

ÖZGE PASİN

İSTANBUL ÜNİVERSİTESİ SAĞ. BİL. ENST.

DOKTORA TEZİ

İSTANBUL-2019

**T.C.
İSTANBUL ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ**

(DOKTORA TEZİ)

**MEME KANSERLİ HASTALARDA ÖLÜM RİSKİNİN BİRLEŞİK
ZAYIFLIK MODELLERİ İLE İNCELENMESİ**

ÖZGE PASİN

**DANIŞMAN
PROF.DR. AHMET DİRİCAN**

**İKİNCİ DANIŞMAN
PROF. DR. HANDAN ANKARALI**

**BİYOİSTATİSTİK ANABİLİM DALI
BİYOİSTATİSTİK VE TIP BİLİŞİMİ PROGRAMI**

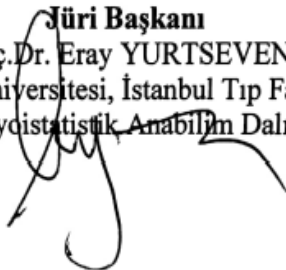
İSTANBUL-2019

TEZ ONAYI

İstanbul Üniversitesi Sağlık Bilimleri Enstitüsü İstanbul Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Biyoistatistik ve Tıp Bilişimi Programında Doktora öğrencisi Özge PASİN tarafından Prof.Dr. Ahmet DİRİCAN ve Prof.Dr. Handan ANKARALI'nın danışmanlığında hazırlanan "Meme Kanseri Hastalarda Ölüm Riskinin Birleşik Zayıflık Modelleri ile İncelenmesi" başlıklı tez aşağıdaki jüri üyeleri tarafından 4/12/2019 tarihinde yapılan Tez Savunma Sınavında başarılı bulunmuş ve Doktora Tezi olarak kabul edilmiştir.

Jüri Başkanı

Doç.Dr. Eray YURTSEVEN
İstanbul Üniversitesi, İstanbul Tıp Fakültesi,
Biyoistatistik Anabilim Dalı



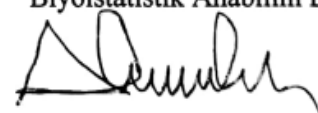
Jüri-Danışman

Prof.Dr. Ahmet DİRİCAN
İÜ-Cerrahpaşa, Cerrahpaşa Tıp Fakültesi,
Biyoistatistik Anabilim Dalı



Jüri

Prof.Dr.Abdülbari BENER
İÜ-Cerrahpaşa, Cerrahpaşa Tıp Fakültesi,
Biyoistatistik Anabilim Dalı



Jüri

Doç.Dr. Atilla USLU
İstanbul Üniversitesi, İstanbul Tıp Fakültesi,
Fizyoloji Anabilim Dalı



Jüri

Dr.Öğr.Üyesi Ömer UYSAL
Bezmialem Vakıf Üniversitesi, Tıp Fakültesi,
Biyoistatistik ve Tıp Bilişimi Anabilim Dalı



BEYAN

Bu tez çalışmasının kendi çalışmam olduğunu, tezin planlanmasından yazımına kadar bütün safhalarda etik dışı davranışımın olmadığını, bu tezdeki bütün bilgileri akademik ve etik kurallar içinde elde ettiğimi, bu tez çalışmayla elde edilmeyen bütün bilgi ve yorumlara kaynak gösterdiğimi ve bu kaynakları da kaynaklar listesine aldığımı, yine bu tezin çalışılması ve yazımı sırasında patent ve telif haklarını ihlal edici bir davranışımın olmadığı beyan ederim.

Özge PASİN



İTHAF

“Canım annem NEŞE PASİN’E ithaf ediyorum.”

TEŞEKKÜR

Mesleğimizin en önemli basamaklardan biri olan doktora tezimin her aşamasında yardımlarını esirgemeyen hocalarım Prof. Dr. Ahmet Dirican'a, Prof. Dr. Handan Ankaralı' ya ve Prof. Dr. Rian Dişçi'ye, Doç.Dr. Eray Yurtseven'e, mesai arkadaşım Dr. Başak Gürtekin'e, verilerin elde edilmesi aşamasında yardımcı olan Prof. Dr. Hasan Karanlık'a teşekkür ederim. Her zaman yanımda olan, bugünlere gelmemi sağlayan, beni yüreklendiren çok değerli annem Neşe Pasin'e, babam Fatih Pasin'e ve ablam Tuğçe Pasin'e sonsuz teşekkürlerimi ve saygılarımı sunarım.



İÇİNDEKİLER

TEZ ONAYI	İİ
BEYAN	İİİ
İTHAF	İV
TEŞEKKÜR.....	V
İÇİNDEKİLER	VI
TABLolar LİSTESİ	İX
ŞEKİLLER LİSTESİ.....	X
SEMBOLLER / KISALTMALAR LİSTESİ	Xİ
ÖZET.....	Xİİ
ABSTRACT.....	Xİİİ
1. GİRİŞ VE AMAÇ	1
2. GENEL BİLGİLER.....	4
2.1. Sağkalım Çözümlemesi ve Temel Kavramlar.....	4
2.1.1. Sağkalım Fonksiyonu ve Hazard Fonksiyonu	6
2.1.2. Parametrik Sağkalım Dağılımları	7
2.1.2.1. Üstel Dağılım	8
2.1.2.2. Weibull Dağılım.....	8
2.1.2.3. Lognormal Dağılım	9
2.1.2.4. Gamma Dağılım	10
2.1.2.5. Log-lojistik Dağılım	11
2.1.2.6. Gompertz Dağılım.....	11
2.1.2.7. Pareto Dağılım	12
2.1.3. Sağkalım ve Hazard Fonksiyonlarının Tahmini	13
2.1.3.1. Kaplan Meier Tahmincisi	14
2.1.3.2. Nelson-Aalen Tahmincisi	15
2.1.4. Hazard Fonksiyonunu Modelleme	16
2.1.4.1. Genelleştirilmiş Orantılı Hazard Modeli (Cox Regresyon Modeli).....	16
2.1.4.2. Modele Kovaryat Değişkeninin Eklenmesi	17
2.1.4.3. Modele Faktör Değişkeninin Eklenmesi	18
2.1.4.4. Cox Modeline Zamana Bağımlı Kovaryat Değişken Eklenmesi	18
2.1.4.5. Modele Etkileşim Etkisinin Eklenmesi	19

2.1.4.6. Cox Regresyon Modelinin Tahmini	20
2.1.4.7. β Katsayılarının Hipotez Testi ve Güven Aralıkları.....	22
2.1.4.8. Hazard Oranlarının Standart Hatası ve Güven Aralıkları.....	22
2.1.5. Cox Regresyon Analizinin Varsayımları ve Orantılı Hazard Varsayımı	22
2.1.5.1. Grafiksel Yöntemler	23
2.1.5.2. Cox Orantılı Hazard Modeline Zamana Bağımlı Kovaryat Eklenmesi	24
2.1.5.3. Schoenfeld Artıkları ile Test Etme	24
2.1.6. Cox Regresyon Modelinin Artıkları	25
2.2. Zayıflık Modelleri (Frailty Models)	27
2.2.1. Kesikli Zayıflık Modeli	31
2.2.2. Gamma Zayıflık Modeli	33
2.2.3. Lognormal Zayıflık Modeli	35
2.2.4. Ters Gauss Zayıflık Modeli	36
2.2.5. Pozitif Durağan Zayıflık Modeli.....	38
2.2.6. Power Varians Function (PVF) Zayıflık Modeli	39
2.2.7. Bileşik Poisson Zayıflık Modeli	41
2.2.8. Kuadratik Hazard Zayıflık Modeli.....	43
2.2.9. Lèvy Zayıflık Modeli	45
2.2.10. Log-t Zayıflık Modeli.....	45
2.2.11. Paylaşılmış Zayıflık Modeli (Shared Frailty Model)	45
2.2.12. İlişkili Zayıflık Modeli (Correlated Frailty Model)	48
2.2.13. İyileşme Zayıflık Modeli (Cure Frailty Model).....	50
2.2.14. Birleşik Zayıflık Modeli (Joint Frailty Model).....	52
2.3. Dinamik Tahmin Yöntemi	56
2.3.1. Birinci Tahmin Yöntemi ($P^1(t, t + w, \xi)$).....	58
2.3.2. İkinci Tahmin Yöntemi ($P^2(t, t + w, \xi)$)	58
2.3.3. Üçüncü Tahmin Yöntemi ($P^3(t, t + w, \xi)$)	59
2.4. Meme Kanseri ve Etki Eden Faktörler	59
3. GEREÇ VE YÖNTEM.....	64
3.1. Veri	64
3.2. İstatistiksel Analiz	64
4. BULGULAR	66

5. TARTIŞMA	86
KAYNAKLAR	92
İNTİHAL ORANI.....	97
ÖZGEÇMİŞ	98



TABLOLAR LİSTESİ

Table 2-1 : A ve B faktörlerinin seviyeleri	19
Table 2-2: Meme kanseri evreleri	61
Table 2-3: Meme kanseri evrelerinin detaylı gruplandırılması	62
Table 4-1: Meme kanseri hastalarına ait tanımlayıcı istatistikler	66
Table 4-2: Cox regresyon analizine ait tahmin sonuçları	68
Table 4-3: Splines yöntemi kullanılarak birleşik zayıflık modeli ile elde edilen tahmin sonuçları.....	72
Table 4-4: Splines-per yöntemi kullanılarak birleşik zayıflık modeli ile elde edilen tahmin sonuçları.....	76
Table 4-5: $t=30$ durumunda dinamik tahmin ile elde edilen ölüm risklerine ait tahmin değerleri	81
Table 4-6: $t=30$ durumunda dinamik tahmin ile elde edilen yeni bir nüks olayını yaşama riskine ait tahmin değerleri	82
Table 4-7: $t=30$ durumunda dinamik tahmin ile elde edilen tahmin değerlerinin Brier skor ile değerlendirilmesi	84
Table 4-8: $t=40$ durumunda dinamik tahmin ile elde edilen tahmin değerlerinin Brier skor ile değerlendirilmesi	85
Table 4-9: $t=50$ durumunda dinamik tahmin ile elde edilen tahmin değerlerinin Brier skor ile değerlendirilmesi	85

ŞEKİLLER LİSTESİ

Şekil 4-1: Kaplan Meier eğrisi	67
Şekil 4-2: Cox regresyon analizi sonucunda elde edilen tahminler için Martingale artıkları	69
Şekil 4-3: Meme kanserli hastaların zayıflık tahminleri	70
Şekil 4-4: Splines yöntemi kullanıldığında nüks olayı için Martingale artıkları	73
Şekil 4-5: Splines yöntemi kullanıldığında ölüm olayı için Martingale artıkları	73
Şekil 4-6: Splines ve Splines-per yöntemlerinin karşılaştırılması	77
Şekil 4-7: Splines-per yöntemi kullanıldığında nüks olayı için Martingale artıkları	78
Şekil 4-8: Splines-per yöntemi kullanıldığında ölüm olayı için Martingale artıkları	79
Şekil 4-9: Tahmin zamanı 30 ay için veri setinde yer alan beş bireye ait tahmini birikimli ölüm riskleri.....	83

SEMBOLLER / KISALTMALAR LİSTESİ

CVPOL	: Cross-Validated Prognosis Observed Loss
EPOCE	: Expected Prognostic Observed Cross-Entropy
ER	: Östrojen Reseptörü
İDK	: İnvaziv Duktal Karsinom
MCMC	: Markov Chain Monte Carlo
MKC	: Meme Koruyucu Cerrahi
MPLE	: Maximum Penalized Likelihood Estimation
MRM	: Modified Radical Mactectomy
NAKT	: Neoadjuvan Kemoterapi
PR	: Progesteron Reseptörü
PVF	: Power Variance Function
TNM	: Tümör Nod Metastaz

ÖZET

PASİN, Ö. (2019). Meme Kanserli Hastalarda Ölüm Riskinin Birleşik Zayıflık Modelleri ile İncelenmesi. İstanbul Üniversitesi Sağlık Bilimleri Enstitüsü, Biyoistatistik Anabilim Dalı. Doktora Tezi. İstanbul

Meme kanseri, kadınlarda en sık görülen kanser türleri arasında yer almaktadır ve hastalığın görülme oranı gün geçtikçe artmaktadır. Dolayısıyla bu alanda yapılan çalışmalarda artmaktadır. Yapılan çalışmaların birçoğunda olguların sağkalım sürelerini etkileyen risk faktörleri araştırılırken, bireyler arasındaki heterojenlikler dikkate alınmamaktadır ve nüks,ölüm olayları birbirinden bağımsız değerlendirilmektedir. Ancak ölüm, nüks ve metastaz olaylarının birbirinden bağımsız değerlendirilerek klasik sağkalım analizlerinin yapılması elde edilen parametre tahminlerinin yanıltıcı olmasına neden olacaktır. Bu durumda birleşik zayıflık modellerinin kullanımı önerilmektedir. Birleşik zayıflık modelleri ile bireyler arasındaki heterojenlikler dikkate alınmakta ve bireylerin ölüm riskleri modellenirken tekrarlayan olaylar da değerlendirilmektedir. Yapılan bu çalışmanın amacı, meme kanserli hastalarda nüks ile ölüm arasındaki ilişkiyi, bireyler arasındaki heterojenliği araştırarak ve dikkate alarak, ölüm ve nüks üzerine etkili olan faktörleri birleşik zayıflık modeli ile incelemektir. Çalışmanın ikincil amacı ise, dinamik tahmin yöntemleri yardımıyla hastanın tüm hikayesini değerlendirilerek ölüm risklerini tahmin etmektir. Çalışmada 2009-2016 yılları arasında İstanbul Tıp Fakültesi Onkoloji Enstitüsünde ameliyat olan 465 hastanın verileri kullanılmıştır. Çalışmanın sonucunda, nüks ve ölüm arasında pozitif ilişki gözlenmiştir. Ayrıca bireyler arasında heterojenlik gözlenmiştir. Nüks için tümör tipi, ameliyattan sonraki lenf nodu sayısı, neoadjuvan kemoterapi alma durumu, klinik evre, ameliyat tipi faktörlerinin etkileri istatistiksel olarak anlamlı bulunurken, ölüm için ise Cerb-B2, tümör tipi, ameliyattan sonra lenf nodu sayısı, neoadjuvan kemoterapi alma durumu faktörlerinin etkileri anlamlı bulunmuştur. Dinamik tahmin yöntemi ile elde edilen tahminlerin değerlendirilmesinde kullanılan Brier skorlarının değerleri 0,30'un altında bulunmuştur.

Sonuç olarak, zayıflık etkisi anlamlı olduğu durumlarda, aynı özelliklere sahip bireylerin dahi ölüm riskleri birbirinden farklılık göstermektedir. Nüks ve ölüm olaylarının birbiri ile ilişkili olduğu ve heterojenlik etkilerinin saptandığı durumlarda birleşik zayıflık modellerinin kullanımı önerilmektedir.

Anahtar Kelimeler: Hazard Oranı, Heterojenlik, Meme Kanseri, Nüks, Ölüm, Sağkalım, Zayıflık

ABSTRACT

PASİN, Ö. (2019). Assessment of Death Risk of Breast Cancer Patients with Joint Frailty Models. Istanbul University Institute of Health Sciences, Department of Biostatistics. PhD Thesis. Istanbul

Breast cancer is among the most common cancers in women and prevalence of disease has currently been increasing. Therefore, the number of studies on this area has been increasing. In most studies heterogeneities between the individuals, the dependency between recurrence and death are not considered while investigating the risk factors affecting the survival time of patients. However, conducting standard survival analyses by evaluating the death, recurrence and metastases independently will result biased parameter estimates. The use of joint frailty models are recommended in such cases. The heterogeneities between the individuals are considered, and recurrent events may also be evaluated in modelling the death risks of individuals with joint frailty models. The aim of the present study was to investigate the relationship between recurrence and death in breast cancer patients considering the heterogeneity between individuals and to investigate the factors effective on death and recurrence using the joint frailty model. The secondary aim of the study was to predict the death risks using the dynamic estimation methods after evaluating all history of the patient. The data of 465 patients who underwent surgery in Oncology Institute of Istanbul Faculty of Medicine between 2009 and 2016 were used in the study. A positive correlation was detected between recurrence and death at the end of the study. The effects of tumor type, the number of lymph node after surgery, neoadjuvant chemotherapy receiving status, clinical stage, and surgery type factors were found statistically significant for recurrence, and the effects of Cerb-B2, tumor type, the number of lymph nodes after surgery, and neoadjuvant chemotherapy receiving factors were found significant for death. Brier score values used in the evaluation of the obtained estimation using the dynamic estimation method were found below 0.30.

In conclusion, the death risks of even the individuals with the same characteristics are different from each other in patients when the frailty effect are significant. The use of joint frailty models is recommended when the recurrence, and death cases are associated, and heterogeneity effects are detected.

Keywords: Hazard Ratio, Heterogeneity, Breast Cancer, Recurrence, Death, Survival, Frailty

1. GİRİŞ VE AMAÇ

Geçmişten günümüze kanser hastalığı en önemli sağlık sorunu haline gelmiştir. Dünyadaki ölüm nedenlerine ait istatistikler incelendiğinde, kanserin ölüme sebebiyet veren ikinci en önemli faktör olduğu belirlenmiştir (Çevik ve Pirinçci 2017). 2018 yılındaki istatistikler incelendiğinde dünyada yaklaşık 9,6 milyon kişinin kanser hastalığı nedeniyle hayatını kaybettiği görülmüştür (WHO 2018). Dolayısı ile her 6 ölümden birinin nedeninin kanser olduğu söylenebilmektedir. Bu rakamların ise ilerideki yıllarda daha da artacağı düşünülmektedir. Meme kanseri ise bu rakamların oldukça büyük bir kısmını oluşturmaktadır. Dünya sağlık örgütü verilerine göre ise dünyada en sık görülen kanser türleri arasında meme kanseri ikinci sırada yer almaktadır. Meme kanseri kadınlarda ise en sık karşılaşılan kanser türüdür. Meme kanserine yakalanma sıklığının ilerleyen yıllarda daha da artacağı düşünülmektedir. 2018 yılında toplam kanser olgularının %11,6'sının meme kanseri olduğu belirlenmiştir (Bray ve ark. 2018). Ülkemizde de benzer şekilde meme kanseri sıklığı son 20 yılda iki katından fazla artış göstermektedir. Ayrıca gelişmekte olan ülkelerde özellikle hastaların çoğunluğunu genç, tümör boyutu büyük, evresi ileri ve prognozu kötü olan hastalar oluşturmaktadır. Dolayısıyla riski bu kadar yüksek olan bir hastalık için de teşhis ve tedaviye yönelik çalışmalar da hızla devam etmektedir.

Günümüz koşullarında meme kanseri hücrelerinin oluşumunu tamamen engellemek çok da olanaklı değildir. Ancak gelişen tedavi yöntemleri ile birlikte hastalık erken evrelerde teşhis edilebilmekte ve gerekli tedavi yöntemleri uygulanabilmektedir. Böylece meme kanserinin erken evrelerinde, tedavi uygulamalarında başarı elde edilmekte ve hastaların uzun süre yaşamaları sağlanmaktadır. Hastalığın, ileri evrelerde teşhisi ise bu durumu daha zorlaştıracaktır ve hastalara verilen tedavinin şekli ve yolu çok önemli bir seçim haline gelecektir.

Dünyada meme kanserine yakalanan bireylerdeki artış ile birlikte, bu alanda yapılan çalışmalarda önemli gelişmeler yaşanmaktadır. Her geçen yıl yeni bir ilaç veya yeni bir tedavi yöntemi geliştirilmektedir. Bu yöntemlerin kullanılabilmesi için hangi hastalık grubu için hangi yöntemin fayda sağladığı veya yöntemin bilinen, uygulanan yöntemlerden farklı olup olmadığı araştırılmaktadır. Yapılan araştırmalarda ana amaç, hastanın daha kaliteli, sağlıklı ve uzun bir yaşam sürmesini sağlayacak yöntemlerin

bulunmasıdır. Bu aşamada araştırmacılar tedavilerin etkilerini ve hastalığa etki eden risk faktörlerinin incelenmesinde sağkalım analizlerinden yararlanmaktadır. Yapılan araştırmaların birçoğunda meme kanserine etki eden faktörler değerlendirilirken, yalnızca ölüm veya yalnızca tekrarlayan olaylar dikkate alınmaktadır (Lafourcade ve ark. 2018). Ancak nüks ve ölüm birbiri ile ilişkili iki durumdur, bu iki olayı birbirinden bağımsız düşünmek doğru bir yaklaşım olmayacaktır. Nüks gerçekleşen hastaların ölüm ile karşılaşma riski, nüks gerçekleşmeyen hastalara göre daha yüksek olması beklenmektedir. Ayrıca bazı bireylerde nüks gerçekleşme olasılığı diğer bireylere göre daha fazladır. Bu farklılıklar yani bireyler arasındaki heterojenlikler, kişinin genetik yapısına, yaşadığı çevresel faktörlerin etkisine, beslenme alışkanlıklarına ve başka birçok faktöre bağlıdır. Buna karşın hastalık üzerine etkisi olabilecek tüm faktörleri bir model içerisinde değerlendirmek mümkün değildir. Çünkü hastalığa etkisi olabilecek ancak bilinmeyen birçok faktör söz konusudur. Dolayısı ile dikkate alınmayan veya etkisi bilinmeyen faktörler sebebiyle yapılan araştırmaların birçoğunda bireyler arasında bir heterojenite ortaya çıkmaktadır. Bireyler arasındaki bu heterojenite dikkate alınmaz ise parametre tahminleri yanlı olacaktır. Son yıllarda bireyler arasındaki heterojenite, nüks ve ölüm arasındaki bağımlılığı dikkate alarak faktörlerin etkilerini inceleyen birleşik zayıflık modelleri önerilmiştir (Abdulkarimova 2013; Rondeau ve ark. 2007).

Bu tez çalışmasında, bireyler arasındaki heterojeniteyi dikkate alan ve birbiri ile bağımlı sonlanma olaylarını birlikte değerlendiren birleşik zayıflık modelinin teorik özelliklerini açıklamak ve uygulama alanlarını göstermek amaçlanmıştır. Dolayısı ile yapılan çalışmada meme kanseri hastalarının sonlanma olayına etki eden faktörler incelenirken nüks ve ölüm arasındaki ilişki dikkate alınarak, ölüm ve nüks üzerine etkili olan faktörler birleşik zayıflık modeli ile incelenecektir. Uygulama amacıyla meme kanseri hastalarına ait veriler kullanılacak ve birleşik zayıflık modelinin sonuçları ile klasik sağkalım analiz sonuçları karşılaştırılacaktır. Çalışmanın ikincil amacı ise, hastaların tüm geçmiş hikayelerini değerlendiren dinamik tahmin yöntemleri kullanılarak ölüm risklerini yansız bir şekilde tahmin etmektir. Son aşamada ise dinamik tahmin yöntemleri ile elde edilen ölüm risklerinin tahmin başarılarının Brier skoru ile değerlendirilmesi amaçlanmıştır.

“Meme Kanserli Hastalarda Ölüm Riskinin Birleşik Zayıflık Modelleri ile İncelenmesi” başlıklı tez çalışması beş bölüm altında incelenmiştir.

“Giriş” adlı ilk bölümde hastalık ile ilgili temel bilgiler ve kanser araştırmalarında kullanılan istatistiksel yöntemlerin özellikleri ile birlikte tez çalışmasının amacına ait açıklamalar yer almaktadır.

İkinci bölüm olan genel bilgiler kısmında ise sağkalım çözümlemesi ile ilgili genel bilgilere, çözümlemede kullanılan fonksiyonlara ve bu fonksiyonların dağılımlarına ait bilgilere yer verilmektedir. Ardından hazard fonksiyonunu modelleme süreçleri anlatılarak, tezin ana konusu olan zayıflık modelleri ile ilgili teorik bilgiler ayrıntılar ile açıklanmaktadır. Tezin ikincil amacı olan ölüm risklerinin tahmininde kullanılan dinamik tahmin yöntemlerinin teorik özellikleri verilerek, yöntem detayları ile açıklanmaktadır. Genel bilgilerin son kısmında ise meme kanseri ile ilgili tanımlayıcı açıklamalara yer verilmektedir.

Üçüncü bölüm olan gereç ve yöntemler kısmında ise, ele alınan gerçek verinin tanımlamaları ve çalışma periyotundan bahsedilerek istatistiksel bazı açıklamalara da yer verilmektedir.

Dördüncü bölüm olan bulgular kısmını ise hedeflenen amaçlar doğrultusunda gerçekleştirilen tüm analizlerin sonuçları oluşturmaktadır.

Son bölümde ise, sonuçlar literatür bilgileri ile karşılaştırılarak tartışılmakta ve çalışmanın kısıtlılıkları belirtilmektedir.

2. GENEL BİLGİLER

2.1. Sağkalım Çözümlemesi ve Temel Kavramlar

Sağkalım analiz yöntemleri sağlık alanı başta olmak üzere ziraat, eğitim bilimleri, sosyal bilimler, pazar araştırmaları ve daha birçok alanda kullanılmaktadır. Örneğin, bir bitkinin bulunduğu ortamda ne kadar sürede ve nasıl bir ortamda yaşayabileceğinin belirlenmesinde, bir otomobilin lastiklerinin aşınma süresinin tahmininde, organ naklinin yapıldığı hastalarda 24 saat sonra organın red edilme durumunun değerlendirilmesinde, evli çiftlerin boşanma sürelerinin tahmininde, sektörde yer alan şirketlerin iflas etme durumlarının tahmininde veya hapisten çıkan suçluların tekrar suç işlemesine etki eden risk faktörlerinin araştırılması gibi birçok farklı alanda sağkalım analizlerinden yararlanılabilmektedir (İnceoğlu 2013; Kaya 2017).

Sağlık araştırmalarında, sağkalım çözümlemeleri özellikle kanser araştırmalarında sıklıkla kullanılmaktadır. Yeni bir tedavi yönteminin hastalar üzerine etkisinin araştırılmasında, hastalık seyrinin incelenmesinde, belirli bir hastalığa ait ortalama sağkalım süresinin tahmininde, iki veya daha fazla tedavi yönteminin sağkalım süresi üzerine etkilerinin araştırılmasında ve hastalığa etki eden risk faktörlerinin belirlenmesi gibi konularda sağkalım çözümlemelerinden sıklıkla yararlanılmaktadır. Ayrıca tedaviye ne zaman başlanması gerektiğine ve uygulanacak tedavinin süresinin tahmininde yine sağkalım analizleri yardımıyla karar verilebilmektedir (Dişçi 2015).

Gelişen tedavi yöntemleri ile birçok kanser türü erken evrelerde önlenebilmekte ve hastalar iyileşebilmektedir. İleri evrelerde hastalığın teşhisi ise bu durumu daha zorlaştıracaktır ve hastalara verilen tedavinin şekli ve yolu çok önemli bir seçim haline gelecektir. Örneğin, hastalığı evre 3'de olan karaciğerinde nöroendokrin tümörüne sahip bir hasta için kemoterapinin mi lutesyum tedavisinin mi sağkalım süresini uzatmada etkili olabileceği, hastaya en fazla fayda sağlayacak yöntemin hangisi olabileceğinin seçimi hastalık seyri açısından oldukça önemlidir. Bu durumda klinikçinin başvuracağı yöntem sağkalım analizi sonuçları olacaktır.

Hayatta kalma analizi olarak da bilinen sağkalım analizi, herhangi bir olayın (sonuç değişkeninin) meydana gelmesine kadar geçen sürede elde edilen verilerin analizi olarak tanımlanabilmektedir. Bu nedenle araştırma için belirli bir başlangıç ve sonlanma noktası tanımlanmaktadır. Kanser araştırmaları için sağkalım analizinde başlangıç zamanı; hastaların çalışmaya dâhil edildiği randomizasyon tarihi, tanı anı, ameliyat tarihi

veya tedavinin başlangıç tarihidir. Sonlanma olayı (terminal event) ise ölüm, hastalığın başlaması, nüks ve metastaz gibi olaylardır (Dişçi 2015).

Sağkalım sürelerinin analiz edilmesinde klasik istatistiksel yöntemlerin kullanılamamasının ana sebebi, verinin durdurulmuş (sansürlü) yapıda olmasıdır. Sansürlü veri ise temelde üç durumda ortaya çıkabilmektedir. İlk durum, çalışmanın sonlanmasına rağmen bireyin hala yaşıyor olmasıdır (sonuç olayın gözlenmemesi). Bu durumda kişinin sağkalım süresi çalışma süresinden daha uzun olmaktadır. İkinci durum ise, çalışma esnasında belirli sebepler ile gönüllüye ulaşılamaması veya bilgi alınamamasıdır (kişinin başka bir şehre taşınıp haber vermemesi gibi, lost to follow up). Üçüncü durum ise gönüllünün başka sebeplerden ötürü hayatını kaybetmesi veya çalışmadan çekilmesidir (withdraw) (Wienke 2011).

Genel olarak sağdan, soldan ve aralık sansürlü olmak üzere üç tip sansürlü veri tanımlanmaktadır. Sağdan sansürlü veriler, gönüllünün gerçek sağkalım zamanı tam olarak bilinmediğinde ortaya çıkmaktadır. Örneğin, kişinin çalışmadan çekilmesi, kaybolması veya çalışma bitmesine rağmen sonuç olayın gözlenmemesi gibi durumlarda sağdan sansür söz konusudur. Sağdan sansürlü veriler sağkalım analizlerinde oldukça sık görülmektedir. Soldan sansürlü veri, ilgilenilen olayın belli bir zamandan önce gerçekleştiğinin bilindiği fakat kesin olarak zamanının bilinmediği durumlarda ortaya çıkmaktadır. Bu tip verilerde, kişinin gözlenen sağkalım süresi, gerçek sağkalım süresinden daha uzundur. Örneğin, AIDS hastalığına yakalanmış, ancak hastalık başladığında tanı konulamamış bir hastaya, hastalığın etkileri görülmeye başlandığında test yapılırsa ve sonuç pozitif bulunursa hastanın gerçek sağkalım süresi hastalığa maruz kaldığı zamandan ölüme kadar geçen süre olarak tanımlanmaktadır. Buna karşın gözlenen sağkalım süresi test sonucunun pozitif çıkmasından ölümüne kadar geçen süredir. Aralıklı sansürlü veri, çalışma periyodu içerisinde ilgilenilen olayın iki zaman arasında gerçekleştiği bilindiğinde fakat kesin olarak hangi zamanda meydana geldiği bilinmediğinde kullanılan bir tanımlamadır. Kısaca ilgilenilen olayın tam gerçekleşme zamanı bilinmemekte ancak belirli bir aralıkta olayın gerçekleştiği bilinmektedir. Örneğin periyodik olarak izlenen metastaz durumları. Karaciğer kanseri bir hastanın operasyon sonrası 6 ayda bir takip edildiği bir durumda, hastanın ilk kontrolünün 6. ayda yapıldığı ve sağlık durumunun iyi olduğu, herhangi bir yeni oluşumun meydana gelmediği ve ikinci kontrole gittiğinde ise akciğerinde metastazı geliştiği varsayıldığında, metastazın ortaya

çıkış zamanı kesin olarak bilinemez ancak operasyon sonrası 6. ay ile 12. ay arasında geliştiği bilinmektedir. Bu veri tipi aralık sansürlü veri olarak tanımlanmaktadır. Benzer şekilde bir hastaya AIDS pozitif tanısı konulduğu an, hastalığın başlangıç anı olarak değerlendirilemez. Bunun yerine hastanın geçmiş kan sonuçları ve HIV testi sonucu dikkate alınarak en son negatif bulunan sonuç ile ilk pozitif bulunduğu sonuç arasında geçen zaman aralık sansürlü veriye örnek olarak verilebilmektedir (Kleinbaum ve Klein 2010).

2.1.1. Sağkalım Fonksiyonu ve Hazard Fonksiyonu

Sağkalım süresi, pozitif ve sıfır değerlerini alabilen rasgele bir değişken olmak üzere T olarak ifade edilmektedir. t değeri ise, T 'nin spesifik bir değeridir. Örneğin, kanser tedavisinden sonra hastanın 3 yıldan fazla yaşayıp yaşamadığı ile ilgilenildiği bir durumda $t=3$ olacaktır. Araştırılmak istenen hipotez ise $T > t=3$ olup olmadığıdır. Sağkalım analizlerinde kullanılan bir başka terim ise rasgele bir değişken olan gösterge değişkeni terimidir. Bu rasgele değişken, δ olarak ifade edilmekte ve başarısızlık veya sansür durumunu göstermektedir. δ , başarısızlık durumunda bir, sansür durumunda ise sıfır değerini almaktadır.

Sağkalım fonksiyonu, herhangi bir t noktasına kadar ki hayatta kalma olasılığını tanımlamaktadır. Sağkalım fonksiyonunun formülü aşağıdaki gibi ifade edilmektedir (Moore 2016).

$$S(t) = P(T > t), 0 < t < \infty \quad (1-1)$$

Bu fonksiyon $t=0$ anında bir değerini almakta, zaman ilerledikçe ise sağkalım süresi azalmaktadır, ancak hiçbir zaman sıfırın altına düşmemektedir. Sağkalım fonksiyonu genellikle hazard fonksiyonu açısından tanımlanmaktadır. Hazard fonksiyonu anlık başarısızlık oranıdır. Bu fonksiyon, yoğunluk fonksiyonu veya mortalite gücü olarak da bilinmektedir (Moore 2016).

Hazard fonksiyonu aşağıdaki formül yardımıyla elde edilmektedir (Moore, 2016).

$$h(t) = \lim_{\delta \rightarrow 0} \frac{P(t < T < t + \delta | T > t)}{\delta} \quad (1-2)$$

Hazard fonksiyonunun payı, bir olasılık değeri olup negatif değer almamaktadır. Paydası ise sıfır ile sonsuz arasında değer alabilmektedir. Sağkalım dağılımını tanımlamak için sağkalım ve hazard fonksiyonu dışında birikimli dağılım fonksiyonu da kullanılmaktadır. Birikimli dağılım fonksiyonu aşağıdaki eşitlik yardımıyla tanımlanmaktadır.

$$F(t)=P(T\leq t), 0<t<\infty \quad (1-3)$$

Bu fonksiyon sağkalım fonksiyonunun tamamlayıcısıdır. Sağkalım analizinde, bu fonksiyon birikimli risk fonksiyonu olarak da bilinmektedir. Birikimli risk fonksiyonundaki değişimin hızını gösteren olasılık yoğunluk fonksiyonu ise aşağıdaki gibi tanımlanmaktadır.

$$f(t)=-\frac{d}{dt}S(t)=\frac{d}{dt}F(t) \quad (1-4)$$

Hazard fonksiyonu, $h(t)=\frac{f(t)}{S(t)}$ olmak üzere, olasılık fonksiyonu ile ilişkilidir. Birikimli hazard fonksiyonu, $H(t)=\int_0^t h(u)du$ olmak üzere, sağkalım süresi ile hazard fonksiyonu arasında aşağıdaki gibi bir ilişki söz konusudur (Moore 2016).

$$S(t)=\exp(\int_0^t h(u)du) = \exp(-H(t)) \quad (1-5)$$

2.1.2. Parametrik Sağkalım Dağılımları

Sağkalım verisini modellemede kullanılan birçok sağkalım dağılımı vardır. Bunların en başında Üstel, Gamma, Weibull ve Lognormal gibi dağılımlar yer almaktadır (Hanagal 2011).

2.1.2.1. Üstel Dağılım

Üstel dağılım diğer dağılım modelleri arasında en basit olanı olmak üzere modelde yalnızca bir bilinmeyen parametre vardır. Hazard oranı $h(t)=\lambda$ olmak üzere,

birikimli dağılım fonksiyonu

$$F(t)=1-e^{-\lambda t}, \quad (1-6)$$

sağkalım fonksiyonu

$$S(t)=e^{-\lambda t}, \quad (1-7)$$

olasılık dağılım fonksiyonu

$$f(t)=\lambda e^{-\lambda t} \quad (1-8)$$

formülleri ile elde edilmektedir. Üstel dağılımın ortalaması $1/\lambda$, varyansı $\frac{1}{\lambda^2}$ 'dir.

Üstel dağılım sabit hazard oranına sahip tek dağılım olma özelliğini taşımaktadır (Hanagal 2011).

2.1.2.2. Weibull Dağılım

Weibull dağılımı iki parametrelidir. Dağılıma ait fonksiyon ve tanımlayıcı değerler aşağıdaki fonksiyonlar yardımıyla elde edilmektedir (Hanagal 2011).

