



Sigortacılık sektöründe yaşam çözümlemesi: Birikimli hayat sigortaları ile ilgili bir uygulama

Uğur Karabey

Hacettepe Üniversitesi
Aktüerya Bilimleri Bölümü
06800-Beytepe, Ankara, Türkiye.
ukarabey@hacettepe.edu.tr

Nihal Ata

Hacettepe Üniversitesi
İstatistik Bölümü
06550-Beytepe, Ankara, Türkiye
nihalata@hacettepe.edu.tr

Özet

Yaşam çözümlemesinde başarısızlık olarak adlandırılan olaylarla ilgilenilmektedir. Örneğin tıbbi çalışmalarda ölüm ya da hastalığın tekrar nüks etmesi başarısızlık olarak ele alınırken, sigortacılık sektöründe de poliçenin iptal ettirilmesi başarısızlık olarak nitelendirilebilmektedir. Bu çalışmada öncelikle Türkiye'deki sigorta faaliyetleri ile ilgili kısa bilgiler verilmiş, daha sonra yaşam çözümlemesi teorisi anlatılmıştır. Son olarak ise türetilen bir örneklem üzerinden yaşam çözümlemesinde kullanılan yarı parametrik ve parametrik regresyon modelleri, birikimli hayat sigortası poliçe sahiplerine ait örneklem uygulanarak poliçe iptallerini etkileyen faktörler incelenmiştir.

Anahtar sözcükler: Birikimli hayat sigortası; Gompertz; Hazard fonksiyonu; Parametrik dağılımlar; Yaşam modelleri.

Abstract

Survival analysis in insurance sector: An application to long-term life insurance

Survival analysis is interested in the events called as failure. While death or recurrency of illness called as failure in medical research, cancellation of policy can be taken as failure in insurance sector. In this study, firstly information about the activities of insurance sector in Turkey is given. Later survival analysis theory is explained. Finally semi-parametric and parametric regression models are applied to generated sample and factors that effect cancellation of policy are analysed.

Keywords: Long-term life insurance; Hazard function; Gompertz, parametric distributions; Survival models.

1. Giriş

Ülkemizde sigortacılık faaliyetleri, hayat ve hayat dışı dallar olmak üzere iki ana grupta toplanmıştır. Hayat dışı ana dalları sırasıyla; yangın, nakliyat, kaza, makina-montaj, dolu, hayvan hayat, hastalık, hukuksal koruma, ferdi kaza, kredi ve trafik iken hayat grubu yalnızca hayat ana dalından ibarettir[14].

Türkiye, 2005 yılında dünya prim üretiminde % 0,17 pay ile 88 ülke içerisinde 35. sırada yer almıştır. Prim üretiminin GSYİH'ye oranına göre yapılan sıralamada 69. sırada yer alan Türkiye, kişi başı prim üretiminde ise 61.'liğe yükselmiştir. Toplam prim üretimine göre Avrupa ülkeleri arasında bir değerlendirme yapıldığında ise Türkiye, % 0.41 pay ile 19. sırada yer almıştır.

Türkiye’de hayat dışı şirketlerin poliçe sayısı yıllar itibariyle düzenli bir artış göstermekte iken, hayat şirketlerinin poliçe sayısı son yıllarda azalmaktadır. Hayat dışı sigorta dallarında poliçe sayısı 2006 yılında % 24.76 oranında artış göstermişken, hayat dalındaki poliçe sayısı % 14.09 oranında azalmıştır[13]. Bu durumun sebeplerinden biri, bireysel emeklilik sisteminin 2003 yılında fiilen yürürlüğe girmesi sonucu hayat sigortası yaptırmak isteyenlerin emeklilik sistemine dahil olmayı tercih etmesi olarak görülmektedir. Diğer önemli bir etken ise poliçelerini iptal ettirmek isteyen poliçe sahipleri olarak görülmektedir. Prim üretimindeki azalmanın en önemli nedeni ise, ödeme düzenine uymayan poliçe sahipleri olarak görülmektedir. Poliçe sahiplerinin poliçelerini iptal ettirmesini etkileyen faktörlerin belirlenmesinde yaşam çözümlenmesi yöntemleri kullanılabilir. Bu çalışmada bir birikimli hayat sigortası sistemi üzerinde, yaşam çözümlenmesi yöntemlerinin uygulanabilirliği incelenmiştir.

