

**T.C.
ESKİŞEHİR OSMANGAZİ ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
BİYOİSTATİSTİK ANABİLİM DALI**

**BAYESGİL YAŞAM ANALİZİ VE COX REGRESYON YAŞAM
ANALİZİ'NİN TÜRETİLMİŞ VE GERÇEK VERİ SETLERİNDE
UYGULANMASI**

DOKTORA TEZİ

İMİRAN KURT

**DANIŞMANLAR
PROF.DR. KAZIM ÖZDAMAR
DOÇ.DR. MEVLÜT TÜRE**

MAYIS 2008

T.C.
ESKİŞEHİR OSMANGAZİ ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
BİYOİSTATİSTİK ANABİLİM DALI

BAYESGİL YAŞAM ANALİZİ VE COX REGRESYON YAŞAM
ANALİZİ'NİN TÜRETİLMİŞ VE GERÇEK VERİ SETLERİNDE
UYGULANMASI

DOKTORA TEZİ

İMİRAN KURT

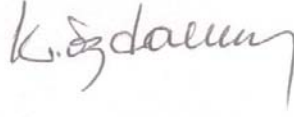
DANIŞMANLAR
PROF.DR. KAZIM ÖZDAMAR
DOÇ.DR. MEVLÜT TÜRE

KABUL VE ONAY SAYFASI

İmran KURT'un Doktora Tezi olarak hazırladığı "Bayesgil Yaşam Analizi ve Cox Regresyon Yaşam Analizi'nin Türetilmiş ve Gerçek Veri Setlerinde Uygulanması" başlıklı bu çalışma Eskişehir Osmangazi Üniversitesi Lisansüstü Eğitim ve Öğretim Yönetmeliği'nin ilgili maddesi uyarınca değerlendirilerek "KABUL" edilmiştir.

09.05.2008

Üye: Prof.Dr. Kazım ÖZDAMAR



Üye: Doç.Dr. Mevlüt TÜRE



Üye: Yrd.Doç.Dr. Setenay ÖNER




Üye: Yrd.Doç.Dr. Fezan MUTLU



Üye: Yrd.Doç.Dr. Canan BAYDEMİR



Eskişehir Osmangazi Üniversitesi Sağlık Bilimleri Enstitüsü Yönetim Kurulu'nun 15.05./2008.. tarih ve 742./2433.. sayılı kararı ile onaylanmıştır.



Prof.Dr. Ferruh YÜCEL

Enstitü Müdürü

İÇİNDEKİLER

İÇİNDEKİLER.....	i
ÖZET.....	iii
SUMMARY.....	iv
TABLO DİZİNİ.....	v
ŞEKİL DİZİNİ.....	viii
SİMGE VE KISALTMALAR DİZİNİ.....	xii
1. GİRİŞ VE AMAÇ.....	1
2. GENEL BİLGİLER.....	3
2.1. Cox Regresyon (Orantısal Hazard Regresyon) Analizi.....	3
2.1.1. Cox Regresyon Modeli.....	4
2.1.2. En Çok Olabilirlik Parametre Tahmin Yöntemi.....	6
2.1.3. Katsayıların Önemliliğinin Test Edilmesi.....	9
2.1.4. Regresyon Katsayılarının Yorumlanması.....	10
2.2. Bayesgil Yaşam Analizi.....	15
2.2.1. Bayes Teoremi.....	15
2.2.2. Bayesgil Çıkarsama.....	17
2.2.3. Markov Zinciri Monte Carlo Yöntemi.....	19
2.2.4. Katsayıların Prior Dağılımları.....	23
2.2.5. Markov Zincirinin Yakınsamasının Değerlendirilmesi.....	25
2.2.6. Bayesgil Analizin Özet İstatistikleri.....	28
2.2.7. Bayesgil Analizin Uygunluk İstatistikleri.....	29
3. GEREÇ VE YÖNTEM.....	31
3.1. Kullanılan Programlar.....	31
3.2. Simülasyon Çalışması.....	31
3.2.1. Simülasyon Algoritmaları.....	32
3.2.2. Simülasyon Parametreleri.....	33
3.2.3. Karşılaştırma Ölçütleri.....	34
3.3. Gerçek Veri Seti Çalışması.....	35
3.3.1. Modele Girecek Bağımsız Değişkenlerin Aşamalı Seçimi.....	37

3.3.2. Yaşam Fonksiyonlarının Hesaplanması	38
4. BULGULAR.....	39
4.1. Simülasyon Sonuçları.....	39
4.1.1. Açıklayıcı Olmayan Prior Bilgiye Dayalı Simülasyon Sonuçları	39
4.1.2. Açıklayıcı Prior Bilgiye Dayalı Simülasyon Sonuçları.....	48
4.2. Gerçek Uygulama Sonuçları	71
5. TARTIŞMA	88
6. SONUÇ VE ÖNERİLER.....	93
KAYNAKLAR DİZİNİ.....	96
ÖZGEÇMİŞ	101

ÖZET

Bu çalışmada, Bayesgil Yaşam Analizi (BYA) ile Cox Regresyon Analizi (CRA) yöntemlerinin performanslarının, farklı koşullar altında yapılan simülasyon çalışmaları ve gerçek bir uygulamayla karşılaştırılması amaçlandı.

Simülasyon çalışması, açıklayıcı olan ve olmayan prior bilgiye dayalı iki farklı algoritmaya göre veri türetimi biçiminde yapıldı. Uygulamada kullanılan gerçek veri seti ise Trakya Üniversitesi Tıp Fakültesi Radyasyon Onkolojisi Anabilim Dalı'na 1998-2007 yılları arasında başvuran ve meme kanseri tanısı konulan 423 hastanın yinelemesiz yaşam sürelerine ilişkin veri seti idi.

Simülasyon uygulamasında, açıklayıcı olmayan prior bilgili veri setlerinde BYA ve CRA yöntemlerinin simülasyon parametresine yakınsama bakımından benzer performans sergilediği gözlemlendi. Açıklayıcı prior bilgili simülasyon uygulamasında ise veri yapısına uygun ve gerçeği yansıtan prior bilgi kullanılan BYA, oldukça küçük yanlılıkla iyi bir performans gösterdi. Her koşulda prior bilgi gerçeği yansıtmaktan uzaklaştıkça, BYA'nın yanlılığının arttığı belirlendi. Ayrıca BYA'da, gerçeği yansıtan prior bilgi ışığında hem az birim içeren hem de çok sayıda birim içeren örneklerde CRA'ya göre daha küçük yanlılık ve standart hatalı tahminler elde edildi.

Gerçek veri seti uygulamasında, aşamalı CRA, açıklayıcı prior bilgili BYA ve açıklayıcı olmayan prior bilgili BYA yöntemlerinde yaş, tümör büyüklüğü, hormon terapisi ve aksiller lenf nodu tutulumunun yinelemesiz yaşam süresi için önemli risk faktörleri olduğu bulundu. Ayrıca açıklayıcı prior bilgili BYA'da, parametre tahminlerinin standart hatalarının biraz daha küçük olduğu gözlemlendi.

Sonuç olarak; BYA, tahmin edilecek parametreler hakkında var olan bilgiler ve uzman görüşleri hesaba katılarak subjektif veri analizi gerçekleştirildiğinde, CRA'ya göre daha iyi performans göstermektedir. Bu nedenle yaşam sürelerine ilişkin verilerin analizinde, veri yapısına uygun ve gerçeği yansıtan prior bilgi olduğunda BYA yöntemi aksi durumda ise CRA yöntemi tercih edilmelidir.

Anahtar Kelimeler: Cox Regresyon, Bayes, Bayesgil Yaşam Analizi, Yaşam Süresi, Meme Kanseri, Markov Zinciri Monte Carlo, Prior, Posterior, Simülasyon

SUMMARY

In this study, it is aimed to compare the performance of Cox Regression (CRA) and Bayesian Survival Analyses (BSA) by using simulations that performed in different conditions and a real application.

Simulation study was carried out with two different algorithms that were informative and noninformative priors. Moreover, in a real data set application, a data set related to recurrence-free survivals that were obtained from 423 breast cancer patients diagnosed between years of 1998-2007 in Trakya University Medical Faculty Department of Radiation Oncology was used.

In the simulation application, it was observed that BSA with noninformative priors and CRA methods were showed similar performances in point of convergence to simulation parameter. In the informative priors' simulation application, BSA with reliable informative prior showed a good performance with too little bias. It was found out that bias of BSA increased while priors were becoming distant from reliability in all conditions. In addition, BSA obtained predictions with more little bias and standard error than the CRA in both of small and big samples in the light of reliable priors.

In the real data set application, age, tumor size, hormonal therapy, axillary nodal status were found statistically significant prognostic factors for recurrence-free survival in stepwise CRA, BSA with informative and noninformative priors. Furthermore, standard errors of predictions in BSA with informative priors were observed slightly little.

As a result, BSA shows better performance than CRA, when subjective data analysis performed by considering of expert opinions and historical knowledge about parameters. Consequently, BSA should be preferred in existence of reliable informative priors, in the contrast cases, CRA should be preferred.

Keywords: Cox Regression, Bayes, Bayesian Survival Analysis, Survival Time, Breast Cancer, Markov Chain Monte Carlo, Prior, Posterior, Simulation

TABLO DİZİNİ

Tablo 2.1. Dört kategorili bir değişkene ilişkin kukla değişkenler.....	13
Tablo 3.1. Meme kanserli hastalara ilişkin veri setinde yineleme durumuna göre bağımsız değişkenlerin tanımlayıcı istatistikleri	36
Tablo 3.2. Meme kanserli hastalara ilişkin veri setinde bağımsız değişkenlere ilişkin prior bilgiler	37
Tablo 4.1. $\beta=0$ için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri, standart hatalar ve yanlılıklar	40
Tablo 4.2. $\beta=1$ için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri, standart hatalar ve yanlılıklar	41
Tablo 4.3. $\beta=\ln(2)$ (0.6931) için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri, standart hatalar ve yanlılıklar	42
Tablo 4.4. Uniform ve binom dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=30$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri, standart hataları ve yanlılıkları.....	50
Tablo 4.5. Uniform ve binom dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=50$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri, standart hataları ve yanlılıkları.....	53
Tablo 4.6. Uniform ve binom dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=100$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri, standart hataları ve yanlılıkları.....	57

Tablo 4.7. Uniform ve binom dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=250$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri, standart hataları ve yanlılıkları.....	61
Tablo 4.8. Uniform ve binom dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=500$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri, standart hataları ve yanlılıkları.....	65
Tablo 4.9. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin CRA sonuçları.....	72
Tablo 4.10. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin açıklayıcı prior bilgili BYA sonuçları	74
Tablo 4.11. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin açıklayıcı olmayan prior bilgili BYA sonuçları.....	75
Tablo 4.12. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin CRA, açıklayıcı olan ve olmayan prior bilgili BYA için uyum istatistikleri	76
Tablo 4.13. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin açıklayıcı prior bilgili BYA için MCMC'nin yakınsama istatistikleri	77
Tablo 4.14. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin açıklayıcı olmayan prior bilgili BYA için MCMC'nin yakınsama istatistikleri	78
Tablo 4.15. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin aşamalı CRA sonuçları	80
Tablo 4.16. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin açıklayıcı prior bilgili aşamalı BYA sonuçları	80
Tablo 4.17. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin açıklayıcı olmayan prior bilgili aşamalı BYA sonuçları	81

Tablo 4.18. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin aşamalı CRA, açıklayıcı olan ve olmayan prior bilgili BYA için uyum istatistikleri.....	82
Tablo 4.19. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin açıklayıcı prior bilgili aşamalı BYA için MCMC'nin yakınsama istatistikleri	83
Tablo 4.20. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin açıklayıcı olmayan prior bilgili aşamalı BYA için MCMC'nin yakınsama istatistikleri.....	83

ŞEKİL DİZİNİ

Şekil 4.1. CRA ve BYA yöntemlerinin $\beta=0$ ve uniform dağılım için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminlerinin örneklem büyüklüğüne göre değişimi.....	43
Şekil 4.2. CRA ve BYA yöntemlerinin $\beta=0$ ve binom dağılım için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminlerinin örneklem büyüklüğüne göre değişimi.....	43
Şekil 4.3. CRA ve BYA yöntemlerinin $\beta=0$ ve uniform dağılım için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminlerinin yanlılıklarının örneklem büyüklüğüne göre değişimi.....	44
Şekil 4.4. CRA ve BYA yöntemlerinin $\beta=0$ ve binom dağılım için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminlerinin yanlılıklarının örneklem büyüklüğüne göre değişimi.....	44
Şekil 4.5. CRA ve BYA yöntemlerinin $\beta=1$ ve uniform dağılım için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminlerinin örneklem büyüklüğüne göre değişimi.....	45
Şekil 4.6. CRA ve BYA yöntemlerinin $\beta=1$ ve binom dağılım için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminlerinin örneklem büyüklüğüne göre değişimi.....	45
Şekil 4.7. CRA ve BYA yöntemlerinin $\beta=1$ ve uniform dağılım için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminlerinin yanlılıklarının örneklem büyüklüğüne göre değişimi.....	45
Şekil 4.8. CRA ve BYA yöntemlerinin $\beta=1$ ve binom dağılım için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminlerinin yanlılıklarının örneklem büyüklüğüne göre değişimi.....	46

Şekil 4.9. CRA ve BYA yöntemlerinin $\beta=ln(2)$ ve uniform dağılım için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminlerinin örneklem büyüklüğüne göre değişimi.....	46
Şekil 4.10. CRA ve BYA yöntemlerinin $\beta=ln(2)$ ve binom dağılım için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminlerinin örneklem büyüklüğüne göre değişimi.....	47
Şekil 4.11. CRA ve BYA yöntemlerinin $\beta=ln(2)$ ve uniform dağılım için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminlerinin yanlılıklarının örneklem büyüklüğüne göre değişimi.....	47
Şekil 4.12. CRA ve BYA yöntemlerinin $\beta=ln(2)$ ve binom dağılım için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminlerinin yanlılıklarının örneklem büyüklüğüne göre değişimi.....	47
Şekil 4.13. Uniform dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=30$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri ve standart hatalarının grafiksel gösterimi.....	51
Şekil 4.14. Binom dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=30$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri ve standart hatalarının grafiksel gösterimi.....	52
Şekil 4.15. Uniform dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=50$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri ve standart hatalarının grafiksel gösterimi.....	54
Şekil 4.16. Binom dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=50$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri ve standart hatalarının grafiksel gösterimi.....	55
Şekil 4.17. Uniform dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=100$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo	

simülasyonundan elde edilen parametre tahminleri ve standart hatalarının grafiksel gösterimi	58
Şekil 4.18. Binom dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=100$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri ve standart hatalarının grafiksel gösterimi	59
Şekil 4.19. Uniform dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=250$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri ve standart hatalarının grafiksel gösterimi	62
Şekil 4.20. Binom dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=250$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri ve standart hatalarının grafiksel gösterimi	63
Şekil 4.21. Uniform dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=500$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri ve standart hatalarının grafiksel gösterimi	66
Şekil 4.22. Binom dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=500$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri ve standart hatalarının grafiksel gösterimi	67
Şekil 4.23. Uniform dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=30, 50, 100, 250$ ve 500 için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri ve standart hatalarının grafiksel gösterimi	69
Şekil 4.24. Binom dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=30, 50, 100, 250$ ve 500 için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri ve standart hatalarının grafiksel gösterimi	70

Şekil 4.25. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin CRA, BYA-I ve BYA-II'nin parametre tahminleri ve standart hatalarının gösterimi	85
Şekil 4.26. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin aşamalı CRA, BYA-I ve BYA-II'nin parametre tahminleri ve standart hatalarının gösterimi	86
Şekil 4.27. Aşamalı CRA, BYA-I ve BYA-II için meme kanserli hastaların yinelemesiz yaşam sürelerine göre yaşam fonksiyonları grafiği	87

SİMGE VE KISALTMALAR DİZİNİ

n	:	Birim sayısı
p	:	Bağımsız değişken sayısı
t	:	Yaşam süresi
\mathbf{x}	:	Bağımsız değişkenler vektörü
$\boldsymbol{\beta}$:	Parametre vektörü
$\hat{\boldsymbol{\beta}}$:	Parametre tahmini vektörü
$\bar{\boldsymbol{\beta}}$:	Simülasyondan elde edilen $\hat{\boldsymbol{\beta}}$ parametrelerinin ortalaması
$\bar{\boldsymbol{\beta}}_p$:	Prior dağılım ortalaması ya da prior bilgi parametresi
σ_p^2	:	Prior dağılım varyansı
$h_0(t)$:	$\mathbf{x}=\mathbf{0}$ olduğunda temel hazard fonksiyonu
$h(t, \mathbf{x})$:	Cox Regresyon modeli
$S_0(t)$:	Temel yaşam fonksiyonu
$S(t, \mathbf{x})$:	Yaşam fonksiyonu
$\hat{S}_0(t)$:	Temel yaşam fonksiyonu tahmini
$\hat{S}(t, \mathbf{x})$:	Yaşam fonksiyonunun tahmini
$\hat{\Lambda}_0(t)$:	Yığılımlı hazard oranı
$\mathbf{R}(t_{(i)})$:	i . yaşam süresinde ($t_{(i)}$ anındaki) riskteki birimlerin risk seti
$L(\boldsymbol{\beta})$:	Kısmi olabilirlik fonksiyonu
$l(\boldsymbol{\beta})$:	Logaritmik kısmi olabilirlik fonksiyonu
$\hat{\mathbf{V}}(\hat{\boldsymbol{\beta}})$:	$\hat{\boldsymbol{\beta}}$ 'nin varyans kovaryans matrisi
$\mathbf{I}(\boldsymbol{\beta})$:	Gözlenen bilgi matrisi
$\hat{\mathbf{V}}(\hat{\boldsymbol{\beta}}_j)$:	j . parametre tahmininin varyansı ($j=1,2,\dots,p$)
$n_{(i)}$:	$t_{(i)}$ yaşam süreli birim sayısı
$\mathbf{X}_{(i)+}$:	$v_{(i)}$ birimleri için bağımsız değişken değerleri toplamı
G	:	Kısmi olabilirlik oran test istatistiği
W_j	:	j . parametre tahmini için Wald istatistiği
S	:	Skor test istatistiği

$g(t, x)$:	Hazard fonksiyonunun logaritması
$\hat{\sigma}_{\hat{\beta}}$:	$\hat{\beta}$ parametre tahmininin standart hatası
$\pi(\beta)$:	Prior olasılık
$p(x \beta)$ ya da	:	β parametresine bağlı olarak x 'in dağılımı ya da β parametrelili
$L(\beta x) = L(\beta)$:	gözlenen verinin olabilirliği
$p(\beta x)$:	Posterior olasılık
$p(x)$:	Posterior dağılımın normalleştirme sabiti
M	:	İterasyon sayısı ($m=1,2,\dots,M$)
$\hat{\rho}_{(h)}$:	h gecikme zamanı için örneklem otokorelasyonu tahmini
$\hat{\gamma}_{(h)}$:	h gecikme zamanı için örneklem otokovaryans fonksiyonu
z_n	:	Geweke test istatistiği
$\bar{\beta}_1$ ve $\bar{\beta}_2$:	İki alt kısım için β parametresinin posterior ortalamaları
V_1 ve V_2	:	İki alt kısma ayrılmış MCMC zinciri için varyanslar
V_B	:	Zincirler arası varyans
V_W	:	Zincir içi varyans
$\widehat{Var}(\beta x) = \hat{V}$:	Posterior marjinal varyans tahmini
\hat{V}/V_W	:	Gelman-Rubin istatistiği
$E(\beta_j x) \approx \bar{\beta}_j$:	Posterior dağılımın ortalaması
$Var(\beta_j x) \approx s_j^2$:	Posterior dağılımın varyansı
$\overline{D(\beta)}$:	Sapmanın posterior ortalaması
$D(\bar{\beta})$:	$\bar{\beta}$ için ölçülen sapma
N_e	:	Tamamlanmış (uncensored) gözlem sayısı
ABK	:	Akaike Bilgi Kriteri
BBK	:	Bayesgil Bilgi Kriteri
BYA	:	Bayesgil Yaşam Analizi
BYA-I	:	Açıklayıcı Prior Bilgili Bayesgil Yaşam Analizi
BYA-II	:	Açıklayıcı Olmayan Prior Bilgili Bayesgil Yaşam Analizi
CRA	:	Cox Regresyon Analizi
HO	:	Hazard Oranı
MCMC	:	Markov Zinciri Monte Carlo

- SBK : Sapma Bilgi Kriteri
- Pre Menopoz : Adet düzensizliklerinin ve diđer yetmezlik başladığı zamandan son adet kanamasına kadar geçen süre
- Peri Menopoz : Gerçek menopoz, son adet kanamasının başlaması ile başlar, bir yıl sürer
- Post Menopoz : Son adet kanamasından 1 yıl sonra başlar ve 65 yaşına kadar sürer

1. GİRİŞ VE AMAÇ

Yaşam analizleri, yaşam süresine etki eden risk faktörlerinin belirlenmesi açısından önem taşımaktadır. Özellikle yaşam süresine ilişkin çalışmalarda, Cox Regresyon Analizi (CRA) yaygın olarak kullanılmasına rağmen, son yıllarda Bayesgil yaklaşımlara olan ilgi artmış ve CRA'ya alternatif olarak Bayesgil Yaşam Analizi (BYA) geliştirilmiştir.

Bayesgil yaklaşımlar, teorik açıdan oldukça karmaşık oldukları için geçmişte sağlık alanıyla ilgili çalışmalarda yoğun olarak kullanılamamaktaydı. Ancak günümüzde gelişen bilgisayar teknolojisiyle birlikte her alanda olduğu gibi sağlık alanında da Bayesgil yaklaşımların kullanımına ilgi artmaktadır. Bunun temel nedenleri; Bayesgil analizlere ilişkin hesaplamaların kolaylaşması, klasik istatistik metotlardaki kısıtlamaların çokluğu ve tahmin edilecek parametreler hakkındaki var olan bilgilerin (daha önce yapılan çalışmaların sonuçlarının) ve/veya uzman görüşlerinin veri analizine katılmasıdır.

CRA, bilinmeyen parametrelerin sabit olduğunu varsaymasına rağmen; BYA, parametreleri tesadüfi değişkenler, olasılıkları ise “kanaat derecesi (degrees of belief)” olarak tanımlayan alternatif bir yaklaşım önermektedir [22, 40]. Bayesgil yöntemler ve klasik yöntemler arasındaki en büyük fark, Bayesgil yaklaşımların araştırmacının subjektif kanaatlerini yansıtmasıdır [6, 22, 40].

BYA, subjektif düşüncenin temel taşı olarak bahsedilen Bayes teoremine dayanarak geliştirilmiştir. “Bayesgil” kelimesi de, parametre tahminleri için çıkarsama işlemlerinde Thomas Bayes’in teoreminin en önemli rol oynayan unsur olmasından kaynaklanmaktadır [22, 40].

Bayesgil yaklaşımlar, karmaşık veriyi modellemede açıklayıcı olmayan prior bilgiye başvurma gibi esnekliği nedeniyle klasik yöntemlere göre avantajlıdır [22, 47]. Dolayısıyla prior bilginin elde edilmesi, Bayesgil çıkarsamada önemli rol oynar. Bayesgil analiz; önceki bilgiler, geçmiş deneyimler ya da harici bir kanıt gibi bağımsız kaynaklar (açıklayıcı prior bilgiler) ışığında gözlenen verinin subjektif yorumuna

dayanır ve dolayısıyla elde edilen yeni bilginin bileşimine dayanarak gerçekleştirilir [6, 47].

Son yıllarda araştırmacılar, Bayesgil yaklaşımın yaşam analizinde kullanılmasına yönelik çalışmalar yapmışlardır. Calle ve ark. (2006), BYA'yı yiyeceklerin raf ömrünü değerlendirmek için kullanmışlardır. Yağsız ve tam yağlı yoğurt ürünlerinin raf ömrüyle ilgili uygulama çalışmasında, açıklayıcı olan ve olmayan prior bilgiye dayanarak BYA uygulaması yapmışlardır [4]. Wong ve ark. (2005), okul öncesi Çinli çocuklarda aktif diş çürümesini durdurmada silver diamin florid ve sodyum florid parlatıcısının etkinliğini araştırmak için klinik bir çalışmadan elde edilen tamamlanmamış veriyi Bayesgil yaklaşımla incelemişlerdir [47]. Yin ve Ibrahim (2006), BYA'yı simülasyon çalışması ve cilt kanseriyle ilgili gerçek bir veri seti üzerinde kullanmışlardır [48].

Bu çalışmada,

- Bayesgil yaklaşımın yaşam analizinde kullanımını göstermek,
- CRA ve BYA yöntemlerinin değişik algoritmalara göre türetilen veri setlerinde parametre tahminleri, yanlışlıklar ve standart hatalar kullanılarak performanslarını karşılaştırmak,
 - CRA ve BYA yöntemlerinin farklı gözlem sayısı, farklı değişken yapısı ve farklı parametre değerlerine sahip türetilmiş veri setlerinde performanslarını belirlemek,
 - Farklı prior bilgiler kullanılarak BYA yönteminin performansını Monte Carlo simülasyon yöntemi ile incelemek,
 - Hangi koşullarda hangi yöntemin diğer yöntemlere göre performansının daha iyi olduğunu belirlemek,
- Gerçek veri setinde CRA ve BYA yöntemlerinin uygulamasını göstermek, gerçek veri setlerinde bilimsel çıkarımlar yapmak ve meme kanserli hastaların yinelemesiz yaşam sürelerini etkileyen risk faktörlerine ilişkin analiz sonuçlarını ortaya koymak

amaçlandı.

2. GENEL BİLGİLER

Yaşam analizi, belirli bir hastalığa maruz kalan bir bireyin uygulanan bir girişimden sonra daha ne kadar yaşayabileceğini ya da hastalığın ne kadar sürede tekrarlanabileceğini tahmin etmek, tedavi tiplerinin ve diğer faktörlerin yaşam süresine etkilerini incelemek amacıyla geliştirilmiş yöntemler ailesidir.

Yaşamsal verilerin analizinde Yaşam Tablosu (Life Table, LT), Kaplan-Meier (KM), CRA ve BYA gibi yöntemler kullanılmaktadır. Bu yöntemler içinde yaşamsal verilerin karmaşık risk faktörlerini içermesi, oransallık koşulları, incelenen olayla ilgili prior olasılıkların kullanılması gibi çok karmaşık özelliklerin dikkate alınması gerektiğinde CRA ve BYA analizlerinin diğer yöntemlere göre daha etkin sonuçlar verdiği ileri sürülmüştür. Karmaşık veri yapıları ve değişik koşulların dikkate alınması gerektiği durumlarda CRA ve BYA yöntemleri daha sık kullanılan yaşam analizi yöntemleridir.

2.1. Cox Regresyon (Orantısal Hazard Regresyon) Analizi

CRA, yaşamsal verilerin nedensellik analizinde yaygın olarak kullanılan bir yöntemdir. CRA'nın doğrusal regresyon analizinden en önemli farkı, bağımlı değişkenin normal dağılım göstermemesi ve gözlemlerin izlem süresi içinde kesin ölüm ya da yaşama biçiminde olmamasıdır [31].

CRA'nın teorik yapısına geçmeden önce bazı tanımlamalara yer verilmesi gerekmektedir:

- Yaşam süresi (izlem süresi): Bir bireyin belirli bir girişime ya da etkene maruz kaldıktan sonra iyileşmesine, hastalığının tekrarlamasına ya da ölümüne kadar geçen süreye yaşam süresi denir. Yaşam analizlerinde yaşam süresi, bağımlı değişken olarak tanımlanır. Yaşam süresi genellikle Üstel, Weibull ya da Gompertz dağılır.

- Bağımsız değişkenler (prognostik faktörler): Yaşam süresi üzerinde etkide bulunan faktör değişkenler olarak tanımlanır. Bağımsız değişkenlerin dağılımı konusunda herhangi bir kısıtlama yoktur.

Yaşam süresine ilişkin veri setlerinde en önemli nokta, diğer istatistiksel analiz yöntemlerinde ele alınan veriler gibi tamamlanmış olmamasıdır. Yaşam süresine ilişkin veri setlerinde izlem süresi içinde bazı birimlerin ölmesi ya da yaşıyor olması, çeşitli nedenlerle kaybolması (örneğin izlenen hastalar takip edilemedikleri ya da başka bölgelere nakil oldukları için ölmüş ya da yaşıyor olabilir), izlem dışı bırakılması (örneğin izlenen hastalar incelenen neden dışında başka bir nedenden ölmüş olabilir), araştırma süresi sona erdiği için izlendikleri halde hala yaşayan hastaların olması gibi durumlarla karşılaşılır. Böylece verilerin bazıları diğer nedenlerden birimlerin araştırma dışı bırakıldıkları zaman olarak belirlenirken, diğerleri de kesin ölüm zamanıyla ilgilidir. Bu nedenle izlem süreleri, tamamlanmış ve tamamlanmamış olarak tanımlanır [21, 24, 31, 41]:

- Tamamlanmamış veri (censored, incomplete): Araştırma sonlandırıldığı anda takip edilen tüm bireyler ölmeyebilir, araştırma sürerken takip edilen bireylerden bazıları araştırmadan çıkarılabilir ya da incelenen hastalık dışında başka bir nedenden dolayı ölüm gözlenebilir. Bu tür veriler, tamamlanmamış veri olarak tanımlanır.

- Tamamlanmış veri (uncensored, complete): İncelenen olayla ilgili ölüm anına ilişkin veriler, tamamlanmış veri olarak tanımlanır.

CRA, tamamlanmış izlem verilerinde yaşam süresi ile bağımsız değişkenler arasındaki neden-sonuç bağıntısını ortaya koyan bir yaşam analizi yöntemidir [24, 31, 46].

2.1.1. Cox Regresyon Modeli

CRA modeli, hazard modeline dayanarak yazılır. $\mathbf{x} = (x_1, x_2, \dots, x_p)$ bağımsız değişkenler vektörünü göstermek üzere CRA modeli,

$$h(t, \mathbf{x}) = h_0(t)e^{\sum_{j=1}^p \beta_j x_j} \quad (2.1)$$

eşitliğiyle gösterilir [2, 12, 21, 24, 29, 31, 45]. Bu model, orantısal hazard model olarak da bilinir. Eşitlik (2.1), matris biçiminde aşağıdaki gibi gösterilir:

$$h(t, \mathbf{x}) = h_0(t)e^{\beta' \mathbf{x}} \quad (2.2)$$

Bu model, bir birimin bağımsız değişkenler setine dayanarak t anındaki hazard olasılığını ifade eder. Bir başka ifadeyle, bağımsız değişkenler setine dayanarak bir birimin hazard olasılığı modellenir [24, 29].

Eşitlik (2.2)'de $\beta' = (\beta_1, \beta_2, \dots, \beta_p)$ regresyon katsayıları vektörü, $h_0(t)$ $\mathbf{x}=\mathbf{0}$ olduğunda temel hazard fonksiyonu ve t yaşam süresidir. Bu modelde $h_0(t)$, t'nin bir fonksiyonu olup \mathbf{x} vektöründen bağımsızdır ve $e^{\beta' \mathbf{x}}$, yaşam süresinden bağımsızdır.

Temel hazard fonksiyonu $h_0(t)$, bilinmeyen bir fonksiyondur. Bu nedenle, CRA modeli yarı parametrik bir modeldir. CRA modelinin kullanımının popüler olması yarı parametrik olmasından kaynaklanmaktadır [21, 24]. Ayrıca modelde, temel hazard fonksiyonu belirlenememesine rağmen farklı veri yapıları için optimum parametre tahminleri ve hazard oranları vermesi nedeniyle tercih edilmektedir.

CRA modelinin yaşam fonksiyonu,

$$S(t, \mathbf{x}) = [S_0(t)]^{e^{\sum_{j=1}^p \beta_j x_j}} \quad \text{ya da} \quad S(t, \mathbf{x}) = [S_0(t)]^{e^{\beta' \mathbf{x}}} \quad (2.3)$$

eşitliğiyle gösterilir. Eşitlik (2.3)'de $S_0(t)$, temel yaşam fonksiyonudur ve daima 0 ile 1 arasında değer alır [21, 24, 27, 46].

CRA modelinin:

1. Bağımsız değişkenlerin hazard fonksiyonu üzerine etkilerinin logaritmik lineer olması,
2. Bağımsız değişkenlerin logaritmik lineer fonksiyonu ile hazard fonksiyonu arasında çarpımsal bir ilişki olması

varsayımları dışında herhangi bir varsayımı yoktur [21, 24, 31, 44].

2.1.2. En Çok Olabilirlik Parametre Tahmin Yöntemi

CRA, parametrelerin tahmin edilmesinde en çok olabilirlik yöntemini kullanır. En çok olabilirlik yöntemi, bağımsız değişkenler hakkında herhangi bir kısıtlama gerektirmeyen bir tahmin yöntemidir. Bu yöntem, veri setinden elde edilen olasılığı maksimum yapan bilinmeyen parametre değerlerini verir [21, 25]. Bu yöntemin uygulanması için öncelikle olabilirlik fonksiyonu olarak isimlendirilen bir fonksiyon elde edilir. En çok olabilirlik tahminçileri, bu fonksiyonu maksimum yapan parametre tahminlerini belirler. Böylece gözlenen veriyle en iyi örtüşen parametre tahminleri elde edilir [21, 25].

CRA modelinin parametre tahmininde kullanılan olabilirlik fonksiyonu, kısmi olabilirlik fonksiyonu olarak isimlendirilir. Kısmi olabilirlik kavramı, olabilirlik formülünde olasılıkların sadece tamamlanmış olan birimler için ele alınmasından kaynaklanmaktadır [24].