Birikimli dağılım fonksiyonu,

$$F(t)=1-e^{-\left(\frac{t}{\lambda}\right)^\gamma}, 0 < t < \infty \quad (1-9)$$

sağkalım fonksiyonu

$$S(t) = e^{-\left(\frac{t}{\lambda}\right)^\gamma}, \quad (1-10)$$

olasılık dağılım fonksiyonu

$$f(t) = \frac{\gamma}{t} \left(\frac{t}{\lambda}\right)^{\gamma-1} e^{-\left(\frac{t}{\lambda}\right)^\gamma}, \quad (1-11)$$

Hazard oranı $\frac{\gamma}{t} \left(\frac{t}{\lambda}\right)^{\gamma-1}$, dağılımın ortalaması $\lambda \Gamma\left(1 + \frac{1}{\gamma}\right)$, dağılımın varyansı ise,

$$\lambda^2 \Gamma\left(1 + \frac{2}{\gamma}\right) - [\lambda \Gamma\left(1 + \frac{1}{\gamma}\right)]^2 \text{ dir.}$$

Yukarıdaki formüllerde geçen λ parametresi ölçüm, γ ise dağılım parametresi olarak tanımlanmakta ve Γ ise $\Gamma(N) = (N-1)!$ olmak üzere bir Gamma fonksiyonudur. $\gamma = 1$ olduğunda, zaman ilerledikçe hazard fonksiyonu sabit kalmaktadır ve Weibull dağılımı Üstel dağılım haline gelmektedir. $\gamma > 1$ durumunda, başarısızlık hızı artmakta iken tam tersi durumda yani $\gamma < 1$ durumunda ise başarısızlık hızı azalmaktadır. Dolayısı ile Weibull dağılımı popülasyondaki her üç durum içinde yani riskin artan, azalan ve sabit olduğu durumlarda kullanılabilmektedir. Riskin arttığı duruma (hazard oranı artan) örnek verilirse, bağırsak kanserine yakalanan hastalar, riskin azaldığı duruma ise beyin tümörü olan hastaların başarılı bir ameliyat ile tümörünün alınması örnek olarak verilebilir (Hanagal 2011).

2.1.2.3. Lognormal Dağılım

Lognormal dağılımda ölçek parametresi medyan ($T_{50} = e^\mu$), varyasyon parametresi ise varyanstır.

Logaritması alınmış başarısızlık süresinin ortalaması $\ln T_{50}$, standart sapması ise σ olmak üzere normal dağılım göstermektedir. Lognormal dağılıma ait fonksiyon özellikleri aşağıdaki eşitlikler yardımıyla elde edilmektedir (Hanagal 2011).

Birikimli dağılım fonksiyonu,

$$F(t) = \int_0^t \frac{1}{\sigma x \sqrt{2\pi}} e^{-\left(\frac{1}{2\sigma^2}\right)(\ln x - \ln T_{50})^2} dx \quad (1-12)$$

sağkalım fonksiyonu,

$$S(T) = 1 - F(t) \quad (1-13)$$

olasılık dağılım fonksiyonu,

$$f(t) = \frac{1}{\sigma t \sqrt{2\pi}} e^{-\left(\frac{1}{2\sigma^2}\right)(\ln t - \ln T_{50})^2}, 0 < t < \infty \quad (1-14)$$

hazard oranı (başarısızlık hızı),

$$h(t) = \frac{\left(\frac{1}{t\sigma\sqrt{2\pi}}\right) \exp\left[-\frac{(\ln t - \ln T_{50})^2}{2\sigma^2}\right]}{1 - G[(\ln t - \ln T_{50})/\sigma]} \quad (1-15)$$

olarak ifade edilmektedir.

Dağılımın ortalaması $T_{50}e^{\sigma^2/2}$, medyanı T_{50} , varyansı ise $T_{50}^2 e^{\sigma^2}(e^{\sigma^2} - 1)$ 'dir (Hanagal 2011).

2.1.2.4. Gamma Dağılım

Literatürde Gamma dağılımı için farklı ifade şekilleri mevcuttur. Ölçek ve varyasyon parametreleri için farklı yazarların farklı sembolleri kullandığı görülmektedir. Gamma dağılımında γ varyasyon parametresi, λ ise ölçek parametresi olarak tanımlandığında, olasılık dağılım fonksiyonu aşağıdaki gibi ifade edilmektedir (Hanagal 2011).

$$f(t, \gamma, \lambda) = \lambda^\gamma / (\Gamma(\gamma)) t^{\gamma-1} e^{(-\lambda t)}, 0 < t < \infty \quad (1-16)$$

Dağılımın ortalaması γ/λ , varyansı ise γ/λ^2 olarak belirtilmektedir.

2.1.2.5. Log-lojistik Dağılım

$\ln(T)$ lojistik dağılım gösterdiği durumlarda T sağkalım zamanı Log-lojistik dağılıma sahip olmaktadır. Log-lojistik dağılımı α ve γ olmak üzere iki parametre ile tanımlanmaktadır. Dağılıma ait olasılık dağılımı, sağkalım fonksiyonları ve hazard oranı aşağıdaki eşitlikler yardımıyla elde edilmektedir (Hanagal 2011).

Olasılık dağılım fonksiyonu,

$$\frac{\alpha \gamma t^{\gamma-1}}{(1+\alpha t^\gamma)^2}, 0 < t < \infty \quad (1-17)$$

sağkalım fonksiyonu,

$$\frac{1}{1+\alpha t^\gamma} \quad (1-18)$$

hazard oranı,

$$\frac{\alpha \gamma t^{\gamma-1}}{1+\alpha t^\gamma} \quad (1-19)$$

eşitlikleri ile elde edilmektedir.

2.1.2.6. Gompertz Dağılım

Gompertz dağılımı sağkalım analizlerinde kullanılan bir başka dağılım çeşididir. Dağılım Üstel dağılımının bir türüdür. Dağılımda, hazard oranı zamana dayalı bir üstel fonksiyona bağlı olarak değişmektedir.

Gompertz dağılımı λ ve γ parametreleri ile tanımlanmaktadır. Hazard fonksiyonu

Gompert dağılımı için aşağıdaki gibi ifade edilmektedir.

$$h(t)=exp(\lambda + \gamma t) \quad (1-20)$$

$\gamma=0$ olduğunda, $h(t)$ fonksiyonu e^λ sabitine eşit olmaktadır.

Sağkalım fonksiyonu ise;

$$S(t)=exp[-\frac{e^\lambda}{\gamma}(e^{\gamma t} - 1)] \quad (1-21)$$

olasılık yoğunluk fonksiyonu;

$$f(t)=exp[(\lambda + \gamma t) - \frac{1}{\gamma}(e^{\lambda + \gamma t} - e^\lambda)] \quad (1-22)$$

olarak ifade edilmektedir.

$G(x)=e^x \int_x^\infty y^{-1} e^{-y} dy$ olmak üzere, Gompertz dağılımının ortalaması

$G(e^\lambda/\gamma)/e^\lambda$ dir.

Gompertz dağılımı, 30-80 yaşları arasındaki bireylerde ölüm durumunu tanımlamaktadır. İleri yaşlardaki bireylerin ölme hızları ölüm yasasına göre tahmin edildiği gibi hızlı artmamaktadır, dolayısı ile ileri yaşlardaki bireylerin ölüm hızlarının yavaşlaması gibi bir durum söz konusudur ki bu durumda zayıflık modellerinin geliştirilmesindeki başlangıç noktalarından biridir (Wienke 2011).

2.1.2.7. Pareto Dağılım

Pareto dağılımı 19. yüzyıl sonlarında popülasyonun gelir dağılımını açıklamak amacıyla Vilfredo Pareto tarafından 1987 yılında kullanılmıştır. Dağılımın şekli çarpıktır ve $\omega > 0$ ve $\zeta > 0$ olmak üzere iki parametresi vardır. Dağılımın hazard fonksiyonu ise, monoton olarak azalmaktadır.

Pareto dağılımının olasılık yoğunluk fonksiyonu,

$$f(t)=\frac{\zeta}{\omega} \left(\frac{\omega}{\omega+t}\right)^{\zeta+1} \quad (1-23)$$

sağkalım fonksiyonu,

$$S(t)=\left(\frac{\omega}{\omega+t}\right)^{\zeta} \quad (1-24)$$

hazard oranı ise

$$h(t)=\frac{\zeta}{\omega+t} \quad (1-25)$$

eşitliği yardımıyla elde edilmektedir.

Pareto dağılımı yaşam sürelerinin Üstel dağılımının Gamma karması olarak da tanımlanabilmektedir. Genellikle popülasyonların dağılımlarının araştırılmasında, sosyo-ekonomik çalışmalarda ve petrol sahası alanlarında kullanılmaktadır (Wienke 2011).

2.1.3. Sağkalım ve Hazard Fonksiyonlarının Tahmini

Tahminlerde parametrik yöntem kullanılması durumunda, başarısızlık zamanlarının dağılımı hakkında bir takım varsayımlar kullanılmaktadır. Bu durum hastalığın ve yaşın doğal sürecinin bilindiği durumlarda olanaklıdır. Ancak bazı durumlarda olanaklı değildir. Bu durumda parametrik olmayan modeller kullanılmaktadır. En basit parametrik olmayan tahminin dağılımı ise yapay dağılım fonksiyonudur. Bu fonksiyon, dağılımın sürekli olması durumunda da tahminlerin kesikli olması anlamına gelmektedir. Sansürlü veri içeren sağkalım analizinde uygun parametrik olmayan veya yarı parametrik yöntemlerin geliştirilmesinde en önemli adım Kaplan Meier tahmincisidir ve bu tahmincinin uzantılarıdır (Wienke 2011).

Parametrik olmayan modeller, parametrik modellere göre daha esnektir ve hangi olasılık dağılımı olursa olsun çözüm gerçekleştirebilmektedir. Ancak parametrik olmayan modellerin güvenilir sonuçlar üretebilmesi için daha çok veriye ihtiyaç vardır ve bu modellerde hazard fonksiyonu tahminlerinin elde edilmesi oldukça zordur. Parametrik modellerde ise, seçilen modele bağlı olarak hazard ve sağkalım fonksiyonları kapalı-form tanımlamalarına imkan sağlamaktadır. Parametrik modeller birkaç parametre değeri ile tanımlanmaktadır. Küçük örneklem genişliklerinde bile oldukça iyi sonuçlar vermektedir. Varsayılan model doğru ise, parametrik modellerin tahmini parametrik olmayan modellerin tahminine göre daha güçlüdür (Wienke 2011).

2.1.3.1. Kaplan Meier Tahmincisi

Gruplardaki bireylerin sağkalımlarının tanımlanabilmesi için yapay sağkalım fonksiyonlarının hesaplanması ve grafiklendirilmesi kullanışlı bir yaklaşımdır. Örneklemde sansürlü veri olmadığı durumlarda, t anındaki yapay sağkalım fonksiyonu, t anındaki sağkalanların örneklem genişliğine oranlanması ile elde edilmektedir. Her bir gözlemlenen başarısızlığın ardından bu fonksiyon $1/n$ 'e kadar azalan eğilim göstermektedir. Ancak pratikte bazı gözlemler sansürlüdür. Bu durumu kolaylık ile değerlendirmek amacıyla bazı metodolojiler gereklidir. $T_i = \min\{T_i^*, C_i\}$ ve $\Delta_i = 1$ ($T_i^* \leq C_i$) durumlarında $(T_1, \Delta_1), \dots, (T_n, \Delta_n)$ çiftleri gözlemlenmektedir. $T_1, T_2, \dots, T_n$ 'lerin sıralanmış istatistikleri $T_{(1)} < T_{(2)} < \dots < T_{(n)}$ olmak üzere $T_{(i)}$ ile ilişkilendirilerek $\Delta_{(i)}$, Δ 'nın değeri olarak tanımlanmaktadır. Bu ifade, $T_{(i)} = T_j$ durumunda, $\Delta_{(i)} = \Delta_j$ eşit olduğunu anlamına gelmektedir. Ancak $\Delta_{(1)}, \Delta_{(2)}, \dots, \Delta_{(n)}$ sıralı değildir (Wienke 2011).

Kaplan Meier tahmincisi aşağıdaki formül yardımıyla elde edilmektedir.

$$\hat{S}(t) = \prod_{i: t_{(i)} < t} \left(1 - \frac{\delta_{(i)}}{n - i + 1}\right) \quad (1-26)$$

Pratikte, ölçümlerin hassas yapılmaması nedeniyle tekrar eden değerler ortaya çıkmaktadır. Bu tekrarların üstesinden gelmek için, $r \leq n$ olmak üzere t_i , r farklı gözlemlenme zamanlarını ($t_1 < t_2 < \dots < t_r$) belirtmektedir. Tekrarları dikkate alan Kaplan Meier tahmincisi ise aşağıdaki formül yardımıyla elde edilmektedir.

$$\hat{S}(t) = \prod_{i \in R(t)} \left(1 - \frac{d_i}{\#R(t_i)}\right) \quad (1-27)$$

d_i , t_i zamanındaki olayların sayısıdır. $R(t)$, t zamanındaki risk altındaki tüm bireylerin indeks setidir. $\#R(t)$, t zamanında risk setindeki gözlemlerin sayısıdır. Kaplan Meier tahmincisi, sadece bir olay sırasında değişen azalan aşama fonksiyonudur. En son gözlem sansürlü olduğunda, $\hat{S}$ en uzun gözlem zamanından sonra tanımlanamamaktadır. Bu durumda ortalama sağkalım süresi tahmin edilememektedir. Bu sorunu çözenin yolu ise, sağkalım fonksiyonunun en uzun zamandan sonra sıfır olduğunu varsaymaktır ki bu durum da tahminler yanlı olacaktır. Dolayısı ile en iyi çözüm yolu, medyanı sağkalım zamanı olarak kullanmaktır (Wienke 2011).

Kaplan Meier tahmininin varyansı ise,

$$V\left(\hat{M}(t)\right) = \widehat{S^2(t)} \sum_{i \in R(t)} \frac{d_i}{\#R(t_i) - d_i} \quad (1-28)$$

eşitliği yardımıyla elde edilmektedir.

2.1.3.2. Nelson-Aalen Tahmincisi

Nelson-Aalen tahmincisi Nelson tarafından 1972 yılında önerilmiş ve 1978 yılında Aalen tarafından yeniden keşfedilmiştir. Bu tahmin edici birikimli hazard fonksiyonu $M(t)$ 'nin bir tahmincisidir ve aşağıdaki gibi formülize edilmiştir.

$$\hat{M}(t) = \sum_{i \in R(t)} \frac{d_i}{\#R(t_i)} \quad (1-29)$$

Formülde geçen simgelerin tanımlanması Kaplan Meier tahmin edicisi ile benzerdir. Eşitlikten de anlaşılacağı üzere, her bir ayrı ölüm zamanında hazard oranının tahmin edilebilmesi için, ölüm sayılarının risk altındaki sayıya oranlanması gerekmektedir. Birikimli hazard değeri, t zamanına kadar ki ölüm zamanlarının hazard toplamlarıdır. Nelson-Aalen tahmininin varyansı ise,

$$V(\hat{M}(t)) = \sum_{i \in R(t)} \frac{d_i}{(\#R(t_i))^2} \quad (1-30)$$

olarak ifade edilmektedir.

Kaplan Meier ve Nelson Aalen tahmin edicileri asimptotik olarak eşittir ve özellikle risk altındaki gözlem sayısı ölüm sayısından küçük olduğu durumlarda sonuçlar birbirlerine oldukça yakın elde edilmektedir (Wienke 2011).

2.1.4. Hazard Fonksiyonunu Modelleme

Sağkalım verilerini modelleme yaklaşımında, bireylerin veya grupların sağkalımlarının bağımsız değişken/lerden nasıl etkilendiğine açıklık getirilmektedir. Bunun için de hazard fonksiyonu kullanılmaktadır. Model katsayılarının tahmin süreci ve elde edilen sonuçlar doğrusal modellerden farklılık göstermektedir. Ancak model kurulurken doğrusal modellerin bazı temel prensip ve prosedürlerinden yararlanılmaktadır (Collett 2003).

Sağkalım verilerini modelleme de; modele alınan açıklayıcı değişkenlerin hazard fonksiyonuna etkisini belirlemek, bireyler için hazard fonksiyonu tahminini elde etmek ve dolaylı olarak sağkalım fonksiyonunu tahmin etmek amaçlanmaktadır.

Sağkalım verilerini modellemede orantılı hazard modeli kullanılmaktadır. Bu model Cox tarafından 1972 yılında geliştirilmiştir. Model aynı zamanda Cox regresyon modeli olarak da bilinmektedir. Model orantılı hazard varsayımına dayanmaktadır. Sağkalım zamanları için özel bir dağılım varsayımı bulunmamaktadır. Dolayısı ile yarı parametrik bir yöntem olarak bilinmektedir.

2.1.4.1. Genelleştirilmiş Orantılı Hazard Modeli (Cox Regresyon Modeli)

Bireylerin herhangi bir zamandaki ölüm riski, p tane $X_1, X_2, \dots, X_p$ açıklayıcı değişkenin $x_1, x_2, \dots, x_p$ değerlerine bağlıdır. Çalışmanın başlangıcında bu değerlerin kaydedildiği varsayılmaktadır. Orantılı hazard modelinde açıklayıcı değişkenlerin tüm seti $x = (x_1, x_2, \dots, x_p)'$ vektörü ile tanımlanmaktadır. $h_0(t)$ ifadesi, bir bireyin x vektörünün sıfır olduğu durumda bireylerin hazard fonksiyonunu ifade etmektedir. i . bireyin hazard fonksiyonu aşağıdaki gibi ifade edilmektedir (Collett 2003).

$$h_i(t) = \psi(x_i) h_0(t) \quad (1-31)$$

Yukarıdaki formülde geçen $\psi(x_i)$ ifadesi, i . birey için açıklayıcı değişkenlerin vektör değerlerinin bir fonksiyonudur. Rölatif risk (relative hazard) ($\psi(x_i)$), negatif değer alamamaktadır. Dolayısı ile bu ifadeyi η_i , x_i 'deki p açıklayıcı değişkenin doğrusal kombinasyonu olmak üzere $\exp(\eta_i)$ olarak yazmak daha uygun olmaktadır. Böylece

$\eta_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi}$ olarak gösterilmektedir. Matris notasyonunda $\eta_i = \beta' x_i$ olarak gösterilmektedir. β , x_i açıklayıcı değişkenlerinin katsayılarına ait vektördür. η_i değeri, modelin doğrusal bileşeni olarak adlandırılmakta, aynı zamanda risk skoru veya prognostik indeks olarak da bilinmektedir. Böylece genelleştirilmiş orantılı hazard modeli aşağıdaki gibi tanımlanmaktadır (Collett 2003).

$$h_i(t) = \exp(\beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi}) h_0(t) \quad (1-32)$$

Orantılı hazard modeli, aynı zamanda hazard oranlarının logaritmasının doğrusal modellenmesi olarak da aşağıdaki gibi ifade edilmektedir.

$$\log \left\{ \frac{h_i(t)}{h_0(t)} \right\} = \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi} \quad (1-33)$$

Cox regresyon modelinde, hazard fonksiyonunu etkileyen kovaryat ve faktör olmak üzere iki farklı değişken türü vardır. Faktör değişkeni kategorik yapıda iken, kovaryat değişkeni niceliksel yapıdadır. Bu değişkenlerin modele ekleniş şekilleri farklılık göstermektedir.

2.1.4.2. Modele Kovaryat Değişkeninin Eklenmesi

Kovaryatlar, Cox regresyon modeline β katsayıları ile dâhil edilmektedir. Modelde her bir kovaryat değişkeni için bir β katsayısı eklenmektedir. Örneğin, hazard fonksiyonunda X_1 ve X_2 olmak üzere iki kovaryat olduğu düşünülürse, bu kovaryatların

i . birey için değerleri sırasıyla x_{1i} ve x_{2i} olarak gösterildiğinde, i . birey için Cox regresyon modeli aşağıdaki eşitlikteki gibi ifade edilmektedir (Collett 2003).

$$h_i(t) = \exp(\beta_1 x_{1i} + \beta_2 x_{2i}) h_0(t) \quad (1-34)$$

$h_0(t)$, modelde yer alan tüm kovaryat değerlerinin sıfır olarak alındığı hazard fonksiyonudur.

2.1.4.3. Modele Faktör Değişkeninin Eklenmesi

Modeli etkileyeceği düşünülen α seviyeli A faktörünün olduğu varsayıldığında, A faktörünün j . seviyesinin modele dâhil edilebilmesi için, α_j teriminin hazard fonksiyonuna eklenmesi gerekmektedir. $\alpha_1, \alpha_2, \dots, \alpha_\alpha$ terimleri, A faktörünün ana etkileridir.

Cox regresyon modeline göre A faktörünün j . seviyesindeki hazard fonksiyonu $\exp(\alpha_j) h_0(t)$ 'dir. Fonksiyonda yer alan $h_0(t)$, açıklayıcı değişkenlerin hepsinin sıfıra eşit olduğu temel hazard fonksiyonudur. Bilgisayar programlarında modelin çözümünün gerçekleştirilmesi için faktörler modele kukla değişken olarak dâhil edilmektedir. A faktörünün j . seviyesi ($j=2,3,\dots,\alpha$) için X_j 'nin değerinin x_j olduğu durumda α_j terimi, $\alpha_2 x_2 + \alpha_3 x_3 + \dots + \alpha_\alpha x_\alpha$ şeklinde ifade edilmektedir. Dolayısı ile A faktörünün ana etkisi ile ilişkili $\alpha-1$ parametre vardır. Bir başka ifadeyle A faktörünün serbestlik derecesi $\alpha - 1$ 'dir (Collett 2003).

2.1.4.4. Cox Modeline Zamana Bağımlı Kovaryat Değişken Eklenmesi

Model oluşturulurken, tahmin edicilerin ve sağkalım zamanı fonksiyonunun etkileşimleri kullanılarak zamana bağımlı kovaryat değişkeni eklenebilmektedir. Örneğin, X_j tahmin edicisi için modelde zamana bağımlı kovaryat değişkeni $X_j(t)$ ile ifade edildiğinde, $X_j(t) = X_j g(t)$ 'dir. Burada $g(t)$, zamanın bir fonksiyonudur (Collett 2003).

X_j için kovaryatlara göre düzeltilmiş, orantılı hazard varsayımının değerlendirilmesinde kullanılan model aşağıdaki gibi tanımlanmaktadır (Collett 2003).

$$h(t, x(t)) = h_0(t) \exp[\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_j x_j + \dots + \beta_p x_p + \delta x_j * g(t)] \quad (1-35)$$

2.1.4.5. Modele Etkileşim Etkisinin Eklenmesi

Faktör sayısı birden fazla olduğu durumlarda, iki veya daha fazla faktörün seviyelerinin kombinasyon etkileri modele dahil edilebilmektedir. Bu etkiler etkileşim etkisi olarak tanımlanmaktadır. Örneğin cinsiyet ve tümör evresi gibi iki faktör olduğu durumda, tümör evresinin ölüm üzerine etkisi cinsiyetlere göre farklılık gösteriyorsa, cinsiyet ve tümör evreleri arasında anlamlı bir etkileşim vardır ve bu etki modele dahil edilerek incelenmelidir (Collett 2003).

A ve B faktörlerinin kombinasyon seviyelerinin hazard fonksiyonunu etkilediği varsayıldığı durumda, faktörlerin seviyeleri α ve b olarak tanımlanmakta iken, $(\alpha b)_{jk}$ ise A ve B faktörlerinin etkileşimini ifade etmektedir ($j=1,2,\dots,\alpha, k=1,2,\dots,k$).

İstatistiksel modellemede, etkileşim etkisinin varlığı ana etkilerin yanında etkileşim teriminin de modele eklenmesi ile araştırılmaktadır. Modele ana etkiler ile ilişkili olarak kukla değişken oluşturulmaktadır. Örneğin, A ve B faktörlerinin sırasıyla seviyeleri 2 ve 3 olduğu durumda, U_2, V_2 ve V_3 kukla değişkenleri Tablo 2-1'deki gibi tanımlanmaktadır (Collett 2003).

Table 2-1 : A ve B faktörlerinin seviyeleri

A Faktörü		B faktörü		
A Faktörünün seviyesi	U_2	B Faktörünün seviyesi	V_2	V_3
1	0	1	0	0
2	1	2	1	0
		3	0	1

$j=2; k=2,3$ olduğunda U_j ve V_k 'nin değerleri u_j ve v_k olmak üzere, $(\alpha\beta)_{jk} = (\alpha\beta)_{22}u_2v_2 + (\alpha\beta)_{23}u_2v_3$ şeklinde tanımlanmaktadır.

Dolayısı ile A ve B arasındaki etkileşim ile ilişkili olan iki parametre vardır. Genel olarak A ve B faktörlerinin α ve b seviyeleri olduğunda, AB etkileşimi $(\alpha - 1)(b - 1)$ parametresi ile ilişkilidir. Bir başka ifadeyle AB etkileşiminin serbestlik derecesi $(\alpha - 1)(b - 1)$ 'dir (Collett 2003).

2.1.4.6. Cox Regresyon Modelinin Tahmini

Cox regresyon modelini tahmin etmek için, $X_1, X_2, \dots, X_p$ bağımsız değişkenlerinin bilinmeyen katsayıları ve $h_0(t)$ fonksiyonunun tahmin edilmesi gerekmektedir. Bu amaçla ilk olarak katsayılar tahmin edilmektedir ve bu katsayılar kullanılarak da $h_0(t)$ fonksiyonunu tahmin edilmektedir (Collett 2003).

Cox modelindeki β katsayıları maksimum olabilirlik yöntemi kullanılarak tahmin edilmektedir. Yöntemin uygulanabilmesi için ilk olarak örnek verinin olabilirliği elde edilmektedir. Bu olabilirlik gözlenen verinin birlikte olasılığıdır. Maksimum olabilirlik tahmincileri, olabilirlik fonksiyonunu maksimize eden değerlerdir.

Gözlem sayısı n olan bir veri seti için, r ölüm zamanlarını, $n-r$ sağdan sansür zamanlarını ifade etmek üzere, sıralanmış r ölüm zamanları $t_{(1)} < t_{(2)} < \dots < t_{(r)}$ olarak tanımlanmaktadır. Dolayısı ile j . sıralanmış ölüm zamanı $t_{(j)}$ ile ifade edilmektedir. $t_{(j)}$ zamanında risk altındaki bireyler $R(t_{(j)})$, $t_{(j)}$ 'den önceki zamanlardaki hayatta kalan ve sansürlü olmayan bireylerdir (Collett 2003).

Cox regresyon modeli için olabilirlik fonksiyonu aşağıdaki gibi ifade edilmektedir.

$$L(\beta) = \prod_{j=1}^r \frac{\exp(\beta' x_{(j)})}{\sum_{l \in R(t_{(j)})} \exp(\beta' x_l)} \quad (1-36)$$

Yukarıdaki formüldeki $x_{(j)}$ terimi, $t_{(j)}$ zamanında ölen bireylerin kovaryat vektörüdür. Olabilirlik fonksiyonunun payda toplamı ise, $t_{(j)}$ zamanında risk altındaki tüm bireylerin $\exp(\beta' x)$ değerleri toplamıdır. Sağkalım zamanları sansürlü olan gözlemler olabilirlik fonksiyonunun payına dâhil edilmemekte ancak sansür zamanından önce meydana gelen ölüm zamanlarındaki risk setlerinin toplamına dâhil edilmektedir. Olabilirlik fonksiyonu ölüm zamanlarının sırasına bağlıdır. Dolayısıyla bağımsız değişkenlerin hazard fonksiyonu üzerine etkisinin tahmini, ölüm zamanlarının sırasına dayanmaktadır (Collett 2003).

Verinin gözlemlenen sağkalım süreleri $t_1, t_2, \dots, t_n$ olmak üzere, δ_i olayı gösteren gösterge değişkenidir. i . sağkalım zamanı sağdan sansürlü olduğunda, δ_i sıfır değerini almaktadır ($i=1,2,\dots,n$).

Olabilirlik fonksiyonu aşağıdaki eşitlik yardımıyla da ifade edilmektedir.

$$\prod_{i=1}^n \left\{ \frac{\exp(\beta' x_i)}{\sum_{l \in R(t_i)} \exp(\beta' x_l)} \right\}^{\delta_i} \quad (1-37)$$

Log-olabilirlik fonksiyonu ise,

$$\log L(\beta) = \sum_{i=1}^n \delta_i \{ \beta' x_i - \log \sum_{l \in R(t_i)} \exp(\beta' x_l) \} \text{ dir (Collett 2003).}$$

Cox regresyon modelindeki β katsayılarının maksimum olabilirlik tahmini, log olabilirlik fonksiyonunun Newton-Raphson yöntemi yardımıyla maksimize edilmesi ile elde edilmektedir.

Newton Raphson yöntemi sansürlü sağkalım veri modellerinde olabilirlik fonksiyonunu maksimize etme süreci için kullanılmaktadır. β parametrelerine dayanılarak, log-olabilirlik fonksiyonunun ilk türevi $p \times 1$ boyutlu $u(\beta)$ vektörü etkili skor vektörü olarak bilinmektedir. Log olabilirliğin ikinci türevinin negatifi ise $p \times p$ boyutlu $I(\beta)$ matrisidir ve $-\frac{\partial^2 \log L(\beta)}{\partial \beta_j \partial \beta_k}$ olarak ifade edilmektedir. $I(\beta)$ gözlenen bilgi matrisi olarak bilinmektedir (Collett 2003).

Newton Raphson yöntemine göre iteratif sürecin $(s+1)$. döngüsündeki β parametre vektörünün tahmini,

$$\widehat{\beta}_{s+1} = \widehat{\beta}_s + I^{-1}(\widehat{\beta}_s) u(\widehat{\beta}_s) \quad (1-38)$$

eşitliği yardımıyla elde edilmektedir.

Formülde yer alan $u(\hat{\beta}_s)$ terimi etkili skor vektörüdür, $I^{-1}(\hat{\beta}_s)$ ise bilgi matrisinin tersidir. Süreç $\hat{\beta}_0 = 0$ ile başlar iken, log olabilirlik fonksiyonunda çok küçük değişiklikler olana kadar iteratif süreç devam etmektedir (Collett 2003).

2.1.4.7. β Katsayılarının Hipotez Testi ve Güven Aralıkları

Model sonuçları verildiğinde, parametre tahminleri standart hata değerleri ile birlikte verilmektedir. Standart hatalar bilinmeyen beta parametreleri için yaklaşık güven aralıklarının bulunması amacıyla kullanılmaktadır. Pratikte, $100(1-\alpha)\%$ güven düzeyinde, bilinmeyen β parametresinin güven aralıkları $\hat{\beta} \pm z_{\alpha/2} * \text{standart hata}(\hat{\beta})$ 'dır. β parametresinin güven aralıkları sıfırı içermediğinde, β parametresinin sıfıra eşit olmadığı hipotezi kabul edilmektedir. $\beta=0$ 'a eşit olan yokluk hipotezi ($\hat{\beta} / \text{standart hata}(\hat{\beta})$) hesaplanması ile test edilmektedir. Bu istatistiğin karesi, bir serbestlik dereceli ki-kare dağılımlı Wald testi ile karşılaştırılmaktadır (Collett 2003).

Cox regresyon modelinde β parametrelerinin tahminleri birbirinden bağımsız değildir ve parametrelerin yorumları kolay değildir. Örneğin, X_1, X_2 ve X_3 bağımsız değişkenlerinin olduğu bir modelde standart hatalar ile karşılaştırılarak $\hat{\beta}_1$ ve $\hat{\beta}_2$ parametrelerinin, test sonucunda sıfıra eşit olduğu görüldüğü zaman bu sonuç modelde sadece X_3 değişkeninin yer alması gerektiği anlamına gelmemektedir. Örneğin, X_1 değişkeni modelden çıkartıldığında, X_2 değişkeninin katsayı değeri büyük bir değişim gösterebilmektedir. Bu değişim özellikle bu iki değişken arasında yüksek korelasyon olduğunda mevcut olmaktadır.

2.1.4.8. Hazard Oranlarının Standart Hatası ve Güven Aralıkları

İki grupta sağkalım sürelerine ait bir veri seti düşünüldüğünde, tek bir gruptaki β parametresi diğer grup ile ilişkili olarak, t zamanındaki ölüme ait hazard oranının logaritmasıdır. Dolayısı ile hazard oranı $\psi = e^{\beta}$ 'dir. Hazard oranının tahmini,

$\hat{\psi} = \exp(\hat{\beta})$, $\hat{\psi}$ 'nin varyansı ise $\{\exp(\hat{\beta})\}^2 \text{var}(\hat{\beta})$ 'dır (Collett 2003).

2.1.5. Cox Regresyon Analizinin Varsayımları ve Orantılı Hazard Varsayımı

Cox regresyon modelinin kullanılabilmesi için gerekli varsayımlar; gözlemlerin bağımsızlığı, bağımsız değişkenlerin hazard fonksiyonu üzerine etkilerinin loglineer olması ve bu loglineer fonksiyon ile risk fonksiyonu arasındaki ilişkinin çarpımsal olmasıdır (Yay ve ark. 2007). En temel varsayımı ise, hazard oranının zamana bağlı

olarak değişmemesi olup bu varsayım “orantılı hazard” olarak adlandırılmaktadır. Orantılı hazard, bir bireyin hazard fonksiyonunun diğer bireyin hazard fonksiyonu ile orantılı olmasıdır ve hazard oranının zaman boyunca değişmemesidir. Bu varsayımın kontrolünde çeşitli yöntemler kullanılmaktadır. Aşağıda bu yöntemlerin açıklamaları yer almaktadır.

2.1.5.1. Grafiksel Yöntemler

Cox orantılı hazard sağkalım fonksiyonu, hazard ve sağkalım fonksiyonu ile ilişkili olarak aşağıdaki gibi tanımlanmaktadır (Qi 2009).

$$S(t, x) = S_0(t)^{\exp(\sum_{i=1}^p \beta_i x_i)} \quad (1-39)$$

Yukarıdaki eşitlikte yer alan $x = (x_1, x_2, \dots, x_p)'$ bireyin açıklayıcı değişkenlerine ait değerlerinin vektörüdür. $S(t, x)$ eşitliğinin iki kere logaritması alındığında,

$$\ln[-\ln S(t, x)] = \sum_{i=1}^p \beta_i x_i + \ln[-\ln S_0(t)] \quad (1-40)$$

elde edilmektedir.

Sırasıyla $x_1 = (x_{11}, x_{12}, \dots, x_{1p})$ ve $x_2 = (x_{21}, x_{22}, \dots, x_{2p})$ değişken değerlerine sahip iki gözlemin log-log eğrilerinin farkı ise,

$$\ln[-\ln S(t, x_1)] - \ln[-\ln S(t, x_2)] = \sum_{i=1}^p \beta_i (x_{1i} - x_{2i}) \quad (1-41)$$

'dir (Qi 2009).

Bu eşitlik orantılı hazard varsayımının değerlendirilmesinde oldukça önemlidir. Hazard oranları orantılı olduğu durumda, $\log(-\log(\text{sağkalım}))$ 'a karşı çizilen sağkalım süresi eğrisi gruplar için paralellik göstermektedir. Ancak bu yöntem çok sayıda kategori içeren kategorik değişkenler ve sürekli değişkenler için kullanışlı değildir (Qi 2009).

Log(-log(sağkalım)) grafiğine direk bakarak bu varsayımın sağlanıp sağlanmadığı hakkında kesin yargıya varmak güçtür. Ancak yöntem ön değerlendirme için oldukça kullanışlı bilgiler vermektedir.

