

MEZUNİYET SÜRELERİNİ ETKİLEYEN FAKTÖRLERİN YAŞAM ÇÖZÜMLEMESİ İLE İNCELENMESİ

Hatice IŞIK¹

Nihal ATA TUTKUN²

Durdu KARASOY³

ÖZ

Öz; Sağlık bilimlerinde kullanımı yaygın olan yaşam çözümlemesinin sosyal bilimlerdeki kullanımı son yıllarda önem kazanmıştır. İşletme, ekonomi, eğitim gibi alanlarda uygulamaları görülmektedir. Eğitim alanında öğrencilerin mezuniyet süreleri ya da belli bir konuyu kavrama süreleri gibi konularda yaşam çözümlemesi yöntemlerinden yararlanılmaktadır. Bu çalışmada Hacettepe Üniversitesi İstatistik Bölümü'nden 1967 yılından 2015 yılına kadar mezun olan 3182 öğrencinin mezuniyet süreleri ve etkilediği düşünülen faktörler incelenmiştir. Normal süresi içinde eğitimini tamamlamayanlar başarısız gözlem olarak nitelendirilmiştir. Kaplan-Meier yaşam olasılıkları ve Tarone-Ware testi sonuçları elde edilmiştir. Yarı parametrik bir yöntem olan Cox regresyon modeli uygulanmış, ancak modelin varsayımı olan orantılı tehlikeler varsayımı sağlanmadığından kullanılamamıştır. Bu modele alternatif olan hızlandırılmış başarısızlık süresi modelleri ve sonrasında da zayıflık modeli uygulanmıştır. Veri için en uygun modelin zayıflık terimi içeren log-lojistik hızlandırılmış başarısızlık süresi modeli olduğu görülmüştür. Bu modele göre akademik ortalama (p -değeri=0.007), İngilizce hazırlık (p -değeri=0.003) ve tercih sırası (p -değeri=0.001) değişkenleri mezuniyet sürelerini etkileyen önemli faktörler olarak belirlenmiştir.

Anahtar Kelimeler: Cox regresyon modeli, durdurma, hızlandırılmış başarısızlık süresi modelleri, mezuniyet süresi, zayıflık.

INVESTIGATION OF THE FACTORS AFFECTING GRADUATION TIME THROUGH SURVIVAL ANALYSIS

ABSTRACT

The use of survival analysis, a popular method in health sciences, has recently attracted scholarly attention in the context of social sciences. Studies based on survival analyses continue to proliferate in

¹ Hacettepe Üniversitesi Fen Fakültesi İstatistik Bölümü Arş. Gör. Hatice Işık. mail: haticeyeniay@hacettepe.edu.tr

² Hacettepe Üniversitesi Fen Fakültesi İstatistik Bölümü Doç. Dr. Nihal Ata Tutkun. mail: nihalata@hacettepe.edu.tr

³ Hacettepe Üniversitesi Fen Fakültesi İstatistik Bölümü Prof. Dr. Durdu Karasoy. mail: durdu@hacettepe.edu.tr

fields such as business administration, economics, and education. In the field of education, survival analysis is used to investigate topics such as graduation time or the period of comprehension of a certain subject. In this study, graduation time and factors that are likely to affect the graduation time of 3182 students who graduated from Hacettepe University Department of Statistics from 1967 to 2015 are investigated. Those who did not complete their education within a regular time period were considered as failed observations. The results of Kaplan-Meier survival probabilities and Tarone-Ware test were obtained. A semi-parametric method, the Cox regression model, was used but it proved to be ineffective because of the violation of a basic assumption known as proportional hazard assumption. Instead, an alternative model, the accelerated failure time model and then the frailty model were introduced. The most saturated model for the data was found to be log-logistic accelerated failure time model with frailty. According to this model, academic average (p -value = 0.007), English preparatory education (p -value = 0.003) and order of preference (p -value = 0.001) were identified as significant factors affecting graduation times.

Keywords: *Cox regression model, censoring, accelerated failure time models, graduation time, frailty.*

1.GİRİŞ

Öğrencilerin normal eğitim süreleri içinde mezun olmaları ya da okulu bırakmaları birçok araştırmaya konu olmaktadır. Sprady (1970), Tinto (1975), Bean (1980), Terenzini ve Pascarella (1980) lisans öğrencilerinin bir dönem ya da bir yıl sonra okuldan atılıp atılmamalarını, dört yıl sonra mezun olup olmamalarını ya da eğitimlerine devam edip etmemelerini inceleyen modelleri araştırmışlardır. Bu konudaki çalışmalar özellikle Amerika Birleşik Devletleri'ndeki eğitim sistemini ele alarak incelenmiştir (Levin, 1972; Tinto, 1982; Caterall, 1987; Natriello, 1987; Weiss ve diğ., 1989; Porter, 1990; Ziolkowski, 1990). Amerika Birleşik Devletleri dışında da birçok ülkede benzer konuda çalışmalar (Kanada (Watson, 1974), İngiltere (Buckley ve Hooley, 1988; Johnes ve Taylor, 1989; Bennett, 2003), Almanya (Blossfeld, 1990)) yapılmıştır. İlk yapılan çalışmalarda korelasyon ya da regresyon çözümlemesi kullanılarak incelemeler yapılmıştır. Yaşam çözümlemesi yöntemleri ise ilk olarak Civian (1990)'ın çalışmasında kullanılmıştır. Bu çalışmada Harvard Üniversitesi Eğitim Bilimleri öğrencilerinin doktoralarını tamamlama

