problemini ortadan kaldırarak modelden rastgele dağılım fonksiyonunun integralinin alınmasına dayanmaktadır³². Bunlar genellikle “marjinal” yöntemler olarak adlandırılır. Rastgele dağılım fonksiyonlarını da muhafaza eden bu yeni algoritmanın temel fikri, bir Gibbs örnekleyicisinin her yinemesinde örneklenecek nesnelerin bilinen sonlu bir sayısına izin veren gizli bir değişkenin tanıtılmasıdır. Gizli değişkenlerin tanıtılması, bir Gibbs örnekleyici aracılığıyla yineme yapmak için gerekli olan rastgele dağılım fonksiyonunun parçasını sonlu yapar. Ayrıca, tüm koşullu dağılımların örneklenmesi kolaydır ve hiçbir kabul veya red yöntemine gerek yoktur. Çubuk-kırma temsili ile bir Dirichlet karışım modeli genel olarak şu şekilde yazılabilir:

$$f_{w,\theta}(y) = \sum_{k=1}^{\infty} w_k N(y | \theta_k) \quad (3.45)$$

Burada temel nokta, doğru durağan dağılıma sahip geçerli bir Markov zinciri üretmek için tam olarak hangi (sonlu sayıda) değişkenin örneklenmesi gerektiğini bulmaktır. Modeli Gibbs örnekleme yaklaşımı ile tahmin etmek için, $(w, \theta = (\mu, \sigma^2))$ verilğinde (y, u) ortak yoğunluğu aşağıdaki gibi tanımlanacak bir gizli u değişkeni tanımlanır:

$$f_{w,\theta}(y, u) = \sum_{k=1}^{\infty} 1(u < w_k) N(y | \theta_k) \quad (3.46)$$

Bu ortak yoğunlukla birlikte u için marjinal bir yoğunluk da olacaktır³³. u , verildiğinde bileşenlerin sayısı sonludur ve indeksler $A_u = \{k: w_k > u\}$ olur. Burada A_u tüm $u > 0$ için sonlu bir kümedir ve A_u boyutu $\sum_{k=1}^{\infty} 1(w_k > u)$ olur. Böylece u verildiğinde y 'nin koşullu yoğunluğu şu şekilde verilir:

$$f_{w,\theta}(y | u) = f_w(u)^{-1} \sum_{k \in A_u} N(y | \theta_k) \quad (3.47)$$

³² Diğer algoritmalar modelde rastgele dağılım fonksiyonunun tutulmasına odaklanmaktadır, ancak örnekleme algoritmaları, bu dağılımların sahip olduğu sayılabilir sonsuz temsilden zarar görmektedir (Walker 2007: 45).

³³ Lebesgue ölçüsüne göre u üzerinden integral alındığında istenen $f_{w,\theta}(y)$ yoğunluğu elde edilmektedir.

Burada $f_w(u) = \sum_{k=1}^{\infty} 1(w_k > u)$, u için marjinal yoğunluktur³⁴. Böylece u verildiğinde sonsuz toplamdan sonlu toplama geçilir ve $1/f_w(u)$ eşit ağırlıklara sahip sonlu bir karışım modeli elde edilir. Daha sonra, y 'nin alınacağı karışım bileşenini tanımlayacak olan başka bir gizli gösterge değişkeni tanımlanır. Bu nedenle, aşağıdaki ortak yoğunluk dikkate alınır:

$$f_{w,\mu,\sigma^2}(\{y_i, u_i, d_i\}_{i=1}^n) = \prod_{i=1}^n 1(u_i < w_{d_i}) N(y_i | \mu_{d_i}, \sigma_{d_i}^2) \quad (3.48)$$

Böylece, (v, μ, σ^2) için, n boyutundaki bir örneğe dayalı tam olabilirlik fonksiyonu terimlerin basit bir çarpımı olarak mevcuttur ve d sonludur. Daha önce bahsedildiği gibi, w ve $\theta = (\mu, \sigma^2)$ için ilk dağılımlar zaten biliniyor. Örnekleme söz konusu olduğunda w 'ler yerine v 'ler kullanılmaktadır (Walker 2007: 45-48).

Walker (2007) kesit örneklemesine dayanan daha genel bir kesit örnekleme sınıfı şu şekilde tanımlanabilir:

$$f_{v,\mu,\sigma^2}(y, u, d) = \xi_d^{-1} 1(u < \xi_d) w_d N(y | \mu_d, \sigma_d^2) \quad (3.49)$$

Burada $\xi_1, \xi_2, \xi_3, \dots$ herhangi bir pozitif dizidir. $\xi_1, \xi_2, \xi_3, \dots$ seçimi önemlidir ve herhangi bir seçim, verimlilik ile hesaplama süresini dengelemek zorundadır. Karıştırma, i artış hızına sahip $r_i = E[w_i]/\xi_i$ oranına bağlıdır³⁵. Daha hızlı artış oranları, daha iyi karıştırma ile ilişkili olur. Ancak ortalama A_u boyutu arttığı için daha uzun çalışma süreleri ile ilişkilidir. Gibbs örnekleme adımının her yenilenmesinde örnekleme gereken değişkenler şu şekildedir: $\{(\mu_k, \sigma_k^2, v_k), k = 1, 2, \dots; (d_i, u_i), i = 1, \dots, n\}$ ve ortak son dağılım aşağıdaki eşitlikle orantılıdır:

$$\prod_{i=1}^n 1(u_i < \xi_{d_i}) w_{d_i} / \xi_{d_i} N(y_i | \mu_{d_i}, \sigma_{d_i}^2) \quad (3.50)$$

Kesit örnekleme, her yinelemede kullanılacak bir sonlu karışım modeli için Gibbs örnekleme adımına izin vermektedir. Kalli vd. (2011) tarafından geliştirilen bu kesit örnekleme algoritması şu şekilde özetlenebilir.

Algoritma 5: ξ ve v koşullu olarak bağımsız olmak üzere, aşağıdaki adımlar tekrarlanarak örnekleme yapılır:

³⁴ Burada y ve u , w_k olasılıkla bağımsızdır ve sırasıyla normal ve düzgün (uniform) dağılıma sahiptir. Dolayısıyla, u için marjinal yoğunluk şu şekilde verilir: $f_w(u) = \sum_{k=1}^{\infty} w_k U(u|0, w_k) = \sum_{k=1}^{\infty} 1(w_k > u)$.

³⁵ Örneklerde $r_i = (1.5)^i$ iyi bir denge sağladığı görülmektedir.

- 1- Adım: $\pi(\mu_k, \sigma_k^2 | \dots) \propto G_0(\mu_k, \sigma_k^2) \prod_{d_i=k} N(y_i | \mu_k, \sigma_k^2)$ yazılır.
- 2- Adım: $\pi(v_k) \propto Be(v_k | a_k, b_k)$, burada $a_k = 1 + \sum_{i=1}^n 1(d_i = k)$ ve $a_k = \alpha + \sum_{i=1}^n 1(d_i > k)$ çekilir.
- 3- Adım: $\pi(u_i | \dots) \propto 1(0 < u_i < \xi_{d_i})$.
- 4- Adım: $G(d_i = h | \dots) \propto 1(h: \xi_h > u_i) w_h / \xi_h N(y_i | \mu_h, \sigma_h^2)$.

Görüldüğü gibi bu algoritma (μ_k, σ_k^2, v_k) değişkenlerin tamamını örneklemmez. Ancak zinciri ilerletmek için bu gerekli değildir. 4. Adımı tam olarak gerçekleştirmek için sadece uygun olan tüm h 'ların bulunan N tamsayısına kadar örneklenmesi gerekmektedir. Gerekli (h) kümesini bulmak basittir, çünkü bu $\{1, \dots, N\}$ türünden olacaktır, burada $N = \max_i \{N_i\}$ eşittir ve $N_i, \xi_l > u_i$ için en büyük l tamsayıdır. Bu algoritma, karışım sayısını sonlu yapan anahtar gizli değişkenin eklenmesiyle, rastgele olasılık ölçütlerine dayalı hem çubuk- kırma hem de normalleştirilmiş türlerdeki karışım modelleri için uygulanabilmektedir³⁶ (Kalli vd. 2011: 94-95).

4.UYGULAMA: OECD ÜLKELERİNDE DAYANIKLI ÖĞRENCİLERİN MATEMATİK BAŞARISINI ETKİLEYEN FAKTÖRLERİN BELİRLENMESİ

Bu bölümde önceki bölümde tanıtılan nonparametrik Bayesyen regresyon modeli kullanılarak OECD ülkelerinde dayanıklı öğrencilerin matematik başarısını etkileyen faktörler incelenecektir. Bölümün devamında öncelikle akademik dayanıklılık kavramı tanıtılacak ve ilgili literatüre yer verilecektir. Ardından analizde kullanılacak veri setine ait bilgiler ayrıntılı olarak sunulacaktır. Son olarak nonparametrik Bayesyen regresyon tahmin sonuçlarına ilişkin bulgulara ve optimal kümeleme sonuçlarına ilişkin değerlendirmelere yer verilecektir.

³⁶ Bu algoritma, bağımsız olarak $v_k \sim Be(\alpha_k, \beta_k)$ çekimleri ile inşa edilen daha genel çubuk-kırma süreçleri için bir uzantı sağlamaktadır. Daha genel ilk dağılımlara bu kolay genişletme, marjinal örnekleme algoritmalarının bir özelliği değildir (Kalli vd. 2011).

4.1. Akademik Dayanıklılık

Eğitim, bireylerin ve toplumların yaşam kalitesini yükseltebilir. Eğitim sistemleri için tüm öğrencilerin akademik olarak tam potansiyellerine ulaşmalarını sağlamak hem eşitlik hem de verimlilik nedenleriyle önemli bir politika hedefidir (OECD 2011: 29). Eğitimde eşitlik, eğitim sistemlerinin tüm öğrencilere eşit öğrenme fırsatları sağlaması anlamına gelmektedir. Böylece eğitim sistemleri arasında farklı sosyoekonomik statüye, cinsiyete veya aile geçmişine sahip öğrenciler temel bilişsel alanlarda benzer düzeyde akademik performansa ve benzer düzeyde sosyal ve duygusal refaha ulaşırlar. Eşitlik tüm öğrencilerin eşit eğitim sonuçları elde etmesi anlamına gelmez, bunun yerine sonuçlardaki farklılıkların öğrencilerin kontrol sahibi olmadığı ekonomik ve sosyal koşullarla ilgili olmaması anlamına gelmektedir. Rapor sonuçlarının da gösterdiği gibi eğitimdeki sosyoekonomik eşitsizliklerin tamamen ortadan kalktığı hiçbir ülke yoktur (OECD 2018: 13). Son yıllardaki ekonomik eşitsizlikteki keskin artışa bakıldığında, (OECD ülkelerinde gelir eşitsizliği 1980'lerden bu yana en yüksek seviyededir) ülkeler eğitimde eşitliğin iyileştirilmesi konusunda önceki yıllarda olduğundan çok daha fazla endişe duymaktadır (OECD 2019).

Eğitimde eşitlik, Birleşmiş Milletler tarafından 2015 yılında belirlenen 17 Sürdürülebilir Kalkınma Hedefleri içerisinde 7'sinde eğitimle ilgili hedefleri içeren 2030 gündeminde merkezi bir rol oynamaktadır. Eğitimle ilgili sürdürülebilir kalkınma hedefi 2030 yılına kadar "*kapsayıcı ve eşitlikçi kaliteli eğitimi sağlamayı ve herkes için yaşam boyu öğrenme fırsatlarını teşvik etmeyi*" amaçlamaktadır (UNESCO, 2018). Kapsayıcı eğitim öğrencinin sosyoekonomik geçmişinden bağımsız asgari yeterliliğe ulaşması ve kişisel durumları ne olursa olsun tüm çocukların eğitime erişiminin sağlanması olarak tanımlanmakta ve eğitimde eşitliğin sağlanmasının temel şartı olarak gösterilmektedir (OECD 2019: 54).

Uluslararası büyük ölçekli eğitim değerlendirmeleri, zaman içinde ülkeler arasında öğrenci performansındaki farklılıkları anlamak için kullanılmaktadır. Başarıdaki farklılıklar, öğrencilerin temel özellikleriyle ilişkilidir ve en göze çarpan arka plan değişkenleri eğitim sistemleri arasında farklılık gösterse de öğrenci başarısındaki önemli farklılık sosyoekonomik statü (SES) ile açıklanmaktadır (Broer vd. 2019: 2).

Sosyoekonomik statü (SES) eğitim çalışmalarında en yaygın kullanılan değişken olsa da tanımı konusunda çok az fikir birliği vardır (Şirin 2005: 417). Sosyoekonomik statü terimi genel olarak bireyin servet, güce erişim veya bunlar üzerindeki kontrolüne dayalı olarak sosyal konumunun bir ölçüsüdür. Sosyoekonomik statü aynı zamanda bireyin eğitim fırsatlarını, belirli işgücü piyasalarına erişimlerini, ve ömür boyu kazançlarını etkiler (Willms, Tramonte 2015: 16). Ulusal Eğitim İstatistikleri Merkezi

(NCES), tarafından sosyoekonomik statü değişkeninin tanımı ve Ulusal Eğitim İlerleme Değerlendirme programı (NAEP) için kullanımı konusunda fikir birliğine varmak için bir panel düzenlenmiştir. Uzmanlar sosyoekonomik statüyü kişinin finansal sosyal kültürel ve beşerî sermaye kaynaklarına erişimi olarak tanımlamıştır. Panelde bir öğrencinin düşük sosyoekonomik statü ölçüsünün aile geliri, ebeveyn eğitim düzeyi ve ebeveyn mesleki durumu olmak üzere üç temel bileşenden oluştuğu sonucuna varılmıştır (NCES 2012: 4).

Uluslararası Matematik ve Fen Eğilimleri Araştırması (TIMSS) ve Uluslararası Öğrenci Değerlendirme Programı (PISA) gibi uluslararası eğitim değerlendirmeleri sosyoekonomik statünün eğitim sürecinde öğrenci performansını etkilediğini göstermektedir. TIMSS sekizinci sınıf öğrenci anketi sosyoekonomik statüyü ebeveyn eğitimi, evdeki kitaplar ve ev eşyaları (evde bilgisayar, internet bağlantısı, çalışma masası ve öğrencinin kendi odası) gibi üç bileşeni içeren evde eğitim kaynakları (HER) ölçeği ile tanımlamaktadır (Erberber vd. 2015; Martin vd. 2020; Mullis vd. 2019).

Birleşmiş Milletler Kalkınma Programına (UNDP) göre düşük sosyoekonomik statüye sahip öğrenciler okul içinde ve okul dışında eşitsizlik unsurlarından etkilenmektedir. Sosyal bilimciler ve eğitim teorisyenleri tarafından düşük sosyoekonomik statü (“dezavantajlı” etiketi de kullanılmaktadır), düşük gelirli ailelerden gelen çocukları ifade etmektedir (Wang vd. 1998: 5). Araştırmalar düşük sosyoekonomik statüye sahip aile ve topluluklardan gelen çocukların akademik becerilerinin yüksek sosyoekonomik gruplardan gelen çocuklara göre daha yavaş geliştiğini göstermektedir (APA 2017).

Araştırmacılar ve politika yapıcılar yaygın olarak sosyoekonomik statünün öğrencilerin akademik başarısını nasıl etkilediğini incelemektedir. 1960’larda eşitlik ve eğitim fırsatları üzerinde yayınlanan Coleman raporundan sonra uzun süredir devam eden uluslararası ve ülkeye özgü çalışmalar öğrencilerin sosyoekonomik geçmişleri ile okuldaki başarıları arasında önemli bir ilişki olduğunu ortaya koymaktadır (Coleman 1968; OECD 2019). OECD ülkeleri genelinde, öğrenciler arasındaki performans farklılıklarının yaklaşık %15’i öğrencilerin sosyoekonomik durumlarındaki eşitsizlikler tarafından açıklanmaktadır (OECD 2012: 35). Sosyoekonomik statünün, öğrencilerin akademik başarısının güçlü bir açıklayıcısı olması, eğitim sistemlerinin “fırsatları eşitleyici” olma yeteneğini sorgulatmaktadır (Agasisti vd. 2021: 750). Ayrıca, farklı eğitim sistemlerinde çeşitli ampirik çalışmalar düşük sosyoekonomik statüye sahip öğrencilerin, yüksek sosyoekonomik statüye sahip akranlarına göre genellikle daha düşük akademik başarıya sahip olma olasılığının yüksek olduğunu bulmuştur (Coleman 1968; Finn, Rock 1997; OECD 2011a; OECD 2012; OECD 2018; OECD 2019; Şirin 2005; UNDP 2019;TIMSS 2020). OECD ülkeleri genelinde yüksek sosyoekonomik statüye sahip bir öğrenci düşük sosyoekonomik statüye sahip öğrenciye göre matematikte 39 puan (bir yıllık örgün eğitime eşdeğer) daha yüksek puan almaktadır (OECD 2012: 36). Düşük sosyoekonomik statüye sahip öğrenciler ve

performans arasındaki bu ilişkiden dolayı bu öğrenciler akademik anlamda başarılı olabilmek için çok daha fazla çaba harcamak zorunda kalmaktadır.

Düşük sosyoekonomik statüye sahip öğrenciler genellikle daha düşük başarı şansı ile ilişkilendirilmesine rağmen bu öğrencilerin tamamı sosyoekonomik dezavantajla ilişkili koşullara karşı eşit derecede savunmasız değildir. Bazı öğrenciler sosyoekonomik olarak dezavantajlı bir geçmişe sahip olmalarına rağmen “zorlukların üstesinden gelerek” yüksek başarı elde etmektedir. Akademik dayanıklı öğrenciler, düşük sosyoekonomik statüye rağmen, onlara karşı olan zorlukların üstesinden gelen ve yüksek akademik performansı sürdürebilen öğrencilerdir (Agasisti vd. 2016; Agasisti, Longobardi 2014; Borman, Overman 2004; Erberber vd. 2015; Kalender 2015; OECD 2010; OECD 2011a; OECD 2019; Sandoval-Hernández, Białowolski 2016; Sandoval-Hernández, Cortes 2012).

Dayanıklılık öncelikle psikiyatristler, klinik psikologlar ve gelişim psikologları tarafından incelenmiştir. Başlangıçta, araştırmalar dayanıklı çocukların özelliklerini belirlemeye odaklanmıştır. 1980'lerde eğitim araştırmalarında paralel bir gelişme ortaya çıkmıştır. Araştırmacılar, birden çok risk faktörü göz önüne alındığında, "zorlukların üstesinden gelen" öğrencilerinin öğrenme başarılarında beklenenden daha yüksek düzeylere ulaşmada daha etkili olduklarını fark etmeye başladılar. Eğitim araştırmalarında dayanıklılık genellikle “akademik dayanıklılık” veya “eğitimsel dayanıklılık” bağlamında ele alınmaktadır. Wang vd. (1994) akademik dayanıklılığı, koşullar ve deneyimlerin getirdiği kişisel zayıflıklar ve çevresel olumsuzluklara rağmen okulda başarı olasılığının artması şeklinde tanımladılar (Wang vd. 1994: 6).

Dayanıklılık terimi eğitim araştırmalarında Uluslararası Öğrenci Değerlendirme Programı PISA (2009)'dan beri kullanılmaktadır (Agasisti vd. 2018; OECD 2010). Akademik dayanıklı öğrencilerin tanımı eğitim literatüründe farklılık gösterse de akademik dayanıklılık üzerine yapılan çalışmalarda dayanıklı öğrencileri belirlemek için öğrencilerin sahip olduğu risklerin veya dezavantajlı durumunun bir göstergesi olarak sosyoekonomik statü (SES) değişkeni ve pozitif uyum göstergesi olarak eğitimsel çıktılar (başarı testlerindeki göreceli sonuçlar) kullanılmaktadır (OECD 2011a; Ye vd. 2021: 170). Amerikan Araştırma Enstitüsü tarafından hazırlanan raporda Erberber vd. (2015) TIMSS değerlendirmesinde akademik dayanıklı öğrencileri tanımlamak amacıyla düşük sosyoekonomik statülü öğrenciler için evde eğitim kaynakları (HER) endeksinin “az kaynak” kategorisini ve pozitif uyum göstergesi olarak “orta düzey” uluslararası bir başarı ölçüsünü kesme noktası olarak belirlemiştir. Böylece hem sosyoekonomik statü göstergesi hem de olumlu uyum göstergesi açısından uluslararası bir akademik dayanıklı öğrenci tanımı yapılmıştır (Erberber vd. 2015: 3).

OECD ülkelerindeki öğrencilerin yaklaşık %6'sı (yaklaşık bir milyon öğrenci) dayanıklıdır (OECD 2013: 36). Performans açısından değerlendirildiğinde OECD ülkelerindeki dayanıklı öğrenciler aynı

ülkedeki ortalama bir öğrenciden daha iyi performans göstermektedir (OECD 2011a: 26). Ülkeler içindeki dayanıklı öğrencilere bakıldığında, öğrenci performansında eşitliğin daha fazla olduğu eğitim sistemlerinde, daha dezavantajlı öğrenciler bulundukları ülkelerdeki en iyi öğrenciler arasında performans gösterdikleri görülmektedir (OECD 2018: 29). Raporlar tüm ülkelerde düşük sosyoekonomik statüye sahip öğrencilerin ekonomik ve sosyal dezavantajlarını aşma ve daha yüksek sosyoekonomik statüye sahip akranlarına benzer hatta daha iyi performans gösterme potansiyeline sahip olduğunu göstermektedir.

Öğrencilerin geçmişinin eğitim çıktıları üzerindeki etkisini azaltarak eğitimde eşitliği sağlamak önemli bir konu olduğu için akademik dayanıklılığın anlaşılması eğitimcilerin ve politika yapıcılarının, dezavantajlı bir geçmişe sahip öğrencileri akademik performanslarını geliştirmede desteklemelerine yardımcı olabilir (Ye vd. 2021: 170). Akademik dayanıklılık sosyoekonomik eşitsizliklerin neden olduğu başarı boşluklarını kapamak açısından önemli bilgiler sağlamaktadır bu nedenle ülkelerde öğrencilerin akademik dayanıklılığını destekleyen faktörlerin belirlenmesi kritik öneme sahiptir (OECD 2011). Günümüzde 18 yaşın altında yaklaşık 1,2 milyar çocuk yoksulluk içinde yaşadığı için bu konu özellikle önemlidir (UNICEF 2021). Akademik dayanıklılığın desteklenmesinin eğitimde fırsat eşitliğinin sağlanmasına yardımcı olabileceği düşünülmektedir (Chirkina vd. 2020: 1245). Eğer politika yapıcılar ve eğitimciler bu öğrencilerin sosyoekonomik dezavantaja rağmen başarılı olmalarını sağlayan faktörleri anlayabilirse benzer düşük sosyoekonomik statüye sahip öğrencilerin performanslarını iyileştirmek için daha doğru adımlar atabilirler. Düşük sosyoekonomik statüye sahip öğrencilerin performanslarının artırılması hem öğrencilerin hem de eğitim sistemlerin eşitliğine fayda sağlayacağı düşünülmektedir (Erberber vd. 2015).

Dezavantajlı öğrencilerin zorlukların üstesinden gelerek yüksek akademik sonuçlara ulaşımaya ulaşamayacaklarını belirleyen faktörleri anlamak, uzun süredir araştırmacıların ilgisini çekmektedir. Eğitim psikolojisindeki araştırmaların düşük sosyoekonomik statüye sahip öğrencilerin iyi bir akademik performans elde etme yetenekleriyle ilişkili kişilik özelliklerini inceledikleri görülmektedir. Finn ve Rock (1997), Amerika Birleşik Devletleri'nde düşük gelirli evlerden gelen öğrenci grubunu incelemiş ve öz saygı ve kontrol duygusunun öğrencilerin dayanıklı olma olasılığıyla güçlü ilişki gösteren kişisel özellikler olduğunu tespit etmiştir. Ayrıca öğrenme etkinliklere aktif katılımının akademik dayanıklılığın önemli bir bileşeni olduğunu vurgulamaktadır. Borman ve Overman (2004), düşük sosyoekonomik statüye sahip matematikte başarılı ilköğrencilerini dayanıklı olmayan akranlarından ayıran dört temel bireysel özellik tanımlamaktadır: bağlılık, öz yeterlik, okula karşı olumlu tutum ve öz saygı. Martin ve Marsh (2006), düşük sosyoekonomik statüye sahip ancak başarılı öğrenciler tarafından paylaşılan psikolojik faktörü modelleyerek (5-C modeli olarak bilinen)

dayanıklılıkla ilişkili beş bireysel faktör tanımladılar: güven (öz yeterlilik), koordinasyon (planlama), kontrol, soğukkanlılık (düşük kaygı) ve bağlılık (sebat).