Gözlenen n birimlik veri setinde, tamamlanmış k tane farklı yaşam süresi olsun. k farklı yaşam süresi, $\mathbf{x}_{(1)}, \mathbf{x}_{(2)}, \dots, \mathbf{x}_{(k)}$ bağımsız değişkenleriyle ilişkilendirilerek $t_{(1)}, t_{(2)}, \dots, t_{(k)}$ biçiminde sıralansın. Bunun yanı sıra $\mathbf{R}(t_{(i)})$, i . yaşam süresindeki birimlerin risk seti, yani $t_{(i)}$ anında gözlenen $\mathbf{x}_{(i)}$ bağımsız değişken vektörü olarak tanımlansın. $\mathbf{R}(t_{(i)})$ risk setinde i 'inci yaşam süresindeki riskler setinin genel risk içindeki oranı,

$$\frac{\exp\left(\sum_{j=1}^p \beta_j x_{j(i)}\right)}{\sum_{l \in \mathbf{R}(t_{(i)})} \exp\left(\sum_{j=1}^p \beta_j x_{jl}\right)} = \frac{\exp(\boldsymbol{\beta}' \mathbf{x}_{(i)})}{\sum_{l \in \mathbf{R}(t_{(i)})} \exp(\boldsymbol{\beta}' \mathbf{x}_l)} \quad (2.4)$$

eşitliğiyle hesaplanır. Bu oran, riskler oranı olarak da isimlendirilir. Her bir ölüme, her faktör katkıda bulunur ve böylece kısmi olabilirlik fonksiyonu,

$$L(\boldsymbol{\beta}) = \prod_{i=1}^k \frac{\exp\left(\sum_{j=1}^p \beta_j x_{j(i)}\right)}{\sum_{l \in \mathbf{R}(t_{(i)})} \exp\left(\sum_{j=1}^p \beta_j x_{jl}\right)} = \prod_{i=1}^k \frac{\exp(\boldsymbol{\beta}' \mathbf{x}_{(i)})}{\sum_{l \in \mathbf{R}(t_{(i)})} \exp(\boldsymbol{\beta}' \mathbf{x}_l)} \quad (2.5)$$

eşitliğiyle gösterilirken logaritmik kısmi olabilirlik fonksiyonu,

$$\begin{aligned} l(\boldsymbol{\beta}) &= \log L(\boldsymbol{\beta}) = \sum_{i=1}^k \sum_{j=1}^p \beta_j x_{ji} - \sum_{i=1}^k \log \left[\sum_{l \in R(t_{(i)})} \exp(\sum_{j=1}^p \beta_j x_{jl}) \right] \\ &= \sum_{i=1}^k \left\{ \boldsymbol{\beta}' \mathbf{X}_{(i)} - \log \left[\sum_{l \in R(t_{(i)})} \exp(\sum_{j=1}^p \boldsymbol{\beta}' \mathbf{X}_l) \right] \right\} \end{aligned} \quad (2.6)$$

eşitliğiyle gösterilir. Burada amaç, logaritmik olabilirlik fonksiyonunu en büyük yapan $\hat{\boldsymbol{\beta}}$ parametre tahminlerinin bulunmasıdır. $l(\boldsymbol{\beta})$ 'nın $\boldsymbol{\beta}$ 'ya göre birinci dereceden türevinin sifıra eşitlenmesiyle,

$$\frac{\partial l(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = 0$$

ya da

$$\frac{\partial l(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \sum_{i=1}^k [x_{u(i)} - A_{ui}(\boldsymbol{\beta})] = 0 \quad u=1,2,\dots,p \quad (2.7)$$

olabilirlik eşitlikleri elde edilir. Burada,

$$A_{ui}(\boldsymbol{\beta}) = \frac{\sum_{l \in R(t_{(i)})} x_{ul} \exp(\sum_{j=1}^p \beta_j x_{jl})}{\sum_{l \in R(t_{(i)})} \exp(\sum_{j=1}^p \beta_j x_{jl})} = \frac{\sum_{l \in R(t_{(i)})} x_{ul} \exp(\boldsymbol{\beta}' \mathbf{X}_l)}{\sum_{l \in R(t_{(i)})} \exp(\boldsymbol{\beta}' \mathbf{X}_l)} \quad (2.8)$$

olarak tanımlanır. Bu eşitliklerin çözümünden $\hat{\boldsymbol{\beta}}$ tahmini elde edilir. Olabilirlik eşitliklerinin çözümlenerek parametre tahminlerinin belirlenmesinde iteratif yol izleyen Newton-Raphson algoritması kullanılır. Burada elde edilen kısmi logaritmik olabilirlik fonksiyonu, veri setinde benzer süre gözlemleri olmadığı varsayımı altında parametre değerlerini tahmin etmektedir [3, 8, 27, 45].

Newton-Raphson yöntemi, eşitlik (2.7) kullanılarak olabilirliği maksimum yapan parametre tahminlerini bulmak için başlangıç değeri olarak $\hat{\boldsymbol{\beta}}^0 = 0$ değerini atar. Algoritmada, m iterasyonundan m+1'inci iterasyona geçişte,

$$\hat{\boldsymbol{\beta}}^{m+1} = \hat{\boldsymbol{\beta}}^m - \left[\frac{\partial^2 l(\hat{\boldsymbol{\beta}}^m)}{\partial \boldsymbol{\beta}^2} \right]^{-1} \frac{\partial l(\hat{\boldsymbol{\beta}}^m)}{\partial \boldsymbol{\beta}} \quad (2.9)$$

eşitliği kullanılır.

Kısmi olabilirlik tahmin teorisine dayanarak parametre tahminlerinin, varyans ve kovaryansları da tahmin edilir. Varyans ve kovaryans tahminleri logaritmik kısmi olabilirlik fonksiyonunun β 'ya göre ikinci dereceden kısmi türevinin alınmasıyla elde edilir. $\hat{\beta}$ 'nin varyans kovaryans matrisi,

$$\hat{V}(\hat{\beta}) = \left[-\frac{\partial^2 l(\beta)}{\partial \beta \partial \beta'} \right]^{-1} \quad (2.10)$$

eşitliğiyle elde edilir. Burada $I(\beta) = -\partial^2 l(\beta) / \partial \beta \partial \beta'$, gözlenen bilgi matrisi olarak isimlendirilir. Böylece eşitlik (2.10),

$$\hat{V}(\hat{\beta}) = I(\beta)^{-1} \quad (2.11)$$

olarak yazılabilir.

CRA modelinde regresyon katsayılarının güven aralığı,

$$\hat{\beta}_j \pm z_{\alpha/2} \sqrt{\hat{V}(\hat{\beta}_j)} \quad (j=1,2,\dots,p) \quad (2.12)$$

biçiminde elde edilir.

Veri setinde benzer süre gözlemleri olduğunda logaritmik kısmi olabilirlik fonksiyonunun maksimum yapılmasında, farklı yaklaşımlar önerilmiştir. Bunlar, Breslow (1974) (ya da Breslow-Peto) yaklaşımı, Efron (1977) yaklaşımı ve Kalbfleisch ve Prentice (1980) tarafından tanımlanan Exact yaklaşımdır. Efron ve Breslow yaklaşımlarının hesaplanması, Exact yaklaşımdan daha kolaydır. Efron yaklaşımının, Exact yaklaşıma göre biraz daha iyi olduğundan bahsedilmektedir. Ayrıca birçok uygulamada, Breslow ve Efron yaklaşımlarından elde edilen parametre tahminleri arasında oldukça küçük ve önemsiz farklılıklar bulunmuştur [21, 24, 27]. Bu nedenle çalışmamızda, bunlar arasında daha yaygın bir yaklaşım olan Breslow kullanılacaktır.

Breslow yaklaşımıyla kısmi olabilirlik fonksiyonu,

$$L(\beta) = \prod_{i=1}^k \frac{\exp(x'_{(i)}\beta)}{[\sum_{l \in R(t_{(i)})} \exp(x'_l \beta)]^{n_{(i)}}} \quad (2.13)$$

eşitliğiyle hesaplanır. Burada $n_{(i)}$, yaşam süresi $t_{(i)}$ olan birimlerin sayısını gösterir ve $X_{(i)+}$, $n_{(i)}$ birimleri için bağımsız değişken değerlerinin toplamına eşittir ($X_{(i)+} = \sum_{l \in D(t_{(i)})} X_l$; burada $D(t_{(i)})$, yaşam süresi $t_{(i)}$ olan birimleri gösterir.) [21].

2.1.3. Katsayıların Önemliliğinin Test Edilmesi

CRA'da katsayılar tahmin edildikten sonra, katsayıların önemliliği için $H_0: \beta=0$ hipotezi kısmi olabilirlik oran, Wald ve skor testleri ile test edilir [21, 24, 27, 31].

1. Kısmi Olabilirlik Oran Testi

Kısmi olabilirlik oran test istatistiği, bağımsız değişken içeren modelin logaritmik kısmi olabilirlik değeri ve bağımsız değişken içermeyen modelin hesaplanan logaritmik kısmi olabilirlik değeri arasındaki farkın iki katı olarak hesaplanır:

$$G = 2\{l(\hat{\beta}) - l(0)\} \quad (2.14)$$

Burada $l(0) = -\sum_{i=1}^k \ln(n_{(i)})$, $n_{(i)}$ ise gözlenen $t_{(i)}$ yaşam süreli risk setindeki birim sayısını ifade eder. Eşitlik (2.14), modele bağımsız değişken/değişkenlerin dahil edilmesinden kaynaklanan G 'deki değişimi gösterir. Bir ya da daha fazla bağımsız değişkenin anlamlılığını test etmek için eşitlik (2.14)'den bağımsız değişkenli ve bağımsız değişkensiz G test istatistiklerinin değerleri karşılaştırılır.

Eşitlik (2.14)'de yer alan $\hat{\beta}$ vektöründeki parametre sayısı p ise kısmi olabilirlik oran test istatistiği, p serbestlik dereceli ki-kare dağılır [21, 24, 27, 31].

2. Wald Testi

Wald test istatistiği, katsayı tahmininin standart hatasına oranlanmasıyla elde edilir.

$$W_j = \frac{\hat{\beta}_j}{\hat{\sigma}_{\hat{\beta}_j}} \quad (j=1,2,\dots,p) \quad (2.15)$$

Wald istatistiđi, standart normal dađılım gosterir. Aynı zamanda Wald istatistiđinin karesi, 1 serbestlik dereceli ki-kare dađılımı gosterir [21, 24, 27, 31].

3. Skor Testi

Skor testi, logaritmik olabilirlik istatistiklerinden yararlanılarak hesaplanır. Skor test istatistiđi, gozlenen bilgi matrisinin karekokunun, logaritmik kısmi olabilirliđin turevине oranlanmasıyla elde edilir:

$$S = \frac{\partial l(\beta)/\partial \beta}{\sqrt{I(\beta)}} \quad (2.16)$$

Skor istatistiđi, standart normal dađılım gosterir. Ayrıca bu istatistiđin karesi, 1 serbestlik dereceli ki-kare dađılımı gosterir [21, 24, 27, 31].

2.1.4. Regresyon Katsayılarının Yorumlanması

Butun regresyon modellerinde bir bađımsız deđiřken iin tahmin edilen katsayı, bađımsız deđiřkendeki bir birimlik deđiřimin bađımlı deđiřkende ka birimlik deđiřime yol aıtıđını gosterir. Katsayıların dođru olarak yorumlanması iin bađımlı ve bađımsız deđiřkenler arasındaki fonksiyonel iliřkiyi tanımlamak gerekir. CRA'da kullanılan hazard fonksiyonu, yařam suresi uzerine bir ya da daha fazla bađımsız deđiřkenin etkisini arařtırmak iin kullanılır. Katsayıları yorumlama surecinde oncelikle, hazard fonksiyonunu dođrusal hale getirmek gerekmektedir. Genelleřtirilmiř dođrusal modeller ailesinde (lineer, lojistik, Poisson ve diđer regresyon modellerinde), dođrusallařtırma donuřumu bir link fonksiyonuyla sađlanmaktadır. Aynı durum CRA modeli iin de gecerlidir [21].

İzleyen aıklamalarda, tek bađımsız deđiřken ieren CRA modeli ele alınacaktır. Tek bađımsız deđiřkenli model,

$$h(t, x) = h_0(t)e^{x\beta} \quad (2.17)$$

eşitliğiyle gösterilir. CRA modeli için link fonksiyonu, logaritmik dönüşümdür. Hazard fonksiyonunun logaritması, $g(t, x) = \ln [h(t, x)]$ olarak tanımlanır. Böylece logaritmik hazard fonksiyonu,

$$g(t, x) = \ln[h_0(t)] + x\beta \quad (2.18)$$

eşitliğiyle tanımlanır. $x = a$ değerinden $x = b$ değerine geçişteki değişimden kaynaklanan logaritmik hazard fonksiyonundaki fark,

$$\begin{aligned} [g(t, x = a) - g(t, x = b)] &= \{\ln[h_0(t)] + a\beta\} - \{\ln[h_0(t)] + b\beta\} \\ &= a\beta - b\beta = (a - b)\beta \end{aligned} \quad (2.19)$$

eşitliğiyle gösterilir. Eşitlik (2.19)'un en son halinde temel hazard fonksiyonları birbirini götürür ve böylece logaritmik hazardların farkı, zamana bağımlı olmayan bir yapıya dönüşür [21].

Logaritmik hazard, bağımsız değişkendeki değişimin etkisini değerlendirmek için uygun bir yaklaşım olmasına rağmen yorumlanması zordur. Bu nedenle CRA modelinde etkinin ölçüsü olarak katsayıların yorumlanmasında hazard oranı (HO) kullanılır [21, 24, 29]. HO, Lojistik Regresyon Analizi'nde kullanılan odds oranı gibi yaşam analizinde sonuçların yorumlanması ve açıklanmasında önemli bir rol oynar ve

$$\begin{aligned} HO(t, a, b, \beta) &= \exp[g(t, x = a) - g(t, x = b)] \\ &= \frac{h(t, a, \beta)}{h(t, b, \beta)} \\ &= e^{(a-b)\beta} \end{aligned} \quad (2.20)$$

eşitliğiyle elde edilir.

CRA modelinde değişkenlerin ölçüm düzeyi konusunda herhangi bir kısıt olmadığı için, aynı modelde karışık ölçekli bağımsız değişkenler yer alabilir. Katsayıların yorumlanmasında, hazard oranlarının regresyon katsayıları yardımıyla hesaplanması bağımsız değişkenin iki kategorili, ikiden fazla kategorili ve sürekli olmasına göre farklılık göstermektedir [21, 25].

1. İki kategorili bağımsız değişken

Bağımsız değişkenin, 0 ve 1 olarak kodlanan iki kategorili durumunu ele alalım. Eşitlik (2.18) dikkate alındığında x bağımsız değişkenini yorumlayabilmek için öncelikle, bağımsız değişkendeki bir birimlik değişim için logaritmik hazarddaki farkın hesaplanması gerekmektedir. Bu durumda,

$$[g(t, x = 1, \beta) - g(t, x = 0, \beta)] = (1 - 0)\beta = \beta \quad (2.21)$$

elde edilir. Eşitlik (2.21)'de, logaritmik hazard fonksiyonundaki farkın değeri üs olarak alınırsa HO,

$$HO(t, 1, 0, \beta) = e^{\beta} \quad (2.22)$$

olur. HO, odds oranından farklı bir ölçü olmasına rağmen yorumlanması benzerdir. Örneğin bağımsız değişkenin cinsiyet (referans kategori=kadın) olduğunu ve HO'nun 2 olarak elde edildiğini varsayalım. Bu durumda HO "erkekler kadınlardan 2 kat daha fazla ölüm riski taşımaktadır" şeklinde yorumlanır.

HO, kolay yorumlanması açısından CRA modelinde ilgilenilen bir parametredir. Teorik olarak, örneklem büyüklüğü arttıkça \widehat{HO} 'nın dağılımının normal dağılıma yaklaştığı kabul edilmektedir. Fakat yeterli örneklem büyüklüğüne çoğu çalışmada ulaşamamaktadır. Bu nedenle hesaplamalar, çok daha küçük örneklem büyüklüğü için normal dağılıma uyma eğiliminde olan $\ln(\widehat{HO}) = \hat{\beta}$ 'nin örnekleme dağılımına dayanır.

HO için güven aralığı tahmini,

$$\exp[\hat{\beta} \pm z_{1-\alpha/2} \hat{\sigma}_{\hat{\beta}}] \quad (2.23)$$

biçiminde elde edilir.

2. İki den çok kategorili bağımsız değişken

CRA modelinde bazı bağımsız değişkenlerin ikiden fazla kategorili isimsel ölçekli olması durumunda bu değişkenlerin sürekli değişkenler gibi modele dahil edilmesi uygun değildir. Çünkü isimsel ölçekli değişkenlerin farklı kategorilerini göstermek için

kullanılan rakamların sayısal anlamları yoktur. İki kategori yerine bağımsız değişken $K > 2$ kategoriye sahip olduğunda kategorilerin kukla değişkenler kullanılarak yeniden kodlanması gerekmektedir. K kategorili bir bağımsız değişken için $K-1$ tane kukla değişkene ihtiyaç vardır.

Bu kukla değişkenleri kodlamada sıklıkla kullanılan yöntem, referans hücre kodlamasıdır. Bu yöntemde, referans düzeyi olarak bağımsız değişkenin bir kategorisi seçilerek (genellikle birinci ya da sonuncu kategori) diğer kategorilerle karşılaştırılır. Referans kategori dışında kalan diğer kategoriler için HO hesaplanır ve her bir grubun HO'su referans kategoriye göre kıyaslanarak yorumlanır.

Örneğin, modele alınan bağımsız değişkenlerden birisi A, B, C ve D olarak kodlanan bir değişken olsun. Bu durumda, üç kukla değişkene ihtiyaç vardır. Değişken cevabı A olduğunda kodlama stratejilerinin biri, üç kukla değişkeninde (K_1 , K_2 ve K_3) sıfıra eşit olmasıdır; cevap B olduğunda $K_1=1$ $K_2=0$ ve $K_3=0$ 'a eşit olabilir; cevap C olduğunda ise $K_1=0$, $K_2=1$ ve $K_3=0$ (Tablo 2.1) [21, 25]. Ayrıca genellikle üç kukla değişkenin değerinin de sıfıra eşit olduğu A kategorisi, referans kategori olarak kodlanır.

Tablo 2.1. Dört kategorili bir değişkene ilişkin kukla değişkenler

Kategoriler	Kukla Değişkenler		
	K_1	K_2	K_3
A	0	0	0
B	1	0	0
C	0	1	0
D	0	0	1

Tahminlenen HO'lar, bu dört kategorinin yaşam süresini karşılaştırmada kolaylıkla yorumlanır. Örneğimizde, referans kategori olarak A'yı kullanarak her bir kategori için HO'ları hesaplayalım. HO tahminlerini bulmadan önce logaritmik hazard fonksiyonunun elde edilmesi gerekir. Logaritmik temel hazard fonksiyonunu göz ardı

eden (iki temel hazard fonksiyonu birbirine eşit olduğu için) logaritmik hazard fonksiyonu,

$$g(t, x, \beta) = \beta_1 K_1 + \beta_2 K_2 + \beta_3 K_3$$

eşitliğiyle gösterilir. x değişkeninin B kategorisini A kategorisiyle karşılaştıran HO'nun tahmini, eşitlik (2.18)'de görüldüğü gibi logaritmik hazard fonksiyonlarının tahminlerindeki fark hesaplanarak elde edilir:

$$[g(t, x = B, \hat{\beta}) - g(t, x = A, \hat{\beta})] = (\hat{\beta}_1 1 + \hat{\beta}_2 0 + \hat{\beta}_3 0) - (\hat{\beta}_1 0 + \hat{\beta}_2 0 + \hat{\beta}_3 0) = \hat{\beta}_1$$

Elde edilen parametre tahmini ($\hat{\beta}_1$) üs olarak alınır,

$$\widehat{HO}(B, A) = e^{\hat{\beta}_1}$$

eşitliği elde edilir. x değişkeninin C ve D kategorileri için de A referans kategorisine göre HO tahminleri,

$$\widehat{HO}(C, A) = e^{\hat{\beta}_2}$$

$$\widehat{HO}(D, A) = e^{\hat{\beta}_3}$$

eşitlikleriyle elde edilir.

İkiden fazla kategorili sıralı ölçekli değişkenler için de regresyon katsayılarının yorumlanması aynıdır.

3. Sürekli bağımsız değişken

Sürekli değişkenler için hesaplanmış katsayıların kullanımı, sürekli olmayan değişkenlerden biraz farklıdır. Kategoriler girilmediği için kategorik değişkenlere göre sürekli bağımsız değişkenin katsayısının yorumu daha kolaydır.

Çoğunlukla, bağımsız değişkendeki değişimin, yaşam süresi üzerindeki etkisini göstermede "1" değeri uygun değildir. Örneğin yaştaki 1 yıllık artış ya da sistolik kan basıncındaki 1 mmHg'lik artış önemli sayılabilecek bir değişim değildir. Bunun yerine

10 yıllık ya da 10 mmHg'lik bir değişimin daha anlamlı olacağı düşünülebilir. Buna karşın, bağımsız değişkenin değerleri 0 ile 1 arasında değişiyorsa 1 birimlik değişim oldukça büyük olduğundan 0.1'lik değişim daha gerçekçi olacaktır. Bu nedenle, sürekli bağımsız değişken katsayılarının yorumlanması için “c” değişim düzeyi belirlenmelidir [21, 25].

$a=x+c$ ve $b=x$ için eşitlik (2.19) ve (2.20) kullanılarak, sürekli bağımsız değişkende c birimlik bir değişim için logaritmik hazard fonksiyonundaki değişim bulunur. Buna göre logaritmik hazard fonksiyonundaki değişim,

$$\begin{aligned} [g(t, x + c) - g(t, x)] &= \{ \ln[h_0(t)] + (x + c)\beta \} - \{ \ln[h_0(t)] + x\beta \} \\ &= (x + c)\beta - x\beta = c\beta \end{aligned} \quad (2.24)$$

eşitliğiyle elde edilir. Eşitlik (2.24)'e göre, bağımsız değişkende meydana gelecek 1 birimlik değişim logaritmik hazard fonksiyonunda meydana gelecek değişimi verir. Burada HO tahmincisi,

$$\widehat{HO}(c) = e^{c\hat{\beta}} \quad (2.25)$$

eşitliğiyle elde edilir. HO'nun güven aralığı tahmini,

$$\exp[c\hat{\beta} \pm z_{1-\alpha/2}|c|\hat{\sigma}_{\hat{\beta}}] \quad (2.26)$$

biçiminde elde edilir.

2.2. Bayesgil Yaşam Analizi

2.2.1. Bayes Teoremi

Son yıllarda klasik istatistiksel yaşam analizi yöntemlerine alternatif olabilecek farklı yeni bir yaklaşım kullanılmaya başlanmıştır. Bu yaklaşımın temelini, 1763'te Thomas Bayes tarafından ortaya konulan basit Bayes teoremi oluşturmaktadır. Thomas Bayes, kendisine ait olasılık kuramını “Essay towards solving a problem in the doctrine of chances” adlı makalesinde yayınlamıştır. Bu makale, “Philosophical Transactions of

the Royal Society of London” dergisinde de yayınlanmıştır. Bu makale döneminin saygın matematikçileri tarafından ilgiyle karşılanırsa da büyük çapta bir yankı uyandırmamış ve bu konuda geniş araştırmalar yapılamamıştır. Bayes’in yargıları 1781’de Laplace tarafından kabul görmüştür. Aslında Bayes kuramını genelleştirerek daha geniş bir kullanım kazandıran, genelleştirdiği bu kuramı mekanik, tıp gibi alanlarda kullanan Laplace’dır [6, 22, 51].

Bilgisayar donanım ve yazılımlarındaki gelişme, Bayes teoreminin uygulanabilirliğini ortaya çıkarmış ve günümüzde standart bir çıkarım yöntemi olarak kabul görmüş ve istatistiksel paket programlarda yer almaya başlamıştır [6, 22].

Bayes teoremi, koşullu olasılıkların hesaplanmasında kullanılan basit bir matematiksel formüldür. Bayes teoreminden, subjektif düşüncenin temel taşı olarak bahsedilir.

Bir A olayının ortaya çıkmasında ikiden daha fazla olayın ya da faktörün etkisi varsa A olayı meydana geldiğinde r . faktörün gözlenme koşullu olasılığı Bayes teoremine dayanarak hesaplanır.

Bir örnek uzayında olaylar seti B_1, B_2, \dots, B_k olarak tanımlansın. Burada $P(B_r) \neq 0$ ($r=1, \dots, k$) olmak üzere A, örnek uzayında ilgilenilen herhangi bir olay olarak tanımlansın ($P(A) \neq 0$). A olayı meydana geldiğinde herhangi bir B_r olayının olması olasılığı,

$$P(B_r/A) = \frac{P(B_r \text{ ve } A)}{P(A)} = \frac{P(B_r \text{ ve } A)}{\sum_{r=1}^k P(B_r \text{ ve } A)} = \frac{P(A/B_r)P(B_r)}{\sum_{r=1}^k P(A/B_r)P(B_r)} \quad (2.27)$$

eşitliğiyle yazılır [5, 6, 10, 22, 23, 31, 38, 43, 46]. Eşitlik (2.27) genel olarak,

$$\text{Posterior olasılık} = \frac{\text{Koşullu olasılık} * \text{Prior olasılık}}{\text{Marjinal olasılık}} \quad (2.28)$$

olarak da hesaplanır. Bayes teoreminde eşitlik (2.28)’de verilen bazı özel kavramlar (prior ve posterior olasılık) ortaya çıkmaktadır.

Prior dağılım: Prior olasılıklar, prior dağılımdan hesaplanmaktadır. Prior bilginin elde edilmesi, Bayesgil çıkarsamada önemli bir rol oynar. Prior dağılım, bilinmeyen parametre hakkında ön bilgi verir [18, 22].

Posterior dağılım: Posterior olasılıkların hesaplandığı posterior dağılım, Bayesgil analizde bilinmeyen bütün nicelikler (gözlenmeyen parametreler) hakkında bilginin güncel durumunu ortaya koyar. Posterior dağılım, prior dağılım ile olabilirlik fonksiyonundan elde edilir [17].

2.2.2. Bayesgil Çıkarsama

Klasik yöntemlerde olasılıklar objektif ve parametreler sabit olduğu için parametreler hakkında olasılıksal yorumların yapılamaması, bilinmeyen parametrelerin sabit (değişmez) olduğu varsayımından kaynaklanmaktadır. Bu nedenle Bayesgil yöntemler, parametreleri tesadüfi değişkenler olarak ele alarak ve olasılıkları “kanaat derecesi” (bir olayın olasılığı, olayın doğru olduğuna inanma derecesi) olarak tanımlayan alternatif bir yaklaşım sunmaktadır [22, 40].

$p(x|\beta)$ yoğunluğuyla tanımlanan bir istatistiksel model için $x = [x_1, x_2, \dots, x_n]$ verisinden β parametresini tahmin edelim. Bayesgil felsefe, β 'nin tam olarak tahmin edilemediğini düşünür. Bu düşünceye dayanarak, olasılık durumları ve dağılımları yardımıyla parametre hakkındaki belirsizliği ortadan kaldırmaya çalışır. Örneğin parametre hakkındaki belirsizliğin normal dağılımla daha iyi açıklanacağına inanılıyorsa, β parametresinin ortalaması 0 ve varyansı 1 olan normal dağılım gösterdiği söylenebilir.

Bayesgil çıkarsamanın temel esasları aşağıda açıklanmaktadır:

1. β için olasılık dağılımı, “prior dağılım” ya da sadece prior olarak isimlendirilir ve $\pi(\beta)$ olarak gösterilir. Prior dağılım, veriyi incelemenden önceki parametre hakkındaki kanaatleri ya da bilinenleri (örneğin ortalama, yaygınlık, çarpıklık v.s.) ifade eder.

2. x değişkeni için, β parametresine bağlı olarak x 'in dağılımı $p(x|\beta)$ biçiminde gösterilir.

3. Prior dağılım ile veriden elde edilen bilgi birleştirilerek $p(\beta|x)$ posterior dağılımının hesaplanması yoluyla β hakkındaki bilgiler güncelleştirilir.

Son maddede yer alan prior dağılım ile modelin birleştirilmesi Bayes teoremi kullanılarak gerçekleştirilir.

En çok olabilirlik tahmin yöntemi gibi klasik yaklaşımlarda parametre tahmini için çıkarsama yapmak, sadece verinin olabilirliğine dayanmaktadır. Bayesgil modellerde ise, β parametrelili gözlenen verinin olabilirlik fonksiyonu ($L(\beta|x)$ ya da $p(x|\beta)$), $p(\beta|x)$ posterior yoğunlukla ortaya konulmak istenen güncelleştirilmiş bilgi elde etmek için $\pi(\beta)$ prior bilgilerini değiştirmede kullanılır. Bu yoğunluklar arasındaki ilişki, standart olasılık denklemlerinden elde edilir:

$$p(x, \beta) = p(x|\beta)\pi(\beta) = p(\beta|x)p(x) \quad (2.29)$$

ve buradan posterior yoğunluk,

$$p(\beta|x) = \frac{p(x, \beta)}{p(x)} = \frac{p(x|\beta)\pi(\beta)}{p(x)}$$

ya da

$$p(\beta|x) = \frac{L(\beta)\pi(\beta)}{p(x)} \quad (2.30)$$

eşitliğiyle gösterilir. Burada $p(x)$, verinin marjinal olasılığıdır ve posterior dağılımın normalleştirme sabitidir. $p(x)$, prior yoğunlukla olabilirliğin integralinden elde edilir:

$$p(x) = \int p(x|\beta)\pi(\beta)d\beta \quad (2.31)$$

Bu nicelik, Bayesgil model seçimi için bazı yaklaşımlarda önemli rol oynar. Fakat şu anki amaç için sadece normalleştirme sabiti olduğu için integralin herhangi bir değeri, posterior dağılım hakkında ek bir bilgi sağlamaz. Böylece $p(\beta|x)$ eşitliği orantılı biçimiyle,

$$p(\beta|x) \propto L(\beta)\pi(\beta) \quad (2.32)$$

eşitliğiyle gösterilir. Burada $L(\beta)$ gözlenen veriden elde edilen katkıyı gösterirken, $\pi(\beta)$ ise ölçülen prior bilgiden gelen katkıyı gösterir. Ayrıca eşitlik (2.32)'de $p(x)$ bilinmediği için posterior olasılıklar orantılı olarak tahmin edilir. Esasen Bayes teoremi, yeni bilgiyle var olan bilginin nasıl güncelleştirileceğini ifade eder [5-7, 10, 22, 35, 40, 42, 43].

BYA'da da, posterior olasılıklar tahmin edilirken olabilirlik fonksiyonu olarak kısmi olabilirlik fonksiyonu kullanılır [22, 40].

BYA, CRA modelindeki parametreler hakkında parametrenin posterior dağılımına dayanarak çıkarsamada bulunur. Klasik yöntemler parametre tahmini için en çok olabilirlik ya da momentler tekniğini kullanırken, BYA'da

$$E(\beta|x) = \int \beta p(\beta|x) d\beta \quad (2.33)$$

eşitliğiyle elde edilen posterior ortalama kullanılır [6, 22, 40].

2.2.3. Markov Zinciri Monte Carlo Yöntemi

BYA'da analizlerin gerçekleştirilmesi analitik olarak oldukça güç olduğundan posterior tahminleri bulmak için Markov Zinciri Monte Carlo (Markov Chain Monte Carlo, MCMC) simülasyon algoritması kullanılır [6, 11, 12, 19, 22, 28, 35, 40].

MCMC yöntemi, olasılık teorisi üzerine kurulu bir sistemdir. Bu yöntemde amaç, istatistiksel ve matematiksel tekniklerle bir deneyi veya çözülmesi gereken bir olayı rastgele sayıları defalarca kullanarak simülasyonla çözmektir. Bu yöntemin bir probleme uygulanması, problemin rastgele sayıları kullanarak simülasyonla tahmin edilmek istenen parametrenin bu simülasyonlarının sonuçlarına bakılarak yaklaşık

olarak hesaplanması sürecine dayanır. MCMC, basit sayısal integral hesaplama yöntemleri yanında günümüz istatistik teorisinin yoğun hesaplama gerektiren Bayesgil çıkarsama yöntemlerini pratik ve rutin olarak uygulanabilir hale getiren modern bir simülasyon yöntemidir. MCMC'den elde edilen parametre tahmin sonuçları diğer yöntemlerle karşılaştırıldığında riski daha iyi temsil etmesi nedeniyle; mühendislik, eğitimde ölçme ve değerlendirme, askeri savunma teknolojisi, fen ve mühendislik alanında, nükleer teknoloji ve uzay sisteminde, istatistiksel analiz ve sosyoekonomik alanlarda sıklıkla kullanılan bir yöntemdir [10, 20, 28, 36, 37].

MCMC yöntemi, posterior dağılımdan örneklem alan ve ilgilenilen posterior nicelikleri hesaplayan genel bir simülasyon yöntemidir. MCMC yöntemi içerisinde Markov zinciri yöntemi, modern Bayesgil hesaplamalarda kullanılan oldukça iyi bir yöntemdir. Basit Bayesgil modellerde, posterior dağılımların analitik biçimleri tanımlanabilir ve doğrudan çıkarsamalar yapılabilir. Fakat karmaşık modellerde, posterior yoğunlukları doğrudan belirlemek oldukça zordur. MCMC yöntemiyle keyfi bir $p(\beta|x)$ posterior yoğunluğundan örneklem oluşturmak mümkündür ve ilgilenilen niceliğin beklenenlerini yaklaşık olarak tahmin etmek için bu örneklem kullanılır [40].

Bayesgil yöntemler, istatistiksel çıkarsama için bütün çıkarsamaları $p(\beta|x)$ posterior dağılımından elde edilen bir alternatif önermektedir. Bu yaklaşım, simülasyon yöntemlerinin kullanımını içeren ayrıntılı ve karmaşık hesaplamalar gerektirir. Bu hesaplamalarda, posterior dağılımdan örneklem meydana getirilir ve ilgilenilen nicelikleri tahmin etmek için bu örneklem kullanılır. Dolayısıyla MCMC, yüksek boyutlu posterior integralleri değerlendiren örnekleme yöntemlerine dayanır. MCMC yöntemini içeren örnekleme yöntemleri; Gibbs örnekleme, Metropolis-Hastings örnekleme ve diğer hibrid algoritmalarından oluşmaktadır. Bu algoritmalar içinden genellikle, MCMC yöntemlerini kapsayan Gibbs örnekleme kullanılır. Gibbs örnekleme, $p(x)$ normalleştirme sabiti bilinmeksizin $p(\beta|x)$ 'den örneklem almayı sağlayan çok güçlü bir simülasyon algoritmasıdır [6, 7, 11, 14, 22, 36, 40].

MCMC yönteminin temel amacı, her bir örneklemin bir öncekine bağlı olarak çekilmesiyle hedef bir dağılımdan art arda örneklem almaktır. Markov zincirlerinin en

önemli özelliği, sistemin zaman içerisinde bulunabileceği tüm olası durumların listesini oluşturmasıdır. Markov zincirinde; posterior dağılıma yakınsamak amacıyla türetilen β^m , önceki β^{m-1} değerine bağlı olmaktadır [40, 49].

MCMC yönteminin en önemli özelliği, simülasyon algoritması doğru bir şekilde uygulandıysa Markov zinciri, zincirin başlangıç değerine bakmaksızın $p(\beta|x)$ hedef dağılımına yakınsamayı garanti etmektedir. Markov zinciri çok uzun bir simülasyon sürecinde çalıştırılırsa $p(\beta|x)$ 'yi daha yüksek doğrulukla elde edebilir. Genellikle uygulamalarda yüksek boyuttan kaynaklanan problemler olmasına rağmen MCMC yönteminde simülasyon algoritması, çok sayıda parametre içeren ya da aşırı karmaşık modeller için kolaylıkla genişletilebilir ve kullanılabilir [40].

Gibbs Örnekleme:

Daha öncede bahsedildiği gibi Bayesgil yaklaşımlarda, posterior olasılık ya da dağılımların belirlenmesi üzerinde durulmaktadır. Fakat bazı durumlarda özellikle posterior momentlerin hesaplanması için gerekli olan integrallerin analitik olarak çözümleri mümkün olmamakta ya da güç olmaktadır. Bu durumlarda, Markov zinciri türetme ve yakınsaklık özellikleri ile posterior dağılımı elde etme yaklaşımları kullanılmaktadır. Bu yaklaşımlar MCMC başlığı altında toplanmaktadır.