2. Hayat sigortaları

Sigorta, konusu insan hayatı olan, genellikle uzun vadeli menfaat veya meblağ sigortalarıdır. Öncelikle, poliçe sahibinin sigorta süresi içerisinde ölmesi halinde, poliçeden yararlanan kişi veya sigortalının yasal varislerine toplu para ödenmesi amacıyla düzenlenen hayat poliçeleri, zaman içerisinde kişilerin emeklilik programlarına bir destek niteliği kazanmış son olarak da, yatırım amaçlı olarak kullanılmaya başlanmıştır. Hayat sigortaları iki ana başlık altında incelenebilir.

2.1. Risk ağırlıklı hayat sigortaları

Risk ağırlıklı hayat sigortalarında, sigorta süresi içinde (genellikle bir yıl), sigortalının yaşamını kaybetmesi durumunda belirlenen tazminat tutarı kanuni varislere ya da poliçede lehdar olarak belirlenen kişilere ödenmektedir.

2.2. Birikimli hayat sigortaları

Birikimli hayat sigortaları, sigorta süresi en az 10 yıl olan uzun dönemli sigortalardır. Yaşam kaybı, ferdi kaza ve maluliyet teminatlarının yanı sıra toplanan primlerden risk primi, genel gider ve komisyon kesintileri yapıldıktan sonra kalan tutar yatırıma yönlendirilmektedir. Birikimli hayat sigortaları iki temel güvence sağlamaktadır:

- Sigorta süresince prim ödeyen sigortalının, sigorta süresi dolmadan ölümü halinde poliçede belirlenen tazminat tutarına, kar paylı birikim tutarı eklenerek varislerine ya da poliçede lehdar olarak gösterilen kişilere toplu para ödenmektedir.
- Poliçe süresi içinde teminat altında olan herhangi bir riskin gerçekleşmemesi ve en az 10 yıl prim ödemek koşuluyla vade sonunda toplu bir para veya gelir ödemesi şeklinde poliçeden yararlanılabilmektedir[15].

3. Yaşam çözümlenmesi

Yaşam çözümlenmesi, makina parçalarının bozulması, hastalığın başlangıcı, depremler, trafik kazaları, borsada hisse senetlerinin düşmesi, doğum, ölüm, boşanma, terfi, emeklilik, işten çıkarma, vb. olayları incelemek için yararlı bir istatistiksel yöntemdir. Bu çözümlenme yöntemi, farklı alanlardaki uygulamalar için araştırmacılar tarafından farklı isimler ile adlandırılmaktadır. Örneğin, sosyolojide “olay tarihi çözümlenmesi (event history analysis)”, mühendislikte “güvenilirlik kuramı (reliability theory)” ya da “başarısızlık zamanı çözümlenmesi (failure time analysis)”, ekonomide “süreklilik çözümlenmesi (duration analysis)” ya da “geçiş çözümlenmesi (transition analysis)” ve klinik denemelerde “yaşam çözümlenmesi (survival analysis)” olarak adlandırılmaktadır.

Yaşam çözümlemesi, başarısızlık olarak adlandırılan bir nokta olayı ile ilgilenmektedir. Yaşayan bir organizmanın ya da cansız bir nesnenin belirli bir başlangıç zamanı ile başarısızlığı arasında geçen zamana “yaşam süresi” ya da “başarısızlık süresi” adı verilmektedir. Yaşam süresini etkileyen faktörler yarı parametrik ya da parametrik regresyon modelleri kullanılarak incelenebilmektedir. Yaşam çözümlemesinde kullanılan regresyon modellerinin diğer istatistiksel modellerden temel farkı durdurulmuş (censored) veri için tasarlanmış istatistiksel yöntemler bütünü olmasıdır. Gözlem süresi boyunca çalışmada yer alan birimlerin tamamı başarısızlık ile karşılaşmamış olabilir. Bu birimler yaşam çözümlemesinde durdurulmuş olarak nitelendirilmektedir[4]. Yaşam çözümlemesinin farklı alanlarda kullanımına ilişkin çalışmalar son yıllarda da literatürde yer almaktadır. Örneğin, Keiding v.d (1998)’de hayat dışı sigortalarda, Box-Steffensmeier v.d. (2003)’de uluslararası ilişkilerde ve Evrensel (2007)’de uluslararası banka krizi verilerinin analizinde yaşam çözümlemesi yöntemlerini incelemişlerdir[3,6,7]. Abone tabanlı ilişkilere sahip işletmeler için yaşam çözümlemesi müşteri profilini belirlemek, müşteri sadakatini ölçmek için kullanılmaktadır[9]. Burada ilgilenilen olay müşteri kaybıdır. Müşteri ilişkilerini iyi tanımlamış, bir başlangıç ve bitiş zamanına sahip işletmeler için yaşam çözümlemesi yöntemleri kolaylıkla uygulanabilir[12]. Abone tabanlı işletmelerde müşteri veri tabanı, hizmet ya da ürünün kullanımına başlanması ya da iptal edilmesi gibi birçok önemli olaya ait verileri içermektedir. Bu veriler, yaşam çözümlemesinde kullanılan hazard fonksiyonu, yaşam fonksiyonu ve regresyon modelleri yardımı ile incelenebilir ve müşteriyi kaybetmemek ve pazarlama stratejileri belirlemek amacıyla kullanılabilir.