Psikoloji literatüründe dayanıklılık genellikle çocuk gelişimi araştırmaları bağlamında kullanılmıştır. OECD (2010) raporu, sosyoekonomik olarak dezavantajlı bir geçmişe sahip ancak yüksek performans gösteren öğrencilerin akademik başarılarını inceleyerek, performans ve sosyoekonomik geçmiş arasında ilişki kuran bir literatür bağlamında akademik dayanıklılık terimini eğitim araştırmalarında popüler hale getirdiği görülmektedir. Daha sonra OECD (2011) raporunda öğrencilerinin dayanıklılığının bir tanımının yapılmasıyla birlikte eğitim araştırmalarında büyüyen bir literatür oluşmuştur. Diğer taraftan, TIMSS değerlendirmesinde uluslararası bir dayanıklı öğrenci tanımı yapan Erberber vd. (2015) raporda 28 eğitim sistemi için dayanıklılıkla ilişkili faktörleri lojistik regresyon yardımı ile incelemiştir. Bulgular öğrencilerin eğitim beklentileri, matematiğe değer vermeleri, daha az akran zorbalığına maruz kalması ve öğretmenin öğrenciye olan güvenine ilişkin inançlarının dayanıklılığı açıklayan önemli faktörler olduğunu göstermiştir. TIMSS 2011 değerlendirmesi için Erberber vd. (2015) dayanıklılık tanımını takip eden Sandoval-Hernández ve Białowolski (2016) çalışmalarında beş Asya eğitim sistemini ele almışlardır. Lojistik regresyon sonuçları, öğrencilerin matematiğe yönelik olumlu tutumlarının, yüksek eğitim beklentilerinin ve matematik ödevi için harcanan sürenin daha yüksek akademik başarı olasılığı ile ilişkili olduğunu göstermiştir. Benzer bölgeler, yedi Doğu Asya ülkesi için Clavel vd. (2022), PISA 2015 değerlendirmesini kullanarak akademik dayanıklılık ve bilişsel olmayan beceriler arasındaki ilişkileri hiyerarşik lojistik regresyon yardımıyla incelediği çalışmalarında, bilimden zevk almanın ve bilime ilgi duymanın dayanıklılık ile olumlu yönde ilişkili olduğu vurgulanmıştır. Cheung (2017) ise PISA 2012 değerlendirmesinde yüksek performans gösteren ilk beş Asya ekonomisi için öğrenme değişkenlerinde dayanıklılığın matematik okuryazarlığı performansı üzerindeki etkilerini lojistik regresyon yardımı ile incelediği çalışmada matematik öğrenmede öz-yeterlik ve öz-kavram değişkenlerinin akademik dayanıklılıkla güçlü ilişkisi olduğu tespit etmiştir. Wang vd. (2022), Hong Kong PISA 2018 değerlendirmesini kullanarak, makine öğrenimi (rassal orman sınıflandırması) yöntemiyle otuz farklı kişisel ve sosyal-bağlamsal faktör içerisinde dayanıklılığı en iyi tahmin eden değişkenlerin okumadan keyif alma, öğretmen yönlendirmeli öğretim, okula ait hissetme, okul iklimi, öz-yeterlik ve algılanan öğretmen ilgisi olduğunu belirlemişlerdir.

İtalyan öğrencilerin akademik dayanıklılığı etkileyen okul düzeyindeki değişkenlerin rolüne odaklanan Agasisti ve Longobardi (2014), öğrenciler ve öğretmenler arasındaki ilişkinin (öğretmenler ve öğrenciler arasındaki) dayanıklılığı açıklayan güçlü bir değişken olduğunu ve bireysel faktörler olarak özgüven ve motivasyonun yüksek başarı ile ilişkili olduğunu tespit etmişlerdi. Benzer şekilde PISA 2009 verilerini kullandıkları diğer çalışmalarında Agasisti ve Longobardi (2016) uluslararası bir bakış açısı ile on beş Avrupa ülkesinde müfredat dışı etkinliklerin ve olumlu bir okul ikliminin dayanıklılıkla

ilişkili en önemli değişkenleri olduğunu bulmuşlardır. Güncel çalışmalarında Agasisti vd., (2021), 18 ülke için dayanıklı öğrencilerin başarılı olma olasılığı ile ilişkili okul faktörlerini hiyerarşik lojistik regresyon analizi sonucunda okul politikalarının ve uygulamalarının, dezavantajlı öğrencilerin başarılı olma olasılığını etkileyebileceği ve dayanıklı olma olasılıklarını artırabileceği vurgulanmaktadır. García-Crespo vd. (2022), TIMSS 2019'a katılan 22 Avrupa Birliği ülkesini dayanıklılıkla ilişkili faktörleri üç düzeyli hiyerarşik lojistik model ile incelediği çalışmalarında bu ülkelerde matematik ve fen bilimlerinde akademik dirençli öğrencilerin ortalama yüzdeleri benzer olduğunu ve öğrenci güveni, cinsiyet, öğrencilerin göçmenlik geçmişinin dayanıklılığı etkileyen en önemli faktörler olduğunu vurgulamışlardır. Norveç ve İsveç iki İskandinav ülkesindeki akademik dayanıklı öğrencilere odaklanan Radišić ve Pettersen (2021), TIMSS 2015 verilerini kullanarak dayanıklı öğrencilerin öz inanç profillerini dayanıklı olmayan farklı öğrenci gruplarıyla karşılaştırmıştır. Her iki ülkede de dayanıklı öğrencilerin büyük bir kısmı matematikte kendine güvenmektedir ve güven değişkenin dayanıklılıkla ilgili önemli bir faktör olduğu vurgulanmıştır.

Türkiye, dayanıklı öğrenci oranlarının en yüksek olduğu ilk sekiz ülke arasında yer almaktadır bu bakımdan dayanıklı öğrencilerin bir incelenmesinin sunulması için iyi bir örnek olabilir (Erberber vd. 2015; OECD 2013). Fındık ve Kavak (2013), dayanıklı öğrencilerin okuma başarısını PISA 2009 değerlendirmesi ile genel tarama modeli kullanılarak incelemişlerdir. Genel olarak dayanıklı öğrencilerin üçüncü yeterlilik düzeyine dezavantajlı öğrencilerin ise ikinci yeterlilik düzeyine ulaştığını tespit etmişlerdir. Ayrıca çalışmada dayanıklı kız öğrenci oranının erkek öğrenci oranından fazla olduğu bulgusuna ulaşılmıştır. Kalender, (2015) düşük sosyoekonomik statüye sahip düşük başarılı ve dayanıklı öğrenciler arasındaki performans farklılıklarını etkileyen okul ve öğretmen faktörlerini incelediği çalışmada, lojistik regresyon analizi sonuçlarına göre dayanıklı öğrencilerin okul ve öğretmene karşı olumlu tutumlara sahip olduğunu tespit etmiştir. Ardından dayanıklı öğrenciler ve PISA 2012 değerlendirmesinde ölçülen aidiyet duygusu değişkeni arasındaki ilişkiyi karar ağacı algoritmalarından biri olan Ki-kare Otomatik Etkileşim Dedektörü (CHAID) analizi ile inceleyen Aydın ve Kalender (2015) akademik dayanıklılık ile aidiyet duygusu arasında ilişki olduğunu göstermişlerdir. Benzer şekilde PISA 2012 değerlendirmesini kullanan Önder ve Uyar (2018), matematik başarısı açısından düşük başarılı ve dayanıklı öğrencilerinin öğrenci özelliklerini yapısal eşitlik modeli yarımıyla inceledikleri çalışmalarında düşük başarılı öğrenci performansını okula yönelik tutumun, dayanıklı öğrencilerin performansını ise matematiğe yönelik duyuşsal özelliklerin pozitif yönde etkilediğini bulmuşlardır. Özden ve Atasoy (2020), dayanıklı öğrencilerin okul hakkındaki beklentilerini nitel araştırma yöntemleri ve görüşme teknikleri kullanarak incelemişlerdir. Dayanıklı öğrencilerin zorlukların üstesinden nasıl geldiklerini anlamak için içsel motivasyon, bağlılık, öz güven ve öz yönetim gibi psikolojik çerçeve arka planlarına sahip özelliklerine odaklanmanın önemi tespit edilmiştir. Son

olarak Aydın ve Erdem (2022), TIMSS 2015 matematik puanlarına dayalı olarak dayanıklı Türk ve İtalyan öğrencilerini inceledikleri çalışmalarında lojistik regresyon sonuçları akademik beklentiler ve matematiğe olan güvenin bir öğrencinin dayanıklı olma ihtimalini artıran en önemli iki değişken olduğunu göstermiştir.

Literatürde görüldüğü gibi akademik dayanıklılık üzerine hem teorik hem de ampirik birçok çalışma mevcuttur. Çalışmaların genel olarak benzer metodolojik yaklaşım, klasik hiyerarşik modellerden hareket ettiği görülmektedir. Bu çalışmada literatürdeki benzer metodolojik yaklaşımdan farklı olarak nonparametrik Bayesyen regresyon modeli olan lineer bağımlı Dirichlet süreç modeli kullanılmıştır. Dirichlet süreç karışım modelleri, özellikle gruplanma problemlerinde kullanılan olasılıksal bir modeldir. Bu modeller farklı gruplar üzerinde modeller oluşturulmasına, parametrelerin gruplar arasında farklılaşmasına ve grup düzeyindeki etkileri tahmin ederken belirsizliğin hesaba katılmasına izin vermektedir. Ayrıca Bayesyen yaklaşım model parametreleriyle birlikte rastgele etkileri de tahmin etmektedir. Böylece nonparametrik Bayesyen yaklaşım sayesinde tüm rastgele etkiler desteklenir ve model tabanlı bir kümeleme analizi sağlanır. Nonparametrik Bayesyen model $Y|x$, x 'e bağlıdır böylece yalnızca $Y|x$ ortalamasına odaklanan klasik yaklaşımına kıyasla veri analizine daha zengin bir yaklaşım ve rastgele etkilerin esnek modellemesini sağlamaktadır (Karabatsos 2015).

Diğer taraftan akademik dayanıklılık üzerinde yapılan araştırmalarda, özel olarak seçilmiş bir öğrenci grubu ile ilgilenildiği için genellikle küçük örneklem boyutlarıyla çalışılmaktadır. Bu durum mevcut çalışmalar için bir sınırlılık olarak görülmektedir (Ye vd. 2021). Bu nedenle Bayesyen yaklaşımının kullanılmasının bir nedeni de büyük örneklem teorisine dayanmaması, küçük örneklem için daha doğru tahminler sağlaması ve karmaşık veri yapılarıyla birlikte karmaşık modellerin daha esnek bir şekilde tahmin edilmesine olanak sunmasıdır (König, van de Schoot 2018: 497). Bu özellikle karışım modelinde bazı grupların çok daha küçük örneklem büyüklüğüne sahip olduğu durumlarda fayda sağlamaktadır. Ayrıca çalışma 16 OECD ülkesini ele alarak burada dayanıklılık konusuna uluslararası bir bakış açısı sağlamaktadır. Akademik dayanıklılığın ülkeler arası analizi, farklı ülkelerdeki öğrencilerin istikrarını ve dayanıklı olma özelliklerinin ülkeler arasında farklılık gösterip göstermediğini belirleyebilir.

Çalışmanın literatürden temel metodolojik farkı doğrultusunda bu yaklaşımın diğer yaklaşımlardan farkı ve üstünlükleri maddeler halinde sunulmuştur.

1- Bu tez çalışmasında, genellikle okulda kötü başarı sonuçlarıyla ilişkilendirilen düşük sosyoekonomik statüye sahip olmalarına rağmen zorlukları yenerek akademik olarak başarılı olan dayanıklı öğrenci grubuna odaklanılmaktadır. Dayanıklı öğrenciler düşük sosyoekonomik statüye sahip

öğrencilerin bir alt kümesi olduğu için mevcut çalışmalarda küçük örneklem boyutu bir sınırlılık olarak görülmektedir. İlk olarak Bayesyen yaklaşımın tercih edilmesinin nedeni büyük örneklem teorisine dayanmaması, küçük örneklem için daha doğru tahminler sağlamasıdır.

2- Hiyerarşik modellerin önemli bir özelliği belirli bir gruba atıfta bulunan her parametrenin benzer özelliklere sahip diğer gruba karşılık gelen parametrelerinden güç alabilmesidir. Bu özellikle karışım modellerinde bazı grupların çok daha az gözleme sahip olduğu durumlarda fayda sağlar.

3- Ayrıca tüm Bayes yöntemlerinin altında yatan temel varsayım, modeli belirleyen parametrelerin rastgele bir değişken olmasıdır. Böylece Bayesyen yaklaşım, klasik yaklaşımdan farklı olarak model parametreleriyle birlikte rastgele etkilerin tahmin edilmesine olanak sağlar.

4- Nonparametrik Bayesyen modeller, parametrik modeller gibi veri üretme mekanizması hakkında kısıtlayıcı varsayımlara sahip değildir. Böylece nonparametrik Bayesyen yaklaşım sayesinde tüm rastgele etkiler desteklenir ve esnek model tabanlı bir kümeleme analizi sağlanır.

5- Dirichlet süreç karışım modelleri, özellikle gruplanma problemlerinde kullanılan olasılıksal bir modeldir. Bu karışım modelleri farklı gruplar üzerinde modeller oluşturulmasına, parametrelerin gruplar arasında farklılaşmasına ve grup düzeyindeki etkileri tahmin ederken belirsizliğin hesaba katılmasına izin vermektedir. Böylece klasik yöntemlerin hesaplama sorunuyla karşılaştıkları çok sayıda rastgele etki veya hiyerarşik yapı içeren karmaşık modellerin herhangi bir hesaplama sorunuyla karşılaşmadan analiz edilmesine olanak sağlanır.

Bu bağlamda Çalışmanın amacı, 16 OECD ülkesindeki dayanıklı öğrencilerin akademik başarısını etkileyen faktörleri karmaşık modellerin esnek bir tahmini sunan nonparametrik Bayesyen regresyon modelleri yardımıyla belirlemektir. Böylece daha fazla düşük sosyoekonomik statüye sahip öğrencinin zorlukları yenmesi için neler yapılabileceği konusunda çıkarım yapılacaktır.

4.2. Veri Seti ve Değişkenler

1995'ten beri Uluslararası Eğitim Başarısı Değerlendirme Derneği'nin (IEA) tarafından dört yılda bir yürütülen TIMSS, dördüncü ve sekizinci sınıflarda matematik ve fen başarısındaki uluslararası eğilimlerin bir değerlendirmesini sunmaktadır. Bu tez çalışmasında kullanılan veriler, Uluslararası Matematik ve Fen Eğilimleri Araştırması (TIMSS) 2019 değerlendirmesinde öğrencilerin sosyo-demografik özellikleri ve başarı testi puanlarına ilişkin maddeleri içeren (BSGM7) öğrenci dosyasından elde edilmiştir. Çalışma uluslararası bir bakış açısı sağladığından dayanıklılık için uluslararası bir tanım

seçilmiştir³⁷. Dayanıklı öğrenciler, TIMSS bağlamsal çerçevesine (Mullis vd. 2012) uygun olarak ve Erberber vd. (2015) çalışması takip edilerek akademik dayanıklılık literatürüne uygun olacak şekilde seçilmiştir. Buna göre matematikte başarılı öğrenciler TIMSS değerlendirmesinde “orta düzey” 475 veya üzerinde puan alan öğrenciler olarak tanımlanmaktadır.³⁸ Bu puanı alan öğrenciler ondalık sayıları, kesirleri, oranları ve yüzdeleri içeren problemleri çözebilir, basit cebirsel ilişkileri anlayabilir, iki boyutlu bir çizimi üç boyutlu bir nesneyle ilişkilendirebilir ve temel olasılık kavramını tanımlayabilmektedir (Mullis vd. 2012: 125). Diğer taraftan düşük sosyoekonomik statüye sahip öğrenciler için Evde Eğitim Kaynakları (HER) endeksinin “az kaynak” kategorisi seçilmiştir. Burada az kaynağa sahip öğrenciler, 25 veya daha az kitaba sahip olduklarını ve evde çalışırken desteklenmediğini ve hiçbir ebeveynin ortalama olarak lise eğitiminin ötesine geçmediğini bildirmişlerdir (Mullis vd. 2020). Bu iki değişkenin birleştirilmesi dayanıklı öğrencileri temsil etmektedir. Akademik dayanıklı öğrenciler matematikte başarılı dezavantajlı öğrenciler olduğu için düşük sosyoekonomik statüye sahip öğrencilerin bir alt kümesidir (Erberber vd. 2015).

Analizde ülke grubu olarak TIMSS 2019 sekizinci sınıf matematik değerlendirmesine katılan 16 OECD ülkesi (Avustralya, Şili, Finlandiya, Fransa, Macaristan, İrlanda, İtalya, Japonya, Kore, Litvanya, Yeni Zelanda, Norveç, Portekiz, İsveç, Türkiye ve Amerika Birleşik Devletleri) seçilmiştir.³⁹ Mevcut dayanıklılık tanımı ve eksik veriler için satır silme işlemi uygulandıktan sonra 16 OECD için toplam 1250 dayanıklı öğrenci olduğu tespit edilmiştir. Her ülkede farklı sayıda gözlem bulunmaktadır ve eşit olarak temsil edilmesi gereken alt gruplar olduğundan tabakalı örnekleme yöntemi kullanarak alt örneklem n=610 olarak belirlenmiştir.

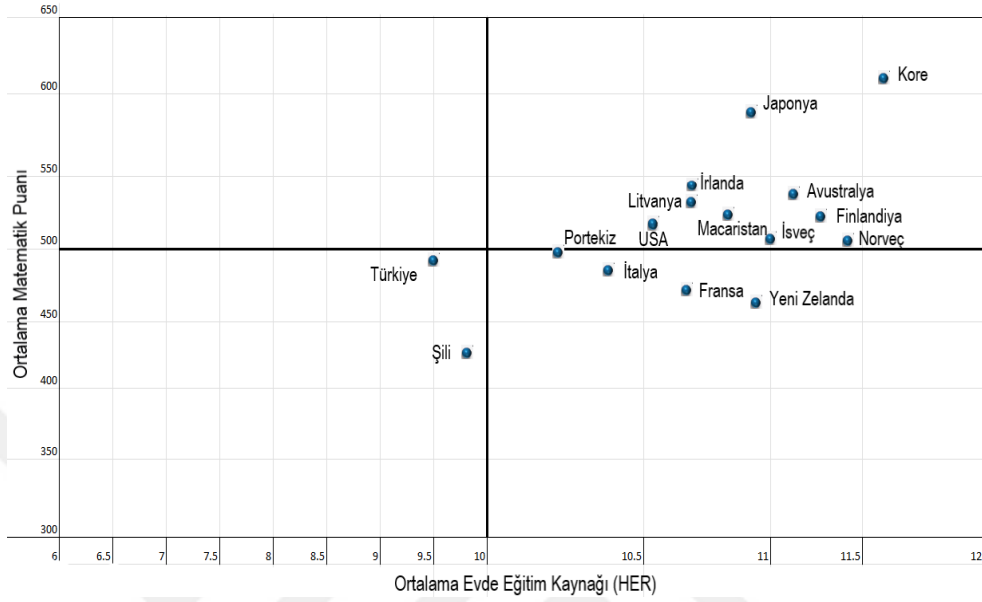
TIMSS 2019 sonuçları, öğrencilerin sosyoekonomik statüsü ile eğitim başarıları arasında güçlü bir pozitif ilişki olduğunu gösteren araştırmaları desteklemektedir. Şekil 2 sekizinci sınıf öğrencilerinin evde eğitim kaynakları ile ülkeler genelinde ortalama matematik başarıları arasında pozitif ilişki olduğunu göstermektedir.

³⁷ Uluslararası büyük ölçekli diğer veri kümelerini kullanan farklı çalışmalar akademik dayanıklılık tanımı için benzer yaklaşımlar kullanmışlardır (Agasist, Longobardi 2014; OECD 2011).

³⁸ Kore ve Japonya için düşük sosyoekonomik statüye sahip öğrencilerin ortalama matematik puanı orta düzey (475) puandan yüksektir. Dayanıklılığın ortalamadan daha iyi bir başarı puanı yansıttığını göstermek için bu iki ülkede dayanıklı öğrencileri belirlemek için düşük düşük sosyoekonomik statüye sahip öğrencilerin ortalama puanı (orta düzey 475 puan yerine) kullanılmıştır.

³⁹ <https://nces.ed.gov/timss/participation.asp#fr3>. İsrail, öğrenci dosyasında eksiklikler olduğu için analize dahil edilmemiştir.

Şekil 2: Evde Eğitim Kaynaklarına Göre Ortalama Matematik Başarısı



Kaynak: TIMSS 2019 sekizinci sınıf matematik değerlendirmesi (Mullis vd. 2020).

Sosyoekonomik statü göstergesi olarak evde eğitim kaynakları ülkeler arasında büyük farklılık göstermektedir. Şekil 2 incelendiğinde 16 OECD ülkesi içerisinde Türkiye'nin evde eğitim kaynakları bakımından dezavantajlı konumda olmasına rağmen diğer ülkelere göre nispeten yüksek başarı gösterdiği söylenebilir. TIMSS 2019 sekizinci sınıf matematik değerlendirmesi incelendiğinde çok kaynağa sahip öğrencilerin ortalama matematik başarısının az kaynağa sahip öğrencilerle karşılaştırıldığında 113 puanlık bir fark olduğu görülmektedir (Mullis vd. 2020).

Tablo 3, 16 OECD ülkesi için düşük sosyoekonomik statüye sahip sekizinci sınıf öğrenci yüzdesini ve bu öğrenciler arasındaki akademik dayanıklı öğrenci yüzdeslerini sunmaktadır. Dayanıklı öğrenci oranı eğitimde eşitlik bakımından değerlendirildiğinde genellikle etkili bir eğitim sistemi göstergesi olarak görülmektedir (Erberber vd. 2015; OECD 2016).

Tablo 3: Ülkelere Göre Düşük SES'e Sahip Öğrencilerin ve Akademik Dayanıklı Öğrencilerin Yüzdesi

Eğitim Sistemi	TIMSS 2019 Ortalama Matematik Puanı	Düşük SES'e sahip öğrenci yüzdesi	Düşük SES'e sahip öğrenciler arasında dayanıklı öğrencilerin yüzdesi	Düşük- SES% Akademik Dayanıklı Öğrenci%
Avustralya (AUS)	517	4	28	
Şili (CHL)	441	16	9	
Finlandiya(FIN)	509	2	30	
Fransa (FRA)	483	7	23	
Macaristan (HUN)	517	5	16	
İrlanda (IRL)	524	6	36	
İtalya (ITA)	497	13	34	
Japonya (JPN)*	594	3	52	
Kore (KOR)*	607	2	50	
Litvanya (LTU)	520	3	25	
Yeni Zelanda (NZL)	482	6	28	
Norveç (NOR)	503	2	28	
Portekiz (PRT)	500	16	38	
İsveç (SWE)	503	4	29	
Türkiye (TUR)	496	33	32	
ABD (USA)	515	8	39	

Kaynak: TIMSS 2019 sekizinci sınıf matematik değerlendirmesine katılan öğrenci verileri kullanılarak yazar tarafından oluşturulmuştur. * Kore ve Japonya için düşük sosyoekonomik statüye sahip öğrencilerin ortalama matematik puanı orta düzey (475) puandan yüksektir. Erberber vd. (2015) takip edilerek dayanıklılığın ortalamadan daha iyi bir başarı yansıttığını göstermek için bu iki ülkede dayanıklı öğrencileri belirlemek için düşük sosyoekonomik statüye sahip öğrencilerin ortalama puanı (orta düzey 475 puan yerine) kullanılmıştır.

