

**GÖZLEMLENEMEYEN RİSK FAKTÖRLERİNİN ÖLÜMLÜLÜK
ÜZERİNDEKİ ETKİSİ**

THE EFFECT OF UNOBSERVED RISK FACTORS ON MORTALITY

FUNDA KUL

Hacettepe Üniversitesi

Lisansüstü Eğitim – Öğretim ve Sınav Yönetmeliğinin

AKTÜERYA BİLİMLERİ Anabilim Dalı İçin Öngördüğü

YÜKSEK LİSANS TEZİ

olarak hazırlanmıştır.

2011

Fen Bilimleri Enstitüsü Müdürlüğü'ne,

Bu çalışma jürimiz tarafından **AKTÜERYA BİLİMLERİ ANABİLİM DALI** 'nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Başkan


.....
(Yrd. Doç. Dr. Şeref HOŞGÖR)

Üye (Danışman)


.....
(Doç. Dr. Meral SUCU)

Üye


.....
(Dr. Murat BÜYÜKYAZICI)

ONAY

Bu tez Hacettepe Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliği'nin ilgili maddeleri uyarınca yukarıdaki jüri üyeleri tarafından/...../..... tarihinde uygun görülmüş ve Enstitü Yönetim Kurulunca/...../..... tarihinde kabul edilmiştir.

Prof.Dr. Adil DENİZLİ

Fen Bilimleri Enstitüsü Müdürü

GÖZLEMLENEMEYEN RİSK FAKTÖRLERİNİN ÖLÜMLÜLÜK ÜZERİNDEKİ ETKİSİ

Funda KUL

ÖZ

Aktüerler, sigortalı popülasyonun aynı risk faktörlerine maruz kalan bireylerden oluşmadığının, diğer bir deyişle, sigortalı popülasyonun homojen bir yapıya sahip olmadığının farkındadırlar. Bu heterojenlik, popülasyonun ölüm zamanı verisi kullanılarak elde edilen ölüm oranları arasında bağımlılığa neden olmaktadır.

Hayat sigortası ürünlerinin doğru fiyatlandırılabilmesi için popülasyonun ölümlülüğüne etki eden heterojenlik yapısının iyi anlaşılması gerekmektedir. Sigortalı popülasyonun heterojenliğinin temel olarak iki kaynağı bulunmaktadır. Bunlar; gözlemlenebilir ve gözlemlenemeyen risk faktörlerinden kaynaklanan heterojenliktir. Gözlemlenebilir risk faktörlerinden kaynaklanan heterojenlik poliçe düzenlenirken risk kabul sürecinde dikkate alınabilirken, gözlemlenemeyen risk faktörlerinden kaynaklanan heterojenlik ise dikkate alınamamaktadır. Gözlemlenemeyen risk faktörlerinden kaynaklanan heterojenlik risk faktörüne ilişkin bilginin bulunmaması veya var olan bilginin yetersiz olmasından kaynaklanmaktadır.

Bu çalışmada, Türkiye’de faaliyet gösteren 5 büyük sigorta şirketinden elde edilen ölüm verisi cinsiyete göre ayrılmış ve heterojenlik yapısının belirlenebilmesi için hassasiyet modelleri incelenmiştir. Hassasiyet modeli parametrelerinin tahmini için ise Genelleştirilmiş Doğrusal Modelleme Yöntemi kullanılmıştır. Elde edilen sonuçlara göre kadın ve erkek sigortalı veri kümeleri için heterojenliğin etkisinin önemli olduğu görülmüştür.

Anahtar Kelimeler:

Hassasiyet Modelleri, Risk Faktörleri, Ölüm Hızı, Gompertz Ölümlülük Yasası, Genelleştirilmiş Doğrusal Modeller.

Danışman: Doç. Dr. Meral SUCU, Hacettepe Üniversitesi, Aktüerya Bilimleri Bölümü, Aktüerya Bilimleri Anabilim Dalı

THE EFFECT OF UNOBSERVED RISK FACTORS ON MORTALITY

Funda KUL

ABSTRACT

Actuaries know that people, who bring into being insured population, does not exposure same risk factors, in other words, insured population is not homogeneous. This inhomogeneity, causes dependence between mortality rates which is obtained from population mortality time data.

In order to fairly price life insurances, a better understanding of the extent of heterogeneity in population mortality is required. Basically, there are two reasons for heterogeneity in insured population. They are heterogeneity causes from observable risk factors and heterogeneity causes from unobservable risk factors. As, heterogeneity causes from observable risk factors is taken into consideration at underwriting process when issuing a policy, heterogeneity causes from unobservable risk factors is not taken. Heterogeneity causes from unobservable risk factors results from being uninformed or not having enough knowledge about risk factor.

This paper applies well established frailty models to quantify the extent of heterogeneity in Turkey insured population death data which is taken from 5 active insurance company in Turkey and seperated by gender. To estimate parameters is used Generalized Linear Modelling Method. The results confirm significant heterogeneity effect exists for male and female insured population.

Keywords:

Frailty Models, Risk Factors, Force of Mortality Rate, Gompertz Mortality Law, Generalized Linear Models.

Advisor: Assoc. Prof. Dr. Meral SUCU, Hacettepe University, Department of Actuarial Sciences, Actuarial Sciences Section.

TEŞEKKÜR

Tez çalışmamın her aşamasında değerli katkı ve eleştirileriyle bana yol gösteren, karşılaştığım zorlukların aşılmasında yardımcı olan danışmanım Sayın Doç. Dr. Meral SUCU'ya,

katkılarından dolayı Zoltan Butt, Steven Haberman ve Annamaria Olivieri'ye,

çalışma süresince yardım ve hoşgörülerini esirgemeyen tüm hocalarıma ve çalışma arkadaşlarıma,

teşvik ve desteklerini hiç esirgemeyen Arş. Gör. Ezgi NEVRUZ, Arş. Gör. Başak BULUT ve Arş. Gör. Tuğba TUNÇ'a,

iyi zamanlarımı daha iyi, zor anlarımı daha kolay hale getiren tüm arkadaşlarıma,

her zaman yanımda olan, bana güç veren, sonsuz anlayışları, sevgileri, hoşgörülerini ve destekleri ile bu günlere gelebildiğim annem Kadiriye KARAMAN, babam Şerafettin KARAMAN, erkek kardeşim Ferhat KARAMAN, kız kardeşlerim Melek Gizem KARAMAN ve Sinem KARAMAN'a,

hayatımın en güzel anlarını paylaştığım, her anımda yanımda olan, bana her konuda destek veren birtanecik eşim Serdar KUL'a

en içten teşekkürlerimi sunarım.

İÇİNDEKİLER DİZİNİ

Sayfa

ÖZ	i
ABSTRACT	ii
TEŞEKKÜR	iii
İÇİNDEKİLER DİZİNİ.....	iv
ŞEKİLLER DİZİNİ	vii
ÇİZELGELER DİZİNİ.....	ix
SİMGELER VE KISALTMALAR DİZİNİ	x
1. GİRİŞ ve ÖNCEKİ ÇALIŞMALAR.....	1
1.1. Giriş.....	1
1.2. Önceki Çalışmalar	5
2. YAŞAM ANALİZİ.....	8
2.1. Giriş.....	8
2.2. Yaşam Analizinde Kullanılan Fonksiyonlar.....	8
2.3. Sansürleme ve Kesme	10
2.3.1. Sansürleme	10
2.3.2. Kesme	12
2.4. Yaşam Modellemesi	14
2.4.1. Parametrik yaklaşım	15
2.4.1.1. Parametreler yardımıyla modelleme	15
2.4.1.2. Bilinen dağılımlar yardımıyla modelleme	21
2.4.2. Parametrik Olmayan Yaklaşım	22
2.5. Hassasiyet Modelleri	24
3. HASSASIYET MODELLERİ.....	26
3.1. Giriş.....	26
3.2. Hassasiyet Dağılımlarının Temel Özellikleri	27
3.3. Hassasiyet Model Türleri.....	30
3.3.1. Kesikli yaklaşım.....	30

3.3.2.	Sürekli yaklaşım	34
3.4.	Tek Değişkenli Hassasiyet Modelleri	35
3.4.1.	Gamma dağılımı	38
3.4.2.	Pozitif Durağan dağılım	40
3.4.3.	Ters Gauss dağılımı	41
3.4.4.	Güç Varyans Fonksiyonu dağılımı	42
3.4.5.	Bileşik Poisson dağılımı	43
3.5.	Çok Değişkenli Hassasiyet Modelleri	44
3.5.1.	Paylaşımlı hassasiyet modelleri	45
3.5.2.	İlişkili hassasiyet modelleri	47
4.	UYGULAMA	48
4.1.	Modelleme	48
4.2.	Kaba Ölüm Hızlarının Elde Edilmesi	49
4.2.1.	2004-2008 yılları arasındaki Türkiye sigortalı verisi	49
4.2.2.	Eksik ve hatalı veri sorununun ortadan kaldırılması	49
4.2.3.	Tekrarlama hatasının ortadan kaldırılması	49
4.2.4.	Riske maruz kalan birim sayısının elde edilmesi	49
4.2.5.	Ölen kişi sayısının bulunması	51
4.2.6.	Gözlenen ölüm hızının elde edilmesi	51
4.3.	Parametre Tahmini İçin Kullanılacak Yöntem	53
4.3.1.	Genelleştirilmiş Doğrusal Modeller	54
4.3.1.1.	Giriş	54
4.3.1.2.	GLM'in yapısı	54
4.3.1.3.	GLM için olabilirlik fonksiyonu	54
4.3.1.4.	GLM için artıkların incelenmesi	60
4.3.2.	Taylor açılımı ile GLM	61
4.4.	Uyum İyiliği Testleri	63
4.5.	Kadın ve Erkek Sigortalı Veri Setinin İncelenmesi	64

4.5.1. Kadın sigortalı veri setinin incelenmesi.....	64
4.5.2. Erkek sigortalı veri setinin incelenmesi	70
4.6. Ölüm Hızlarının Uzatılması.....	75
5. SONUÇ VE ÖNERİLER.....	77
5.1. Sonuçlar	77
5.2. Öneriler	84
KAYNAKÇA.....	85
EKLER DİZİNİ	93
ÖZGEÇMİŞ	103

ŞEKİLLER DİZİNİ

Sayfa

Şekil 2.1 Sansürleme.....	11
Şekil 2.2 Kesme	13
Şekil 4.1 Sansürleme ve Kesme.....	50
Şekil 4.2 Kadın Sigortalı Bireyler İçin Yaşa Göre Ölüm Hızları.....	52
Şekil 4.3 Erkek Sigortalı Bireyler İçin Yaşa Göre Ölüm Hızları.....	52
Şekil 4.4 Kadın Sigortalı Bireyler İçin Yaşa Göre Gözlenen ve Modele Göre Ölüm Hızları	65
Şekil 4.5 Kadın Sigortalı Ölüm Hızları İçin Artıkların İncelenmesi	66
Şekil 4.6 Kadın Sigortalılar İçin Yaşa Göre Gözlenen ve İkinci Modele Göre Ölüm Hızları	67
Şekil 4.7 İkinci Modellemede Kadın Sigortalı Ölüm Hızları İçin Artıkların İncelenmesi	68
Şekil 4.8 Erkek Sigortalı Bireyler İçin Yaşa Göre Gözlenen ve Beklenen Ölüm Hızları	70
Şekil 4.9 Erkek Sigortalı Ölüm Hızları İçin Artıkların İncelenmesi.....	71
Şekil 4.10 Erkek Sigortalı Bireyler İçin İkinci Modellemede Yaşa Göre Gözlenen ve Beklenen Ölüm Hızları.....	72
Şekil 4.11 İkinci Modelleme Erkek Sigortalı Ölüm Hızları İçin Artıkların İncelenmesi	73
Şekil 5.1 Kadın Sigortalılar İçin Hassasiyet Etkisi.....	77
Şekil 5.2 Erkek Sigortalılar İçin Hassasiyet Etkisi.....	78
Şekil 5.3 (25)-(50) Yaş Aralığı İçin Kadın Sigortalı Ölüm Hızlarının TRSH-2010 ve Hassasiyet Modellemesine Göre Karşılaştırılması	79
Şekil 5.4 (51)-(80) Yaş Aralığı İçin Kadın Sigortalı Ölüm Hızlarının TRSH-2010 ve Hassasiyet Modellemesine Göre Karşılaştırılması	80
Şekil 5.5. (81)-(110) Yaş Aralığı İçin Kadın Sigortalı Ölüm Hızlarının TRSH-2010 ve Hassasiyet Modellemesine Göre Karşılaştırılması	81

Şekil 5.6. (25)-(50) Yaş Aralığı İçin Erkek Sigortalı Ölüm Hızlarının TRSH-2010 ve Hassasiyet Modellemesine Göre Karşılaştırılması	81
Şekil 5.7. (51)-(80) Yaş Aralığı İçin Erkek Sigortalı Ölüm Hızlarının TRSH-2010 ve Hassasiyet Modellemesine Göre Karşılaştırılması	82
Şekil 5.8. (81)-(110) Yaş Aralığı İçin Erkek Sigortalı Ölüm Hızlarının TRSH-2010 ve Hassasiyet Modellemesine Göre Karşılaştırılması	82

ÇİZELGELER DİZİNİ

Sayfa

Çizelge 4.1. Üstel Dağılım Ailesi İçin Fonksiyonlar..... 57

Çizelge 4.2. Bağ Fonksiyonları ve Bağ Fonksiyonlarının Ters Fonksiyonları..... 58

SİMGELER VE KISALTMALAR DİZİNİ

PH Modeli	Bölüşmeli Tehlike Modeli
AFT Modeli	Hızlandırılmış Başarısızlık Zamanı Modeli
GLM	Genelleştirilmiş Doğrusal Modeller
He	Heterojen
Ho	Homojen
HPC Yöntemi	İlişkisel Yöntem
TRSH	Türkiye Sigortalı Hayat Tablosu

1. GİRİŞ ve ÖNCEKİ ÇALIŞMALAR

1.1. Giriş

Risk, gerçekleşmesi kesin olmayan; ancak gerçekleştiğinde, zarar veya kayıp durumlarına yol açabilecek bir olayın ortaya çıkması anlamına gelir. Hasta olma riski, trafik kazası yapma riski ya da faiz riski genel anlamıyla risk kavramına örnek olarak verilebilir. Hasta olma riski, ilaç ya da aşı gibi korunma araçları ile en aza indirgenebileceği gibi, maddi zararlara yol açabilecek risklerden de sigorta ürünleri satın alınarak korunulabilir.

Sigorta, kişilerin karşılaşılabilecekleri zarar ve gelir kaybına yol açabilecek olayların ekonomik sonuçlarından korunmak için aynı riski paylaşan bireylerin risklerini, belli bir prim karşılığında transfer etme sistemidir. Yangın, deprem, trafik kazası gibi çeşitli risklerin ev ya da otomobil üzerinde yol açabileceği finansal zararlardan korunmak için yangın sigortası, doğal afet sigortası, kasko ya da trafik sigortası yaptırabileceği gibi kişilerin hayatlarını kaybetmesi ya da malul olması gibi durumlarda geride bırakılacak aile ya da lehdarların finansal kayba uğraması riskinden korunmak için de hayat sigortası yaptırabilir (Yıldırım, 2010). Gelişmiş ülkelerde hayat sigortaları sadece riskten korunma amaçlı değil yatırım yapma amacıyla da kullanılmaktadır (Atkinson ve Dallas, 2000).

Hayat sigortası yaptırmadan önce bireyin risk kabul sürecinden (underwriting) geçmesi gerekmektedir. Risk kabul sürecinde; yaş, cinsiyet, yaşanılan bölge gibi risk faktörleri dikkate alınarak bireyin ait olduğu risk sınıfı belirlenir ve bu birey için o risk sınıfına ilişkin prim oranı uygulanır. Cummins et al. (1983), hayat sigortaları için risklerin sınıflandırılmasını ayrıntılı olarak incelemişlerdir. Günümüzde, sigorta şirketleri 'Sayısal Fiyatlandırma Sistemi (Numerical Rating System)' olarak adlandırılan yöntemi kullanarak risk sınıflandırması yapmaktadır. Bu yöntem, ilk olarak 1919 yılında Amerika'da uygulanmaya başlanmıştır (Olivieri, 2006).

Hayat sigortalarında aktüeryal denge eşitliği ilkesine göre prim hesaplanırken iki temel varsayım kullanılmaktadır. Bunlar; faiz oranı varsayımı ve hayat tablosu varsayımıdır. Faiz oranı varsayımı, teminatın zaman değerinin sürekli değişmesi nedeniyle, hayat tablosu varsayımı ise yaşa göre ölümlülük eğiliminin ileriye yönelik analizinin yapılabilmesi için kullanılmaktadır. Hayat tabloları, her yaşa göre yaşayan

kiři sayısı, ölen kiři sayısı ve ölüm oranlarının bulunduđu tablolardır. Tosetti et al. (2000), hayat sigortası poliçelerinin fiyatlandırılmasında ölüm oranları ve faiz oranlarının yanlış seçiminin etkilerini ayrıntılı olarak incelemişlerdir. Birçok hayat sigortası ürününün finansal yapısı önemli olmasına rağmen son yıllarda ölümlülük eğilimine bağılı olarak ölüm oranlarının değışiminin incelenmesi daha büyük önem taşımaktadır (Pitacco, 2003). Bu nedenle hayat sigortalarının değerlemesinde ölüm oranlarının zaman veya yaşa göre değışiminin analiz edilmesi oldukça önem kazanmıştır.

Olivieri ve Pitacco (2002) ölümlülük eğilimine ilişkin aşağıdaki üç durumun gözlemlendiğini belirtmişlerdir :

- Genişleme (Expansion)
- Dikdörtgenleşme (Rectangularization)
- Ölümlülük dalgalanması (Mortality fluctuation)

Genişleme, yaşam eğrisinin en büyük yaşının daha ileri yaşlara uzaması; dikdörtgenleşme, yaşam eğrisinin ileri yaşlarda yoğunlaşması, erken yaşlarda yoğunluğunun azalması ve ölümlülük dalgalanması ise, ölüm oranlarının yaştan yaşa farklı şekillerde sapma göstermesi olarak açıklanmaktadır. Yukarıda açıklanan ölümlülük eğilimlerinden genişleme ve dikdörtgenleşmenin modellenenebilmesi için uzun zaman boyunca kaydedilmiş ölüm kesit verisine ihtiyaç duyulurken ölümlülük dalgalanmasının modellenenebilmesi için ise kısa zaman boyunca kaydedilmiş ölüm kesit verisi yeterli olmaktadır.

Ölümlülük dalgalanmalarına neden olan riskler farklı yaklaşımlar kullanılarak modellenenebilmektedir. Bu modelleme yaklaşımlarının bazılarında uzun ömürlülüğün (longevity) artma eğilimi incelenirken diğeri modelleme yaklaşımlarında ise ölümlülük dalgalanmalarının yapısı incelenmektedir (Lee ve Carter, 1992; Pitacco, 2003).

Ölümlülük dalgalanmasının dört temel nedeni vardır. Bunlardan ilki, ölüm tablosunun ya da olasılıksal ölümlülük modelinin yanlış olarak belirlenmesidir. Bir diğeri neden ise popülasyondaki gözlem sayısının yeterli olmamasıdır. Çünkü gözlem sayısı az olan popülasyonlarda Büyük Sayılar Kanunu doğrulanamamakta ve bu yüzden bu popülasyonlarda rastgele sapmaların aksine sistematik sapmalar ortaya çıkmaktadır (Olivieri, 2002). Bir diğeri neden ise Yaş Kaçması (Age Heaping)'dir. Yaş göre

popölasyon verisi arařtırması yapılırken, bireylerin yařlarını en yakın beř ve beřin katlarına yuvarlayarak vermelerinden kaynaklanmaktadır. Son neden ise popölasyonu etkileyen risk faktörlerinin yeterli olarak belirlenememesidir. Portföy, homojen ve yeterli gözlem sayısına sahip olduėunda, öölümlölük dalgalanmaları doėal bir yapıya sahip olmaktadır. Bu dalgalanmaların kaynaėı öölümlölüėü etkileyen risk faktörleri ile açıklanabilmektedir. Ölüm oranlarına iliřkin projeksiyonun gerçeėe en uygun řekilde yapılabilmesi için ise risk faktörlerinin yeterli ve doėru bir řekilde tanımlanması gerekmektedir (Pitacco, 2004; Pitacco ve Olivieri, 2005). Fakat ölüm verisi toplanırken risk faktörlerine iliřkin ayrıntılı bilgiye ulařmak her zaman mümkün olamamaktadır. Bu nedenden dolayı, risk faktörlerine göre oluřturulmuř alt popölasyonlar, heterojen bir yapıya sahip olur. Alt popölasyonların heterojen olması ise ölüm oranları arasında baėımlılık yapısı oluřturur. Bu baėımlılık yapısı, ortalamaya baėlı tüm istatistiksel bilginin yorumlanmasını zorlařtırır (Keyfitz, 1985; Hougaard, 2000). Bundan dolayı, Yashin et al. (1995) öölümlölüėü etkileyen risk faktörlerini gözlemlenebilir (observable) risk faktörleri ve gözlemlenemeyen (unobservable) risk faktörleri olmak üzere ikiye ayırmıřlardır.

Pitacco et al. (2009) ise gözlemlenebilir risk faktörlerini temel olarak 5 ana bařlık altında incelemiřlerdir:

- i. Yař, cinsiyet gibi biyolojik ve fizyolojik özelliklere iliřkin risk faktörleri
- ii. İklim ve kirlilik gibi yařanılan çevreye iliřkin risk faktörleri
- iii. Mesleėe iliřkin risk faktörleri
- iv. Bireysel yařam řekline iliřkin risk faktörleri
- v. Saėlık durumuna iliřkin risk faktörleri

Bu risk faktörlerinden (ii) bařlıėı tüm popölasyonun öölümlölüėünü etkilemektedir. Bu nedenle hayat tabloları bölge ayırımında oluřturulur. Geriye kalan faktörler ise bireysel öölümlölüėü etkilemektedir. Hayat sigortası řirketleri, bu risk faktörlerini poliçe düzenlendiėinde gözlemleyebilmektedir. Aktüerya literatüründe, gözlemlenebilir risk faktörlerine göre oluřturulan deėerleme modellerine 'bireysel deėerleme modelleri (individual valuation models)' adı verilmektedir (Olivieri, 2006).

Gözlemlenemeyen risk faktörleri, bir risk faktörüne iliřkin yeterli bilginin bulunmaması veya var olan bilginin yetersiz olması durumunda ortaya çıkmaktadır. Gözlemlenemeyen risk faktörlerinden kaynaklanan heterojenliėin deėerlendirildiėi modellere ise aktüeryal literatürde 'birikimli deėerleme modelleri (collective valuation models)' adı verilmektedir (Olivieri, 2006).

Yaşam modellemesinde gözlemlenemeyen risk faktörlerinin etkisini inceleyen modellerden birisi, Hassasiyet Modelleri (Frailty Models)'dir. Hassasiyet modelleri, popülasyonun ölümlülüğüne etki eden gözlemlenemeyen ortak risk faktörlerinin etkisi bilindiğinde ölüm oranlarının birbirinden bağımsız olduğu varsayımını kullanmaktadır. Diğer bir deyişle; hassasiyet modelleri, ölüm oranlarının koşullu bağımsızlık özelliğine sahip olduğu düşünülerek oluşturulmuştur.

Hassasiyet modelleri, risk faktörlerine göre sınıflandırılmış bireyler için tanımlanan dayanak ölüm hızının (baseline force of mortality), negatif olmayan ve gözlemlenemeyen risk faktörlerinin etkisini gösteren bir raslantı değişkeni ile çarpılarak oluşturulmaktadır (Duchateau ve Janssen, 2008). Hassasiyet modellerindeki temel düşünce, popülasyondaki bireylerin yaşları ilerledikçe, hassasiyet seviyesi yüksek olan bireyin ölme riskinin yüksek olması nedeniyle, sistemden daha erken ayrılmasıdır. Böylece hassasiyet etkisi seviyesi yüksek olan birey, sistemden daha erken ayrılacağı için popülasyonun ortalama hassasiyet seviyesi düşecek ve hassasiyet dağılımının sol tarafına kayacaktır. Sonuç olarak, popülasyonda hayatta kalan bireyler daha düşük ortalama hassasiyet etkisi seviyesine sahip olacaktır.

Hassasiyet modelleri için popülasyonun heterojenlik yapısı, hassasiyet etkisinin dağılımı ile belirlenmektedir. Hassasiyet etkisi dağılımı ne kadar basık ise popülasyonun da o derece heterojen olduğu düşünülmektedir. Popülasyonun heterojenliğinden kaynaklanan ölüm oranları arasındaki bağımlılık ise hassasiyet etkisi dağılımının varyansı ile ölçülebilmektedir. Hassasiyet etkisi dağılımının varyansı büyük ise heterojenlikten kaynaklanan bağımlılık yapısı da güçlü olmaktadır (Duchateau ve Janssen, 2008).

Bu çalışmada, 2004-2008 yılları arasında Türkiye kadın ve erkek sigortalı verisi kullanılarak ölüm oranlarındaki dalgalanmalar, hassasiyet modelleri kullanılarak modellenmeye çalışılacaktır. Çalışmanın Birinci Bölümü'nde tezin amacı, kapsamı ve önceki çalışmalar verilmiştir. İkinci Bölüm'de ise yaşam modellerinin tarihçesi, temel özellikleri ve uygulandığı temel çalışmalar ele alınacaktır. Üçüncü Bölüm'de hassasiyet modelleri ayrıntılı olarak incelenecektir. Dördüncü Bölüm'de modele ilişkin uygun parametrelerin belirlenmesi için kullanılacak Genelleştirilmiş Doğrusal Modelleme Yöntemi ile ilgili bilgi verilerek model algoritması tanıtılacak ve 2004-2008 Türkiye Sigortalı verisine ilişkin uygun hassasiyet modeli araştırması yapılacaktır.

Ayrıca, ilişkisel yöntem kullanılarak; modelleme sonucunda elde edilen ölüm hızları için uzatma işlemi yapılacaktır. Beşinci Bölüm'de ise elde edilen sonuçlar değerlendirilecek ve 2010 Türkiye Sigortalı Hayat Tablosu (TRSH-2010) ile karşılaştırılacaktır.

1.2. Önceki Çalışmalar

Ölüm verisi için gözlemlenemeyen risk faktörlerinin etkisi, ilk olarak Beard (1959) çalışmasında kullanılmıştır. Bu çalışmada gözlemlenemeyen risk faktörlerinin etkisinin tanımlanabilmesi için uzun ömürlülük etkisi kullanılmış ve popülasyon için dayanak ölüm hızının Makeham Ölümlülük Yasası'na uyduğu kabul edilerek modelleme yapılmıştır.

Gözlemlenemeyen risk faktörlerinin etkisini tanımlamak için hassasiyet etkisi ifadesi ilk olarak, Vaupel et al. (1979) tarafından kullanılmıştır. Bu çalışmada, gözlemlenemeyen risk faktörlerinin etkisini belirlemek için Z rastlantı değişkeni tanımlanmış ve bu rastlantı değişkeninin Gamma dağılımı'na uyduğu varsayılarak hassasiyet modellemesi yapılmıştır. Oluşturulan hassasiyet modelinde her bireye özgü hassasiyet etkisi tanımlanarak bireylerin ölüm hızları arasında farklılık olmasına izin verilmiştir. Ayrıca, Z rastlantı değişkeninin farklı değerleri kullanılarak standart ölüm oranlarına ilişkin çıkarımlar yapılmıştır.

Manton et al. (1981), Vaupel et al. (1979) çalışmasında tanımlanan modele takvim yılının etkisini de ekleyerek yeniden modelleme yapmışlardır. Bu çalışmada, düzeltme yapmak amacıyla 1850-1975 yılları arasında Amerika ve İsveç için ulusal ölüm oranlarını kullanarak hassasiyet modeli oluşturulmuştur.

Vaupel ve Yashin (1985a) birbirleriyle heterojen ve kendi içlerinde homojen olan iki alt popülasyon için inceleme yapmışlardır. Bu çalışmada, bu iki alt popülasyon için ayrı ayrı ölüm hızları elde edilmiş ve bu alt popülasyonların bir bütün olarak ele alınmasının ölüm hızlarında beklenmedik sonuçlara neden olabileceği gösterilmiştir. Bunu gösterebilmek için bu iki alt popülasyonu farklı oranlarda karıştırmışlardır. Ayrıca hassasiyet etkisi için Kesikli, Tekdüze, Gamma, Weibull ve Lognormal dağılımı kullanarak farklı hassasiyet dağılımı seçiminin etkilerini incelemişlerdir.

Vaupel ve Yashin (1985b), yaş ve zaman faktörünün ölümlülük modelleri üzerindeki etkisini incelemişler ve nedene bağlı (kaza nedeniyle ölüm gibi) ölümlülük oranları için hassasiyet modelleri oluşturmaya çalışmışlardır.

Yashin et al. (1985), risk kabul sürecindeki riskleri gözlemlenebilir ve gözlemlenemeyen risk faktörleri olmak üzere ikiye ayırmış, yaş ve zaman için stokastik değişime izin veren hassasiyet modellerini incelemişlerdir.

Manton et al. (1986), Amerika Birleşik Devletleri'nin toplam ölümlülük verisini ve akciğer kanseri ölümlülük verisini kullanarak hassasiyet dağılımı için Gamma, Ters Gauss ve Dejenere dağılımı seçerek bu hassasiyet dağılımları için uyum iyiliği testi elde etmişlerdir. Sonuç olarak ise dayanak ölüm hızı seçiminin, hassasiyet dağılımının seçiminden daha önemli olduğu sonucuna ulaşmışlardır.

Jones (1988), sigortalı bireyler için risk kabul sürecinde seçilmiş ayrılma oranlarının etkisini incelemek için hassasiyet modellerini kullanmıştır.

Congdon (1993), (25)-(90) yaş aralığındaki İngiltere ölüm verisini kullanarak Gompertz/Gamma hassasiyet modeli ve ELT (English Life Tables) 12 düzeltme formüllerine olan uyumluluğu incelemiştir.

Congdon (1994), Gompertz/Gamma hassasiyet modelinde diğer parametreler sabit kaldığında Gamma dağılımının varyansındaki değişimin doğrudan beklenen yaşam süresi ile ilişkili olduğunu göstermiştir. Ayrıca bu çalışmada Gompertz/Ters Gauss hassasiyet modeli kullanılmış ve Gompertz/Gamma hassasiyet modeline ilişkin olarak parametre tahmininde farklı bir yöntem geliştirilmeye çalışılmıştır. Manton et al. (1994) tarafından belirtilen dayanak ölüm hızı seçiminin hassasiyet dağılımının seçiminden daha önemli olduğu sonucunun tersinin doğru olduğuna ilişkin bulgular elde edilmiştir.

Yashin et al. (1994), risk faktörlerine göre sınıflandırma yapılmadan hassasiyet modellemesinin doğru sonuçlar vermeyeceğini göstermişlerdir.

Damaskos (1998), İngiliz Hayat Tabloları (ELT)'nı kullanarak Gompertz/Gamma hassasiyet modeli için ağırlıklandırılmış en küçük kareler yöntemi ile modele ilişkin parametre değerlerini tahmin etmiştir.

Wang ve Brown (1998), Amerika Aktüerler Derneği (Society of Actuaries) tarafından oluşturulmak istenen 1994 Grup Hayat Annüitesi Rezerv Tablosu'nda dayanak ölüm hızı için Gompertz Ölümlülük Yasası'nı ve hassasiyet dağılımı için ise Gamma dağılımını kullanmışlardır.

Vaupel (1999), 1850-1975 Amerika ve İsveç ölüm verisini kullanılarak yaş ve zamana göre dinamik bir hassasiyet modeli oluşturmuştur.

Butt ve Haberman (2002) sigorta tabanlı iki ölüm verisinde hassasiyet etkisinin olup olmadığını sorgulamışlardır. Bunun için dayanak ölüm hızı için Makeham/Gompertz Ölümlülük Modeli'ni ve hassasiyet dağılımı olarak ise Gamma ve Ters Gauss dağılımlarını kullanmışlardır. Bu modellere ilişkin parametre tahminlerinin elde edilebilmesinde Genelleştirilmiş Doğrusal Modeller (Generalized Linear Models) ve Ağırlıklandırılmış Doğrusal Olmayan Modeller (Weighted Non-Linear Models) kullanılmıştır. 2002 yılında yapılan bu araştırma projesi, 2004 yılında makale olarak yayınlanmıştır (Butt ve Haberman,2004).

Olivieri (2006) hassasiyet modellerinin kullanılması sonucunda hayat annüitesi portföyünün riskini değerlendirmiştir. Bu çalışmada Butt ve Haberman (2004) tarafından oluşturulan hassasiyet modeli parametrelerinden yararlanılmıştır.

Willemse ve Kaas (2007) ölüm hızının Gompertz Ölümlülük Yasası'na ve hassasiyet etkisi dağılımının Gamma dağılımı'na uyması durumunun Genelleştirilmiş Gompertz Dağılımı'na uyan iki rastlantı değişkeninin farkı şeklinde ifade edilebileceğini ispatlamışlardır.

Martine et al. (2009), ileri yaşlardaki bireyler için bazı ortak risk faktörlerinin etkisini hesaba katarak hassasiyet etkisinin olup olmadığını incelemişlerdir.

Su ve Sherris (2011) ise Avustralyalı kadın ve erkek annüitantlar için Gompertz/Gamma hassasiyet modeli, Gompertz/Ters Gauss hassasiyet modeli ve Markov Modeli'ni kullanmışlardır. Bu modellere ilişkin parametrelerin log-olabilirlik fonksiyonu ile basit bir şekilde tahmin edilebilmesi için popülasyonun ölüm hızının Normal dağılıma uyduğu varsayılmıştır.

2. YAŞAM ANALİZİ

2.1. Giriş

Bireylerin gelecek yaşam sürelerinin modellenmesi, hayat sigortası şirketleri ve sosyal güvenlik kurumları açısından oldukça önemlidir. Bireylerin beklenen gelecek yaşam sürelerinin ne yönde ve ne kadar değişeceğinin öngörülebilmesi için ise geçmiş yaşam deneyiminin doğru olarak yorumlanıp, modellenmiş olması gerekmektedir. Yapılan modelleme doğru ise belli bir güven düzeyinde geleceğe ilişkin doğru tahminler yapılabilir.

Yaşam analizini klasik istatistiksel analizden ayıran en önemli özellik veri kümesindeki gözlem sansürlenmesi (data-censoring) ve gözlem kesilmesi (data-truncation)dir. Sansürleme ve kesme, klasik istatistiksel analizdeki temel varsayımları kullanmayı oldukça zorlaştırmaktadır (Karim, 2008). Belirlenmiş bir zaman noktasından önce, sonra veya belirlenmiş iki nokta arasındaki olayların dikkate alındığı duruma 'sansürleme (censoring)' ve herhangi bir zamandan önce gerçekleşen olayların dikkate alınmadığı durum için ise 'kesme (truncation)' kavramı kullanılmaktadır. Sansürleme işlemi rastgele iken kesme işlemi rastgele değildir.

Yaşam analizini klasik istatistiksel analizden ayıran bir diğer önemli özellik ise varsayımların farklılık göstermesidir. Örneğin; klasik istatistiksel analizde model artıklarının Normal dağılıma, yaşam analizinde ise ölüm sayısı artıklarının Poisson dağılımına uyduğu varsayılmaktadır. Ölüm sayısı artıklarının Poisson dağılımına uyduğu varsayımı, Forfar et al. (1988) ve Renshaw (1991,1995) tarafından ayrıntılı olarak incelenmiştir.

2.2. Yaşam Analizinde Kullanılan Fonksiyonlar

Bu kesimde risk sınıflandırılması, sansürleme ve kesme olmadığında yaşam analizi ile ilgili bazı önemli gösterimler ele alınacaktır.

