

T.C.
MARMARA ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
EKONOMETRİ ANABİLİM DALI
İSTATİSTİK BİLİM DALI

OLAY ZAMANI ANALİZİNDE TESADÜFİ ETKİLER
POISSON REGRESYON MODELİ İLE
GÖZLEMLENEMEYEN HETEROJENLİĞİN İNCELENMESİ

DOKTORA TEZİ

Nilay KÖLEOĞLU

Danışman: DOÇ. DR. AHMET METE ÇİLİNGİRTÜRK

İSTANBUL, 2006

İÇİNDEKİLER

	Sayfa No.
TABLO LİSTESİ	III
ŞEKİL LİSTESİ	IV
1. GİRİŞ	1
2. OLAY ZAMANI VERİLERİNİN KAPSAMI VE KULLANILAN ANALİZ YÖNTEMLERİ	5
2.1 Olay Zamanı Analizinin Kullanım Amacı.....	8
2.2 Olay Zamanı Analizi Veri Türü.....	12
2.2.1 Durdurma.....	13
2.2.1.1 I.Tip Durdurma.....	14
2.2.1.2 II. Tip Durdurma.....	16
2.2.1.3 Tesadüfi Durdurma.....	17
2.2.1.4 Sağdan Durdurma.....	18
2.2.1.5 Soldan Durdurma.....	20
2.2.1.6 Aralıklı Durdurma.....	21
2.2.2 Kesme.....	22
2.2.2.1 Soldan Kesme.....	23
2.2.2.2 Sağdan Kesme.....	24
2.2.3 Olay Zamanı Analizinde Kullanılan Değişkenler.....	24
2.3 Olay Zamanı Analizinde Kullanılan Yöntemler.....	26
2.3.1 Durdurulmamış Veriler İçin Yaşam Fonksiyonu Yöntemleri.....	30
2.3.2 Durdurulmuş Veriler İçin Yaşam Fonksiyonu Yöntemleri.....	32
2.3.3 Kesilmiş Veriler İçin Yaşam Fonksiyonu Yöntemleri.....	33
2.4 Olay Zamanı Analizinde Kullanılan Regresyon Modelleri.....	33
2.4.1 Üssel Regresyon Modeli.....	35
2.4.2 Weibull Regresyon Modeli.....	38
2.4.3 Oransal Hazard Modeli.....	39
2.4.4 Cox Oransal Hazard Regresyon Modeli.....	43
2.4.5 Poisson Regresyon Modeli.....	45
3. POISSON REGRESYON	48
3.1 Poisson Dağılımı.....	50
3.2 Poisson Regresyon Modeli.....	52
3.2.1 Modelin En Çok Benzerlik Tahminçileri.....	57
3.2.2 Modelin Artıkları.....	61
3.2.3 Aşırı Yayılımın Test Edilmesi.....	63
3.3 Poisson-Gamma Regresyon Modeli.....	65
3.3.1 Yarı Benzerlik Tahminçileri.....	68
3.3.2 EM Algoritması.....	70
3.3.3 Gözlemlenemeyen Heterojenlik.....	72
3.4 Tesadüfi Etkiler Poisson-Gamma Regresyon Modeli.....	80
3.4.1 En Çok Benzerlik Tahminçisi.....	81
3.4.2 Yarı Benzerlik Tahminçisi.....	82

3.5 Modelde Yer Alan Parametre Tahmincilerinin ve Modelin Uyum İyiliğinin Sınanması.....	83
3.5.1 Parametre Tahmincilerinin Sınanması.....	84
3.5.1.1 t-Testi Sınanması.....	84
3.5.1.2 Wald Testi.....	85
3.5.1.3 Benzerlik Oran Testi.....	86
3.5.2 Modelin Uyum İyiliğinin Sınanması.....	87
3.5.2.1 Pearson χ^2 Sınanması.....	89
3.5.2.2 Sapmalardaki Farklılıklar ve Model χ^2 Sınanması.....	89
3.5.2.3 Sahte-R ² Sınanması.....	90
3.5.2.4 Bayes Bilgi Ölçütü.....	92
3.5.2.5 Akaiki Bilgi Ölçütü.....	94
4.ÇANAKKALE'DE PROFESYONEL YÖNETİCİLİK SÜRESİNİ ETKİLEYEN FAKTÖRLER	96
4.1 Teorik Çerçeve.....	96
4.2 Araştırmanın Kapsam ve Kısıtları.....	103
4.3 Yöneticilik Süresine Etki eden Faktörler.....	107
4.4 Parametrelerin Tahmini.....	112
5. SONUÇ	121
EKLER	124
<i>EK 1. Soru Formu</i>	125
KAYNAKÇA	128

1. GİRİŞ

Hızla gelişen bilgi çağında verileri toplamak ve bunlardan sonuç çıkarmak araştırmacıların hedeflerinden biridir. Araştırmalarda kullanılan veriler toplama türlerine ve içerdikleri ölçme düzeylerine göre bir takım sınıflandırmalara tabi tutulurlar. Bu sınıflandırmalardan biri; kesit verileri (cross-sectional data) ve uzun kesit verileri (longitudinal data) şeklinde yapılabilir.

Uzun kesit verileri, araştırmadaki birimleri bir zaman aralığında incelenmiş olmasından dolayı kesit verilerine göre daha geniş kapsamlıdır. Uzun kesit verilerinin kendi içinde bilgilendirme düzeylerinde de bir takım farklılaşmalar görülmektedir. Sosyal bilimlerde sıklıkla, uzun kesit veri türlerinden biri olan olay zamanı (event history) verileri kullanılmaktadır. Bu veri türü bilgilendirici bir niteliğe de sahiptir.

Bir çok bilim dalında kullanılan olay zamanı verileri, kullanıldıkları bilim dalındaki amaçlar doğrultusunda analizlere isimlerini vermişlerdir. İlk çıkış noktası tıp bilimindeki çalışmalardır. Bu verilerin yaşam analizi veya sağ kalım çözümlemesi (survival analysis) olarak bilinmelerinin en büyük sebebi, tıp alanındaki çalışmaların diğer bilim dallarındaki çalışmalara göre daha önce başlamış olması ve daha yaygın görülmesidir. Mühendislik biliminde bu verilerle yapılan analizler başarısızlık zamanı analizi (failure time analysis) olarak isim alırken sosyal bilimlerde ise olay zamanı analizi (event-history analysis) olarak isimlendirilmiştir. Tıp biliminde önem verilen konu, kişilerin yaşamlarına devam edip etmeyeceğidir. Mühendislikte ise inceleme konusu makine ve makine parçaları olduğu için buradaki çözümlenmesi istenilen durum başarı yada başarısızlık olarak ifade edilmiştir. Sosyal bilimciler için önemli olan ise olaylardır. Olayların ortaya çıkış sebepleri veya ortaya çıkmış olayların bitiş sebepleri bu alanda çalışanların merak konusu olmuştur. Bu nedenle olay zamanı analizi olarak isimlendirilmiştir.

Olay zamanı verilerinin, sosyal bilimlerdeki kullanımı hızla artmaktadır. Yapılan araştırmalardaki veri yapısı, olay zamanı veri yapısına uyum göstermektedir. Olay zamanı analizi diğer analizlerde olmayan özelliklere sahiptir. Bazı gözlemler

arařtırmacının isteęi dıřında gözlem süresi bitmeden izlemeden ıkabilir. Bu tür gözlemler durdurulmuř veri (censored data) olarak isimlendirilmektedir. Olay zamanı dıřındaki analiz yöntemleri, durdurulmuř verileri eksik veri olarak düşünmektedir. Durdurulmuř veriler çoęu zamanda dięer yöntemlerde analiz dıřında kalmaktadır. Olay zamanı analizinin bir bařka özellięi ise, zamana göre farklılařan deęiřkenleri bünyesinde barındırmasıdır. Bazı deęiřkenler incelenen gözlem aralıęında deęiřimlere uğrarlar. Medeni durum deęiřkeni bu tür deęiřkenlerin en bilinen örneęidir. Kiřilerin arařtırmanın bařladıęı noktadaki medeni durumları ile bittięindeki medeni durumları aynı kalmayabilir. Deęiřkenlerin zaman iindeki bu tür farklılařmalarına izin veren analiz olarak, olay zamanı analizi dięer yöntemlere göre tercih sebebi olmaktadır.

Olay zamanı analizinde yakın zamana kadar göz ardı edilen bir konu gözlemlenemeyen heterojenlik (unobserved heterogeneity) kavramıdır. Bu kavram arařtırmacılar tarafından gizli heterojenlik (hidden heterogeneity) olarak da ifade edilmektedir. Bazı açıklayıcı deęiřkenler, ölçümlerinde yařanan sorunlar yüzünden ihmal edilebilirler veya aslında ok önemli olan bir açıklayıcı deęiřken modele dahil edilmeyebilir. Bunlar ve daha sonra detaylarıyla açıklanacak olan bir ok neden gözlemlenemeyen heterojenlięin temel kaynakları olarak karřımıza ıkmaktadır. Farklı nedenlerden dolayı modele alınmayan açıklayıcı deęiřkenler yanlı parametre tahmincilerine sebep olmaktadır. Bunun sonucunda da hazard oranı gereęinden farklı bir řekilde ortaya ıkmaktadır.

Gözlemlenemeyen heterojenlik kavramı ok uzun zaman önce bilinmesine raęmen, kullanımı son yıllara kadar görülmemiřtir. Karmařık bir yapıya sahip olmaları ve bilgisayar paket programlarında analiz yer almaması kullanımı geciktiren etmenlerdendir. Son yıllardaki yapılan alıřmalarda gözlemlenemeyen heterojenlięi kontrol etmeye yarayan yöntemler geliřtirilmiřtir. Sabit etki yöntemleri ve tesadüfi etki yöntemleri bařlıkları altında incelenen yöntemler bir sorun olarak görülen gözlemlenemeyen heterojenlięi kontrol etmeye alıřmaktadır. Sabit etki yöntemi zamana göre farklılařan deęiřkenleri iermemesi ve bir takım varsayımlara baęlı kalmasından dolayı tesadüfi etki yöntemine göre daha pasif kalmaktadır.

Yaşayan tüm varlıkların süreç içerisinde yaşama devam edip etmeyecekleri ne kadar önemliyse, işletmelerde ayakta kalmaları gereken birer varlık olarak düşünüldüğünde, yaşamlarını sürdürmeleri bir o kadar önemlidir. İşletmelerin yaşamlarını sürdürmelerini sağlayan bir takım etmenler vardır. Bunların rekabetçi çevre koşullarında varlıklarına devam etmemesi veya işletme politikalarındaki yanlışlar işletmelerin yaşamlarının bitmesine neden olabilir. İflas yada işletme başarısızlığı ile isimlendirilen bu durum sonuçları açısından işletme paydaşları tarafından hiç istenilmeyen bir durumdur.

Organizasyon başarısızlıkları teorik olarak bir çok açıdan değerlendirilseler bile analiz aşamasında ölçümlerden kaynaklanan sorunlar yaşanmaktadır. Bugün konunun incelenmesinde birçok farklı modelin olmasına rağmen gerçekçi bir şekilde başarısızlıkları ortaya koyan modeller geliştirilememiştir. Birçok organizasyon, finans ve maliyet değerleri iyi olmasına rağmen başarısızlıklarla karşı karşıya kalmaktadır. Bu da konunun daha derin ve detaylı değişkenler açısından ele alınmasını gerekli kılmaktadır. Organizasyonlar açısından bir çok faktör ölçülemedikleri için veya ölçümleri en az aralıklı ölçekle sağlanamadığı için analize dahil edilememektedir. Bu çerçevede kullanılacak modelin organizasyon başarısızlıklarına farklı boyutlardan bakabilme olanağı sağlayacak bir modelin geliştirilmesi gerekliliğini ortaya çıkarmaktadır.. Böylece paydaşlar ve yöneticiler açısından başarısızlık sürecinin incelenmesinde etkili olan ancak dikkate alınamayan unsurların etkisi ortaya konarak, karar süreçlerinin etkinleştirilmesi sağlanmış olacaktır.

Organizasyonlarda yönetim teorisi, organizasyonların rekabetçi ortamda ayakta kalmalarını sağlayacak, değişen piyasa koşullarını takip etmesine önderlik edecek ve sonuç olarak başarılı organizasyon olma yolunda yapılması gereken her türlü adımı içinde barındırmaktadır. Bu anlamda organizasyonların tepe yöneticileri zorlu bir görevi üstlenmektedirler.

Organizasyonlarda yönetim değişimi bünyesinde bir çok neden barındırmaktadır. Tepe yöneticilerin değişiminin organizasyon üzerinde farklı etkileri bulunduğu kaçınılmazdır. Bu kapsamda yapılacak olan bir araştırmada açıklayıcı değişkenlerin ölçümlerinden kaynaklı sorunlar gözlemlenemeyen heterojenlik teorisi

çerçevesinde incelendiğinde karma modeller gündeme gelmektedir. Tepe yöneticilerin değişiminin Poisson Regresyon Analizi ile modellenmesi mümkün olmaktadır. Poisson regresyon analizi sonuçlarında aşırı yayılım ile karşılaşılması, gözlemlenemeyen heterojenlik kavramının göz önünde bulundurulması gerektiğini göstermektedir. Modelin gözlemlenemeyen heterojenlik barındırması durumunda ise Tesadüfi Etkiler Poisson-Gamma Regresyon Analizinin kullanımını ön plana çıkarmaktadır.

Çalışma dört bölümden oluşmaktadır. İlk bölüm giriş bölümüdür. İkinci bölümünde olay zamanı verileri ve kullanılan analiz türleri ile ilgili ayrıntılı olarak literatür taraması yapılmıştır. Üçüncü bölümde olay zamanı analizinde kullanılan poisson regresyon modeli, poisson regresyon modelinde aşırı yayılımla karşılaşıldığında kullanılan poisson-gamma regresyon modeli ve olay zamanı verilerinde gözlemlenemeyen heterojenlikle karşılaşıldığında kullanılan tesadüfi etkiler poisson-gamma regresyon modeli incelenmiştir. Çalışmanın son bölümünde Tesadüfi Etkiler Poisson-Gamma Regresyon Analizi kullanılarak işletmelerin tepe yöneticilerin değişiminde ortaya çıkan gözlemlenemeyen heterojenlik kavramı kontrol edilmeye çalışılmıştır. Tesadüfi Etkiler Poisson-Gamma Regresyondan elde edilen sonuçlar, Poisson-Gamma Regresyon Analizi ile elde edilen sonuçlarla karşılaştırılarak gözlemlenemeyen heterojenlik kavramına çözüm getirilmeye çalışılacaktır.

2. OLAY ZAMANI VERİLERİNİN KAPSAMI VE KULLANILAN ANALİZ YÖNTEMLERİ

Araştırmacılar, araştırma sorusuna cevap aramak için bireyler yada olgularla ilgili olarak belirli bir zaman noktasında çalışma yapmak isteyebilirler. Bir araştırmacı, belediye hizmetlerinin belediye başkanıyla olan etkisini inceleyebilir. Toplayacağı veri, yönetim değişmeden önce verilen hizmetten memnuniyet ve değiştikten sonra verilen hizmetten memnuniyet olacaktır. Veriler, zaman içerisindeki iki noktadan toplanmaktadır. Yapılan çalışma kesit (cross sectional, one shut) fakat süreçsel olarak zaman içerisinde yapılmaktadır. Çalışmadaki bağımlı değişkenlere ait veriler zaman içerisindeki iki yada daha fazla noktadan elde ediliyorsa ve araştırma sorusuna cevap aranıyorsa toplanan veri tipi uzun kesit verisi (longitudinal data) olarak isimlendirilir¹.

Uzun kesit çalışmalarının temel özelliği tanımlandığında, bir bireyin zaman içinde tekrarlanan durumunun ölçümünü ifade etmektedir. Uzun kesit çalışmalarının tam tersi olan kesit verisi çalışmaların da ise bir kez ölçüm gerçekleşir. Kesit verisi analizinde ve uzun kesit verisi analizinde en temel sorun kohort etkisidir. Kohort araştırmalarında, araştırmanın yönü sebepten sonuca doğrudur. Bir başka ifade ile, araştırma grupları bağımsız değişkene göre oluşturulur. Araştırma esnasında bağımsız değişkenle karşılaşan grup “kohort grubu” (araştırma kohortu) olarak isimlendirilebilir². Uzun kesit veri analizi, zaman içerisinde ortaya çıkan değişim ve farklılaşmaları çok net bir biçimde ortaya koymaktadır. Bazı çalışmalarda üçüncü zaman ölçeği bu periyot yada takvim tarihinin ölçümünde önem taşımaktadır. Herhangi iki yaş periyodu yada kohortu üçüncüyü belirlemektedir. Örneğin, bireysel yaş ve doğum kohortu ölçülen durumun tarihini bize vermektedir.

Uzun kesit veriler geleceğe yönelik yada geçmişe yönelik toplanabilir. Uzun kesit veriler daha çok geleceğe yönelik olarak toplanır. Çünkü geçmiş kayıtlardan elde

¹ Uma Sekaran, **Research Methods for Business a Skill-Building Approach**. New York: John Willey ve Sons, inc, 2000, s. 139.

² Banu Çakır, Tıbbi Araştırma teknikleri: Temel Özellikler, Yarar ve Kısıtlılar, Olası Hata Kaynakları, http://www.toraks.org.tr/kisokulu3-ppt-pdf/Banu_cakir.pdf, (2 Eylül 2002).

edilen veriler genelde sağlıklı değildir. Uzun kesit veri setleri, bir değişken üzerinde zaman içerisinde farklılıkları değerlendirmemize olanak sağlamakta ve bu model özel istatistiksel teknikler gerektirmektedir. Çünkü gözlemlenen konudaki olaylar arasında korelasyon vardır. Bunun bilimsel olarak geçerli yöntemlerle açıklanması gerekir. Bu korelasyon konusunun ortaya çıkması konusunda zaman serilerinin ölçümü karşımıza çıkmaktadır.

Uzun kesit verileri özellikle klinik deneylerde yoğun olarak kullanılan bir yöntemdir. Tıp bilimindeki araştırmalarda bir hastalıktaki tedavi etkisinin ortaya çıkarılması ile yapılacak olan bir çalışmada, birkaç ölçüm sonucunun birey üzerindeki etkisi ancak zaman içerisinde ortaya konulabilir. Sosyal bilimlerde ise özellikle davranış değişimi yada bireyin zaman içerisinde karşı karşıya kaldığı değişim süreciyle ilgilenilmektedir. Uzun kesit veri yapısı sosyal olayların değerlendirilmesi sırasında değişim süreçlerinin analiz edilmesinde büyük bir kolaylık sağlar³.

Uzun kesit verileri kapsadıkları bilgilendirme özelliğine bağlı olarak, zaman serisi, panel veri ve olay zamanı (event-history) verisi olmak üzere üç temel sınıflandırılma ile ifade edilebilir⁴.

Uzun kesit verilerinin en az bilgi verici tipi, zaman serisi veya trend bilgisidir. Zaman serilerinde aynı tip veriler farklı zaman noktalarında toplanabilir. Uzun kesit verileri kesit verilerinin diğer tiplerinden farklı olarak, zamanın herhangi bir noktasında toplanan birimleri farklı fakat veri yapısı aynıdır. Sonuç olarak; birimler üzerindeki değişiklikler, değişimler veya olaylar hakkında hiç bir bilgi olmamasına rağmen çalışılan anakütledeki bütün değişiklikler incelenmektedir.

Panel veri, aynı zamandaki farklı noktalarda aynı birimlerin incelenmesiyle elde edilir. Bir olaydaki değişken değerinin, önceki olay üzerindeki değerine eşit olup olmadığı gözlem yapmanın amacını oluşturmaktadır. Bu değer eşit olmadığı durumda en az bir değişim veya deneyimle incelenen birim ilişkilendirilirse bu durumda olayın

³ Peter J. Diggle, Kung- Yee Liang ve Scott L. Zger, **Analysis of Longitudinal Data**, Oxford: Oxford University Press, 1994, s. 1.

Geert Verbeke, **Linear Mixed Models for Longitudinal Data**, New York: Springer-Verlag, 2000, s. 1-3.

⁴ Jeroen K. Vermunt, **Log Linear Models for Event Histories**, California: Sage Publication, 1997, s. 2.

gerçekleşmesi kesindir. Diğer bir deyişle, eğer iki değer eşit değilse incelenen olay gerçekleşebilir veya gerçekleşmeyebilir. Sonuç olarak; panel veri net değişimler üzerinde bilgi sağlamasa da brüt veya birim değişimleri üzerinde kısmi bilgi sağlar.

Olay zamanı verileri, diğer veri tiplerine göre daha da fazla bilgi vericidir. Çünkü bu tip veriler değişiklik, değişim veya olay zamanlamasına ilişkin bilgi içerir. Diğer bir deyişle, olay zamanı verisi olay sayısını, onların ardışıklıklarını ve meydana gelme zamanlarını kayıt eder. Bu durum daha çok konularını sürekli olarak güncelleyen deneysel çalışmalarda, gözlem yapılmasına olanak tanıdığı için tercih edilmektedir. Yapılan bir çalışmada, olayın zamanlanmasının kayıt edilmesi özellikle deneysel araştırmalar için büyük bir avantaj sağlamaktadır. Deneysel olmayan çalışmalarda olay zamanı verilerini toplamak hayli zordur. Verilerin toplanmasındaki ilk yol, arşivlerden yararlanmaktır. Düzenli olarak tutulmuş arşiv verisi bulunamazsa diğer bir yol ise, olaya şahit olmuş kişilerden geçmişini hatırlamaları istenerek veri toplanabilir.

Diğer bilimlerde olduğu gibi sosyal bilimlerin her alanında da olaylar ve onların sebepleri arasında önemli ilişkiler vardır. Medikal sosyologlar, hastaneye yatırımlar ve hekim viziteleriyle meşgul olurlar. İş ve kariyer çalışmalarında dikkat; meslek değişimleri, promosyon, geçici iş bırakımı ve emeklilik üzerindedir. Politik bilimlerde ise isyanlar, devrimler ve sakin hükümet değişimleri ile ilgilenir. Demografi bilimcileri doğumlar, ölümler, evlenme, boşanma ve göçler üzerine odaklanır. Farklı alanlarda da olsa incelenecek olan her konuda bazı niteliksel değişimleri içeren bir olay, zamanın herhangi bir noktasında meydana gelmektedir. Bazı nitel değişkenlerdeki kademeli olarak değişimleri tanımlayan “olay” dönemleri kullanılmayabilir. Değişimler, hem olay anındaki izleme hem de olayın tamamlandığı süreçten sonraki zamanda izleme arasında nispeten keskin ayrımlardan meydana gelmek zorundadır. Çünkü olaylar, incelenen sürecin bitişinden sonraki zamandaki (overtime) değişim dönemlerinde tanımlanmaktadır. Bu durum gitgide olayı çalışmanın en iyi yolu olarak kabul edilmektedir. Araştırmacının amacı olay zamanı verisi toplamak olarak belirginleşmektedir. Anketörün, cevaplayıcıya evlilik zamanlarını sorduğu bir araştırmada, eğer araştırma konusu olay sebebiyse, olay zamanı mümkün açıklayıcı

değişkenler üzerindeki verileri kapsar. Bu değişkenlerden bazıları ırk gibi olay bittikten sonraki zamanda sabit kalabilir veya gelir gibi değişebilir⁵.

2.1 Olay Zamanı Analizinin Kullanım Amacı

Olay zamanı veri analizi, sosyal bilimlerdeki dinamiklerin belirlenmesini sağlayan araştırmalarda yaygın olarak kullanılan bir yaklaşımdır. Olay zamanı verilerinin temel kullanım amacı; doğumlar, ölümler, evlilikler, suçlular veya suçlu kanunları, işgücü girişi, iş terfileri, örgüt kurulması ve dağılması gibi süreksiz durumlar arasında yapılan bireysel geçişlerdeki zaman üzerine bilgi sağlamaktır. Bu gibi veri analizlerinin merkezi her bir durumdaki harcanan sürenin ne olduğunun belirlenmesi ve bir sonraki duruma girildiğinde ne olduğunun belirlenmesidir. Araştırmacılar, özellikle bir sonraki duruma girişte ve bulunulan durum için harcanan süredeki açıklayıcı değişken etkilerini araştırırlar⁶.

Olay zamanı analizi, sadece sosyal bilimlerde kullanılan bir analiz değildir. Diğer bilim dallarında da kullanılmaktadır. Bu analizin kullanımı tıp ve mühendislik bilim dallarında da sıkça görülebilir. Bu durum büyük bir karmaşanın doğmasına sebep olmaktadır. Çünkü aynı görüş ve düşünceler özellikle bu üç bilim dalında çoğunlukla farklı sözcüklerle ifade edilmektedir⁷.

Tıpta olay zamanı analizi kullanımı, yaşam analizi (survival analysis) olarak isimlendirilmiştir. Biyoistatistikçilerin incelenen olaydaki en önemli ilgisi, ölüm üzerine odaklanmış olmalarından kaynaklanmaktadır. Hastalığın tedavisi sürecindeki hastaların tümü iyileşmeyi beklemektedirler. Fakat varılan son nokta, ilgilenilen hastalıktan kaynaklı olarak hastanın yaşamının sona ermesidir.

⁵ Paul D. Allison, **Event History Analysis Regression for Longitudinal Event Data**, California: Sage Publication, 1984, s. 9.

⁶Kenneth C. Land, Daniel S. Nagin ve Patricia L. McCall, “Discrete Time Hazard Regression Models with Hidden Heterogeneity The Semiparametric Mixed Poisson Regression Approach”, **Sociological Methods&Research**, No: 29, S: 3 , Şubat 2001, s. 342-343.

⁷ Harvey Goldstein, **Multilevel Statistical Models**, London: Hodder Arnold Publication, 2003, s. 6.

Mühendislikte olay zamanı analizi, endüstriyel makinelerin ve elektronik aletlerin bozulmasında kullanılmıştır. Bu yüzdendir ki, başarısızlık zamanı analizi (failure time analysis) olarak isimlendirilmektedir. İlgilenilen olayın meydana gelme zamanı, güvenilirlik (reliability) ve başarısızlık zamanı (failure time) olarak ifade edilmiştir⁸.

Olay zamanı analizinin en iyi bilinen şekli günümüzde de kullanılmakta olan yaşam tablosu analizidir. Bu analiz, 18. yüzyıldan beri kullanılmasına rağmen tüm olay tiplerini kapsayamadığı için yetersiz kalmıştır. Olay zamanı modellerinin yapısı, yaşam tabloları analizinin genişlemesi üzerine temellendirilmiştir. Yaşam tabloları analizi; bir anakütledeki bireylerin yaşam oranı (survival rate) olarak isimlendirdikleri, yaşam beklentilerini hesaplayan demograficiler tarafından yaygın olarak kullanılmaktadır⁹. Yaşam tablosu analizi regresyon modeli içermemesinden ötürü yeni yöntemlerin gelişmesine olanak sağlamıştır. 1972’ de Cox tarafından geliştirilen kısmi benzerlik yöntemi bir çok temel fikri geliştirmede yaşam tablosu yönteminin önüne geçmiştir. Mühendislik ve tıp bilimleri birbirlerini takip halinde olduklarından son yıllarda birbirleriyle entegre olmuşlardır.

Sosyal bilimciler ise bu iki bilimin iletişiminden habersizdirler. Markov sürecinin 1960’lı yılların sonunda ve 1970’li yılların başında kullanıldığı görülmektedir. Biyoistatistikçiler bu modeli önceden beri kullanıyor olmalarına rağmen bu yenilik, sosyal bilimler ve tıp arasında köprü görevini görmektedir. Bu durum, çok uzun süre habersiz kalınmasına rağmen tıp, mühendislik ve sosyal bilimler arasındaki ilişkilerin gelişmesine ve entelektüel ortaklığın ortaya çıkmasını sağlamıştır¹⁰.

“Sosyal bilimlerde olay zamanı verileriyle ilgili bir problem, olaydan ayrılanların olayla ilişkili olabileceğidir. Örneğin, boşanmaları tamamlanmadan bireyler, çalışmadan ayrılmak eğilimde olabilirler. Bu durum boşanma oranları tahminlerine çok zarar vermektedir. Problem ilk bakışta hemen hemen açıklanamaz gibi görünmektedir fakat Hill (1997), Hill, Axin ve Thornton’un (1993) paylaşılan

⁸ Allison, s. 11.

⁹David Rose ve Oriel Sullian, **Introducing Data Analysis for Social Scientists**, Buckingham: Open Uni. Press, 1998, s. 219-220.

¹⁰Allison, s. 12.

ölçülmemiş risk faktörü kullanımını, önemli bir çözüm olarak ortaya koymuştur. Temel özellik gözlem yapmak olduğundan ayrılan insanların çok yakında boşanacakları bilinmemesine rağmen, kimlerin yüksek boşanma riski altında oldukları tahmin edilebilir”¹¹. Olaydan ayrılan bireylerin, çalışma henüz bitmeden gözlem dışı kalmaları veri kaybına sebep olmaktadır. Bu durum ise durdurma (censoring) olarak isimlendirilmektedir.

Politik bilimlerde de olay zamanı analizinin kullanımı hızla yaygınlaşmaktadır. Politik olayların önemli bir özelliği tekrarlanabilir olmalarıdır. Olay zamanı veri analizinin tekrarlanan politik olaylarda modellenmesi Box-Steffensmeier ve Zorn¹² tarafından incelenmiştir.

Sosyoloji alanında da olay zamanı analizi, olayların incelenmesinde farklı bir bakış açısı getirmiştir. Yoder, Whitbeck ve Hoyt¹³ ilk kez evden kaçan gençleri ve ilk kez sokakta yaşayan gençleri iki ayrı olay zamanı modeli geliştirerek incelemiştir. Geliştirdikleri iki süreksiz zaman olay zamanı modeli üç farklı tip açıklayıcı değişken kullanılarak analiz edilmiştir. Bunlardan ilki zaman ölçümleridir. İlk kez evden kaçış yaşı bu tür bir değişken olarak modele katılmıştır. İkinci tür değişken; aile yapısındaki değişimlerin kümülatif sayısı gibi, zamana göre farklılaşan değişkenlerdir. Cinsiyet ve ırk gibi değişkenlerde zamana göre farklılaşmayan açıklayıcı değişken olarak modelde yer almışlardır. Detang ve Molho¹⁴, bireylerin okuldan mezun olduktan sonra iş bulmak için göç etmeleri veya ikametlerine devam etmelerini kadın-erkek karşılaştırması üzerinde odaklanarak, yaptıkları çalışma ile ortaya koymuşlardır. Garasky, Haurin ve Haurin¹⁵ gençlerin yetişkinliğe geçiş aşamasında ailesi ile beraber veya yalnız yaşamak yerine gençlerden oluşan bir grup içinde yaşama isteklerini olay zamanı analizi kullanarak incelemiştir.

¹¹ Adrian E. Raftery, “*Statistics in Sociology, 1950-2000: A Selective Review*”, **Sociological Methodology**, No: 31, S: 1, Ocak 2001, s. 14-15.

¹² Janet M. Box-Steffensmeier ve Christopher Zorn, “*Duration for Repeated Events*”, **The Journal of Politics**, No: 64, S: 4, Kasım 2002, s. 1069.

¹³ Kevin A. Yoder, Les B. Whitbeck ve Dan R. Hoyt, “*Event History Analysis of Antecedents to Running Away From Home and Being on the Street*”, **The American Behavioral Scientist**, No: 45, S: 1, 2001, s. 55-56.

¹⁴ Cecile Détang-Dessendre ve Ian Molho, “*Residence Spells and Migration: A Comparison for Men and Women*”, **Urban Studies**, No: 37, S: 2, Şubat 2000, s. 251.

¹⁵ Steven Garasky, R Jean Haurin ve Donald R. Haurin, “*Group Living as Youths Transition to Adulthood*”, **Journal of Population Economics**, No: 14, 2001, s. 329-331.

Eđitim alanında yapılan alıřmalarda da olay zamanı analizinin kullanıldıđı grlmektedir. Chizmar¹⁶, ekonomi eđitimi alan niversite đrencilerinin blme devamlılıđını sreksiz olay zamanı analizi kullanarak incelemiřtir.

Hazard regresyon analizi olay zamanı analizinde kullanılan bir tekniktir. Evlilik gibi medeni durumdaki deđiřimler, ocuk dođurma bariteleri, iřgc aktiviteleri, eđitim becerisini geliřtirme, eř seimi, hastalık riski ve sađlık kontrol gibi bireylerin temsil ettiđi davranıř tiplerini belirlemeye ynelik alıřmalar iin alıřılmıř bir aratır. Hazard fonksiyonu ile deđiřimlerin meydana geldiđi zamandaki deđiřim llmektedir¹⁷.

Bir olay belli bir zaman iinde meydana gelirken bir ok faktre bađlı olarak gerekleřmektedir. Bu durumda her bir faktrn greceli katsayısını gsteren katsayılar tahmin edilmeye alıřılacaktır. Modellerin bir ođu tahmin yapabilmek iin, en ok benzerlik tahmincisini kullanmaktadır. Bu sre log-lineer analize benzetilebilir. Log-lineer analizde mmkn olduđu lde gzlemlenen hcre frekanslar kmesi yaratılmaktadır. En ařırı u katsayı kombinasyonunu kullanarak apraz tablolardan gzlemlenen hcre frekanslar kmesi yaratmak log-lineer analizin amacını oluřturmaktadır. Herhangi bir model iin gzlemlenen frekansları retmeye en fazla katkısı olan bu katsayılar, aıklayıcı faktrler olarak en nemli deđiřkenleri temsil ederler. Olay zamanı analizi log-lineer analize benzetmekle beraber ayırt edici bir farka sahiptir. Olay zamanı analizinde bađımlı deđiřken apraz tablolařmadaki frekanslar kmesi deđildir. Tam tersine belli bir zaman iinde meydana gelme olasılıđıdır¹⁸. Olay zamanı analizinde kullanılacak modeller bu zelliđi dikkate alarak oluřturulmuřtur. Ayrıca olay zamanları verilerinde karřılařılan zel bir durum da verilerin durdurma ve kesme (truncated) zelliklerine sahip olmasıdır.

Olay zamanları, olay sebeplerinin incelenmesinde ideal olmasına rađmen durdurma ve kesme gibi zellikleri barındırmasından tr standart istatistiksel

¹⁶ John F Chizmar, "A Discrete-Time Hazard Analysis of Role of Gender in Persistence in Economics Major", **Journal of Economic Education**, No: 2, 2000, s. 108.

¹⁷M. Jan Hoem, "Systematic Patterns of 2000 Exposures in Event-History Analysis", **Sociological Methodology**, No: 30, 2000, s. 38.

¹⁸Rose ve Sullian, s. 220.

tekniklerin kullanımına izin vermemektedirler. Bilinen çoklu regresyon modelleri, olay zamanı verilerinin bu iki özelliğinden dolayı, kullanıldığında problem yaratmaktadır.

Olay zamanı analizinde kullanılan regresyon modellerini standart regresyon modellerinden ayıran diğer bir önemli nokta ise, içerdikleri açıklayıcı değişken yapısıdır. Çoklu regresyon modellerinde açıklayıcı değişkenler tek tip olarak denkleme katılmaktadır. Zaman serileri analizinde ise çoklu regresyon modelindeki açıklayıcı değişkenlere zaman unsurunu da açıklayıcı bir değişken olarak ekleyip denklem oluşturulmaktadır. Olay zamanı analizi bu açıklayıcı değişkenleri içermenin yanında diğer analizlerde kullanılmayan zamana göre farklılaşan değişkenleri de (time-varying variables) analize dahil etmektedir. Durdurma veya kesme ile zamana göre farklılaşan değişkenleri modele dahil etmesi olay zamanı analizinin etkin olarak kullanılmasındaki en önemli iki sebebi oluşturmaktadır.

2.2 Olay Zamanı Analizi Veri Türü

Olay zamanı analizinin diğer analizlerden tercih edilmesini sağlayan temel nedenlerinden biri, içerdiği veri yapısı olarak ifade edilebilir. Olay zamanı verileri için durdurma ve kesme özellikleri, diğer analizlerde kullanılmayacak olan ve hatalı olarak görülen verilerin analize katılmalarına imkan tanımaktadır.

Durdurma, araştırmacının isteği dışındaki sebeplerden dolayı gözlemlerin izlenememesi olarak ifade edilebilir. “I. tip durdurma” (type I censoring), “II. Tip durdurma” (type II censoring), “tesadüfi durdurma” (random censoring), “sağdan durdurma” (right censoring), “soldan durdurma” (left censoring) ve “aralıklı durdurma” (interval censoring) olmak üzere altı durdurma türü bulunmaktadır.

Kesme, arařtırmacı tarafından belirlenen bir zaman aralıęında gözlemlenen olaylar olarak ifade edilebilir. Kesme, soldan kesme (left truncated) ve saędan kesme (right truncated) olmak üzere iki farklı biçimde ifade edilebilir¹⁹.

2.2.1 Durdurma

Arařtırmada kullanılacak olan veri setinin incelenmesindeki zorluklardan bir tanesi, bazı örneklerin arařtırma süresi boyunca başarısızlıklarının gözlemlenememesidir. İncelenen duruma maruz kalan birimler arařtırmacının planlamadıęı bir şekilde gözlemden çıkabilirler. Dięer analizlerde problem olarak karřılařılan bu durum, olay zamanı analizine dahil edilen bir konu olmaktadır. Olay zamanı analizinin bu özellięi içermesi ile dięer analizlerden ayırt edici bir avantaj saęladıęı söylenebilir.

Olay zamanı analizinin kullanıldıęı tüm bilim dallarında durdurulmuş verilerle karřılařmak mümkündür. Endüstriyel alandaki güvenilirlik çalıřmalarına bakıldıęında, yařam testleri deneylerinin, bütün öęeleri başarısız olmayabilir. Tıp alanındaki çalıřmalara baktıęımızda, klinik uygulamalar sonunda bir çok hasta yařamayı ümit etmesine raęmen, bazı hastaların farklı bir sebepten dolayı yařamı sona erebilir. Örneęin, akcięer kanseri ile ilgili klinik bir çalıřmada hastanın kalp krizinden yařamının sona ermesi akcięer kanserinden ölmesine sebep deęildir. Ölüm sebebi arařtırmanın dışında bir durumdur²⁰.

Sosyal bilimlerde yapılan arařtırmalarda da olay zamanı modelleri, arařtırmadaki durdurulmuş gözlemler (censored observation) hakkındaki bilgileri kullanabilmektedir. Bir arařtırmada ilgilenilen olay üzerinde bireyler bu olayı tecrübe etmemiřlerse yada bu olay bu kiřilerin bařından geçmemiřse, gözlem periyodu

¹⁹ Guang Guo, “Event-history Analysis for Left Truncated Data”, **Sociological Methodology**, No: 23, 1993, s. 217-222.

²⁰ Laszlo Gyorfı, **Distribution-Free Theory of Nonparametric Regression**, New York: Springer-Verlag, 2002, s. 540-541.

bittiğinde durdurulmuş bir gözlem olarak değerlendirilmektedir. Örneğin evlilikteki yaşın belirlenmesi için yapılan bir çalışmada henüz evlenmemiş bir çok kişi bulunabilir. Evlilik yapmaktaki yaşı tahmin etmede kullanılacak olan standart regresyon modellerinde bu hiç evlenmemiş kişiler hatalı veri olarak değerlendirilir ve çalışmanın dışında tutulurlar. Bu verilerin silinmesi yanlış sonuçlar doğurmaktadır. Çünkü şu an evli olmayan kişiler daha sonraki yaşlarında evlilik yapacaklardır. Olay zamanı analizi standart regresyon analizi metotlarından bu noktada farklılık göstermektedirler²¹.