3.1. Yaşam fonksiyonu ve Hazard fonksiyonu

Yaşam modeli, $P(T>t)$ olasılığı ile temsil edilen bir yaşam fonksiyonu $S(t)$ olarak tanımlanmaktadır. Yaşam fonksiyonu, T raslantı değişkeninin t 'den daha büyük olma olasılığıdır ve

$$S(t) = P(T > t) = \int_t^{\infty} f(x)dx \quad 0 < t < \infty \quad (1)$$

biçiminde gösterilmektedir. Yaşam fonksiyonunun $t=0$ ve $t=\infty$ için aldığı değerler sırasıyla $S(0) = \lim_{t \rightarrow 0} S(t) = 1$ ve $S(\infty) = \lim_{t \rightarrow \infty} S(t) = 0$ biçimindedir[10]. t_j zamanında riskte olan birimlerin sayısı n_j , t_j 'deki başarısızlıkların sayısı d_j ve $t_1 < \dots < t_k$ sıralı başarısızlık süreleri olmak üzere yaşam fonksiyonunun Kaplan-Meier tahmini,

$$\hat{S}(t) = \prod_{j < t} \left(\frac{n_j - d_j}{n_j} \right) \quad (2)$$

biçiminde tanımlanmaktadır. Durdurma ve durdurmanın olmadığı durumların her ikisinde de yaşam fonksiyonu tahmini bir adım fonksiyonudur ve her bir yaşam süresi t_j 'den hemen sonra $(n_j - d_j)/n_j$ kadar azalmaktadır.

Başarısızlık süresi için kullanılan fonksiyonlardan bir diğeri, t anındaki başarısızlığın koşullu yoğunluğu olarak tanımlanan hazard fonksiyonudur. Hazard fonksiyonu, t zamanına kadar yaşayan birimin $[t, t+\Delta t]$ aralığında yaşamının sona ermesi riskinin bir tanımıdır ve

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t / T > t)}{\Delta t} \quad (3)$$

biçimindedir. Sürekli dağılımlar için $h(t) > 0$ ve $\int_0^{\infty} h(t)dt = \infty$ özelliklerini sağlamaktadır. Hazard fonksiyonu, aynı zamanda $h(t) = f(t)/S(t)$ olarak da ifade edilmektedir[10].

3.2. Regresyon modelleri

Bağımlı değişken olan yaşam süresi üzerinde açıklayıcı değişkenlerin etkilerinin araştırıldığı regresyon modelleri yaşam çözümlemesinde önemli bir yere sahiptir.

Yaşam süresi dağılımının belirlenebildiği durumlarda yarı parametrik regresyon modelinin, kesin olarak belirlenebildiği durumlarda ise parametrik regresyon modellerinin kullanılması uygun olmaktadır. Bu modeller aşağıdaki bölümlerde kısaca ele alınmıştır.

3.2.1. Yarı parametrik regresyon modeli

Yaşam süresine ilişkin faktörlerin hazard fonksiyonu üzerindeki etkilerinin çarpımsal olduğu modeller, yaşam süresi verilerinin çözümlemesinde önemli bir rol oynamaktadır. İlk kez 1972’de Cox tarafından ele alınan Cox regresyon modeli de bu modellerden biridir.

