

AKDENİZ ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ

Aygül ANAVATAN

**ZAMAN BAĞIMLI COX REGRESYONU: GENÇ FİRMALARIN
HAYATTA KALMA BAŞARISI ÜZERİNE
BİR UYGULAMA**

Ekonometri Ana Bilim Dalı
Yüksek Lisans Tezi

Antalya, 2011

AKDENİZ ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ

Aygül ANAVATAN

**ZAMAN BAĞIMLI COX REGRESYONU: GENÇ FİRMALARIN
HAYATTA KALMA BAŞARISI ÜZERİNE
BİR UYGULAMA**

Danışman

Prof. Dr. Murat KARAÖZ

Ekonometri Ana Bilim Dalı

Yüksek Lisans Tezi

Antalya, 2011

Akdeniz Üniversitesi
Sosyal Bilimler Enstitüsü Müdürlüğüne,

Aygül ANAVATAN'ın bu çalışması, jürimiz tarafından Ekonometri Ana Bilim Dalı Yüksek Lisans Programı tezi olarak kabul edilmiştir.

Başkan : Prof. Dr. İbrahim Garpör
Üye (Danışmanı) : Prof. Dr. Murat Karaöz
Üye : İrd. Doç. Dr. Halcan Demirçel

Tez Konusu: Zaman Bağımlı Cox Regresyonu : Genç Firmaların Hayatta Kalma Başarısı Üzerine Bir Uygulama

Onay : Yukarıdaki imzaların, adı geçen öğretim üyelerine ait olduğunu onaylarım.

Tez Savunma Tarihi : 7.9./2011

Mezuniyet Tarihi : 12/09/2011

Prof. Dr. Mehmet ŞEN
Müdür

.....

İÇİNDEKİLER

ŞEKİLLER LİSTESİ.....	iv
TABLOLAR LİSTESİ.....	viii
KISALTMALAR LİSTESİ.....	ix
ÖZET.....	x
ABSTRACT.....	xi
TEŞEKKÜR.....	xii
GİRİŞ.....	1

BİRİNCİ BÖLÜM

HAYATTA KALMA ANALİZİ

1.1. Hayatta Kalma Analizi Nedir?	4
1.1.1. Gözlenen Süre ve Gözlemlerin Sansürlenmesi	7
1.1.1.1. Gözlenen Süre	10
1.1.1.2. Gözlemlerin Sansürlenmesi.....	10
1.1.2. Hayatta Kalma Analizinin Amaçları.....	12
1.1.3. Karıştırıcı Faktörler ve Etkileşim	14
1.1.3.1. Karıştırıcı Faktörler.....	14
1.1.3.2. Etkileşim	15
1.2. Hayatta Kalma Analizinde Kullanılan Fonksiyonlar.....	16
1.2.1. Olasılık Yoğunluk Fonksiyonu.....	18
1.2.2. Hayatta Kalma Fonksiyonu	18
1.2.3. Tehlike Fonksiyonu.....	20
1.2.3.1. Baz Tehlike Fonksiyonu	23
1.2.3.2. Birikimli Tehlike Fonksiyonu.....	24
1.3. Hayatta Kalma Analizinde Kullanılan Fonksiyonlar Arasındaki İlişkiler.....	25
1.3.1. Tehlike Oranı ve Eğrisi	26
1.3.2. Tehlike Orantısı ve Eğrisi	27

İKİNCİ BÖLÜM

COX ORANTILI TEHLİKE REGRESYON MODELİ

2.1. Cox Regresyon Modeli ve Varsayımları.....	32
2.2. Cox Regresyon Modelinin Genel Özellikleri ve Tahminlerin Direnci.....	33

2.3. Cox Regresyon Modeli ve Test İstatistikleri	35
2.4. Cox Orantılı Tehlike Regresyon Modelinin Yapısı ve PH Varsayımı	37
2.5. Orantılı Tehlike Modeli için Olabilirlik Fonksiyonunun Elde Edilmesi	39
2.5.1. Newton-Raphson Yöntemi	41

ÜÇÜNCÜ BÖLÜM

ORANTILI TEHLİKE VARSAYIMININ TEST EDİLMESİ

3.1. Orantılı Tehlike Varsayımının Test Edilmesinde Kullanılan Grafikselle Yöntemler	45
3.1.1. Kaplan-Meier ve Log-Log Grafikleri	46
3.1.2. Gözlenene Karşı Beklenen Çizimler	51
3.1.3. Uyum İyiliği Araştırması.....	53
3.2. Orantılı Tehlike Varsayımının Test Edilmesinde Kullanılan İstatistiksel Yöntemler ...	54
3.2.1. Schoenfeld Artıkları ve PH Testi.....	54
3.2.2. Log-Rank Testi.....	58
3.2.3. Log-Rank Testinin Alternatifleri	61
3.3. Cox PH Modelini Kullanarak Hayatta Kalma Eğrilerinin Ayarlanması	64

DÖRDÜNCÜ BÖLÜM

ZAMANA BAĞLI DEĞİŞKENLER VE COX ORANTILI TEHLİKE REGRESYON MODELİNİN GENİŞLETİLMESİ

4.1. Cox Orantılı Tehlike Regresyon Modeli	66
4.2. Zamana Bağlı Değişkenlerin Tanımı	69
4.3. Zamana Bağlı Değişkenler ve Genişletilmiş Cox Model.....	74
4.4. Genişletilmiş Cox Model için Tehlike Orantısı Formülü	77
4.5. Orantılı Tehlike Varsayımını Sağlamayan Zamandan Bağımsız Değişkenlerin Değerlendirilmesi	80
4.6. Zamana Bağlı Değişkenler için Cox Modelin Çözüm Yöntemleri.....	82

BEŞİNCİ BÖLÜM

UYGULAMA

5.1. Araştırmada Kullanılan Değişkenler	88
5.2. Araştırmada Kullanılan Değişkenlerin Özellikleri	92
5.3. Cox Orantılı Tehlike Regresyon Modeli: Temel Model Tahmin Sonuçları.....	93
5.4. Modelde Yer Alan Değişkenlerin PH Varsayımının Denetlenmesi	101

5.4.1. Grafiksel Yaklaşım 1: Log-Log KM Hayatta Kalma Eğrileri ile PH Varsayımının Denetlenmesi.....	101
5.4.2. Grafiksel Yaklaşım 2: Gözlenene Karşı Beklenen Hayatta Kalma Eğrileri ile PH Varsayımının Denetlenmesi.....	109
5.4.3. İstatistiksel Test ile PH Varsayımının Denetlenmesi	115
5.4.4. Schoenfeld Artıkları ile PH Varsayımının Denetlenmesi.....	117
5.4.5. Zamana Bağlı Değişkenleri İçeren Cox Model ile PH Varsayımının Denetlenmesi	128
SONUÇ VE ÖNERİLER.....	134
KAYNAKÇA	136
ÖZGEÇMİŞ.....	139

ŞEKİLLER LİSTESİ

Şekil 1.1: Sansürlü ve başarısız deneklerin gösterimi.....	8
Şekil 1.2: Soldan sansürlü yaşam süresi	9
Şekil 1.3: Sağdan sansürlü yaşam süresi.....	11
Şekil 1.4: Hayatta kalma fonksiyonu için örnekler.....	13
Şekil 1.5: Kontrol ve deney grubu için hayatta kalma fonksiyonlarının karşılaştırılması.....	13
Şekil 1.6: Karıştırıcı faktörlerin kontrol ve deney grubuna olan etkisi.....	14
Şekil 1.7: X bağımsız değişkeninin etkileşim etkisi	15
Şekil 1.8: Hayatta kalma fonksiyonunun grafiği	20
Şekil 1.9: Tahminlenen hayatta kalma fonksiyonunun grafiği	20
Şekil 1.10: Farklı tipteki tehlike fonksiyonlarının grafikleri	22
Şekil 1.11: Üç farklı tehlike fonksiyonu için grafik.....	26
Şekil 1.12: Tahminlenmiş hayatta kalma eğrilerinden medyan yaşam süresinin elde edilmesi	27
Şekil 1.13: Zamana göre sabit tehlike orantısı grafiği	29
Şekil 1.14: Zamana göre sabit olmayan tehlike oranı grafiği	29
Şekil 1.15: Tehlike orantısının zaman içindeki değişimi	30
Şekil 2.1: Sansürlü denekler için kullanılan yaşam süresi bilgisi	42
Şekil 3.1: Log-log hayatta kalma eğrisi grafiği.....	45
Şekil 3.2: Gözlenen ve tahminlenen hayatta kalma eğrilerinin grafiği.....	45
Şekil 3.3: Gözlenen ve tahminlenen hayatta kalma eğrilerinin grafiği.....	49
Şekil 3.4: y eksen ölççeği	49
Şekil 3.5: Gözlenen ve tahminlenen eğriler arasındaki uzaklık.....	50
Şekil 4.1: Kısmen içsel kısmen de yardımcı olabilen değişken için bir örnek	73
Şekil 4.2: Değişkeni kısmen içsel kısmen de yardımcı yapan faktörler	74
Şekil 4.3: Genişletilmiş Cox model için bir örnek.....	79
Şekil 4.4: Zamana göre artan tehlike orantısı	82
Şekil 5.1: Firmaların kuluçkada kalma sürelerinin dağılımı (ay)	92
Şekil 5.2: Kuluçka içerisindeki firmaların hayatta kalma eğrisi (ay)	93
Şekil 5.3: (tekgelir) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi	101
Şekil 5.4: (gender) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi	102

Şekil 5.5: (enteduuni) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi	102
Şekil 5.6: (ailecev) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi	103
Şekil 5.7: (ortaksay) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi	103
Şekil 5.8: (export) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi	104
Şekil 5.9: (onlyselffin) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi	104
Şekil 5.10: (networking) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi	105
Şekil 5.11: (yenilik) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi	105
Şekil 5.12: (reklam) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi	106
Şekil 5.13: (marka) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi	106
Şekil 5.14: (buro) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi	107
Şekil 5.15: (whenest) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi	107
Şekil 5.16: (sektor) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi	108
Şekil 5.17: (cycle) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi	108
Şekil 5.18: (gender) değişkeni için gözlenene karşı beklenen hayatta kalma olasılıkları ile PH varsayımının değerlendirilmesi	109
Şekil 5.19: (enteduuni) değişkeni için gözlenene karşı beklenen hayatta kalma olasılıkları ile PH varsayımının değerlendirilmesi	110
Şekil 5.20: (ailecev) değişkeni için gözlenene karşı beklenen hayatta kalma olasılıkları ile PH varsayımının değerlendirilmesi	110
Şekil 5.21: (ortaksay) değişkeni için gözlenene karşı beklenen hayatta kalma olasılıkları ile PH varsayımının değerlendirilmesi	111

Şekil 5.22: (export) değişkeni için gözlenene karşı beklenen hayatta kalma olasılıkları ile PH varsayımının değerlendirilmesi	112
Şekil 5.23: (yenilik) değişkeni için gözlenene karşı beklenen hayatta kalma olasılıkları ile PH varsayımının değerlendirilmesi	112
Şekil 5.24: (reklam) değişkeni için gözlenene karşı beklenen hayatta kalma olasılıkları ile PH varsayımının değerlendirilmesi	113
Şekil 5.25: (marka) değişkeni için gözlenene karşı beklenen hayatta kalma olasılıkları ile PH varsayımının değerlendirilmesi	114
Şekil 5.26: (whenest) değişkeni için gözlenene karşı beklenen hayatta kalma olasılıkları ile PH varsayımının değerlendirilmesi	114
Şekil 5.27: (cycle) değişkeni için gözlenene karşı beklenen hayatta kalma olasılıkları ile PH varsayımının değerlendirilmesi	115
Şekil 5.28: (tekgelir) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi	117
Şekil 5.29: (gender) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi	118
Şekil 5.30: (Inentage) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi	118
Şekil 5.31: (enteduuni) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi	119
Şekil 5.32: (workexp) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi	119
Şekil 5.33: (ailecev) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi	120
Şekil 5.34: (ortaksay) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi	120
Şekil 5.35: (export) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi	121
Şekil 5.36: (Inempini) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi	121
Şekil 5.37: (onlyselffin) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi	122
Şekil 5.38: (networking) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi	122

Şekil 5.39: (yenilik) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi	123
Şekil 5.40: (reklam) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi	123
Şekil 5.41: (marka) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi	124
Şekil 5.42: (buro) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi	124
Şekil 5.43: (whenest) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi	125
Şekil 5.44: (incubsize) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi	125
Şekil 5.45: (sektor) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi	126
Şekil 5.46: (compete) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi	126
Şekil 5.47: (ilrank) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi	127
Şekil 5.48: (cycle) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi	127

TABLOLAR LİSTESİ

Tablo 1.1: Sansürlü ve başarısız deneklerin özet bir gösterimi	9
Tablo 1.2: Genel ve Alternatif Veri Düzeni	12
Tablo 1.3: Belirli bir zamandaki hayatta kalma olasılığının tahminlenmesi	12
Tablo 1.4: Farklı model tipleri için çıktı türleri	16
Tablo 1.5: Çeşitli t değerleri için hayatta kalma fonksiyonları	19
Tablo 2.1: Tehlike fonksiyonunu oluşturan unsurlar arasındaki farklar.....	38
Tablo 2.2: Bir katsayının anlamlılığını değerlendirmede kullanılan yöntemler.....	43
Tablo 3.1: Alternatif veri düzeninin genişletilmesi	60
Tablo 3.2: Çeşitli test istatistikleri için kullanılan ağırlıklar	63
Tablo 4.1: Zamana bağlı ve zamandan bağımsız değişkenler	70
Tablo 5.1: İller itibariyle araştırmaya dahil olan firma sayıları.....	87
Tablo 5.2: Analizde kullanılan değişkenler ve tanımlayıcı istatistikler i	89
Tablo 5.3: Analizde kullanılan kuluçka hizmetleri değişkenlerinin (buro, ortakhiz, danismanlik) alt-unsurları	91
Tablo 5.4: Hayatta kalma üzerinde etkili olan girişimci tecrübelerini, sermaye yoğunluğunu, kuluçka hizmetlerini tek tek ele alan Cox model tahminleri	94
Tablo 5.5: Hayatta kalmayı etkileyen girişimci tecrübesi (workexp) olarak alındığında, sermaye yoğunluğu, kuluçka hizmetlerini aynı anda ele alan Cox model tahminleri.....	97
Tablo 5.6: Hayatta kalmayı etkileyen girişimci tecrübesi (entexp) olarak alındığında, sermaye yoğunluğu, kuluçka hizmetlerini aynı anda ele alan Cox model tahminleri.....	98
Tablo 5.7: Hayatta kalmayı etkileyen girişimci tecrübesi (mangexp) olarak alındığında, sermaye yoğunluğu, kuluçka hizmetlerini aynı anda ele alan Cox model tahminleri.....	99
Tablo 5.8: PH varsayımının istatistiksel bir yöntemle test edilmesi	116
Tablo 5.9: Modeldeki değişkenler için PH varsayımının topluca değerlendirilmesi	128
Tablo 5.10: Zamana bağlı değişkenleri içeren Cox model tahminleri.....	128
Tablo 5.11: Zamana bağlı değişkenlerin belirlenmesinde kullanılan yöntemlerin özet sonuçları	131

KISALTMALAR LİSTESİ

KM: Kaplan-Meier

PH: Orantılı Tehlike

HR: Tehlike orantısı

OR: Odds (bahis) oranı

ML: Ençok olabilirlik

LR: Olabilirlik istatistiği

GOF: Uyum iyiliği

ÖZET

Hayatta kalma verisinin çok deęişkenli analizlerinde en çok kullanılan model Cox tarafından önerilen orantılı tehlike (hazard) modelidir. Modele ilişkin sonuçların elde edilmesi ve yorumlanması kolay olmakla birlikte orantılı tehlike modelinin temel varsayımı, yaşam süresi üzerinde etkili olduęu düşünölen deęişkene dayalı olarak elde edilen tehlike orantısının gözlem periyodu boyunca sabit olmasıdır. Bu varsayımın geçerli olmadığı durumlarda model, sapmalı sonuçlar verebilmektedir. Orantılı tehlike varsayımının sağlanmadığı durumda tehlike orantısını modellemek üzere kullanılan metotlardan biri tahmin edici ile zamanın parametrik bir fonksiyonu arasındaki etkileşimi gösteren zamana baęlı bir deęişkenin modelde yer almasıdır. Bu çalışmada, Türkiye’de faaliyet gösteren önemli iş kuluçkaları uygulamalarından birisi olan İş Geliştirme Merkezleri (İŞGEM)’nde yer alan firmaların hayatta kalma süresi üzerinde etkili olan faktörler tespit edilmeye çalışılmıştır. Çalışmada kullanılan veri seti, “İş Kuluçkalarında Yeni Kurulan Girişimlerin Hayatta Kalma ve Büyüme Performansını Etkileyen Faktörler: KOSGEB İş Geliştirme Merkezleri (İŞGEM) Üzerine Bir Araştırma” konulu TÜBİTAK projesinden elde edilmiştir.

Anahtar Kelimeler: Hayatta kalma analizleri, Cox regresyon modeli, orantılı tehlike (hazard) varsayımı, genç firmalar

ABSTRACT

The most widely used model in multivariate analysis of survival data is proportional hazards model proposed by Cox. While it is easy to get and interpret the results of the model, the basic assumption of proportional hazards model is that the hazard ratio obtained from the variable thought to have an effect on survival time is constant throughout the observation period. Model can give biased results in cases which this assumption is violated. One of the methods used to model the hazard ratio in the cases that the proportional hazard assumption is not met is that to add a time-dependent variable showing the interaction between the predictor variable and a parametric function of time. In this study, it is tried to determine the factors that affect the survival time of the firms are included in Business Development Centers (İŞGEM) which is one of the practices of an important business incubators operating in Turkey. The data set used in this study has been obtained from the TÜBİTAK project on “The Factors Affecting Survival and Growth Performance of Newly Established Enterprises in Business Incubators: A Survey on the KOSGEB Business Development Centers (İŞGEM)”.

Key Words: Survival analysis, Cox regression model, propotional hazard assumption, young firms

TEŞEKKÜR

Hayatımın başlangıcından şu anki zamana ulaşmama dek geçen süre içinde, benliğimin oluşmasına tüm varlığıyla emek verip sonsuz fedakârlıklar gösteren canım anneme ve canım babama,

Ekonometriyi bana tanıtan, bu alanda ilerlemem için gerekli doneleri veren, her zaman örnek aldığım, ufkumu aydınlatan çok saygıdeğer hocalarım; M. Vedat Pazarlıođlu, Şenay Üçdođruk ve Hamdi Emeç'e,

Bu tezi oluşturmada benden hiçbir emeđini esirgemeyen, çalışmadaki veri setini kullanmama olanak sağlayan, eşsiz deneyim ve tecrübelerini paylaşarak beni yönlendiren ve kısa sürede bana çok değer katan, çalışma azmi ve başarılarını gıpta ile örnek aldığım kıymetli hocam Murat Karaöz'e,

Bir parçası olmaktan gurur ve mutluluk duyduğum bölümümüzdeki değerli hocalarım; İbrahim Güngör, Emre İpekçi Çetin, Fahriye Uysal ve Çiğdem Demir'e,

Bugüne kadar bana emek veren diđer bütün hocalarıma,

Ayrıca bu çalışma TÜBİTAK-1001 araştırma destek programı kapsamında, bütün yönleriyle TÜBİTAK Sosyal ve Beşeri Bilimler Araştırma Grubu (SOBAG) tarafından desteklenmiş olan bir araştırma projesinin sonuçlarının bir kısmını oluşturduğundan projeye sağladığı katkı ve bana verdiği burs imkanı için TÜBİTAK'a,

Teşekkürlerimi sunarım.

GİRİŞ

Cox model yaygın olarak kullanılmasına rağmen (Cox (1972) tarafından orijinal makalenin yayınlanmasından beri 25000'den fazla alıntı yapılmıştır.) güncel yayınlar, Cox model uygulamalarının kalitesinde artan bir paya sahiptir (Bellera vd., 2010).

İnsan ve çevresindeki birimlerin başarısızlık (failure) ya da ölüm zamanının istatistiksel analizine ilişkin çalışmalar ilk olarak yaşam tablosu (life-table) analizi ile ele alınmaya başlamıştır. Bu çalışmalar daha sonra geliştirilerek başarısızlık modeli ya da tehlike (hazard) modeli olarak adlandırılmıştır. Bu modeller, iyi tanımlanmış herhangi bir olayın olma ya da gözlemlene zamanının analiz edilmesinde kullanılmıştır (Sertkaya, Ata ve Sözer, 2005).

20. yüzyılda başlayan hayatta kalma analizi çalışmaları, bu yüzyılın ikinci yarısı boyunca büyük ilerlemeler göstermiştir. Bu alandaki en etkili gelişmeler (Sertkaya, Ata ve Sözer, 2005);

- Hayatta kalma fonksiyonunun tahmini için kullanılan “Kaplan-Meier Metodu”,
- İki hayatta kalma dağılımının karşılaştırılması için kullanılan “Log-Rank Test İstatistiği”,
- Hayatta kalma süresi üzerinde açıklayıcı değişkenlerin etkilerini ölçebilmek için kullanılan “Cox Regresyon Modeli”,
- Hayatta kalma analizi istatistiklerinin küçük ve büyük örneklem özelliklerini ortaya koyabilmek için birleşik bir yapı sağlayan “Sayma Süreci Martingale Kuramı” ile başlamış ve günümüze kadar gelmiştir.

Hayatta kalma analizi, belirli bir olayın belli bir zamanda gerçekleşme olasılığını gösteren analizdir (Cox ve Oakes, 1984). Küçük işletme ekonomisi ve girişimcilik literatüründe yeni doğan firmaların yıllar itibariyle hayatta kalma ya da kapanmama oranlarını gösteren çalışmalarda kullanılmaktadır. Örneğin ampirik olarak gözlenen bir grup yeni kurulmuş işyerinin kurulduğu günden itibaren yıllar boyunca kapanıp kapanmadığı ve eğer kapanmamış ise hangi yıl yüzde kaçının hayatta kalmaya devam ettiği hayatta kalma eğrileri ile gösterilmektedir. Gözlenen grupta yer alan işyerlerinin bir kısmı birinci yıl içerisinde kapanırken, devam eden yıllar içerisinde de gruptaki işyerlerinin bir kısmı peyderpey kapanmaya devam etmektedir. Dolayısıyla ele alınan grup için doğumdan itibaren, yıllar itibariyle, her bir yıl için hayatta kalma oranı elde edilmektedir. Hayatta kalma oranı grubun

ilk kurulduğu yıl %100 olurken yıllar geçtikçe %0 değerine basamak basamak yaklaşmaktadır. Dolayısıyla hayatta kalma eğrisi basamaklı bir eğri olup hangi yıl grupta yer alan işyerlerinin % kaçının hayatta kalmaya devam ettiğini göstermektedir (Karaöz ve Albeni, 2011, s.10)

Hayatta kalma verisi ile yapılan birçok çalışmada, uygulanan herhangi bir işlemin etkili olup olmadığı ya da hayatta kalma üzerinde etkili olduğu düşünülen ve çalışmanın başında kaydedilen değişkenlerin ya da risk faktörlerinin düzeyleri arasında farklılık olup olmadığı araştırılmak istenmektedir. Cox orantılı tehlike (PH) modeli hayatta kalma süresi üzerinde değişkenlerin etkisini araştırmak üzere en çok tercih edilen modeldir. Cox modelin en önemli varsayımı, faktörlerin farklı düzeylerine ilişkin tehlike orantısının izleme dönemi boyunca sabit olmasıdır. Bu oran, değişkenin bir düzeyinin diğerine göre göreceli riski (relative risk) veya tehlike orantısı (hazard ratio) olarak tek bir sayıyla ifade edilebilmektedir (Başar, 2006).

Faktörlerin farklı düzeylerine ilişkin yaşam sürdürme hızlarının eşit olup olmadığını test etmek için genellikle log-rank testi kullanılmaktadır. Cox PH modelinde olduğu gibi kesin bir varsayım olmamakla birlikte log-rank istatistiğinin testi, risk faktörlerinin farklı düzeylerine ilişkin tehlike hızlarının zaman boyunca değişmez olduğu varsayımına duyarlıdır (Başar, 2006).

Orantılı tehlike varsayımı altında hayatta kalma eğrilerinin kesişmesi mümkün değildir. Faktörlerin düzeyleri bakımından farklılık görülmediği durumlarda hayatta kalma eğrilerinin kesişmesi orantılı tehlike varsayımının bozulduğuna işaret etmektedir. Böyle bir durumda faktör düzeyleri bakımından hayatta kalma eğrilerindeki farklılığa ilişkin log-rank testi yapmak muhtemelen anlamlı olmayacaktır. Eğer bir ya da daha çok değişkene ilişkin orantılı tehlike varsayımı bozulursa, çok değişkenli Cox PH regresyon modeli tahminleri yanlış değerlendirmelere neden olacaktır. Ayrıca elde edilen tahminler yanlış olabilmektedir. Bu durumda özel bir değişkene ilişkin tehlike orantısı, göreceli risk olarak ifade edilememektedir (Başar, 2006).

Orantılı tehlike varsayımının bozulduğu bir değişkenin etkisini modellemek için kullanılan yöntemlerden biri, değişken ile zamanın bir fonksiyonu olan $f(t)$ arasındaki etkileşimi gösteren zamana bağlı (time-dependent) bir değişkenin modelde yer almasıdır. Bu fonksiyonun biçimi, izleme dönemi boyunca tehlike orantısındaki değişimi yansıtmaktadır. Moreau (1985) tarafından önerilen bir diğer yöntem parçalı (piecewise) orantılı tehlike modeli kullanmaktır. Bu modelde tehlike orantısı önceden belirlenen her bir zaman aralığında sabit, ancak aralıklar arasında farklı olan bir adım (step) fonksiyonudur (Başar, 2006).

Bu çalışmanın birinci bölümünde; hayatta kalma analizinin tanımı yapılarak analizde kullanılan değişkenlerin özellikleri ve gözlem türleri açıklanmıştır. Ayrıca hayatta kalma analizinin amaçları, hayatta kalma süresi üzerinde etkili olabilen karıştırıcı faktörler ve hayatta kalma analizinde kullanılan fonksiyonlar tanımlanmıştır.

İkinci bölümde; Cox orantılı tehlike modeli tanımlanmış ve modelin yapısı, varsayımları, model için kullanılan test istatistikleri açıklanmıştır.

Üçüncü bölümde; Cox orantılı tehlike modelinin temel varsayımlarından biri olan PH varsayımının önemi vurgulanarak bu varsayımın test edilmesinde kullanılan grafiksel ve istatistiksel yöntemler ele alınmıştır.

Dördüncü bölümde; PH varsayımını sağlamayan zamana bağlı ve zamandan bağımsız değişkenler için Cox modelin nasıl genişletileceği açıklanmıştır.

Beşinci bölümde; İş Geliştirme Merkezlerinde yer alan firmaların hayatta kalma süresini etkileyen faktörleri ve bu faktörlerin zamana bağlı etkisini içeren bir genişletilmiş Cox model kurulmuş ve yorumlanmıştır.

Son olarak altıncı bölümde; İŞGEM'de yer alan firmaların hayatta kalma sürelerini etkileyen faktörler özetlenmiş ve analiz sonuçlarına göre firmaların yaşam süresini arttırmak için yapılması gerekenler belirtilmiştir.

BİRİNCİ BÖLÜM

HAYATTA KALMA ANALİZİ

Birçok çalışmada ilgilenilen başlıca değişken, kanser teşhisinden ilgilenilen belirli bir “olay” a kadar geçen zaman süresi gibi bir vadedir. Bu olay sıklıkla ölüm olabilmektedir. Bu nedenle böyle verilerin analizi, hayatta kalma (yaşam sürdürme) analizi olarak adlandırılmaktadır (Bellera vd., 2010).

Tıbbi araştırmalarda klasik analiz yöntemleri yetersiz kalmaktadır. Bunun birinci nedeni, çoğu kez araştırmanın değerlendirilmesinin tüm hastalar ölmeden veya incelenen sonuç ortaya çıkmadan yapılması gerektiği, aksi durumda hangi tedavi yönteminin daha iyi olduğunun ve hastalığı etkileyen faktörlerin saptanmasının yıllar alabileceğidir. İkinci nedeni ise, hastalara uygulanan tedavilerin aynı zamanda başlamamasıdır. Bu tip çalışmalarda hayatta kalma analizi yöntemleri daha uygun sonuçlar vermektedir (Sertkaya, Ata ve Sözer, 2005).

1.1. Hayatta Kalma Analizi Nedir?

Hayatta kalma analizi genel olarak, ilgilenilen çıktı değişkeni *bir olay gerçekleşene kadarki zaman* olan verilerin analizi için istatistiksel yöntemlerinin bir toplamıdır (Harrell, 2001, s.389).

Zaman ile; bir bireyin takibinin başından bir olay olana kadarki yıllar, aylar, haftalar ya da günler kastedilmektedir. Alternatif olarak zaman, bir olayın olduğu zamanda bireyin yaşını da gösterebilmektedir (Kleinbaum ve Klein, 2005, s.4).

Hayatta kalma analizinde, zaman değişkeni genellikle **yaşam süresi** (hayatta kalma veya başarısızlık süresi) olarak adlandırılmakta ve bir bireyin takip periyodunda “sağ kaldığı” zamanı vermektedir (Geiss vd, 2009). Ayrıca olay, çoğunlukla bir **başarısızlık** olarak da ifade edilebilmektedir. Çünkü bu durumda, ilgilenilen olay genellikle ölüm, hastalık sıklığı ya da diğer bazı olumsuz kişisel tecrübelerdir. Ancak yaşam süresi, başarısızlık durumunun olumlu bir olay olduğu “cerrahi bir müdahaleden sonra işe geri dönme zamanı” da olabilmektedir (Kleinbaum ve Klein, 2005, s.4).

Olay ile; ölüm, hastalık sıklığı, hastalığın nüksetmesi, iyileşme (örneğin; işe dönme) ya da bir bireyde gerçekleşip gerçekleşmemesiyle ilgilenilen kararlaştırılmış herhangi bir deneyim ifade edilmektedir (Kleinbaum ve Klein, 2005, s.4).

Aynı analizde birden daha fazla olay ele alınabilmesine rağmen, ilgilenilmesi kararlaştırılan sadece bir olay olduğu varsayılmaktadır. Birden daha fazla olay ele alındığında (örneğin; herhangi birkaç sebebe bağlı ölüm) istatistiksel problem, bir tekrarlanan olay (recurrent events) ya da bir yarışan risk (competing risk) problemi olarak karakterize edilebilmektedir (Kleinbaum ve Klein, 2005, s.4).

Hayatta kalma analizi, belirli bir başlangıç noktasından başlayan (hastalığın başlangıcı, tedavinin başlangıcı) izleme sürecindeki olgunun sonuca varmasıdır. Analiz; ölüm veya tedaviye yanıt almak gibi farklı durumla sonuçlanabilmektedir. İzleme sürecinde, izlenen gözlemlerden bazılarının hastalık dışındaki nedenler yüzünden ölmesi veya araştırmayı yapan kişi tarafından izlenmekten vazgeçilmesi durumları söz konusu olabilmektedir. Bu gibi durumlarda analizin sonuçlanamaması son derece olağandır. Bu tür gözlemlere **sansürlü (censored)** gözlemler denilmektedir. Sansürlü veriler gözlemlerin tamamının değil bir kısmının bilindiği veriler olarak da açıklanabilmektedir (Hougaard, 2000).

Hayatta kalma analizi, ilgilenilen olayın gerçekleşmesine kadar geçen süreyi inceleyen verilerin analizinde kullanılmaktadır. Analizin bağımlı değişkeni; çalışmanın sona ermesi, kişinin takibinin kaybedilmesi ya da kişinin çalışmadan ayrılması durumlarından birinin gerçekleşmesine kadar geçen süre olan “yaşam süresi”dir.

Bağımlı değişken, ölüme kadar geçen süre, hamile kalıncaya kadar geçen süre vb. ile örneklenebilir. Olağandışı bir durumla karşılaşmadıkça bağımlı değişkenin gözlenememesi (sansürlü olması) söz konusu olmamaktadır. Bununla birlikte hayatta kalma analizi, bazı gözlemler sansürlü olsa da uygulanabilmektedir. Bu, hayatta kalma analizinin güçlü yanlarından biri olarak kabul edilmektedir (Yay, Çoker ve Uysal, 2007).

Hayatta kalma analizi problemlerine çeşitli örnekler verilebilir. İlk olarak; hastalığın gerileme dönemindeki lösemi hastalarını hafta bazında gerileme döneminde ne kadar kalacağını saptamak için takip eden bir çalışma örnek olarak verilebilir. İkinci örnek, bir hastalığın takip edilmesidir; kalp hastalığının nüksetme durumunu karşılaştırmak için yıllara göre serbest bir birey grubunu takip eden bir çalışma yapılabilir. Üçüncü örnek, yaşlı bir popülasyonun (yaşları 60’tan büyük olan) hayatta kalma süresini izlemek için 13 yıl boyunca takip eden bir çalışmadır. Dördüncü örnek şartlı salınan kişilerin tekrar tutuklanıp tutuklanmayacağını görmek için birkaç hafta takip edilmesidir. Bu problem türü, yeniden suça

karışma çalışması olarak adlandırılmaktadır. Beşinci örnek; kalp naklinden sonra hastaların ne kadar yaşayacağıının izlenmesi olarak verilebilir (Kleinbaum ve Klein, 2005, s.5).

Yukarıdaki örneklerin tümü hayatta kalma analizidir. Çünkü çıktı değişkeni bir olay olana kadar geçen zamandır. Lösemi hastalarını içeren ilk örnekte, ilgilenilen olay (yani başarısızlık) “hastalığın gerileme dönemi dışına çıkması”dır ve çıktı “bir kişinin hastalığının gerileme dönemi dışına çıkmasına kadarki haftalık süre”dir. İkinci örnekte, olay “kalp hastalığının nüksetmesi”dir ve çıktı “bir kişinin kalp hastalığı nüksedinceye kadar geçen süre”dir. Üçüncü örnekte, olay “ölüm”dür ve çıktı “ölüme kadarki yıllık süre”dir. Örnek dört yeniden suça karışma olayını (yani; tekrar tutuklanmayı) ele almaktadır ve çıktı “tekrar tutuklanana kadarki haftalık süre”dir. Son olarak, beşinci örnek çıktısı “ölüme kadar geçen süre (kalp naklinden sonraki aylarda)” olan “ölüm” olayını ele almaktadır (Kleinbaum ve Klein, 2005, s.5).

Hayatta kalma analizinin bağımsız değişkenleri normal dağılmazlar; ayrıca bu değişkenler arasında oransal ilişkiler söz konusudur. Yani değişkenler bağımsız değildir. Değişkenlerin bu özelliklerinden dolayı hayatta kalma verilerinin analizinde klasik çoklu regresyon değil, Cox regresyon yöntemi kullanılmaktadır (Yay, Çoker ve Uysal, 2007).

Hem Kaplan-Meier (1958) metodu hem de Cox orantılı tehlike (PH) modeli sansürlü veriyi analiz etmeye ve bir deneğin gelecekteki bir t zamanından sonra hayatta kalmasının olasılığı olan $S(t)$ hayatta kalma olasılığını tahmin etmeye izin vermektedir. İstatistiksel olarak bu olasılık, hayatta kalma fonksiyonu $S(t) = P(T > t)$ tarafından sağlanmaktadır. Burada T , hayatta kalma süresidir. KM metodu, hayatta kalma olasılıklarını parametrik olmayarak yani temel bir fonksiyon kalıbı varsaymadan tahmin etmektedir. Gruplar arasındaki hayatta kalma dağılımlarını karşılaştırmak için, log-rank ve Mann-Whitney-Wilcoxon gibi çeşitli testler (Gehan, 1965; Mentel, 1966) mevcuttur. Cox PH modeli çarpımsal risk faktörlerini eşanlı olarak hesaplamaktadır. Cox PH modeli hiçbir dağılım varsayımı yapmamaktadır. Ancak olayın anlık gerçekleşme sıklığı düzeyi, zamanın ve risk faktörlerinin bir fonksiyonu olarak modellenmektedir (Bellera vd., 2010).

Hayatta kalma analizinin amaçları; farklı zamanlarda hayatta kalma olasılığı tahminlerinin elde edilmesi, yaşam süresinin dağılımının tahmin edilmesi, farklı hasta gruplarının yaşam süresi dağılımlarının karşılaştırılması olarak sıralanabilmektedir (Collett, 1994).

Hayatta kalma analizinin temel amaçları aşağıdaki gibi üç başlık altında toplanabilmektedir (Yay, Çoker ve Uysal, 2007):

- Hayatta kalma karakteristiklerini tahmin etmek ve yorumlamak için Kaplan Meier (KM) yöntemi ve grafikleri, medyan tahmini ve güven aralıkları,
- Farklı gruptaki yaşam sürelerini karşılaştırmak için log-rank testi,
- Bağımsız değişkenler ile yaşam süresi arasındaki ilişkinin incelenmesi için Cox regresyon modeli.

1.1.1. Gözlenen Süre ve Gözlemlerin Sansürlenmesi

Hayatta kalma analizlerinin çoğu **sansürleme** (censoring) olarak adlandırılan temel bir analitik problemi göz önünde bulundurmalıdır. Özünde, sansürleme bireyin yaşam süresi hakkında bazı bilgilere sahip olduğunda fakat **yaşam süresi kesin olarak bilinmediğinde** olmaktadır (Kleinbaum ve Klein, 2005, s.5).

İlgilenilen olay istatistiksel analiz süresinde meydana gelmeyebilmekte ya da benzer olarak olay gözlenmeden önce bir deneğin takibi kaybedilebilmektedir. Böyle bir durumda, analiz zamanında ya da hastanın takibi kaybedildiği zamanda veri, sansürlü olarak adlandırılmaktadır. Kesin tarihi bilinmemesine rağmen, olayın sansürleme zamanından daha sonra olduğu bilindiğinden sansürlü veri, hala biraz bilgi vermektedir (Bellera vd., 2010).

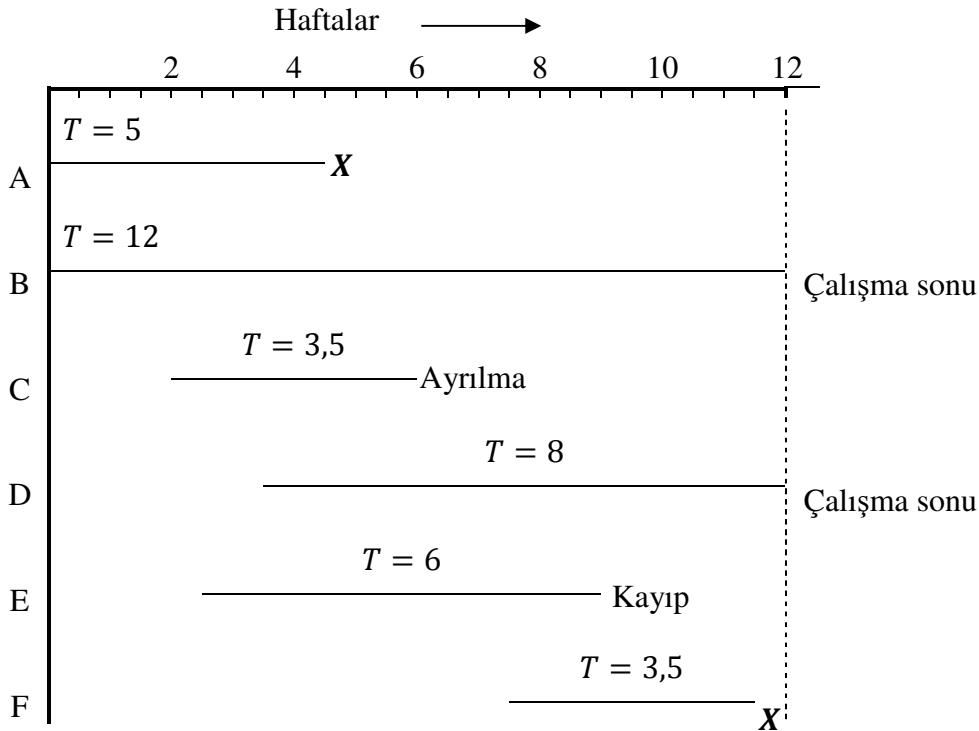
Sansürlemeye basit bir örnek olarak, hastalığı gerileme dönemi dışına çıkana kadar takip edilen lösemi hastaları ele alınabilir. Eğer belirli bir hasta için, kişi hala hastalığın gerileme döneminde iken (yani; olay gerçekleşmeden) çalışma sonlarsa, o hastanın hayatta kalma süresi sansürlü olarak ele alınmaktadır. Bu kişi için yaşam süresinin, en az kişinin takip edildiği periyot kadar uzun olduğu bilinmektedir fakat kişi hastalığın gerileme dönemi dışına çalışma sonlandıktan sonra çıkarsa tam yaşam süresi bilinmeyecektir (Kleinbaum ve Klein, 2005, s.6).

Genel olarak sansürlemeye neden olabilen üç sebep vardır (Kleinbaum ve Klein, 2005, s.6);

- (1) kişinin **çalışma sonlanmadan** önce olayı yaşamaması;
- (2) kişinin çalışma periyodu süresinde **takibinin kaybedilmesi**;
- (3) bir kişinin ölüm nedeniyle (eğer ölüm ilgilenilen olay değilse) ya da başka sebeplerle (örneğin; ilacın yan etkisi ya da diğer yarışan riskler) **çalışmadan ayrılması**

Bu durumlar grafiksel olarak Şekil 1.1’de gösterilmektedir. Grafik, takip edilen çeşitli bireylerin zaman içinde karşılaştığı olayları tanımlamakta ve **X**, yaşayan bir bireyi göstermektedir (Kleinbaum ve Klein, 2005, s.6);

Şekil 1.1: Sansürlü ve başarısız deneklerin gösterimi



Örneğin, çalışmanın başından 5. haftada olay olana kadar takip edilen A kişinin yaşam süresi 5 haftadır ve sansürlü değildir. B kişisi de çalışmanın başından itibaren gözlenmekte fakat 12 haftalık çalışma periyodunun sonuna kadar olay olmadan takip edilmektedir; burada yaşam süresi sansürlüdür çünkü yaşam süresi için sadece *en az* 12 hafta olduğu söylenebilir. C kişisi çalışmaya ikinci ve üçüncü haftalar arasında girmekte ve 6. haftada çalışmadan ayrılana kadar takip edilmektedir; bu kişinin yaşam süresi 3,5 haftadan sonra sansürlenmektedir. D kişisi 4. haftadan itibaren çalışmanın geri kalanında olay olmadan takip edilmiştir; bu kişinin sansürlü süresi 8 haftadır. E kişisi çalışmaya 3. haftada girmekte ve kişi kaybedildiği 9. haftaya kadar takip edilmektedir; bu kişinin sansürlü süresi 6 haftadır. F kişisi 8. haftada çalışmaya girmekte ve 11,5. haftada olay olana kadar takip edilmektedir. A kişisinde olduğu gibi, burada da sansürlenme yoktur; yaşam süresi 3,5 haftadır (Kleinbaum ve Klein, 2005, s.6-7).

Özet olarak, gözlenen altı kişinin ikisinde (A ve F kişileri) olay gerçekleşmektedir, dördünde ise (B,C,D ve E kişileri) olay sansürlüdür (Kleinbaum ve Klein, 2005, s.7).

Şekil 1.1'deki altı kişi için yaşam süresi verileri, Tablo 1.1'de özetlenmektedir. İkinci sütunda olayın ortaya çıkmasına ya da sansürlü olmasına bağlı olarak her bireyin yaşam süresi verilmektedir. Son sütunda bu sürenin sansürlü olup olmadığı (başarısızlığı 1 ile, sansürlenmeyle 0 ile göstererek) belirtilmektedir. Örneğin, C ve F kişilerinin her ikisinin de

yaşam süresi 3,5 olmasına karşın; C kişinin yaşam süresi sansürlü, F kişinin yaşam süresi sansürlü değildir (Kleinbaum ve Klein, 2005, s.7).

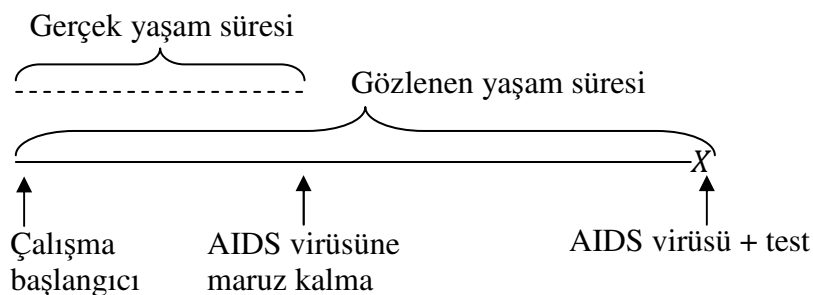
Tablo 1.1: Sansürlü ve başarısız deneklerin özet bir gösterimi

Kişi	Yaşam süresi	Başarısız olma (1); Sansürlü olma (0)
A	5	1
B	12	0
C	3,5	0
D	8	0
E	6	0
F	3,5	1

Örnekte dört kişinin her biri için sansürlenme; çalışma sona erdiğinde, kişinin takibi kaybedildiğinde ya da kişi çalışmadan ayrıldığında ortaya çıkmaktadır. Kişinin kesin yaşam süresinin takip edilen periyodun **sağ** tarafında eksik olduğu görülmektedir. Bu veri çeşidi genel olarak **sağdan sansürlü** olarak adlandırılmaktadır (Oakes, 2000). Bu veriler için bilinmeyen tam yaşam süresi aralığı, gözlenen yaşam süresinin sağ tarafında kesilmektedir (yani sansürlenmektedir). Veriler **soldan sansürlü** de olabilmesine rağmen, hayatta kalma verilerinin çoğu sağdan sansürlüdür (Vittinghoff vd., 2005, s. 211-212).

Soldan sansürlü veriler; kişinin gerçek yaşam süresinin, gözlenen yaşam süresinden daha az ya da gözlenen yaşam süresine eşit olduğu durumda ortaya çıkmaktadır. Örneğin; AIDS virüsü pozitif olana kadar kişiler takip edilirse, bir deneğin virüs için pozitif olduğu ilk test başarısızlık olarak kayıt edilebilir. Ancak, virüse ilk maruz kalma zamanı kesin olarak bilinemez ve dolayısıyla başarısızlığın ne zaman meydana geldiği de kesin olarak bilinmez. Bu yüzden virüse maruz kalındığında sona eren gerçek yaşam süresi, deneğin testi pozitif olduğunda sona eren takip süresinden daha kısa olduğundan; yaşam süresi sol tarafta sansürlenmektedir (Kleinbaum ve Klein, 2005, s.8).

Şekil 1.1: Soldan sansürlü yaşam süresi



Sansürlü gözlemler, bir kişinin yaşam süresi kesin olarak bilinmediği için tamamlanmamış olmasına rağmen, bu sansürlü kişinin izi kaybedilene kadar sahip olunan bilgiden yine de yararlanılabilmektedir. Sansürlü bir kişinin bilgisini tümüyle atmaktan ziyade, böyle bir kişinin sansürlenme zamanına kadar sahip olunan bilgisinin tümü kullanılmaktadır (Kleinbaum ve Klein, 2005, s.23).

1.1.1.1. Gözlenen Süre

Hayatta kalma analizi, her biri için genellikle başarısızlık olarak adlandırılan bir nokta olayı olan bireyler grubu ya da grupları ile ilgilenmektedir. Başarısızlık belirli bir zaman aralığından sonra oluşmakta ve buna başarısızlık zamanı (failure time) denilmektedir. Yaşayan bir organizmanın ya da cansız bir nesnenin belirli bir başlangıç zamanı ile ölümü (başarısızlığı) arasında geçen zamana “yaşam süresi” ya da “başarısızlık süresi” adı verilmekte ve genellikle T ile gösterilmektedir. Başarısızlık süresine örnek olarak, makine bileşenlerinin yaşam süreleri, ekonomide işsizlik dönemleri, psikolojik bir deneyde deneyin belirlenen görevi tamamlama süresi ve klinik bir deneyde hastaların yaşam süreleri gösterilebilir (Sertkaya, Ata ve Sözer, 2005).

Hayatta kalma analizinde temel olan, gözlenen başarısızlık sürelerinin incelenmesi olduğundan, bu değişkenin iyi tanımlanması gerekmektedir. İlgilenilen olaya göre farklı şekillerde ortaya çıkabilen bu değişkenin duyarlı olarak ölçülmesi için (Le, 1997, s.3-4);

- Başlangıç zamanı her bir birim ya da birey için şüpheye yer vermeyecek şekilde tanımlanmalıdır,
- Geçen süreyi ölçmek için bir zaman ölçeği kabul edilmelidir,
- Her bir birim ya da birey için ömrün sona erdiği ya da başarısızlığın meydana geldiği an tamamen net olmalıdır. Diğer bir deyişle, başarısızlığın tam olarak tanımlanması gerekmektedir.

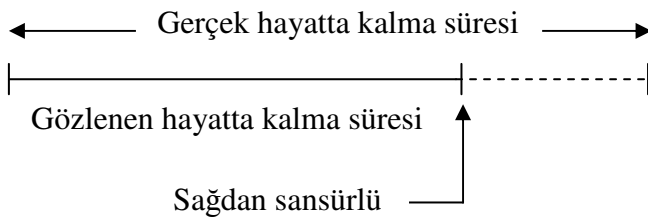
1.1.1.2. Gözlemlerin Sansürlenmesi

Hayatta kalma verilerinin analizinde karşılaşılan temel güçlük, gözlem altına alınan bazı birim ya da bireylerin başarısızlık zamanlarının gözlenememiş olmasıdır. Örneğin, tıbbi bir çalışmada gözlem altına alınan hastaların bazıları deney sonunda hala yaşamlarını sürdürüyor olabilir. Ayrıca, gözlem altındaki bir birey, bazı nedenlerden dolayı gözlemden

çıkabilmektedir. Eğer başarısızlık zamanı, bu gibi nedenlerden dolayı tamamlanamamış ise sansürlü (censored) durum söz konusudur (Sertkaya, Ata ve Sözer, 2005).

Yaşam süresi verilerinin çoğu sağdan sansürlüdür, çünkü kesin olarak bilinmeyen gerçek yaşam süresi aralığı, gözlenen zaman aralığının sağ tarafında kesilmektedir (yani sansürlenmektedir). Bu gözlenen zaman aralığı, gerçek yaşam süresinden daha kısa olan gözlenen yaşam süresini vermektedir. Gerçek yaşam süresi hakkında çıkarım yapmak için gözlenen yaşam süresi kullanılmak istenmektedir (Kleinbaum ve Klein, 2005, s.48).

Şekil 1.2: Sağdan sansürlü yaşam süresi



Hayatta kalma analizi için genel ve alternatif veri düzeni aşağıdaki gibidir. Tablo 1.2'de genel veri düzeninin ilk sütunu çalışma deneklerini tanımlamaktadır. İkinci sütun gözlenen yaşam süresi bilgisini vermektedir. Üçüncü sütun ise sansür durumunu gösteren ikili değişken δ için bilgi vermektedir. Genel veri düzenindeki bilgilerin kalanı, ilgilenilen açıklayıcı değişkenlerin değerlerini göstermektedir. Alternatif veri düzeninde ise; ilk sütun küçükten büyüğe sıralanmış yaşam sürelerini, ikinci sütun her ayrı başarısızlık zamanında, başarısızlığın sıklık sayısını vermektedir. q_j ile gösterilen üçüncü sütun; $t_{(j)}$ başarısızlık zamanıyla başlayan fakat $t_{(j+1)}$ tarafından gösterilen bir sonraki başarısızlık zamanını içermeyen zaman aralığındaki, sansürlü kişilerin sayısını göstermektedir. Son sütun, en az $t_{(j)}$ zamanına kadar yaşayan bireylerin toplamını gösteren **risk setini** vermektedir. Alternatif veri düzeni **Kaplan-Meier** hayatta kalma eğrilerinin türetilmesine temel oluşturmaktadır. (Kleinbaum ve Klein, 2005, s.50).

Tablo 1.2: Genel ve Alternatif Veri Düzeni

Genel Veri Düzeni:					Alternatif (Sıralı) Veri Düzeni:			
Birey	t	δ	X_1	$X_2 \dots X_p$	Sıralı başarısızlık zamanları, $t_{(j)}$	Başarısızlık sayıları, m_j	$[t_{(j)}, t_{(j+1)})$ aralığındaki sansür sayısı, q_j	Risk seti, $R(t_{(j)})$
1	t_1	δ_1	X_{11}	$X_{12} \dots X_{1p}$	$t_{(0)} = 0$	$m_0 = 0$	q_0	$R(t_{(0)})$
2	t_2	δ_2	X_{21}	$X_{22} \dots X_{2p}$	$t_{(1)}$	m_1	q_1	$R(t_{(1)})$
\vdots	\vdots	\vdots	\vdots	\vdots	$t_{(2)}$	m_2	q_2	$R(t_{(2)})$
n	t_n	δ_n	X_{n1}	$X_{n2} \dots X_{np}$	\vdots	\vdots	\vdots	\vdots
					$t_{(k)}$	m_k	q_k	$R(t_{(k)})$

k = deneklerin başarısız olduğu farklı zamanların sayısı; n = deneklerin sayısı ($k \leq n$);

Risk seti $R(t_{(j)})$, yaşam süresi en az $t_{(j)}$ olan bireylerin setidir. ($j \leq k$)

Belirli bir zamandaki hayatta kalma olasılığını tahminlemek için; tüm bilgiyi tümüyle çıkarmaktan ziyade, sansürlenmiş bir kişinin sansürlenme zamanına kadar sahip olduğu bilgiyi içeren zamandaki risk seti kullanılmaktadır (Kleinbaum ve Klein, 2005, s.50);

Tablo 1.3: Belirli bir zamandaki hayatta kalma olasılığının tahminlenmesi

$t_{(j)}$	m_j	q_j	$n_j = n - (m_j + q_j),$ yani $R(t_{(j)})$	$\hat{S}(t_{(j)}) = R(t_{(j)})/n$
$t_{(0)} = 0$	$m_0 = 0$	δ_1	n_0	$\hat{S}(t_{(0)})$
$t_{(1)}$	m_1	δ_2	n_1	$\hat{S}(t_{(1)})$
$t_{(2)}$	m_2	δ_3	n_2	$\hat{S}(t_{(2)})$
\vdots	\vdots	\vdots	\vdots	\vdots
$t_{(k)}$	m_k	δ_k	n_k	$\hat{S}(t_{(k)})$

1.1.2. Hayatta Kalma Analizinin Amaçları

Hayatta kalma analizlerinin temel amaçları şunlardır (Kleinbaum ve Klein, 2005, s.15):

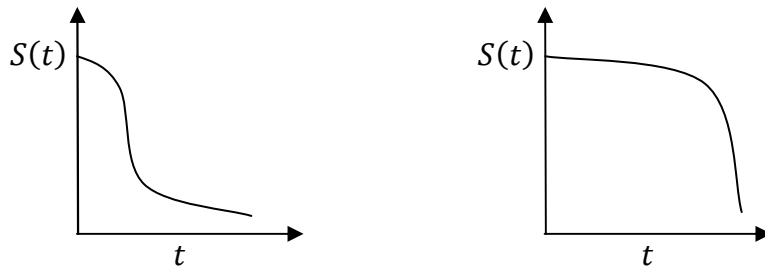
Amaç 1: Yaşam verisinden hayatta kalma ve/veya tehlike fonksiyonlarını tahmin etmek ve yorumlamaktır.

Amaç 2: Hayatta kalma ve/veya tehlike fonksiyonlarını karşılaştırmaktır.

Amaç 3: Açıklayıcı değişkenlerin yaşam süresi ile ilişkisini değerlendirmektir.

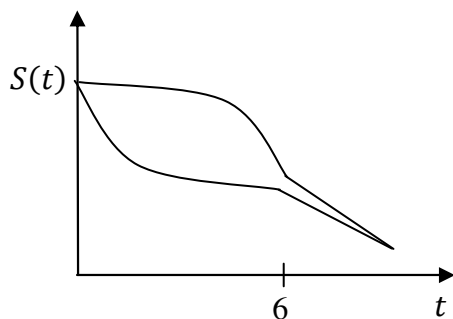
Birinci amaca ilişkin olarak, çok farklı yorumlar veren Şekil 1.4'teki iki fonksiyon ele alınmaktadır. Soldaki fonksiyon, takibin başında hayatta kalma olasılıklarında hızlı bir düşüş göstermekte fakat daha sonra düzleşmektedir. Sağdaki fonksiyon, tam tersine hayatta kalma olasılıklarında takibin başında çok yavaş, daha sonra ise keskin bir düşüş göstermektedir (Kleinbaum ve Klein, 2005, s.15).

Şekil 1.4: Hayatta kalma fonksiyonu için örnekler



Şekil 1.5'teki grafikte, aynı eksen üzerinde çizilen bir deney grubunun hayatta kalma fonksiyonu ile ilaç olarak tesirsiz bir maddenin verildiği üyelerden oluşan bir kontrol (plasebo) grubunun hayatta kalma fonksiyonu karşılaştırılmaktadır. Deney grubu için hayatta kalma fonksiyonunun 6. haftaya kadar kontrol grubu için olan fonksiyonun üzerinde dağıldığı görülmektedir. Fakat daha sonra, bu iki fonksiyon yaklaşık olarak aynı seviyede ilerlemektedir. Şekil 1.5'teki ikili grafik, 6. haftaya kadar hayatta kalma için deneyin plasebodan **daha etkili** olduğunu fakat daha sonra yaklaşık olarak aynı etkiye sahip olduklarını göstermektedir (Kleinbaum ve Klein, 2005, s.15).

Şekil 1.5: Kontrol ve deney grubu için hayatta kalma fonksiyonlarının karşılaştırılması



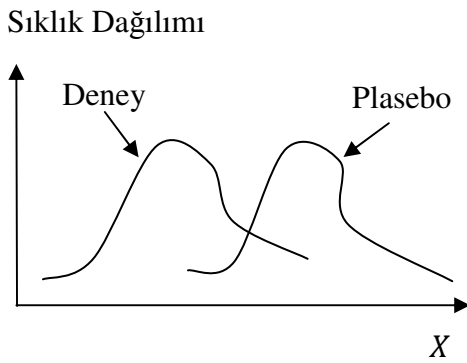
Üçüncü amaç, Cox PH yaklaşımı gibi bir matematiksel modelleme kalıbı kullanmayı gerektirmektedir (Kleinbaum ve Klein, 2005, s.15).

