Available online at www.alphanumericjournal.com

alphanumeric journal

The Journal of Operations Research, Statistics, Econometrics and Management Information Systems

Volume 6, Issue 2, 2018

Received: June 28, 2017
Accepted: November 15, 2018
Published Online: December 31, 2018

AJ ID: 2018.06.02.STAT.03
DOI: 10.17093/alphanumeric.323904
Research Article

Discrete Survival Time Models: An Application on Marriage Duration

Hilal Ölmez Hosta
M.Sc., Ankara, Turkey, hilolmez@hotmail.com

Nihal Ata Tutkun, Ph.D.
Assoc. Prof., Department of Statistics, Faculty of Science, Hacettepe University, Ankara, Turkey, nihalata@hacettepe.edu.tr

* Hacettepe Üniversitesi, Fen Fakültesi, İstatistik Bölümü, 06800-Beytepe, Ankara, Türkiye

ABSTRACT

In survival analysis which is used in the social and physical sciences, it is usually assumed that the observed process is continuous. Since this assumption is not appropriate for most of the survival time data structure, survival times are measured wrongly and unreliable results are obtained for the discrete survival time data. Continuous time survival models used for the time data have represented the structure of data in the studies regarding health sciences. The usage of the discrete time survival models in social sciences is more common since the structure of the studied data is more appropriate for the discrete time models. In this study, discrete time survival models are examined theoretically and were applied to “Research on Domestic Violence against Women in Turkey, 2008” data received from Turkish Statistical Institute. In order to examine the factor effecting the duration of marriage, discrete time survival models have been used and achieved results have been interpreted.

Keywords: Cox Regression, Discrete Time Survival Models, Logit Model, Non-Proportional Hazards, Complementary Log-Log Model

Kesikli Yaşam Süresi Modelleri: Evlilik Süreleri Üzerine Bir Uygulama

ÖZ

Fen ve sosyal bilimlerde kullanılabilen yaşam modellerinde genellikle ilgilenilen sürecin sürekli olduğu varsayılmaktadır. Ancak böyle bir varsayım bazı yaşam verilerinin yapısına uygun olmadığından yaşam süreleri hatalı ölçülmekte ve kesikli yaşam süresi verileri için güvenilir olmayan sonuçlar elde edilmektedir. Sürekli veriler için kullanılan sürekli yaşam modelleri, sağlık bilimlerinde yer alan uygulamalardaki verilerin yapısını yansıtabilir. Fakat kesikli zaman verilerinin en çok kullanıldığı sosyal bilimler alanında, mevcut verilerin yapısı kesikli modellere daha uygun olduğu için özellikle bu alanda kesikli yaşam süresi modellerinin kullanımı daha yaygındır. Bu çalışmada, kesikli yaşam süresi modelleri teorik açıdan incelenmiş ve Türkiye İstatistik Kurumu’ndan alınan “Türkiye'de Kadına Yönelik Aile İçi Şiddet Araştırması, 2008” verisine uygulanmıştır. Araştırmada yer alan kadınların evli kalma sürelerine etki eden faktörlerin incelenmesinde kesikli yaşam süresi modelleri kullanılmış ve sonuçlar yorumlanmıştır.

Anahtar Kelimeler: Cox Regresyon, Kesikli Yaşam Süresi Modelleri, Logit Model, Orantısız Tehlikeler, Tamamlayıcı Log-Log Modeli

1. Giriş

Yaşam çözümlemesi çalışmaya konu olan bir birimin gözlemlenmeye başladığı andan ilgilenilen olayı yaşayana kadar geçen süre ile ilgilenen istatistiksel bir araştırma yöntemidir. Günümüzde yaşam çözümlemesi bir hastalığın görülmesine kadar geçen süreyi, bir ekipmanın bozulmasına kadar geçen süreyi ya da deprem olana kadar geçen süreyi incelemek için birçok alanda kullanılabilen istatistiksel bir yöntem olmuştur. Başarısızlık süresi verilerine dayanan bu araştırma yönteminin kökeni yüzyıllar önce yapılan mortality (ölüm) ya da yaşam tablosuna dayanmaktadır. Ancak yaşam çözümlemesi ikinci dünya savaşına kadar bir araştırma yöntemi olarak ortaya konulamamıştır. Yaşam çözümlemesi yöntemi asıl olarak askeri malzemelerin bozulmalarına kadar geçen süreler ile ilgilenilmesiyle ortaya çıkmıştır. Geliştirilen bu yöntem savaşın sonunda ölüm verileri üzerine yapılan çalışmadan başarısızlık süresinin ilgilenildiği çalışmalara kadar pek çok alanda kullanılmaya başlanmış ve hızla yayılmıştır (Smith and Smith, 1972).

İlgilenilen olayın ortaya çıkma zamanının yani başarısızlık süresinin bağımlı değişken olarak ele alındığı yaşam modellerinin kullanımı Cox (1972) tarafından geliştirilen regresyon modeli ile yaygınlaşmıştır. Collett (1994) ve Kalbfleisch ve Prentice (1980)'in çalışmaları ile de ilerleme göstermiştir.

Yaşam çözümlemesinin başlıca uygulama alanları,

- Eğitim: Zorunlu eğitimin bitmesinden itibaren tam gün eğitimi bırakma zamanı, öğretmenlik mesleğini bırakma zamanı,
- Ekonomi: İşsizlik ya da çalışma süresi,
- Demografi: İlk doğum zamanı, ilk evlilik zamanı, ilk boşanma zamanı,
- Psikoloji: Uyarılara tepki süresi,
- Sağlık bilimleri: Ölüm, hastalığın nüks etmesi, ameliyat sonrası iyileşme süresi,
- Aktüerya: Sigortalı kalma süresi
- Mühendislik: Makine parçalarının bozulma süresi
- Uluslararası İlişkiler: Savaş sonrası barış süreci

biçimindedir. Yaşam çözümlemesi kullanılarak yapılan sağlık bilimlerine yönelik çalışmalarda, hava kirliliğinin ölüm (Pope ve diğ.., 1995) ve AIDS (Geskus, 2000) hastalığı üzerine etkisi, kalp nakli bekleyen hastaların bekleme süresinin hayatta kalmaları üzerine etkisi (Crowley ve Hu, 1977) gibi çeşitli konular incelenmiştir. Bunun dışında sosyal bilimlerde örgütsel değişim çalışmasında (Hannan ve Carroll, 1981), sigortalı kalm süresi çalışmasında Usui (1994), siyaset bilimlerindeki çeşitli uygulamalarda (Box-Steffensmeier and Jones, 1997), toplumsal hareketlerin gelişimi çalışmasında (Olzak, 1989), iş hareketliliği (Mills ve diğ.., 2006) ve evlilik (Blossfeld ve Mills, 2001) çalışmalarında da yaşam çözümlemesinden yararlanılmıştır.

Cox (1972)'de yaşam süresi "yaşam çözümlemesinde canlı ya da cansız bir nesnenin belirli bir başlangıç zamanı ile başarısızlığı arasında geçen süre" olarak tanımlanmıştır. Yaşam modellerinin temelinde yaşam süresine etki eden faktörlerin belirlenmesi yer almaktadır. Bu amaçla en yaygın kullanıma sahip modeller Cox orantılı tehlikeler modeli ve hızlandırılmış başarısızlık süresi modelleridir. Bu modellerde sürecin sürekli olduğu varsayımı yapılmaktadır. Ancak bu varsayım çoğu yaşam süresi verilerinin yapısına uygun olmadığından bazı çalışmalarda yaşam süreleri kesikli olmasına rağmen süreklilik varsayımının sağlanması için sürekli yapıya dönüştürülmektedir. Yaşam süresi kesikli olduğunda kesikli yaşam süresi modellerinin kullanılması daha uygun olacaktır. Bu modeller, riskin veya bir olayın gerçekleşme olasılığının modellenmesi ile ilgilenmektedir ve sürekli yaşam süresi modellerine göre bazı avantajları vardır. Kesikli yaşam süresi modelleri, kesikli zaman aralıklarında ölçülen ve özellikle klinik araştırmalarda kullanılan birçok yaşam süresi verisine daha uygun olmaktadır. Bu modeller ayrıca tehlike fonksiyonunun biçiminin incelenmesine de imkan vermektedir. Sürekli yaşam süresi için kullanılan Cox orantılı tehlikeler modelinde ise tehlike fonksiyonunun biçimi gözardı edilerek ilgilenilen olaya etki eden açıklayıcı değişkenler incelenmektedir. Bununla birlikte, tehlike fonksiyonu ilgilenilen olayın gerçekleşip gerçekleşmediğini, ne zaman ortaya çıkabileceğini ve olayların ortaya çıkışının zamanla nasıl değiştiğini belirttiğinden tehlike fonksiyonunun incelenmesi yaşam çözümlemesinde önemli bir yere sahiptir.

Bu çalışmanın amacı, kesikli yaşam süresi modellerini inceleyerek gerçek bir veri üzerinde uygulanabilirliğini göstermektir. Bu kapsamda makalenin ikinci bölümünde kesikli yaşam süresi ile ilgili genel bilgiler verilerek kesikli yaşam süresi modelleri anlatılmıştır. Üçüncü bölümde ise Türkiye İstatistik Kurumundan alınan “Türkiye'de Kadına Yönelik Aile İçi Şiddet Araştırması, 2008 (TÜİK, 2008)” veri kümesindeki kadınların evlilik sürelerinin incelenmesinde klasik yaşam çözümlemesi yöntemleri ve kesikli yaşam süresi modelleri uygulanarak elde edilen sonuçlar değerlendirilmiştir.