T, bir birimin yaşam süresini temsil eden sürekli raslantı değişkeni ve x bu birimle ilgili açıklayıcı değişkenler vektörü olmak üzere orantılı hazard varsayımı altında x bilindiğinde T’nin hazard fonksiyonu,

$$h(t/x) = h_0(t) \exp(\beta'x) \quad (4)$$

biçimindedir. Burada, β regresyon katsayıları vektörü ve $h_0(t)$ ise $x = 0$ olan bir birimin temel hazard fonksiyonu olarak tanımlanmaktadır[5]. Cox regresyon modeli için parametreler olabilirlik fonksiyonu (L) maksimize edilerek tahmin edilmektedir.

Cox (1972)’un önerdiği yaklaşım, $h_0(t)$ için özel bir biçimin varsayılmadığı, dağılımdan bağımsız bir yaklaşımdır. Yani bu yöntem, ilgili yaşam süresi dağılımına bir başka deyişle temel hazard fonksiyonu $h_0(t)$ ’ye dayanmamaktadır. Bu model hazard fonksiyonun fonksiyonel biçimini belirlemeden hazard fonksiyonu üzerindeki açıklayıcı değişkenlerin etkilerini incelemek için tasarlandığından temel hazard fonksiyonunun tahmin edilmesine gerek yoktur.

3.2.2. Parametrik regresyon modelleri

Cox regresyon modeli, orantılı hazardlar varsayımına karşı duyarlıdır. Temel hazard fonksiyonunun biçimi tahmin edicilerin özelliklerini etkileyebilir[2]. Temel hazard fonksiyonu belirli bir dağılım ile hesaplanmadığından Cox regresyon modeli parametrik regresyon modellerine göre daha avantajlıdır. Çünkü yaşam sürelerinin olasılık dağılımının belirli bir biçimi olmadığından esnek ve yaygın kullanıma sahiptir. Ancak veri kümesi için belirli bir olasılık dağılımı varsayımı geçerli ise, bu varsayımaya dayalı çıkarsamalar daha kesindir. Ayrıca parametre tahminleri ve göreceli hazard ya da ortanca yaşam süresi gibi ölçümlerin tahminleri daha küçük standart hataya sahip olur[4]. Efron (1977) and Oakes (1977) bu durumda parametrik regresyon modellerinin Cox regresyon modeline göre daha etkili parametre tahminlerine sahip olduklarını göstermişlerdir[11].

Açıklayıcı değişkenler vektörü $X=(X_1, \dots, X_p)$ ile gösterilsin. İstatistiksel literatürde, yaşam süresini etkileyen açıklayıcı değişkenleri parametrik regresyon modelleri ile modellemek için iki yaklaşım vardır. Birinci yaklaşım, klasik doğrusal regresyon yaklaşımına benzemektedir. Bu yaklaşımda, yaşam süresinin doğal logaritması $Y=\ln(T)$ modellenir. Y için doğrusal model,

$$Y = \mu + \beta'X + \sigma W \quad (5)$$

biçimindedir. Burada $\beta' = (\beta_1, \beta_2, \dots, \beta_p)$ regresyon katsayıları vektörü ve W hata dağılımıdır. Hata dağılımı için en sık kullanılan seçenekler log-normal regresyon modelini veren standart normal dağılım, Weibull regresyon modelini veren uçdeğer dağılımı (extreme value) ya da log-lojistik regresyon modelini veren lojistik dağılımdır. Bu yaklaşım için kullanılan parametrik yaşam modelleri hızlandırılmış

başarısızlık zamanı modelleri olarak da adlandırılmaktadır. Ancak bu yaklaşım hata dağılımları için varsayım içerdiğinden kullanım alanları daha sınırlıdır. Bu durumda ikinci yaklaşımın kullanılması daha uygun olabilir. Açıklayıcı değişkenlerin yaşam fonksiyonu üzerindeki etkilerini modellemek için kullanılan ikinci yaklaşım, açıklayıcı değişkenlerin bir fonksiyonu olarak koşullu hazard fonksiyonunu modellemektir. Buna göre koşullu hazard fonksiyonu,

$$h(t/x) = h_0(t)\varphi(\beta'x) \quad (6)$$

biçiminde yazılır[8,10]. Burada $\varphi(\beta'x)$ açıklayıcı değişkenlerin negatif olmayan bir fonksiyonudur. Birçok uygulamada $\varphi(\beta'x) = \exp(\beta'x)$ olarak seçilmektedir. Parametrik regresyon modeli için koşullu hazard fonksiyonu belirlenirken, yaşam süresi hangi dağılıma uygunsa o dağılıma ait temel hazard fonksiyonu kullanılır.

