

YAŞAM ÇÖZÜMLEMESİNDE İYİLEŞME MODELLERİ

CURE MODELS IN SURVIVAL ANALYSIS

PINAR KARA

DOÇ. DR. NİHAL ATA TUTKUN

Tez Danışmanı

Hacettepe Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin
İstatistik Anabilim Dalı için Öngördüğü
YÜKSEK LİSANS TEZİ olarak hazırlanmıştır.

2017

PINAR KARA' nın hazırladığı “**Yaşam Çözümlemesinde İyileşme Modelleri**” adlı bu çalışma aşağıdaki jüri tarafından **İSTATİSTİK ANABİLİM DALI'** nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Prof. Dr. Mehtap AKÇİL OK
Başkan



Doç. Dr. Nihal ATA TUTKUN
Danışman



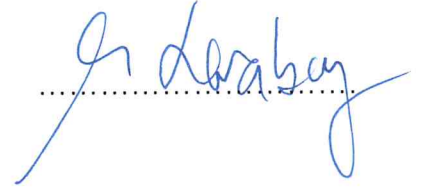
Prof. Dr. Durdu KARASOY
Üye



Doç. Dr. Derya ERSEL
Üye



Yrd. Doç. Dr. Uğur KARABEY
Üye



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

Prof. Dr. Menemşe GÜMÜŞDERELİOĞLU
Fen Bilimleri Enstitüsü Müdürü

YAYINLAMA VE FİKRİ MÜLKİYET HAKLARI BEYANI

Enstitü tarafından onaylanan lisansüstü tezimin/raporumun tamamını veya herhangi bir kısmını, basılı (kağıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanması zorunlu metinlerin yazılı izin alarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

- Tezimin/Raporumun tamamı dünya çapında erişime açılabilir ve bir kısmı veya tamamının fotokopisi alınabilir.**
(Bu seçenikle teziniz arama motorlarında indekslenebilecek, daha sonra tezinizin erişim statüsünün değiştirilmesini talep etmeniz ve kütüphane bu talebinizi yerine getirirse bile, tezinin arama motorlarının önbelleklerinde kalmaya devam edebilecektir.)
- Tezimin/Raporumun tarihine kadar erişime açılmasını ve fotokopi alınmasını (İç Kapak, Özet, İçindekiler ve Kaynakça hariç) istemiyorum.**
(Bu sürenin sonunda uzatma için başvuruda bulunmadığım takdirde, tezimin/raporumun tamamı her yerden erişime açılabilir, kaynak gösterilmek şartıyla bir kısmı ve ya tamamının fotokopisi alınabilir)
- Tezimin/Raporumun 01.07.2018 tarihine kadar erişime açılmasını istemiyorum, ancak kaynak gösterilmek şartıyla bir kısmı veya tamamının fotokopisinin alınmasını onaylıyorum.**
- Serbest Seçenek/Yazarın Seçimi**

05 / 06 / 2017



Pınar Kara

ETİK

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

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

beyan ederim.

05/09/2017



PINAR KARA

ÖZET

YAŞAM ÇÖZÜMLEMESİNDE İYİLEŞME MODELLERİ

Pınar KARA

Yüksek Lisans, İstatistik Bölümü

Tez Danışmanı: Doç. Dr. Nihal ATA TUTKUN

Haziran 2017, 86 sayfa

Yaşam çözümlemesi, bir kitlede ilgilenilen olayın meydana gelme sürelerinin araştırılması için kullanılan istatistiksel yöntemler topluluğu olarak tanımlanmaktadır. Yaşam çözümlemesinde, klasik istatistiksel yöntemlerden farklı birçok istatistiksel model kullanılmaktadır.

Yaşam çözümlemesinde sıklıkla kullanılan Cox regresyon modeli, çalışmadaki tüm birimlerin ilgilenilen olayı yaşayacağı varsayımına dayanmaktadır. Bu varsayım, kitlede ilgilenilen olayı yaşamayan çok sayıda birim bulunduğunda geçerli olmayabilir. Özellikle son yıllarda kanser tedavilerindeki gelişmeler sayesinde, çalışma grubundaki birçok hasta iyileşmektedir. Bu nedenle de iyileşmiş kısmı inceleyen “iyileşme modelleri”nin kullanımı önem kazanmıştır.

Bu çalışmada iyileşme modelleri teorik açıdan incelenmiş ve uygulama bölümünde Glioma ve mide kanseri verilerine uygulanmıştır. Glioma ve mide kanseri veri kümeleri için orantılı tehlikeler varsayımı incelenmiş, glioma veri kümesi için orantılı tehlikeler varsayımı sağlanırken, mide kanseri veri kümesi için bu varsayımın sağlanmadığı görülmüştür. Orantılı tehlikeler varsayımını sağlayan glioma veri kümesine Cox karma iyileşme modeli uygulanmıştır. Lojit bağlantı fonksiyonu kullanıldığında iyileşmiş kısım için hiçbir değişken anlamlı bulunmaz iken iyileşmemiş kısım için malignite düzeyi ve yaş değişkenlerinin istatistiksel olarak anlamlı değişkenler olduğu görülmüştür. Cloglog

bağlantı fonksiyonu kullanıldığında ise malignite düzeyi ve yaş değişkenlerinin hem iyileşmemiş hem de iyileşmiş kısım için istatistiksel olarak anlamlı değişkenler olduğu görülmüştür. Orantılı tehlikeler varsayımını sağlamayan mide kanseri veri kümesi için hızlandırılmış başarısızlık süresi karma iyileşme modeli kullanılmıştır. Cloglog bağlantı fonksiyonu kullanıldığında hem iyileşmiş hem de iyileşmemiş kısım için metastaz değişkeni istatistiksel olarak anlamlı bulunmuştur. İki veri kümesinde de elde edilen sonuçlar kullanılarak farklı değişken düzeyleri için iyileşme oranları elde edilmiştir. İyileşme modellerinden elde edilen sonuçlar ile klasik yaşam modellerinden elde edilen sonuçlar incelenmiş ve yorumlanmıştır.

Anahtar Kelimeler: Cox regresyon modeli, hızlandırılmış başarısızlık süresi modeli, iyileşme modelleri, orantılı tehlikeler, yaşam çözümlenmesi.

ABSTRACT

CURE MODELS IN SURVIVAL ANALYSIS

Pınar KARA

Master of Science, Department of Statistics

Supervisor: Assoc. Prof. Dr. Nihal ATA TUTKUN

June 2017, 86 pages

Survival analysis is generally defined as a group of statistical methods for analyzing the data where the outcome variable is the time until the occurrence of an event of interest. Various statistical models differ than classical models have been used in survival analysis.

The Cox regression model, which is commonly used in survival analysis, assumes that every subject in the study will eventually experience the event of interest. However, this assumption may not be true when there are a lot patients in the population who are cured from their diseases and never experienced the event of interest. Especially with the improvements at cancer therapies in the recent years, lots of patients could be survivors of their diseases and because of this using “cure models” to analyze this cured fraction has become important.

In this study, cure models were examined theoretically and were applied to the glioma and stomach cancer data. First, to decide which type of cure models should be used, proportional hazards assumption was examined for both glioma and stomach cancer data sets. The assumption of proportional hazards was valid for the glioma data, whereas it is not valid for stomach cancer data. Cox cure model was considered for the glioma data set. When logit link function was selected for this Cox cure model, none of

the covariates were found to be statistically significant for the cured part while malignancy grade and age were statistically significant covariates for uncured part. When cloglog link function was used, malignancy grade and age covariates were statistically significant for both cured and uncured parts. For stomach cancer data, accelerated failure time model is considered due to the violation of proportional hazards assumption. Cloglog link function was selected for this model, it was seen that metastasis covariate was statistically significant for cured and uncured parts. The results which were obtained from these cure models and the ones that were obtained from classical survival models were examined and interpreted.

Keywords: Cox regression model, accelerated failure time model, cure models, proportional hazards, survival analysis.

TEŐEKKÜR

Tezimin oluŐmasında bana en büyük desteęi veren, alıŐmalarımda bana yol gÖsteren ve teŐvik eden, bilgisini ve tecrübelerini benimle paylaŐan danıŐmanım Sayın Do. Dr. Nihal Ata Tutkun'a,

Her konuda yanımda olan, hibir desteęi ve yardımı esirgemeyen annem Sevgi Kara'ya,

Deęerli önerileri ile tezime katkıda bulunan Sayın Prof. Dr. Durdu Karasoy, Sayın Prof. Dr. Mehtap Akil Ok, Sayın Do. Dr. Derya Ersel ve Sayın Yrd. Do. Dr. Uęur Karabey'e,

tüm samimiyetimle teŐekkür ederim.

İÇİNDEKİLER

	<u>Sayfa</u>
ÖZET.....	i
ABSTRACT.....	iii
TEŞEKKÜR.....	v
İÇİNDEKİLER.....	vi
ÇİZELGELER.....	viii
ŞEKİLLER	x
KISALTMALAR	xi
1. GİRİŞ.....	1
2. YAŞAM ÇÖZÜMLEMESİ	3
2.1. Giriş	3
2.2. Yaşam Tehlike Fonksiyonları	4
2.3. Kaplan-Meier Yaşam Tahmini	5
2.4. Tehlike ve Birikimli Tehlike Fonksiyonları	6
2.5. Yaşam Fonksiyonlarını Karşılaştırmada Parametrik Olmayan Testler	7
2.6. Cox Regresyon Modeli	8
2.6.1. Olabilirlik Fonksiyonunun Elde Edilmesi	9
2.6.1.1. Kısmi Olabilirlik Yöntemi	9
2.6.1.2. Breslow Yöntemi	10
2.6.1.3. Efron Yöntemi	11
2.6.2. Temel Tehlike ve Yaşam Fonksiyonlarının Elde Edilmesi	12
2.7. Hızlandırılmış Başarısızlık Süresi Modelleri	14
2.7.1. Üstel Dağılım	15
2.7.2. Weibull Dağılımı	15
2.7.3. Gamma Dağılımı	15
2.7.4. Genelleştirilmiş Gamma Dağılımı	16
2.7.5. Log-Normal Dağılım	17
2.7.6. Log-Lojistik Dağılım	17
3. İYİLEŞME MODELLERİ.....	19
3.1. Giriş	19

3.2. İyileşme Modeli ve Cox Regresyon Modeli Arasındaki Farklar	22
3.3. İyileşme Modelleri	23
3.3.1. Karma İyileşme Modelleri.....	25
3.3.1.1. Cox Karma İyileşme Modeli.....	27
3.3.1.2. Yarı Parametrik Hızlandırılmış Başarısızlık Süresi Karma İyileşme Modeli...	31
3.3.1.3. Yarı Parametrik Hızlandırılmış Tehlikeler Karma İyileşme Modeli.....	33
3.3.1.4. Parametrik Karma İyileşme Modeli.....	39
3.3.2. Karma Olmayan İyileşme Modelleri.....	42
4. UYGULAMA.....	45
4.1. Glioma Veri Kümesi.....	45
4.1.1. Parametrik Olmayan Yaşam Çözümlemesi Sonuçları.....	46
4.1.2. Cox Regresyon Modeli Sonuçları	48
4.1.2.1. Orantılı Tehlikeler Varsayımının İncelenmesi.....	49
4.1.3. Cox İyileşme Modeli Sonuçları.....	53
4.2. Mide Kanseri Veri Kümesi.....	60
4.2.1. Parametrik Olmayan Yaşam Çözümlemesi Sonuçları.....	60
4.2.2. Cox Regresyon Modeli Sonuçları.....	63
4.2.2.1 Orantılı Tehlikeler Varsayımının İncelenmesi.....	64
4.2.3. Tabakalı Cox Regresyon Modeli Sonuçları.....	66
4.2.4. Genişletilmiş Cox Regresyon Modeli Sonuçları.....	68
4.2.5. Hızlandırılmış Başarısızlık Süresi Modelleri Sonuçları.....	69
4.2.6. Hızlandırılmış Başarısızlık Süresi Karma İyileşme Modeli Sonuçları.....	71
5. SONUÇ VE ÖNERİLER.....	75
KAYNAKLAR.....	81
ÖZGEÇMİŞ.....	86

ÇİZELGELER

	<u>Sayfa</u>
Çizelge 4.1. Glioma çalışmasında kullanılan değişkenler ve düzeyleri.....	45
Çizelge 4.2. Glioma verisi için log-rank testi sonuçları.....	48
Çizelge 4.3. Glioma verisi için CRM'den elde edilen sonuçlar.....	49
Çizelge 4.4. Glioma verisi için Schoenfeld artıkları ile yaşam süresinin rankı arasındaki korelasyon testi.....	51
Çizelge 4.5. Cinsiyet değişkeni için CRM'den elde edilen sonuçlar.....	52
Çizelge 4.6. Zamana bağlı cinsiyet değişkeni için CRM'den elde edilen sonuçlar.....	52
Çizelge 4.7. Glioma verisinde tedavi değişkeni için CRM sonuçlar.....	53
Çizelge 4.8. Glioma verisinde zamana bağlı tedavi değişkeni için CRM sonuçları.....	53
Çizelge 4.9. Glioma verisinde iyileşmiş kısım için lojit Cox iyileşme modeli sonuçları...	54
Çizelge 4.10. Glioma verisinde iyileşmemiş kısım için lojit Cox iyileşme modeli sonuçları.....	55
Çizelge 4.11. Glioma verisinde iyileşmiş kısım için cloglog Cox iyileşme modeli sonuçları.....	58
Çizelge 4.12. Glioma verisinde iyileşmemiş kısım için cloglog Cox iyileşme modeli sonuçları.....	58
Çizelge 4.13. Mide kanseri çalışmasında kullanılan değişkenler ve düzeyleri.....	60
Çizelge 4.14. Mide kanseri verisi için log-rank testi sonuçları.....	63
Çizelge 4.15. Mide kanseri verisi için CRM'den elde edilen sonuçlar.....	63
Çizelge 4.16. Mide kanseri verisi için Schoenfeld artıkları ile yaşam süresinin rankı arasındaki korelasyon testi.....	64
Çizelge 4.17. Metastaz değişkeni için CRM sonuçları.....	65
Çizelge 4.18. Zamana bağlı metastaz değişkeni için CRM sonuçları.....	65
Çizelge 4.19. Etkileşimsiz model için CRM sonuçları.....	66
Çizelge 4.20. Metastazı olmayan hastalar için etkileşimli CRM sonuçları.....	67
Çizelge 4.21. Metastazı olan hastalar için etkileşimli CRM sonuçları.....	67
Çizelge 4.22. Mide verisi için genişletilmiş CRM sonuçları.....	69
Çizelge 4.23. HBS modelleri için model karşılaştırma kriterleri.....	71
Çizelge 4.24. Mide kanseri verisi için Weibull HBS modeli sonuçları.....	71

Çizelge 4.25. Mide kanseri verisinde iyileşmiş kısım için cloglog HBS iyileşme modeli sonuçları.....	72
Çizelge 4.26. Mide kanseri verisinde iyileşmemiş kısım için cloglog HBS iyileşme modeli sonuçları.....	72
Çizelge 4.27. Sadece metastaz ve radyoterapi değişkenleri kullanıldığında iyileşmiş kısım için elde edilen cloglog HBS iyileşme modeli sonuçları.....	73
Çizelge 4.28. Sadece metastaz ve radyoterapi değişkenleri kullanıldığında iyileşmemiş kısım için elde edilen cloglog HBS iyileşme modeli sonuçları.....	73

ŞEKİLLER

	<u>Sayfa</u>
Şekil 3.1. Yaşam eğrisi örneği.....	21
Şekil 3.2. Olay zamanlarının kanser ve diğer akut hastalıklar için dağılımı.....	24
Şekil 3.3. CRM, HBS ve HT modelleri için tehlike eğrileri.....	34
Şekil 4.1. Glioma verisi için Kaplan-Meier yaşam eğrisi.....	46
Şekil 4.2. Glioma verisinde kullanılan açıklayıcı değişkenlerin Kaplan-Meier eğrileri...	47
Şekil 4.3. Glioma verisinde açıklayıcı değişkenler için $\ln(-\ln)$ yaşam eğrileri.....	50
Şekil 4.4. Seçilen değişken düzeyleri için yaşam eğrileri 1.....	56
Şekil 4.5. Seçilen değişken düzeyleri için yaşam eğrileri 2.....	57
Şekil 4.6. Mide kanseri verisi için Kaplan-Meier yaşam eğrisi.....	61
Şekil 4.7. Mide kanseri çalışmasındaki açıklayıcı değişkenler için Kaplan-Meier yaşam eğrileri.....	62
Şekil 4.8. Mide kanseri verisinde metastaz değişkeni için $\ln(-\ln)$ grafiği.....	65
Şekil 4.9. Mide kanseri verisi için Cox-Snell artıkları.....	70
Şekil 4.10. Farklı değişken düzeyleri için yaşam eğrileri 3.....	74

KISALTMALAR

Kısaltmalar

AIC	Akaike Bilgi Kriteri
BIC	Bayesci Bilgi Kriteri
CRM	Cox Regresyon Modeli
EÇO	En Çok Olabilirlik
HT	Hızlandırılmış Tehlikeler
HBS	Hızlandırılmış Başarısızlık Süresi
TO	Tehlike Oranı

1. GİRİŞ

Yaşam çözümlemesi, bir kitlede ilgilenilen olayın meydana gelme sürelerinin araştırılması için kullanılan istatistiksel yöntemler topluluğudur. Olay gerçekleşene kadar geçen süre çözümlemesi (time-to-event analysis) olarak da adlandırılmaktadır. İlgilenilen olayın sonucu, iyi tanımlanmış bir başlangıç noktası ile iyi tanımlanmış bir bitiş noktası arasında geçen süredir. Yaşam çözümlemesi ölçümleri için,

1. Sürenin ölçülmesi için kararlaştırılmış bir ölçek,
2. Sürenin ölçülmesi için kesin bir başlangıç noktası,
3. İlgilenilen olaya ait net bir tanım

gerekmektedir. Yaşam çözümlemesinin sağlık bilimleri, mühendislik, sosyoloji, işletme, ekonomi, uluslararası ilişkiler, eğitim başta olmak üzere birçok alanda uygulamaları görülebilir. Yaşam çözümlemesinin en sık kullanıldığı sağlık bilimleri alanındaki çalışmalar için bazı örnekler;

- Kanserin teşhisinden hastanın ölümüne kadar geçen süre,
- Bir kanser araştırmasında, nüksetme gerçekleşene kadar geçen süre,
- HIV bulaşmasından AIDS hastalığına kadar geçen süre,
- Yeni doğanların sütten kesilmesine kadar geçen süre,
- Ağır yanıklar nedeniyle hastaneye kaldırılan hastaların enfekte olmasına kadar geçen süre,
- İki kan bağışlama girişimi arasında geçen süre

biçiminde verilebilir [1].

Kanser hastaları için yaşam çözümlemesi çalışmalarında ilgilenilen, çoğunlukla kanserin teşhisinden ölüme kadar geçen süredir. Bu durumda başlangıç noktası kanserin teşhisi, ilgilenilen olay ise ölümdür. İlgilenilen olay, çalışmada araştırılan kanser nedeniyle ölüm ya da herhangi bir nedenden dolayı ölüm olarak tanımlanabilir. Zaman ölçüsü, teşhisten itibaren geçen süredir ve bu, her birey çalışmaya sıfır noktasında gireceği için, en kolay ölçüdür. Kanser çalışmalarını içeren yaşam çözümlemesi uygulamalarında, teşhisten itibaren geçen süre seçimi çok belirgin olduğu için zamanın ölçümü genellikle bir sorun teşkil etmemektedir. Ancak yaş gibi bir ölçek

kullanmak istenilirse sorunlar ortaya çıkabilmektedir. Bu durum, çalışmayı karmaşık bir hale getirebilir çünkü artık bireyler çalışmaya sıfır zaman noktasından girmeyeceklerdir ve de bireyler arasında çalışmaya girme zamanı farklılık gösterecektir [1].

Yaşam çözümlemesinin temel özelliği, çalışmadaki her birey için ilgilenilen olayın gözlemlenememesidir. Bu nedenden dolayı, yaşam çözümlemesinde farklı istatistiksel yöntemler kullanılması gerekmektedir [1].

Yaşam çözümlemesinin temel kavramlarından biri olan “yaşam süresi” ya da “başarısızlık süresi” canlı ya da cansız bir nesnenin, belirli bir başlangıç noktasından, ilgilenilen olayı yaşamasına kadar geçen süre olarak tanımlanmaktadır. Yaşam süresi üzerinde etkili olan faktörleri belirlemek amacıyla en sık tercih edilen yaşam modeli Cox regresyon modeli (CRM)’dir. Bu model orantılı tehlikeler modeli (proportional hazards model) olarak da adlandırılmaktadır.

CRM, kitledeki her birimin ilgilenilen olaya yatkın olduğu ve ilgilenilen olayı elbet yaşayacağı varsayımına dayanmaktadır. Bu varsayım her zaman doğru olmayabilir. Özellikle son yıllarda geliştirilen tedaviler sayesinde birçok hastanın hayatta kalabildiği kanser türlerinde, CRM gibi iyileşme olasılığını hesaba katmayan modelleri kullanmak uygun olmayabilir. İyileşmiş kısmı içeren yaşam modelleri “iyileşme modelleri (cure models)” olarak adlandırılmaktadır ve son yıllarda yaşam verilerini çözümlmek için kullanılan önemli bir istatistiksel model olmuştur [1].

Bu çalışmanın amacı, yaşam çözümlemesinde iyileşme modellerini incelemek ve gerçek veri kümelerini kullanarak konu ile ilgili uygulama yapmaktır. Çalışmanın ikinci bölümünde yaşam çözümlemesine ilişkin temel bilgiler verilmiştir. Üçüncü bölümde yaşam çözümlemesinde iyileşme modellerinin kullanımı incelenmiş, ve iyileşme modelleri anlatılmıştır. Dördüncü bölümde ise glioma (beyin tümörü) ve mide kanseri verilerine iyileşme modelleri ve diğer yaşam modelleri uygulanmış ve bu modeller için elde edilen sonuçlar yorumlanmıştır.

2. YAŞAM ÇÖZÜMLEMESİ

2.1. Giriş

Sağlık bilimleri alanında, özellikle kanser konusunda yapılan araştırmalar gün geçtikçe hız kazanmaktadır. Kanser çalışmalarının çoğunda, değerlendirilen asıl çıktı ilgilenilen olayın meydana gelme süresidir. Bu süre genellikle “yaşam süresi” olarak adlandırılmaktadır. Eğer ilgilenilen olay çalışma grubundaki bütün bireylerde meydana gelirse, birçok istatistiksel çözümleme yöntemi kullanılabilir. Ancak, bazı bireylerin gözlem süresinin sonunda ilgilenilen olayı yaşamamış olması ve bundan dolayı bu bireylerde ilgilenilen olayın gerçekleşme sürelerinin bilinmemesi durumuyla sıklıkla karşılaşmaktadır. Ayrıca, yaşam verileri genellikle normal dağılıma uygunluk göstermeyen çarpık dağılıma sahiptir. Verilerin bu özellikleri, yaşam çözümlemesinin ortaya çıkmasına neden olmuştur [2].

Tıbbi çalışmaların çoğunda, “ölüm” ilgilenilen olaydır. Kanser çalışmalarındaki önemli ölçülerden biri ise, bireyin tedaviye cevap vermesi ile hastalığın nüksetmesi arasında geçen süre ya da nüksetmesiz hayatta kalımdır, bu durum hastaliksız yaşam süresi olarak da adlandırılmaktadır. Yaşam çözümlemesinde ilgilenilen olayın ne olduğunun ve de gözlem süresinin başlangıç ve bitiş zamanlarının net olarak belirtilmesi önemlidir.

Yaşam çözümlemesinde karşılaşılan zorluklar büyük ölçüde, çalışma grubunda sadece bir kısım bireyin ilgilenilen olayı yaşaması ve olayı yaşamayan alt küme için yaşam sürelerinin bilinmemesinden kaynaklanmaktadır. Bu durum ‘durdurma (censoring)’ olarak adlandırılır ve meydana gelme nedenleri şu şekilde sıralanabilir; (a) birey nüksetme ya da ölüm gibi ilgilenilen olayı çalışma süresi boyunca yaşamamış olabilir, (b) çalışma sürecinde hastanın takibi kaybedilebilir, (c) hasta, takibin devam edilmesini imkansız hale getiren bir olay yaşayabilir. Durdurulmuş yaşam süreleri, gerçek (ama bilinmeyen) olayın meydana geliş sürelerinin olduğundan daha az tahmin edilmesine yol açabilir. Bir hastanın yaşam sürecini ele alırsak, ilgilenilen olayın gerçekleşme zamanı (gerçekleşeceği varsayılarak) takip süresinin bitiminden sonra olabilir. Bu durum ‘sağdan durdurma’ olarak adlandırılır. Ayrıca bir olayın ya da durumun varlığını gözlemleyebiliyor ancak tam olarak nerede başladığını bilemiyorsak da durdurma gerçekleşir. Örneğin, birincil tümör ameliyat ile alındıktan sonra, kanserin nüksetme

zamanını inceleyen bir araştırma ele alınsın. Hastalar nüksetme olup olmadığını gözlemek amacıyla, ameliyattan 3 ay sonra muayane edildiğinde kanserin nüksettiği hastalar 'soldan durdurulmuş' olarak alınır, çünkü nüksetme süresi 3 aydan kısadır. Ayrıca 'aralıklı durdurulmuş gözlemler' olabilir, yani hasta önce gözlem içi sonra gözlem dışı olmuştur. Önceki örnek tekrar ele alınmış ve hastalar ameliyattan sonraki 6. ayda da muayane edilmiş olsun. 3. ayda nüksetme gözlemlenmeyen hastalısız bireyin, 3. ve 6. ay arasında takibi kaybedilirse, bu gözlem 'aralıkta durdurulmuş' olarak değerlendirilir. Çoğu yaşam verisi sağdan durdurulmuş gözlemleri içermektedir, ancak aralıklı ve soldan durdurulmuş veriler için de yöntemler mevcuttur [3]. Veri kümesinde durdurulmuş gözlemlerin varlığı, yaşam çözümlemesinin kullanılması gerektiğini göstermektedir.

2.2. Yaşam ve Tehlike Fonksiyonları

Yaşam verileri genellikle yaşam (survival) ve tehlike (hazard) fonksiyonları kullanılarak tanımlanır ve modellenir. Yaşam fonksiyonu hayatta kalım fonksiyonu olarak da adlandırılmaktadır. Yaşam fonksiyonu $S(t)$, T rastlantı değişkeni yaşam süresini göstermek üzere T 'nin t 'den daha büyük olma olasılığı olarak tanımlanmaktadır. t 'nin farklı değerleri için yaşam olasılıkları, yaşam verileri için önemli bilgiler sağlamaktadır. Bu değerler, bir kuşak (cohort) çalışması için yaşam deneyimini doğrudan açıklayabilmektedirler. Yaşam fonksiyonu;

$$S(t) = P(T > t) = \int_t^{\infty} f(x)dx, \quad 0 < t < \infty \quad (2.1)$$

olarak ifade edilmektedir. $f(t)$, yaşam süresi T 'nin olasılık yoğunluk fonksiyonu, $F(t)$ ise dağılım fonksiyonudur ve sırasıyla aşağıdaki biçimde gösterilmektedir;

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T \leq t + \Delta t)}{\Delta t}, \quad 0 < t < \infty \quad (2.2)$$

$$F(t) = P(T \leq t) = \int_0^t f(x)dx, \quad 0 < t < \infty$$

Tehlike fonksiyonu, t zamanına kadar yaşayan birimin $[t, t + \Delta t]$ aralığında yaşamının sona ermesi riskinin bir tanımıdır. Herhangi bir zamandaki başarısızlık riski olarak

tanımlanmaktadır ve genellikle $h(t)$ ya da $\lambda(t)$ olarak gösterilmektedir. Yaşam fonksiyonunun aksine, tehlike fonksiyonu ilgilenilen olayı yaşama olasılığına odaklanılır. Bu durum hem koşullu başarısızlık hızı hakkında bilgi sahibi olunmasını sağlar hem de yaşam modelinin belirlenmesine yardımcı olur. Kısacası yaşam fonksiyonu, olayın gerçekleşmemesini yansıtırken, tehlike fonksiyonu, olayın o anda meydana gelmesi ile ilişkilendirilmektedir [2]. Tehlike fonksiyonu;

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T \leq t + \Delta t / T > t)}{\Delta t} \quad (2.3)$$

biçimindedir.

2.3. Kaplan-Meier Yaşam Tahmini

Yaşam olasılığı, Kaplan-Meier (ya da çarpım-limit (product-limit)) yöntemi ile hem durdurulmuş hem de durdurulmamış yaşam süreleri kullanılarak parametrik olmayan bir biçimde tahmin edilebilir [4]. k tane hastanın, takip sürecinde, $t_1 < t_2 < \dots < t_k$ ayrık zamanlarında ilgilenilen olayı yaşadığı varsayalım. Olayların birbirinden bağımsız olarak gerçekleştiği varsayıldığından, birikimli yaşam olasılığı bir zaman aralığından bir sonraki zaman aralığına kadar yaşam olasılıklarının tamamı çarpılarak elde edilebilir. t_j anındaki yaşam fonksiyonu $S(t_j)$, t_{j-1} zamanında yaşam fonksiyonu $S(t_{j-1})$ kullanılarak hesaplanır. n_j ; t_j anından hemen önce hayatta olan hastaların sayısı ve d_j ise t_j anında gerçekleşen olayların sayısıdır. Buna göre,

$$S(t_j) = S(t_{j-1}) \left(1 - \frac{d_j}{n_j} \right) \quad (2.4)$$

biçiminde elde edilebilir. Burada $t_0 = 0$ ve $S(0) = 1$ 'dir. $S(t)$ 'nin aldığı değer iki olay arasında sabittir ve bu nedenle tahmin olasılığı, sadece olayın meydana geldiği anlarda değeri değişen bir adım fonksiyonudur. Her birey ilgilenilen olayı yaşıyorsa (durdurma yok ise), bu tahmin edici t zamanında ilgilenilen olayı yaşamamış birey sayısının, çalışmaya katılan toplam birey sayısına oranına indirgenebilir.

Kaplan-Meier yaşam eğrisi ise Kaplan-Meier yaşam olasılıklarının zamana karşı çizilen bir grafiğidir ve ortanca yaşam süresi gibi ölçülerin tahmin edilmesinde kullanılabilecek faydalı bir bilgi özeti sunmaktadır [2].

2.4. Birikimli Tehlike Fonksiyonu

Yaşam fonksiyonu $S(t)$ ve tehlike fonksiyonu $h(t)$ arasında, Eşitlik 2.5 ile ifade edilebilen bir ilişki vardır;

$$h(t) = -\frac{d}{dt} [\log S(t)]. \quad (2.5)$$

Buradaki önemli nokta, $S(t)$ ya da $h(t)$ 'den biri biliniyorsa diğerinin kolaylıkla hesaplanabileceğidir. Dolayısıyla, bu iki fonksiyondan herhangi biri, istatistiksel çözümlemenin temeli olabilir.