Görüldüğü gibi OECD ülkeleri arasında hem düşük sosyoekonomik statüye sahip öğrenci yüzdesi hem de dayanıklı öğrenci yüzdesi bakımından büyük farklılıklar görülmektedir. Örneğin, Finlandiya, Kore ve Norveç'te düşük sosyoekonomik statüye sahip öğrenci oranının %2 olduğu diğer taraftan Türkiye'ye bakıldığında bu oranın %33'ye kadar arttığı kadar görülmektedir. Genel olarak, dayanıklı öğrencilerin en yüksek payına sahip olduğu ülkelerde (Japonya, Kore, ABD, Finlandiya, İrlanda, Litvanya, Avustralya, Norveç ve İsveç), düşük sosyoekonomik statüye sahip öğrenci yüzdesinin düşük olduğu

görülmektedir. Bu bağlamda eğitim sistemleri için dayanıklı öğrenci oranı eğitimde eşitliğin bir göstergesi olarak değerlendirilebilir.

Diğer taraftan Portekiz ve Şili benzer düşük sosyoekonomik statüye sahip öğrenci yüzdesine (%16) sahipken bu ülkelerde dayanıklı öğrenci yüzdeleri (sırasıyla %38 ve %9) arasında büyük farklılıklar olduğu görülmektedir. Ayrıca dayanıklı öğrencilerin yüksek yüzdesine sahip eğitim sistemlerinin bazı istisnalar olsa da genel olarak daha yüksek ortalama matematik puanlarına sahip olduğu görülmektedir. 16 OECD ülkesi içerisinde Türkiye'nin en yüksek dezavantajlı öğrenci yüzdesine sahip (%33) olduğu ancak bu öğrencilerin nispeten yüksek bir yüzdesinin (%32) dayanıklı olduğu görülmektedir. Benzer şekilde Portekiz nispeten yüksek dezavantajlı yüzdesine (%16) sahipken yüksek bir dayanıklılık yüzdesine (%38) sahiptir. Genellikle eğitim sistemleri için yüksek dezavantajlı öğrenci yüzdesine sahip olmak bir zorluk gibi görülse de bu zorluğun üstesinden geldiği durumların olduğu görülmektedir. Bu bağlamda bu tez çalışmasında 16 OECD ülkesinde akademik dayanıklılığı etkileyen faktörler araştırılmaktadır.

Kapsamı farklı çalışmalar okullarla ilgili faktörlerin ortalama olarak öğrenci başarısı üzerinde öğrencilerin bireysel özelliklerinden çok daha düşük bir etkiye sahip olduğunu göstermektedir (Agasisti vd. 2016: 919). Diğer taraftan okul farklılıkları genel öğrenci performans değişikliğinin %30'unu temsil ederken öğrenci farklılıkları performans değişikliğinin %45'ni temsil ettiği görülmektedir (OECD 2010). Bu nedenle çalışma temelde dayanıklılığı açıklayan faktörler olarak öğrenci bireysel özelliklerine odaklanmaktadır.

Matematik başarısı STEM (bilim, teknoloji, mühendislik ve matematik) konularının merkezindedir ve teknoloji, ekonomi, sağlık ve çevre gibi alanlarda sürdürülebilir kalkınma için hayati önem taşımaktadır (Scherer, Nilsen 2016). Sosyoekonomik geçmişe bağlı başarı farklılıklarının öğrenci yaşamları boyunca (ergenlik ve çocukluk döneminde) arttığı görülmektedir (OECD 2018: 54). Bu çalışmada genel olarak ortaokullardaki matematik başarı farkı ilkokullara göre daha geniş olduğu görülmektedir (Takashiro 2015). Bu bağlamda sekizinci sınıf öğrencilerinin matematik başarısı bağımlı değişken olarak belirlenmiştir. TIMSS matematik başarısı için 500 ortalama ve 100 standart sapma olacak şekilde BSMMAT01'den BSMMAT05'e beş makul değer hesaplanmaktadır. Matematik başarısı olarak kullanılan bağımlı değişken (MAT) bu beş puanın ortalaması alınarak oluşturulmuştur.

Akademik dayanıklılığı etkileyen değişkenlerin seçimi literatür taramasına dayanmaktadır (Clavel vd. 2022; Erberber vd. 2015; Sandoval-Hernández ve Białowolski 2016; Sandoval-Hernández ve Cortes 2012). Bu bağlamda çalışmada kullanılan bağımsız değişkenler Tablo 4'te ayrıntılı olarak açıklanmıştır.

Tablo 4: Değişken Tanımları

Değişken	Tanım	Değer
Bağımlı Değişken		
MAT	Ortalama matematik puanı	(BSMMAT01- BSMMAT05 ortalama değer)
Bağımsız Değişkenler		
ITSEX	Öğrencinin cinsiyeti	1(Kız)-0(Erkek)
BSDAGE	Öğrencinin yaşı	Minimum (9)-Maksimum (19)
BSBG07	Öğrencinin akademik beklentisi	1(Orta öğretim)-6(Yüksek Lisans/Doktora)
		Eğitiminizde ne kadar ilerlemeyi düşünüyorsunuz?
BSBM19G	Pozitif öğretmen beklentisi	1(Pek katılmıyorum)-4(Çok katılıyorum)
		Öğretmenim bana matematikte iyi olduğumu söyler
BSDGSSB	Öğrencinin okula aidiyet duygusu	Ölçek değeri
		1) Okulda olmayı seviyorum
		2) Okuldayken kendimi güvende hissediyorum
		3) Kendimi bu okula ait hissediyorum
	Beş maddeye verilen yanıtlar kullanılarak oluşturulmuş ölçek	4) Okulumdaki öğretmenler bana karşı adil davranıyor
		5) Bu okula gitmekten gurur duyuyorum
BSBSGSB	Öğrencinin akran zorbalığa maruz kalması	Ölçek değeri
		1) Fiziksel görünüşüm hakkında kötü şeyler söylenir
		2) Hakkımda yalanlar yayıldı
		3) Sırlarım başkalarıyla paylaşıldı
		4) Benimle konuşmayı reddetti
		5) Ailemden birine hakaret edildi
		6) Benden bir şey çalındı
		7) Bana yapmak istemediğim şeyleri yaptırdı
	On dört maddeye verilen yanıtlar kullanılarak oluşturulmuş ölçek	8) Çevrimiçi olarak bana kötü veya incitici mesajlar gönderdi
		9) Çevrimiçi olarak hakkımda kötü veya incitici şeyler paylaştı
		10) İnternette utanç verici fotoğraflarım paylaşıldı
		11) Beni tehdit etti
		12) Fiziksel olarak beni incitti
		13) Beni gruplarından dışladılar (partiler, mesajlaşma)
		14) Bana ait bir şeye bilerek zarar verildi
BSBGSCM	Öğrencinin matematiğe olan güveni	Ölçek değeri

Dokuz maddeye verilen yanıtlar kullanılarak oluşturulmuş ölçek	1) Genelde matematikte başarılıyım
	2) Matematik benim için birçok sınıf arkadaşşıma göre daha zor
	3) Matematik güçlü yanlarımdan biri deęil
	4) Matematikte her şeyi çabuk öğrenirim
	5) Matematik beni geriyor
	6) Zor matematik problemlerini çözmekte iyiyimdir
	7) Öğretmenim bana matematikte iyi olduğumu söylüyor
	8) Matematik benim için dięer konulardan daha zor
	9) Matematik kafamı karıştırıyor

Kaynak: Ölçeklere ilişkin tanımlamalar TIMSS 2019 Teknik rapordan alınmıştır (Martin vd. 2020).

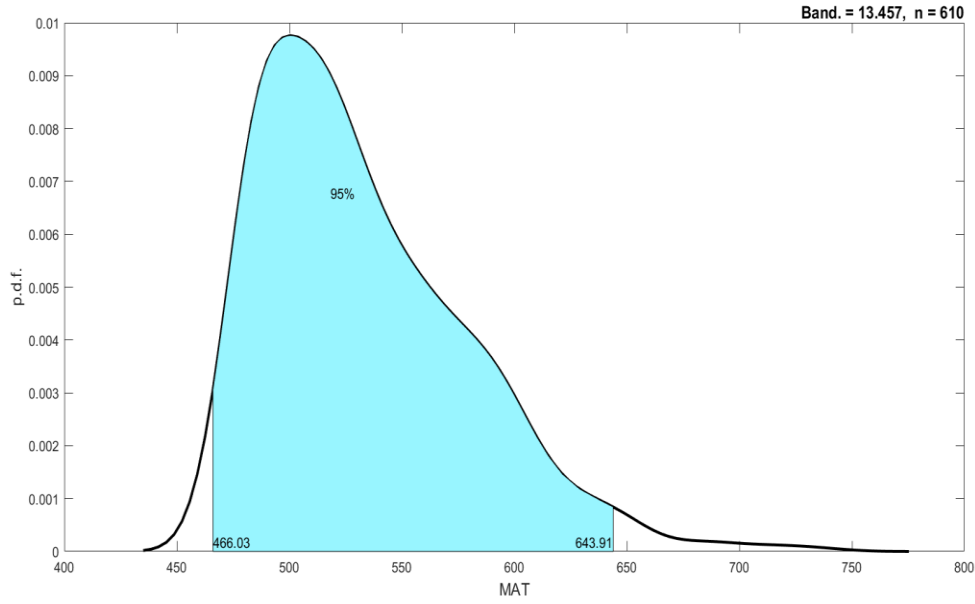
Analizde kullanılacak bağımlı deęişken ortalama matematik başarışı (MAT) ve yukarıda tanımlı yapılan bağımsız deęişkenlere ait tanımlayıcı istatistikler Tablo 5’te sunulmaktadır.

Tablo 5: Tanımlayıcı İstatistikler

Deęişkenler n=610	Ortalama	Standart Sapma	Min.	Q25	Medyan	Q75	Maks.	Çarpıklık	Basıklık
MAT	533.204	45.82	475.45	496.61	522.1	560.48	734.94	1.058	4.125
ITSEX	0.489	0.5	0	0	0	1	1	0.046	1.002
BSDAGE	14.123	0.572	12.92	13.75	14.08	14.42	18	1.448	9.063
BSBG07	4.377	1.519	1	3	5	6	6	-0.672	2.152
BSBM19G	3.32	0.821	1	3	4	4	4	-1.094	3.568
BSBGSSB	9.813	1.976	3.924	8.461	9.408	10.71	13.272	0.102	2.65
BSBGSB	10.074	1.732	1.953	8.847	9.928	11.188	12.937	0.287	2.86
BSBGSCM	10.287	2.043	3.278	9.226	10.121	11.418	15.852	0.333	4.075

Bağımsız deęişkenlerin bağımlı deęişken üzerindeki etkisinin karşılaştırılabilirliğini ve tutarlılığını sağlamak için veri setindeki deęişkenler için (sıfır ortalama ve bir standart sapma) standartlaştırma işlemi yapılmıştır. Bu dönüşüm karışım modellerinde eğim katsayılarını ortak bir ölçekte yorumlanabilir hale getirmek için yapılmıştır (Karabatsos 2017, 2020; Sanyal, Ehlen 2017). 16 OECD ülkesi için akademik dayanlı öğrencilerin matematik başarılarının kernel yoğunluk tahmini Şekil 3’te görülmektedir.

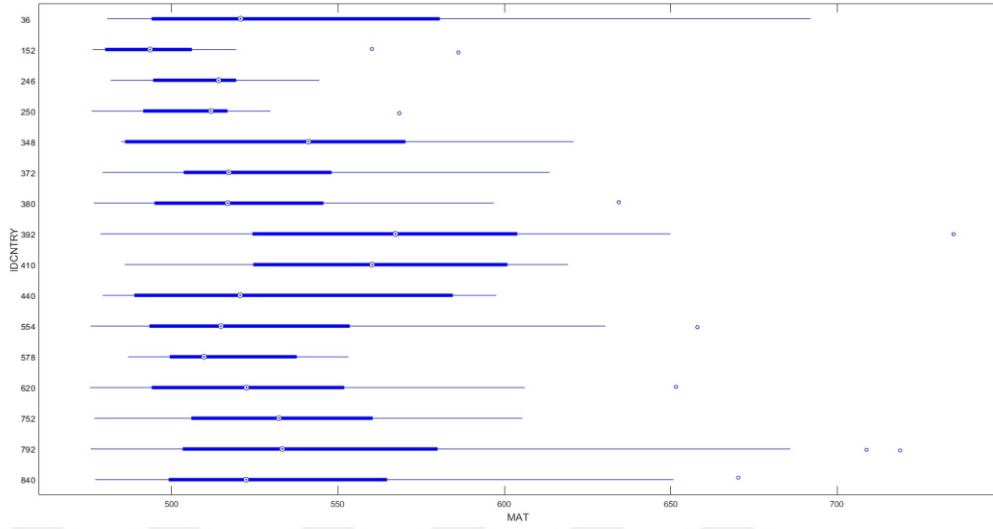
Şekil 3: Dayanıklı Öğrencilerin Matematik Başarısının Kernel Yoğunluk Tahmini



Kernel yoğunluk tahminine bakıldığında verilerin normal dağılmadığı, sağa çarpık (pozitif) bir dağılım takip ettiği görülmektedir.⁴⁰ Bu durum birçok dayanıklı öğrencilerin nispeten düşük puana sahip olduğunu ve giderek daha az sayıda öğrencinin yüksek puan aldığını göstermektedir. Dayanıklı öğrencilerin matematik puanlarının 533 puan civarında yoğunlaştığı ve matematik başarısı için güven aralığının [466, 643] olduğu görülmektedir. Dayanıklı öğrencilerin matematik puanlarının ülkelere göre dağılımı Şekil 4’te sunulmuştur. Şekil incelendiğinde bazı ülkelerde aykırı değerlerin olduğu görülmektedir.

⁴⁰ Ek 1’de Normallik testine ilişkin Kolmogorov-Smirnov testi ve Q-Q grafiği çıktıları sunulmaktadır.

Şekil 4: Ülkelere Göre Dayanıklı Öğrencilerin Matematik Puan Dağılımı



Ayrıca analizde incelenen 16 OECD ülkesi için dayanıklı öğrencilere ait özet istatistikler Tablo 6’da sunulmuştur.

Tablo 6: Ülkelere Göre Tanımlayıcı İstatistikler

Ülkeler	Standart						
	Ortalama	Sapma	Medyan	Min.	Q25	Q75	Maks.
Avustralya (AUS)	537.755	54.038	520.735	480.62	493.99	580.56	691.93
Şili (CHL)	499.192	26.589	493.535	476.27	480.03	506.07	586.17
Finlandiya (FIN)	510.553	18.235	514.11	481.66	494.43	519.37	544.33
Fransa (FRA)	508.536	21.758	511.805	475.98	491.45	516.79	568.41
Macaristan (HUN)	537.458	48.899	541.095	484.84	486	570.25	620.79
İrlanda (IRL)	524.609	33.679	517.145	479.21	503.62	548.03	613.61
İtalya (ITA)	522.504	34.61	516.82	476.65	494.825	545.615	634.35
Japonya (JPN)	569.318	58.572	567.26	478.65	524.27	603.85	734.94
Kore (KOR)	558.204	46.553	560.235	485.97	524.535	600.88	619.08
Litvanya (LTU)	528.824	46.159	520.6	479.3	488.8	584.47	597.56
Yeni Zelanda (NZL)	526.941	47.432	514.83	475.62	493.325	553.53	657.97
Norveç (NOR)	516.387	22.569	509.725	486.9	499.44	537.585	553.15
Portekiz (PRT)	525.511	37.169	522.46	475.45	494.018	551.865	651.57
İsveç (SWE)	535.952	40.812	532.22	476.74	505.917	560.393	605.37
Türkiye (TUR)	543.224	50.663	533.235	475.66	503.31	579.91	718.91
ABD (USA)	534.735	45.006	522.37	477.05	499.12	564.72	670.25

Analizde ele alınan dayanıklı öğrencilerin Japonya ve Kore’de ortalama matematik başarısının beklendiği gibi yüksek olduğu, Şili’de ise dayanıklı öğrencilerin 16 OECD ülkesi içerisinde en düşük başarı puanına sahip olduğu görülmektedir.

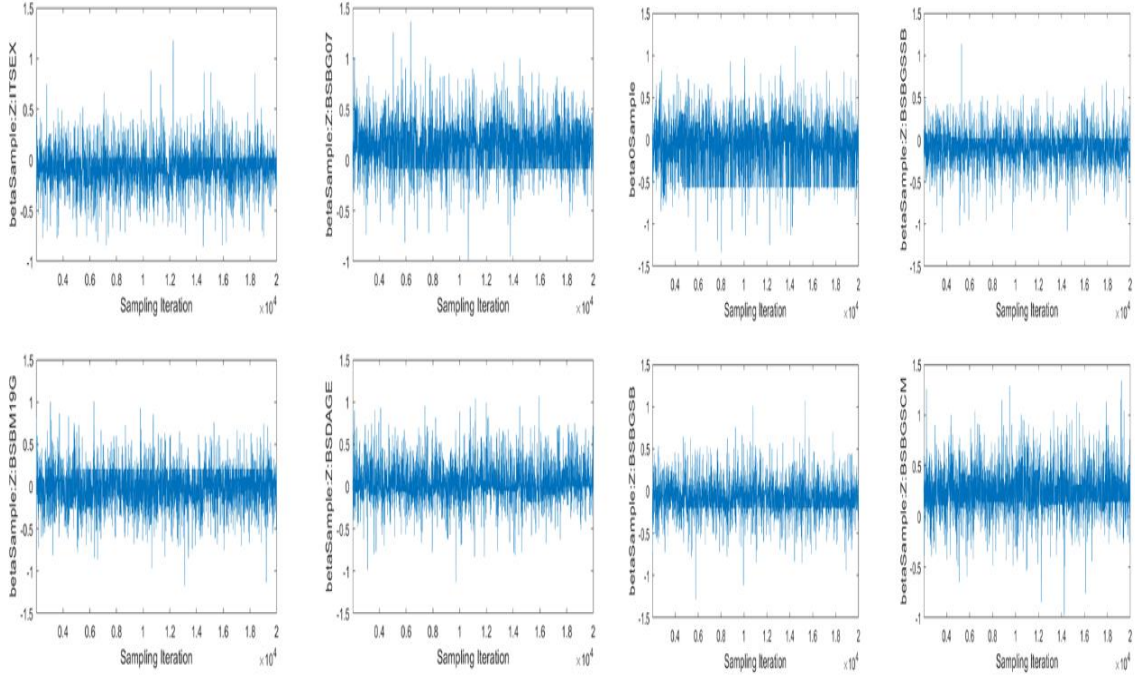
4.3. Bulgular

Bu tez çalışmasında 16 OECD ülkesindeki dezavantajlı sosyoekonomik geçmişlerine rağmen yüksek performans gösteren sekizinci sınıf öğrencilerin matematik başarısını etkileyen faktörler nonparametrik Bayesyen regresyon modeli olan lineer bağımlı Dirichlet süreç regresyon modeli kullanılarak tahmin edilmektedir. Nonparametrik Bayesyen regresyon her ülke grubu için öğrenci dayanıklılığı ile ilişkili değişkenlerin etkisinin belirlenmesini sağlamaktadır. Elde edilen bulgularla, düşük sosyoekonomik statüye sahip öğrencilerin zorlukların üstesinden gelmesine yardımcı olmak için neler yapılacağı konusunda çıkarımlar yapılmaktadır.

16 OECD ülkesi için veri seti ve değişkenler bölümünde açıklanan değişkenler kullanılmıştır. Lineer bağımlı Dirichlet süreç modeli (3.35) için bilgi vermeyen ilk dağılım varsayılmış ve şu hiperparametreler takip edilmiştir: $r_0 = 10$, $s_0 = 10$, $a_0 = 5$, $a_\alpha = b_\alpha = 1$. Nonparametrik Bayesyen model parametrelerinin son dağılımı hakkında çıkarımlar yapmak için son dağılım algoritmaları bölümünde açıklanan kesit örnekleme (Kalli vd. 2011) kullanılmıştır. Çalışmadaki analizler için MATLABR2022b ve BayesRegression programları kullanılmıştır. Model parametreleri 2000 (burn-in) ve 5 inceltme aralığı ile 20.000 Markov zinciri Monte Carlo örnekleme yinelemesi kullanılarak tahmin edilmiştir.⁴¹ MCMC simülasyonunun çıktısının analizi, yeterince büyük sayıda yineleme, MCMC örneğinin, çalıştırılıp çalıştırılmadığını kontrol etmek için yakınsama analizi yapılmaktadır. Bu amaçla ilk olarak model parametrelerinin MCMC örnekleme yinelemeleri üzerinden parametre örneklerinin tek değişkenli iz grafiği incelenmektedir. MCMC yinelemelerinde kararlı ve "tüylü" görünen tek değişkenli bir iz grafiği önerilmektedir.

⁴¹ Sonuçlar, üretilen toplam 20.000 örnekten 3600 Monte Carlo örneğine dayanmaktadır. Sonuçlar, burn-in veya inceltilmiş örnekler dışında üretilen MCMC numunelerinden hesaplanır.

Şekil 5: Model Parametrelerine Ait İz Grafiği



Şekil 5’te model parametrelerinin MCMC örneklerinin iz grafikleri kararlı ve "tüylü" görüldüğünden, bu parametreler için yakınsamanın sağlandığını söylenebilir. Bir model parametresinin iz grafiğinin “tüylülüğünün” bir ölçüsü olarak Brooks (1998), 0 ile 1 arasında CUSUM istatistiği olması gerektiğini önermektedir. Bu amaçla tahmin sonuçlarında CUSUM istatistik değerleri sunulmaktadır. Son dağılım nokta tahmini için, %95 MCCI (Monte Carlo Confidence Interval) Monte Carlo güvenilir aralık yarı genişlik boyutu (toplu ortalama analizi), Monte Carlo örnekleme hatasından kaynaklanan tahminin kesinliğini yansıtmaktadır. Bu bağlamda son dağılım nokta tahminlerin kesinliğinin değerlendirilmesi için Monte Carlo standart hatası (MCSE, Monte Carlo Standard Error) sunulmuştur (Flegal, Jones 2011: 182). Sonuçlar, parametrelerin son dağılım ortalama, medyan, standart sapma, son dağılım kantiller arası güvenilir aralığı %25-%75 (Interquartile Range Credible Interval)⁴², MCSE (Monte Carlo Standart Hatası) ve CUSUM istatistikleri çerçevesinde sunulacaktır.

Lineer DDP modeli için model son dağılım uyum istatistikleri model son dağılım hata kareler toplamı ve R^2 değerleri Tablo 7’de sunulmuştur.

⁴² Simülasyon sonuçları son dağılım kantiller arası güvenilir aralığın kullanılmasının uygun olduğunu göstermektedir (Li, Liny 2010). Son dağılım kantiller arası güvenilir aralık, beta katsayısının marjinal sonsal dağılımının % 25 ve %75’i ile tanımlanan kantiller arası güvenilir aralığı ifade etmektedir. Ayrıca veri normal dağılmıyorsa, veri dağılımı bilinmiyorsa, çarpıkça veya iki modlu bir son dağılım düşünülürse medyan ve kantiller arası güvenilir aralıkların rapor edilmesi önerilmektedir (Habibzadeh 2017; Kruschke 2010; Pupovac, Petroveckı 2011).

Tablo 7: Model Uyum İstatistikleri

		İstatistik	95% MCCIhw
Model son dağılım SSE	D(m)	196,198	0,448
Modelin uyum iyiliği	Gof (m)	58,908	0,115
Model karmaşıklığı cezası	P(m)	137,29	0,172
	R^2	0,492	0,001

*95%MCCIhw: %95 Monte Carlo güvenilir aralığın yarım genişliğini ifade etmektedir.