MCMC yaklaşımları kullanılarak sonlu sayıda gözlem değeriyle sonsuz sayıda veri elde etmek mümkündür. Böylece çözümü analitik olarak zor olan bazı problemlerin, benzetim teknikleri ve bilgisayar yazılımları sayesinde hızlı biçimde çözülmesi mümkün olmaktadır [49].

Gibbs örnekleme, Bayesgil hesaplamalarda kullanılan MCMC örnekleme algoritmalarından en yaygın olarak kullanılanıdır. Gibbs örneklemede; \mathbf{x} bağımsız değişkenler vektörü $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$, parametre vektörünün olabilirliği $p(\mathbf{x}|\boldsymbol{\beta})$ ve prior dağılımı $\pi(\boldsymbol{\beta})$ olarak tanımlansın. Markov zincirinin, $p(\boldsymbol{\beta}|\mathbf{x})$ dağılımına yakınsaması için $\pi(\beta_i|\beta_j, i \neq j, \mathbf{x})$ tam posterior koşullu dağılımından türetilen örneklemler kullanılmaktadır. $\pi(\beta_i|\beta_j, i \neq j, \mathbf{x})$ tam posterior koşullu dağılımı, birleşik posterior yoğunluğa orantılıdır ve

$$\pi(\beta_i | \beta_j, i \neq j, \mathbf{x}) \propto p(\mathbf{x} | \boldsymbol{\beta}) \pi(\boldsymbol{\beta}) \quad (2.34)$$

formülüyle gösterilir. Örneğin $\pi(\beta_1)$ 'in bir boyutlu koşullu dağılımı,

$$\pi(\beta_1 | \beta_j = \beta_j^*, 2 \leq j \leq p, \mathbf{x}) = p(\mathbf{x} | (\boldsymbol{\beta} = (\beta_1, \beta_2^*, \dots, \beta_k^*)')) \pi(\boldsymbol{\beta} = (\beta_1, \beta_2^*, \dots, \beta_k^*)')$$

eşitliğiyle hesaplanır.

Gibbs örnekleme aşağıdaki adımlar izlenerek uygulanır [5-7, 10, 22, 36, 38, 40, 43, 48]:

1. M iterasyon sayısı, m her bir iterasyon olmak üzere, m=0 iterasyonunda keyfi başlangıç değeri olarak $\boldsymbol{\beta}^{(0)} = \{\beta_1^{(0)}, \beta_2^{(0)}, \dots, \beta_p^{(0)}\}'$ alınır.
2. $\boldsymbol{\beta}$ parametresinin her bir bileşenini $\boldsymbol{\beta}^{(m+1)} = \{\beta_1^{(m+1)}, \beta_2^{(m+1)}, \dots, \beta_p^{(m+1)}\}'$ olarak meydana getirmek için aşağıdaki yol izlenir:
 - $\beta_1^{(m+1)}$ için $\pi(\beta_1 | \beta_2^{(m)}, \dots, \beta_p^{(m)}, \mathbf{x})$ olasılık dağılımından $\beta_1^{(m+1)}$ çekilir.
 - $\beta_2^{(m+1)}$ için $\pi(\beta_2 | \beta_1^{(m+1)}, \beta_3^{(m)}, \dots, \beta_p^{(m)}, \mathbf{x})$ olasılık dağılımından $\beta_2^{(m+1)}$ çekilir.
 -
 - $\beta_p^{(m+1)}$ için $\pi(\beta_p | \beta_1^{(m+1)}, \beta_2^{(m+1)}, \dots, \beta_{p-1}^{(m+1)}, \mathbf{x})$ olasılık dağılımından $\beta_p^{(m+1)}$ çekilir.
3. m=m+1 alınır ve m < M (m=1,2,...,M) ise 2. adıma gidilir.

Gibbs örnekleme, $\beta^{(m)}$ 'den $\beta^{(m+1)}$ 'e geçiş adımlarını tamamlayarak parametreden parametreye güncellemeleri yapar. Yakınsama sağlandığında $\beta^{(m)}$ değerleri $\pi(\beta_1, \dots, \beta_p)$ dağılımdan alınmış değerlere karşılık gelmektedir [6, 22].

Bir iterasyondan sonra $\boldsymbol{\beta}^{(1)} = \{\beta_1^{(1)}, \beta_2^{(1)}, \dots, \beta_p^{(1)}\}'$ parametre vektörü elde edilir. M iterasyondan sonra $\boldsymbol{\beta}^{(M)} = \{\beta_1^{(M)}, \beta_2^{(M)}, \dots, \beta_p^{(M)}\}'$ olarak M'inci parametre vektörü elde edilir. Bu adımlar, tam koşullu dağılımdan örnekleme çekilerek gerçekleştirilir. Bu örneklemler birbirinden bağımsız olarak çekilir [40].

Gibbs örneklemede iterasyon sayısına dayalı bazı terimlerin tanımı aşağıda verilmiştir:

Burn-in uzunluğu: Posterior çıkarsama yapmada, başlangıç değerlerinin etkisini minimum yapmak için bir Markov zinciri örnekleminin başlangıç kısmını çıkartmak amacıyla kullanılan iterasyon sayısıdır. Örneğin hedef posterior dağılımın $N(0,1)$ olduğunu ve Markov zincirinin 10^6 başlangıç değeriyle başladığını varsayalım. Markov zinciri, birkaç iterasyon sonunda 0 civarında değer almaya başlayacaktır. Bununla birlikte posterior ortalama hesaplamasında 10^6 civarında değer alan örneklemler, ortalama tahmininde iterasyonun başlangıcında güçlü bir bias meydana getirebilir. Aslında teoride, Markov zinciri sonsuz iterasyon sayısında çalışırsa, başlangıç değerlerinin etkisinin 0 değerine doğru yaklaşacağı varsayılır. Fakat uygulamada sonsuz iterasyon sayısından bahsetmek mümkün değildir. Bu nedenle bu varsayım altında, belli bir iterasyon sayısından sonra zincirin hedef dağılıma ulaşması beklenir. Ayrıca zincirin ilk kısmının atılması ve posterior çıkarsama için sadece iyi örneklemlerin kullanılması istenir. Uygulamada bu durumu sağlayan iterasyon sayısına “burn-in uzunluğu” adı verilir. Genellikle burn-in sayısı, 2000 olarak alınır [6, 22, 40].

MCMC iterasyon sayısı: Burn-in uzunluğundan sonraki iterasyon sayısıdır. Genellikle iterasyon sayısı, 10000 olarak alınır [6, 7, 22, 40].

Markov zincirinin başlangıç değerleri: Markov zincirinde posterior dağılımın başlangıç değerlerinin belirlenmesinde iki durum söz konusudur: (i) araştırmacı tarafından atanır ya da (ii) en çok olabilirlik yönteminden elde edilen parametre tahminleri başlangıç değeri olarak dikkate alınır. Genellikle uygulamalarda, en çok olabilirlik parametre tahmin yönteminden elde edilen değerler Markov zinciri için başlangıç değeri olarak kullanılır.

2.2.4. Katsayıların Prior Dağılımları

Prior dağılım, parametrenin posterior dağılımının elde edilmesinde kullanıldığı için Bayesgil çıkarsamanın en önemli öğelerinden biridir. Prior dağılım ve olabilirlik fonksiyonunun çarpımı sonucu, parametrenin posterior dağılımı elde edilir. Prior

dağılım kullanılmaksızın herhangi bir Bayesgil çıkarsama yürütülemez ya da herhangi bir modelleme yapılamaz. Prior dağılım, bilinmeyen β parametresi hakkındaki ön bilgiyi gösterir [19, 22, 35, 38, 40, 42, 43].

Prior bilginin elde edilmesi, Bayesgil çıkarsamada önemli rol oynar. Temel olarak Bayesgil analiz; bir çalışma, önceki bilgi, geçmiş deneyimler ya da harici kanıt gibi bağımsız kaynaklar (açıklayıcı prior bilgiler) ışığında gözlenen verinin subjektif (bireyin düşünce ve duygularına dayanan) yorumuna dayanır ve dolayısıyla elde edilen yeni bilginin bileşimine dayanarak gerçekleştirilir [6, 47]. Dolayısıyla Bayesgil çıkarsamada en önemli dönemeç, açıklayıcı prior dağılımın elde edilmesi ve Bayesgil analizlere alınmasıdır. Açıklayıcı prior bilginin mevcut olmadığı ya da elde edilemediği durumlarda ise açıklayıcı olmayan prior bilgiye başvurulur. Bu dağılımlar, objektif ya da açıklayıcı olmayan prior dağılımlar olarak isimlendirilir.

1. *Açıklayıcı olmayan (noninformative) prior dağılım:* Açıklayıcı olmayan prior dağılımlar, uniform ve normal dağılıma dayanarak elde edilen prior dağılımlardır.

- Uniform prior dağılım için birleşik prior yoğunluk,

$$p(\beta_1, \dots, \beta_p) \propto 1 \quad -\infty < \beta_i < \infty \quad (2.35)$$

olarak gösterilir. Uniform prior dağılım, regresyon katsayılarının alabileceği mümkün tüm değerlerinin olasılığı eşit alınarak, parametrenin konumu hakkında yetersiz bilgiyi yansıtan bir dağılım olduğu için “düz (flat) prior” olarak isimlendirilir. Genellikle uniform prior dağılım kullanılan Bayesgil çıkarsamaların sonuçları, klasik yöntemlerde olabilirliğin maksimum yapılmasıyla elde edilen parametre tahmin sonuçlarına benzemektedir [40].

- Normal prior dağılımda, β parametre vektörünün β_0 ortalama vektörlü ve Σ_0 kovaryans matrisli çok değişkenli normal dağılım gösterdiği varsayılır. Birleşik prior yoğunluk,

$$p(\beta) \propto e^{-\frac{1}{2}(\beta-\beta_0)'\Sigma_0^{-1}(\beta-\beta_0)} \quad (2.36)$$

olarak gösterilir. Burada, veri setindeki tüm değişkenler için β_0 ortalama vektörüne ve Σ_0 kovaryans matrisine rasgele sabit değer atanarak açıklayıcı olmayan prior bilgi elde edilir.

Uygun olmayan prior dağılımlar, posterior dağılımda da uygunsuzluğa yol açabilir ve bu durumda da uygun olmayan posterior dağılıma göre yapılan çıkarımlar da geçersiz olacaktır [40]. Bu nedenle prior dağılımın belirlenmesi, bilgi ve deneyim gerektiren bir konudur.

2. Açıklayıcı (informative) prior dağılım: Posterior dağılım üzerinde etkili olan ve olabilirlik fonksiyonunun egemenliği altında kalmayan prior bilgidir. Bu tip prior dağılımlar, ayrıntıya girilen ve özenle yapılan gerçek uygulamalardan belirlenebilir. Önceki çalışmalar, geçmiş deneyim ya da uzman düşüncenin birleşimiyle elde edilen bilgilerin prior dağılımının kullanılması, Bayesgil yöntemin gücünü gösterir [40].

2.2.5. Markov Zincirinin Yakınsamasının Değerlendirilmesi

Bayesgil çıkarımda Markov zinciri yakınsamasını değerlendirmek için genellikle Otokorelasyon, Geweke ve Gelman-Rubin testleri kullanılır. Bu testler, özellikle Markov zincirinin durağanlığa (belli bir değere yakınsama) ya da istenen posterior dağılıma ulaşmış ulaşmadığını değerlendirir.

1. Otokorelasyon

Zaman serilerinde otokorelasyon, bir zaman serisinin belirli bir gecikme zamanı ya da zaman aralığıyla birbirinden ayrık değerler arasındaki ilişkiyi ölçer. Bayesgil analizlerde ise otokorelasyon katsayısı, Markov zinciri örneklemeleri arasındaki bağımlılığı ölçmektedir.

Gecikme zamanı (lag) h için örneklem otokorelasyonu, örneklem otokovaryans fonksiyonuna dayanarak

$$\hat{\rho}_{(h)} = \frac{\hat{\gamma}_{(h)}}{\hat{\gamma}_{(0)}} , \quad |h| < M \quad (2.37)$$

eşitliğiyle elde edilir. β_j^m 'nin gecikme zamanı h için örneklem otokovaryans fonksiyonu,

$$\hat{Y}_{(h)} = \frac{1}{M-h} \sum_{m=1}^{M-h} (\beta_j^{m+h} - \bar{\beta}_j)(\beta_j^m - \bar{\beta}_j) \quad (2.38)$$

eşitliğiyle tanımlanır. Uzun gecikme zamanları arasındaki yüksek korelasyonlar, Markov zincirlerinin yavaş yakınsadığını gösterir. Yavaş yakınsama, Markov zincirinin parametre uzayından yavaş bir şekilde geçtiği ve zincirin aşırı bağımlı olduğu anlamına gelir. Örneklem otokorelasyon değerinin büyük olması, hesaplanan Monte Carlo standart hatasının yanlı olmasına neden olur [40].

2. Geweke Testi

Geweke tarafından 1992 yılında geliştirilen test, yakınsamanın başarısızlığını denetlemek için iki kısma ayrılmış Markov zincirinin ilk kısmındaki değerlerle zincirin ikinci kısmındaki değerleri karşılaştırır. Genellikle zincirin başlangıç kısmından %10 ve son kısmından %50'lik bölümü alınarak, bu iki kısmın iki örneklem t testi ile karşılaştırması yapılır.

Geweke testinde, β^m Markov zinciri β_1^m ($m=1, \dots, n_1$) ve β_2^m ($m=n_a, \dots, M$) olarak tanımlanan iki alt kısma ayrılır. Burada $1 < n_1 < n_a < M$ olacaktır. $n_2 = M - n_a + 1$ olarak tanımlanırsa iki alt kısım için β parametresinin posterior ortalamaları,

$$\begin{aligned} \bar{\beta}_1 &= \frac{1}{n_1} \sum_{m=1}^{n_1} \beta^m \\ \bar{\beta}_2 &= \frac{1}{n_2} \sum_{m=n_a}^M \beta^m \end{aligned} \quad (2.39)$$

eşitlikleriyle elde edilir. n_1/M ve n_2/M oranları sabit, $(n_1 + n_2)/M < 1$ ve zincir durağan ise Geweke test istatistiği,

$$Z_n = \frac{\bar{\beta}_1 - \bar{\beta}_2}{\sqrt{V_1 + V_2}} \quad (2.40)$$

eşitliğiyle elde edilir. Burada V_1 ve V_2 , sırasıyla iki MCMC zinciri için ortalamaların varyanslarıdır. Test istatistiği, asimptotik olarak standart normal dağılıma yakınsar.

$p > 0.05$ ise Markov zincirinin istenen posterior dağılıma yakınsadığı ya da durağanlaşarak yakınsama gösterdiği ifade edilir. Eşitlik (2.40)'da varyanslar,

$$V_1 = (1/n_1) \left[\gamma_{(0)} \sum_{m=1}^{n_1-1} \gamma_{(m)} \left(\frac{n_1-m}{n_1} \right) \right]$$

ve

$$V_2 = (1/n_2) \left[\gamma_{(0)} \sum_{m=n_a}^{n_2-1} \gamma_{(m)} \left(\frac{n_2-m}{n_2} \right) \right] \quad (2.41)$$

eşitlikleriyle hesaplanır. Burada $\gamma_{(0)}$, 0 gecikme zamanındaki otokovaryanstır [7, 40].

3. Gelman-Rubin Testi

Gelman-Rubin testi, her bir zincir içi ve zincirler arası varyansları karşılaştırarak çoklu tekrarlanan MCMC zincirlerinin analizine dayanır. Bu iki varyans arasında büyük sapma olması, yakınsama olmadığını gösterir [40].

Tek bir Markov zincirinin m 'inci örneklem çıktısı, β^m olarak tanımlansın ($m=1, \dots, M$). K , hedef dağılımın farklı kısımlarından başlatılan paralel Markov zincir sayısını gösterebilir. Her bir zincir M uzunluğunda tanımlansın (burn-in uzunluğu dışında kalan iterasyon sayısı). Her bir β^m için simülasyonlar, β_k^m olarak tanımlansın ($m=1, \dots, M$ ve $k=1, \dots, K$). Zincirler arası varyans,

$$V_B = \frac{M}{K-1} \sum_{k=1}^K (\bar{\beta}_k - \bar{\beta})^2 \quad (2.42)$$

eşitliğiyle gösterilir. Burada,

$$\bar{\beta}_k = \frac{\sum_{m=1}^M \beta_k^m}{M} \text{ ve } \bar{\beta} = \frac{\sum_{k=1}^K \bar{\beta}_k}{K} \quad (2.43)$$

şeklinde hesaplanır. Zincir içi varyans,

$$V_W = \frac{1}{K} \sum_{k=1}^K s_k^2 \quad (2.44)$$

eşitliğiyle gösterilir.

Burada,

$$s_k^2 = \frac{1}{M-1} \sum_{m=1}^M (\beta_k^m - \bar{\beta}_k)^2 \quad (2.45)$$

biçiminde elde edilir.

$Var(\beta|\mathbf{x})$ posterior marjinal varyans, V_B ve V_W varyanslarının ağırlıklı ortalamasıdır. Posterior marjinal varyans tahmini,

$$\widehat{Var}(\beta|\mathbf{x}) = \hat{V} = \frac{M-1}{M} V_W + \frac{K+1}{MK} V_B \quad (2.46)$$

eşitliğiyle elde edilir. Gelman-Rubin istatistiği, \hat{V}/V_W oranı şeklinde hesaplanır. K zincirin hepsi hedef dağılıma ulaşmışsa eşitlik (2.46)'daki posterior varyans tahmini, V_W zincir içi varyansa çok yakın olacaktır. Bu nedenle \hat{V}/V_W oranının, 1'e yakın olması beklenir. Bu oranın karekökü, *potansiyel ölçek indirgeme faktörü* (*potential scale reduction factor*) olarak isimlendirilir. *Potansiyel ölçek indirgeme faktörü* 1'e yakın çıkmışsa, M zincirin her birinin durağanlaştığı ve hedef dağılıma ulaştığı söylenebilir [7, 40].

2.2.6. Bayesgil Analizin Özet İstatistikleri

p boyutlu parametre vektörü $\boldsymbol{\beta} = [\beta_1, \dots, \beta_p]$ ve her bir j değişkeni ($j \in \{1, \dots, p\}$) için M sayıda sonuç olduğunda ($\beta_j = \{\beta_j^m, m = 1, \dots, M\}$), posterior dağılımın ortalaması,

$$E(\beta_j|\mathbf{x}) \approx \bar{\beta}_j = \frac{1}{M} \sum_{m=1}^M \beta_j^m \quad (2.47)$$

eşitliğiyle elde edilir.

Posterior dağılımın varyansı,

$$Var(\beta_j|\mathbf{x}) \approx s_j^2 = \frac{1}{M-1} \sum_{m=1}^M (\beta_j^m - \bar{\beta}_j)^2 \quad (2.48)$$

eşitliğiyle elde edilir [22, 40].

2.2.7. Bayesgil Analizin Uygunluk İstatistikleri

Bayesgil analizde, modelin uygunluğunun değerlendirilmesi için sapma bilgi kriteri (deviance information criterion), Akaike bilgi kriteri (Akaike's information criterion) ve Bayesgil bilgi kriteri (Bayesian information criterion ya da Schwarz criterion) kullanılır. Bir veri seti için bu kriterlerin değerleri ne kadar küçük olursa, uyum o kadar iyi olacaktır.

1. Sapma Bilgi Kriteri:

Bu kriter, posterior yoğunlukları kullanır. Sapma bilgi kriteri (SBK),

$$SBK = 2\overline{D(\boldsymbol{\beta})} - D(\bar{\boldsymbol{\beta}}) \quad (2.49)$$

eşitliğiyle elde edilir. Burada,

$\bar{\boldsymbol{\beta}}$: Posterior ortalama vektörüdür. $\bar{\boldsymbol{\beta}} = \frac{1}{M} \sum_{m=1}^M \boldsymbol{\beta}^m$ eşitliğiyle hesaplanır.

$\overline{D(\boldsymbol{\beta})}$: Sapmanın posterior ortalamasıdır. $\overline{D(\boldsymbol{\beta})} = \frac{1}{M} \sum_{m=1}^M D(\boldsymbol{\beta}^m)$ eşitliğiyle hesaplanan beklenen sapma, verinin uyumunun ne kadar iyi olduğunu ölçer.

$D(\bar{\boldsymbol{\beta}})$: $\bar{\boldsymbol{\beta}}$ için ölçülen sapmayı gösterir. $D(\bar{\boldsymbol{\beta}}) = -2\log(p(\mathbf{x}|\bar{\boldsymbol{\beta}}))$ eşitliğiyle hesaplanan sapma, en iyi posterior tahmini değerlendirir [6, 40].

2. Akaike Bilgi Kriteri:

Akaike bilgi kriteri (ABK), model seçim kriteri olup, yaygın olarak kullanılan bir ölçüdür [6, 40]. Bu kriter,

$$ABK = -2\log(L(\hat{\boldsymbol{\beta}})) + 2p \quad (2.50)$$

eşitliğiyle elde edilir. Burada p , parametre sayısı; $\hat{\boldsymbol{\beta}}$, $\boldsymbol{\beta}$ parametresinin en çok olabirlik tahmini ve $L(\hat{\boldsymbol{\beta}})$, olabirlik fonksiyonudur.

3. Bayesgil Bilgi Kriteri:

Bayesgil bilgi kriteri (BBK), model seçim kriteri olup, yaygın olarak kullanılan bir diğer ölçüdür [6, 40]. Bu kriter,

$$\text{BBK} = -2\log(L(\hat{\beta})) + p\log(N_e) \quad (2.51)$$

eşitliğiyle elde edilir. Burada N_e , tamamlanmış (uncensored) gözlemlerin sayısıdır.

3. GEREÇ VE YÖNTEM

Bu çalışmanın uygulama bölümü, iki farklı aşamadan oluşmuştur. Birinci aşamasında, yöntemlerin performanslarının karşılaştırılması amacıyla Monte Carlo simülasyon yöntemi ile türetilmiş veriler üzerinde CRA ve BYA yöntemleri uygulanmış ve yöntemlerin parametre tahminleri, yanlılıkları ve standart hataları hesaplanmıştır. Bu kriterler kullanılarak değişik birim içeren ve farklı değişken yapılarına sahip verilerde yöntemlerin performansları karşılaştırılmıştır. İkinci aşamasında ise yöntemlerin uygulanabilirliğini göstermek amacıyla, 1998-2007 yıllarına ilişkin meme kanserli 423 hastanın yinelemesiz yaşam süresine etki eden risk faktörlerinin incelenmesi yapılmış ve her iki yönteme göre gerçek verilerde parametre tahminleri hesaplanmış ve risk faktörleri değerlendirilmiştir.

3.1. Kullanılan Programlar

Analizler, SAS 9.1.3 programında gerçekleştirildi. CRA için PROC PHREG ve PROC TPHREG, BYA için PROC BPHREG ve belirtici istatistiklerin hesaplanmasında ise PROC MEANS prosedürleri kullanıldı [9, 39, 40]. Veri setlerinin türetilmesinde SAS programlama dilinden yararlanıldı. Ayrıca PROC PHREG, PROC TPHREG ve PROC BPHREG prosedürlerinde Monte Carlo simülasyonu yapabilmek için SAS dilinde MACRO yazıldı.

3.2. Simülasyon Çalışması

Yaşam sürelerine ilişkin veri setlerinde kullanılan CRA ve BYA yöntemlerinin karşılaştırılmasında 1000 tekrarlı Monte Carlo simülasyon yöntemi kullanıldı. Karşılaştırmalar; parametre tahminleri, parametre yanlılıkları, parametre tahminlerinin standart hataları ve önemlilik olasılığı olmak üzere dört temel ölçüt kullanılarak yapıldı.

Simülasyonlarda kullanılmak üzere tek bağımsız değişken içeren CRA modeli $h(t, x) = h_0(t)e^{\beta'x}$ kullanıldı.

Sabit temel hazard fonksiyonlu bir CRA modeli için yaşam süreleri,

$$t = -\frac{\log(U)}{h_0(t)e^{\beta'x}} = \frac{E}{h_0(t)e^{\beta'x}} \quad (3.1)$$

eşitliği kullanılarak türetilmiştir. Burada $U \sim \text{Uniform}(0,1)$ ve $E \sim \text{Üstel}(1)$ olarak tanımlıdır. U değişkeni 0 ve 1 parametreleriyle uniform dağılım gösterdiğinde $-\log(U)$, 1 parametrelili üstel dağılım gösterir. Bu nedenle yaşam sürelerinin türetilmesinde, üstel ya da uniform dağılımdan yararlanılır [1, 26].

3.2.1. Simülasyon Algoritmaları

Bu çalışmada, CRA ile açıklayıcı olmayan prior bilgili BYA yönteminin karşılaştırılması için hazırlanan algoritma (Algoritma-I) ve CRA ile açıklayıcı prior bilgili BYA yönteminin karşılaştırılması için hazırlanan algoritma (Algoritma-II) olmak üzere iki tür algoritma kullanıldı.

1. Algoritma-I

CRA ile açıklayıcı olmayan prior bilgili BYA yönteminin karşılaştırılması için hazırlanan algoritma aşağıdaki adımlardan oluşmaktadır.

1. Tahmin edilecek olan simülasyon parametresi β 'ya sabit bir değer atandı.
2. Örneklem büyüklüğü için sabit bir değer atandı.
3. Tamamlanmış gözlem oranını belirlemek için temel hazard fonksiyonuna ($h_0(t)$) sabit bir değer atandı.
4. U değişkeni için $U \sim \text{Uniform}(0,1)$ olacak biçimde veri üretimi yapıldı.
5. n gözlem içeren veri seti oluşturmak için bağımsız değişkene ilişkin iki farklı dağılıma göre veri üretimi yapıldı:

$$x \sim \text{Uniform}(0,1) \text{ ve } x \sim \text{Binom}(1,0.5)$$

6. İlk 4 adımda elde edilen deęerler, eřitlik (3.1)'de yerine konularak t yařam sũreleri elde edildi.

7. Tamamlanmıř verileri elde etmek iin 1 parametrelili ũstel daęılımdan t_sũre olarak tanımlanan deęiřken iin veri tũretildi. Bu veriler, 100 ile arpılarak deęerleri yařam sũresi deęiřkeninin deęerlerine yakın tutuldu. Eęer $t \leq t_sũre$ ise 1 (tamamlanmıř) aksi durumda ise 0 (tamamlanmamıř) olacak řekilde kořullu olarak durum deęiřkeni elde edildi.

8. Elde edilen veriler, CRA iin PROC PHREG ve PROC TPHREG, BYA iin PROC BPHREG prosedũrleri ile analiz edildi.

9. Yapılan analizler sonucunda her bir yũntemden elde edilen parametre tahminleri ve parametre tahminlerine ait standart hatalar kaydedildi.

10. 4-9. adımlar 1000'er kez tekrarlandı.

2. Algoritma-II

Bu algoritmada, simũlasyon parametresine baęlı olarak belirlenen aıklayıcı prior bilgilere dayanarak Algoritma-I'deki adımlara gũre BYA ile CRA uygulandı.

3.2.2. Simũlasyon Parametreleri

Algoritma-I iin simũlasyon parametreleri:

Simũlasyon alıřmasında ȳrnekleme bũy¼kl¼ęũ (n) olarak 30, 50, 100, 250 ve 500 kullanıldı.

Algoritmanın 1. adımımda bahsedilen β parametresi iin yanlılıkların ve standart hataların elde edilmesinde ũ farklı deęer ($\beta = 0, 1$ ve $\ln 2$) kullanıldı.

Temel hazard fonksiyonu ($h_0(t)$), tamamlanmamıř gũzlem oranı %70 civarında olacak řekilde 0.005 olarak alındı.

Aıklayıcı olmayan prior bilgi olarak, normal prior daęılım kullanıldı.

Algoritma-II için simülasyon parametreleri:

Simülasyon çalışmasında örneklem büyüklüğü olarak, $n= 30, 50, 100, 250$ ve 500 kullanıldı. Algoritma-I'in 1. adımında bahsedilen β parametresi için sadece $\beta = 1$ değeri kullanıldı.

Algoritma-I'de olduğu gibi temel hazard fonksiyonu, tamamlanmamış gözlem oranı %70 civarında olacak şekilde 0.005 olarak alındı.

BYA'da simülasyon parametresi olarak verilen $\beta = 1$ değerine göre belirlenen, prior dağılım ortalaması $\bar{\beta}_p = 1, 1.1, 1.5, 2$ ve prior dağılım varyansı $\sigma_p^2 = 0.01, 0.05, 0.1, 0.5$ olan açıklayıcı prior bilgiler alındı.

Her iki algoritmada da BYA'da kullanılan Gibbs örnekleme yaklaşımında burn-in uzunluğu için 2000 ve posterior hesaplamalar için 10000 iterasyon yapıldı.

3.2.3. Karşılaştırma Ölçütleri

Parametre tahminleri: Her bir yöntem için, 1000 farklı parametre tahmininin ortalaması alındı ($\bar{\beta}$).

Parametre yanlışlıkları: Parametre ile ortalama arasındaki fark alınarak ($\bar{\beta} - \beta$) her bir yöntem için yanlışlıklar hesaplandı.

Parametre tahminlerinin standart hataları: Her bir yöntem için, 1000 farklı parametre tahmininin standart hatasının ortalaması alındı ($\hat{\sigma}_{\bar{\beta}}$).

Önemlilik olasılığı (p): Parametre tahminlerinin simülasyon parametresine yakınsamanın sağlanıp sağlanmadığının test edilmesinde tek örneklem t testi kullanıldı ve p olasılıklarına göre karar verildi.

Algoritma-I için 60000 ve Algoritma-II için 170000 olmak üzere toplam 230000 simülasyon yapıldı.

3.3. Gerçek Veri Seti Çalışması

Trakya Üniversitesi Tıp Fakültesi Radyasyon Onkolojisi Anabilim Dalı'na 1998-2007 yılları arasında başvuran ve meme kanseri tanısı konulan 547 hastanın yinelemesiz yaşam sürelerine ilişkin veri seti kullanıldı.

Meme kanserinde yinelemeyi etkileyen risk faktörleri olarak; yaş, menopoz durumu, ilk doğum yaşı, aile hikayesi, tümör kadranı, tümör büyüklüğü, progesteron reseptör durumu, östrojen reseptör durumu, aksiller lenf nodu tutulumu, kapsül dışı lenf nodu tutulumu, histolojik tip ve tedavi (radyoterapi ve hormon terapisi) alındı. İlk doğum yaşı ≥ 30 olanların meme kanserine yakalanma ve hastalığın yinelemesi bakımından daha riskli olduğu için çalışmamızda da ilk doğum yaşı değişkeni < 30 ve ≥ 30 olmak üzere iki alt kategoriye ayrıldı [33]. Veri setinde eksik gözlemler birimler çalışmadan çıkarıldı. Böylece analizler, 423 birim üzerinden yürütüldü.

Yaş ve tümör büyüklüğü değişkenlerinin normal dağılıma uygunluğu Kolmogorov Smirnov testi ile incelendi (Tablo 3.1). Her iki nicel değişkende normal dağılıma uygunluk göstermediği için yineleme olan ve olmayan gruplara göre karşılaştırmada Mann-Whitney U testi kullanıldı ve sonuçlar medyan (%25-%75 persantil) şeklinde gösterildi. Kategorik yapıdaki değişkenlerin yineleme olan ve olmayan gruplara göre karşılaştırmasında ise Ki-kare analizi kullanıldı.

Tablo 3.1 incelendiğinde, meme kanserli hastalarda kanserin yineleme durumu üzerinde tümör büyüklüğü, progesteron reseptör durumu, hormon terapisi, aksiller lenf nodu tutulumu, tümör kadranı ve kapsül dışı lenf nodu tutulumu değişkenlerinin önemli etkisi olduğu belirlendi.

BYA uygulamasında, ele alınan bağımsız değişkenler için geçmişte yapılan çalışmalardan prior bilgi elde edildi [13, 15, 32, 34, 50]. Kaynak taraması sonucunda, bazı değişkenler için daha önce yapılmış benzer bir çalışmaya ulaşılamaması nedeniyle bu değişkenler için açıklayıcı olmayan normal prior bilgi kullanıldı. Her bir değişken ve kategorik değişkenlerin referans dışındaki kategorileri için parametre ($\hat{\beta}_p$) ve standart hatasına ($\hat{\sigma}_{\hat{\beta}_p}$) ilişkin prior bilgiler, Tablo 3.2'de verildi.

Tablo 3.1. Meme kanserli hastalara ilişkin veri setinde yineleme durumuna göre bağımsız değişkenlerin tanımlayıcı istatistikleri

Bağımsız Değişken		Yineleme		z	p
		Yok (n=303)	Var (n=120)		
Yaş (yıl) medyan (%25-%75)		49 (43-58)	52 (43-60)	-1.029	0.304
Tümör Büyüklüğü (cm) medyan (%25-%75)		3 (2-4)	3.5 (2.1-5)	-3.769	<0.001
		n (%)	n (%)	χ^2	p
İlk Doğum Yaşı (yıl)	≥30	13 (4.3)	7 (5.8)	0.176	0.675
	<30	290 (95.7)	113 (94.2)		
Menopoz Durumu	Post Menopoz	142 (46.9)	64 (53.3)	1.440	0.230
	Pre+Peri Menopoz	161 (53.1)	56 (46.7)		
Progesteron Reseptör Durumu	Negatif	76 (25.1)	46 (38.3)	7.354	0.007
	Pozitif	227 (74.9)	74 (61.7)		
Östrojen Reseptör Durumu	Negatif	73 (24.1)	40 (33.3)	3.749	0.053
	Pozitif	230 (75.9)	80 (66.7)		
Radyoterapi	Yok	60 (19.8)	17 (14.2)	1.474	0.225
	Var	243 (80.2)	103 (85.8)		
Hormon Terapisi	Yok	59 (19.5)	44 (36.7)	13.796	<0.001
	Var	244 (80.5)	76 (63.3)		
Aile Hikayesi	Yok	215 (71.0)	77 (64.2)	1.858	0.395
	Meme Kanseri	25 (8.3)	12 (10.0)		
	Diğer Kanserler	63 (20.8)	31 (25.8)		
Aksiller Lenf Nodu Tutulumu	Negatif	139 (45.9)	27 (22.5)	32.613	<0.001
	Pozitif Lenf Nodu 1-3	85 (28.1)	28 (23.3)		
	Pozitif Lenf Nodu ≥4	79 (26.1)	65 (54.2)		
Tümör Kadranı	Tek Merkezli	276 (91.1)	96 (80.0)	8.950	0.003
	Çok Merkezli	27 (8.9)	24 (20.0)		
Kapsül Dışı Lenf Nodu Tutulumu	Negatif	215 (71.0)	55 (45.8)	23.501	<0.001
	Pozitif	88 (29.0)	65 (54.2)		
Histolojik Tip	Duktal	243 (80.2)	104 (86.7)	2.021	0.155
	Diğer	60 (19.8)	16 (13.3)		

Tablo 3.2. Meme kanserli hastalara ilişkin veri setinde bağımsız değişkenlere ilişkin prior bilgiler

Bağımsız Değişken		$\hat{\beta}_p$	$\hat{\sigma}_{\beta_p}$	Kaynak
Yaş (yıl)		-0.654	0.316	Yu ve ark. (1995)
Tümör Büyüklüğü (cm)		0.351	0.077	Foekens ve ark. (1992)
İlk Doğum Yaşı (yıl)	≥ 30	-	-	-
Menopoz Durumu	Pre+Peri Menopoz	0.163	0.285	Gasparini ve ark. (1997)
Progesteron Reseptör Durumu	Negatif	0.626	0.285	Pinto ve ark. (1999)
Östrojen Reseptör Durumu	Negatif	0.610	0.296	Pinto ve ark. (1999)
Radyoterapi	Var	-0.062	0.102	Rowlings ve ark. (1999)
Hormon Terapisi	Var	-	-	-
Aile Hikayesi	Meme Kanseri Diğer Kanserler	-	-	-
Aksiller Lenf Nodu Tutulumu	Pozitif Lenf Nodu 1-3 Pozitif Lenf Nodu ≥ 4	0.698 1.203	0.167 0.171	Foekens ve ark. (1992)
Tümör Kadranı	Çok Merkezli	-	-	-
Kapsül Dışı Lenf Nodu Tutulumu	Pozitif	-	-	-
Histolojik tip	Duktal	-0.035	0.335	Gasparini ve ark. (1997)

3.3.1. Modele Girecek Bağımsız Değişkenlerin Aşamalı Seçimi

Bu çalışmada, her iki yöntemde de (CRA ve BYA) modele girecek bağımsız değişkenlerin seçiminde aşamalı seçim uygulandı. Gerçek veri setine yönelik uygulamada her iki yöntem için de modele girecek değişkenlerin seçiminde kritik önemlilik düzeyi olarak 0.05 olasılığı alınırken, modelden atılacak değişkenlerin seçiminde kritik önemlilik düzeyi olarak 0.10 olasılığı alındı.