$S(t)$ 'nin aksine, birikimli tehlike fonksiyonu olarak bilinen $H(t)$ 'yi tahmin etmenin basit bir yolu yoktur. Birikimli tehlike fonksiyonu,

$$H(t) = \int_0^t h(x) dx = -\log S(t)$$

olarak ifade edilir ve bu fonksiyon artan ve sağdan sürekli bir fonksiyondur. $H(t)$ 'nin yorumlanması zordur ancak $H(t)$ 'yi düşünmenin en kolay yolu ölümlülüğün birikimli gücü ya da olay tekrarlanabilir bir süreç ise, her bir birey için t zamanına kadar gözlemlenmesi beklenen olay sayısıdır. $H(t)$, $h(t)$ 'yi tahmin etmek için aracı bir ölçü ve modelin geçerliliğini değerlendirmek için bir araç olarak kullanılır. $H(t)$ 'yi tahmin etmek için kullanılan temel bir parametrik olmayan yöntem Nelson-Aalen tahmin edicisidir [3].

Tehlike fonksiyonlarını tahmin etmek için kullanılan bir başka yaklaşım, yaşam sürelerinin belirli bir istatistiksel dağılımı izlediğini varsaymaktır. Fonksiyonun tamamen parametrik dağılımlar kullanılarak belirlenmesi önemli bir modelleme yöntemidir. Bu amaçla, sabit, artan, azalan ya da U-biçimli tehlike fonksiyonlarına izin veren üstel, Weibull, lognormal, loglojistik ve Gamma dağılımları kullanılabilir.

2.5. Yaşam Fonksiyonlarının Karşılaştırılmasında Kullanılan Parametrik Olmayan Testler

İki ya da daha fazla grup arasındaki yaşam olasılığı, parametrik olmayan bir test yardımıyla karşılaştırılabilir. Log-rank testi [5] iki ya da daha fazla grubun yaşam olasılıklarını karşılaştırmak amacıyla kullanılan bir yöntemdir. Gruplar, tedavi yöntemleri ya da tanısal gruplar (kanser evresi) olabilir. Bu yöntem, her bir olay zamanında, her bir grup için, gruplar arasında bir farklılık yok ise bir önceki olaydan sonra meydana gelmesi beklenen olayların sayısını hesaplar. Elde edilen değerler, her bir grup için beklenen olay sayısını vermesi için toplanır. E_i , i . grup için beklenen olay sayısı ve O_i gözlemlenen olay sayısı olmak üzere, log-rank test istatistiği yaklaşık olarak,

$$\chi^2 = \sum_{i=1}^g \frac{(O_i - E_i)^2}{E_i}$$

ile hesaplanır. Bu değer, g grup sayısını göstermek üzere, $g-1$ serbestlik dereceli ki-kare dağılımı ile karşılaştırılır. Böylelikle, yaşam olasılıkları arasındaki farkın istatistiksel olarak anlamlı olup olmadığını değerlendirmek için bir p değeri elde edilebilir [2].

Sadece iki grup karşılaştırıldığında, log-rank testi iki grubun tehlike oranının 1'e eşit olduğu sıfır hipotezini test eder. Tehlike oranı (TO) iki gruptaki göreceli yaşam deneyiminin ölçüsüdür ve,

$$TO = \frac{O_1/E_1}{O_2/E_2} \quad (2.6)$$

ile tahmin edilebilir. Burada O_i/E_i i grubu için tahmin edilen göreceli (artık) tehlikedir [6]. Tehlike oranının bire eşit olması yaşam olasılıkları açısından gruplar arasında bir farklılık olmadığını ifade eder. Ancak, tehlike oranını CRM gibi bir regresyon modelleme yöntemi kullanarak tahmin etmek daha iyi sonuçlar vermektedir [2].

Grupları yaşam olasılıkları bakımından karşılaştırmak için başka parametrik olmayan testler de kullanılabilir [7]. Ancak orantılı tehlikeler varsayımının sağlandığı durumlarda log-rank yönteminin, diğer yöntemlere göre daha sağlam olduğu kabul edilmektedir [3].

2.6. Cox Regresyon Modeli

İstatistikçi Sir David Cox tarafından 1972 yılında önerilen CRM istatistik literatüründe çok önemli bir yere sahiptir. Cox'un orijinal çalışmasından elde edilen sonuçlar, farklı alanlardaki birçok probleme uygulanmaktadır. Cox'un bu çalışmaları, 20. yüzyılın en önemli istatistiksel başarılarından biri olarak görülmektedir. CRM'nin mantığı basit ve sadedir. $h_0(t)$ temel tehlike fonksiyonu, β regresyon parametreleri ve x açıklayıcı değişkenler iken, i . birey için tehlike fonksiyonu;

$$h_i(t) = h_0(t) \exp(\beta' x_i) \quad (2.7)$$

biçiminde ifade edilir. CRM için tehlike fonksiyonlarının oranının orantılı olduğu varsayımı vardır. i . ve j . bireye ait iki tehlike fonksiyonu arasındaki oran;

$$\frac{h_i(t)}{h_j(t)} = \exp(\beta'(x_i - x_j)) \quad (2.8)$$

olarak yazılabilir ve bu oran zaman boyunca sabittir.

CRM'de orantılı tehlikeler varsayımı bulunmaktayken, temel tehlike oranı $h_0(t)$ 'nin bilinmediği ve parametreleştirilmediği varsayılır. Bu nedenle CRM, yarı parametrik bir modeldir [8].

CRM aşağıdaki gibi de yazılabilir;

$$h_i(t) = \exp(\beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki}) h_0(t).$$

Eğer tehlike oranını logaritma kullanarak ifade edilirse,

$$\log\left(\frac{h_i(t)}{h_0(t)}\right) = \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki}$$

biçiminde elde edilir. CRM'nin farklı biçimlerine bakıldığında hiçbirinin β_0 sabit terimini içermediği görülmektedir. Bu terim, CRM'de temel tehlike fonksiyonunun içine katılmıştır. CRM'den temel tehlike fonksiyonu doğrudan tahmin edilemez [8].

2.6.1. Olabilirlik Fonksiyonunun Elde Edilmesi

2.6.1.1. Kısmi Olabilirlik

Kısmi olabilirlik yöntemi, ilgilenilen olayın meydana geliş zamanları (ya da başarısızlık zamanları) arasındaki aralıkların, açıklayıcı değişkenler ve tehlike oranı arasındaki ilişki hakkında herhangi bir bilgi sağlamadığı varsayımına dayanmaktadır [6]. Bu, CRM için temel tehlike fonksiyonunun doğrudan parametreleştirilememesinin sebebidir. Bu nedenle, bu oranın rastgele bir biçimi olduğu ve Collett'in de belirttiği gibi [6] iki başarısızlığın gerçekleşmesi arasındaki süre boyunca sıfır olabileceği varsayılır. Ayrıca kısmi olabilirlik fonksiyonuna bilgi sağlayan, başarısızlıklar arasındaki aralıklar değil, sıralı başarısızlık zamanlarıdır. Bu, olabilirlik fonksiyonlarının türetilmesi için gerçek yaşam sürelerinin kullanıldığı parametrik yöntemlere terstir. CRM mevcut verilerin sadece bir kısmını kullandığından ($h_0(t)$ tahmin edilmez), CRM için olabilirlik fonksiyonu bir 'kısmi' olabilirlik fonksiyonudur [8].

k ayrık başarısızlık zamanlı, n boyutlu bir veri kümesinin kısmi olabilirlik fonksiyonunu elde etmek için, t_i 'ler i. birey için başarısızlık zamanlarını ifade etmek üzere veriler ilk olarak $t_1 < t_2 < \dots < t_k$ gibi, sıralı başarısızlık süreleri biçiminde sıralanılır. Her bir durdurulmamış gözlemin, ilgilenilen olayı farklı bir zamanda yaşadığı, aynı anda gerçekleşen olay bulunmadığı varsayılır. Durdurma gösterge değişkeni δ_i olarak tanımlanmaktadır ve δ_i sağdan durdurulma durumunda 0, durdurulmamış olması halinde ise 1 değerini almaktadır. Sıralı başarısızlık süreleri, x açıklayıcı değişkenlerinin bir fonksiyonu olarak modellenir [8].

t zamanında ilgilenilen olayı yaşama riski olan bireylerin sayısı, $R_i(t)$ risk kümesi ile gösterilirse, j. bireyin T_i anında başarısızlığı yaşama olasılığı;

$$\Pr(t_j = T_i | R(t_i)) = \frac{e^{\beta x_i}}{\sum_{j \in R(t_i)} e^{\beta x_j}} \quad (2.9)$$

biçiminde ifade edilebilir. Burada toplam operatörü kullanılarak risk kümesindeki tüm gözlemlerin toplamı alınmıştır. Bu eşitlikteki koşullu olasılıkların çarpımı, kısmi olabilirlik fonksiyonunun Eşitlik 2.10.'daki gibi elde edilmesini sağlar:

$$L = \prod_{i=1}^k \left[\frac{e^{\beta' x_i}}{\sum_{j \in R(t_i)} e^{\beta' x_j}} \right] \quad (2.10)$$

ve buna karşılık gelen log-olabilirlik fonksiyonu;

$$\log L = \sum_{i=1}^k \left[\beta' x_i - \log \sum_{j \in R(t_i)} e^{\beta' x_j} \right] \quad (2.11)$$

olur. Log-olabilirlik fonksiyonu maksimize edilerek, β 'ların tahmini elde edilebilir. Bu yöntem, kısmi en çok olabilirlik tahmini olarak da adlandırılmaktadır. Bu olabilirlik fonksiyonu, gerçek olabilirlik fonksiyonu değildir [6]. Bunun nedeni, durdurulmuş ve durdurulmamış gözlemlerin gerçek yaşam sürelerinin olabilirliğe doğrudan dahil edilmemesidir. Bu durum, başarısızlık süreleri arasındaki aralıkların, açıklayıcı değişkenler arasındaki ilişkiye ve sıralı olay sürelerine herhangi bir bilgi sağlamadığını gösterdiğinden dolayı önemlidir. Cox [9, 10] kısmi en çok olabilirlik tahmininin, en çok olabilirlik (EÇO) tahminleri ile aynı özelliklere sahip parametre tahminleri ürettiğini göstermiştir. EÇO tahmini ile aynı koşullar altında, kısmi olabilirliklerden elde edilen parametre tahminleri, asimptotiksel olarak normal, sabit ve değişmezdir. Böylece standart hipotez testleri, CRM modeli için de doğrudan geliştirilebilmiştir [8].

2.6.1.2. Breslow Yöntemi

Breslow yöntemi [11], gözlemlerin aynı başarısızlık süresine sahip olması (ilgilenilen olayı aynı zaman noktasında yaşayan gözlemler) sorununu ortadan kaldırmak için en sık kullanılan yöntemdir. Bu yaklaşımın sıkça kullanılmasının nedeni diğer yöntemlerden daha basit olmasıdır. Aynı anda gerçekleşen olaylarda, olay meydana geliş sıralamasına karar vermek mümkün olmadığı için Breslow yöntemi, hangi olayın daha önce gerçekleştiğinden bağımsız olarak risk kümesi büyüklüğünü aynı kabul eder. Örneğin, olay meydana geliş zamanları $t = 5, 5, 8, 14$ olan dört gözlem olduğu varsayalım. Gözlemlerin ikisi 5 zamanında beklenen olayı yaşamıştır. Bu iki gözlem arasında sıralama yapılamayacağından, Breslow yöntemi bu iki gözlemin, tüm dört gözlemi de içeren risk kümesinde başarısız olduğunu varsayarak kısmi olabilirlik fonksiyonunu elde eder. Böylece, kısmi olabilirlik, birinci gözlem için;

$$l_1 = \frac{\psi_1}{\psi_1 + \psi_2 + \psi_3 + \psi_4}$$

ve ikinci gözlem için;

$$l_2 = \frac{\psi_2}{\psi_1 + \psi_2 + \psi_3 + \psi_4}$$

olarak düşünülebilir. İki durum için de risk kümesi değiştirilmemiştir ve hem birinci hem de ikinci gözlem için maksimum büyüklüktedir. Breslow yöntemi, aynı zamanda gerçekleşen olayları, başarısızlık zamanında riskteki tüm gözlemleri içeren bir risk kümesinden, sıralı olarak meydana geliyormuş gibi (sırasının bilinmemesi ve eşit görünüyorsa rağmen) varsayarak, kısmi olabilirlik fonksiyonunu tahmin eder. Breslow'un önerdiği tahmini olabilirlik fonksiyonu;

$$L_{\text{Breslow}} = \prod_{i=1}^k \frac{e^{\beta' s_i}}{\left[\sum_{j \in R(t_i)} e^{\beta' x_j} \right]^{d_i}} \quad (2.12)$$

biçiminde elde edilir. Eşitlik 2.12'de s_i , aynı zamanda gerçekleşen olaylar için x açıklayıcı değişkenlerinin toplamı ve d_i 'ler, t_i aynı zamanında gerçekleşen gözlemlerin sayısıdır. Breslow tahmin yönteminin kullanılması, verilen zaman noktası için aynı anda gerçekleşen olay sayısı az iken daha uygundur, ancak aynı zamanlı olayların sayısı arttıkça, her bir period için risk kümesi büyüklüğü çok büyük olacak ve tahmin daha duyarsız olacaktır [6, 12].

2.6.1.3. Efron Yöntemi

Alternatif bir yöntem olarak Efron [13], aynı anda gerçekleşen olayların sırasına bağlı olarak, risk kümesinin nasıl değiştiğini hesaba katan bir tahmin yöntemi önermiştir. Breslow yöntemindeki örnek tekrar ele alındığında, eğer birinci gözlem, ikinci gözlemden önce başarısızlığı yaşamış ise, bu veriler için kısmi olabilirlik fonksiyonu aşağıdaki biçimde olacaktır;

$$l_1 = \frac{\psi_1}{\psi_1 + \psi_2 + \psi_3 + \psi_4} \times \frac{\psi_2}{\psi_2 + \psi_3 + \psi_4}.$$

Ancak, eğer ikinci gözlem, birinci gözlemden önce başarısızlığı yaşamış ise, kısmi olabilirlik fonksiyonu;

$$l_2 = \frac{\psi_2}{\psi_1 + \psi_2 + \psi_3 + \psi_4} \times \frac{\psi_1}{\psi_1 + \psi_3 + \psi_4}$$

olarak elde edilecektir. İkinci risk kümesinin yapısı olayların olası sıralamasına göre değişmektedir. Efron, kısmi olabilirlik fonksiyonu için aşağıdaki yaklaşımı önermiştir;

$$L_{\text{Efron}} = \prod_{i=1}^k \left(\frac{e^{\beta' s_i}}{\prod_{r=1}^{d_i} \left[\sum_{j \in R(t_i)} e^{\beta' x_j} - (r-1) d_i^{-1} \sum_{j \in D(t_i)} e^{\beta' x_j} \right]} \right). \quad (2.13)$$

Burada r , aynı anda meydana gelen olayların sayısını ve $D(t_i)$, t_i zamanında risk kümesindeki aynı zamanlı olayların sayısını göstermektedir. Bu yöntem t_i zamanında risk kümesinin değişik yapılarını göz önüne aldığından, Breslow yöntemine göre daha uygun sonuç verdiği söylenebilir [8].

2.6.2. Temel Tehlike ve Yaşam Fonksiyonlarının Elde Edilmesi

Temel tehlike ve temel yaşam fonksiyonlarının tahminlerinin elde edilmesi için, başarısızlık zamanları $t_1 < t_2 < \dots < t_k$ biçiminde sıralanır. Collett'in çalışmasından [6] yola çıkılarak i . gözlem için tehlike fonksiyonu tahmini;

$$\hat{h}_i(t) = e^{\hat{\beta}' x_i} \hat{h}_0(t)$$

olarak elde edilmiştir. Burada $\hat{h}_0(t)$ tahmini temel tehlike fonksiyonudur. Bu fonksiyon kısmi olabilirlik fonksiyonundan doğrudan tahmin edilemez ancak, Kalbfleisch ve Prentice [14, 15] bu fonksiyonu elde etmek için bir yöntem önermişlerdir. Olay zamanlarının sıralanabileceği bilindiğinden, verilen herhangi bir başarısızlık zamanında risk kümesinin büyüklüğünü saptamak kolaydır. Bir gözlemin t_i anında risk kümesinde olmasının tahmini olabilirliği;

$$\hat{\alpha}_i = \sum_{j \in R(t_i)} e^{\beta' x_j}$$

biçimindedir. Bu ifade, Eşitlik 2.10'daki kısmi olabilirlik fonksiyonunun paydasıdır. Kalbfleisch ve Prentice [14, 15] t_i zamanında temel tehlike fonksiyonunun bir tahmininin;

$$\hat{h}_0(t_i) = 1 - \hat{\alpha}_i \quad (2.14)$$

biçiminde elde edilebileceğini göstermiştir. Eğer başarısızlık zamanları arasındaki tehlike oranının sabit olduğu varsayılıyor ise, $\hat{\alpha}_i$ bir gözlemin $t_i + t_{(i+1)}$ aralığı boyunca yaşam olasılığının bir tahminidir. $\hat{\alpha}_i$ 'den, temel yaşam fonksiyonun bir tahmini;

$$\hat{S}_0(t) = \prod_{i=1}^k \hat{\alpha}_i \quad (2.15)$$

biçiminde elde edilir ve bu bir adım fonksiyonudur. $\hat{\alpha}_0$ için, $\hat{S}_0(t) = 1$ 'dir ve bu ilk başarısızlık zamanından önceki aralıkta yaşam olasılığının 1'e eşit olduğunu ifade eder.

Yaşam fonksiyonu tahmin edildikten sonra, birikimli tehlike fonksiyonu, $\hat{H}_0(t)$ tahmini;

$$\hat{H}_0(t) = -\log \hat{S}_0(t) = -\sum_{i=1}^n \log \hat{\alpha}_i$$

biçiminde elde edilmiştir. Yaşam fonksiyonu ve birikimli tehlike fonksiyonu arasındaki ilişki nedeniyle, i. bireyin yaşam fonksiyonunu,

$$\hat{S}_i(t) = \exp^{-H_i(t)}$$

olarak elde edilebilir.

Bu eşitlik,

$$\hat{S}_i(t) = \hat{S}_0(t)^{\exp(\hat{\beta}'x)} \quad (2.16)$$

biçiminde de yazılabilir [8].

2.7. Hızlandırılmış Başarısızlık Süresi Modelleri

Klasik CRM için, Eşitlik 2.7,

$$h(t, \mathbf{x}_i) = h_0(t) \exp\left(\sum_{j=1}^p \beta_j x_{ij}\right) \quad (2.17)$$

biçiminde de ifade edilebilmektedir. Burada $h_0(t)$ temel tehlike fonksiyonudur ve tüm açıklayıcı değişken değerleri sıfıra eşit olan bir birey için tehlike fonksiyonunu ifade eder. $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})$, $j = (1, \dots, p)$ değerlerini alıyor iken \mathbf{X}_j açıklayıcı değişkenleri için i. bireyden elde edilen değerleri ve β_j , $j = 1, \dots, p$ ise modelde bilinmeyen parametreleri göstermektedir. Bu modelin tahmini, kısmi olabilirliğin maksimizasyonu ile gerçekleşir [16].

Hızlandırılmış başarısızlık süresi (HBS) modelinde tehlike fonksiyonu, $h_0(\cdot)$, t , \mathbf{X} ve β 'nin bir fonksiyonu iken;

$$h(t, \mathbf{x}_i) = h_0\left(t \exp\left(\sum_{j=1}^p \beta_j X_{ij}\right)\right) \exp\left(\sum_{j=1}^p \beta_j X_{ij}\right) \quad (2.18)$$

biçiminde tanımlanır. Bu nedenle, bu modelde, açıklayıcı değişkenlerin yaşam süresi üzerindeki etkisi doğrudandır ve başarısızlık zamanını hızlandırır ya da yavaşlatır. Bu model çoklu regresyon modelinde olduğu gibi, yaşam süresinin logaritması açıklayıcı değişkenlerle ilişkilendirilerek belirtilebilir. Bu durum,

$$\ln T = \mathbf{X}_i \boldsymbol{\gamma} + W$$

biçiminde ifade edilebilir. Burada $\mathbf{X} = [\mathbf{X}_1, \dots, \mathbf{X}_p]$, $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_p)$ ve $j = 1, \dots, p$ için $\gamma_j = -\beta_j$ 'dir. Birçok durumda, bu modelin tahmini, süre için bir dağılım varsayılarak ve log-olabilirlik fonksiyonunun maksimizasyonu ile sağlanır [16].

Parametrik yaşam modelleri için en çok kullanılan dağılımlar ve özellikleri Alt bölüm 2.7.1-2.7.6'da özetlenmiştir.

2.7.1. Üstel Dağılım

Üstel dağılım, $h(t) = \lambda$ sabit tehlikesine sahiptir. Yaşam fonksiyonu $S(t) = \exp\{-\lambda t\}$ ve yoğunluk fonksiyonu $f(t) = \lambda \exp\{-\lambda t\}$ ile ifade edilebilir. $E(T) = 1/\lambda$ ve $\text{var}(T) = 1/\lambda^2$ olduğu gösterilebilir. Üstel dağılım, uç değer dağılımı (extreme value distribution) ile ilişkilidir. Eğer,

$$Y = \log T = \alpha + W$$

ise λ parametresiyle üstel dağılıma sahiptir ve $T \approx E(\lambda)$ olarak gösterilebilir. Burada $\alpha = -\log \lambda$ 'dır ve W , $f_w(w) = e^{-w-e^{-w}}$ yoğunluğu ile standart uç değer dağılımına sahiptir [17].

2.7.2. Weibull Dağılımı

Eğer $T^p \approx E(\lambda)$ ise T , λ ve p parametreleriyle Weibull dağılımına sahiptir ve $T \approx W(\lambda, p)$ ile gösterilir. Birikimli tehlike fonksiyonu $H(t) = (\lambda t)^p$, yaşam fonksiyonu $S(t) = \exp\{-(\lambda t)^p\}$ ve tehlike fonksiyonu;

$$h(t) = \lambda^p p t^{p-1}$$

biçiminde ifade edilebilir. Weibull tehlike fonksiyonunun logaritması, $p \log \lambda + \log p$ sabiti ve $p-1$ eğimi ile zamanın logaritmasının doğrusal bir fonksiyonudur. p değeri >1 ise tehlike fonksiyonun değeri artar, p değeri $=1$ ise sabittir ve p değeri <1 ise tehlike fonksiyonunun değeri azalır.

Weibull dağılımı da uç değer dağılımıyla ilişkilidir; eğer

$$Y = \log T = \alpha + \sigma W$$

ise $T \approx W(\lambda, p)$ 'dir. Burada, $\alpha = -\log \lambda$ ve $p = 1/\sigma$ 'dir [17].

2.7.3. Gamma Dağılımı

$\Gamma(\lambda, k)$ ile tanımlanan, λ ve k parametrelili Gamma dağılımı için olasılık yoğunluk fonksiyonu;

$$f(t) = \frac{\lambda(\lambda t)^{k-1} e^{-\lambda t}}{\Gamma(k)}, \quad 0 < t < \infty$$

ve yaşam fonksiyonu;

$S(t) = 1 - I_k(\lambda t)$ biçiminde ifade edilebilir. Burada $I_k(x)$ tamamlanmamış (incomplete) gamma fonksiyonudur ve $I_k(x) = \int_0^x \lambda^{k-1} e^{-x} dx / \Gamma(k)$ olarak tanımlanır.

Yaşam fonksiyonu için kapalı biçimli bir ifade yoktur ancak hesaplanması için algoritmalar bulunmaktadır. Tehlike fonksiyonu için de açık bir formül bulunmamaktadır ancak tehlike fonksiyonu, olasılık yoğunluk fonksiyonunun yaşam fonksiyonuna oranı $h(t) = f(t)/S(t)$ biçiminde hesaplanabilir. Gamma tehlike fonksiyonu;

- Eğer $k > 1$ ise başlangıç noktasında 0'dan, λ maksimum değerine kadar monoton olarak artar. Eğer k birden büyük bir tamsayı ise Gamma dağılımı, 'Erlang dağılımı' olarak adlandırılabilir ve bir Poisson sürecinde k tane olay meydana gelene kadar geçen bekleme zamanı olarak ifade edilebilir.
- $k=1$ ise sabittir. Bu durumda, Gamma dağılımı üstel dağılıma indirgenebilir. Bu, Poisson sürecinde bir olayın meydana gelmesi için bekleme zamanı olarak tanımlanabilir.
- $k < 1$ ise başlangıç noktasında sonsuzdan, λ 'nın bir asimptotik değerine kadar monoton olarak azalır.

Tamsayı olmayan k değeri için de dağılımlar mevcuttur.

Gamma dağılımı aynı zamanda, $\log T$ 'nin bir dağılımı biçiminde de ifade edilebilir. Basit bir düzenleme ile, $Y = \log T = \alpha + W$ ise, $T \approx \Gamma(\lambda, k)$ olarak gösterilebilir. Burada W , k

parametrelili $f_w(w) = \frac{e^{kw - e^w}}{\Gamma(k)}$ yoğunluk fonksiyonlu bir genelleştirilmiş uç değer

dağılımına sahiptir. Bu yoğunluk fonksiyonu, $k=1$ iken klasik uç değer dağılımına indirgenir [17].

2.7.4. Genelleştirilmiş Gamma Dağılımı

Genelleştirilmiş Gamma dağılımının olasılık yoğunluk fonksiyonu, $p = 1/\sigma$ iken aşağıdaki biçimde yazılabilir;

$$f(t) = \frac{\lambda p (\lambda t)^{pk-1} e^{-(\lambda t)^p}}{\Gamma(k)}, \quad 0 < t < \infty.$$

Genelleştirilmiş Gamma dağılımının özel durumları, diğer dağılımları da kapsamaktadır. Genelleştirilmiş gamma dağılımı;

- $p=1$ iken Gamma dağılımına,
- $k=1$ iken Weibull dağılımına,
- $p=1$ ve $k=1$ iken üstel dağılıma

dönüşmektedir. Ayrıca, bu dağılım k 'nın sonsuza gittiği durumlarda log-normal dağılımı da kapsamaktadır [17].

2.7.5. Log-Normal Dağılım

Eğer, $Y = \log T = \alpha + \sigma W$ eşitliğinde W standart normal dağılıma sahip ise, T log-normal dağılıma sahiptir. Log-normal dağılım için;

$$f(t) = \frac{1}{t\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(\log t - \mu)^2\right], \quad t > 0, \sigma > 2,$$

$$S(t) = \frac{1}{\sigma\sqrt{2\pi}} \int_t^{\infty} \frac{1}{x} \exp\left[-\frac{1}{2\sigma^2}(\log x - \mu)^2\right] dx \text{ ve}$$

$$h(t) = \frac{(1/t\sigma\sqrt{2\pi}) \exp\left[-\log(e^{-\mu}t)^2 / 2\sigma^2\right]}{1 - \Phi[\log(e^{-\mu}t/\sigma)]}.$$

biçiminde ifade edilmektedir. Log-normal dağılımın tehlike fonksiyonu, sıfırdan bir maksimum noktaya kadar artar ve sonra monoton biçimde azalır. k değeri sonsuza yaklaşırken genelleştirilmiş uç değer dağılımı standart normal dağılıma yakınsar ve genelleştirilmiş gamma dağılımı ise log-normal dağılıma yakınsar [17].

2.7.6. Log-Lojistik Dağılım

Eğer $Y = \log T = \alpha + \sigma W$ eşitliğinde W rastlantı değişkeni, olasılık yoğunluk fonksiyonu;

$$f_w(w) = \frac{e^w}{(1+e^w)^2}$$

ve birikimli dağılım fonksiyonu;

$$F_w(w) = \frac{e^w}{1+e^w}$$

ile bir standart lojistik dağılıma sahip ise, T log-lojistik dağılıma sahiptir. Yaşam fonksiyonu;

$$S_w(w) = \frac{1}{1+e^w}$$

biçiminde ifade edilebilir. Buradaki değişkenleri T 'ye dönüştürerek yaşam fonksiyonunu

$S(t) = \frac{1}{1+(\lambda t)^p}$ biçiminde de elde edebiliriz. Bu eşitlikte $\alpha = -\log \lambda$ ve $p = 1/\sigma$ 'dir.

Birikimli tehlike fonksiyonu ise $H(t) = \frac{\lambda p (\lambda t)^{p-1}}{1+(\lambda t)^p}$ olarak elde edilebilir. Log-lojistik

dağılımda tehlike fonksiyonu;

- $p < 1$ ise, sonsuzdan monoton olarak azalır.
- $p = 1$ ise, λ 'dan monoton olarak azalır.
- $p > 1$ ise, log-normal dağılıma benzer biçimdedir [17].

3. İYİLEŞME MODELLERİ

3.1. Giriş

Yaşam çözümlenmesi alanında yapılan araştırmalar, önerilen yeni modeller ve yaklaşımlar ile gelişimini sürdürmektedir. Yaşam çözümlenmesinde en çok kullanılan model olan CRM, kitledeki her bir bireyin ilgilenilen olayı yaşayacağı, araştırılan olaya göre başarısızlığa (ölüme, metastaza veya nüksetmeye) yatkın olduğu varsayımına dayanmaktadır [18]. Ancak bu varsayım, yeterli bir gözlem süresinin ardından hastaların büyük bir bölümünün hayatta kalması durumunda doğru olmayabilir. Özellikle son yıllarda, meme kanseri [19], çocukluk lenfomaları ve lösemileri [20] dahil olmak üzere bazı kanser türlerindeki araştırmalar, bu hastalıklara sahip hastaların önemli bir bölümünün, uygulanan tedavilerden sonra iyileştiğini göstermiştir. Dolayısıyla, bu durumu dikkate alan ve bu amaçla geliştirilen iyileşme modelleri, kanser hastalarının yaşam verilerini analiz etmek için kullanılan önemli bir istatistiksel model haline gelmiştir [18].

İyileşme modelleri ile ilgili literatürde yer alan çalışmaların bir kısmı aşağıda özetlenmiştir:

Boag (1949)'ın çalışmasında iyileşme modelleri ilk defa ele alınmıştır. Standart karma iyileşme modelleri incelenmiştir. Kitle düşük risk grubu (ilgilenilen olayı yaşama olasılığı düşük olan bireyler) ve yüksek risk grubu (ilgilenilen olayı yaşama olasılığı yüksek olan bireyler) olmak üzere iki alt kümeye ayrılarak iyileşmiş kısmın oranı tahmin edilmiştir [21].

Berkson ve Gage (1952), Boag'ın ele aldığı karma iyileşme modellerini geliştirilmiş, daha kapsamlı bir çözümlenmesini yapmışlardır [22].

Farewell (1977) çalışmasında, karma iyileşme modellerinde, yaşam süresini modelleyen, üstel ve Weibull gibi dağılımları da içeren çeşitli parametrik dağılımlar kullanılmıştır [23].

Kuk ve Chen (1992) çalışmasında, lojistik regresyon modeli ile CRM'yi birleştiren bir karma model ele alınmıştır [24].

Yakovlev ve diğerleri (1993)'nin çalışmasında, standart karma iyileşme modellerine bir alternatif olarak "sınırlı birikimli tehlikeler modeli" ele alınmıştır [25].

Taylor (1995) çalışmasında, karma iyileşme modellerinde, yarı parametrik tahminleri incelemiştir [26].

Maller ve Zhou (1996)'nun çalışmasında, çeşitli parametrik regresyon modelleri başta olmak üzere, iyileşme modelleri konusuna kapsamlı bir yaklaşım önerilmiş ve parametrik olmayan modeller üzerine de çalışmalar yapılmıştır [27].