1.1.3. Karıştırıcı Faktörler ve Etkileşim

1.1.3.1. Karıştırıcı Faktörler

Bazı durumlarda ilgilenilen olay birden fazla faktörün etkisi altında olabilmektedir. Aynı zamanda bu faktörler tüm deneklerde aynı etkiyi göstermeyebilir ya da aynı etkiyi göstermesine rağmen aynı bulgular meydana gelmeyebilir. Bu gibi durumlarda **karıştırıcı faktörlerin etkisi** (confounding) söz konusudur. Bu **karıştırıcı faktörler** (confounders) ilgilenilen değişkeni değerlendirmede sistematik hatalara neden olabilmektedir (Çakır, 2002). Karıştırıcı faktörün X olarak adlandırıldığı örnek bir grafik Şekil 1.6'daki gibidir (Kleinbaum ve Klein, 2005, s.27):

Şekil 1.6: Karıştırıcı faktörlerin kontrol ve deney grubuna olan etkisi



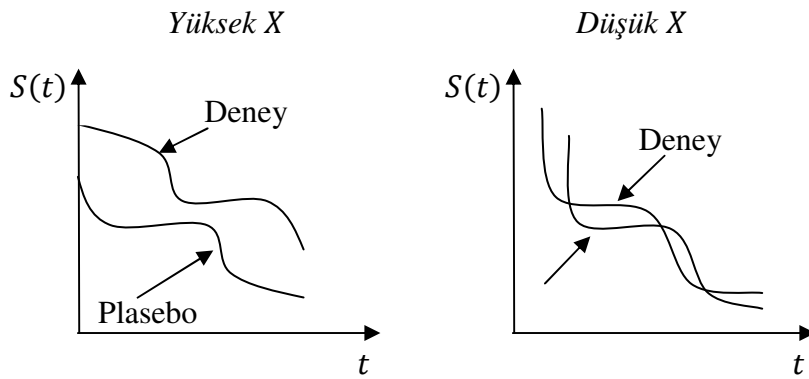
Karıştırıcı faktörler olayın olmadığı durumlarda dahi, etken ile ve etkenden bağımsız olarak olay ile ilgili olan, ancak “etkenin olaya yol açtığı sebeplere bağlı olmayan faktörler” olarak tanımlanmaktadır. Bu tür karıştırıcı faktörlerin en sık gözlenenleri ırksal ve etnik farklılıklar, yaş ve cinsiyettir. Her çalışmada olası karıştırıcı faktörler için mutlaka kontrol sağlanmalıdır. Örneğin, iş ortamındaki stres ile alkol bağımlılığı arasındaki ilişki incelenirken, cinsiyet için kontrol sağlanmalıdır. Erkeklerde, iş ortamına bağlı stresten bağımsız olarak, alkol bağımlılığı kadınlardakinden daha yüksek olabilir. Benzer şekilde, alkol alma durumu ne olursa olsun, erkekler kadınlara kıyasla çalıştıkları işler nedeniyle daha fazla stres duyuyor olabilirler. Kişilerin iş yeri ortamına bağlı stres duymalarının alkol bağımlılığına etkisi, stresin cinsiyet üzerinde yarattığı etkiler üzerinden olamayacağı için, bu araştırmada cinsiyet “karıştırıcı” bir faktör olarak düşünülmeli ve cinsiyet için kontrol sağlanmalıdır. Cinsiyet için ayarlama yapılmaz ise, farklı çalışmalarda araştırma grubundaki erkek/kadın oranına bağlı olarak, işyeri ortamındaki stres ile alkol bağımlılığı arasındaki ilişki farklı bulunacaktır (Çakır, 2002).

Yapılan çalışmalarda bu tür karıştırıcı faktörlerin varlığı incelenmez ve kontrol edilemez ise gerçekte olmadığı halde değişken-olay arasında ilişki olduğu gözlenebilir, gerçekte ilişki olduğu halde çalışma grubunda tespit edilemeyebilir ya da gerçekte var olan bir ilişki saptanır, ancak saptanan olayın toplumdaki gerçek boyutunu yansıtmayabilir. Bu duruma karıştırıcı faktörlerin etkisi (confounding) adı verilmektedir. Karıştırıcı faktörlerin etkisi istatistiksel testlerin kullanılması ile değerlendirilebilmekte ve standardizasyon, tabakalama, kovaryans analizi ve lojistik regresyon benzeri çok değişkenli analizlerin kullanımı ile kontrol edilebilmektedir (Çakır, 2002).

1.1.3.2. Etkileşim

X bağımsız değişkeninin etkisine ilişkin göz önüne alınacak diğer bir problem **etkileşim**dir (interaction). Etkileşim, X bağımsız değişkeninin düzeyine bağlı olarak deneyin etkisinin değişebilmesidir. Etkileşim içine girilen değişken X olarak adlandırıldığında, etkileşim olması (soldaki grafik) ve olmaması (sağdaki grafik) durumuna Şekil 1.7'deki grafikler örnek olarak verilebilir (Kleinbaum ve Klein, 2005, s.27);

Şekil 1.7: X bağımsız değişkeninin etkileşim etkisi



Etkileşim, olayın oluşumunda çok sayıda faktörün rol oynadığı durumlarda, bu faktörlerin aynı anda var olması durumunda olay üzerindeki etkilerinin, her birinin tek başına bulunduğu anda gözlenen etkilerden daha farklı etki oluşturması demektir. İki tür etkileşim söz konusudur (Çakır, 2002):

Çarpımsal (multiplicative, istatistiksel) etkileşim: Her iki faktörün ortak etkisi ayrı ayrı olan etkilerinin çarpımına eşittir veya fazladır. İstatistiksel etkileşim matematiksel olarak modellemede kolaylıkla değerlendirilebilmekte ve açıklanabilmekte; hangi durumlarda etkileşimlerin belirtilmesi gerektiği saptanabilmektedir (Çakır, 2002).

Toplamsal (additive) etkileşim: Her iki faktörün birlikte etkisi, tek tek olan etkilerinden daha büyüktür ancak çarpımlı düzeye varan bir etkileşim yoktur (Çakır, 2002).

Açıklayıcı değişkenleri kapsayan karıştırıcılara ek olarak analizde etkileşimin değerlendirilmesi de her durumda ele alınacak bir konudur (Kleinbaum ve Klein, 2005, s.28).

X bağımsız değişkeninin olası karıştırıcı ya da etkileşim etkisini değerlendirmede, X 'in ayarlanması ve/veya etkisinin değerlendirilmesi için alternatif metotlar aşağıdaki gibidir (Kleinbaum ve Klein, 2005, s.29);

- X bağımsız değişkeni üzerinde tabakalandırma ve farklı tabakalar için yaşam eğrilerini karşılaştırmak ya da
- Orantılı tehlike veya diğer hayatta kalma modelleri gibi matematiksel modelleme yöntemlerini kullanmaktır.

1.2. Hayatta Kalma Analizinde Kullanılan Fonksiyonlar

Yaşam sürdürme analizinde temel olarak; olasılık yoğunluk fonksiyonu, hayatta kalma fonksiyonu, tehlike fonksiyonu kullanılmaktadır (Göz Çekçeki, 2007, s.9).

Modelleme metotları arasındaki önemli fark, kullanılan çıktı değişkeninin türüdür. Hayatta kalma analizinde, çıktı değişkeni “bir olayın gerçekleşmesi için geçen süre” veya sansürlü veridir. Doğrusal regresyon modellemesinde, çıktı değişkeni genellikle kan basıncı gibi sürekli bir değişkendir. Lojistik modellemede, çıktı değişkeni evet ya da hayır gibi ikili bir değişkendir. Doğrusal ve lojistik modellemede, genellikle mevcut takip süresi hakkında bilgi sahibi olunmamaktadır (Kleinbaum ve Klein, 2005, s.32).

Tablo 1.4: Farklı model tipleri için çıktı türleri

Model	Çıktı
Hayatta kalma analizi	Olaya kadarki süre (sansürleyerek)
takip süresi bilgisi kullanılmaz { Doğrusal regresyon Lojistik regresyon	Sürekli
	İkili

Doğrusal ve lojistik modellemede olduğu gibi hayatta kalma analizinin istatistiksel bir amacı, ilgili dışsal değişkenler için ayarlanan olay-çıktı ilişkisini tanımlayan etkilerin bazı ölçümlerini elde etmektir (Kleinbaum ve Klein, 2005, s.32).

Doğrusal regresyon modellemesinde, etkinin ölçümü genellikle bir regresyon katsayısı β 'dır. Lojistik modellemede etkinin ölçümü, modelde bir ya da daha fazla regresyon katsayısının üstü olarak ifade edilen (örneğin; e^β) bir bahis oranı (odds ratio (OR), risk oranı, şans oranı)'dır. Hayatta kalma analizinde, genellikle elde edilen ölçümün etkisi bir **tehlike orantısı** olarak adlandırılmaktadır. Lojistik modeldeki gibi, bu tehlike orantısı modeldeki bir ya da daha fazla regresyon katsayısının üstü olarak ifade edilmektedir (Kleinbaum ve Klein, 2005, s.32).

Etkinin ölçümü:

Doğrusal regresyon: regresyon katsayısı β

Lojistik regresyon: bahis oranı e^β

Hayatta kalma analizi: tehlike orantısı e^β

Tehlike orantısı, bahis oranından biraz farklı bir ölçüm olmasına rağmen etkinin gücüyle ilgili benzer bir yoruma sahiptir. Tehlike orantısının 1 olması, hiçbir etki olmadığı anlamına gelmektedir; yani 1, olay-çıktı ilişkisi için etkisiz bir değerdir. Tehlike orantısının 10 olması; “olayın meydana geldiği grup, olayı tecrübe etmeyen grubun 10 katı kadar tehlikeye sahiptir” şeklinde yorumlanmaktadır. Benzer olarak, tehlike orantısının 1/10 olması olayı tecrübe eden grubun, olayı geçirmeyen grubun onda biri kadar tehlikeye sahip olduğunu kastetmektedir (Kleinbaum ve Klein, 2005, s.33).

HR'nin yorumu (OR gibi):

$HR = 1 \Rightarrow$ ilişki yok

$HR = 10 \Rightarrow$ olayın gerçekleşmesindeki tehlike, gerçekleşmemesi durumundakinin 10 katı

$HR = 1/10$ olayın gerçekleşmesindeki tehlike, gerçekleşmemesi durumundakinin 1/10 katı

Birçok durumda incelenen hayatta kalma süresinin başka faktörler tarafından etkilenebileceği göz önünde bulundurulduğunda, bağımlı değişken olan hayatta kalma süresi üzerinde açıklayıcı değişkenlerin de etkilerinin modellendiği regresyon modelleri hayatta kalma analizinde önemli bir yer tutmaktadır. Örneğin, akciğer kanseri olan hastaların hayatta kalma süreleri üzerine yapılan bir çalışmada hastanın yaşı, hastanın genel fiziksel durumu, tümörün büyüklüğü, hastalığın teşhisinden sonra geçen süre ve diğer ilgili değişkenler açıklayıcı değişkenler olarak alınabilmektedir (Sertkaya, Ata ve Sözer, 2005).

Hayatta kalma analizinde en çok kullanılan regresyon modeli Cox regresyon modelidir. Son yıllarda yapılan çalışmalarda Cox regresyon modeli sabit açıklayıcı

değişkenlerin yanı sıra zamana bağlı açıklayıcı değişkenleri de içeren Cox regresyon modeline genişletilmiştir (Sertkaya, Ata ve Sözer, 2005).

1.2.1. Olasılık Yoğunluk Fonksiyonu

i. birey için yaşam sürdürme zamanı T rassal değişken olmak üzere, $T = t$ 'ye ilişkin olasılık yoğunluk fonksiyonu aşağıdaki şekilde gösterilmekte ve bireyin t anındaki istenilen durumun koşulsuz ortaya çıkma yoğunluğunu vermektedir (Cox ve Oakes, 1984);

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t)}{\Delta t} \quad (1.1)$$

1.2.2. Hayatta Kalma Fonksiyonu

Her hayatta kalma analizinde ele alınan iki sayısal terim bulunmaktadır. Bunlar: $S(t)$ ile gösterilen **hayatta kalma fonksiyonu** ve $h(t)$ ile gösterilen **tehlike fonksiyonudur** (Kleinbaum ve Klein, 2005, s.8).

Hayatta kalma fonksiyonu, hayatta kalma analizi için bir temeldir çünkü t 'nin farklı değerleri için hayatta kalma olasılıklarını elde etmek, yaşam verisinden çok önemli özet bilgiler vermektedir (Kleinbaum ve Klein, 2005, s.9).

Hayatta kalma analizinde T yaşam süresi içinde herhangi bir t_i zamanında ölümle (başarısızlıkla) karşılaşılacağı varsayılmaktadır. Ayrıca her bir gözlem için başarısızlık olarak adlandırılan ve başarısızlık zamanı (failure time) olarak tanımlanan bir nokta olay söz konusudur (Yay, Çoker ve Uysal, 2007).

Bir kişinin yaşam süresi için rassal değişken T ile gösterilmektedir. T zamanı gösterdiğinden, T 'nin mümkün değerleri negatif olmayan tüm sayıları içermekte, yani T sıfıra ya da sıfırdan büyük herhangi bir sayıya eşit olabilmektedir (Kleinbaum ve Klein, 2005, s.8).

Rassal değişken T 'nin ilgilenilen herhangi bir değeri t ile ifade edilmektedir. Örneğin, bir kişinin kanser tedavisi gördükten sonra 5 yıldan daha fazla yaşayıp yaşamadığıyla ilgileniliyorsa, t 5'e eşit olur; o halde T 'nin 5'ten büyük olup olmadığı ($T > t = 5$) araştırılır (Kleinbaum ve Klein, 2005, s.8).

Başarısızlığı ya da sansürlü olmayı gösteren (0,1) rassal değişkeni δ ile gösterilmektedir. Yani, çalışma periyodu boyunca olay olursa başarısızlık için $\delta = 1$ ya da çalışma periyodunun sonuna kadar yaşam süresi sansürlendiyse $\delta = 0$ 'dır. Eğer bir kişi

başarısızlığa uğramadıysa yani çalışma periyodu boyunca olay olmadıysa, sansür **sadece** o kişinin yaşam süresini sürdürme ihtimalidir. Yani $\delta = 0$ ancak ve ancak; bir kişinin çalışma sonuna kadar hayatta kalması, çalışma periyodu boyunca takibinin kaybedilmesi ya da çalışmadan ayrılması durumlarından biri gerçekleşirse olmaktadır (Kleinbaum ve Klein, 2005, s.8).

Yaşam süresi (T), birikimli dağılım fonksiyonu $P(t) = Pr(T \leq t)$ ve olasılık yoğunluk fonksiyonu $p(t) = dP(t)/d(t)$ olan bir rassal değişken olarak tanımlandığında, tehlike fonksiyonunun tümleyeni olan yaşam fonksiyonu $S(t)$ Denklem (1.2) ile açıklanmaktadır (Yay, Çoker ve Uysal, 2007);

$$S(t) = P(T > t) = 1 - P(t) \quad (1.2)$$

Hayatta kalma fonksiyonu $S(t)$, bir kişinin belirli bir t zamanından daha uzun hayatta kalması olasılığını vermektedir. Yani $S(t)$, tanımlanmış bir t zamanını aşan rassal değişken T 'nin olasılığını ifade etmektedir (Kleinbaum ve Klein, 2005, s.9).

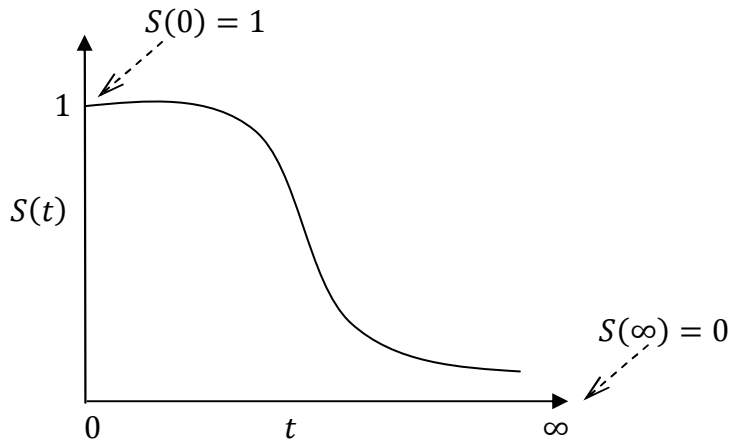
Tablo 1.5: Çeşitli t değerleri için hayatta kalma fonksiyonları

t	$S(t)$
1	$S(1) = P(T > 1)$
2	$S(2) = P(T > 2)$
3	$S(3) = P(T > 3)$
\vdots	\vdots

t , 0'dan sonsuza gittiği için, hayatta kalma fonksiyonu teorik olarak düzgün bir eğri olarak çizilebilmektedir. t 'nin, X eksenini tanımladığı Şekil 1.7'de gösterildiği gibi, tüm hayatta kalma fonksiyonları aşağıdaki özelliklere sahiptir (Kleinbaum ve Klein, 2005, s.9):

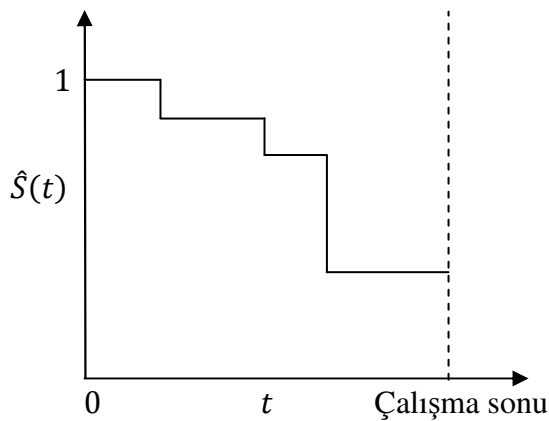
- Artmayan fonksiyonlardır; yani t arttıkça aşağı doğru kaymaktadır.
- $t = 0$ zamanında, $S(t) = S(0) = 1$ 'dir; yani çalışmanın başında henüz hiçbir birey olayı yaşamadığı için, geçen 0 zamanının hayatta kalma olasılığı 1'dir.
- $t = \infty$ zamanında, $S(t) = S(\infty) = 0$ 'dır; yani teorik olarak çalışma periyodu sınırsız olarak artırılırsa, nihayetinde kimse hayatta kalmayacaktır bu yüzden hayatta kalma eğrisi er ya da geç sıfıra düşmektedir.

Şekil 1.8: Hayatta kalma fonksiyonunun grafiği



Bu özelliklerin hepsi uygulamada sağlanmamaktadır. Gerçek veriler kullanıldığında, genellikle düzgün eğrilerden ziyade aşağıda gösterildiği gibi **adım fonksiyonu** olan grafikler elde edilmektedir. Üstelik çalışma periyodunun uzunluğu asla sonsuz olmadığından ve başarısızlık için yarışan riskler olabildiğinden çalışmadaki herkesin olayı yaşamaması mümkündür. Bu yüzden tahminlenen hayatta kalma fonksiyonu, çalışmanın sonunda tamamıyla sıfıra düşmeyebilmektedir. Tahminlenen hayatta kalma fonksiyonu Şekil 1.9'da $\hat{S}(t)$ ile gösterilmektedir (Kleinbaum ve Klein, 2005, s.9).

Şekil 1.9: Tahminlenen hayatta kalma fonksiyonunun grafiği



1.2.3. Tehlike Fonksiyonu

Anlık gerçekleşme sıklığı, ölüm ya da başarısızlık oranı; ya da riski olarak adlandırılan t zamanındaki anlık tehlike oranı, henüz gerçekleşmemiş bir olayın t zamanında tecrübe edilmesinin anlık olasılığıdır. Bu, zamanın her birimi için olayın bir oranıdır ve zamanla

değişmesine izin verilmektedir. Her birim zaman için olayların gerçekleşme riski, zamanın her biriminde kat edilen uzaklığı gösteren bir arabanın hız göstergesi tarafından verilen hız ile benzerlik göstermektedir (Bellera vd., 2010).

$h(t)$ ile gösterilen tehlike fonksiyonu, Denklem (1.3) tarafından verilmektedir. Burada Δt , küçük bir zaman aralığını ifade etmektedir. Bu matematik formülünü pratik terimlerle açıklamak zordur (Kleinbaum ve Klein, 2005, s.9);

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \quad (1.3)$$

Tehlike fonksiyonu $h(t)$; bireyin t zamanına kadar hayatta kaldığı farz edilerek, olayın meydana gelmesi için birim zaman başına anlık potansiyeli vermektedir (Tabatabai vd, 2007). Başarısız olmamaya odaklanan hayatta kalma fonksiyonunun tersine, tehlike fonksiyonu başarısızlığa yani olayın meydana gelmesine odaklanmaktadır. Bu yüzden tehlike fonksiyonu bir bakıma, hayatta kalma fonksiyonu tarafından verilenin ters yönünde bilgi verdiği şeklinde düşünülebilmektedir (Kleinbaum ve Klein, 2005, s.10).

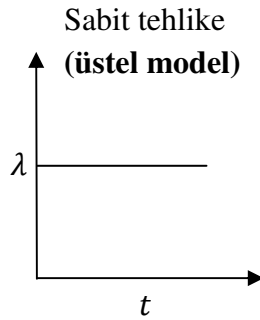
Tehlike fonksiyonu $h(t)$; bireyin t zamanına kadar hayatta kaldığı farz edilerek, ölüm ya da ilgilenilen bir hastalık gibi bir olayın gerçekleşmesi için t zamanındaki anlık potansiyeli vermektedir. Tanımdaki t zamanına kadar hayatta kalma koşulu, hız örneğindeki tanımlamayla benzerdir. Çünkü zaman içinde bir noktadaki hız göstergesi değeri doğasında, göstergenin yazdığı ölçüme kadar zaten biraz mesafe alınmış olduğunu (yani hayatta kalındığını) varsaymaktadır (Kleinbaum ve Klein, 2005, s.10).

Matematiksel anlamda, tehlike fonksiyonu formülünün koşullu kısmı olasılık ifadesinde bulunmaktadır. Bu ifade koşullu bir olasılıktır çünkü “*B gerçekleştiğinde A’nın olasılığı*” biçimindedir ($P(A|B)$). Burada P olasılığı, B’den A’yı ayıran uzun dikey çizgi ise “koşul”u göstermektedir. Tehlike formülünde koşullu olasılık; bir kişinin yaşam süresi T ’nin, t ’ye eşit ya da t ’den daha büyük olması koşuluyla t ve $t + \Delta t$ arasındaki zaman aralığında uzanma olasılığını vermektedir. Formüldeki koşul işaretinden dolayı, tehlike oranı bazen **koşullu başarısızlık oranı** olarak adlandırılmaktadır (Kleinbaum ve Klein, 2005, s.11).

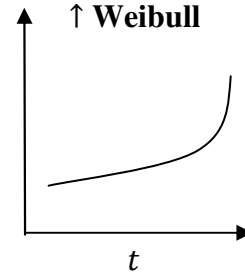
Tehlike fonksiyonlarının farklı tipteki bazı grafikleri Şekil 1.10’daki gibidir (Kleinbaum ve Klein, 2005, s.12);

Şekil 1.10: Farklı tipteki tehlike fonksiyonlarının grafikleri

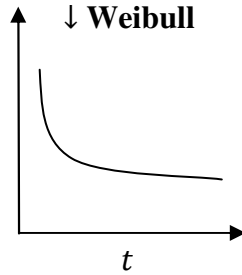
Grafik 1: Sağlıklı kişiler için $h(t)$



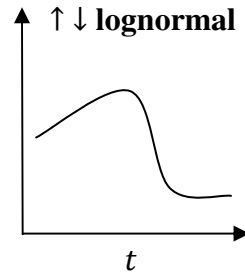
Grafik 2: Lösemi hastaları için $h(t)$



Grafik 3: Ameliyatı atlatan kişiler için $h(t)$



Grafik 4: Verem hastaları için $h(t)$



Grafik 1’de sağlıklı kişilerin bir çalışması için sabit tehlike gösterilmektedir. Bu grafikte, belirlenen t ’nin değeri ne olursa olsun $h(t)$ aynı değere (bu örnekte, λ) eşit olmaktadır. Çalışma periyodu boyunca sağlıklı olmaya devam eden bir kişinin, periyot içinde herhangi bir zamanda hasta olmasının anlık potansiyeli takip edilen periyot boyunca sabit kalmaktadır. Tehlike fonksiyonunun sabit olması, hayatta kalma modelinin **üstel** olduğu anlamına gelmektedir (Kleinbaum ve Klein, 2005, s.12).

Grafik 2, zamanla artan bir tehlike fonksiyonunu göstermektedir. Bu grafik türü, **artan Weibull** modeli olarak adlandırılmaktadır. İlgilenilen olay ölüm olduğunda, tedaviye yanıt vermeyen verem hastaları için böyle bir grafik beklenebilmektedir. Böyle bir hasta için yaşam süresi arttıkça buna bağlı olarak hastalığın seyri kötüleşmekte ve ayrıca hastanın bu hastalıktan ölme potansiyeli de artmaktadır (Kleinbaum ve Klein, 2005, s.13).

Grafik 3’te tehlike fonksiyonu zamana göre azalmaktadır. Bu grafik türü, **azalan Weibull** modeli olarak adlandırılmaktadır. Olay ameliyat geçiren kişilerdeki ölüm olduğunda, böyle bir grafik beklenebilmektedir. Çünkü ameliyattan sonra geçen süre arttıkça ölüm potansiyeli genellikle düşmektedir (Kleinbaum ve Klein, 2005, s.13).

Grafik 4’te grafikte tehlike fonksiyonunun önce arttığı sonra azaldığı görülmektedir. Bu grafik türü, **lognormal hayatta kalma** modeli olarak adlandırılmaktadır. Ölüm potansiyeli

hastalığın başında artıp daha sonra azaldığı için, verem hastaları için böyle bir grafik beklenebilmektedir (Kleinbaum ve Klein, 2005, s.13).

Ele alınan $S(t)$ ve $h(t)$ fonksiyonlarından, hayatta kalma fonksiyonu bir çalışma grubunun hayatta kalma deneyimini doğrudan tanımlamaktadır. Sırf bu yüzden $S(t)$, yaşam verisi analizleri için normal olarak daha cazip olmaktadır (Kleinbaum ve Klein, 2005, s.13).

Ancak tehlike fonksiyonu da aşağıdaki nedenlerden dolayı önemlidir (Kleinbaum ve Klein, 2005, s.13):

- Yaşam eğrisi zaman üzerine birikimli bir ölçüm olmasına karşın tehlike fonksiyonu anlık potansiyelin bir ölçüsüdür.
- Tehlike fonksiyonu; kişiye ait verinin uygun olduğu üstel, Weibull ya da lognormal eğri gibi belirli bir model kalıbı tanımlamak için kullanılabilir.
- Tehlike fonksiyonu; yaşam verilerinin matematiksel modellemesini yapmayı içeren bir araçtır. Yani, hayatta kalma modeli genellikle tehlike fonksiyonu cinsinden yazılmaktadır.

1.2.3.1. Baz Tehlike Fonksiyonu

Tehlike fonksiyonu $h(t)$, koşullu bir başarısızlık oranıdır. Bu fonksiyon, çok kısa bir zaman aralığındaki başarısızlık olasılığı olarak da tanımlanabilmektedir. Fonksiyonda, her bir bireyin aralığın başlangıcından itibaren gözlemlendiği varsayılmaktadır. Çok kısa bir zaman aralığı $(t, t + \Delta t)$ olarak ifade edildiğinde tehlike fonksiyonu aşağıdaki gibidir (Yay, Çoker ve Uysal, 2007);

$$h(t) = \frac{\lim_{\Delta t \rightarrow 0} P \left[\begin{array}{c} \text{bir bireyin } t \text{ zamanına kadar yaşadığı bilindiğine göre,} \\ \text{bu bireyin } (t, t + \Delta t) \text{ zaman aralığına düşmesi} \end{array} \right]}{\Delta t} \quad (1.4)$$

Tehlike fonksiyonu, birikimli fonksiyon $F(t)$ ve olasılık yoğunluk fonksiyonu $f(t)$ cinsinden, aşağıdaki gibi de tanımlanabilmektedir (Yay, Çoker ve Uysal, 2007);

$$h(t) = \frac{f(t)}{1-F(t)} \quad (1.5)$$

\mathbf{X} , $(n \times p)$ boyutlu bağımsız değişkenler matrisi ve X_1, X_2, \dots, X_p , $(n \times 1)$ boyutlu p tane bağımsız değişken vektörü olmak üzere Cox PH modelinde, bağımsız değişkenlerin t zamanındaki tehlike fonksiyonu (Başar, 2006);

$$h(t) = h_0(t) \exp \sum \beta_j X_j, \quad j = 1, \dots, p \quad (1.6)$$

biçiminde tanımlanmaktadır. $h_0(t)$, t zamanındaki tanımlanmamış temel tehlike fonksiyonudur. β_j tehlikenin logaritması üzerinde X_j değişkenin etkisini yansıtan log tehlike orantısıdır. Değişkenlerin anlamlılığı p serbestlik dereceli ençok olabilirlik oran testi ile test edilmektedir (Başar, 2006).

Temel tehlike oranı $h_0(t)$, zamanın negatif olmayan belirsiz bir fonksiyonudur. Temel tehlike fonksiyonu, tehlikenin zamana bağlı kısmıdır ve tüm değişken değerleri sıfıra eşit olduğunda tehlike oranı $h(t)$ 'ye karşılık gelmektedir. $\beta_1, \beta_2, \dots, \beta_p$, $\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p$ regresyon fonksiyonunun katsayılarıdır. Tek bir değişkenle ilgilenildiği varsayıldığında, tehlike aşağıdaki gibi olmaktadır (Bellera vd., 2010);

$$h(t) = h_0(t) \exp \sum \beta X \quad (1.7)$$

$p (> 1)$ tane bağımsız değişkenin varlığında ise,

$$h(t) = h_0(t) \exp(\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p) \quad (1.8)$$

ile tahmin edilmektedir. Yukarıdaki eşitlikte yer alan X_1, X_2, \dots, X_p bağımsız değişkenleri, T yaşam süresi üzerinde etkili olan kan basıncı, sıcaklık, yaş gibi sürekli değişkenler veya histolojik evreler, hastalık evresi gibi kategorik değişkenler olabilmektedir (Yay, Çoker ve Uysal, 2007).

1.2.3.2. Birikimli Tehlike Fonksiyonu

Birikimli tehlike fonksiyonu, araştırma dönemi içerisinde belirli bir t anı için hesaplanmış olan başarısızlık oranlarının birikimli fonksiyonudur. Birikimli tehlike fonksiyonu $H(t)$ ile gösterilir ve aşağıdaki şekilde tanımlanır (Cox ve Oakes, 1984; Rose vd., 2006);

$$H(t) = \int_0^t h(u) du \quad (1.9)$$

Ayrıca $\lim_{t \rightarrow \infty} H(t) = \infty$ olur yani $H(t)$ fonksiyonu artan bir fonksiyondur (Göz Çekçeki, 2007, s.14).

1.3. Hayatta Kalma Analizinde Kullanılan Fonksiyonlar Arasındaki İlişkiler

Tehlike fonksiyonu koşullu olasılık tanımından aşağıdaki şekilde elde edilebilir (Göz Çekçeki, 2007, s.14);

$$h(t) = \frac{f(t)}{S(t)} \quad (1.10)$$

Olasılık yoğunluk fonksiyonu hayatta kalma fonksiyonu cinsinden ifade edilirse aşağıdaki eşitlik elde edilir;

$$f(t) = \frac{dF(t)}{dt} = \frac{d(1 - S(t))}{dt} = -S'(t) \quad (1.11)$$

$S(t)$ ve $h(t)$ 'nin ilişkisi:

$S(t)$ ya da $h(t)$ 'den hangi fonksiyonun tercih edildiğine bakılmaksızın, **ikisi arasında açıkça tanımlanmış bir ilişki bulunmaktadır.** $S(t)$ 'nin yapısı biliniyorsa, ilgili $h(t)$ türetilbilmekte, tam tersi durum da geçerli olmaktadır. Örneğin, tehlike fonksiyonu sabit (yani, λ belirli bir değer olmak üzere $h(t) = \lambda$) ise ilgili hayatta kalma fonksiyonu aşağıdaki formül tarafından verilmektedir (Kleinbaum ve Klein, 2005, s.14):

$$h(t) = \lambda \text{ ancak ve ancak } S(t) = e^{-\lambda t}$$

$S(t)$ ya da $h(t)$ arasındaki ilişki daha genel olarak, aşağıda gösterilen iki hesaplama formülünün herhangi biriyle eşdeğer şekilde ifade edilebilmektedir (Kleinbaum ve Klein, 2005, s.14):

$$S(t) = \exp\left[-\int_0^t h(u) du\right] \quad (1.12)$$

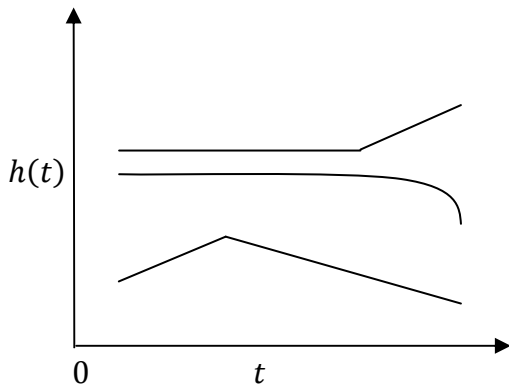
$$h(t) = -\left[\frac{dS(t)/dt}{S(t)}\right] \quad (1.13)$$

Denklem (1.12) hayatta kalma fonksiyonu $S(t)$ 'nin, tehlike fonksiyonunu içeren bir integral cinsinden nasıl yazılabileceğini tanımlamaktadır. Denklem (1.13) tehlike fonksiyonu $h(t)$ 'nin hayatta kalma fonksiyonunu içeren bir türev cinsinden nasıl yazılabileceğini tanımlamaktadır. Sonuç olarak, $S(t)$ ya da $h(t)$ 'den herhangi biri biliniyorsa, diğeri doğrudan elde edilebilmektedir (Kleinbaum ve Klein, 2005, s.14).

1.3.1. Tehlike Oranı ve Eğrisi

Bir hayatta kalma fonksiyonunda olduğu gibi tehlike fonksiyonu $h(t)$, t 'nin çeşitli değerler üzerine dağıldığı bir grafikte gösterilebilmektedir. Şekil 1.11'de üç farklı tehlike fonksiyonu gösterilmektedir. Hayatta kalma fonksiyonunun aksine, $h(t)$ 'nin grafiği 1'den başlamak ve sıfıra düşmek zorunda değildir. Ayrıca herhangi bir yerden başlayarak zaman üzerinde herhangi bir yönde artıp azalabilmektedir (Kleinbaum ve Klein, 2005, s.12).

Şekil 1.11: Üç farklı tehlike fonksiyonu için grafik



t 'nin belirli bir değeri için tehlike fonksiyonu $h(t)$ aşağıdaki özelliklere sahiptir (Kleinbaum ve Klein, 2005, s.12):

- Negatif olmayandır; yani her zaman sıfıra eşit ya da sıfırdan büyüktür. $h(t) \geq 0$
- Hiçbir üst sınıra sahip değildir.

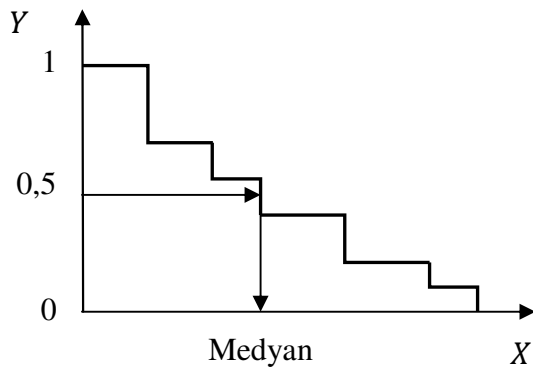
Bu iki özellik $h(t)$ formülünde oran ifadesinin sonucunu belirlemektedir. Çünkü hem paydaki olasılık hem de paydadaki Δt negatif olmayandır ve Δt , 0 ile ∞ arasında dağılabilmektedir (Kleinbaum ve Klein, 2005, s.12).

Hayatta kalma olayının tanımlayıcı ölçüleri, yaşam süresinin basit ortalaması ve ortalama tehlike oranıdır. Sansürlü olma durumu göz ardı edilerek hesaplanan ortalama yaşam süresi (\bar{T}), gerçek yaşam süresi ortalamasının altında tahmin edilmektedir. Çünkü formül sansürlü zamanları da içermektedir. Yaşam süresi ortalaması büyük olan grubun, hayatta kalma konusunda daha etkili olduğu yorumu yapılmaktadır. Her grubun **ortalama tehlike oranı** (\bar{h}) ise toplam başarısızlık sayısının gözlenen yaşam süresi toplamına bölünmesi olarak tanımlanmaktadır. Tehlike oranı hayatta kalma olasılığından ziyade başarısızlık potansiyelini gösterdiğinden ortalama tehlike oranı yükseldikçe grupların hayatta kalma olasılığı düşmektedir (Kleinbaum ve Klein, 2005, s.24-25);

$$\bar{T} = \frac{\sum_{i=1}^n t_i}{n}, \quad \text{Ortalama tehlike oranı } (\bar{h}) = \frac{\# \text{ başarısızlık sayısı}}{\sum_{i=1}^n t_i} = \frac{\sum_{j=0}^k m_j}{\sum_{i=1}^n t_i} \quad (1.14)$$

Ayrıca tahminlenmiş yaşam süresi eğrilerinden, hayatta kalma olasılığının 0,5 olduğu anda medyan yaşam süresinin tahmini elde edilebilmektedir. Grafikselsel olarak medyan, Y eksenindeki 0,5 noktasından yaşam eğrisine ulaşana kadar yatay olarak daha sonra aşağı doğru dikey olarak X eksenini medyan yaşam süresinde kesene kadar ilerleyerek elde edilmektedir (Kleinbaum ve Klein, 2005, s.26).

Şekil 1.12: Tahminlenmiş hayatta kalma eğrilerinden medyan yaşam süresinin elde edilmesi



Tanımlayıcı ölçüler, kontrol grubuyla deney grubunun bütün olarak karşılaştırılmasını sağlamaktadır. Bu ölçüler takip edilen zamandaki farklı noktalarda iki grubu karşılaştırmamaktadır. Böyle bir karşılaştırma hayatta kalma eğrilerinin grafikleri tarafından verilmektedir (Kleinbaum ve Klein, 2005, s.25).

1.3.2. Tehlike Orantısı ve Eğrisi

Genel olarak, bir tehlike orantısı (HR), bir bireyin tehlikesinin farklı bir bireyin tehlikesine bölünmesi olarak tanımlanmaktadır. Karşılaştırılan iki birey, kestirici setinin (yani X 'lerin) değerleri tarafından ayırt edilebilmektedir. Tehlike orantısı, X^* 'in bir birey için kestiricilerin setini gösterdiği, X 'in diğer bireyin kestiricilerinin setini gösterdiği aşağıdaki formülle elde edilmektedir (Kleinbaum ve Klein, 2005, s.100);

$$\widehat{HR} = \frac{\hat{h}(t, X^*)}{\hat{h}(t, X)} \quad (1.15)$$

$X^* = (X_1^*, X_2^*, \dots, X_p^*)$ ve $X = (X_1, X_2, \dots, X_p)$ iki birey için X 'lerin setini göstermektedir.

Bahis oranında (OR) olduğu gibi, etkisiz değer 1'i aşan bir HR 'yi yorumlamak 1'den daha az olan bir HR 'yi yorumlamaktan daha kolaydır, yani $\hat{h}(t, X^*) \geq \hat{h}(t, X)$ olması

istenmektedir. Bu yüzden X 'ler geleneksel olarak, daha büyük tehlikeye sahip olan grup X^* 'a ve daha küçük tehlikeye sahip grup X 'e uyacak şekilde kodlanmaktadır. Örnek olarak, daha önce tanımlanan veri için plasebo grubu $X_1^* = 1$ ve deney grubu $X_1 = 0$ olarak kodlanmaktadır (Kleinbaum ve Klein, 2005, s.100).

Tehlike orantısı ifadesinin pay ve paydasında Cox model formülünü yerine koyarak HR formülü için regresyon katsayıları cinsinden bir ifade elde edilebilmektedir. Aşağıdaki yerine koyma işleminde, pay ve paydadaki tek fark X^* 'lara karşı X 'lerin olmasıdır. Ayrıca temel tehlike fonksiyonları sadeleşecektir (Kleinbaum ve Klein, 2005, s.101).

PH varsayımı, HR 'nin zaman göre sabit olmasını ya da eşdeğer anlamda orantılılık sabitinin zamandan bağımsız olduğu durum olan, bir bireyin tehlikesinin herhangi diğer bireyin tehlikesine orantılı olmasını ($\hat{h}(t, X^*) = \text{sabit} \times \hat{h}(t, X)$) gerektirmektedir (Kleinbaum ve Klein, 2005, s.107).

HR formülünün pay ve paydasındaki temel tehlike fonksiyonu $\hat{h}_0(t)$ sadeleştiğinden, eşitliğin son ifadesinin t 'yi içermediği yani tehlike orantısının zamandan bağımsız olduğu görülmektedir (Kleinbaum ve Klein, 2005, s.107-108);

$$\widehat{HR} = \frac{\hat{h}(t, X^*)}{\hat{h}(t, X)} = \frac{\hat{h}_0(t) \exp[\sum \hat{\beta}_i X_i^*]}{\hat{h}_0(t) \exp[\sum \hat{\beta}_i X_i]} = \exp \left[\sum_{i=1}^p \hat{\beta}_i (X_i^* - X_i) \right] \quad (1.16)$$

$X^* = (X_1^*, X_2^*, \dots, X_p^*)$ ve $X = (X_1, X_2, \dots, X_p)$ iki birey için X 'lerin setini göstermektedir.

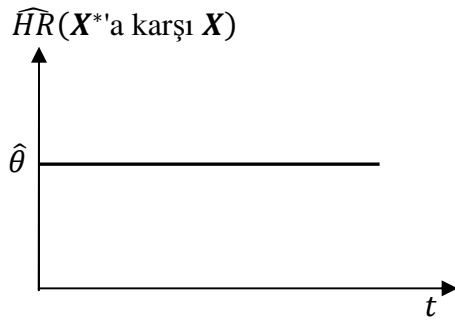
Bu yüzden, modelin kurulması ve X^* ile X değerlerinin tanımlanmasından sonra tahmin edilen tehlike orantısı için üstel ifadenin değeri, zamana bağlı olmayan bir sabittir. Bu sabit $\hat{\theta}$ tarafından gösterilirse, tehlike orantısı aşağıdaki gibi yazılabilmektedir. Bu orantılı tehlike varsayımını gösteren matematiksel bir ifadedir (Kleinbaum ve Klein, 2005, s.108);

$$\hat{\theta} = \exp[\sum_{i=1}^p \hat{\beta}_i (X_i^* - X_i)] \quad (1.17)$$

$$\frac{\hat{h}(t, X^*)}{\hat{h}(t, X)} = \hat{\theta} \quad (1.18)$$

Grafiksel olarak bu ifade, herhangi iki kişiyi karşılaştıran tehlike orantısı tahmininin zaman boyunca bir sabit olarak dağıldığını belirtmektedir (Kleinbaum ve Klein, 2005, s.108);

Şekil 1.13: Zamana göre sabit tehlike oranısı grafiği

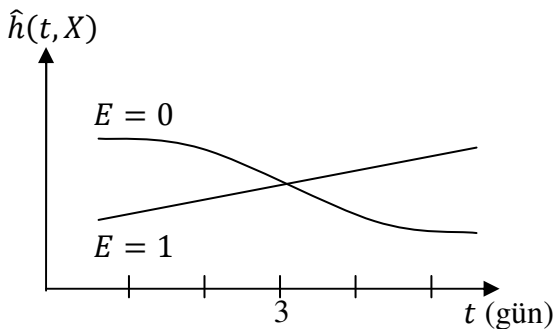


Matematiksel olarak X^* bireyi için tehlike fonksiyonunun, X bireyi için tehlike fonksiyonunun $\hat{\theta}$ katı olduğunu ifade eden orantılı tehlike varsayımını, yazmanın başka bir yolu aşağıdaki gibidir. Orantılılık katsayısının zamana bağlı olmayan $\hat{\theta}$ olduğu bu ifade, bir bireyin tehlike fonksiyonunun başka bir bireyin tehlike fonksiyonu ile orantılı olduğunu belirtmektedir (Elandt-Johnson ve Johnson, 1999, s.280);

$$\hat{h}(t, X^*) = \hat{\theta} \hat{h}(t, X) \quad (1.19)$$

Orantılı tehlike kavramını daha ayrıntılı açıklamak için, orantılı tehlike varsayımını *sağlamayan* duruma varsayımsal bir örnek verilebilir. Şekil 1.14'te yer alan grafikte, ilgilenilen sadece bir açığa çıkma değişkeni E olduğu ve E 'nin ilgilenilen olayın gerçekleşme durumunda 1, ilgilenilen olayın gerçekleşmemesi durumunda 0 değerini alan bir $(0,1)$ değişkeni olduğu varsayılmaktadır (Kleinbaum ve Klein, 2005, s.109);

Şekil 1.14: Zamana göre sabit olmayan tehlike oranı grafiği

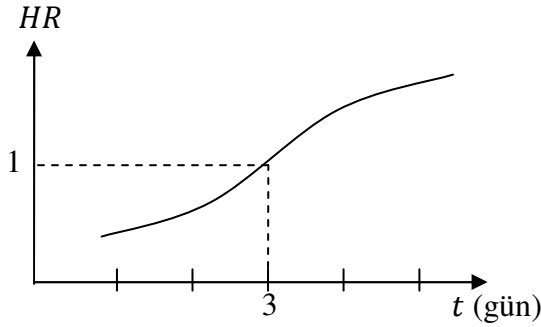


Yaklaşık üçüncü günde bu iki fonksiyonun kesiştiği ve üç günden önce $E = 0$ olan grubun tehlikesinin $E = 1$ olan grubun tehlikesinden daha yüksek olduğu, üçüncü günden sonra tam tersi durumun söz konusu olduğu görülmektedir (Kleinbaum ve Klein, 2005, s.109).

Böylece, her grup için tehlike fonksiyonları yukarıdaki gibiyse, **tehlike orantıları zaman içinde sabit değildir**. Yani, tehlike orantısı üç günden önce 1'den daha küçük, üç günden sonra 1'den daha büyük bir sayıdır (Kleinbaum ve Klein, 2005, s.110).

Tehlike orantısının zaman içindeki değişimi Şekil 1.15'teki gibidir;

Şekil 1.15: Tehlike orantısının zaman içindeki değişimi



Bu yüzden, bu durum için bir Cox PH modelini kullanmak uygunsuzdur çünkü PH modeli, zamanın karşısında sabit bir tehlike orantısını varsaymaktadır. Oysa yukarıdaki durum, zamanla değişen bir tehlike orantısı vermektedir (Kleinbaum ve Klein, 2005, s.110).

Aslında, Cox PH modeli yine de kullanılırsa, herhangi bir zamanda olayın meydana geldiği ve gelmediği kişileri karşılaştıran tehlike orantısı tahmini, zamana göre değişmeyen sabit değer $e^{\hat{\beta}}$ tarafından verilecektir (Kleinbaum ve Klein, 2005, s.110);

$$h(t, \mathbf{X}) = h_0(t)e^{\beta E} \quad (1.20)$$

Tehlike oranı Denklem (1.20)'deki gibi iken tehlike orantısı aşağıdaki gibi olacaktır;

$$\widehat{HR} = \frac{\hat{h}(t, \mathbf{E} = 1)}{\hat{h}(t, \mathbf{E} = 0)} = e^{\hat{\beta}} \quad (1.21)$$

Bu örnek, tehlike fonksiyonlarının kesiştiği durumda, PH varsayımının sağlanamadığı genel kuralını göstermektedir. Bu yüzden bir Cox PH modeli uygun olmamaktadır (Kleinbaum ve Klein, 2005, s.110).

Aslında, bu örneği analiz etmek için birkaç olası seçenek vardır. Bunlar (Kleinbaum ve Klein, 2005, s.110);

- Açığa çıkma değişkenini tabakalandırarak analiz; yani açığa çıkma değişkeni hiçbir modele uymaz, bunun yerine her bir açığa çıkma grubu için ayrı ayrı Kaplan-Meier eğrileri elde edilir.

- Analize üçüncü günde başlanır ve üç gün hayatta kalanlar üzerine bir Cox model uygulanır.
- Üç günden daha az ve daha fazla olan süre için iki farklı tehlike orantısı tahmin etmek amacıyla, bu iki zaman periyodunun her biri için iki farklı Cox model kurulur.
- Zaman ile açığa çıkan olayın etkileşimini ölçen zamana bağlı bir değişken içeren değiştirilmiş bir Cox model kurulur. Bu model **genişletilmiş Cox model** olarak adlandırılmaktadır.

Farklı seçenekler farklı sonuçlara yol açabilmektedir. Bu yüzden en iyi olarak belirli bir seçeneğe karar vermeden önce, elde edilen verinin ışığında her seçeneğin görece avantajlarını tartmak zorunda kalınmaktadır (Kleinbaum ve Klein, 2005, s.111).

İKİNCİ BÖLÜM

COX ORANTILI TEHLİKE REGRESYON MODELİ

Cox regresyon yöntemi, bir hastanın yaşam süresi (bağımlı değişken) ile birden fazla bağımsız değişken arasındaki ilişkiyi ortaya çıkaran istatistiksel bir yöntemdir. Cox regresyon analizinde hedef, hayatta kalma verilerinin genel durumunu yansıtacak bir model oluşturmaktır. Bu yolla, hayatta kalma süresi üzerinde etkili olduğu düşünülen bağımsız değişkenlerin etkileri eşanlı olarak açıklanabilmekte, yani ölçülebilmektedir. Örneğin; Cox regresyon modeli, klinik bir çalışmada hastaların yaşamını analiz etmek için kullanıldığında, tedavinin etkisini diğer değişkenlerin etkisinden ayırmaya yardımcı olmaktadır (Yay, Çoker ve Uysal, 2007).

Cox model, araştırılan faktörlerin tehlike (ya da risk) üzerinde zaman içinde sabit bir etkiye sahip olduğu anlamına gelen orantılı tehlike varsayımına dayanmaktadır (Bellera vd., 2010).

2.1. Cox Regresyon Modeli ve Varsayımları

Parametrik modellerin gerektirdiği varsayımların (normallik, bağımsızlık vb.) sağlanmadığı durumlarda Cox regresyon analizi, parametrik analizlerden daha etkili olmaktadır. Cox regresyon modelinin temel varsayımları şu şekilde açıklanabilir (Yay, Çoker ve Uysal, 2007):

1. Bağımsız değişkenlerin tehlike fonksiyonu üzerindeki etkileri loglineerdir.
2. Bağımsız değişkenlerin loglineer fonksiyonu ile tehlike fonksiyonu arasındaki ilişki çarpımsaldır.
3. Bu iki varsayıma ek olarak gözlemlerin birbirinden bağımsız olmaları ve tehlike orantısının zamana göre değişmemesi, yani sabit olması gerekmektedir. Tehlike orantısı ile ilgili bu varsayım orantılı tehlike varsayımı olarak bilinmektedir.

Cox PH modeli, araştırılan faktörlerin tehlike (ya da risk) üzerinde zaman içinde sabit bir etkiye sahip olduğu anlamına gelen tehlikelerin orantılılığı gibi temel bir varsayıma dayanmaktadır (Ata ve Sözer, 2007). Eğer uygun modelleme olmadan zamana bağlı değişkenler modele dahil edilirse, orantılı tehlike varsayımı ihlal edilmektedir. Sonuç olarak, yanıltıcı etki tahminleri türetilen periyodun başlangıcındaki (ya da

sonundaki) anlamlı etki kaçırılmamaktadır. Parametre tahminleri üzerindeki olası sonuçları göz önünde tutulursa, tehlikelerin orantılılığını kontrol etme Cox modele dayanan bir hayatta kalma analizinin ayrılmaz bir parçası olmalıdır. Zamanla değişken risklere sahip değişkenlerin varlığında, sonuçları kanıtlamak ve orantısızlığın nerede bulunduğunu göstermek için grafikler kullanılmalıdır (Bellera vd., 2010).

Orantılı tehlike varsayımı sağlanmadığında, zamana bağlı bağımsız değişkenli Cox regresyon modeline başvurulmaktadır (Yay, Çoker ve Uysal, 2007). İncelenen örnekte bağımsız değişkenler yoksa veya sadece bir kategorik bağımsız değişken varsa tehlike fonksiyonları için yaşam tablolarına ve Kaplan-Meier yöntemine başvurulabilir (Parker vd, 2002).

2.2. Cox Regresyon Modelinin Genel Özellikleri ve Tahminlerin Direnci

Cox modelin popülerliğinin en önemli sebebi; tehlike fonksiyonu tanımsız olmasına rağmen, çok çeşitli veri durumu için regresyon katsayılarının, ilgilenilen tehlike oranı değerlerinin ve ayarlanmış hayatta kalma eğrilerinin oldukça iyi tahminlerinin elde edilebilmesidir. Başka bir deyişle Cox PH modeli, “dirençli” (robust) bir modeldir. Yani Cox model kullanarak elde edilen sonuçlar, doğru parametrik modelin sonuçlarına çok yaklaşıktır (Kleinbaum ve Klein, 2005, s.96).

Örneğin, doğru parametrik model Weibull ise, Cox model kullanmak, Weibull modeli kullanılarak elde edilenlere sıklıkla yakın sonuçlar verecektir. Ya da doğru model üstel ise, Cox model sonuçları, kurulan üstel model sonuçlarına çok yaklaşıktır (Kleinbaum ve Klein, 2005, s.96).

Eğer doğru modelden emin olunursa, bir parametrik model kullanmak tercih edilebilmektedir. Bir parametrik yöntemin uyum iyiliğini değerlendirmek için birçok yöntem olmasına rağmen, belirli bir parametrik modelin uygun olduğuna tam anlamıyla emin olunamamaktadır (Kleinbaum ve Klein, 2005, s.96).

Bu yüzden şüpheye düşüldüğünde, “güvenli” bir model seçimi olduğundan ve yanlış parametrik modelin seçilip seçilmediği hakkında endişelenmeye gerek olmadığından, Cox model yeterince güvenilir sonuçlar verecektir (Kleinbaum ve Klein, 2005, s.97).

Cox modelin genel “dirençliliği”ne (robustness) ek olarak modelin özel yapısı, birçok sebepten dolayı caziptir (Kleinbaum ve Klein, 2005, s.97);

- Tehlike modelinin üstel kısmı $e^{\sum_{i=1}^p \beta_i X_i}$, kurulan modelin her zaman negatif olmayan tehlike tahminleri vermesini sağlamaktadır. Tahminlerin negatif olmaması istenir çünkü herhangi bir tehlike fonksiyonunun değeri sıfır ile artı sonsuz arasında dağılmalıdır. Yani bir tehlike her zaman negatif olmayandır ($0 \leq h(t, \mathbf{X}) < \infty$). Üstel ifadenin yerinde, X 'lere göre doğrusal olan modeldeki X kısmı olsaydı, göz önüne alınmayan negatif tehlike tahminleri elde edilebilirdi (Kleinbaum ve Klein, 2005, s.97).

$$h(t, \mathbf{X}) = h_0(t) \underbrace{\sum_{i=1}^p \beta_i X_i}_{\text{Doğrusal} \Rightarrow < 0 \text{ olabilir}} \quad (2.1)$$

- Cox modelin diğer bir başvuru özelliği, modelin temel tehlike kısmı tanımsız olmasına rağmen, modelin üstel kısmındaki β 'ları tahmin etmenin yine de mümkün olmasıdır. İlgilenilen açıklayıcı değişkenin etkisini değerlendirmek için tek yapılması gereken β 'ların tahmin edilmesidir. Tehlike orantısı olarak adlandırılan etkinin ölçüsü, temel tehlike fonksiyonunu tahmin etmek zorunda olmadan hesaplanmaktadır (Kleinbaum ve Klein, 2005, s.97).

- Temel tehlike fonksiyonu tanımlanmamasına rağmen, Cox model için tehlike fonksiyonu $h(t, \mathbf{X})$ ve onun ilgili hayatta kalma eğrisi $S(t, \mathbf{X})$ tahmin edilebilmektedir. Bu yüzden Cox model ile en az varsayımı kullanarak, hayatta kalma analizinden istenen esas bilgi yani tehlike orantısı ve hayatta kalma eğrisi elde edilebilmektedir (Kleinbaum ve Klein, 2005, s.97).

- Cox modelin popülerliği hakkındaki son husus, yaşam süresi bilgisi mevcut olduğunda ve sansürleme olduğunda lojistik modelin geliştirilmesine yardım etmesidir. Yani Cox model, bir (0,1) çıktısını inceleyen ve yaşam süresi ile sansürlenmeyi göz ardı eden lojistik modelden daha çok bilgi (yaşam süreleri) kullanmaktadır (Kleinbaum ve Klein, 2005, s.98).

Aşağıda bir Cox modelde katsayılarının yorumlanması, bir değişkenin karıştırıcı ve/veya etkileşim etkisinin değerlendirilmesi ile ilgili varsayımsal bir örnek verilmektedir.

İlgilenilen temel sorun; deney grubu için $X_1 = 0$, plasebo grubu için $X_1 = 1$ değerlerini alan iki farklı grup üzerinde etkisi olduğu düşünülen herhangi bir dışsal değişkenin (burada X_2), olası karıştırıcı ve/veya etkileşim etkisi için ayarlanan bu grupların hayatta kalma deneyimlerini karşılaştırmayla alakalıdır (Kleinbaum ve Klein, 2005, s.86).

Burada X_1 , ilgilenilen asıl çalışma değişkeni ya da açığa çıkma değişkenidir. Deney grubu için 0, plasebo grubu için 1 değerini alan X_1 ; gözlemlerin ait olduğu grubun deney

durumunu göstermektedir. X_2 ise olası karıştırıcı ya da etki değiştirici olarak düşünülen dışsal değişkeni göstermektedir. Etkisi araştırılmak istenen dışsal değişkenin (X_2), gruplar (deney ve plasebo) üzerindeki olası etkileşim etkisi değerlendirilmek istenirse; X_1 ve X_2 'nin çarpımı olan üçüncü bir değişkeni hesaba katma ihtiyacı duyulacaktır. (Kleinbaum ve Klein, 2005, s.87).

2.3. Cox Regresyon Modeli ve Test İstatistikleri

Bir katsayı kendi standart hatasına bölüldüğünde, bulunan sayının yaklaşık olarak standart normal ya da Z değişkeni olduğu varsayılmaktadır. Bu Z istatistiği, en çok olabilirlik (ML) tahminlerini kullanan geleneksel iki testten biri olan **Wald istatistiği** olarak bilinmektedir. **Olabilirlik oranı** ya da LR istatistiği olarak adlandırılan diğer test istatistiği, log olabilirlik istatistiğini kullanmaktadır. Bilgisayar paket programı çıktılarından elde edilebilen log olabilirlik istatistiği (L), aşağıdaki formülde yerine koyularak olabilirlik oranı hesaplanmaktadır (Kleinbaum ve Klein, 2005, s.89);

$$LR \text{ istatistiği} = -2 \ln L \quad (2.2)$$

Olabilirlik oranı (LR) istatistiğini etkileşim teriminin anlamlılığını test etmede kullanmak için, etkileşim terimini içermeyen indirgenmiş modelin olabilirlik istatistiği ve etkileşim terimini içeren genişletilmiş modelin olabilirlik istatistiği arasındaki farkın hesaplanması gerekmektedir. Genel olarak LR istatistiği aşağıdaki formülle ifade edilmektedir (Kleinbaum ve Klein, 2005, s.89);

$$LR \text{ istatistiği} = -2 \ln L_R - 2(-\ln L_F) \quad (2.3)$$

R = indirgenmiş model, F = genişletilmiş (tam) model

Etkileşim etkisi olmadığını öne süren sıfır hipotezi altında, test istatistiği p serbestlik dereceli ki-kare dağılımına sahiptir. Burada p değerlendirilen kestirici değişkenlerin sayısını göstermektedir. Bu test istatistiği için de elde edilen p değerinin 0,05'ten büyük olması, anlamlı bir etkileşim olmadığını göstermektedir (Kleinbaum ve Klein, 2005, s.90).

Genellikle LR ve Wald istatistikleri tam olarak aynı sonucu vermemektedir, ancak her iki değer de aynı sonucu doğurmaktadır. Şüpheye düşülen durumlarda ise LR istatistiği kullanılmalıdır. Çünkü istatistik uzmanları LR istatistiğinin daha iyi istatistiksel özelliklere sahip olduğu kanıtlamışlardır (Kleinbaum ve Klein, 2005, s.90).

