

**BİLGİ İÇEREN DURDURMA VARLIĞINDA YİNELEMELİ
OLAY SÜRECİ**

**RECURRENT EVENT PROCESS WITH INFORMATIVE
CENSORING**

HANDE ÜNLÜ

Doç. Dr. Ayten YİĞİTER

Tez Danışmanı

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin

İstatistik Anabilim Dalı İçin Öngördüğü

DOKTORA TEZİ

olarak hazırlanmıştır.

2013

**BİLGİ İÇEREN DURDURMA VARLIĞINDA YİNELEMELİ
OLAY SÜRECİ**

**RECURRENT EVENT PROCESS WITH INFORMATIVE
CENSORING**

HANDE ÜNLÜ

Doç. Dr. Ayten YİĞİTER

Tez Danışmanı

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin

İstatistik Anabilim Dalı İçin Öngördüğü

DOKTORA TEZİ

olarak hazırlanmıştır.

2013

Biricik Anneme ve Sevgili Eşime...

ETİK

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

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

beyan ederim.

09.09.2013

Hande ÜNLÜ

ÖZET

BİLGİ İÇEREN DURDURMA VARLIĞINDA YİNELEMELİ OLAY SÜRECİ

HANDE ÜNLÜ

Doktora, İstatistik Bölümü

Tez Danışmanı: Doç. Dr. AYTEN YİĞİTER

Eylül 2013, 101 sayfa

Bu tez çalışmasının amacı, bilgi içeren durdurma varlığında yinelemeli olay verisini modellemek ve parametre tahminlerini elde etmektir.

Bilgi içeren durdurma, yinelemeli olay verisinin modellenmesinde dikkat edilmesi gereken önemli bir konudur. Bilgi içeren durdurma, ölüm gibi bir başarısızlık olayı ya da bireylerin çalışmadan çekilmesi sebepleriyle çalışmanın sona ermesi sonucu meydana gelir. Bu nedenle yinelemeli olay zamanları ile bilgi içeren durdurma zamanı ilişkilidir. Yinelemeli olay verisinde bu ilişkinin göz ardı edilmesi model parametrelerinin yanlış tahmin edilmesine yol açmaktadır.

Tez çalışmasında, bilgi içeren durdurmanın ölüm olayından kaynaklandığı durum incelenmiş ve ölüm olayı varlığında yinelemeli olay süreci homojen Poisson süreci kullanılarak modellenmiştir. Yinelemeli olay sürecine ilişkin yoğunluk fonksiyonları için iki farklı model önerilmiş ve yineleme zamanları ile ölüm zamanı arasındaki ilişki yapısı, paylaşılmış zayıflık modelleri kullanılarak oluşturulmuştur. Parametrelerin en çok olabirlik tahminleri EM ve Metropolis-Hastings algoritmaları kullanılarak elde edilmiştir. Önerilen modeller, Byar (1976)'ın yapmış olduğu çalışmada yer alan mesane kanseri verilerine uygulanmıştır. Önerilen modellerin geçerliliğinin gösterilmesi için benzetim çalışması yapılmıştır.

Anahtar Kelimeler: Yinelemeli olay verisi, bilgi içeren durdurma, sayma süreci, homojen Poisson süreci, yoğunluk fonksiyonu, EM algoritması, Metropolis-Hastings algoritması.

ABSTRACT

RECURRENT EVENT PROCESS WITH INFORMATIVE CENSORING

HANDE UNLU

Doctor of Philosophy, Department of Statistics

Supervisor: Associated Professor Ayten Yigiter

September 2013, 101 pages

The aim of this study is to model recurrent event data in the presence of informative censoring and obtain the estimate of parameters.

Informative censoring is an important issue which have to be taken into account while modelling recurrent event data. Informative censoring might happen as a result of terminal event such as death or drop out of individuals in the study. Therefore, there is correlation between recurrent event time and terminal event time. Ignoring this correlation in the recurrent event data, biased estimates of model parameters could be obtained.

In this study, we focused on the situation in which informative censoring happen as a result of death and modelled recurrent event data using homogeneous Poisson Process. We proposed two different models based on intensity functions and constructed the structure of correlations between recurrent event time and death time via shared frailty. Maximum likelihood estimates of parameters are obtained using EM and Metropolis-Hastings algorithms. The proposed models are applied on the well known bladder tumor study by Byar (1976). To show the validity of our proposed models, the simulation study is conducted for different scenarios.

Keywords: Recurrent event data, informative censoring, counting process, homogeneous Poisson process, intensity function, EM algorithm, Metropolis-Hastings algorithm.

TEŞEKKÜR

Çalışmamın başlangıcından bitimine kadar, benden yardımlarını esirgemeyen pek çok kişiye teşekkür borçluyum.

Tez çalışmamın son döneminde beraber olma fırsatını yakaladığım sabrı, içten tavırları ve güven veren duruşu ile çalışmamın sonlandırılmasında büyük pay sahibi olan değerli danışmanım Doç. Dr. Ayten YİĞİTER' e,

Tez çalışmam sırasında beni Amerika'ya davet eden ve Amerika'da kaldığım süre boyunca yaptığı bilimsel katkılarla tez konumun şekillenmesini sağlayan başta Prof. Dr. Debashis GHOSH olmak üzere The Penn State Üniversitesi İstatistik Bölümü hocalarına,

Çalışma sırasında bilimsel katkıları ile bana yardımcı olan, eğitimim süresince yardımlarını esirgemeyen, eski tez danışmanım ve hocam Prof. Dr. Öniz TOKTAMIŞ'a,

Hacettepe Üniversitesi İstatistik Bölümü'nde geçirdiğim şu zamana kadar desteklerini ve samimiyetlerini yürekten hissettiğim sevgili bölüm başkanımız Prof. Dr. Hülya ÇINGI' ya ve başta Doç. Dr. Serpil AKTAŞ ALTUNAY, Doç. Dr. Meral Candan ÇETİN, Yrd. Doç. Dr. Canan HAMURKAROĞLU ve Doç. Dr. Çağdaş Hakan ALADAĞ olmak üzere tüm hocalarıma,

Benden desteklerini ve yardımlarını esirgemeyen Hacettepe Üniversitesi İstatistik Bölümü araştırma görevlileri başta Onur TOKA, Nurbanu BURSA, Özge KARADAĞ ve İrem TATAR olmak üzere tüm çalışma arkadaşlarıma,

Hacettepe Üniversitesi İstatistik Bölümü yüksek lisans öğrencilerinden Bülent ALPTEKİN'e,

Desteğini benden hiçbir zaman esirgemeyen değerli arkadaşım Sibel ALADAĞ'a

Bugün sahip olduğum her şeyimi borçlu olduğum sevgili annem ve babama, en zor zamanlarımda sabrı ve desteği ile yanımda olan sevgili eşime,

Tez çalışmamın şekillenmesinde tecrübeleri ile bana ışık tutan sevgili Tez İzleme Komitesi üyeleri ve Tez Jürisi üyelerine,

Doktora eğitimimi destekleyerek bana maddi olanak sağlayan ve daima bilime katkıda bulunan TÜBİTAK kurumuna,

içten teşekkürlerimi sunarım.

İÇİNDEKİLER

	<u>Sayfa</u>
ÖZET	i
ABSTRACT	ii
TEŞEKKÜR	iii
İÇİNDEKİLER.....	iv
ŞEKİLLER	vii
ÇİZELGELER	viii
Raporda Geçen Terimlerin İngilizce ve Türkçe Karşılıkları	x
1. GİRİŞ.....	1
2. TEK DEĞİŞKENLİ YAŞAM ANALİZİ	3
2.1. Yaşam Analizinde Kullanılan Fonksiyonlar	3
2.1.1. Olasılık Yoğunluk Fonksiyonu	4
2.1.2. Yaşam Fonksiyonu.....	4
2.1.3. Tehlike Fonksiyonu.....	5
2.1.4. Birikimli Tehlike Fonksiyonu	6
2.2. Durdurma (Sansür).....	7
2.2.1. Sağdan Durdurma.....	7
2.2.2. Soldan durdurma	8
2.2.3. Aralık durdurması	8
2.3. Regresyon Modelleri	8
2.3.1. Cox Orantılı Tehlike Modeli	9
2.3.2. Parametrik Orantılı Tehlike Modelleri.....	13
2.3.2.1. Weibull Orantılı Tehlike Modeli	14
2.3.2.2. Üstel Orantılı Tehlike Modeli.....	16
2.3.2.3. Gompertz Orantılı Tehlike Modeli	17

2.3.3.	Hızlandırılmış Başarısızlık Zamanı Modelleri.....	18
2.4.	Zayıflık Modelleri	22
2.4.1.	Paylaşılmamış Zayıflık Modeli	24
2.4.2.	Paylaşılmış Zayıflık Modeli.....	24
2.5.	Olabilirlik Fonksiyonu ve Parametre Tahminleri	25
3.	YİNELEMELİ OLAY ANALİZİ.....	29
3.1.	Sayma Süreci Notasyonu	32
3.2.	Yinelemeli Olay Analizinde Kullanılan Temel Fonksiyonlar	35
3.3.	Yinelemeli Olay Analizinde Kullanılan Temel Modeller.....	36
3.3.1.	Prentice, William ve Peterson Yaklaşımı.....	36
3.3.2.	Wei, Lin ve Weissfeld Yaklaşımı	39
3.3.3.	Andersen ve Gill Yaklaşımı	41
3.3.4.	Pepe ve Cai Tarafından Önerilen Oran Modeller	42
3.3.5.	Hızlandırılmış Başarısızlık Zamanı Modelleri.....	43
3.3.6.	Paylaşılmış Zayıflık Modelleri.....	45
4.	POISSON SÜREÇLERİ.....	48
4.1.	Homojen Poisson Süreci	48
4.2.	Homojen Olmayan Poisson Süreci.....	49
4.3.	Homojen Poisson Süreci için Olabilirlik Fonksiyonunun Çıkarımı	50
4.4.	Poisson Süreçlerinin Kullanıldığı Yinelemeli Olay Modelleri	53
5.	BİLGİ İÇEREN DURDURMA VARLIĞINDA YİNELEMELİ OLAY ANALİZİ	56
5.1.	Bilgi İçeren Durdurma Varsayımı Altında Yapılan Çalışmalar.....	56
5.2.	Liu, Wolfe ve Huang Tarafından Önerilen Yöntem	60
5.3.	Önerilen Modeller	65
5.3.1.	Önerilen Birinci Model	65
5.3.2.	Önerilen İkinci Model	70

6. UYGULAMA VE BENZETİM ÇALIŞMASI.....	81
6.1. Mesane Kanseri Verisi Üzerine Bir Uygulama.....	81
6.2. Benzetim Çalışması.....	83
7. SONUÇLAR VE TARTIŞMA.....	93
ÖZGEÇMİŞ.....	100

ŞEKİLLER

	<u>Sayfa</u>
Şekil 2.1. Teorik olarak elde edilen yaşam fonksiyonu.....	5
Şekil 2.2. Uygulamada elde edilen yaşam fonksiyonu.....	5
Şekil 3.1. Sayma sürecinin grafiksel gösterimi.....	33
Şekil 3.2. Risk sürecinin grafiksel gösterimi.....	34
Şekil 4.1 $[a, b]$ aralığının çok küçük aralıklara parçalanışı.....	51

ÇİZELGELER

	<u>Sayfa</u>
Çizelge 1. Önerilen Birinci Model İçin Mesane Kanseri Verisi Sonuçları.....	82
Çizelge 2. Önerilen İkinci Model İçin Mesane Kanseri Verisi Sonuçları.....	82
Çizelge 3. n=40 ve Durdurma Oranı=0.01 Alındığında Önerilen Birinci Model İçin Benzetim Sonuçları.....	84
Çizelge 4. n=40 ve Durdurma Oranı =0.05 Alındığında Önerilen Birinci Model İçin Benzetim Sonuçları.....	84
Çizelge 5. n=40 ve Durdurma Oranı =0.10 Alındığında Önerilen Birinci Model İçin Benzetim Sonuçları.....	85
Çizelge 6. n=60 ve Durdurma Oranı =0.01 Alındığında Önerilen Birinci Model İçin Benzetim Sonuçları.....	85
Çizelge 7. n=60 ve Durdurma Oranı =0.05 Alındığında Önerilen Birinci Model İçin Benzetim Sonuçları.....	85
Çizelge 8. n=60 ve Durdurma Oranı =0.10 Alındığında Önerilen Birinci Model İçin Benzetim Sonuçları.....	86
Çizelge 9. n=100 ve Durdurma Oranı =0.01 Alındığında Önerilen Birinci Model İçin Benzetim Sonuçları.....	87
Çizelge 10. n=100 ve Durdurma Oranı =0.05 Alındığında Önerilen Birinci Model İçin Benzetim Sonuçları.....	87
Çizelge 11. n=100 ve Durdurma Oranı =0.10 Alındığında Önerilen Birinci Model İçin Benzetim Sonuçları.....	87
Çizelge 12. n=40 ve Ölüm Oranı=0.20 Alındığında Önerilen İkinci Model İçin Benzetim Sonuçları.....	89
Çizelge 13. n=40 ve Ölüm Oranı=0.50 Alındığında Önerilen İkinci Model İçin Benzetim Sonuçları.....	89
Çizelge 14. n=40 ve Ölüm Oranı=0.70 Alındığında Önerilen İkinci Model İçin Benzetim Sonuçları.....	89

Çizelge 15. n=60 ve Ölüm Oranı =0.20 Alındığında Önerilen İkinci Model İçin Benzetim Sonuçları.....	90
Çizelge 16. n=60 ve Ölüm Oranı =0.50 Alındığında Önerilen İkinci Model İçin Benzetim Sonuçları.....	90
Çizelge 17. n=60 ve Ölüm Oranı =0.70 Alındığında Önerilen İkinci Model İçin Benzetim Sonuçları.....	90
Çizelge 18. n=100 ve Ölüm Oranı =0.20 Alındığında Önerilen İkinci Model İçin Benzetim Sonuçları.....	91
Çizelge 19. n=100 ve Ölüm Oranı =0.50 Alındığında Önerilen İkinci Model İçin Benzetim Sonuçları.....	91
Çizelge 20. n=100 ve Ölüm Oranı =0.70 Alındığında Önerilen İkinci Model İçin Benzetim Sonuçları.....	91

Raporda Geçen Terimlerin İngilizce ve Türkçe Karşılıkları

Survival Analysis	:	Yaşam Analizi
Failure Time Analysis	:	Başarısızlık Zamanı Analizi
Event Time Analysis	:	Olay Zamanı Analizi
Failure	:	Başarısızlık
Censored	:	Durdurulmuş ya da Sansürlü
Left Censoring	:	Soldan Durdurma
Right Censoring	:	Sağdan Durdurma
Interval Censoring	:	Aralık Durdurması
Frailty	:	Zayıflık
Shared Frailty Models	:	Paylaşılmış Zayıflık Modelleri
Unshared Frailty Models	:	Paylaşılmamış Zayıflık Modelleri
Noninformative Censoring	:	Bilgi İçermeyen Durdurma
Informative Censoring	:	Bilgi İçeren Durdurma
Competing Risk Data	:	Yarışan Risk Verisi
Recurrent Event Data	:	Yinelemeli Olay Verisi
Intensity Function	:	Yoğunluk Fonksiyonu
Rate Function	:	Oran Fonksiyonu
Mean Function	:	Ortalama Fonksiyonu
Product Integral	:	Çarpım İntegrali
Cumulative Intensity Function	:	Birikimli Yoğunluk Fonksiyonu
Multiple Event Data	:	Çoklu Olay Verisi
Coverage Probability	:	Kapsama Olasılığı
Quasi Likelihood	:	Yarı Olabilirlik

1. GİRİŞ

Klinik arařtırmalarda sıklıkla kullanılan yařam analizi, belli bir bařlangıç anından itibaren ilgilenilen olay gerekleřene dek geen zamanın incelenmesinde kullanılan istatistiksel bir yontemdir. Yařam analizinde zaman, yařam zamanı, yařam sũrdũrme zamanı ya da bařarisızlık zamanı olarak da adlandırılır. Yařam analizini diđer istatistiksel analizlerden ayıran en ˆnemli ˆzellik, bu analizde durdurulmuř ya da sansũrlũ gˆzlemlerin yer almasıdır.

Yařam verisinin Cox (1972) tarafından geliřtirilen regresyon modelleri ile modellenmesi, daha sonraki teorik ve uygulamalı arařtırmalar iin bir temel oluřturmuřtur [1]. Literatũrde, Cox tarafından ˆnerilen regresyon modelinin geliřtirilmesiyle oluřan ok sayıda alıřma bulunmaktadır. Ayrıca yařam verisi iinde durdurma mekanizmasının bulunması, modellemede eřitliliđe neden olmaktadır.

Yinelemeli olay analizi son yıllarda ˆzerinde alıřılan bir konudur. Yinelemeli olaylarda bireyler aynı olayı birden fazla deneyimleyebilirler. Bu nedenle aynı bireyden elde edilen yineleme zamanları iliřkilidir. Yinelemeli olaylara ˆrnek olarak epilepsi hastalarındaki nˆbetler, kanser vakalarında tekrarlayan tũmˆrler ya da ˆretim donanımlarında tekrarlayan arızalar verilebilir.

Yinelemeli olay verileri, Cox tipi regresyon modellerinin geliřtirilmesine dayanmakla birlikte; bu verilerin karakteristiklerini yansıtacak derecede tam bir modelleme yapılması olanaksızdır. Aynı zamanda bu verilenin modellenmesi ve analizi istatistik biliminin gũndeminde olan konularından biridir.

Bunun yanı sıra, yinelemeli olay verisinin modellenmesinde durdurma olayının yapısı ok iyi irdelenmelidir. ođu arařtırmada, bařarisızlık zamanı; ˆlũm gibi bir bařarisızlık olayı, alıřmanın sona ermesi ya da izlenme dˆneminde bireyin alıřmadan ekilmesi sebepleriyle durdurmaya maruz kalabilir. Eđer durdurma, alıřmanın sona ermesinden kaynaklanıyorsa, durdurma zamanı bađımsız ya da bilgi iermeyen řeklinde kabul edilir.

Birçok uygulamada yinelemeli olaylar, bireyin çalışmadan çekilmesi ya da bireyin ölümü ile sonlandırılır. Bu şekilde meydana gelen durdurma zamanı bilgi içerir ve yinelemeli olay zamanları ile ilişkilidir. Söz konusu ilişkiyi göz ardı ederek modelleme yapmak yanlış sonuçlar elde edilmesine neden olmaktadır.

Tez çalışmasının amacı, bilgi içeren durdurma varlığında yinelemeli olay verisini modellemek ve parametre tahminlerini elde etmektir. Çalışmanın İkinci Bölümü'nde, tek değişkenli yaşam analizinde kullanılan modeller açıklanmış ve gerekli tanımlamalar verilmiştir. Üçüncü Bölüm' de, yinelemeli olay verisi hakkında bilgi verilmiş ve bu verinin modellenmesinde kullanılan yarı parametrik ve parametrik olmayan tahmin yöntemlerinden bahsedilmiştir. Dördüncü Bölüm' de, Poisson süreçleri hakkında bilgi verilmiş ve yinelemeli olay verisini Poisson süreçleri kullanarak modelleyen çalışmalar anlatılmıştır. Beşinci Bölüm' de, bilgi içeren durdurma varlığında yinelemeli olay verisi için önerilen modeller detaylı bir biçimde açıklanmıştır. Altıncı Bölüm' de, önerilen modellerin gerçek bir veri kümesi üzerindeki sonuçlarına ve benzetim çalışmasına yer verilmiştir. Yedinci Bölüm ise sonuç ve tartışma bölümüdür.

2. TEK DEĞİŞKENLİ YAŞAM ANALİZİ

Yaşam analizi, belli bir başlangıç anından itibaren ilgilenilen olay gerçekleşene dek geçen zamanın incelenmesinde kullanılan istatistiksel bir yöntemdir. Yaşam analizine, yaşam sürdürme analizi, başarısızlık zamanı analizi ya da olay zamanı analizi de denilmektedir. Burada olay, ölüm, hastalık ya da araştırmacının ilgilendiği herhangi bir olay olabilir. Olay, genellikle ölüm, hastalık gibi olumsuz bir deneyim olduğu için “başarısızlık” olarak adlandırılır. Fakat bazı durumlarda olay ya da başarısızlık olumlu bir deneyim de olabilir. Örneğin; hastaneden ayrılma, hastalığın iyileşmesi ya da bir sınavın bitmesi gibi. Yaşam analizinde zaman, yaşam zamanı, yaşam sürdürme zamanı ya da başarısızlık zamanı olarak da adlandırılır. Başarısızlık zamanı denmesinin nedeni, belli bir başlangıç anından olayın ya da başarısızlığın gerçekleştiği zamana dek geçen süreyi vermesidir. Burada zaman, yıl, hafta, gün veya olayın ortaya çıktığı andaki yaş olabilir. Başarısızlık zamanına örnek olarak, herhangi bir fabrikada makine parçalarının bozulmasına dek geçen zaman ya da sağlıklı insanların kalp krizi geçirdiği andaki yaşları verilebilir [2].

Yaşam analizinde temel olan gözlenen başarısızlık sürelerinin incelenmesi olduğundan bu değişkenin iyi tanımlanması gerekmektedir. Başlangıç zamanı her bir birim ya da birey için kuşkuyla yer bırakmayacak şekilde belirlenmeli ve başarısızlığın ortaya çıktığı an net olmalıdır [3].

Yaşam analizinde ilk adım, birim ya da bireylerin başarısızlık zamanlarını sayısal ya da grafiksel olarak belirlemektir. İkinci adım, başarısızlık zamanlarına ilişkin yaşam fonksiyonunu ya da tehlike fonksiyonunu tahmin etmektir.

2.1. Yaşam Analizinde Kullanılan Fonksiyonlar

Yaşam analizinde olasılık yoğunluk, yaşam, tehlike ve birikimli tehlike fonksiyonu olmak üzere dört temel fonksiyon kullanılmaktadır.

2.1.1. Olasılık Yoğunluk Fonksiyonu

T , bir bireyin başarısızlık zamanını gösteren negatif olmayan bir rastlantı değişkeni olmak üzere; T rastlantı değişkeninin olasılık yoğunluk fonksiyonu,

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T \leq t + \Delta t)}{\Delta t} \quad (2.1)$$

biçiminde gösterilir. Bir birey ya da birimin t anındaki başarısız olma yoğunluğunu verir [4].

2.1.2. Yaşam Fonksiyonu

Yaşam fonksiyonu $S(t)$, bir birimin ya da bireyin belirli bir t anından daha uzun süre yaşama olasılığını verir. Yaşam fonksiyonu,

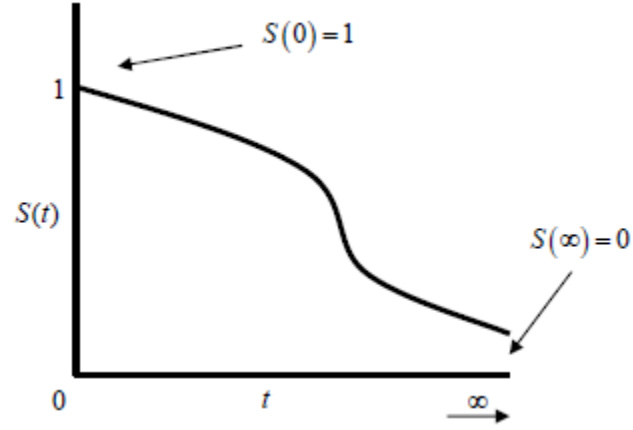
$$S(t) = P(T > t) = \int_t^{\infty} f(t) dt \quad (0 < t < \infty) \quad (2.2)$$

eşitliğindeki gibi yazılabilir. $S(t)$, $[0, 1]$ aralığında değer alan bir fonksiyon olup,

$$S(t) = 1 - F(t) \quad (2.3)$$

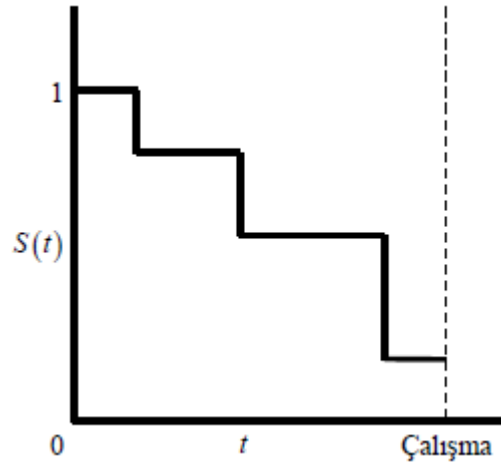
olarak yazılabilir.

Yaşam fonksiyonu monoton azalan bir fonksiyondur. Teoride yaşam fonksiyonu $S(t)$, eğrisel olarak Şekil 2.1' deki gibidir:



Şekil 2.1. Teorik olarak elde edilen yaşam fonksiyonu

Uygulamada ise, yaşam fonksiyonu basamak fonksiyonu şeklinde olup, Şekil 2.2' de gösterildiği gibidir:



Şekil 2.2. Uygulamada elde edilen yaşam fonksiyonu

2.1.3. Tehlike Fonksiyonu

Tehlike fonksiyonu, t anına dek yaşadığı bilinen bir bireyin $[t, t + \Delta t)$ aralığında başarısızlık yaşama riskini gösterir. Tehlike fonksiyonu;

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \quad (2.4)$$

biçiminde ifade edilir.

2.1.4. Birikimli Tehlike Fonksiyonu

Birikimli tehlike fonksiyonu aşağıdaki eşitlikte verildiği gibidir:

$$H(t) = \int_0^t h(u) du. \quad (2.5)$$

Eşitlik (2.1), (2.2) ve (2.4) ile verilen fonksiyonlar, birbirleriyle ilişkili fonksiyonlar olup aralarındaki ilişkiler Eşitlik (2.6) ve (2.7) ile verilmiştir.

Koşullu olasılık tanımından tehlike fonksiyonu,

$$h(t) = P\{t < T < t + \Delta t \mid T > t\} = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{S(t)} \quad (2.6)$$

biçiminde yazılabilir.

Başarısızlık zamanına ilişkin olasılık yoğunluk fonksiyonu $f(t)$ ile yaşam fonksiyonu $S(t)$ arasındaki bağıntı ise,

$$f(t) = \frac{dF(t)}{dt} = \frac{d(1 - S(t))}{dt} = -S'(t) \quad (2.7)$$

eşitliğindeki gibi ifade edilebilir [2, 5].

2.2. Durdurma (Sansür)

Yaşam analizini diğer istatistiksel analizlerden ayıran en önemli özellik, bu analizde durdurulmuş ya da sansürlü gözlemlerin yer almasıdır. Yaşam analizinde, ilgilenilen olay önceden belirlenen zaman aralığında ortaya çıkmayabilir ya da bireyin (birimin) takibi yapılamamış olabilir. Bu durumda bu gözlemlere ilişkin gerçek başarısızlık zamanı elde edilemez. Bu şekilde ortaya çıkan gözlemlere, sansürlü ya da durdurulmuş gözlem denir. İncelenen olayın yapısına ve elde edilen verilerin türüne göre birçok durdurma türünden bahsedilebilir. Bunlar,

- Sağdan durdurma,
- Soldan durdurma
- Aralık durdurmasıdır.

2.2.1. Sağdan Durdurma

Sağdan durdurma en çok karşılaşılan durdurma türüdür. Bu durdurma türünde birey ya da birime ait gerçek başarısızlık zamanı, çalışmanın süresinden başka bir deyişle takip süresinden fazla olduğu bilinmektedir. Sağdan durdurma,

- Çalışma bitene kadar başarısızlığın gözlenememiş olması sebebiyle,
- Bireyin çalışmadan çekilmesi sebebiyle
- Bireyin başka bir olay (başka nedenden ölüm) sebebi ile çalışmadan çekilmesi sonucu ortaya çıkabilir.

Sağdan durdurma, 1. Tip durdurma ve 2. Tip durdurma olarak ikiye ayrılır:

1. Tip durdurma, çalışmanın önceden belirlenen bir takip süresi sonunda sonlandırıldığı durdurma türü,

2. Tip durdurma, çalışmanın, önceden belirlenen sayıda başarısızlığın ortaya çıkmasıyla sonlandırıldığı durdurma türüdür.

2.2.2. Soldan durdurma

İlgilenilen olayın başlangıç zamanının tam olarak bilinmediği durdurma türüdür. Soldan durdurulmuş veride, başarısızlık zamanı takip süresinden küçüktür. Bu durdurma türünü daha iyi açıklayabilmek için şu örnek verilebilir: Belirli bir kanser türüne sebep olan primer bir tümörün, cerrahi bir müdahale ile alındıktan sonra kanserin yinelenmesine dek geçen zamanın incelendiği bir çalışma düşünelim. Cerrahi müdahaleden üç ay sonra, kanserin nüksedip nüksetmediğini anlamak için hastalar muayene edilmiş ve bazı hastalarda kanserin nüksettiği tespit edilmiştir. Bu hastalar için kanserin gerçek nüksetme zamanı üç aydan az olup, nüksetme zamanları soldan durdurulmuştur [5, 6].