Gözlemlenemeyen veya tamamlanamayan olayın başarısızlık zamanı durdurma olarak isimlendirilir. Bir olayın başarısızlığı ile ilgili bir araştırma da, bu olay için başarısızlık bir nokta olarak düşünülürse aynı şekilde durdurmada, bir nokta olaydır. Durdurulmuş bireyler için gözlem periyodu kayıt edilmektedir²².

Araştırma yapılan konuya ve birimlerin türüne göre, durdurma, farklı çeşitlerde ortaya çıkmaktadır. “I. tip durdurma”, “II. Tip durdurma”, “tesadüfi durdurma”, “sağdan durdurma”, “soldan durdurma” ve “aralıklı durdurma” olmak üzere durdurma türlerini ifade etmek mümkündür²³.

2.2.1.1 I. Tip Durdurma

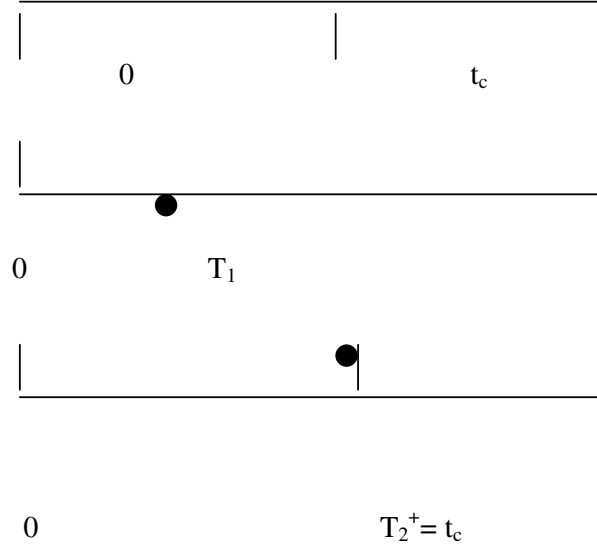
I. tip durdurma, daha çok mühendislik uygulamalarında meydana gelmektedir. Örneğin, Transistor, borular ve ciplerin dayanıklılıklarının incelendiği bir kalite kontrol araştırmasında $t=0$ zamanında başarısızlık zamanları kayıt edilmeye başlanmıştır. Bazı birimlerin tamamen yok olduğunu görmek uzun zaman alabilir. Araştırmacı bu kadar uzun zaman bekleyemeyebilir. Bu durumda önceden belirtilmiş t_c zamanında deneme

²¹ Tim B Heaton ve Vaughin R.A. Call, “Modelling Family Dynamics With Event History Techniques”, **Journal of Marriage and Family**, S: 57, Kasım 1995, s. 1078.

²² D. R. Cox ve D. Oakes, **Analysis of Survival Data**, London: Chapman and Hall, 1984, s. 4.

²³ M. Tableman ve J.S. Kim, **Survival Analysis Using S Analysis of Time-to-Event Data**, New York: Chapman and Hall, 2004, s. 9.

sonlandırılır. Başarısızlık zamanlarının gözlem sayısı tesadüfidir. Teste tabi olan birimlerin sayısı eğer n ise, n adet farklı başarısızlık zamanı gözlemlenebilir²⁴.



Şekil 1. I. Tip Durdurma

Kaynak : M. Tableman ve J.S. Kim, Survival Analysis Using S Analysis of Time-to-Event Data, New York: Chapman and Hall, 2004, s. 9.

n birimlik bir örnekte i. birimin başarısızlık zamanı T_i tesadüfü değişkeni ile gösterilmektedir. Şekil 1’ de görüldüğü gibi, t_c zamanına kadar herhangi bir başarısızlık gözlemlenmemiştir. Oluşabilecek bir başka durum ise başarısızlık, hemen gözlemlenmiştir. En son durumda ise deneme sonuna gelindiğinde başarısızlık gözlemlenmiştir. Bu durum gözlemler için aşağıdaki şekilde özetlenebilir.

$$Y_i = \min(T_i, t_c) = \begin{cases} T_i & \text{eger } T_i \leq t_c \\ t_c & \text{eger } t_c < T_i \end{cases} \quad (1)$$

$T_i \leq t_c$ ise “durdurulmamış” (uncensored) yada “gerçekleşmeyen durdurma zamanı” (unrealized censoring time) olarak isimlendirilmektedir²⁵.

²⁴ Tableman ve Kim, s. 9.

²⁵ Cox ve Oakes, s. 5.

II. tip durdurmada r deki başarısızlık zaman sayısı sabit deęer son nokta ise (t_r) tesadüfi gözlemdir. Araştırmacı, r' in olası başarısızlığını gözlemlemek için uzun süre beklemek zorunda kalabilir veya tam terside söz konusu olabilir.

I. ve II tip durdurma farklı biçimlerde tasarlanmış olsalar da, gözlemlenen benzerlik fonksiyonu her iki tip durdurma içinde aynı olmaktadır²⁷.

2.2.1.3 Tesadüfi Durdurma

Tesadüfi durdurma daha çok tıp alanındaki klinik deneylerinde kullanılmaktadır. Klinik çalışmalarında hastalar farklı zamanlarda deneye tabi olmaktadır. Her tedavi için farklı terapi söz konusudur. Başarısızlık gözlemlenmek istendiğinde durdurma aşağıda belirtilmiş olan üç durumdan biri olarak karşımıza çıkmaktadır.

1. Takibin yitilmesi durumunda durdurma söz konusudur. Hasta çeşitli sebeplerden dolayı ayrılır ve bir daha görmek mümkün değildir. Bu durumda,

Yaşam zamanı \geq gözlemlenen deęer

olmaktadır.

2. Tedavinin uygulama aşamasın kötü bir zorlayıcı boyuttur. Hasta bazı nedenlerden dolayı tedaviyi ret eder ve çalışmadan ayrılır.

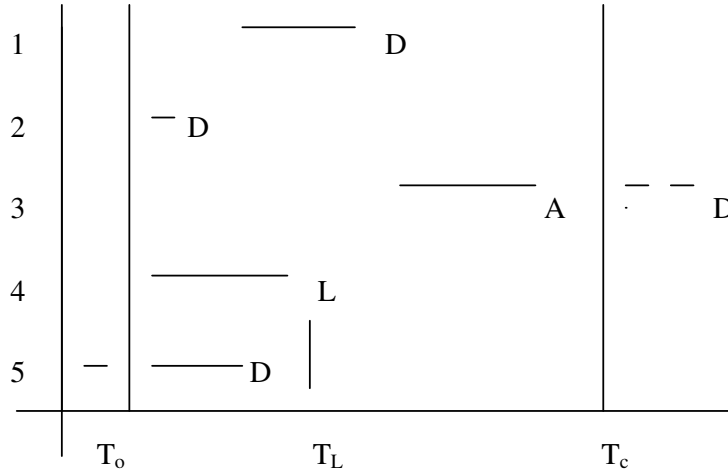
3. Çalışmanın bitişinde hasta hala yaşamını sürdürmektedir²⁸.

²⁷ Tableman ve Kim, s. 10-11.

²⁸ Tableman ve Kim, s. 12; Shoukri ve Pause, s. 1.

2.2.1.4 Sağdan Durdurma

Sağdan durdurma, daha çok mühendislik ve tıp alanlarında yapılan çalışmalarda görülmektedir. Çalışmanın incelenmek istenen zaman periyodu içerisinde, gözlem süreci başladığından bir süre sonra önceden bilinmeyen sebeplerden dolayı bazı birimler hakkında veriye ulaşılamayabilir. Araştırma kapsamındaki birimler gözlem süresi başladıktan sonra kendi istekleri doğrultusunda oluşturdukları mazeretlerden kaynaklı olarak çalışmaya devam etmek istemeyebilirler. Bir başka neden ise, birimler çalışma başladıktan sonra farklı sebeplerden dolayı başka yerlere gitmek zorunda kaldıklarından çalışma sonuna kadar iletişim kurulamayabilir. Verilen örnekler dışında daha farklı sebeplerden kaynaklı olarak da bazı birimlere çalışma başladıktan sonra ulaşılması mümkün olmayabilir. Olay zamanı analizinde bu tür veriler sağdan durdurulmuş olarak isimlendirilmektedir.



Şekil 3. Sağdan ve Soldan Durdurma Tipleri

Kaynak: Annette J. Dobson, **An Introduction to Generalized Linear Models**, London: Chapman ve Hall/CRC, Second Edition, 2002, s. 105.

Şekil 3 incelendiğinde T_0 çalışmanın başlangıç noktası T_C ise çalışmanın bitiş noktasını göstermektedir. D ölmüş veya başarısız olmuş olayları ve A ise çalışmanın sonunda yaşayan olayları göstermektedir. L ise çalışmada kaldığı sürece yaşadığı veya başarısız olmadığı bilinen olaylardır. Bu tür olaylar çeşitli sebeplerden dolayı çalışmadan ayrılmak zorunda kalmış yani durdurulmuş bireyi simgelemektedir. 1 ve 2 olayları, çalışma periyodu içerisinde yaşamlarını yitirmişler veya başarısız olmuşlardır. 3 olayı ise çalışma süresince yaşamda kalmış veya başarısız olmamış ancak ne zaman başarısız olduğu veya öldüğü bilinmeyen bir olaydır. Bu tür olaylar sağdan durdurulmuş olay olarak ifade edilebilir²⁹.

Sağdan durdurma ile karşılaşıldığında üç durum söz konusu olabilir³⁰. İlk durum; durdurulmuş gözlemler ihmal edilir yani ilgilenilen olaya maruz kalan kişilerle çalışmaya devam edilir. Olaya maruz kalmayan kişilerin bilgileri ihmal edildiğinden dolayı bu çözüm, bilgi kaybına sebep olmaktadır. Örnek seçiminde yanlılığa sebep olur çünkü durdurulmuş gözlemler daha fazla sürekliliğe sahiptir. Başka bir ifadeyle, risk periyodundaki durdurulmuş birimler için hazard oranı, durdurulmamış birimler için hazard oranından daha düşük olacaktır. Silinen durdurulmuş birimler, çok fazla durdurulmuş birim içeren gruplar için tahmin edilen hazard oranının yanlılığını arttırabilir.

İkinci durum, durdurulmuş zamanlarda meydana gelen olaylar hakkında durdurulmuş zaman olarak çalışmaya devam edilir. İlgilenilen meydana gelmiş olaydaki zamanlar, durdurulmuş zamanların başarısız olması olarak değerlendirilir. Bu yüzden kayıtlanan meydana gelmemiş olaylar anlamsız olarak analiz edilmektedir.

Son durum ise, durdurma mekanizması ile ilişkili varsayımları olan kısmi gözlemlenmiş bilgileri kullanan tahmin yöntemlerini kullanmaktır. Sağdan durdurulmuş veriler için bu en iyi çözüm olabilir. Bu tip gözlem verileri içeren çalışmalar, kabul edilmiş kesin varsayımlar altında tahmin teknikleri kullanılarak analiz edilebilir.

²⁹ Annette J. Dobson, **An Introduction to Generalized Linear Models**, London: Chapman ve Hall/CRC, Second Edition, 2002, s. 2.

³⁰ Vermunt, s. 119.

2.2.1.5 Soldan Durdurma

Bireylerin ilgilenilen olayla ilgili yaşam zamanları, çalışmaya başlamadan önce başlamışsa kayıt edilmemiş yaşam zamanları soldan durdurulmuş olarak tanımlanmaktadır. Şekil 3’ de görüldüğü gibi “5” olayı soldan durdurulmuştur³¹.

Soldan durdurma, sağdan durdurmadan daha karmaşık bir yapıya sahiptir. Soldan durduruluş gözlemlerde, üç mümkün durumdan biri ile analize devam edilir.

İlk durum, soldan durdurulmuş tüm gözlemlerin silinmesidir. En kolay yapılması düşünülen durum budur. Sağdan durdurulmuş gözlemlerde olduğu gibi yanlış parametre tahmincileri ile karşılaşılmaktadır. Ortalama yaşam zamanlarının karşılaştırılmasında gözlem periyodu kısaysa çok büyük bilgi kaybına sebep olabilir.

İkinci durum ise, birimin risk setine girişini gözleme başlama zamanı olarak kabul etmektir. Soldan durdurulmuş bir birim, olaya maruz kalmamış olarak değerlendirilir. Eğer hazard oranı zamandan bağımsızsa bu doğru bir çözüm olacaktır. Ne yazık ki, üssel yaşam dağılımının parametre tahminlerindeki birkaç yanlışlığa izin verdiği hatalı olarak varsayılsa bile sabit hazard oranının varsayımları veya üssel yaşam gerçekçi olmayacaktır.

Üçüncü durum ise sağdan durdurmada olduğu gibi soldan durdurulmuş gözlemleri kullanmaktır. Soldan durdurulmuş gözlemlerin kullanılmasındaki dezavantaj ise parametre tahmincileri için yanlış seçim probleminin ortaya çıkmasıdır. Gözlemlerin risk setine giriş zamanları bilinmiyorsa yanlış seçim olduğunun ispatlamak zor olabilir³².

³¹ Dobson, s. 2.

³² Vermunt, s. 125.

2.2.1.6 Aralıklı Durdurma

Aralıklı durdurmada (interval-censored), olay zamanı sadece bir aralık içerisinde bilinebilir. Bu durumda hastanın klinik deneylerde yada periyodik olarak takibi sonucunda gerçekleşmektedir.

Hastanın çeşitli sebeplerden dolayı ayrılması durumunda, laboratuvar testleri olayın meydana gelmediğini gösterir. Olay zamanı tam olarak bilinmeyecektir. Bu tür klinik deneylerde olayın iki zaman noktası arasında meydana geldiği bilinmektedir. Bu noktalar, hastanın son negatif laboratuvar test zamanı ve ilk pozitif laboratuvar test zamanıdır. Olay zamanı verileri bu tip durdurmaya maruz kaldıklarında, aralıklı durdurma olarak isimlendirilir³³.

İlgilenilen olay zamanının $[L,R]$ aralığında meydana geldiği bilindiğinde durdurma $L \leq T \leq R$ olarak ifade edilmektedir. Aralıklı durdurma verilerinin özel bir durumu “mevcut durum verileri” (current status data) olarak adlandırılmaktadır. Burada T' in, özel ziyaret zamanı U' dan öncemi veya sonramı meydana gelip geldiği bilinmektedir³⁴.

Sosyal bilimlerde de aralıklı durdurulmuş verilerin kullanımı mümkün olmaktadır. Örneğin, gelir düzeyi incelemesi ve katılım programlarının dört aylık süre içerisinde mülakatlar yapılması sonucunda bir takım veriler elde edilmesi istenmiştir. Her mülakatta dört ay içerisinde bir ay öncesine göre gelirin önemli bir kısmını nerden elde ettiği sorulmuştur. İlk ve üçüncü ay mülakatları yapılmışken çeşitli sebeplerden dolayı ikinci ay görüşme yapılmamış olabilir. Aynı zamanda gelir düzeyine sahip olmayan gözlem birimi çalışmada bulunabilir. Burada geliri olan ve olmayanın durumu dört ay ile sınırlandırılmıştır. Bu takip edilemeyen ayda içermektedir. Burada veri tamamlanmış olmasına rağmen cevap hataları olaylar içinden çok, olaylar arasında

³³ Diane M. Finkelstein, William B. Goggins, David A. Schoenfeld, “Analysis of Failure Time Data with Dependent Interval Censoring”, **Biometrics**, No: 58, Haziran 2002, s. 298.

³⁴ Tianxi Chai, Rebecca A. Betensky, “Hazard Regression for interval-Censored Data with Penalized Spline”, **Biometrics**, No: 59, Eylül 2003, s. 570.

önemli bir geçiş yaratmaktadır. Olası bir hatanın yansıması durumu, zaman aralığında olaylarla ilgili olarak net raporların ortaya konması şeklinde belirlenebilir³⁵.

Durdurulmuş veriler içeren çalışmalarda karşılaştırma yapmak için testler kullanılmaktadır. Sağdan durdurulmuş iki olayın karşılaştırılmasında yaygın olarak kullanılan long-rank testidir. Sun³⁶ tarafından yapılan çalışmada long-rank testi ile aralıklı durdurulmuş iki olayın karşılaştırılması incelenmiştir.

2.2.2 Kesme

Kesme, olay zamanı analizinde bir süreç olarak değerlendirilebilir. Medikal izlemelerde ve mühendislikteki yaşam testi çalışmalarında sıkça kullanılmaktadır. Bunun yanında ekonomide de karşılaşılan bir durum olmaktadır. Medikal izlemelerde, bir hastanın izlenmesinde temel olay olarak ele alınmaktadır. Eğer hasta, kesme noktasında ise olay daha önce araştırmacı tarafından gözlemlenmiştir. Sonuç olarak konunun ret edilmesi söz konusudur. Araştırmacı çalışmanın başında bu durumun varlığından kesinlikle haberdar değildir. Kesme durumu bazı kesin hastalıklarda ortaya çıkabilir. Özellikle hastalığın orta evresindeki hastaların ölmeden önceki hallerinde karşılaşılmaktadır. Ekonomide ise, bağımlı değişkenler farklı yollarla sınırlandırılmaktadır. Örneğin dayanıklı tüketim mallarının satın alınması ile ilgili bir araştırmada dayanıklı tüketim malları için yapılmış olan harcama miktarı negatif olmama koşulunu taşımak zorundadır. Bu tür modeller sınırlandırılmış bağımlı değişken modelleridir. Probit modellere benzemelerinden dolayı Tobit modeller olarak isimlendirilmektedirler. İstatistiksel olarak bu tür modeller “kesme regresyon” modelleri veya “durdurulmuş regresyon” modelleri olarak bilinmektedir. Belirli bir

³⁵James Trussell, Richard Hankinson ve Judith Tilton, **Demographic Applications of Event History Analysis**, Oxford: Clarendon Press, 1992, s. 209-210.

³⁶Jianguo Sun, “*Nonparametric Test for Doubly Interval-Censored Failure Time Data*”, **Lifetime Data Analysis**, No:7, 2001, s. 363-375.

aralığın dışındaki gözlemlerin tümü yitirilirse bu durum kesme olarak isimlendirilmektedir³⁷.

Kesme, soldan kesme (left truncated) ve sağdan kesme (right truncated) olmak üzere iki farklı biçimde ifade edilmektedir.

2.2.2.1 Soldan Kesme

Yaşam zamanlarının zaman başlangıcı, çalışmanın zaman başlangıcından önce gelirse soldan kesme meydana gelmektedir. Soldan kesmenin olduğu çalışmalarda, sadece çalışma başladıktan sonraki başarısızlıklar izlenmektedir. Başka bir ifadeyle bu tür veriler soldan kesilmişlerdir³⁸.

Örneğin, emekli olmuş yaşlıların ölüm zamanları, ay olarak değerlendirilen bir çalışma yapılmıştır. Emekli olmuş kişilerin ne zaman öldüğü sol merkez (yada sağdan durdurma) olarak ve emekli olma yaşı da olayı kesme zamanı olarak kaydedilmiştir. Bireyler emekli olduktan sonra belirli bir yaşa kadar yaşamlarını sürdürmektedirler. Ve ileri yaşlarda ölenler çalışmadan çıkarılmaktadır. Sonuç olarak veri setinde yaşam uzunluğu soldan kesmedir. Bu kesmeden habersiz olmak örneklem uzunluğu önyargısı olarak ortaya çıkmaktadır³⁹.

³⁷Cheng Hsiao, **Analysis of Panel Data**, West Nyack, NY, USA: Cambridge University Press, 2002, s. 225.

³⁸ Liunquan Sun ve Xian Zhou, "Survival Function and Density Estimation for Truncated Dependent Data", **Statistics & Probability Letters**, No: 52, 2001, s. 47.

³⁹ Tableman ve Kim, s. 17.

2.2.2.2 Sağdan Kesme

Sağdan kesme, sadece olayla ilgili olan durumları kapsamaktadır. Olayla ilgili olmayan tüm faktörler dışarıda tutulmaktadır. Ölümle ilgili çalışmalarda kullanılmaktadır. Özellikle ölüm kayıtları sağdan kesme olarak değerlendirilmektedir.

HIV virüsünden etkilenmiş bir AIDS hastasının bekleme süresi içerisinde ölümü söz konusu olan bir çalışmayı örnek olarak ele alalım. Bu örnekteki plana göre, bireyde daha önceden AIDS geliştiği için araştırma içerisinde yer almaktadır. Yeni AIDS mikrobi alanlar ise örneklemin dışında tutulmaktadır. Bu araştırmacı tarafından bilinmediği için sağdan kesmedir⁴⁰.

2.2.3 Olay Zamanı Analizinde Kullanılan Değişkenler

Olay zamanı verileri beş farklı açıklayıcı değişken yardımıyla modellenirler. Bu değişkenler; zaman değişkenleri (time variables), zamana göre farklılaşan değişkenler (time-varying variables), zamana göre sabit değişkenler (time-constant variables), içsel değişkenler (internal variables), ve dışsal değişkenlerdir (external variables)⁴¹.

Zaman değişkenlerinde (time variables), değişkenler doğrudan zamanla ilişkilendirilir. Örneğin yaş veya takvim yılı gibi değişkenler zaman değişkeni olarak isimlendirilirler. Bir bireyin yada olayın zaman içindeki ilerleme ölçümü zaman değişkeni yardımıyla yapılmaktadır.

Zamana göre farklılaşan değişken (time-varying variables), olayların arasındaki kadar olayların içindeki farklılaşmayı ifade eden değişkendir. Her değer ileriki aşama

⁴⁰ Tableman ve Kim, s. 17.

⁴¹ Jörg Blasius ve Michael Geenacre, **Visualization of Categorical Data**, London: Academic Press, 1998, s. 49.

için farklı zaman noktalarında ele alınır. Deneysel süreç içerisinde kontrol altındadırlar ve zaman değişkeni ile zamana göre sabit değişken arasındaki etkileşim olarak ifade edilebilirler. Farklı ölçek türleriyle de ölçülmeleri mümkündür. Örneğin, haftalık çalışma saati oranlı ölçekle ölçülür bunun yanında yüksek akademik statü sıralayıcı ölçekle ölçülür ve medeni durum sınıflayıcı ölçekle ölçülmektedir.

Zamana göre sabit değişken (time-constant variables) ise, zamana göre farklılaşan değişkene benzerlik göstermekle beraber, gözlemlenen zaman içerisinde değerleri değişmemektedir. Zamana göre sabit değişkende, her olay kendi içinde farklılık gösterirken zamanın değişmesi bu farklılaşmayı etkilememektedir. Cinsiyet, ırk ve doğum yeri zamana göre sabit değişken kapsamında incelenmektedir.

İçsel değişkenler (internal variables), bir birey üzerindeki zaman faktörü olarak ele alınmaktadır. Bu durum bireyin yaşamsallığını gerekli kılmaktadır. Olaylar doğrudan bireylerle ilişkilendirilir ve bireyden bireye farklılık gösterir. Anne ve babanın ölüm günleri ve sürücü belgesinin alınma günü içsel değişken olarak isimlendirilir. İçsel değişkenlerde dikkat edilmesi gereken bir durum, ters nedenselliğe veya bir takım önyargılara sebep olabilmeleridir. İçsel değişkenler diğer değişkenlerden etkilenmektedir. Örneğin medikal bir tedavide farklı müdahalelerle ölüm oranının hazardı özel vakalardan etkilenebilir. Fakat zamana göre farklılaşan değişkenler kontrol edildiğinde yok olur. Bu durum tedavinin ölüm üzerinde doğrudan etkisi olmadığı sonucu olarak değerlendirilir. Tedavi genel durumu iyileştirerek ölüm riskini azaltabilir. İçsel değişkenler süreci iyi anlamada daha etkili olabilir fakat devam eden süreçteki değişimi daha iyi anlayabilmek için içsel değişkenlerin araştırılması gerekmektedir.

Dışsal değişkenler (external variables), bireyin yaşamsallığı içerisinde bağımsız olarak değer aldıklarında ortaya çıkmaktadır. Dışsal değişkenlerin değer almaları olayın olması yada olmamasına bağlıdır. Zaman değişkeni olarak incelenen takvim yılı çalışmalarında olduğu gibi dışsal değişkende bütün örnekleri etkiler. Örneğin, işsizlikle ilgili yapılan bir çalışmada hükümet değişimleri veya incelenen bölgedeki temel fabrikaların kapanması dışsal değişken olarak ele alınmaktadır. İçsel

değişkenlerin dezavantajı olan ters nedensellik ve bir takım önyargılara bağlı olmaları dışsal değişkenlerde görülmemektedir⁴².

2.3 Olay Zamanı Analizinde Kullanılan Yöntemler

Sosyal bilimlerde bazı uygulamalar, olay zamanı verilerinin modellenmesini içerir. Bu tür veriler iyi tanımlanmış bir başlangıç zamanından “başarısızlık” olarak isimlendirilen bazı olayların meydana gelişine kadar olan zamandır. “T” ile gösterilen başarısızlığa kadarki zaman verileri iki önemli özelliğe sahiptirler⁴³.

1. Bu zamanlar negatif olamazlar ve özellikle uzun kuyruklu asimetrik dağılımlara sahiptirler.

2. Birimlerin bazıları gerçek olay zamanları bilinmemelerine rağmen yaşamlarını sürdürebilirler. Başarısızlık zamanlarının tamamen bilinmediği bu veriler “durdurulmuş” olarak ifade edilmektedir. Üzerinde değerlendirme yapılacak tüm olguların araştırmaya aynı anda hepsi için geçerli bir başlangıç noktası ile katılmaları nerede ise olanaksızdır, dolayısıyla araştırma için belirlenmiş bir süre içinde, çeşitli anlarda araştırmaya izlenmeye katılan yeni olgular olabileceği gibi belli bir özelliğe erişmesi (ölmesi) nedeni ile izlemeden çıkan örneklerde bulunacaktır.⁴⁴

T dağılımı dört fonksiyon ile ifade edilmektedir.

1. Kümülatif dağılım fonksiyonu $F(t)$ ile gösterilmektedir. Bu miktar t zamanından önce başarısızlık olasılığını tanımlamaktadır.

$$\begin{aligned} F(t) &= P(\text{bireylerin } t \text{ zamanından önce başarısızlığı}) \\ &= P(T < t) \end{aligned} \quad (2)$$

⁴²Blasius ve Geenacre, s. 49; Vermunt, s. 132.

⁴³Dobson, s. 3.

⁴⁴Mustafa Şenocak, **Biyoistatistik**, İstanbul: Dilek Ofset Matbaacılık, 1998, s. 253.

2. Olasılık yoğunluk fonksiyonu $f(t)$ ile gösterilir. Bu miktar, kümülatif dağılım fonksiyonunun türevidir ve bireylerin küçük bir aralığın her birim zamanında ki başarısızlığa düşme olasılığıdır.

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{P(\text{bir bireyin}(t, \Delta t)\text{'deki ölümü})}{\Delta t} \quad (3)$$

Tüm yoğunluk fonksiyonlarının $f(t)$, aşağıdaki özelliklere sahip olduğu varsayılmıştır.

a. Eğri altındaki alan 1'e eşittir.

b. Yoğunluk fonksiyonu negatif olmayan bir fonksiyondur.⁴⁵

$$f(t) \geq 0, \text{ ise } t \geq 0$$

$$f(t) = 0, t < 0 \quad (4)$$

3. Yaşam fonksiyonu $S(t)$ ile gösterilmektedir. Bu fonksiyon t zamanından daha uzun yaşama olasılığını verir.

$$S(t) = P(\text{bir bireyin } t \text{'den daha uzun yaşaması})$$

⁴⁵ Shoukri ve Pause, s. 2.

$$=P(T>t) \quad (5)$$

olasılık yoğunluk fonksiyonu ile ilişkilendirilirse;

$$\begin{aligned} S(t) &= 1 - P(T>t) \\ &= 1 - F(t) \end{aligned} \quad (6)$$

Yaşam fonksiyonunun aşağıdaki özelliklere sahip olduğu varsayılmıştır.

- a. Sıfır zamandaki yaşam olasılığı birdir. $S(t)=1$, $t=0$ için
- b. Sonsuz yaşam olasılığı sıfırdır. $S(t)=0$, $t=\infty$ için
- c. Yaşam fonksiyonu artmayan bir fonksiyondur.

T sürekli bir tesadüfi değişkense $S(t)$ sürekli monoton azalan bir fonksiyondur ve olasılık yoğunluk fonksiyonu $f(t)$ 'in integralidir.

$$S(t) = P(T \geq t) = \int_t^{\infty} f(u) d_u \quad (7)$$

Böylece izleyen ilişki aşağıdaki gibi olmaktadır.

$$f(t) = - \frac{dS(t)}{dt} \quad (8)$$

$f(t)\Delta t$, t zamanında meydana gelen olayın “yaklaşık” olasılığı olarak ifade edilebilir ve bire eşit olan $f(x)$ 'in altındaki bölge ile negatif olmayan bir fonksiyondur.

T süreksiz bir değişken olduğunda, artmayan soldan devam eden adım fonksiyondur. Eğer T kesikli tesadüfi değişkenin olasılık dağılımı ise,

$$P(t_j) = P(T=t_j), \quad (9)$$

j =1,2,..... ile $t_0 < t_1 < t_2, \dots$ değerlerini alırsa,

$$S(t) = P(x \geq t) = \sum_{j:t_j \geq t} P(t_j) \quad (10)$$

olmaktadır⁴⁶.

4.Hazard fonksiyonu h(t) ile gösterilmektedir. Bu fonksiyon, aralığın başlangıcındaki bireylerin yaşamları üzerinde koşullu, küçük bir zaman aralığındaki bir bireyin başarısızlığa düşme olasılığını verir⁴⁷.

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P((t + \Delta t) > T \geq t / T \geq t)}{\Delta t} \quad (11)$$

Önceki tanımlanmış olan miktarlar açısından, hazard fonksiyonu aşağıdaki gibi yazılabilir.

$$h(t) = \frac{f(t)}{1 - F(t)} \quad (12)$$

⁴⁶J.P Klein, **Survival Distributions and Their Characteristics**. [www.biostat.mcw.edu/homepgs/klein/Teach Report #22](http://www.biostat.mcw.edu/homepgs/klein/Teach%20Report%20#22), (28 Kasım 2002).

⁴⁷ Vera Bröcheler, Steven Maijoor ve Arjen Van Witteloostuijin, "Auditor Human Capital and Audit Firm Survival The Dutch Audit Industry in 1930-1992", **Accounting, Organizations and Society**, No: 29, 2004, s. 637.

Uygulamada, hazard fonksiyonu ölümlülük zorlayıcısı veya anlık başarısızlık oranı olarak tanımlanabilir.

Kümülatif hazard fonksiyonu $H(t)$ ile gösterilir ve integral alınmış hazard fonksiyonudur.⁴⁸

$$H(t) = \int_0^t h(x) dx \quad (13)$$

İncelenen dört miktardan birisi verildiğinde diğer üçü kolaylıkla elde edilebilir. Yaşam fonksiyonu belirtildiğinde kümülatif dağılım fonksiyonu, olasılık yoğunluk fonksiyonu ve hazard fonksiyonu belirlenebilir

Yaşam verileri analizi için farklı metodlar bulunmaktadır. Olasılık yoğunluk fonksiyonu $f(t)$ için hiçbir dağılım varsayımları bulunmayan parametrik olmayan yaklaşımlar ve dağılım kısıtları bulunan parametrik yaklaşımlardır.⁴⁹

Bu çerçevede içerisinde olay zamanı analizinde kullanılan yöntemler durdurulmamış veriler için yaşam fonksiyonu yöntemleri ve durdurulmuş veriler için yaşam fonksiyonu yöntemleri olarak iki farklı sınıflandırma ile incelenebilir.

2.3.1 Durdurulmamış Veriler İçin Yaşam Fonksiyonu Yöntemleri

Yaşam fonksiyonu tahminleri, olasılık yoğunluk fonksiyonu ve hazard fonksiyonu, durdurulmamış veriler için aşağıdaki gibi gösterilmektedir.⁵⁰

1. Durdurulmamış veri için yaşam fonksiyonu tahmini

⁴⁸ Ross A. Maller ve Zhou Xian, **Survival Analysis with Long-Term Survivors-Wiley Series in Probability and Statistics**, New York: John Wiley ve Sons, Ltd.(UK), 1996, s. 28.

⁴⁹Shoukri ve Pause, s. 103-108.

⁵⁰Shoukri ve Pause, s. 120.

$$S(t) = \frac{n_s}{N} \quad (14)$$

n_s = t zamanda daha uzun yaşıyan kiři sayısı ve

N= Toplam kiři sayısı

olarak gösterilebilir.

2.Durdurulmamıř veri için olasılık yoğunluk fonksiyonu tahmini

$$f(t) = \frac{n_t}{N \cdot \Delta t} \quad (15)$$

n_t =t zamanında başlayan aralıktaki ölenlerin sayısı ve

Δt =aralık genişliđi

olarak eşitlik (15)' de yer almaktadır.

3. Durdurulmamıř veri için hazard fonksiyonu tahmini

$$h(t) = \frac{n_i}{n(t)\Delta t} \quad (16)$$

n_i = t zamanındaki aralık başlangıcında ölen kiři sayısı ve

$n(t)$ =t zamanında yaşıyan kiři sayısı

olarak ifade edilebilir.

2.3.2 Durdurulmuş Veriler İçin Yaşam Fonksiyonu Yöntemleri

Yaşam fonksiyonun durdurulmuş veriler için tahminleri Kaplan Meier tarafından geliştirilmiştir. Çalışılan anakütledeki bireylere $t_1 < t_2 < \dots < t_k$ zamanlarında değer biçilebiliyorsa, t_i ($1 \leq i \leq k$), zamanındaki yaşam fonksiyonu aşağıdaki gibi olmaktadır⁵¹.

$$S(t_i) = P(t_1 \text{ zamanında yaşama}) * P(t_2 \text{ zamanında yaşama} \mid t_1 \text{ zamanında yaşama}) \dots * P(t_j \text{ zamanında yaşama} \mid t_{j-1} \text{ zamanında yaşama}) * P(t_i \text{ zamanında yaşama} \mid t_{i-1} \text{ zamanında yaşama}) \quad (17)$$

t_j ($1 \leq i \leq k$) zamanında, i 'nci kişi çalışma yapılan olay sonucu başarısızlığa uğrayanlar (inceleme konusu hastalık ise hastalıktan ölenler) veya çalışmadan ayrılanlardır. Ayrılanlar maruz kalınan durum dışındaki ölümlerdir veya çalışmaya devam etmekte isteksizdirler. Verilerin bu tipi sağdan durdurulmuş veri olarak isimlendirilerek yapılandırılmaktadır. Yaşam analizinde sağdan durdurulmuş verilerle sıkça karşılaşmaktadır.

S_{j-1} kişilerin, t_{j-1} zamanı boyunca yaşayan (başarısızlığa maruz kalmayan) ve tamamlanmamış verilere sahip olduğu varsayılmıştır. Bu kişiler arasında, S_j kişiler başarısız olan kişiler, d_j başarısızlığa uğrayanlar ve l_j kişiler t_j zamanındaki durdurulmuş (çalışmadan ayrılanlar) kişilerdir.

$$S_{j-1} = S_j + d_j + l_j \quad (18)$$

Böylece yaşam fonksiyonunun Kaplan-Meier tahmincisi $\hat{S}(t_i)$, aşağıdaki gibi olmaktadır.

⁵¹ Ying Zhang, **Logistic Regression Model**, <http://pegasus.cc.ucf.edu/~zhang/Sta6938/topic.pdf>, (10 Ocak 2004).

$$\hat{S}(t_i) = \left(1 - \frac{d_1}{S_0}\right) x \left(1 - \frac{d_2}{S_1}\right) x \dots x \left(1 - \frac{d_i}{S_{i-1}}\right)^{52}. \quad (19)$$

2.3.3 Kesilmiş Veriler İçin Yaşam Fonksiyonu Yöntemleri

Kesilmiş veriler için yaşam fonksiyonunu Kaplan-Meier tahmincileri ile tahmin edilebilmesi için bir takım düzenlemelerin yapılması gerekmektedir. Elde edilen tahminci “Product-Limit” tahmincisi olarak isimlendirilmektedir. Kesilmiş veriler için yaşam fonksiyonu aşağıdaki gibi gösterilebilir⁵³.

$$S(t) = \prod \left[1 - \frac{d_i}{Y_i}\right] \quad (20)$$

Eşitlikteki Y_i risk altındaki kişi sayısı olarak ifade edilebilir.

2.4 Olay Zamani Analizinde Kullanılan Regresyon Modelleri

Tüm parametrik fonksiyonlar olasılık yoğunluk fonksiyonu $f(t)$ için, dağılım şekli şartları içerir. Bunlar, yaşam fonksiyonu $S(t)$ ve hazard fonksiyonu $h(t)$ için üstel ve weibull dağılımına göre aşağıdaki gibi olmaktadır.

Üstel bir dağılımın yoğunluk fonksiyonu⁵⁴;

$$f(t) = \lambda e^{-\lambda t} \quad t \geq 0, \lambda > 0$$

⁵² M.J. Campbell, **Statistics at Square One**, London: BMJ Publishing Group, 2002, s. 126.

⁵³ Cheng, s. 225-226.

⁵⁴ Fikret Er ve Harun Sönmez, **İstatistik Formülleri ve Olasılık Dağılımları**, Eskişehir: Anadolu Üniversitesi Fen Fakültesi Yayınları No. 21, 2001, s. 43.

$$=0 \quad t < 0 \quad (21)$$

olarak ifade edilebilir. Yaşam fonksiyonu⁵⁵;

$$S(t) = e^{-\lambda t} \quad , t \geq 0 \quad (22)$$

ve hazard fonksiyonu,

$$h(t) = \lambda \quad , t \geq 0 \quad (23)$$

olmaktadır.

Hazard fonksiyonu zamandan bağımsızdır. Ani koşullu başarısızlık oranları, bir yaşam süresi içinde değişmezler. $\lambda=1$ olduğunda dağılım birim üstel dağılımdır. Uygulamada, bir çok başarısızlık zamanı sabit hazard başarısızlığıyla karşılaşmamaktadır. Bu yüzden yaşam analizleri için sabit hazard uygulamaları sınırlı olmaktadır. Üstel model, geniş bir uygulama alanına sahip olan weibull modelinin özel bir durumudur⁵⁶.

Weibull dağılımının yoğunluk fonksiyonu,

$$f(t) = \lambda \gamma t^{\gamma-1} e^{-\lambda t^\gamma} \quad t \geq 0, \gamma, \lambda \geq 0 \quad (24)$$

olarak ifade edilirse, Yaşam fonksiyonu

$$S(t) = e^{-\lambda t^\gamma} \quad (25)$$

olmaktadır. Hazard fonksiyonu

$$h(t) = \lambda \gamma t^{\gamma-1} \quad (26)$$

olmaktadır.

⁵⁵Chap T. Le, **Applied Survival Analysis**, New York: John Willey ve Sons, Inc., 1997, s. 9.

⁵⁶Tableman ve Kim, s. 56.