süreleri yaşam çözümlemesi yöntemleri ile incelemiştir. Daha sonra Willett ve Singer (1991), öğrencilerin okula devam süreleri ve öğretmenlerin çalıştıkları okuldaki hizmet süreleri gibi konuları incelemek için yaşam çözümlemesi yöntemlerinin uygun olduğunu belirtmişlerdir. DesJardins, et. al. (1994) çalışmalarında öğrencilerin demografik özelliklerini, başarı durumlarını ve maddi durumlarını içeren bir yaşam modelini incelemiştir. Zwick ve Sklar (2005), üniversite öğrencilerinin mezun olma olasılıklarını kesikli yaşam süresi modelleri ile değerlendirmişlerdir. Radcliffe, et al. (2006), spor bilimleri fakültesi öğrencilerinin okuldan mezun olma durumlarını yaşam çözümlemesinden yararlanarak incelemiştir.

Eğitim bilimleri alanında yaşam çözümlemesinin kullanımına ilişkin çalışmalar oldukça azdır. Türkiye’de ise bu konuda uygulamalı bir çalışma yapılmamıştır. Bu çalışmada, Hacettepe Üniversitesi İstatistik Bölümü öğrencilerinin mezuniyet sürelerini etkileyen faktörleri belirlemek için yaşam çözümlemesi yöntemleri kullanılmıştır. Çalışmanın ikinci bölümünde yaşam çözümlemesi ile ilgili kısaca bilgiler verilmiştir. Üçüncü bölümde yöntem açıklanmış ve dördüncü bölümde bulgular verilmiştir. Beşinci bölümde ise sonuçlar yorumlanarak öneriler sunulmuştur.

2. YAŞAM ÇÖZÜMLEMESİ

Yaşam çözümlemesi, tanımlanan herhangi bir olayın ortaya çıkmasına kadar geçen sürenin çözümlenmesinde kullanılan istatistiksel teknikler bütünüdür. Olayın gerçekleşmesi başarısızlık olarak tanımlanmaktadır. Başarısızlık süresi; bir birimin belirli bir başlangıç zamanı ile ilgilenilen olayın gerçekleşmesi arasında geçen süreye denir ve genellikle T ile gösterilir. Olay başarısızlık olarak ifade edilir. Yaşam çözümlemesinde temel olan, gözlenen başarısızlık sürelerinin incelenmesidir. Her bir birime ait yaşam süresi T, tanımı gereği sürekli ve pozitif bir değere sahiptir.

Yaşam çözümlemesi ile yapılan araştırmalarda bazı birimler gözlem süresi içinde ilgilenilen olay ile karşılaşmayabilir. Herhangi bir nedenden dolayı verinin izlenmesi yapılamamış ya da çeşitli nedenlerle gözlem dışı bırakılmış olabilir. Bu veriler yaşam çözümlemesinde, “durdurulmuş (censored)” gözlemler olarak ifade edilir. Yaşam çözümlemesinin klasik istatistiksel yöntemlerden farkı durdurulmuş gözlemleri içeren bir çözümleme yöntemi olmasıdır (Cox, 1972; Klein ve Moeschberger, 1997; Collett, 2003; Lee ve Wang, 2003, Kleinbaum ve Klein, 2005).

Yaşam çözümlemesinde kullanılan üç temel fonksiyon bulunmaktadır. Bunlar; olasılık yoğunluk fonksiyonu, yaşam fonksiyonu ve tehlike (hazard) fonksiyonudur. Bu fonksiyonlar birbiri ile ilişkili olan fonksiyonlardır (Cox, 1972).

Yaşam süresi, T ile gösterilen rastgele bir değişkendir. Bir bireyin belli bir t zamanından daha fazla yaşaması olasılığına yaşam fonksiyonu denir ve

$$S(t) = P(T > t) = \int_t^{\infty} f(x)dx, \quad 0 < t < \infty$$

biçiminde ifade edilir. Yaşam fonksiyonu monoton azalan soldan sürekli bir fonksiyondur ve t=0 iken S(0)=1, t=∞ iken S(∞) = 0 olur.

Tehlike fonksiyonu h(t), t zamanına kadar yaşayan bir birimin (t+Δt) zamanına kadar yaşamının sona ermesi riskidir. Birimin ilgilenilen özellik bakımından başarısızlık eğiliminin bir ölçüsüdür. Tehlike fonksiyonu;

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t \mid T > t)}{\Delta t} = \frac{f(t)}{S(t)}$$

biçimindedir. Bu fonksiyon bir olasılık fonksiyonu olmadığından zamana göre artabilir, azalabilir veya sabit kalabilir (Cox, 1972; Lee ve Wang, 2003).