2.2.3. Aralık durdurması

Aralık durdurması, birey ya da birimlerin, başarısızlığı deneyimlediğinin bilindiği fakat gerçek başarısızlık zamanlarının tam olarak gözlenemeyip, bir zaman aralığında ortaya çıktığı durdurma türüdür. Soldan durdurma için verilen kanser örneği tekrar ele alınsın. Hastanın, cerrahi müdahaleden sonraki üç aylık zaman diliminde kanserin nüksetmediği bilindiğinde, ancak cerrahi müdahaleden altı ay sonra muayene edildiği zaman kanserin tekrar nüksettiği tespit edilirse, gerçek nüksetme zamanının cerrahi müdahaleden sonraki üç ay ile altı ay arasında bir zamanda gerçekleştiği söylenebilir [5, 6].

2.3. Regresyon Modelleri

Yaşam verisinin modellenmesinde regresyon modelleri yoğun biçimde kullanılmaktadır. Bu modellerden, Cox orantılı tehlike modeli, parametrik orantılı tehlike modelleri ve hızlandırılmış başarısızlık zamanı modelleri bu kesimde açıklanmıştır.

2.3.1. Cox Orantılı Tehlike Modeli

Başarısızlık zamanını modellemek için kullanılan en temel model, orantılı tehlike modelidir. Cox tarafından önerildiğinden Cox orantılı tehlike modeli ya da Cox regresyon modeli olarak da adlandırılmaktadır [1, 3, 7].

Cox orantılı tehlike modeli,

$$h(t|\mathbf{Z}) = h_0(t) \exp(\beta_1 Z_1 + \beta_2 Z_2 + \dots + \beta_p Z_p) = h_0(t) \exp(\boldsymbol{\beta}^T \mathbf{Z}) \quad (2.8)$$

biçimindedir.

$h(t|\mathbf{Z})$, \mathbf{Z} p boyutlu açıklayıcı değişken vektörü verildiğinde, başarısızlık zamanı raslantı değişkeni T 'nin tehlike fonksiyonu olarak tanımlanır. Eşitlik (2.8)'deki $\boldsymbol{\beta}$, regresyon parametrelerini ve $h_0(t)$, temel tehlike fonksiyonunu göstermektedir.

Cox orantılı tehlike modeli varsayımı altında i . bireye ait yaşam fonksiyonu,

$$S(t|\mathbf{Z}) = [S_0(t)]^{\exp(\boldsymbol{\beta}^T \mathbf{Z})} \quad (2.9)$$

biçimindedir.

Açıklayıcı değişken değerleri, sırasıyla Z_1 ve Z_2 olan iki bireyin tehlike oranı,

$$\frac{h(t|Z_1)}{h(t|Z_2)} = \frac{h_0(t) \exp(\beta Z_1)}{h_0(t) \exp(\beta Z_2)} = \exp(\beta(Z_1 - Z_2)) \quad (2.10)$$

biçiminde olup, Eşitlik (2.10)' dan tehlike oranının zamandan bağımsız olduğu görülebilir. Bu modelin, temel varsayımı tehlike oranlarının zaman boyunca sabit kaldığı ya da bir bireye ilişkin tehlikenin diğer bireye ilişkin tehlikeye orantılı olduğudur. Bu yüzden, Cox orantılı tehlike modeli olarak da adlandırılmaktadır.

Cox orantılı tehlike modeli, orantılı tehlike varsayımına dayanmasına rağmen başarısızlık zamanı, T raslantı değişkeni için belirli bir olasılık dağılımı yoktur. Bu nedenle de Cox orantılı tehlike modeli, yarı parametrik bir model olarak ele alınmaktadır [1, 2, 5].

Cox orantılı tehlike modelinin üstün tarafı, başarısızlık zamanı, T raslantı değişkeni için herhangi bir dağılım varsayımı yapılmamasına rağmen parametre tahminlerinin yapılabilmesidir.

Cox orantılı tehlike modelinde parametre tahminlerinin elde edilişi açıklanmadan önce bazı temel kavramlar tanıtılacaktır.

$a \wedge b = \min(a, b)$ olsun. T_i^* , i . bireyin gerçek başarısızlık zamanını, C_i durdurma zamanını ve \mathbf{Z} , p boyutlu açıklayıcı değişken vektörünü gösterebilir. Bu durumda bir birey için $\{T, \delta, \mathbf{Z}\}$ kümesinin gözlenen $\{t_i, \delta_i, \mathbf{Z}_i\}$ ($i = 1, 2, \dots, n$) değerlerini oluşturur. Burada $T = T^* \wedge C$ ve $\delta = I(T^* \leq C)$ dir. Ayrıca \mathbf{Z} verildiğinde T^* ve C 'nin bağımsız olduğu varsayılmaktadır.

n bireye ait başarısızlık zamanları (t_1, t_2, \dots, t_n) olsun. r bireye ait sıralanmış başarısızlık zamanları, $t_{(1)} < t_{(2)} < \dots < t_{(r)}$ ve $t_{(j)}$ zamanına dek hayatta kalan ve durdurulmamış bireylerden oluşan risk kümesi $R(t_{(j)})$ ile gösterilsin. $R(t_{(j)})$ kümesinden $t_{(j)}$ anında bir bireyin başarısız olduğu verildiğinde i . bireyin $t_{(j)}$ anında başarısız olmasının koşullu olasılığı,

$P(i. \text{ bireyin } t_{(j)} \text{ anında başarısız olması} \mid t_{(j)} \text{ anında } R(t_{(j)}) \text{ kümesinden bir bireyin başarısız olması})$

$$\begin{aligned}
&= \frac{P(i. \text{ bireyin } t_{(j)} \text{ anında başarısız olması})}{P(t_{(j)} \text{ anında } R(t_{(j)}) \text{ kümesinde bir başarısızlık olması})} \\
&= \frac{P(i. \text{ bireyin } t_{(j)} \text{ anında başarısız olması})}{\sum_{k \in R(t_{(j)})} P(t_{(j)} \text{ anında } R(t_{(j)}) \text{ kümesinde } k. \text{ bireyin başarısız olması})} \\
&\approx \frac{P(i. \text{ bireyin } (t_{(j)}, t_{(j)} + \Delta t) \text{ aralığında başarısız olması}) / \Delta t}{\sum_{k \in R(t_{(j)})} P(t_{(j)}, t_{(j)} + \Delta t \text{ aralığında } R(t_{(j)}) \text{ kümesinde } k. \text{ bireyin başarısız olması}) / \Delta t} \\
&= \frac{\lim_{\Delta t \downarrow 0} P(i. \text{ bireyin } (t_{(j)}, t_{(j)} + \Delta t) \text{ aralığında başarısız olması}) / \Delta t}{\lim_{\Delta t \downarrow 0} \sum_{k \in R(t_{(j)})} P(t_{(j)}, t_{(j)} + \Delta t \text{ aralığında } R(t_{(j)}) \text{ kümesinde } k. \text{ bireyin başarısız olması}) / \Delta t} \\
&= \frac{h_i(t_{(j)})}{\sum_{k \in R(t_{(j)})} h_k(t_{(j)})} \\
&= \frac{h_0(t_{(j)}) \exp\{\boldsymbol{\beta}^T \mathbf{Z}_i(t_{(j)})\}}{\sum_{k \in R(t_{(j)})} h_0(t_{(j)}) \exp\{\boldsymbol{\beta}^T \mathbf{Z}_k(t_{(j)})\}} \\
&= \frac{\exp\{\boldsymbol{\beta}^T \mathbf{Z}_i(t_{(j)})\}}{\sum_{k \in R(t_{(j)})} \exp\{\boldsymbol{\beta}^T \mathbf{Z}_k(t_{(j)})\}}
\end{aligned}$$

olarak yazılabilir.

r birey üzerinden bu olasılıkların çarpımı kısmi olasılık fonksiyonunu gösterir ve kısmi olabilirlik fonksiyonu,

$$\mathbf{L}(\boldsymbol{\beta}) = \prod_{j=1}^r \frac{\exp\{\boldsymbol{\beta}^T \mathbf{Z}_i(t_{(j)})\}}{\sum_{k \in R(t_{(j)})} \exp\{\boldsymbol{\beta}^T \mathbf{Z}_k(t_{(j)})\}} \quad (2.11)$$

biçimindedir.

Eşitlik (2.11) durdurulmamış gözlemler içindir. δ_i , i . bireye ait başarısızlık zamanı durdurulmuş ise 0, değilse 1 değerini alan gösterge fonksiyonu olmak üzere, n bireye ait başarısızlık zamanları (t_1, t_2, \dots, t_n) için, Eşitlik (2.11)'de verilen olabilirlik fonksiyonu,

$$L(\boldsymbol{\beta}) = \prod_{i=1}^n \left[\frac{\exp\{\boldsymbol{\beta}^T \mathbf{Z}_i(t_{(j)})\}}{\sum_{k \in R(t_{(j)})} \exp\{\boldsymbol{\beta}^T \mathbf{Z}_i(t_{(j)})\}} \right]^{\delta_i} \quad (2.12)$$

biçiminde ifade edilebilir.

t_i anındaki risk kümesi $Y_j(t_i) = I(k \in R(t_i))$ olarak alındığında Eşitlik (2.12),

$$L(\boldsymbol{\beta}) = \prod_{i=1}^n \left[\frac{\exp\{\boldsymbol{\beta}^T \mathbf{Z}_i(t_{(j)})\}}{\sum_{j=1}^n Y_j(t_i) \exp\{\boldsymbol{\beta}^T \mathbf{Z}_i(t_{(j)})\}} \right]^{\delta_i} \quad (2.13)$$

biçiminde bulunur.

Cox orantılı tehlike modelinde, parametre tahminleri kısmi olabilirlik fonksiyonunu maksimum yapan değerlerdir. Başka bir deyişle, parametre tahminleri, Cox orantılı tehlike fonksiyonuna ait skor fonksiyonunu sıfır yapan değerlerdir. Skor fonksiyonu,

$$U(\boldsymbol{\beta}) = \frac{\partial \log L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \sum_{i=1}^n \delta_i [\mathbf{Z}_i - \bar{\mathbf{Z}}_i(\boldsymbol{\beta})] = 0 \quad (2.14)$$

olmak üzere,

$$\bar{Z}(\boldsymbol{\beta}) = \sum_{i=1}^n Y_j(t_i) \mathbf{Z}_j p_j(\boldsymbol{\beta}; t_i) \quad (2.15)$$

biçimindedir. Eşitlik (2.15)'te $p_j(\boldsymbol{\beta}; t_i)$ koşullu olasılığı,

$$p_j(\boldsymbol{\beta}, t_i) = \frac{\exp\{\boldsymbol{\beta}^T \mathbf{Z}_j(t_j)\}}{\sum_{l=1}^n Y_l(t_i) \exp\{\boldsymbol{\beta}^T \mathbf{Z}_l(t_l)\}} \quad (2.16)$$

olarak yazılır [8, 9].

Tehlike modellerinde başarısızlık zamanı, T raslantı değişkeni için belirli bir olasılık dağılımı tanımlanırsa, bu modeller parametrik tehlike modelleri olarak adlandırılır.

2.3.2. Parametrik Orantılı Tehlike Modelleri

Yaşam analizinde kullanılan Cox orantılı tehlike modelinde, başarısızlık zamanları için herhangi bir dağılım varsayımı yoktur. Bu nedenle tehlike fonksiyonu herhangi bir fonksiyonel yapı ile kısıtlanmamıştır. Bu yüzden model geniş uygulama alanına ve esnekliğe sahiptir. Diğer yandan, başarısızlık zamanı için geçerli bir olasılık dağılımı varsa çıkarsamalar daha kesin olacaktır. Göreli tehlikeler ve başarısızlık zamanlarının medyanı gibi niceliklerin tahminleri, başarısızlık zamanı için herhangi bir dağılım varsayımının yapılmadığı modellere (Örneğin Cox orantılı tehlike modeline) göre daha küçük standart hatalara sahip olacaktır [5].

Orantılı tehlike yapısına uyan ve yaşam zamanları için kullanılan çeşitli olasılık dağılımlarının yer aldığı parametrik modeller aşağıda verilmektedir:

Parametrik orantılı tehlike modeli, Cox orantılı tehlike modelinin parametrik uyarlamasıdır. p boyutlu açıklayıcı değişken vektörüne $\mathbf{Z}^T = (Z_1, Z_2, \dots, Z_p)$ sahip bir bireyin t zamanındaki tehlike fonksiyonu,

$$h(t|\mathbf{Z}) = h_0(t) \exp(\beta_1 Z_1 + \beta_2 Z_2 + \dots + \beta_p Z_p) = h_0(t) \exp(\boldsymbol{\beta}^T \mathbf{Z})$$

biçiminde ifade edilir.

Parametrik orantılı tehlike modeliyle, Cox orantılı tehlike modeli arasındaki temel fark, parametrik orantılı tehlike modelinde temel tehlike fonksiyonun belirli bir dağılıma sahip olmasıdır. Cox orantılı tehlike modelinde parametreler, kısmi olabilirlik yöntemi ile kestirilirken, parametrik orantılı tehlike modelinde en çok olabilirlik yöntemi ile kestirilir. Bunun dışında, bu iki model arasında fark yoktur. Tehlike oranlarının yorumları ve tehlikelerin orantılılığı aynıdır.

Farklı tehlike fonksiyonları kullanılarak, birçok parametrik model elde edilebilir. En çok kullanılanları Weibull, Üstel ve Gompertz modelleridir [5].

2.3.2.1. Weibull Orantılı Tehlike Modeli

Başarısızlık zamanının, ölçek parametresi λ ve biçim parametresi γ olan Weibull dağılımına sahip olduğu varsayalım. $Weibull(\lambda, \gamma)$ dağılımına ait tehlike ve yaşam fonksiyonu sırasıyla,

$$h(t) = \lambda \gamma (t)^{\gamma-1}, \quad S(t) = \exp(-\lambda t^\gamma) \quad \lambda > 0, \gamma > 0$$

biçimindedir.

$\gamma > 1$ için zaman geçtikçe tehlike hızı artarken, $\gamma < 1$ için tehlike hızı azalmaktadır. $\mathbf{Z}^T = (Z_1, Z_2, \dots, Z_p)$ açıklayıcı değişken vektörüne sahip bir birey için, Weibull orantılı tehlike modeli varsayımı altında tehlike fonksiyonu,

$$h(t|\mathbf{Z}) = \lambda \gamma t^{\gamma-1} \exp(\beta_1 Z_1 + \beta_2 Z_2 + \dots + \beta_p Z_p) = \lambda \gamma (t)^{\gamma-1} \exp(\boldsymbol{\beta}^T \mathbf{Z}) \quad (2.17)$$

eşitliğindeki gibi bulunur. Eşitlik (2.17)' den başarısızlık zamanının $\lambda \exp(\boldsymbol{\beta}^T \mathbf{Z})$ ölçek parametresi ve γ biçim parametresi ile Weibull dağılımına sahip olduğu görülebilir.

Bundan dolayı sabit γ değerine sahip Weibull dağılımı aileleri, orantılı tehlike özelliğine sahiptir. Bu açıklayıcı değişkenler, dağılımının ölçek parametresini değiştirir; ancak biçim parametresi üzerinde hiçbir etkiye sahip değildir.

Eşitlik (2.17)' deki tehlike fonksiyonuna karşılık gelen yaşam fonksiyonu,

$$S(t|\mathbf{Z}) = \exp\{-\exp(\boldsymbol{\beta}^T \mathbf{Z}) \lambda t^\gamma\} \quad (2.18)$$

biçimindedir. Weibull tehlike fonksiyonu, şekil parametresi γ ' nın değerlerine göre değişik formlar alabildiği ve uygun özet istatistiklerinin kolayca elde edilebildiği için yaşam verisinin parametrik analizinde yaygın olarak kullanılmaktadır [5].

Birikimli tehlike fonksiyonu, $H(t) = -\log S(t)$ ' dir. Weibull dağılımına sahip başarısızlık zamanına ilişkin yaşam fonksiyonunun logaritmik dönüşümü, birikimli tehlike fonksiyonunun logaritmik dönüşümünden yararlanarak,

$$\log\{-\log[S(t)]\} = \log \lambda + \gamma \log t$$

olarak elde edilir. $\log\{-\log[S(t)]\}$ ' nin $\log t$ ' ye karşı grafiği logaritmik birikimli tehlike fonksiyonunun grafiğidir. $\log\{-\log[S(t)]\}$ ' nin $\log t$ ' ye karşı grafiği çizildiğinde başlangıç noktasından geçen bir doğru elde ediliyorsa, başarısızlık zamanının Weibull dağıldığı varsayımı geçerlidir. İki grup için elde edilen doğrular paralel ise, orantılı tehlike model varsayımı geçerlidir [4, 5, 8].

2.3.2.2. Üstel Orantılı Tehlike Modeli

Üstel orantılı tehlike modeli, Weibull modelinde $\gamma = 1$ olduğu duruma karşılık gelir. Bu model altında, zaman boyunca tehlike fonksiyonunun sabit olduğu varsayılır. Üstel orantılı tehlike modelinde, başarısızlık zamanına ilişkin olasılık yoğunluk fonksiyonu,

$$f(t) = \lambda \exp(-\lambda t), \quad t \geq 0$$

dağılım fonksiyonu,

$$F(t) = \int_0^t \lambda \exp(-\lambda Z) dz = 1 - \exp(-\lambda t)$$

yaşam fonksiyonu,

$$S(t) = 1 - F(t) = \exp(-\lambda t)$$

biçimindedir. Böylece tehlike fonksiyonu,

$$h(t) = \frac{f(t)}{S(t)} = \frac{\lambda \exp(-\lambda t)}{\exp(-\lambda t)} = \lambda$$

olarak elde edilir. Üstel orantılı tehlike modeli varsayımı altında, bir bireyin tehlike fonksiyonu,

$$h(t|\mathbf{Z}) = \lambda \exp(\beta_1 Z_1 + \beta_2 Z_2 + \dots + \beta_p Z_p) = \lambda \exp(\boldsymbol{\beta}^T \mathbf{Z}). \quad (2.19)$$

eşitliğindeki gibi bulunur.

Bu model varsayımı altında, başarısızlık zamanına ilişkin yaşam fonksiyonunun logaritmik dönüşümünden sonra,

$$-\log[S(t)] = -\lambda t$$

olarak elde edilir. Eğer, $\log S(t)$ ' ye karşı t ' nin grafiği çizildiğinde başlangıç noktasından geçen bir doğru elde ediliyorsa başarısızlık zamanının Üstel dağıldığı varsayımı geçerlidir denir [5, 8].

2.3.2.3. Gompertz Orantılı Tehlike Modeli

Başarısızlık zamanı Gompertz dağılımına sahip ise tehlike ve yaşam fonksiyonları sırasıyla,

$$h(t) = \lambda \exp(\gamma t),$$

$$S(t) = \exp\left(\frac{\lambda}{\gamma}(1 - e^{\gamma t})\right), \quad 0 \leq t < \infty; \lambda > 0$$

biçimindedir

γ parametresi, tehlike fonksiyonunun biçimini belirler. $\gamma = 0$ olduğu zaman tehlike fonksiyonu $h(t) = \lambda$ olduğundan, başarısızlık zamanlarının dağılımının üstel dağılıma sahip olduğu söylenir. Ayrıca üstel dağılım, Gompertz dağılımının da özel bir durumudur. Weibull tehlike fonksiyonu gibi Gompertz tehlike fonksiyonu da düzgün bir şekilde artar ya da azalır. Gompertz tehlike fonksiyonunun logaritması alınır ise,

$$\log h(t) = \log \lambda + \gamma t$$

elde edilir. Yukarıdaki eşitlikten görüldüğü gibi $\log h(t)$, t ' ye göre doğrusaldır. Eğer $\log h(t)$ ' nin t ' ye göre grafiği çizildiğinde bir doğru elde ediliyor ise, başarısızlık zamanları için Gompertz dağılımının kullanılması uygundur.

Gompertz orantılı tehlike modeli altında, bir bireye ait tehlike fonksiyonu,

$$h(t|\mathbf{Z}) = \lambda \exp(\gamma t) \exp(\beta_1 Z_1 + \beta_2 Z_2 + \dots + \beta_p Z_p) = \lambda \exp(\gamma t) \exp(\boldsymbol{\beta}^T \mathbf{Z}) \quad (2.20)$$

biçimindedir. Bu eşitlikten, başarısızlık zamanının $\lambda \exp(\boldsymbol{\beta}^T \mathbf{Z})$ ve γ parametreleriyle Gompertz dağılımına sahip olduğu görülmektedir. Gompertz dağılımı da, orantılı tehlike özelliğine sahiptir [8].

2.3.3. Hızlandırılmış Başarısızlık Zamanı Modelleri

Yaşam analizinde, parametrik orantılı tehlike modelleri, kolay uygulanabilir olmasına karşın, orantılılık varsayımını sağlayan çok az olasılık dağılımı vardır. Bu özelliği sağlayan Weibull ve Gompertz gibi olasılık dağılımları ise, tehlike fonksiyonunu artan ya da azalan şekilde modellemeye olanak sağlamaktadır. Yaşam analizinde başarısızlık zamanları, farklı olasılık dağılımları kullanılarak daha geniş kapsamda modellenmek istendiğinde hızlandırılmış başarısızlık zamanı modelleri de kullanılabilir. Hızlandırılmış başarısızlık zamanı modellerinde, açıklayıcı değişkenlerinin başarısızlık zamanı üzerindeki doğrudan

etkisi ölçülür. Bu durum, ortalama başarısızlık zamanı üzerindeki açıklayıcı değişkenlerin etkisinin daha kolay yorumlanmasını sağlar. Parametrik orantılı tehlike modellerine benzer olarak hızlandırılmış başarısızlık zamanı modelleri de bir grup açıklayıcı değişken ile yaşam olasılıkları arasındaki ilişkiyi tanımlar.

$\mathbf{Z}^T = (Z_1, Z_2, \dots, Z_p)$ açıklayıcı değişken vektörü olsun. Hızlandırılmış başarısızlık zamanı modeline göre, i . bireye ait yaşam fonksiyonu,

$$S_i(t) = S_0 \left(\frac{t}{\exp(\eta_i)} \right) \quad (2.21)$$

biçiminde yazılır. Eşitlik (2.21)' de $S_0(t)$, temel yaşam fonksiyonu, $\boldsymbol{\alpha}^T = (\alpha_1, \alpha_2, \dots, \alpha_p)$ regresyon parametreleri, $Z(\cdot)$ açıklayıcı değişken vektörü ve $\eta_i = \alpha_1 Z_{1i} + \alpha_2 Z_{2i} + \dots + \alpha_p Z_{pi}$ olmak üzere $\exp(\eta_i)$ hızlandırma faktörüdür. Hızlandırma faktörü, $S(t)$ 'nin sabit bir değerine karşılık gelen başarısızlık zamanlarının bir oranıdır. Bu faktör, bilinmeyen temel başarısızlık zamanını genişletir ya da daraltır. [2, 10]

Hızlandırılmış başarısızlık zamanı modellerinde, açıklayıcı değişken etkisi zaman ölçeği üzerinde sabit ve çarpımsaldır. Yani, açıklayıcı değişkenler, yaşam zamanına hızlandırma faktörü adı verilen sabit bir faktör kadar etki eder.

Yaşam fonksiyonu ve tehlike fonksiyonu arasındaki ilişkiden yararlanarak, i . bireye ait tehlike fonksiyonu,

$$h_i(t) = \exp(-\eta_i) h_0 \left(\frac{t}{\exp(\eta_i)} \right) \quad (2.22)$$

eşitliğindeki gibi elde edilir. i . bireye ait başarısızlık zamanı T_i raslantı değişkenine karşılık gelen log-doğrusal model,

$$\log T_i = \mu + \alpha_1 Z_{1i} + \alpha_2 Z_{2i} + \dots + \alpha_p Z_{pi} + \sigma \varepsilon_i \quad (2.23)$$

eşitliğindeki gibidir. Burada μ , sabit terim, σ , ölçek parametresidir. ε_i raslantı değişkeni, $\log T_i$ değerlerinin modelin doğrusal kısmından sapmasını modellemek için kullanılır. ε_i 'nin ya da $\log T_i$ 'nin dağılımından ziyade T_i 'nin dağılımıyla adlandırılır [5, 8].

Eşitlik (2.21) ve (2.23)'deki modeller arasındaki ilişkiyi göstermek için, i . bireye ait başarısızlık zamanı T_i raslantı değişkeni ele alındığında yaşam fonksiyonu,

$$S_i(t) = P(T_i \geq t) = P\left(\exp\{\mu + \boldsymbol{\alpha}^T \mathbf{Z}_i + \alpha \varepsilon_i\} \geq t\right) \quad (2.24)$$

olarak yazılabilir. Eşitlik (2.24)'te $\boldsymbol{\alpha}^T \mathbf{Z}_i = \alpha_1 Z_{1i} + \alpha_2 Z_{2i} + \dots + \alpha_p Z_{pi}$ dir.

Eşitlik (2.24)'ten

$$S_i(t) = P\left(\frac{\exp\{\mu + \sigma \varepsilon_i\} \geq t}{\exp(\boldsymbol{\alpha}^T \mathbf{Z}_i)}\right) \quad (2.25)$$

olarak elde edilir.

Temel tehlike fonksiyonu $S_0(t)$, açıklayıcı değişken vektörü $Z = 0$ olan bir bireyin yaşam fonksiyonudur.

Buna göre,

$$S_0(t) = P(\exp\{\mu + \sigma \varepsilon_i\} \geq t) \quad (2.26)$$

eşitliği yazılabilir. Eşitlik (2.25) yeniden düzenlenirse,

$$S_i(t) = S_0\left(\frac{t}{\exp(\boldsymbol{\alpha}^T \mathbf{Z}_i)}\right) \quad (2.27)$$

biçiminde elde edilir. Eşitlik (2.27), i . bireye ait yaşam fonksiyonunun, hızlandırılmış başarısızlık zamanı modellerindeki genel gösterimidir. Modelin bu gösteriminde i . birey için hızlandırma faktörü $\exp(-\boldsymbol{\alpha}^T \mathbf{Z}_i)$ ve tehlike fonksiyonu ise,

$$h_i(t) = \exp(-\boldsymbol{\alpha}^T \mathbf{Z}_i) h_0\left(\frac{t}{\exp(\boldsymbol{\alpha}^T \mathbf{Z}_i)}\right) \quad (2.28)$$

olarak elde edilir. Eşitlik (2.22)' de $\eta_i = \boldsymbol{\alpha}^T \mathbf{Z}_i$ dir. Hızlandırılmış başarısızlık zamanı modelinin log-doğrusal ifadesi, i . bireye ait yaşam fonksiyonunu genel bir biçimde ifade etmekte de kullanılır. i . bireye ait yaşam fonksiyonu,

$$S_i(t) = P(T_i \geq t) = P(\log T_i \geq \log t) \quad (2.29)$$

biçiminde yazılabilir.

Eşitlik (2.23)' ten,

$$\begin{aligned}
S_i(t) &= P(\mu + \boldsymbol{\alpha}^T \mathbf{Z}_i + \sigma \varepsilon_i \geq \log t) \\
&= P\left(\varepsilon_i \geq \frac{\log t - \mu - \boldsymbol{\alpha}^T \mathbf{Z}_i}{\sigma}\right) \\
&= S_{\varepsilon_i}\left(\frac{\log t - \mu - \boldsymbol{\alpha}^T \mathbf{Z}_i}{\sigma}\right)
\end{aligned} \tag{2.30}$$

elde edilir. Bu sonuç, ε_i raslantı değişkeninin yaşam fonksiyonunun dağılımından yararlanarak T_i raslantı değişkeninin yaşam fonksiyonunun dağılımının nasıl bulunabileceğini gösterir. $H_i(t) = -\log S_i(t)$ eşitliğinden yararlanarak birikimli tehlike fonksiyonu,

$$H_i(t) = -\log S_{\varepsilon_i}\left(\frac{\log t - \mu - \boldsymbol{\alpha}^T \mathbf{Z}_i}{\sigma}\right) = H_{\varepsilon_i}\left(\frac{\log t - \mu - \boldsymbol{\alpha}^T \mathbf{Z}_i}{\sigma}\right) \tag{2.31}$$

biçiminde elde edilir. Eşitlik (2.31)' de $H_{\varepsilon_i} = -\log S_{\varepsilon_i}, \varepsilon_i$ ' nin birikimli tehlike fonksiyonudur. Eşitlik (2.31)' in t ' ye göre türevi alınırsa, i . birey için tehlike fonksiyonu,

$$h_i(t) = \frac{1}{\sigma t} h_{\varepsilon_i}\left(\frac{\log t - \mu - \boldsymbol{\alpha}^T \mathbf{Z}_i}{\sigma}\right) \tag{2.32}$$

biçiminde elde edilir. Eşitlik (2.32)' deki $h_{\varepsilon_i}(\varepsilon)$, ε_i 'nin tehlike fonksiyonudur [5, 8].