2. Kesikli Yaşam Süresi

Yaşam modellerinin çoğunda ilgilenilen sürecin sürekli olduğu varsayılarak modeller oluşturulmaktadır. Ancak bu varsayım kimi yaşam verisinin yapısına uygun olmamaktadır. Bu nedenle ilgilenilen süreler hatalı ölçülmekte ve kesikli yaşam süresi verileri için güvenilir olmayan sonuçlar elde edilmektedir.

Kesikli yaşam süresi verileri iki farklı şekilde gözlemlenebilmektedir:

1. Birinci durum, yaşam sürelerinin ay ya da yıl gibi kesikli zaman aralıkları biçiminde gruplanabildiği durumdur. Bu durumda dönem uzunlukları pozitif tam sayılar ile özetlenebilir ve böylece geçiş sürecindeki (transition process) gözlemler sürekli değil kesikli olmuş olurlar. Yani ilgilenilen geçiş süreci aslında sürekli zamanda meydana gelmiş olsa da, veriler sürekli yapıda gözlemlenemezler. Veri kümesinde eş zamanlı (bağlı) gözlemlerin olması durumunda şüphelenilmesi gereken bu durum “aralıklı durdurma” olarak adlandırılmaktır. Fakat bazı sürekli yaşam modelleri, geçişlerin (transition) yalnızca farklı zamanlarda meydana gelebileceğini varsaymaktadır. Bu nedenle veri kümesinde aynı yaşam süresine sahip kişiler varsa, eş zamanlılığın gerçek olup olmadığı ya da bu eş zamanlılığın yalnızca yaşam sürelerinin gözlemlendiği aşamada gruplandırılmasından mı kaynaklanmış olduğu sorgulanabilir.

2. Kesikli yaşam sürelerinin gözlemlenebildiği ikinci durum ise, esas geçiş sürecinin yapısal olarak kesikli olduğu durumdur.

Sürekli veriler için en çok kullanılan ve uygulaması en yaygın olan sürekli yaşam modelleri, sağlık bilimlerinde yeralan uygulamalardaki verilerin yapısını yansıtabilir. Fakat kesikli zaman verilerinin en çok kullanıldığı sosyal bilimler alanında, mevcut verilerin yapısı kesikli modellere daha uygun olduğu için özellikle bu alanda kesikli yaşam süresi modellerinin kullanımı daha yaygındır (Box-Steffensmeier ve Jones, 2004; Jenkins, 2005).

Kesikli zaman yaklaşımının avantajları:

- Veriler özellikle geriye dönük bir şekilde toplandığında, olay zamanları genelde kesikli zaman birimleri ile ölçülür,
- Orantılı olmayan tehlikelerin modellenmesi için de kolaylık sağlar,
- Kesikli verilerin modellenmesinde kolaylık sağlar. Bu durum karışık veri yapıları ve süreçlerinin analiz edilebilmesi için oldukça önemlidir.
- Kesikli zaman yaklaşımının dezavantajları:
- Her bir zaman aralığında olay meydana gelene ya da durdurulana kadar gözlem dizisine sahip olabilmek için öncelikle veriler her bir veri için yeniden düzenlenmelidir.
- Gözlem periyotları yaşam sürelerinin ölçüldüğü zaman aralıklarının genişliklerine göre daha uzun ise veri seti çok büyük bir hale gelebilir (Box-Steffensmeier ve Jones, 2004; Jenkins, 2005).

Kesikli yaşam çözümlemesi yaklaşımı kullanılarak yapılan çalışmalardan bazıları Xie ve diğ. (2003), Yang (2004), Eleuteri ve diğ. (2007), Rubenbauer (2011) tarafından yapılan çalışmalardır.

2.1. Yaşam Sürelerinin Kesikli Zaman Aralıkları Biçiminde Gruplanabildiği Durumda (Aralıklı Durdurma) Tehlike ve Yaşam Fonksiyonları

Yaşam süresi ekseninin birbiri ile çakışmayan ve sınırlarının a0 = 0, a1, a2, a3, ..., ak zaman noktaları olduğu ardışık aralıklara bölünmüş olduğu varsayılsın. Bu durumda zaman aralıkları Eşitlik 1'de belirtildiği gibi tanımlanabilir:

$$\left[0 = a_0; a_1\right],\ \left(a_1; a_2\right],\ \left(a_2; a_3\right],\ \ldots,\ (a_{k-1}; a_k = \infty] \tag{1}$$

Bu tanımlama, $(a_{j-1};a_j]$ aralığının işaret edilen başlangıç tarihinden hemen sonra başladığını ve aralığın sonundaki aj tarihinin bu aralığın içine dahil olduğunu varsaymaktadır. Zaman aralıklarının birbirine eşit uzunlukta olmak zorunda olmadığı bu tanımda, a_1, a_2, ..., a_k zaman noktalarını yani süreleri göstermektedir. Buna göre j. aralığın başlangıcı için yaşam fonksiyonu;

$$P_r\left(T > a_{j-1}\right) = 1 - F\left(a_{j-1}\right) = S\left(a_{j-1}\right) \tag{2}$$

ile ifade edilmektedir. Eşitlik 2'de belirtilen F fonksiyonu başarısızlık fonksiyonudur. j-1 ile j. aralığın içinde olma olasılığı ise Eşitlik 3 ile verilmektedir.

$$P_r\left(a_{j-1} < T < a_j\right) = F\left(a_j\right) - F\left(a_{j-1}\right) = S\left(a_{j-1}\right) - S\left(a_j\right) \tag{3}$$

j. aralığın dışında olma olasılığı

$$P_r\left(T > a_j\right) = 1-F\left(a_j\right) = \left(a_j\right) = S\left(a_j\right) \quad (4)$$

biçiminde ifade edilmektedir. Buna göre, kesikli tehlike fonksiyonu (discrete hazard rate) olarak da tanımlanan aralıklı tehlike fonksiyonu (interval hazard rate), $h(a_j)$) (a_{j-1};a_j] aralığının içinde kalma olasılığına eşit olmaktadır ve Eşitlik 5'deki gibi ifade edilmektedir:

$$h\left(a_j\right)=P_r\left(a_{j-1} < T < a_j \setminus T > a_{j-1}\right)= P_r\,\frac{(a_{j\text{-}1} < T < a_j)}{(T > a_{j\text{-}1})} = 1 - \frac{S(a_{j\text{-}1}) - S(a_j)}{S(a_{j\text{-}1})} \quad (5)$$

Aralıklı tehlike fonksionu koşullu olasılık olduğu için değer aralığı 0 ile 1 arasındadır [$0\leq h(a_j)\leq 1$]. Buna göre de kesikli tehlike fonksiyonu, sürekli tehlike fonksiyonundan farklı olmaktadır.

Eşitlik 1' de verilen zaman aralıkları tanımı temel olarak eşit uzunlukta olmayan zaman aralıkları için kullanılsa da, uygulamada aralıkların bir hafta ya da bir ay gibi eşit uzunlukta olduğu varsayılmaktadır. Bu durumda zaman aralıkları pozitif tam sayılar ile gösterilebilir. (a_{j-1};a_j] aralığı a_j=1,2,3,… değerleri için (a_{j-1},a_j] şeklinde yeniden tanımlanarak j. aralığı temsil edebilir. Böylece kesikli tehlike oranı $h(a_j)$ yerine h(j) olarak gösterilebilir.

Aralıkların birbirine bir birim uzaklıkta olduğu durumda, yaşam olasılığı j. aralığın sonuna kadar her bir aralık için olayın meydana gelmemesi olasılıklarından oluşmaktadır. Örneğin; 3. aralıkta yaşam olasılığı, S_3 = (1. aralıkta yaşam olasılığı) x (1. aralıkta yaşadığı bilindiğine göre 2. aralıkta yaşama olasılığı) x (2. aralıkta yaşadığı bilindiğine göre 3. aralıkta yaşama olasılığı) biçiminde hesaplanır. Bu hesaplamanın genelleştirilmiş biçimi Eşitlik 6 ile verilmiştir:

$$\mathrm{S(j)} \equiv \mathrm{S_j} = (1-\mathrm{h_1})(1-\mathrm{h_2})\ldots(1-\mathrm{h_{j-1}})(1-\mathrm{h_j}) = \prod_{\mathrm{k=1}}^{\mathrm{j}}(1\text{-}\mathrm{h_k}) \quad (6)$$

Eşitlik 6'da aralıklı tehlike fonksiyonlarına göre yazılan S(j) kesikli yaşam fonksiyonunu ifade etmektedir. Tehlike oranının zaman içinde sabit olduğu yani yaşam sürelerinin geometrik dağılıma sahip olduğu özel durumlar için (örneğin bütün j değerleri için h_j=h olduğunda) yaşam fonksiyonu Eşitlik 7'de verilmiştir:

$$\mathrm{S_j} = (1\text{-}\mathrm{h})^{\mathrm{j}} \quad (7)$$

Kesikli zamanlı dağılım fonksiyonu ise;

$$\mathrm{F(j)} \equiv \mathrm{F_j} = 1\text{-}\mathrm{S(j)} = 1\text{-}\prod_{\mathrm{k=1}}^{\mathrm{j}}(1\text{-}\mathrm{h_k}). \quad (8)$$

biçimindedir Aralıklı durdurma durumunda kesikli zamanlı yoğunluk fonksiyonu f(j), j. aralığın içinde kalma olasılığıdır ve Eşitlik 9 ile gösterilir:

$$\mathrm{f(j)} = \mathrm{P_r}(\mathrm{a_{j\text{-}1}} < \mathrm{T} < \mathrm{a_j}) = \mathrm{S(j\text{-}1)\text{-}S(j)} = \left(\frac{1}{1-\mathrm{h_j}} - 1\right)\mathrm{S(j)}. \quad (9)$$

Bu nedenle, kesikli zamanlı yoğunluk fonksiyonu j-1 aralığının sonuna kadar hayatta kalma olasılığını ifade etmektedir. Kesikli zamanlı yoğunluk fonksiyonunun değer aralığı 0 ile 1 arasındadır (0≤ f(j) ≤1) (Box-Steffensmeier ve Jones, 2004; Jenkins, 2005).