2.1. Kaplan-Meier tahmini

Kaplan-Meier (KM) tahmini, yaşam fonksiyonunu tahmin etmek için en çok kullanılan parametrik olmayan yöntemlerden biridir. Bu yöntemle yaşam fonksiyonunun tahmini

$$\hat{S}_{KM}(t) = \prod_{j=1}^k \left(\frac{n_j - d_j}{n_j} \right) \quad k \leq n, \quad t_{(j)} \leq t < t_{(j+1)}$$

biçiminde tanımlanır. Burada,

d_j : t_j 'deki başarısızlıkların sayısını,

n_j : t_j 'de riskte olan birimlerin sayısını, yani t_j 'den hemen önce durdurulmamış ve yaşayan birimlerin sayısını,

k: sıralı başarısızlık süresi sayısını,

n: toplam birim sayısını

gösterir (Fleming ve Harrington, 1991; Leton ve Zuluaga, 2001).

KM tahmini ile elde edilen iki ya da daha fazla yaşam olasılığının karşılaştırılmasında parametrik olmayan testler kullanılmaktadır. Bu testler, orantılı tehlikeler varsayımı sağlandığı durumda log-rank testi ve sağlanmadığı durumda ise ağırlıklandırılmış log-rank testlerinden

Breslow, Tarone-Ware v.b. testleridir (Klein ve Moesberger, 1997; Kleinbaum ve Klein, 2005). Ağırlıklandırılmış log-rank testleri, her bir durdurulmamış gözlem süresi için tanımlanan w_j ağırlıklarına dayanır. Bu testler ki-kare testine dayanmaktadır ve test istatistiği diğer ki-kare testlerinde olduğu gibi sonuç değişkeninin kategorileri üzerinden gözlenen değerler (d_{ij}) ile beklenen değerler (e_{ij}) arasındaki fark alınarak hesaplanmaktadır ve

$$U = \sum_{j=1}^k w_j (d_{ij} - e_{ij}) = \sum_{j=1}^k w_j (d_{ij} - d_j \frac{n_{ij}}{n_j})$$

$$E(U)=0$$

$$V(U) = \sum_{j=1}^k w_j^2 V(d_{ij}) = \sum_{j=1}^k w_j^2 v_{ij} = \sum_{j=1}^k w_j^2 \frac{n_{ij} n_{2j} d_j (n_j - d_j)}{n_j^2 (n_j - 1)}$$

$$\chi^2 = \frac{U^2}{V(U)} \sim \chi_1^2$$

biçiminde ifade edilir. Burada,

d_j : t_j 'deki başarısızlıkların sayısını,

d_{ij} : t_j 'deki 1. gruba ait başarısızlıkların sayısını,

n_j : t_j 'de riskte olan birimlerin sayısını, yani t_j 'den hemen önce durdurulmamış ve yaşayan birimlerin sayısını,

n_{1j} : 1. grupta t_j 'de riskte olan birimlerin sayısını, yani t_j 'den hemen önce durdurulmamış ve yaşayan birimlerin sayısını,

w_j : t_j 'deki ağırlığı,

göstermektedir. Ağırlıklandırılmış log-rank testleri $w_j=1$ olması durumunda log-rank testine dönüşmektedir. Literatürde farklı ağırlıklar önerilmiştir, ancak bu testlerden $w_j = \sqrt{n_j}$ olarak tanımlayan Tarone-Ware testi daha güçlüdür (Kleinbaum ve Klein, 2005).

2.2. Cox regresyon modeli

Cox regresyon modeli, yaşam verileri için en çok kullanılan yarı parametrik bir regresyon modelidir (Cox, 1972). Cox orantılı tehlikeler modeli olarak da ifade edilir. Bu model, yaşam sürelerinin olasılık dağılımının belirli bir biçimi olmaması durumunda parametrik regresyon modellerine göre daha avantajlıdır.. Cox regresyon modeli aşağıdaki gibidir:

$$h(t, \mathbf{x}) = h_0(t) \exp(\boldsymbol{\beta}' \mathbf{x})$$

Burada, β regresyon katsayıları vektörü, x açıklayıcı değişkenler vektörü ve $h_0(t)$ ise temel tehlike fonksiyonu olarak tanımlanmaktadır.

Cox regresyon çözümlemesinde değişken düzeylerinden biri referans kategorisi olarak alınmakta ve değişken düzeylerinin yorumlanması buna göre yapılmaktadır. β parametresinin pozitif değer olması bu düzeyin referans kategorisine göre daha fazla riskli olduğunu, β parametresinin negatif değer olması ise bu düzeyin referans kategorisine göre daha az riskli olduğunu göstermektedir. Tehlike oranı olan $\exp(\beta)$ değeri ise önemli bulunan düzeyin, referans kategorisine göre kaç kat daha riskli olduğu yorumunu getirmektedir.