2.4. Zayıflık Modelleri

Zayıflık kavramı, yaşam analizi verisinin modellenmesinde gözlenemeyen heterojenliği modele dâhil eden uygun bir yoldur. En basit anlamda zayıflık, bir bireye ya da ilişkili bireylere (örneğin aynı aileye mensup bireylerde olduğu gibi) ait tehlike fonksiyonunu

değiřtiren gözlenemeyen rasgele bir faktördür. Zayıflık kavramı esas olarak Greenwood ve Yule'nin (1920) "kaza yapma oranı" çalışmasına dayanır. Clayton (1978) çalışmasında ailelerdeki kronik hastalık verisini zayıflık kavramını kullanarak modellemiř; ancak zayıflık terimi ilk kez Vaupel ve arkadaşları (1979) tarafından tek deęiřkenli yařam analizinde benzer özelliklere sahip bireyler arasında başarısızlık zamanlarındaki farklılıkları açıklamak için kullanılmıřtır [11, 12].

Zayıflık modelleri, yařam analizinde en sık kullanılan model olan Cox (1972) modeli olarak da bilinen orantılı tehlike modelinin bir uzantısıdır. Genellikle klinik arařtırmalar, yařam analizine konu olan kitlenin homojen olduęu varsayımına dayanır. Bunun anlamı çalışmaya seçilen tüm bireylerin prensip olarak aynı riske sahip olmasıdır (örneğin, ölüm riski, belirli bir hastalığın tekrarlanması riski vb.).

Birçok uygulamada çalışmaya konu olan kitle homojen olarak deęil heterojen olarak deęerlendirilmelidir (örneğin, farklı tehlikelere sahip bireylerin oluřturduęu kitle). Örneğin birçok çalışmada ilgilenilen hastalığa iliřkin deęiřkenlerin tümünün, bazı deęiřkenlerin öneminin hala bilinmemesinden ya da ekonomik nedenlerden dolayı ölçülmesi olanaksızdır. Zayıflık yaklařımı, ölçülemeyen deęiřkenlerin varlıęı nedeniyle oluřan heterojenlięi hesaba katmayı hedefleyen istatistiksel bir yaklařımdır. Zayıflık modeli, yařam analizinde rasgele etkinin, temel tehlike fonksiyonu üzerinde çarpımsal bir etkiye sahip olduęu bir modelleme türüdür [13].

Zayıflık modellerini daha iyi anlayabilmek için belirli bir kanser hastalığı için geliřtirilen tedavi yöntemlerinin karşılařtırıldıęı bir çalışma düşünölsün. Normalde aynı tedaviyi gören, aynı yařta ve cinsiyete sahip olan bireylerin aynı başarısızlık zamanına sahip olması beklenir ancak bu durum her zaman bu şekilde gerçekleřmeyebilir. Başarısızlık zamanlarındaki oluřan bu farklılıęı yaratan temel sebep, ölçülebilen açıklayıcı deęiřkenlere ek olarak başarısızlık zamanını etkileyen gözlenemeyen faktörler bulunmasıdır.

Bireylerin başarısızlık zamanları arasındaki farklılığa neden olan bu faktörler "Zayıflık", bireyler arasındaki heterojenliği açıklamak için ölçülemeyen rasgele etkiyi tehlike fonksiyonuna dâhil edilmesiyle oluşan modeller de "Zayıflık Modelleri" olarak adlandırılır.

Zayıflık modeli, paylaşılmamış zayıflık modeli ve paylaşılmış zayıflık modeli olarak ikiye ayrılır.

2.4.1. Paylaşılmamış Zayıflık Modeli

Zayıflık modeli, Kesim (2.4)' te bahsedilen gözlenemeyen faktörleri tehlike fonksiyonuna dâhil ederek başarısızlık zamanları arasında oluşan farklılığı açıklar. $h(t)$, t zamanındaki tehlike fonksiyonu ve ν gözlenemeyen değişken (zayıflık terimi) olmak üzere zayıflık modeli, n birey için,

$$h(t|v_i) = v_i h(t), \quad i = 1, 2, \dots, n \quad (2.33)$$

biçiminde gösterilir. Zayıflık terimi ν ' nin dağılımı $E(\nu^2) < \infty$ olacak şekilde alınır [11].

2.4.2. Paylaşılmış Zayıflık Modeli

Paylaşılmış zayıflık modeli bireylerin gözlenemeyen faktörleri (zayıflık) paylaştığı varsayımı ile hareket etmektedir. Literatürde paylaşılmış zayıflık terimi sıklıkla kullanılmaktadır. Örneğin, çok merkezli (hastane gibi) bir klinik araştırmanın yürütüldüğü bir çalışma düşünelim. Aynı hastanede tedavi gören hastalar, o hastanenin fiziki koşullardan yararlanmakta ve bu fiziki koşullar hastaların başarısızlık zamanlarını etkilemektedir. Fiziki koşulların hastaneden hastaneye değiştiği göz önüne alındığında, aynı hastanede tedavi gören hastalar, hastane etkisini gösteren ortak bir terimle ifade edilir. Bu terim ise paylaşılmış zayıflık terimi olarak adlandırılır. Aynı zayıflık terimini paylaşmak, bireyler arasındaki bağımlılığı ortaya çıkarmaktadır.

k tane grup ve n_i , i . gruptaki birey sayısını göstermek üzere, paylaşılmış zayıflık modelinin tehlike fonksiyonu,

$$h_{ij}(t|v_i) = v_i h_{ij}(t), \quad i=1,2,\dots,k ; j=1,2,\dots,n_i \quad (2.34)$$

biçiminde gösterilir. Zayıflık terimi için değişik dağılımlar kullanılabilir. Bunlardan en yaygın kullanılanları Gamma, log-normal ve ters-Gaussian dağılımlarıdır [13].

2.5. Olabilirlik Fonksiyonu ve Parametre Tahminleri

Parametrik modellerde parametre tahminleri, en çok olabilirlik yöntemi kullanılarak elde edilebilir. İlk olarak yaşam verisinde, durdurulmuş gözlemlerin olmadığını, yani n bireye ilişkin gerçek başarısızlık zamanlarının gözlemlendiği varsayalım. Başarısızlık zamanları, (T_1, T_2, \dots, T_n) raslantı değişkenlerinin aynı olasılık yoğunluk fonksiyonu $f(t)$ ' ye sahip olduğu varsayımı altında, gözlenen başarısızlık zamanları (t_1, t_2, \dots, t_n) ' e ilişkin olabilirlik fonksiyonu,

$$L = \prod_{i=1}^n f(t_i) \quad (2.35)$$

biçimindedir. Olabilirlik fonksiyonu, gözlem değerleri verildiğinde, bilinmeyen parametrelerin bir fonksiyonudur. Bu parametrelerin en çok olabilirlik tahminleri, olabilirlik fonksiyonunu maksimum yapan değerleridir.

Yaşam verisinde bir ya da daha fazla durdurulmuş gözlem olduğu durumu ele alalım. n bireyin r tanesinin (t_1, t_2, \dots, t_r) zamanlarında başarısızlığa uğradığı ve kalan $(n-r)$ bireye ilişkin $(t_1^*, t_2^*, \dots, t_{n-r}^*)$ başarısızlık zamanlarının sağdan durdurulmuş olduğu varsayalım. r tane başarısızlık zamanları (t_1, t_2, \dots, t_r) ' nin olabilirlik fonksiyonuna katkısı,

$$L = \prod_{j=1}^r f(t_j) \quad (2.36)$$

biçimindedir.

Eğer bir bireyin başarısızlık zamanı, t^* anında durdurulmuş ise, bu kişinin gerçek başarısızlık zamanı en azından t^* kadardır ve bu olayın olasılığı $P(T \geq t^*)$ yani $S(t^*)$ 'dir. Bu yüzden her durdurulmuş gözlem, olabilirlik fonksiyonuna bu terim kadar katkı sağlar. $(n-r)$ bireye ait durdurulmuş başarısızlık zamanlarının bilgisi de olabilirlik fonksiyonuna dahil edildiğinde toplam olabilirlik fonksiyonu,

$$L = \prod_{j=1}^r f(t_j) \prod_{l=1}^{n-r} S(t_l^*) \quad (2.37)$$

olarak bulunur.

Yaşam verisinin, n çift gözlemden oluştuğu düşünülebilir. Burada i . birey için (t_i, δ_i) çifti sırasıyla t_i , başarısızlık zamanı ve δ_i , birey t_i anında durdurulmuş ise 0, aksi halde 1 değerini alan gösterge değişkenidir. Buna göre, olabilirlik fonksiyonu,

$$L = \prod_{i=1}^n [f(t_i)]^{\delta_i} [S(t_i)]^{1-\delta_i} \quad (2.38)$$

biçimde yazılabilir. Eşitlik (2.38)' de verilen olabilirlik fonksiyonunun, maksimum yapılmasıyla, parametre tahminleri elde edilir [5].

Olabilirlik fonksiyonunun çıkarsaması yapılırken verideki durdurma mekanizması göz önünde bulundurulmalıdır. Eşitlik (2.38) ile verilen olabilirlik fonksiyonu bilgi içermeyen

durdurma varsayımı altında elde edilmiştir. Bunu daha iyi açıklayabilmek için şu tanımlamalar yapılsın.

C_i ve T_i^* sırasıyla i . bireye ait durdurma zamanını ve başarısızlık zamanını gösterebilir. i . bireye ait durdurma zamanı bir rastlantı değişkenidir. C_i , rastlantı değişkenine ait olasılık yoğunluk fonksiyonu ve yaşam fonksiyonu sırasıyla $g_i(t)$ ve $G_i(t)$ olarak gösterilsin.

Kesim (2.3.1)' de ifade edildiği gibi $T = T_i^* \wedge C_i$ ve $\delta_i = I(T_i^* \leq C_i)$ ' dir. Bu bilgilerden yararlanarak,

$$\begin{aligned} P(T_i \in (t, t + \Delta t), \delta_i = 1 | \mathbf{Z}_i, \boldsymbol{\beta}) &= P(T_i^* \in (t, t + \Delta t), C_i > t | \mathbf{Z}_i, \boldsymbol{\beta}) \\ &= G_i(t) f(t | \mathbf{Z}_i, \boldsymbol{\beta}) dt \end{aligned}$$

ve

$$\begin{aligned} P(T_i \in (t, t + \Delta t), \delta_i = 0 | \mathbf{Z}_i, \boldsymbol{\beta}) &= P(C_i \in (t, t + \Delta t), T_i^* > t | \mathbf{Z}_i, \boldsymbol{\beta}) \\ &= g_i(t) S(t | \mathbf{Z}_i, \boldsymbol{\beta}) dt \end{aligned}$$

yazılabilir [14, 15]. Açıklayıcı değişken $\mathbf{Z} = (Z_1, Z_2, \dots, Z_p)$ değerleri verildiğinde (T_i, δ_i) $i = 1, 2, \dots, n$ gözlem çiftleri bağımsızdır. Eğer durdurma, bilgi içermiyorsa ki bu parametre $\boldsymbol{\beta}$ ' ya ilişkin bilgi içermemesi anlamına gelir, olabilirlik fonksiyonu,

$$L \propto \prod_{i=1}^n [f(t_i | \boldsymbol{\beta}, \mathbf{Z}_i)]^{\delta_i} [S(t_i | \boldsymbol{\beta}, \mathbf{Z}_i)]^{1-\delta_i}. \quad (2.39)$$

biçiminde elde edilir.

Eşitlik (2.39)' daki olabilirlik fonksiyonunun, Eşitlik (2.38)' de elde edilen olabilirlik fonksiyonuyla aynı olduğu görülebilir.

Bilgi içeren durdurma olduğu durumda olabilirlik fonksiyonunun nasıl elde edileceğini Bölüm 5'te anlatılmıştır.

3. YİNELEMELİ OLAY ANALİZİ

Yaşam analizinde daha fazla bilgi verebilecek çoklu olay verileri son yıllarda üzerinde çalışılan bir konudur. Çoklu olay verisinde, olaylar aynı bireyler üzerinde gözlemlendiğinden, olay zamanı çoğu kez sıralı ve ilişkilidir. Çoklu olay verisi iki kategoriye ayrılır. Bunlar yinelemeli olaylar ve yarışan risklerdir.

Başarısızlığın nedeni, birden fazla etkenden olabilir ve başarısızlığı oluşturan bu etkenlerden sadece birisidir. Bu tip yaşam analizi verisine yarışan riskler verisi denir.

Bireyler ya da birimler, aynı olayı birden fazla kez deneyimleyebilirler. Bu durum için elde edilen verilere, yinelemeli başarısızlık zamanı verileri denir [16].

Yinelemeli olaylara ilişkin verilerin istatistiksel modellemesini anlamak için veri yapısının karakteristiklerini bilmek önemlidir. Yinelemeli olay verisinin üç önemli özelliği bulunmaktadır.

Yinelemeli olay verilerinde bireyler bağımsız olarak örneklenmiştir, ancak aynı bireylerden elde edilen yineleme zamanları ilişkilidir. Bu ilişki, parametre tahminlerine ilişkin varyansın olduğundan daha düşük tahmin edilmesine neden olur.

Yinelemeli olayların sayısı, yinelemeli olay süreci hakkında bilgi vericidir. Örneğin, bir birey için belli bir zaman periyodu içinde az sayıda yinelemeli olayın gözlenmesi, olayın oluş hızının yavaş olduğunu, olaylar arasında geçen zamanın uzun olduğunu belirtir.

Yinelemeli olaylarda, birey ya da birim aynı olayı değişik zamanlarda tekrar yaşayabilir. Bu nedenle, yinelemeli olay analizinde ilgilenilen olay hiçbir zaman ölüm olamaz.

Yinelemeli olayların stokastik sıralı yapısı istatistiksel modellerin gelişiminde çeşitli güçlüklerle neden olur. Örneğin, bir birey için son tekrar zamanı, sağdan durdurmaya maruz kalabilir. Yinelemeli olaylara örnek olarak aşağıdaki durumlar verilebilir:

Örnek 1 (Evre 1 Mesane Kanseri Klinik Çalışması):

Bu veri, Gaziler Birliği Ürolojik Araştırma Grubu (Veterans Administration Cooperative Urological Research Group) adına Byar ve Blackard (1976) tarafından toplanmıştır. Çalışmanın amacı, plaseboya karşı nükseden mesane tümörlerinde piridoksin (Vitamin B₆) ve tiyotepa etkisini belirlemektir. Bu çalışmada Kasım 1971 ve Ağustos 1976 yıllarında 10 Gazi hastanesinden 121 hasta takip edilmiştir. Tümörü olan, nüks etmiş ya da yeni tümör gözlenmiş tüm hastalar potansiyel olarak çalışmaya dâhil edilmiştir. Ayrıca papilomatoz (dışı pürtüklü ve düzensiz, karnıbahar görünümünde) tümör durumunda bulunan hastalar eğer tümörleri transüretral rezeksiyon (TUR) ile rezeke edilebilir ise çalışmaya kabul edilmişlerdir. Çalışmaya, mesane kanseri için daha önce radyoterapi ya da kemoterapi almış olan, Evre 1' den daha ileri düzeyde mesane kanserine sahip olan, mesanede kanser olmayan tümörü bulunan ve TUR yöntemiyle rezeke edilemeyecek durumdaki mesane papilomatoz durumunda olanlar dahil edilmemişlerdir. Hastalar rasgele olarak plasebo (1 tablet/gün), oral piridoksin (1.25mg tablet/gün) ve tiyotepa içeren daha karmaşık bir tedavi rejimi olmak üzere üç tedavi grubundan birine dâhil edilmişlerdir [17].

Literatürde bu veri kullanılarak yapılan çalışmalarda araştırmacılar, piridoksin ve tiyotepa tedavisini alanların, plasebo alanlara göre, önemli derecede düşük oranlarda nüks etme yaşadığını bildirmişler ve 118 hastadan 30' unun öldüğünü rapor etmişlerdir. Bu çalışmalarda araştırmacılar, ölümü bilgi içermeyen durdurma olayı olarak kabul etmişler ve modellerini bilgi içermeyen durdurma varsayımı altında oluşturmuşlardır [16].

Örnek 2 (Büyük Bir Yazılım Sisteminde Test ve Hata Ayıklama Çalışması):

Dalal ve McIntosh (1994), kaynak kodunda yaklaşık 7 milyon çözümlenmemiş satır (Noncommentary Source Lines: NCLS) bulunan büyük bir yazılım sistemi için bir test ve hata ayıklama süreci tanımlamışlardır. Farklı kişiler tarafından yürütülen test süreci sırasında, testi yapan kişiler her gün test için harcadıkları saati ve buldukları hata sayısını kaydetmişlerdir. Sorunlar çözümlenirken ya da başka sebeplerle yazılım sistemine zaman içerisinde fazladan kod (NCLS) eklemesi yapılmıştır. Test 160 takvim günü kadar sürmüş, 870 hata bulunmuş ve 342,000 yeni NCLS eklenmiştir. Testi yapan her bir kişinin bulduğu hatalar yinelemeli olay olarak kabul edilmiş ve bu hataları bulana kadar geçen süre gün olarak kaydedilmiştir [18].

Yukarıdaki örneklerin her birinde ilgilenilen olay farklıdır, fakat hepsinin ortak noktası aynı olayın birden çok tekrarlanmış olmasıdır. Yinelemeli olay analizinde amaç, açıklayıcı değişkenler ile ilgilenilen olayın tekrarlama hızı arasındaki ilişkiyi değerlendirmektir [2].

Yinelemeli olayların analizi için birçok model kullanılmıştır. Modelleme yöntemlerinden en önemli üç tanesi: yinelemeli olaylar arasında geçen zamanı modellemeye dayanan yöntem [19], her bir yinelemeye dek geçen zamanı modellemeye dayanan marjinal tehlike yöntemi [20], ve yinelemeli olay verilerini modellemeye dayanan yoğunluk/oran modeli [21, 22] yöntemleridir.

Yinemeli olayları modellenmesi ve bu modellerin çözümüne ilişkin çok sayıda çalışma vardır: Örneğin, Andersen ve Gill (1982), yinelemeli olay sürecini bağımsız artmalı sayma süreci kabul ederek, açıklayıcı değişkenler yardımıyla Cox-tipi modeli kullanmışlardır [21]. Pepe ve Cai (1993), Poisson tipi süreçlerde bağımsız artma varsayımını gevşetip koşullu oran fonksiyonunu modelleyerek yeni bir yaklaşım önermişlerdir [22]. Lin, Wei, Yang ve Ying (2000) yinelemeli olaylarda ortalama ve oran fonksiyonlarını Poisson varsayımı olmaksızın yarı parametrik olarak modellemişler ve modelin yeterliliğini test etmek için sayısal yöntemler önermişlerdir [23].

Yinelemeli olay analizinde, hızlandırılmış başarısızlık zamanı modelleri de çalışılmıştır. Lin ve Wei (1992), her yineleme zamanını, hızlandırılmış başarısızlık zamanı modelleri kullanarak modellemiştir [24]. Lin, Wei ve Ying (1998), sayma sürecini kullanarak ortalama fonksiyonlarını hızlandırılmış başarısızlık zamanı modelleri ile modellemiştir [25]. Pan (2001), hızlandırılmış başarısızlık zamanı modellerine zayıflık terimini ekleyerek modellemiştir [26].

Bu tez çalışmasında ilgilenilen olaylar, yinelemeli olaylardır. Bu nedenle bu kesimde yinelemeli olayların modellenmesinde araç olarak kullanılan sayma süreci hakkında bilgi verilmiş, yinelemeli olay analizinde kullanılan fonksiyonlar açıklanmış ve literatürde yer alan önemli modeller hakkında detaylı bilgi verilmiştir.

3.1. Sayma Süreci Notasyonu

(T_i, C_i) ($i=1,2,\dots,n$) bağımsız ve aynı dağılımlı rastlantı değişkenleri olmak üzere i . bireye ilişkin sırasıyla başarısızlık ve durdurma zamanlarını gösterir. Veride sağdan durdurulmuş gözlemler yer aldığı için $T_i = T_i^* \wedge C_i$ zamanı gözlenir. $\delta_i = I(T_i^* \leq C_i)$, i . birey başarısızlık deneyimlemişse 1, değilse 0 değerini alan gösterge değişkenidir.

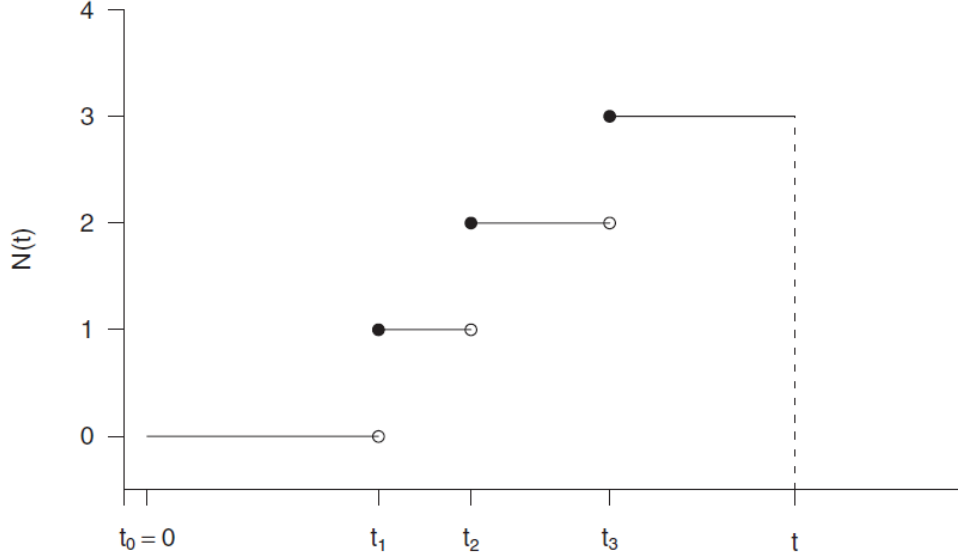
$N(t)$, sayma süreci, $N(0)=0$ ve $N(t) < \infty$ değerini alan stokastik bir süreçtir. $N(t)$, iki sıçrama noktası arasında sabit değer alan, sıçrama noktalarında +1 büyüklüğünde sıçrama yapan sağdan sürekli bir fonksiyondur. $\{N_i(t), t \geq 0\}$ sayma süreci,

$$N_i(t) = I(T_i^* \leq t, \delta_i = 1) \quad (3.1)$$

eşitliğindeki gibi gösterilir. $N_i(t)$, i . birey başarısızlık deneyimleyene dek 0 değerini alır, başarısızlık anında bir kez sıçrama yaparak 1 değerini alır ve daha sonra sabit kalır.

$N(t) = \sum_{i=1}^n N_i(t)$, t zamanına dek çalışmadaki toplam başarısızlık sayısını göstermektedir.

$\{N_i(t), t \geq 0\}$ sayma sürecinin daha iyi anlaşılması için Şekil (3.1)'de grafiksel gösterimi verilmiştir.

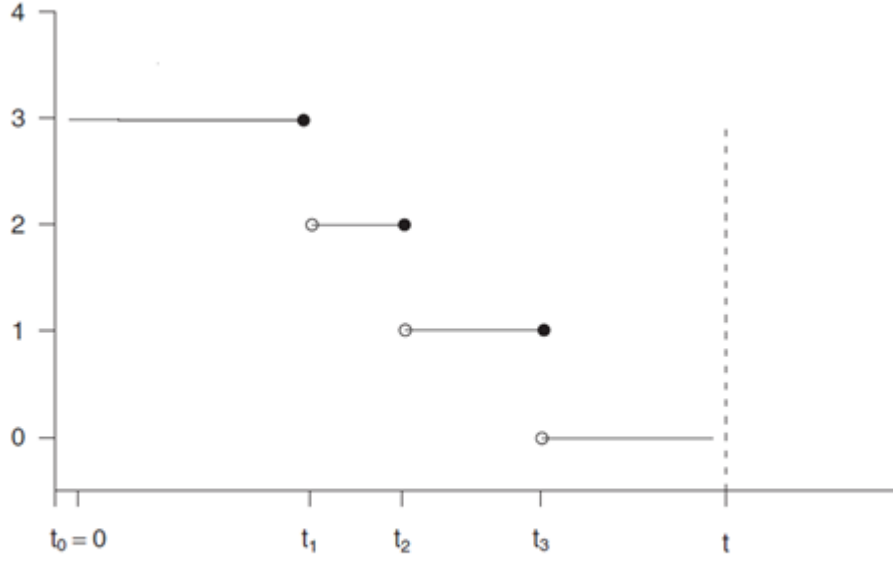


Şekil 3.1. Sayma sürecinin grafiksel gösterimi

$Y(t)$ sayma süreci, $Y(0)=1$ değerini alan stokastik bir süreçtir. $Y(t)$, iki sıçrama noktası arasında sabit değer alan, sıçrama noktalarında 1 büyüklüğünde azalan, soldan sürekli bir fonksiyondur. $\{Y_i(t), t > 0\}$ sayma süreci,

$$Y_i(t) = I(T_i^* \geq t) \quad (3.2)$$

eşitliğindeki gibi gösterilir ve i . bireyin t anında hayatta olup olmadığını ifade eden risk süreci olarak adlandırılır. $\{Y_i(t), t > 0\}$ risk sürecinin daha iyi anlaşılması için Şekil (3.2)'de grafiksel gösterimi verilmiştir.



Şekil 3.2. Risk sürecinin grafiksel gösterimi

Yaşam analizine önemli bir yenilik getiren sayma süreci yaklaşımı ilk olarak Aalen (1975) tarafından tanıtılmıştır [27].

Aalen (1975) tarafından tek değişkenli yaşam analizi modelleri için önerilen sayma süreci yaklaşımı yinelemeli olay analizi gibi çok değişkenli yaşam analizi modelleri için de aşağıdaki şekilde kullanılabilir.

$(T_{i,1}, T_{i,2}, \dots, T_{i,k})$ ve C_i , sırasıyla i . ($i = 1, 2, \dots, n$) bireye ait yinelemeli olay zamanlarını ve durdurma zamanını gösterebilir. $(T_{i,1}, T_{i,2}, \dots, T_{i,k})$ olay zamanları çalışmanın başlangıcından k . yineleme gerçekleşene dek geçen zamanı belirtmektedir. i . bireye ait $[0, t]$ zaman aralığındaki yineleme sayısını gösteren sayma süreci,

$$N_i(t) = \int_0^t dN_i(s) \quad (3.3)$$

eşitliğindeki gibi ifade edilir. Eşitlik (3.3)' te $dN_i(s)$, $[s, s + \Delta s]$ gibi çok küçük bir zaman aralığındaki ortaya çıkan olay sayısını göstermektedir. Sayma sürecinin özellikleri:

- (i) $N_i(t) \geq 0$,
- (ii) $N_i(t)$, kesikli değer alan bir süreç,
- (iii) $s < t$ zaman noktası için $N_i(s) \leq N_i(t)$,
- (iv) (s, t) zaman aralığında ortaya çıkan olay sayısı $N_i(t) - N_i(s)$

biçimindedir.

3.2. Yinelemeli Olay Analizinde Kullanılan Temel Fonksiyonlar

Bölüm 2' de verilen tehlike fonksiyonu tanımı kullanılarak yinelemeli olay analizinde kullanılan yoğunluk, oran ve ortalama fonksiyonları sayma süreci notasyonu yardımıyla aşağıdaki biçimde yazılabilir:

Geçmiş bilgisi $\mathcal{H}(t^-)$ verildiğinde, t anında anlık olay olması olasılığını veren yoğunluk fonksiyonu,

$$\lambda(t | \mathcal{H}(t^-)) = \lim_{\Delta t \rightarrow 0} \frac{P(\Delta N(t) = 1 | \mathcal{H}(t^-))}{\Delta t} \quad (3.4)$$

biçimindedir. Eşitlik (3.4)' te $\mathcal{H}(t^-) = \{N(s), Z(s), 0 \leq s < t\}$ t anına kadar ortaya çıkan olay ve açıklayıcı değişken bilgisini, yani geçmiş bilgisini ifade etmektedir.

Sayma süreci $\{N_i(t), t > 0\}$ için ortalama fonksiyonu,

$$\mu(t) = E[N(t)] \quad (3.5)$$

eşitliğindeki gibi yazılır ve t anına dek ortaya çıkması beklenen olay sayısı olarak yorumlanır. Ortalama fonksiyonundan yararlanarak oran fonksiyonu $\rho(t) = \frac{d\mu(t)}{dt}$, birim zamanda ortaya çıkan beklenen olay sayısı olarak ifade edilir ve

$$\rho(t) = \lim_{\Delta t \rightarrow 0} \frac{E[N(t, t + \Delta t)]}{\Delta t} \quad (3.6)$$

biçiminde tanımlanır [18].