D_n verileri ve $m=1,...,M$ için model son dağılım ortalama hata kareler toplamı şu şekilde ifade edilmektedir:

$$D(m) = \sum_{i=1}^n \{y_i - E_n(Y | x_i, m)\}^2 + \sum_{i=1}^n V_n(Y | x_i, m) \quad (4.1)$$

Burada ilk terim, D_n verileri için modelin uyum iyiliğini ölçer ve ikinci terim, bir model karmaşıklığı cezasını göstermektedir. Karşılaştırılan $m=1,...,M$ regresyon modeli arasında D_n verisi için en iyi tahmine sahip model son dağılım ortalamama hata kareler toplamının minimum değere sahip olduğu model olarak belirlenmektedir (Gelfand, Ghosh 1998).

Bu kapsamda 16 OECD ülkesi için dayanıklı öğrencilerin matematik başarısını etkileyen faktörlerin LDDP model analiz sonuçları bağımsız değişkenler çerçevesinde sunulmaktadır. 16 OECD ülkesine ait tahmin sonuçları ilk olarak öğrenci özellikleri başlığı altında; cinsiyet, öğrencinin yaşı, ikinci olarak öğrenci tutumu ve beklentisi başlığı altında; öğrencinin akademik beklentisi, öğrencinin matematiğe olan güveni ve son olarak okul iklimi başlığı altında; pozitif öğretmen beklentisi, okula ait hissetme ve zorbalığa maruz kalma faktörleri çerçevesinde değerlendirilmektedir.

4.3.1. Öğrenci Özellikleri

Öğrenci özellikleri doğuştan gelen özellikler olup cinsiyet ve yaş değişkeni çerçevesinde ele alınmaktadır. İlk olarak 16 OECD ülkesindeki öğrenciler için cinsiyet parametresinin son dağılımına ilişkin bulgular Tablo 8’de sunulmaktadır. Dayanıklılı öğrenciler için tahmin sonuçları incelendiğinde 16 OECD ülkesi içerisinde yalnızca dört (Finlandiya, Norveç, Fransa ve Şili) ülkedeki kızların erkeklere

göre matematikte daha başarılı olduğu görülmektedir. Ayrıca Japonya ve Yeni Zelanda'daki dayanıklı öğrenciler arasında cinsiyet etkisinin olmadığı görülmektedir. Diğer ülkelerde kızlar matematik başarıları açısından erkeklere göre daha düşük performans sergilemektedir. Cinsiyet parametresinin son dağılım tahminine bakıldığında cinsiyet farklılığının en büyük olduğu ülkenin ABD olduğu görülmektedir. Bulgular dayanıklı öğrenciler arasında da cinsiyetin öğrenci performansı üzerinde etkili olduğunu ve çoğu OECD ülkesinde kızların erkeklere göre matematikte daha düşük performans sergilediklerini göstermektedir.

Tablo 8: Cinsiyet Parametresinin Marjinal Son Dağılım Tahmini

Parametre	Ortalama	Medyan	Standart Sapma	25%	75%	MCSE	CUSUM
ITSEX:ID=36(AUS)*	-0,086	-0,103	0,109	-0,163	-0,026	0,002	0,178
ITSEX:ID=152(CHL)*	0,131	0,127	0,062	0,058	0,199	0,012	0,003
ITSEX:ID=246(FIN)*	0,115	0,128	0,09	0,059	0,198	0,001	0,052
ITSEX:ID=250(FRA)*	0,125	0,128	0,07	0,059	0,198	0,028	0,011
ITSEX:ID=348(HUN)*	-0,126	-0,133	0,133	-0,186	-0,065	0,009	0,306
ITSEX:ID=372(IRL)*	-0,12	-0,116	0,076	-0,146	-0,067	0,007	0,1
ITSEX:ID=380(ITA)*	-0,129	-0,127	0,061	-0,146	-0,088	0,008	0,045
ITSEX:ID=392(JPN)	-0,059	-0,062	0,149	-0,156	0,041	0,017	0,189
ITSEX:ID=410(KOR)*	-0,111	-0,12	0,134	-0,181	-0,041	0,006	0,296
ITSEX:ID=440(LTU)*	-0,164	-0,146	0,129	-0,226	-0,094	0,004	0,316
ITSEX:ID=554(NZL)	-0,066	-0,086	0,129	-0,152	0,001	0,001	0,295
ITSEX:ID=578(NOR)*	0,13	0,128	0,067	0,059	0,198	0,022	0,012
ITSEX:ID=620(PRT)*	-0,119	-0,115	0,061	-0,146	-0,079	0,007	0,026
ITSEX:ID=752(SWE)*	-0,105	-0,119	0,125	-0,173	-0,062	0,004	0,311
ITSEX:ID=792(TUR)*	-0,141	-0,136	0,068	-0,184	-0,094	0,003	0,029
ITSEX:ID=840(USA)*	-0,192	-0,176	0,082	-0,249	-0,14	0,011	0,089

Not: Nokta tahminleri için son dağılım medyan değerleri yorumlanmaktadır. Sıfırı içermeyen %25-%75 son dağılım kantiller arası güvenilir aralık için anlamlı parametreler * ile gösterilmektedir. Son dağılım nokta tahminleri için küçük MCSE (0.1 veya 0.01) önerilmektedir (Flegal, Jones 2011). 0 ile 1 arasındaki CUSUM istatistiği parametreler için yakınsamanın sağlandığını ifade etmektedir (Brooks 1998).

Dayanıklı olmayan öğrenciler açısından bakıldığında da cinsiyet farklılıklarının öğrenci performansı üzerinde etkili olduğu görülmektedir. OECD ülkeleri genelinde ortalama olarak erkekler matematiksel problem becerileri açısından kızlardan yaklaşık olarak 16 puan daha iyi performans gösterdiği bilinmektedir. Buna ek olarak yüksek başarı gösteren kız ve erkekler arasında matematik performans farkının (19 puan) arttığı görülmektedir. Bu farklılığın sebebi olarak kızların matematik problemlerini

özme konusundaki düşük özgüveni gösterilmektedir. Genel olarak kızların hatta yüksek başarı gösteren kızların bile matematik problemlerini özme konusunda erkeklerden daha yüksek kaygı düzeyine ve daha az özgüvenleri olduđu görölmektedir. Bu cinsiyet eşitsizliğinin bir sonucu olarak kadınların STEM (bilim, teknoloji, mühendislik ve matematik) alanlarında yeterince temsil edilmediğı görölmektedir. OECD ülkeleri genelinde ortalama olarak kızların %5'inden azı bu alanlarda çalışmak isterken, bu alanlarda kariyer yapmayı planlayan erkeklerin oranı kızlardan fazladır⁴³ (OECD 2015: 13). Bu sonuçlar küçük yaşlarda başlayan cinsiyet eşitsizliklerinin ileri yaşlarda da bireylerin çalışma koşullarını ve seçimlerini etkilediğini göstermektedir. Ayrıca ülkelerin sürdürülebilir kalkınmasının sağlanmasında önemli bir yere sahip olduđu düşünülen STEM ile ilgili alanlara katılımın azalması uluslararası bir endişe olarak görölmektedir (Scherer, Nilsen 2016). Bu açıdan bakıldığında matematik performansındaki cinsiyet eşitsizliğinin negatif etkisinin STEM ile ilgili alanlarda kariyer yapmak isteyen öğrencilerin seçimlerini de etkileyeceğı düşünülmektedir.

Sürdürülebilir kalkınma hedeflerine bakıldığında Hedef 4.5, genel olarak 2030 yılına kadar, eğitimde cinsiyet eşitsizliklerini ortadan kaldırmak ve eğitim seviyelerine eşit erişimi sağlamaktır. Erkek öğrencilerin lehine büyüyen cinsiyet farkı, eğitim sistemlerinin cinsiyet eşitsizliği için yeniden odaklanan bir stratejiye ihtiyaç olduğunu göstermektedir (IEA 2020: 6). Bizim bulgularımız özellikle dayanıklı öğrencilerin de cinsiyet eşitsizliklerini aşamadıklarını göstermektedir. Cinsiyet farklılığının giderilmesi adına ebeveynlerin kız ve erkek çocuklarına tüm eğitim hayatlarında eşit destek ve cesaret vermesi beklenmektedir. Öğretmenlerin öğrencilere not verirken kendi cinsiyet önyargılarının daha fazla farkına varması ve öğrencilere eşit davranması (özellikle düşük sosyoekonomik statüye sahip öğrenciler) cinsiyet eşitsizliğinin öğrenci performansı üzerindeki etkisini azaltacağı düşünülmektedir. Buna ek olarak cinsiyet eşitsizliği için öğretmenlerin daha fazla talep eden öğretim stratejilerini takip etmesi önerilmektedir. Çünkü öğretmenler öğrencilerden matematik problemlerini bağımsız olarak çözmeye çalışmalarını istediğinde kızların matematikte daha iyi performans gösterdikleri görölmüştür (OECD 2015). Son olarak bizim sonuçlarımızla tutarlı olarak en iyi performans gösteren bazı ülkelerde (Japonya) kızların matematikte erkek sınıf arkadaşlarıyla arasında performans açısından anlamlı bir farklılık olmadığı görölmektedir. Bu sonuç eşit fırsatlar verildiğinde kız ve erkek öğrenciler yüksek başarı elde etme konusunda eşit şansa sahip olabildiklerini göstermektedir (OECD 2015; Sandoval-Hernández, Białowolski 2016).

⁴³ 2012'de üniversiteye giren kadınların sadece %14'ü bu alanlarda çalışmayı seçerken erkeklerin %39'u bu çalışma alanını seçmiştir (OECD 2015).

Öğrenci özellikleri içerisinde incelenilen diğer bir değişken öğrenci yaşıdır. Öğrenci özellikleri olarak yaş değişkeninin matematik başarıları üzerindeki etkisinin araştırılması özellikle dayanıklı öğrenciler için genç ergenler üzerindeki yaş etkisini değerlendirmek bakımından önemli olduğu düşünülmektedir. Modele sürekli değişken olarak eklenen yaş parametresine ait son dağılım özeti 16 OECD ülkesi için Tablo 9’da sunulmaktadır. Sonuçlar incelendiğinde Şili, Finlandiya, Fransa, Japonya ve Norveç’te yaş değişkeninin anlamlı ve dayanıklı öğrencilerin matematik başarılarını pozitif etkilediği görülmektedir. Ayrıca yaş faktörünün Japonya’da matematik performansı üzerinde en büyük etkiye (0.316) sahip olduğu görülmektedir. Diğer taraftan İrlanda, İtalya ve Portekiz’de yaş faktörünün dayanıklı öğrencilerin matematik performansı üzerinde negatif etkiye sahip olduğu tespit edilmiştir.

Tablo 9: Yaş Parametresinin Marjinal Son Dağılım Tahmini

Parametre	Ortalama	Medyan	Standart Sapma	25%	75%	MCSE	CUSUM
BSDAGE:ID=36(AUS)	0,089	0,032	0,196	-0,057	0,229	0,021	0,197
BSDAGE:ID=152(CHL)*	0,066	0,057	0,031	0,036	0,095	0,014	0,002
BSDAGE:ID=246(FIN)*	0,056	0,059	0,049	0,042	0,095	0,008	0,058
BSDAGE:ID=250(FRA)*	0,065	0,059	0,036	0,042	0,095	0,019	0,012
BSDAGE:ID=348(HUN)	0,043	0,005	0,159	-0,062	0,125	0,008	0,341
BSDAGE:ID=372(IRL)*	-0,062	-0,058	0,08	-0,1	-0,035	0,011	0,11
BSDAGE:ID=380(ITA)*	-0,054	-0,058	0,063	-0,087	-0,032	0,016	0,038
BSDAGE:ID=392(JPN)*	0,32	0,316	0,202	0,187	0,454	0,003	0,164
BSDAGE:ID=410(KOR)	0,103	0,05	0,208	-0,05	0,235	0,017	0,323
BSDAGE:ID=440(LTU)	-0,009	-0,029	0,134	-0,081	0,05	0,005	0,4
BSDAGE:ID=554(NZL)	0,13	0,073	0,201	-0,024	0,285	0,009	0,302
BSDAGE:ID=578(NOR)*	0,065	0,059	0,037	0,042	0,095	0,018	0,013
BSDAGE:ID=620(PRT)*	-0,074	-0,062	0,054	-0,108	-0,046	0,015	0,022
BSDAGE:ID=752(SWE)	0,003	-0,017	0,129	-0,078	0,055	0,006	0,356
BSDAGE:ID=792(TUR)	-0,059	-0,057	0,081	-0,122	0,005	0,013	0,023
BSDAGE:ID=840(USA)	0,06	0,046	0,113	-0,04	0,148	0,019	0,086

Not: Nokta tahminleri için son dağılım medyan değerleri yorumlanmaktadır. Sıfırı içermeyen %25-%75 son dağılım kantiller arası güvenilir aralık için anlamlı parametreler * ile gösterilmektedir. Son dağılım nokta tahminleri için küçük MCSE (0.1 veya 0.01) önerilmektedir (Flegal, Jones 2011). 0 ile 1 arasındaki CUSUM istatistiği parametreler için yakınsamanın sağlandığını ifade etmektedir (Brooks 1998).

Öğrenciler belirli bir yaşın üstünde (TIMSS için minimum ortalama yaş 9,5) bu eğitim değerlendirmelerine katılmaktadır. Ancak öğrenci performansındaki eşitsizliklerin erken yaşlarda geliştiği ve sonraki aşamalarda da artmaya devam ettiği görülmektedir (OECD 2018). TIMSS, katılımcı ülkelerdeki öğrencileri, test sırasındaki ortalama yaşın 13,5 (8.sınıf) olması koşuluyla örgün eğitimin

sekizinci yılında değerlendirmektedir. Eğitim sistemleri, okula başlama yaşı, politika ve uygulamalar açısından farklılık gösterdiğinden, yaş faktörünün öğrenci performansı üzerindeki etkisi konusunda ülkeler arasında farklılıklar olduğu görülmektedir (Mullis vd. 2020: 511). Bu çalışmanın bulguları, dayanıklı öğrenciler açısından yaş faktörünün farklı eğitim sistemlerinde eğitim çıktıları üzerinde farklı etkilere sahip olduğunu göstermektedir.

Çalışmalar genellikle öğrenci özellikleri ve aile geçmişinin öğrencinin performansını etkilediğini göstermektedir ancak, öğrencinin eğitim sonuçları ile öğrenci yaşı arasındaki ilişkiye çok fazla dikkat çekilmediği görülmektedir. Bu konudaki bazı teorilerin küçük yaştaki çocukların daha büyük olanlara göre öğrenmeye daha açık olduğunu ileri sürerken, diğer teoriler çocukların daha karmaşık yapıları öğrenmek için belirli bir yaşa ulaşmaları gerektiğini savunduğu görülmektedir (Strøm 2004: 3). Diğer taraftan bu yaş grubundaki öğrencilerin (ergenlik dönemi olarak düşünüldüğünde) olumlu öğrenme duygularında düşüşün gözlemlendiği bir dönemi temsil etmektedir. Olumlu öğrenme duygularının okul yılları boyunca azaldığı ve en önemli düşüşün ergenlik döneminde olduğu belirtilmektedir (Hagenauer, Hascher 2014: 20). Yaş faktörünün sekizinci sınıf öğrencilerin matematik başarıları üzerindeki negatif etkisi bu bağlamda değerlendirilebilir.

Aynı yılda doğan tüm çocukların (ay olarak farklılıklar mevcutken) aynı anda okula başlayacağını söyleyen kuralın takip edilmesiyle birlikte, yaş değişkeninin ilerleyen sınıflarda öğrencilerin performansları üzerinde etkili olduğu görülmektedir. Öğrencilerin yaşı ile öğrenme yeteneği arasında bir ilişki olduğu düşünülürse, bu faktörün öğrencinin performansını da etkilemesi beklenmektedir. Bu sonuçlar, politika yapımcılar için farklı eğitim sistemlerinde çocukların fırsatlarını eşitlemek için daha esnek kayıt kurallarının dikkate alınmasının önemli olabileceğini göstermektedir (Strøm 2004: 1-6).

4.3.2.Öğrenci Tutumu ve Beklentisi

Öğrencilerin eğitim ve mesleki hedeflerine ulaşmalarında yüksek akademik beklentilere ve yüksek özgüvene sahip olmanın önemli faktörler olduğu düşünülmektedir. Özellikle dayanıklı öğrencilerin dezavantajlı geçmişlerine rağmen yüksek başarı sergilemesinde gelecekteki hedeflere ulaşma çabası ve daha yüksek özgüvene sahip olması performansını etkileyen en önemli faktörlerden olarak görülmektedir (Erberber vd. 2015).

Bu bağlamda ilk olarak 16 OECD ülkesindeki dayanıklı öğrenciler için akademik beklenti parametresine ilişkin son dağılım tahmin sonuçları Tablo 10'da sunulmaktadır. Bulgular 4 OECD ülkesi (Şili, Finlandiya, Fransa ve Norveç) dışında bütün ülkelerde akademik beklentinin dayanıklı öğrencilerin

matematik performansı üzerinde anlamı ve pozitif bir etkiye sahip olduğu göstermektedir. Bu sonuçlar özellikle dezavantajlı öğrencilerin zorlukların üstesinden gelerek matematikte yüksek performans göstermesinde akademik beklentinin önemli bir faktör olduğunu vurgulamaktadır.

Tablo 10: Akademik Beklenti Parametresinin Marjinal Son Dağılım Tahmini

Parametre	Ortalama	Medyan	Standart Sapma	25%	75%	MCSE	CUSUM
BSBG07:ID=36(AUS)*	0,286	0,284	0,101	0,216	0,364	0,001	0,178
BSBG07:ID=152(CHL)	-0,003	-0,008	0,034	-0,022	0,027	0,013	0,003
BSBG07:ID=246(FIN)	0,008	-0,006	0,053	-0,006	0,026	0,002	0,065
BSBG07:ID=250(FRA)	-0,001	-0,006	0,039	-0,022	0,026	0,004	0,014
BSBG07:ID=348(HUN)*	0,296	0,282	0,153	0,199	0,385	0,013	0,334
BSBG07:ID=372(IRL)*	0,114	0,134	0,097	0,076	0,165	0,013	0,109
BSBG07:ID=380(ITA)*	0,142	0,135	0,061	0,084	0,193	0,007	0,041
BSBG07:ID=392(JPN)*	0,161	0,187	0,167	0,045	0,281	0,013	0,173
BSBG07:ID=410(KOR)*	0,295	0,283	0,164	0,206	0,387	0,027	0,296
BSBG07:ID=440(LTU)*	0,21	0,206	0,146	0,134	0,3	0,002	0,411
BSBG07:ID=554(NZL)*	0,235	0,233	0,127	0,165	0,327	0,001	0,293
BSBG07:ID=578(NOR)	-0,001	-0,006	0,036	-0,022	0,026	0,005	0,016
BSBG07:ID=620(PRT)*	0,138	0,135	0,064	0,083	0,191	0,002	0,026
BSBG07:ID=752(SWE)*	0,226	0,216	0,136	0,135	0,317	0,005	0,377
BSBG07:ID=792(TUR)*	0,298	0,294	0,091	0,221	0,364	0,008	0,024
BSBG07:ID=840(USA)*	0,126	0,135	0,098	0,061	0,194	0,009	0,083

Not: Nokta tahminleri için son dağılım medyan değerleri yorumlanmaktadır. Sıfırı içermeyen %25-%75 son dağılım kantiller arası güvenilir aralık için anlamlı parametreler * ile gösterilmektedir. Son dağılım nokta tahminleri için küçük MCSE (0.1 veya 0.01) önerilmektedir (Flegal, Jones 2011). 0 ile 1 arasındaki CUSUM istatistiği parametreler için yakınsamanın sağlandığını ifade etmektedir (Brooks 1998).

Dayanıklı öğrencilerin sağlıklı beklentiler sürdürdüğü ve gelecek hakkında net bir amaç duygusuna sahip oldukları bilinmektedir (Wang vd. 1994). Sosyoloji ve gelişim psikolojisi alanındaki araştırmalar erken kariyer planlamalarının nihai kazanımlar için çok önemli olduğunu dair kanıtlar sunmaktadır. Yüksek akademik beklenti yetişkinliğe geçen gençler (yaklaşık 14 veya 15 yaş) açısından avantajlıdır çünkü bu faktörler motivasyonu ve hedeflerine doğru çok çalışma isteğini temsil etmektedir (Reynolds vd. 2006: 187). Diğer taraftan ergenlik döneminde geleceğe yönelik eğitim planları yapmak öğrencileri hedeflerine ulaşma konusunda motive ettiği görülmektedir. Çünkü ergenlerin hedeflerini belirlemeleri ve sosyal çevre kısıtlamalarla birlikte genellikle bu hedefi rasyonel bir şekilde takip etmektedirler (Saha, Sikora 2012: 187). Genellikle yüksek akademik beklenti veya daha iddialı eğitim beklentileri olan

öğrencilerin eğitimleri için daha fazla çaba gösterme eğilimde olduğu görülmektedir (OECD 2018: 72). Son yıllarda, ebeveynlerinin deneyimlerine de bağlı olarak, önceki nesillere kıyasla öğrenciler daha fazla eğitim almakta ve beklemektedir (Goyette 2008). Bu bağlamda bazı araştırmalar da öğrencilerin son yıllarda daha hırslı hale geldiklerini vurgulamaktadır (Goyette 2008; Reynolds vd. 2006). Özellikle dezavantajlı geçmişlere sahip öğrenciler için zorlukların üstesinden gelme açısından bakıldığında yüksek akademik beklentilerin eğitim çıktısı üzerindeki pozitif etkisi önemlidir.

Ayrıca sonuçlar incelendiğinde akademik beklentinin sırasıyla en fazla Türkiye ve Macaristan'daki dayanıklı öğrenciler üzerinde etkili olduğu görülmektedir. Sikora ve Saha (2007) çalışmalarında daha az gelişmiş ve eşitsizliğin fazla olduğu ülkelerde eğitim getirilerinin daha yüksek olduğunu ve bu nedenle bu ülkedeki öğrencilerin genel başarı düzeyi gelişmiş ülkelere göre düşük olmasına rağmen, daha yüksek eğitim beklentilerine sahip olduğunu vurgulamaktadırlar. 16 OECD ülkesi içerisinde Türkiye ve Macaristan'ın düşük insani gelişmişlik endeksi (sırasıyla 0,82; 0,85) değerlerine, büyük cinsiyet eşitsizlik endeksi (sırasıyla 0,27; 0,27) değerlerine ve düşük kişi başı gelire (27,70; 31,32) sahip olduğu görülmektedir.⁴⁴ Bu durum düşük sosyoekonomik statüye sahip öğrencilerin başarılı olmak için normalde yapabileceklerinden daha fazla eğitim beklentisine girmesine ve sonuç olarak daha fazla çabalamasına neden olmaktadır.

Dayanıklı öğrencilerin akademik başarısını etkileyen en önemli faktörlerden biri de öğrencinin yüksek özgüvene sahip olmasıdır. Dayanıklı öğrencilerin matematik başarısının analiz edildiği bu çalışmada özgüven göstergesi olarak öğrencilerin matematikte kendine güvenmesi faktörü ele alınmıştır. Tablo 11'de parametreye ait son dağılım tahmin sonuçları incelendiğinde, 16 OECD ülkesi için de öğrencinin matematikte kendine güvenmesi dayanıklı öğrencilerin matematik başarısı üzerinde anlamlı ve pozitif bir etkiye sahip olduğu görülmektedir. İncelenen tüm ülkelerde dayanıklı öğrenciler için etki büyüklükleri farklı olmasına rağmen ortak pozitif ilişkinin bulunması önemli bir sonuçtur. İlk olarak farklı ülkelerdeki öğrenciler için benzer ilişkinin bulunması özgüven değişkeninin akademik başarı üzerinde güçlü ve tutarlı bir faktör olduğunu göstermektedir. Bu sonuçlar matematik becerileri ile ilgili özgüven düzeyinin aslında dayanıklılığın en güçlü açıklayıcılarından biri olduğunu göstermektedir.