2.1.5.2. Cox Orantılı Hazard Modeline Zamana Bağımlı Kovaryat Eklenmesi

Tahmin edicilerin etkileşimleri ve sağkalım süresi fonksiyonu ile zamana bağlı kovaryat oluşturulur ve oluşturulan bu kovaryat modele eklenir. Örneğin, X_j için zamana bağlı kovaryat $X_j(t) = X_j * g(t)$ ile elde edilir. $g(t)$; $\log t$, t veya Heaviside fonksiyonuna bağlı zamanın bir fonksiyonudur. Bu durumda orantılı hazard varsayımının değerlendirilmesinde kullanılan model aşağıdaki gibidir (Qi 2009).

$$h(t, x(t)) = h_0(t) \exp[\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_j x_j + \dots + \beta_p x_p + \delta x_j * g(t)] \quad (1-42)$$

Yokluk hipotezi $\delta = 0$ şeklinde olmak üzere, test istatistiği Wald veya olabilirlik oran testleri ile elde edilmektedir. Wald testi için test istatistiği, $W = \left(\frac{\hat{\delta}}{se(\hat{\delta})}\right)^2$ 'dir. Yokluk hipotezi altında hesaplanan olabilirlik oranı L_0 , alternatif hipotez için ise olabilirlik değeri L_α olmak üzere $LR = -2 \ln \left(\frac{L_0}{L_\alpha}\right)$ 'dır. Her iki test istatistiği de bir serbestlik dereceli ki-kare dağılımı göstermektedir. Zamana bağlı kovaryat değişkeninin anlamlı bulunduğu durumlarda, yokluk hipotezi red edilmektedir. Bir başka ifade ile orantılı hazard varsayımı sağlanmadığı anlamına gelmektedir (Qi 2009).

2.1.5.3. Schoenfeld Artıkları ile Test Etme

Orantılı hazard varsayımının incelenmesinde kullanılan diğer istatistiksel test Schoenfeld artığına dayanmaktadır. Başarısız olduğu gözlemlenen her birey için Schoenfeld artıkları tanımlanmaktadır. Orantılı hazard varsayımının sağlandığı durumlarda, kovaryat için elde edilen Schoenfeld artıkları sağkalım zamanı ile ilişkili değildir. Dolayısıyla bu test, belirli bir kovaryat değişken için Schoenfeld artıkları ile bireysel hayatta kalma sürelerinin sıralaması arasındaki korelasyonun bulunmasıyla elde edilmektedir (Qi 2009).

2.1.6. Cox Regresyon Modelinin Artıkları

Modelin oluşturulmasının ardından, bu modelin yeterliliğinin değerlendirilmesi gerekmektedir. Modelin değerlendirilmesi artıklar yardımıyla gerçekleştirilmektedir. Sağkalım analizinde, özellikle Cox orantılı hazard modelinde değerlendirilen artıklar farklı amaçlar için kullanılmaktadır. Artıklar, çalışmadaki her bir birey için hesaplanmaktadır. Cox regresyon modelinin değerlendirilmesinde Martingale, Sapma, Schoenfeld artıkları ve Skor olmak üzere başlıca dört artık kullanılmaktadır. Bu artıklardan Schoenfeld artığı kovaryat için orantılı hazard varsayımını test etmede kullanılırken, Martingale artığı Cox modelinin uygunluğunu test etmek için kullanılır ve sapma artığı ise tahmin edilen gözlemler içinden tahmin gücü zayıf olan gözlemleri belirlemede yararlanılırken, Skor artığı etkili gözlemlerin belirlenmesinde kullanılmaktadır. Schoenfeld artığı, orantılı hazard varsayımının değerlendirilmesinde kullanılan Schoenfeld tarafından önerilen bir artıktır. Artıkların zamana bağlılığını test etmek için i .artığın t_i 'ye karşı grafiği çizilir. Değerlendirme sürecinde gözlemlenen X_i değeri ve R_i riskinde koşullu beklenen değeri arasındaki farkın kısmi artığı kullanılmaktadır (Therneau ve Grambsch 2000; Fitrianto 2013).

Schoenfeld artık değeri;

$$\widehat{r}_{ik} = X_{ik} - \widehat{E}(X_{ik}|R_i) \quad (1-43)$$

eşitliği ile elde edilmektedir. Artık değeri gözlenen yaşam süresine bağlı olarak hesaplanmamaktadır. Ayrıca hesaplamalarda birikimli hazard fonksiyonunun tahminine ihtiyaç duyulmaktadır. Bu özellikler Schoenfeld artığını diğer yöntemlere göre avantajlı hale getirmektedir.

Elde edilen kısmi artık değerinin zamana karşı grafiği, orantılı hazard varsayımı sağlandığı durumda, sıfır etrafında rasgele dağılım göstermektedir. Bu durumda artıkların koşullu beklenen değerleri sıfıra yaklaşmaktadır.

Barlow ve Prentice birkaç yıl sonrasında yeni bir artık türü olan Martingale artığını önermişlerdir. Martingale artığı aşağıdaki gibi tanımlanmaktadır.

$$\widehat{M}_i = \delta_i - \widehat{H}_0(\tau_i)e^{\widehat{\beta}'X_i} \quad (1-44)$$

Yukarıdaki formülde geçen $\widehat{H}_0$ terimi, birikimli temel hazard olmak üzere, τ_i i . bireyin gözlem zamanını, δ_i ise gözlemin durumunu ifade etmektedir. Eğer gözlem sansürlü ise δ_i sıfır değerini, sansürlü olmadığında ise δ_i bir değerini almaktadır. Martingale artıkları gözlenen ve beklenen değerleri karşılaştırmaktadır. Eşitlikte geçen ilk kısım gözlenen değeri ikinci kısım ise beklenen değeri ifade etmektedir. Ancak Therneau'ya göre Martingale artıklarının dezavantajları vardır. Martingale artıkları aşırı derecede sağa çarpıklık göstermektedir. Bu çarpıklık özelliği, sapan değerlerin bulunmasını zorlaştırmaktadır. Martingale artıklarının diğer bir dezavantajı ise, en fazla 1, en az ise $-\infty$ değerini almasıdır. Bu dezavantajları ortadan kaldırmak adına Therneau ve arkadaşları, Cox regresyon için sapma artıklarını önermişlerdir. Bu artık Martingale artığına göre daha simetrik bir şekilde dağılım göstermekte olup,

$$d_i = \text{sign}(\widehat{M}_i)[-2\{\widehat{M}_i + \delta_i \log(\delta_i - \widehat{M}_i)\}]^{1/2} \quad (1-45)$$

eşitliği ile hesaplanmaktadır. İşaret fonksiyonu Martingale artığının işaretidir. Dolayısı ile Martingale artığı ile sapma artığı aynı işarete sahiptir. Sapma artıkları simetrik olarak dağılmakta olup standart sapması bire eşittir. Dolayısı ile artık grafiklerinin yorumlanması oldukça kolay olup sapma artıkları grafiği yardımıyla diğer bireylere göre çok uzun veya çok kısa yaşayanlar kolay bir şekilde belirlenebilmektedir (Therneau ve Grambsch 2000).

Skor artığı kullanılan bir başka artık türüdür. i . gözlemin j . açıklayıcı değişkeni için skor artığı

$$L_{ij}(\hat{\beta}) = \int_0^\infty \{X_{ij}(t) - \bar{X}_j(\hat{\beta}, t)\} d\widehat{M}_i(t) \quad (1-46)$$

'dır.

Formülde geçen $\bar{X}_j(\hat{\beta}, t)$ terimi,

$$\bar{X}_j(\hat{\beta}, t) = \frac{\sum Y_i(t) \exp(b'X_i(t))X_{ij}(t)}{\sum X_i(t) \exp(b'X_i(t))} \quad (1-47)$$

yardımıyla elde edilmektedir. Skor artığı her bir gözlem için ayrı ayrı elde edilememektedir ancak her bir kovaryat için ayrı ayrı elde edilebilmektedir. Bu artıklar ayrıca beklenen değer için sıfır değerine sahiptir ve birbiriyle ilişkisizdir (Therneau ve Grambsch 2000).

2.2. Zayıflık Modelleri (Frailty Models)

Sağkalım analizlerinde, belirli bir hastalık bakımından benzer bağımsız faktör özelliklerine sahip bireylerin risklerinin aynı olduğu varsayılmaktadır. Ancak genellikle bilinmeyen faktörler sebebiyle bazı bireyler diğer bireylere göre yüksek derecede riske sahiptir. Bireyler arasındaki bilinmeyen bu heterojenlikler zayıflık olarak adlandırılmaktadır (Hanagal 2011).

Gözlemler arasındaki heterojenliklerin değerlendirilmesi oldukça önemli bir konudur. Örneğin, stres ile baş etme de risk faktörlerinin ve sağkalım sürelerinin değerlendirileceği bir araştırmada iki farklı popülasyonun olduğu düşünölsün. Heterojenliğin değerlendirilmesi adına bu popölasyonlardan birinin homojen diğ erinin ise heterojen olduđu varsayölsin. Homojen olan popölasyonda strese bağılı olan öölümlerin rasgele olarak gerçekleştiğı görölürken (bireyler arasında zayıflık yoktur) , heterojen olan popölasyonda homojen gruba göre daha az öölümler gerçekleşmektedir ve ilk olarak daha zayıf olan bireyler öölmemektedir. Ancak değı erlendirmelerde zayıflık modelinde kişilerin zaman içinde stres ile baş edebilme yeteneklerini kazanması dikkate alınmamaktadır.

Sağkalım analizlerinde zayıflık terimi oldukça önemli bir yere sahiptir ve bireyler arasındaki zayıflıklar mutlaka araştırılmalıdır. Zayıflık kelime anlamından da anlaşılacağı üzere, bazı bireylerin diğ er bireylere göre riskinin daha fazla olması veya bazı bireylerin sonuca (ölüm, metastaz gibi) daha ç abuk ulaş ma eğ iliminde olmasıdır. Zayıflık modelleri günümüzde çok sıklıkla kullanölmassa da modellerin literatüre kazandırılması 1979 yılında gerçekleşmiştir. Vaupel 1979 yılında yaptığı çalış mada, bireylerin yaş, cinsiyet, ağırlık gibi özelliklerinin benzer olmasına rağmen farklı bireylerde farklı risklerin olduğunu belirtmiştir. Yani her ne kadar sonuç değı şkeni ile ilgili bağımsız değı şkenlerin özellikleri

bireylerde benzer olsa da her bir bireyin ilgili sonuç değişkeni ile karşılaşma riski farklıdır. Bireyler kendi içinde heterojenlik göstermektedir. Dolayısı ile bireylere ait değişken özellikleri bireyden bireye farklılık göstermektedir veya gözlemler farklı dağılım özelliklerine sahip olabilmektedir. Vaupel, bireylerin ölüm riskini analiz etmek için, bireylerin paylaştığı gözlenemeyen veya ölçülmeyen/ölçülemeyen bir rasgele etkinin olduğunu ve bu rasgele etkiyi simgeleyecek bir zayıflık terimi kullanmışlardır (Hanagal 2011).

Zayıflık varyasyonu, bireyler arasındaki risklerin farklılıkları, gözlenemeyen heterojenlikleridir. Bu varyasyon rastgelelik, çevresel veya/ve genetik faktörlerden kaynaklanmaktadır. Literatürdeki bazı biyomedikal çalışmalar bireylerin kansere yakalanma hassasiyetlerinin yüksek derecede varyasyon gösterdiğini belirtmişlerdir. Yapılan çalışmaların birçoğunda, ilgili kansere etki eden birçok önemli risk faktörünün olması durumunda bile bireyler arasında yüksek derecede heterojenliğin olduğu gözlemlenmiştir. Çalışmalar ayrıca, popülasyondaki bireyler arasındaki heterojenliğin kişilerin geçmiş hayatlarındaki epigenetik değişikliklerden kaynaklandığını göstermiştir. Bireyler arasındaki heterojenlik için bir başka bakış açısı ise, aslında bireylerin dünyaya gelmeden önce bireysel farklılıklarının var olmasıdır. Ancak çoğu araştırmacılar bireyler arasındaki heterojenliğin zaman ilerledikçe hayatın zorluğu ile birlikte farklılaştığını düşünmektedir. Bireyler arasındaki heterojenliğin bir başka sebebi ise aile öyküsünün olup olmadığıdır. Aile öyküsü olan bireylerin hastalığa yakalanma riski, aile öyküsü olmayan bireylere göre oldukça yüksek olduğu bilinmektedir (Aalen ve ark. 2015).

Özet olarak, sağkalım analizlerinde açıklayıcı değişkenler bilindiğinde orantılı hazard modeli kullanılarak bilinen yöntemler uygulanabilir. Ancak tüm önemli risk faktörlerinin modele dâhil edilmesi imkânsızdır. Çünkü faktörlerin bazı durumlarda gerçekten var olup olmadığı veya risk faktörü olarak değerlendirilmesi gerektiği bilinmeyebilir. Ayrıca risk faktörü olarak değerlendirilmesi gerektiği bilinse bile, risk faktörünün ölçülmesi büyük ölçüde finansal destek ve zaman gerektirmesi sebebiyle hesaplanması veya ölçülmesi imkânsız olduğu durumlarda söz konusu olabilir. Dolayısı ile bireysel risk hızının tahmininde, gözlenemeyen zayıflıkların değerlendirilmesi gerekmektedir. Bu zayıflıklar dikkate alınmadığında hazard fonksiyonu gerçekte olması gerekenden daha düşük tahmin sonuçları üretmekte ve bu durum sonuçların doğruluğunu azaltmaktadır (Wienke 2011).

İzleyen bölümde ise, zayıflık modellerinin çeşitleri ve modellerin teorik özellikleri açıklanmaktadır. Zayıflık modellerinde bireyin hazard fonksiyonu, gözlemlenemeyen zamana bağımlı z rasgele değişkenine dayanmaktadır. z rasgele değişkeni, zamandan bağımsız olmak üzere temel hazard fonksiyonuna aşağıdaki gibi çarpımsal bir etkiyi yaratmaktadır (Wienke 2011).

$$\mu(t|z) = z\mu_0(t) \quad (1-48)$$

Burada z terimi, popülasyon boyunca farklılık gösteren negatif olmayan rasgele karma değişkendir. Beklenen zayıflık dağılımı mevcut olduğunda zayıflık dağılımı $E(Z)=1$ şeklinde standartlaştırılmaktadır. $\sigma^2 = V(Z)$ varyansı popülasyon boyunca heterojenliğin bir ölçüsü olarak yorumlanmaktadır. σ^2 küçük olduğunda, z değeri bir etrafında değer almaktadır. σ^2 büyük olduğunda, $z\mu_0(t)$ bireysel hazard fonksiyonunda oldukça büyük heterojenlik göstermektedir.

Model içindeki gözlemlenen açıklayıcı değişkenler, Cox modeline benzer olarak $X = (X_1, X_2, \dots, X_k)$ olmak üzere, bu değişkenlerin modele katkılarını gösteren katsayılar $\beta = (\beta_1, \beta_2, \dots, \beta_k)$ katsayıları ise aşağıdaki gibi tanımlanmaktadır (Wienke 2011).

$$\mu(t|X, z) = z\mu_0(t)e^{\beta'X} \quad (1-49)$$

Tüm bireyler için $Z=1$ olduğunda, orantılı hazard modeli elde edilmektedir. $(T_i, \Delta_i, X_i, Z_i)$ verisinin olabilirliği aşağıdaki eşitlik yardımıyla elde edilmektedir (Wienke 2011).

$$\prod_{i=1}^n (Z_i \mu_0(t_i) e^{\beta'X_i})^{\delta_i} \exp(-Z_i M_0(t_i) e^{\beta'X_i}) \quad (1-50)$$

Bu olabilirlik formu parametre tahminlerinde kullanılmaktadır. Z bir bireyin zayıflığı olmak üzere, $S(t|z)$ sağkalım fonksiyonu aşağıdaki gibidir (Wienke 2011).

$$S(t|Z) = e^{-\int_0^t \mu(s|Z) ds} = e^{-Z \int_0^t \mu_0(s) ds} = e^{-Z M_0(t)} \quad (1-51)$$

Yukarıdaki eşitlikte yer alan $M_0(t) = \int_0^t \mu_0(s) ds$ fonksiyonu, birikimli temel hazard fonksiyonunu göstermektedir.

Popülasyonda sağkalım fonksiyonu ise, bilinen zayıflık olasılık yoğunluk fonksiyonu dağılımı altında koşullu sağkalım fonksiyonunun ağırlıklandırılmış ortalamasıdır ve aşağıdaki gibi ifade edilmektedir (Wienke 2011).

$$S(t) = ES(t|Z) = E e^{-Z M_0(t)} = L(M_0(t)) \quad (1-52)$$

Laplace dönüşümü koşulsuz sağkalım dağılımının sonuçlarını elde etmek için kullanılabilir. Olay zamanlarının olasılık ve hazard fonksiyonları, zayıflığın beklenen ve varyans değerleri Laplace dönüşümü yardımıyla aşağıdaki gibi elde edilmektedir (Wienke 2011).

$$f(t) = -\mu_0(t) L'(M_0(t)) \quad (1-53)$$

$$\mu(t) = -\mu_0(t) \frac{L'(M_0(t))}{L(M_0(t))} \quad (1-54)$$

$$E(Z) = -L'(0) \quad (1-55)$$

$$V(Z) = L''(0) - (L'(0))^2 \quad (1-56)$$

Yukarıdaki formüllerde geçen L' ve L'' terimleri Laplace dönüşümünün birinci ve ikinci türevlerini göstermektedir.

Zayıflık dağılımı, izlenme sürecinin başında popülasyonun zayıflığını tanımlamaktadır. Zayıflığın her birey için zaman boyunca sabit olduğu ancak popülasyonda zaman boyunca değişebileceği varsayılmaktadır. t zamanında sağkalanlar arasındaki zayıflığın dağılımı,

$$f(z|T > t) = \frac{f_z(z)S(t|z)}{S(t)} \quad (1-57)$$

ile elde edilir ve parametre tahminlerinde EM algoritması kullanılmaktadır. t zamanında ölen bireylerin zayıflık dağılımı ise Bayes formülü yardımıyla elde edilmektedir (Wienke 2011).

$$f(z|t) = \frac{f(z,t)}{f(t)} = \frac{f(t|z)f(z)}{f(t)} \quad (1-58)$$

Sonraki bölümde farklı zayıflık dağılımlarına sahip zayıflık modelleri açıklanmaktadır. Bu modeller arasında en çok kullanılan zayıflık modelleri ise Gamma ve Lognormal dağılımlarıdır.

2.2.1. Kesikli Zayıflık Modeli

Tek değişkenli zayıflık modellerinde çalışma popülasyonu iki alt popülasyondan oluşmaktadır. Bu alt popülasyonların farklı koşullar için homojen bir yapı sergilediği varsayılmaktadır. Bu model ikili zayıflık modeli olarak adlandırılmaktadır. Örneğin, herhangi bir hastalık için sağkalım süreleri analiz edilmek istendiğinde ve genetik yapının etkisi ölçülmediği durumlarda, genetik etki zayıflık terimi içerisinde incelenebilmektedir. η_1 hastalık genotipini taşıyan bireylerin oranının π olduğu varsayıldığında, $P(Z=\eta_1) = \pi$ ve $P(Z = \eta_2) = 1 - \pi$ 'dir (Wienke 2011).

Laplace dönüşümleri ise aşağıdaki gibi ifade edilmektedir.

$$L(u) = \pi e^{-u\eta_1} + (1 - \pi)e^{-u\eta_2} \quad (1-59)$$

$$L'(u) = -\pi\eta_1 e^{-u\eta_1} - (1 - \pi)\eta_2 e^{-u\eta_2} \quad (1-60)$$

$$L''(u) = -\pi\eta_1^2 e^{-u\eta_1} + (1 - \pi)\eta_2^2 e^{-u\eta_2} \quad (1-61)$$

İkili dağılıma sahip zayıflık değişkeninin beklenen ve varyans değeri ise Laplace dönüşümü yardımıyla kolaylıkla elde edilmektedir.

$$E(Z) = \pi\eta_1 + (1 - \pi)\eta_2 \quad (1-62)$$

$$V(Z) = \pi(1 - \pi)(\eta_1 - \eta_2)^2 \quad (1-63)$$

Sağkalım ve olasılık fonksiyonları ise,

$$S(t) = L(M_0(t)) = \pi e^{-\eta_1 M_0(t)} + (1 - \pi)e^{-\eta_2 M_0(t)} \quad (1-64)$$

$$f(t) = \pi\eta_1 \mu_0(t) e^{-\eta_1 M_0(t)} + (1 - \pi)\eta_2 \mu_0(t) e^{-\eta_2 M_0(t)} \quad (1-65)$$

‘dir.

Hazard değeri ise,

$$\mu(t) = \frac{\pi\eta_1 e^{-\eta_1 M_0(t)} + (1 - \pi)\eta_2 e^{-\eta_2 M_0(t)}}{\pi e^{-\eta_1 M_0(t)} + (1 - \pi)e^{-\eta_2 M_0(t)}} \mu_0(t) \quad (1-66)$$

‘dir.

$\eta = \eta_1$ ikili zayıflık modelinde olabilirlik fonksiyonu,

$$L(\theta, \eta, \pi) = \prod_{i=1}^n \left(\frac{\pi \eta e^{-\eta M_0(t_i; \theta)} + (1-\pi) e^{-\frac{1-\pi\eta}{1-\pi} M_0(t_i; \theta)}}{\pi e^{-\eta M_0(t_i; \theta)} + (1-\pi) e^{-\frac{1-\pi\eta}{1-\pi} M_0(t_i; \theta)}} \mu_0(t_i; \theta) \right)^{\delta_i} * \left(\pi e^{-\eta M_0(t_i; \theta)} + (1-\pi) e^{-\frac{1-\pi\eta}{1-\pi} M_0(t_i; \theta)} \right) \quad (1-67)$$

ile belirtilmektedir (Wienke 2011).

t zamanında hastalık genotipine sahip bireylerin oranı $\pi(t)$ olarak ifade edildiğinde, bu oranın hesaplanabilmesi için hastalık genotipine sahip olan ve olmayan bireylerin koşullu sağkalım fonksiyonları dikkate alınmalıdır.

$\pi(t)$ oranı ise aşağıdaki eşitlik yardımıyla elde edilmektedir (Wienke 2011).

$$\pi(t) = \frac{\pi e^{-\eta_1 M_0(t)}}{\pi e^{-\eta_1 M_0(t)} + (1-\pi) e^{-\eta_2 M_0(t)}} \quad (1-68)$$

İkili zayıflık modellerindeki formüller k nokta dağılımları için genelleştirildiğinde,

$$\mu(t|X, Z) = Z \mu_0(t) e^{\beta' X} \quad (1-69)$$

$$EZ = \sum_{j=1}^k P(Z = \eta_j) \eta_j = 1 \quad (1-70)$$

elde edilmektedir.

2.2.2. Gamma Zayıflık Modeli

Gamma dağılımı $\Gamma(k, \lambda)$, k parametresine bağlı olarak farklı dağılımlara yakınsayan esnek bir dağılımdır. Gamma dağılımlı rasgele değişkenin olasılığı,

$$f(z) = \frac{1}{\Gamma(k)} \lambda^k z^{k-1} e^{-\lambda z} \quad (1-71)$$

eşitliği yardımıyla elde edilmektedir (Wienke 2011).

Gamma dağılımına sahip z rasgele değişkeninin yoğunluk fonksiyonu,

$$f(z) = \frac{1}{\Gamma(\frac{1}{\sigma^2})} \left(\frac{1}{\sigma^2}\right)^{\frac{1}{\sigma^2}} z^{\frac{1}{\sigma^2}-1} \exp\left(-\frac{z}{\sigma^2}\right) \quad (1-72)$$

'dir.

Koşulsuz sağkalım fonksiyonu ise Laplace dönüşümü ile aşağıdaki gibi elde edilmektedir.

$$S(t) = L(M_0(t)) = \frac{1}{(1+\sigma^2 M_0(t))^{1/\sigma^2}} \quad (1-73)$$

Koşulsuz olasılık yoğunluk fonksiyonu,

$$f(t) = \frac{\mu_0(t)}{(1+\sigma^2 M_0(t))^{\frac{1}{\sigma^2}+1}} \quad (1-74)$$

koşulsuz hazard fonksiyonu ise,

$$\mu(t) = \frac{\mu_0(t)}{1+\sigma^2 M_0(t)} \quad (1-75)$$

eşitliği yardımıyla elde edilmektedir (Wienke 2011).

Popülasyon zayıflığına ait dağılım, bireylerin çalışmaya seçim zamanı farklılıkları dolayısıyla zamana bağlı olarak değişkenlik gösterebilmektedir. Dolayısı ile, herhangi bir zamanda sağkalanlar arasındaki zayıflık yine Gamma dağılımına sahip olacak ancak σ^2 ve birikimli hazard fonksiyonuna bağlı olarak yeni parametre gösterimleri kullanılacaktır (Wienke 2011).

Herhangi bir zamanda $T > t$ olacak şekilde sağkalanlar arasındaki zayıflık dağılımının olasılık fonksiyonu aşağıdaki gibi elde edilmektedir (Wienke 2011).

$$f(z|X, T > t) = \frac{s(t|X, z)f(z)}{s(t|X)} = \frac{(\frac{1}{\sigma^2} + M_0(t)e^{\beta'X})^{\frac{1}{\sigma^2}}}{\Gamma(\frac{1}{\sigma^2})} z^{\frac{1}{\sigma^2}-1} \exp(-z(\frac{1}{\sigma^2} + M_0(t)e^{\beta'X})) \quad (1-76)$$

t zamanında sağkalanlar arasındaki ortalama zayıflık,

$$E(Z|X, T > t) = \frac{1}{1+\sigma^2 M_0(t)e^{\beta'X}} \quad (1-77)$$

varyans ise,

$$V(Z|X, T > t) = \frac{\sigma^2}{(1 + \sigma^2 M_0(t) e^{\beta' X})^2} \quad (1-78)$$

eşitliği ile belirlenmektedir.

Benzer şekilde t zamanında ölen bireylerin zayıflık dağılımı,

$$f(z|X, T = t) = \frac{f(t|X, Z) f(z)}{f(t|X)} = \frac{(\frac{1}{\sigma^2} + M_0(t) e^{\beta' X})^{\frac{1}{\sigma^2} + 1}}{\Gamma(\frac{1}{\sigma^2} + 1)} z^{\frac{1}{\sigma^2} + 1 - 1} \exp(-z (\frac{1}{\sigma^2} + M_0(t) e^{\beta' X})) \quad (1-79)$$

formülleri ile elde edilmektedir. Yukarıdaki fonksiyonda Gamma dağılımı $\frac{1}{\sigma^2} + M_0(t) e^{\beta' X}$ ve $\frac{1}{\sigma^2} + 1$ parametrelerine sahip olması sebebiyle t zamanında ölen bireyler arasındaki ortalama ve varyans değerlerine ait hesaplamalar aşağıdaki eşitliklerde verilmektedir (Wienke 2011).

$$E(Z|X, T = t) = \frac{1 + \sigma^2}{1 + \sigma^2 M_0(t) e^{\beta' X}} \quad (1-80)$$

$$V(Z|X, T = t) = \frac{\sigma^2 (1 + \sigma^2)}{(1 + \sigma^2 M_0(t) e^{\beta' X})^2} \quad (1-81)$$

2.2.3. Lognormal Zayıflık Modeli

Lognormal dağılım, zayıflık dağılımları arasında en sık kullanılan dağılımlar arasında yer almaktadır. Gelişen bilgi teknolojileri ile birlikte bu dağılımın hesaplama zorlukları da ortadan kaldırılmıştır. Ayrıca Lognormal zayıflık modelleri, çok değişkenli problemlerde yapılar arasındaki bağımlılıkları da dikkate alarak kullanışlı bir model kurma imkânı sağlamaktadır. Lognormal zayıflık modeli, ters Gauss zayıflık modeline oldukça benzerlik göstermektedir (Wienke 2011).

Lognormal zayıflık modelinin iki parametresi vardır. $Z = e^W$ eşitliğini üretmek için rasgele değişken $W \sim N(m, s^2)$ 'nin normal dağıldığı varsayıldığında, bu normal dağılımın parametreleri ise zayıflık parametreleri olan μ ve σ^2 'nin bir fonksiyonu olarak tanımlanmaktadır. Bu fonksiyonlar ise aşağıda ifade edilmektedir (Wienke 2011).

$$\mu = EZ = e^{m+\frac{s^2}{2}} \quad (1-82)$$

$$\sigma^2 = V(Z) = e^{2m+s^2}(e^{s^2} - 1) \quad (1-83)$$

m, s^2 parametrelerine bağılı olarak zayıflık dağılımının olasılık fonksiyonu,

$$f(z) = \frac{1}{\sqrt{2\pi s^2}} e^{-\frac{(\ln(z))^2}{2s^2}} \quad (1-84)$$

'dir. Bu modelde standart bir bireyin hazard oranı $\mu_0(t)$ 'dir.

2.2.4. Ters Gauss Zayıflık Modeli

Ters Gauss dağılımının, zayıflıkta kullanımının tanıtılması Hougaard tarafından yapılmış ardından da Manton, Klein gibi araştırmacılar da bu dağılımı kullanmışlardır. Gamma zayıflık modeline benzer olarak bu modelde basit kapalı formlar kullanıldığı için oldukça kullanışlı bir dağılım çeşididir. $\mu > 0$ ve $\lambda > 0$ parametrelili Ters Gauss dağılımına sahip rasgele değişkenin olasılık fonksiyonu,

$$f(z) = \frac{\sqrt{\lambda}}{\sqrt{2\pi z^3}} \exp\left(-\frac{\lambda}{2\mu^2 z}(z - \mu)^2\right) \quad (1-85)$$

eşitliği yardımıyla elde edilmektedir (Wienke 2011).

Yukarıda verilen olasılık fonksiyonu yardımıyla elde edilen Laplace dönüşümü ise,

$$L(u) = Ee^{-uz} = \exp\left(-\frac{\lambda\sqrt{1+\frac{2\mu^2 u}{\lambda}}}{\mu} + \frac{\lambda}{\mu}\right) * \int \frac{\sqrt{\lambda}}{\sqrt{2\pi z^3}} \exp\left(-\frac{\lambda z}{2} \frac{1+2\mu^2 u}{\mu^2} + \frac{\lambda\sqrt{1+\frac{2\mu^2 u}{\lambda}}}{\mu} - \frac{\lambda}{2z}\right) dz \quad (1-86)$$

ile ifade edilmektedir.

$$-\frac{\lambda z}{2} \frac{1+2\mu^2 u}{\mu^2} + \frac{\lambda\sqrt{1+\frac{2\mu^2 u}{\lambda}}}{\mu} - \frac{\lambda}{2z} = -\frac{\lambda}{2} \frac{\mu^2}{1+\frac{2\mu^2 u}{\lambda}} z \left(z - \frac{\mu}{\sqrt{1+\frac{2\mu^2 u}{\lambda}}}\right)^2 \quad (1-87)$$

ilişkisi sebebiyle dönüşüm,

$$\frac{\sqrt{\lambda}}{\sqrt{2\pi z^3}} \exp\left(-\frac{\lambda z}{2} \frac{1+2\mu^2 u}{\mu^2} + \frac{\lambda \sqrt{1+\frac{2\mu^2 u}{\lambda}}}{\mu} - \frac{\lambda}{2z}\right) = \frac{\sqrt{\lambda}}{\sqrt{2\pi z^3}} \exp\left(-\frac{\lambda}{2} \frac{\mu^2}{1+\frac{2\mu^2 u}{\lambda}} \left(z - \frac{\mu}{\sqrt{1+\frac{2\mu^2 u}{\lambda}}}\right)^2\right) \quad (1-88)$$

şeklini almaktadır. Buradan anlaşılabacağı üzere, ters Gauss dağılımının yoğunluk parametreleri $\frac{\mu}{\sqrt{1+\frac{2\mu^2 u}{\lambda}}}$ ve λ 'dır. Dolayısı ile integral bir değerine eşit olmaktadır. Sonuç olarak, Laplace dönüşümü,

$$L(u) = \exp\left(-\frac{\lambda \sqrt{1+\frac{2\mu^2 u}{\lambda}}}{\mu} + \frac{\lambda}{\mu}\right) = \exp\left(\frac{\lambda}{\mu} \left(1 - \sqrt{1 + \frac{2\mu^2 u}{\lambda}}\right)\right) \quad (1-89)$$

eşitliği ile elde edilmektedir. Dönüşümün birinci ve ikinci türevi ise,

$$L'(u) = -\frac{\mu}{\sqrt{1+\frac{2\mu^2 u}{\lambda}}} \exp\left(\frac{\lambda}{\mu} \left(1 - \sqrt{1 + \frac{2\mu^2 u}{\lambda}}\right)\right) \quad (1-90)$$

$$L''(u) = \frac{\mu^3}{\lambda(1+\frac{2\mu^2 u}{\lambda})^{3/2}} \exp\left(\frac{\lambda}{\mu} \left(1 - \sqrt{1 + \frac{2\mu^2 u}{\lambda}}\right)\right) + \frac{\mu^2}{1+\frac{2\mu^2 u}{\lambda}} \exp\left(\frac{\lambda}{\mu} \left(1 - \sqrt{1 + \frac{2\mu^2 u}{\lambda}}\right)\right) \quad (1-91)$$

'dir. Dağılımın beklenen değeri ise μ , varyansı $\frac{\mu^3}{\lambda}$ 'dır. $E(Z)=1$ ve $V(Z)=1/\lambda$ olduğunda, Laplace dönüşümü $e^{\frac{1}{\sigma^2}(1-\sqrt{1+2\sigma^2 u})}$ basit şeklini almaktadır ve sağkalım, hazard fonksiyonları aşağıdaki gibi ifade edilmektedir (Wienke 2011).

$$S(t) = e^{\frac{1}{\sigma^2}(1-\sqrt{1+2\sigma^2 M_0(t)})} \quad (1-92)$$

$$\mu(t) = \frac{\mu_0(t)}{(1+2\sigma^2 M_0(t))^{1/2}} \quad (1-93)$$

Ters Gauss zayıflık modelinde, regresyon parametreleri üzerine gözlenemeyen heterojenliğin etkisinin araştırılması için modele $e^{\beta'X}$ regresyon parametrelerinin de eklenmesi gerekmektedir. İkili bir açıklayıcı değişken için iki bireyin marjinal hazardları incelenecek olunursa, açıklayıcı değişken değeri bir ve sıfır olduğunda iki birey için hazard oranları,

$$\frac{\mu(t|X=1)}{\mu(t|X=0)} = \frac{(1+2\sigma^2 M_0(t))^{1/2}}{(1+2\sigma^2 M_0(t)e^\beta)^{1/2}} e^\beta \quad (1-94)$$

'dir (Wienke 2011).

$t=0$ zamanında hazard oranı e^β iken, $t \rightarrow \infty$ iken bu oran $e^{\beta/2}$ olmaktadır. Dolayısı ile Gamma zayıflık modeline benzer olarak, bu marjinal model artık bir orantılı hazard modeli değildir. Gözlenemeyen heterojenliğin etkisi ise σ^2 , β ve $M_0(t)$ parametrelerinin büyük değerlerinde daha fazla olmaktadır (Wienke 2011).

Herhangi bir t zamandaki sağkalanlar arasındaki zayıflık dağılımının olasılık fonksiyonu,

$$f(z|X, T > t) = \frac{s(t|X, z)f(z)}{s(t|X)} = \frac{1}{\sqrt{2\pi^2 z^3}} \exp\left(-\frac{(z - \frac{(1+2\sigma^2 M_0(t)e^{\beta'X})^{-1/2}}{2\sigma^2 z})^2}{1+2\sigma^2 M_0(t)e^{\beta'X}}\right) \quad (1-95)$$

'dir.

Beklenen değer ve varyans ise,

$$E(Z|X, T > t) = \frac{1}{\sqrt{1+\sigma^2 M_0(t)e^{\beta'X}}} \quad (1-96)$$

$$V(Z|X, T > t) = \frac{\sigma^2}{(1+\sigma^2 M_0(t)e^{\beta'X})^2} \quad (1-97)$$

eşitlikleri yardımıyla elde edilmektedir (Wienke 2011).

2.2.5. Pozitif Durağan Zayıflık Modeli

Normalize edilmiş bir dağılımdan alınan n tane bağımsız rasgele değişkenin toplamının dağılımı, $D(Z_1 + Z_2 + \dots + Z_n)$, bir rasgele değişkenin ölçek faktörü (γ) ile

çarpımının dağılımı ile, $D(\gamma n Z_1)$, aynı dağılımı gösterdiği zaman bu dağılıma durağan dağılım adı verilmektedir (Wienke 2011).