2.2. Yaşam Süresinin Kesikli Olduğu Durumda Tehlike ve Yaşam Fonksiyonları

Yaşam sürelerinin yapısal olarak kesikli olduğu durumda; t yaşam süresi Eşitlik 10 ile tanımlanan f(j) olasılığına sahip kesikli rastlantı değişkenidir

$$f(j) \equiv f_j = P_r(T = j). \tag{10}$$

Eşitlik 10'da j pozitif tam sayılar kümesinin bir elemanıdır. j'nin eşit uzunluktaki aralıklar şeklinde ifade edildiği yaşam sürelerinin kesikli zaman aralıkları biçiminde gruplanabildiği (aralıklı durdurma) durumunun aksine bu yaklaşımda j döngüleri indekslemektedir. Ancak her iki durumda da yaşam süreleri için pozitif tam sayılar kullanıldığından aynı gösterimler kullanılmaktadır. j döngüsü için kesikli zamanlı yaşam fonksiyonu (Sj) ile gösterilmektedir ve Eşitlik 11'de verildiği gibidir:

$$S(j) = P_r(T \geq j) = \sum_{k=j}^{\infty} f_k. \tag{11}$$

j döngüsündeki kesikli zamanlı tehlike oranı, h(j), j zamanındaki olayın koşullu olasılığıdır. Bu oran Eşitlik 12 ile ifade edilmektedir:

$$h(j) = P_r(T = j \backslash T \geq j) = \frac{f(j)}{S(j\text{-}1)}. \tag{12}$$

Kesikli zamanlı yaşam fonksiyonun, süre değişkeninin eşit uzunlukta aralıklar olarak gruplandırıldığı durumdaki yaşam fonksiyonuna benzer biçimde yazılması çoğu durumda daha bilgi verici ve doğru olmaktadır. Bu durumda yaşam fonksiyonu Eşitlik 6'da verilmiş olduğu gibi ifade edilmektedir. Kesikli zamanlı başarısızlık fonksiyonu ise Eşitlik 8 ile ile yazılmaktadır. Benzer şekilde kesikli zamanlı yoğunluk fonksiyonu f(j) aralıklı durdurma durumundaki yoğunluk fonksiyonuna benzer şekilde yazılabilmektedir (Box-Steffensmeier ve Jones, 2004; Jenkins, 2005).

2.3. Yaşam Çözümlemesinde Kesikli Zaman ile Sürekli Zaman Arasındaki İlişki

Kesikli zaman yaklaşımında, Eşitlik 6'dan yararlanılarak;

$$\log S(j) = \sum_{k=1}^{j} \log(1 - h_k) \tag{13}$$

elde edilir. h_k'nın küçük değerleri için 1. dereceden Taylor serisi yaklaşımı kullanılarak Eşitlik 14'deki gibi elde edilebilir:

$$\log(1\text{-}h_k) \approx \text{-}h_k \tag{14}$$

Eşitlik 13 yeniden düzenlenir ise;

$$logS(j) \approx \sum_{k=1}^{j} h_k \tag{15}$$

elde edilir.

Eşitlik 11 sürekli zamanlı durum ile karşılaştırıldığında, kesikli tehlike oranları üzerinden toplam ile sürekli tehlike oranları üzerinden integral arasındaki paralellik formülize edilirse Eşitlik 16 elde edilmektedir:

$$\log S(t) = \text{-}H(t) = \text{-}\int_0^t \theta(u)du \tag{16}$$

Eşitlik 16'ya göre hk değeri küçüldükçe kesikli zamanlı tehlike oranı hj sürekli zamanlı tehlike oranı θ(t)'ye daha çok yaklaşır. Buna bağlı olarak da kesikli zamanlı yaşam fonksiyonu, sürekli zamanlı yaşam fonksiyonuna yaklaşma eğilimi gösterir.

Kesikli zaman ile sürekli zaman arasında kesin bir ayrım yapılamamaktadır. Bu nedenle çalışmalarda hangi veri türü (kesikli-sürekli) ile çalışılacağına karar vermek zorlaşmaktadır. Genel olarak yaşam süresini yaratan davranışsal süreç ile verilerin kaydedildiği sürecin yapısına bakarak çalışmanın yapılacağı veri türüne açık bir şekilde karar verilebilmektedir. Sosyal bilimlerde genel olarak üzerinde çalışılan davranışsal süreç sürekli zaman biçiminde meydana gelmektedir ve süre uzunlukları gruplandırılmış veri biçiminde kaydedilmektedir. Bu çalışmalarda gün ya da saat birimi ile kaydedilen veriler dahil olmak üzere bütün veriler gruplandırılmış olarak kaydedilmektedir. Bu nedenle, asıl önemli olan konu süre uzunluğuna göre gruplandırılmak için kullanılan aralık uzunluğu olmaktadır (Jenkins, 2005).

Eğer çalışma döneminin başladığı gün/ay/yıl biliniyorsa ya da birimin son olarak gözlemlendiği gün/ay/yıl biliniyorsa ve dönem uzunluğu birkaç ay ya da yıl ise bu durumda yaşam süresinin sürekli rastlantı değişkeni olarak düşünülmesi daha doğru olmaktadır. Ancak, dönem uzunluklarının yalnızca birkaç gün olması durumunda yaşam sürelerinin gün birimi ile gruplandırılarak kaydedilmesi ve aralıklı durdurma olarak ölçülebilen bir özelliğin seçilmesi daha anlamlı olmaktadır. Bu özellik seçilirken eş zamanlı yaşam sürelerinin varlığı da dikkate alınmalıdır. Eş zamanlı gözlemlerin fazla olması, özellik seçilirken yaşam sürelerinin de dikkate alınması gerektiğine işaret edebilir. Gruplama etkileri ile ilgili ilk çalışmalar Bergstrom ve Edin (1992), Petersen (1991) ve Petersen ve Koput (1992) tarafından yapılmıştır (Box-Steffensmeier ve Jones, 2004; Jenkins, 2005).

2.4. Kesikli Yaşam Süresi Modelleri

Kesikli yaşam süresinin modellenebilmesi için temel olarak iki model ele alınmaktadır. Bu modellerden ilki yaşam süresinin yapısal olarak kesikli olduğu durumlarda da kullanılan tamamlayıcı log-log modelidir (complementary log-log model). Bu model sürekli yaşam süresinin modellemesinde kullanılan orantılı tehlikeler modelinin kesikli yaşam süresinin modellenmesindeki karşılığıdır. Diğer model ise, yaşam sürelerinin yapısal olarak kesikli olduğu durumlar için geliştirilen ancak yaşam sürelerinin kesikli zaman aralıkları biçiminde gruplanabildiği durumda (aralıklı durdurma) da kullanılabilen logit modeldir. Bu model ile başarısızlık odds'larının orantılı tahminleri de elde edilebilmektedir. Bu bölümde modellerin incelenebilmesi için eşdeğişkenlerin sabit olduğu varsayılacaktır.

Tamamlayıcı log-log modeli

Elde edilen yaşam süresi verileri aralıklı durdurulmuş veya aralıklarla gruplandırılmış olmasına rağmen, sürekli yaşam modelleri θ(t, X) biçimindeki tehlike oranı ile ifade edilmektedir. Bu durum, bazı zaman aralıklarının içinde kalan yaşam sürelerinin kesin olarak bilinememesine neden olmaktadır. Tamamlayıcı log-log modeli kullanılarak, yaşam süresi verisinin yapısına uygun sürekli tehlike oranını tanımlayan bir parametre tahmini elde edilir. Eşitlik 17 ile tanımlanan aj zamanındaki yaşam fonksiyonu,

$$S(a_j, X) = \exp\left[-\int_0^{a_j} \theta(u, X) du \right]. \quad (17)$$

biçimindedir. Orantılı tehlikeler varsayımının sağlandığı düşünülürse,

$$\theta(t, X) = \theta_0(t) e^{X\beta'} = \theta_0(t)\lambda \quad (18)$$

elde edilebilir. Eşitlik 18'de $\beta' X \equiv \beta_o + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k$ ve $\lambda \equiv \exp(\beta' X)$ 'dir. Buna göre yaşam fonksiyonu Eşitlik 19'daki gibi yeniden yazılabilir:

$$S(a_j, X) = \exp\left[-\int_0^{a_j} \theta_0(t)\lambda\, du \right] = \exp\left[-\lambda \int_0^{a_j} \theta_0(t) du \right] = \exp\left[-H_j \lambda \right]. \quad (19)$$