Ele alınan grupların etkisini değerlendirmede üç istatistiksel amaç bulunmaktadır; İlgilenilen dışsal değişken X_2 için ayarlanan deney durumu değişkeni, X_1 'in **anlamlılığını test etmek**, deney durumunun (X_1) hayatta kalma üzerindeki **etkisinin nokta tahminini** ve **bu etkinin güven aralığını** elde etmektir (Kleinbaum ve Klein, 2005, s.90).

Deney durumu değişkeni X_1 'in anlamlılığına, kendisine ait katsayının p değerine bakarak karar verilmektedir. Deneyin etkisinin nokta tahmini; deney durumu değişkeni X_1 'in katsayısı $\hat{\beta}_1$ 'in, $\widehat{HR} = e^{\hat{\beta}}$ formülünde yerine koyulmasıyla elde edilmektedir. Bu değer, deneyin etkisi için tehlike orantısı tahminini (\widehat{HR}) vermektedir ve plasebo grubunun olay riskinin deney grubunun sahip olduğu olay riskinin kaç katı olduğunu ifade etmektedir (Kleinbaum ve Klein, 2005, s.90-91).

Deney durumunun etkisi için güven aralığı, tehlike orantısı HR 'nin nokta tahmininin etrafını çevrelemektedir. Güven aralığının çok geniş olması, nokta tahmininin biraz güvenilmez olduğunu göstermektedir. p değeri küçük olduğunda HR 'nin güven aralığı, etkisiz değer 1'i içermemektedir. HR 'nin güven aralığının hesaplanması aşağıdaki şekilde yürütülmektedir (Kleinbaum ve Klein, 2005, s.91);

1. İlgilenilen değişkenin regresyon katsayısı β 'nin %95 güven aralığı hesaplanır.

$$(\hat{\beta} \pm z_{\alpha/2} \times s_{\hat{\beta}})$$
2. Elde edilen güven aralığının e tabanında üstü alınır. $(e^{\hat{\beta} \pm z_{\alpha/2} \times s_{\hat{\beta}}})$

İlgilenilen dışsal değişken bir karıştırıcı değişken olarak kabul edilirse, amaç yine bu dışsal değişken için ayarlanan deney durumu değişkeni X_1 'in etkisini tanımlamak olacaktır (Kleinbaum ve Klein, 2005, s.89).

Eğer dışsal değişken X_2 'nin modelde bulunduğu ve bulunmadığı durumlar arasındaki deney durumu değişkeninin ($X_1 = 0$ ve $X_1 = 1$ için) HR değerleri birbirinden anlamlı derece farklı ise, dışsal değişken X_2 'nin karıştırıcı etkisi olduğu söylenebilecektir (Kleinbaum ve Klein, 2005, s.92).

Karıştırıcı etkinin var olduğuna karar verdikten sonra, etkinin geçerli bir tahminini elde etmek amacıyla model karıştırıcı için denetlenmelidir. Karıştırıcı etkiyi kontrol ettiği için, dışsal değişken X_2 'yi içeren model tercih edilecektir (Kleinbaum ve Klein, 2005, s.92).

Eğer “anlamlı” bir karıştırıcı etki olmadığına karar verilirse, geçerli bir sonuç elde etmek için dışsal değişken X_2 'nin kontrol edilmesine gerek yoktur. Yine de tehlike orantısının daha kesin bir sonucunu elde etmek için, değeri ne olursa olsun dışsal değişken X_2 kontrol edilmek istenebilir. Dışsal değişken X_2 kontrol için modele eklendiğinde HR 'nin güven

aralığı daralır, **kesinlik** kazanmak için güven aralığı daha dar olan bu model tercih edilecektir (Kleinbaum ve Klein, 2005, s.92).

Cox PH modelinin analiz stratejisi ve yöntemleri, lojistik ve klasik doğrusal modeller için olan yöntemlerle benzerlik göstermektedir (Kleinbaum ve Klein, 2005, s.93).

2.4. Cox Orantılı Tehlike Regresyon Modelinin Yapısı ve PH Varsayımı

Değişken değerleri X_1 ve X_2 olan iki deneğin tehlikesi sırasıyla $h_{X_1}(t) = h_0(t) \exp(\beta X_1)$ ve $h_{X_2}(t) = h_0(t) \exp(\beta X_2)$ tarafından verilmektedir ve tehlike oranısı aşağıdaki gibi ifade edilmektedir (Bellera vd., 2010);

$$HR = \frac{h_{X_2}(t)}{h_{X_1}(t)} = \exp[\beta(X_2 - X_1)] \quad (2.4)$$

$X_2 = X_1 + 1$ alınır, tehlike oranısı $HR = \exp(\beta)$ haline gelecektir. Tehlike oranısı açıklayıcı değişken X 'teki bir birimlik artışın olayın riski üzerindeki etkisine karşılık gelmektedir. $\beta = \log(HR)$ olduğundan β , log tehlike oranısı olarak adlandırılmaktadır. Tehlike oranı $h(t)$ 'nin zamanla değişmesi kabul edilmesine rağmen, tehlike oranısı HR sabittir; bu da *orantılı tehlike varsayımı*dır. Değişken değeri X_2 olan deneklerin değişken değeri X_1 olan deneklere nazaran olay riski HR , 1'den büyük ($\beta > 0$) ise artmaktadır. Oysa 1'den küçük ($\beta < 0$) bir HR azalan bir riski işaret etmektedir. HR zamana göre sabit olmadığında, değişkenin zamana bağlı bir etkiye sahip olduğu söylenmektedir. Örneğin; bir deneyin etkisi, deneyden hemen sonra güçlü olabilmekte fakat zamanla zayıflayabilmektedir. Bu durum, sigara kullanma durumu gibi değeri zamana göre sabit olmayan bir değişken olan, zamana bağlı değişken ile karıştırılmamalıdır. Gerçekten de bir kişi; önce sigara kullanmayan biri, sonra sigara kullanan biri, sonra da yine sigara kullanmayan biri olabilmektedir. Ancak, bir değişken hem zamana bağlı olabilmekte hem de zamana göre değişen bir etkiye sahip olabilmektedir (Bellera vd., 2010).

Bir Cox PH modelinde tehlike oranısı, bir olayın meydana geldiği tüm t zamanlarını ele alarak tahmin edilmektedir. Takip edilen periyodun tamamı üzerindeki toplam tehlike oranısı tahmin edilirken, neredeyse tüm bireyleri etkileyen çok baştaki tehlike oranlarına ve hala risk altında olan sadece çok az bireyi etkileyen çok sondaki tehlike oranlarına aynı ağırlıklar verilmektedir. Bu yüzden tehlike oranısı, olay zamanlarının *ortalaması*dır. Tehlikelerin orantılı olması durumunda, toplam tehlike oranısı bu ağırlıklandırma yönteminden etkilenmemektedir. Öte yandan, tehlike oranısı zamana göre değişiyorsa yani

tehlike değerleri orantılı değil ise, bu durumda eşit ağırlıklandırma temsil edici olmayan bir tehlike orantısı ile sonuçlanabilmekte ve yanlış sonuçlar üretebilmektedir. Tehlike orantısı, takip edilen zamandan ziyade olay zamanlarının ortalamasıdır. Olayların sıralaması bozulmadan zaman ölçeği değiştirilirse, tehlike orantısı değişmeyecektir (Bellera vd., 2010).

Cox PH modeli genellikle aşağıda gösterildiği gibi tehlike modeli formülü cinsinden yazılmaktadır. Bu model \mathbf{X} ile gösterilen bir açıklayıcı değişkenler setinin belirlenmiş tanımı vasıtasıyla, bir bireyin t zamanındaki tehlikesi için bir formülasyon vermektedir. Yani \mathbf{X} , bir bireyin tehlikesini tahmin etmek için modellenen kestirici değişkenlerin bir yığını (bazen bir “vektör” olarak adlandırılan) temsil etmektedir (Kleinbaum ve Klein, 2005, s.94).

$$h(t, \mathbf{X}) = h_0(t) e^{\sum_{i=1}^p \beta_i X_i} \quad (2.5)$$

$\mathbf{X} = (X_1, X_2, \dots, X_p)$ Açıklayıcı / kestirici değişkenler

Cox model, t zamanındaki tehlikenin iki niceliğın çarpımı olduğunu ifade etmektedir. Bunlardan ilki $h_0(t)$, **temel tehlike** fonksiyonu olarak adlandırılmaktadır. İkinci nicelik, p tane açıklayıcı X değişkeni üzerine $\beta_i X_i$ 'nin e üssünde doğrusal toplamı olan üstel ifadedir (Kleinbaum ve Klein, 2005, s.94).

Tablo 2.1: Tehlike fonksiyonunu oluşturan unsurlar arasındaki farklar

$h_0(t)$	×	$e^{\sum_{i=1}^p \beta_i X_i}$
Temel tehlike fonksiyonu t 'yi içerir, X 'leri içermez.		Üstel kısım X 'leri içerir, t 'yi içermez. X 'ler zamana bağlı değildir.

Bu formülün PH varsayımıyla ilgili önemli bir özelliği; temel tehlike fonksiyonunun t 'nin bir fonksiyonu olup X 'leri içermemesidir. Aksine, üstel ifade X 'leri içerip t 'leri içermemektedir. X 'ler burada **zamandan bağımsız** değişkenlerdir (Kleinbaum ve Klein, 2005, s.94).

Ama yine de, t 'yi içeren X 'leri hesaba katmak mümkündür. Böyle X 'ler **zamana bağlı değişkenler** olarak adlandırılmaktadır. Eğer zamana bağlı değişkenler ele alınırsa, Cox model yapısı yine de kullanılabilir fakat böyle bir model artık PH varsayımını sağlamamaktadır ve **genişletilmiş Cox model** olarak adlandırılmaktadır (Kleinbaum ve Klein, 2005, s.95).

Zamandan bağımsız değişken, belirli bir birey için değeri zamanla değişmeyen bir değişken olarak tanımlanmaktadır. Örnek olarak cinsiyet *GEN* değişkeni ve sigara kullanma durumu *SMK* verilebilir. Aslına bakılırsa bir kişinin sigara kullanma durumu zamanla değişebilir fakat analizin kısıtlarından dolayı, *SMK* değişkeninin ölçüldükten sonra değişmediği varsayılmaktadır yani her birey için sadece bir değer kullanılmaktadır (Kleinbaum ve Klein, 2005, s.95).

Ayrıca, yaş (*AGE*) ve ağırlık (*WGT*) gibi değişkenler zamanla değişmesine rağmen, eğer böyle değişkenlerin değeri zamanla çok fazla değişmiyorsa ya da hayatta kalma riski üzerindeki etkileri temel olarak sadece bir ölçüm değerine dayanıyorsa, bu tür değişkenler analizde zamandan bağımsız gibi davranmaya uygun olabilmektedir (Kleinbaum ve Klein, 2005, s.95).

Cox modelin diğer önemli özelliği, temel tehlike $h_0(t)$ 'nin tanımsız bir fonksiyon olmasıdır. Bu özellik, Cox modeli **yarı parametrik** bir model yapmaktadır (Scheike, 2004, s.83).

Buna karşılık, **parametrik** bir model bilinmeyen parametre değerleri dışında fonksiyonel yapısı tamamen tanımlanan bir modeldir. Örneğin Weibull tehlike modeli parametrik bir modeldir ve bilinmeyen parametreleri λ , p ve β_i olan aşağıdaki yapıya sahiptir. p , parametre sayısını ifade etmektedir (Kleinbaum ve Klein, 2005, s.96);

$$h(t, \mathbf{X}) = \lambda p t^{p-1} = h_0(t) e^{\sum_{i=1}^p \beta_i X_i} \quad (2.6)$$

Burada, $\lambda = e^{\sum_{i=1}^p \beta_i X_i}$ ve $h_0(t) = p t^{p-1}$ 'dir.

Cox modelin bu kadar popüler olmasının sebeplerinden biri yarı parametrik olmasıdır (Kleinbaum ve Klein, 2005, s.96).

2.5. Orantılı Tehlike Modeli için Olabilirlik Fonksiyonunun Elde Edilmesi

Genel olarak bir olabilirlik fonksiyonunun formülü, çıktıların dağılımına dayanmaktadır. Ancak Cox modelin en önemli özelliklerinden biri, çıktı değişkeni için varsayılan bir dağılım olmamasıdır (yani olay zamana bağlıdır). Bu yüzden parametrik modelin aksine, Cox model için çıktı dağılımına dayanan tam bir olabilirlik formüle edilememektedir. Bunun yerine Cox olabilirliğinin yapısı olayların birleşik dağılımından ziyade, **olayların gözlenen sırasına dayanmaktadır**. Bu yüzden Cox olabilirliği, “kısmi” olabilirlik olarak adlandırılmaktadır (Kleinbaum ve Klein, 2005, s.111).

β bilinmeyen katsayılar vektörünün tahmini için kısmi olabilirlik fonksiyonundan faydalanılmaktadır. Cox regresyon analizinde kullanılan kısmi olabilirlik fonksiyonu için öncelikle gözlenen n tane yaşam süresi içinden k tanesinin $t_1 < t_2 < t_3 < \dots < t_k$ şeklinde sıralanmış başarısızlık süreleri olduğu düşünülmektedir (Yay, Çoker ve Uysal, 2007).

k başarısızlık zamanına sahip bir veri seti için; L_j 'nin, j . başarısızlık zamanına ait olabilirliğin katkısını gösterdiği Cox olabilirliğinin genel formülü aşağıdaki gibidir. Cox olabilirliği her bir k teriminin bir çarpımı olarak formüle edilmektedir (Kleinbaum ve Klein, 2005, s.115).

$$L = L_1 \times L_2 \times L_3 \times \dots \times L_k = \prod_{j=1}^k L_j \quad (2.7)$$

Cox modelde gözlenen sıralı olayların olabilirliği, her bireyin değişkenlerinin davranışı (açığa çıkma değişkeninin incelenen bireyde meydana gelmesi veya gelmemesi) tarafından etkilenmektedir (Kleinbaum ve Klein, 2005, s.112).

Cox olabilirliğinin en önemli özelliği, her olabilirlik teriminde temel tehlike fonksiyonunun sadeleşmesidir. Bu yüzden, bir Cox modelin temel tehlikenin yapısını tanımlamaya ihtiyaç duyulmamaktadır. Çünkü temel tehlike, regresyon parametrelerinin tahmininde hiçbir rol oynamamaktadır yani L , $h_0(t)$ 'ye bağlı değildir (Kleinbaum ve Klein, 2005, s.114);

$$L_j = \frac{j \text{ zamanında bireyin tehlikesi}}{j \text{ zamanındaki toplam tehlike}} \quad (2.8)$$

$$L_j = \frac{h_0(t)e^{\hat{\beta}_j}}{h_0(t)e^{\hat{\beta}_j} + h_0(t)e^{\hat{\beta}_{j+1}} + \dots + h_0(t)e^{\hat{\beta}_k}} = \frac{h_0(t)e^{\hat{\beta}_j}}{h_0(t) [e^{\hat{\beta}_j} + e^{\hat{\beta}_{j+1}} + \dots + e^{\hat{\beta}_k}]} = \frac{e^{\hat{\beta}_j}}{\sum_j^k e^{\hat{\beta}_j}} \quad (2.9)$$

$$L = \prod_{j=1}^k \frac{e^{\hat{\beta}_j}}{\sum_j^k e^{\hat{\beta}_j}} \quad (2.10)$$

Cox olabilirliği olayların ve sansürlerin sırası tarafından belirlenmektedir, yani çıktı değişkeninin dağılımı tarafından belirlenmemektedir. Bu yüzden, veri setinde zaman değişkeninin değerleri değişse bile Cox olabilirliği aynı kalacaktır. Çünkü çıktının (zaman) sırası değişmeden kalmaktadır (Kleinbaum ve Klein, 2005, s.114).

Olabilirlik formüle edildikten sonra, *hangi regresyon parametresi değerinin L 'yi maksimize ettiği* hesaplanmaktadır. Olabilirliği maksimize etme süreci, genel olarak aşağıdaki formülle yürütülmektedir ve (skor denklemleri olarak adlandırılan) denklem sistemleri çözülmektedir (Kleinbaum ve Klein, 2005, s.115).

$$\frac{\partial \ln L}{\partial \beta_i} = 0, \quad i = 1, 2, \dots, p \quad (2.11)$$

p = parametre sayısı

Kısmi olabilirlik fonksiyonunun logaritmik formu aşağıdaki gibidir (Yay, Çoker ve Uysal, 2007).

$$\log L = \sum_{j=1}^k \{ \hat{\beta}_j - \log[\sum_j^k \exp(\hat{\beta}_j)] \} \quad (2.12)$$

Herhangi bir bağımsız değişken için elde edilen pozitif bir katsayı, riskin yüksek olduğunu işaret etmektedir. Negatif bir katsayı ise, gözlemlerin hayatta kalma üzerinde daha başarılı olduğu şeklinde yorumlanmaktadır.

2.5.1. Newton-Raphson Yöntemi

“Skor testi” olarak da bilinen bu yöntem, β katsayılarının en çok olabilirlik tahminleri için kısmi olabilirlik fonksiyonunun maksimize edilmesinde kullanılmaktadır.

Cox model parametrelerinin ML tahminleri lojistik regresyondaki gibi, genellikle L ile gösterilen bir olabilirlik fonksiyonunu maksimize ederek türetilmektedir. Olabilirlik fonksiyonu; ele alınan modeldeki bilinmeyen parametrelerin (β 'lar) bir fonksiyonu olarak, çalışmadaki denekler üzerinde gerçekten gözlenen veriyi elde etmenin birleşik olasılığını tanımlayan matematiksel bir ifadedir. L bazen, β 'nın bilinmeyen parametrelerin yığını gösterdiği $L(\beta)$ gibi notasyon şeklinde yazılmaktadır (Kleinbaum ve Klein, 2005, s.98).

$L = L(\beta) =$ gözlenen verinin birleşik olasılığı

Cox model olabilirlik fonksiyonunun formülü, aslında (tam) bir olabilirlik fonksiyonundan ziyade “kısmi” olabilirlik fonksiyonu olarak adlandırılmaktadır. Olabilirlik fonksiyonu sadece başarısız olan deneklerin olasılıklarını hesaba kattığından ve sansürlenmiş deneklerin olasılıklarını tam ve açık biçimde dikkate almadığından “kısmi” olabilirlik terimi kullanılmaktadır (Kleinbaum ve Klein, 2005, s.99).

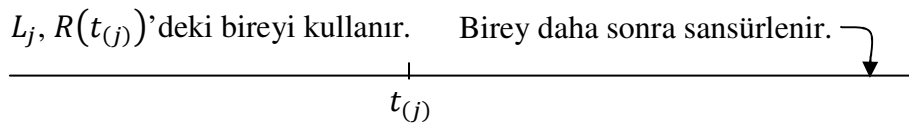
Kısmi olabilirlik özel olarak, her biri k (keyfi bir değer) başarısızlık zamanlarından biri olan birçok olabilirliğin çarpımı şeklinde yazılabilmektedir. Böylece j . başarısızlık zamanında, bu zamana kadar hayatta kalma koşuluyla L_j ; j . başarısızlık zamanında başarısızlığın olabilirliğini göstermektedir. j . başarısızlık zamanında risk altındaki bireylerin seti, “risk seti” $R(t_{(j)})$ olarak adlandırılmakta ve bu setin büyüklüğü, başarısızlık zamanı arttıkça değişmektedir (azalmaktadır) (Kleinbaum ve Klein, 2005, s.99).

$$L = L_1 \times L_2 \times L_3 \times \dots \times L_k = \prod_{j=1}^k L_j \quad (2.13)$$

Burada $L_j =$ belirli bir risk seti $R(t_{(j)})$ 'de j .başarısızlık için L 'nin bir bileşeni, $k =$ başarısızlık zamanı sayısı

Bu yüzden kısmi olabilirlik başarısız olan deneklere odaklanmasına rağmen, sansürlenmiş denekler için sansürlenmeden önceki yaşam süresi bilgisi kullanılmaktadır. Yani, j . başarısızlık zamanından sonra sansürlenmiş bir kişi, L_j 'yi hesaplamak için kullanılan risk setinin bir parçasıdır (Kleinbaum ve Klein, 2005, s.99).

Şekil 2.1: Sansürlü denekler için kullanılan yaşam süresi bilgisi



Belirli bir model için olabilirlik fonksiyonu yapılandırıldıktan sonra, bir sonraki adım bu fonksiyonu maksimize etmektir. Bu genellikle, hesaplanması daha kolay olan L 'nin doğal logaritmasını maksimize ederek yapılmaktadır (Kleinbaum ve Klein, 2005, s.99).

$$\frac{\partial \ln L}{\partial \beta_i} = 0, \quad i = 1, 2, \dots, p \text{ (parametre sayısı)} \quad (2.14)$$

Maksimizasyon süreci, modeldeki her bir parametreye göre L 'nin logaritmasının kısmi türevini alarak yürütülmekte ve daha sonra denklem sistemleri çözülmektedir. Bu çözüm iterasyon kullanarak yürütülmektedir. Yani çözüm, çözümün tahmin edilen bir değeri ile başlayan adımsal bir şekilde elde edilmekte ve tahmin edilen değer son çözüm elde edilene kadar arka arkaya modifiye edilmektedir. İterasyon çözümü şu şekilde özetlenebilir (Kleinbaum ve Klein, 2005, s.99);

- Çözümde bir tahmin elde edilir.
- Art arda gelen aşamalarda tahmin ıslah edilir.
- Çözüm elde edildiğinde durulur.

Log olabilirlik fonksiyonundaki değişim yeterince küçük olduğunda ya da katsayı tahminlerinin değerlerindeki göreceli değişim yeterince küçük olduğunda hesaplama süreci sona erdirilir (Collett, 2003).

ML tahminleri elde edildikten sonra, genellikle bu tahminler cinsinden tanımlanan tehlike oranları hakkında istatistiksel çıkarımlar yapmaya çalışılmaktadır (Kleinbaum ve Klein, 2005, s.100).

Tablo 2.2: Bir katsayının anlamlılığını değerlendirmede kullanılan yöntemler

Test hipotezleri	Güven aralıkları
Wald testi	Büyük örneklem %95 güven aralığı
LR testi	

Bu üç test içinde özel bir yeri olan skor testi ile iki grubu karşılaştırmada kullanılan Mantel-Haenszel Log-rank testi sonuçları benzerdir. Olabilirlik oran testinin, skor testinden daha iyi sonuçlar verdiği durumlarla da karşılaşılabilir. Skor testi, birden fazla değişken ve sürekli değişkenler için log-rank testinin genelleştirilmiş şeklidir. İki değerli değişkenin gözlenme oranı “ p ” olduğunda, bu orana göre ilgilenilen değişkenin r kez gözlenme olasılığını kullanarak en çok olabilirlik fonksiyonu aşağıdaki gibi yazılmaktadır (Yay, Çoker ve Uysal, 2007):

$$\log L = r \log p + (n - r) \log(1 - p) \quad (2.15)$$

Skor test istatistiğini elde etmek için, $\log L$ 'nin ‘ p ’ye göre alınan birinci ve ikinci türevleri sıfıra eşitlenerek çözülmektedir (Yay, Çoker ve Uysal, 2007):

$$S = (\log L')_2 / (\log L'') \quad (2.16)$$

Skor test istatistiği, serbestlik derecesi tahmin edilen parametre sayısı olmak üzere χ^2 dağılımına uymaktadır. S için hesaplanan değer, χ^2 tablo değeri ile karşılaştırılarak istatistiğin anlamlılığı test edilmektedir (Yay, Çoker ve Uysal, 2007).

ÜÇÜNCÜ BÖLÜM

ORANTILI TEHLİKE VARSAYIMININ TEST EDİLMESİ

Orantılı tehlike varsayımının sağlandığını temin etmek, modelleme sürecinin bir parçası olmasının yanında zamanla değişen etkiler üzerinde faydalı bilgi sağlamada da işe yaramaktadır (Bellera vd., 2010).

Temelini oluşturan varsayımların geçerli kılındığından emin olmaksızın bir Cox modelin uygulanması, oluşan tahminler üzerinde negatif sonuçlara sebep olabilmektedir. Orantılılık varsayımını yerine getirmeyen değişkenler için uygun testlerin sınama gücü azalmakta yani anlamlı bir etki gerçekten var olduğunda, bu karara varmak pek mümkün olmamaktadır. Eğer tehlike orantısı zamanla artıyorsa, tehlike orantısını göz önünde tutarak tahmin edilen katsayılar başlangıçta olduğundan daha yüksek, daha sonra olduğundan daha düşük tahmin edilmektedir. Sabit bir tehlike orantısına sahip modelin değişkenleri için, modelin düşük uyumunun bir sonucu olarak testlerin sınama gücü de azalmaktadır (Bellera vd., 2010).

Cox PH modelinin orantılı tehlike (PH) varsayımıyla ilgili önemli bir özelliği; temel tehlikenin t 'nin bir fonksiyonu olup X 'leri içermemesi, buna karşın üstel ifadenin X 'leri içerip t 'leri içermemesidir. Burada X 'ler, **zamandan bağımsız** X 'ler olarak adlandırılmaktadır (Kleinbaum ve Klein, 2005, s.134).

Yine de, t 'leri içeren X 'leri ele almak mümkündür. Böyle X 'ler **zamana bağlı** değişkenler olarak adlandırılmaktadır. Eğer zamana bağlı değişkenler ele alınırsa, Cox model yine de kullanılabilir fakat böyle bir model artık PH varsayımını sağlamamakta ve **genişletilmiş Cox model** olarak adlandırılmaktadır (Kleinbaum ve Klein, 2005, s.134).

Cox PH modelinden X^* ve X olarak tanımlanan X 'lerin iki özelliğini karşılaştıran bir tehlike orantısını tahmin etmek için genel bir formül elde edilebilmektedir (Kleinbaum ve Klein, 2005, s.135).

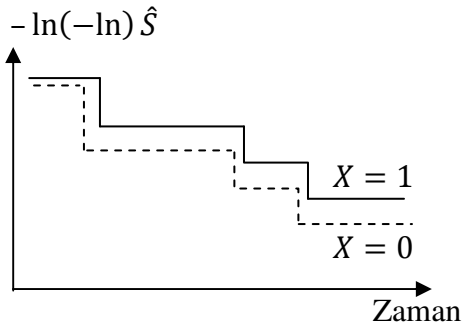
PH varsayımını değerlendirmek için üç yaklaşım bulunmaktadır. Bunlar (Kleinbaum ve Klein, 2005, s.135);

- 1) Grafikselleme
- 2) Uyum iyiliği (GOF) testi
- 3) Zamana bağlı değişkenler

3.1. Orantılı Tehlike Varsayımının Test Edilmesinde Kullanılan Grafikselle Yöntemler

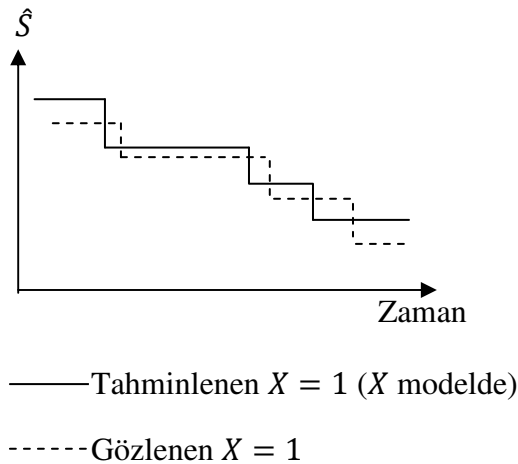
Grafikselle tekniklerin yararlanılan iki çeşidi bulunmaktadır. Bunlardan en popüler olanı, araştırılan değişkenlerin farklı kategorileri(nin kombinasyonu) üzerinde **tahminlenmiş $-\ln(-\ln)$ hayatta kalma eğrileri** karşılaştırmayı içermektedir. Paralel eğriler, PH varsayımının sağlandığını göstermektedir. Şekil 3.1’de tek bir X değişkenini inceleyen bir model için varsayımsal bir log-log grafiği gösterilmektedir (Kleinbaum ve Klein, 2005, s.136);

Şekil 3.1: Log-log hayatta kalma eğrisi grafiği



Alternatif bir grafikselle yaklaşım, **gözlenen ve tahminlenen** hayatta kalma eğrilerini karşılaştırmaktır. Değerlendirilen değişken kategorileri için, PH modeline bu değişkeni koymadan gözlenen eğriler türetilmektedir. Tahminlenen eğriler PH modeline bu değişkeni dahil ederek türetilmektedir. Gözlenen ve tahminlenen eğriler yakınsa, PH varsayımı sağlanmaktadır (Kleinbaum ve Klein, 2005, s.136).

Şekil 3.2: Gözlenen ve tahminlenen hayatta kalma eğrilerinin grafiği



PH varsayımını değerlendirmede, ikinci yaklaşım uyum iyiliği (goodness-of-fit) testlerini içermektedir. Bu yaklaşım, modeldeki diğer değişkenler için ayarlanmış ve modeldeki her değişken için hesaplanabilen büyük örneklem Z ya da ki-kare istatistiklerini vermektedir. Ayrıca, her değişken için standart normal bir istatistikten türetilen p değeri de verilmektedir. Bu p değeri ilgili değişken için PH varsayımını değerlendirmede kullanılmaktadır. Anlamlı olmayan (yani büyük, örneğin 0,10'dan büyük) bir p değeri PH varsayımının sağlandığını, buna karşın küçük bir p değeri (örneğin 0,05'ten küçük) test edilen değişkenin bu varsayımı sağlamadığını göstermektedir (Kleinbaum ve Klein, 2005, s.136).

Zamandan bağımsız bir değişken için PH varsayımını değerlendirmede zamana bağlı değişkenler kullanıldığında Cox model, değerlendirilen zamandan bağımsız değişkeni ve zamanın bir fonksiyonunu içeren **çarpım** (yani etkileşim) **terimlerini** içine alacak şekilde genişletilmektedir (Kleinbaum ve Klein, 2005, s.136).

Örneğin, PH varsayımı cinsiyet değişkeni *GEN* için değerlendiriliyorsa, Cox model *GEN* değişkenine ek olarak "*GEN* \times *t*" değişkenini içerecek şekilde genişletilmektedir. Eğer çarpım teriminin katsayısı anlamlı çıkarsa, PH varsayımının *GEN* değişkeni için ihlal edildiği sonucuna varılmaktadır (Kleinbaum ve Klein, 2005, s.137).

Uyum iyiliği yaklaşımı değerlendirilen her değişken için tek bir test istatistiği sağlamaktadır. Bu yaklaşım, ne grafiksel yaklaşım kadar öznel değildir ne de zamana bağlı değişken yaklaşımı kadar sayısal olarak külfetlidir. Yine de bir uyum iyiliği (GOF) testi diğer iki yaklaşımdan gözlenebilen PH varsayımının belirli bir ihlalini tespit edemeyebildiği için çok "genel" olabilmektedir (Kleinbaum ve Klein, 2005, s.137).

3.1.1. Kaplan-Meier ve Log-Log Grafikleri

Gözlem grubunda hiç sansürlü veri olmaması durumunda farklı yaşam süreleri için hayatta kalma olasılığı aşağıdaki formülü kullanarak tahmin edilebilmektedir (Kleinbaum ve Klein, 2005, s.53).

$$\hat{S}(t_{(j)}) = \frac{t_{(j)} \text{ geçtikten sonra hayatta kalanların sayısı}}{\text{takibin başlangıcındaki denek sayısı}} \quad (3.1)$$

Ancak gözlem grubunda sansürlü veri olması durumunda hayatta kalma olasılıklarını hesaplamak için alternatif bir formül gerekmektedir. Bu alternatif formül Kaplan-Meier (KM, Edward L. Kaplan ve Paul Meier) yaklaşımı olarak adlandırılmakta ve hiç sansürlü verinin olmadığı durumlarda da kullanılabilir (Kleinbaum ve Klein, 2005, s.53).

Kaplan-Meier formülü koşullu olasılık terimlerinin ($Pr(T > t_{(j)} | T \geq t_{(j)})$) çarpımını içermektedir. Yani çarpımdaki her terim, deneğin belirli bir sıralı başarısızlık zamanı $t_{(j)}$ 'ye kadar hayatta kalması koşuluyla bahsedilen başarısızlık zamanını aşma olasılığıdır. Daha genel olarak, bir hayatta kalma olasılığı için herhangi bir KM formülü belirlenen yaşam süresine kadarki çarpım terimleriyle sınırlıdır. KM formülünden sıkça “çarpım-limit formülü” (product limit formula) olarak bahsedilmesinin sebebi budur (Kleinbaum ve Klein, 2005, s.54).

$t_{(0)}$ zamanındaki hayatta kalma tahmini her zaman $\hat{S}(0) = 1$ 'dir. Çünkü geçen zaman sıfır olduğundaki hayatta kalma olasılığını vermektedir (Kleinbaum ve Klein, 2005, s.54).

Bir denek iki sebepten birinden dolayı $t_{(j)}$ zamanında mevcut olmayabilmektedir: (1) $t_{(j)}$ zamanından önce başarısız olması, (2) $t_{(j)}$ zamanından önce sansürlenmiş olması (Kleinbaum ve Klein, 2005, s.55).

Kategorik bir değişkenin varlığında, değişkenin her seviyesi için hayatta kalma zamanının bir fonksiyonu olarak Kaplan-Meier hayatta kalma dağılımı $S(t)$ çizilebilmektedir. Orantılı tehlike varsayımı sağlanmazsa, eğriler muntazam olarak uzaklaşmalıdır. Ayrıca Kaplan-Meier hayatta kalma eğrilerine bir dönüşüm uygulanabilmekte ve log'un doğal logaritma fonksiyonunu temsil ettiği, log hayatta kalma zamanının bir fonksiyonu olarak $\log(-\log(\hat{S}))$ fonksiyonu çizilebilmektedir. Eğer tehlikeler orantılı ise, belirli tabakadaki log-log çizimleri yaklaşık olarak paralel olan sabit farklar sergilemelidir (Bellera vd., 2010).

Bu görsel metotların uygulanması basit olmakla birlikte bazı kısıtlamaları bulunmaktadır. Değişkenler ikiden fazla kategoriye sahip olduğunda, Kaplan-Meier grafikleri orantısızlığı ayırt etmek için kullanışlı değildir. Çünkü böyle bir durumda grafikler kalabalıklaşmaktadır. Aynı şekilde, orantılı tehlike varsayımı ihlal edilmemesine rağmen, uygulamada log-log eğrileri nadiren kusursuz biçimde paralel olabilmektedir. Ayrıca log-log eğrileri, daha uzun zaman noktalarında dağılma eğilimindedir ve bu yüzden daha az kesindir. Yeterli yakınlığın paralele ne kadar yakın olduğunu ve bu yüzden tehlikelerin ne kadar orantılı olduğunu sayısallaştırmak mümkün değildir. Orantılı tehlike hipotezini kabul etme kararı, sıklıkla bu eğrilerin birbiriyle kesişip kesişmediğine bağlıdır. Sonuç olarak orantılı tehlike hipotezini kabul etme kararı öznel ve tutucu olabilmektedir, bu yüzden orantılı tehlike varsayımının ihlal edildiği sonucuna varmak için güçlü kanıt (kesişen doğrular) sahip olunmalıdır. Bu kısıtlardan dolayı, bu çizimler için standart hataların hesaplanması önerilmektedir. Ancak bu yaklaşım hesaplama olarak karmaşık olabilmektedir ve standart

bilgisayar programlarında doğrudan doğruya mevcut değildir. Kaplan-Meier ve log-log çizimleri çoğu standart istatistiksel pakette mevcuttur (Bellera vd., 2010).

Sürekli verilerin KM eğrilerini elde edebilmek ve bu tür verilere log-rank testi uygulayabilmek için ilgilenilen değişkenin kategorize edilmesi gerekmektedir (Kleinbaum ve Klein, 2005, s.62).

$t_{(j)}$ başarısızlık zamanında KM hayatta kalma olasılığı için genel formül aşağıda gösterilmektedir. Bu formül; önceki başarısızlık zamanı $t_{(j-1)}$ 'den sonra hayatta kalma olasılığı ile *en az* $t_{(j)}$ zamanında hayatta olma koşuluyla geçen zaman $t_{(j)}$ için koşullu hayatta kalma olasılığının çarpımını vermektedir (Kleinbaum ve Klein, 2005, s.56).

$$\hat{S}(t_{(j)}) = \hat{S}(t_{(j-1)}) \times \hat{Pr}(T > t_{(j)} | T \geq t_{(j)}) \quad (3.2)$$

Ayrıca yukarıdaki KM formülü; $t_{(j-1)}$ ve daha önceki başarısızlık zamanları için koşullu olasılıkları tahmin eden tüm kesirlerin çarpımı olan $\hat{S}(t_{(j-1)}) = \prod_{i=1}^{j-1} \hat{Pr}(T > t_{(i)} | T \geq t_{(i)})$ hayatta kalma olasılığı yerine koyulursa, çarpım limiti olarak ifade edilebilmektedir (Kleinbaum ve Klein, 2005, s.56);

$\hat{S}(t_{(j)}) = \prod_{i=1}^j \hat{Pr}(T > t_{(i)} | T \geq t_{(i)})$ (Bu formül, $Pr(A \cap B) = Pr(A) \times Pr(B|A)$ olasılık kuralından türetilmektedir.)

Bu alternatif iki formül genel olarak aşağıdaki gibi gösterilmektedir (Kleinbaum ve Klein, 2005, s.71);

$$S_{(j)} = S_{(j-1)} Pr(T > t_{(j)} | T \geq t_{(j)}) \quad (3.3)$$

$$S_{(j)} = \prod_{i=1}^j Pr(T > t_{(i)} | T \geq t_{(i)}) \quad (\text{çarpım limit formülü}) \quad (3.4)$$

Bir log-log hayatta kalma eğrisi, tahminlenmiş bir hayatta kalma eğrisinin sadece bir dönüşümüdür. Tahminlenmiş hayatta kalma eğrisinin sonuçları, tahminlenmiş bir hayatta kalma olasılığının *iki kez* doğal logaritmasının alınmasından meydana gelmektedir (Kleinbaum ve Klein, 2005, s.137).¹

¹ $S(t, \mathbf{X})$, herhangi bir t değeri ve \mathbf{X} vektörünün herhangi bir tanımı için değeri 0 ile 1 arasında bir sayı olan bir hayatta kalma olasılığını göstermektedir. 0 ile 1 arasında herhangi bir sayının doğal logaritması negatif bir tamsayıdır, bu yüzden hem $S_0(t)$ 'nin logaritması hem de $S(t, \mathbf{X})$ 'in logaritması negatif sayıdır. İkinci defa log almadan önce bu ifadenin önüne eksi işareti koymak zorunda olunmasının sebebi, negatif bir sayının logaritması diye bir şeyin olmamasıdır (Kleinbaum ve Klein, 2005, s.139).

1. log:

$$\ln S(t, \mathbf{X}) = e^{\sum_{i=1}^p \beta_i X_i} \times \ln S_0(t)$$

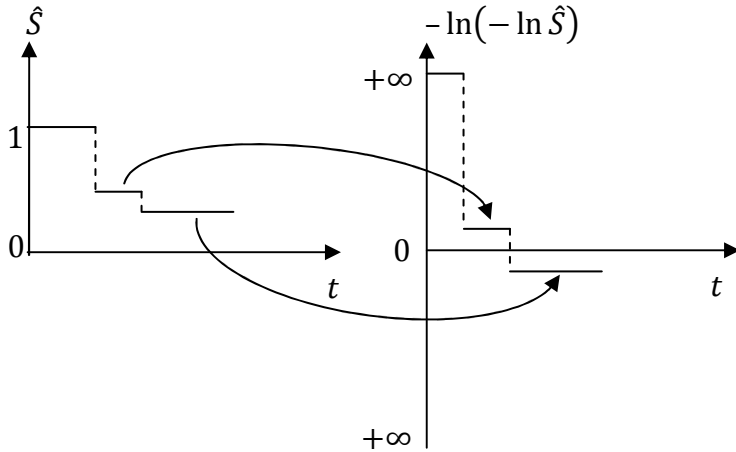
$$0 \leq S(t, \mathbf{X}) \leq 1$$

$\ln(\text{olasılık}) = \text{negatif değer}$, bu yüzden $\ln S(t, \mathbf{X})$ ve $\ln S_0(t)$ negatiftir.

İkinci logaritma alındıktan sonra, bu ifade iki terimin toplamı olarak tekrar yazılabilir (Kleinbaum ve Klein, 2005, s.139);

Hayatta kalma eğrisi genellikle bir adım fonksiyonu olarak çizildiğinden, log-log eğrisi de bir adım fonksiyonu olarak çizilecektir (Kleinbaum ve Klein, 2005, s.138).

Şekil 3.3: Gözlenen ve tahminlenen hayatta kalma eğrilerinin grafiği



Tahminlenen bir hayatta kalma eğrisinin y eksenindeki ölçeği 0 ile 1 aralığında seyretmektedir. Buna karşın, bu hayatta kalma eğrisine ait bir $-\ln(-\ln)$ eğrisinin ölçeği $-\infty$ ile $+\infty$ aralığındadır (Kleinbaum ve Klein, 2005, s.138).

Şekil 3.4: y eksenini ölçeği

$$\begin{array}{c} 1 \\ 0 \end{array} \left| \begin{array}{c} \hat{S} \\ \\ \\ \end{array} \right. \begin{array}{c} +\infty \\ \\ \\ -\infty \end{array} \left| \begin{array}{c} -\ln(-\ln) \hat{S} \\ \\ \\ \end{array} \right.$$

İki farklı bireyin \mathbf{X}_1 ve \mathbf{X}_2 'ye tekabül ettiği \mathbf{X} vektörünün iki farklı tanımı ele alındığında, herhangi bir \mathbf{X} bireyi için önceki log-log eğrisi ifadesinde \mathbf{X} yerine \mathbf{X}_1 ve \mathbf{X}_2

2. log:

$$\ln[-\ln S(t, \mathbf{X})] = \ln \left[-e^{\sum_{i=1}^p \beta_i X_i} \times \ln S_0(t) \right]$$

$$\ln[-\ln S(t, \mathbf{X})] = \ln \left[e^{\sum_{i=1}^p \beta_i X_i} \right] + \ln[-\ln S_0(t)]$$

$$\ln[-\ln S(t, \mathbf{X})] = \sum_{i=1}^p \beta_i X_i + \ln[-\ln S_0(t)]$$

veya

$$-\ln[-\ln S(t, \mathbf{X})] = -\sum_{i=1}^p \beta_i X_i - \ln[-\ln S_0(t)]$$

İkinci log hem pozitif hem de negatif olabilmektedir. Daha fazla log alınmayacağından, ikinci bir negatif alınmasına gerek yoktur. Ancak, tutarlılık amacıyla ortak bir uygulama $-\ln -\ln$ ifadesini elde etmek için ikinci logaritmanın önüne bir eksi işareti koymaktır (Kleinbaum ve Klein, 2005, s.139).

koyularak, bu bireyler için log-log eğrileri aşağıdaki gibi elde edilmektedir (Kleinbaum ve Klein, 2005, s.140);

$$\begin{cases} \mathbf{X}_1=(X_{11},X_{12},\dots,X_{1p}) \\ \mathbf{X}_2=(X_{21},X_{22},\dots,X_{2p}) \end{cases} \begin{cases} -\ln[-\ln S(t, \mathbf{X}_1)] = -\sum_{i=1}^p \beta_i X_{1i} - \ln[-\ln S_0(t)] \\ -\ln[-\ln S(t, \mathbf{X}_2)] = -\sum_{i=1}^p \beta_i X_{2i} - \ln[-\ln S_0(t)] \end{cases} \quad (3.5)$$

Birinci log-log eğrisinden ikincisi çıkarılırsa, aşağıdaki ifade elde edilecektir (Kleinbaum ve Klein, 2005, s.140);

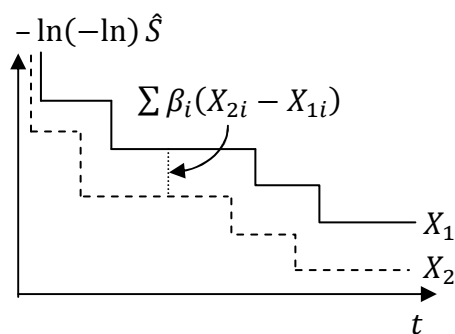
$$-\ln[-\ln S(t, \mathbf{X}_1)] - (-\ln[-\ln S(t, \mathbf{X}_2)]) = \sum_{i=1}^p \beta_i (X_{2i} - X_{1i}), t'yi \text{ i\c{c}ermez.} \quad (3.6)$$

Bu ifade iki bireyin kestirici deęerler arasındaki farkın doęrusal bir toplamıdır. Temel hayatta kalma fonksiyonu sadeleşmektedir, bu yüzden log-log eğrileri arasındaki fark t zamanını içermeyen bir ifadedir. Alternatif olarak, Denklem (3.6)'daki eşitlik aşağıdaki şekilde de yazılabilir (Kleinbaum ve Klein, 2005, s.140);

$$-\ln[-\ln S(t, \mathbf{X}_1)] = -\ln[-\ln S(t, \mathbf{X}_2)] + \sum_{i=1}^p \beta_i (X_{2i} - X_{1i}) \quad (3.7)$$

Denklem (3.6)'daki formül, bir Cox PH modeli kullanıldığında ve bireyler için tahminlenmiş log-log hayatta kalma eğrileri aynı grafik üzerinde çizildiğinde, iki çizimin yaklaşık olarak paralel olacağını göstermektedir. İki eğri arasındaki uzaklık, kestirici deęerler arasındaki farkın zamanı içermeyen doęrusal bir ifadesidir. Genel olarak iki eğri arasındaki dikey uzaklık sabitse, eğriler paraleldir (Kleinbaum ve Klein, 2005, s.140).

Şekil 3.5: Gözlenen ve tahminlenen eğriler arasındaki uzaklık



Cox PH modeli için log-log hayatta kalma çizimlerinin paralellięi, PH varsayımını deęerlendirmede grafiksel bir yaklařım sağlamaktadır. Yani bir PH modeli kestiricilerin belirli bir seti için uygun ise, farklı bireyler için log-log hayatta kalma eğrilerinin deneysel çizimlerinin yaklaşık olarak paralel olması beklenmelidir (Kleinbaum ve Klein, 2005, s.140).

PH varsayımını deęerlendirmede bu grafiksel yaklařım ile ilgili bazı problemler bulunmaktadır. Temel problem, “ne kadar paralel, paraleldir?”e nasıl karar verileceęi ile

ilgilidir. Bu karar, özellikle çalışma hacmi görece olarak küçükse belirli bir veri seti için oldukça öznel olabilmektedir. Bu karar için, log-log eğrilerinin paralel olmadığına dair **güçlü** bir kanıt olmadıkça, PH varsayımının sağlandığını varsayan koruyucu bir strateji önerilmektedir (Kleinbaum ve Klein, 2005, s.142).

Diğer bir problem sürekli bir değişkenin nasıl sınıflandırılacağı ile ilgilidir. Eğer çok sayıda kategori seçilirse, her kategoride veri “seyrekleşecektir” bu da farklı eğrileri karşılaştırmayı güçleştirecektir. Ayrıca farklı sınıflandırmalar, farklı grafiksel şekiller verebilmektedir (Kleinbaum ve Klein, 2005, s.143).

Sürekli değişkenleri sınıflandırmada; eğer mümkünse kategori sayısının küçük (örneğin; iki ya da üç) tutulması ve kategori seçiminin mümkün olduğunca anlamlı olması ve ayrıca makul bir sayı dengesinin sağlanması tavsiye edilmektedir (Kleinbaum ve Klein, 2005, s.143).

Bu iki probleme ek olarak başka bir sorun, log-log hayatta kalma eğrilerini kullanarak aynı anda birçok değişken için PH varsayımının nasıl değerlendirileceğidir. Eşanlı karşılaştırmalar için bir strateji, tüm değişkenleri ayrı ayrı sınıflandırıp kategorilerin kombinasyonlarını oluşturmak ve daha sonra tüm kombinasyonların log-log eğrilerini aynı grafik üzerinde karşılaştırmaktır (Kleinbaum ve Klein, 2005, s.143).

Bu stratejinin bir dezavantajı, kombinasyonların sayısı makul olarak artsa bile verinin yine “seyrelmeye” meyilli olmasıdır. Ayrıca her kombine edilmiş kategori için yeterli sayıda gözlem olsa bile, eğrilerin paralel olmaması durumundan hangi değişkenlerin sorumlu olduğunu belirlemek genellikle zor olmaktadır (Kleinbaum ve Klein, 2005, s.143).

Birçok kestiriciyi birlikte ele almada alternatif bir grafiksel yöntem; PH varsayımını sağladığı varsayılan diğer kestiriciler için düzeltilen bir kestiricinin PH varsayımını değerlendirmektir. Bu yöntem Kaplan-Meier eğrilerini kullanmaktan ziyade, ayarlanmış log-log hayatta kalma eğrilerinin bir karşılaştırmasını içermektedir (Kleinbaum ve Klein, 2005, s.144).

3.1.2. Gözlenene Karşı Beklenen Çizimler

PH varsayımını değerlendirmek için gözlenene karşı beklenen çizimlerin kullanımı, uyum iyiliği (GOF) testi yaklaşımının grafiksel halidir. Bu yüzden log-log hayatta kalma eğrisi yaklaşımına makul bir alternatiftir (Kleinbaum ve Klein, 2005, s.145).

Gözlenene karşı beklenen yöntemi, log-log yaklaşımındaki gibi iki yöntemin ikisini ya da herhangi birini kullanarak yürütülebilmektedir: 1) herhangi bir zamanda değişkenler için PH varsayımını değerlendirme 2) diğer değişkenler için ayarladıktan sonra PH varsayımını değerlendirme. Herhangi bir zamanda değişkenler için PH varsayımını değerlendirmede; öncelikle değerlendirilecek kestiricinin kategorileri için veri tabakalandırılmakta, ardından her kategori için ayrı ayrı KM eğrileri türetilerek “gözlenen” çizimler elde edilmektedir. “Beklenen” çizimleri elde etmek için değerlendirilen kestiriciyi içeren bir Cox PH modeli tahmin edilmektedir. Tahminlenmiş hayatta kalma eğrileri için kestiricinin her bir kategorisinin değerini formülde ayrı ayrı yerine koyarak beklenen çizimler elde edilmektedir. Dolayısıyla her kategori için ayrı bir tahminlenmiş hayatta kalma eğrisi elde edilmektedir (Kleinbaum ve Klein, 2005, s.146).

Eğer değerlendirilen kestiricinin her bir kategorisi için gözlenen ve beklenen çizimler birbirine “yakın”sa, PH varsayımının sağlandığı sonucuna ulaşılabilmektedir. Ancak bir ya da daha fazla kategori oldukça farklı gözlenen ve beklenen çizimler gösterirse, PH varsayımının ihlal edildiği sonucuna varılacaktır (Kleinbaum ve Klein, 2005, s.147).

Bu grafiksel yaklaşımın belirgin bir dezavantajı, belirli bir kategori için gözlenene karşı beklenen eğriler karşılaştırıldığında “ne kadar yakın, yakındır”a karar vermektir. Bu log-log hayatta kalma eğrilerini karşılaştırmada “ne kadar paralel, paraleldir”e karar vermeye benzerdir. Burada da, sadece gözlenen ve beklenen çizimler aşırı derece farklı olduğunda, PH varsayımının sağlanmıyor olarak ele alınması tavsiye edilmektedir (Kleinbaum ve Klein, 2005, s.147).

Sürekli bir değişkenin PH varsayımını değerlendirmek için gözlenene karşı beklenen çizimler kullanıldığında, kategorik değişkenlerde olduğu gibi sürekli değişkenin kategorilerinden tabaka oluşturup her kategori için KM eğrilerini elde ederek gözlenen çizimler türetilir (Kleinbaum ve Klein, 2005, s.148).

Ancak kesikli kestiriciler için beklenen çizimleri hesaplamada mevcut iki seçenek vardır. Bir seçenek, k kategorileri göstermek üzere $k - 1$ kukla değişken içeren bir Cox PH modeli kullanmaktır. Belirli bir c kategorisi için beklenen çizim, o kategoriyi tanımlayan kukla değişkenlerin değerlerini tahminlenmiş hayatta kalma eğrileri formülünde yerine koyarak ayarlanmış bir hayatta kalma eğrisi elde edilmesidir (Kleinbaum ve Klein, 2005, s.148).

$$h(t, \mathbf{X}) = h_0(t) \exp\left(\sum_{i=1}^{k-1} \beta_i X_i\right) \quad (3.8)$$

$$\hat{S}(t, \mathbf{X}_c) = [\hat{S}_0(t)]^{\exp(\sum \hat{\beta}_i X_{ci})} \quad (3.9)$$

Burada $\mathbf{X}_c = (X_{c1}, X_{c2}, \dots, X_{c,k-1})$ c kategorisi için kukla değişkenlerin değerini vermektedir.

İkinci seçenek, değerlendirilen sürekli kestiriciyi içeren bir Cox PH modeli kurmaktır. Daha sonra ayarlanmış hayatta kalma eğrileri gibi farklı kategoriler için kestirici değerlerini tanımlayarak beklenen çizimler elde edilir. Kestirici değerleri tanımlamak için her kategorinin ortalama değeri alınabilir (Kleinbaum ve Klein, 2005, s.149).

$$h(t, X) = h_0(t) \exp(\beta X) \quad (3.10)$$

↓
Sürekli

$$\hat{S}(t, \bar{X}_c) = [\hat{S}_0(t)]^{\exp(\hat{\beta} \bar{X}_c)} \quad (3.11)$$

\bar{X}_c : c kategorisi içindeki X değişkeninin ortalama değerini göstermektedir.

Diğer değişkenler için ayarlama stratejisi ise, gözlenen çizimleri oluşturmak için tabakalandırılmış bir Cox PH modeli kullanılmaktadır. Gözlenen çizimlerde, PH modeli ayarlanmış değişkenleri içermektedir ve tabakalandırılan değişken ise değerlendirilen kestiricidir (Kleinbaum ve Klein, 2005, s.146).

3.1.3. Uyum İyiliği Araştırması

Bu yaklaşım modeldeki her değişken için hesaplanabilen ki-kare istatistiklerine dayanan testler yardımıyla orantılı tehlike varsayımını test eder. Uyum iyiliği (GOF) testi, orantılı tehlike varsayımının değerlendirilmesinde istatistiksel güvenilirliği yüksek bir test sağlar. Bu nedenle uyum iyiliği testi kullanarak grafiksel yaklaşımlardan daha kesin ve objektif karar verilebilir (Göz Çekçeki, 2007, s.32).

Uyum iyiliğinin araştırılmasında, modeldeki her bir değişken için olayın gerçekleştiği her gözlemin Schoenfeld artıkları hesaplanır. Uyum iyiliği araştırılan değişkenin t_j zamanında i . gözlemi olaya maruz kaldığında hesaplanan Schoenfeld artığı; gözlenen değer, t_j zamanında hala risk altında olan diğer gözlemler için ağırlıklı ortalamasıdır. Ağırlıklar diğer gözlemlerin tehlike oranlarıdır. Belirli bir değişken için PH varsayımı sağlanıyorsa, o değişkenin Schoenfeld artıkları yaşam süresi ile ilişkili olmayacaktır (Kleinbaum ve Klein, 2005, s.151).

Uyum iyiliği testinin adımları; 1) her değişken için Schoenfeld artıklarının elde edilmesi, 2) başarısızlık zamanlarının gerçekleşme zamanına göre sıralanıp numaralandırılması yoluyla yeni bir değişken elde edilmesi 3) ilk iki adımda elde edilen değişkenlerin aralarında fark olmadığı sıfır hipotezi altında, ilişkili olup olmadıklarının test

edilmesidir. Dolayısıyla sıfır hipotezinin reddedilmesi PH varsayımının ihlal edildiği anlamına gelecektir (Kleinbaum ve Klein, 2005, s.152).

Testin uygulanmasında bahsedilen üçüncü adım grafiksel olarak da değerlendirilebilir. Sıralanmış başarısızlık zamanlarına karşı Schoenfeld artıklarının çizimi yapıldığında elde edilen eğri yatay eksene paralel (eğrinin eğimi sıfır) ise bu iki değişken arasında ilişki olmadığı sonucuna ulaşılabilecektir. Başka bir değişle, değişkenin her bir başarısızlık zamanı için hesaplanan artıkları, başarısızlık zamanlarına bağlı olmayacaktır.

Uyum iyiliği testlerinin avantajlarının yanı sıra dezavantajları da bulunmaktadır. İlk olarak, uyum iyiliği testi genel anlamda orantılı tehlike varsayımından sapmaları tespit etmek için yapılmış genel bir kontrol mekanizmasıdır. Ancak bu test orantılı tehlike varsayımında özel bir sapmayı tespit edemeyebilir. Daha önce anlatılan grafiksel testler bu tarz özel sapmaların tespitinde daha kullanışlıdır. Sonuç olarak orantılı tehlike varsayımının değerlendirilmesinde son karar verilirken hem grafiksel yöntemlerin hem de uyum iyiliği testinin kullanılması tavsiye edilmektedir (Göz Çekçeki, 2007, s.32).

3.2. Orantılı Tehlike Varsayımının Test Edilmesinde Kullanılan İstatistiksel Yöntemler

Orantılı tehlike varsayımını kontrol eden grafiksel metotlar muntazam bir tanımlayıcı test sağlamamakta ve doğrulayıcı yaklaşımlar gerektirmektedir. Orantısızlığın test edilmesi ve hesaplanması için birçok seçenek mevcuttur (Bellera vd., 2010).

Cox, zamana bağlı yapıda bir değişkeni Cox modele katarak yani zamanı içeren bir etkileşim terimi ekleyerek ve onun anlamlılığını test ederek, orantısızlığın giderilmesini amaçlamıştır. Bazı X değişkenlerinin zamanla değişen bir etkiye sahip olup olmadığı değerlendirilmek istendiğinde, tehlike fonksiyonunun yapısı değişmektedir. X (sürekli veya kategorik) kestiricisi ve t zamanının bir fonksiyonu ($f(t) = t, t^2, \log(t), \dots$) arasında bir etkileşim terimi oluşturarak zamana bağlı bir değişken meydana getirilebilmektedir. Bu etkileşim Denklem (1.7)'ye eklenirse, tehlike şu şekilde olmaktadır (Bellera vd., 2010);

$$h_X(t) = h_0(t) \exp[\beta X + \delta X f(t)] \quad (3.12)$$

X değişkenindeki bir birimlik artış için tehlike oranı $HR(t) = h_{X+1}(t)/h_X(t) = \exp[\beta + \delta X f(t)]$ tarafından verilmektedir ve bu oran $f(t)$ fonksiyonu boyunca zamana bağlıdır. Eğer $\delta > 0$ ($\delta < 0$) ise, tehlike oranı zamanla artacaktır (azalacaktır). PH testi, δ 'nın sıfırdan anlamlı derecede farklı olup olmadığının testine eşdeğerdir. Polinom ya da üstel

bozunma gibi farklı zaman fonksiyonları kullanılabilen fakat sıklıkla zamanın çok basit durağan fonksiyonları tercih edilmektedir. Ayrıca t zaman değerleri tehlike orantısı fonksiyonuna yerleştirilebildiğinden, bu modelleme yaklaşımı farklı zaman noktalarında tehlike orantısının tahminlerini vermektedir (Bellera vd., 2010).

Zamana bağlı değişkenler, PH varsayımını değerlendirmek için esnek bir metot ve zamanla göreceli riskin bağımlılığına dair model kurmak için bir yaklaşım sağlamaktadır. Ancak bu yaklaşım dikkatle kullanılmalıdır. Eğer seçilen zaman fonksiyonu yanlış tanımlanırsa, sonuca varılan model isabetli olmayacaktır. Bu, daha esnek yaklaşım üzerine bu metodun dezavantajıdır (Bellera vd., 2010).