3.3. Yinelemeli Olay Analizinde Kullanılan Temel Modeller

Cox tarafından verilen orantılı tehlike modeli, yinelemeli olay verileri için temel olarak Prentice, William ve Peterson (1981), Wei, Lin ve Weissfeld (1989), Andersen ve Gill (1982) tarafından genişletilmiştir. Bu üç temel modelin farklı bir uyarlaması Pepe ve Cai (1993) tarafından önerilmiştir. Yinelemeli olay verileri hızlandırılmış başarısızlık zamanı modelleri ve paylaşılmış zayıflık modelleri kullanılarak da modellenenabilir. Bu modeller ilerleyen kesimlerde ayrıntılı bir biçimde açıklanmıştır.

3.3.1. Prentice, William ve Peterson Yaklaşımı

$(T_{i,1}, T_{i,2}, \dots, T_{i,K})$, i . bireye ait sıralı başarısızlık zamanlarını ve C_i , i . bireye ait durdurma zamanını gösterebilir. $\mathbf{Z}(t) = (Z_1(t), Z_2(t), \dots, Z_p(t))$, $t \geq 0$, t anındaki zamana bağlı açıklayıcı değişken vektörünü ve $N(t)$, $[0, t]$ zaman aralığındaki sayma sürecini gösterebilir.

Buna göre, açıklayıcı değişken ve sayma süreci bilgisi verildiğinde t anındaki anlık başarısızlık olması olasılığını veren yoğunluk fonksiyonu,

$$\lambda(t|N(t), \mathbf{Z}(t)) = \lim_{\Delta t \rightarrow 0} \frac{P\{T_{n(t)+1} \in [t, t + \Delta t | N(t), \mathbf{Z}(t)]\}}{\Delta t} \quad (3.7)$$

biçimindedir. Prentice ve arkadaşları, Cox orantılı tehlike modelinden yararlanarak Eşitlik (3.7)' deki yoğunluk fonksiyonunu aşağıdaki biçimlerde modellemişlerdir:

$$\lambda_{ik}(t|N_i(t), \mathbf{Z}_i(t)) = \lambda_{0k}(t) \exp(\boldsymbol{\beta}_k^T \mathbf{Z}_{ik}(t)) \quad (3.8)$$

$$\lambda_{ik}(t|N_i(t), \mathbf{Z}_i(t)) = \lambda_{0k}(t - T_{i,k-1}) \exp(\boldsymbol{\beta}_k^T \mathbf{Z}_{ik}(t)). \quad (3.9)$$

Eşitlik (3.8) ve (3.9)' da $\lambda_{0k}(\cdot)$ bilinmeyen temel tehlike fonksiyonunu, k tabaka değişkenini β_k , k . tabakaya özgü regresyon parametresini göstermektedir. Tabaka değişkeni k , her bir yineleme tarafından belirlenmektedir. Eşitlik (3.8)' de verilen tehlike fonksiyonu, çalışmanın başından t anına dek geçen zamanı modellerken, Eşitlik (3.9)' da verilen tehlike fonksiyonu ardışık iki başarısızlık arasında geçen zamanı modellemektedir. Ayrıca $Y_{ik}(t) = I(T_{i,k-1} \leq t \leq T_{i,k})$ ve $Y_{ik}(t) = I(T_{i,k} \geq T_{i,k-1} + t)$ sırasıyla Eşitlik (3.8) ve (3.9) için risk süreçlerini ifade etmektedir [19].

Prentice ve arkadaşları tarafından önerilen bu modelde, $(k+1)$. yineleme için risk kümesini oluşturan bireyler, k . yinelemeyi deneyimleyen bireylerden oluşur kısıtı vardır. Bu yüzden çalışmanın başındaki risk kümesi korunmamaktadır.

Eşitlik (3.8)' in özel bir durumu $Z(t) = Z$ (Örneğin, 0-1 tedavi değişkeni) ve $\beta_k = \beta$ alınsın. Bu durumda β , birinci yineleme için log tehlike oranını; birinci yinelemeyi

deneyimleyenler arasında ikinci yineleme için log tehlike oranını ve bu şekilde devam edildiğinde k . yinelemeyi deneyimleyenler arasında $(k+1)$. yineleme için log tehlike oranını ifade eder. Prentice ve arkadaşları tarafından önerilen modelin birinci yineleme için elde edilen sonuçları, Cox orantılı tehlike modeli kullanılarak elde edilen sonuçlarla aynıdır.

Bu modelin getirdiği kısıtın bir sonucu olarak, çalışmanın başında tedavi ve kontrol grubundaki risk kümesi çalışmanın sonunda aynı kalmayacaktır. Eşitlik (3.8)' de verilen Prentice ve arkadaşlarının modeli için skor fonksiyonu,

$$U_{TT}(\boldsymbol{\beta}) = \sum_{i=1}^n \int_0^{\tau} [\mathbf{Z}_{ik}(t) - \bar{\mathbf{Z}}(\boldsymbol{\beta}, t)] dN_i(t), \quad i = 1, 2, \dots, n; \quad k = 1, 2, \dots, K \quad (3.10)$$

biçimindedir. Eşitlik (3.10)' da,

$$\bar{\mathbf{Z}}(\boldsymbol{\beta}, t) = \frac{S_k^1(\boldsymbol{\beta}_k, t)}{S_k^0(\boldsymbol{\beta}_k, t)}$$

ve

$$S_k^{(j)}(\boldsymbol{\beta}_k, t) = \frac{1}{n} \sum_{i=1}^n Y_{ik}(t) \mathbf{Z}_{ik}(t)^{\otimes j} \exp(\boldsymbol{\beta}_k^T \mathbf{Z}_{ik}(t))$$

biçimindedir. Δ_{ik} , i . bireyin k . yineleme yaşama durumunu ifade eden gösterge değişkeni ve $N_i(t) = I(T_{i,k} \leq t, \Delta_{ik} = 1)$ i . bireyin t zamanına dek yaşadığı yineleme sayısını gösteren sayma sürecidir. \mathbf{Z} açıklayıcı değişken vektörü için $\mathbf{Z}^{\otimes 0} = 1$, $\mathbf{Z}^{\otimes 1} = \mathbf{Z}$ ve $\mathbf{Z}^{\otimes 2} = \mathbf{Z}\mathbf{Z}^T$ ifade etmektedir.

Prentice ve arkadaşlarının Eşitlik (3.9)' da önerdikleri model için skor fonksiyonu,

$$U_{GT} = \sum_{i=1}^n \int_0^{\tau} [\mathbf{Z}_{ik}(t+T_{i,k-1}) - \bar{\mathbf{Z}}(\boldsymbol{\beta}, t)] d\tilde{N}_i(t), \quad i=1,2,\dots,n; \quad k=1,2,\dots,K. \quad (3.11)$$

biçimindedir. Eşitlik (3.11)' de,

$$\bar{\mathbf{Z}}(\boldsymbol{\beta}, t) = \frac{R_k^1(\boldsymbol{\beta}_k, t)}{R_k^0(\boldsymbol{\beta}_k, t)}$$

ve

$$R_k^{(j)}(\boldsymbol{\beta}_k, t) = \frac{1}{n} \sum_{i=1}^n Y_{ik}(t) \mathbf{Z}_{ik}(T_{i,k-1} + t)^{\otimes j} \exp(\boldsymbol{\beta}_k^T \mathbf{Z}_{ik}(T_{i,k-1} + t))$$

biçimindedir. $\tilde{N}_i(t) = I(G_{i,k} \leq t, \Delta_{ik} = 1)$, i . bireyin t zamanına dek yaşadığı yineleme sayısını gösteren sayma sürecidir. Burada $G_{i,k}$, i . bireyin $(k-1)$. ve k . yinelemeleri arasında geçen zamanı göstermektedir.

Eşitlik (3.10) ve (3.11)' deki skor fonksiyonlarını sıfır yapan $\boldsymbol{\beta}_k$ değeri, parametre tahminini vermektedir [16, 19, 28].

3.3.2. Wei, Lin ve Weissfeld Yaklaşımı

Yinelemeli olay verisini modellemek için kullanılan diğer bir yöntem Wei ve arkadaşları tarafından önerilen “Marjinal Tehlike Modeli” dir [20]. Wei ve arkadaşlarının yaklaşımları, k . yineleme gerçekleşene dek geçen zamanı modellemektedir.

$(T_{i,1}, T_{i,2}, \dots, T_{i,K})$, i . bireye ait sıralı başarısızlık zamanlarını, C_i , i . bireye ait durdurma zamanını, $\mathbf{Z}(t) = (Z_1(t), Z_2(t), \dots, Z_p(t))$, $t \geq 0$, t anındaki zamana bağlı açıklayıcı değişken

vektörünü ve $N(t)$, $[0,t]$ zaman aralığındaki sayma sürecini göstermek üzere Wei ve arkadaşlarının, k . ($k = 1, 2, \dots, K$) yineleme için önerdikleri model,

$$\lambda_{ik}(t|\mathbf{Z}) = \lambda_{0k}(t) \exp(\boldsymbol{\beta}_k^T \mathbf{Z}_{ik}(t)) \quad (3.12)$$

biçimindedir [20]. Eşitlik (3.12)'deki $\lambda_{ik}(t|\mathbf{Z}(t))$, i . birey için k . yineleme gerçekleşene dek geçen zamanın tehlike fonksiyonunu ve $\lambda_{0k}(t)$ bilinmeyen temel tehlike fonksiyonunu, k tabaka değişkenini ifade etmek üzere $\boldsymbol{\beta}_k$, k . tabakaya özgü parametre vektörünü göstermektedir. Bu modelde yinelemeler arası bağımlılık yapısı modellenmemiştir.

Bu modelin çok sayıda üstünlüğü vardır. Birincisi, bu model sadece marjinal tehlikeleri modellediğinden, yinelenen olaylar arasındaki bağımlılığa karşı sağlamdır. İkincisi ise, model k . yinelemeye dek geçen zamanı modellediğinden, çalışmanın başındaki risk kümesi korunur.

Eşitlik (3.12)'de verilen model altında gözlenen başarısızlık zamanlarına ilişkin olabirlilik fonksiyonunun logaritmasınının $\boldsymbol{\beta}_k$ 'ya göre birinci türevi alınırsa, k . yineleme için aşağıdaki skor fonksiyonu elde edilir:

$$U_k(\boldsymbol{\beta}) = \sum_{i=1}^n \int_0^{\tau} [\mathbf{Z}_{ik}(t) - \bar{\mathbf{Z}}(\boldsymbol{\beta}, t)] dN_{ik}(t), \quad i = 1, 2, \dots, n; \quad k = 1, 2, \dots, K. \quad (3.13)$$

Eşitlik (3.13)'te,

$$\bar{\mathbf{Z}}(\boldsymbol{\beta}, t) = \frac{S_k^{(1)}(\boldsymbol{\beta}_k, t)}{S_k^{(0)}(\boldsymbol{\beta}_k, t)}$$

ve

$$S_k^{(j)}(\boldsymbol{\beta}_k, t) = \frac{1}{n} \sum_{i=1}^n Y_{ik}(t) \mathbf{Z}_{ik}(t)^{\otimes j} \exp(\boldsymbol{\beta}_k^T \mathbf{Z}_{ik}(t))$$

biçimindedir. $Y_{ik}(t) = I(T_{i,k} \geq t)$, i . bireyin k . yineleme için risk kümesinde olup olmadığını gösteren gösterge değişkeni ve $\Delta_{ik} = I(T_{i,k} \leq C_i)$ ise i . bireyin k . yineleme zamanında durdurulmuş olup olmadığını gösteren gösterge değişkenidir. Eşitlik (3.13)' ü sıfır yapan değerler parametre tahminlerini vermektedir [16, 20, 28].

3.3.3. Andersen ve Gill Yaklaşımı

Cox tarafından önerilen tahmin yönteminin yinelemeli olay analizi için uyarlaması Andersen ve Gill (1982) tarafından yapılmıştır [21].

Andersen ve Gill, sayma süreci gösterimlerini kullanarak Cox orantılı tehlike modelindeki kısmi olabilirlik fonksiyonun birinci türevini yeniden formüle etmiş ve bundan yararlanarak yinelemeli olay verilerini modellemişlerdir.

$\lambda_0(t)$, temel tehlike fonksiyonunu, $\boldsymbol{\beta}$, $p \times 1$ boyutlu parametre vektörünü ve $\mathbf{Z}(t) = (Z_1(t), Z_2(t), \dots, Z_p(t))$ açıklayıcı değişken vektörünü göstermek üzere Eşitlik (2.8)' deki Cox orantılı tehlike modeli,

$$\lambda(t|\mathbf{Z}) = \lambda_0(t) \exp(\boldsymbol{\beta}^T \mathbf{Z}(t)) \quad (3.14)$$

biçimindedir. Eşitlik (3.14)' te verilen yoğunluk fonksiyonuna ilişkin kısmi olabilirlik fonksiyonu Eşitlik (3.15)' te verilmiştir.

$$\log L(\boldsymbol{\beta}) = \sum_{i=1}^n \left[\boldsymbol{\beta}^T \mathbf{Z}_i(t) - \log \left(\sum_{j=1}^n Y_j(t) \exp(\boldsymbol{\beta}^T \mathbf{Z}_j(t)) \right) \right] \Delta_i \quad (3.15)$$

Eşitlik (3.15)' te $Y_j(t) = I(C_j \geq t)$, i . bireyin t anında risk kümesinde olup olmadığını gösteren gösterge değişkeni ve $\Delta_i = I(T_{i,k} \leq C_i)$ i . bireyin k . yineleme zamanında durdurulmuş olup olmadığını gösteren gösterge değişkenidir.

Eşitlik (3.15)' te verilen kısmi olabilirlik fonksiyonun birinci türevi, sayma süreci gösterimine göre yeniden ifade edildiğinde,

$$U^{AG}(\boldsymbol{\beta}) = \sum_{i=1}^n \int_0^{\infty} \left[\mathbf{Z}_i(t) - \frac{\sum_{j=1}^n Y_j(t) \mathbf{Z}_j(t) \exp(\boldsymbol{\beta}^T \mathbf{Z}_j(t))}{\sum_{j=1}^n Y_j(t) \exp(\boldsymbol{\beta}^T \mathbf{Z}_j(t))} \right] dN_i(t) \quad (3.16)$$

eşitliği elde edilir. Eşitlik (3.16)' da $N_i(t)$, i . bireyin t zamanına dek yaşadığı yineleme sayısını gösteren sayma sürecidir. Eşitlik (3.16)' yı sıfır yapan değerler $\boldsymbol{\beta}$ parametre vektörünün tahminini vermektedir [21].

3.3.4. Pepe ve Cai Tarafından Önerilen Oran Modelleri

Pepe ve Cai, (1993)' te, yinelemeli olay verisini modellemek için bilgi içermeyen durdurma varsayımı altında, zamana bağlı açıklayıcı değişkenlerin yer aldığı oran modellerini önermişlerdir. İlk yineleme ve ilk yinelemeyi deneyimleyenler içinde diğer yinelemeleri iki ayrı oran fonksiyonu ile modellemişlerdir: Bu oranlar,

$$r^F(t, \alpha) = \lim_{\Delta \rightarrow \infty} \frac{1}{\Delta} P \left\{ (t, t + \Delta) \text{ zaman aralığında yineleme yaşanması} \mid \text{hiç başarısızlık} \right. \\ \left. \text{yaşamamış ve risk altındaki bireyler} \right\}$$

$$r^R(t; \beta) = \lim_{\Delta \rightarrow \infty} \frac{1}{\Delta} P \{ (t, t + \Delta) \text{ zaman aralığında yineleme yaşanması} \mid \text{bir kez yineleme} \\ \text{yaşamış ve risk altındaki bireyler} \}$$

biçiminde tanımlanmış; hiç başarısızlık yaşamamış ve risk altındaki bireyler verildiğinde, birinci yinelemeye ilişkin olabilirlik fonksiyonu,

$$L(\alpha) = \prod_{t_i \in D} r^F(t_i; \alpha) \exp \left\{ - \int_0^{t_i} r^F(u; \alpha) du \right\} \prod_{t_j \in O} \exp \left\{ - \int_0^{t_j} r^F(u; \alpha) du \right\}$$

biçiminde elde edilmiştir. Burada D , ilk yineleme zamanlarının yer aldığı ve O ise, durdurulma zamanının ya da takip süresinin yer aldığı kümeyi gösterir. Benzer şekilde $L(\beta)$ elde edilebilir ve buradan parametrelerin en çok olabilirlik tahminleri elde edilmiştir [22].

3.3.5. Hızlandırılmış Başarısızlık Zamanı Modelleri

Bilgi içeren durdurma olmadığında yinelemeli olay verilerinin başka bir modellemesi Lin, Wei ve Ying (1998) tarafından önerilen hızlandırılmış başarısızlık zamanı modelleridir [25].

$T_{i,k}$, ($i=1,2,\dots,n$ ve $k=1,2,\dots,K$) i . bireyin k . yineleme zamanını ve C_i , i . bireyin durdurma zamanını gösterebilir. $N_i(t)$, i . bireyin $t \wedge C_i$ zamanına dek (takip süresi boyunca) deneyimlediği olay sayısını göstermek üzere aşağıdaki eşitlikteki gibi ifade edilir:

$$N_i(t) = \sum_{k=1}^{\infty} I(T_{ik} \leq t \wedge C_i).$$

$\mu(t | \mathbf{Z}) = E(N(t) | \mathbf{Z})$, yinelemeli olay sürecinin ortalama fonksiyonu olmak üzere bu fonksiyon,

$$E\{N_i(t) | \mathbf{Z}_i\} = \mu_0(e^{\boldsymbol{\beta}^T \mathbf{Z}_i t}) \quad (3.17)$$

biçiminde tanımlanır. Eşitlik (3.17)' de $\boldsymbol{\beta}$ regresyon parametrelerini $\mu_0(\cdot)$ bilinmeyen temel ortalama fonksiyonunu ve \mathbf{Z} açıklayıcı değişken vektörünü göstermektedir. Burada açıklayıcı değişkenlerin zamandan bağımsız olduğu varsayılmıştır.

Eğer $\tilde{T}_{ik}(\boldsymbol{\beta}) = T_{ik} e^{\boldsymbol{\beta}^T \mathbf{Z}_i}$ ve $\tilde{C}_i(\boldsymbol{\beta}) = C_i e^{\boldsymbol{\beta}^T \mathbf{Z}_i}$ yazılırsa, i . bireyin $t \wedge \tilde{C}_i$ zamanına dek deneyimlediği olay sayısını gösteren sayma süreci,

$$\tilde{N}_i(t; \boldsymbol{\beta}) = \sum_{k=1}^{\infty} I(\tilde{T}_{ik}(\boldsymbol{\beta}) \leq t \wedge \tilde{C}_i(\boldsymbol{\beta}))$$

olarak elde edilir. Yukarıdaki eşitlikten $\tilde{N}_i(t; \boldsymbol{\beta}) = N_i(t e^{-\boldsymbol{\beta}^T \mathbf{Z}_i})$ elde edileceği açıktır. Bu durumda Eşitlik (3.17),

$$E\{\tilde{N}_i(t; \boldsymbol{\beta})\} = \mu_0(t) \quad (3.18)$$

biçiminde ifade edilebilir. Eşitlik (3.17), $\mathbf{Z}_i = \mathbf{Z}$ verildiğinde t zamanına dek beklenen olay sayısının, $\mathbf{Z}_i = 0$ verildiğinde $t e^{\boldsymbol{\beta}^T \mathbf{Z}_i}$ zamanına dek beklenen olay sayısına eşit olduğunu ifade etmektedir. Başka bir deyişle, \mathbf{Z}_i açıklayıcı değişken vektörü, $\mathbf{Z}_i = 0$ ' a göre zaman ölçeğini $t e^{\boldsymbol{\beta}^T \mathbf{Z}_i}$ kadar genişleterek ya da daraltarak zaman boyunca meydana gelen yineleme sıklığını etkilemektedir.

β regresyon parametrelerinin tahmini için Lin ve arkadaşları tarafından önerilen hızlandırılmış başarısızlık zamanı modelleri kullanılarak elde edilen tahmin denklemi,

$$U(\beta) = \sum_{i=1}^n \int_0^{\infty} Q(t; \beta) (Z_i - \bar{Z}(t; \beta)) d\tilde{N}_i(t; \beta) \quad (3.19)$$

gibidir. Eşitlik (3.19)' da $Q(t; \beta)$, önceden belirlenen ağırlık fonksiyonu ve

$$\bar{Z}(t; \beta) = \frac{\sum_{i=1}^n Y_i(t; \beta) Z_i}{\sum_{i=1}^n Y_i(t; \beta)}$$

olarak bulunur. Yukarıdaki eşitlikte $Y_i(t; \beta) = I(C_i \geq te^{-\beta^T Z_i})$ ' dir. $Q(t; \beta) = 1$ alınırsa

Eşitlik (3.19), log-rank tahmin fonksiyonu olarak; $Q(t; \beta) = n^{-1} \sum_{i=1}^n Y_i(t; \beta)$ alınırsa

Eşitlik (3.19), Gehan tahmin fonksiyonu olarak adlandırılır. Eşitlik (3.19)' u sıfır yapan değerler β parametre vektörünün tahminini vermektedir. Ancak, Eşitlik (3.19)' un β ' ya göre türevi süreksiz olduğundan Newton-Raphson algoritması kullanılarak çözülemez. Dolayısıyla bu eşitliğin çözümü, farklı algoritmalar kullanılmayı gerektirir [25].

3.3.6. Paylaşılmış Zayıflık Modelleri

Aynı bireyden elde edilen yinelenmeli olay verileri bireye özgü faktörler nedeniyle ilişkili olmaktadır. Paylaşılmış zayıflık modelinde, aynı kişiden elde edilen yinelenmeler bir küme oluşturur ve her küme aynı zayıflık düzeyini paylaşır.

Önceki modellerin aksine paylaşılmış zayıflık sadece kestirilmiş varyansları düzeltmez, modele eklenerek parametre kestirimlerini de etkiler.

Literatürde paylaşılmış zayıflık modeline ilişkin çok sayıda çalışma vardır. Zayıflık terimi, Cox orantılı tehlike modeli, hızlandırılmış başarısızlık zamanı modeli gibi pek çok modele entegre olabilir. Zayıflık teriminin orantılı tehlike modeline eklenmesiyle oluşturulan çalışmalardan bazıları aşağıda açıklanmıştır.

McGilchrist ve Aisbett (1991)' de, yinelemeli olay verisini Cox orantılı tehlike modeline zayıflık terimini ekleyerek modellemiştir. Bu modelde yoğunluk fonksiyonu,

$$\lambda(t) = \nu \lambda_0(t) \exp(\beta^T \mathbf{Z})$$

biçiminde alınmış; regresyon ve zayıflık parametrelerinin tahminleri en çok olabilirlik yöntemini kullanarak elde edilmiştir [29].

McGilchrist ve Aisbett (1991) tarafından önerilen modelin genişletilmiş biçimi Nielsen ve arkadaşları (1992) tarafından önerilmiştir [30]. Nielsen ve arkadaşları, yinelemeli olay verisini sayma süreci notasyonu kullanarak modellemiştir.

Zayıflık terimini ekleyerek oluşturdukları modeller,

- Tek örneklem için,

$$\lambda_{il}(t) = \nu_i Y_{il}(t) \lambda_0(t), \quad i = 1, 2, \dots, n ; l = 1, 2, \dots, n_i$$

- k örneklem için,

$$\lambda_{ihl}(t) = \nu_i Y_{ihl}(t) \lambda_{0h}(t), \quad i = 1, 2, \dots, n ; h = 1, 2, \dots, k ; l = 1, 2, \dots, n_{ih}$$

- k örneklem için orantılı tehlike modeli,

$$\lambda_{ihl}(t) = \nu_i Y_{ihl}(t) \exp(\beta_h) \lambda_0(t), \quad i = 1, 2, \dots, n ; h = 1, 2, \dots, k ; l = 1, 2, \dots, n_{ih}$$

- Regresyon modeli,

$$\lambda_i(t) = \nu_i Y_i(t) \exp(\boldsymbol{\beta}^T \mathbf{Z}_i(t)) \lambda_0(t), \quad i = 1, 2, \dots, n$$

biçimindedir.

Zayıflık terimi eklenerek oluşturulan bu modellerde, $\lambda_0(t)$ ve $\lambda_{0h}(t)$ temel yoğunluk fonksiyonlarını; n , çalışmadaki birey sayısını; k , çalışmadaki tabaka (örneğin tedavi grupları) sayısını; n_i , aynı zayıflık terimini paylaşan birey sayısını; $Y_i(t)$, t anında i . bireyin hayatta olup olmadığını gösteren risk sürecini; $\boldsymbol{\beta}$ açıklayıcı değişken vektörünü ve ν_i , i . bireye ait zayıflık terimini göstermektedir. Parametre tahminleri için, parametrik, yarı-parametrik ve parametrik olmayan yöntemler kullanılmıştır [30].

4. POISSON SÜREÇLERİ

Poisson süreçleri, olayların zaman boyunca rasgele ortaya çıktığı ve birbirleriyle ara kesiti olmayan zaman aralıklarında ortaya çıkan olay sayılarının birbirinden bağımsız olduğu süreçlerdir. Poisson süreçleri, Bernoulli denemelerinin oluşturduğu sayma sürecinin sürekli zaman parametrelili bir türüdür [31].

Bu bölümde, yinelenmeli olay verisinin modellenmesinde kullanılan homojen ve homojen olmayan Poisson süreçleri açıklanmıştır.

4.1. Homojen Poisson Süreci

Homojen bir Poisson süreci, birim zamanda ortaya çıkması beklenen olay sayısının zaman boyunca değişmediği stokastik bir süreçtir. $N(t, t + \Delta t)$, $[t, t + \Delta t)$ gibi çok küçük zaman aralığında ortaya çıkan olay sayısını ve $N(t) = N[0, t)$ de $[0, t)$ zaman aralığında ortaya çıkan olay sayısını gösterebilir. Bu durumda, $\{N(t), t \geq 0\}$ süreci aşağıda verilen aksiyomlar altında λ yoğunluğuyla (oranıyla) Poisson sürecine sahiptir ve bu aksiyomlar aşağıda verilmiştir:

- $N(0) = 0$
- $\{N(t), t \geq 0\}$ süreci, bağımsız artışlara sahip bir süreçtir.
- $P(N(t, t + \Delta t) = 1) = \lambda \Delta t + o(\Delta t)$
- $P(N(t, t + \Delta t) = 0) = 1 - \lambda \Delta t + o(\Delta t)$
- $P(N(t, t + \Delta t) > 1) = o(\Delta t)$
- $\lim_{\Delta t \rightarrow 0} \frac{o(\Delta t)}{\Delta t} = 0$

Yukarıdaki aksiyomlarda yer alan λ , sürecin yoğunluk fonksiyonu olmak üzere sürekli bir fonksiyondur. Böylece,

$$\Lambda(t) = \int_0^t \lambda \, du = \lambda t \quad (4.1)$$

yazılabilir ve bu ifade birikimli yoğunluk ya da ortalama fonksiyonu olarak adlandırılır [31].

Çalışmada n bireyin olduğu ve süreçte i . ($i = 1, 2, \dots, n$) bireye ilişkin sayma sürecinin λ_i yoğunluğuyla homojen Poisson sürecine sahip olduğu varsayımı altında,

$$N_i(t) \sim \text{Poisson}(\lambda_i) \quad (4.2)$$

yazılabilir.

Her bir bireye ait yoğunluklar aynı olabileceği gibi farklı da olabilir. Her bir bireye ilişkin açıklayıcı değişken $Z(\cdot)$ bilgisinin de olduğu durumda, i . bireye ilişkin sayma süreci, parametrik orantılı tehlike modellerinden yararlanarak $\lambda_i \exp(\boldsymbol{\beta}^T \mathbf{Z}_i)$ yoğunluğuyla homojen Poisson sürecine sahiptir ve

$$N_i(t) \sim \text{Poisson}(\lambda_i \exp(\boldsymbol{\beta}^T \mathbf{Z}_i)) . \quad (4.3)$$

biçiminde gösterilir.