Parametrik regresyon modellerinden en önemlileri üstel, Weibull, log-lojistik, log-normal, Gompertz, Genelleştirilmiş Gamma dağılımlarını içermektedir. Bu dağılımlara ait temel hazard fonksiyonları ($h_0(t)$), olasılık yoğunluk fonksiyonları ($f(t)$) ve yaşam fonksiyonları ($S(t)$) Çizelge 1'de verilmiştir.

Çizelge 1. Parametrik dağılımlar için temel hazard fonksiyonları, olasılık yoğunluk fonksiyonları ve yaşam fonksiyonları

Dağılım	Parametre	$h_0(t)$	$f(t)$	$S(t)$
Üstel	$\lambda > 0$	λ	$\lambda \exp(-\lambda t)$	$\exp(-\lambda t)$
Weibull	$\lambda > 0$ $\alpha > 0$	$\lambda \alpha t^{\alpha-1}$	$\lambda \alpha t^{\alpha-1} \exp(-\lambda t^\alpha)$	$\exp(-\lambda t^\alpha)$
Log-lojistik	$\alpha, \lambda > 0$	$\frac{\alpha t^{\alpha-1} \lambda}{1 + \lambda t^\alpha}$	$\frac{\alpha t^{\alpha-1} \lambda}{[1 + \lambda t^\alpha]^2}$	$\frac{1}{1 + \lambda t^\alpha}$
Log-normal	$\alpha > 0$	$\frac{f(t)}{S(t)}$	$\frac{\exp\left[-\frac{1}{2}\left(\frac{\ln t - \mu}{\sigma}\right)^2\right]}{t(2\pi)^{1/2}\sigma}$	$1 - \Phi\left[\frac{\ln t - \mu}{\sigma}\right]$
Gompertz	$\lambda > 0$ $\alpha \in (-\infty, \infty)$	$\lambda \exp(\alpha t)$	$\lambda \exp(\alpha t) \exp\left(\frac{\lambda}{\alpha}(1 - \exp(\alpha t))\right)$	$\exp\left[\frac{\lambda}{\alpha}(1 - e^{\alpha t})\right]$
Genelleştirilmiş Gamma	$\alpha, \beta, \lambda > 0$	$\frac{f(t)}{S(t)}$	$\frac{\alpha \lambda^\beta t^{\alpha\beta-1} \exp(-\lambda t^\alpha)}{\Gamma(\beta)}$	$1 - I[\lambda t^\alpha, \beta]$

I : Tamamlanmamış Gama fonksiyonu

Φ : Standart normal bir değişkenin birikimli dağılım fonksiyonu

μ : Ortalama, σ : Varyans

Uygun parametrik regresyon modeline karar verebilmek için $-2\log L$ istatistiği ya da Akaike bilgi kriteri (AIC) kullanılmaktadır. En küçük $-2\log L$ istatistiğine ya da AIC değerine sahip model en uygun model olarak belirlenmektedir. Parametrik modeller için AIC değeri Eşitlik 7'deki gibi hesaplanmaktadır:

$$AIC = -2 \log L + 2(p + 2 + k) \quad (7)$$

AIC değerleri hesaplanırken üstel model için $k=0$, Weibull, log-lojistik ve log-normal modelleri için $k=1$ ve genelleştirilmiş Gamma model için $k=2$ olarak alınmaktadır[8].

4. Uygulama

Sigorta şirketleri istatistiksel analizleri kullanarak poliçesi devam eden, poliçesini iptal ettiren ya da iptal ettirebilecek poliçe sahiplerinin özelliklerini belirleyerek, hem yeni müşteriler elde etmek için, hemde halihazırda bulunan müşterilerini kaybetmemek için çeşitli stratejiler geliştirebilirler.