⁴⁴ Göstergeler için 2019 yılı baz alınmıştır. İnsani Gelişmişlik Endeksi (HDI), Cinsiyet eşitsizliği (GII) ve Kişi Başı Gayri Safi Milli Gelir, 2017 PPP\$ (GNI) UNDP (2020) raporundan alınmıştır.

Tablo 11: Özgüven Parametresinin Marjinal Son Dağılım Tahmini

Parametre	Ortalama	Medyan	Standart Sapma	25%	75%	MCSE	CUSUM
BSBGSCM:ID=36(AUS)*	0,391	0,376	0,107	0,313	0,451	0,011	0,2
BSBGSCM:ID=152(CHL)*	0,264	0,286	0,044	0,223	0,288	0,001	0,003
BSBGSCM:ID=246(FIN)*	0,261	0,287	0,053	0,224	0,289	0,011	0,062
BSBGSCM:ID=250(FRA)*	0,264	0,287	0,044	0,224	0,289	0,001	0,013
BSBGSCM:ID=348(HUN)*	0,32	0,315	0,141	0,243	0,397	0,008	0,372
BSBGSCM:ID=372(IRL)*	0,194	0,201	0,107	0,142	0,249	0,001	0,123
BSBGSCM:ID=380(ITA)*	0,163	0,165	0,085	0,116	0,222	0,009	0,037
BSBGSCM:ID=392(JPN)*	0,489	0,482	0,154	0,397	0,578	0,017	0,188
BSBGSCM:ID=410(KOR)*	0,347	0,358	0,166	0,277	0,428	0,013	0,318
BSBGSCM:ID=440(LTU)*	0,289	0,289	0,178	0,177	0,383	0,02	0,402
BSBGSCM:ID=554(NZL)*	0,384	0,377	0,149	0,293	0,466	0,005	0,306
BSBGSCM:ID=578(NOR)*	0,263	0,287	0,047	0,224	0,289	0,001	0,015
BSBGSCM:ID=620(PRT)*	0,18	0,184	0,086	0,141	0,243	0,009	0,033
BSBGSCM:ID=752(SWE)*	0,283	0,289	0,161	0,181	0,368	0,018	0,356
BSBGSCM:ID=792(TUR)*	0,34	0,344	0,064	0,293	0,38	0,007	0,029
BSBGSCM:ID=840(USA)*	0,115	0,11	0,099	0,03	0,179	0,005	0,091

Not: Nokta tahminleri için son dağılım medyan değerleri yorumlanmaktadır. Sıfırı içermeyen %25-%75 son dağılım kantiller arası güvenilir aralık için anlamlı parametreler * ile gösterilmektedir. Son dağılım nokta tahminleri için küçük MCSE (0.1 veya 0.01) önerilmektedir (Flegal, Jones 2011). 0 ile 1 arasındaki CUSUM istatistiği parametreler için yakınsamanın sağlandığını ifade etmektedir (Brooks 1998).

Öğrencilerin akademik becerilerine dair inançları akademik performanslarını destekleyebilir veya engelleyebilir. Çalışmalar dayanıklı öğrencilerin “azim, güçlü irade ve olumlu eğilim” gibi özellikleri vurgulamaktadır (Borman, Overman 2004). Dayanıklı öğrencilerin önemli bireysel özelliklerinden biri de yüksek özgüvene sahip olmalarıdır (Wang vd. 1994). Araştırmalar dayanıklı öğrencilerin düşük sosyoekonomik statüye sahip ve düşük başarılı öğrencilere göre yeteneklerine daha fazla güvendiklerini göstermektedir (Borman, Overman 2004; OECD 2011). Daha fazla motivasyona sahip, yeteneklerine daha fazla güvenen öğrenciler daha yüksek başarı puanlarına sahip olma eğilimindedir (Michaelides vd. 2019; OECD 2011). Yüksek özgüven ile birlikte bu öğrenciler daha güçlü olumlu öğrenme yaklaşımlarına sahip olurlar ve dersler için daha fazla çaba göstererek daha yüksek başarı puanı elde edebilirler.

Ayrıca ülkeler açısından sonuçlar değerlendirildiğinde Japonya ve Kore’de özgüvenin diğer ülkelere göre matematik başarısını daha fazla etkilediği görülmektedir. Bu sonuçlar Japonya ve Kore gibi Konfüçyüsçü Asya kültürlerinin öğrenmeye yüksek saygı ve başarı çabasına vurgu yapan bir geleneğe sahip olması ile bağlantılı olabilir (Sandoval-Hernández, Białowolski 2016; Stankov 2010). Asya eğitim sistemi, geçmişte ve günümüzde Konfüçyüsçü prensiplere dayanan değerlere ve uygulamalara atıfta bulunmaktadır. Konfüçyüsçülük, çalışma ve rekabetin önemini vurgulayan bir öğretilerdir. Stankov (2010) çalışmasında, Konfüçyüsçü Asyalı öğrencilerin akademik başarıya yönelik güçlü dürtülere sahip olduğunu vurgulamaktadır. Böylece Japonya ve Kore’deki dayanıklı öğrencilerin akademik başarıya olan inançlarının yüksek başarı ile sonuçlandığı görülmektedir.

Tüm ülkelerdeki dayanıklı öğrenciler için özgüven değişkeni ile matematik başarısı arasındaki anlamlı ve pozitif ilişki düşük sosyoekonomik statüye sahip başarısız öğrencilerin performanslarını artırmanın bir yolu olarak görülebilir. Yani düşük sosyoekonomik statüye sahip öğrencilerin özgüvenlerini destekleyen politikaların geliştirilmesi, bu öğrencilerin dezavantajlı durumun üstesinden gelerek başarılı olmalarına katkı sağlayabilir.

Düşük sosyoekonomik statüye sahip öğrencilerin motivasyonlarını ve yeteneklerine olan güvenlerini artıran uygulamalar geliştirilmesi bu öğrencilerin başarılı olmasını sağlayabilir. Özellikle düşük sosyoekonomik statüye sahip öğrenciler arasında olumlu öğrenmeye yönelik yaklaşımları teşvik eden politikaların ve bu öğrencilerin akademik becerilerine olan güvenlerini artırmak için planlanmış programların uygulanmasının önemli olacağı düşünülmektedir. Diğer taraftan düşük sosyoekonomik statüye sahip öğrencilerin böyle programlara katılımının sağlanması diğer öğrencilere göre daha zor olabilir. Tüm sosyoekonomik grupları ortak olarak hedef alan politikalar aslında avantajlı ve dezavantajlı öğrenciler arasındaki uçurumun artmasına neden olabilir (OECD 2011: 80). Bu nedenle politika yapımcıların politikaları belirlerken ve uygularken sosyoekonomik gruplar arasındaki farklılığı dikkate alması beklenmektedir.

Temel olarak akademik beklenti ve özgüven ile ilgili sonuçlar, düşük sosyoekonomik statüye sahip öğrencilerin bu iki faktörü geliştirerek başarılı olabileceklerini vurgulamaktadır. Bulgular en dezavantajlı öğrenciler arasında bile yüksek başarı puanlarının elde edilmesini sağlayan koşulların oluşturulabileceğini göstermektedir.

4.3.3.Okul İklimi

Eğitim araştırmalarında öğrencilerin okula aidiyet duygusu, pozitif öğretmen beklentisi ve okulda akran zorbalığına maruz kalması öğrenci başarısını etkileyen önemli faktörler olarak görülmektedir. Bu değişkenler düşük sosyoekonomik statüye sahip öğrencilerin okullarda geliştirdikleri sosyal ortam, akranlarıyla ve öğretmenlerle kurdukları ilişkilerin akademik başarısı üzerindeki etkileri bağlamında değerlendirilecektir.

Okula aidiyet duygusu öğrencilerin okuldaki akranları tarafından bireysel olarak kabul gördüğünü, okul sosyal ortamına kabul edildiğini ve desteklendiğini hissetmesi olarak tanımlanmıştır (Ma 2003). Okula ait hissetme parametresine ait son dağılım tahmin sonuçları Tablo 12’de sunulmuştur. Sonuçlar okula aidiyet duygusunun Finlandiya ve İrlanda dışında diğer ülkelerdeki öğrenciler için matematik başarısı üzerinde anlamlı bir etkiye sahip olduğunu göstermektedir. Diğer taraftan okula aidiyet duygusu yalnızca Şili, Fransa ve Norveç’teki dayanıklı öğrencilerin matematik başarısı üzerinde pozitif etkiye sahipken, diğer ülkelerde negatif etkiye sahip olduğu görülmektedir.

Tablo 12: Okula Ait Hissetme Parametresinin Marjinal Son Dağılım Tahmini

Parametre	Ortalama	Medyan	Standart Sapma	25%	75%	MCSE	CUSUM
BSBGSSB:ID=36(AUS)*	-0,159	-0,154	0,09	-0,214	-0,113	0,001	0,163
BSBGSSB:ID=152(CHL)*	0,011	0,009	0,032	0,007	0,032	0,001	0,005
BSBGSSB:ID=246(FIN)	0,004	0,009	0,046	-0,024	0,032	0,015	0,063
BSBGSSB:ID=250(FRA)*	0,01	0,009	0,034	0,009	0,032	0,007	0,021
BSBGSSB:ID=348(HUN)*	-0,159	-0,149	0,132	-0,217	-0,1	0,006	0,321
BSBGSSB:ID=372(IRL)	-0,05	-0,054	0,097	-0,122	0,001	0,003	0,103
BSBGSSB:ID=380(ITA)*	-0,08	-0,112	0,079	-0,128	-0,022	0,01	0,031
BSBGSSB:ID=392(JPN)*	-0,128	-0,14	0,153	-0,223	-0,032	0,008	0,186
BSBGSSB:ID=410(KOR)*	-0,171	-0,163	0,14	-0,228	-0,105	0,007	0,293
BSBGSSB:ID=440(LTU)*	-0,114	-0,125	0,137	-0,187	-0,047	0,001	0,368
BSBGSSB:ID=554(NZL)*	-0,128	-0,141	0,118	-0,203	-0,07	0,011	0,267
BSBGSSB:ID=578(NOR)*	0,011	0,009	0,034	0,009	0,032	0,003	0,018
BSBGSSB:ID=620(PRT)*	-0,09	-0,112	0,081	-0,138	-0,025	0,002	0,026
BSBGSSB:ID=752(SWE)*	-0,109	-0,128	0,132	-0,18	-0,051	0,003	0,336
BSBGSSB:ID=792(TUR)*	-0,154	-0,148	0,066	-0,207	-0,114	0,006	0,026
BSBGSSB:ID=840(USA)*	-0,155	-0,136	0,098	-0,216	-0,1	0,011	0,086

Not: Nokta tahminleri için son dağılım medyan değerleri yorumlanmaktadır. Sıfırı içermeyen %25-%75 son dağılım kantiller arası güvenilir aralık için anlamlı parametreler * ile gösterilmektedir. Son dağılım nokta tahminleri için küçük MCSE (0.1 veya

0.01) önerilmektedir (Flegal, Jones 2011). 0 ile 1 arasındaki CUSUM istatistiği parametreler için yakınsamanın sağlandığını ifade etmektedir (Brooks 1998).

Öğrencilerin okuldaki etkinliklere katılmadığında veya kendine saygı duyulduğunu hissetmediğinde, yani kendini okula ait hissetmediğinde bu durumun derslere katılmasında sınırlamalara neden olacağı için başarı açısından zararlı etkileri olabileceği vurgulanmaktadır (Finn 1989). Diğer taraftan bazı çalışmalar okula aidiyet duygusu arttığında öğrencilerin matematik başarısının düştüğünü tespit etmişlerdir (Topal 2021; Topçu vd. 2016).

Bu çalışmada elde edilen bulgular çoğu eğitim sistemi için okula aidiyet duygusunun dayanıklı öğrencilerin matematik başarısını düşürdüğünü vurgulamaktadır. Genel olarak 28 OECD ülkesinde son yıllarda kendilerini okula ait hissettiklerini söyleyen öğrencilerin oranının azaldığı görülmektedir. Diğer taraftan yüksek sosyoekonomik statüye sahip öğrencilerin düşük sosyoekonomik statüye sahip öğrencilere göre okulda daha güçlü bir aidiyet duygusuna sahip oldukları bilinmektedir. OECD ülkelerinde ortalama olarak düşük sosyoekonomik statüye sahip öğrencilerin yaklaşık %28 ve yüksek sosyoekonomik statüye sahip öğrencilerin %33'ü okulda sosyal olarak bütünleşmiş hissetmektedirler (OECD 2018: 96). Topçu vd. (2016) aidiyet duygusunun başarı üzerindeki negatif etkisini öğrencilerin okulları baskı ortamı olarak görmeleri ile ilişkilendirmektedir. Buna ek olarak dayanıklı öğrenciler için aidiyet duygusunun akademik başarı üzerindeki bu negatif etkisi düşük sosyoekonomik statüye sahip öğrencilerin okula karşı olumsuz tutumları ile ilgili olabilir. Ayrıca bu çalışmada sekizinci sınıf öğrencileri ele alınmaktadır ve ortaokul sürecinin aidiyet duygusunun gelişimi için en kritik aşama olduğu belirtilmektedir (Ma 2003). Öğrencilerin çocukluktan ergenliğe geçiş süreci içerisindeki bu dönemde farklı tutumları akademik başarıları üzerinde olumsuz etkilere neden olmuş olabilir.

Okula aidiyet duygusunun öğrencilerin kendi davranışları yanı sıra, okul arkadaşlarının ve öğretmen tutumları ile ilgili olduğu vurgulanmaktadır (Ma 2003). Öğretmenlerin öğrencilerin aidiyet duygularını şekillendirmede önemli bir rol oynadığı düşünülürse, özellikle düşük sosyoekonomik statüye sahip öğrenciler için bu durumun koruyucu bir faktör olarak tersine çevrilebileceği düşünülmektedir.

Okul iklimi çerçevesinde ele alınan ikinci faktör pozitif öğretmen beklentisidir ve pozitif öğretmen beklentisi öğrenci performansına öğretmenin güvenini ifade etmektedir. Pozitif öğretmen beklentisi yardımıyla öğretmen tutumunun dayanıklı öğrencilerin matematik başarısı üzerindeki etkisi incelenebilir. Bu bağlamda, pozitif öğretmen beklentisi parametresine ait son dağılım tahmin sonuçları Tablo 13' te sunulmuştur. Sonuçlar incelendiğinde pozitif öğretmen beklentisinin Avustralya, Japonya, Kore, Yeni Zelanda ve Türkiye'deki dayanıklı öğrencilerin matematik başarısı üzerinde negatif bir etkiye sahip olduğu tespit edilmiştir. Diğer taraftan öğrencilerin öğretmenlerinin kendilerine olan

güvenlerine ilişkin inançlarının Şili, Finlandiya, Fransa ve Norveç'teki dayanıklı öğrencilerin matematik başarıları üzerinde ise pozitif etkiye sahip olduğu görülmektedir.

Tablo 13: Pozitif Öğretmen Beklentisi Parametresinin Marjinal Son Dağılım Tahmini

Parametre	Ortalama	Medyan	Standart Sapma	25%	75%	MCSE	CUSUM
BSBM19G:ID=36(AUS)*	-0,184	-0,146	0,156	-0,302	-0,068	0,017	0,199
BSBM19G:ID=152(CHL)*	0,074	0,074	0,022	0,058	0,088	0,006	0,003
BSBM19G:ID=246(FIN)*	0,067	0,074	0,043	0,056	0,088	0,026	0,042
BSBM19G:ID=250(FRA)*	0,072	0,074	0,024	0,056	0,088	0,003	0,016
BSBM19G:ID=348(HUN)	-0,048	-0,047	0,152	-0,137	0,038	0,004	0,344
BSBM19G:ID=372(IRL)	-0,022	-0,006	0,119	-0,046	0,05	0,031	0,132
BSBM19G:ID=380(ITA)	0,041	0,047	0,07	-0,019	0,065	0,001	0,029
BSBM19G:ID=392(JPN)*	-0,334	-0,335	0,132	-0,423	-0,245	0,005	0,192
BSBM19G:ID=410(KOR)*	-0,16	-0,142	0,154	-0,253	-0,055	0,005	0,32
BSBM19G:ID=440(LTU)	-0,035	-0,033	0,154	-0,126	0,05	0,006	0,372
BSBM19G:ID=554(NZL)*	-0,18	-0,15	0,168	-0,307	-0,05	0,02	0,299
BSBM19G:ID=578(NOR)*	0,073	0,074	0,024	0,056	0,088	0,004	0,017
BSBM19G:ID=620(PRT)	0,026	0,04	0,071	-0,028	0,053	0,007	0,026
BSBM19G:ID=752(SWE)	0,011	-0,006	0,15	-0,078	0,065	0,022	0,367
BSBM19G:ID=792(TUR)*	-0,065	-0,068	0,085	-0,132	-0,002	0,001	0,027
BSBM19G:ID=840(USA)	0,017	0,039	0,103	-0,041	0,071	0,002	0,088

Not: Nokta tahminleri için son dağılım medyan değerleri yorumlanmaktadır. Sıfırı içermeyen %25-%75 son dağılım kantiller arası güvenilir aralık için anlamlı parametreler * ile gösterilmektedir. Son dağılım nokta tahminleri için küçük MCSE (0.1 veya 0.01) önerilmektedir (Flegal, Jones 2011). 0 ile 1 arasındaki CUSUM istatistiği parametreler için yakınsamanın sağlandığını ifade etmektedir (Brooks 1998).

Bu çalışmada elde edilen sonuçlar farklı eğitim sistemlerinde öğretmen beklentileri açısından dayanıklı öğrenciler için ortak bir yaklaşım ortaya konulmadığını göstermektedir. Çünkü ülkeler arasında öğretmen nitelikleri, işe alma uygulamaları ve aldıkları eğitim farklılık göstermektedir. Buna bağlı olarak öğretmenin öğrenci başarısı üzerindeki etkisinin okullar ve eğitim sistemleri arasında farklılık gösterdiği söylenebilir.

Öğretmen beklentilerinin okulun ilk yıllarından itibaren öğrenme ve akademik başarı üzerinde önemli ve destekleyici etkiye sahip olduğunu gösteren çalışmalar mevcuttur (Agasisti, Longobardi 2014;

Bowles 1970; Sandoval-Hernández, Białowolski 2016; Wang vd. 1998). Literatürle uyumlu olarak çalışmanın bulguları öğretmenlerin pozitif beklentisinin, özellikle düşük sosyoekonomik statüye sahip öğrenci oranının çok düşük olduğu Finlandiya, Fransa ve Norveç'teki öğrencilerin matematik başarıları üzerinde pozitif etkiye sahip olduğunu göstermektedir.

Diğer taraftan düşük sosyoekonomik statüye sahip öğrencilerin performansı üzerindeki öğretmen etkisinin azaldığı bulunmuştur. Yüksek sosyoekonomik statüye sahip öğrencilerin diğer öğrencilere göre öğretmen etkilerden daha fazla yararlandığı gözlenmiştir (Konstantopoulos 2009). Bu durumun düşük sosyoekonomik statüye sahip öğrencilerin genellikle daha az yetenekli ve az deneyime sahip öğretmenlerin çoğunun bulunduğu okullara gitmesinden kaynaklanıyor olabilir (Lankford vd. 2002). Ayrıca düşük sosyoekonomik statüye sahip öğrencilerin başarabileceğini söylemek öğretmenler açısından zor olabilir (Wang vd. 1994). Tüm bu sebepler pozitif öğretmen beklentisinin düşük sosyoekonomik statüye sahip başarılı öğrencilerin matematik performansı üzerindeki negatif etkisini açıklayabilir. Beş OECD ülkesi (Avustralya, Japonya, Kore, Yeni Zelanda ve Türkiye) için elde edilen negatif etki pozitif öğretmen beklentisinin dezavantajlı öğrenciler için koruyucu bir faktör olmadığı şeklinde yorumlanabilir.

Sürdürülebilir kalkınma hedefleri kapsamında öğretmenler tüm SDG 4 eğitim hedeflerine ulaşmanın anahtarı olarak gösterilmektedir (Martin, Mullis 2017). Eğitim hedeflerinden Hedef 4.c 2030'a kadar gelişmekte olan ülkelerde, özellikle en az gelişmiş ülkelerde ve gelişmekte olan küçük ada devletlerinde öğretmen eğitimi için uluslararası iş birliği de dahil olmak üzere, nitelikli öğretmen arzını önemli ölçüde artırmayı içermektedir (IEA 2020). Bulgularımız öğretmenlerin yalnızca yüksek beklentilerle düşük sosyoekonomik statüye sahip öğrencilerin yüksek performanslar elde etmesini sağlayamayacağını göstermektedir. Eğitim sistemlerinde öğretmenin öğrenci potansiyeline olan inancının, akademik dayanıklılığı destekleyen bir faktör olabilmesi için öğretmen tutumunun öğrenmeyi kolaylaştıran güçlü öğretim uygulamaları ile bileştirilmesi gerekmektedir. Ayrıca öğretmenler farklı stratejiler geliştirerek, dezavantajlı öğrencileri yüksek beklentilere tabi tutar ve tüm öğrencilerin öğrenebileceğine dair inancını gösterebilir. Böylelikle düşük sosyoekonomik statüye sahip öğrenciler de pozitif öğretmen beklentilerden faydalanabilirler (Wang vd. 1998).

Diğer taraftan okul ikliminin bir bütün olarak şiddet ve zorbalıktan etkilendiği, genellikle korku ve güvensizlik ikliminin öğrencilerin eğitim kalitesini düşürdüğü bilinmektedir (UNESCO 2017). Bu nedenle okulda şiddet ve zorbalığın öğrenciler üzerindeki eğitimsel etkisi önemlidir. Bu bağlamda dayanıklı öğrenciler için zorbalık parametresine ait son dağılım tahminleri sunulmaktadır. Tablo 14'te zorbalık parametresine ait son dağılım tahminleri incelendiğinde, zorbalığın tüm ülkelerdeki dayanıklı öğrencilerin matematik başarıları üzerinde negatif etkiye sahip olduğu görülmektedir. İrlanda da diğer

ülkelerdeki dayanıklı öğrencilere göre zorbalığın matematik başarısı üzerinde daha büyük bir etkisi olduğu tespit edilmiştir.

Tablo 14: Zorbalık Parametresinin Marjinal Son Dağılım Tahmini

Parametre	Ortalama	Medyan	Standart Sapma	25%	75%	MCSE	CUSUM
BSBGSB:ID=36(AUS)	-0,059	-0,078	0,101	-0,125	0,015	0,003	0,167
BSBGSB:ID=152(CHL)*	-0,074	-0,072	0,033	-0,089	-0,068	0,007	0,003
BSBGSB:ID=246(FIN)*	-0,078	-0,071	0,038	-0,09	-0,066	0,014	0,068
BSBGSB:ID=250(FRA)*	-0,076	-0,071	0,03	-0,09	-0,066	0,007	0,021
BSBGSB:ID=348(HUN)	-0,067	-0,079	0,144	-0,133	0,009	0,002	0,322
BSBGSB:ID=372(IRL)*	-0,117	-0,106	0,087	-0,149	-0,055	0,028	0,132
BSBGSB:ID=380(ITA)*	-0,064	-0,069	0,067	-0,107	-0,042	0,009	0,041
BSBGSB:ID=392(JPN)	0,012	0,007	0,152	-0,086	0,099	0,02	0,193
BSBGSB:ID=410(KOR)	-0,07	-0,079	0,132	-0,136	0,016	0,009	0,289
BSBGSB:ID=440(LTU)	-0,024	-0,045	0,133	-0,107	0,039	0,005	0,355
BSBGSB:ID=554(NZL)	-0,067	-0,079	0,109	-0,129	0,002	0,004	0,274
BSBGSB:ID=578(NOR)*	-0,076	-0,071	0,032	-0,09	-0,066	0,007	0,016
BSBGSB:ID=620(PRT)*	-0,081	-0,077	0,055	-0,114	-0,044	0,002	0,031
BSBGSB:ID=752(SWE)	-0,048	-0,059	0,114	-0,107	0,012	0,01	0,336
BSBGSB:ID=792(TUR)*	-0,079	-0,086	0,077	-0,131	-0,023	0,008	0,025
BSBGSB:ID=840(USA)	-0,001	-0,023	0,088	-0,053	0,057	0,008	0,096

Not: Nokta tahminleri için son dağılım medyan değerleri yorumlanmaktadır. Sıfırı içermeyen %25-%75 son dağılım kantiller arası güvenilir aralık için anlamlı parametreler * ile gösterilmektedir. Son dağılım nokta tahminleri için küçük MCSE (0.1 veya 0.01) önerilmektedir (Flegal, Jones 2011). 0 ile 1 arasındaki CUSUM istatistiği parametreler için yakınsamanın sağlandığını ifade etmektedir (Brooks 1998).