Şu anda (x) yaşında olan bir bireyin gelecek yaşam süresi, T rastlantı değişkeni ile gösterilmektedir. Bu durumda (x) yaşındaki birey öldüğü anda $(x+T)$ yaşında olacaktır (Gerber, 1997). Burada, T , negatif olmayan bir rastlantı değişkeni olarak tanımlanmakta ve aksi belirtilmediği sürece bu rastlantı değişkeninin tanım aralığının $[0, \infty)$ olduğu kabul edilmektedir (Duchateau ve Janssen, 2008). T rastlantı değişkeninin birikimli dağılım fonksiyonu,

$$G(t) = \Pr(T \leq t) = \int_0^t g(u)du, \quad t \geq 0 \quad (2.1)$$

şeklinde tanımlanmaktadır. Eş. (2.1), bir bireyin t yıl içerisinde ölmesi olasılığını göstermektedir. T rastlantı değişkenine ilişkin G birikimli fonksiyonunun bilindiği ve sürekli olduğu varsayıldığında olasılık yoğunluk fonksiyonu, $g(t) = G'(t)$ olarak elde edilmektedir ve

$$g(t)dt = \Pr(t < T < t + dt) \quad (2.2)$$

şeklinde tanımlanmaktadır (Gerber,1997). Bu bağlamda, bir bireyin t anından sonra yaşaması olasılığını gösteren fonksiyon, yaşam fonksiyonu olarak adlandırılmakta ve

$$S(t) = P(T \geq t) = 1 - G(t) = \int_t^{\infty} g(u)du \quad (2.3)$$

eşitliği ile ifade edilmektedir. Yaşam analizindeki bir diğer önemli fonksiyon ise, tehlike hızı fonksiyonudur ve $h(t)$ ile gösterilmektedir. Tehlike hızı fonksiyonu, olayın t anından önce gerçekleşmediği bilindiğinde, $[t, t + \Delta t]$ aralığında gerçekleşmesinin koşullu olasılığından elde edilmektedir. Tehlike hızı fonksiyonu bu koşullu olasılığın Δt sıfıra giderken limitinin alınması ile bulunmaktadır. Bu durumda tehlike hızı fonksiyonu ,

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

şeklinde elde edilmekte ve birikimli tehlike hızı fonksiyonu ise

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

olarak tanımlanmaktadır. Yaşam fonksiyonu, tehlike hızı fonksiyonunun integrali alınarak,

$$S(t) = \exp\left(-\int_0^t h(u)du\right) = \exp(-H(t)) \quad (2.6)$$

şeklinde yazılabilmektedir. Aynı şekilde tehlike hızı fonksiyonu da yaşam fonksiyonu kullanılarak,

$$h(t) = -\frac{d}{dt} \ln(S(t)) \quad (2.7)$$

eşitliğinden elde edilebilmektedir (Duchateau ve Janssen, 2008).

Hayat Sigortaları Matematiği'nde tehlike hızı fonksiyonu μ_x veya $\mu(x)$ ile gösterilmekte ve ölüm hızı (force of mortality rate) olarak adlandırılmaktadır (Bowers et al., 1997). Hayat Sigortaları Matematiği'nde bir diğer önemli fonksiyon ise q_x ile gösterilen fonksiyondur. Bu fonksiyon, (x) yaşındaki bireyin bir yıl içerisinde ölmesi olasılığını göstermektedir.

l_x , (x) yaşında yaşayan kişi sayısını göstermek üzere $l_x \mu_x$ eğrisi ölüm eğrisi (curve of deaths) olarak adlandırılmaktadır. Bu eğrinin (x)'e göre türevinin sıfır olduğu noktalar, büküm noktasını göstermektedir. Bazı yaşlar için ölüm oranı olan q_x ve ölüm hızı olan μ_x arasında $q_x > \mu_x$ veya $q_x < \mu_x$ ilişkisi bulunmaktadır. $l_x \mu_x$ eğrisinin artan olduğu noktalarda $q_x > \mu_x$ ve $l_x \mu_x$ eğrisinin azalan olduğu noktalarda $q_x < \mu_x$ olmaktadır (Neill, 1983).

2.3. Sansürleme ve Kesme

Sansürlenmiş veya kesilmiş gözlem, ilgilenilen duruma ilişkin kısmi bilgi içermektedir. Bundan dolayı sansürleme ve kesmeye neden olan noktaların dikkatlice belirlenmesi gerekmektedir. Bu kesimde sansürleme ve kesme işlemlerinin, yaşam verisinin birikim fonksiyonunu nasıl değiştirdiği incelenecektir.

2.3.1. Sansürleme

Gelecek yaşam süresinin tam değerinin bilinmediği fakat sadece u değerine eşit ve bu değerden büyük olduğu bilindiğinde gelecek yaşam süresini gösteren T rastlantı değişkeni soldan sansürlüdür. Aksine, gelecek yaşam süresinin r değerine eşit ve bu değerden küçük olduğu bilindiğinde gelecek yaşam süresini gösteren T rastlantı değişkeni sağdan sansürlüdür. Eğer, $[r, u)$ aralığı biliniyorsa gelecek yaşam süresini gösteren rastlantı değişkeni aralık sansürlüdür (Duchateau ve Janssen, 2008).

Bu çalışmada temel olarak sağdan sansürleme kullanılacağı için sağdan sansürlemenin gelecek yaşam süresi rastlantı değişkeninin olasılık yoğunluk fonksiyonunu nasıl etkilediği incelenecektir. Şekil 2.1'de sağdan sansürleme gösterilmiştir:



Şekil 2.1. Sansürleme

Burada, t_1 ilgili gözlem döneminin başladığı ve t_2 ise ilgili gözlem döneminin bittiğini noktayı göstermektedir. Bu iki nokta arasında geçen süre genelde bir yıl olarak belirlenmektedir. Bu şekilde (x) incelemeye alınan yaşam sürelerini ve (o) ile gösterilenler ise sağdan sansürlenmiş gözlemleri tanımlamaktadır.

$X_1, X_2, \dots, X_n$ birikimli dağılım fonksiyonu F olan bağımsız ve aynı dağılıma sahip rastlantı değişkenleri ve $Y_1, Y_2, \dots, Y_n$ ise birikimli dağılım fonksiyonu G olan birbirinden bağımsız ve aynı dağılıma sahip sansürlenme zamanlarını gösteren rastlantı değişkenleri olsun. Burada, X ve Y rastlantı değişkenlerinin bağımsız ve aynı dağılıma sahip olmaları varsayımı, sansürlenmiş veriden yapılacak dağılım çıkarımının elde edilebilmesi için yeterlidir (Fleming ve Harrington, 1991). Böylece, f ve g olasılık yoğunluk fonksiyonları, sırasıyla, X ve Y rastlantı değişkenlerine aittir. Burada sadece $(T_1, \Delta_1), (T_2, \Delta_2), \dots, (T_n, \Delta_n)$ değerleri gözlemlenebilmekte, $T_i = \min\{X_i, Y_i\}$ olarak tanımlanmakta ve Δ_i ,

$$\Delta_i = \begin{cases} 1 : X_i \leq Y_i, T_i \text{ sansürlü olmayan gözlem} \\ 0 : X_i > Y_i, T_i \text{ sansürlü gözlem} \end{cases}$$

ile gösterilmektedir.

Sansürlü olmayan gözlem için gelecek yaşam süresini gösteren rastlantı değişkeninin birikim fonksiyonu,

$$\begin{aligned}
 P(T < t, \Delta = 1) &= P(X < t, X \leq Y) = \int \int_{x < t, x \leq y} f(x)g(y)dx dy \\
 &= \int_{x < t} f(x) \left(\int_{x \leq y} g(y)dy \right) dx = \int_{x < t} f(x)(1 - G(x)) dx \neq F(t)
 \end{aligned} \tag{2.8}$$

olarak elde edilmektedir. Sansürlü gözlem için ise gelecek yaşam süresini gösteren rastlantı değişkeninin birikimli fonksiyonu,

$$\begin{aligned}
 P(T < t, \Delta = 0) &= P(X < t, X > Y) = \int \int_{x < t, x > y} f(x) g(y)dx dy \\
 &= \int_{x < t} f(x) \left(\int_{x > y} g(y)dy \right) dx = \int_{x < t} f(x)G(x) dx \neq F(t)
 \end{aligned} \tag{2.9}$$

şeklinde elde edilir (Wienke,1996).

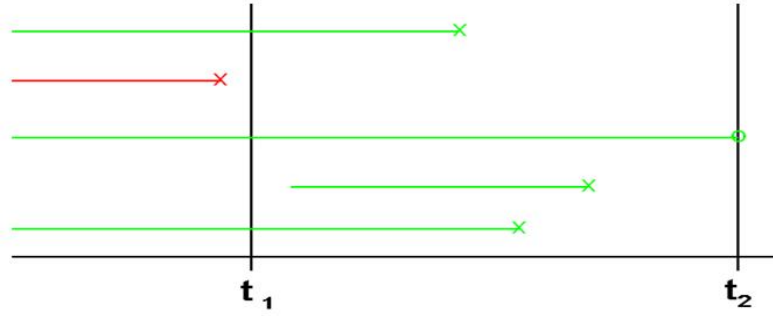
Wienke (2006) tarafından oluşturulan teoreme göre (T, Δ) verisinin olasılık yoğunluk fonksiyonu,

$$f(t, \delta) = (f(t)(1 - G(t)))^\delta (g(t)(1 - F(t)))^{1-\delta} \tag{2.10}$$

şeklinde göstermektedir. Bu teoremden sansürleme zamanı ve gelecek yaşam sürelerinin birbirinden bağımsız olduğu varsayımı kullanılmaktadır. Sansürleme ve gözlem bağımsız değilse olabilirlik fonksiyonuna sansürleme etkisi dahil edilmez.

2.3.2. Kesme

Birçok yaşam analizi uygulamasında, soldan kesme kullanılmaktadır. Sansürleme, rastgele olarak tanımlanmasına rağmen; kesme, rastgele değildir. Soldan kesmede çalışmanın doğal başlangıcından sonra, gözlemlerin bilinen bir zamanda sisteme girdiği bilinmektedir. Söz konusu kesme zamanından önce bireyin başarısız olması durumunda bireyler kayıt altına alınmamaktadır. Bu çalışmada soldan kesme kullanılacağı için soldan kesmenin T rastlantı değişkeninin olasılık yoğunluk fonksiyonunu nasıl etkilediği incelenecektir.



Şekil 2.2. Kesme

Burada, t_1 ilgili gözlem döneminin başladığı ve t_2 ise ilgili gözlem döneminin bittiğini noktayı göstermektedir. Bu iki nokta arasında geçen süre genelde bir yıl olarak alınmaktadır.

Şekil 2.2’de ikinci gözlem yaşam analizi çalışması başlamadan önce sistemden ayrıldığı için inceleme altına alınmamıştır. Burada, çalışma yürütücüsü bu bireyin farkında değildir. Diğer bir deyişle, kesme, koşullu dağılımdan yapılan örnekleme olarak düşünülebilmektedir. Gözlemlerin yaşam süresi ve sansürlenme durumunu içerdigi düşünölsün ve (T^*, Δ^*) şeklinde gösterilsin. Bilinen kesme zamanı t^* ile gösterildiğinde gözlemlerin soldan kesilmiş olması durumunda yaşam fonksiyonu,

$$P(T^* \geq t, \Delta^* = \delta) = P(T \geq t, \Delta = \delta | T^* \geq t)$$

$$= \frac{P(T \geq t, \Delta = \delta)}{P(T^* \geq t)} = \frac{P(T \geq t, \Delta = \delta)}{(1 - F(t^*))(1 - G(t^*))} \quad (2.11)$$

şeklinde tanımlanmaktadır. Bu durumda hem soldan kesilmiş hem de sağdan sansürlenmiş gözlem için olasılık yoğunluk fonksiyonu,

$$\begin{aligned} f(t, \delta, t^*) &= \frac{h(t, \delta)}{(1 - F(t^*))(1 - G(t^*))} \\ &= \delta \frac{f(t)(1 - G(t))}{(1 - F(t^*))(1 - G(t^*))} + (1 - \delta) \frac{g(t)(1 - F(t))}{(1 - F(t^*))(1 - G(t^*))} \end{aligned} \quad (2.12)$$

olarak tanımlanır. Eş. (2.12)'de sansürleme zamanı, kesme zamanı ve yaşam sürelerinin birbirinden bağımsız olduğu varsayımı kullanılmaktadır.

2.4. Yaşam Modellemesi

Yaşam modellemelerinde temel olarak iki yaklaşım kullanılmaktadır. Bunlar; 'parametrik yaklaşım' ve 'parametrik olmayan yaklaşım'dır. Parametrik yaklaşım kullanmanın avantajları şunlardır :

- Elde edilen sonuçlar yaşlara göre düzeltilebilir, interpolasyon ve ekstrapolasyon yapılabilir.
- Yaşam modeli oluşturularak hayat tablosu oluşturulabilir.
- Parametreler kolaylıkla yorumlanabilir.
- Çok fazla sayıdaki oran ve olasılıklar, birkaç parametreyle açıklanabilir (Keyfitz, 1985; Congdon,1993; Chang,1998).
- Yaşam modeline bağlı olan fonksiyonlar analitik olarak kolaylıkla incelenebilir.
- Farklı popülasyonlar için ölümlülük modelleri arasında karşılaştırma yapılabilir.
- Ölümlülük eğiliminin değerlemesi ve ele alınan zaman içerisinde öngörüsü yapılabilir.

Aktüeryal ve demografik literatüre bakıldığında parametrik modellemenin kullanılmasının amacı, daha az parametre ile veriye en iyi uyumu sağlayan modellemenin yapılmasıdır (Butt ve Haberman, 2002).

Parametrik olmayan yaklaşımın avantajları da şu şekilde sıralanabilir (Wolfgang, 1990; John, 2000):

- Güçlü varsayımlarda bulunulmasına gerek yoktur.
- Sabit parametrik modele bağlı kalınmamaktadır.
- İki değişken arasında var olan ilişkinin açıklanmasında esneklik sağlamaktadır.
- Ardışık değerler arasında interpolasyon yapılmasına izin vermektedir.
- Kayıp gözlemlerin yerine konulması için esnek yöntemler bulmaya izin vermektedir.
- Aykırı değerlerin bulunduğu verinin analizinde en uygun yaklaşımdır.

Bu kesimde yaşam modellemesi için kullanılan parametrik ve parametrik olmayan modellere ilişkin kısa bir literatür incelemesi yapılacaktır.

2.4.1. Parametrik yaklaşım

Yaşam modellemesinde parametrik yaklaşım için iki farklı yaklaşım kullanılmaktadır. İlk yaklaşımda veri setine uygun model birkaç parametre yardımı ile açıklanır. İkinci yaklaşımda ise bilinen parametrik dağılımlar kullanılarak modelleme yapılır.

2.4.1.1. Parametreler yardımı ile yapılan modelleme

Yaşam analizi ile ilgili ilk önemli çalışma, Graund (1662) tarafından yapılmıştır. Bu çalışma, Londra'nın ölümlülüğüne ilişkin ilk ayrıntılı çalışmadır. İkinci önemli çalışma ise 1693 yılında Halley tarafından yapılmıştır. Bu çalışmada ise Breslau için ilk defa hayat tablosu oluşturulmuştur.

Ölümlülük yasası olarak nitelendiren ilk çalışma, 1725 yılında De Moivre tarafından yapılmıştır. Bu modelde, maksimum yaşı (w) olarak bilindiği ve (x) yaşındaki bir birey için T rastlantı değişkeninin ($0, w-x$) aralığında sürekli Tekdüze dağılıma sahip olduğu varsayılmaktadır. Bu model için olasılık yoğunluk fonksiyonu,

$$f(x) = \frac{1}{w-x}, 0 < x < w-x \quad (2.13)$$

şeklinde tanımlanmaktadır.

Gompertz Modeli, 1825 yılında Benjamin Gompertz tarafından oluşturulmuştur. Bu çalışmada oldukça büyük bir yaş aralığı için (bebek ölümlülüğü ve ileri yaş ölümlülüğü hariç) yaşa göre ölüm hızlarının durağan bir üstel oran ile artan bir şekilde modellenebileceği sonucu elde edilmiştir. Bu modelde yaşlanma süreci De Moivre Yasası'ndan daha iyi tanımlamakta olup en son yaşı (w) olduğu varsayımı ortadan kalkmaktadır. Bu model, demografi ve diğer disiplinlerde yaygın olarak kullanılmaktadır. Gözlemlerin matematiksel olarak Gompertz Modeli'ne uyduğunun araştırma dünyasında kabul görmesi nedeniyle ölümlülük yasası olarak düşünülmektedir. Bu modele göre ölüm hızı,

$$\mu_x = Bc^x \quad (2.14)$$

olarak tanımlanmıştır. Burada, $B > 0$ temel ölümlülüğü, $c > 0$ ise yaşa göre ölümlülükteki sabit artış oranını göstermektedir. Bu modelde ölümlülük eğilimi sadece yaşa göre tanımlanmıştır. Bu yüzden Gompertz Ölümlülük Yasası belirli bir yaşa kadar yetişkin ölümlülüğünü tanımlamak için kullanılmaktadır. Bazı ileri yaş aralığı içinde bu modelin uygun olduğu varsayılır. İkinci Gompertz Ölümlülük Yasası (1860), 10 parametreden oluşan bir formüldür. Bu formül,

$$\ln(l_x) = -bc^x + gh^x - xdf^x - jk^{m^x(x-n)} \quad (2.15)$$

şeklindedir. Bu modeldeki amaç, Gompertz Ölümlülük Yasası'nda yeterince tanımlanamayan ölümlülüğün tüm yaş aralığı için modellenmesidir. Eş. (2.15)'de tanımlanan fonksiyon, ölüm hızına eşittir. Üçüncü Gompertz Ölümlülük Yasası ise 1862 yılında oluşturulmuştur.

1867 yılında Makeham tarafından yapılan çalışmada Gompertz Ölümlülük Yasası ile tanımlanan ölüm hızına kaza nedeniyle genç ölümlülüğünü tanımlayan sabit bir parametre eklenerek Makeham Ölümlülük Yasası oluşturulmuştur. Bu modelde ölüm hızı,

$$\mu_x = A + Bc^x \quad (2.16)$$

şeklinde tanımlanmaktadır. 1871 yılında Thiele ise ölüm hızlarını modellemek için 7 parametrelili,

$$\mu_x = A \exp(-Bx) + C \exp(-D(x - E)^2) + FG^x \quad (2.17)$$

şeklinde bir model geliştirmiştir. Bu modelde yaşam süresinin tüm yaşlar için modellenmesi amaçlanmıştır. Bu modeldeki her bir parametre insan ölümlülüğüne ilişkin birer özelliği tanımlamaktadır. İlk terim, doğumdan sonra hızla artan bebek ölümlülüğünü, ikinci terim kaza nedeniyle gerçekleşen ölümlülüğü ve üçüncü terim ise ileri yaşlar için ölümlülüğü göstermektedir. İkinci Makeham modeli (1890) ise dört parametrelidir ve

$$\mu_x = A + Hx + Bc^x \quad (2.18)$$

şeklinde gösterilmektedir. Perks (1932), lojistik eğri olarak bilinen dört parametrelili bir model oluşturmuştur. Bu modele, Birinci Perks Modeli adı verilmekte ve yaşa göre ölüm hızı,

$$\mu_x = \frac{A + Bc^x}{1 + Dc^x} \quad (2.19)$$

olarak tanımlanmıştır. Bu modelde A, B ve c parametrelili Makeham Ölümlülük Yasası ile 1, D ve c parametrelili Makeham Ölümlülük Yasası oranlanmıştır. D parametresi sıfır olarak alındığında bu model Makeham Ölümlülük Yasası'na, A ve D parametreleri sıfır olduğunda ise Gompertz Ölümlülük Yasası'na uymaktadır. Böylece Perks (1932) tarafından oluşturulan bu lojistik model, Gompertz ve Makeham Ölümlülük Yasası'nın hangi yaş aralıkları için uygun olduğunun incelenmesinde kullanılabilmektedir. Birinci Perks Modeli heterojenliğin ölüm oranları üzerindeki etkisini inceleyen modellerden birisidir. Birinci Perks Modeli, İngiltere'de annüitantlar için oluşturulan hayat tablosunun düzeltilmesinde kullanılmıştır. Aynı yıl içerisinde Perks, Eş. (2.19)'da tanımlanan modele benzeyen beş parametrelili yeni bir modeli,

$$\mu_x = \frac{A + Bc^x}{1 + Kc^{-x} + Dc^x} \quad (2.20)$$

eşitliği ile ifade etmiştir. Eş. (2.20)'de tanımlanan model İkinci Perks Modeli olarak adlandırılmaktadır. Bu modelde paydanın etkisi, (80) ve üzeri yaşlar için ölümlülüğü çok iyi tanımlayan Gompertz Ölümlülük Yasası'ndaki üstel artışı düzeltmek amacıyla kullanılmıştır.

1951 yılında Weibull tarafından yapılan çalışmada ise ölüm hızının bulunulan yaşın bir kuvvet olarak büyüdüğü bir model oluşturulmuştur. Weibull (1951) ölüm hızı için

$$\mu_x = Ax^B \quad (2.21)$$

iki parametrelili bir model kullanmıştır. Eş. (2.21)'de tanımlanan modele göre A ve B parametreleri kullanılmıştır. A parametresi, başlangıç yaş için temel ölümlülüğü tanımlamaktadır. B parametresi ise yaş ilerledikçe yaşın ölüm hızı üzerindeki etkisini ölçeklendirmek amacıyla kullanılmıştır.

ELT (English Life Tables) 11-12 çalışmalarında (27) ve üzerindeki yaşlar için ölüm oranlarında düzeltme yapmak amacıyla; $a, b, c, x_1, x_2, \alpha$ ve β parametrelerine sahip bir eğri oluşturmuştur. Bu model,

$$m_x \approx \mu_{x+\frac{1}{2}} = a + c \exp(-\beta(x - x_2)^2) + \frac{b}{1 + \exp(-\alpha(x - x_1))} \quad (2.22)$$

eşitliği ile ifade edilmiştir. Eş. (2.22)'de tanımlanan modele göre üç terim bulunmaktadır. İlk terim her bir yaş için sabit ölümlülüğü tanımlamaktadır. İkinci terim, (x_2) yaşından uzaklaştıkça, β durağan üstel artış oranını göstermek üzere, ölüm hızındaki azalmayı tanımlamaktadır. İleri yaşlarda bu terim sıfıra oldukça yakın bir değer almaktadır. Üçüncü terim ise (x_1) yaşından uzaklaştıkça ölüm hızındaki artışı tanımlamaktadır. İleri yaşlarda ise bu terim b 'ye yakınsamaktadır. Beard (1971), beş parametrelili bir model kurmuştur. Modelde (x) yaşındaki bir bireyin bir yıl içinde ölme olasılığı,

$$q_x = A + \frac{Bc^x}{Ec^{-2x} + 1 + Dc^x} \quad (2.23)$$

eşitliği ile modellenmiştir. Eş. (2.23)'te tanımlanan modele göre A parametresi sabit ölümlülüğü; B, C, D ve E parametreleri ise yaşlara göre ölümlülüğü detaylı olarak tanımlamak için kullanılmıştır. Bu model, İngiltere için 1949-1952 yılları arasında sigortalı bireyler için hayat tablosunun düzeltilmesinde kullanılmıştır. Barnett (1974) yıllık ölüm oranları için dört parametrelili formül elde etmiştir. Bu formül,

$$q_x = \frac{f(x)}{1 + f(x)} \quad (2.24)$$

şeklinde ve $f(x) = A - Hx + Bc^x$ şeklinde tanımlanmaktadır. Bu formül 1967-1970 yılları arasındaki İngiltere sigortalı deneyiminden elde edilen ölüm oranlarını düzeltmek için kullanılmıştır. Eş. (2.24)'te tanımlanan modele göre A parametresi sabit ölümlülüğü; H , parametresi yaşa göre ölüm oranında azalmayı; B ve c parametreleri ise yaşa göre ölüm oranlarındaki üstel artışı açıklamaktadır. Bu modele göre yaş ilerledikçe ölüm oranı artmaktadır.

Birinci ve İkinci Heligman-Pollard (1980) modellerinde tüm yaşlar için ölümlülüğü modellemek amacıyla A, B, C, D, E, F, G ve H parametrelerinden oluşan ve üç terimli modeller oluşturulmuştur.

Birinci Heligman-Pollard Modeli'nde $q_x = \frac{f(x)}{1+f(x)}$ olmak üzere,

$$f(x) = A^{(x+B)^C} + De^{-E\left(\log_e \frac{x}{F}\right)^2} + GH^x \quad (2.25)$$

şeklindedir. İkinci Heligman-Pollard Modeli ise,

$$q_x = A^{(x+B)^C} + De^{-E\left(\log_e \frac{x}{F}\right)^2} + \frac{GH^x}{1+GH^x} \quad (2.26)$$

eşitliği ile verilir. Eş. (2.26)' de tanımlanan model, biraz daha değiştirilerek Üçüncü Heligman-Pollard Modeli elde edilmiştir. Bu model,

$$q_x = A^{(x+B)^C} + D \exp\left(-E\left(\log_e \frac{x}{F}\right)^2\right) + \frac{GH^x}{1+KGH^x} \quad (2.27)$$

şeklindedir. İleri yaşlar için ilk iki terimin çok küçük bir değer alması nedeniyle Eş. (2.25) ve Eş.(2.26)'da tanımlanan modeller benzer sonuçlar vermektedir. ELT (English Life Tables) 15 tablosu Eş. (2.27)'deki model kullanılarak oluşturulmuştur. Dördüncü Heligman-Pollard Modeli (1980) ise 9 parametrelidir ve

$$q_x = A^{(x+B)^C} + D \exp\left(-E\left(\log_e \frac{x}{F}\right)^2\right) + \frac{GH^{x^k}}{1+KGH^{x^k}} \quad (2.28)$$

şeklindedir. Heligman-Pollard Modelleri'nde temel olarak üç terim bulunmaktadır. Bu modellerde ilk terim çocuk ölümlülüğünü, ikinci terim kazaen ölümlülüğü (özellikle 20 yaş civarında) ve üçüncü terim ise ileri yaş ölümlülüğünü tanımlamaktadır.

Heligman-Pollard Modeli'nde ilk terim içerisinde üç parametre bulunmaktadır. A parametresi, kabaca (1) yaşındaki ölümlülüğü vermekte olup çocuk ölümlülüğüne ilişkin seviye veya yoğunluk ölçümü olarak alınabilmektedir (McNown ve Rogers, 1989; Rogers ve Gard, 1991; Hartmann, 1987). B parametresi, yaş uzaklık değişkeni (age displacement variable)dir ve (0) ile (1) yaş ölümlülükleri arasındaki farkı

göstermektedir (Rogers ve Gard, 1991). B parametresinin değeri arttıkça (0) yaşındaki bireyin bir yıl içinde ölmesi olasılığının değeri 0,5'in altına düşmekte ve (1) yaşındaki bireyin bir yıl içinde ölmesi olasılığı olasılığına yakınlaşmaya başlamaktadır. Son olarak C parametresi çocukluk dönemi ve genç yetişkinlik dönemi boyunca ölüm oranının yaşa göre nasıl bir hızla azaldığını göstermek için kullanılmaktadır. A'nın değerindeki azalma çocuk ölümlülüğündeki azalma ile tutarlıdır. A,B ve C parametrelerinin hepsi sıfır ile bir aralığında yer almaktadır.

Heligman-Pollard Modeli'nde ikinci terim içerisinde de üç parametre bulunmaktadır. D parametresi, yetişkin ölümlülüğündeki yoğunluğu ve seviyeyi tanımlamaktadır. E parametresi kazaen ölümlülükteki sıçramayı tanımlamakta olup sıçramanın yayılımı ile ters ilişkilidir. F parametresi ise bulunulan pozisyonu göstermektedir (Heligman ve Pollard, 1980; McNown ve Rogers, 1989; Rogers ve Gard, 1991). D parametresi sıfır ile bir aralığında yer almaktadır. E parametresi ise sıfır ile sonsuz aralığında bir değer almaktadır. F parametresinin tanım aralığı tam olarak belirlenememektedir.

Heligman-Pollard Modeli'nin üçüncü teriminde Eş. (2.25) ve Eş. (2.26)'da iki parametre, Eş. (2.27) ve Eş. (2.28)'de üç parametre bulunmaktadır. G parametresi (0) yaşındaki temel ölümlülüğü ve H parametresi artış oranını göstermektedir (Rogers ve Gard, 1991). G parametresi sıfır ile bir aralığında yer almakta iken H parametresi sıfır ile sonsuz aralığında değer almaktadır. Eş. (2.27) ve Eş. (2.28)'deki K parametresi ise ağırlıklandırma yapmak amacıyla kullanılmaktadır.

Forfar, McCutcheon, Wilkie (1988), Eş. (2.29) ve Eş. (2.30) ile tanımlanan modelleri oluşturmuşlardır. Bu modellerde (r+s) kadar parametrenin bulunduğu genel bir model elde edilmiştir. Bu modeller,

$$\mu_x = GM^{r,s} = \sum_{i=1}^r \alpha_i x^{i-1} + \exp \left(\sum_{i=r+1}^{r+s} \alpha_i x^{i-r-1} \right) \quad (2.29)$$

$$q_x = \frac{GM^{r,s}}{1 + GM^{r,s}} = LGM^{r,s} \quad (2.30)$$

şeklinde. α parametreleri her bir yaş için ağırlıklandırma yapma amacıyla kullanılmıştır. Bu modelde r, sıfır değerini aldığı polinomal terim bulunmamakta ve s, sıfır değerini aldığı üstel terim bulunmamaktadır.

Eş. (2.29)'da tanımlanan modelde GM(0,2) Gompertz Ölümlülük Yasası'nı, GM(1,2) Birinci Makeham Ölümlülük Yasası'nı ve GM(2,2) İkinci Makeham Ölümlülük Yasası'nı göstermektedir. İngiltere'de 1991-1994 dönemi için iki yıllık seçim periyoduna sahip ve seçim periyoduna sahip olmayan emekli bireyler için bu model düzeltme amaçlı kullanılmıştır. Bu analizde r parameresi iki ve s parametresi üç olarak alınmıştır.

2.4.1.2. Bilinen dağılımlar yardımı ile yapılan modelleme

Bu modelleme için bilinen ilk ve en basit model, Üstel modeldir. Üstel modelde riskin zaman boyunca sabit kaldığı varsayılmaktadır. Bu nedenden dolayı hafızasız olarak nitelendirilmektedir. Diğer bir deyişle, bir zaman aralığında ölme olasılığı; bulunulan ana değil, sadece o aralığın uzunluğuna bağlıdır. Yani, $X \geq x$ olduğu bilindiğinde $X - x$ 'in dağılımı orjinal dağılım ile aynıdır:

$$P(x \leq X < x + \delta | X \geq x) = Pr(X < \delta) \quad (2.31)$$

Burada, δ herhangi bir pozitif değerdir. Sonuç olarak, üstel dağılım bulunulan zamanın sıfır olmasından etkilenmemektedir. Üstel model, bir t parametresinin olması nedeniyle küçük bir değişime bile çok duyarlı olduğundan daha esnek dağılımlar kullanılmıştır.

Weibull modeli, 1951 yılında iki pozitif parametrelili üstel dağılımın genişletilmiş hali olarak, Waloddi Weibull tarafından oluşturulmuştur. Modelin ikinci parametresinin esneklik sağlaması nedeniyle farklı şekillerde ölüm hızı fonksiyonu elde edilebilmektedir. Weibull Dağılımı'nın olasılık yoğunluk fonksiyonu,

$$f(x) = \lambda \gamma x^{\gamma-1} \exp(-\lambda x^\gamma) \quad (2.32)$$

şeklindedir. Pike (1966) kanserin diğer hastalıklardan bağımsız olduğu varsayımı ile kanser hastaları için Weibull tehlike hızı fonksiyonunun kuramsal olarak uyum sağladığını göstermiştir. Weibull Dağılımı, tehlike hızlarının tek modlu ve geniş U şekline benzediği durumlarda uygun değildir. Weibull dağılımı genişletilerek değişik şekiller için incelemeler Mudholkar ve Hutson (1996) tarafından yapılmıştır.

Log-lojistik dağılımın tehlike hızı fonksiyonu, Lognormal dağılımın tehlike hızı fonksiyonunun şekline benzemektedir. Log-lojistik dağılım, Gompertz Ölümlülük

Yasası ve ortalaması ile varyansı bir olan Gamma dağılımının karmasından oluşmaktadır.

Ölümlülük modellemesinde kullanılan bir başka dağılım ise Gamma dağılımıdır. Gamma dağılımı, Üstel dağılımın özel bir durumudur. Gamma dağılımı yaşam analizinde yaşam fonksiyonu ve tehlike hızı fonksiyonlarının kapalı bir fonksiyon olması nedeniyle yaygın olarak kullanılamamaktadır.

2.4.2. Parametrik olmayan yaklaşım

Parametrik modelleme yaklaşımında model parametrelerinin tahmin edilmesindeki zorluklar nedeniyle parametrik olmayan yaklaşım kullanılmaktadır. Bu modellerden en önemlileri Kaplan-Meier tahmin edicisi, Hızlandırılmış Başarısızlık Zamanı Modeli (Accelerated failure time model) ve Orantılı Tehlike Modeli (The Proportional Hazard Model)'dir.

Kaplan-Meier tahmin edicisi (Kaplan ve Meier, 1958), yapay yaşam fonksiyonuna dayanan yaşam fonksiyonunun parametrik olmayan en çok olabilirlik tahmin edicisidir. Homojen bireylerden oluşan bir gruba ilişkin yaşam biçiminin karakterize edilebilmesi için yapay yaşam fonksiyonunun yaşlara göre değeri hesaplanır. Eğer örnekleme sansürlenmiş gözlemler yoksa t anındaki yapay yaşam fonksiyonu, t anında yaşayan bireylerin sayısının örneklem büyüklüğünün n değerine bölünmesi ile elde edilir. Sansürlenmiş gözlemler ile ilgilenildiğinde ise yaşam fonksiyonunun sansürleme durumuna uygunlaştırılması gerekmektedir.

Diğer önemli parametrik olmayan model ise orantılı tehlike modelidir ve PH modeli olarak kullanılmaktadır. Orantılı tehlike modeli, bağımlı değişkenin durasyonu ile ilgilenen oldukça basit bir regresyon modelidir. Orantılı tehlike modelinde bireylerin yaşam süresini etkileyen ortak risk faktörleri eşdeğişkenler (covariates) olarak adlandırılmaktadır. Bu durumda X eşdeğişken vektörü ile $h_0(t)$, t anında (yaşında) bireysel tehlike hızı fonksiyonunu göstermek üzere popülasyonun tehlike hızı fonksiyonu,

$$h(t, X) = h_0(t) \exp \left(\sum_{i=1}^k \beta_i X_i \right) \quad (2.33)$$

olarak tanımlanmaktadır. Bu modelde eşdeğişkenlerin bağımsız olduğu ve tüm gözlemlerin aynı dayanak tehlike hızı fonksiyonuna sahip olduğu varsayımı kullanılmaktadır. Ayrıca bu modelde, eş değişkenler dayanak tehlike hızı fonksiyonunu çarpımsal olarak etkilemektedir.

Yaşam analizinde eşdeğişken etkisini modellemek için kullanılan bir diğer önemli parametrik olmayan model ise hızlandırılmış başarısızlık zamanı modelidir. Hızlandırılmış başarısızlık zamanı modeli, AFT model olarak adlandırılmaktadır (Kalbfleisch and Prentice, 1980). AFT modelleri, veriye daha iyi uyum sağlaması nedeniyle PH modellerinden daha uygulanabilir (Reid, 1994). AFT modeli,

$$\log(t) = -\beta' x + \varepsilon \quad (2.34)$$

şeklinde ifade edilir. Burada; x , sabit veya rastgele olarak tanımlanmış eşdeğişken vektörünü; β , eşdeğişken etki vektörünü ve ε ise hata terimini göstermektedir.

AFT modeli, parametrik ve parametrik olmayan AFT modeli olmak üzere ikiye ayrılmaktadır. Parametrik AFT modelinde, hata teriminin dağılımının bilindiği ve bu hata terimlerinin dağılımlarının birbirinden bağımsız ve aynı dağılıma sahip olduğu varsayılmaktadır. Parametrik olmayan AFT modelinde ise hata teriminin dağılımı belirlenmemektedir (Duchateau ve Janssen, 2008).