Değişkenlerin zamanla değişen etkilerini test etmek için diğer bir basit yaklaşım, farklı zaman periyotları için farklı Cox modeller kurmayı gerektirmektedir. Nitekim, orantılı tehlike varsayımı takip edilen tüm periyot boyunca geçerli olmayabilmekte, daha kısa bir zaman aralığında sağlanabilmektedir. Belirli bir kesme zamanı değeri, bir ilgi konusu olmadığı sürece medyan olay zamanına dayanarak verinin iki altkümüsi oluşturulabilir. Yani, ilk analiz bu zaman noktasını aşan risk altında olan herkesi sansürleyerek ve ikinci analiz ise bundan sonra risk altında olan sadece bu denekleri alarak yürütülür. Bu durumda, modellerin yorumu hayatta kalma süresinin uzunluğuna bağlı olmaktadır ve bu nedenle sonuçlar dikkatle yorumlanmalıdır. Analizin periyodu kısaltılsa bile, yine de bu azaltılan zaman periyotları içinde orantılı tehlike varsayımının ihlal edilmediğinden emin olunmalıdır. Dahası, daha az sayıda olay zamanı ele alındığından, analiz azalan sınama gücünden zarar görmektedir. Sonuç olarak, her şeyden önce bu metodu uygulamak basittir ve bu metot bazı durumlarda yani kısa bir zaman aralığı ile ilgileniliyorsa yeterli bilgi sağlayabilmektedir. Böyle bir testi gerçekleştirmek için farklı bir parametre grubu oluşturmak gerekmektedir. Bu metot orantılı tehlike varsayımını doğrudan test etmemektedir (Bellera vd., 2010).

Moreau vd. (1985) tarafından önerildiği gibi zaman eksenini kısımlara ayırarak da orantısızlığı hesaplamak mümkündür. Zaman eksenini kısımlara ayırmakta ve zaman eksenini kısımlara ayırdıktan sonra her aralığın içinde tehlike oranları tahmin edilmektedir. Bu yüzden orantısızlığı test etme, belirli bir anda başlayan tehlike oranlarının anlamlı derecede farklı olup olmadığını test etmeye eşdeğerdir. Ancak sonuçlar bazen zaman aralıklarının sayısı tarafından sürüklenebilmektedir, bu yüzden zaman aralıkları dikkatli biçimde seçilmelidir (Bellera vd., 2010).

Parçalı orantılı tehlike modeli, orantılı olmayan tehlikelerin bir adım fonksiyonu olarak gösterilmesidir. Tehlike oranları önceden belirlenen zaman aralarında sabit, ancak

aralıklar arasında farklıdır. Zaman aralığı sayısı $= 1, \dots, r$ olmak üzere i 'nci aralığın içinde tehlike (Başar, 2006);

$$h(t) = h_0(t) \exp[\sum(\beta_j + \gamma_{ij})X_j] \quad (3.13)$$

biçiminde tanımlanmaktadır. Burada $\gamma_{ij} = 0$ 'dır. Log tehlike orantısı ilk aralıkta β_j 'ye, sonraki aralıklarda $(\beta_j + \gamma_{ij})$ 'ye eşittir. $\gamma_2 = \dots = \gamma_r = 0$ iken model (2), model (1)'in özel bir durumu olmaktadır. $p \geq 1$ sayıda değişken kümesi için orantılı tehlike hipotezi $p(r - 1)$ serbestlik dereceli olabilirlik oran testi ile test edilmektedir (Başar, 2006).

Cox modelde olduğu gibi orantılı tehlike varsayımından feragat etmek, diğer bir seçenektir. Aslında zamanla değişen etkileri hesaplamak için; toplamsal modelleri, hızlandırılmış başarısızlık zamanı modellerini, regresyon kama modellerini ya da kesirli polinomları içeren başka güçlü istatistiksel modeller mevcuttur (Bellera vd., 2010).

Son olarak, zamanla değişen bir etkiye sahip olduğundan şüphelenilen değişkenin tabakalandırıldığı istatistiksel bir analiz gerçekleştirilebilir. Bu değişken kategorik olmalı ya da kategorize edilmelidir. Her k tabakası, farklı bir temel tehlike fonksiyonuna sahiptir, fakat katsayı vektörü β için ortak değerlere sahiptir. Yani k . tabakadaki bir birey için tehlike $h_k(t) = \exp(\beta X)$ 'tir. Tabakalandırma, her tabakada diğer değişkenlerin aynı yönde hareket ettiğini varsaymaktadır yani tehlike oranları tabaka boyunca benzerdir. Tabakalandırma, orantısızlık problemini kaldırmada etkili olmasına ve uygulaması basit olmasına rağmen bazı sakıncalara sahiptir. Bunlardan en önemlisi orantısız bir değişkenin tabakalandırılması, değişkenin gücünün tahmin edilmesini ve Cox model dahilinde test edilmesini engellemektedir. Bu yüzden, tabakalandırma için kullanılan değişkenin etkisini ölçmeyle doğrudan ilgilenilmiyorsa, bu yaklaşım seçilmelidir. Dahası tabakalandırılmış bir Cox model, güç kaybına yol açabilmektedir çünkü ayrı tehlike fonksiyonlarını tahmin etmek için daha fazla veri kullanılmaktadır. Bu etki; denek ve tabaka sayısına bağlı olacaktır. Zamanla değişen riske sahip birkaç değişken varsa, büyük ihtimalle toplam sınama gücünü tekrar azaltacak olan bu çoklu faktörler üzerine modelin tabakalandırılması gerekecektir (Bellera vd., 2010).

3.2.1. Schoenfeld Artıkları ve PH Testi

PH varsayımı, modelin artıklarını kullanarak da araştırılabilmektedir. Bir artık, gözlenen veri ve model varsayımı altında beklenen veri arasındaki farkı ölçmektedir. Bu yöntemde, Schoenfeld artıkları hesaplanmakta ve orantılı tehlike varsayımı altında her başarısızlık zamanında rapor edilmektedir. Bu anlamda sansürlü denekler için Schoenfeld

artıkları tanımlanmamaktadır. Schoenfeld artıkları, her birey için ortak değişim değeri olarak tanımlanmaktadır ki bu modelin dayandığı hipotezi göz önünde tutarak *başarısızlıktan* onun beklenen değerinin çıkarılmasıdır. Her değişken için her bireyin ayrı bir artığı vardır. O halde, Schoenfeld artıklarının düzgünleştirilmiş bir çizimi, log tehlike orantısını görselleştirmede direk olarak kullanılabilir. Tehlikelerin orantılı olduğu varsayılırsa, Schoenfeld artıkları zamandan bağımsızdır. Bu yüzden, zamana karşı rassal olmayan bir kalıbı ortaya koyan bir çizim orantısızlığın kanıtıdır. Grafikselleştirilerek, bu metot $\log(-\log(\hat{S}))$ fonksiyonunun çiziminden daha güvenilir ve yorumlanması kolaydır. Zaman ile doğrusal bir ilişkinin varlığı, basit doğrusal bir regresyon ve bir test trendi uygulayarak test edilebilmektedir. Bir eğimin sıfırdan anlamlı derecede farklı olması, orantılılığa karşıt bir kanıt olabilmektedir: artan (azalan) bir trend zamanla artan (azalan) bir tehlike orantısını gösterebilmektedir. Bu testin uygulanmasına ek olarak artık çizimine de dikkatli bakılması tavsiye edilmektedir. Çünkü bazı kalıplar, çizimler (ikinci dereceden, logaritmik) üzerinde görünür olabilmekte fakat istatistiksel testler tarafından saptanamayabilmektedir. Dahası, uç değerlerin aşırı etkisi açık olabilmektedir (Bellera vd., 2010).

Düzenleştirilmiş Schoenfeld artıklarına dayanan bu metot, zamana bağlı tahminler sağlanmasına rağmen bazı sakıncalara sahiptir. Sonuçta oluşan zamana bağlı tahminler ile ilgili kesin olmayan tahminlerin, uygulamada kullanılması zor olabilmekte ve bulunan tahminleyici, tutarlılık gibi iyi istatistiksel özelliklere sahip olmayabilmektedir. Daha da önemlisi, Schoenfeld artıklarına dayanan trend testlerinin p değerleri, modelin diğer değişkenleri için Cox modelin düzeltildiğini varsayarak modelin her değişkeni için ayrı ayrı elde edilmektedir, bu anlamda sonuçlar dikkatli bir biçimde yorumlanmalıdır. Schoenfeld artıklarına dayanan testler, çoğu standart istatistik paketinde kolayca uygulanabilmektedir (Bellera vd., 2010).

Orantılı tehlikeleri değerlendirmek için, Schoenfeld artıklarının birikimli toplamı ya da eşdeğer şekilde gözlenen skor süreci de kullanılabilir. Grafikselleştirilerek gözlenen skor süreci, temel Cox modelinin doğru olduğunu yani orantılı tehlikeleri varsayan simüle edilmiş süreçlerle birlikte, modelin her değişkeni için zamana karşı çizilmektedir. Simüle edilen süreçlerden gözlenen skor sürecinin herhangi bir sapması orantılılığa karşıt bir kanıttır. Bu çizimler, uyum eksikliği mevcut olduğunda da değerlendirme için kullanılabilir. Özellikle, simüle edilmiş sürecin oldukça üzerinde gözlenen bir skor, ortalamadan daha yüksek bir etkinin göstergesidir ve bu durumun tam tersi de geçerlidir. Uyum iyiliği testleri, birikimli artıklar baz alınarak uygulanabilmektedir. Schoenfeld artıkları ile karşılaştırıldığında birikimli artıklara dayanan yaklaşım, bazı sorunların üstesinden gelmektedir çünkü sonuçta

oluşan tahminler daha iyi istatistiksel özelliklere sahip olmaya ve türetilen p değerleri doğru olmaya meyillidir. Birikimli artıklar yaklaşımı, bazı standart istatistiksel paketlerde uygulanabilmektedir (Bellera vd., 2010).

3.2.2. Log-Rank Testi

İki ya da daha fazla grubun KM eğrilerinin istatistiksel olarak eşit olup olmadığını değerlendirmede, en fazla rağbet gören yöntem log-rank testidir. İki KM eğrisinin “istatistiksel olarak eşit” olması ifadesi ile, “toplu anlam”da iki eğriyi karşılaştıran bir test yönteminin uygulanması kastedilmektedir (Kleinbaum ve Klein, 2005, s.57);

Log-rank testi, karşılaştırılan KM eğrilerinin toplu olarak ilişkisini veren bir istatistik test kriterini kullanan bir büyük örneklem ki-kare testidir. Diğer ki-kare testi çeşitlerinde kullanılan birçok istatistik gibi bu istatistik, sonuçların kategorileri üzerinde gözleneneye karşı beklenen sıklık oranlarını (expected cell counts) kullanmaktadır. Log-rank istatistiği için kategoriler, analiz edilen veri setinin tümü için sıralanmış her bir başarısızlık zamanı tarafından tanımlanmaktadır. İlgilenilen grubun log-rank test istatistiği; başarısızlık zamanlarının her birinde gözlenenenden beklenen sıklık sayıları çıkarılıp, bulunan değerler toplanarak elde edilmektedir. İki grup için beklenen sıklık sayıları aşağıdaki gibi hesaplanmaktadır (Kleinbaum ve Klein, 2005, s.58-59);

$$e_{1j} = \left(\frac{n_{1j}}{n_{1j} + n_{2j}} \right) \times (m_{1j} + m_{2j}) \quad \text{ve} \quad e_{2j} = \left(\frac{n_{2j}}{n_{1j} + n_{2j}} \right) \times (m_{1j} + m_{2j}) \quad (3.14)$$

\uparrow \uparrow
 Risk setindeki İki grubun
 oran başarısızlık sayıları

1. grup için bu formül, j zamanındaki beklenen olay sıklığını (yani e_{1j}) vermektedir. Hesaplama; j zamanında risk altında olan grubun, her iki gruptaki toplam denek sayısına oranı (yani $n_{1j}/n_{1j} + n_{2j}$) ile iki grup için aynı zamandaki toplam başarısızlık sayısı (yani $(m_{1j} + m_{2j})$) çarpılarak yapılmaktadır. 2. grup için de e_{2j} benzer şekilde hesaplanmaktadır. İki gruptan oluşan veri setleri için, başarısızlığın beklenen sıklık değeri genel olarak aşağıdaki formülle hesaplanmaktadır (Kleinbaum ve Klein, 2005, s.58);

$$e_{ij} = \left(\frac{n_{ij}}{n_{1j} + n_{2j}} \right) \times (m_{1j} + m_{2j}), \quad i = 1,2 \text{ için} \quad (3.15)$$

İki gruptan oluşan veri seti için log-rank istatistiği aşağıdaki şekilde hesaplanabilmektedir. Formülde $O_i - E_i$ ifadesi; i . grup için gözlenen eksi beklenen skorların toplamıdır (Kleinbaum ve Klein, 2005, s.59);

$$\text{Log-rank istatistiği} = \frac{(O_i - E_i)^2}{\text{Var}(O_i - E_i)}, \quad i = 1,2 \text{ için} \quad O_i - E_i = \sum_j (m_{ij} - e_{ij}) \quad (3.16)$$

Tahminlenen varyans aşağıdaki şekilde hesaplanmaktadır. Grup sayısı 2 olduğunda, varyans formülü her grup için aynı olmaktadır. Varyans formülü, her gruptaki risk seti sayısını (n_{ij}) ve j zamanında her gruptaki başarısızlık sayısını (m_{ij}) içermektedir. Toplama işlemi, tüm farklı başarısızlık zamanları üzerinedir (Kleinbaum ve Klein, 2005, s.60);

$$\text{Var}(O_i - E_i) = \sum_j \frac{n_{1j}n_{2j}(m_{1j} + m_{2j})(n_{1j} + n_{2j} - m_{1j} - m_{2j})}{(n_{1j} + n_{2j})^2(n_{1j} + n_{2j} - 1)}, \quad i = 1,2 \text{ için} \quad (3.17)$$

Test edilen sıfır hipotezi, iki hayatta kalma eğrisi arasında toplu fark olmadığıdır. Bu sıfır hipotezi altında, log-rank istatistiği bir serbestlik dereceli ki-kare dağılımına yaklaşmaktadır. Bu yüzden, log-rank testinin p değeri ki-kare dağılımı tablosundan belirlenmektedir. p değeri 0,05'ten küçük olursa sıfır hipotezi reddedilmekte ve iki grup için KM hayatta kalma eğrilerinin $\alpha = 0,05$ önem düzeyinde anlamlı derecede farklı olduğu sonucuna ulaşılmaktadır. (Kleinbaum ve Klein, 2005, s.60).

Varyans formülünü hesaplamak zorunda olmadan her grup için gözlenen ve beklenen değerleri kullanarak, log-rank istatistiğinin yaklaşık bir değeri hesaplanabilmektedir. Yaklaşık formül, klasik ki-kare yapısındadır (Kleinbaum ve Klein, 2005, s.61).

Yaklaşık formül:

$$\chi^2 \approx \sum_i^G \frac{(O_i - E_i)^2}{E_i}, \quad \text{burada } G = 2 = \text{grup sayısı} \quad (3.18)$$

Log-rank testi ikiden fazla hayatta kalma eğrisini karşılaştırmak için de kullanılabilir. Daha genel bu durum için sıfır hipotezi, tüm hayatta kalma eğrilerinin aynı olduğudur. Uygun matematiksel formül, matris terimleriyle verilebilmektedir (Kleinbaum ve Klein, 2005, s.61).

Eğer karşılaştırılan grup sayısı $G(\geq 2)$ ise, log-rank istatistiği yaklaşık olarak $G - 1$ serbestlik dereceli büyük örneklem ki-kare dağılımına sahiptir. Bundan dolayı anlamlılık hakkındaki karar, uygun serbestlik dereceli ki-kare tablosu kullanılarak verilmektedir (Kleinbaum ve Klein, 2005, s.62).

Log-rank değerini hesaplamak için alternatif veri düzeni Tablo 3.1'teki gibi genişletilmektedir (Kleinbaum ve Klein, 2005, s.59);

Tablo 3.1: Alternatif veri düzeninin genişletilmesi

j	$t_{(j)}$	m_{ij}	n_{ij}	e_{ij}	Gözlenen - Beklenen $m_{ij} - e_{ij}$
1	$t_{(1)}$	m_{i1}	n_{i1}	e_{i1}	$m_{i1} - e_{i1}$
2	$t_{(2)}$	m_{i2}	n_{i2}	e_{i2}	$m_{i2} - e_{i2}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
k	$t_{(k)}$	m_{ik}	n_{ik}	e_{ik}	$m_{ik} - e_{ik}$
Toplam		$\sum_{j=1}^k m_{ij}$			$\sum_{j=1}^k (m_{ij} - e_{ij})$

j zamanında i . grup için; m_{ij} = başarısızlığın gözlenen sıklık sayısı, n_{ij} = risk setindeki sıklık değeri, e_{ij} = başarısızlığın beklenen sıklık sayısıdır. Burada e_{ij} = (risk setindeki oran) \times (tüm gruptaki başarısızlık sayısı), yani ikiden fazla grup için genel formül $e_{ij} = \left(\frac{n_{ij}}{\sum_{i=1}^{grup\ sayısı} n_i} \right) \times \sum_{j=1}^k m_{ij}$ 'dir (Kleinbaum ve Klein, 2005, s.71).

Varyans ya da kovaryans hesaplamaları olmadan sadece gözlenen ve beklenen sıklık değerlerini içeren yaklaşık formül, ikiden fazla karşılaştırılan grup olduğunda da kullanılabilir. Ancak log-rank istatistiğini tam olarak hesaplayan bilgisayar programları olduğundan bu formüle gerek duyulmamaktadır (Kleinbaum ve Klein, 2005, s.62).

Yaklaşık formül:

$$\chi^2 = \sum_i^{grup\ sayısı} \frac{(O_i - E_i)^2}{E_i} \quad (3.19)$$

İkiden fazla grup için log-rank istatistiğinin matris formülü aşağıdaki gibidir (Kleinbaum ve Klein, 2005, s.82);

G = grup sayısı, k = farklı başarısızlık zamanlarının sayısı olmak üzere $i = 1, 2, \dots, G$ ve $j = 1, 2, \dots, k$ için;

n_{ij} = j . sıradaki başarısızlık zamanında i . grupta risk altında olan deneklerin sayısı

m_{ij} = j . sıradaki başarısızlık zamanında i . grupta gözlenen başarısızlık sayısı

e_{ij} = j . sıradaki başarısızlık zamanında i . grupta beklenen başarısızlık sayısı

$$n_j = \sum_{i=1}^G n_{ij} , \quad m_j = \sum_{i=1}^G m_{ij}$$

$$O_i - E_i = \sum_{j=1}^k (m_{ij} - e_{ij}) \quad (3.20)$$

$$\text{Var}(O_i - E_i) = \sum_{j=1}^k \frac{n_{ij}(n_j - n_{ij})m_{ij}(m_j - m_{ij})}{n_j^2(n_j - 1)} \quad (3.21)$$

$$\text{Cov}(O_i - E_i, O_l - E_l) = \sum_{j=1}^k \frac{-n_{ij}n_{lj}m_j(n_j - m_j)}{n_j^2(n_j - 1)} \quad (3.22)$$

$$\mathbf{d} = (O_1 - E_1, O_2 - E_2, \dots, O_{G-1} - E_{G-1})' \quad (3.23)$$

$$\mathbf{V} = v_{il} \quad (3.24)$$

Burada $i = 1, 2, \dots, G - 1$, $l = 1, 2, \dots, G - 1$ olmak üzere $v_{ii} = \text{Var}(O_i - E_i)$ ve $v_{il} = \text{Cov}(O_i - E_i, O_l - E_l)$ 'dir (Kleinbaum ve Klein, 2005, s.82).

Tüm grupların ortak bir hayatta kalma eğrisine sahip olduğu sıfır hipotezi altında $G - 1$ serbestlik dereceli ki-kare dağılımına yaklaşan log-rank istatistiği, aşağıdaki çarpım matrisi tarafından verilmektedir (Kleinbaum ve Klein, 2005, s.82);

$$\text{Log-rank istatistiği} = \mathbf{d}'\mathbf{V}^{-1}\mathbf{d} \quad (3.25)$$

Log-rank istatistiği ile ki-kare tablo değeri karşılaştırılarak gruplar arasında anlamlı bir fark olup olmadığına karar verilmektedir. Test istatistiği tablo değerinden büyük ya da p değeri 0,05'ten küçük olursa sıfır hipotezi reddedilmektedir ve karşılaştırılan gruplar için KM hayatta kalma eğrilerinin $\alpha = 0,05$ önem düzeyinde anlamlı derecede farklı olduğu sonucuna ulaşılmaktadır. (Kleinbaum ve Klein, 2005, s.63).

3.2.3. Log-Rank Testinin Alternatifleri

İki ya da daha fazla hayatta kalma eğrisinin eşit olması hipotezini test etmek için log-rank testine alternatif olarak **Wilcoxon**, **Tarone-Ware**, **Peto** ve **Fleminton-Harrington** gibi birçok test bulunmaktadır. Bu testlerin hepsi log-rank testinin bir çeşididir (Kleinbaum ve Klein, 2005, s.63).

Log-rank testinde her gruptaki $O - E$ başarısızlık sayıları toplandığında, bu basit toplam tüm başarısızlık zamanları için aynı ağırlığı göstermektedir. Wilcoxon, Tarone-Ware, Peto ve Fleminton-Harrington test istatistikleri; log-rank test istatistiğinin varyasyonlarıdır ve j . başarısızlık zamanında farklı ağırlıklar kullanarak türetilmektedirler. İki grup için test istatistiğini ağırlıklandırma aşağıdaki gibi yapılmaktadır (Kleinbaum ve Klein, 2005, s.63);

Test istatistiği:

$$\frac{(\sum_j w(t_j)(m_{ij}-e_{ij}))^2}{\text{var}(\sum_j w(t_j)(m_{ij}-e_{ij}))} \quad (3.26)$$

$i = 1,2 ; j = j.$ başarısızlık zamanı ; $w(t_j) = j.$ başarısızlık zamanındaki ağırlık

Wilcoxon Testi

Breslow testi olarak da adlandırılan Wilcoxon testi; t_j zamanında tüm gruptaki $O - E$ skorlarını n_j risk seti sayısı ile ağırlıklandırmaktadır. Risk seti sayısı, baştaki başarısızlıkların daha sonraki başarısızlıklardan daha fazla ağırlık almasına yol açmaktadır. Bu yüzden Wilcoxon testi, hayatta kalma eğrisinin başlangıçtaki bilgisine daha fazla önem vermektedir. Bu ağırlıklandırma çeşidi, bir deneyin hayatta kalma üzerindeki etkisinin uygulama safhasının başında daha güçlü olup olmadığını ve zamanla etkisinin azalma eğiliminde olup olmadığını değerlendirmek için kullanılabilir (Kleinbaum ve Klein, 2005, s.64).

Tarone-Ware Testi

Tarone-Ware test istatistiği de başlangıçtaki başarısızlık zamanlarına daha fazla ağırlık vermektedir. Bu test, risk seti sayısının karekökü $\sqrt{n_j}$ tarafından t_j zamanındaki $O - E$ skorunu ağırlıklandırarak uygulanmaktadır (Kleinbaum ve Klein, 2005, s.64).

Peto Testi

Peto testi, tüm grupların toplamları üzerine hesaplanan hayatta kalma tahmini $\tilde{s}(t_j)$ tarafından $j.$ başarısızlık zamanını ağırlıklandırmaktadır. Bu hayatta kalma tahmini $\tilde{s}(t_j)$ Kaplan-Meier hayatta kalma tahminine benzemektedir, fakat bu tahmine tam olarak eşit değildir (Kleinbaum ve Klein, 2005, s.64).

Fleminton-Harrington Testi

Fleminton-Harrington testi, $j.$ başarısızlık zamanına ait ağırlıkları hesaplamak için tüm gruplar üzerindeki Kaplan-Meier hayatta kalma tahmini $\hat{s}(t)$ 'yi kullanmaktadır. Ağırlıklar, $\hat{s}(t_{j-1})^p [1 - \hat{s}(t_{j-1})]^q$ formülü ile hesaplanmaktadır (Kleinbaum ve Klein, 2005, s.64).

Bu test, ağırlıkların seçimi bakımından en çok esnekliği tanımaktadır çünkü p ve q değerleri yöntemi kullanan kişi tarafından belirlenmektedir. Aşağıda p ve q 'nin alabileceği değerler için üç farklı örnek gösterilmektedir. Birinci durumda; $w(t) = \hat{s}(t_{j-1})$, 1'e yakın olduğunda başlangıçtaki hayatta kalma sürelerine daha fazla ağırlık verilmektedir. Ancak sonlardaki hayatta kalma sürelerinin daha fazla ağırlık alması durumunda, ikinci örnekte

gösterilen $w(t) = 1 - \hat{s}(t_{j-1})$ uygulanmaktadır. Eğer üçüncü örnekteki gibi $w(t) = 1$ olursa Fleminton-Harrington testi, log-rank testine indirgenecektir (Kleinbaum ve Klein, 2005, s.64);

$$w(t) = \hat{s}(t_{j-1})^p [1 - \hat{s}(t_{j-1})]^q \quad (3.27)$$

Eğer $p = 1$ ve $q = 0$ ise $w(t) = \hat{s}(t_{j-1})$

Eğer $p = 0$ ve $q = 1$ ise $w(t) = 1 - \hat{s}(t_{j-1})$

Eğer $p = 0$ ve $q = 0$ ise $w(t) = 1$ (log-rank testi)

Bu test istatistiklerinin her biri için ağırlıklar, Tablo 3.2’de özetlenmektedir (Kleinbaum ve Klein, 2005, s.64);

Tablo 3.2: Çeşitli test istatistikleri için kullanılan ağırlıklar

Test İstatistiği	$w(t_j)$
Log-rank	1
Wilcoxon	n_j
Tarone-Ware	$\sqrt{n_j}$
Peto	$\tilde{s}(t_j)$
Fleminton-Harrington	$\hat{s}(t_{j-1})^p [1 - \hat{s}(t_{j-1})]^q$

Kullanılacak Test İstatistiğinin Seçimi

Genel olarak çeşitli ağırlıklar, benzer sonuçları vermelidir ve sıfır hipotezini reddedip reddetmeme ile ilgili aynı karara yol açmalıdır. Ağırlıklandırmada hangi test istatistiğinin (örneğin, log-rank ya da Wilcoxon) kullanılacağı, hangi testin en büyük istatistiksel gücü sağlayacağını düşünülmesine bağlıdır. İstatistiksel güç ise testlerde sıfır hipotezinin ihlal edildiğine ne kadar inanıldığına bağlıdır (Kleinbaum ve Klein, 2005, s.66).

Eğer bir olayın etkisinin, hayatta kalma fonksiyonunun başlangıcına (ya da sonuna) doğru daha belirgin olduğuna karar vermek için teorik bir sebep varsa, belirli bir ağırlıklandırma değeri seçmek makuldür. Ancak istenilen p değerinin aranmasından ziyade, **hangi istatistiksel testin kullanılacağına dair öncül (a priori, kuramsal) bir karar verilmelidir.** İstenilen sonucun aranması sapmaya yol açabilmektedir (Kleinbaum ve Klein, 2005, s.66).

3.3. Cox PH Modelini Kullanarak Hayatta Kalma Eğrilerinin Ayarlanması

Bir hayatta kalma analizi açısından istenilen iki temel nicelik, tehlike orantısı tahminleri ve hayatta kalma eğrisi tahminleridir (Kleinbaum ve Klein, 2005, s.103).

Eğer yaşam verisine uyan bir model kullanılmıyorsa, Kaplan-Meier metodunu kullanarak bir hayatta kalma eğrisi tahmin edilebilmektedir. Böyle KM eğrileri adım fonksiyonu olarak çizilmektedir (Kleinbaum ve Klein, 2005, s.104).

Yaşam verisine uyan bir Cox model kullanıldığında, kestirici olarak kullanılan açıklayıcı değişkenler için ayarlanan hayatta kalma eğrileri elde edilebilmektedir. Bu eğriler, ayarlanmış hayatta kalma eğrileri olarak adlandırılmakta ve KM eğrileri gibi adım fonksiyonu olarak çizilmektedir. Yani Cox model, ayarlanmış hayatta kalma eğrilerini vermektedir, ayrıca adım fonksiyonudur (Kleinbaum ve Klein, 2005, s.104).

Cox PH modeli için tehlike fonksiyonu formülü, ilgili hayatta kalma fonksiyonu formülüne dönüştürülebilmektedir. Bu hayatta kalma fonksiyonu formülü, ayarlanmış hayatta kalma eğrilerini belirlemede temeldir. Kestirici olarak \mathbf{X} vektörüne sahip bir denek için t zamanında hayatta kalma fonksiyonunun formülü aşağıdaki gibidir (Kleinbaum ve Klein, 2005, s.104);

Cox model tehlike fonksiyonu:

$$h(t, \mathbf{X}) = h_0(t) e^{\sum_{i=1}^p \beta_i X_i} \quad (3.28)$$

Cox model hayatta kalma fonksiyonu:

$$S(t, \mathbf{X}) = [S_0(t)] e^{-\sum_{i=1}^p \beta_i X_i} \quad (3.29)$$

Tahmin edilmiş tehlike fonksiyonu:

$$\hat{S}(t, \mathbf{X}) = [\hat{S}_0(t)] e^{-\sum_{i=1}^p \hat{\beta}_i X_i} \quad (3.30)$$

Tipik olarak, ayarlanmış hayatta kalma eğrileri hesaplanırken bir değişkeni ayarlamak için seçilen değer, aritmetik ortalama ya da medyan ($X = \bar{X}$ ya da X_{medyan}) gibi bir ortalama değerdir. Cox model için çoğu bilgisayar programı ise ayarlanan her bir değişken için, otomatik olarak tüm deneklerin ortalama değerini kullanmaktadır (Kleinbaum ve Klein, 2005, s.105).

Genel olarak, bir açığa çıkma değişkeninin iki düzeyi için hayatta kalma eğrileri karşılaştırılmak ve birkaç değişken için ayarlanmak isteniyorsa, her bir eğri için aşağıda

gösterilen formül yazılabilmektedir. Burada açığa çıkma değişkeninin katsayı tahmini $\hat{\beta}_1$ olan X_1 değişkeni olduğu ve X_1 'in değerinin, olayın gerçekleştiği denekler için 1, gerçekleşmediği denekler için 0 olduğu varsayılmaktadır. İki grubu karşılaştıran ayarlanmış hayatta kalma eğrileri için genel formül (Kleinbaum ve Klein, 2005, s.106);

Olayın meydana geldiği denekler;

$$\hat{S}(t, \mathbf{X}_1) = [\hat{S}_0(t)]^{\exp[\hat{\beta}_1(1) + \sum_{i \neq 1} \hat{\beta}_i \bar{X}_i]} \quad (3.31)$$

Olayın meydana gelmediği denekler;

$$\hat{S}(t, \mathbf{X}_0) = [\hat{S}_0(t)]^{\exp[\hat{\beta}_1(0) + \sum_{i \neq 1} \hat{\beta}_i \bar{X}_i]} \quad (3.32)$$

Ayrıca modeldeki tüm değişkenleri ayarlayan ayarlanmış hayatta kalma eğrisini elde etmek isteniyorsa, her bir değişken için ortalama değeri kullanan aşağıdaki genel formül kullanılmaktadır. Bu formül her açığa çıkma grubu için farklı eğriler yerine, ayarlanmış tek bir hayatta kalma eğrisi verecektir (Kleinbaum ve Klein, 2005, s.106);

$$\hat{S}(t, \bar{\mathbf{X}}) = [\hat{S}_0(t)]^{\exp[\sum \hat{\beta}_i \bar{X}_i]} \quad (3.33)$$

Hayatta kalma eğrisi için bu ifadede, tanımlanan herhangi bir t değeri için bir hayatta kalma olasılığı hesaplanabilmektedir. Bir bilgisayar paket kullanarak bu hayatta kalma eğrisinin grafiği çizildiğinde, seçilen t değerleri çalışmada olayı geçiren tüm kişilerin başarısızlık zamanlarıdır. Bu işlem, kullanıcının her bir başarısızlık zamanını belirtmesine gerek kalmadan bilgisayar tarafından otomatik olarak yapılmaktadır (Kleinbaum ve Klein, 2005, s.106).

DÖRDÜNCÜ BÖLÜM

ZAMANA BAĞLI DEĞİŞKENLER VE COX ORANTILI TEHLİKE REGRESYON MODELİNİN GENİŞLETİLMESİ

Bu bölüm Cox orantılı tehlike modelinin, zamana bağlı değişkenleri kestirici olarak kullanmak için nasıl genişletilebileceğini tanımlamaktadır. Burada, modelin formu ve karakteristikleri, tehlike orantısı için kullanılacak formüle ve tehlike orantısının yorumlanmasına odaklanılacaktır. Son olarak, genişletilmiş Cox olabilirlik oranı ve Cox oranlı tehlike olabilirlik fonksiyonu ile nasıl zıt olduğu tanımlanacaktır (Kleinbaum ve Klein, 2005, s.214).

Şu ana kadar, tüm değişken değerlerinin her denek üzerinde takibe başlanılan noktada (sıfır zamanında) belirlendiği ve bu değerlerin gözlem periyodu içinde değişmediği varsayılmıştır. Takip edilen periyot boyunca bir ya da daha fazla değişkenin ölçüldüğü ve bunların değerlerinin değiştiği durumlar olabilmektedir. Bu durumlarda olay için tehlike orantısı değerinin, sıfır zamanındaki değerlerinden daha çok bu değişkenlerin mevcut değerlerine dayandığı örnekler olabilmektedir (Hosmer, Lemeshow ve May, 2008, s.213).

4.1. Cox Orantılı Tehlike Regresyon Modeli

Cox PH modelinin genel gösterimi şu şekildedir (Kleinbaum ve Klein, 2005, s.214);

$$h(t, \mathbf{X}) = h_0(t) \exp\left[\sum_{i=1}^p \beta_i X_i\right] \quad (4.1)$$

$\mathbf{X} = (X_1, X_2, \dots, X_p)$ Açıklayıcı / kestirici değişkenler

Bu model, \mathbf{X} ile gösterilen açıklayıcı değişkenlerin verilen bir seti ile bir birey için t zamanındaki tehlikenin bir ifadesini vermektedir (Kleinbaum ve Klein, 2005, s.214). Yani, t zamanındaki bir gözlem için “başarısızlığın” gerçekleşme ihtimalini göstermektedir.

Modelde t , bir gözlemin yaşam süresini; β regresyon katsayıları vektörünü; \mathbf{X} ise bir gözlemin tehlike orantısını öngörmek için modellenen açıklayıcı değişkenlerin bir yığını temsil etmektedir (Kleinbaum ve Klein, 2005, s.214).

Tehlike fonksiyonu; $h_0(t)$ ve $\exp\left[\sum_{i=1}^p \beta_i X_i\right]$ değerlerinin çarpımından oluşmaktadır. İlk çarpan $h_0(t)$, temel tehlike fonksiyonu olarak adlandırılmaktadır. Temel tehlike fonksiyonu t 'yi içerip X 'leri içermezken, ikinci çarpan olan üstel ifade için tam tersi bir

durum söz konusudur. Yani, X 'ler zamandan bağımsızdır (Kleinbaum ve Klein, 2005, s.214). Bu durum Tablo 2.1'de özetlenmektedir.

Birçok durumda açıklayıcı değişken verisi uzun dönemde elde edilmektedir. Örneğin kan basıncı, CD4 sayımı, göreceli ağırlık ve hastalık öyküsü seçilen periyodik zaman noktalarında elde edilebilmektedir. Tedavi ya da diğer etkiler zaman içinde değişebilmektedir. Uygun istatistiksel modelde zamanla değişen açıklayıcı değişken bilgisini kullanmak gerekmektedir. Bunu yapmanın yollarından biri de zamana bağlı açıklayıcı değişkenli Cox regresyon modelini kullanmaktır (Sertkaya, Ata ve Sözer, 2005).

Cox orantılı tehlike modelindeki zamana bağlı etkilerin, X_j değişkeni ile önceden belirlenen parametrik bir $f(t)$ fonksiyonu arasındaki etkileşimle gösterilmesidir (Başar, 2006). Zamana bağlı açıklayıcı değişkenler olduğu durumda, Cox regresyon modeli zamandan bağımsız değişkenleri ve zamanın bazı fonksiyonları ile bu değişkenlerin çarpımını içeren bir modele genişlemektedir. X_1, X_2, \dots, X_p zamandan bağımsız değişkenler $X_1(t), X_2(t), \dots, X_p(t)$ zamana bağlı değişkenler olmak üzere açıklayıcı değişkenler (Sertkaya, Ata ve Sözer, 2005);

$$X(t) = (X_1, X_2, \dots, X_p, X_1(t), X_2(t), \dots, X_p(t))$$

biçiminde gösterilmektedir. Buna göre β ve δ açıklayıcı değişkenlerin katsayılar vektörü olmak üzere Cox regresyon modeli (Sertkaya, Ata ve Sözer, 2005);

$$h(t, X(t)) = h_0(t) \exp\left[\sum_{i=1}^{p_1} \beta_i X_i + \sum_{j=1}^{p_2} \delta_j X_j g(t)\right] \quad (4.2)$$

biçiminde yazılmaktadır. Burada $g(t)$ zamanın bir fonksiyonu olarak tanımlanmaktadır. $g(t)$ 'nin seçimi kullanılan değişkenlerin durumuna göre ve araştırmacının bilgisine göre değişkenlik göstermektedir. Bu fonksiyon genellikle t , $\log(t)$, $\ln(t)$ ya da adım fonksiyonları (step function) biçiminde tanımlanmaktadır (Sertkaya, Ata ve Sözer, 2005).

Genişletilmiş Cox model, orantılı tehlike varsayımını sağlamamaktadır. Cox PH modelinden tehlike orantısı formülü elde edilebilmektedir (Kleinbaum ve Klein, 2005, s.215):

$$\widehat{HR} = \exp\left[\sum_{i=1}^p \hat{\beta}_i (X_i^* - X_i)\right]$$

Burada $X^* = (X_1^*, X_2^*, \dots, X_p^*)$ ve $X = (X_1, X_2, \dots, X_p)$ olmak üzere X 'in iki özelliğini göstermektedir.

PH varsayımı:

$$\frac{\widehat{h}(t, X^*)}{\widehat{h}(t, X)} = \hat{\theta} \quad (t \text{ 'ye göre bir sabit}) \quad (4.3)$$

$$\text{Yani, } \hat{h}(t, \mathbf{X}^*) = \hat{\theta} \hat{h}(t, \mathbf{X}) \quad (4.4)$$

\mathbf{X} açıklayıcı değişkenlerinin herhangi iki özelliğini karşılaştıran Cox PH modeline dayanan orantılı tehlike varsayımı, zamana göre sabittir. Bu demektir ki, bir bireyin tehlikesi diğer herhangi bir bireyin tehlikesine orantılıdır (Kleinbaum ve Klein, 2005, s.215).

İki ya da daha fazla grubun zaman göre grafiği çizildiğinde tehlikeler kesişiyorsa, orantılı tehlike varsayımı sağlanmamaktadır. Ancak tehlike fonksiyonları kesişmese bile orantılı tehlike varsayımının sağlanmaması mümkündür (Kleinbaum ve Klein, 2005, s.215).

PH varsayımını değerlendirmek için üç genel yaklaşım vardır (Kleinbaum ve Klein, 2005, s.215);

- Grafik yaklaşımı
- Genişletilmiş Cox modelde zamana bağlı değişkenlerin kullanılması
- Uyum iyiliği testi uygulanması

Değişken-zaman etkileşiminin parametrik olarak modellenmesi iki alternatif arasında seçim yapılmasını gerektirmektedir; ilk alternatif, analizlerin önceden seçilen tek bir parametrik fonksiyonla sınırlanmasıdır. Bu durumda her bir değişkenin etkisi “homojen”dir. Tehlike orantısındaki değişimin gerçek yapısı, seçilen fonksiyonun biçimi ile tutarlı değilse analizin tek bir fonksiyonla sınırlanması yanlış tahminlere neden olabilmektedir. İkinci alternatif, farklı parametrik fonksiyonların kullanılması ile tahminlerin elde edilmesi ve sonradan en iyi uyan modelin oluşturulmasıdır. Model sonradan seçilen farklı fonksiyonlardan oluştuğu için her bir değişkenin etkisi “heterojendir”. Bu durumda özellikle çok değişkenli analizlerde beklenenden yüksek tahminler elde edilebilmekte ayrıca istatistiksel olarak sonuç çıkarımı da güçleşmektedir (Başar, 2006).

Zamandan bağımsız bir değişkenin orantılı tehlike varsayımını değerlendirmek için Cox model, zamandan bağımsız değişken olarak değerlendirilen ve zamanın bazı fonksiyonlarından oluşan çarpım (yani etkileşim) terimiyle genişletilmektedir (Kleinbaum ve Klein, 2005, s.215).

Örneğin, orantılı tehlike varsayımı cinsiyet değişkeni GEN için değerlendirilirse, Cox model GEN değişkenine ek olarak $(GEN \times t)$ değişkenini içerecek şekilde genişletilebilmektedir. Çarpım teriminin katsayısı anlamsız çıkarsa, $(GEN \times t)$ değişkeni zamana bağlı değişkenin uygun bir çeşidi olması koşuluyla, orantılı tehlike varsayımının GEN değişkeni için sağlandığı sonucuna ulaşılmaktadır. Model şu şekilde kurulmaktadır (Kleinbaum ve Klein, 2005, s.216);

$$h(t, \mathbf{X}) = h_0(t) \exp[\beta_1 GEN + \beta_2 (GEN \times t)]$$

$H_0: \beta_2 = 0$ ise orantılı tehlike varsayımı sağlanmaktadır.

Modelde bir veya daha fazla kestirici için orantılı tehlike varsayımı sağlanmazsa, ele alınacak iki seçenek vardır. Birisi, orantılı tehlike varsayımını sağlayan değişkenleri modelde tutarken tehlike varsayımını sağlamayan değişken(ler)i tabakalandıran bir tabakalı Cox model kullanmaktır. Bir diğer seçenek ise zamana bağlı değişkenleri kullanmaktır (Kleinbaum ve Klein, 2005, s.216).

Kestiriciler, gerçekten zamana bağlı olabilirler ya da buna ek olarak orantılı tehlike varsayımını sağlamayan zamandan bağımsız bir değişkeni analiz etmek için zamana bağlı değişken olarak ele alınabilirler. Burada özünde zamana bağlı olan değişkenler ele alınacaktır (Kleinbaum ve Klein, 2005, s.216).

4.2. Zamana Bağlı Değişkenlerin Tanımı

Zamana bağlı bir değişken, belirli bir denek için değeri zaman (t) üzerinde farklı olabilen bir değişken olarak tanımlanmaktadır. Tam tersi olarak, zamandan bağımsız bir değişken ise, belirli bir denek için değeri zamana göre sabit kalan bir değişkendir (Kleinbaum ve Klein, 2005, s.216).

Denklem (2.5)'te verilen modelde temel tehlike fonksiyonu zamanın bir fonksiyonunu içerirken, üstel kısımda yer alan değişkenler zamanın bir fonksiyonu değildir, yani zamandan bağımsızdır. Modelde zamanı içeren değişkenler de olabilmektedir. Bu değişkenler zamana bağlı açıklayıcı değişkenler olarak adlandırılmaktadır. Zamana bağlı değişken, incelenen birey için değeri zamanla değişebilen herhangi bir değişken olarak tanımlanmaktadır (Neffke vd, 2011). Zamandan bağımsız değişken ise incelenen birey için değeri zaman içinde sabit kalan bir değişkendir. En yaygın tanımlanan zamana bağlı değişken, zamandan bağımsız bir değişken ile zamanın ya da zamanın bir fonksiyonunun çarpımı biçimindedir (Pettitt, 1990; Therneau, 2000).

Basit bir örnek olarak, ($IRK \times zaman$) değişkeni zamana bağlı bir değişken olmasına karşın, IRK değişkeni zamandan bağımsız bir değişkendir (Kleinbaum ve Klein, 2005, s.216).

Tablo 4.1: Zamana bağılı ve zamandan bağımsız değişkenler

Zamana bağılı	Zamandan bağımsız
Değişkenin değeri zamana göre değişir. ($IRK \times t$)	Değişkenin değeri zamana göre sabittir. IRK

Zamanla değişen (ya da zamana bağılı) değişkenler genellikle *içsel* ya da *dışsal* olarak sınıflandırılmaktadır. Zamana bağılı bir içsel değişkenin değeri deneğe özgü bir özelliktir ve deneğin periyodik gözlem altında olması gerekmektedir. Örneğin, bitiş noktası kanserden ölüm olan yeni bir kanser tedavisi için klinik bir deney ele alınsın. Başlangıçta ölçülen bir değişkene sahip olduğu fakat değerinin zamana göre değişebildiği varsayılınsın. Tehlike orantısının başlangıç değerinden daha çok mevcut değerine bağılı olduğu durumlar olabilmektedir. Takip eden periyot boyunca bu değişkenin değerinin ölçülebilmesi için çalışmada ilgilenilen denekler doğrudan gözlem altında olmalıdır (Hosmer, Lemeshow ve May, 2008, s.214).

Buna karşın, zamana bağılı dışsal bir değişkenin belirli bir zamandaki değerinin ilgilenilen denek için direk gözlem altında olması gerekmemektedir. Tipik olarak, bu değişkenler gözlem altında tüm deneklere uygulanan çalışma ya da çevresel faktörlerdir. Örneğin, belirtilerden kurtulana kadarki saatlerin sayısının çıktı kayıtları olarak alındığı bir zamana bağılı değişkeni içeren saman nezlesinin belirtilerinden kurtulmak için yeni bir tedavi çalışması ele alınsın. Bu çalışmada, dışsal değişkene bir örnek saat başına ortalama polen sayısıdır. Deneğe bağılı bir dışsal değişken deneğin yaşıdır. Yeterince uzun bir zaman periyodu için denekler takip edilirse onların şimdiki yaşı, çalışma başladığı zamandaki yaşlarından daha çok hayatta kalma üzerinde bir etkiye sahip olabilmektedir. Ancak, bir deneğin doğum tarihi öğrenildikten sonra deneğin hala gözlem altında olup olmadığına bakılmaksızın zamanın herhangi bir noktasında yaşı hesaplanabilmektedir. Diğer bir önemli zamana bağılı dışsal değişken zamanın kendisidir (Hosmer, Lemeshow ve May, 2008, s.214).

Modelde zamana bağılı değişkenleri içeren gösterimin genelleştirilmesi gerekmektedir. t zamanında ölçülen X değişkeninin değeri $X(t)$ ile gösterilmektedir. Gerçek zaman ya da takvim zamanı (calendar time) kullanılmasına karşın sıfır zamanından başlamak için her deneğin takip süresinin düzeltildiği varsayılınsın. Böylece, aynı t farklı takvim zamanlarında meydana gelebilmesine rağmen, t 'nin $X(t)$ 'nin değerini belirlediği varsayılmaktadır.

Örneğin, hastanede kalma uzunluğu gibi zamana bağlı bir değişken tanımlandığı farz edilsin. Denek 1, 1 Mayıs'ta bir hastaneye alınabilir ve hastanede 60 gün kalabilir. Denek 2, 1 Temmuz'da bir hastaneye alınabilir ve hastanede 60 gün kalabilir. Deneklerin farklı zamanlarda hastaneye alınması değil, her iki deneğin de 60 gün hastanede olması önemlidir. Kısmi olabilirlikte ve onların çok değişkenli denklemlerinde böyle bir değişken eklemek için, belirli bir sürenin yanı sıra denek için hesap yapmak gerekmektedir. t_i zamanındaki j deneği için değişkenin değeri $X_j(t_i)$ ile gösterilsin. Çok değişkeni hesaba katmak için t_i zamanındaki j deneği için k . değişkenin değeri $x_{jk}(t_i)$, $k = 1, 2, \dots, p$ olarak gösterilir ve değişkenlerin vektörü aşağıdaki gibidir (Hosmer, Lemeshow ve May, 2008, s.214-215);

$$X'_j(t_i) = [x_{j1}(t_i), x_{j2}(t_i), \dots, x_{jp}(t_i)] \quad (4.5)$$

(4.4) denklemindeki gösterim, belirli bir x_k değişkeninin sabit (yani zamana göre değişmeyen) olması durumunda tamamen geneldir (Hosmer, Lemeshow ve May, 2008, s.215);

$$x_{jk}(t_i) = x_{jk}(t = 0) = x_{jk} \quad (4.6)$$

Zamana bağlı açıklayıcı değişkenin üç türü bulunmaktadır. Bunlar; tanımlanmış (defined) değişken, içsel (internal) değişken ve yardımcı (ancillary) değişkendir (Sertkaya, Ata ve Sözer, 2005);

i. Tanımlanmış Değişken

Bu değişken, sabit bir açıklayıcı değişken ile zamanın bir fonksiyonunun çarpımı şeklinde elde edilmektedir (Sertkaya, Ata ve Sözer, 2005). ($IRK \times zaman$), 'tanımlanmış' zamana bağlı değişken olarak adlandırılan değişkenlere bir örnektir. Örnekteki değişken, zamandan bağımsız bir değişkenin (örneğin; IRK) zaman ya da zamanın bir fonksiyonu ile çarpılması formundadır. Verilen bir denek için IRK belirlendikten sonra, ($IRK \times zaman$) değişkeninin tüm değerleri, çalışmanın tanımlı zaman aralığı üzerinde tamamen tanımlanır (Kleinbaum ve Klein, 2005, s.217).

Tanımlanmış değişken: $IRK \times t$

Zamandan bağımsız 

$$IRK=1 \Rightarrow IRK \times t = t$$

$$IRK=0 \Rightarrow IRK \times t = 0 \text{ (tüm } t \text{ değerlerinde)}$$

Tanımlanmış değişkenin ikinci örneği $E \times (\log t - 3)$ tarafından verilebilir, burada E , çalışmaya birinin girmesine göre $(0,1)$ olarak belirlenen açığa çıkan durum değişkenini göstermektedir. Zamanı tek başına kullanmaktan ziyade, $(\log t - 3)$ gibi zamanın bir fonksiyonu kullanıldığına dikkat edilmelidir (Kleinbaum ve Klein, 2005, s.217).

Tanımlanmış değişkenin diğer bir örneği, zamanın bir fonksiyonunu içeren $E \times g(t)$ verilebilir, burada $g(t)$; t, t_0 olarak adlandırılan t 'nin tanımlanmış bir değerine eşit ya da daha büyük ise 1 ve t_0 'dan daha küçük ise 0 değerini almaktadır (Kleinbaum ve Klein, 2005, s.217).

$$E \times g(t) \text{ burada } g(t) = \begin{cases} t \geq t_0 \text{ ise } 1 \\ t < t_0 \text{ ise } 0 \end{cases}$$

Koşullu fonksiyon (Heavyside function)

$g(t)$ 'koşullu' fonksiyon olarak adlandırılmaktadır. t, t_0 'a eşit ya da daha büyük olduğunda $g(t)$ 1'e eşit yani $E \times g(t) = E$ olur; ancak t, t_0 'dan daha küçük olduğunda $g(t) = 0$ 'dır ve $E \times g(t)$ 'nin değeri her zaman 0'dır. E gibi zamandan bağımsız bir değişken, orantılı tehlike varsayımını sağlamadığında analiz için bir yöntem olarak koşullu fonksiyonlar kullanılabilir (Kleinbaum ve Klein, 2005, s.217).

$$t \geq t_0: E \times g(t) = E$$

$$t < t_0: E \times g(t) = 0$$

ii. İçsel Değişken

İçsel değişkenler için değişkenin değerindeki değişimin nedeni içsel özelliklere, yani bireyin kendi davranışlarına ve özelliklerine dayanmaktadır. Bu gibi değişkenlere örnek olarak, t zamanındaki sigara kullanma durumu (SMK) ve t zamanındaki obezite seviyesi (OBS) verilebilir. İçsel değişkenler için değişkenin değerindeki değişimin nedeni içsel özelliklere, yani bireyin kendi davranışlarına ve özelliklerine dayanmaktadır (Sertkaya, Ata ve Sözer, 2005).

İçsel değişken örnekleri: $SMK(t), OBS(t)$

Bu örnekler, çalışılan herhangi bir denek için zamana göre değişebilen değerleri olan değişkenleri ele almaktadır. Dahası, içsel değişkenler için değerindeki bir değişikliğin sebebi bireye özgü ‘içsel’ özellik ya da davranışa dayanmaktadır (Kleinbaum ve Klein, 2005, s.217).

iii. Yardımcı Değişken

Değişkenin değerindeki değişimin nedeni dışsal, çevreye ait özelliklerdir. Özel bir coğrafi alan için t zamanındaki hava kirliliği indeksi dışsal değişkene bir örnek olarak verilebilir (Sertkaya, Ata ve Sözer, 2005). Başka bir örnek, birinin istihdam edilip edilmediğinin esas sebebinin kişisel özelliklerinden çok genel ekonomik şartlara dayandığı t zamanındaki istihdam durumu (*EMP*)’dir (Kleinbaum ve Klein, 2005, s.218).

Kısmen içsel, kısmen dışsal olabilecek değişkenler de vardır. Buna örnek olarak ciddi bir kalp rahatsızlığı olan bir hastanın kalp nakli için uygun olup olmaması durumu düşünülebilir. Kalp nakli durumuna ilişkin değişken (*HT*), bireyin ayırt edici özelliklerinden dolayı içsel olarak, tıbbi vericinin elde edilebilirliğinden dolayı yardımcı değişken olarak adlandırılmaktadır (Sertkaya, Ata ve Sözer, 2005). t zamanındaki *HT* değişkeninin değeri, kişi herhangi bir zamanda organ nakli yaptırdıysa; yani t_0 , t ’den önce ise 1’dir. t zamanındaki *HT* değeri, kişi t zamanına kadar bir nakil yapılmadıysa 0’dır (Kleinbaum ve Klein, 2005, s.218).

Şekil 4.1: Kısmen içsel kısmen de yardımcı olabilen değişken için bir örnek

$$\begin{array}{c} \text{Nakil varsa } HT(t) = 0000 \dots 0111111111 \\ \hline t \rightarrow \qquad \qquad \qquad t_0 \end{array}$$

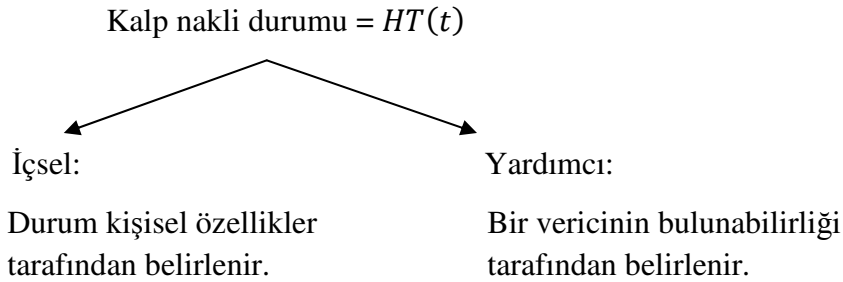
$$\begin{array}{c} \text{Nakil yoksa } HT(t) = 0000 \dots 00000 \\ \hline t \rightarrow \end{array}$$

Bir kişi, t_0 zamanında nakil ameliyatı yaptırdıktan sonra, sonraki zamanların tümü için *HT*’nin değeri 1 olarak kalmaktadır. Bu yüzden, nakil ameliyatı olan bir kişi için, *HT*’nin değeri nakil zamanına kadar 0, nakil zamanından sonra 1 olarak kalmaktadır. Buna karşın, hiç nakil yaptırmayan bir kişi, çalışmada olduğu periyot boyunca tüm zamanlar için 0’a eşit *HT*’ye sahip olmaktadır (Kleinbaum ve Klein, 2005, s.218).

‘Kalp nakli durumu’ değişkeni *HT(t)*, aslında içsel bir değişken olarak ele alınabilir, çünkü aday nakil alıcısının birey özellikleri, nakil ameliyatını yürütmek için kararın önemli

belirleyicileridir. Bununla beraber, aday alıcı ile doku eşleştirmesi ve diğer karşılaştırmaların yapılmasından önce bir vericinin bulunabilirliği, alıcı için dışsal bir yardımcı özellik olarak ele alınabilir (Kleinbaum ve Klein, 2005, s.218).

Şekil 4.2: Değişkeni kısmen içsel kısmen de yardımcı yapan faktörler



Tanımlanmış, içsel ya da yardımcı değişkenler arasında ayırım yapmanın başlıca sebebi, genişletilmiş Cox modelde kullanmak üzere değişkenleri tanımlamak için gerekli bilgisayar komutlarının, kullanılan bilgisayar programına bağlı olarak farklı değişken tipleri için biraz farklı olmasıdır. Bununla beraber, genişletilmiş Cox modelin formu; değişken tipine ve regresyon katsayılarının ve diğer parametrelerin tahminlerini elde etmek için kullanılan prosedürlere bakılmaksızın aynıdır. Ayrıca ek olarak, istatistiksel yorumların yürütülmesi de aynıdır (Kleinbaum ve Klein, 2005, s.219).

4.3. Zamana Bağlı Değişkenler ve Genişletilmiş Cox Model

Cox regresyon modeli, zamanla değişen açıklayıcı değişkenlerle de kullanılabilir. Uygulamada kullanımları, sabit (zamandan bağımsız) açıklayıcı değişkenlere göre daha karmaşıktır. Ayrıca hatalı çıkarımlara ve modelleme için potansiyel gittikçe artmaktadır (Sertkaya, Ata ve Sözer, 2005).

Orantılı tehlike regresyon modelinin geliştirilmesi ve onun zamana bağlı değişkenleri içeren kısmi olabilirliğini genellemek zor değildir. Ancak, kavramsal bakış açısından model daha fazla karmaşık olmaktadır ve modelde önceden bulunan herhangi bir zamana bağlı değişkenin niteliğine ciddi önem gösterilmesi gerekmektedir. Özellikle analiz zamanı sıfır olduğunda değişkene yakın dikkat gösterilmelidir çünkü analiz zamanının başlangıcı farklı denekler için farklı takvim zamanında olabilmektedir. Ancak herhangi bir zamana bağlı değişkenin değeri takvim zamanına değil sadece çalışma zamanına bağlı olmalıdır. Başka bir endişe, zamana bağlı değişkenler kullanıldığında bir modelin aşırı

uyumlama potansiyelidir. Zamana bağılı değişkenlerin dahil olduğu tüm bu örnekler, güçlü bir klinik kanıtı dayandırılmalıdır (Hosmer, Lemeshow ve May, 2008, s.213).