Zorbalık tüm başarı düzeylerinde ve “tüm sosyoekonomik statüdeki” öğrencileri etkileyen küresel bir olgu olarak tanımlanmaktadır (Rutkowski, Rutkowski 2016). Son yirmi yılda tüm dünyada zorbalığa maruz kaldığını bildiren öğrenci oranının ve bu durumun olumsuz sonucunun etkisinin arttığı görülmektedir. UNESCO her yıl 246 milyon çocuğun okul çevresinde şiddet ve zorbalığa maruz kaldığını tahmin etmektedir (UNESCO 2017). Çalışmalar zorbalık ile öğrenci başarısı arasında güçlü bir ilişki olduğunu göstermiştir (Gardella vd. 2017; Rutkowski, Rutkowski 2018; Rutkowski, Rutkowski 2016; Wagemaker, Mirazchiyski 2023). Zorbalığa daha fazla maruz kalan öğrencilerin depresyon, kaygı ve psikososyal zorluklar yaşadığı görülmektedir (Eyüboğlu vd. 2021). Yani zorbalık öğrencilerin ruh

sağlığını etkileyebilir. Gardella vd. (2017) çalışmalarında, öğretilen konu ne olursa olsun zorbalığın eğitim çıktılarıyla olumsuz bir ilişkisi olduğunu vurgulamaktadır. Uluslararası değerlendirme ve raporlar zorbalığın öğrencilerin matematik gibi temel konulardaki performanslarını düşürdüğünü göstermektedir (UNESCO 2017). Ayrıca, bu çalışmanın bulguları ile tutarlı olarak Rutkowski ve Rutkowski (2018) zorbalığa uğrayan öğrencilerin matematikte daha az başarılı olma eğiliminde olduğunu tespit etmiştir.

Sonuçlar incelendiğinde diğer ülkelerdeki öğrencilere göre zorbalığın İrlanda'daki dayanıklı öğrenciler üzerinde daha büyük negatif etkiye sahip olduğu görülmektedir. UNESCO (2017) gelişmiş ülkelerden elde ettiği verilerden yola çıkarak bu ülkelerde zorbalığın fiziksel şiddete göre daha yaygın olduğunu vurgulamaktadır. Diğer ülkelerde ise fiziksel şiddetin daha yaygın olduğu görülmektedir. 16 OECD ülkesi arasında İrlanda en yüksek İnsani Gelişim Endeks (0,955) değerine sahiptir ve bu bağlamda düşünüldüğünde bu büyük negatif etki zorbalığın fiziksel şiddetten daha yaygın olması ile ilişkili olabilir. Diğer taraftan sonuçlar bazı ülkelerde zorbalığın dayanıklı öğrencilerin matematik performansı üzerinde anlamlı bir etkisinin olmadığını göstermektedir. Bu durum eğitim sistemleri için zorbalık mağduriyetinin coğrafya, dil, baskın ırk, kültür ve ekonomik gelişmeyle ilişkili olabileceğini gösterebilir.

Sonuçlar, zorbalığın öğrencilerin eğitim haklarını tehdit edebileceğini ve bunun da öğrencinin akademik kazanım üzerinde olumsuz bir etkiye neden olacağını göstermektedir. Tüm öğrencilere güvenli bir öğrenme ortamı sağlamak, zorbalığı azaltmak ve başarıyı artırmak için politika yapıcılarının bu konuyu dikkate alması gerekmektedir (Rutkowski, Rutkowski 2018). Bu bağlamda sürdürebilir kalkınma kapsamında eğitim hedeflerinden Hedef 4.a herkes için güvenli, şiddet içermeyen, kapsayıcı ve etkili eğitim ortamları sağlamayı içermektedir (IEA 2020). Birçok ülkede okulda zorbalıkla mücadele etmek için birtakım girişimler uygulansa da bu girişimlerin kapsamlı bir yaklaşım olmadığı vurgulanmaktadır. Bu nedenle daha özel olarak ulusal yasaların ve politikaların, okul politikalarının ve davranış kurallarının yasalaştırılması; öğretmenler ve diğer okul personeli için olumlu disiplin biçimleri konusunda eğitim sağlanması konularının dikkate alınması gerektiği belirtilmektedir (UNESCO 2017). Ayrıca, öğretmenlerin zorbalık konusunda da önemli rol oynadığı düşünülmektedir. Öğretmen desteğinin öğrenci saldırganlığını azaltarak okullarda zorbalık yaygınlığını azalttığı ve zorbalık meydana geldiğinde mağduriyet oluşmasını önlediği görülmüştür (Rutkowski, Rutkowski 2016). Bulgularımız zorbalığın uluslararası bir olgu olduğunu ve düşük sosyoekonomik statüye sahip öğrencilerin başarılı olmak için üstesinden gelmesi gereken bir zorluk olduğunu vurgulamaktadır.

16 OECD ülkesi için nonparametrik Bayesyen regresyon tahmin sonuçları dayanıklı öğrencilerin akademik beklenti ve özgüveninin matematik başarısını pozitif etkileyen en önemli ve tutarlı faktörler olduğunu göstermektedir. Ayrıca öğrencilerin daha fazla zorbalığa maruz kalması bütün ülkelerdeki

dayanıklı öğrencilerin matematik performansı üzerinde negatif bir etkiye sahiptir. Cinsiyet parametresine ilişkin bulgular çoğu eğitim sisteminde matematikte erkeklerin kızlara göre daha başarılı olduğunu gösterir. Bu sonuçlar dayanıklı öğrenciler için cinsiyet etkisinin varlığını göstermektedir. Diğer faktörlerde birçok ülkede dayanıklı öğrencilerin matematik başarısını etkilemektedir. Bu sonuçlar düşük sosyoekonomik geçmişten gelen öğrencilerin mevcut zorluklarının belirlenmesi ve bu öğrencilerin başarı şanslarının olduğunu göstermesi açısından önemli olabilir. Diğer taraftan ülkeler arası elde edilen farklı bulgular akademik dayanıklılığı artırmaya yönelik politikaların ülkeler bağlamında özel olarak ele alması gerektiğini göstermektedir.

4.3.4. Optimal Kümeleme Algoritması

Yukarıda 16 OECD ülkesi için lineer bağımlı Dirichlet süreç regresyon model parametrelerine ait son dağılım sonuçları ve bulgular sunulmuştur. Dirichlet süreç karışım modellerinin model tabanlı bir kümeleme analizi sağladığı bilinmektedir. LDDP model analiz son dağılım örnekleri (MCMC) üzerinden ülkelerin optimal kümelemesi hesaplanabilmektedir. Burada toplayıcı kümeleme algoritmasına dayalı Fritsch ve Ickstadt (2009) optimal kümeleme yöntemi kullanılarak ülkelerin en uygun kümeleme sonuçları sunulmaktadır.

Kümeleme algoritma çıktısı analizde ele alınan 16 OECD ülkesinin optimal bir kümelemesini göstermektedir. Aynı kümeleme değerlerine sahip herhangi iki ülkenin aynı küme grubuna ait olduğunu ve benzer rastgele katsayı parametre değerlerine sahip olduğu söylenmektedir. Aksi takdirde, iki ülke farklı küme gruplarına aittir ve rastgele parametreler farklı değerlere sahip olmaktadır. Tablo 15 incelendiğinde kümelenen bu ikili ülke gruplarının benzer parametre işaretlerine ve anlamlılık durumuna sahip olduğu görülmektedir.

Tablo 15: Benzer Parametre Tahmin Sonuçları

Ülke	ITSEX	BSBG07	BSBM19G	BSDAGE	BSBGSSB	BSBGSB	BSBGSCM
Avustralya	-	+	-	Ø	-	Ø	+
Yeni Zelanda	Ø	+	-	Ø	-	Ø	+
İtalya	-	+	Ø	-	-	-	+
Portekiz	-	+	Ø	-	-	-	+
Norveç	+	Ø	+	+	+	-	+
Finlandiya	+	Ø	+	+	Ø	-	+
Litvanya	-	+	Ø	Ø	-	Ø	+
İsveç	-	+	Ø	Ø	-	Ø	+

*(Ø), parametrenin nokta tahmininin (medyan) istatistiksel olarak anlamlı olmadığını; (-), parametrenin nokta tahmininin (medyan) negatif işaretli olduğunu; (+), parametrenin nokta tahmininin (medyan) pozitif işaretli olduğunu ifade etmektedir.

Sonuçlar incelendiğinde 16 OECD ülkesi içinde Avustralya ve Yeni Zelanda, İtalya ve Portekiz, Norveç ve Finlandiya, İsveç ve Litvanya birlikte kümелendiği görülmektedir. Birlikte kümelenen bu ikili ülke grupları için dayanıklı öğrencilerin matematik başarısını etkileyen faktörlerin (parametre değerlerinin) benzerlik gösterdiği söylenebilir.

Ayrıca birlikte kümelenen ülkelere ait ortalama matematik puanları, dayanıklı öğrenci oranları ve önemli sosyoekonomik göstergeler Tablo 16’da sunulmaktadır.

Tablo 16: Optimal Kümeleme Sonuçları

Ülkeler	TIMSS 2019 Ortalama Matematik Puanı	Dayanıklı Öğrenci Yüzdesi	İnsani Gelişmişlik Endeksi (HDI)	Eğitim Harcaması (GSYİH’nin yüzdesi)	Ortalama Eğitim Yılı (MYS)	Demokrasi Endeksi (DI)
	(8.Sınıf)	(%)	(İndeks)	(%)	(Yıl)	(İndeks)
Avustralya	517	28	0,944	5,135	12,7	9,09
Yeni Zelanda	482	28	0,931	5,164	12,8	9,26
İtalya	497	34	0,892	4,095	10,4	7,52
Portekiz	500	38	0,864	4,630	9,3	8,03
Norveç	503	28	0,957	7,938	12,9	9,87
Finlandiya	509	30	0,938	6,417	12,8	9,25
İsveç	503	29	0,945	7,639	12,5	9,39
Litvanya	520	25	0,882	3,972	13,1	7,5

*Göstergeler için 2019 yılı baz alınmıştır. İnsani Gelişim Endeksi ve Ortalama Eğitim yılı göstergeleri UNDP (2020) raporundan alınmıştır. Eğitim Harcaması Dünya Bankası veri tabanından (WDI) ve Demokrasi Endeksi EIU (2020) raporundan

derlenmiştir. Ortalama matematik puanı TIMSS 2019 sekizinci sınıf matematik değerlendirmesinden alınmıştır (Mullis vd. 2020).

Öncelikle optimal kümeleme sonuçları ülkelerdeki dayanıklı öğrenci oranları açısından tutarlıdır. Birlikte kümelenen ikili ülke gruplarının dayanıklı öğrenci oranlarının benzer olduğu görülmektedir. Bu sonuçlar dayanıklılığın uluslararası bir değerlendirmesini sunması açısından önemli bir yere sahiptir. Çünkü sonuçlar benzer dayanıklılık oranlarına sahip ülkelerin dayanıklılığı etkileyen faktörler açısından da benzer olduğunu göstermektedir. Bulgular Avustralya ve Yeni Zelanda, İtalya ve Portekiz, Norveç ve Finlandiya, İsveç ve Litvanya ikili kümelenmiş ülkelerde bu faktörlerin benzer etkilere sahip olduğunu göstermektedir.

Diğer taraftan, insani gelişmişlik endeksi, eğitim harcaması, ortalama eğitim yılı ve demokrasi endeksi açısından incelendiğinde birlikte kümelenen ülkelerin çok benzer (İsveç ve Litvanya dışında) değerlere sahip olduğu görülmektedir. İsveç ve Litvanya dayanıklı öğrenci oranı ve diğer sosyoekonomik göstergeler açısından incelendiğinde iki ülke göstergeleri arasında İsveç lehine önemli farklar bulunduğu tespit edilmiştir. Ancak ortalama matematik puanının Litvanya da İsveç'e göre daha yüksek olduğu görülmektedir. Eşit bir eğitim sisteminin öğrencilerin sosyoekonomik geçmişlerinden bağımsız olması gerektiği bilinmektedir. Çalışmalar ve uluslararası değerlendirmeler önceden dünyanın en eşit eğitim sistemlerinden birine sahip olan İsveç'ten son dönemde eğitim sistemindeki örgütsel değişikliklerin eğitim eşitsizliğinin arttığını vurgulamaktadır (OECD 2019b; Siebecke, Jarl 2022; Yang Hansen, Gustafsson 2019: 81). İsveç'in gelişmişlik ve eğitim göstergeleri açısından yüksek değerlere sahipken daha düşük göstergelere sahip Litvanya ile kümelenmesi bu bağlamda değerlendirilebilir.

Optimal kümeleme sonuçları dayanıklılığı etkileyen faktörlerin 16 OECD içerisinde yalnızca birlikte kümelenmiş bu sekiz ülke için benzer ilişkilere sahip olduğunu göstermektedir. Dolayısıyla bu sonuçlar gelişmekte olan ülkelerdeki politika yapıcılarının eğitim politikası ve reformlarını sadece gelişmiş ülkelerin araştırma sonuçlarına göre şekillendirmemesi gerektiğini göstermektedir. Bunun yerine bu politikaların ülkeye özgü bulgulara göre belirlenmesi gerektiğini vurgulamaktadır. Başka bir deyişle, bir ülke için bir faktörün faydalı olması bu faktörün başka bir ülke için otomatik olarak uygulanıp faydalı olacağı anlamına gelmez. Steiner-Khamsi (2004), bu durumu “*eğitim politikası borçlanması*” olarak adlandırmaktadır. Çalışmasında bir ülke için başarılı olan politikanın başka bir ülke tarafından kopyalanmasının ülkelerin benzersiz özelliklerinin hesaba katmadığı için sorun teşkil edebileceğini vurgulamaktadır. Örneğin, Finlandiya eğitim sistemindeki eşitlik ve yüksek başarı puanları nedeniyle genellikle politika borçlanması için hedef olan ülkelere biridir. Bu çalışmanın bulguları ise dayanıklı öğrencilerin matematik başarısı açısından 16 OECD ülkesi arasında Finlandiya'nın yalnızca Norveç ile benzer ilişkilere sahip olabileceğini göstermektedir.

5.SONUÇ

Eğitimin, bireylerin ve toplumların yaşam kalitesinin yükselmesinde önemli bir yere sahip olduğu bilinmektedir. Tüm öğrencilerin akademik olarak tam potansiyellerine ulaşmalarını sağlamak hem eşitlik hem de verimlilik nedenleriyle ülkeler için önemli bir konudur. Tarihsel olarak düşük sosyoekonomik statüye sahip çocukların orantısız bir şekilde eşitsizlik unsurlarından etkilendiği ve akademik başarısızlık riskiyle karşı karşıya kaldığı görülmektedir. Ancak bazı öğrenciler sosyoekonomik olarak dezavantajlı geçmişe sahip olmalarına rağmen zorlukların üstesinden gelerek yüksek başarı elde etmektedir. Bu öğrenciler düşük sosyoekonomik geçmişleri nedeniyle kötü eğitim sonuçları beklenen ancak akademik olarak başarılı olan öğrenciler olduğu için dayanıklı öğrenciler olarak tanımlanmaktadır. Bu öğrenci grubunun incelenmesi ve akademik başarılarını etkileyen faktörlerinin belirlenmesi önemlidir. Eğer politika yapımcılar ve eğitimciler bu öğrencilerin sosyoekonomik dezavantaja rağmen başarılı olmalarını sağlayan faktörleri anlayabilirse benzer sosyoekonomik geçmişlere sahip öğrencilerin performanslarını iyileştirmek için daha doğru adımlar atabilirler.

Dayanıklı öğrencilerin akademik başarısını inceleyen hem teorik hem de ampirik birçok çalışma mevcuttur. Bu çalışmada mevcut literatürdeki yaklaşımlardan farklı olarak 16 OECD ülkesinde dayanıklı öğrencilerin matematik başarısını etkileyen faktörleri nonparametrik Bayesyen regresyon yardımıyla analiz edilmiştir. Dayanıklı öğrenciler düşük sosyoekonomik statüye sahip öğrencilerin bir alt kümesi olduğu için mevcut çalışmalarda küçük örneklem boyutu bir sınırlılık olarak görülmektedir. İlk olarak Bayesyen yaklaşımın tercih edilmesinin nedeni büyük örneklem teorisine dayanmaması, küçük örneklem için daha doğru tahminler sağlaması ve karmaşık veri yapılarıyla birlikte karmaşık modellerin daha esnek bir şekilde tahmin edilmesine olanak sunmasıdır. Bu, özellikle karışım modelinde bazı grupların çok daha küçük örneklem büyüklüğüne sahip olduğu durumlarda fayda sağlamaktadır. Ayrıca tüm Bayes yöntemlerinin altında yatan temel varsayım, modeli belirleyen parametrelerin rastgele bir değişken olmasıdır. Böylece Bayesyen yaklaşım model parametreleriyle birlikte rastgele etkileri tahmin etmektedir.

Diğer taraftan nonparametrik Bayesyen modeller temel olarak sonsuz sayıda parametreye izin veren esnek modellemeye ve üretici yaklaşımlara odaklanmaktadır. Bu modeller parametrik modeller gibi veri

retme mekanizması hakkında kısıtlayıcı varsayımlara sahip deęillerdir. Bylece nonparametrik Bayesyen yaklaşıma sayesinde tm rastgele etkiler desteklenir ve esnek model tabanlı bir kmeleme analizi saęlanır. Dirichlet sreç karışım modelleri, zellikle gruplanma problemlerinde kullanılan olasılıksal bir modeldir. Bu karışım modelleri farklı gruplar zerinde modeller oluřturulmasına, parametrelerin gruplar arasında farklılaşmasına ve grup dzeyindeki etkileri tahmin ederken belirsizlięin hesaba katılmasına izin vermektedir. Bylece klasik yntemlerin hesaplama sorunuyla karşılařtıkları ok sayıda rastgele etki veya hiyerarşik yapı ieren karmařık modellerin herhangi bir hesaplama sorunuyla karřılamadan analiz edilmesine olanak saęlanır. Son olarak 16 OECD lkesi iin akademik dayanıklılıęın lkeler arası analizinin sunulması farklı lkelerdeki ęrencilerin dayanıklılık zelliklerinin farklılık gsterip gstermedięi ve bu eęitim sistemleri iin dayanıklılık kalıpları hakkında bilgi vereceęi dřnlmektedir. Bu baęlamda alıřma bildięimiz kadarıyla farklı eęitim sistemlerindeki dayanıklı ęrencilerin matematik bařarısını inceleyen ve nonparametrik Bayesyen modeller yardımıyla ele alan ilk alıřmadır. alıřma kapsamında 16 OECD lkesindeki dayanıklı ęrencilerin matematik bařarısını etkileyen faktrler incelenerek ıkarımlar elde edilmiř ve uluslararası bir bakıř aısıyla dayanıklılık literatrne katkı saęlanmıřtır.

Bu alıřmanın sonuları, dayanıklı ęrencilerin matematik bařarısını etkileyen faktrlerin anlařılması ve dřk sosyoekonomik statye sahip ęrencilerin bařarılı olmasına yardımcı olabilecek faktrlerin varlıęı ile ilgili uluslararası ıkarımlara sahiptir. ncelikle cinsiyet parametresine iliřkin bulgular birok lkede matematikte erkeklerin kızlara gre daha bařarılı olduęunu gstermektedir. Bu sonular dayanıklı ęrenciler iin cinsiyet etkisinin varlıęını kanıtlamaktadır. Erkek ęrencilerin lehine byyen cinsiyet farkı, eęitim sistemlerinin cinsiyet eřitsizlięini derinlemesine dikkate alan yeni stratejilere ihtiyaı olduęunu gstermektedir. Bir dięer ęrenci zellięi olarak, dayanıklı ęrenciler aısından yař faktrnn farklı lkelerde eęitim ıktıları zerinde farklı etkilere sahip olduęu tespit edilmiřtir. Yař faktr ile ilgili bu sonular, politika yapıcılar iin farklı eęitim sistemlerinde ocukların fırsatlarını eřitlemek iin daha esnek kayıt kurallarının dikkate alınmasının nemli olabileceęini gstermektedir.

16 OECD lkesi iin nonparametrik Bayesyen regresyon tahmin sonularından dayanıklı ęrencilerin akademik beklenti ve zgveninin matematik bařarısını pozitif etkileyen en nemli ve tutarlı faktrler olduęu tespit edilmiřtir. Yksek zgven ile bu ęrenciler daha gl olumlu ęrenme yaklaşımlarına sahip olurlar ve dersler iin daha fazla aba gstererek daha yksek bařarı puanı elde edebilirler. Dięer taraftan dayanıklı ęrencilerin dezavantajlı gemiřlerine raęmen yksek bařarı sergilemesinde gelecekteki hedeflere ulařma abası ve eęitim beklentilerinin nemli olduęu grlmektedir. Farklı lkelerdeki dayanıklı ęrenciler iin akademik bařarı ile bu anlamlı ve pozitif sonular dřk sosyoekonomik statye sahip ęrenciler arasında da yksek bařarı puanlarının elde edilmesini saęlayan kořulların oluřturulabileceęini gstermektedir. Bu sonular zellikle dřk sosyoekonomik statye

sahip öğrenciler arasında olumlu öğrenmeye yönelik yaklaşımları teşvik eden politikaların ve bu öğrencilerin akademik becerilerine olan güvenlerini artırmak için planlanmış programların uygulanmasının önemli olduğunu göstermektedir.

Okula aidiyet duygusu, pozitif öğretmen beklentisi ve okulda akran zorbalığına maruz kalma düşük sosyoekonomik statüye sahip öğrencilerin okullarda geliştirdikleri sosyal ortam, akranlarıyla ve öğretmenlerle kurdukları ilişkilerin akademik başarısı üzerindeki etkileri bağlamında değerlendirilmiştir. Çoğu eğitim sistemi için okula aidiyet duygusunun dayanıklı öğrencilerin matematik başarısı üzerinde negatif etkiye sahip olduğu görülmektedir. Dayanıklı öğrenciler için aidiyet duygusunun akademik başarı üzerindeki bu negatif etkisi düşük sosyoekonomik statüye sahip öğrencilerin okula karşı olumsuz tutumları olduğunu göstermektedir. Öğretmenlerin öğrencilerin aidiyet duygularını şekillendirmede önemli bir rol oynadığı düşünülürse, özellikle düşük sosyoekonomik statüye sahip öğrenciler için bu durumun koruyucu bir faktör olarak tersine çevrilebileceği düşünülmektedir.

Pozitif öğretmen beklentisi yardımıyla öğretmen tutumunun dayanıklı öğrencilerin matematik başarısı üzerindeki etkisi incelendiğinde, sonuçlar farklı ülkelerde öğretmen beklentileri açısından dayanıklı öğrenciler için ortak bir yaklaşım ortaya konulmadığını göstermektedir. Öğretmenler farklı stratejiler geliştirerek, dezavantajlı öğrencileri yüksek beklentilere tabi tutarak ve tüm öğrencilerin öğrenebileceğine dair inancını gösterebilir. Böylece düşük sosyoekonomik statüye sahip öğrencilerin pozitif öğretmen beklentilerden faydalanabileceği düşünülmektedir.