Aktüeryal literatür incelendiğinde ise parametrik olmayan düzeltme için genel olarak iki yöntem kullanıldığı görülmektedir. Bunlardan birisi eğri uydurma (splines) yöntemidir. Bu yöntem, ELT (English Life Tables) 13-14 çalışmalarında kullanılmıştır. Eğri uydurma yönteminde parçalı polinomlar kullanılmaktadır. $[a,b]$ aralığında $m+2$ tane düğüm (knot) adı verilen reel sayılar

$$a = \xi_0 < \xi_1 < \dots < \xi_m < \xi_{m+1} = b \quad (2.35)$$

şeklinde belirlenerek eğri uydurma fonksiyonu, s , her bir aralık için $p_0, p_1, \dots, p_m$ polinomları ile,

$$s(x) = \begin{cases} p_0(x); & \xi_0 \leq x < \xi_1 \\ p_1(x); & \xi_1 \leq x < \xi_2 \\ \dots & \\ p_m(x); & \xi_m \leq x < \xi_{m+1} \end{cases} \quad (2.36)$$

şeklinde tanımlanmaktadır. Belirlenen düğüm noktalarının sıralanışına göre eğri uydurmanın şekli belirlenmektedir (Chan et al., 1984).

Parametrik olmayan düzeltmeye ilişkin aktüerler tarafından en yaygın olarak kullanılan yöntem Whittaker-Henderson modelidir. Bu model, amaç fonksiyonunun minimize edilmesine dayanarak düzeltme yapmak amacıyla kullanılmaktadır. $z_1, z_2, \dots, z_n$ gözlem değerlerine karşılık gelen düzeltilmiş değerler $y_1, y_2, \dots, y_n$ ile gösterildiğinde amaç fonksiyonu,

$$F(y_1, y_2, \dots, y_n) = \sum_{h=1}^n w_h (y_h - z_h)^2 + \lambda \sum_{h=1}^{n-k} w_h (\Delta_k y_h)^2 \quad (2.37)$$

şekllindedir. Burada λ sabit bir parametre w değerleri ağırlıkları ve $\Delta_k y_h$ ise

$$\Delta_k y_h = \sum_{i=0}^k (-1)^i \binom{k}{i} y_{h+k-i}$$

şeklinde gösterilen y_h 'nin k 'inci dereceden ileri farkı olarak tanımlanmaktadır. Türkiye Kadın ve Erkek Sigortalı Hayat Tabloları'nın (TRSH-2010) düzeltilmesinde bu yöntem kullanılmıştır.

2.5. Hassasiyet Modelleri

Yaşam analizinde gelecek yaşam süresi dağılımını analiz etmek için kullanılan karışık (mixture) dağılımların en önemlilerinden birisi hassasiyet modelleridir (Klugman et al., 2008). Hassasiyet modellerinin temelini gözlemlenebilen ve gözlemlenemeyen risk faktörleri ayrımı oluşturmaktadır. Hassasiyet modellerinde gözlemlenemeyen risk faktörlerini tanımlamak için dayanak tehlike hızı fonksiyonuna Z rastlantı değişkeninin etkisi, çarpımsal olarak eklenmektedir.

Gözlemlenemeyen risk faktörlerinin etkisi de eklendiğinde tehlike hızı fonksiyonu,

$$h_x(z) = \lim_{t \rightarrow 0} \frac{\Pr(T_x \leq t | Z_x = z)}{t} = z h_x(t) \quad (2.38)$$

şeklinde tanımlanmaktadır. Bu durumda Z bilindiğinde koşullu yaşam fonksiyonu,

$$S_{T|Z}(t|z) = \exp \left[- \int_0^t h_{T|Z}(t|z) dt \right] = \exp(-zH(t)) = (S(t))^z \quad (2.39)$$

şeklinde elde edilir. Hassasiyet etkisi, genel olarak, parametrik dağılımlarla ifade edilmektedir.

Hassasiyet modellerini parametrik yaklaşım veya parametrik olmayan yaklaşım başlıkları altında incelemek doğru değildir. Bunun temel nedeni ise dayanak tehlike hızı fonksiyonunun, $h_0(t)$, parametrik veya parametrik olmayan bir model olarak tanımlanmasından kaynaklanmaktadır.

Dayanak tehlike hızı fonksiyonunda ortak risk faktörlerinin etkisi tanımlanmadığında, hassasiyet modeli; parametrik bir model olmaktadır.

Dayanak tehlike hızı fonksiyonunda ortak risk faktörlerinin etkisi hesaba katıldığında ise dayanak tehlike hızı fonksiyonu parametrik olmayan bir model olmakta ve hassasiyet modeli de parametrik olmayan bir model olmaktadır.

Sağlık alanı ile ilgili çalışmalarda parametrik olmayan hassasiyet modelleri kullanılırken, demografik ve aktüeryal çalışmalarda parametrik hassasiyet modelleri kullanılmaktadır. Üçüncü Bölüm'de hassasiyet modelleri ayrıntılı olarak incelenecektir.

3. HASSASİYET MODELLERİ

3.1. Giriş

Yaşam analizi, ölüm zamanı verisi kullanılarak yapılmaktadır. Geleneksel yaşam modellerinde veri kümesindeki gözlemlerin gelecek yaşam sürelerinin bağımsız ve aynı dağılımlı olduğu, diğer bir deyişle, incelenen popülasyonun homojen olduğu varsayılmaktadır. Fakat bu varsayımın sağlık ve ölümlülük istatistiklerine bakıldığında doğru olmadığı gözlenmektedir. Bu durumda, gözlemlerin gelecek yaşam sürelerinin, birbirinden bağımsız olduğu varsayımı yanlış olmaktadır. Çünkü popülasyonun heterojen olması gözlemlerin gelecek yaşam süreleri arasında bağımlılığa neden olmaktadır. Popülasyonun heterojen olması, bazı beklenmedik sonuçların açıklanmasını ve elde edilen sonuçlar için alternatif yorumlar yapılabilmesine olanak sağlamaktadır (Hougaard, 1991).

Popülasyonun heterojenliğinin iki temel kaynağı olduğu söylenebilir :

- Gözlemlenebilir risk faktörlerinden kaynaklanan heterojenlik
- Gözlemlenemeyen risk faktörlerinden kaynaklanan heterojenlik

Bu çalışmada ikinci heterojenlik kaynağına değinilecek ve gözlemlenebilir risk faktörlerinden kaynaklanan heterojenlik ise bütünlük sağlayabilmek açısından incelenecektir. Popülasyonun heterojen olması nedeniyle bazı bireylerin yüksek ölüm riskine sahip olduğu düşünüldüğünde, diğer bireylerin bu bireylerden daha az veya daha yüksek ölüm riskine sahip olmaktadır. Bu şekildeki bir popülasyonun homojen olduğu düşünülerek yaşam modellemesi yapılması durumunda, popülasyonun ölümlülüğünün yanlış tanımlanması sorunu ortaya çıkar.

Popülasyon için gözlemlenemeyen risk faktörleri nedeniyle heterojenlik söz konusu ise yaşam modellemesi yapılırken popülasyonun değişik risk seviyelerine sahip bireylerden oluştuğu varsayımı altında karışık modeller kullanılmaktadır. Bu modellerde gözlemlenemeyen risk faktörlerinin etkisi, rastlantı değişkenleri ile tanımlanmakta ve hassasiyet etkisi (frailty effect) olarak adlandırılmaktadır.

Hassasiyet modellerinde bireylerin farklı hassasiyet değerlerine sahip olduğu ve hassasiyet seviyesi yüksek olan bireyin hassasiyeti daha düşük olan bireye göre popülasyondan daha erken ayrılacağı varsayımında bulunmaktadır. Böylece zaman

veya yaş ilerledikçe hassas bireylerin sistemden erken ayrılması nedeniyle daha güçlü bireyler sistemde kalacak ve buna bağlı olarak ölüm hızı azalacaktır.

Hassasiyet etkisi, bir veri kümesindeki gözlem veya veri kümesindeki her bir risk sınıfı için ayrı ayrı tanımlanabilir. Genel bir tanım yapılmak istenirse, hassasiyet modelleri, bireyler veya risk sınıfları arasındaki bağımlılığı gösteren rastgele etkiyi içeren genişletilmiş parametrik veya orantılı tehlike modelidir (Duchateau ve Janssen, 2008). Hassasiyet modelleri için popülasyonun heterojenlik durumu, hassasiyet teriminin dağılımı ile belirlenmektedir.

Popülasyon için (0) yaşındaki bireylerin, düşük hassasiyet seviyesinde yoğun ve sağa doğru uzun kuyruklu olduğu ve (0) yaşındaki bireyler için ortalama hassasiyet seviyesinin bir olduğu düşünülmektedir. Yaş ilerledikçe, hassasiyet seviyesi yüksek olan bireylerin ölme riskinin yüksek olması nedeniyle bu bireyler sistemden daha erken ayrılacaklardır. Bu durumda; hassasiyet seviyesi yüksek olan bireyler sistemden daha erken ayrılacağı için ortalama hassasiyet seviyesi düşecek ve yaşayan bireylerin hassasiyet dağılımının sol tarafına kaymasına neden olacaktır. Böylece hayatta kalan bireyler, daha düşük ortalama hassasiyet seviyesine sahip olacaktır.

Hassasiyet etkisini gösteren rastlantı değişkeni için farklı parametrik dağılımlar kullanılmaktadır. Bu dağılımların bazı özellikleri taşıması gerekmektedir. Bir sonraki kesimde, hassasiyet dağılımlarının özellikleri ayrıntılı olarak incelenecektir.

3.2. Hassasiyet Dağılımlarının Temel Özellikleri

Hassasiyet dağılımı, söz konusu veri kümesindeki heterojenlikten kaynaklanan bağımlılık yapısını en iyi şekilde tanımlayabilmelidir. Hassasiyet dağılımının kullanılmasındaki temel amaç, gelecek yaşam süreleri arasındaki bağımlılık yapısının açıklanmasına kolaylık sağlamaktır.

Hassasiyet dağılımı olarak kullanılacak dağılımın uygunluğunu belirleyebilmek için bazı fonksiyonlar ve bu fonksiyonların temel özelliklerinden yararlanılması gerekmektedir.

Bu fonksiyonlardan ilki Laplace Dönüşümü'dür. Laplace Dönüşümü ile bağlantı ilk olarak Hougaard (1984, 1986a, 1986b) tarafından dikkate alınmıştır. Doğal olarak hassasiyet dağılımı belirlenirken kullanılacak dağılımın Laplace Dönüşümü'nün olup

olmaması önemli olmaktadır. Bu sayede parametre tahmini yapılırken geleneksel en çok olabilirlik yöntemleri kullanılabilmekte ve hassasiyet dağılımının yaşam fonksiyonu ve ölüm hızı fonksiyonu nasıl etkilediği Laplace Dönüşümü kullanılarak sorgulanabilmektedir. Literatürde Gamma dağılımı, Pozitif durağan dağılım (Positive Stable Distribution), Ters gauss dağılımı, Güç varyans fonksiyonu (Power Variance Function) dağılımı ve Bileşik poisson dağılımı, hassasiyet dağılımı olarak kullanılmıştır. Literatürde kullanılan bir başka dağılım ise Lognormal dağılımdır. Lognormal dağılımın Laplace Dönüşümü'nün yapılamaması nedeniyle bu bölümde ele alınmayacaktır.

Burada, i risk sınıfında n_i tane bireyin olduğunu düşünölsün. Hassasiyet rastlantı değışkeni Z , bu risk sınıfındaki tüm bireyler için birikimli dayanak tehlike ölüm hızı fonksiyonu $H_0(t)$ ve eşdeğışken etkisi $\exp(x^t\beta)$ şeklinde tanımlansın. Bu durumda bu risk sınıfı için koşullu yaşam fonksiyonu,

$$S_i(t_{n_i}) = \exp \left[-z_i(H_0(t_1)\exp(x_{i1}^t\beta) + \dots + H_0(t_{n_i})\exp(x_{in_i}^t\beta)) \right] \quad (3.1)$$

şeklinde yazılır. Dayanak ölüm hızı fonksiyonunun risk sınıfındaki tüm bireyler için aynı olması durumunda, tehlike hızı fonksiyonu bireyler arasında hassasiyet terimi kadar fark gösterecektir. Eş. (3.1)'de dayanak tehlike hızı fonksiyonunun her bir birey için aynı olduğu ve ortak risk faktörü etkisini olmadığı varsayıldığında, bu risk sınıfı için çoklu yaşam fonksiyonu,

$$S_i(t_{n_i}) = \exp(-z_i H_0(t_1)) \dots \exp(-z_i H_0(t_{n_i})) \quad (3.2)$$

olarak yazılabilmektedir. Eş. (3.2)'de

$$L(s) = E(\exp(-Zs))$$

Laplace Dönüşümü kullanıldığında bu risk sınıfı için çoklu yaşam fonksiyonu,

$$S_i(t_{n_i}) = L(H_0(t_1)) \dots L(H_0(t_{n_i})) = \prod_{s=1}^{n_i} L(H_0(t_s)) \quad (3.3)$$

şeklinde tanımlanmaktadır. Laplace Dönüşümü, hassasiyet dağılımları için önemli olup, Laplace Dönüşümü'nün türevi alınarak hassasiyet dağılımının beklenen değer ve varyans gibi momentleri bulunabilmektedir.

Bir rastlantı değişkeninin Laplace Dönüşümü yapılabiliyorsa, bu rastlantı değişkeninin Laplace Dönüşümü'nün sıfır noktasında komşuluğu olduğu anlamına gelmektedir. Bu durumda Z rastlantı değişkeninin beklenen değeri,

$$E(Z) = -L'(0)$$

şeklindedir. Bir rastlantı değişkeninin sıfıra göre momentlerinin Laplace Dönüşümü ile bulunmasında,

$$E(Z^k) = (-1)^k L^{(k)}(0)$$

eşitliğinden yararlanılır. Böylece, hassasiyet rastlantı değişkeninin varyansı Laplace Dönüşümü kullanılarak,

$$\text{Var}(Z) = L''(0) - (L'(0))^2 \quad (3.4)$$

şeklinde elde edilir. Hassasiyet rastlantı değişkeninin olasılık yoğunluk fonksiyonunu gösteren f_z , ilgilenilen popülasyon için çalışmanın başında belirlenmekte ve daha sonra zamanla değiştiği varsayılmaktadır. Bu değişim, ortalamadan daha hassas bireylerin popülasyondan erken ayrılması şeklinde olmaktadır. Basitleştirmek için ortak risk faktörü bilgisinin olmadığı varsayıldığında t anında yaşayan bir birey için olasılık yoğunluk fonksiyonu,

$$f_z(z|T > t) = \frac{\exp(-zH_0(t))f_z(z)}{L(H_0(t))} \quad (3.5)$$

şeklindedir. Koşullu hassasiyet teriminin başlangıçta beklenen değerinin bir olduğu ve zaman ilerledikçe hassasiyet seviyesi yüksek olan bireylerin, sistemden ayrılması beklendiği için bu değer azalacağı varsayılmaktadır. Benzer şekilde, t anında ölen bir birey için hassasiyet dağılımının olasılık yoğunluk fonksiyonu,

$$f_z(z|T = t) = \frac{z \exp(-zH_0(t))f_z(z)}{L'(H_0(t))} \quad (3.6)$$

şeklinde gösterilir. Bir diğer önemli fonksiyon ise hassasiyet teriminin bağımlılık derecesinin ölçülmesi için kullanılan Kendall'ın τ fonksiyonudur (Duchateau ve Janssen, 2008). Bu yöntemi basitçe anlatabilmek için ikişer gözlemlili iki risk sınıfı olduğu ve eşdeğişken faktörlerinin etkisinin sıfır olduğu varsayalım. İlk risk sınıfı i, bu

risk sınıfındaki olay zamanları (T_{i1}, T_{i2}) ve ikinci risk sınıfı k , bu risk sınıfındaki olay zamanları (T_{k1}, T_{k2}) olarak tanımlansın. Bu durumda Kendall'ın τ değeri (Kendall, 1938),

$$\tau = E[\text{sign}((T_{i1} - T_{k1})(T_{i2} - T_{k2}))] \quad (3.7)$$

şeklindedir. Sürekli dağılımlar için diğer bir formül de Genest ve MacKay (1986) tarafından elde edilmiştir. Bu çalışmaya göre τ değeri,

$$\begin{aligned} \tau &= P((T_{i1} - T_{k1})(T_{i2} - T_{k2}) > 0) - P((T_{i1} - T_{k1})(T_{i2} - T_{k2}) < 0) \\ &= 2P((T_{i1} - T_{k1})(T_{i2} - T_{k2}) > 0) - 1 \\ &= 2p - 1 \end{aligned}$$

şeklinde elde edilmektedir. Laplace Dönüşümü kullanılarak Kendall'ın τ değeri ise,

$$\tau = 4 \int_0^\infty \int_0^s sL(s)L^{(2)}sds - 1 \quad (3.8)$$

şeklinde elde edilir (Duchateau ve Janssen, 2008).

3.3. Hassasiyet Model Türleri

Literatürde, heterojenliğin etkisini hesaba katan kesikli ve sürekli yaklaşım olmak üzere iki tür hassasiyet modellemesi yaklaşımı bulunmaktadır.

3.3.1. Kesikli yaklaşım

Kesikli yaklaşımda, bazı özel risk faktörlerinin etkisinin belirlenmesi amaçlanmış ve heterojenlik basit fonksiyonlar yardımıyla tanımlanmıştır. 1959 yılında Levinson tarafından yapılan çalışmada, gelecek yaşam süresinin olasılık yoğunluk fonksiyonunun kesikli olduğu belirlenmiş ve farklı yaşlardan bireylerden oluşan heterojen bir popülasyon m adet homojen risk sınıfına bölündüğünde i 'inci grup için yıllık ölümlülük oranı $q_{(i)}$ olarak tanımlanmıştır. Popülasyonun homojen risk sınıflarına ayrılmasından dolayı her bir risk sınıfındaki bireylerin yaşlarının aynı olması beklenir. Risk sınıfları için yıllık ölümlülük oranları,

$$q_{(1)} < q_{(2)} < \dots < q_{(m)}$$

şeklinde sıralansın. Zaman ilerledikçe, ölümler ve yaşam koşullarının değişiklik göstermesi nedeniyle risk sınıflarının tüm popülasyona olan oranlarının farklılaşması beklenmektedir. Fakat bu modelde, her bir risk sınıfı için bu risk faktörlerin değişmediği varsayımı kullanıldığında, zamana bağlı olarak $q_{(i)}$ yıllık ölümlülük oranlarının değişmeyeceği varsayılmıştır.

$P_{ij}(x)$, (x) yaşında i 'inci sınıfta bulunan bireyin $(x+1)$ yaşında j 'inci sınıfa geçmesi olasılığını göstermek üzere ve bir bireyin bir yıl içerisinde ölüm dışında en fazla bir geçiş yapılabildiği varsayıldığında

$$\begin{aligned} P_{ij}(x) &\geq 0 ; j \geq i \\ P_{ij}(x) &= 0 ; j < i \end{aligned} \quad (3.9)$$

olur ve bu geçiş olasılıklarının toplamı bire eşit olmaktadır. Bu çalışmada P_{ij} geçiş olasılıklarına bozulma olasılıkları adı verilmektedir. 1969 yılında Redington tarafından yapılan çalışmada parametrik yaklaşım kullanılmış ve

$$\mu_x = c^{x-z} \quad (3.10)$$

şeklinde tanımlanmıştır. Burada, $c = e^\beta$ ve $z = -\frac{\ln \alpha}{\beta}$ olarak kullanılmıştır.

Bu modelde z , ölüm hızının bir olduğu yaşı göstermektedir. Z 'nin oldukça büyük hatta son yaşa (maximum age) yakın bir yaş olması ve verilen z yaşı için ölüm hızının artış oranını gösteren c 'nin değerinin bire yaklaşması gerekmektedir. Z değeri için c ne kadar büyükse; ölüm hızı da o kadar küçük olmakta ve bu nedenle c , ölümlülüğe direnç parametresi (resistance to death) olarak adlandırılmaktadır. Bu tanım, Gompertz Ölümlülük Yasası'nda ölümden kaçınma gücü (strength to avoid death) olarak tanımlanmıştır.

Bir bireyin ölümlülüğünü etkileyen faktörlerin gözlemlenemediği bilgisi verildiğinde c ve z parametrelerinin bu bireye özgü olduğu söylenebilir ve popülasyondaki heterojenliğin tanımlanabilmesi için ise c ve z birer rastlantı değişkeni olarak belirlenebilir.

Redington (1969), bir popülasyonu m homojen risk grubuna ayırmış ve her bir grup için c, z ve ölüm hızları ayrı ayrı belirlemiştir. Heterojenlik nedeniyle model kesikli olarak oluşturulmuş ve c ile z için simetrik dağılımlar kullanılmıştır. Böylece popülasyon için ortalama ölüm hızı,

$$\bar{\mu}_x = \sum_{i=1}^m a_{x(i)} \mu_{x(i)} \quad (3.11)$$

şeklinde tanımlanmış ve $\bar{\mu}_x$ merkezi ölüm hızı olarak adlandırılmıştır. Bu bilgilere göre popülasyonun ortalama ölüm hızı,

$$\mu'_x = c^{(x-z')}$$

olarak modellenmiştir. Heterojenlikten dolayı ileri yaşlar için $\bar{\mu}_x < \mu'_x$ olması beklenir. Redington(1969) c ve z parametrelerinin simetrik olması nedeniyle standart sapmalarının sonlu olması dışında özel varsayımlara gerek olmadığını vurgulamıştır.

Keyfitz ve Littman (1979) heterojenliği uygun şekilde tanımlayabilen önemli ve basit modeller geliştirmişlerdir. Popülasyonun homojen olduğu varsayılın ve bu popülasyon için yaşa bağlı yıllık ölümlülük oranları q_x , $x = 0, 1, 2, \dots$ olarak tanımlansın. Bu durumda (x) yaşındaki birey için beklenen tam yıl yaşam süresi,

$$e_x = \sum_{h=0}^{\infty} h \cdot q_x = \sum_{h=1}^{\infty} h p_x \quad (3.12)$$

şeklinde tanımlanır. Yıllık ölümlülük oranlarının yaştan bağımsız olduğu varsayıldığında ise beklenen tam yıl yaşam süresi,

$$e_x = e = \sum_{h=1}^{\infty} p^h = \frac{1}{1-p} = \frac{1}{q}$$

olur. Burada, hepsi (x_0) yaşındaki bireylerden oluşan heterojen bir popülasyon olduğu varsayılın ve bu popülasyon bazı risk faktörlerine göre m tane homojen risk sınıfına ayrılın. Ayrıca, her bir risk sınıfı için yıllık ölümlülük oranlarının sabit olduğu düşünölsün. Bu durumda $q_{(i)}$, i'inci risk sınıfı için yıllık ölümlülük oranını ve $a_{x_0(i)}$ i'inci risk sınıfının tüm popülasyona olan oranını göstermek üzere (x_0) yaşındaki bireyler için ortalama yıllık ölümlülük oranı,

$$\bar{q}_{|x_0} = \sum_{i=1}^m a_{x_0(i)} q_{(i)} \quad (3.13)$$

şeklinde elde edilir. Bu durumda tüm popülasyon için beklenen tam yıl yaşam süresi,

$$e_{x_0}^{[HO]} = \frac{1}{\bar{q}_{|x_0}} = \frac{1}{\sum_{i=1}^m a_{x_0(i)} q_{(i)}} \quad (3.14)$$

şeklinde ifade edilir. Aynı ayrı düşünüldüğünde ise i'inci risk sınıfındaki birey için beklenen yaşam süresi,

$$e_{(i)} = \frac{1}{q_{(i)}}$$

şeklinde tanımlanır. Burada, her bir risk sınıfının toplam popülasyona olan oranının farklı olması nedeniyle, bazı risk sınıfları için tanımlanan beklenen tam yıl yaşam süresi daha az tanımlanırken, diğer risk sınıfları için tanımlanan beklenen tam yıl yaşam süresi de fazla belirlenmektedir. Bu yüzden her bir risk sınıfı için beklenen tam yıl yaşam süreleri ayrı ayrı hesaplanıp ağırlıklandırıldığında tüm popülasyon için ortalama beklenen yaşam süresi,

$$e_{x_0}^{[HE]} = \sum_{i=1}^m a_{x_0(i)} e_{(i)} = \sum_{i=1}^m \frac{a_{x_0(i)}}{q_{(i)}} \quad (3.15)$$

şeklinde tanımlanır. Her bir risk sınıfı için yıllık ölümlülük oranlarının Eş. (3.15)'deki gibi aritmetik ortalaması yerine harmonik ortalaması alınarak popülasyonun homojen olduğu varsayıldığında elde edilen beklenen tam yıl yaşam süresinin popülasyonun heterojen olduğunda elde edilen beklenen tam yıl yaşam süresinden küçük olması beklenir. Ancak bu durum, hayat sigortası şirketlerinin yükümlülükler için yetersiz rezerv ayırmasına neden olacağından istenmemektedir.

Popülasyonun heterojen olduğu varsayıldığında ise her bir risk sınıfında beklenen kişi sayısının,

$$\frac{a_{x_0(i)}}{q_{(i)}}$$

olması beklenir. Bu oranlar kullanılarak popülasyonun bileşimi,

$$a_{x(i)} = \frac{a_{x_0(i)} / q_{(i)}}{\sum_{i=1}^m a_{x_0(i)} / q_{(i)}} \quad (3.16)$$

şeklinde elde edilir. Bu durumda popülasyon için beklenen yaşam süresi,

$$e_x^{[HE]} = \sum_{i=1}^m a_{x(i)} e_{(i)} = \sum_{i=1}^m \frac{a_{x_0(i)} / q_{(i)}}{\sum_{i=1}^m a_{x_0(i)} / q_{(i)}} e_{(i)} = \sum_{i=1}^m \left(a_{x_0(i)} \frac{e_{(i)}}{e_{x_0}^{[HE]}} \right) e_{(i)} \quad (3.17)$$

şeklinde dir. Eş. (3.17), Eş. (3.15) eşitliği ile karşılaştırıldığında başlangıçta belirlenmiş ağırlıklar bireylerin beklenen yaşam süreleri arasındaki orana bağlı olarak düzenlenmiştir. Düşük $q_{(i)}$ değeri için yüksek $e_{(i)}$ değeri elde edileceğinden $e_x^{[HE]} > e_{x_0}^{[HO]}$ olur. Yani, popülasyonun homojen olduğu varsayımı altında elde edilen beklenen tam yıl yaşam süresi tahmini zaman ilerledikçe doğru ve yeterli olamayacaktır.

3.3.2. Sürekli yaklaşım

Bir önceki kısımda da değinildiği gibi popülasyon için heterojenlik durumu, kesikli olarak modellenmek istenildiğinde, çok fazla sayıda parametrenin kullanılması gerekmektedir. Kesikli yaklaşımda, popülasyon m risk sınıfına ayrılarak her bir risk sınıfının büyüklüğüne göre başlangıç oranları tanımlanmakta ve bu risk sınıflarının her biri için gelecek yaşam süresini tanımlayan olasılık dağılımı belirlenmektedir. Risk sınıfları büyüklüklerinin popülasyon büyüklüğüne olan oranın sabit olması nedeniyle, elde edilen sonuçlar gerçeği yansıtmayabilir. Analitik sonuçlar elde edilmek isteniyorsa,, yaşam dağılımı için önemli basitleştirmelerin yapılması gerekmektedir (Olivieri, 2006).

Yaşam analizinde parametre sayısının fazla olması sorunu, sürekli yaklaşım kullanılarak çözülebilir. Parametre sayısının az olması ve değişen ölümlülük eğilimi dikkate alındığında, yaşa göre ölümlülüğün tanımlanabilmesi için sürekli modelleme tercih edilebilir. Hayat sigortası şirketleri, risk kabul sürecinde gözlemlenebilir risk faktörlerine göre risk sınıflandırması yapabilmektedirler. Daha öncede belirtildiği üzere hayat sigortası şirketleri, fiyatlandırma yaparken ölümlüğe etki eden gözlemlenemeyen risk faktörlerinin etkisini de dikkate alması gerekmektedir. Bu

yüzden gözlemlenemeyen risk faktörlerinden kaynaklanan heterojenliğin etkisini göstermek amacıyla negatif olmayan Z rastlantı değişkeni kullanılmaktadır (Butt ve Haberman, 2004).

Sürekli yaklaşımda hassasiyet modelleri temel olarak ikiye ayrılmaktadır. Bunlardan birisi, tek değişkenli hassasiyet modelleri; diğeri ise çok değişkenli hassasiyet modelleridir. Çok değişkenli hassasiyet modelleri de kendi içerisinde ikiye ayrılmaktadır. Bunlar, paylaşımlı hassasiyet modelleri (shared frailty models) ve ilişkili hassasiyet modelleri (correlated frailty models) olarak adlandırılmaktadır.

Bu çalışmada tek değişkenli hassasiyet modelleri kullanılacağından, tek değişkenli hassasiyet modelleri ayrıntılı olarak ele alınacaktır. Çok değişkenli hassasiyet dağılımlarına ilişkin olarak ise tanım, varsayımlar ve kullanılan alanlar konusunda genel bilgi verilecektir.

3.4. Tek Değişkenli Hassasiyet Modelleri

Gözlemlenemeyen risk faktörlerinin etkisini içeren hassasiyet modellerinin anlatımını basitleştirmek için gözlemlenebilir risk faktörlerinden kaynaklanan heterojenlik etkisinin olmadığı varsayılacaktır. Popülasyondaki gözlemlenemeyen risk faktörlerinden kaynaklanan heterojenlik probleminin çözülebilmesi için ise ilk çalışma Vaupel et al. (1979) tarafından yapılmıştır. Bu çalışmada temel varsayım, gözlemlenemeyen risk faktörlerinin etkisi olarak tanımlanan ve yaştan bağımsız olan Z rastlantı değişkeninin dayanak tehlike hızı fonksiyonunu çarpımsal olarak etkilemesidir. Bu varsayım,

$$h_0(t, Z) = Zh_0(t) \quad (3.18)$$

şeklinde gösterilmektedir. Gözlemlenemeyen risk faktörleri popülasyondaki tüm bireyleri etkilediğinden, Z rastlantı değişkeninin; beklenen değerinin bir, varyansının ise sonlu olması beklenmektedir. Hassasiyet dağılımı rastlantı değişkeninin varyansı, popülasyondaki heterojenliğin ölçüm değerini göstermektedir. σ^2 küçük bir değer ise, Z 'nin değerleri bir etrafında yoğun olarak dağılmaktadır. Eğer σ^2 büyük bir değer alıyorsa Z 'nin değerleri oldukça yayılmıştır. Popülasyondaki bütün bireyler veya risk sınıfları dayanak tehlike hızı fonksiyonundan Z sabit oranı ile ayrılmaktadır.

$h_0(t, Z)$ bireysel tehlike hızı fonksiyonunu göstermek üzere, bu fonksiyon Taylor Serisi ile açıldığında,

$$h_0(t, Z) = h_0(t, 0) + h_0'(t, 0) + o(Z)$$

şeklinde gösterilebilmektedir. Burada, $o(Z)$, birinci dereceden büyük dereceler için Taylor açılımını ifade etmektedir. Bu terim ihmal edildiğinde, bireysel tehlike hızı fonksiyonu,

$$h_0(t, Z) \approx h_0(t, 0) + Zh_0'(t, 0)$$

şeklinde yakınsamaktadır. Doğal bir varsayım yaparak sıfır rastgele etki seviyesindeki bireyler için ölümlülüğün sıfır olduğu varsayıldığında,

$$h_0(t, Z) \approx Zh_0'(t, 0) \approx Z \frac{\partial h_0(t, Z)}{\partial Z} \Big|_{Z=0} \approx Zh_0(t) \quad (3.19)$$

olur. Z hassasiyet seviyesi bilindiğinde koşullu yaşam fonksiyonu,

$$S(t|Z) = \exp \left(- \int_0^t h_0(s, Z) ds \right) = \exp \left(- Z \int_0^t h_0(s) ds \right) = \exp(-ZH_0(t)) \quad (3.20)$$

şeklinde elde edilir. $f_0(t|z) = S(t|z)h_t(z)$ bilindiğinde z rastgele etki seviyesinde T_0 'ın koşullu olasılık yoğunluk fonksiyonu,

$$f_0(t, z) = e^{-ZH(t)} z h_t = - \frac{d}{dt} S(t|z) \quad (3.21)$$

şeklindedir. Buradan T ve Z rastlantı değişkenlerinin ikili tehlike hızı fonksiyonu,

$$h_0(t, z) = f_0(t, z)g_0(z) = S(t|z)h_t(z) g_0(z) \quad (3.22)$$

şeklindedir. Şu ana kadar verilen modelde popülasyondaki bir birey tanımlanmıştır. Buna rağmen, bireysel olarak model gözlemlenemez. Sonuç olarak, modelin popülasyon için dikkate alınması gerekmektedir. Tüm popülasyon için yaşam fonksiyonu, hassasiyet dağılımı bakımından bireysel yaşam fonksiyonlarının ortalamasına eşittir. Böylece popülasyon için ortalama yaşam fonksiyonu,

$$\bar{S}(t) = \int_0^{\infty} S(t|z)g_0(z)dz \quad (3.23)$$

şeklinde yazılabilir. Burada, $\bar{S}(t)$, yeni doğan bireylerden (t) yaşına ulaşanların oranını göstermektedir.

Popülasyonda bireyin (x) yaşında olduğu düşünülüğünde, bu birey için hassasiyet seviyesi ile yeni doğan bir birey için hassasiyet seviyesi arasında;

$$Z_x = Z_0 | T_0 > x$$

ilişkisi bulunmaktadır. Bu durumda (x) yaşındaki birey için hassasiyet rastlantı değişkeninin olasılık yoğunluk fonksiyonu,

$$\begin{aligned} g_x(z) &= \lim_{\Delta z \rightarrow 0} \frac{\Pr(z < Z_x \leq z + \Delta z)}{\Delta z} \\ &= \lim_{\Delta z \rightarrow 0} \frac{\Pr(T_0 > x | z < Z_x \leq z + \Delta z) \Pr(z < Z_x \leq z + \Delta z)}{\Delta z \Pr(T_0 > x)} \end{aligned}$$

şeklinde elde edilir ve olasılık ifadeleri kısaltılarak yazıldığında,

$$g_x(z) = \frac{S(x|z)g_0(z)}{\bar{S}(x)} = \frac{S(x|z)g_0(z)}{\int_0^{\infty} S(x|z)g_0(z)dz} \quad (3.24)$$

olur. Eş. (3.23) ve Eş. (3.24) karşılaştırıldığında (0) yaşındaki birey için hassasiyet rastlantı değişkeninin olasılık yoğunluk fonksiyonu verildiğinde, (x) yaşındaki birey için hassasiyet dağılımının olasılık yoğunluk fonksiyonu arasında,

$$\frac{S(x|z)}{\bar{S}(x)}$$

ilişkisi vardır. (x) yaşındaki bireylerden oluşan popülasyon için tehlike hızı fonksiyonu,

$$\bar{h}_x(z) = \frac{\int_0^{\infty} h_x(z)S(x|z)g_0(z)dz}{\int_0^{\infty} S(x|z)g_0(z)dz} = \frac{\int_0^{\infty} h_0(x,z)dz}{\bar{S}(x)}$$

şeklindedir ve bu popülasyon için tehlike hızı fonksiyonu,

$$\begin{aligned}
\bar{h}_x(z) &= \frac{\int_0^\infty h_x(z)S(x|z)g_0(z)dz}{\int_0^\infty S(x|z)g_0(z)dz} = \frac{\int_0^\infty z h_x(z)S(x|z)g_0(z)dz}{\int_0^\infty S(x|z)g_0(z)dz} = \frac{h_x \int_0^\infty z S(x|z)g_0(z)dz}{\int_0^\infty S(x|z)g_0(z)dz} \\
&= \frac{h_x \int_0^\infty z \bar{S}(x)g_x(z)dz}{\bar{S}(x)} = \frac{h_x(z)\bar{S}(x) \int_0^\infty z g_x(z)dz}{\bar{S}(x)} = h_x \int_0^\infty z g_x(z)dz = h_x \bar{z}_x \quad (3.25)
\end{aligned}$$

şeklinde basit bir şekilde ifade edilir. Popülasyon için ortalama hassasiyet dağılımının bulunulan yaşa göre türevi alındığında,

$$\frac{d}{dx} \bar{h}_x(z) = -h_x \text{Var}(Z_x) < 0$$

sonucu elde edilmektedir. Bu durumda; popülasyon için ölüm hızı, popülasyondaki bireylerin ölüm hızlarından daha hızlı değişmektedir. Bu durum, hassasiyet seviyesi popülasyonun ortalama hassasiyet seviyesinden büyük olan bireyin popülasyondan erken ayrılması nedeniyle, popülasyonun ortalama hassasiyet seviyesinin yaş ilerledikçe azalmasına neden olmaktadır.