4.2. Homojen Olmayan Poisson Süreci

Homojen Poisson sürecinde λ yoğunluğu zaman boyunca değişmez iken, homojen olmayan Poisson sürecinde λ yoğunluğu zamanın bir fonksiyonudur. Homojen olmayan

Poisson süreci, homojen Poisson sürecine benzer şekilde bağımsız artmalara sahiptir. $\{N(t), t \geq 0\}$ sürecinde $N(t)$ 'nin dağılımı,

$$\Lambda(t) = \int_0^t \lambda(u) du \quad (4.4)$$

ortalamasıyla Poisson dağılımı gösterir. Çalışmada n bireyin olduğu ve süreçte i . ($i=1,2,\dots,n$) bireye ilişkin sayma sürecinin $\Lambda_i(t)$ yoğunluğuyla homojen olmayan Poisson sürecine sahip olduğu varsayımı altında, $N_i(t)$ 'nin dağılımı,

$$N_i(t) \sim \text{Poisson}(\Lambda_i(t)) \quad (4.5)$$

biçimindedir.

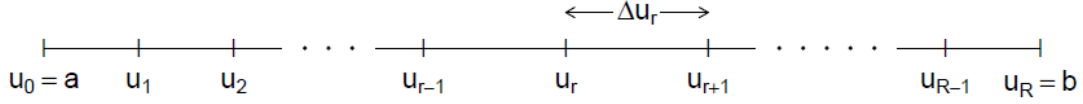
Her bir bireye ilişkin açıklayıcı değişken $Z(\cdot)$ bilgisinin de olduğu durumda, i . bireye ilişkin sayma sürecinin, parametrik orantılı tehlike modellerinden yararlanarak $\Lambda_i(t)\exp(\boldsymbol{\beta}^T \mathbf{Z}_i)$ yoğunluğuyla homojen olmayan Poisson sürecine sahip olduğu söylenir ve aşağıdaki eşitlikteki gibi ifade edilir:

$$N_i(t) \sim \text{Poisson}(\Lambda_i(t)\exp(\boldsymbol{\beta}^T \mathbf{Z}_i)). \quad (4.6)$$

4.3. Homojen Poisson Süreci için Olabilirlik Fonksiyonunun Çıkarımı

$\{N_i(t), t \geq 0\}$ sürecine ilişkin parametre tahminleri en çok olabilirlik tahmin yöntemi kullanılarak elde edilebilir. Sürece ilişkin olabilirlik fonksiyonunu yazabilmek için çarpım integralinden yararlanır. Bu nedenle çarpım integrali aşağıda açıklanmıştır.

$g(u)$, $[a, b]$ aralığında sürekli, integrallenebilir bir fonksiyon olsun. $[a, b]$ aralığı $a = u_0 < u_1 < \dots < u_R = b$ olacak şekilde parçalsın. $\Delta u_r = u_{r+1} - u_r$, $r = 0, 1, \dots, R$ olarak tanımlansın ve $u_{R+1} = u_{R^+}$ olsun. Bu parçalama aşağıdaki Şekil (4.1)'de gösterilmektedir.



Şekil 4.1 $[a, b]$ aralığının çok küçük aralıklara parçalanışı

$g(u)$ 'nin $[a, b]$ aralığındaki çarpım integrali,

$$\mathcal{P}_a^b \{1 + g(u) du\} = \lim_{R \rightarrow \infty} \prod_{r=0}^R \{1 + g(u_r) \Delta u_r\}. \quad (4.7)$$

biçiminde tanımlanır. Eşitlik (4.7)'de $R \rightarrow \infty$ için $\max(\Delta u_r) \rightarrow 0$ olur. $\log \{1 + g(u) \Delta u\} = g(u) \Delta u + o(\Delta u)$ eşitliğinden yararlanarak Eşitlik (4.7)'nin logaritması limit durumunda Riemann integraline yakınsar ve

$$\mathcal{P}_a^b \{1 + g(u) du\} = \exp \left\{ \int_a^b g(u) du \right\} \quad (4.8)$$

eşitliği elde edilir. Eşitlik (4.8)'ten,

$$\mathcal{P}_a^b \{1 + g(u) du + o(du)\} = \exp \left\{ \int_a^b g(u) du \right\} \quad (4.9)$$

olduğu kolayca görülebilir [18].

Süreçteki her i . bireyin τ_i zamanına dek gözlemlendiği varsayalım. i . bireyin τ_i zamanına dek deneyimlediği olay sayısı $N_i(\tau_i)$; i . bireyin t^- zamanına kadar olan geçmiş bilgisi $\mathcal{H}_i(t^-) = \{N(s), Z(s), 0 \leq s < t\}$ ile gösterilsin. i . bireye ilişkin olabilirlik fonksiyonu,

$$\begin{aligned} P(\mathcal{H}_i(t^- + \Delta t) | \mathcal{H}_i(t^-)) &= P(\Delta N_i^R(t), \Delta N_i^C(t) | \mathcal{H}_i(t^-)) \\ &= P(\Delta N_i^R(t) | \mathcal{H}_i(t^-)) P(\Delta N_i^C(t) | \Delta N_i^R(t), \mathcal{H}_i(t^-)) \end{aligned} \quad (4.10)$$

terimlerinin çarpılmasından elde edilir Eşitlik (4.10)' da $\Delta N_i^R(t)$, i . bireyin $[t, t + \Delta t)$ zaman aralığında başarısızlık (yineleme) yaşaması ve $\Delta N_i^C(t)$, i . bireyin $[t, t + \Delta t)$ zaman aralığında durdurma yaşaması durumlarını göstermektedir.

Bilgi içermeyen durdurma varsayımı altında Eşitlik (4.10),

$$P(\mathcal{H}_i(t^- + \Delta t) | \mathcal{H}_i(t^-)) \propto P(\Delta N_i^R(t) | \mathcal{H}_i(t^-)). \quad (4.11)$$

biçiminde ifade edilebilir. $0 \leq T_{i,1} < T_{i,2} < T_{i,3} < \dots < T_{i,R} < \tau_i$ i . bireye ait sıralı başarısızlık zamanlarını gösterebilir. i . bireye ait takip süresi olan $[0, \tau_i)$ zaman aralığı $0 = u_0 < u_1 < \dots < u_R = \tau_i$ olacak şekilde ara kesiti olmayan küçük aralıklara parçalsın. $[u_r, u_{r+1})$ zaman aralığında 0 ya da 1 olayın ortaya çıkması beklenildiğinden bu aralıkta ortaya çıkan olay sayısı Bernoulli dağılımından yararlanarak elde edilebilir. Homojen Poisson sürecinin *iii.* ve *iv.* aksiyomlarından yararlanarak i . bireye ait olabilirlik fonksiyonu aşağıdaki gibi elde edilir:

$$\begin{aligned} L_i &= \prod_{j=1}^{N_i(\tau_i)} \prod_{r=0}^R \left\{ \lambda(u_r | \mathcal{H}_i(u_{r-})) \Delta u_r + o(\Delta u_r) \right\}^{\Delta N_i^R(u_r)} \times \\ &\quad \left\{ 1 - \lambda(u_r | \mathcal{H}_i(u_{r-})) \Delta u_r + o(\Delta u_r) \right\}^{1 - \Delta N_i^R(u_r)} \end{aligned} \quad (4.12)$$

Eşitlik (4.12)' de $R \rightarrow \infty$ için $\max(\Delta u_r) \rightarrow 0$ ' dır. i . bireye ait sıralı başarısızlık zamanlarını $(T_{i,1}, T_{i,2}, T_{i,3}, \dots, T_{i,R})$ içeren R tane aralık için $\Delta N^R(u_r) = 1$, diğerleri için $\Delta N^R(u_r) = 0$ değerini almaktadır. Çarpım integralinden yararlanılarak Eşitlik (4.12),

$\prod_{r=0}^R (\Delta u_r)^{\Delta N^R(u_r)}$ ifadesine bölüldüğünde $R \rightarrow \infty$ için,

$$L_i = \prod_{j=1}^{N_i(\tau_i)} \left\{ \lambda(t_{ij} | \mathcal{H}_i(t_{ij})) \right\} \exp \left\{ - \int_0^{\tau_i} \lambda(u | \mathcal{H}_i(u^-)) du \right\} \quad (4.13)$$

elde edilir. Eşitlik (4.13)' te verilen ifade n birey için yazılırsa olabilirlik fonksiyonu,

$$L = \prod_{i=1}^n \left\{ \prod_{j=1}^{N_i(\tau_i)} \left\{ \lambda(t_{ij} | \mathcal{H}_i(t_{ij})) \right\} \right\} \exp \left\{ - \int_0^{\tau_i} \lambda(u | \mathcal{H}_i(u^-)) du \right\} \quad (4.14)$$

biçiminde bulunur. Parametre tahminleri Eşitlik (4.14)' te verilen olabilirlik fonksiyonunun maksimum yapılmasıyla elde edilir [14, 18, 28].

4.4. Poisson Süreçlerinin Kullanıldığı Yinelemeli Olay Modelleri

Literatürde yinelemeli olayları, bilgi içermeyen durdurma varsayımı altında Poisson süreçleri ile modelleyen çok sayıda çalışma bulunmaktadır.

Lawless (1987), yinelemeli olayları, aşağıda verilen yoğunluk fonksiyonu ile homojen olmayan Poisson süreci ile modellemiştir.

$$\lambda(t) = \lambda_0(t, \theta) \exp(\beta^T \mathbf{Z})$$

$$\Lambda(t) = \int_0^t \lambda(u) du .$$

Yukarıdaki eşitliklerde $\lambda_0(t; \theta)$, temel yoğunluk fonksiyonunu, $\Lambda(t)$, birikimli yoğunluk fonksiyonunu ve $Z(\cdot)$, her bir bireye ilişkin sabit açıklayıcı değişken vektörünü göstermektedir.

Çalışmada m birey olduğu ve i . bireye ait τ_i takip süresi boyunca n_i olayın ortaya çıktığı varsayalım. $(t_{i1} < t_{i2} \dots < t_{in_i})$, i . bireye ilişkin sıralı başarısızlık zamanlarını göstermek üzere, m birey için olabilirlik fonksiyonu,

$$L(\theta, \beta) = \prod_{i=1}^m \left[\prod_{j=1}^{n_i} \lambda_i(t_{ij}) \right] \exp[-\Lambda_i(T_i)] \quad (4.15)$$

biçiminde elde edilir. Eşitlik (4.15)' te verilen olabilirlik fonksiyonu,

$$L(\theta, \beta) = \left\{ \prod_{i=1}^m \prod_{j=1}^{n_i} \frac{\lambda_0(t_{ij}; \theta)}{\Lambda_0(T_i; \theta)} \right\} \prod_{i=1}^m \exp[-\Lambda_0(T_i; \theta) e^{\beta^T Z}] \exp[-\Lambda_0(T_i; \theta) e^{\beta^T Z}]^{n_i}$$

olarak yeniden düzenlenebilir. Böylece olabilirlik fonksiyonu,

$$L(\theta, \beta) = L_1(\theta) L_2(\theta, \beta) \quad (4.16)$$

biçiminde çarpanlara ayrılabilir. Bu çarpanlar yardımı ile parametrelerin en çok olabilirlik tahminlerini elde edilmiştir. Ayrıca çalışmada, zayıflık terimi v_i ,

$$\lambda(t) = \lambda_0(t; \theta) \exp(v_i + \beta^T Z)$$

modele dahil edilmiş ve parametre tahminleri için Eşitlik (4.16)' da verilen çarpanlara ayırmaya benzer bir yaklaşımla elde edilmiştir [32].

Balshaw (1997), tez çalışmasında yinelemeli olay sürecini homojen olmayan Poisson süreci olarak ele almış ve yarı olabilirlik tahmin yöntemini kullanarak parametre tahminlerini elde etmiştir [33].

Nielsen (2007), tez çalışmasında, yinelemeli olay sürecini karışık etkili homojen olmayan Poisson süreci ile modellemiştir. Bireye özgü heterojenlik ise zayıflık terimini modele dahil ederek açıklamıştır. Sayma sürecine ilişkin yoğunluk fonksiyonunun, eğrisel çizgili düzleştirme fonksiyonu olduğunu varsaymıştır. Açıklayıcı değişken etkilerini de eğrisel çizgileri kullanarak modellemiştir. Bu sayede model, zamana bağlı açıklayıcı değişken kullanımına olanak sağlamaktadır [34].

Fredette (2011), tez çalışmasında yinelemeli olay sürecini homojen ve homojen olmayan Poisson süreci olarak almış ve çeşitli yoğunluk fonksiyonları kullanarak parametre tahminlerini elde etmiştir. Fakat kullanılan yoğunluk fonksiyonlarında açıklayıcı değişken bilgisi yer almamaktadır [35].

5. BİLGİ İÇEREN DURDURMA VARLIĞINDA YİNELEMELİ OLAY ANALİZİ

Yinelemeli olay verileri çoğu kez uzun dönem izlenen çalışmalarla toplanır ve çalışmadaki başarısızlık zamanı; ölüm gibi bir başarısızlık olayı, çalışmanın sona ermesi ya da izlenme döneminde bireyin çalışmadan çekilmesi sebepleriyle durdurmaya maruz kalabilir [2, 5]. Eğer durdurma, çalışmanın sona ermesinden kaynaklanıyorsa, durdurma zamanı bağımsız ya da bilgi içermeyen şekilde kabul edilir. Birçok uygulamada yinelemeli olaylar, bireyin çalışmadan çekilmesi ya da bireyin ölümü ile sonlandırılır. Bu şekilde meydana gelen durdurma zamanı bilgi içerir ve yinelemeli olay zamanları ile ilişkilidir [36-41]. Durdurma zamanı ile yineleme zamanları arasındaki ilişki daha açık bir şekilde şöyle ifade edilebilir: Bir bireye ilişkin yineleme sıklığı beklenmedik şekilde artarsa (azalırsa), o bireye ilişkin ölüm riski de artar (azalır). Örneğin, kalp krizi gibi hayati önem taşıyan olayların yinelenmesi ölüm riskinin artmasına sebep olur. Ölüm olayı, yineleme yaşanması engellediği için bağımlı durdurma söz konusudur [42]. Söz konusu ilişkiyi göz ardı ederek modelleme yapmak yanlış tahminler elde edilmesine neden olmaktadır [43-45].

Literatürde bilgi içeren durdurma varlığı altında yinelemeli olay verisini modellemek için önerilen çok sayıda çalışma vardır. Bu çalışmalardan önemlileri aşağıda verilen kesimde detaylı olarak açıklanmıştır.

5.1. Bilgi İçeren Durdurma Varsayımı Altında Yapılan Çalışmalar

Cook ve Lawless, (1997)' de yinelemeli olayları, bilgi içeren durdurma varlığı altında ele almışlar; yinelemeli olaylar ile bilgi içeren olay arasındaki bağımlılık yapısını bileşik modeller kullanarak modellemişlerdir. $N_i(t)$, i . ($i = 1, 2, \dots, n$) bireyin t zamanına dek gözlemlediği başarısızlık sayısını, $T_{i,k}$, i . bireyin k . ($k = 1, 2, \dots, K$) yineleme zamanını, $Z_i = (Z_{1i}, \dots, Z_{pi})$ i . bireyin zamandan bağımsız açıklayıcı değişken vektörünü göstermek üzere, modellerinde yoğunluk fonksiyonlarını,

$$\phi_i(t; H_i(t)) = \phi_0(t) \exp\{\beta_1^T \mathbf{Z}_{1i} + \beta_2^T \mathbf{Z}_{2i}(t)\},$$

$$\lambda_i(t; H_i(t)) = \lambda_0(t) \exp\{\beta_3^T \mathbf{Z}_{3i}\}$$

olarak tanımlanmıştır. Burada $Y_i(t) = I(T_i \geq t)$, i . bireyin t anında hayatta olup olmadığını ifade eden risk sürecini göstermek üzere, $\mathcal{H}_i(t) = \{N_i(s), Y_i(s), s < t\}$, t zamanına dek sürecin geçmiş bilgisini göstermektedir. β_1 ve β_3 sırasıyla, bilgi içeren durdurma olayı ve yineleme olayı için yoğunluk fonksiyonlarındaki açıklayıcı değişken etkisini; β_2 , t zamanına dek gözlenen yinelemeli olay sayısının etkisini göstermektedir. Parametre tahminleri için yarı parametrik yöntemleri kullanmışlardır [46].

Huang ve Wolfe (2002), yinelemeli olay verisinde, başarısızlık zamanı ile durdurma olayı arasındaki ilişkiyi zayıflık terimi (ν) ile modellemişlerdir i . ($i = 1, 2, \dots, n$) bireyin k . ($k = 1, 2, \dots, K$) yineleme zamanı $T_{i,k}$, i . bireyin bilgi içeren durdurma zamanı D_i ; bilgi içermeyen durdurma zamanı C_i ve $\mathbf{Z}_i = (Z_{1i}, \dots, Z_{pi})$ i . bireyin zamandan bağımsız açıklayıcı değişken vektörü olmak üzere, modellerinde, yoğunluk fonksiyonunu,

$$\lambda_{ik}^R(t | \mathbf{Z}_{ik}^R, \mathbf{Z}_{ik}^D, \nu_i, D_{ik}) = \lambda_0^R(t) \exp\left\{(\beta^R)^T \mathbf{Z}_{ik}^R + \nu_i\right\}$$

$$\lambda_{ik}^D(t | \mathbf{Z}_{ik}^R, \mathbf{Z}_{ik}^D, \nu_i, D_{ik}) = \lambda_0^D(t) \exp\left\{(\beta^D)^T \mathbf{Z}_{ik}^D + \alpha \nu_i\right\}$$

biçimlerinde tanımlamışlardır. Burada β^R ve β^D sırasıyla yineleme olayı ve bilgi içeren durdurma olayı için yoğunluk fonksiyonlarındaki açıklayıcı değişken etkisini, α ise bilgi içeren durdurma olayı için zayıflık terimini arttıran (azaltan) bir faktördür. Zayıflık teriminin dağılımının normal dağılıma sahip olduğu varsayılmıştır. Parametre tahminlerini EM ve Metropolis-Hastings algoritmaları kullanarak elde etmişlerdir [38].

Huang ve Wang (2004) çalışmalarında $N(t)$ sayma sürecini homojen olmayan Poisson süreciyle modellemiş, yineleme süreci ile durdurma süreci arasındaki bağımlılık yapısını paylaşılmış zayıflık terimi kullanarak bağlamışlardır. Yinelemeli olay sürecine ilişkin yoğunluk fonksiyonunu,

$$\lambda^R(t) = \nu \lambda_0(t) \exp(\boldsymbol{\beta}^T \mathbf{Z}), \quad 0 < t < \tau$$

ve durdurma sürecine ilişkin yoğunluk fonksiyonunu da,

$$\lambda^D(t) = \nu \lambda_0(t) \exp(\boldsymbol{\alpha}^T \mathbf{Z}), \quad 0 < t < \tau$$

biçiminde tanımlamışlardır. Burada ν , zayıflık terimini, $\boldsymbol{\beta}$ ve $\boldsymbol{\alpha}$ sırasıyla yineleme olayı ve bilgi içeren durdurma olayı için açıklayıcı değişkenlere ilişkin parametre vektörünü göstermektedir. Huang ve Wang modellerinde önce birikimli temel yoğunluk fonksiyonunun tahminini, Kaplan-Meier tahmin yöntemi kullanarak tahmin etmişlerdir. Daha sonra birikimli temel yoğunluk fonksiyonunun tahmini kısmi olabilirlik fonksiyonunda yerine yazılarak öteki parametrelerin tahminleri elde edilmiştir [39].

Huang, Qin ve Wang (2010), yinelemeli olay verisinde, başarısızlık zamanı ile durdurma olayı arasındaki ilişkiyi zayıflık terimi (ν) ile modellemişlerdir. Modellerinde zayıflık teriminin dağılımına ilişkin bir varsayımda bulunmamışlardır. ν_i , i . ($i = 1, 2, \dots, n$) bireye ilişkin zayıflık terimini ve $N_i(t)$, i . bireyin t zamanına dek yaşadığı başarısızlık sayısını gösterebilir. ν_i bilindiğinde, $\{N_i(t), t > 0\}$ homojen olmayan Poisson süreci olduğu varsayımı altında Huang ve arkadaşlarının önermiş olduğu yoğunluk fonksiyonu aşağıda verilmiştir:

$$\lambda_i(t) = \nu_i \lambda_0(t) \exp(\mathbf{X}_i(t)^T \boldsymbol{\beta} + \mathbf{W}_i^T \boldsymbol{\gamma}), \quad 0 \leq t \leq \tau.$$

Burada X_i ve W_i , sırasıyla i . bireye ait zamana bağlı ve zamandan bağımsız açıklayıcı değişken vektörlerini; β ve γ sırasıyla zamana bağlı ve zamandan bağımsız açıklayıcı değişkenlere ilişkin parametre vektörlerini göstermektedir. Huang ve arkadaşları, model parametrelerini, yineleme zamanları ve zayıflık terimi için herhangi bir dağılım varsayımında bulunmadan yarı parametrik yöntemler kullanarak tahmin etmişlerdir [47].

Yukarıda açıklanan modellerden farklı olarak yinelemeli olay verisini parametrik olmayan yöntemler kullanarak modelleyen çalışmalar da bulunmaktadır. Bunlardan bazıları aşağıda özetlenmiştir.

Ghosh ve Lin (2000), durdurma olayının varlığı altında birikimli yinelemeli olay sayısının marjinal ortalamalarına dayanan tek ve iki örnekleme ilişkin parametrik olmayan bir yöntem geliştirmişlerdir [48].

Ghosh ve Lin (2002)' de, yaşam verisinde ölüm olayı gibi bilgi içeren durdurma olduğu zaman yinelemeli olayların ortaya çıkış zamanlarını modellemede marjinal ortalama fonksiyonunu ele almışlar; olayların stokastik yapısını dikkate almadan açıklayıcı değişkenlere ilişkin parametrelerin tahminlerini parametrik olmayan yöntemler kullanarak elde etmişlerdir [49].

Ghosh ve Lin (2003), daha sonra yinelemeli olay ve bağımlı durdurma süreçlerinin marjinal modellerini hızlandırılmış başarısızlık zamanı modelleriyle formüle eden yarı parametrik bileşik bir model önermişlerdir [50].

Bu tez çalışmasında, bilgi içeren durdurma varlığı altında yinelemeli olaylar homojen Poisson süreci kullanarak modellenmiş; yinelemeli olaylar ile durdurma olayı arasındaki bağıntıyı kurmak için iki yeni model önerilmiştir. Çalışmada önerilen ikinci model için Liu, Wolfe ve Huang (2004) tarafından yapılan çalışma temel alındığından dolayı bundan sonraki kesimde bu çalışma ayrıntılı biçimde anlatılmıştır.

5.2. Liu, Wolfe ve Huang Tarafından Önerilen Yöntem

Liu ve arkadaşları (2004), yinelemeli olay verisini bilgi içeren durdurma varsayımı altında Poisson süreçlerini kullanarak modellemiştir. Çalışmalarında ölüm olayını, bilgi içeren durdurma olarak ele almışlar; durdurma olayı ile yinelemeli olaylara ilişkin bileşik olabilirlik fonksiyonunu paylaşılmış zayıflık modellerinden yararlanarak elde etmişlerdir [37].

Liu ve arkadaşları tarafından önerilen modeli açıklayabilmek için aşağıdaki tanımlamalar verilsin:

C_i , i . bireye ait bilgi içermeyen durdurma zamanını, D_i , i . bireye ait bilgi içeren durdurma zamanını ve $T_{i,k}$, i . bireyin k . yineleme zamanını gösterebilir. Liu ve arkadaşları, bilgi içeren durdurma olayını, ölüm olarak ele almışlardır.

Tez çalışmasında bilgi içeren durdurma olarak özellikle ölüm olayı ile ilgilenildiği için bundan sonraki kesimlerde “ölüm” denildiğinde “bilgi içeren durdurma”; “durdurma” denildiği zaman ise, “bilgi içermeyen durdurma” anlaşılacaktır.

$\tau_i = \min(D_i, C_i)$, i . bireye ilişkin takip süresi ve $\Delta_i = I(D_i \leq C_i)$, i . bireyin ölüm olayını yaşama durumunu gösteren gösterge değişkenidir. $Y_i(t) = I(\tau_i \geq t)$, i . bireyin t zamanında hayatta olup olmadığını gösteren risk süreci ve $N_i^D(t) = I(\tau_i \leq t, \Delta_i = 1)$, i . bireyin t zamanına dek ölme olayını yaşayıp yaşamadığını gösteren sayma sürecidir. $dN_i^R(t) = \lim_{\Delta t \rightarrow 0} N_i^R(t + \Delta t^-) - N_i^R(t^-)$, i . bireyin $(0, \Delta t]$ zaman aralığında deneyimlediği yineleme sayısını gösteren sayma süreci olmak üzere i . bireyin t zamanına dek yaşadığı yinelemeleri sayan sayma süreci $N_i^R(t) = \int_0^t Y_i(t) dN_i^R(t)$ olarak ifade edilir. i . bireye ait geçmiş bilgisi $\mathcal{H}_i(t) = \{N_i^R(u), N_i^D(u), Z_i(u), Y_i(u), 0 \leq u \leq t\}$ ile gösterilmektedir. O_i , i . bireye ait gözlenen veriyi; ν_i , ise, i . bireye ait zayıflık terimini göstermektedir.

Liu ve arkadaşlarının, önerdikleri model için yaptıkları varsayımlar aşağıda verilmiştir:

1. Yinelemeli olay, ölüm ve durdurma süreçleri sürekli dağılıma sahiptir. Ölüm, durdurma ve yinelemeli olaylar aynı anda meydana gelemez. $[t, t + \Delta t)$ çok küçük zaman aralığında önce ölüm olayının gerçekleştiği varsayılmıştır. Bu yüzden i . bireye ait yineleme süreci $N_i(t)$, ölüm olayından sonra değişmez. Durdurma sürecinde ise, yinelemeli olaylar durdurma zamanından sonra ortaya çıkabilir ancak gözlenememektedir. Dolayısıyla, ölüm süreci ile yinelemeli olay süreci bağımsız değildir.
2. Ölüm olayına ilişkin yoğunluk fonksiyonu,

$$\begin{aligned} P(dN^D(t) = 1 | \mathcal{H}_i(t^-)) &= Y_i(t) d\Lambda_i(t) \equiv Y_i(t) \lambda_i(t) dt \\ d\Lambda_i(t) &= P(dN^D(t) = 1 | Z, v, D \geq t) \end{aligned} \quad (5.1)$$

biçiminde tanımlanmıştır. Eşitlik (5.1)' de $\lambda_i(t)$, i . bireye ilişkin ölüm süreci için temel yoğunluk fonksiyonunu göstermektedir.

3. Ölüm olayının varlığı altında yinelemeli olay sürecine ilişkin yoğunluk fonksiyonunu,

$$\begin{aligned} P(dN^R(t) = 1 | \mathcal{H}_i(t^-), D \geq t) &= Y_i(t) dR_i(t) \equiv Y_i(t) r_i(t) dt \\ dR_i(t) &= P(dN_i^R(t) = 1 | Z_i, v_i, D_i \geq t) \end{aligned} \quad (5.2)$$

biçiminde tanımlanmıştır. Eşitlik (5.2)' de $r_i(t)$, i . bireye ilişkin yinelemeli olay süreci için temel yoğunluk fonksiyonunu göstermektedir.

4. Durdurma, zayıflık terimi v_i ' ye ilişkin bilgi içermemektedir.

5. Zayıflık teriminin dağılımı, $v_i \sim \text{Gamma}\left(\frac{1}{\theta}, \theta\right)$ olarak alınmıştır. v_i ' nin olasılık yoğunluk fonksiyonu

$$f(v_i, \theta) = \frac{1}{\Gamma(1/\theta)\theta^{1/\theta}} v_i^{\frac{1}{\theta}-1} e^{-v_i/\theta}, \quad v_i \geq 0 \quad (5.3)$$

$$= 0, \quad \text{ö.d}$$

biçimindedir.

Yukarıda yapılan varsayımlar ve tanımlamalar altında, O_i gözlenen değerleri için olabilirlik fonksiyonu çarpım integralinden yararlanarak,

$$L = L(\mathcal{H}(0)) L(O | \mathcal{H}(0)) \quad (5.4)$$

$$= \mathcal{F}_0^\infty \left\{ L \left[\mathcal{H}(t^- + \Delta t) | \mathcal{H}(t^-) \right] \right\}$$

yazılabilir [14]. Eşitlik (5.4)' te $\mathcal{H}(0)$, sürecin $t=0$ anındaki bilgisini göstermektedir. Eşitlik (5.4)' teki olabilirlik fonksiyonu durdurma sürecinin bilgisi göz ardı edildiğinde,

$$L(O_i | \mathcal{H}(0)) \propto \mathcal{F}_0^\infty \left\{ L \left(dN_i^R(t), dN_i^D(t) | \mathcal{H}(t^-) \right) \right\} \quad (5.5)$$

biçiminde yazılabilir.