Son olarak öğrencilerin daha fazla zorbalığa maruz kalmasının bütün ülkelerdeki öğrencilerin matematik performansı üzerinde negatif bir etkiye sahip olduğu tespit edilmiştir. Bu sonuçlar zorbalığın tüm sosyoekonomik statüdeki öğrencileri etkileyen küresel bir olgu olduğunu göstermektedir. Zorbalık için özel olarak ulusal yasaların ve politikaların, okul politikalarının ve davranış kurallarının yasalaştırılması; öğretmenler ve diğer okul personeli için olumlu disiplin biçimleri konusunda eğitim sağlanması konularının dikkate alınması gerektiği düşünülmektedir.

Optimal kümeleme sonuçları 16 OECD ülkesi içinde Avustralya ve Yeni Zelanda, İtalya ve Portekiz, Norveç ve Finlandiya, İsveç ve Litvanya birlikte kümelendiğini tespit etmiştir. Bu sonuçlar dayanıklılığı etkileyen faktörlerin 16 OECD ülkesi içerisinde yalnızca birlikte kümelenmiş bu sekiz ülkede benzer ilişkilere sahip olduğunu ifade etmektedir. Buradan gelişmekte olan ülkelerdeki politika yapıcılarının eğitim politikası ve reformlarını sadece gelişmiş ülkelerin araştırma sonuçlarına göre şekillendirmemesi gerektiği sonucuna ulaşılabilir. Bunun yerine bu politikaların ülkeye özgü bulgular ve yerel koşullar tarafından şekillendirilmesinin daha faydalı olacağı düşünülmektedir.

Sonular ilk olarak dřk sosyoekonomik gemiřten gelen ğrencilerin mevcut zorluklarının belirlenmesi ve bu ğrencilerin başarı řanslarının olduėunu gstermesi aısından önemlidir. Ayrıca bu sonuların politika yapıcıların dřk sosyoekonomik statye sahip ğrencilerin zorlukların stesinden gelebilmesine katkı saėlayabilecek faktrleri anlaması ve bu ğrenciler iin neyin fark yaratacaėını belirlemesi aısından nemli olduėu dřnlmektedir. Diėer taraftan lkeler arası elde edilen farklı bulgular akademik dayanıklılıėı artırmaya ynelik politikaların eėitim sistemi baėlamında zel olarak ele alması gerektiėini gstermektedir.



KAYNAKÇA

- Agasisti, T., Avvisati, F., Borgonovi, F., & Longobardi, S. (2018). Academic Resilience: What Schools and Countries Do to Help Disadvantaged Students Succeed in PISA. İçinde *OECD Publishing* (C. 3, Sayı 167). <http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED581639&site=eds-live>
- Agasisti, T., Avvisati, F., Borgonovi, F., & Longobardi, S. (2021). What School Factors are Associated with the Success of Socio-Economically Disadvantaged Students? An Empirical Investigation Using PISA Data. İçinde *Social Indicators Research* (C. 157, Sayı 2). Springer Netherlands. <https://doi.org/10.1007/s11205-021-02668-w>
- Agasisti, T., & Longobardi, S. (2014). Inequality in Education: Can Italian Disadvantaged Students Close the Gap? *Journal of Behavioral and Experimental Economics*, 52, 8–20. <https://doi.org/10.1016/j.socec.2014.05.002>
- Agasisti, T., & Longobardi, S. (2016). Equality of Educational Opportunities, Schools' Characteristics and Resilient Students: An Empirical Study of EU-15 Countries Using OECD-PISA 2009 Data. *Social Indicators Research*, 134(3), 917–953. <https://doi.org/10.1007/s11205-016-1464-5>
- Agasisti, T., Soncin, M., & Valenti, R. (2016). School Factors Helping Disadvantaged Students to Succeed: Empirical Evidence from Four Italian Cities. *Policy Studies*, 37(2), 147–177. <https://doi.org/10.1080/01442872.2015.1127341>
- Akay, E. Ç., & Uyar, S. K. (2017). *R Uygulamalı Nonparametrik Ekonometri*. Der Yayınları.
- Ando, T. (2010). *Bayesian Model Selection and Statistical Modeling*. CRC Press.
- Antoniak, C. E. (1974). Mixtures of Dirichlet Processes with Applications to Bayesian Nonparametric Problems. *The Annals of Statistics*, 2(6), 1152–1174.
- APA. (2017). Education and Socioeconomic Status. İçinde *APA Publishing*. <https://www.apa.org/pi/ses/resources/publications/factsheet-education.pdf>

- Aydin, M., & Erdem, C. (2022). Profiles of Academically Resilient Students: An Examination on TIMSS Mathematics Data. *Current Psychology*. <https://doi.org/10.1007/s12144-022-03055-5>
- Aydiner, A., & Kalender, İ. (2015). Student Segments Based on the Factors Related to Sense of Belonging Across Disadvantaged and Resilient Groups in PISA 2012. *Procedia - Social and Behavioral Sciences*, 174, 3299–3305. <https://doi.org/10.1016/j.sbspro.2015.01.997>
- Berger, J. O. (1980). *Statistical Decision Theory: Foundations, Concepts and Methods*. Springer New York.
- Bernardo, Jose M. (1979). Reference Posterior Distributions for Bayesian Inference. *Journal of the Royal Statistical Society: Series B (Methodological)*, 41(2), 113–147. <https://doi.org/10.1111/j.2517-6161.1979.tb01066.x>
- Bernardo, José M., & Smith, A. F. M. (2000). *Bayesian Theory*. Wiley.
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer Science+Business Media, LLC. <https://doi.org/10.13109/9783666604409.185>
- Blei, D. M. (2014). Build, Compute, Critique, Repeat: Data Analysis with Latent Variable Models. *Annual Review of Statistics and Its Application*, 1(2014), 203–232. <https://doi.org/10.1146/annurev-statistics-022513-115657>
- Blei, D. M., & Jordan, M. I. (2006). Variational Inference for Dirichlet Process Mixtures. *International Society for Bayesian Analysis*, 1(1), 121–144.
- Bolstad, W. M., & Curran, J. M. (2017). *Introduction to Bayesian Statistics*. John Wiley & Sons.
- Borman, G. D., & Overman, L. T. (2004). Academic Resilience in Mathematics among Poor and Minority Students. *Elementary School Journal*, 104(3), 177–195. <https://doi.org/10.1086/499748>
- Bowles, S. (1970). Education, Income, and Human Capital. İçinde *NBER*.
- Broer, M., Bai, Y., & Fonseca, F. (2019). Socioeconomic Inequality and Educational Outcomes. İçinde *Life Sciences* (C. 16, Sayı 12). Springer. <https://doi.org/10.1016/0024->

- Brooks, S. P. (1998). Quantitative Convergence Assessment for Markov Chain Monte Carlo via Cusums. *Statistics and Computing*, 8(3), 267–274. <https://doi.org/10.1023/A:1008965613031>
- Carlin, B. P., & Louis, T. A. (2000). *Bayes and Empirical Bayes Methods for Data Analysis*. CRC Press.
- Carlin, B. P., & Louis, T. A. (2009). *Bayesian Methods for Analysis*. CRC Press.
- Cheung, K. C. (2017). The Effects of Resilience in Learning Variables on Mathematical Literacy Performance: A Study of Learning Characteristics of the Academic Resilient and Advantaged Low Achievers in Shanghai, Singapore, Hong Kong, Taiwan and Korea. *Educational Psychology*, 37(8), 965–982. <https://doi.org/10.1080/01443410.2016.1194372>
- Chirkina, T., Khavenson, T., Pinskaya, M., & Zvyagintsev, R. (2020). Factors of Student Resilience Obtained from TIMSS and PISA Longitudinal Studies. *Issues in Educational Research*, 30(4), 1245–1263.
- Clavel, J., García-Crespo, F., & Sanz San Miguel, L. (2022). Rising Above Their Circumstances: What Makes Some Disadvantaged East and South-East Asian Students Perform Far Better in Science than Their Background Predicts? *Asia Pacific Journal of Education*, 42(4), 714–729. <https://doi.org/10.1080/02188791.2021.1886905>
- Coleman, J. S. (1968). Equality of Educational Opportunity. *Equity and Excellence in Education*, 6(5), 19–28. <https://doi.org/10.1080/0020486680060504>
- Congdon, P. (2006). *Bayesian Statistical Modelling*. John Wiley & Sons.
- Congdon, P. (2020). *Bayesian Hierarchical Models With Applications Using R*. CRC Press.
- Cowles, M. K. (2013). *Applied Bayesian Statistics With R and OpenBUGS Examples*. Springer.
- Cruz -Marcelo, A., Rosner, G. R., Mueller, P., & Stewart, C. (2010). Modeling Covariates with Nonparametric Bayesian Methods. *SSRN Electronic Journal*, 1–28. <https://doi.org/10.2139/ssrn.1576665>

- De Iorio, M., Johnson, W. O., Müller, P., & Rosner, G. L. (2009). Bayesian Nonparametric Nonproportional Hazards Survival Modeling. *Biometrics*, 65(3), 762–771. <https://doi.org/10.1111/j.1541-0420.2008.01166.x>
- De Iorio, M., Müller, P., Rosner, G. L., & MacEachern, S. N. (2004). An ANOVA Model for Dependent Random Measures. *Journal of the American Statistical Association*, 99(465), 205–215. <https://doi.org/10.1198/016214504000000205>
- Donovan, T. M., & Mickey, R. M. (2019). *Bayesian Statistics for Beginners*. Oxford University Press.
- EIU. (2020). Democracy Index 2019: A Year of Democratic Setbacks and Popular Protest. İçinde *The Economist Intelligence Unit*. <http://www.eiu.com/topic/democracy-index>
- Erberber, E., Stephens, M., Mamedova, S., Ferguson, S., & Kroeger, T. (2015). Socioeconomically Disadvantaged Students Who are Academically Successful: Examining Academic Resilience Cross-Nationally. İçinde *IEA's Policy Brief Series No. 5* (C. 56, Sayı 2). <https://doi.org/10.1177/0022146515584605>
- Escobar, M. D., & West, M. (1995). Bayesian Density Estimation and Inference Using Mixtures. *Journal of the American Statistical Association*, 90(430), 577–588. <https://doi.org/10.1080/01621459.1995.10476550>
- Eyüboğlu, M., Eyüboğlu, D., Pala, S. C., Oktar, D., Demirtaş, Z., Arslantaş, D., & Ünsal, A. (2021). Traditional School Bullying and Cyberbullying: Prevalence, the Effect on Mental Health Problems and Self-Harm Behavior. *Psychiatry Research*, 297(January), 113730. <https://doi.org/10.1016/j.psychres.2021.113730>
- Ferguson, R. F. (2003). Teachers' Perceptions and Expectations and the Black-White Test Score Gap. *Urban Education*, 38(4), 460–507. <https://doi.org/10.1177/0042085903254970>
- Ferguson, T. S. (1973). A Bayesian Analysis of Some Nonparametric Problems. *The Annals of Statistics*, 1(2), 209–230.
- Finn, J. D. (1989). Withdrawing from School. *Review of Educational Research*, 59(2), 117–142. <https://doi.org/10.3102/00346543059002117>

- Finn, J. D., & Rock, D. A. (1997). Academic Success among Students at Risk for School Failure. *Journal of Applied Psychology*, 82(2), 221–234. <https://doi.org/10.1037/0021-9010.82.2.221>
- Fındık, L., & Kavak, Y. (2013). Türkiye’deki Sosyo-Ekonomik Açıdan Dezavantajlı Öğrencilerin PISA 2009 Başarılarının Değerlendirilmesi. *Kuram ve Uygulamada Eğitim Yönetimi*, 19(2), 249–273.
- Flegal, J. M., & Jones, G. L. (2011). Implementing MCMC: Estimating with Confidence. İçinde *Handbook of Markov Chain Monte Carlo* (ss. 175–197). <https://www.ptonline.com/articles/how-to-get-better-mfi-results>
- Fritsch, A., & Ickstadt, K. (2009). Improved Criteria for Clustering Based on the Posterior Similarity Matrix. *Bayesian Analysis*, 4(2), 367–392. <https://doi.org/10.1214/09-BA414>
- Gamerman, D., & Lopes, H. F. (2006). *Markov Chain Monte Carlo Stochastic Simulation for Bayesian Inference*. CRC Press.
- García-Crespo, F. J., Suárez-álvarez, J., Fernández-Alonso, R., & Muñiz, J. (2022). Academic Resilience in Mathematics and Science: Europe TIMSS-2019 Data. *Psicothema*, 34(2), 217–225. <https://doi.org/10.7334/psicothema2021.486>
- Gardella, J. H., Fisher, B. W., & Teurbe-Tolon, A. R. (2017). A Systematic Review and Meta-Analysis of Cyber-Victimization and Educational Outcomes for Adolescents. *Review of Educational Research*, 87(2), 283–308. <https://doi.org/10.3102/0034654316689136>
- Gelfand, A. E., & Ghosh, S. K. (1998). Model Choice: A Minimum Posterior Predictive Loss Approach. *Biometrika*, 85(1), 1–11.
- Gelman, A., B.Carlin, J., S.Stern, H., & B.Rubin, D. (2004). *Bayesian Data Analysis*. Chapman and Hall/CRC. <http://arxiv.org/abs/2009.00781>
- Gelman, A., & Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.
- Gershman, S. J., & Blei, D. M. (2012). A Tutorial on Bayesian Nonparametric Models. *Journal of Mathematical Psychology*, 56(1), 1–12. <https://doi.org/10.1016/j.jmp.2011.08.004>

- Geweke, J. (2005). *Contemporary Bayesian Econometrics and Statistics*. John Wiley & Sons.
- Ghosal, S., & van der Vaart, A. (2017). *Fundamentals of Nonparametric Bayesian Inference*. Cambridge University Press.
- Gilks, W. R., Richardson, S., & Spiegelhalter, D. J. (1997). *Markov Chain Monte Carlo in Practice*. Springer Science+Business Media Dordrecht 1996.
- Gill, J. (2015). *Bayesian Methods: A Social and Behavioral Sciences Approach*. CRC Press.
<https://doi.org/10.1198/000313008x370915>
- Goldstein, M., & Wooff, D. (2007). *Bayes Linear Statistics*. Wiley.
- Goodman, S. N. (1999a). Toward Evidence-Based Medical Statistics. 1: The P Value Fallacy. *American College of Physicians–American Society of Internal Medicine*, 130, 995–1004.
- Goodman, S. N. (1999b). Toward Evidence-Based Medical Statistics. 2: The Bayes Factor. *American College of Physicians–American Society of Internal Medicine*, 130, 1005–1013.
- Goyette, K. A. (2008). College for Some to College for All: Social Background, Occupational Expectations, and Educational Expectations Over Time. *Social Science Research*, 37(2), 461–484. <https://doi.org/10.1016/j.ssresearch.2008.02.002>
- Habibzadeh, F. (2017). How to Report the Results of Public Health Research. *Journal of Public Health and Emergency*, 1, 90–90. <https://doi.org/10.21037/jphe.2017.12.02>
- Hagenauer, G., & Hascher, T. (2014). Early Adolescents' Enjoyment Experienced in Learning Situations at School and Its Relation to Student Achievement. *Journal of Education and Training Studies*, 2(2), 20–30. <https://doi.org/10.11114/jets.v2i2.254>
- Hamada, M. S., Wilson, A. G., Reese, C. S., & Martz, H. F. (2008). *Bayesian Reliability*. Springer Science+Business Media, LLC.
- Hardle, W. (1994). *Applied Nonparametric Regression*. <https://doi.org/10.2307/2533418>
- Held, L., & Bové, D. (2014). *Applied Statistical Inference: Likelihood and Bayes*. Springer.
- Held, L., & Ott, M. (2018). On p-Values and Bayes Factors. *Annual Review of Statistics and Its Application*, 5, 1–27.

- Hjort, N. L., Holmes, C., Müller, P., & Walker, S. G. (2010). *Bayesian Nonparametrics*. Cambridge University Press.
- Ibrahim, J. G., & Chen, M. H. (2000). Power Prior Distributions for Regression Models. *Statistical Science*, 15(1), 46–60. <https://doi.org/10.1214/ss/1009212673>
- IEA. (2020). Measuring Global Education Goals: How TIMSS Helps. *IEA-TIMSS*, 15. http://timssandpirls.bc.edu/timss2019/international-results/%0Ahttps://unesdoc.unesco.org/notice?id=p::usmarcdef_0000375119
- Jackman, S. (2009). *Bayesian Analysis for the Social Sciences*. A John Wiley and Sons.
- Jaynes, E. T. (1983). *Papers on Probability, Statistics and Statistical Physics*. Kluwer Academic Publisher. <https://doi.org/10.1007/bf00046910>
- Johnson, V. E., & Albert, J. H. (1999). *Statistics for Social Science and Public Policy*. Springer.
- Kalender, İ. (2015). An Analysis of the Profile of Resilient Students based on PISA 2012. *Journal of Measurement and Evaluation in Education and Psychology*, 6(1), 158–172. <https://doi.org/10.21031/epod.16925>
- Kalli, M., Griffin, J. E., & Walker, S. G. (2011). Slice Sampling Mixture Models. *Statistics and Computing*, 21(1), 93–105. <https://doi.org/10.1007/s11222-009-9150-y>
- Karabatsos, G. (2015). *Bayesian Nonparametric Regression for Educational Research*.
- Karabatsos, G. (2017). A Menu-Driven Software Package of Bayesian Nonparametric (and Parametric) Mixed Models for Regression Analysis and Density Estimation. *Behavior Research Methods*, 49(1), 335–362. <https://doi.org/10.3758/s13428-016-0711-7>
- Karabatsos, G. (2020). Fast Search and Estimation of Bayesian Nonparametric Mixture Models Using a Classification Annealing EM Algorithm. *Journal of Computational and Graphical Statistics*, 30(1), 236–247. <https://doi.org/10.1080/10618600.2020.1807995>
- Kass, R. E., & Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, 90(430), 773–795.
- Keele, L. (2008). *Semiparametric Regression for the Social Sciences*. John Wiley & Sons Inc. <https://doi.org/10.1002/9780470998137>

- Koch, K.-R. (2007). *Introduction to Bayesian Statistics*. Springer.
- König, C., & van de Schoot, R. (2018). Bayesian Statistics in Educational Research: A Look at the Current State of Affairs. *Educational Review*, 70(4), 486–509. <https://doi.org/10.1080/00131911.2017.1350636>
- Konstantopoulos, S. (2009). Effects of Teachers on Minority and Disadvantaged Students' Achievement in the Early Grades. *Elementary School Journal*, 110(1), 92–113. <https://doi.org/10.1086/598845>
- Koop, G. (2003). *Bayesian Econometrics*. John Wiley & Sons.
- Koop, G., Poirier, D. J., & Tobias, J. L. (2007). *Bayesian Econometric Methods*. Cambridge University Press.
- Kruschke, J. K. (2010). *Doing Bayesian Data Analysis: A Tutorial with R and BUGS*.
- Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher Sorting and the Plight of Urban Schools: A Descriptive Analysis. *Educational Evaluation and Policy Analysis*, 24(1), 37–62. <https://doi.org/10.3102/01623737024001037>
- Lee, J., & MacEachern, S. N. (2014). Inference Functions in High Dimensional Bayesian Inference. *Statistics and its Interface*, 7(4), 477–486. <https://doi.org/10.4310/SII.2014.v7.n4.a5>
- Li, Q., & Liny, N. (2010). The Bayesian Elastic Net. *Bayesian Analysis*, 5(1), 151–170. <https://doi.org/10.1214/10-BA506>
- Lynch, S. M. (2007). *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists*. Springer.
- Ma, X. (2003). Sense of Belonging to School: Can Schools Make a Difference? *Journal of Educational Research*, 96(6), 340–349. <https://doi.org/10.1080/00220670309596617>
- Maceachern, S. N. (2007). Comment on Article by Jain and Neal. *Bayesian Analysis*, 2(3), 483–494.
- MacEachern, S. N. (1999). Dependent Nonparametric Processes. In *Bayesian Methods with Applications to Science, Policy, and Official Statistics*, 1–6.

- MacEachern, S. N. (2000). Dependent Dirichlet Processes. *In Bayesian Methods with Applications to Science, Policy, and Official Statistics*, 1–41.
- MacEachern, S. N. (2001). Decision Theoretic Aspects of Dependent Nonparametric Processes. *In Bayesian Methods with Applications to Science, Policy, and Official Statistics*, 551–560.
- MacEachern, S. N. (2016). Nonparametric Bayesian Methods: A Gentle Introduction and Overview. *Communications for Statistical Applications and Methods*, 23(6), 445–466. <https://doi.org/10.5351/csam.2016.23.6.445>
- Maceachern, S. N., & Müller, P. (1998). Estimating Mixture of Dirichlet Process Models. *Journal of Computational and Graphical Statistics*, 7(2), 223–238. <https://doi.org/10.1080/10618600.1998.10474772>
- Martin, A., & Marsh, H. (2006). Academic Resilience and Its Psychological and Educational Correlates: A Construct Validity Approach. *Psychology in the Schools*, 43(3), 267–281. <https://doi.org/10.1002/pits.20149>
- Martin, M., Davier, M. von, & Mullis, I. V. S. (2020). *TIMSS 2019 Technical Report*. TIMSS & PIRLS International Study Center, Lynch School of Education and Human Development, Boston College and International Association for the Evaluation of Educational Achievement (IEA). <https://timssandpirls.bc.edu/timss2019/methods/>
- Martin, M. O., & Mullis, I. V. . (2017). TIMSS Assessment Frameworks. *TIMSS & PIRLS International Study Center*, 1–40.
- McCullagh, P., & Nelder, J. A. (1983). *Generalized Linear Models*. Chapman and Hall. [https://doi.org/10.1016/0377-2217\(84\)90282-0](https://doi.org/10.1016/0377-2217(84)90282-0)
- McElreath, R. (2019). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*.
- McNeish, D. (2016). On Using Bayesian Methods to Address Small Sample Problems. *Structural Equation Modeling: A Multidisciplinary Journal*, 23(5), 750–773.
- Michaelides, M. P., Brown, G. T. L., Eklof, H., & Papanastasiou, E. C. (2019). The Relationship of Motivation with Achievement in Mathematics. *İçinde Motivational Profiles in TIMSS Mathematics* (ss. 9–23). <https://doi.org/10.1007/978-3-030-26183-2>

- Müller, P., Quintana, F. A., Jara, A., & Hanson, T. E. (2015). *Bayesian Nonparametric Data Analysis*. Springer Series.
- Mullis, I. V. S., Martin, M., Foy, P., & Arora, A. (2012). TIMSS 2011 International Results in Mathematics. İçinde *TIMSS & PIRLS International Study Center*. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3295935&tool=pmcentrez&rendertype=abstract>
- Mullis, I. V. S., Martin, M., Foy, P., Kelly, D. L., & Fishbein, B. (2020). TIMSS 2019 International Results in Mathematics and Science. İçinde *TIMSS & PIRLS International Study Center*. <https://timssandpirls.bc.edu/timss2019/international-results/>
- Mun, E.-Y., Eye, A. von, Bates, M. E., & Vaschillo, E. G. (2008). Finding Groups Using Model-based Cluster Analysis: Heterogeneous Emotional Self-regulatory Processes and Heavy Alcohol Use Risk. *International Journal of Obesity*, 32(6), 481–495. <https://doi.org/10.1037/0012-1649.44.2.481>.Finding
- NCES. (2012). Improving the Measurement of Socioeconomic Status for the National Assessment of Educational Progress. İçinde *National Center for Education Statistics*. <https://doi.org/10.1016/b978-0-407-72903-2.50010-1>
- Neal, R. M. (1995). *Bayesian Learning for Neural Networks*.Doktora Tezi. [Toronto Üniversitesi]. <https://doi.org/10.2307/2965731>
- Neal, R. M. (2000). Markov Chain Sampling Methods for Dirichlet Process Mixture Models. *Journal of Computational and Graphical Statistics*, 9(2), 249–265. <https://doi.org/10.1080/10618600.2000.10474879>
- Neal, R. M. (2011). MCMC Using Hamiltonian Dynamics. İçinde *Handbook of Markov Chain Monte Carlo* (ss. 113–162). Chapman and Hall/CRC.
- Ntzoufras, I. (2009). *Bayesian Modeling Using WinBUGS*. John Wiley & Sons.
- O’Hagan, A. (2004). *Bayesian Statistics: Principles and Benefits*. University of Sheffield, Department of Probability and Statistics.
- OECD. (2010). PISA 2009 Results : Overcoming Social Background - Equity in Learning Opportunities and Outcomes(Volume II). İçinde *OECD Publishing*.