Hassasiyet modellerinde kullanılan en önemli varsayım, hassasiyet dağılımı için kullanılacak dağılımın belirlenmesidir. Literatüre bakıldığında; Gamma, Ters Gauss, Pozitif Durağan Dağılım, Ters Gauss Dağılımı, Güç Varyans Fonksiyonu Dağılımı, Bileşik Poisson Dağılımı ve Lognormal dağılım yaygın olarak kullanılmaktadır. Lognormal dağılımın Laplace Dönüşümü'nün olmaması nedeniyle, bu kesimde Lognormal dağılım hariç diğer dağılımlar incelenecektir.

3.4.1. Gamma dağılımı

Hassasiyet dağılımı olarak, Gamma dağılımı çok yaygın bir şekilde kullanılmaktadır. Greenwood ve Yule (1920), Congdon (1995), dos Santos et al. (1995) ve Hougaard (2000) örnek olarak verilebilir. Analitik açıdan bakıldığında; koşulsuz yaşam, olasılık yoğunluk ve ölüm hızı fonksiyonlarının kolay elde edilmesi nedeniyle ölüm zamanı verisine uyum sağlamaktadır. Bu durum Laplace Dönüşümü'nün basit olarak elde edilmesinden kaynaklanmaktadır. Bu dağılımın kullanılmasının bir diğer önemli nedeni de biçim parametresi için farklı değerler kullanılarak, olasılık yoğunluk fonksiyonunun değişik şekillerde elde edilmesidir. İki parametrelili Gamma dağılımı'nın olasılık yoğunluk fonksiyonu,

$$f_z(z) = \frac{\gamma^\delta z^{\delta-1} \exp(-\gamma z)}{\Gamma(\delta)}$$

şeklindedir. Burada $\delta > 0$ biçim parametresini ve $\gamma > 0$ ölçek parametresini göstermektedir. Laplace Dönüşümü ise,

$$\begin{aligned} L(s) &= \int_0^\infty \exp(-ts) f_z(z) dz \\ &= \gamma^\delta (s + \gamma)^{-\delta} \end{aligned}$$

ile tanımlanmaktadır. Laplace Dönüşümü sıfır komşuluğa sahip olduğundan, Laplace Dönüşümü kullanılarak hassasiyet etkisi dağılımının beklenen değer ve varyansı;

$$\begin{aligned} E(Z) &= \delta / \gamma \\ \text{Var}(Z) &= \delta / \gamma^2 \end{aligned}$$

şeklinde elde edilir. Burada (0) yaşı için hassasiyet etkisi dağılımının beklenen değerinin bire eşit olması gerekliliği nedeniyle, biçim ve ölçek parametrelerinin birbirine eşit ve varyansının sonlu olması gerekmektedir. Bu durumda $\delta = \gamma$ olacak ve Gamma dağılımı'nın varyansı,

$$\text{Var}(Z) = \frac{1}{\gamma} = \theta \quad (3.26)$$

şeklinde tanımlandığında Laplace Dönüşümü

$$L(s) = (1 + \theta s)^{-1/\theta} \quad (3.27)$$

olacaktır. Bu durumda, hassasiyet dağılımının Gamma dağılıma uyması durumunda; popülasyonun koşulsuz ölüm hızı fonksiyonu,

$$h(t) = \frac{h_0(t)}{1 + \theta H_0(t)} \quad (3.28)$$

ve yaşam fonksiyonu,

$$S(t) = (1 + \theta H_0(t))^{-1/\theta} \quad (3.29)$$

şeklindedir. Hassasiyet etkisi dağılımı, Gamma dağılımı olarak belirlendiğinde popülasyonun ortalama hassasiyet etkisi seviyesi $\delta/\gamma(x)$ olmaktadır. Yaş ilerledikçe

ortalama hassasiyet etkisi seviyesi azalmaktadır. Burada, δ 'nın küçük olması heterojenlik seviyesinin yüksek olması ve popülasyonun ortalama hassasiyet etkisi seviyesinin daha çabuk düşmesi anlamına gelmektedir. Bu durum, hassasiyet modellerinin seçim etkisini daha önemli kılmaktadır. Bu durumda hassasiyet etkisi seviyesinin varyansı da $\delta/(\gamma(x))^2$ yaş ilerledikçe azalmasına neden olmaktadır. Fakat Gamma hassasiyet dağılımı için değişim katsayısı, $\sqrt{1/\delta}$, sabittir ve yaş ile değişmemektedir. Bu durum sadece Gamma hassasiyet dağılımı'na özgü bir özelliktir. Diğer hassasiyet dağılımları için değişim katsayısı yaş ilerledikçe, azalma eğilimi göstermektedir.

Gamma dağılımının hassasiyet etkisi dağılımı olarak kullanımı, literatürde çok fazla yer almaktadır. Vaupel et al. (1979), heterojen popülasyonlar için hayat tablosunun doğrulanmasında; Manton et al. (1981), heterojen popülasyonun ölüm verisini kullanarak ölümlülük yapısını karşılaştırmak için; Manton, Stallard ve Vaupel (1981), Amerika'daki siyah ve beyaz bireylerin ölümlülükleri arasındaki ilişkiyi bulmak için; Manton et al. (1986), sağlık verisini kullanarak ileri yaşlara ilişkin yaşam analizi yapmak için ve Jones (1988), hayat sigortacılığında seçilmiş ayrılma etkisinin ölümlülük üzerindeki etkisini incelemek için kullanmışlardır.

3.4.2. Pozitif Durağan dağılım

Aynı dağılımdan elde edilmiş n tane birbirinden bağımsız rastlantı değişkeninin toplamının standartlaştırılmış hali, pozitif durağan dağılım olarak adlandırılmaktadır. Burada, standartlaştırma θ 'nın $(0,1]$ aralığında olması koşulunda $n^{1/\theta}$ ile yapılmaktadır. Olasılık yoğunluk fonksiyonu ve yaşam fonksiyonu kapalı bir formda olmamasına rağmen Laplace Dönüşümü,

$$L(s) = \exp(-s^\theta) \quad (3.30)$$

şeklinde yazılabilmektedir. Pozitif durağan dağılımların Laplace Dönüşümü'nün sıfırda komşuluğu bulunmamaktadır. Çünkü Laplace Dönüşümü'nün sağdan limiti alındığında, Laplace Dönüşümü'nün değeri eksi sonsuz değerine ulaşmaktadır. Yani iraksaktır. Bu yüzden, bu dağılımın tüm momentleri sonsuz değerini almaktadır. Tek değişkenli hassasiyet modellerinden hassasiyet etkisini gösteren rastlantı değişkeninin değerinin dörde eşit veya küçük olması olasılığının 0,90'dan büyük olması durumunda, beklenen değer momenti sonsuz değerini almamaktadır. Bu

durum Pozitif durağan dağılımın, Normal dağılıma göre daha sivri bir dağılım olmasından kaynaklanmaktadır. Çünkü pozitif durağan dağılımının kuyruk ağırlığının sıfıra çok yakın bir değer alması beklenmektedir. Bu durumda pozitif durağan dağılım için popülasyonun koşulsuz yaşam fonksiyonu,

$$S(t) = L(H_0(t)) = \exp(-H_0(t)^\theta) \quad (3.31)$$

olasılık yoğunluk fonksiyonu,

$$f(t) = \theta h_0(t) H_0(t)^{\theta-1} \exp(-H_0(t)^\theta) \quad (3.32)$$

ve tehlike hızı fonksiyonu,

$$h(t) = \theta h_0(t) H_0(t)^{\theta-1} \quad (3.33)$$

şeklindedir.

Pozitif durağan hassasiyet dağılımı, Hougaard (1986b) çalışmasında tanıtılmış ve Manatunga ve Oakes (1999) çalışmasında uygulanmıştır.

3.4.3. Ters Gauss dağılımı

Ters gauss dağılımının olasılık yoğunluk fonksiyonu,

$$f_z(z) = \sqrt{\frac{\alpha}{2\pi}} z^{-3/2} \exp\left(\frac{-\alpha}{2z\mu^2} (z - \mu)^2\right)$$

şeklindedir. Burada $\mu > 0$ ve $\alpha > 0$ olarak tanımlanmaktadır. Ters Gauss dağılımının Laplace Dönüşümü,

$$L(s) = \exp\left(\frac{\alpha}{\mu} - \left(\frac{\alpha^2}{\mu^2} + 2\alpha s\right)^{1/2}\right)$$

beklenen değer ve varyansı ise,

$$E(Z) = \mu \text{ ve } \text{Var}(Z) = \frac{\mu^3}{\alpha}$$

şeklindedir. Hassasiyet etkisi dağılımının birinci moment değerinin bir olması beklendiği için $\mu = 1$ olarak alınmaktadır. Bu durumda α sonsuza giderken,

heterojenlik durumunun ortadan kalktığı sonucuna ulaşılır. Ters gauss dağılımının beklenen değer ve varyansından $\mu = 1$ olarak alındığında,

$$E(Z) = 1 \text{ ve } \text{Var}(Z) = \frac{1}{\alpha} = \sigma^2 \quad (3.34)$$

elde edilir. Bu durumda, popülasyon için koşulsuz yaşam fonksiyonu;

$$S(t) = \exp\left(\frac{1}{\sigma^2}(1 - \sqrt{1 + 2\sigma^2 H_0(t)})\right) \quad (3.35)$$

ve ölüm hızı fonksiyonu;

$$h(t) = \frac{h_0(t)}{(1 + 2\sigma^2 H_0(t))^{1/2}} \quad (3.36)$$

şeklinde elde edilir.

Ters Gauss hassasiyet etkisi dağılımı, Hougaard (1984) çalışmasında tanıtılmış ve Manton ve Vaupel (1995), Keiding et al. (1997), Klein ve Moeschberger (1997) ile Price ve Manatunga (2001) çalışmalarında kullanılmıştır.

3.4.4. Güç Varyans Fonksiyonu dağılımı

Gamma, Ters Gauss ve Pozitif Durağan dağılımları da içeren, oldukça geniş bir hassasiyet dağılımı ailesidir. Bu dağılımın üç parametresi bulunmakta ve Laplace Dönüşümü,

$$L(s) = \exp\left(-\frac{k}{\gamma} - ((\lambda + s)^\gamma - \lambda^\gamma)\right)$$

şeklindedir. Buradan Güç Varyans Fonksiyonu dağılımının beklenen değeri,

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

ve varyansı,

$$\text{Var}(Z) = k(1 - \gamma)\lambda^{\gamma-2}$$

şeklinde elde edilir. Buradan da popülasyonun koşulsuz yaşam fonksiyonu,

$$S(t) = \exp\left(-\frac{k}{\gamma}((\lambda + H_0(t))^\gamma - \lambda^\gamma)\right)$$

ve koşulsuz ölüm hızı fonksiyonu,

$$h(t) = k h_0(t)(\lambda + H_0(t))^{\gamma-1}$$

elde edilir. Hassasiyet etkisi dağılımının beklenen değerinin bir olması durumunda,

$$E(Z) = k \lambda^{\gamma-1} = 1 \quad (3.37)$$

$$\text{Var}(Z) = k(1-\gamma)\lambda^{\gamma-2} = \frac{k\lambda^{\gamma-1}}{\lambda}(1-\gamma) = \frac{1-\gamma}{\lambda} = \sigma^2 \quad (3.38)$$

yazılır. Bu durumda koşulsuz ölüm hızı fonksiyonu,

$$h(t) = \frac{h_0(t)}{(1 + \frac{\sigma^2}{1-\gamma} H_0(t))^{\frac{1}{1-\gamma}}} \quad (3.39)$$

olur. Güç varyans fonksiyonu hassasiyet dağılımı, Tweedy (1984) çalışmasında önerilmiş ve Hougaard (1986a) çalışmasında kullanılmıştır. Hougaard et al. (1992) Danimarkalı ikizlerin gelecekteki yaşam sürelerinin modellenmesinde hassasiyet dağılımı olarak, bu dağılımı kullanmışlardır.

3.4.5. Bileşik Poisson dağılımı

Bileşik Poisson hassasiyet dağılımı, X rastlantı değişkenlerinin birbirinden bağımsız ve Gamma dağılımına ve N rastlantı değişkeninin de Poisson dağılımına uyduğu ve X rastlantı değişkenlerinden bağımsız olduğu dağılımdır ve

$$Z = \begin{cases} X_1 + X_2 + \dots + X_N & , N > 0 \\ 0 & , N = 0 \end{cases}$$

şeklinde tanımlanmaktadır. Gamma ve Poisson dağılımlarının Laplace Dönüşüm'leri kullanılarak, Bileşik Poisson dağılımının Laplace Dönüşümü,

$$L(s) = \exp\left\{-\rho + \rho\left(1 + \frac{s}{\lambda}\right)^{-k}\right\}$$

şeklinde elde edilir. Burada, ρ Poisson dağılımının ortalamasını ve (k, λ) , Gamma dağılımının parametrelerini göstermektedir.

Hassasiyet etkisi dağılımı olarak Bileşik Poisson dağılımı kullanıldığında koşulsuz yaşam fonksiyonu,

$$S(t) = \exp\left(-\frac{k}{\gamma}((\lambda + H_0(t))^\gamma - \lambda^\gamma)\right) \quad (3.40)$$

ve tehlike hızı fonksiyonu ise ,

$$h(t) = k h_0(t)(\lambda + H_0(t))^{\gamma-1} \quad (3.41)$$

şeklinde olur.

Parametre değerleri $\rho = -\frac{k \lambda^\gamma}{\gamma}$ $\lambda = \lambda$ $k = -\gamma$ olarak değiştirildiğinde Laplace Dönüşümü,

$$L(s) = \exp\left\{-\frac{k}{\gamma}((\lambda + s)^\gamma - \lambda^\gamma)\right\} \quad (3.42)$$

olur. Burada γ , farklılık yaratan bir parametredir. $\gamma \geq 0$ olduğunda dağılım güç varyans dağılımına; $\gamma < 0$ olduğunda bileşik poisson dağılımına ve $\gamma = 0$ ise Gamma dağılımına uymaktadır.

Bileşik Poisson hassasiyet etkisi dağılımı, ilk olarak Aalen (1988,1992) tarafından kullanılmıştır. Aalen ve Tretli (1999), testis kanseri olanların, gelecek yaşam sürelerinin modellenmesinde, Bileşik Poisson dağılımını kullanmışlardır.

3.5. Çok Değişkenli Hassasiyet Modelleri

Hassasiyet modellerinin en çok uygulandığı diğer bir önemli alan, çok değişkenli yaşam verisi analizidir. Bu analizler için bazı örnekler aşağıdadır:

- Birbirleriyle genetik olarak bağlı olan bireylerin yaşam sürelerinin analizi,
- Birbirleriyle genetik olarak bağlı olan bireylerin aynı hastalığa tekrar yakalanma zamanının analizi,
- Bir bireyin aynı hastalığa tekrar yakalanma zamanının analizi.

Çok değişkenli modellerde; olay zamanları arasındaki var olan bağımlılık yapısı da dikkate alınmaktadır. Yaygın olarak kullanılan yaklaşım; gözlemlenemeyen veya gizli değişkenler olması durumunda gözlemlenen verinin özelliklerine göre yapılan sınıflar arasında, bağımsızlığın tanımlanmasıdır (Hougaard, 2000). Çok değişkenli yaşam modellerinde bağımlılık yapısı, gözlemlenen çoklu yaşam süresi için koşullu modellerdeki gizil değişkenden ileri gelmektedir. Örneğin; gözlemlenen eşdeğişken bilgileri farklı, genetik olarak bağlı olan bireylerin yaşam süreleri gibi. Gizil değişken için varsayılan dağılımın ortalaması alındıktan sonra gözlemlenen veri için çok değişkenli bir model oluşturulmaktadır.

Bir risk sınıfında akraba olan iki birey olsun ve eşdeğişken vektörleri birbirinden farklı X_1 ile X_2 vektörleri ile gösterilsin. Bu risk sınıfı için iki boyutlu yaşam fonksiyonu,

$$S(t_1, t_2) = \int_0^{\infty} S(t_1|z, X_1)S(t_2|z, X_2)f(z)dz$$

şeklinde tanımlanmaktadır. Burada; f , Z rastlantı değişkeninin olasılık yoğunluk fonksiyonunu göstermektedir. Çok değişkenli hassasiyet modelleri, paylaşımlı ve ilişkili hassasiyet modelleri olmak üzere iki başlık altında incelenmektedir.

3.5.1. Paylaşımlı hassasiyet modelleri

Paylaşımlı hassasiyet modelleri; genetik bağımlılığı olan bireyler, benzer organlar ve tekrar eden ölçümlerin olay zamanlarının analiz edilmesinde kullanılmaktadır. Bir risk sınıfı içerisindeki bireyler, aynı hassasiyet değerine sahiptirler. Bu yüzden bu modellere ‘paylaşımlı hassasiyet modelleri’ adı verilmektedir. Bu modeller, Clayton (1978) tarafından tanıtılmış ve Hougaard (2000) tarafından genişletilmiştir. Anlatımı basitleştirmek amacıyla, popülasyonun iki risk sınıfına ayrıldığı ve bu risk sınıflarının her birinde iki birey olduğu düşünölsün. Z hassasiyet koşulunda, herhangi bir risk sınıfı içindeki iki bireyden birisi için ölüm hızı fonksiyonu $Zh_0(t)\exp(\beta^T X)$ şeklindedir. Bu risk sınıfındaki bireyler için hassasiyetin Z değerine eşit olması, bu bireylerin yaşam sürelerinde bağımlılığa neden olmaktadır. Burada risk sınıfı içerisindeki bireylerin, gelecek yaşam süreleri bağımsız olarak varsayıldığında; bozulmuş hassasiyet modelleri kullanılmaktadır. Bozulmuş hassasiyet modellerinde, hassasiyet etkisini gösteren Z rastlantı değişkeninin ortalaması bir, varyansı ise sıfır değerine eşit olmaktadır. Hassasiyet rastlantı değişkeninin varyansının sıfırdan büyük olması

durumunda ise bu risk sınıfı içerisindeki iki birey için pozitif bağımlılık söz konusudur ve Z hassasiyet etkisi rastlantı değişkeni bilindiğinde bu risk sınıfındaki bireyler için ikili yaşam fonksiyonu,

$$S(t_1, t_2 | Z) = S_1(t_1)^Z S_2(t_2)^Z$$

şeklindedir. Paylaşımlı hassasiyet etkisi olan Z, bir risk sınıfı içerisindeki bireyler arasındaki ilişkiyi tanımlamaktadır. Bununla birlikte paylaşımlı hassasiyet modellerinin bazı sınırlamaları bulunmaktadır. Bunlar şu şekildedir :

- Herhangi bir risk sınıfı içerisindeki bireyler için gözlemlenemeyen risk faktörlerinin aynı olduğu varsayımı, her zaman gerçeği yansıtmamaktadır.
- Herhangi bir risk sınıfı içerisindeki yaşam sürelerinin bağımlılığı, yaşam sürelerinin marjinal dağılımlarından kaynaklanmaktadır.
- Tek boyutlu hassasiyet etkisi, risk sınıfı içerisinde pozitif bağımlılık oluşturmaktadır.

Paylaşımlı hassasiyet modelleri için bu kısıtlamalardan ilişkili hassasiyet modelleri kullanılarak kurtulunabilmektedir.

McGilchrist ve Aisbett (1991) katatere bağlı enfeksiyon verisini modellemek için paylaşımlı Lognormal hassasiyet modeli kullanmışlardır.

Dos Santos et al. (1995) göğüs kanserinin yeniden tekrarlanabilirliğinin modellenmesi için Gamma ve Lognormal dağılımlı paylaşımlı hassasiyet modeli kullanmışlardır.

Manatunga ve Oakes (1999) diyabet çalışmasını modellemek için Pozitif durağan dağılım kullanmışlardır.

Andersen et al. (1999) tedavi görülen yerin etkisini test etmek için paylaşımlı hassasiyet modellerini kullanarak bir çalışma yapılmıştır.

3.5.2. İlişkili hassasiyet modelleri

İlişkili hassasiyet modelleri ilk olarak iki değişkenli başarısızlık zamanı verisinin analizi için kullanılmıştır. Bir risk sınıfı için iki birey olduğu düşünüldüğünde ve bir bireyin hassasiyet rastlantı değişkeninin dağılımı bilindiğinde, diğer birey için de hassasiyet rastlantı değişkeninin bilindiği varsayılmaktadır. Literatür incelendiğinde, ilişkili hassasiyet modellerinde genel olarak; Gamma dağılımı kullanılmakta ve ilişkili hassasiyet modelleri için yaşam fonksiyonu,

$$S(t_1, t_2 | Z_1, Z_2) = S_1(t_1)^{Z_1} S_2(t_2)^{Z_2}$$

şeklinde gösterilmektedir.

Pickles et al. (1994) İngiliz ikizleri arasındaki ergenlik ve anti-sosyal davranışların başlangıç yaşının belirlenmesinde, ilişkili Gamma Hassasiyet Modeli'ni kullanmıştır.

Yashin ve Iachine (1995) ve Yashin et al. (1995) Danimarkalı ikizlerin ölümlülüğünü analiz etmek için ilişkili Gamma Hassasiyet Modeli'ni kullanmışlardır.

Wienke et al. (2001) ve Zdravkovic et al. (2002) İsveç ve Danimarkalı ikizler için koroner kalp hastalığından kaynaklanan ölümlülüğü etkileyen genetik faktörlerini analiz etmek için ilişkili Gamma Hassasiyet Modeli'ni kullanmışlardır.

Wienke et al. (2002) Danimarkalı ikizler için koroner kalp hastalığından kaynaklanan ölüme ilişkin faktörlerin analiz edilmesi için genişletme yapmıştır.

Zahl (1997) Norveç'te nedene bağlı kanser ölümlülüğünü modellemek için ilişkili Gamma Hassasiyet Modeli'ni kullanmıştır.

4. UYGULAMA

Bu bölümde, 2004-2008 yılları arasındaki Türkiye kadın ve erkek sigortalı ölüm verisi kullanılarak, Hassasiyet Modeli araştırması yapılacaktır. Modelin doğruluğu uyum iyiliği testleri ile sınanacak ve modelleme sonucunda elde edilen ölüm oranları (110) yaşına kadar uzatılacaktır.

4.1. Modelleme

Bu kesimde, Üçüncü Bölüm'de incelenen tek değişkenli hassasiyet modelinde farklı dağılımlar için tanımlanan ölüm hızı, dayanak ölüm hızı fonksiyonunun Gompertz Ölümlülük Yasası'na

$$\mu_x = b(\exp(px)) \quad , x > 0$$

uyduğu varsayıldığında; tek bir formül ile kısa bir şekilde elde edilmiştir. Burada, z hassasiyet seviyesine sahip (x) yaşındaki birey için

$$\bar{\mu}_x(z) = \frac{\exp(-g + px)}{(1 + \exp(-s + px))^\psi} \quad (4.1)$$

şeklinde tanımlanmıştır. Burada, ψ 'nin aldığı değerlere göre hassasiyet teriminin dağılımı, aşağıda verildiği gibi değişmektedir:

- $\psi = 1$ değerini aldığı anda, hassasiyet dağılımı; Gamma dağılıma uymakta ve Eş.(4.1), Gompertz/Gamma hassasiyet modelini tanımlamaktadır. Eş.(3.28) ile Eş.(4.1) arasındaki uyum, Ek 1'de verilmiştir.
- $\psi = 1/2$ değerini aldığı anda, hassasiyet dağılımı; Ters Gauss dağılıma uymakta ve Eş.(4.1), Gompertz/Ters Gauss hassasiyet modelini tanımlamaktadır. Eş.(3.36) ile Eş.(4.1) arasındaki uyum, Ek 2'de verilmiştir.
- $0 \leq \psi < 1$ değerini aldığı anda, hassasiyet dağılımı; Pozitif Durağan dağılıma uymakta ve Eş.(4.1), Gompertz/Pozitif Durağan hassasiyet modelini tanımlamaktadır. Eş.(3.33) ile Eş.(4.1) arasındaki uyum, Ek 3'te verilmiştir.
- $\psi > 1$ değerini aldığı anda, hassasiyet dağılımı; Bileşik Poisson dağılıma uymakta ve Eş.(4.1), Gompertz/Bileşik Poisson hassasiyet modelini tanımlamaktadır. Eş.(3.41) ile Eş.(4.1) arasındaki uyum, Ek 4'te verilmiştir.

- $\psi < 1$ değerini aldığında, hassasiyet dağılımı; Güç Varyans Fonksiyonu dağılımına uymakta ve Eş.(4.1), Gompertz/Güç Varyans Fonksiyonu hassasiyet modelini tanımlamaktadır. Eş.(3.39) ile Eş.(4.1) arasındaki uyum, Ek 5'te verilmiştir.

4.2. Kaba Ölüm Hızlarının Elde Edilmesi

Bu kesimde 2004-2008 yılları arasındaki kadın ve erkek sigortalı verisinden, kaba ölüm hızlarının elde edilmesi için sigortalı veri setine uygulanacak işlemler ele alınacaktır.

4.2.1. 2004-2008 yılları arasındaki Türkiye sigortalı verisi

Bu çalışmada Sigorta Bilgi Merkezi (SBM) aracılığı ile Türkiye'de faaliyet gösteren beş önemli hayat sigortası şirketinden elde edilen sigortalı verisi kullanılmıştır. Bu veri, 1 Ocak 2004 tarihinden 31 Aralık 2008 tarihine kadar düzenlenmiş toplam **13 332 687** poliçeden oluşmaktadır.

4.2.2. Eksik ve hatalı veri sorununun ortadan kaldırılması

Veri kümesi denetimden geçirilerek, tekrarlı ve eksik bilgiler belirlenmiştir. Eksik ve hatalı veri sorununun ortadan kaldırılabilmesi için ilk olarak veri kaynağına geri dönüş yapılarak, yeniden sorgulama yapılmıştır. Bu sorgulama sonucunda, düzeltilemeyen gözlemler sigortalı veri kümesinden çıkarılmıştır.

4.2.3. Tekrarlama hatasının ortadan kaldırılması

Tekrarlama hatasının ortadan kaldırılabilmesi için T.C. Kimlik Numaraları'na göre sorgulama yapılmıştır. Eksik ve hatalı gözlemler sigortalı veri setinden çıkarılıp tekrarlama hataları da ortadan kaldırıldıktan sonra sigortalı veri setinde **3 803 022** erkek sigortalı birey ve **999 716** kadın sigortalı birey kalmıştır.

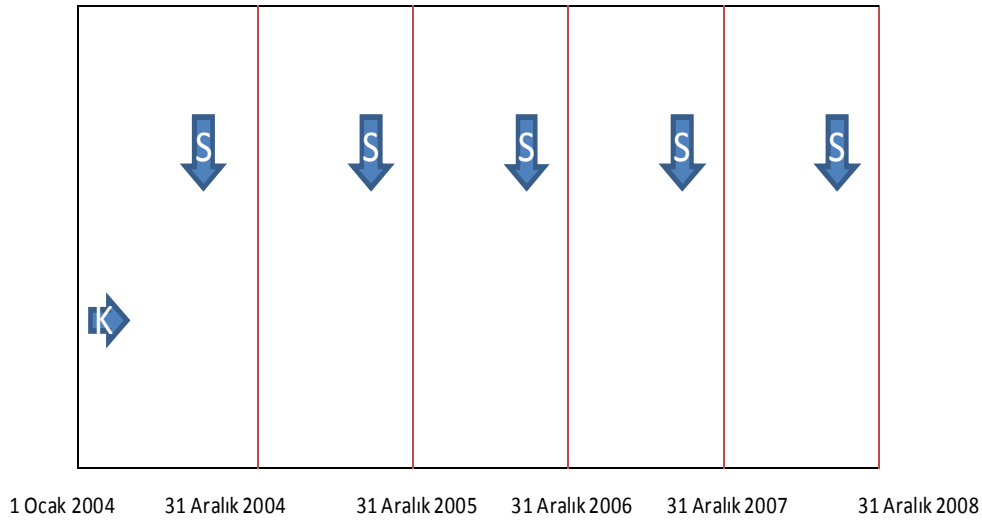
4.2.4. Riske maruz kalan birim sayısının elde edilmesi

Düzeltilmelerden sonra kalan sigortalı verisinden, yaşa göre ölüm hızlarının, μ_x , elde edilmesi gerekmektedir.

Burada, her bir takvim yılı; gözlem dönemi (observation period) olarak adlandırılmaktadır. Herhangi bir gözlem döneminde gözlemlerin tam olması, (x)

yaşındaki bireyin $(x+1)$ yaşına ulaşana kadar veya bu gözlem dönemi içerisinde ölene kadar gözlemlendiği anlamına gelmektedir.

1 Ocak 2004, çalışmanın başladığı tarih olarak kabul edilmiştir. Yani, 1 Ocak 2004 tarihinden önce gözlemlenen ölümler ve sigorta sistemine giriş-çıkış yapan gözlemler dikkate alınmayarak soldan kesme işlemi yapılmıştır. 31 Aralık 2004, 31 Aralık 2005, 31 Aralık 2006, 31 Aralık 2007 ve 31 Aralık 2008 tarihlerinde sağdan sansürleme yapılmıştır. Gözlemlere sağdan sansürleme yapılarak her bir gözlem dönemi için yaşlara göre ölüm hızları elde edilmiştir. Kesme ve sansürleme işlemleri Şekil 4.1’de gösterilmiştir:



Şekil 4.1. Sansürleme ve Kesme

Şekil 4.1’de yatay çizgiler, gözlem altındaki bireyleri yaşlara göre sınıflandırmak için, dikey çizgiler ise her bir gözlem dönemini birbirinden ayırmak için kullanılmıştır.

Şekil 4.1’deki gözlem dönemlerinden herhangi birinde n adet gözlemin bulunduğu ve gözlemlerin $i = 1, 2, \dots, n$ şeklinde etiketlendiği düşünölsün. Burada i ’inci gözlemin $(x + t_i)$ yaşında gözlem dönemine girdiği ve $(x + s_i)$ yaşında gözlem döneminden ölüm veya başka bir nedenden dolayı ayrıldığı düşünöldüğünde, $(0 \leq t_i < s_i \leq 1)$ o gözlem dönemi için (x) yaşında riske maruz kalan birim sayısı,

$$E_x = (s_1 - t_1) + (s_2 - t_2) + \dots + (s_n - t_n) \quad (4.2)$$

eşitliğinden hesaplanmaktadır. Burada, $0 < s_i - t_i \leq 1$, $i = 1, 2, \dots, n$ olmaktadır.

4.2.5. Ölen kişi sayısının bulunması

Herhangi bir gözlem döneminde, gözlemlenen her ölüm dikkate alınmaktadır. D_x , herhangi bir gözlem döneminde (x) yaşında ölen kişi sayısını göstermektedir.

4.2.6. Gözlenen ölüm hızının elde edilmesi

Bu çalışmada ölümlerin yıl içerisinde tekdüze dağıldığı varsayımı kullanılarak, (x) yaşındaki bireyler için kaba ölüm hızı,

$$\mu_x = \frac{D_x}{E_x - 0,5D_x}$$

şeklinde hesaplanmaktadır (Butt ve Haberman, 2002).

Her bir gözlem dönemi için elde edilen riske maruz kalan birim sayılarının ortalaması alınarak, tüm çalışma için kullanılacak ortalama riske maruz kalan birim sayıları, $\bar{E}_x$, elde edilir. Aynı şekilde her bir gözlem dönemi için elde edilen ölüm sayılarının ortalaması alınarak, tüm çalışma için kullanılacak ortalama ölüm sayıları, $\bar{D}_x$, değerleri elde edilir.

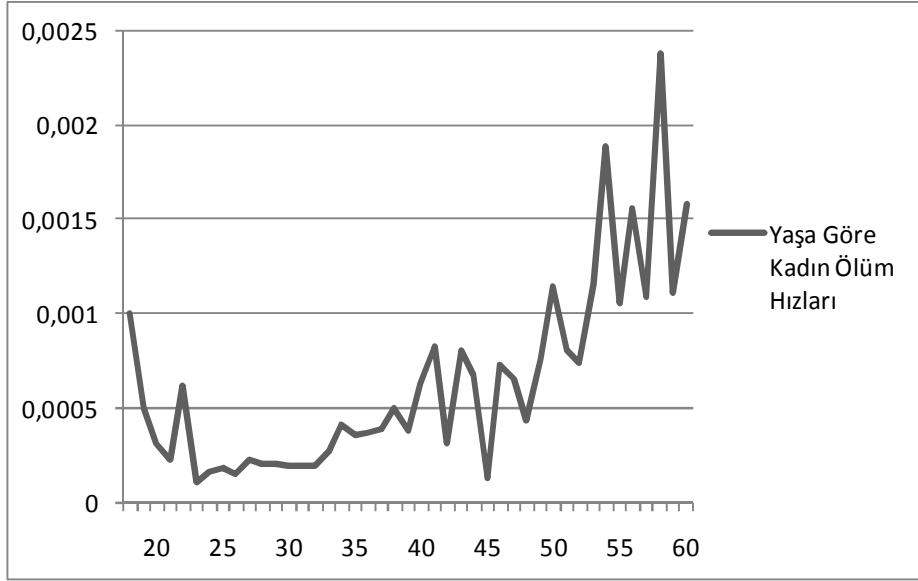
Kadın ve erkek sigortalılar için elde edilen ortalama ölüm sayıları ve ortalama riske maruz kalan birim sayıları kullanılarak, kadın ve erkek sigortalı bireyler için yaşa göre ölüm hızları,

$$\bar{\mu}_x = \frac{\bar{D}_x}{\bar{E}_x - 0,5\bar{D}_x} \quad (4.3)$$

şeklinde hesaplanır.

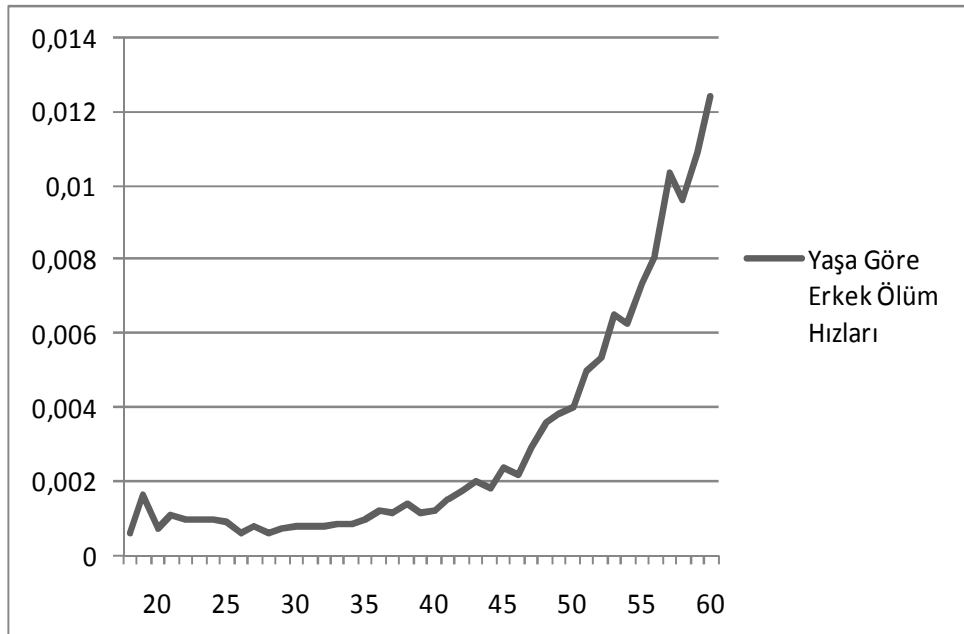
Kadın ve erkek sigortalıların ölüm hızları incelenirken, tek ortak risk faktörü yaş olarak alındığından; yaşın ölüm hızlarını açıklama oranlarının da incelenmesi gerekmektedir. Bu incelemenin yapılabilmesi için Spearman'ın sıra korelasyon katsayısından (Spearman, 1904) yararlanılmaktadır. Bu değer, iki sıralı değişken arasındaki ilişki derecesinin ölçümünü göstermektedir.