Eşitlik (5.5), aynı zamanda

$$\begin{aligned} \mathcal{P}_0^\infty \left\{ L\left(dN_i^R(t), dN_i^D(t) \middle| \mathcal{H}(t^-) \right) \right\} = \\ \mathcal{P}_0^\infty \left\{ L\left(dN_i^R(t) \middle| \mathcal{H}(t^-), D > t \right) \right\} \times \mathcal{P}_0^\infty \left\{ L\left(dN_i^D(t) \middle| \mathcal{H}(t^-) \right) \right\} \end{aligned} \quad (5.6)$$

biçiminde de ifade edilebilir.

Liu ve arkadaşları, ölüm olayının varlığı altında Eşitlik (5.6)'daki olabilirlik fonksiyonunu,

$$\begin{aligned} L\left(dN_i^R(t), dN_i^D(t) \middle| \mathcal{H}_i(t^-) \right) = [Y_i(t) d\Lambda_i(t)]^{dN_i^D(t)} [1 - Y_i(t) d\Lambda_i(t)]^{1 - dN_i^D(t)} \\ \times \left\{ [Y_i(t) dR_i(t)]^{dN_i^R(t)} [1 - Y_i(t) dR_i(t)]^{1 - dN_i^R(t)} \right\}^{1 - dN_i^D(t)} \end{aligned} \quad (5.7)$$

biçimde formüle etmişlerdir. Yineleme zamanlarının sürekli olduğu durumda, $0^0 = 1$ özelliğinden yararlanarak, Eşitlik (5.7),

$$\begin{aligned} L\left(dN_i^R(t), dN_i^D(t) \middle| \mathcal{H}_i(t^-) \right) = [Y_i(t) d\Lambda_i(t)]^{dN_i^D(t)} [1 - Y_i(t) d\Lambda_i(t)]^{1 - dN_i^D(t)} \\ \times \left\{ [Y_i(t) dR_i(t)]^{dN_i^R(t)} [1 - Y_i(t) dR_i(t)] \right\}^{1 - dN_i^D(t)} \end{aligned} \quad (5.8)$$

biçiminde ifade edilebilir.

Liu ve arkadaşları ölüm sürecine ilişkin yoğunluk fonksiyonunu,

$$\lambda_i(t) = v_i^\gamma \exp(\boldsymbol{\alpha}^T \mathbf{Z}_i) \lambda_0(t) \quad (5.9)$$

ve ölüm olayı bilindiğinde yinelemeli olay sürecine ilişkin yoğunluk fonksiyonunu,

$$r_i(t) = r_0(t) v_i \exp(\boldsymbol{\beta}^T \mathbf{Z}_i) \quad (5.10)$$

biçiminde tanımlamışlardır. Eşitlik (5.9) ve (5.10)' da, $\lambda_0(t)$ ve $r_0(t)$ sırasıyla ölüm sürecine ilişkin temel yoğunluk fonksiyonunu ve ölüm olayı varlığında yinelemeli olay sürecine ilişkin temel yoğunluk fonksiyonunu göstermektedir. Eşitlik (5.9) ve (5.10)' dan yararlanarak Eşitlik (5.6)' da verilen olabilirlik fonksiyonu,

$$\begin{aligned} L(O_i, v_i | Z_i) = & \prod_{i=1}^n \left[\exp \left\{ - \int_0^\infty Y_i(t) v_i \exp(\boldsymbol{\beta}^T \mathbf{Z}_i) dR_0(t) \right\} \times \prod_k \left[v_i \exp(\boldsymbol{\beta}^T \mathbf{Z}_i) dR_0(T_{i,k}) \right]^{\delta_{ik}} \right] \times \\ & \left[\exp \left\{ - \int_0^\infty Y_i(t) v_i^\gamma \exp(\boldsymbol{\alpha}^T \mathbf{Z}_i) d\Lambda_0(t) \right\} \times \left[v_i^\gamma \exp(\boldsymbol{\alpha}^T \mathbf{Z}_i) d\Lambda_0(\tau_i) \right]^{\Delta_i} \right] \times \\ & \frac{1}{\Gamma(1/\theta) \theta^{1/\theta}} v_i^{\frac{1}{\theta}-1} e^{-v_i/\theta} \end{aligned} \quad (5.11)$$

biçiminde elde edilir. Eşitlik (5.11)' de δ_{ik} , i . bireyin $T_{i,k}$ anında yineleme yaşamasını ifade eden gösterge değişkenidir. Eşitlik (5.11)' deki olabilirlik fonksiyonunun logaritması alındığında log-olabilirlik fonksiyonu, aşağıdaki biçimde elde edilir.

$$\begin{aligned} l = \log \prod_{i=1}^n L(O_i, v_i | Z_i) &= \log \left\{ \prod_{i=1}^n L(O_i | v_i, Z_i) f_\theta(v_i) \right\} \\ &= \sum_{i=1}^n \left\{ \sum_k \delta_{ik} \left[\log v_i + \boldsymbol{\beta}^T \mathbf{Z}_i + \log dR_0(T_{i,k}) \right] - \int_0^\infty Y_i(t) v_i \exp(\boldsymbol{\beta}^T \mathbf{Z}_i) dR_0(t) \right\} \\ &+ \sum_{i=1}^n \left\{ \Delta_i \left[\gamma \log v_i + \boldsymbol{\alpha}^T \mathbf{Z}_i + \log d\Lambda_0(\tau_i) \right] - \int_0^\infty v_i^\gamma \exp(\boldsymbol{\alpha}^T \mathbf{Z}_i) d\Lambda_0(t) \right\} \quad (5.12) \\ &+ \sum_{i=1}^n \left\{ \left(\frac{1}{\theta} - 1 \right) \log v_i + \frac{v_i}{\theta} - \frac{1}{\theta} \log \theta - \log \left(\Gamma \left(\frac{1}{\theta} \right) \right) \right\}. \end{aligned}$$

Eşitlik (5.12), gözlenen veri ile gözlenemeyen rasgele etkinin (zayıflık teriminin) bileşik olabilirlik fonksiyonunu vermektedir. Eğer her bir bireye ait zayıflık değerleri gözlenebilmiş olsaydı, Eşitlik (5.12)' deki log-olabilirlik fonksiyonunu maksimum yapan değerler parametre tahminlerini verecekti. Bireylere ilişkin zayıflık değerleri gözlenemediği için Eşitlik (5.12)' de verilen olabilirlik fonksiyonunun, zayıflık parametresine göre integrali alınarak gözlenen verinin marjinal olabilirlik fonksiyonuna geçilir ve marjinal olabilirlik fonksiyonunu maksimum yapan parametre tahminleri elde edilir. Ancak, bu integrali almak her zaman kolay olmayabilir. Bu durumda, Gibbs örnekleme, Metropolis-Hastings vb. yöntemlerden yararlanır.

5.3. Önerilen Modeller

Bu tez çalışmasında, yinelemeli olaylar Poisson süreci ile modellenmiş; bilgi içeren durdurma ve bilgi içermeyen durdurma olayının varlığı altında yinelemeli olay süreci için yoğunluk fonksiyonları önerilmiştir. Bundan sonraki kesimde önerilen modeller ve parametre tahminleri için kullanılan yöntemler açıklanmıştır.

5.3.1. Önerilen Birinci Model

Önerilen bu modelde ölüm olayı, açıklayıcı değişken olarak ele alınarak bilgi içeren durdurma varsayımı altında yinelemeli olay süreci modellenmiştir.

C_i , i . bireye ait bilgi içermeyen durdurma zamanını, D_i , i . bireye ait bilgi içeren durdurma zamanını gösterebilir. $\tau_i = \min(D_i, C_i)$, i . bireye ilişkin takip süresi ve $\Delta_i = I(D_i \leq C_i)$, i . bireyin ölüm olayını deneyimleme durumunu gösteren gösterge değişkenidir. $Y_i(t) = I(\tau_i \geq t)$, i . bireyin t zamanında hayatta olup olmadığını gösteren risk sürecidir. $N_i^D(t) = I(\tau_i \leq t, \Delta_i = 1)$, i . bireyin t zamanına dek ölüm olayını yaşayıp yaşamadığını gösteren sayma sürecidir. $dN_i^R(t) = \lim_{\Delta t \rightarrow 0} N_i^R(t + \Delta t) - N_i^R(t)$, i . bireyin $(0, \Delta t]$ zaman aralığında deneyimlediği yineleme sayısını gösteren sayma süreci olmak üzere i . bireyin t zamanına dek yaşadığı yinelemeleri sayan sayma süreci $N_i^R(t) = \int_0^t Y_i(t) dN_i^R(t)$ olarak

ifade edilir. i . bireye ait geçmiş bilgisi $\mathcal{H}_i(t^-) = \{N_i^R(u), \mathbf{Z}_i(u), Y_i(u), 0 \leq u \leq t^-\}$ ve $\mathbf{Z} = (Z_1, Z_2, \dots, Z_p)$ zamandan bağımsız açıklayıcı değişken vektörüdür. O_i , i . bireye ait gözlenen veriyi; ν_i , ise i . bireye ilişkin zayıflık terimini göstermektedir.

Önerilen birinci model için yapılan varsayımlar şunlardır.

1. Yinelemeli olay, ölüm ve durdurma süreçlerinin sürekli dağılıma sahiptir;, ölüm, durdurma ve yinelemeli olaylar aynı anda meydana gelemez. $[t, t + \Delta t)$ çok küçük zaman aralığında önce ölüm olayı gerçekleşir.

2. Yinelemeli olay sürecine ilişkin yoğunluk fonksiyonu,

$$P(dN^R(t) = 1 | \mathcal{H}_i(t^-), D \geq t) = Y_i(t) dR_i(t) \equiv Y_i(t) r_i(t) dt$$

$$dR_i(t) = P(dN_i^{R*}(t) = 1 | Z_i, \nu_i, D_i \geq t)$$

biçimindedir.

3. i . bireye ait sayma süreci $N_i(t)$, $\lambda_i(\cdot)$ yoğunluğuyla homojen Poisson sürecine sahiptir.

4. Durdurma, zayıflık terimi ν_i 'ye ilişkin bilgi içermemektedir.

5. Zayıflık teriminin dağılımı, Eşitlik (5.3)' te verildiği gibi $\nu_i \sim \text{Gamma}(\frac{1}{\theta}, \theta)$ olarak alınmıştır.

Yukarıda verilen varsayımlar kullanılarak ölüm olayı varlığı altında yinelemeli olay sürecine ilişkin yoğunluk fonksiyonu,

$$\lambda_i(t | \mathcal{H}_i(t^-)) = r_0(t) \nu_i \exp(\boldsymbol{\beta}_1^T \mathbf{Z}_{1i} + \boldsymbol{\beta}_2^T \mathbf{Z}_{2i}) \quad (5.13)$$

biçiminde tanımlanmıştır. Eşitlik (5.13)' te $r_0(t)$, ölüm olayı varlığı altında yinelemeli olay sürecine ilişkin temel yoğunluk fonksiyonunu, \mathbf{Z}_{1i} , zamandan bağımsız açıklayıcı değişken vektörünü, \mathbf{Z}_{2i} , i . bireyin ölüm olayını deneyimleyip deneyimlemediği bilgisini gösteren açıklayıcı değişken değerini, β_1 ve β_2 sırasıyla \mathbf{Z}_{1i} ve \mathbf{Z}_{2i} açıklayıcı değişkenlerine karşılık gelen parametreleri göstermektedir.

Eşitlik (5.6)' dan yararlanarak

$$L(dN_i^R(t), dN_i^D(t) | \mathcal{H}_i(t^-)) = [Y_i(t) d\Lambda_i(t)]^{dN_i^R(t)} [1 - Y_i(t) d\Lambda_i(t)]^{1 - dN_i^R(t)}$$

yazılabilir. Eşitlik (5.13)' te verilen yoğunluk fonksiyonu kullanılarak, i . bireye ait olabilirlik fonksiyonu,

$$\begin{aligned} L_i(O_i, v_i | z_i) &= L_i(O_i | v_i, \mathbf{Z}_i) f_\theta(v_i) \\ &= \exp\left\{-\int_0^\infty Y_i(t) v_i \exp(\beta_1^T \mathbf{Z}_{1i} + \beta_2 \mathbf{Z}_{2i}) dR_0(t)\right\} \\ &\quad \times \prod_k \left\{v_i \exp(\beta_1^T \mathbf{Z}_{1i} + \beta_2 \mathbf{Z}_{2i}) dR_0(t)\right\}^{\delta_{ik}} \\ &\quad \times \frac{1}{\Gamma(1/\theta) \theta^{1/\theta}} v_i^{\frac{1}{\theta}-1} e^{-v_i/\theta} \end{aligned} \quad (5.14)$$

biçiminde yazılabilir. Burada δ_{ik} , i . bireyin $T_{i,k}$ anında yineleme yaşamasını ifade eden gösterge değişkenidir. Eşitlik (5.13)' te verilen yoğunluk fonksiyonunda, temel yoğunluk fonksiyonu $r_0(t) = \lambda$ ve β_1 zamandan bağımsız açıklayıcı değişkenine karşılık gelen parametre olarak alınırsa, Eşitlik (5.14)' te verilen olabilirlik fonksiyonu aşağıdaki biçimde yazılabilir.

$$L_i(O_i, v_i | \mathbf{Z}_i) = \left(\prod_{k=1}^{N_i(\tau_i)} \lambda v_i \exp(\beta_1 \mathbf{Z}_{1i} + \beta_2 \mathbf{Z}_{2i}) \right) \exp \left(- \int_0^{\tau_i} \lambda v_i \exp(\beta_1 \mathbf{Z}_{1i} + \beta_2 \mathbf{Z}_{2i}) du \right) \\ \times \frac{1}{\Gamma(1/\theta) \theta^{1/\theta}} v_i^{\frac{1}{\theta}-1} e^{-v_i/\theta}. \quad (5.15)$$

Eşitlik (5.15)' te $1/\theta = \theta^*$ olarak alınırsa,

$$L_i(O_i, v_i | \mathbf{Z}_i) = [\lambda \exp(\beta_1 \mathbf{Z}_{1i} + \beta_2 \mathbf{Z}_{2i})]^{N_i(\tau_i)} \exp[-\lambda v_i \tau_i \exp(\beta_1 \mathbf{Z}_{1i} + \beta_2 \mathbf{Z}_{2i})] \\ \times \frac{(\theta^*)^{\theta^*}}{\Gamma(\theta^*)} v_i^{\theta^*-1} e^{-\theta^* v_i} \quad (5.16)$$

biçiminde yeniden ifade edilebilir. Eşitlik (5.16)' da verilen olabilirlik fonksiyonu, yinelemeli olay sürecinin ve gözlenemeyen zayıflık parametresinin bileşik olabilirlik fonksiyonudur. Bu yüzden zayıflık parametresine göre integrali alınarak n birey için marjinal olabilirlik fonksiyonu;

$$L_{\text{marg}}(O | v, \mathbf{Z}) = \prod_{i=1}^n L_{i,\text{marg}} = \prod_{i=1}^n \int_0^{\infty} L_i(O_i | v_i, \mathbf{Z}_i) f_{\theta}(v_i) dv_i \\ = \prod_{i=1}^n \frac{(\theta^*)^{\theta^*} \times [\lambda \exp(\beta_1 \mathbf{Z}_{1i} + \beta_2 \mathbf{Z}_{2i})]^{N_i(\tau_i)}}{\Gamma(\theta^*)} \\ \times \frac{\Gamma(N_i(\tau_i) + \theta^*)}{[\lambda \tau_i \exp(\beta_1 \mathbf{Z}_{1i} + \beta_2 \mathbf{Z}_{2i}) + \theta^*]^{N_i(\tau_i) + \theta^*}} \quad (5.17)$$

biçiminde elde edilir. Eşitlik (5.17)' deki $\Gamma(\cdot)$ fonksiyonu için,

$$\Gamma(N_i(\tau_i) + \theta^*) = \prod_{j=1}^{N_i(\tau_i)} \{N_i(\tau_i) - j + \theta^*\} \Gamma(\theta^*)$$

eşitliğinden yararlanılarak marjinal olabilirlik fonksiyonu,

$$L_{\text{marg}}(O | \nu, \mathbf{Z}) = \prod_{i=1}^n \left\{ \prod_{j=1}^{N_i(\tau_i)} \{N_i(\tau_i) - j + \theta^*\} \right\} \frac{(\theta^*)^{\theta^*} [\lambda \exp(\beta_1 \mathbf{Z}_{1i} + \beta_2 \mathbf{Z}_{2i})]^{N_i(\tau_i)}}{[\lambda \tau_i \exp(\beta_1 \mathbf{Z}_{1i} + \beta_2 \mathbf{Z}_{2i}) + \theta^*]^{N_i(\tau_i) + \theta^*}} \quad (5.18)$$

ve log olabilirlik fonksiyonu,

$$\begin{aligned} \log L_{\text{marg}}(O | \nu, \mathbf{Z}) &= \sum_{i=1}^n \left\{ \sum_{j=1}^{N_i(\tau_i)} \log(N_i(\tau_i) - j + \theta^*) \right\} + N_i(\tau_i) \{ \log(\lambda) + \beta_1 \mathbf{Z}_{1i} + \beta_2 \mathbf{Z}_{2i} \} \\ &\quad + \theta^* \log(\theta^*) - \{N_i(\tau_i) + \theta^*\} \log(\lambda \tau_i \exp(\beta_1 \mathbf{Z}_{1i} + \beta_2 \mathbf{Z}_{2i}) + \theta^*) \end{aligned} \quad (5.19)$$

biçiminde elde edilir. Eşitlik (5.19)' da verilen log-olabilirlik fonksiyonunun bilinmeyen λ , β_1 , β_2 ve θ^* parametrelerine göre türevleri aşağıda verilmiştir:

$$\frac{\partial L_{\text{marg}}(O | \nu, \mathbf{Z})}{\partial \lambda} = \sum_{i=1}^n \frac{N_i(\tau_i)}{\lambda} - \frac{(N_i(\tau_i) + \theta^*)(\tau_i \exp(\beta_1 \mathbf{Z}_{1i} + \beta_2 \mathbf{Z}_{2i}))}{(\lambda \tau_i \exp(\beta_1 \mathbf{Z}_{1i} + \beta_2 \mathbf{Z}_{2i}) + \theta^*)},$$

$$\frac{\partial L_{\text{marg}}(O | \nu, \mathbf{Z})}{\partial \beta_1} = \sum_{i=1}^n N_i(\tau_i) \mathbf{Z}_{1i} - \frac{(N_i(\tau_i) + \theta^*)(\lambda \tau_i \mathbf{Z}_{1i} \exp(\beta_1 \mathbf{Z}_{1i} + \beta_2 \mathbf{Z}_{2i}))}{(\lambda \tau_i \exp(\beta_1 \mathbf{Z}_{1i} + \beta_2 \mathbf{Z}_{2i}) + \theta^*)},$$

$$\frac{\partial L_{\text{marg}}(O | \nu, \mathbf{Z})}{\partial \beta_2} = \sum_{i=1}^n N_i(\tau_i) \mathbf{Z}_{2i} - \frac{(N_i(\tau_i) + \theta^*)(\lambda \tau_i \mathbf{Z}_{2i} \exp(\beta_1 \mathbf{Z}_{1i} + \beta_2 \mathbf{Z}_{2i}))}{(\lambda \tau_i \exp(\beta_1 \mathbf{Z}_{1i} + \beta_2 \mathbf{Z}_{2i}) + \theta^*)}$$

$$\begin{aligned} \frac{\partial L_{\text{marg}}(O | \nu, \mathbf{Z})}{\partial \theta^*} &= \sum_{i=1}^n \left\{ \sum_{k=1}^{N_i(\tau_i)} \frac{1}{N_i(\tau_i) - j + \theta^*} \right\} + \log(\theta^*) + 1 - \\ &\quad \left\{ \log(\lambda \tau_i \exp(\beta_1 \mathbf{Z}_{1i} + \beta_2 \mathbf{Z}_{2i}) + \theta^*) + \frac{(N_i(\tau_i) + \theta^*)}{(\lambda \tau_i \exp(\beta_1 \mathbf{Z}_{1i} + \beta_2 \mathbf{Z}_{2i}) + \theta^*)} \right\}. \end{aligned}$$

Yukarıda verilen skor fonksiyonlarını sıfır yapan değerler parametre kestirimlerini vermektedir. Bu skor fonksiyonlarının eşanlı çözümü ancak denklemler doğrusal ise olasıdır. Bu nedenle doğrusal olmayan denklem sisteminin çözümü için iteratif (sayısal) yöntemlere başvurulur. Yukarıda yer alan denklem sistemlerinin çözümü için Newton-Raphson algoritması kullanılmıştır.

5.3.2. Önerilen İkinci Model

Önerilen ikinci modelde, Liu ve arkadaşları (2004) tarafından önerilen yöntemden yararlanılmıştır. Ölüm olayı ve durdurma olayının varlığında yinelemeli olay sürecine ilişkin farklı yoğunluk fonksiyonları tanımlanmış; yinelemeli olayların Poisson sürecine uygun olarak çıktığı varsayımı altında yinelemeli olaylar modellenmiştir.

Liu ve arkadaşları, yinelemeli olay süreci ile bağımlı durdurma sürecini modellerken bilgi içermeyen durdurma sürecinin bilgisini modele katmamışlardır. Önerilen ikinci modelde Liu ve arkadaşlarının modellerinden farklı olarak bilgi içermeyen durdurma olayına ilişkin de bir yoğunluk fonksiyonu önerilerek bilgi içermeyen durdurma bilgisi de modele katılmıştır.

Önerilen ikinci model için yapılan tanımlamalar aşağıdadır:

C_i , i . bireye ait bilgi içermeyen durdurma zamanını, D_i , i . bireye ait bilgi içeren durdurma zamanını gösterebilir. $\tau_i = \min(D_i, C_i)$, i . bireye ilişkin takip süresi ve $\Delta_i = I(D_i \leq C_i)$, i . bireyin ölüm olayını deneyimleme durumunu gösteren gösterge değişkenidir. $Y_i(t) = I(\tau_i \geq t)$, i . bireyin t zamanında hayatta olup olmadığını gösteren risk sürecidir. $N_i^D(t) = I(\tau_i \leq t, \Delta_i = 1)$, i . bireyin t zamanına dek ölüm olayını yaşayıp yaşamadığını gösteren sayma sürecidir. $dN_i^{RD}(t) = \lim_{\Delta t \rightarrow 0} N_i^{RD}(t + \Delta t^-) - N_i^{RD}(t^-)$, ölüm olayını deneyimleyen i . bireyin $(0, \Delta t]$ zaman aralığında deneyimlediği yineleme sayısını gösteren sayma süreci olmak üzere i . bireyin t zamanına dek yaşadığı yinelemeleri sayan sayma süreci $N_i^{RD}(t) = \int_0^t Y_i(t) dN_i^{RD}(t)$ olarak ifade edilir. $dN_i^{RC}(t) = \lim_{\Delta t \rightarrow 0} N_i^{RC}(t + \Delta t^-) - N_i^{RC}(t^-)$,

durdurma yaşayan i . bireyin $(0, \Delta t]$ zaman aralığında deneyimlediği yineleme sayısını gösteren sayma süreci olmak üzere i . bireyin t zamanına dek yaşadığı yinelemeleri sayan sayma süreci $N_i^{R|C}(t) = \int_0^t Y_i(u) dN_i^{R|C}(u)$ olarak ifade edilir. i . bireye ait geçmiş bilgisi $\mathcal{H}_i(t) = \{N_i^R(u), N_i^D(u), \mathbf{Z}_i(u), Y_i(u), 0 \leq u \leq t\}$ ve $\mathbf{Z} = (Z_1, Z_2, \dots, Z_p)$ zamandan bağımsız açıklayıcı değişken vektörüdür. O_i , i . bireye ait gözlenen veriyi; v_i , ise i . bireye ait zayıflık terimini göstermektedir.

Önerilen ikinci model için yapılan varsayımlar şunlardır:

1. Yinelemeli olay, ölüm ve durdurma süreçlerinin sürekli dağılıma sahiptir; ölüm, durdurma ve yinelemeli olaylar aynı anda meydana gelemez. $[t, t + \Delta t)$ çok küçük zaman aralığında önce ölüm olayı gerçekleşir.
2. Ölüm sürecine ilişkin yoğunluk fonksiyonu,

$$\begin{aligned} P(dN^D(t) = 1 | \mathcal{H}_i(t^-)) &= Y_i(t) d\Lambda_i(t) \equiv Y_i(t) \lambda_i(t) dt \\ d\Lambda_i(t) &= P(dN^{D_i}(t) = 1 | \mathbf{Z}_i, v_i, D_i \geq t) \end{aligned} \quad (5.20)$$

biçimindedir.

3. Ölüm olayı varlığı altında yinelemeli olay sürecine ilişkin yoğunluk fonksiyonu,

$$\begin{aligned} P(dN^{R|D}(t) = 1 | \mathcal{H}_i(t^-), D_i \geq t) &= Y_i(t) dR_i^{R|D}(t) \equiv Y_i(t) r_i^{R|D}(t) dt \\ dR_i^{R|D}(t) &= P(dN_i^{R|D}(t) = 1 | \mathbf{Z}_i, v_i, D_i \geq t) \end{aligned} \quad (5.21)$$

biçimindedir.

4. Durdurma varlığı altında yinelemeli olay sürecine ilişkin yoğunluk fonksiyonu,

$$\begin{aligned} P(dN^{RiC}(t) = 1 | \mathcal{H}_i(t^-), D \geq t) &= Y_i(t) dR_i^{RiC}(t) \equiv Y_i(t) r_i^{RiC}(t) dt \\ dR_i^{RiC}(t) &= P(dN_i^{RiC}(t) = 1 | \mathbf{Z}_i, \nu_i, C_i \geq t) \end{aligned} \quad (5.22)$$

biçimindedir.

5. Ölüm olayını yaşayan bireylere ilişkin sayma süreci $N_i^{RiD}(t)$, $\lambda_i^{RiD}(\cdot)$ yoğunluğuyla homojen Poisson sürecine sahiptir.
6. Durdurma olayını yaşayan bireylere ilişkin sayma süreci $N_i^{RiC}(t)$, $\lambda_i^{RiC}(\cdot)$ yoğunluğuyla homojen Poisson sürecine sahiptir.
7. Durdurma, zayıflık terimi ν_i 'ye ilişkin bilgi içermemektedir.
8. Zayıflık terimi ν , verildiğinde yinelemeli olay süreci ile ölüm süreci bağımsızdır.
9. Zayıflık teriminin dağılımı, Eşitlik (5.3)' te verildiği gibi $\nu_i \sim \text{Gamma}(\frac{1}{\theta}, \theta)$ olarak alınmıştır.

Önerilen ikinci modelde yukarıda verilen varsayımlar kullanılarak ölüm olayına ilişkin yoğunluk fonksiyonu,

$$\lambda_i^D(t) = \lambda_0^D(t) \nu_i \exp(\boldsymbol{\beta}^T \mathbf{Z}_i) \quad (5.23)$$

ölüm olayının varlığı altında yinelemeli olay sürecine ilişkin yoğunluk fonksiyonu,

$$\lambda^{RiD}(t) = r_0^{RiD}(t) \nu_i \left(\exp(\boldsymbol{\beta}^T \mathbf{Z}_i) \right)^\alpha \quad (5.24)$$

ve durdurma olayının varlığı altında yinelemeli olay sürecine ilişkin yoğunluk fonksiyonu,

$$\lambda^{R^1 C}(t) = r_0^{R^1 C}(t) v_i \exp(\boldsymbol{\beta}^T \mathbf{Z}_i) \quad (5.25)$$

olarak alınmıştır. Eşitlik (5.23), (5.24) ve (5.25)' te $\boldsymbol{\beta}$, $p \times 1$ boyutlu parametre vektörü ve $\mathbf{Z} = (Z_1, Z_2, \dots, Z_p)$ zamandan bağımsız açıklayıcı değişken vektörüdür. λ_0^D , $r_0^{R^1 D}(t)$ ve $r_0^{R^1 C}(t)$ sırasıyla, ölüm sürecine ilişkin temel yoğunluk fonksiyonunu, ölüm olayı varlığı altında yinelemeli olay sürecine ilişkin temel yoğunluk fonksiyonunu ve durdurma varlığında yinelemeli olay sürecine ilişkin temel yoğunluk fonksiyonunu göstermektedir.