Cox regresyon modelinin kullanılabilmesi için orantılı tehlikeler varsayımının sağlanması gerekmektedir. Tehlike oranı, bir birimin tehlike fonksiyonunun diğer birimin tehlike fonksiyonuna oranı olarak tanımlanmaktadır. Tehlike oranının zamana karşı sabit olması durumunda orantılı tehlikeler varsayımı sağlanmaktadır. Bu varsayım; grafiksel yöntemlerle, istatistiksel testlerle (Schoenfeld artıkları) ya da zamana bağlı değişken yaklaşımı ile incelenebilir. Bunlardan en çok kullanılanı ise Grambsch ve Therneau (1994) tarafında geliştirilen Schoenfeld (1982) artıklarına dayanan test istatistiğidir. Belirli bir değişken için Schoenfeld artıkları ile bireylerin yaşam sürelerinin rankı (yaşam süreleri) arasındaki korelasyon kullanılarak orantılı tehlikeler varsayımı incelenmiştir. Bu teste göre, orantılı tehlike varsayımının sağlanması için korelasyonun sıfıra yakın olması beklenmektedir. Yokluk hipotezi "orantılı tehlike varsayımı sağlanmaktadır" biçimindedir. Bu test istatistiği, orantılı tehlike varsayımının incelenmesi için kullanılan grafiksel yöntemlere göre daha nesnel bir kriter sağlamaktadır (Klein ve Moesberger, 1997; Kleinbaum ve Klein, 2005)

2.3. Hızlandırılmış yaşam süresi modelleri

Veri kümesi için belirli bir olasılık dağılımı varsayımı geçerli ise, bu durumda parametrik regresyon modelleri Cox regresyon modeline göre daha etkili parametre tahminleri verir (Nardi ve Schemper, 2003). Yaşam çözümlemesinde parametrik orantılı tehlikeler modelleri uygulanabilir olmasına rağmen, bu modeller ile kullanılacak az sayıda olasılık dağılımı olduğundan hızlandırılmış başarısızlık süresi (HBS) modelleri kullanılır. HBS modeline karşılık gelen logaritmik doğrusal form;

$$\log T_i = \mu + \alpha_1 X_{1i} + \alpha_2 X_{2i} + \dots + \alpha_p X_{pi} + \sigma \varepsilon_i$$

biçimindedir. Burada, T yaşam süresini, X açıklayıcı değişkenleri, μ kesişim, σ ölçek parametresini ve ε_i hata terimini gösterir. HBS modellerinde, tehlike yerine yaşam süresi üzerine açıklayıcı değişkenlerin doğrudan etkileri ölçüldüğünden açıklayıcı değişkenlerin yaşam süresi üzerine etkisini yorumlamak daha kolaydır (Wei, 1992; Orbe, et al., 2002; Qui, 2009; Karasoy ve Sezayi, 2014)

2.4. Zayıflık modelleri

Zayıflık (frailty) modeli ile ilgili ilk çalışmalar Vaupel ve diğ. (1979) tarafından yapılmış ve “zayıflık” kavramı özelliklere sahip birimler arasında yaşam sürelerindeki farklılıkları açıklamak için kullanılmıştır. Zayıflık modeli, birimler arasındaki heterojenliği açıklamak için ölçülemeyen rasgele etkiyi tehlike fonksiyonuna dahil eden bir modeldir. $h(t)$, t zamanındaki tehlike fonksiyonu ve α ölçülemeyen açıklayıcı değişken (zayıflık) olmak üzere zayıflık modeli, $h(t / \alpha) = \alpha h(t)$

biçiminde ifade edilmektedir (Aalen, 1994; Hougaard, 1995; Gutierrez, 2002; O’Quigley ve Stare, 2002). Zayıflık teriminin (α) birim ortalamaya ve sonlu varyansa (θ) sahip olduğu varsayılmaktadır. Zayıflık modeli ile ilgili çalışmaların birçoğu zayıflık teriminin dağılımının seçimini içermektedir. Çalışmalarda zayıflık terimi için en çok kullanılan dağılım, Gamma ya da ters Gaussian dağılımlarıdır. $\alpha > 1$ 'e sahip olan birimlerin açıklayıcı değişkenler tarafından açıklanamayan nedenlerden dolayı artan başarısızlık riskine ve $\alpha < 1$ 'e sahip olan birimlerin ise daha uzun yaşama eğilimine sahip olduğu belirtilir (Gutierrez, 2002).

3. YÖNTEM

3.1. Araştırma Verileri

Hacettepe Üniversitesi İstatistik Bölümü mezunlarının mezuniyet sürelerini etkileyen faktörler yaşam çözümlemesi yöntemleri ile incelenmiştir. Hacettepe Üniversitesi İstatistik Bölümü 1967 yılında kurulmuştur. Kurulduğu yıldan 2015 yılına kadar mezun olan 3182 mezun çalışmaya alınmıştır. Eksik tutulan ve doğruluğundan şüphe duyulan kayıtlar, yatay ve dikey geçişli öğrencilerin kayıtları ve af ile dönen ancak bölümde bulunduğu süre tam olarak bilinmeyen toplam 295 gözlem çalışma dışı bırakılmış ve 2947 gözlem üzerinden analizler yapılmıştır. Çalışmada SPSS ve

STATA istatistiksel paket programlarından yararlanılarak analizler yapılmıştır.

3.2. Verilerin Analizi

Araştırmada verilerin istatistiksel çözümlemesinde yaşam çözümlemesi yöntemleri kullanılmıştır. Bu çalışmada H.Ü. İstatistik Bölümü öğrencilerinin lisans eğitimlerini normal süresinde tamamlamayanlar başarısız, normal süresinde tamamlayanlar ise durdurulmuş gözlem olarak ele alınmıştır. Buna göre durdurma oranı %57.9 olarak bulunmuştur. Çalışmada öncelikle KM yaşam olasılıkları, yaşam eğrileri elde edilmiş ve yaşam olasılıklarını karşılaştırmak için Tarone-Ware testi uygulanmıştır. Orantılı tehlikeler varsayımı incelenmiş ve bu varsayım sağlanmadığından Cox regresyon modeli uygulanmamış, HBS modelleri elde edilmiştir. HBS modellerinden en uygun olan model belirlendikten sonra da zayıflık teriminin modele dahil edilip edilemeyeceği incelenmiş ve en uygun model sonuçları yorumlanmıştır.