Çalışmamızın yapısına uygun gerçek bir veri kümesine ulaşılamadığından, birikimli hayat poliçesi sahiplerine ait örnek bir veri kümesi üretilmiş ve poliçe sahiplerinin, poliçelerini iptal ettirmelerini etkileyen faktörler yaşam çözümlemesi yöntemleri kullanılarak belirlenmeye çalışılmıştır. Veri üretme sürecinde başarısızlık süresi için Gompertz dağılımı, durum değişkeni(durdurulmuş, başarısız) için ise tekbiçimli dağılım kullanılmıştır. Gompertz dağılımı, tıp bilimleri, aktüerya bilimleri ve güvenilirlik kuramında yaşam süresinin dağılımı için sıklıkla kullanılan bir dağılımdır[1]. Gompertz dağılımına uygun veri üretmek için Bender v.d. (2005)'nin çalışmalarında kullanılan veri üretme süreci kullanılmıştır. Çalışmada poliçe sahiplerinin poliçeyi yaptırdığı tarihten, poliçeyi iptal ettirdiği tarihe kadar geçen süre (yıl olarak) başarısızlık süresi olarak alınmıştır. Poliçesini 10 yıldan önce iptal ettiren poliçe sahipleri başarısız, poliçesi devam ettiren poliçe sahipleri ise durdurulmuş olarak tanımlanmıştır. Poliçe sahiplerinin izlenme süresi sona erdiğinde 2000 poliçe sahibinden 1582'sinde (% 79.1) başarısızlık ve 418'inde (% 20.9) durdurma gözlenmiştir. Uygulamada cinsiyet, medeni durum, meslek, yaş, gelir, ödeme türü ve prim ödeme dönemi değişkenleri çözümlemeye alınmıştır. Bu değişkenlerin düzeyleri Çizelge 2'de verilmiştir. Çalışmada kullanılan değişkenler için 3, 5 ve 10 yıllık yaşam olasılıkları elde edilmiş ve Çizelge 2'de verilmiştir. Yıllar ilerledikçe poliçe sahiplerinin poliçelerini iptal ettirme riskinin arttığı, yani yaşam olasılığının azaldığı görülmüştür. Poliçe sahiplerinin başarısızlık sürelerini etkileyen faktörler, yarı parametrik (Cox) ve parametrik (üstel, Weibull, Log-lojistik, Log-normal, Gama, Gompertz) regresyon modelleri kullanılarak belirlenmeye çalışılmıştır.

Çizelge 2. Kaplan-Meier yaşam olasılıkları sonuçları

Değişkenler ve Düzeyleri		Birikimli Yaşam Olasılığı		
		3YIL	5 YIL	10 YIL
Cinsiyet	1.Erkek	0.5515	0.4047	0.2110
	2.Kadın	0.5182	0.3876	0.2071
Medeni Durum	1.Bekar	0.4769	0.3402	0.1573
	2.Evli	0.5931	0.4537	0.2625
Meslek	1. Memur	0.5552	0.4044	0.2100
	2. İşçi	0.5386	0.4090	0.2068
	3. Serbest	0.5152	0.3804	0.2151
	4. İşsiz	0.3548	0.2258	0.0645
	5. Emekli	0.6667	0.5833	0.4583
Yaş	1. 18-25	0.3200	0.1938	0.0277
	2. 26-40	0.4424	0.2944	0.0822
	3. 41-55	0.5924	0.4297	0.2196
	4. 56+	0.7350	0.6359	0.5069
Gelir	1. ≤ 500	0.3810	0.2976	0.2143
	2. ≤ 1000	0.4873	0.3527	0.1782
	3. ≤ 1500	0.5519	0.4148	0.2222
	4. ≤ 2000	0.5655	0.4045	0.2434
	5. ≤ 2500	0.5878	0.4466	0.1985
	6. 2500 +	0.5585	0.3907	0.2067
Ödeme Türü	1. YTL	0.4903	0.3369	0.1254
	2. Dolar	0.9800	0.9800	0.9800
	3. Euro	0.2609	0.0870	0.0435
Prim Ödeme Dönemi	1. 1 aylık	0.5427	0.4047	0.2267
	2. 3 aylık	0.8889	0.5778	0.0778
	3. 6 aylık	0.2113	0.2113	0.2113
	4. 1 yıllık	0.0385	0.0395	0.3850

Modellerin anlamlığını test etmek için olabilirlik oranı (LR) test istatistiği kullanılmış ve model seçim kriteri olarak da $-2\log L$ kullanılmıştır. Bu nedenle çalışmada incelenen modellere ait olabilirlik oranı test istatistikleri ve $-2\log L$ değerleri elde edilmiş ve Çizelge 3’de verilmiştir.