Eşitlik 19'da $H_j \equiv H(a_j) = \int_0^{a_j} \theta_0(u, X) du$ şeklinde elde edilmektedir. Buna göre aj zamanındaki temel yaşam fonksiyonu $S_0(a_j) = \exp(-H_j)$ biçiminde yazılabilmektedir. Kesikli zaman aralıklı tehlike fonksiyonu, $h(a_j, X) \equiv h_j(X)$ Eşitlik 20 ile tanımlanmaktadır:

$$h_j(X) = \frac{S(a_{j-1}, X) - S(a_j, X)}{S(a_{j-1}, X)} = 1 - \frac{S(a_j, X)}{S(a_{j-1}, X)} = 1 - \exp\left[\lambda(H_{j-1} - H_j) \right] \quad (20)$$

(Jenkins, 2005). Eşitlik 20'nin logaritması alınınca $\log(1 - h_j(X)) = \lambda(H_{j-1} - H_j)$ eşitliği elde edilmektedir. Bu eşitlik tekrar düzenlenirse $\log(-\log\left[\lambda(H_{j-1} - H_j) \right]) = \beta' X + \log(H_j - H_{j-1})$ olur. Benzer olarak (a_{j-1}, a_j) aralığı için kesikli zaman aralıklı temel tehlike oranı,

$$1 - h_{0j} = \exp(H_{j-1} - H_j) \quad (21)$$

biçiminde elde edilir. Eşitlik 21'in logaritması alınıp düzenlenirse;

$$\log\left[-\log(1 - h_{0j}) \right] = \log(H_j - H_{j-1}) = \log\left[\int_{a_{j-1}}^{a_j} \theta_0(u) du \right] = \gamma_j \quad (22)$$

elde edilir. Eşitlik 22'de γ_j (a_{j-1}, a_j) aralığın sonundaki bütünleştirilmiş tehlike oranı ($\theta_0(t)$) ile aralığın başındaki tehlike oranı farkının logaritmasını ifade etmektedir. Bu ifade daha önce h(a_j, X) olarak ifade edilen tehlike oranı yerine kullanıldığında, aralıklı tehlike oranı için;

$$\log(-\log\left[1 - h_j(X) \right]) = \beta' X + \gamma_j \quad (23)$$

$$h(a_j, X) = 1 - \exp\left[-\exp(\beta' X + \gamma_j) \right] \quad (24)$$

yazılabilir. Log(-log(.)) dönüşümü tamamlayıcı log-log dönüşümü olarak adlandırılır. Bu nedenle kesikli zamanlı orantılı tehlikeler modeli genellikle "cloglog modeli" olarak adlandırılır (Box-Steffensmeier ve Jones, 2004; Jenkins, 2005).

Eğer bütün aralıklar eşit uzunlukta ise, zaman aralıklarını aralık numaraları yerine doğrudan her bir aralığının sonundaki zaman ile gösterebiliriz. Bu durumda kesikli zamanlı tehlike oranı,

$$h(j, X) = 1 - \exp\left[-\exp(\beta' X + \gamma_j) \right] \quad (25)$$

biçimindedir. Eşitlik 25'den de görülebileceği gibi cloglog modeli, belirli bir doğrusal bağlantı fonksiyonunun genelleştirilmiş biçimidir. Aralıklı durdurulmuş yaşam verileri kullanılarak yapılan tahminler regresyon katsayılarının (β) ve parametrelerinin (γ_j) tahminlerinin elde edilmesini sağlar. β katsayıları sürekli zamanlı tehlike fonkiyonunu $\theta(t)=\theta_0(t)\exp(\beta^{'}X)$) tanımlayan katsayılar ile aynıdır. Ancak, temel tehlike fonksiyonunu tanımlayan parametreler ek varsayımlar olmadan belirlenemezler. Bu varsayımlardan ilki regresyon parametrelerinin (β_j) bütünleştirilmiş tehlike fonksiyonu değerleri arasındaki farkı göstermesidir. İkinci varsayım ise, regresyon parametrelerinin (β_j) her bir aralık içinde farklı tehlike fonksiyonu yapılarına sahip olmasıdır. Başka bir ifade ile regresyon katsayıları bir aralıktaki tehlike oranının zamana bağımlı modelini özetlemektedir. Ancak bu model sürekli zamanlı tehlike oranı için ilave varsayımlar olmaksızın kesin bir biçimde belirlenememektedir.

Regresyon katsayıları üzerine konulan kısıtlamalar kesikli zamanlı modellerin doğrudan parametrik orantılı tehlikeler modeline karşılık gelmesine neden olabillir. Ancak uygulamada genellikle regresyon katsayıları üzerine yukarıda değinilen kısıtlamalar getirilmemektedir. Bunun yerine uygulamalarda regresyon katsayıları sürekli zamanlı tehlike oranı yerine kesikli zamanlı tehlike oranındaki zamana bağımlılığı belirtmektedir. Başka bir ifade ile aralıklar arasındaki regresyon katsayısının değişimi parametrik bir fonksiyonel biçim kullanılarak elde edilir.

Tamamlayıcı log-log modeli, sürekli zamanlı model ve aralıklı durdurulmuş yaşam süresi verileri ile uyumlu tek model değildir. Sueyoshi (1995) araştırmasında, izleyen başlıkta verilen lojistik tehlike modelinin, log-lojistik dağılıma sahip olan sürekli zamanlı modeli ile uyumlu olabileceğini göstermiştir (Jenkins, 2005).

Logit model

Yaşam sürelerinin yapısal olarak kesikli olduğunda, sonuçların yorumlanmasında fark olsa da önceki başlıkta değinildiği gibi tamamlayıcı log-log modeli kullanılabilir. Alternatif olarak yaşam süreleri kesikli olduğunda uygulamada genellikle orantılı odds modeli (proportional odds model) olarak adlandırılan model kullanılmaktadır. Bu modelde odds oranları tehlike oranlarına işaret etmektedir (Box-Steffensmeier ve Jones, 2004; Jenkins, 2005).

Yaşam sürelerinin aylık olarak kaydedilmiş olması durumunda, orantılı odds modeline göre j ayından bir önceki ayın sonuna kadar hayatta kalma olasılığını veren göreli (relative) odds oranı Eşitlik 26 ile ifade edilmektedir

$$\frac{h(j,X)}{1\text{-}h(j,X)}=\left[\frac{h_0(j)}{1-h_0(j)}\right]\exp(\beta'X) \tag{26}$$

Eşitlik 26'da h(j,X) j ayı için kesikli zamanlı tehlike oranını, $h_0(j,X)$ ise temel tehlike oranını göstermektedir. Verilen bir zamanda geçiş yapmanın göreli odds oranı iki bileşenin toplamı olarak ifade edilmekte ve Eşitlik 27'deki gibi elde edilmektedir. Bu bileşenlerden ilki bütün birimler için ortak olan göreli odds oranı, ikincisi ise birime özgü ölçeklendirme faktörüdür.

$$\operatorname{logit}\left[h(j,X)\right]=\log\left[\frac{h(j,X)}{1\text{-}h(j,X)}\right]=a_j+\beta^{'}X \tag{27}$$

Eşitlik 27'de $a_j=\operatorname{logit}\left[h_o(j)\right]$ olarak tanımlanmaktadır ve Eşitlik 28'deki gibi de yazılabilir:

$$h(j,X) = \frac{1}{1+\exp(-\ a_j -\ \beta^{'}X)} \tag{28}$$

Eşitlik 28 ile verilen model lojistik tehlike modeli ya da logit model olarak adlandırılır. Bu model genişletilerek orantılı odds modeli elde edilmektedir. Teorik olarak aj değişkeni yaşam süresinin ölçüldüğü her bir ay için farklı değerler alabilir. Ancak, genellikle aj'deki değişimin yapısı j değişkeninin bazı fonksiyonları kullanılarak tanımlanır (Box-Steffensmeier ve Jones, 2004; Jenkins, 2005).

Kesikli zaman modellerinde süre bağımlılığının modellenmesi için kullanılan fonksiyonlar

Süre bağımlılığı modelde kullanılan parametrelerin zamanın bir fonksiyonu olması anlamına gelmektedir. Süre bağımlılığına örnek olarak aşağıda açıklanan durumlar verilebilir:

- rlog(j) fonksiyonu ile r>0 iken tehlike oranı monoton olarak arttığından, r<0 iken monoton olarak azaldığından ve r=0 olduğunda sabit kaldığından sürekli zamanlı Weibull modelinin kesikli zamandaki benzeri olarak düşünülebilir. Bu model lojistik tehlike modeli ile birleştirildiğinde, elde edilen model $\text{logit}\left[h(j,X)\right] = r\log j + \beta^{'}X$ olur. Bu modelde r parametresi β vektöründeki sabit terim ve eğim parametreleri ile birlikte tahmin edilebilen bir parametredir.
- Biçim parametreleri z_1, z_2, z_3,..., z_p olan zamanın p. dereceden polinom fonksiyonu $z_1 j + z_2 j^2 + z_3 j^3 + \ldots + z_p j^p$ ile verilsin. Zamanın karesel fonksiyonunda (p=2 olduğunda) aralık tehlike oranı U biçiminde ya da ters U biçiminde olmaktadır. Bu model cloglog tehlike modeli ile birleştirildiğinde elde edilen model $c\log\log\left[h(j,X)\right] = z_1 j + z_2 j^2 + \beta^{'}X$ olur. Elde edilen modelde, z_1 ve z_2 parametreleri β vektöründeki sabit terim ve eğim parametreleri ile birlikte tahmin edilebilen bir parametredir.