Normalleştirme γ , $(0,1]$ aralığında karakteristik üs olarak belirlenen bir parametre olmak üzere $n^{1/\gamma}$ ile verilmektedir. Bir parametrelili pozitif durağan rasgele dağılımın olasılık yoğunluk fonksiyonu aşağıdaki gibi tanımlanmaktadır (Wienke 2011).

$$f(z) = \frac{1}{\pi} \sum_{k=1}^{\infty} (-1)^{k+1} \frac{\Gamma(k\gamma+1)}{k!} z^{-k\gamma-1} \sin(k\gamma\pi) \quad (1-98)$$

Olasılık yoğunluk fonksiyonunda yer alan z, γ parametreleri $z \geq 0$ ve $0 \leq \gamma \leq 1$ aralıklarında değer almaktadır. Laplace dönüşümü ise $L(u) = e^{-u^\gamma}$ dir. Bu dağılımın tüm momentleri sonsuzdur. Dolayısı ile zayıflığın beklenen değeri sonsuz iken varyansı tanımlanamamaktadır (Wienke 2011).

Laplace dönüşümü yardımıyla elde edilen koşulsuz sağkalım fonksiyonu $S(t) = e^{-M_0(t)^\gamma}$ ve yoğunluk fonksiyonu ise $f(t) = \gamma M_0(t)^{\gamma-1}$ dir.

Pozitif durağan dağılım, zayıflık dağılımları içinde zayıflığın entegre edilmesinden sonra koşulsuz hazardlarda orantılı hazard koşulunu sağlayan tek dağılım özelliğini taşımaktadır.

Gözlenen kovaryatların etkilerinin görülmesi adına model incelendiğinde, ikili bir açıklayıcı değişken değerlerine sahip iki hazardın oranı $e^{\gamma\beta}$ dir. Marjinal hazardların oranı e^β yerine $e^{\gamma\beta}$ ile orantılıdır. Dolayısı ile ilişkili kovaryatlar modele dâhil edilmediğinde orantılı hazard modelindeki parametre tahminleri yanlı olacaktır. $0 < \gamma < 1$ arasında olduğu için, popülasyon hazard oranı bire yaklaşacaktır. γ parametresi ne kadar bir değerinden sapma gösteriyorsa, o kadar popülasyonun hazard oranı koşullu hazard oranında sapma gösterecektir (Wienke 2011).

2.2.6. Power Varians Function (PVF) Zayıflık Modeli

PVF zayıflık modeli $\mu > 0, \lambda > 0$ ve $0 < \gamma \leq 1$ olmak üzere olasılık yoğunluk fonksiyonu,

$$f(z) = e^{-\lambda(1-\gamma)\left(\frac{z}{\mu}\right)^{\frac{1}{1-\gamma}}} * \frac{1}{\pi} \sum_{k=1}^{\infty} (-1)^{k+1} \frac{(\lambda(1-\gamma))^{\frac{k(1-\gamma)}{\gamma}} \mu^{k\gamma}}{\gamma^k} \frac{\Gamma(k\gamma+1)}{k!} z^{-k\gamma-1} \sin(k\gamma\pi) \quad (1-99)$$

eşitliği ile ifade edilmektedir (Wienke 2011).

Laplace dönüşümü ise,

$$L(u) = e^{\frac{\lambda(1-\gamma)}{\gamma}(1-(1+\frac{\mu u}{\lambda(1-\gamma)})^\gamma)} \quad (1-100)$$

olarak ifade edilmektedir ve bu dönüşümlerin birinci ve ikinci türevleri aşağıdaki gibi ifade edilmektedir.

$$L'(u) = -\mu(1 + \frac{\mu u}{\lambda(1-\gamma)})^{\gamma-1} e^{-\frac{\lambda(1-\gamma)}{\gamma}(1-(1+\frac{\mu u}{\lambda(1-\gamma)})^\gamma)} \quad (1-101)$$

$$L''(u) = \frac{\mu^2}{\lambda}(1 + \frac{\mu u}{\lambda(1-\gamma)})^{\gamma-2} e^{-\frac{\lambda(1-\gamma)}{\gamma}(1-(1+\frac{\mu u}{\lambda(1-\gamma)})^\gamma)} + \mu^2(1 + \frac{\mu u}{\lambda(1-\gamma)})^{2\gamma-2} e^{-\frac{\lambda(1-\gamma)}{\gamma}(1-(1+\frac{\mu u}{\lambda(1-\gamma)})^\gamma)} \quad (1-102)$$

Gamma modeline benzer olarak zayıflık rasgele değişkeninin beklenen değeri μ , varyansı ise μ^2/λ 'dır. Dolayısı ile PVF zayıflık modeli, pozitif durağan dağılımın özel bir formu olarak bilinmektedir (Wienke 2011).

Gözlenemeyen heterojenliğin belirlenebilmesi için gözlenen açıklayıcı değişkenlerinde modele dâhil edilmesi gerekmektedir. Yine diğer zayıflık modellerine benzer olarak iki kategoriden oluşan niteliksel bir açıklayıcı değişken düşünüldüğünde, marjinal hazardların oranı,

$$\frac{\mu(t|X=1)}{\mu(t|X=0)} = \frac{(1+\frac{\sigma^2}{1-\gamma} M_0(t))^{1-\gamma}}{(1+\frac{\sigma^2}{1-\gamma} M_0(t)e^\beta)^{1-\gamma}} e^\beta \quad (1-103)$$

olarak ifade edilmektedir. Pozitif zayıflık modelinin tersine bu modelde marjinal hazardlar paralellik göstermemektedir (Wienke 2011).

Herhangi bir zamandaki, sağkalanlar arasındaki zayıflık dağılımının hesaplanması için $\mu=1$ 'e eşitlenir ve aşağıdaki formül yardımıyla olasılıklar hesaplanmaktadır.

$$f(Z|T > t) = \frac{S(t|Z)f(Z)}{S(t)} e^{\frac{-(1-\gamma)}{\sigma^2 A^{-\gamma}}(\frac{Z}{A^{\gamma-1}} - \frac{1}{\gamma})} * \frac{1}{\pi} \sum_{k=1}^{\infty} (-1)^{k+1} \frac{(1-\gamma)^{k(1-\gamma)} A^{k\gamma(\gamma-1)} \Gamma(k\gamma+1)}{(\sigma^2 A^{-\gamma})^{k(1-\gamma)} \gamma^k k!} Z^{-k\gamma-1} \sin(k\gamma\pi) \quad (1-104)$$

Yukarıda ifade edilen A terimi, $A=1 + \frac{\sigma^2 M_0(t)}{1-\gamma}$ 'e eşittir ve olasılık fonksiyonunun parametreleri A , $\sigma^2 A^{-\gamma}$ ve γ 'dır (Wienke 2011).

PVF zayıflık modeli diğer dağılımlara dönüşebilme özelliği ile esnek bir dağılım özelliği taşımaktadır. Dolayısı ile kullanışlılığı sayesinde bu modellerden sıklıkla yararlanılmaktadır.

2.2.7. Bileşik Poisson Zayıflık Modeli

Bileşik zayıflık modeli Aalen tarafından 1988 ve 1992 yıllarında tanıtılmıştır. PVF modeline benzer olarak bileşik zayıflık modelinin olasılık yoğunluk fonksiyonu ($\mu > 0, \lambda > 0, \gamma < 0$ olmak üzere) aşağıdaki eşitlikteki gibi ifade edilmektedir (Wienke 2011).

$$f(z) = e^{-\lambda(1-\gamma)\left(\frac{z}{\mu} - \frac{1}{\gamma}\right)} * \frac{1}{\pi} \sum_{k=1}^{\infty} \frac{(\lambda(1-\gamma))^{k(1-\gamma)} \mu^{k\gamma} \Gamma(k\gamma+1)}{\gamma^k k!} (-z)^{-k\gamma-1} \sin(k\gamma\pi) \quad (1-105)$$

Modelde ilgilenilen olay ile karşılaşmayan bireyler için sıfır zayıflığına sahip bir alt grup vardır. Bu özellik ise bu modelin ilginç bir özelliğidir. Toplam ölüm oranının hesabı durumunda ise, insanın doğası gereği ölüm ile karşılaşma gerçeği sebebiyle ilgilenilen olay ile karşılaşmama durumu söz konusu değildir. Ancak sağkalım analizlerinde ilgilenilen olay belirli bir hastalık sebebiyle ölüm olduğu için bu modeller kullanım alanına sahiptir.

Bileşik Poisson dağılımı adında anlaşılaacağı üzere birbirinden bağımsız aynı Gamma dağılımına sahip rasgele değişkenlerin Poisson dağılımlarının toplamıdır. Z değeri $N > 0$ olduğu durumlarda ρ beklenen değerine sahip N 'in Poisson dağıldığı varsayıldığında, $V_i \sim \Gamma(k, \lambda)$ olmak üzere Gamma dağılımlı ve $V_1, V_2, \dots, V_N$ birbirinden bağımsız olduğunda V_i 'lerin toplamına eşit iken, N sıfıra eşit olduğu durumda ise Z sıfır değerini almaktadır (Wienke 2011).

Bileşik Poisson dağılımın zayıflığının Laplace dönüşümü,

$$L(u) = e^{-\rho\left(1 - \left(1 + \frac{u}{\lambda}\right)^{-k}\right)} \quad (1-106)$$

'dir. $\rho = -\frac{k\lambda^\gamma}{\gamma}$, $k = -\gamma$ eşitliklerindeki ilişkiler yardımıyla bileşik Poisson dağılımlı rasgele değişkenin Laplace dönüşümü,

$$L(u) = \exp\left[\frac{k\lambda^\gamma}{\gamma} \left\{1 - \left(1 + \frac{u}{\lambda}\right)^\gamma\right\}\right] \quad (1-107)$$

basit şekline dönüşmektedir. Dönüşümün birinci ve ikinci türevi, beklenen ve varyans değerleri ise aşağıdaki eşitlikler yardımıyla ifade edilmektedir (Wienke 2011).

$$L'(u) = -k\lambda^{\gamma-1} \left(1 + \frac{u}{\lambda}\right)^{\gamma-1} \exp \left[\frac{k\lambda^\gamma}{\gamma} \left\{ 1 - \left(1 + \frac{u}{\lambda}\right)^\gamma \right\} \right] \quad (1-108)$$

$$L''(u) = -k\lambda^{\gamma-2}(\gamma-1) \left(1 + \frac{u}{\lambda}\right)^{\gamma-2} \exp \left[\frac{k\lambda^\gamma}{\gamma} \left\{ 1 - \left(1 + \frac{u}{\lambda}\right)^\gamma \right\} \right] + k^2\lambda^{2\gamma-2} \left(1 + \frac{u}{\lambda}\right)^{2\gamma-2} \exp \left[\frac{k\lambda^\gamma}{\gamma} \left\{ 1 - \left(1 + \frac{u}{\lambda}\right)^\gamma \right\} \right] \quad (1-109)$$

$$E(Z) = k\lambda^{\gamma-1} \quad (1-110)$$

$$V(Z) = k(1-\gamma)\lambda^{\gamma-2} \quad (1-111)$$

Dağılımın marjinal sağkalım ve hazard fonksiyonu,

$$S(t) = e^{\frac{1-\gamma}{\gamma\sigma^2} \left(1 - \left(1 + \frac{\sigma^2 M_0(t)}{1-\gamma} \right)^\gamma \right)} \quad (1-112)$$

$$\mu(t) = \frac{\mu_0(t)}{\left(1 + \frac{\sigma^2}{1-\gamma} M_0(t) \right)^{1-\gamma}} \quad (1-113)$$

gibi elde edilmektedir.

PVF ve bileşik Poisson dağılımlarından elde edilen Laplace dönüşümleri benzer olduğundan, her iki dağılımında Laplace dönüşüm özellikleri benzerdir. İkili kategorik değişken için kovaryatlı model,

$$\frac{\mu(t|X=1)}{\mu(t|X=0)} = \frac{\left(1 + \frac{\sigma^2}{1-\gamma} M_0(t) \right)^{1-\gamma}}{\left(1 + \frac{\sigma^2}{1-\gamma} M_0(t)e^\beta \right)^{1-\gamma}} e^\beta \quad (1-114)$$

'dir (Wienke 2011).

γ parametresi dağılım çeşidini belirlemektedir. $\gamma \geq 0$ koşulunda PVF dağılımı özellikleri sağlanırken, $\gamma < 0$ durumunda bileşik Poisson dağılımı özellikleri sağlanmaktadır ve her iki dağılım ($\gamma=0$) Gamma dağılımı ile ayrışmaktadır.

2.2.8. Kuadratik Hazard Zayıflık Modeli

Lognormal zayıflık modelinde sağkalım ve hazard fonksiyonlarının tam olarak tanımlamaları mevcut değildir. Dolayısı ile bu dezavantajı ortadan kaldırmak ve normal dağılımın esneklik özelliğini kombine eden kuadratik hazard zayıflık modeli Aalen tarafından 1987 yılında, Yashin ve Iachine tarafından ise 1996 yılında önerilmiştir. Model rasgele etkilerin normal dağılımına ve kuadratik koşullu hazard temeline dayanmaktadır.

$W \sim N(m, s^2)$ normal dağılıma sahip rasgele değişken olmak üzere, $Z = W^2$ olduğu varsayıldığında, koşulsuz hazard fonksiyonu açısından model, $\mu(t|Z) = Z\mu_0(t) = W^2\mu_0(t)$ şeklinde ifade edilmektedir (Wienke 2011).

Laplace dönüşümü,

$$L(u) = E \exp(-uW^2) = \frac{1}{\sqrt{1+2s^2u}} \exp\left(-\frac{m^2u}{1+2s^2u}\right) \quad (1-115)$$

şeklinde gösterilmektedir. Yukarıdaki eşitlikten görüleceği üzere $m = 0$ ve $k = 1/2$, $\lambda = 1/2s^2$ parametrelerine sahip $Z = W^2$ zayıflığında Gamma dağılımı özelliği taşımaktadır. Laplace dönüşümünün birinci ve ikinci türevi, zayıflığın beklenen ve varyans değerleri aşağıdaki eşitlikler yardımıyla elde edilmektedir (Wienke 2011).

$$L'(u) = -s^2(1+2s^2u)^{-\frac{3}{2}} e^{-\frac{m^2u}{1+2s^2u}} - m^2(1+2s^2u)^{-\frac{5}{2}} e^{-\frac{m^2u}{1+2s^2u}} \quad (1-116)$$

$$L''(u) = m^4(1+2s^2u)^{-\frac{11}{2}} e^{-\frac{m^2u}{1+2s^2u}} \quad (1-117)$$

$$E(Z) = m^2 + s^2 \quad (1-118)$$

$$V(Z) = 2s^4 + 4m^2s^2 \quad (1-119)$$

Koşulsuz sağkalım ve hazard fonksiyonları ise sırasıyla,

$$S(t) = \frac{1}{\sqrt{1+2s^2M_0(t)}} \exp\left(-\frac{m^2M_0(t)}{1+2s^2M_0(t)}\right) \quad (1-120)$$

$$m(t) = m_0(t) \left(\frac{m^2}{(1+2s^2M_0(t))^2} + \frac{s^2}{1+2s^2M_0(t)} \right) \quad (1-121)$$

'dir. t zamanında sağkalanlar arasındaki W parametresi normal dağılmaktadır.

$m^2 + s^2 = 1$ eşitliğinden dolayı koşulsuz sağkalım ve hazard fonksiyonu,

$$S(t) = \frac{1}{\sqrt{1+2s^2M_0(t)}} e^{-\frac{(1-s^2)M_0(t)}{1+2s^2M_0(t)}} \quad (1-122)$$

$$\mu(t) = \mu_0(t) \frac{1+2s^4M_0(t)}{(1+2s^2M_0(t))^2} \quad (1-123)$$

gibi yazılabilmektedir (Wienke 2011).

Kuadratik hazard zayıflık modelinde, gözlenmemiş heterojenliğin etkisinin görülmesi adına gözlemlenen kovaryat etkileri dâhil edilmektedir. Dolayısı ile ikili yapıya sahip bir açıklayıcı değişken için marjinal hazard oranı,

$$\frac{\mu(t|X=1)}{\mu(t|X=0)} = \frac{1+2s^4M_0(t)e^\beta}{(1+2s^2M_0(t)e^\beta)^2} \frac{(1+2s^2M_0(t))^2}{1+2s^4M_0(t)} e^\beta \quad (1-124)$$

'dir (Wienke 2011).

Bu modelde marjinal hazardlar, pozitif durağan zayıflık modelindeki gibi zamandan bağımsız değildir. Ancak marjinal ve koşullu hazard oranları sıfır noktasında çakışmaktadır, izlem süresinin ilk bölümlerinde ise giderek farklılaşmakta ve t zamanı sonsuza yaklaştığında oranlar birbirine yaklaşmaktadır. Kuadratik hazard zayıflık modelinde, W rasgele etkisinin varyansı s^2 iken, bu varyans değeri zayıflık terimine ait değildir. Bu farklılığın görülmesi adına s^2 rasgele etkilerin varyans parametresi olarak kullanılırken, σ^2 zayıflık teriminin varyans parametresi olarak kullanılmaktadır. Zayıflığın varyansı rasgele değişkenin normal dağılımının dördüncü momentinden ve $m^2 + s^2 = 1$ eşitliği yardımıyla kolaylıkla elde edilmektedir (Wienke 2011).

$$V(Z) = 2s^2(2 - s^2) \quad (1-125)$$

2.2.9. Lèvy Zayıflık Modeli

Herhangi bir zayıflık modeli uygulandığında, genellikle bütün grupların aynı zayıflık dağılımına sahip olduğu varsayılmaktadır. Dolayısı ile bir gruptaki bireylerin koşullu hazard değeri diğer gruplara göre orantılı olarak daha büyük olacaktır. Zayıflık ve koşulsuz hazardlar modele entegre edildiğinde ise, hazard oranları orantılı olmayacaktır.

Lèvy sürecine dayanan zayıflık dağılımları kullanılması durumunda popülasyonun hazardı orantılı olacaktır. Bu özelliği ile Lèvy modelleri oldukça avantajlıdır. Zayıflık modeli oluşturulurken negatif olmayan $D = \{D(s), s \geq 0\}$ Lèvy sürecinden dağılımlar oluşturulmaktadır. Sürecin Laplace dönüşümü aşağıdaki Lèvy-Chinchin formülü yardımıyla elde edilmektedir (Wienke 2011).

$$L(u; s) = e^{-s\psi(u)} \quad (1-126)$$

Modelin sağkalım fonksiyonu ve koşulsuz hazard fonksiyonu aşağıdaki gibidir (Wienke 2011).

$$S(t) = e^{-s\psi(M_0(t))} \quad (1-127)$$

$$\mu(t) = s\mu_0(t) \psi'(M_0(t)) \quad (1-128)$$

2.2.10. Log-t Zayıflık Modeli

Zayıflık modellerinde t dağılımının logaritması oldukça nadir kullanılmaktadır. Çalışmaların çoğunda modelin çözümlemeleri için Bayes yaklaşımı ve Markov Chain Monte Carlo (MCMC) yöntemi kullanılmaktadır (Wienke 2011).

2.2.11. Paylaşılmış Zayıflık Modeli (Shared Frailty Model)

Zayıflık modelleri gözlenemeyen açıklayıcı değişkenlerden kaynaklanan heterojenlik için kullanılmaktadır. Ancak bazı durumlarda, olay zamanları arasındaki ilişkilerin modellenmesinde de zayıflık modelleri kullanılmaktadır. Paylaşılmış zayıflık modelleri karma ve çok değişkenli bir modeldir. Çünkü her bir kümedeki genel riskin rasgele olduğu varsayılmaktadır. Modelde aynı kümedeki tüm bireyler için zayıflık geneldir ve koşullu bağımsız bir modeldir. Dolayısı ile olay zamanları arasındaki

bağımlılıktan model etkilenmektedir. Bu sebep ile model paylaşılmış zayıflık modeli olarak anılmaktadır. Modelde hem bireyler arası hem gruplar arası varyasyonlar hazard fonksiyonunda tanımlanmaktadır. Farklı kümelerdeki bireylerin bağımsız olduğu varsayılmaktadır. Zayıflık değişkeninin varyasyonu sıfır olduğu durumlarda kümelerdeki olay zamanlarının bağımsız olduğu anlamına gelmektedir, olay zamanları bağımlı olduğunda ise zayıflık değeri pozitif olacaktır. Sağkalım modellerinde olay zamanları arasındaki bağımlılık daha karmaşık sorulara cevap vermesi açısından (yaşlanma gibi) daha kullanışlıdır. Bu bağımlılık birbiri ile ilişkili bireyler için (aile) veya tekrar eden başarısızlıklar (kalp krizi) olaylarında ortaya çıkmaktadır. Böyle durumlarda paylaşılmış zayıflık modellerinin kullanımı önerilmektedir. Her birey için hazard modeli, tek değişkenli zayıflık modeli ile aynı gibi görünmesine rağmen paylaşılmış zayıflık modelinde zayıflık aynı grubu paylaşan bireylerin rölatif riskinin bir ölçümü olarak tanımlanmaktadır (Wienke 2011).

Paylaşılmış zayıflık modelinde belirli bir zayıflıkta, aynı kümedeki tüm başarısızlık zamanlarının koşullu olarak bağımsız olduğu varsayılmaktadır. Zayıflık terimi tüm zamanda sabit ve aynı kümedeki tüm bireylerin aynı zayıflığı paylaştığı bilinmektedir. Dolayısı ile kümedeki olay zamanları arasındaki bağımlılık pozitifdir. Paylaşılmış zayıflık modelleri bireylerden elde edilen gözlem değerlerinin bağımsız olduğu durumlarda oldukça kullanışlıdır ve Cox modelinin bir uzantısı olarak düşünülmektedir. Bu model kümelenebilir olay zamanlarında oluşan veri setleri için uygundur. Kümeler arasındaki korelasyon yapıları simetriktir. Özellikle çok merkezli klinik araştırmalarda bu korelasyon yapıları ile sıkça karşılaşilmektedir (Wienke 2011).

Paylaşılmış zayıflık modelinin incelenmesinde n kümenin olduğu ve i . kümenin n_i gözlemden oluştuğu düşünüldüğünde bu gözlemler gözlenmeyen zayıflık $Z_i (i \leq n)$ ile ilişkilendirilmektedir. X_{ij} vektörü i . kümedeki j . gözlemin T_{ij} zamandaki açıklayıcı değişkenlerin bilgisini içermektedir. Zayıflık terimine bağlı olarak, i . kümenin sağkalım zamanları bağımsızdır ve hazard fonksiyonu aşağıdaki fonksiyon yardımıyla elde edilmektedir (Wienke 2011).

$$\mu(t|X_{ij}, Z_i) = Z_i \mu_0(t) e^{\beta' X_{ij}} \quad (1-129)$$

Bu fonksiyonda $\mu_0(t)$ temel hazard fonksiyonunu, β tahmin edilecek olan sabit etki parametrelerini göstermektedir. Modelde yer alan zayıflıkların birbirinden bağımsız

olduğu varsayılmaktadır. Zayıflık teriminin yoğunluğu tahmin edilecek olan bilinmeyen parametrelere bağlıdır.

Paylaşılmış zayıf modelinin temel varsayımı, aynı kümeyi paylaşan bireylerin zayıflık değerinin de aynı olmasıdır. Farklı kümelerdeki bireylerin olay zamanları ise birbirinden bağımsızdır.

Zayıflığa bağlı olarak sağkalım fonksiyonu ise,

$$S(t_{i1}, \dots, t_{in_i} | X_i, Z_i) = S(t_{i1} | X_{i1}, Z_i) \dots S(t_{in_i} | X_{in_i}, Z_i) = \exp(-Z_i \sum_{j=1}^{n_i} M_0(t_{ij}) e^{\beta' X_{ij}}) \quad (1-130)$$

ile elde edilmektedir (Wienke 2011).

Fonksiyonda yer alan $M_0(t) = \int_0^t \mu_0(s) ds$ eşitliği birikimli temel hazard fonksiyonu, $X_i = (X_{i1}, \dots, X_{in_i})$ ise i . kümedeki bireylerin açıklayıcı değişken matrisidir.

Marjinal sağkalım fonksiyonu ise,

$$S(t_{i1}, \dots, t_{in_i} | X_i) = ES((t_{i1}, \dots, t_{in_i} | X_i, Z_i)) = E \exp(-Z_i \sum_{j=1}^{n_i} M_0(t_{ij}) e^{\beta' X_{ij}}) = L(\sum_{j=1}^{n_i} M_0(t_{ij}) e^{\beta' X_{ij}}) \quad (1-131)$$

'dir. Eşitlikte yer alan L terimi, zayıflık değişkeninin Laplace dönüşümüdür. Laplace dönüşümü yardımıyla elde edilen koşulsuz sağkalım fonksiyonu ise,

$$S(t_{ij} | X_{ij}) = ES(t_{ij} | X_{ij}, Z_i) = E \exp(-Z_i M_0(t_{ij}) e^{\beta' X_{ij}}) = L(M_0(t_{ij}) e^{\beta' X_{ij}}) \quad (1-132)$$

ile ifade edilmektedir (Wienke 2011).

L^{-1} terimi ise, Laplace dönüşümünün ters fonksiyonudur.

$$M_0(t_{ij}) e^{\beta' X_{ij}} = L^{-1}(S(t_{ij} | X_{ij})) \quad (1-133)$$

'dir ve i . kümenin koşulsuz sağkalım fonksiyonu,

$$S(t_{i1}, \dots, t_{in_i} | X_i) = L(L^{-1}(S(t_{i1} | X_{i1})) + \dots + L^{-1}(S(t_{in_i} | X_{in_i}))) \quad (1-134)$$

'dir (Wienke 2011).

Daha önceki dağılım özelliklerine benzer olarak paylaşılmış zayıflık modellerinde de Gamma, Lognormal, Bileşik Poisson, PVF, Pozitif durağan model çeşitleri mevcuttur.

Paylaşılmış zayıflık modelinin bazı kısıtlılıkları vardır. Bunlardan biri, aynı kümedekilerin gözlenmeyen faktör etkilerinin aynı olduğunun varsayılmasıdır. Ancak genel olarak böyle bir varsayımın uygulamada gerçekleşmesi oldukça zordur. İkinci olarak, kümeler içindeki ilişkilerin pozitif olduğu varsayılmaktadır. Ancak bazı durumlarda kümeler içindeki bireylerin sağkalım zamanları arasında negatif yönde bir ilişki olabilmektedir. Özellikle organ nakil durumlarında böyle bir ilişki söz konusudur. Örneğin, daha uzun ömürlü olan hasta nakli beklemektedir ancak bu hastaların nakilden sonra ise yaşama şansı oldukça kısadır. Üçüncü kısıtlılık ise, küme içindeki sağkalım zamanları arasındaki bağımlılığın marjinal dağılımlara dayanmasıdır (Wienke 2011).

Paylaşılmış zayıflık modellerinde, aynı ailedeki ilişkili çiftler arasındaki korelasyonun aynı olduğu varsayılmaktadır. Ancak bu varsayım çok da geçerli değildir. Genetik araştırmaların çoğunda aileler arasındaki ilişkilerin farklı olduğu belirlenmiştir. Böyle bir durumda ise paylaşılmış zayıflık modellerinin yerine ilişkili zayıflık modellerinin kullanılması önerilmektedir.

2.2.12. İlişkili Zayıflık Modeli (Correlated Frailty Model)

İlişkili zayıflık modelleri her bireyin zayıflığının rasgele olduğunu varsaydığından karma bir model olarak tanımlanmaktadır. Herhangi bir kümedeki olay zamanlarının birbirinden bağımsız olduğu varsayılmaktadır. Herhangi bir kümedeki tüm bireyler için genel olmayan ancak birbiriyle ilişkili zayıflık değişkenlerinin olduğu koşullu bağımsız bir modeldir. Dolayısı ile olay zamanları arasındaki bağımlılıklara karşı da duyarlı bir modeldir. Paylaşılmış zayıflık modeli, zayıflıklar arasındaki korelasyonun bire eşit olduğu durumda ilişkili zayıflık modelinin özel bir durumdur. İlişkili zayıflık modelinde zayıflık vektörünün dağılımında hem grup hem de birey varyasyonları dâhil edilmektedir. Dolayısı ile genetik araştırmalarda bu modelden yararlanılmaktadır. Bu tür çalışmalarda zayıflıklar arasındaki korelasyon önemlilik arz etmektedir. Özellikle genetik araştırmalarda genlerin mi çevrenin mi hastalığa veya ölüme etkisinin daha çok olduğu araştırılmaktadır. Bu modeller daha çok aile ve ikiz olay zamanı veri setlerinde kullanılmaktadır (Wienke 2011).

Model farklı kümeler arasındaki gözlemlerin bağımsız olduğunu varsaymaktadır. Eğer zayıflık değişkenlerinin varyansı sifıra eşitse, kümelerdeki olay zamanları arasında

bağımsızlık söz konusu olacaktır. İlişkili zayıflık modeli, bağımlı yarışan risklerin modellenmesinde de kullanılmaktadır. Model için farklı dağılım çeşitleri mevcuttur. Ancak Lognormal ilişkili zayıflık modeli kümelerdeki olay zamanları arasındaki bağımlılık için oldukça geniş bir aralık imkânı sağlamasından dolayı esneklik sağlamaktadır. Gamma ilişkili zayıflık modeli ise olay zamanları arasında sadece pozitif bağımlılıkları dikkate almaktadır (Wienke 2011).

Modelin koşullu sağkalım fonksiyonu,

$$S(t_1, t_2 | Z_1, Z_2) = S_1(t_1 | Z_1) S_2(t_2 | Z_2) = e^{-Z_1 M_{01}(t_1)} e^{-Z_2 M_{02}(t_2)} \quad (1-135)$$

olarak ifade edilmektedir. Z_1 ve Z_2 , ilişkili iki zayıflık terimidir. (Z_1, Z_2) vektörünün dağılımının belirlenmesi gerekmektedir. Daha sonrasında ise modeldeki olay zamanları arasındaki ilişkilerin yapıları belirlenmelidir.

Modelde her bir bireyin zayıflığı rölatif riskin bir ölçüsü olarak tanımlanmaktadır. Zayıflıkların temel hazard fonksiyonu üzerine çarpımsal etkisi vardır ve gözlem çiftleri (sağ ve sol böbrek yetmezliği gibi olay) arasında koşullu bağımsızlık vardır. Dolayısı ile i . çiftte ($i=1, \dots, n$) j . bireyin hazardı,

$$\mu(t | X_{ij}, Z_{ij}) = Z_{ij} \mu_{0j}(t) e^{\beta' X_{ij}} \quad (1-136)$$

ile belirlenmektedir. X_{ij} , gözlemlenen açıklayıcı değişkenler vektörü iken, β açıklayıcı değişkenlerin etkilerini gösteren regresyon parametrelerinin vektörüdür. $\mu_{0j}(\cdot)$ ise temel hazard fonksiyonunu, Z_{ij} zayıflıklarına ait vektörü ifade etmektedir (Wienke 2011).

İkili değişken varlığında ilişkili zayıflık modellerinde, iki boyutlu zayıflık vektörlerinin (Z_{i1}, Z_{i2}) bileşik dağılımlarından yararlanılmaktadır. İki zayıflık birbirinden bağımsız ise, olay zamanları da birbirinden bağımsızdır ve modelde herhangi bir küme varlığından bahsedilemez. İki zayıflık birbirine eşit ise, paylaşımlı zayıflık modeli elde edilmektedir.

Model tahminleri olabilirlik fonksiyonları yardımıyla elde edilmektedir. i . çiftte ($i=1, \dots, n$) j . bireyin ($j=1, 2$) sansür durumunu gösteren δ_{ij} terimi bir değerine eşit ise bireyin ilgilenilen olay ile karşılaştığı, sıfır değerine eşit ise bireyin ilgilenilen olay ile

karşılaşmadığı anlamına gelmektedir. i . çiftte ($i=1, \dots, n$) j . bireyin koşullu sağkalım fonksiyonu,

$$S(t|X_{ij}, Z_{ij}) = e^{Z_{ij}M_{0j}(t)e^{\beta'X_{ij}}} \quad (1-137)$$

olmak üzere, $M_{0j}(t)$ ise birikimli temel hazard fonksiyonudur. i . çiftte ($i=1, \dots, n$) j . bireyin ($j=1, 2$) koşullu olabilirliğe katkısı,

$$(Z_{ij}\mu_{0j}(t_{ij})e^{\beta'X_{ij}})^{\delta_{ij}} e^{Z_{ij}M_{0j}(t_{ij})e^{\beta'X_{ij}}} \quad (1-138)$$

t_{ij} , i çiftindeki j . bireyin gözlem zamanıdır. Marjinal olabilirlik fonksiyonu ise,

$$\prod_{i=1}^n \int_{R^+} \int_{R^+} (z_{i1}\mu_{01}(t_{i1})e^{\beta'X_{i1}})^{\delta_{i1}} e^{z_{i1}M_{01}(t_{i1})e^{\beta'X_{i1}}} * (z_{i2}\mu_{02}(t_{i2})e^{\beta'X_{i2}})^{\delta_{i2}} e^{z_{i2}M_{02}(t_{i2})e^{\beta'X_{i2}}} f(z_{i1}, z_{i2}) dz_{i1} dz_{i2} \quad (1-139)$$

ile ifade edilmektedir. $f(\cdot, \cdot)$ karşılık gelen zayıflık dağılımının olasılık yoğunluk fonksiyonudur. Paylaşılmış modele benzer olarak, ilişkili zayıflık modellerinde de zayıflığın dağılımına bağlı olarak Gamma, Lognormal, Bileşik Poisson gibi ilişkili zayıflık modelleri mevcuttur (Wienke 2011).

2.2.13. İyileşme Zayıflık Modeli (Cure Frailty Model)

Sağkalım verilerinin analizinde sıklıkla Cox regresyon analizinden yararlanılmaktadır. Ancak bu analizin temel varsayımlarından biri kişilerin ilgilenilen olayı yaşayacağı ve yaşamaya duyarlı olmalarıdır. İzlem süreci gereğinden çok uzunsa, kişiler ilgilenilen olayı yaşayacaktır. Sonuçta ilgilenilen olay örneğin ölüm ise, her kişi ölüm ile karşılaşacaktır. Ancak yeteri miktarda bir gözlem süresinden sonra çalışmadaki bazı kişiler ilgilenilen olayı yaşamamaktadır. Bu bireyler iyileşmiş olarak tanımlanmaktadır. Buradaki iyileşme terimi hastanın klinik olarak tam bir iyileşme göstermesi değil, uzun sürede sağkalmasıdır. Dolayısı ile veri setinde iyileşen bireyler söz konusu olduğu durumlarda Boag ve Berkson, Gage tarafından önerilen iyileşme fonksiyonu kullanılmaktadır. Bu modellerde en önemli nokta yeteri kadar bir izlem süresinin olmasıdır. İyileşen bireylerin ilgilenilen olayı yaşamayacağı belirlendiğinden, bu bireylerin sağkalım zamanı sonsuz olarak tanımlanmaktadır (Rondeau 2010).

Kanser tedavilerindeki ilerlemeler, istatistiksel araştırmalar arasında yer alan iyileşme modellerinin geliştirilmesine neden olmuştur. Sağkalım verilerinin birçoğunun analizi genel sağkalım veya progresyonsuz sağkalıma dayanmaktadır. Ancak hastaların birçoğu uzun dönemde ilgilenilen olayı yaşamamış olabilir. Dolayısı ile iyileşme modelleri, uzun dönemde sağkalan ve kalmayan hastalar arasındaki heterojenliği araştırmak için kullanılabilecek bir yöntemdir. Belirli bir veri setinin uzun vadede hayatta kalanların bir alt kümesine sahip olup olmayacağını tanımlamanın basit bir yolu sağkalım eğrisine bakmaktır. Sağkalım eğrisi çalışmanın sonunda belirli bir noktadan sonra durağanlığa sahipse iyileşme modeli, verileri analiz etmek için uygun ve kullanışlı bir yol olabilmektedir. Klinik anlamda iyileşme hastalığın ve semptomlarının tamamen yok olmasıdır. Ancak istatistik anlamda iyileşme ise, çalışmadaki ölüm oranının beklenen ölüm oranına eşit olması ile sağlanmaktadır (Boag 1949; Berkson ve Gage 1952; Othus ve ark. 2012; Kara ve Tutkun 2017).

İyileşme modelleri veri setinde sonuç olayını yaşamayacak bireyler olduğunda iyileşme olasılıklarını da dikkate alarak sağkalım modellemelerini oluşturmaktadır. Karma iyileşme modeli ilk olarak Boag ve Gage tarafından önerilmiştir. Geliştirdikleri modelde hem tedaviden iyileşenlerin hızı hem de iyileşmeyen hastaların sağkalım hızları aynı anda tahmin edilebilir hale gelmiştir (Boag 1949; Berkson ve Gage 1952; Cai ve ark. 2012).