Chen ve diğerleri (1999), iyileşme modelleri için yeni bir Bayesci yaklaşım önermişlerdir [28].

Sy ve Taylor (2000)'in çalışmasında, Cox iyileşme modeli incelenmiştir [29].

Peng ve Dear (2000), karma iyileşme modellerinde, yaşam süresini modelleyen parametrik olmayan yöntemler kullanılmıştır [30].

Chatterjee ve Shih (2001), çok değişkenli iyileşme modellerini önermiştir. Kitlede iyileşmiş birimler var iken, ilişkili yaşam verilerini çözümlmek için bir yöntem geliştirilmiştir [31].

Tsodikov ve diğerleri (2003) çalışmasında, karma ve karma olmayan iyileşme modelleri arasındaki ilişkiyi incelemişlerdir [32].

Wienke ve diğerleri (2003), Chatterjee ve Shih [31] tarafından önerilen modelin özel bir biçimi geliştirilmiştir [33].

Shao ve Zhou (2004)'nin çalışmasında, üç parametrelili Burr XII dağılımını kullanarak yeni bir parametrik iyileşme modeli geliştirilmiştir [34].

Yin ve İbrahim (2005), sağdan durdurulmuş veriler için yeni bir iyileşme modeli önermişlerdir. Bu model karma iyileşme modelleri ve karma olmayan iyileşme modellerini iki özel durum olarak içermektedir [35].

Yu (2008), iyileşmiş kısmı bulunan tekrarlı ölçümlü verilerinin çözümlenmesi için, karma iyileşme modeline zayıflık teriminin dahil edilmesini incelemiştir [36].

Zhang ve Peng (2009)'in çalışmasında, hızlandırılmış tehlikeler karma iyileşme modeli için yarı parametrik bir tahmin yöntemi geliştirilmiştir [37].

Lu (2010), iyileşmiş kısma sahip hızlandırılmış başarısızlık süresi modelleri için etkili tahminlerin elde edilmesini incelemiştir [38].

Tong, Mues ve Thomas (2012)'in çalışmasında, karma iyileşme modellerinin kredi derecelendirmede kullanımı incelenmiştir [39].

Cai, Zou, Peng ve Zhang (2012), yarı parametrik karma iyileşme modellerinin tahmini için bir R paketi (smcure) önermişlerdir [40].

Maetani ve Gamel (2013)'in çalışmasında, parametrik iyileşme modelleri ile CRM karşılaştırılmıştır [41].

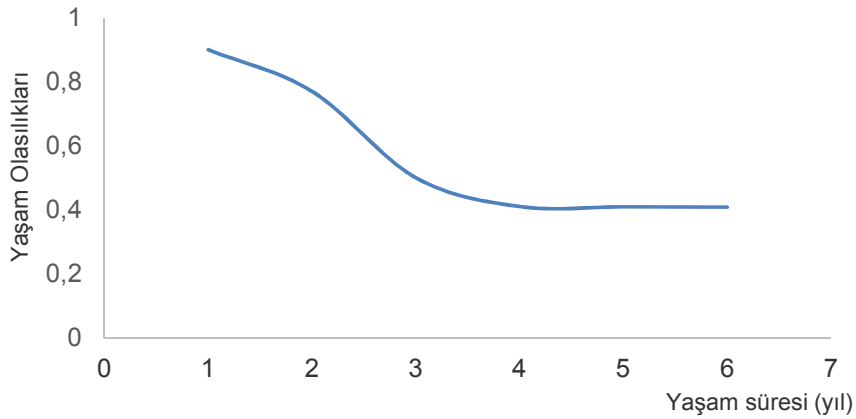
İyileşmenin Tıbbi ve İstatistiksel Tanımı

Tıbbi anlamda 'iyileşme', hastada kanserin bütün belirtilerinin, göstergelerinin yok olmasıdır. Bu da genellikle tüm kötü huylu hücreler yok edildiği zaman gözlemlenir. Bu tanım, iyileşmenin tıbbi ve hasta (birey) düzeyindeki tanımıdır. Ancak bir hastanın tıbbi olarak iyileştiğini varsaymak zor olabilir [49].

İstatistiksel anlamda 'iyileşme' ise, kitlede hastalıkla ilişkili ölüm oranının, kitlede beklenen ölüm oranına eşit olması ile gerçekleşir. Bu durumda ek ölüm oranı (excess mortality rate) sıfıra yaklaşır. Bu, iyileşmenin kitle düzeyindeki tanımıdır [49]. Gözlenen ölüm oranı aşağıdaki gibi ifade edilebilir:

$$\text{Gözlenen Ölüm Oranı} = \text{Beklenen Ölüm Oranı} + \text{Ek Ölüm Oranı}$$

Ek ölüm oranı sıfıra yaklaştığında yaşam eğrisi, Şekil 3.1'deki gibi düzleşerek düz bir kuyruğa sahip olur.



Şekil 3.1. Yaşam eğrisi örneği

Şekil 3.1'deki gibi yaşam eğrilerine sahip veri kümelerinde, iyileşme modellerinin kullanılması literatürde tavsiye edilmektedir.

3.2. İyileşme Modeli ve Cox Regresyon Modeli Arasındaki Farklar

Cox [9] belli bir aralığın başında hayatta olan bireylerin, bu aralık sürecinde ölmesi olasılığını dikkate alarak bir model geliştirmiştir. Herhangi bir zaman noktasında, bir gruptaki tehlike fonksiyonunun, aynı zaman noktası için başka bir gruptaki tehlike fonksiyonu ile orantılı olacağı varsayımında bulunmuştur. Ancak, iki grup arasındaki tehlike oranı sabit kalırken, mutlak tehlike fonksiyonunun zaman içerisinde değişebileceğini dikkate almak önemlidir. Bu, gruplar arasındaki yaşam olasılıklarının karşılaştırılmasına olanak sağlamaktadır. Örneğin, iki grup arasındaki ortalama tehlike oranı 0.5 ise, bir gruptaki ölüm oranının, diğer gruptaki ölüm oranının yarısı kadar olduğu varsayılır [42].

Genellikle, iyileşmiş hastaların oranının, klasik yaşam çözümlemesi yöntemleri ile hesaplanan “1-TO” ile tahmin edilebileceği varsayılır. Örneğin, transtuzumablı ve transtuzumabsız adjuvan kemoterapiyi karşılaştıran bir çalışma ele alınsın ve HER2-pozitif göğüs kanseri hastaları için tehlike oranı 0.67 olarak elde edilmiş olsun. Bu sonuç, ‘transtuzumab tedavisi, ölüm oranını %33 (1-0.67) azaltmıştır’ ya da ‘transtuzumab tedavisi, tedavi olmadan meydana gelebilecek ölümlerin %33’ünü engellemiştir’ biçiminde yorumlanabilir. Böyle yorumlar, sağlık bilimleri yayınlarında sıkça görülebilir. Bu bilgilere göre, tedavi için TO değeri 1’den küçük ise, hastanın iyileşme şansının artacağına inanılır [41]. Ancak, Peto ve diğerleri [5] ve Clark ve diğerlerine [2] göre, “1-TO” tedavi sayesinde ölümü ya da nüksetmesi önlenen hastaların oranı değil, önlenen ya da ertelenen ölümlerin oranıdır. Dolayısıyla, TO ölümü önleyen (iyileştirici tedavi) ve ölümü erteleyen (yaşam süresini uzatan veya ölümü erteleyen tedavi) arasında ayırım yapamaz [42].

İyileşme oranı ve tehlike oranı arasındaki ilişki, Cox’un çalışmasında görülmektedir [9]. 1959 ve 1960 yılları arasında, çocuklardan oluşan hasta grubu ile 6-Merkaptopürin (6-MP) ve placebo grupları olacak biçimde kontrollü bir deney yapılmıştır. CRM kullanılarak veriler tekrar analiz edildiğinde, 6-MP grubunun placebo grubuna göre tehlike oranı 0.22 (%95 güven düzeyinde güven aralığı ise 0.10-0.49’dır) bulunmuştur. Eğer tehlike oranı ile iyileşme oranı arasındaki ilişki geçerli olsaydı, bulunan sonuç şunu ifade edecekti: placebo grubunda meydana gelebilecek nüksetmelerin %78’i

($1 - TO = 0.78$), 6-MP tedavisi sayesinde önlenmiştir. Ancak, hastalar daha uzun süre takip edildiğinde, 6-MP ile tedavisi yapılan çocukların neredeyse hepsinin hastalığın nüksetmesinden dolayı öldüğü görülmüştür [42].

İyileşme modellerinin klasik CRM'ye tercih edildiği durumlar aşağıdaki biçimde özetlenebilir;

1. Yaşam eğrisi çizildiğinde grafik düz bir kuyruğa sahip ise orantılı tehlikeler varsayımı sağlanmıyor olabilir. Bu durumda iyileşme modellerin kullanılması daha iyi sonuç verebilmektedir.
2. Çalışma grubu içerisinde heterojenlik var ise iyileşme modellerin kullanılması önerilebilmektedir [43].

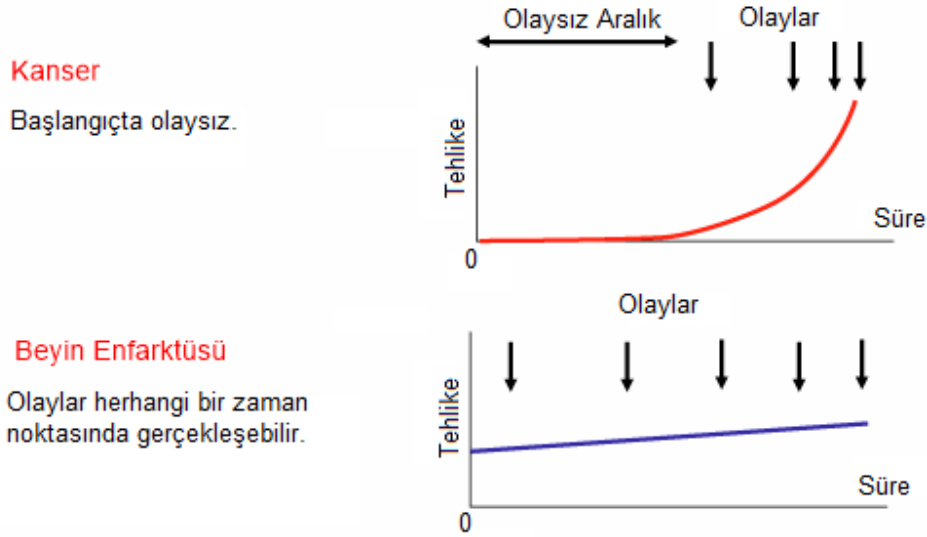
3.3. İyileşme Modelleri

Son yıllarda, kanser hastalığının teşhisi ve tedavisindeki gelişmeler sayesinde kanser hastalarının önemli bir bölümü iyileşebilmektedir. Son yıllardaki bir diğer gelişme ise hasta odaklı tedavidir, bu tedavide hastalar, uygulanacak en uygun tedavinin seçiminde önemli bir rol oynamaktadırlar. Bu hedefi gerçekleştirmek için, gerekli bilgilerin anlaşılır hale getirilmesi gerekmektedir [42].

Hastalar, kanser teşhisi konulduktan sonra genellikle 'iyileşme şansım ne kadar?' ve 'iyileşmezsem ne kadar süre daha yaşayabilirim?' sorularını sormaktadırlar. İyileşme olsa da olmasa da tedavi, hem hastaların hayat kalitesi hem de nüksetmeye bağlı fiziksel ve ruhsal değişimleri yaşamaması açısından çok önemlidir ve hastanın hayatını yıllarca uzatabilir. Örneğin, çocuklara uygulanan tedaviler ile yaşam süresi ortalama 50 yıl kadar uzatabilmektedir. Bu nedenle yaşam çözümlenmesinde, hastaları iyileştiren tedaviler ile ölümü erteleyen tedaviler arasında ayırım yapabilmek önemlidir. Bu temel soru, CRM'ye dayanan klasik yaşam modellerinde ele alınmamıştır. Gamel [44, 45], Frankel ve Longmate [46], Spoto [20] ve Maetani, Segawa ve diğerleri [47] çalışmalarında bu soruna değinmişlerdir.

İyileşme modellerinin en sık kullanıldığı alanlardan biri olan kanser, adım adım ilerleyen kronik bir hastalıktır. Kanser bu yönüyle, beyin enfarktüsü ya da anevrizması gibi ölümün her an gerçekleşebileceği diğer akut hastalıklardan, Şekil 3.2'de de ifade edildiği gibi farklılık göstermektedir. Bu hastalıkların tersine, kanser hastalığında genellikle

çalışmanın başında olay gözlemlenmez. Ölümün gerçekleşmesi yakın olan ölümcül durumlar genellikle çalışmalara dahil edilmediğinden, daha ileri kanserlere sahip olan hastalar için de bu durum geçerlidir. Kanser ve diğer akut hastalıklar arasındaki bu fark, uygun yaşam modelini formüle etmek açısından önemlidir [42].



Şekil 3.2. Olay zamanlarının kanser ve diğer akut hastalıklar için dağılımı [42].

Silverstri ve diğerleri [48] tarafından yayınlanan bir çalışmada, öncesinde kemoterapi görmüş 3. ve 4. evredeki küçük olmayan hücreli akciğer kanseri olan 73 hasta ile görüşülmüştür. Metastazlı, küçük olmayan hücreli akciğer kanseri hastalarına, kemoterapinin hayat uzatma etkisi ve buna karşılık yan etkisi arasında nasıl bir seçim yapacağı sorulmuştur. Sorulardan biri, hastaların eğer hayatlarını üç aya kadar uzatacak ise cisplatin bazlı kemoterapiyi kabul edip etmeyeceğidir. Hastaların sadece 18'i (%22'si) kemoterapiyi tercih etmiştir. Hastalara, hayatı uzatmayan ancak hastalığın semptomlarını yok eden bir tedavi teklif edildiğinde, 55 hasta (hastaların %68'i) kemoterapi almayı kabul etmiştir. Dolayısıyla, kanser tedavilerinde yaşam süresini uzatmak hasta için her zaman fayda anlamına gelmemektedir. Bu ikilemi çözebilmek için, tedavilerin uzun dönemli etkilerinin ölçülmesini sağlayacak daha iyi yöntemlere ihtiyaç doğmuştur. 1949 yılında Boag, hastalara en uygun tedavinin seçilmesi için gerekli bilgileri sağlayan bir parametrik yaşam modeli önermiştir.

Ancak, yüksek iyileşme oranının, daima uzun süre yaşamı garanti etmediğine dikkat edilmelidir. Hasta, ilgilenilen kanserden iyileşse bile başka nedenlerden ölebilir. Bu durumun, yaşlı hastalarda ve beklenen yaşam süresinin daha kısa olduğu ülkelerde yaşayan hastalarda daha sık gözlenmesi beklenir. Bu nedenle, kanser tedavilerinin uzun dönemli etkilerini araştırırken, sadece çalışılan kanserden değil, tüm nedenlerden meydana gelen ölümleri dikkate almak önemlidir [42].

3.3.1. Karma İyileşme Modelleri

Karma iyileşme modelleri, ilk olarak Boag [21] tarafından önerilmiş ve Berkson ve Gage'in çalışmaları [22] ile geliştirilmiştir. Karma iyileşme modelleri, kitlede hem iyileşmiş kısmı, hem de iyileşmemiş kısmı dikkate alan modellerdir. Karma iyileşme modellerinde kitle iki alt gruba ayrılabilir; kitledeki bireyler, z açıklayıcı değişkenler vektörü ve $0 \leq \pi \leq 1$ olmak üzere ya $1 - \pi(z)$ oranıyla iyileşmiştir ya da $\pi(z)$ oranıyla iyileşmemiştir.

$S_u(t | x) = S(t | Y = 1, x)$ t anında iyileşmemiş hastaların yaşam fonksiyonu, x belirli bir açıklayıcı değişken vektörü iken, karma iyileşme modelleri en genel biçimiyle aşağıdaki biçimde ifade edilebilir;

$$S(t | x, z) = \pi(z)S_u(t | x) + (1 - \pi(z))$$

ya da

$$S(t | x, z) = \pi(z)S(t | Y = 1, x) + (1 - \pi(z)). \quad (3.1)$$

Karma iyileşme modellerinde iyileşmiş kısım ve iyileşmemiş kısım ayrı ayrı modellenmektedir. Karma modellerin bu özelliği açıklayıcı değişkenlerin, iyileşmiş ve iyileşmemiş hastalar üzerindeki etkilerinin farklı olmasına olanak sağlamaktadır. Örneğin, yapılan bir tedavi, iyileşmiş hastaların oranını arttırabilir ancak bunun yanı sıra iyileşmemiş hastaların yaşamında herhangi bir etkisi olmayabilir. Karma iyileşme modelleri, bunun gibi farklı etkilerin görülebilmesine imkan sağlamaktadır [43].

Karma iyileşme modellerinde, z açıklayıcı değişkenler vektörünün iyileşmiş kısımdaki etkisini modellemek için bazı bağlantı (link) fonksiyonları kullanılmaktadır. İyileşmiş

kısım oranı $1 - \pi(\mathbf{z})$, açıklayıcı değişkenler ile farklılık gösterebilir ve farklı durumlarda farklı bağlantı fonksiyonlarının kullanılması bir avantaj olabilir.

Bu bağlantı fonksiyonları aşağıdaki gibi özetlenebilir:

Lojit bağlantı fonksiyonu (Logit link function), iyileşmiş kısmı modellemek için en çok tercih edilen bağlantı fonksiyonudur ve aşağıdaki biçimde ifade edilir;

$$\pi(\mathbf{z}) = \frac{e^{b\mathbf{z}}}{1 + e^{b\mathbf{z}}}.$$

Burada \mathbf{b} , açıklayıcı değişken vektörü ile ilişkili bilinmeyen parametrelerin vektörüdür.

Tamamlayıcı log-log bağlantı fonksiyonu (Complementary log-log link function), lojit bağlantı fonksiyonundan sonra en çok tercih edilen ikinci bağlantı fonksiyonudur.

$\log(-\log(1 - \pi(\mathbf{z}))) = b\mathbf{z}$ biçiminde ifade edilmektedir.

Probit bağlantı fonksiyonu (Probit link function), Φ , standart normal birikimli dağılım fonksiyonu iken,

$$\Phi^{-1}(\pi(\mathbf{z})) = b\mathbf{z}$$

biçiminde ifade edilir [49].

Karma iyileşme modellerinde iyileşmemiş kısım için yaşam fonksiyonu $S_u(t | \mathbf{x})$ ise kullanılan karma iyileşme modeline göre farklılık göstermektedir. $S_{u0}(t)$ iyileşmemiş

hastalar için temel yaşam fonksiyonu iken, $S_u(t | \mathbf{x}) = S_{u0}(t)^{e^{b\mathbf{x}}}$ ise bu karma iyileşme modeli “Cox karma iyileşme modeli” ya da “orantılı tehlikeler karma iyileşme modeli”,

$S_u(t | \mathbf{x}) = S_{u0}(te^{b\mathbf{x}})$ ise “yarı parametrik hızlandırılmış başarısızlık süresi karma iyileşme modeli”, $S_u(t | \mathbf{x}) = S_{u0}(te^{b\mathbf{x}})^{\exp(-b\mathbf{x})}$ ise “hızlandırılmış tehlikeler karma iyileşme modeli” olarak adlandırılmaktadır.

n tane hastanın katıldığı bir çalışma ele alınsın. $i=1,2,\dots,n$ iken, i . hasta için gözlemlenen değerler $\mathbf{O} = (t_i, \delta_i, \mathbf{z}_i, \mathbf{x}_i)$ olsun. Burada t_i yaşam süresi, δ_i durdurma gösterge değişkeni olarak ifade edilmiştir. $\Theta = (\mathbf{b}, \boldsymbol{\beta}, S_0(t))$ tahmin edilmek istenilen bilinmeyen parametreler ve Y , $Y=1$ iken bireyin ilgilenilen olayı (başarısızlığı) yaşayacağı, $Y=0$ iken bireyin ilgilenilen olayı yaşamayacağını belirten gösterge olsun.

$y = (y_1, y_2, \dots, y_n)$ vektörü ve gözlemlenen veriler \mathbf{O} biliniyorken olabilirlik fonksiyonu aşağıdaki biçimde ifade edilebilir;

$$L(\mathbf{b}, \boldsymbol{\beta}) = \prod_{i=1}^n [1 - \pi_i(\mathbf{z}_i)]^{1-y_i} \pi_i(\mathbf{z}_i)^{y_i} h(t_i | Y = 1, \mathbf{x}_i)^{\delta_i y_i} S(t_i | Y = 1, \mathbf{x}_i)^{y_i} . \quad (3.2)$$

Bu eşitlikte $h(\cdot)$, $S(\cdot)$ 'ye karşılık gelen tehlike fonksiyonudur. Bu tam olabilirlik fonksiyonu iki parça olarak da ifade edilebilir;

$$l_{c1}(\mathbf{b}; \mathbf{0}, y) = \sum_{i=1}^n y_i \log[\pi(\mathbf{z}_i)] + (1 - y_i) \log[1 - \pi(\mathbf{z}_i)] \quad (3.3)$$

ve

$$l_{c2}(\boldsymbol{\beta}; \mathbf{0}, y) = \sum_{i=1}^n y_i \delta_i \log[h(t_i | Y = 1, \mathbf{x}_i)] + y_i \log[S(t_i | Y = 1, \mathbf{x}_i)]. \quad (3.4)$$

Tam olabilirlik ve log-olabilirlik fonksiyonları bu şekilde ifade edildikten sonra, farklı tahmin yöntemleri ele alınabilir. EM algoritması, y 'nin modele dahil olması nedeniyle kullanışlıdır. Burada gözlemlenemeyen değişken, EM algoritmasındaki beklenen değerle yer değiştirecektir.

3.3.1.1. Cox Karma İyileşme Modeli

Karma iyileşme modelleri için çeşitli parametrik ve parametrik olmayan yöntemler önerilmiştir. Farewell (1982) çalışmasında iyileşmiş kısım için lojistik regresyon ve iyileşmemiş hastaların yaşam fonksiyonu için Weibull regresyon modeli kullanmıştır. Kuk ve Chen, ilgilenilen olaya duyarlı olan grupta CRM kullanarak, Farewell'in modelinin yarı parametrik bir genellemesini önermiştir. Bu modele "Cox karma iyileşme modeli" ya da "orantılı tehlikeler karma iyileşme modeli" denilmektedir [29].

Y , bir bireyin, ilgilenilen olayı kesinlikle yaşayacağını ($Y=1$) ya da hiç yaşamayacağını ($Y=0$) gösteren gösterge değişken olsun. T , ilgilenilen olayın gerçekleşme zamanını göstermektedir ve sadece $Y=1$ iken tanımlıdır ve olasılık yoğunluk fonksiyonu $f(t | Y = 1)$, yaşam fonksiyonu ise $S(t | Y = 1)$ ile gösterilmiştir. Durdurulmuş gözlemler için Y gözlemlenemez. T 'nin marjinal yaşam fonksiyonu, $t < \infty$ iken, $S(t) = (1 - \pi) + \pi S(t | Y = 1)$ 'dir. t sonsuza yakınsarken, $S(t)$, $1 - \pi$ 'ye yakınsayacaktır.

Bağımsız, bilgilendirici olmayan ve rasgele bir durdurulma modeli varsayılmıştır ve durdurulma Y'den istatistiksel olarak bağımsızdır [29].

Farewell [50], iyileşmiş kısmın hesaplanması için lojistik regresyon modelini kullanmıştır ve iyileşme oranı $1 - \pi(z) = 1 - P(Y = 1; z) = 1 - [\exp(z'b)/(1 + \exp(z'b))]$ biçiminde ifade edilebilir. Burada açıklayıcı değişken vektörü \mathbf{z} , $S(t | Y = 1)$ için parametrik bir yaşam modeli ve kesişim terimi içermektedir. Kuk ve Chen [24] \mathbf{x} kesişim değişkeni dışındaki açıklayıcı değişkenler vektörü ve $h_0(t | Y = 1)$ temel tehlike fonksiyonu olmak üzere, $h(t | Y = 1; \mathbf{x}) = h_0(t | Y = 1) \exp(\mathbf{x}'\beta)$ tehlike fonksiyonuna sahip CRM'yi kullanarak bu modeli genelleştirmiştir.

Bu model \mathbf{b} ve β aracılığı ile açıklayıcı değişkenlerin iyileşmiş ve iyileşmemiş kısım üzerindeki etkilerini ayırabilme özelliğine sahiptir. Bu nedenle de ilgilenilen olaya duyarlı olmayan bir grupta önsel bilgi olduğunda, esnek bir model sağlar. Kuk ve Chen [24] modeli için, koşullu birikimli tehlike fonksiyonu $H(t | Y = 1; \mathbf{x}) = H_0(t | Y = 1) \exp(\mathbf{x}'\beta)$ olur.

Burada $H_0(t | Y = 1; \mathbf{x}) = \int_0^t H_0(u | Y = 1) du$ 'dur. Koşullu temel yaşam fonksiyonu

$S_0(t | Y = 1)$ iken, koşullu yaşam fonksiyonu $S(t | Y = 1; \mathbf{x}) = S_0(t | Y = 1) \exp(\mathbf{x}'\beta)$ olarak ifade edilir.

Orantılı tehlikeler fonksiyonlarının bir karışımının, artık orantılı olmadığı ve hatta iki elemanlı bir açıklayıcı değişken için, Cox karma iyileşme modelinin, kesişen marjinal yaşam eğrilerine sahip olabildiği gözlemlenebilir. Ancak, klasik CRM, tüm \mathbf{z} 'ler için, $\pi(z) = 1$ olan Cox karma iyileşme modelinin özel bir biçimidir.

Cox iyileşme modeli, bir birey için tehlike fonksiyonunun $h(t | Y; \mathbf{x}) = Yh(t | Y = 1; \mathbf{x})$ olarak ifade edilebildiği, çarpımsal zayıflık modelinin özel bir biçimidir [29].

Cox karma iyileşme modellerinin EÇÖ tahminleri, log-olabilirlik fonksiyonu maksimize edilerek elde edilebilir. Eğer fonksiyonun kapalı biçimi yok ise Newton-Raphson iterasyon yöntemi kullanılarak sonuçlar elde edilebilir [51].

Cox Karma İyileşme Modeli için En Çok Olabilirlik Tahmini

i bireyi için gözlenen veriler $\mathbf{O} = (t_i, \delta_i, \mathbf{x}_i, \mathbf{z}_i)$ olarak ifade edilirse, burada t_i olayın meydana geliş zamanı ya da durdurulma zamanı, \mathbf{x}_i ise açıklayıcı değişkenler vektörüdür. Eğer t_i durdurulmamışsa $\delta_i = 1$, durdurulmuş ise $\delta_i = 0$ değerini alır. $t_1 < t_2 < \dots < t_k$ k tane ayrık olayın meydana geliş zamanları olsun. Burada y_i , Y_i rastgele değişkeninin aldığı değer iken, eğer $\delta_i = 1$ ise $y_i = 1$ olacağı ve $\delta_i = 0$ ise y_i 'nin gözlemlenemeyeceği sonucuna ulaşılabilir. i . bireyin olabilirlik fonksiyonuna katkısı, $\pi_i = P(Y_i = 1; \mathbf{x}_i)$ iken, $\delta_i = 1$ için $\pi_i f(t_i | Y = 1; \mathbf{z}_i)$, $\delta_i = 0$ için ise $(1 - \pi_i(z)) + \pi_i(z)S(t_i | 1; \mathbf{x}_i)$ 'dir. Böylece, Cox karma iyileşme modeli için gözlenen tam olabilirlik fonksiyonu;

$$L(\mathbf{b}, \beta, h_0) = \prod_{i=1}^n \left\{ \pi_i h_0(t_i | Y = 1) e^{x_i' \beta e^{-H_0(t_i | Y=1) \exp(x_i' \beta)}} \right\}^{\delta_i} \times \left\{ (1 - \pi_i) + \pi_i e^{-H_0(t_i | Y=1) \exp(x_i' \beta)} \right\}^{1 - \delta_i} \quad (3.5)$$

olarak ifade edilir [29]. Burada $L(\mathbf{b}, \beta, H_0)$ 'yi maksimize eden $\hat{\mathbf{b}}$ ve $\hat{\beta}$ tahminleri elde edilmek istenmektedir. Klasik CRM'de, kısmi olabilirliği kullanmak için klasik çözümlenme, $h_0(t)$ 'ye bağlı değildir. Breslow [52] yarı parametrik bir tam olabilirlik yapısı ve $H_0(t)$ 'nin tahmin edicisinin Aalen-Nelson tahmin edicisi olduğu profil olabilirlik yöntemini kullanmıştır. Breslow, β 'nin bu kısmen maksimize edilmiş olabilirlik fonksiyonunun, kısmi olabilirlik ile orantılı olduğunu göstermiştir. Ancak bu yaklaşım, Cox iyileşme modeli için uygun değildir. $S_0(t)$ 'yi eleyerek sadece az miktarda bilginin kaybedildiği klasik CRM'nin aksine, tahmindeki $S_0(t | Y = 1)$, \mathbf{b} hakkındaki bilgiler kaybedilmeden modelden çıkartılamaz [29].

EM Algoritması

Cox karma iyileşme modelinde, EM algoritması kullanılmaktadır. EM algoritmasındaki E adımı, gözlemlenen veri \mathbf{O} ve parametrelerin θ andaki tahminleri $\Theta^{(m)} = (\mathbf{b}^{(m)}, \beta^{(m)}, S_0^{(m)}(t))$ biliniyorken, y_i 'lere göre tam log-olabilirliklerin koşullu beklentisini hesaplar. Bu adım için Eşitlik 3.1 ve Eşitlik 3.2 y_i 'nin doğrusal fonksiyonları

olduğu için bu adımı tamamlamak için y_i 'nin koşullu beklentisini elde etmek yeterlidir.

$E(y_i | O, \Theta^{(m)})$ aşağıdaki biçimde ifade edilebilir;

$$w_i^{(m)} = E(y_i | O, \Theta^{(m)}) = \bar{\delta}_i + (1 - \bar{\delta}_i) \frac{\pi(\mathbf{z}_i) S(t_i | Y = 1, \mathbf{x}_i)}{1 - \pi(\mathbf{z}_i) + \pi(\mathbf{z}_i) S(t_i | Y = 1, \mathbf{x}_i)} \Bigg|_{(O, \Theta^{(m)})} \quad (3.6)$$

Bu eşitlikten, eğer $\bar{\delta}_i = 1$ ise $w_i^{(m)} = 1$ ve $\bar{\delta}_i = 0$ ise $w_i^{(m)}$ 'nin iyileşmemiş hastaların oranı olacağı görülmektedir. $E(y_i | O, \Theta^{(m)})$ eşitliğinin ikinci kısmı, i. hastanın iyileşmemiş olarak kalması koşullu olasılığı olarak yorumlanabilir. $\bar{\delta}_i \log w_i^{(m)} = 0$ ve $\bar{\delta}_i w_i^{(m)} = \bar{\delta}_i$ olduğu için Eşitlik 3.3 ve Eşitlik 3.4'ün beklenen değerleri sırasıyla aşağıdaki gibi yazılabilir;

$$E(l_{c1}) = \sum_{i=1}^n w_i^{(m)} \log[\pi(\mathbf{z}_i)] + (1 - w_i^{(m)}) \log[1 - \pi(\mathbf{z}_i)], \quad (3.7)$$

$$E(l_{c2}) = \sum_{i=1}^n \bar{\delta}_i \log[w_i^{(m)} h(t_i | Y = 1, \mathbf{x}_i)] + w_i^{(m)} \log[S(t_i | Y = 1, \mathbf{x}_i)]. \quad (3.8)$$

EM algoritmasında M adımı, bilinmeyen parametrelere göre Eşitlik 3.7 ve Eşitlik 3.8'in maksimize edilmesi ile elde edilir. İyileşmiş kısım için parametrelerin tahmininde, R programında 'glm' fonksiyonunun içindeki 'link' seçeneği kullanılır. Peng ve Dear [30] ve Sy ve Taylor [29] çalışmalarında β 'yi temel tehlike fonksiyonu belirtmeden tahmin etmek için bir kısmi olabilirlik yöntemi ele alınmıştır. Bunun için tahmin eşitliği;

$$\log \prod_{i=1}^n \left[h_0(t_i) \exp(\beta \mathbf{x}_i + \log(w_i^{(m)})) \right]^{\bar{\delta}_i} S_0(t_i)^{\exp(\beta \mathbf{x}_i + \log(w_i^{(m)}))} \quad (3.9)$$

biçimindedir.