Örneğin, ayların aynı tehlike oranına sahip olduğu ancak tehlikenin gruplar arasında değişken olduğu parçalı sabit model verilsin. Bu model lojistik tehlike modeli ile birleştirildiğinde, elde edilen model $c\log\log\left[h(j,X)\right] = \gamma_1 D_1 + \gamma_2 D_2 + \ldots + \gamma_j D_j + \beta^{'}X$ olur. Burada D_l, j=l iken 1 değerini alan, j'nin diğer bütün değerleri için 0 değerini alan iki sonuçlu bir değişkendir. Örneğin araştırmacı her bir aralığa ya da aralık gruplarına karşılık gelen kukla değişkenler oluştursun. Bu durumda, modelin tamamının tahmini yapılırken β vektörü sabit terim içermez aksi halde model eşdoğrusal olur. Alternatif olarak sabit terim modelde yer alırsa, kukla değişkenlerden biri modelden çıkarılabilir (Box-Steffensmeier ve Jones, 2004; Jenkins, 2005).

Tehlike fonksiyonunun biçiminin seçilmesi sürekli zamanlı modellerde araştırmacının farklı parametrik fonksiyon biçimlerini seçebilmesi gibi kesikli zaman modellerde de araştırmacıya bağlıdır. (Box-Steffensmeier ve Jones, 2004; Jenkins, 2005).

Uygulamada, cloglog ve logit modelleri aynı süre bağımlılığı özelliklerini taşımaktadır. Bu modellerde tehlike oranları göreceli olarak küçük olduğu sürece aynı açıklayıcı değişken değerleri benzer tahminlerin elde edilmesini sağlamaktadır. Bu durumun nedeni Eşitlik 29 ile açıklanmaktadır:

$$\text{logit}(h) = \log\left[\frac{h}{1-h}\right] = \log(h) - \log(1-h) \tag{29}$$

Eşitlik 29'da h →0 iken logaritma (1-h) değeri de 0'a yaklaşır. Yani, tehlike oranı h'nin küçük değerleri için logit (h) değeri log(h) değerine yaklaşır (logit (h)≈log(h)).

Yeterince küçük tehlike oranına sahip olunduğunda, süre bağımlılığı doğrusal bir fonksiyon ise orantılı odds modeli tehlike oranının logaritmasının bağımlı değişken olduğu model ile benzer sonuçlar verir. Buna göre, β nın kesikli zamanlı orantılı tehlike modelden elde edilen tahmini log(θ)'nın doğrusal bir fonkiyona sahip olduğu sürekli zamanlı modelden elde edilen tahmine karşılık gelmektedir (Box-Steffensmeier ve Jones, 2004; Jenkins, 2005).

3. Uygulama

3.1. Veri Yapısı

Bu bölümde Türkiye İstatistik Kurumundan alınan "Türkiye'de Kadına Yönelik Aile İçi Şiddet Araştırması, 2008 (TÜİK, 2008)" verilerinin bir bölümü düzenlenerek incelenen modellere uygulanarak, sonuçlar değerlendirilmiştir.

Ergöçmen ve diğ. (2009)'da belirtildiği gibi "Türkiye'de Kadına Yönelik Aile İçi Şiddet Araştırması 2008 (TÜİK, 2008)" çalışması "kadınların yaşadığı aile içi şiddetin büyüklüğü, içeriği, neden ve sonuçları ile risk faktörlerinin anlaşılması amacıyla ülke çapında yürütülmüş kapsamlı bir araştırmadır". Kadına yönelik aile içi şiddet çok boyutlu bir sorundur. Sadece kadınları değil, çocuklarını, ailelerini ve toplumu da ilgilendirmektedir. "Kadına yönelik şiddetle daha etkili bir şekilde mücadele etmek için hedeflenen politika ve programların oluşturulmasına ve mevcut politika ve programların geliştirilmesine imkan sağlamak ve kadına yönelik şiddet ile ilgili ülke düzeyinde veri oluşturmak amacı ile Türkiye İstatistik Kurumu tarafından yapılan araştırmanın kapsamının, Türkiye sınırları dahilinde bulunan tüm yerleşim yerlerindeki hanehalkları, kapsanan kitlenin ise hanede bulunan kadınlar" olduğu araştırmada belirtilmiştir. Dünya Sağlık Örgütü'nün "Multi-country Study on Women's Health and Domestic Violence Against Women (Garia-Moreno ve diğ., 2005)" çalışmasında kullanılan anket soru kağıdı Türkiye'ye uyarlanarak 2008'deki araştırmada kullanılmıştır. Ankette kullanılan kitle, örneklem bilgilerine, anket sorularına ve uygulanmasına ilişkin detaylı bilgilere yer verilmiştir.

Bu çalışmada evlilik süresi yaşam süresi olarak ele alınarak, yaşam çözümlemesi yöntemleri ile incelenmiştir. TÜİK verisi ayrıntılı bir biçimde incelenerek yaşam çözümlemesi yöntemleri kullanılacak biçime dönüştürülmüştür.

TÜİK tarafından yapılan çalışmada %86'lık bir cevaplama oranı ile 12.795 kadın ile görüşülmüştür. Çalışmamıza ilk evliliğini yapan kadınlar içinden evlilik süresi 10 yıl ve daha az olanlar dahil edilmiş, eksik gözlemlerin çıkarılması ile yaşam çözümlemesi için incelenebilen örneklem büyüklüğü 2627 olarak belirlenmiştir. Çalışmada yer alan analizler için SPSS ve STATA programlarından yararlanılmıştır.

Yapılan bu çalışmada kadınların ilk evlilikleri ele alınmıştır. Çalışmaya konu olan kadınlar için evlilik sürelerini etkileyen faktörler incelenmiştir. Çalışmada, evlilik süreleri (yıl) yaşam süresi olarak alınmıştır. Boşanma durumu ise başarısızlık olarak ifade edilmiştir. Boşanma durumunun gerçekleşmediği gözlemler durdurulmuş olarak tanımlanmıştır. Durdurulmuş gözlemler için evlilik tarihinden çalışmanın yapıldığı 2008 yılına kadar olan süre yaşam süresi olarak tanımlanmıştır. Çalışmaya dahil edilen 2627 kadından 77'sinde (%2.9) başarısızlık ve 2550'sinde (%97.1) durdurma gözlenmiştir. Uygulamada kadının evlendiği yaş (kadın yaş), erkeğin evlendiği yaş (erkek yaş), araştırmanın uygulandığı bölgeler (bölge), yerleşim yeri, düzenli olarak çalışılan iş (kadın_iş), elde edilen kazanç ya da gelir (kadın_gelir), rahatlıkla ziyaret edilebilecek yakınlıkta aile (yakında aile), nikah türü, kadının evliliğe rızası (evliliğe rıza), evlenilirken alınan başlık parası (başlık parası), eşin düzenli olarak çalıştığı iş (erkek iş), eşin başka bir kadın ile ilişkisi olması (ilişki), eşin kadının arkadaşlarıyla görüşmesini kısıtlaması (arkadaş), eşin kadının ailesi ile görüşmesini kısıtlaması (aile), eşin kadın istemediği halde gelirini elinden alması (elden gelir alma), eşin karısını korkutması/tehdit etmesi (tehdit), kadının eşi ile akrabalık durumu (akrabalık), çocuk sayısı (çocuk), eşin alkol kullanması (alkol), eşin kumar oynaması (kumar), eşin uyuşturucu kullanması (uyuşturucu), yaşanılan şiddet olayı sonucunda polise başvurulma durumu (polis), kadının eğitim düzeyi (kadın_eğitim), kadının sosyal güvenlik durumu (kadın sosyal güvenlik), erkeğin eğitim düzeyi (erkek_eğitim), erkeğin sosyal güvenlik durumu (erkek sosyal güvenlik) ve kadının eşinden şiddet görmesi (şiddet) değişkenleri ele alınmıştır. Bu değişkenler Tablo 1'de verilmiştir.

Değişken	**Değişken Düzeyleri**	**Toplam Denek Sayısı (n)**	**%**	**Olay Sayısı**	**Durdurulmuş Denek Sayısı**
Kadın Yaş	7-17	373	14.2	14	359
	18-25	1867	71.1	52	1815
	26-30	298	11.3	8	290
	31-35	68	2.6	2	66
	36+	21	0.8	1	20
Erkek Yaş	7-17	12	0.5	2	10
	18-25	1382	52.6	41	1341
	26-30	981	37.3	25	956
	31-35	183	7.0	3	180
	36+	69	2.6	6	63
Bölge	Batı	714	27.2	26	688
	Güney	248	9.4	8	240
	Merkez	607	23.1	24	583
	Kuzey	344	13.1	6	338
	Doğu	714	27.2	13	701
Yerleşim Yeri	Kent	2090	79.6	65	2025
	Kır	537	20.4	12	525
Kadın_İş	İş Yok	2506	95.4	71	2435
	İş Var	121	4.6	6	115
Kadın Gelir	Gelir Yok	2155	82.0	34	2121
	Gelir Var	472	18.0	43	429
Nikah Türü	Resmi	66	2.5	8	58
	Dini	81	3.1	10	71
	Resmi ve Dini	2480	94.4	59	2421
Evliliğe Rıza	Rıza Yok	66	2.5	9	57
	Rıza Var	1038	39.5	24	1014
	Cevapsız	1523	58.0	44	1479
Başlık Parası	Yok	196	7.5	3	193
	Var	2431	92.5	74	2357
Yakında Aile	Yok	769	29.3	10	759
	Var	1858	70.7	67	1791
Erkek-İş	İş Yok	855	32.5	50	805
	İş Var	32	1.2	2	30
	Cevapsız	1740	66.2	25	1715
İlişki	Yok	2430	92.5	34	2396
	Var	109	4.1	29	80
	Olabilir	26	1.0	8	18
	Cevapsız	62	2.4	6	56
Arkadaş	Kısıtlamaz	2261	86.1	46	2215
	Kısıtlar	366	13.9	31	335
Aile	Kısıtlamaz	2393	91.1	41	2352
	Kısıtlar	234	8.9	36	198
Elden Gelir Alma	Cevapsız	2536	96.5	58	2478
	Evet	91	3.5	19	72
Tehdit	Yok	2504	95.3	51	2453
	Var	123	4.7	26	97
Akrabalık	Yok	2211	84.2	67	2144
	Var	213	8.1	5	208
	Cevapsız	203	7.7	5	198