λ ve γ ölçek ve şekil parametreleridir. Özel bir durum olarak $\gamma=1$ olduğunda, sabit hazard'lı üstel model tanımlanmaktadır.

$\gamma>1$ olduğunda zamana göre artan bir hazard ve $\gamma<1$ olduğunda ise zamana göre azalan bir hazard olmaktadır⁵⁷.

Olay zamanı analizinde kullanılan regresyon modelleri; üssel regresyon modeli, weibull regresyon modeli, oransal hazard modeli, cox oransal hazard regresyon modeli ve poisson regresyon modeli olarak ifade edilebilir.

2.4.1 Üssel Regresyon Modeli

Üssel dağılım, hazard regresyon modeli için genelleştirilebilir. Üssel dağılım için hazard fonksiyonu;

$$h(t) = \lambda \quad (27)$$

olmaktadır. Burada λ zamana göre bir sabittir ve

$$E(t) = \frac{1}{\lambda} \quad (28)$$

olmaktadır.

Hazard oranı λ , eş değişken vektörü x' in bir fonksiyonu olarak modellenmektedir. Bir birey için t zamanındaki hazard fonksiyonu,

⁵⁷Wayne Nelson, "Theory and Applications of Hazard Plotting for Censored Failure Data", **Technometrics**, No: 42, (2000), s. 12-25; Fikret Er ve Harun Sönmez, **İstatistik Formülleri ve Olasılık Dağılımları**, Eskişehir: Anadolu Üniversitesi Fen Fakültesi Yayınları, No. 21, 2001, s. 44; Chap T. Le, **Applied Survival Analysis**, New York: John Willey ve Sons, 1997, s. 10-11.

$$h(t / x) = h_0(t).k(x\beta) = \lambda.k(x\beta) = \lambda.k(\beta_1 x_1 + \dots + \beta_m x_m) \quad (29)$$

biçiminde ifade edilirse $\beta = (\beta_1, \beta_2, \dots, \beta_m)$ regresyon katsayıları vektörüdür. $\lambda > 0$ bir sabittir ve k tanımlanan özel bir fonksiyondur. Eş değişken vektörü $x = 0$ ve $\beta = 0$ olduğunda, $h_0(t)$ fonksiyonuna eşit olur ve temel hazard olarak isimlendirilir. Hazard fonksiyonu t zamanına göre sabittir fakat x 'e bağlıdır.

$$\begin{aligned} h(t / x) &= \lambda \cdot \exp(x\beta) \\ &= \lambda \cdot \exp(\beta_1 x_1 + \dots + \beta_m x_m) \\ &= \lambda \cdot \exp(\beta_1 x_1) \cdot \exp(\beta_2 x_2) \cdot \dots \cdot \exp(\beta_m x_m) \end{aligned} \quad (30)$$

Yukarıdaki eşitlikte, eş değişkenlerin hazard oranı üzerinde çarpımsal etkisi olduğu görülmektedir. Eşitliğin logaritması alındığında,

$$\log(h(t / x)) = \log(\lambda) + \eta = \log(\lambda) + (x\beta) = \log(\lambda) + \beta_1 x_1 + \dots + \beta_m x_m \quad (31)$$

biçiminde yazılabilir. $\eta = x\beta$ log-hazardın doğrusal tahmincisi olarak isimlendirilir⁵⁸.

k fonksiyonu birkaç farklı fonksiyon olarak ele alınabilir. Üç farklı şekilde k fonksiyonu ifade edilebilir⁵⁹.

⁵⁸ Tableman ve Kim, s. 96.

$$(1) k(\eta) = 1+\eta \quad (32)$$

$$(2) k(\eta) = 1/(1+\eta) \quad (33)$$

$$(3) k(\eta) = \exp(\eta) \quad (34)$$

Üç farklı fonksiyon için model ayrı ayrı düzenlenirse birinci fonksiyon için model;

$$h(t/x) = \lambda (1 + x\beta) \quad (35)$$

x 'in doğrusal bir fonksiyonu olur. İkinci fonksiyon için düzenleme yapıldığında;

$$E(T/x) = 1/h(t/x) = (1 + x\beta) / \lambda \quad (36)$$

biçiminde ifade edilir. İki fonksiyonda da β değerlerinin mümkün olan tüm birimler için $k(x\beta) > 0$ olması koşulu sağlanmalıdır. Koşul sağlanmazsa hazard değeri negatif olabilir. Bundan dolayı ilk iki fonksiyon pek kullanılmamaktadır. $k(\eta) = \exp(\eta)$ fonksiyonu $k(x\beta) > 0$ koşulunu kendiliğinden sağladığı için daha çok kullanılan bir fonksiyon olmuştur. Modelde, $k(\eta) = \exp(\eta)$ göre düzenleme yapıldığında yaşam fonksiyonu,

$$S(t/x) = \exp(-h(t/x)t) = \exp(-\lambda \exp(x\beta)t) \quad (37)$$

olur. Böylece, T 'nin koşullu olasılık yoğunluk fonksiyonu,

$$f(t/x) = h(t/x) \cdot S(x\beta) = \lambda \cdot \exp(x\beta) \cdot \exp(-\lambda \cdot \exp(x\beta)t) \quad (38)$$

⁵⁹ J.D. Kalbfleisch ve R.L. Prentice, **The Statistical of Failure Time Data**, New York: John Willey and Sons, 1980, s. 31.

şeklinde yazılabilir. Eğer T üstel dağılmış bir değişken ise, $Y = \log(T)$ ölçek parametresi ($\sigma=1$) olan uç değerler dağılımı olarak dağılmıştır. $Y = \log(T)$ dönüşümü yapılırsa,

$$\mu = -\log(h(t/x)) = -\log(\lambda \exp(x\beta)) = -\log(\lambda) - x\beta \quad (39)$$

$$Y = \log(T) = \mu + \sigma Z = \beta_0^* + x\beta^* + Z$$

olarak yazılabilir. Eşitlik (39)'da $\beta_0^* = -\log(\lambda)$, $\beta^* = -\beta$ ve $Z \sim f(z) = (\exp z - e^{-z})$ olmaktadır⁶⁰.

2.4.2 Weibull Regresyon Modeli

Weibull dağılımı, olay zamanı verilerinin modellenmesinde oldukça uygun bir dağılım olarak görülebilir. Weibull regresyon modellemesi, üssel regresyon modellemesi ile aynı biçimde yapılmaktadır. Başka bir ifade ile, weibull regresyon modeli weibull dağılımının bir genelleştirmesidir⁶¹. Hazard fonksiyonu,

$$h(t) = \alpha \lambda^\alpha t^{\alpha-1} \quad (40)$$

olarak ifade edilirse koşullu hazard fonksiyonu,

$$\begin{aligned} h(t/x) &= h_0(t) \cdot \exp(x\beta) \\ &= \alpha \lambda^\alpha t^{\alpha-1} \cdot \exp(x\beta) \end{aligned} \quad (41)$$

⁶⁰Tableman ve Kim, s. 96-97.

⁶¹M. Moeschberger, *Survival Analysis*, New York: Springer-Verlag, 1997, s. 375.

olarak yazılabilir. $Y = \log T$ dönüşümü uygulandığında

$$Y = \beta_0^* + x \beta^* + \sigma Z \quad (42)$$

biçiminde gösterilebilir. Eşitlik (42) weibull regresyon modelini ifade etmektedir.

Modeldeki $\beta_0^* = -\log(\lambda)$ ve $\beta^* = -\sigma\beta$ olmaktadır. Üssel ve weibull regresyon modellerinin her ikisi de log-lineer modellerdir. Her iki modelde de Z hata değişkenidir. Z değişkeni uç değerler dağılımına sahiptir⁶².

2.4.3 Oransal Hazard Modeli

Parametrik metotların olumlu bir özelliği, benzerlikle yapılandırılmayı kabul eden olasılık yoğunluk fonksiyonu için bir form belirtmesidir. Maksimum benzerlik tahminleri ve standart hatalar modeldeki tüm parametreler için elde edilebilir. Buna rağmen, model parametrelerindeki engel, bir kısım veri setleri için bir olasılık dağılım fonksiyonu belirtmeyi istememesine uzanmaktadır. Bu durum, parametrik olmayan hesaplamaları daha cazip hale getirmektedir. İdeal durum, yoğunluk fonksiyonunda dağılım kısıtlayıcıları olmamasını içermektedir. Bununla birlikte, regresyon parametrelerinin maksimum benzerlik tahminleri kolayca elde edilebilmektedir. Bu gibi özellikleri içeren bir analiz, cox oransal hazard modeli kullanımıyla ifade edilebilmektedir.

Oransal hazard modeli, i 'nci bireyin hazard fonksiyonu tüm bireyleri kapsayan temel hazardın bir parçasıdır varsayımıyla yapılandırılmıştır ve $h_0(t)$ ile gösterilmektedir. Ve açıklayıcı değişken vektörünün bir fonksiyonu, x_i bireyleri gösterdiğinde, $\zeta(x_i)$ şeklinde ifade edilmektedir.

⁶² Tableman ve Kim, s. 96-97.

$$h_i(t) = h_o(t) \zeta (x_i) \quad (43)$$

Hazard oranı veya göreli hazard negatif olmayan bir fonksiyondur⁶³. Daha genel olarak ifade etmek gerekirse, belli bir zaman diliminde ortaya çıkan bir olayın olasılığı hazard oranı (hazard rate) olarak isimlendirilir⁶⁴.

Olay zamanı analizinde, olay zamanları devam eden tesadüfi bir değişken olarak varsayılmıştır. Oransal hazard modeli, açıklayıcı değişkenlerin hazard fonksiyonuna çarpım esaslı zaman-bağımsız olarak dahil edildiğini varsayar⁶⁵.

$$\zeta (x_i) = \frac{h_i(t)}{h_o(t)} \quad (44)$$

Açıklayıcı değişken fonksiyonu da negatif olmayan bir fonksiyondur.

$$\zeta (x_i) = e^{x_i \beta} \quad (45)$$

şeklinde gösterilmektedir.

$x_i = (x_{i1}, x_{i2}, \dots, x_{im})$ bireylerin i'nci açıklayıcı değişken vektörleri yazıldığında, tekrar bir düzenleme yapıldığında oransal hazard modelinin log hazard oranı için doğrusal bir model olduğu aşağıdaki gibi gösterilmektedir.

$$\log \left[\frac{h_i(t)}{h_o(t)} \right] = \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_m x_{im} \quad (46)$$

⁶³ Kalbfleisch ve Prentice, s. 71-84.

⁶⁴ Rose ve Sullivan, s. 220.

⁶⁵ H. Kozumi, "Bayesian Analysis of Discrete Survival Data with a Hidden Markov Chain", **Biometrics**, S. 56, Aralık 2000, s. 1002.

Oransal hazard için kısmi benzerlikler

$$L(\beta) = \prod_{j=1}^r \frac{e^{(\beta'x_{(j)})}}{\sum_{i \in R(t_{(j)})} e^{(\beta'x_{(i)})}} \quad (47)$$

$R(t_{(j)})$ j'nci düzenli ölüm zamanındaki $t_{(j)}$ risk grubudur (bireylerin ayrılışı). Böylece benzerlik formülün j'nci düzenli ölüm zamanları yeniden sonuçlanır.

$$L_i(\beta) = \frac{e^{(\beta'x_{(i)})}}{\sum_{i \in R(t_{(i)})} e^{(\beta'x_{(i)})}} \quad (48)$$

=P ($t_{(i)}$ ' deki bir başarısızlık ve $R(t_{(i)})$ risk grubunda olan gözlemlenmiş bireylerde başarısızlığın meydana gelmesi)⁶⁶.

Modelde sürekli zaman hazard oranı; sosyal araştırmalarda kullanılan olay zamanı analizinde, yaygın olarak kullanılan bir teknik olmuştur. Bu modelin ismi oransal hazard (Proportional hazard) olarak isimlendirilir. Oransal ifadesinin kullanılma sebebi; zaman içerisinde herhangi bir nokta için, bir bağımsız değişken kategorisindeki etki oranının sabit olmasından kaynaklanmaktadır. Özellikle sosyal bilimciler arasında tercih edilmekte ve yaygın olarak kullanılmaktadır. Oransal hazard üzerindeki aralık (range) etkisini ölçmesi, modelin tercih sebebi olmasını sağlamaktadır.

Oransal hazard modelleri iki önemli avantaj sunmaktadırlar. Birincisi hem sabit hem de zamana ait değişkenlerin aynı modelde kullanılabilmesidir. İkincisi zaman ve hazard oranı arasındaki ilişkiyi belirlemek araştırmacıya bağlı değildir. Araştırmacıdan bağımsızdır.⁶⁷

⁶⁶ Kalbfleisch ve Prentice, s. 71-84.

⁶⁷ Rose ve Sullivan, s. 220.

Oransal hazard Fonksiyonu Nadas⁶⁸ tarafından geliştirilmiştir. Son yıllarda oransal hazard modeliyle ilgili gelişmeler görülmüştür. Andersen⁶⁹ yapmış olduğu çalışmasında, oransal hazard modelini test merkezleri açısından değerlendirmiştir. Burada iki tür yaklaşım kullanılmaktadır. Birinci yaklaşım, her merkezde varolan oransal hazard modelinin sabit açıklayıcı değişkenlerle birlikte yer almasıdır. Merkezin özel olarak uyumu iki skorda kullanılarak sağlanır. Bunlar Hipotez testleri açısından Wald ve benzerlik oranıdır. Bütün özel açıklayıcı değişkenler merkezde sıfırdır. Alternatif bir yaklaşım olarak tesadüfi etki yada hata modelidir. Son zamanlarda bazı araştırmacılar tesadüfi etki modelini skor testleri açısından kullanmayı amaçlamaktadırlar. Ortalama örnekler sabit etkinin özel değerden daha yüksek olduğunu göstermektedir. Ancak tesadüfi etki testi sıfır hipotezinin altında gerçekleşmektedir. Alternatif hipotezlerde tesadüfi etki testi, göreceli olarak küçük parçaların yada tesadüfi merkez etkisinin ortaya konmasında etkilidir. Aynı zamanda eğer merkez etkisinden habersiz olunursa tahmincinin etkisinin yanlışlık gösterdiği ve tahmincinin tutarsız olduğu ifade edilebilir. Liu⁷⁰, sosyal araştırmalardaki açıklayıcı değişkenler arasındaki etkilerin, hazard oranı üzerindeki olumsuz etkilerini gidermek için yapısal hazard oranını önermiştir. Yang ve Ying⁷¹ ise çoklu olay zamanı verileri için marjinal oransal hazard modeli geliştirmişlerdir. Farrington⁷² aralıklı durdurulmuş veriler için oransal hazard modelinde kullanılan artıkları incelemiştir.

⁶⁸ Arthur Nadas, "On Proportional Hazard Functions", **Technometrics**, No: 12, S: 2, Mayıs 1970, s. 413.

⁶⁹ Perprag Andersen, John P. Klein ve Mei-Jie Zhang, "Testing for Centers effects in Multicenter Survival studies: A Monte Carlo Comparison of Fixed and Random Effects Tests", **Statistics and Medicine**, No: 18, S: 12, Ocak 1999, s. 1489-1500.

⁷⁰Xian Liu, "Development of a Structural Hazard Rate Model in Sociological Research", **Sociological Methods and Research**, No: 29, S:1, Ağustos 2000, s. 77-79.

⁷¹Yaning Yang ve Zhiliang Ying. "Marjinal Proportional Hazards Models for Multiple Event-Time Data", **Biometrika**, No: 88, 2001, s. 581.

⁷²C.P. Farrington. "Residuals for Proportional Hazards Models with Interval-Censored Survival Data", **Biometrics**, No: 56, Temmuz 2000, s. 473-482.

2.4.4 Cox Oransal Hazard Regresyon Modeli

Dr. Cox tarafından geliştirilmiş olan ve kendi adıyla anılan Cox oransal hazard modeli; lineer ve genelleştirilmiş lineer modellere benzeyen regresyon modeliyle yaşam verilerini analiz eden bir model olmaktadır. Lineerlik varsayımını sağlayan ölçek log hazard ölçeğidir. Model, Cox'un maksimum benzerlik tahmincilerini kullanarak uygulanmaktadır. Modelin katsayıları, iki grup arasındaki hazard oranının logaritması olarak tahmin edilmektedir⁷³.

Cox oransal hazard modeli, yaşam verilerinin analizinde en çok kullanılan modellerden biridir. Yarı parametrik olan cox oransal hazard modeli, hazard fonksiyonu tahmincilerinin etkileri ile ilgili parametrik varsayımlar içerir. Bunun yanında doğal hazard fonksiyonunun kendisi hakkında hiçbir varsayım içermemektedir. Cox oransal hazard modeli, tahmincilerin hazard fonksiyonu üzerinde çarpımsal davrandığını varsayar fakat hazard fonksiyonunu (üstel modelde olduğu gibi), weibull veya herhangi özel formlarda varsaymaz. Modelin regresyon kısmı tam parametrikdir. Açıklayıcı değişkenler doğrusal olarak log hazard veya log-kümülatif hazardla ilişkilidir. Bir çok durumda, gerçek olan hazard formları bilinmez veya karmaşıktır. Bu yüzden cox modelin avantajları sınırlıdır.

Cox oransal hazard modeli, sadece başarısızlık sıra verisi (the rank ordering of the failure) ve durdurulmuş zamanları kullanır. Bu yüzden başarısızlık zamanlarının uç değerlerinin tam parametrik metotlarından daha az etkilenir. Model özel bir durum olarak iki grup arasındaki yaşamı karşılaştırmak için iyi bilinen log-rank testini içerir. Tahminler ve regresyon katsayıları testleri için, parametrik modelin tüm varsayımları yerine getirilmiş olsa bile cox model parametrik model kadar etkili olmaktadır. Parametrik bir modelin varsayımları doğru olmadığı zaman (örneğin weibull modeli

⁷³ Peter Dalgaard, **Introductory Statistics with R**, New York: Springer-Verlag, 2002, s. 218.

kullanıldığında ve anakütle weibull yaşam dağılım formunda değilse yani model seçimi doğru değilse), cox analizi parametrik analizlerden çok daha etkilidir⁷⁴.

Cox regresyon modeli aşağıdaki eşitlik ile gösterilebilir.

$$h(t)=h_0(t)\exp(x\beta) \quad (49)$$

Cox regresyon modeli, etkili bir analiz olmasından dolayı farklı çalışmalarda araştırmacılara yol gösteren bir model olmaktadır. Ma, Krewski ve Burnet⁷⁵ tesadüfi etkiler Cox regresyon modeline bir poisson regresyon modeli yaklaşımını geliştirmişlerdir. Bu yaklaşımın önemli bir özelliği olarak, asıl sonuçların gözlemlenemeyen tesadüfi etkilerin ilk ve ikinci momentlere bağlı olduğunu göstermişlerdir. Önerilen metot, büyük ölçekli kohort araştırmalardaki verilerin tekrar analizini izah etmektedir. Fisher ve Lin⁷⁶ yaptıkları çalışmada, cox oransal hazard regresyon modelinin zamana göre değişen açıklayıcı değişkenlerin kullanımında karşılaşılan durumları incelemişlerdir.

Sosyal bilimler alanında yapılan ilginç bir çalışmada, organizasyonların finansal başarısızlıkları cox regresyon ile modellenmiştir. Cox regresyon analizi ile elde edilen sonuçlar çoklu diskriminant analizi ve logistik regresyon analizi ile elde edilen sonuçlarla karşılaştırılarak organizasyonların finansal başarısızlıklarına çözüm üretilmeye çalışılmıştır⁷⁷.

⁷⁴ D. Y. Lin, “ *On Fitting Cox’s Proportional Hazards Models to Survey Data*”, **Biometrika**, No: 87, S: 1, 2000, s. 37-38.

⁷⁵ Renjun Ma, Daniel Krewski ve Richard T. Burnett, “*Random Effects Cox Models: A Poisson Modelling Approach*”, **Biometrika**, No: 90, S: 1, Mart 2003, s. 157-158.

⁷⁶ Lloyd D. Fisher ve D.Y.Lin, “*Time Dependent Covariates in The Cox Proportional-Hazards Regression Model*”, **Annu.Rev. Public Health**, No: 20, 1999, s. 145-157.

⁷⁷ Serpil Ergün Bülbül. “*İstatistiksel Başarısızlık Zaman Modelleri ve Finanssal Analizlerde uygulanması*”, MÜ Sosyal Bilimler Enstitüsü (Yayınlanmamış Doktora Tezi), İstanbul, 1998.

2.4.5 Poisson Regresyon Modeli

Poisson regresyon modeli, cevap deęişkeni veya baęımlı deęişkenin sayılabilir bir deęişken olduęu durumlarda kullanılan bir regresyon modeli olmaktadır. İlgilenilen bir olayın meydana gelme oranları veya bir olayın meydana gelme sayıları poisson regresyon ile modellenebilir⁷⁸. Öğrencilerin okula gitmedikleri gün sayısı, doktor ziyareti sayısı ve bakteriler için koloni sayısı gibi farklı örneklerin modellenmesinde poisson regresyon modeli uygun bir model olabilir.

y pozitif deęerler alan sayılabilir bir tesadüfi deęişken olarak ele alındığında, istatistiksel teori çoęunlukla poisson daęılımının kullanıldığı bazı süreksiz daęılımların kullanımını sağlamaktadır. Poisson daęılımı,

$$P(y)=f(y)=\frac{e^{-\lambda} \cdot \lambda^y}{y!}, \lambda > 0 \quad (50)$$

olarak gösterilebilir. y tesadüfi deęişkeni 0,1,2,3... gibi tam sayı deęerlere sahiptir. Formüldeki λ daęılımın parametresidir. Poisson daęılımının ortalaması ve varyansı λ eşittir. Bu parametre, birimler arasındaki gözlemlenmiş heterojenlik ile ifade edilen regresyon modelinin açıklayıcı deęişkenleri x' le deęişebilir⁷⁹. Poisson regresyon modeli,

$$\lambda = e^{-\beta} \quad (51)$$

ifade edilmesi mümkündür. β , λ 'daki x'in etkilerini ölçen bir vektördür. Modelin üstel fonksiyon olarak gösterilmesinin nedenlerinden biri, λ 'nın pozitif olmasının garantisi olabilir.

Poisson regresyon modelinin kullanılmasındaki en büyük kısıtlayıcı tek bir parametreye sahip olmasıdır. Bu sebeple ortalama ve varyans arasındaki oran "1"e

⁷⁸ William H. Greene, **Limdep Version 7.0 User's Manual**, Australia: Econometric Software, Inc, 1995, s. 545.

⁷⁹ Peter Berkhout ve Erik Plug, "A Bivariate Poisson Count Data Model Using Conditional Probabilities", **Statistica Neerlandica**, No: 58, S: 3, 2004, s. 350-351.

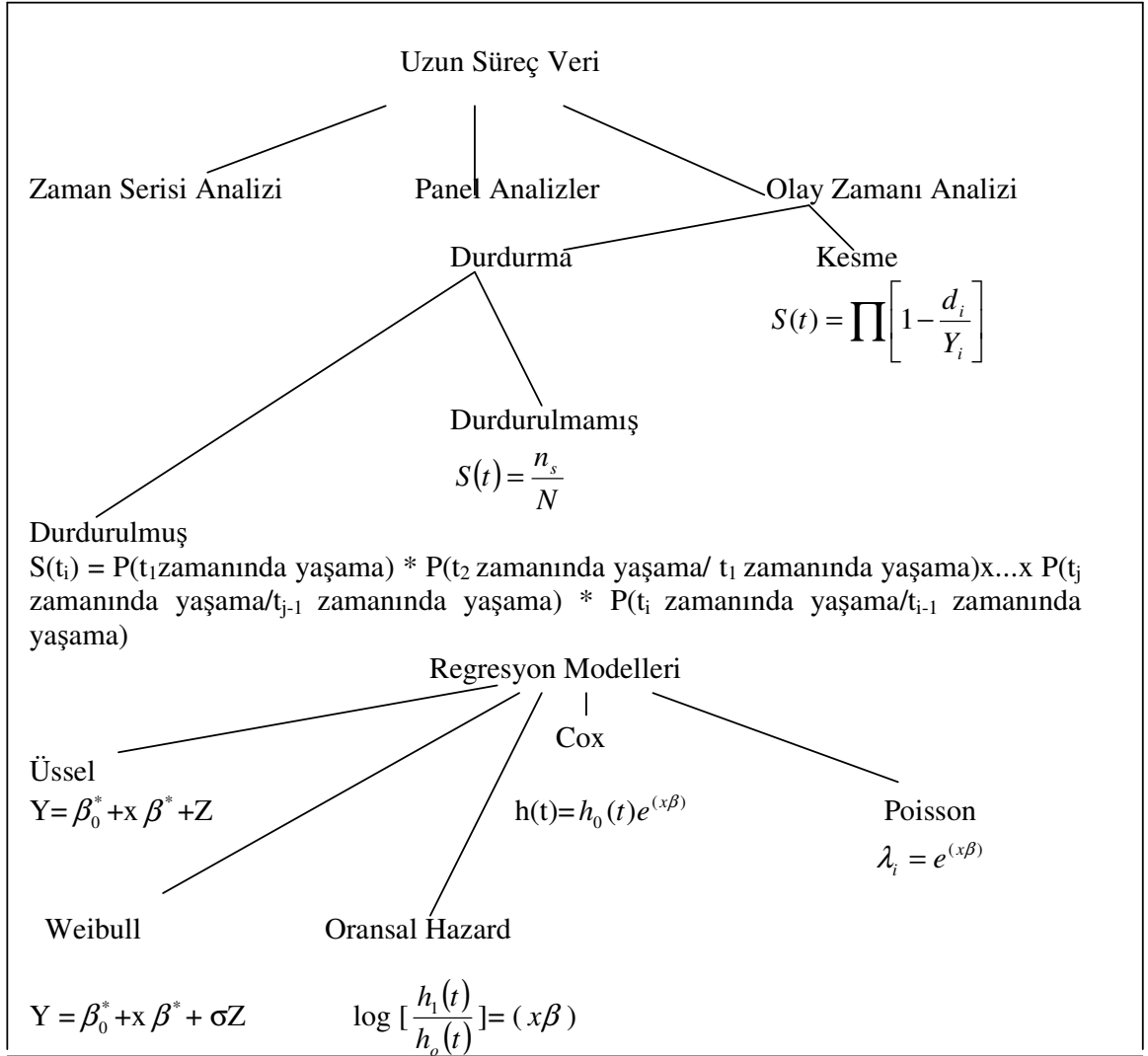
eşittir. Modeldeki bu durum eş yayılım (equidispersion) anlamına gelmektedir. Araştırmada kullanılan verilerde, varyans ve ortalamanın birbirine eşit olduğu çok nadir olarak görülebilir. Veriler poisson regresyon modelinin bu varsayımını sağlamıyorlarsa, bu durumun iki kaynağının olduğu düşünülebilir.

Sorunun ilk kaynağı olarak; başarılı olayların bağımsız olduğu poisson varsayımı altında ilişkilendirilir ve yanlış belirleme olduğu düşünülür. Bağımsızlık koşulunun sağlanamadığı durumda, sayılabilir verilerin analiz edilmesinde poisson modeli istatistiksel olarak bir araç olarak görülmeyebilir. Sorunun ikinci kaynağı ise, gözlemlenemeyen heterojenlik olabilir. Ortalamanın varyansı aştığı durumlarda aşırı yayılım (overdispersion) söz konusu olabilir. Modele hata teriminin (ε) eklenmesiyle, bu sorunun üstesinden gelinebilir.

$$\lambda = e^{x\beta + \varepsilon} \quad (52)$$

λ ve ε için eşzamanlı dağılımın seçiminde model eşitlik (51)'de gösterildiği gibi kapalı formda yazılabilir. Bu model poisson-gamma veya diğer ismiyle negatif binom modelidir.

Olay zamanı analizinde kullanılan yöntemler araştırmanın amacına ve kullanılan veri yapısına göre farklılıklar gösterebilir. Şekil 4'de veri yapısına göre kullanılacak olan yöntemler görülmektedir.



Şekil 4. Uzun Süreç Veri Yapısı ve Olay Zamanı Analizi Yöntemleri

Şekil 4’de görüldüğü gibi veri tiplerine uygun olarak analiz yöntemleri seçilebilir. Olay zamanı analizi uygulamalarında, veriler için durdurma ve kesme sözkonusu olduğunda farklı teknikler kullanılmaktadır. Araştırma için bir regresyon modellenmesi yapılaması uygun görülürse, araştırmanın amacına ve veri yapısına uygun olarak üssel, weibull, oransal hazard, cox ve poisson regresyon modellerinden biri seçilerek analize devam edilir.

3. POISSON REGRESYON

Genelleştirilmiş lineer modeller gözlemlerin olasılık dağılımları kullanılarak oluşturulmaktadır. Kullanılan bağ fonksiyonu ise regresyon parametrelerinin ortalamalarıyla ilişkilidir. Bağ fonksiyonu gözlemlerin ortalamaları üzerinde şartlı olarak istatistiksel bağımsız varsayılmaktadır. Standart regresyon modeli, normal dağılıma sahip gözlemlerin özdeşlik bağı ile (identity link) genelleştirilmiş doğrusal modelidir. Sayılabilir veriler için log-linear model ise, poisson dağılımı ve logaritmik bağ ile genelleştirilmiş lineer model olmaktadır. Lojistik regresyon modeli, lojit bağ ve binomial dağılımlı genelleştirilmiş doğrusal modeldir. Bu üç özel durumda, esnek teoriksel özelliklere sahip olan standart genelleştirilmiş doğrusal modeller kullanılmaktadır⁸⁰.

Lojistik regresyon, Poisson regresyon ve Cox regresyon, ikili olay serileri için olasılık modelleri olarak görülebilir. Poisson ve Cox regresyon için şartlı olay olasılıkları küçüktür ve risk büyüktür. Buna rağmen, her iki analiz çoğu çalışmada durdurulmuş açıklayıcı değişken bilgileri içermektedir⁸¹.

Poisson regresyon, incelenmek istenen zaman periyodundaki meydana gelmiş olan olaylarda belirlenen açıklayıcı değişkenlerde uygulanmaktadır. Model; süresiz ve negatif olmayan, sayılabilir verileri içermelerinden dolayı, beklenen sayıların logaritmasının açıklayıcı değişkenlerin doğrusal bir fonksiyonu olduğu varsayımından yola çıkmaktadır. Özel açıklayıcı değişkenlerin regresyon katsayıları, diğer tüm açıklayıcı değişkenlerin sabit tutulduğu durumda açıklayıcı değişkendeki bir birimin artışından önce ve sonra beklenen sayıların oranının logaritması olarak yorumlanabilir.

⁸⁰ Robert Haining, **Spatial Data Analysis: Theory and Practice**, West Nyack: Cambridge University Press., 2003, s. 317-318.

⁸¹ Juni Palmgren ve Samuli Ripatti, "Fitting Exponential Family Mixed Models", **Statistical Modelling**, No: 2 , 2002, s. 23-25.

Doğrusal, Lojistik ve Poisson regresyon modelleri genelleştirilmiş lineer modellerdir. Poisson dağılımı üstel ailesinin özel bir durumudur. Bu ailenin diğer dağılımları ise gauss veya normal dağılım, binomial dağılım ve iki parametrelili gamma dağılımıdır.⁸² .

Lojistik regresyon modeli, normallik ve ortak kovaryansa sahip olma gibi varsayımların sağlanamaması durumunda diskriminant analizi ve çapraz tablolara bir alternatif model olmaktadır. Modelde bağımlı değişkenin ikili (binary) yada ikiden çok düzey (polychotomous) içeren kesikli değişken olması durumunda normallik varsayımı kısıtlayıcısı olmaması nedeniyle kullanım esnekliği sağlamaktadır. Bazı durumlarda ise diskriminant analizi lojistik regresyon analizinden daha iyi sonuçlar vermektedir. Çinko⁸³, un çalışması, kredi riski modelinde bootstrap tekniğinin kullanıldığı diskriminant ve lojistik regresyon analizlerinin karşılaştırılmasını içermektedir. Bootstrap tekniği ile yapılan diskriminant analizi içerdiği bir çok kısıtlayıcı varsayıma rağmen lojistik regresyona göre daha iyi sonuç vermiştir.

Doğrusal regresyon modelinde, bağımlı değişkenin sürekli olma koşulu sağlanmalıdır. Lojistik regresyonda ise bağımlı değişken 0-1 arasında sınırlı olasılık değerleri almaktadır. Açıklayıcı değişkenin sınırsız değerler alması nedeniyle söz konusu eşitlik her zaman sağlanamamaktadır. Böyle bir durumla karşılaşılması için, en iyi çözüm sonuç değeri olarak bağımlı değişkenin olasılık değerinin çeşitli dönüşümlerle $-\infty$, $+\infty$ arasında tanımlı hale getirilmesidir. Bu amaçla yapılan lojit dönüşümde modeldeki olasılık değerleri üzerinde,

$$\frac{P}{1-P} \quad (53)$$

⁸²Diggle, Liang ve Zeger, s. 242-243.

Daniel Zelterman, **Models for Discrete Data**, Oxford: Clarendon Press, 1999, s. 138-140.

⁸³Murat Çinko, “*Diskriminant ve Lojistik Regresyon Analizlerinde Bootstrap Tekniğinin Kullanımı ve Kredi Riski Modeli Oluşturulması.*”, MÜ Sosyal Bilimler Enstitüsü (Yayınlanmamış Doktora Tezi), İstanbul, 2003, s. 85.

dönüşümü yapılarak sonuç değerinin sınırları $0, +\infty$ tanımlı hale getirilmektedir. Bu oranın doğal logaritmasının alınmasıyla bağımlı değişkenin sınırları $-\infty, +\infty$ tanımlı hale gelmektedir⁸⁴.

Poisson regresyon analizi, bir zaman periyodu içerisinde sayılabilen gözlem değerleri için kullanılmaktadır. Bu analiz, zaman içerisinde düzenli olarak gerçekleşen olayların oranlarının hesaplanılmasında kullanılır. Örneğin aylık başarısızlık verisinin bir ekipmanın anakütlesine ilişkin olarak toplandığı bir çalışma yapılmaktadır. Bu çalışmada analiz, sezon ve diğer faktörlerin başarısızlık üzerindeki etkilerini ortaya çıkarmak için kullanılmaktadır. Bunun yanında araştırmacı eğer bütün bir yılı gözlemeleme periyodu içerisine dahil ederse mevsimsel etkiyi ortadan kaldırabilir⁸⁵.

Poisson regresyonun bir başka kullanım alanı ise tekrarlanmış süreçlerdir. Tekrarlanan süreçlerde bozulan birim veya makineler tamir edildikten sonra tekrar kullanılabilir. Tekrarlanan süreçlerdeki heterojenlik olarak isimlendirilen bu durumda, Weibull dağılımı kullanılmasıyla azalan hazard oranı ile karşılaşılmaktadır. Hazard oranındaki azalma sorununu ortadan kaldırmak için poisson dağılımı ve poisson-gamma dağılımı kullanılarak çözüm üretilmeye çalışılmaktadır⁸⁶.

Brouhns, Denuit ve Vermunt⁸⁷, demografide ölüm değişimlerinde kullanılan “Lee Carter” modelini poisson regresyon modeliyle geliştirmişlerdir.

3.1 Poisson Dağılımı

Poisson dağılımı belirli bir zaman, belirli bir alan veya hacimde gerçekleşebilecek olay sayısını tanımlayan bir dağılımdır. Dağılım 18. yüzyılda yaşamış

⁸⁴ Hüseyin Tatlıdil, **Uygulamalı Çok Değişkenli İstatistiksel Analiz**, Ankara: Cem Web Ofset Ltd. Şti., 1996, s. 289-292.

⁸⁵ Wayne Nelson, **Applied Life Data Analysis**, New York: Wiley, 1981, s. 445.

⁸⁶ Dean A. Follmann ve Mathew S. Goldberg, “*Distinguishing Heterogeneity from Decreasing Hazard Rates*”, **Tecnometrics**, No: 30, S: 4, Kasım 1988, s. 389-390.

⁸⁷ Natacha Brouhns, Michel Denuit ve Joren K. Vermunt, “*A Poisson Log-Bilinear Regression Approach to The Construction of Projected Lifetables*”, **Insurance: Mathematics and Economics**, No: 31, 2002, s. 375-376.

olan Fransız matematikçi Poisson tarafından ilk olarak çalışılmıştır. Sonraki yıllarda farklı bilim adamları tarafından üzerinde çalışılan ve genişletilerek bugünkü halini alan dağılım ilk temellerini atan ünlü matematikçinin adını taşımaktadır⁸⁸. Poisson dağılımı, nadiren ortaya çıkan olayların olasılıklarının hesaplanmasında kullanılır. Bu özelliğinden dolayı, küçük olasılıklar dağılımı olarak da bilinmektedir. İki sonuçlu olaylara uygulanan bir dağılım olan poisson dağılımı, binom dağılımı koşullarında başarı olasılığının (p) küçüldüğü, deneme sayısının (n) büyüdüğü durumlardaki yaklaşımıdır.

İki sonuçlu olaylarda, denemelerin birbirinden bağımsız olduğu, denemelerde başarı ve başarısızlık olasılıklarının değişmediği durumlarda, başarı sayısını gösteren y tesadüfi değişkeninin dağılımı poisson dağılımıdır. Poisson dağılımını genel olarak aşağıdaki gibi göstermek mümkündür.

$$P(y;\lambda) = \frac{e^{-\lambda} \cdot \lambda^y}{y!}, y= 0,1,2,\dots \text{ için}$$

$$=0, \text{ diğer } y \text{ değerleri için.} \quad (54)$$

Poisson dağılımı tek parametrelili bir dağılımdır. Dağılımın parametresi λ 'dır. Dağılımın parametresi olan $\lambda >0$ olmalıdır⁸⁹.

Belirli bir zaman aralığında gözlemlenen olay sayısı y'nin poisson rassal değişkeni olarak belirlenebilmesi için aşağıdaki şartları taşıması gerekmektedir⁹⁰.

1. Farklı zaman aralıklarında veya farklı alanlarda ortaya çıkan olaylar bağımsızdır. Olayların meydana gelmeleri arasında hiçbir ilişki bulunmamaktadır.
2. Küçük bir zaman aralığında olay bir defa meydana gelebilmektedir. Olayın birden fazla ortaya çıkması mümkün olmamaktadır.

⁸⁸ Necmi Gürsakal, **Bilgisayar Uygulamalı İstatistik 1**, Bursa: Marmara Kitabevi, 1997, s. 238.

⁸⁹ Selahattin Güriş ve Şahamet Bülbül, **Olasılık**, İstanbul: Yayılım Matbaası, 1995, s. 422-429.

⁹⁰ Özer Serper, **Uygulamalı İstatistik 1**, Bursa: Ezgi Kitabevi, 2000, s. 296.

3. Küçük bir zaman aralığında ilgilenilen olayın ortaya çıkma olasılığı (p) değişmemektedir. Bu durumda $p < 0,05$ eşitsizliği sağlanmaktadır.

4. Deney sayısı sonsuza yaklaşmaktadır. $n \rightarrow \infty$ olarak ifade edilebilir.

5. Belirli bir zaman aralığında ilgilenilen olayın ortaya çıkma sayısı (λ) sabittir. Ortalama niteliğindeki λ kesirli bir değer olarak bulunabilir.

Poisson dağılımının ortalaması,

$$E(y) = \lambda \quad (55)$$

olmaktadır. Varyansı ise ,

$$E[(y - E(y))^2] = \lambda \quad (56)$$

olarak ifade edilmektedir⁹¹.