Algoritmanın M adımı ise l_c 'nin maksimizasyonunu içerir. $H_0(t | Y = 1)$ ya da $S_0(t | Y = 1)$ için, profil olabilirlik yöntemlerini kullanarak M adımına ek bir maksimizasyon adımı uygulanır. CRM'den $H_0(t | Y = 1)$ için Breslow-tipi tahmin edici ve $S_0(t | Y = 1)$ için çarpım-limit tahmin edicisi olmak üzere iki yöntem genişletilebilir [29].

E adımına geri dönmek için, tahmin edilen yaşam fonksiyonları güncellenmelidir. $t_1 < t_2 < \dots < t_k$ ayrık, durdurulmamış başarısızlık zamanları olsun. d_{t_j} ilgilenilen olayların sayısı, $R(t_j)$, t_j zamanındaki risk kümesi olsun. $S_0(t | Y = 1)$ için Breslow tipi tahmin edici Eşitlik 3.10'da verilmektedir;

$$\hat{S}_0(t | Y = 1) = \exp \left(- \sum_{j: t_{(j)} \leq t} \frac{d_{t_{(j)}}}{\sum_{i \in R(t_{(j)})} w_i^{(m)} e^{\hat{\beta}_m x_i}} \right). \quad (3.10)$$

Kitlede iyileşmiş kısmın varlığından dolayı, t sonsuza giderken $\hat{S}_0(t | Y = 1)$ sifıra yakınsamayabilir. Bu nedenle, tanımlanma sorunlarını ortadan kaldırmak için, t_k son gözlemlenen başarısızlık zamanını ifade ediyor iken, $t > t_k$ için $\hat{S}_0(t | Y = 1) = 0$ olarak ayarlanabilir. Bunun sonucunda tahmin edilen yaşam fonksiyonu $\hat{S}_0(t | Y = 1) = \hat{S}_0(t_k | Y = 1)^{\exp(\hat{\beta}_m x)}$ olur.

3.3.1.2. Yarı Parametrik Hızlandırılmış Başarısızlık Süresi Karma İyileşme Modeli

HBS modelleri, yaşam çözümlenmesinde kullanılan önemli modellerden biridir. HBS modelinin kullanılmasının en büyük avantajı, açıklayıcı değişkenlerin başarısızlık süreleri üzerinde olan etkilerinin doğrudan modellenebilmesi ve de bu etkilerin yorumlanmasının CRM'den elde edilen sonuçlara kıyasla daha basit olmasıdır. Genellikle, modeldeki regresyon parametrelerinin tahminleri, eğer modeldeki hata teriminin dağılımı belirtilmişse, olabilirlik fonksiyonunun maksimize edilmesiyle elde edilir. Tsiatis [53] hata terimi için parametrik bir varsayım kullanmamış, rank tahmin yöntemi önermiştir. Ritov [54] genel en küçük kareler (EKK) tahmin yöntemini ele alarak rank tahmin yöntemi ile genel EKK tahmin yöntemi arasındaki asimptotik eş değeri göstermiştir.

HBS modelinin önemli bir varsayımı, yeterli bir takip süresinin sonunda, çalışmadaki her birimin mutlaka ilgilenilen olayı yaşayacak olmasıdır. Bu varsayım, kitlede olayı yaşamayan bireyler olduğunda sağlanmayacaktır.

HBS karma iyileşme modelleri, Li ve Taylor [67] ve Zhang ve Peng [68] tarafından karma iyileşme modelindeki iyileşmemiş bireylerin başarısızlık zamanlarının modellenmesi için önerilmiştir.

HBS karma iyileşme modelinde, iyileşmemiş hastalar için yaşam fonksiyonu,

$$S(t | Y = 1, x) = S_0(te^{x\beta}) \quad (3.11)$$

biçimindedir. Temel dağılım için genellikle az sayıda bilinmeyen parametrelili bir parametrik dağılım varsayılır ve modeldeki parametreler EÇÖ yaklaşımı kullanılarak tahmin edilir [50, 55, 56].

Bu modelin açıkça belirtilmemiş varsayımı, iyileşmemiş hastaların tehlike oranındaki açıklayıcı değişken etkilerinin ani olmasıdır. Eğer yeni bir tedavi uygulanıyorsa 1, klasik tedavi uygulanıyorsa 0 değerini alan, tek bir açıklayıcı değişkene sahip durum ele alınsın. Karma iyileşme modelinde, açıklayıcı değişken hem x hem de z 'de değerlendirilsin ve klasik tedavi grubundaki hastaların tehlikesi $h_0(0) > 0$ koşulunu sağlasın. İyileşmemiş hastalar için Cox karma iyileşme modelinde, yeni tedavi grubundaki hastaların, klasik tedavi grubundaki hastalara göre tehlike oranının t sıfır iken $e^{\beta x}$ olduğu ve her $t > 0$ için aynı kaldığı görülmektedir. HBS karma iyileşme modeli ise her ne kadar tehlike oranı zaman aralığı süresince sabit olmasa da, Cox karma iyileşme modelindeki gibi $t = 0$ anında $e^{\beta x}$ ile başlamaktadır.

HBS karma iyileşme modellerindeki parametreleri tahmin etmek için son dönemlerde birçok araştırmacı yarı parametrik tahmin yöntemleri ile ilgilenmiştir. Li ve Taylor [57] HBS karma iyileşme modellerinde bilinmeyen parametrelerin tahmini için M-tahmin yöntemini [54] kullanmıştır. Zhang ve Peng [58] HBS karma iyileşme modeli için yarı parametrik bir tahmin yöntemi geliştirmek amacıyla bir rank tahmin yöntemini [59] modele uyarlamıştır.

HBS karma iyileşme modelinde parametre tahmini için, iyileşmiş kısım Cox karma iyileşme modelinde olduğu gibi tahmin edilir, ancak iyileşmemiş kısım farklı biçimde elde edilmek zorundadır. M adımı β 'nin tahmini için ranka dayalı bir tahmin yöntemi, Zhang ve Peng tarafından önerilmiştir. Bu yöntem, Eşitlik 3.9'u, $w_i^{(m)}$ ek sabit terimini

içeren yarı parametrik klasik bir HBS modeli için bir log olabilirlik fonksiyonu biçiminde tekrar yazmayı içerir. Bu, Eşitlik 3.12'de verilmiştir;

$$\log \prod_{i=1}^n \left[w_i^{(m)} h(\log(t_i) - \beta x_i) \right]^{\delta_i} \left[S(\log(t_i) - \beta x_i)^{w_i^{(m)}} \right] \quad (3.12)$$

Bu, HBS modeli için mevcut yarı parametrik tahmin yöntemlerini kullanarak M adımındaki β parametresini tahmin etmemize olanak sağlar. Zhang ve Peng çalışmalarında $G(\beta)$ konveks fonksiyonunun maksimize edilmesini önermişlerdir;

$$G(\beta) = n^{-1} \sum_{i=1}^n \sum_{j=1}^n \delta_i w_i^{(m)} |\varepsilon_i - \varepsilon_j| I(\varepsilon_i - \varepsilon_j) \quad (3.13)$$

Eşitlik 3.9'un maksimize edilmesi, Eşitlik 3.13'ün maksimize edilmesiyle gerçekleşir.

Yukarıda açıklanan rank yönteminden elde edilen β ile, yaşam fonksiyonu için $\tau_i = \log t_i - \beta x_i$, $i = 1, \dots, n$ artıklarına dayalı bir tahmin edici elde edilebilir. $\tau_1 < \tau_2 < \dots < \tau_k$ ayırık durdurulmamış başarısızlık artıkları olsun. d_{τ_j} başarısızlıkların sayısı ve $R(\tau_j)$, τ_j 'deki risk kümesi olsun. $S_0(\varepsilon | Y = 1)$ 'in tahmini;

$$\hat{S}_0(\varepsilon | Y = 1) = \exp \left(- \sum_{j: t_{(j)} < \varepsilon} \frac{d_{\tau_j}}{\sum_{i \in R(\tau_j)} w_i^{(m)}} \right) \quad (3.14)$$

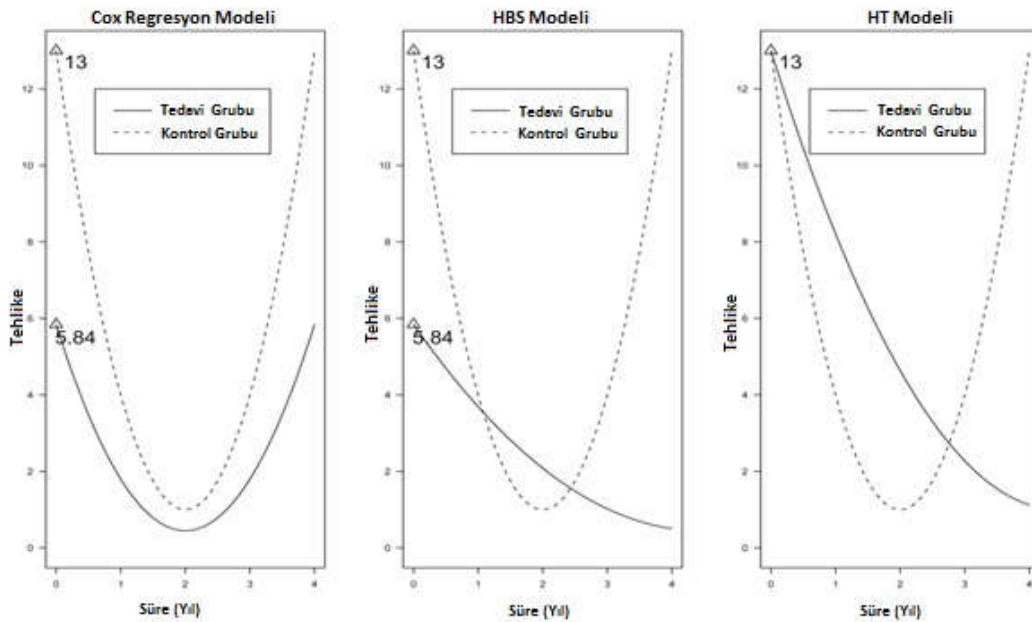
Cox karma iyileşme modeline benzer olarak, $\varepsilon > \tau_k$ için, $\hat{S}_0(\varepsilon | Y = 1) = 0$ olarak ayarlanabilir. Bundan dolayı, $\hat{S}(t | Y = 1) = \hat{S}_0(\varepsilon | Y = 1)$ olacaktır.

3.3.1.3. Yarı Parametrik Hızlandırılmış Tehlikeler Karma İyileşme Modeli

İyileşme modellerinde, Cox karma iyileşme modeli ve yarı parametrik HBS karma iyileşme modeli, tahmin yöntemlerinin ve elde edilen sonuçların yorumlanmasının kolaylığı nedeniyle oldukça popülerdir. Ancak, bu modellerdeki varsayımlar her zaman sağlanamayabilir. Varsayımların uygun olmadığı durumlar için kullanılması önerilen alternatif bir model, hızlandırılmış tehlikeler (HT) karma iyileşme modelidir. HBS karma iyileşme modelinde bulunan ani etki varsayımı, tedavinin etkisinin zamanla birlikte

kademeli olarak arttığı bazı kanser çalışmaları için tercih edilen bir varsayım olmayabilir. Örneğin, antidepresanların etkisini test ederken, ilacın, tedavinin erken aşamalarında etkili olduğunu değil, $t=0$ zamanında etkisiz olduğunu ve zaman ilerledikçe tedavinin etkisinin kademeli olarak arttığını varsaymak daha uygun olabilmektedir [37]. Bu model, tedavinin etkisinin sıfır zamanından, zaman ilerledikçe kademeli olarak arttığı kanser tedavileri için de kullanışlıdır. CRM ve HBS modellerinden farklı olarak, HT modelinin, özellikle tedavinin ya da diğer açıklayıcı değişken etkilerinin zaman ile kademeli olarak arttığı durumları modellemek için yararlı olduğu görülmüştür [37].

Hızlandırılmış tehlikeler modeli, ilk olarak Chen ve Wang tarafından [60] iyileşmiş kısmı olmayan yaşam verilerinde ele alınmıştır. Diğer karma iyileşme modellerinin aksine bu yeni model, iyileşmemiş hastaların başarısızlık zamanlarının dağılımında, açıklayıcı değişken etkilerinin başlangıç zamanında göz ardı edilebilmesine ve bu açıklayıcı değişken etkilerinin zaman ilerledikçe artmasına olanak sağlamaktadır. Daha önce önerilen klasik modeller bunun gibi durumlarda çok başarılı sonuçlar verememektedir.



Şekil 3.3. CRM, HBS ve HT modelleri için tehlike eğrileri

Şekil 3.3'de, farklılıkları daha iyi ifade edebilmek için, aynı veri kümesine uygulanan üç farklı modelden elde edilen tehlike fonksiyonu grafikleri verilmiştir [37]. Kontrol grubu için $x = 0$ ve tedavi grubu için $x = 1$ olmak üzere iki grup ele alınmıştır. Temel tehlike fonksiyonu, sağlık araştırmalarında sıklıkla kullanılan U-biçimli bir fonksiyondur. β değeri -0.8 'e ayarlanmıştır. İki grubun tehlike eğrilerini karşılaştırsak, CRM'nin, uygulanan tedavi ile tehlike oranının tüm zaman aralığı için $e^{-0.8} = 0.45$ azaldığını gösterdiğini söyleyebiliriz. HBS modelinde ise iki grubun tehlike oranları arasındaki ilişki daha karmaşıktır. Tedavi, başlangıçta daha düşük tehlike oranına sahip iken ilerleyen zamanda daha yüksek tehlike oranına ulaşmış, iki periyod sonra ise tekrar daha düşük tehlike oranına sahip olmuştur. HT modeline bakıldığında, tedavi grubu ile kontrol grubu aynı tehlike oranı ile başlar, çalışmanın erken aşamalarında tedavi grubu, uygulanan tedavinin toksikliği gibi nedenlerle, kontrol grubundan daha yüksek tehlike oranına sahip olduğu görülür. Ancak belirli bir zaman noktasından sonra tedavinin pozitif etkisi, tedavi grubunun kontrol grubuna kıyasla çok daha düşük tehlike oranına sahip olması ile ispatlanmıştır [37].

İyileşmiş kısma sahip olmayan bir veri kümesinde kademeli tedavi etkisini modellemek için Chen tarafından geliştirilen [61] HT modelinin tehlike fonksiyonu;

$$h(t | x) = h_{u0}(te^{\beta x}) \quad (3.15)$$

ile ifade edilebilir.

Chen ve Wang [60], Eşitlik 3.15'deki parametreleri, yarı parametrik olarak tahmin etmek için tahmin eşitliklerini önermiştir. İyileşmemiş hastaların başarısızlık zamanlarındaki açıklayıcı değişkenlerin kademeli etkilerine imkan sağlamak için, karma iyileşme modelinde iyileşmemiş hastalar için yaşam fonksiyonunu ifade eden $S_u(t)$ 'nin hızlandırılmış tehlikeler modeli kullanılarak modellenmesi önerilmiştir [61]. $h_{u0}(t)$ temel tehlike fonksiyonu ve $S_{u0}(t)$ buna karşılık gelen iyileşmemiş hastalar için yaşam fonksiyonu iken,

$$h_u(t | x) = h_{u0}(te^{x\beta})$$

ya da

$$S_u(t | x) = S_{u0}(te^{x\beta})^{\exp(-x\beta)} \quad (3.16)$$

ise bu model hızlandırılmış tehlikeler karma iyileşme modeli olarak adlandırılmaktadır [37].

Eğer $h_{u0}(t)$, hızlandırılmış tehlikeler karma iyileşme modelinde az sayıda bilinmeyen parametreler ile belirtilebilirse, modeldeki parametreler en çok olabilirlik yaklaşımı kullanılarak tahmin edilebilir.

Chen [61] Eşitlik 3.15'deki parametrelerin tahminleri için yarı parametrik bir model önermiştir. İyileşmiş hastaların varlığı sebebiyle Chen'in bu yöntemi uygulanabilir değildir ve yeni bir tahmin yöntemi gerekmektedir. HT karma iyileşme modelindeki parametreleri tahmin etmek için, EM algoritmasına dayalı yarı parametrik bir yöntem önerilmiştir [37].

$i = 1, \dots, n$ olmak üzere t_i i. hasta için gözlemlenen yaşam süresi (durdurulmuş olabilir), δ_i , durdurulmuş t_i için 0, durdurulmamış t_i için 1 değerini alan gösterge değişkeni, x_i ve z_i ise i. birey için x ve z açıklayıcı değişkenlerinin gözlemlenen değerleri iken, $(t_i, \delta_i, z_i, x_i)$, i. bireyin gözlemlenen verileri olsun. Durdurulmuş bir gözlem için y 'nin gizli değişken olduğu ve değerinin gözlemlenemeyeceği açıktır. $y = (y_1, \dots, y_n)$ olarak ifade edilsin. y 'nin tüm değerlerinin mevcut olduğunu varsaydığımızda EM algoritmasındaki log olabilirlik fonksiyonu;

$$l_{c1}(y; y) = \sum_{i=1}^n y_i \log[\pi(Z_i)] + (1 - y_i) \log[1 - \pi(Z_i)]$$

ve

$$l_{c2}(\beta, h_{u0}(\cdot); y) = \sum_{i=1}^n y_i \delta_i \log[h_{u0}(e^{\beta^T x_i} t_i)] + y_i \log[S_{u0}(e^{\beta^T x_i} t_i)^{\exp(-\beta^T x_i)}]$$

iken,

$$l(\beta, h_{u0}(t), y; y) = l_{c1}(y; y) + l_{c2}(\beta, h_{u0}(t); y) \quad (3.17)$$

olarak verilmiştir.

E adımı, mevcut tahminler $\Theta^{(m)} = \{\gamma^{(m)}, \beta^{(m)}, h_{u_0}^{(m)}(t)\}$ verilmiş iken, tam log-olabilirliğin y 'ye göre koşullu tahmini $E[l_c(\gamma, \beta, h_{u_0}(t); y | \Theta^{(m)})]$ 'yi hesaplar.

$$w_i^{(m)} = E(y_i | \Theta^{(m)}) = \delta_i + (1 - \delta_i) \frac{\pi(z_i) S_{u_0}(e^{\beta^T x_i} t_i) e^{-\beta^T x_i}}{1 - \pi(z_i) + S_{u_0}(e^{\beta^T x_i} t_i) e^{-\beta^T x_i}} \quad (3.18)$$

olmak üzere $w^{(m)} = (w_1^{(m)}, \dots, w_n^{(m)})$ olarak alınırsa,

$$\begin{aligned} E[l_c(\gamma, \beta, h_{u_0}(t); y) | \Theta^{(m)}] &= l_c[\gamma, \beta, h_{u_0}(t); w^{(m)}] \\ &= l_{c1}[\gamma; w^{(m)}] + l_{c2}[\beta, h_{u_0}(t); w^{(m)}] \end{aligned} \quad (3.19)$$

olduğu görülebilir [37].

M basamağı, $l_{c1}[\gamma; w^{(m)}]$ ve $l_{c2}[\beta, h_{u_0}(t); w^{(m)}]$ 'yi, γ , β ve $h_{u_0}(t)$ 'ye göre maksimize eder.

$l_{c1}[\gamma; w^{(m)}]$ γ 'ya göre Newton-Raphson yöntemi kullanılarak kolaylıkla maksimize edebilebilir. $l_{c2}[\beta, h_{u_0}(t); w^{(m)}]$ 'yi β ve $h_{u_0}(t)$ 'ye göre maksimize etmek ise daha zordur. β ve $h_{u_0}(t)$ 'yi güncellemek için rank benzeri tahmin yöntemi önerilmiştir. $\delta_i \log w_i^{(m)} \equiv 0$ olduğundan, $l_{c2}[\beta, h_{u_0}(t); w^{(m)}]$ aşağıdaki gibi yazılabilir;

$$\sum_{i=1}^n \delta_i \log[w_i^{(m)} h_{u_0}(t_i e^{\beta^T x_i})] + [\log(S_{u_0}(e^{\beta^T x_i} t_i) w_i^{(m) \exp(-\beta^T x_i)})] \quad (3.20)$$

ve Chen [61] tarafından ele alınan $w_i^{(m)} h_{u_0}(t_i e^{\beta^T x_i})$ tehlike fonksiyonuna sahip HT modelin log-olabilirlik fonksiyonu gibi davranılabilir. Chen [61] ve Zhang ve Peng [58]'in çalışmalarından yola çıkarak, $k(\cdot)$ genel (öngörülebilir) ağırlık fonksiyonu iken, β 'nın rank tipi tahmin denklemi aşağıdaki gibi yazılabilir;

$$\Psi(\beta; k(\cdot)) = \sum_{i=1}^n \delta_i k(t_i e^{\beta^T x_i}) \times \left(x_i - \frac{\sum_{j=1}^n x_j w_j^{(m)} e^{-\beta^T x_j} I(t_j e^{\beta^T x_j} \geq t_i e^{\beta^T x_i})}{\sum_{j=1}^n w_j^{(m)} e^{-\beta^T x_j} I(t_j e^{\beta^T x_j} \geq t_i e^{\beta^T x_i})} \right). \quad (3.21)$$

Burada $k(\cdot)$ ağırlığı için Gehan tipi ağırlık;

$$k(u) = \frac{1}{n} \sum_{j=1}^n w_j^{(m)} e^{-\beta^T x_j} I(t_j e^{\beta^T x_j} \geq u) \quad (3.22)$$

tercih edildiğinde, buna karşılık gelen tahmin eşitliği aşağıdaki gibi yazılabilir;

$$\Psi(\beta) = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n \delta_i(x_i - x_j) w_j^{(m)} e^{-\beta^T x_j} I(t_j e^{\beta^T x_j} \geq t_i e^{\beta^T x_i}) \quad (3.23)$$

Eşitlik 3.22'de verilen Gehan tipi ağırlık kullanmanın avantajı, Eşitlik 3.23'ün, β 'nin süreksiz ama monoton bir fonksiyonu olmasıdır. $k(\cdot)$ için diğer ağırlık fonksiyonları da dikkate alınabilir, ancak bunlara karşılık gelen tahmin denklemi $\Psi(\beta; k(\cdot))$ β 'nin monoton bir fonksiyonu değildir ve köklerini bulmak zor olabilir.

$\beta^{(m+1)}$ verildiğinde, β 'nin güncellenmiş tahmini, $h_{u_0}(t)$ 'nin parametrik olmayan bir tahmini, $t_i e^{\beta^{(m+1)T} x_i}$ artıklarına dayanarak elde edilebilir [62, 11]. Örneğin, $\tau_1 < \tau_2 \dots < \tau_k$, ayrıntı durdurulmamış artıklar olsun, $R(\tau_j)$ risk kümesini gösterebilir. Geçerli M basamağındaki $H_{u_0}(t)$ tahmini,

$$\hat{H}_{u_0}^{(m+1)}(t) = \exp \left(- \sum_{\substack{j: \tau_j < t \\ i \in R(\tau_j)}} \frac{d_{\tau_j}}{\sum w_i^{(m)} \exp(-\beta^{(m+1)T} x_i)} \right) \quad (3.24)$$

ve

$$\hat{S}_{u_0}^{(m+1)}(t_i e^{\beta^{(m+1)T} x_i}) = \exp \left[- \hat{H}_{u_0}^{(m+1)}(e^{\beta^{(m+1)T} x_i} t_i) \right] \quad (3.25)$$

olarak elde edilir ve $t_i e^{\beta^{(m+1)T} x_i} > \tau_k$ ise 0'dır. $\beta^{(m+1)}$ ve $\hat{S}_{u_0}^{(m+1)}(t_i e^{\beta^{(m+1)T} x_i})$ ile Eşitlik 3.18'deki w_i güncellenebilir ve EM algoritması, yakınsama gerçekleşene kadar çalışmaya devam eder.

HT karma iyileşme modelinde tahmin edilen parametrelerin varyanslarını elde etmek, Eşitlik 3.23'e karşılık gelen tam log olabilirlik fonksiyonu mevcut olmadığı için açık ve net değildir. EM algoritması için önerilen klasik yöntemler [63, 64], varyansların tahminlerini elde etmek için kullanılamaz. Bootstrap yöntemi tahminlerin varyanslarını

tahmin etmek için daha kolay bir hesaplama yöntemi elde edilene kadar kullanılabilir [37].

3.3.1.4. Parametrik Karma İyileşme Modeli

En genel haliyle karma iyileşme modelleri;

$$S(t) = S^*(t)((1-\pi) + \pi S_u(t)) \quad (3.26)$$

olarak ifade edilmektedir. Burada önceki modellerde belirttiğimiz gibi $(1-\pi)$ iyileşmiş kısmın oranı, $S_u(t)$ ya da $S(t|Y=1)$ iyileşmemiş hastalar için yaşam fonksiyonudur.

$S^*(t)$ beklenen yaşam fonksiyonu, $h^*(t)$ beklenen tehlike fonksiyonudur, bu fonksiyonlar genellikle ulusal ya da bölgesel yaşam çizelgeleri gibi veri kaynaklarından elde edilir ve bilindiği varsayılır [65].

Bu modele karşılık gelen tehlike fonksiyonu, $f(t)$, $F(t)$ 'ye karşılık gelen yoğunluk fonksiyonu iken;

$$h(t) = \frac{\pi f(t)}{S(t)} \quad (3.27)$$

olarak ifade edilebilir ve;

$$f(t) = \pi f_u(t) \quad (3.28)$$

biçimindedir. $F(0)=0$ ve $F(\infty)=1$ olduğu ve bu nedenle $S(0)=1$ ve $S(\infty)=1-\pi$ olacağı dikkate alınmalıdır. Bu eşitlikte $f_u(t)$ iyileşmemiş kısım için olasılık yoğunluk fonksiyonudur.

Parametrik karma iyileşme modelleri için çeşitli parametrik dağılımlar değerlendirilebilir. İyileşmiş kısmın tahmini, hangi parametrik dağılımın seçileceğine karşı duyarlıdır. Weibull dağılımı en sık kullanılan dağılımdır ve çoğu veri kümesi için uygun sonuçlar verecektir. Weibull dağılımı iki durum için; iyileşme oranı çok yüksek olduğunda (>%80) ya da takip süresinin ilk haftalarında yüksek ölüm oranı gözlemlenmesi durumunda çok doğru sonuç vermeyebilir. Bu ikinci sorun genellikle ileri yaşlardaki hastalarda gözlenmektedir. Lognormal dağılım, kanser çalışmalarında iyileşme modelleri için nadiren iyi bir tahmin sağlar. Genelleştirilmiş gamma dağılımı, Weibull, üstel ve

lognormal dağılımlarını kapsadığı için, genellikle kullanışlıdır. Ancak, yakınsama sorunlarıyla karşılaşılabilir ve bu sorun genellikle Weibull dağılımı, iyileşmiş kısım için iyi bir tahmin sağlamadığında meydana gelmektedir [65].

İyileşmemiş hastalar için yaşam fonksiyonu $S_u(t)$ 'nin modellenmesi için en sık kullanılan dağılımlar üstel, Weibull, log-normal ve log-lojistik dağılımlardır. $S_u(t)$ tercih edilen dağılıma göre aşağıdaki gibi ifade edilebilmektedir;

$$S_u(t) = \begin{cases} \text{üstel dağılım için, } \exp[-\exp(\log t - \mu)], \\ \text{Weibull dağılımı için, } \exp\left[-\exp\left(\frac{\log t - \mu}{\sigma}\right)\right], \\ \text{log-normal dağılım için, } 1 - \Phi\left(\frac{\log t - \mu}{\sigma}\right), \\ \text{log-lojistik dağılım için } \left[1 + \exp\left(\frac{\log t - \mu}{\sigma}\right)\right]^{-1}. \end{cases}$$

μ 'nün, γ bilinmeyen regresyon parametrelerini ifade ederken $\mu = \gamma'x$ gibi parametreleştirilmesi ile açıklayıcı değişkenler modele dahil edilebilir. Bu modeller 'parametrik HBS iyileşme modelleri' olarak da adlandırılmaktadır [66].

Parametrik karma iyileşme modellerindeki $F(t)$ fonksiyonu, parametrik iyileşme modellerinin çekirdeği (kerneli) olarak adlandırılmaktadır. $(0, \infty)$ aralığındaki herhangi bir uygun dağılım fonksiyonu, parametrik iyileşme modellerinin çekirdeği olarak alınabilir [50]. Genellikle EÇÖ kriteri kullanılarak en iyi uyum sağlayan dağılım seçilir [67]. Örneğin, çekirdek $f(t)$, $\Phi(\cdot)$ standart normal dağılım fonksiyonu olmak üzere, log-normal dağılım ile aşağıdaki gibi tanımlanabilir;

$$F(t) = \Phi(\ln(\lambda t)^\gamma) \quad (3.29)$$

Benzer biçimde $f(t)$, λ ölçü parametresi ve γ şekil parametresi olmak üzere, Weibull dağılımı;

$$F(t) = 1 - e^{-(\lambda t)^\gamma} \quad (3.30)$$

ile de ifade edilebilir. Bu modeller, standart EÇO yöntemleri kullanılarak veriye uydurulur [67].

Hem iyileşme oranı $1-\pi$, hem de çekirdek $F(t)$, açıklayıcı değişkenlere bağlı olabilir. Eğer z_1, z_2, \dots, z_p açıklayıcı değişkenleri temsil ediyorsa (örneğin cinsiyet, hastalık evresi, yaş gibi), bu değişkenler iyileşen kısmı,

$$1 - \pi = \exp(-\exp\{a_0 + a_1 z_1 + a_2 z_2 + \dots + a_p z_p\}) \quad (3.31)$$

fonksiyonunda da görülebileceği gibi etkileyecektir.

Burada kısmen karmaşık bir fonksiyon seçilmiştir. Çünkü a_1, \dots, a_p parametreleri, klasik CRM çözümlerinden elde edilen parametreler ile benzer ve orantılı tehlikeler varsayımı altındaki karma olmayan modellerden elde edilenler ile de kesinlikle benzer olacaktır [20].