4. BULGULAR

Çalışmada kullanılan nitel değişkenlere ilişkin bilgiler Tablo 1’de, nicel değişken olan akademik ortalama ve mezuniyet süresine ait tanımlayıcı istatistikler ise Tablo 2’de verilmiştir.

Tablo 1: Nitel değişkenlere ilişkin bilgiler

Değişken	Düzeyler	n	%
İngilizce Hazırlık	1. Okumadı 2. Okudu <i>Toplam</i>	1787 1160 2947	60.6 39.4
Cinsiyet	1. Kadın 2. Erkek <i>Toplam</i>	1494 1312 2806	53.2 46.8
İl	1. Ankara’dan 2. Ankara dışından <i>Toplam</i>	794 411 1205	65.9 34.1
Tercih Sırası	1. <=5 2. >5 <i>Toplam</i>	438 767 1205	36.3 63.7

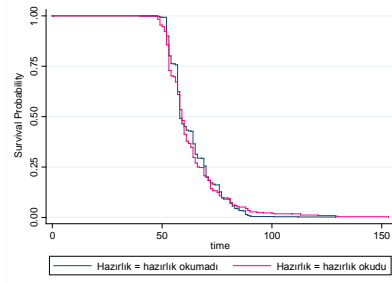
Mezuniyet Sürelerini Etkileyen Faktörlerin Yaşam Çözümlemesi İle İncelenmesi

Durum	0. Normal süresinde mezun olan	1706	57.9
	1. Normal süresinden daha uzun sürede mezun olan	1241	42.1
	<i>Toplam</i>	2947	

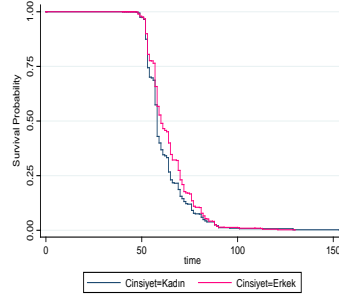
Tablo 2: Nicel değişkenler için tanımlayıcı istatistikler

Değişken	N	Min. Değer	Max. Değer	Ortalama	Std. sapma
Akademik ortalama	2947	1.83	3.90	2.60	0.352
Mezuniyet süresi (ay)	2947	40	153	52.78	11.357

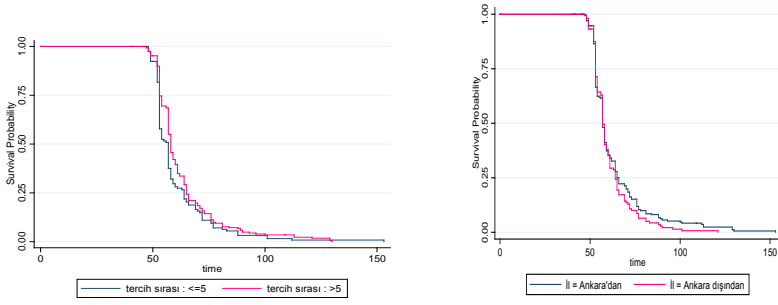
Çalışmada yer alan değişkenler için yaşam olasılıklarını gösteren KM yaşam eğrileri Şekil 1’de verilmiştir. Orantılı tehlikeler varsayımı sağlanmadığından yaşam olasılıklarını karşılaştırmak için Tarone-Ware testi ile incelendiğinde cinsiyet (p-değeri=0.001) ve tercih sırası (p-değeri=0.003) değişkenlerinin düzeyleri arasında anlamlı fark olduğu, hazırlık (p-değeri=0.09) ve il (p-değeri=0.619) değişkenlerinin düzeyleri arasında fark olmadığı %95 güven düzeyinde söylenebilir. Kadınlar erkeklere göre daha kısa sürede mezun olmuştur. Tercih sırası 5’den küçük olanların mezuniyet süresi, 5’den büyük olanlara göre daha kısadır.



(a) Hazırlık okuma



(b) Cinsiyet



(c) Tercih sırası

(d) İl

Şekil 1. Kaplan-Meier yaşam eğrileri

Orantılı tehlikeler varsayımının sağlanıp sağlanmadığı Schoenfeld (1982) tarafından önerilen Schoenfeld artıkları kullanılarak test edilmiş ve hazırlık (p-değeri=0.0014), cinsiyet (p-değeri=0.0172) ve tercih sırası (p-değeri=0.0216) değişkenleri için orantılı tehlikeler varsayımının sağlanmadığı görülmüştür.