İyileşme zayıflık modellerinde bireyler arasındaki korelasyonlar dikkate alınarak rasgele etkiler incelenmektedir. İyileşme, zayıflık modelinde $\pi(\cdot)$ daha öncesinde gelişen olay sonrasında yeni bir olayın gelişme olasılığıdır. $\pi(\cdot)$, $\pi_{ij}(Z_{ij}^*|\omega_i)$ ile modellenmektedir ve fonksiyonlar aşağıdaki gibi elde edilmektedir.

$$S(t_{ij}|\omega_i) = 1 - \pi_{ij}(Z_{ij}^*|\omega_i) + \pi_{ij}(Z_{ij}^*|\omega_i) \exp(-\Lambda_0(t_{ij}|U_{ij} = 1) \exp(\beta'Z_{ij} + \omega_i)) \quad (1-140)$$

$$\pi_{ij}(Z_{ij}^*|\omega_i) = P(U_{ij} = 1|Z_{ij}^*) = \frac{\exp(b'Z_{ij}^* + \alpha\omega_i)}{(1 + \exp(b'Z_{ij}^* + \alpha\omega_i))}, \omega_i \sim N(0; \theta^2) \quad (1-141)$$

α terimi, zayıflık sebebiyle iyileşme fonksiyonu ile tekrarlayan olayın hızı arasındaki ilişkiyi tanımlamaktadır. $\alpha=0$ durumunda, iyileşme fonksiyonu ve tekrarlayan olaylar birbirinden bağımsızdır. $\alpha=1$ durumunda, zayıflığın etkisi yeni bir olayın gelişme olasılığı için ve tekrarlama riski için aynıdır. $\alpha>0$ durumunda ise, tekrarlama hızı ile

iyileşme fonksiyonu pozitif ilişkilidir. Zayıflık terimi ne kadar büyükse yeni tekrarın veya yeni bir olayın gerçekleşme olasılığı da o kadar fazladır. Modeldeki parametreleri tahmin etmek için maksimum olabilirlik yöntemi, parçalı sabit temel hazard fonksiyonlarını kullanmaktadır. Temel hazard fonksiyonu $\lambda_0(t^*|U_{ij} = 1)$ olarak tanımlandığından maksimum olabilirlik tahminleri Newton Raphson iteratif süreci ve Gauss kuadratik entegresyon tekniği ile elde edilmektedir. Asimptotik olarak standart hatalar ise, gözlenen bilgi matrisinin tersi alınarak elde edilmektedir (Rondeau 2010).

2.2.14. Birleşik Zayıflık Modeli (Joint Frailty Model)

Kanserden meydana gelen ölüm olayına etki eden faktörler değerlendirilirken çalışmaların bir çoğunda tekrarlanan olaylar (nüks veya metastaz) ve sonlanma olayı (ölüm) birlikte dikkate alınmamaktadır ve bağımsızmış gibi modeller oluşturulmaktadır. Ancak birçok kanser türünde nüks/metastaz olayları ölüm ile ilişkilidir. Hastanın nüks/metastaz olayını yaşaması genellikle ölüm riskini artırmakta ve genellikle iki olay arasında pozitif bir ilişki söz konusudur. Nüks/metastaz gerçekleşen hastaların ölüm ile karşılaşma riski, nüks gerçekleşmeyen hastalara göre daha yüksek olması beklenmektedir. Dolayısıyla nüks/metastaz olayını yaşayan bireyler ölüm olayını yaşama konusunda daha zayıftırlar ve bireyler arasında heterojenlikler söz konusudur. Bu heterojenlikler hem nüks/metastaz durumlarındaki farklılıklardan hem de gözlenemeyen faktörlerin etkilerinden kaynaklanmaktadır. Ayrıca eğer ölüm gerçekleşmiş ise kişinin nüks veya metastaz olaylarını yaşama imkanı yoktur. Nüks/metastaz ve ölüm olaylarını birbirinden bağımsız gibi değerlendirmek parametre sonuçlarının yanıltıcı olmasına neden olacaktır. Bu dezavantajları ortadan kaldırmak amacıyla birleşik zayıflık modeli önerilmektedir. Birleşik zayıflık modeli hem bireyler arasındaki heterojenlikleri ve hem de tekrarlayan olaylar ile ölüm arasındaki bağımlılığı dikkate alarak faktörlerin etkilerini incelemektedir. Bu modelin kullanılması ile parametre tahminlerinde oluşabilecek yanıltıcılıklar da ortadan kaldırılmaktadır (Rondeau ve ark 2007; Mauguen 2015).

Birleşik zayıflık modellerinin uygulandığı bir veri seti için veriler belirli bir zaman aralığında sürekli olarak ölçülmektedir. Dolayısıyla aynı bireylerden birden fazla ölçüm elde edilmektedir. Bireyin zaman içerisinde karşılaştığı tekrarlayan olay/lar ve sonlanma zamanı birbiri ile ilişkilidir ki bu durumda elde edilecek hata terimleri de bağımlı olacaktır. Böyle bir durumda ise parametrik varsayımlar sağlanmamakta, yarı parametrik yöntemlere başvurulmaktadır. Koşullar sağlanmadığı için parametrik

olmayan yöntemlere başvurulamamaktadır. Çünkü bu yöntemler de bağımsız değişken sayısının ikiden fazla olmaması gerekmektedir. Tahminler yapılırken yarı parametrik bir yaklaşım olan cezalandırılmış (penalized) olabilirlik tahmini yaklaşımı kullanılmaktadır. Yöntemde hazard fonksiyonu için yumuşatma parametresi kullanılmaktadır. Hazard fonksiyonunu yumuşatma (smoothing) işlemi, olay dışındaki zamanlarda hazard fonksiyonu sıfır olması sebebiyle uygulanmakta ve birçok noktada oluşan karmaşık yapıda olan eğriler için uygun modellerin elde edilmesine imkan sağlamaktadır. Ancak yumuşatma parametresinin seçimi de modellerin oluşturulmasında ve tahmin güçlerinde oldukça önemlidir. Hazard fonksiyonunu yumuşatma ile ilgili öncü bilgiyi ifade etmenin bir yolu ise olabilirliği bir terim ile cezalandırmaktır. Cezalandırılmış log olabilirliğin maksimize edilmesi ise maksimum cezalandırılmış olabilirlik tahmin edici ile tanımlanmaktadır. Ancak bu tahmin edicilerin elde edilmesi oldukça zordur ve açık değildir. Bu tahminleri elde etmek amacıyla ise eğrisel çizgiler (spline) kullanılarak tahminler yaklaşık bir şekilde hesaplanmaktadır. Çalışmada paket programı içerisinde kullanılan eğrisel çizgi yöntemi ise kübik M-spline yöntemidir. Bu yöntem B-spline yönteminin bir çeşidi olup belirli bir aralık içerisinde doğrusal olarak birleştirilen parçalı polinom fonksiyonlardır. Spline'lar için belirlenen aralıklar ise iki şekilde oluşturulabilmektedir. İlk olarak spline denilen yöntemde zaman aralıkları her biri eşit olacak şekilde bölünerek parametre sonuçları elde edilirken, spline-per yönteminde ise zaman aralıkları her alt ara birimde aynı sayıda olay (event) gözlenecek şekilde parametre sonuçları elde edilmektedir (Tabakan 2009; Commenges ve Jacqmin-Gadda 2015).

Birleşik zayıflık modelinde, i ($i=1, \dots, N$) bireyi için j . tekrar zamanları (event), sansür ve ölüm zamanları sırasıyla X_{ij} , C_i ve D_i olarak ifade edilmektedir. Her birey için izlem süresi, $\min(X_{ij}, C_i, D_i)$ olmak üzere X_{ij} gözlendiğinde δ_{ij} ($\delta_{ij} = I_{(T_{ij}=X_{ij})}, I_{(.)}$ gösterge fonksiyonunu ifade etmektedir) tekrar zamanları için ölümün varlığını veya yokluğunu gösteren gösterge değişkenidir. Ölüm durumunda bir, yaşanılan durumda ise sıfır değerini almaktadır. i bireyi için son izlenme süresi ise T_i^* olarak tanımlanmaktadır ($T_i^* = \min(C_i, D_i)$) ve $\delta_i^* = I_{(T_i^*=X_i)}$ 'dir. Dolayısı ile gerçekten gözlenenler ($T_{ij}, \delta_{ij}, \delta_i^*$)'dir. $N_i^{R*}(t)$, i . birey için $(0, t]$ zaman aralığındaki tekrarlayan olayların gerçek sayısıdır. Ancak sansür durumu söz konusu olduğu için, $N_i^{R*}(t)$ 'nin gerçek değerinin gözlenmesi imkânsızdır. Benzer şekilde gözlenen ölüm göstergesi $N_i^D(t)$ olmak üzere gerçek ölüm gösterge değişkeni ise $N_i^{D*}(t) = I(D_i \leq t)$ 'dir. Ayrıca bireyin t zamanında

izlem altında olup olmadığını belirten süreç $Y_i(t) = I(T_i^* \geq t)$ ile tanımlanmaktadır. $[t, t+dt)$ zaman aralığında i . birey için meydana gelen tekrar sayısı $dN_i^{R*}(t) = N_i^{R*}((t+dt)^-) - N_i^{R*}(t^-)$ ve $dN_i^R(t) = Y_i(t) dN_i^{R*}(t)$ 'dir (Rondeau ve ark 2007).

F_t , tüm gözlenen veriyi ve gözlenmeyen zayıflığı ω içerecek şekilde $F_t = \sigma\{Y_i(u), N_i^R(u), N_i^D(u), Z_i(u), 0 \leq u \leq t, \omega_i, i = 1, \dots, n\}$ ile tanımlanmaktadır. F_t herhangi bir bireyin t zamanına kadar ki tüm geçmişini içermektedir

Modelin varsayımları ise aşağıdaki gibidir (Rondeau ve ark. 2007).

1. Tekrarlayan olayların, ölümün ve sansürleme sürecinin sürekli olduğu varsayılmaktadır. Tekrarlayan olaylar ve ölüm aynı anda gerçekleşmemektedir.
2. $N_i^{R*}(t)$, D_i zamanından sonra sabit iken, C_i 'den sonra artış gösterebilmektedir. Bu varsayım, ölümün yeni tekrarlayan olayların meydana gelmesini engellediği anlamına gelmektedir.
3. $(F_t)_{t \geq 0}$ filtrasyonunda t zamanında, $r_i(t)dt = dR_i(t) = P(dN_i^{R*}(t) = 1 | Z_i(t), \omega_i, D_i \geq t)$ kullanıldığında $Y_i(t)r_i(t)$ tekrarlayan olay sürecinin yoğunluğudur. Hayatta kalanlar için tekrarlayan olayın oranı tanımlanırken, $P(dN_i^R(t) = 1 | F_{t-}) = Y_i^{(t)} dR_i(t) = Y_i(t)r_i(t)dt$ bağımsız sansürün olduğu varsayılmaktadır.
4. t zamanında ölüm yoğunluk süreci $Y_i(t)\lambda_i(t)$ olarak tanımlanmakta ve $\lambda_i(t)dt = d\Lambda_i(t) = P(dN_i^D(t) = 1 | Z_i(t), \omega_i, D_i \geq t)$ zayıflığı kullanılmaktadır. Dolayısı ile ölüm için bağımsız sansürleme $P(dN_i^D(t) = 1 | F_{t-}) = Y_i(t)d\Lambda_i(t) = Y_i(t)\lambda_i(t)dt$ eşitliğini sağlaması gerekmektedir.

Tekrarlayan olay ($r_i(\cdot)$) ve ölüm ($\lambda_i(\cdot)$) için hazard fonksiyonu birleşik modeller ile aşağıdaki gibi ifade edilmektedir (olaya kadar geçen süre için (calendar times, time to event)).

$$r_i(t|\omega_i) = \omega_i r_0(t) \exp(\beta_1' Z_i(t)) = \omega_i r_i(t) \quad (1-142)$$

$$\lambda_i(t|\omega_i) = \omega_i^\alpha \lambda_0(t) \exp(\beta_2' Z_i(t)) = \omega_i^\alpha \lambda_i(t) \quad (1-143)$$

Yukarıdaki modelde görüldüğü gibi tekrarlayan olayın ve ölüm zamanı için açıklayıcı değişkenlerin etkileri farklılık gösterebilmektedir. Ayrıca iki model için farklı

açıklayıcı değişkenler kullanılabilir. Bireylerin geçmiş olay hikâyesine ve hayatta kalma durumuna bağlı olarak tekrarlayan olayların ve ölüm olayının anlık oluşma olasılığı açısından β_1 ve β_2 parametreleri yorumlanabilmektedir. Model ayrıca zamana bağlı açıklayıcı değişkenler ile yansız tahminler üretmektedir.

Birleşik zayıflık modelinde rasgele etkilerin ω_i (zayıflıkların) birbirinden bağımsız olduğu varsayılmaktadır. Gözlenemeyen ω_i etkisi hem tekrarlayan olayın zamanını hem de ölüm zamanını etkilediği için, T_i^* ve T_{ij} arasındaki bağımlılık $Z_i(t)$ 'ye bağlıdır. Zayıflık parametresi (ω_i), veri setindeki heterojenlikleri dikkate alarak gözlenemeyen kovaryatlar ile ilişkilendirmektedir (Rondeau ve ark. 2007).

Modelde, $\alpha=0$ eşitliği test edilmektedir. $\alpha=0$ olması, $\lambda_i(t)$ 'nin ω_i 'ye bağlı olmaması ve bu sebeple de tekrarlayan olayın hızının bilgisinin ölüm için bilgilendirici olmaması anlamına gelmektedir. Bir başka ifadeyle kovaryatlara bağlı olarak $\lambda_i(t)$ ve $r_i(t)$ hızları birbiriyle ilişkili değildir. $\alpha = 1$ 'a eşit olduğunda, zayıflığın etkisi tekrarlayan olaylar ve ölüm için aynıdır. $\alpha > 0$ olduğunda ise, tekrarlayan olayın hızı ile ölüm hızı arasında pozitif bir ilişki vardır. Bu durumda yüksek zayıflık, yüksek tekrarlama riski ve daha yüksek ölüm riski ile sonuçlanacaktır (Rondeau ve ark. 2007).

Tekrarlayan olaylar arasındaki zaman aralığına (gap times (time-between-events)) göre ise, tekrarlayan olaylar için T_{ij} değeri $S_{ij} = T_{ij} - T_{ij-1}$ ($T_{i0} = 0$) ile yer değiştirir ve birleşik modelde hazard fonksiyonları aşağıdaki gibi elde edilmektedir.

$$r_i(s|\omega_i) = \omega_i r_0(s) \exp(\beta_1' Z_i(t)) = \omega_i r_i(s) \quad (1-144)$$

$$\lambda_i(s|\omega_i) = \omega_i^\alpha \lambda_0(t) \exp(\beta_2' Z_i(t)) = \omega_i^\alpha \lambda_i(t) \quad (1-145)$$

Tekrarlayan olayın hazard fonksiyonu ($r_0(t)$) ve ölüm zamanları için temel hazard fonksiyonu ($\lambda_0(t)$), β , α ve θ parametreleri için yarı parametrik cezalandırılmış olabilirlik yöntemi kullanılmaktadır (Rondeau ve ark. 2007).

Çoğu durumda, hazard fonksiyonlarını parça parça modellemek genellikle gerçekçi bir yaklaşım değildir. Bu tanımlamaları kullanmak için olabilirliği kaba fonksiyonlar için büyük değerleri olan bir terimle cezalandırılmaktadır. Dolayısıyla

cezalandırılmış log olabilirlik fonksiyonu aşağıdaki gibi tanımlanmaktadır (Rondeau ve ark. 2007).

$$pl(r_0(.), \lambda_0(.), \beta, \alpha, \theta) = l(\phi) - k_1 \int_0^\infty r_0''^2(t) dt - k_2 \int_0^\infty \lambda_0''^2(t) dt \quad (1-146)$$

$l(\lambda_0(.), \beta, \alpha, \eta)$ tam olabilirlik olarak tanımlanırken, $k \geq 0$ veri uyumu ile fonksiyonların düzgünlüğü arasındaki dengeyi gösteren pozitif bir düzeltme parametresidir. Cezalandırılmış log olabilirliğin maksimizasyonu maksimum cezalandırılmış olabilirlik tahmini (Maximum Penalized Likelihood Estimation (MPLE)) olarak tanımlanmaktadır ve parametre tahminleri $\hat{r}_0(t), \hat{\lambda}_0(t), \hat{\beta}, \hat{\alpha}$ ve $\hat{\theta}$ ile ifade edilmektedir. Varyans tahminleri için $\hat{H}^{-1}$ matrisi kullanılmaktadır. Birleşik zayıflık modellerine ait tahminler yumuşatma adı verilen yöntemler ile gerçekleştirilmektedir. Fonksiyonlarda yumuşatma yöntemi olarak spline kullanılmaktadır. $\hat{r}_0(t)$ ve $\hat{\lambda}_0(.)$ tahminleri açık bir şekilde hesaplanamamakta ancak spline'lar yardımıyla yaklaşık olarak elde edilmektedir. Spline'lar herhangi bir aralıkta fonksiyonu yaklaşık olarak belirlemek için doğrusal olarak birleştirilen parçalı polinom fonksiyonlarıdır. Splines yöntemlerinde eğer yarı parametrik hazard fonksiyonu olan ve eşit aralıkları kullanan fonksiyon kullanılırsa Splines yöntemi, yüzdeler kullanılırsa Splines-per yöntemi kullanılmaktadır. Cezalandırılmış log olabilirlik fonksiyonunda yer alan k_1 ve k_2 parametreleri, pozitif yumuşatma parametreleridir. Nüks için yumuşatma parametresi (k_1) paylaşılmış zayıflık modeli ile ölüm için yumuşatma parametresi ise (k_2) Cox regresyon ile elde edilebilmektedir. Tahmin edilen parametreler ise Newton Raphson algoritması ve dereceli azalma (steepest descent) algoritmasının kombinasyonlarından oluşan dayanıklı (robust) Marquardt algoritması ile elde edilmektedir. İntegrallerin çözümleri ise Gauss-Laguerre quadrature ile çözülmektedir (Rondeau ve ark. 2007).

2.3. Dinamik Tahmin Yöntemi

Onkolojik araştırmalarda risk tahminlerinin yapılması önemli bir konu haline gelmiştir. Son yıllarda kişiselleştirilmiş tedavi yöntemlerine ilgi oldukça fazladır. Hastalara geleneksel tedavi yöntemlerinden ziyade hastanın klinik özellikleri, hastalığının seyri gibi özelliklerine bakılarak kişiselleştirilmiş tedavi yöntemleri uygulanmaktadır. Kişiselleştirilmiş tedavi yöntemlerinin uygulanabilmesi için hastaların

risk analizlerinin oluşturulması gerekmektedir. Klinik araştırmalarda son yıllarda risk tahminleri ayrı bir bilimsel araştırma konusu haline gelmiştir. Uzman kişi hastanın risklerini değerlendirerek uygun tedavi yöntemlerini ve tedavi kombinasyonlarını geliştirmek istemektedir. Oluşturulan bir tahmin modelinin doğruluğu ise, farklı popülasyonlar için ve farklı zamanlarda aynı model denendiğinde test edilmektedir. Kullanılan risk modellerinin çoğu farklı popülasyonlara uygulanamamakta ve zaman sürecinin ilerlemesini de göz önüne alarak güncellenememektedir. Dolayısı ile risk ölçütünün kalitesi tartışmalı hale gelmekte ve gerçeği ne kadar yansıttığı düşündürmektedir. Tahmin modeli, farklı popülasyonlara uygulanabildiği ve zaman içindeki değişimleri de dikkate aldığı durumda genellenebilir tahmin değerleri verecektir. Dolayısı ile zaman içindeki değişen yapıları kontrol eden ve zaman içinde kendini güncelleyebilen tahmin modellerinin kullanılması gerekmektedir. Bu modeller dinamik tahmin modelleri olarak bilinmektedir. Dinamik tahmin modelleri, zaman içinde verinin güncellenmesi ile meydana gelen değişikliklere adapte edilebilmekte dolayısıyla hasta riskinin belirlenmesinde yansız tahminler üretebilmektedir (Kimmel 2018).

Onkolojik araştırmalarda dinamik tahmin modelleri kullanılırken, hastanın nüks/metastaz geçmişleri dikkate alınmaktadır. Tahminler yapılırken üç farklı yöntem kullanılmaktadır. İlk tahmin yönteminde, hastanın kesin nüks sayısı geçmişi dikkate alınmaktadır. Bir başka ifadeyle tahmin zamanı t 'den önce j tane nüks meydana gelmiş ise j 'den fazla nüks oluşmadığı varsayılarak modellemeler yapılmaktadır. Dolayısı ile hasta $j+1$. nüksü tahmin zamanı t 'den sonra yaşayabilmektedir. Örneğin, 3 yıl önce karaciğer kanser tanısı almış bir hasta, bu süre içinde üç kere nüks olayı yaşasın. Bu hastanın bir sonraki 4 yıl içinde oluşan üç nüks olayı dikkate alınarak (3 ile 7 yıl arasındaki yıllarda) ölüm olasılığı tahmin edilmek istendiğinde bu yöntem kullanılmaktadır. İkinci tahmin yönteminde ise, tahmin zamanından önce en az j gözlenen nüks olayı olduğu varsayılarak, tahmin zamanı ile j . nüks arasında ne yaşandığı göz ardı edilerek ölüm riskleri tahmin edilmektedir. Örneğin hasta tanı konulduktan beş yıl sonra üç kere nüks olayını yaşamış olduğu varsayılsın. Hastanın 8. yılda hala yaşadığı biliniyorsa bu hastanın 10 yıl yaşama olasılığının ne olacağı sorularına üç yıldan sonra ne kadar nüks olayının yaşandığı göz ardı edilerek yanıt aranıyorsa ikinci tahmin yönteminden yararlanılmaktadır. Son tahmin yönteminde ise, nüks geçmişi sadece model tahmininde yer alan parametrelerde yer alarak kullanılmaktadır. Tahmin hesaplamalarında nüks olayı dikkate alınmamaktadır.

Örneğin yeni tanısı konulmuş bir hastanın ilk beş yılda ölme riski tahmin edilmek isteniyorsa üçüncü tahmin modelinden yararlanılmaktadır (Mauguen ve ark. 2013).

Tahmin zamanı t , tahmin penceresi (aralığı) w olmak üzere dinamik modellemede t ve $t+w$ arasındaki ölüm olasılıkları ile ilgilenilmektedir. D_i , i . bireyin ölüm zamanı ve X_{ij} , i . bireyin j . gözlemlenen nüks sayısı olmak üzere, Z_{ij}^R nüks riski ile ilişkili kovaryatları, Z_i^D ölüm ile ilişkili kovaryatları ifade etmektedir. $S_{ij}^R(t) = P(X_{ij} \geq t)$ ve $S_i^D(t) = P(D_i \geq t)$ nüks ve ölüme ait sağkalım fonksiyonları olmak üzere tüm parametreleri içeren vektör $\xi = (\lambda_0^R(\cdot), \lambda_0^D(\cdot), \beta, \alpha, \theta)$ ile ifade edilmektedir. Dinamik tahmin modellemesinde, hastanın nüks hikayesine bağlı olarak ölüm risklerinin tanımlanması gerekmektedir. Bunun için iki farklı nüks geçmişi vardır. Biri tamamlanmış nüks hikayesi iken, diğeri kısmi nüks hikayesidir. t zamanına kadar i . hastanın tamamlanmamış nüks hikayesi $H_i^{j,1}(t) = \{N_i^R(t) = j, X_{i1} < \dots < X_{ij} \leq t\}$ fonksiyonu ile tanımlanmakta iken, t zamanına kadar hastanın kısmi nüks hikayesi $H_i^{j,2}(t) = \{N_i^R(t) \geq j, X_{i1} < \dots < X_{ij} \leq t\}$ fonksiyonu ile tanımlanmaktadır. Yukarıda açıklanan tahmin yöntemlerinin hesaplama aşamaları ise aşağıdaki bölümde ayrıntılı bir şekilde gösterilmektedir (Mauguen ve ark. 2013).

2.3.1. Birinci Tahmin Yöntemi ($P^1(t, t + w, \xi)$)

Birinci tahmin yönteminde, t zamanı öncesinde kesin j tane nüks geçiren hastaların t ve $t+w$ arasındaki ölüm riskleri hesaplanmaktadır. t ve $t+w$ arasındaki ölüm olasılıkları aşağıdaki fonksiyon yardımıyla elde edilmektedir (Mauguen ve ark. 2013).

$$P^1(t, t + w, \xi) = P(D_i \leq t + w | D_i > t, H_i^{j,1}(t), Z_{ij}^R, Z_i^D, \xi) \quad (1-147)$$

$$= \int_0^\infty P(D_i \leq t + w | D_i > t, H_i^{j,1}(t), Z_{ij}^R, Z_i^D, u_i, \xi) * g(u_i | D_i > t, Z_{ij}^R, Z_i^D, \xi) du_i$$

Eşitlikte yer alan $(u_i | D_i > t, Z_{ij}^R, Z_i^D, \xi)$ terimi, u_i zayıflığının koşullu olasılık fonksiyonu olarak tanımlanmaktadır.

2.3.2. İkinci Tahmin Yöntemi ($P^2(t, t + w, \xi)$)

İkinci tahmin yönteminde ise, tahmin zamanından önce en az j gözlenen nüks olayı olduğu varsayılarak, tahmin zamanı ile j . nüks arasında ne yaşandığı göz ardı

edilerek ölüm riskleri tahmin edilmektedir. İkinci tahmin yöntemine göre ölüm riskleri aşağıdaki fonksiyon yardımıyla elde edilmektedir (Mauguen ve ark. 2013).

$$P^2(t, t + w, \xi) = P(D_i \leq t + w | D_i > t, H_i^{j,2}(t), Z_{ij}^R, Z_i^D, \xi) \quad (1-148)$$

$$= \frac{\int_0^\infty [s_i^D(t | Z_i^D, u_i, \xi) - s_i^D(t + w | Z_i^D, u_i, \xi)] (u_i)^J s_{ij}^R(X_{ij} | Z_{ij}^R, \xi, u_i) g(u_i) du_i}{\int_0^\infty s_i^D(t | Z_i^D, \xi, u_i) (u_i)^J s_{ij}^R(X_{ij} | Z_{ij}^R, \xi, u_i) g(u_i) du_i}$$

2.3.3. Üçüncü Tahmin Yöntemi ($P^3(t, t + w, \xi)$)

Nüks zamanı ve geçmişi sadece model tahmininde yer alan parametre tahminleri için kullanılırken, ölüm risklerinin tahmininde doğrudan kullanılmamaktadır. Üçüncü tahmin yöntemine göre ölüm riskleri ise aşağıdaki fonksiyon yardımıyla elde edilmektedir (Mauguen ve ark. 2013).

$$P^3(t, t + w, \xi) = P(D_i \leq t + w | D_i > t, Z_i^D, \xi) \quad (1-149)$$

$$= \frac{\int_0^\infty [s_i^D(t | Z_i^D, u_i, \xi) - s_i^D(t + w | Z_i^D, u_i, \xi)] g(u_i) du_i}{\int_0^\infty s_i^D(t | Z_i^D, \xi, u_i) g(u_i) du_i}$$

Sonraki bölümde, meme kanseri ile ilgili bilgilere ve meme kanserine etkili olan risk faktörleri ve prognostik faktörler açıklanmaktadır.

2.4. Meme Kanseri ve Etki Eden Faktörler

Dünya sağlık örgütü verilerine göre, dünyada en sık görülen kanser türleri arasında meme kanseri ikinci sırada yer almaktadır. Meme kanseri kadınlarda en sık karşılaşılan ve ölüm ile sonuçlanabilen önemli bir kanser türüdür. Geçmişten bu yana değişen çevre ve iklim koşulları ile birlikte de meme kanserine yakalanma sıklığı artmaktadır ve ilerleyen yıllarda daha da artacağı düşünülmektedir. 2018 yılındaki istatistikler incelendiğinde, kansere yakalanan bireylerin %11,6'sının meme kanseri olduğu belirlenmiştir (Bray ve ark. 2018). Ülkemizde de benzer şekilde meme kanseri sıklığı son 20 yılda iki katından fazla artış göstermektedir. Dolayısı ile riski bu kadar yüksek olan bir hastalık için de teşhis ve tedavilere yönelik çalışmalar da hızla devam etmektedir.

Meme kanseri farklı organlara sıçrayan ve tekrarlayan bir kanser türüdür. Özellikle ilk iki yıl metastaz ve nüks olasılığı oldukça yüksektir. Ancak şunu da belirtmekte fayda vardır ki erken evrede bile teşhis edilen bir hastanın tümör geçmişi yaklaşık 6-10 yıla dayanmaktadır. Meme kanseri hastalığı genellikle beyin, kemik, karaciğer, akciğer gibi uzak organlara sıçramaktadır ve hastalığın ileri evrelerinde birden fazla nüks olayı gerçekleşmektedir. Metastazın gerçekleşme durumu ile tümör boyutu arasında ise pozitif bir ilişki mevcuttur. Metastatik bir kanser olan meme kanserinin hastalığın ilerlemesinden 10-15 yıl sonrasında kemik, karaciğer gibi organlara sıçraması nadir görülen bir durum değildir. Ancak hastalık taramalar sonucunda erken teşhis edilirse, sağkalım olasılığı yüksektir ve hastalığın prognozu da iyidir. İleri evrelerde ise hastalığın tedavisi de zorlaşmaktadır. Hastaların mutlaka sıkı takip altına alınması gerekmektedir. Meme kanseri takip gerektiren bir hastalık olması sebebiyle hastalar sürekli izlem altındadır. Hastaya uygulanan tedavilerin sonrasında da hastanın uzun bir dönemde izlenmesi gerekmekte ve gerektiği durumlarda hastaya farklı tedavi yöntemleri denenmektedir. Meme kanseri için hangi tedavinin daha etkili olduğu konusunda belirgin bir sonuç yoktur. Çünkü genellikle tedaviler hastanın hikayesi değerlendirilerek farklı kombinasyonlar ile verilmektedir (Saphner ve ark. 1996; Haydaroglu 2007).

Meme kanseri için erken tanı oldukça önemlidir. Erken tanıda kişinin kendi kendine yapacağı muayene erken tanıya imkan sağlamaktadır. Kişi kendi kendine eli ile muayene ettiğinde ele kitle geliyorsa, memede boyut olarak değişiklik olmuşsa, meme başında çekilmeler, deride morluklar, yaralar olmuş veya koltuk altında şişlikler gözlenmiş ise kişinin mutlaka bir hekime başvurması gerekmektedir. Teşhisi konulan hasta için meme kanseri tedavisine başlanmaktadır. Bu tedaviler cerrahi yöntem, kemoterapi, radyoterapi, hormonal tedaviler, biyolojik tedaviler gibi yöntemler olacaktır. Tedavi seçenekleri hastanın hikayesine bağlı olarak farklı şekillerde ve birleşimlerde uygulanabilmektedir (Balcı 2017).

Meme kanserinin duktal kanser, duktal karsinoma in situ, invaziv duktal karsinom (İDK), lobüler kanser, lobüler karsinoma in situ, invaziv lobüler kanser gibi türleri vardır. Bu türler arasında en sık karşılaşılan kanser türü invaziv duktal karsinomdur. İDK, süt kanallarından duvarı aşarak, meme çevresine dahi yayılım göstermektedir. Ayrıca tedavi edilmediğinde ölüm ile sonuçlanabilmektedir (Gürtekin 2011).

Meme kanserinde uygun tedavi seçeneklerinin seçiminde ve hastanın ölüm risklerinin değerlendirilmesinde klinik evreleme süreci önemlidir. Hastalığın evrelendirilmesi tümörde yayılım olmasına ve bu yayılımın seviyesine bağlı olarak değişiklik göstermektedir. Dolayısı ile hastalığın evresi kanserin ilerleme sürecini göstermektedir. Meme kanserinde evreleme işleminde, TNM (Tümör Nod Metastaz) üçlemesi kullanılmaktadır. T, tümörün büyüklüğü ve nerede bulunduğu ile ilgili iken, N hastalığın lenf noduna sıçrayıp sıçramadığı, sıçramış ise kaç tane olduğu ile ilgilidir. M ise, hastalığın farklı organa sıçraması olayı ile değerlendirme yapmaktadır. Genel olarak bu evreleme işlemi Tablo 2-2’de gösterilmektedir. Tablo 2-2’de görüldüğü gibi, evre 1 ve 2’lerde tümör boyutu 5 cm’den küçük, en fazla üç lenf nodu olan ve metastaz olmayan hastalardan oluşmaktadır. Evre 3 ise, T’nin boyutuna bağlı olmaksızın lenf nodu sayısı ile ilgilidir. Lenf nodu sayısı 4’den büyükse, aksilla dışında başka yerde görülmüş ise, metastaz yaşanmamış ise hasta evre 3’de gruplandırılmaktadır veya tümör hacmi 5 cm’den büyük ise, cilde veya göğüs duvarına yayılım gösteren ancak uzak metastaz olmayan durumlar da hasta evre 3 olarak gruplandırılmaktadır. Metastaz olduğu durumda ise, hastanın evresi evre 4 olarak tanımlanmaktadır.

Table 2-2: Meme kanseri evreleri

Evre 1 ve 2	Evre 3	Evre 4
Tümör hacmi 5’den küçük, en fazla 3 lenf nodu, metastaz yok	1.Lenf nodu sayısı 4’den fazla, aksilla dışında başka yerlerde lenf nodu görülmesi, metastaz yok 2.Tümör hacmi 5’den büyük, cilt veya göğüs duvarına yayılım görülmesi, metastaz yok	Metastaz var

Daha detaylı olarak evreleme işlemi yapılmak istendiğinde ise Tablo 2-3’deki gibi gruplandırmalar yapılabilmektedir (Hammer ve ark. 2008).

Table 2-3: Meme kanseri evrelerinin detaylı gruplandırılması

	Tümör	Lenf Nodu	Metastaz
Evre 0	Karsinoma in situ	Lenf nodu hakkında bilgi yok	Metastaz yok
Evre 1	Tümör hacmi 2 cm'den küçük	Lenf nodu hakkında bilgi yok	Metastaz yok
Evre 2A	Primer tümör hakkında bilgi yok	Lenf nodu 1-3 arasında	Metastaz yok
	Tümör hacmi 2 cm'den küçük	Lenf nodu 1-3 arasında	Metastaz yok
	Tümör hacmi 2 ile 5 cm arasında	Lenf nodu hakkında bilgi yok	Metastaz yok
Evre 2B	Tümör hacmi 2 ile 5 cm arasında	Lenf nodu 1-3 arasında	Metastaz yok
	Tümör hacmi 5 cm'den büyük	Lenf nodu hakkında bilgi yok	Metastaz yok
Evre 3A	Primer tümör hakkında bilgi yok	Lenf nodu sayısı 4-10 arasında	Metastaz yok
	Tümör hacmi 2 cm'den küçük	Lenf nodu sayısı 4-10 arasında	Metastaz yok
	Tümör hacmi 2 ile 5 cm arasında	Lenf nodu sayısı 4-10 arasında	Metastaz yok
	Tümör hacmi 5 cm'den büyük	Lenf nodu 1-3 arasında	Metastaz yok
	Tümör hacmi 5 cm'den büyük	Lenf nodu sayısı 4-10 arasında	Metastaz yok
Evre 3B	Tümör hacmi herhangi bir boyutta ancak cilt veya göğüs duvarına yayılım görülmesi	Lenf nodu hakkında bilgi yok	Metastaz yok
	Tümör hacmi herhangi bir boyutta ancak cilt veya göğüs duvarına yayılım görülmesi	Lenf nodu 1-3 arasında	Metastaz yok
	Tümör hacmi herhangi bir boyutta ancak cilt veya göğüs duvarına yayılım görülmesi	Lenf nodu sayısı 4-10 arasında	Metastaz yok
Evre 3C	Tümör boyutunun önemi yok	Lenf nodu sayısı 10'dan fazla	Metastaz yok
Evre 4	Tümör boyutunun önemi yok	Lenf nodu sayısının önemi yok	Metastaz var

Meme kanserinde prognostik ve risk faktörlerinin bilinmesi hastalığın erken evrede teşhis edilmesinde ve hem ileri hem de erken evrelerinin tedavi aşamalarında oldukça önemli rol oynamaktadır. Meme kanserinin ortaya çıkış nedeni tam olarak bilinmemektedir. Ancak hastalığın ortaya çıkmasında birçok çevresel ve genetik faktörlerin etkisi olmaktadır.