Eş. (4.3) kullanılarak (18)-(60) yaş aralığındaki kadın ve erkek sigortalılar için elde edilen yaşa göre ölüm hızları, sırasıyla, Şekil 4.2 ve Şekil 4.3'te gösterilmiştir :



Şekil 4.2. Kadın sigortalı bireyler için yaşa göre ölüm hızları

Şekil 4.2 incelendiğinde (18) yaşından (25) yaşına kadar kadın sigortalılar için ölüm hızlarında hızlı bir düşüş görülmektedir. (25) yaşından (40) yaşına kadar ölüm hızlarından hafif bir artış eğilimi vardır. (40) yaşından (60) yaşına kadar ise, ölüm hızları fazla dalgalanmış hem de ölümlülük eğilimi artmıştır.



Şekil 4.3. Erkek sigortalı bireyler için yaşa göre ölüm hızları

Şekil 4.3 incelendiğinde erkek sigortalıların ölüm hızlarının, kadın sigortalıların ölüm hızlarına göre daha durağan bir yapı izlediği görülmektedir. Fakat (45) yaşından sonra ölümlülük eğilimi hızlı bir şekilde artmaktadır.

Kadın sigortalılar için Spearman'ın sıra korelasyon katsayısı 0,8736 ve erkek sigortalı bireyler için Spearman'ın sıra korelasyon katsayısı 0,9802 olarak elde edilmiştir. Diğer bir deyişle, kadın sigortalıların ölüm hızlarındaki değişimi yaş faktörü % 87,36 ve erkek sigortalıların ölüm hızlarındaki değişimi yaş faktörü % 98,02 oranında açıklayabilmektedir. Bu oranlar kabul edilebilir olduğundan sadece yaş faktörü dikkate alınarak ölüm hızlarına ilişkin model araştırması yapılabilir.

4.3. Parametre Tahmini İçin Kullanılacak Yöntem

Dayanak ölüm hızı ve hassasiyet dağılımının parametrelerinin tahmin edilebilmesi için kullanılan temel yaklaşım, gözlenen popülasyon ölümlülük verisi için ortalama hassasiyet etkisi yaklaşımıdır (Vaupel et al., 1985a ; Butt ve Haberman, 2002).

Ortalama hassasiyet etkisi yaklaşımında, tüm popülasyon için gözlenen ölüm hızının; her bir risk sınıfı için elde edilmiş ölüm hızlarının ortalamasına eşit olduğu varsayılmaktadır.

Ortalama hassasiyet etkisi yaklaşımında, diğer bir önemli varsayım ise ölümlülüğün yıl içerisinde $\bar{\mu}_x \bar{E}_x$ ortalama ile Poisson dağılıma sahip olmasıdır. Bu varsayım ile her bir yaş için riske maruz kalan birim sayılarının, yeterli büyüklüğe sahip olduğu düşünülmektedir.

Hassasiyet modellerinde dayanak ölüm hızı ve hassasiyet etkisi dağılımının ayrı ayrı belirlenmesi oldukça zordur (Elbers ve Ridder, 1982). Ortalama hassasiyet etkisi yaklaşımında hassasiyet modelinin parametre tahminleri, dayanak ölüm hızının seçimine bağlıdır. Hassasiyet modellerinde hassasiyet etkisinin dağılımı ise, dayanak ölüm hızı ile kaba ölüm hızları arasında bağlantı kurulmasında kullanılmaktadır.

Bu çalışmada Eş. (4.1)'de tanımlanmış hassasiyet modelinin parametrelerinin tahmin edilebilmesi için Genelleştirilmiş Doğrusal Modeller kullanılacağından Genelleştirilmiş Doğrusal Modeller'e ilişkin bilgi verilecektir.

4.3.1. Genelleştirilmiş Doğrusal Modeller

4.3.1.1. Giriş

Genelleştirilmiş doğrusal modeller klasik doğrusal modellerin genelleştirilmiş halidir (Haberman ve Renshaw, 1996). Bir bireyin veya popülasyonun ölümlülüğünü etkileyen çok fazla sayıda risk faktörü bulunmaktadır. Yaşam analizinde bazı risk faktörlerinin etkisi dikkate alınırken diğer risk faktörlerinin etkisi göz ardı edilmektedir. Etkisinin dikkate alındığı faktörler açıklayıcı değişken olarak adlandırılmaktadır. Açıklayıcı değişkenler de ikiye ayrılmaktadır. Bunlar nitel ve nicel açıklayıcı değişkenlerdir. Bunlardan nicel açıklayıcı değişkenler, eşdeğişken ve nitel açıklayıcı değişkenler ise risk faktörü düzeyi olarak adlandırılmaktadır. Yaşam analizinde açıklayıcı değişkenler ve açıklayıcı değişkenlerden kaynaklanan bağımlılığın bulunarak analiz edilebilmesi için genelleştirilmiş doğrusal modeller gibi çok değişkenli modeller kullanılmaktadır.

Genelleştirilmiş doğrusal modellerin amacı, bağımlı değişken Y ile bağımsız değişkenler olarak tanımlanan X 'ler arasındaki ilişkinin açıklanmasıdır. Aktüerya bilimlerinde genelleştirilmiş doğrusal modeller, 1800'lü yılların başından bugüne kadar kullanılmaktadır (Johnson ve Hey, 1971; Grimes, 1971; Bennett, 1978; Baxter et al., 1980; Coutts, 1984). Genelleştirilmiş doğrusal modellerde temel varsayım verinin tek parametrelili üstel dağılım ailesine ait bir dağılımdan alınmasıdır (Haberman ve Renshaw, 1996). GLM yönteminin diğer varsayımları ise

- Gözlemler birbirinden bağımsız veya en azından ilişkili olmaması
- Modelde tek hata terimi olması

şeklindedir (McCullagh ve Nelder, 1989). Genelleştirilmiş doğrusal modellerde ilk olarak veri seti kümesi, açıklayıcı değişkenler ve modele ilişkin parametrelere ilişkin eşitlikler oluşturulur. Bu şekilde oluşturulmuş eşitlik sistemi optimal çözüme en uygun değer elde edilene kadar adimsal olarak çözülür. Genelleştirilmiş doğrusal modellerin bazı dezavantajları aşağıda verilmiştir :

- Açıklayıcı değişkenlerin, modelin parametre tahminleri üzerindeki etkisinin anlamlı olup olmadığına ilişkin testler yapılamaz.
- Model parametre tahminleri için güven aralığı oluşturulamaz.

Genelleştirilmiş doğrusal modellerin dezavantajlarının yanında avantajları da vardır:

- Ölümlülük modeline ilişkin parametrelerin tahmin edicilerine ilişkin bir takım varsayımlar yapılmasına olanak sağlar.
- Genelleştirilmiş doğrusal modeller ile yapılan çözümlemeler iteratif yöntemlere göre daha etkindir.
- Genelleştirilmiş doğrusal modeller, yalnızca anlamlı değişkenlerin seçilmesine ve model varsayımlarının doğrulanmasına yardımcı olan istatistiksel kontrol mekanizmasına sahiptir.
- Klasik doğrusal modellerde kullanılan bağımsız ve açıklayıcı değişkenlerin varyansı, sabit Normal dağılıma uyması kısıtı bulunmamaktadır.
- Açıklayıcı değişkenlerin toplanabilir olması varsayımı kullanılmamaktadır.
- Bağımlı değişken Y rastlantı değişkeninin negatif olan veya negatif olmayan değerler alması zorunluluğu bulunmamaktadır.

4.3.1.2. GLM'in yapısı

Y_i , i ' nci bağımlı değişkeni göstermek üzere birbirinden bağımsız n sayıda bağımlı değişkenden oluşan vektör Y ile tanımlanmaktadır. Doğrusal modelde, bağımlı değişkenin aldığı değerlerden oluşan Y vektörü, açıklayıcı değişkenlerden oluşan X vektörü ve bilinmeyen parametreler β vektörü ile gösterildiğinde, bağımlı değişken ile açıklayıcı değişkenler arasında

$$Y = X\beta + \varepsilon \quad (4.4)$$

biçiminde bir ilişki tanımlanır. Burada, ε hata terimleri vektörünü göstermektedir. Genelleştirilmiş doğrusal modellerin temel unsuru, bağımsız değişken vektörü X ile, bağımlı değişken Y ile açıklayıcı değişkenler arasındaki bağlantının kurulmasıdır.

Genelleştirilmiş doğrusal modellerdeki temel amaç, ε hata terimi vektörü elemanlarının kareler toplamını en küçük yapan β vektörünün elemanlarının bulunmasıdır. Gözlem sayısı n ve modeldeki parametre sayısı k ise, β vektörü k

sayıda, ε vektörü n sayıda bileşenden oluşur. Burada modelin parametre sayısının gözlem sayısından küçük olması gerekmektedir.

Üstel dağılım ailesi, iki parametrelili fonksiyonlar kümesi olup, Y bağımlı değişkeninin i sayıdaki elemanlarından her birine ilişkin olasılık yoğunluk fonksiyonu,

$$f(y_i; \theta, \phi) = \exp((y_i \theta_i - b(\theta_i))/a(\phi) + c(y_i, \phi_i)) \quad (4.5)$$

ile tanımlanmaktadır. Eşitlikteki $a(\cdot)$, $b(\cdot)$ ve $c(\cdot)$ fonksiyonları tanımlı fonksiyonlar olup, bu fonksiyonlara ilişkin birtakım varsayımlar yapılmaktadır. Bu fonksiyonların tanımlanabilmesi için yeterli koşul; bağımlı değişkenin olasılık yoğunluk fonksiyonunun, bağımlı değişkenin aldığı tüm değerler için integrali alındığında elde edilen değerlerin bir eşi olmasıdır. Bu durumda farklı $a(\cdot)$, $b(\cdot)$ ve $c(\cdot)$ fonksiyonları için farklı çözümler elde edilebilmektedir. $a(\cdot)$, $b(\cdot)$ ve $c(\cdot)$ fonksiyonlarının tüm gözlem değerleri için aynı olduğu varsayıldığında, tüm gözlemlerin de aynı dağılıma sahip olduğu kabul edilmektedir. Burada $a(\cdot)$ fonksiyonu, pozitif ve sürekli; $b(\cdot)$ fonksiyonu, θ 'ya göre ikinci türevi pozitif olmak koşuluyla iki kez türevlenebilir bir fonksiyon ve $c(\cdot)$ fonksiyonu, θ parametresinden bağımsız bir fonksiyondur.

Eş. (4.5)' de yer alan θ parametresi doğal parametre ve ϕ parametresi ölçek veya yayılım parametresi olarak tanımlanmaktadır. Her bir gözlem değeri için doğal parametre değişebildiği halde, ölçek parametresi tüm gözlemler için sabit olmaktadır. Doğal parametrenin gözlemden gözleme değişmesi, ortalamanın da gözlemden gözleme değişmesine olanak sağlamaktadır. Poisson dağılımı için ölçek parametresi 1 olarak alınmaktadır. Üstel dağılım ailesine ait diğer dağılımlar için ise ölçek parametresi bilinmemekte bu yüzden veri kümesi kullanılarak en çok olabilirlik yöntemiyle tahmin edilmesi ya da genelleştirilmiş Pearson ki-kare istatistiğinin kullanılması gerekmektedir.

Üstel dağılım fonksiyonunun yapısında var olan $a(\phi)$ fonksiyonu, ϕ/w kullanılmaktadır. Yani ölçek parametresinin başlangıçta tanımlanan ağırlığa olan oranıdır.

Üstel dağılım ailesine mensup dağılımlardan birine sahip olan bağımlı değişkene ilişkin ortalama

$$E(Y_i) = \mu_i = b'(\theta) \quad (4.6)$$

ve varyans

$$\text{Var}(Y_i) = b''(\theta)a(\phi) \quad (4.7)$$

şeklinde ifade edilmektedir. Varyans eşitliğinden de görüldüğü üzere bağımlı Y değişkenine ilişkin varyans, ortalamanın bir fonksiyonu olarak tanımlanabilmektedir. Bu yüzden bağımlı Y değişkenine ilişkin varyans $V(\mu)$ ile gösterilmektedir (De Jong ve Heller, 2008). Genelleştirilmiş doğrusal modelin biçimini belirleyen faktörler, bağ fonksiyonu, varyans fonksiyonu ve tahmin edilen veya bilinen ölçek parametresidir. Üstel dağılım ailesi içinde yer alan bazı dağılımlar için $a(\phi)$, $b(\theta)$ ve $c(y, \phi)$ fonksiyonları Çizelge 4.1’ de verilmiştir:

Çizelge 4.1. Üstel Dağılım Ailesi İçin Fonksiyonlar

Dağılım	$a(\phi)$	$b(\theta)$	$c(y, \phi)$
Normal Dağılım	ϕ/w	$\theta^2/2$	$-1/2(wy^2 / \phi + \ln(2\pi\phi / w))$
Poisson Dağılımı	ϕ/w	$\exp(\theta)$	$-\ln y!$
Gamma Dağılımı	ϕ/w	$-\ln(-\theta)$	$(w/\phi)\ln(wy/\phi) - \ln y - \ln(\Gamma(w/\phi))$
Binom Dağılımı	ϕ/w	$m \ln(1 + e^\theta)$	$\ln \binom{m}{y}$
Ters Gauss Dağılımı	ϕ/w	$-\sqrt{-2\theta}$	$-1/2(\ln(2\pi\phi y^3 / w) + w / (\phi y))$

Genelleştirilmiş doğrusal modeller, veri yapısına bağlı olarak seçilen üstel dağılım ailesindeki dağılımlardan biri kullanılarak oluşturulur. Y_i , i’ nci gözleme ilişkin bağımlı değişkeni göstermek üzere bu bağımlı değişkenin ortalaması;

$$\mu_i = E(Y_i) = g^{-1}(\sum_j X_{ij}\beta_j + \xi_i) \quad i = 1, 2, \dots, n ; j = 1, 2, \dots, k \quad (4.8)$$

ve varyansı;

$$\text{Var}(Y_i) = \frac{\phi V(\mu_i)}{w_i} \quad (4.9)$$

şeklinde ifade edilir. Genelleştirilmiş doğrusal modellerin yapısında yer alan ξ_i , i' nci gözleme ilişkin bilinen hata terimidir. Bu terim, herhangi bir açıklayıcı değişkenin bağımlı değişken üzerindeki etkisinin bilindiği durumlarda kullanılır. Bilinen hata teriminin ilgili modelin doğrusal tahmin edici fonksiyonuna eklenmesiyle

$$\eta = X\beta + \xi \quad (4.10)$$

eşitliği elde edilir. Doğrusal tahmin edici fonksiyonu Eş. (4.4) ile tanımlanan bağımlı değişkenin beklenen değeri,

$$E(Y) = \mu = g^{-1}(\eta) = g^{-1}(X\beta + \xi) \quad (4.11)$$

eşitliği ile elde edilmektedir. i'inci gözleme ilişkin ortalama, i'inci gözleme ilişkin doğrusal tahmin edicinin bir fonksiyonu olarak ifade edilmektedir. Doğrusal tahmin edici için belirlenen bu fonksiyon genellikle bağ fonksiyonunun ters fonksiyonudur ve

$$\mu_i = g^{-1}(\eta_i) \quad (4.12)$$

şeklinde ifade edilir. Genelleştirilmiş doğrusal modellerde kullanılan bağ fonksiyonunun, türevlenebilirlik ve tekdüzelik koşullarını sağlaması gerekmektedir. Çizelge 4.2' de genelleştirilmiş doğrusal modellerde kullanılan bazı bağ fonksiyonları ve bu bağ fonksiyonlarının tersi gösterilmiştir.

Çizelge 4.2. Bağ Fonksiyonları ve Bağ Fonksiyonlarının Ters Fonksiyonları

Bağ Fonksiyonu	$g(x)$	$g^{-1}(x)$
Birim fonksiyon	X	X
Logaritmik fonksiyon	$\ln(x)$	$\exp(x)$
Logit fonksiyon	$\ln(x / (1 - x))$	$\exp(x)/(1 + \exp(x))$
Ters fonksiyon	$1/x$	$1/x$

Genelleştirilmiş doğrusal modellerde kullanılan bağ fonksiyonlarından biri olan logaritmik bağ fonksiyonu, açıklayıcı değişkenlerin bağımlı değişken üzerindeki etkisinin çarpımsal olması özelliğine sahip olması durumunda kullanılmaktadır (McCullagh ve Nelder, 1989).

4.3.1.3. GLM için olabilirlik fonksiyonu

GLM için y yanıt vektörü, birbirinden bağımsız olabilirlik fonksiyonu,

$$f_Y(y, \theta, \varphi) = \exp \left\{ \frac{y\theta - b(\theta)}{k(\varphi)} + c(y, \varphi) \right\} \quad (4.13)$$

şeklinde tanımlanır. Burada, k , b ve c fonksiyonları verilen GLM yapısına göre belirlenmektedir. φ biliniyorsa, rastgele kısım θ parametresi ile üstel dağılım ailesinden olan bir dağılıma uymaktadır. φ bilinmiyorsa, rastgele kısım iki parametrelili üstel dağılım ailesine ait bir dağılıma uyabilir veya uymayabilir.

Bu olabilirlik fonksiyonunun doğal logaritması alındığında

$$l_Y(y, \theta, \varphi) = \frac{y\theta - b(\theta)}{a(\varphi)} + c(y, \varphi) \quad (4.14)$$

elde edilir. Burada

$$E\left(\frac{\partial l}{\partial \theta}\right) = 0 \quad \text{ve} \quad E\left(\frac{\partial^2 l}{\partial \theta^2}\right) + E\left(\frac{\partial l}{\partial \theta}\right)^2$$

durumları sağlanmalıdır. Buradan

$$E\left(\frac{\partial l}{\partial \theta}\right) = E\left(\frac{y - b'(\theta)}{a(\varphi)}\right) = \frac{\mu - b'(\theta)}{a(\varphi)} = 0$$

$$\mu = b'(\theta) = E(Y) \quad (4.15)$$

ve

$$E\left(\frac{\partial^2 l}{\partial \theta^2}\right) + E\left(\frac{\partial l}{\partial \theta}\right)^2 = -\frac{b''(\theta)}{a(\varphi)} + \frac{\text{Var}(Y)}{a^2(\varphi)} = 0$$

$$b''(\theta)a(\varphi) = \text{Var}(Y) \quad (4.16)$$

sonuçları elde edilir. Genel olarak $a(\varphi)$ fonksiyonu,

$$a(\varphi) = \frac{\varphi}{w_i} = \frac{\sigma^2}{\mu} \quad i = 1, \dots, n$$

şeklindedir. Bu oran, φ dağılım parametresi ile önceden belirlenmiş ve her bir gözleme göre farklı w_i ağırlıklarına göre belirlenmiş bir sabittir. Burada a fonksiyonu bir ölçek fonksiyondur.

4.3.1.4. GLM için artıkların incelenmesi

Gözlem sayısı n olan bir veri seti verildiğinde bunun n parametrelili bir model ile modellenmesi beklenmektedir. Uygun olarak seçilen model için uyum iyiliği testi gözlemlerle tam olarak eşleşen ve uygun model arasındaki farkın artık değerleri incelenerek yapılmaktadır. Bu artık değerler, Sapma artıkları, Pearson artıkları ve Anscombe artıkları olarak adlandırılmakta ve

$$\text{- Sapma artıkları} \quad r_D = \text{sign}(y - \hat{m})\sqrt{d} \quad (4.17)$$

$$\text{- Pearson artıkları} \quad r_P = (y - \hat{m})\sqrt{\frac{w}{V(\hat{m})}} \quad (4.18)$$

$$\text{- Anscombe artıkları} \quad r_a = \frac{\ln(y) - \ln(\hat{m})}{\sqrt{\hat{m}}} \quad (4.19)$$

şeklinde elde edilmektedir. Modelin uygun olup olmadığının incelenebilmesi için artık yapılarının belli özelliklere sahip olması gerekmektedir. Sapma artıkları incelendiğinde düzenli bir dağılım göstermemesi beklenmektedir (Butt ve Haberman, 2004).

Uygulamalarda yukarıda tanımlanan artık türlerinden en çok Pearson artıkları incelenmektedir. Pearson artıklarının $(-3, +3)$ arasında olması beklenmektedir.

Sapma artıkları ile Pearson artıkları arasında belirgin bir fark bulunmamaktadır. Anscombe artıklarının ise bir doğru etrafında dağılması beklenmektedir.

4.3.2. Taylor açılımı ile GLM

Hassasiyet modellerinde tanımlanan kaba ölüm hızlarının parametrelerinden bazıları doğrusal değildir. Bu yüzden, o yaş veya yıl için tanımlanmış kaba ölüm hızları ile o yaş veya yıl için tanımlanmış riske maruz kalan birim sayısı çarpıldığında o yıl veya yaş için ortalama ölüm sayısı elde edilmektedir. Ölüm sayılarının her bir yıl veya yaş için Poisson dağılımı'na sahip olduğunu varsayıldığında,

$$\bar{D}_x \sim \text{Poisson}(\bar{E}_x \bar{\mu}_x) \Rightarrow m_x = E[\bar{D}_x] = \bar{E}_x \bar{\mu}_x \quad (4.20)$$

şeklinde tanımlanmaktadır. Burada ölüm sayıları $\bar{E}_x \bar{\mu}_x$ ortalama ile Poisson dağılımına uyduğu varsayımında bulunmaktadır. $\bar{E}_x \bar{\mu}_x$ ortalamaya sahip Poisson Genelleştirilmiş Doğrusal Modelleme'de log olabilirlik fonksiyonu,

$$\text{LnL}(\beta; y) = y \ln(\bar{E}_x \bar{\mu}_x) - (\bar{E}_x \bar{\mu}_x) - \ln y!$$

şeklindedir. Burada, genelleştirilmiş doğrusal modeller fonksiyonları,

$$\theta = \ln(\bar{E}_x \bar{\mu}_x) \quad (4.21)$$

$$b(\theta) = \exp(\theta) \quad (4.22)$$

$$a(\phi) = 1 \quad (4.23)$$

$$c(y, \phi) = -\ln(y!) \quad (4.24)$$

şeklinde elde edilir.

Hassasiyet modelleri ile tanımlanmış ölüm hızı modelinin doğrusal olmaması nedeniyle GLM yönteminin uygulanabilmesi için bu modelin doğrusallaştırılması gerekmektedir. Renshaw (1991) tarafından önerilen yöntem uygulandığında ölüm hızlarının logaritmik dönüşümünü yapılarak doğrusal olmayan kısmı için Taylor açılımı kullanıldığında Eş (4.1),

$$\begin{aligned} \ln(\bar{E}_x \bar{\mu}_x) &= \ln \left(\bar{E}_x \frac{\exp(-g + px)}{(1 + \exp(-s + px))^\psi} \right) \\ &= \ln(\bar{E}_x) - g + px - \psi \ln(1 + \exp(-s + px)) \end{aligned}$$

$$= \ln(\bar{E}_x) - g + px - \psi(c_x + Sk_x + \gamma v_x) \quad (4.25)$$

şeklinde elde edilmektedir. Burada,

$$c_x = \ln(1 + \exp(-s + px)) \quad , k_x = \exp(-s + px)(1 + \exp(-s + px)) \quad , v_x = xk_x$$

$$S = s - s_0 \quad \text{ve} \quad \gamma = p - p_0$$

şeklinde tanımlanmıştır. Bu model, MATLAB programlama dili kullanılarak yazılmıştır. MATLAB programının algoritması ise şu şekildedir:

- 1. Adım:** ψ değerleri için eşit olasılıklarla rastgele değerler elde edilir.
- 2. Adım:** ψ değerleri elde edildikten sonra p_0 ve s_0 başlangıç değerleri kullanılmıştır. Bu başlangıç değerleri regresyon analizi ile Gompertz Ölümlülük Yasası'na uyum yapılarak elde edilir.
- 3. Adım:** Üçüncü adımda p_0 ve s_0 başlangıç değerleri kullanılarak yaşa göre c_x , k_x ve v_x değerleri elde edilir.
- 4. Adım:** Ölüm sayılarının Poisson dağılıma uyduğu varsayımı ile genelleştirilmiş doğrusal modelleme yönteminde parametre tahminlerinin elde edilebilmesi için en küçük kareler yöntemi kullanılarak g , s ve p parametrelerinin tahmin değerleri elde edilir.
- 5. Adım:** Analizin durdurulacağı noktanın belirlenebilmesi için kısıt olarak Pearson'un Ki-Kare Test İstatistiği kullanılmıştır. 0,99 güven düzeyi ve 34 serbestlik derecesi için Ki-Kare Tablo Değeri **50,892180** olarak alınmıştır. Modelin anlamlılığının sağlanabilmesi için ise bu değer üst sınır olarak alınmıştır.
- 6. Adım:** Oluşturulan model 5'inci Adım'daki kısıtı sağlıyorsa ekler dizininde(Ek 1, Ek 2, Ek 3, Ek 4, Ek 5) verilen dönüşümlerin tersi uygulanarak Üçüncü Bölüm'de tanıtılan hassasiyet modellerinin parametre değerleri elde edilerek 7'inci Adım'a geçilebilir. Oluşturulan modelin 5'inci adımdaki kısıtı sağlamaması durumunda ise sıfır değerine oldukça yakın s ve p değerleri elde edilerek Üçüncü Adım'a geri dönüş yapılmıştır.

7. Adım : Uyum iyiliği testleri ve duyarlılık analizinin yapılabilmesi için tahmin edilen parametreler kullanılarak yaşa göre kaba ölüm hızları ve modele göre ölüm hızları elde edilir.

MATLAB programlama dilinde yazılan program kullanılarak ele edilen kaba ve beklenen ölüm hızları kullanılarak uyum iyiliği testlerinin yapılabilmesi ve grafiklerin çizilebilmesi için Microsoft Excel, SPSS ve R programı kullanılmıştır.

4.4. Uyum İyiliği Testleri

Modele göre ölüm hızları ve kaba ölüm hızları arasındaki uyumun test edilmesi gerekmektedir. Bu uyumun test edilebilmesi için hipotez testi,

H_0 : Model uygundur.

H_s : Model uygun değildir.

biçiminde kurulur. İstatistiksel testler kullanılarak α yanılma düzeyinde test istatistiğinin değeri tablo değerleriyle karşılaştırılarak yokluk hipotezinin kabul edilebilir olup olmadığı test edilir. Bu amaçla kullanılan testler,

1. Wilcoxon İşaret Sıralama Testi
2. Serisel Korelasyon Testi
3. Belirtme Katsayısı
4. Pearson'un χ^2 test istatistiği

şeklindedir.

Wilcoxon İşaret Sıralama Testi, kaba ölüm hızları ile modelleme ile elde edilmiş ölüm hızlarının ortanca değerleri arasında fark olup olmadığının belirlenmesine ilişkin yapılan testtir (Siegel, 1956). Bu testin yapılabilmesi için Standart Normal Dağılım Tablosu'ndan α yanılma düzeyine karşılık gelen tablo değeri kullanılmaktadır. Z hesap değerinin mutlak değeri , Z tablo değerinden küçük ise yokluk hipotezinin α yanılma düzeyinde kabul edilebilir olduğu sonucuna ulaşılır.

Serisel Korelasyon Testi, artıklar için otokorelasyon sorunun olup olmadığının test edilmesi için kullanılmaktadır (Durbin ve Watson,1951). Sıra korelasyon test değeri sıfıra yakın bir değer aldığında artıklar arasında otokorelasyon problemi olmadığı sonucu elde edilir.

Belirtme katsayısı, yaşam analizinde uygun modelin kabul edilebilirliğini gösterebilecek bir kriter olarak kabul edilmektedir. Bu çalışmada iki farklı belirtme katsayısı hesabı dikkate alınmıştır. Bunlardan ilki regresyon analizinde kullanılan belirtme katsayısıdır (Anscombe, 1973). İkinci yöntemde ise belirtme katsayısı, ortalama ölüm oranı ve modele göre elde edilen ölüm oranlarına göre artık kareleri yaşam fonksiyonu ile ağırlıklandırılarak elde edilmektedir (Su ve Sherrish, 2011). Her iki yöneme göre hesaplanan belirtme katsayılarının % 100'e yakın bir değer olması istenmektedir.

Pearson'un Ki-Kare Test İstatistiği ise kadın ve erkek sigortalıların ölüm hızları için oluşturulan modelin doğru olup olmadığının belirlenebilmesi için kullanılmaktadır (Plackett, 1983). Bu istatistik değeri, her bir gözlem için Pearson artıklarının karelerinin toplamı şeklinde elde edilmektedir. Yokluk hipotezinin kabul edilebilir olması için hesaplanan bu istatistik değerinin α yanılma düzeyinde $n - k - 1$ serbestlik derecesindeki Ki-Kare Tablo Değeri'nden küçük olması gerekmektedir. Burada; n , gözlem sayısını ve k , bağımsız değişken sayısını göstermektedir.

4.5. Kadın ve Erkek Sigortalı Veri Kümesinin İncelenmesi

Genelleştirilmiş Doğrusal Modelleme Yöntemi kullanılarak kadın ve erkek sigortalıların ölüm hızları için uygun hassasiyet modeli araştırması yapılmıştır. Bu kesimde kadın ve erkek sigortalılar için ölümlülük modellemesinde yapılan işlemler ayrıntılı olarak ele alınacaktır.

4.5.1. Kadın sigortalı veri kümesinin incelenmesi

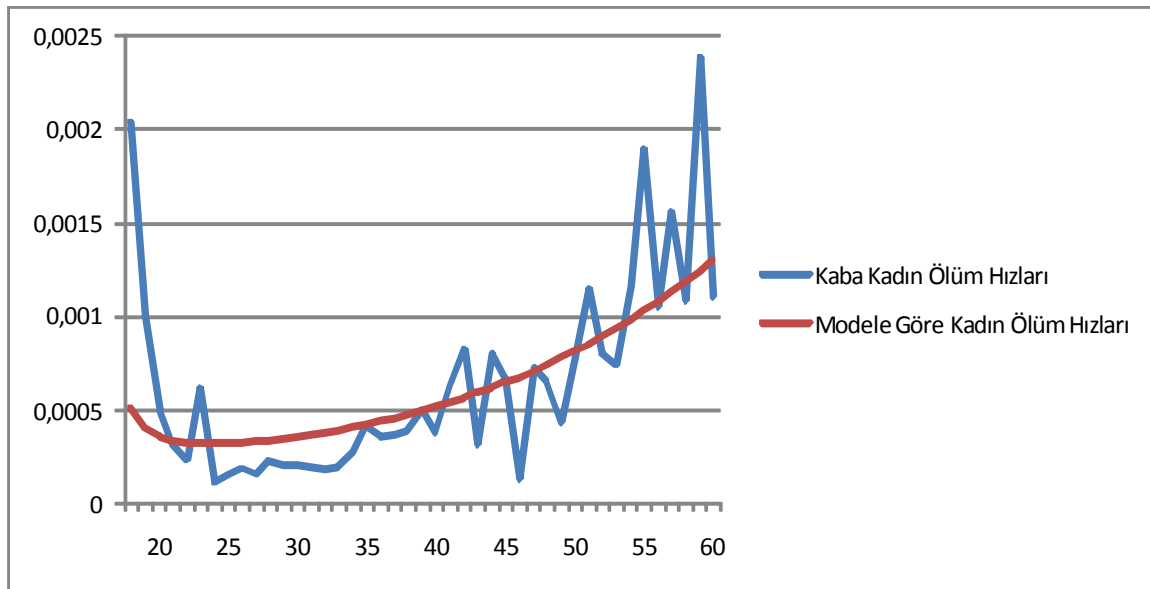
Kadın sigortalıların ölüm hızları incelendiğinde ölüm hızlarının Gompertz/Ters Gauss Hassasiyet Modeli'ne uyum sağladığı sonucuna ulaşılmış, diğer bir deyişle, Eş. (4.1)'de tanımlanan modele göre $\psi = 0,5$ olarak bulunmuştur.

Üçüncü Bölüm'de hassasiyet dağılımının Ters Gauss dağılımına uyması durumunda Eş. (3.36)'da tanımlanan modele göre $\sigma^2 = 833,335$ olarak elde edilmiştir. Dayanak ölüm hızı fonksiyonunun Gompertz Ölümlülük Yasası'na uyduğu varsayılmış model parametreleri $b=0,00051$ ve $p=0,097$ olarak tahmin edilmiştir. Bu durumda kadın sigortalıların ölüm hızları için model,

$$\bar{\mu}_x = \frac{0,00051 \exp(0,097x')}{\left(1 + 1666,67 \left(\frac{0,00051}{0,097} (\exp(0,097x') - 1) \right)\right)^{0,5}} \quad (4.26)$$

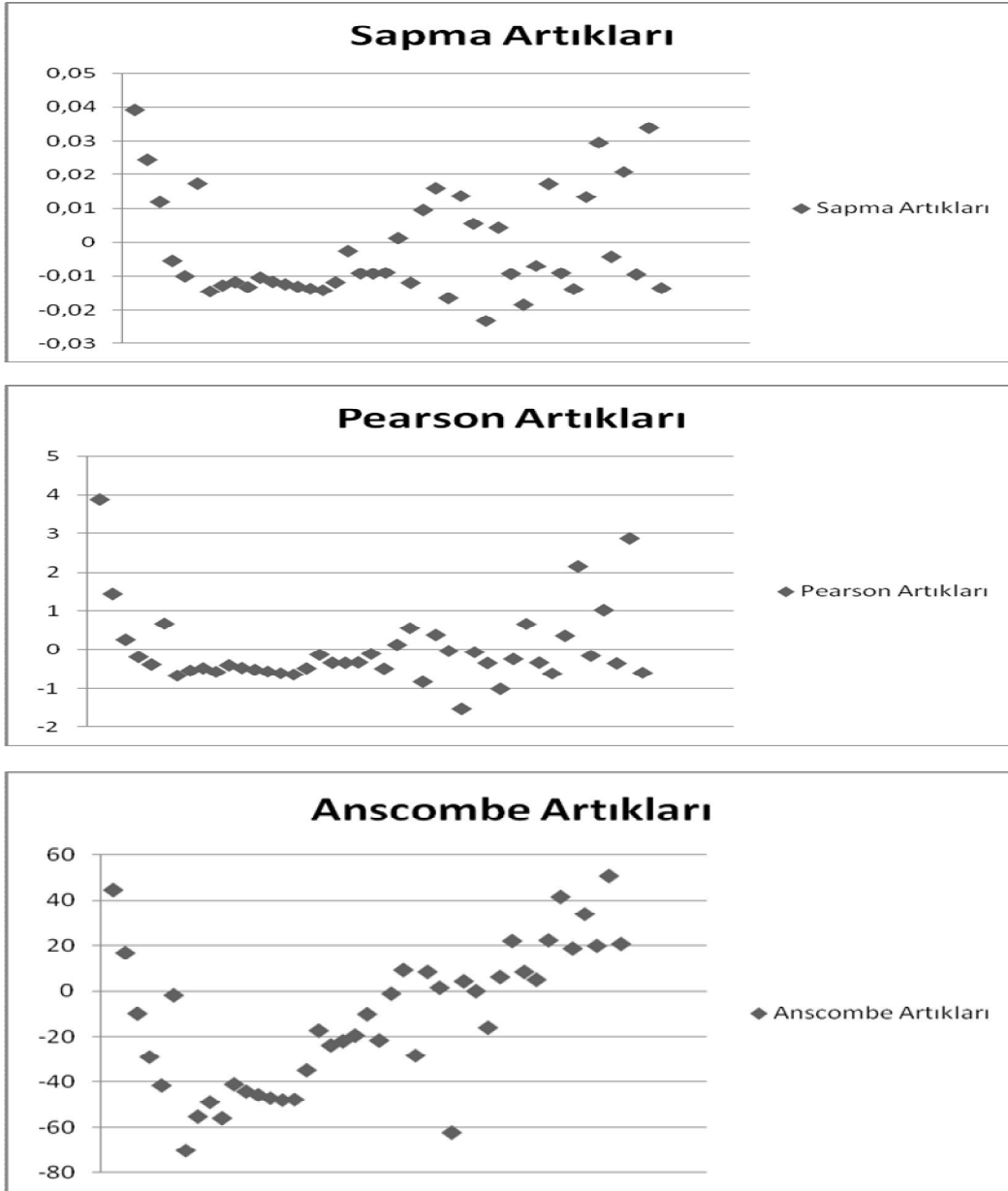
şeklinde elde edilmiştir. Burada, $x' = x - 18$, $18 \leq x \leq 60$ olarak tanımlanmıştır.