Denklem (4.5)'teki durum Denklem (4.6)'daki zamana bağılı gösterimi kullanmaya temel oluşturmuştur. Çok değişkenli zamana bağılı değişkenleri imkan dahilinde hesaba katmak için $h(t, \mathbf{X}) = h_0(t)e^{\beta \mathbf{X}}$ (Hosmer, Lemeshow ve May, 2008, s.70) olan orantılı tehlike regresyon modeli fonksiyonunun genelleştirilmiş şu şekildedir (Hosmer, Lemeshow ve May, 2008, s.215);

$$h(t, \mathbf{X}(t)) = h_0(t) \exp[\beta \mathbf{X}(t)] \quad (4.7)$$

(4.7) denklemindeki tehlike fonksiyonu, temel tehlike fonksiyonundan başka şekillerde zamana bağılı olabilmektedir. Zamana bağılı değişkenleri içeren birçok yerde, model ayrıca sabit değişkenleri (örneğin; *cinsiyet*) de kapsayacaktır. t zamanında deneğin X_1 ile gösterilen kan basıncını ve X_2 ile gösterilen cinsiyetini içeren bir modele sahip olduğu düşünölsün. Bu model için tehlike fonksiyonu (4.7) denkleminde şu şekilde olacaktır (Hosmer, Lemeshow ve May, 2008, s.215);

$$h(t, \mathbf{X}(t)) = h_0(t) \exp[\beta_1 X_1(t) + \beta_2 X_2] \quad (4.8)$$

Cinsiyet için tehlike orantısı zamana bağılı değildir (Hosmer, Lemeshow ve May, 2008, s.215);

$$HR(t, X_2 = 1, X_2 = 0 | X_1(t)) = \frac{h_0(t) \exp[\beta_1 X_1(t) + \beta_2]}{h_0(t) \exp[\beta_1 X_1(t)]} = \exp(\beta_2) \quad (4.9)$$

Yani bir bakıma, tehlike bir oransallık özelliğine sahiptir fakat payda zamanla ilişkili olan $h_0(t) \exp[\beta_1 X_1(t)]$ 'dir. Bu ayırım, (4.7)'deki model kullanıldığında "orantılı" teriminden vazgeçmek için bazı yazarlara (örneğin; Collett (2003)) yol göstermiştir. Ancak bu modele orantılı tehlike modeli denilmeye devam edilmelidir. Çünkü temel tehlike fonksiyonundan potansiyel olarak çok daha karışık olmasına rağmen, model zamanın bir fonksiyonu için hala orantılıdır (Hosmer, Lemeshow ve May, 2008, s.215-216).

Modele zamana bağılı bir değişken dahil edildiğinde ciddi bir sapma olabilmektedir ve bir deneyin çıktısı üzerindeki etkisi, bu zamana bağılı değişken tarafından dolaylı olarak etkilenmektedir. Örneğin, *kan basıncı* zamana göre değişir ve deneyin çıktısı üzerinde dolaylı bir etkisi olabilir. Bu durumda, zamana bağılı bir değişken olarak kan basıncının dahil edilmesi, deney etkisinin tahminini yanlış yapabilmektedir. Çünkü var olsa bile bir deney etkisini tanımlamada başarısız olunabilmektedir. Bu olası problem ayrıca zamana bağılı değişkenleri modellemenin kavramsal olarak neden karmaşıklaştırılabileceği göstermektedir (Hosmer, Lemeshow ve May, 2008, s.216).

Hem zamandan bağımsız hem de zamana bağlı kestirici değişkenleri içeren belirli bir hayatta kalma analizi durumu, iki tipi de birleştiren genişletilmiş Cox model tarafından aşağıdaki gibi de yazılabilmektedir (Kleinbaum ve Klein, 2005, s.219);

$$h(t, \mathbf{X}(t)) = h_0(t) \exp\left[\sum_{i=1}^{p_1} \beta_i X_i + \sum_{j=1}^{p_2} \delta_j X_j(t)\right] \quad (4.10)$$

$$\mathbf{X}(t) = \left(\underbrace{X_1, X_2, \dots, X_{p_1}}_{\text{Zamandan bağımsız}}, \underbrace{X_1(t), X_2(t), \dots, X_{p_2}(t)}_{\text{Zamana bağlı}} \right)$$

Cox PH modelindeki gibi, genişletilmiş model de üstel bir fonksiyon ile çarpılan temel tehlike fonksiyonunu $h_0(t)$ içermektedir. Ancak genişletilmiş modelde üstel kısım, hem X_i değişkenleri tarafından gösterilen zamandan bağımsız kestiricileri hem de $X_j(t)$ değişkenleri tarafından gösterilen zamana bağlı kestiricileri kapsamaktadır. t zamanındaki tüm kestiricilerin yığını $\mathbf{X}(t)$ ile gösterilmektedir (Kleinbaum ve Klein, 2005, s.219).

i. Regresyon Katsayılarının Yorumlanması

Basit Cox PH modelinde olduğu gibi, genişletilmiş Cox modeldeki regresyon katsayıları ‘ençok olabilirlik yöntemi’ni kullanarak tahminlenmektedir. Ençok olabilirlik tahminleri, (kısmi) olabilirlik fonksiyonu L 'yi maksimize ederek elde edilmektedir. Ancak genişletilmiş Cox model için hesaplamalar, Cox PH modeli için yapılan hesaplamalardan daha karmaşıktır. Çünkü olabilirlik fonksiyonunu oluşturmak için kullanılan risk setleri zamana bağlı değişkenler ile daha karmaşık olmaktadır (Kleinbaum ve Klein, 2005, s.219).

Paydada, t_i zamanında risk setindeki tüm denekler üzerine toplamının $R(t_i)$ ile gösterildiği $L = \prod_{i=1}^n \left[\frac{e^{x_i \beta}}{\sum_{j \in R(t_i)} e^{x_j \beta}} \right]^{c_i}$ (Hosmer, Lemeshow ve May, 2008, s.75) olan kısmi olabilirlik fonksiyonunun genelleştirilmiş şu şekildedir (Hosmer, Lemeshow ve May, 2008, s.216);

$$L = \prod_{i=1}^n \left[\frac{e^{x'_i(t_i) \beta}}{\sum_{l \in R(t_i)} e^{x'_l(t_i) \beta}} \right]^{c_i} \quad (4.11)$$

Burada c_i , $i = 1, 2, \dots, n$ olan sansürlü değişkendir (censoring variable).

ii. Genişletilmiş Cox Model için Bilgisayar Programları

Sıklıkla uygulamadaki en büyük problem, zamana bağlı değişkenlerin değerlerini yazılıma tanımlamak için veri yönetimi gerekliliğidir. Veri toplanmaya başlanmadan önce,

modelin geliştirilmesi için hangi yazılım paketinin kullanılacağına, veri yönetimindeki uzmanlara danışılmalıdır. Hangi zamana bağlı değişkenin ele alındığı ile kolaylığına göre yazılım paketleri arasında önemli farklılıklar vardır (Hosmer, Lemeshow ve May, 2008, s.216).

Genişletilmiş Cox modele uyan Stata, SAS ve SPSS gibi programlar mevcuttur. Her programın ilgili komutu parantez içindeki gibidir; Stata (Stcox), SAS (PHREG) ve SPSS (COXREG).

iii. İstatistiksel Yorumlar

Esasında istatistiksel yorum yapma metotları, orantılı tehlike model ile aynıdır. Wald ve/veya olabilirlik oranı (LR) testi ve büyük örneklem güven aralığı metotları kullanılabilir (Kleinbaum ve Klein, 2005, s.220).

4.4. Genişletilmiş Cox Model için Tehlike Orantısı Formülü

Bu bölümde genişletilmiş Cox modelden türetilen tehlike orantısı formülü tanımlanmaktadır. Bu formülün en önemli özelliği, genişletilmiş Cox model kullanıldığında orantılı tehlike varsayımının artık sağlanıyor olmasıdır (Kleinbaum ve Klein, 2005, s.221).

Genişletilmiş Cox model için genel tehlike orantısı formülü aşağıda gösterilmektedir (Kleinbaum ve Klein, 2005, s.221);

$$\widehat{HR}(t) = \frac{\widehat{h}(t, X^*(t))}{\widehat{h}(t, X(t))} = \exp\left[\sum_{i=1}^{p_1} \hat{\beta}_i [X_i^* - X_i] + \sum_{j=1}^{p_2} \delta_j [X_j^*(t) - X_j(t)]\right] \quad (4.12)$$

Bu formül belirli bir t zamanındaki tehlikelerin oranını tanımlamaktadır ve t zamanındaki iki kestirici setinin tanımını gerektirmektedir. Bu iki set $X^*(t)$ ve $X(t)$ olarak gösterilmektedir (Kleinbaum ve Klein, 2005, s.221).

$X^*(t)$ ve $X(t)$ kestiricilerinin iki seti, hem zamandan bağımsız hem de zamana bağlı değişkenleri içeren kestiricilerin seti için t zamanındaki iki özelliği tanımlamaktadır. Kestiricilerin her bir seti için ayrı ayrı bileşenler aşağıda gösterilmektedir (Kleinbaum ve Klein, 2005, s.221);

Kestiricilerin iki seti:

$$X^*(t) = \left(X_1^*, X_2^*, \dots, X_{p_1}^*, X_1^*(t), X_2^*(t), \dots, X_{p_2}^*(t) \right)$$

$$\mathbf{X}(t) = (X_1, X_2, \dots, X_{p_1}, X_1(t), X_2(t), \dots, X_{p_2}(t))$$

Basit bir örnek olarak modelin; açığa çıkma durumu E olarak adlandırılan bir (0,1) değişkeni olan, zamandan bağımsız sadece bir kestirici içerdiği ve $E \times t$ olarak adlandırılan bir zamana bağlı kestirici içerdiği varsayalım. $E = 1$ olduğu olayın gerçekleştiği kişilerle $E = 0$ olduğu olayın gerçekleşmediği kişileri t zamanında karşılaştırmak için; kestiricilerin $\mathbf{X}^*(t)$ seti, $E = 1$ ve $E \times t = t$ olan iki bileşene sahiptir, $\mathbf{X}(t)$ seti ise $E = 0$ ve $E \times t = 0$ olan iki bileşene sahiptir (Kleinbaum ve Klein, 2005, s.222).

$$h(t, \mathbf{X}(t)) = h_0(t) \exp[\beta E + \delta(E \times t)] \quad (4.13)$$

$$E = \begin{cases} 1 & \text{olay meydana geliyorsa} \\ 0 & \text{olay meydana gelmiyorsa} \end{cases}$$

$$\mathbf{X}^*(t) = (E = 1, E \times t = t)$$

$$\mathbf{X}(t) = (E = 0, E \times t = 0)$$

t zamanında olayın meydana geldiği kişilerle olayın meydana gelmediği kişileri karşılaştıran tahminlenmiş tehlike orantısı hesaplanırsa, aşağıdaki formül elde edilecektir (Kleinbaum ve Klein, 2005, s.222);

$$\widehat{HR}(t) = \frac{\hat{h}(t, E = 1)}{\hat{h}(t, E = 0)} = \exp[\hat{\beta}(1 - 0) + \hat{\delta}((1 \times t) - (0 \times t))] = \exp[\hat{\beta} + \hat{\delta}t] \quad (4.14)$$

Bu formül, tehlike orantısının zamanın bir fonksiyonu olduğunu, özellikle $\hat{\delta}_j$ pozitif ise zaman arttıkça tehlike orantısının artacağını söylemektedir. Yani bu örnekteki model için orantılı tehlike varsayımı sağlanmamaktadır, bu yüzden tehlike orantısı kesinlikle sabit değildir (Kleinbaum ve Klein, 2005, s.222).

$$\hat{\delta} > 0 \Rightarrow t \uparrow \text{ iken } \widehat{HR}(t) \uparrow$$

Orantılı tehlike varsayımı sağlanmamaktadır.

Daha genel olarak, t zamanındaki zamana bağlı değişkenlerin değerleri farklılıklar içerdiğinden genel tehlike orantısı formülü zamanın bir fonksiyonudur. Bu yüzden genelde, herhangi bir δ_j sifıra eşit olmazsa genişletilmiş Cox model orantılı tehlike varsayımını sağlamayacaktır (Kleinbaum ve Klein, 2005, s.222).

Tehlike orantısı formülünde, j . zamana bağlı değişkenin değerleri arasındaki farkın katsayısı $\hat{\delta}_j$ 'nin kendisi zamana bağlı değildir. Bu yüzden bu katsayı, zamana bağlı ilgili değişkenin çalışmada ölçüldüğü tüm zamanları hesaba katan 'toplam' etkisini temsil etmektedir (Kleinbaum ve Klein, 2005, s.222).

$$\widehat{HR}(t) = \exp\left[\sum_{i=1}^{p_1} \hat{\beta}_i [X_i^* - X_i] + \sum_{j=1}^{p_2} \boxed{\hat{\delta}_j} \boxed{[X_j^*(t) - X_j(t)]}\right] \quad (4.15)$$

↑
Zamanın bir fonksiyonu

- Genelde, genişletilmiş model için orantılı tehlike varsayımı sağlanmamaktadır.
- $\hat{\delta}_j$, zamana bağlı değildir.
- $\hat{\delta}_j$, $X_j(t)$ üzerindeki 'toplam' etkiyi temsil etmektedir (Kleinbaum ve Klein, 2005, s.222).

Tehlike orantısı formülünü açıklamak için başka bir örnek olarak, t zamanında kimyasal maddelere maruz kalma durumunu haftalık olarak ölçen sadece bir değişken içeren genişletilmiş bir Cox model ele alınsın. Bu değişkenin; $E(t)$ olarak gösterildiği, sırayla verilen haftalık ölçümlerde bir kişinin kimyasal maddeye maruz kalıp kalmadığına bağlı olarak 0 ya da 1 değerlerinden birini alabildiği varsayılınsın (Kleinbaum ve Klein, 2005, s.223).

Tanımlandığı gibi, $E(t)$ farklı denekler için farklı değer kalıpları alabilmektedir. Örneğin, beş haftalık bir periyot için Denek A'nın değerleri 01011, buna karşın Denek B'nin değerleri 11010 olabilir (Kleinbaum ve Klein, 2005, s.223).

Şekil 4.3: Genişletilmiş Cox model için bir örnek

Denek A:	$E(t)$	0 1 0 1 1
	t	1 2 3 4 5...
Denek B:	$E(t)$	1 1 0 1 0
	t	1 2 3 4 5...

Bu örnekte, çalışma boyunca bir grubun her zaman olaya maruz kaldığı, diğer grubun hiçbir zaman olaya maruz kalmadığı iki ayrı denek grubu ele alınmamaktadır. Grubun hiçbir zaman olaya maruz kalmadığı durum, olayın meydana gelmesi için zamandan bağımsız bir (0,1) değişkenini gerektirebilmektedir, buna karşın buradaki örnek zamana bağlı bir açığa çıkma değişkenini (exposure variable) içermektedir (Kleinbaum ve Klein, 2005, s.223).

Genişletilmiş Cox model aşağıda gösterildiği gibi sadece $E(t)$ değişkenini içermektedir. Bu modelde açığa çıkma değişkeninin değerleri, farklı denekler için zamana göre değişebilmektedir fakat modeldeki bir değişkenle ilgili sadece bir δ katsayısı bulunmaktadır. Bu yüzden δ , zamana bağlı $E(t)$ değişkeninin hayatta kalma süresi üzerindeki toplam etkisini temsil etmektedir (Kleinbaum ve Klein, 2005, s.223).

$$h(t, \mathbf{X}(t)) = h_0(t) \exp[\delta E(t)]$$

↑
bir katsayı

t zamanında olaya maruz kalan bir kişiyle maruz kalmayan bir kişiyi karşılaştıran tehlike oranıtısı formülü, aşağıdaki ifadeyi vermektedir (Kleinbaum ve Klein, 2005, s.223);

$$\widehat{HR}(t) = \frac{\widehat{h}(t, E(t)=1)}{\widehat{h}(t, E(t)=0)} = \exp[\widehat{\delta}(1 - 0)] = e^{\widehat{\delta}}, \text{ sabit bir sayı} \quad (4.16)$$

Bu sonuç sabit bir sayı olmasına rağmen, orantılı tehlike varsayımı sağlanmamaktadır. Sabit sayı, olayın bahsi geçen zamanda açığa çıkma durumunun pay kısmında 1 olduğunu ve açığa çıkmama durumunun payda kısmında 0 olduğunu varsayan belirli bir zamandaki tehlike oranıtısını ifade etmektedir. Bu yüzden tehlike oranıtısı zamana bağlıdır çünkü formül yalnızca tek bir sabit sayı vermesine rağmen, olayın açığa çıkma durumu zamana bağlıdır (Kleinbaum ve Klein, 2005, s.223).

4.5. Orantılı Tehlike Varsayımını Sağlamayan Zamandan Bağımsız Değişkenlerin Değerlendirilmesi

Bu bölümde, genişletilmiş bir Cox modelde zamandan bağımsız değişkenler için orantılı tehlike varsayımının kontrol edilmesinin ve orantılı tehlike varsayımını sağlamayan bir değişkenin etkisinin değerlendirmesinin nasıl yapılacağı anlatılmaktadır (Kleinbaum ve Klein, 2005, s.224).

Daha önce tanımlandığı gibi orantılı tehlike varsayımını değerlendirmek için ortak olarak kullanılan üç yöntem vardır: (1) grafiksel yöntemi kullanma- log-log hayatta kalma eğrileri (2) genişletilmiş Cox model kullanma ve (3) uyum iyiliği testi uygulama. Bu bölümde, ikinci yöntem üzerine odaklanılacaktır (Kleinbaum ve Klein, 2005, s.224).

Eğer çalışmanın veri seti p tane zamandan bağımsız değişken içeriyorsa, aşağıda gösterildiği gibi bu değişkenlerin her birini içeren bir Cox PH modeli kurmak istenilebilir (Kleinbaum ve Klein, 2005, s.224);

p tane zamandan bağımsız X için Cox PH modeli:

$$h(t, \mathbf{X}) = h_0(t) \exp\left[\sum_{i=1}^p \beta_i X_i\right] \quad (4.17)$$

Genişletilmiş Cox model:

Ancak, bir orantılı tehlike modeline uygun olup olmadığını değerlendirmek için, zamanın bazı fonksiyonlarıyla her bir zamandan bağımsız değişkeni içeren birçok çarpım terimi tanımlanarak bu model genişletilmektedir. Zamandan bağımsız i . değişken X_i olarak gösterildiğinde, i . çarpım terimi $X_i \times g_i(t)$ olarak tanımlanmaktadır. Burada $g_i(t)$ i . değişken için zamanın bir fonksiyonudur (Kleinbaum ve Klein, 2005, s.224).

İlgilenilen tüm bağımsız değişkenleri aynı anda ele alan genişletilmiş Cox model şu şekildedir (Kleinbaum ve Klein, 2005, s.224);

$$h(t, \mathbf{X}(t)) = h_0(t) \exp\left[\sum_{i=1}^p \beta_i X_i + \sum_{j=1}^p \delta_j X_j g_j(t)\right] \quad (4.18)$$

Bu genişletilmiş modeli kullanmada can alıcı karar, $g_i(t)$ fonksiyonlarının alması gereken formdur. $g_i(t)$ için en basit form, herhangi bir zamanda tüm $g_i(t)$ 'lerin özdeş olarak sıfır olduğu formdur. Bu, hiç zamana bağlı terim içermeyen orijinal orantılı tehlike modelini kurmanın başka bir yoludur (Kleinbaum ve Klein, 2005, s.225).

Örnek:

Bu teste bir örnek olarak, E 'nin temel etkisine ek olarak çarpım terimi $E \times t$ 'yi içeren genişletilmiş Cox model ele alınsın. Burada E , (0,1) olan zamandan bağımsız meydana gelen değişkeni ifade etmektedir (Kleinbaum ve Klein, 2005, s.226).

Genişletilmiş Cox Model:

$$h(t, \mathbf{X}(t)) = h_0(t) \exp[\beta E + \delta(E \times t)] \quad (4.19)$$

$H_0: \delta = 0$ (yani PH varsayımı sağlanmaktadır.)

Bu modelin orantılı tehlike varsayımını sağlayıp sağlamadığı için bir test, $\delta = 0$ olan sıfır hipotezini test etmek demektir. Bu hipotez altında indirgenmiş model, sadece E 'nin temel etkisini içeren orantılı tehlike modeli tarafından verilmektedir. Aşağıda gösterildiği gibi tam (yani genişletilmiş) model ve indirgenmiş (yani orantılı tehlike) model için log olabilirlik oranı istatistikleri arasındaki fark olan olabilirlik oranı istatistiği, büyük örneklerde bir serbestlik dereceli ki-kare dağılımına yaklaşmaktadır (Kleinbaum ve Klein, 2005, s.226).

İndirgenmiş Model:

$$h(t, \mathbf{X}) = h_0(t) \exp[\beta E] \quad (4.20)$$

$LR = -2 \ln L_R - (-2 \ln L_F)$ (H_0 altında 1 serbestlik dereceli χ^2 dağılımına yaklaşmaktadır.)

F = tam (genişletilmiş) model, R = indirgenmiş (orantılı tehlike) model

Bu iki farklı model tipinin testinde hesaplamaları yürütmek için, bir orantılı tehlike modeli ve bir genişletilmiş Cox model kurulmasına ihtiyaç duyulmaktadır (Kleinbaum ve Klein, 2005, s.226).

Orantılı tehlike varsayımı için testin sonucu anlamlı ise, orantılı tehlike modeli yerine genişletilmiş Cox model tercih edilmektedir. Dolayısıyla, ilgilenilen açığa çıkma değişkeninin etkisi için elde edilen HR değeri zamana bağlıdır. Çıktı üzerinde meydana gelen etki tek bir HR değeri tarafından özetlenememekte ancak sadece zamanın bir fonksiyonu olarak ifade edilebilmektedir (Kleinbaum ve Klein, 2005, s.226).

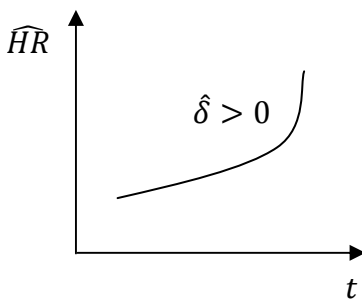
Örnek:

Genişletilmiş Cox modeli içeren önceki örnek aşağıda tekrar ele alınmaktadır. Bu model için, meydana gelen etkinin tahminlenmiş tehlike orantısı $e^{\hat{\beta} + \hat{\delta}t}$ ifadesi tarafından verilmektedir. Böylece, $\hat{\delta}$ 'ın pozitif ya da negatif olup olmasına bağlı olarak tahminlenmiş tehlike orantısı, t arttıkça üstel olarak artacak veya azalacaktır. Şekil 4.4, $\hat{\delta}$ pozitif olduğunda tehlike orantısı değerlerinin nasıl değiştiğinin bir taslağını vermektedir (Kleinbaum ve Klein, 2005, s.227).

$$h(t, \mathbf{X}(t)) = h_0(t) \exp[\beta E + \delta(E \times t)] \quad (4.21)$$

$$\widehat{HR} = \exp[\hat{\beta} + \hat{\delta}t] \quad (4.22)$$

Şekil 4.4: Zamana göre artan tehlike orantısı



4.6. Zamana Bağlı Değişkenler için Cox Modelin Çözüm Yöntemleri

Cox regresyon modeli için dört çeşit artık tanımlanmaktadır. Bu bağlamda dört temel artık analizinden söz edilmektedir. Bunlar; martingale, sapma, Cox-Snell ve Schoenfeld artıklarıdır (Yay, Çoker ve Uysal, 2007).

Cox-Snell, martingale ve sapma artıklarının iki önemli dezavantajı söz konusudur. Bu dezavantajlar, artıkların ağırlıklı olarak gözlenen hayatta kalma süresine bağlı olmalarını ve birikimli tehlike fonksiyonunun tahminini gerektirmesidir. Schoenfeld tarafından önerilen, skor artıkları olarak da adlandırılan Schoenfeld artıklarında bu sorunlar giderilmiştir. Bu yönüyle Schoenfeld artıkları, diğer artıklardan önemli bir farklılık göstermektedir. Bu artıklarda, her bireyin artığı için tek bir değer yerine, tahmin edilmiş olan Cox regresyon modelinde yer alan her bir bağımsız değişken için birer tane olmak üzere değerler kümesi yer almaktadır (Yay, Çoker ve Uysal, 2007).

Bilindiği gibi Schoenfeld artığı, değişkenin gerçek değeri ile ağırlıklı tehlike oranı ortalaması arasındaki fark şeklinde tanımlanmaktadır. İki kategori söz konusu olduğu durumda değişken için artık hesaplarırken, gerçek değer yerine geçecek iki ayrı (kategorik) değer olacağından anlamsızlık kaçınılmaz olacaktır. Bu nedenle kategorik değişkenler için Schoenfeld grafiğinin çizilmesi önerilmemektedir (Yay, Çoker ve Uysal, 2007).

Model içeriğindeki j . bağımsız değişken X_j 'nin i . Schoenfeld artığı aşağıdaki formül yardımıyla hesaplanmaktadır (Yay, Çoker ve Uysal, 2007):

$$r_{s_{ji}} = \delta_i \{x_{ji} - a_{ji}\} \quad (4.23)$$

Burada x_{ji} , $j = 1, 2, \dots, p$ olmak üzere çalışmadaki i . birey için j . bağımsız değişkenin değeridir (Yay, Çoker ve Uysal, 2007).

$$a_{ji} = \frac{\sum_{l \in R(t_i)} x_{jl} \exp(\hat{\beta}' x_l)}{\sum_{l \in R(t_i)} \exp(\hat{\beta}' x_l)} \quad (4.24)$$

a_{ji} , ağırlıklı tehlike ortalaması olarak da adlandırılmaktadır. Formüldeki $R(t_i)$, t_i zamanında risk altındaki tüm bireylerin oluşturduğu gruptur. Bu artıklar için sıfırdan farklı değerler, yalnızca sansürlü gözlemler için ortaya çıkmaktadır. Schoenfeld artıkları kısaca, bağımsız değişkenin gerçek değeri ile ağırlıklı tehlike oranı ortalaması arasındaki fark olarak tanımlanmaktadır (Yay, Çoker ve Uysal, 2007).

Bu artığın elde edilmesi için birkaç yol vardır. Schoenfeld artığı, modeldeki j . parametre için etkin skorun i . bileşeninin bir tahminidir. Etkin skor, β_j 'ye ilişkin kısmi olabilirlik fonksiyonunun logaritmasının birinci türevine eşittir (Yay, Çoker ve Uysal, 2007):

$$\frac{\partial \log L(\beta)}{\partial \beta_j} = \sum_{i=1}^n \left\{ x_{ji} - \frac{\sum_l x_{jl} \exp(\hat{\beta}' x_l)}{\sum_l \exp(\hat{\beta}' x_l)} \right\} \quad (4.25)$$

β_j 'nin tahminindeki i . terim, X_j için Schoenfeld artığını oluşturmaktadır (Yay, Çoker ve Uysal, 2007).

Orantılı tehlike varsayımının geçerliliği zamana bağlı çizilen Schoenfeld artıkları grafiği ile kontrol edilebilmektedir. Schoenfeld artıkları grafiğinde, artıkların yatay bir simetri göstermesi beklenmektedir. Çizilen Schoenfeld artıkları grafiği, yatay bir doğru etrafında seyrediyorsa, orantılı tehlike varsayımının sağlandığı söylenecektir (Yay, Çoker ve Uysal, 2007).

t_i zamanında risk grubunda bulunan l . birey için kullanılan ağırlık $\exp(\beta'x_l)$ 'dir ve Denklem (4.23)'teki x_{jl} katsayısı, bu birey için maksimize edilmiş kısmi olabilirlik fonksiyonunun bileşeni olan; $\frac{\exp(\hat{\beta}'x_l)}{\sum_{l \in R(t_i)} \exp(\hat{\beta}'x_l)}$, ye eşittir (Yay, Çoker ve Uysal, 2007).

β 'ların tahmini için; $\left. \frac{\partial \log L(\beta)}{\partial \beta_j} \right|_{\hat{\beta}} = 0$ olduğundan, Schoenfeld artıklarının toplamı sıfır olmalıdır. Bu artıklar için, büyük örneklerde r_{sji} 'nin beklenen değerinin sıfıra eşit olduğu ve artıkların bağımsız olduğu söylenebilmektedir (Yay, Çoker ve Uysal, 2007).

Bu artıkları kullanırken, t_i zamanında ölüm olasılıkları düşük olan bireyler, t_i zamanında ölüm riski altında olan bireylere göre, daha düşük değerli artıklara sahip olacaklardır. Dolayısıyla, ölüm riski altında olan bireylere ilişkin artıklar, ölüm riski az olan bireylere ilişkin artıklara oranla daha büyük olacaktır (Yay, Çoker ve Uysal, 2007).

Orantısızlık tespit edildikten sonra, zaman bağımlılığı farklı yollarla açıklanabilmektedir. Strateji, çalışma hedeflerine bağlı olacaktır. Eğer uzun zaman periyotlarının bir önemi yoksa orantısızlığın kısa zaman aralıklarında sorun olması pek muhtemel olmadığından, takip edilen zaman kısaltılabilmektedir. Zamanla değişken etkiye sahip bir değişkene özel bir ilgi yoksa bu değişken istatistiksel analizlerde tabakalandırılabilir. Ancak tabakalandırılmış değişken ve hayatta kalma arasında hiç bir ilişki test edilememektedir. Eğer zaman boyuca değişkenin etkisi tanımlanmak istenirse, değişken etkileşimleri tarafından zamana veya farklı zaman noktalarında göreceli riskleri tahmin etmek için artıkların grafiğine güvenmek mümkündür (Bellera vd., 2010).

Orantısızlığı test etmek için en iyi stratejinin kesin kurallarını ileri sürmek zordur. Her yöntem avantajlara ve kısıtlamalara sahiptir ve bazı yaklaşımlar çalışmanın amacına bağlı olarak tercih edilebilmektedir. İstatistiksel modelleme yapmadan önce kullanılacak istatistiksel testler gibi çalışma amaçları, önceden açıkça saptanmalıdır. PH varsayımı grafiksel ya da sayısal yaklaşımlar kullanarak araştırılabilmektedir. Grafik metotları, orantısızlığı test eden değişken için Kaplan-Meier hayatta kalma eğrilerinin görselleştirilmesini içermektedir. Bu grafiksel yöntem kategorik değişkenler gerektirmektedir ve özellikle ikili verilere uygundur, ancak bu veriler biçimsel tanımlayıcı testleri

vermemektedir. Sayısal testler örneğin, değişken-zaman etkileşimlerini (covariate-by-time interactions) ya da modelin artıklarında bir trendin varlığını test etmeyi içermektedir. Bir değişken-zaman etkileşimini Cox modelin içine katmak her şeyden önce basittir, ancak sonuçlar zaman fonksiyonunun fonksiyonel formunun seçimine bağlıdır. Birikimli artıklara dayanan testler, Schoenfeld artıklarına dayananlardan daha iyi istatistiksel özelliklere sahip olmaya eğilimlidir. Sonuç olarak, birikimli artıklara dayanan bir testi yapmak, zamanla değişen etkilere sahip değişkenleri ortaya çıkarmada daha güçlü bir yaklaşım gibi görünmektedir (Bellera vd., 2010).

İstatistiksel test, sınama gücü problemini yani gerçek etkileri bulmada testlerin yeterliliğini arttırmaktadır. Gözlem periyodunu kısaltma gibi bazı basit yöntemler, daha az olay ele alındığı için azalan sınama gücünden zarar görebilmektedir. Bu durum küçük veri setlerinde bir kısıtlama olabilmektedir. Tabakalandırılmış Cox modellemesi, genellikle daha geniş güven aralıkları vermektedir yani tabakalandırılmamış analizlere nazaran testin sınama gücü azalmaktadır. Zamanla değişen etkilerin istatistiksel testleri, orantısızlığı tespit etmek için farklı bir güce sahiptir. Zamana bağlı değişkenlere ya da Schoenfeld artıklarına dayanan testler orantılı olmayan tehlikelerdeki bir değişikliğin orantısızlığını saptamada aynı derecede iyi ve neredeyse birbirine eşit sınama gücüne sahiptir. Ancak başarısızlık zamanının bölümlere ayrılmasını gerektiren testler diğer testlerden daha az sınama gücüne sahiptir. Sınama gücü problemi, örneklem büyüklüğü sorununu doğurmaktadır. Klinik çalışmalar genellikle tedavinin etkisini saptamak için sadece yeterli sınama gücü ile yapılandırılmaktadır. Gözlemsel çalışmalar genellikle keşif amaçlı analizler için yapılandırılmaktadır ve örnek hacminin şekilsel (formal) bir tahminine güvenilmemektedir. Belirli bir zaman trendini tespit etmede yeterli sınama gücü her zaman olmayabilmektedir. Sınama gücü yetersizliği sorunu, orantısızlığı test etmeye karşı bir delil olarak yorumlanmamalıdır. Diğer istatistiksel modellerde olduğu gibi, temel varsayımların ihlal edilmediğinden emin olunmalıdır (Bellera vd., 2010).

BEŞİNCİ BÖLÜM

UYGULAMA

Bu çalışmada; Zonguldak, Tarsus, Ereğli, Eskişehir, Adana, Mersin, Van, Avanos, Samsun, Elazığ, Yozgat ve Diyarbakır'da bulunan İŞGEM'lerde faaliyetlerine devam eden, daha önce İŞGEM'e kabul edilerek mezun olmuş ya da süresi dolmadan ayrılmış firmaların verisi kullanılmaktadır. Ancak İŞGEM'den herhangi bir sebeple ayrılan firmalarla yüz yüze görüşme imkanı sağlanamadığından, bu firmalarla ilgili sadece mevcut bilgiler kullanılabilmiştir.

Uygulamada kullanılan veri seti, Prof.Dr.Murat Karaöz'ün yürütücülüğünü yaptığı "İş Kuluçkalarında Yeni Kurulan Girişimlerin Hayatta Kalma ve Büyüme Performansını Etkileyen Faktörler: KOSGEB İş Geliştirme Merkezleri (İŞGEM) Üzerine Bir Araştırma" konulu TÜBİTAK projesi için yapılan anketlerden elde edilmiştir.

İş kuluçkaları; ortak olarak kullanılan ofis alanı, uzman personel, makine, donanım, tesis ve iş yardımı sunarak firmaların ilk dönem gelişimini kolaylaştıran bir organizasyon olarak tanımlanabilir (Aernoudt, 2004). İş kuluçkası çatı bir organizasyon olup bünyesinde pek çok yeni kurulmuş girişimleri barındırmaktadır. İŞGEM ise, Türkiye'de faaliyet gösteren önemli iş kuluçkaları uygulamalarından birisidir. İŞGEM genellikle küçük girişimcileri mevcut imkanlarıyla bir iş kurmaları konusunda destekleyen, iş kurulduktan sonra bu işin gelişmesi ve hayatta kalması konusunda işletmelere destek sağlamak amacıyla oluşturulmuş bir kuruluştur.

Genç girişimlerin başlangıçtaki başarılarını etkileyen faktörleri şu şekilde sıralamak mümkündür (Hackett ve Dilts, 2004);

1. Yeni girişim sahibinin eğitim düzeyi ve sektör tecrübesi gibi insan sermayesi özellikleri,
2. Ölçek, yaş ve insan sermayesi gibi firma özellikleri,
3. Pazar büyüme hızı ve giriş engelleri gibi endüstri özellikleri,
4. Kuluçka özellikleri,
5. Makroekonomik dalgalanmalar, bölgesel faktörler ve kamusal politikalar gibi diğer dışsal faktörler.

Holmes vd. (2008); firmaya özgü, makroekonomik ve endüstriye özgü değişkenlerin firmaların hayatta kalması üzerinde önemli bir etkiye sahip olduğunu bulmuştur.

İnsan sermayesi ile ilgili olarak; yeni firma kurulması ve firmanın hayatta kalması üzerinde girişimcinin yaşı, cinsiyeti, eğitim seviyesi, ailesi, gelir kaynakları ve kariyer geçmişinin etkileri ele alınmaktadır.

Firma özellikleri; kuruluştaki firma büyüklüğü, ortak sayısı, sermaye yoğunluğu, ihracat düzeyi, firmada yenilik yapılıp yapılmaması ve firmanın reklama ve ağbağlaşmaya önem verip vermemesi ve marka sahibi olup olmaması gibi faktörler ile değerlendirilmektedir.

Endüstri özellikleri için; firmanın bulunduğu sektör ve o sektördeki rekabet yoğunluğu dikkate alınmaktadır.

Kuluçka özellikleri ise; firmanın ait olduğu kuluçkanın yaşı ve işlik kapasitesi, kuluçkadaki büro hizmetleri, ortak hizmetler ve danışmanlık hizmetleri olarak gruplandırılan değişkenler kullanılarak incelenmektedir.

Dışsal faktörlerde ise; firmanın bir kriz zamanı tecrübe edip etmemesi ve kuluçkanın bulunduğu ilin gelir düzeyi ele alınmaktadır.

İllerde yer alan toplam firma sayıları ve anket uygulaması yapılan firma sayıları özet olarak aşağıdaki tabloda gösterilmektedir.

Tablo 5.1: İller itibariyle araştırmaya dahil olan firma sayıları

No	İŞGEM	İŞGEM'deki Mevcut+Mezun Toplam Firma Sayısı	Anket Uygulanan Mevcut+Mezun Toplam Firma Sayısı
1	Adana	87	57
2	Diyarbakır	15	14
3	Elazığ	29	20
4	Ereğli	37	35
5	Zonguldak	31	31
6	Mersin	23	23
7	Tarsus	119	119
8	Samsun	32	32
9	Van	38	38
10	Yozgat	24	17
11	Eskişehir	31	21
12	Nevşehir	Bilinmiyor	7
	Genel Toplam	466	414

Kaynak: Karaöz M. ve Albeni M., "İş Kuluçkalarında Yeni Kurulan Girişimlerin Hayatta Kalma ve Büyüme Performansını Etkileyen Faktörler: KOSGEB İş Geliştirme Merkezleri (İŞGEM) Üzerine Bir Araştırma", 109K139 nolu TÜBİTAK Projesi, Isparta, Mart 2011.

5.1. Araştırmada Kullanılan Değişkenler

Bu uygulamada firmaların hayatta kalma süresini etkileyen faktörleri analiz etmek için kullanılacak değişkenlerin tümü Karaöz ve Albeni (2011)'nin çalışması baz alarak belirlenmiştir.

Girişimcinin yaşı, cinsiyeti, eğitim durumu, mesleki kariyer geçmişi ve birikimi ile aile tecrübesi faktörleri literatürde firmaların hayatta kalması açısından temel faktörlerdir (Karaöz ve Albeni, 2011, s.26). Bu faktörler dikkate alınarak girişimcinin cinsiyeti (**gender**), yaşı (**lnentage**), eğitim durumu (**enteduuni**), çalışma tecrübesi (**workexp**), girişimcilik tecrübesi (**entexp**), yöneticilik tecrübesi (**mangexp**) ile aile ve çevre etkisi (**aillecev**) değişkenleri analize dahil edilecektir. Girişimcinin iş tecrübesiyle ilgili (**workexp**), (**entexp**), (**mangexp**) olmak üzere 3 adet değişken oluşturulmuştur. Bunlar sırasıyla; girişimcinin başka firmalarda çalışan olarak kaç yıl iş tecrübesi olduğunu, girişimci olarak geçmişte tecrübesinin kaç yıl olduğunu ve yönetici olarak tecrübesinin kaç yıl olduğunu ifade etmektedir. Girişimcinin aile özellikleri ile ilgili olarak kullanılan (**aillecev**) değişkeni, girişimcinin ailesinde ve yakın çevresinde kendisinden başka girişimci / firma sahibi olup olmadığını veya girişimcinin ailesinde veya yakın çevresinde deneyimlerinden yararlandığı girişimcilerin olup olmadığını göstermektedir. Ayrıca girişimcinin sadece yeni kurduğu firmadan elde ettiği gelire bağımlı olup olmadığını gösteren (**tekgelir**) değişkenine de yer verilmiştir.

Firma özellikleri ile ilgili olarak firmadaki ortak sayısı (**ortaksay**), firmanın ihracat yapıp yapmadığı (**export**), firmanın piyasaya giriş ölçeği (**lnempini**), firmanın sadece borç alınarak kurulup kurulmadığı (**onlyloan**), firmanın sadece girişimcinin kendi finansal imkanları ile kurulup kurulmadığı (**onlyselffin**), firma sahibinin kuluçka içi ve dışı aktörlerle ağyapı ilişkisi kurup kurmadığı (**networking**), firmanın yenilik yapıp yapmadığı (**yenilik**), reklam yapıp yapmadığı (**reklam**) ve marka sahibi olup olmadığı (**marka**) değişkenleri analize dahil edilmiştir.

Kuluçkada sunulan hizmetlerle ilgili (**buro**), (**ortakhiz**) ve (**danismanlik**) değişkenleri modele alınmıştır. Aynı zamanda 36 aylık dönemde kuluçkalara uygulanan özel bu sübvansiyonların kuluçka performansını ve dolayısıyla firma performansını etkileyip etkilemediği araştıran (**whenest**) değişkeni ve kuluçkanın kapasitesini gösteren (**incubsize**) değişkeni ele alınmıştır.

Endüstri özelliklerini yansıtan, firmanın içinde bulunduğu sektörü belirten (**sektor**) değişkeni ve o sektördeki rekabet yoğunluğunu gösteren (**compete**) değişkeni kullanılmıştır. Dışsal faktörlerle ilgili ise kuluçkanın bulunduğu ilin kişi başı GSMH'nin Türkiye'nin kişi

başı GSMH'sı içerisindeki payını (yüzde cinsinden) gösteren (**ilrank**) değişkeni ile kuluçkadaki firmanın ekonomik kriz dönemini tecrübe edip etmediğini gösteren (**cycle**) değişkeni kullanılmaktadır.

Tablo 5.2: Analizde kullanılan değişkenler ve tanımlayıcı istatistikler

DEĞİŞKEN	TANIMI	Gözlem Sayısı	Ortalama	Minimum	Maksimum	
	İLGİLENİLEN OLAY					
firmexit	Eğer firma kuluçkada bulunduğu süre içerisinde kapanmış ise 1, kapanmamış ise 0	414	-	0	1	
	BAĞIMLI DEĞİŞKEN					
incubtime	Firmanın kuruluşundan kuluçkadan ayrılma tarihine kadar geçen süre (ay)	404	29,12	0	158	
	BAĞIMSIZ DEĞİŞKENLER					
GİRİŞİMCİ ÖZELLİKLERİ	tekgelir	Girişimcinin geliri sadece genç firmadan ise 1, değilse 0	414	-	0	1
	gender	Girişimci bayan ise 1, erkek ise 0 (hem bay hem de bayan ortak var ise 0)	414	-	0	1
	lnentage	Girişimcinin yaşı (Ortaklık var ise en yaşlı girişimci-logaritmik ölçek)	367	3,64	3	4,25
	enteduuni	Girişimci üniversite mezunu ise 1, değilse 0 (ortaklık varsa ve ortaklardan birisi üniversite mezunu ise 1)	414	-	0	1
	workexp	Girişimcinin İŞGEM'e gelmeden önceki dönemde işgören olarak çalışma tecrübesi (yıl)	414	8,21	0	53
	entexp	Girişimcinin İŞGEM'e gelmeden önceki dönemdeki girişimcilik tecrübesi (yıl)	414	5,83	0	40
	mangexp	Girişimcinin İŞGEM'e gelmeden önceki dönemdeki yöneticilik tecrübesi (yıl)	414	6,98	0	40
	ailecev	Girişimcinin aile ve/veya çevresinde girişimci olmasını etkileyen	414	-	0	1

		birisinin olması durumunda 1, yoksa 0				
FİRMA ÖZELLİKLERİ	ortaksay	Girişimdeki ortak sayısı	414	1,24	1	4
	export	Firma ihracat yapmış ise 1, yapmamışsa 0	414	-	0	1
	lnempini	İlk kuruluştaki firma büyüklüğü (logaritmik ölçek)	392	1,31	0	5,70
	onlyloan	Girişimin kuruluşundaki sermaye tamamıyla borç ise 1, değilse 0	414	-	0	1
	onlyselffin	Girişimin kuruluşundaki sermaye tamamıyla kendi parası ise 1, değilse 0	414	-	0	1
	networking	Girişimci kuluçka içinde ve dışında paydaşlarla işbirliği içerisinde ise 1, değilse 0	414	-	0	1
	yenilik	Girişimci yenilik yapmış ise 1, yapmamış ise 0	414	-	0	1
	reklam	Firma reklam yapmış ise 1, yapmıyor ise 0	414	-	0	1
	marka	Firma marka sahibi ise 1, değilse 0	414	-	0	1
KULUÇKA HİZMETLERİ VE ÖZELLİKLERİ	buro	Girişimci kuluçkadan sunulan büro hizmetlerinden en az birisini almış ise 1, almamış ise 0	414	-	0	1
	ortakhiz	Girişimci kuluçkadan sunulan ortak hizmetlerinden en az birisini almış ise 1, almamış ise 0	414	-	0	1
	danismanlik	Girişimci kuluçkadan sunulan danışmanlık hizmetlerinden en az birisini almış ise 1, almamış ise 0	414	-	0	1
	whenest	Firma kuluçkanın kurulduğu il 36 ay içerisinde kurulmuş ise 1, değilse 0	414	-	0	1
	incubsize	Kuluçkadaki işlik sayısı	414	43,14	14	84
ENDÜSTRİ ÖZELLİKLERİ	sektor	Firma imalat sanayinde ise 1, hizmet sektöründe ise 0	411	-	0	1
	compete	Sektördeki rekabet yoğunluğu (1-5 likert ölçeği)	410	-	1	5

DIŞSAL ÖZELLİKLER	ilrank	Kuluçkanın bulunduğu ilin KBGSMH'nın Türkiye KBGSMH'sı içerisindeki payı (%)	414	1,51	0,59	2,07
	cycle	Firma kriz zamanını tecrübe etmiş ise 1, etmemişse 0	414	-	0	1

Kaynak: Karaöz M. ve Albeni M., "İş Kuluçkalarında Yeni Kurulan Girişimlerin Hayatta Kalma ve Büyüme Performansını Etkileyen Faktörler: KOSGEB İş Geliştirme Merkezleri (İŞGEM) Üzerine Bir Araştırma", 109K139 nolu TÜBİTAK Projesi, Isparta, Mart 2011.

Kuluçka hizmetleri (**bu**ro), (**danisman**) ve (**ortakhiz**) olmak üzere üç değişkende toplanmıştır. Her bir değişkenin alt-unsurları Tablo 5.3'te sunulmaktadır. Her bir grupta var olan alt-unsurların her biri için yapılan regresyon analizleri, ayrı olarak bir arada sunulmaktadır.

Tablo 5.3: Analizde kullanılan kuluçka hizmetleri değişkenlerinin (**bu**ro, **ortakhiz**, **danismanlik**) alt-unsurları

BÜRO HİZMETLERİ (GENEL İDARİ HİZMETLER)	
Büro hizmetleri ve sekreterlik hizmeti varsa yararlanıyor musunuz?	
Muhasebe ve defter tutma hizmeti varsa yararlanıyor musunuz?	
Bilgisayar ve ekipmanı sağlama hizmeti varsa yararlanıyor musunuz?	
İnternet erişimi hizmeti varsa yararlanıyor musunuz?	
Web sitesi hazırlama-hosting hizmeti varsa yararlanıyor musunuz?	
Telefon hizmetleri (santral, telesekreter) varsa yararlanıyor musunuz?	
Mobilya (masa, sandalye, kitaplık vb) sağlama hizmeti varsa yararlanıyor musunuz?	
Fotokopi hizmetleri varsa yararlanıyor musunuz?	
Faks hizmetleri varsa yararlanıyor musunuz?	
ORTAK HİZMETLER	
Posta / kargo hizmetleri varsa yararlanıyor musunuz?	
Kütüphane hizmeti varsa yararlanıyor musunuz?	
Güvenlik hizmetleri varsa yararlanıyor musunuz?	
Çay salonu-yemek salonu hizmetleri varsa yararlanıyor musunuz?	
Toplantı / konferans salonu hizmeti varsa yararlanıyor musunuz?	
Ofis büyüklüğü konusunda esnek davranma imkanı varsa yararlanıyor musunuz?	
Eğitim (kurs, seminer) hizmetleri varsa yararlanıyor musunuz?	
DANIŞMANLIK HİZMETLERİ	
Yönetim danışmanlığı hizmeti varsa yararlanıyor musunuz?	
Personel yönetimi – İKY konusunda danışmanlık hizmeti varsa yararlanıyor musunuz?	
Hukuki temsil ve danışmanlık hizmeti varsa yararlanıyor musunuz?	
Pazarlama / reklam danışmanlığı hizmeti varsa yararlanıyor musunuz?	
İhracat / ithalat danışmanlığı hizmeti varsa yararlanıyor musunuz?	
Banka kredileri / finansman danışmanlığı hizmeti varsa yararlanıyor musunuz?	
Vergi, sigorta, prim ödemeleri konusunda danışmanlık hizmeti varsa yararlanıyor musunuz?	
Ulusal kurumlardan (KOSGEB vd) kredi / hibe / destekler ile ilgili danışmanlık hizmeti varsa yararlanıyor musunuz?	
Patent başvurusu konusunda danışmanlık hizmeti varsa yararlanıyor musunuz?	

Kaynak: Karaöz M. ve Albeni M., "İş Kuluçkalarında Yeni Kurulan Girişimlerin Hayatta Kalma ve Büyüme Performansını Etkileyen Faktörler: KOSGEB İş Geliştirme Merkezleri (İŞGEM) Üzerine Bir Araştırma", 109K139 nolu TÜBİTAK Projesi, Isparta, Mart 2011.

5.2. Araştırmada Kullanılan Değişkenlerin Özellikleri

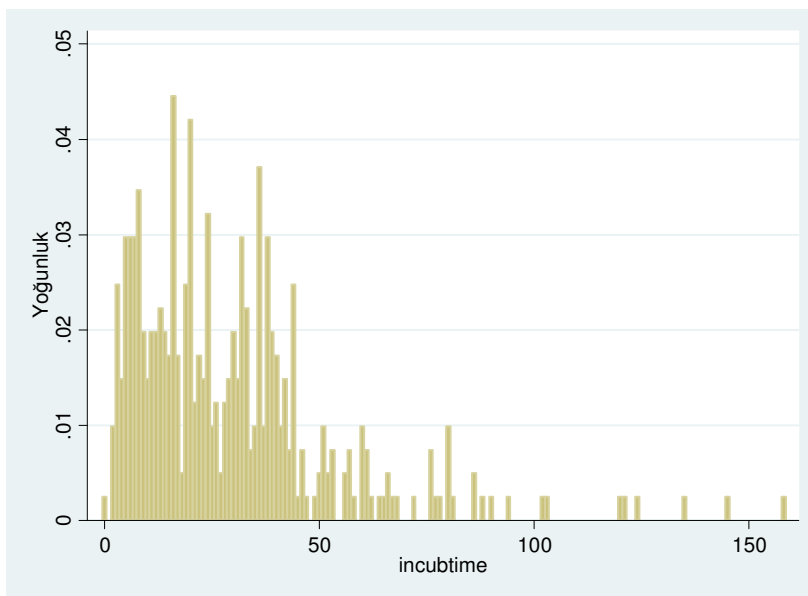
Bağımlı değişken olarak (**firmexit**) değişkeni kullanılmıştır. Firma eğer kuluçkada bulunduğu süre içerisinde kapanmışsa 1, kapanmamış ve / veya mezun olmuşsa 0 değerini almaktadır. Firmanın kuluçkadan çıktıktan sonra kapanıp kapanmadığı analizde dikkate alınmamıştır.

Modelde bağımsız değişken olarak kullanılacak olan kategorik değişkenler; (**tekgelir**), (**gender**), (**enteduuni**), (**ailceev**), (**ortaksay**), (**export**), (**onlyloan**), (**onlyselffin**), (**networking**), (**yenilik**), (**reklam**), (**marka**), (**buro**), (**ortakhiz**), (**danismanlik**), (**whenest**), (**sektor**), (**compete**), (**cycle**)'dır. Nümerik değişkenler ise; (**lnentage**), (**workexp**), (**entexp**), (**mangexp**), (**lnempini**), (**incubsize**), (**ilrank**)'tır.

Kuluçka hizmetleri değişkenleri olan (**buro**), (**danisman**) ve (**ortakhiz**) değişkenleri arasındaki yüksek korelasyon nedeniyle her birisi için ayrı ayrı regresyon tahminleri yapılmıştır. Böylece her bir açıklayıcı değişkenin bireysel etkilerini ayrı ayrı görmek mümkün olabilecektir. Benzer durum (**workexp**), (**entexp**) ve (**mangexp**) değişkenleri ve (**onlyloan**), (**onlyselffin**) değişkenleri arasında da geçerlidir.

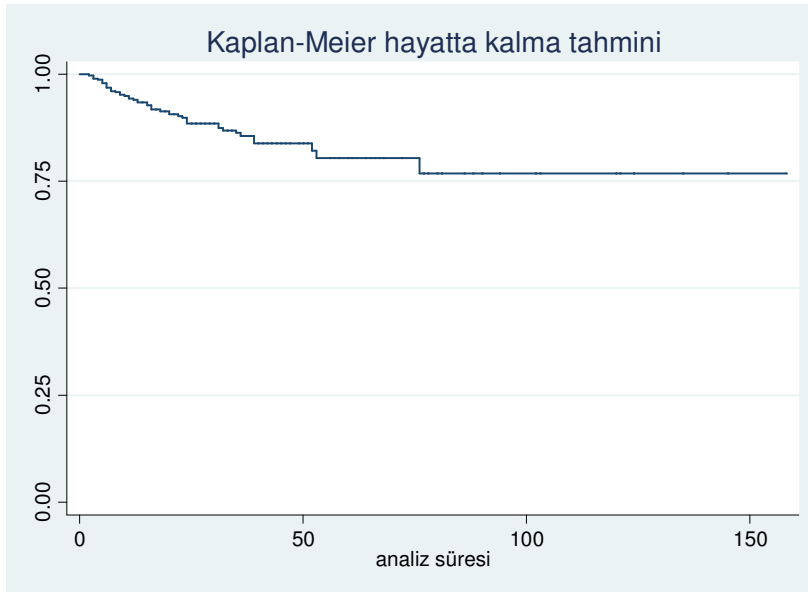
Kuluçkada kalma süresi ay cinsinden ölçülmektedir. Anket verilerine göre ortalama kuluçkada kalma süresi 29,12 aydır. Maksimum kalma süresi 158 ay olarak gerçekleşmiştir. Kuluçkada kalma süreleri (**incubtime**) ile ilgili histogram aşağıda Şekil 5.1'de sunulmaktadır.

Şekil 5.1: Firmaların kuluçkada kalma sürelerinin dağılımı (ay)



Firmaların kuluçka içerisinde iken kapanmalarının zaman içerisindeki yayılımını gösteren hayatta kalma eğrisi Şekil 2’de aşağıda sunulmaktadır. Buna göre kuluçkadaki firmaların önemli bir kısmının kurulduktan itibaren ilk 35-40 ay içerisinde kapandığı gözlenmektedir. İlerleyen sürelerde hala kuluçkada kalmaya devam eden firmaların kapanma riskleri hemen hemen kalmamaktadır.

Şekil 5.2: Kuluçka içerisindeki firmaların hayatta kalma eğrisi (ay)



Hayatta kalma tahminlerinin bu kadar yüksek olmasının sebebi ise, sadece kuluçka içindeyken kapanan firmaların ele alınmasıdır. Yani firmalar kuluçkadan mezun olduktan sonra kapansa dahi, kuluçkadan başarılı bir şekilde ayrıldığı için analizde başarılı olarak ele alınmıştır.

5.3. Cox Orantılı Tehlike Regresyon Modeli: Temel Model Tahmin Sonuçları

Cox model tahmin değerlerini sunan tablolarda, hem katsayı tahminleri hem de tehlike orantıları sunulmuştur. Tehlike orantılarını yorumlamak nispeten daha kolay olmaktadır. Ancak (0,1) şeklindeki iki kategorili değişkenlerin doğrudan doğruya katsayı ve tehlike orantılarını gözleyerek pratik olarak yorumlama imkanı olduğundan modellerde kullanılan çoğu açıklayıcı değişken mümkün olduğunca kategorik olarak alınmıştır.

Cox regresyon analizinde değişken düzeylerinden biri referans kategorisi olarak alınmakta ve değişken düzeylerinin yorumlanması buna göre yapılmaktadır. Bu çalışmada, modeldeki değişken için β parametresi ve standart hatası (S.H.), p değeri, tehlike orantısı ($\exp(\beta)$) ile değişken düzeyleri için tehlike orantısının alt ve üst sınırları verilmektedir. β

parametresinin pozitif değer olması bu düzeyin referans kategorisine göre daha fazla riskli olduğunu, β parametresinin negatif değer olması ise bu düzeyin referans kategorisine göre daha az riskli olduğunu göstermektedir. Tehlike orantısı olan $\exp(\beta)$ değeri ise önemli bulunan düzeyin, referans kategorisine göre kaç kat (ya da % ne kadar) daha riskli olduğu yorumunu getirmektedir. Önemli bulunan değişkenlerin her bir düzeyine karşılık gelen p değerlerine bakılarak önemli değişken düzeyleri belirlenmektedir (Sertkaya, Ata ve Sözer, 2005).

Tehlike orantısı 0'dan başlayarak sonsuza doğru gitmektedir. Bu çerçevede tehlike orantılarının 1'den büyük olması değişkenin 1 kategorisi özelliğini taşıyan grubun 0 kategorisi gruba göre daha yüksek olasılıkla olaya maruz kalacağını gösterecektir (Kleinbaum ve Klein, 2002).

Sürekli değişkenleri değerlendirmek, kategorik değişkenleri değerlendirmek kadar kolay değildir. Ancak istendiğinde, bu tür değişkenler genellikle ortalama değer ya da aranan diğer bazı özel değerlere ilişkin karşılaştırmalı hesaplar yapılarak yorumlanabilmektedir.

Hayatta kalma üzerinde en etkili olan girişimci tecrübesi olarak ifade edilen (**workexp**), (**entexp**) ve (**mangexp**) değişkenleri kendi arasında; girişimcinin sermayesinin kendisine ait olup olmadığını belirten (**onlyloan**) ve (**onlyselffin**) değişkenleri kendi arasında; kuluçka hizmetleri ile ilgili (**buro**), (**ortakhiz**) ve (**danismanlik**) değişkenlerini de kendi arasında yüksek korelasyonludur. Bu nedenle bu değişkenlerin aynı modelde yer almaması ve firmaların hayatta kalmasını en iyi açıklayan değişkenlerin seçilip nihai modele koyulması gerekmektedir. Bu amaçla, hayatta kalma üzerinde en etkili olan girişimci tecrübesi için sırasıyla (**workexp**), (**entexp**) ve (**mangexp**) değişkenlerini içeren Model 1, Model 2 ve Model 3 kurulmuştur. Girişimcinin sermaye yoğunluğunu belirten (**onlyloan**) ve (**onlyselffin**) değişkenleri ile Model 4 ve Model 5 oluşturulmuştur. Model 6, Model 7 ve Model 8 ise kuluçka hizmetleri ile ilgili (**buro**), (**ortakhiz**) ve (**danismanlik**) değişkenlerini ayrı ayrı incelemek için oluşturulmuştur. Her bir değişken için yer alan bilgilerden ilki tehlike orantısını, ikincisi değişkenin katsayı tahminini ve sonuncusu da p değerini göstermektedir.