3.2 Poisson Regresyon Modeli

Poisson regresyon modeli; sayılabilir verilerin ortalamasını içeren çarpımsal modellerde, açıklayıcı değişkenler arasında şartlı veya marjinal bir bağımlılık olduğunda esnek bir model olmasından dolayı araştırmacılara kolaylık sağlayan bir model olmaktadır⁹².

$Y = 1, 2, 3, \dots$ tesadüfi değişkeni bir zaman diliminde olayın gerçekleşme sıklığını ifade etmektedir. $y(t+\Delta t)$, $(t+\Delta t)$ kısa zaman aralığında gözlemlenen olay sayısını göstermektedir. Belirtilen aralıktaki yaşanan olay sayısının olasılık fonksiyonu

$$P(Y=y) = f(y; \lambda) = e^{-\lambda} \left[\frac{\lambda^y}{y!} \right] \quad y=0, 1, 2, \dots, \lambda > 0 \quad (57)$$

⁹¹Paul Newbold, **İşletme ve İktisat İçin İstatistik**, Çev. Ümit Şenesen, İstanbul: Literatür yayıncılık, 2000, s. 181.

⁹²Chris J.Lloyd, **Statistical Analysis of Categorical Data**, New York: John Wiley ve Sons, 1999, s. 298.

şeklinde ifade edilmiştir.

Δt zaman aralığında λ pozitif bir değer ise aşağıdaki durumlar ortaya çıkmaktadır⁹³.

1. Aralık yaklaşık olarak $1-\lambda\Delta t$ eşit olduğunda sıfır olayın gerçekleşme olasılığı,
 $1-\lambda\Delta t$

olmaktadır. Artık değer olarak $r_1(\Delta t)/\Delta t$, Δt 'den sıfıra zaman aralığında sıfıra yaklaşmaktadır.

2. $\lambda\Delta t$ zaman aralığında gerçekleşen tek olayın olasılığı,
 $\lambda\Delta t + r_2(\Delta t)$

olmaktadır. Artık değer $r_2(\Delta t)/\Delta t$, Δt 'den sıfıra doğru zaman aralığında sıfıra yaklaşmaktadır.

3. İki yada daha fazla olayın zaman aralığı içinde meydana gelme olasılığı,
 $r_3(\Delta t)$

değerine eşit olmaktadır. Bununla birlikte artık değer $r_3(\Delta t)/\Delta t$, Δt değeri Δt aralığında sıfıra yaklaştığında artık değerde sıfıra yaklaşmaktadır.

Y bağımsız tesadüfi bir değişken olduğunda, Y'in N gözlemi bağımsızdır ve (y_1, y_2, \dots, y_N) değerlerini almaktadır. Genelleştirilmiş regresyon modelleri için tanımlanan eşitlik aşağıdaki gibi olmaktadır.

$$f(y_i; \theta_i, \phi) = \exp\{[y_i \theta_i - b(\theta_i)]/a(\phi) + c(y_i, \phi)\}. \quad (58)$$

θ parametresi doğal parametre olarak isimlendirilmektedir. $a(\phi)$ fonksiyonu ω ile ağırlıklandırılmış olan $a(\phi) = \phi/\omega$ bir yapıya sahip olmaktadır. ϕ dağılım (dispersion) parametresi olarak isimlendirilmektedir.

Modelin ortalaması ve varyansı aşağıdaki gibi olmaktadır.

$$\lambda_i = E(Y_i) = b'(\theta_i) \quad (59)$$

$$\text{Var}(Y_i) = b''(\theta_i) a(\phi) \quad (60)$$

⁹³Land, Nagin ve McCall, s. 351.

λ ortalama sahip olan poisson dağılmış bir Y_i tesadüfi değişkeni için genelleştirilmiş lineer model aşağıdaki gibi olmaktadır.

$$\begin{aligned}
 f(y_i; \lambda_i) &= \frac{e^{-\lambda_i} \cdot \lambda_i^{y_i}}{y_i!} \\
 &= \exp[y_i \log(\lambda_i) - \lambda_i - \log(y_i!)] \\
 &= \exp[y_i \theta_i - \exp(\theta_i) - \log(y_i!)] \tag{61}
 \end{aligned}$$

(61)'nci Eşitlikte θ_i değeri $\theta_i = \log(\lambda_i)$ olarak ifade edilmiştir. Eşitlik (58)'deki $b(\theta_i) = \exp(\theta_i)$, $a(\phi) = 1$ ve $c(y_i, \phi) = -\log(y_i!)$ olmaktadır. Doğal parametre $\theta_i = \log(\lambda_i)$ olarak gösterilmiştir. Modelin ortalaması,

$$E(Y_i) = b'(\theta_i) = \exp(\theta_i) = \lambda_i \tag{62}$$

ve varyansı da,

$$\text{Var}(Y_i) = b''(\theta_i) = \exp(\theta_i) = \lambda_i \tag{63}$$

olarak gösterilebilir.

$g(\lambda_i) = \theta_i$ olarak tanımlanırsa “g” bir bağ fonksiyonu olarak isimlendirilir. Bu durumda kanonik bağ, log bağ olarak

$$\eta_i = \log(\lambda_i) \tag{64}$$

şeklinde ifade edilebilir. Log lineer model için bağı fonksiyonunu, aşağıdaki eşitlikle göstermek mümkündür⁹⁴.

$$\log \lambda = X\beta. \quad (65)$$

Poisson regresyon modelinin formüle edilmesinde üç önemli nokta göz önünde bulundurulur. Bunlardan ilki; olaylar sabit bir zaman aralığı içinde meydana geliyorsa, bu durumda ilgilenilen sabit zaman aralığı içinde olayların tekrar sayısı önem kazanır. Bu ilk durumda, olayların sabit zaman aralığında ortaya çıkış sayıları bağımlı değişken olarak ele alınır.

y_i ; N birimden i 'nci birim için, verilen zaman aralığında meydana gelen olay sayısını ifade etmektedir. Her i birey, oran parametresi λ ve x ($i=1, \dots, m$) açıklayıcı değişkenlerle içermektedir. Bu durumda klasik poisson regresyon denklemi aşağıdaki gibi olmaktadır

$$\lambda_i = \exp(x_i \beta) \quad (66)$$

$$\log \lambda_i = x_i \beta. \quad (67)$$

Burada x_i , m açıklayıcı değişkenlerin X ile örtüşen m sütun açıklayıcı matrisinin i 'nci satırındaki i 'yi açıklamaktadır. β ise modeldeki regresyon katsayılarının birinci kolon vektörünün m satırındadır. $\log \lambda_i$, daha çok λ_i olarak kullanılır ve her $\lambda_i > 0$ için x_i 'e lineerdir. Eşitlik (57)'deki poisson dağılımının ortalama oran parametrelerini kullanılarak geliştirilmesi eşitlik (65) ve (66)'deki gibidir. Bu durum regresyon veya şartlı beklenti fonksiyonu (regression or conditional expectation function) olarak ifade edilebilir. $E[y_i / x_i]$ gösterimi; i 'nci birey için gözlemlenmiş açıklayıcı değişkenlerin

⁹⁴ Alan Agresti, **Categorical Data Analysis**, New York: Wiley, 1990, s. 446-453.

özel değerleri üzerinde şartlı olan her bir i birey için, her birim zamanda bağımlı ve sayılabilir değişkenlerin beklenen değerlerini vermektedir⁹⁵.

Poisson regresyonun modellenmesinde göz önünde bulundurulmuş ikinci durum, birinci durumda gerçekleşen olay zamanı verilerine benzerlik göstermektedir. Farklı zaman noktalarında gerçekleşen olaylar için olayın ortaya çıkma sayısı önem kazandığı çalışmalarda kullanılmaktadır. Olayın ortaya çıkma sayısı bağımlı değişken olarak kabul edilirken, olayın ortaya çıkma uzunluğunun da önemli olduğu kabul edildiği için zaman değişkeni olarak modelde yer almaktadır. Bu tür veriler için poisson regresyon modelini aşağıdaki gibi göstermek mümkün olmaktadır⁹⁶.

$$\log (E(y_i)) = \log t_i + \beta x_i \quad (68)$$

Poisson regresyonun modellenmesinde göz önünde bulundurulmuş üçüncü durum ise, olayların birbirlerinden farklı zamanlarda ortaya çıkmalarıdır. Bu olaylar, aynı zamanda kümelenmiş de olabilirler. Bu durumda, farklı zamanlarda ortaya çıkan olayların tekrar sayıları ile değil ortaya çıkma uzunluklarıyla ilgilenilir. Olay zamanı verileri olarak isimlendirilen bu tür verilerin poisson regresyon ile modellenmesi aşağıdaki gibi olmaktadır.

Olay sayısı t zamana kadar gözlemlenmiştir ve $t \in (0, T_i]$. Böylece i birey T_i zamanı için takip edilir. T_i 'nin, $Y_i(t)$ olarak ifade edilen sayılabilir süreçten bağımsız olduğu varsayılmaktadır. Devam eden süreç ve T_i 'nin dağılımı hakkında güçlü varsayımlar varsa bağımsızlık varsayımı aranmaz. Bu durum r 'nci başarısızlıkta durdurma gibi, daha çok genel geri çekilme planlarına izin vermektedir.

⁹⁵ Brian Everitt, **Statistical Aspects of the Design and Analysis of Clinical Trials**, London: Imperial College Press, 2004, s. 76-78.

⁹⁶ Alex Pedan, **Analysis of Count Data Using the SAS System**, <http://www2.sas.com/proceedings/sugi26/p247-26.pdf>, (3 Ekim 2005).

Olay zamanı verileri için poisson regresyon modeli aşağıdaki gibi olmaktadır⁹⁷.

$$\lambda_i(t) = e^{x_i\beta}. \quad (69)$$

3.2.1 Modelin En Çok Benzerlik Tahmincileri

Log benzerlik fonksiyonu genelleştirilmiş regresyon modeli olan eşitlik (58) kullanılarak elde edilmektedir. N bağımsız gözlem için log benzerlik aşağıdaki gibi olmaktadır⁹⁸.

$$L(\beta) = \sum_i L_i = \sum_i \log f(y_i; \theta_i; \phi) = \sum_i \frac{y_i \theta_i - b(\theta_i)}{a(\phi)} + \sum_i c(y_i, \phi) \quad (70).$$

$L(\beta)$, modelin β parametresi üzerinde θ 'nın bağımsızlığını ifade etmektedir. Benzerlik eşitliği bütün j için,

$$\partial L(\beta) / \partial \beta_j = \sum_i \frac{\partial L_i}{\partial \beta_j} = 0 \quad (71)$$

olmaktadır. Log benzerlik fonksiyonunun zincir kuralına göre diferansiyeli alındığında,

$$\frac{\partial L_i}{\partial \beta_j} = \frac{\partial L_i}{\partial \theta_i} \frac{\partial \theta_i}{\partial \lambda_i} \frac{\partial \lambda_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_j} \quad (72)$$

⁹⁷ C.B. Dean ve R. Balshaw, "Efficiency Lost By Analyzing Counts Rather Than Event Times in Poisson and Overdispersed Poisson Regression Models", **Journal of the American Statistical Association**, No: 92, S: 440, Aralık 1997, s. 1387.

M. Zhou, D. Yang, Y. Wang ve S. Nadarajah, "Conditional Poisson Distributions", **Probability in the Engineering and Informational Sciences**, No: 20, 2006, s. 95-102.

⁹⁸ Alan Agresti, **Categorical Data Analysis**, New York: Wiley, 2002, s. 135-136.

ifade edilebilir.

$$\frac{\partial L_i}{\partial \theta_i} = \frac{[y_i - b'(\theta_i)]}{a(\phi)} \text{ olduğundan}$$

(54) ve (55) eşitliklerdeki $\lambda_i = b'(\theta_i)$ ve $\lambda_i = b''(\theta_i)$ kullanılarak

$$\frac{\partial L_i}{\partial \theta_i} = \frac{(y_i - \lambda_i)}{a(\phi)} \quad (73)$$

$$\frac{\partial \lambda_i}{\partial \theta_i} = b''(\theta_i) = \frac{Var(Y_i)}{a(\phi)} \quad (74)$$

olmaktadır. $\eta_i = \sum_j \beta_j x_{ij}$ kullanıldığında ise

$$\frac{\partial \eta_i}{\partial \beta_j} = x_{ij} \quad (75)$$

olarak yazılabilir. $\eta_i = g(\lambda_i)$ olduğundan, modeli için bağ fonksiyonu $\frac{\partial \lambda_i}{\partial \eta_i}$ değerine bağlı

olduğundan eşitliği aşağıdaki gibi yazmak mümkündür.

$$\frac{\partial L_i}{\partial \beta_j} = \frac{(y_i - \lambda_i)}{a(\phi)} \frac{a(\phi)}{Var(Y_i)} \frac{\partial \lambda_i}{\partial \eta_i} x_{ij} \quad (76)$$

Genelleştirilmiş regresyon modeli için benzerlik eşitliği aşağıdaki gibi olmaktadır.

$$\sum_{i=1}^N \frac{(y_i - \lambda_i)x_{ij}}{Var(Y_i)} \cdot \frac{\partial \lambda_i}{\partial \eta_i} = 0, j=1, \dots, t. \quad (77)$$

Benzerlik eşitliği β 'nın doğrusal olmayan bir fonksiyonudur. Çözümlemede $\hat{\beta}$ için iteratif metotlara gereksinim duyulmaktadır. $\hat{\beta}$ 'nin β 'ya yaklaşım oranı bilgi matrisine (information matrix) bağlıdır.

$$E\left(\frac{\partial^2 L_i}{\partial \beta_h \partial \beta_j}\right) = -E\left(\frac{\partial L_i}{\partial \beta_h}\right)\left(\frac{\partial L_i}{\partial \beta_j}\right) \quad (78)$$

$$\begin{aligned} &= -E\left[\frac{(Y_i - \lambda_i)x_{ih}}{Var(Y_i)} \frac{\partial \lambda_i}{\partial \eta_i} \frac{(Y_i - \lambda_i)x_{ij}}{Var(Y_i)} \frac{\partial \lambda_i}{\partial \eta_i}\right] \\ &= \frac{-x_{ih}x_{ij}}{Var(Y_i)} \left(\frac{\partial \lambda_i}{\partial \eta_i}\right)^2 \end{aligned}$$

$$E\left(\frac{\partial^2 L(\beta)}{\partial \beta_h \partial \beta_j}\right) = -\sum_{i=1}^N \frac{x_{ih}x_{ij}}{Var(Y_i)} \left(\frac{\partial \lambda_i}{\partial \eta_i}\right)^2 \quad (79)$$

$E\left[-\frac{\partial^2 L(\beta)}{\partial \beta_h \partial \beta_j}\right]$ ögesine sahip bilgi matrisi

$$\text{Inf} = X'WX \quad (80)$$

eşit olmaktadır. W , temel diyagonal üzerinde

$$w_i = (\partial \lambda_i / \partial \eta_i)^2 / \text{Var}(Y_i) \quad (81)$$

ögesi ile diyagonal matris olmaktadır⁹⁹. Log olabilirlik fonksiyonunun ikinci derece türevlerinden elde edilen matrise bilgi matrisi adı verilmekte ve kovaryans matrisi özelliklerini göstermektedir.¹⁰⁰

Poisson dağılımına sahip Poisson regresyon modeli için eşitlik (78) kullanılarak parametre tahmincilerinin elde edilmesi mümkündür.

$$\sum_i (y_i - \lambda_i) x_{ij} = 0. \quad (82)$$

$$w_i = (\partial \lambda_i / \partial \eta_i)^2 / \text{Var}(Y_i) = \lambda_i \quad (83)$$

$\hat{\beta}$ 'nin tahmini kovaryans matrisi $(X' \hat{W} X)^{-1}$ olmaktadır. \hat{W} , $\hat{\lambda}$ 'nin temel diyagonal olduğu diyagonal matristir. Bu durumda Poisson log benzerlik aşağıdaki gibi olmaktadır¹⁰¹.

$$L = \sum [y_i \log(\lambda_i) - \lambda_i] \quad (84)$$

⁹⁹ Agresti, s. 447-449.

¹⁰⁰ Tathdil, s. 307.

¹⁰¹ Agresti, s. 454.

Bary Arnold, **Conditional Specification of Statistical Models**, New York: Springer-Verlag, 1999, s. 217-218.

3.2.2 Modelin Artıkları

Regresyon çözümlemesindeki model bozukluklarını arařtırmak ve aykırı deęerleri belirlemek için kullanılan en etkin yöntemlerden biri artıkların incelenmesidir. Artıklar, hataların gerçekleşen yada gözlenen deęerleri olarak düşünülebilir. Bu nedenle hatalar için söz konusu olan varsayımlara ilişkin bozulmalar artıklar yardımıyla incelenebilir. Bu incelemede, işlenmemiş artıklar (raw residuals) ve standartlaştırılmış artıklar (Standardized residuals) olmak üzere iki tür artıktan yararlanılmaktadır.¹⁰²

Genelleştirilmiş regresyon modellerinde model seçimi için, varyans analizindeki kareler toplamına benzer yöntemle sapmalar kullanımı önerilmektedir. Sapma analizi (deviance of analysis) olarak isimlendirilen bu analiz asimptotik dağılım teorisini kullanmaktadır. Bilindięi gibi asimptotik dağılım teorisi: uygun bir serbestlik derecesiyle χ^2 dağılımı gösteren sapmalar arasındaki farklılık varsayımıyla açıklanmaktadır. Uyum iyilięi testlerinde, verilerin ve modelin uygunluęunun her ikisinde yapılmasında artıkların analizi önemli yer tutmaktadır. Uygun bir model olarak elde edilmiş lineer regresyondaki artıklar, farklı varyansa sahip ve korelasyonludurlar¹⁰³.

Poisson regresyon modelinin işlenmemiş artıkları; gözlem deęerleri ile uygunlaştırılmış deęerler (observed and fitted values) arasındaki fark,

$$y_i - \hat{\lambda}_i$$

olarak ifade edilebilir¹⁰⁴. “D*” olarak gösterilen sapma artıklar, log benzerlik oran istatistięi (log-likelihood-ratio-statistic)

$$D^* = 2[L_{\max} - L(\hat{\beta})] \quad (85)$$

¹⁰² Reha Alpar, **Uygulamalı Çok deęişkenli İstatistiksel Yöntemlere Giriş 1**, Ankara: Nobel Yayın Dağıtım, 2003, s. 266.

¹⁰³ P. J. Green ve B. W. Silverman, **Nonparametric Regression and Generalized Linear Models**, London: Chapman ve Hall, 1995, s. 96.

¹⁰⁴ Lloyd, s. 306.

olarak tanımlanabilir. Model ϕ parametresi içerdiğinden artıklar ölçeklendirilmeden de kullanılabilir. Poisson dağılımı için $\phi=1$ olduğundan,

$$D = \phi D^* = 2 \sum [y_i \log(y_i / \lambda_i) - y_i + \lambda_i] \quad (86)$$

şeklinde ifade edilebilir¹⁰⁵.

Artıkların modellenmesi log-ölçekleme (log-scale) kullanılarak yapılmaktadır. Bu durum aşağıdaki gibi olmaktadır.

$$\text{Log}(y_i) - \text{log}(\hat{\lambda}_i) \approx \frac{y_i - \hat{\lambda}_i}{\hat{\lambda}_i}. \quad (87)$$

Sağ taraf, birinci sıra Taylor dağılımıdır (first-order Taylor expansion) ve $y_i=0$ olduğu zaman anlamlı olmaktadır. Paydadaki $\hat{\lambda}_i$ değeri kareköklü bir ifade değildir. Tahminciler göre işlenmemiş artıkların grafiklerinde, modelin sistematik parçalarındaki hatalar ortaya çıkabilir¹⁰⁶.

Modeldeki artıklar için alternatif olan bir yaklaşım Pearson χ^2 istatistiği olmaktadır. Pearson artıkları,

$$e_i = \frac{y_i - \hat{\lambda}_i}{\sqrt{\hat{\lambda}_i}} \quad (88)$$

olarak ifade edilebilir. Model doğru olduğunda artıkların çoğunluğu ± 2 arasında olmaktadır. e_i 'nin büyüklüğü, e_i 'nin $\hat{\lambda}_i$ 'den farklılığına bağlı olarak artmaktadır. e_i 'nin

¹⁰⁵ Green ve Silverman, s. 96.

¹⁰⁶ Lloyd, s. 306.

değeri büyük olduğunda, modelin verilere zayıf bir uygunluk gösterdiğinin işareti olmaktadır¹⁰⁷. Pearson artıklarının $n_i \rightarrow \infty$ varsayımı altında varyansı “1” olmaktadır. Artıklar asimptotik olarak $N(0,1)$ olarak dağılım göstermektedirler¹⁰⁸.

$$\chi^2 = \phi \sum_{i=1}^n \frac{\{y_i - E(y_i)\}^2}{\text{var}(y_i)} = \sum_{i=1}^n \frac{(y_i - \hat{\lambda}_i)^2}{\hat{\lambda}_i} \quad (89)$$

olarak gösterilebilir Artıkların hipotez testleri Pearson artıkları ve sapma artıkları için ayrı ayrı yapılabilmektedir. Sapma artıkları G^2 istatistiği kullanılarak test edilirken, Pearson artıkları χ^2 dağılımına sahip olduğundan χ^2 testi kullanılmaktadır¹⁰⁹.

3.2.3 Aşırı Yayılımın Test Edilmesi

Poisson veya çoklu nominal dağılımlarda büyük varyansla karşılaşılabilir. Örneklem birimleri arasında pozitif korelasyon bulunması bağımsızlık varsayımını ihlal etmektedir. Bu durum için temel neden örneklemin küme-yapısıdır. Örneklemdeki büyük değişkenliğin varlığı, örneklem modelinde aşırı yayılım (overdispersion) olarak isimlendirilir. Bir başka ifadeyle aşırı yayılım, gözlemlerin gerçek varyansının, gerçek dağılımın nominal varyansını geçmesi olarak tanımlanabilir. Genel olarak aşırı yayılımın ortaya çıkış sebebinin, kümelenmiş verilerde küme yapısı ile ilgili olduğu düşünülür. Başarı oranlarının tesadüfiliği, korelasyonun varlığı olarak yorumlanabilir. Küme içindeki skorların kümeler arasındakine göre daha çok birbirine

¹⁰⁷ William D. Dupont, **Statistical Modeling for Biomedical Researches: A Simple Introduction to the Analysis of Complex Data**, West Nyack, NY, USA: Cambridge University Press, 2002, s. 315-316.

¹⁰⁸ Clive Loader, **Local Regression & Likelihood**, New York: Springer- Verlag, 1999, s. 67.

¹⁰⁹ Green ve Silverman, s. 96-97.

benzediği görülebilir. Böyle bir durumda karma dağılımlar göz önünde bulundurulabilir¹¹⁰.

Aşırı yayılım basit olarak, $\phi > 1$ sabitiyle varyansın çarpılması şeklinde modellenmektedir. Yayılım parametresi olarak ifade edilen ϕ normal dağılımda olduğu gibi $\phi = \sigma^2$ olarak biliniyordur yada örnekten tahmin edilmektedir¹¹¹.

Yayılım parametresi ϕ 'nın değeri sıfıra eşit olduğunda, Poisson regresyon modelinin etkisini azaltmaktadır. Poisson regresyon modelinin yayılımını değerlendirmek için hipotez testine ihtiyaç duyulmaktadır.

$$H_0: \phi = 1$$

$$H_1: \phi > 1$$

Yayılım parametresi ϕ 'nın regresyon modeli içinde yer alması sıfır hipotezinin ret edilip alternatif hipotezin kabul edilmesiyle olmaktadır. Bu test yapılabilmesi için asimptotik normal wald t istatistiği kullanılmaktadır. Bu değer, tahmini ϕ oranının standart hatası olarak tanımlanmaktadır. Eğer t istatistiği (-1.96, 1.96) değerlerinin dışındaysa %5 anlamlılık düzeyinde poisson regresyon modelinin ret edildiği anlamına gelmektedir¹¹². Aşırı yayılımın test edilmesi için Famoye ve Wang¹¹³ tarafından önerilen bir başka yaklaşım ise benzerlik oran test istatistiğidir. Sıfır hipotezi doğru olduğunda, test istatistiği bir serbestlik derecesi ile yaklaşık olarak χ^2 dağılımına sahip olmaktadır.

¹¹⁰ K. Poortema, "On Modelling Overdispersion of Counts", **Statistica Neerlandica**, No: 53, S: 1, 1999, s. 5-20.

¹¹¹ Radhakrishna C. Rao, **Linear Models: Least Squares and Alternatives**, New York: Springer-Verlag, 1999, s. 315-316.

¹¹² Wang Weiren ve Felix Famoye, "Modeling Household Fertilty Decisions with Generalized Poisson Regression", **Journal of Economics**, No:10, 1997, s. 280.

¹¹³ Felix Famoye ve Weiren Wang, "Censored Generalized Poisson Regression Model", **Computational Statistics&Data Analysis**, No: 46, 2004, s. 552.

3.3 Poisson-Gamma Regresyon Modeli

Sonlu bir karma dağılım, birkaç basit dağılımın ağırlıklandırılmış toplamı olarak ifade edilebilir. Karma olarak oluşturulan model çoğunlukla iki dağılım içermektedir. Karma modellerde belirli bir bağ kullanılmaktadır. Oluşturulan regresyon modeli doğrusal değildir¹¹⁴.

Poisson regresyondaki temel kısıtlamalar eşitlik (55) ve (56) belirtilen ortalamanın varyansa eşit olduğu durumlardaki gibidir. Regresyon fonksiyonunun şartlı ortalama ve varyansına uygulanmaktadır. Bir çok veri seti için bu tür bir kısıtlama çok sınırlandırıcı olmaktadır. Bu kısıtlamayı esnek hale getirmek için (66) ve (67)'nci eşitlik aşağıdaki gibi düzenlenebilir.

$$\lambda_i(t) = \exp[x_i \beta] * \exp[\varepsilon_i] \quad (90)$$

veya

$$\log \lambda_i(t) = x_i \beta + \varepsilon_i. \quad (91)$$

ε_i , hata terimi olarak ele alınmakta ve $\log \lambda_i$ 'deki açıklanamayan tesadüfiliği içermektedir. Bu hata terimi, içsel tesadüfilikler veya gözlemlenemeyen açıklayıcı değişkenlerin ihmal edilmesi yüzünden oluşan özel bir hata olarak eşitliğe yansiyabilir. Bu hata kavramı, poisson ortalama oran parametrelerindeki farklı heterojenlik kaynaklarını eşitliğe katmak amacıyla yapılandırılmıştır. Hazard regresyon modellerinde olduğu gibi (68)'nci eşitlikte ifade edilmiş olan geleneksel poisson regresyon modelinde de doğrudan açıklayıcı değişken etkisi ifade edilmektedir. Poisson oran parametrelerinin değişkenlerinin kaynağı, gözlemlenmiş heterojenlik yada değişkenler arasındaki farklılıklar şeklinde ifade edilebilir. Sayısı ne kadar olursa olsun

¹¹⁴ J. K. Lindsey, **Modelling Frequency and Count Data**, Oxford: Clarendon Press, 1997, s. 136.

açıklayıcı değişkenlerin sayısı (91)'nci eşitliğin sağ tarafında yer alsa da her zaman için gözlemlenemeyen yada gizli heterojenlik kaynaklarının bireysel ortalama oranını etkileme olasılığı mevcut olmaktadır. Eşitlik (90)'n özel hale getirilmesi iki tür heterojenlik için tasarlanmıştır. (91) eşitlikte hata kavramının katılmasıyla birlikte dağılımının özel hale getirilmesine ihtiyaç duyulmaktadır. ε_i ' nin olasılık yoğunluğunu $g(\varepsilon_i)$ açıklamaktadır. ε_i ' nin x_i 'den bağımsız olması varsayımına dayanarak, Y_i 'nin marjinal yoğunluğu aşağıdaki eşitlikle gösterilebilir.

$$\begin{aligned} \Pr[Y_i=y_i] &= [\Pr[Y_i=y_i \mid x_i, \varepsilon_i] g(\varepsilon_i) d\varepsilon_i \\ &= \int [\exp[\exp(x_i \beta + \varepsilon_i)]] [\exp(x_i \beta + \varepsilon_i)]^{y_i} / y_i! g(\varepsilon_i) d\varepsilon_i \end{aligned} \quad (92)$$

$g(\varepsilon_i)$ 'nin karma dağılımıyla (mixing distribution) oluşan eşitlik (90), bileşik veya karma poisson dağılımı (compound or mix poisson distribution) olarak tanımlanabilir. Eşitlik (90)'in oluşturulması, $g(\varepsilon_i)$ 'nin yada alternatif olarak $g[\exp(\varepsilon_i)]$ özel olarak seçimine bağlı olabilir. ε_i 'nin bu dağılımının kesin parametrik formu için (90)'inci eşitlikte gösterilen kapalı form deyimiyile elde edilebilir. En iyi bilinen parametrik özellik, varyansı δ 'ya eşit ve ortalaması 1 olan gamma dağılımına sahip olan $\exp(\varepsilon_i)$ 'dir.

γ ve δ parametrelili gamma dağılımı "Gamma (γ , δ)" olarak gösterildiğinde yoğunluk fonksiyonu aşağıdaki gibi ifade edilebilir¹¹⁵.

$$f(x; \gamma, \delta) = \frac{1}{\Gamma(\gamma)\delta^\gamma} x^{\gamma-1} e^{-\frac{x}{\delta}} \quad (93)$$

¹¹⁵ Güriş ve Bülbül, s. 524.

Eşitlik (92) gamma fonksiyonuna uygulandığında Poisson-Gamma karma dağılımı ortaya çıkmaktadır. Bu durum şartlı varyansa eşit olan şartlı ortalama açısından bir sınırlandırma getirmemektedir. Poisson-Gamma karma dağılımının diğer bir ismi “Negatif Binom Dağılım” olarak da ifade edilmektedir. Poisson-gamma fonksiyonun olasılık fonksiyonu eşitlik (91)’ de gösterilmektedir¹¹⁶.

$$P(y) = \frac{\Gamma(y + \alpha^{-1})}{\Gamma(y + 1)\Gamma(\gamma)} \left(\frac{\alpha^{-1}}{\mu + \alpha^{-1}} \right)^{\alpha^{-1}} \left(\frac{\mu}{\mu + \alpha^{-1}} \right)^y \quad (94)$$

Poisson-gamma dağılımın ortalaması ve varyansı eşitlik (95) ve eşitlik (96)’ de olduğu gibi gösterilmesi mümkündür.

$$E(y) = \mu \quad (95)$$

$$\text{Var}(y) = \mu(1 + \alpha\mu) \quad (96)$$

$\delta \rightarrow \infty$ değeri için poisson-gamma dağılımı Poisson dağılımına yaklaşmaktadır. Poisson-gamma regresyon modelinde $\gamma = \delta$ olarak ifade edildiğinde $E(y) = 1$ olarak gösterilmesi mümkün olmaktadır.

Bireyler arasındaki gözlemlenemeyen heterojenlik için kullanılan bir model poisson-gamma regresyon modeli olmaktadır. Bu model bireyler arasındaki diğer heterojenlik kaynaklarının modellenmesine izin vermektedir¹¹⁷.

Eğer gözlemlenemeyen heterojenlikte gamma dağılımı doğru değilse regresyon parametrelerinin en çok benzerlik tahmincilerinin istatistiksel olarak istikrarlılığını etkilememesine rağmen varyans tahmincisi olabilir. Böyle bir durum regresyon katsayısının t oranını yükseltebilir ve yanlış çıkarımlara neden olabilir. Bununla beraber daha esnek bir şekilde karma poisson regresyon modeli için özel

¹¹⁶A. Colin Cameron ve Pravin K. Trivedi, **Essential of Count Data Regression**, <http://cameron.econ.uc.davis.edu/research/CTE01preprint.pdf>, s. 7, (21 Eylül 2003).

¹¹⁷Peter GM Van Der Heijden, Rami Bustami, Maarten JLF Cruyff, Godfried Engbersen ve Hans C. van Houwelingen, “Point and Interval Estimatin of The Populatin Size Using The Truncated Poisson Regression Model”, **Statistical Modelling**, No: 3, 2003, s. 305.

karma dağılımlar ortaya konmaktadır. Poisson dağılımının gamma ile yapılan karmasından başka, birde “ters gauss” (inverse gaussien) dağılımının karmasıyla oluşan farklı bir dağılımı daha bulunmaktadır. Bu model “karma poisson-ters gauss regresyon” olarak bilinmektedir. Log normal dağılım kullanılarak yapılan karma ise “karma poisson-log normal regresyon” modeli olarak ifade edilebilir. Ayrıca en çok benzerlik tahmincisi ve karma poisson regresyonunun yarı parametrik uygunluk modelleri de gelişmektedir¹¹⁸. Land, McCall ve Nagin¹¹⁹ yarı parametrik karma en çok benzerlik tahmincisi ve bunun en çok benzerlik fonksiyonunu ortaya koymuştur. Bu modellerde hata terimi, uzun kesit örneklerinde her birey için ayrı bir şekilde düşünülür. Üzerinde çalışılan örnekleme zaman içerisinde bir farklılık göstermez. Buradaki hata teriminin dağılımı parametrik olmayan bir şekilde yaklaştırılmıştır.

3.3.1 Yarı Benzerlik Tahmincileri

Poisson-gamma regresyon modelinin parametre tahmincilerinde en çok kullanılan teknikler; yarı benzerlik tahmincileri (quasi likelihood) ve en çok benzerlik tahmincilerini elde etmede kullanılan EM algoritması olmaktadır.

Genelleştirilmiş modeller, veriler için tesadüfi bileşenli doğal üssel ailesinin bir dağılımını varsaymaktadır. Bu varsayım yerine gelmediğinde ortalama ve varyans arasındaki ilişkilerin fonksiyonunun elde edilmesinde alternatif bir yaklaşım olan yarı benzerlik kullanılabilir. Yarı benzerlik fonksiyonu yayılım parametresi ile seçilen V fonksiyonunun çarpımı olarak aşağıdaki şekilde gösterilebilir.

$$\text{Var}(y) = \phi V(\lambda) \quad (97)$$

¹¹⁸ Kenneth C. Land, Patricia L. McCall ve Daniel S. Nagin, “How Many Latent Classes of Delinquent/Criminal Careers? Results from Mixed Poisson Regression Analyses”, *AJS*, No: 103, S. 6, Mayıs 1998, s. 1599-1600; Dimitris Karlis, “A General EM Approach for Maximum Likelihood Estimation in Mixed Poisson Regression Models”, *Statistical Modelling*, No: 1, 2001, s. 309.

¹¹⁹ Kenneth C. Land, Patricia L. McCall ve Daniel S. Nagin, “A Comparison of Poisson, Negatif Binomial, and Semiparametric Mixed Regression Models”, *Sociological Methods & Research*, No: 24, S: 4, Mayıs 1996, s. 403-408.

β regresyon katsayısı için yarı benzerlik tahmin eşitliğini aşağıdaki şekilde gösterilmesi mümkün olmaktadır.

$$\sum_{i=1}^N \frac{y_i - \lambda_i}{V(\lambda_i)} x_{ij} \frac{\partial \lambda_i}{\partial \eta_j} = 0 \quad (j=1, \dots, p) \quad (98)$$

Poisson regresyon modeli için yarı benzerlik fonksiyonun elde edilmesinde yayılım parametresi $\phi=1$ olarak kabul edilmektedir. y_i ' nin dağılımı doğal üssel ailesinin bir dağılımına sahip olduğunda tahmin eşitliği en çok benzerlik tahmincileriyle aynı olmaktadır. Bu yüzden yarı benzerlik, poisson regresyon modeli için alternatif sonuçlar vermemektedir¹²¹.

Karma poisson regresyon modeli üssel bir dağılım olan poisson ve sürekli bir dağılım olan gamma dağılımını kapsadığından, yarı benzerlik poisson-gamma regresyon modelinde benzerlik tahmincilerine alternatif bir tahminci eşitliği olarak ifade edilmektedir. Yarı benzerlik fonksiyonun elde edilmesinde yayılım parametresi aşağıdaki eşitlik yardımıyla tahmin edilmektedir¹²².

$$\phi = \frac{n}{n - 2v_1 + v_2} \frac{\sum_{i=1}^n I(y_i, \lambda(x_i))^2}{\sum I(y_i, \lambda(x_i))} \quad (99)$$

Son yıllarda Poisson-gamma dağılımıyla ilgili yapılan araştırmalara bakıldığında yarı benzerlik ve EM algoritması dışında farklı modeller üzerinde çalışıldığı görülmektedir. Minimum Hellinger aralık tahminleri en çok benzerlik tahminlerine önemli bir alternatif olmaktadır. Lu, Hui ve Lee yaptıkları çalışmalarında poisson regresyon modelinin sonlu karmaları için Minimum Hellinger¹²³ aralık tahminlerini geliştirmişlerdir.

¹²⁰ Rao, s. 301-302.

¹²¹ Brajendra C. Sutradhar ve Vandna Jowaheer, "On Familial Longitudinal Poisson Mixed Models with Gamma Random Effects", **Journal of Multivariate Analysis**, No: 87, 2003, s. 398-399.

¹²² Loader, s. 71.

¹²³ Zudi Lu, Yer Van Hui ve Andy H. Lee, "Minimum Hellinger Distance Estimation for Finite Mixtures of Poisson Regression Models and Its Applications", **Biometrics**, No: 59, 2003, s. 1016.

3.3.2 EM Algoritması

“EM” algoritması hatalı değer içeren verilerde en çok benzerlik tahmincisi için güçlü bir algoritma olarak ifade edilebilir. Bu algoritma; karma regresyon modellenmesi için yapılan karma işlemi sırasında veri kaybı olacağı göz önüne alınarak karma dağılımlar için uygun bir teknik olmaktadır. Karma dağılım fonksiyonu üzerinde gerekli işlemler açıklayıcı değişkenlerin olmadığı varsayımı altında yapılmaktadır. Karma dağılımdaki hatalı veriler, y_i noktasındaki her bir gözlem için gözlemlenemeyen karma parametre θ_i olarak ifade edilmiştir. Bu yüzden E adımında, bu parametrenin bazı fonksiyonları için şartlı beklentilerin hesaplanılmasına ihtiyaç duyulmaktadır. Bu işlemden sonra, karma dağılımın yoğunluğunun maksimizasyonunu azaltmak için tamamlanmış modelin (complete model) benzerliği maksimum yapılmaktadır. Üssel ailesindeki karma modellerde şartlı beklentiler yeterli istatistiklerle tesadüfleştirilmektedir. Modelin açıklayıcı değişkenlere sahip olduğu varsayıldığında yeterli istatistiklerin sonraki beklentileri E adımında tekrar hesaplanılmaktadır¹²⁴.

“EM” algoritması hatalı veriler için kullanılan bir algoritma olmakla beraber karma dağılımlarda en çok benzerlik tahmincilerinin elde edilmesinde kullanılan bir algoritma olmaktadır. EM algoritması, en çok benzerlik tahmincilerinin iki adım içeren iteratif bir süreçten geçmesiyle elde edilmektedir. İlk adım olan E adımında, gözlem verilerinin şartlı beklentileriyle hatalı veriler yer değiştirmektedir. İkinci adım olan M adımında ise, beklenen log benzerliği maksimum yapan parametre tahminleri elde edilebilir¹²⁵.