3.3.2. Yaşam Fonksiyonlarının Hesaplanması

BYA için SAS paket programında yaşam fonksiyonlarının belirlenmesi için hazır bir prosedür olmadığı için yaşam fonksiyonları, Excel programında hesaplandı ve grafikleri çizildi.

Eşitlik (2.3)'de verilen yaşam fonksiyonlarının hesaplanmasında aşağıdaki adımlar izlendi [21, 24, 30, 46]:

1. Büyükten küçüğe sıralanmış her bir yaşam süresi için $e^{\hat{\beta}'x}$ tahminleri ve bu tahminlerin yığılımlı toplamları ($\sum_{l \in R(t_{(i)})} \exp(\hat{\beta}'x_l)$) elde edildi. Burada $R(t_{(i)})$, $t_{(i)}$ anında gözlenen $x_{(i)}$ bağımsız değişken vektörüdür.

2. Gözlenen n birimlik veri setindeki k farklı yaşam süresi ($t_{(1)}, t_{(2)}, \dots, t_{(k)}$) küçükten büyüğe sıralandı ve her bir yaşam süresinde tamamlanmış gözlem sayıları belirlendi ($n_{(1)}, n_{(2)}, \dots, n_{(k)}$).

3. Her bir $t_{(i)}$ anındaki yığılımlı hazard oranını hesaplamak için

$$\hat{\Lambda}_0(t) = \sum_{t_{(i)} \leq t} \frac{n_{(i)}}{\sum_{l \in R(t_{(i)})} \exp(\hat{\beta}'x_l)} \quad (3.2)$$

eşitliğiyle gösterilen Breslow'un kestiricisi hesaplandı.

4. Eşitlik (2.3)'de yer alan temel yaşam fonksiyonu $S_0(t)$ 'nin tahmini,

$$\hat{S}_0(t) = \exp[-\hat{\Lambda}_0(t)] \quad (3.3)$$

eşitliğiyle elde edildi.

5. Bağımsız değişkenler setinin ortalama vektörü için düzeltilmiş yaşam eğrileri,

$$\hat{S}(t, \mathbf{x}) = [e^{-\hat{\Lambda}_0(t)}]^{e^{\hat{\beta}'x}} = [\hat{S}_0(t)]^{e^{\hat{\beta}'x}} \quad (3.4)$$

eşitliğiyle elde edildi.

6. CRA ve BYA için elde edilen düzeltilmiş yaşam fonksiyonlarının grafiği çizildi.

4. BULGULAR

4.1. Simülasyon Sonuçları

4.1.1. Açıklayıcı Olmayan Prior Bilgiye Dayalı Simülasyon Sonuçları

Algoritma-I kullanılarak yapılan simülasyonlar sonucunda yöntemlerin karşılaştırılmasında kullanılan ölçütlere göre (birim sayısı, parametre değeri ve değişken tipi) her bir yöntemden elde edilen parametre tahminleri, standart hatalar ve yanlılıklara ilişkin elde edilen bulgular sırasıyla aşağıda verildi.

Her bir yöntemden elde edilen parametre tahminleri ve yanlılıkları farklı birim, parametre ve değişken tipine göre Tablo 4.1, Tablo 4.2 ve Tablo 4.3’de verildi. Tablo 4.1’de sunulan bulgular $\beta = 0$, Tablo 4.2’de sunulan bulgular $\beta = 1$ ve Tablo 4.3’de sunulan bulgular $\beta = \ln(2)$ parametre değeri için elde edildi. Yanlılık değerleri $\bar{\beta} - \beta$ biçiminde hesaplandı. Tablo 4.1, 4.2 ve 4.3’de her bir yöntem için verilen parametre tahminlerinin ve parametre tahminlerinin yanlılıklarının grafiksel gösterimleri Şekil 4.1-4.12’de verildi.

Tablo 4.1-4.3 ve Şekil 4.1-4.12’de görüldüğü gibi BYA ve CRA yöntemlerinde verilen parametre değerine yakın bir tahmin değerine ulaşıldı. Fakat BYA’da genel olarak parametreler, CRA’ya göre daha düşük yanlılıkla tahmin edildi. Her iki yöntemde de parametre 0, 1 ve $\ln(2)$ olduğunda ve bağımsız değişken uniform dağılım gösterdiğinde standart hata değerleri, binom dağılıma göre biraz daha büyük bulundu. Bunun yanı sıra BYA ve CRA yöntemlerinden elde edilen parametre tahminlerinin, tüm koşullarda simülasyon parametresine önemli düzeyde yakın tahmin yaptığı ve yakınsamanın sağlandığı gözlemlendi ($p > 0.05$). Buna göre simülasyon parametresini tahminde açıklayıcı olmayan prior bilgi kullanılan BYA’nın performansının CRA’ya göre üstünlüğü saptanmadı.

Tablo 4.1. $\beta=0$ için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri, standart hatalar ve yanlılıklar

Değişken tipi	n	CRA				BYA			
		$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$	Yanlılık	p	$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$	Yanlılık	p
Uniform	30	0.0523	1.2214	0.0523	0.681	0.0454	1.2613	0.0454	0.972
	50	0.0557	0.8929	0.0557	0.536	0.0438	0.9058	0.0438	0.962
	100	0.0164	0.6206	0.0164	0.978	0.0086	0.6180	0.0086	0.989
	250	0.0130	0.3821	0.0130	0.973	0.0079	0.3817	0.0079	0.984
	500	0.0105	0.2694	0.0105	0.969	0.0068	0.2686	0.0068	0.980
Binom	30	0.0264	1.5592	0.0264	0.987	0.0178	0.7489	0.0178	0.981
	50	0.0183	0.5386	0.0183	0.973	0.0137	0.5189	0.0137	0.979
	100	0.0171	0.3591	0.0171	0.962	0.0117	0.3558	0.0117	0.974
	250	0.0044	0.1556	0.0044	0.977	0.0023	0.1554	0.0023	0.988
	500	0.0037	0.2212	0.0037	0.987	0.0010	0.2210	0.0010	0.996

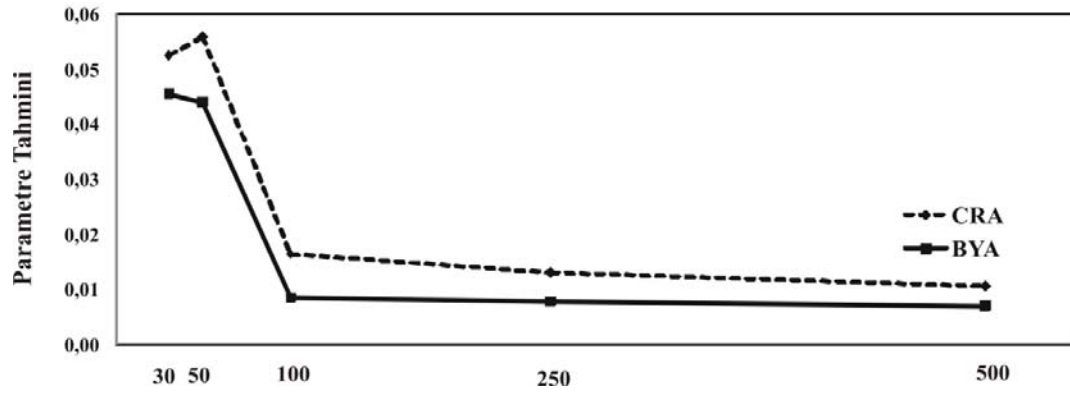
Tablo 4.2. $\beta=1$ için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri, standart hatalar ve yanlılıklar

Değişken tipi	n	CRA				BYA			
		$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$	Yanlılık	p	$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$	Yanlılık	p
Uniform	30	1.0515	0.7495	0.0515	0.709	1.0567	0.7551	0.0567	0.941
	50	1.0437	0.5659	0.0437	0.939	1.0429	0.5368	0.0429	0.937
	100	1.0433	0.5653	0.0433	0.939	1.0417	0.5367	0.0417	0.938
	250	1.0165	0.3349	0.0165	0.961	1.0151	0.3337	0.0151	0.964
	500	1.0100	0.2360	0.0100	0.966	1.0083	0.2350	0.0083	0.972
Binom	30	1.0491	0.4552	0.0491	0.915	1.0311	0.4634	0.0311	0.947
	50	1.0281	0.3436	0.0281	0.935	1.0217	0.3415	0.0217	0.950
	100	1.0126	0.2369	0.0126	0.958	1.0141	0.2366	0.0141	0.953
	250	1.0087	0.1478	0.0087	0.953	1.0082	0.1471	0.0082	0.956
	500	1.0023	0.1040	0.0023	0.982	1.0017	0.1035	0.0017	0.987

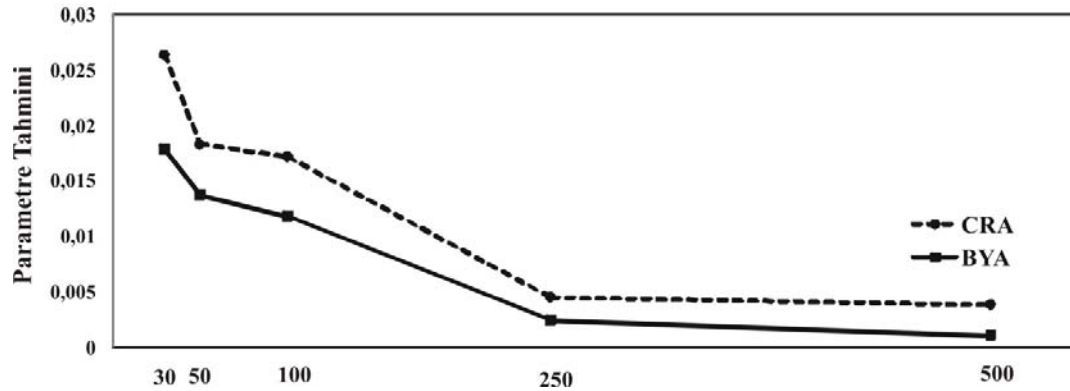
Tablo 4.3. $\beta=\ln(2)$ (0.6931) için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri, standart hatalar ve yanlılıklar

Değişken tipi	n	CRA				BYA			
		$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$	Yanlılık	p	$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$	Yanlılık	p
Uniform	30	0.6859	0.8020	-0.0072	0.993	0.6884	0.8100	-0.0047	0.995
	50	0.6880	0.7452	-0.0051	0.995	0.6885	0.7509	-0.0046	0.995
	100	0.6963	0.3867	0.0032	0.993	0.6918	0.2393	-0.0013	0.996
	250	0.6948	0.1692	0.0017	0.992	0.6925	0.1685	-0.0006	0.997
	500	0.6947	0.3860	0.0016	0.997	0.6929	0.2402	-0.0002	0.999
Binom	30	0.7285	0.4427	0.0354	0.937	0.7207	0.4358	0.0276	0.950
	50	0.7075	0.4851	0.0144	0.976	0.6865	0.3261	-0.0066	0.984
	100	0.6881	0.2473	-0.0050	0.984	0.6923	0.1430	-0.0008	0.996
	250	0.6945	0.1426	0.0014	0.992	0.6934	0.1421	0.0003	0.998
	500	0.6921	0.3285	-0.0010	0.998	0.6933	0.1437	0.0002	0.999

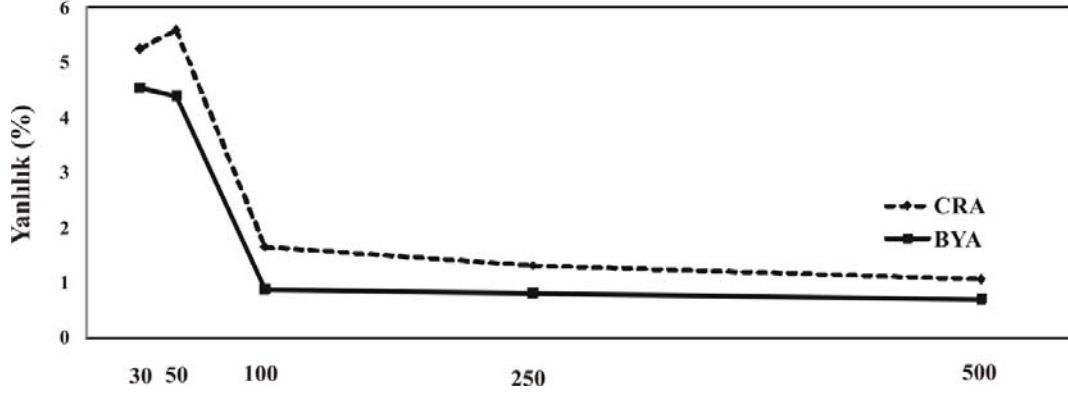
$\beta=0$ için; Şekil 4.1 ve 4.2’de, sırasıyla uniform ve binom dağılımlarına ve farklı örneklem büyüklüklerine göre 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri, Şekil 4.3 ve 4.4’de ise parametre tahminlerinin yanlılıkları gösterildi. Bu şekiller incelendiğinde, CRA ve BYA yöntemlerinin örneklem büyüklüğü arttıkça simülasyon parametresini tahminde yanlılıklarının giderek azaldığı gözlemlendi. Ayrıca bağımsız değişkenin dağılımı binom olduğunda parametre tahminlerinin uniform dağılıma göre daha az yanlılıkla elde edildiği gözlemlendi.



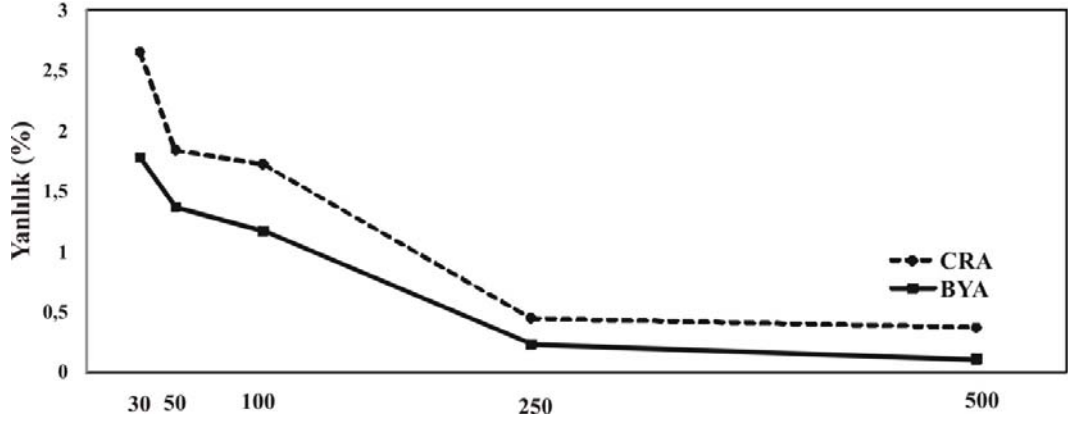
Şekil 4.1. CRA ve BYA yöntemlerinin $\beta=0$ ve uniform dağılım için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminlerinin örneklem büyüklüğüne göre değişimi



Şekil 4.2. CRA ve BYA yöntemlerinin $\beta=0$ ve binom dağılım için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminlerinin örneklem büyüklüğüne göre değişimi

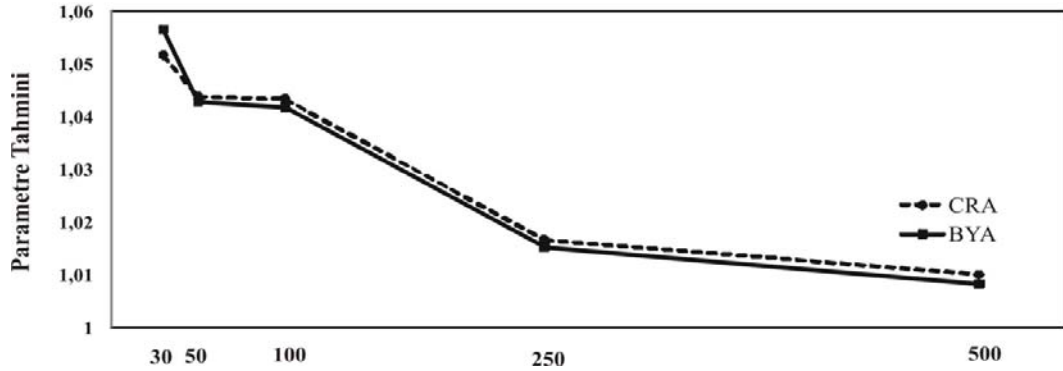


Şekil 4.3. CRA ve BYA yöntemlerinin $\beta=0$ ve uniform dağılım için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminlerinin yanlılıklarının örneklem büyüklüğüne göre değişimi

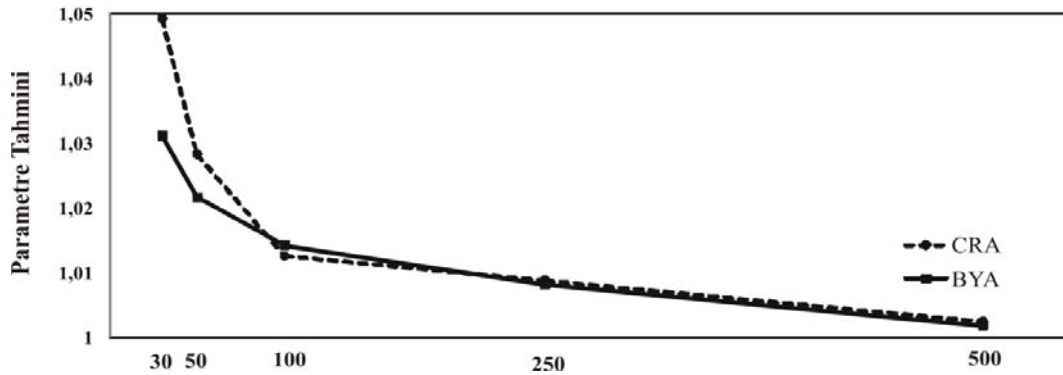


Şekil 4.4. CRA ve BYA yöntemlerinin $\beta=0$ ve binom dağılım için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminlerinin yanlılıklarının örneklem büyüklüğüne göre değişimi

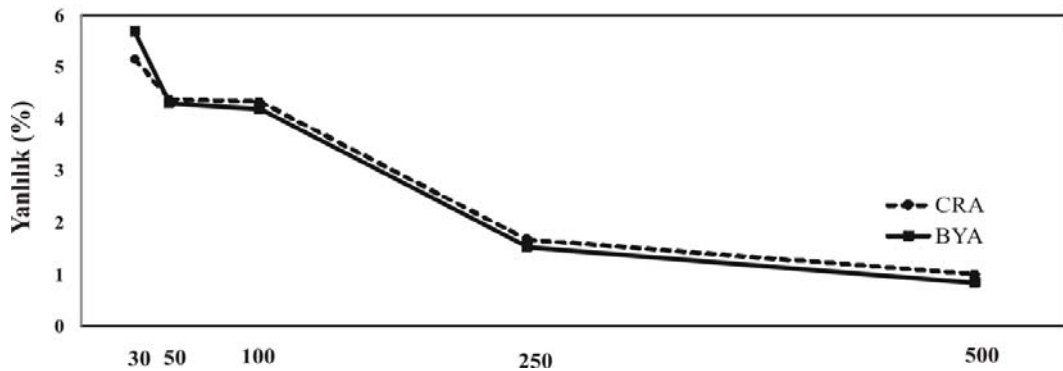
$\beta=1$ için; Şekil 4.5 ve 4.6'da, sırasıyla uniform ve binom dağılımlarına ve farklı örneklem büyüklüklerine göre 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri, Şekil 4.7 ve 4.8'de ise parametre tahminlerinin yanlılıkları gösterildi. Bu şekiller incelendiğinde, bağımsız değişkenin dağılımı binom olduğunda BYA'nın, CRA'ya göre $n=30$ örneklem büyüklüğü için simülasyon parametresini yaklaşık %2'lik yanlılıkla tahmin ettiği, BYA ve CRA yöntemlerinin örneklem büyüklüğü arttıkça simülasyon parametresini tahmininde yanlılıklarının azaldığı görüldü.



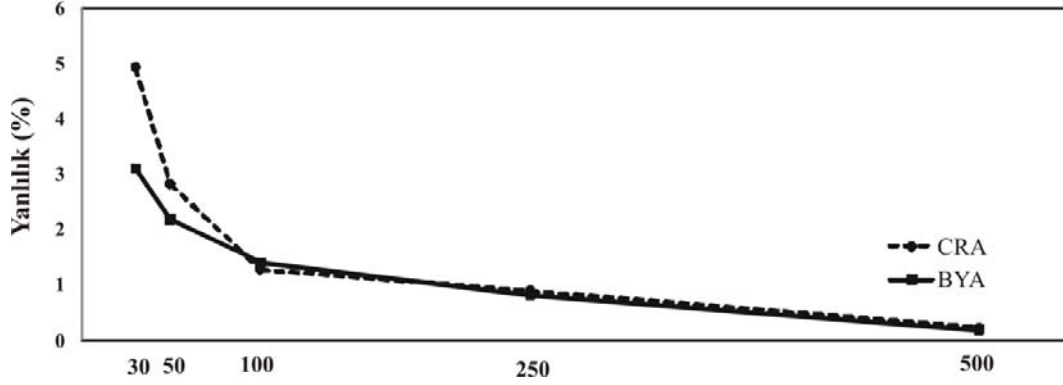
Şekil 4.5. CRA ve BYA yöntemlerinin $\beta=1$ ve uniform dağılım için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminlerinin örneklem büyüklüğüne göre değişimi



Şekil 4.6. CRA ve BYA yöntemlerinin $\beta=1$ ve binom dağılım için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminlerinin örneklem büyüklüğüne göre değişimi

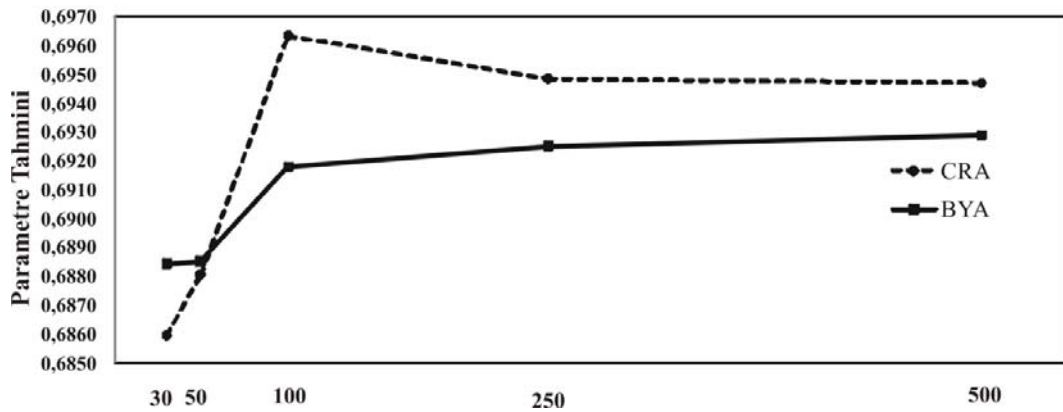


Şekil 4.7. CRA ve BYA yöntemlerinin $\beta=1$ ve uniform dağılım için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminlerinin yanlılıklarının örneklem büyüklüğüne göre değişimi

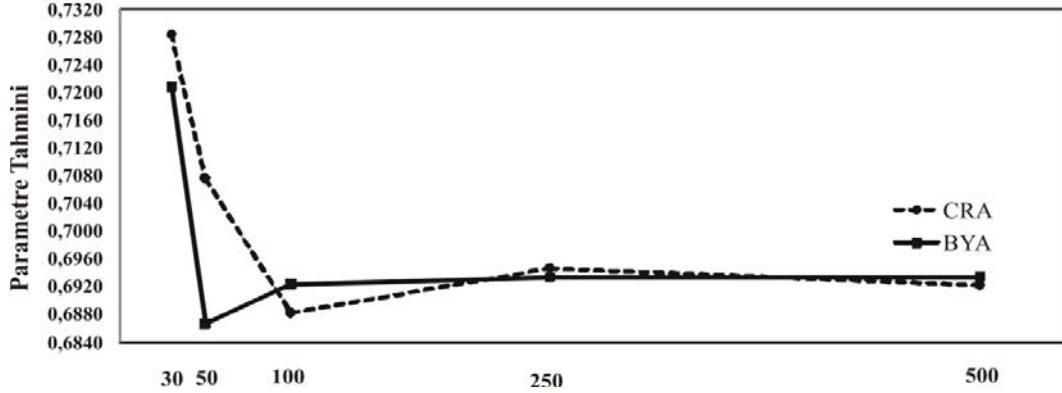


Şekil 4.8. CRA ve BYA yöntemlerinin $\beta=1$ ve binom dağılım için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminlerinin yanlılıklarının örneklem büyüklüğüne göre değişimi

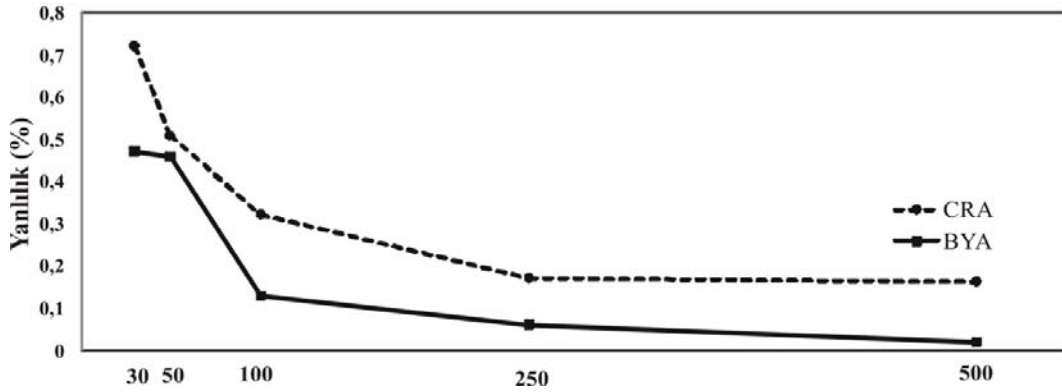
$\beta=\ln(2)$ (0.6931) için; Şekil 4.9 ve 4.10'da, sırasıyla uniform ve binom dağılımlarına ve farklı örneklem büyüklüklerine göre 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri, Şekil 4.11 ve 4.12'de parametre tahminlerinin yanlılıkları gösterildi. Bu şekiller incelendiğinde, bağımsız değişkenin dağılımı uniform olduğunda parametre tahminlerinin küçük örneklemelerde oldukça az yanlılıkla elde edildiği ve özellikle binom dağılan bağımsız değişken için örneklem büyüklüğü $n=30$ olduğunda BYA ve CRA'nın yanlılığının fazla olduğu gözlemlendi. Ayrıca örneklem büyüklüğü arttıkça CRA ve BYA yöntemlerinin, simülasyon parametresini tahminlemede yanlılıklarının giderek azaldığı gözlemlendi.



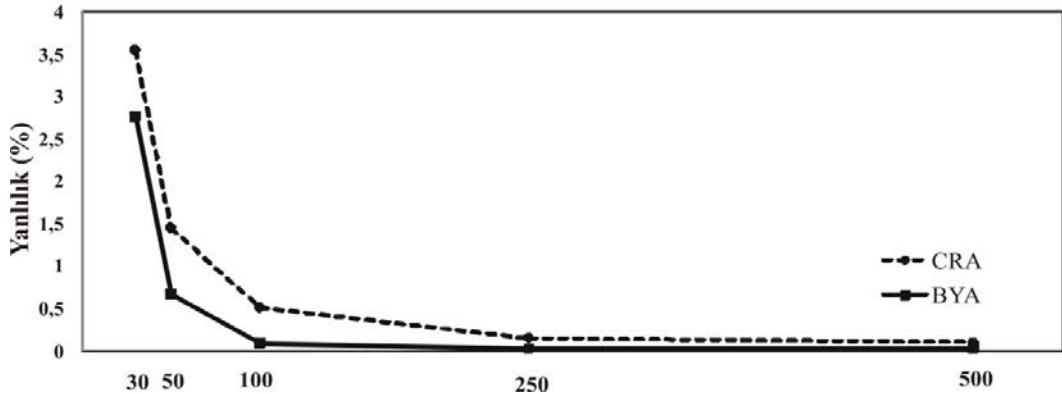
Şekil 4.9. CRA ve BYA yöntemlerinin $\beta=\ln(2)$ ve uniform dağılım için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminlerinin yanlılıklarının örneklem büyüklüğüne göre değişimi



Şekil 4.10. CRA ve BYA yöntemlerinin $\beta=\ln(2)$ ve binom dağılım için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminlerinin örneklem büyüklüğüne göre değişimi



Şekil 4.11. CRA ve BYA yöntemlerinin $\beta=\ln(2)$ ve uniform dağılım için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminlerinin yanlılıklarının örneklem büyüklüğüne göre değişimi



Şekil 4.12. CRA ve BYA yöntemlerinin $\beta=\ln(2)$ ve binom dağılım için 1000 Monte Carlo simülasyonundan elde edilen parametre tahminlerinin yanlılıklarının örneklem büyüklüğüne göre değişimi

4.1.2. Açıklayıcı Prior Bilgiye Dayalı Simülasyon Sonuçları

Algoritma-II kullanılarak yapılan simülasyonlar sonucunda yöntemlerin karşılaştırılmasında kullanılan ölçütlere göre (parametre değeri, birim sayısı, prior bilgi ve değişken tipi) her bir yöntemden elde edilen parametre tahminleri, standart hatalar ve yanlılıklara ilişkin elde edilen bulgular incelendi.

Farklı birim sayısı, prior bilgi ve değişken tiplerine göre her bir yöntemden elde edilen parametre tahminleri, parametre tahminlerinin standart hataları ve yanlılıkları Tablo 4.4-4.8 ve Şekil 4.13-4.22’de verildi.

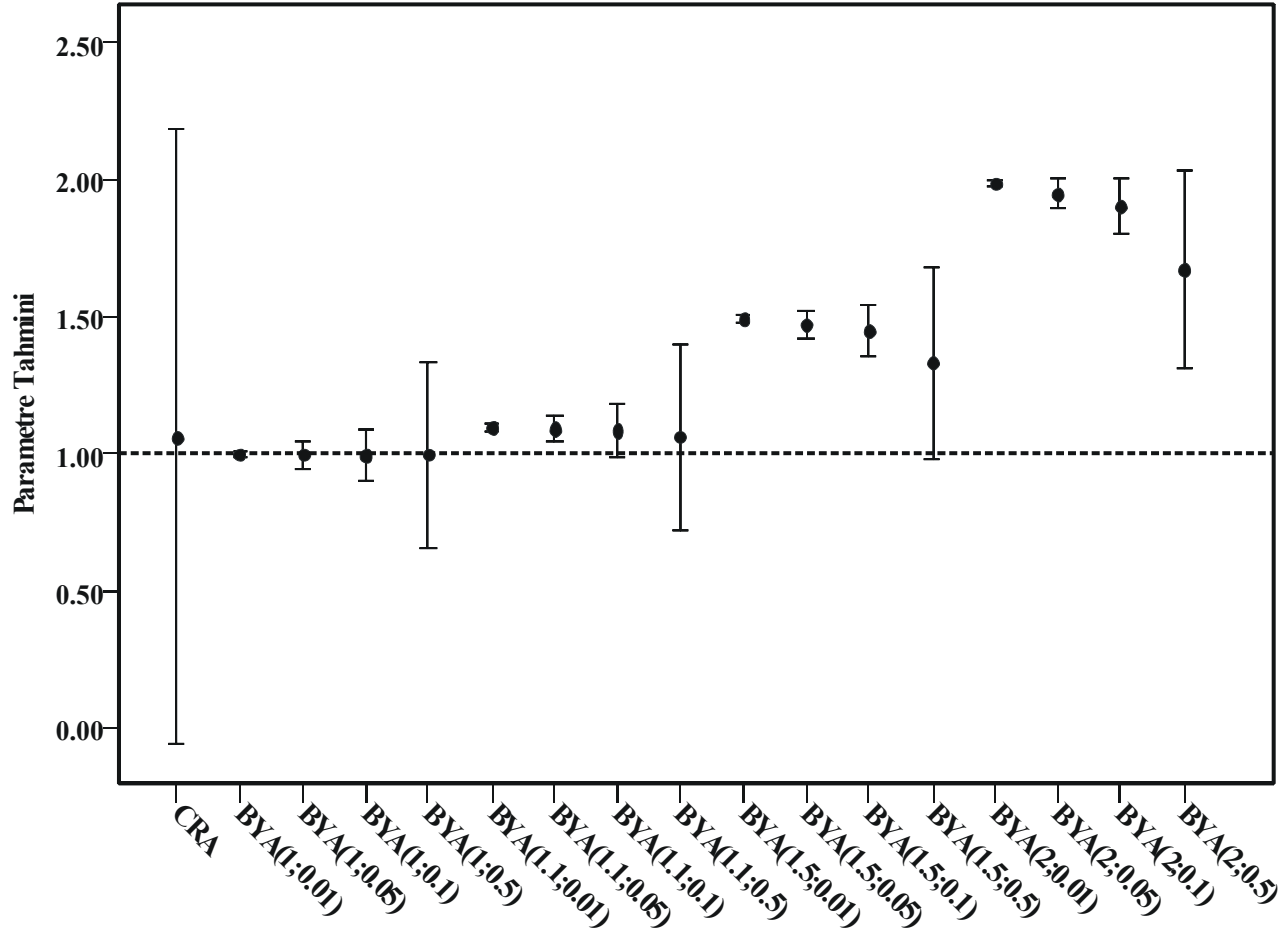
Tablo 4.4, Şekil 4.13 ve Şekil 4.14’de $n=30$ için CRA ve farklı prior bilgili BYA sonuçları incelendiğinde, veri yapısına uygun ve gerçeği yansıtan prior bilgi olduğunda ($\bar{\beta}_p = 1$ ve $\sigma_p^2=0.01, 0.05, 0.01$ ve 0.5) her iki yöntemde de yakınsamanın sağlandığı belirlendi ($p>0.05$). Ayrıca prior bilgiye ilişkin varyansın da BYA’nın performansı üzerinde önemli etkisinin olduğu ve prior bilginin parametresi sabit tutulup varyans artırıldığında simülasyon parametresinin tahmininde önemli değişme olduğu gözlemlendi. BYA’da gerçeği yansıtmayan prior bilgi kullanıldığında ($\bar{\beta}_p = 1.1, 1.5, 2$) ise simülasyon parametresine daha uzak ve prior bilgi parametresine daha yakın bir tahmin değeri elde edildi. Gerçeği yansıtmayan prior bilginin varyansı artırıldığında parametre tahmininin, simülasyon parametresine biraz daha yakınsadığı gözlemlendi. $\bar{\beta}_p = 1.1$ ve $\sigma_p^2=0.01$ için simülasyon parametresinden önemli düzeyde farklı tahmin yapıldığı ve yakınsamanın olmadığı belirlendi ($p<0.001$). $\bar{\beta}_p = 1.1$ için varyans değeri arttıkça istatistiksel olarak önemli düzeyde yakınsama sağlandığı belirlendi ($p>0.05$). $\bar{\beta}_p = 1.5, 2$ ve $\sigma_p^2=0.01, 0.05$ ve 0.1 için simülasyon parametresinden önemli düzeyde farklı tahmin yapıldığı ve yakınsamanın olmadığı ($p<0.01$ ve $p<0.001$) ve en büyük varyans değerine ulaşıldığında ($\sigma_p^2=0.5$) önemli düzeyde yakınsama sağlandığı belirlendi ($p>0.05$).