Tablo 5.4: Hayatta kalma üzerinde etkili olan girişimci tecrübelerini, sermaye yoğunluğunu, kuluçka hizmetlerini tek tek ele alan Cox model tahminleri

Değişken	Tehlike Orantısı / Katsayı							
	[P > z]							
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
tekgelir	19.317	10.486	17.116	62.178	54.055	46.063	40.447	33.115
	2.961	2.350	2.840	4.130	3.990	3.830	3.700	3.500
	[0.001]***	[0.004]***	[0.002]***	[0.000]***	[0.000]***	[0.001]***	[0.001]***	[0.000]***

gender	1.014 0.014 [0.988]	0.938 -0.064 [0.945]	0.807 -0.215 [0.819]	0.389 -0.944 [0.324]	0.453 -0.792 [0.421]	0.515 -0.664 [0.483]	0.527 -0.641 [0.518]	0.501 -0.691 [0.487]
lnentage	34.467 3.540 [0.016]**	87.357 4.470 [0.017]**	138.380 4.930 [0.009]***	186.793 5.230 [0.005]***	699.244 6.550 [0.002]***	550.045 6.310 [0.002]***	387.610 5.960 [0.002]***	550.045 6.310 [0.002]***
enteduuni	0.729 -0.316 [0.665]	1.573 0.453 [0.514]	1.474 0.388 [0.571]	0.769 -0.263 [0.713]	1.015 0.014 [0.984]	1.224 0.202 [0.780]	1.443 0.367 [0.605]	1.007 0.007 [0.993]
workexp	0.908 -0.097 [0.038]**	-	-	0.899 -0.106 [0.054]*	0.886 -0.121 [0.034]**	0.899 -0.107 [0.058]*	0.902 -0.103 [0.064]*	0.888 -0.119 [0.033]**
entexp	-	0.921 -0.083 [0.102]	-	1.012 0.012 [0.925]	0.937 -0.065 [0.631]	0.945 -0.056 [0.684]	0.950 -0.052 [0.687]	0.918 -0.086 [0.492]
mangexp	-	-	0.907 -0.098 [0.042]**	0.891 -0.115 [0.289]	0.941 -0.061 [0.605]	0.929 -0.074 [0.536]	0.955 -0.046 [0.688]	0.982 -0.018 [0.867]
ailecev	1.172 0.159 [0.789]	1.387 0.327 [0.582]	1.192 0.176 [0.763]	0.852 -0.160 [0.787]	1.132 0.124 [0.845]	1.060 0.059 [0.924]	0.992 -0.008 [0.988]	1.340 0.293 [0.644]
ortaksay	0.013 -4.360 [0.009]***	0.014 -4.300 [0.019]**	0.011 -4.480 [0.017]**	0.013 -4.320 [0.006]***	0.004 -5.480 [0.004]***	0.006 -5.080 [0.004]***	0.010 -4.620 [0.006]***	0.006 -5.150 [0.004]***
export	4.855 1.580 [0.277]	4.953 1.600 [0.272]	6.172 1.820 [0.200]	8.331 2.120 [0.144]	6.619 1.890 [0.224]	6.050 1.800 [0.262]	5.641 1.730 [0.259]	6.424 1.860 [0.204]
lnempini	1.190 0.174 [0.519]	1.075 0.072 [0.795]	0.977 -0.024 [0.935]	0.815 -0.205 [0.472]	0.884 -0.123 [0.674]	0.929 -0.074 [0.807]	0.974 -0.027 [0.928]	0.955 -0.046 [0.874]
onlyloan	2.181 0.780 [0.476]	1.660 0.507 [0.639]	1.303 0.265 [0.810]	0.564 -0.572 [0.558]	-	1.126 0.119 [0.917]	0.896 -0.110 [0.921]	1.735 0.551 [0.620]
onlyselffin	5.989 1.790 [0.066]*	6.959 1.940 [0.037]**	5.104 1.630 [0.074]*	-	5.585 1.720 [0.051]*	5.207 1.650 [0.071]*	7.463 2.010 [0.032]**	9.583 2.260 [0.029]**
networking	0.017 -4.090 [0.000]***	0.038 -3.280 [0.001]***	0.030 -3.500 [0.000]***	0.032 -3.440 [0.000]***	0.018 -4.030 [0.000]***	0.025 -3.680 [0.001]***	0.022 -3.800 [0.000]***	0.014 -4.260 [0.000]***
yenilik	0.039 -3.250 [0.004]***	0.092 -2.390 [0.030]**	0.079 -2.540 [0.023]**	0.079 -2.540 [0.028]**	0.063 -2.760 [0.027]**	0.069 -2.680 [0.029]**	0.058 -2.840 [0.016]**	0.046 -3.080 [0.009]***
reklam	1.198 0.181 [0.789]	0.962 -0.039 [0.955]	0.961 -0.040 [0.954]	1.175 0.161 [0.825]	1.194 0.177 [0.772]	1.129 0.121 [0.867]	1.284 0.250 [0.720]	1.058 0.056 [0.931]
marka	4.904 1.590 [0.233]	4.437 1.490 [0.279]	6.360 1.850 [0.173]	4.953 1.600 [0.195]	7.389 2.000 [0.143]	8.248 2.110 [0.121]	8.758 2.170 [0.111]	6.297 1.840 [0.159]
buro	0.310 -1.170 [0.155]	0.278 -1.280 [0.181]	0.239 -1.430 [0.140]	0.145 -1.930 [0.037]**	0.217 -1.530 [0.122]	0.156 -1.860 [0.003]***	-	-
ortakhiz	1.962 0.674 [0.586]	1.274 0.242 [0.846]	1.301 0.263 [0.831]	2.974 1.090 [0.372]	1.745 0.557 [0.660]	-	0.181 -1.710 [0.028]**	-
danismanlik	0.221 -1.510 [0.098]*	0.370 -0.995 [0.249]	0.413 -0.885 [0.307]	0.340 -1.080 [0.211]	0.289 -1.240 [0.155]	-	-	0.137 -1.990 [0.009]***
whenest	0.095 -2.350 [0.012]**	0.061 -2.790 [0.005]***	0.061 -2.790 [0.004]***	0.089 -2.420 [0.007]***	0.046 -3.090 [0.004]***	0.065 -2.730 [0.003]***	0.045 -3.100 [0.001]***	0.037 -3.290 [0.001]***
incubsize	0.932	0.923	0.919	0.927	0.922	0.928	0.930	0.921

	-0.070 [0.003]***	-0.080 [0.006]***	-0.084 [0.003]***	-0.076 [0.002]***	-0.081 [0.002]***	-0.075 [0.002]***	-0.072 [0.002]***	-0.082 [0.003]***
sektor	1.362 0.309 [0.793]	2.945 1.080 [0.399]	2.620 0.963 [0.446]	1.556 0.442 [0.706]	2.326 0.844 [0.509]	2.413 0.881 [0.504]	1.726 0.546 [0.646]	2.210 0.793 [0.497]
compete	1.028 0.028 [0.925]	1.010 0.010 [0.974]	0.978 -0.022 [0.941]	0.928 -0.075 [0.797]	0.972 -0.028 [0.923]	0.985 -0.015 [0.958]	1.050 0.049 [0.858]	0.979 -0.021 [0.936]
ilrank	0.267 -1.320 [0.056]*	0.576 -0.552 [0.454]	0.462 -0.772 [0.286]	0.295 -1.220 [0.089]*	0.373 -0.986 [0.199]	0.428 -0.849 [0.272]	0.470 -0.756 [0.308]	0.357 -1.030 [0.143]
cycle	0.740 -0.301 [0.593]	0.883 -0.124 [0.825]	1.140 0.131 [0.824]	1.186 0.171 [0.758]	1.069 0.067 [0.912]	1.163 0.151 [0.801]	1.064 0.062 [0.915]	0.812 -0.208 [0.726]
Gözlem Sayısı	338	338	338	338	338	338	338	338
Kuluçka İçinde iken Kapanan Firma Sayısı	20	20	20	20	20	20	20	20
Log-Likelihood	-70.753	-71.642	-70.866	-69.865	-67.907	-68.877	-71.153	-69.525
chi2	69.41	67.63	69.18	71.18	75.10	73.16	68.61	71.86
Prob>chi2	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
*: %10 önem düzeyinde anlamlılığı, **: %5 önem düzeyinde anlamlılığı ve ***: %1 önem düzeyinde anlamlılığı göstermektedir. (p değerleri köşeli parantez içinde verilmiştir.)								

Model 2’de yer alan (**entexp**) değişkeni anlamsız çıkmıştır. Dolayısıyla girişimci tecrübesi için (**workexp**) ve (**mangexp**) değişkenleri göz önüne alınabilir. Model 3’te Model 1’den farklı olarak (**ilrank**) değişkeni anlamlı çıkmıştır. Girişimcinin sermayesi ile ilgili değişkenler karşılaştırıldığında Model 5’te yer alan (**onlyselffin**) değişkeninin anlamlı olduğu, (**onlyloan**) değişkeninin ise Model 4 dahil olmak üzere hiçbir modelde anlamlı olmadığı görülmektedir. Ancak Model 5’ten farklı olarak Model 4’te (**burol**) ve (**ilrank**) değişkenlerinin anlamlıdır. Model 6, Model 7 ve Model 8’de ise değerlendirilen (**burol**), (**ortakhiz**) ve (**danismanlik**) değişkenlerinin her birinin de anlamlı olduğu görülmektedir. Ayrıca modellerdeki diğer değişkenlerin açıklama güçleri arasında da belirgin bir fark bulunmamaktadır.

Korelasyonları yüksek olan değişkenlerin modelde birlikte meydana getireceği etkileri de göz önünde bulundurmak amacıyla Model 9’dan Model 26’ya kadar değişkenlerin mümkün bütün kombinasyonları ile farklı modeller kurulmuştur. Aşağıdaki tabloda girişimci tecrübesini en iyi açıklayan ve modeldeki diğer bağımsız değişkenlerle en az ilişkili olan değişkenin (**workexp**) olduğu varsayıldığında ortaya çıkacak modeller özetlenmektedir.

Tablo 5.5: Hayatta kalmayı etkileyen girişimci tecrübesi (**workexp**) olarak alındığında, sermaye yoğunluğu, kuluçka hizmetlerini aynı anda ele alan Cox model tahminleri

Değişken	Model 9		Model 10		Model 11		Model 12		Model 13		Model 14	
	Tehlike Orantısı	Katsayı	Tehlike Orantısı	Katsayı	Tehlike Orantısı	Katsayı	Tehlike Orantısı	Katsayı	Tehlike Orantısı	Katsayı	Tehlike Orantısı	Katsayı
tekgelir	22.647	3.12***	24.245	3.19***	24.123	3.18***	16.615	2.81***	19.232	2.96***	22.426	3.11***
gender	0.907	-0.098	0.974	-0.027	0.902	-0.104	1.153	0.143	1.188	0.172	1.189	0.173
lnentage	17.044	2.84**	16.321	2.79**	18.835	2.94**	21.221	3.06**	29.918	3.4**	39.446	3.67***
enteduuni	0.785	-0.242	1.108	0.102	0.839	-0.175	0.978	-0.022	1.377	0.32	1.047	0.046
workexp	0.924	-0.079*	0.929	-0.074*	0.921	-0.082*	0.923	-0.08*	0.920	-0.084*	0.912	-0.092**
ailecev	0.935	-0.068	0.760	-0.274	0.844	-0.17	1.081	0.078	0.872	-0.137	0.983	-0.017
ortaksay	0.040	-3.23**	0.054	-2.92**	0.047	-3.05**	0.020	-3.93***	0.023	-3.79***	0.014	-4.25***
export	4.765	1.56	4.787	1.57	5.354	1.68	4.230	1.44	4.483	1.5	4.771	1.56
lnempini	1.122	0.115	1.106	0.101	1.072	0.070	1.193	0.176	1.211	0.191	1.172	0.159
onlyloan	0.606	-0.501	0.425	-0.855	0.565	-0.572	-	-	-	-	-	-
onlyselffin	-	-	-	-	-	-	3.572	1.27*	5.612	1.72***	5.310	1.67**
networking	0.043	-3.14***	0.048	-3.04***	0.040	-3.21***	0.027	-3.6***	0.027	-3.63***	0.018	-4.02***
yenilik	0.053	-2.94***	0.044	-3.12***	0.049	-3.02***	0.038	-3.27***	0.039	-3.25***	0.035	-3.36***
reklam	1.566	0.448	1.467	0.383	1.496	0.403	1.606	0.474	1.440	0.365	1.410	0.344
marka	4.538	1.51	4.380	1.48	3.468	1.24	5.872	1.77	6.515	1.87	5.557	1.71
buro	0.221	-1.51**	-	-	-	-	0.239	-1.43**	-	-	-	-
ortakhiz	-	-	0.385	-0.955	-	-	-	-	0.252	-1.38*	-	-
danismanlik	-	-	-	-	0.259	-1.35**	-	-	-	-	0.204	-1.59**
whenest	0.141	-1.96***	0.100	-2.31***	0.101	-2.3***	0.124	-2.09***	0.070	-2.67***	0.063	-2.76***
incubsize	0.937	-0.066***	0.934	-0.068***	0.934	-0.069***	0.937	-0.065***	0.934	-0.069***	0.929	-0.074***
sektor	1.289	0.254	1.376	0.319	1.567	0.449	1.508	0.411	1.329	0.284	1.538	0.43
compete	1.042	0.041	0.988	-0.012	0.938	-0.064	1.088	0.085	1.073	0.070	0.996	-0.004
ilrank	0.284	-1.26*	0.303	-1.19*	0.272	-1.3**	0.287	-1.25*	0.340	-1.08	0.261	-1.34**
cycle	0.892	-0.115	0.899	-0.107	0.785	-0.242	0.813	-0.207	0.829	-0.188	0.733	-0.311
Gözlem Sayısı	338		338		338		338		338		338	
Log-Likelihood	-73.769		-75.974		-74.515		-72.242		-73.579		-72.092	
chi2	63.37		58.96		61.88		66.43		63.76		66.73	
Prob>chi2	0.000***		0.000***		0.000***		0.000***		0.000***		0.000***	

*: %10 önem düzeyinde anlamlılığı, **: %5 önem düzeyinde anlamlılığı ve ***: %1 önem düzeyinde anlamlılığı göstermektedir.

Model 9, Model 10 ve Model 11’de (**onlyloan**) değişkeni anlamsız olarak bulunmuştur. (**ilrank**) değişkeninin ise sadece Model 13’te anlamsız olduğu görülmektedir. Model 12 ve Model 14’teki (**buro**) ve (**danismanlik**) değişkenleri arasında seçim yapabilmek için aşağıdaki modeller oluşturulmuştur. Bu modellerde girişimcilik tecrübesinin (**entexp**) değişkeni tarafından açıkladığı varsayılmaktadır;

Tablo 5.6: Hayatta kalmayı etkileyen girişimci tecrübesi (**entexp**) olarak alındığında, sermaye yoğunluğu, kuluçka hizmetlerini aynı anda ele alan Cox model tahminleri

	Model 15		Model 16		Model 17		Model 18		Model 19		Model 20	
Değişken	Tehlike Orantısı	Katsayı	Tehlike Orantısı	Katsayı	Tehlike Orantısı	Katsayı	Tehlike Orantısı	Katsayı	Tehlike Orantısı	Katsayı	Tehlike Orantısı	Katsayı
tekgelir	13.033	2.57***	12.877	2.56***	11.340	2.43***	10.934	2.39***	10.928	2.39***	10.576	2.36***
gender	0.840	-0.174	0.783	-0.244	0.850	-0.163	0.940	-0.062	0.786	-0.240	1.020	0.019
lnentage	23.931	3.18**	23.377	3.15**	20.115	3.00**	81.646	4.40**	93.013	4.53**	77.427	4.35**
enteduuni	1.399	0.336	1.641	0.495	1.389	0.329	1.870	0.626	1.996	0.691	1.777	0.575
entexp	0.938	-0.064	0.939	-0.063	0.935	-0.067	0.918	-0.086*	0.926	-0.077	0.922	-0.081
ailecev	1.137	0.128	1.146	0.136	1.250	0.223	1.304	0.265	1.304	0.265	1.470	0.385
ortaksay	0.039	-3.25**	0.062	-2.78**	0.057	-2.86**	0.015	-4.22**	0.027	-3.60**	0.018	-4.04**
export	4.304	1.460	4.139	1.420	4.572	1.520	4.007	1.390	4.441	1.490	5.651	1.730
lnempini	1.051	0.050	1.038	0.037	1.020	0.020	1.049	0.048	1.050	0.049	1.058	0.056
onlyloan	0.461	-0.774	0.325	-1.130	0.453	-0.791	-	-	-	-	-	-
onlyselffin	-	-	-	-	-	-	4.890	1.59*	6.448	1.86***	5.673	1.74**
networking	0.091	-2.39***	0.084	-2.47***	0.080	-2.52***	0.050	-3.00***	0.042	-3.16***	0.034	-3.37***
yenilik	0.109	-2.21**	0.083	-2.49**	0.094	-2.37**	0.094	-2.37**	0.088	-2.43**	0.073	-2.62**
reklam	1.424	0.354	1.659	0.507	1.479	0.391	1.143	0.134	1.278	0.245	1.118	0.112
marka	4.509	1.510	4.369	1.470	3.280	1.190	6.189	1.820	6.095	1.810	4.103	1.410
burol	0.229	-1.48**	-	-	-	-	0.191	-1.66***	-	-	-	-
ortakhiz	-	-	0.385	-0.954	-	-	-	-	0.199	-1.61**	-	-
danismanlik	-	-	-	-	0.294	-1.23**	-	-	-	-	0.208	-1.57**
whenest	0.112	-2.19***	0.100	-2.3***	0.105	-2.25***	0.074	-2.6***	0.057	-2.87***	0.055	-2.9***
incubsize	0.933	-0.07***	0.934	-0.068***	0.934	-0.069***	0.927	-0.076***	0.929	-0.073***	0.923	-0.081***
sektor	2.123	0.753	1.962	0.674	2.427	0.887	2.990	1.100	2.085	0.735	2.773	1.020
compete	0.998	-0.002	0.970	-0.030	0.943	-0.059	1.007	0.007	1.014	0.014	0.967	-0.034
ilrank	0.583	-0.539	0.569	-0.563	0.506	-0.681	0.603	-0.505	0.667	-0.404	0.483	-0.728
cycle	0.873	-0.135	0.825	-0.192	0.737	-0.305	0.925	-0.078	0.863	-0.148	0.782	-0.246
Gözlem Sayısı	338		338		338		338		338		338	
Kuluçka İçinde İken Kapanan Firma Sayısı	20		20		20		20		20		20	
Log-Likelihood	-74.682		-76.620		-75.373		-73.579		-74.076		-72.881	
chi2	61.55		57.67		60.17		66.18		62.76		65.15	
Prob>chi2	0.000***		0.000***		0.000***		0.000***		0.000***		0.000***	

*: %10 önem düzeyinde anlamlılığı, **: %5 önem düzeyinde anlamlılığı ve ***: %1 önem düzeyinde anlamlılığı göstermektedir.

Model 15, Model 16 ve Model 17’de yer alan (**onlyloan**) değişkeni tekrar anlamsız olarak bulunmuştur. (**entexp**) değişkeni sadece, (**onlyselffin**) ve (**burol**) değişkenlerini içeren

Model 18’de anlamlıdır. Son olarak girişimci tecrübesi için (**mangexp**) değişkenini kabul eden modeller aşağıdaki gibi oluşturulmuştur;

Tablo 5.7: Hayatta kalmayı etkileyen girişimci tecrübesi (**mangexp**) olarak alındığında, sermaye yoğunluğu, kuluçka hizmetlerini aynı anda ele alan Cox model tahminleri

Değişken	Model 21		Model 22		Model 23		Model 24		Model 25		Model 26	
	Tehlike Orantısı	Katsayı	Tehlike Orantısı	Katsayı	Tehlike Orantısı	Katsayı	Tehlike Orantısı	Katsayı	Tehlike Orantısı	Katsayı	Tehlike Orantısı	Katsayı
tekgelir	23.870	3.17***	20.620	3.03***	17.360	2.85***	17.802	2.88***	15.066	2.71***	14.469	2.67***
gender	0.671	-0.398	0.634	-0.455	0.709	-0.343	0.779	-0.250	0.693	-0.366	0.894	-0.113
lnentage	62.497	4.14**	48.022	3.87**	37.777	3.63**	148.383	5.00***	131.065	4.88***	100.698	4.61***
enteduuni	1.413	0.345	1.568	0.450	1.357	0.305	1.676	0.517	1.826	0.602	1.605	0.473
mangexp	0.905	-0.099**	0.914	-0.090**	0.912	-0.092*	0.902	-0.103**	0.915	-0.088*	0.914	-0.090*
ailecev	1.043	0.043	1.111	0.105	1.181	0.167	1.154	0.144	1.202	0.184	1.307	0.268
ortaksay	0.024	-3.73**	0.050	-3.00**	0.046	-3.08**	0.012	-4.39***	0.027	-3.63**	0.017	-4.07**
export	5.330	1.670	5.015	1.610	5.118	1.630	5.394	1.690	5.736	1.750	6.323	1.840
lnempini	0.932	-0.071	0.934	-0.069	0.908	-0.097	0.968	-0.033	0.984	-0.016	0.976	-0.024
onlyloan	0.416	-0.876	0.292	-1.230	0.425	-0.856	-	-	-	-	-	-
onlyselffin	-	-	-	-	-	-	4.103	1.41**	5.592	1.72**	4.777	1.56**
networking	0.063	-2.77***	0.060	-2.82***	0.054	-2.91***	0.037	-3.29***	0.034	-3.37***	0.026	-3.64***
yenilik	0.090	-2.41**	0.071	-2.65***	0.075	-2.59**	0.080	-2.53**	0.078	-2.56**	0.062	-2.79**
reklam	1.222	0.200	1.665	0.510	1.376	0.319	1.040	0.039	1.281	0.248	1.101	0.096
marka	6.523	1.880	5.483	1.700	4.327	1.460	7.984	2.080	7.031	1.950	5.425	1.690
buro	0.194	-1.64**	-	-	-	-	0.167	-1.79***	-	-	-	-
ortakhiz	-	-	0.333	-1.100	-	-	-	-	0.191	-1.66**	-	-
danismanlik	-	-	-	-	0.276	-1.29**	-	-	-	-	0.203	-1.59**
whenest	0.094	-2.37***	0.091	-2.39***	0.093	-2.37***	0.072	-2.63***	0.060	-2.81***	0.058	-2.85***
incubsize	0.927	-0.075***	0.932	-0.070***	0.930	-0.078***	0.922	-0.081***	0.927	-0.076***	0.921	-0.083***
sektor	2.028	0.707	1.879	0.631	2.291	0.829	2.607	0.958	1.901	0.643	2.524	0.926
compete	1.007	0.007	0.999	-0.001	0.946	-0.056	0.987	-0.013	1.003	0.003	0.939	-0.063
ilrank	0.469	-0.758	0.469	-0.758	0.393	-0.935	0.477	-0.740	0.555	-0.588	0.385	-0.954
cycle	1.077	0.074	0.948	-0.053	0.855	-0.157	1.195	0.178	1.030	0.029	0.923	-0.080
Gözlem Sayısı	338		338		338		338		338		338	
Kuluçka İçinde İken Kapanan Firma Sayısı	20		20		20		20		20		20	
Log-Likelihood	-73.132		-75.402		-74.227		-71.397		-73.456		-72.333	
chi2	64.65		60.11		62.46		68.12		64.00		66.25	
Prob>chi2	0.000***		0.000***		0.000***		0.000***		0.000***		0.000***	

*: %10 önem düzeyinde anlamlılığı, **: %5 önem düzeyinde anlamlılığı ve ***: %1 önem düzeyinde anlamlılığı göstermektedir.

(**mangexp**) deęişkenini içeren Model 24, Model 25 ve Model 26 modelleri ile (**workexp**) deęişkenini içeren Model 12 karşılaştırıldığında, dięerlerinden farklı olarak Model 12’de (**ilrank**) deęişkeninin anlamlı olduęu görülmektedir. Açıklama gücü daha yüksek olduğundan (**workexp**), (**onlyselffin**) ve (**burol**) deęişkenlerine yer veren Model 12 tercih edilmektedir.

Tehlike orantılarının 1’den büyük olması deęişkenin 1 kategorisi özelliğini taşıyan grubun 0 kategorisi grubuna göre daha yüksek olasılıkla ilgilenilen olaya maruz kalacağını göstermektedir. Eęer tehlike oranıtısı 1 ise iki grubun kapanma şansının eşit olduğunu, eęer 0 ile 1 deęerleri arasında ise 0 kategori deęeri alan grubun 1 kategori deęerli olana kıyasla kapanma olasılığının daha düşük olduğunu ifade etmektedir. Buna göre tercih edilen Model 12’deki katsayılar şöyle yorumlanabilir;

Model 12’de (**gender**), (**enteduuni**), (**ailecev**), (**export**), (**reklam**), (**marka**), (**sektor**), (**compete**), (**cycle**) deęişkenleri anlamsız bulunmuştur. Yani girişimcinin cinsiyeti, eğitim durumu, aile çevresinden etkilenmesi; firmanın ihracat ve reklam yapıp yapmaması, marka sahibi olup olmaması, hangi sektörde bulunduęu; sektördeki rekabet yoğunluğu ve firmanın kriz atlatıp atlatmaması hayatta kalma süresi üzerinde etkili deęildir.

Literatüre bakıldığında genel olarak eğitim seviyesinin kapanma üzerine olan azaltıcı etkilerinin yüksek bilgi düzeyi ve uzmanlık gerektiren sektörlerde öne çıktığı anlaşılmaktadır. Yüksek bilgi düzeyi ve uzmanlık gerektirmeyen sektörlerde ve karmaşık iş yükü olmayan küçük ölçekli firmalarda ise üniversite mezunu olup olmamanın firmaların hayatta kalma yeteneğine katkısı her zaman ortaya çıkmamaktadır. Bunun yerine, sektöre has iş tecrübesi ve genel olarak edinilmiş olan yöneticilik tecrübesinin bu firmalarda daha önemli olduğu ancak bunların genelde üniversitelerde verilen eğitimle deęil, bunun yerine geçmişteki iş tecrübeleri ile geliştięi bilinmektedir. Sonuç olarak, girişimcinin üniversitede almış olduğu formel eğitimin, sahip olduğu firmanın kuluçka içerisinde iken hayatta kalmasına bir etkisinin olmadığını göstermiştir (Karaöz ve Albeni, 2011, s.29).

Yapılan çalışmalarda girişimcinin iş tecrübesinin, kazanılan bilgi, beceri ve deneyimlerinin firmanın performansı ve hayatta kalması üzerine pozitif bir etkisi bulunduğu sonucuna ulaşılmıştır. Halen keşfedilmemiş karlı piyasa fırsatlarını yakalama ve piyasa koşullarıyla ilgili uygun bilgilere kolaylıkla ulaşma imkanına sahip girişimciler, bu avantajları geçmiş tecrübelerine dayanarak elde etmektedir (Demirgil, 2008).

Geliri sadece mevcut firmadan olan girişimciye sahip firmaların başarısızlık ihtimali, dięer firmaların yaklaşık 17 katıdır. Bu durumda ek gelire sahip girişimcilerin firmaların daha başarılı olduğu anlaşılmaktadır. Girişimcinin yaşı arttıkça firmanın hayatta kalma ihtimalinin

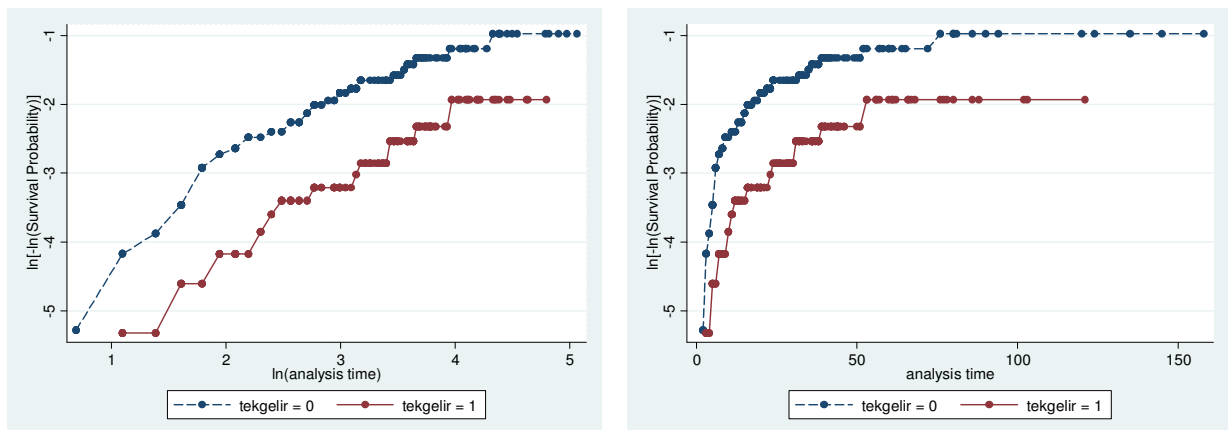
azaldığı görülmektedir. Girişimcinin çalışma tecrübesi ve firmadaki ortak sayısı arttıkça, firmaların başarısızlık ihtimalinin azaldığı söylenebilmektedir. Girişimin kuruluşundaki sermaye tamamıyla kendisine ait olan girişimcilerin firmalarının başarısızlık riski, diğerlerinin yaklaşık 3,5 katıdır. Girişimci kuluçka içinde ve dışında paydaşlarla işbirliği içerisinde ise, bu firmanın hayatta kalma ihtimali diğerlerinin yaklaşık 37 katıdır. Yine firmanın yenilik yapması hayatta kalma ihtimalinin, yenilik yapmayan firmalardan yaklaşık 26 kat daha yüksek olduğunu görülmektedir. Kuluçka hizmetlerinden yararlanan firmaların hayatta kalma ihtimali, diğer firmaların yaklaşık 4 katıdır. Kuluçkanın kurulduğu ilin son 36 ay içinde kurulmuş olması firmaların hayatta kalma ihtimalini diğerlerinin yaklaşık 8 katı kadar olmasını sağlamaktadır. Kuluçkadaki işlik sayısı ve kuluçkanın bulunduğu ilin KBGSMH'sının payı arttıkça başarısızlık riskinin azaldığı görülmektedir.

5.4. Modelde Yer Alan Değişkenlerin PH Varsayımının Denetlenmesi

5.4.1. Grafikselsel Yaklaşım 1: Log-Log KM Hayatta Kalma Eğrileri ile PH Varsayımının Denetlenmesi

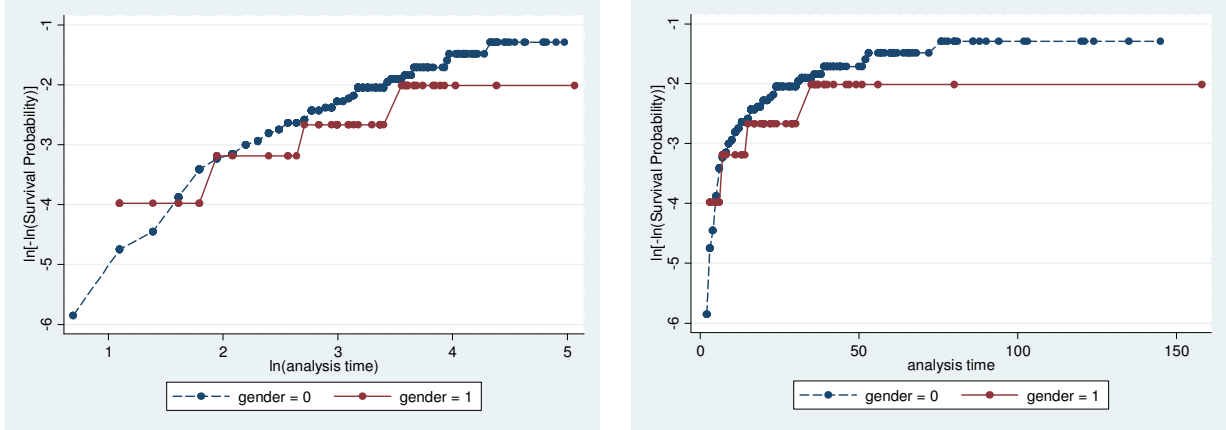
PH varsayımını test etmek için öncelikle grafikselsel yaklaşımla log-log KM hayatta kalma eğrileri değerlendirilmelidir. PH varsayımının sağlanması için her bir kategori için log-log hayatta kalma olasılığı çizimlerinin birbirine paralel olması gerekmektedir. Model 12'de yer alan kategorik değişkenler için log-log KM hayatta kalma çizimleri aşağıdaki grafiklerde verilmiştir. Her bir değişken için soldaki grafikte yatay ekseninde zamanın logaritması, sağdaki grafikte yatay ekseninde zamanın kendisi yer almaktadır. Bu iki tür çizim için de paralellik koşulu bulunmaktadır. Ancak iki eğrinin paralel olmadığına dair güçlü bir kanıt olmadıkça, PH varsayımının sağlanmadığı sonucuna varılmamalıdır.

Şekil 5.3: (tekgelir) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi



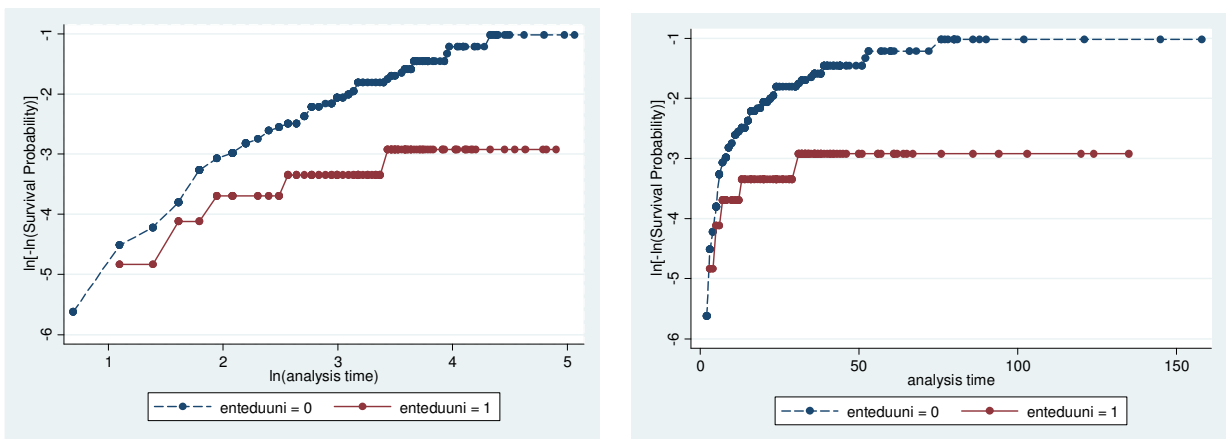
(**tekgelir**) değişkeni her iki grafikte de paralel olduğundan PH varsayımını sağlamaktadır. Yani girişimcinin geliri sadece genç firmadan olup olmadığını gösteren (**tekgelir**) değişkeninin, firmanın hayatta kalması üzerindeki etkisi zaman içinde sabittir.

Şekil 5.4: (**gender**) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi



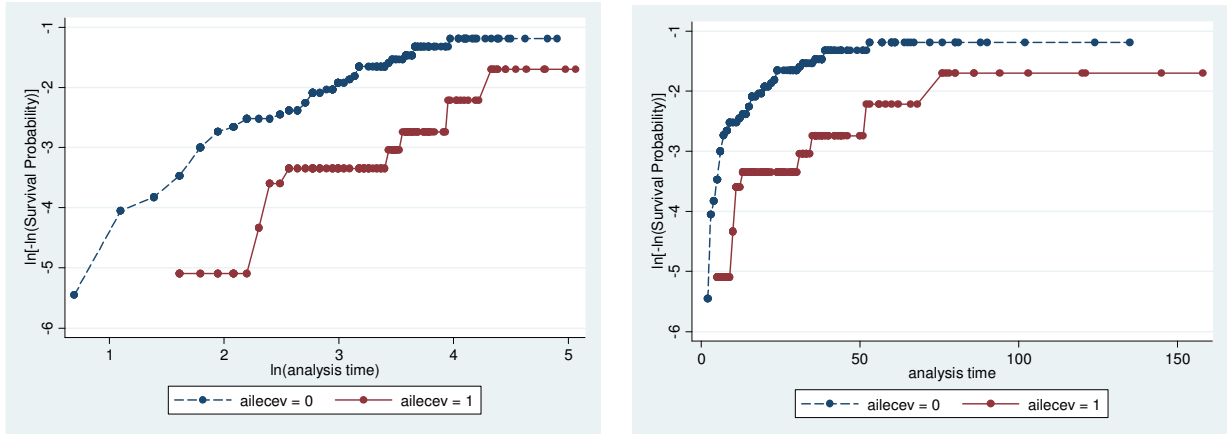
(**gender**) değişkeni iki grafikte de paralel değildir ve 1 değeri verilen kadın girişimcilerin hayatta kalma eğrisi, erkek girişimcilerin hayatta kalma eğrisi ile bazı noktalarda kesişmektedir. Bu değişken PH varsayımını sağlamadığından, cinsiyetin firmaların hayatta kalması üzerindeki etkisinin zaman içinde sabit olmadığı söylenebilir.

Şekil 5.5: (**enteduuni**) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi



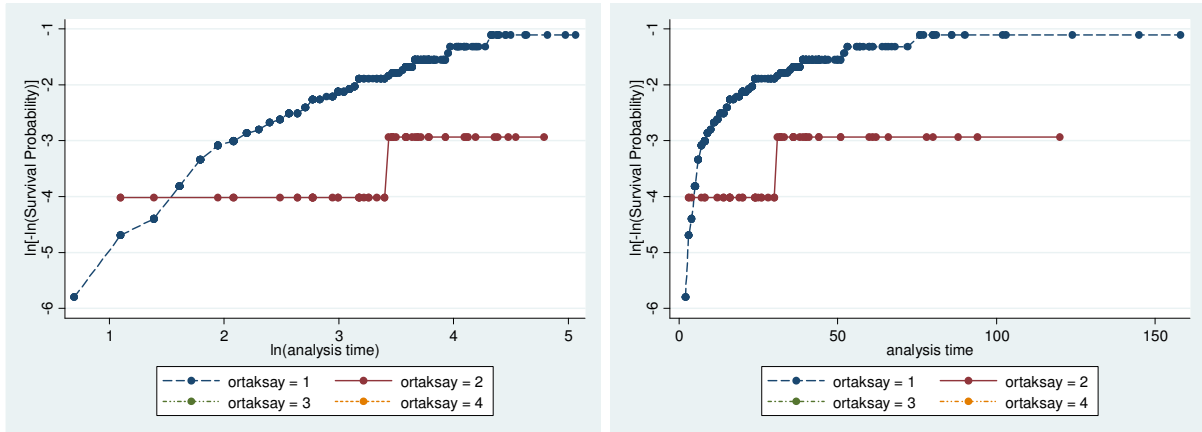
(**enteduuni**) değişkeni iki grafikte de birbirinden uzaklaşmaktadır yani eğriler paralel değildir. Bu değişken PH varsayımını sağlamadığından, girişimcinin eğitim düzeyinin firmaların hayatta kalma süresi üzerindeki etkisi zamanla değişmektedir.

Şekil 5.6: (ailecev) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi



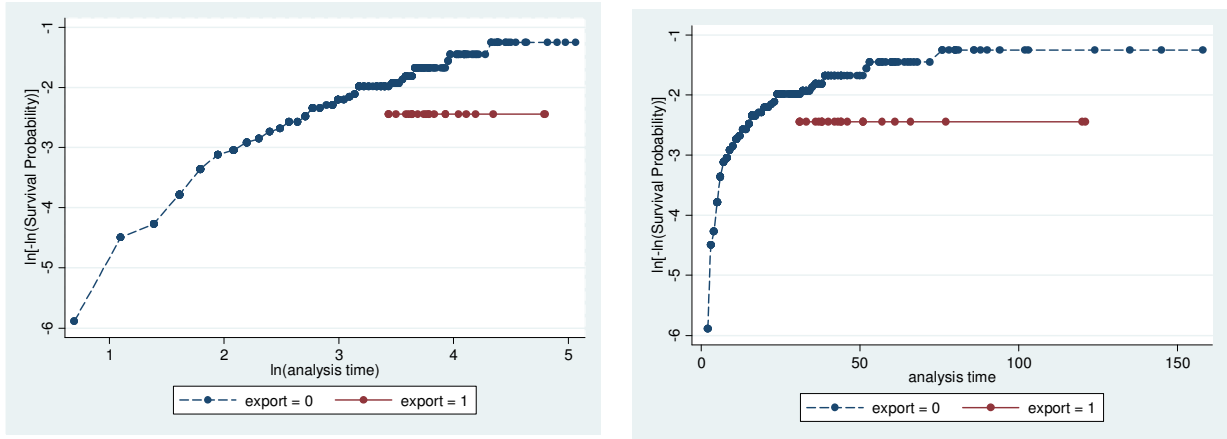
(ailecev) değişkeni iki grafikte de birbirine yaklaşımaktadır yani eğriler paralel değildir. Bu değişken PH varsayımını sağlamadığından girişimcinin aile ve/veya çevresinden birinden etkilenip etkilenmemesinin, firmanın hayatta kalması üzerindeki etkisi sabit değildir.

Şekil 5.7: (ortaksay) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi



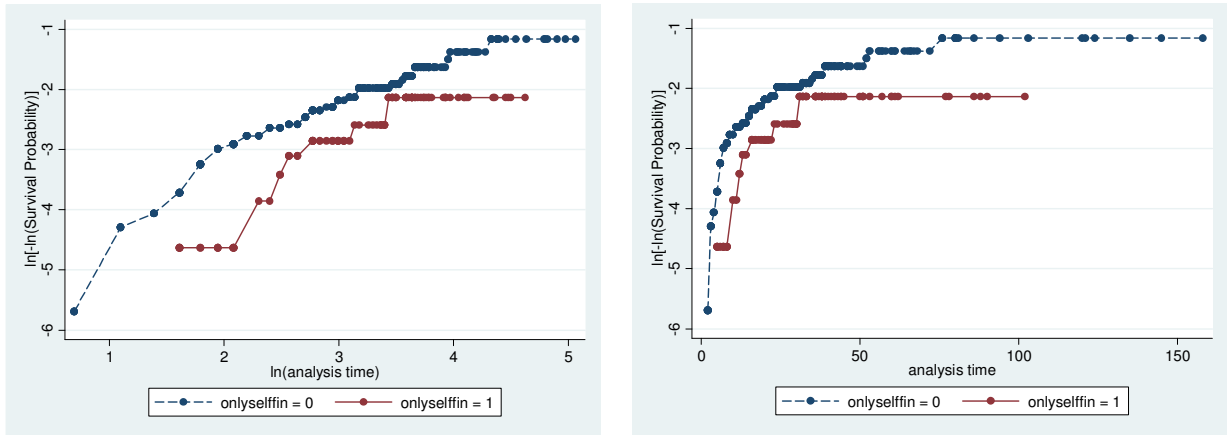
(ortaksay) değişkeninin ilk iki kategorisi iki grafikte de birbiriyle kesişmektedir ve çizimler paralel değildir. Ortak sayısının, firmanın hayatta kalmasına olan etkisi için PH varsayımı ihlal edilmektedir.

Şekil 5.8: (**export**) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi



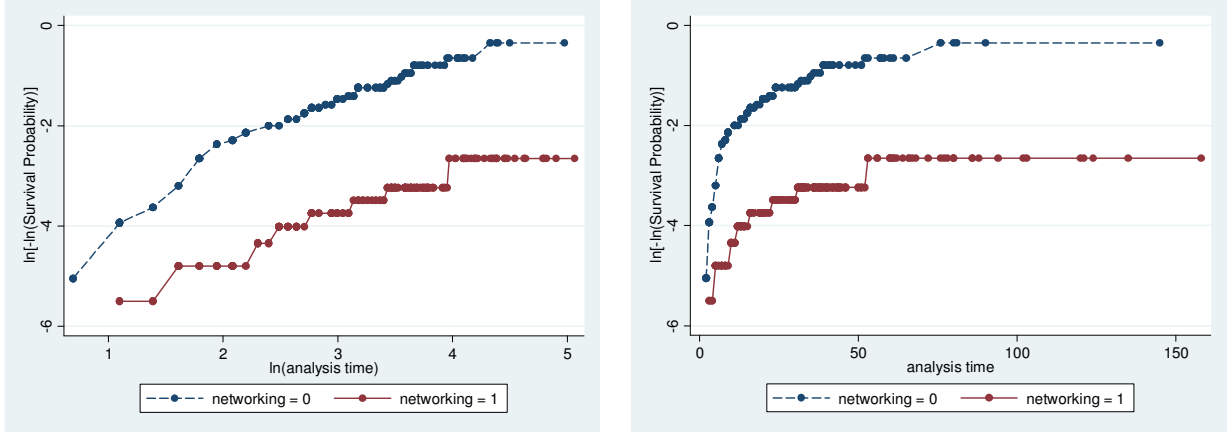
(**export**) değişkeninin log-log KM hayatta kalma çizimleri paralel değildir. Yani (**export**) değişkeni de PH varsayımını sağlamamaktadır ve hayatta kalma üzerine olan etkisi sabit değildir.

Şekil 5.9: (**onlyselffin**) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi



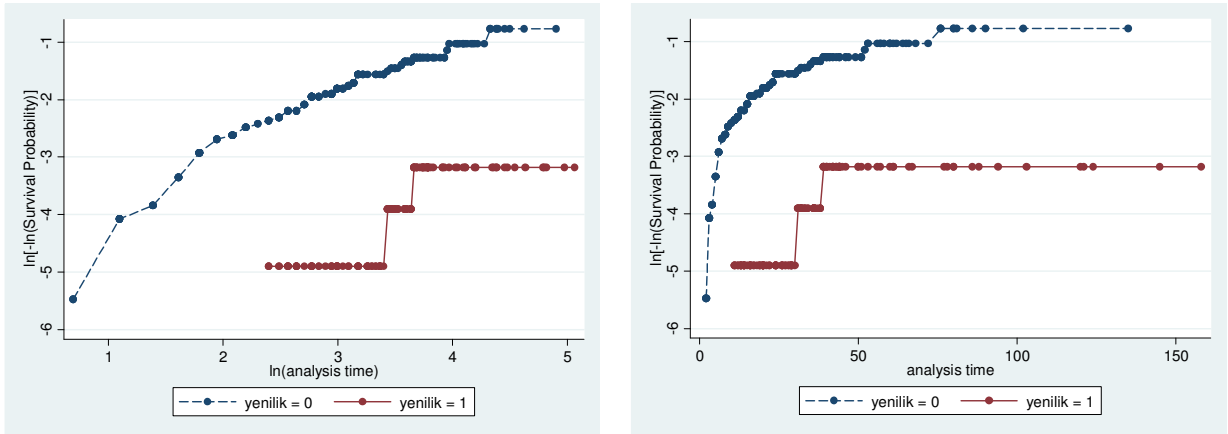
(**onlyselffin**) değişkeni için log-log KM hayatta kalma çizimlerinin kabaca paralel olduğu söylenebilir. Yani (**onlyselffin**) değişkeni PH varsayımını sağlamamaktadır. Buna göre firmanın kuruluş aşamasındaki sermaye yapısının, firmanın hayatta kalma ihtimali üzerine etkisi zamana göre sabittir.

Şekil 5.10: (**networking**) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi



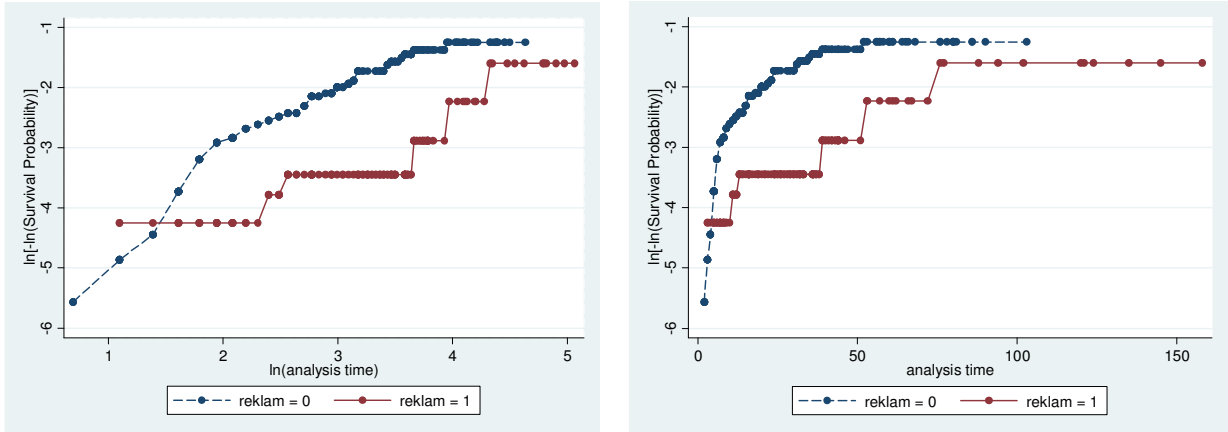
(**networking**) değişkeni için hayatta kalma eğrilerinin paralel olduğu görülmektedir. Bu yüzden, PH varsayımı sağlandığı söylenebilir. Girişimcinin kuluçka içinde ve dışında paydaşlarıyla işbirliği içinde olup olmaması, firmaların hayatta kalma ihtimali üzerinde zamana göre sabit bir etkiye sahiptir.

Şekil 5.11: (**yenilik**) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi



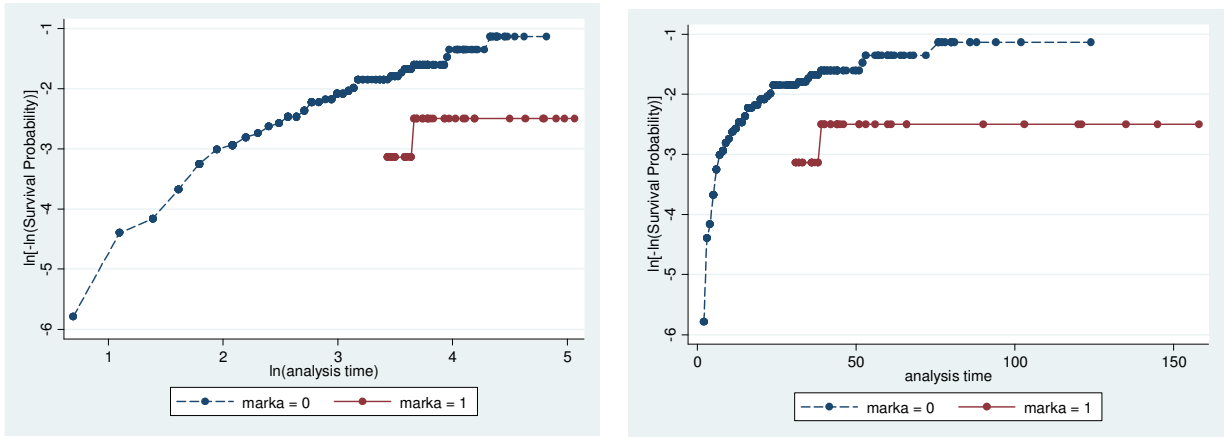
(**yenilik**) değişkeni için hayatta kalma eğrileri paralel değildir. Bu yüzden, PH varsayımı ihlal edilmektedir. Girişimcinin firmada yenilik yapıp yapmamasının; firmaların hayatta kalma ihtimali üzerindeki etkisi zaman içinde sabit değildir.

Şekil 5.12: (reklam) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi



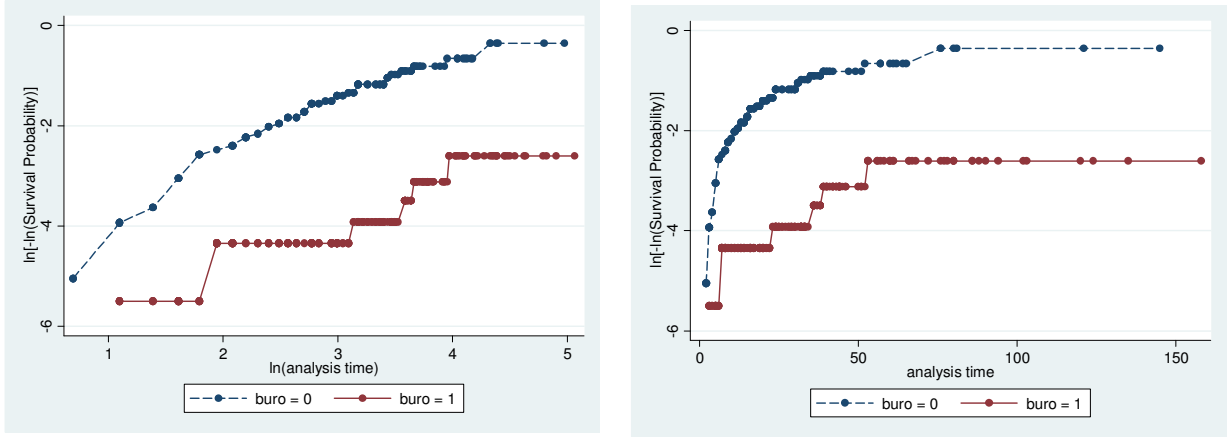
(reklam) değişkeni için hayatta kalma eğrileri başlangıçta bir noktada kesişmektedir ve aralarındaki uzaklık önce artıp sonra azalmaktadır. Bu yüzden, PH varsayımı ihlal edilmektedir. Firmanın reklam yapıp yapmamasının, firmaların hayatta kalma ihtimali üzerindeki etkisi zaman içinde sabit değildir.

Şekil 5.13: (marka) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi



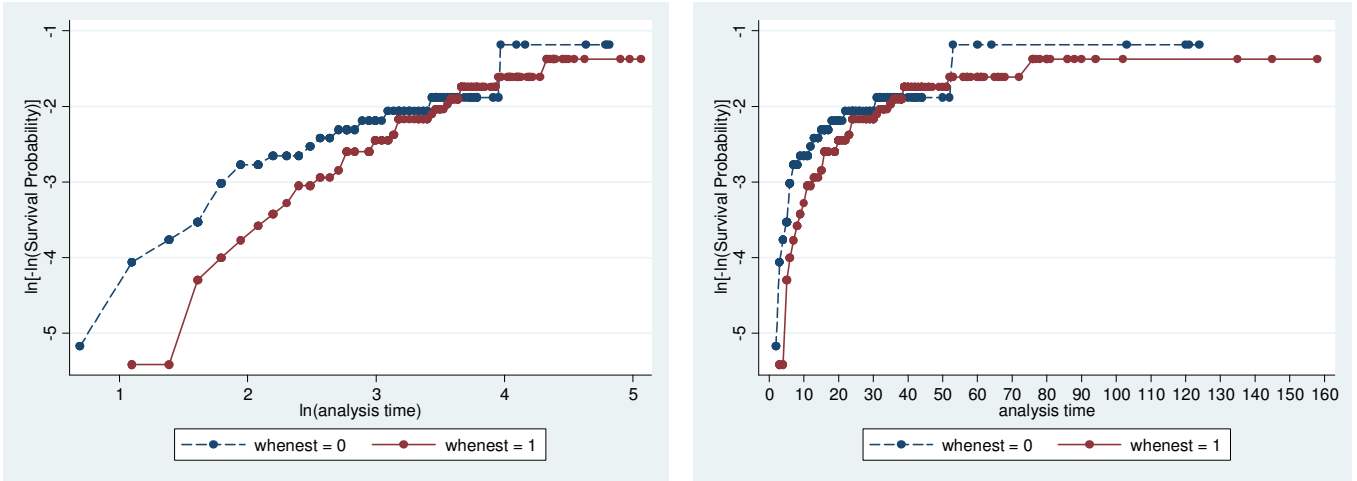
(marka) değişkeninin hayatta kalma eğrileri paralel değildir. Bu yüzden, PH varsayımı ihlal edilmektedir. Firmanın marka sahibi olup olmasının; hayatta kalma ihtimali üzerine etkisi zaman içinde değişmektedir.

Şekil 5.14: (**bu**) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi



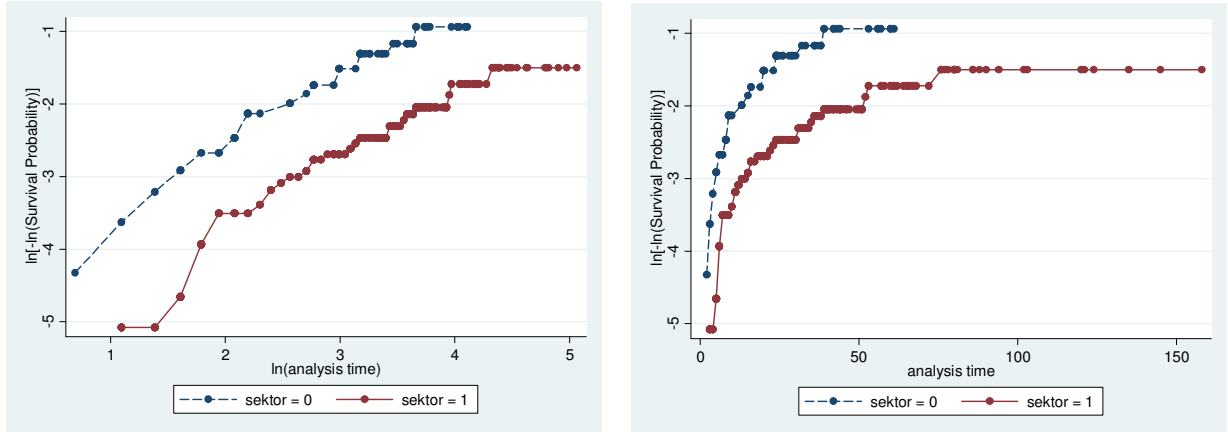
(**bu**) değişkeni için hayatta kalma eğrilerinin kabaca paralel olduğu söylenebilir. Bu değişken, PH varsayımını sağlamaktadır. Girişimcinin kuluçka tarafından sunulan büro hizmetlerinden yararlanıp yararlanmamasının, firmaların hayatta kalma ihtimali üzerine zaman içinde sabit bir etkisi vardır.

Şekil 5.15: (**whonest**) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi



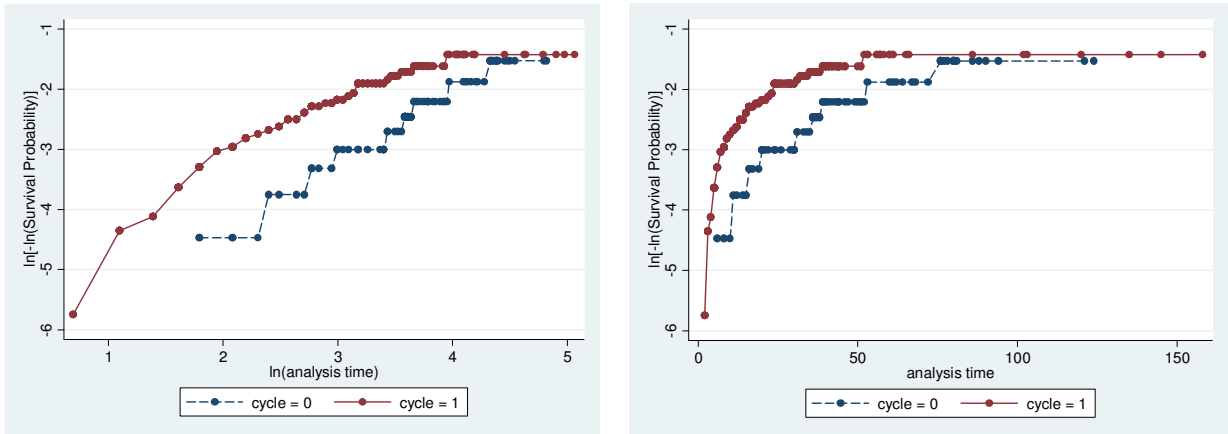
(**whonest**) değişkeninin hayatta kalma eğrileri iki noktada kesişmekte ve başlangıçta log-log hayatta kalma ihtimalleri, 1 değerini taşıyan gözlemler için yüksek iken yaklaşık 40. ayda 0 değeri atan gözlemler için daha yüksek olmakta ve yaklaşık 50. ayda tekrar ilk haline dönmektedir. Bu değişken, PH varsayımını açıkça ihlal etmektedir. Firmanın, kuluçkanın kurulduğu ilk 36 ay içerisinde kurulmuş olup olmamasının, hayatta kalma ihtimali üzerine zaman içinde sabit olmayan bir etkisi vardır.