Meme kanseri yalnızca kadınlarda görülmemektedir. Hastaların yaklaşık %1'ini erkek hastalar oluşturmaktadır. Meme kanseri için kişinin erken menarş olması, geç menopoza girmesi bir risk faktörüdür. Ayrıca kişinin doğum yapıp yapmaması, doğum yaptığı yaşı hatta doğum sayısında meme kanserine yakalanma riski açısından değerlendirilmektedir. Meme kanseri riskine etkili olabilecek bir başka faktör ise yaştır. 35 yaş ve öncesinde kişinin hastalığa yakalanma olasılığı nadir iken, yaş ilerledikçe özellikle 65 yaş sonrasında risk artmaktadır. Başlarda belirttiğimiz gibi kişinin genetik yapısı da oldukça önemli rol oynamaktadır. Eğer ailede meme kanserine yakalanma geçmişi var ise, kişinin de bu hastalığa yakalanma riski artacaktır. Bu faktörün yanı sıra kişi daha öncesinde yumurtalık, bağırsak gibi kanserlere yakalanmış ise yine meme kanserine yakalanma riski diğer bireylere göre daha fazla olacaktır. Aşırı derecede alkol tüketimi östrojen seviyesini artıracığı için fazla alkol tüketen bireylerde risk artacaktır. Sigara içenler için ise bazı çalışmalar meme kanseri ile ilişki bulurken, bazı çalışmalar sigaranın meme kanseri için bir risk faktörü olmadığını belirtmişlerdir. Kilo ise riski artıran başka bir etkidir. Menopoz sonrası kilo alımının, kişinin fiziksel aktivite yapmamasının veya aşırı obezitenin riski oldukça artırdığı belirlenmiştir (Gürtekin 2011; Tanas 2013).

Meme kanseri tedavisinde son yıllarda oldukça ilerlemeler meydana gelmiştir. Hastalığa ait mortalite oranı da ciddi şekilde azalmıştır. Ancak her hasta için hastalığın seyri farklılık göstermektedir ki bu durum hastalığın heterojen olduğunu göstermektedir. Dolayısı ile kişiye özgü tedavilerin oluşturulması son yıllarda hastalığın tedavisi açısından yardımcı olmaktadır (Öksüz 2015). Kişiye özgü tedavilerin oluşturulması bakımından ise prognostik faktörler önemlidir. Kişinin klinik evresi, histopatolojik durumu, tedavi şekilleri, tümör sayısı ve boyutları, lenf nodu ve sayıları gibi faktörler mortalite oranı üzerine oldukça etkilidir (Akbari ve ark. 2016).

3. GEREÇ VE YÖNTEM

3.1. Veri

Çalışmaya, 2009-2016 yılları arasında İstanbul Üniversitesi Onkoloji Enstitüsü’de meme kanseri sebebiyle ameliyat olan 465 hasta alınmıştır. Hastalar daha öncesinde başka bir kanser hastalığı sebebiyle tanı almamıştır. Hasta verileri geriye dönük olarak elde edilmiştir. Hastalar ameliyat oldukları tarihlerden itibaren en az takip süresi 2 ay olmak üzere, en fazla 10 yıla kadar takip edilmişlerdir. Hastalar hastaneye belirli aralıklar ile kontrole gelmişlerdir. Hastaların bazıları izleme sürecinde hiç nüks olayı yaşamamış iken bazıları ise birden fazla nüks olayı ile karşılaşmıştır. Hastaların çalışmaya dâhil edilme kriterleri, hastanın kesin olarak meme kanseri tanısı alarak ameliyattan sonra en az 2 ay izlenme süresinin olmasıdır. Çalışmada açıklayıcı değişken bilgileri eksik olan gözlemler ise analiz dışı bırakılmıştır. Çalışmadaki hastaların klinik evreleri 1, 2 ve 3’den oluşmaktadır. Evre 4 hastaları çalışmada yer almamaktadır. İzleme süresi, ameliyat olma tarihinden nüks veya ölüme kadar geçen süre olarak tanımlanmıştır. Çalışmada incelenen faktörler ise, yaş (≤ 40 , >40), Cerb-B2 (pozitif, negatif), tümör tipi (İDK, diğerleri), ameliyat sonrası lenf nodu sayısı (≤ 5 , >5), Neoadjuvan kemoterapi (NAKT) alma durumu (hayır, evet), klinik evre (evre 1-2 (evre 1, evre 2A-2B), evre 3 (evre 3A-3B)), Ki-67 (≤ 25 , >25), ameliyat çeşidi (Meme koruyucu cerrahi (MKC), Modifiye radikal mastektomi (MRM)), Hormon reseptör (ER (Östrojen reseptörü) veya/ve PR (Progesteron reseptörü) pozitifliği, ER ve PR negatif)’dir.

3.2. İstatistiksel Analiz

Meme kanserinde ölüm gerçekleşmeden önce genellikle birden fazla sonlanım noktası meydana gelmektedir. Klinik karar sürecinin etkililiğini artırmak için tüm bu farklı sonlanım noktalarının aynı anda değerlendirilmesi gerekmektedir. Meme kanserinde ölümler değerlendirilirken metastaz, nüks olaylarının ölümden bağımsız gibi değerlendirilmesi sonuçların yanlış çıkmasına neden olacaktır. Yanlılığı ortadan kaldırmak için, çalışmamızda birleşik zayıflık modellerden yararlanılmıştır. Birleşik zayıflık modelleri ile ölüm süreçleri ve nüks zamanları eş zamanlı olarak incelenmiş, nüks ve ölüm arasındaki bağımlılıklar analizlerde dikkate alınmıştır. Değerlendirmelerde ayrıca nüks olayları arasındaki bağımlılıklar da dikkate alınmaktadır. Gözlenemeyen

heterojenliğin varlığı araştırılarak, gözlenmeyen etkilerin varlığında heterojenliğin kaynağı olarak rasgele etkiler modele alınmıştır.

Modelde parametre tahminleri cezalandırılmış olabilirliği yöntemi ile elde edilmiştir. Modelde yer alan yumuşatıcı parametreler için iki farklı model kurulmuştur. Nüks için yumuşatma parametresi (k_1), paylaşılmış zayıflık modeli ile ölüm için yumuşatma parametresi ise (k_2) Cox regresyon ile elde edilmiştir. Kesim noktası (knots) sayısına ise hazard fonksiyona ait grafikler çizilerek karar verilmiştir. Kesim noktası sayısı modelde 4 ile 20 arasında değişen bir değer almaktadır. Modelde hazard fonksiyonları elde edildiğinde 10'dan sonrasında eğride değişiklikler gözlenmemiştir. Bu sebeple knots sayısı 10 olarak elde edilmiştir. Parametre tahminleri ise Newton Raphson algoritması ve steepest descent algoritmasının kombinasyonlarından oluşan robust Marquardt algoritması ile elde edilmiştir. Model uyumu Martingale artıkları değerlendirilerek analiz edilmiştir. Birleşik zayıflık modellerinde tahminleri elde etmek için hem Spline ve hem de Splines-per yöntemleri kullanılarak farklılıklar değerlendirilmiştir. Birleşik zayıflık modellerinde modeller oluşturulduktan sonra ölüm risklerinin tahmininde hastaların nüks hikayelerini değerlendiren dinamik tahmin yöntemleri kullanılmıştır. Dinamik tahmin yöntemi ile elde edilen tahminlerin değerlendirilmesinde ise Brier skor değerleri kullanılmıştır.

Çalışmadaki değişkenlere ait tanımlayıcı istatistikler ise sayı ve yüzde olarak verilmiştir. Bireyler arasındaki heterojenliğin ve nüks ile ölüm arasındaki bağımlılıkların dikkate alınmasının önemini vurgulamak adına, nüks ve ölüm birbirinden bağımsız düşünülerek ölüm riski üzerine faktörlerin etkileri Cox regresyon analizi kullanılarak incelenmiştir. Değişken seçim yöntemi olarak enter metodu kullanılmıştır. Sağkalım hızları Kaplan-Meier yöntemiyle elde edilmiştir.

İstatistik anlamlılık düzeyi olarak 0,05 alınmış ve $p < 0,05$ istatistiksel olarak anlamlı kabul edilmiş ve çalışmadaki analizler SPSS (versiyon 21) paket programı ve R.3.5.3'deki frailtypack paketi kullanılarak gerçekleştirilmiştir (Rondeau ve ark. 2012).

4. BULGULAR

Uygulama verisi olarak kullanılan ve meme kanseri tanısı nedeniyle opere edilen 465 kadının izlem süresi 2 ay ile 127 ay arasında değişmektedir. Hastaların yaş aralığı 20 ile 89 arasında olup izlem süresi boyunca meme kanseri nüksetmeyen hastaların oranı %77 (n=361), operasyon sonrası bir defa nüks geçiren hastaların oranı %18 (n=83) ve iki defa nüks geçiren hasta oranı ise %0,05 (n=21)' dir. Ayrıca hastaların %33,3'ü (n=155) izlem süresi sonuna kadar ex (ölüm) olmuştur.

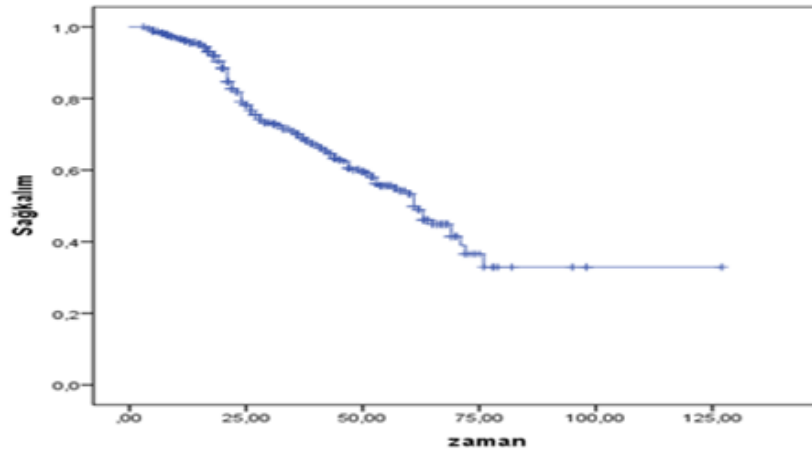
Çalışmadaki hastalara ait tanımlayıcı istatistikler Tablo 4-1'de özetlenmiştir. Tablo 4-1 incelendiğinde, çalışmadaki hastaların %24,3'ü (n=113) 40 yaş altıdır. Cerb-B2 değeri pozitif olan hasta sayısı 101 (%21,7)' dir. Ameliyattan sonra toplam lenf nodu sayısı 5'in altında olan hastaların oranı %74,4 (n=346) ve kemoterapi alan hastaların oranı ise %64,3'tür (n=299). Hastaların %82,8'sinin (n=385) tümör tipi invaziv duktal karsinom olup büyük çoğunluğu erken evre tanısı almıştır (%71,4, n=332). Ayrıca Ki-67 değeri 25'in altında olanların oranı %55,5 (n=258) ve hormon reseptörü pozitif olanların oranı %71,6 (n=333) olarak elde edilmiştir. Hastaların %50,8'i (236 hasta) meme korucu cerrahi operasyonu geçirirken, %49,2'si (229 hasta) modifiye radikal mastektomi ameliyatı geçirmiştir.

Table 4-1: Meme kanseri hastalarına ait tanımlayıcı istatistikler

Değişkenler		Sayı	%
Yaş	≤40	113	24,3
	>40	352	75,7
Cerb-B2	Negatif	364	78,3
	Pozitif	101	21,7
Ameliyattan sonra toplam lenf sayısı	≤5	346	74,4
	>5	119	25,6
Tümör tipi	İDK	385	82,8
	Diğer	80	17,2
Neoadjuvan Kemoterapi	Hayır	166	35,7
	Evet	299	64,3
Klinik Evre	1-2	332	71,4
	3	133	28,6
Ki-67	≤25	258	55,5
	>25	207	44,5

Ameliyat	MKC	236	50,8
	MRM	229	49,2
Hormon Reseptör	Negatif	132	28,4
	Pozitif	333	71,6

Çalışmadaki hastaların ölüm durumlarına göre oluşturulan sağkalım eğrisi Şekil 4-1’de yer almaktadır. Hastaların iki yıllık sağkalım oranları 0,60’ın üzerinde olduğu gözlenmiştir. İlk 75 ayda önemli düşüşler gözlenirken, 75. ay sonrasında sağkalım oranlarında önemli bir değişim meydana gelmemiştir. Bu da meme kanseri hastaları için ameliyattan sonraki yılların önemli olduğunu göstermiştir. Çalışmadaki verilere göre, hastaların uzun dönemde sağkalma olasılıkları yüksek olduğu sonucuna varılmıştır (%0,30).



Şekil 4-1: Kaplan Meier eğrisi

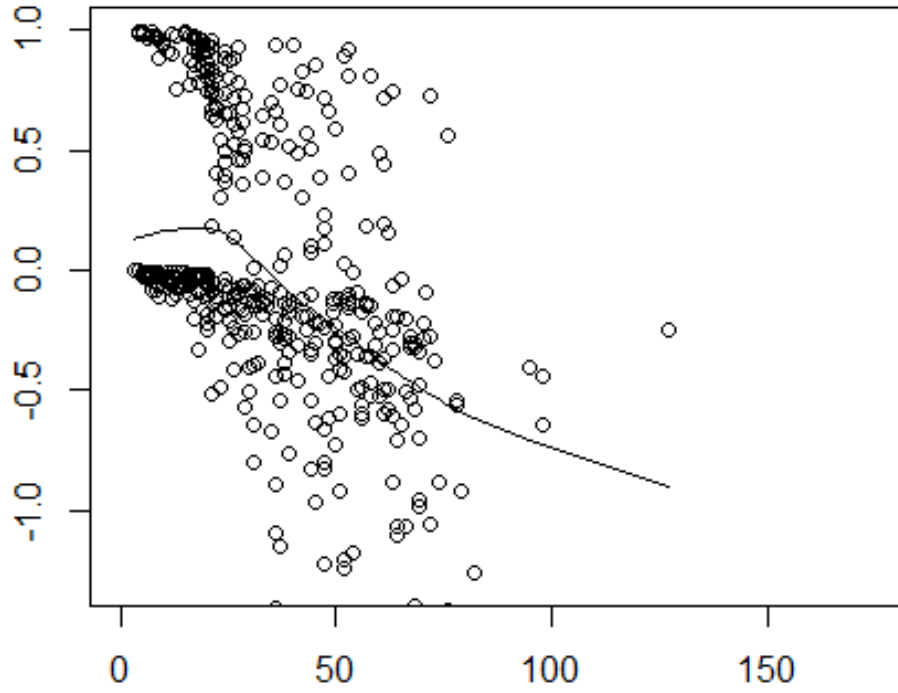
Bireyler arasındaki heterojenliği ve nüks ile ölüm arasındaki bağımlılıkların dikkate alınmasının önemini vurgulamak adına, nüks ve ölüm birbirinden bağımsız düşünülerek ölüm riski üzerine etkili olabilecek faktörler Cox regresyon analizi kullanılarak incelenmiştir. Cox regresyon analizi sonuçları Tablo 4-2’de verilmiştir. Tablo 4-2’deki sonuçlar incelendiğinde, yaş, ameliyattan sonra toplam lenf nodu sayısı, Ki-67, ameliyat tipi ve hormon reseptör değişkenleri bakımından ölüm riskleri farklılık

göstermemiştir. Bir başka ifadeyle bu değişkenlerin ölüm riski üzerine etkileri istatistiksel olarak anlamlı bulunmamıştır ($p>0,05$). Ancak ameliyattan sonraki lenf nodu sayısı değişkeninin p değeri incelendiğinde, sonuçların anlamlılığa yakın olduğu gözlenmiştir ($p=0,068$). Cerb-B2 bakımından sonuçlar değerlendirildiğinde, Cerb-B2 pozitifliğinin ölüm riski üzerine etkisinin oldukça yüksek olduğu gözlenmiştir. Cerb-B2 pozitif olan hastaların ölüm riski, negatif olan hastalara göre 2,704 kat daha fazladır ($p<0,001$). Diğer tümör tiplerine sahip bireylerin ölüm riski İDK tümörlerine sahip bireylere göre 2,120 kat fazladır ($p<0,001$). Neoadjuvan kemoterapinin etkisini incelediğimizde ise, kemoterapi almayanların alanlara göre ölüm riskinin 2,505 kat daha fazla olduğu görülmüştür ($p<0,001$). Hastaların klinik evrelerinin ölüm riski üzerine etkileri istatistiksel olarak anlamlı bulunmuştur ($p=0,003$). Hazard oranı değerlendirildiğinde ise, klinik evresi 3 olan hastaların klinik evresi 1 veya 2 olanlara göre ölüm riskinin 1,716 kat daha fazla olduğu belirlenmiştir.

Table 4-2: Cox regresyon analizine ait tahmin sonuçları

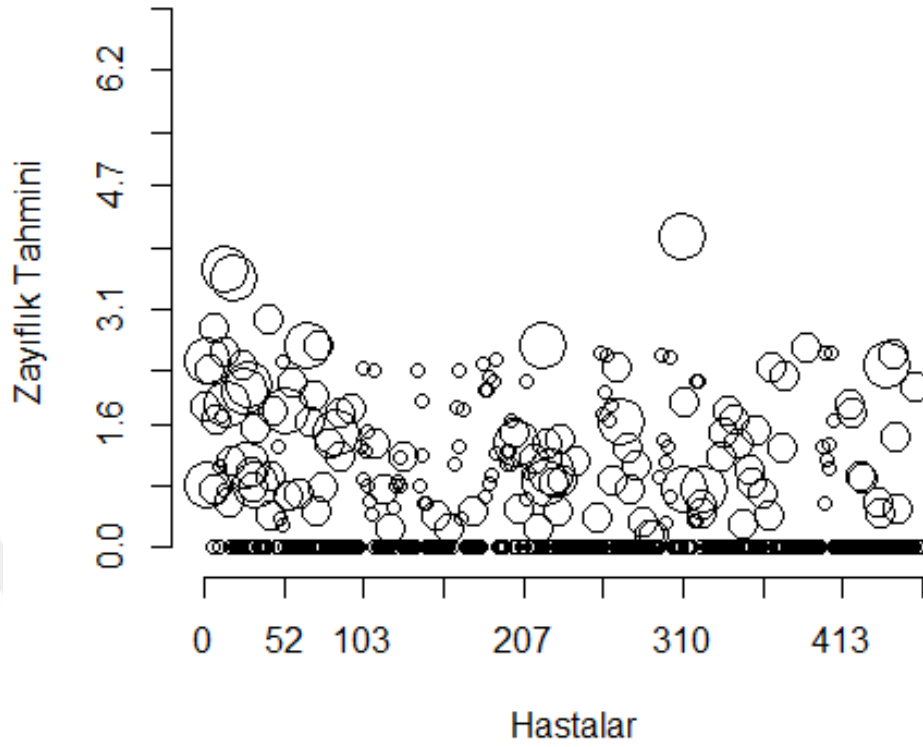
Değişkenler	Hazard Oranı	Hazard oranına ait %95 güven aralıkları		p
		Alt sınır	Üst sınır	
Yaş	1,321	0,904	1,931	0,150
Cerb-B2	2,704	1,872	3,907	<0,001
Ameliyattan sonra toplam lenf sayısı	1,425	0,974	2,086	0,068
Tümör Tipi	2,120	1,409	3,192	<0,001
Neoadjuvan Kemoterapi	2,505	1,706	3,679	<0,001
Klinik evre	1,716	1,200	2,456	0,003
Ki-67	1,341	0,937	1,918	0,109
Ameliyat Tipi	1,307	0,939	1,820	0,113
Hormon Reseptör	0,787	0,548	1,130	0,194

Meme kanserli hastalar üzerine uygulanan Cox regresyon modelinin değerlendirilmesi için Martingale artıkları incelenmiştir. Şekil 4-2’de zamana bağlı olarak çizilen Martingale artık tahmin değerlerine ait grafik yer almaktadır. Artıkların sıfır olmadığı belirlenmiştir. Bu sonuç tahminlerin başarılı olmadığını göstermektedir.



Şekil 4-2: Cox regresyon analizi sonucunda elde edilen tahminler için Martingale artıkları

Bireyler arası heterojenliği göstermek amacıyla, her bir birey için nüks sayısı ile ilişkilendirilen zayıflık tahminleri Şekil 4-3’de verilmektedir. Şekilde yer alan yuvarlak noktalar zayıflık derecesini göstermektedir. Nüks sayıları arttıkça noktaların büyüklükleri de artmaktadır. Bu sonuç, noktaları büyük olan yani daha fazla nüks geçiren hastaların sonlanma olayı (ölüm) ile karşılaşma riskinin daha fazla olduğunu göstermektedir. Bir başka ifadeyle, büyük noktalara sahip olan hastalar biyolojik olarak daha zayıf kişilerdir. Uygulama verisinde yer alan hastalar, zayıflık açısından heterojenlik göstermektedir. Bu sonuç, Cox regresyon modelinin yanı sıra tahminler verdiğinin bir başka göstergesidir. Bu durumda heterojenlikleri, ölüm ve nüks arasındaki bağımlılıkları dikkate alan birleşik zayıflık modellerinin kullanımı önerilmektedir.



Şekil 4-3: Meme kanserli hastaların zayıflık tahminleri

Bir sonraki aşamada, dikkate alınan risk faktörlerinin etkileri, birleşik zayıflık modeli kullanılarak incelenmiştir. Ancak birleşik zayıflık modeli uygulamadan önce modelde kullanılacak olan yumuşatma parametrelerine karar verilmiştir. Bu amaçla nüks için paylaşılmış zayıflık modeli ve ölüm içinde Cox modeli kullanılmıştır. Model sonucunda kappa 1 ve 2 değerleri sırasıyla 120000 ve 85000 olarak elde edilmiştir. Modellerden elde edilen yumuşatma parametreleri birleşik zayıflık modelleri için kullanıldığında ise Tablo 4-3'deki sonuçlar elde edilmiştir. Tablo 4-3'de elde edilen tahmin değerleri nüks ve ölüm arasındaki bağımlılıklar dikkate alınarak hesaplanmıştır. Tablodaki birleşik zayıflık modeli sonuçları, eşit aralıkları kullanan yarıparametrik hazard fonksiyonu olan Splines yöntemi kullanılarak elde edilmiştir. Tabloda yer alan θ parametresi, bilinen zayıflıkta nüks ve ölüm arasındaki bağımlılıkları dikkate almaktadır. Parametreler değerlendirildiğinde, ölüm ile nüks arasında pozitif anlamlı bir ilişki olduğu gözlenmiştir. Hastalarda nüks sayısı arttıkça, ölüm riskinin de anlamlı derecede arttığı belirlenmiştir. Ölümden önceki nüks olayı verilen zayıflıklar dikkate alınarak ölüm riskini etkilemektedir (rasgele etki). Ayrıca değerlendirmeler sonucunda hastaların

heterojenlik gösterdiği gözlenmiştir. Hastalar arasındaki heterojenlik istatistiksel olarak anlamlı bulunmuştur ($p < 0,001$, $\theta = 1,004$, $\alpha = 2,542$).

Faktör etkileri değerlendirildiğinde ise, Cerb-B2, tümör tipi, ameliyattan sonraki lenf nodu sayısı, neoadjuvan kemoterapi alma durumu, klinik evre ve hormon reseptör değişkenlerinin nüks oranı üzerine etkilerinin istatistiksel olarak anlamlı olduğu sonucuna varılmıştır ($p = 0,034$, $p = 0,006$, $p = 0,008$, $p < 0,001$, $p = 0,041$, $p < 0,001$). Nüks bakımından yaş grubu, Ki-67 ve ameliyat tipi değişkenlerinin ise etkileri istatistiksel olarak anlamlı bulunmamıştır ($p > 0,05$). Anlamlı çıkan faktörlerin etkileri değerlendirildiğinde ise, Cerb-B2 pozitif olan hastaların nüks riski 2,39 kat fazla bulunurken, tümör tipi diğer olanların nüks riski, tümör tipi İDK olanlara göre 2,064 kat daha fazla olduğu gözlenmiştir. Ameliyattan sonraki lenf nodu sayısı bakımından incelendiğinde ise, ameliyattan sonraki lenf nodu sayısı beşten büyük olanların nüks riski, beşten küçük olanlara göre 1,869 kat daha fazla olduğu belirlenmiştir. Neoadjuvan kemoterapi alma durumlarına göre değerlendirildiğinde ise, neoadjuvan kemoterapi almayanların alanlara göre nüks riskinin 4,373 kat daha fazla olduğu bulunmuştur. Dolayısı ile hastanın nüks riski üzerine ameliyattan sonra neoadjuvan kemoterapi almasının oldukça büyük etkisi olduğu görülmüştür. Klinik evre açısından nüks sonuçları değerlendirildiğinde ise, klinik evresi 3 olan hastaların, klinik evresi 1 veya 2 olanlara göre nüks riskinin 1,956 kat daha fazla olduğu belirlenmiştir. Hormon reseptör pozitif olanlar ise nüks riski üzerine 0,466 kat koruyuculuk özelliği taşımaktadır.

Birleşik zayıflık modeli kullanılarak ölüm riski üzerine etkili faktörler araştırıldığında; yaş, klinik evre, Ki-67, ameliyat tipi ve hormon reseptör değişkenlerinin ölüm riski üzerine etkilerinin istatistiksel olarak anlamlı olmadığı sonucuna varılmıştır ($p > 0,05$). Buna karşın, Cerb-B2, ameliyattan sonraki lenf nodu sayısı, tümör tipi, neoadjuvan kemoterapi alma durumunun ölüm riski üzerine etkileri istatistiksel olarak anlamlı bulunmuştur ($p < 0,001$, $p = 0,041$, $p = 0,008$, $p < 0,001$). Anlamlı bulunan riskler ayrıntılı olarak incelendiğinde, Cerb-B2 pozitif olan hastaların ölüm riski anlamlı düzeyde 9,911 kat yüksek bulunmuştur. Ameliyattan sonra lenf nodu sayısı beşin üzerinde olan hastaların ölüm riski, anlamlı düzeyde 2,757 kat daha fazla bulunmuştur. Operasyon sonrası neoadjuvan kemoterapi almayan hastaların ölüm riski 8,734 kat daha yüksek çıkmıştır. Bunun yanı sıra tümör tipinin İDK olması ölüm riskini anlamlı düzeyde azalttığı belirlenmiştir (Risk = 1/ 2,600).

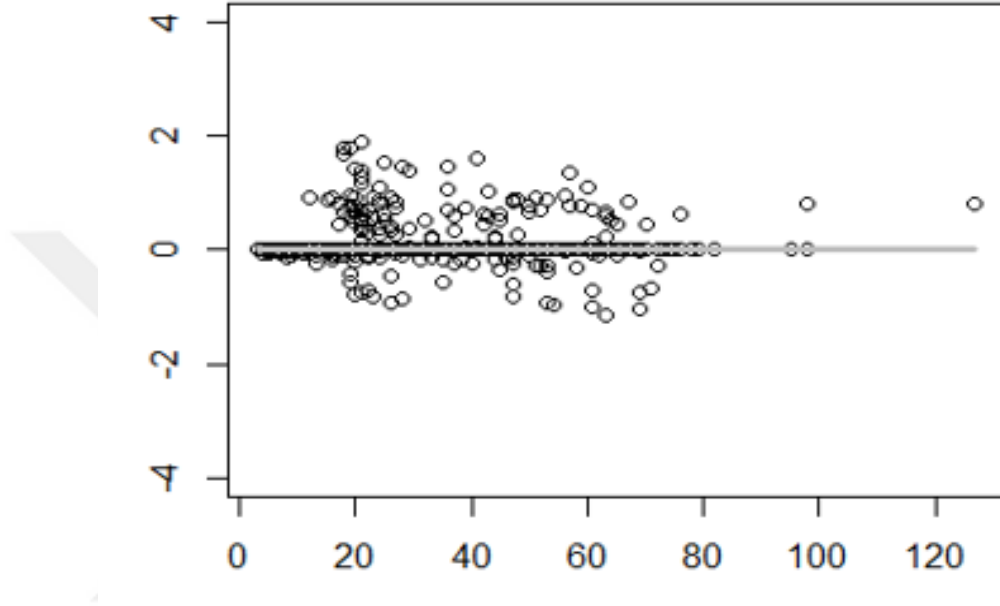
Sonuç olarak Cerb-B2, tümör tipi, ameliyat sonrası lenf nodu sayısı ve ameliyat sonrası neoadjuvan kemoterapi alma durumu hem nüks gelişimini hem de ölüm riskini anlamlı düzeyde etkilemektedir. Ayrıca lenf nodu sayısı, tümör tipi ve neoadjuvan kemoterapi tedavisinin ölüm riski üzerine etkilerinin nüks riski üzerine etkilerinden daha fazla olduğu gözlenmiştir. Nüks için ameliyat etkisi anlamlılığa yakın bulunurken, ölüm için ise Ki-67 değişkeninin etkisi anlamlılığa yakın bulunmuştur.

Table 4-3: Splines yöntemi kullanılarak birleşik zayıflık modeli ile elde edilen tahmin sonuçları

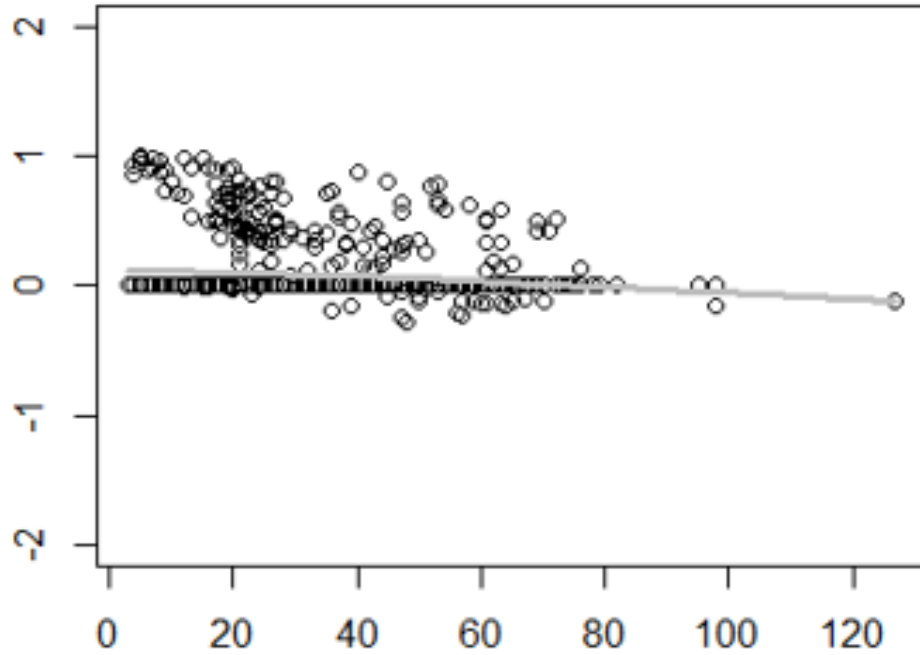
	Değişkenler	Risk	%95 Güven Aralıkları	p
Nüks	Yaş	1,013	(0,63-1,62)	0,956
	Cerb-B2	2,390	(1,06-5,26)	0,034
	Tümör tipi	2,064	(1,23-3,48)	0,006
	Ameliyattan sonra toplam lenf sayısı	1,869	(1,18-2,97)	0,008
	Neoadjuvan Kemoterapi	4,373	(2,59-7,37)	<0,001
	Klinik Evre	1,956	(1,03-3,73)	0,041
	Ki-67	1,207	(0,77-1,90)	0,412
	Ameliyat	1,494	(0,97-2,30)	0,068
	Hormon Reseptör	0,466	(0,30-0,73)	<0,001
Ölüm	Yaş	1,274	(0,63-2,59)	0,502
	Cerb-B2	9,911	(3,61-27,24)	<0,001
	Tümör Tipi	2,600	(1,04-6,51)	0,041
	Ameliyattan sonra toplam lenf sayısı	2,757	(1,30-5,85)	0,008
	Neoadjuvan Kemoterapi	8,734	(3,78-20,20)	<0,001
	Klinik Evre	1,274	(0,55-2,95)	0,571
	Ki67	1,936	(0,98-3,83)	0,058
	Ameliyat	1,739	(0,89-3,41)	0,106
	Hormon Reseptör	0,787	(0,41-1,52)	0,476
θ (Standart hata)		1,004(0,100)		<0,001
α (Standart hata)		2,542(0,447)		<0,001

Birleşik zayıflık modelinin uyum iyiliğinin değerlendirilmesinde Martingale artıkları kullanılmıştır. Martingale artıkları, gözlemlenen olay sayısının model tarafından doğru tahmin edilip edilmediğini modellemektedir. Nüks ve ölüm sonlanma zamanları için modelin gözlemlenen olay sayısını doğru tahmin edip etmediğinin incelenmesi sonucunda elde edilen grafikler Şekil 4-4 ve Şekil 4-5’de verilmiştir. Şekil 4-4 nüks tahminleri için, Şekil 4-5 ise ölüm tahminlerine ait artıkların grafiğini göstermektedir. Her iki artık grafikleri içinde, artıkların yaklaşık sıfıra eşit bir ortalamaya sahip olduğu görülmüştür. Ayrıca noktaların $y=0$ doğrusu etrafında saçıldığı belirlenmiştir. Yalnızca

ölüm için takip süresi 127 ay olan bir hasta için artıklarda biraz sapma meydana geldiği görülmüştür. Ancak bu hasta modelden çıkarıldığında da tahminlerde önemli bir değişim olmadığı gözlenmiştir. Bu sebeple hasta çalışma dışı bırakılmamıştır. Sonuç olarak, modelin uyum iyiliği değerlendirmeleri sonucunda birleşik zayıflık modelinin başarılı olduğu sonucuna varılmıştır.



Şekil 4-4: Splines yöntemi kullanıldığında nüks olayı için Martingale artıkları



Şekil 4-5: Splines yöntemi kullanıldığında ölüm olayı için Martingale artıkları

Birleşik zayıflık modellerinde farklı hazard fonksiyonlarının seçiminin sonuçları değiştirip değiştirmediği incelenmiştir. İkinci aşamada birleşik zayıflık modelleri kurulurken yüzdeleri kullanan Splines-per yöntemi kullanılmıştır. Modelde alınan parametreler bir önceki oluşturulan model ile aynıdır. Splines-per yöntemi kullanılarak birleşik zayıflık modeline ait sonuçlar Tablo 4-4’de verilmektedir. Model sonuçları değerlendirilirken ilk bakılması gereken θ ve α parametrelerinin anlamlılıklarıdır. Parametrelerin sonuçları değerlendirildiğinde her iki parametrenin de anlamlı olduğu ve pozitif olduğu görülmektedir. Bu ilişkiler ölüm ile nüks arasında pozitif anlamlı bir ilişkinin olduğunu göstermektedir. Bir başka deyişle, nüks olayı yaşayan veya nüks olayını daha çok yaşayan hastaların ölüm ile karşılaşma riski nüks olayı yaşamayanlara göre daha yüksektir ve bireyler arasında heterojenlikler söz konusudur. Dolayısı ile klasik sağkalım analiz yöntemlerinin uygulanması yanlı sonuçlara neden olacağından birleşik zayıflık modellerinin kullanımının doğru olacağına karar verilmiştir ($p < 0,001$, $\theta = 1,008$, $\alpha = 2,945$). Model sonuçları incelendiğinde, yaş ve Ki-67’nin nüks etme üzerine anlamlı etkisi olmadığı görülmüştür ($p > 0,05$). Cerb-B2, tümör tipi, ameliyattan sonraki toplam lenf sayısı, neoadjuvant kemoterapi alma durumu, klinik evre, ameliyat tipi ve hormon reseptör değişkenlerinin nüks riski üzerine anlamlı etkileri olduğu belirlenmiştir (sırasıyla p değerleri $p = 0,026$, $p = 0,004$, $p = 0,008$, $p < 0,001$, $p = 0,050$, $p = 0,050$ $p < 0,001$).