Tablo 4.5, Şekil 4.15 ve Şekil 4.16’da $n=50$ için CRA ve farklı prior bilgili BYA sonuçları incelendiğinde, veri yapısına uygun ve gerçeği yansıtan açıklayıcı prior bilgi olduğunda ($\bar{\beta}_p = 1$ ve $\sigma_p^2=0.01, 0.05, 0.01$ ve 0.5), her iki yöntemde de yakınsamanın sağlandığı ($p>0.05$) ve bunun yanı sıra BYA’nın CRA’ya göre simülasyon

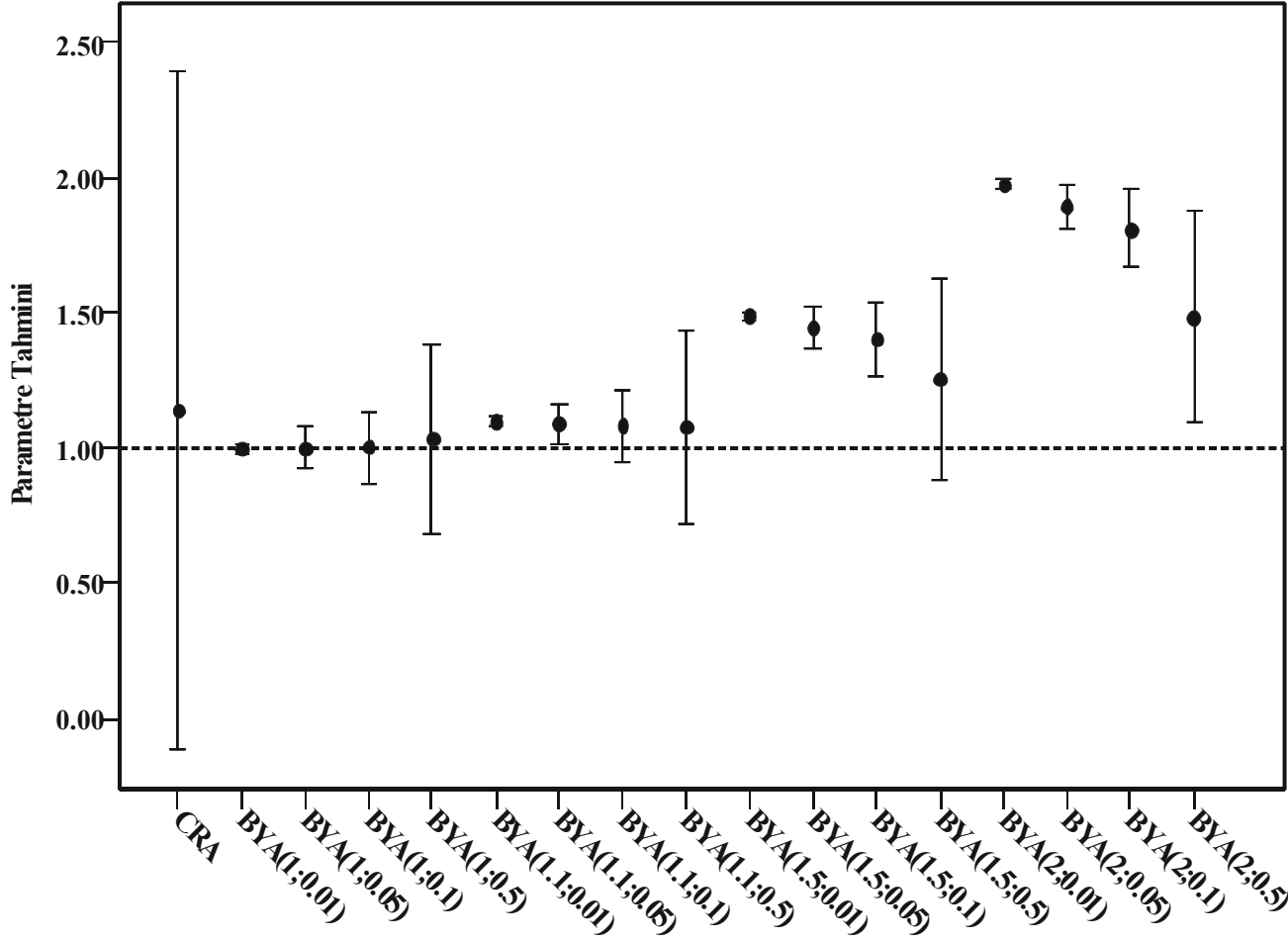
parametresini daha küçük standart hatayla tahmin ettiği bulundu. Ayrıca prior bilgiye ilişkin varyansın da BYA'nın performansı üzerinde önemli etkisinin olduğu ve prior bilginin parametresi sabit tutulup varyans artırıldığında simülasyon parametresine yakınsamada önemli değişme olduğu gözlemlendi. BYA'da gerçeği yansıtmayan prior bilgi kullanıldığında ($\bar{\beta}_p = 1.1, 1.5, 2$) ise simülasyon parametresine daha uzak ve prior bilgi parametresine daha yakın bir tahmin değeri elde edildi. Gerçeği yansıtmayan prior bilginin varyansı artırıldığında parametre tahmininin, simülasyon parametresine biraz daha yakınsadığı ve yanlılığın azaldığı gözlemlendi. $\bar{\beta}_p = 1.1$ ve $\sigma_p^2=0.01$ için simülasyon parametresinden önemli düzeyde farklı tahmin yapıldığı ve iyi bir yakınsamanın sağlanmadığı belirlendi ($p<0.001$). $\bar{\beta}_p = 1.1$ için varyans değeri arttıkça istatistiksel olarak önemli düzeyde yakınsama sağlandığı belirlendi ($p>0.05$). $\bar{\beta}_p = 1.5$ ve $\sigma_p^2=0.01, 0.05$ ve 0.1 için simülasyon parametresinden önemli düzeyde farklı tahmin yapıldığı ve yakınsamanın olmadığı (varyansa göre uniform dağılım için sırasıyla $p<0.001, p<0.001, p<0.01$ ve binom dağılım için sırasıyla $p<0.001, p<0.001, p<0.05$) ve en büyük varyans değerine ulaşıldığında ($\sigma_p^2=0.5$) önemli düzeyde yakınsama sağlandığı belirlendi ($p>0.05$). Aynı şekilde $\bar{\beta}_p = 2$ ve $\sigma_p^2=0.01, 0.05$ ve 0.1 için simülasyon parametresinden önemli düzeyde farklı tahmin yapıldığı ve yakınsamanın sağlanmadığı ($p<0.001$) ve en büyük varyans değerine ulaşıldığında ($\sigma_p^2=0.5$) önemli düzeyde yakınsama sağlandığı bulundu ($p>0.05$).

Tablo 4.4. Uniform ve binom dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=30$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri, standart hataları ve yanlılıkları

Yöntem	Prior Bilgi		Bağımsız Değişkenin Dağılımı							
	$\bar{\beta}_p$	σ_p^2	Uniform				Binom			
			$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$	Yanlılık	p	$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$	Yanlılık	p
CRA	-	-	1.0604	1.1207	0.0604	0.957	1.1399	1.2590	0.1399	0.912
	1	0.01	0.9985	0.0103	-0.0015	0.885	0.9992	0.0166	-0.0008	0.962
	1	0.05	0.9965	0.0493	-0.0035	0.944	1.0005	0.0744	0.0005	0.995
	1	0.1	0.9953	0.0938	-0.0047	0.960	1.0038	0.1320	0.0038	0.977
	1	0.5	0.9968	0.3408	-0.0032	0.993	1.0322	0.3556	0.0322	0.928
BYA	1.1	0.01	1.0974	0.0103	0.0974	<0.001	1.0965	0.0166	0.0965	<0.001
	1.1	0.05	1.0917	0.0495	0.0917	0.074	1.0884	0.0747	0.0884	0.246
	1.1	0.1	1.0862	0.0941	0.0862	0.367	1.0823	0.1328	0.0823	0.540
	1.1	0.5	1.0640	0.3419	0.0640	0.853	1.0759	0.3581	0.0759	0.834
	1.5	0.01	1.4935	0.0105	0.4935	<0.001	1.4865	0.0168	0.4865	<0.001
	1.5	0.05	1.4728	0.0504	0.4728	<0.001	1.4429	0.0763	0.4429	<0.001
	1.5	0.1	1.4502	0.0958	0.4502	<0.001	1.4005	0.1365	0.4005	0.006
	1.5	0.5	1.3338	0.3476	0.3338	0.345	1.2536	0.3698	0.2536	0.498
	2	0.01	1.9888	0.0109	0.9888	<0.001	1.9762	0.0172	0.9762	<0.001
	2	0.05	1.9502	0.0520	0.9502	<0.001	1.8940	0.0794	0.8940	<0.001
	2	0.1	1.9071	0.0989	0.9071	<0.001	1.8094	0.1436	0.8094	<0.001
	2	0.5	1.6741	0.3575	0.6741	0.069	1.4845	0.3894	0.4845	0.223



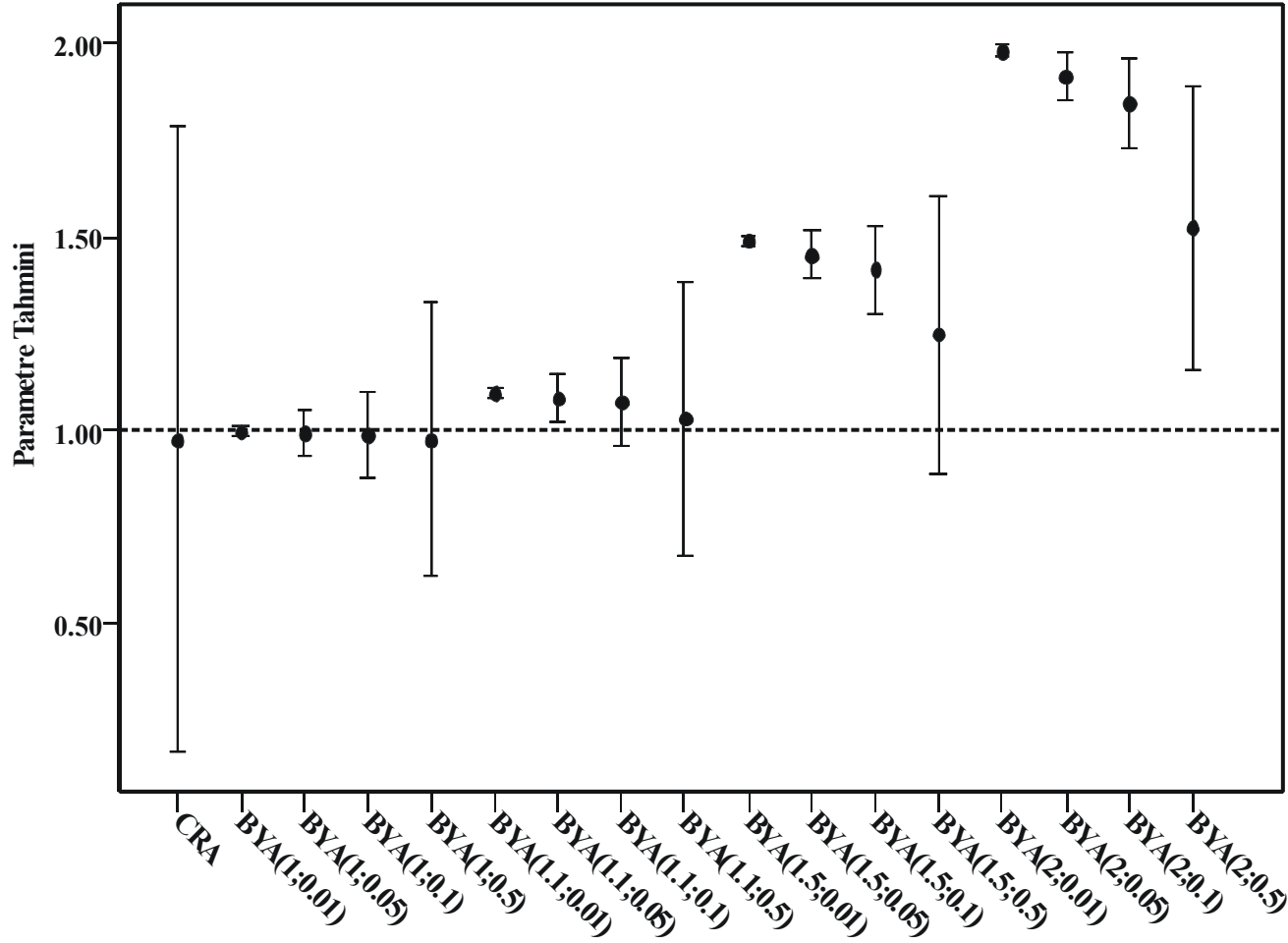
Şekil 4.13. Uniform dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=30$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri ve standart hatalarının grafiksel gösterimi



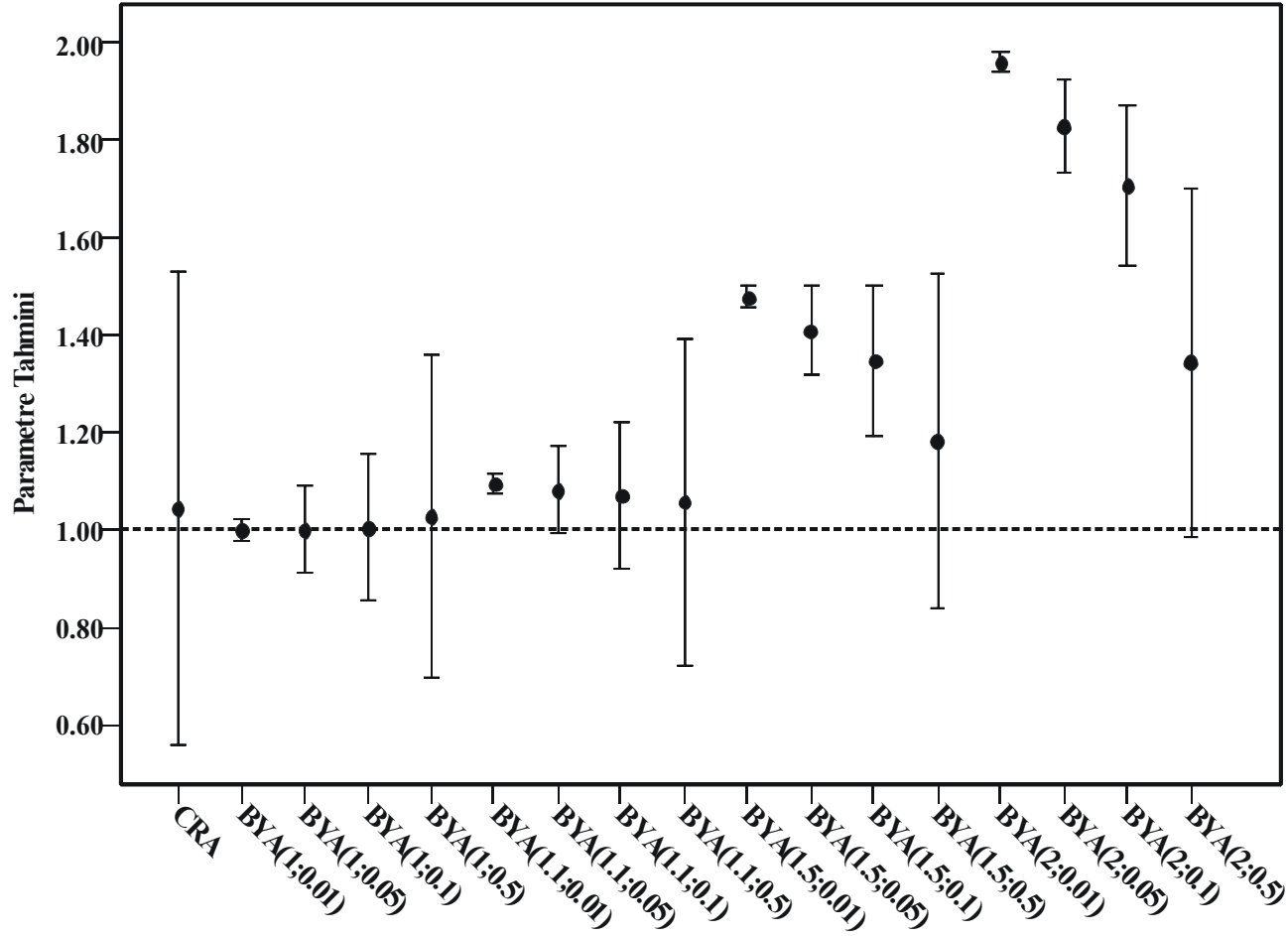
Şekil 4.14. Binom dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=30$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri ve standart hatalarının grafiksel gösterimi

Tablo 4.5. Uniform ve binom dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=50$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri, standart hataları ve yanlılıkları

Yöntem	Prior Bilgi		Bağımsız Değişkenin Dağılımı							
	$\bar{\beta}_p$	σ_p^2	Uniform				Binom			
			$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$	Yanlılık	p	$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$	Yanlılık	p
CRA	-	-	0.9771	0.8083	-0.0229	0.978	1.0446	0.4878	0.0446	0.928
	1	0.01	0.9977	0.0129	-0.0023	0.859	0.9992	0.0214	-0.0008	0.970
	1	0.05	0.9929	0.0603	-0.0071	0.907	1.0008	0.0899	0.0008	0.993
	1	0.1	0.9887	0.1116	-0.0113	0.920	1.0043	0.1503	0.0043	0.977
	1	0.5	0.9764	0.3537	-0.0236	0.947	1.0275	0.3328	0.0275	0.934
BYA	1.1	0.01	1.0960	0.0129	0.0960	<0.001	1.0947	0.0214	0.0947	<0.001
	1.1	0.05	1.0850	0.0604	0.0850	0.166	1.0817	0.0902	0.0817	0.369
	1.1	0.1	1.0741	0.1119	0.0741	0.511	1.0723	0.1511	0.0723	0.634
	1.1	0.5	1.0308	0.3546	0.0308	0.931	1.0582	0.3349	0.0582	0.863
	1.5	0.01	1.4893	0.0131	0.4893	<0.001	1.4779	0.0214	0.4779	<0.001
	1.5	0.05	1.4538	0.0613	0.4538	<0.001	1.4093	0.0919	0.4093	<0.001
	1.5	0.1	1.4165	0.1136	0.4165	0.001	1.3491	0.1555	0.3491	0.029
	1.5	0.5	1.2492	0.3592	0.2492	0.491	1.1830	0.3444	0.1830	0.598
	2	0.01	1.9814	0.0135	0.9814	<0.001	1.9606	0.0217	0.9606	<0.001
	2	0.05	1.9165	0.0629	0.9165	<0.001	1.8298	0.0956	0.8298	<0.001
	2	0.1	1.8469	0.1166	0.8469	<0.001	1.7081	0.1635	0.7081	<0.001
	2	0.5	1.5248	0.3670	0.5248	0.159	1.3444	0.3591	0.3444	0.342



Şekil 4.15. Uniform dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=50$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri ve standart hatalarının grafiksel gösterimi

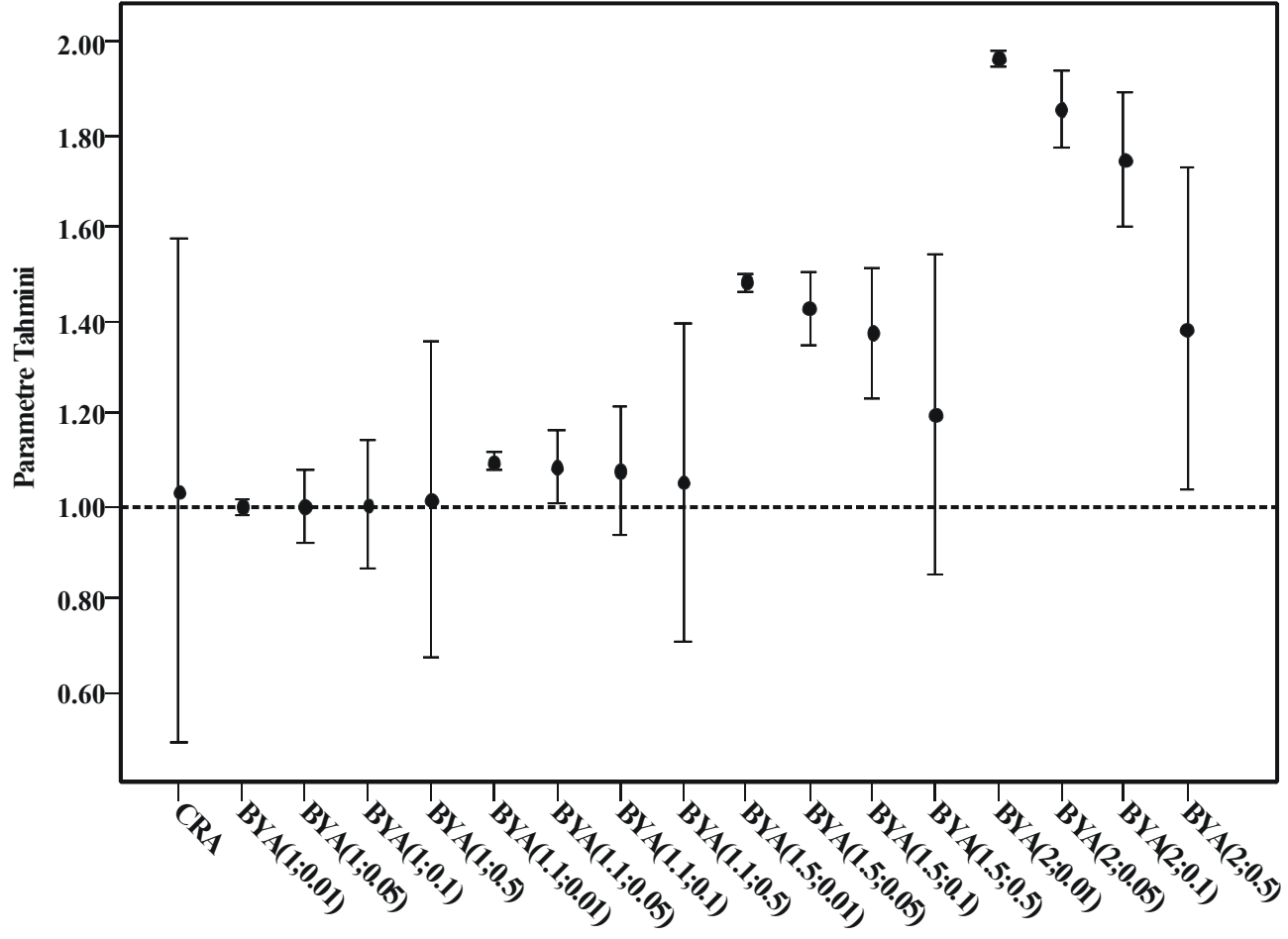


Şekil 4.16. Binom dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=50$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri ve standart hatalarının grafiksel gösterimi

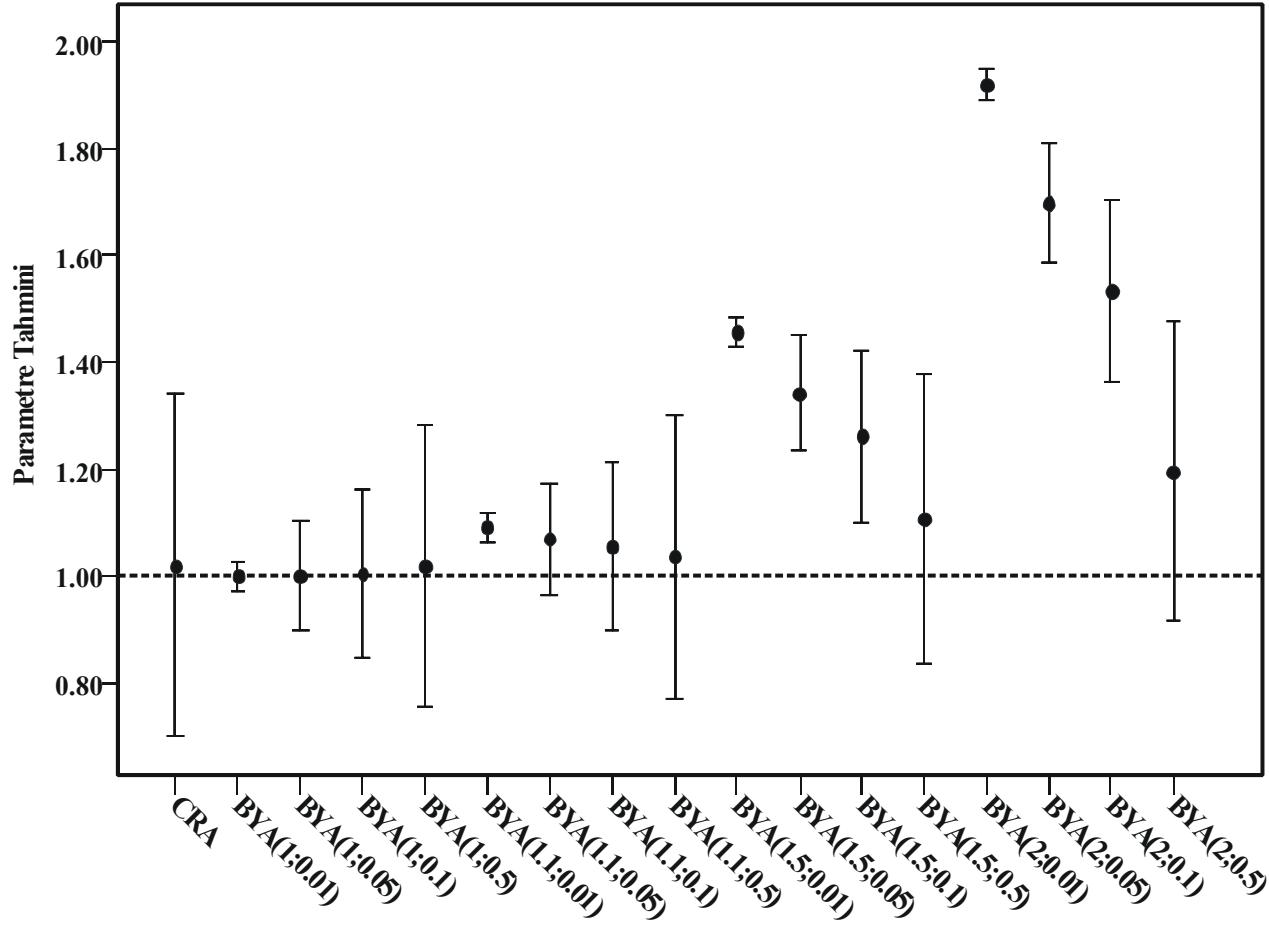
Tablo 4.6, Şekil 4.17 ve Şekil 4.18’de n=100 için CRA ve farklı prior bilgili BYA sonuçları incelendiğinde, veri yapısına uygun ve gerçeği yansıtan açıklayıcı prior bilgi olduğunda ($\bar{\beta}_p = 1$ ve $\sigma_p^2 = 0.01, 0.05, 0.01$ ve 0.5), BYA ve CRA’da yakınsamanın sağlandığı ($p > 0.05$) ve bunun yanı sıra BYA’nın CRA’ya göre simülasyon parametresini daha küçük standart hatayla tahmin ettiği bulundu. Ayrıca prior bilgiye ilişkin varyansın da BYA’nın performansı üzerinde önemli etkisinin olduğu ve prior bilginin parametresi sabit tutulup varyans artırıldığında yanlılığın arttığı gözlemlendi. BYA’da gerçeği yansıtmayan prior bilgi kullanıldığında ($\bar{\beta}_p = 1.1, 1.5, 2$), simülasyon parametresine daha uzak ve prior bilgi parametresine daha yakın bir tahmin değeri elde edildi. Gerçeği yansıtmayan prior bilginin varyansı artırıldığında parametre tahmininin, simülasyon parametresine biraz daha yakınsadığı ve yanlılığın azaldığı gözlemlendi. $\bar{\beta}_p = 1.1$ ve $\sigma_p^2 = 0.01$ için simülasyon parametresinden önemli düzeyde farklı tahmin yapıldığı ve iyi bir yakınsamanın sağlanmadığı belirlendi (uniform dağılım için $p < 0.001$ ve binom dağılım için $p < 0.01$). $\bar{\beta}_p = 1.1$ için varyans değeri arttıkça önemli düzeyde yakınsama sağlandığı belirlendi ($p > 0.05$). Uniform dağılım için $\bar{\beta}_p = 1.5$ ve $\sigma_p^2 = 0.01, 0.05$ ve 0.1 prior bilgileri ele alındığında simülasyon parametresinden önemli düzeyde farklı tahmin yapıldığı ve yakınsamanın sağlanmadığı (sırasıyla $p < 0.001, p < 0.001$ ve $p < 0.01$) ve en büyük varyans değeri ele alındığında ($\sigma_p^2 = 0.5$) önemli düzeyde yakınsamanın sağlandığı bulundu ($p > 0.05$). Binom dağılım için $\bar{\beta}_p = 1.5$ ve $\sigma_p^2 = 0.01$ ve 0.05 prior bilgileri ele alındığında simülasyon parametresinden önemli düzeyde farklı tahmin yapıldığı ve yakınsamanın sağlanmadığı (sırasıyla $p < 0.001$ ve $p < 0.01$) ve $\sigma_p^2 = 0.1$ ve 0.5 varyans değerleri ele alındığında önemli düzeyde yakınsamanın sağlandığı bulundu ($p > 0.05$). Uniform ve binom dağılım için $\bar{\beta}_p = 2$ ve $\sigma_p^2 = 0.01, 0.05$ ve 0.1 değerleri ele alındığında simülasyon parametresinden önemli düzeyde farklı tahmin yapıldığı ve yakınsamanın sağlanmadığı ($p < 0.001$ ve $p < 0.01$) ve en büyük varyans değeri ele alındığında ise ($\sigma_p^2 = 0.5$) önemli düzeyde yakınsama sağlandığı bulundu ($p > 0.05$).