Poisson-gamma regresyon modelinde “EM” algoritması, $E(\theta)=1$ değeri $\gamma=\delta$ olduğu varsayımı altında gerçekleşmektedir. Karma dağılımın parametresi γ ve β regresyon katsayısını ifade ettiğinde yeni tahminler k 'ninci iterasyondan sonraki mevcut tahminlerden $(\gamma_{(k)}, \beta_{(k)})$ elde edilmektedir. Böylece “E” adım, aşağıdaki gibi hesaplanılmaktadır.

¹²⁴ Karlis, s. 307-308.

¹²⁵ Peiming Wang, Iain M. Cockburn ve Martin L. Puterman, “Analysis of Patent-a Mixed-Poisson-Regression-Model Approach”, **American Statistical Association**, No: 16, S: 1, Ocak 1998, s. 40.

$$\lambda_i = \exp(\beta_{(k)} x_i) \quad i= 1, \dots, n \quad (100)$$

eşitlik (97) kullanılarak sahte değer aşağıdaki eşitlik yardımıyla elde edilebilir.

$$w_i = E(\theta_i, y_i, x_i, \gamma_{(k)}, \beta_{(k)}) = \frac{\gamma_{(k)} + y_i}{\gamma_{(k)} + \lambda_i} \quad (101)$$

ve

$$s_i = E(\theta_i \square y_i, x_i, \gamma_{(k)}, \beta_{(k)}) = \Psi(\gamma_{(k)} + y_i) - \ln(\gamma_{(k)} + \lambda_i) \quad (102)$$

$i=1, \dots, n$ için $\Psi(x)$ digama fonksiyonunu göstermektedir.

“M” adım aşağıdaki işlemlerden oluşmaktadır.

- ❖ Sahte değer w_i kullanımıyla ve basit poisson regresyon modelinin uyum iyiliğiyle regresyon parametresi β güncellenmektedir.
- ❖ Newton iterasyonu ile

$$\gamma_{(k+1)} = \gamma_{(k)} - \frac{\Psi(\gamma_{(k)}) - \ln(\gamma_{(k)}) - \bar{s} + \bar{w} - 1}{\Psi_3(\gamma_{(k)}) - \frac{1}{\gamma_{(k)}}} \quad (103)$$

γ güncellenmektedir. $\Psi_3(x)$ trigama fonksiyonunu göstermektedir.

Eğer kriter koşulları sağlamazsa iterasyon durur ve “E” adıma tekrar geri dönülmektedir¹²⁶.

¹²⁶ Karlis, s. 307-310.

3.3.3 Gözlemlenemeyen Heterojenlik

Son birkaç yıldır dikkatleri çeken bir konu bazı araştırmacılar tarafından gizli heterojenlik (hidden heterogeneity) olarak da isimlendirilen gözlemlenemeyen heterojenlik (unobserved heterogeneity) problemidir. Örneklem birimleri arasındaki tesadüfi farklılaşmalardan dolayı gözlemlenemeyen heterojenlik ortaya çıkabilir. Örneğin, açıklayıcı değişkenlerdeki ölçüm yanlışları, açıklayıcı anahtar faktörlerin çıkarılması yada çevre üzerindeki basit tesadüfi şoklar gözlemlenemeyen heterojenliğe neden olur. Gözlemlenemeyen heterojenlik ve zaman bağımlı parametrik şekil yada formlar arasındaki farklılıkların tanımlanması zordur. Çünkü yaşamla ilgili teoriler uygun fonksiyonel formlarla ilgili zaman bağımlılığı yada gözlemlenemeyen heterojenlik kaynaklarıyla ilgili olarak tatmin edici hipotezler sunmamaktadır¹²⁷. Modelde önemli faktörlerin tümünün yer almaması gözlemlenemeyen heterojenlik kaynaklarından biri olarak sayılabilir. Bu durumun bir sonucu da uygun verilerin sınırlandırılmış olmasıdır. Bazı önemli değişkenler modelde bulunmuş olmasına rağmen onlar hakkında bilgiye sahip olamayabiliriz. Böylece, hangi değişkenin model için önemli olduğuna karar verilmesi zor olmaktadır¹²⁸.

Gözlemlenemeyen heterojenlik, hazard modelindeki parametrelerin yanlış bir şekilde farklı yollarla etkilenmesi olarak ifade edilebilir. Vermunt'a¹²⁹ göre yanlışlığı ortaya çıkaran sebepleri aşağıdaki gibi özetlemek mümkündür.

Olay zamanı analizinin en önemli avantajlarından biri açıklayıcı değişkenlerin gözlemlenen zaman içerisindeki değişmelerini içermesidir. Zamana göre farklılaşan değişkenlerin etkisi olarak ifade edilen nedensel algılamaların sebep olduğu hatalar, olay zamanı analizinin zorluk boyutlarından biridir. Zamana göre farklılaşan değişkenlerin ilk dağılımlarında yapılan seçimde önyargı etkilidir. Bu durum önemli açıklayıcı değişkenlerin model içerisinde yer almamasından kaynaklanmaktadır. Farklı

¹²⁷ Karl Ulrich Mayer ve Nancy Brandon Tuma, "Life Course Research and Event History Analysis: An Overview", **Event History Analysis in Life Course Research**, 1990, s. 3-20.

¹²⁸ Hans-Peter Blossfeld ve Götz Rohwer, **Techniques of Event History Modelling: New Approaches to Causal Analysis**, Mahwah: NJ Lawrence Erlbaum Associates, Inc, 2002, s. 255-256.

¹²⁹ Vermunt, s.125-126.

olarak seçim önyargısı zamana göre farklılaşan değişkenlerde tesadüfi örneklem söz konusu ise $T=0$ olmaktadır. Eğer açıklayıcı değişkenlerin $T=0$ ile bir korelasyonu yoksa, olası gözlemlenemeyen faktörlerle ilişkisi vardır. Gözlem süresi içerisinde bireyler özel olarak açıklayıcı değişken değeri üzerinde durabilirler. Daha genel olarak ifade etmek gerekirse; zamana göre farklılaşan değişkenler ve ilgilenilen olayın hazard oranı gözlenemeyen bu faktörlerden etkilenmektedir. Sonuç olarak; örneklemin bir kısmı, hem belirli açıklayıcı değişken durumunu oluşturmada yüksek bir olasılığa hem de daha yüksek bir hazard oranına sahiptir. Zamana göre farklılaşan değişkenlerin hazard oranı üzerindeki etkisi kısmidir. Diğer olası olumsuz değerlendirme ise zamana göre farklılaşan değişkenlerin ters nedenselliğidir. Açıklayıcı değişkenlerin ve bağımlı süreçlerin nedensellik üzerinde etkisinin olduğu beklenir. Bu “durum bağıllığı” olarak isimlendirilmektedir. Farklı olarak, bu sadece gözlem zamanı içinde zamana göre değişmeyen değişkenler için geçerlidir. Zamana göre farklılaşan değişkenler ters nedenselliğe bağlıdır. Bu durum, araştırmalardaki açıklayıcı değişken sürecini etkileyebilir. Açıklayıcı değişken süreci hem zaman içerisinde farklı noktadaki durumlardan hem de hazard oranının büyüklüğünden etkilenir. Bu durum “oran bağımlılığı” olarak adlandırılmaktadır.

Parametrelerdeki yanlılığın bir başka sebebi ise bazı açıklayıcı değişkenlerin hangi kategoride inceleneceğine tam olarak karar verilememesinden kaynaklanmaktadır. Bazı açıklayıcı değişkenler vardır ki, bireysel olaylar üzerinde dışsal faktörler olarak tanımlanmayabilir ve içsel olması da gerekli değildir. Örneğin işsizlik çalışmasında bir bireyin medeni durumu hem içsel hem de dışsal bir durum olabilir. Medeni durum, birey işsiz kaldıktan sonrada tanımlanabileceğinden içsel bir değişken değildir. Medeni durumun içsel ve dışsal olması ($T=t_{i-1}$) bağlıdır. Bireyin durum statüsü (t_i) zamanı içerisinde durumunun tahmin edilmesine yardımcı olmaktadır. İçsellik, çalışma ve medeni durum arasındaki ters nedensellik sonucunda da olabilir. Çalışmaya başlamanın az da olsa boşanmaya neden olacağı görüşü bu duruma örnek olarak gösterilebilir. Bir başka sebepte gözlenemeyen açıklayıcı değişkenlerdeki ortak bir takım sonuçların, işsiz kalmanın hazard oranını ve bireysel yaşamın düzenliliğini etkilemesinden dolayı boşanma hazardını etkilemesidir.

Gözlemlenemeyen heterojenliğin zaman bağıllığı üzerinde aşağı doğru bir etkisi söz konusudur. $T=0$ zamanında modelin açıklayıcı değişken değerlerinin ilişkisi gözlemlenemeyen değerler arasında olmamasına rağmen, gözlemlenemeyen heterojenliğin zaman bağıllığı konusunda aşağı doğru bir etkisi söz konusudur. Daha fazla açıklayıcı değişken etkileri yanlı bir şekilde ortaya çıkabilir. Çünkü gözlemlenemeyen değişkenler ve gözlemlenen değişkenler $T=0$ ilişkisine sahip olabilirler. $T=0$ zamanında diğer açıklayıcı değişkenlerle gözlemlenemeyen değişken ilişki içinde olduğu zaman, bazı yanlılık ve şüpheli (spurious) ilişki etkilerinin T ve x arasında varolduğu görülebilir. Sonuç olarak zamana göre farklılaşan açıklayıcı değişkenlerin şu anki gözlemlenemeyen risk faktörlerinin bir sonucunda gerçek olmayan bir şekilde varolduğu söylenebilir. Bu durum açıklayıcı değişken sürecini ve bağıllık sürecini etkilemektedir.

Sürekli zaman hazard regresyon ve süreksiz zaman hazard regresyon modellerinde hazard oranındaki açıklanan değişim için ilişkili olan açıklayıcı değişkenlerin tümünün gözlemlendiği ve ölçüldüğü varsayılmıştır. Sonuç olarak, ihmal edilen değişkenler olmadığı için gözlemlenemeyen heterojenlik yoktur. Sosyal olaylar ise çoklu nedensel etkilere maruz kalmaktadır. Bu durum ancak, gözlemlenemeyen heterojenliğe sahip olmayan sosyal değişim verileri için nadir olabilir¹³⁰.

Gözlemlenemeyen heterojenlikteki problemin, hazard modelinde tesadüfi etki ve gizli değişken (latent variable) içerdiği inceleme konusu olmuştur. Normal şartlarda gizli değişkenin parametrik ve parametrik olmayan dağılım fonksiyonunun $T=0$ da gözlemlenebilen açıklayıcı değişken değerleriyle korelasyonsuz olduğu düşünülmektedir. Zayıf bir ihtimalle, bu tür gizli değişkenler şüpheli durumu gerçek zaman bağıllığından kurtarmaktadır. Gerçek olmayan zaman bağıllığı sadece girişteki gözlemlenemeyen bir faktördeki yanlılıkta görülebilir. Sözü edilen yaklaşımla sınırlı olarak karşılaşılmaktadır. Çünkü, gizli değişkenin gözlemlenen açıklayıcı değişkenle arasındaki ilişkiyi inceleme olanağı sağlamamaktadır. Sonuç olarak, bu durum, yanlılığı elimine etmekte kullanılamaz¹³¹.

¹³⁰ Land, Nagin ve McCall, s. 343.

¹³¹ Vermunt, s. 125-126.

Galler ve Poetter¹³² çalışmalarında gözlemlenemeyen heterojenliği işsizlik süreciyle ilişkilendirmişlerdir. Özellikle anakütlede gözlemlenemeyen heterojenlik kavramının önemli bir konu olduğu görülmektedir. Hata kavramının basitleştirilmiş modeli, gerçek ilişkilerden modeldeki sapmaları ifade etmektedir. Açıklamalarda değişkenlerin çıkarılması çok özel hatalardan kaynaklanabilir. Bu çok özel hatalar önyargı yada taraflı tahminci sonuçlarının etkisi altında kalmaktadır ve gerçek hazard oranının yanlış yorumlanmasına yol açmaktadır. İşsizlik sürecinde zamanla ilgili olarak en fazla ihmal edilen kavram heterojenlik kavramıdır. $R(i,t)$ olarak oluşturulmuş işsizlik fonksiyonunda açıklayıcı değişkenlerin çıkarılması, oranlardan uzaklaşmak yada zaman bağıllığının yetersiz modellenmesi zaman bağıllığı formunun tahmincilerinin ve açıklayıcı değişkenlerin etkisinde yanlı bir durum yaratmaktadır. Bireylerin istihdam piyasasında kalmasında bir takım faktörler vardır. Bunlardan birisi yabancıların istihdam piyasasından ayrılışı o ülke vatandaşlarına göre daha yavaş olmaktadır. Bir başka unsur ise gençler orta yaşlılara göre daha hızlı iş değiştirmektedirler. Yaşlı insanlar bu konuda çok daha yavaşlardır özellikle erkeklerde profesyonel bir eğitimi olmayanların oranlarına göre daha yavaş piyasadan ayrıldıkları görülmektedir. Bütün bu faktörler istihdam piyasasında bireylerin işte kalış ve ayrılış süreçleri üzerinde önemli bir etki yaratmaktadır. Tüm bu faktörler gözlemlenemeyen heterojenliği oluşturmaktadır.

Lomi¹³³ 1964-1988 yılları arasında İtalya'nın gelişmemiş bölgelerinde bankalarla ilgili olarak lokasyon bağıllığı ve gözlemlenemeyen heterojenliği tespit etmeye çalışmıştır. Lokasyonun yoğunluk ve örgüt kuruluşları arasındaki rolü, dikkatlerin doğrudan örgüt popülasyonundaki heterojenlik konusuna çekmektedir. Yapısal parametreleri teorik olarak etkileyen lokasyon faktörleri gözlemlenemeyen heterojenlik unsuru olarak değerlendirilmektedir. Çalışmanın modellenmesi aşamasında, poisson regresyon ve poisson-gamma regresyon analizleri kullanılarak, örgütlerin kurulmasında içsel farklılıkların etkili olduğu yapılan çalışmada ortaya konulmuştur.

¹³² Heinz P. Galler ve Ulrich Poetter, "Unobserved Heterogeneity in Models of Unemployment Duration", **Event History Analysis in Life Course Research**, 1990, s. 241-252.

¹³³ Alessandro Lomi, "The Population Ecology of Organizational Founding: Location Dependence and Unobserved Heterogeneity", **Administrative Science Quarterly**, No: 40, 1995, s. 113-114.

Baker, Faulkner ve Fisher¹³⁴ çalışmalarında organizasyonlar arası pazar ilişkilerindeki gözlemlenemeyen heterojenliği incelemiştir. Organizasyonlar arası pazarın bitiş sebebini, reklam ajansları ve onların müşterileri arasındaki bağlantıları örneğinde üç önemli faktörle açıklamışlardır. Bu faktörler; pazardaki rekabet durumu, pazarda güçlü olma ve kurumsal zorlayıcılar olarak belirlenmiştir. Pazar kurumsal bir yapı olarak mükemmel olmayan bir şekilde yapısını değiştirmektedir. Çünkü, rekabet ve değişen normlar pazar bağlantıları sürecinde pazar ilişkilerindeki dengeleri bozmaktadır. Toplanan olay zamanı verileri, olay zamanı analizlerinden biri olan poisson regresyon analizi ile değerlendirilmiştir. Poisson regresyon analizi sonucunda, aşırı yayılımla karşılaşılmasından dolayı poisson-gamma regresyon kullanılarak pazar riski kontrol edilmeye çalışılmıştır.

Blosfeld ve Hamerle¹³⁵ sürekli zaman-hazard oran modelindeki gözlemlenemeyen heterojenliği incelemiştir. Bireylerin işten ayrılmalarında etkili olan gözlemlenemeyen heterojen unsurlar üzerinde çalışmalarını gerçekleştirmişlerdir. Çalışmalarında; insan sermayesine yapılan yatırım ve sosyal akışkanlık gibi konular temel heterojenlik faktörleri olarak belirlenmiştir.

Gözlemlenemeyen heterojenlikle ilgili olarak bir başka yaklaşım ise; tahminci geliştirilerek sorun çözülmeye çalışılmasıdır. Bilindiği gibi açıklayıcı değişkenlerin doğrusal ve karma modellerdeki regresyon katsayılarıyla ilgilenildiği durumlarda regresyon analizi yapılır. Zayıf bir yaklaşım olarak her konunun tekrar ölçüldüğü durumlar söz konusu olabilir. Bunlar doğrusal modeller tarafından takip edilir. Ve sonuçta konu ile ilgili özel tahminci katsayısı açıklayıcı değişkenler yerine ikame edilir. Eğer temel regresyon denklemi katsayılardan oluşturulmuş ise ki bunlar önyargılı tahmincilerden meydana gelebilir, bu durumda gözlemlenemeyen açıklayıcı değişkenlerle ikame edilebilir. Bu önyargıyı düzeltmek için beklenen sahte tahminci

¹³⁴ Wayne E. Baker, Robert R. Faulkner ve Gene A. Fisher, “Hazards of The Market: The Continuity and Dissolution of Interorganizational Market Relationships”, **American Sociological Association Review**, No: 63, S: 2, Nisan 1998, s. 154-165.

¹³⁵ Hans-Peter Blosfeld ve Alfred Hamerle, “Unobserved Heterogeneity in Hazard Rate Models: A Test and an Illustration from A Study of Career Mobility”, **Event History Analysis in Life Course Research**, 1990, s. 241-252.

eşitliği geliştirilmiştir¹³⁶. Açıklayıcı değişkenlerin seçimindeki eksikliklerden dolayı oluşan hataları önlemek için Kulich ve Lin¹³⁷ toplamsal hazard regresyon modeli için tahminci geliştirmişlerdir.

Yamaguchi çalışmasında¹³⁸ gözlemlenemeyen heterojenlikle ilgili olarak genel değerlendirmelere yer vermiştir. Kısaca bu değerlendirmeler aşağıdaki şekilde özetlenebilir. Gözlemlenemeyen heterojenliğin kontrol edilmesinde iki farklı grup yöntem kullanılmaktadır. Bunlardan ilki, sabit etki yöntemi (fix effect) ve ikincisi ise tesadüfi etki (random effect) yöntemidir.

Sabit etki yönteminde, sabit bireylerde özel etkiler olduğu düşünülmektedir. Sabit etki yöntemi, yanlış seçimi ortadan kaldırmaktadır. Fakat sadece tekrar edilebilir olaylarda uygulanabilir olması bu yöntemin dezavantajlarından biridir. Ayrıca yöntemin uygulanması için çok güçlü gereksinimlere ihtiyaç duyulması bu yöntemin önemli seviyede kısıtlara sahip olduğunu ifade etmektedir. Sabit etki yönteminin birinci türü, yenilenen süreçlerden (renewal process) durum verilerine (duration data) kadar uygulanabilir olmasıdır. Bu yöntem, iki tane durdurulmuş zaman olduğunda kullanılabilir. Bu modelde, incelenecek olan her konu için çoklu zamanlardaki olayların bağımsız sonuçlar olduğu düşünülür. Aynı durum oransal hazard modelinde de görülmektedir. Oransal hazard modelinde, cox kısmi benzerlik yöntemi uygulanmaktadır. Aynı şekilde açıklayıcı değişkenler için (covariate), parametreler olaylar arasındaki bir durumdur. Bu durum, olaylar arasında özel olmayan temel hazard fonksiyonunun farklılaşması olarak ifade edilebilir. Bu aşamayı tabakalı modeller için kısmi benzerlik tahmincileri takip eder. Bunun için tamamen durdurulmuş olaylardaki bağımsız gözlem verilerinden yararlanılır ve her olaya farklı bir veri olarak yaklaşılır. Bu süreç sadece bireyler arasında yada durum verilerinin karşılaştırılması halinde tahminci parametresi için kullanılır. Eğer veri seti her olayla ilgili durdurulmuş çift zamana sahipse ve açıklayıcı değişkenler zaman sürecinde farklılık gösteriyorsa bu

¹³⁶ C.Y. Wang, Naisyin Wang ve Suojin Wang, “Regression Analysis When Covariates are Regression Parameters of a Random Effects Model for Observed Longitudinal Measurements”, **Biometrics**, No: 56, 2000, s. 487-489.

¹³⁷ Michal Kulich ve D.Y. Lin. “Additive Hazards Regression with Covariate Measurement Error”, **Journal of American Statistical Association**, No: 95, S: 449, Mart 2000, s. 238-248.

¹³⁸ Kazuo Yamaguchi, **Event History Analysis**, California: Sage Publication, Inc, 1991, s. 131.

durumda kısa zaman için “bir” atayarak ve diğerlerine de “sıfır” atayarak iki uçlu bağımlı değişkenlerin değerlerine lojistik regresyon modeli uygulanabilir.

Sabit etki yöntemi, genelleştirilmiş tahmin denklemi (generalized estimating equations) yaklaşımını kapsamaktadır. Genelleştirilmiş tahmin denklemi yaklaşımı, çalışma korelasyon yapısının (working correlation structure) seçilmesine gereksinim duymaktadır. Çalışma korelasyon yapısının seçimine yol gösterici olarak gözlemlerin korelasyon yapısını kullanmak, iki düzeyli sonuç değişkeni içeren modellerde problem yaratmaktadır. Bu soruna çözüm olarak, değiştirilebilir bir korelasyon yapısı kullanmak önerilmektedir¹³⁹.

Horowitz¹⁴⁰ tarafından yapılan çalışmada oransal hazard, sabit etki modeliyle analiz edilmektedir. Çalışma daha önceden bilinmeyen temel model ve birleştirilmiş temel hazard fonksiyonunun incelemektedir. Varolan tahminciler temel hazard fonksiyonunun bilinen parametrik değerleri üzerine kuruludur. Sonuçta Horowitz’in ortaya koyduğu tahminci, varolanlardan daha genel durumu ifade etmektedir. Aynı zamanda sağdan durdurulmuş veriler için parametrik oransal hazard tahmincisini de ortaya koymuştur.

Abrevaya¹⁴¹ tarafından yapılan çalışmada ise sabit etki modeline göre genelleştirilmiş regresyon modelinin tahmincisi üzerine odaklanılmıştır. Bu model, hata dağılımı üzerindeki parametrik olmayan durumlarda durdurmaya izin vermektedir. Yapılan çalışma, sabit etkinin açıklayıcı değişkenlerle korelasyonunu ortaya koymaktadır. Çalışmada öncelikle genellenen sabit etki regresyon modelindeki tahminciler açıklanmıştır. İkili seçimdeki sabit etki modelinde maksimum skor tahmincisi ele alınmıştır. Modelin bu türü kesilmiş verinin kullanımına da olanak sağlamaktadır. Bu durumda tahminci açısından hata dağılımında uyumlu bir şekilde hazard fonksiyonunun artmasına yol açmaktadır.

¹³⁹ Jos W. R. Twisky, **Applied Longitudinal Data Analysis for Epidemiology: A Practical Guide**, West Nyack, NY, USA: Cambridge University Press, 2003, s. 157-159.

¹⁴⁰ Joell Horowitz ve Sobae Lee. “*Semiparametric Estimation of a Panel Data Proportional Hazard model With Fixed*”, **Biometrics**, January 2002, s. 1-2.

¹⁴¹ Jason Abrevaya, “*Rank Estimation of a Generalized Fixed- Effects Regression Model*”, **Journal of Econometrics**, No: 95, 2000, s. 1-23.

Zamana göre sabit deęişkenler (time constant covariate) sabit etki yöntemiyle kullanılmazlar. Çünkü sabit deęişkenlerin etkileri anakütlenin heterojenliğindeki gibi kontrol edilememektedir.

Tesadüfi etki yöntemi, tesadüfi hata kavramını modelde açıklamaktadır. Sabit etki yönteminden farklı olarak tesadüfi etki yöntemi tekrarlanan olaylar kadar tekrarlanmayan olaylarda da kullanılabilir. Sabit etki yöntemine göre, kullanımında çok az sınırlandırmalar vardır. Tesadüfi etki yönteminde sabit etki yönteminden farklı olarak yanlış seçim elimine edilememektedir. Bunun sebebi de, açıklayıcı deęişkenler için tesadüfi hatanın bağımsızlığının bu modelde yer almamasından kaynaklandığı söylenebilir. Bu yöntemin avantajlarından biri, olası yanlışlıkları da düzeltici bir niteliğe sahip olmasıdır¹⁴².

Tesadüfi etki yöntemleri tesadüfi hata kavramları için karma dağılım kullanmaktadır. Karma dağılım anakütledeki tesadüfi hatanın dağılımı için fonksiyonel yapıyı özel hale getirmektedir. Heckman¹⁴³ parametre tahmincilerinin karma (mix) dağılımlara bağlı olduğunu görmüştür. Bunun sonucunda da parametrik olmayan karma tahminci dağılımlarını geliştirmiştir. Gözlemlenemeyen heterojenlik konusu olay zamanı analizinin kullanımını ima etmektedir. Bilgisayar paket programlarının yetersizliğinden kaynaklı olarak gözlemlenemeyen heterojenlik kontrol edilememektedir. Bir başka nedense sofistike matematiğin arkasındaki güvene sahip olunamamasıdır.

Gözlemlenemeyen heterojenlikteki yanlışlık, modelin yanlış belirlenmesindeki olasılıklardan bir tanesidir. Bunun dışında bir çok nedenin, modeli yanlış kurmada etkisinin olduğu görülmektedir. Bir örnekle açıklamak gerekirse; açıklayıcı deęişkenler arasında önemli etkileşim söz konusu olduğunda ve bu durum dikkate alınmadığında model yanlış yönlendirilebilir.

¹⁴² Charles N. Halaby, “Panel Models in Sociological Research: Theory into Practice”, *Annu. Rev. Socio.*, No: 30, 2004, s. 517-531.

¹⁴³ Charles F. Manski ve Daniel L. McFadden, *Structural Analysis of Discrete Data with Econometric Applications*, Cambridge: The MIT Press, 1981, s. 114-119.

Powers ve Xie' e göre¹⁴⁴ sabit etki yönteminin en büyük avantajı; gözlemlenemeyen heterojen dağılım hakkında varsayımlar yapılabilmesinden dolayı tedbirli bir yöntem oluşudur. Bununla beraber, bir çok verinin kaybedilmesi bu avantajı azaltan en büyük etmendir. Bu yüzden, çalışmadaki bütün verileri kullanan yaklaşımlara göre daha az etkili bir yöntemdir. Panel verilerin analizinde çok kullanılan bir model olmasına rağmen; gözlemlenemeyen heterojen dağılımın varsayımlarının sağlanmadığı durumlarda, gözlem süresi içerisinde değişime uğramayan açıklayıcı değişkenleri dışlayan bir modeldir. Böylece sabit açıklayıcı değişken etkileri tahmin edilememektedir. Halbuki tesadüfi etkiler modelinde sabit ve zamana göre farklılaşan değişkenlerin her ikisi de kullanılmaktadır. Tesadüfi etki yönteminin sabit etki yöntemine göre tercih edilmesinin en büyük sebebi her iki değişkeni de kapsamından kaynaklanmaktadır.

3.4 Tesadüfi Etkiler Poisson-Gamma Regresyon Modeli

Tesadüfi etki modelleri, tekrarlı ölçümler içeren çalışmalarda ve gözlemlerin grup içinde bağımsız olarak gruplandırıldığı kümelenmiş tesadüfi denemelerin analizinde kullanılmaktadır¹⁴⁵.

Tesadüfi etki modelleri sosyal bilimlerde yapılan çalışmalarda oldukça öneme sahiptir. Birimlerin farklı kümelerde gruplaşmış olmasından kaynaklanan mevcut bağımlılığa imkan tanınması, bu modelin yaygın olarak kullanılmasına olanak vermiştir. Oldukça esnek bir model olmakla beraber zamana göre değişmeyen tesadüfi

¹⁴⁴Daniel A. Powers ve Xie Yu, **Statistical Methods for Categorical Data Analysis**, San Diego: Academic Press, 2000, s. 180-181.

¹⁴⁵ M. J. Campbell, **Statistics at Square Two: Understanding Modern Statistical Applications in Medicine**, London: BMJ Publishing Group, 2001, s. 74.

etkiler, modeli sınırlandırmaktadır. Biyoistatistik alanında yapılan çalışmalarda aynı model için, bozulma modelleri (frailty models) terimi kullanılmaktadır¹⁴⁶.

Poisson modelinin en basit tesadüfi etkiler genişletilmiş, oran parametresi için bir gamma dağılımı varsayımıdır. Bu model poisson-gamma regresyon modeli olarak da isimlendirilmektedir.

Herhangi bir birim için belirtilmiş olan tesadüfi etki “ v_i ” ile gösterildiğinde model aşağıdaki biçimde yazılabilir.

$$\lambda_i(t) = v_i \rho(t, \phi) e^{x_i \beta}. \quad (104)$$

Modelde yer alan v_i belirli bir dağılıma sahiptir. Poisson regresyon modelinde aşırı yayılım söz konusu olduğunda v_i ’nin dağılımı karma bir dağılım olmaktadır. ϕ , $\rho(t, \phi)$ yoğunluk fonksiyonunun şekil parametresidir. Sabit yoğunluk fonksiyonu varsayımı geçerli olduğu durumlarda $\rho(t, \phi)$, 1’e eşit olarak tanımlanabilir ve ϕ modelde yer almaz. Karma modelin ortalaması “1” ve varyansı “ τ ” olarak gösterilebilir. Benzerlik analizleri için tesadüfi etkilerin dağılımı $p(v, \tau)$ olarak yazılabilir¹⁴⁷.

3.4.1 En Çok Benzerlik Tahmincisi

Modelin parametre tahmincilerinin bulunmasında tüm veriler için benzerlik fonksiyonun aşağıdaki gibi yazılması mümkündür.

$$L_{\text{tüm}}(\beta, \alpha, \tau) = L_{\alpha}(\alpha) L(\beta, \alpha, \tau). \quad (105)$$

¹⁴⁶ R. Crouchley ve R. B. Davies, “A Comparison of Population Average and Random-Effect Models for The Analysis of Longitudinal Count Data with Base-Line Information”, **J.R. Statistical Society A**, No: 162, 1999, s. 332.

¹⁴⁷ Dean ve Balshaw, s. 1387-1388.

$L_\alpha(\alpha)$ için aşağıdaki eşitlik (103) yazıldığında

$$L_\alpha(\alpha) = \prod_{i=1}^M \prod_{s=1}^{n_i} \frac{\rho(\omega_{is}; \phi)}{R(T_i; \phi)} \quad (106)$$

biçiminde gösterilebilir. $n_i = Y_i(t)$ olarak tanımlandığında $\{\omega_{is}, s = 1 \dots n_i\}$ i'inci birim için olay zamanları olarak ifade edilebilir.

Karma poisson regresyon modeli için benzerlik fonksiyonu aşağıdaki gibi yazılabilir.

$$L(\beta, \alpha, \tau) = \prod_{i=1}^M \int_0^\infty (v_i \lambda_i)^{n_i} e^{-v_i \lambda_i} (n_i!)^{-1} p(v_i) \partial v_i \quad (107)$$

Eğer v_i gamma dağılımına sahip ise, $L(\beta, \alpha, \tau)$ için poisson-gamma benzerlik

$$L_{P-G}(\beta, \phi, \tau) = \prod_{i=1}^M \frac{\Gamma(n_i + \tau^{-1})}{(n_i!) \Gamma(\tau^{-1})} \left(\frac{\tau \lambda_i}{1 + \tau \lambda_i} \right)^{n_i} \left(\frac{1}{1 + \tau \lambda_i} \right)^{\tau^{-1}} \quad (108)$$

olarak ifade edilebilir¹⁴⁸.

3.4.2 Yarı Benzerlik Tahmincisi

Yarı benzerlik tahmincilerinin yaygın olarak kullanılmasında önemli sebepler bulunmaktadır. β 'nin asimptotik varyansı, τ için tahmin fonksiyonunun seçiminden bağımsızdır ve sadece sayılabilir dağılımın ilk iki momentine bağlıdır. Bu asimptotik varyans, τ tahmincisinin tutarlılığı kadar tutarlıdır. τ için genel tahmin eşitliği, tahmincinin tutarlılığını sağlayan ilk ve ikinci moment varsayımı için gerekli görülebilir. Bu durum yarı benzerlik tahmincilerine robust bir nitelik sağlayabilir. Başka

¹⁴⁸ Dean ve Balshaw, s. 1391.

bir şekilde ifade etmek gerekirse, τ 'nın en çok benzerlik tahmincisinin tutarlılığı dağılım varsayımının doğrulanmasını gerektirebilir.

β regresyon katsayısı için yarı benzerlik tahmin eşitliğinin aşağıdaki şekilde gösterilmesi mümkün olmaktadır.

$$\sum_{i=1}^N \frac{(n_i - \mu_i)x_{ir}}{(1 + \tau\mu_i)} = 0 \quad r=1, \dots, k \quad (109)$$

Yarı benzerlik fonksiyonun belirlenmesinde, yayılım parametresi aşağıdaki eşitlik yardımıyla tahmin edilebilir.

$$\phi = \sum_{i=1}^N \frac{(n_i - \mu_i)}{(1 + \tau\mu_i)} \left(\frac{\partial \log R_i}{\partial \phi} \right) = 0. \quad (110)$$

Eşitlikteki R_i ,

$$R_i(T_i; \phi) = \int_0^{T_i} \rho(\varepsilon; \phi) du = R_i \quad (111)$$

olarak gösterilebilir.

3.5 Modelde Yer alan Parametre Tahmincilerinin ve Modelin Uyum İyiliğinin Sınanması

Regresyon çözümlemesi sonucunda ulaşılan modelin doğru olduğu biliniyorsa, daha ileri bir analize gerek kalmadan çalışma sonlandırılabilir. Ancak, yinede elde edilen sonuçlar üzerinde dikkatli bir denetim yapmadan sonuç modeli kullanmamak gerekmektedir. Bu denetim süreci, genellikle modelin yeterliliğinin saptanması süreci olarak bilinmektedir¹⁴⁹. Modelin yeterliliğinin sınanması, parametre tahmincilerinin

¹⁴⁹ Alpar, s. 259.

sınanması ve modelin uyum iyiliğinin test edilmesi sürecinden oluşmaktadır. Parametre tahmincilerinin sınanması ve modelin uyum iyiliğinin sınanmasında farklı testler bulunmaktadır. Bu testler sonucunda verilerin modele uyum gösterdiği ve parametre tahmincilerinin modelde uygun tahminciler olduğuna karar verilebilir.

3.5.1 Parametre Tahmincilerinin Sınanması

Poisson regresyon ve karma poisson regresyon modelinde yer alan parametre tahmincilerinin sınanmasında, lineer modellerde olduğu gibi hipotez testi sınamaları yapılmaktadır. Kullanılan sınamalar; “t-testi”, “wald testi” ve “benzerlik oran testi” (the likelihood-ratio test) şeklinde ifade edilebilir. Parametre tahmincilerinin sınanmasında kullanılan hipotezlerin

$$H_0 : \beta_m = 0$$

$$H_1 : \beta_m \neq 0$$

olarak gösterilmesi mümkündür.

3.5.1.1 t-Testi Sınaması

Bireysel parametrelerin hipotezlerinin test edilmesi için t testi veya güven aralığı yaklaşımı hesaplamada kolaylık sağladığı için en kolay olarak nitelendirilen yaklaşımlardan biridir. Maksimum benzerlik tahmincileri, kesin genel şartlar uygulandığında istenen bir çok özelliğe sahip olmaktadır. Model için belirlenen şartlar; bağımsız ve tanımlanmış dağılımların gözlemlerinden oluşmuş bir veri seti olması ve x_i 'nin bağımsız olması ayrıca model hatası e_i içermesi şeklinde belirlenebilir. Bu durum tatmin edici olduğunda maksimum benzerlik tahmincisi asimptotik olarak yansızdır, normal dağılım göstermektedir ve asimptotik normal tahminciler arasında en küçük

varyanslıdır. Ters bilgi matrisi, maksimum benzerlik tahmincisinin ulaşılabilir aralıkta olmasını sağlamaktadır. Bilgi matrisi bir çok sayıdaki istatistiksel testi ortaya koymaktadır. Sıfır hipotezi açısından t oranları, tahmincinin standart hatadan elde edilen ters bilgi matrisinin k'nıncı diyagonal elementinin kareködür. Bu değer daha sonra standart normal tablolar ile karşılaştırılır. Bir çok bilgisayar programı çıktısı t oranını, sıfır hipotezinin doğru olduğu durumlardaki varsayımlarından elde etmektedir¹⁵⁰.

3.5.1.2 Wald Testi

Wald testi bir yada daha çok modelin test edilmesinde kullanılan bir testtir. Bu yüzden diğer testlere göre daha genel bir test olmaktadır. Benzerlik oran testinden farklı olarak tahmini birkaç nest model gerektirmektedir. Yada modelde kısıt olarak belirlenen iki parametrenin karşılaştırılmasında kullanılmaktadır.

$\beta_m = 0$ hipotezi için, wald χ^2 istatistiği aşağıdaki gibi olmaktadır.

$$W_i = \hat{\beta}'_k V_m^{-1} \hat{\beta}_m. \quad (112)$$

V , $\hat{\beta}$ 'nin asimptotik varyans kovaryans matrisidir. V_m , V 'nin alt matrisini ifade etmektedir. Wald istatistiği, serbestlik derecesinin sınırlandırılmış parametrelere eşit olduğu durumlarda χ^2 olarak dağılmaktadır. Tek parametre ile wald istatistiği basit olarak t oranının karesi olmaktadır. Tek parametreden daha fazla parametre olduğunda, wald istatistiği ilgili matrislerin ve vektör sayılarının çıkarılması ve yönlendirilmesiyle değerlendirilir. Wald istatistiği diğer kısıtların test edilmesine izin vermek için genelleştirilebilir. Bu durumda sıfır hipotezi aşağıdaki gibi olmaktadır.

$$H_0 = R \beta_m = q.$$

¹⁵⁰ Powers ve Xie, s. 71-72.

Sıfır hipotezinde yer alan R her satırın tek kısıt olarak katsayı vektörüyle kısıt matrisidir. q sıfırların alt matrisidir¹⁵¹.

Wald testinin benzerlik oran testinden farkı, kısıtlanmamış parametreler ve kısıtlanmış parametreler arasındaki yatay (horizontal) farklılıklar üzerine temellendirilmiş olmasıdır. Bu ters bilgi, parametre tahminleri için varyans olarak isimlendirilebilir. Bilgi kısıtlanmamış parametre tahmin değerleriyle hesaplanmaktadır. Test, kısıtlanmış parametre sayısına eşit serbestlik derecesiyle χ^2 dağılımına sahip olmaktadır¹⁵².

3.5.1.3 Benzerlik Oran Testi

Modelin parametre tahmincisi olan log-benzerlik tek başına uygunluk indeksi olarak kullanılamamaktadır. Çünkü, örneklem büyüklüğünden bağımsız değildir. Log benzerlik değerinin küçük olması ise örneklem büyüklüğüne göre değişim göstermektedir¹⁵³.

$$\lambda = \frac{\max L_c}{\max L_f} \quad (113)$$

olarak ifade edildiğinde L_c ; H_1 hipotezi altında maksimum log-benzerlik ve L_0 ; H_0 hipotezi altında maksimum log-benzerlik benzerliği ifade etmektedir. L_c göreceli olarak L_f 'dan küçükse modelin veri ile yeterli bir şekilde uyum göstermediği söylenebilir. Başka bir ifadeyle zayıf model uygunluğu söz konusudur. L_c değeri L_f değeri ile aynı olduğunda modelin uygun bir model olduğu söylenebilir. Genel olarak modelin uyum iyiliğinin göstergesini benzerlik oran istatistiği olan G^2 ortaya çıkarmaktadır. Benzerlik oran istatistiği,

¹⁵¹ Powers ve Xie, s. 72-73.