Şekil 5.16: (**sektor**) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi



(**sektor**) değişkeninin hayatta kalma eğrileri paraleldir, yani PH varsayımını sağlamaktadır. Firmanın bulunduğu sektörün; hayatta kalma ihtimali üzerine etkisi zaman içinde sabittir.

Şekil 5.17: (**cycle**) değişkeni için log-log KM hayatta kalma grafikleri ile PH varsayımının denetlenmesi



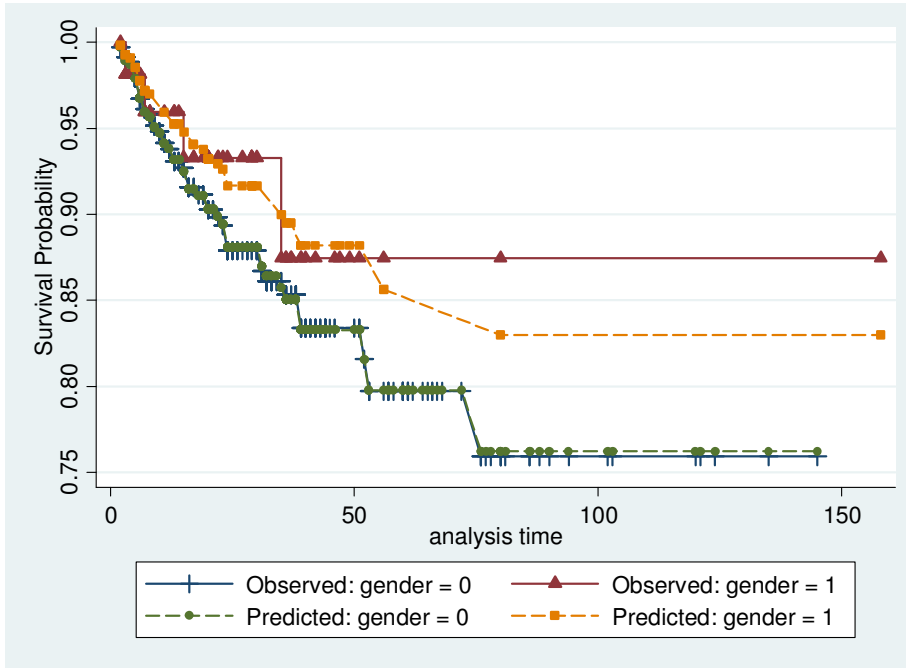
(**cycle**) değişkeninin hayatta kalma eğrileri zamanla birbirine yaklaşmaktadır, yani PH varsayımı ihlal edilmektedir. Firmanın bir kriz zamanı tecrübe edip etmemesinin; hayatta kalma ihtimali üzerine etkisi zaman içinde sabit değildir.

Log-log KM hayatta kalma eğrileri değerlendirildiğinde; (**gender**), (**enteduuni**), (**ailecev**), (**ortaksay**), (**export**), (**yenilik**), (**reklam**), (**marka**), (**whenest**), (**cycle**) değişkenlerinin PH varsayımını ihlal ettiği sonucuna ulaşılmaktadır.

5.4.2. Grafiksel Yaklaşım 2: Gözlenene Karşı Beklenen Hayatta Kalma Eğrileri ile PH Varsayımının Denetlenmesi

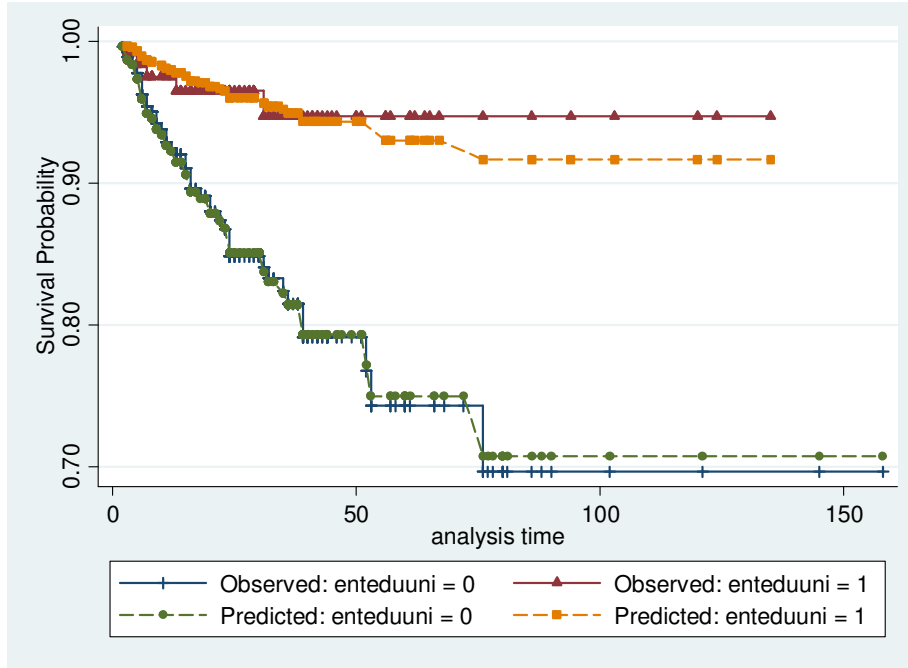
Bu değişkenler, modelde yer alan diğer bağımsız değişkenler için ayarlanarak tekrar çizim yapılırsa ayarlanmış Cox hayatta kalma eğrileri elde edilir. Gözlenene karşı beklenen hayatta kalma olasılıklarını karşılaştıran bu yöntemde, log-log KM hayatta kalma eğrileri Cox ayarlaması tarafından paralel olmaya zorlanmaktadır. Bu yöntemde PH varsayımının sağlanması için ayarlanmış Cox hayatta kalma eğrileri birbirine yakın olmalıdır. PH varsayımını ihlal ettiğinden şüphelenilen değişkenler için ayarlanmış Cox hayatta kalma tahminleri ile log-log KM hayatta kalma tahminleri aynı grafikler üzerinde aşağıda gösterilmektedir.

Şekil 5.18: (**gender**) değişkeni için gözlenene karşı beklenen hayatta kalma olasılıkları ile PH varsayımının değerlendirilmesi



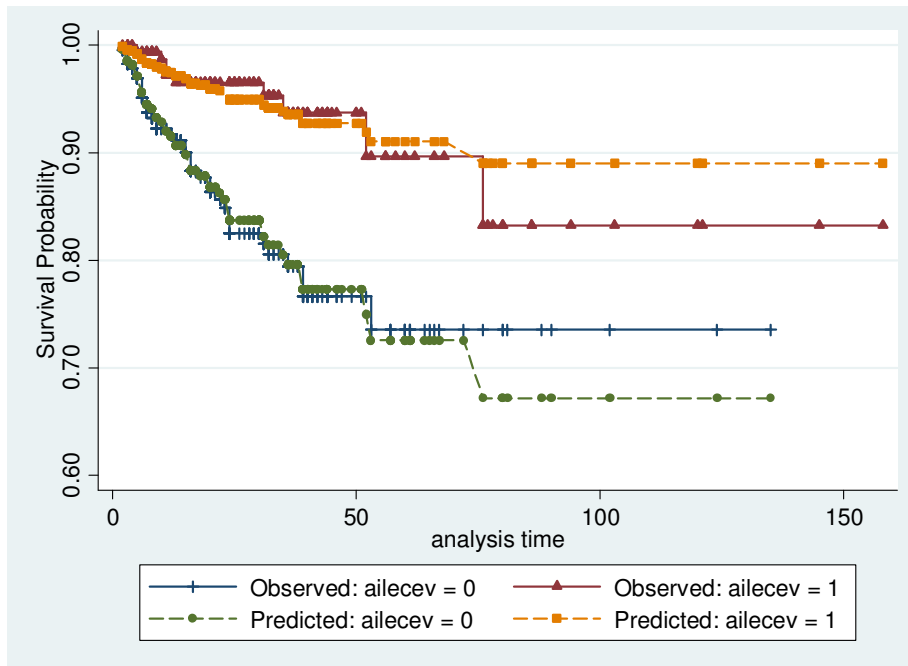
(**gender**) değişkenininin 0 ile gösterilen erkek girişimciler için gözlenen ve beklenen çizimleri birbirine çok yakın çıkmış ancak 1 ile gösterilen kadın girişimciler için ise birbirine yakın çıkmamıştır. Bu nedenle bu değişken hala PH varsayımını sağlamamaktadır.

Şekil 5.19: (**enteduuni**) değişkeni için gözlenene karşı beklenen hayatta kalma olasılıkları ile PH varsayımının değerlendirilmesi



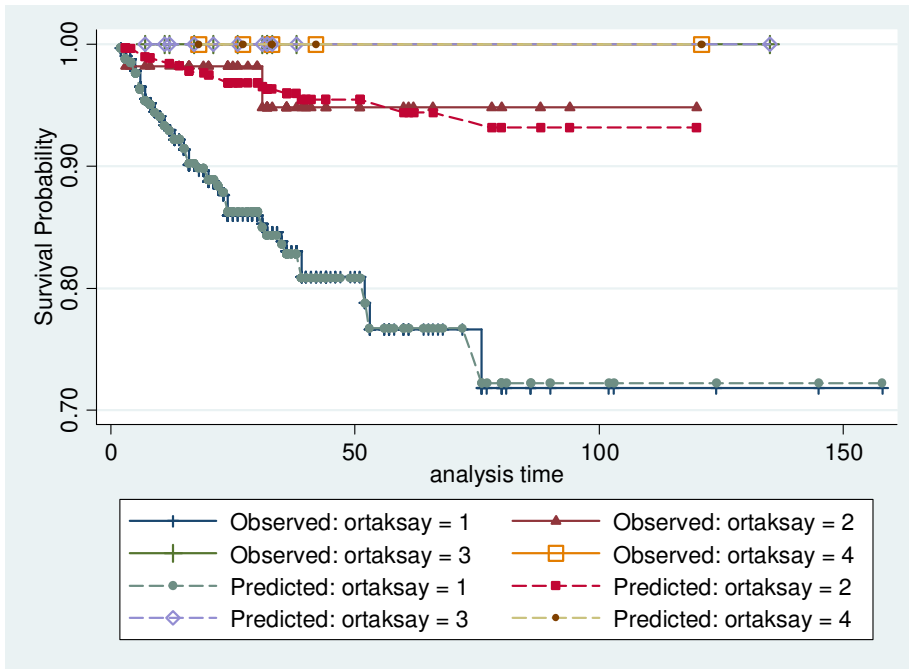
(**enteduuni**) değişkeninde de 1 kategorisi için hayatta kalma eğrileri birbirinden uzaklaşmaktadır yani eğriler birbirine yakın değildir. Bu değişken PH varsayımını sağlamamaktadır.

Şekil 5.20: (**aillecev**) değişkeni için gözlenene karşı beklenen hayatta kalma olasılıkları ile PH varsayımının değerlendirilmesi



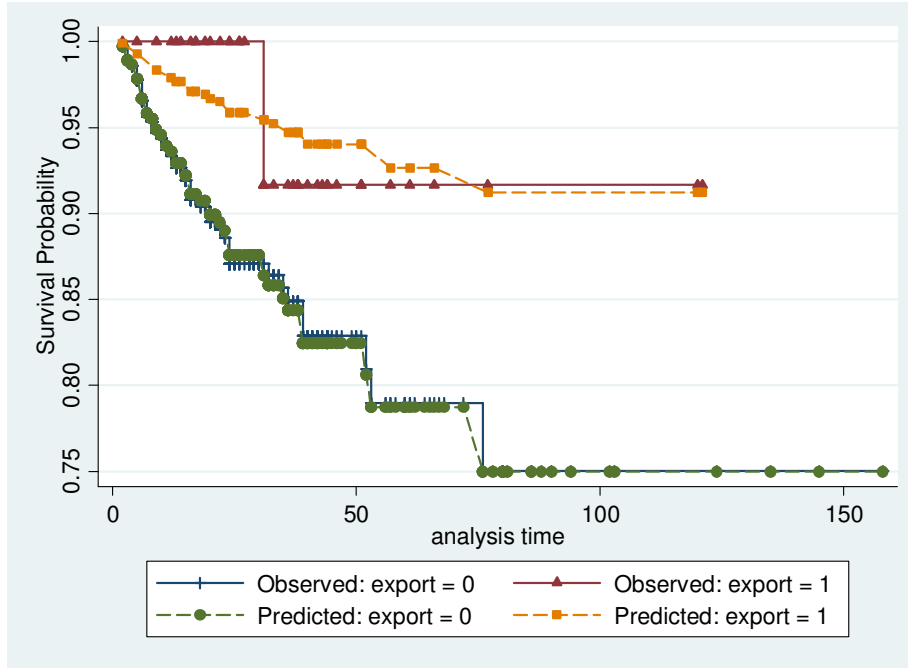
(**aillecev**) deęiřkeni iin her iki kategorinin de gzlenen ve beklenen hayatta kalma eęrileri birbirine yakın deęildir. Grafikselsel yntemle bu deęiřkenin de PH varsayımını saęlamadıęı sylenebilir.

řekil 5.21: (**ortaksay**) deęiřkeni iin gzlenene karřı beklenen hayatta kalma olasılıkları ile PH varsayımının deęerlendirilmesi



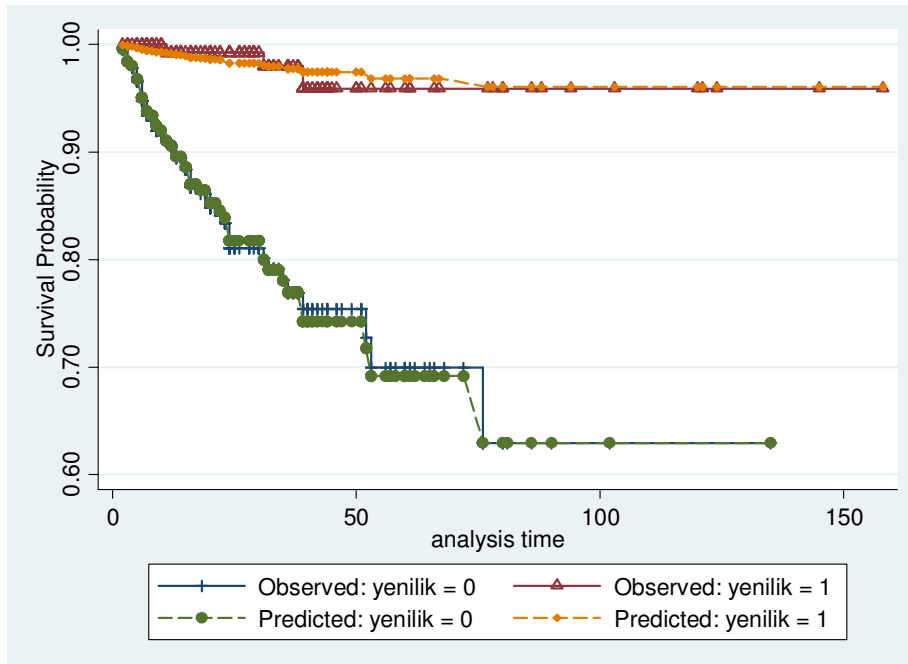
(**ortaksay**) deęiřkeninin 1,3 ve 4. kategoriler iin gzlenen ve beklenen hayatta kalma olasılıęı eęrileri neredeyse tamamen rtřmektedir. İkinci kategorinin gzlenen ve beklenen hayatta kalma olasılıęı eęrileri kesiřmesine raęmen birbirine yakın olduęundan PH varsayımının saęlandıęı sonucuna ulařılabilir.

Şekil 5.22: (**export**) değişkeni için gözlenene karşı beklenen hayatta kalma olasılıkları ile PH varsayımının değerlendirilmesi



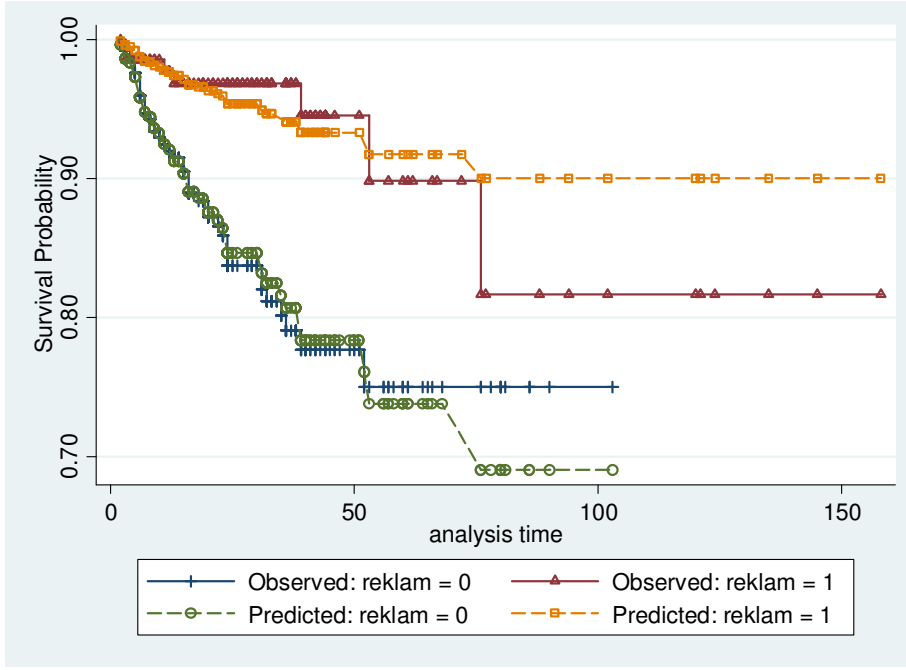
(**export**) değişkeninin ayarlanmış Cox hayatta kalma olasılığı tahminleri ihracat yapan (**export** = 1) firmalar için birbirine yakın değildir. Yani (**export**) değişkeni de PH varsayımını sağlamamaktadır.

Şekil 5.23: (**yenilik**) değişkeni için gözlenene karşı beklenen hayatta kalma olasılıkları ile PH varsayımının değerlendirilmesi



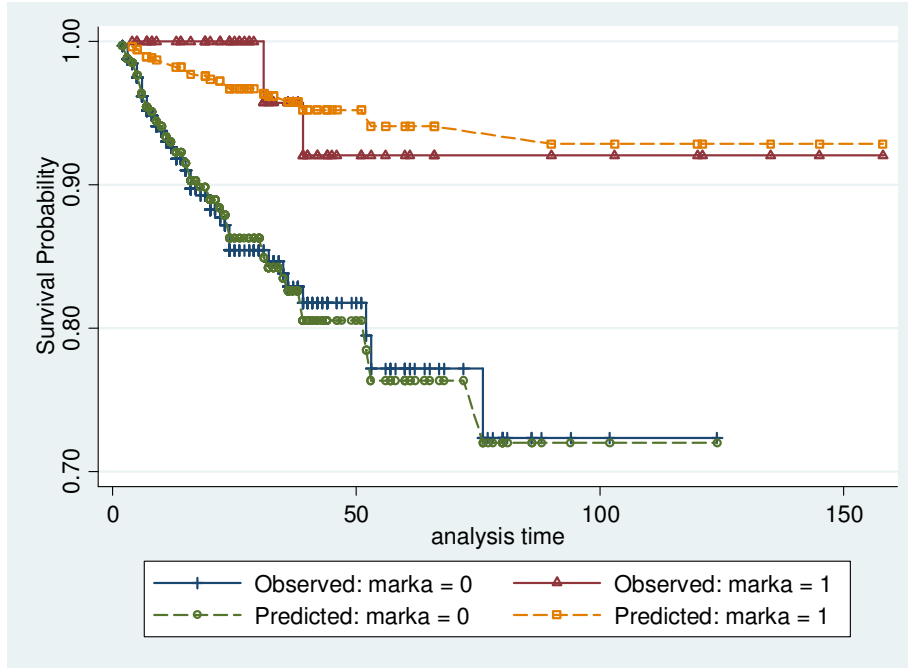
(yenilik) değişkeni için hayatta kalma çizimleri birbirine yakındır. Ayarlanmış Cox hayatta kalma tahminleri, PH varsayımı sağlanmaktadır.

Şekil 5.24: (reklam) değişkeni için gözlenene karşı beklenen hayatta kalma olasılıkları ile PH varsayımının değerlendirilmesi



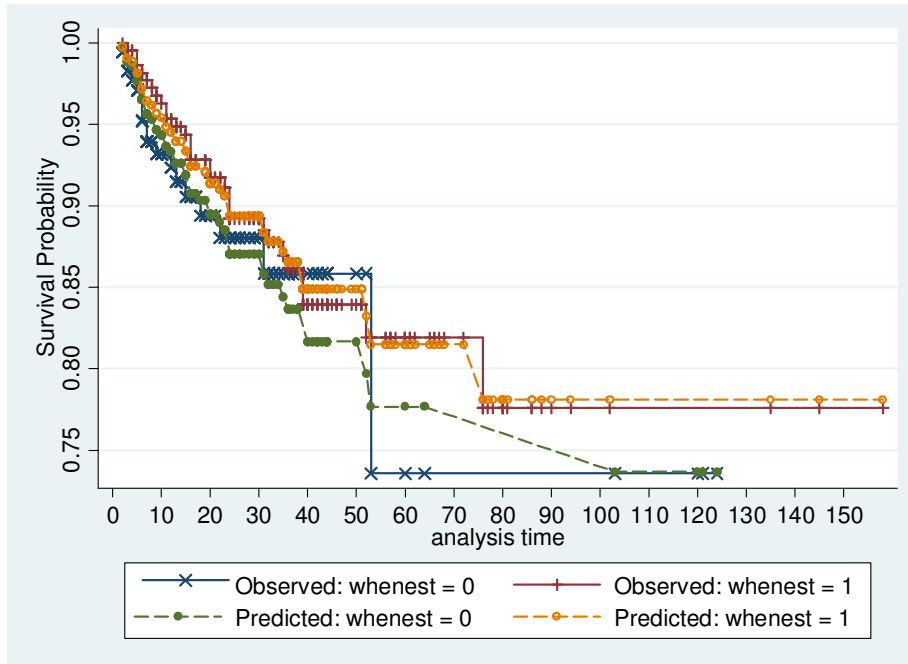
(reklam) değişkeni için gözlenen ve beklenen hayatta kalma eğrileri başlangıçta yakın olmasına rağmen yaklaşık 75. aydan sonra aralarındaki mesafe oldukça artmaktadır. Bu yüzden, PH varsayımı ihlal edilmektedir.

Şekil 5.25: (**marka**) değişkeni için gözlenene karşı beklenen hayatta kalma olasılıkları ile PH varsayımının değerlendirilmesi



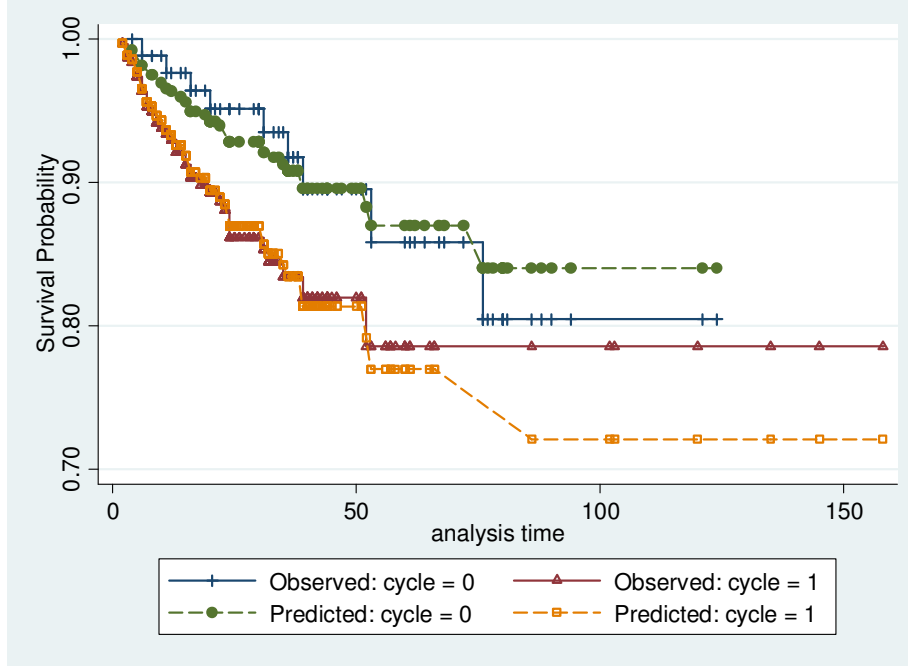
(**marka**) değişkeninde 1 kategorisine dahil olan gözlemlerin (marka sahibi olan firmalar) hayatta kalma eğrileri birbirine yakındır. Bu yüzden, PH varsayımı ihlal edilmektedir.

Şekil 5.26: (**whenest**) değişkeni için gözlenene karşı beklenen hayatta kalma olasılıkları ile PH varsayımının değerlendirilmesi



(**whenest**) değişkeninin 0 kategorisinde yer alan gözlemlerin gözlenen ve beklenen hayatta kalma eğrileri 30. aya kadar birbirine yakındır ancak gözlem süresi boyunca iki defa birbirlerine uzaklaşıp yaklaşmaktadır. Bu değişken için PH varsayımı ihlal edilmektedir.

Şekil 5.27: (**cycle**) değişkeni için gözlenene karşı beklenen hayatta kalma olasılıkları ile PH varsayımının değerlendirilmesi



(**cycle**) değişkeninin hayatta kalma eğrileri yaklaşık 70. aya kadar birbirine yakın bir seyir izlemesine rağmen, birbirlerinden uzaklaşmışlardır, yani PH varsayımı ihlal edilmektedir.

Gözlenene karşı beklenen hayatta kalma eğrileri ile PH varsayımı değerlendirildiğinde; (**gender**), (**enteduuni**), (**ailecev**), (**export**), (**reklam**), (**marka**), (**whenest**), (**cycle**) değişkenlerinin hala PH varsayımını sağlamadığı sonucuna ulaşılmaktadır. (**ortaksay**) ve (**yenilik**) değişkenleri ise Cox ayarlaması yapıldıktan sonra PH varsayımını sağlamıştır.

5.4.3. İstatistiksel Test ile PH Varsayımının Denetlenmesi

Ayarlanmış Cox hayatta kalma eğrileri tahmin sonuçlarına göre sonra PH varsayımını sağlamayan değişkenler; (**gender**), (**enteduuni**), (**ailecev**), (**export**), (**reklam**), (**marka**), (**whenest**) ve (**cycle**)'dir. Ancak grafik yöntemler daha öznel sonuçlara dayanmaktadır. Bu

yüzden değişkenlerin PH varsayımını sağlayıp sağlamadığını test etmek için Schoenfeld artıklarını kullanan bir istatistiksel test olan PH testi uygulanmıştır.

Bu test, grafiksel yöntemlerden daha iyi değildir ancak daha objektiftir. Bu testi uygulayabilmek için; genel test sonucu için Schoenfeld artıklarının, her bir değişkenin ayrı ayrı testi için ise ölçeklendirilmiş Schoenfeld artıklarının elde edilmesi gerekmektedir. PH varsayımının sağlanması için artıkların yaşam süresi ile ilişkili olmaması gerekir. Artıklar göreceli olarak daha önce meydana gelen olaylar için pozitif, daha sonra meydana gelen olaylar için negatif olmaya meyilli (ya da tam tersi) ise böyle bir durum tehlike oranlarının zamana göre sabit olmadığına (yani PH varsayımının ihlal edildiğine) dair bir kanıttır.

Tablo 5.8: PH varsayımının istatistiksel bir yöntemle test edilmesi

Değişken	rho	chi2	prob>chi2
tekgelir	0.268	1.360	0.243
gender	-0.413	6.200	0.013**
lnentage	0.356	1.860	0.173
enteduuni	-0.320	2.740	0.098*
workexp	-0.061	0.070	0.785
ailecev	-0.123	0.450	0.505
ortaksay	-0.086	0.070	0.795
export	-0.148	0.410	0.524
lnempini	-0.396	4.850	0.028**
onlyselffin	-0.142	0.480	0.487
networking	-0.393	1.830	0.176
yenilik	0.100	0.230	0.631
reklam	-0.147	0.630	0.427
marka	0.354	2.660	0.103*
buro	-0.210	1.630	0.201
whenest	0.244	1.260	0.261
incubsize	0.389	1.750	0.187
sektor	0.009	0.000	0.965
compete	-0.233	2.220	0.136
ilrank	-0.016	0.000	0.947
cycle	-0.258	1.300	0.254
Genel test		29.34	0.106

*: %10 önem düzeyinde anlamlılığı, **: %5 önem düzeyinde anlamlılığı ve ***: %1 önem düzeyinde anlamlılığı göstermektedir.

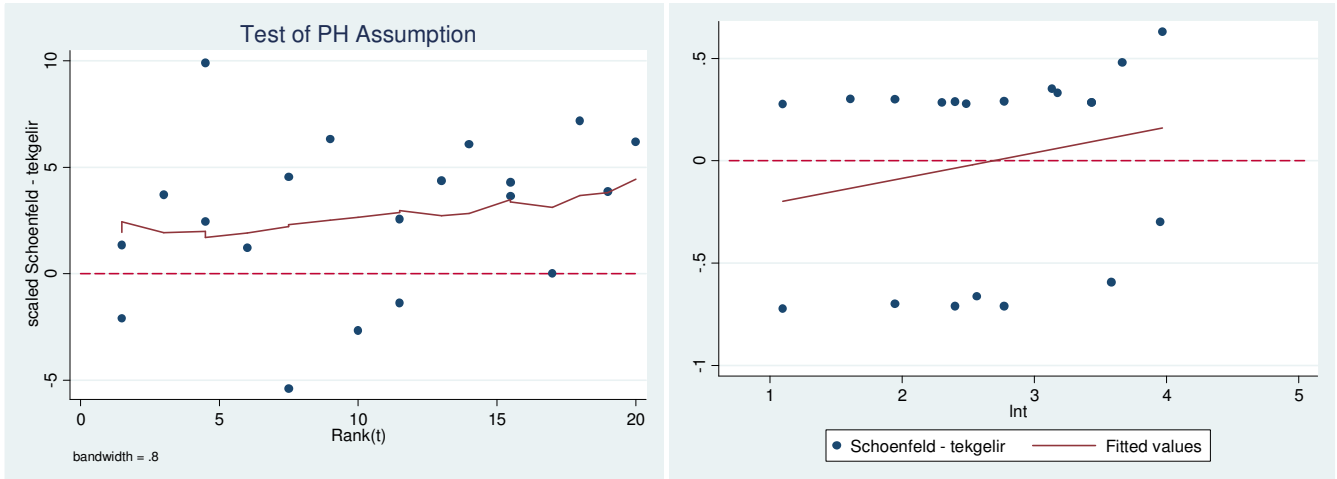
Genel test sonucu modelde aynı anda tüm değişkenlerin yer aldığı durumdaki PH varsayımını değerlendirmektedir. Modeldeki değişkenler toplu olarak PH varsayımını ihlal etmemektedir. Her bir değişkenin katsayısı (rho) ve p değerleri ise, değişkenlerin ayrı ayrı PH varsayımını ihlal edip etmediğini göstermektedir. (**gender**), (**enteduuni**) ve (**lnempini**)

değişkenleri için PH varsayımı ihlal edilmektedir. (**marka**) değişkeni ise kritik değere çok yakın olduğundan (0.103) %11 önem düzeyinde anlamlı olarak ele alınabilir. Kategorik olan (**gender**), (**enteduuni**) ve (**marka**) değişkenlerinin PH varsayımını ihlal ettiği sonucu grafiksel yöntemle de bulunmuştu.

5.4.4. Schoenfeld Artıkları ile PH Varsayımının Denetlenmesi

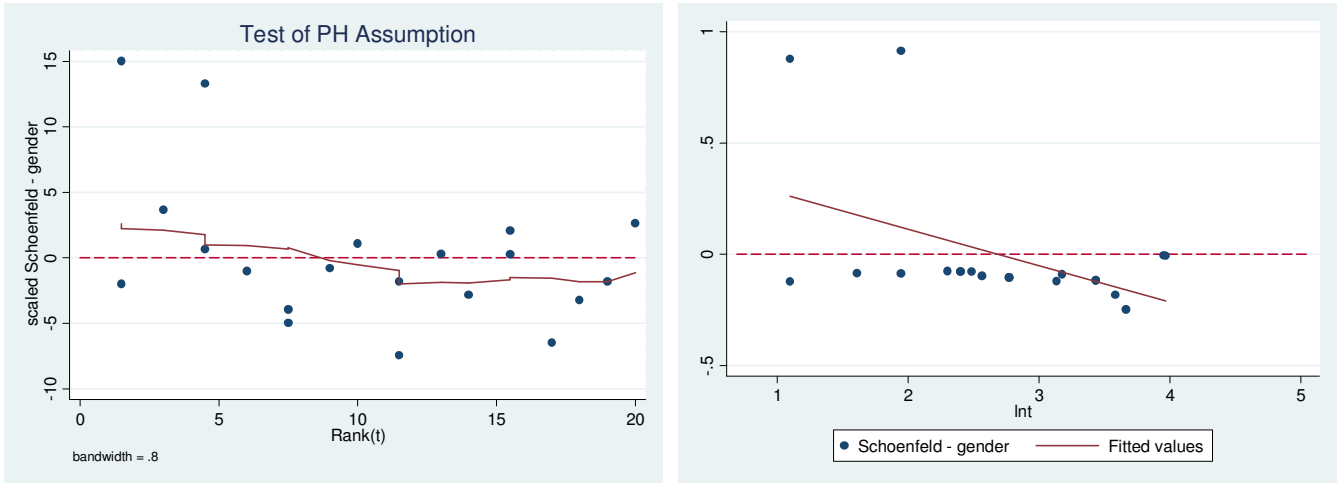
Bu istatistik sonuçları grafiksel olarak ifade edildiğinde; yatay eksende sıralanmış başarısızlık zamanları, dikey eksende değişkene ait ölçeklendirilmiş Schoenfeld artıkları yer alır. Değişkenin PH varsayımını sağlaması için, artıklar için tahminlenen eğrinin yatay eksene paralel olması gerekir çünkü ölçeklendirilmiş Schoenfeld artıkları yaşam süresinden bağımsız olmalıdır. Aşağıda modelde yer alan her bir değişken için ayrı ayrı PH varsayımını denetleyen grafikler, sol tarafta verilmektedir. Sağ tarafta yer alan grafikler ise, her bir değişkenin $\ln(t)$ boyunca değişen Schoenfeld artıkları için tahminlenen doğruları göstermektedir. Bu doğruların da yatay eksene paralel olması, Schoenfeld artıkları ile zaman arasında ilişki olmadığı yani ilgilenilen değişkenin PH varsayımını sağladığı anlamına gelmektedir.

Şekil 5.28: (**tekgelir**) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi



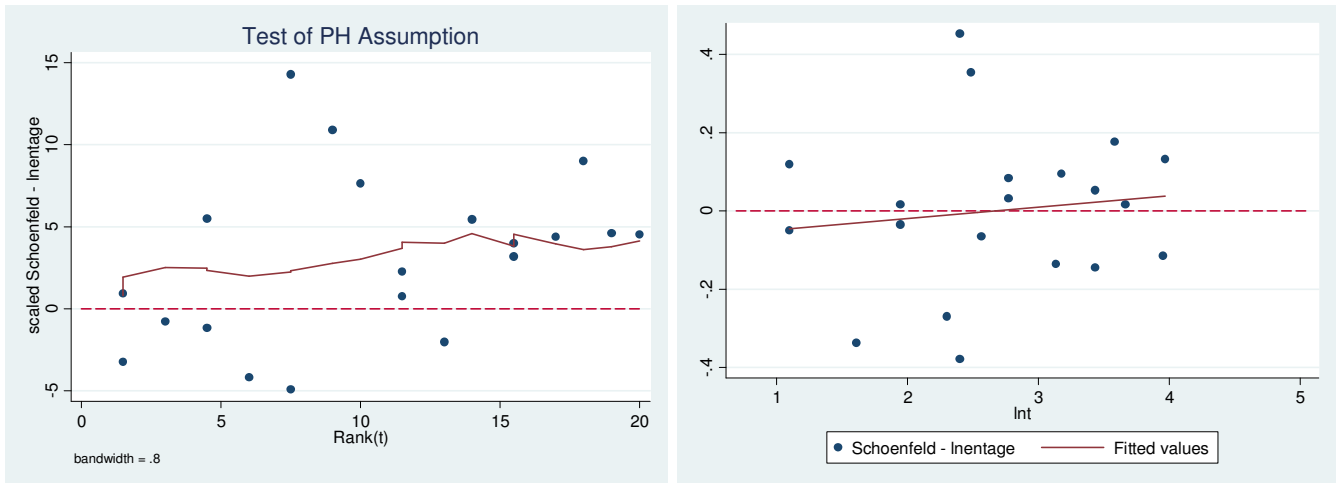
(**tekgelir**) değişkeni için artıkların yatay bir seyir izlediği yani PH varsayımının sağlandığı söylenebilir.

Şekil 5.29: (**gender**) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi



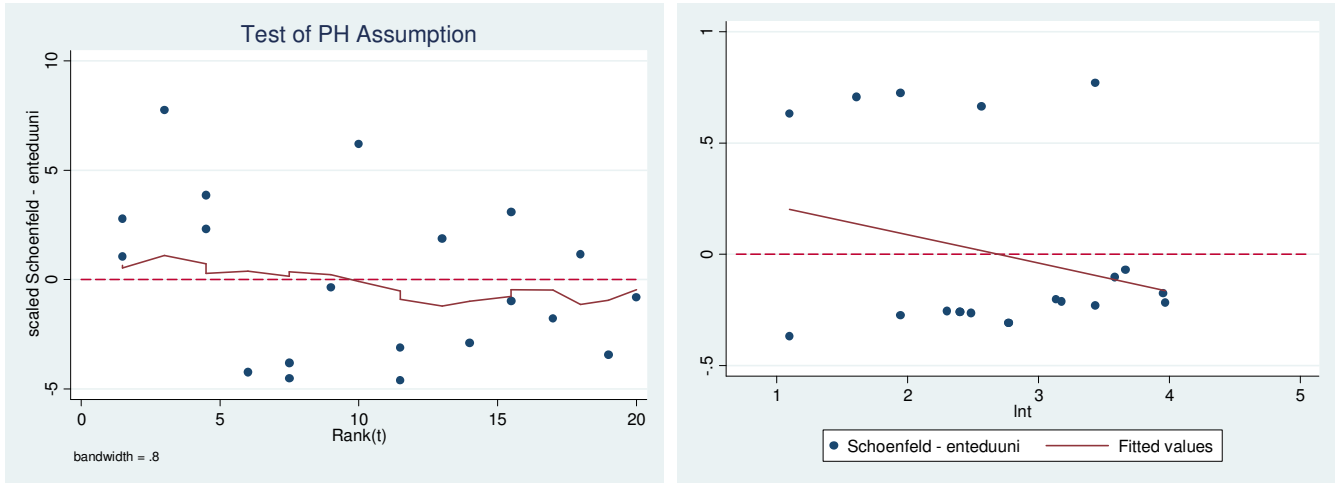
(**gender**) değişkeninin artıkları için tahminlenen eğri başlangıçta pozitif, daha sonra ise negatiftir. Artıkların değeri zamanla azalmaktadır, bu nedenle PH varsayımı ihlal edilmektedir.

Şekil 5.30: (**lnentage**) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi



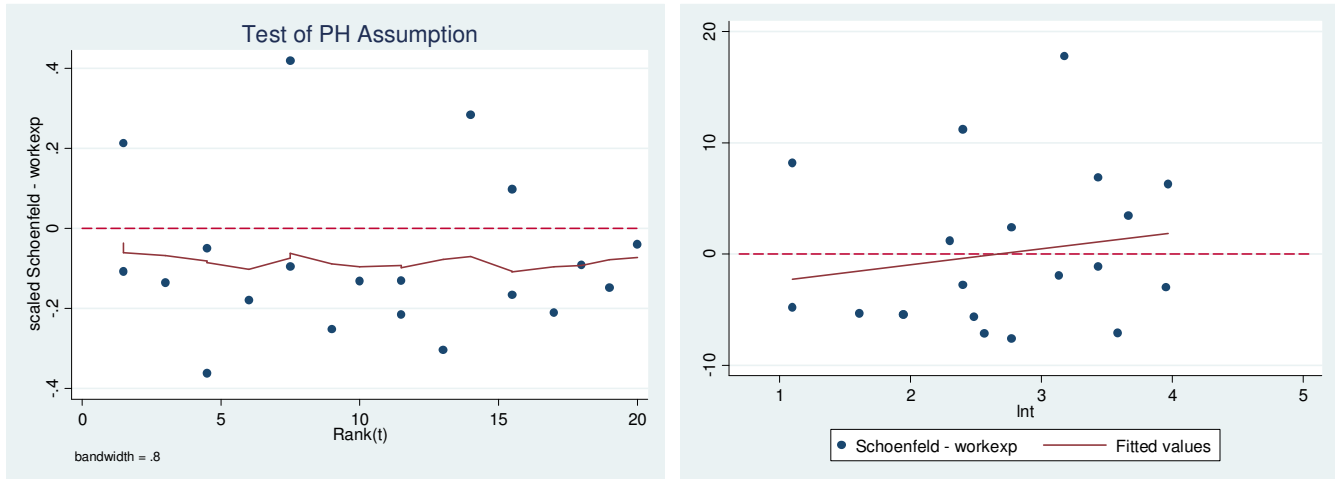
(**lnentage**) değişkeninin artıkları da kabaca yatay bir seyir izlemektedir. Yani (**lnentage**) değişkeni, PH varsayımını sağlamaktadır.

Şekil 5.31: (**enteduuni**) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi



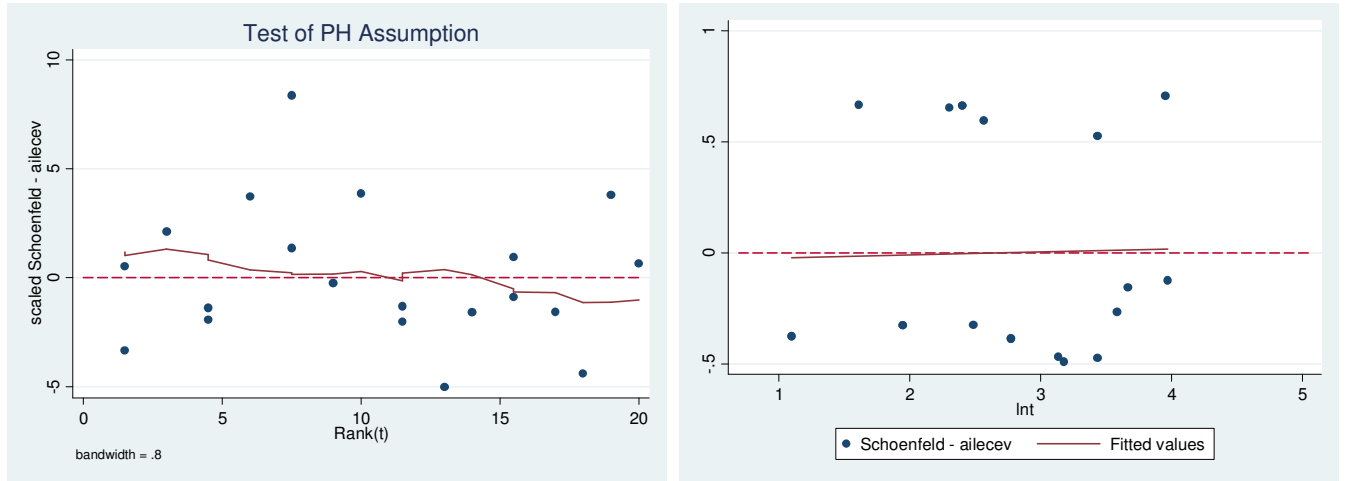
(**enteduuni**) değişkeninin artıkları için tahminlenen eğri başlangıçta pozitif, daha sonra ise negatif artıklara sahiptir. Tahminlenen eğri yatay eksene paralel olmadığından, PH varsayımı ihlal edilmektedir.

Şekil 5.32: (**workexp**) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi



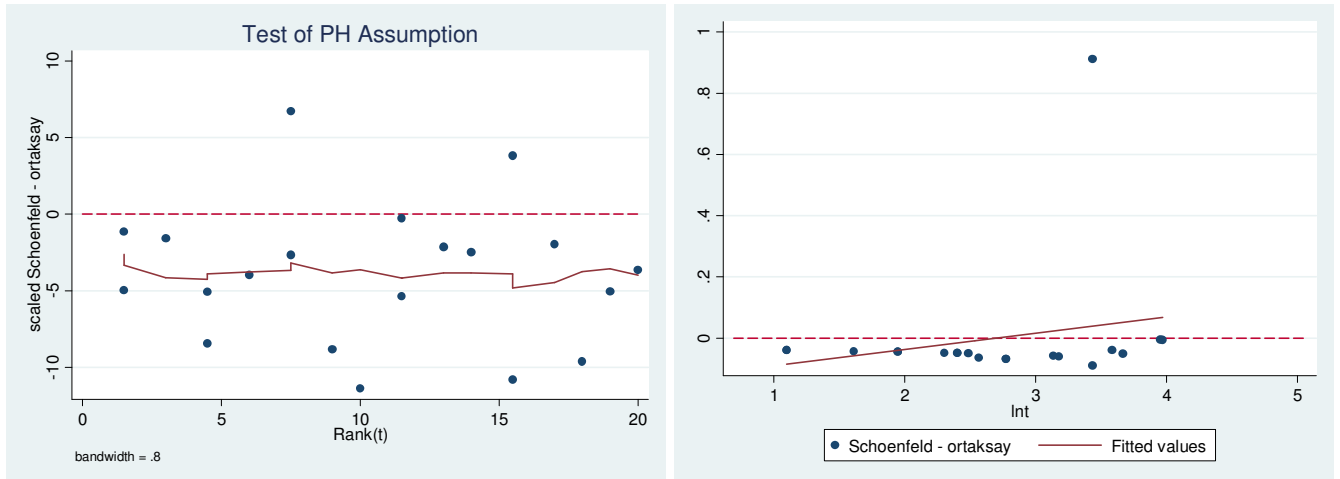
(**workexp**) değişkeninin artıkları, neredeyse tüm başarısızlık zamanlarında negatiftir ve tahminlenen eğri yatay eksene paraleldir. Bu nedenle (**workexp**) değişkeni için PH varsayımı sağlanmaktadır.

Şekil 5.33: (ailecev) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi



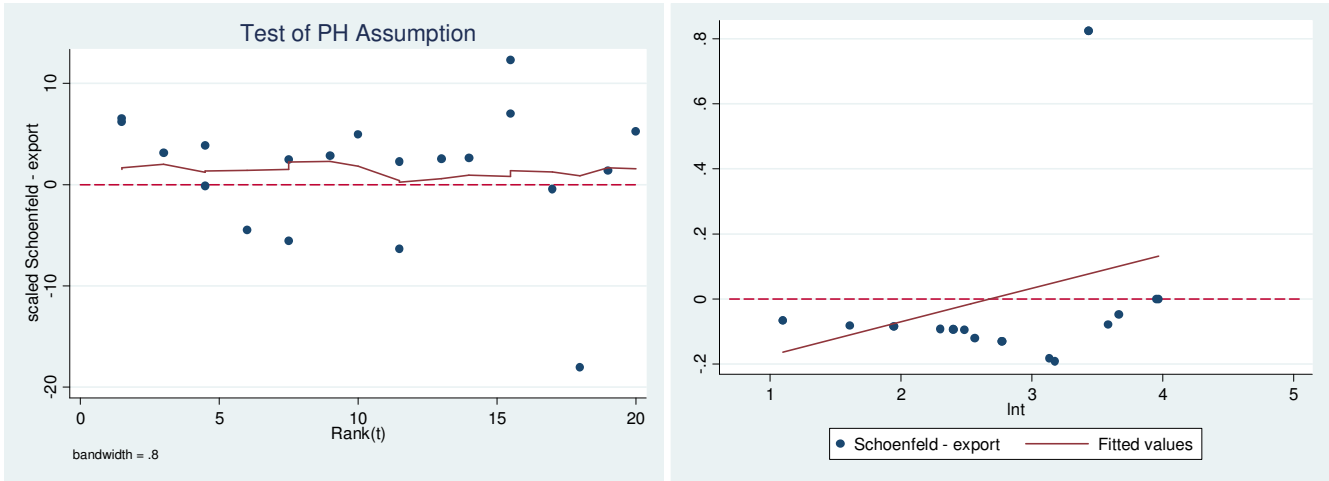
(ailecev) değişkeninin artıkları için tahminlenen eğrinin önce pozitif, sonra negatif, daha sonra tekrar sırasıyla pozitif ve negatif bölgede yer aldığı görülmektedir. Eğri, yatay eksene paralel olmadığından PH varsayımı ihlal edilmektedir.

Şekil 5.34: (ortaksay) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi



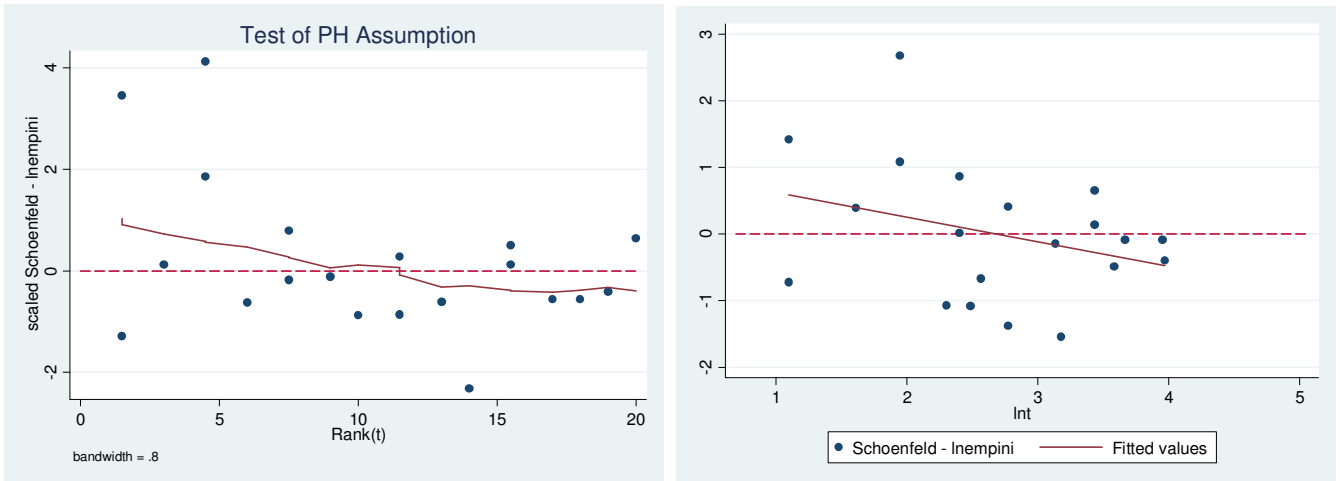
(ortaksay) değişkeninin artıkları, neredeyse tüm başarısızlık zamanlarında negatiftir ve tahminlenen eğri, yatay eksene paraleldir. Yani bu değişken için PH varsayımı sağlanmaktadır.

Şekil 5.35: (**export**) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi



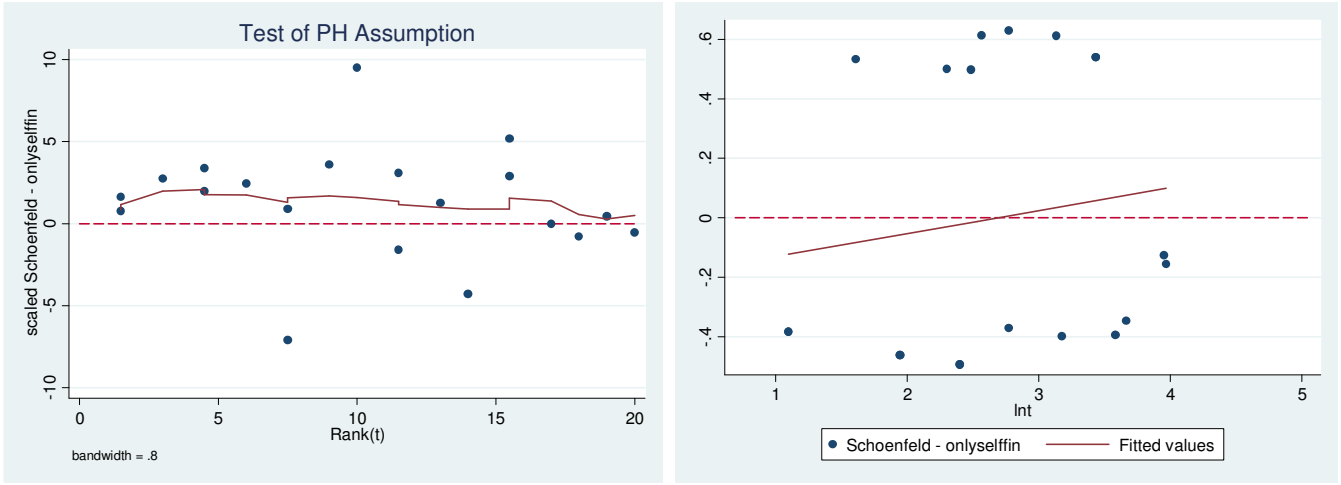
Sağdaki grafite, (**export**) değişkeninin artıkları için tahminlenen eğri, pozitif bölgede yer almakta ve yatay bir seyir izlemektedir. Soldaki grafikte ise, eğri önce negatif sonra pozitif bölgede yer almaktadır, ancak bunun sebebi 0,8 değerine sahip olan uç değerdir. Bu nedenle, (**export**) değişkeni PH varsayımını sağlamaktadır.

Şekil 5.36: (**lnempini**) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi



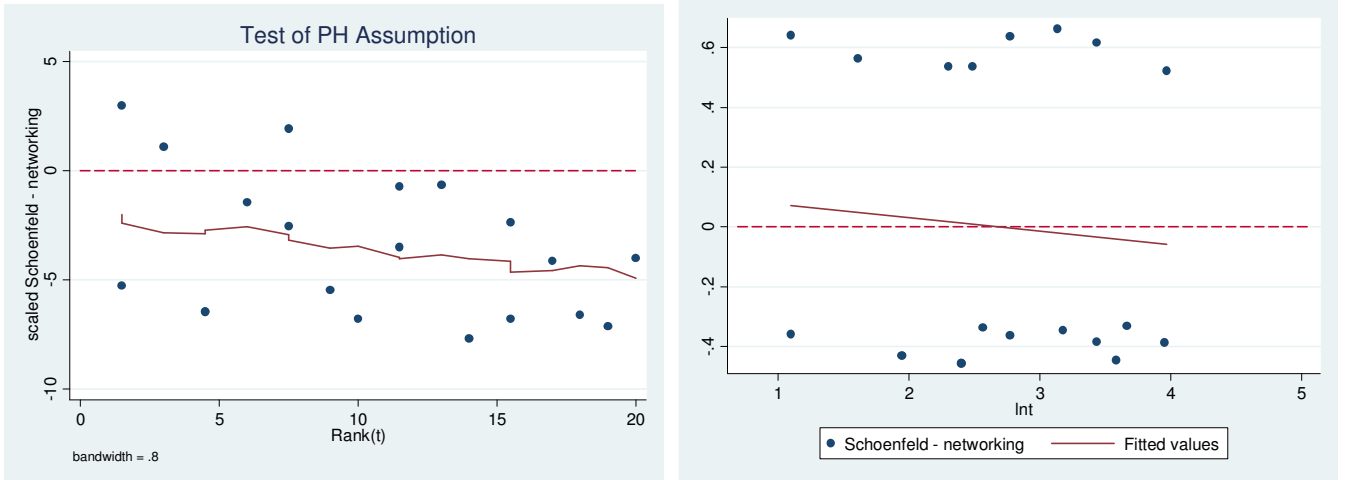
(**lnempini**) değişkeni için artıkların tahminlenen eğrisi önce pozitif, sonra negatif bölgede yer almaktadır. Eğri negatif eğime sahip olduğundan, PH varsayımı sağlanmamaktadır.

Şekil 5.37: (**onlyselffin**) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi



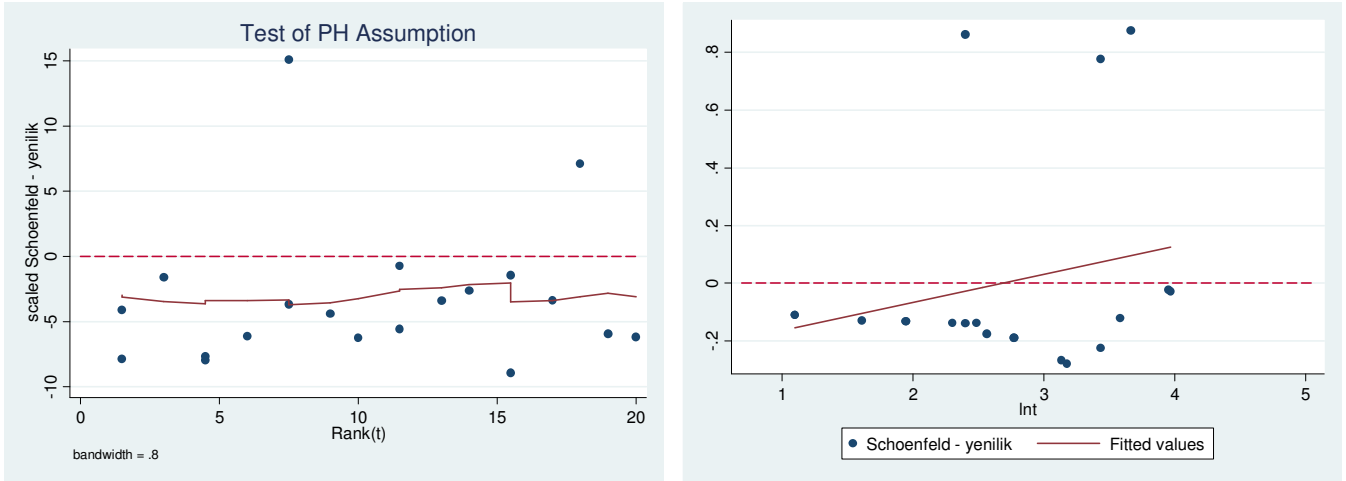
(**onlyselffin**) değişkeni için artıkların tahminlenmiş eğrisi, tüm başarısızlık zamanları için pozitif bölgededir ve yataydır. Bu nedenle, PH varsayımı sağlanmaktadır.

Şekil 5.38: (**networking**) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi



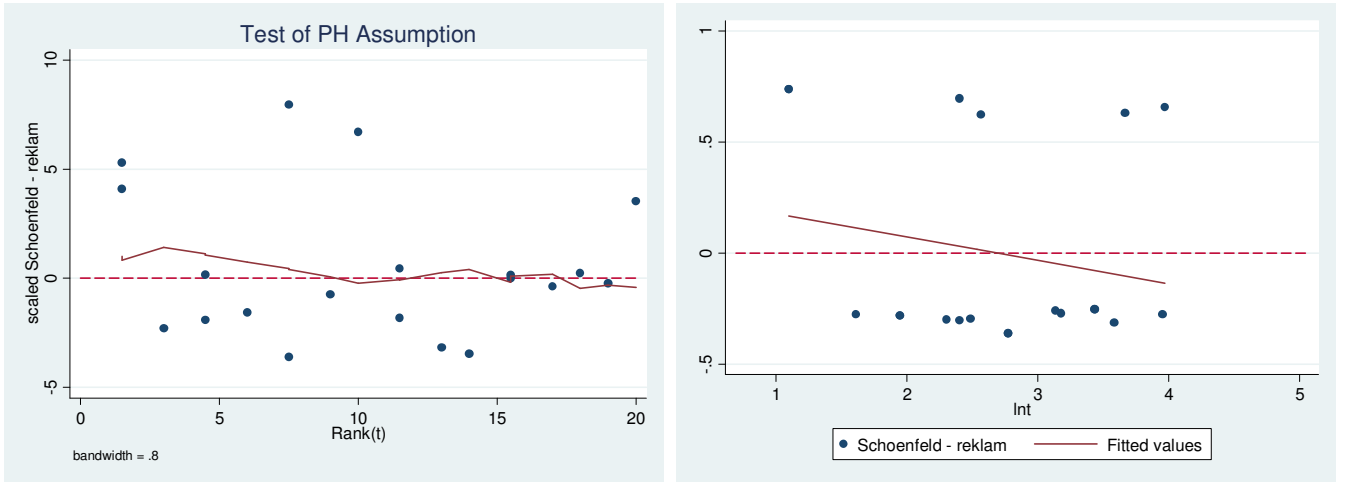
(**networking**) değişkenine ait artıkların tahminlenmiş eğrisi, tüm başarısızlık zamanları için negatif bölgededir. Görel olarak başlangıçta meydana gelen üç tane başarısızlığa ait pozitif artıkların eğriyi yukarıya çekebileceği göz önünde bulundurulduğunda, eğrinin kabaca yatay bir seyir izlediği söylenebilir. Yani PH varsayımının sağlandığı sonucuna ulaşılır.