Tablo 4.6. Uniform ve binom dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=100$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri, standart hataları ve yanlışlıkları

Yöntem	Prior Bilgi		Bağımsız Değişkenin Dağılımı							
	$\bar{\beta}_p$	σ_p^2	Uniform				Binom			
			$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$	Yanlılık	p	$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$	Yanlılık	p
CRA	-	-	1.0321	0.5435	0.0321	0.953	1.0195	0.3212	0.0195	0.952
	1	0.01	0.9993	0.0181	-0.0007	0.969	0.9993	0.0281	-0.0007	0.980
	1	0.05	1.0003	0.0795	0.0007	0.997	1.0014	0.1035	0.0014	0.989
	1	0.1	1.0022	0.1383	0.0022	0.987	1.0048	0.1560	0.0048	0.976
	1	0.5	1.0137	0.3412	0.0137	0.968	1.0183	0.2663	0.0183	0.945
	1.1	0.01	1.0958	0.0181	0.0958	<0.001	1.0904	0.0282	0.0904	0.002
	1.1	0.05	1.0853	0.0797	0.0853	0.287	1.0688	0.1040	0.0688	0.510
	1.1	0.1	1.0762	0.1386	0.0762	0.584	1.0557	0.1570	0.0557	0.724
	1.1	0.5	1.0503	0.3418	0.0503	0.883	1.0357	0.2674	0.0357	0.894
	BYA	1.5	0.01	1.4824	0.0184	0.4824	<0.001	1.4575	0.0284	0.4575
1.5		0.05	1.4261	0.0806	0.4261	<0.001	1.3432	0.1072	0.3432	0.002
1.5		0.1	1.3732	0.1403	0.3732	0.009	1.2636	0.1621	0.2636	0.107
1.5		0.5	1.1972	0.3447	0.1972	0.569	1.1064	0.2722	0.1064	0.697
2		0.01	1.9663	0.0187	0.9663	<0.001	1.9227	0.0294	0.9227	<0.001
2		0.05	1.8545	0.0825	0.8545	<0.001	1.6990	0.1130	0.6990	<0.001
2		0.1	1.7474	0.1435	0.7474	<0.001	1.5340	0.1711	0.5340	0.002
2		0.5	1.3823	0.3496	0.3823	0.277	1.1966	0.2790	0.1966	0.483



Şekil 4.17. Uniform dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=100$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri ve standart hatalarının grafiksel gösterimi

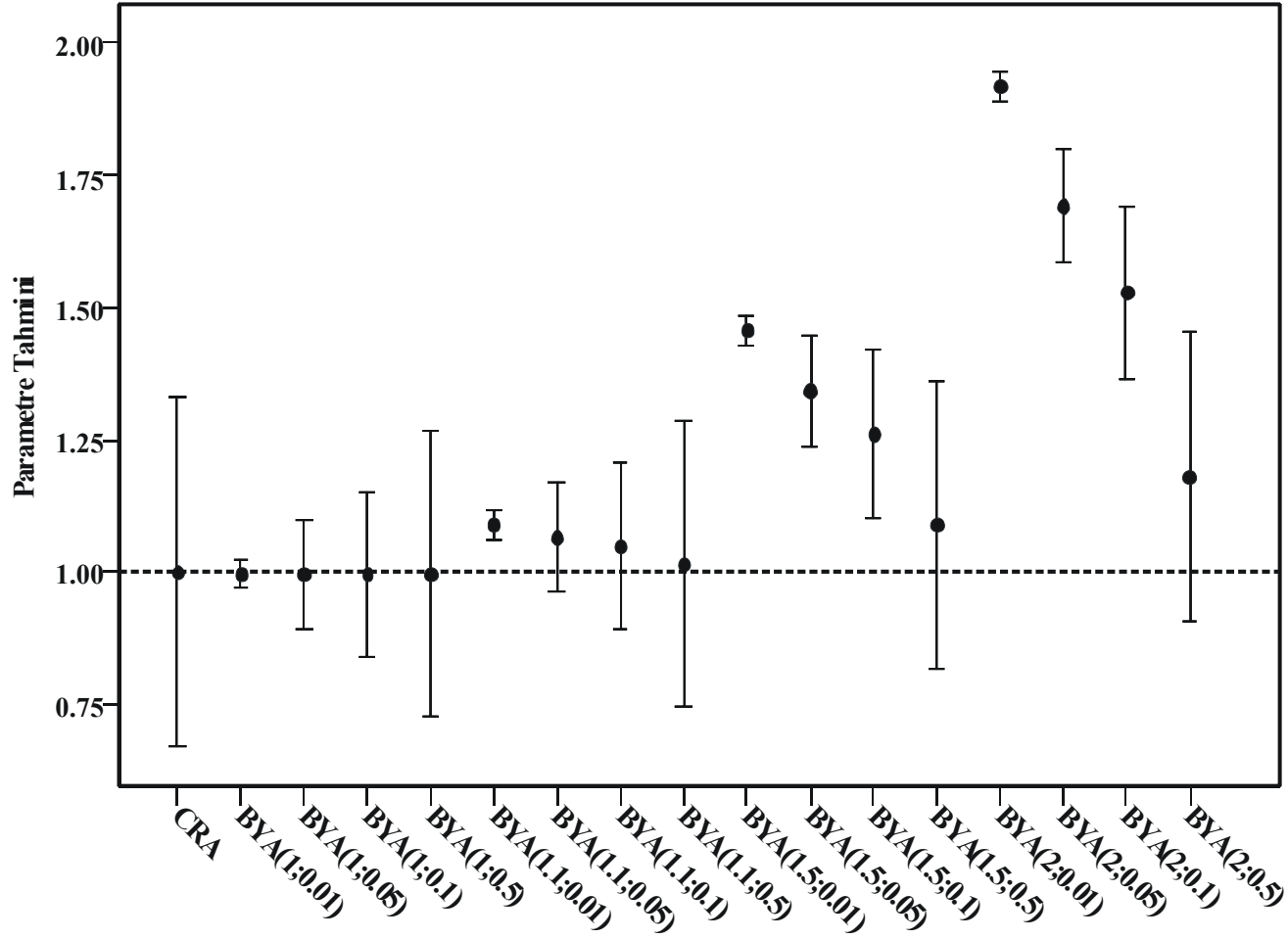


Şekil 4.18. Binom dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=100$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri ve standart hatalarının grafiksel gösterimi

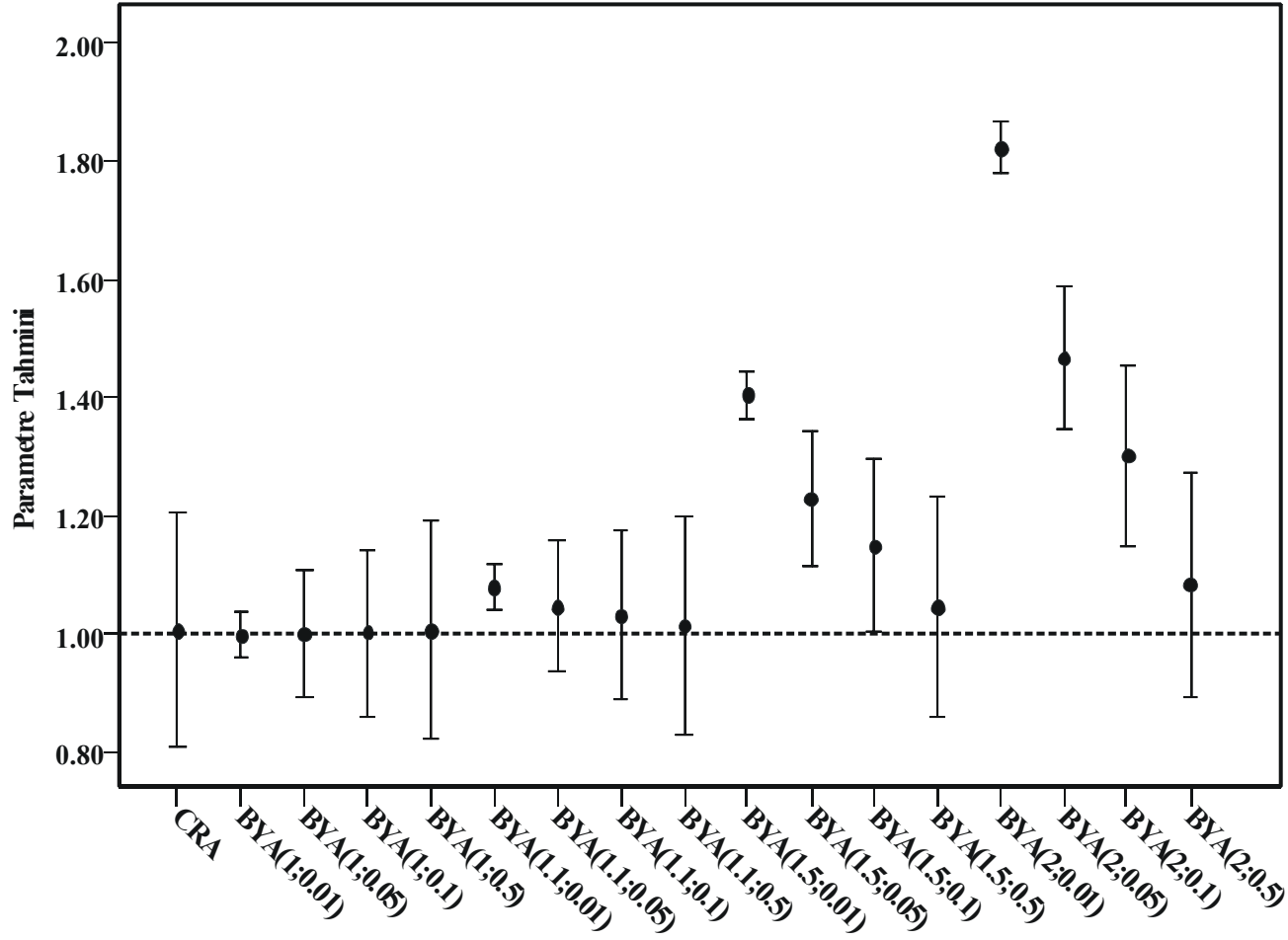
Tablo 4.7, Şekil 4.19 ve Şekil 4.20’de $n=250$ için CRA ve farklı prior bilgili BYA sonuçları incelendiğinde, veri yapısına uygun ve gerçeği yansıtan açıklayıcı prior bilgi olduğunda ($\bar{\beta}_p = 1$ ve $\sigma_p^2=0.01, 0.05, 0.01$ ve 0.5), her iki yöntemde de yakınsamanın sağlandığı ($p>0.05$) ve bunun yanı sıra BYA’nın CRA’ya göre simülasyon parametresini daha küçük standart hatayla tahmin ettiği bulundu. BYA’da gerçeği yansıtmayan prior bilgi kullanıldığında ($\bar{\beta}_p = 1.1, 1.5, 2$), simülasyon parametresine daha uzak ve prior bilgi parametresine daha yakın bir tahmin değeri elde edildi. Gerçeği yansıtmayan prior bilginin varyansı artırıldığında parametre tahmininin, simülasyon parametresine biraz daha yakınsadığı ve yanlılığın azaldığı gözlemlendi. $\bar{\beta}_p = 1.1$ ve $\sigma_p^2=0.01$ için simülasyon parametresinden önemli düzeyde farklı tahmin yapıldığı ve iyi bir yakınsamanın sağlanmadığı belirlendi (uniform dağılım için $p<0.01$ ve binom dağılım için $p<0.05$). $\bar{\beta}_p = 1.1$ için varyans değeri arttıkça ($\sigma_p^2=0.05, 0.1$ ve 0.5) önemli düzeyde yakınsama sağlandığı belirlendi ($p>0.05$). $\bar{\beta}_p = 1.5$ ve $\sigma_p^2=0.01, 0.05$ prior bilgileri ele alındığında simülasyon parametresinden önemli düzeyde farklı tahmin yapıldığı ve yakınsamanın sağlanmadığı (uniform dağılım için sırasıyla $p<0.001, p<0.01$ ve binom dağılım için sırasıyla $p<0.001, p<0.05$) ve varyans değeri $\sigma_p^2=0.1$ ve 0.5 değerlerine ulaştığında önemli düzeyde yakınsamanın sağlandığı bulundu ($p>0.05$). Uniform ve binom dağılım için $\bar{\beta}_p = 2$ ve $\sigma_p^2=0.01, 0.05$ ve 0.1 değerleri ele alındığında simülasyon parametresinden önemli düzeyde farklı tahmin yapıldığı ve yakınsamanın sağlanmadığı (uniform dağılım için sırasıyla $p<0.001, p<0.001, p<0.01$ ve binom dağılım için sırasıyla $p<0.001, p<0.001, p<0.05$) ve en büyük varyans değeri ele alındığında ise ($\sigma_p^2=0.5$) önemli düzeyde yakınsama sağlandığı bulundu ($p>0.05$). Ayrıca her koşulda BYA’nın CRA’ya göre standart hatası daha küçük tahminler yaptığı ve sonuçlara genel olarak bakıldığında $n=250$ örneklem büyüklüğü için her iki yöntemde yanlılıklarının daha küçük olduğu bulundu.

Tablo 4.7. Uniform ve binom dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=250$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri, standart hataları ve yanlışlıkları

Yöntem	Prior Bilgi		Bağımsız Değişkenin Dağılımı							
	$\bar{\beta}_p$	σ_p^2	Uniform				Binom			
			$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$	Yanlılık	p	$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$	Yanlılık	p
CRA	-	-	1.0007	0.3308	0.0007	0.998	1.0069	0.1982	0.0069	0.972
	1	0.01	0.9982	0.0272	-0.0018	0.947	0.9993	0.0394	-0.0007	0.986
	1	0.05	0.9962	0.1022	-0.0038	0.970	1.0014	0.1095	0.0014	0.990
	1	0.1	0.9957	0.1560	-0.0043	0.978	1.0035	0.1411	0.0035	0.980
	1	0.5	0.9972	0.2703	-0.0028	0.992	1.0079	0.1842	0.0079	0.966
BYA	1.1	0.01	1.0899	0.0273	0.0899	0.001	1.0796	0.0396	0.0796	0.045
	1.1	0.05	1.0653	0.1026	0.0653	0.525	1.0463	0.1102	0.0463	0.675
	1.1	0.1	1.0485	0.1565	0.0485	0.757	1.0325	0.1419	0.0325	0.819
	1.1	0.5	1.0155	0.2707	0.0155	0.954	1.0155	0.1846	0.0155	0.933
	1.5	0.01	1.4575	0.0279	0.4575	<0.001	1.4050	0.0406	0.4050	<0.001
	1.5	0.05	1.3427	0.1044	0.3427	0.001	1.2296	0.1140	0.2296	0.045
	1.5	0.1	1.2607	0.1589	0.2607	0.102	1.1505	0.1456	0.1505	0.302
	1.5	0.5	1.0892	0.2724	0.0892	0.744	1.0460	0.1861	0.0460	0.805
	2	0.01	1.9186	0.0287	0.9186	<0.001	1.8229	0.0427	0.8229	<0.001
	2	0.05	1.6927	0.1075	0.6927	<0.001	1.4676	0.1204	0.4676	<0.001
	2	0.1	1.5288	0.1628	0.5288	0.001	1.3023	0.1516	0.3023	0.047
	2	0.5	1.1818	0.2749	0.1818	0.509	1.0844	0.1881	0.0844	0.654



Şekil 4.19. Uniform dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=250$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri ve standart hatalarının grafiksel gösterimi

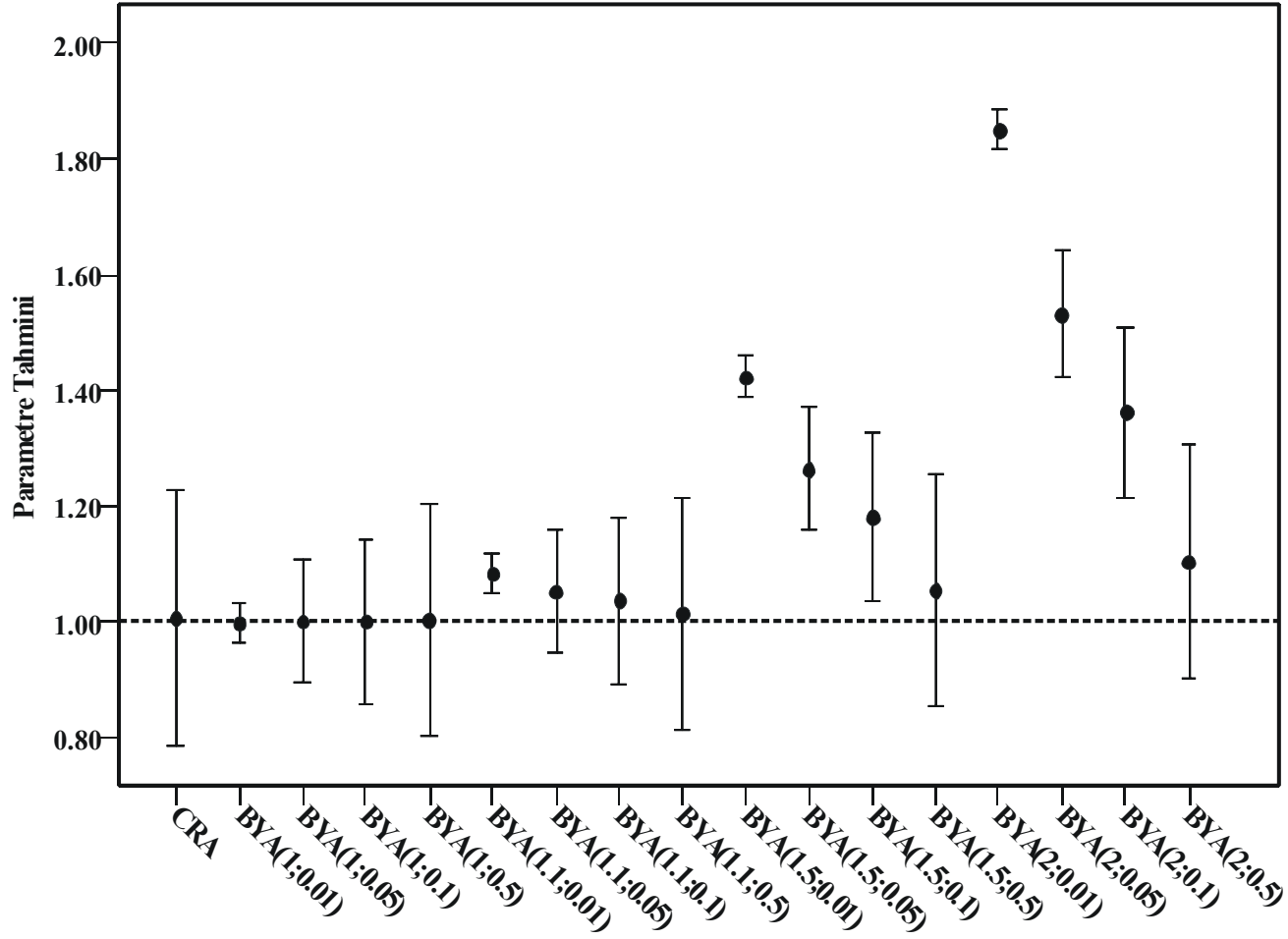


Şekil 4.20. Binom dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=250$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri ve standart hatalarının grafiksel gösterimi

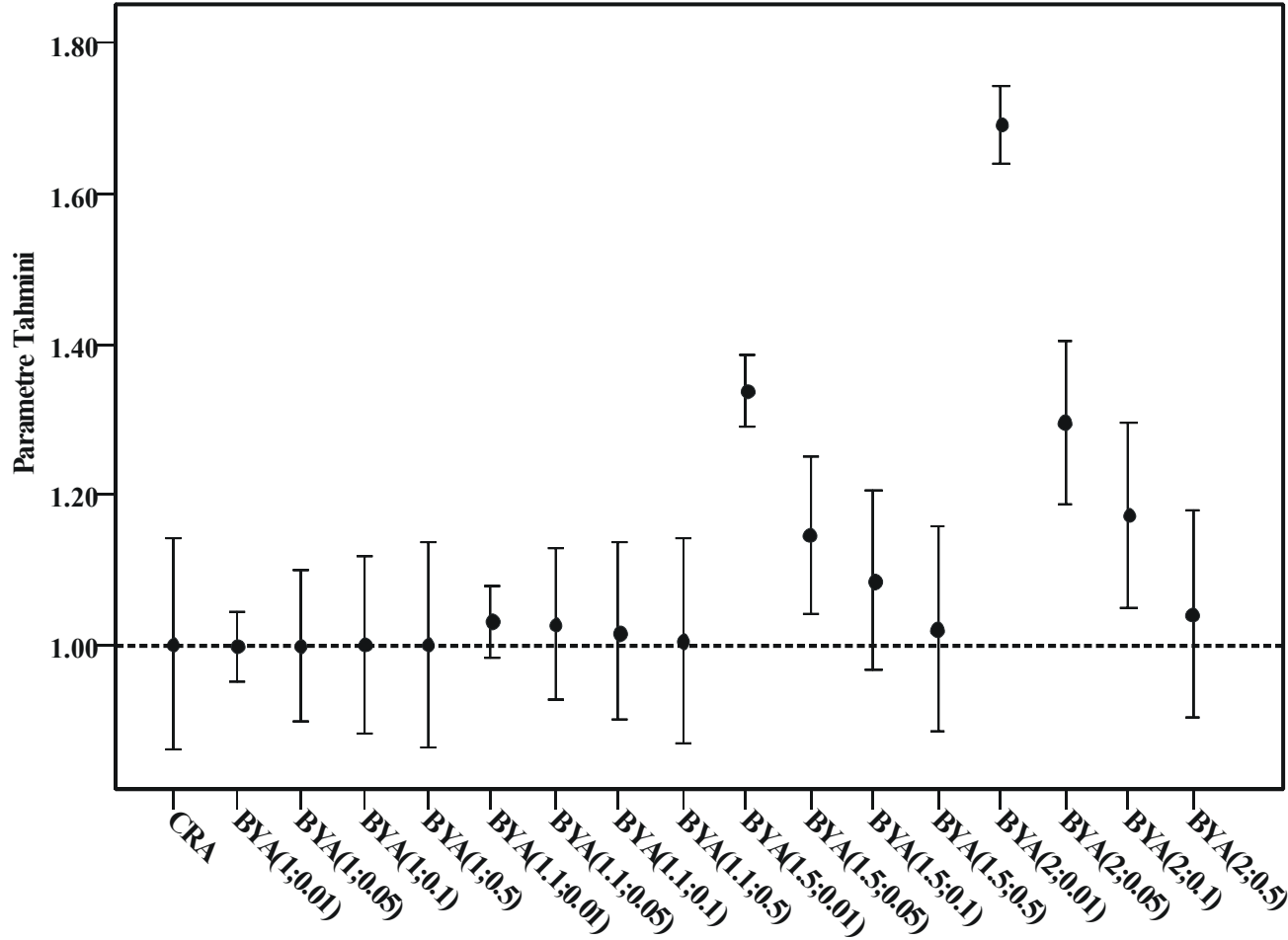
Tablo 4.8, Şekil 4.21 ve Şekil 4.22’de n=500 için CRA ve farklı prior bilgili BYA sonuçları incelendiğinde, veri yapısına uygun ve gerçeği yansıtan açıklayıcı prior bilgi olduğunda ($\bar{\beta}_p = 1$ ve $\sigma_p^2 = 0.01, 0.05, 0.01$ ve 0.5), yöntemlerde yakınsamanın sağlandığı ($p > 0.05$) ve bunun yanı sıra BYA’nın CRA’ya göre simülasyon parametresini daha küçük standart hatayla tahmin ettiği bulundu. BYA’da gerçeği yansıtmayan prior bilgi kullanıldığında ($\bar{\beta}_p = 1.1, 1.5, 2$), simülasyon parametresine daha uzak ve prior bilgi parametresine daha yakın bir tahmin değeri elde edildi. Gerçeği yansıtmayan prior bilginin varyansı artırıldığında parametre tahmininin, simülasyon parametresine biraz daha yakınsadığı ve yanlılığın azaldığı gözlemlendi. $\bar{\beta}_p = 1.1$ ve $\sigma_p^2 = 0.01$ prior bilgileri kullanıldığında, uniform dağılım için simülasyon parametresinden önemli düzeyde farklı tahmin yapıldığı ve iyi bir yakınsamanın sağlanmadığı ($p < 0.01$) ve binom dağılım için önemli derecede yakınsamanın sağlandığı ($p > 0.05$) belirlendi. Her iki dağılımda da $\bar{\beta}_p = 1.1$ ve $\sigma_p^2 = 0.05, 0.1, 0.5$ değerleri için önemli düzeyde yakınsamanın sağlandığı belirlendi ($p > 0.05$). Uniform dağılımda $\bar{\beta}_p = 1.5$ ve $\sigma_p^2 = 0.01, 0.05$ prior bilgileri ele alındığında simülasyon parametresinden önemli düzeyde farklı tahmin yapıldığı ve yakınsamanın sağlanmadığı (sırasıyla $p < 0.001, p < 0.01$) ve binom dağılımda ise sadece $\bar{\beta}_p = 1.5$ ve $\sigma_p^2 = 0.01$ prior bilgileri için yakınsamanın sağlanmadığı ($p < 0.001$) belirlendi. Uniform dağılım için $\bar{\beta}_p = 2$ ve $\sigma_p^2 = 0.01, 0.05$ ve 0.1 değerleri ele alındığında simülasyon parametresinden önemli düzeyde farklı tahmin yapıldığı ve yakınsamanın sağlanmadığı (sırasıyla $p < 0.001, p < 0.001, p < 0.05$) ve binom dağılım için sadece $\bar{\beta}_p = 2$ ve $\sigma_p^2 = 0.01, 0.05$ değerleri ele alındığında yakınsamanın sağlanmadığı (sırasıyla $p < 0.001, p < 0.01$) bulundu. Uniform dağılım için $\bar{\beta}_p = 2$ ve $\sigma_p^2 = 0.5$ ile binom dağılım için $\bar{\beta}_p = 2$ ve $\sigma_p^2 = 0.1, 0.5$ prior bilgileri ele alındığında ise önemli düzeyde yakınsama sağlandığı bulundu ($p > 0.05$). Ayrıca yine her koşulda BYA’nın CRA’ya göre standart hatası daha küçük tahminler yaptığı bulundu. Sonuçlara genel olarak bakıldığında örneklem büyüklüğü büyük olduğunda yanlılıkların, gerçeği yansıtan prior bilgi ve özellikle gerçeği yansıtmayan prior bilgiyle çalışıldığında daha az olduğu gözlemlendi.

Tablo 4.8. Uniform ve binom dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=500$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri, standart hataları ve yanlışlıkları

Yöntem	Prior Bilgi		Bağımsız Değişkenin Dağılımı							
	$\bar{\beta}_p$	σ_p^2	Uniform				Binom			
			$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$	Yanlılık	p	$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$	Yanlılık	p
CRA	-	-	1.0069	0.2221	0.0069	0.975	1.0031	0.1411	0.0031	0.982
	1	0.01	0.9993	0.0338	-0.0007	0.983	0.9994	0.0468	-0.0006	0.990
	1	0.05	1.0003	0.1049	0.0003	0.998	1.0010	0.1006	0.0010	0.992
	1	0.1	1.0014	0.1424	0.0014	0.992	1.0020	0.1176	0.0020	0.986
	1	0.5	1.0040	0.1999	0.0040	0.984	1.0034	0.1361	0.0034	0.980
BYA	1.1	0.01	1.0840	0.0340	0.0840	0.014	1.0332	0.0471	0.0332	0.481
	1.1	0.05	1.0529	0.1053	0.0529	0.616	1.0297	0.1012	0.0297	0.769
	1.1	0.1	1.0371	0.1429	0.0371	0.795	1.0188	0.1180	0.0188	0.873
	1.1	0.5	1.0140	0.2001	0.0140	0.944	1.0073	0.1362	0.0073	0.957
	1.5	0.01	1.4238	0.0348	0.4238	<0.001	1.3387	0.0488	0.3387	<0.001
	1.5	0.05	1.2645	0.1072	0.2645	0.014	1.1467	0.1038	0.1467	0.158
	1.5	0.1	1.1807	0.1448	0.1807	0.213	1.0867	0.1201	0.0867	0.471
	1.5	0.5	1.0542	0.2009	0.0542	0.787	1.0229	0.1368	0.0229	0.867
	2	0.01	1.8510	0.0360	0.8510	<0.001	1.6921	0.0517	0.6921	<0.001
	2	0.05	1.5317	0.1103	0.5317	<0.001	1.2970	0.1081	0.2970	0.006
	2	0.1	1.3619	0.1476	0.3619	0.015	1.1733	0.1230	0.1733	0.159
	2	0.5	1.1046	0.2020	0.1046	0.605	1.0425	0.1376	0.0425	0.758



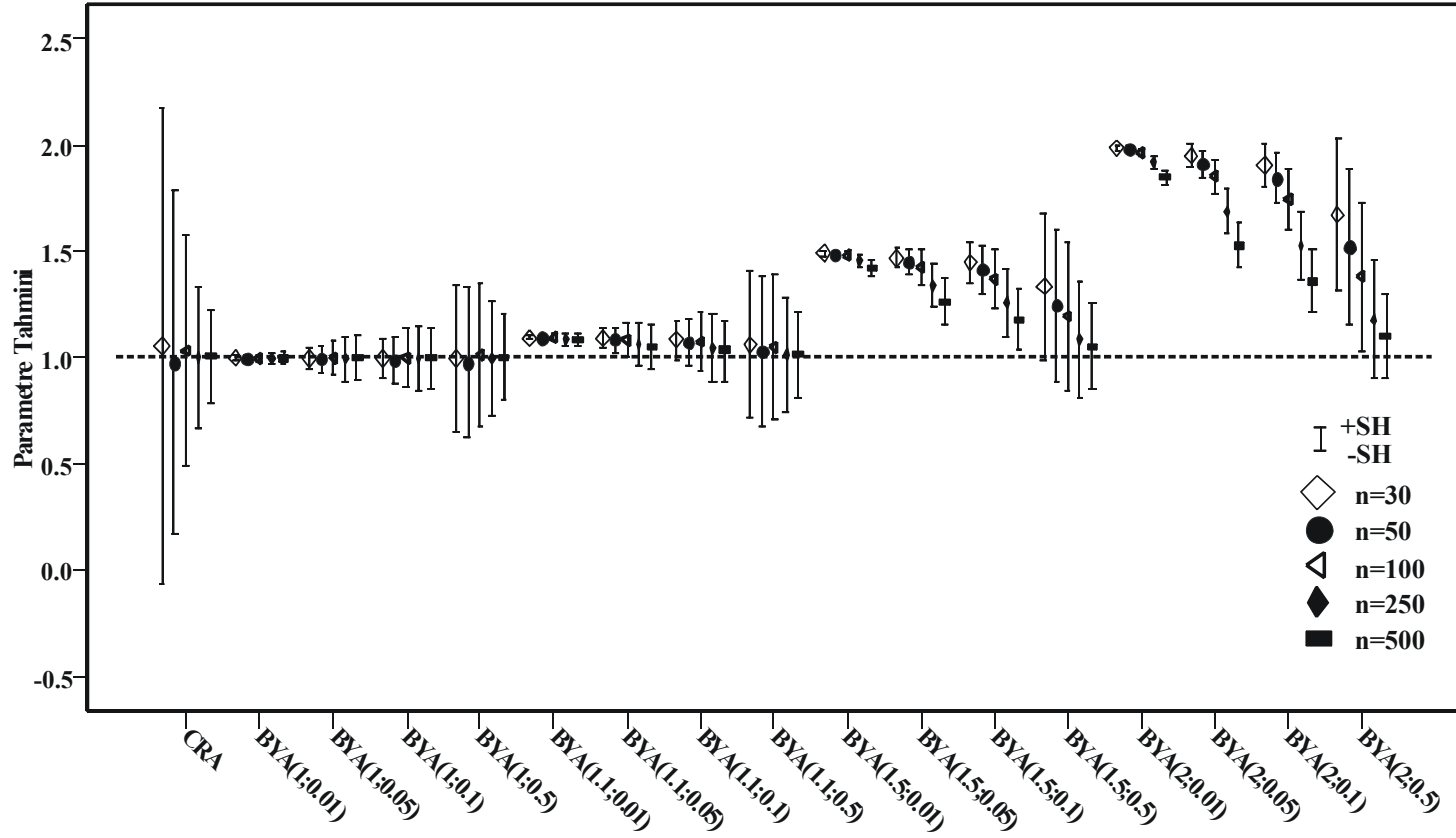
Şekil 4.21. Uniform dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=500$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri ve standart hatalarının grafiksel gösterimi



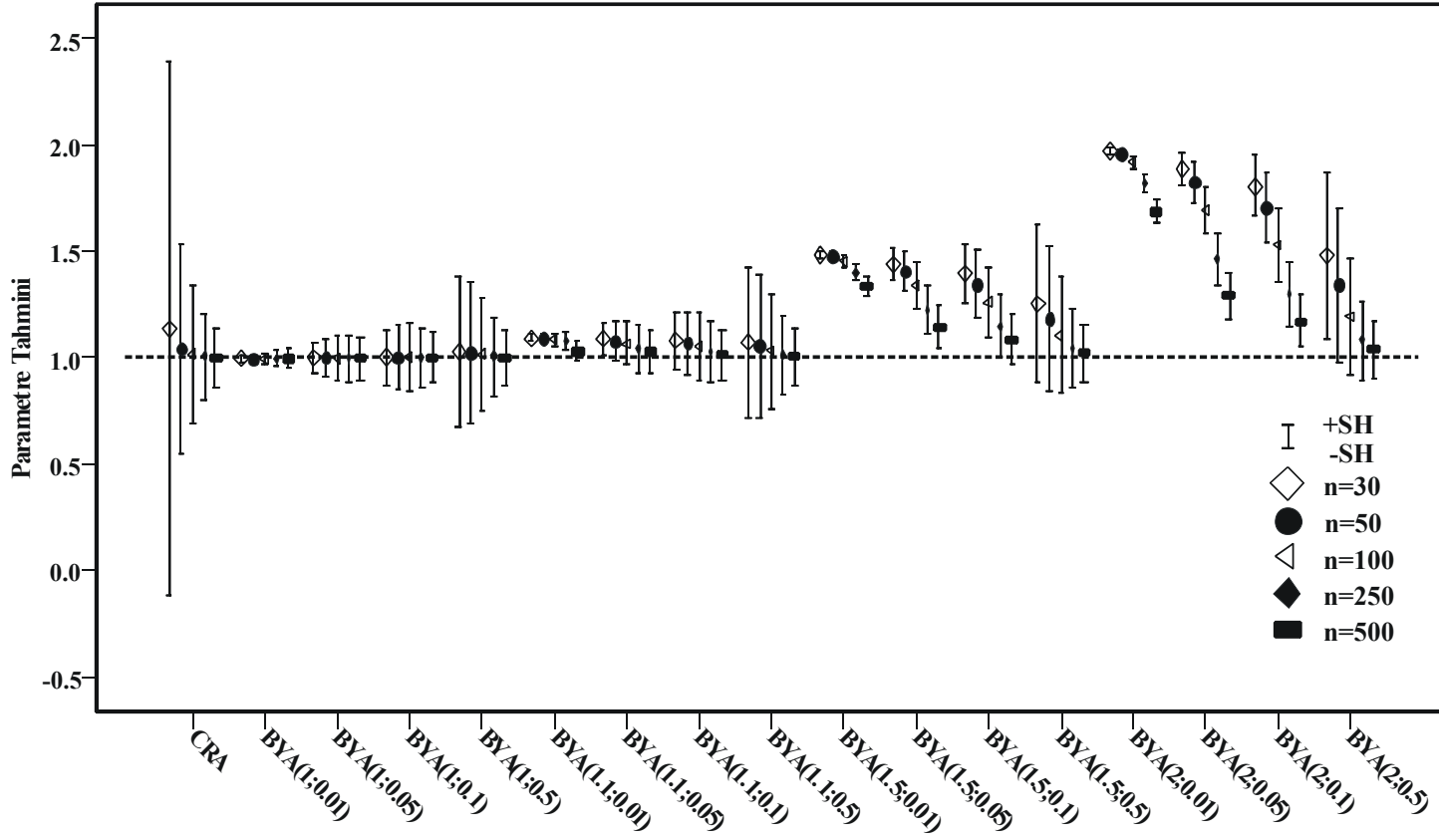
Şekil 4.22. Binom dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=500$ için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri ve standart hatalarının grafiksel gösterimi

Şekil 4.23’de $n=30, 50, 100, 250$ ve 500 örneklem büyüklükleri ve uniform dağılılan bağımsız değişkene ilişkin CRA ve farklı prior bilgili Bayesgil yaklaşımdan elde edilen parametre tahmini ve standart hatalarının birlikte grafiksel gösterimi verildi. CRA’da, örneklem büyüklüğü arttıkça standart hatası daha düşük parametre tahminleri elde edildi. BYA’da veriye uygun prior bilgi olduğunda ($\bar{\beta}_p = 1$) örneklem büyüklüğü arttıkça, simülasyon parametresine CRA’ya göre oldukça iyi bir yakınsama gerçekleşti. Özellikle de prior bilginin varyansı küçük olduğunda yanlılığın azaldığı gözlemlendi. BYA’da gerçeğe uygun olmayan prior bilgi olduğunda ($\bar{\beta}_p = 1.1, 1.5, 2$) farklı sonuçlarla karşılaşıldı. Veri yapısına uygun ve gerçeği yansıtan prior bilgiden uzaklaşıldıkça, simülasyon parametresine yakınsamanın güçleştiği gözlemlendi. Buna karşın örneklem büyüklüğü ve prior bilginin varyansı arttıkça, uygun olmayan prior bilgiye rağmen yanlılığın azaldığı belirlendi.

Şekil 4.24’de farklı örneklem büyüklükleri ve binom dağılılan bağımsız değişkene ilişkin CRA ve açıklayıcı prior bilgili BYA yaklaşımından elde edilen parametre tahmini ve standart hatalarının grafiksel gösterimi verilmiştir. Uniform dağılımda olduğu gibi CRA’da örneklem büyüklüğü arttıkça, standart hatası daha düşük parametre tahminleri elde edildi. Veriye uygun prior bilgi olduğunda ($\bar{\beta}_p = 1$) örneklem büyüklüğü arttıkça, simülasyon parametresine BYA’da CRA’ya göre oldukça iyi bir yakınsama gerçekleşti. Özellikle de prior bilginin varyansı küçük olduğunda yanlılığın azaldığı gözlemlendi. BYA’da gerçeğe uygun olmayan prior bilgi olduğunda ($\bar{\beta}_p = 1.1, 1.5, 2$), uniform dağılım gösteren verilerden elde edilen sonuçlara benzer sonuçlar elde edildi. Veri yapısına uygun ve gerçeği yansıtan prior bilgiden uzaklaşıldıkça, simülasyon parametresine yakınsamanın güçleştiği gözlemlendi. Buna karşın örneklem büyüklüğü ve prior bilginin varyansı arttıkça, uygun olmayan prior bilgiye rağmen yanlılığın azaldığı belirlendi. Ayrıca binom dağılımından elde edilen parametre tahminlerinin yanlılığının, uniform dağılımdan elde edilen parametre tahminlerinin yanlılığından daha küçük olduğu bulundu.