Değişken	**Değişken Düzeyleri**	**Toplam Denek Sayısı (n)**	**%**	**Olay Sayısı**	**Durdurulmuş Denek Sayısı**
Çocuk	Yok	468	17.8	27	441
	Kız çocuk	684	26.0	15	669
	Erkek çocuk	878	33.4	26	852
	Hem kız hem erkek çocuk	597	22.7	9	588
Alkol	Kullanmıyor	2418	92.0	51	2367
	Kullanıyor	209	8.0	26	183
Kumar	Oynamıyor	2587	98.5	62	2525
	Oynuyor	40	1.5	15	25
Uyuşturucu	Kullanmıyor	2610	99.4	66	2544
	Kullanıyor	17	0.6	11	6
Polis	Başvurulmadı	824	31.4	47	777
	Başvuruldu	29	1.1	9	20
	Cevapsız	1774	67.5	21	1753
Kadın_Eğitim	İlkokul	1342	51.1	35	1307
	Ortaokul+İlköğretim	338	12.9	13	325
	Lise	655	24.9	18	637
	Üniversite	292	11.1	11	281
Kadın_Sosyal Güvenlik	Yok	2333	88.8	58	2275
	Var	294	11.2	19	275
Erkek_Eğitim	İlkokul	810	30.8	24	786
	Ortaokul+İlköğretim	391	14.9	18	373
	Lise	906	34.5	19	887
	Üniversite	505	19.2	13	492
	Cevapsız	15	0.6	3	12
Erkek_Sosyal Güvenlik	Yok	590	22.5	13	577
	Var	1875	71.4	43	1832
	Bilinmiyor	162	6.2	21	141
Şiddet	Yok	1772	67.5	20	1752
	Fiziksel	556	21.2	34	522
	Fiziksel/Cinsel	214	8.1	20	194
	Cinsel	85	3.2	3	82

Tablo 1. Kullanılan değişkenler ve düzeyleri

3.2. Klasik Yaşam Çözümlemesi Sonuçları

Klasik yaşam çözümlemesi yöntemlerini kullanmadan önce orantılı tehlikeler varsayımının incelenmesi kullanılacak test istatistiklerine ya da modellerine karar vermede önemlidir. Belirli bir değişken için Schoenfeld artıkları ile bireylerin başarısızlık sürelerinin rankı arasındaki korelasyon kullanılarak orantılı tehlikeler varsayımı incelenebilmektedir (Kleinbaum ve Klein, 2005). Bu çalışmada da orantılı tehlikeler varsayımı Schoenfeld artıkları ile yaşam süresinin rankı arasındaki korelasyon testi ile incelenmiştir. Bu test istatistiği için yokluk hipotezi "orantılı tehlikeler varsayımı sağlanmaktadır" biçimindedir. Yerleşim yeri ($p=0.0024$), başlık parası ($p=0.0154$), kız çocuk ($p=0.0024$), erkek çocuk ($p=0.009$) alkol ($p=0.046$) değişkenleri için orantılı tehlikeler varsayımının sağlanmadığı, diğer değişkenler için ise varsayımın sağlandığı ($p>0.05$) sonucuna ulaşılmıştır.

Yaşam eğrileri arasındaki farklılığın olup olmaması durumu orantılı tehlikeler varsayımını sağlayan değişkenler için log-rank testi, varsayımı sağlamayan değişkenler için Breslow testi kullanılarak incelenebilir. Bu testler için değişkenler tek tek ele alınır ve değişken düzeyleri arasında fark olup olmadığı incelenir. Bu çalışmada da orantısız tehlikelere sahip değişkenler için log-rank yerine Breslow testi kullanılmıştır. Bu testlerin sonuçlarına göre erkek yaş ($p=0.002$), kadın gelir ($p=0.000$), nikah türü ($p=0.000$), evliliğe rıza ($p=0.000$), erkek iş ($p=0.000$), ilişki ($p=0.000$), arkadaş ($p=0.000$), aile ($p=0.000$), elden gelir alma ($p=0.000$), tehdit ($p=0.000$), çocuk ($p=0.000$), alkol ($p=0.000$), kumar ($p=0.000$), uyuşturucu ($p=0.000$), polis ($p=0.000$), kadın sosyal güvenlik ($p=0.000$), erkek eğitim ($p=0.001$), erkek sosyal güvenlik ($p=0.000$), şiddet ($p=0.000$) ve yakında aile ($p=0.002$) değişkenlerinin düzeyleri arasında yaşam olasılıkları açısından %95 güven düzeyinde fark olduğu görülmüştür. Diğer değişkenlerin düzeyleri arasında yaşam olasılıkları açısından anlamlı bir farklılık görülmemiştir.

Yaşam modellerinden Cox regresyon modeli veri için uygulanabilir. Ancak bu model orantılı tehlikeler varsayımına sahiptir ve ilgili varsayımının incelenmesi gerekmektedir. Verinin tamamı için orantılı tehlikeler varsayımı test edildiğinde p=0.022 olarak elde edildiğinden orantılı tehlikeler varsayımının %95 güven düzeyinde sağlanmadığı sonucuna ulaşılmıştır. Bu nedenle, orantılı tehlikeler varsayımı sağlanmadığından klasik yaşam modellerinden olan Cox regresyon çözümlemesinin sonuçlarını kullanmak yanıltıcı sonuçlar elde edilmesine neden olabilir. Bu durumda Cox regresyon modeli yerine tabakalandırılmış ya da genişletilmiş Cox regresyon modeli kullanılabilir. Ancak, orantılı tehlikeler varsayımını sağlamayan çok fazla değişken olduğu için orantısız tehlikeler için kullanılabilen tabakalandırılmış Cox regresyon modelinin kurulmasında güçlükler olmaktadır. Çünkü verinin 10 tabakadan oluşan gruplara ayrılması bilgi kaybına neden olmaktadır. Genişletilmiş Cox regresyon modelinde özellikle birden fazla açıklayıcı değişkende orantısız tehlikeler görülüyorsa, zamanın fonksiyonunu belirlemek için kesin kurallar bulunmamaktadır. Bu nedenle buradaki kullanımı uygun olmayacaktır.

Veri yapısı kesikli olduğundan ve Cox regresyon modeli ve uzanımlarını kullanmak doğru olmadığından kesikli yaşam modellerinin kullanımını uygundur.

3.3. Kesikli Yaşam Süresi Modeli Sonuçları

Verilere kesikli yaşam süresi modellerinden tamamlayıcı log-log ve logit modeler uygulanmış ve sonuçlar elde edilmiştir. Modellerin anlamlılığını test etmek için olabilirlik oranı (LR) test istatistiği kullanılmış ve tüm modellerin istatistiksel olarak anlamlı olduğu ($p < 0.05$) görülmüştür. Model seçiminde ise yöntemleri karşılaştırmak için Akaike Bilgi Kriteri (AIC) ve Bayesci Bilgi Kriteri (BIC) kullanılmış ve elde edilen sonuçlar Tablo 2'de verilmiştir:

Model	-2Log(L)	AIC	BIC
Tamamlayıcı Log-Log	293.3	419.26	789.20
Logit	295.6	421.58	791.53

Tablo 2. Kesikli yaşam süresi modellerinin karşılaştırılması

Tablo 2 incelendiğinde AIC ve BIC değerlerine göre aralarında çok büyük farklılıklar olmamakla birlikte tamamlayıcı log-log modelininin en uygun model olduğu görülmektedir. Buna göre tamamlayıcı log-log modeline ait sonuçları yorumlamak daha doğru olduğundan modele ait sonuçlar Tablo 3'de verilmiştir.