<http://dx.doi.org/10.1787/9789264091504-en>

OECD. (2011). *Against the Odds - Disadvantaged Students Who Succeed in School*. İçinde *OECD Publishing*. http://www.oecd-ilibrary.org/education/against-the-odds_9789264090873-en

OECD. (2013). *PISA 2012 Results: Excellence Through Equity: Giving Every Student the Chance to Succeed (Volume II)*. İçinde *OECD Publishing*. <https://doi.org/10.1787/9789264201132-en>

OECD. (2015). *The ABC of Gender Equality in Education*. İçinde *OECD Publishing*. www.oecd.org/pisa

OECD. (2016). *PISA 2015 Results: Policies and Practices for Successful Schools(Volume II)*. İçinde *OECD Publishing*. <http://dx.doi.org/10.1787/9789264267510-en>

OECD. (2018). *Equity in Education: Breaking Down Barriers to Social Mobility*. İçinde *OECD Publishing*. <http://dx.doi.org/10.1787/9789264073234-en> https://www.oecd-ilibrary.org/education/equity-in-education_9789264073234-en

OECD. (2019a). *PISA 2018 Results (Volume II): Where All Students Can Succeed*. İçinde *OECD Publishing* (C. 2). https://www.oecd.org/pisa/publications/PISA2018_CN_IDN.pdf

OECD. (2019b). *Sweden Country Note*. İçinde *OECD Publishing*. https://www.oecd.org/pisa/publications/PISA2018_CN_IDN.pdf

Önder, E., & Uyar, Ş. (2018). Factors Affecting the Academic Achievement in Socioeconomically Disadvantaged Students. *Pegem Eğitim ve Öğretim Dergisi*, 8(2), 253–280. <https://doi.org/10.14527/pegegog.2018.011>

Orbanz, P. (2014). *Lecture Notes on Bayesian Nonparametrics*. İçinde *Columbia University*. https://doi.org/10.1057/978-1-349-95189-5_2596

Orbanz, P., & Teh, Y. W. (2010). Bayesian Nonparametric Models. *Encyclopedia of Machine Learning*, 1–14. https://doi.org/10.1007/978-0-387-30164-8_66

Özden, C., & Atasoy, R. (2020). Socioeconomically Disadvantaged and Resilient Students'

- Characteristics and Their Perceptions about School. *International Journal of Progressive Education*, 16(5), 70–89. <https://doi.org/10.29329/ijpe.2020.277.5>
- Press, S. J. (2003). *Subjective and Objective Bayesian Statistics*. John Wiley & Sons.
- Pupovac, V., & Petrovecki, M. (2011). Summarizing and Presenting Numerical Data. *Biochemia Medica*, 21(2), 106–110.
- Quintana, F. A., Mueller, P., Jara, A., & MacEachern, S. N. (2020). The Dependent Dirichlet Process and Related Models. *Submitted to Statistical Science*, 1–33. <http://arxiv.org/abs/2007.06129>
- Radišić, J., & Pettersen, A. (2021). Resilient and Nonresilient Students in Sweden and Norway- Investigating the Interplay Between Their Self-Beliefs and the School Environment. *Içinde Equity, Equality and Diversity in the Nordic Model of Education* (ss. 273–304). https://doi.org/10.1007/978-3-030-61648-9_11
- Raftery, A. E., & Dean, N. (2006). Variable Selection for Model-Based Clustering. *Journal of the American Statistical Association*, 101(473), 168–178. <https://doi.org/10.1198/016214506000000113>
- Reich, B. J., & Ghosh, S. K. (2019). *Bayesian Statistical Methods*. CRC Press.
- Reynolds, J., Stewart, M., MacDonald, R., & Sischo, L. (2006). Have Adolescents Become Too Ambitious? High School Seniors' Educational and Occupational Plans, 1976 to 2000. *Social Problems*, 53(2), 186–206. <https://doi.org/10.1525/sp.2006.53.2.186>
- Richardson, R., & Hartman, B. (2018). Bayesian Nonparametric Regression Models for Modeling and Predicting Healthcare Claims. *Insurance: Mathematics and Economics*, 83, 1–24. <https://doi.org/10.1016/j.insmatheco.2018.06.002>
- Robert, C. P. (2007). *The Bayesian Choice From Decision-Theoretic Foundations to Computational Implementation*. Springer Science+Business Media, LLC.
- Robert, C. P., & Casella, G. (2004). *Monte Carlo Statistical Methods*. Springer Science+Business Media Inc.
- Robert, C. P., & Casella, G. (2008). *Introducing Monte Carlo Methods with R*. Springer.

- Rutkowski, D., & Rutkowski, L. (2018). How Systematic is International Bullying and What Relationship Does It Have with Mathematics Achievement in 4th Grade? İçinde *IEA Compass: Briefs in Education*.
- Rutkowski, L., & Rutkowski, D. (2016). The Relation Between Students' Perceptions of Instructional Quality and Bullying Victimization. İçinde *IEA Research for Education* (C. 2, ss. 115–133). https://doi.org/10.1007/978-3-319-41252-8_6
- Saha, L. J., & Sikora, J. (2012). The Career Aspirations and Expectations of School Students: From Individual to Global Effects. *Education and Society*, 26(2), 5–22. <https://doi.org/10.7459/es/26.2.02>
- Sandoval-Hernández, A., & Białowolski, P. (2016). Factors and Conditions Promoting Academic Resilience: A TIMSS-Based Analysis of Five Asian Education Systems. *Asia Pacific Education Review*, 17(3), 511–520. <https://doi.org/10.1007/s12564-016-9447-4>
- Sandoval-Hernández, A., & Cortes, D. (2012). Factors and Conditions that Promote Academic Resilience: A Cross-Country Perspective. *Conference: International Congress for School Effectiveness Improvement (ICSEI), January 2012*.
- Sanyal, P., & Ehlen, M. (2017). The Interactions of Public Debt, Income Inequality and Economic Growth for U.S. States: A Bayesian Non-Parametric Analysis. *Empirical Economics Review*, 7(1). <https://www.researchgate.net/publication/316091923>
- Scherer, R., & Nilsen, T. (2016). The Relations Among School Climate, Instructional Quality, and Achievement Motivation in Mathematics. İçinde *Teacher Quality, Instructional Quality and Student Outcomes*, *IEA Research for Education* (C. 2, ss. 51–80). https://doi.org/10.1007/978-3-319-41252-8_3
- Schervish, M. J. (1995). *Theory of Statistics*. Springer-Verlag New York. <https://doi.org/10.2307/2988561>
- Shemyakin, A., & Kniazev, A. (2017). *Introduction to Bayesian Estimation and Copula Models of Dependence*. John Wiley & Sons.
- Siebecke, D. E., & Jarl, M. (2022). Does the Material Well-Being at Schools Successfully Compensate for Socioeconomic Disadvantages? Analysis of Resilient Schools in Sweden.

- Large-Scale Assessments in Education*, 10(1). <https://doi.org/10.1186/s40536-022-00130-y>
- Sikora, J., & Saha, L. J. (2007). Corrosive Inequality? Structural Determinants of Educational and Occupational Expectations in Comparative Perspective. *International Education Journal*, 8(3), 57–78.
- Şirin, S. R. (2005). Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research. *Review of Educational Research*, 75(3), 417–453. <https://doi.org/10.3102/00346543075003417>
- Sivia, D. S., & J. Skilling. (2006). *Data Analysis: A Bayesian Tutorial*. Oxford University Press.
- Stankov, L. (2010). Unforgiving Confucian Culture: A Breeding Ground for High Academic Achievement, Test Anxiety and Self-Doubt? *Learning and Individual Differences*, 20(6), 555–563. <https://doi.org/10.1016/j.lindif.2010.05.003>
- Steiner-Khamsi, G. (2004). *Global Politics of Educational Borrowing and Lending*. Teachers College Press.
- Strøm, B. (2004). Student Achievement and Birthday Effects. *Norwegian University of Science and Technology working paper*, 2004.
- Takashiro, N. (2015). *A Multilevel Analysis of Japanese Middle School Student and School Socioeconomic Status Influence on Mathematics Achievement*. Yayınlanmış Doktora Tezi.
- Teh, Y. W. (2011). Dirichlet Process. *Encyclopedia of Machine Learning*, 280–287. https://doi.org/10.1007/978-0-387-30164-8_219
- Teh, Y. W., Jordan, M. I., Beal, M. J., & Blei, D. M. (2006). Hierarchical Dirichlet processes. *Journal of the American Statistical Association*, 101(476), 1566–1581. <https://doi.org/10.1198/016214506000000302>
- Topal, K. H. (2021). Variable Selection via the Adaptive Elastic Net: Mathematics Success of the Students in Singapore and Turkey. *Journal of Applied Microeconometrics*, 1(1), 41–55. <https://doi.org/10.53753/jame.1.1.04>
- Topçu, M. S., Erbilgin, E., & Arıkan, S. (2016). Factors Predicting Turkish and Korean

- Students' Science and Mathematics Achievement in TIMSS 2011. *Eurasia Journal of Mathematics, Science and Technology Education*, 12(7), 1711–1737. <https://doi.org/10.12973/eurasia.2016.1530a>
- Türkman, M. A. A., Paulino, C. D., & Müller, P. (2019). *Computational Bayesian Statistics*. Cambridge University Press.
- UNDP. (2019). Human Development Report 2019: beyond income, beyond averages, beyond today. İçinde *United Nations Development Programme*.
- UNDP. (2020). Human Development Report 2020. İçinde *United Nations Development Programme*. <https://hdr.undp.org/sites/default/files/hdr2020.pdf>
- UNESCO. (2017). School violence and bullying: Global status report. İçinde *United Nations Educational, Scientific and Cultural Organization*. <https://healtheducationresources.unesco.org/library/documents/school-violence-and-bullying-global-status-and-trends-drivers-and-consequences>
- UNESCO. (2018). Quick Guide to Education Indicators for SDG 4. İçinde *UNESCO Institute for Statistics*. <http://uis.unesco.org/sites/default/files/documents/quick-guide-education-indicators-sdg4-2018-en.pdf>
- Wagemaker, H., & Mirazchiyski, P. (2023). Cyberbullying and Student Learning. İçinde *IEA Compass briefs in education*.
- Walker, S. G. (2007). Sampling the Dirichlet Mixture Model with Slices. *Communications in Statistics: Simulation and Computation*, 36(1), 45–54. <https://doi.org/10.1080/03610910601096262>
- Wang, F., King, R., & Leung, S. L. (2022). Beating the Odds: Identifying the Top Predictors of Resilience among Hong Kong Students. *Child Indicators Research*, 15(5), 1921–1944. <https://doi.org/10.1007/s12187-022-09939-z>
- Wang, M. C., Haertel, G. D., & Walberg, H. J. (1994). Educational Resilience in Inner Cities. İçinde *Educational resilience in inner-city America: Challenges and prospects*. (ss. 45–72).
- Wang, M. C., Haertel, G. D., & Walberg, H. J. (1998). Educational Resilience. İçinde *National*

Research Center on Education in the Inner Cities. https://doi.org/10.1007/978-0-387-71799-9_155

Willms, J. D., & Tramate, L. (2015). Towards the Development of Contextual Questionnaires for the PISA for Development Study. İçinde *OECD Education Working Papers No. 118*. <https://doi.org/https://dx.doi.org/10.1787/5js1kv8crsjf-en>

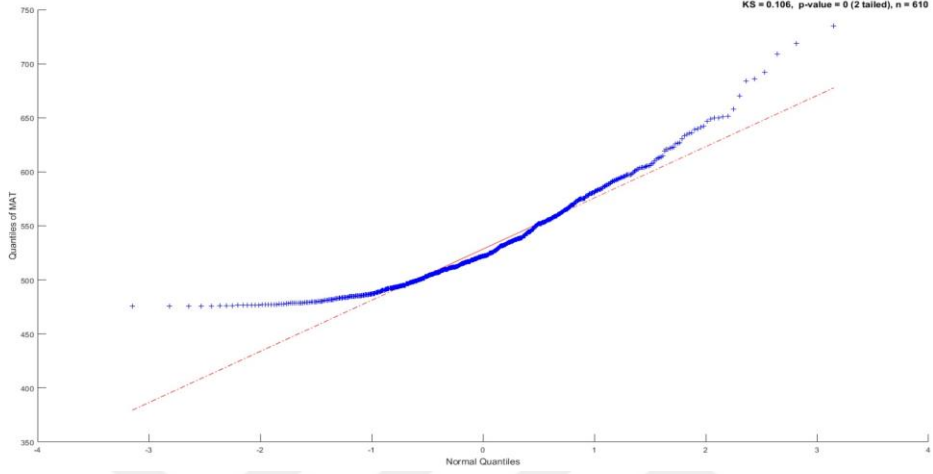
Yang Hansen, K., & Gustafsson, J. E. (2019). Identifying the Key Source of Deteriorating Educational Equity in Sweden Between 1998 and 2014. *International Journal of Educational Research*, 93(September 2018), 79–90. <https://doi.org/10.1016/j.ijer.2018.09.012>

Yatchew, A. (1998). Nonparametric Regression Techniques in Economics. *Journal of Economic Literature*, 36(2), 669–721.

Ye, W., Strietholt, R., & Blömeke, S. (2021). Academic Resilience: Underlying Norms and Validity of Definitions. *Educational Assessment, Evaluation and Accountability*, 33(1), 169–202. <https://doi.org/10.1007/s11092-020-09351-7>

6.EKLER

Ek 1: Kolmogorov-Smirnov testi ve Q-Q grafiği



Ek 2: Dirichlet Süreç Karışım Model Parametrelerinin Marjinal Son Dağılım Özeti

Parametre	Ortalama	Medyan	Standart Sapma	25%	75%	MCSE	CUSUM
beta0Sample	-0,137	-0,094	0,33	-0,364	0,099	0,004	0,475
betaSample:Z:ITSEX	-0,057	-0,075	0,208	-0,172	0,09	0,003	0,476
betaSample:Z:BSBG07*	0,151	0,148	0,217	0,01	0,272	0,004	0,503
betaSample:Z:BSBM19G	-0,048	-0,02	0,217	-0,159	0,074	0,004	0,481
betaSample:Z:BSDAGE	0,059	0,039	0,214	-0,058	0,143	0,005	0,484
betaSample:Z:BSBGSSB	-0,084	-0,079	0,198	-0,187	0,015	0,003	0,479
betaSample:Z:BSBGSB	-0,057	-0,071	0,197	-0,131	0,022	0,009	0,422
betaSample:Z:BSBGSCM*	0,287	0,287	0,213	0,184	0,393	0,016	0,426
sigma^2Sample*	1,701	0,943	3,428	0,566	1,325	0,272	0,477
ICC*	0,129	0,108	0,088	0,066	0,176	0,009	0,476
Alpha (α)*	12,038	11,082	5,805	7,854	15,162	0,329	0,474

Not: Dirichlet süreç karışım modelinin karışım parametrelerinin son dağılım tahmininden örnekleri, rastgele etkilerin son dağılımını göstermektedir. Sıfırı içermeyen %25-%75 son dağılım kantiller arası güvenilir aralık için anlamlı parametreler * ile gösterilmektedir. ICC, gruplar arası korelasyon katsayısının son dağılımı gruplar arası varyansın toplam varyansa oranını ifade etmektedir. Alpha (α), Dirichlet süreç kesinlik parametresinin marjinal son dağılım tahminin göstermektedir.

Ek 3: Model Sabit Terimi Son Dağılım Özeti

Parametre	Ortalama	Medyan	Standart Sapma	25%	75%	MCSE	CUSUM
beta0:ID=36*	0,134	0,111	0,125	0,055	0,188	0,023	0,178
beta0:ID=152*	-0,631	-0,643	0,055	-0,643	-0,589	0,002	0,003
beta0:ID=246*	-0,591	-0,643	0,148	-0,643	-0,589	0,001	0,04
beta0:ID=250*	-0,624	-0,643	0,078	-0,643	-0,589	0,061	0,008
beta0:ID=348	0,001	0,024	0,168	-0,105	0,109	0,01	0,36
beta0:ID=372*	-0,107	-0,118	0,098	-0,16	-0,056	0,012	0,135
beta0:ID=380*	-0,133	-0,142	0,071	-0,172	-0,093	0,013	0,033
beta0:ID=392*	0,208	0,2	0,209	0,079	0,33	0,024	0,171
beta0:ID=410*	0,149	0,131	0,186	0,048	0,222	0,021	0,293
beta0:ID=440	-0,031	-0,028	0,17	-0,137	0,092	0,002	0,409
beta0:ID=554	0,078	0,092	0,147	-0,019	0,172	0,018	0,274
beta0:ID=578*	-0,625	-0,643	0,077	-0,643	-0,589	0,07	0,011
beta0:ID=620*	-0,128	-0,142	0,068	-0,172	-0,085	0,01	0,03
beta0:ID=752	0,007	0,022	0,16	-0,105	0,107	0,006	0,369
beta0:ID=792*	0,076	0,086	0,085	0,023	0,134	0,009	0,025
beta0:ID=840	-0,055	-0,066	0,095	-0,133	0,016	0,011	0,096

Not: Sıfırı içermeyen %25-%75 son dağılım kantiller arası güvenilir aralık için anlamlı parametreler * ile gösterilmektedir.

Ek 4: Regresyon Hata Varyans Parametresi İçin Son Dağılım Özeti

Parametre	Ortalama	Medyan	Standart Sapma	25%	75%	MCSE	CUSUM
s2:ID= 36*	0,987	0,974	0,146	0,912	1,042	0,016	0,165
s2:ID=152*	0,325	0,317	0,028	0,317	0,318	0,009	0,002
s2:ID=246*	0,343	0,318	0,082	0,317	0,318	0,032	0,065
s2:ID=250*	0,327	0,317	0,039	0,317	0,318	0,022	0,007
s2:ID=348*	0,864	0,912	0,263	0,655	1,003	0,055	0,367
s2:ID=372*	0,621	0,587	0,14	0,529	0,655	0,011	0,116
s2:ID=380*	0,609	0,593	0,101	0,558	0,653	0,053	0,044
s2:ID=392*	1,206	1,12	0,344	0,967	1,354	0,094	0,161
s2:ID=410*	1,049	0,992	0,331	0,909	1,104	0,08	0,306
s2:ID=440*	0,908	0,921	0,325	0,655	1,016	0,089	0,428
s2:ID=554*	0,974	0,966	0,19	0,889	1,048	0,03	0,265
s2:ID=578*	0,328	0,317	0,044	0,317	0,318	0,024	0,009
s2:ID=620*	0,606	0,583	0,102	0,558	0,653	0,059	0,028
s2:ID=752*	0,89	0,921	0,276	0,676	1,011	0,059	0,375
s2:ID=792*	0,987	0,982	0,087	0,94	1,04	0,007	0,023
s2:ID=840*	0,904	0,92	0,169	0,799	1,011	0,017	0,085

Not: Sıfırı içermeyen %25-%75 son dağılım kantiller arası güvenilir aralık için anlamlı parametreler * ile gösterilmektedir.

Ek 5: Dirichlet Süreç Temel Dağılım Parametrelerinin Son Dağılım Özeti

Parametre	Ortalama	Medyan	Standart Sapma	MCSE
mu:beta0	-0,093	-0,094	0,154	0,002
mu:Z:ITSEX	-0,069	-0,069	0,133	0,006
mu:Z:BSBG07	0,157	0,156	0,134	0,001
mu:Z:BSBM19G	-0,063	-0,062	0,135	0,004
mu:Z:BSDAGE	0,071	0,071	0,139	0,009
mu:Z:BSBGSSB	-0,097	-0,099	0,132	0,003
mu:Z:BSBGSB	-0,047	-0,047	0,128	0,001
mu:Z:BSBGSCM	0,294	0,292	0,136	0,007
tau:beta0	0,119	0,107	0,054	0,014
tau:beta0:beta:Z:ITSEX	-0,01	-0,008	0,033	0,002
tau:beta0:beta:Z:BSBG07	0,011	0,01	0,031	0,001
tau:beta0:beta:Z:BSBM19G	-0,013	-0,01	0,031	0,001
tau:beta0:beta:Z:BSDAGE	0,004	0,003	0,032	0,001
tau:beta0:beta:Z:BSBGSSB	-0,007	-0,006	0,031	0,002
tau:beta0:beta:Z:BSBGSB	0,002	0,002	0,03	0,001
tau:beta0:beta:Z:BSBGSCM	0,008	0,006	0,031	0,002
tau:beta:Z:ITSEX	0,088	0,079	0,039	0,011
tau:beta:Z:ITSEX:beta:Z:BSBG07	-0,004	-0,003	0,026	0,002
tau:beta:Z:ITSEX:beta:Z:BSBM19G	0,001	0,001	0,026	0,001
tau:beta:Z:ITSEX:beta:Z:BSDAGE	0,003	0,002	0,027	0,002
tau:beta:Z:ITSEX:beta:Z:BSBGSSB	0,002	0,002	0,026	0,001
tau:beta:Z:ITSEX:beta:Z:BSBGSB	-0,001	-0,001	0,025	0,001
tau:beta:Z:ITSEX:beta:Z:BSBGSCM	0,001	0,001	0,027	0,002
tau:beta:Z:BSBG07	0,088	0,08	0,038	0,009
tau:beta:Z:BSBG07:beta:Z:BSBM19G	-0,003	-0,002	0,027	0,002
tau:beta:Z:BSBG07:beta:Z:BSDAGE	-0,001	0,001	0,026	0,001
tau:beta:Z:BSBG07:beta:Z:BSBGSSB	-0,004	-0,003	0,026	0,001
tau:beta:Z:BSBG07:beta:Z:BSBGSB	-0,001	0,001	0,026	0,001
tau:beta:Z:BSBG07:beta:Z:BSBGSCM	0,003	0,002	0,026	0,001
tau:beta:Z:BSBM19G	0,092	0,083	0,041	0,006
tau:beta:Z:BSBM19G:beta:Z:BSDAGE	-0,006	-0,005	0,028	0,001
tau:beta:Z:BSBM19G:beta:Z:BSBGSSB	0,001	0,001	0,026	0,002
tau:beta:Z:BSBM19G:beta:Z:BSBGSB	0,001	0,001	0,026	0,001
tau:beta:Z:BSBM19G:beta:Z:BSBGSCM	-0,008	-0,006	0,027	0,001

tau:beta:Z:BSDAGE	0,091	0,082	0,04	-0,005
tau:beta:Z:BSDAGE:beta:Z:BSBGSSB	0,001	0,001	0,026	0,001
tau:beta:Z:BSDAGE:beta:Z:BSBGSB	0,002	0,002	0,026	0,001
tau:beta:Z:BSDAGE:beta:Z:BSBGSCM	0,005	0,004	0,027	0,001
tau:beta:Z:BSBGSSB	0,085	0,077	0,037	0,008
tau:beta:Z:BSBGSSB:beta:Z:BSBGSB	-0,002	-0,002	0,024	0,001
tau:beta:Z:BSBGSSB:beta:Z:BSBGSCM	-0,001	-0,001	0,025	0,001
tau:beta:Z:BSBGSB	0,083	0,075	0,037	0,009
tau:beta:Z:BSBGSB:beta:Z:BSBGSCM	0,001	0,001	0,025	0,001
tau:beta:Z:BSBGSCM	0,089	0,081	0,039	0,008