Orantılı tehlikeler varsayımı sağlanmadığından Cox regresyon modeli yerine HBS modelleri uygulanmıştır. En uygun HBS modelini belirlemek için yaşam süresinin dağılımını incelemek gerekmektedir. Dağılıma uygunluğu değerlendirmek için Kolmogorov-Smirnov testinden yararlanılmış ve birinci sırada verilerin log-lojistik dağılıma uygun olduğu sonucuna varılmıştır. Log-lojistik HBS modeli verilere uygulanmıştır. Zayıflık teriminin modele dahil edilip edilmemesine karar vermek için olabilirlik oran testi kullanılmış ve ki-kare değeri 188.06 ve p-değeri 0.000 bulunmuştur. Buradan zayıflık teriminin model üzerinde etkisi olduğu ve modele dahil edilmesi gerektiği sonucuna ulaşılmıştır. Buna göre log-lojistik HBS modeli yerine zayıflık terimi içeren log-lojistik HBS modelinin bu veri için kullanılmasının daha uygun olduğu söylenebilmektedir.

Buna göre zayıflık terimi içeren log-lojistik regresyon modeli sonuçları Tablo 3'de verilmiştir.

Tablo 3: Zayıflık terimi içeren log-lojistik regresyon modeli sonuçları

Değişken	Tahmin	Std. Hata	%95 Güven Aralığı		p-değeri
			Alt Sınır	Üst Sınır	
Akademik ortalama	-0.032	0.015	-0.061	-0.004	0.026*

Mezuniyet Sürelerini Etkileyen Faktörlerin Yaşam Çözümlemesi İle İncelenmesi

Hazırlık=Okumadı	0.027	0.009	0.010	0.045	0.002*
Cinsiyet=Kadın	-0.011	0.009	-0.029	0.007	0.216
Tercih sırası <=5	-0.029	0.009	-0.046	-0.011	0.001*
İl=Ankara'dan	-0.004	0.009	-0.022	0.014	0.649
Sabit	4.108	0.036	4.038	4.177	0.000

Zayıflık terimlerinin yorumlanmasında zayıflık teriminin (θ) heterojenliğin bir ölçüsü olduğu bilgisi altında Pankratz ve diğ. (2005) yaklaşımına göre zayıflık teriminin karekökünün üstel ifadesi $\exp(\sqrt{\theta})$ rastgele etkinin yani zayıflığın etkisini göstermektedir. Zayıflık terimi içeren log-lojistik HBS model için $\exp(\sqrt{\theta})$ değeri $\exp(\sqrt{0.89}) = 2.57$ olarak elde edilmiştir. Bu değer herhangi bir öğrencinin, çalışma sürecinde seçilen bir noktada, tüm çalışma grubuna ait riskten ± 2.57 kata kadar daha fazla ya da az risk taşıyabileceğini ifade etmektedir.

Tablo 3'de verilen model incelendiğinde akademik ortalama, hazırlık ve tercih sırası değişkenlerinin %95 güven düzeyinde anlamlı olduğu görülmüştür. Akademik ortalama arttıkça, öğrencilerin mezuniyet sürelerinin kısaldığı görülmüştür. Hazırlık okumayanların hazırlık okuyanlara göre 1.03 ($\exp(0.027) \cong 1.03$) kat daha uzun sürede mezun olduğu söylenebilir. Tercih sırası 5 ve daha az olanların, tercih sırası 5'ten büyük olanlara göre ($\exp(0.029) \cong 1.03$) 1.03 kat daha kısa sürede mezun olduğu söylenebilir.

Modelde adımsal (stepwise) yöntem kullanılarak değişken seçimi yapıldığında akademik ortalama, hazırlık ve tercih sırası değişkenlerini içeren modelin veri için en uygun model olduğu sonucuna varılmıştır. Elde edilen sonuçlar Tablo 4'de verilmiştir.

Tablo 4: Değişken seçimi yapıldığında zayıflık terimi içeren log-lojistik regresyon modeli sonuçları

Değişken	Tahmin	Std. Hata	%95 Güven Aralığı		p-değeri
			Alt Sınır	Üst Sınır	
Akademik ortalama	-0.037	0.013	-0.065	-0.01	0.007*
Hazırlık=Okumadı	0.027	0.008	0.0093	0.044	0.003*
Tercih sırası <=5	-0.028	0.009	-0.046	-0.011	0.001*

Sabit	4.110	0.036	4.042	4.181	0.000
-------	-------	-------	-------	-------	-------

Tablo 4’de elde edilen sonuçlara göre,

- Akademik ortalama arttıkça, mezuniyet süresinin kısaldığı,
- Hazırlık okumayanların hazırlık okuyanlara göre 1.03 kat daha uzun sürede mezun olduğu,
- Tercih sırası 5 ve daha az olanların, tercih sırası 5’ten büyük olanlara göre 1.03 kat daha kısa sürede mezun olduğu söylenebilir.