İyileşen kısım başka biçimlerde de modellenebilir, örneğin iyileşmiş kısım;

$$1 - \pi = a_0 + a_1 z_1 + \dots + a_p z_p \quad (3.32)$$

ile verilen basit doğrusal bir fonksiyon ile de modellenebilir. Bu fonksiyon, açıklayıcı değişkenlerin etkilerini görmemize olanak sağlamaktadır. Ayrıca, iyileşmiş kısım Farewell tarafından [50] parametrik iyileşme modellerinde kullanılan olasılıkların bir dönüşümü olan lojistik fonksiyon;

$$\pi = \frac{\exp\{a_0 + a_1 z_1 + \dots + a_p z_p\}}{1 + \exp\{a_0 + a_1 z_1 + \dots + a_p z_p\}} \quad (3.33)$$

ile de modellenebilir. İyileşmiş kısmı da içeren modellerdeki parametreler, CRM'deki parametrelere doğrudan benzer değildir ve çok değişkenli çözümlemede, farklı gruplar için iyileşen kısmın tahmininde farklı sonuçlar verebilirler. Çünkü, ölçüm birimlerinde farklılık söz konusudur [20].

Parametrik karma iyileşme modellerinde parametreler EÇO yöntemi kullanılarak tahmin edilir. t_i , i. bireyin yaşam/durdurulma zamanı, d_i durdurma göstergesi, t_{0i} i. bireyin çalışmaya girme zamanı ise, i. bireyin log-olabilirlik fonksiyonuna katkısı;

$$\ln L_i = d_i \ln[h(t_i)] + \ln[S(t_i)] - \ln[S(t_{0i})]$$

olarak gösterilebilir.

Parametrik karma iyileşme modelleri için tanımlanmış tehlike ve yaşam fonksiyonları, i. bireyin log-olabilirliğe katkısını hesaplamakta kullanılır ve;

$$\ln L_i = d_i \ln \left[h^*(t_i) + \frac{\pi f_u(t_i)}{(1-\pi) + \pi S_u(t_i)} \right] + \ln[S^*(t_i)] + \ln[(1-\pi) + \pi S_u(t_i)] - \ln[S^*(t_{0i})] - \ln[(1-\pi)S_u(t_{0i})]$$

biçimindedir. Parametrik karma iyileşme modelinin log-olabilirlik fonksiyonu için $\ln(S^*(t_i))$ ve $\ln(S^*(t_{0i}))$ terimleri model parametrelerine bağlı değildir ve parametrelerin tahminini elde etmek için maksimizasyon yaparken kullanılması gerekli değildir.

3.3.2. Karma Olmayan İyileşme Modelleri

Karma olmayan iyileşme modelleri, 'sınırlı birikimli tehlike modeli' ya da 'ilerleme zamanlı (promotion time) iyileşme modeli' olarak da adlandırılmaktadır ve standart karma iyileşme modellerine alternatif olarak Yakovlev ve diğerleri tarafından [25] önerilmiştir. Bu model, Chen ve diğerleri [28], Ibrahim ve diğerleri [68], Yin [69], Yin ve Ibrahim [35] ve Zeng ve diğerleri'nin [70] çalışmalarında incelenmiştir.

Kitledeki bir birey ele alındığında, N, ilk tedavinin ardından, bu hastada aktif halde kalan metastaz özellikli tümör hücrelerini ifade etmektedir. Metastaz özellikli tümör hücreleri, metastaza meyilli olan tümör hücreleridir. N'in θ parametresi ile Poisson dağılıma sahip olduğu varsayılmıştır. Z_i , i. metastaza meyilli tümör hücresinin, fark edilebilir bir hastalığa neden olduğu rastgele zamanı göstermektedir. Burada Z_i , i. tümör hücresi için ilerleme zamanı (promotion time) olarak düşünülebilir. N verildiğinde, Z_i , $i=1, 2, \dots$, rastgele değişkenlerinin, N'e bağımlı olmayan $f(t) = 1 - S(t)$ dağılım fonksiyonu ile bağımsız ve aynı dağıldığı varsayılmıştır. Kanserin nüksetme zamanı, $P(Z_0 = \infty) = 1$ iken, $T = \min\{Z_i, 0 \leq i \leq N\}$ olarak tanımlanmıştır. T için yaşam fonksiyonu ve kitle için yaşam fonksiyonu,

$$S(t) = P(\text{t zamanına kadar hiç metastaz olmaması}) \\ = P(N=0) + P(Z_1 > t, \dots, Z_N > y, N > 1)$$

olur. Burada bazı düzenlemeler yapılırsa aşağıdaki eşitlikler elde edilir:

$$S(t) = \exp(-\theta) + \sum_{k=1}^{\infty} S(t)^k \frac{\theta^k}{k!} \exp(-\theta) \quad (3.34) \\ = \exp(-\theta + \theta S(t)) = \exp(-\theta F(t))$$

Burada $\theta = \exp(\beta' x)$ 'dir.

$S = \exp(-\theta) > 0$ olduğu için, Eşitlik 3.34 uygun bir yaşam fonksiyonu değildir. Yakovlev ve Tsodikov [71] Eşitlik 3.34'ün, tümör büyümesinin iki farklı karakteristiğinin (metastaza meyilli tümör hücrelerinin başlangıçtaki sayısı ve ilerleme hızı), başarısızlık zamanına olan katkısını açıkça gösterdiğini belirtmişlerdir. Buradan, modelin biyolojik olarak anlam taşıyan parametreleri içerdiği söylenebilir. Bu biyolojik etkenin yanı sıra, Eşitlik 3.34'deki model, iyileşmiş kısmı bulunan bütün yaşam verilerine uygundur. Böylece, verinin iyileşen kısma sahip olduğu ve bilinmeyen N tane gizli yarışan riskler (Z_i 'ler) ile üretilebildiği düşünülürse, biyolojik tanıma uymayan yaşam verileri de Eşitlik 3.34 ile modellenenbilmektedir. Böylece bu model, kanserin nüksetme zamanı, ölüm zamanı, ilk enfeksiyon zamanı vb. içeren, değişik biçimlerdeki yaşam verilerini modellemek için yararlı olmaktadır.

Eşitlik 3.34'den iyileşmiş kısmın aşağıdaki gibi verildiği görülebilir;

$$S(\infty) = P(N=0) = \exp(-\theta) \quad (3.35)$$

İyileşmiş kısım, $\theta \rightarrow \infty$ iken sıfıra $\theta \rightarrow 1$ iken ise bire yakınsar. Eşitlik 3.35'e karşılık gelen olasılık yoğunluk fonksiyonu;

$$f(t) = \theta f(t) \exp(-\theta F(t))$$

biçimindedir. $t \rightarrow \infty$ iken $\lambda \rightarrow 0$ olacağını ve $\int \lambda(t) < \infty$ olduğunu da dikkate almak gerekmektedir. Kitlenin tehlike fonksiyonu,

$$h(t) = \frac{f(t)}{S(t)} = \theta f(t)$$

olarak elde edilmiştir. Buradaki tehlike fonksiyonu $h(t)$, yaşam çözümlenmesinde istenilen bir özellik olan orantılı tehlikeler yapısına sahiptir. Karma iyileşme modellerinden farklı olarak, karma olmayan iyileşme modelleri, orantılı tehlikeleri özel bir durum olarak modele dahil etme avantajına sahiptirler [59] .

İlerleme zamanlı iyileşme modelleri, karma iyileşme modelleri gibi de yazılabilir. Örneğin, Eşitlik 3.34;

$$\begin{aligned} S(t) &= \exp(-\theta) + \exp(-\theta)\exp(S(t)\theta - 1) \\ &= \exp(-\theta) + (1 - \exp(-\theta)) \frac{\exp(-\theta F(y)) - \exp(-\theta)}{(1 - \exp(-\theta))} \end{aligned}$$

olarak da ifade edilebilir. Kitledeki iyileşmemiş kısım için yaşam fonksiyonu,

$$S(t) = P(T > t | N \geq 1) = \frac{\exp(-\theta F(y)) - \exp(-\theta)}{(1 - \exp(-\theta))}$$

olarak elde edilmiştir.

Kitle uygun olmayan bir yaşam fonksiyonuna sahip olduğundan ve θ parametresini içerdiğinden, ilerleme zamanlı iyileşme modelleri ile ilgili yapılan çalışmalar genellikle Bayesci yaklaşım kapsamındadır. Yin ve Ibrahim [35], Box-Cox dönüşümünü kullanarak, iyileşme modellerinin genel bir sınıfını geliştirmişlerdir. Bu model, $a \in [0, 1]$ olmak üzere,

$$\frac{S(t | \mathbf{X}_i, \mathbf{Z}_i)^a - 1}{a} = -\theta(a, \mathbf{Z}_i)F(t | X_i)$$

biçiminde verilmiştir. Burada, \mathbf{X}_i ve \mathbf{Z}_i i. bireye karşılık gelen açıklayıcı değişkenlerin vektörleri, a ise dönüşüm parametresidir. Bu model, $a=1$ iken karma iyileşme modeline, $a=0$ iken ise ilerleme zamanlı iyileşme modeline dönüşmektedir [51].

4. UYGULAMA

Çalışmada IBM SPSS Statistics 23 ve R programı kullanılarak sonuçlar elde edilmiştir.

4.1. Glioma Veri Kümesi

Çalışmada, agresif bir beyin kanseri türü olan glioma hastası 411 bireyi içeren veri kümesi [72] ele alınmıştır. Hastalara iki çeşit kemoterapi tedavisi uygulanmıştır. Çalışma sürecinde 411 hastanın 274'ü hayatını kaybetmiştir ve bu hastalar başarısız olarak ele alınmıştır. Çalışmadaki diğer hastalar ise durdurulmuş olarak nitelendirilmiştir. Buna göre 411 hastadan 274'ünde (%66.7) başarısızlık ve 137'sinde (%33.3) durdurma gözlenmiştir. Yaşam süresi hastaların ölümüne ya da durdurulmasına kadar geçen süre (gün) olarak tanımlanmıştır. Çalışma için takip süresi 1617 gündür. Tedavi değişkeni dışında, yaşam süresini etkilemesi olası olan dört değişken daha çalışmaya dahil edilmiştir. Çalışmada kullanılan değişkenler ve düzeyleri Çizelge 4.1'de verilmiştir. Yaş değişkeni ortanca yaşa göre sınıflandırılarak kullanılmıştır.

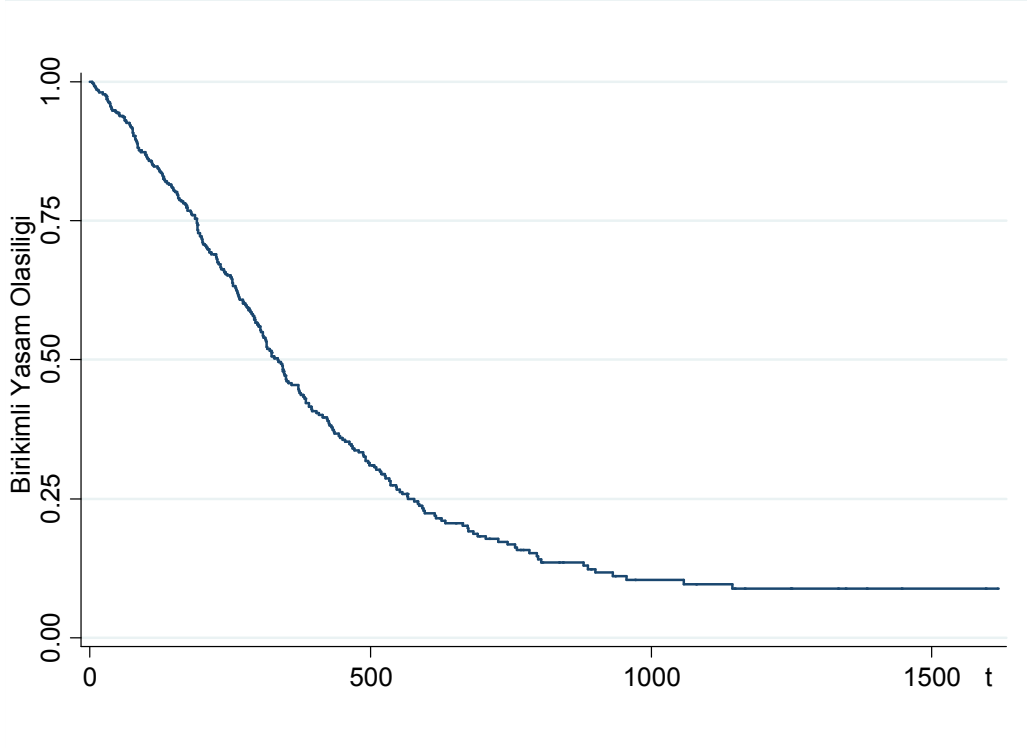
Çizelge 4.1. Glioma çalışmasında kullanılan değişkenler ve düzeyleri

Değişken	Değişken Düzeyleri	n (%)	Başarısız Gözlemlerin Sayısı	Durdurulmuş Gözlemlerin Sayısı
Malignite Düzeyi	0: Malignite düzeyi 1 ve 2 olanlar 1: Malignite düzeyi 3 olanlar	139 (%33.8) 272 (%66.2)	74 200	65 72
Cinsiyet	0 : Erkek 1 : Kadın	233 (%56.7) 178 (%43.3)	158 116	75 62
Yaş	0: Yaş ≤ 54 olanlar 1: Yaş > 54 olanlar	212 (%51.6) 199 (%48.4)	123 151	89 48
Karnofsky İndeksi	0: Karnofsky indeksi ≤ 80 olanlar 1: Karnofsky indeksi > 80 olanlar	312 (%75.9) 99 (%24.1)	219 55	93 44
Tedavi	0 : Standart 1 : Yeni	204 (%49.6) 207 (%50.4)	137 137	67 70

Çalışma için ortanca yaşam süresi 335 ± 17.332 gündür. 5 yıllık yaşam olasılığı ise 0.088 ± 0.021 olarak elde edilmiştir.

4.1.1. Parametrik Olmayan Yaşam Çözümlemesi Sonuçları

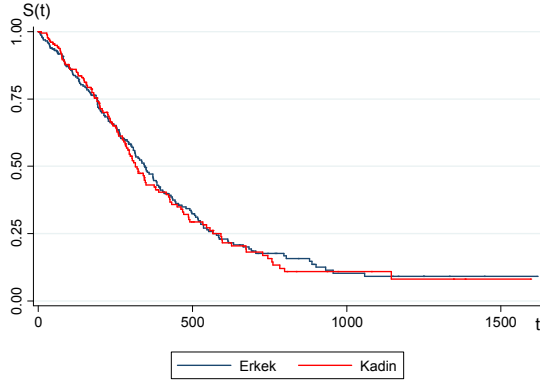
Çalışmada kullanılan glioma veri kümesine ait Kaplan-Meier yaşam eğrisi Şekil 4.1'deki biçimde elde edilmiştir.



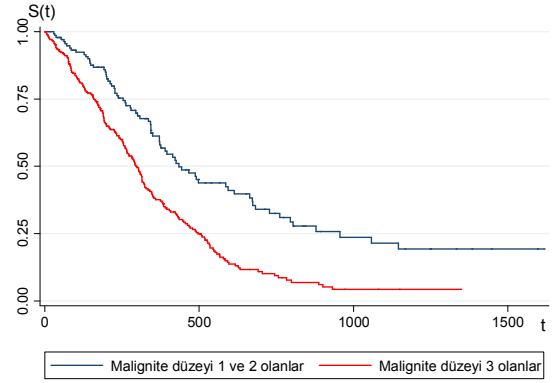
Şekil 4.1. Glioma verisi için Kaplan-Meier yaşam eğrisi

Glioma veri kümesi için elde edilen Kaplan-Meier yaşam eğrisine bakıldığında eğrinin düz bir kuyruğa sahip olduğu görülmektedir. Kitlede ilgilenilen olayı, yani ölümü, yaşamayan bireylerin varlığı düşünülebilir ve iyileşme modellerinin kullanımının değerlendirilmesi daha iyi sonuçlar elde edilmesini sağlayabilir.

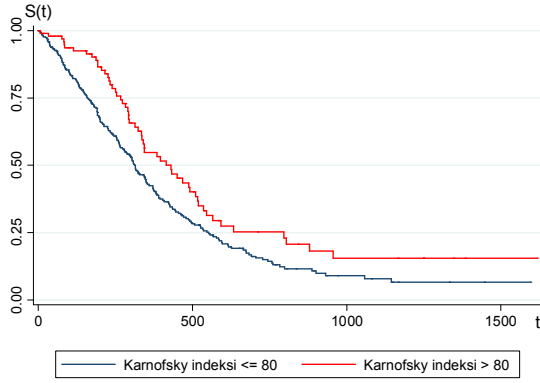
Çalışmada kullanılan her bir açıklayıcı değişken için Kaplan-Meier yaşam eğrileri ise Şekil 4.2'de verilmiştir.



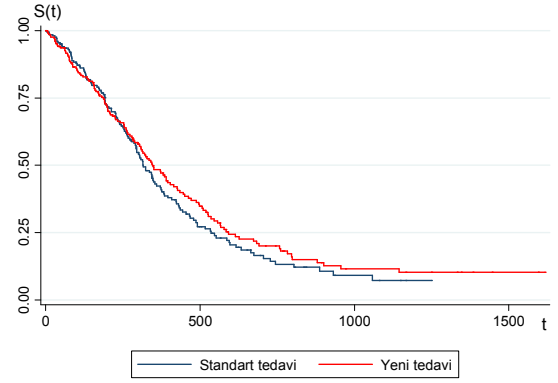
(a) Cinsiyet



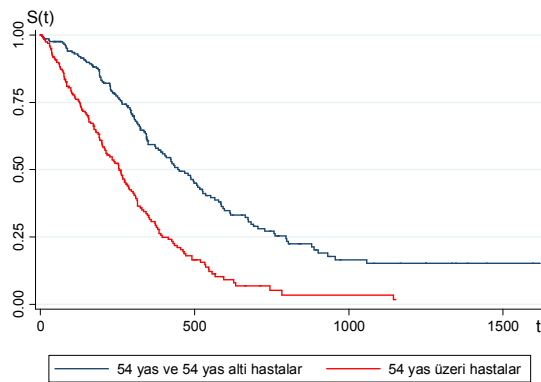
(b) Malignite Derecesi



(c) Karnofsky İndeksi



(d) Tedavi



(e) Yaş

Şekil 4.2. Glioma verisinde kullanılan açıklayıcı değişkenlerin Kaplan-Meier eğrileri

Şekil 4.2 incelendiğinde, cinsiyet ve tedavi değişkenlerinde değişken düzeylerinin yaşam eğrileri birbirine çok yakın olarak elde edilmiştir. Diğer değişkenler için değişken düzeyleri arasında farklılığın daha fazla olduğu görülmektedir. Malignite düzeyi 1 ve 2 olan hastaların yaşam olasılıklarının, malignite düzeyi 3 olan hastalara göre daha yüksek olduğu söylenebilir. Yaşı 54 ve 54'den küçük olan hastaların yaşam olasılıklarının, yaşı 54 üzeri olan hastaların yaşam olasılıklarından yüksek olduğu görülmektedir. Ayrıca Karnofsky indeksi 80'den büyük olan hastaların, Karnofsky indeksi 80 ve 80'den küçük olan hastalara göre daha yüksek yaşam olasılıklarına sahip oldukları görülmektedir.

Açıklayıcı değişken düzeyleri için yaşam olasılıkları bakımından istatistiksel olarak anlamlı bir farklılık olup olmadığı ise log-rank testi kullanılarak incelenmiş ve Çizelge 4.2'deki sonuçlar elde edilmiştir.

Çizelge 4.2. Glioma verisi için log-rank testi sonuçları

Değişken	Ki-kare	p değeri
Cinsiyet	0.10	0.755
Malignite Derecesi	26.51	0.000
Karnofsky İndeksi	7.71	0.005
Tedavi	0.77	0.380
Yaş	49.60	0.000

Log-rank testi sonuçlarına göre, malignite derecesi ($p=0.0000$), karnofsky indeksi ($p=0.005$) ve yaş ($p=0.000$) değişkenlerinin düzeyleri arasında, yaşam olasılıkları bakımından istatistiksel olarak anlamlı bir farklılık olduğu %95 güven düzeyinde söylenebilir.

4.1.2. Cox Regresyon Modeli Sonuçları

Cox regresyon çözümlemesinde, değişken düzeylerinin yorumlanması amacıyla düzeylerden biri referans kategorisi olarak alınıp yorumlar yapılmaktadır. Bu çalışmada, tüm değişkenler için birinci düzey referans kategorisi olarak alınmıştır.

Glioma veri kümesi için CRM kullanıldığında p değeri 0.0000 olarak elde edildiğinden modelin anlamlı olduğu %95 güven düzeyinde söylenilebilmektedir. CRM için AIC

değeri 2739.576, BIC değeri 2759.576 ve -2LogL değeri 2729.576 olarak elde edilmiştir. Glioma verisi için CRM sonuçları Çizelge 4.3'te verilmiştir.

Çizelge 4.3. Glioma verisi için CRM'den elde edilen sonuçlar

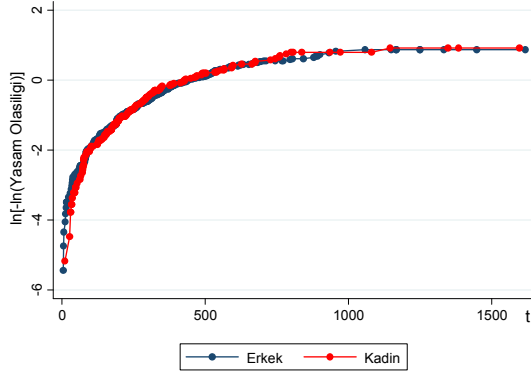
Değişken	$\hat{\beta}$	Std. hata	p-değeri	$\exp(\hat{\beta})$	$\exp(\hat{\beta})$ için %95 güven aralığı
Cinsiyet	-0.128	0.125	0.308	0.880	0.688, 1.125
Malignite Derecesi	0.569	0.139	0.000	1.766	1.344, 2.320
Yaş	0.781	0.129	0.000	2.183	1.696, 2.810
Karnofsky İndeksi	-0.292	0.153	0.056	0.746	0.553, 1.007
Tedavi	-0.086	0.121	0.479	0.918	0.724, 1.164

Çizelge 4.3'de açıklayıcı değişkenler için p değerleri incelendiğinde, yaş ($p=0.000$) ve malignite derecesi ($p=0.000$) değişkenlerinin, glioma kanseri veri kümesi için yaşam süresi üzerinde etkili olan önemli risk faktörleri olduğu görülmüştür. CRM'den elde edilen sonuçlara göre, malignite derecesi 3 olan hastaların, malignite derecesi 1 ve 2 olan hastalara göre başarısızlık riskinin 1.766 kat ($e^{0.569}=1.766$) daha fazla olduğu söylenebilmektedir. Aynı biçimde, 54 yaş üzeri hastaların, yaşı 54 ve 54'den küçük olan hastalara göre başarısızlık riskinin 2.183 kat ($e^{0.781}=2.183$) daha fazla olduğu ifade edilebilmektedir.

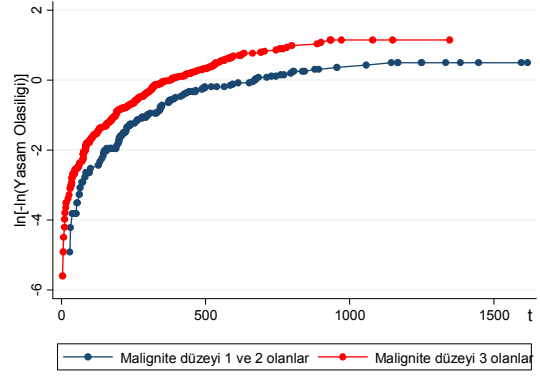
4.1.2.1. Orantılı Tehlikeler Varsayımının İncelenmesi

Ln(-ln) Yaşam Eğrileri Yöntemi

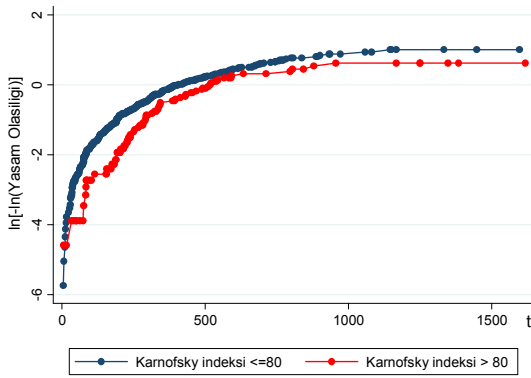
Ln(-ln) yaşam eğrileri, orantılı tehlikeler varsayımının incelenmesi için grafiksel bir yaklaşım sağlamaktadır [15]. Glioma verisi için açıklayıcı değişkenlere ait Ln(-ln) yaşam eğrileri elde edilmiş ve Şekil 4.3'te verilmiştir.



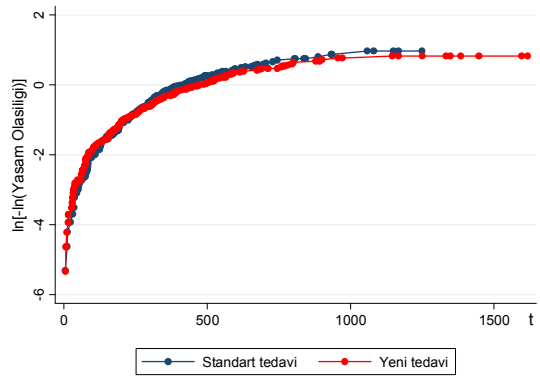
(a) Cinsiyet



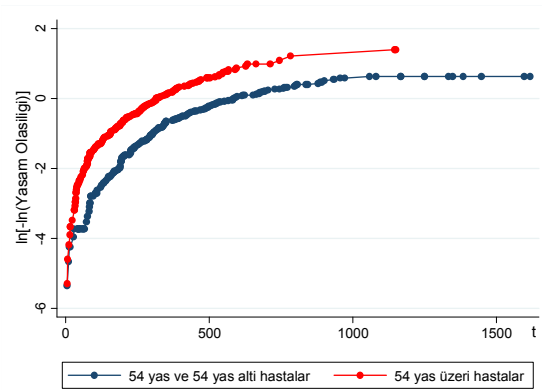
(b) Malignite Derecesi



(c) Karnofsky İndeksi



(d) Tedavi



(e) Yaş

Şekil 4.3. Glioma verisinde açıklayıcı değişkenler için $\ln(-\ln)$ yaşam eğrileri

Veri kümesi için orantılı tehlikeler varsayımı sağlanıyorsa, ln(-ln) eğrilerinin çakışmaması beklenmektedir. Şekil 4.3 incelendiğinde cinsiyet ve tedavi değişkenleri için ln(-ln) yaşam eğrilerinde çakışma olduğu görülmektedir. Buna göre ln(-ln) yaşam eğrileri yöntemine göre cinsiyet ve tedavi değişkenlerinin orantılı tehlikeler varsayımını sağlamadığı söylenebilmektedir.

Tedavi ve cinsiyet değişkenlerin orantılı tehlikeler varsayımını sağlayıp sağlamadığını kesin olarak söyleyebilmek için Schoenfeld artıkları yöntemi ve zamana bağlı açıklayıcı değişkenler yöntemi de kullanılmıştır.

Schoenfeld Artıkları Yöntemi

Glioma veri kümesindeki değişkenler için orantılı tehlikeler varsayımının sağlanıp sağlanmadığı, Schoenfeld artıkları ile bireylerin başarısızlık sürelerinin rankı arasındaki korelasyon testi ile incelenebilmektedir [73,74]. Bu test istatistiği için yokluk hipotezi; “ H_0 : Orantılı tehlikeler varsayımı sağlanmaktadır” biçimindedir. Çalışmada orantılı tehlikeler varsayımının sağlanıp sağlanmadığını incelemek için kullanılan Schoenfeld artıkları ile yaşam sürelerinin rankı arasındaki korelasyon testine ait sonuçlar Çizelge 4.4’de verilmiştir.

Çizelge 4.4. Glioma verisi için Schoenfeld artıkları ile yaşam süresinin rankı arasındaki korelasyon testi

Değişken	r*	Ki-kare	Sd	p değeri
Cinsiyet	0.026	0.20	1	0.658
Malignite Derecesi	-0.000	0.00	1	0.990
Yaş	-0.085	2.04	1	0.153
Karnofsky İndeksi	0.112	3.50	1	0.061
Tedavi	-0.023	0.14	1	0.705

* r: Schoenfeld artıkları ile logt arasındaki Pearson çarpım-moment korelasyon katsayısı

Glioma veri kümesindeki tüm açıklayıcı değişkenler için p değerleri 0.05’den büyük olarak elde edildiği için bu modelde orantılı tehlikeler varsayımının sağlandığı söylenebilmektedir.

Zamana Bağlı Açıklayıcı Değişkenler Yöntemi

Orantılı tehlikeler varsayımının incelenmesi için kullanılan bir diğer yöntem zamana bağlı açıklayıcı değişkenler yöntemidir [74]. Bu çalışmada orantılı tehlikeler varsayımını sağlamayabileceği düşünülen cinsiyet değişkeni için zamana bağlı değişkenler yöntemi ele alınmıştır. Bu amaçla ilk olarak sadece cinsiyet değişkenini ele alan bir CRM kurulmuş ve bu model için elde edilen sonuçlar Çizelge 4.5.'de verilmiştir. Bu modele ait -2LogL değeri 2800.9358 olarak elde edilmiştir. İkinci aşamada cinsiyet değişkeni ve zamana bağlı bir fonksiyon olan cinsiyetxlogt değişkeni için elde edilen CRM sonuçları Çizelge 4.6.'da verilmiştir. Bu model için de -2LogL değeri 2799.975'dir.

Çizelge 4.5. Cinsiyet değişkeni için CRM'den elde edilen sonuçlar

Değişken	$\hat{\beta}$	Std. hata	p değeri	$\exp(\hat{\beta})$	$\exp(\hat{\beta})$ için %95 güven aralığı
Cinsiyet	0.038	0.122	0.755	1.039	0.817, 1.321

Çizelge 4.6. Zamana bağlı cinsiyet değişkeni için CRM'den elde edilen sonuçlar

Değişken	$\hat{\beta}$	Std. hata	p değeri	$\exp(\hat{\beta})$	$\exp(\hat{\beta})$ için %95 güven aralığı
Cinsiyet	-0.626	0.694	0.367	0.534	0.137, 2.083
Cinsiyetxlogt	0.124	0.127	0.330	1.132	0.882, 1.453

Orantılı tehlikeler varsayımının sağlanıp sağlanmadığını incelemek amacıyla kullanılan zamana bağlı açıklayıcı değişkenler yönteminde yokluk hipotezi $H_0 : \delta = 0$ biçiminde kurulmuştur. CRM ile zamana bağlı CRM arasındaki log-olabilirlik oranı istatistiğinin farkı test istatistiğini vermektedir. Test istatistiği yokluk hipotezi altında 1 serbestlik dereceli ki-kare dağılımına sahiptir. Yokluk hipotezi kabul edilirse, cinsiyet değişkeni için orantılı tehlikeler varsayımının sağlandığı söylenebilmektedir [74]. Bu model için test istatistiği,

$$LR = -2\ln\hat{L}_{CRM} - (-2\ln\hat{L}_{Zamana\ bağılı\ CRM}) = 2800.9358 - 2799.975 = 0.9608$$

biçiminde elde edilmiştir. Elde edilen test istatistiği, $\chi^2 = 3.86$ 'dan küçük olduğu için yokluk hipotezi %95 güven düzeyinde kabul edilmiş ve cinsiyet değişkeni için orantılı tehlikeler varsayımının sağlandığı sonucuna ulaşılmıştır.