Eşitlik (5.6)' dan yararlanarak

$$\begin{aligned} L(dN_i^{R^1 D}(t), dN_i^{R^1 C}(t), dN_i^D(t) | \mathcal{H}_i(t^-)) = \\ \prod_{i=1}^n \left\{ \left[v_i \exp(\boldsymbol{\beta}^T \mathbf{Z}_i) d\Lambda_0^D(t) \right]^{dN_i^D(t)} \times \left[1 - v_i \exp(\boldsymbol{\beta}^T \mathbf{Z}_i) d\Lambda_0^D(t) \right]^{1-dN_i^D(t)} \right\}^{\Delta_i} \times \\ \left\{ \left[v_i (\exp(\boldsymbol{\beta}^T \mathbf{Z}_i))^\alpha dR_0^{R^1 D}(t) \right]^{dN_i^{R^1 D}(t)} \times \left[1 - v_i (\exp(\boldsymbol{\beta}^T \mathbf{Z}_i))^\alpha dR_0^{R^1 D}(t) \right]^{1-dN_i^{R^1 D}(t)} \right\}^{\Delta_i} \times \\ \left\{ \left[v_i \exp(\boldsymbol{\beta}^T \mathbf{Z}_i) dR_0^{R^1 C}(t) \right]^{dN_i^{R^1 C}(t)} \times \left[1 - v_i \exp(\boldsymbol{\beta}^T \mathbf{Z}_i) dR_0^{R^1 C}(t) \right]^{1-dN_i^{R^1 C}(t)} \right\} \end{aligned}$$

yazılabilir.

Yukarıdaki eşitlikte $1/\theta = \theta^*$ olarak alındığında, Eşitlik (5.23), (5.24) ve (5.25)' te verilen yoğunluk fonksiyonları kullanılarak olabilirlik fonksiyonu,

$$\begin{aligned}
L(O, v | \mathbf{Z}) &= \prod_{i=1}^n L_i(O_i | v_i, \mathbf{Z}_i) f_{\theta}(v_i) \\
&= \prod_{i=1}^n \left\{ \left[v_i \exp(\boldsymbol{\beta}^T \mathbf{Z}_i) d\Lambda_0^D(\tau_i) \right] \exp \left[- \int_0^{\infty} Y_i(t) v_i \exp(\boldsymbol{\beta}^T \mathbf{Z}_i) d\Lambda_0^D(t) \right] \right\}^{\Delta_i} \times \\
&\quad \left\{ \left[\prod_{k=1}^{N_i(\tau_i)} \left[v_i (\exp(\boldsymbol{\beta}^T \mathbf{Z}_i))^{\alpha} dR_0^{R|D}(T_{i,k}) \right]^{\delta_{ik}} \right] \times \exp \left[- \int_0^{\infty} Y_i(t) v_i (\exp(\boldsymbol{\beta}^T \mathbf{Z}_i))^{\alpha} dR_0^{R|D}(t) \right] \right\}^{\Delta_i} \times \\
&\quad \left\{ \left[\prod_{k=1}^{N_i(\tau_i)} \left[v_i (\exp(\boldsymbol{\beta}^T \mathbf{Z}_i)) dR_0^{R|C}(T_{i,k}) \right]^{\delta_{ik}} \right] \times \exp \left[- \int_0^{\infty} Y_i(t) v_i (\exp(\boldsymbol{\beta}^T \mathbf{Z}_i))^{\alpha} dR_0^{R|C}(t) \right] \right\}^{1-\Delta_i} \times \\
&\quad \frac{(\theta^*)^{\theta^*}}{\Gamma(\theta^*)} v_i^{\theta^*-1} e^{-\theta^* v_i}
\end{aligned} \tag{5.26}$$

biçiminde elde edilir. Eşitlik (5.26)' daki, δ_{ik} , i . bireyin $T_{i,k}$ anında yineleme yaşamasını ifade eden gösterge değişkenidir.

Eşitlik (5.26)' daki olabilirlik fonksiyonunun logaritması alındığında, log-olabilirlik fonksiyonu aşağıdaki biçimde elde edilir:

$$\begin{aligned}
l &= \log \{ L(O, v | \mathbf{Z}) \} = \log \left\{ \prod_{i=1}^n L_i(O_i | v_i, \mathbf{Z}_i) f_{\theta}(v_i) \right\} \\
&= \sum_{i=1}^n \Delta_i \left\{ \left[\log v_i + \boldsymbol{\beta}^T \mathbf{Z}_i + \log(d\Lambda_0^D(\tau_i)) \right] - \int_0^{\infty} Y_i(t) v_i \exp(\boldsymbol{\beta}^T \mathbf{Z}_i) d\Lambda_0^D(t) \right\} \\
&\quad + \sum_{i=1}^n \Delta_i \left\{ \sum_k \delta_{ik} \left[\log v_i + \alpha \boldsymbol{\beta}^T \mathbf{Z}_i + \log(dR_0^{R|D}(T_{i,k})) \right] - \int_0^{\infty} Y_i(t) v_i (\exp(\boldsymbol{\beta}^T \mathbf{Z}_i))^{\alpha} dR_0^{R|D}(t) \right\} \\
&\quad + \sum_{i=1}^n (1-\Delta_i) \left\{ \sum_k \delta_{ik} \left[\log v_i + \boldsymbol{\beta}^T \mathbf{Z}_i + \log(dR_0^{R|C}(T_{i,k})) \right] - \int_0^{\infty} Y_i(t) v_i \exp(\boldsymbol{\beta}^T \mathbf{Z}_i) dR_0^{R|C}(t) \right\} \\
&\quad + \sum_{i=1}^n \left\{ \theta^* \log(\theta^*) - \log(\Gamma(\theta^*)) + (\theta^* - 1) \log v_i - \theta^* v_i \right\}.
\end{aligned} \tag{5.27}$$

Önerilen ikinci modelin özel bir durumu olarak, ölüm sürecine ilişkin temel yoğunluk fonksiyonu μ , ölüm olayı varlığında yinelemeli olay sürecine ilişkin temel yoğunluk fonksiyonu λ^α ve durdurma olayı varlığında yinelemeli olay sürecine ilişkin temel yoğunluk fonksiyonu λ olarak alındığı varsayalım. Modelde tek bir açıklayıcı değişken kullanılsın ve bu açıklayıcı değişken zamandan bağımsız olsun. Bu durumda Eşitlik (5.23), (5.24) ve (5.25)' te verilen yoğunluk fonksiyonları sırasıyla aşağıdaki biçimlerde tanımlanmıştır:

$$\lambda_i^D(t) = v_i \mu \exp(\beta \mathbf{Z}_i) \quad (5.28)$$

$$\lambda_i^{RD}(t) = v_i (\lambda \exp(\beta \mathbf{Z}_i))^\alpha \quad (5.29)$$

$$\lambda_i^{RC}(t) = v_i \lambda \exp(\beta \mathbf{Z}_i) \quad (5.30)$$

Eşitlik (5.28), (5.29) ve (5.30)' da verilen yoğunluk fonksiyonlarından yararlanılarak Eşitlik (5.27)' de verilen log-olabilirlik fonksiyonu,

$$\begin{aligned} l &= \log \{L(O, v | \mathbf{Z})\} = \log \left\{ \prod_{i=1}^n L_i(O_i | v_i, \mathbf{Z}_i) f_\theta(v_i) \right\} \\ &= \sum_{i=1}^n \Delta_i \{ \log v_i + \beta \mathbf{Z}_i + \log \mu - v_i \tau_i \mu \exp(\beta \mathbf{Z}_i) \} \\ &+ \sum_{i=1}^n \Delta_i \{ N_i^{RD}(\tau_i) [\log v_i + \alpha (\beta \mathbf{Z}_i + \log \lambda)] - v_i \tau_i (\lambda \exp(\beta \mathbf{Z}_i))^\alpha \} \quad (5.31) \\ &+ \sum_{i=1}^n (1 - \Delta_i) \{ N_i^{RC}(\tau_i) [\log v_i + \beta \mathbf{Z}_i + \log \lambda] - v_i \tau_i \lambda \exp(\beta \mathbf{Z}_i) \} \\ &+ \sum_{i=1}^n \theta^* \log(\theta^*) - \log(\Gamma(\theta^*)) + (\theta^* - 1) \log v_i - \theta^* v_i \end{aligned}$$

biçimde elde edilir. Eşitlik (5.31)' de verilen log-olabilirlik fonksiyonu gözlenen veri ile gözlenemeyen rasgele etkinin (zayıflık parametresi) bileşik olabilirlik fonksiyonunu

vermektedir. Eđer her bir bireye ait zayıflık deęerleri gözlenebilmiş olsaydı, Eşitlik (5.31)'deki log-olabilirlik fonksiyonunu maksimum yapan deęerler parametre tahminlerini verecekti. Liu ve arkadaşları (2004) tarafından önerilen modelin çözümlemesinde de ifade edildięi gibi bireylere ilişkin zayıflık deęerleri gözlenemedięi için Eşitlik (5.31)'de verilen olabilirlik fonksiyonunun, zayıflık terimine göre integrali alınırsa gözlenen verinin marjinal olabilirlik fonksiyonu aşığıdaki gibi bulunur:

$$\begin{aligned} L_{\text{marg}}(O|\nu, \mathbf{Z}) &= \int_0^{\infty} L(O|\nu, \mathbf{Z}) f_{\theta}(\nu) d\nu \\ &= E_{\nu}(L(O|\nu, \mathbf{Z})). \end{aligned} \quad (5.32)$$

Ancak bu integrali almak her zaman kolay olmayabilir. Dolayısıyla bu integrali çözümleyebilmek için, Kesim 5.2'de belirtildięi gibi alternatif yöntemlere başvurulur. Eşitlik (5.31)'de görüldüğü gibi marjinal olabilirlik fonksiyonuna geçiş zayıflık parametresine göre $L(O|\nu, \mathbf{Z})$ 'nin beklenen deęerini almaya eş deęerdir. Ancak zayıflık katsayıları gözlenemedięi için bu beklenen deęer hesaplanamaz. Bu durum, bir kayıp gözlem problemi olarak düşünölüp çözümlenmelidir. Bu yüzden Eşitlik (5.30)'da verilen beklenen deęer problemi için EM algoritması kullanılmıştır. Bu nedenle EM algoritması aşığıda kısaca açıklanmıştır:

EM algoritması, iki adımdan oluşmaktadır. Bunlar, E ve M adımlarıdır. E adımında gözlenemeyen verinin tahmini yapılır. M adımında ise gözlenemeyen veriye ilişkin elde edilen tahmin deęerleri olabilirlik fonksiyonunda yerine konularak bu fonksiyonu maksimum yapan parametre deęerleri kestirilir. EM algoritmasının adımları aşığıda verilmiştir:

Parametrelere başlangıç deęeri verilerek başlanır. Parametre vektörü γ ile gösterilmek üzere $\gamma^{(0)}$, başlangıçtaki parametre deęerlerini içeren vektörü göstermektedir.

E-Adımı:

1. Başlangıçta verilen parametre değerleri $\gamma^{(0)}$ ve i . bireye ilişkin gözlenen veri değerleri O_i , kullanılarak v_i 'nin koşullu olasılık yoğunluk fonksiyonundan $f(v_i | O_i, \gamma^{(0)})$, $v_i^{(1)}$ değeri bulunur.
2. Birinci adımda bulunan $v_i^{(1)}$ değerleri olabilirlik fonksiyonunda yerine yazılır.

M-Adımı:

E-adımının ikinci basamağında elde edilen olabilirlik fonksiyonu maksimum yapılarak yeni parametre vektörü $\gamma^{(1)}$ hesaplanır.

M-adımında hesaplanan yeni parametre vektörü $\gamma^{(1)}$ kullanılarak tekrar E-adımına dönülerek yeni $v_i^{(2)}$ değerleri hesaplanır. Hesaplanan $v_i^{(2)}$ değerleri olabilirlik fonksiyonunda yerine yazılarak maksimum yapan parametre vektörü $\gamma^{(2)}$ elde edilir. EM algoritmasının bu döngüsü, ardışık iterasyonlarda elde edilen parametre vektörleri arasındaki fark $\gamma^{(i)} - \gamma^{(i-1)} < \varepsilon$ olana kadar devam ettirilir. Burada ε , araştırmacının belirleyeceği hoşgörü miktarıdır.

EM algoritmasının E-adımında gözlenemeyen değişken değerlerinin beklenen değerini hesaplamak için kullanılan $f(v_i | O_i, \gamma)$ koşullu olasılık yoğunluk fonksiyonu için kapalı bir form mevcut değildir. Bu yüzden EM algoritmasının E-adımında v_i değerlerinin hesaplanmasında Metropolis-Hastings algoritmasından yararlanıldığı için, Metropolis-Hastings algoritmasının adımları aşağıda verilmiştir:

1. v_i değerlerinin hesaplanması için öneri dağılımı belirlenir. Öneri dağılımı olarak, Eşitlik (5.3)' te verildiği gibi $v_i \sim \text{Gamma}(\theta, \frac{1}{\theta})$ olarak alınmıştır.

EM algoritmasının başında belirlenen parametre vektörü $\gamma^{(0)}$ kullanılarak rasgele veri türetilir ve üretilen bu rasgele veriler, aday noktalar v_i^{yeni} olarak belirlenir.

2. Aday noktaların kabul edilip edilmeyeceğini belirlemek için ölüm süreci, ölüm olayı varlığında yinelemeli olay süreci ve durdurma varlığında yinelemeli olay sürecine ilişkin Eşitlik (5.28), (5.29) ve (5.30)' da önerilen yoğunluk fonksiyonları kullanılarak yazılan $f(O_i | v_i, \gamma)$ koşullu olasılık yoğunluk fonksiyonu değeri, aday ve eski nokta için hesaplanır:

$$f(O_i | v_i, \gamma) = \prod_{i=1}^n \left\{ [v_i \mu \exp(\beta \mathbf{Z}_i)]^{\Delta N_i^{D(t)}} \times [1 - v_i \mu \exp(\beta \mathbf{Z}_i)]^{1 - \Delta N_i^{D(t)}} \right\}^{\Delta_i} \times \\ \left\{ [v_i (\lambda \exp(\beta \mathbf{Z}_i))^\alpha]^{\Delta N_i^{R1D(t)}} \times [1 - v_i (\lambda \exp(\beta \mathbf{Z}_i))^\alpha]^{1 - \Delta N_i^{R1D(t)}} \right\}^{\Delta_i} \times \\ \left\{ [v_i \lambda \exp(\beta \mathbf{Z}_i)]^{\Delta N_i^{R1C(t)}} \times [1 - v_i \lambda \exp(\beta \mathbf{Z}_i)]^{1 - \Delta N_i^{R1C(t)}} \right\}$$

3. Daha sonra U~tekbüçimli (0,1) dağılımından rasgele veri türetilir ve aday noktanın kabul edilip edilmemesine karar verilir.

$$v_i^{(m+1)} = \begin{cases} v_i^{yeni}, & \text{eğer } u \leq \min(1, \frac{f(O_i | v_i^{yeni})}{f(O_i | v_i^{(m)})}) \\ v_i^{(m)}, & \text{öteki durumda} \end{cases}$$

Kabul edilen ya da edilmeyen noktalar her i . birey için zincirin bir adımını oluşturur. Sonra adım 1'e tekrar geri dönülerek işlemler tekrar edilir. Bu şekilde m sayıda zincir oluşturulur. Her i . birey için oluşturulan zincir değerlerinin ortalaması gözlenemeyen değişken değerlerine karşılık gelir ve EM algoritmasının E-adımı tamamlanmış olur. Bu değerler

olabilirlik fonksiyonunda yerine yazılarak maksimum yapan parametre değerleri elde edilir.

Eşitlik (5.28), (5.29) ve (5.30)' da önerilen yoğunluk fonksiyonlarından elde edilen olabilirlik fonksiyonunu çözümlmek için yukarıda açıklanan EM ve Metropolis-Hastings algoritmaları kullanılarak gözlenemeyen zayıflık değişkenine ilişkin değerler elde edildikten sonra olabilirlik fonksiyonunun maksimizasyonu kısmına geçilir. $\theta^* = 1/\theta$ Eşitlik (5.31)' de verilen log-olabilirlik fonksiyonunun parametrelere göre türevlerinden oluşan skor fonksiyonları aşağıda verilmiştir:

$$\frac{\partial \log(O, v | \mathbf{Z})}{\partial \lambda} = \sum_{i=1}^n \Delta_i \left\{ \frac{N_i^{RD}(\tau_i) \alpha}{\lambda} - \frac{E(v_i) \tau_i \alpha}{\lambda} [\lambda \exp(\beta \mathbf{Z}_i)]^\alpha \right\} \\ + (1 - \Delta_i) \left\{ \frac{N_i^{RC}(\tau_i)}{\lambda} - E(v_i) \tau_i \exp(\beta \mathbf{Z}_i) \right\}$$

$$\frac{\partial \log(O, v | \mathbf{Z})}{\partial \beta} = \sum_{i=1}^n \Delta_i \left\{ \mathbf{Z}_i - E(v_i) \tau_i \mu \mathbf{Z}_i \exp(\beta \mathbf{Z}_i) + N_i^{RD}(\tau_i) \alpha \mathbf{Z}_i \right. \\ \left. - E(v_i) \tau_i \alpha \mathbf{Z}_i [\lambda \exp(\beta \mathbf{Z}_i)]^\alpha \right\} + (1 - \Delta_i) \left\{ N_i^{RC}(\tau_i) \mathbf{Z}_i - E(v_i) \tau_i \lambda \mathbf{Z}_i \exp(\beta \mathbf{Z}_i) \right\}$$

$$\frac{\partial \log(O, v | \mathbf{Z})}{\partial \alpha} = \sum_{i=1}^n \Delta_i \left\{ N_i^{RD}(\tau_i) \{ \log \lambda + \beta \mathbf{Z}_i \} - \right. \\ \left. E(v_i) \tau_i \left\{ \log(\lambda \exp(\beta \mathbf{Z}_i)) \times (\lambda \exp(\beta \mathbf{Z}_i))^\alpha \right\} \right\}$$

$$\frac{\partial \log(O, v | \mathbf{Z})}{\partial \mu} = \sum_{i=1}^n \Delta_i \left\{ \frac{1}{\mu} - E(v_i) \tau_i \exp(\beta \mathbf{Z}_i) \right\}$$

$$\frac{\partial \log(O, v | \mathbf{Z})}{\partial \theta^*} = \sum_{i=1}^n \Delta_i \left\{ \log \theta^* + 1 - \psi(\theta^*) + E(\log v_i) - E(v_i) \right\}.$$

Burada $\psi(\theta)$ digamma fonksiyonudur. Yukarıda verilen skor fonksiyonlarını sıfır yapan değerler parametre kestirimlerini vermektedir. Bu skor fonksiyonları eşanlı şekilde çözümlenmelidir. Doğrusal olmayan denklem sistemi olduğu için iteratif yöntemlerden yararlanır. Bu nedenle bu tez çalışmasında yukarıda yer alan denklem sistemleri Newton-Raphson algoritması kullanılarak çözümlenmiştir.

6. UYGULAMA VE BENZETİM ÇALIŞMASI

Çalışmanın amacı doğrultusunda bu bölümde benzetim çalışması ve gerçek bir veri kümesinde bir uygulama yapılmıştır.

6.1. Mesane Kanseri Verisi Üzerine Bir Uygulama

Uygulamada, Bölüm 3' te örnek olarak verilen Üroloji Araştırma Grubu tarafından elde edilen mesane kanseri çalışmasındaki veriler kullanılmıştır [17]. Bu çalışmada, çalışmaya giren tüm hastalarda mesane tümörü vardır. Bu tümörler alınıp hastalar rasgele olarak tedavi (piridoksin ve tiyotepa) ve kontrol grubu olmak üzere üç gruba atanmış ve hastalar izlenmeye başlanmıştır. Birçok hastada, çalışma boyunca tümör yeniden oluşmuştur. Bu tümörler tekrar alınmış ve hastalar izlenmeye devam edilmiştir. Burada, her hastanın tümör yineleme zamanları, tedavinin başlangıcından itibaren ölçülmüştür.

Çalışmanın amacı, plaseboya karşı nükseden mesane tümörlerinde piridoksin (Vitamin B₆) ve tiyotepa etkisini belirlemektir. Bu veri kümesi literatürde birçok çalışmada kullanılmıştır. Ancak bu çalışmalarda kontrol ile tiyotepa grupları karşılaştırıldığı için tez çalışmasında bu iki grup alınmıştır. Çalışma, kontrol grubunda 48 hasta, tedavi (tiyotepa) grubunda 38 hasta olmak üzere toplamda 76 hastadan oluşmaktadır. Çalışmanın sonunda kontrol grubundaki hastalardan 11 kişi, tedavi grubundaki hastalardan ise 10 kişinin hastalığından dolayı öldüğü gözlenmiştir. Dolayısıyla bu çalışmada bilgi içeren durdurma söz konusudur.

Bu verinin modellenmesinde, tez çalışmasında önerilen iki yeni model kullanılmış ve modellerde yer alan parametrelerin tahmini R programında kod yazılarak elde edilmiştir [51]. R programında yazılan kod aracılığıyla her iki modele ilişkin parametrelerin tahmin değerleri, tahminlerin standart hataları (SH) ve Akaike bilgi kriteri (AIC) değerleri Çizelge 1 ve Çizelge 2' de verilmiştir.

Çizelge 1. Önerilen Birinci Model İçin Mesane Kanseri Verisi Sonuçları

		λ	β_1	β_2	θ^*
Önerilen Birinci Model	TAHMİN	0.0489	-0.3454	0.5968	0.5452
	SH	0.0119	0.3643	0.4230	0.0690
	AIC	980.2474			

Önerilen birinci modelde Eşitlik (5.13)' te yer alan zayıflık terimine göre integral alınabildiği için parametre yorumları Cox orantılı tehlike modeline benzer şekilde yapılabilir. Çizelge 1'e göre çalışmanın sonunda ölüm olayını yaşamayan hastalar için kontrol grubunun tedavi grubuna göre kanser yineleme riski $\exp(\beta_1)=1.41$ kat daha fazladır. Bu durum çalışmanın sonunda ölüm olayını yaşayan hastalar için tedavi grubunun kontrol grubuna göre kanser yineleme riski $\exp(\beta_1 + \beta_2)=1.28$ kat fazladır.

Çizelge 2. Önerilen İkinci Model İçin Mesane Kanseri Verisi Sonuçları

		λ	β	α	μ	θ^*
Önerilen İkinci Model	TAHMİN	0.0097	0.8402	0.5017	0.0499	0.4980
	SH	0.00109	0.1942	0.0427	0.0122	0.0626
	AIC	1113.7810				

Önerilen ikinci modelde ise zayıflık terimine göre integral alınmadığı için orantılılık varsayımı geçerli değildir. Bu yüzden parametre tahminlerinin yorumları tehlike oranları türünden yorumlanamaz. Çizelge 2' den β parametresine göre tedavinin kanser yineleme sıklığı üzerinde arttırıcı ($0.84>0$) bir etkisi olduğu görülmektedir.

Çizelge 1 ve Çizelge 2' deki sonuçlar karşılaştırıldığında; tedavinin (tiyotepa) kanser yineleme sıklığı üzerinde Çizelge 1' e göre azaltıcı Çizelge 2' ye göre ise arttırıcı etkisinin olduğu görülmektedir. Bu çelişkinin nedeni, aynı verinin incelendiği Ghosh (2000)' un tez çalışmasında da belirtildiği gibi tedavinin kanser yineleme sıklığı üzerinde azaltıcı bir etkisinin olmasının yanı sıra ölüm riskini arttırıcı bir etkiye de sahip olmasıdır [16].

Önerilen yeni modeller, AIC kriterine göre karşılaştırıldığında bu veri kümesi için önerilen birinci modelin daha uygun bir model olduğu söylenebilir.

6.2. Benzetim Çalışması

Benzetim çalışmasında, bilgi içeren durdurma varlığında yinelemeli olay verisi homojen Poisson süreci kullanılarak farklı örneklem büyüklükleri ve durdurma oranları için türetilmiştir

Örneklem büyüklükleri 40, 60 ve 100 olarak alınmıştır. Modeldeki açıklayıcı değişken, 0,1 değerlerini alan kategorik bir değişkendir ve açıklayıcı değişken değerleri $p = 0.50$ olasılıkla Bernoulli dağılımdan türetilmiştir. Regresyon parametreleri $\beta_1 = -1$, $\beta_2 = 1$ ve modeldeki temel yoğunluk fonksiyonunun değeri $\lambda = 2$ olarak alınmıştır.

Zayıflık teriminin dağılımı için $\theta^* = 2$ ile $Gamma(\theta^*, \frac{1}{\theta^*})$ dağılımı düşünülmüş ve bireylere özgü zayıflık değerleri bu dağılımdan türetilmiştir. Çalışmadaki bireylerin her biri en az bir kez, en çok dört kez yineleme yaşamıştır. Bireylerin durdurma ya da ölüm zamanları birbirinden farklı olacak şekilde türetilmiştir. Ölüm oranı olarak 0.10, her yinelemeden sonra durdurma oranı ise 0.01, 0.05 ve 0.10 olarak alınmıştır. Her bir örneklem büyüklüğü ve ölüm oranları için 500 tekrar yapılmıştır.

Farklı örneklem büyüklükleri ve durdurma oranları için elde edilen parametrelerin tahmin değerleri, hata kareler ortalamaları (HKO), kapsama olasılıkları (KO) ve Akaike bilgi kriteri değerleri (AIC) Çizelge 3-11' de verilmiştir. Ayrıca çizelgelerde yer alan azaltılmış model, önerilen birinci modelde $\beta_2 = 0$ ve $\theta^* = 0$ durumuna karşılık gelmektedir. Bunun anlamı veride bilgi içeren durdurmanın dikkate alınmamasıdır.

Çizelge 3. n=40 ve Durdurma Oranı=0.01 Alındığında Önerilen Birinci Model İçin Benzetim Sonuçları

		λ	β_1	β_2	θ^*
Önerilen Birinci Model	TAHMİN	1.5665	-1.1831	1.0389	1.5520
	HKO	0.5159	0.0968	0.0648	0.6887
	KO	0.6525	0.7258	0.9034	0.7142
	AIC	827.4091			
Azaltılmış Model	TAHMİN	0.3134	0.1530	-	-
	HKO	4.4458	1.3396	-	-
	KO	0.092664	0	-	-
	AIC	3951.1422			
Ölüm Oranı		0.1000			

Çizelge 4. n=40 ve Durdurma Oranı =0.05 Alındığında Önerilen Birinci Model İçin Benzetim Sonuçları

		λ	β_1	β_2	θ^*
Önerilen Birinci Model	TAHMİN	1.5959	-1.1826	1.0352	1.5264
	HKO	0.5164	0.0977	0.0689	0.7455
	KO	0.7564	0.7980	0.9198	0.6908
	AIC	788.5604			
Azaltılmış Model	TAHMİN	0.3176	0.1421	-	-
	HKO	4.4496	1.3143	-	-
	KO	0.0992	0	-	-
	AIC	3491.9802			
Ölüm Oranı		0.1000			

Çizelge 5. n=40 ve Durdurma Oranı =0.10 Alındığında Önerilen Birinci Model İçin Benzetim Sonuçları

		λ	β_1	β_2	θ^*
Önerilen Birinci Model	TAHMİN	1.5821	-1.1790	1.0118	1.5605
	HKO	0.5268	0.1056	0.1002	0.7487
	KO	0.7426	0.8069	0.9314	0.7248
	AIC	745.2377			
Azaltılmış Model	TAHMİN	0.2198	0.1445	-	-
	HKO	7.2018	1.3191	-	-
	KO	0.076305	0	-	-
	AIC	3255.1475			
Ölüm Oranı		0.1000			

Çizelge 6. n=60 ve Durdurma Oranı =0.01 Alındığında Önerilen Birinci Model İçin Benzetim Sonuçları

		λ	β_1	β_2	θ^*
Önerilen Birinci Model	TAHMİN	1.5584	-1.1696	1.0524	1.7537
	HKO	0.3925	0.0750	0.0490	0.5221
	KO	0.7827	0.8418	0.9174	0.7783
	AIC	1141.1738			
Azaltılmış Model	TAHMİN	1.6415	-0.0049	-	-
	HKO	1.6150	0.9934	-	-
	KO	0.761589	0	-	-
	AIC	3309.6809			
Ölüm Oranı		0.1000			

Çizelge 7. n=60 ve Durdurma Oranı =0.05 Alındığında Önerilen Birinci Model İçin Benzetim Sonuçları

		λ	β_1	β_2	θ^*
Önerilen Birinci Model	TAHMİN	1.5282	-1.1567	1.0653	1.8359
	HKO	0.3969	0.0686	0.0484	0.5693
	KO	0.7582	0.8674	0.9011	0.8599
	AIC	1052.5197			
Azaltılmış Model	TAHMİN	1.7317	-0.0109	-	-
	HKO	1.8445	0.9808	-	-
	KO	0.802721	0	-	-
	AIC	2951.4873			
Ölüm Oranı		0.1000			

Çizelge 8. n=60 ve Durdurma Oranı =0.10 Alındığında Önerilen Birinci Model İçin Benzetim Sonuçları

		λ	β_1	β_2	θ^*
Önerilen Birinci Model	TAHMİN	1.4940	-1.1414	1.0806	1.7745
	HKO	0.4511	0.0649	0.0513	0.4772
	KO	0.7256	0.8590	0.8815	0.8125
	AIC	1006.8226			
Azaltılmış Model	TAHMİN	1.6502	-0.0085	-	-
	HKO	2.0832	0.9857	-	-
	KO	0.7578	0	-	-
	AIC	2880.4654			
Ölüm Oranı		0.1000			

Çizelge 9. n=100 ve Durdurma Oranı =0.01 Alındığında Önerilen Birinci Model İçin Benzetim Sonuçları

		λ	β_1	β_2	θ^*
Önerilen Birinci Model	TAHMİN	1.4250	-1.1323	1.0897	1.6431
	HKO	0.4169	0.0412	0.0318	0.2986
	KO	0.6705	0.8553	0.8845	0.7246
	AIC	1903.9224			
Azaltılmış Model	TAHMİN	2.0797	-0.0211	-	-
	HKO	0.6594	0.9584	-	-
	KO	0.9784	0	-	-
	AIC	5522.9343			
Ölüm Oranı		0.1000			

Çizelge 10. n=100 ve Durdurma Oranı =0.05 Alındığında Önerilen Birinci Model İçin Benzetim Sonuçları

		λ	β_1	β_2	θ^*
Önerilen Birinci Model	TAHMİN	1.4006	-1.1405	1.0815	1.7478
	HKO	0.4434	0.0426	0.0295	0.3159
	KO	0.6645	0.8286	0.8449	0.8062
	AIC	1808.9962			
Azaltılmış Model	TAHMİN	2.1240	-0.0214	-	-
	HKO	0.9605	0.9577	-	-
	KO	1	0	-	-
	AIC	5108.5693			
Ölüm Oranı		0.1000			

Çizelge 11. n=100 ve Durdurma Oranı =0.10 Alındığında Önerilen Birinci Model İçin Benzetim Sonuçları

		λ	β_1	β_2	θ^*
Önerilen Birinci Model	TAHMİN	1.4067	-1.1399	1.0821	1.7513
	HKO	0.4404	0.0437	0.0309	0.3037
	KO	0.6512	0.8246	0.8648	0.7567
	AIC	1692.4875			
Azaltılmış Model	TAHMİN	2.1168	-0.0214	-	-
	HKO	0.8290	0.9577	-	-
	KO	1	0	-	-
	AIC	4716.6303			
Ölüm Oranı		0.1000			

Çizelge 3-11'den görüldüğü gibi λ ve β_1 parametrelerine ilişkin hata kareler ortalamaları, önerilen birinci modelde azaltılmış modele göre daha düşük çıkmıştır. Parametrelere ilişkin kapsama olasılıklarına bakıldığında örneklem büyüklüğü arttıkça azaltılmış modelde λ parametresi için sonuçlar iyileşmekte ancak β_1 parametresi önerdiğimiz birinci model için azaltılmış modele göre her zaman daha iyi sonuç vermektedir. Ayrıca örneklem büyüklüğünün 60' dan küçük olduğu durumlarda da λ değeri önerdiğimiz birinci modelde daha iyidir. β_2 ve θ^* parametreleri azaltılmış modelde yer almadığından, bu parametreler bakımından iki model karşılaştırılmaz. Bu yüzden model karşılaştırması AIC değerlerine bakılarak yapılmıştır. AIC değerleri modeldeki parametre sayısına duyarlı olmasına

rağmen önerdiğimiz modelde daha düşük çıkmıştır. Her bir yineleme için durdurma oranlarının artması sonuçlar üzerinde önemli bir etki yaratmamıştır.