5. SONUÇ ve ÖNERİLER

Bu çalışmada eğitim bilimleri alanında lisans öğrencilerinin mezuniyet süreleri yaşam çözümlemesi yöntemleri ile incelenmiştir. Hacettepe Üniversitesi Fen Fakültesi İstatistik Bölümü öğrencilerine ait 2947 mezuna ait bilgiler kullanılarak mezuniyet süreleri incelenmiş ve mezuniyet sürelerini etkileyen faktörler belirlenmeye çalışılmıştır. Bu amaçla, akademik ortalama, cinsiyet, hazırlık okuma, tercih sırası ve mezunların lise eğitimlerini tamamladıkları il değişkenleri ele alınmıştır. Öncelikle değişkenlere ilişkin bilgiler verilmiş, Kaplan-Meier yaşam eğrileri çizilmiş ve kategorik değişken düzeyleri arasında zamanında mezun olma (yaşam) olasılıkları açısından fark olup olmadığını incelemek için Tarone-Ware testine ait sonuçlar elde edilmiştir. %95 güven düzeyinde cinsiyet (p-değeri=0.001) ve tercih sırası (p-değeri=0.003) değişkenlerinin düzeyleri arasında anlamlı fark bulunurken il ve hazırlık değişkenleri için yaşam olasılıkları arasında fark olmadığı görülmüştür. Yaşam modelleri içinde en yaygın kullanıma sahip olan Cox regresyon modeli orantılı tehlikeler varsayımının sağlanmamasından dolayı kullanılamamıştır. Yaşam çözümlemesinde orantılı tehlikeler varsayımı sağlanmadığı durumda kullanılan HBS modellerinden log-lojistik HBS modelinin veri için uygun olduğu görülmüştür. Daha sonra ise çalışmaya dâhil edilen bireyler arasındaki heterojenliğin önemli olup olmadığını araştırmak için zayıflık modeli kullanılmış ve zayıflık teriminin modelde bulunmasının %95 güven düzeyinde anlamlı olduğuna karar verilmiştir. Son olarak, klasik regresyon modellerinde de kullanılan adimsal değişken seçim yöntemi kullanılarak en uygun log-lojistik HBS modelinin akademik ortalama (p-değeri=0.007), hazırlık (p-değeri=0.003) ve tercih sırası (p-değeri=0.001) değişkenlerini içeren model olduğu görülmüştür. Akademik ortalama

arttıkça, mezuniyet süresinin kısaldığı, hazırlık okumayanların hazırlık okuyanlara göre 1.03 kat daha uzun sürede mezun olduğu, tercih sırası 5 ve daha az olanların, tercih sırası 5'ten büyük olanlara göre 1.03 kat daha kısa sürede mezun olduğu sonucuna ulaşılmıştır.

Gelecek çalışmalarda veri kümesi genişletilerek Hacettepe Üniversitesi Fen Fakültesi öğrencilerinin tamamı ya da üniversitede tüm öğrenciler çalışmaya dâhil edilebilir. Mezuniyet sürelerine etki eden faktörleri belirlemek için çalışmaya dâhil edilen değişkenler çeşitlendirilebilir. Eğitim-öğretim yılı başında lisans eğitimlerine başlayan öğrencilere anket uygulanarak elde edilen bilgiler arttırılabilir ve öğrenciler eğitim süreleri boyunca gözlemlenip eğitim süreleri sonundaki durumları araştırılabilir. Ayrıca ilköğretim, lise, önlisans, lisans ya da lisansüstü programlarda da öğrencilerin mezuniyet süreleri ya da eğitimlerini bırakma eğilimleri gerek kayıtlardaki veriler gerekse anket çalışması ile elde edilen veriler kullanılarak yaşam çözümlemesi yöntemleri ile incelenebilir.

6. KAYNAKLAR

AALEN, O.O. (1994), "Effects of Frailty in Survival Analysis", *Statistical, Methods in Medical Research*, S:3, ss.227-243.

BEAN, J.P. (1980), "Dropouts and Turnover: The synthesis and Test of a Causal Model of Student Attrition", *Research in Higher Education*, S:12, ss.155-187.

BENNETT, R. (2003), "Determinants of Undergraduate Student Drop Out Rates in a University Business Studies Department", *Journal of Further and Higher Education*, 27(2), 123-140.

BLOSSFELD, H. P. (1990), "Changes in educational careers in the Federal Republic of Germany", *Sociology of Education*, 63, 165-177.

BUCKLEY, P.J., ve HOOLEY, G.J. (1988), "The Non-Completion of Doctoral Research in Management: Symptoms, Causes and Cures", *Educational Research*, S:30, ss.110-119.

CATERALL, J. S. (1987), "On the Social Costs of Dropping Out of School", *The High School Journal*, S:71, ss.19-30.

CIVIAN, J. T. (1990), *Using Proportional Hazard Models to Examine Time to Doctorate*, (Yayınlanmamış doktora tezi), Harvard University, Cambridge, MA.

COLLET, D. (2003), *Modelling Life Data in Medical Research*, London: Chapman&Hall.

COX, D.R. (1972), "Regression models and life-tables", *Journal of the Royal Statistical Society, Series B*, S:34(2), ss.187-220.

DESJARDINS, S. L., AHLBURG, D., ve MCCALL, B. (1994), "Studying the Determinants of Student Stopout: Identifying "True" from Spurious Time-Varying Effects", Paper presented at the 34th Annual Forum of the Association for Institutional Research, New Orleans, LA.

FLEMING, T.R. ve HARRINGTON, D.P. (1991), *Counting Processes and Survival Analysis*, New York: Wiley

GRAMBSCH, P.M. ve THERNEAU, T.M. (1994), "Proportional hazards tests and diagnostics based on weighted residuals", *Biometrika*, S:81, ss. 515-526.

GUTIERREZ, R.G. (2002), "Parametric Frailty and Shared Frailty Models", *The Stata Journal*, S:2(1), ss.22-44.

HOUGAARD, P. (1995), "Frailty Models for Survival Data", *Lifetime Data Analysis*, S:1, ss.255-273.