Orantılı tehlikeler varsayımını sağlamayabileceği düşünülen tedavi değişkeni için de yukarıda belirtilen yöntem kullanılmış ve modeller için elde edilen sonuçlar Çizelge 4.7 (-2logL= 2800.2638) ve Çizelge 4.8 (-2logL= 2799.1406)'de verilmiştir.

Çizelge 4.7. Glioma verisinde tedavi değişkeni için CRM sonuçları

Değişken	$\hat{\beta}$	Std. Hata	p-değeri	$\exp(\hat{\beta})$	$\exp(\hat{\beta})$ için %95 güven aralığı
Tedavi	-0.106	0.12118	0.380	0.899	-0.344, 0.131

Çizelge 4.8. Glioma verisinde zamana bağlı tedavi değişkeni için CRM sonuçları

Değişken	$\hat{\beta}$	Std. Hata	p değeri	$\exp(\hat{\beta})$	$\exp(\hat{\beta})$ için %95 güven aralığı
Tedavi	0.601	0.680	0.377	1.824	-0.732, 1.934
Tedavixlogt	-0.132	0.125	0.290	0.876	-0.377, 0.113

Olabilirlik oran test istatistiği,

$$LR = -2\ln\hat{L}_{CRM} - (-2\ln\hat{L}_{Zamana\ bağıli\ CRM}) = 2800.2638 - 2799.1406 = 1.1232$$

olarak elde edilmiştir. Buna göre $LR=1.1232 < \chi_1^2 = 3,86$ olduğu için tedavi değişkeninin orantılı tehlikeler varsayımını sağladığı %95 güven düzeyinde söylenebilmektedir.

Glioma veri kümesi için yapılan incelemeler sonucundan orantılı tehlikeler varsayımının sağlandığı sonucuna ulaşılmıştır. Buna göre CRM sonuçları kullanılarak yorum yapılabilmektedir.

4.1.3. Cox İyileşme Modeli Sonuçları

Karma iyileşme modelleri, kitlede ilgilenilen olayı yaşamayan çok sayıda birey var iken, kitledeki her bireyin başarısızlığa yatkın olduğunu varsayan yaşam modellerine göre daha doğru sonuçlar verebilmektedir. Kaplan-Meier yaşam eğrisi incelenmekte ve eğri

düzleşmiş bir kuyruğa sahip ise iyileşme modellerinin kullanılması önerilmektedir. Glioma verisi için Kaplan-Meier eğrisi düzleşen bir kuyruğa sahip olduğundan bu veri kümesi için iyileşme modellerinin değerlendirilmesi mantıklıdır. En sık kullanılan karma iyileşme modelleri Cox karma iyileşme modeli (orantılı tehlikeler karma iyileşme modeli) ve hızlandırılmış başarısızlık süresi karma iyileşme modelidir. Glioma veri kümesi için orantılı tehlikeler varsayımı sağlandığından bu çalışmada Cox iyileşme modeli kullanılmıştır.

Karma iyileşme modellerinin en büyük avantajı, iyileşmiş ve iyileşmemiş kısımlarının ayrı ayrı modellenmesi ve böylece açıklayıcı değişkenlerin iyileşmiş kısım ve iyileşmemiş kısım üzerindeki etkilerinin farklı olmasına imkân sağlamasıdır.

Karma iyileşme modellerinde iyileşme oranının modellenmesinde üç temel bağlantı fonksiyonu kullanılmaktadır, bunlar lojit bağlantı fonksiyonu, probit bağlantı fonksiyonu ve tamamlayıcı log-log bağlantı fonksiyonlarıdır. Glioma verisi için probit bağlantı fonksiyonu kullanıldığında model yakınsama sorunları nedeni ile elde edilemediği için çalışmada sonuçlara yer verilememiştir.

Glioma verisinde iyileşme oranının tahmini için lojit bağlantı fonksiyonu kullanıldığında elde edilen sonuçlar Çizelge 4.9'da, iyileşmemiş kısım için elde edilen sonuçlar ise Çizelge 4.10'da verilmiştir.

Çizelge 4.9. Glioma verisinde iyileşmiş kısım için lojit Cox iyileşme modeli sonuçları

Değişken	$\hat{\beta}$	Std hata	z değeri	p değeri
Sabit	1.938	1.221	1.587	0.112
Cinsiyet	0.273	0.816	0.335	0.738
Malignite düzeyi	2.007	2.720	0.738	0.460
Yaş	1.887	3.945	0.478	0.632
Karnofsky indeksi	-1.185	1.322	-0.897	0.370
Tedavi	-1.092	1.280	-0.853	0.394

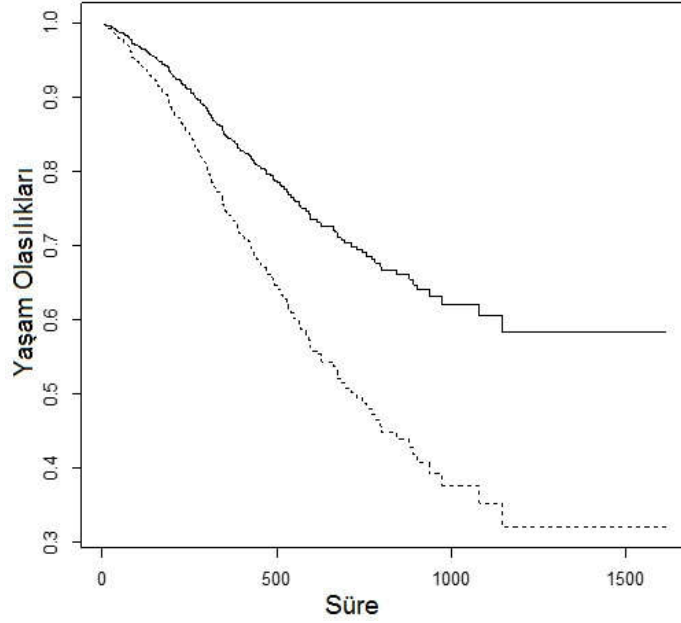
Çizelge 4.10. Glioma verisinde iyileşmemiş kısım için lojit Cox iyileşme modeli sonuçları

Değişken	$\hat{\beta}$	Std hata	z değeri	p değeri
Cinsiyet	-0.215	0.160	-1.342	0.180
Malignite düzeyi	0.379	0.157	2.413	0.016
Yaş	0.705	0.137	5.164	0.000
Karnofsky indeksi	-0.141	0.177	-0.798	0.425
Tedavi	-0.027	0.140	-0.195	0.845

İyileşmiş kısım için değişken katsayıları ve düzeyleri kullanılarak, istenilen iyileşme oranları elde edilebilmektedir. İyileşmiş kısmın oranını hesaplarken, parametre tahmininde kullanılan bağlantı fonksiyonu göz önüne alınmakta ve hesaplamalar buna göre yapılmaktadır. Lojit bağlantı fonksiyonu için, iyileşme oranı tahmin edilmek istenilen değişken düzeyleri seçildikten sonra $1 - \pi(\mathbf{z}) = e^{bz} / (1 + e^{bz})$ biçiminde hesaplanmaktadır.

Glioma veri kümesi için, malignite düzeyi 1 ya da 2 olan, karnofsky indeksi 80'den büyük, yeni tedavi uygulanmış ve yaşı 54 ya da 54'den küçük olan erkek hastaların iyileşme oranı; $1 - \pi(\mathbf{z}) = 1 - \frac{e^{(1.938+(2.207 \times 0)+(-1.185 \times 1)+(-1.092 \times 0)+(1.887 \times 0)+(0.273 \times 0))}}{1 + e^{(1.938+(2.207 \times 0)+(-1.185 \times 1)+(-1.092 \times 0)+(1.887 \times 0)+(0.273 \times 0))}}$ biçiminde hesaplanmış ve %58.4 olarak elde edilmiştir. Bu değişken düzeyleri için iyileşme oranı %58.4'dür. Aynı değişken düzeyleri alınıp sadece tedavi düzeyi değiştirilip, standart tedavi seçildiğinde iyileşme oranının $1 - e^{1.938-1.185} / (1 + e^{1.938-1.185}) = \%32$ 'ye düştüğü görülmüştür. Bu değişken düzeyleri için yaşam eğrileri Şekil 4.4'de verilmiştir.

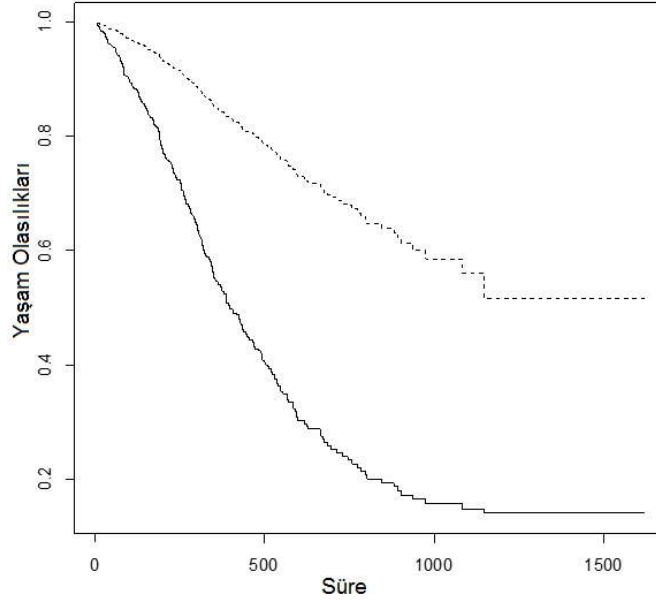
Şekil 4.4.'de, kesiksiz eğri malignite düzeyi 1 veya 2 ve karnofsky indeksi 80'den büyük olan, yaşı 54 ya da 54'den küçük, yeni tedavi türü uygulanmış erkek hastaların yaşam eğrisi verilmiştir. Kesikli eğri ise aynı değişken düzeyleri için standart tedavi uygulandığında elde edilen yaşam eğrisidir. Şekil 4.4'den bu değişken düzeyleri için sadece uygulanan tedavinin türü değiştiğinde, yaşam olasılıklarında değişme olduğu ve 'yeni tedavi' uygulandığında iyileşme oranının arttığı görülmektedir.



Şekil 4.4. Seçilen değişken düzeyleri için yaşam eğrileri 1

Aynı biçimde malignite düzeyi 1 ya da 2 olan, karnofsky indeksi 80'den büyük, yeni tedavi uygulanmış ve yaşı 54 ya da 54'den küçük olan kadın hastaları alırsak, bu grup için iyileşme oranı %51.7 bulunmuştur. Bu değişken düzeyleri için tedavi türü değiştirilip standart tedavi tercih edildiğinde ise iyileşmiş kısmın oranı sadece %26.4 olarak tahmin edilmiştir. Bu veri kümesi için, malignite düzeyi 3 olan, 54 yaş üzeri, karnofsky indeksi 80'den küçük olan ve standart tedavi uygulanan kadınlar için iyileşme oranı en düşük olarak tahmin edilmiştir ve sadece %2'dir.

Malignite düzeyi 1 veya 2 olan, karnofsky indeksi 80'den büyük, yeni tedavi uygulanmış 54 yaş üzeri kadın hastaların iyileşme oranı hesaplanmış ve %13.9 olarak elde edilmiştir. Malignite düzeyi 1 veya 2 olan, karnofsky indeksi 80'den büyük, yeni tedavi uygulanmış 54 yaş ve 54 yaş altı kadın hastalar için ise bu oran %51.6 olarak hesaplanmıştır. Bu değişken düzeyleri için yaşam eğrileri Şekil 4.5'de verilmiştir.



Şekil 4.5. Seçilen değişken düzeyleri için yaşam eğrileri 2

Şekil 4.5’de kesikli eğri malignite düzeyi 1 veya 2 olan, karnofsky indeksi 80’den büyük, yeni tedavi uygulanmış 54 yaş ve 54 yaş altı kadın hastalar yaşam eğrisi, sürekli eğri ise aynı değişken düzeyleri için yaş 54 üzeri alındığında elde edilen yaşam olasılıklarını vermektedir. Bu grafikten de yaşın küçük olmasının, iyileşme oranını arttıran bir faktör olduğu görülmektedir.

Çizelge 4.10 kitledeki iyileşmemiş hastalar için değişken düzeylerinin tehlike oranlarının yorumlanmasını sağlamaktadır. Bu model için malignite düzeyi ($p=0.016$) ve yaş değişkeni ($p=0.000$) istatistiksel olarak anlamlı bulunmuştur. Çizelge 4.10’dan elde edilen sonuçlara dayanarak, malignite düzeyi 3 olan hastaların riskinin, malignite düzeyi 1 ve 2 olan hastalara göre 1.46 kat ($e^{0.379}=1.46$) daha fazla olduğu söylenebilmektedir. Yaşı 54 üzeri olan hastaların, yaşı 54 ve 54’den küçük hastalara göre 2.02 kat ($e^{0.705}=2.02$) daha fazla risk altında olduğu ifade edilebilmektedir.

Glioma verisi için Cox iyileşme modelinin elde edilmesinde tamamlayıcı log-log (cloglog) bağlantı fonksiyonu kullanıldığında elde edilen sonuçlar Çizelge 4.11 ve Çizelge 4.12’de verilmiştir.

Çizelge 4.11. Glioma verisinde iyileşmiş kısım için cloglog Cox iyileşme modeli sonuçları

Değişken	$\hat{\beta}$	Std hata	z değeri	p değeri
Sabit	0.159	0.139	1.139	0.255
Cinsiyet	-0.131	0.125	-1.052	0.293
Malignite düzeyi	0.425	0.134	3.165	0.001
Yaş	0.360	0.124	2.891	0.004
Karnofsky indeksi	-0.195	0.137	-1.422	0.155
Tedavi	-0.012	0.123	-0.096	0.923

Çizelge 4.12. Glioma verisinde iyileşmemiş kısım için cloglog Cox iyileşme modeli sonuçları

Değişken	$\hat{\beta}$	Std hata	z değeri	p değeri
Cinsiyet	-0.062	0.134	-0.463	0.643
Malignite düzeyi	0.411	0.152	2.705	0.006
Yaş	0.651	0.133	4.913	0.000
Karnofsky indeksi	-0.206	0.147	-1.407	0.159
Tedavi	-0.136	0.136	-0.996	0.319

Tamamlayıcı log-log bağlantı fonksiyonu kullanıldığında, hem iyileşmiş hastaların iyileşme oranı için, hem de iyileşmemiş hastaların yaşam fonksiyonu için malignite düzeyi ve yaş değişkenlerinin %95 güven düzeyinde anlamlı olduğu ($p < 0.05$) görülmektedir.

İstenilen değişken düzeyleri için iyileşme oranı $1 - \pi$, tamamlayıcı log-log bağlantı fonksiyonunu kullanılarak ($\log(-\log(1 - \pi)) = \mathbf{bz}$) elde edilmiştir. Buna göre, malignite düzeyi 1 ya da 2 olan, karnofsky indeksi 80'den büyük, yeni tedavi türü uygulanmış ve yaşı 54 ya da 54'den küçük olan erkek hastaların iyileşme oranı %38.6 olarak bulunmuştur. Aynı değişken düzeyleri alınıp sadece uygulanan tedavi türü standart tedavi olarak değiştirildiğinde iyileşme oranı %38.1 olmuştur. Bu model için, tedavi türündeki değişikliğin, iyileşmiş kısmın oranını etkileyen önemli bir değişken olmadığı görülmektedir. Yaş değişkeninin iyileşmiş kısma olan etkisini incelemek için, malignite düzeyi 1 ya da 2 olan, karnofsky indeksi 80'den büyük, yeni tedavi türü uygulanmış erkek hastalardan, yaşı 54 üzeri olan hastalar için iyileşme oranı hesaplanmış ve %25.5 olarak elde edilmiştir. Bu modelde, düşük yaş grubundaki hastaların iyileşme oranının,

yüksek yaş grubundaki hastaların iyileşme oranından daha yüksek olduğu ve yaş değişkeninin istatistiksel olarak anlamlı olduğu görülmektedir. Yüksek yaş, bu model için iyileşme oranını azaltan bir faktördür. Malignite düzeyinin iyileşmiş kısımdaki etkisini incelemek için, Karnofsky indeksi 80'den büyük, yeni tedavi türü uygulanmış ve yaşı 54 ya da 54'den küçük olan erkek hastalardan malignite düzeyi 3 olanlar alınmıştır. Bu grup için iyileşme oranı %23.3 olarak elde edilmiştir. Aynı değişken düzeylerinde, sadece malignite 1 veya 2 değerini aldığında bu oran %38.6 olarak elde edilmişti. Buradan da, bu grup için malignite düzeyi 3 olanların iyileşme oranının, malignite düzeyi 1 ve 2 olanlardan daha düşük olduğu görülebilmektedir ve malignite düzeyi bu model için istatistiksel olarak anlamlı ($p=0.001$) bir değişkendir.

Glioma verisinde tamamlayıcı log-log Cox karma iyileşme modeli sonuçlarından, iyileşmemiş kısım için malignite düzeyi ($p=0.06$) ve yaş ($p=0.000$) değişkenlerinin istatistiksel olarak anlamlı olduğu %95 güven düzeyinde söylenebilmektedir. Malignite düzeyi 3 olan hastalar, malignite düzeyi 1 ve 2 olan hastalara göre 1.51 kat ($e^{0.411}=1.51$) daha fazla risk altında bulunduğu ifade edilebilir. Yaşı 54'ün üzerinde olan hastaların başarısızlık riskinin, yaşı 54 ve 54'ün altında olan hastaların başarısızlık riskine göre 1.92 kat ($e^{0.651}=1.92$) daha fazla olduğu söylenebilmektedir.

İstenilen tüm gruplar için iyileşme oranı, parametre tahminleri ve değişken düzeyleri kullanılarak bu biçimde yorumlanabilmektedir. Daha önce de belirttiğimiz gibi, karma iyileşme modellerinin, açıklayıcı değişken etkilerinin iyileşmiş kısım ve iyileşmemiş kısımda farklı olmasına olanak sağladığı elde edilen sonuçlarda da görülmektedir.

Glioma verisi için lojit Cox karma iyileşme modeli ve cloglog Cox karma iyileşme modeli sonuçlarına bakıldığında, cloglog bağlantı fonksiyonu kullanıldığında elde edilen modelin p değerlerinin ve standart hatalarının çok daha düşük olduğu görülmüştür ve bu nedenle cloglog Cox karma iyileşme modelinin kullanılması daha uygun olabilir.

4.2. Mide Kanseri Veri Kümesi

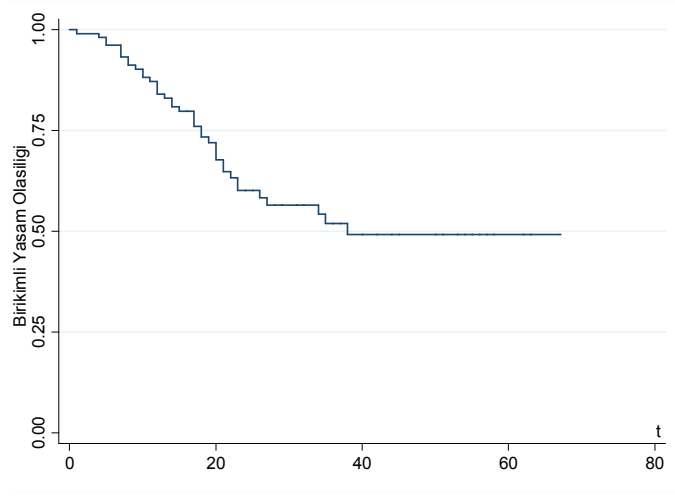
Bu çalışmada, Hacettepe Üniversitesi Biyoistatistik Ana Bilim Dalı'ndan elde edilen mide kanseri hastalarına ait veriler kullanılmıştır [75]. Bu çalışmada, klasik CRM, HBS modelleri ve iyileşme modellerinden elde edilen sonuçları karşılaştırmak ve bu veri kümesi için en uygun modeli bulmak amacıyla 106 mide kanseri hastasına ait bilgiler ele alınmıştır. Hastaların ölümü başarısızlık olarak ifade edilmiştir. Çalışma süresinin sonunda başarısızlık ile karşılaşmayan hastalar ise durdurulmuş olarak tanımlanmıştır. Çalışmada hastaların takip süresi sona erdiğinde, 106 hastadan 67'sinde (%63.2) başarısızlık ve 39'unda (%36.8) durdurma gözlenmiştir. Bu çalışmada kullanılan değişkenler ve değişkenlerin düzeyleri Çizelge 4.13'de verilmiştir. Takip süresi 67 aydır. Ortanca yaşam süresi 38 ay ve 5 yıllık yaşam olasılığı 0.492 ± 0.064 olarak elde edilmiştir.

Çizelge 4.13. Mide kanseri çalışmasında kullanılan değişkenler ve düzeyleri

Değişken	Değişken Düzeyleri	n (%)	Başarısız Gözlemlerin Sayısı	Durdurulmuş Gözlemlerin Sayısı
Metastaz	0: Yok	74 (%69.8)	16	58
	1: Var	32 (%30.2)	23	9
Alkol	0: Yok	94 (%88.7)	35	59
	1: Var	12 (%11.3)	4	8
Aile öyküsü	0: Yok	79 (%74.5)	27	52
	1: Var	27 (%25.5)	12	15
Radyoterapi	0: Yok	34 (%32.1)	14	20
	1: Var	72 (%67.9)	25	47
Yaş	0: ≤ 56	54 (%50.9)	21	33
	1: > 56	52 (%49.1)	18	34

4.2.1 Parametrik Olmayan Yaşam Çözümlemesi Sonuçları

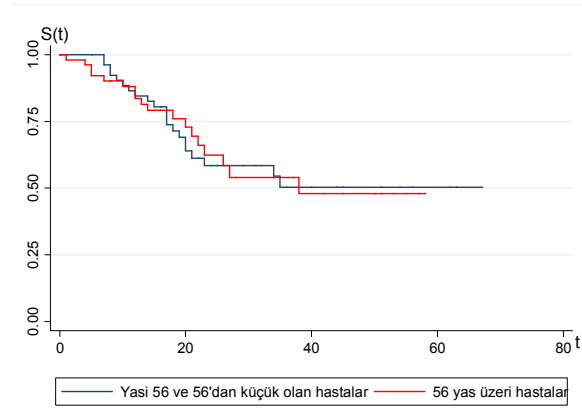
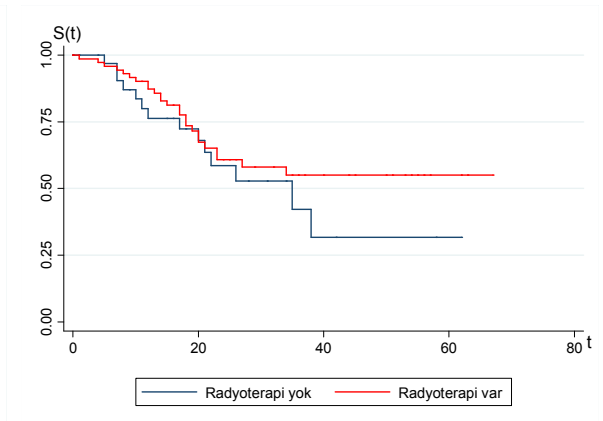
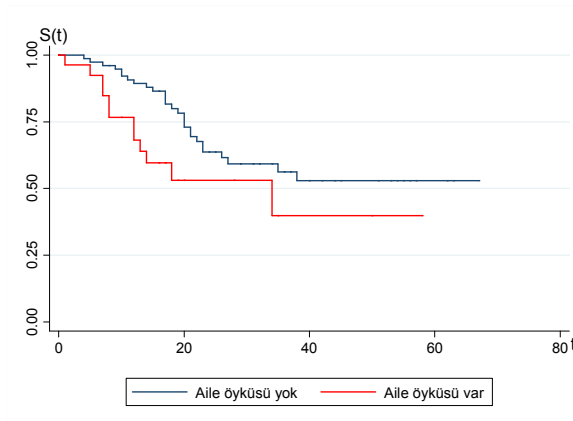
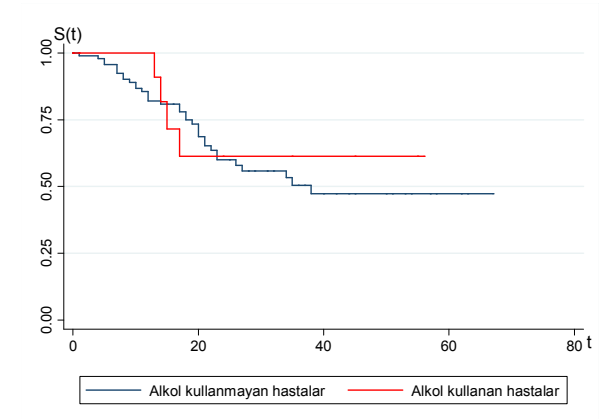
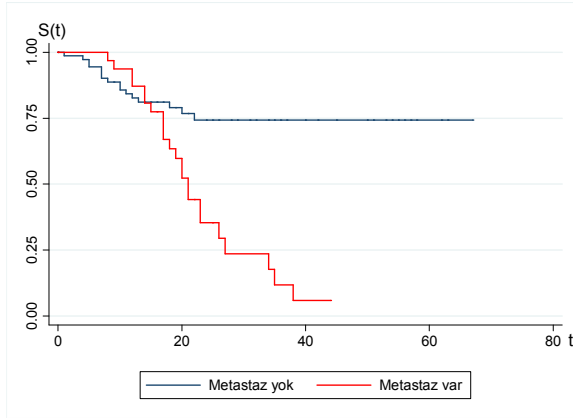
Mide kanseri verileri için Kaplan-Meier yaşam eğrisi Şekil 4.6'da verilmiştir. Bu eğri incelendiğinde, eğrinin yaşam olasılığı 0.50 civarında düzleştiği ve düz bir kuyruğa sahip olduğu görülmüştür. Buna göre Kaplan-Meier yaşam eğrisi düz kuyruğa sahip olduğundan mide kanseri verisinin çözümlenmesinde, iyileşme modellerinin kullanılması önerilmektedir.



Şekil 4.6. Mide kanseri verisi için Kaplan-Meier yaşam eğrisi

Her bir açıklayıcı değişken için Kaplan-Meier yaşam eğrileri de elde edilmiş ve Şekil 4.7'de verilmiştir. Bu değişkenlerin düzeylerine göre yaşam olasılıklarındaki farklılıklar grafiksel olarak görülmektedir.

Elde edilen Kaplan-Meier yaşam grafiklerinden, açıklayıcı değişkenlerin düzeyleri yaşam olasılıkları bakımından karşılaştırılmıştır. Mide kanseri veri kümesi için, metastaz yaşamamış hastaların yaşam olasılıklarının, metastaz yaşamış hastalardan daha yüksek olduğu görülmektedir. Bu veri kümesi için, yaşam olasılıkları bakımından alkol kullanan hastalar ve alkol kullanmayan hastalar arasında anlamlı bir farklılık görülmemiştir. Mide kanseri verisi için, aile öyküsü bulunan hastaların yaşam olasılıklarının, aile öyküsü bulunmayan hastaların yaşam olasılıklarından yüksek olduğu görülmektedir. Aynı biçimde, radyoterapi almış hastaların, radyoterapi almamış hastalara göre daha yüksek yaşam olasılıklarına sahip oldukları ifade edilebilmektedir.



Şekil 4.7. Mide kanseri çalışmasındaki açıklayıcı değişkenler için Kaplan-Meier yaşam eğrileri

Yaşam olasılıkları arasında istatistiksel olarak anlamlı bir farklılık olup olmadığını inceleyebilmek için de log-rank testi uygulanmış ve Çizelge 4.14'deki sonuçlar elde edilmiştir.

Çizelge 4.14. Mide kanseri verisi için log-rank testi sonuçları

Değişken	Ki-kare	p değeri
Metastaz	16.37	0.000
Alkol Kullanımı	0.09	0.761
Aile Öyküsü	4.58	0.032
Radyoterapi	0.90	0.343
Yaş	0.00	0.994

Log-rank testi sonuçlarına göre metastaz ($p=0.0001$) ve aile öyküsü ($p=0.0324$) değişkenlerinin düzeyleri arasında yaşam olasılıkları bakımından, istatistiksel olarak anlamlı bir farklılık olduğu söylenebilmektedir.

4.2.2. Cox Regresyon Modeli Sonuçları

Cox regresyon çözümlemesi için elde edilen sonuçlar Çizelge 4.15'de verilmiştir. Bu modelin anlamlılığı test etmek için elde edilen p değeri 0.0013 olduğundan mide kanseri verisi için CRM'nin kullanılabileceği söylenebilir. Modelin AIC değeri 312.892, BIC değeri 326.209 ve $-2\text{LogL} = 302.892$ olarak elde edilmiştir.

Çizelge 4.15. Mide kanseri verisi için CRM sonuçları

Değişken	$\hat{\beta}$	Std. hata	p değeri	$\exp(\hat{\beta})$	$\exp(\hat{\beta})$ için %95 güven aralığı
Yaş	-0.150	0.342	0.661	0.861	0.440, 1.682
Metastaz	1.280	0.335	0.000	3.595	1.864, 6.934
Alkol	-0.034	0.551	0.951	0.966	0.328, 2.847
Aile Öyküsü	0.848	0.357	0.018	2.335	1.160, 4.701
Radyoterapi	-0.229	0.354	0.518	0.795	0.397, 1.592

Çizelge 4.15'deki p değerleri incelendiğinde, metastaz ve aile öyküsü değişkenlerinin, mide kanserinden ölümü etkileyen önemli değişkenler olduğu %95 güven düzeyinde söylenebilmektedir. Çizelge 4.15'e göre, metastaz yaşamış hastaların, metastaz yaşamamış hastalara göre ölüm riski 3.595 kat ($e^{1.280}=3.595$) daha fazladır. Aile öyküsü

bulunan hastaların, bulunmayan hastalara göre ölüm riskinin 2.335kat ($e^{0.848}=2.335$) kat daha fazla olduğu %95 güven düzeyinde söylenebilmektedir.

4.2.2.1. Orantılı Tehlikeler Varsayımının İncelenmesi

Çalışmada orantılı tehlikeler varsayımı Schoenfeld artıkları yöntemi kullanılarak incelenmiştir. Elde edilen sonuçlar Çizelge 4.16'da verilmiştir.