Şekil 5.39: (**yenilik**) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi



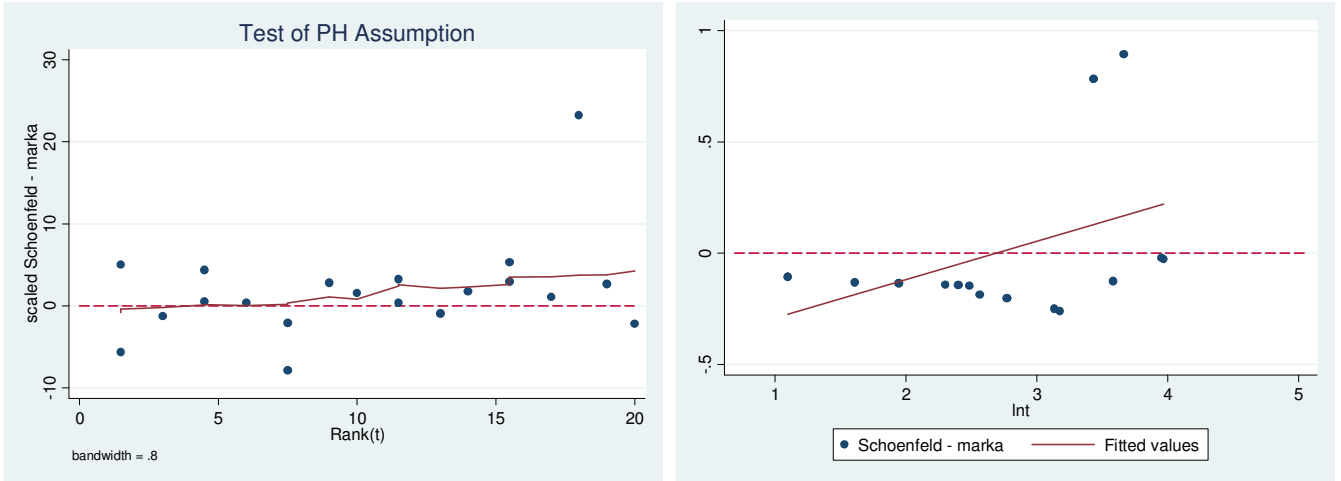
(**yenilik**) değişkeninin artıkları için tahminlenen eğri, negatif bölgede yer almakta ve yatay bir seyir izlemektedir. Bu nedenle, (**yenilik**) değişkeni PH varsayımını sağlamaktadır.

Şekil 5.40: (**reklam**) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi



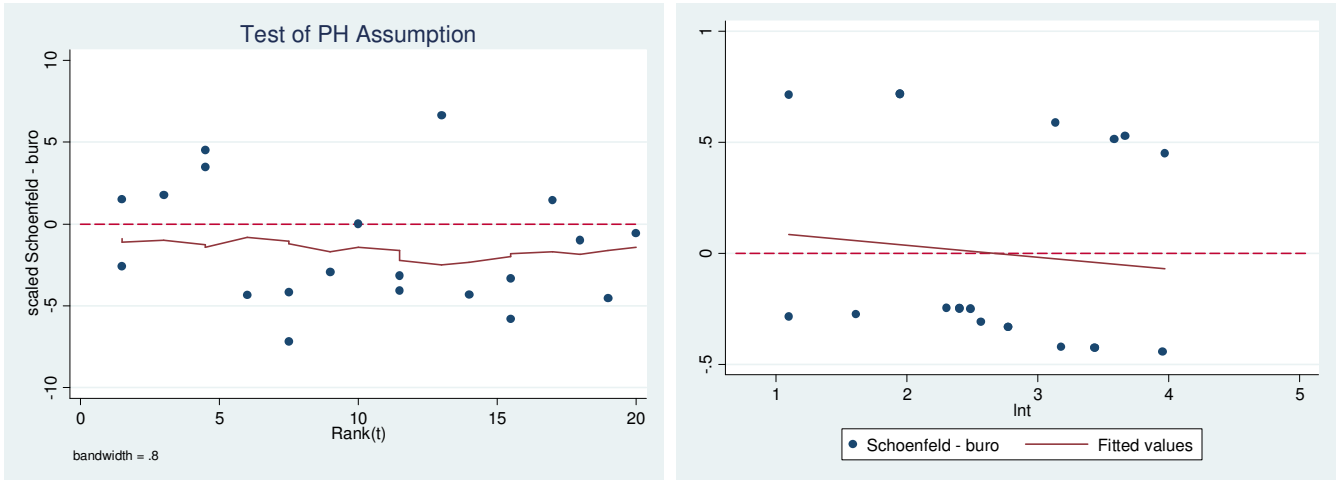
(**reklam**) değişkeninin artıkları için tahminlenen eğri, negatif bir eğime sahiptir. Ayrıca, birçok defa pozitif ve negatif bölgeler arasında yer değiştirerek ilerlemiştir. Bu nedenle, (**reklam**) değişkeni için PH varsayımı ihlal edilmektedir.

Şekil 5.41: (**marka**) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi



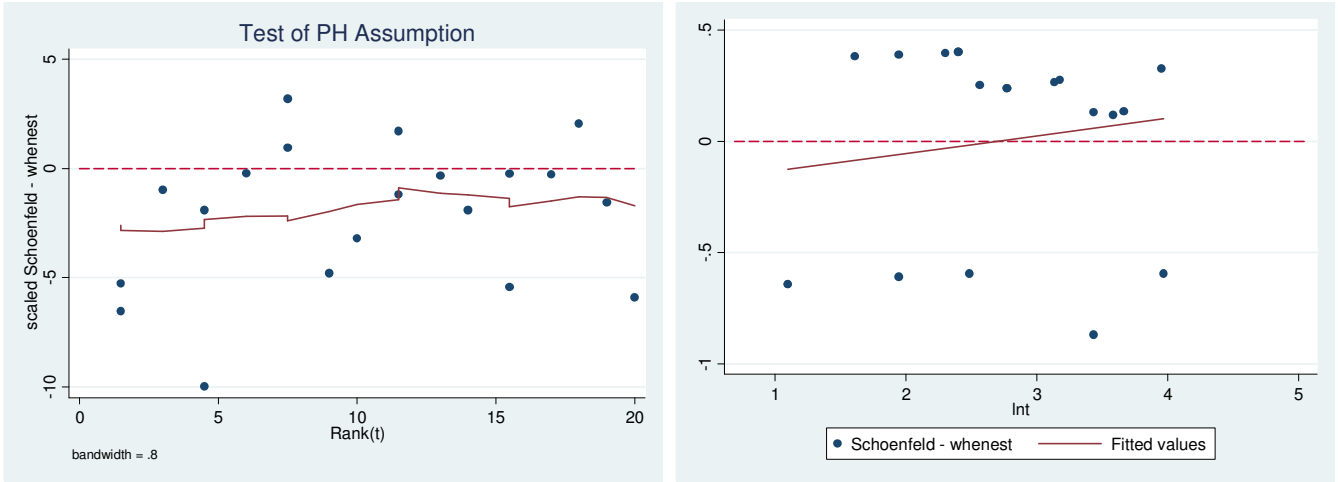
(**marka**) değişkeninin artıkları için tahminlenen eğri, pozitif bir eğime sahip olmakla beraber önce negatif, sonra pozitif bölgede yer almaktadır. Bu nedenle, (**marka**) değişkeni için PH varsayımı ihlal edilmektedir.

Şekil 5.42: (**buuro**) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi



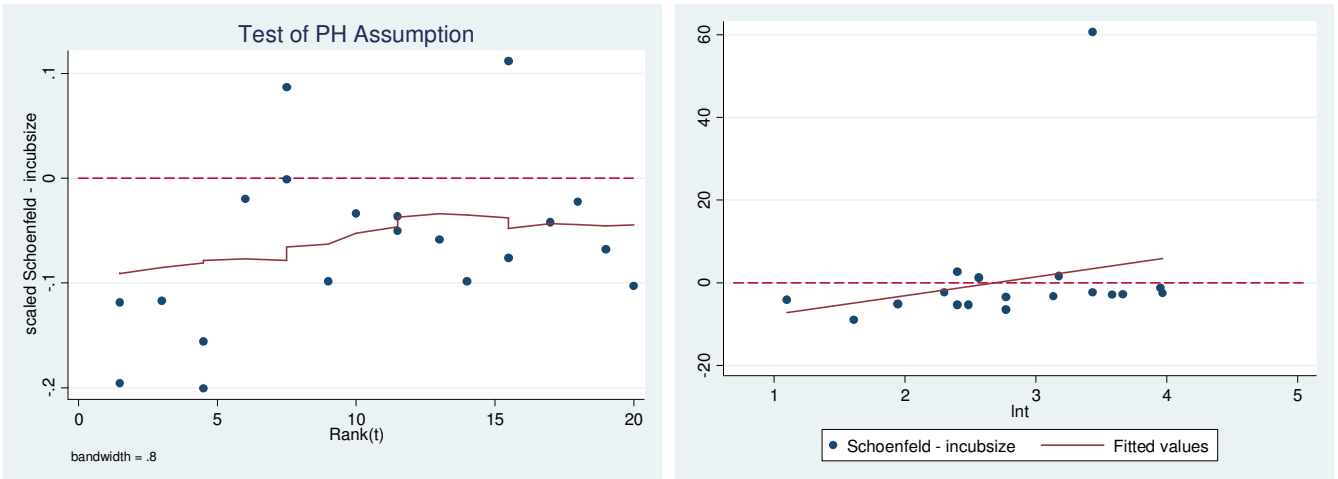
(**buuro**) değişkeninin artıkları için tahminlenen eğri, negatif bölgede yer almakta ve yatay bir seyir izlemektedir. Bu nedenle, (**buuro**) değişkeni için PH varsayımı sağlanmaktadır.

Şekil 5.43: (**whenest**) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi



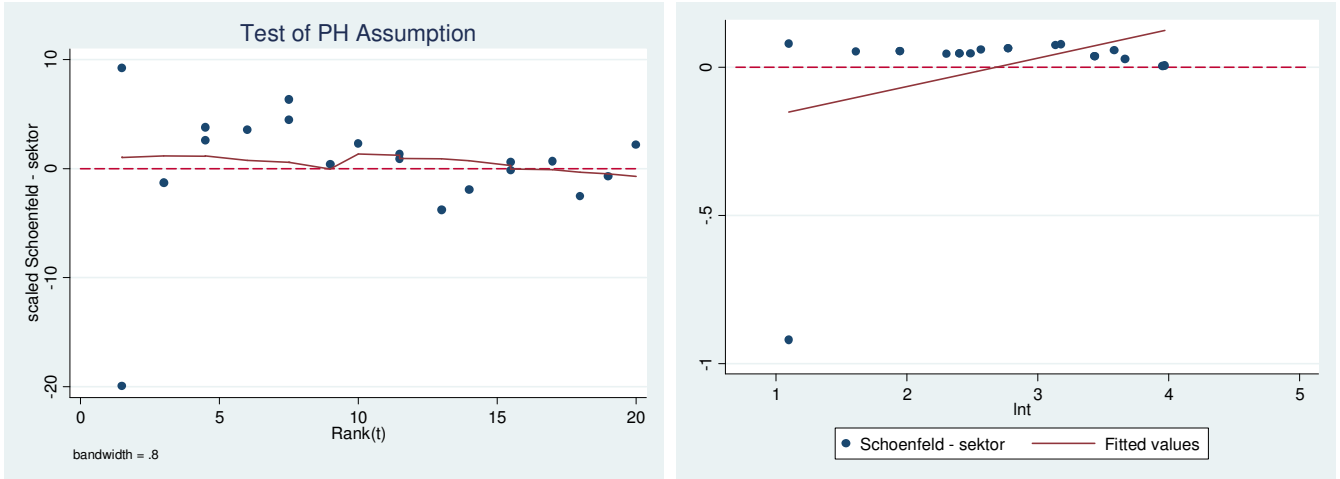
(**whenest**) değişkeninin artıkları için tahminlenen eğri de, negatif bölgede yer almaktadır. Ayrıca, kabaca yatay bir seyir izlediği söylenebilir. Bu nedenle, (**whenest**) değişkeni için PH varsayımı sağlanmaktadır.

Şekil 5.44: (**incubsize**) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi



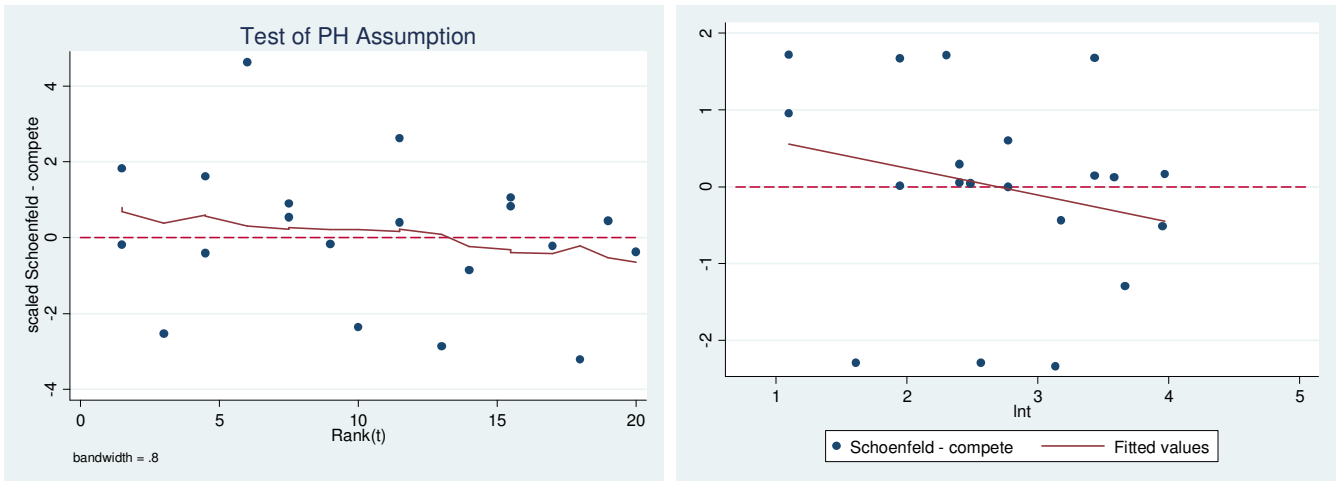
(**incubsize**) değişkenine ait artıkların tahminlenmiş eğrisi de, negatif bölgede yer almaktadır. Ancak ilk beş başarısızlık zamanındaki artıkların, eğriyi başlangıçta aşağıya çektiği söylenebilir. Bu nedenle, eğrinin kabaca yatay olduğunu söylemek mümkündür. Yani, (**incubsize**) değişkeni için PH varsayımı sağlanmaktadır.

Şekil 5.45: (**sektor**) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi



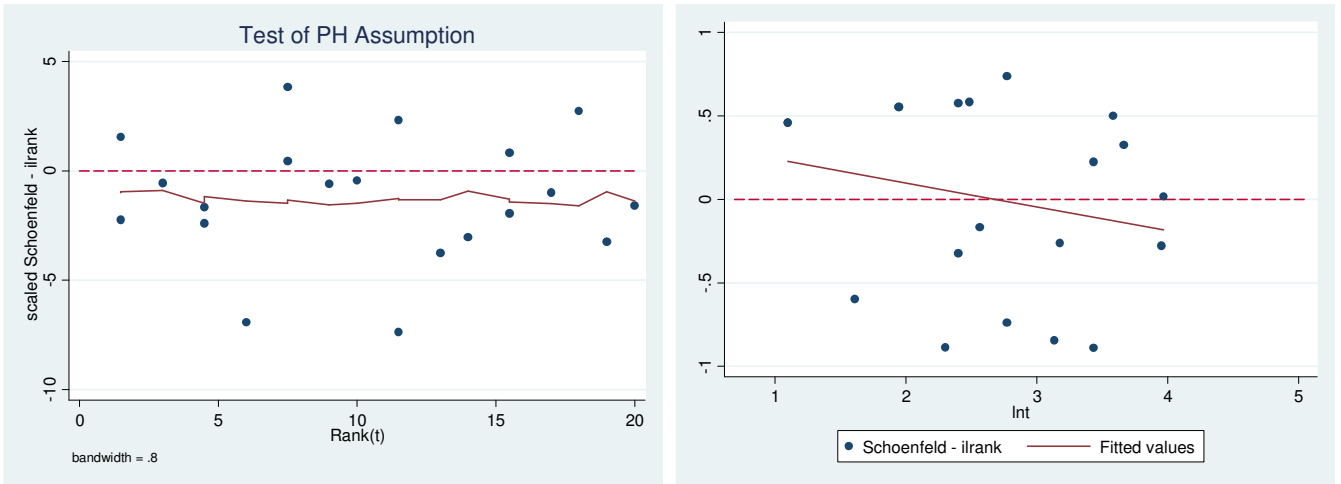
(**sektor**) değişkenine ait artıkların tahminlenmiş eğrisi yataydır. Eğri, genel olarak pozitif bölgede olmasına rağmen son üç başarısızlık zamanında negatif bölgede yer almaktadır. Ancak, işaret değişiklikleri sıfıra çok yakın değerlerde gerçekleştiği için göz ardı edilebilecek düzeydedir. Bu nedenle, (**sektor**) değişkeninin PH varsayımını sağladığı sonucuna ulaşılmaktadır.

Şekil 5.46: (**compete**) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi



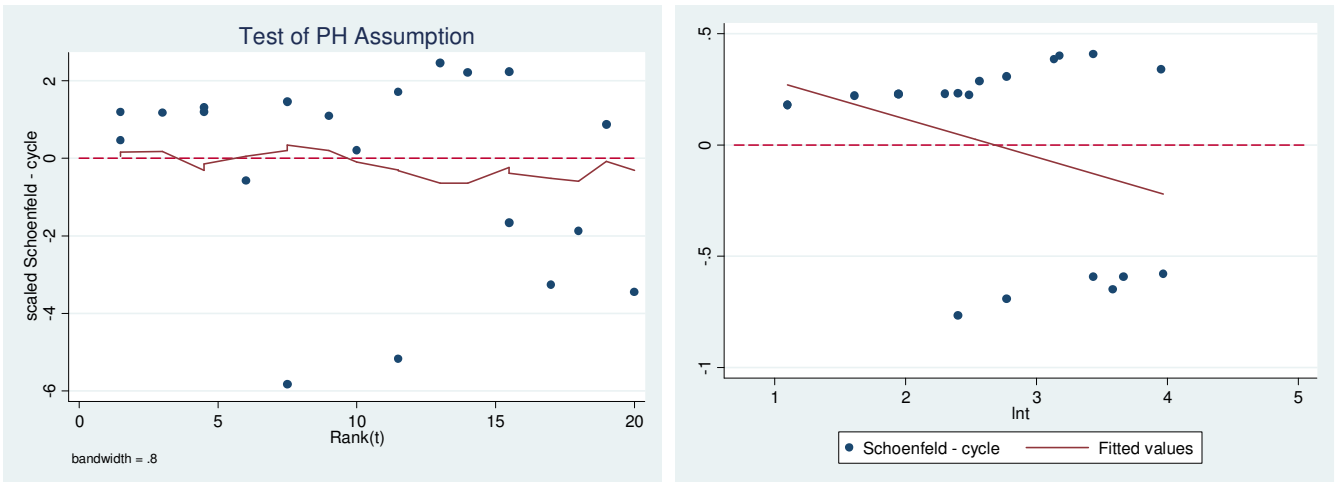
(**compete**) değişkenine ait eğri, negatif eğime sahiptir ve önce pozitif, sonra negatif bölgede yer almaktadır. Bu yüzden, PH varsayımı ihlal edilmektedir.

Şekil 5.47: (**ilrank**) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi



(**ilrank**) değişkeninin artıkları için tahminlenen eğri, negatif bölgede yer almakta ve yatay bir seyir izlemektedir. Bu nedenle, (**ilrank**) değişkeni için PH varsayımı sağlanmaktadır.

Şekil 5.48: (**cycle**) değişkeni için Schoenfeld artıkları ile PH varsayımının değerlendirilmesi



(**cycle**) değişkenine ait artıkların tahminlenmiş eğrisi yataydır. Eğri, iki defa sırasıyla pozitif ve negatif bölgeler arasında yer değiştirerek ilerlemiştir. Ancak, işaret değişiklikleri sifira çok yakın değerlerde gerçekleştiği için göz ardı edilebilecek düzeydedir. Bu nedenle, (**cycle**) değişkeninin PH varsayımını sağladığı sonucuna ulaşılmaktadır.

Sonuç olarak PH varsayımının testi için çizilen grafiklerde; (**gender**), (**enteduuni**), (**aillecev**), (**Inempini**), (**reklam**), (**marka**) ve (**compete**) değişkenlerinin bu varsayımı ihlal ettiği bulunmuştur.

Yapılan analizlerden bazı değişkenler için PH varsayımının ayrı ayrı sağlandığı görülmektedir. Modelde PH varsayımı tüm değişkenler için aynı anda değerlendirildiğinde, modelin bir bütün olarak PH varsayımını sağladığı sonucu elde edilmektedir.

Tablo 5.9: Modeldeki değişkenler için PH varsayımının topluca değerlendirilmesi

Zaman: <i>t</i>		
	chi2	Prob>chi2
genel test	18.740	0.602

5.4.5. Zamana Bağlı Değişkenleri İçeren Cox Model ile PH Varsayımının Denetlenmesi

PH varsayımı sağlanmadığında tabakalandırılmış Cox modele diğer bir alternatif, zamanla değişen değişkenleri içeren bir Cox model kurmaktır. Genişletilmiş Cox model kurmanın amacı, modele dahil edilecek yaşam süresi için uygun fonksiyonu seçmektir. Değişkenlerin zamanın logaritması $\ln(t)$ ile etkileşimi ele alındığında ve zaman bağımlılığı ayrı ayrı değerlendirildiğinde; (**gender**), (**marka**) ve (**sektor**) değişkenleri zamana bağlı çıkmaktadır. Ancak en iyi modeli seçebilmek için genelden özele seçim metodu kullanılarak aşağıdaki modeller kurulmuştur. Tablo 5.10'da her değişken için ilk satır değişkenin katsayı değerini, ikinci satır p değerini göstermektedir.

Tablo 5.10: Zamana bağlı değişkenleri içeren Cox model tahminleri

Değişken	Model 27	Model 28	Model 29	Model 30	Model 31	Model 32	Model 33	Model 34	Model 35	Model 36	Model 37	Model 38	Model 39	Model 40	Model 41
tekgelir	-1.160	-1.150	-1.380	-0.677	-0.017	0.362	2.890	3.020	2.930	2.970	2.980	3.170	3.160	3.170	2.860
	0.790	0.791	0.743	0.863	0.996	0.917	0.006***	0.005***	0.005***	0.004***	0.004***	0.001***	0.001***	0.001***	0.001***
gender	13.300	13.300	13.000	13.500	14.100	13.300	13.700	13.800	13.900	14.600	14.400	14.000	13.900	12.500	9.850
	0.014**	0.014**	0.012**	0.010***	0.007***	0.007***	0.003***	0.003***	0.004***	0.002***	0.004***	0.003***	0.003***	0.003***	0.008***
lnentage	0.185	0.175	0.039	3.460	3.550	3.670	3.530	3.570	3.440	3.600	3.580	4.090	4.150	4.220	3.920
	0.979	0.980	0.995	0.043**	0.036**	0.031**	0.036**	0.030**	0.036**	0.028**	0.028**	0.011**	0.008***	0.007***	0.008***
enteduuni	6.940	6.940	6.860	7.170	7.090	6.900	6.900	6.520	6.780	6.930	6.380	6.240	5.900	6.360	5.820
	0.036**	0.036**	0.035**	0.023**	0.027**	0.026**	0.025**	0.033**	0.038**	0.028**	0.033**	0.038**	0.038**	0.019**	0.036**
workexp	-0.422	-0.422	-0.423	-0.433	-0.466	-0.454	-0.495	-0.461	-0.467	-0.482	-0.472	-0.501	-0.385	-0.132	-0.130
	0.162	0.162	0.161	0.139	0.097*	0.100*	0.064*	0.075*	0.076*	0.062*	0.072*	0.051*	0.088*	0.012**	0.012**
ailecev	5.940	5.920	5.820	5.980	6.300	5.860	5.820	5.370	5.710	5.290	5.340	5.830	6.210	5.860	5.430
	0.067*	0.062*	0.061*	0.051*	0.036**	0.043**	0.047**	0.058*	0.036**	0.043**	0.037**	0.021**	0.013**	0.018**	0.025**
ortaksay	-5.490	-5.270	-5.290	-5.110	-5.120	-5.450	-5.280	-5.000	-4.790	-4.560	-4.470	-4.750	-4.440	-4.870	-4.650
	0.424	0.005***	0.005***	0.004***	0.004***	0.003***	0.002***	0.002***	0.001***	0.001***	0.002***	0.002***	0.004***	0.003***	0.004***
export	-7.140	-7.190	-7.210	-6.640	-6.470	-7.260	-6.050	0.352	0.898	1.250	1.100	1.170	1.280	1.510	1.680

	0.484	0.478	0.473	0.491	0.518	0.457	0.528	0.841	0.587	0.430	0.485	0.420	0.380	0.312	0.283
lnempini	3.780	3.770	3.790	3.910	3.980	3.810	3.660	3.500	3.500	3.550	3.410	3.750	3.770	3.260	2.520
	0.010**	0.010**	0.009***	0.007***	0.006***	0.007***	0.008***	0.010**	0.010**	0.008***	0.009***	0.004***	0.004***	0.006***	0.012**
onlyselffin	3.280	3.260	3.290	3.650	3.390	1.500	1.540	1.660	1.640	1.660	1.600	1.680	1.580	1.650	1.560
	0.350	0.345	0.334	0.266	0.295	0.071*	0.058*	0.045**	0.051*	0.045**	0.052*	0.041**	0.050*	0.040**	0.048**
networking	3.590	3.610	3.910	3.330	3.300	4.150	2.900	1.700	1.120	-0.055	-0.896	-1.710	-4.910	-4.970	-4.780
	0.476	0.463	0.398	0.450	0.467	0.328	0.450	0.630	0.749	0.985	0.746	0.515	0.000***	0.000***	0.000***
yenilik	-2.750	-2.710	-2.410	-2.120	-2.390	-1.090	-2.270	-3.300	-5.980	-5.810	-5.600	-5.700	-5.420	-5.700	-5.810
	0.599	0.595	0.613	0.652	0.618	0.795	0.549	0.415	0.002***	0.002***	0.002***	0.001***	0.001***	0.001***	0.001***
reklam	7.350	7.340	7.370	7.030	7.350	6.950	6.620	6.000	6.230	6.330	5.270	5.760	6.710	4.980	4.660
	0.043**	0.042**	0.038**	0.040**	0.031**	0.036**	0.037**	0.054*	0.045**	0.041**	0.055*	0.039**	0.014**	0.025**	0.020**
marka	-35.200	-35.400	-35.600	-36.800	-38.400	-41.800	-41.000	-42.300	-36.300	-35.700	-33.400	-34.200	-31.400	-34.200	-33.100
	0.147	0.130	0.129	0.116	0.105	0.077*	0.073*	0.048**	0.047**	0.045**	0.051*	0.042**	0.039**	0.039**	0.042**
buro	-1.390	-1.380	-1.440	-1.840	-3.120	-3.220	-3.270	-3.310	-3.320	-3.400	-3.290	-3.400	-3.170	-2.870	-2.620
	0.667	0.667	0.648	0.550	0.002***	0.002***	0.001***	0.002***	0.002***	0.001***	0.001***	0.001***	0.001***	0.001***	0.001***
whenest	0.880	0.906	0.924	0.537	0.400	1.170	0.403	-0.554	-0.949	-1.530	-3.840	-4.140	-4.050	-4.160	-3.960
	0.834	0.825	0.818	0.889	0.920	0.760	0.913	0.864	0.766	0.613	0.001***	0.000***	0.000***	0.000***	0.001***
incubsize	-0.158	-0.158	-0.117	-0.112	-0.110	-0.110	-0.111	-0.111	-0.102	-0.100	-0.096	-0.094	-0.091	-0.093	-0.090
	0.391	0.390	0.003***	0.002***	0.002***	0.002***	0.002***	0.001***	0.001***	0.001***	0.000***	0.000***	0.000***	0.000***	0.000***
sektor	-52.600	-52.300	-52.400	-52.500	-52.400	-52.600	-54.200	-53.900	-52.700	-52.100	-51.300	-51.500	-49.500	-49.500	0.917
	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.507
compete	2.610	2.610	2.570	2.450	2.590	2.380	2.100	1.840	2.000	2.190	1.920	1.990	2.160	2.000	2.170
	0.104	0.101	0.103	0.115	0.082*	0.090*	0.106	0.138	0.106	0.072*	0.105	0.075*	0.055*	0.075*	0.047**
ilrank	-0.004	0.012	0.387	0.090	0.088	0.453	0.381	-0.039	-0.619	-2.320	-2.320	-2.480	-2.280	-2.240	-2.220
	0.999	0.997	0.903	0.976	0.977	0.878	0.897	0.990	0.838	0.017**	0.017**	0.008***	0.012**	0.012**	0.013**
cycle	3.250	3.260	3.370	3.470	3.590	4.020	4.110	4.530	3.750	3.380	3.080	-0.810	-0.713	-0.649	-0.682
	0.381	0.376	0.352	0.337	0.321	0.255	0.246	0.233	0.288	0.326	0.376	0.232	0.279	0.328	0.299
tekgelir × ln(t)	1.590	1.590	1.670	1.400	1.190	1.040	-	-	-	-	-	-	-	-	-
	0.340	0.341	0.301	0.349	0.400	0.449	-	-	-	-	-	-	-	-	-
gender × ln(t)	-5.630	-5.630	-5.530	-5.740	-5.950	-5.650	-5.740	-5.820	-5.830	-6.130	-6.130	-6.010	-6.020	-5.380	-4.250
	0.021**	0.021**	0.020**	0.018**	0.014**	0.015**	0.009***	0.009***	0.010**	0.006***	0.011**	0.009***	0.009***	0.010**	0.018**
lnentage × ln(t)	1.240	1.250	1.310	-	-	-	-	-	-	-	-	-	-	-	-
	0.638	0.636	0.619	-	-	-	-	-	-	-	-	-	-	-	-
enteduuni × ln(t)	-3.080	-3.080	-3.060	-3.180	-3.160	-3.110	-3.070	-2.980	-3.060	-3.140	-2.870	-2.900	-2.780	-2.840	-2.560
	0.024**	0.024**	0.023**	0.015**	0.018**	0.018**	0.017**	0.018**	0.022**	0.016**	0.020**	0.019**	0.016**	0.012**	0.020**
workexp × ln(t)	0.092	0.092	0.091	0.096	0.105	0.103	0.122	0.109	0.115	0.120	0.118	0.125	0.085	-	-
	0.336	0.336	0.339	0.300	0.238	0.245	0.138	0.166	0.147	0.126	0.139	0.114	0.230	-	-
ailecev × ln(t)	-1.940	-1.930	-1.900	-1.950	-2.050	-1.830	-1.830	-1.640	-1.760	-1.610	-1.620	-1.820	-2.020	-1.900	-1.820
	0.099*	0.093*	0.093*	0.077*	0.060*	0.074*	0.076*	0.097*	0.065*	0.077*	0.070*	0.036**	0.019**	0.026**	0.028**
ortaksay × ln(t)	0.068	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	0.973	-	-	-	-	-	-	-	-	-	-	-	-	-	-
export × ln(t)	2.660	2.680	2.640	2.440	2.400	2.510	2.110	-	-	-	-	-	-	-	-
	0.391	0.385	0.387	0.407	0.431	0.400	0.473	-	-	-	-	-	-	-	-
lnempini × ln(t)	-1.300	-1.300	-1.300	-1.340	-1.340	-1.280	-1.240	-1.170	-1.180	-1.200	-1.140	-1.240	-1.280	-1.100	-0.888
	0.015**	0.015**	0.015**	0.011**	0.011**	0.013**	0.014**	0.019**	0.017**	0.014**	0.014**	0.006***	0.007***	0.009***	0.015**

onlyselffin × ln(t)	-0.603	-0.595	-0.610	-0.760	-0.690	-	-	-	-	-	-	-	-	-	-
	0.630	0.628	0.615	0.508	0.542	-	-	-	-	-	-	-	-	-	-
networking × ln(t)	-3.140	-3.160	-3.260	-3.040	-3.040	-3.370	-2.870	-2.530	-2.180	-1.760	-1.420	-1.230	-	-	-
	0.103	0.094*	0.070*	0.075*	0.081*	0.042**	0.052*	0.065*	0.083*	0.090*	0.136	0.175	-	-	-
yenilik × ln(t)	-1.170	-1.180	-1.300	-1.400	-1.360	-1.820	-1.360	-1.150	-	-	-	-	-	-	-
	0.569	0.557	0.496	0.459	0.479	0.298	0.393	0.485	-	-	-	-	-	-	-
reklam × ln(t)	-2.490	-2.480	-2.500	-2.380	-2.490	-2.300	-2.120	-1.910	-1.980	-2.040	-1.640	-1.730	-2.170	-1.630	-1.580
	0.069*	0.067*	0.061*	0.067*	0.053*	0.061*	0.066*	0.092*	0.080*	0.071*	0.098*	0.081*	0.022**	0.045**	0.036**
marka × ln(t)	12.000	12.000	12.100	12.500	13.000	14.100	13.800	14.300	12.200	12.000	11.300	11.500	10.700	11.500	11.200
	0.107	0.093*	0.091*	0.083*	0.075*	0.052*	0.049**	0.030**	0.025**	0.024**	0.026**	0.021**	0.018**	0.019**	0.020**
buro × ln(t)	-0.608	-0.613	-0.594	-0.451	-	-	-	-	-	-	-	-	-	-	-
	0.582	0.574	0.581	0.662	-	-	-	-	-	-	-	-	-	-	-
whenest × ln(t)	-1.830	-1.840	-1.830	-1.680	-1.640	-1.960	-1.700	-1.370	-1.150	-0.934	-	-	-	-	-
	0.261	0.246	0.241	0.257	0.279	0.180	0.229	0.268	0.328	0.403	-	-	-	-	-
incubsize × ln(t)	0.012	0.012	-	-	-	-	-	-	-	-	-	-	-	-	-
	0.816	0.816	-	-	-	-	-	-	-	-	-	-	-	-	-
sektor × ln(t)	44.100	43.800	44.000	43.800	43.800	43.900	45.200	45.300	44.500	44.400	44.000	44.500	43.000	43.000	-
	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
compete × ln(t)	-0.862	-0.859	-0.843	-0.798	-0.852	-0.782	-0.667	-0.543	-0.597	-0.678	-0.564	-0.585	-0.681	-0.647	-0.731
	0.153	0.151	0.153	0.171	0.129	0.141	0.174	0.230	0.186	0.124	0.180	0.136	0.077*	0.099*	0.049**
ilrank × ln(t)	-0.872	-0.879	-1.020	-0.929	-0.962	-1.110	-1.070	-1.000	-0.655	-	-	-	-	-	-
	0.546	0.538	0.426	0.456	0.438	0.359	0.372	0.401	0.553	-	-	-	-	-	-
cycle × ln(t)	-1.300	-1.310	-1.360	-1.420	-1.470	-1.630	-1.680	-1.850	-1.510	-1.390	-1.290	-	-	-	-
	0.308	0.301	0.270	0.246	0.229	0.169	0.158	0.142	0.179	0.205	0.245	-	-	-	-
Gözlem Sayısı	338	338	338	338	338	338	338	338	338	338	338	338	338	338	338
Kuluçka İçinde İken Kapanan Firma Sayısı	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20
Log-Likelihood	-49.376	-49.376	-49.406	-49.535	-49.629	-49.819	-50.126	-50.532	-50.761	-50.943	-51.303	-52.122	-53.086	-53.842	-56.197
chi2	112.16	112.16	112.10	111.84	111.65	111.27	110.66	109.85	109.39	109.03	108.31	106.87	104.74	103.23	98.52
Prob>chi2	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***

*: %10 önem düzeyinde anlamlılığı, **: %5 önem düzeyinde anlamlılığı ve ***: %1 önem düzeyinde anlamlılığı göstermektedir.

Model 27, Model 12’de yer alan değişkenleri ve değişken-zaman etkileşimini ele almak için bu değişkenlerin her birinin zamanın bir fonksiyonu olan $\ln(t)$ ile çarpılmasıyla oluşturulan etkileşim terimlerinin tümünü içermektedir. Bir değişkenin zamana bağlı olup olmadığına, değişkene ait etkileşim teriminin katsayı anlamlılığına bakarak karar verilmektedir. Katsayı anlamsız ise, değişkenin zamandan bağımsız olduğu sonucuna ulaşılır ve değişkene ait etkileşim terimi modelden çıkarılır. Katsayı anlamlı ise, söz konusu değişkenin zamana bağlı olduğu yorumu yapılır ve zaman etkisini de göz önüne almak için etkileşim terimi modelden çıkarılmaz.

Model 27'den Model 40'a kadar, en yüksek p değerine sahip olan etkileşim terimleri adım adım modellerden çıkarılmıştır. Model 27'den sırasıyla (**ortaksay**), (**incubsize**), (**lnentage**), (**buro**), (**onlyselffin**), (**tekgelir**), (**export**), (**yenilik**), (**ilrank**), (**whenest**), (**cycle**), (**networking**) ve (**workexp**) değişkenlerine ait etkileşim terimleri çıkarılmıştır. Yani, bu değişkenler ile zaman arasında anlamlı bir ilişki yoktur, başka bir deyişle bu değişkenler zamana bağlı değildir. Model 40'ta (**gender**), (**enteduuni**), (**ailecev**), (**lnempini**), (**reklam**), (**marka**), (**sektor**) ve (**compete**) değişkenlerine ait etkileşim terimleri bulunmaktadır ve tüm terimler anlamlıdır. Ancak (**sektor**) değişkeni ile kendisine ait etkileşim terimi aynı anda modelde yer aldığına, değişkenin p değeri hesaplanamamaktadır. Grafikselleştirme yöntemlerinde (**sektor**) değişkeninin zamana bağlı olmadığı bulunduğundan, nihai modelde (**sektor**) değişkenine ait etkileşim terimine yer verilmemiştir. Sonuç olarak, nihai model olarak değişkenler arasındaki ilişkileri ve değişkenlerin zamanla etkileşimlerini göz önüne alarak firmaların hayatta kalma süresini etkileyen faktörleri en iyi temsil eden Model 42 seçilmelidir.

Değişkenlerin zamana bağlı etkilerini değerlendiren yöntemler ve PH varsayımını ihlal eden değişkenler aşağıda özetlenmektedir. Tablonun son sütunu nihai modelde ilgili değişkene ait etkileşim teriminin yer alıp almadığını göstermektedir. Etkileşim terimi nihai modelde yer alıyorsa “var”, yer almıyorsa “yok” olarak ifade edilmektedir. Tabloda yer almayan modeldeki diğer değişkenler için tüm yöntemlerde PH varsayımı sağlanmaktadır.

Tablo 5.11: Zamana bağlı değişkenlerin belirlenmesinde kullanılan yöntemlerin özet sonuçları

Değişken	Kullanılan Yöntem					Nihai Model
	Log-log KM hayatta kalma eğrileri	Ayarlanmış Cox hayatta kalma eğrileri	PH testi	PH testi için çizilen grafikler	Genişletilmiş Cox model	
gender	+	+	+	+	+	var
enteduuni	+	+	+	+	+	var
ailecev	+	+	-	+	+	var
ortaksay	+	-	-	-	-	yok
export	+	+	-	-	-	yok
lnempini	-	-	+	+	+	var
yenilik	+	-	-	-	-	yok
reklam	+	+	-	+	+	var
marka	+	+	+	+	+	var
whenest	+	+	-	-	-	yok
sektor	-	-	-	-	+	yok
compete	-	-	-	+	+	var
cycle	+	+	-	-	-	yok

+: PH varsayımının ihlal edildiğini, **-**: PH varsayımının sağlandığını göstermektedir.

Zamana bağı etkileri dikkate alan nihai modelde, (**export**), (**sektor**) ve (**cycle**) değişkenleri anlamsız çıkmıştır. Bu değişkenler zamana bağı değişkenleri içermeyen Model 12'de de anlamsız çıkmıştır. Firmanın ihracat yapıp yapmamasının, üretim veya hizmet sektöründe olmasının ve ekonomik kriz atlatıp atlatmamasının hayatta kalma süresi üzerinde anlamlı bir etkisi bulunmamaktadır.

Geliri sadece mevcut firmadan olan girişimciye sahip firmaların başarısızlık ihtimali, zaman etkilerini dikkate almayan modeldeki gibi diğer firmaların yaklaşık 17 katıdır. Girişimcisi bayan olan firmaların, girişimcisi erkek olan firmalardan yaklaşık 270 ($e^{9.850+(-4.250)}$) kat fazla başarısızlık riskine sahip olduğu söylenebilir. Girişimcinin yaşı arttıkça firmanın hayatta kalma ihtimalinin azaldığı görülmektedir. Girişimcisi üniversite mezunu olan firmaların başarısızlık riski diğer firmaların yaklaşık 26 ($e^{5.820+(-2.560)}$) katıdır. Girişimcinin çalışma tecrübesi ve firmadaki ortak sayısı arttıkça, firmaların başarısızlık ihtimalinin azaldığı söylenebilir. Aile veya çevresinden girişimcilik yönünden etkilenen girişimcilere ait firmaların hayatta kalma riskinin diğer firmalarınkinden yaklaşık 37 ($e^{5.430+(-1.820)}$) katı olduğu bulunmuştur. Kuruluştaki firma büyüklüğünün hayatta kalma ihtimalini yaklaşık 5 kat ($e^{2.520+(-0.888)}$) olumsuz etkilediği tespit edilebilir. Girişimin kuruluşundaki sermaye tamamıyla kendisine ait olan girişimcilerin firmalarının başarısızlık riski, diğerlerinin yaklaşık 5 katı olduğu bulunmuştur. Girişimci kuluçka içinde ve dışında paydaşlarla işbirliği içerisinde ise, firmanın hayatta kalma ihtimali arttığı bu modelle de bulunmuştur ancak diğer modelin aksine (37 kat) değişkenin hayatta kalma ihtimali üzerine etkisinin az ($e^{5.430} \cong 0.01$) olduğu görülmektedir. Yine firmanın yenilik yapmasının hayatta kalma ihtimalini arttırdığı söylenebilir. Reklam yapan firmaların hayatta kalma ihtimalinin yaklaşık 22 ($e^{4.660+(-1.580)}$) kat daha fazla olduğu bulunmuştur. Marka sahibi olmanın ise firmanın hayatta kalma üzerine çok fazla olmamakla birlikte olumlu bir etkisinin olduğu bulunmuştur. Kuluçka hizmetlerinden yararlanan firmaların hayatta kalma ihtimali, diğer firmaların yaklaşık 14 katıdır. Kuluçkanın son 36 ay içinde kurulmuş olan bir ilde olması firmaların hayatta kalma ihtimali diğerlerinin yaklaşık 50 katı kadar arttırmaktadır. Kuluçkadaki işlik sayısı ve kuluçkanın bulunduğu ilin KBGSMH'sının payı arttıkça başarısızlık riskinin azaldığı görülmektedir. Son olarak rekabet yoğunluğu fazla olan firmaların başarısız olma ihtimali diğerlerinin yaklaşık 4 ($e^{2.170+(-0.731)}$) katıdır.

Sonuç olarak (**workexp**), (**ortaksay**), (**networking**), (**yenilik**), (**marka**), (**buro**), (**whenest**), (**incubsize**), (**ilrank**) firmaların hayatta kalma ihtimalini olumlu olarak etkileyen değişkenler olarak; (**tekgelir**), (**gender**), (**lnentage**), (**enteduuni**), (**ailecev**), (**lnempini**),

(onlyselffin), **(reklam)**, **(compete)** firmaların başarısız olma ihtimalini arttıran deęişkenler olarak bulunmuştur.

SONUÇ VE ÖNERİLER

Cox orantılı tehlike modeli, tercih edilen pek çok özelliklerinin yanında temel bir varsayım olan tehlikelerin orantılı olmasına dayalıdır. Değişken-zaman etkileşiminin parametrik olarak modellenmesinde ise, değişimin gerçek yapısının seçilen fonksiyonun biçimi ile tutarlı olmaması durumunda yanlış tahminler elde edilebilmektedir.

Zamanla değişen etkileri araştırma, Cox hayatta kalma analizlerinin ayrılmaz bir parçası olmalıdır. Zamana bağlı etkilerin tespit edilmesi ve hesaplanması, dikkate alınmadığı takdirde gözden kaçırılabilen bazı özel zaman kalıplarını ve değerli bilgileri kavramayı sağlamaktadır.

Uygulamada, öncelikle tüm açıklayıcı değişkenlerin zaman içinde sabit olduğu düşünülerek Cox regresyon analizi yapılmıştır. Daha sonra, açıklayıcı değişkenlerden zamana bağlı olanlar modele dahil edilerek genişletilmiş Cox regresyon analizi yapılmıştır. Zamanla değişen açıklayıcı değişken olduğunda sabit değişkenli Cox regresyon analizi yerine zamana bağlı açıklayıcı değişkenli Cox regresyon analizinin yapılması gerekliliği vurgulanmaya çalışılmıştır. Elde edilen katsayılar ile İŞGEM’lerde yer alan firmaların hayatta kalma ihmali üzerinde etkili olan zamandan bağımsız ve zamana bağlı değişkenler tespit edilmiştir.

Firmanın ihracat yapıp yapmamasının, imalat sanayinde veya hizmet sektöründe bulunmasının ve bir kriz zamanını tecrübe edip etmemesinin hayatta kalma süresi üzerinde anlamlı bir etkisi olmadığı bulunmuştur.

Girişimcinin cinsiyetinin, eğitim durumunun, aile çevresinden etkilenmesinin; kuruluştaki firma büyüklüğünün, firmanın reklam yapıp yapmamasının, firmanın marka sahibi olup olmamasının; firmanın bulunduğu sektördeki rekabet yoğunluğunun hayatta kalma üzerine etkisinin zamana bağlı olduğu saptanmıştır.

Girişimci özelliklerinden, girişimcinin yıl olarak çalışma tecrübesi firmanın hayatta kalma ihtimalini arttırmaktadır. Firma özelliklerinden; firmadaki ortak sayısının artmasının hayatta kalma ihtimalini arttırdığı söylenebilmektedir. Girişimcinin kuluçka içindeki veya dışındaki paydaşlarla işbirliği içinde olmasının ve firmanın yenilik yapmasının hayatta kalma ihtimali üzerinde olumlu etkisi bulunmaktadır. Firmanın marka sahibi olması da hayatta kalma ihtimalini arttırmaktadır ancak bu etki zamanla azalmaktadır.

Kuluçkadaki büro hizmetlerinden yararlanmanın ve firmanın bulunduğu kuluçkanın kurulduğu ilin son 36 ay içinde il olmasının da hayatta kalma ihtimalini arttırdığı söylenebilmektedir. Aynı zamanda kuluçkadaki işlik sayısı arttıkça firmanın hayatta kalma

ihtimali de artmaktadır. Kuluçkanın bulunduğu ilin KBGSMH'sının Türkiye'nin KBGSMH'sı içindeki payının artması da firmanın başarısızlık riskini azaltmaktadır.

Girişimcinin tüm geliri sadece genç firmadan ise firmanın hayatta kalma ihtimali de azalmaktadır. Girişimcileri arasında en az bir bayan olan firmaların başarısızlık riski diğerlerinden oldukça yüksektir. Ancak cinsiyetin zamana bağlı olan etkisi göz önüne alınırsa, bayan girişimciye sahip firmaların başarısızlık riskinin zaman içinde diğer firmalardan daha az arttığı söylenebilmektedir.

Girişimcinin yaşı arttıkça firmanın başarısızlık riski de artmaktadır. Girişimcisi en az üniversite mezunu olan firmaların hayatta kalma ihtimali diğerlerinden düşüktür fakat hayatta kalma ihtimali üzerindeki bu olumsuz etki zaman içinde azalmaktadır. Girişimcinin aile veya çevresinde deneyimlerinden yararlandığı kişi veya kişilerin olması başlangıçta firmanın başarısızlık riskinin yüksek olmasına yol açmaktadır. Ancak aile çevresi etkisi başarısızlık riskini zaman geçtikçe daha az arttırmaktadır.

Kuruluştaki firma ölçeğinin büyüklüğü, zamanla azalarak da olsa firmanın başarısızlık riskini arttırmaktadır. Girişimin kuruluşundaki sermayenin tamamıyla girişimcinin kendisine ait olması da firmanın başarısızlık riskini arttırmaktadır. Firmanın reklam yapması da giderek azalan oranda hayatta kalma ihtimalini düşürmektedir. Firmanın bulunduğu sektördeki rekabet yoğunluğu arttıkça başarısızlık riski artmaktadır. Fakat firmanın yaşam süresi arttıkça sektördeki rekabetten daha az etkilenmektedir.

Kuluçkada yer alan firmaların hayatta kalma ihtimali artırmak için iş tecrübesi yüksek olan firma dışından da sabit gelire sahip olan girişimcilere sahip olması gerekmektedir. Firmaların çok ortaklı olması da firmanın hayatta kalmasına olumlu etkide bulunacaktır. Hayatta kalma ihtimalini arttırmak için kuluçka içinde ve dışında paydaşlarla işbirliği içinde olunmalı, kuluçkanın olanaklarından etkin biçimde yararlanılmaya çalışılmalı, yenilik faaliyetlerine önem verilmeli, marka değeri oluşturulmaya çalışılmalıdır.

Firmanın dahil olacağı kuluçkanın bulunduğu ilin şartları değerlendirilmeli, KBGSMH'sı yüksek olan ve sübvansiyonlardan yararlanma imkanı olan iller ve işlik sayısı fazla olan kuluçkalar tercih edilmelidir. Kuruluş aşamasında sermaye için finansal borç kaynakları bulunmalıdır. Ayrıca rekabetin az olduğu sektörlere yönelmek de firmaların hayatta kalma ihtimalini arttırmaya yardımcı olacaktır.

İŞGEM'lerde sunulan hizmetler firmaların hayatta kalma olasılıkları üzerine olumlu etkilere sahip olduğu görülmektedir (Karaöz, Albeni ve Demirgil, 2011, s.81). Bu nedenle, kuluçka sayısının ve niteliklerinin artırılması büyük önem taşımaktadır.

KAYNAKÇA

- Aernoudt, R, (2004),“Incubators: Tool for Entrepreneurship?”, *Small Business Economics*, 23, s,127.
- Ata N. ve Sözer M. T.; “Cox Regression Models With Nonproportional Hazards Applied to Lung Cancer Survival Data”, *Hacettepe Journal of Mathematics and Statistics*, No: 36 (2), (2007); 157-167.
- Başar E., “Orantılı Olmayan Tehlike Üzerine Bir Çalışma”, 5. İstatistik Günleri Sempozyumunda sunulan bildiri, Ankara Üniversitesi Fen Fakültesi İstatistik Bölümü, Antalya, (Mayıs 2006); 111-116.
- Bellera C.A., MacGrogan G., Debled M., De Lara C.T., Brouste V. ve Mathoulin-Pélissier S., “Variables with time-varying effects and the Cox model: Some statistical concepts illustrated with a prognostic factor study in breast cancer”, *BMC Medical Research Methodology*, No. 10:20, (Mart 2010), 1-12.
- Collett D., *Modelling Survival Data in Medical Research*, London: Chapman & Hall; 1994.
- Collett D., *Modelling Survival Data in Medical Research, Second Edition*, Chapman Hall, London, 2003.
- Cox D.R. ve Oakes D., *Analysis of Survival Data*, Chapman and Hall, London, 1984.
- Cox D.R., “Regression Models and Life-Tables”, *Journal of the Royal Statistical Society, Series B*, Cilt:34, No:2, (1972); 187-220.
- Çakır B., “Genetik Epidemiyoloji: Psikiyatrik Araştırmalarda Yeri ve Kullanımı”, *Türk Psikiyatri Dergisi*, No: 13(4), (2002); 300-311.
- Demirgil H., “Firmaların Hayatta Kalma ve Büyüme Performanslarını Belirleyen Faktörler: Göller Bölgesi Üzerine Bir Araştırma”, Süleyman Demirel Üniversitesi Sosyal Bilimler Enstitüsü Doktora Tezi, 2008.
- Elandt-Johnson R.C. ve Johnson N.L., *Survival Models and Data Analysis*, Wiley Classics Library Edition, John Wiley & Sons, Inc., New York, 1999.
- Gehan E.A.; “A generalized Wilcoxon test for comparing arbitrarily singly censored samples.”, *Biometrika*, 1965, 52; 203-223.

Geiss K., Meyer M., Radespiel-Tröger M. ve Gefeller O., “SURVSOFT—Software for nonparametric survival analysis”, *Computer Methods and Programs in Biomedicine*, Elsevier Ireland LTD., 96, (Nisan2009); 63–71.

Göz Çekçeki Ö., “Orantılı Hazard Modelinin Zamana Bağlı Değişkenlerle Genişletilmesi ve Çocuk Suçluluğu Üzerine Bir Uygulama”, Gazi Üniversitesi Fen Bilimleri Enstitüsü Yüksek Lisans Tezi, 2007.

Hackett M, ve Dilts D.M., (2004), “A Systematic Review of Business Incubation Research”, *Journal of Technology Transfer*, 29, ss, 55-82.

Harrell F.E. Jr., *Regression Modelling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Models*, Springer-Verlag New York, Inc., New York, 2001.

Holmes P., Hunt A. ve Stone I., “An analysis of new firm survival using a hazard function”, *Applied Economics*, No: 42, 2010, 185–195.

Hosmer D.W., Lemeshow S. ve May S., *Applied Survival Analysis: Regression Modelling Time-to-Event Data*, Second Edition, Wiley, Hoboken, 2008.

Hougaard, P., *Analysis of Multivariate Survival Data*, Springer-Verlag, New York. 2000.

Kaplan E.L. ve Meier P., “Nonparametric Estimation From Incomplete Observations,” *Journal of the American Statistical Association*, No: 53, 1958, 457-481.

Karaöz M. ve Albeni M., "İş Kuluçkalarında Yeni Kurulan Girişimlerin Hayatta Kalma ve Büyüme Performansını Etkileyen Faktörler: KOSGEB İş Geliştirme Merkezleri (İŞGEM) Üzerine Bir Araştırma", 109K139 nolu TÜBİTAK Projesi, Isparta, Mart 2011.

Karaöz M., Albeni M. ve Demirgil H., "İş Kuluçkalarında Yeni Kurulan Girişimlerin Hayatta Kalmalarını Etkileyen Faktörler", *Finans & Politik ve Ekonomik Yorumlar*, No: 557, (Temmuz 2011); 69-84.

Kleinbaum D.G. ve Klein M., *Survival Analysis: A Self-Learning Text*, Second Edition, Springer, New York, 2005.

Le C.T., *Applied Survival Analysis*, John Wiley & Sons, Inc., New York, 1997.

Mantel N.; “Evaluation of survival data and two new rank order Statistics arising in its consideration”, *Cancer Chemother Rep*, 1966, 50; 163-170.

Moreau T, O’Quigley J, Mesbah M, “A global goodness of-fit statistics for the proportional hazards model”, *Appl. Stat.*, 1985, No: 34, 212-218.

Neffke F.M H., Henning M. ve Boschma R., “The impact of aging and technological relatedness on agglomeration externalities: a survival analysis”, *Journal of Economic Geography*, Mart 2011, s. 1-33.

Oakes D., “Survival Analysis”, *Journal of the American Statistical Association*, Vol. 95, No. 449, (Mart 2000); 282-285.

Parker S., Peters G.F. ve Turetsky H.F., “Corporate Governance and Corporate Failure: A Survival Analysis”, *Corporate Governance*, 2.02.2002, 4-12.

Pettitt A.N., Daud I.B., “Investigating time dependence in Cox’s proportional hazards model”, *Applied Statistics*, 1990; 39: 313-329.

Rose C.E. Jr., Hall D. B., Shiver B.D., Clutter M.L. ve Borders B., “A Multilevel Approach to Individual Tree Survival Prediction”, *Forest Science*, No: 52(1) (Nisan 2006), s. 31-43.

Scheike T.H., *Advances in Survival Analysis*, der. Balakrishnan N. ve Rao C.R., Elsevier North-Holland, Amsterdam, 2004.

Sertkaya D., Ata N. ve Sözer M. T., “Yaşam çözümlemesinde zamana bağlı açıklayıcı değişkenli Cox regresyon modeli”, *Ankara Üniversitesi Tıp Fakültesi Mecmuası*, No: 58, (Aralık 2005); 153-158.

Tabatabai M.A., Bursac Z., Williams D.K. ve Singh K. P., “Hypertabastic survival model”, *Theoretical Biology and Medical Modelling*, 4:40, 2007; 1-13.

Therneau T.M. ve Grambsch P.M., “Modelling Survival Data: Extending the Cox Model”, New York, Springer; 2000.

Vittinghoff E., Glidden D.V., Shiboski S.C. ve McCulloch C.E., *Regression Methods in Biostatistics: Linear, Logistic, Survival and Repeated Measures Models*, Springer, New York, 2005.

Yay M., Çoker E. ve Uysal Ö., “Yaşam Analizinde Cox Regresyon Modeli ve Artıkların İncelenmesi”, *Cerrahpaşa Tıp Dergisi*, No: 38, (2007); 139-145.

Ö Z G E Ç M İ Ş

Adı ve Soyadı : Aygöl ANAVATAN

Doğum Tarihi ve Yeri : Kırcaali / 13.01.1986

Medeni Durumu : Bekar

Eğitim Durumu

Mezun Olduğu Lise : Konak Anadolu Lisesi - 2004

Lisans Diploması : Dokuz Eylül Üniversitesi İktisadi ve İdari Bilimler Fakültesi
Ekonometri Bölümü - 2008

Yabancı Dil / Diller : İngilizce

İş Deneyimi

Stajlar : Goksu Food Ltd. Şti. - 2007 (1 ay)
Dimes İzmir Fabrikası - 2006 (20 gün)

Çalıştığı Kurumlar : Akdeniz Üniversitesi Sosyal Bilimler Enstitüsü Ekonometri
ABD - Araştırma Görevlisi (25.08.2010 - Devam)
Dokuz Eylül Üniversitesi Eğitim Bilimleri Enstitüsü - Memur
(04.09.2009 - 24.08.2010)