Şekil 4.23. Uniform dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=30, 50, 100, 250$ ve 500 için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri ve standart hatalarının grafiksel gösterimi



Şekil 4.24. Binom dağılımlı bağımsız değişkene göre $\beta=1$ ve $n=30, 50, 100, 250$ ve 500 için farklı prior bilgi koşullarında BYA ve CRA'nın 1000 Monte Carlo simülasyonundan elde edilen parametre tahminleri ve standart hatalarının grafiksel gösterimi

4.2. Gerçek Uygulama Sonuçları

Meme kanserli hastaların yinelemesiz yaşam süresine ilişkin verinin 303'ü (%71.63) tamamlanmamış gözlemlerden oluşmaktadır. Meme kanserinde yinelemeyi etkileyen risk faktörleri olarak; yaş, menopoz durumu, ilk doğum yaşı, aile hikayesi, tümör kadranı, tümör büyüklüğü, progesteron reseptör durumu, östrojen reseptör durumu, aksiller lenf nodu tutulumu, kapsül dışı lenf nodu tutulumu, histolojik tip ve tedavi (radyoterapi ve hormon terapisi) değişkenleri alınarak yapılan analiz sonuçları Tablo 4.9, 4.10 ve 4.11'de verildi. Tablo 4.15, 4.16 ve 4.17'de ise aşamalı olarak yapılan analiz sonuçları verilmiştir.

Tablo 4.9'da, CRA sonuçları verildi. Bu sonuçlar incelendiğinde, meme kanserli hastaların yinelemesiz yaşam süresi üzerinde yaş, ilk doğum yaşı, menopoz durumu, aile hikayesi, tümör kadranı, progesteron reseptör durumu, östrojen reseptör durumu, kapsül dışı lenf nodu tutulumu, histolojik tip ve radyoterapi değişkenlerinin önemli etkilerinin olmadığı bulundu ($p>0.05$).

Tümör büyüklüğü ($p<0.001$), hormon terapisi ($p<0.05$) ve aksiller lenf nodu tutulumu ($p<0.01$) değişkenlerinin yinelemesiz yaşam süresi üzerinde önemli etkilerinin olduğu belirlendi (Tablo 4.9).

CRA sonuçlarına göre; tümör büyüklüğü arttıkça yineleme riskinin 1.147 kat arttığı, hormon terapisi alanların almayanlara göre 0.508 oranında yineleme riski taşıdığı ve *pozitif lenf nodu* ≥ 4 olanların *negatif* olanlara göre 2.501 kat yineleme riski taşıdığı bulundu (Tablo 4.9).

Tablo 4.9. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin CRA sonuçları

Bağımsız Değişkenler		$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$	Wald Ki-kare	p	HO
Yaş (yıl)		0.0151	0.0122	1.5263	0.217	1.015
Tümör Büyüklüğü (cm)		0.1369	0.0371	13.6235	<0.001	1.147
İlk Doğum Yaşı (yıl)	≥30	0.4840	0.4134	1.3708	0.242	1.622
Menopoz Durumu	Pre+Peri Menopoz	-0.1746	0.3016	0.3350	0.563	0.840
Progesteron Reseptör Durumu	Negatif	0.3629	0.2946	1.5172	0.218	1.437
Östrojen Reseptör Durumu	Negatif	-0.5545	0.3582	2.3964	0.122	0.574
Radyoterapi	Var	-0.1928	0.2849	0.4579	0.499	0.825
Hormon Terapisi	Var	-0.6765	0.2947	5.2691	0.022	0.508
Aile Hikayesi	Meme Kanseri	0.5364	0.3234	2.7509	0.097	1.710
	Diğer Kanserler	0.2248	0.2226	1.0199	0.313	1.252
Aksiller Lenf Nodu Tutulumu	Pozitif Lenf Nodu 1-3	0.3024	0.3049	0.9838	0.321	1.353
	Pozitif Lenf Nodu ≥4	0.9168	0.3263	7.8970	0.005	2.501
Tümör Kadranı	Çok Merkezli	0.2490	0.2443	1.0385	0.308	1.283
Kapsül Dışı Lenf Nodu Tutulumu	Pozitif	0.3326	0.2434	1.8680	0.172	1.395
Histolojik tip	Duktal	0.2109	0.2818	0.5600	0.454	1.235

Ulaşılabilen açıklayıcı prior bilgiler kullanılan BYA sonuçları, Tablo 4.10'da verildi. Tablo 4.10 incelendiğinde, meme kanserli hastaların yinelemesiz yaşam süresi üzerinde yaş ($p<0.05$), tümör büyüklüğü ($p<0.001$) ve aksiller lenf nodu tutulumu ($p<0.001$) değişkenlerinin önemli etkilerinin olduğu, menopoz durumu, ilk doğum yaşı, aile hikayesi, tümör kadranı, progesteron reseptör durumu, östrojen reseptör durumu, kapsül dışı lenf nodu tutulumu, histolojik tip, hormon terapisi ve radyoterapi değişkenlerinin önemli etkilerinin olmadığı bulundu ($p>0.05$).

Ulaşılabilen açıklayıcı prior bilgiler kullanılan BYA sonuçlarına göre; yaş arttıkça yineleme riskinin 1.024 kat arttığı, tümör büyüklüğü arttıkça yineleme riskinin 1.186 kat arttığı, aksiller lenf nodu tutulumu *pozitif lenf nodu 1-3* olanların *negatif* olanlara göre yineleme riskinin 1.847 kat arttığı ve *pozitif lenf nodu ≥ 4* olanların *negatif* olanlara göre yineleme riskinin 3.189 kat arttığı bulundu (Tablo 4.10).

Açıklayıcı olmayan prior bilgili BYA sonuçları, Tablo 4.11'de verildi. Tablo 4.11 incelendiğinde, meme kanserli hastaların yinelemesiz yaşam süresi üzerinde tümör büyüklüğü ($p<0.001$), hormon terapisi ($p<0.05$) ve aksiller lenf nodu tutulumu ($p<0.001$) değişkenlerinin önemli etkilerinin olduğu; yaş, menopoz durumu, ilk doğum yaşı, aile hikayesi, tümör kadranı, progesteron reseptör durumu, östrojen reseptör durumu, kapsül dışı lenf nodu tutulumu, histolojik tip ve radyoterapi değişkenlerinin ise önemli etkilerinin olmadığı bulundu ($p>0.05$).

Açıklayıcı olmayan prior bilgili BYA sonuçlarına göre; tümör büyüklüğü arttıkça yineleme riskinin 1.144 kat arttığı, hormon terapisi alanların almayanlara göre 0.510 oranında yineleme riski taşıdığı ve *pozitif lenf nodu ≥ 4* olanların *negatif* olanlara göre yineleme riskinin 2.464 kat arttığı bulundu (Tablo 4.11).

Tablo 4.10. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin açıklayıcı prior bilgili BYA sonuçları

Bağımsız Değişkenler		$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$	Wald Ki-kare	p	HO
Yaş (yıl)		0.0233	0.0101	5.3409	0.021	1.024
Tümör Büyüklüğü (cm)		0.1707	0.0316	29.2239	<0.001	1.186
İlk Doğum Yaşı (yıl)	≥30	0.4545	0.4270	1.1329	0.287	1.575
Menopoz Durumu	Pre+Peri Menopoz	-0.0006	0.2070	0.00001	0.998	0.999
Progesteron Reseptör Durumu	Negatif	0.3415	0.1946	3.0811	0.079	1.407
Östrojen Reseptör Durumu	Negatif	0.0181	0.2119	0.0073	0.932	1.018
Radyoterapi	Var	-0.0898	0.0970	0.8559	0.355	0.914
Hormon Terapisi	Var	-0.2603	0.2635	0.9764	0.323	0.771
Aile Hikayesi	Meme Kanseri	0.4590	0.3265	1.9773	0.160	1.583
	Diğer Kanserler	0.227	0.2211	1.0517	0.305	1.254
Aksiller Lenf Nodu Tutulumu	Pozitif Lenf Nodu 1-3	0.6134	0.1375	19.9118	<0.001	1.847
	Pozitif Lenf Nodu ≥4	1.1596	0.1433	65.5252	<0.001	3.189
Tümör Kadranı	Çok Merkezli	0.1476	0.2411	0.3748	0.540	1.159
Kapsül Dışı Lenf Nodu Tutulumu	Pozitif	0.2475	0.2036	1.4778	0.224	1.281
Histolojik tip	Duktal	0.1161	0.2092	0.3078	0.579	1.123

Tablo 4.11. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin açıklayıcı olmayan prior bilgili BYA sonuçları

Bağımsız Değişkenler		$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$	Wald Ki-kare	p	HO
Yaş (yıl)		0.0151	0.0121	1.5553	0.212	1.015
Tümör Büyüklüğü (cm)		0.1347	0.0376	12.8341	<0.001	1.144
İlk Doğum Yaşı (yıl)	≥30	0.4103	0.4280	0.9188	0.338	1.507
Menopoz Durumu	Pre+Peri Menopoz	-0.1689	0.2968	0.3239	0.569	0.845
Progesteron Reseptör Durumu	Negatif	0.3666	0.2956	1.5382	0.215	1.443
Östrojen Reseptör Durumu	Negatif	-0.5586	0.3542	2.4876	0.115	0.572
Radyoterapi	Var	-0.1659	0.2886	0.3303	0.566	0.847
Hormon Terapisi	Var	-0.6725	0.2942	5.2242	0.022	0.510
Aile Hikayesi	Meme Kanseri	0.5097	0.3297	2.3908	0.122	1.665
	Diğer Kanserler	0.2228	0.2253	0.9781	0.323	1.250
Aksiller Lenf Nodu Tutulumu	Pozitif Lenf Nodu 1-3	0.2770	0.3014	0.8443	0.358	1.319
	Pozitif Lenf Nodu ≥4	0.9017	0.3231	7.7885	0.005	2.464
Tümör Kadranı	Çok Merkezli	0.3519	0.2436	2.0862	0.149	1.422
Kapsül Dışı Lenf Nodu Tutulumu	Pozitif	0.2468	0.2425	1.0355	0.309	1.280
Histolojik tip	Duktal	0.2362	0.2869	0.6775	0.410	1.266

Tablo 4.12’de, CRA, açıklayıcı prior bilgili BYA ve açıklayıcı olmayan prior bilgili BYA yöntemlerine ilişkin uyum istatistikleri verildi. ABK ve BBK, her üç yaklaşım için aynı değerde elde edildi. En büyük $-2x$ Logaritmik Olabilirlik istatistiği açıklayıcı prior bilgili BYA’da elde edilirken, en küçük değer $-2x$ Logaritmik Olabilirlik istatistiği CRA’da bulundu (Tablo 12).

Tablo 4.12. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin CRA, açıklayıcı olan ve olmayan prior bilgili BYA için uyum istatistikleri

Uyum İstatistiği	CRA	BYA-I	BYA-II
ABK	1234.540	1234.540	1234.540
BBK	1276.353	1276.353	1276.353
$-2x$ Logaritmik Olabilirlik	1204.540	1242.138	1219.509

*BYA-I: Açıklayıcı prior bilgili BYA ve BYA-II: Açıklayıcı olmayan prior bilgili BYA

Tablo 4.13 ve 4.14’de BYA yaklaşımlarında Markov zincirinin posterior dağılıma yakınsama performansına ilişkin Gelman-Rubin testi, Geweke testi ve posterior dağılımın otokorelasyon sonuçları verildi. BYA yaklaşımlarının her ikisinde de modele giren her bir değişkenin Gelman-Rubin test istatistiklerinin 1’e yakın olduğu ve Geweke testine göre de her bir değişken için Markov zincirinin iyi bir şekilde durağanlaşarak yakınsadığı ve Markov zincirinin ilgilenilen posterior dağılıma ulaştığı belirlendi ($p>0.05$). Aynı zamanda posterior dağılımın otokorelasyon değerleri çok büyük olmadığı için Markov zincirinin iyi bir şekilde durağanlaşarak yakınsadığı gözlemlendi.

Tablo 4.13. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin açıklayıcı prior bilgili BYA için MCMC'nin yakınsama istatistikleri

Bağımsız Değişkenler	Gelman-Rubin		Geweke		Posterior Dağılım Otokorelasyonları				
	\hat{V}/V_w	%97.5 Limiti	z_n	p	Lag1	Lag5	Lag10	Lag50	
Yaş (yıl)	1.0004	1.0010	0.5845	0.559	0.4428	0.0233	0.0100	0.0104	
Tümör Büyüklüğü (cm)	1.0001	1.0004	-1.8608	0.063	0.0847	0.0083	0.0057	-0.0000	
İlk Doğum Yaşı (yıl)	≥30	0.9999	1.0000	-1.9027	0.057	0.0827	-0.0066	0.0068	0.0084
Menopoz Durumu	Pre+Peri Menopoz	1.0006	1.0021	0.3748	0.708	0.4381	0.0150	0.0234	0.0034
Progesteron Reseptör Durumu	Negatif	1.0002	1.0002	-0.2873	0.774	0.3593	-0.0057	-0.0177	0.0110
Östrojen Reseptör Durumu	Negatif	1.0000	1.0001	-1.1182	0.264	0.4347	-0.0104	-0.0117	-0.0002
Radyoterapi	Var	1.0002	1.0007	-0.5848	0.559	0.0256	0.0056	-0.0021	0.0095
Hormon Terapisi	Var	1.0003	1.0008	-0.9754	0.329	0.4722	0.0087	-0.0120	0.0079
Aile Hikayesi	Meme Kanseri	1.0001	1.0002	-0.1555	0.876	0.0938	0.0087	-0.0155	0.0102
	Diğer Kanserler	1.0000	1.0000	-0.0942	0.925	0.0918	-0.0009	-0.0038	0.0198
Aksiller Lenf Nodu Tutulumu	Pozitif Lenf Nodu 1-3	1.0000	1.0003	-0.4078	0.683	0.1119	-0.0027	-0.0069	-0.0069
	Pozitif Lenf Nodu ≥4	1.0000	1.0000	1.6524	0.098	0.1980	0.0015	0.0008	-0.0114
Tümör Kadranı	Çok Merkezli	1.0000	1.0002	1.9056	0.056	0.1218	-0.0017	-0.0201	0.0096
Kapsül Dışı Lenf Nodu Tutulumu	Pozitif	1.0001	1.0004	-1.9101	0.053	0.1518	-0.0085	-0.0004	-0.0007
Histolojik tip	Duktal	1.0000	1.0000	1.9333	0.052	0.0254	-0.0198	0.0031	0.0023

Tablo 4.14. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin açıklayıcı olmayan prior bilgili BYA için MCMC'nin yakınsama istatistikleri

Bağımsız Değişkenler	Gelman-Rubin		Geweke		Posterior Dağılım Otokorelasyonları				
	\hat{V}/V_W	%97.5 Limiti	z_n	p	Lag1	Lag5	Lag10	Lag50	
Yaş (yıl)	1.0008	1.0021	-1.6904	0.091	0.6263	0.0821	-0.0010	-0.0265	
Tümör Büyüklüğü (cm)	1.0007	1.0020	-1.7271	0.084	0.1661	0.0014	0.0153	0.0208	
İlk Doğum Yaşı (yıl)	≥30	1.0000	1.0002	0.8547	0.393	0.1035	-0.0015	-0.0084	-0.0071
Menopoz Durumu	Pre+Peri Menopoz	1.0006	1.0012	-1.7499	0.080	0.6216	0.0749	-0.0024	-0.0256
Progesteron Reseptör Durumu	Negatif	1.0004	1.0010	-1.3008	0.193	0.5728	0.0510	-0.0028	0.0166
Östrojen Reseptör Durumu	Negatif	1.0003	1.0014	1.7040	0.088	0.6943	0.1199	0.0040	0.0036
Radyoterapi	Var	1.0001	1.0002	-1.6492	0.099	0.1408	0.0012	-0.0106	0.0011
Hormon Terapisi	Var	1.0002	1.0006	1.1911	0.234	0.5761	0.0518	-0.0077	-0.0027
Aile Hikayesi	Meme Kanseri	1.0000	1.0001	0.4397	0.660	0.1103	-0.0026	-0.0036	-0.0028
	Diğer Kanserler	0.9999	0.9999	-0.6498	0.516	0.1338	-0.0137	0.0055	-0.0204
Aksiller Lenf Nodu Tutulumu	Pozitif Lenf Nodu 1-3	1.0007	1.0027	-0.4089	0.683	0.5049	0.1031	0.0093	-0.0003
	Pozitif Lenf Nodu ≥4	1.0008	1.0031	0.1016	0.919	0.6845	0.1391	-0.0062	-0.0062
Tümör Kadranı	Çok Merkezli	1.0000	1.0001	0.2385	0.812	0.1282	0.0027	-0.0137	0.0008
Kapsül Dışı Lenf Nodu Tutulumu	Pozitif	1.0002	1.0010	1.2808	0.200	0.4346	0.0850	-0.0052	-0.0111
Histolojik tip	Duktal	1.0001	1.0005	-0.4009	0.689	0.0886	0.0096	0.0071	-0.0146

Tablo 4.15, 4.16 ve 4.17’de, CRA, açıklayıcı olan ve olmayan prior bilgili BYA yöntemleri için aşamalı değişken seçimi yaklaşımı kullanılarak elde edilen analiz sonuçları verildi.

Tablo 4.15’de, aşamalı CRA sonuçları verildi. Bu sonuçlara göre; yaş ($p<0.01$), tümör büyüklüğü ($p<0.001$), hormon terapisi ($p<0.01$) ve aksiller lenf nodu tutulumu ($p<0.001$) değişkenlerinin yinelemesiz yaşam süresi üzerinde önemli etkilerinin olduğu belirlendi. Aşamalı CRA sonuçları incelendiğinde; yaş arttıkça yineleme riskinin 1.023 kat arttığı, tümör büyüklüğü arttıkça yineleme riskinin 1.140 kat arttığı, hormon terapisi alanların almayanlara göre 0.604 oranında yineleme riski taşıdığı ve *pozitif lenf nodu ≥ 4* olanların *negatif* olanlara göre yineleme riskinin 3.261 kat arttığı bulundu (Tablo 4.15).

Tablo 4.16’da, açıklayıcı prior bilgili aşamalı BYA sonuçları verildi. Bu sonuçlara göre; yaş ($p<0.01$), tümör büyüklüğü ($p<0.001$), hormon terapisi ($p<0.05$) ve aksiller lenf nodu tutulumu ($p<0.001$) değişkenlerinin yinelemesiz yaşam süresi üzerinde önemli etkilerinin olduğu belirlendi. Ulaşılabilen açıklayıcı prior bilgiler kullanılan aşamalı BYA sonuçları incelendiğinde; yaş arttıkça yineleme riskinin 1.024 kat arttığı, tümör büyüklüğü arttıkça yineleme riskinin 1.180 kat arttığı, hormon terapisi alanların almayanlara göre 0.613 oranında yineleme riski taşıdığı, *pozitif lenf nodu 1-3* olanların *negatif* olanlara göre yineleme riskinin 1.908 kat arttığı ve *pozitif lenf nodu ≥ 4* olanların *negatif* olanlara göre yineleme riskinin 3.322 kat arttığı bulundu (Tablo 4.16).

Tablo 4.17’de, açıklayıcı olmayan prior bilgili aşamalı BYA sonuçları verildi. Bu sonuçlara göre; yaş ($p<0.01$), tümör büyüklüğü ($p<0.001$), hormon terapisi ($p<0.01$) ve aksiller lenf nodu tutulumu ($p<0.001$) değişkenlerinin yinelemesiz yaşam süresi üzerinde önemli etkilerinin olduğu belirlendi. Açıklayıcı olmayan prior bilgili aşamalı BYA sonuçları incelendiğinde; yaş arttıkça yineleme riskinin 1.023 kat arttığı, tümör büyüklüğü arttıkça yineleme riskinin 1.138 kat arttığı, hormon terapisi alanların almayanlara göre 0.606 oranında yineleme riski taşıdığı ve *pozitif lenf nodu ≥ 4* olanların *negatif* olanlara göre yineleme riskinin 3.314 kat arttığı bulundu (Tablo 4.17).

Tablo 4.15. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin aşamalı CRA sonuçları

Bağımsız Değişkenler		$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$	Wald Ki-kare	p	HO
Yaş (yıl)		0.0226	0.0076	8.8274	0.003	1.023
Tümör Büyüklüğü (cm)		0.1314	0.0360	13.3168	<0.001	1.140
Hormon Terapisi	Var	-0.5037	0.1917	6.9051	0.009	0.604
Aksiller Lenf Nodu Tutulumu	Pozitif Lenf Nodu 1-3	0.5194	0.2731	3.6178	0.057	1.681
	Pozitif Lenf Nodu ≥ 4	1.1819	0.2430	23.6517	<0.001	3.261

Tablo 4.16. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin açıklayıcı prior bilgili aşamalı BYA sonuçları

Bağımsız Değişkenler		$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$	Wald Ki-kare	p	HO
Yaş (yıl)		0.0237	0.0075	10.0554	0.002	1.024
Tümör Büyüklüğü (cm)		0.1655	0.0303	29.9290	<0.001	1.180
Hormon Terapisi	Var	-0.4893	0.1908	6.5785	0.010	0.613
Aksiller Lenf Nodu Tutulumu	Pozitif Lenf Nodu 1-3	0.6462	0.1364	22.4580	<0.001	1.908
	Pozitif Lenf Nodu ≥ 4	1.2006	0.1330	81.4679	<0.001	3.322

Tablo 4.17. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin açıklayıcı olmayan prior bilgili aşamalı BYA sonuçları

Bağımsız Değişkenler		$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$	Wald Ki-kare	p	HO
Yaş (yıl)		0.0225	0.0077	8.5867	0.003	1.023
Tümör Büyüklüğü (cm)		0.1289	0.0365	12.4488	<0.001	1.138
Hormon Terapisi	Var	-0.5005	0.1920	6.7933	0.009	0.606
Aksiller Lenf Nodu Tutulumu	Pozitif Lenf Nodu 1-3	0.5252	0.2724	3.7190	0.054	1.691
	Pozitif Lenf Nodu ≥ 4	1.1982	0.2433	24.2468	<0.001	3.314

Tablo 4.18’de, aşamalı CRA, açıklayıcı olan ve olmayan prior bilgili BYA yöntemlerine ilişkin uyum istatistikleri verildi. ABK ve BBK, her üç yaklaşım için aynı değerde elde edildi. En büyük -2xLogaritmik Olabilirlik istatistiği açıklayıcı prior bilgili BYA’da elde edilirken en küçük değer -2xLogaritmik Olabilirlik istatistiği CRA’da bulundu. Bu değerler arasında da çok büyük değer farklılıkları söz konusu değildi (Tablo 18).

Tablo 4.18. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin aşamalı CRA, açıklayıcı olan ve olmayan prior bilgili BYA için uyum istatistikleri

Uyum İstatistiği	CRA	BYA-I	BYA-II
ABK	1227.214	1227.214	1227.214
BBK	1241.152	1241.152	1241.152
-2xLogaritmik Olabilirlik	1217.214	1243.573	1222.218

*BYA-I: Açıklayıcı prior bilgili BYA ve BYA-II: Açıklayıcı olmayan prior bilgili BYA

Tablo 4.19 ve 4.20’de aşamalı BYA yaklaşımlarında Markov zincirinin posterior dağılıma yakınsama performansına ilişkin Gelman-Rubin testi, Geweke testi ve posterior dağılımın otokorelasyon sonuçları verildi. BYA yaklaşımlarının her ikisinde de modele giren her bir değişkenin Gelman-Rubin test istatistiklerinin 1’e yakın olduğu ve Geweke testine göre de her bir değişken için Markov zincirinin iyi bir şekilde durağanlaşarak yakınsadığı ve Markov zincirinin ilgilenilen posterior dağılıma ulaştığı belirlendi ($p>0.05$). Aynı zamanda posterior dağılımın otokorelasyon değerleri çok büyük olmadığı için Markov zincirinin iyi bir şekilde durağanlaşarak yakınsadığı gözlemlendi.

Tablo 4.19. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin açıklayıcı prior bilgili aşamalı BYA için MCMC'nin yakınsama istatistikleri

Bağımsız Değişkenler	Gelman-Rubin		Geweke		Posterior Dağılım Otokorelasyonları				
	\hat{V}/V_W	%97.5 Limiti	z_n	p	Lag1	Lag5	Lag10	Lag50	
Yaş (yıl)	1.0002	1.0007	-0.3694	0.712	0.0556	-0.0017	-0.0127	0.0019	
Tümör Büyüklüğü (cm)	1.0001	1.0006	1.9274	0.054	0.0843	0.0071	-0.0114	-0.0014	
Hormon Terapisi	Var	1.0005	1.0018	0.3126	0.108	0.0120	0.0063	0.0064	-0.0002
Aksiller Lenf Nodu Tutulumu	Pozitif Lenf Nodu 1-3	1.0000	1.0000	-0.7866	0.755	0.0859	-0.0116	-0.0083	0.0069
	Pozitif Lenf Nodu ≥ 4	1.0000	1.0001	-1.6078	0.432	0.1128	-0.0030	-0.0154	-0.0011

Tablo 4.20. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin açıklayıcı olmayan prior bilgili aşamalı BYA için MCMC'nin yakınsama istatistikleri

Bağımsız Değişkenler	Gelman-Rubin		Geweke		Posterior Dağılım Otokorelasyonları				
	\hat{V}/V_W	%97.5 Limiti	z_n	p	Lag1	Lag5	Lag10	Lag50	
Yaş (yıl)	1.0003	1.0012	0.0318	0.975	0.0759	-0.0007	0.0063	0.0186	
Tümör Büyüklüğü (cm)	1.0001	1.0006	-1.4505	0.147	0.1333	-0.0054	-0.0074	0.0133	
Hormon Terapisi	Var	1.0003	1.0009	0.1305	0.896	0.0284	0.0008	-0.0161	0.0189
Aksiller Lenf Nodu Tutulumu	Pozitif Lenf Nodu 1-3	1.0000	1.0001	0.7888	0.430	0.3855	-0.0039	-0.0015	0.0018
	Pozitif Lenf Nodu ≥ 4	1.0001	1.0005	0.8045	0.421	0.4228	0.0018	-0.0073	0.0116

Tablo 4.9, 4.10 ve 4.11’de elde edilen bulgulara dayanarak yaş, menopoz durumu, ilk doğum yaşı, aile hikayesi, tümör kadranı, tümör büyüklüğü, progesteron reseptör durumu, östrojen reseptör durumu, aksiller lenf nodu tutulumu, kapsül dışı lenf nodu tutulumu, histolojik tip, radyoterapi ve hormon terapisi değişkenlerinin parametre tahminleri ve standart hatalarının gösterimi Şekil 4.25’de verildi.

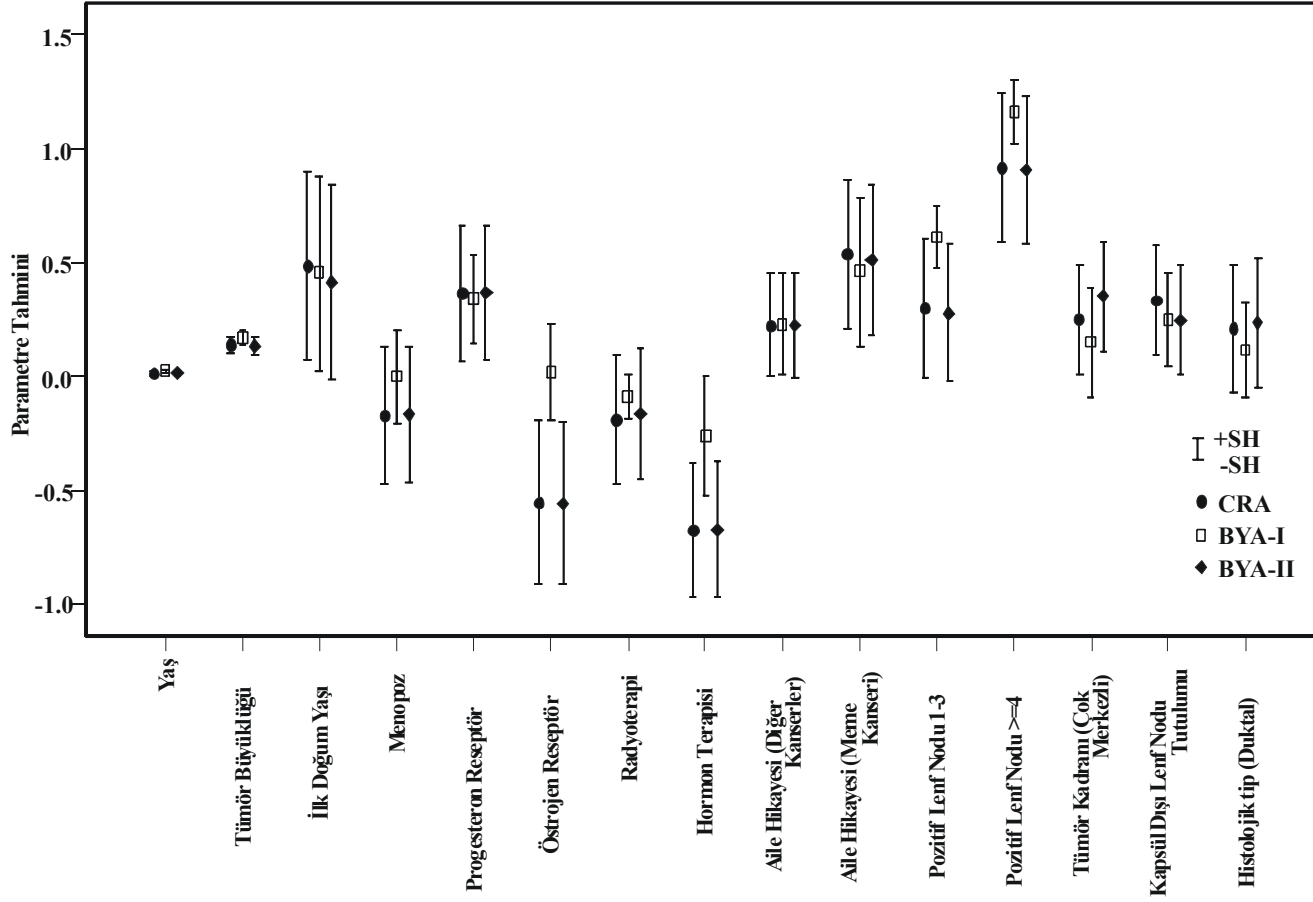
Şekil 4.25 incelendiğinde, özellikle prior bilgi kullanılan BYA’da parametre tahminlerinin standart hatalarının, açıklayıcı olmayan prior bilgi kullanılan BYA ile CRA’nın parametre tahminlerinin standart hatalarından daha küçük olduğu belirlendi.

Tablo 4.15, 4.16 ve 4.17’de elde edilen bulgulara dayanarak yaş, tümör büyüklüğü, aksiller lenf nodu tutulumu ve hormon terapisi değişkenlerinin parametre tahminleri ve standart hatalarının gösterimi Şekil 4.26’da verildi.

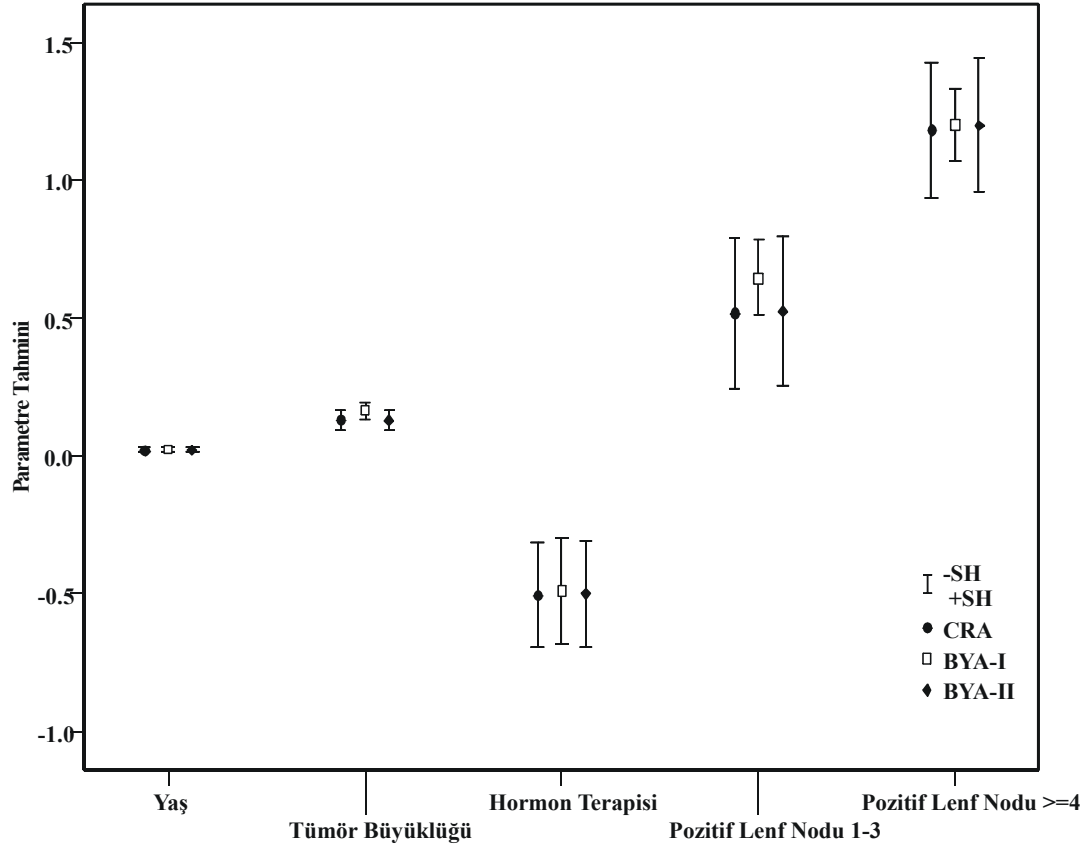
Şekil 4.26 incelendiğinde, prior bilgi kullanılan BYA’da parametre tahminlerinin standart hatalarının, CRA ve açıklayıcı olmayan prior bilgili BYA’nın parametre tahminlerinin standart hatalarından biraz daha küçük olduğu belirlendi.

Aşamalı CRA ve BYA sonucunda elde edilen parametre tahminlerine göre düzeltilmiş yaşam fonksiyonlarının grafiği Şekil 4.27’de verildi.