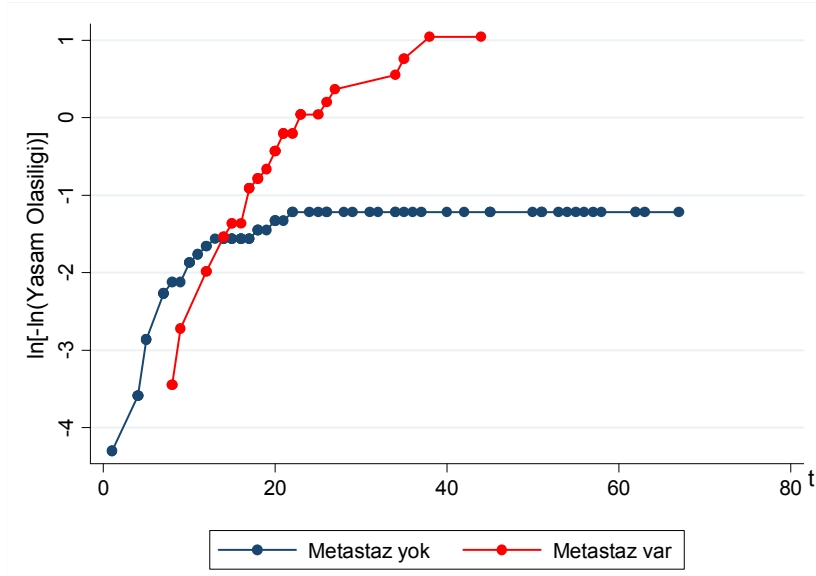
Çizelge 4.16. Mide kanseri verisi için Schoenfeld artıkları ile yaşam süresinin rankı arasındaki korelasyon testi

Değişken	r*	Ki-kare	sd	p değeri
Yaş	0.023	0.02	1	0.879
Metastaz	0.650	20.31	1	0.000
Alkol	0.168	1.18	1	0.277
Aile Öyküsü	-0.252	2.20	1	0.138
Radyoterapi	0.013	0.01	1	0.931

* r: Schoenfeld artıkları ile logt arasındaki Pearson çarpım-moment korelasyon katsayısı

Bu yöntem için yokluk hipotezi “orantılı tehlikeler varsayımı sağlanmaktadır” biçimindedir. Metastaz değişkeni için orantılı tehlikeler varsayımının sağlanmadığı görülmektedir ($p= 0.0000$). Diğer değişkenler için orantılı tehlikeler varsayımının sağlandığı %95 güven düzeyinde görülmüştür.

Metastaz değişkeninin orantılı tehlikeler varsayımını sağlayıp sağlamadığı, ln(-ln) yaşam eğrileri yöntemi kullanılarak da test edilmiştir. ln(-ln) yaşam eğrisi Şekil 4.8'de verilmiştir. Metastaz görülmeyen ve görülen hastalar için yaşam eğrilerinde çakışma olduğundan metastaz değişkeninin orantılı tehlikeler varsayımını sağlamadığı söylenebilir.



Şekil 4.8. Mide kanseri verisinde metastaz değişkeni için ln(-ln) grafiği

Metastaz değişkeninin orantılı tehlikeler varsayımını sağlayıp sağlamadığı, zamana bağlı açıklayıcı değişkenler yöntemi kullanılarak da test edilmiştir. Bu yöntem için elde edilen modellere ait sonuçlar Çizelge 4.17 (-2LogL = 308.308) ve Çizelge 4.18 (-2LogL = 284.283)'de verilmiştir. Bu model için p değeri 0.001 olarak elde edilmiştir.

Çizelge 4.17. Metastaz değişkeni için CRM sonuçları

Değişken	$\hat{\beta}$	Std. hata	p değeri	$\exp(\hat{\beta})$	$\exp(\hat{\beta})$ için %95 güven aralığı
Metastaz	1.252	0.331	0.000	3.497	1.827, 6.692

Çizelge 4.18 Zamana bağlı metastaz değişkeni için CRM sonuçları

Değişken	$\hat{\beta}$	Std. hata	p değeri	$\exp(\hat{\beta})$	$\exp(\hat{\beta})$ için %95 güven aralığı
Metastaz	-8.217	2.754	0.003	0.0002	0.000, 0.059
Metastazxlogt	3.617	1.036	0.000	37.217	4.880, 283.814

Olabilirlik oran test istatistiği,

$$LR = -2\ln\hat{L}_{CRM} - (-2\ln\hat{L}_{ZamanabađlıCRM}) = 308.308 - 284.283 = 24.025 \text{ olarak elde edilmiştir.}$$

Buna göre $LR = 24.205 > \chi^2 = 3.86$ olduđu için metastaz deđişkeninin orantılı tehlikeler varsayımını sağlamadıđı sonucuna ulaşılmıştır.

Elde edilen bu sonuçlara göre, modelin orantılı tehlikeler varsayımını sağlamadıđı sonucuna ulaşılmıştır. Bu nedenle klasik CRM yerine tabakalı CRM, zamana bađlı açıklayıcı deđişkenli CRM, HBS, zayıflık modeli ya da iyileşme modellerinin bu veri kümesi için daha uygun sonuçlar verebileceđi düşünölmüş, bu modeller ele alınmıştır.

4.2.3. Tabakalı Cox Regresyon Modeli Sonuçları

Orantılı tehlikeler varsayımının sağlanmadıđı durumlarda, klasik CRM yerine orantılı tehlikeler varsayımını sağlamayan deđişken tabaka deđişkeni olarak alınıp etkileşimli ve etkileşimsiz tabakalandırılmış CRM veri kümesine uygulanabilir.

Etkileşimsiz Model Sonuçları

Modelde orantılı tehlikeler varsayımının bozulmasına neden olan metastaz deđişkeni tabaka deđişkeni olarak ele alınarak etkileşimsiz model için CRM modeli sonuçları elde edilmiştir. Sonuçlar Çizelge 4.19'da verilmiştir. Etkileşimsiz model için p deđeri 0.322, AIC deđeri 253.190, BIC deđeri 263.844, -2LogL deđeri ise 245.190 olarak elde edilmiştir.

Çizelge 4.19. Etkileşimsiz model için CRM sonuçları

Deđişken	$\hat{\beta}$	Std. hata	p deđeri	$\exp(\hat{\beta})$	$\exp(\hat{\beta})$ için %95 güven aralığı
Alkol	0.200	0.561	0.721	1.222	0.407, 3.670
Aile Öyküsü	0.787	0.364	0.031	2.197	1.076, 4.488
Radyoterapi	-1.119	0.362	0.742	0.887	0.436, 1.804
Yaş	-0.073	0.344	0.832	0.930	0.474, 1.824

Çizelge 4.19'daki p deđerleri incelendiđinde aile öyküsü deđişkeninin başarısızlığı etkileyen önemli bir risk faktörü olduđu %95 güven düzeyinde söylenebilmektedir. Aile

öyküsü bulunan mide kanseri hastalarının, aile öyküsü bulunmayan hastalara göre 2.197 kat daha fazla risk altında olduğu görülmektedir.

Etkileşimli Model Sonuçları

Modelde orantılı tehlikeler varsayımının bozulmasına neden olan metastaz değişkenine göre metastazı olan ve metastazı olmayan olmak üzere veri kümesi iki tabakaya ayrılmıştır. Buna göre metastazı olmayan hastalar için CRM yapılmış ve sonuçlar Çizelge 4.20’de verilmiştir. Metastazı olmayan hastalar için p değeri 0.222, AIC değeri 131.522, BIC değeri 140.739, -2LogL değeri ise 123.522 olarak elde edilmiştir.

Çizelge 4.20. Metastazı olmayan hastalar için etkileşimli CRM sonuçları

Değişken	$\hat{\beta}$	Std. hata	p değeri	$\exp(\hat{\beta})$	$\exp(\hat{\beta})$ için %95 güven aralığı
Alkol	-0.910	1.073	0.397	0.403	0.049, 3.299
Aile Öyküsü	1.036	0.527	0.049	2.819	1.003, 7.923
Radyoterapi	-0.490	0.536	0.360	0.613	0.214, 1.750
Yaş	-0.114	0.521	0.828	0.893	0.321, 2.479

Çizelge 4.20’den elde edilen p değerleri incelendiğinde, aile öyküsü değişkeninin, başarısızlığı etkileyen önemli bir risk faktörü olduğu %95 güven düzeyinde söylenebilmektedir.

Çizelge 4.21. Metastazı olan hastalar için etkileşimli CRM sonuçları

Değişken	$\hat{\beta}$	Std. hata	p değeri	$\exp(\hat{\beta})$	$\exp(\hat{\beta})$ için %95 güven aralığı
Alkol	1.746	0.777	0.025	5.732	1.249, 26.298
Aile Öyküsü	0.469	0.535	0.381	1.598	0.559, 4.564
Radyoterapi	0.530	0.509	0.298	1.699	0.626, 4.610
Yaş	0.292	0.493	0.553	1.340	0.509, 3.523

Metastazı olan hastalar için etkileşimli CRM için p değeri 0.188, AIC değeri 122.491, BIC değeri 128.354, -2LogL değeri ise 114.491 olarak elde edilmiştir. Metastazı olan hastalar için CRM sonuçları Çizelge 4.21’de verilmiştir.

Çizelge 4.21'deki p değerleri incelendiğinde bu model için, alkol değişkeninin (p=0.025) başarısızlığı etkileyen önemli bir risk faktörü olduğu %95 güven düzeyinde söylenebilmektedir.

Etkileşimsizlik Varsayımının İncelenmesi

Etkileşimsizlik varsayımını incelenmek için yokluk hipotezi “H₀: Etkileşim yoktur” biçiminde kurulmaktadır. Etkileşimsiz model ile etkileşimli model arasındaki log-olabilirlik istatistiğinin farkı test istatistiğini vermektedir. Test istatistiği yokluk hipotezi altında 4 serbestlik dereceli (etkileşimli modelde 4 tane çarpım terimi olduğundan) ki-kare dağılımına sahiptir. Yokluk hipotezinin kabul edilmesi durumunda, etkileşimin olmadığını söyleyebiliriz. R indirgenmiş (etkileşimsiz) modeli ve F tam (etkileşimli) modeli göstermek üzere olabilirlik oranı (LR) test istatistiği,

$$\begin{aligned} LR &= -2\ln\hat{L}_R - (-2\ln\hat{L}_F) \\ &= 245.190 - 238.013 = 7.177 \end{aligned}$$

$LR = 7.177 < \chi^2_{4;0.04} = 9.49$ olduğu için yokluk hipotezi kabul edilir. Etkileşimin olmadığı %95 güven düzeyinde söylenebilmektedir. Bu sonuca göre veri kümemiz için orantısızlığa neden olan değişkene göre tabakalandırma yaparken etkileşimsiz modelin etkileşimli modele tercih edilebileceğini söyleyebiliriz. Ancak tabakalandırılmış CRM modellerinin tamamında modelin anlamlılığına bakıldığında modellerin anlamsız olduğu görülmüştür. Bu nedenle de mide kanseri veri kümesi için tabakalandırılmış CRM sonuçlarının kullanılması doğru değildir.

4.2.4. Genişletilmiş Cox Regresyon Modeli Sonuçları

Veri kümesi için orantılı tehlikeler varsayımı sağlanmadığında, genişletilmiş CRM kullanılabilir. Bu modelde klasik CRM, zamana bağlı açıklayıcı değişkenlerin olduğu bir modele genişlemektedir. Çalışmamızda zamana bağlı fonksiyon olarak $g(t)=\text{Int}$ fonksiyonu kullanılmıştır. Genişletilmiş CRM için AIC değeri 291.604, BIC değeri 307.584, -2LogL ise 279.604 olarak elde edilmiştir ve modelin anlamlılığı için elde edilen p değeri 0.000'dır. Sonuçlar Çizelge 4.22'de verilmiştir.

Çizelge 4.22. Mide verisi için genişletilmiş CRM sonuçları

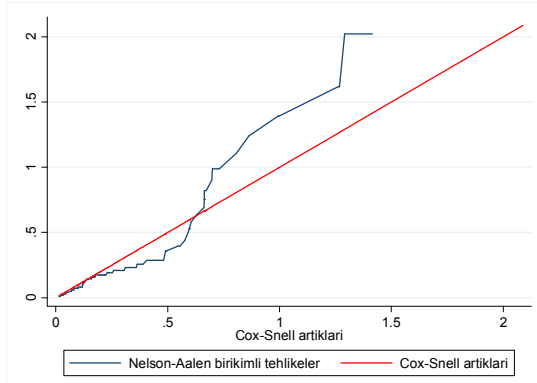
Değişken	$\hat{\beta}$	Std. hata	p değeri	$\exp(\hat{\beta})$	$\exp(\hat{\beta})$ için %95 güven aralığı
Alkol	0.223	0.562	0.692	1.249	0.415, 3.762
Aile Öyküsü	0.778	0.364	0.033	2.176	1.067, 4.441
Radyoterapi	-0.136	0.362	0.707	0.873	0.429, 1.775
Yaş	-0.096	0.343	0.780	0.908	0.464, 1.780
Metastaz	-8.093	2.756	0.003	0.0003	0.000, 0.068
Metastazxg(t)	3.591	1.039	0.001	36.269	4.732, 277.980

Çizelge 4.22'deki p değerleri incelendiğinde aile öyküsü ve metastaz değişkenlerinin başarısızlığı etkileyen önemli risk faktörleri olduğu %95 güven düzeyinde söylenebilmektedir. Aile öyküsü olan hastalar, aile öyküsü olmayan hastalara göre 2.176 kat daha fazla risk altındadır. Zamana bağlı metastaz değişkeni, herhangi bir t zamanı için yorumlanabilmektedir. Örneğin, 12 aydır çalışmada olan metastaz yaşamış hastaların, metastaz yaşamamış hastalara göre başarısızlık riskinin $1/\exp(-8.093 + (3.591 \times \log(12))) = 67.9$ olarak elde edilmiştir. 24 aydır çalışmada olan metastaz yaşamış hastaların, metastaz yaşamamış hastalardan $1/\exp(-8.093 + (3.591 \times \log(24))) = 23$ kat daha fazla risk altında olduğu söylenebilmektedir. Benzer yorumlar her bir yaşam süresi için benzer biçimde yapılabilir.

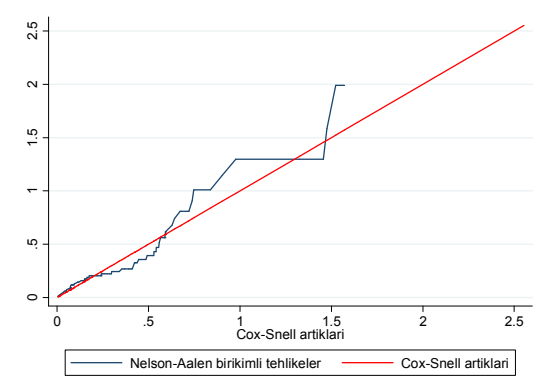
4.2.5. Hızlandırılmış Başarısızlık Süresi Modelleri Sonuçları

HBS modelleri parametrik yaşam modelleridir. Parametrik yöntemler, yaşam süresi bilinen bir dağılıma uygunluk gösteriyor ise, parametrik olmayan ya da yarı parametrik yöntemlere göre daha iyi sonuçlar vermektedir.

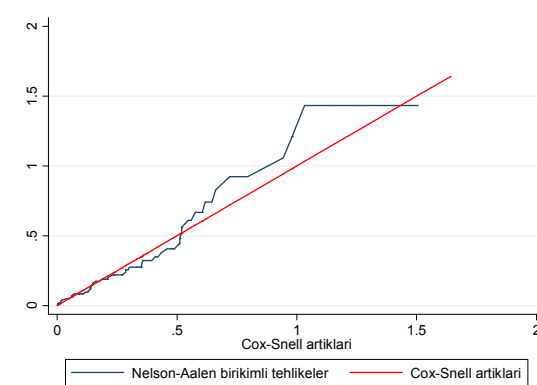
Cox-Snell artıkları, parametrik regresyon modellerinde uyum iyiliğini incelemek için kullanılmaktadır. Mide kanseri verisi için Üstel, Weibull, Log-normal, Log-lojistik parametrik regresyon modellerine ait Cox-Snell grafikleri Şekil 4.9'da verilmiştir. Gamma regresyon modeli, yakınsama sağlanamadığı için elde edilememiştir.



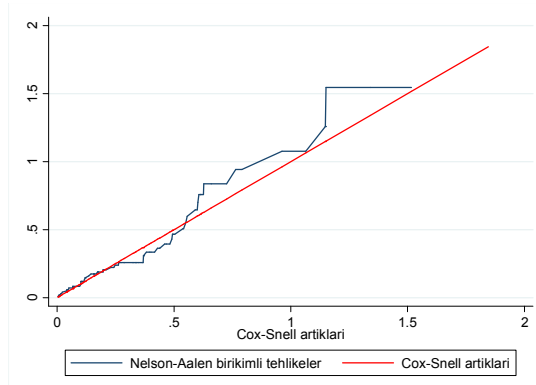
(a) Üstel



(b) Weibull



(c) Log-normal



(d) Log-lojistik

Şekil 4.9. Mide kanseri verisi için Cox-Snell artıkları

Elde edilen Cox-Snell artıkları grafiklerine göre, üstel HBS modeli dışında diğer modellerin veri kümesi için uygun olduğu söylenebilir ancak en uygun parametrik regresyon modelini tespit edebilmek için model karşılaştırma kriterlerinden yararlanılmıştır.

Mide kanseri veri kümesi için HBS modelleri kapsamında üstel ($p=0.0003$), Weibull ($p=0.0001$), log-lojistik ($p=0.0016$), log-normal ($p=0.0073$), Gamma regresyon modelleri ile çözümlenmiştir. Gamma HBS modeli için yakınsama sağlanamadığından, bu modele ait sonuçlar elde edilememiştir. Gamma HBS modeli hariç tüm modeller istatistiksel olarak anlamlı ($p < 0.05$) bulunmuştur. Model seçimi yapabilmek için incelenen HBS modellerine ait $-2\log L$, AIC ve BIC değerleri Çizelge 4.23'de verilmiştir.

Çizelge 4.23. HBS modelleri için model karşılaştırma kriterleri

HBS Modeli	-2LogL	AIC	BIC
Üstel	177.648	189.648	205.628
Weibull	172.648	186.648	205.292
Log-normal	178.141	192.141	210.785
Log-lojistik	175.755	189.755	208.399

Çizelge 4.23'de yer alan sonuçlara göre veriye en uygun HBS modelinin, en düşük AIC ve BIC değerlerine sahip olan Weibull regresyon modeli olduğu görülmektedir. Mide kanseri veri kümesi için Weibull HBS modeli sonuçları Çizelge 4.24'de verilmiştir.

Çizelge 4.24. Mide kanseri verisi için Weibull HBS modeli sonuçları

Değişken	$\hat{\beta}$	Std. hata	$\hat{\beta}$ için %95 güven aralığı	Z değeri	p değeri
Sabit	4.406	0.322	3.775, 5.036	13.69	0.000
Yaş	0.594	0.248	-0.426, 0.545	0.24	0.811
Metastaz	-1.133	0.256	-1.634, -0.631	-4.42	0.000
Alkol	0.003	0.395	-0.771, 0.777	0.01	0.994
Aile Öyküsü	-0.609	0.261	-1.120, -0.097	-2.33	0.020
Radyoterapi	0.183	0.256	-0.320, 0.686	0.71	0.475
Ölçek	0.724	0.097	0.557, 0.942		

Çizelge 4.24 incelendiğinde, metastaz ($p=0.000$) ve aile öyküsü ($p=0.020$) değişkenlerinin, başarısızlık süresini etkileyen önemli faktörler olduğu %95 güven düzeyinde söylenebilmektedir. Metastaz yaşamamış hastaların ortalama yaşam süresi, metastaz yaşamış olan hastaların yaşam süresine göre yaklaşık $1/\exp(-1.133)=3.1$ kat daha uzundur. Aile öyküsü bulunmayan hastaların ortalama yaşam süresi, aile öyküsü bulunan hastalara göre $1/\exp(-0.609)=1.84$ kat daha uzundur.

4.2.6. Hızlandırılmış Başarısızlık Süresi Karma İyileşme Modeli Sonuçları

Mide kanseri verisi için elde edilen Kaplan-Meier yaşam eğrisi düzleşen bir kuyruğa sahip olduğundan karma iyileşme modellerinin kullanımı bu veri kümesi için faydalı olabilir. Mide kanseri verisi için orantılı tehlikeler varsayımı test edilmiş, orantılı tehlikeler varsayımının sağlanmadığı görülmüştür. Bu varsayımın sağlanmaması

nedeniyle, HBS karma iyileşme modelinin kullanımının uygun olduğu sonucuna ulaşılmış ve R programı kullanılarak bu model için sonuçlar elde edilmiştir.

HBS karma iyileşme modeli için lojit ve probit bağlantı fonksiyonları modeldeki yakınsama sorunları nedeni ile hata vermiştir. cloglog bağlantı fonksiyonu kullanıldığında elde edilen HBS karma iyileşme modeli sonuçları Çizelge 4.25 ve Çizelge 4.26'da verilmiştir.

Çizelge 4.25. Mide kanseri verisinde iyileşmiş kısım için cloglog HBS iyileşme modeli sonuçları

Değişken	$\hat{\beta}$	Std hata	z değeri	p değeri
Sabit	-1.066	0.429	-2.488	0.013
Metastaz	2.001	0.458	4.365	0.000
Alkol	-0.146	2.841	-0.051	0.959
Aile Öyküsü	0.600	0.484	1.241	0.215
Radyoterapi	-0.408	0.446	-0.914	0.361
Yaş	-0.221	0.395	-0.559	0.576

Çizelge 4.26. Mide kanseri verisinde iyileşmemiş kısım için cloglog HBS iyileşme modeli sonuçları

Değişken	$\hat{\beta}$	Std hata	z değeri	p değeri
Sabit	-1.141	0.367	-0.384	0.701
Metastaz	0.757	0.271	2.792	0.005
Alkol	-0.118	0.365	-0.323	0.746
Aile Öyküsü	-0.303	0.253	-1.199	0.231
Radyoterapi	-0.212	0.194	-1.091	0.275
Yaş	-0.113	0.207	-0.545	0.586

Karma HBS modeli sonuçları kullanılarak, alkol kullanmayan, aile öyküsü bulunmayan, radyoterapi uygulanmış, yaşı 56 ve 56'dan küçük ve metastazı bulunmayan hastalar için iyileşme oranı %79.5 olarak elde edilmiştir. Aynı değişken düzeylerine sahip ancak metastazı bulunan hastalar için bu iyileşme oranı sadece %18.4 olarak hesaplanmıştır. Buradan, metastaz değişkeninin ($p=0.000$) iyileşme oranı üzerinde etkili bir değişken olduğu ve metastazı bulunan hasta grubu için iyileşme oranının çok daha düşük olduğu görülmektedir. Aynı değişken düzeyleri (alkol kullanmayan, aile öyküsü bulunmayan, yaşı 56 ve 56'dan küçük olan ve metastazı bulunmayan hastalar) alınıp bu defa

radoterapi uygulanmayanlar seçildiğinde iyileşme oranı %70 bulunmuştur. Bu sonuç, iyileşme oranı %79.5 olan radyoterapi almış olan grupla karşılaştırıldığında, radyoterapi uygulanmasının iyileşmiş kısmın oranını arttırdığını düşündürmektedir. İstenilen her bir değişken düzeyi için elde edilen parametreler kullanılarak benzer yorumlar elde edilebilir.

İyileşmemiş kısım için elde edilen ve Çizelge 4.26'da verilen sonuçlara göre metastaz değişkeninin ($p=0.005$) iyileşmemiş kısmın yaşam süresini etkileyen önemli bir faktör olduğu görülmektedir. Metastazı olan hastaların, metastazı olmayan hastalara göre ortanca yaşam süresi $\exp(0.757) = 2.3$ kat daha fazladır.

Bu veri kümesi için metastaz değişkeni anlamlı olduğundan, bu değişkenin radyoterapi değişkeni ile birlikte etkisi ayrıca incelenmek istenmiştir. Bu amaçla da bu değişkenle için de model kurulmuş ve iyileşmiş kısım ve iyileşmemiş kısım için elde edilen sonuçlar Çizelge 4.27 ve Çizelge 4.28'de verilmiştir.

Çizelge 4.27. Sadece metastaz ve radyoterapi değişkenleri kullanıldığında iyileşmiş kısım için elde edilen cloglog HBS iyileşme modeli sonuçları

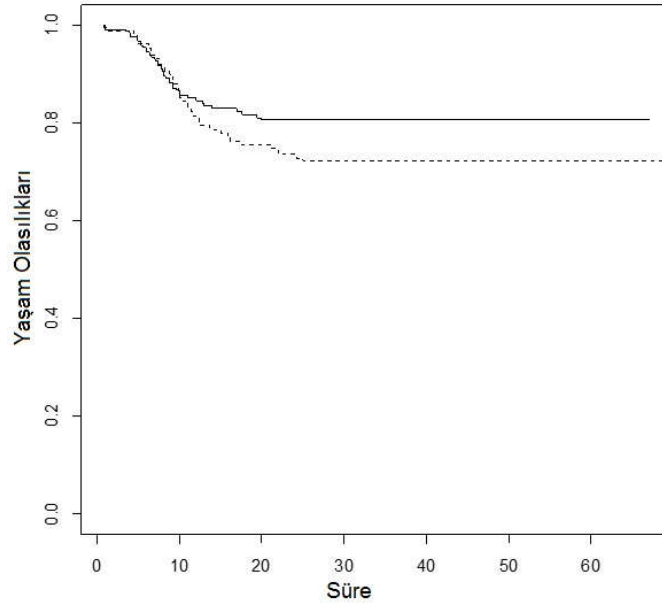
Değişken	$\hat{\beta}$	Std hata	z değeri	p değeri
Sabit	-0.957	0.309	-2.488	0.000
Metastaz	1.926	0.348	5.530	0.000
Radyoterapi	-0.468	0.364	-1.287	0.198

Çizelge 4.28. Sadece metastaz ve radyoterapi değişkenleri kullanıldığında iyileşmemiş kısım için elde edilen cloglog HBS iyileşme modeli sonuçları

Değişken	$\hat{\beta}$	Std hata	z değeri	p değeri
Sabit	0.065	0.237	0.272	0.786
Metastaz	0.819	0.225	3.634	0.000
Radyoterapi	-0.221	0.189	-1.167	0.243

Çizelge 4.27'de elde edilen parametreler kullanılarak, metastaz ve radyoterapi değişkenlerinin farklı düzeyleri için iyileşme oranları tahmin edilebilmektedir. Metastazı olmayan ve radyoterapi uygulanmış hastaların iyileşme oranı %78.6 olarak elde edilmiştir. Metastazı bulunmayan ve radyoterapi uygulanmamış hastalar için ise

iyileşme oranı %68.1'dir. Seçilmiş bu değişken düzeyleri için elde edilen yaşam olasılıkları Şekil 4.10'da verilmiştir.



Şekil 4.10. Farklı değişken düzeyleri için yaşam eğrileri 3

Şekil 4.10'da kesiksiz eğri, metastazı olmayan ve radyoterapi almış hastaların yaşam eğrisi, kesikli eğri ise metastazı olmayan ve radyoterapi uygulanmamış hastaların yaşam eğrisidir. Buradan, metastazı olmayan hastalar arasında radyoterapi uygulanan hastaların yaşam olasılıklarının, radyoterapi uygulanmayanlara göre daha yüksek olduğu görülmektedir. Ancak radyoterapi değişkeni ($p=0.198$) istatistiksel olarak anlamlı bulunmamıştır.

Metastazı olan ve radyoterapi uygulanmış hastalar için iyileşme oranı sadece %19.2 olarak bulunmuşken, metastazı olmayan ve radyoterapi almış hastalar için iyileşme oranı %78.6 olarak elde edilmiştir. Metastazı olmayan hastaların iyileşme oranının, metastazı olan hastalara göre çok daha yüksek olduğu ve iyileşme oranları arasındaki bu farklılığın istatistiksel olarak anlamlı olduğu ($p=0.000$) %95 güven düzeyinde söylenebilmektedir.

İstenilen tüm değişkenler ve değişken düzeyleri için benzer biçimde iyileşme oranları yorumlanabilmektedir.

5. SONUÇ VE ÖNERİLER

Sağlık bilimleri alanında yaşam çözümlenmesi önemli bir yere sahiptir. Günümüzde, özellikle kanser ile ilgili araştırmalarda yeni yaşam modellerinin önerilmesi ve bu modellerin uygulama alanlarının belirlenmesi üzerinde çalışmalar hız kazanmıştır. Yaşam çözümlenmesi çalışmalarının çoğunda ilgilenilen olay (ölüm, nüks etme, vb.) ortaya çıkana kadar geçen süre ve bu süreyi etkileyen faktörlerin incelenmesi için çalışmalar yapılmaktadır. Bu amaçla, en yaygın kullanıma sahip ve literatürde temel olarak kabul edilen yaşam modeli Cox regresyon modelidir. Bu model kitledeki her bir bireyin ilgilenilen olayı yaşayabileceği ve başarısızlığa (ölüme, nüksetmeye, vb.) yatkın olduğu varsayımına dayanmaktadır. Yeterli bir gözlem sürecinin sonunda hastaların büyük bir bölümünün hayatta kalması durumunda bu varsayım sağlanmayacaktır. Meme kanseri, çocukluk lenfomaları ve lösemileri gibi bazı kanser türlerindeki araştırmalar, bu hastalıklara sahip hastaların önemli bir bölümünün, uygulanan tedavilerden sonra iyileştiğini göstermiştir. Ayrıca son yıllarda hasta odaklı tedavi ön plana çıkmıştır ve hastalar, uygulanacak en uygun tedavinin seçiminde önemli bir rol oynamaktadırlar. Bu özellikler dikkate alınarak geliştirilen iyileşme modelleri, kanser hastalarının yaşam verilerini incelemek için kullanılan önemli bir istatistiksel model haline gelmiştir. İyileşme modellerinin özellikle orantılı tehlikeler varsayımı sağlanmıyorsa ya da çalışma grubu içerisinde heterojenliğin olduğu düşünülüyorsa klasik Cox regresyon modeline tercih edildiği literatürde belirtilmektedir.

İyileşme modelleri karma ve karma olmayan iyileşme modelleri olmak üzere iki biçimde ele alınmaktadır. Karma iyileşme modellerinde, kitle iki alt gruba ayrılabilir, kitledeki bireyler z açıklayıcı değişkenler vektörü iken, ya $1 - \pi(z)$ oranıyla iyileşmiştir ya da $\pi(z)$, ($0 \leq \pi \leq 1$) oranıyla iyileşmemiştir. Karma iyileşme modellerinde iyileşmiş kısım ve iyileşmemiş kısım ayrı ayrı modellenmektedir. Karma modellerin bu özelliği açıklayıcı değişkenlerin, iyileşmiş ve iyileşmemiş hastalar üzerindeki etkilerinin farklı olmasına olanak sağlamaktadır. Karma modellerin iyileşmiş kısmında modelleme yaparken lojit, tamamlayıcı log-log ya da probit gibi bazı bağlantı fonksiyonları kullanılmaktadır.

Karma iyileşme modelleri için çeşitli parametrik ve parametrik olmayan yöntemler önerilmiştir. Karma iyileşme modelinin iyileşmemiş kısmını modellemek için ise Cox

regresyon modeli ya da hızlandırılmış başarısızlık süresi modelleri kullanılmakta ve elde edilen karma iyileşme modeli, bu kısımda kullanılan modele göre adlandırılmaktadır. Karma iyileşme modeli başlığı altında Cox iyileşme modeli, yarı parametrik hızlandırılmış başarısızlık süresi iyileşme modeli, yarı parametrik hızlandırılmış tehlikeler karma iyileşme modeli ve parametrik karma iyileşme modeli incelenmiştir.