JOHNES, J. ve TAYLOR, J. (1989), "Undergraduate Non-Completion Rates: Differences Between UK Universities", *Higher Education*, S:18, ss.209-225.

KARASOY, D. ve SEZAYİ, S. (2014), "Yaşam çözümlemesinde hızlandırılmış başarısızlık süresi modelleri ve bir uygulama", *Suleyman Demirel University Journal of Natural and Applied Science*, S:18(1), ss.1-7.

KLEIN, J.P. ve MOESCHBERGER, M.L. (1997), *Survival Analysis: Techniques for Censored and Truncated Data*, Springer-Verlag, New York.

KLEINBAUM, D.G. ve KLEIN, M. (2005), *Survival Analysis: A Self-Learning text*, New York, Springer.

LEE, E.T. ve Wang, J.W. (2003), *Statistical Methods for Survival Data Analysis*, New York: Wiley.

LETON, E. ve ZULUAGA, P. (2001), "Equivalence Between Score and Weighted Tests for Survival Curves," *Commun. Statist.-Theory Meth.*, S:30(4), ss.591-608.

LEVIN, H. M. (1972), "The costs to the nation of inadequate education" (Report to the Select Committee on Equal Educational Opportunity of the United States Senate, January). Washington, DC: U.S. Government Printing Office.

NARDI, M., ve SCHEMPER, M. (2003), "Comparing Cox and Parametric Models in Clinical Studies", *Statistics in Medicine*, S:22, ss.3597-3610.

NATRIELLO, G. (1987), *School Dropouts: Patterns and Policies*, New York: Teachers College Press.

O'QUIGLEY, J. ve STARE, J. (2002), "Proportional Hazards Models with Frailties and Random Effects", *Statistics in Medicine*, S:21, ss.3219-3233.

ORBE, J., FERREIRA, E. ve NUNEZ-ANTON, V. (2002), "Comparing Proportional Hazards and Accelerated Failure Time Models for Survival Analysis", *Statistics in Medicine*, S:21, ss.3493-3510.

QUI, J. (2009), "Comparison of proportional hazards and accelerated failure time models" University of Saskatchewan, Department of Mathematics and Statistics, Saskatoon, Saskatchewan, Canada.

PANKRATZ, V. S., DE ANDRADE, M. ve THERNAU, T. M. (2005), "Random Effects Cox Proportional Hazard Model: General Variance Components Methods for Time-To-Event Data", *Genetic Epidemiology*, S:28, ss.97-109.

PORTER, A. M. (1990), "Undergraduate Completion and Persistence at Four-Year Colleges and Universities", Reports-Research, Washington, DC: National Institute of Independent Colleges and Universities.

RADCLIFFE, P. M., HUESMAN, R.L. Jr. ve KELLOGG, J.P. (2006), "Identifying Students at Risk: Utilizing Survival Analysis to Study Student Athlete Attrition", Retrieved from the University of Minnesota Digital Conservancy, <http://purl.umn.edu/159769>.

SCHOENFELD, D. (1982), "Partial Residuals For The Proportional Hazards Model", *Biometrika*, S:69, ss.551-55.

SPADY, W. (1970), "Drop-Outs From Higher Education: An Interdisciplinary Review And Synthesis", *Interchange*, S:1, ss.64-85.

TARONE, R.E. ve WARE, J. (1977), "On distribution-free tests for equality of survival distributions", *Biometrika*, S:64(1), ss.156-160.

TERENZINI, P. T. ve PASCARELLA, E. T. (1980), "Towards The Validation of Tinto's Model of College Student Attrition: A Review Of Recent Studies", *Research in Higher Education*, S:12(3), ss.271-282

TINTO, V. (1975), "Dropout From Higher Education: A Theoretical Synthesis Of Recent Research", *Review of Educational Research*, S:45(1), ss.89-125.

TINTO, V. (1982), "Limits in Theory and Practice in Student Attrition", *Journal of Higher Education*, S:53, ss.687-700.

VAUPEL, J.W., MANTON, K. ve STALLARD, E. (1979), "The Impact of Heterogeneity in Individual Frailty On The Dynamics of Mortality", *Demography*, S:16, ss.439-454.

WATSON, C. (1974), "Focus on dropouts", Toronto, Canada: Ontario Institute for the Study of Education.

WEI, L.J. (1992), "The Accelerated Failure Time Model: A Useful Alternative To The Cox Regression Model In Survival Analysis", *Statistics in Medicine*, S:11, ss.1871-1879.

WEIS, L., Farrar, E., & Petrie, H. G. (1989), "Dropouts From School: Issues, Dilemmas, And Solutions", Albany, NY: State University of New York.

WILLETT, J.B. ve SINGER, J.D. (1991), "From Whether To When: New Methods For Studying Student Dropout And Teacher Attrition", *Review of Educational Research*, S:61, ss.407-450.

ZIOLKOWSKI, T. (1990), "The PhD Squid", *American Scholar*, S:59(2), ss.177-195.

ZWICK, R. ve SKLAR, J. (2005), "Predicting College Grades And Degree Completion Using High School Grades And Sat Scores: The Role Of Student Ethnicity And First Language", *American Educational Research Journal*, S:42, ss.439-464.