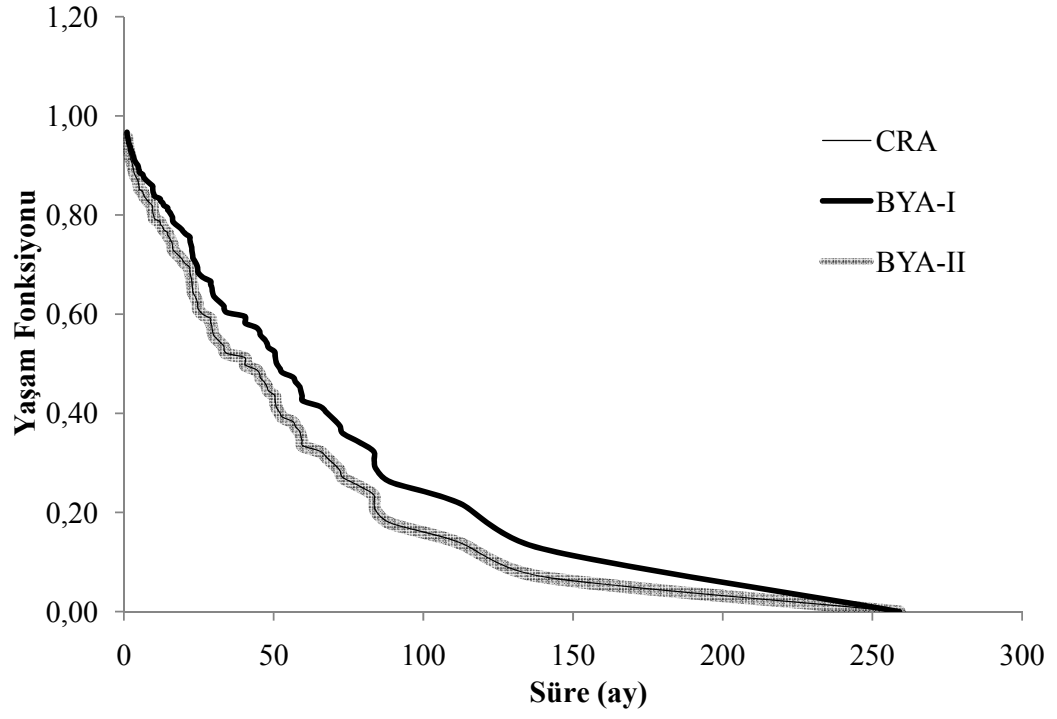
Şekil 4.27 incelendiğinde, CRA ve açıklayıcı olmayan prior bilgili BYA’nın yaşam fonksiyonlarının hemen hemen aynı değerleri alarak eğrilerin örtüştüğü gözlemlendi. Açıklayıcı prior bilgili BYA’nın ise yaşam fonksiyonunun diğerlerine göre farklı bir görüntü sergileyerek, CRA ve açıklayıcı olmayan BYA’ya göre yaşama olasılıklarının her bir yaşam süresinde biraz daha büyük olduğu gözlemlendi.



Şekil 4.25. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin CRA, BYA-I ve BYA-II'nin parametre tahminleri ve standart hatalarının gösterimi



Şekil 4.26. Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin aşamalı CRA, BYA-I ve BYA-II'nin parametre tahminleri ve standart hatalarının gösterimi



Şekil 4.27. Aşamalı CRA, BYA-I ve BYA-II için meme kanserli hastaların yinelemesiz yaşam sürelerine göre yaşam fonksiyonları grafiği

5. TARTIŞMA

Sağlık alanında yaşam sürelerine ilişkin verilerin incelenmesi amacıyla kullanılan yöntemler her zaman güncelliğini korumaktadır. Özellikle birden çok risk faktörünü ele alarak yaşam süresi üzerinde etkili olan faktörlerin belirlenmesinde kullanılan CRA, hemen hemen bütün istatistiksel paket programlarda yer alan bir yöntem olduğu için araştırmacılar tarafından tercih edilmektedir. CRA'ya alternatif olarak geliştirilmiş olan BYA ise son yıllarda kullanılmaya başlanmıştır.

BYA, gözlenen veriden elde edilen yeni bilgi ile diğer kaynaklardan elde edilen bilgilerin birleştirilmesine dayanan bir yöntemdir. Dolayısıyla prior bilginin elde edilmesi, Bayesgil çıkarsamada önemli bir rol oynar. Bayesgil analiz; önceki bilgiler ve geçmiş deneyimler gibi açıklayıcı prior bilgiler ile gözlenen verinin birleştirilerek subjektif yorumuna dayanır [6, 47]. Bayesgil ve klasik yöntemler arasındaki en önemli fark, BYA'nın prior bilgiyi hesaba katması ve dolayısı ile araştırmacının subjektif kanaatlerini de yansıtmasıdır.

İncelenen olayla ilgili geçmişte herhangi bir çalışma yapılmamış olabilir ya da yapılan çalışmaların sonuçlarının güvenilir olmaması nedeniyle Bayesgil analizlerde, uygun prior bilgiye ulaşılamayabilir. Bayesgil yaklaşımlarda, veri yapısına uygun ve gerçeği yansıtan açıklayıcı prior bilgiye ulaşılamadığı durumlarda, açıklayıcı olmayan prior bilgi kullanılabilmesine rağmen [22, 39], Calle ve ark. (2006) ve İbrahim ve ark. (2001), Bayesgil ve klasik yaklaşımların genellikle benzer sonuçlar verdiğini ve açıklayıcı olmayan prior bilgi kullanılarak yapılan BYA'nın klasik yöntemlerin özel bir durumu olduğunu belirtmişlerdir [4, 22]. Çalışmamızda da açıklayıcı olmayan prior bilgiye dayalı farklı koşullarda (farklı değişken yapısı, farklı parametre ve farklı örneklem büyüklüğü) yapılan simülasyon sonuçları incelendiğinde, genel olarak BYA'nın CRA'ya göre performansının önemli derecede yüksek olmadığı ve sadece biraz daha küçük yanlılıkla simülasyon parametresini tahmin ettiği belirlendi. İbrahim ve ark. (2001), açıklayıcı olmayan ya da açıklayıcı olup da uygun olmayan prior bilgilerin kullanılmasının değişkenlerin önemliliğinin belirlenmesinde yararlı olabileceğini fakat model seçimi ve model karşılaştırması gibi uygulamalarda uygun

olmayacağını belirtmişlerdir. Ayrıca, posterior model olasılıklarının hesaplanmasında uygun prior bilginin kullanılması gerektiğini vurgulamış ve açıklayıcı olmayan prior bilgiyle yürütülen çalışmalardan elde edilen parametre tahminlerinin gelecekte yapılacak çalışmalar için uygun olmayacağını ifade etmişlerdir [22]. Benzer şekilde çalışmamızda da açıklayıcı olmayan prior bilgilerin kullanılmasının, posterior tahminlerde kararsızlığa ve Gibbs örnekleme yönteminde yakınsama problemlerine neden olduğu gözlenmiştir. Gelfand ve Mallick (1995), büyük hacimli veri setlerinde, açıklayıcı olmayan prior bilgi kullanılan analizlerden elde edilen parametre tahminlerinin klasik en çok olabilirlik yönteminden elde edilen parametre tahminlerine yakın olacağını ifade etmişlerdir [16]. Çalışmamızda açıklayıcı olmayan prior bilgi kullanıldığında BYA, CRA'ya göre biraz daha küçük yanlılık ve standart hatalı simülasyon parametresi tahmin etti ve açıklayıcı olmayan prior bilgi kullanıldığında BYA'nın, çok kararlı bir yapı sergilemediği gözlendi.

Yin ve Ibrahim (2006), BYA'yı cilt kanseriyle ilgili gerçek bir veri seti üzerinde ve 500 tekrarlı, $n=300$, 500 ve 1000 örneklem büyüklüğüne göre bir simülasyon çalışması ile karşılaştırmış, açıklayıcı olmayan prior bilgi için ölüm oranı farklı olan veri setlerinde ölüm oranı arttıkça posterior dağılımın standart sapmasının arttığını tespit etmişlerdir [48]. Calle ve ark. (2006), yağsız ve tam yağlı yoğurt ürünlerinin raf ömrüyle ilgili uygulama çalışmasında, açıklayıcı olan ve olmayan prior bilgiye dayanarak BYA uygulaması yapmışlar ve açıklayıcı olan ve olmayan prior bilgidен benzer sonuçlar elde etmişlerdir [4]. Wong ve ark. (2005), okul öncesi Çinli çocuklarda aktif diş çürümesini durdurmada silver diamin florit ve sodyum florit parlaticısının etkinliğini araştırmak için klinik bir çalışmadan elde edilen tamamlanmamış verinin analizinde açıklayıcı olmayan prior bilgiye dayalı BYA'yı kullanmışlar ve Bayesgil yaklaşımların doğru bir şekilde kullanılmaması durumunda, yanlılığı büyük olan tahminler ve dolayısıyla yanlış sonuçlar ortaya koyacağını belirtmişlerdir [47]. Simülasyon çalışmamızda da BYA için simülasyon parametresinin tahmininde, prior bilginin çok önemli bir belirleyici olduğu ortaya konuldu. Özellikle açıklayıcı prior bilgi kullanılmasının, BYA'nın performansını önemli derecede etkilediği gözlendi. Veri yapısına uygun ve gerçeği yansıtan prior bilgi kullanılması durumunda BYA'nın oldukça küçük yanlılıkla çok iyi performans gösterdiği, özellikle gerçeği yansıtan prior

bilginin varyansının küçük olmasının BYA'nın performansını iyi yönde artırdığı gözlemlendi. Uygun olmayan açıklayıcı prior bilgi kullanıldığında ise BYA'nın performansının kötü yönde etkilendiği gözlemlendi. Farklı örneklem büyüklüğü ve değişken yapısında açıklayıcı prior bilgi gerçeği yansıtmaktan uzaklaştıkça, Bayesgil yaklaşımın yanlılığının arttığı gözlemlendi. Ayrıca gerçeğe uygun olmayan prior bilginin varyansının küçük olmasının, yanlılığın daha da artmasına neden olduğu belirlendi. Bunun dışında yöntemlerin tahmin performansının örneklem büyüklüğünden de etkilendiği gözlemlendi. Gelfand ve Mallick (1995) ve Gelman (2002), farklı örneklem büyüklüğündeki veri setlerini karşılaştırmışlar ve parametreleri doğru bir şekilde tanımlanan büyük örneklem hacimli veri setleriyle çalışıldığında mantıklı bir prior dağılımın seçilmesinin, posterior çıkarımlar üzerinde küçük etkilere sahip olduğunu ve buna karşın örneklem büyüklüğü küçük ya da mevcut verinin ilgilenilen parametre hakkında sadece dolaylı bilgi sağladığı durumlarda, prior dağılımın daha önemli olduğunu belirtmişlerdir [16, 18]. Benzer şekilde çalışmamızda da, özellikle veri yapısına uygun ve gerçeği yansıtan prior bilgi kullanılması durumunda örneklem büyüklüğü küçük olduğunda bile BYA'nın iyi performans gösterdiği belirlendi. Örneklem büyüklüğü artırıldığında ise simülasyon parametresini tahminde CRA'nın yanlılığı azalmasına rağmen, açıklayıcı prior bilgili BYA'nın hem yanlılığının hem de standart hatasının azaldığı gözlemlendi.

Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin araştırmalar incelendiğinde; Pinto ve ark. (2007), 308 meme kanserli hastanın yinelemesiz yaşam süresi üzerinde yaş, grade, tümör büyüklüğü, aksiler lenf nodu tutulumu, DNA tetraploidy tümör aralığı, S phase fraction, östrojen ve progesteron reseptör durumu değişkenlerinin etkilerini incelemek amacıyla CRA uygulamış ve yinelemesiz yaşam süresi üzerinde ele alınan bütün risk faktörlerinin etkisinin önemli olduğunu belirlemişlerdir [32]. Yu ve ark. (1995), 174 meme kanserli hastanın yinelemesiz yaşam süresi üzerine etki eden risk faktörlerini belirlemek amacıyla yaptıkları çalışmalarında prostat spesifik antijeni, yaş, evre, aksiler lenf nodu tutulumu, tümör büyüklüğü, histolojik grade, östrojen ve progesteron reseptör değişkenlerinden sadece progesteron reseptör dışında kalan risk faktörlerinin yinelemesiz yaşam süresi üzerinde önemli etkilerinin olduğunu belirlemişlerdir [50]. Foekens ve ark. (1992), 671 meme kanserli

hastanın yinelemesiz yaşam süresi üzerine etki eden risk faktörlerini belirlemek amacıyla yaptıkları çalışmalarında menopoz durumu, tümör büyüklüğü, aksiller lenf nodu tutulumu, östrojen/progesteron reseptör oranı, veurokinase tip plasminogen aktivator (uPA) değişkenlerinin önemli etkilerinin olduğunu CRA ile belirlemişlerdir [13]. Çalışmamızda ise CRA sonucunda tümör büyüklüğü, hormon terapisi ve aksiller lenf nodu tutulumu değişkenlerinin yinelemesiz yaşam süresi üzerinde önemli etkilerinin olduğu belirlendi. Açıklayıcı olmayan prior bilgili BYA sonucunda da yine tümör büyüklüğü, hormon terapisi ve aksiller lenf nodu tutulumu değişkenlerinin yinelemesiz yaşam süresi üzerinde önemli etkilerinin olduğu belirlendi. Açıklayıcı prior bilgili BYA sonucunda ise yaş, tümör büyüklüğü ve aksiller lenf nodu tutulumu değişkenlerinin önemli risk faktörleri olduğu belirlendi. Bu sonuçlara göre tümör büyüklüğü ve aksiller lenf nodu tutulumu değişkenlerinin, yöntemlerin hepsinde yinelemesiz yaşam süresi üzerinde etkili olduğu bulundu. Aşamalı yaklaşım kullanıldığında hem CRA hem de açıklayıcı olmayan prior bilgili BYA sonucunda yaş, tümör büyüklüğü, hormon terapisi ve aksiller lenf nodu tutulumu (*pozitif lenf nodu ≥ 4*) değişkenlerinin yinelemesiz yaşam süresi üzerinde önemli etkilerinin olduğu belirlendi. Açıklayıcı prior bilgili aşamalı BYA sonucunda ise yaş, tümör büyüklüğü, hormon terapisi değişkenlerine ve aksiller lenf nodu tutulumuna (*pozitif lenf nodu ≥ 4*) ek olarak aksiller lenf nodu tutulumu (*pozitif lenf nodu 1-3*) kategorisinin de yinelemesiz yaşam süresi üzerinde önemli etkisinin olduğu belirlendi. Sonuçlara genel olarak bakıldığında, açıklayıcı prior bilgili BYA ile elde edilen parametre tahminlerinin standart hatalarının, açıklayıcı olmayan prior bilgili BYA ile CRA'dan elde edilen parametre tahminlerinin standart hatalarından biraz daha küçük olduğu gözlemlendi.

İbrahim ve ark. (2001) ve Wong ve ark. (2005), BYA'nın karmaşık veriyi modellemede prior bilgiyi kullanması nedeniyle klasik yöntemlere göre avantajlı olduğunu belirtmelerine rağmen simülasyon çalışmamızda, BYA'nın, özellikle gerçeği yansıtan prior bilgiye dayanarak uygulandığında CRA'ya göre daha avantajlı olduğu, gerçeği yansıtmayan ya da rasgele prior bilgi kullanılmasının avantajdan daha çok dezavantaj getirdiği gözlemlendi. BYA uygulanacaksa hangi koşulda olursa olsun, veri yapısına uygun ve gerçeği yansıtan prior bilgiyle çalışılmalıdır. BYA'da açıklayıcı olmayan ya da otomatik prior bilgi kullanıldığında, CRA'nın sonuçlarına benzer

sonular elde edilmektedir. Fakat veri yapısına uygun ve gereęi yansıtan prior bilgiyle alıřıldığında, BYA'nın performansı oldukça artmaktadır. Bu durum Bayesgil yaklařımlar kullanılacaksa gereęi yansıtan prior bilginin elde edilmesini zorunlu kılmaktadır.

6. SONUÇ VE ÖNERİLER

BYA ve CRA yöntemleri, farklı koşullarda gerçekleştirilen simülasyon uygulamaları ve meme kanserli hastalara ilişkin bir veri seti uygulamasıyla karşılaştırıldı.

Simülasyon uygulamasında, açıklayıcı olmayan prior bilgiye dayalı BYA ve CRA yöntemlerinin simülasyon parametresine yakınsama bakımından benzer performans sergilediği ($p>0.05$) ve BYA yönteminin CRA'ya göre sadece biraz daha küçük yanlılıkla simülasyon parametresini tahmin ettiği belirlendi. Açıklayıcı prior bilgiye dayalı simülasyon uygulamasında ise BYA'nın performansında göze çarpan bir değişim gözlemlendi. Gerçeği yansıtan prior bilgi kullanılması durumunda BYA, oldukça küçük yanlılıkla iyi bir performans gösterdi. Özellikle gerçeği yansıtan prior bilginin varyansının küçük olmasının BYA'nın yanlılığını azalttığı gözlemlendi. Uygun olmayan açıklayıcı prior bilgi kullanıldığında ise BYA'nın prior bilgiden etkilendiği ortaya çıktı. Her koşulda açıklayıcı prior bilgi gerçeği yansıtmaktan uzaklaştıkça, BYA'nın yanlılığının arttığı ve simülasyon parametresinden uzaklaştığı belirlendi. Ayrıca gerçeğe uygun olmayan prior bilginin varyansının küçük olmasının, yanlılığın daha da artmasına neden olduğu bulundu. Bunun dışında yöntemlerin tahmin performansının örneklem büyüklüğünden de etkilendiği gözlemlendi. Özellikle gerçeği yansıtan prior bilgi kullanılması durumunda, BYA küçük hacimli örneklemde de CRA'ya göre daha küçük standart hatalı tahminler elde ederek iyi performans gösterdi. Örneklem büyüklüğü artırıldığında CRA'nın da yanlılığının ve standart hatasının azaldığı, fakat açıklayıcı prior bilgili BYA'nın hem yanlılığı hem de standart hatası daha küçük tahminler ortaya koyduğu bulundu.

Meme kanserli hastaların yinelemesiz yaşam sürelerine ilişkin uygulamada CRA sonucunda tümör büyüklüğü ($H_0=1.147$, $p<0.001$), hormon terapisi ($H_0=0.508$, $p<0.05$) ve aksiller lenf nodu tutulumu (*pozitif lenf nodu* ≥ 4 için $H_0=2.501$, $p<0.01$) değişkenlerinin yinelemesiz yaşam süresi üzerinde önemli etkilerinin olduğu belirlendi. Açıklayıcı olmayan prior bilgili BYA sonucunda da yine tümör büyüklüğü ($H_0=1.144$, $p<0.001$), hormon terapisi ($H_0=0.510$, $p<0.05$) ve aksiller lenf nodu tutulumu (*pozitif*

lenf nodu ≥ 4 için HO=2.464, $p<0.01$) deęişkenlerinin yinelemesiz yaşam süresi üzerinde önemli etkilerinin olduęu belirlendi. Açıklayıcı prior bilgili BYA sonucunda ise yaş (HO=1.024, $p<0.05$), tümör büyüklüęü (HO=1.186, $p<0.001$) ve aksiller lenf nodu tutulumu (*pozitif lenf nodu 1-3* için HO=1.847, $p<0.001$ ve *pozitif lenf nodu* ≥ 4 için HO=3.189, $p<0.001$) deęişkenlerinin yinelemesiz yaşam süresi üzerinde önemli etkilerinin olduęu belirlendi. Ayrıca açıklayıcı olmayan prior bilgili BYA ile CRA benzer sonuçlar verirken, açıklayıcı prior bilgili BYA'da, hormon terapisinin önemlilięi ortadan kalkarak yaş deęişkeninin yinelemesiz yaşam süresi üzerinde etkili olduęu sonucu bulundu. Bu sonuçlara göre tümör büyüklüęü ve aksiller lenf nodu tutulumu deęişkenlerinin yöntemlerin hepsinde yinelemesiz yaşam süresi üzerinde etkili olduęu bulundu.

Aşamalı CRA yönteminde yaş (HO=1.023, $p<0.01$), tümör büyüklüęü (HO=1.140, $p<0.001$), hormon terapisi (HO=0.604, $p<0.01$) ve aksiller lenf nodu tutulumunun (*pozitif lenf nodu* ≥ 4 için HO=3.261, $p<0.001$) yinelemesiz yaşam süresi üzerinde önemli etkilerinin olduęu belirlendi. Aşamalı açıklayıcı olmayan prior bilgili BYA yönteminde yaş (HO=1.023, $p<0.01$), tümör büyüklüęü (HO=1.138, $p<0.001$), hormon terapisi (HO=0.606, $p<0.01$) ve aksiller lenf nodu tutulumunun (*pozitif lenf nodu* ≥ 4 için HO=3.314, $p<0.001$) yinelemesiz yaşam süresi üzerinde önemli etkilerinin olduęu bulundu. Aşamalı açıklayıcı prior bilgili BYA yönteminde ise yaş (HO=1.024, $p<0.01$), tümör büyüklüęü (HO=1.180, $p<0.001$), hormon terapisi (HO=0.613, $p<0.05$) ve aksiller lenf nodu tutulumunun (*pozitif lenf nodu 1-3* için HO=1.908, $p<0.001$ ve *pozitif lenf nodu* ≥ 4 için HO=3.322, $p<0.001$) yinelemesiz yaşam süresi üzerinde önemli etkilerinin olduęu bulundu. Böylece aşamalı yaklaşım kullanıldığında hem CRA hem de BYA yöntemleriyle yaş, tümör büyüklüęü, hormon terapisi ve aksiller lenf nodu tutulumu deęişkenlerinin yinelemesiz yaşam süresi üzerinde önemli etkilerinin olduęu belirlendi. Fakat açıklayıcı prior bilgili BYA'da, *pozitif lenf nodu 1-3* kategorisinin de yinelemesiz yaşam süresi üzerinde önemli etkisinin olduęu belirlendi. Bulgular ayrıntılı olarak incelendiğinde, açıklayıcı prior bilgili BYA'da parametre tahminlerinin standart hatalarının, açıklayıcı olmayan prior bilgili BYA ile CRA'nın parametre tahminlerinin standart hatalarından biraz daha küçük olduęu bulundu.

Çalışmamızdan elde edilen bulgulara dayanarak yaşam verilerinin analizinde hangi yöntemin tercih edileceği konusunda, öncelikle veri analiziyle ne amaçlandığının sorgulanmasının gerekliliği ortaya çıkmaktadır. Bir araştırmada ilgilenilen değişkenler hakkında prior bilgi varsa BYA'nın tercih edilmesi uygun olacaktır. BYA sadece bir nokta tahminine dayanarak değil, posterior dağılıma dayanarak incelenmek istenen verilerin analizi için uygun bir yapı sağlamaktadır. Diğer taraftan sadece olabilirliğe dayanarak parametrelerin tahmini yeterliyse, Newton-Raphson yöntemi gibi nümerik optimizasyon yöntemlerini kullanan CRA'nın tercih edilmesi uygun olacaktır. Ayrıca hangi koşulda olursa olsun BYA uygulanacaksa, prior bilginin veri yapısına uygun ve gerçeği yansıtan değerde olması gerekmektedir. Açıklayıcı olmayan prior bilgili BYA ile CRA, yaklaşık olarak benzer sonuçlar vermektedir. Fakat veri yapısına uygun ve gerçeği yansıtan prior bilgi olduğunda BYA'nın performansında iyi yönde önemli değişim gözlenmektedir. Bu durum, BYA kullanılacaksa veri yapısına uygun ve gerçeği yansıtan prior bilginin elde edilmesini zorunlu kılmaktadır. Bu nedenle araştırmacılar, veri yapısına uygun ve gerçeği yansıtan açıklayıcı prior bilgiye ulaşmak için iyi bir kaynak taraması yapmalıdır. Prior bilginin elde edilmesinde, özellikle aynı toplumdaki gelen benzer özellik taşıyan birimler üzerinde yapılan çalışmalar tercih edilmeli ya da bu tür çalışmalara ulaşılamıyor ve prior bilgi gerçeği yansıtmıyorsa, prior bilgi olarak büyük varyanslı değişkenler kullanılmalıdır. Bunun yanı sıra, özellikle aynı konuyla ilgili birçok çalışma yapılmış ise, meta analizi uygulamalarına dayanarak belirlenecek prior bilgi veri yapısına daha uygun ve geçerli olacaktır. Ayrıca ilgilenilen olayların zaman içerisinde değişim gösterebilmesi nedeniyle, prior bilginin elde edilmesinde ulaşılan çalışmaların güncelliği de dikkate alınmalıdır.

Sonuç olarak, gerçeği yansıtan prior bilgiye ulaşamadığında BYA yönteminin uygulanmasının, fazla bilgisayar zamanı gerektirdiği ve açıklayıcı olmayan prior bilgi kullanımının CRA'ya göre bir avantaj sağlamadığı ortaya çıkmıştır. Bu nedenle yaşam sürelerine ilişkin verilerin analizinde araştırmacılar, veri yapısına uygun ve gerçeği yansıtan prior bilgi olduğunda BYA yöntemini aksi durumda ise CRA yöntemini tercih etmelidir.

KAYNAKLAR DİZİNİ

1. Bender, R., Augustin, T., Blettner, M., 2005, Generating survival times to simulate Cox proportional hazards models, *Statistics in Medicine*, 24(11), 1713-1723 p.
2. Billingham, L.J., Abrams, K.R., Jones, D.R., 1999, Methods for the analysis of quality-of-life and survival data in health technology assessment, *Health Technology Assessment*, 3(10), 45-49 p.
3. Broström, G., 2002, Cox regression; ties without tears, *Communications in Statistics - Theory and Methods*, 31(2), 285-297 p.
4. Calle, M.L., Hough, G., Curia, A., Gómez, G., 2006, Bayesian survival analysis modeling applied to sensory shelf life of foods, *Food Quality and Preference*, 17(3-4), 307-312 p.
5. Carlin, B.P. and Louis, T.A., 2000, *Bayes and empirical Bayes methods for data analysis*, Chapman and Hall/CRC, Hertfordshire, U.K.
6. Congdon, P., 2006, *Bayesian statistical modelling*, John Wiley & Sons, England.
7. Congdon, P., 2003, *Applied bayesian modelling*, John Wiley & Sons, England.
8. De Blasi, P. and Hjort, N.L., 2007, Bayesian survival analysis in proportional hazard models with logistic relative risk, *Scand J Statist*, 34, 229-257 p.
9. Delwiche, L.D. and Slaughter, S.J., 1996, *The little SAS book*, SAS Institute Inc., Cary, NC, USA.
10. Dempster, A.P., 2005, Bayesian methods, In: Armitage, P., Colton, T. (eds.), *Encyclopedia of Biostatistics*, Wiley, New York.
11. Doss, H. and Huffer, F.W., 2003, Monte Carlo methods for bayesian analysis of survival data using mixtures of dirichlet process priors, *Journal of Computational and Graphical Statistics* 12(2), 282-307 p.

12. Figini, S., Giudici, P., Brooks, S.P., Bayesian feature selection to estimate customer survival, ISBA/8th World Meeting on Bayesian Statistics, Valencia, June 5-6 2006.
13. Foekens, J.A., Schmitt, M., Van Putten, W.L.J., Peters, H.A., Bontenbal, M., Janicke, F., Klijn, J.G.M., 1992, Prognostic value of urokinase-type plasminogen activator in 671 primary breast cancer patients, *Cancer Research*, 52, 6101-6105.
14. Forrester, M.L., Pettitt, A.N., Gibson, G.J., 2007, Bayesian inference of hospital-acquired infectious diseases and control measures given imperfect surveillance data, *Biostatistics*, 8(2), 383-401.
15. Gasparini, G., Toi, M., Gion, M., Verderio, P., Dittadi, R., Hanatani, M., Matsubara, I., Vinante, O., Bonoldi, E., Boracchi, P., Gatti, C., Suzuki, H., Tominaga, T., 1997, Prognostic significance of vascular endothelial growth factor protein in node-negative breast carcinoma, *Journal of the National Cancer Institute*, 89(2), 139-147.
16. Gelfand, A.E. and Mallick, B.K., 1995, Bayesian analysis of proportional hazards models built from monotone functions, *Biometrics*, 51(3), 843-852 p.
17. Gelman, A., 2002, Posterior distribution, *Encyclopedia of Environmetrics*, 3, 1627-1628 p.
18. Gelman, A., 2002, Prior distribution, *Encyclopedia of Environmetrics*, 3, 1634-1637 p.
19. Ghosh, J.K. and Ramamoorthi, R.V., 2003, *Bayesian nonparametrics*, Springer, New York.
20. Hançerlioğulları, A., 2006, Monte Carlo simülasyon metodu ve MNCP kod sistemi, *Kastamonu Eğitim Dergisi*, 14(2), 545-556 s.
21. Hosmer, D.W. and Lemeshow, S., 1999, *Applied survival analysis: regression modeling of time to event data*, John Wiley & Sons, Canada.

22. Ibrahim, J.G., Chen, M.H., Sinha, D., 2001, Bayesian survival analysis, Springer-Verlag, New York.
23. Kara, İ., 2000, Olasılık, Bilim Teknik Yayınevi, İstanbul.
24. Kleinbaum, D.G. and Klein, M., 1996, Survival analysis: a self-learning text, Springer, USA.
25. Kurt, İ., 2004, Yapay Sinir Ağları ile Lojistik Regresyon Analizi'nin Karşılaştırılması, Yüksek Lisans Tezi, Trakya Üniversitesi, Sağlık Bilimleri Enstitüsü.
26. Langner, I., Bender, R., Lenz-Tönjes, R., Küchenhoff, H., Blettner, M., 2003, Bias of Maximum-Likelihood estimates in logistic and Cox regression models: A comparative simulation study. Sonderforschungsbereich 386, Paper 362. Online <http://epub.ub.uni-muenchen.de/archive/00001737/>.
27. Lee, E.T. and Wang, J.W., 2003, Statistical methods for survival data analysis, John Wiley & Sons, Canada.
28. Matos, C.A., Ritter, C., Gianola, D., Thomas, D.L., 1993, Bayesian analysis of lamb survival using Monte Carlo numerical integration with importance sampling, J Anim Sci, 71(8), 2047-2054 p.
29. Newman, S.C., 2001, Biostatistical methods in epidemiology, John Wiley & Sons, Inc., New York.
30. Nieto, F.J., Coresh, J., 1996, Adjusting survival curves for confounders: a review and a new method, American Journal of Epidemiology, 143, 1059-1068 p.
31. Özdamar, K., 2003, SPSS ile Biyoistatistik, Kaan Kitabevi, Eskişehir.
32. Pinto, A.E., Andre, S., Soares, J., 1999, Short-term significance of DNA ploidy and cell proliferation in breast carcinoma: a multivariate analysis of prognostic markers in a series of 308 patients, J Clin Pathol, 52, 604-611 p.

33. Ramon, J.M., Escriba, J.M., Casas, I., Benet, J., Iglesias, C., Gavaldà, L., Torras, G., Oromi, J., 1996, Age at first full-term pregnancy, lactation and parity and risk of breast cancer: a case-control study in Spain, *Eur J Epidemiol*, 12(5), 449-53 p.
34. Rowlings, P.A., Williams, S.F., Antman, K.H., Fields, K.K., Fay, J.W., Reed, E., et. al., 1999, Factors correlated with progression-free survival after high-dose chemotherapy and hematopoietic stem cell transplantation for metastatic breast cancer, *JAMA*, 282, 1335-1343 p.
35. Robert, C.P., 2007, *The Bayesian choice from decision-theoretic foundations to computational implementation*, Springer, New York.
36. Robert, C.P. and Casella, G., 2004, *Monte Carlo statistical methods*, Springer, New York.
37. Ross, S.M., c2002, *Simulation*, Academic Press, Amsterdam.
38. Rowe, D.B., 2003, *Multivariate Bayesian statistics models for source separation and signal unmixing*, Chapman and Hall/CRC, New York.
39. SAS Institute, 2002, *SAS for Windows Version 9.0, User Manual*, SAS Institute Inc., Cary, NC, USA.
40. SAS Institute, 2006, *Preliminary capabilities for bayesian analysis in SAS/STATR Software*, SAS Institute Inc., Cary, NC, USA.
41. Şenocak, M., 1998, *Biyostatistik*, İstanbul, İstanbul Üniversitesi.
42. Sinha, D., Ibrahim, J.G., Chen, M.H., 2003, A Bayesian justification of Cox's partial likelihood, *Biometrika* 90(3), 629-641 p.
43. Spiegelhalter, D.J., Abrams, K.R., Myles, J.P., 2004, *Bayesian approaches to clinical trials and health-care evaluation*, John Wiley & Sons, England.

44. Terzi, Y., Cengiz, M.A., Bek, Y., 2005, Cox regresyon modelinde oransal hazard varsayımının artıklarla incelenmesi ve akciğer kanseri hastaları üzerinde uygulanması, Türkiye Klinikleri, 25, 770-775 s.
45. Therneau, T.M. and Grambsch, P.M., 2000, Modelling survival data: extending the Cox model, Springer-Verlag, New York.
46. Van Belle, G., Fisher, L.D., Heagerty, P.J., Lumley, T.S., 2004, Biostatistics: A Methodology for the Health Sciences, John Wiley & Sons, Inc., New York.
47. Wong, M.C.M., Lam, K.F., Lo, E.C.M., 2005, Bayesian analysis of clustered interval-censored data, J Dent Res 84(9), 817-821 p.
48. Yin, G. and Ibrahim, J.G., 2006, Bayesian transformation hazard model, IMS Lecture Notes–Monograph Series, 2nd Lehmann Symposium–Optimality, Vol. 49 170-182 p.
49. Yardımcı, A. ve Erar, A., 2005, Aykırı değer varlığında doğrusal regresyonda değişken seçimine Gibbs örnekleme yaklaşımı, G.Ü. Fen Bilimleri Dergisi, 18(4), 603-611 s.
50. Yu, H., Giai, M., Diamandis, E.P., Katsaros, D., Sutherland, D.J., Levesque, M.A., Roagna, R., Ponzzone, R., Sismondi, P., 1995, Prostate-specific antigen is a new favorable prognostic indicator for women with breast cancer. Cancer Research, 55(10), 2104-2110 p.
51. Yücebaş, S.C., 2006, Hipokrat-I: Bayes ağı tabanlı tıbbi teşhis destek sistemi, Yüksek Lisans Tezi, Başkent Üniversitesi, Fen Bilimleri Enstitüsü.

ÖZGEÇMİŞ

Bireysel Bilgiler

Adı-Soyadı : İmran KURT
Doğum tarihi ve yeri : 1978 Eskişehir
Uyruğu : T.C.

Eğitim Durumu

Lisans	Eskişehir Osmangazi Üniversitesi Fen Edebiyat Fakültesi İstatistik Bölümü	1996-2000
Yüksek Lisans	Trakya Üniversitesi Sağlık Bilimleri Enstitüsü Biyostatistik Yüksek Lisans Programı	2001-2004
Doktora	Eskişehir Osmangazi Üniversitesi Sağlık Bilimleri Enstitüsü Biyostatistik Doktora Programı	2005-

Yabancı Dil İngilizce

Mesleki Deneyim

Görev Unvanı	Görev Yeri	Yıl
Araştırma Görevlisi	Trakya Üniversitesi Tıp Fakültesi	2001-2005
Araştırma Görevlisi	Eskişehir Osmangazi Üniversitesi Tıp Fakültesi	2005-