Kadın sigortalılar için kaba ölüm hızları ile modele göre ölüm hızları, Şekil 4.4'te gösterilmiştir:



Şekil 4.4. Kadın Sigortalı Bireyler İçin Yaşa Göre Gözlenen ve Modele Göre Ölüm Hızları

Kadın sigortalılar için Sapma artıkları, Pearson artıkları ve Anscombe artıkları ise Şekil 4.5'te gösterilmiştir:



Şekil 4.5. Kadın Sigortalı Ölüm Hızları İçin Artıkların İncelenmesi

Şekil 4.5 incelendiğinde Sapma artıklarının rastgele bir dağılım gösterdiği görülmektedir. Pearson artıklarının ise sadece (18) yaşındaki artık değeri hariç diğer artıklar $(-3,+3)$ aralığında bulunduğu ve Anscombe artıklarının ise (18)-(25) yaş aralığı haricinde pozitif eğimli bir doğrunun etrafında dağıldığı görülmektedir.

Kadın sigortalıların ölüm hızları için elde edilen Anscombe artıklarının bir doğru etrafında dağılmaması nedeniyle (18)-(25) yaş aralığındaki kaba ölüm hızları göz ardı

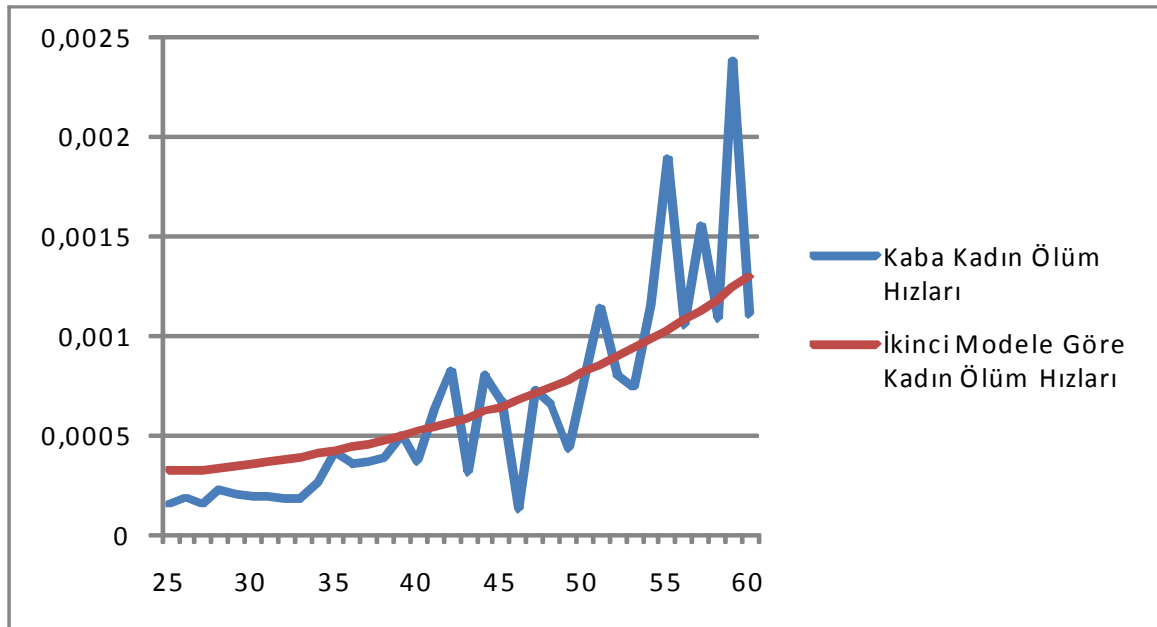
edilerek ve $\psi = 0,5$ olarak alınarak yeniden modelleme yapıldığında parametre değerlerinin değiştiği görülmektedir.

Kadın sigortalıların ölüm hızları için yeniden modelleme yapıldığında Üçüncü Bölüm'de hassasiyet dağılımının Ters Gauss dağılımına uyumu durumunda Eş. (3.36)'da tanımlanan modele göre $\sigma^2 = 1000$ olarak bulunmuştur. Dayanak ölüm hızı fonksiyonu Gompertz Ölümlülük Yasası'na uyduğu varsayıldığında model parametreleri $b=0,00051$ ve $p=0,12178$ olarak tahmin edilmiştir. Bu durumda kadın sigortalılar için model,

$$\bar{\mu}_x = \frac{0,00051 \exp(0,12178x')}{\left(1 + 1000 \left(\frac{0,00051}{0,12178} (\exp(0,12178x') - 1)\right)\right)^{0,5}} \quad (4.27)$$

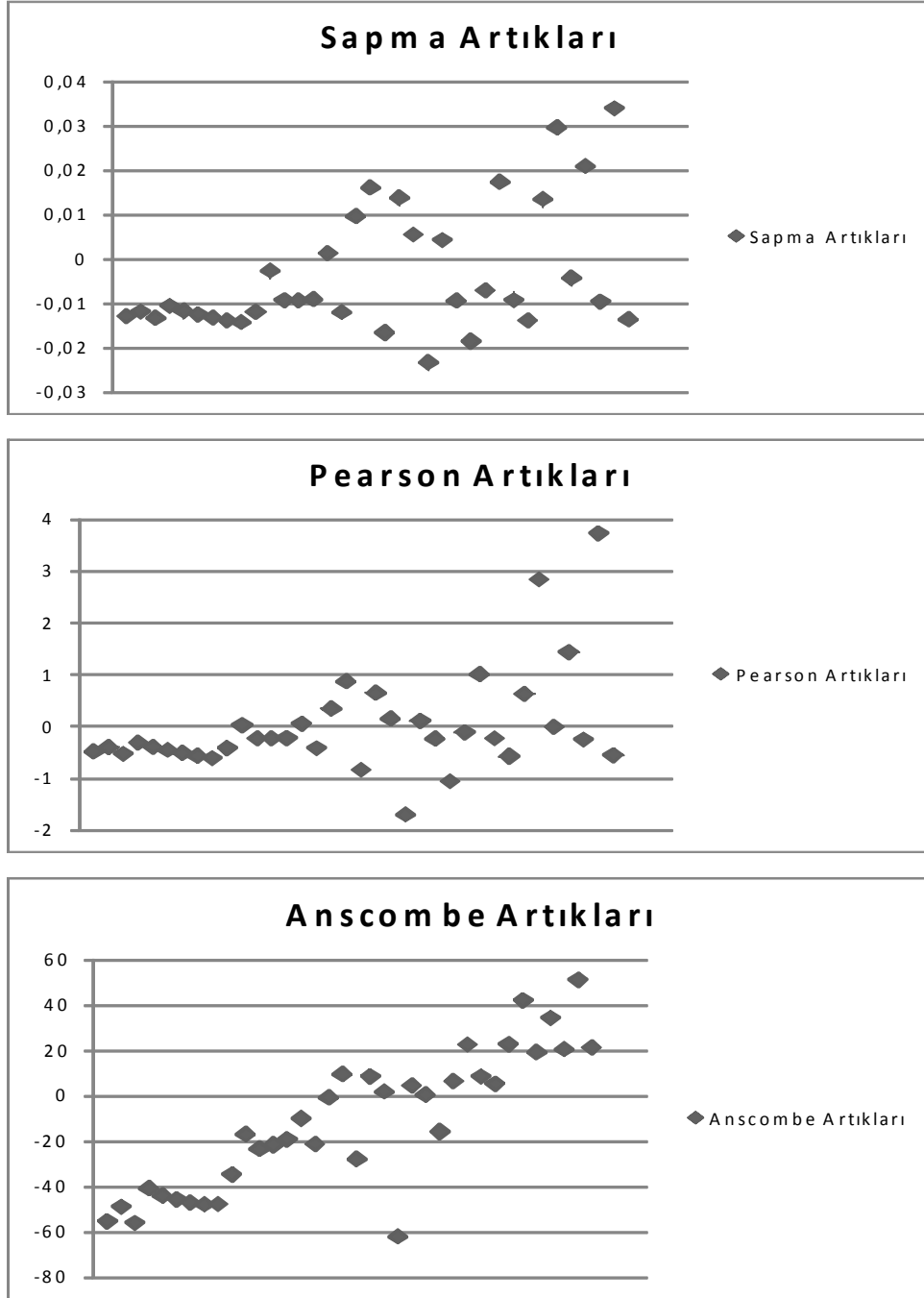
olarak elde edilmiştir. Burada, $x' = x - 25$, $25 \leq x \leq 60$ şeklinde tanımlanmıştır.

Kadın sigortalıların ölüm hızları için ikinci modelleme yapıldığında, elde edilen modele göre kadın ölüm hızları ve kaba kadın ölüm hızları Şekil 4.6'da gösterilmiştir:



Şekil 4.6. Kadın Sigortalılar İçin Yaşa Göre Gözlenen ve İkinci Modele Göre Beklenen Ölüm Hızları

Kadın sigortalıların ölüm hızları için ikinci modelleme sonucunda artıklar yeniden incelendiğinde elde edilen sonuçlar Şekil 4.7’de gösterilmiştir:



Şekil 4.7. İkinci Modellemede Kadın Sigortalı Ölüm Hızları İçin Artıkların İncelenmesi

Şekil 4.7 incelendiğinde Sapma artıklarının rastgele dağıldığı görülmektedir. Sapma artıkları arasında otokorelasyon sorununun olup olmadığının belirlenebilmesi için Serisel Korelasyon testi yapılmıştır. Serisel korelasyon testi yapıldığında -0,0364

değeri elde edilmiştir. Bu değerın sıfıra yakın bir değeri olması nedeniyle artıklar arasında otokorelasyon sorununun olmadığı görülmüştür.

Pearson artıkları incelendiğinde ise (59) yaşı hariç aykırı değeri görülmemektedir. Anscombe artıklarının da pozitif eğimli bir doğru etrafında dağıldığı görülmektedir. Pearson artıklarındaki aykırı değeri veri kümesinden çıkarılıp yeniden modelleme yapıldığında ise tahmin edilen model parametrelerinin değişmediği görülmüştür.

Anscombe artıklarında (44) yaşındaki artık değeri hariç diğer artıkların pozitif eğimli bir doğru etrafında dağıldığı görülmektedir. (44) yaşındaki aykırı değeri veri kümesinden çıkarılıp yeniden modelleme yapıldığında da model parametrelerinin değişmediği görülmüştür.

Kadın sigortalıların ölüm hızları için uyum iyiliği testlerinden Wilcoxon İşaret Sıralama Testi yapıldığında -1,1469 değeri elde edilmiştir. 0,01 yanılma düzeyinde standart normal dağılım tablo değeri olan 2,575 ile karşılaştırılmış ve $1,1469 < 2,575$ olduğu görülmüştür. Bu durumda yokluk hipotezinin, yani modelin uygunluğunun kabul edilebileceği sonucuna varılmıştır.

Kadın sigortalıların ölüm hızları için belirtme katsayısına bakıldığında ise; birinci yonteme göre hesaplanan belirtme katsayısı değeri 0,65 ve ikinci yonteme göre hesaplanan belirtme katsayısı değeri 0,87 olarak elde edilmiştir. Kadın sigortalıların ölüm hızları için belirtme katsayısının düşük olmasının nedeni kadın sigortalıların ölüm hızlarındaki aşırı dalgalanma durumudur.

Kadın sigortalıların ölüm hızları için Pearson'un Ki-Kare Test İstatistik değeri ise 35 olarak elde edilmiştir. Hesap değeri, $35 < 50,892180$ olduğundan modelin uygunluğunu gösteren yokluk hipotezinin %99 güven düzeyinde kabul edilebilir olduğu sonucuna ulaşılmıştır.

Kadın sigortalıların ölüm hızları için artık değeri incelendiğinde ve uyum iyiliği testleri sonuçlarına bakıldığında kadın sigortalı bireylerin ölüm hızları için Eş. (4.27)'de tanımlanan modelin uygunluğu kabul edilmiştir.

4.5.2. Erkek sigortalı veri kümesinin incelenmesi

Erkek sigortalıların ölüm hızları incelendiğinde ölüm hızlarının Gompertz/Gamma Hassasiyet Modeli'ne uyum sağladığı sonucuna ulaşılmıştır. Diğer bir deyişle Eş. (4.1)'deki modelde $\psi = 1$ olarak bulunmuştur.

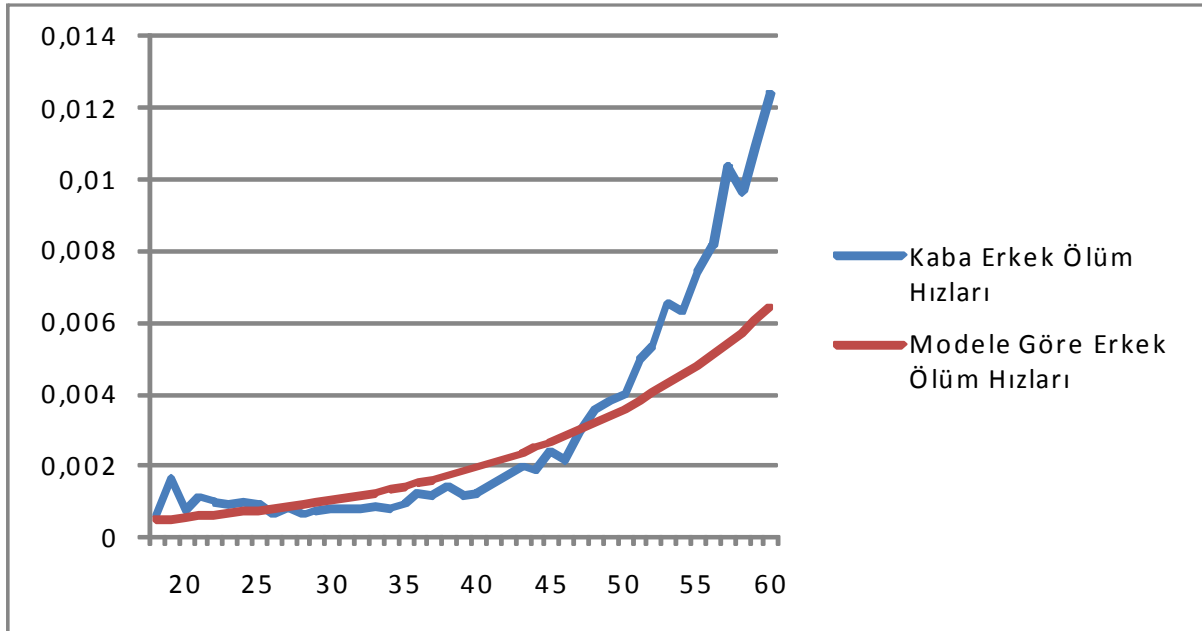
Üçüncü Bölüm'de hassasiyet dağılımının Gamma dağılımına uyması durumunda, Eş. (3.36)'da tanımlanan modele göre $\theta=1,02290$ olarak bulunmuştur. Dayanak ölüm hızı fonksiyonu Gompertz Ölümlülük Yasası'na uyduğu varsayıldığından, modelin parametreleri $b=0,000493$ ve $p=0,06343$ olarak tahmin edilmiştir.

Erkek sigortalılar için Gompertz/Gamma hassasiyet modeli,

$$\bar{\mu}_x = \frac{0,000493 \exp(0,06343x')}{1 + 1,0229 \left(\frac{0,000493}{0,06343} (\exp(0,06343x') - 1) \right)} \quad (4.28)$$

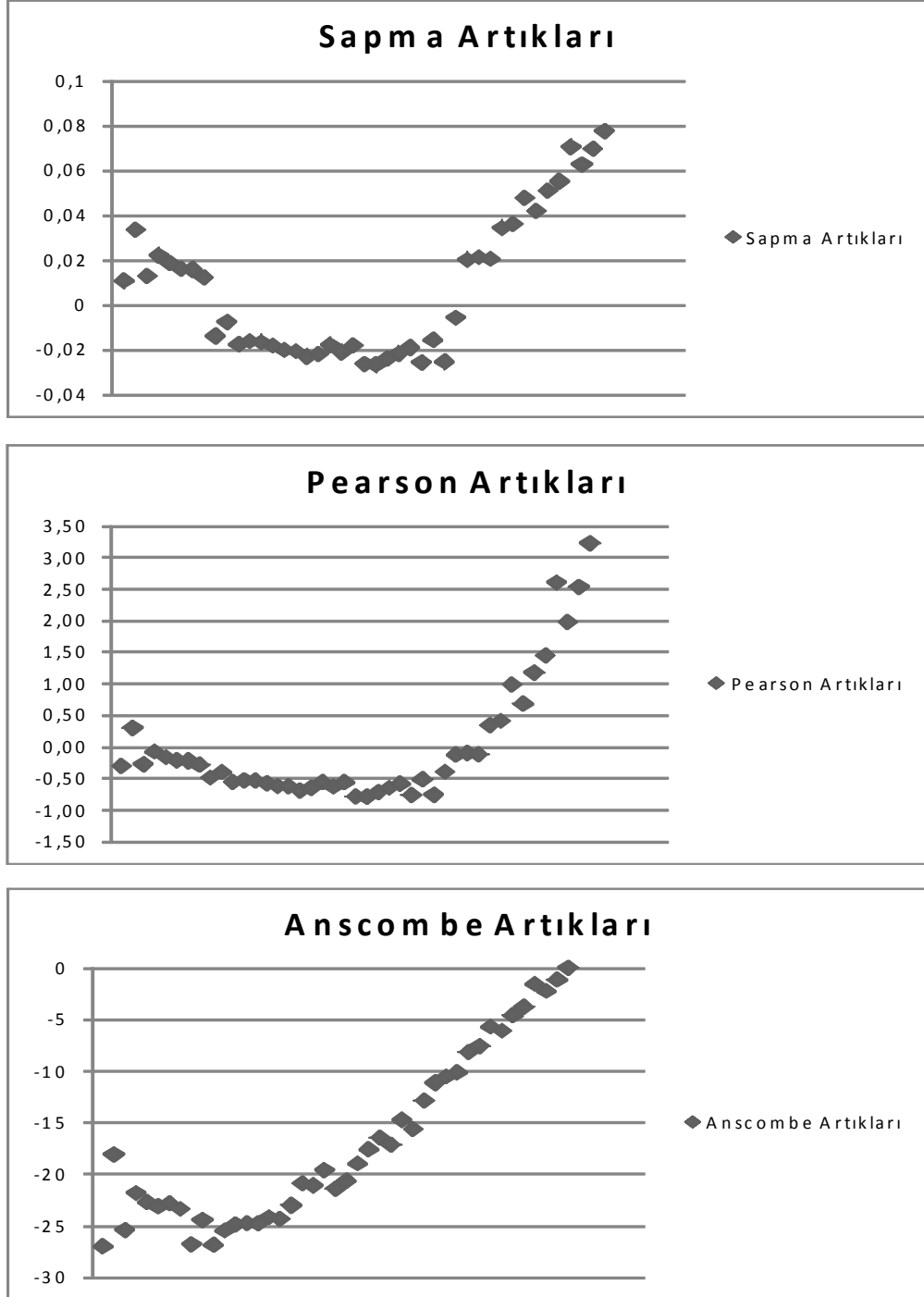
şeklinde elde edilmiştir. Burada, $x' = x - 18$, $18 \leq x \leq 60$ olarak tanımlanmıştır.

Erkek sigortalılar için kaba ölüm hızları ile modele göre ölüm hızları, Şekil 4.8'de gösterilmiştir:



Şekil 4.8. Erkek Sigortalı Bireyler İçin Yaşa Göre Gözlenen ve Beklenen Ölüm Hızları

Erkek sigortalılar için Sapma artıkları, Pearson artıkları ve Anscombe artıkları ise Şekil 4.9' da gösterilmiştir:



Şekil 4.9. Erkek Sigortalı Ölüm Hızları İçin Artıkların İncelenmesi

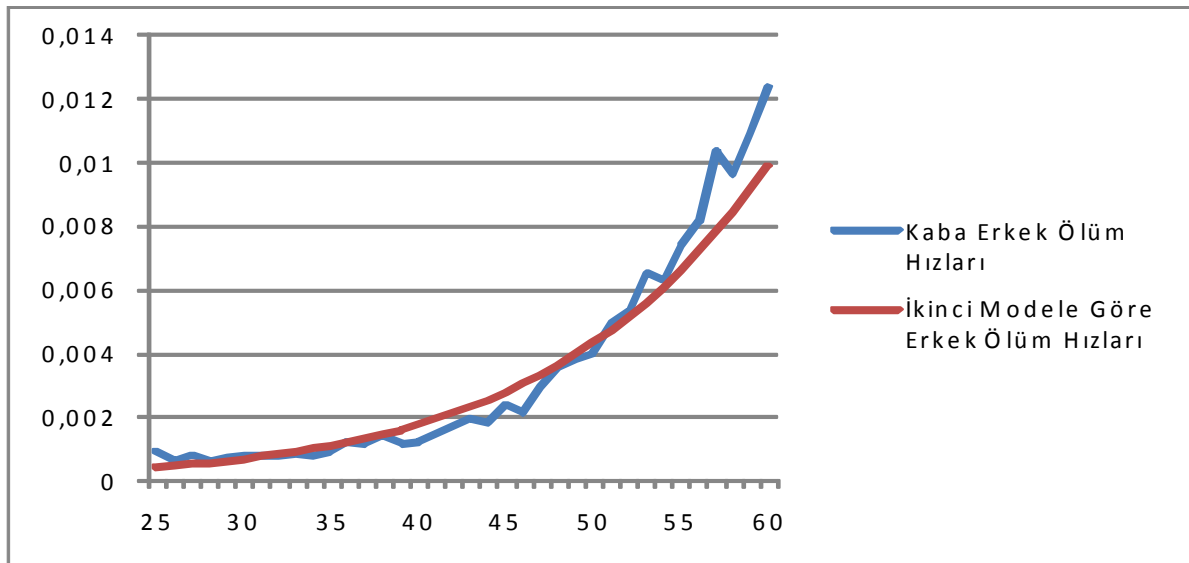
Şekil 4.9 incelendiğinde Sapma artıklarının sistematik olmayan bir dağılıma sahip olduğu görülmektedir. Pearson artıkları incelendiğinde, (-3,+3) aralığında yer almayan bir tane artık değeri bulunduğu ve Anscombe artıklarına bakıldığında ise (18)-(25) yaşları arasındaki artıklar hariç diğer artıkların pozitif eğimli bir doğru etrafında dağıldığı gözlemlenmektedir.

Erkek sigortalılar için (18)-(25) yaş aralığındaki ölüm hızı değerleri veri kümesinden çıkartılarak, yeniden Gompertz/Gamma hassasiyet modeli analizi yapılmıştır. Üçüncü Bölüm’de hassasiyet dağılımının Ters Gauss dağılımına uyması durumunda Eş. (3.36)’ da tanımlanan modele göre $\theta=1,507983$ olarak bulunmuştur. Dayanak ölüm hızı fonksiyonu Gompertz Ölümlülük Yasası’na uyduğu varsayıldığından model parametreleri $b=0,000443$ ve $p=0,09351$ olarak tahmin edilmiştir.

Erkek sigortalılar için Gompertz/Gamma hassasiyet modeli,

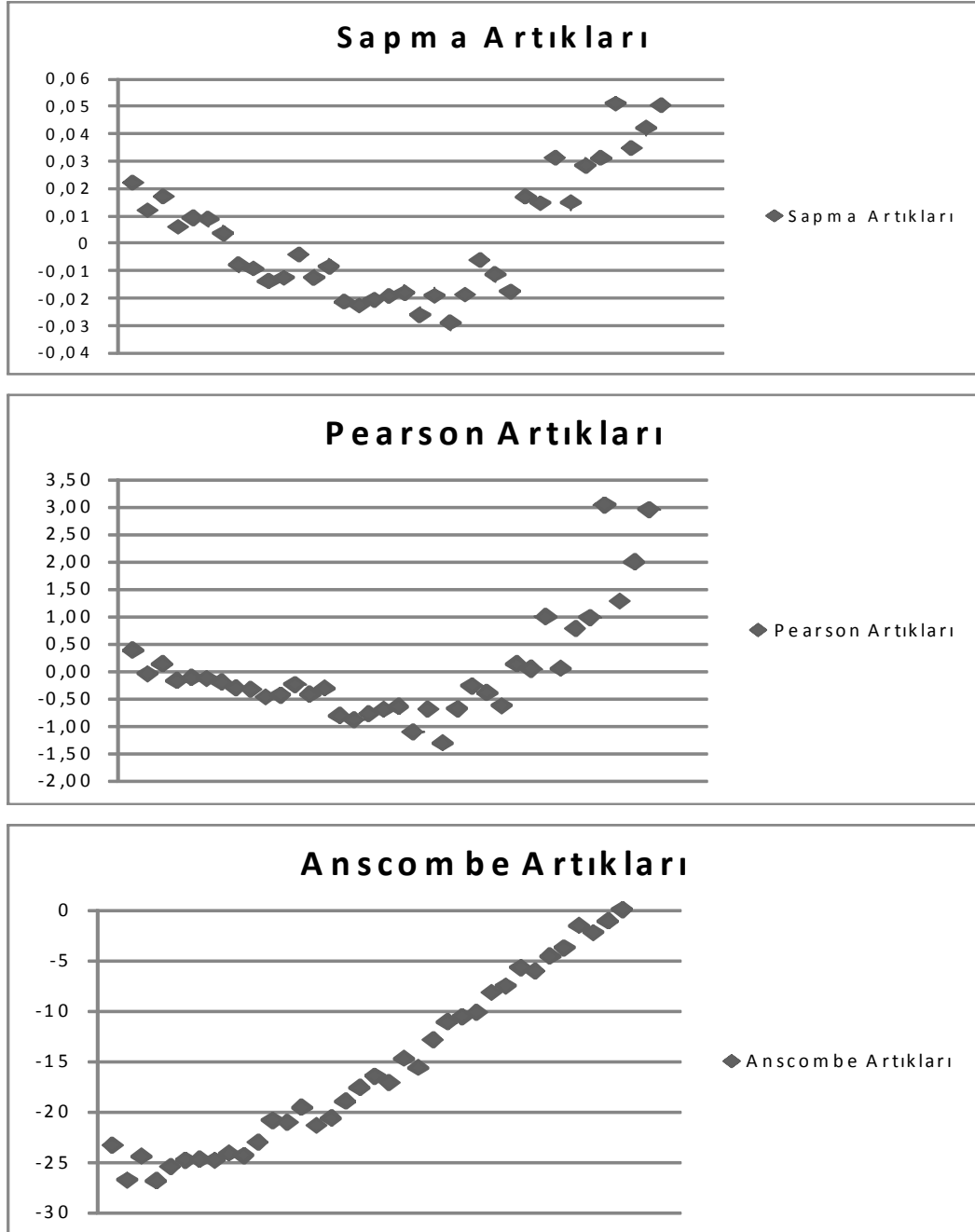
$$\bar{\mu}_x = \frac{0,000443 \exp(0,09351x')}{1 + 1,507983 \left(\frac{0,000443}{0,09351} (\exp(0,09351x') - 1) \right)} \quad (4.29)$$

şeklinde elde edilmiştir. Burada, $x' = x - 25$, $25 \leq x \leq 60$ olarak tanımlanmıştır. Şekil 4.10’da ikinci Gompertz/Gamma hassasiyet modellemesinden elde edilen erkek ölüm hızları ile kaba erkek ölüm hızları verilmiştir:



Şekil 4.10. Erkek Sigortalı Bireyler İçin İkinci Modellemede Yaşa Göre Gözlenen ve Beklenen Ölüm Hızları

Erkek sigortalıların ölüm hızları için Sapma artıkları, Pearson artıkları ve Anscombe artıkları Şekil 4.11’de gösterilmiştir:



Şekil 4.11. İkinci Modellemede Erkek Sigortalı Ölüm Hızları İçin Artıkların İncelenmesi

Şekil 4.11 incelendiğinde Sapma artıklarının sistematik bir dağılım göstermediği görülmektedir. Artıklar arasında otokorelasyon sorunun olup olmadığının belirlenebilmesi amacıyla erkek sigortalı ölüm hızlarına Serisel Korelasyon testi

yapılmıştır. Serisel korelasyon testi değeri 0,0103 olarak elde edilmiştir. Bu değer sifıra yakın olduğundan, artıklar arasında otokorelasyon sorununun olmadığı görülmüştür.

Pearson artıklarında ise biri tam (+3) değerinin üzerinde olan, tek bir artık bulunmakta olup; diğer artıklar (-3,+3) aralığında yer almaktadır. Anscombe artıklarının ise pozitif eğimli bir doğru etrafında dağıldığı görülmektedir. Pearson artıklarında görülen aykırı değer veri setinden çıkarıldığında, model parametre tahminlerinin değişmediği görülmüştür.

Erkek sigortalıların ölüm hızları için uyum iyiliği testlerinde Wilcoxon İşaret Sıralama Testi yapıldığında 0,0943 değeri elde edilmiştir. 0,01 yanılma düzeyinde standart normal dağılım tablo değeri 2,575 olarak elde edilmiş ve $0,0943 < 2,575$ olduğu görülmüştür. Bu durumda modelin uygunluğunu gösteren yokluk hipotezinin kabul edilebilir olduğu sonucuna varılmıştır.

Erkek sigortalıların ölüm hızları için yapılan ikinci modellemede, birinci yöneme ve ikinci yöneme göre hesaplanan belirtme katsayısı 0,99 olarak hesaplanmıştır. Pearson'un Ki-Kare Test İstatistik değeri ise 35 olarak elde edilmiştir.

Erkek sigortalı bireylerin ölüm hızları için Pearson artıklarının toplamının 35 olduğu ve bu hesap değeri tablo değerinden küçük olduğu için ($35 < 50,892180$) yokluk hipotezinin 0,01 yanılma düzeyinde kabul edilebilir olduğu görülmüştür.

Erkek sigortalıların ölüm hızları için ikinci modelleme denemesinden sonra artık yapıları, Wilcoxon İşaret Testi, Serisel Korelasyon Testi, Pearson'un Ki-Kare Test İstatistiği ve Belirtme Katsayısı değerleri dikkate alındığında Eş. (4.34)'de tanımlanan modelin uygun olduğu sonucuna ulaşılmıştır.

Ölümlülük modellemesinde bir modelin uygun olup olmadığının belirlenebilmesi için aşağıda verilen iki önemli kriterin sağlanması gerekmektedir:

- Modelin parametre tahminlerinin aykırı değerlerden etkilenmemesi
- Hata varyanslarının homojen olmamasına izin veren bir model olması, diğer bir deyişle hata terimi varyansının yaşlara göre değişimine izin vermesi (Butt ve Haberman, 2002).

Kadın ve erkek sigortalıların ölüm hızları incelendiğinde birinci kriterin sağlandığı görülmüştür. İkinci kriterin sağlanıp sağlanmadığının incelenebilmesi için ise Goldfeld Quandt Testi yapılabilir. Bu test büyük örneklerle uygulanabilen bir testtir ve kıyaslama yapılabilmesi için F tablosu kullanılmaktadır (Goldfeld ve Quandt, 1965).

Kadın sigortalıların ölüm hızları için hesap değeri 31,32 ve erkek sigortalıların ölüm hızları için hesap değeri 35,64 olarak elde edilmiştir. Pay serbestlik derecesi 11 ve payda serbestlik derecesinin 11 olması durumunda 0,01 yanılma düzeyinde F istatistik değeri 4,47 olduğu görülmüştür. Kadın ve erkek sigortalıların ölüm hızları için elde edilen hesap değerleri 4,47'den büyük olması nedeniyle hata terimleri varyanslarının aynı olmadığı sonucuna ulaşılmıştır.

Kadın ve erkek sigortalılar için hassasiyet etkisinin bağımlılık derecesinin ölçülebilmesi için Kendall'ın τ değerinden yararlanılmaktadır. Eş. (3.8) kullanılarak kadın sigortalıların ölüm hızları için Kendall'ın τ değeri 499 ve erkek sigortalıların ölüm hızları için ise 0,33838 olarak elde edilmiştir. Bu durumda gözlemlenemeyen risk faktörlerinin kadın sigortalıların ölüm hızlarında oluşturduğu bağımlılığın erkek sigortalıların ölüm hızlarında oluşturduğu bağımlılık derecesinden daha büyük olduğu sonucuna varılmaktadır.

Kadın ve erkek sigortalıların ölüm hızlarındaki bağımlılığın ölçümü ile ilgili kullanılabilecek bir başka ölçüt ise hassasiyet etkisi dağılımının varyansıdır. Kadın sigortalıların ölüm hızları için uygun bulunan hassasiyet etkisinin varyansı, erkek sigortalıların ölüm hızları için uygun bulunan hassasiyet etkisinin varyansından büyük olduğu için kadın sigortalıların daha heterojen bir yapıya sahip oldukları söylenebilmektedir.

4.6. Ölüm Hızlarının Uzatılması

Kadın ve erkek sigortalıların ölüm hızları için yapılan modelleme sonucunda elde edilen ölüm hızlarına İlişkisel Model (HPC), uzatma yöntemi uygulanmıştır (Himes, Preston ve Coudran, 1994). Bu yöntemin uygulanmasının nedeni, İlişkisel Model'in diğer uzatma yöntemlerine göre daha uygun sonuçlar vermesidir. Himes, Preston ve Coudran (1994) çalışmasında, (45) yaşından (99) yaşına kadar ölüm oranları düşük 82 farklı ülkenin ölümlülük yapısı dikkate alınarak standart hayat tablosu elde edilmiştir. Bu standart hayat tablosundan elde edilen regresyon denklemi kullanılarak

modellenen ölüm hızları uzatılabilmektedir. Modele göre ölüm hızlarının İlişkisel Model kullanılarak uzatılabilmesi için aşağıdaki adımlar izlenmiştir:

- 1. Adım:** Uzatılmak istenen tabloyla ilişkilendirilmek üzere standart hayat tablo seçilir.
- 2. Adım:** Standart tablonun istenilen son yaşa kadar uzatılması amacıyla, standart tablodaki Y bağımlı değişkenine logit dönüşümü kullanılarak basit doğrusal regresyon analizi yapılır.
- 3. Adım:** Uzatılmak istenen tablo için bağımsız değişken standart tablodan elde edilen Y bağımlı değişkeni ile uzatılmak istenen tablonun Y bağımlı değişken arasında basit doğrusal regresyon analizi yapılır. Kurulan regresyon denklemi,

$$\ln\left(\frac{\mu_x}{1-\mu_x}\right) = \delta + \gamma \ln\left(\frac{\mu_x^s}{1-\mu_x^s}\right) \quad (4.30)$$

şeklindedir.

- 4. Adım:** Uzatılmak istenen tabloya ilişkin regresyon katsayıları elde edilerek, istenilen yaşa kadar uzatma işlemi yapılır.

Kadın sigortalıların ölüm hızları için yapılan uzatma işlemi sonucunda (60)-(110) yaşları için elde edilen ölüm hızları Ek 6'da verilmiştir. Erkek sigortalıların ölüm hızları için yapılan uzatma işlemi sonucunda (60)-(110) yaşları için elde edilen yıllık ölüm oranları ise Ek 7'de verilmiştir.