¹⁵² A. Buse, "The Likelihood Ratio, Wald, and Lagrange Multiplier Tests: An Expository Note", *American Statistician*, No: 36, S: 3, Ağustos, 1982, s. 153-157.

¹⁵³ Powers ve Xie, s. 66-67.

$$G^2 = -2\log(L_c / L_f)$$

$$G^2 = -2(\log L_c - \text{Log}L_f)$$

$$G^2 = 2\sum_{i=1}^n y_i \log\left(\frac{y_i}{\hat{\lambda}_i}\right) - 2\sum_{i=1}^n (y_i - \hat{\lambda}_i) \quad (114)$$

k serbestlik derecesi ile χ^2 dağılımı göstermektedir. Hiyerarşik model çiftlerinin log benzerliği (-2) çarpımıyla sıfır hipotezi altında yaklaşık olarak bir χ^2 dağılımı olmaktadır¹⁵⁴.

Benzerlik oran testinin serbestlik derecesi, logit sayısı eksi modeldeki bağımsız parametre sayısına eşit olmaktadır. Logit sayısı ve bağımsız parametre sayısı birbirine eşit olduğunda serbestlik derecesi sıfır olarak hesaplanır. Bu modeller tam uygunluğa sahiptir. Bağımsız parametre ve gözlem sayılarının eşit olduğu bu tür modeller, doyurulmuş modeller (saturated model) olarak isimlendirilmektedir¹⁵⁵.

3.5.2 Modelin Uyum İyiliğinin Sınanması

Lineer regresyon modellerinde örneklem doğrusunun verilere ne kadar iyi olduğunun ölçütü, bir veri kümesine uyarlanmış regresyon doğrusunun uyum iyiliği olarak isimlendirilebilir¹⁵⁶. Parametreler tahmin edildikten sonra gözlemlerin modelin şekli etrafındaki dağılımlarının ölçülmesi gerekmektedir. Çünkü gözlemler tahmin edilen modelin çizilen şekline ne kadar yakınsa modelin uyum iyiliği de o kadar yüksek

¹⁵⁴ Michael T. Hannan ve John Freeman, **Organizational Ecology**, Cambridge : Harward University Press, 1989, s. 195-197.

¹⁵⁵ Helge Toutenburg, **Statistical Analysis of Designed Experiments**, New York: Springer –Verlag, 2002, s. 261-262.

Duncan Sprott, **Statistical Inference in Science**, New York: Springer Verlag, 2000, s. 90.

¹⁵⁶ Damador N Gujaratti, **Temel Ekonometri**, Çev. Ümit Şenesen ve Gülay Günlük Şenesen, İstanbul: Literatür Yayıncılık, 1999, s. 74.

olacaktır. Bir başka deyişle, Y'deki değişimin açıklayıcı değişkenlerdeki değişmelerle açıklanması o kadar iyi olacaktır¹⁵⁷.

Poisson regresyon modeli ve karma poisson regresyon modelinin uyum iyiliğinin sınanmasında kullanılan sınamalar lineer regresyona göre değişmekle birlikte aynı mantıktan hareket edilmektedir. Modelin uyum iyiliğinin sınanmasında; Pearson χ^2 sınaması, sapmalardaki farklılıklar ve model χ^2 sınaması, sahte- R^2 (Pseudo- R^2) sınaması, Bayes Bilgi Ölçütü (Bayes Information Criteria BIC) ve Akaiki Bilgi Ölçütü (Akaiki Information Criteria AIC) yaygın olarak kullanılan teknikler olmaktadır. Son yıllarda model seçiminde kullanılan yeni yaklaşımlar geliştirilmiştir. Charnigo ve Sun¹⁵⁸ karma modellerin seçiminde kullanılan “ L^2 Uzaklık Testini” önermiştir. Miller¹⁵⁹ benzerlik oran istatistiğini temel alarak “Bütün Mümkün Karşılaştırmalar” yaklaşımını geliştirmiştir. Geliştirilen testlerin uygulanması özel bir yazılım programı gerektirdiğinden bu çalışmada kullanılamamaktadır.

Modelin uyum iyiliği testlerinde;

H_1 = Tam modele (full model) karşı

H_0 = Model doğrudur

hipotezleri modelin uyum iyiliğinin sınanmasında kullanılan hipotezler olmaktadır.

¹⁵⁷ A.Koutsoyiannis, **Ekonometri Kuramı**, Çev. Ümit Şenesen ve Gülay Günlük Şenesen, Ankara: Verso Yayıncılık, 1989, s. 69-70.

¹⁵⁸ Richard Charnigo ve Jiang Sun, “Testing Homogeneity in a Mixture Distribution Via The L^2 Distance Between Competing Models”, **Journal of The American Statistical Association**, No: 99, S: 466, Haziran 2004, s. 488.

¹⁵⁹ Arden Miller, “The Analysis of Unreplicated Factorial Experiments Using All Possible Comparisons”, **Technometrics**, No: 47, S: 1, Şubat 2005, s. 51.

3.5.2.1 Pearson χ^2 Sınaması

Pearson χ^2 istatistiği, incelenen model altında beklenen frekanslar ile gözlenen frekanslar göz önünde bulundurularak hesaplanılmaktadır.

Pearson χ^2 istatistiği,

$$\chi^2 = \phi \sum_{i=1}^n \frac{\{y_i - E(y_i)\}^2}{\text{var}(y_i)} = \sum_{i=1}^n \frac{(y_i - \hat{\lambda}_i)^2}{\hat{\lambda}_i} \quad (115)$$

olarak ifade edilebilir. χ^2 küçük değeri modelin verilere uyumunu yada tam uygunluğu gözlenen ve beklenen değerler arasında vermektedir. Farklı olarak büyük değerler modelin verilere uyumsuzluğunu yada uygunluk eksikliğini ifade etmektedir. Hesaplanan χ^2 değerinin serbestlik derecesine oranının 1'den daha büyük değer alması (χ^2 /serbestlik derecesi >1) verilerin modele uygun olmadığını ve aşırı yayılımın söz konusu olduğunu ifade etmektedir¹⁶⁰.

3.5.2.2 Sapmalardaki Farklılıklar ve Model χ^2 Sınaması

Modelin uyum iyiliğinin belirlenmesinde, genellikle veriler açısından en iyi model oluşturmaya çalışılmaktadır. Modelin uyum iyiliğinin değerlendirilmesinde farklı bir yaklaşım alternatif modellerin sapmalarının karşılaştırılmasıdır. Yarışan nest (competing nested) modellerin sapmalarının karşılaştırılması uygunluğun artırılması için bir yaklaşım sunmaktadır. Sapmalardaki farklılıklar, yarışan modellerdeki farklı

¹⁶⁰ Sprott, s. 90; Allan McCutcheon , **Advances in Latent Class Models**, West Nyack, NY, USA: Cambridge University Press, 2002, s. 348-349.

sayıdaki parametrelere eşit serbestlik derecesiyle χ^2 dağılımı göstermektedir. G_1^2 , birinci model için sapmayı ve G_2^2 , ikinci model için sapmayı ifade ettiğinde farklar aşağıdaki gibi olmaktadır.

$$\Delta G^2 = G_1^2 - G_2^2 \quad (116)$$

ifadesi χ^2 olarak dağılmakta ve serbestlik derecesi her ek parametre açısından daha az kısıtlayıcı bir model olmaktadır. Sapma veya G^2 istatistiği, benzerlik oran istatistiği olarak bilinen genel bir ölçümün özel bir durumu olmaktadır. Paket programlarda sapma $\log L_0$ veya $-2 \log L_0$ modeliyle hesaplama yapılmaktadır.

Model χ^2 yaklaşımında, güncel model (current model) ile sıfır modeli (null model) uygunluk açısından karşılaştırılmaktadır.

$$\text{Model } \chi^2 = -2 \log L_0 - (-2 \log L_c) \quad (117)$$

ifadesinde “ L_0 ” sıfır modelin benzerliğini, “ L_c ” ise güncel modelin benzerliğini ifade etmektedir. Model χ^2 , güncel ve sıfır modelindeki parametre sayısındaki farklılara eşit serbestlik derecesiyle χ^2 dağılımına sahiptir¹⁶¹.

3.5.2.3 Sahte- R^2 Sınaması

Sınırlandırılmış ve nitel bağımsız değişkenli modellerde kullanılan bir ölçüm olarak sahte- R^2 önerilmektedir. Ayrıca modellerin değerlendirmelerinde de uygun bir

¹⁶¹ Powers ve Xie, s. 69-70.

araç olarak kullanılabilir. Sahte- R^2 , geleneksel R^2 gibi aynı asimptotik sınırlara sahiptir¹⁶².

Sahte- R^2 'nin hesaplanılmasında farklı hesaplama yöntemleri mevcuttur. Bu yöntemlerden ilki aşağıdaki eşitlikte gösterildiği gibidir.

$$R_G^2 = \frac{L(\phi, \lambda_i) - L(\phi, \bar{y})}{L(\phi, y_i) - L(\phi, \bar{y})} \quad (118)$$

Eşitlikte yer alan $L(\phi, \lambda_i)$, $L(\phi, \bar{y})$ ve $L(\phi, y_i)$ hesaplanılması aşağıda gösterildiği gibi olmaktadır.

$$L(\phi, \lambda_i) = \sum_{i=1}^n \left\{ y_i \log\left(\frac{\lambda_i}{1 + \phi \lambda_i}\right) + (y_i - 1) \log(1 + \phi y_i) - \frac{\lambda_i (1 + \phi y_i)}{1 + \phi \lambda_i} - \log(y_i!) \right\} \quad (119)$$

$$L(\phi, \bar{y}) = \sum_{i=1}^n \left\{ y_i \log\left(\frac{\bar{y}}{1 + \phi \bar{y}}\right) + (y_i - 1) \log(1 + \phi y_i) - \frac{\bar{y} (1 + \phi y_i)}{1 + \phi \bar{y}} - \log(y_i!) \right\} \quad (120)$$

$$L(\phi, y_i) = \sum_{i=1}^n \left\{ y_i \log\left(\frac{y_i}{1 + \phi y_i}\right) + (y_i - 1) \log(1 + \phi y_i) - \frac{y_i (1 + \phi y_i)}{1 + \phi y_i} - \log(y_i!) \right\} \quad (121)$$

R_G^2 log-benzerlikteki mümkün olan maksimum artışın açıklanmasını ifade etmektedir. Sahte R^2 'in hesaplanılmasında kullanılan ikinci yöntem aşağıdaki gibi olmaktadır.

¹⁶²Thomas Laitila, "A Pseudo- R^2 Measure for Limited and Qualitative Dependent Variable Models", *Journal of Econometrics*, No: 56, S: 3, Nisan 1993, s. 341.

$$R_p^2 = 1 - \frac{\chi_p^2}{\chi_0^2} \quad (122)$$

$$\chi_0^2 = \frac{(n \text{ var } (y))}{\bar{y}} \quad (123)$$

var (y) ve \bar{y} , Y'nin varyansı ve ortalamasını ifade etmektedir.

$$\chi_p^2 = \frac{\sum (y_i - \lambda_i)^2}{\lambda_i} \quad (124)$$

χ_p^2 , Pearson χ^2 istatistiğidir. Yüksek değerdeki R_G^2 ve R_p^2 modelin en iyi model olduğunu göstermektedir.

3.5.2.4 Bayes Bilgi Ölçütü

Benzerlik Oran testi (G^2), parametre tahmincilerinin sınamalarında yaygın olarak kullanılan bir test olmaktadır. Benzerlik Oran testi uyum iyiliği sınamalarında kullanıldığında, çalışılan örneklem büyük bir örneklem ise daha karmaşık bir modeli kabul etmek bu test için daha kolay olmaktadır. Çünkü, Benzerlik Oran testi gözlem verileri ve model arasındaki herhangi bir ayrılığı ortaya çıkarmayı amaçlamaktadır. Modele daha fazla terim ilave edilirse modelin uygunluğu artacaktır. Fakat büyük bir örneklemde önemsiz bir uygunluk olduğunda, modelin uygulandığı artışın gerçek bir artış olduğunu tartışmak oldukça zor olmaktadır. Bu düşünceyle, Benzerlik Oran testi

¹⁶³ Wang ve Famoye, s. 279.

Heinzl, Harald ve Martina Mittlböck, "Pseudo R-Squared Measures for Poisson Regression Models with Over- or Underdispersion", *Computational Statistics and Data Analysis*, No: 44, 2003, s. 253-258.

kabul edilebilir modelleri sıklıkla ret etmektedir. Bu probleme bir çözüm olarak verilere yeterli bir uygunluk sağlayan aşırı tutumlu modellerde (parsimonious models) Bayes Bilgi Ölçütü kullanılması önerilebilir¹⁶⁴.

Bayes Bilgi Ölçütü, veriler ve model arasında uygunluğu ölçen bir terim ortaya koyulmasını amaçlamaktadır. Böylece bu ölçütün maksimizasyonu, karmaşık ve uygunluk arasında uzlaşma sunan bir modelin seçimine olanak tanımaktadır¹⁶⁵.

Bayes Bilgi Ölçütü, büyük örneklemin Bayes Bilgi Ölçütü ve Bayes faktörünün logaritmasının iki katına yaklaştırılması olarak tanımlanabilir.

Bayes Bilgi Ölçütü = -2 (modelin benzerliğinin log maksimizasyonu) + 2 (parametre sayısı) * log (n)

$$\text{Bayes Bilgi Ölçütü} = G^2(X) + 2 (\text{parametre sayısı}) * \log (n) \quad (125)$$

Bayes Bilgi Ölçütünün özünde yatan düşünce, Bayes faktörüne bir Laplace yaklaşımı içermesi olarak ifade edilebilir. Aynı zamanda bu ölçüt, modelin sabit terimini göz ardı etmektedir. Bayes Bilgi Ölçütü ve Bayes faktörünün logaritmasının iki katı arasındaki farklılık genellikle asimptotik olarak yok olmaktadır. Hiçbir kısıtı olmayan düzenli şartlar altında Bayes Bilgi Ölçütü ve Bayes faktörünün logaritmasının iki katı arasındaki farklılık, ön parametrelerin seçiminde sıfır eğiliminde olmaktadır¹⁶⁶.

¹⁶⁴ Powers ve Xie, s. 106.

¹⁶⁵ Jean-Baptiste Duran ve Oliver Gaudoin, "Software Reliability Modelling and Prediction with Hidden Markov Chains", **Statistical Modelling**, No: 5, 2005, s. 83.

¹⁶⁶ Chris T. Volinsky ve Adrian E. Raftery, "Bayesian Information Criterion for Censored Survival Models", **Biometrics**, No: 56, Mart 2000, s. 256-257.

Sik-Yum Lee ve Xi-Yuan Song, "Maximum Likelihood Estimation and Model Comparison for Mixtures of Structural Equation Models with Ignorable Missing Data" **Journal of Classification**, No: 20, 2003, s. 232-233.

3.5.2.5 Akaiki Bilgi Ölçütü

İstatistiksel modelleri kapsayan bir bilgi ölçütü Akaiki tarafından önerilmiştir. Akaiki'nin yol gösterici çalışması, model seçimi için temel kaynak olarak Kullback-Leibler bilgi veya uzaklık kullanımını önermektedir. Aday modellerin her birindeki parametre ve f (tam gerçeklik) hakkında tam bir bilgi olmadan hesaplanamamaktadır. Akaiki, modelin maksimum noktasında deneysel benzerlik fonksiyonunu temel alarak K-L bilgisini tahmin etmenin güçlü bir yolunu bulmuştur. Verilen parametrik yapıya sahip bir modelde K-L aralığını minimize eden parametre değeri tektir. Bilinmeyen parametrenin minimizasyonu; f , parametre ve örneğe bağlı olmaktadır. Bu düşüncede maksimum benzerlik tahminleri altında parametre tahmincisinin değeri doğru olarak kabul edilmektedir.

Akaiki model seçim hedefinin bir tahmincisi olarak, log benzerliğin maksimizasyonunu artı yönde hata (Upward bias) olarak ifade etmiştir. Artı yönde hata yaklaşık olarak, yaklaştırılan modeldeki tahmin edilebilir parametre sayısı K 'ya eşit olarak kabul edilmektedir. Akaiki'nin log benzerliğin maksimizasyonu ve nispi beklenen K-L uzaklığı arasındaki ilişkiye ait bulguları, karmaşık veri setleri analizi ve model seçiminde uygulamalı ve teorik ilerlemeleri kabul etmektedir¹⁶⁷.

Akaiki Bilgi Ölçütü (AIC), bilgiyi maksimum tutan model seçimini savunmaktadır. Akaiki Bilgi Ölçütünün maksimizasyonu, model X 'in en aza indirgenmiş seçimine bağlıdır.

$$\text{Akaiki Bilgi Ölçütü} = G^2(X) - [q - 2r] \quad (126)$$

¹⁶⁷ Kenneth Burnham, **Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach**, New York: Springer-Verlag, 2002, s. 60; Geert Verbeke, **Linear Mixed Models for Longitudinal Data**, New York: Springer-Verlag, 2000, s. 74 -75.

Modeldeki “ $G^2(X)$ ” doyurulmuş modele karşı X model için benzerlik oran test istatistiğidir. “r”, X model için serbestlik derecesidir. Doyurulmuş modelin serbestlik derecesi “q” ile ifade edilmektedir.

$$\begin{aligned} \text{Akaiki Bilgi Ölçütü } -q &= G^2(X) - 2 [q - r] \\ &= G^2(X) - 2 (\text{testin serbestlik derecesi}) \quad (127) \end{aligned}$$

“q” değerinin model X’e bağlı olmaması yüzünden “Akaiki Bilgi Ölçütü – q” değerinin minimizasyonu, Akaiki Bilgi Ölçütünün minimizasyonuna eşit olmaktadır. Doyurulmuş modelde “Akaiki Bilgi Ölçütü – q = 0” olmaktadır¹⁶⁸.

$$\text{Akaiki Bilgi Ölçütü} = G^2(X) + 2 (\text{modelin parametre sayısı}) \quad (128)$$

Bu ölçüt, modele ilave edilen değişken sayısıyla azalmaktadır. Aynı zamanda modele ilave edilen değişken gereksiz olduğunda ölçüt artış göstererek bir sinyal vermektedir. Akaiki bilgi ölçütü, en küçük kareler regresyon denklemindeki düzeltilmiş R^2 ’yi hatırlatmaktadır. Her ikisinde de modellerdeki parametre sayısı ile uyum iyiliği denmektedir¹⁶⁹.

¹⁶⁸ Ronald Christensen, **Log-Linear Models & Logistic Regression**, New York: Springer-Verlag, 1997, s. 106.

¹⁶⁹ Melvin Moeschberger, **Survival Analysis**, New York: Springer-Verlag, 1997, s. 253-254.

4.ÇANAKKALE'DE YÖNETİCİLİK SÜRESİNİ ETKİLEYEN FAKTÖRLER

Bu uygulamanın amacı, yöneticilerin hizmet verdikleri organizasyonlardaki görev sürelerinin devamlılığına ilişkin bir olay zamanı modeli geliştirmek ve görev yapılan organizasyondaki yöneticiliğe devam etmede etkili olan faktörleri belirleyebilmektir.

Yöneticiler kurumları açısından hayati öneme sahiptirler. Kurumlar yöneticilerin başarısı ve kararları ile rekabet gücü kazanırlar yada kaybederler. İşletmelerde özellikle yöneticilerin görevlerini bırakmalarını yada devam etmelerini sağlayan birçok faktör mevcuttur. Bu faktörlerin bir kısmı kurumsal faktörlerdir. Diğer kısmı ise, yönetici ile ilgili faktörler olarak karşımıza çıkmaktadır. Bu çalışmada, organizasyonların piyasadaki faaliyetlerini etkileyen yöneticilerin, organizasyonla olan bağlarının devamlılığı, birçok boyut açısından ele alınmaya çalışılacaktır. Hangi faktörlerin yöneticilerin görevlerini sürdürmelerinde yada bırakmalarında etkili olduğu ortaya konmaya çalışılacaktır.

4.1 Teorik Çerçeve

Organizasyonların piyasada ortaya çıkışı veya yok olması farklı dinamiklerin etkisi altında kalır. Organizasyonların resmi olarak ortaya çıkması, doğum olarak ifade edilirken, organizasyon ölümü ise farklı tüzel kişi olarak organizasyonun, faaliyetini durdurması olarak tanımlanmaktadır. Piyasalarda, son zamanlarda çok yaygın olan birleşmeler (mergers), bir organizasyon ölümü olarak adlandırılabilir¹⁷⁰. Bu süreçlerin gelişmesinde ve yönetilmesinde yöneticilerin büyük rollerinin oldu görülmektedir.

¹⁷⁰ T. L. Amburgey ve H. Rao, “*Organizational Ecology: Past, Present and Future Direction*”, **Academy of Management Journal**, No: 39, S: 5, 1996, s. 1265-86.

İşletmelerin başarısızlığına yönelik yapılmış çok çeşitli çalışmalar mevcuttur. Yapılan çalışmalarda örgütsel sistemin farklı parametreleri üzerine odaklanarak, örgütsel süreçlerin faaliyetini olumsuz etkileyen faktörler, çeşitli yönleri ile ortaya konulmaya çalışılmıştır. Longenecker, Simonetti ve Sharkey¹⁷¹ yaptıkları çalışmalarında organizasyon başarısızlıklarının altında yatan nedenleri “tepe yönetimin başarısızlığı”, “tüketici ve pazarın taleplerini karşılamada kalınan yetersizlikler”, “finansal başarısızlıklar” ve “sistem başarısızlıkları” olmak üzere dört ana başlıkta toplamışlardır.

Tepe yönetimin başarısızlığı, organizasyonların başarısız olmasındaki en önemli sebep olarak görülmektedir. Organizasyon için alınan kararlarda ve yeni vizyon yaratılmasında karşılaşılan başarısızlık, kurumların başarısızlığının en önemli gerekçelerinden biri olarak karşımıza çıkabilir. Birçok örgütün başarısızlığının temel nedeni, sorunlar karşısında önceden harekete geçememe ve organizasyondaki değişim ihtiyacının zamanında algılanmaması sonucunda, başarısızlıkla karşı karşıya kalması olarak görülebilir. Bununla birlikte birçok örgütün etkin stratejiler geliştirmede yetersiz oldukları görülmektedir. İş dünyasındaki veya pazardaki gerçekleri görmedeki ilgisizlik, iş dünyası ile ilgili tahmin ve planların yapılmasında yetersiz kalınması, etkin organizasyon kültürünü güçlendirmek için yapılan çalışmalardaki yeteneksizlik, kendi halinde kalan ve kendine yeten bir organizasyon olarak kalmayı kabullenmek, politik rekabete izin vermek ve yetki problemleri gibi bir çok neden yönetim başarısızlıklarına sebep olabilmektedir. Bu faktörler genel olarak örgütlerin piyasadaki faaliyetlerini olumsuz yönde etkileyerek örgütsel operasyonların olumsuz yönde etkilenmesine yol açabilir.

Organizasyon başarısızlıklarını belirlemede etkili olan bir başka sebep, tüketici ve pazarın taleplerini karşılamada kalınan yetersizliklerdir. Özellikle müşterilerin beklentilerinin karşılanmaması ve rakiplerin stratejilerine gerekli tepkinin verilememesi, örgütler açısından birçok alanda olumsuzluklar yaratacaktır. Tüketici isteklerini anlamadaki başarısızlık, rekabetteki bilgisizlik, zayıf fiyat politikaları, etkin olmayan bir pazarlama planı, pazar geliştirmede hızlı karar alamama, kar yerine satış rakamları üzerinde odaklanma yeni ürün geliştirilmesinde ve pazara tanıtılmasında yaşanan

¹⁷¹ Clinton O Longenecker, Jack L. Simonetti ve Thomas W. Sharkey, “Why Organizations Fail: The View from The Front-Line”, **Management Decision**, No: 37, S: 6, 1999, s. 503-506.

sorunlar ve buna benzer pek çok neden organizasyonun tüketici portföyünde ve pazar yapısında başarısızlıklara sebep olabilmektedir.

Finanssal yönetiminin de zayıf olması, organizasyonun başarısızlığına sebep olan bir başka neden olarak ifade edilebilir. Nakit akışı problemleri, etkisini yitirmiş olan finanssal kontrol sistemlerinin kullanılması, zayıf kredi değerlendirmeleri, kar marjlarındaki yıpranma ve büyüme için yapılan planlamadaki yetersizlik ve daha pek çok sebep finanssal yönetim başarısızlığı olarak değerlendirilebilir.

Son olarak, sistem ve yapı başarısızlıkları da organizasyon başarısızlıklarında incelenen bir konu olmaktadır. Teknolojik problemler, etkin olmayan bir bilgi sistemi yönetimine sahip olmak, zayıf kontrol sistemleri, iç rekabet ve çözümsüz olarak kalan performans engelleri organizasyonun sistem ve yapı başarısızlığını sağlayan nedenlerden bazıları olarak gösterilebilir.

Organizasyonlar yaşayan varlıklar olarak kabul edildiklerinde; rekabet ortamında bulunan organizasyonların yaşamlarını sürdürebilmeleri yönetsel başarıları ile mümkün olmaktadır. Yöneticiler ile organizasyonlar arasındaki bağın devam etmesi karşılıklı olarak memnuniyetlerinin sürmesi ile sağlanabilir. Konuya organizasyonların rekabet ortamında yaşamlarını sürdürmeleri olarak bakıldığında; pazardaki paylarını arttırmak isteyen organizasyonların değişen koşullara göre olumlu yönde değişim göstermelerini sağlayan yöneticilerle çalışma istekleri görülebilir.

Organizasyonlardaki değişim farklı şekillerde tanımlanabilir. Fossum¹⁷² organizasyonlardaki değişimi; organizasyon içerisinde yapılan önemli faaliyetlerin farklılaştırılması olarak ifade etmiştir.

Hannan ve Freeman¹⁷³ organizasyon değişimlerinde, organizasyonların amaç değişimi ve yapısal değişimin önemi üzerinde yoğunlaşmışlardır. Sing ve House¹⁷⁴ yaptıkları çalışmada organizasyon değişimlerini altı grupta değerlendirmişlerdir.

¹⁷² Lynn. B. Fossum, **Understanding Organizational Change Converting Theory to Practice**, USA: Crisp Publications, 1989, s. 3.

¹⁷³ Michael T. Hannan ve John Freman, **Organizational Ecology**, Cambridge: Harvard University Press, 1989, s. 10.

1.Organizasyondaki yönetici değişimi: Farklı seviyelerde organizasyonlardaki yöneticilerin değişmesi ve bunların yerlerine yeni yöneticilerin gelmesi genel olarak yönetici değişimi olarak ifade edilebilir.

2.Organizasyonun hizmet alanındaki değişim: Rekabet koşulları yada stratejik kararlar organizasyonların hizmet alanlarını farklılaştırabilirler.

3.Organizasyonun amaçlarındaki değişim: Amaçlar organizasyonların temel faaliyetlerini ve bunların şeklini ortaya koyan faktörlerdir. Zaman zaman organizasyonların amaçlarında farklılaşmalar söz konusu olmaktadır.

4. Organizasyonun sponsorluklarındaki değişim: Organizasyonu destekleyen ve desteklediği faaliyetler zaman içerisinde koşullara bağlı olarak değişebilmektedir.

5. Organizasyonun coğrafik faaliyetindeki değişim: Organizasyonlarda büyüyen yapı ve farklı pazarlardaki fırsatlar coğrafik olarak faaliyet alanının genişlemesine neden olmaktadır. Böylece çok çeşitli coğrafik alanlara yönelen organizasyonlar önemli bir değişim süreci yaşamış olurlar.

6. Organizasyondaki yapısal değişim: Organizasyonlar daha yaratıcı olmak, verimliliği ve rekabet gücünü geliştirmek için dikey bir örgütlenme yerine daha basık bir örgütlenmeyi tercih ederler. Bu tür yapılanmalar organizasyonlarda önemli değişimlere kaynaklık da etmiş olur.

Burada hatırlanması gereken temel nokta, örgütlerdeki değişimin farklı alanlarda da olabileceğidir. Ancak yapılan uygulamada, temel olarak yönetsel alana odaklandığı için daha çok yönetsel süreçlerle ilgili değişim üzerinde etkili olan faktörler vurgulanmaya çalışılacaktır.

Yöneticilik modellerinin farklı şekilde incelenmesine yönelik olarak çalışmalar incelendiğinde konunun farklı boyutlardan ele alındığı görülmektedir. Zaleznik¹⁷⁵, yöneticiler ile ilgili çalışmasında en önemli noktanın hayal kurma olduğunu

¹⁷⁴ Jitendra V. Singh ve Robert J. House, “*Organizational Change and Organizational Mortality*”, **Administrative Science Quarterly**, No: 31, 1986, s. 587-611.

¹⁷⁵A. Zaleznik, “*Managers and Leaders Are They Different*”, **Inside The Mind Of The Leader Harvard Business Review**, Ocak 2004, s. 74-81.

vurgulamaktadır. Bu kavram da iş dünyası ile ilgili konuları ve problemleri fırsata dönüştürme yeteneği şeklinde açıklanabilir. Bu süreç basit olarak çalışma ortamında bulunan elamanları problemleri çözmeye motive etme şeklinde tanımlanmaktadır. Sorunlar karşısında fırsat yaratma ve bunlara çözüm bulma en önemli noktalardır. Fırsatçılık, yöneticilere vizyonlarını pratik olarak ifade etme olanağı sağlar. Hayal kurma yeteneği, neyin olması gerektiğini aynı zamanda görselleştirme çabası olarak da değerlendirilmektedir. Yöneticiler, iyimser olmak durumundadır. Çünkü organizasyonda görev alan birçok bireyi yöneticilerin vizyonları şekillendirmektedir.

Konu ile ilgili yapılan başka bir önemli çalışmada ise¹⁷⁶, yöneticilik ile ilgili olarak günümüzün kurumlarında yöneticiliğe daha çok ihtiyaç duyulduğu ortaya çıkarılmıştır. Yöneticilerin faaliyetleri gerçekleştirmek için genel olarak taşınmaları gereken özellikler yapılan bu çalışmalarda kapsamlı bir şekilde açıklanmaktadır.

Aynı konu ile ilgili bir başka çalışmaya bakıldığında ise¹⁷⁷, konunun değerlendirilmesi sağlık sektörü çalışanları tarafından ele alınarak yapılmıştır. Hemşirelere yönelik olarak yapılan araştırmada değişim yönetimi arasındaki ilişkiye vurgu yapılmaktadır. Yöneticilik konusu bir başka çalışmada, kadınlar açısından kapsamlı bir şekilde incelemiştir¹⁷⁸. Bu incelemede kadın yöneticilerin, daha önce erkeklerin çalıştıkları pozisyonlarda hızla ilerleyerek yer aldıkları görülmektedir. Son araştırmalar ise özellikle cinsiyet ile yöneticilik arasında bir ilişkinin olup olmadığını ortaya çıkarmaya yönelmiştir. Bu durum özellikle cinsiyetin organizasyon içindeki durumunu zenginleştirmektedir. Bir başka çalışma¹⁷⁹, konuya ruhsal ve psikoloji açısından dikkat çekerek yöneticilik sürecinin farklı bir boyutu üzerinde durmaktadırlar. Bu çalışmada, diğer çalışmalarda olduğu gibi yöneticiliğin moral ve motivasyon açısından önemini ifade etmektedir. Organizasyonlar açısından konuyu değerlendiren

¹⁷⁶ M. N. Tichy ve A. M. Devanna, “*The Transformational Leader*”, **The Manager’s Bookshelf: A Mosaic of Contemporary Views**, 1994, s. 244-255.

¹⁷⁷ J. Apker, “*Sensemaking Of Change In The Managed Care Era: A Case Of Hospital-Based Nurses*”, **Journal Of Organizational Change Management**, No: 17, S: 2, 2004, s. 211-227.

¹⁷⁸ R. Kark, “*The Transformational Leader: Who is (s)he? A Feminist Perspective*”, **Journal of Organizational Change Management**, No: 17, S: 2, 2004, s. 160-176.

¹⁷⁹ S. K. Chakraborty ve D. Chakraborty, “*The Transformed Leader And Spiritual Psychology: A Few Insights*”, **Journal of Organizational Change Management**, No: 2, S: 17, 2004, s. 194-210.

bir diğ er ç alıřmada ise¹⁸⁰, yöneticilerin ilgi alanları beř faktör etrafında toplanmıřtır. Bunlar; yöneticinin kendisi, organizasyonlar, içerik, iliřkiler ve deęiřimdir.

Goleman ise¹⁸¹ yaptıęı deęerlendirmede her hangi bir yöneticilik modelinin kesin sonuçları doęurduęuna ait kanıtların olmadıęını ifade etmektedir. Yöneticilięin deęiřen rollerini ele alan ç alıřmaya göre de¹⁸², zaman içerisinde organizasyon ile ilgili bir ç ok alanda yařanan geliřmelerin yöneticilerin rollerini yeniden tanımlamalarını gerekli kılmaktadır. Arařtırmada; ç alıřanlara, organizasyonel yapıya ve bilgi akıřına yönelik geliřmeler yöneticilik açısından ele alınmaktadır. Yöneticilerin faaliyetlerini süreçsel olarak deęerlendiren ç alıřma ise¹⁸³, konuyu zaman süreci içerisinde liderlerin davranıř, tutum ve yöneltme anlayıřlarını ele almaktadır. Yöneticilerin faaliyetlerini inceleyen bařka bir ç alıřmada¹⁸⁴, yöneticilerin gerçekten neler yaptıęı açıklanmaya ç alıřılmaktadır. Yöneticiler, organizasyonlarda özellikle deęiřimi yaratan ve gerçekteřiren unsur olarak deęerlendirilmektedir. Ayrıca liderler ile organizasyonel unsurlar arasında baęlantı kurularak bunları ortak amaçlara yönelik olarak planladıęı ortaya konmaktadır.

Conger ve Fulmer¹⁸⁵ ç alıřmalarında, organizasyonları bařarı ve bařarısızlıęa yönlendiren unsurları dikkate alarak, dinamik ve geliřtirici bařarı planlarına sahip organizasyonların mekanik süreç ve yapılara göre daha bařarılı olduklarını ortaya koymuřlardır. Organizasyonel ekolojide rekabet ve yönetim stratejisinin belirlenmesi ve bařarılı bir řekilde uygulanması, liderlerin önemli ç alıřma alanları arasında yer almaktadır. Doęru strateji, rekabet ve yeniliklere karřı örgütlerin alternatif yaratmalarına olanak saęlamaktadır. Konu ile ilgili ç ok ç eřitli ç alıřmaların bu alanda varolduęu görölmektedir. Özellikle yöneticilerin kiřilik özelliklerinin iř yerinde diğ er

¹⁸⁰ J. Gosling ve H. Mintzberg, “*The Five Minds of A Leader*”, **Harvard Business Review**, Kasım 2003, s. 54-61.

¹⁸¹ D. Goleman, “*What Makes A Leader?*”, **Harvard Business Review**, (Kasım-Aralık 1998), s. 93-102.

¹⁸² A. C. Barlett ve L. S. Ghosha, “*Changing The Role Of Top Management: Beyond Systems To People*”, **Harvard Business Review**, Mayıs-Haziran, 1995, s. 132-142.

¹⁸³ B. Bennis, “*Why Leaders Can't Lead*”, **The Manager's Bookshelf: A Mosaic of Contemporary Views**, 1994, s. 229-236.

¹⁸⁴ J. P. Kotter, “*What Leaders Really Do*”, **Breakthrough Leadership**, **Harvard Business Review**, Aralık 2001, s. 85-97.

¹⁸⁵ A. J. Cognier ve M. R. Fulmer, “*Developing Your Leadership Pipeline*”, **Harvard Business Review**, Aralık 2003, s. 76-84.

bireyleri hangi alanlarda etkilediği¹⁸⁶; örgütsel performans ve yönetici arasındaki ilişkileri değerlendiren çalışmaların sonucunda ortaya çıkan faktörler arasında pozitif ilişkilerin varolduğu¹⁸⁷; tepe yöneticiler ve çevre arasındaki ilişkilerinin ele alındığı çalışmada, stratejik basamağın örgütsel sistemin işleyişini yönlendirmedeki etkileri değerlendirilmektedir. Yöneticilerin uzmanlık alanlarının strateji ve organizasyonel başarı süreci üzerindeki etkileri, dünyadaki önemli şirketlerde çalışan yöneticiler açısından ele alınarak değerlendirilmeye çalışılmıştır¹⁸⁸. Başka bir çalışmada tepe yönetimin demografik yapısının sosyal entegrasyon ve iletişim ile ilişkisi incelenmiştir¹⁸⁹. Tepe yönetim basamağının heterojenliğini inceleyen çalışma, bu yapının organizasyonel rekabet süreci üzerindeki etkilerini ortaya koymaya çalışmaktadır¹⁹⁰.

Yapılan literatür taraması sonucunda, yöneticilerin organizasyondaki görev sürelerinin devamlılığına ilişkin bir olay zamanı modellemesine ilişkin bir araştırmaya rastlanmamıştır. Yöneticilerin çalıştıkları organizasyonla olan bağları incelenirken gözlemlenemeyen bir heterojenlik olup olmadığı bu alanda bir eksiklik olarak görülülerek bu konunun çalışılması uygun görülmüştür. Ayrıca, yukarıdaki araştırmalarda da vurgulandığı gibi yöneticiler kurumlar açısından büyük bir öneme sahiptirler. Kurumlar yöneticilerin vereceği kararlar ile geleceğe yönelik olarak çalışmalarını yönlendirirler. Konunun farklı açılardan ele alınması hem organizasyonel yaşam açısından büyük bir önem taşımaktadır, hem de konunun metodolojik olarak geliştirilmesinin alana önemli düzeyde katkı vereceği düşünülmektedir. Böylece yöneticilerin çok yönlü olarak değerlendirilmesinin ortaya çıkan sonuçlar çerçevesinde mümkün olabileceği düşünülmektedir.

¹⁸⁶ M. Lansiti ve R. Levien, “*Strategy as Ecology*” **Harvard Business Review**, Mart 2004, s. 68-78.

¹⁸⁷ A. B. Thomas, “*Does Leaders Make A Difference To Organizational Performance?*”, **Administrative Science Quaterly**, No: 33, S: 3, Kasım 1988, s. 388-400.

¹⁸⁸ A. M. Geletkanycz ve C. D. HambricK, “*The External Ties Of Top Executives: Implications For Strategic Choice And Performance*”, **Administrative Science Quaterly**, No: 42, S: 4, Aralık 1997, s. 654-681.

¹⁸⁹ M. Smith ve C. M. White, “*Strategy, CEO Specialization, and Succession*”, **Administrative Science Quaterly**, No: 32, S: 2, Haziran 1987, s. 263-280.