Değişken	$\widehat{\beta}$	Std. hata	p-değeri	$exp(\widehat{\beta})$	$\widehat{\beta}$ için güven aralıkları
Yıl(2)	-0.600	0.577	0.298	0.549	-1.730 ; 0.530
Yıl(3)	-2.067	0.810	0.011*	0.127	-3.656 ; -0.479
Yıl(4)	-2.169	0.850	0.011*	0.114	-2.836 ; -0.503
Yıl(5)	-1.173	0.666	0.078	0.309	-2.480 ; 0.133
Yıl(6)	-1.421	0.748	0.057	0.241	-2.887 ; 0.044
Yıl(7)	-1.511	0.781	0.053	0.221	-3.040 ; 0.019
Yıl(8)	-0.967	0.663	0.144	0.380	-2.266 ; 0.332
Yıl(9)	-0.973	0.766	0.204	0.378	-2.474 ; 0.528
Yıl(10)	-2.328	0.852	0.006*	0.097	-3.998 ; -0.658
Kadın Yaş(2)	0.678	0.462	0.143	1.069	-0.229 ; 1.584
Kadın Yaş (3)	0.316	0.698	0.650	1.372	-1.051 ; 1.683
Kadın Yaş (4)	0.573	1.033	0.579	1.773	-1.452 ; 2.598
Kadın Yaş (5)	0.834	1.725	0.629	2.303	-2.547 ; 4.216
Erkek Yaş(2)	-3.668	1.215	0.003*	0.026	-6.048 ; -1.287
Erkek Yaş (3)	-3.492	1.228	0.004*	0.030	-5.899 ; -1.085
Erkek Yaş (4)	-3.879	1.433	0.007*	0.021	-6.687 ; -1.072
Erkek Yaş (5)	-3.856	1.658	0.020*	0.021	-7.106 ; -0.606
Bölge(2)	-1.104	0.582	0.058	0.332	-2.244 ; 0.037
Bölge(3)	-0.010	0.421	0.981	0.990	-0.835 ; 0.816
Bölge(4)	-0.087	0.522	0.868	0.917	-1.110 ; 0.937
Bölge(5)	-1.700	0.548	0.002*	0.183	-2.775 ; -0.625
Yerleşim Yeri (2)	-0.242	0.424	0.567	0.785	-1.073 ; 0.588
Kadın İş (2)	-0.414	0.673	0.539	0.661	-1.733 ; 0.906
Kadın Gelir (2)	2.082	0.403	0.000*	8.019	1.292 ; 2.872
Nikah Türü(2)	0.544	0.872	0.533	1.723	-1.165 ; 2.252
Nikah Türü(3)	-1.403	0.720	0.052	0.246	-2.814 ; 0.009
Evliliğe Rıza(2)	-2.017	0.662	0.002*	0.133	-3.315 ; -0.719
Evliliğe Rıza(3)	-1.847	0.659	0.005*	0.158	-3.139 ; -0.555
Başlık Parası	0.104	0.732	0.887	1.110	-1.330 ; 1.539
Yakında Aile	1.059	0.443	0.017*	2.884	0.191 ; 1.927
Erkek İş(2)	-1.182	0.946	0.211	3.262	-0.671 ; 3.036
Erkek İş(3)	-1.456	0.448	0.001*	0.233	-2.335 ; -0.577
İlişki(2)	2.332	0.381	0.000*	10.294	1.585 ; 3.078
İlişki(3)	1.149	0.756	0.129	3.156	-0.333 ; 2.632
İlişki(4)	1.341	0.608	0.027*	3.823	0.150 ; 2.532
Arkadaş (2)	-0.658	0.470	0.161	0.518	-1.579 ; 0.263
Aile (2)	1.376	0.454	0.002*	3.957	0.485 ; 2.266
Elden Gelir Alma (2)	0.623	0.500	0.213	1.865	-0.357 ; 1.604
Tehdit (2)	0.652	0.458	0.155	1.919	-0.246 ; 1.549
Akrabalık(2)	0.471	0.673	0.484	1.601	-0.848 ; 1.789
Akrabalık(3)	-0.745	0.706	0.292	0.475	-2.129 ; 0.640
Çocuk(2)	-0.838	0.541	0.122	0.433	-1.898 ; 0.223
Çocuk(3)	-0.512	0.485	0.292	0.600	-1.463 ; 0.440
Çocuk(4)	-1.088	0.666	0.102	0.337	-2.394 ; 0.218
Alkol (2)	1.022	0.410	0.013*	2.778	0.218 ; 1.825
Kumar (2)	0.883	0.653	0.176	2.418	-0.397 ; 2.162
Uyuşturucu (2)	0.689	0.790	0.383	1.992	-0.859 ; 2.238
Polis(2)	0.700	0.663	0.291	2.014	-0.600 ; 2.000
Polis(3)	1.621	3.586	0.651	5.057	-5.408 ; 8.650
Kadın Eğitim(2)	0.211	0.490	0.667	1.235	-0.750 ; 1.171
Kadın Eğitim(3)	-0.410	0.469	0.383	0.664	-1.329 ; 0.510
Kadın Eğitim(4)	-0.175	0.713	0.807	0.840	-1.571 ; 1.222
Kadın SGK (2)	0.087	0.567	0.878	1.091	-1.025 ; 1.198
Erkek Eğitim(2)	-0.653	1.107	0.555	0.520	-2.824 ; 1.517
Erkek Eğitim(3)	-0.802	1.101	0.467	0.448	-2.960 ; 1.357
Erkek Eğitim(4)	-1.092	1.090	0.316	0.336	-3.228 ; 1.044
Erkek Eğitim(5)	-1.398	1.170	0.232	0.247	-3.692 ; 0.895
Erkek SGK(2)	0.805	0.449	0.073	2.236	-0.075 ; 1.684
Erkek SGK(3)	0.646	0.573	0.260	1.907	-0.478 ; 1.769
Şiddet(2)	3.398	3.592	0.344	29.912	-3.642 ; 10.439
Şiddet(3)	2.307	3.614	0.523	10.044	-4.777 ; 9.391
Şiddet(4)	2.048	3.675	0.577	7.750	-5.155 ; 9.251

Tablo 3. Tamamlayıcı log-log modelinin sonuçları

Tablo 3'deki p değerleri incelendiğinde yıl(3), yıl(4), yıl(10) erkek yaş(2), erkek yaş(3),erkek yaş(4), erkek yaş(5), bölge(5), kadın gelir, evliliğe rıza(2), evliliğe rıza(3), yakında aile, erkek iş (3), ilişki (2), ilişki(4), aile ve alkol değişkenlerinin boşanmayı etkileyen önemli risk faktörleri olduğu %95 güven düzeyinde söylenebilmektedir (p<0.05 olduğundan). Buna göre aşağıdaki yorumlar elde edilmiştir:

- Bir yıllık evli olan çiftler, üç yıldır evli olan çiftlere göre yaklaşık 8 kat , dört yıldır evli olan çiftlere göre yaklaşık 9 kat ve 10 yıldır evli olan çiftlere göre yaklaşık 10 kat daha az boşanma riskine sahiptir.
- Yaşı 7-17 yaş aralığında olan erkekler, yaşı 18-25 yaş arası olan erkeklere göre yaklaşık 38 kat, 26-30 yaş arasında olan erkeklere gore yaklaşık 33 kat, yaşı 31-35 yaş arasında olan erkeklere gore yaklaşık 48 kat ve 36 yaş ve üzeri olan erkeklere göre yaklaşık 48 kat daha az boşanma riskine sahiptir. Ancak yaş değişkeni için düzeylerdeki sıklık dağılımı nedeni ile bu sonuçları yorumlamak çok doğru olmayacaktır.
- Batı bölgesinde oturan kadınların, doğu bölgesinde oturan kadınlara göre yaklaşık 5 kat daha az boşanma riskine sahip olduğu da Çizelge 3'deki sonuçlara bakılarak elde edilmektedir.
- Düzenli bir geliri olan kadınlar, düzenli bir geliri olmayan kadınlara göre yaklaşık 8 kat daha fazla boşanma riskine sahiptir.
- Evliliğe rızası olmayan kadınlar rızası olan kadınlara göre yaklaşık 8 kat, rızası olup olmadığını cevaplamayan kadınlara göre yaklaşık 6 kat daha az boşanma riskine sahiptir.
- Ailesi yakınında olan kadınlar, yakında ailesi olmayan kadınlara göre yaklaşık 3 kat daha fazla boşanma riskine sahiptir.
- Eşinin işi olmayan kadınlar, eşinin işi olup olmadığını cevaplamayan kadınlara göre yaklaşık 4 kat daha az boşanma riskine sahiptir.
- Eşinin başka bir kadın ile ilişkisi olan kadınlar, eşinin başka bir kadın ile ilişkisi olmayan kadınlara gore yaklaşık 10 kat daha fazla boşanma riskine sahiptir. Eşinin başka bir kadın ile ilişkisi olup olmadığına cevep vermeyen kadınlar ise eşinin başka bir kadın ile ilişkisi olmayan kadınlara göre yaklaşık 4 kat daha fazla boşanma riskine sahiptir.
- Eşi tarafından ailesi ile görüşmesi kısıtlanan kadınlar ailesi ile görüşmesi kısıtlanmayan kadınlara göre yaklaşık 4 kat daha fazla boşanma riskine sahiptir.
- Son olarak, eşi alkol kullanan kadınların eşi alkol kullanmayan kadınlara göre yaklaşık 3 kat daha fazla boşanma riskine sahip olduğu da Çizelge 3'deki sonuçlara bakılarak elde edilen yorumlar arasındadır.

4. Sonuç ve Öneriler

Bu çalışmada yaşam çözümlemesi hakkında genel bilgiler verilmiş ve yaşam çözümlemesinde kesikli yaşam süresi modelleri incelenmiş ve kullanılabileceği durumlar ortaya koyulmuştur. Kesikli yaşam süresi modelleri, ilk defa 1972 yılında Cox tarafından kesikli ve eş zamanlı gözlemler için önerilmiş ve bu modeller lojistik regresyonun bir benzeri olarak tanımlanmıştır. Genel olarak bilinen ve kullanılan iki kesikli yaşam süresi modeli vardır. Bunlar logit model ve tamamlayıcı log-log modelidir.