Anlamlı çıkan faktörlerin etkileri değerlendirildiğinde ise, Cerb-B2 pozitif olan hastaların ölüm riskinin, Cerb-B2 negatif olan hastalara göre riskinin 2,5 kat fazla olduğu belirlenmiştir. Tümör tipi diğer olanların nüks riski ise, tümör tipi İDK olanlara göre 2,065 kat daha fazla olduğu gözlenmiştir. Ameliyattan sonraki lenf nodu sayısı bakımından incelendiğinde ise, ameliyattan sonraki lenf nodu sayısı beşten büyük olanların nüks riski, beşten küçük olanlara göre 1,852 kat daha fazla olduğu belirlenmiştir. Neoadjuvan kemoterapi alma durumlarına göre değerlendirildiğinde ise, neoadjuvan kemoterapi almayanların alanlara göre nüks riskinin 4,211 kat daha fazla olduğu bulunmuştur. Dolayısı ile hastanın nüks riski üzerine ameliyattan sonra neoadjuvan kemoterapi almasının oldukça büyük etkisi olduğu görülmüştür. Klinik evre açısından nüks sonuçları değerlendirildiğinde ise, klinik evresi 3 olan hastaların klinik evresi 1 veya 2 olanlara göre nüks riskinin 1,889 kat daha fazla olduğu belirlemiştir. Ameliyat tipi MRM olan hastalarda nüks riski, MKC ameliyatı olanlara göre 1,521 kat daha fazladır. Hormon reseptör pozitif olma durumunun ise nüks riski üzere 0,460 kat koruyuculuğu vardır.

Ölüm riski üzerine etkili faktörler incelendiğinde; yaş, klinik evre ve hormon reseptörünün anlamlı düzeyde etkili olmadığı ($p>0,05$); buna karşın Cerb-B2, tümör tipi, ameliyattan sonraki lenf sayısı, neoadjuvant kemoterapi alma durumu, Ki-67 ve ameliyat tipi değişkenlerinin ölüm riskini anlamlı düzeyde değiştirdiği görülmüştür (sırasıyla p değerleri $p<0,001$, $p=0,017$, $p=0,007$, $p<0,001$, $p=0,046$, $p=0,044$). Anlamlı bulunan riskler ayrıntılı olarak incelendiğinde, Cerb-B2 pozitif olan hastaların ölüm riskinin, Cerb-B2 negatif olan hastalara göre 11,860 kat fazla olduğu belirlenmiştir. Cerb-B2 pozitifliğin ölüm riski üzerine etkisinin oldukça yüksek olduğu gözlenmiştir. Ameliyattan sonra lenf nodu sayısı beşin üzerinde olan hastaların ölüm riski, lenf nodu sayısı beşin altında olan hastalara göre 3,113 kat daha fazladır. Ayrıca tümör tipi İDK olmasının ise ölüm riski üzerine azaltıcı etkisi olduğu belirlenmiştir (Risk=1/2,798). Ameliyattan sonra neoadjuvan kemoterapi almayan hastaların ölüm riskinin, neoadjuvan kemoterapi alan hastalara göre 9,366 kat daha fazla olduğu gözlenmiştir. Ki-67 değeri 25'in üstünde olanların ölüm riski ise 2,029 kat daha fazla bulunmuştur. Yani kişinin Ki-67 değerinin yüksek olması meme kanseri ölüm riski üzerine etkisi anlamlı olarak elde edilmiştir. Ameliyat tipi MRM olan hastalarda nüks riski, MKC ameliyatı olanlara göre 1,991 kat daha fazladır.

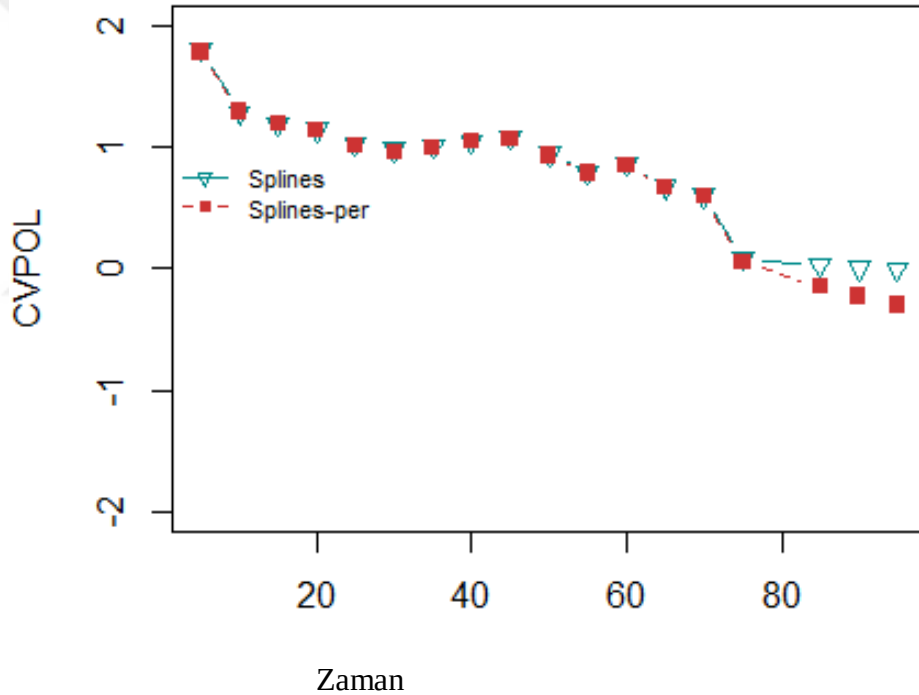
Tablo 4-3 ile Tablo 4-4 sonuçları karşılaştırıldığında ise, faktörlere ait risk değerlerinin birbirine benzer olduğu gözlenmiştir. Ancak nüks için faktör etkileri Splines yönteminde, Splines-per yöntemine göre biraz daha yüksek çıkarken, ölüm için Splines-per yöntemine ait sonuçlarda ise risk değerlerinin splines yöntemine göre biraz daha yüksek olduğu gözlenmiştir. Faktörlerin ölüm ve nüks üzerine etkilerinin anlamlılıkları bakımından değerlendirildiğinde ise sonuçların Splines ve Splines-per yöntemlerinde aynı olduğu belirlenmiştir. Yalnızca ölüm riski üzerine Ki-67 değişkeninin etkisi Splines yönteminde anlamlı bulunmaz iken, Splines-per yöntemine göre anlamlı bulunmuştur. Splines yönteminde Ki-67 değişkeninin ölüm riski üzerine etkisi anlamlılığa yakın (0,05'e yakın) bir değer aldığı görülürken, bu değer istatistiksel olarak anlamlı bulunmamış ve riske ait güven aralıkları da "1" değerini kapsamaktadır.

Table 4-4: Splines-per yöntemi kullanılarak birleşik zayıflık modeli ile elde edilen tahmin sonuçları

	Değişkenler	Risk	%95 Güven Aralıkları	p
Nüks	Yaş	1,013	(0,64-1,61)	0,955
	Cerb-B2	2,500	(1,11-5,55)	0,026
	Tümör tipi	2,065	(1,25-3,42)	0,004
	Ameliyattan sonra toplam lenf sayısı	1,852	(1,17-2,94)	0,008
	Neoadjuvan Kemoterapi	4,211	(2,49-7,11)	<0,001
	Klinik Evre	1,889	(0,99-3,60)	0,050
	Ki-67	1,201	(0,77-1,88)	0,423
	Ameliyat	1,521	(1-2,32)	0,050
	Hormon reseptör	0,460	(0,30-0,72)	<0,001
	Yaş	1,302	(0,65-2,60)	0,455
Ölüm	Cerb-B2	11,860	(5,60-25,10)	<0,001
	Tümör Tipi	2,798	(1,20-6,54)	0,017
	Ameliyattan sonra toplam lenf sayısı	3,113	(1,37-7,10)	0,007
	Neoadjuvan Kemoterapi	9,366	(4,43-19,78)	<0,001
	Klinik Evre	1,199	(0,52-2,78)	0,600
	Ki-67	2,029	(1,01-4,08)	0,046
	Ameliyat	1,991	(1,02-3,90)	0,044
	Hormon Reseptör	0,744	(0,40-1,40)	0,359
	0 (Standart hata)	1,008(0,099)		<0,001
	α (Standart hata)	2,945(0,432)		<0,001

Splines ve Splines-per yöntemi ile elde edilen birleşik zayıflık modellerinde yer alan katsayı tahminlerinin doğruluğu EPOCE (expected prognostic observed cross-entropy) tahmin edicisi ile incelenmiştir. EPOCE tahmin edicisi, birleşik zayıflık modellerinin parametre tahminlerinin doğruluklarının araştırılmasında ve iki veya daha fazla birleşik zayıflık modellerinin karşılaştırılmasında kullanılmaktadır. EPOCE, Cox regresyon analizi sonucunda elde edilen model tahminlerinin değerlendirilmesinde

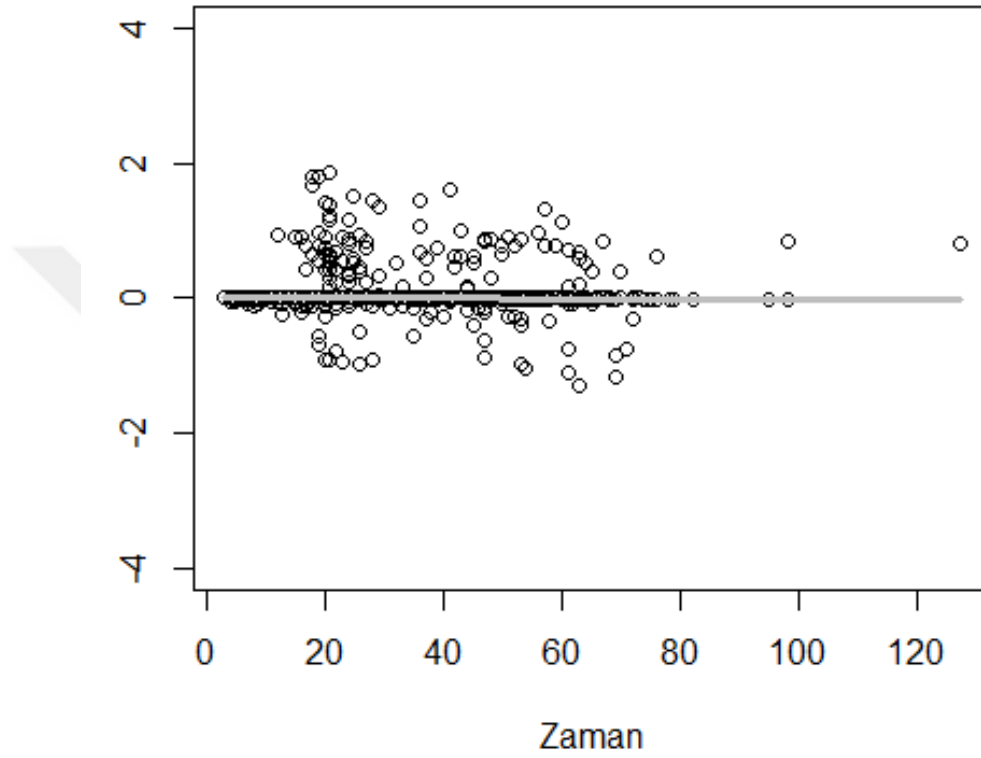
kullanılamamaktadır. EPOCE, prognostik koşullu log-olabilirliğin bir köküdür. Sonuçlarda EPOCE tahmin edicileri arasında yer alan CVPOL (cross-validated prognosis observed loss) kullanılmıştır. Splines ve Splines-per yöntemlerinin sonuçlarını gösteren Şekil 3-6 grafiği incelendiğinde yaklaşık 80.aya kadar her iki yöntemde CVPOL değerlerinin birbirine oldukça yakın olduğu görülmekte iken, yaklaşık 80.ay sonrasında Splines-per CVPOL değerinin, Splines yöntemine göre daha düşük olduğu gözlenmiştir. Dolayısı ile 80.ay ve sonrasında Splines-per yönteminin parametre sonuçlarına ait değerlerin doğruluğunun daha yüksek olduğu görülmüştür. İki yöntem arasında kısa dönemde tahmin farklılıkları gözlenmez iken, uzun dönemde yöntem sonuçlarının farklılaştığı görülmüştür.



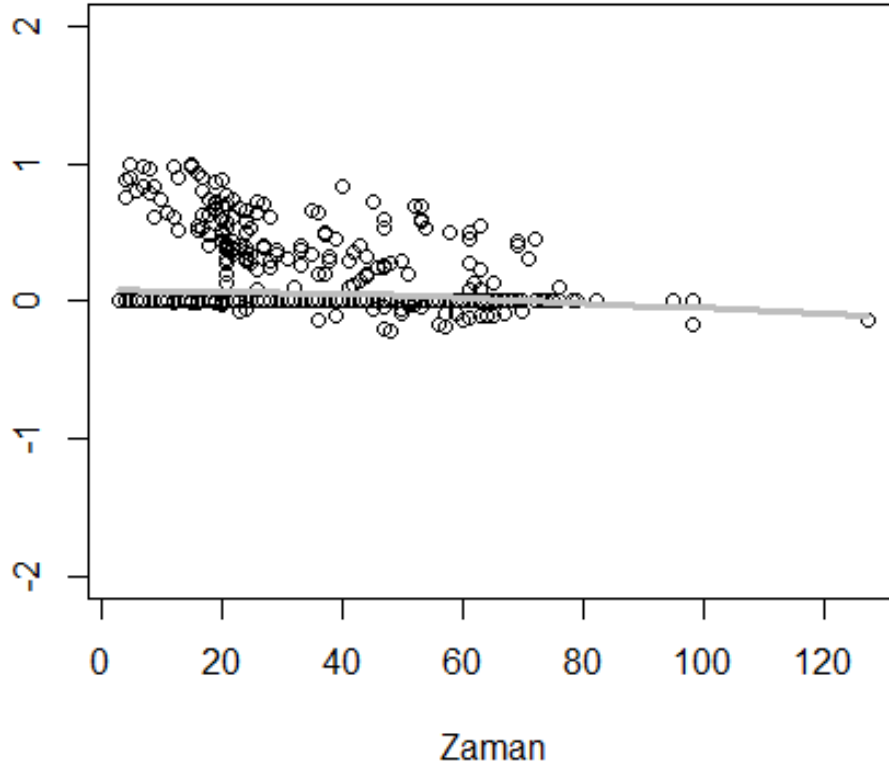
Şekil 4-6: Splines ve Splines-per yöntemlerinin karşılaştırılması

Splines-per yöntemi kullanılarak elde edilen sonuçlara ait nüks ve sağkalım süreleri için modelin gözlemlenen olay sayısını doğru tahmin edip etmediğini gösteren Martingale artıklarının incelenmesi sonucunda elde edilen grafikler Şekil 4-7 ve Şekil 4-8’de verilmiştir. Şekil 4-7 nüks tahminleri için, Şekil 4-8 ise ölüm tahminlerine ait artıkların grafiği gösterilmiştir. Her iki artık grafikleri içinde, artıkların yaklaşık sıfıra eşit bir ortalamaya sahip olduğu görülmüştür. Ayrıca noktaların $y=0$ doğrusu etrafında

saçıldığı belirlenmiştir. Bir önceki modele göre çizilen Martingale grafikleri ile sonuçların çok benzer olduğu görülmektedir. Dolayısı ile Splines ve Splines-per yöntemlerine göre kurulan birleşik zayıflık modelleri sonucunda elde edilen tahminlere ait Martingale artıkları değerleri birbirine oldukça yakın olduğu gözlenmiştir.



Şekil 4-7: Splines-per yöntemi kullanıldığında nüks olayı için Martingale artıkları



Şekil 4-8: Splines-per yöntemi kullanıldığında ölüm olayı için Martingale artıkları

Modellerin oluşturulmasının ardından, bir sonraki aşama da hastaların ölüm riskleri tahmin edilmiştir. Bilindiği üzere her bir hasta için ayrı ayrı ölüm risklerinin tahmini hastaya ait tedavi seçeneklerinden uygun olanı seçmede ve kişiye özel tedavi kararlarının alınmasında oldukça büyük önem taşımaktadır. Hastanın hastalığına ait ilerleme sürecinin bilinmesi bakımından doğru tedavi çeşitlerine karar vermek oldukça önemlidir. Kişiselleştirilmiş tıp alanındaki ilerlemeler ile birlikte her bir hasta için ayrı ayrı ölüm risklerinin tahmini oldukça önemli hale gelmiştir. Bu bağlamda yaptığımız çalışmada, ölüm risklerinin tahmin edilmesinde, bilinen tahmin yöntemlerinden ziyade dinamik tahmin yöntemi kullanılmıştır. Dinamik tahmin yöntemlerinin kullanılmasının nedeni t zamanında yapılacak olan tahminlerde t zamanından önceki hastanın bilgilerinin (bu bilgilerden en önemlisi hastanın nüks hikayesidir) kullanılması ve yeni bir olay meydana geldiğinde de tahminlerin kendini yenileyebilmesidir. Veri setinde yer alan tüm kişiler için bu riskler hesaplanabilmektedir. Yapılan tahminler için örnek teşkil etmesi bakımından çalışmamızdaki ilk beş hastaya ait tahminlere yer verilmiştir. Tahmin sonuçlarına göre üç farklı ölüm riskleri elde edilmiştir. Birinci yöntemde risk tahmini, t

zamanından önceki kesin olan j adet nüks tekrarı (exactly j recurrences) dikkate alınarak yapılmaktadır. İkinci tahmin yönteminde ise, tahmin zamanı olan t zamanından önce hastanın geçmişindeki en az j nüks olayını dikkate almakta iken, j . nüks ile tahmin zamanı arasında meydana gelen olaylar dikkate alınmamaktadır. Üçüncü tahmin yönteminde ise, nükse ait geçmiş bilgi dikkate alınmamaktadır ve yalnızca birleşik zayıflık modellerinden elde edilen parametre tahminleri kullanılarak hesaplamalar yapılmaktadır.

Gerçekleştirilen tahminler için $t=30$ ay alınmış, horizon zamanı olarak ise 35, 45, 55 ve 70 aylar alınmıştır. Horizon zamanı gelecekte belirli süreçlerin değerlendirileceği veya biteceği kabul edilen sabit bir zaman noktasıdır ve gelecekteki tahminleri elde etmede kullanılan bir aralıktır. Risk tahminleri yapılan hastaların özellikleri ise şöyledir, birinci hastanın özellikleri ise, yaşı 40'dan küçük, Cerb-B2 değeri negatif, ameliyattan sonraki lenf nodu sayısı beşden küçük, tümör tipi diğer, neoadjuvan kemoterapi tedavisi almış, klinik evre 1-2, Ki-67 değeri 25'den küçük, ameliyat tipi MKC ve hormon reseptörü negatiftir. İkinci hastanın özellikleri ise yaşı 40'dan büyük, Cerb-B2 değeri negatif, ameliyattan sonraki toplam lenf sayısı beşden fazla, tümör tipi İDK, neoadjuvan kemoterapi almış, klinik evresi 3, Ki-67 değeri 25'den küçük, ameliyat tipi MRM ve hormon reseptörü negatiftir. Üçüncü hastanın ise yaşı 40'dan büyüktür, Cerb-B2 değeri negatiftir, ameliyattan sonraki lenf sayısı beşden büyüktür, tümör tipi İDK'dır, neoadjuvan kemoterapi almıştır, klinik evresi 3'tür, Ki-67 değeri 25'in altındadır, ameliyat tipi MRM'dir ve hormon reseptörü negatiftir. Dördüncü hasta için özellikler şöyledir, yaş 40'dan büyük, Cerb-B2 değeri negatif, ameliyattan sonra lenf sayısı beşin altında, tümör tipi İDK, Neoadjuvan kemoterapi almış ve klinik evresi 1-2, Ki-67 değeri 25'in altında, ameliyat tipi MKC ve hastanın hormon reseptör durumu negatiftir. Son olarak beşinci hastanın özellikleri ise, yaşı 40'dan büyük, Cerb-B2 değeri negatif, ameliyattan sonra lenf sayısı beşden küçük, tümör tipi İDK, neoadjuvan kemoterapi almamış, klinik evresi 1-2 kategorisinde, Ki-67 değeri 25'in altında, ameliyat tipi MKC ve hastanın hormon reseptörü durumu negatiftir. Dinamik tahmin yöntemi ile beş bireye ait tahmin edilen ölüm ve nüks riskleri Tablo 4-5 ve Tablo 4-6'da gösterilmiştir. Tablo 4-5'de üç farklı tahmin yöntemleri ile değerlendirilen sonuçlar yer almaktadır ve elde edilen ölüm riskleri $t=30$ ile horizon zamanları arasında hesaplanan değerlerdir. Örneğin ilk bireyin 30 ile 35 ayları arasında ölüm riski birinci tahmin yöntemine göre %14 iken, 30 ile 45 ayları arasında ölüm riski yaklaşık %40'a kadar çıkmıştır. Üç farklı tahmin yöntemleri arasında yer alan birinci tahminde $t=30$ tahmin zamanından önceki net nüks

sayıları dikkate alınarak ölüm riskleri hesaplanmaktadır. İkinci tahmin yönteminde, tahmin zamanı $t=30$ iken, nüks sayımı ile tahmin zamanı $t=40$ ay arasındaki zamanı göz ardı ederek, $t=30$ zamanından önce ne kadar nüks gerçekleşmiş ise en az o kadar nüks olayının olduğunu dikkate alarak tahminler yapılmaktadır. Üçüncü tahmin yöntemi ise hastaların nüks sayısını doğrudan dikkate almamaktadır, bu tahmin yöntemine göre sadece parametre tahminlerinden elde edilen değerler yardımıyla riskler tahmin edilmektedir. Dolayısı ile ilk iki tahmin yöntemi daha yansız risk tahminleri üretecektir.

Table 4-5: $t=30$ durumunda dinamik tahmin ile elde edilen ölüm risklerine ait tahmin değerleri

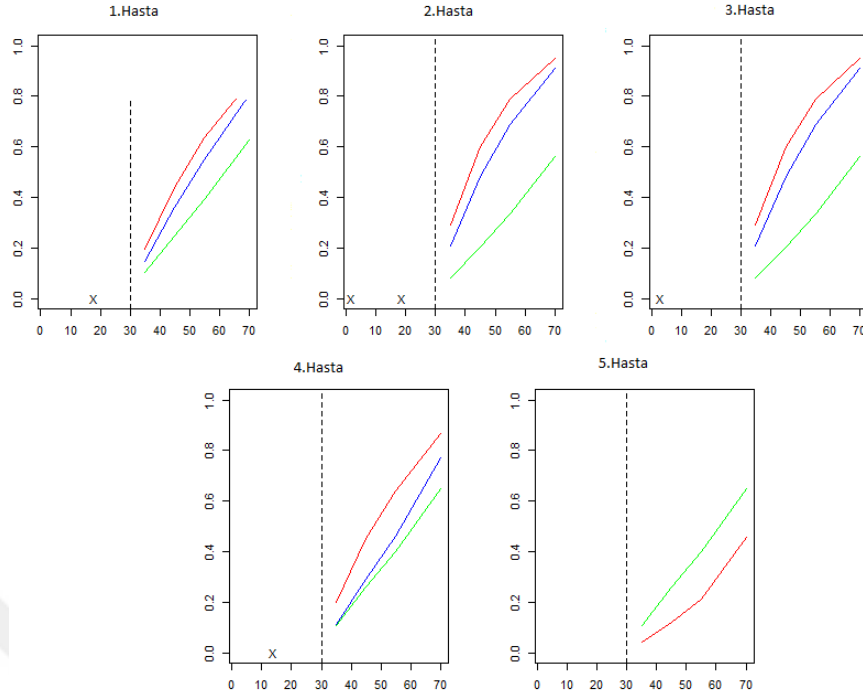
Tahmin	Hasta no	Zaman			
		35 ay	45 ay	55 ay	70 ay
Tahmin 1	1	0,147	0,359	0,550	0,860
	2	0,210	0,483	0,690	0,911
	3	0,210	0,483	0,690	0,911
	4	0,112	0,293	0,463	0,773
	5	0,042	0,120	0,214	0,459
Tahmin 2	1	0,195	0,438	0,635	0,854
	2	0,292	0,600	0,788	0,951
	3	0,291	0,599	0,787	0,951
	4	0,201	0,452	0,642	0,869
	5	0,042	0,120	0,214	0,459
Tahmin 3	1	0,1018	0,249	0,390	0,627
	2	0,0799	0,205	0,334	0,560
	3	0,0799	0,205	0,334	0,560
	4	0,1073	0,261	0,401	0,648
	5	0,1073	0,261	0,401	0,648

Tablo 4-6’da ise seçilen beş hastanın tahmin zamanı ve horizon zamanları arasındaki tahmini nüks riskleri yer almaktadır. Örneğin ilk hasta için 30 ve 70. aylar arasında nüks olayını geçirme riski 0,713 olarak tahmin edilmiştir.

Table 4-6: $t=30$ durumunda dinamik tahmin ile elde edilen yeni bir nüks olayını yaşama riskine ait tahmin değerleri

Hasta No	Zaman			
	35 ay	45 ay	55 ay	70 ay
1	0,181	0,418	0,536	0,713
2	0,238	0,524	0,654	0,825
3	0,238	0,524	0,654	0,824
4	0,260	0,544	0,664	0,816
5	0,114	0,274	0,361	0,507

Veri setinde yer alan bu beş bireye ait tahmini ölüm risklerinin grafikleri Şekil 4-9’da verilmiştir. Grafikler de x ekseninde zaman, y ekseninde ise tahmin zamanı 30 ile horizon zamanları arasındaki tahmini birimli ölüm riskleri değerleri yer almaktadır. Şekilde gösterilen kırmızı çizgi birinci tahmin yöntemine göre elde edilen riskleri, mavi çizgi ile belirtilen riskler ikinci tahmin yöntemine göre elde edilen riskleri ve yeşil çizgi ile gösterilen ölüm riskleri ise üçüncü tahmin yöntemini göstermektedir.



Şekil 4-9: Tahmin zamanı 30 ay için veri setinde yer alan beş bireye ait tahmini birikimli ölüm riskleri

Yukarıda örnek olarak beş bireye ait riskler ve grafikler verilmiştir. Ayrıca dinamik tahmin yöntemleri ile veri setindeki 465 kişiye ait ölüm riskleri hesaplanmıştır. Ancak 465 kişiye ait ölüm riskleri ve tanımlayıcı parametreleri vermek oldukça yer kapladığı için, tahmin risklerine ait değerler verilmemiştir. Elde edilen tahmin sonuçları Splines-per yöntemine aittir. Splines yöntemine göre elde edilen tahminler benzer olduğundan bu yöntemin sonuçlarına yer verilmemiştir. Dinamik tahmin yöntemleri ile üç farklı tahmin yöntemleri kullanılarak elde edilen ölüm risklerinin doğruluğunu test etmek amacıyla ise Brier skor değerleri hesaplanmıştır. Ayrıca tahmin zamanları arasındaki farklılıkları da değerlendirmek adına farklı tahmin zamanları değerlendirilmiştir. Tablo 4-7, Tablo 4-8 ve Tablo 4-9’da tahmin zamanları sırasıyla $t=30$, $t=40$ ve $t=50$ alındığında dinamik tahmin yöntemleri ile elde edilen ölüm risklerinin doğruluğunun denetlenmesi adına elde edilen Brier skor değerleri yer almaktadır. Brier skor değeri sıfır ile bir arasında değer almaktadır. Skor değeri sıfıra yaklaştıkça, risk tahminlerinin doğruluğu da artmaktadır. Skorun bire yakınlığı ise, tahminlerin tutarlı olmadığını göstermektedir. Bir başka ifade ile ölüm risklerinin tahmin gücü, skor değeri ne kadar sıfıra yakınsa o derecede güçlüdür. Tablo 4-7, Tablo 4-8 ve Tablo 4-9 sonuçları incelendiğinde Brier skor değerlerinin hepsinin 0,30’un altında olduğu görülmektedir.

Özellikle 3. horizon zamanlarında (55 ay, 65 ay, 75 ay) Brier skor değerleri birinci ve ikinci tahmin yönteminde en düşük seviyede olduğu belirlenmiştir. $t=30$ tahmin zamanı için en düşük Brier skor değerinin ikinci tahmin yöntemi ile elde edilen 40 ile 55 ay arasında olduğu gözlenmiştir (Brier skor değeri =0,180). $t=40$ için en düşük Brier skor değeri olarak ise tahmin zamanı 40 ay ile 65 ay arasında , ikinci tahmin yöntemi ile hesaplanan 0,175 değeri olarak belirlenmiştir. Son olarak $t=50$ için ikinci tahmin yönteminde ile elde edilen ve horizon zamanı 65 ay alındığı durumda Brier skoru en düşük seviyededir (Brier skoru=0,176). Üçüncü tahmin yöntemine göre yapılan tahminler için hesaplanan Brier skor değerleri incelendiğinde, birinci ve ikinci tahmin yöntemlerine göre daha yüksek olduğu gözlenmiştir. Dolayısı ile birinci ve ikinci tahmin yöntemlerindeki sonuçların doğruluğu üçüncü tahmin yöntemine göre daha yüksek bulunmuştur. Bu sonuçların nedeni ise, üçüncü tahmin yöntemi ile ölüm riskleri tahmin edilirken geçmişte meydana gelen nüks hikayelerinin dikkate alınmaması, ölüm risklerinin yalnızca parametre tahminlerine dayalı olarak tahmin etmesidir. Cox regresyon analizi sonucunda elde edilen tahminlere göre hesaplanan Brier skoru değerinin ise 0,330 olduğu belirlenmiştir. Dolayısı ile aşağıdaki tablolardan görüldüğü üzere birleşik zayıflık modelleri ile elde edilen tahminlerin hiçbirinde skor değeri 0,30'un üzerine çıkmaz iken, Cox regresyon analizi sonucunda elde edilen tahminlere ait Brier skorunun 0,30'un üzerinde olduğu gözlenmiştir. Bu da birleşik zayıflık modellerindeki tahminlerin doğruluğunu göstermektedir.

Table 4-7: $t=30$ durumunda dinamik tahmin ile elde edilen tahmin değerlerinin Brier skor ile değerlendirilmesi

Tahmin	Zaman			
	35 ay	45 ay	55 ay	70 ay
1	0,284	0,220	0,185	0,219
2	0,278	0,212	0,180	0,218
3	0,296	0,248	0,228	0,272

Table 4-8: $t=40$ durumunda dinamik tahmin ile elde edilen tahmin değerlerinin Brier skor ile değerlendirilmesi

Tahmin	Zaman			
	45 ay	55 ay	65 ay	80 ay
1	0,279	0,197	0,182	0,225
2	0,269	0,185	0,175	0,224
3	0,289	0,230	0,234	0,287

Table 4-9: $t=50$ durumunda dinamik tahmin ile elde edilen tahmin değerlerinin Brier skor ile değerlendirilmesi

Tahmin	Zaman			
	55 ay	65 ay	75 ay	90 ay
1	0,256	0,187	0,184	0,223
2	0,242	0,176	0,180	0,223
3	0,270	0,225	0,239	0,287

5. TARTIŞMA

Meme kanseri nedeniyle meydana gelen ölümlerin birçoğunda ölüm gerçekleşmeden önceki zamanlarda genellikle hastalık tekrar etmektedir. Yapılan çalışmaların bir çoğunda, hastalığın tekrar ettiği (nüks veya metastaz) ve ölüm olaylarının tekrarlardan ayrı olarak değerlendirildiği görülmektedir (Lafourcade ve ark. 2018). Nüks ve ölüm olaylarının ayrı değerlendirilmesi ise yanlı sonuçların ortaya çıkmasına neden olmaktadır. Yanlı sonuçların ortadan kaldırılması ve yapılan tahminlerin doğruyu yansıtması amacıyla son yıllarda tekrarlayan olaylar ve ölüm arasındaki bağımlılıkları dikkate alan ve bireyler arasındaki heterojenlikleri değerlendiren birleşik zayıflık modellerinin kullanımı önerilmektedir (Rondeau ve ark. 2007). Gerçekleştirdiğimiz çalışmada, birleşik zayıflık modellerinin özellikleri açıklanarak, nüks ve ölümlerin değerlendirilmesinde birleşik zayıflık modelleri kullanılmış, bireyler arası heterojenlikler ve olaylar arası bağımlılıklar dikkate alınarak parametre tahminleri elde edilmiştir. Kurulan birleşik zayıflık modeli yardımıyla iki farklı sağkalım sonlanım noktaları için etkili olabilecek açıklayıcı değişkenlerin etkileri aynı anda tahmin edilmiştir. Model tüm hasta bilgilerini kullanarak tek bir seferde parametre tahminlerini vermektedir. Ayrıca modelde nüks ve ölüm arasındaki bağımlılıklar ve bireyler arasındaki heterojenlikler dikkate alındığı için, regresyon katsayılarına ait tahminler de yanlı olmayacaktır. Birleşik zayıflık modelleri ayrıca, iki farklı sonuç değişkeni için farklı açıklayıcı değişken etkilerinin araştırılmasına imkan sağlamaktadır. Araştırılan bu değişkenler zamana bağlı olarak da farklılık gösterebilmekte veya zamandan bağımsız da olabilmektedir.

Birleşik zayıflık modelinin uygulanmasının ardından bir sonraki adım olan bireysel risk tahminlerinin elde edilmesi için ise dinamik tahmin yöntemi kullanılmıştır. Ölümden önce tekrarlayan olay meydana geldiği durumlarda (nüks gibi) hastaların bireysel ölüm risklerinin tahmininde dinamik tahmin yöntemlerinin kullanımı oldukça avantaj sağlamaktadır. Dinamik tahmin yöntemleri ile yapılan tahminlerde tüm hastaların ve tüm parametrelerin bilgileri kullanıldığı için yansız sonuçlar elde edilmektedir. Dolayısı ile çalışmamızda hastaların tüm geçmiş (hikaye) bilgilerini kullanarak tahminler üreten dinamik tahmin yöntemleri kullanılmıştır. Dinamik tahmin yöntemleri ile elde edilen sonuçların kullanılması, hastaya uygun tedavi seçiminde ve kişiselleştirilmiş tıp yöntemlerinin uygulanmasında klinisyenlere çok önemli katkılar sağlamaktadır. Dinamik

tahmin yöntemlerinin kullanılması yaptığımız çalışmanın diğer çalışmalardan avantajlı uygulamaları barındıran bir yaklaşım olduğunu göstermektedir.

Çalışmada kullanılan verilerin analizi sonucunda, nüks ve ölüm arasında pozitif bir ilişkinin olduğu ve ölüm riskinin nüks olayının meydana gelmesinden etkilendiği belirlenmiştir. Elde ettiğimiz sonuç birçok araştırmayı da destekler niteliktedir. O'Shaughnessy metastatik olay ile ölüm arasında önemli bir ilişki olduğunu gözlemlerken, Elkhuisen ve arkadaşları tarafından yapılan çalışmada da benzer şekilde bölgesel nüksler ve ölüm arasında önemli ilişkiler bulunmuştur (Elkhuisen ve ark. 1998; O'Shaughnessy 2005).

Borow ve arkadaşları 2019 yılında yayınladıkları çalışmalarında, halen çalışılmaya devam edilen kalp krizi hastalığı ile ilgili olan DREAM-HF denemesini çalışmaktadırlar. Yaptıkları çalışmada birleşik zayıflık modellerinin öneminden de bahsetmişlerdir. Ölümcül olmayan kalp krizi olaylarının birbirinden bağımsız olmadığını, bireyin daha öncesinde kalp krizi geçmişi olduğunda sonraki zamanlarda kalp krizi riskinin ve kalp krizinden ölme riskinin de arttığını ve bireyler arasındaki kalp krizi risklerinde farklılıklar olduğunu belirtmişlerdir. Daha öncesinde tekrarlayan olaylar için Poisson, negatif binom ve Anderson Gill modellerinin kullanıldığını ancak birleşik zayıflık modellerinin sonuç olay ve tekrarlayan olay arasındaki ilişkileri kullanarak iki farklı model kurması ve bireyler arasındaki farklılıkları dikkate alması nedeniyle avantajlı olduğu ifade edilmiştir. Ayrıca birleşik zayıflık modellerinde tekrarlayan ölümcül olmayan olayların oranlarının zaman içinde sabit olmadığı durumların değerlendirildiği de belirtilmiştir. Orantılı hazard modelindeki parametreler ile birleşik zayıflık modelinde elde edilen hazard oranları karşılaştırılmamıştır. Çünkü yaptıkları çalışmada tekrarlayan olay hazardlarının birey zayıflıklarına bağlı olduğunu, dolayısı ile bireylerin tekrarlayan olay riski gözlenen olay hızına ve birey zayıflıklarına bağlı olarak hesaplandığını, orantılı hazard modelinde ise bireyler arası farklılıklar dikkate alınmadan ortalama bir hesaplama yapıldığı ifade edilmiştir (Borow ve ark. 2019).