¹⁹⁰ G. K. Smith, K. A. Smith, D. J. Olian, Jr. P. H. Sims, P. D. O’bannon ve A. J. Scully, “*Top Management Team Demography and Process: The Role of Social Integration and Communication*”, **Administrative Science Quaterly**, No: 39, S: 3, Kasım 1994, s. 412-438.

4.2 Araştırmanın Kapsam ve Kısıtları

Bütün bilimsel çalışmalarda özellikle verilerin toplanması ve değerlendirilmesi aşamasında bazı temel zorlukların olduğu görülmektedir. Verilerin düzgün ve etkin bir şekilde kayıt işleminin yapılmaması, bununla birlikte veri yaratma sürecinin kendi yapısından kaynaklanan zorluklar dikkate alınarak ve konu ile ilgili duyarlılık çerçevesinde araştırmamıza konu olan veriler üzerinde büyük titizlikle durulmuştur. Veri yapısında üst düzey yöneticiler, işletme sahipleri ve işletme bölümü öğretim üyelerinden yoğun bir şekilde destek alınarak çalışmanın bu alandaki gereksinimleri ortaya konmaya çalışılmıştır. Ayrıca verilerin daha sağlıklı olabilmesi için mümkün olduğunca yerel düzeydeki veri kaynaklarından yararlanılma yoluna gidilmiştir. Yerel çevredeki veri kaynakları ile birlikte uzun vadeli ve kalıcı ilişkiler kurularak kendilerine güven verilmeye çalışılmıştır. Bunun sonucu olarak da verilerin güvenilirlik ve geçerlilik seviyesinin yükseltilmesi amaçlanmıştır.

Araştırmamızın veri yapısını oluşturan olay zamanı analizinde, olaylar belirli bir zaman aralığı içinde gözlemlenerek yada bu zaman aralığı içinde olaylar hakkında kayıtlı veri varsa onlar kullanılarak analiz yapılabilir. Analiz için uzun bir zaman aralığı inceleniyorsa, geçmişle ilgili olan bu tür verilere her zaman ulaşmak bu kadar kolay olmayabilir. Böyle durumlarda, kayıtlı bilgilerin yanı sıra olayın gerçekleşmesi sırasında olaya şahit olan kişi veya kişilerle de görüşme yapılarak veri derleme işlemi tamamlanabilir.

Uygulamaya başlamadan önce ön çalışma olarak Çanakkale ve ilçelerinde bulunan beş işletmenin sahipleri ve yönetim kurulu üyeleri ile görüşülerek, üst düzey yönetim ve üst düzey yöneticiler hakkında genel bir değerlendirme yapılmıştır. Bu değerlendirmelerle birlikte, yönetim alanında yapılan çalışmalar için yapılan literatür incelemesi ile organizasyon teorisi çerçevesinde birleştirilerek bir görüşme formu hazırlanmıştır (Bkz. EK-1). Uygulama, Çanakkale merkez ve tüm ilçelerinde bulunan profesyonel olarak üst düzey yönetici çalıştıran 120 işletme üzerinde yapılmıştır. Veriler olay zamanı veri seti olarak düzenlenmiş olup, işletmelerin kuruluşundan bu yana

çalıştırdıkları yöneticilere ilişkin sorular hazırlanan görüşme formu aracılığı ile organizasyon sahipleri ve genel kurul üyeleriyle görüşme yapılarak elde edilmiştir. İşletmelerden randevu alınarak görüşmeler gerçekleştirilmiştir. Çalışmadaki en büyük zorluklardan birisi veri toplama aşamasında olmuştur. İşletmelerle yapılan görüşmeler, işletmenin kuruluşundan bu yana çalıştırdıkları tüm yöneticileri kapsadığı için görüşmeler yaklaşık olarak 45-50 dakika sürmüştür. Böylece her işletme ile ilgili olarak çok detaylı bir veri yapısı oluşturulmaya çalışılmıştır. Daha sonraki aşamada da bu verilerin kodlanması ve değerlendirilmesi yapılarak analiz için hazır hale gelmeleri sağlanmıştır.

Uygulamada kullanılan veriler SAS 8.0 programında değerlendirilirken veri girişleri şekil 5’de görüldüğü gibi olmaktadır. Tüm değişkenler, her işletme için gruplandırılmış veri olarak kodlanmıştır. İşletme de çalışmış olan tüm yöneticilere ait veriler aynı işletmeye verilen kod altında gösterilmiştir. Şekil 5’ in ilk sütunu işletmeye verilen sıra numarası olarak kodlanmıştır. Bu sıralamanın kodlamada herhangi bir özel durumu olmayıp tamamen tesadüfi olarak kodlama yapılmıştır.

```
cards;
```

	x_1	x_2	x_3	λ	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}
1	50	17	47	2	1	2	1	2	1	0	4	3	3	1
1	40	19	40	1	1	2	1	2	1	2	4	3	4	1
1	60	20	37	1	1	1	1	2	3	0	4	3	3	1
2	45	7	32	7	1	1	20	1	4	4	4	4	5	2
2	50	14	40	8	1	1	20	1	4	4	4	4	5	5
3	70	5	40	5	2	2	1	3	1	0	4	3	3	2
3	80	10	43	5	2	2	1	3	2	0	4	3	3	2
3	90	15	45	5	2	2	1	3	2	0	4	2	2	2
3	70	20	48	5	2	1	1	3	5	1	4	3	1	2
4	38	3	28	3	2	1	20	1	1	0	4	3	1	3
4	45	4	35	1	1	2	20	1	3	1	4	3	1	3
4	50	11	30	7	1	1	25	3	4	2	4	2	1	3
4	55	12	48	1	1	1	25	1	4	1	4	3	2	3
.....														
.....														
.....														
.....														

Şekil 5. SAS Programına Veri Giriş Sayfası

Şekil 5 incelendiğinde, ilk işletmeye ait üç yönetici bulunduğu için veri sayfasının ilk üç satırı “1” kodu ile gösterilmiştir. Modelde yer alan bağımlı değişken, on üç açıklayıcı değişken ve işletmenin kodu ile birlikte veri sayfası on beş sütundan oluşmuştur. İkinci sütun işletme büyüklüğüdür. İşletmede çalışan toplam personel sayısı büyüklük ölçütü olarak belirlenmiştir. Üçüncü sütun işletme yaşıdır. İlk yöneticinin işletme ile bağları tamamlandığındaki yaşı olarak alınmıştır. Dördüncü sütun yöneticinin işletmeden ayrılma yaşıdır. Beşinci sütunda yöneticilerin işletmede çalışma süreleri gösterilmiştir. Toplantı yapılıp yapılmadığı altıncı sütunda “1=evet”, “2=hayır” olarak kodlanmıştır. Danışmanlık hizmeti alınıp alınmadığı yedinci sütunda “1=evet”, “2=hayır” olarak gösterilmiştir. Sekizinci sütunda işletmenin bölgedeki satış payı yüzde olarak yer almıştır. Dokuzuncu sütunda işletmenin yönetici dönemindeki rekabet durumu “1=artma”, “2=azalma”, “3=sabit” şeklinde kodlama yapılmıştır. İşletmenin faaliyet gösterdiği il sayısı onuncu sütunda görülmektedir. Onbirinci sütunda, işletmenin işbirliği içinde bulunduğu diğer işletmelerin sayısı yer almaktadır. Onikinci sütunda likert ölçeği ile ölçülmüş olan yöneticinin yönetim yaklaşımı “1=istismarcı-otokratik”, “2=yardımsever-otokratik”, “3=katılımcı” ve “4=demokratik” olarak gösterilmiştir. Yöneticinin eğitim durumu onüçüncü sütunda “1=orta öğretim”, “2=lise”, “3=önlisans-lisans” ve “4=lisansüstü” olarak

Çalışmada kullanılacak olan açıklayıcı değişkenler yönetim ve organizasyon teorisi çerçevesinde, yönetsel unsurların yönetici değişimi üzerinde etkilerini ortaya koymak amacıyla belirlenmiştir. Uygulamada kullanılan değişkenler Tablo 1’de görülebilir. X değişkenleri modelde açıklayıcı değişken olarak yer almıştır. Tablo 1’de açıklayıcı değişkenlerin hangi ölçekle ölçüldükleri ifade edilmiştir. λ modelde yer alan bağımlı değişken olmaktadır. Olay zamanı modellerinde bağımlı değişken beklenen yönetici süresi olmaktadır.

Tablo 1

Uygulamada Kullanılan Değişkenler

λ : Beklenen Yöneticilik Süresi
X_1 : İşletmenin Büyüklüğü (Aralıklı ölçek)
X_2 : İşletmenin Yaşı (Aralıklı ölçek)
X_3 : Yönetici Yaşı (Aralıklı ölçek)
X_4 : İşletmenin Üst Düzey Yönetim Olarak Toplantı Yapıp Yapmadığı (Sınıflandırıcı ölçek)
X_5 : Yönetici Döneminde İşletmenin Danışmanlık Hizmeti Alıp Almadığı (Sınıflandırıcı ölçek)
X_6 : İşletmenin Bölgedeki Satış Payı (Aralıklı ölçek)
X_7 : İşletmenin Yönetici Döneminde Rekabetinin Artma, Azalma veya Sabit Kalması (Sınıflandırıcı ölçek)
X_8 : İşletmenin Faaliyet Gösterdiği İl Sayısı (Aralıklı ölçek)
X_9 : İşletmenin İşbirliği İçinde Bulunduğu Diğer İşletmelerin Sayısı (Aralıklı ölçek)
X_{10} : Yöneticinin Yönetim Yaklaşımı (Likert ölçeği)
X_{11} : Yöneticinin Eğitim Durumu (Sınıflandırıcı ölçek)
X_{12} : Yöneticinin Proje Üretimi (Sıralayıcı ölçek)
X_{13} : Yöneticinin İşletmeden Ayrılma Sebebi (Sınıflandırıcı ölçek)

Literatür incelendiğinde de olay zaman modeli ile ilgili olarak yapılmış olan simülasyon çalışmalarına da rastlanmaktadır. Ng ve Cook¹⁹¹ tekrar eden olaylar için Gamma, Log-Normal ve parametrik olmayan dağılımlar için karma Poisson tesadüfi

¹⁹¹ E.T.M.Ng ve R.J. Cook, "A Comparison of Some Random Effect Models for Parameter Estimation in Recurrent Events", **Mathematical and Computer Modelling**, No: 32, 2000, s. 11-26.

etki modellerinin karşılaştırılmasında Monte Carlo simülasyonunu kullanmışlardır. Munkin ve Trivedi¹⁹² ise çok değişkenli karma Poisson regresyon modellerini simülasyon çalışması altında değerlendirmişlerdir.

4.3 Yöneticilik Süresine Etki Eden Faktörler

Uzun süreç verilerinde bağımlı değişken, sayılabilir olay verilerinde ve olay zamanı verilerinde iki farklı şekilde tanımlanabilmektedir. Sayılabilir olay verilerinde olayın tekrar sayısı bağımlı değişken olarak incelenirken, olay zamanı verilerinde olayın ortaya çıkma uzunluğu bağımlı değişken olarak ele alınmaktadır.

Sayılabilir olay verilerinde, incelenen zaman aralığında olayın “hangi sıklıkla” meydana geldiği incelenmektedir. Bir ülkenin belirli bir zaman aralığında kaç kere savaşa katıldığı, herhangi bir hükümet döneminde meclis kararlarının cumhurbaşkanı tarafından kaç kez veto edildiği ve belirli yıllar arasında verilen yargı kararlarının sayıları, sayılabilir olay verileri olarak araştırma konusu olmaktadır.

Olay zamanı verilerinde ise, incelenen olay tipinin “hangi uzunlukta” meydana geldiği incelenmektedir. Bir ülkenin belirli bir zaman aralığında yaptığı savaşlarda savaşların süresi, belirli bir zaman aralığında herhangi bir üniversitedeki öğrencilerin yemekleri beğenmeme sebebi ile yaptıkları boykotun süresi ve özel sektörde çalışan mühendislerin kariyer uzunluğu gibi olayın süresini ifade eden pek çok araştırma konusu olay zamanı verileri olarak ifade edilebilir¹⁹³.

Çalışmada kullanılan bağımlı değişken işletmedeki üst düzey yöneticilerinin işletme ile olan bağlarının sürekliliğidir. Bağların sürekliliği, sonlanma ve başlama arasındaki zaman uzunluğuyla belirtilmektedir. Olay zamanı modelleri, başarısızlık olana kadarki zamanı belirtmek için kullanılmaktadır. İşletme ile yönetici arasındaki

¹⁹² Murat K. Mungin ve Pravin K. Trivedi, “*Simulated Maximum Likelihood Estimation of Multivariate Mixed-Poisson Regression Models, with Application*”, **Econometrics Journal**, 1999, No: 2, s. 29-48.

¹⁹³ Janet M. Box-Steffensmeier ve Bradford S. Jones, **Event History Modelling: A Guide for Social Scientists**, Cambridge: Cambridge University Press, 2004, s. 7-9.

bağ, bitişe kadar devam etmesine rağmen uygulamada süreksiz olay zamanı modeli kullanılmıştır.

Bu çalışma iki önemli nokta gerekçesinde, süreksiz olay zamanı modelinin kullanılmasını uygun görmektedir. Bunlardan ilki, bir yönetici ile işletme arasındaki bağ bir olay olarak düşünülürse, herhangi bir olay süreksiz zaman aralığı içinde gözlemlenmektedirler. Süreksiz olay zamanı modellerinde, olayın bitişine kadar olan zaman 1,2,3.... diye devam eden değerlerle süreksiz bir dağılımdır.

İkinci önemli nokta ise, uygulamada kullanılan verilerin zaman aralıklarıyla gruplandırılmış olmasıdır. Olay zamanı verilerinin analiz edilmesinde, sürekli zaman modellerinden iyi bilinen “Cox Oransal Hazard” modeli de kullanılabilirdi. Gruplanan veriler birden fazla bağıın görüntüsü yarattığı için bu çalışmada “Cox Oransal Hazard” modelinin kullanılması uygun bir model olarak görülmemiştir.

Literatür taraması yapıldığında; Kalbfleisch ve Prentice¹⁹⁴ olay zamanları sıklıkla aynı olduğu zaman, süreksiz olay zamanı modellerinin tercih edileceğini ifade etmişlerdir.

Baker, Faulkner ve Fisher¹⁹⁵,de çalışmalarında reklam ajansları ve onların müşterileri arasındaki bağlantıların sürekliliğini incelerken, süreksiz olay zamanı modelini kullanmışlardır.

Olay zamanı verilerinin modellenmesinde beş farklı açıklayıcı değişken yapısı ile karşılaşılabılır. Modeldeki değişkenler bu çerçevede içinde değerlendirildiklerinde “işletme yaşı” zaman değişkeni olarak modelde yer almaktadır. Çünkü, uygulamada kullanılan verilerin gruplandırılmış veriler olmasından kaynaklı olarak, olayın zaman içindeki ilerleme ölçütü işletmenin yaşı ile takip edilebilir.

“Yöneticinin yönetim yaklaşımı, yöneticinin eğitim durumu, yöneticinin proje üretimi, yöneticinin işletmeden ayrılma sebebi, yöneticinin üst düzey yönetim olarak toplantı yapıp yapmadığı ve yönetici döneminde işletmenin danışmanlık hizmeti alıp

¹⁹⁴ Kalbfleisch ve Prentice, s. 35.

¹⁹⁵ Baker, Faulkner ve Fisher, s. 164.

almadığı” değişkenleri sadece yöneticiden yöneticiye farklılık göstermekte, zamanın değişmesi bu değişkenler üzerinde etki yaratmamaktadır. Bu yüzden zamana göre sabit değişkenler olarak ifade edilebilir.

Zamana göre farklılaşan değişkenler ise “işletme büyüklüğü, işletmenin yönetici döneminde rekabetinin artma, azalma veya sabit kalması, işletmenin bölgedeki satış payı, işletmenin faaliyet gösterdiği il sayısı ve işletmenin işbirliği içinde bulunduğu diğer işletmelerin sayısı” olarak belirlenmiştir .

“Yönetici yaşı” zamana göre farklılaşan değişken olarak görülmesine rağmen yöneticilerin işletme ile olan bağlarının bitişindeki yaş olarak sabit alındığı için yada başka bir ifade ile zamana göre yönetici yaşı değişmediğinden dolayı, modelde içsel değişken olarak değerlendirilmiştir. Çünkü yöneticiler üzerindeki etkili zaman faktörü, yöneticiden yöneticiye farklılık göstermektedir. Uygulamada kullanılan değişkenler arasında dışsal değişken yapısına uygun değişken modelde yer almamaktadır.

1- İşletme büyüklüğü, belirgin ve tahmin edilebilir bir organizasyonel özelliktir¹⁹⁶. Yapılan çalışmalarda işletmenin büyüklüğü arttıkça ölme riskinin o oranda azalmakta olduğudur ortaya çıkarılmıştır¹⁹⁷. Çünkü, büyüklük işletmelerde rutinliği arttıracak seçilimin bu tür işletmelerin lehine olmasına neden olacaktır¹⁹⁸. Böylece büyük işletmelerin başarısızlıkla daha az karşı karşıya kalacağı düşünülebilir. Büyük işletmelerin, daha karmaşık yapılara, daha spesifik işlere, birimlerin farklılaşmasına ve yönetsel bileşenlerinin geliştiğine rastlanmaktadır¹⁹⁹. Yöneticilerin devamlılığına ilişkin yapılan çalışmada da, işletme büyüklüğünün etkisinin olup olmadığı modelde incelenmesi gerektiği düşünülmüştür.

2- İşletme Yaşı, kurumun kurulduğu yıl dikkate alınarak bugünkü yaşı olarak ele alınmıştır.

¹⁹⁶ Parac,J.F.,H. Thomas, F.Wilson, D. Paton ve A. Konter, “*Rivalry and the Industry Model of Scottish Knitwear Producers*”, **Administrative Science Quarterly**, No: 40, 1995, s. 209.

¹⁹⁷ J.A.C. Baum ve C. Oliver, “*Institutional Linkages and Organizational Mortality*”, **Administrative Science Quarterly**, No: 36, 1991, s. 191.

¹⁹⁸ Hannan ve Freeman, s. 149-64.

¹⁹⁹ Mintzberg, s. 230.

3- Yönetici Yaşı, yöneticinin biyolojik yaşının kaç olduğu araştırmanın bir başka değişkeni olarak ele alınmıştır.

4- İşletmenin Üst Düzey Yönetim Olarak Toplantı Yapıp Yapmadığı, yöneticinin işini profesyonel olarak algılayıp algılamadığını ölçen bir değişken olabilir. Ayrıca toplantılardan elde edilen sonuçlarla yapılacak olan ileriye dönük atılımlar veya problemlerin çözümünde kullanılacak olan geri dönüşümler, yöneticinin ölçme ve değerlendirme yeteneğini de ortaya koyabilir. Bu değişken yardımıyla, özellikle kurumsal olarak örgütte belirli zaman aralığında değerlendirmelerin olup olmadığı görülmeye çalışılacaktır.

5- Yönetici Döneminde İşletmenin Danışmanlık Hizmeti Alıp Almadığı, işletmenin kurumsallaşma düzeyini gösteren bir değişken olarak gösterilebilir. İşletmeler kurumsallaşmak ve rekabet güçlerini geliştirmek için uzman firmalardan destek alma yoluna giderler. Birçok alanda işletmeler bu örgütlerin görüş ve önerilerini dinleyerek zaman içinde stratejik planlamalarını yapabilirler.

6- İşletmenin Bölgedeki Satış Payı, işletmenin faaliyetlerini gerçekleştirdiği pazardaki payı bu değişkenle ele alınmıştır.

7- İşletmenin Yönetici Döneminde Rekabetinin Artma, Azalma veya Sabit Kalması, işletmelerde görev alan yöneticilerin rekabet gücü üzerindeki etkisinin olup olmadığını ölçen bir değişken olarak düşünülmüştür. Bu değişkenle yapısal rekabet sistemi üzerindeki etki görülmeye çalışılmıştır.

8- İşletmenin Faaliyet Gösterdiği İl Sayısı, işletmenin yönetici döneminde farklı pazarlara açılıp açılmadığı bu değişkenle ele alınmıştır.

9- İşletmenin İşbirliği İçinde Bulunduğu Diğer İşletmelerin Sayısı, işletmenin kaç tane farklı işletmenin ile işbirliği içerisinde faaliyetlerini sürdürdüğü ortaya çıkarılmaya çalışılmıştır.

10- Yöneticinin Yönetim Yaklaşımı, bir işletmedeki yöneticinin aynı zamanda lider özelliklerini de göstermesi işletme açısından çok önemli bir özellik olarak düşünüldüğünden modelde yer almıştır. Yönetici davranışı ile lider davranışı birbirini

tamamlayan bir unsur olarak düşünülürken, bir arada incelenmesi gerekliliği ortaya çıkmaktadır. Yönetici ve lider davranışlarının bir arada anlaşılmasını sağlayan ve aynı zamanda gruplaşma özelliğine de sahip olan bir model Rensis Likert'in geliştirdiği sistem1-sistem4 modelidir. Rensis Likert etkin olan organizasyonlarla etkin olmayan organizasyonları birbirinden ayırmak amacıyla ilgili olarak yaptığı çalışmalarda, bu model aracılığıyla organizasyonların yapısal ve davranışsal faktörlerini ortaya koymaya çalışmıştır. Bir yöneticinin davranışlarını; sistem1 olarak adlandırdığı "istismarcı-otokratik", sistem2 olarak "yardımsever-otokratik", sistem3 "katılımcı" ve son olarak da sistem4 "demokratik" olmak üzere 4 farklı şekilde gruplama yapmanın mümkün olduğunu göstermiştir. Likert'in geliştirdiği bu modele göre yöneticilerinin davranış biçimleri dört grup altında toplanabilir. Her grup belirli davranışları içerir. Yöneticiler ile ilgili yapılan bu dört grup sınıflandırma kısaca aşağıdaki gibi özetlenebilir.

Likert'in araştırmaları verimliliği yüksek olan organizasyonların sistem3 ve sistem4 tipi bir yönetim biçimi ile yönetildiğini, verimliliği düşük olan organizasyonların ise sistem1 ve sistem2 tipi bir yönetim tarzını seçtiklerini göstermişlerdir. Sistem 4 modeli ile ilgili olarak, bu sistemin organizasyonlarda uygulanmasının öngörülmesi ve her yerde daima geçerli en etkin yönetim tarzı olarak gösterilmesine ilişkin görüşleri eleştiri konusu olmuştur²⁰⁰.

11- Yöneticinin Eğitim Durumu, Eğitim seviyesi orta öğretim, lise, önlisans-lisans ve lisans üstü olarak değerlendirilmiştir. Yöneticinin eğitim durumunun bir değişken olarak modele katılması, üst düzey yöneticiler arasında eğitim farklılığının olmaması gerektiği düşüncesine yol açabilir. Fakat uygulamaya başlamadan önce yapılan ön çalışma neticesinde, organizasyon sahipleri tarafından iş dünyasında tecrübeli olarak değerlendirilen kişilerin üniversite eğitimi almamalarına rağmen yönetimde daha iyi oldukları iddiası bu değişkeninde modelde yer almasını sağlayan bir etken olmuştur.

12- Yöneticinin Proje Üretimi, işletmelerin gelecek için planlarının olup olmadığını ölçmek için modelde yer alan bir değişkendir.

²⁰⁰ Tamer Koçel, **İşletme Yöneticiliği**, İstanbul: Has Matbaası, 2005, s. 595-596.

13- Yöneticinin İşletmeden Ayrılma Sebebi, yöneticinin işletmeden ayrılmasında etkili olan faktör, yada faktörlerin ortaya çıkarılması düşünülmüştür.

4.4 Parametrelerin Tahmini

Poisson regresyon modelinde gözlemlenemeyen heterojenlik ile karşılaşılması durumunda gözlemlenemeyen heterojenliğin kontrol altında tutulması için sabit etki ve tesadüfi etki yöntemlerinden birinin kullanılması gerekliliği ortaya çıkmaktadır. Sabit etki yönteminde zamana göre sabit değişkenlerin modelde yer alması bu değişkenlerin etkilerinin kontrol edilmesini zorlaştırabileceğinden tesadüfi etki yönteminin kullanılması bu uygulamanın tercih sebebini oluşturmuştur.

Poisson regresyon modeli gözlemlerin birbirleriyle bağımsız olduğunu varsaymaktadır. Buna rağmen poisson regresyon modelindeki başlangıç nokta, bir bireyle ortaya konmuş olan bir olay olabilir. Bu gibi olaylar bireylerin birbirleriyle korelasyonlu olması yüzünden dışlanabilir. Bir başka ifadeyle anlatılmak istenirse, bir olay meydana geldiğinde ondan sonraki olayın meydana gelme olasılığı artabilir. Olaylar arasındaki bu korelasyon ekonomide durum bağılılığı (state dependence), biyometrik ve sosyolojik literatürde geçiş modelleri (contagion model) olarak isimlendirilir. Durum bağılılığının sebebi verilerin kümelenmiş olmasından kaynaklanabilir. Bu durumun sonucunda modelde aşırı yayılımla karşılaşılmasına sebep olabilir²⁰¹.

Poisson regresyon modelinin kısıtlayıcı bir varsayımı varyansla ortalamanın birbirine eşit olduğu varsayımdır. Bu varsayım; her bir tahmin oranları etrafındaki artıkların varyansının tahmin oranlarına eşit olduğu şeklinde ifade edilebilir. Bu varsayımın sağlanabilmesi için, bireyler arasındaki ortalama oranlardaki sistematik değişimin tümünün, tahminci setiyle açıklanması gerekir. Diğer potansiyel

²⁰¹ Jacob Cohen, **Applied Multiple Regression–Correlation Analysis for The Behavioral Sciences**, Mahwah: Lawrence Erlbaum Associates, incorporate, 2003, s. 528-532.

tahmincilerin hiçbiri toplamsal varyansı açıklamayabilir. Oranlarda sistematik deęişim tahminci setiyle açıklanamazsa; her bir tahminci oranı etrafındaki artıklar için, poisson regresyon modelinin izin verdiğiinden daha büyük bir deęişimin olduęu söylenebilir. Bu durum aşırı yayılım olarak ifade edilebilir²⁰².

Çanakkale Sanayici ve İşadamları odasında kayıtlı tüm işletmeler çalışmamızın anakütlesini oluşturması gerekirken, odaya kayıtlı tüm işletmeler üst düzey yönetici çalıştırmadıkları için sadece üst düzey yönetici çalıştıran işletmeler çalışmanın anakütlesini oluşturmuştur. Üst düzey yönetici çalıştırmayan, geleneksel olarak yaşamını devam ettiren işletmeler anakütleyle dahil edilmemiştir.

Çanakkale merkezi ve tüm ilçelerindeki üst düzey yönetici çalıştıran tüm işletmelerin sayısı 200 olarak tespit edilmiştir. Bunlardan ancak 120 tanesine ulaşılabilmiş ve randevu alınarak görüşme sağlanmıştır. Çalışmada yapılan tüm analizler, SAS 8.0 paket programı kullanılarak yapılmıştır. Poisson regresyon modellemesi yapılırken SAS paket programında bulunan Genmod prosedürü aracılığıyla, modele uygun bir makro hazırlanmıştır. Yönetim teorisi çerçevesinde belirlenen deęişkenlerin tümü modele katılarak analize başlanmıştır. SAS 8.0 paket programının Genmod prosedüründe hazırlanan makro, şekil 6’da görülebilir.

```
proc genmod data=d;
    model yoneticici= buyukluk syas age top danis
pay rekabet bol isbirlięi yaklasim egit proje
sebep / dist=poisson
    link = log dscale
    type3 wald;
```

Şekil 6. Genmod Prosedüründe Tüm Deęişkenler İçin Poisson Regresyon Makrosu

Tablo1’de belirtilen 13 açıklayıcı deęişken şekil 6’da verilen makro kullanılarak poisson regresyon modeli ile analiz edilmiştir. Tablo 2’de poisson

²⁰² Peter Berkhout ve Eric Plug, A Bivariate Poisson Count Data Model Using Conditional Probabilities, *Statistica Neerlandica*, No: 58, S: 3, 2004, s. 349-350.

regresyon modelinin parametre tahminleri ve parametre tahminlerinin anlamlı olup olmadığının belirlenmesi için yapılan wald testi sonuçları görülmektedir.

Tablo 2

Tüm Değişkenler İçin Poisson Regresyon Modelinin Parametre Tahminleri ve Wald Testi Sonuçları

Değişkenler	Tahmin	Standart Hata	Wald testi Alt Güven Sınırları	Wald testi Üst Güven Sınırları	Ki-Kare	Pr> χ^2
X_1	0.0196	0.0072	0.0054	0.0337	7.35	0.0067
X_2	0.0032	0.0091	-0.0146	0.0211	0.13	0.7216
X_3	-0.0577	0.0168	-0.0906	-0.0248	11.78	0.0006
X_4	0.7835	0.2422	0.3088	1.2583	10.46	0.0012
X_5	-0.5301	0.2531	-1.0262	-0.0340	4.39	0.0362
X_6	-0.0067	0.0682	-0.1403	0.1270	0.01	0.9219
X_7	-0.1826	0.2022	-0.5790	0.2138	0.82	0.3666
X_8	-0.0067	0.0682	-0.1403	0.1270	0.01	0.9219
X_9	-0.0067	0.0435	-0.1013	0.0693	0.14	0.7130
X_{10}	-0.1112	0.1034	-0.3140	0.0915	1.16	0.2821
X_{11}	0.0006	0.1517	-0.2967	0.2979	0.00	0.9967
X_{12}	0.2462	0.1086	0.0334	0.4590	5.14	0.0234
X_{13}	0.0747	0.0841	-0.2394	0.0901	0.79	0.3744

Wald testi sonuçlarına göre; X_1 , X_3 , X_4 , X_5 ve X_{12} değişkenleri %95 güvenle istatistiksel olarak anlamlı bulunmuştur. Anlamlı bulunan beş değişken şekil 7’de gösterilen makro ile tekrar analize edilmiştir.

```
proc genmod data=d;  
model yoneticici= buyukluk age top danis proje  
      / dist=poisson  
        link    = log    dscale  
        type3 wald;
```

Şekil 7. Genmod Prosedüründe Beş Değişken İçin Poisson Regresyon Makrosu

Poisson regresyon modelinde istatistiksel olarak anlamlı bulunan beş değişken tekrar analiz edildiklerinde %95 güvenle istatistiksel olarak anlamlı buldukları görülmüştür. Tablo 3’de parametre tahminlerinin anlamlılığının sınanmasında kullanılan wald testi sonuçları görülmektedir. Tablo 3 ve tablo 2 birlikte incelendiğinde, tablo 2’de anlamlı bulunan beş değişkenin tahmini değerleri ile tablo 3’de elde edilen tahmin değerleriyle birbirine yakın sonuçlar verdiği görülmektedir. Örnek olarak X_1 değişkeni gözönüne alındığında, tablo 2’deki tahmin değeri “0.0196” olarak bulunmuştur. Tekrar analiz edildiğinde elde edilen tahmin değeri ise “0.0122” olarak ilk analiz sonucuna çok yakın bir değer olduğu görülebilir.

Tablo 3

Beş Değişken için Poisson Regresyon Modelinin Parametre Tahminleri ve Wald Testi Sonuçları

	Tahmin	Standart Hata	Wald testi %95 Alt Güven Sınırları	Wald testi %95 Üst Güven Sınırları	Ki-Kare	Pr> χ^2
X_1	0.0122	0.0053	0.0019	0.0225	5.43	0.0198
X_3	-0.0564	0.0097	-0.0754	-0.0374	33.75	<.0001
X_4	0.7147	0.2083	0.3064	1.1230	11.77	0.0006
X_5	-0.5650	0.1861	-0.9299	-0.2002	9.21	0.0024
X_{12}	0.2269	0.0831	0.0640	1.3875	7.46	0.0063

Tablo3' den yararlanılarak poisson regresyon modeli,

$$\log \lambda_i(t) = 0.0122 X_1 - 0.0564 X_3 + 0.7147 X_4 - 0.5650 X_5 + 0.2269 X_{12}$$

olarak yazılabilir.

Regresyon modelinin uyum iyiliğinin sınanması için pearson χ^2 test istatistiği sonuçları değerlendirilmiştir. Verilerin model için uygun veriler olmadığına ve aşırı yayılım söz konusu olduğuna %95 güvenle karar verilebilir ($\chi^2 /$ serbestlik derecesi = 2.3527 >1).

Analiz sonuçlarında aşırı yayılım olduğu görülmektedir. Aşırı yayılım, gözlemlenemeyen bir heterojenlik olduğunu göstermektedir. Bu durumda, araştırmada kullanılan veriler için poisson regresyon modelinin uygun bir model olmadığı görülmektedir.

Aşırı yayılım olduğunda, karma bir dağılım olan poisson-gamma regresyon modeli oluşturma gerekliliği ortaya çıkmıştır. Poisson-gamma regresyon modeli için SAS 8.0 programının genmod prosedürü kullanılarak yeni bir makro hazırlanmıştır. Belirlenen beş değişken şekil 8’de gösterilen makro ile yeni bir analize tabi tutulmuştur.

```
proc genmod data=d;  
  
model yonetici = buyukluk age top danis proje  
          / dist=nb  
          link = log dscale  
          type3 wald;
```

Şekil 8. Genmod Prosedüründe Beş Değişken İçin Poisson-Gamma Regresyon Makrosu

Şekil 8’de görülen makro kullanılarak yapılan analiz sonucunda, poisson-gamma regresyon modeli için parametre tahminleri elde edilmiştir. Parametre tahminleri ve parametre tahminleri için yapılan wald testi analizi sonuçları tablo 4’de görülmektedir.

Tablo 4

Poisson-Gamma Regresyon Modeli İçin Parametre Tahminleri ve Wald Testi Sonuçları

	Tahmin	Standart Hata	Wald testi %95 Alt Güven Sınırları	Wald testi %95 Üst Güven Sınırları	Ki-Kare	$Pr>\chi^2$
X_1	0.0162	0.0060	0.0044	0.0279	7.31	0.0068
X_3	0.0570	0.0120	-0.0806	-0.0335	22.50	0.0001
X_4	0.6846	0.2185	0.2563	1.1129	9.81	0.0017
X_5	-0.5299	0.1960	-0.9141	-0.1457	7.31	0.0069
X_{12}	0.2179	0.0865	0.0484	0.3873	6.35	0.0117
ε_i	1.3875	0.0000	1.3875	1.3875		

Poisson-gamma karma regresyon modeli için yapılan wald testi sonuçlarına göre, analize giren X_1 , X_3 , X_4 , X_5 ve X_{12} değişkenleri %95 güvenle istatistiksel olarak anlamlı bulunmuşlardır. Tablo4'den yararlanılarak modelin,

$$\log \lambda_i(t) = 0.0162 X_1 + 0.0570 X_3 + 0.6846 X_4 - 0.5299 X_5 + 0.2179 X_{12} + 1.3875$$

olarak ifade edilmesi mümkündür.

Analizin son aşamasında, poisson-gamma regresyon modelinde karşılaşılan gözlemlenemeyen heterojenliği kontrol edebilmek amacıyla tesadüfi etkiler poisson-

gamma regresyon modeli uygulanması sonucuna varılmıştır. Tesadüfi etkiler poisson-gamma regresyon modelinin uygulanabilmesi için, SAS paket programının nlmixed prosedüründe yeni bir makro hazırlanmıştır. Poisson-gamma regresyon modelinde anlamlı bulunan beş değişken şekil 9’gösterilen SAS makrosu kullanılarak tesadüfi etkiler poisson-gamma regresyon modeli ile analiz edilmiştir.

```
proc nlmixed data=d;
log_mu=b1*buyukluk+b2*age+b3*top+b4*danis+b5*proje+u;
model yoneticici~negbin(1,p);
random u~normal(0,exp(2*log_sigma)) subject=id;
predict log_mu out=log_mu;
```

Şekil 9. Tesadüfi Etkiler Poisson-Gamma Regresyon Makrosu

Tablo 5’de tesadüfi etkiler poisson-gamma regresyon modeli için parametre tahminleri ve tahminlerin anlamlılığının sınanmasında kullanılan t testi sonuçları yer almaktadır.

Tablo 5

Tesadüfi Etkiler Poisson-Gamma Regresyon Modeli İçin Parametre Tahminleri ve t Testi Sonuçları

Değişkenler	Tahmin	Standart Hata	P
X_1	0.01465	0.01009	0.5942
X_3	-0.05924	0.02049	0.0080
X_4	0.6612	0.3559	0.1596
X_5	-0.4710	0.3368	0.0755
X_{12}	0.3397	0.1539	0.0371
$\log \rho(t, \phi)$	-8.5710	1011.21	0.9933
$\log v_i$	1.5562	0.6967	0.0351

Tablo 5 incelendiğinde, X_3 ve X_{12} değişkenleri modelde %95 güvenle anlamlı bulunurken diğer değişkenler model için anlamlı değildir. Poisson-gamma regresyonda anlamlı bulunan değişkenlerden sadece X_3 ve X_{12} değişkenleri tesadüfi etki analizi için anlamlı bulunmuştur. Anlamlı bulunan değişkenlerin tahmini değerleride farklılaşmıştır. X_3 değişkeni poisson gamma regresyon modelinde pozitif bir etkiye sahipken tesadüfi etki analizinde tahmini değeri negatif olmuştur. X_{12} değişkenin tahmini ise poisson-gamma regresyom modelinde tahmin edilen değere göre biraz daha artış göstermiştir.

$\log \rho(t, \phi)$ negatif bir değer olarak bulunması $\rho(t, \phi)$ değerinin 0-1 arasında bir sayı olduğunu ifade etmektedir. Tahmin edilen $\log(-8.5710)$ sayısı sifıra çok yakın bir sayıdır. Model için anlamlılığı sınıandığında anlamlı bulunmamıştır.

Tablo 5’de görüldüğü gibi gözlemlenemeyen heterojenliğin modele katkısı ise 1.5562 birim olarak ifade edilebilir.

SONUÇ

Herhangi bir arařtırmada toplanan veriler kesit verisi ve uzun kesit verisi olarak toplanabilirler. Uzun kesit verileri kendi içlerinde sınıflandırmaya tabi tutulduklarında en kapsamlı olanı olay zamanı verileri olmaktadır. Olay sayısını, onların ardışıklıklarını, meydana gelme zamanlarını ve sürelerini kayıt etmelerinden dolayı olay zamanı verileri, diđer veri tiplerine göre daha avantajlı olmaktadır.

Olay zamanı verilerinin modellenmesinde bağımlı deęişkenin aldığı deęerlere göre farklı modellemelerin yapılması mümkündür. İncelenen bağımlı deęişken herhangi bir olayın ortaya çıkma uzunluęu olarak tanımlandığında bu tür bir olayın modellenmesi poisson regresyon modellemesi ile yapılabilmektedir. Poisson regresyon, ortalama ve varyansın birbirine eřit olduęu varsayımının üzerine yapılandırılmıştır. Bu varsayımını sağlanamaması durumunda ortalama varyanstan büyükse aşırı yayılımla karşılaşılmaktadır. Aşırı yayılım karşısında karma bir daęılım kullanılması uygun görülmektedir.

Bu çalışma çok uzun bir süreç sonucunda teorik ve uygulama alanında yaşanan eksikliklere yönelik olarak, konuların daha etkili bir şekilde anlaşılması için bir model geliştirme çabası içerisinde oluşturulmuştur. Bu çerçeve içerisinde çalışmada öncelikle konu ile ilgili modeller tartışılmış ve daha sonra yönetim alanında konuya uygun deęişkenler belirlenerek modelleme yapılmaya çalışılmıştır.