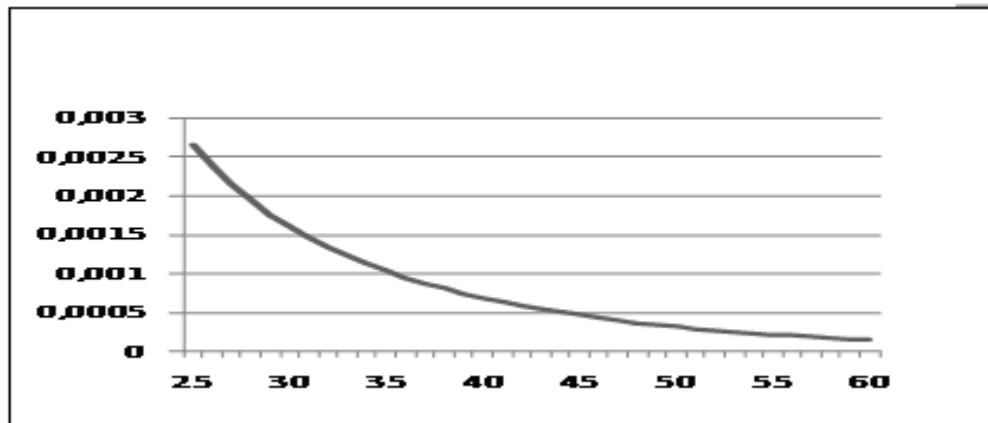
5. SONUÇ VE ÖNERİLER

5.1. Sonuçlar

Herhangi bir ölüm zamanı verisinden elde edilen ölüm oranları modellenirken, popülasyonun heterojen olmasından kaynaklanan bağımlılık durumuna ilişkin; aktüeryal, demografik ve istatistiksel literatür incelendiğinde, çok fazla sayıda çalışma yapıldığı görülmektedir. Popülasyondaki heterojenlikten kaynaklanan bağımlılık etkisinin ölçülebilmesi için kesikli ve sürekli yaş formülleri kullanılarak birçok model oluşturulmuştur.

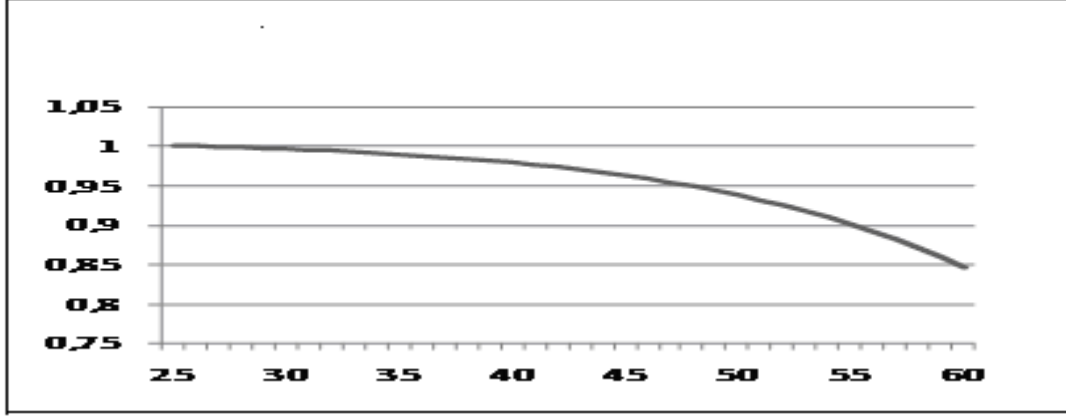
Bu çalışmada, 2004-2008 Türkiye sigortalı verisi kullanılarak dayanak ölüm hızının Gompertz Modeli'ne uyduğu varsayılmış, hassasiyet dağılımı için Gamma, Ters Gauss, Pozitif Durağan, Güç Varyans Fonksiyonu ve Bileşik Poisson dağılımlarından en uygun hassasiyet modeli araştırması yapılmıştır. Parametre tahmini yapılırken Genelleştirilmiş Doğrusal Modelleme Yöntemi kullanılmıştır. Genelleştirilmiş Doğrusal Modeller, Klasik Doğrusal Modeller'e göre daha kararlı parametre tahminleri elde edilmesi ve az sayıda aykırı değer kabul edilebilirliği açısından daha uygun bir yöntem olmaktadır (Butt ve Haberman, 2004). Kadın sigortalıların ölüm hızları için Gompertz/Ters Gauss hassasiyet modelinin ve erkek sigortalıların ölüm hızları için Gompertz/Gamma hassasiyet modelinin daha uygun olduğu sonucuna varılmıştır.

Kadın sigortalıların ölüm hızları için yaşa göre hassasiyet etkisi değerleri Şekil 5.1'de gösterilmiştir:



Şekil 5.1 Kadın Sigortalılar İçin Hassasiyet Etkisi

Şekil 5.1’de kadın sigortalılar için hassasiyet etkisinin oldukça küçük bir değer aldığı ve yaş ilerledikçe azalan hızla azaldığı görülmektedir. Burada, gözlemlenemeyen risk faktörlerinin kadın ölüm hızları üzerindeki etkisinin az olduğu sonucuna ulaşılmaktadır. Erkek sigortalıların ölüm hızları için yaşa göre hassasiyet etkisi değerleri ise Şekil 5.2’de gösterilmektedir :



Şekil 5.2 Erkek Sigortalılar İçin Hassasiyet Etkisi

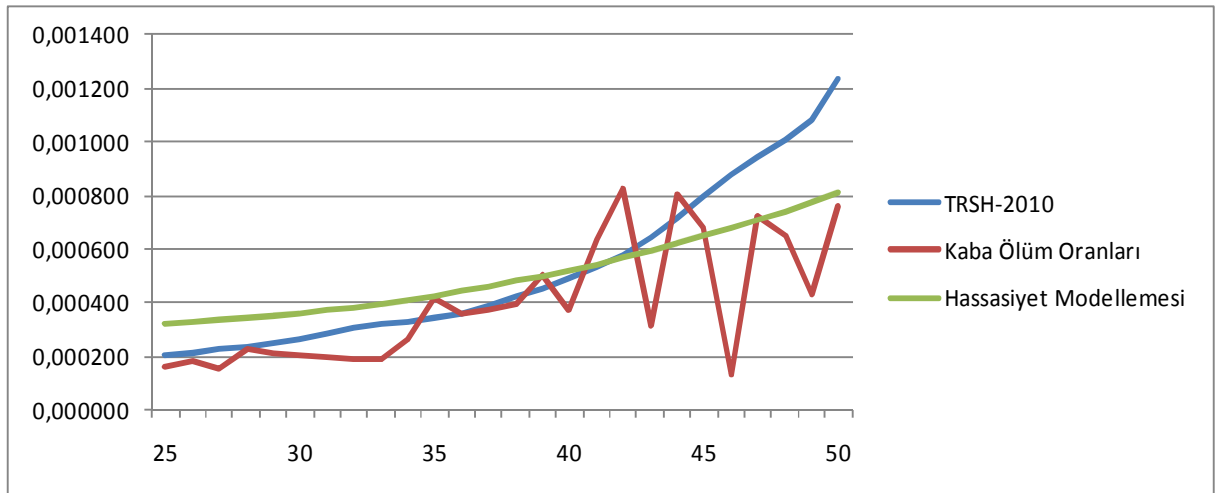
Şekil 5.2’de erkek sigortalılar için hassasiyet etkisinin (25)-(30) yaşları arasında bire yakın değer aldığı ve yaş ilerledikçe artan hızla azaldığı görülmektedir. Burada, gözlemlenemeyen risk faktörlerinin erkek ölüm hızları üzerindeki etkisinin fazla olduğu sonucuna ulaşılmaktadır.

Şekil 5.1 ve Şekil 5.2’de, hassasiyet modellerinin temel varsayımlarında olduğu gibi, yaş ilerledikçe kadın ve erkek sigortalıların ölüm hızları için hassasiyet seviyesinin azaldığı görülmektedir. Diğer bir önemli durum ise erkek sigortalı popülasyonun, kadın sigortalı popülasyona göre daha hızlı bir şekilde homojen olmasıdır.

Hassasiyet etkisinin bire yakın değer alması ilgilenilen popülasyonun homojen olduğunu göstermemektedir. Popülasyonun ne derecede heterojen yapıya sahip olduğu hassasiyet etkisinin varyansı ile belirlenebilmektedir. Dördüncü Bölüm’de kadın ve erkek sigortalılar için hassasiyet etkisinin varyansı karşılaştırıldığında, kadın sigortalıların, erkek sigortalılara göre daha heterojen yapıya sahip olduğu sonucuna ulaşılmıştır.

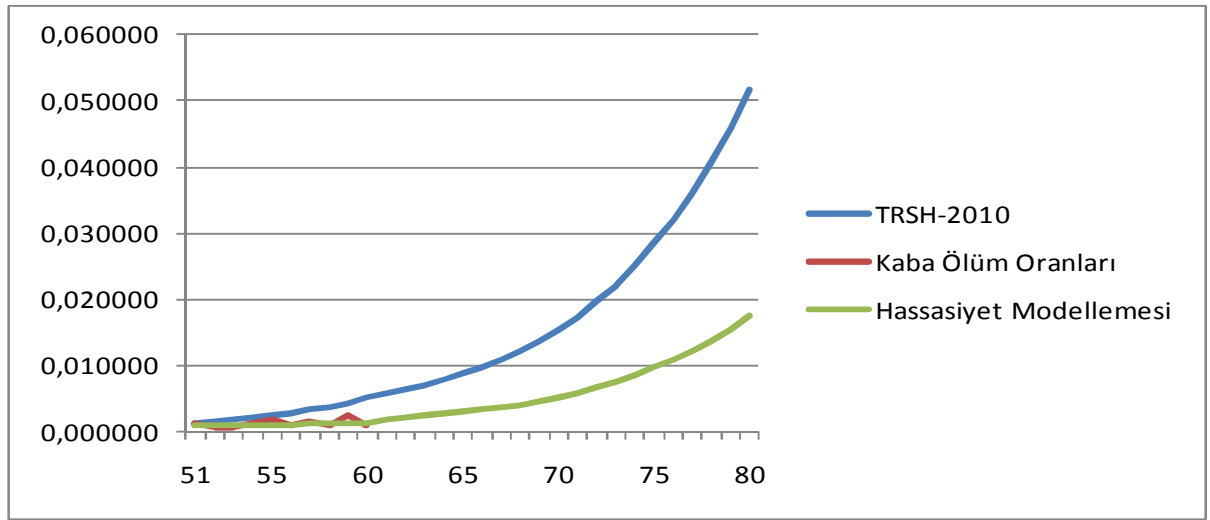
Kadın ve erkek sigortalıların hassasiyet etkisi değerleri incelendiğinde, erkek sigortalıların, kadın sigortalılara göre daha çabuk sistemden ayrılması beklenmektedir.

Kadın ve erkek sigortalılar için Hassasiyet modellemesi sonucunda elde edilen ölüm hızları ile Türkiye Kadın ve Erkek Sigortalı Hayat Tablosu'nun (TRSH-2010) karşılaştırılması amacıyla ölüm oranlarının tekdüze dağıldığı varsayımı kullanılmıştır. Türkiye Kadın ve Erkek Sigortalı Hayat Tablosu oluşturulurken, 2004-2008 yılları arası için sigortalı popülasyonun ölüm verisi kullanılarak kadın ve erkek alt popülasyonları için yaşa göre kaba ölüm oranları elde edilmiştir. Kadın sigortalı ölüm oranlarında, (18)-(60) yaş aralığı için Whittaker-Henderson yöntemi kullanılarak düzeltme işlemi yapılmış, (60)-(110) yaş aralığı için ise İlişkisel Model kullanılarak uzatma işlemi yapılmıştır. Erkek sigortalı ölüm oranlarında ise (18)-(70) yaş aralığı için Whittaker-Henderson yöntemi kullanılarak düzeltme işlemi yapılmış, (70)-(110) yaş aralığı için ise İlişkisel Model kullanılarak uzatma işlemi yapılmıştır. TRSH-2010 Kadın Hayat Tablosu Ek 8'de ve TRSH-2010 Erkek Hayat Tablosu Ek 9'da verilmiştir. Hassasiyet modellemesi ve TRSH-2010 hayat tablosunun karşılaştırılabilmesini sağlamak amacıyla kadın ve erkek sigortalılar için (25)-(110) yaş aralığı üçe bölünmüştür. Şekil 5.3'te kadın sigortalıların (25)-(50) yaş aralığı için ölüm oranları gösterilmiştir:



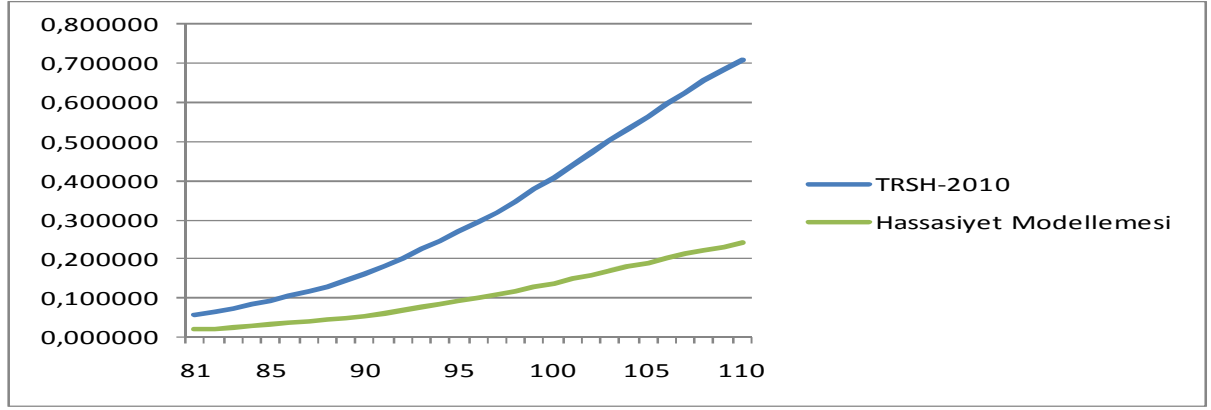
Şekil 5.3. (25)-(50) Yaş Aralığı İçin Kadın Sigortalı Ölüm Hızlarının TRSH-2010 ve Hassasiyet Modellemesine Göre Karşılaştırılması

Şekil 5.3 incelendiğinde; Whittaker-Henderson parametrik olmayan düzeltme yöntemi ile elde edilen ölüm oranlarının, Hassasiyet modellemesi parametrik düzeltme yöntemine göre (25)-(40) yaş aralığı için daha düşük ve (40)-(50) yaş aralığı için ise daha yüksek olduğu görülmektedir. Hassasiyet modellemesi ile Whittaker-Henderson yöntemi sonucu elde edilen ölüm oranlarının arasındaki farkın çok olmasının temel nedeni, kaba ölüm oranlarının fazla dalgalanma göstermesidir. Kadın sigortalıların ölüm oranlarının; Hassasiyet modellemesi ve TRSH-2010 tablosunun karşılaştırmaları (51)-(80) yaş aralığı için Şekil 5.4'te, (81)-(110) yaş aralığı için ise Şekil 5.5'te verilmiştir :



Şekil 5.4. (51)-(80) Yaş Aralığı İçin Kadın Sigortalı Ölüm Hızlarının TRSH-2010 ve Hassasiyet Modellemesine Göre Karşılaştırılması

Şekil 5.4 incelendiğinde (51)-(60) yaş aralığı için kadın kaba ölüm oranları ile Hassasiyet modellemesi uyumunun, Whittaker-Henderson yöntemine göre daha iyi olduğu görülmektedir. Bunun temel nedeni, Whittaker-Henderson düzeltme yönteminde en uygun ölüm oranları eğrisi elde edilmeye çalışılırken, sapmalara olan duyarlılığın daha fazla olması nedeniyle tahmin edilen ölüm oranlarının yüksek olmasıdır.

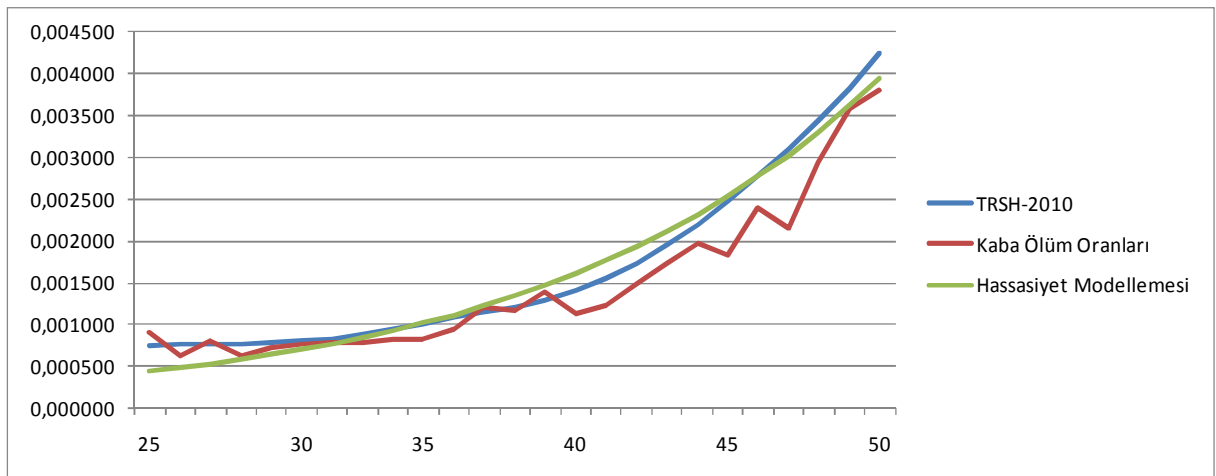


Şekil 5.5. (81)-(110) Yaş Aralığı İçin Kadın Sigortalı Ölüm Hızlarının TRSH-2010 ve Hassasiyet Modellemesine Göre Karşılaştırılması

Şekil 5.5'te Whittaker-Henderson düzeltme yöntemine göre elde edilen ölüm oranlarının Hassasiyet modellemesine göre daha yüksek olması nedeniyle, TRSH-2010 için yapılan uzatma işlemi sonucunda elde edilen ölüm oranlarının da Hassasiyet modellemesi için yapılan uzatma işlemi sonucunda elde edilen ölüm oranlarından yüksek olmaktadır.

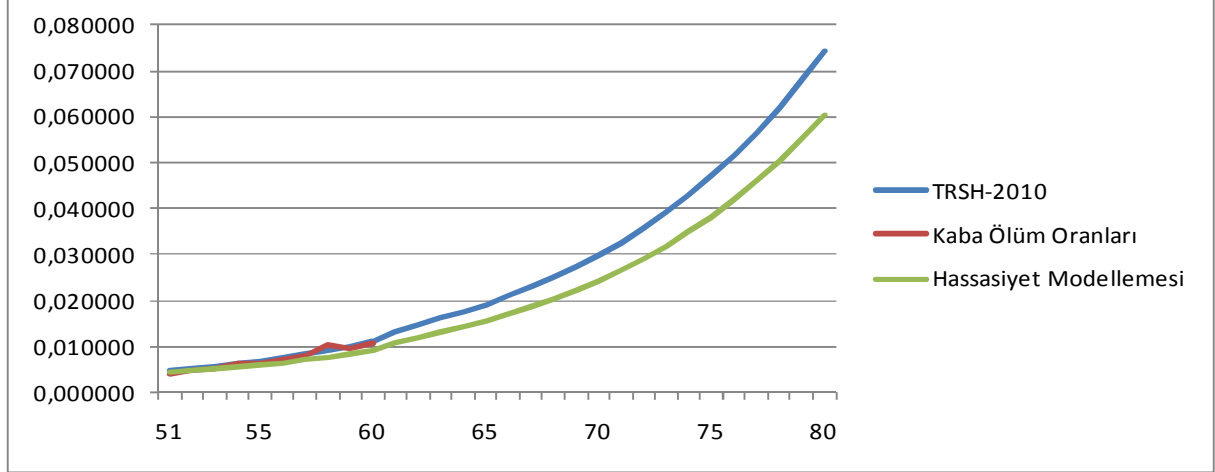
Kadın sigortalılar için kaba ölüm oranları ile modellemeler sonucunda elde edilen ölüm oranları karşılaştırıldığında parametrik düzeltmenin parametrik olmayan düzeltme yöntemine göre daha iyi sonuç verdiği görülmektedir.

Erkek sigortalıların ölüm oranlarının; Hassasiyet modellemesi ve TRSH-2010 tablosunun karşılaştırılmaları (25)-(50) yaş aralığı için Şekil 5.6'da, (51)-(80) yaş aralığı için Şekil 5.7'de, (81)-(110) yaş aralığı için ise Şekil 5.8'de verilmiştir:

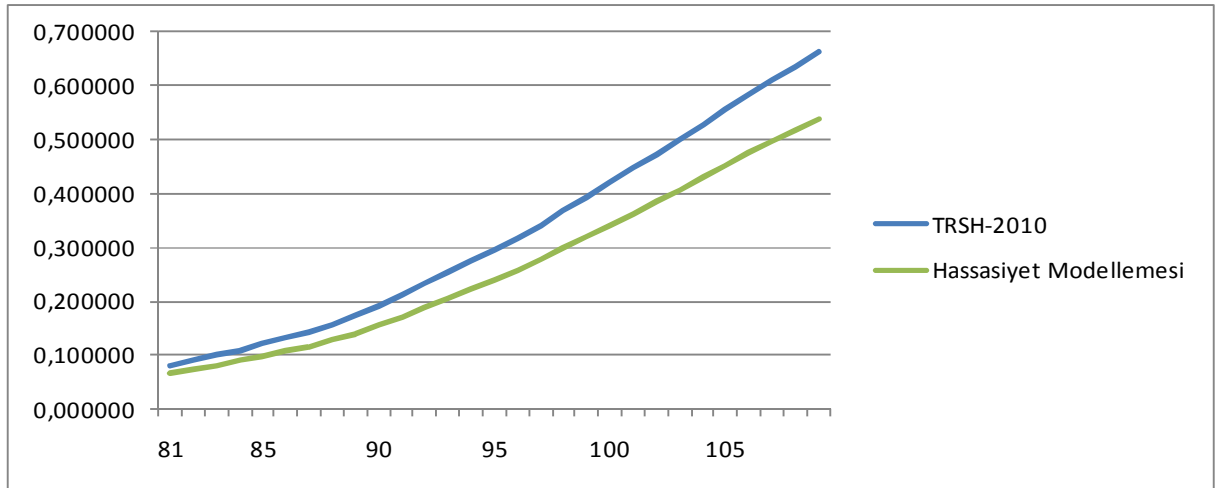


Şekil 5.6. (25)-(50) Yaş Aralığı İçin Erkek Sigortalı Ölüm Hızlarının TRSH-2010 ve Hassasiyet Modellemesine Göre Karşılaştırılması

Şekil 5.6 incelendiğinde, (25)-(45) yaş aralığı için Whittaker-Henderson parametrik olmayan düzeltme yöntemiyle elde edilen ölüm oranları, Hassasiyet modellemesi parametrik düzeltme yöntemine göre elde edilen ölüm oranlarına göre daha yüksek; (46)-(50) yaş aralığı için ise daha düşük olduğu görülmektedir. Bunun temel nedeni, ölüm oranı eğrisindeki sapmaların (45) yaşından sonra artış göstermesidir.



Şekil 5.7. (51)-(80) Yaş Aralığı İçin Erkek Sigortalı Ölüm Hızlarının TRSH-2010 ve Hassasiyet Modellemesine Göre Karşılaştırılması



Şekil 5.8. (81)-(110) Yaş Aralığı İçin Erkek Sigortalı Ölüm Hızlarının TRSH-2010 ve Hassasiyet Modellemesine Göre Karşılaştırılması

Şekil 5.7 ve Şekil 5.8 incelendiğinde ise Whittaker-Henderson düzeltme yöntemi ve Hassasiyet modellemesi sonucunda elde edilen ölüm oranları, İlişkisel Yöntem ile uzatıldığında, Whittaker-Henderson yöntemine göre elde edilen ölüm oranlarının

daha yüksek olduđu gör÷lmektedir. Bunun temel nedeni ise parametrik olmayan düzeltme ile elde edilen ölüm oranlarının parametrik düzeltme yöntemine göre daha yüksek olmasıdır.

Erkek sigortalıların yıllık ölüm oranları incelendiğinde; Whittaker-Henderson parametrik olmayan düzeltme yöntemi ve Hassasiyet modellemesi parametrik düzeltme yönteminin yaklaşık sonuçlar verdiği gör÷lmektedir. Kadın sigortalıların ölüm oranlarında olduđu gibi erkek sigortalıların ölüm oranlarında da ölümlülük eğiliminin Hassasiyet modellemesinde daha az olduđu sonucuna ulaşılmaktadır.

Bu çalışmada elde edilen sonuçlar incelendiğinde, Türkiye kadın ve erkek sigortalılardan oluşan pop÷lasyonlar için heterojenliğin etkisinin önemli olduđu gör÷lmüştür.

Kadın ve erkek sigortalıların ölüm hızları için yapılan modelleme sonucunda ölüm oranlarında sapmanın fazla olduđu durumlarda, parametrik düzeltme yönteminin; parametrik olmayan düzeltme yöntemine göre daha iyi sonuçlar verdiği gör÷lmektedir.

5.2. Öneriler

Hassasiyet modelleri, sağlık durumu ve sosyo-ekonomik durum gibi risk faktörlerinin farklı ölüm seviyeleri ile ifade edilebilmesi durumunda, hayat sigortalarının fiyatlandırılmasında kullanılabilir.

Temel hassasiyet modelleri ile modellenmiş popülasyonlar için tüm yaşam süresi boyunca, bireylerin hassasiyet seviyelerinin sabit olması; bireyler arasındaki ölümlülük eğilimi farklılığını açıklamaya yeterli olmayabilir. Yaş ve cinsiyet dışındaki diğer risk faktörlerinin de (psikolojik değişimler, sosyolojik değişimler ve zaman faktöründeki değişim gibi), hassasiyet üzerindeki etkisinin hesaba katıldığı modeller oluşturulabilir.

Hassasiyet modellerinde, her bir gözlemin sisteme girdiği anda, hassasiyet seviyesinin belirlendiği ve daha sonraki zamanda, bunun sabit kaldığı varsayımının doğru olamayacağı, insan popülasyonundaki yaşlanma süreci incelendiğinde görülmektedir. Stokastik süreçler kullanılarak, hassasiyet seviyesinin zaman içerisindeki değişiminin de incelendiği hassasiyet modelleri oluşturulabilir.

KAYNAKÇA

- Aalen O.O., 1988, Heterogeneity In Survival Analysis, *Statistics in Medicine*, Vol 7, 1121-1137.
- Aalen O.O., 1992, Modelling Heterogeneity In Survival Analysis By The Compound Poisson Distribution, *Annals of Applied Probability*, Vol. 4, 951-972.
- Aalen O.O., Tretli S., 1999, Analysis Incidence Of Testis Cancer By Means Of A Frailty Model, *Cancer Causes and Control*, Vol 10, 285-292.
- Aim S., Fulla S., Laurent J.P., 2005, Mortality Fluctuations Modelling With A Shared Frailty Approach, *Paris Actuarial Congree*, France.
- Andersen P.K., Klein J.P., Zhang M.J., 1999, Testing For Centre Effects In Multi-Centre Survival Studies: A Monte Carlo Comparison Of Fixed And Random Effects Tests, *Statistics in Medicine*, Vol 18, 1489-1500.
- Anscombe F.J., 1973, Graphs in statistical analysis, *American Statistician*, Vol. 27, 17-21.
- Atkinson D.B., Dallas J.W., 2000, *Life Insurance Products and Finance*, Society of Actuaries, U.S.A.
- Baxter L.A., Coutts S.M., Ross G.A.F., 1980, Applications Of Linear Models In Motor Insurance, in *Proct. 21st International Congress Actuaries*, Vol 2, 11-29.
- Beard R.E., 1959, Appendix: Note on some mathematical mortality models, *CIBA Foundation Colloquia on Aging*, Vol 5, 302-311.
- Beard R.E, 1971, Some Aspects Of Theories Of Mortality, Cause Of Death Analysis, Forecasting And Stochastic Processes, In *Biological Aspects Of Demography*, W. Brass, ed., Taylor & Francis, London.
- Bennett M., 1978, Models in Motor Insurance, *Journal Institute Actuarial Student Society*, Vol 22, 87-148.
- Bowers N.L, Gerber H.U., Hickman J.C., Jones D.A., Nesbitt C.J., 1997, *Actuarial Mathematics*, The Society of Actuaries, Schamburg, Illinois.
- Butt Z., Haberman S., 2002, Application Of Frailty-based Mortality Models To Insurance Data: Actuarial Research Paper No. 142, Department of Actuarial Science & Statistics , City University, London.
- Butt Z., Haberman S., 2004, Application of Frailty-based Mortality Models Using Generalized Linear Models, *ASTIN Bulletin*, Vol 34, Issue 1, 175-197.
- Chan F.Y., Chan L.Y., Yu M.H., 1984, A Generalization of Whittaker-Henderson Graduation, *Transactions of Society of Actuaries*, Vol 36.
- Chang S-C., 1998, Using Parametric Statistical Models To Estimate Mortality Structure: The Case of Taiwan, *Journal of Actuarial Practice*, Vol. 6, 197-219.

- Clayton D.G., 1978, A Model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence, *Biometrika*, Vol. 65, 141-151.
- Congdon P., 1993, Statistical Graduation In Local Demographic Analysis And Projection, *Journal of the Royal Statistical Society, Series A*, Vol. 156, 237-270.
- Congdon P., 1994, Analyzing Mortality In London: Life Tables With Frailty, *The Statistician*, Vol. 43, 277-308.
- Congdon P., 1995, Modelling Frailty In Area Mortality, *Statistics in Medicine*, Vol 14, 1859-1874.
- Coutts S.M., 1984, Motor Insurance Rating, An Actuarial Approach, *Journal Institute of Actuaries*, Vol 111, 87-148.
- Cummins J.D., Smith B.D., Vance R.N., VanDerhei J.L., 1983, Risk Classification In Life Insurance, Kluwer-Nijhoff Publishing, Boston.
- Damaskos E., 1998, Investigating Models Of Frailty Of Mortality, MSc Dissertation, Department of Actuarial Science & Statistics, City University, London.
- De Jong P., Heller G.Z., 2008, Generalized Linear Models For Insurance Data, Cambridge University Press.
- De Moivre, A. 1725, in *Annuities on Lives*, H. & S. London, 3–125.
- dos Santos D.M., Davies R.B., Francis B., 1995, Nonparametric Hazard versus Nonparametric Frailty Distribution In Modelling Recurrence Of Breast Cancer, *Journal of Statistical Planning and Inference*, Vol 47, 111-127.
- Duchateau L., Janssen P., 2008, *The Frailty Model*, Springer.
- Durbin J., Watson G.S., 1951, Testing For Serial Correlation in Least Squares Regression II, *Biometrika*, Vol. 38, Issue 1-2, 159-178.
- Elbers C., Ridder G., 1982, True and Spurious Duration Dependence: The Identifiability of the Proportional Hazard Model, *The Review of Economic Studies*, Vol. 49, No.3: 403-409.
- Fleming T., Harrington D., 1991, *Counting Processes And Survival Analysis*, Wiley, Chichester.
- Forfar D.O., McCutcheon J.J., Wilkie A.D., 1988, On Graduation By Mathematical Formula, *Journal of the Institute of Actuaries*, Vol. 115, 1-149.
- Forfar D.O., Smith D.M., 1988, The Changing Shape Of English Life Tables, *Transactions of the Faculty of Actuaries*, Vol 40, 98-134.
- Genest C., MacKay J., 1986, The Joy Of Copulas: bivariate distributions with uniform marginals, *American Statistics*, Vol 40, 280-283.

- Gerber H.U., 1997, *Life Insurance Mathematics*, Springer, Berlin.
- Goldfeld S.M., Quandt R.E., 1965, Some Tests for Homoscedasticity, *Journal of the American Statistical Association*, Vol. 60, 539-547.
- Gompertz B., 1825, On The Nature Of The Function Expressive Of The Law Of Human Mortality etc., *Phil. Trans. Roy. Soc.*, Vol 115, 513-585.
- Gompertz B., 1860, On One Uniform Law Of Mortality From Birth To Extreme Old Age And On The Law Of Sickness, presented to International statistical congress 1860 and reproduced in 1871, *Journal of the Institute of Actuaries*, Vol 16, 329-344.
- Gompertz, B., 1862, On One Uniform Law Of Mortality From Birth To Extreme Old Age And On The Law Of Sickness, *Philosophical Transactions of Royal Society (Series A)* 152, 511-559.
- Graund J., 1662, *Natural and Political Observations Made upon the Bills of Mortality*, London.
- Greenwood M., Yule G.U., 1920, An Inquiry into The Nature Of Frequency Distributions Representative Of Multiple Happenings With Particular Reference To The Occurence Of Multiple Attacks Of Disease Or Of Repeated Accidents, *Journal of the Royal Statistical Society*, Vol 83, 255-279.
- Grimes T., 1971, Claim Frequency Analysis In Motor Insurance, *Journal Institute Actuarial Student Society*, Vol 19, 147-154.
- Haberman S., Renshaw A.E., 1996, Generalized Linear Models And Actuarial Science, *The Statistician*, Vol 45, No:4, 407-436.
- Halley, E., 1693, An Estimate Of The Degrees Of The Mortality Of Mankind, drawn from curious tables of the births and funerals at the city of Breslaw; with an attempt to ascertain the price of annuities upon lives. *Phil. Trans. Roy. Soc. London* Vol 17, 596-610.
- Hartmann, M., 1987, Past and Recent Attempts to Model Mortality at All Ages, *Journal of Official Statistics*, Vol. 3, 19-36.
- Heligman L., Pollard J.H., 1980, The Age Pattern Of Mortality, *Journal of the Institute of Actuaries*, Vol 107, 49-80.
- Himes C.L., H.P. Preston, Condran G.A., 1994, A Relational Model of Mortality at Older Ages in Low Mortality Countries, *Population Studies*, Vol.48, 269-291.
- Hougaard P., 1984, Life Table Methods For Heterogeneous Populations, *Biometrika*, Vol 71, 75-83.
- Hougaard P., 1986a, Survival Models For Heterogeneous Populations Derived From Stable Distributions, *Biometrika*, Vol 73, 387-396.

- Hougaard P., 1986b, A Class Of Multivariate Failure Time Distributions, *Biometrika*, Vol 73, 671-678.
- Hougaard P., 1991, Modelling Heterogeneity In Survival Data, *Journal of Applied Probability*, Vol 28, 695-701.
- Hougaard P., Harvald B., Holm N.V., 1992, Measuring The Similarities Between The Lifetimes Of Adult Danish Twins Born 1881-1930, *Journal of the American Statistical Association*, Vol 87, 17-24.
- Hougaard P., 2000, *Analysis of Multivariate Survival Data*, Springer, New York.
- John F., 2000, *Multiple and Generalized Nonparametric Regression*, Sage University Paper, 131, s. 20
- Johnson P.D., Hey G.B., 1971, Statistical studies in motor insurance (with discussion), *Journal Institute of Actuaries*, Vol 97, 199-232.
- Jones B.L., 1988, A model for analyzing the impact of selective lapsation on mortality, *North American Actuarial Journal*, Vol 2, No.1, 79-86.
- Karim M.E., 2008, *Frailty Models*, Institute of statistical research and training, University of Dhaka, Dhaka, Bangladesh.
- Keyfitz N., 1985, Choice of functions for mortality analysis: Effective forecasting depends on a minimum parameter representation, *Theoretical Population Biology*, Vol. 21, 329-352.
- Kalbfleisch J.D., Prentice R.L., 1980, *The statistical analysis of failure time data*, John Wiley & Sons, New York.
- Kaplan E.L., Meier P., 1958, Nonparametric estimation from incomplete observations, *Journal of the American Statistical Association*, Vol 53, 457-481.
- Keiding N., Andersen P., Klein J., 1997, The role of frailty models and accelerated failure time models in describing heterogeneity due to omitted covariates, *Statistics in Medicine*, Vol 16, 215-224.
- Kendall M.G., 1938, A new measure of rank collection, *Biometrika*, Vol 30, 81-93.
- Keyfitz N., Littman G., 1979, Mortality in heterogeneous populations: Distributions describing heterogeneity, *Biometrika*, Vol 71, 75-83.
- Keyfitz N., 1985, Choice of Functions for Mortality Analysis: Effective Forecasting Depends on a Minimum Parameter Representation, *Theoretical Population Biology*, Vol 21, 329-352.
- Klein J.P., Moeschberger M.L., 1997, *Survival analysis-techniques for censored and truncated data*, Springer, New York.