Mauguen ve arkadaşları 2013 yılında gerçekleştirdikleri çalışmada, kanser hastalığında ölüm olayından önce bir ya da birden fazla olayın meydana geldiğini ve bu olayların tümünün hastanın hikayesini dikkate alarak ölüm risklerinin tahmininde kullanılmasının klinik kararların doğruluğunu arttırdığını belirtmişlerdir. Dinamik tahmin yöntemi ile tahmin zamanı t ile horizon zamanı arasındaki ölüm olasılıkları tahmin

edilmiştir. Bu amaç ile 1989 ve 1993 yılları arasında meme kanseri operesi olan 1161 hasta verisi kullanılmıştır. Çalışmada faktör olarak yaş, tümör hacmi, HER2, hormon reseptör, nodal tutulum, histopatolojik evre, peritümoral vasküler invazyon alınmıştır. Bireyler arası zayıflık etkisi 1,04 olarak elde edilmiş, bizim çalışmamıza benzer olarak nüks ve ölüm riskleri arasında pozitif ilişki saptanmış ve zayıflık teriminin ölüm riski üzerine etkisinin büyük olduğu görülmüştür. Çalışmalarında aldıkları ve bizim çalışmamızda da yer alan tümör hacminin hem ölüm hem de nüks üzerine etkisi istatistiksel olarak olduğu saptanmıştır. Yaş için ise 55 yaş ve üstü bireylerin ölüm risklerinin, 40-55 yaş arasındaki bireylere göre yüksek olduğu bulunurken, 40 yaş altı bireylere göre ise nüks riski anlamı bulunmuştur. Yaptığımız çalışmada ise yaş ile riskler arasında anlamlı ilişki bulunmamış, hormon reseptör değişkeninin ise nüks üzerine istatistiksel olarak etkisi olduğu gözlenmiştir. Bu farklılıklar modele farklı değişkenlerin alınmasından kaynaklı da olabilmektedir. Dinamik tahmin sonuçları değerlendirildiğinde ise, bizim çalışmamıza benzer olarak birinci ve ikinci tahmin yöntemlerinden elde edilen tahminlerin birbirine yakın olduğu, üçüncü tahmin yönteminde tahmin hatasının büyük olduğu belirlenmiştir. Ayrıca Mauguen ve arkadaşları yaptıkları simülasyon çalışmalarında farklı zayıflık dağılımlardan elde ettikleri tahmin hatalarının benzer olduğunu ve birleşik zayıflık modellerinde kullanılabilecek zayıflık dağılımlarının sınırlı olduğu vurgulamışlardır. Bu sonuçtan hareketle yaptığımız çalışmada, farklı zayıflık dağılımlarının etkileri araştırılmamıştır (Mauguen ve ark. 2013).

Mauguen ve arkadaşları 2015 yılında yaptıkları çalışmada üç farklı ülkenin (Fransa, İngiltere ve Hollanda) verisi üzerine birleşik zayıflık modellerini uygulamışlardır. Çalışmada ölüm risklerini tahmin ederken, ölümden önce meydana gelen nüks olaylarının da dikkate alınması gerektiğini vurgulamışlardır. Bu amaçla, çalışmalarında Landmark yöntemi ile birleşik zayıflık modellerini kullanmışlardır. Ölüm risklerini tahmin ederken nüks bilgilerinin de dikkate alınmasının tahmin hatasını azalttığını belirtmişlerdir. Landmark yöntemi sadece ölüme dayalı oluşturulması sebebiyle elde edilen tahminlerin birleşik zayıflık modellerinden düşük olduğu gözlenmiştir. Birleşik zayıflık modelleri hem nüks hem de ölüm için iki ayrı model tahminleri oluşturmaktadır. Biz de benzer olarak gerçekleştirdiğimiz çalışmamızda hem nüks hem de ölüm olayları ile ilgilendiğimiz için çalışmada birleşik zayıflık modellerini kullandık. Mauguen ve arkadaşları bizim çalışmamızdaki sonuçları destekleyecek şekilde, rasgele etkinin sıfırdan anlamlı derecede farklı olduğunu ve ölüm ile nüks

arasında pozitif yönde anlamlı bir ilişki olduğu sonucuna varmışlardır. Bireysel ölüm risklerinin dinamik tahmin yöntemleri ile elde etmişler ve birleşik zayıflık modelinin tüm bireyleri ve parametreleri tek bir seferde alarak modeli oluşturduğunu ve tahminlerin dinamik tahmin yöntemleri ile yapmanın oldukça avantajlı olduğunu belirtmişlerdir. Çalışmalarında dinamik tahminler yapılırken tahmin zamanının iki yıldan az alınmasının güvenilir sonuçlar vermediğini belirtmişlerdir. Bizim gerçekleştirdiğimiz çalışmada ise dinamik tahmin yöntemleri ile tahmin yapılırken tahmin zamanı iki yıldan fazla alınmıştır ve tahmin zamanı arttıkça Brier skorlarının da azaldığı gözlenmiştir (Mauguen ve ark. 2015).

Osmani ve arkadaşları 2018 yılında yayınladıkları çalışmada meme kanserli 342 hastanın verilerini geriye dönük tarama yaparak elde etmiş ve nüks ile ölüm arasındaki ilişkileri birleşik zayıflık modelleri ile incelemişlerdir. Araştırmalarında, hem uzak organ hem lokal nüks olaylarını değerlendirmişlerdir. Çalışmalarında lenf nodu gözlenme durumu, yaş, hormon reseptör durumu, tümör genişliği, HER2+ ve histopatolojik evre faktörleri dahil edilmiştir. Modelde ölüm ve nüks riskleri üzerine yaşın önemli ölçüde farklılık yaratmadığını belirlemişlerdir (40 yaş ve altı ile diğer yaş grupları karşılaştırıldığında). Çalışmada lokal nüks ile ölüm arasında anlamlı ilişki bulunmamış iken, Wapnir ve arkadaşlarının yaptıkları çalışmada lokal nüks ile ölüm arasında anlamlı ilişkilerin gözlemlendiğini vurgulamışlardır ki bu sonuç bizim çalışmamızda elde edilen sonuçları desteklemektedir (Osmani ve ark. 2018a).

Osmani ve arkadaşları 2018 yılında yaptıkları çalışmada, 342 meme kanserli hastanın retrospektif olarak verilerini kullanarak lokal nüks ile metastaz arasındaki ilişkileri araştırmışlardır. Gerçekleştirdikleri çalışmada ölüm ile nüks arasındaki ilişkiler kullanılmamış yalnızca iki tekrarlayan olay arasındaki ilişkiler birleşik zayıflık modelleri ile incelenmiştir. Çalışma sonucunda aynı değişken özelliklerine sahip bireylerin dahi risklerinin farklı olduğu gözlenmiştir (Osmani ve ark. 2018b).

Lafourcade ve arkadaşları 2018 yılında yaptıkları çalışmada, meme kanserinde tekrarlayan olay ve ölüm için prognostik faktörleri araştırmışlardır ve hastaların geçmiş bilgilerini kullanarak bireysel ölüm risk tahminlerini elde etmişlerdir. Çalışmalarında E3N((Etude Epidémiologique auprès de Femmes de la Mutuelle Générale de l'Education Nationale) adında bir kohort çalışmanın verilerini kullanmışlar ve toplamda 4926 meme kanserli hasta üzerinde çalışmışlardır. Çalışmalarında bizim çalışmamızda aldığımız

prognostik faktörlerden farklı faktör etkileri araştırılmıştır. Örneğin tanıdan sonra sigara içme durumu gibi. Dolayısıyla elde ettikleri faktör etkileri ile bizim çalışmamızdaki faktör etkileri karşılaştırılamamaktadır. Ancak Lafourcade ve arkadaşları da çalışmalarında tekrarlayan olayların etkilerinin ölüm üzerine anlamlı olduğunu bulmuşlar ve sağkalım fonksiyonunun sonuç değişkeninin (ölümün) tekrar eden olayın meydana gelmesinden etkilendiği ve dinamik tahmin yöntemlerinin kullanılmasının pratikte oldukça kullanışlı ve değerli olabileceğini vurgulamışlardır (Lafourcade ve ark. 2018).

Çalışmada yaptığımız tahminlerde, zayıflık teriminin zaman boyunca değişmediği varsayılmıştır. Ancak gözlenemeyen faktörlerin etkileri zaman boyunca değişiklik gösterebilir ki bu durumda zayıflık terimi zamana bağlı değişiklik gösterir. Yapılan bu çalışmada ise, zayıflık teriminin izlem süresince sabit kaldığı varsayılarak tahminler gerçekleştirilmiştir.

Bu çalışmanın kısıtlılığı, birleşik zayıflık modellerinin avantajlarını göstermek adına veya başka modeller ile karşılaştırmak adına bir simülasyon çalışmasının yapılmamasıdır. Tez çalışması tamamlandıktan sonra her bir modelin uygulanabileceği veri setlerini ortaya koyabilecek bir simülasyon çalışması yürütülmesi düşünülmektedir. Ancak literatürde bu modeller üzerine yapılan bazı simülasyon çalışmaları mevcut olup elde edilen sonuçlar şöyle özetlenebilmektedir. Rondeau ve arkadaşları 2007 yılında gerçekleştirdikleri çalışmada tahmin edicilerin performansını değerlendirmek ve birleşik zayıflık modellerini basit bir zayıflık modeli ile karşılaştırmak amacıyla bir simülasyon çalışması yapmışlardır. Çalışmalarında örneklem genişliklerinin de etkisinin görülmesi amacıyla iki farklı örneklem genişliği ve farklı sayıda nüks geçmişi olan örnekler kullanılmışlardır. Ayrıca çalışmalarında nüks ve ölüm arasındaki ilişkiler de farklı şekillerde incelenmiştir. Bağımlılıkların hem negatif olduğu hem de pozitif olduğu durumlar basit zayıflık modeli ile birleşik zayıflık modelleri arasında sonuçlar bakımından oldukça farklılık olduğu gözlenirken, bağımlılıkları dikkate almayan basit zayıflık modelinin oldukça yanlış tahminler ürettiği gözlemlenmiştir (Rondeau ve ark. 2007). Huang ve Liu 2007 yılında yaptıkları çalışmada sonlanma olayı ve tekrarlayan olay arasındaki bağımlılıklar dikkate alınmadığında yanlış sonuçların ortaya çıktığını yaptıkları simülasyon çalışmaları ile göstermişlerdir. Ayrıca zayıflık dağılımlarının farklı seçilmesinin sonuçlar üzerine etkisinin olmadığını ve zayıflık dağılımının hatalı seçilmesi durumunda ise modelin dayanıklı olduğunu belirtmişlerdir ki bu sonucu Pickles ve

Crochly ve O'Quigley ve Store daha önceki yıllarda göstermiştir (Huang ve Liu 2007). Mazroui ve arkadaşları benzer şekilde 2012 yılında bir simülasyon çalışmasında zayıflık dağılımının yanlış seçilmesi durumuna karşı birleşik zayıflık modellerinin dayanıklı olduğunu bulmuşlardır (Mazroui ve ark. 2012).

Sonuç olarak, onkolojik araştırmalarda ölümden önce meydana gelen metastaz ve nüks gibi olaylar genellikle ölüm ile ilişkilidir ve bu olaylar bakımından bireyler arasında heterojenlik söz konusudur. Ölüm veya nüks gelişmesi üzerine etkili faktörlerin incelenmesi amacıyla kurulacak modelin, bu varyasyonları dikkate alması gerekmektedir. Aksi takdirde yapılacak risk tahminleri yanıltıcı olacaktır. Literatür incelendiğinde sağkalım verilerinin genellikle Cox regresyon modeli ile değerlendirildiği gözlenmektedir. Ancak bu model yukarıda tanımlanan varyasyon kaynaklarını dikkate almaz. Gerçekleştirdiğimiz çalışma ile birleşik zayıflık modelleri kullanılarak bu varyasyonlar ve bağımlılıklar dikkate alınmış, ayrıca hastaların tüm hikayelerini kullanan dinamik tahmin yöntemleri ile ölüm riskleri yansız bir şekilde tahmin edilmiştir.

KAYNAKLAR

- Aalen, O., Valberg, M., Grotmol, T., Tretli, S. (2015). Understanding Variation In Disease Risk: The Elusive Concept Of Frailty. *International Journal of Epidemiology*, 44, 1408-1421.
- Abdulkarimova, U. (2013). *Frailty Models For Modelling Heterogeneity* (Yüksek Lisans Tezi, McMaster University, Hamilton)
- Akbari, M. E., Akbari, A., Nafissi, N., Shormeij, Z., Sayad, S., Rasaf, M. R., Shojaee, L. (2016). Prognostic Factors of Recurrence (Early And Late) and Death in Breast Cancer Patients In Iranian Women. *Iranian Journal of Cancer Prevention*, 9, e5747.
- Aygün Çevik, B., Pirinçci, E. (2017). Beslenme ve Kanser. *Fırat Tıp Dergisi*, 22, 1-7.
- Balcı, H.D. (2017). *Kadınlarda Meme Kanseri Risk Değerlendirme Araştırması*. (Uzmanlık Tezi, Sağlık Bilimleri Üniversitesi Şişli Etfal Sağlık Uygulama ve Araştırma Merkezi, İstanbul).
- Berkson, J., Gage, R. (1952). Survival Curve for Cancer Patients Following Treatment. *Journal of the American Statistical Association*, 47,501–515.
- Boag, J. (1949). Maximum Likelihood Estimates of The Proportion of Patients Cured by Cancer Therapy. *Journal of the Royal Statistical Society Series B (Methodological)*, 11,15–53.
- Borow, K. M., Yaroshinsky, A., Greenberg, B., Perin, E. C. (2019). Phase 3 DREAM-HF Trial of Mesenchymal Precursor Cells in Chronic Heart Failure: A Review of Biological Plausibility and Implementation of Flexible Clinical Trial Design. *Circulation Research*, 125, 265-281.
- Bray, F., Ferlayi J., Soerjomataram, I., Siegel, R., Torre, L., Jemal, A (2018). Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal of Clinicians*, 68, 394-424.
- Cai, C., Zou, Y., Peng, Y., Zhang J. (2012). Smcure: An R-package for Estimating Semiparametric Mixture Cure Models. *Computer Methods and Programs in Biomedicine*, 108,1255-1260.

- Collett, D. (2003). *Modelling Survival Data in Medical Research*. Boca Raton: Chapman & Hall/CRC.
- Commenges, D., Jacqmin-Gadda, H. (2015). *Dynamical Biostatistical Models*. Boca Raton: Chapman & Hall/CRC.
- Dişçi, R. (2015). *Temel ve Klinik Biyoistatistik*. İstanbul: İstanbul Tıp Kitabevi.
- Elkhuizen, P. H., van de Vijver, M. J., Hermans, J. O., Zonderland, H. M., van de Velde, C. J., Leer, J. W. H. (1998). Local Recurrence After Breast-Conserving Therapy for Invasive Breast Cancer: High Incidence in Young Patients and Association with Poor Survival. *International Journal of Radiation Oncology, Biology, Physics*, 40, 859-867.
- Fitrianto, A. (2013). Several Types of Residuals in Cox Regression Model: An Empirical Study. *International Journal of Mathematical Analysis*, 7, 2645-2654.
- Gürtekin, B. (2011). *Propensity Skor Yönteminin Meme Kanserinde Yaş ile Dokudaki ER, PR, CEB-B2 Değişiminin Analizinde Kullanılması* (Doktora tezi, İstanbul Üniversitesi, İstanbul).
- Hammer, C., Fanning, A., Crowe, J. (2008). Overview of Breast Cancer Staging and Surgical Treatment Options. *Cleveland Clinic Journal of Medicine*, 75, 10-16.
- Hanagal, D.D. (2011). *Modelling Survival Data Using Frailty Models*. Boca Raton: Chapman & Hall/CRC.
- Haydaroğlu, A. (2007). Meme Kanserinde İzlem. *Meme Sağlığı Dergisi*, 3, 95-99.
- Huang, X., Liu, L. (2007). A Joint Frailty Model for Survival and Gap Times Between Recurrent Events. *Biometrics*, 63, 389-397.
- İnceoğlu, F. (2013). *Sağkalım Analiz Yöntemleri ve Karaciğer Nakli Verileri ile Bir Uygulama* (Yüksek Lisans Tezi, İnönü Üniversitesi, Malatya).
- Kara, P., Tutkun, NA. (2017). Cox Karma İyileşme Modeli ve Glioma Veri Kümesi Üzerine Bir Uygulama. *Türkiye Klinikleri Journal of Biostatistics*, 9, 241-256.
- Kaya, M.E. (2017). *Sağkalım Analizi ve Zootekni Alanında Kullanımı* (Yüksek Lisans Tezi, Namık Kemal Üniversitesi, Tekirdağ)
- Kimmel, S.E. (2018). *Dynamic Prediction Modelling to Improve Clinical Predictions*. Erişim 01.12.2018: <http://grantome.com/grant/NIH/R01-HL141294-01>.

- Kleinbaum, D.G., Klein, M. (2010). *Survival Analysis*. New York: Springer.
- Lafourcade, A., His, M., Baglietto, L., Boutron-Ruault, MC., Dossus, L., Rondeau, V. (2018). Factors Associated with Breast Cancer Recurrences of Mortality and Dynamic Prediction of Death Using History of Cancer Recurrence:The French E3N Cohort. *BMC Cancer*, 18,1-9.
- Lee, E.T., Wang, J.W. (2003). *Statistical Methods for Survival Data Analysis*. New Jersey: John Wiley & Sons.
- Mauguen, A., Rachet, B., Mathoulin-Pélissier, S., MacGrogan, G., Laurent, A., Rondeau, V. (2013). Dynamic Prediction of Risk of Death Using History of Cancer Recurrences in Joint Frailty Models. *Statistics in Medicine*, 32, 5366-5380.
- Mauguen, A. (2015). *Prognosis of Cancer Patients: Input of Standard and Joint Frailty Models* (Doktora Tez, Bordeaux Üniversitesi, Fransa).
- Mauguen, A., Rachet, B., Mathoulin-Pélissier, S., Lawrence, G. M., Siesling, S., MacGrogan, G., Laurent, A., Rondeau, V. (2015). Validation of Death Prediction After Breast Cancer Relapses Using Joint Models. *BMC Medical Research Methodology*, 15, 27-36.
- Mazroui, Y., Mathoulin-Pelissier, S., Soubeyran, P., Rondeau, V. (2012). General Joint Frailty Model for Recurrent Event Data with A Dependent Terminal Event: Application to Follicular Lymphoma Data. *Statistics in Medicine*, 31, 1162-1176.
- Moore, D.F. (2016). *Applied Survival Analysis Using R*. New York: Springer.
- O'Shaughnessy, J.(2005). Extending Survival with Chemotherapy in Metastatic Breast Cancer. *Oncologist*. 2005,10, 20–29.
- Osmani, F., Hajizadeh, E., Rasekhi, A., Akbari, M.E. (2018a). Analyzing Relationship Between Local and Metastasis Relapses with Survival of Patients with Breast Cancer: A Study Using Joint Frailty Model. *International Journal of Cancer Management*, 11, e81783.
- Osmani, F., Hajizadeh, E., Rasekhi, A.A. (2018b). Joint Frailty Modeling For Multiple Recurrent Events and Its Application in Patients with Breast Cancer. *Journal of Obstetrics, Gynecology and Cancer Research*, 3, 53-58.

- Othus, M., Barlogie, B., Leblanc, ML., Crowley, JJ.(2012). Cure Models As A Useful Statistical Tool For Analyzing Survival. *Clinical Cancer Research*, 18,3731-3736.
- Öksüz, D.Ç. (2015). Meme Kanserinde Prognostik ve Prediktif Faktörler. *Türkiye Klinikleri Radyasyon Onkolojisi-Özel Konular*, 1, 7-12.
- Qi, J. (2009). *Comparison of Proportional Hazards and Accelerated Failure Time Models* (Doktora tezi).
- Rondeau, V., Mathoulin-Pelissier, S., Jacqmin-Gadda, H., Brouste, V., Soubeyran, P. (2007). Joint Frailty Models For Recurring Events and Death Using Maximum Penalized Likelihood Estimation: Application on Cancer Events. *Biostatistics*, 8, 708-721.
- Rondeau, V. (2010). Statistical Models For Reccurent Events and Death: Application to Cancer Events. *Mathematical and Computer Modelling*, 52, 949-955.
- Rondeau, V., Yassin M., Juan R. G. (2012). frailtypack: An R package for The Analysis of Correlated Survival Data with Frailty Models Using Penalized Likelihood Estimation or Parametrical Estimation. *Journal of Statistical Software*, 47, 1-28.
- Saphner, T., Tormey, D.C., Gray, R. (1996). Annual Hazard Rates of Recurrence for Breast Cancer After Primary Therapy. *Journal of Clinical Oncology*, 14, 2738-2746.
- Tabakan, G. (2009). *Yarı Parametrik Regresyonda Tahmin Metodları* (Doktora tezi, Çukurova Üniversitesi, Adana).
- Tanas, S. (2013). *Meme Kanseri Olgularının Demografik Özellikleri* (Uzmanlık Tezi, Atatürk Üniversitesi, Erzurum).
- Therneau, T.M., Grambsch, P.M. (2000). *Modelling Survival Data Extending The Cox Model*. New York: Springer.
- Yay, M., Çoker, E., Uysal, Ö. (2007). Yaşam Analizinde Cox Regresyon Modeli Ve Artıkların İncelenmesi. *Cerrahpaşa Tıp Dergisi*, 38, 139-145.
- WHO (2018, Eylül). Latest Global Cancer Data: Cancer Burden Rises to 18.1 Million New Cases and 9.6 Million Cancer Deaths in 2018. WHO Reports. Erişim 14.02.2015, <https://www.who.int/cancer/PRGlobocanFinal.pdf>.

Wienke, A. (2011). *Frailty Models in Survival Analysis*. Boca Raton: Chapman & Hall/CRC.



İNTİHAL ORANI

MEME KANSERLİ HASTALARDA ÖLÜM RİSKİNİN BİRLEŞİK ZAYIFLIK MODELLERİ İLE İNCELENMESİ

ORJİNALLİK RAPORU

%3	%2	%1	%2
BENZERLİK ENDEKSİ	İNTERNET KAYNAKLARI	YAYINLAR	ÖĞRENCİ ÖDEVLERİ

BİRİNCİL KAYNAKLAR

1	www.biyoistatistikkongresi.org İnternet Kaynağı	%1
2	polen.itu.edu.tr İnternet Kaynağı	<%1
3	www.cinematography.net İnternet Kaynağı	<%1
4	Submitted to Uludag University Öğrenci Ödevi	<%1
5	acikerisim.istanbul.edu.tr İnternet Kaynağı	<%1
6	halksagligiokulu.org İnternet Kaynağı	<%1
7	www.f2bclan.com İnternet Kaynağı	<%1
8	Submitted to Konya Necmettin Erbakan University Öğrenci Ödevi	<%1

ÖZGEÇMİŞ

Kişisel Bilgiler

Adı	Özge	Soyadı	Pasin
Doğ.Yeri	Kadıköy	Doğ.Tar.	09.02.1990
Uyruğu	TC	TC Kim No	15206077862
Email	ozgepasin90@yahoo.com.tr	Tel	05374020606

Eğitim Düzeyi

	Mezun Olduğu Kurumun Adı	Mez. Yılı
Doktora	İstanbul Üniversitesi(Biyoistatistik)	2019
Yük.Lis.	Düzce Üniversitesi (Biyoistatistik ve Tıp Bilişimi)	2015
Lisans	Gazi Üniversitesi (İstatistik Bölümü)	2011
Lise	Üsküdar Fen Lisesi	2007

İş Deneyimi (Sondan geçmişe doğru sıralayın)

	Görevi	Kurum	Süre (Yıl - Yıl)
1.	Araştırma Görevlisi	İstanbul Üniversitesi	2016-
2.	Araştırma Görevlisi	Düzce Üniversitesi	2013-2016

Yabancı Dilleri	Okuduğunu Anlama*	Konuşma*	Yazma*	KPDS/ÜDS/ Yökdil Puanı	(Diğer) Puanı
İngilizce	Çok İyi	Orta	İyi	83,75	

*Çok iyi, iyi, orta, zayıf olarak değerlendirin

	Sayısal
ALES Puanı	87,21

Bilgisayar Bilgisi

Program	Kullanma becerisi
SPSS. Medcalc. NCSS.	Çok İyi
SAS. Stata. R.	İyi
Excel. Word.	Çok İyi

Yayınları/Tebliğleri Sertifikaları/Ödülleri

ULUSLAR ARASI YAYINLAR:

1. Pasin Özge, Ankaralı Handan, Cangür Şengül (2014). Ankilozan Spondilitli Hastaların Yersel Kümelenmelerinin İncelenmesi. Türkiye Klinikleri J Biostat, 6(2): 74-82.
2. Cangür Şengül, Pasin Özge, Ankaralı Handan (2015). Comparison of Sampling Distributions and Performances of Minitab and Freud Perles Quartiles. Pakistan Journal of Statistics 31(1): 1-20.
3. Pasin Ozge, Ankaralı Handan (2015). Usage of Kernel K-means and DBSCAN cluster Algorithms in Health Studies: An Application. Clinical Research and Trials, 1(3):63-69.
4. Cangür Şengül, Ankaralı Handan, Pasin Özge (2016). Comparing Performances of Multiple Comparison Methods in Commonly Used 2*C Contingency Tables. Interdisciplinary Science: Computational Life Sciences, 8(4):337-345.
5. Pasin Ozge, Ankaralı Handan, Cangür Şengül, Sungur Mehmet Ali (2016). Parametrik Olmayan Çok Değişkenli Varyans Analizi ve Sağlık Alanında Bir Uygulaması. International Journal of Informatics Technologies, 9(1): 13-20.
6. Pasin Ozge, Ankaralı Handan (2016). Usage of Cluster Algorithms in Health Studies: An Application. Türkiye Klinikleri Journal of Medical Sciences, 36(1): 40-52.
7. Ankaralı Handan, Pasin Ozge, Seyit Ankaralı (2017). A New Index for Evaluating Academic Performance: HOS-Index. International Journal of Computer Science Issues, 14(3): 78-86.
8. Pasin Ozge, Ankaralı Handan, Yeşilyurt Levent (2017). Usage of Heckman Sample Selection Model in Health Studies: An Application of Prostate Cancer Patients. Turkish Journal of Medical Science, 37(4):147-154.
9. Alpay Merve, Kısmalı Gökem, Ögünç Meral, Sel Tevhide, Özmerdivenli Recep, Pasin Özge (2017). Antioxidant Therapy Impresses in Oxidative Stress-induced kidney cells. Bratislava Medical Journal, 118(2):89-94.
10. Alpay Merve, Kısmalı Gökem, Sel Tevhide, Pasin Özge, Karagül Hilal (2017). The Investigation Antiproliferative and Apoptotic Effects of Capsaicin and Alpha Lipoic Acid on Colon Adenocarcinoma Cells. International Journal of Current Medical and Pharmaceutical Research, 3(11): 2612-2620.

11. Pasin Ozge, Ankaralı Handan (2017). Comparison of EM and Two-step Cluster Method for Mixed Data: An Application. *International Journal of Medical Science and Clinical Inventions*, 4(3):2768-2773.
12. Bahar Ölmez Safiye, Ataoğlu Büşra Bahar, Başar Kocagöz Zehra, Pasin Özge (2018). An Investigation of Childhood Trauma in Patients with Panic Disorder. *Nobel Medicus*, 14(1): 39-48.
13. Özkuvancı Ünsal, Dönmez Muhammet İrfan, Özgör Faruk, Erbin Akif, Pasin Özge, Yase Müslümanoğlu Ahmet (2018). Durasphere® EXP: A Non-biodegradable Agent for Treatment of Primary Vesico-Ureteral Reflux in Children. *Int Braz J Urol*, 44(3): 585-590.
14. Ataoğlu Safinaz, Ankaralı Handan, Ankaralı Seyit, Pasin Özge (2018). Determination of The Appropriate Quality of Life Scale for Patients with Rheumatoid Arthritis and Osteoarthritis. *Ankara Medical Journal*, 18(4): 612-27.
15. Pasin Tuğçe, Ataoğlu Safinaz, Pasin Özge, Ankaralı Handan (2019). Comparison of the Effectiveness of Platelet-Rich Plasma, Corticosteroid, and Physical Therapy in Subacromial Impingement Syndrome. *Arch Rheumatol*, 34(3):308-316.
16. Ankaralı Handan, Ataoğlu Safinaz, Ankaralı Seyit, Pasin Özge (2019). Severity of Pain, and Sensory Changes in Patients with Rheumatoid Arthritis: A New SH-Pain Scale. *EPMR*, 2(3): 1-7.
17. Yıldırım İlknur, Çelik İ Ayça, Bay B Sema, Pasin Özge, Tütüncü Ç Ayşe (2019). Propofol-based Balanced Anesthesia is Safer in Pediatric Radiotherapy. *J Oncol Pharm Practise*, 0(0): 1-6.

ULUSAL YAYINLAR:

1. Ankaralı Handan, Pasin Özge, Karacan Büşra, Tokar Mehdi, Künüroğlu Merve, Çaça Muhammed, Özislam Muhammed Enes, Şahingöz Nur Banu (2017). Üniversite Öğrencilerinin Türkiye'deki Suriyeli Sığınmacılara Bakış Açısı. *Journal of Duzce University Health Sciences Institute*, 7(3): 122-132.
2. Çelebi Esra, Ataoğlu Safinaz, Ankaralı Handan, Özşahin Mustafa, Pasin Özge (2016). Fibromiyaljili Hastalarda Kişilik Özelliklerine Göre Pregabalin ve Duloksetin Tedavilerinin Etkinliğinin Karşılaştırılması 18(2):47-53.

BİLDİRİLER:

1. Pasin Özge, Ankaralı Handan, Cangür Şengül (2013). Düzce Üniversitesi Araştırma Hastanesinde Tedavi Gören Ankilozan Spondilitli Hastaların Yersel Kümelenmelerinin İncelenmesi. 8. Uluslararası İstatistik Kongresi.
2. Cangür Şengül, Ankaralı Handan, Pasin Özge (2013). R*C Boyutlu Çapraz Tablolarda Çoklu Karşılaştırma Yöntemlerinin İncelenmesi Bir Uygulama. 8. Uluslararası İstatistik Kongresi.
3. Cangür Şengül, Pasin Özge, Ankaralı Handan (2013). Minitab ve Freud Perles Kartil Tahmin Yöntemlerinin Örneklem Dağılımları ve Performanslarının Karşılaştırılması: Bir Simülasyon Çalışması. 15. Ulusal Biyoistatistik Kongresi.
4. Erdem Havva, Başar Feyza, Kadioğlu Nilüfer, Belada Abdullah, Pasin Özge, Bahadır Anzel, İnce Nevin, Gündoğdu Betül, Şener Ebru (2014). CerB2 Ekspresyon of Gland Cell and Surface Cell at the Nasal Polyps. 17. International of Congress on Oral Patology and Medicine.
5. Pasin Özge, Ankaralı Handan, Cangür Şengül (2014). Tobit Regresyon Analizi ve Sağlık Alanında Bir Uygulama. XVI. Ulusal Biyoistatistik Kongresi.
6. Pasin Özge, Ankaralı Handan, Cangür Şengül, Sungur Mehmet Ali (2014). An Application of Non-parametric Factorial Manova in Health Research. International Society for Clinical Biostatistics Congree.
7. Cangür Şengül, Ankaralı Handan, Pasin Özge (2015). Comparing Performances of Multiple Comprasion Methods in Commonly Used 2*C Contingency Tables. IX. International Confrence on Multiple Comparion Procedures.
8. Pasin Özge, Ankaralı Handan, Cangür Şengül, Sungur Mehmet Ali (2015). Sağlık Alanında Yapılan Araştırmalarda Kümeleme Algoritmalarının Kullanımı: Bir Uygulama. XVII. Ulusal Biyoistatistik Kongresi.
9. Gümüş Zahide, İskender Abdülkadir, Özlü Onur, Şeker İlknur Suidiye, Demiraran Yavuz, Özkan Aybars, İlhan Ethem, Ünlü İlhan, Konuk İsmail Doğan, Pasin Özge, Türkoğlu Kübra (2015). Ketamin Anestezisi Altında Sünnet Operasyonu Geçirecek Çocuklarda Orta Kulak Basıncı Değişikliklerinin Değerlendirilmesi. Türk Anesteziyoloji ve Reanimasyon Derneği 49. Ulusal Kongresi.

10. Ankaralı Handan, Pasin Özge (2016). The Usage of Analytic Hierarchy Process in Health Studies. International Society for Clinical Biostatistics Congree.
11. Ankaralı Handan, Pasin Özge (2016). Comparison of EM and Two-Step Cluster Method for Mixed Data An Application. International Society for Clinical Biostatistics Congree.
12. Ankaralı Handan, Ankaralı Seyit, Cangür Şengül, Pasin Özge (2016). Az Denekle Yürütülen Çalışmalar için Özel Deney Tasarımları Klinik Öncesi Çalışmalar. 18 Ulusal Biyoistatistik Kongresi ve 1. Uluslararası Biyoistatistik Kongresi.
13. Alpay Merve, Li Chun, Özmerdivenli Recep, Ataoğlu Safinaz, Pasin Özge (2017). Moleküler Hastalıkların Teşhis ve Tedavisi için Kök Hücre ve Nanopartiküllerin İşbirliği. Uluslararası Katılımlı Türk Romatoloji Kongresi.
14. Çelebi Esra, Ataoğlu Safinaz, Ataoğlu Büşra Bahar, Ankaralı Handan, Pasin Özge, Ölmez Safiye Bahar (2017). The Investigation of Effects of Pregabalin and Duloxetine Treatment According to Personality Characteristics with Fibromyalgia Patients. 9. International Congress on Psychopharmacology, 5. International Symposium on Child and Adolescent Psychopharmacology.
15. Pasin Özge, Ankaralı Handan, Ataoğlu Safinaz (2017). Three Different Approaches to Diagnose of Rheumatic Diseases by Using Quality of Life Score. 9. EMR-IBS and Italian Region Conference.
16. Ankaralı Handan, Pasin Özge, Ankaralı Seyit (2017). A New HOS Index for Academic Performance. 9. EMR-IBS and Italian Region Conference.
17. Pasin Özge, Diricen Ahmet, Gürtekin Başak, Özel Yıldız Sevdâ, Dişçi Rian (2017). Evaluation of Different Statistical Methods Used to Test for Agreement Between the Results of Continuous Measurements. 9. EMR-IBS and Italian Region Conference.
18. Ataoğlu Safinaz, Ankaralı Handan, Pasin Özge (2017). Reliability and Validity of Short Forms of Some Scales for Used in Quality of Life in Rheumatic Disease. 9. EMR-IBS and Italian Region Conference.
19. Sever Ülgen, Uysal Mehmet Atilla, Pasin Özge, Working Group Yedikule Smoking Cessation Polyclinic, Pehlivan Sacide (2017). The Role of DAT, 5-HTT, MAOA and Nicotine Dependence in Turkish Population. VI. International Congress of Molecular Medicine.

20. Ankaralı Handan, Pasin Özge, Ankaralı Seyit (2018). Sampling Distributions of Four Test Statistics for Posttest-PreTest Ratio: A Simulation Study. 4. International Researchers, Statisticians and Young Statisticians Congress.
21. Ankaralı Handan, Pasin Özge, Ankaralı Seyit (2018). Comparison of Four Different Test Statistics for Percent Changes: A Simulation Study. 4. International Researchers, Statisticians and Young Statisticians Congress.
22. Pasin Özge, Ankaralı Handan (2018). Meme Kanserli Hastalar Üzerine Orantısız Tehlikeler Karma İyileşme Modelinin Uygulanması. 20. Ulusal ve 3. Uluslararası Biyoistatistik Kongresi.
23. Pasin Özge, Gürtekin Başak, Gülen Onur Elif (2018). Heckman Örneklem Seçim Modelinin Sigara Bağımlılığı Üzerine Uygulanması. 20. Ulusal ve 3. Uluslararası Biyoistatistik Kongresi.
24. Çetin Şirin, Demir Osman, Dede İsa, Pasin Özge, Kuyucu Emre (2019). Pankreas Kanseri Verilerinin Paylaşılmış Zayıflık Modeli ile Analizi. 21. Ulusal ve 4. Uluslararası Biyoistatistik Kongresi.