Çizelge 3. Yarı parametrik ve parametrik regresyon modelleri için $-2\log L$ değerleri

Kriter	Yarı Parametrik Regresyon Modeli	Parametrik Regresyon Modeli					
	Cox	Üstel	Weibull	Log-lojistik	Log-normal	Gama	Gompertz
LR (p değeri)	966.16 (0,00)	1414.33 (0,00)	1153.47 (0,00)	1043.54 (0,00)	692.23 (0,00)	1053.93 (0,00)	974.68 (0,00)
$-2\log L$	192601.62	6066.50	6030.39	5969.9	6080.43	5987.7	5949.65

Çizelge 3’te yer alan sonuçlar incelendiğinde tüm modellerinin istatistiksel olarak anlamlı olduğu (p değeri < 0.05) görülmüştür. $-2\log L$ değerinin küçük olması modelin veri kümesine uygunluğunu gösterdiğinden, bu çalışmaya en uygun modelin parametrik regresyon modellerinden Gompertz regresyon modelinin olduğu sonucuna ulaşılmıştır.

Çalışmadaki açıklayıcı değişkenlerle Gompertz regresyon çözümlemesi yapıldığında elde edilen sonuçlar Çizelge 4’de verilmiştir.

Çizelge 4. Gompertz regresyon çözümlemesinin sonuçları

Değişken		Exp(β)	S.H.	p-değeri	Alt sınır – Üst sınır
Cinsiyet	1. Erkek				
	2. Kadın	1.0878	0.0553	0.098	0.9846 – 1.2019
Medeni durum	1. Bekar				
	2. Evli	1.0242	0.0556	0.660	0.9208 – 1.1392
Meslek	1. Memur				
	2. İşçi	0.9631	0.6078	0.551	0.8511 – 1.0899
	3. Serbest	1.0159	0.6523	0.805	0.8958 – 1.1522
	4. İşsiz	1.0099	0.1614	0.516	0.8250 – 1.4664
	5. Emekli	1.0433	0.3045	0.884	0.5889 – 1.8485
Yaş	1. 18-25				
	2. 26-40	0.4691	0.0396	0.000	0.3975 – 0.5535
	3. 41-55	0.3107	0.0273	0.000	0.2616 – 0.3691
	4. 56+	0.3250	0.0347	0.000	0.2642 – 0.3999
Gelir	1. ≤ 500				
	2. ≤ 1000	1.1537	0.1665	0.322	0.8695 – 1.5307
	3. ≤ 1500	1.3649	0.2066	0.040	1.0145 – 1.8362
	4. ≤ 2000	1.3673	0.2074	0.039	1.0156 – 1.8407
	5. ≤ 2500	1.2710	0.1937	0.116	0.9428 – 1.7134
	6. 2500 +	1.3403	0.1877	0.036	1.0186 – 1.7634
Ödeme Türü	0. YTL				
	1. Dolar	0.0099	0.0050	0.000	0,3712 – 0.0269
	2. Euro	1.0192	0.2240	0.931	0.6625 – 1.5678
Prim Ödeme Dönemleri	0. 1 aylık				
	1. 3 aylık	0.3179	0.0416	0.000	0.2460 – 0.4108
	2. 6 aylık	2.6189	0.8003	0.002	1.4388 – 4.7670
	3. 1 yıllık	2.6645	0.3547	0.000	2.0524 – 3.4589
Gamma		-0.0822	0.0082	0.000	-0.0984 – -0.0661

Yaşam çözümlemesinde kullanılan regresyon modellerinde değişken düzeylerinden biri referans kategorisi olarak alınmakta ve değişken düzeylerinin yorumlanması buna göre yapılmaktadır. Bu

çalışmada, modeldeki değişken için standart hata (S.H.), p değeri, hazard oranı ($\exp(\beta)$) ile hazard oranının alt ve üst sınırları verilmiştir. Her bir değişken için ilk düzeyler referans kategorisi olarak alınmıştır. Başarısızlık süresini etkileyen faktörleri belirlemek için Çizelge 4 incelendiğinde, yaş ve prim ödeme dönemleri değişkenlerinin, gelir değişkeninin 3., 4., 6. düzeylerinin, ödeme türü değişkeninin 2.düzeyinin önemli olduğu (p -değeri <0.05) %0,5 yanılma düzeyi ile söylenebilirken, medeni durum, cinsiyet ve meslek değişkenlerinin % 0,5 yanılma düzeyinde önemli olmadığı görülmüştür. Yaş arttıkça poliçeyi iptal ettirme riski azalmaktadır. Ayrıca prim ödeme dönemleri uzadıkça poliçeyi iptal ettirme riskinin arttığı görülmüştür. Primi ödeme dönemi 6 aylık ve 1 yıllık olan poliçe sahiplerinin, prim ödeme dönemi 1 aylık olan poliçe sahiplerine göre poliçelerini iptal ettirme risklerinin 2,6 kat daha fazla olduğu görülmüştür.