- Klugman S.A., Panjer H.H., Willmot G.E., 2008, Loss models from data to decisions, Third Edition, John Wiley & Sons, New York.
- Lee R.D., Carter L., 1992, Modelling and Forecasting The time series of US Mortality, Journal of the American Statistical Association, Vol. 87, 659-671.
- Levinson L.H., 1959, A theory of mortality classes, Transactions of the Society of Actuaries, Vol 13, 325-358.
- Makeham W.M., 1867, On the law of mortality, Journal of the Institute of Actuaries, Vol 8, 301-310.
- Makeham W.M., 1890, On the further development of Gompertz's law, Journal of the Institute of Actuaries, Vol 28, 316-332.
- Manatunga A.K., Oakes D., 1999, Parametric analysis of matched pair survival data, Lifetime Data Analysis, Vol. 5, 371-387.
- Manton K.G., Stallard E., Vaupel J.W., 1981, Methods for comparing the mortality experience of heterogeneous populations, Demography, Vol 18, 389-410.
- Manton K.G., Vaupel J.W., 1986, Alternative models for the heterogeneity of mortality risks among the aged, Journal of the American Statistical Association, Vol.81, 635-644.
- Manton K.G., Vaupel J.W., 1995, Survival after the age of 80 in the United States, Sweden, France, England and Japan, New England Journal of Medicine, Vol 333, 1232-1235.
- Martine T.E.P., Shekary N., Widdershoven G., Heldens J., Deeg D.J.H., The meaning of frailty according to Dutch older frail and non-frail persons, Journal of Aging Studies, Volume 23, Issue 4, 258-266.
- McCullagh P., Nelder J.A., 1989, Generalised linear models, Chapman and Hall 2-6 Boundary Row, London.
- McGilchrist C.A., Aisbett C.W., 1991, Regression with frailty survival analysis, Biometrics, Vol 47, 461-466.
- McNown R., Rogers A., 1989, Forecasting Mortality: A parametrized Time Series Approach, Demography, Vol. 26, 645-660.
- Mudholkar G.S., Hutson A.D., 1996, The exponentiated Weibull distribution, Metrika, Vol. 62, Part 1, 17-22.
- Neill A., 1983, Life Contingencies, William Heinemann, London.
- Olivieri A., 2002, Risk and Reinsurance in life annuity portfolios, Quaderni del Dipartimento di Matematica Applicata Alle Scienze Economiche Statistiche e Attuariali, Università di Trieste.

- Olivieri A., Pitacco E., 2002, Inference about mortality improvements in life annuity portfolios, Transactions of the 27th International Congress of Actuaries, Cancun, Mexico.
- Olivieri A., 2006, Heterogeneity in survival models : Application to pensions and life annuities, University of Parma, Italy.
- Perks W., 1932, On some experiments on the graduation of mortality statistics, Journal of the Institute of Actuaries, Vol 63, 12-40.
- Pickles A., Crouchley R., Simonoff E., Eaves L., Meyer J., Rutter M., Hewitt J., Silberg J., 1994, Survival Models for Developmental Genetic Data: Age of Onset of Puberty and Antisocial Behavior in Twins, Genetic Epidemiology, Vol 11, 155 – 170.
- Pike M.C., 1966, A method of analysis of a certain class of experiments in carcinogenesis, Biometrics, Vol. 22, 142-161.
- Pitacco E., 2003, Survival models in actuarial mathematics : from Halley to Longevity, Transactions of the 7th International Congress of Insurance: Mathematics and Economics.
- Pitacco E., 2004, Survival Models in a dynamic context: a survey, Insurance: Mathematics and Economics. Vol 35, Part 2.
- Pitacco E., Olivieri A., 2005, Forecasting mortality : an introduction, CERAP Working Paper No. 35, Universita Bocconi, Milano.
- Pitacco E., Denuit M., Haberman S., Olivieri A., 2009, Modelling Longevity Dynamics for Pensions and Annuity Business, Oxford University Press, New York, U.S.A.
- Plackett R.L., 1983, Karl Pearson and The Chi-Squared Test, International Statistical Review, Vol 51, 59-72.
- Price D.L., Manatunga A.K., 2001, Modelling survival data with a cured fraction using frailty models, Statistics in Medicine, Vol 20, 1515-1527.
- Renshaw A.E., 1991, Actuarial graduation practice and generalized linear and non-linear models, Journal of the Institute of Actuaries, Vol. 118, 295-312.
- Renshaw A.E., 1995, Graduation and generalized linear models: An overview, Department of Actuarial Science and Statistics, City University, London, Actuarial Research Paper No:73.
- Redington F.M., 1969, An exploration into patterns of mortality, Journal of the Institute of Actuaries, Vol 95, 243-298.
- Reid N., 1994, A conversation with Sir David Cox, Statistical Science, Vol 9, 439-455.
- Rogers A., Gard K., 1991, Applications of the Heligman-Pollard Model Mortality Schedule, Population Bulletin of the United Nations, Vol. 30, 79-105.

- Siegel S., 1956, Non-parametric statistics for the behavioral sciences, New York: McGraw-Hill, 75-83.
- Spearman C., 1904, The proof and measurement of association between two things, American Journal of Psychology, Vol.15, 72-101.
- Su S., Sherris M., 2011, Heterogeneity of Australian Population Mortality and Implications for a Viable Life Annuity Market, Australian School of Business, University of New South Wales, Australia.
- Thiele, P.N., 1871, On a mathematical formula to express the rate of mortality throughout the whole of life, Journal of the Institute of Actuaries, Vol 16, 313–329.
- Tosetti A., Behar T., Fromenteau M., Menart S., 2000, Assurance Comptabilité, Réglementation et Actuariat, Economica.
- Tweedy M., 1984, An index which distinguishes between some important exponential families, Statistics: Applications and New Directions, Proceedings of the Indian Statistical Institute Golden Jubilee International Conference J. Ghosh, J. Roy (eds), 579-604.
- Vaupel J.W., Manton K.G., Stallard E., 1979, The impact of heterogeneity in individual frailty on the dynamics of mortality, Demography, Vol 16, 439-454.
- Vaupel J.W., Yashin A.I., 1985a, The deviant dynamics of death in heterogeneous populations, Sociology Methodology, 179-211, Edited N.B. Tuma, Jossey-Bas, San Francisco.
- Vaupel J.W., Yashin A.I., 1985b, Heterogeneity's ruses: Some surprising effects of selection on population dynamics, The American Statistician, Vol 39, 176-185.
- Vaupel J.W., 1999, Contribution to the discussion of 'The long-term pattern of adult mortality and the highest attained age, by A.R. Thatcher, Journal of the Royal Statistical Society, Series A, Vol 162, 31-32.
- Venables W.N., Ripley B.D., 1999, Modern applied statistics with S-Plus, Springer-Verlag, New York, Second Edition.
- Wang S.S., Brown R.L., 1998, A frailty model for projection of human mortality improvement, Journal of Actuarial Practice 6, 221-241.
- Weibull W., 1951, A statistical distribution function with wide applicability. Appl.Mech., Vol 18, 293–297.
- Wienke A., 1996, Asymptotisch optimale adaptive Auswahlverfahren in semiparammetrischen modellen. Hans Jacobs, Lage.
- Wienke A., Holm N., Skytthe A., Yashin A.I., 2001, The heritability of mortality due to heart diseases: a correlated frailty model applied to Danish twins, Twin Research 4, 266 – 274.

- Wienke A., Christensen K., Skytthe A., Yashin A.I., 2002, Genetic analysis of cause of death in a mixture model with bivariate lifetime data, *Statistical Modelling*, Vol 2, 89 – 102.
- Wienke A., 2006, *Frailty models*, Güstrow, Deutsch.
- Willemse W.J., Kaas R., 2007, Rational reconstruction of frailty-based mortality models by a generalisation of Gompertz' law of mortality, *Insurance: Mathematics and Economics* , Vol 40, 468-484.
- Wolfgang H., 1990, *Applied Nonparametric Regression*, Cambridge University Press, 4-5.
- Yashin A.I., Manton K.G., Vaupel J.W., 1985, Mortality and aging in a heterogeneous population: A stochastic process model with observed and unobserved variables, *Theoretical Population Biology*, Vol 27, 154-175.
- Yashin A. I., Manton K. G., 1994, Modifications of the EM algorithm for survival influenced by an unobserved stochastic process, *Stochastic Process. Appl.*, Vol 54, 257-274.
- Yashin A.I., Iachine I.A., 1995, Genetic Analysis of Durations: Correlated Frailty Model Applied to Survival of Danish Twins, *Genetic Epidemiology*, Vol 12, 529 – 538.
- Yashin A.I., Vaupel J.W., Iachine I.A., 1995, Correlated individual frailty: An advantageous approach to survival analysis of bivariate data, *Mathematical Population Studies*, Vol 5, 145-159.
- Yıldırım F., 2010, Türkiye ölüm yapısının Lee-Carter ve Bulanık Lee-Carter ile modellenmesi, Hacettepe Üniversitesi, Ankara, Türkiye.
- Zahl P.H., 1997, Frailty modelling for the excess hazard, *Statistics in Medicine*, Vol 16, 1573-1585.
- Zdravkovic S., Wienke A., Pedersen N.L., Marenberg M.E., Yashin A.I., de Faire U. , 2002, Heritability of death from coronary heart disease: a 36-year follow-up of 20,966 Swedish twins, *Journal of Internal Medicine*, Vol 252, 247 – 254.

EKLER DİZİNİ

EK 1. GOMPERTZ/GAMMA HASSASİYET MODELİ

EK 2.GOMPERTZ/TERS GAUSS HASSASİYET MODELİ

EK 3. GOMPERTZ/ POZİTİF DURAĞAN DAĞILIM HASSASİYET MODELİ

EK 4. GOMPERTZ/ BİLEŞİK POISSON DAĞILIM HASSASİYET MODELİ

EK 5. GOMPERTZ/ GÜÇ VARYANS FONKSİYONU HASSASİYET MODELİ

EK 6. KADIN SİGORTALILAR İÇİN GOMPERTZ/TERS GAUSS HASSASİYET
MODELLEMESİNE GÖRE ÖLÜM HIZLARI

EK 7. ERKEK SİGORTALILAR İÇİN GOMPERTZ/GAMMA HASSASİYET
MODELLEMESİNE GÖRE İÇİN ÖLÜM HIZLARI

EK 8. TRSH-2010 KADIN HAYAT TABLOSU

EK 9. TRSH-2010 ERKEK HAYAT TABLOSU

EK 1. GOMPERTZ/GAMMA HASSASIYET MODELİ

$$\mu_x = \left(\frac{b \exp(px)}{1 + \theta \frac{b}{p} (\exp(px) - 1)} \right) \left(\frac{1 - \theta \frac{b}{p}}{1 - \theta \frac{b}{p}} \right) = \frac{\frac{b}{1 - \theta \frac{b}{p}} b \exp(px)}{\left(1 + \frac{\left(\frac{\theta b}{p} \right)}{\left(\frac{p - \theta b}{p} \right)} \exp(px) \right)}$$

$$= \frac{\frac{bp}{(p - \theta b)} b \exp(px)}{\left(1 + \frac{\left(\frac{\theta b}{p} \right)}{\left(\frac{p - \theta b}{p} \right)} \exp(px) \right)} = \frac{\frac{bp}{(p - \theta b)} b \exp(px)}{\left(1 + \left(\frac{\theta b}{p - \theta b} \right) \exp(px) \right)}$$

$$= \frac{\exp(-g + px)}{1 + \exp(-s + px)}$$

$$\exp(-g) = \frac{bp}{(p - \theta b)}$$

$$\exp(-s) = \frac{\theta b}{p - \theta b}$$

EK 2. GOMPERTZ/TERS GAUSS HASSASIYET MODELİ

$$\mu_x = \frac{b \exp(px)}{\left(1 + \frac{2b}{p\alpha}(\exp(px) - 1)\right)^{1/2}} = \frac{b \exp(px)}{\left(1 + \frac{2b}{p\alpha} \exp(px) - \frac{2b}{p\alpha}\right)^{1/2}}$$

$$= \left(\frac{b \exp(px)}{\left(1 + \frac{2b}{p\alpha} \exp(px) - \frac{2b}{p\alpha}\right)^{1/2}} \right) \frac{\sqrt{1 - \frac{2b}{p\alpha}}}{\sqrt{1 - \frac{2b}{p\alpha}}}$$

$$= \frac{b \exp(px) / \sqrt{1 - \frac{2b}{p\alpha}}}{\left(\left(1 + \frac{2b}{p\alpha} \exp(px) - \frac{2b}{p\alpha}\right) / \left(1 - \frac{2b}{p\alpha}\right) \right)^{1/2}}$$

$$= \frac{\frac{b}{\sqrt{1 - \frac{2b}{p\alpha}}} \exp(px)}{\left(\left(1 + \frac{2b}{p\alpha} \exp(px) - \frac{2b}{p\alpha}\right) / \left(1 - \frac{2b}{p\alpha}\right) \right)^{1/2}} = \frac{\exp(-g + px)}{(1 + \exp(-s + px))^{1/2}}$$

$$\exp(-g) = \frac{b}{\sqrt{1 - \frac{2b}{p\alpha}}} \quad , \quad \frac{2b}{p\alpha} < 1$$

$$\exp(-s) = \frac{2b}{p\alpha - 2b}$$

EK 3. GOMPERTZ/ POZİTİF DURAĞAN DAĞILIM HASSASİYET MODELİ

$$\begin{aligned}\mu_x &= \theta b \exp(px) \left[\frac{b}{p} (\exp(px) - 1) \right]^{\theta-1} = \frac{\theta b \exp(px)}{\left[\frac{b}{p} (\exp(px) - 1) \right]^{1-\theta}} \\&= \left(\frac{\theta b \exp(px)}{\left(\frac{b}{p} (\exp(px) - 1) \right)^{1-\theta}} \right) \frac{(-b/p)^{1-\theta}}{(-b/p)^{1-\theta}} \\&= \frac{(\theta b / (-b/p)^{1-\theta}) \exp(px)}{\left[\left(\frac{b}{p} (\exp(px) - 1) \right) / \left(-\frac{b}{p} \right) \right]^{1-\theta}} \\&= \frac{(\theta b / (-b/p)^{1-\theta}) \exp(px)}{(1 - \exp(px))^{1-\theta}} \\&= \frac{\exp(-g + px)}{(1 + \exp(-s + px))^{1-\theta}}\end{aligned}$$

$$\exp(-g) = \theta b / (-b/p)^{1-\theta}, -p < b$$

$$\exp(-s) = -1$$

Burada $\theta \approx (0,1]$ olması gerekmektedir.

EK 4. GOMPERTZ/ BİLEŞİK POISSON DAĞILIM HASSASİYET MODELİ

$$\begin{aligned}\mu_x &= \frac{kb \exp(px)}{\left(\lambda + \frac{b}{p} - \frac{b}{p} \exp(px)\right)^{1-\gamma}} = \left[\frac{kb \exp(px)}{\left(\lambda + \frac{b}{p} - \frac{b}{p} \exp(px)\right)^{1-\gamma}} \right] \frac{\left(\lambda + \frac{b}{p}\right)^{1-\gamma}}{\left(\lambda + \frac{b}{p}\right)^{1-\gamma}} \\ &= \frac{kb \exp(px) \left(\lambda + \frac{b}{p}\right)^{\gamma-1}}{\left(1 + \left(\frac{-b}{\lambda p + b}\right) \exp(px)\right)^{1-\gamma}} = \frac{\exp(-g + px)}{(1 + \exp(-s + px))^{1-\gamma}}\end{aligned}$$

$$\exp(-g) = kb \left(\lambda + \frac{b}{p}\right)^{\gamma-1}, \lambda + \frac{b}{p} > 0$$

$$\exp(-s) = \frac{-b}{\lambda p + b}$$

Burada $\gamma < 0$ olması gerekmektedir.

EK 5. GOMPERTZ/ GÜÇ VARYANS FONKSİYONU HASSASİYET MODELİ

$$\begin{aligned}
 \mu_x &= \frac{b \exp(px)}{\left(1 + \frac{\sigma^2}{1-\gamma} \frac{b}{p} (\exp(px) - 1)\right)^{1-\gamma}} = \left[\frac{b \exp(px)}{\left(1 + \frac{\sigma^2}{1-\gamma} \frac{b}{p} (\exp(px) - 1)\right)^{1-\gamma}} \right] \frac{\left(1 - \frac{\sigma^2 b}{(1-\gamma)p}\right)^{1-\gamma}}{\left(1 - \frac{\sigma^2 b}{(1-\gamma)p}\right)^{1-\gamma}} \\
 &= \frac{b \left(1 - \frac{\sigma^2 b}{(1-\gamma)p}\right)^{\gamma-1} \exp(px)}{\left(\left(1 + \frac{\sigma^2}{1-\gamma} \frac{b}{p} (\exp(px) - 1)\right) / \left(1 - \frac{\sigma^2 b}{(1-\gamma)p}\right)\right)^{1-\gamma}} \\
 &= \frac{b \left(1 - \frac{\sigma^2 b}{(1-\gamma)p}\right)^{\gamma-1} \exp(px)}{\left(1 + \left(\frac{\sigma^2 b}{(1-\gamma)p - \sigma^2 b}\right) \exp(px)\right)^{1-\gamma}} \\
 &= \frac{\exp(-g + px)}{(1 + \exp(-s + px))^{1-\gamma}}
 \end{aligned}$$

$$\exp(-g) = b \left(1 - \frac{\sigma^2 b}{(1-\gamma)p}\right)^{\gamma-1}, 1 > \frac{\sigma^2 b}{(1-\gamma)p}$$

$$\exp(-s) = \left(\frac{\sigma^2 b}{(1-\gamma)p - \sigma^2 b}\right)$$

Burada $\gamma > 0$ olması gerekmektedir.

**EK 6. KADIN SİĞORTALILAR İÇİN GOMPERTZ/TERS GAUSS HASSASIYET
MODELLEMESİNE GÖRE ÖLÜM HIZLARI**

Yaş	Ölüm Hızı	Yaş	Ölüm Hızı
25	0,000182	68	0,004113
26	0,000188	69	0,004622
27	0,000196	70	0,005201
28	0,000206	71	0,005861
29	0,000216	72	0,006639
30	0,000228	73	0,007517
31	0,000241	74	0,008512
32	0,000256	75	0,009601
33	0,000272	76	0,010829
34	0,000289	77	0,012203
35	0,000308	78	0,013783
36	0,000329	79	0,015543
37	0,000351	80	0,017556
38	0,000375	81	0,019833
39	0,000401	82	0,022484
40	0,000429	83	0,025403
41	0,000459	84	0,028581
42	0,000492	85	0,031968
43	0,000527	86	0,03568
44	0,000564	87	0,03978
45	0,000605	88	0,044301
46	0,000648	89	0,049354
47	0,000695	90	0,055002
48	0,000745	91	0,061578
49	0,000799	92	0,068852
50	0,000857	93	0,076532
51	0,00092	94	0,084021
52	0,000986	95	0,091638
53	0,001058	96	0,099795
54	0,001136	97	0,108747
55	0,001219	98	0,118295
56	0,001308	99	0,128216
57	0,001403	100	0,138448
58	0,001506	101	0,14892
59	0,001616	102	0,159554
60	0,001735	103	0,170269
61	0,001916	104	0,180979
62	0,002132	105	0,191598
63	0,002376	106	0,202045
64	0,002646	107	0,212244
65	0,002943	108	0,222123
66	0,003276	109	0,231624
67	0,003665	110	0,240696

**EK 7. ERKEK SİGORTALILAR İÇİN GOMPERTZ/GAMMA HASSASİYET
MODELLEMESİNE GÖRE ÖLÜM HIZLARI**

Yaş	Ölüm Hızı	Yaş	Ölüm Hızı
25	0,000443	68	0,020359
26	0,000487	69	0,022242
27	0,000534	70	0,024291
28	0,000586	71	0,026545
29	0,000643	72	0,029103
30	0,000705	73	0,031902
31	0,000773	74	0,034981
32	0,000848	75	0,038270
33	0,000930	76	0,041897
34	0,001019	77	0,045866
35	0,001117	78	0,050299
36	0,001225	79	0,055059
37	0,001342	80	0,060322
38	0,001471	81	0,066212
39	0,001611	82	0,073146
40	0,001765	83	0,080892
41	0,001932	84	0,089238
42	0,002115	85	0,097809
43	0,002315	86	0,106902
44	0,002532	87	0,116708
45	0,002769	88	0,127770
46	0,003028	89	0,140383
47	0,003309	90	0,154963
48	0,003614	91	0,171654
49	0,003946	92	0,189108
50	0,004306	93	0,206550
51	0,004696	94	0,222805
52	0,005118	95	0,239691
53	0,005575	96	0,257669
54	0,006068	97	0,277376
55	0,006599	98	0,297946
56	0,007171	99	0,319131
57	0,007786	100	0,340827
58	0,008445	101	0,362913
59	0,009151	102	0,385263
60	0,009905	103	0,407742
61	0,010851	104	0,430214
62	0,011927	105	0,452541
63	0,013112	106	0,474590
64	0,014367	107	0,496234
65	0,015685	108	0,517356
66	0,017079	109	0,537852
67	0,018639	110	0,557627

EK 8. TRSH-2010 KADIN HAYAT TABLOSU

TRSH - 2010 KADIN											
YAŞ	q _x	p _x	l _x	d _x	e _x	YAŞ	q _x	p _x	l _x	d _x	e _x
0	0,008161	0,991839	1000000,00	8161,00	81,62	56	0,002863	0,997137	966220,51	2766,29	27,33
1	0,000266	0,999734	991839,00	263,70	81,29	57	0,003300	0,996700	963454,22	3179,40	26,41
2	0,000214	0,999786	991575,30	212,49	80,31	58	0,003823	0,996177	960274,83	3671,13	25,49
3	0,000171	0,999829	991362,81	169,72	79,33	59	0,004415	0,995585	956603,70	4223,41	24,59
4	0,000136	0,999864	991193,09	134,39	78,34	60	0,005113	0,994887	952380,29	4869,52	23,70
5	0,000105	0,999895	991058,70	104,51	77,35	61	0,005648	0,994352	947510,77	5351,54	22,82
6	0,000093	0,999907	990954,19	91,99	76,36	62	0,006285	0,993715	942159,23	5921,47	21,94
7	0,000078	0,999922	990862,21	77,78	75,37	63	0,007002	0,992998	936237,76	6555,54	21,08
8	0,000064	0,999936	990784,43	63,78	74,37	64	0,007800	0,992200	929682,22	7251,52	20,22
9	0,000052	0,999948	990720,65	51,88	73,38	65	0,008675	0,991325	922430,70	8002,09	19,38
10	0,000044	0,999956	990668,77	43,99	72,38	66	0,009657	0,990343	914428,61	8830,64	18,54
11	0,000039	0,999961	990624,78	38,83	71,38	67	0,010802	0,989198	905597,98	9782,27	17,72
12	0,000036	0,999965	990585,96	35,14	70,39	68	0,012124	0,987876	895815,71	10860,87	16,91
13	0,000035	0,999965	990550,82	34,82	69,39	69	0,013623	0,986377	884954,84	12055,74	16,11
14	0,000040	0,999960	990516,00	39,75	68,39	70	0,015329	0,984671	872899,10	13380,67	15,32
15	0,000052	0,999948	990476,25	51,85	67,39	71	0,017275	0,982725	859518,43	14848,18	14,56
16	0,000077	0,999923	990424,40	75,99	66,40	72	0,019568	0,980432	844670,25	16488,51	13,80
17	0,000112	0,999888	990348,41	110,93	65,40	73	0,022158	0,977842	828141,74	18349,96	13,07
18	0,000151	0,999849	990237,48	149,29	64,41	74	0,025090	0,974910	809791,77	20317,68	12,35
19	0,000186	0,999814	990088,18	183,73	63,42	75	0,028300	0,971700	789474,10	22342,12	11,66
20	0,000209	0,999791	989904,45	206,89	62,43	76	0,031918	0,968082	767131,98	24485,32	10,98
21	0,000223	0,999777	989697,56	220,70	61,44	77	0,035968	0,964032	742646,66	26711,52	10,33
22	0,000225	0,999775	989476,86	222,63	60,46	78	0,040627	0,959373	715935,15	29086,30	9,69
23	0,000220	0,999780	989254,23	217,64	59,47	79	0,045813	0,954187	686848,85	31466,61	9,08
24	0,000210	0,999790	989036,59	207,70	58,48	80	0,051747	0,948253	655382,24	33914,06	8,50
25	0,000205	0,999795	988828,89	202,71	57,50	81	0,058458	0,941542	621468,18	36329,79	7,93
26	0,000212	0,999788	988626,18	209,59	56,51	82	0,066273	0,933727	585138,39	38778,88	7,39
27	0,000225	0,999775	988416,59	222,39	55,52	83	0,074878	0,925122	546359,52	40910,31	6,88
28	0,000239	0,999761	988194,20	234,20	54,53	84	0,084245	0,915755	505449,21	42581,57	6,40
29	0,000248	0,999752	987960,00	245,01	53,55	85	0,094229	0,905771	462867,64	43615,55	5,94
30	0,000265	0,999735	987714,98	261,74	52,56	86	0,105169	0,894831	419252,08	44092,32	5,51
31	0,000289	0,999711	987453,24	285,37	51,57	87	0,117254	0,882746	375159,76	43988,98	5,10
32	0,000310	0,999690	987167,87	306,02	50,59	88	0,130580	0,869420	331170,78	43244,28	4,71
33	0,000324	0,999676	986861,84	319,74	49,60	89	0,145476	0,854524	287926,50	41886,40	4,34
34	0,000332	0,999668	986542,10	327,53	48,62	90	0,162124	0,837876	246040,10	39889,01	3,99
35	0,000343	0,999657	986214,57	338,27	47,63	91	0,181507	0,818493	206151,10	37417,87	3,67
36	0,000362	0,999638	985876,30	356,89	46,65	92	0,202947	0,797053	168733,23	34243,90	3,37
37	0,000388	0,999612	985519,41	382,38	45,67	93	0,225584	0,774416	134489,33	30338,64	3,11
38	0,000421	0,999579	985137,03	414,74	44,68	94	0,247657	0,752343	104150,69	25793,65	2,86
39	0,000457	0,999543	984722,29	450,02	43,70	95	0,270110	0,729890	78357,04	21165,02	2,64
40	0,000492	0,999508	984272,27	484,26	42,72	96	0,294153	0,705847	57192,02	16823,20	2,44
41	0,000532	0,999468	983788,01	523,38	41,74	97	0,320539	0,679461	40368,82	12939,78	2,24
42	0,000581	0,999419	983264,63	571,28	40,77	98	0,348684	0,651316	27429,04	9564,07	2,07
43	0,000642	0,999358	982693,35	630,89	39,79	99	0,377927	0,622073	17864,97	6751,65	1,90
44	0,000716	0,999284	982062,46	703,16	38,81	100	0,408085	0,591915	11113,32	4535,18	1,76
45	0,000800	0,999200	981359,31	785,09	37,84	101	0,438952	0,561048	6578,14	2887,49	1,62
46	0,000876	0,999124	980574,22	858,98	36,87	102	0,470299	0,529701	3690,65	1735,71	1,50
47	0,000944	0,999056	979715,24	924,85	35,90	103	0,501881	0,498119	1954,94	981,15	1,40
48	0,001011	0,998989	978790,39	989,56	34,94	104	0,533448	0,466552	973,79	519,47	1,30
49	0,001083	0,998917	977800,83	1058,96	33,97	105	0,564750	0,435250	454,33	256,58	1,21
50	0,001234	0,998766	976741,87	1205,30	33,01	106	0,595544	0,404456	197,75	117,77	1,13
51	0,001415	0,998585	975536,57	1380,38	32,05	107	0,625604	0,374396	79,98	50,04	1,05
52	0,001632	0,998368	974156,19	1589,82	31,09	108	0,654725	0,345275	29,94	19,61	0,97
53	0,001881	0,998119	972566,36	1829,40	30,14	109	0,682730	0,317270	10,34	7,06	0,86
54	0,002168	0,997832	970736,97	2104,56	29,20	110	0,709471	0,290529	3,28	2,33	0,65
55	0,002490	0,997510	968632,41	2411,89	28,26						

EK 9. TRSH-2010 ERKEK HAYAT TABLOSU

TRSH - 2010 ERKEK											
YAŞ	q _x	p _x	l _x	d _x	e _x	YAŞ	q _x	p _x	l _x	d _x	e _x
0	0,019533	0,980467	1000000,00	19533,00	74,32	56	0,008392	0,991608	895900,43	7518,40	22,94
1	0,000849	0,999151	980467,00	832,89	74,79	57	0,009229	0,990771	888382,03	8198,88	22,13
2	0,000757	0,999243	979634,11	741,36	73,85	58	0,010148	0,989852	880183,16	8932,10	21,33
3	0,000656	0,999344	978892,75	642,47	72,90	59	0,011125	0,988875	871251,06	9692,67	20,55
4	0,000559	0,999441	978250,28	546,97	71,95	60	0,012183	0,987817	861558,39	10496,37	19,77
5	0,000473	0,999527	977703,32	462,75	70,99	61	0,013347	0,986653	851062,02	11359,12	19,01
6	0,000389	0,999611	977240,57	379,88	70,03	62	0,014670	0,985330	839702,90	12318,44	18,26
7	0,000327	0,999673	976860,69	318,95	69,05	63	0,016128	0,983872	827384,46	13344,06	17,53
8	0,000283	0,999717	976541,74	276,41	68,08	64	0,017672	0,982328	814040,40	14385,72	16,81
9	0,000255	0,999745	976265,34	248,71	67,09	65	0,019293	0,980707	799654,68	15427,74	16,10
10	0,000238	0,999762	976016,62	232,35	66,11	66	0,021008	0,978992	784226,94	16475,04	15,41
11	0,000235	0,999765	975784,28	229,63	65,13	67	0,022926	0,977074	767751,90	17601,48	14,73
12	0,000249	0,999751	975554,65	242,88	64,14	68	0,025042	0,974958	750150,42	18785,27	14,06
13	0,000275	0,999725	975311,77	268,58	63,16	69	0,027358	0,972642	731365,15	20008,69	13,41
14	0,000311	0,999689	975043,18	303,22	62,18	70	0,029879	0,970121	711356,47	21254,62	12,77
15	0,000352	0,999648	974739,96	343,28	61,19	71	0,032651	0,967349	690101,85	22532,52	12,15
16	0,000406	0,999594	974396,67	395,39	60,22	72	0,035798	0,964202	667569,33	23897,65	11,54
17	0,000474	0,999526	974001,29	461,84	59,24	73	0,039241	0,960759	643671,68	25258,32	10,95
18	0,000547	0,999453	973539,44	532,67	58,27	74	0,043028	0,956972	618413,36	26609,09	10,38
19	0,000614	0,999386	973006,78	597,91	57,30	75	0,047073	0,952927	591804,27	27858,00	9,82
20	0,000666	0,999334	972408,87	647,62	56,33	76	0,051535	0,948465	563946,27	29062,97	9,28
21	0,000686	0,999314	971761,25	666,63	55,37	77	0,056416	0,943584	534883,30	30175,98	8,76
22	0,000706	0,999294	971094,62	685,59	54,41	78	0,061869	0,938131	504707,32	31225,74	8,25
23	0,000721	0,999279	970409,02	699,66	53,45	79	0,067724	0,932276	473481,59	32066,07	7,77
24	0,000734	0,999266	969709,36	711,77	52,49	80	0,074198	0,925802	441415,52	32752,15	7,29
25	0,000754	0,999246	968997,59	730,62	51,52	81	0,081443	0,918557	408663,37	33282,77	6,84
26	0,000771	0,999229	968266,97	746,53	50,56	82	0,089972	0,910028	375380,60	33773,74	6,40
27	0,000776	0,999224	967520,43	750,80	49,60	83	0,099500	0,900500	341606,86	33989,88	5,98
28	0,000774	0,999226	966769,64	748,28	48,64	84	0,109766	0,890234	307616,97	33765,88	5,59
29	0,000778	0,999222	966021,36	751,56	47,68	85	0,120308	0,879692	273851,09	32946,48	5,22
30	0,000800	0,999200	965269,79	772,22	46,71	86	0,131493	0,868507	240904,61	31677,27	4,86
31	0,000837	0,999163	964497,58	807,28	45,75	87	0,143554	0,856446	209227,34	30035,42	4,52
32	0,000886	0,999114	963690,29	853,83	44,79	88	0,157161	0,842839	179191,92	28161,98	4,20
33	0,000950	0,999050	962836,46	914,69	43,83	89	0,172675	0,827325	151029,94	26079,09	3,89
34	0,001018	0,998982	961921,77	979,24	42,87	90	0,190609	0,809391	124950,84	23816,76	3,59
35	0,001080	0,998920	960942,53	1037,82	41,91	91	0,211139	0,788861	101134,09	21353,35	3,32
36	0,001140	0,998860	959904,72	1094,29	40,96	92	0,232608	0,767392	79780,74	18557,64	3,08
37	0,001207	0,998793	958810,42	1157,28	40,00	93	0,254063	0,745937	61223,10	15554,52	2,86
38	0,001294	0,998706	957653,14	1239,20	39,05	94	0,274057	0,725943	45668,58	12515,79	2,66
39	0,001409	0,998591	956413,94	1347,59	38,10	95	0,294827	0,705173	33152,78	9774,34	2,48
40	0,001556	0,998444	955066,35	1486,08	37,15	96	0,316941	0,683059	23378,45	7409,59	2,30
41	0,001740	0,998260	953580,27	1659,23	36,21	97	0,341181	0,658819	15968,86	5448,27	2,14
42	0,001957	0,998043	951921,04	1862,91	35,27	98	0,366482	0,633518	10520,59	3855,61	1,99
43	0,002202	0,997798	950058,13	2092,03	34,34	99	0,392541	0,607459	6664,98	2616,28	1,85
44	0,002471	0,997529	947966,10	2342,42	33,42	100	0,419227	0,580773	4048,70	1697,33	1,72
45	0,002778	0,997222	945623,68	2626,94	32,50	101	0,446394	0,553606	2351,38	1049,64	1,61
46	0,003102	0,996898	942996,73	2925,18	31,59	102	0,473885	0,526115	1301,74	616,87	1,50
47	0,003448	0,996552	940071,56	3241,37	30,68	103	0,501535	0,498465	684,86	343,48	1,41
48	0,003824	0,996176	936830,19	3582,44	29,79	104	0,529176	0,470824	341,38	180,65	1,32
49	0,004237	0,995763	933247,75	3954,17	28,90	105	0,556639	0,443361	160,73	89,47	1,24
50	0,004690	0,995310	929293,58	4358,39	28,02	106	0,583760	0,416240	71,26	41,60	1,16
51	0,005182	0,994818	924935,19	4793,01	27,15	107	0,610383	0,389617	29,66	18,11	1,09
52	0,005724	0,994276	920142,18	5266,89	26,29	108	0,636364	0,363636	11,56	7,35	1,01
53	0,006310	0,993690	914875,29	5772,86	25,44	109	0,661574	0,338426	4,20	2,78	0,89
54	0,006947	0,993053	909102,42	6315,53	24,60	110	0,685898	0,314102	1,42	0,98	0,66
55	0,007628	0,992372	902786,89	6886,46	23,76						

ÖZGEÇMİŞ

Adı Soyadı : Funda KUL

Doğum Yeri : Kayseri

Doğum Yılı : 1985

Medeni Hali : Evli

Eğitim ve Akademik Durumu:

Lise : 2000 - 2003 Ankara Sokullu Mehmet Paşa Anadolu Lisesi

Lisans : 2003 - 2007 Hacettepe Üniversitesi Aktüerya Bilimleri Bölümü

Yabancı Dil : İngilizce, İspanyolca

İş Tecrübesi :

2008 - ... Hacettepe Üniversitesi Aktüerya Bilimleri Bölümü