Uygulamada, Türkiye İstatistik Kurumunun 2015 yılında da gerçekleştirdiği Türkiye'de kadına yönelik aile içi şiddet çalışması geç açıklandığı için çalışmanın başladığı yıl elde edilebilir olan "Türkiye'de Kadına Yönelik Aile İçi Şiddet Araştırması 2008" verileri ele alınmıştır. Çalışmada boşanmayı etkileyen faktörlerin belirlenmesi için yaşam modelleri içerisinden öncelikle yaygın kullanıma sahip olan Cox regresyon çözümlemesi kullanılmıştır. Ancak, veri kümesinde orantılı tehlikeler varsayımının sağlanmadığı, bu nedenle klasik yaşam modellerinden olan Cox regresyon çözümlemesinin sonuçlarını kullanmanın yanıltıcı sonuçlara sebep olabilceği görülmüştür.

Uygulamanın üçüncü bölümünde incelenen verilere kesikli yaşam süresi modelleri olan tamamlayıcı log-log ve logit modelleri uygulanmıştır. Tamamlayıcı log-log ve logit modelleri karşılaştırıldığında, beklenildiği üzere aralarında çok büyük farklılıkların olmadığı ancak tamamlayıcı log-log modelinin veri yapısına daha uygun olduğu görülmüştür. Tamamlayıcı log-log modelinden elde edilen sonuçlara göre; bir yıllık evli olan çiftler, üç yıldır evli olan çiftlere göre yaklaşık 8 kat, dört yıldır evli olan çiftlere göre yaklaşık 9 kat ve 10 yıldır evli olan çiftlere göre yaklaşık 10 kat daha az boşanma riskine sahiptir. Ayrıca, yaşı 7-17 yaş aralığında olan erkekler, yaşı 18-25 yaş arası olan erkeklere gore yaklaşık 38 kat, 26-30 yaş arasında olan erkeklere gore yaklaşık 33 kat, yaşı 31-35 yaş arasında olan erkeklere gore yaklaşık 48 kat ve 36 yaş ve üzeri olan erkeklere göre yaklaşık 48 kat daha az boşanma riskine sahiptir. Batı bölgesinde oturan kadınların, doğu bölgesinde oturan kadınlara göre yaklaşık 5 kat daha az boşanma riskine sahip olduğu da tamamlayıcı log-log modelinin sonuçlarına bakılarak elde edilmektedir. Düzenli bir geliri olan kadınlar, düzenli bir geliri olmayan kadınlara göre yaklaşık 8 kat daha fazla boşanma riskine sahiptir. Evliliğe rızası olmayan kadınlar rızası olan kadınlara göre yaklaşık 8 kat, rızası olup olmadığını cevaplamayan kadınlara göre yaklaşık 6 kat daha az boşanma riskine sahiptir. Ailesi yakınında olan kadınlar, yakında ailesi olmayan kadınlara göre yaklaşık 3 kat daha fazla boşanma riskine sahip iken eşinin işi olmayan kadınlar, eşinin işi olup olmadığını cevaplamayan kadınlara göre yaklaşık 4 kat daha az boşanma riskine sahiptir. Eşinin başka bir kadın ile ilişkisi olan kadınlar, eşinin başka bir kadın ile ilişkisi olmayan kadınlara gore yaklaşık 10 kat daha fazla boşanma riskine sahiptir. Eşinin başka bir kadın ile ilişkisi olup olmadığına cevep vermeyen kadınlar ise eşinin başka bir kadın ile ilişkisi olmayan kadınlara göre yaklaşık 4 kat daha fazla boşanma riskine sahiptir. Eşi tarafından ailesi ile görüşmesi kısıtlanan kadınlar ailesi ile görüşmesi kısıtlanmayan kadınlara göre yaklaşık 4 kat daha fazla boşanma riskine sahiptir. Son olarak, eşi alkol kullanan kadınların eşi alkol kullanmayan kadınlara göre yaklaşık 3 kat daha fazla boşanma riskine sahip olması da elde edilen sonuçlar arasında yer almaktadır.

Ülkemizde, boşanmaya ilişkin çalışmalarda boşanmanın etkileri ve nedenlerine ilişkin sosyolojik yorumlar birçok çalışmada görülmektedir. Ancak bu çalışmada amacımızın sosyolojik yorumlarda bulunmak değil boşanma süresini incelediğimiz verilerin yapısının kesikli yaşam modellere daha uygun olduğunu gösterebilmek olduğunu belirtmemiz gerekmektedir.

Sürekli yaşam modelleri sağlık bilimleri alanında kullanılan durdurulmuş veri yapısına daha uygun olabilirken, kesikli yaşam modelleri özellikle sosyal bilimler alanındaki durdurulmuş veri yapısına daha uygun sonuçlar ve yorumlar getirebilmektedir.

Kaynakça

Bergstrom, R., Edin, P-A, 1992, Time aggregation and the distributional shape of unemployment duration, Journal of Applied Econometrics, 7, 5-30.

Blossfeld, H., P., Mills, M., 2001, A Causal Approach to Interrelated Family Events: A Cross-national Comparison of Cohabitation, Nonmarital Conception, and Marriage, Special Issue on Longitudinal Methodology, Canadian Studies in Population, 28(2), 409- 437.

Box-Steffensmeier, J., Jones, B., 1997, Time Is of the Essence: Event History Models in Political Science, American Journal of Political Science, 41, 1414-1461.

Box-Steffensmeier, M., J., Jones, S. B., 2004, Event History Modelling, Cambridge Universtiy Press.

Collett, D., 1994, Modelling Survival Data in Medical Research, London, Chapman&Hall.

Cox, D.R., 1972, Regression models and life-tables, Journal of the Royal Statistical Society,Series B, 34, 187-220.

Crowley, J., Hu, M., 1977, Covariance Analysis of Heart Transplant Survival Data, Journal of the American Statistical Association, 72, 27–36.

Eleuteri, A., Aung, M.S.H., Taktak, A.F.G., Damato, B., Lisboa, P.J.G., 2007, Continuous and discrete time survival analysis: neural network approaches, Engineering in Medicine and Biology Society, 5420-5423.

Ergöçmen B., Üner S., Yiğit E., 2009, Türkiye'de Kadına Yönelik Şiddet, Ankara..

García-Moreno, C., Jansen, H. A. F. M., Ellsberg, M., Heise, L., Watts, C, 2005, WHO multi-country study on women's health and domestic violence against women: initial results on prevalance, health outcomes and women's responses, World Health Organisation.

Geskus, R.B., 2000, On the inclusion of prevalent cases in HIV/AIDS natural history studies through a marker-based estimate of time since seroconversion, Statistics in Medicine, 19, 1753–1769.

Hannan, M., Carroll, G.R., 1981, Dynamics of formal political structure, Sociological Rev, 46, 19-35.

Hess, W., Persson, M., Rubenbauer, S., Gertheiss, J., 2011, The Varying Effects of Distance on The Survival of Trade Flows, paper presented at European Trade Study Group (ETSG) 13th Annual Conference, Copenhagen, Denmark.

Jenkins, S., 2005, Survival Analysis, Institute for Social and Economic Research, University of Essex, Colchester, UK.

Kalbfleisch, JD, Prentice, RL., 1980, The Statistical Analysis of Failure Time Data, New York , Wiley.

Mills, M., Johnston, A. D., DiPrete, T. A., 2006, Globalization and Men's Job Mobility in the United States, Uncertainty and Men's Careers: An International Comparison, 328–62.

Olzak, S., 1989, Analysis of Events in the Study of Collective Action, Annual Review of Sociology, 15, 119–41.

Petersen, T, 1991, Time-aggregation bias in continuous-time hazard-rate models, Sociological Methodology,, Blackwell, Cambridge MA.

Petersen, T., Koput K.W., 1992, Time-aggregation hazard-rate models with covariates, Sociological Methods and Research, 21, 2551.

Pope, C. A., Thun, M., Namboodiri, M., Dockery, D., Evans, J., Speizer, F., and Heath, C., 1995, Particulate air pollution as a predictor of mortality in a prospective study of U.S. adults, American Journal Respiratory Critical Care Medicine, 151, 669–674.

Kleinbaum DG, Klein M. Survival Analysis: A Self-Learning Text, Second Edition, Springer-Verlag, New York, 2005.

Smith, T., Smith, B., 1972, Survival Analysis And The Application Of Cox's Proportional Hazards Modeling Using SAS, Department of Defense Center for Deployment Health Research, Naval Health Research Center, San Diego, CA.

Sueyoshi, G.T., Edin, P-A, 1995, A class of binary response models for grouped duration data, Journal of Applied Econometrics, 10, 411-43.

Türkiye İstatistik Kurumu (TÜİK), 2008, Türkiye'de Kadına Yönelik Aile İçi Şiddet Araştırması.

Usui, C., 1994, Welfare State Development in a World System Context: Event History Analysis of First Social Insurance Legislation, Cambridge University Press, 254-77.

Xie, H., McHugo, G., Sengupta, A., Drake, R. 2003, Using discrete-time analysis to examine patterns of remission from substance use disorder among persons with severe mental illness, Mental Health Services Research, 5, 55-64.

Yang, C.C., 2004, Bayesian Discrete Time Survival Analysis of Multivariate Reoccurrable Events: Surviving Early Depressive Moods, Journal of Research on Measurement & Statistics, 12, 1-18.