İyileşme modellerinde, Cox karma iyileşme modeli ve yarı parametrik hızlandırılmış başarısızlık süresi karma iyileşme modeli tahmin yöntemlerinin ve elde edilen sonuçların yorumlanmasının kolaylığı nedeniyle daha çok tercih edilmektedir. Bu modellerdeki varsayımların sağlanmadığı durumlar için de hızlandırılmış tehlikeler karma iyileşme modeli önerilmiştir. Hızlandırılmış başarısızlık süresi karma iyileşme modelinde bulunan ani etki varsayımı, tedavinin etkisinin zamanla birlikte kademeli olarak arttığı bazı kanser çalışmaları için tercih edilen bir varsayım olmayabilir. Hızlandırılmış tehlikeler karma modelinin, özellikle tedavinin ya da diğer açıklayıcı değişken etkilerinin zaman ile kademeli olarak arttığı durumları modellemek için yararlı olduğu görülmüştür. Parametrik karma iyileşme modelleri için de iyileşmemiş kısım için üstel, Weibull, log-normal, log-lojistik ve genelleştirilmiş Gamma gibi dağılımlar kullanılmaktadır. Ancak bu modele ait sonuçlar parametrik dağılımın seçimine ve doğruluğuna karşı duyarlıdır. Bu nedenle de kullanım alanı daha kısıtlıdır.

Karma olmayan iyileşme modelleri ise literatürde teorik açıdan yer almasına rağmen, uygulamada kullanımlarına rastlanmamaktadır.

İyileşme modelleri ile ilgili benzetim çalışmalarını içeren yayınlar literatürde olmakla birlikte uygulamalı çalışmalar daha nadir yer almaktadır. Bunun nedeni ise olabilirlik fonksiyonlarının karmaşıklığı ve buna bağlı olarak gerçek veriler için istatistiksel paket programlarındaki kullanım güçlüğüdür. Son yıllarda STATA, R ve SAS gibi istatistiksel paket programlarında hazırlanan kodları içeren yayınlar yapılmakta ve bu yazılımlar geliştirilmeye çalışılmaktadır. Ancak bu yayınlarda da modelin parametre tahminlerinde yakınsama başta olmak üzere bazı problemler olduğu belirtilmektedir. Bu da iyileşme modellerinin uygulamalı çalışan araştırmacılar açısından kullanımını zorlaştırmaktadır.

Bu tez çalışmasında karma iyileşme modelleri incelenmiş ve iyileşme modelinin kullanımına uygun olduğu düşünülen iki kanser verisi üzerinde uygulama yapılmıştır. Elde edilen iyileşme modellerine ait sonuçlar literatürde yer alan klasik yaşam çözümlerinden elde edilen sonuçlar ile birlikte değerlendirilmiştir.

İlk olarak bir beyin kanseri türü olan glioma hastası 411 bireyi içeren veri kümesi [72] ele alınmıştır. Çalışmada %66.7 başarısızlık (ölüm) ve %33.3 durdurma gözlenmiştir. Tedavi değişkeni olarak radyoterapi değişkeni ele alınmış ve bu değişken dışında da cinsiyet, yaş, malignite düzeyi ve karnofsky indeksi incelenen modellere dahil edilmiştir. Kaplan-Meier yaşam eğrisine bakıldığında eğrinin düz bir kuyruğa sahip olduğu görülmüş ve bu da kitlede ilgilenilen olayı yaşamayan bireylerin olabileceği sonucunu düşündürmüştür. Bu nedenle de iyileşme modellerinin kullanımının incelenmesinin veri için daha uygun olacağı sonucuna ulaşılmıştır. Veri için öncelikli olarak orantılı tehlikeler varsayımı incelenmiş ve varsayımın sağlandığı görülmüştür. Malignite derecesi, karnofsky indeksi ve yaş değişkenlerinin düzeyleri arasında yaşam olasılıkları açısından anlamlı bir farklılık olduğu %95 güven düzeyinde söylenebilmektedir. Cox regresyon modeli sonuçlarına göre yaş ve malignite derecesi değişkenlerinin yaşam süresini etkileyen önemli risk faktörleri olduğu %95 güven düzeyinde söylenebilmektedir. Elde edilen sonuçlara göre, malignite derecesi 3 olan hastaların, malignite derecesi 1 ve 2 olan hastalara göre başarısızlık riskinin 1.766 kat daha fazla olduğu; 54 yaş üzeri hastaların, yaşı 54 ve 54'den küçük olan hastalara göre başarısızlık riskinin 2.183 kat daha fazla olduğu görülmüştür. Orantılı tehlikeler varsayımı sağlandığı için Cox karma iyileşme modeli kullanılmış, lojit ve cloglog bağlantı fonksiyonları için model sonuçları elde edilmiştir. Lojit bağlantı fonksiyonu kullanıldığında, glioma veri kümesi için, malignite düzeyi 1 ya da 2 olan, karnofsky indeksi 80'den büyük, yeni tedavi uygulanmış ve yaşı 54 ya da 54'den küçük olan erkek hastaların iyileşme oranı %58.4 bulunmuştur. Aynı değişken düzeyleri için standart tedaviye ait iyileşme oranı hesaplandığında bu oranın %32'ye düştüğü görülmüştür. Aynı biçimde malignite düzeyi 1 ya da 2 olan, karnofsky indeksi 80'den büyük, yeni tedavi uygulanmış ve yaşı 54 ya da 54'den küçük olan kadın hastaları alırsak, bu grup için iyileşme oranı %51.7; aynı değişken düzeyleri için standart tedavi kullanıldığında iyileşmiş kısmın oranı sadece %26.4 olarak tahmin edilmiştir. Bu veri kümesi için, en

düşük iyileşme oranı malignite düzeyi 3 olan, 54 yaş üzeri, karnofsky indeksi 80'den küçük olan ve standart tedavi uygulanan kadınlar için %2 olarak elde edilmiştir. İyileşmemiş kısım için elde edilen sonuçlar, kitledeki iyileşmemiş hastalar için değişken düzeylerinin tehlike oranlarının yorumlanmasını sağlamaktadır. Bu model için malignite düzeyi ve yaş değişkeni istatistiksel olarak anlamlı bulunmuştur. Elde edilen sonuçlara dayanarak, malignite düzeyi 3 olan hastaların başarısızlık riskinin, malignite düzeyi 1 ve 2 olan hastalara göre 1.46 kat daha fazla olduğu söylenebilmektedir. Yaşı 54 üzeri olan hastaların, yaşı 54 ve 54'den küçük hastalara göre 2.02 kat daha fazla risk altında olduğu ifade edilebilmektedir.

Glioma verisi için cloglog bağlantı fonksiyonu kullanıldığında hem iyileşmiş hastaların iyileşme oranı için, hem de iyileşmemiş hastaların yaşam fonksiyonu için malignite düzeyi ve yaş değişkenlerinin %95 güven düzeyinde anlamlı olduğu görülmüştür. Malignite düzeyi 1 ya da 2 olan, karnofsky indeksi 80'den büyük, yeni tedavi türü uygulanmış ve yaşı 54 ya da 54'den küçük olan erkek hastaların iyileşme oranı %38.6; aynı değişken düzeyleri alınarak standart tedavi için hesaplama yapıldığında iyileşme oranı %38.1 olarak bulunmuştur. Bu model için, tedavi türündeki değişikliğin, iyileşmiş kısmın oranını etkileyen önemli bir değişken olmadığı görülmektedir. Yaş değişkeninin iyileşmiş kısma olan etkisini incelemek için, malignite düzeyi 1 ya da 2 olan, karnofsky indeksi 80'den büyük, yeni tedavi türü uygulanmış erkek hastalardan, yaşı 54 üzeri olan hastalar için iyileşme oranı %25.5 olarak elde edilmiştir. Bu modelde, düşük yaş grubundaki hastaların iyileşme oranının, yüksek yaş grubundaki hastaların iyileşme oranından daha yüksek olduğu ve yaş değişkeninin istatistiksel olarak anlamlı olduğu görülmektedir. Yüksek yaş, gruptaki iyileşme oranını azaltan bir faktördür. Malignite düzeyinin iyileşmiş kısımdaki etkisini incelemek için, Karnofsky indeksi 80'den büyük, yeni tedavi türü uygulanmış ve yaşı 54 ya da 54'den küçük olan erkek hastalardan malignite düzeyi 3 olanlar alınmış ve iyileşme oranı %23.3 olarak elde edilmiştir. Aynı değişken düzeylerinde, sadece malignite 1 veya 2 değerini aldığı bu oran %38.6 olarak elde edilmişti. Buna göre, malignite düzeyi 3 olanların iyileşme oranının, malignite düzeyi 1 ve 2 olanlardan daha düşük olduğu görülebilmektedir. İyileşmemiş kısım için ise malignite düzeyi 3 olan hastaların, malignite düzeyi 1 ve 2 olan hastalara göre 1.51 kat daha fazla risk altında bulunduğu ifade edilebilir. Yaşı 54'ün üzerinde olan

hastaların başarısızlık riskinin, yaşı 54 ve 54'ün altında olan hastaların başarısızlık riskine göre 1.92 kat daha fazla olduğu söylenebilmektedir. İstenilen tüm gruplar için iyileşme oranı, parametre tahminleri ve değişken düzeyleri kullanılarak benzer biçimde yorumlanabilmektedir. Bu veri için lojit Cox karma iyileşme modeli ve cloglog Cox karma iyileşme modeli sonuçlarına bakıldığında, cloglog bağlantı fonksiyonu kullanıldığında elde edilen modelin p değerlerinin ve standart hatalarının çok daha düşük olduğu görülmüştür ve bu nedenle cloglog Cox karma iyileşme modelinin kullanılması daha uygun olabilir.

İkinci olarak, Hacettepe Üniversitesi Biyoistatistik Ana Bilim Dalı'ndan elde edilen 106 mide kanseri hastasına ait veri kümesi kullanılmıştır [75]. Çalışmada hastaların takip süresi sona erdiğinde, 106 hastadan %63.2'sinde başarısızlık (ölüm) ve %36.8'inde durdurma gözlenmiştir. Tedavi değişkeni olarak radyoterapi değişkeni ele alınmış ve bu değişken dışında da metastaz, alkol kullanımı, aile öyküsü ve yaş değişkenleri modele dahil edilmiştir. Kaplan-Meier yaşam eğrisi incelendiğinde iyileşme modellerinin kullanımına uygun olduğu görülmüştür. Kullanılacak olan karma iyileşme modellerinin türüne karar verebilmek için orantılı tehlikeler varsayımı incelenmiş ve bu varsayımın sağlanmadığı sonucuna ulaşılmıştır.

Orantılı tehlikeler varsayımı sağlanmadığı için literatürde bu amaçla kullanılan tabakalandırılmış CRM, genişletilmiş CRM ve HBS modelleri öncelikli olarak incelenmiştir. Tabakalandırılmış CRM sonuçları modellerin anlamsız bulunması nedeni ile kullanılamamıştır. Genişletilmiş CRM sonuçlarına göre aile öyküsü ve metastaz değişkenlerinin başarısızlığı etkileyen önemli risk faktörleri olduğu %95 güven düzeyinde görülmüştür. Weibull HBS modelinin mide kanser verisi için en uygun HBS modeli olduğu sonucuna varılmıştır. Bu modele göre de aile öyküsü ve metastaz değişkenlerinin, başarısızlık süresini etkileyen önemli faktörler olduğu görülmüştür. Metastaz yaşamamış hastaların ortanca yaşam süresi, metastaz yaşamış olan hastaların yaşam süresine göre yaklaşık 3.1 kat daha uzun; aile öyküsü bulunmayan hastaların ortanca yaşam süresi, aile öyküsü bulunan hastalara göre 1.84 kat daha fazla olduğu görülmüştür. HBS karma iyileşme modeli için lojit ve probit bağlantı fonksiyonları ile sonuç alınmaz iken, cloglog bağlantı fonksiyonu kullanıldığında HBS karma iyileşme modeli sonuçları elde edilmiştir. Modelde anlamlı bulunan tek değişken

metastaz deęişkeni olmuştur. Ancak iyileşme modelleri farklı açılardan yorum yapmaya imkan sağladığından dięer deęişkenler de kullanılarak yorumlar elde edilmiştir. Alkol kullanmayan, aile öyküsü bulunmayan, radyoterapi uygulanmış, yaşı 56 ve 56'dan küçük ve metastazı bulunmayan hastalar için iyileşme oranı %79.5 olarak elde edilirken aynı deęişken düzeylerine sahip ancak metastazı bulunan hastalar için iyileşme oranı sadece %18.4 olarak hesaplanmıştır. Buna göre, metastaz deęişkeninin iyileşme oranı üzerinde etkili bir deęişken olduğu ve metastazı bulunan hasta grubu için iyileşme oranının çok daha düşük olduğu görülmektedir. Aynı deęişken düzeyleri için radyoterapi uygulanmayanlar seçildiğinde iyileşme oranı %70 bulunmuştur. Bu sonuca göre radyoterapi uygulanmasının iyileşmiş kısmın oranını arttırdığı düşünülmektedir. İyileşmemiş kısım için HBS sonuçları incelendiğinde metastaz deęişkeninin yaşam süresini etkileyen önemli bir faktör olduğu görülmüştür.

İyileşme modellerinin sağlık bilimlerinde kullanıma ilişkin Türkiye'de bir çalışmaya rastlanmamıştır. Benzer biçimde modelin uygulanması ve sonuçların elde edilmesinde zorluklar nedeni ile yurt dışında yapılan uygulamalı çalışmalara çok rastlanmamıştır. İyileşme modelleri, klasik yaşam modellerinden farklı yorumların da yapılabilmesine imkan sağladığı için kullanışlı bir modeldir. Bununla birlikte, uygulamaki güçlüklerinin ortadan kaldırılması gerekmektedir. Bu nedenle de özellikle istatistiksel paket programlardaki yazılımların geliştirilmesi konusunda çalışmalar yapılabilir.

KAYNAKLAR

- [1] Dickman, P.W., Hakulinen, T., *Population-Based Cancer Survival Analysis*, http://www.pauldickman.com/teaching/tampere2004/book_draft.pdf, **2003**. Erişim tarihi: Haziran 2016.
- [2] Clark, T.G., Bradburn, M.J., Love, S.B., Altman, D.G., Survival analysis part I: Basic concepts and first analyses, *British Journal of Cancer*, **2003**.
- [3] Hosmer, D.W., Lemeshow, S., *Applied Survival Analysis: Regression Modeling of Time to Event Data*, John Wiley & Sons, New York, **1999**.
- [4] Kaplan, E.L., Meier, P., Nonparametric Estimation from Incomplete Observations, *Journal of the American Statistical Association*, 53(282), 457-481, **1958**.
- [5] Peto, R., Pike, M.C., Armitage, P., Breslow, N.E., Cox, D.R., Howard, S.V., Mantel, N., McPherson, K., Peto, J., Smith, P.G., Design and analysis of randomized clinical trial requiring prolonged observation of each patient. II. Analysis and examples, *British Journal of Cancer*, 35(1), 1-39, **1977**.
- [6] Collett, D., *Modelling Survival Data in Medical Research*, Taylor & Francis, **1994**.
- [7] Altman, D.G., Bland, J.M., Interaction revisited: The difference between two estimates, *The BMJ*, 326(7382), 219, **2003**.
- [8] Box-Steffensmeier, J.M., Jones, B.S., *Event History Modelling: A Guide for Social Scientists*, Cambridge University Press, **2004**.
- [9] Cox, D.R., Regression models and life-tables, *Journal of the Royal Statistical Society*, 34(2), 187-220, **1972**.
- [10] Cox, D.R., Partial likelihood, *Biometrika*, 62(2), 269-276, **1975**.
- [11] Breslow, N.E., Covariance analysis of censored survival data, *Biometrics*, 30(1), 89-99, **1974**.
- [12] Vermunt, J.D., Metacognitive, cognitive and affective aspects of learning styles and strategies: A phenomenographic analysis, *Higher Education*, 31(1), 25-50, **1996**.
- [13] Efron, B., The efficiency of Cox's likelihood function for censored data, *Journal of the American Statistical Association*, 72(359), 312-319, **1977**.
- [14] Kalbfleisch, J.D., Prentice, R. L., Marginal likelihoods based on Cox's regression and life model, *Biometrika*, 60(2), 267-278, **1973**.
- [15] Kalbfleisch, J.D., Prentice, R.L., *The Statistical Analysis of Failure Time Data*, Wiley, New York, **1980**.
- [16] Orbe, J., Ferreira, E., Núñez-Antón, V., Comparing proportional hazards and accelerated failure time models for survival analysis, *Statistics in Medicine*, 21(22), 3493-3510, **2002**.
- [17] Rodriguez, G., *Parametric Survival Models*, Princeton University, <http://data.princeton.edu/wws509/notes/c7.pdf>, **2010**. Erişim tarihi: Ekim 2016.

- [18] Rama, R., Swaminathan, R., Venkatesan, P., Cure models for estimating hospital-based breast cancer survival, *Asian Pacific Journal of Cancer Prevention*, 11(2), 387-391, **2010**.
- [19] Tai, P., Yu, E., Cserni, G., Vlastos, G., Royce, M., Kunkler, I., Vinh-Hung, V., Minimum follow-up time required for the estimation of statistical cure of cancer patients: Verification using data from 42 cancer sites in the SEER database, *BMC Cancer*, 17(5), 48-57, **2005**.
- [20] Sposto, R., Cure model analysis in cancer: An application to data from the children's cancer group, *Statistics in Medicine*, 21(2), 293-312, **2002**.
- [21] Boag, J.W., Maximum Likelihood Estimates of the Proportion of Patients Cured by Cancer Therapy, *Journal of the Royal Statistical Society*, 11(1), 15-44, **1949**.
- [22] Berkson, J., Gage, R.P., Survival curve for cancer patients following treatment, *Journal of the American Statistical Association*, 47(259), 501-515, **1952**.
- [23] Farewell, V.T., A Model for a Binary Variable with Time-Censored Observations, *Biometrika*, 64(1), 43-46, **1977**.
- [24] Kuk, A.Y.C., Chen, C.H., A mixture model combining logistic regression with proportional hazards regression, *Biometrika*, 79(3), 531-541, **1992**.
- [25] Yakovlev, A.Y., Asselain, B., Bardou, V.J., Fourquet, A., Hoang, T., Rochefedieri, A., Tsodikov, A.D., A simple stochastic model of tumor recurrence and its applications to data on pre-menopausal breast cancer, *Biometrie et Analyse de Dormees Spatio-Temporelles*, 12, 66-82, **1993**.
- [26] Taylor, J.M.G., Semi-parametric estimation in failure time mixture models, *Biometrics*, 51(3), 899-907, **1995**.
- [27] Maller, R.A., Zhou, S., *Survival Analysis with Long-Term Survivors*, New York, Wiley, **1996**.
- [28] Chen, M., Ibrahim, J.G., Sinha, D., A new bayesian model for survival data with a surviving fraction, *Journal of American Statistical Association*, 94(447), 909-919, **1999**.
- [29] Sy, J.P., Taylor, J.M.G., Estimation in a proportional hazards cure model, *Biometrics*, 56(1), 227-336, **2000**.
- [30] Peng, Y., Dear, K.B., A nonparametric mixture model for cure rate estimation, *Biometrics*, 56(1), 237-243, **2000**.
- [31] Chatterjee, N., Shih, J., A bivariate cure-mixture approach for modeling familial association in diseases, *Biometrics*, 57(3), 779-786, **2001**.
- [32] Tsodikov, A.D., Ibrahim, J.G., Yakovlev, A.Y., Estimating cure rates from survival data: An alternative to two-component mixture models, *Journal of the American Statistical Association*, 98(464), 1063-1078, **2003**.
- [33] Wienke, A., Lichtenstein, P., Yashin, A.I., A bivariate frailty model with a cure fraction for modelling familial correlations in diseases, *Biometrics*, 59(4), 1178-1183, **2003**.

- [34] Shao, Q., Zhou, X., A New Parametric Model for Survival Data with Long-Term Survivors, *Statistics in Medicine*, 23(22), 3525-3543, **2004**.
- [35] Yin, G., Ibrahim, J.G., Cure Rate Models: A Unified Approach, *The Canadian Journal of Statistics*, 33(4), 559-570, **2005**.
- [36] Yu, G., A Frailty Mixture Cure Model with Application to Hospital Admission Data, *Biometrical Journal*, 50(3), 394, **2008**.
- [37] Zhang, J., Peng, Y., Accelerated Hazards Mixture Cure Model, *Lifetime Data Analysis*, 15(4), 455-467, **2009**.
- [38] Lu, W., Efficient estimation for an accelerated failure time model with a cure fraction, *Stastica Sinica*, 20(2), 661-674, **2010**.
- [39] Tong, E.N., Mues, C., Thomas, L.C., Mixture cure models in credit scoring: If and when borrowers default, *European Journal of Operational Research*, 218(1), 132-139, **2012**.
- [40] Cai, C., Zou, Y., Peng, Y., Zhang, J., Smcure: An R-package for estimating semiparametric mixture cure models, *Comput Methods Programs Biomed*, 108(3), 1255-1260, **2012**.
- [41] Maetani, S., Gamel, J.W., Parametric cure model versus proportional hazards model in survival analysis of breast cancer and other malignancies, *Advances in Breast Cancer Research*, 2(4), 119-125, **2013**.
- [42] Maetani, S., Gamel, J.W., Parametric cancer survival analysis based on the Boag model: A clinician's view, *Journal of Biometrics & Biostatistics*, 3(3), 7-17, **2012**.
- [43] Othus, M., Barlogie, B., LeBlanc, M.L., Crowley, J.J., Cure models as a useful statistical tool for analyzing survival, *Clinical Cancer Research*, 18(14), 3731-3736, **2012**.
- [44] Gamel, J.W., Vogel, R.L., Valagussa, P.V., Bonadona, G., Parametric survival analysis of adjuvant therapy for stage II breast cancer, *Cancer*, 74(9), 2483-2490, **1994**.
- [45] Gamel, J.W., Vogel, R.L., Comparison of parametric and non-parametric survival methods using simulated clinical data, *Statistics in Medicine*, 16(14), 1629-1643, **1997**.
- [46] Frankel, P., Longmate, J., Parametric models for accelerated and long-term survival: A comment on proportional hazards, *Statistics in Medicine*, 21(21), 3279-3289, **2002**.
- [47] Maetani, S., Segawa, Y., Banja, H., Obayashi, H., Simulation of randomized controlled trials for cancer therapy using the Boag model and the competing risk model, *Tenri Medical Bulletin*, 11(1), 23-48, **2008**.
- [48] Silvestri, G.A., Handy, J., Lackland, D., Corley, E., Reed, C.E., Specialists achieve better outcomes than generalists for lung cancer surgery, *Chest*, 114(3), 675-680, **1998**.

- [49] Cai, C., Advanced Methodology Developments in Mixture Cure Models, University of South Carolina, <http://scholarcommons.sc.edu/etd/544>, **2013**.
- [50] Farewell, V.T., The use of mixture models for the analysis of survival data with long-term survivors, *Biometrics*, 38(4), 1041–1046, **1982**.
- [51] Sumathi, K., Aruna, R.K., Cure rate models: A partial review with an application to recurrent event or count data, *InterStat: Statistics on the Internet*, 11(1), 1-25, **2008**.
- [52] Breslow, N.E., Analysis of survival data under the proportional hazards model, *International Statistical Review*, 43(1), 45-57, **1975**.
- [53] Tsiatis, A., Estimating regression parameters using linear rank tests for censored data, *The Annals of Statistics*, 18(1), 354-372, **1990**.
- [54] Ritov, Y., Estimation in a linear regression model with censored data, *The Annals of Statistics*, 18(1), 303-328, **1990**.
- [55] Yamaguchi, K., Accelerated failure-time regression-models with a regression-model of surviving fraction: An application to the analysis of permanent employment in Japan, *Journal of the American Statistical Association*, 87(418), 284–292, **1992**.
- [56] Peng, Y.W., Dear, K.B.G., Denham, J.W., A generalized F mixture model for cure rate estimation, *Statistics in Medicine*, 17(8), 813–830, **1998**.
- [57] Li, C.S., Taylor, J.M., A semi-parametric accelerated failure time cure model, *Statistics in Medicine*, 21(21), 3235-3247, **2002**.
- [58] Zhang, J., Peng, Y., A new estimation method for the semiparametric accelerated failure time mixture cure model, *Statistics in Medicine*, 26(16), 3157–3171, **2007**.
- [59] Jin, Z., Lin, D.Y., Wei, L.J., Ying, Z., Rank-Based Inference for the Accelerated Failure Time Model, *Biometrika*, 90(2), 341-353, **2003**.
- [60] Chen, Y.Q., Wang, M., Analysis of accelerated hazards models, *Journal of the American Statistical Association*, 95(450), 608–618, **2000**.
- [61] Chen, Y.Q., Accelerated hazards regression model and its adequacy for censored survival data, *Biometrics*, 57(3), 853–860, **2001**.
- [62] Kalbfleisch, J.D., Prentice, R.L., *The Statistical Analysis of Failure Time Data*, Second Edition, John Wiley & Sons, New Jersey, **2002**.
- [63] Louis, T.A., Finding the observed information matrix when using the EM algorithm, *Journal of the Royal Statistical Society, Series B*, 44(2), 226-233, **1982**.
- [64] Meilijson, I., A fast improvement to the EM algorithm on its own terms, *Journal of the Royal Statistical Society, Series B*, 51(1), 127-138, **1989**.
- [65] Lambert, P.C., Modeling of the cure fraction in survival studies, *The Stata Journal*, 7(3), 351–375, **2007**.

- [66] Corbiere, F., Joly, P., A SAS macro for parametric and semiparametric mixture cure models, *Computer Methods and Programs in Biomedicine*, 85(2), 173-180, **2007**.
- [67] Lindsey, J.K., *Parametric Statistical Inference*, Oxford University Press, USA, **1996**.
- [68] Ibrahim, J.G., Chen, M.H., Sinha, D., Bayesian semiparametric models for survival data with a cure fraction, *Biometrics*, 57(2), 383-388, **2001**.
- [69] Yin, G., Bayesian cure rate frailty models with applications to a root canal therapy study, *Biometrics*, 61(2), 552-558, **2005**.
- [70] Zeng, D., Yin, G., Ibrahim, J.G., Semiparametric transformation models for survival data with a cure fraction, *Journal of the American Statistical Association*, 101(474), 670-684, **2006**.
- [71] Yakovlev, A.Y., Tsodikov, A.D., Stochastic models of tumor latency and their biostatistical application, *World Scientific*, Singapore, **1996**.
- [72] Royston, P., Sauerbrei, W., *Fractional Polynomials for One Variable, in Multivariable Model-Building: A Pragmatic Approach to Regression Analysis Based on Fractional Polynomials for Modelling Continuous Variables*, John Wiley & Sons, Chichester, UK, **2008**.
- [73] Schoenfeld, D., Partial residuals for the proportional hazards regression model, *Biometrika*, 69(1), 239-241, **1982**.
- [74] Kleinbaum, D.G., Klein, M., *Survival Analysis: A Self-Learning Text*, Second Edition, Springer-Verlag, New York, **2005**.
- [75] Akkaya, E., *Mide Kanseri Verileri İçin Cox Regresyon Çözümlemesi*, Rapor, Hacettepe Üniversitesi, Fen Fakültesi İstatistik Bölümü, İleri İstatistik Projeleri, Ankara, **2008**.

ÖZGEÇMİŞ

Kimlik Bilgileri

Adı Soyadı : Pınar KARA
Doğum Yeri : Ankara
Medeni Hali : Bekar
E-posta : pinar.kara@hacettepe.edu.tr
Adresi : 1522.sokak 10/16 100.YIL ANKARA

Eğitim

Lise : 2002-2006 Dr. Binnaz Rıdvan Ege Anadolu Lisesi
Lisans : 2007-2012 Hacettepe Üniversitesi İstatistik Bölümü
Yüksek Lisans : 2013-2017 Hacettepe Üniversitesi İstatistik Bölümü

Yabancı Dil ve Düzeyi

İngilizce, çok iyi
Almanca, orta

İş Deneyimi

Deneyim Alanları

Tezden Üretilmiş Projeler ve Bütçesi

Tezden Üretilmiş Yayınlar

Tezden Üretilmiş Tebliğ ve/veya Poster Sunumu ile Katıldığı Toplantılar



HACETTEPE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
YÜKSEK LİSANS/DOKTORA TEZ ÇALIŞMASI ORJİNALLİK RAPORU

HACETTEPE ÜNİVERSİTESİ
FEN BİLİMLER ENSTİTÜSÜ
İSTATİSTİK ANABİLİM DALI BAŞKANLIĞI'NA

Tarih: 05/06/ 2017

Tez Başlığı / Konusu : YAŞAM ÇÖZÜMLEMESİNDE İYİLEŞME MODELLERİ

Yukarıda başlığı/konusu gösterilen tez çalışmamın a) Kapak sayfası, b) Giriş, c) Ana bölümler ve d) Sonuç kısımlarından oluşan toplam 81 sayfalık kısmına ilişkin, 05/06/2017 tarihinde şahsım/tez danışmanım tarafından Turnitin adlı intihal tespit programından aşağıda belirtilen filtrelemeler uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı %7 'dir.

Uygulanan filtrelemeler:

- 1- Kabul/Onay ve Bildirim sayfaları hariç,
- 2- Kaynakça hariç
- 3- Alıntılar hariç/~~dâhil~~
- 4- 5 kelimedenden daha az örtüşme içeren metin kısımları hariç

Hacettepe Üniversitesi Fen Bilimleri Enstitüsü Tez Çalışması Orijinallik Raporu Alınması ve Kullanılması Uygulama Esasları'nı inceledim ve bu Uygulama Esasları'nda belirtilen azami benzerlik oranlarına göre tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Gereğini saygılarımla arz ederim.

05.06.2017

Adı Soyadı: Pınar Kara
Öğrenci No: N13120630
Anabilim Dalı: İstatistik
Programı: Yüksek Lisans
Statüsü: Y.Lisans Doktora Bütünleşik Dr.

DANIŞMAN ONAYI

UYGUNDUR.

Doç. Dr. Nihal Ata Tutkun



HACETTEPE UNIVERSITY
GRADUATE SCHOOL OF SCIENCE AND ENGINEERING
THESIS/DISSERTATION ORIGINALITY REPORT

HACETTEPE UNIVERSITY
GRADUATE SCHOOL OF SCIENCE AND ENGINEERING
TO THE DEPARTMENT OF STATISTICS

Date: 05/06/2017

Thesis Title / Topic: CURE MODELS IN SURVIVAL ANALYSIS

According to the originality report obtained by myself/my thesis advisor by using the Turnitin plagiarism detection software and by applying the filtering options stated below on 05/06/2017 for the total of 81 pages including the a) Title Page, b) Introduction, c) Main Chapters, and d) Conclusion sections of my thesis entitled as above, the similarity index of my thesis is 7%.

Filtering options applied:

1. Approval and Declaration sections excluded
2. Bibliography/Works Cited excluded
3. Quotes excluded
4. Match size up to 5 words excluded

I declare that I have carefully read Hacettepe University Graduate School of Social Sciences Guidelines for Obtaining and Using Thesis Originality Reports; that according to the maximum similarity index values specified in the Guidelines, my thesis does not include any form of plagiarism; that in any future detection of possible infringement of the regulations I accept all legal responsibility; and that all the information I have provided is correct to the best of my knowledge.

I respectfully submit this for approval.

05.06.2017

Name Surname: Pinar Kara
Student No: N13120630
Department: Statistics
Program: Masters
Status: Masters Ph.D. Integrated Ph.D.

ADVISOR APPROVAL

APPROVED.

Assoc. Prof. Dr. Nihal Ata Tutkun