Önerilen ikinci model için yapılan benzetim çalışmasında ise, bilgi içeren durdurma varlığında yinelemeli olay verisi önerilen birinci modeldekine benzer olarak homojen Poisson süreci kullanılarak farklı örneklem büyüklükleri, ölüm ve durdurma oranları için türetilmiştir. Örneklem büyüklükleri yine 40, 60 ve 100 olarak alınmıştır. Modeldeki açıklayıcı değişken, 0,1 değerlerini alan kategorik bir değişkendir ve açıklayıcı değişken değerleri $p = 0.50$ olasılıkla Bernoulli dağılımdan türetilmiştir. Regresyon parametresi $\beta = 1$ olarak alınmıştır. Zayıflık teriminin dağılımı için $\theta^* = 2$ ile $Gamma\left(\theta^*, \frac{1}{\theta^*}\right)$

dağılımı düşünülmüş ve bireylere özgü zayıflık değerleri bu dağılımdan türetilmiştir. Ölüm süreci için temel yoğunluk fonksiyonu $\mu = 5$, ölüm olayı varlığında yinelemeli olay sürecine ilişkin yoğunluk fonksiyonunda yer alan $\alpha = 2$ ve yinelemeli olay sürecine ilişkin temel yoğunluk fonksiyonu $\lambda = 2$ olarak alınmıştır. Çalışmadaki bireylerin her biri en az bir kez, en çok dört kez yineleme yaşamıştır. Bireylerin durdurma ya da ölüm zamanları birbirinden farklı olacak şekilde türetilmiştir. Ölüm oranları olarak 0.20, 0.50 ve 0.70; her yinelemeden için durdurma oranı ise 0.01 olarak alınmıştır. Her bir örneklem büyüklüğü ve ölüm oranları için 500 tekrar yapılmıştır.

Farklı örneklem büyüklükleri ve ölüm oranları için elde edilen parametrelerin tahmin değerleri, hata kareler ortalamaları (HKO), kapsama olasılıkları (KO) ve Akaike bilgi kriteri değerleri (AIC) Çizelge 12-20' de verilmiştir. Ayrıca çizelgelerde yer alan azaltılmış model, bilgi içeren durdurmanın dikkate alınmadığı durumu ifade ettiği için azaltılmış modelde $\alpha = 1$, $\mu = 0$ ve $\theta^* = 0$ durumuna karşılık gelmektedir.

Çizelge 12. n=40 ve Ölüm Oranı=0.20 Alındığında Önerilen İkinci Model İçin Benzetim Sonuçları

		λ	β	α	μ	θ^*
Önerilen İkinci Model	TAHMİN	1.9711	0.9614	1.9358	4.9483	1.9814
	HKO	0.1055	0.0825	0.1743	4.1510	0.2651
	KO	0.9494	0.9691	0.9663	1.0000	0.9579
	AIC	280.5430				
Azaltılmış Model	TAHMİN	1.9253	0.3051	-	-	-
	HKO	0.0778	0.6330	-	-	-
	KO	0.9270	0.1348	-	-	-
	AIC	227.0402				

Çizelge 13. n=40 ve Ölüm Oranı=0.50 Alındığında Önerilen İkinci Model İçin Benzetim Sonuçları

		λ	β	α	μ	θ^*
Önerilen İkinci Model	TAHMİN	1.9401	0.9493	1.8958	4.9633	1.8971
	HKO	0.0888	0.1112	0.1822	3.4593	0.2999
	KO	0.9233	0.9443	0.9268	1.0000	0.9059
	AIC	144.1925				
Azaltılmış Model	TAHMİN	1.9205	0.4568	-	-	-
	HKO	0.0930	0.5691	-	-	-
	KO	0.9059	0.3206	-	-	-
	AIC	211.0169				

Çizelge 14. n=40 ve Ölüm Oranı=0.70 Alındığında Önerilen İkinci Model İçin Benzetim Sonuçları

		λ	β	α	μ	θ^*
Önerilen İkinci Model	TAHMİN	1.8637	0.8849	1.7974	4.9019	1.6527
	HKO	0.1661	0.1874	0.3799	3.4658	0.4472
	KO	0.8103	0.8735	0.8379	1.0000	0.7115
	AIC	87.5674				
Azaltılmış Model	TAHMİN	1.9566	0.6732	-	-	-
	HKO	0.1002	0.5164	-	-	-
	KO	0.9368	0.4190	-	-	-
	AIC	178.2271				

Çizelge 15. n=60 ve Ölüm Oranı =0.20 Alındığında Önerilen İkinci Model İçin Benzetim Sonuçları

		λ	β	α	μ	θ^*
Önerilen İkinci Model	TAHMİN	1.9665	0.9936	1.9427	4.9675	2.0601
	HKO	0.0767	0.0632	0.0889	2.7656	0.2134
	KO	0.9467	0.9594	0.9315	1.0000	0.9619
	AIC	401.6538				
Azaltılmış Model	TAHMİN	1.9215	0.3375	-	-	-
	HKO	0.0657	0.5553	-	-	-
	KO	0.8858	0.1015	-	-	-
	AIC	322.9334				

Çizelge 16. n=60 ve Ölüm Oranı =0.50 Alındığında Önerilen İkinci Model İçin Benzetim Sonuçları

		λ	β	α	μ	θ^*
Önerilen İkinci Model	TAHMİN	1.9707	0.9708	1.9439	4.9566	1.9375
	HKO	0.0655	0.0577	0.1065	1.8021	0.1856
	KO	0.9424	0.9741	0.9280	1.0000	0.9366
	AIC	196.9906				
Azaltılmış Model	TAHMİN	1.9344	0.4459	-	-	-
	HKO	0.0673	0.5056	-	-	-
	KO	0.8934	0.2709	-	-	-
	AIC	301.3031				

Çizelge 17. n=60 ve Ölüm Oranı =0.70 Alındığında Önerilen İkinci Model İçin Benzetim Sonuçları

		λ	β	α	μ	θ^*
Önerilen İkinci Model	TAHMİN	1.9129	0.9106	1.8712	4.9212	1.7309
	HKO	0.0796	0.0816	0.1374	1.4958	0.2895
	KO	0.8527	0.8876	0.8566	1.0000	0.7907
	AIC	100.7805				
Azaltılmış Model	TAHMİN	1.9544	0.6452	-	-	-
	HKO	0.0758	0.4541	-	-	-
	KO	0.9147	0.4147	-	-	-
	AIC	265.4166				

Çizelge 18. n=100 ve Ölüm Oranı =0.20 Alındığında Önerilen İkinci Model İçin Benzetim Sonuçları

		λ	β	α	μ	θ^*
Önerilen İkinci Model	TAHMİN	1.9843	1.0050	1.9662	4.9782	2.0874
	HKO	0.0391	0.0269	0.0492	1.6562	0.1384
	KO	0.9555	0.9488	0.8820	1.0000	0.9688
	AIC	645.6856				
Azaltılmış Model	TAHMİN	1.9146	0.3601	-	-	-
	HKO	0.0560	0.4856	-	-	-
	KO	0.8441	0.0312	-	-	-
	AIC	523.2573				

Çizelge 19. n=100 ve Ölüm Oranı =0.50 Alındığında Önerilen İkinci Model İçin Benzetim Sonuçları

		λ	β	α	μ	θ^*
Önerilen İkinci Model	TAHMİN	1.9811	0.9957	1.9493	4.9921	1.9452
	HKO	0.0836	0.4593	0.9588	0.97613	0.1059
	KO	0.9337	0.9821	0.8929	1.0000	0.9464
	AIC	308.5778				
Azaltılmış Model	TAHMİN	1.9170	0.4530	-	-	-
	HKO	0.0579	0.4238	-	-	-
	KO	0.8520	0.1582	-	-	-
	AIC	480.4181				

Çizelge 20. n=100 ve Ölüm Oranı =0.70 Alındığında Önerilen İkinci Model İçin Benzetim Sonuçları

		λ	β	α	μ	θ^*
Önerilen İkinci Model	TAHMİN	1.9240	0.9271	1.8572	4.9300	1.7265
	HKO	0.0731	0.0609	0.1124	0.8608	0.2286
	KO	0.8492	0.8810	0.8294	1.0000	0.7381
	AIC	142.4132				
Azaltılmış Model	TAHMİN	1.9578	0.6167	-	-	-
	HKO	0.0551	0.3310	-	-	-
	KO	0.9087	0.3294	-	-	-
	AIC	413.0457				

Çizelge 12-20' den görüldüğü gibi, ölüm oranlarının arttığı sabit örneklem büyüklüğünde önerilen ikinci modeldeki μ parametresinin tahminlerinde iyileşme gözlenmiştir. Örneklem büyüklüğü arttığı zaman genel olarak parametrelerin HKO değerlerinde düşme gözlenmiştir. Azaltılmış modelde α , μ ve θ^* parametreleri yer almadığından model karşılaştırmaları AIC değerleri bakımından yapılmıştır. Buna göre önerilen ikinci modeldeki AIC değerleri daha düşük çıkmıştır.

7. SONUÇLAR VE TARTIŞMA

Yinelemeli olay verisinin modellenmesinde bilgi içeren durdurma olayının göz önünde bulundurulması gerekmektedir. Çünkü veride ölüm gibi bir durdurma söz konusu olduğunda yineleme zamanları ile ölüm zamanları ilişkilidir; bu ilişkinin modelde dikkate alınmaması parametrelere ilişkin yanlış tahminler elde edilmesine yol açmaktadır. Gerçekte bu ilişkinin tam olarak modellenmesi mümkün değildir. Ancak, literatürde bilgi içeren durdurma varlığı altında yinelemeli olay verisinin modellenmesi üzerine yapılan çok sayıda çalışma vardır. Tez çalışmasında ayrıntılı olarak anlatılan modeller, bilgi içeren durdurma varlığında yinelemeli olay verisinin modellenmesi üzerine değişik yaklaşımları içerir.

Tez çalışmasında, yinelemeli olay süreci ile ölüm süreci arasında güçlü bir ilişkinin olduğu düşünülerek bu ilişkinin yapısını da modelde dikkate alan iki yeni model önerilmiştir. Önerilen her iki model de Cox orantılı tehlike modelinin, yinelemeli olay süreci için genişletilmiş biçimi olarak düşünülebilir. Bilgi içeren durdurma varlığında, yinelemeli olay süreci, homojen Poisson süreci ile modellenmiş; bilgi içeren durdurma olarak ise ölüm olayı alınmıştır.

Önerilen birinci modelde ölüm olayının, yinelemeli olay sürecine ilişkin yoğunluk fonksiyonuna açıklayıcı değişken olarak eklenmesi düşünülmüştür. Aynı biçimde yinelemeli olaylar arasındaki ilişki ise modele gözlenemeyen rasgele etkiyi ifade eden zayıflık teriminin eklenmesiyle oluşturulmuştur.

Önerilen ikinci modelde, Liu ve arkadaşlarının (2004) çalışması temel alınmıştır. Liu ve arkadaşlarının çalışmasında ölüm olayı ve ölüm olayı bilindiğinde yinelemeli olay sürecine ilişkin iki farklı yoğunluk fonksiyonu önerilirken; tez çalışmasında ölüm olayı bilindiğinde yinelemeli olay süreci ile durdurma varlığında yinelemeli olay süreci iki farklı süreç olduğu için ölüm olayına ve bu iki farklı sürece ilişkin üç farklı yoğunluk fonksiyonu önerilmiştir. Ayrıca önerilen bu yoğunluk fonksiyonları ortak bir zayıflık terimi ile

birbirlerine bağlanmıştır. Zayıflık terimleri gözlenemediği için parametrelerin en çok olabirlik tahminini elde edilmesinde EM algoritmasından yararlanılmıştır.

Tez çalışmasında önerilen her iki model için gerçek bir veri kümesi kullanılarak bilgi içeren durdurma varlığında yinelemeli olay sürecine ilişkin parametre tahminleri elde edilmiş; daha sonra da farklı senaryolar için benzetim çalışması yapılmıştır. Benzetim çalışması R programında kod yazılarak yürütülmüş ve her bir senaryoya ilişkin sonuçlar yaklaşık 20 gün sonunda elde edilebilmiştir.

Benzetim çalışmasında önerilen modellerin geçerliliğini göstermek için, ölüm olayının dikkate alınmadığı ve bilgi içermeyen durdurma olarak düşünüldüğü azaltılmış model ile karşılaştırma yapılmış, bunun için AIC kriteri kullanılmıştır. Çalışmada önerilen birinci ve ikinci modellerin, AIC değerlerine göre azaltılmış model ile yapılan karşılaştırılmasında genel olarak daha iyi sonuçlar verdiği görülmüştür. Bu sonuç, ölüm olayı bilindiğinde yinelemeli olay süreci ve durdurma varlığında yinelemeli olay sürecinin modele etkisinin önemli olduğunu düşündürmektedir.

Bu tez çalışmasında önerilen modeller, literatürde yer alan modellerle karşılaştırılamamıştır. Bunun nedeni literatürde yer alan çalışmalarda geçen modellerin yapısının ve modellerdeki parametrelerin anlamsal olarak birbirinden farklılık göstermesidir.

Buradan yola çıkarak daha sonra yapılacak çalışmalarda,

- Bilgi içeren durdurma varlığında yinelemeli olay sürecinin modellenmesinde zamandan bağımsız açıklayıcı değişkenle birlikte zamana bağlı açıklayıcı değişkenlerin de kullanılması düşünülebilir.

- Tez çalışmasında homojen Poisson süreci varsayımı yapılmakla beraber temel yoğunluk fonksiyonu, zamanın bir fonksiyonu olarak düşünülerek yinelemeli homojen olmayan Poisson süreci olarak da modellenebilir.
- Ayrıca tez çalışmasında, önerilen iki yeni model için açıklayıcı değişkenlerin yoğunluk fonksiyonu üzerindeki çarpımsal etkisi ele alınmıştır. Bunun yanı sıra açıklayıcı değişkenlerin yoğunluk fonksiyonu üzerinde toplamsal etkisinin olduğu modeller de düşünülebilir.
- Çalışmada önerilen her iki modelde, bilgi içeren durdurma ile yineleme süreci arasındaki ilişki zayıflık terimi kullanılarak modellenmiştir. Ancak gerçekte bilgi içeren durdurma ile yinelemeli olay süreci arasındaki ilişkinin tam olarak modellenmesinin mümkün olmadığı literatürde yer alan çoğu çalışmada belirtilmiştir. Bu dikkate alındığından bilgi içeren durdurma ve yinelemeli olay süreci arasındaki ilişkinin fonksiyonel yapısı üzerine de modellerin geliştirilmesi düşünülebilir.

KAYNAKLAR

1. Cox, D.R., Regression models and life tables. *Journal of the Royal Statistical Society, Series B* , 34 (2), 187-220, **1972**.
2. Kleinbaum, D., M. Klein, *Survival Analysis: A Self Learning Text*, Springer, New York, **2005**.
3. Ata, N., *Yaşam Çözümlemesinde Orantısız Hazard Modeli*, Yüksek Lisans Tezi, Hacettepe Üniversitesi Fen Bilimleri Enstitüsü, Ankara, **2005**.
4. Başar, E., *Yaşam Tabloları Analizinde Kullanılan Bazı İstatistiksel Tekniklerin Böbrek Nakli Verilerine Uygulanması*, Doktora Tezi, Hacettepe Üniversitesi Fen Bilimleri Enstitüsü, Ankara, **1993**.
5. Collett, D., *Modelling Survival Data in Medical Research*, Chapman & Hall, London, **2003**.
6. Özkartal, B., *Yinelemeli Olay Analizi ve Çocuk Suçluluğu Üzerine Bir Uygulama*, Yüksek Lisans Tezi, Marmara Üniversitesi Fen Bilimleri Enstitüsü, İstanbul, **2009**.
7. Ata, N., Karasoy, D., Sözer, T., Orantısız hazardlar için tabakalandırılmış Cox regresyon modeli ve meme kanseri hastaları üzerine bir uygulama, *Türkiye Klinikleri*, 28, 327-332, **2008**.
8. Qi, J., *Comparison of Proportional Hazards and Accelerated Failure Time Models*, MS Thesis, University of Saskatchewan Department of Mathematics and Statistics, **2009**.
9. Zhang, D., *Marginal Quantile Regression Methods for Censored Multiple Event Times*, PhD Thesis, Graduate School of Arts and Science, Columbia University, **2006**.
10. Chen, Y., Q., Wang, M .C., Analysis of accelerated failure hazards models, *Journal of the American Statistical Association*, 95(450), 608-618, **2000**.
11. Vaupel, J.W., Manton, K.G., Stallard E., The impact of heterogeneity in individual frailty on the dynamics of mortality, *Demography*, 16, 439-454, **1979**.
12. Ata, N., Karasoy, D., Sağkalım çözümlemesi için zayıflık modeli ve mide kanseri hastalarına ilişkin verilerle bir uygulama, *Çankaya University Journal of Science and Engineering*, 8 (2), 225-235, **2011**.
13. Wienke, A., Frailty Models, [http://www.ressources-actuarielles.net/EXT/ISFA/1226.nsf/0/37d8815030298678c125737b002e1b13/\\$FILE/wp-2003-032.pdf](http://www.ressources-actuarielles.net/EXT/ISFA/1226.nsf/0/37d8815030298678c125737b002e1b13/$FILE/wp-2003-032.pdf) (Haziran, **2013**).
14. Kalbfleisch, J.D., Prentice, R. L., *The Statistical Analysis of Failure Time Data*, Second Edition, John Wiley & Sons, Inc, Canada, **2002**.

15. Duchateau, L., Janssen, P., *The Frailty Model*, Springer, New York, **2008**.
16. Ghosh, D., *Nonparametric and Semiparametric Analysis of Recurrent Events in the Presence of Terminal Events and Dependent Censoring*, Phd Dissertation, Washington University, United States, **2000**.
17. Byar, D.P., Blackard, C., Comparisons of placebo, pyridoxine, and topical thiotepa in preventing recurrence of stage 1 bladder cancer, *Urology*, 10, 556-561, **1976**.
18. Cook, R.J., Lawless, J.F., *The Statistical Analysis of Recurrent Events*, Springer, New York, **2007**.
19. Prentice, P.L., Williams, B.J., Peterson, A.V., On the regression analysis of multivariate failure time data, *Biometrika* 68(2), 373-379, **1981**.
20. Wei, L.J., Lin, D.Y., Weissfeld, L., Regression analysis of multivariate incomplete failure time data by modelling marginal distributions, *Journal of American Statistical Associations*, 84(408), 1065-1073, **1989**.
21. Andersen, P.K., Gill, R.D., Cox's regression model for counting processes: A large sample study, *The Annals of Statistics*, 10(4), 1100-1120, **1982**.
22. Pepe, M.S., Cai, J., Some graphical displays and marginal regression analyses for recurrent failure times and time dependent covariates, *Journal of the American Statistical Association*, 88(423), 811-820, **1993**.
23. Lin, D.Y., Wei, L.J., Yang, I., Ying, Z., Semiparametric regression for the mean and rate functions of recurrent events, *Journal of Royal Statistical Society*, 62(4), 711-730, **2000**.
24. Lin, J.S., Wei, L.J., Linear regression analysis for multivariate failure time observations, *Journal of the American Statistical Association*, 87, 1091-1097, **1992**.
25. Lin, D.Y., Wei, L.J., Ying, Z., Accelerated failure time models for counting processes, *Biometrika*, 85, 605-618, **1998**.
26. Pan, W., Using frailties in the accelerated failure time model, *Lifetime Data Analysis*, 7, 55-64, **2001**.
27. Aalen, O.O., *Statistical Inference for a Family of Counting Processes*, PhD Thesis, University of California, Berkeley, CA, **1975**.
28. Viswanathan, S., *Statistical Methods for Recurrent Event Data in the Presence of a Terminal Event and Incomplete Covariate Information*, PhD Thesis, University of North Carolinas Department of Biostatistics, **2011**.
29. McGilchrist, C.A., Aisbett, C.W., Regression with frailty in survival analysis, *Biometrics*, 47(2) 461-466, **1991**.
30. Nielsen, G.G., Gill, R.D., Andersen, P.K., Sorensen, T.I.A., A counting process approach to maximum likelihood estimation in frailty models, *Scandinavian Journal of Statistics*,. 19(1), 25-43, **1992**.

31. Çinlar, E., *Introduction to Stochastic Processes*, Prentice-Hall, Inc., New Jersey, USA, **1975**.
32. Lawless, J.F., Regression methods for poisson process data, *Journal of the American Statistical Association*, 82(399), 808-815, **1987**.
33. Balshaw, R.F., *Analysis of Recurrent Event Data*, PhD Thesis, Simon Fraser University Department of Mathematics and Statistics, **1997**.
34. Nielsen, J.D., *Spline Models for the Analysis of Recurrent Event Panel Data*, Phd Thesis, Simon Fraser University Department of Statistics and Actuarial Science, **2007**.
35. Fredette, M., *Prediction of Recurrent Events*, Phd Thesis, Department of Statistics, University of Waterloo, Waterloo, Ontario, Canada, **2004**.
36. Li, Q., Lagakos, S., Use of the Wei-Lin-Weissfeld method for the analysis of a recurring and a terminating event, *Statistics in Medicine*, 16, 925-940, **1997**.
37. Liu, L., Wolfe, R.A., Huang, X., Shared frailty models for recurrent events and a terminal event, *Biometrics*, 60(3), 747-756, **2004**.
38. Huang, X., Wolfe, R.A., A frailty model for informative censoring, *Biometrics*, 58(3), 510-520, **2002**.
39. Huang, C.Y., Wang, M.C., Joint modelling and estimation for recurrent event processes and failure time data, *Journal of the American Statistical Association*, 99(468), 1153-1165, **2004**.
40. Ye, Y., Kalbfleisch, J.D., Schaubel, D.E., Semiparametric analysis of correlated recurrent and terminal events, *Biometrics*, 63, 78-87, **2007**.
41. Zeng, D., Lin, D.Y., Semiparametric transformation models with random effects for joint analysis of recurrent and terminal events, *Biometrics*, 65, 746-752, **2009**.
42. Liu, L., *Modelling Recurrent Events and Medical Cost Data in the Presence of a Correlated Terminal Event*, Phd Thesis, The University of Michigan Department of Biostatistics, **2004**.
43. Wang, M.C., Qin, J., Chiang, C.T., Analyzing recurrent event data with informative censoring, *Journal of the American Statistical Association*, 96(455), 1057-1065, **2001**.
44. Wang, M.C., Chiang, C.T., Non-parametric methods for recurrent event data with informative and non-informative censorings, *Statistics In Medicine*, 21, 445-456, **2002**.
45. Luo, X., Wang, M.C., *Recurrent event models in the presence of a terminal event: comparison, inference and data analysis*, Working Papers, Johns Hopkins University, Department of Biostatistics, **2006**.

46. Cook, R. J., Lawless, J.F., Marginal analysis of recurrent events and a terminating event. *Statistics in Medicine*, 16, 911-924, **1997**.
47. Huang, C.Y., Qin, J., Wang, M.C., Semiparametric analysis for recurrent event data with time-dependent covariates and informative censoring, *Biometrics*, 66, 39-49, **2010**.
48. Ghosh, D., Lin, D.Y., Nonparametric analysis of recurrent events and death, *Biometrics*, 56, 554-562, **2000**.
49. Ghosh, D., Lin, D.Y., Marginal regression models for recurrent and terminal events, *Statistica Sinica*, 12, 663-688, **2002**.
50. Ghosh, D., Lin, D.Y. Semiparametric analysis of recurrent events data in the presence of dependent censoring, *Biometrics*, 59, 877-885, **2003**.
51. Gentleman, R., Ihaka, R. R Programming Language, <http://www.r-project.org/> (Şubat, **2012**)

ÖZGEÇMİŞ

Kimlik Bilgileri

Adı Soyadı : Hande ÜNLÜ
Doğum Yeri : Ankara
Medeni Hali : Evli
E-posta : handekonsuk@gmail.com
Adresi : Hacettepe Üniversitesi Fen Fakültesi İstatistik Bölümü
Beytepe / ANKARA

Eğitim

Lise : 1998 - 2002 Ayrancı Süper Lisesi
Lisans : 2002 - 2007 Hacettepe Üniversitesi Fen Fakültesi
İstatistik Bölümü
Yüksek Lisans
& Doktora : 2007 - 2013 Bütünleşik Yüksek Lisans ve Doktora Programı

Yabancı Dil ve Düzeyi

İngilizce (KPDS: 84/100)

İş Deneyimi

2007 - ... Araştırma Görevlisi, Hacettepe Üniversitesi,
İstatistik Bölümü
1 Şubat - 1 Ağustos 2012 Ziyaretçi Bilim Adamı (Visiting Scholar), Penn
State Üniversitesi, *İstatistik Bölümü*, State
College, ABD. (Prof. Dr. Debashis Ghosh ile)

Deneyim Alanları

Yaşam Çözümlemesi, Olasılıksal Süreçler, Regresyon Çözümlemesi

Tezden Üretilmiş Projeler ve Bütçesi

Tezden Üretilmiş Yayınlar

Tezden Üretilmiş Tebliğ ve/veya Poster Sunumu ile Katıldığı Toplantılar

Koşuk, H., Toktamış, Ö., Accelerated Failure Time Model of Recurrent Events with Informative Censoring, *International 7th Statistics Congress*, Antalya, Turkey, 28th April - 01st May 2011.

Koşuk, H., Yinelemeli Olayların Analizinde Kullanılan Yöntemlerin Karşılaştırılması, *13. Ulusal Biyoistatistik Kongresi*, Ankara, 12-14 Eylül 2011.