5. Sonuç

Bu çalışmada yaşam modellerinden yarı parametrik ve parametrik regresyon modelleri incelenmiş, birikimli hayat sigortası ile ilgili bir uygulama üzerinden bu modellerin sigortacılık sektöründe kullanımı ele alınmıştır. Uygulamada yaşam olasılıkları 3, 5 ve 10 yıllık dönemler için elde edilmiş ve poliçe sahiplerinin poliçelerini iptal ettirmelerini etkileyen faktörler yaşam modelleri kullanılarak belirlenmeye çalışılmıştır. Bu amaçla elde edilen veri kümesi için Cox regresyon modeli ve parametrik regresyon modellerine ait sonuçlar elde edilmiş ve Gompertz regresyon modelinin veri kümesi için uygun olduğu sonucuna ulaşılmıştır. İncelenen yaşam verisinin belirli bir dağılıma sahip olması durumunda, parametrik regresyon modellerinin Cox regresyon modelinden daha iyi sonuç verdiği bu çalışmada da görülmüştür.

Sigorta şirketleri poliçe sahiplerine ait bilgileri yaşam modellerini kullanarak inceleyebilir ve sektöre özel modeller oluşturularak ileriye dönük tahminler elde edebilirler. Sigorta şirketlerinden poliçe sahiplerine ait gerçek veriler alınabilmesi halinde, ilgilenilen başarısızlık olayını etkileyen değişkenler incelenerek sigortacılık sektörüne faydalı olabilecek sonuçlar ortaya konulabilecektir.

Kaynaklar

- [1] Ananda, M.M., Dalpatadu, R.J., Singh, A.K., "Adaptive Bayes estimators for parameters of the Gompertz survival model", Applied Mathematics and Computation, vol.75(2), pp.167-177, 1996.
- [2] Bender R., Augustin, T., Blettner, M. "Generating survival times to simulate Cox proportional hazards models", Statistics in Medicine, vol.24, pp.1713-1723, 2005.
- [3] Box-Steffensmeier, J.M., Reiter, D., Cgristopher, Z., "Nonproportional Hazards and Event History Analysis in International Relations", The Journal of Conflict Resolution, vol.27(1), pp.33-53, 2003.
- [4] Collett, D., "Modelling Survival Data in Medical Research", Chapman&Hall, UK, 1994.
- [5] Cox, D.R., "Regression models and life-tables", Journal of the Royal Statistical Society, Series B, 34, 187-220, 1972.
- [6] Evrensel, A.Y., "Banking Crisis and Financial Structure: A Survival Time Analysis", International Review of Economics and Finance, 2007.
- [7] Keiding, N., Andersen, C., Fledulus, P., "The Cox Regression Model for Claims Data in Non-life Insurance", Astin Bulletin, vol.28(1), pp.95-118, 1998.
- [8] Klein, John P., Moeschberger, Melvin L., 1997, Survival Analysis Techniques for Censored and Truncated Data, Springer, New York.
- [9] Linoff, G.S., "Survival Data Mining for Customer Insight", Intelligent Enterprise, 2004.
- [10] London, D., "Survival Models and Their Estimation", Actex Publications, USA, 1997.
- [11] Nardi, A., Schemper, M., "Comparing Cox and Parametric Models in Clinical Studies", Statistics in Medicine, vol.22, pp.3597-3610, 2003.
- [12] Potts, W., "Survival data mining", SUGI Technical Report, 2004.
- [13] T.C. BHMSDK (2006), Sigortacılık ve BES Faaliyet Raporu-2006, T.C. Başbakanlık Hazine Müsteşarlığı Sigorta Denetleme Kurulu, 1-7.
- [14] T.C. BHMSGM (2005), Türk Sigortacılık Sektörünün Temel Göstergeleri 4, T.C. Başbakanlık Hazine Müsteşarlığı Sigortacılık Genel Müdürlüğü, 1-3.
- [15] TSRŞB (2008), Türkiye Sigorta ve Reasürans Şirketleri Birliği Web Sayfası, <http://www.tsrbsb.org.tr/tsrbsb/Sigorta/Branşlar/Hayat+Sigortalari.htm>