Yöneticilik alanında çok çeşitli arařtırmalar yapılmış olmasına rağmen yöneticilerin organizasyon ile olan bağlarının devamlılıęı ile ilgili olarak olay zamanı veri yapısına uygun çalışma yapılmaması dikkat çeken bir unsur olmuştur. Yönetim bilimi çerçevesinde belirlenen on üç deęişken poisson regresyon ile modellendiklerinde veri yapısının modele uygun olmadığı ve aşırı yayılım olduęu görülmüştür. Bunun sebebi ise, yöneticilerin organizasyonlar içinde gruplandırılmasından kaynaklanmış olduęu söylenebilir.

Verilerin poisson regresyon modelinde aşırı yayılım göstermesi sonucunda karma bir dağılım olan poisson-gamma regresyon modeli ile analize devam edilmiştir. Poisson-gamma regresyon modeli ile yapılan analiz sonucunda gözlemlenemeyen heterojenlikle karşılaşılması bu konunun tesadüfi etkiler veya sabit etkiler analizi ile kontrol altında tutulması gerekliliğini gündeme getirmiştir. Sabit etki analizleri, zamana göre sabit değişkenleri modelde buldurmasında bir takım sorunlar ortaya çıkarmaktadır. Bu durumdan kaynaklı olarak tesadüfi etkiler poisson-gamma regresyon modeli uygun bir model olarak görülmüştür.

Araştırmanın modellenmesinde “Tesadüfi Etkiler Poisson-Gamma Regresyon” modeli kullanılarak işletmeler ve yöneticiler arasındaki ilişkiler incelenmiştir. Ortaya çıkan sonuçların işletmeler açısından ilginç olduğu düşünülmektedir. Çünkü yöneticiler organizasyonlar için büyük bir öneme sahiptirler. Onların kurumlardaki varlığı ve kararları birçok açıdan sistemi önemli ölçüde etkilemektedir. Kurumlardan ayrılmalarının da örgütler açısından büyük öneminin olduğunu söyleyebiliriz. Burada belirlediğimiz değişkenlerden bazılarının anlamlı çıkmasının organizasyonlar ve yöneticiler açısından önemli olduğu düşünülmektedir.

Poisson-gamma regresyon modelini oluşturan değişkenlere bakıldığında hem organizasyonla ilgili hemde yönetici ile ilgili değişkenlerin anlamlı olduğu görülmektedir. Organizasyonun büyüklüğü ve organizasyonun danışmanlık hizmeti alıp-almama değişkeni organizasyonla ilgili değişkenlerdir. Modelde anlamlı çıkan yönetici yaşı ve yöneticinin üst düzey olarak toplantı yapıp yapmaması değişkenleri ise yönetici ile ilgili değişkenler olmaktadır. Tesadüfi etkiler poisson-gamma regresyonda ise organizasyonla ilgili olan bu iki değişken model için anlamlı olmamaktadır. Bu durum tesadüfi etkiler poisson-gamma regresyon modelinde gözlemlenemeyen heterojenliğin yüksek bir değer olarak ortaya çıkmasıyla da desteklenmektedir. Her iki modelinde sonuçları karşılaştırıldığında tesadüfi etkiler poisson-gamma regresyon modelindeki gözlemlenemeyen heterojenliğin modele katkısı, poisson-gamma regresyon modelindeki hata payından daha yüksek çıktığı görülmektedir.

Ortaya çıkan sonuçlar, araştırma yapılan şirketlerin kurumsal yapıda oluşmaması sonucunun bir göstergesi olarak ele alınabilir. Özellikle Türkiye’de ve

bizim arařtırmamızda yer alan rneklerde grldę gibi kk ve orta lekli sayılabilecek řirketler genel olarak aile bireyleri tarafından ynlendirilmektedirler. Btn karar srelerinde aile bireyleri rol almaktadır. Bu yzden organizasyonlar ve profesyonel ynetici kadroları arasındaki kalıcı ve uzun sreli alıřmaların devam etmesinde sıkıntılar ortaya ıkabilmektedir.

Tesadfi etkiler poisson-gamma regresyon modeli uygulanması sonucunda, yneticiler ve rgtler arası iliřkinin incelenmesinde iki nemli deęiřken karřımıza ıkmıřtır. Bunlar yneticilerin proje retkenlięi ve yneticilerin yařlarıdır. Gerektede iř dnyası aısından ayakta kalabilmek ve rekabet edebilmek iin en nemli faktrlerden bir tanesi kurumların yeniliki ve yaratıcı olarak ortaya koydukları proje sayılarıdır. Kurumlar ancak bunlarla piyasadaki rekabet glerini srdrlebilir kılmaktadırlar. Bu arařtırmanın bulgularının da bu durumu desteklemesinin nemli olduęu dřnlmektedir. Yneticiler aısından ise proje retmenin kariyerlerinde ne kadar nemli olduęunu kavramaları byk nem tařımaktadır. Ayrıca yneticilerin yařı da nemli bir faktr olarak karřımıza ıkmaktadır. Gen yařlarda mesleęe yeni girmenin getirmiř olduęu motivasyon ile iřlerine karřı daha aktif bir davranıř ve tutum ierisinde olmalarının bir sonucu olarak, nemli bir faktr olduęunu ifade etmek mmkndr. Yneticinin iře alınmasında ve deęerlendirilmesinde bu faktrn de dikkate alınması zellikle rgt ve ynetici arasındaki iliřkilerin srdrlebilirlięi aısından byk bir nem tařımaktadır.

Organizasyonun ynetim kurulunda yer alan kiřilerin tutum ve davranıřları ile yneticinin tutum ve davranıřları bu arařtırma kapsamında llemedięi iin gzlemlenemeyen heterojenlięin yksek ıkmasında etkili bir unsur olduęu dřnlebilir. Tutum ve davranıřla ilgili iř tatmin lekleri bulunmakla beraber organizasyondan ayrılan yneticilere ulařmadaki sorun, bu deęiřkenlerin incelenmesini olanaksız kılmıřtır. Bu konularda yaratılacak veriler bundan sonraki arařtırmalar aısından nemli olacaktır.

EKLER

EK 1: SORU FORMU

Bu veri formuna katkı verdiğinizden dolayı öncelikle size teşekkür ederiz. Bu araştırmanın amacı işletmelerde yöneticilerin çalışma süreleri ve şirketten ayrılma süreleri ile ilgili faktörleri ortaya çıkarmaktır.

Vereceğiniz cevaplar araştırmanın analizi ve bilimsel güvenilirliği açısından büyük bir öneme sahiptir. Bu nedenle mümkün olduğunca bütün soruların cevaplarının tam ve eksiksiz olarak doldurulması büyük bir önem taşımaktadır. Burada ortaya konan bilgiler anonim olarak değerlendirilecektir. Sizin ve şirketiniz ile ilgili olarak özel bir değerlendirme yapılmayacağını bilgilerinize sunmak isterim.

Temel Sorular

Lütfen veri toplama formumuzun diğer bölümüne geçmeden önce aşağıdaki temel demografik bilgileri doldurunuz.

1-Şirkette kaç yıldan beri görev yapmaktasınız?-----

2-Şirketteki unvanınız nedir?-----

3-Cinsiyetiniz? 1 -Erkek 2-Bayan

4-Kaç yaşındasınız?-----

5- Eğitim düzeyiniz? 1-İlköğretim 2-Lise 3-Lisans 4-Y.Lisans -Doktora

ŞİRKETİNİZ İLE İLGİLİ TEMEL SORULAR

1-Şirketin kuruluş tarihi? -----/-----/-----

2-Şirketin faaliyette bulunduğu sektör ?-----

3-Şirketinizde toplam çalışan sayısı?-----

4-Şirketinizde çalışan yönetici sayısı?-----

5-Şirketinizde düzenli olarak (aylık) hangi toplantılar yapılmaktadır?

- a-Tepe yönetim toplantısı
- b-Pazarlama toplantısı
- c-Üretim toplantısı
- d-Finansal değerlendirme toplantısı
- e-Stratejik insan kaynakları toplantısı

6-Şirkette yeniden yapılanma çalışmaları var mı? a-Evet b-Hayır

A-Cevabınız evet ise danışmanlık hizmeti aldınız mı ? a-Evet b-Hayır

B-Aldıysanız yaklaşık olarak ne kadar süreden beri almaktasınız?

-----/-----/-----

7-Şirketinizin bulunduğu bölgede faaliyette bulunduğu sektör içerisindeki payı?

9- Şirketinizin son on yılda sektörde rekabet gücünde yaşanan eğilim?

-95	96	97	98	99	2000	01	02	03	04	05
arttı	arttı	arttı	arttı	arttı	arttı	arttı	arttı	arttı	arttı	arttı
azaldı	azaldı	azaldı	azaldı	azaldı	azaldı	azaldı	azaldı	azaldı	azaldı	azaldı
	sabit	sabit	sabit	sabit	sabit	sabit	sabit	sabit	sabit	sabit

10-Şirketiniz kaç farklı bölgede faaliyette bulunmaktadır? -----

11-Şirketiniz şuana kadar yurt içinden ve yurt dışından şirketler yaptığı işbirliği sayısı ve alanları nedir

ŞİRKETİNİZDE ÇALIŞAN YÖNETİCİLER İLE İLGİLİ BİLGİLER

Aşağıda işletmenizde çalışan yöneticiler ile ilgili temel sorular söz konusudur. Lütfen şirketinize uygun bir şekilde cevaplandırınız.

1-Şirketinizin kuruluşundan bu yana görev alan genel müdürlerin işe başlama ve ayrılış tarihleri

	İŞE BAŞLAMA TARİHİ	AYRILIŞ TARİHİ
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		

	YÖNETİCİLER					
	1	2	3	4	5	6
1-Yönetici döneminde çalışan eleman sayısı:						
2-Şirketinizde Yönetici döneminde düzenli olarak hangi toplantılar yapılmaktadır?						
3-Şirkette yeniden yapılanma için Yönetici danışmanlık hizmeti aldınız mı?						
4-ŞirketinizYönetici döneminde kaç farklı bölgede faaliyette bulunmaktadır?						
5-ŞirketinizinYönetici döneminde yurt içinden ve yurt dışından şirketler ile yaptığı iş birliği sayısı						
6- Şirketinizin Yöneticisi şirketinizde göreve başlamadan önce farklı şirketlerde kaç yıl yöneticilik yaptı						
7- ŞirketinizinYöneticisinin yönetim yaklaşımı nasıldı?	a-istismarcı-otokratik b-yardımsever otokratik c-katılımcı d-demokratik	a-istismarcı-otokratik b-yardımsever otokratik c-katılımcı d-demokratik	a-istismarcı-otokratik b-yardımsever otokratik c-katılımcı d-demokratik	a-istismarcı-otokratik b-yardımsever otokratik c-katılımcı d-demokratik	a-istismarcı-otokratik b-yardımsever otokratik c-katılımcı d-demokratik	a-istismarcı-otokratik b-yardımsever otokratik c-katılımcı d-demokratik
8-ŞirketinizinYöneticisinin İşe başlama yaşı nedir						
9- ŞirketinizinYöneticisinin eğitim durumu nedir?						
10- ŞirketinizinYöneticisinin çalıştığı süre içerisinde ürettiği projeler						
11- ŞirketinizinYöneticisinin şirketinizden ayrılma sebebi nedir						

KAYNAKÇA

Kitaplar:

- Agresti, Alan. **Categorical Data Analysis**, New York: Wiley,1990.
- Agresti, Alan. **Categorical Data Analysis**, New York: Wiley, 2002.
- Allison, D. Paul. **Event History Analysis Regression for Longitudinal Event data**, California: Sage Publication, 1984.
- Alpar, Reha. **Uygulamalı Çok deęişkenli İstatistiksel Yöntemlere Giriş 1**, Ankara: Nobel Yayın dağıtım, 2003.
- Bary, Arnold. **Conditional Specification of Statistical Models**, New York: Springer-Verlag, 1999.
- Blasius, Jörg ve Michael Geenacre. **Visualization of Categorical Data**, London: Academic Press, 1998.
- Blossfeld, Hans-Peter ve Götz Rohwer. **Techniques of Event History Modelling: New Approaches to Causal Analysis**, Mahwah: NJ Lawrence Erlbaum Associates, Inc, 2002.
- Burnham, Kenneth. **Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach**, New York: Springer-Verlag, 2002.
- Cameron, A.Colin. **Regression Analysis of Count Data**, West Nyack, NY, USA: Cambridge University Press, 1998.
- Campbell, M .J. **Statistics at Square One**, London: BMJ Publishing Group, 2002.
- Campbell, M. J. **Statistics at Square Two: Understanding Modern Statistical Applications in Medicine**, London: BMJ Publishing Group, 2001.
- Christensen, **Ronald. Log-Linear Models&Logistic Regression**, New York: Springer-Verlag, 1997.
- Cox, D.R. ve D. Oakes. **Analysis of Survival Data**, London: Chapman and Hall, 1984.
- Dalgaard, Peter. **Introductory Statistics with R**, New York: Springer-Verlag, 2002.
- Diggle, J. Peter, Kung- Yee Liang ve Scott L. Zger. **Analysis of Longitudinal Data**, Oxford: Oxford University Press, 1994.
- Dobson, J. Annette. **An Introduction to Generalized Linear Models**, London: Chapman ve Hall/CRC, 2002.

- Dupont, D. William. **Statistical Modeling for Biomedical Researches: A Simple Introduction to the Analysis of Complex Data**, West Nyack, NY, USA: Cambridge University Press, 2002.
- Er, Fikret ve Harun Sönmez. **İstatistik Formülleri ve Olasılık Dağılımları**, Eskişehir: Anadolu Üniversitesi Fen Fakültesi Yayınları No. 21, 2001.
- Everitt, Brian. **Statistical Aspects of the Design and Analysis of Clinical Trials**, London: Imperial College Press, 2004.
- Fossum, Lynn. B. **Understanding Organizational Change Converting Theory to Practice**, USA: Crisp Publications, 1989.
- Goldstein., Harvey. **Multilevel Statistical Models**, London: Hodder Arnold Publication, 2003.
- Green, P.J. ve B.W. Silverman. **Nonparametric Regression and Generalized Linear Models**, London: Chapman ve Hall, 1995.
- Greene, H. William. **Limdep Version 7.0 User's Manual**, Australia: Econometric Software, Inc, 1995.
- Gujaratti, Damador N. **Temel Ekonometri**, Çev. Ümit Şenesen ve Gülay Günlük Şenesen, İstanbul: Literatür Yayıncılık, 1999.
- Güriş, Selahattin ve Şahamet Bülbül. **Olasılık**, İstanbul: Yaylım Matbaası, 1995.
- Gürsakal, Necmi. **Bilgisayar Uygulamalı İstatistik 1**, Bursa: Marmara Kitabevi, 1997.
- Gyorfi, Laszlo. **Distribution- Free Theory of Nonparametric Regression**, New York: Springer-Verlag, 2002.
- Haining, Robert. **Spatial Data Analysis: Theory and Practice**, West Nyack: Cambridge Universtiy Press, 2003.
- Hannan, T. Michael ve John Freeman. **Organizational Ecology**, Cambridge: Harward University Press, 1989.
- Harell, E. Frank. **Regression Modeling Strategies with Applications to Linear Models, Logistic Regression and Survival Analysis**, New York: Springer-Verlag, 2001.
- Hsiao, Cheng. **Analysis of Panel Data**, West Nyack, NY, USA: Cambridge University Press, 2002.
- Kalbfleisch, J.D. ve R.L. Prentice. **The Statistical of Failure Time Data**, New York: John Willey and sons, 1980.
- Koçel, Tamer. **İşletme Yöneticiliği**, İstanbul: Has Matbaası, 2005.

- Koutsoyiannis, A. **Ekonometri Kuramı**, Çev. Ümit Şenesen ve Gülay Günlük Şenesen, Ankara: Verso Yayıncılık, 1989.
- Le, T . Chap. **Applied Survival Analysis**, New York: John Willey ve Sons, 1997.
- Lindsey, J.K.. **Modelling Frequency and Count Data**, Oxford: Clarendon Press, 1997.
- Lloyd, Chris J. **Statistical Analysis of Categorical Data**, New York: John Wiley ve Sons, 1999.
- Loader, Clive. **Local Regression & Likelihood**, New York: Springer- Verlag, 1999.
- Maller, A.Ross ve Zhou Xian. **Survival Analysis with Long-Term Survivors-Wiley Series in Probability and Statistics**, New York: John Wiley ve Sons, 1996.
- McCutcheon, Allan. **Advances in Latent Class Models**, West Nyack, NY, USA: Cambridge University Press, 2002.
- Moeschberger, Melvin. **Survival Analysis**, New York: Springer-Verlag, 1997.
- Nelson, Wayne. **Applied Life Data Analysis**, New York: Wiley, 1981.
- Newbold, Paul. **İşletme ve İktisat İçin İstatistik**, Çev. Ümit Şenesen, İstanbul: Literatür Yayıncılık, 2000.
- Powers, A.Daniel ve Xie Yu. **Statistical Methods for Categorical Data Analysis**, San Diego: Academic Press, 2000.
- Rao, C. Radhakrisna. **Linear Models: Least Squares and Alternatives**, New York: Springer-Verlag, 1999.
- Rose, David ve Oriel Sullian. **Introducing Data Analysis for Social Scientists**, Buckingham, Open Uni. Press, 1998.
- Sekaran, Uma. **Research Methods for Business a Skill-Building Approach**, New York: John Wiley ve Sons, 2000.
- Serper, Özer. **Uygulamalı İstatistik 1**, Bursa: Ezgi Kitabevi, 2000.
- Shoukri, M. Mohamed ve C.A. Pause. **Statistical Methods For Healt Sciences**, New York: CRC Press LCC, 1999.
- Sprott, Duncan. **Statistical Inference in Science**, New York: Springer Verlag, 2000.
- Şenocak, Mustafa. **Biyoistatistik**, İstanbul: Dilek Ofset Matbaacılık,1998.
- Tableman, Mara ve Jong Sung Kim. **Survival Analysis Using S Analysis of Time-to-Event Data**, New York: Chapman and Hall, 2004.

- Tatlídil, Hüseyin. **Uygulamalı Çok Değişkenli İstatistiksel Analiz**, Ankara: Cem Web Ofset Ltd.Şti., 1996.
- Trussell, James, Richard Hankinson ve Judith Tilton, **Demographic Applications of Event History Analysis**, Oxford: Clarendon Press, 1992.
- Toutenburg, Helge. **Statistical Analysis of Designed Experiments**, New York: Springer –Verlag, 2002.
- Twisky, W. R. Jos. **Applied Longitudinal Data Analysis for Epidemiology: A Practical Guide**, West Nyack, NY, USA: Cambridge University Press, 2003.
- Verbeke, Geert. **Linear Mixed Models for Longitudinal Data**, New York: Springer-Verlag, 2000.
- Vermunt, K. Jeroen. **Log Linear Models for Event Histories**, California: Sage Publication, 1997.
- Yamaguchi, Kazuo. **Event History Analysis**, California: Sage Publication, 1991.
- Zelterman, Daniel. **Models for Discrete Data**, Oxford: Clarendon Press, 1999.

Sürekli Yayınlar:

- Abrevaya, Jason. “*Rank Estimation of a Generalized Fixed- Effects Regression Model*”, **Journal of Econometrics**, No: 95, 2000, s. 1-23.
- Agresti, Alan, G. James P. Hobert ve Brian Caffo. “*Random-Effects Modelling of Categorical Response Data*”, **Sociological Methodology**, No: 30, S: 1, 2000, s. 27-80.
- Amburgey, T. L. ve H. Rao. “*Organizational Ecology: Past, Present and Future Direction*”, **Academy of Management Journal**, No: 39, S: 5, 1996, s. 1265-86.
- Andersen, Perprag, John P. Klein ve Mei-Jie Zhang. “*Testing for Centers Effects in Multicenter Survival Studies: A Monte Carlo Comparison of Fixed and Random Effects Tests*”, **Statistics and Medicine**, No: 18, S: 12, Ocak 1999, s. 1489-1500.
- Apker, J. “*Sensemaking Of Change In The Managed Care Era: A Case Of Hospital-Based Nurses*”, **Journal Of Organizational Change Management**, No: 2, S: 17, 2004, s. 211-227.

- Bae, S., F. Famoye, J.T. Wulu, A. A. Bartolucci ve K.P. Singh. “*A Rich Family of Generalized Poisson Regression Models with Applications*”, **Mathematics and Computers in Simulation**, No: 69, 2005, s. 4-11.
- Barlett, A. C. ve L. S. Ghosha. “*Changing The Role Of Top Management: Beyond Systems To People*”, **Harvard Business Review**, Mayıs-Haziran 1995, s. 132-142.
- Baker, E. Wayne, Robert R. Faulkner ve Gene A. Fisher. “*Hazards of The Market: The Continuity and Dissolution of Interorganizational Market Relationships*”, **American Sociological Association Review**, No: 63, S: 2, Nisan 1998, s. 147-177.
- Baum, J. A. C. ve C. Oliver, “*Institutional Linkages and Organizational Mortality*”, **Administrative Science Quarterly**, No: 36, 1991, s. 187-218.
- Beise, Jan ve Voland Eckart. “*A Multilevel event History Analysis of the Effects of Grandmothers on Child Mortality in A Historical German Population (Krummhoörn, Ostfriesland, 1720-1874)*”, **Demographic Research**, No: 7, 2002, s. 470-494.
- Bennis, B. “*Why Leaders Can't Lead*”, **The Manager's Bookshelf: A Mosaic of Contemporary Views**, 1994, s. 229-236.
- Berkhout, Peter ve Eric Plug. “*A Bivariate Poisson Count Data Model Using Conditional Probabilities*”, **Statistica Neerlandica**, No: 58, S: 3, 2004, s. 349-364.
- Blofeld, Hans-Peter ve Alfred Hamerle. “*Unobserved Heterogeneity in Hazard Rate Models: A Test and an Illustration from A Study of Career Mobility*”, **Event History Analysis in Life Course Research**, 1990, s. 241-252.
- Both, G. James, George Casella, Herwig Friedl ve James P. Hobert. “*Negative Binomial Loglinear Mixed Models*”, **Statistical Modelling**, S: 3, 2003, s. 179-191.
- Box-Steffensmeier, M. Janet ve Christopher Zorn. “*Duration for Repeated Events*”, **The Journal of Politics**, No: 64, S: 4, Kasım 2002, s. 1069-1094.
- Bröcheler, Vera, Steven Maijoor ve Arjen Van Witteloostuijin. “*Auditor Human Capital and Audit Firm Survival The Dutch Audit Industry in 1930-1992*”, **Accounting, Organizations and Society**, No: 29, 2004, s. 627-646.
- Chai, Tianxi ve Rebecca A. Betensky. “*Hazard Regression for Interval-Censored Data with Penalized Spline*”, **Biometrics**. No: 59, Eylül 2003, s. 570-579.
- Chakraborty, S. K. ve D. Chakraborty. “*The Transformed Leader And Spiritual Psychology: A Few Insights*”, **Journal Of Organizational Change Management**, No: 2, S: 17, 2004, s. 194-210.

- Charnigo, Richard ve Jiang Sun. “Testing Homogeneity in a Mixture Distribution Via The L^2 Distance Between Competing Models”, **Journal of The American Statistical Association**. No: 99, S: 466, Haziran 2004, s. 488-498.
- Chizmar, F. John. “A Discrete-Time Hazard Analysis of Role of Gender in Persistence in Economics Major”, **Journal of Economic Education**, No: 2, 2000, s. 107-118.
- Choi, Yunhee, Hongshik Ahn ve James J. Chen. “Regression Trees for Analysis of Count Data with Extra Poisson Variation”, **Computational Statistics and Data Analysis**, No: 3, S: 49, Haziran 2005, s. 893-915.
- Crouchley, R. ve R. B. Davies. “A Comparison of Population Average and Random-Effect Models for The Analysis of Longitudinal Count Data with Base-Line Information”, **J.R. Statistical Society A.**, No: 162, 1999, s. 331-347.
- Cogner, A. J. ve M. R. Fulmer. “Developing Your Leadership Pipeline”, **Harvard Business Review**, Aralık 2003, s. 76-84.
- Dean, C.B. ve R. Balshaw. “Efficiency Lost By Analyzing Counts Rather Than Event Times in Poisson and Overdispersed Poisson Regression Models”, **Journal of the American Statistical Association**. No: 92, S: 440, Aralık 1997, s. 1387-1398.
- Détang-Dessendre, Cecile ve Ian Molho. “Residence Spells and Migration: A Comparison for Men and Women”, **Urban Studies**. No: 37, S: 2, Şubat 2000, s. 247-260.
- Duran, Jean-Baptiste ve Oliver Gaudoin. “Software Reliability Modelling and Prediction with Hidden Markov Chains”, **Statistical Modelling**, No: 5, 2005, s. 75-93.
- Famoye, Felix ve Weiren Wang. “Censored Generalized Poisson Regression Model”, **Computational Statistics and Data Analysis**, No: 46, 2004, s. 547-560.
- Farrington, C. P. “Residuals for Proportional Hazards Models with Interval-Censored Survival Data”, **Biometrics**, No: 56, Temmuz 2000, s. 473-482.
- Finkelstein, Dianne M., William B. Goggins ve David A. Schoenfeld. “Analysis of Failure Time Data with Dependent Interval Censoring”, **Biometrics**. No: 58, Haziran 2002, s. 298-304.
- Fisher, Lloyd D. ve D.Y.Lin. “Time Dependent Covariates in The Cox Proportional-Hazards Regression Model”, **Annual Reviews Public Health**, No: 20, 1999, s. 145-157.
- Follmann, Dean A. ve Mathew S. Goldberg. “Distinguishing Heterogeneity from Decreasing Hazard Rates”, **Tecnometrics**, No: 30, S: 4, Kasım 1988, s. 389-396.

- Galler, P. Heinz ve Poetter Ulrich. “*Unobserved Heterogeneity in Models of Unemployment Duration*”, **Event History Analysis in Life Course Research**, 1990, s. 241-252.
- Garasky, Steven, R Jean Haurin ve Donald R. Haurin. “*Group Living as Youths Transition to Adulthood*”, **Journal of Population Economics**, No: 14, 2001, s. 329-349.
- Goleman, D. “*What Makes A Leader?*”, **Harvard Business Review**, Kasım-Aralık 1998, s. 93-102.
- Gosling, J. ve H. Mintzberg. “*The Five Minds of A Leader*”, **Harvard Business Review**, Kasım 2003, s. 54-61.
- Guo, Guang. “*Event-history Analysis for Left Truncated Data*”, **Sociological Methodology**, No: 23, 1993, s. 217-243.
- Halaby, Charles N. “*Panel Models in Sociological Research: Theory into Practice*”, **Annu. Rev. Socio.**, No: 30, 2004, s. 507-544.
- Heaton, Tim. B. ve R.A. Call Vaughin. “*Modelling Family Dynamics With Event History Techniques*”, **Journal of Marriage and Family**, No: 57, Kasım 1995, s. 1078-1090.
- Heijden, Peter GM van der, Rami Bustami, Maarten JLF Cruyff, Godfried Engbersen ve Hans C. van Houwelingen. “*Point and Interval Estimatin of The Populatin Size Using The Truncated Poisson Regression Model*”, **Statistical Modelling**, No: 3, 2003, s. 305-323.
- Heinzl, Harald ve Martina Mittlböck. “*Pseudo R- Squared Measures for Poisson Regression Models with Over–or Underdispersion*”, **Computational Statistics and Data Analysis**, No: 44, 2003, s. 253-271.
- Hoem, M. Jan. “*Systematic Patterns of 2000 Exposures in Event-History Analysis*”, **Sociological Methodology**, No: 30, 2000, s. 237-259.
- Horowitz, L. Joell ve Sobae Lee. “*Semiparametric Estimation of a Panel Data Proportional Hazard Model with Fixed*”, Ocak 2002, s.1-59.
- Geletkanycz, A. M. ve C. D. HambricK. “*The External Ties Of Top Executives: Implications For Strategic Choice And Performance*”, **Administrative Science Quaterly**, No: 42, S: 4, Aralık 1997, s. 654-681.
- Kark, R. “*The Transformational Leader: Who is (s)he? A Feminist Perspective*”, **Journal of Organizational Change Management**, No: 2, S: 17, 2004, s. 160-176.
- Karlis, Dimitris. “*A General EM Approach for Maximum Likelihood Estimatin in Mixed Poisson Regression Models*”, **Statistical Modelling**, No: 1, 2001, s. 305-318.

- Kotter, J. P. “*What Leaders Really Do*”, **Breakthrough Leadership, Harvard Business Review**, Aralık 2001, s. 85-97.
- Kozumi, Hideo. “*Bayesian Analysis of Discrete Survival Data with a Hidden Markov Chain*”, **Biometrics**, No: 56, Aralık 2000, s. 1002-1006.
- Kulich, Michal ve D.Y. Lin. “*Additive Hazards Regression with Covariate Measurement Error*”, **Journal of American Statistical Association**, No: 95, S: 449, Mart 2000, s. 238-248.
- Laitila, Thomas. “*A Pseudo-R² Measure for Limited and Qualitative Dependent Variable Models*”, **Journal of Econometrics**, No: 56, S: 3, Nisan 1993, s. 341-355.
- Land, C. Kenneth, Patricia L. McCall ve Daniel S. Nagin. “*A Comparison of Poisson, Negatif Binomial, and Semiparametric Mixed Regression Models*”, **Sociological Methods and Research**, No: 24, S: 4, Mayıs 1996, s. 387-442.
- Land, C. Kenneth, Patricia L. McCall ve Daniel S. Nagin. “*How Many Latent Classes of Delinquent/ Criminal Careers? Results from Mixed Poisson Regression Analyses*”, **AJS**, No: 103, S: 6, Mayıs 1998, s. 1593-1630.
- Land, C. Kenneth, S. Daniel Nagin, ve L. Patricia McCall. “*Discrete Time Hazard Regression Models with Hidden Heterogeneity The Semiparametric Mixed Poisson Regression Approach*”, **Sociological Methods and Research**, No: 29, S: 3, Şubat 2001, s. 342-373.
- Lansiti, M. ve R. Levien. “*Strategy as Ecology* ” **Harvard Business Review**, Mart 2004, s. 68-78.
- Lee, Sik-Yum ve Xin-Yuan Song. “*Maximum Likelihood Estimation and Model Comparison for Mixtures of Structural Equation Models with Ignorable Missing Data*” **Journal of Classification**, No: 20, 2003, s. 221-255.
- Liu, Xian. “*Development of a Structural Hazard Rate Model in Sociological Research*”, **Sociological Methods and Research**, No: 29, S:1, Ağustos 2000, s. 77-117.
- Lomi, Alessandro. “*The Population Ecology of Organizational Founding: Location Dependence and Unobserved Heterogeneity*”, **Administrative Science Quarterly**, No: 40, Mart 1995, s.111-144.
- Longenecker, O. Clinton, L. Jack Simonetti ve Thomas W. Sharkey. “*Why Organizations Fail: The View from The Front-Line*”, **Management Decision**, No: 37, S: 6, 1999, s. 503-513.
- Lu, Zudi, Yer Van Hui ve Andy H. Lee. “*Minimum Hellinger Distance Estimation for Finite Mixtures of Poisson Regression Models and Its Applications*”, **Biometrics**, No: 59, 2003, s.1016-1026.

- Ma, Renjun, Daniel Krewski ve Richard T. Burnett. “*Random Effects Cox Models: A Poisson Modelling Approach*”, **Biometrika**, No: 90, S: 1, Mart 2003, s. 157-169.
- Mayer, Karl Ulrich ve Nancy Brandon Tuma. “*Life Course Research and Event History Analysis: An Overview*”, **Event History Analysis in Life Course Research**, 1990, s. 3-20.
- Miller, Arden. “*The Analysis of Unreplicated Factorial Experiments Using All Possible Comparisons*”, **Technometrics**, No: 47, S: 1, Şubat 2005, s. 51-63.
- Minkof, Debra C. “*Interorganizational Influences on the Founding of African American Organizations, 1955-1985*”, **Sociological Forum**, No: 10, S: 1, 1995, s. 51-79.
- Mittlböck, M. ve T. Waldhör. “*Adjustments for R^2 -Measures for Poisson Regression Models*”, **Computational Statistics & Data Analysis**, No: 34, 2000, s. 461-472.
- Mungin, Murat K. ve Pravin K. Trivedi. “*Simulated Maximum Likelihood Estimation of Multivariate Mixed-Poisson Regression Models, with Application*”, **Econometrics Journal**, No: 2, 1999, s. 29-48.
- Nadas, Arthur. “*On Proportional Hazard Functions*”, **Technometrics**, No: 12, S: 2, Mayıs 1970, s. 413-425
- Nelson, Wayne. “*Theory and Applications of Hazard Plotting for Censored Failure Data*”, **Technometrics**, No: 42, 2000, s. 12-25.
- Ng, E. T. M.ve R. J. Cook. “*A Comparison of some Random Effect Models for Parameter Estimation in Recurrent Events*”, **Mathematical and Computer Modelling**, No: 32, 2000, s. 11-26.
- Palmgren, Juni ve Samuli Ripatti. “*Fitting Exponential Family Mixed Models*”, **Statistical Modelling**, No: 2, 2002, s. 23-38.
- Poortema, K. “*On Modelling Overdispersion of Counts*”, **Statistica Neerlandica**, No: 53, S: 1, 1999, s. 5-20.
- Raftery, E. Adrian. “*Statistics in Sociology, 1950-2000: A Selective Review*”, **Sociological Methodology**, No: 31, S: 1, Ocak 2001, s.1-45.
- Spinell, J. John, Richard A. Lockhart ve Michael A. Stephens. “*Test for The Response Distribution in a Poisson Regression Model*”, **Journal of Statistical Planning and Inference**, No: 108, 2002, s. 137-154.
- Singh, V. Jitendra ve Robert J. House. “*Organizational Change and Organizational Mortality*”, **Administrative Science Quarterly**, No: 31, 1986, s. 587-611.

- Smith, G. K., K. A. Smith, D. J. Olian, Jr. P. H. Sims, P. D. O'bannon ve A. J. Scully. "Top Management Team Demography and Process: The Role of Social Integration and Communication", **Administrative Science Quaterly**, No: 39, Sayı: 3, Kasım 1994, s. 412-438.
- Smith, M. ve C. M. White. "Strategy, CEO Specialization, and Succession", **Administrative Science Quaterly**, No: 32, S: 2, Haziran 1987, s. 263-280.
- Sun, Liuquan ve Zhou Xian. "Survival Function and Density Estimation for Truncated Dependent Data", **Statistics and Probability Letters**, No: 52, 2001, s. 47-57.
- Sun, Jianguo. "Nonparametric Test for Doubly Interval-Censored Failure Time Data", **Lifetime Data Analysis**, No: 7, 2001, s. 363-375.
- Sutradhar, Brajendra C. ve Vandna Jowaheer. "On Familial Longitudinal Poisson Mixed Models with Gamma Random Effects", **Journal of Multivariate Analysis**, No: 87, 2003, s. 398-412.
- Tichy, M. N. ve A. M Devanna. "The Transformational Leader", **The Manager's Bookshelf: A Mosaic of Contemporary Views**, 1994, s. 244-255.
- Thomas, A. B. "Does Leaders Make A Difference To Organizational Performance?", **Administrative Science Quaterly**, No: 33, S: 3, Kasım 1988, s. 388-400.
- Volinsky, T. Chris ve Adrian E. Raftery. "Bayesian Information Criterion for Censored Survival Models", **Biometrics**, No: 56, Mart 2000, s. 256-262.
- Yoder, Kevin A., Les B. Whitbeck ve Dan R. Hoyt. "Event History Analysis of Antecedents to Running Away From Home and Being on the Street", **The American Behavioral Scientist**, No: 45, S: 1, 2001, s. 51-65.
- Wang, C.Y., Naisyin Wang ve Suojin Wang. "Regression Analysis When Covariates Are Regression Parameters of a Random Effects Model for Observed Longitudinal Measurements", **Biometrics**, No: 56, 2000, s. 487-495.
- Wang, Peiming, Iain M. Cockburn ve Martin L. Puterman. "Analysis of Patent-a Mixed-Poisson-Regression-Model Approach", **American Statistical Association**, No: 16, S: 1, ocak 1998, s. 40.
- Wang, Yanping, Raymond H. Myers, Eric P. Smith ve Keying Ye. "D-Optimal Desings for Poisson Regression Models", **Journal of Statistical Planning and Inference**, No: 136, 2006, s. 2831-2845.
- Weiren, Wang ve Felix Famoye. "Modeling Household Fertilty Decisions with Generalized Poisson Regression", **Journal of Economics**, No: 10, 1997, s. 273-283.
- Yang, Yaning ve Zhiliang Ying. "Marjinal Proportional Hazards Models for Multiple Event-Time Data", **Biometrika**, No: 88, 2001, s. 581-586.

A. Zaleznik. “*Managers and Leaders Are They Different*”, **Inside The Mind Of The Leader Harvard Business Review**, Ocak 2004, s. 74-81.

Zhou, M. D., Yang, Y. Wang ve S. Nadarajah. “*Conditional Poisson Distributions*”, **Probability in the Engineering and Informational Sciences**, No: 20, 2006, s. 95-102.

İnternet'ten Alınan Kaynaklar:

Cameron, A.C. ve Pravin K. Trivedi. **Essential of Count Data Regression**. <http://cameron.econ.uc.davis.edu/research/CTE01preprint.pdf>, (21 Eylül 2003).

Çakır, Banu. **Tıbbi Araştırma teknikleri: Temel Özellikler, Yarar ve Kısıtlılar, Olası Hata Kaynakları**. [http : // www . toraks . org . tr / kisokulu3 – pp - pdf/Banu_cakir.pdf](http://www.toraks.org.tr/kisokulu3-pp-pdf/Banu_cakir.pdf), (2 Eylül 2002).

Klein, J.P. **Survival Distributions and Their Characteristics**. [www.biostat.mcw.edu / homepgs / klein / Teach Report #22](http://www.biostat.mcw.edu/homepgs/klein/Teach%20Report%22.pdf), (28 Kasım 2002).

Pedan, Alex. **Analysis of Count Data Using the SAS System**, [http://www.2 sas.Com/proceedings /sugi26/p247-26.pdf](http://www2.sas.com/proceedings/sugi26/p247-26.pdf), (3 Ekim 2005).

Zhang, Ying. **Logistic Regression Model**. [http://pegasus.cc.ucf.edu./zhang/Sta6938/topic.pdf](http://pegasus.cc.ucf.edu/~zhang/Sta6938/topic.pdf), (10 Ocak 2004).

Yayınlanmamış Tezler:

Bülbül, Serpil Ergün. “*İstatistiksel Başarısızlık Zaman Modelleri ve Finanssal Analizlerde Uygulanması*”, MÜ Sosyal Bilimler Enstitüsü (Yayınlanmamış Doktora Tezi), İstanbul, 1998.

Çinko, Murat. “*Diskriminant ve Lojistik Regresyon Analizlerinde Bootstrap Tekniğinin Kullanımı ve Kredi Riski Modeli Oluşturulması.*”, MÜ Sosyal Bilimler Enstitüsü (Yayınlanmamış Doktora Tezi), İstanbul, 2003.